The volume-preserving Willmore flow

Fabian Rupp[*][†]

January 31, 2023

Abstract: We consider a closed surface in $\mathbb{R}^3$ evolving by the volume-preserving Willmore flow and prove a lower bound for the existence time of smooth solutions. For spherical initial surfaces with Willmore energy below 8π we show long time existence and convergence to a round sphere by performing a suitable blow-up and by proving a constrained Łojasiewicz–Simon inequality.

Keywords: Willmore flow, fixed volume, blow-up, Łojasiewicz–Simon inequality, non-local geometric evolution equation.
MSC(2020): 53E40 (primary), 35B40, 35K41 (secondary).

1. Introduction and main results

For an immersion $f\colon \Sigma \to \mathbb{R}^3$ of a compact, connected and oriented surface Σ without boundary, its *Willmore energy* is defined by

$$\mathcal{W}(f) := \frac{1}{4}\int_\Sigma H^2 \,\mathrm{d}\mu. \tag{1.1}$$

Here $\mu = \mu_f$ denotes the area measure, induced by the pull-back of the Euclidean metric $g_f := f^*\langle\cdot,\cdot\rangle$, and $H = H_f := \langle \vec{H}_f, \nu_f\rangle$ denotes the (scalar) mean curvature with respect to $\nu = \nu_f\colon \Sigma \to \mathbb{S}^2$, the unique unit normal along f induced by the chosen orientation on Σ, see (2.1) below. By the Gauß–Bonnet theorem, the Willmore energy (1.1) only differs by a topological constant from the squared L^2-norm of A^0, the trace-free part of the second fundamental form. Indeed, we have

$$\overline{\mathcal{W}}(f) := \int_\Sigma |A^0|^2 \,\mathrm{d}\mu = 2\mathcal{W}(f) - 4\pi\chi(\Sigma), \tag{1.2}$$

[*]Institute of Applied Analysis, Ulm University, Helmholtzstraße 18, 89081 Ulm, Germany.
[†]Present address: Faculty of Mathematics, University of Vienna, Oskar-Morgenstern-Platz 1, 1090 Vienna, Austria. `fabian.rupp@univie.ac.at`.

where $\chi(\Sigma)$ denotes the Euler characteristic. Note that both energies are not only *geometric*, i.e. invariant under diffeomorphisms of Σ, but also *conformally invariant*, i.e. invariant under rigid motions and inversions provided the center of inversion does not lie on $f(\Sigma)$. As already observed in [41], $\mathcal{W}(f) \geq 4\pi$ with equality only for round spheres. Therefore, $\mathcal{W}$ and hence also $\overline{\mathcal{W}}$ are a natural way to measure the total bending of an immersed surface with various applications also beyond differential geometry, for instance in the study of biological membranes [9, 16], general relativity [15] and image restoration [13].
The analysis of the *Willmore flow*, i.e. the L^2-gradient flow associated to the energy $\overline{\mathcal{W}}$, started with the work of Kuwert and Schätzle. In [21], they proved a lifespan theorem under the assumption that the *concentration of curvature* of the initial datum is controlled. In [20], this was used to set up a blow-up procedure and to prove convergence to a round sphere if the energy is sufficiently small. Then in [22], long-time existence and convergence was shown for the flow of spherical immersions with initial datum $f_0 \colon \mathbb{S}^2 \to \mathbb{R}^3$ satisfying $\mathcal{W}(f_0) \leq 8\pi$. The threshold 8π already appears in the celebrated *Li–Yau inequality for the Willmore energy* [27, Theorem 6], yielding that f is an embedding if $\mathcal{W}(f) < 8\pi$. Furthermore, this threshold is in fact sharp for the convergence result in [22], see [28] for numerical experiments and [4] for an analytic proof. It remains an open problem to prove or disprove whether this singularity happens in finite time.
Recently, similar convergence results have been established for the Willmore flow of *tori of revolution* [12] with the same energy threshold and also for the Willmore flow of *Hopf-tori* in the three-sphere $\mathbb{S}^3$ [19].
Moreover, various authors have extended the methods of Kuwert and Schätzle to related geometric evolution equations also involving constraints, including, for instance, the *surface diffusion flow* [32, 39, 40], *Helfrich-type flows* [31, 5] and other higher order flows [3, 29]. In [18], the *area-preserving Willmore flow* was studied. Related constrained evolution problems for the *elastic energy* of curves have been considered in [14], [11] and [35], for instance.
In this article, we introduce a *constrained gradient flow*, which evolves an initial immersion $f \colon \Sigma \to \mathbb{R}^3$ such that $\overline{\mathcal{W}}$ decreases as fast as possible, while $\mathcal{V}$, the *signed volume* of $f(\Sigma)$, defined by

$$\mathcal{V}(f) := -\frac{1}{3} \int_\Sigma \langle f, \nu \rangle \, \mathrm{d}\mu, \tag{1.3}$$

is kept constant. The analogous problem for the *mean curvature flow* was introduced by Huisken [17]. More explicitly, we say that a smooth family of immersions $f \colon [0,T) \times \Sigma \to \mathbb{R}^3$ is a *volume-preserving Willmore flow*, if it satisfies the geometric evolution equation

$$\partial_t f = \left(-\Delta H - |A^0|^2 H + \lambda \right) \nu, \tag{1.4}$$

where $\Delta = \Delta_f$ is the Laplace–Beltrami operator on (Σ, g_f) and the Lagrange multiplier $\lambda := \lambda(t) := \lambda(f_t)$ depends on the immersion $f_t := f(t, \cdot)$ and is given by

$$\lambda := \frac{\int_\Sigma |A^0|^2 H \, \mathrm{d}\mu}{\mathcal{A}(f)}, \tag{1.5}$$

where $\mathcal{A}(f) := \int_\Sigma \mathrm{d}\mu_f$ denotes the total area of Σ. In Section 2.3, we will prove that (1.4) actually decreases $\overline{\mathcal{W}}$, hence also $\mathcal{W}$ by (1.2), while keeping $\mathcal{V}$ fixed. Note that the energies defined in (1.1) and (1.2) do not change, if we reverse the orientation on Σ. While the normal ν and hence the volume (1.3) change sign, the flow equation (1.4) with λ as in (1.5) is also invariant under reversing the orientation.
The *stationary solutions* to the volume-preserving Willmore flow (1.4) are characterized as solutions of the PDE

$$\Delta H + |A^0|^2 H = \lambda \quad \text{for some } \lambda \in \mathbb{R}. \tag{1.6}$$

Formally, (1.6) is the Euler–Lagrange equation of critical points of $\mathcal{W}$ (and hence of $\overline{\mathcal{W}}$) subject to a volume constraint, so we refer to solutions of (1.6) as *(volume)-constrained Willmore immersions,* see Lemma 2.4 and Lemma 2.6 below. By changing from λ to $-\lambda$, (1.6) is preserved under reversing the orientation on Σ. Note that while the energies $\mathcal{W}$ and $\mathcal{V}$ may not be well-defined, (1.6) makes sense even if Σ is not compact, and we still term noncompact solutions of (1.6) constrained Willmore immersions.
Our first main contribution extends the energy concentration-based lower bound on the lifespan for the Willmore flow [21] to the volume-preserving Willmore flow.

Theorem 1.1. *There exists an absolute constant $\bar{\varepsilon} > 0$ such that if $f_0 \colon \Sigma \to \mathbb{R}^3$ is an immersion with $\mathcal{W}(f_0) \leq K$ and $\rho > 0$ is chosen such that*

$$\int_{B_\rho(x)} |A_0|^2 \,\mathrm{d}\mu_0 \leq \varepsilon < \bar{\varepsilon} \quad \textit{for all } x \in \mathbb{R}^3,$$

then the maximal existence time T of the volume-preserving Willmore flow with initial data f_0 satisfies

$$T > \hat{c}\rho^4,$$

for some $\hat{c} = \hat{c}(K, \chi(\Sigma)) > 0$ and furthermore for all $0 \leq t \leq \hat{c}\rho^4$ it holds

$$\int_{B_\rho(x)} |A|^2 \,\mathrm{d}\mu \leq \hat{c}^{-1}\varepsilon \quad \textit{for all } x \in \mathbb{R}^3. \tag{1.7}$$

Following the notation of [21], the integrals above have to be understood over the preimages under f_0 and f_t, respectively.
The proof of Theorem 1.1 follows the concentration-compactness strategy developed by Kuwert and Schätzle for the Willmore flow in [21]. Since this method is relying on smallness of the curvature in small balls, it is *intrinsically local,* making the *nonlocal* nature of the Lagrange multiplier a major difficulty. To compensate this, the $L^{4/3}$-norm of λ naturally appears in these estimates, a scale-invariant quantity (see Remark 2.2) which we can control under certain assumptions, see Section 4.1 below. In particular, we do not assume a-priori L^∞-bounds on λ as in [32, 39]. However, to show this correct integrability, we have to allow the constant $\hat{c}$ to depend on an upper bound for the initial energy, as well as on the topology of Σ, in contrast to [21, Theorem 1.2].
Our second main contribution extends the convergence result of Kuwert and Schätzle [22, Theorem 5.2] in the following

Theorem 1.2. *Let $f_0 \colon \mathbb{S}^2 \to \mathbb{R}^3$ be a smooth immersion such that $\mathcal{W}(f_0) \leq 8\pi$ and $\mathcal{V}(f_0) \neq 0$. Then the volume-preserving Willmore flow with initial datum f_0 exists for all times and converges smoothly, after reparametrization, to a round sphere with radius $R = \sqrt[3]{\frac{3|\mathcal{V}(f_0)|}{4\pi}}$ as $t \to \infty$.*

As in [20, 22], the strategy to prove Theorem 1.2 is a blow-up construction based on the lifespan bound in Theorem 1.1. The blow-up limit is a constrained Willmore immersion and in general noncompact. However, apart from a small energy regime [30], for $\lambda \neq 0$ there is no classification of solutions to (1.6). Nonetheless, under an L^2-integrability assumption on the Lagrange multiplier, we are able to conclude that the blow-up is a Willmore immersion, i.e. $\lambda = 0$ in (1.6). In the energy regime of Theorem 1.2, this integrability can be deduced from a reverse isoperimetric inequality [36, 4]. Together with the removability result [22] and the classification of Willmore spheres [8], we then conclude that the blow-up limit is compact. A result of independent interest is that the volume-constrained Willmore functional satisfies an appropriate constrained version of the Łojasiewicz–Simon gradient inequality in the sense of [34]. Finally, this inequality yields a stability result in the spirit of [10] from which we conclude global existence and convergence of the flow if a blow-up is compact.

Note that in view of [22, Theorem 5.2], we believe that the reparametrization in Theorem 1.2 is not necessary, but a common consequence when relying on the Łojasiewicz–Simon gradient inequality, cf. [10, Lemma 4.1].

This article is structured as follows. First, we recall some definitions and compute the evolution of relevant quantities in Section 2. Section 3 is devoted to proving a key ingredient of the paper: localized integral estimates in the spirit of [21], which now require an $L^{4/3}$ in time integrability of the Lagrange multiplier. Combined with the careful a-priori estimates of λ which we establish in Section 4, they are then used to prove the lifespan bound, Theorem 1.1, in Section 5. In Section 6 we construct a blow-up limit and study its properties. We then deduce a convergence result for compact blow-ups in the spirit of [10] by proving a *constrained Łojasiewicz–Simon gradient inequality* in Section 7, before proving Theorem 1.2 in Section 8. For the sake of readability, some details and well-known arguments have been moved to the appendix and may be skipped by the eager or experienced reader.

2. Preliminaries

In this section, we will review the geometric and analytic background and prove some first properties of the flow (1.4). In the following, Σ will always denote a compact and connected oriented surface without boundary. Note that in contrast to [21, 20], we work exclusively in codimension one, which simplifies the relevant geometric objects.

2.1. Immersed and embedded surfaces in $\mathbb{R}^3$

An immersion $f\colon \Sigma \to \mathbb{R}^3$ induces the pullback metric $g = f^*\langle\cdot,\cdot\rangle$ on Σ, which in local coordinates is given by

$$g_{ij} := \langle \partial_i f, \partial_j f\rangle,$$

where $\langle\cdot,\cdot\rangle$ denotes the Euclidean metric. The chosen orientation on Σ determines a unit normal field $\nu\colon \Sigma \to \mathbb{S}^2$ along f, which in local coordinates in the orientation is given by

$$\nu = \frac{\partial_1 f \times \partial_2 f}{|\partial_1 f \times \partial_2 f|}. \tag{2.1}$$

We will always work with this unit normal vector field. The second fundamental form of f is then given by projecting the second derivatives of f in normal direction, i.e. in local coordinates we define $A_{ij} := \langle \partial_i\partial_j f, \nu\rangle$. The mean curvature and the tracefree part of the second fundamental form are

$$H := g^{ij}A_{ij} \text{ and } A^0_{ij} := A_{ij} - \frac{1}{2}Hg_{ij},$$

where $g^{ij} := (g_{ij})^{-1}$. Note that g, A and A^0 are scalar valued $\binom{2}{0}$-tensors. Moreover, for any vector field X along f, we have the tangential and normal projections

$$\begin{aligned} P^\top X &:= P^{\top_f} X := g^{ij}\langle X, \partial_i f\rangle \partial_j f \\ P^\perp X &:= P^{\perp_f} X := X - P^\top X. \end{aligned}$$

The Levi-Civita connection $\nabla = \nabla_f$ induced by g extends uniquely to a connection on tensors, which we also denote by ∇. For an orthonormal basis $\{e_1, e_2\}$ of the tangent space, the Codazzi–Mainardi equations then yield

$$\nabla_i H = (\nabla_j A)(e_i, e_j) = 2(\nabla_j A^0)(e_i, e_j). \tag{2.2}$$

The *Laplace–Beltrami operator* on (Σ, g) is given by

$$\Delta_g \xi = g^{ij}\nabla_i\nabla_j \xi, \quad \text{for } \xi \in C^\infty(\Sigma).$$

For a $\binom{2}{0}$- tensor T_{ij}, its tensor norm is $|T|^2 := g^{ij}g^{k\ell}T_{ik}T_{j\ell}$ and hence we get

$$|A|^2 = |A^0|^2 + \frac{1}{2}H^2. \tag{2.3}$$

Consequently, using (1.2), we find

$$\int_\Sigma |A|^2 \,\mathrm{d}\mu = \overline{\mathcal{W}}(f) + 2\mathcal{W}(f) = 4\mathcal{W}(f) - 4\pi\chi(\Sigma). \tag{2.4}$$

2.2. The PDE perspective

Note that in the general context of *constrained gradient flows* on Hilbert spaces, cf. [34, Section 5], the flow in a Hilbert space H associated to the energy $\mathcal{E} = \overline{\mathcal{W}}$ with constraint $\mathcal{G} = -\mathcal{V} \equiv$ *constant* is formally given by

$$\begin{cases} \partial_t f &= -\nabla_H \overline{\mathcal{W}}(f) - \lambda(f)\nabla_H \mathcal{V}(f), \quad t > 0 \\ f(0) &= f_0, \end{cases}$$

where the Lagrange multiplier is defined by the formula

$$\lambda(f) = -\frac{\langle \nabla \overline{\mathcal{W}}(f), \nabla \mathcal{V}(f)\rangle_H}{\|\nabla \mathcal{V}(f)\|_H^2}.$$

If we choose $H := L^2(\mathrm{d}\mu_f)$, by the explicit form of the L^2-gradients (see Lemma 2.4 below), the divergence theorem and the fact that $\partial\Sigma = \emptyset$, this definition coincides with the flow in (1.4) with Lagrange multiplier λ as in (1.5). In particular, λ does not contain any derivatives of the curvature and is therefore of lower order compared to the leading term $-\Delta H$ in (1.4). This will significantly simply the analysis of λ later on.
Despite that, the flow equation (1.4) is still a *quasilinear, degenerate parabolic* PDE of 4^{th} order which is *nonlocal* due to the Lagrange multiplier. Hence, even short time existence and uniqueness is not immediate. However, as λ is of lower order, for smooth initial data one can show the following local well-posedness result by using an appropriate fixed-point argument in parabolic Hölder spaces, see for instance [33, Section 7] and [23, Section 3.1].

Proposition 2.1. *Let $f_0\colon \Sigma \to \mathbb{R}^3$ be a smooth immersion. Then there exist $T \in (0,\infty]$ and a unique, nonextendable smooth solution $f\colon [0,T)\times\Sigma \to \mathbb{R}^3$ of the volume-preserving Willmore flow with initial datum $f(0) = f_0$.*

An important property of the volume-preserving Willmore flow is the following *parabolic scaling*, which directly follows from the scaling behavior of the geometric quantities.

Remark 2.2. *If $f\colon [0,T)\times\Sigma \to \mathbb{R}^3$ is a volume-preserving Willmore flow, then for any $\rho > 0$ the family of immersions $\tilde{f}(t,p) := \rho^{-1} f(\rho^4 t, p)$ is also a volume-preserving Willmore flow on $[0,\tilde{T})\times\Sigma$ with $\tilde{T} = \rho^{-4}T$. Moreover, by a direct computation we have*

$$\int_0^T |\lambda(t)|^{\frac{4}{3}}\,\mathrm{d}t = \int_0^{\tilde{T}} |\tilde{\lambda}(t)|^{\frac{4}{3}}\,\mathrm{d}t, \quad \int_0^T |\lambda(t)|^2 \mathcal{A}(f_t)\,\mathrm{d}t = \int_0^{\tilde{T}} |\tilde{\lambda}(t)|^2 \mathcal{A}(\tilde{f}_t)\,\mathrm{d}t.$$

Note that the power $p = \frac{4}{3}$ is the only exponent for which the L^p-norm of λ behaves correctly with respect to the rescaling above and will naturally show up in Section 3 below. In Section 4, we show how to control both of these integrals.

2.3. Evolution of the geometric quantities

In this section, we recall the variation of the relevant geometric quantities and the (localized) evolution of the energy. The proofs are standard and can be found in [21, 20], for instance, or follow from direct computations.

