A Phase Field based PDE Constrained Optimization Approach to Time Discrete Willmore Flow

Martina Franken, Martin Rumpf and Benedikt Wirth

A novel phase field model for Willmore flow is proposed based on a nested variational time discretization. Thereby, the mean curvature in the Willmore functional is replaced by an approximate speed of mean curvature motion, which is computed via a fully implicit variational model for time discrete mean curvature motion. The time discretization of Willmore flow is then performed in a nested fashion: in an outer variational approach a natural time discretization is setup for the actual Willmore flow, whereas for the involved mean curvature the above variational approximation is taken into account. Hence, in each time step a PDE-constrained optimization problem has to be solved in which the actual surface geometry as well as the geometry resulting from the implicit curvature motion time step are represented by phase field functions. The convergence behavior is experimentally validated and compared with rigorously proved convergence estimates for a simple linear model problem. Computational results in 2D and 3D underline the robustness of the new discretization, in particular for large time steps and in comparison with a semi-implicit convexity splitting scheme. Furthermore, the new model is applied as a minimization method for elastic functionals in image restoration.

Key words. phase field approach, Willmore flow, image restoration, PDE constraint optimization.

AMS subject classification. 35R35, 49J40, 60G40.

1 Introduction

In this paper a new phase field model for the time discretization of Willmore flow, also known as elastic flow, is proposed. Willmore flow is the L^2-gradient flow for the Willmore energy

$$w[x] = \frac{1}{2} \int_{\Gamma[x]} \mathbf{h}^2 \mathrm{d}\mathcal{H}^{d-1} \tag{1}$$

on hypersurfaces $\Gamma[x] \subset \mathbb{R}^d$ parametrized over itself by the identity mapping x, where $\mathbf{h}$ is the mean curvature of $\Gamma[x]$ and $\mathcal{H}^{d-1}$ represents the $(d-1)$-dimensional Hausdorff measure. Physically, this energy reflects an approximation of the stored energy in a thin elastic shell. Applications of the Willmore energy and Willmore flow range from the modeling of edge sets in imaging [40, 39, 57, 8] to applications in surface modeling [54, 5, 4, 47, 56]. An extension of the Willmore energy, the Helfrich model, is used to describe elastic cell membranes in biology [30, 51, 21].

Willmore surfaces, defined as minimizers of the Willmore energy [55], and Willmore flow have attracted a lot of attention over the last decade. Simonett proved in [52] the existence of a unique and locally smooth solution of Willmore flow for sufficiently smooth initial surfaces. Furthermore, he proved exponential convergence to a sphere for initial surfaces close to a sphere. The analytic treatment of Willmore flow of curves and surfaces was investigated by Polden [45, 46] already in 1996. Kuwert and Schätzle treated long time existence and regularity of solutions in [31, 32, 33]. Recently, Rivière [48] extended results of Kuwert and Schätzle [34] for co-dimension 1 to arbitrary co-dimension.

A theoretical and numerical treatment of Willmore flow of curves was presented by Dziuk, Kuwert and Schätzle in [24]. Concerning the numerical approximation of parametric Willmore flow of surfaces Rusu [50] proposed a mixed method for the surface parametrization x and the mean curvature vector $\mathbf{h}\, n$ (with n being the surface normal) as independent variables, see also [11] for the application to surface restoration. A level set formulation was given in [20] based on a different type of splitting, involving the level set function ϕ and a curvature density function $\mathbf{h}\,|\nabla\phi|$. An error analysis for spatially discretized, time-continuous Willmore flow for graphs was presented by Deckelnick and Dziuk in [15]. They used an analogous splitting in the context of piecewise linear finite elements and proved $L^\infty(L^2)$- as well as $L^2(L^2)$-error bounds of $O(h^2 \log h)$ for the discretized graph solution. Deckelnick and Schieweck demonstrated convergence of a conforming finite element approximation in case of axially symmetric surfaces [17]. An error analysis in the case of the elastic flow of curves was recently presented by Dziuk and Deckelnick in [18]. Barrett, Garcke and Nürnberg [2] and Dziuk [25] presented alternative finite element algorithms for parametric Willmore flow. The Willmore functional is invariant with respect to Möbius transformations. In [6] Bobenko and Schröder proposed a discrete Willmore flow scheme which takes into account a circle pattern on the surface, whose temporal evolution directly reflects these invariances.

In this paper we discuss Willmore flow in the context of a phase field model. In their pioneering paper [38] Modica and Mortola proved the Γ-convergence of $a^\varepsilon[u] = \frac{1}{2}\int_\Omega \varepsilon|\nabla u|^2 + \frac{1}{\varepsilon}\Psi(u)\,\mathrm{d}x$ to the area functional, where Ψ is a properly chosen double well function. This motivated the use of a corresponding phase field model for the mean curvature motion as the L^2-gradient flow for the area functional [37]. Nochetto, Paolini and Verdi treated in [42, 41] the error between the exact evolution of an interface under mean curvature flow and the evolution of a diffusive interface computed via a phase field mean curvature motion model. They proved an optimal error estimate of order $O(\varepsilon^2)$. More recently, Evans, Soner and Souganidis proved in [28] that a scaled Allen–Cahn equation leads to a generalized motion by mean curvature. De Giorgi conjectured that the functional $w^\varepsilon[u] = \frac{1}{2\varepsilon}\int_\Omega \left(-\varepsilon\Delta u + \frac{1}{2\varepsilon}\Psi'(u)\right)^2 \mathrm{d}x$, whose integrand is the squared first variation of $a^\varepsilon[u]$, Γ-converges to the Willmore functional [14]. This functional has been investigated analytically by Loreti and March in [35] and Bellettini and Mugnai in [3]. Du et al. proved in [23] by formal asymptotics that the Euler–Lagrange equation of the phase field formulation converges to the Euler–Lagrange equation of the Willmore energy (1). For a modified functional a corresponding Γ-convergence result could finally be established by Röger and Schätzle [49]. Dondl, Mugnai and Röger used a phase field model for minimizing Euler's elastica energy of non-overlapping curves in a bounded domain [19]. Concerning numerically discretized phase field models, Chen et al. proved in [9] that the zero level set of the solution of the Allen–Cahn equation converges to the mean curvature flow as ε goes to zero if h, $\sqrt{\tau} = O(\varepsilon^p)$ for $p > 1$ for the grid size h and the time step size τ. In [16] Deckelnick, Dziuk and Elliott

provided a review of different discretization methods for mean curvature motion and compared the parametric, graph, level set, and phase field approach. With respect to the simulation of fourth order PDEs based on phase field models, Elliott et al. [26] proved (among other estimates) for a splitting scheme for the Cahn–Hilliard evolution an $O(h^2)$ estimate for the $L^\infty(L^2)$-error. A corresponding time-discretized backward Euler scheme was analyzed in [27], verifying an $O(h^2 + \tau)$ error estimate. In [21, 22] Du et al. presented a discrete semi-implicit scheme for Willmore flow based on the above phase field energy $w^\varepsilon[\cdot]$. We will compare their model with the one presented here in Section 5.2.

Our phase field model is based on an approximation of the mean curvature $\mathbf{h}$ by a time-discrete, approximate speed of the mean curvature evolution, which is described by a fully implicit mean curvature motion time step. Applying the concept of natural time discretization of gradient flows to Willmore flow, we take into account an outer variational problem which reflects the balance between the L^2-distance of the surfaces at two consecutive time steps and the decay of the Willmore energy. To evaluate the Willmore energy at the next time step, an inner variational problem for the natural time discretization of mean curvature motion is solved, and the resulting time-discrete speed is taken as the mean curvature argument $\mathbf{h}$. Hence, time discrete Willmore flow is phrased as a nested variational problem leading to a PDE constraint optimization problem to be solved in each time step.

The structure of both minimization problems is closely related to the variational time discretization of mean curvature motion based on the work of Almgren, Taylor and Wang [1] and Luckhaus and Sturzenhecker [36]. In [7] Chambolle presented a corresponding level set algorithm for this type of variational time stepping. Our inner variational problem represents a phase field analogue of these approaches, and also the outer variational problem, representing the actual Willmore flow time step, is based on the same approach using a Modica–Mortola type phase field description. We aim for a stable time discretization, which allows for large time steps. For a related model in case of explicitly parametrized curves and surfaces based on finite elements on a triangulation of the evolving geometry we refer to Olischläger and Rumpf [44].

This paper is structured as follows. In Section 2 the model is derived, including the nested time discretization, the approximation of the mean curvature, and the optimization aspect. The next Section focuses on the finite element discretization and the Newton method which is necessary to solve the problem. In Section 4 we present a compact discussion of the convergence analysis for the analogues discretization of a linear model problem and validate the result with test simulations. After this we compute the experimental error of our Willmore flow scheme for radially symmetric Willmore flow and compare our model with the one from Du et al. [21, 22] in Section 5. In the last section we consider an elastic model for edge restoration in images [40] as an exemplary application of our model.

2 Derivation of the model

In this section we will derive the nested variational scheme in the phase field context and deduce the Lagrangian of the resulting constrained optimization problem in the continuous setup to prepare the later formulation of a Newton scheme based on a finite element discretization.

2.1 Nested time discretization of Willmore flow

We consider a hypersurface Γ contained in some computational domain $\Omega \subset \mathbb{R}^d$ and evolving under Willmore flow. Let us denote by $x : \Gamma \to \Gamma$ the identity on the surface Γ and introduce the associated Willmore energy $w[x] = \frac{1}{2} \int_\Gamma \mathbf{h}^2(x) \, \mathrm{d}\mathcal{H}^{d-1}$, where $\mathbf{h}$ represents the mean curvature. Then, Willmore flow is defined as the L^2-gradient flow for the Willmore energy, and we obtain the evolution problem

$$(\partial_t x, \vartheta)_{L^2(\Gamma)} = -\partial_x w[x](\vartheta)$$

for all test functions $\vartheta \in C^\infty(\Gamma, \mathbb{R}^d)$, defining a one parameter family of hypersurfaces $\{\Gamma(t)\}_{t \geq 0}$ with $x(t)$ being the identity mapping on $\Gamma(t)$. Here, $(\cdot, \cdot)_{L^2(\Gamma)}$ denotes the L^2-metric on the interface Γ and $\partial_x w[x](\vartheta)$ the variation of the Willmore energy in a direction ϑ. Now, we approximate the mean curvature by the time-discrete propagation speed of mean curvature motion. In fact, if the parametrization $y : \Gamma \to \mathbb{R}^d$ of a hypersurface is the approximate solution of the mean curvature evolution at time $\tilde{\tau}$ with initial parametrization $x : \Gamma \to \Gamma$ being the identity, then we consider $\frac{|y-x|}{\tilde{\tau}}$ as an approximation of $\mathbf{h}$. Therefore, following the idea of Luckhaus and Sturzenhecker [36], we define y for given x as a minimizer of the functional

$$e_{\mathrm{in}}[x, y] = \|y - x\|^2_{L^2(\Gamma)} + 2\tilde{\tau} \mathcal{H}^{d-1}[y(\Gamma)], \tag{2}$$

arising from the natural time discretization approach of mean curvature motion. Here, $\mathcal{H}^{d-1}[y(\Gamma)]$ denotes the area functional. Next, based on this approximation of the mean curvature we approximate the Willmore functional by $\frac{1}{2} \int_\Gamma \frac{(y-x)^2}{\tilde{\tau}^2} \mathrm{d}\mathcal{H}^{d-1}$ and take this into account in the following approach for the natural time discretization of Willmore flow. Given an approximation x^k to $x(k\tau)$ we define the solution x^{k+1} at time $(k+1)\tau$ as the minimizing x of the functional

$$e_{\mathrm{out}}[x^k, x, y] = \|x - x^k\|^2_{L^2(\Gamma^k)} + \tau \int_\Gamma \frac{(y-x)^2}{\tilde{\tau}^2} \mathrm{d}\mathcal{H}^{d-1} \tag{3}$$

under the constraint that for given x the mapping y is a minimizer of $e_{\mathrm{in}}[x, \cdot]$. Here, x^k is the identity map on Γ^k and with a slight misuse of notation we write x instead of $x \circ x^k$. We refer to [44, 43] for further details on this nested time stepping approach for Willmore flow in the parametric context.

