A convergent evolving finite element algorithm for Willmore flow of closed surfaces

**Balázs Kovács · Buyang Li ·
Christian Lubich**

Abstract A proof of convergence is given for a novel evolving surface finite element semi-discretization of Willmore flow of closed two-dimensional surfaces, and also of surface diffusion flow. The numerical method proposed and studied here discretizes fourth-order evolution equations for the normal vector and mean curvature, reformulated as a system of second-order equations, and uses these evolving geometric quantities in the velocity law interpolated to the finite element space. This numerical method admits a convergence analysis in the case of continuous finite elements of polynomial degree at least two. The error analysis combines stability estimates and consistency estimates to yield optimal-order H^1-norm error bounds for the computed surface position, velocity, normal vector and mean curvature. The stability analysis is based on the matrix–vector formulation of the finite element method and does not use geometric arguments. The geometry enters only into the consistency estimates. Numerical experiments illustrate and complement the theoretical results.

Keywords Willmore flow · surface diffusion flow · geometric evolution equations · evolving surface finite elements · stability · convergence analysis

Mathematics Subject Classification (2000) 35R01 · 65M60 · 65M15 · 65M12

B. Kovács
Mathematisches Institut, Universität Tübingen, Auf der Morgenstelle 10, 72076 Tübingen, Germany;
Fakultät für Mathematik, Universität Regensburg, 93040 Regensburg, Germany,
E-mail: balazs.kovacs@mathematik.uni-regensburg.de

Ch. Lubich
Mathematisches Institut, Universität Tübingen,
Auf der Morgenstelle 10, 72076 Tübingen, Germany
E-mail: lubich@na.uni-tuebingen.de

B. Li
Department of Applied Mathematics, Hong Kong Polytechnic University,
Kowloon, Hong Kong
E-mail: buyang.li@polyu.edu.hk

1 Introduction

The elastic bending energy or Willmore energy (named after Thomas Willmore, see [Wil65, Wil93]) of a surface Γ is given as

$$W(\Gamma) = \frac{1}{2} \int_\Gamma H^2,$$

where H is the mean curvature of the surface. *Willmore flow* is the L^2 gradient flow of surfaces for the elastic bending energy. It plays an important role in modelling lipid bilayers [Hel73], biomembranes [ES10], vesicles [BGN15], regularization of phase-field systems [CLS$^+$18], and in the analysis of curvature on surfaces; see [MN14] proving the Willmore conjecture.

The negative L^2 gradient of the Willmore energy W for a two-dimensional surface Γ in $\mathbb{R}^3$ has no tangential contribution and its normal component equals

$$V = \Delta_\Gamma H + H(\tfrac{1}{2}H^2 - 2K) \quad \text{on } \Gamma, \tag{1.1}$$

with mean curvature $H = \kappa_1 + \kappa_2$ (here taken without a factor $1/2$) and Gaussian curvature $K = \kappa_1\kappa_2$, where κ_1 and κ_2 are the principal curvatures on Γ. Willmore [Wil93] gives a proof of this result and attributes the formula to Thomsen [Tho24] (who mentions Schadow in 1922) and Blaschke [Bla29].

For Willmore flow, V of (1.1) is taken as the normal velocity of the evolving surface. This yields a fourth-order geometric evolution equation.

Numerical methods for Willmore flow (1.1) have been proposed in many articles. Algorithms based on evolving surface finite element methods for Willmore flow were studied by Rusu [Rus05], Dziuk [Dzi08] and Barrett, Garcke & Nürnberg [BGN07a, BGN08] based on different variational formulations of (1.1). Bonito, Nochetto & Pauletti [BNP10] considered surface finite elements for biomembranes described by Willmore flow under area and volume constraints (the Helfrich model). Pozzi [Poz15] studied a numerical method for anisotropic Willmore flow. Numerical simulations indicate that these methods apparently converge in practical computations — but to our knowledge, no proof of convergence has so far been obtained for any numerical method for the Willmore flow of closed surfaces.

For the Willmore flow of *curves*, Bartels [Bar13] and Deckelnick & Dziuk [DD09] proved convergence of finite element semi-discretizations. Pozzi & Stinner [PS18] proved convergence of a semi-discrete surface finite element method for elastic flow of curves coupled to a lateral diffusion process. For the Willmore flow of *graphs*, Deckelnick & Dziuk proved convergence of a (non-evolving) C^0 mixed finite element method in [DD06], and a convergence result for C^1 finite elements was given in [DKS15]. However, convergence of a surface finite element method for the Willmore flow of *closed surfaces* has remained an open problem.