Lemma 2.3. *Let $f\colon [0,T) \times \Sigma \to \mathbb{R}^3$ be a smooth family of immersions with normal velocity $\partial_t f = \xi\nu$. For an orthonormal basis $\{e_1, e_2\}$ of the tangent space, the geometric quantities induced by f satisfy*

$$(\partial_t g)(e_i, e_j) = -2A_{ij}\xi, \tag{2.5}$$

$$\partial_t(\mathrm{d}\mu) = -H\xi\,\mathrm{d}\mu, \tag{2.6}$$

$$\partial_t H = \Delta\xi + |A^0|^2\xi + \frac{1}{2}H^2\xi, \tag{2.7}$$

$$(\partial_t A)(e_i, e_j) = \nabla^2_{ij}\xi - A_{ik}A_{kj}\xi, \tag{2.8}$$

$$(\partial_t A^0)(e_i, e_j) = \left(\nabla^2_{ij}\xi\right)^0 - g_{ij}|A^0|^2\xi, \tag{2.9}$$

$$\partial_t \nu = -\operatorname{grad}_g \xi =: g^{ij}\partial_i\xi\partial_j f. \tag{2.10}$$

As a consequence, we find the first variation of the energy and the volume.

Lemma 2.4. *Let $f\colon \Sigma \to \mathbb{R}^3$ be an immersion. Then, the first variations of $\overline{\mathcal{W}}$ and $\mathcal{V}$ are given by*

$$\mathcal{V}'(f)\varphi = -\int \langle \nu, \varphi\rangle\,\mathrm{d}\mu,$$

$$\overline{\mathcal{W}}'(f)\varphi = \int \langle (\Delta H + |A^0|^2 H)\nu, \varphi\rangle\,\mathrm{d}\mu, \tag{2.11}$$

for all $\varphi \in C^\infty(\Sigma;\mathbb{R}^3)$ normal along f. Here and in the following, we always integrate over the whole surface Σ if the domain of integration is not specified.

This means that the $L^2(\mathrm{d}\mu_f)$-*gradients* of $\overline{\mathcal{W}}$ and $\mathcal{V}$ are given by the identities

$$\nabla\overline{\mathcal{W}}(f) = (\Delta H + |A^0|^2 H)\nu,$$

$$\nabla\,\mathcal{V}(f) = -\nu.$$

These gradients are purely normal, so we will often work with the *scalar* $L^2(\mathrm{d}\mu_f)$*-gradient*

$$\nabla_{sc}\overline{\mathcal{W}}(f) := \Delta H + |A^0|^2 H. \tag{2.12}$$

By direct computation, along a solution of (1.4) the volume is indeed preserved since

$$\frac{\mathrm{d}}{\mathrm{d}t}\,\mathcal{V}(f) = 0, \tag{2.13}$$

whereas by (2.11) and (2.13) the energy decreases by

$$\frac{\mathrm{d}}{\mathrm{d}t}\overline{\mathcal{W}}(f) = -\int |\partial_t f|^2\,\mathrm{d}\mu \leq 0. \tag{2.14}$$

Remark 2.5. *The computation in* (2.14) *implies that $\overline{\mathcal{W}}$ is a* strict Lyapunov function*, i.e. $\overline{\mathcal{W}}$ is strictly decreasing unless f is constant. By* (1.2) *this also holds for $\mathcal{W}$.*

It is now easy to prove the following rigidity result for constrained Willmore immersions.

Lemma 2.6. *Let Σ be a compact, oriented surface without boundary and let $f\colon \Sigma \to \mathbb{R}^3$ be a solution to* (1.6) *with $\mathcal{V}(f) \neq 0$. Then f is a Willmore immersion, i.e. a solution to* (1.6) *with $\lambda = 0$.*

We note that the assumptions in Lemma 2.6 are automatically satisfied if f is an embedding of a compact surface, since then $\mathcal{V}$ is exactly the volume of the domain enclosed by $f(\Sigma)$ by the divergence theorem. On the other hand, the example of an infinitely long cylinder shows that the statement of the lemma is no longer true without the compactness assumption.

Proof of Lemma 2.6. We observe that by the scaling invariance of the Willmore energy, we have $\overline{\mathcal{W}}(f+tf) = \overline{\mathcal{W}}(f)$ for all $|t| < 1$. Hence, by (2.11) (which also holds for variations which are not necessarily normal, see for instance [23, p. 11]), we find

$$0 = \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} \overline{\mathcal{W}}(f+tf) = \int \langle \nabla \overline{\mathcal{W}}(f), f\rangle \,\mathrm{d}\mu = \int \langle (\Delta H + |A^0|^2 H)\langle \nu, f\rangle \,\mathrm{d}\mu = 3\lambda\, \mathcal{V}(f),$$

using (1.6) and then (1.3) in the last step. As $\mathcal{V}(f) \neq 0$, this yields the claim. ☐

3. Localized energy estimates

In this section, we will use the interpolation inequalities developed in [21, 20]. As we shall see, control over the concentration of curvature and λ enables us to estimate derivatives of arbitrary order of the second fundamental form.

In the following, we restrict to a particular class of test functions. Let $\tilde{\gamma} \in C_c^\infty(\mathbb{R}^3)$ with $0 \leq \tilde{\gamma} \leq 1$ and $\|D\tilde{\gamma}\|_\infty \leq \Lambda, \|D^2\tilde{\gamma}\|_\infty \leq \Lambda^2$ for some $\Lambda > 0$. Then setting

$$\begin{aligned} &\gamma := \tilde{\gamma} \circ f \colon [0,T) \times \Sigma \to \mathbb{R} \text{ we find} \\ &|\nabla \gamma| \leq C\Lambda \text{ and } |\nabla^2 \gamma| \leq C\Lambda^2 + C|A|\Lambda, \end{aligned} \tag{3.1}$$

for a universal constant $C \in (0, \infty)$. Note that γ_t has compact support in space for all $0 \leq t < T$. The estimates in (3.1) follow by the identities

$$\begin{aligned} \nabla \gamma &= (D\tilde{\gamma} \circ f) Df \\ \nabla^2 \gamma &= (D^2 \tilde{\gamma} \circ f)(Df\cdot, Df\cdot) + (D\tilde{\gamma} \circ f) A(\cdot,\cdot). \end{aligned}$$

Unless specified otherwise, constants $C \in (0,\infty)$ are always universal and are allowed to change from line to line.

Following the strategy in [21, Secion 3], we can prove the following

Lemma 3.1. *Let $f\colon [0,T) \times \Sigma \to \mathbb{R}^3$ be a volume-preserving Willmore flow. We have*

$$\partial_t \int \frac{1}{2}|H|^2 \gamma^4 \,\mathrm{d}\mu + \frac{1}{2}\int |\nabla \overline{\mathcal{W}}(f)|^2 \gamma^4 \,\mathrm{d}\mu \leq C\Lambda^2 \int |A|^2 H^2 \gamma^2 \,\mathrm{d}\mu + C\Lambda^4 \int_{[\gamma>0]} H^2 \,\mathrm{d}\mu$$

$$+\lambda\int|A^0|^2H\gamma^4\,\mathrm{d}\mu+C\Lambda|\lambda|\int H^2\gamma^3\,\mathrm{d}\mu,$$

for some universal constant C with $0<C<\infty$ and

$$\partial_t\int|A^0|^2\gamma^4\,\mathrm{d}\mu+\frac{1}{2}\int|\nabla\overline{\mathcal{W}}(f)|^2\gamma^4\,\mathrm{d}\mu\leq C\Lambda^2\int|A^0|^2|A|^2\gamma^2\,\mathrm{d}\mu+C\Lambda^4\int_{[\gamma>0]}|A|^2\,\mathrm{d}\mu$$
$$+\lambda\int|A^0|^2H\gamma^4\,\mathrm{d}\mu+C\Lambda|\lambda|\int|A^0|^2\gamma^3\,\mathrm{d}\mu.$$

Proof. See Appendix B. ☐

Under the assumption of non-concentrated curvature, the following estimate by Kuwert and Schätzle allows us to locally control derivatives up to second order of the second fundamental form by the localized Willmore gradient and the localized energy. In the form stated below, it follows directly from [20, Proposition 2.6 and Lemma 4.2].

Proposition 3.2 ([20])**.** *There exist absolute constants $\varepsilon_0, C\in(0,\infty)$ such that if $f\colon\Sigma\to\mathbb{R}^3$ is an immersion with*

$$\int_{[\gamma>0]}|A|^2\,\mathrm{d}\mu<\varepsilon_0,$$

for some γ as in (3.1)*, then we have*

$$\int\left(|\nabla^2A|^2+|A|^2|\nabla A|^2+|A|^6\right)\gamma^4\,\mathrm{d}\mu\leq C\int|\nabla\overline{\mathcal{W}}(f)|^2\gamma^4\,\mathrm{d}\mu+C\Lambda^4\int_{[\gamma>0]}|A|^2\,\mathrm{d}\mu.$$

This will be the crucial tool in studying the volume-preserving Willmore flow if the *concentration of curvature* is controlled, cf. Sections 4 to 6. Note that in Lemma 3.1, a term involving λ and a cubic power of A occur. However, the energy decay only allows us to control square powers of A, hence we have to pay the price in terms of a higher power of the Lagrange multiplier.

Proposition 3.3. *Suppose $f\colon[0,T)\times\Sigma\to\mathbb{R}^3$ is a volume-preserving Willmore flow. If*

$$\int_{[\gamma>0]}|A|^2\,\mathrm{d}\mu<\varepsilon_0\quad\textit{at time } t\in[0,T),$$

where $\varepsilon_0>0$ is as in Proposition 3.2 and γ is as in (3.1)*, then we have*

$$\partial_t\int|A|^2\gamma^4\,\mathrm{d}\mu+c_0\int\left(|\nabla^2A|^2+|A|^2|\nabla A|^2+|A|^6\right)\gamma^4\,\mathrm{d}\mu$$
$$\leq C\Lambda^4\int_{[\gamma>0]}|A|^2\,\mathrm{d}\mu+C|\lambda|^{\frac{4}{3}}\int|A|^2\gamma^4\,\mathrm{d}\mu$$

at time t for some universal constants $c_0, C\in(0,\infty)$.

Proof. Combining Lemma 3.1, Proposition 3.2 and (2.3), we find

$$\partial_t \int |A|^2\gamma^4 \,\mathrm{d}\mu + c_0 \int \left(|\nabla^2 A|^2 + |A|^2|\nabla A|^2 + |A|^6\right)\gamma^4 \,\mathrm{d}\mu$$
$$\leq C\Lambda^2 \int |A|^4\gamma^2 \,\mathrm{d}\mu + C\Lambda^4 \int_{[\gamma>0]} |A|^2 \,\mathrm{d}\mu + C|\lambda| \int |A|^3\gamma^4 \,\mathrm{d}\mu + C\Lambda|\lambda| \int |A|^2\gamma^3 \,\mathrm{d}\mu.$$

For the first term on the right hand side above, for $\varepsilon > 0$ we estimate

$$\Lambda^2 \int |A|^4\gamma^2 \,\mathrm{d}\mu \leq \varepsilon \int |A|^6\gamma^4 \,\mathrm{d}\mu + C(\varepsilon)\Lambda^4 \int_{[\gamma>0]} |A|^2 \,\mathrm{d}\mu.$$

For the third term, we use Young's inequality with $p = 4, q = \frac{4}{3}$ to obtain

$$\lambda \int |A|^{\frac{3}{2}+\frac{3}{2}}\gamma^4 \,\mathrm{d}\mu \leq \varepsilon \int |A|^6\gamma^4 \,\mathrm{d}\mu + C(\varepsilon)|\lambda|^{\frac{4}{3}} \int |A|^2\gamma^4 \,\mathrm{d}\mu.$$

Similarly for the fourth term, we find

$$\Lambda|\lambda| \int |A|^{\frac{1}{2}+\frac{3}{2}}\gamma^3 \,\mathrm{d}\mu \leq C\Lambda^4 \int_{[\gamma>0]} |A|^2 \,\mathrm{d}\mu + C|\lambda|^{\frac{4}{3}} \int |A|^2\gamma^4 \,\mathrm{d}\mu.$$

Taking $\varepsilon > 0$ small enough and absorbing yields the claim. □

The integrated form of Proposition 3.3 will be particularly useful.

Corollary 3.4. *Let $f\colon [0,T) \times \Sigma \to \mathbb{R}^3$ be a volume-preserving Willmore flow such that for $\varepsilon_0 > 0$ as in Proposition 3.2 and γ as in (3.1) we have*

$$\int_{[\gamma>0]} |A|^2 \,\mathrm{d}\mu \leq \varepsilon < \varepsilon_0 \quad \textit{for all } 0 \leq t < T.$$

Then there exist universal constants $c_0, C \in (0,\infty)$ such that for all $0 \leq t < T$ we have

$$\int_{[\gamma=1]} |A|^2 \,\mathrm{d}\mu + c_0 \int_0^t \int_{[\gamma=1]} \left(|\nabla^2 A|^2 + |A|^2|\nabla A|^2 + |A|^6\right) \mathrm{d}\mu$$
$$\leq \int_{[\gamma_0>0]} |A_0|^2 \,\mathrm{d}\mu_0 + C\Lambda^4\varepsilon t + C\varepsilon \int_0^t |\lambda(\tau)|^{\frac{4}{3}} \,\mathrm{d}\tau. \tag{3.2}$$

Here we used the notation $\int_{[\gamma_0>0]} |A_0|^2 \,\mathrm{d}\mu_0 = \int_{[\gamma>0]} |A|^2 \,\mathrm{d}\mu\Big|_{t=0}$.

Note that in order to bound the left hand side of (3.2) up to time $t = T$, the control of the curvature concentration alone does not suffice. Recalling the nonlocal nature of the evolution (1.4), this is not entirely surprising. However, the above result shows that this lack of control can be compensated, if in addition we can bound the $L^{4/3}(0,T)$-norm of λ, a spatially global quantity, which behaves correctly under parabolic rescaling by Remark 2.2. We will discuss under which assumptions $\lambda \in L^{4/3}(0,T)$ can be guaranteed in Section 4.

As in [20], an appropriate higher order version of Corollary 3.4 can be used to prove higher order interior estimates.

Proposition 3.5. *Let $f\colon [0,T)\times\Sigma \to \mathbb{R}^3$ be a volume-preserving Willmore flow. Suppose $\rho > 0$ is chosen such that $T \leq T^*\rho^4$ for some $0 < T^* < \infty$ and*

$$\int_{B_\rho(x)} |A|^2 \,\mathrm{d}\mu \leq \varepsilon < \varepsilon_0 \quad \text{for all } 0 \leq t < T,$$

where $x \in \mathbb{R}^3$, $\varepsilon_0 > 0$ is as in Proposition 3.2 and $\int_0^T |\lambda|^{\frac{4}{3}} \,\mathrm{d}t \leq L < \infty$. Then, for all $m \in \mathbb{N}_0$ and $t \in (0,T)$ we have the estimates

$$\|\nabla^m A\|_{L^2(B_{\rho/2}(x))} \leq C(m, T^*, L)\sqrt{\varepsilon} t^{-\frac{m}{4}},$$
$$\|\nabla^m A\|_{L^\infty(B_{\rho/2}(x))} \leq C(m, T^*, L)\sqrt{\varepsilon} t^{-\frac{m+1}{4}}.$$

The proof of Proposition 3.5 is essentially the same as in [20, Theorem 3.5], so it is moved to Appendix C.

4. Integral estimates for the Lagrange multiplier

In Proposition 3.5, we were able to control all derivatives of the second fundamental form, if the concentration of curvature is sufficiently small and the Lagrange multiplier has some sort of integrability. This section is devoted to showing time integrability of λ under certain assumptions.

4.1. The $L^{4/3}(0,T)$-norm of λ in the case of non-concentration

First, we will control the $L^{4/3}$-norm of λ, which will be the key ingredient in the proof of the lifespan result in Theorem 1.1. We begin by making the following observation for immersions with non-concentrated curvature.

Lemma 4.1. *There exists an absolute constant $0 < \varepsilon_1 < 8\pi$ such that if $f\colon \Sigma \to \mathbb{R}^3$ is an immersion, $x_0 \in f(\Sigma)$ and $\rho > 0$ satisfies*

$$\int_{B_\rho(x_0)} |A|^2 \,\mathrm{d}\mu \leq \varepsilon < \varepsilon_1,$$

then we have $\rho \leq C\mathcal{A}(f)^{\frac{1}{2}}$, where $0 < C < \infty$ is an absolute constant.