Now, we proceed with the transfer of this variational approach to the phase field framework. We suppose the hypersurfaces to be represented by a Modica–Mortola type phase-field function [38]. In detail, we assume that u^k, u, and v are phase field representations of the hypersurfaces Γ^k with the identity map x^k as parametrization, Γ with x as parametrization, and $y(\Gamma)$, respectively. We consider the interfacial energy

$$a^\varepsilon[v] = \frac{1}{2} \int_\Omega \varepsilon |\nabla v|^2 + \frac{1}{\varepsilon} \Psi(v) \, \mathrm{d}x \tag{4}$$

with the double well potential $\Psi(v) = (1 - v^2)^2$. It is well-known that the Γ-limit of $a^\varepsilon[\cdot]$ is $\frac{4}{3} \mathcal{H}^{d-1}[\cdot]$, where $\mathcal{H}^{d-1}[\cdot]$ is the area functional on the sharp interface limit of the phase field model [38], and $\frac{4}{3} = \int_{-1}^1 \sqrt{\Psi(s)} \, \mathrm{d}s$. For sgndist$(\cdot, \Gamma)$ being the signed distance function of the interface Γ and for fixed ε the optimal profile of a phase field representation u_ε of this interface is given by $\tanh\left(\frac{1}{\varepsilon} \mathrm{sgndist}(y, \Gamma)\right)$. This is due to the observation that $\tilde{v} : \mathbb{R} \to \mathbb{R}; t \mapsto \tanh(t)$ solves $\tilde{v}'(t) = \sqrt{\Psi(\tilde{v}(t))}$ and thus $t \mapsto \tilde{v}_\varepsilon(t)$ with $\tilde{v}_\varepsilon(t) = \tilde{v}(\frac{t}{\varepsilon})$ balances the two energy terms in $a^\varepsilon[\cdot]$ on normal slices across the

4

interface, leading to the total energy $\int_{-1}^{1} \sqrt{\Psi(s)}\,\mathrm{d}s$ and thereby to the above constant $\frac{4}{3}$ in front of the area functional. Now, we compare—at first in one dimension—a shifted profile with the optimal profile and obtain

$$
\varepsilon \int_{\mathbb{R}} \left(\tilde{v}_\varepsilon(t) - \tilde{v}_\varepsilon(t-\delta)\right)^2 \mathrm{d}t = \varepsilon \int_{\mathbb{R}} \left(\tilde{v}_\varepsilon'(t)\delta + \delta^2 \int_0^1 (s-1)\tilde{v}_\varepsilon''(t-s\delta)\,\mathrm{d}s\right)^2 \mathrm{d}t
$$

$$
= \varepsilon \int_{\mathbb{R}} (\tilde{v}_\varepsilon'(t))^2\delta^2 + 2\left(\tilde{v}_\varepsilon'(t)\int_0^1 (s-1)\tilde{v}_\varepsilon''(t-s\delta)\,\mathrm{d}s\right)\delta^3
$$

$$
+ \left(\int_0^1 (s-1)\tilde{v}_\varepsilon''(t-s\delta)\,\mathrm{d}s\right)^2 \delta^4\,\mathrm{d}t = \frac{4}{3}\delta^2(1+\theta(\delta,\varepsilon))\,,
$$

where $\theta(\delta,\varepsilon)$ represents the last two terms in the above formula devided by $\frac{4}{3}\delta^2$ and can be estimated as follows,

$$
\theta(\delta,\varepsilon) \leq \frac{3}{2}\varepsilon\delta \left(\int_{\mathbb{R}} (\tilde{v}_\varepsilon'(t))^2\,\mathrm{d}t\right)^{\frac{1}{2}} \left(\int_{\mathbb{R}} \left(\int_0^1 (s-1)\,\tilde{v}_\varepsilon''(t-s\delta)\,\mathrm{d}s\right)^2 \mathrm{d}t\right)^{\frac{1}{2}}
$$

$$
+ \frac{3}{4}\varepsilon\delta^2 \int_{\mathbb{R}} \int_0^1 \left(\tilde{v}_\varepsilon''(t-s\delta)\right)^2 \mathrm{d}s\,\mathrm{d}t
$$

$$
\leq C\left(\frac{\delta}{\varepsilon} + \frac{\delta^2}{\varepsilon^2}\right) =: \Theta(\delta,\varepsilon)\,.
$$

Here, we have used that $\varepsilon \int_{\mathbb{R}} (\tilde{v}_\varepsilon''(t))^2\,\mathrm{d}t \leq C\varepsilon^{-2}$ and once again that $\varepsilon \int_{\mathbb{R}} (\tilde{v}_\varepsilon'(t))^2\,\mathrm{d}t = \frac{4}{3}$. For $\delta = \varepsilon^\beta$, $\beta > 1$, we obtain $\Theta(\delta,\varepsilon) \leq C\delta^{(1-\frac{1}{\beta})}$. From this, we deduce for a function u_ε in higher dimensions taking the optimal profile normal to the interface Γ,

$$
\varepsilon \int_\Omega \left(u_\varepsilon(x+\delta(x)n(x)) - u_\varepsilon(x)\right)^2 \mathrm{d}x = \frac{4}{3}\int_\Gamma \delta^2(x)\mathrm{d}\mathcal{H}^{d-1}\left(1 + O(\Theta(\|\delta\|_\infty,\varepsilon))\right)\,,
$$

where now δ is some function on Γ and n is the normal on Γ (both δ and n are assumed to be extended constantly in normal direction to Γ). Next, assuming optimal profiles of all involved phase field functions u^k, u, and v we observe that

$$
\varepsilon\|v-u\|_{L^2(\Omega)}^2 = \frac{4}{3}\|y-x\|_{L^2(\Gamma)}^2 \left(1 + O(\Theta(\|y-x\|_{L^\infty(\Gamma)},\varepsilon))\right)\,,
$$

$$
\varepsilon\|u-u^k\|_{L^2(\Omega)}^2 = \frac{4}{3}\|x-x^k\|_{L^2(\Gamma^k)}^2 \left(1 + O(\Theta(\|x-x^k\|_{L^\infty(\Gamma^k)},\varepsilon))\right)\,.
$$

Finally, in the spirit of these approximations we can rephrase the energies $e_{\mathrm{in}}[\cdot]$ in (2) and $e_{\mathrm{out}}[\cdot]$ in (3) in terms of these phase field functions,

$$
e_{\mathrm{out}}^\varepsilon[u^k,u,v] = \varepsilon\|u-u^k\|_{L^2(\Omega)}^2 + \frac{\tau\varepsilon}{\tilde{\tau}^2}\|v-u\|_{L^2(\Omega)}^2\,, \tag{5a}
$$

$$
e_{\mathrm{in}}^\varepsilon[u,v] = \varepsilon\|v-u\|_{L^2(\Omega)}^2 + 2\tilde{\tau}a^\varepsilon[v]\,. \tag{5b}
$$

Altogether for sufficiently small phase field parameter ε and sufficiently small time steps τ, $\tilde{\tau}$ this leads to the following nested variational time discretization of Willmore flow:

Definition 1 (Nested variational time discretization of Willmore flow)**.** *Given a phase field u^k at time $k\tau$ define the phase field u^{k+1} at time $(k+1)\tau$ by*

$$
u^{k+1} = \operatorname*{argmin}_{u \in L^2(\Omega)} e_{\mathit{out}}^\varepsilon[u^k,u,v]\,, \quad \text{where} \tag{6a}
$$

$$
v = \operatorname*{argmin}_{\tilde{v} \in H^1(\Omega)} e_{\mathit{in}}^\varepsilon[u,\tilde{v}]\,. \tag{6b}
$$

5

Furthermore, we denote the solution v of the inner variational problem (6b) *for u =*
u^{k+1} *by* v^{k+1}.

In our above approximation arguments δ plays the role of local distance between the
evolving hypersurfaces at two consecutive time steps either of mean curvature motion
or of Willmore flow. Hence, the approximation is effectively valid if these distances
are sufficiently smaller than the phase field parameter ε. The time step $\tilde{\tau}$ in the inner
minimization problem is associated with the accuracy with which the mean curvature
in the Willmore functional of the outer minimization problem is evaluated, while the
outer time step τ defines the actual time scale at which the Willmore flow is resolved.
Thus, from a modeling perspective $\tilde{\tau}$ can be assumed to be significantly smaller than
τ.

The constraint (6b) can also be expressed in terms of the Euler–Lagrange equation
to (5b), that is, v has to solve (in a weak sense)

$$\varepsilon \frac{v-u}{\tilde{\tau}} + \frac{1}{2\varepsilon}\Psi'(v) - \varepsilon\Delta v = 0\,.$$

The following theorem states an existence result for this time-discrete Willmore flow
model:

Theorem 2 (Existence of a time-discrete phase field solution)**.** *Let Ω be a bounded set
in $\mathbb{R}^d$ with Lipschitz boundary and suppose $u^0 \in L^2(\Omega)$. Then there exists a sequence
$((u^k, v^k))_{k=1,\dots} \subset L^2(\Omega) \times H^1(\Omega)$ of solutions of* (6a) *and* (6b).

Proof. At first we consider the variational problem (6b) for given $u^k \in L^2(\Omega)$. By
the direct method in the calculus of variations we immediately obtain a minimizer
$v^k = v^k[u^k] \in H^1(\Omega)$. Now, we consider for fixed $k > 0$ and given $u^{k-1} \in L^2(\Omega)$ a
minimizing sequence $(u_j^k, v_j^k)_{j=1,\dots}$ with v_j^k being a minimizer of $e_{\mathrm{in}}^\varepsilon[u_j^k, \cdot]$. We obtain
that u_j^k and v_j^k are uniformly bounded in $L^2(\Omega)$ and $H^1(\Omega)$, respectively. Thus, we
can extract a subsequence, again denoted by $(u_j^k, v_j^k)_{j=1,\dots}$, such that u_j^k converges
weakly in $L^2(\Omega)$ to some u^k and v_j^k converges weakly in $H^1(\Omega)$ to some v^k. By a
standard convexity argument we then obtain $e_{\mathrm{out}}^\varepsilon[u^{k-1}, \cdot, \cdot]$ in (6a) is weakly lower
semicontinuous, i. e. $\liminf_{j\to\infty} e_{\mathrm{out}}^\varepsilon[u^{k-1}, u_j^k, v_j^k] \geq e_{\mathrm{out}}^\varepsilon[u^{k-1}, u^k, v^k]$.

It remains to prove that $e_{\mathrm{in}}^\varepsilon[u^k, v^k] = \min_{v\in H^1(\Omega)} e_{\mathrm{in}}^\varepsilon[u^k, v]$. Let us assume that there
is a v^* with $e_{\mathrm{in}}^\varepsilon[u^k, v^k] - e_{\mathrm{in}}^\varepsilon[u^k, v^*] > 0$, then we obtain

$$0 < e_{\mathrm{in}}^\varepsilon[u^k, v^k] - e_{\mathrm{in}}^\varepsilon[u^k, v^*] = \varepsilon\int_\Omega (u^k - v^k)^2 - (u^k - v^*)^2\,\mathrm{d}x + 2\tilde{\tau}(a^\varepsilon[v^k] - a^\varepsilon[v^*])$$

$$= \varepsilon\int_\Omega 2u^k(v^* - v^k) + (v^k)^2 - (v^*)^2\,\mathrm{d}x + 2\tilde{\tau}(a^\varepsilon[v^k] - a^\varepsilon[v^*])$$

$$\leq \liminf_{j\to\infty} \varepsilon\int_\Omega 2u_j^k(v^* - v_j^k) + (v_j^k)^2 - (v^*)^2\,\mathrm{d}x + 2\tilde{\tau}(a^\varepsilon[v_j^k] - a^\varepsilon[v^*])$$

$$= \liminf_{j\to\infty}\left(e_{\mathrm{in}}^\varepsilon[u_j^k, v_j^k] - e_{\mathrm{in}}^\varepsilon[u_j^k, v^*]\right)\,,$$

where we have used the lower semi-continuity of $a^\varepsilon[\cdot]$, the weak convergence of v_j^k to
v^k in $H^1(\Omega)$ and the weak convergence of u_j^k to u^k in $L^2(\Omega)$, and that by Rellich's em-
bedding theorem v_j^k already converges strongly in $L^2(\Omega)$, which implies $v_j^k u_j^k \rightharpoonup v^k u^k$
weakly in $L^1(\Omega)$. Hence, there exists an index j with $e_{\mathrm{in}}^\varepsilon[u_j^k, v_j^k] - e_{\mathrm{in}}^\varepsilon[u_j^k, v^*] > 0$, con-
tradicting our assumption that $e_{\mathrm{in}}^\varepsilon[u_j^k, v_j^k] = \min_{v\in H^1(\Omega)} e_{\mathrm{in}}^\varepsilon[u_j^k, v]$. Thus, $e_{\mathrm{in}}^\varepsilon[u^k, v^k] =
\min_{v\in H^1(\Omega)} e_{\mathrm{in}}^\varepsilon[u^k, v]$ and (u^k, v^k) solves the variational problem (6a) and (6b). $\quad\square$

Let us remark, that because of the non-convexity of Ψ the solutions of (6a) and (6b) are not necessarily unique.