Closely related to Willmore flow is the *surface diffusion flow*, which is the H^{-1} gradient flow of the area functional $|\Gamma| = \text{area}(\Gamma)$. The normal velocity

of the evolving surface then becomes

$$V = \Delta_\Gamma H. \tag{1.2}$$

The evolving surface described by (1.2) is the limit of the zero level set of the solution to the Cahn–Hilliard equation with a concentration dependent mobility (when the parameter representing the thickness of phase transition zone tends to zero), see [CENC96]. This model is often used to describe the diffusion-driven motion of the surface of a crystal, the motion of an interface in an alloy, and dewetting of thin solid films deposited on substrates; see [Mul57, BJSW17]. Since the equations of surface diffusion flow (1.2) and Willmore flow (1.1) have similar structure, they can often be approximated by similar numerical methods based on variational formulations of the equations. Numerical approximation to the surface diffusion flow (1.2) has been studied by Bänsch, Morin & Nochetto [BMN05] by a surface finite element method with mesh regularization, and by Barrett, Garcke & Nürnberg [BGN07a] based on a similar variational formulation as for their Willmore flow algorithm. Recently, Bao et al. [BJSW17,BJWZ17] and Zhao et al. [ZJB20] proposed finite element methods for the anisotropic surface diffusion flow with contact line migration in studying the evolution of solid-thin films on a substrate. A surface-decreasing and volume-conserving algorithm was recently proposed in [BZ21], based on the weak formulation of [BGN07a]. A perimeter-decreasing and area-conserving algorithm for *curve diffusion* was recently given in [JL21].

Similar to the case of Willmore flow, convergence of finite element semi-discretizations for the surface diffusion flow of *graphs* and axially symmetric surfaces has been proved; see [BMN04,DDE03,DDE05b]. However, convergence of a surface finite element algorithm for the surface diffusion flow of closed surfaces is still an open problem.

The objective of this article is to construct an evolving surface finite element algorithm that can be proved to be convergent for the Willmore flow of closed two-dimensional surfaces in three-dimensional space, and similarly for the surface diffusion flow.

The key idea is to use fourth-order parabolic evolution equations for the mean curvature H and the normal vector ν along the Willmore flow. We derive an algorithm based on the system that couples these evolution equations to the velocity law (1.1). Here, H and ν are considered to be independently evolving unknowns that are not directly extracted from the surface at any given time. This is different from the previously mentioned approaches.

This approach is motivated by our recent work on mean curvature flow in [KLL19]. There we used the coupled system with the evolution equations for H and ν, which were derived by Huisken [Hui84] along with evolution equations for other geometric quantities, to construct the first provably convergent finite element algorithm for mean curvature flow of closed surfaces. For Willmore flow, the corresponding evolution equations of H and ν are derived in this paper. Their general structure as fourth-order quasilinear parabolic equations was already derived by Kuwert & Schätzle [KS01,KS02] for proving existence results for Willmore flow, but the precise form of the equations as needed for

computations was not given and it was not evident that the equations for H and ν form a closed system that does not involve further geometric quantities.

The velocity equation and the coupled fourth-order evolution equations for the normal vector and mean curvature are rewritten as a second-order differential–algebraic system. The geometric evolution equations have a particular *anti-symmetric* structure. Based on a mixed variational formulation, we use evolving surface finite elements (of degree at least 2) for the semi-discretization of the coupled system, which preserves the anti-symmetry of the second-order system. The velocity law is approximated simply by finite element interpolation, in contrast to enforcing it using a Ritz projection as in [KLL19] for mean curvature flow.

Optimal-order H^1-norm semi-discrete convergence estimates are shown for all variables X, v, ν, H by clearly separating the issues of stability and consistency. Convergence is shown towards sufficiently regular solutions of Willmore flow, which excludes the formation of singularities (within the considered time interval).

The main issue in the paper is proving *stability*, which is here understood as bounding the errors in terms of consistency defects and initial errors. For the velocity law, stability is shown using a stability bound for the interpolation of products of surface finite element functions. The main idea for the stability estimates for the second-order system for the geometric variables is to exploit the *anti-symmetric* structure of the semi-discrete error equations and combine it with *multiple energy estimates*, testing with both the errors and their time derivative. The structure of the energy estimates is sketched in Figure 5.2. The proof is performed in the matrix–vector formulation of the numerical method, and it uses technical lemmas relating different finite element surfaces that were shown in [KLLP17, Section 4] and [KLL19, Section 7.1]. A key step in the proof is to establish $W^{1,\infty}$-norm error bounds for all variables, which are obtained from the time-uniform H^1-norm error bounds using an inverse estimate. The stability analysis is completely independent of geometric arguments.

The *consistency* analysis, i.e. proving estimates for the defects (the residuals obtained upon inserting appropriate projections of the exact solution into the method) and their time derivatives is based on [KLLP17, Section 8] and [KLL19, Section 9], which in turn are based on geometric estimates in [Dem09, Dzi88, DE07, DE13b, Kov18]. Together with error bounds for the initial values, it uses geometric error estimates, interpolation and Ritz map error bounds. Due to a non-symmetric divergence term for one of the variables a modified Ritz map is used, similar to the one in [LM15].