Proof. By Simon's monotonicity formula [37, (1.4)] and (2.3) for any $x_0 \in f(\Sigma)$ and some universal constant $0 < C < \infty$ we have

$$\pi \leq C\left(\rho^{-2}\mu(f^{-1}(B_\rho(x_0))) + \int_{B_\rho(x_0)} |H|^2 \,\mathrm{d}\mu\right) \leq C\rho^{-2}\mu(f^{-1}(B_\rho(x_0))) + 2C\varepsilon_1.$$

For $\varepsilon_1 := \frac{\pi}{4C} > 0$ we thus find $\frac{\pi}{2} \leq C\rho^{-2}\mu(f^{-1}(B_\rho(x_0))) \leq C\rho^{-2}\mathcal{A}(f)$. □

Proposition 4.2. *Let $f\colon [0,T)\times\Sigma\to\mathbb{R}^3$ be a volume-preserving Willmore flow with $\mathcal{W}(f_0)\le K$ such that $\rho>0$ satisfies*

$$\sup_{0\le t\le T}\int_{B_\rho(x)}|A|^2\,\mathrm{d}\mu\le\varepsilon<\varepsilon_2\qquad\text{for all } x\in\mathbb{R}^3,$$

where $\varepsilon_2:=\min\{\varepsilon_0,\varepsilon_1\}\in(0,8\pi)$, with ε_0 as in Proposition 3.2 and $\varepsilon_1>0$ as in Lemma 4.1. Then, we have

$$\int_0^t|\lambda|^{\frac{4}{3}}\,\mathrm{d}\tau\le C(K,\chi(\Sigma))\left(\frac{t^{\frac{1}{2}}}{\rho^2}+\frac{t}{\rho^4}\right)\qquad\text{for all } 0\le t<T.$$

Proof of Proposition 4.2. First, fix $x\in\mathbb{R}^3$. Let $\tilde\gamma\in C_c^\infty(\mathbb{R}^3)$ be a bump function with $\chi_{B_{\rho/2}(x)}\le\tilde\gamma\le\chi_{B_\rho(x)}$, $\|D\tilde\gamma\|_\infty\le\frac{C}{\rho}$ and $\|D^2\tilde\gamma\|_\infty\le\frac{C}{\rho^2}$. Therefore, $\gamma:=\tilde\gamma\circ f$ is as in (3.1) with $\Lambda=\frac{C}{\rho}$, and thus by integrating Proposition 3.3 from 0 to τ we find

$$\begin{aligned}\int_{B_{\rho/2}(x)}|A|^2\,\mathrm{d}\mu\Big|_{t=\tau}&+c_0\int_0^\tau\int_{B_{\rho/2}(x)}\left(|\nabla^2A|^2+|A|^2|\nabla A|^2+|A|^6\right)\mathrm{d}\mu\,\mathrm{d}t\\&\le\int_{B_\rho(x)}|A_0|^2\,\mathrm{d}\mu_0+\frac{C}{\rho^4}\int_0^\tau\int_{B_\rho(x)}|A|^2\,\mathrm{d}\mu\,\mathrm{d}t+C\int_0^\tau|\lambda|^{\frac{4}{3}}\int_{B_\rho(x)}|A|^2\,\mathrm{d}\mu\,\mathrm{d}t.\end{aligned}\tag{4.1}$$

It is possible to find $(x_\ell)_{\ell\in\mathbb{N}}\subset\mathbb{R}^3$ with $\mathbb{R}^3=\bigcup_{\ell\in\mathbb{N}}B_{\rho/2}(x_\ell)$ such that each point $y\in\mathbb{R}^3$ is contained in at most M of the balls $B_\rho(x_\ell)$, where $M>0$ is a universal constant, in particular independent of $\rho>0$. Therefore, choosing $x=x_\ell$ in (4.1) and summing over $\ell\in\mathbb{N}$ we find

$$\begin{aligned}\int_0^\tau\int|A|^6\,\mathrm{d}\mu\,\mathrm{d}t&\le\sum_\ell\int_0^\tau\int_{B_{\rho/2}(x_\ell)}|A|^6\,\mathrm{d}\mu\,\mathrm{d}t\\&\le M\int|A_0|^2\,\mathrm{d}\mu_0+\frac{CM}{\rho^4}\int_0^\tau\int|A|^2\,\mathrm{d}\mu\,\mathrm{d}t+CM\int_0^\tau|\lambda|^{\frac{4}{3}}\int|A|^2\,\mathrm{d}\mu\,\mathrm{d}t\end{aligned}$$

Now, by (2.4) we have $\int|A|^2\,\mathrm{d}\mu\le C(K,\chi(\Sigma))$ and hence

$$\int_0^\tau\int|A|^6\,\mathrm{d}\mu\,\mathrm{d}t\le C\left(K,\chi(\Sigma)\right)\left(1+\frac{\tau}{\rho^4}+\int_0^\tau|\lambda|^{\frac{4}{3}}\,\mathrm{d}t\right).\tag{4.2}$$

Thus, using (2.3), Hölder's inequality and Lemma 4.1 we find from (1.5)

$$\begin{aligned}\int_0^\tau|\lambda|^{\frac{4}{3}}\,\mathrm{d}t&\le C\int_0^\tau\mathcal{A}(f_t)^{-\frac{4}{3}}\left(\int|A|^3\,\mathrm{d}\mu\right)^{\frac{4}{3}}\mathrm{d}t\le C\int_0^\tau\mathcal{A}(f_t)^{-1}\left(\int|A|^4\,\mathrm{d}\mu\right)\mathrm{d}t\\&\le C\rho^{-2}\left(\int_0^\tau\int|A|^6\,\mathrm{d}\mu\,\mathrm{d}t\right)^{\frac{1}{2}}\left(\int_0^\tau\int|A|^2\,\mathrm{d}\mu\,\mathrm{d}t\right)^{\frac{1}{2}}.\end{aligned}$$

Therefore, using (2.4), the energy decay (2.14), (4.2) and Young's inequality, we find

$$\int_0^\tau|\lambda|^{\frac{4}{3}}\,\mathrm{d}t\le C(K,\chi(\Sigma))\frac{\tau^{\frac{1}{2}}}{\rho^2}\left(1+\frac{\tau^{\frac{1}{2}}}{\rho^2}+\left(\int_0^\tau|\lambda|^{\frac{4}{3}}\,\mathrm{d}t\right)^{\frac{1}{2}}\right)$$

$$\leq C(K, \chi(\Sigma)) \left(\frac{\tau^{\frac{1}{2}}}{\rho^2} + \frac{\tau}{\rho^4} \right) + \frac{1}{2} \int_0^\tau |\lambda|^{\frac{4}{3}} \, \mathrm{d}t. \qquad \square$$

4.2. An L^2-type estimate

In this section, we prove an $L^2(0,T)$-type estimate for λ, which will be crucial in the analysis of the blow-ups in Section 6. Since we rely on a *reverse isoperimetric inequality* [6], this is the first instance where we require the Willmore energy to be below 8π.
As a first step, we want to relate the diameter to the Lagrange multiplier. To that end, we use the different scaling of $\overline{\mathcal{W}}$ and $\mathcal{V}$ to obtain a different representation of λ, cf. [14, pp. 1236 – 1237] and also [30, Proof of Theorem 1.4].

Lemma 4.3. *Let $f\colon [0,T) \times \Sigma \to \mathbb{R}^3$ be a volume-preserving Willmore flow. Then for all $t \in [0,T)$ and any $p \in \mathbb{R}^3$ we have*

$$3\lambda\, \mathcal{V}(f_0) = -\int \langle \partial_t f, f \rangle \, \mathrm{d}\mu = -\int \langle \partial_t f, f - p \rangle \, \mathrm{d}\mu.$$

Proof. Fix $t \in [0,T)$. For $\alpha > 0$, consider the immersion $h_\alpha := p + \alpha(f_t - p)\colon \Sigma \to \mathbb{R}^3$. We then have $\overline{\mathcal{W}}(h_\alpha) = \overline{\mathcal{W}}(f_t)$, whereas $\mathcal{V}(h_\alpha) = \alpha^3\, \mathcal{V}(f_0)$. Thus, we find

$$\frac{\mathrm{d}}{\mathrm{d}\alpha}\bigg|_{\alpha=1} \left(\overline{\mathcal{W}}(h_\alpha) + \lambda(t)\, \mathcal{V}(h_\alpha)\right) = 0 + 3\lambda(t)\, \mathcal{V}(f_0),$$

whereas by the definition of $L^2(\mathrm{d}\mu_f)$-gradients we have

$$\frac{\mathrm{d}}{\mathrm{d}\alpha}\bigg|_{\alpha=1} \left(\overline{\mathcal{W}}(h_\alpha) + \lambda(t)\, \mathcal{V}(h_\alpha)\right) = \int \langle \nabla \overline{\mathcal{W}}(f) + \lambda \nabla\, \mathcal{V}(f), f - p \rangle \, \mathrm{d}\mu \bigg|_t .$$

Therefore, by (1.4) and Lemma 2.4 we have the identity

$$3\lambda\, \mathcal{V}(f_0) = -\int \langle \partial_t f, f - p \rangle \, \mathrm{d}\mu \quad \text{on } [0,T).$$

Picking $p = 0$ yields the first equality. $\square$

This finally enables us to prove the desired L^2-estimate.

Proposition 4.4. *Let $f\colon [0,T) \times \Sigma \to \mathbb{R}^3$ be a volume-preserving Willmore flow with $\mathcal{W}(f_0) \leq 8\pi - \delta$ for $\delta > 0$. Then, for all $0 \leq t < T$ we have*

$$\int_0^t \lambda^2(\tau) \mathcal{A}(f_\tau) \, \mathrm{d}\tau \leq C(\delta) \mathcal{W}(f_0)^2.$$

Proof. Observe that by (2.13), we have $|\mathcal{V}(f)| = |\mathcal{V}(f_0)|$. Picking some $p(t) \in f_t(\Sigma)$ for each t, we find by Lemma 4.3 and Cauchy–Schwarz

$$|\lambda(t)| \leq \frac{1}{3|\mathcal{V}(f_0)|} \int |\partial_t f_t| \, \mathrm{d}\mu_t \operatorname{diam} f_t(\Sigma) \leq \frac{\mathcal{A}(f_t)^{\frac{1}{2}}}{3|\mathcal{V}(f_0)|} \left(\int |\partial_t f_t|^2 \, \mathrm{d}\mu_t \right)^{\frac{1}{2}} \operatorname{diam} f_t(\Sigma).$$

Squaring this inequality, by Simon's diameter estimate [37, Lemma 1.1] we conclude

$$\lambda^2(t)\mathcal{A}(f_t) \leq \frac{C}{|\mathcal{V}(f_0)|^2}\mathcal{A}(f_t)^3\mathcal{W}(f_t)\int |\partial_t f_t|^2 \,\mathrm{d}\mu_t.$$

Now, by the reverse isoperimetric inequality [6, Theorem 1.1] (see also [36, Theorem 1] for the spherical case), and the assumption on the initial energy, we have

$$\lambda^2(t)\mathcal{A}(f_t) \leq C(\delta)\mathcal{W}(f_t)\int |\partial_t f_t|^2 \,\mathrm{d}\mu_t.$$

Integrating in time and using (2.14) and (1.2) we conclude

$$\begin{aligned}\int_0^\tau \lambda^2(t)\mathcal{A}(f_t)\,\mathrm{d}t &\leq C(\delta)\int_0^\tau \mathcal{W}(f_t)\left(-\int \langle \nabla\overline{\mathcal{W}}(f_t), \partial_t f_t\rangle \,\mathrm{d}\mu_t\right)\mathrm{d}t \\ &= -C(\delta)\int_0^\tau \mathcal{W}(f_t)\partial_t\overline{\mathcal{W}}(f_t)\,\mathrm{d}t = C(\delta)\int_0^\tau -\partial_t\left(\mathcal{W}(f_t)\right)^2\mathrm{d}t \leq C(\delta)\mathcal{W}(f_0)^2.\end{aligned}$$

Renaming τ into t yields the claim. □

5. Proof of the lifespan theorem

In this section, we will prove Theorem 1.1, which yields a lower bound on the maximal existence time of the volume-preserving Willmore flow. This will be crucial for the construction of the blow-up in Section 6.
Here we only work with the integrability of the constraint parameter λ which we proved in Section 4 and do not require strong L^∞-type bounds as in [32, (A1)], [39, (7)].

Proof of Theorem 1.1. This can now be achieved in the same fashion as [21, Theorem 1.2], so we focus on the differences arising from the Lagrange multiplier. Without loss of generality, $\rho = 1$, cf. Remark 2.2. If $\Gamma > 1$ denotes the number of radius $\frac{1}{2}$ balls necessary to cover $B_1(0) \subset \mathbb{R}^3$, we set $\bar{\varepsilon} := \frac{\varepsilon_2}{3\Gamma}$ with $\varepsilon_2 > 0$ as in Proposition 4.2. We observe

$$\varepsilon(t) := \sup_{x\in\mathbb{R}^3}\int_{B_1(x)} |A|^2 \,\mathrm{d}\mu \leq \Gamma \cdot \sup_{x\in\mathbb{R}^3}\int_{B_{1/2}(x)} |A|^2 \,\mathrm{d}\mu, \tag{5.1}$$

and, for a parameter $0 < \beta < 1$, to be specified below, define

$$t_0 := \sup\left\{0 \leq t \leq \min\{T, \beta\} \mid \varepsilon(\tau) \leq 3\Gamma\varepsilon \text{ for all } 0 \leq \tau < t\right\} > 0. \tag{5.2}$$

Picking an appropriate test function in Corollary 3.4, we obtain

$$\int_{B_{1/2}(x)} |A|^2 \,\mathrm{d}\mu \leq \int_{B_1(x)} |A_0|^2 \,\mathrm{d}\mu_0 + 3c\Gamma\Lambda^4\varepsilon t + 3c\Gamma\varepsilon\int_0^t |\lambda|^{\frac{4}{3}}(\tau)\,\mathrm{d}\tau$$

for all $0 \leq t < t_0$ where $c, \Lambda \in (0,\infty)$ are universal constants. By the choice of $\bar{\varepsilon}$ and Proposition 4.2, the integral $\int_0^t |\lambda|^{\frac{4}{3}} \,\mathrm{d}\tau$ grows less than linearly in $t \in [0, t_0)$. Hence, by a suitable application of Young's inequality, we find

$$\begin{aligned}\int_{B_{1/2}(x)} |A|^2 \,\mathrm{d}\mu &\leq \int_{B_1(x)} |A_0|^2 \,\mathrm{d}\mu_0 + 3c\Gamma\Lambda^4\varepsilon t + \frac{\varepsilon}{2} + C(K, \chi(\Sigma), c, \Gamma)t\varepsilon \\ &\leq \int_{B_1(x)} |A_0|^2 \,\mathrm{d}\mu_0 + \frac{\varepsilon}{2} + C(K, \chi(\Sigma), c, \Gamma, \Lambda)t\varepsilon. \end{aligned} \tag{5.3}$$

If we choose $\beta^{-1} := 2C(K, \chi(\Sigma), c, \Gamma, \Lambda)$, the assumption $t_0 < \min\{T, \beta\}$ contradicts maximality of t_0 in (5.2). Consequently, $t_0 = \min\{T, \beta\}$ has to hold. If $t_0 = \beta$, we find $T \geq \beta$. In this case, (1.7) then follows from (5.1), (5.3) and the definition of β.
Assume $t_0 = T \leq \beta$. Then from (5.3) we find $\int_{B_{1/2}(x)} |A|^2 \,\mathrm{d}\mu \leq 2\varepsilon$, hence by (5.1), we have

$$\varepsilon(t) \leq 2\Gamma\varepsilon < \varepsilon_0 \quad \text{for all } 0 \leq t < t_0. \tag{5.4}$$

Now, $T \leq \beta$ by assumption and $L = \int_0^T |\lambda|^{\frac{4}{3}} \,\mathrm{d}t \leq C(K, \chi(\Sigma))$ by Proposition 4.2. We may use Proposition 3.5 and argue exactly as in [21, Theorem 1.2] to prove $f(t) \to f(T)$ smoothly as $t \nearrow T$, which enables us to smoothly extend the flow past T. Taking $\hat{c} \in (0, \beta) \subset (0, 1)$ small enough, (5.4) guarantees that (1.7) is satisfied. $\square$

6. Construction of the blow-up

In this section, we will rescale a volume-preserving Willmore flow as we approach the maximal existence time to obtain a blow-up limit, combining the approaches in [20, Section 4] and [22, pp. 348 – 349]. As we shall see, if the Lagrange multiplier has a certain integrability in time, then the limit is not only stationary, but even an *unconstrained* Willmore immersion.

Definition 6.1. *For a smooth family of immersions $f\colon [0,T)\times\Sigma \to \mathbb{R}^3$, $t \in [0,T), r > 0$, we define the* curvature concentration function

$$\kappa(t, r) := \sup_{x\in\mathbb{R}^3} \int_{B_r(x)} |A_t|^2 \,\mathrm{d}\mu_t.$$

Theorem 6.2. *Let $f\colon [0,T) \times \Sigma \to \mathbb{R}^3$ be a maximal volume-preserving Willmore flow with initial energy $\mathcal{W}(f_0) \leq K$. Let $(t_j)_{j\in\mathbb{N}} \subset [0,T), t_j \nearrow T, (r_j)_{j\in\mathbb{N}} \subset (0,\infty)$, $(x_j)_{j\in\mathbb{N}} \subset \mathbb{R}^3$ such that*

$$\kappa(t_j, r_j) \leq \varepsilon_3 := \bar{\varepsilon}\hat{c} \quad \textit{for all } j \in \mathbb{N}, \tag{6.1}$$

where $\bar{\varepsilon} > 0$ and $\hat{c} = \hat{c}(K, \chi(\Sigma)) \in (0,1)$ are as in Theorem 1.1. Then we find

$$t_j + r_j^4\hat{c} < T, \tag{6.2}$$

and after passing to a subsequence, the rescaled and translated immersions

$$\hat{f}_j := r_j^{-1}\left(f(t_j + r_j^4\hat{c},\cdot) - x_j\right)$$

converge as $j \to \infty$ smoothly on compact subsets of $\mathbb{R}^3$, after reparametrization, to a proper constrained Willmore immersion $\hat{f}\colon \hat{\Sigma} \to \mathbb{R}^3$ of (1.6) *with $\mathcal{W}(\hat{f}) \leq K$. Moreover, if $\int_0^T \lambda(t)^2\mathcal{A}(f_t)\,\mathrm{d}t < \infty$, then $\hat{f}$ is an* unconstrained *Willmore immersion.*

Note that while we cannot apply Lemma 2.6 to the limit immersion, under the L^2-integrability condition above, we still find that the Willmore part of the evolution dominates in the blow-up.

Remark 6.3. *By Proposition 4.4, the condition $\int_0^T \lambda(t)^2\mathcal{A}(f_t)\,\mathrm{d}t < \infty$ is automatically satisfied if $\mathcal{W}(f_0) < 8\pi$, i.e. if $K < 8\pi$ in Theorem 6.2.*

Remark 6.4. *For general sequences $(t_j)_{j\in\mathbb{N}}, (r_j)_{j\in\mathbb{N}}$ and $(x_j)_{j\in\mathbb{N}}$, the limit may be trivial, for instance, $\hat{\Sigma} = \emptyset$, if $\hat{f}_j$ parametrizes the round spheres $\partial B_1(x_j)$ with $x_j \to \infty$. In order to make use of the construction, we will select t_j and x_j such that this cannot happen.*

Any constrained Willmore immersion $\hat{f}\colon \hat{\Sigma} \to \mathbb{R}^3$ which arises from the process described in Theorem 6.2 is called a *concentration limit*. More precisely, we call $\hat{f}$ a *blow-up* if $r_j \to 0$, a *blow-down* for $r_j \to \infty$ and a *limit under translation* if $r_j \to r \in (0,\infty)$. Note that by (6.2) the last two can only occur if $T = \infty$.