2.2 The constrained optimization perspective

The time stepping scheme for Willmore flow introduced in Definition 1 requires the solution of a variational problem with a PDE constraint in each time step, which can be phrased in the context of constrained optimization in terms of searching for a saddle point of the associated Lagrangian ℓ. This Lagrangian is defined as the sum of the outer variational functional (6a) and the variation of the inner functional (6b) in the direction of a dual function p which acts as the Lagrange multiplier. Hence, we obtain for the Lagrangian

$$
\ell[u^k, u, v, p] = e_{\text{out}}^\varepsilon[u^k, u, v] + \partial_v e_{\text{in}}^\varepsilon[u, v](p)
$$
$$
= \int_\Omega \varepsilon(u - u^k)^2 + \frac{\varepsilon\tau}{\tilde{\tau}^2}(v - u)^2 \, \mathrm{d}x + \int_\Omega 2\varepsilon(v - u)p + \frac{\tilde{\tau}}{\varepsilon}\Psi'(v)p + 2\tilde{\tau}\varepsilon\nabla v \cdot \nabla p \, \mathrm{d}x \ . \quad (7)
$$

Consequently, to identify a saddle $(u, v, p) \in L^2(\Omega) \times H^1(\Omega) \times H^1(\Omega)$ we have to solve $\nabla_{(u,v,p)}\ell[u^k, u, v, p] = 0$, where the different components read

$$
\partial_u \ell[u^k, u, v, p](\vartheta) = \int_\Omega 2\varepsilon(u - u^k)\vartheta - \frac{2\tau\varepsilon}{\tilde{\tau}^2}(v - u)\vartheta - 2\varepsilon\vartheta p \, \mathrm{d}x \ ,
$$
$$
\partial_v \ell[u^k, u, v, p](\xi) = \int_\Omega \frac{2\tau\varepsilon}{\tilde{\tau}^2}(v - u)\xi + 2\varepsilon p\xi + \frac{\tilde{\tau}}{\varepsilon}\Psi''(v)p\xi + 2\tilde{\tau}\varepsilon\nabla p \cdot \nabla \xi \, \mathrm{d}x \ ,
$$
$$
\partial_p \ell[u^k, u, v, p](\varsigma) = \int_\Omega 2\varepsilon(v - u)\varsigma + \frac{\tilde{\tau}}{\varepsilon}\Psi'(v)\varsigma + 2\tilde{\tau}\varepsilon\nabla v \cdot \nabla \varsigma \, \mathrm{d}x
$$

for all $\vartheta \in L^2(\Omega)$, $\xi \in H^1(\Omega)$, and $\varsigma \in H^1(\Omega)$. Here, $\partial_p \ell = 0$ reflects the Euler–Lagrange equation of the inner variational problem (6b), $\partial_v \ell = 0$ is the dual problem defining a dual variable p, and finally, the primal problem $\partial_u \ell = 0$ is the actual Lagrangian multiplier formulation of the nested variational problem.

Due to the double well potential $\Psi(v) = (1 - v^2)^2$, the variation of the Lagrangian in v and p is nonlinear in v. Thus, in the finite element context we will consider a Newton scheme to solve for a saddle point.

3 A fully practical numerical scheme

Now we introduce a spatial discretization using piecewise affine finite elements and derive a Newton scheme for the corresponding discrete Lagrangian. In what follows we will use lower case letters for continuous and upper case letters for discrete quantities.

3.1 Finite element discretization in space

We take into account an admissible, regular, and uniform simplicial mesh $\mathcal{T}$ covering the computational domain Ω. On this mesh we define the finite element space of continuous, piecewise affine functions $\mathcal{V} := \{\Phi \in C^0(\Omega) \,|\, \Phi|_T \in \mathcal{P}_1 \, \forall \, T \in \mathcal{T}\}$, where $\mathcal{P}_1$ is the space of affine functions. The finite element space $\mathcal{V}$ is spanned by the usual nodal basis $\{\Phi_i\}_{i \in I}$, where I is the vertex index set of $\mathcal{T}$. Moreover the nodal coordinate vector to a discrete scalar function $F = \sum_{i \in I} F_i \Phi_i \in \mathcal{V}$ is denoted by $\overline{F} = (F_i)_{i \in I}$.

To derive a fully discrete problem we define a weighted lumped mass matrix $\mathbf{M}[\lambda] = \left(\int_\Omega \mathcal{I}_h(\lambda\, \Phi_i \Phi_j)\, dx \right)_{i,j}$ with λ being a general continuous weight function and $\mathcal{I}_h$ the Lagrange interpolation, the classical lumped mass matrix $\mathbf{M} = \mathbf{M}[1]$, and the stiffness matrix $\mathbf{L} = \left(\int_\Omega \nabla\Phi_i \cdot \nabla\Phi_j\, dx \right)_{i,j}$. Based on this notation we obtain discrete counterparts

$$\mathcal{E}^\varepsilon_{\mathrm{out}}[U^k, U, V] = \varepsilon\mathbf{M}\left(\overline{U} - \overline{U}^k \right) \cdot \left(\overline{U} - \overline{U}^k \right) + \frac{\tau\varepsilon}{\tilde{\tau}^2}\mathbf{M}\left(\overline{V} - \overline{U} \right) \cdot \left(\overline{V} - \overline{U} \right) \tag{8a}$$

$$\mathcal{E}^\varepsilon_{\mathrm{in}}[U, V] = \varepsilon\mathbf{M}\left(\overline{V} - \overline{U} \right) \cdot \left(\overline{V} - \overline{U} \right) + 2\tilde{\tau} A^\varepsilon[V] \tag{8b}$$

of the functionals $e^\varepsilon_{\mathrm{out}}[\cdot]$ in (5a) and $e^\varepsilon_{\mathrm{in}}[\cdot]$ in (5b). Here, the discrete phase field energy A^ε is given by

$$A^\varepsilon[V] = \frac{1}{2}\int_\Omega \varepsilon\nabla V \cdot \nabla V + \frac{1}{\varepsilon}\mathcal{I}_h\left(\Psi(V) \right) dx = \frac{\varepsilon}{2}\mathbf{L}\overline{V} \cdot \overline{V} + \frac{1}{2\varepsilon}\mathbf{M}\overline{\Psi(V)} \cdot \overline{1}.$$

Finally, we are lead to the following fully discrete minimization problem to be solved in each time step.

Definition 3 (Fully discrete variational time discretization of Willmore flow). *Given a discrete phase field $U^k \in \mathcal{V}$ at time $k\tau$ define the phase field $U^{k+1} \in \mathcal{V}$ at time $(k+1)\tau$ by*

$$U^{k+1} = \operatorname*{argmin}_{U \in \mathcal{V}} \mathcal{E}^\varepsilon_{out}[U^k, U, V]\,, \text{ where} \tag{9a}$$

$$V = \operatorname*{argmin}_{\tilde{V} \in \mathcal{V}} \mathcal{E}^\varepsilon_{in}[\tilde{V}, U]\,, \tag{9b}$$

and denote by V^{k+1} the solution V of (9b) for $U = U^{k+1}$.

Let us remark that existence of a solution (U^{k+1}, V^{k+1}) for this fully discrete problem is straightforward.

3.2 Newton method for the constrained optimization problem

To solve the above fully discrete constrained optimization problem (9a) and (9b) we aim at finding a saddle point of the associated discrete Lagrangian $\mathcal{L}$ using Newton's method, where

$$\mathcal{L}[U^k, U, V, P] = \mathcal{E}^\varepsilon_{\mathrm{out}}[U^k, U, V] + \partial_V \mathcal{E}^\varepsilon_{\mathrm{in}}[U, V](P) \tag{10}$$

$$= \varepsilon\mathbf{M}\left(\overline{U} - \overline{U}^k \right) \cdot \left(\overline{U} - \overline{U}^k \right) + \frac{\tau\varepsilon}{\tilde{\tau}^2}\mathbf{M}\left(\overline{V} - \overline{U} \right) \cdot \left(\overline{V} - \overline{U} \right)$$

$$+ 2\varepsilon\mathbf{M}\left(\overline{V} - \overline{U} \right) \cdot \overline{P} - \frac{4\tilde{\tau}}{\varepsilon}\mathbf{M}[1 - V^2]\overline{V} \cdot \overline{P} + 2\tilde{\tau}\varepsilon\mathbf{L}\overline{V} \cdot \overline{P}.$$

To this end, we define $\mathcal{F}[U, V, P] := \nabla_{(U,V,P)}\mathcal{L}[U^k, U, V, P]$ and obtain for the different components

$$\partial_U \mathcal{L}[U^k, U, V, P](\Theta) = 2\varepsilon\mathbf{M}\left(\overline{U} - \overline{U}^k \right) \cdot \overline{\Theta} - \frac{2\tau\varepsilon}{\tilde{\tau}^2}\mathbf{M}\left(\overline{V} - \overline{U} \right) \cdot \overline{\Theta} - 2\varepsilon\mathbf{M}\overline{P} \cdot \overline{\Theta},$$

$$\partial_V \mathcal{L}[U^k, U, V, P](\Xi) = \frac{2\tau\varepsilon}{\tilde{\tau}^2}\mathbf{M}\left(\overline{V} - \overline{U} \right) \cdot \overline{\Xi} + 2\varepsilon\mathbf{M}\overline{P} \cdot \overline{\Xi}$$

$$- \frac{4\tilde{\tau}}{\varepsilon}\mathbf{M}[1 - 3V^2]\overline{P} \cdot \overline{\Xi} + 2\tilde{\tau}\varepsilon\mathbf{L}\overline{P} \cdot \overline{\Xi},$$

$$\partial_P \mathcal{L}[U^k, U, V, P](\Sigma) = 2\varepsilon\mathbf{M}\left(\overline{V} - \overline{U} \right) \cdot \overline{\Sigma} - \frac{4\tilde{\tau}}{\varepsilon}\mathbf{M}[1 - V^2]\overline{V} \cdot \overline{\Sigma} + 2\tilde{\tau}\varepsilon\mathbf{L}\overline{V} \cdot \overline{\Sigma}.$$

The Hessian of the discretized Lagrangian in matrix form then reads

$$D^2\mathcal{L} = \begin{pmatrix} 2\varepsilon\left(1+\frac{\tau}{\tilde{\tau}^2}\right)\mathbf{M} & -\frac{2\tau\varepsilon}{\tilde{\tau}^2}\mathbf{M} & -2\varepsilon\mathbf{M} \\ -\frac{2\tau\varepsilon}{\tilde{\tau}^2}\mathbf{M} & \frac{2\varepsilon\tau}{\tilde{\tau}^2}\mathbf{M}+\frac{24\tilde{\tau}}{\varepsilon}\mathbf{M}[VP] & 2\varepsilon\mathbf{M}-\frac{4\tilde{\tau}}{\varepsilon}\mathbf{M}[1-3V^2]+2\tilde{\tau}\varepsilon\mathbf{L} \\ -2\varepsilon\mathbf{M} & 2\varepsilon\mathbf{M}-\frac{4\tilde{\tau}}{\varepsilon}\mathbf{M}[1-3V^2]+2\tilde{\tau}\varepsilon\mathbf{L} & 0 \end{pmatrix}.$$

Now, we apply the following Newton iteration to find a root of $\mathcal{F}$ and thus a critical point of $\mathcal{L}$. Given U_i, V_i, and P_i as an approximation for U^{k+1}, V^{k+1}, and the associated dual state P^{k+1} and defining $Z_i := (U_i, V_i, P_i)$ we solve the linear system of equations

$$D\mathcal{F}[Z_i]^T D\mathcal{F}[Z_i](Z_{i+1} - Z_i) = -D\mathcal{F}[Z_i]^T \mathcal{F}[Z_i]$$

for Z_{i+1} with $D\mathcal{F} = D^2\mathcal{L}$. In our implementation we employ the Cholesky solver from the CHOLMOD package [13, 10]. In each time step we take as initial data for the Newton scheme $U_0 = U^k$ and compute V_0 as a root of $\partial_P\mathcal{L}[U_0, \cdot, P] = 0$ and P_0 as the solution of the linear system of equations $\partial_V\mathcal{L}[U_0, V_0, P] = 0$ in P. It turned out that in all our applications a time step control in the Newton scheme is not required.