The paper is organized as follows. Section 2, besides introducing basic notations and geometric concepts, is mainly devoted to deriving the evolution equations for the mean curvature and the normal vector along Willmore flow. The coupled second-order system and its weak formulation, which serves as the basis of the algorithm, is presented there. Section 3 describes the evolving surface finite element semi-discretization, the matrix–vector formulation, and the modified semi-discrete problem to correct the initial values of the algebraic variables. Section 4 states the main result of the paper: optimal-order

semi-discrete error bounds in the H^1-norm for the errors in positions, velocity, normal vector and mean curvature. In Section 5 we prove the stability result after presenting the required auxiliary results. Section 6 contains the consistency analysis: the bounds for the defects and their time derivatives, and estimates for the errors in the initial values. Section 7 presents numerical experiments illustrating and complementing our theoretical results.

2 Evolution equations for Willmore flow

2.1 Basic notions and notation

(The text of this preparatory subsection is taken verbatim from [KLL19, Section 2.1].) We consider the evolving two-dimensional closed surface $\Gamma(t) \subset \mathbb{R}^3$ as the image

$$\Gamma(t) = \{X(p,t) \,:\, p \in \Gamma^0\}$$

of a smooth mapping $X : \Gamma^0 \times [0,T] \to \mathbb{R}^3$ such that $X(\cdot,t)$ is an embedding for every t. Here, Γ^0 is a smooth closed initial surface, and $X(p,0) = p$. In view of the subsequent numerical discretization, it is convenient to think of $X(p,t)$ as the position at time t of a moving particle with label p, and of $\Gamma(t)$ as a collection of such particles. To indicate the dependence of the surface on X, we will write

$$\Gamma(t) = \Gamma[X(\cdot,t)], \quad \text{or briefly} \quad \Gamma[X]$$

when the time t is clear from the context. The *velocity* $v(x,t) \in \mathbb{R}^3$ at a point $x = X(p,t) \in \Gamma(t)$ equals

$$\partial_t X(p,t) = v(X(p,t),t). \tag{2.1}$$

For a known velocity field v, the position $X(p,t)$ at time t of the particle with label p is obtained by solving the ordinary differential equation (2.1) from 0 to t for a fixed p.

For a function $u(x,t)$ ($x \in \Gamma(t)$, $0 \le t \le T$) we denote the *material derivative* (with respect to the parametrization X) as

$$\partial^\bullet u(x,t) = \frac{\mathrm{d}}{\mathrm{d}t}\, u(X(p,t),t) \quad \text{for} \quad x = X(p,t).$$

For the following notions, see the review [DDE05a] or [Eck12, Appendix A] or any textbook on differential geometry. On any regular surface $\Gamma \subset \mathbb{R}^3$, we denote by $\nabla_\Gamma u : \Gamma \to \mathbb{R}^3$ the *tangential gradient* of a function $u : \Gamma \to \mathbb{R}$, and in the case of a vector-valued function $u = (u_1, u_2, u_3)^T : \Gamma \to \mathbb{R}^3$, we let $\nabla_\Gamma u = (\nabla_\Gamma u_1, \nabla_\Gamma u_2, \nabla_\Gamma u_3)$. We thus use the convention that the gradient of u has the gradient of the components as column vectors. We denote by $\nabla_\Gamma \cdot f$ the *surface divergence* of a vector field f on Γ, given by $\mathrm{tr}(\nabla_\Gamma f)$, [DE07, equation (2.7)], or by (2.13), and by $\Delta_\Gamma u = \nabla_\Gamma \cdot \nabla_\Gamma u$ the *Laplace–Beltrami operator* applied to $u : \Gamma \to \mathbb{R}$. In the case of a vector-valued function $u = (u_1, u_2, u_3)^T :$

$\Gamma \to \mathbb{R}^3$, we set componentwise $\Delta_\Gamma u = (\Delta_\Gamma u_1, \Delta_\Gamma u_2, \Delta_\Gamma u_3)^T$. (In the case of a tangential vector field u, this componentwise Laplace–Beltrami operator differs from the intrinsic definition of the Laplace–Beltrami operator on tangential vector fields.)

We denote the unit outer normal vector field to Γ by $\nu : \Gamma \to \mathbb{R}^3$. Its surface gradient contains the (extrinsic) curvature data of the surface Γ. At every $x \in \Gamma$, the matrix of the *extended Weingarten map*,

$$A(x) = \nabla_\Gamma \nu(x),$$

is a symmetric 3×3 matrix (see, e.g., [Wal15, Proposition 20]). Apart from the eigenvalue 0 with eigenvector $\nu(x)$, its other two eigenvalues are the principal curvatures κ_1 and κ_2 at the point x on the surface. They determine the fundamental quantities

$$H := \mathrm{tr}(A) = \kappa_1 + \kappa_2, \qquad |A|^2 = \kappa_1^2 + \kappa_2^2,$$

where $|A|$ denotes the Frobenius norm of the matrix A. Here, H is called the *mean curvature* (as in most of the literature, taken without a factor $1/2$).