Proof of Theorem 6.2. For $j \in \mathbb{N}$, we consider the rescaled and translated flows

$$\begin{aligned} f_j\colon [-r_j^{-4}t_j, r_j^{-4}(T-t_j)) \times \Sigma \to \mathbb{R}^3, \\ f_j(t,p) = r_j^{-1}\left(f(t_j + r_j^4 t, p) - x_j\right) \end{aligned}$$

and observe that f_j is a volume-preserving Willmore flow with initial datum given by $f_j(0) = r_j^{-1}\left(f(t_j,\cdot) - x_j\right)$ and maximal existence time $r_j^{-4}(T-t_j)$. In particular by Remark 2.5 we have $\mathcal{W}(f_j(0)) \leq K$ for any $j \in \mathbb{N}$. Moreover, by (6.1) we have

$$\sup_{x\in\mathbb{R}^3}\int_{B_1(x)} |A_{f_j(0,\cdot)}|^2\,\mathrm{d}\mu_{f_j(0,\cdot)} = \sup_{x\in\mathbb{R}^3}\int_{B_{r_j}(x)} |A_{f_{t_j}}|^2\,\mathrm{d}\mu_{f_{t_j}} \leq \varepsilon_3.$$

Hence, by Theorem 1.1 the maximal existence time of the flow f_j is bounded from below by $\hat{c} = \hat{c}(K,\chi(\Sigma))$ and (6.2) follows. Furthermore, (1.7) yields

$$\sup_{x\in\mathbb{R}^3}\int_{B_1(x)} |A_{f_j(t,\cdot)}|^2\,\mathrm{d}\mu_{f_j(t,\cdot)} \leq \bar{\varepsilon} < \varepsilon_2 \quad \text{for all } 0 \leq t \leq \hat{c},$$

using that $\bar{\varepsilon} < \varepsilon_2$ by definition (cf. Proof of Theorem 1.1), where $\varepsilon_2 > 0$ is as in Proposition 4.2. Consequently, by Proposition 4.2 the $L^{4/3}(0,\hat{c})$-norm of the Lagrange multiplier of f_j is bounded by $C(K,\chi(\Sigma))$ for any $j \in \mathbb{N}$. Therefore, using $\bar{\varepsilon} < \varepsilon_2 \leq \varepsilon_0$ (cf. Proposition 4.2) by Proposition 3.5 we find

$$\|\nabla^m A_{f_j(t,\cdot)}\|_\infty \leq C(m,K,\chi(\Sigma))t^{-\frac{m+1}{4}} \quad \text{for } 0 < t \leq \hat{c}. \tag{6.3}$$

Moreover, using the scale-invariance of the Willmore energy, (2.14) and the a-priori energy bound, we can use Simon's monotonicity formula [37] to conclude that

$$R^{-2}\mu_{f_j(t,\cdot)}(B_R(0)) \leq C(K,\chi(\Sigma)) < \infty \quad \text{for all } 0 < t \leq \hat{c}, R > 0.$$

Thus, we may apply Theorem A.1 and Corollary A.2 to the sequence of immersions $\hat{f}_j := f_j(\hat{c},\cdot)$.
After passing to a subsequence, we thus find a proper limit immersion $\hat{f}\colon \hat{\Sigma} \to \mathbb{R}^3$, where $\hat{\Sigma}$ is a complete surface without boundary, diffeomorphisms $\phi_j\colon \hat{\Sigma}(j) \to U_j$, where $U_j \subset \Sigma$ are open sets and $\hat{\Sigma}(j) = \{p \in \hat{\Sigma} \mid |\hat{f}(p)| < j\}$, and functions $u_j \in C^\infty(\hat{\Sigma}(j);\mathbb{R}^3)$ such that we have

$$\hat{f}_j \circ \phi_j = \hat{f} + u_j \quad \text{on } \hat{\Sigma}(j)$$

as well as $\|\hat{\nabla}^m u_j\|_{L^\infty(\hat{\Sigma}(j),\hat{g})} \to 0$ for $j \to \infty$ for all $m \in \mathbb{N}_0$.
For $j \in \mathbb{N}$, we now define the flows $\tilde{f}_j := f_j \circ \phi_j := f_j(\cdot, \phi_j(\cdot))\colon (0,\hat{c}] \times \hat{\Sigma}(j) \to \mathbb{R}^3$ and observe that they also satisfy the curvature estimates (6.3).
We use (6.3) to estimate

$$|\lambda(f_j)| \leq C\|A_{f_j}\|_\infty^3 \leq C(K,\chi(\Sigma),\xi). \tag{6.4}$$

Now, using (1.5), (6.3) and (6.4), it is not difficult to also bound $|\partial_t\lambda(f_j(t,\cdot))|$ and thus $\|\partial_t\tilde{f}_j(t,\cdot)\|_{L^\infty(\hat{\Sigma}(j))}$ for $t \in [\xi,\hat{c}]$, $j \in \mathbb{N}$ by $C(K,\chi(\Sigma),\xi)$. From here on, it is a standard procedure to establish L^∞-bounds in a local chart (U,ψ) of $\hat{\Sigma}$, i.e. estimates of the form

$$\|\partial^m\partial_t\tilde{f}_j\|_{L^\infty(U)} + \|\partial^{m+1}\tilde{f}_j\|_{L^\infty(U)} \leq C(m,K,\chi(\Sigma),\xi) \qquad \text{for all } t \in [\xi,\hat{c}], m \geq 0,$$

for all $j \geq J$ sufficiently large, where ∂ denotes the coordinate derivative in the chart (U,ψ), see for instance [21, p. 331–332]. By (2.1), this also transfers to estimates of the induced normal field $\nu_{\tilde{f}_j} := \nu_{f_j} \circ \phi_j$.
Moreover, using the scale-invariance, cf. Remark 2.2, and the invariance under reparametrization, we have the evolution

$$\partial_t\tilde{f}_j = -\nabla\overline{\mathcal{W}}(\tilde{f}_j) + \lambda(f_j)\nu_{\tilde{f}_j}. \tag{6.5}$$

Using the established bounds and the evolution (6.5), it is not difficult to see that $\tilde{f}_j$ converges in $C^1([\xi,\hat{c}];C^m(P;\mathbb{R}^3))$ for all $P \subset \hat{\Sigma}$ compact and for all $m \in \mathbb{N}$ to a limit flow $f_{lim}\colon [\xi,\hat{c}] \times \hat{\Sigma} \to \mathbb{R}^3$ and $\lambda(f_j) \to \lambda_{lim}$ in $C^0([\xi,\hat{c}];\mathbb{R})$ as $j \to \infty$, after passing to a subsequence.
Fix $P \subset \hat{\Sigma}$ compact and let $j \in \mathbb{N}$ be large enough. Then, using (6.5), (2.13) and (2.14)

$$\int_\xi^{\hat{c}}\int_P |\partial_t\tilde{f}_j|^2 \,\mathrm{d}\mu_{\tilde{f}_j}\,\mathrm{d}t \leq \int_\xi^{\hat{c}}\int_\Sigma \langle -\nabla\overline{\mathcal{W}}(f_j) + \lambda(f_j)\nu_{f_j}, \partial_t f_j\rangle \,\mathrm{d}\mu_{f_j}\,\mathrm{d}t = \int_\xi^{\hat{c}} (-\partial_t\overline{\mathcal{W}}(f_j))\,\mathrm{d}t.$$

In particular, using the convergence $\tilde{f}_j \to f_{lim}$ in $C^1([\xi,\hat{c}];C^m(P;\mathbb{R}^3))$, we find

$$\int_\xi^{\hat{c}}\int_P |\partial_t f_{lim}|^2 \,\mathrm{d}\mu_{f_{\mathrm{lim}}}\,\mathrm{d}t \leq \lim_{j\to\infty}\left(\overline{\mathcal{W}}(f(t_j + r_j^4\xi,\cdot)) - \overline{\mathcal{W}}(f(t_j + r_j^4\hat{c},\cdot))\right) = 0, \tag{6.6}$$

by scale-invariance and monotonicity of the energy Consequently, f_{lim} is constant in time, hence $f_{lim} \equiv f_{lim}(\hat{c}, \cdot) = \lim_{j\to\infty} f_j(\hat{c}, \phi_j(\cdot)) = \lim_{j\to\infty} \hat{f}_j \circ \phi_j = \hat{f}$. We observe that $\hat{\nu} := \nu_{lim}(\hat{c}, \cdot)$ is a global and smooth normal vector field on $\hat{\Sigma}$ and hence $\hat{\Sigma}$ is orientable. Setting $\hat{\lambda} := \lambda_{lim}(\hat{c})$ and using (6.5) we find

$$-\nabla\overline{\mathcal{W}}(\hat{f}) + \hat{\lambda}\hat{\nu} = \lim_{j\to\infty} \partial_t \tilde{f}_j(\hat{c}, \cdot) = \partial_t f_{lim}(\hat{c}, \cdot) = 0 \quad \text{on } \hat{\Sigma},$$

so $\hat{f}$ solves (1.6) and hence is a constrained Willmore immersion. In addition, the lower semicontinuity of the Willmore functional $\mathcal{W}$ with respect to smooth convergence on compact sets (which is discussed in [12, Appendix B] for instance) yield

$$\mathcal{W}(\hat{f}) \leq \liminf_{j\to\infty} \mathcal{W}(\hat{f}_j) \leq \mathcal{W}(f_0) \leq K.$$

For the "moreover" part of the theorem, we note that since f_j is a volume-preserving Willmore flow we find by (1.4) and (1.5)

$$\begin{aligned}
\int_\xi^{\hat{c}} \int_P |\nabla\overline{\mathcal{W}}(\tilde{f}_j)|^2 \,\mathrm{d}\mu_{\tilde{f}_j} \,\mathrm{d}t &\leq \int_\xi^{\hat{c}} \int_\Sigma |\nabla\overline{\mathcal{W}}(f_j)|^2 \,\mathrm{d}\mu_{f_j} \,\mathrm{d}t \\
&= \int_\xi^{\hat{c}} \int_\Sigma \left\langle -\partial_t f_j + \lambda(f_j)\nu_{f_j}, \nabla\overline{\mathcal{W}}(f_j)\right\rangle \mathrm{d}\mu_{f_j} \,\mathrm{d}t \\
&= -\int_\xi^{\hat{c}} \partial_t \overline{\mathcal{W}}(f_j) \,\mathrm{d}t + \int_\xi^{\hat{c}} |\lambda(f_j)|^2 \mathcal{A}(f_j) \,\mathrm{d}t. \qquad (6.7)
\end{aligned}$$

As in (6.6), the first term goes to zero as $j \to \infty$. For the second term, note that by (1.5), λ scales by $\lambda(r^{-1}f) = r^3\lambda(f)$ for $r > 0$. Therefore, we find

$$\begin{aligned}
&\int_\xi^{\hat{c}} |\lambda(r_j^{-1} f(t_j + r_j^4 t, \cdot))|^2 \mathcal{A}(r_j^{-1} f(t_j + r_j^4 t, \cdot)) \,\mathrm{d}t \\
&\quad = \int_{t_j + r_j^4 \xi}^{t_j + r_j^4 \hat{c}} |r_j^3 \lambda(f(\tau, \cdot))|^2 r_j^{-2} \mathcal{A}(f(\tau, \cdot)) r_j^{-4} \,\mathrm{d}\tau = \int_{t_j + r_j^4 \xi}^{t_j + r_j^4 \hat{c}} |\lambda(f_\tau)|^2 \mathcal{A}(f_\tau) \,\mathrm{d}\tau,
\end{aligned}$$

after a change of variables. Recall that by assumption $\int_0^T \lambda^2 \mathcal{A} \,\mathrm{d}t < \infty$, so the second term in (6.7) also goes to zero as $j \to \infty$ using dominated convergence. Consequently, by (6.7), we have

$$\int_\xi^{\hat{c}} \int_P |\nabla\overline{\mathcal{W}}(f_{lim})|^2 \,\mathrm{d}\mu_{f_{lim}} \,\mathrm{d}t = \lim_{j\to\infty} \int_\xi^{\hat{c}} \int_P |\nabla\overline{\mathcal{W}}(\tilde{f}_j)|^2 \,\mathrm{d}\mu_{\tilde{f}_j} \,\mathrm{d}t = 0.$$

Since $f_{lim}(t, \cdot) \equiv \hat{f}$ and as P was arbitrary, we conclude $\nabla\overline{\mathcal{W}}(\hat{f}) = 0$, so $\hat{f}$ is a Willmore immersion. ☐

We can choose $(t_j)_{j\in\mathbb{N}}, (r_j)_{j\in\mathbb{N}}, (x_j)_{j\in\mathbb{N}}$ such that the concentration limit is nontrivial, even if $T = \infty$. The argument is exactly as in [22, p. 348–349], so the proof can safely be omitted.

Proposition 6.5. *Let $f\colon [0,T) \times \Sigma \to \mathbb{R}^3$ be a volume-preserving Willmore flow with $0 < T \leq \infty$. Then, we can choose sequences $t_j \nearrow T$, $(r_j)_{j\in\mathbb{N}} \subset (0,\infty)$ and $(x_j)_{j\in\mathbb{N}} \subset \mathbb{R}^3$ satisfying* (6.1) *such that the concentration limit $\hat{f}\colon \hat{\Sigma} \to \mathbb{R}^3$ from Theorem 6.2 satisfies*

$$\int_{\overline{B_1(0)}} |A_{\hat{f}}|^2 \, \mathrm{d}\mu_{\hat{f}} > 0,$$

in particular $\hat{\Sigma} \neq \emptyset$.

7. Convergence for compact concentration limits

The main result of this section is the following

Theorem 7.1. *Let $f\colon [0,T) \times \Sigma \to \mathbb{R}^3$ be a volume-preserving Willmore flow and let $\hat{f}\colon \hat{\Sigma} \to \mathbb{R}^3$ be a concentration limit with $\hat{\Sigma} \neq \emptyset$. If $\hat{\Sigma}$ has a compact component and*

(i) $\mathcal{V}(\hat{f}) \neq 0$ or

(ii) $\mathcal{V}(\hat{f}) = \mathcal{V}(f_0) = 0$,

then $\hat{f}$ is a limit under translation. Moreover, the flow exists globally and converges, as $t \to \infty$, after reparametrization by diffeomorphisms, to a constrained Willmore immersion f_∞ with $\mathcal{W}(f_\infty) = \mathcal{W}(\hat{f})$.

Under certain assumptions, the first part of the statement can also be directly obtained from the scaling behavior of the volume.

Remark 7.2. *Under the assumption that $\hat{\Sigma}$ is compact, we have $\mathcal{V}(\hat{f}) = \lim_{j\to\infty} r_j^{-3} V_0$ which immediately yields that*

(i) if $\mathcal{V}(\hat{f}) \neq 0$, then $\hat{f}$ cannot be a blow-up or a blow-down;

(ii) if $V_0 \neq 0$, then $\hat{f}$ cannot be a blow-up.

Clearly, these arguments fail if $\mathcal{V}(\hat{f}) = V_0 = 0$.

The key ingredient to prove the powerful convergence result Theorem 7.1 relies on a suitable extension of the *Łojasiewicz–Simon gradient inequality.*

7.1. The constrained Łojasiewicz–Simon gradient inequality

In this subsection, we will state and prove a *constrained* or *refined* Łojasiewicz–Simon gradient inequality, cf. [34], for the volume-preserving Willmore flow. A similar result for the length-preserving elastic flow of curves was recently proven in [35].
The strategy is the same as in [10, Section 3]. First, in order to get rid of the invariance of the Willmore and volume energy, we restrict ourselves to normal variations. Throughout

this section we will fix some smooth immersion $f\colon \Sigma \to \mathbb{R}^3$. The *normal Sobolev spaces* along f are defined by

$$W^{k,2}(\Sigma;\mathbb{R}^3)^\perp := \{\phi \in W^{k,2}(\Sigma;\mathbb{R}^3) \mid P^\perp\phi = \phi\},$$

for $k \in \mathbb{N}_0$, with $L^2(\Sigma;\mathbb{R}^3)^\perp := W^{0,2}(\Sigma;\mathbb{R}^3)^\perp$. Note that $L^2(\Sigma;\mathbb{R}^3)^\perp$ is a Hilbert space with inner product

$$\langle \phi_1, \phi_2\rangle_{L^2(\Sigma;\mathbb{R}^3)^\perp} = \int_\Sigma \langle \phi_1, \phi_2\rangle \,\mathrm{d}\mu_f \quad \text{for } \phi_1, \phi_2 \in L^2(\Sigma;\mathbb{R}^3)^\perp. \tag{7.1}$$

Remark 7.3. *(i) Note that since we are in codimension one, we have*

$$W^{k,2}(\Sigma;\mathbb{R}^3)^\perp = \{u\nu_f \mid u \in W^{k,2}(\Sigma)\},$$

for $k \in \mathbb{N}_0$, where ν_f is the unit normal to f and $W^{k,2}(\Sigma) := W^{k,2}(\Sigma;\mathbb{R})$. In fact, the map $W^{k,2}(\Sigma) \to W^{k,2}(\Sigma;\mathbb{R}^3)^\perp, u \mapsto \phi = u\nu_f$ is an isomorphism of Banach spaces and for $k = 0$ an isometry between the Hilbert spaces $L^2(\Sigma)$ and $L^2(\Sigma;\mathbb{R}^3)^\perp$.

(ii) Since Σ is compact, the spaces $W^{k,2}(\Sigma;\mathbb{R}^3)$ and $L^2(\Sigma;\mathbb{R}^3)$ do not depend on the metric, cf. [2, Theorem 2.20].

First, we prove a constrained Łojasiewicz–Simon gradient inequality in normal directions in a neighborhood of a constrained Willmore immersion, i.e. a solution to (1.6).

Proposition 7.4. *Let $f\colon \Sigma \to \mathbb{R}^3$ be a smooth constrained Willmore immersion. Then, there exists $C, \sigma > 0$ and $\theta \in (0, \frac{1}{2}]$ such that for all $\phi \in W^{4,2}(\Sigma;\mathbb{R}^3)^\perp$ with $\|\phi\|_{W^{4,2}} \le \sigma$ and $\mathcal{V}(f+\phi) = \mathcal{V}(f)$ we have*

$$|\overline{\mathcal{W}}(f+\phi) - \overline{\mathcal{W}}(f)|^{1-\theta} \le C\|\nabla\overline{\mathcal{W}}(f+\phi) - \lambda(f+\phi)\nu_{f+\phi}\|_{L^2(\mathrm{d}\mu_{f+\phi})},$$

where λ is as in (1.5).