4 Numerical analysis for a related linear model problem

In order to gain insight into the convergence behavior of the proposed scheme, we consider its analogon for the linear model problem, the L^2-gradient flow of the energy $w^l[u] = \frac{1}{2}\int_\Omega(\Delta u)^2\,dx$ on a polygonally bounded, convex domain Ω. As before, upper case letters denote the spatially discretized versions of their lower case, spatially continuous counterparts. Furthermore, we will distinguish between time-discrete and time-continuous quantities, the latter being indicated by a tilde as in $\tilde{u}$. In fact, we are lead to the biharmonic heat equation

$$\partial_t\tilde{u} = -\Delta^2\tilde{u} \quad \text{on } \Omega \tag{11}$$

with initial condition $\tilde{u} = \tilde{u}^0$ at $t = 0$. Any corresponding L^2-gradient flow depends on the imposed boundary conditions, for which there are several possibilities: natural boundary conditions as for the exposition in Section 2, periodic boundary conditions on a fundamental cell $\Omega = (0,1)^d$, or Dirichlet boundary conditions as in Section 4.2 on the numerical validation and in Section 6 on image restoration. In the variational formulation this is reflected by different ansatz spaces $\bar{H}^1$: $\bar{H}^1 = \{u \in H^1 \mid u = 0 \text{ on } \partial\Omega\}$ in case of Dirichlet boundary conditions $u = \Delta u = 0$ on $\partial\Omega$; $\bar{H}^1 = H^1 \cap L_0^2$ with $L_0^2 = \{u \in L^2 \mid \int_\Omega u\,dx = 0\}$ in case of natural boundary conditions $\partial_\nu u = \partial_\nu\Delta u = 0$ on $\partial\Omega$ for the normal derivative ∂_ν on $\partial\Omega$ with outer normal ν; and $\bar{H}^1 = \{u \in H^1 \cap L_0^2 \mid u(x + e_i) = u(x) \text{ on } \partial\Omega\}$. Furthermore, as in [53] we define $\dot{H}^s$ as the subspace of functions with s weak derivatives which is associated with the spectral norm belonging to Δ^s. For example in the case of Dirichlet boundary conditions $\Delta^m u = 0$ on $\partial\Omega$ for $m < \frac{s}{2}$.

4.1 A nested scheme for the biharmonic heat equation

For the variational time discretization, we introduce the Dirichlet energy $a^l[u] = \frac{1}{2}\int_\Omega|\nabla u|^2\,dx$ and in analogy to (5a) and (5b) define

$$e_{\text{out}}^l[u^k, u, v] = \|u - u^k\|_{L^2}^2 + \frac{\tau}{\tilde{\tau}^2}\|v - u\|_{L^2}^2\,, \tag{12a}$$

$$e_{\text{in}}^l[u, v] = \|v - u\|_{L^2}^2 + 2\tilde{\tau}a^l[v]\,. \tag{12b}$$

Hence, as a linear counterpart of the variational time discretization for Willmore flow in Definition 1 we obtain the following time-discrete scheme for the biharmonic heat equation.

Definition 4 (Time-discrete biharmonic heat flow)**.** *For given* $u^0 \in \bar{H}^1$*, let the sequence* $u^k \in \bar{H}^1$*,* $k \in \mathbb{N}$*, be defined as the solution of*

$$u^{k+1} = \operatorname*{argmin}_{u \in \bar{H}^1} e^l_{out}[u^k, u, v], \text{ where} \tag{13a}$$

$$v = \operatorname*{argmin}_{\tilde{v} \in \bar{H}^1} e^l_{in}[u, \tilde{v}]. \tag{13b}$$

Next, introducing the finite element space $\bar{\mathcal{V}} = \mathcal{V} \cap \bar{H}^1$ we derive an associated fully discrete scheme.

Definition 5 (Fully discretized biharmonic heat flow)**.** *For given* $U^0 \in \bar{\mathcal{V}}$*, the sequence* $U^k \in \bar{\mathcal{V}}$*,* $k \in \mathbb{N}$*, is defined as the solution of*

$$U^{k+1} = \operatorname*{argmin}_{U \in \bar{\mathcal{V}}} \mathcal{E}^l_{out}[U^k, U, V], \text{ where} \tag{14a}$$

$$V = \operatorname*{argmin}_{\tilde{V} \in \bar{\mathcal{V}}} \mathcal{E}^l_{in}[U, \tilde{V}]. \tag{14b}$$

As in Section 2 one can also consider $\mathcal{V}$ as the ansatz space in case of natural boundary conditions and for $U^0 \in \bar{\mathcal{V}}$. We will assume $\mathcal{E}^l_{\mathrm{out}} = e^l_{\mathrm{out}}$ and $\mathcal{E}^l_{\mathrm{in}} = e^l_{\mathrm{in}}$ to be evaluated exactly in the fully discrete scheme, which is equivalent to assuming at least second order quadrature. Let us remark, that in this linear problem the inner heat equation time step $\tilde{\tau}$ is associated with the accuracy with which the Laplace operator is approximated, while the outer time step τ defines the time scale at which the bi-Laplace is resolved. Hence, as in the nonlinear case of Willmore flow, it conceptually does not make sense to increase $\tilde{\tau}$ beyond τ. Nevertheless, the scheme stably approximates the biharmonic heat flow for any combination of time steps $\tau, \tilde{\tau}$ (cf. the following theorem).

In what follows, we will investigate the convergence behavior of this fully discrete model. The subsequent numerical analysis reflects some of the qualitative aspects of the nested variational approach in general. Furthermore, the observed experimental convergence behavior of the phase field model for Willmore flow will be in accordance with the results obtained in this linear setup.

We will consider quasiuniform triangulations with grid size h. Let $\mathcal{P}_h$ denote the local L^2-projection onto $\bar{\mathcal{V}}$ [12]. If we denote by $-\Delta_h \Phi \in \bar{\mathcal{V}}$ the discrete Laplacian, i.e. the unique solution of $(-\Delta_h \Phi, \Theta)_{L^2} = (\nabla \Phi, \nabla \Theta)_{L^2} \ \forall \Theta \in \bar{\mathcal{V}}$, then note that for $u \in H^2$, $\Delta_h \mathcal{P}_h(u)$ is bounded in $L^2(\Omega)$, uniformly in h. Indeed, $\|\mathcal{P}_h(u)\|_{L^2}$ and $\|\Delta_h^{\frac{1}{2}} \mathcal{P}_h(u)\|_{L^2} = |\mathcal{P}_h(u)|_{H^1}$ are bounded due to the Bramble–Hilbert lemma, and an inverse inequality implies $\|\Delta_h \mathcal{P}_h(u)\|^2_{L^2} = (\nabla(u - \mathcal{P}_h(u)), \nabla \Delta_h \mathcal{P}_h(u))_{L^2} + (\Delta u, \Delta_h \mathcal{P}_h(u))_{L^2} \leq Ch\|\nabla \Delta_h \mathcal{P}_h(u)\|_{L^2} + \|\Delta u\|_{L^2}\|\Delta_h \mathcal{P}_h(u)\|_{L^2} \leq C\|\Delta_h \mathcal{P}_h(u)\|_{L^2}$.

Finally, let us introduce the classical Ritz projection $R_h u \in \bar{\mathcal{V}}$ for $u \in \bar{H}^1$, defined as the unique finite element solution of $(\nabla R_h u, \nabla \Theta)_{L^2} = (\nabla u, \nabla \Theta)_{L^2}$ for all $\Theta \in \bar{\mathcal{V}}$. Note that $\|\Delta_h^m R_h u\|_{L^2}$ for $m \leq 2$ is bounded uniformly in h if $u \in \dot{H}^4$. For instance, for $m = 2$ we have $\|\Delta_h^2 R_h u\|^2_{L^2} = -(\nabla R_h u, \nabla \Delta_h^3 R_h u)_{L^2} = -(\nabla u, \nabla \Delta_h^3 R_h u)_{L^2} = (\Delta_h \mathcal{P}_h(\Delta u), \Delta_h^2 R_h u)_{L^2} + (\Delta u - \mathcal{P}_h(\Delta u), \Delta_h^3 R_h u)_{L^2} \leq \|\Delta_h \mathcal{P}_h(\Delta u)\|_{L^2}\|\Delta_h^2 R_h u\|_{L^2} + Ch^2 |\Delta u|_{H^2}\|\Delta_h^3 R_h u\|_{L^2} \leq C\|\Delta_h^2 R_h u\|_{L^2}$, where we have exploited the eigenvalue bound $\frac{C}{h^2}$ for $-\Delta_h$. Let us also introduce the projections $\hat{R}_h u, \mathring{R}_h u \in \bar{\mathcal{V}}$ for $u \in \dot{H}^3$ and $u \in \dot{H}^5$, respectively, with $\hat{R}_h u = \Delta_h^{-1} R_h \Delta u$ and $\mathring{R}_h u = \Delta_h^{-2} R_h \Delta^2 u$.

Theorem 6 (Error estimates for fully discrete biharmonic heat flow)**.** *For given* $\tilde{u}^0 \in \dot{H}^8$, *let* $\tilde{u}$ *be the solution to the biharmonic heat equation* (11), *and let* U^k, $k \in \mathbb{N}$, *be the solution of the fully discretized scheme* (14) *with* $U^0 = \hat{R}_h \tilde{u}^0$. *On quasiuniform triangulations with grid size* h *and for uniform time steps* τ *and* $\tilde{\tau}$

$$\|U^k - \tilde{u}(t_k)\|_{L^2} \leq C(t_k)(h^2 + \tau + \tilde{\tau}), \tag{15a}$$

$$\|U^k - \tilde{u}(t_k)\|_{H^1} \leq C(t_k)\left(h + (1 + \tfrac{\sqrt{\tilde{\tau}}}{h})(\tau + \tilde{\tau})\right) \tag{15b}$$

holds for a continuous function $C : \mathbb{R}^+ \to \mathbb{R}^+$.

Proof. For the error analysis, the error $\tilde{u}(t_k) - U^k$ at time $t_k = k\tau$ is split up into the usual difference of a spatially discrete $\tilde{U}(t_k)$ and a spatially continuous solution $\tilde{u}(t_k)$, both time-continuous, and the difference between the time-continuous but spatially discrete $\tilde{U}(t_k)$ and the fully discrete solution U^k. The first error term is rather standard and can be estimated adapting the numerical analysis of the heat equation in [53]. For the sake of completeness, we give a brief exposition here. The variational structure of our scheme is taken into account in the analysis of the second error term.