2.2 Evolution equations for normal vector and mean curvature of a surface moving under Willmore flow

Willmore flow sets the normal velocity $V = v \cdot \nu$ of the surface $\Gamma[X]$ to V of (1.1), that is,

$$V = \Delta_{\Gamma[X]} H + Q \quad \text{with} \quad Q = -\tfrac{1}{2} H^3 + |A|^2 H. \tag{2.2}$$

(In the surface diffusion flow (1.2), we have $Q = 0$.) In this paper we choose the velocity (2.1) to have only a normal component, so that

$$v = V\nu. \tag{2.3}$$

The geometric quantities on the right-hand sides of (2.2)–(2.3) satisfy the following evolution equations.

Lemma 2.1 *For a regular surface $\Gamma(t) = \Gamma[X(\cdot, t)]$ moving under Willmore flow, the outer normal vector ν and the mean curvature H satisfy the following equations, with the extended Weingarten map $A = \nabla_\Gamma \nu$, and with Q of (2.2) (we omit the ubiquitous argument t),*

$$\partial^\bullet H = -\left(\Delta_\Gamma + |A|^2\right)(\Delta_\Gamma H + Q), \tag{2.4}$$

$$\partial^\bullet \nu = \left(-\Delta_\Gamma + (HA - A^2)\right)(\Delta_\Gamma \nu + |A|^2 \nu) + |\nabla_\Gamma H|^2 \nu \tag{2.5}$$
$$\quad - 2\left(\nabla_\Gamma \cdot (A\nabla_\Gamma H)\right)\nu - A^2 \nabla_\Gamma H - \nabla_\Gamma Q.$$

To our knowledge, the exact formulation of these evolution equations has not been reported in the literature before. In [KS02], the general structure of the evolution equations as fourth-order quasilinear parabolic equations is given but the lower-order terms are not stated explicitly and are therefore not available for computations.

A noteworthy feature of these evolution equations is that $\nabla_\Gamma A$ and the Hessian of H do not appear, but only $\nabla_\Gamma H$ and the term $\nabla_\Gamma \cdot (A\nabla_\Gamma H)$ in divergence form. This will allow us to work with the evolution equations for only H and ν in the numerical discretization, without the need for an evolution equation for A or further geometric quantities.

The following commutator formula will be needed to prove Lemma 2.1. Its proof is given in the Appendix.

Lemma 2.2 *For any regular surface Γ and for any regular function f on Γ the following commutator formula holds true:*

$$\Delta_\Gamma \nabla_\Gamma f - \nabla_\Gamma \Delta_\Gamma f$$
$$= (HA - A^2)\,\nabla_\Gamma f + (\Delta_\Gamma \nu \cdot \nabla_\Gamma f)\nu - 2\big(\nabla_\Gamma \cdot (A\nabla_\Gamma f)\big)\,\nu - A^2 \nabla_\Gamma f. \tag{2.6}$$

We remark that for the intrinsic Laplace–Beltrami operator applied to the tangential vector field $\nabla_\Gamma f$ a simpler commutator formula holds, with only the first term on the right-hand side. However, for our numerical purposes we need the componentwise Laplace–Beltrami operator as considered here. Since it can be shown that $HA - A^2 = K(I - \nu\nu^T)$, the first term on the right-hand side actually simplifies to $K\nabla_\Gamma f$. We prefer, however, to work with the first term as stated, since in the numerical method $\nabla_\Gamma f$ will be replaced by a vector field that is only approximately tangential.

Proof (of Lemma 2.1) For a general regular surface flow, the material derivatives of the outer normal vector and mean curvature satisfy, with $\Gamma = \Gamma[X(\cdot, t)]$ for short,

$$\partial^\bullet \nu = -\nabla_\Gamma V, \tag{2.7}$$

$$\partial^\bullet H = -\Delta_\Gamma V - |A|^2 V, \tag{2.8}$$

see [Man12, Eck12], or Lemma 2.37 (ii) and Lemma 2.39 (ii) in [BGN20][1].

The evolution equation (2.4) for H is thus obtained directly by inserting (2.2) into (2.8).