Proposition 7.4 will follow from [10] and [34, Corollary 5.2]. To that end, we need to show the analyticity of certain maps and study their second variations. Most of the results will follow from [10] in the case of codimension one, only the volume needs to be studied in detail.

Lemma 7.5. *Let $U := B_\rho(0) \subset W^{4,2}(\Sigma)$. Then for $\rho > 0$ small enough and writing $f_u := f + u\nu_f$ for $u \in U$ we have*

(i) $f_u\colon \Sigma \to \mathbb{R}^3$ is an immersion and $U \to W^{4,2}(\Sigma;\mathbb{R}^3), u \mapsto f_u$ is analytic;

(ii) the map $U \to C^0(\Sigma;\mathbb{R}^3), u \mapsto \nu_{f_u}$ is analytic.

Proof. (i) Taking $\rho > 0$ small enough and using the Sobolev embedding $W^{4,2}(\Sigma) \hookrightarrow C^1(\Sigma)$ we find that f_u is an immersion for all $u \in U$. The map $u \mapsto f_u$ is linear and bounded, hence analytic.

(ii) In local coordinates (y^1, y^2) in the orientation on Σ, by (i) and the Sobolev embedding theorem $W^{4,2}(\Sigma) \hookrightarrow C^1(\Sigma)$, the map $B_\rho(0) \to C^1(\Sigma; \mathbb{R}^3), u \mapsto \partial_{y^1} f_u \times \partial_{y^2} f_u$ is bilinear and bounded, hence analytic. Moreover, since f_u is a C^1-immersion by (i), the denominator in the definition of ν_{f_u} in (2.1) is uniformly bounded away from zero. Since $\mathbb{R}^3 \setminus B_\delta(0) \to \mathbb{R}^3, x \mapsto \frac{x}{|x|}$ is analytic for any $\delta > 0$ the claim follows from the characterization of analytic Nemytskii operators on $C(\Sigma)$, cf. [1, Theorem 6.8]. □

Let $\tilde{U} := \{\phi \in W^{4,2}(\Sigma; \mathbb{R}^3)^\perp \mid \phi = u\nu_f \text{ with } u \in U\}$. By Remark 7.3 (i), $\tilde{U}$ is open in $W^{4,2}(\Sigma; \mathbb{R}^3)^\perp$. We consider the shifted energies, defined by

$$\begin{aligned} &W \colon \tilde{U} \to \mathbb{R}, W(\phi) := \overline{\mathcal{W}}(f + \phi), \\ &V \colon \tilde{U} \to \mathbb{R}, V(\phi) := \mathcal{V}(f + \phi). \end{aligned}$$

Lemma 7.6. *Under the assumptions of Lemma 7.5, the following maps are analytic:*

(i) the function $\tilde{U} \to C^0(\Sigma), \phi \mapsto \rho_{f+\phi}$, where $\mathrm{d}\mu_{f+\phi} = \rho_{f+\phi}\, \mathrm{d}\mu_f$;

(ii) the function $\tilde{U} \to \mathbb{R}, \phi \mapsto W(\phi)$;

(iii) the function $\tilde{U} \to L^2(\Sigma; \mathbb{R}^3)^\perp, \phi \mapsto P^\perp \nabla \overline{\mathcal{W}}(f + \phi)\rho_{f+\phi}$;

(iv) the function $\tilde{U} \to \mathbb{R}, \phi \mapsto V(\phi)$;

(v) the function $\tilde{U} \to L^2(\Sigma; \mathbb{R}^3)^\perp, \phi \mapsto P^\perp (-\nu_{f+\phi}\rho_{f+\phi})$.

Proof. Statement (i) is [10, Lemma 3.2 (vii)] and (ii) follows from [10, Lemma 3.2 (iv) and (vii)].
By [10, Lemma 3.2 (v) and (vi)], $\tilde{U} \to L^2(\Sigma; \mathbb{R}^3)^\perp, \phi \mapsto \nabla \overline{\mathcal{W}}(f + \phi)$ is analytic, and hence (iii) follows from (i).
Note that by Remark 7.3, $\tilde{U} \to U, \phi \mapsto u = \langle \phi, \nu_f \rangle$ is linear and bounded, thus analytic. Therefore, $\tilde{U} \to C^0(\Sigma; \mathbb{R}^3), \phi \mapsto \nu_{f+\phi}$ is analytic, hence so is V and by (i) statement (v) follows. □

As a last missing ingredient towards proving the constrained Łojasiewicz–Simon gradient inequality, we compute the first and second variations.

Lemma 7.7. *Let $H := L^2(\Sigma; \mathbb{R}^3)^\perp$. Under the assumption of Lemma 7.5, for each $\phi \in \tilde{U}$, the H-gradients of W and V are given by*

$$\begin{aligned} \nabla_H W(\phi) &= P^\perp \nabla \overline{\mathcal{W}}(f + \phi)\rho_{f+\phi}, \\ \nabla_H V(\phi) &= P^\perp (-\nu_{f+\phi}\rho_{f+\phi}) . \end{aligned} \tag{7.2}$$

Moreover, the Fréchet-derivatives of the H-gradient maps of W and V at $u = 0$ satisfy

$$\begin{aligned} (\nabla_H W)'(0) \colon W^{4,2}(\Sigma, \mathbb{R}^3)^\perp \to L^2(\Sigma, \mathbb{R}^3)^\perp \quad &\textit{is a Fredholm operator with index zero,} \\ (\nabla_H V)'(0) \colon W^{4,2}(\Sigma, \mathbb{R}^3)^\perp \to L^2(\Sigma, \mathbb{R}^3)^\perp \quad &\textit{is compact.} \end{aligned}$$

Proof. For $\phi, \psi \in \tilde{U}$, we have by the first variation of the Willmore energy and (7.1)

$$\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} W(\phi + t\psi) = \int_\Sigma \langle \nabla\overline{\mathcal{W}}(f+\phi), \psi\rangle \, \mathrm{d}\mu_{f+\phi} = \left\langle P^\perp \nabla\overline{\mathcal{W}}(f+\phi)\rho_{f+\phi}, \psi\right\rangle_H,$$

where we also used $\psi = P^\perp\psi$.
Similarly, $\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} V(\phi+t\psi) = -\int_\Sigma \langle \nu_{f+\phi}, \psi\rangle \, \mathrm{d}\mu_{f+\phi} = -\left\langle P^\perp \nu_{f+\phi}\rho_{f+\phi}, \psi\right\rangle_H$. The Fredholm property of $(\nabla_H W)'(0)$ follows from (1.2) and [10, Lemma 3.3 and p. 356]. For the last statement, we use (2.10) and Remark 7.3 (i) to obtain for $\phi = u\nu_f$

$$\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \nu_{f+t\phi} = -\operatorname{grad}_g u = -\operatorname{grad}_g\langle \phi, \nu_f\rangle.$$

Now, by (2.6), we find $\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \rho_{f+t\phi}\, \mathrm{d}\mu_f = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (\mathrm{d}\mu_{f+t\phi}) = -\langle H_f\nu_f, \phi\rangle \, \mathrm{d}\mu_f$. Using (7.2) we obtain, since the gradient term is tangential,

$$\begin{aligned}(\nabla_H V)'(0)\phi = -P^\perp \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (\nu_{f+t\phi}\rho_{f+t\phi}) &= P^\perp \operatorname{grad}_g\langle\phi, \nu_f\rangle + P^\perp \nu_f \langle H_f\nu_f, \phi\rangle \\ &= \nu_f\langle H_f\nu_f, \phi\rangle.\end{aligned}$$

As this is only of zeroth order in $\phi \in W^{4,2}(\Sigma;\mathbb{R}^3)^\perp$, the claim follows from the Rellich–Kondrachov Theorem, see for instance [2, Theorem 2.34]. $\square$

Proof of Proposition 7.4. We verify the assumptions of [34, Corollary 5.2] for the Hilbert space $W^{4,2}(\Sigma;\mathbb{R}^3)^\perp$ which embeds densely into $H = L^2(\Sigma;\mathbb{R}^3)^\perp$. The functionals W and V are analytic with analytic H-gradients in a neighborhood $\tilde{U}$ of zero by Lemma 7.6. By Lemma 7.7, the second variation of W at zero is Fredholm of index zero, whereas the second variation of V at zero is compact. Note that $\nabla_H V(0) \neq 0$ since we have

$$\langle \nabla_H V(0), \nu_f\rangle_H = -\int_\Sigma \langle \nu_f, \nu_f\rangle \, \mathrm{d}\mu_f = -\mathcal{A}(f) < 0.$$

Thus, by [34, Corollary 5.2], W satisfies a constrained Łojasiewicz–Simon gradient inequality near $\phi = 0$, i.e. there exist $C, \sigma > 0$ and $\theta \in (0, \frac{1}{2}]$ such that for all $\phi \in \tilde{U}$ with $\|\phi\|_{W^{4,2}} \leq \sigma$ and $V(\phi) = V(0)$, we have

$$|W(\phi) - W(0)|^{1-\theta} \leq C\|P_\phi \nabla_H W(\phi)\|_H,$$

where $P_\phi\colon H \to H$ is the H-orthogonal projection onto $\{y \in H \mid \langle y, \nabla V(\phi)\rangle_H = 0\}$, cf. [34, Proposition 3.3]. Thus, for $\lambda(f+\phi)$ as in (1.5), we find

$$\begin{aligned}\|P_\phi \nabla W(\phi)\|_H^2 = \|P_\phi(\nabla W(\phi) + \lambda(f+\phi)\nabla V(\phi))\|_H^2 &\leq \|\nabla W(\phi) + \lambda(f+\phi)\nabla V(\phi)\|_H^2 \\ = \int_\Sigma |\nabla\overline{\mathcal{W}}(f+\phi) - \lambda(f+\phi)\nu_{f+\phi}|^2 \rho_{f+\phi}\, \mathrm{d}\mu_{f+\phi}. \quad &(7.3)\end{aligned}$$

Now, by the Sobolev embedding theorem $W^{4,2}(\Sigma;\mathbb{R}^3) \hookrightarrow C^1(\Sigma;\mathbb{R}^3)$, we may bound $\|\rho_{f+\phi}\|_\infty$ for all $\|\phi\|_{W^{4,2}} \leq \sigma$. Using (7.3) and the definition of W and V yields the claim. $\square$

This finally yields the inequality for all directions.

Theorem 7.8. *Let $f\colon \Sigma \to \mathbb{R}^3$ be a constrained Willmore immersion. Then, there exist $C, \sigma > 0$ and $\theta \in (0, \frac{1}{2}]$ such that for all $h \in W^{4,2}(\Sigma; \mathbb{R}^3)$ with $\|h - f\|_{W^{4,2}} \leq \sigma$ and $\mathcal{V}(h) = \mathcal{V}(f)$ we have*

$$|\overline{\mathcal{W}}(h) - \overline{\mathcal{W}}(f)|^{1-\theta} \leq C\|\nabla\overline{\mathcal{W}}(h) - \lambda(h)\nu_h\|_{L^2(\mathrm{d}\mu_h)}.$$

Proof. Let C, σ, θ as in Proposition 7.4. Like in [10, p. 357], there exists $\sigma' > 0$ such that every $h \in W^{4,2}(\Sigma; \mathbb{R}^3)$ with $\|h - f\|_{W^{4,2}} \leq \sigma'$ can be written as $h \circ \Phi = f + \phi$ where $\Phi\colon \Sigma \to \Sigma$ is an orientation-preserving diffeomorphism and $\phi \in W^{4,2}(\Sigma; \mathbb{R}^3)^{\perp}$ with $\|\phi\|_{W^{4,2}} \leq \sigma$. Then, we have $\overline{\mathcal{W}}(h) = \overline{\mathcal{W}}(f + \phi)$ and $\mathcal{V}(h) = \mathcal{V}(f + \phi) = V(f)$ by invariance under diffeomorphism, and moreover by the geometric transformation of the L^2-norms

$$\|\nabla\overline{\mathcal{W}}(h) - \lambda(h)\nu_h\|_{L^2(\mathrm{d}\mu_h)} = \|\nabla\overline{\mathcal{W}}(f + \phi) - \lambda(f + \phi)\nu_{f+\phi}\|_{L^2(\mathrm{d}\mu_{f+\phi})}.$$

Renaming σ' into σ, the statement then follows from Proposition 7.4. ☐

7.2. An asymptotic stability result

The following stability result is an analogue of [10, Lemma 4.1].

Lemma 7.9. *Let $f_W\colon \Sigma \to \mathbb{R}^3$ be a constrained Willmore immersion and let $k \in \mathbb{N}$, $k \geq 4$, $\delta > 0$. Then there exists $\varepsilon = \varepsilon(f_W) > 0$ such that if $f\colon [0, T) \times \Sigma \to \mathbb{R}^3$ is a volume-preserving Willmore flow with $\mathcal{V}(f) \equiv \mathcal{V}(f_W)$ satisfying*

(i) $\|f_0 - f_W\|_{C^{k,\alpha}} < \varepsilon$ for some $\alpha > 0$;

(ii) $\overline{\mathcal{W}}(f(t)) \geq \overline{\mathcal{W}}(f_W)$ whenever $\|f(t) \circ \Phi(t) - f_W\|_{C^k} \leq \delta$, for some diffeomorphisms $\Phi(t)\colon \Sigma \to \Sigma$;

then, the flow exists globally, i.e. we may take $T = \infty$. Moreover, it converges, after reparametrization by some diffeomorphisms $\tilde{\Phi}(t)\colon \Sigma \to \Sigma$, smoothly to a constrained Willmore immersion f_∞, satisfying $\overline{\mathcal{W}}(f_W) = \overline{\mathcal{W}}(f_\infty)$ and $\|f_\infty - f_W\|_{C^k} \leq \delta$.

The proof of Lemma 7.9 is essentially a nonlocal version of the one of [10, Lemma 4.1], with the classical Łojasiewicz–Simon inequality replaced with the constrained one. It is thus moved to Appendix D. This finally enables us to prove Theorem 7.1.

Proof of Theorem 7.1. By Theorem 6.2, there are $t_j \nearrow T, r_j \to r \in [0, \infty]$ and $x_j \in \mathbb{R}^3$ for all $j \in \mathbb{N}$ such that $t_j + \hat{c}r_j^4 < T$ and

$$\hat{f}_j := r_j^{-1}\left(f(t_j + \hat{c}r_j^4, \cdot) - x_j\right) \to \hat{f} \tag{7.4}$$

smoothly, after reparametrization, on compact subsets of $\mathbb{R}^3$, where $\hat{f}\colon \hat{\Sigma} \to \mathbb{R}^3$ is a constrained Willmore immersion. By assumption, $\hat{\Sigma}$ contains a compact component and thus, by the same argument as in [20, Lemma 4.3], we may assume $\hat{\Sigma} = \Sigma$ is compact.

Consequently $\hat{f}_j \circ \Phi_j \to \hat{f}$ smoothly on Σ, where $\Phi_j \colon \Sigma \to \Sigma$ are diffeomorphisms. Note that now $\hat{f}$ is a constrained Willmore immersion of the compact surface $\hat{\Sigma} = \Sigma$. Thus, there exists $\varepsilon = \varepsilon(\hat{f})$ as in Lemma 7.9. We would like to apply Lemma 7.9 for the flow with the initial datum $\hat{f}_j \circ \Phi_j$, however, this might not have the correct volume. Under the assumptions of the theorem, we can fix that by another rescaling. Note that by smooth convergence and since Σ is compact, if $\mathcal{V}(\hat{f}) \neq 0$, then also $\mathcal{V}(\hat{f}_j \circ \Phi_j) \neq 0$ and Φ_j is orientation-preserving for all j sufficiently large. For such $j \in \mathbb{N}$, we define

$$v_j := \begin{cases} \left(\frac{\mathcal{V}(\hat{f})}{\mathcal{V}(\hat{f}_j \circ \Phi_j)}\right)^{\frac{1}{3}} & \text{if } \mathcal{V}(\hat{f}) \neq 0, \\ 1 & \text{if } \mathcal{V}(\hat{f}) = \mathcal{V}(f_0) = 0, \end{cases}$$

By smooth convergence and convergence of the volume, we have $v_j \to 1$ as $j \to \infty$, so we may assume $v_j \in (0,2)$ and

$$\|v_{j_0} \hat{f}_{j_0} \circ \Phi_{j_0} - \hat{f}\|_{C^{4,\alpha}} \leq |v_{j_0} - 1| \|\hat{f}_{j_0} \circ \Phi_{j_0}\|_{C^{4,\alpha}} + \|\hat{f}_{j_0} \circ \Phi_{j_0} - \hat{f}\|_{C^{4,\alpha}} < \varepsilon, \tag{7.5}$$

if we choose $j = j_0$ sufficiently large. We define $\bar{r}_{j_0} := v_{j_0}^{-1} r_{j_0} \in (0,\infty)$. By Remark 2.2, the flow

$$h_{j_0}(t,\cdot) := \bar{r}_{j_0}^{-1} \left(f(t_{j_0} + \bar{r}_{j_0}^4 t, \cdot) - x_{j_0} \right) \circ \Phi_{j_0}, \quad t \in [0, \bar{r}_{j_0}^{-4}(T - t_{j_0})),$$

is again a volume-preserving Willmore flow with $h_{j_0}(v_{j_0}^4 \hat{c}) = v_{j_0} \hat{f}_{j_0} \circ \Phi_{j_0}$ and volume $\mathcal{V}(h_{j_0}) \equiv \mathcal{V}(v_{j_0} \hat{f}_{j_0} \circ \Phi_{j_0}) = \mathcal{V}(\hat{f})$ by definition of v_{j_0}. Moreover, for $t \in [0, \bar{r}_{j_0}^{-4}(T - t_{j_0}))$, we have using monotonicity of the energy, the invariances of the Willmore energy and $t_k \nearrow T$

$$\overline{\mathcal{W}}(h_{j_0}(t)) \geq \lim_{s \to \bar{r}_{j_0}^{-4}(T - t_{j_0})} \overline{\mathcal{W}}(f(t_{j_0} + \bar{r}_{j_0}^4 s)) = \lim_{s \to T} \overline{\mathcal{W}}(f(s)) = \lim_{k \to \infty} \overline{\mathcal{W}}(\hat{f}_k) = \overline{\mathcal{W}}(\hat{f}).$$