Step 1 (Time-continuous finite element estimates). Now, we compare $\tilde{u}$ with the spatially discrete solution $\tilde{U}$ of $\partial_t \tilde{U} = -\Delta_h^2 \tilde{U}$, $\tilde{U}(0) = U^0$, which effectively is given as a solution of a system of ODEs. In order to analyse the difference $\tilde{U} - \tilde{u}$ let us introduce $\tilde{w} = -\Delta \tilde{u}$ and $\tilde{W} = -\Delta_h \tilde{U}$. We observe that $(\Delta^2 \tilde{u}, \Theta)_{L^2} = (\nabla R_h \tilde{w}, \nabla \Theta)_{L^2}$ for all $\Theta \in \bar{\mathcal{V}}$ and $(R_h \tilde{w}, \Psi)_{L^2} = (\nabla \hat{R}_h \tilde{u}, \nabla \Psi)_{L^2}$ for all $\Psi \in \bar{\mathcal{V}}$. Now, we can decompose the error according to

$$\tilde{U} - \tilde{u} = (\tilde{U} - \hat{R}_h \tilde{u}) + (\hat{R}_h \tilde{u} - \tilde{u}) =: \theta + \rho, \tag{16a}$$

$$\tilde{W} - \tilde{w} = (\tilde{W} - R_h \tilde{w}) + (R_h \tilde{w} - \tilde{w}) =: \eta + \sigma, \tag{16b}$$

where ρ and σ represent the error of the approximation to the coupled elliptic problems $w = -\Delta u$ and $-\Delta^2 u = \Delta w$. From the definition of the projections R_h and $\hat{R}_h$ and the definition of $\tilde{U}$ we obtain

$$(-\partial_t \theta, \Theta)_{L^2} + (-\partial_t \rho, \Theta)_{L^2} = (\nabla \eta, \nabla \Theta)_{L^2} \qquad \forall \Theta \in \bar{\mathcal{V}}, \tag{17a}$$

$$(\eta, \Psi)_{L^2} = (\nabla \theta, \nabla \Psi)_{L^2} \qquad \forall \Psi \in \bar{\mathcal{V}}. \tag{17b}$$

Testing these two equations and the time derivative of (17b) with $\Theta = \theta$, $\Theta = \partial_t \theta$, $\Theta = \eta$, and $\Psi = \eta$ and $\Psi = \partial_t \theta$ we end up with

$$\frac{\mathrm{d}}{\mathrm{d}t}\|\theta\|_{L^2} \leq \|\partial_t \rho\|_{L^2}, \frac{\mathrm{d}}{\mathrm{d}t}\|\eta\|_{L^2}^2 \leq \|\partial_t \rho\|_{L^2}^2, \frac{\mathrm{d}}{\mathrm{d}t}\|\nabla \theta\|_{L^2}^2 \leq \|\partial_t \rho\|_{L^2}^2 + \|\eta\|_{L^2}^2. \tag{18}$$

Hence, the error analysis of the spatially discretized parabolic problem is reduced to the analysis of a purely elliptic problem. For the error components $\rho = \hat{R}_h \tilde{u} - \tilde{u}$ and $\sigma = R_h \tilde{w} - \tilde{w}$ we obtain

$$0 = (\nabla(R_h \tilde{w} - \tilde{w}), \nabla \Theta)_{L^2} \qquad \forall \Theta \in \bar{\mathcal{V}}, \tag{19a}$$

$$((R_h \tilde{w} - \tilde{w}), \Psi)_{L^2} = (\nabla(\hat{R}_h \tilde{u} - \tilde{u}), \nabla \Psi)_{L^2} \qquad \forall \Psi \in \bar{\mathcal{V}}. \tag{19b}$$

From (19a) we obtain via standard arguments including the usual duality argument on convex domains

$$h\|R_h \tilde{w} - \tilde{w}\|_{H^1} + \|R_h \tilde{w} - \tilde{w}\|_{L^2} \leq Ch^2 \|\tilde{w}\|_{H^2}. \tag{20}$$

11

Furthermore, testing (19b) with $\Psi = \mathcal{P}_h(\tilde{u}) - \hat{R}_h\tilde{u} = \mathcal{P}_h(\tilde{u}) - \tilde{u} + \tilde{u} - \hat{R}_h\tilde{u}$ and taking into account (20), we obtain

$$c\|\tilde{u} - \hat{R}_h\tilde{u}\|_{H^1}^2 \leq \|\hat{R}_h\tilde{u} - \tilde{u}\|_{H^1}h + \|\hat{R}_h\tilde{u} - \tilde{u}\|_{H^1}h^2 + h^4 \qquad (21)$$

for some constant c. From this we deduce $\|\rho\|_{H^1} \leq Ch$ for a different constant C. A duality argument to derive an L^2-bound can also be applied to $\tilde{u} - \hat{R}_h\tilde{u}$ and we achieve $\|\rho\|_{L^2} \leq Ch^2$.

After a differentiation of the elliptic problem with respect to time we obtain by the same sequence of arguments $\|\partial_t\rho\|_{L^2} = \|(\partial_t\tilde{u}) - \hat{R}_h(\partial_t\tilde{u})\|_{L^2} \leq Ch^2$ using the smoothness of the initial data. Altogether, after incorporating (18) and noting that $\|\tilde{U}^0 - \tilde{u}^0\|_{L^2} + h\|\tilde{U}^0 - \tilde{u}^0\|_{H^1} \leq Ch^2$ (iterating once more the arguments for Ritz projections in (19a), (19b), (20)) we obtain

$$\|\tilde{U} - \tilde{u}\|_{L^2} + h\|\tilde{U} - \tilde{u}\|_{H^1} \leq \|\theta\|_{L^2} + \|\rho\|_{L^2} + h(\|\theta\|_{H^1} + \|\rho\|_{H^1}) \leq Ch^2. \qquad (22)$$

From (18) we see that the constant C grows at most quadratically in time. However, this type of behaviour cannot be expected when nonlinear terms are taking into account. Then, indeed exponential growth would result from the application of the usual Gronwall argument.

Step 2 (Estimates for the time discretization). Now we will analyse the error $e^k := U^k - \tilde{U}(t_k)$. By the Euler–Lagrange equations for (14) we obtain $V = (\mathrm{id} - \tilde{\tau}\Delta_h)^{-1}U^k$ and thus

$$\frac{U^k - U^{k-1}}{\tau} = -\left(\frac{(\mathrm{id} - \tilde{\tau}\Delta_h)^{-1} - \mathrm{id}}{\tilde{\tau}}\right)^2 U^k. \qquad (23)$$

Defining a new discrete Laplacian $\Delta_{\tilde{\tau}} = \frac{(\mathrm{id} - \tilde{\tau}\Delta_h)^{-1} - \mathrm{id}}{\tilde{\tau}}$ and $\partial_\tau U^k = \frac{U^k - U^{k-1}}{\tau}$ this can be rewritten as $\partial_\tau U^k = -\Delta_{\tilde{\tau}}^2 U^k$. Based on the identity $(\mathrm{id} - \tilde{\tau}\Delta_h)^{-1} - \mathrm{id} = \tilde{\tau}\Delta_h + \tilde{\tau}^2\Delta_h^2(\mathrm{id} - \tilde{\tau}\Delta_h)^{-1}$ one obtains for the truncation error

$$\omega^k := \partial_\tau\tilde{U}(t_k) + \Delta_{\tilde{\tau}}^2\tilde{U}(t_k) = \partial_t\tilde{U}(t_k) + \frac{1}{\tau}\int_{t_{k-1}}^{t_k}\partial_t^2\tilde{U}(r)(r - t_{k-1})\,\mathrm{d}r + \Delta_{\tilde{\tau}}^2\tilde{U}(t_k)$$

$$= \frac{1}{\tau}\int_{t_{k-1}}^{t_k}\partial_t^2\tilde{U}(r)(r - t_{k-1})\,\mathrm{d}r + (\Delta_{\tilde{\tau}}^2 - \Delta_h^2)\tilde{U}(t_k)$$

$$= \frac{1}{\tau}\int_{t_{k-1}}^{t_k}\Delta_h^4\tilde{U}(r)(r - t_{k-1})\,\mathrm{d}r + \left(\frac{1}{\tilde{\tau}^2}\left(\tilde{\tau}\Delta_h + \tilde{\tau}^2\Delta_h^2(\mathrm{id} - \tilde{\tau}\Delta_h)^{-1}\right)^2 - \Delta_h^2\right)\tilde{U}(t_k)$$

$$= \frac{1}{\tau}\int_{t_{k-1}}^{t_k}\Delta_h^4\tilde{U}(r)(r - t_{k-1})\,\mathrm{d}r + \left(2\tilde{\tau}(\mathrm{id} - \tilde{\tau}\Delta_h)^{-1}\Delta_h^3 + \tilde{\tau}^2\Delta_h^4(\mathrm{id} - \tilde{\tau}\Delta_h)^{-2}\right)\tilde{U}(t_k).$$

Thus, we can estimate

$$\|\omega^k\|_{L^2} \leq C(\tau + \tilde{\tau}). \qquad (24)$$

Here, we have used the boundedness of the discrete operators $(\mathrm{id} - \tilde{\tau}\Delta_h)^{-1}, (\mathrm{id} - \tilde{\tau}\Delta_h)^{-2} : \bar{\mathcal{V}} \to \bar{\mathcal{V}}$ since the eigenvalues of Δ_h lie in $[-\frac{C}{h^2}, 0]$ (cf. [29]) and the estimate $\|\Delta_h^m\tilde{U}(t)\|_{L^2} \leq \|\Delta_h^m\tilde{U}^0\|_{L^2} \leq C$ for $m = 3, 4$. The latter estimate directly follows from an argument in [53] suitably adapted to the representation formula $\tilde{U}(t) = \sum_i \exp(-\lambda_i^2 t)(\tilde{U}(0), V^i)_{L^2}V^i$ for solutions of the spatially discrete biharmonic heat flow, where λ_i are the eigenvalues and V^i the corresponding eigenfunctions of $-\Delta_h$.

From (23) and the definition of the truncation error we obtain

$$\omega^k = \frac{e^k - e^{k-1}}{\tau} + \left(\frac{(\mathrm{id} - \tilde{\tau}\Delta_h)^{-1} - \mathrm{id}}{\tilde{\tau}} \right)^2 e^k . \tag{25}$$

Testing with e^k yields

$$\|e^k\|_{L^2}^2 \leq \|e^k\|_{L^2}^2 + \tau \| \left(\tfrac{(\mathrm{id} - \tilde{\tau}\Delta_h)^{-1} - \mathrm{id}}{\tilde{\tau}} \right) e^k \|_{L^2}^2$$
$$= (e^{k-1}, e^k)_{L^2} + \tau(\omega^k, e^k)_{L^2} \leq \|e^{k-1}\|_{L^2}\|e^k\|_{L^2} + \tau\|\omega^k\|_{L^2}\|e^k\|_{L^2} \tag{26}$$

and thus by induction

$$\|e^k\|_{L^2} \leq \tau \sum_{i=1}^{k} \|\omega^i\|_{L^2} \leq C t_k (\tau + \tilde{\tau}) . \tag{27}$$

For an H^1-error estimate we first aim for an estimate of $\|\Delta_{\tilde{\tau}} e^k\|_{L^2}$, for which purpose we test (25) with $\partial_\tau e^k$ to obtain

$$\|\partial_\tau e^k\|_{L^2}^2 + (\Delta_{\tilde{\tau}}^2 e^k, \partial_\tau e^k)_{L^2} = (\omega^k, \partial_\tau e^k)_{L^2} \leq \frac{1}{2}\|\omega^k\|_{L^2}^2 + \frac{1}{2}\|\partial_\tau e^k\|_{L^2}^2 . \tag{28}$$

Since $\Delta_{\tilde{\tau}}$ is self-adjoint, we readily find

$$(\Delta_{\tilde{\tau}}^2 e^k, \partial_\tau e^k)_{L^2} = (\Delta_{\tilde{\tau}} e^k, \partial_\tau \Delta_{\tilde{\tau}} e^k)_{L^2} \geq \tfrac{1}{2}\partial_\tau \|\Delta_{\tilde{\tau}} e^k\|_{L^2}^2$$

so that altogether $\partial_\tau \|\Delta_{\tilde{\tau}} e^k\|_{L^2}^2 \leq \|\omega^k\|_{L^2}^2$ and thus

$$\|\Delta_{\tilde{\tau}} e^k\|_{L^2}^2 \leq \tau \sum_{i=1}^{k} \|\omega^i\|_{L^2}^2 \leq C t_k (\tau + \tilde{\tau})^2 . \tag{29}$$

As a next step, we deduce from the definition of $\Delta_{\tilde{\tau}}$ that $(\mathrm{id} + \tilde{\tau}\Delta_{\tilde{\tau}})(\mathrm{id} - \tilde{\tau}\Delta_h) = \mathrm{id}$ and thus $\Delta_{\tilde{\tau}} - \Delta_h = \tilde{\tau}\Delta_h\Delta_{\tilde{\tau}}$ so that

$$\|\Delta_{\tilde{\tau}} e^k - \Delta_h e^k\|_{L^2} = \|\tilde{\tau}\Delta_h\Delta_{\tilde{\tau}} e^k\|_{L^2} \leq C \frac{\tilde{\tau}}{h^2} \|\Delta_{\tilde{\tau}} e^k\|_{L^2} . \tag{30}$$

Here we have again used that the spectrum of $-\Delta_h$ is bounded by $C h^{-2}$. Now testing (25) with $-\Delta_h e^k$ we obtain

$$-(\partial_\tau e^k, \Delta_h e^k)_{L^2} - (\Delta_{\tilde{\tau}}^2 e^k, \Delta_h e^k)_{L^2} = -(\omega^k, \Delta_h e^k)_{L^2} \leq \frac{1}{2}\|\omega^k\|_{L^2}^2 + \frac{1}{2}\|\Delta_h e^k\|_{L^2}^2 .$$