For the derivation of the evolution equation for ν we use (2.7) and (2.2), and obtain

$$\partial^\bullet \nu = -\nabla_\Gamma\big(\Delta_\Gamma H + Q\big).$$

We now use the commutator formula (2.6) with $f = H$, together with the following expression for the tangential gradient of mean curvature:

$$\nabla_\Gamma H = \Delta_\Gamma \nu + |A|^2 \nu, \tag{2.9}$$

---

[1] In [BGN20] the opposite sign convention for the normal vector is used.

see [Eck12, (A.9)] or [Wal15, Proposition 24]. This yields

$$
\begin{aligned}
\partial^{\bullet}\nu = {}& -\nabla_\Gamma \Delta_\Gamma H - \nabla_\Gamma Q \\
= {}& \big(-\Delta_\Gamma + (HA - A^2)\big)\nabla_\Gamma H + (\Delta_\Gamma \nu \cdot \nabla_\Gamma H)\nu \\
& - 2\big(\nabla_\Gamma \cdot (A\nabla_\Gamma H)\big)\nu - A^2\nabla_\Gamma H - \nabla_\Gamma Q \\
= {}& \big(-\Delta_\Gamma + (HA - A^2)\big)(\Delta_\Gamma \nu + |A|^2\nu) + ((\nabla_\Gamma H - |A|^2\nu)\cdot \nabla_\Gamma H)\nu \\
& - 2\big(\nabla_\Gamma \cdot (A\nabla_\Gamma H)\big)\nu - A^2\nabla_\Gamma H - \nabla_\Gamma Q,
\end{aligned}
$$

which becomes the stated evolution equation for the normal vector on noting that $\nu \cdot \nabla_\Gamma H = 0$.   □

2.3 The system of equations used for discretization

Collecting the above equations, we have reformulated Willmore flow as the system of non-linear fourth-order parabolic equations (2.4)–(2.5) on the surface coupled to the ordinary differential equations (2.1) for the surface positions via the velocity law (2.2)–(2.3).

By introducing two new variables V and z (we note that V is the normal velocity of the surface from (2.2)), the above system of fourth-order equations is reformulated as a system of formally second-order problems:

$$
\begin{aligned}
\partial^{\bullet}H = {}& -\Delta_{\Gamma[X]}V - |A|^2 V, & \text{(2.10a)} \\
V = {}& \Delta_{\Gamma[X]}H + Q; & \text{(2.10b)}
\end{aligned}
$$

$$
\begin{aligned}
\partial^{\bullet}\nu = {}& -\Delta_{\Gamma[X]}z + (HA - A^2)\,z + |\nabla_{\Gamma[X]}H|^2\nu - 2\big(\nabla_{\Gamma[X]}\cdot(A\nabla_{\Gamma[X]}H)\big)\nu \\
& - A^2\nabla_{\Gamma[X]}H - \nabla_{\Gamma[X]}Q, & \text{(2.10c)} \\
z = {}& \Delta_{\Gamma[X]}\nu + |A|^2\nu. & \text{(2.10d)}
\end{aligned}
$$

This system is coupled with the equations for position and velocity,

$$
\begin{aligned}
\partial_t X = {}& v \circ X, & \text{(2.11a)} \\
v = {}& V\nu. & \text{(2.11b)}
\end{aligned}
$$

Note that the normal velocity V appears twice: in the velocity equation (2.11b) and in the evolution equation for mean curvature (2.10a)–(2.10b).

The numerical discretization is based on a weak formulation of (2.10): we search for (H, V, ν, z) such that

$$\int_{\Gamma[X]} \partial^\bullet H \, \varphi^H - \int_{\Gamma[X]} \nabla_{\Gamma[X]} V \cdot \nabla_{\Gamma[X]} \varphi^H = - \int_{\Gamma[X]} |A|^2 \, V \, \varphi^H, \qquad (2.12a)$$

$$\int_{\Gamma[X]} V \varphi^V + \int_{\Gamma[X]} \nabla_{\Gamma[X]} H \cdot \nabla_{\Gamma[X]} \varphi^V = \int_{\Gamma[X]} Q \, \varphi^V; \qquad (2.12b)$$

$$\int_{\Gamma[X]} \partial^\bullet \nu \cdot \varphi^\nu - \int_{\Gamma[X]} \nabla_{\Gamma[X]} z \cdot \nabla_{\Gamma[X]} \varphi^\nu = \int_{\Gamma[X]} (HA - A^2) z \cdot \varphi^\nu$$
$$+ \int_{\Gamma[X]} \left(|\nabla_{\Gamma[X]} H|^2 \nu + A^2 \nabla_{\Gamma[X]} H \right) \cdot \varphi^\nu$$
$$+ 2 \int_{\Gamma[X]} (A \nabla_{\Gamma[X]} H) \cdot (\nabla_{\Gamma[X]} \varphi^\nu \nu)$$
$$+ \int_{\Gamma[X]} Q \, \nabla_{\Gamma[X]} \cdot \varphi^\nu - \int_{\Gamma[X]} QH \, \nu \cdot \varphi^\nu, \qquad (2.12c)$$