The last equality holds since the convergence $\hat{f}_k \circ \Phi_k \to \hat{f}$ is smooth. This together with (7.5) yields that the assumptions of Lemma 7.9 are satisfied, and thus the flow h_{j_0} exists globally with

$$h_{j_0}(t) \circ \tilde{\Phi}(t) \to f_\infty \quad \text{smoothly as } t \to \infty,$$

where $\tilde{\Phi}(t) \colon \Sigma \to \Sigma$ are diffeomorphisms and f_∞ is a constrained Willmore immersion. Hence, f also exists globally, so we may take $T = \infty$. Moreover, for all $t \geq t_{j_0}$ we have

$$\begin{aligned} &f\left(t, \Phi_{j_0} \circ \tilde{\Phi}(\bar{r}_{j_0}^{-4}(t - t_{j_0}))\right) \\ &= \bar{r}_{j_0} h_{j_0}\left(\bar{r}_{j_0}^{-4}(t - t_{j_0}), \tilde{\Phi}(\bar{r}_{j_0}^{-4}(t - t_{j_0}))\right) + x_{j_0} \to \bar{r}_{j_0} f_\infty + x_{j_0} \end{aligned} \tag{7.6}$$

as $t \to \infty$ smoothly on Σ. It remains to show that $\hat{f}$ is a limit under translation. Let $r_j \to r \in [0,\infty]$. Picking $t := t_k + \hat{c} r_k^4$, $k \in \mathbb{N}$, in (7.6), we obtain for the diameters

$$d_k := \operatorname{diam} f(t_k + \hat{c} r_k^4)(\Sigma) \to \bar{r}_{j_0} \operatorname{diam} f_\infty(\Sigma), \quad \text{as } k \to \infty,$$

whence $\lim_{k\to\infty} d_k \in (0,\infty)$ since Σ is compact. On the other hand, using (7.4) we find

$$\operatorname{diam} \hat{f}(\hat{\Sigma}) = \lim_{k\to\infty} r_k^{-1} d_k \in (0,\infty),$$

as $\hat{\Sigma} \neq \emptyset$ is compact by assumption. Consequently, $\lim_{k\to\infty} r_k \in (0,\infty)$. $\square$

8. Convergence to the sphere

In this section, we will prove our main convergence result. While Σ was a general surface before, in this section we will work exclusively with $\Sigma = \mathbb{S}^2$. The key ingredients in proving Theorem 1.2 are the blow-up procedure, the classification of Willmore spheres in $\mathbb{R}^3$ due to Bryant [8], and a removability result for point singularities [22].

Proof of Theorem 1.2. Let $f\colon [0,T) \times \mathbb{S}^2 \to \mathbb{R}^3$ be a volume-preserving Willmore flow with initial datum f_0 with T maximal and $\mathcal{W}(f_0) \leq 8\pi$. If f_0 is a constrained Willmore immersion, then it is a Willmore immersion by Lemma 2.6, since $\mathcal{V}(f_0) \neq 0$. Hence it has to be a round sphere since by [8, Section 5], the critical values of Willmore immersions of spherical type are $4\pi d$ with $d \in \mathbb{N} \setminus \{2,3\}$ and the global minimizers are the round spheres [41]. In this case the result follows. If f_0 is not a constrained Willmore immersion, then the energy instantaneously drops below 8π by Remark 2.5, so we can assume $\mathcal{W}(f_0) < 8\pi$.
By Theorem 6.2, Remark 6.3 and Proposition 6.5, there exist $t_j \nearrow T, (r_j)_{j\in\mathbb{N}} \subset (0,\infty)$ and $(x_j)_{j\in\mathbb{N}} \subset \mathbb{R}^3$ such that the corresponding concentration limit $\hat{f}\colon \hat{\Sigma} \to \mathbb{R}^3$ is a unconstrained Willmore immersion satisfying

$$\int_{\hat{\Sigma}} |A_{\hat{f}}|^2 \,\mathrm{d}\mu_{\hat{f}} > 0. \tag{8.1}$$

Moreover, by Theorem 6.2 we have $\mathcal{W}(\hat{f}) < 8\pi$. Suppose $\hat{\Sigma}$ is not compact. There is $x_0 \notin \hat{f}(\hat{\Sigma})$ and with the inversion $I(x) := |x-x_0|^{-2}(x-x_0)$, we set $\bar{\Sigma} := I(\hat{f}(\hat{\Sigma})) \cup \{0\}$. By the removability result [22, Lemma 5.1], $\bar{\Sigma}$ is a smooth Willmore surface. Moreover, since $\hat{\Sigma}$ is complete by Corollary A.2, so is $\hat{f}(\hat{\Sigma})$. Hence, $\operatorname{dist}(x_0, \hat{f}(\hat{\Sigma})) > 0$ and consequently $\bar{\Sigma}$ is bounded. Using the definition of $\bar{\Sigma}$ and the completeness of $\hat{f}(\hat{\Sigma})$ again, it is not difficult to show that $\bar{\Sigma}$ is closed in $\mathbb{R}^3$ and thus compact. Furthermore, by [22, Lemma 5.1], we have $\mathcal{W}(\bar{\Sigma}) < 8\pi$ and $g(\bar{\Sigma}) = 0$ and hence $\bar{\Sigma}$ is a Willmore sphere. Using [8, 41] as above, we conclude that $\bar{\Sigma}$ has to be a round sphere. Since $\hat{f}(\hat{\Sigma})$ is not compact by assumption, this yields that $\hat{f}(\hat{\Sigma}) = I^{-1}(\bar{\Sigma})$ is a plane, contradicting (8.1).
Thus, $\hat{\Sigma}$ is compact, hence by arguing as in [20, Lemma 4.3], we can assume $\hat{\Sigma} = \Sigma = \mathbb{S}^2$. By [8, 41] and the Li–Yau inequality [27], we then have that $\hat{f}$ parametrizes an embedded round sphere, in particular $\mathcal{V}(\hat{f}) \neq 0$. Hence, Theorem 7.1 yields global existence and convergence to a constrained Willmore immersion f_∞ with $\mathcal{W}(f_\infty) = \mathcal{W}(\hat{f})$. By [41], f_∞ parametrizes a round sphere. Since the volume is preserved by (2.13), we conclude that $\mathcal{V}(\bar{r}_{j_0} f_\infty + x_{j_0}) = \mathcal{V}(f_0)$ and consequently the radius is $R := (\frac{3|\mathcal{V}(f_0)|}{4\pi})^{\frac{1}{3}} > 0$. $\square$

Appendix A Smooth convergence on compact sets

The essential tool in the construction of the blow-up in Theorem 6.2 was the following local version of Langer's compactness theorem [25] by Kuwert and Schätzle [20], see also [7] and [12, Appendix B] for some consequences of this notion of convergence.

Theorem A.1 ([20, Theorem 4.2])**.** *Let $f_j \colon \Sigma_j \to \mathbb{R}^3$ be a sequence of proper immersions, where Σ_j is a 2-manifold without boundary. Let $\Sigma_j(R) := \{p \in \Sigma_j \mid |f_j(p)| < R\}$ and assume the bounds*

$$\mu_j(\Sigma_j(R)) \leq C(R) \text{ for any } R > 0,$$
$$\|\nabla^m A_j\|_{L^\infty(\Sigma_j)} \leq C(m) \text{ for all } m \in \mathbb{N}_0.$$

Then, there exist a proper immersion $\hat{f} \colon \hat{\Sigma} \to \mathbb{R}^3$, where $\hat{\Sigma}$ is a 2-manifold without boundary, such that after passing to a subsequence we have a representation

$$f_j \circ \phi_j = \hat{f} + u_j \text{ on } \hat{\Sigma}(j) = \{p \in \hat{\Sigma} \mid |\hat{f}(p)| < j\}$$

with the following properties:

$$\begin{aligned}
&\phi_j \colon \hat{\Sigma}(j) \to U_j \subset \Sigma_j \text{ is a diffeomorphism,}\\
&\Sigma_j(R) \subset U_j \text{ if } j \geq j(R),\\
&u_j \in C^\infty(\hat{\Sigma}(j);\mathbb{R}^3) \text{ is normal along } \hat{f},\\
&\|\hat{\nabla}^m u_j\|_{L^\infty(\hat{\Sigma}(j))} \to 0 \text{ as } j \to \infty, \text{ for any } m \in \mathbb{N}_0.
\end{aligned}$$

Corollary A.2. *In Theorem A.1, the manifold $(\hat{\Sigma}, g_{\hat{f}})$ is complete.*

Proof. Suppose $(p_n)_{n\in\mathbb{N}} \subset \hat{\Sigma}$ is a Cauchy-Sequence with respect to the Riemannian distance $\hat{d}$ on $\hat{\Sigma}$. Recall that the metric $g_{\hat{f}} = \hat{f}^*\langle\cdot,\cdot\rangle$ on $\hat{\Sigma}$ induced by the immersion $\hat{f}$ makes $\hat{f}$ an isometry. Now, for any curve $\gamma\colon [0,1] \to \hat{\Sigma}$ such that $\eta(0) = p_n, \gamma(1) = p_m$ we have

$$|\hat{f}(p_n) - \hat{f}(p_m)| \leq \mathcal{L}(\hat{f} \circ \gamma) = \mathcal{L}(\gamma),$$

and hence we find $|\hat{f}(p_n) - \hat{f}(p_m)| \leq \hat{d}(p_n, p_m)$ for all $n, m \in \mathbb{N}$. In particular there exists $R > 0$ such that $(\hat{f}(p_n))_{n\in\mathbb{N}} \subset \overline{B_R(0)}$. As $\hat{f}$ is proper we find $p_n \in \hat{f}^{-1}(\overline{B_R(0)})$ which is compact. Since $(p_n)_{n\in\mathbb{N}}$ is Cauchy, $\lim_{n\to\infty} p_n \in \hat{\Sigma}$ exists. □

Appendix B Proof of Lemma 3.1

This section is devoted to proving Lemma 3.1. First, we compute a localized version of (2.14). Although the calculations are essentially the same as in [21, Section 3], we give some details here how the dependence on λ comes into play.

Lemma B.1. *Let $f\colon [0,T)\times\Sigma\to\mathbb{R}^3$ be a smooth volume-preserving Willmore flow, $\tilde\eta\in C_c^\infty(\mathbb{R}^3)$ and $\eta:=\tilde\eta\circ f$. Then, we have*

$$\partial_t\int\frac{1}{2}H^2\eta\,\mathrm{d}\mu+\int|\nabla\overline{\mathcal{W}}(f)|^2\eta\,\mathrm{d}\mu=\lambda\int|A^0|^2H\eta\,\mathrm{d}\mu-2\int\nabla_{sc}\overline{\mathcal{W}}(f)\langle\nabla H,\nabla\eta\rangle\,\mathrm{d}\mu$$
$$-\int\nabla_{sc}\overline{\mathcal{W}}(f)H\Delta\eta\,\mathrm{d}\mu+\int\frac{1}{2}H^2\partial_t\eta\,\mathrm{d}\mu \tag{B.1}$$

and

$$\partial_t\int|A^0|^2\eta\,\mathrm{d}\mu+\int|\nabla\overline{\mathcal{W}}(f)|^2\eta\,\mathrm{d}\mu=\lambda\int|A^0|^2H\eta\,\mathrm{d}\mu-2\int\nabla_{sc}\overline{\mathcal{W}}(f)\langle\nabla H,\nabla\eta\rangle\,\mathrm{d}\mu$$
$$-2\int\nabla_{sc}\overline{\mathcal{W}}(f)\langle A^0,\nabla^2\eta\rangle\,\mathrm{d}\mu+\int|A^0|^2\partial_t\eta\,\mathrm{d}\mu. \tag{B.2}$$

Proof. We use a (local) orthonormal basis $\{e_i\}_{i=1,2}$. As in [20, (31) and (32)], using (2.6) and (2.7) we find

$$\partial_t\left(\frac{1}{2}H^2\,\mathrm{d}\mu\right)=-|\nabla\overline{\mathcal{W}}(f)|^2\,\mathrm{d}\mu+\lambda\Delta H\,\mathrm{d}\mu+\lambda|A^0|^2H\,\mathrm{d}\mu+\nabla_i\left(H\nabla_i\xi-\xi\nabla_iH\right)\mathrm{d}\mu$$

Consequently, we compute using integration by parts

$$\begin{aligned}&\partial_t\int\frac{1}{2}H^2\eta\,\mathrm{d}\mu+\int|\nabla\overline{\mathcal{W}}(f)|^2\eta\,\mathrm{d}\mu\\&\quad=\lambda\int(\Delta H+|A^0|^2H)\eta\,\mathrm{d}\mu+\int(2\xi\nabla_iH\nabla_i\eta+H\xi\Delta\eta)\,\mathrm{d}\mu+\int\frac{1}{2}H^2\partial_t\eta\,\mathrm{d}\mu.\end{aligned}$$

Now, using (2.12) we observe that

$$\int(2\xi\nabla_iH\nabla_i\eta+H\xi\Delta\eta)\,\mathrm{d}\mu=-2\int\nabla_{sc}\overline{\mathcal{W}}(f)\nabla_iH\nabla_i\eta\,\mathrm{d}\mu+2\lambda\int\nabla_iH\nabla_i\eta\,\mathrm{d}\mu$$
$$-\int\nabla_{sc}\overline{\mathcal{W}}(f)H\Delta\eta\,\mathrm{d}\mu+\lambda\int H\Delta\eta\,\mathrm{d}\mu.$$

Recalling that $\Delta(H\eta)=\Delta H\eta+2\nabla_iH\nabla_i\eta+H\Delta\eta$, the identity (B.1) follows.
For the second identity, we proceed as in [20, p. 423]. Using (2.5) and the identity $A^0_{ik}A^0_{kj}A^0_{ij}=0$ (see [21, (2.5)]), a short computation yields

$$A^0(\partial_te_i,e_j)A^0(e_i,e_j)=\frac{1}{2}|A^0|^2H\xi. \tag{B.3}$$

Applying (2.6), (2.9) and (B.3) yields

$$\partial_t\left(|A^0|^2\,\mathrm{d}\mu\right)=2\nabla_i(\nabla_j\xi A^0(e_i,e_j))\,\mathrm{d}\mu-\nabla_j\xi\nabla_jH\,\mathrm{d}\mu+|A^0|^2H\xi\,\mathrm{d}\mu,$$

where we used (2.2) and the fact that $A^0_{ij}(\nabla^2_{ij}\xi)^0=A^0_{ij}\nabla^2_{ij}\xi$ as A^0 is trace-free. Consequently we find

$$\partial_t\left(|A^0|^2\,\mathrm{d}\mu\right)=2\nabla_i(\nabla_j\xi A^0(e_i,e_j))\,\mathrm{d}\mu-\nabla_j(\xi\nabla_jH)\,\mathrm{d}\mu-|\nabla\overline{\mathcal{W}}(f)|^2\,\mathrm{d}\mu+\lambda\nabla_{sc}\overline{\mathcal{W}}(f)\,\mathrm{d}\mu.$$

Integration by parts and (2.2) then yield

$$\begin{aligned}
&\partial_t \int |A^0|^2 \eta \,\mathrm{d}\mu + \int |\nabla \overline{\mathcal{W}}(f)|^2 \eta \,\mathrm{d}\mu \\
&\quad = -2 \int \nabla_{sc} \overline{\mathcal{W}}(f) \nabla_i H \nabla_i \eta \,\mathrm{d}\mu - 2 \int \nabla_{sc} \overline{\mathcal{W}}(f) A^0_{ij} \nabla^2_{ji} \eta \,\mathrm{d}\mu + \int |A^0|^2 \partial_t \eta \,\mathrm{d}\mu \\
&\qquad + \lambda \left[- \int \nabla_i H \nabla_i \eta \,\mathrm{d}\mu + 2 \int \nabla_i H \nabla_i \eta \,\mathrm{d}\mu + \int \Delta H \eta \,\mathrm{d}\mu + \int |A^0|^2 H \eta \,\mathrm{d}\mu \right]
\end{aligned}$$

The claim follows from integrating by parts in the terms involving λ. ☐

Equipped with this evolution identity, we can now prove Lemma 3.1.

Proof of Lemma 3.1. Again, we use a local orthonormal basis $\{e_i\}_{i=1,2}$. To prove both inequalities in Lemma 3.1, we estimate the evolution in Lemma B.1 with $\eta = \gamma^4$. The last term in (B.1) and (B.2) generates an additional term with λ, since

$$|\partial_t \eta| \leq C\Lambda\gamma^3 |\partial_t f| \leq C\Lambda\gamma^3 \left(|\nabla \overline{\mathcal{W}}(f)| + |\lambda|\right). \tag{B.4}$$

Therefore, both (B.1) and (B.2) contain two terms involving λ. The terms without λ can be estimated exactly as in [20, Lemma 3.2] (with $\rho^{-1} = \Lambda$). The claim follows after we estimate the λ-term generated by $\partial_t \eta$ as in (B.4) and keep the term $\lambda \int |A^0|^2 H \gamma^4 \,\mathrm{d}\mu$. ☐

Appendix C Proof of Proposition 3.5

This section is devoted to proving Proposition 3.5.
Following [21, 20], for tensors ϕ, ψ on Σ, we denote by $\phi * \psi$ any multilinear form, depending on ϕ and ψ in a universal bilinear way. In particular, we have $|\phi * \psi| \leq c|\phi||\psi|$ and $\nabla(\phi * \psi) = \nabla\phi * \psi + \phi * \nabla\psi$. Note that since we are in codimension one, we can work with tensors with scalar values and not with normal values.
Moreover, for $m \in \mathbb{N}_0$ and $r \in \mathbb{N}, r \geq 2$ we denote by $P^m_r(A)$ any term of the type

$$P^m_r(A) = \sum_{i_1 + \cdots + i_r = m} \nabla^{i_1} A * \cdots * \nabla^{i_r} A.$$

In addition, for $r = 1$ we extend this definition by denoting by $P^m_1(A)$ any contraction of $\nabla^m A$ with respect to the metric g. We can now compute the evolution of higher order derivatives of the second fundamental form.