Furthermore, we can estimate $-(\partial_\tau e^k, \Delta_h e^k)_{L^2} = (\partial_\tau \nabla e^k, \nabla e^k)_{L^2} \geq \frac{1}{2}\partial_\tau\|\nabla e^k\|_{L^2}^2$ and $-(\Delta_{\tilde{\tau}}^2 e^k, \Delta_h e^k)_{L^2} = -(\Delta_{\tilde{\tau}} e^k, \Delta_h \Delta_{\tilde{\tau}} e^k)_{L^2} = \|\nabla\Delta_{\tilde{\tau}} e^k\|_{L^2}^2 \geq 0$ so that

$$\partial_\tau \|\nabla e^k\|_{L^2}^2 \leq \|\omega^k\|_{L^2}^2 + 2\|\Delta_{\tilde{\tau}} e^k\|_{L^2}^2 + 2\|\Delta_{\tilde{\tau}} e^k - \Delta_h e^k\|_{L^2}^2 \tag{31}$$
$$\leq C(\tau + \tilde{\tau})^2 + 2C t_k (1 + \tfrac{\tilde{\tau}}{h^2})(\tau + \tilde{\tau})^2$$

By induction, we finally achieve $\|\nabla e^k\|_{L^2} \leq C(1 + \frac{\sqrt{\tilde{\tau}}}{h})(\tau + \tilde{\tau})$. Combined with (27) and (22) this concludes the proof. Revisiting all constants, we see that they scale quadratically in time. However, as mentioned before, with nonlinear terms in the PDE one would expect an exponential growth. □

Note that the above proof would yield the same scaling of the L^2-error if the initial data were $U^0 = \hat{R}_h(\tilde{u}^0)$ and thus only $\|\omega^k\|_{L^2} \leq C t_k^{-\frac{1}{2}}(\tau + \tilde{\tau})$ in (24). The summation in (27) would then yield a scaling $C\sqrt{t_k}(\tau + \tilde{\tau})$.

k	$\tau = h^2, \tilde{\tau} = h^2$		$\tau = h^2, \tilde{\tau} = h$		$\tau = h, \tilde{\tau} = h^2$		$\tau = h, \tilde{\tau} = h$	
	L^2	H^1	L^2	H^1	L^2	H^1	L^2	H^1
6	2.032	1.020	0.886	0.959	1.417	1.113	1.085	1.093
7	2.009	1.005	0.919	0.969	1.259	1.048	1.045	1.050
8	2.002	1.001	0.953	0.979	1.148	1.021	1.023	1.026

Table 1: Experimental order of convergence for the numerical solution of the biharmonic heat equation in the L^2- and H^1-norm $((\ln 2)^{-1}(\ln(e_k(t)) - \ln(e_{k+1}(t))))$ for different choices of τ and $\tilde{\tau}$ and for $h = 2^{-k}$ (for $t = 0.0024$ in the first two cases and $t = 0.1562$ else).

4.2 Numerical validation

To verify the above convergence result experimentally we have solved $\partial_t \tilde{u} + \Delta^2 \tilde{u} = \tilde{f}$ in $\Omega = [0,1]^2$ with Dirichlet boundary conditions $\tilde{u} = 0$ and $\Delta \tilde{u} = 0$ on $\partial\Omega$. Furthermore, we consider a particular right hand side $\tilde{f}(t,x,y) := 20\, e^{20t}(x^4 - 2x^3 + x)\,(y^4 - 2y^3 + y) + 24\,(e^{20t} - 1)\,(y^4 - 2y^3 + y + (x^2 - x)(12y^2 - 12y) + x^4 - 2x^3 + x)$ for which $\tilde{u}(t,x,y) = (e^{20t} - 1)\,(x^4 - 2x^3 + x)\,(y^4 - 2y^3 + y)$ is the solution.

With a nonzero right hand side, (14) in each time step has to be replaced by the minimization of the energy $w_f^l[U^k, U, V] := \mathcal{E}_{\text{out}}^l[U^k, U, V] - 2\tau \int_\Omega \mathcal{I}_h(\tilde{f}U)\,\mathrm{d}x$ under the constraint (14b), which in analogy to (10) is achieved by finding the saddle of the Lagrangian

$$\mathcal{L}^l[U^k, U, V, P] = \mathbf{M}\left(\overline{U} - \overline{U}^k\right) \cdot \left(\overline{U} - \overline{U}^k\right) + \frac{\tau}{\tilde{\tau}^2}\mathbf{M}\left(\overline{V} - \overline{U}\right) \cdot \left(\overline{V} - \overline{U}\right)$$
$$- 2\tau \mathbf{M}\overline{\mathcal{I}_h(\tilde{f})} \cdot \overline{U} + 2\mathbf{M}\left(\overline{V} - \overline{U}\right) \cdot \overline{P} + 2\tilde{\tau}\mathbf{L}\overline{V} \cdot \overline{P}.$$

We consider triangular meshes which are formed by a subdivision of all squares into two triangles in a uniform quadrilateral grid. We employ the grid sizes $h = 2^{-5}, 2^{-6}$, and 2^{-7} and different choices for τ and $\tilde{\tau}$. Error plots are given in Figure 1, and the experimental error order is computed in Table 1. We show the L^2-error for $\tau = \tilde{\tau} = h^2$ as well as for $\tau = h^2$, $\tilde{\tau} = h$, and the H^1-error for $\tau = h$, $\tilde{\tau} = h^2$ and $\tau = \tilde{\tau} = h$. The above numerical analysis predicts an error decay by a factor $\frac{1}{4}$ between two subsequent grids for the first case, a decay by $\frac{1}{\sqrt{2}}$ for the last case, and a decay by $\frac{1}{2}$ for the remaining cases. The decay rates for the L^2-error are confirmed by the numerical experiments, while the convergence rates with respect to the H^1-norm are higher than predicted. Indeed, the factor $\frac{\sqrt{\tilde{\tau}}}{h}$ which results from estimating the difference between the two Laplace operator approximations $\Delta_{\tilde{\tau}}$ and Δ_h in equation (30) via discrete spectral estimates seems not to be needed. Note however that $\tilde{\tau}$ may be chosen small anyway (thereby compensating a factor $\frac{\sqrt{\tilde{\tau}}}{h}$) without compromising the speed of the time stepping scheme for which only the outer time step τ is relevant.

5 Numerical results for Willmore flow

In this section we discuss numerical results for Willmore flow computed via the proposed nested time stepping scheme. In addition we compare our discretization with the one proposed by Du et al. [22, 21].
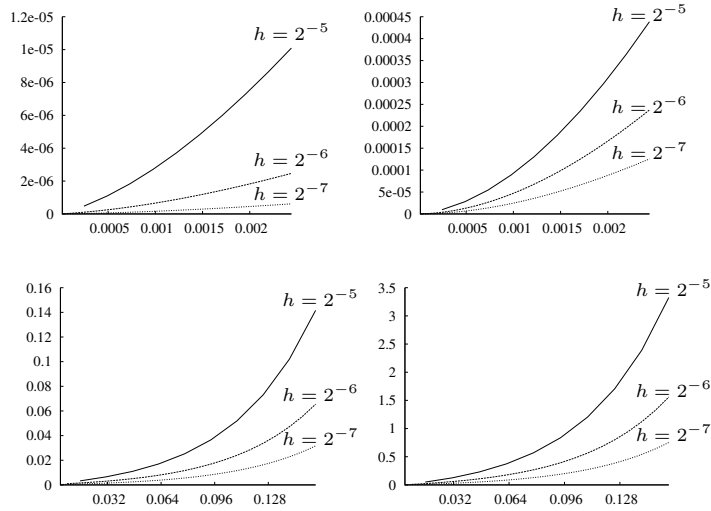
Figure 1: Temporal error evolution for the numerical solution of the biharmonic heat equation. Upper row: L^2-error for $\tau = h^2$ and $\tilde{\tau} = h^2$ (left) as well as $\tilde{\tau} = h$ (right). Lower row: H^1-error for $\tau = h$ and $\tilde{\tau} = h^2$ (left) as well as $\tilde{\tau} = h$ (right).

5.1 Simulation based on the nested time discretization.

At first we consider the evolution of a circle in $\mathbb{R}^2$ under Willmore flow. For an initial radius r_0 the flow results in a family of concentric circles $\Gamma_{r(t)} \subset \mathbb{R}^2$ whose radius $r(t)$ increases in time according to

$$r(t) = \sqrt[4]{2t + r_0^4} \, .$$

At each time step k we compute the L^2-error $\|e^k\|_{L^2} = \|\tilde{u}_{k\tau} - U^k\|_{L^2}$ with U^k the numerical solution and $\tilde{u}_{k\tau}(\cdot) = \tanh(\text{sgndist}(\cdot, \Gamma_{r(k\tau)})/\varepsilon)$ having the expected phase field profile associated with the exact solution. The computations are performed for a quarter circle centered around one corner of the computational domain $\Omega = [0,1]^2$ and for initial radius $r_0 = 0.4$ and diffusive interface thickness $\varepsilon = 2^{-6}$. The error evolution in time is shown in Figure 2 for $\tau = \tilde{\tau} = h^2$ and grid sizes $h = 2^{-7}, 2^{-8}$ and 2^{-9}. The error oscillation in the case $h = 2^{-7}$ indicates that the interface thickness $\varepsilon \sim 2h$ on this grid is still too small. From the analysis of the linear model problem in Section 4 we expect an error decay by a factor $\frac{1}{4}$ when halving the grid size, which is indeed reflected by our numerical results also for Willmore flow.

As a further test simulation, let us consider Willmore flow of two exemplary geometries, one involving corners (cf. Figure 3) and the other involving a topological change in the phase field context starting from two neighboring circles (cf. Figure 4). In both cases we take $\Omega = [0,1]^2$, $h = 2^{-9}$, $\varepsilon = 4h$, and $\tau = \tilde{\tau} = h^2$. The evolution of the rectangle is initially particularly pronounced at the corners leading to a locally concave object, which then gets convex again, and finally evolves to a circle. In the other application the two circles merge into one locally concave object and then evolve similarly as in the first case. In both Figure 3 and 4, the curvature approximation $(U^k - V^k)/\tilde{\tau}$ of the evolving geometries is also displayed.

A three-dimensional simulation of Willmore flow is provided in Figure 5, showing the evolution of a cube and a flat disk under Willmore flow. As in the two-dimensional
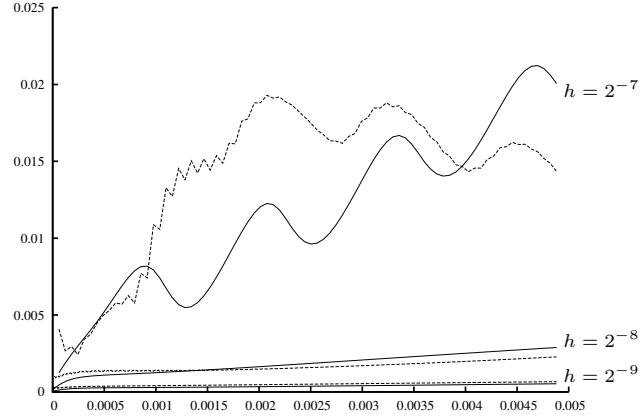
Figure 2: Temporal evolution of the L^2-error for Willmore flow of a quarter circle with initial radius $r_0 = 0.4$, using different grid sizes h and $\varepsilon = 2^{-6}$, $\tau = \tilde{\tau} = h^2$. The solid lines show the error of our scheme, the dotted lines belong to the scheme proposed by Du et al. (cf. Section 5.2).
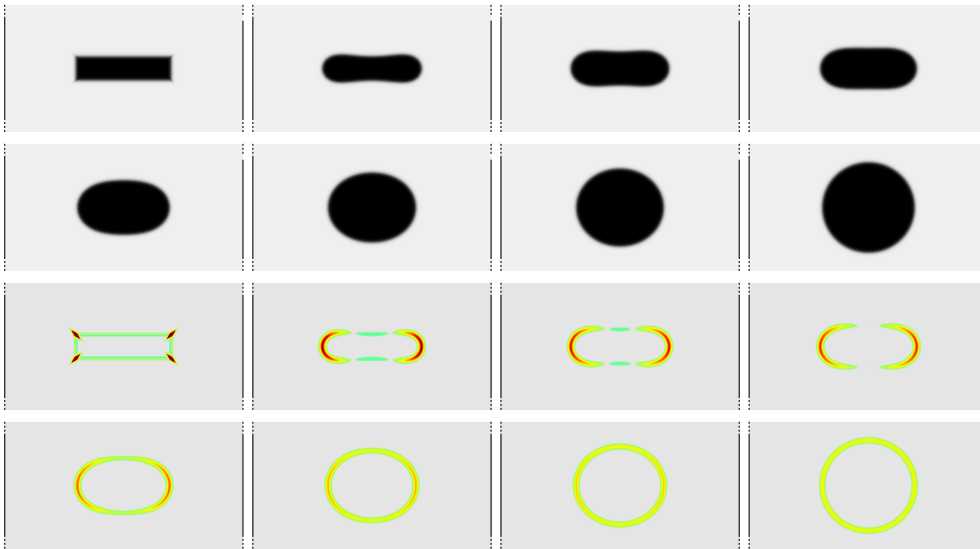