$$\int_{\Gamma[X]} z \cdot \varphi^z + \int_{\Gamma[X]} \nabla_{\Gamma[X]} \nu \cdot \nabla_{\Gamma[X]} \varphi^z = \int_{\Gamma[X]} |A|^2 \nu \cdot \varphi^z, \qquad (2.12d)$$

for all test functions $\varphi^v \in H^1(\Gamma[X])^3$, $\varphi^H \in H^1(\Gamma[X])$, $\varphi^V \in H^1(\Gamma[X])$, and $\varphi^\nu \in H^1(\Gamma[X])^3$, $\varphi^z \in H^1(\Gamma[X])^3$. Here, we use the Sobolev space $H^1(\Gamma) = \{u \in L^2(\Gamma) : \nabla_\Gamma u \in L^2(\Gamma)\}$. The system (2.11)–(2.12) is complemented with the initial data X^0, ν^0 and H^0.

We note that the last two terms on the right-hand side of (2.12c) result from the integration by parts formula (cf. [DE13a, Section 2.3])

$$\int_\Gamma \nabla_\Gamma f \cdot \varphi = - \int_\Gamma f \, \nabla_\Gamma \cdot \varphi + \int_\Gamma f \, H \nu \cdot \varphi, \qquad (2.13)$$

which we use here with $f = Q$. The terms in the second and third lines of (2.12c) result from this integration by parts (used with $\nu \cdot \varphi$ and $A \nabla_\Gamma H$ in the roles of f and φ, respectively) and the product rule:

$$\int_\Gamma -\nabla_\Gamma \cdot (A \nabla_\Gamma H)(\nu \cdot \varphi) = \int_\Gamma (A \nabla_\Gamma H) \cdot \nabla_\Gamma (\nu \cdot \varphi) - \int_\Gamma (A \nabla_\Gamma H) \cdot (H\nu)(\nu \cdot \varphi)$$
$$= \int_\Gamma (A \nabla_\Gamma H) \cdot (A\varphi) + \int_\Gamma (A \nabla_\Gamma H) \cdot (\nabla_\Gamma \varphi \, \nu) - \int_\Gamma (\nabla_\Gamma H) \cdot (HA\nu)(\nu \cdot \varphi)$$
$$= \int_\Gamma (A^2 \nabla_\Gamma H) \cdot \varphi + \int_\Gamma (A \nabla_\Gamma H) \cdot (\nabla_\Gamma \varphi \, \nu),$$

where we used $\nabla_\Gamma \nu = A = A^T$ and $A\nu = 0$.

It is instructive to compare our system (2.11)–(2.12) with the equations on which the algorithms of Dziuk [Dzi08] and of Barrett, Garcke and Nürnberg in [BGN07b, BGN08] are based. They both use the fact that $-H\nu = \Delta_\Gamma x_\Gamma$, with

x_Γ denoting the identity map on Γ, and also employ the above integration by parts formula (for $f = 1$).

Dziuk's algorithm is based on the following weak formulation [Dzi08, Problem 2]: Find $w \in H^1(\Gamma[X])^3$ and the velocity $v \in H^1(\Gamma[X])^3$ such that

$$\int_{\Gamma[X]} w \cdot \varphi^w = -\int_{\Gamma[X]} \nabla_{\Gamma[X]} x_\Gamma \cdot \nabla_{\Gamma[X]} \varphi^w,$$

$$\int_{\Gamma[X]} v \cdot \varphi^v = \tfrac{1}{2} \int_{\Gamma[X]} |w|^2 \nabla_{\Gamma[X]} \cdot \varphi^v + \int_{\Gamma[X]} \nabla_{\Gamma[X]} w \cdot \nabla_{\Gamma[X]} \varphi^v$$

$$+ \int_{\Gamma[X]} (\nabla_{\Gamma[X]} \cdot w)(\nabla_{\Gamma[X]} \cdot \varphi^v) - \int_{\Gamma[X]} D(\varphi^v) \nabla_{\Gamma[X]} x_\Gamma \cdot \nabla_{\Gamma[X]} w,$$

$$(2.14)$$

for all $\varphi^w \in H^1(\Gamma[X])^3$ and $\varphi^v \in H^1(\Gamma[X])^3$, and where $D(\varphi)$ denotes the symmetric tensor $D(\varphi) = \nabla_{\Gamma[X]} \varphi + (\nabla_{\Gamma[X]} \varphi)^T$. The above formulation is based on the idea that if $w = -H\nu = \Delta_\Gamma x_\Gamma$, then the first variation of the Willmore functional (i.e. the negative surface velocity $-v$) can be expressed using the second expression in (2.14), see [Dzi08, Lemma 3].