Lemma C.1. *Let $f\colon [0,T) \times \Sigma \to \mathbb{R}^3$ be a volume-preserving Willmore flow. Then for all $m \in \mathbb{N}_0$ we have*

$$\partial_t(\nabla^m A) + \Delta^2(\nabla^m A) = P^{m+2}_3(A) + P^m_5(A) + \lambda P^m_2(A).$$

Proof. First, we note that H is a contraction of A and hence $H = P_1^0(A)$, and consequently also $A^0 = P_1^0(A)$. Thus, by (1.4), we have

$$\xi = -\Delta H + P_3^0(A) + \lambda. \tag{C.1}$$

For $m = 0$ we insert this into (2.8) to obtain

$$\partial_t A = \nabla^2 \xi + A * A * \xi = -\nabla^2(\Delta H) + P_3^2(A) + P_5^0(A) + \lambda P_2^0(A),$$

Using [21, (2.11)] twice, we find $\nabla^2 \Delta H = \Delta \nabla^2 H + P_3^2(A)$, hence by Simons' identity [38] we have

$$\partial_t A = -\Delta^2 A + P_3^2(A) + P_5^0(A) + \lambda P_2^0(A).$$

Assume the statement is true for $m \geq 1$. Using [21, Lemma 2.3] with $\phi = \nabla^m A$ and the fact that we are in codimension one yields

$$\begin{aligned}
\partial_t \nabla^{m+1} A + \Delta^2 \nabla^{m+1} A &= \nabla \left(P_3^{m+2}(A) + P_5^m(A) + \lambda P_2^m(A) \right) \\
&\quad + \sum_{i+j+k=3} \nabla^i A * \nabla^j A * \nabla^{k+m} A \\
&\quad + A * \nabla \xi * \nabla^m A + \nabla A * \xi * \nabla^m A \\
&= P_3^{m+3}(A) + P_5^{m+1}(A) + \lambda P_2^{m+1}(A),
\end{aligned}$$

where we used (C.1) in the last step. □

In analogy to [20, Proposition 3.3], we have localized energy estimates for higher order derivatives of A.

Lemma C.2. *Let $f\colon [0,T) \times \Sigma \to \mathbb{R}^3$ be a volume-preserving Willmore flow and γ as in (3.1). Then for $\phi = \nabla^m A, m \in \mathbb{N}_0$ and $s \geq 2m+4$ we have*

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t} \int |\phi|^2 \gamma^s \,\mathrm{d}\mu + \frac{1}{2} \int |\nabla^2 \phi| \gamma^s \,\mathrm{d}\mu \\
&\leq C \left(|\lambda|^{\frac{4}{3}} + \|A\|^4_{L^\infty([\gamma>0])} \right) \int |\phi|^2 \gamma^s \,\mathrm{d}\mu + C \left(1 + |\lambda|^{\frac{4}{3}} + \|A\|^4_{L^\infty([\gamma>0])} \right) \int_{[\gamma>0]} |A|^2 \,\mathrm{d}\mu
\end{aligned}$$

where $C = C(s, m, \Lambda) > 0$.

Proof. In the following, note that the value of $C = C(s, m, \Lambda)$ is allowed to change from line to line. Using [21, Lemma 3.2], we find

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t} \int |\phi|^2 \gamma^s \,\mathrm{d}\mu + \int |\nabla^2 \phi| \gamma^s \,\mathrm{d}\mu \\
&\leq 2 \int \langle Y, \phi \rangle \gamma^s \,\mathrm{d}\mu + \int A * \phi * \phi * \xi \gamma^s \,\mathrm{d}\mu + \int |\phi|^2 s \gamma^{s-1} \partial_t \gamma \,\mathrm{d}\mu \\
&\quad + C \int |\phi|^2 \gamma^{s-4} \left(|\nabla \gamma|^4 + \gamma^2 |\nabla^2 \gamma|^2 \right) \mathrm{d}\mu + C \int |\phi|^2 \left(|\nabla A|^2 + |A|^4 \right) \gamma^s \,\mathrm{d}\mu,
\end{aligned} \tag{C.2}$$

where $\partial_t \phi + \Delta^2 \phi = Y$ and $\xi = P_1^2(A) + P_3^0(A) + \lambda$ by (C.1). By Lemma C.1, we have

$$2\int \langle Y, \phi\rangle \gamma^s \,\mathrm{d}\mu + \int A * \phi * \phi * \xi \gamma^s \,\mathrm{d}\mu + C\int |\phi|^2 (|\nabla A|^2 + |A|^4)\gamma^s \,\mathrm{d}\mu$$
$$= \int \left(P_3^{m+2}(A) + P_5^m(A)\right) * \phi\gamma^s \,\mathrm{d}\mu + \lambda \int P_2^m(A) * \phi\gamma^s \,\mathrm{d}\mu. \tag{C.3}$$

Moreover, by (3.1) we find

$$\int |\phi|^2 \gamma^{s-1} \partial_t \gamma \,\mathrm{d}\mu = \int |\phi|^2 \gamma^{s-1} \langle D\tilde{\gamma} \circ f, \nu\rangle \left(-\Delta H - |A^0|^2 H + \lambda\right) \mathrm{d}\mu. \tag{C.4}$$

We proceed by estimating all the terms involving λ in (C.3) and (C.4). For the λ-term on the right hand side of (C.3), using [21, Corollary 5.5] with $k = m$, $r = 3$ we find

$$\lambda \int P_2^m(A) * \phi\gamma^s \,\mathrm{d}\mu \leq C(s, m, \Lambda)|\lambda| \|A\|_{L^\infty([\gamma>0])} \left(\int |\phi|^2 \gamma^s \,\mathrm{d}\mu + \int_{[\gamma>0]} |A|^2 \,\mathrm{d}\mu\right).$$

The λ-term on the right hand side of (C.4) is estimated using Young's inequality with $p = \frac{4}{3}$ and $q = 4$ to obtain

$$C|\lambda| \int |\phi|^2 \gamma^{s-1} \,\mathrm{d}\mu \leq C|\lambda|^{\frac{4}{3}} \int |\phi|^2 \gamma^s \,\mathrm{d}\mu + C\int |\phi|^2 \gamma^{s-4} \,\mathrm{d}\mu.$$

Consequently, we find from (C.2), (C.3), (C.4) and Young's inequality

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t} \int |\phi|^2 \gamma^s \,\mathrm{d}\mu + \int |\nabla^2 \phi| \gamma^s \,\mathrm{d}\mu \\
&\quad\leq \int \left(P_3^{m+2}(A) + P_5^m(A)\right) * \phi\gamma^s \,\mathrm{d}\mu + \int |\phi|^2 \gamma^{s-1} \langle D\tilde{\gamma} \circ f, \nu\rangle \left(-\Delta H - |A^0|^2 H\right) \mathrm{d}\mu \\
&\qquad + C\left(|\lambda|^{\frac{4}{3}} + \|A\|^4_{L^\infty([\gamma>0])}\right) \int |\phi|^2 \gamma^s \,\mathrm{d}\mu + C\left(|\lambda|^{\frac{4}{3}} + \|A\|^4_{L^\infty([\gamma>0])}\right) \int_{[\gamma>0]} |A|^2 \,\mathrm{d}\mu \\
&\qquad + C\int |\phi|^2 \gamma^{s-4} \,\mathrm{d}\mu + \int |\phi|^2 \gamma^{s-4} \left(|\nabla\gamma|^4 + \gamma^2 |\nabla^2 \gamma|^2\right) \mathrm{d}\mu.
\end{aligned} \tag{C.5}$$

Now, all the terms involving λ on the right hand side of (C.5) are as in the statement. For the second and the last term in (C.5), one may proceed exactly as in the proof of [21, Proposition 3.3]. This way, one creates additional terms which can be estimated by

$$\int |\phi|^2 \gamma^{s-4} \,\mathrm{d}\mu + \int |\nabla\phi|^2 \gamma^{s-2} \,\mathrm{d}\mu \leq \varepsilon \int |\nabla^2 \phi|^2 \gamma^s \,\mathrm{d}\mu + C_\varepsilon \int_{[\gamma>0]} |A|^2 \gamma^{s-4-2m} \,\mathrm{d}\mu,$$

for every $\varepsilon > 0$, using twice the interpolation inequality [21, Corollary 5.3] (which trivially also holds in the case $k = m = 0$). The first term on the right hand side of (C.5) can then be estimated by means of [21, (4.15)]. After choosing $\varepsilon > 0$ small enough and absorbing, the claim follows. $\square$

Proposition 3.5 can now be deduced from a Gronwall-type argument exactly as in [20, Theorem 3.5]. To keep track of the role of λ, we give the details here.

Proof of Proposition 3.5. Without loss of generality, after rescaling as in Remark 2.2, we may assume $\rho = 1$.
We pick a cutoff function $\tilde{\gamma} \in C_c^\infty(\mathbb{R}^3)$ with $\chi_{B_{3/4}(x)} \leq \tilde{\gamma} \leq \chi_{B_1(x)}$ such that $\gamma := \tilde{\gamma} \circ f$ is as in (3.1) with a universal constant $\Lambda > 0$. Now, using Corollary 3.4, we deduce

$$\int_0^T \int_{B_{3/4}(x)} \left(|\nabla^2 A|^2 + |A|^6\right) \mathrm{d}\mu \, \mathrm{d}t \leq C\varepsilon + C\Lambda^4 \varepsilon T + C\varepsilon L = C(T^*, L)\varepsilon, \tag{C.6}$$

using $T \leq T^*$. Consequently, by using we find

$$\int_0^T \|A\|^4_{L^\infty(B_{3/4}(x))} \, \mathrm{d}t \leq C(T^*, L)\varepsilon. \tag{C.7}$$

Now, we change to another test function $\tilde{\gamma} \in C_c^\infty(\mathbb{R}^3)$ with $\chi_{B_{1/2}(x)} \leq \tilde{\gamma} \leq \chi_{B_{3/4}(x)}$ and $\gamma := \tilde{\gamma} \circ f$. Note that (3.1) still remains satisfied with a universal $\Lambda > 0$. We now define Lipschitz cutoff functions in time via

$$\xi_j(t) := \begin{cases} 0, & \text{for } t \leq (j-1)\frac{T}{m}, \\ \frac{m}{T}\left(t - (j-1)\frac{T}{m}\right), & \text{for } (j-1)\frac{T}{m} \leq t \leq j\frac{T}{m} \\ 1, & \text{for } t \geq j\frac{T}{m}, \end{cases}$$

where $m \in \mathbb{N}$ and $0 \leq j \leq m$. We also define $\xi_{-1}(t) := 0$ and $\xi_0(t) := 1$ for all $t \in \mathbb{R}$ if $m = 0$. We note that $\xi_m(T) = 1$ and

$$0 \leq \frac{\mathrm{d}}{\mathrm{d}t}\xi_j \leq \frac{m}{T}\xi_{j-1}, \quad \text{for all } j \in \mathbb{N}_0. \tag{C.8}$$

We now define $a(t) = \|A\|^4_{L^\infty(B_{3/4}(x))}$, $E_j(t) = \int |\nabla^{2j} A|^2 \gamma^{4j+4} \, \mathrm{d}\mu$. Then, by Lemma C.2 and using $\gamma \leq 1$ we have

$$\frac{\mathrm{d}}{\mathrm{d}t}E_j(t) + \frac{1}{2}E_{j+1}(t) \leq C(j,m)\left(|\lambda(t)|^{\frac{4}{3}} + a(t)\right)E_j(t) + C(j,m)\left(1 + |\lambda(t)|^{\frac{4}{3}} + a(t)\right)\varepsilon.$$

Therefore, if we define $e_j := \xi_j E_j$ this implies using (C.8)

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}e_j(t) \leq \; & \frac{m}{T}\xi_{j-1}(t)E_j(t) + C(j,m)\left(|\lambda(t)|^{\frac{4}{3}} + a(t)\right)e_j(t) \\ & + C(j,m)\left(1 + |\lambda(t)|^{\frac{4}{3}} + a(t)\right)\varepsilon - \frac{1}{2}\xi_j(t)E_{j+1}(t). \end{aligned} \tag{C.9}$$

We will now show that this implies for $0 \leq j \leq m$ and $t \in (0, T)$

$$e_j(t) + \frac{1}{2}\int_0^t \xi_j(s)E_{j+1}(s) \, \mathrm{d}s \leq \frac{C(j, m, T^*, L)\varepsilon}{T^j}. \tag{C.10}$$

We proceed by induction on j. For $j=0$ we have $\xi_0 \equiv 1$ on $(0,T)$. Therefore, we have $e_0 = \int |A|^2\gamma^4 \,\mathrm{d}\mu \le \varepsilon$ by assumption. Moreover, by (C.6) we find $\int_0^t E_1(s)\,\mathrm{d}s = \int_0^t \int |\nabla^2 A|^2 \gamma^8 \,\mathrm{d}\mu\,\mathrm{d}s \le C(T^*, L)\varepsilon$.
For $j \ge 1$ we have, integrating (C.9) on $[0,t]$ and using $e_j(0) = 0$

$$
\begin{aligned}
&e_j(t) + \frac{1}{2}\int_0^t \xi_j(s) E_{j+1}(s)\,\mathrm{d}s \\
&\quad \le C(j,m)\int_0^t \left(|\lambda(s)|^{\frac{4}{3}} + a(s)\right) e_j(s)\,\mathrm{d}s + C(j,m)\varepsilon \int_0^t \left(1 + |\lambda(s)|^{\frac{4}{3}} + a(s)\right)\mathrm{d}s \\
&\qquad + \frac{m}{T}\int_0^t \xi_{j-1}(s) E_j(s)\,\mathrm{d}s \\
&\le C(j,m)\int_0^t \left(|\lambda(s)|^{\frac{4}{3}} + a(s)\right) e_j(s)\,\mathrm{d}s + C(j,m,T^*,L)\varepsilon + \frac{C(j,m,T^*,L)\varepsilon}{T^{j-1}}\,\frac{m}{T} \\
&\le C(j,m)\int_0^t \left(|\lambda(s)|^{\frac{4}{3}} + a(s)\right) e_j(s)\,\mathrm{d}s + \frac{C(j,m,T^*,L)\varepsilon}{T^j},
\end{aligned}
$$

using (C.7), the induction hypothesis and $T \le T^*$. Therefore, Gronwall's inequality yields using (C.6) and (C.7)

$$
\begin{aligned}
e_j(t) &\le -\frac{1}{2}\int_0^t \xi_j(s) E_{j+1}(s)\,\mathrm{d}s + \frac{C(j,m,T^*,L)\varepsilon}{T^j} \\
&\quad + \int_0^t \frac{C(j,m,T^*,L)\varepsilon}{T^j}\left(|\lambda(s)|^{\frac{4}{3}} + a(s)\right)\exp\left(C(T^*,L)\right)\mathrm{d}s \\
&\le -\frac{1}{2}\int_0^t \xi_j(s) E_{j+1}(s)\,\mathrm{d}s + \frac{C(m,L,T^*)\varepsilon}{T^j},
\end{aligned}
$$

which proves (C.10). Now evaluating at $t = T$ with $j = m$, we find

$$
\int |\nabla^{2m} A|^2 \gamma^{4m+4}\,\mathrm{d}\mu \le \frac{C(m,L,T^*)\varepsilon}{T^m} \quad \text{for all } m \in \mathbb{N}.
$$

The estimate for $\nabla^{2m+1}A$ follows from the interpolation inequality in [21, Lemma 5.1] with $r=1, p=q=2, \alpha=1, \beta=0, s=4m+6$ and $t=\frac{1}{2}\in[-\frac{1}{2},\frac{1}{2}]$. Renaming T into t proves the L^2-estimate. The L^∞-estimate then follows using the L^∞-interpolation estimate in [21, Lemma 2.8], together with [21, Lemma 4.2]. ☐

Appendix D Proof of Lemma 7.9

Proof of Lemma 7.9. We follow [10, Lemma 4.1]. There exists a diffeomorphism $\Phi\colon \Sigma \to \Sigma$, such that for $\varepsilon > 0$ small enough, $f_0 \circ \Phi$ can be written as a normal graph over f_W, i.e.

$$
f_0 \circ \Phi = f_W + \nu_{f_W}\varphi_0 =: \tilde{f}_0,
$$

for some $\varphi_0\colon \Sigma \to \mathbb{R}$, such that

$$\|\varphi_0\|_{C^{4,\alpha}} \leq C\varepsilon, \tag{D.1}$$

for C independent of ε. We now wish to solve the equation

$$\partial_t^\perp \tilde{f}_t = -\nabla\overline{\mathcal{W}}(\tilde{f}_t) + \lambda(\tilde{f}_t)\nu_{\tilde{f}_t}, \tag{D.2}$$

with initial datum $\tilde{f}_0$, where $\partial_t^\perp = P^{\perp_{\tilde{f}_t}}\partial_t$ and $\tilde{f}_t := f_W + \varphi_t\nu_{f_W}$, for smooth functions $\varphi_t\colon \Sigma \to \mathbb{R}$. By (D.2) and (2.12) (as in [10, (4.4)] with codimension one) we compute

$$\begin{aligned}\partial_t(\varphi_t)P^{\perp_{\tilde{f}_t}}\nu_{f_W} &= -(\Delta H_{\tilde{f}_t} + |A^0_{\tilde{f}_t}|^2 H_{\tilde{f}_t})\nu_{\tilde{f}_t} + \lambda(\tilde{f}_t)\nu_{\tilde{f}_t}\\ &= -(g^{ij}_{\tilde{f}_t} g^{k\ell}_{\tilde{f}_t}\partial_{ijk\ell}\varphi_t)P^{\perp_{\tilde{f}_t}}\nu_{f_W}\\ &\quad + \left(1 + \int_\Sigma B_0(\cdot,\varphi_t, D\varphi_t, D^2\varphi_t)\,\mathrm{d}\mu_{f_W}\right) B_1(\cdot,\varphi_t,D\varphi_t,D^2\varphi_t,D^3\varphi_t),\end{aligned} \tag{D.3}$$

using (1.5), where B_0, B_1 are smooth functions depending on f_W. Note that the nonlocal terms appear due to λ. Now, if $\|\tilde{f}_t - f_W\|_{C^1} \leq \delta$ is small enough, $g^{ij}_{\tilde{f}_t} g^{k\ell}_{\tilde{f}_t}$ is uniformly elliptic and we may assume that

$$|P^{\perp_{\tilde{f}_t}}X| \geq |X| - |P^{\top_{f_W}}X - P^{\top_{\tilde{f}_t}}X| \geq \frac{1}{2}|X|, \text{ for all } X \text{ normal along } f_W. \tag{D.4}$$