Figure 3: Computed evolution of a rectangle under Willmore flow for grid size $h = 2^{-9}$, $\varepsilon = 4h$, and $\tau = \tilde{\tau} = h^2$. Results are depicted at time steps $k = 0, 1, 6, 10, 20, 40, 60, 120$. The bottom rows show the underlying discrete curvature $(U^k - V^k)/\tilde{\tau}$, color-coding on the range $[-20, 20]$ as ▬▬▬▬▬.
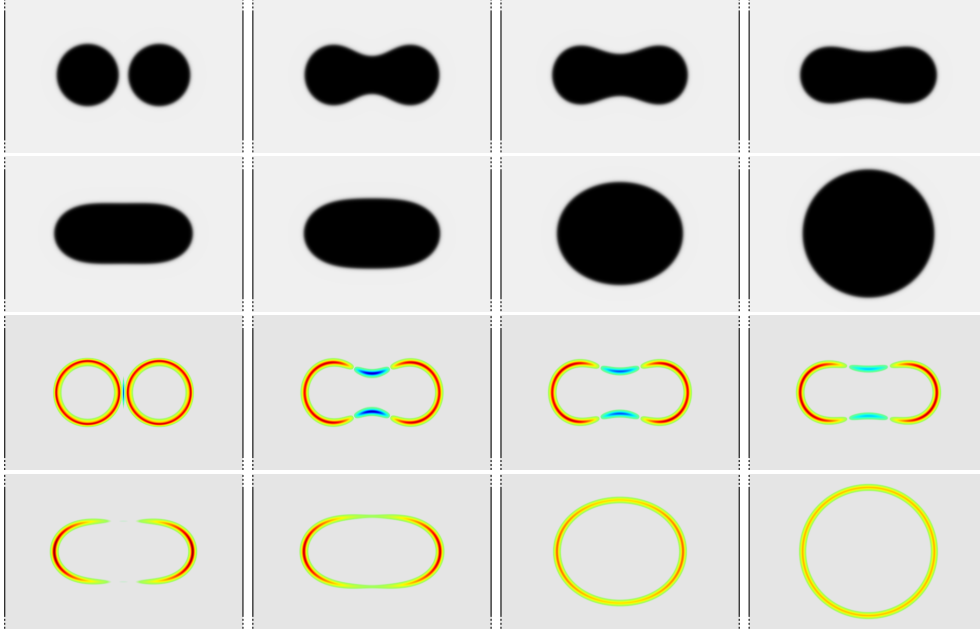
Figure 4: $h = 2^{-9}$, $\varepsilon = 4h$, and $\tau = \tilde{\tau} = h^2$, time steps $k = 0, 1, 2, 4, 20, 40, 160, 450$, discrete curvature $(U^k - V^k)/\tilde{\tau}$ color-coded on the range $[-10, 10]$ as ▆▆▆▆▆.

case in Figure 3, the flow in both examples is most pronounced in regions with high curvature, leading temporarily to concave shapes.

5.2 Comparison with a semi-implicit mixed scheme

In [22] and [21] Du et al. proposed a semi-implicit phase field scheme for Willmore flow, which we now will compare with the variational time discretization presented here. Following [14], the Willmore energy can be approximated as

$$w^\varepsilon[u] = \frac{1}{2\varepsilon} \int_\Omega \left(-\varepsilon \Delta u + \frac{1}{2\varepsilon} \Psi'(u) \right)^2 \mathrm{d}x . \tag{32}$$

Defining $f(u) := - \left(-\varepsilon \Delta u + \frac{1}{2\varepsilon} \Psi'(u) \right) = \varepsilon \Delta u + \frac{2}{\varepsilon} u(1 - u^2)$ and $g(u) := \Delta f(u) + \frac{2}{\varepsilon^2}(1 - 3u^2)f(u)$, its first variation reads

$$\partial_u w^\varepsilon[u](\vartheta) = \int_\Omega g(u)\vartheta \, \mathrm{d}x .$$

Du and Wang [22] used the variable splitting

$$g(u, v) := \frac{1}{2}\Delta \left(f(u) + f(v) \right) + \frac{1}{\varepsilon^2}(1 - u^2 - uv - v^2) \left(f(u) + f(v) \right)$$
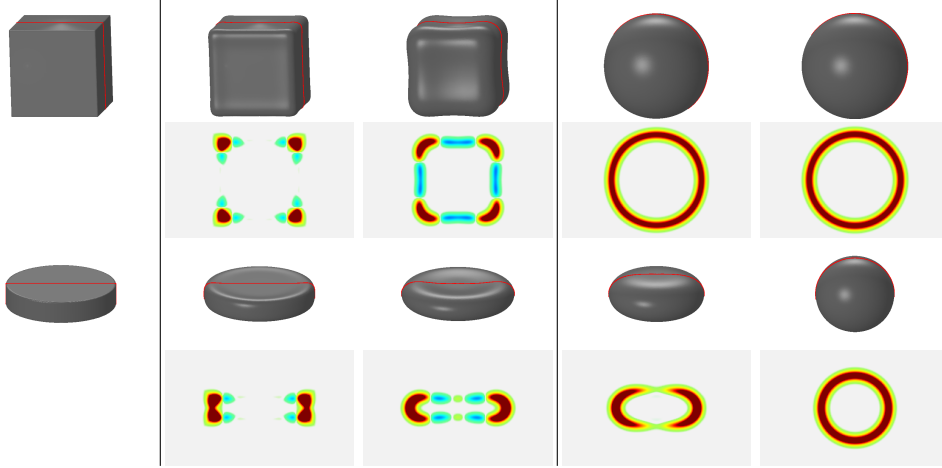
17

Figure 5: Different time steps of the proposed discrete Willmore flow for a cube with edge length 0.4 and a disk of diameter 0.5 and thickness 0.1 (left) as initial surfaces. The computational parameters are $h = 2^{-8}$, $\epsilon = 4h$, $\tilde{\tau} = h^2$. In the middle results are depicted at time steps $k = 1, 20$ with $\tau = h^3$ and on the right at time steps $k = 1, 10$ with $\tau = 0.03h$ in case of the cube and $\tau = h^2$ in case of the disc. In addition to the evolving shapes we render the underlying discrete curvature $(U^k - V^k)/\tilde{\tau}$ on a planar slice indicated in red on the geometries and use a color-coding on the range $[-5, 5]$ as �.

to derive a semi-implicit time discretization ansatz $\varepsilon \frac{u^{k+1}-u^k}{\tau} = -g(u^{k+1}, u^k)$ for Willmore flow, leading in weak form to the system of equations

$$\int_\Omega \varepsilon \frac{u^{k+1} - u^k}{\tau} \vartheta \, \mathrm{d}x = \int_\Omega \frac{1}{2} \nabla \left(f^{k+1} + f^k \right) \cdot \nabla \vartheta$$
$$- \frac{1}{\varepsilon^2} \left(1 - \left(u^{k+1} \right)^2 - u^{k+1} u^k - \left(u^k \right)^2 \right) \left(f^{k+1} + f^k \right) \vartheta \, \mathrm{d}x$$
$$\int_\Omega f^{k+1} \vartheta \, \mathrm{d}x = \int_\Omega -\varepsilon \nabla u^{k+1} \cdot \nabla \vartheta + \frac{2}{\varepsilon} u^{k+1} \left(1 - (u^{k+1})^2 \right) \vartheta \, \mathrm{d}x$$

in u^{k+1} and f^{k+1} for given u^k and $f^k \equiv f(u^k)$. This time-discrete scheme can now be discretized in space using piecewise affine finite elements for the approximation of u^{k+1}, u^k, f^{k+1}, and f^k. To solve this nonlinear system, we apply the same Newton method as in Section 3.2 for our model. Note that the actual model by Du and Wang is more involved due to a volume and a surface constraint, and we here only present a reduced version for pure Willmore flow.

For a direct comparison of both schemes, we simulate Willmore flow of a quarter circle with initial radius $r_0 = 0.4$. Figure 2 compares the respective L^2-errors. In our numerical experiments the new scheme turned out to be significantly more robust. For example, for a grid size $h = 2^{-7}$, the Newton method for the nested time discretization needs about 4 iterations per time step, whereas up to 200 iterations are required in the semi-implicit scheme, resulting in significantly longer computation times. A detailed comparison of the required Newton iterations and the resulting computing times is given in Table 2 based on the simulation of 80 time steps on a single kernel

numerical scheme	$h = 2^{-7}$		$h = 2^{-8}$		$h = 2^{-9}$	
	iter	time [sec]	iter	time [sec]	iter	time [sec]
nested variational	288	1366	961	20651	2562	338602
semi-implicit mixed [22]	6176	18746	3695	121531	5776	1051560

Table 2: Total number of Newton iterations and computation times for the simulations from Figure 2.

of an Intel Xeon E5530 CPU. In our experiments, the nonlinear solver in the nested variational scheme converged stably for step sizes of up to $\tau = 0.03\,h$ whereas we observed convergence problems with the nonlinear solver in the semi-implicit scheme for step sizes of this order.

6 Application to a image restoration

As outlined above the proposed nested variational time discretization allows for large time steps. Thus, it is in particular useful in the context of geometric variational problems where the focus is more on a robust numerical descent scheme than on a high precision time discretization. Here, we will investigate its application to the problem of image restoration.

Picking up the edge restoration approach by Nitzberg et al. [40] we aim at finding a continuation of a given image edge inside a region in which the underlying image is corrupted. In fact, we ask for an edge contour which minimizes the weighted sum of the Willmore energy and the area functional for presribed boundary data on the boundary of the reconstruction region. To this end, we slightly modify the energy (5a) from the outer variational problem,

$$e_{\text{out}}^{\varepsilon,\eta}[u^k, u, v] = e_{\text{out}}^{\varepsilon}[u^k, u, v] + 2\tau\eta a^{\varepsilon}[u]\,, \tag{33}$$

where $a^{\varepsilon}[u]$ is the above-defined phase field approximation of the surface area. Now the constrained minimization amounts to finding a saddle of the Lagrangian

$$\ell[u^k, u, v, p] = e_{\text{out}}^{\varepsilon,\eta}[u^k, u, v] + \partial_v e_{\text{in}}^{\varepsilon}[u, v](p)$$

$$= \int_{\Omega} \varepsilon(u - u^k)^2 \, \mathrm{d}x + \tau\varepsilon \left(\frac{v - u}{\tilde{\tau}}\right)^2 \mathrm{d}x + \tau\eta \int_{\Omega} \varepsilon|\nabla u|^2 + \frac{1}{\varepsilon}\Psi(u) \, \mathrm{d}x$$

$$+ \int_{\Omega} 2\varepsilon(v - u)p + \frac{\tilde{\tau}}{\varepsilon}\Psi'(v)p + 2\tilde{\tau}\varepsilon\nabla v \cdot \nabla p \, \mathrm{d}x$$

using Newton's methods for the associated finite element discretization. In Figure 6 the reconstruction region is outlined in red. The image and the corresponding phase field representation u of the edges outside this region is considered as fixed. To incorporate the desired C^1 boundary condition on the boundary of the reconstruction region, we solve for v in the inner variational problem on an ε-neighborhood of the reconstruction region. Satisfactory reconstruction results are already achieved after a few time steps of the gradient flow associated to the above energy.