The algorithm of Barrett, Garcke and Nürnberg [BGN08] is based on the following weak formulation: Find the velocity $v \in H^1(\Gamma[X])^3$ and $H \in H^1(\Gamma[X])$, $A \in H^1(\Gamma[X])^{3\times3}$ such that

$$\int_{\Gamma[X]} v \cdot (\nu \varphi^v) = -\int_{\Gamma[X]} \nabla_{\Gamma[X]} H \cdot \nabla_{\Gamma[X]} \varphi^v + \int_{\Gamma[X]} Q \varphi^v,$$

$$\int_{\Gamma[X]} H\nu \cdot \varphi^H = \int_{\Gamma[X]} \nabla_{\Gamma[X]} x_\Gamma \cdot \nabla_{\Gamma[X]} \varphi^H,$$

$$(2.15)$$

$$\int_{\Gamma[X]} A \cdot \varphi^A = -\int_{\Gamma[X]} \nu \cdot (\nabla_{\Gamma[X]} \cdot \varphi^A) + \int_{\Gamma[X]} H\nu \cdot (\varphi^A \nu),$$

for all $\varphi^v \in H^1(\Gamma[X])^3$, $\varphi^H \in H^1(\Gamma[X])$ and $\varphi^A \in H^1(\Gamma[X])^{3\times3}$. The first equation in the above weak formulation is obtained by taking the scalar product of (2.2) with ν, the second equation is the weak form of the identity $-H\nu = \Delta_\Gamma x_\Gamma$, while the last equation is obtained by choosing $\varphi = \varphi^A \nu$ in the integration by parts formula and using $A = \nabla_{\Gamma[X]} \nu$.

The key idea for all these approaches is to formulate Willmore flow using additional variables which satisfy differential equations, derived using geometric identities for surfaces or from the properties of Willmore flow.

The characteristic feature of our approach via the evolution equations of Lemma 2.1 is that it has only parabolic equations for the natural geometric variables that already appear in the velocity law for Willmore flow (2.2). This will allow us to prove convergence, which is not known for the previously proposed algorithms.

3 Evolving finite element semi-discretization

3.1 Evolving surface finite elements

(The text of this preparatory subsection is taken verbatim from [KLL19, Section 3.1].) We formulate the evolving surface finite element (ESFEM) discretization for the velocity law coupled with evolution equations on the evolving surface, following the description in [KLLP17, KLL19], which is based on [Dzi88] and [Dem09]. We use triangular finite elements on the surface and continuous piecewise polynomial basis functions of degree k, as defined in [Dem09, Section 2.5].

We triangulate the given smooth initial surface Γ^0 by an admissible family of triangulations $\mathcal{T}_h$ of decreasing maximal element diameter h; see [DE07] for the notion of an admissible triangulation, which includes quasi-uniformity and shape regularity. For a given triangulation of the initial surface Γ^0, we denote by $\mathbf{x}^0$ the vector in $\mathbb{R}^{3N}$ that collects all nodes p_j $(j = 1, \ldots, N)$ of the initial triangulation. By piecewise polynomial interpolation of degree k, the nodal vector defines an approximate surface Γ_h^0 that interpolates Γ^0 in the nodes p_j. We will evolve the jth node in time, denoted $x_j(t)$ with initial condition $x_j(0) = p_j$, and collect the nodes at time t in a column vector

$$\mathbf{x}(t) \in \mathbb{R}^{3N}.$$

We just write $\mathbf{x}$ for $\mathbf{x}(t)$ when the dependence on t is not important.

By piecewise polynomial interpolation on the plane reference triangle that corresponds to every curved triangle of the triangulation, the nodal vector $\mathbf{x}$ defines a closed surface denoted by $\Gamma_h[\mathbf{x}]$. We can then define globally continuous finite element *basis functions*

$$\phi_i[\mathbf{x}] : \Gamma_h[\mathbf{x}] \to \mathbb{R}, \qquad i = 1, \ldots, N,$$

which have the property that on every triangle their pullback to the reference triangle are polynomials of degree k, and which satisfy at the node x_j

$$\phi_i[\mathbf{x}](x_j) = \delta_{ij} \quad \text{for all } i, j = 1, \ldots, N.$$

These functions span the finite element space on $\Gamma_h[\mathbf{x}]$,

$$S_h[\mathbf{x}] = S_h(\Gamma_h[\mathbf{x}]) = \operatorname{span}\big\{\phi_1[\mathbf{x}], \phi_2[\mathbf{x}], \ldots, \phi_N[\mathbf{x}]\big\}.$$

For a finite element function $u_h \in S_h[\mathbf{x}]$, the tangential gradient $\nabla_{\Gamma_h[\mathbf{x}]} u_h$ is defined piecewise on each element.