Therefore, (D.3) is equivalent to

$$\begin{aligned}&\partial_t\varphi_t + g^{ij}_{\tilde{f}_t} g^{k\ell}_{\tilde{f}_t}\partial_{ijk\ell}\varphi_t\\ &\qquad = \left(1 + \int_\Sigma B_0(\cdot,\varphi_t,D\varphi_t,D^2\varphi_t)\,\mathrm{d}\mu_{f_W}\right) B_1(\cdot,\varphi_t,D\varphi_t,D^2\varphi_t,D^3\varphi_t).\end{aligned} \tag{D.5}$$

Since the right hand side of (D.5) is only of third order in φ_t, it is not too difficult to see that the parabolic initial value problem (D.5) with initial datum φ_0 satisfying (D.1) has a unique local solution in the Hölder space $H^{\frac{k+\alpha}{4},k+\alpha}([0,T_1]\times\Sigma;\mathbb{R})$ for some $0 < T_1 \leq T$. This follows from maximal regularity results for linear parabolic problems, in Hölder spaces, cf. [24], and a fixed-point argument using the contraction principle, see also Proposition 2.1 and the corresponding references. Here the order reduction for λ discussed in Section 2.2 is crucial. Now, we apply Theorem 7.8 to f_W. By the embedding $C^4(\Sigma) \hookrightarrow W^{4,2}(\Sigma)$, we may assume that the constrained Łojasiewicz–Simon inequality is satisfied for all $\|h - f_W\|_{C^4} \leq \sigma$ with exponent $\theta \in (0, \frac{1}{2}]$. Choosing $\varepsilon > 0$ sufficiently small, we may without loss of generality assume $C\varepsilon < \sigma < \delta$ with σ as in Theorem 7.8 and that T_1 is the maximal existence interval for (D.5) for which we have (as part of our definition of T_1)

$$\|\tilde{f}_t - f_W\|_{C^k} \leq \sigma < \delta \text{ for all } t \in [0, T_1). \tag{D.6}$$

By parabolic Schauder estimates, from (D.5) and (D.1) we obtain a bound on the parabolic Hölder space norm, i.e. $\|\varphi\|_{H^{\frac{k+\alpha}{4},k+\alpha}} \leq C$, and hence for k as in the statement

$$\|\tilde{f}_t - f_W\|_{C^{k,\alpha}} \leq C \text{ for all } t \in [0, T_1). \tag{D.7}$$

By (D.2), we find

$$\partial_t \tilde{f}_t + \xi_t d\tilde{f}_t = -\nabla\overline{\mathcal{W}}(\tilde{f}_t) + \lambda(\tilde{f}_t)\nu_{\tilde{f}_t},$$

where ξ_t denotes the tangential velocity. Next, by classical flow theory, see for instance [26, Chapter 17], there exists a unique smooth family of diffeomorphisms satisfying

$$\begin{aligned}\partial_t \Phi_t &= \xi_t \circ \Phi_t \text{ on } \Sigma \text{ for } 0 \leq t < T_1\\ \Phi_0 &= \mathrm{Id}_\Sigma\,.\end{aligned}$$

A direct calculation yields $\partial_t(\tilde{f}_t \circ \Phi_t) = -\nabla\overline{\mathcal{W}}(\tilde{f}_t \circ \Phi_t) + \lambda(\tilde{f}_t \circ \Phi_t)\nu_{\tilde{f}_t\circ\Phi_t}$, so

$$[0, T_1) \times \Sigma \to \mathbb{R}^3, (t,p) \mapsto \tilde{f}_t \circ \Phi_t \circ \Phi^{-1}(p)$$

is a smooth volume-preserving Willmore flow with initial data $\tilde{f}_0 \circ \Phi_0 \circ \Phi^{-1} = f_0$. As the solution to the volume-preserving Willmore flow is unique, cf. Proposition 2.1, we conclude $T_1 \leq T$ and

$$f_t = \tilde{f}_t \circ \Phi_t \circ \Phi^{-1} \text{ for all } 0 \leq t < T_1.$$

It suffices to prove that $\tilde{f}$ is global and converges as $t \to \infty$ to a smooth Willmore immersion f_∞ with the desired properties.
First, we show that we may assume $\overline{\mathcal{W}}(f_t) > \overline{\mathcal{W}}(f_W)$ for all $t \in [0, T_1)$. By assumption and (D.6), we have $\overline{\mathcal{W}}(f_t) \geq \overline{\mathcal{W}}(f_W)$. If $\overline{\mathcal{W}}(f_t) = \overline{\mathcal{W}}(f_W)$ for some $t \in [0, T_1)$, then by Remark 2.5, f and $\tilde{f}$ are stationary and the claim follows. Hence, we may indeed assume the strict inequality $\overline{\mathcal{W}}(f_t) > \overline{\mathcal{W}}(f_W)$.
Let θ, C as in Theorem 7.8. By (2.13), (D.6) and since $\mathcal{V}(f_t) = \mathcal{V}(f_W)$, we may apply the constrained Łojasiewicz–Simon gradient inequality to obtain

$$\begin{aligned}
&-\frac{\mathrm{d}}{\mathrm{d}t}\left(\overline{\mathcal{W}}(f_t) - \overline{\mathcal{W}}(f_W)\right)^\theta\\
&= -\theta\left(\overline{\mathcal{W}}(f_t) - \overline{\mathcal{W}}(f_W)\right)^{\theta-1}\langle\nabla\overline{\mathcal{W}}(\tilde{f}_t), \partial_t^\perp \tilde{f}_t\rangle_{L^2(\mathrm{d}\mu_{\tilde{f}_t})}\\
&= -\theta\left(\overline{\mathcal{W}}(f_t) - \overline{\mathcal{W}}(f_W)\right)^{\theta-1}\langle\nabla\overline{\mathcal{W}}(\tilde{f}_t) - \lambda(\tilde{f}_t)\nu_{\tilde{f}_t}, \partial_t^\perp \tilde{f}_t\rangle_{L^2(\mathrm{d}\mu_{\tilde{f}_t})}\\
&= \theta\left(\overline{\mathcal{W}}(f_t) - \overline{\mathcal{W}}(f_W)\right)^{\theta-1}\|\nabla\overline{\mathcal{W}}(\tilde{f}_t) - \lambda(\tilde{f}_t)\nu_{\tilde{f}_t}\|_{L^2(\mathrm{d}\mu_{\tilde{f}_t})}\|\partial_t^\perp \tilde{f}_t\|_{L^2(\mathrm{d}\mu_{\tilde{f}_t})}\\
&\geq \frac{\theta}{C}\|\partial_t^\perp \tilde{f}_t\|_{L^2(\mathrm{d}\mu_{\tilde{f}_t})}
\end{aligned}$$

for $0 \le t < T_1$. Now, using (D.4), (D.6) and the resulting equivalence of the metrics g_{f_W} and $g_{\tilde{f}_t}$, we find

$$\|\partial_t \tilde{f}_t\|_{L^2(\mathrm{d}\mu_{f_W})} \le -\frac{C}{\theta}\frac{\mathrm{d}}{\mathrm{d}t}\left(\overline{\mathcal{W}}(f_t) - \overline{\mathcal{W}}(f_W)\right)^\theta \text{ for every } t \in [0, T_1). \tag{D.8}$$

Integrating in time and using the triangle inequality we find

$$\begin{aligned} \|\tilde{f}_t - f_W\|_{L^2(\mathrm{d}\mu_{f_W})} &\le \|\tilde{f}_0 - f_W\|_{L^2(\mathrm{d}\mu_{f_W})} + C\left(\overline{\mathcal{W}}(\tilde{f}_0) - \overline{\mathcal{W}}(f_W)\right)^\theta \\ &\le C\|\tilde{f}_0 - f_W\|^\theta_{C^2(\Sigma)}, \end{aligned}$$

using the mean value theorem for the Willmore energy and assuming that $\varepsilon > 0$ is small enough. As in [10, p. 361], by interpolation for some $\beta \in (0,1)$ we find for $t \in [0, T_1)$ and k as in the statement, using (D.7) and (D.1)

$$\begin{aligned} \|\tilde{f}_t - f_W\|_{C^k(\Sigma)} &\le C\|\tilde{f}_t - f_W\|^{1-\beta}_{C^{k,\alpha}(\Sigma)}\|\tilde{f}_t - f_W\|^\beta_{L^2(\Sigma, \mathrm{d}\mu_{f_W})} \\ &\le C\|\tilde{f}_0 - f_W\|^{\beta\theta}_{C^2(\Sigma)} \le C\varepsilon^{\beta\theta} \le \frac{\sigma}{2}, \end{aligned} \tag{D.9}$$

if $\varepsilon > 0$ is sufficiently small. Since $T_1 > 0$ is chosen maximal with respect to (D.6), this implies $T_1 = \infty$, which yields that $\tilde{f}$ exist globally and satisfies $\|\tilde{f}_t - f_W\|_{C^k(\Sigma)} \le \sigma$ for all $t \ge 0$. Therefore, (D.8) yields $\partial_t \tilde{f}_t \in L^1([0,\infty); L^2(\Sigma, \mathrm{d}\mu_{f_W}))$, and consequently, there exists $f_\infty := \lim_{t\to\infty} \tilde{f}_t$ in $L^2(\Sigma, \mathrm{d}\mu_{f_W})$. Similar to (D.9), an interpolation argument and (D.7) yield $\lim_{t\to\infty} \tilde{f}_t = f_\infty$ in $C^k(\Sigma)$. By parabolic Schauder estimates, one can then obtain L^∞-bounds on higher order derivatives, such that by interpolation again, one can show that the convergence $\lim_{t\to\infty} \tilde{f}_t = f_\infty$ is even smooth. Since the volume-preserving Willmore flow is a gradient flow, f_∞ is a constrained Willmore immersion. Using that $\|f_\infty - f_W\|_{C^k(\Sigma)} \le \sigma$, we find by Theorem 7.8 that

$$|\overline{\mathcal{W}}(f_\infty) - \overline{\mathcal{W}}(f_W)|^{1-\theta} \le C\|\nabla\overline{\mathcal{W}}(f_\infty) - \lambda(f_\infty)\nu_{f_\infty}\|_{L^2(\mathrm{d}\mu_{f_\infty})} = 0,$$

so $\overline{\mathcal{W}}(f_\infty) = \overline{\mathcal{W}}(f_W)$. □

Acknowledgments

This project has been supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation), project no. 404870139. The author would like to thank Anna Dall'Acqua and Marius Müller for many helpful discussions and comments. The author would also like to thank Mattia Fogagnolo for pointing out that the proof of Theorem 1.2 works for initial energy equal to 8π as well. In addition, the author is grateful to the referee for their careful reading and their valuable comments on the original manuscript.

References

[1] J. Appell and P. P. Zabrejko. *Nonlinear superposition operators*, volume 95 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 1990.

[2] T. Aubin. *Nonlinear analysis on manifolds. Monge–Ampère equations*, volume 252 of *Grundlehren der Mathematischen Wissenschaften*. Springer-Verlag, New York, 1982.

[3] Y. Bernard, G. Wheeler, and V.-M. Wheeler. Concentration-compactness and finite-time singularities for Chen's flow. *J. Math. Sci. Univ. Tokyo*, 26(1):55–139, 2019.

[4] S. Blatt. A singular example for the Willmore flow. *Analysis (Munich)*, 29(4):407–430, 2009.

[5] S. Blatt. A note on singularities in finite time for the L^2 gradient flow of the Helfrich functional. *J. Evol. Equ.*, 19(2):463–477, 2019.

[6] S. Blatt. A reverse isoperimetric inequality and its application to the gradient flow of the Helfrich functional, 2020. arXiv:2009.12273.

[7] P. Breuning. Immersions with bounded second fundamental form. *J. Geom. Anal.*, 25(2):1344–1386, 2015.

[8] R. L. Bryant. A duality theorem for Willmore surfaces. *Journal of Differential Geometry*, 20(1):23–53, 1984.

[9] P. Canham. The minimum energy of bending as a possible explanation of the biconcave shape of the human red blood cell. *J. Theor. Biol.*, 26(1):61–81, 1970.

[10] R. Chill, E. Fašangová, and R. Schätzle. Willmore blowups are never compact. *Duke Math. J.*, 147(2):345–376, 2009.

[11] A. Dall'Acqua, C.-C. Lin, and P. Pozzi. A gradient flow for open elastic curves with fixed length and clamped ends. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)*, 17(3):1031–1066, 2017.

[12] A. Dall'Acqua, M. Müller, R. Schätzle, and A. Spener. The Willmore flow of tori of revolution, 2020. arXiv:2005.13500.

[13] M. Droske and M. Rumpf. A level set formulation for Willmore flow. *Interfaces Free Bound.*, 6(3):361–378, 2004.

[14] G. Dziuk, E. Kuwert, and R. Schätzle. Evolution of elastic curves in $\mathbb{R}^n$: existence and computation. *SIAM J. Math. Anal.*, 33(5):1228–1245, 2002.

[15] S. W. Hawking. Gravitational radiation in an expanding universe. *J. Math. Phys.*, 9(4):598–604, 1968.

[16] W. Helfrich. Elastic properties of lipid bilayers: Theory and possible experiments. *Z. Naturforsch., C, J. Biosci.*, 28(11):693–703, 1973.

[17] G. Huisken. The volume preserving mean curvature flow. *J. Reine Angew. Math.*, 382:35–48, 1987.

[18] F. Jachan. *Area preserving Willmore flow in asymptotically Schwarzschild manifolds.* PhD thesis, FU Berlin, 2014.

[19] R. Jakob. The Willmore flow of Hopf-tori in the 3-sphere, 2020. arXiv:2002.01006.

[20] E. Kuwert and R. Schätzle. The Willmore flow with small initial energy. *J. Differential Geom.*, 57(3):409–441, 2001.

[21] E. Kuwert and R. Schätzle. Gradient flow for the Willmore functional. *Comm. Anal. Geom.*, 10(2):307–339, 2002.

[22] E. Kuwert and R. Schätzle. Removability of point singularities of Willmore surfaces. *Ann. of Math. (2)*, 160(1):315–357, 2004.

[23] E. Kuwert and R. Schätzle. The Willmore functional. In *Topics in modern regularity theory*, volume 13 of *CRM Series*, pages 1–115. Ed. Norm., Pisa, 2012.

[24] O. A. Ladyženskaja, V. A. Solonnikov, and N. N. Ural′ceva. *Linear and quasilinear equations of parabolic type.* Translated from the Russian by S. Smith. Translations of Mathematical Monographs, Vol. 23. American Mathematical Society, Providence, R.I., 1968.

[25] J. Langer. A compactness theorem for surfaces with L_p-bounded second fundamental form. *Math. Ann.*, 270(2):223–234, 1985.

[26] J. M. Lee. *Introduction to smooth manifolds*, volume 218 of *Graduate Texts in Mathematics*. Springer, New York, second edition, 2013.

[27] P. Li and S. T. Yau. A new conformal invariant and its applications to the Willmore conjecture and the first eigenvalue of compact surfaces. *Invent. Math.*, 69(2):269–291, 1982.

[28] U. F. Mayer and G. Simonett. A numerical scheme for axisymmetric solutions of curvature-driven free boundary problems, with applications to the Willmore flow. *Interfaces Free Bound.*, 4(1):89–109, 2002.

[29] J. McCoy, S. Parkins, and G. Wheeler. The geometric triharmonic heat flow of immersed surfaces near spheres. *Nonlinear Anal.*, 161:44–86, 2017.

[30] J. McCoy and G. Wheeler. A classification theorem for Helfrich surfaces. *Math. Ann.*, 357(4):1485–1508, 2013.

[31] J. McCoy and G. Wheeler. Finite time singularities for the locally constrained Willmore flow of surfaces. *Comm. Anal. Geom.*, 24(4):843–886, 2016.

[32] J. McCoy, G. Wheeler, and G. Williams. Lifespan theorem for constrained surface diffusion flows. *Math. Z.*, 269(1-2):147–178, 2011.

[33] J. A. McCoy. Mixed volume preserving curvature flows. *Calc. Var. Partial Differential Equations*, 24(2):131–154, 2005.

[34] F. Rupp. On the Łojasiewicz–Simon gradient inequality on submanifolds. *J. Funct. Anal.*, 279(8):108708, 2020.

[35] F. Rupp and A. Spener. Existence and convergence of the length-preserving elastic flow of clamped curves, 2020. arXiv:2009.06991.

[36] J. Schygulla. Willmore minimizers with prescribed isoperimetric ratio. *Arch. Ration. Mech. Anal.*, 203(3):901–941, 2012.

[37] L. Simon. Existence of surfaces minimizing the Willmore functional. *Comm. Anal. Geom.*, 1(2):281–326, 1993.

[38] J. Simons. Minimal varieties in Riemannian manifolds. *Ann. of Math. (2)*, 88:62–105, 1968.

[39] G. Wheeler. Lifespan theorem for simple constrained surface diffusion flows. *J. Math. Anal. Appl.*, 375(2):685–698, 2011.

[40] G. Wheeler. Surface diffusion flow near spheres. *Calc. Var. Partial Differential Equations*, 44(1-2):131–151, 2012.

[41] T. J. Willmore. Note on embedded surfaces. *An. Şti. Univ. "Al. I. Cuza" Iaşi Secţ. I a Mat. (N.S.)*, 11B:493–496, 1965.