Finally, in Figure 7 we vary the parameter η (all other parameters are kept as before). The results illustrate a different balance between Willmore energy and surface area in the minimization.
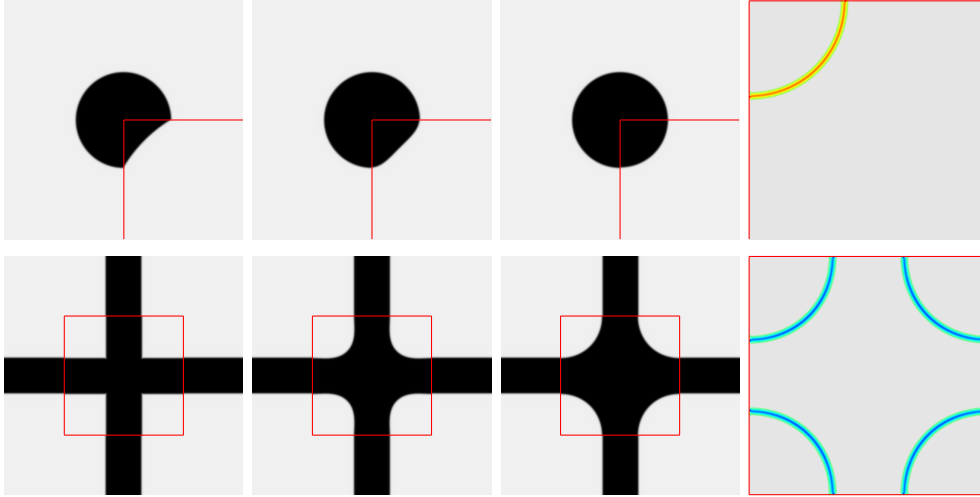
19

Figure 6: Time steps 0, 1, and 30 from two different edge restoration problems. The restoration region is a unit square, outlined in red and resolved by a $2^9 \times 2^9$ regular grid. Parameters are $\varepsilon = 4h$, $\eta = 1$, $\tau = 0.03h$, and $\tilde{\tau} = h^2$. In both rows the rightmost image shows a blowup of the restoration region at the last time step with the curvature color-coded as ████████████ on the range $[-5, 5]$.
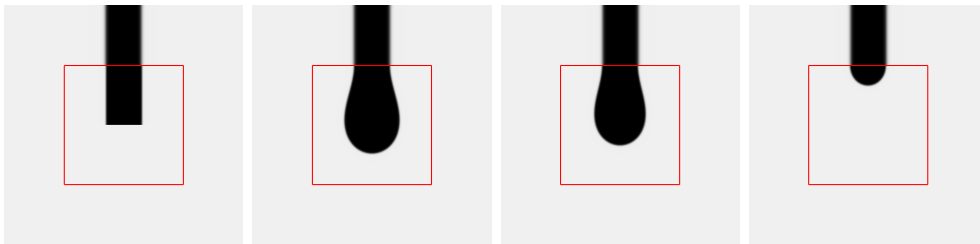

Figure 7: A variational C^1 edge continuation is computed via a gradient flow minimization of the sum of the Willmore energy and a weighted surface area starting from an initial phase field (left). The minimizing phase field is shown for the surface weight $\eta = 0.1$ (second), $\eta = 1$ (third), and $\eta = 10$ (fourth image).

Acknowledgments

References

[1] Fred Almgren, Jean E. Taylor, and Lihe Wang. Curvature-driven flows: a variational approach. *SIAM J. Control Optim.*, 31(2):387–438, 1993.

[2] J. W. Barrett, H. Garcke, and R. Nürnberg. A parametric finite element method for fourth order geometric evolution equations. *J. Comp. Phys.*, 222:441–467, 2007.

[3] G. Bellettini and L. Mugnai. On the approximation of the elastica functional in radial symmetry. *Calculus of Variations and Partial Differential Equations*, 24(1):1–20, 2005.

[4] M. Bertalmio, A.L. Bertozzi, and G. Sapiro. Navier-stokes, fluid dynamics, and image and video inpainting. In *IEEE Proceedings of the International Conference on Computer Vision and Pattern Recognition*, volume 1, pages 355–362, 2001.

[5] M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester. Image inpainting. In *Proc. of SIGGRAPH 2000*, pages 417–424, New Orleans, USA, July 2000.

[6] I. A. Bobenko and P. Schröder. Discrete Willmore flow. pages 101–110. ACM Press, 2005.

[7] Antonin Chambolle. An algorithm for mean curvature motion. *Interfaces and free Boundaries*, 6:195–218, 2004.

[8] T. F. Chan, S. H. Kang, and J. Shen. Euler's elastica and curvature-based inpainting. *SIAM Appl. Math.*, 63(2):564–592, 2002.

[9] Xinfu Chen, Charlie M. Elliott, Andy Gardiner, and Jennifer Jing Zhao. Convergence of numerical solutions to the Allen-Cahn equation. *Appl. Anal.*, 69(1-2):47–56, 1998.

[10] Y. Chen, T. A. Davis, W. W. Hager, and S. Rajamanickam. Algorithm 887: CHOLMOD, supernodal sparse Cholesky factorization and update/downdate. *ACM Transactions on Mathematical Software*, 35(3):22:1–22:14, 2009.

[11] U. Clarenz, U. Diewald, G. Dziuk, M. Rumpf, and R. Rusu. A finite element method for surface restoration with smooth boundary conditions. *Computer Aided Geometric Design*, 21(5):427–445, 2004.

[12] Ph. Clement. Approximation by finite element functions using local regularization. *RAIRO Anal. Numer.*, 9:77–84, 1975.

[13] T. A. Davis and W. W. Hager. Dynamic supernodes in sparse Cholesky update/downdate and triangular solves. *ACM Transactions on Mathematical Software*, 35(4):27:1–27:23, 2009.

[14] Ennio De Giorgi. Some remarks on Γ-convergence and least squares method. In *Composite media and homogenization theory (Trieste, 1990)*, volume 5 of *Progr. Nonlinear Differential Equations Appl.Progr. Nonlinear Differential Equations Appl.*, pages 135–142. Birkhäuser Boston, Boston, MA, 1991.

[15] K. Deckelnick and G. Dziuk. Error analysis of a finite element method for the Willmore flow of graphs. *Interfaces and Free Boundaries*, 8:21–46, 2006.

[16] K. Deckelnick, G. Dziuk, and C. M. Elliott. Computation of geometric partial differential equations and mean curvature flow. *Acta Numerica*, 14:139–232, 2005.

[17] K. Deckelnick and F. Schieweck. Error analysis for the approximation of axisymmetric willmore flow by $c1$-elements. Technical Report 23, Universität Magdeburg, 2009. to appear in Interfaces and Free Boundaries.

[18] Klaus Deckelnick and Gerhard Dziuk. Error analysis for the elastic flow of parametrized curves. *Math. Comp.*, 78(266):645–671, 2009.

[19] Patrick W. Dondl, Luca Mugnai, and Matthias Röger. Confined elastic curves. arXiv:submit/0044093 [math.AP].

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

[21] Qiang Du, Chun Liu, and Xiaoqiang Wang. Simulating the deformation of vesicle membranes under elastic bending energy in three dimensions. *Journal of Computational Physics*, 212(2):757–777, 2006.

[22] Qiang Du and Xiaoqiang Wang. Convergence of numerical approximations to a phase field bending elasticity model of membrane deformations. *International Journal of Numerical Analysis and Modeling*, 4(3–4):441–459, 2007.

[23] Qiang et. al. Du. A phase field formulation of the willmore problem. *Nonlinearity*, 18:1249–1267, 2005.

[24] G. Dziuk, E. Kuwert, and R. Schätzle. Evolution of elastic curves in $\mathbb{R}^n$: existence and computation. *SIAM J. Math. Anal.*, 33, no. 5(5):1228–1245 (electronic), 2002.

[25] Gerhard Dziuk. Computational parametric Willmore flow. *Numer. Math.*, 111(1):55–80, 2008.

[26] C. M. Elliott, D. A. French, and F. A. Milner. A second order splitting method for the Cahn-Hilliard equation. *Numer. Math.*, 54:575–590, 1989.

[27] Charles M. Elliott and Stig Larsson. Error estimates with smooth and nonsmooth data for a finite element method for the Cahn-Hilliard equation. *Math. Comp.*, 58(198):603–630, S33–S36, 1992.

[28] L. C. Evans, H. M. Soner, and P. E. Souganidis. Phase transitions and generalized motion by mean curvature. *Communications on Pure and Applied Mathematics*, 45:1097–1123, 2006.

[29] W. Hackbusch. *Theorie und Numerik elliptischer Differentialgleichungen*. Teubner Studienbücher Mathematik. Teubner, 1986.

[30] W. Helfrich. Elastic properties of lipid bilayers: Theory and possible experiments. *Zeitschrift für Naturforschung*, 28c:693–703, 1973.

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

[32] E. Kuwert and R. Schätzle. Gradient flow for the Willmore functional. *Comm. Anal. Geom.*, 10(5):1228–1245 (electronic), 2002.

[33] E. Kuwert and R. Schätzle. Removability of Point Singularities of Willmore Surfaces. Preprint SFB 611, Bonn, 2002.

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

[35] P. Loreti and R. March. Propagation of fronts in a nonlinear fourth order equation. *European Journal of Applied Mathematics*, 11(2):203–213, 2000.

[36] S. Luckhaus and Th. Sturzenhecker. Implicit time discretization for the mean curvature flow equation. *Calc. Var.*, 3:253–271, 1995.

[37] Luciano Modica. The gradient theory of phase transitions and the minimal interface criterion. *Archive for Rational Mechanics and Analysis*, 98(2):123–142, 1987.

[38] Luciano Modica and Stefano Mortola. Un esempio di Γ^--convergenza. *Boll. Un. Mat. Ital. B (5)*, 14(1):285–299, 1977.

[39] D. Mumford. Elastica and computer vision. In C. Bajaj, editor, *Algebraic Geometry and Its Applications*, pages 491–506. Springer, New York, 1994.

[40] M Nitzberg, D. Mumford, and T. Shiota. *Filtering, Segmentation and Depth (Lecture Notes in Computer Science Vol. 662)*. Springer-Verlag Berlin Heidelberg, 1993.

[41] R. H. Nochetto, M. Paolini, and C. Verdi. Optimal interface error estimates for the mean curvature flow. *Annali della Scuola Normale Superiore di Pisa. Classe di Scienze. Serie IV*, 21(2):193–212, 1994.

[42] R. H. Nochetto, M. Paolini, and C. Verdi. Quadratic rate of convergence for curvature dependent smooth interfaces: a simple proof. *Applied Mathematics Letters. An International Journal of Rapid Publication*, 7(4):59–63, 1994.

[43] Nadine Olischläger and Martin Rumpf. A nested variational time discretization for parametric willmore flow. *Interfaces and Free Boundaries*. submitted.

[44] Nadine Olischläger and Martin Rumpf. Two step time discretization of Willmore flow. 2009. accepted at IMA Conference on the Mathematics of Surfaces.

[45] A. Polden. Closed Curves of Least Total Curvature. *SFB 382 Tübingen, Preprint*, 13:, 1995.

[46] A. Polden. Curves and Surfaces of Least Total Curvature and Fourth-Order Flows. *Dissertation, Universität Tübingen*, page , 1996.

[47] S. D. Rane, J. Remus, and G. Sapiro. Wavelet-domain reconstruction of lost blocks in wireless image transmission and packet-switched networks. In *Image Processing. 2002. Proceedings. 2002 International Conference on 22-25 Sept. 2002, Vol.1*, 2002.

[48] Tristan Rivière. Analysis aspects of Willmore surfaces. *Invent. Math.*, 174(1):1–45, 2008.

[49] Matthias Röger and Reiner Schätzle. On a modified conjecture of De Giorgi. *Mathematische Zeitschrift*, 254(4):675–714, December 2006.

[50] R. Rusu. An algorithm for the elastic flow of surfaces. *Preprint Mathematische Fakultät Freiburg*, 01-35:, 2001.

[51] U. Seifert. Configurations of fluid membranes and vesicles. *Advances in Physics*, 46:13–137, 1997.

[52] G. Simonett. The Willmore Flow near spheres. *Diff. and Integral Eq.*, 14(8):1005–1014, 2001.

[53] Vidar Thomée. *Galerkin Finite Element Methods for Parabolic Problems*, volume 25 of *Springer Series in Computational Mathematics*. Springer, Berlin, 2nd edition, 2006.

[54] William Welch and Andrew Witkin. Variational surface modeling. *Computer Graphics*, 26(2):157–166, 1992.

[55] T.J. Willmore. *Riemannian Geometry*. Claredon Press, Oxford, 1993.

[56] Guoliang Xu and Qing Pan. G^1 surface modelling using fourth order geometric flows. *Computer-Aided Design*, 38(4):392–403, 2006.

[57] Shin Yoshizawa and Alexander G. Belyaev. Fair triangle mesh generation with discrete elastica. In *Proceedings of the Geometric Modeling and Processing; Theory and Applications (GMP'02)*, pages 119–123, Washington, DC, USA, 2002. IEEE Computer Society.