The discrete surface at time t is parametrized by the initial discrete surface via the map $X_h(\cdot, t) : \Gamma_h^0 \to \Gamma_h[\mathbf{x}(t)]$ defined by

$$X_h(p_h, t) = \sum_{j=1}^{N} x_j(t) \, \phi_j[\mathbf{x}(0)](p_h), \qquad p_h \in \Gamma_h^0,$$

which has the properties that $X_h(p_j, t) = x_j(t)$ for $j = 1, \ldots, N$, that $X_h(p_h, 0) = p_h$ for all $p_h \in \Gamma_h^0$, and

$$\Gamma_h[\mathbf{x}(t)] = \Gamma[X_h(\cdot, t)],$$

where the right-hand side equals $\{X_h(p_h, t) : p_h \in \Gamma_h^0\}$ like in Section 2.1.

The *discrete velocity* $v_h(x, t) \in \mathbb{R}^3$ at a point $x = X_h(p_h, t) \in \Gamma[X_h(\cdot, t)]$ is given by

$$\partial_t X_h(p_h, t) = v_h(X_h(p_h, t), t).$$

In view of the transport property of the basis functions [DE07],

$$\frac{\mathrm{d}}{\mathrm{d}t} \Big(\phi_j[\mathbf{x}(t)](X_h(p_h, t)) \Big) = 0,$$

the discrete velocity equals, for $x \in \Gamma_h[\mathbf{x}(t)]$,

$$v_h(x, t) = \sum_{j=1}^{N} v_j(t)\, \phi_j[\mathbf{x}(t)](x) \qquad \text{with} \ \ v_j(t) = \dot{x}_j(t),$$

where the dot denotes the time derivative $\mathrm{d}/\mathrm{d}t$. Hence, the discrete velocity $v_h(\cdot, t)$ is in the finite element space $S_h[\mathbf{x}(t)]$, with nodal vector $\mathbf{v}(t) = \dot{\mathbf{x}}(t)$.

The *discrete material derivative* of a finite element function $u_h(x, t)$ with nodal values $u_j(t)$ is

$$\partial_h^{\bullet} u_h(x, t) = \frac{\mathrm{d}}{\mathrm{d}t} u_h(X_h(p_h, t)) = \sum_{j=1}^{N} \dot{u}_j(t)\phi_j[\mathbf{x}(t)](x) \quad \text{at} \quad x = X_h(p_h, t).$$

3.2 ESFEM spatial semi-discretization

A *preliminary* finite element spatial semi-discretization of the fourth-order parabolic system (2.12) coupled with the velocity and position equations (2.11) reads as follows: Find the unknown nodal vector $\mathbf{x}(t) \in \mathbb{R}^{3N}$ of the finite element surface parametrization $X_h(\cdot, t) \in S_h[\mathbf{x}^0]^3$ and the unknown finite element velocity $v_h(\cdot, t) \in S_h[\mathbf{x}(t)]^3$, and the finite element functions $H_h(\cdot, t) \in S_h[\mathbf{x}(t)]$, $V_h(\cdot, t) \in S_h[\mathbf{x}(t)]$, and $\nu_h(\cdot, t) \in S_h[\mathbf{x}(t)]^3$, $z_h(\cdot, t) \in S_h[\mathbf{x}(t)]^3$ such that

$$\partial_t X_h(p_h, t) = v_h(X_h(p_h, t), t), \qquad p_h \in \Gamma_h^0, \tag{3.1a}$$

with

$$v_h = \widetilde{I}_h(V_h \nu_h), \tag{3.1b}$$

where $\widetilde{I}_h = \widetilde{I}_h[\mathbf{x}] : C(\Gamma_h[\mathbf{x}]) \to S_h(\Gamma_h[\mathbf{x}])$ denotes the finite element interpolation operator on the discrete surface $\Gamma_h[\mathbf{x}]$.

The functions H_h, V_h, ν_h, z_h are determined by the ESFEM semi-discretization of (2.12), denoting by $A_h = \frac{1}{2}(\nabla_{\Gamma_h[\mathbf{x}]}\nu_h + (\nabla_{\Gamma_h[\mathbf{x}]}\nu_h)^T)$ the symmetric part of $\nabla_{\Gamma_h[\mathbf{x}]}\nu_h$, and by $Q_h = -\frac{1}{2}H_h^3 + |A_h|^2 H_h$ the cubic term,

$$\int_{\Gamma_h[\mathbf{x}]} \partial_h^\bullet H_h\, \varphi_h^H - \int_{\Gamma_h[\mathbf{x}]} \nabla_{\Gamma_h[\mathbf{x}]}V_h \cdot \nabla_{\Gamma_h[\mathbf{x}]}\varphi_h^H = -\int_{\Gamma_h[\mathbf{x}]} |A_h|^2\, V_h\, \varphi_h^H, \qquad (3.2a)$$

$$\int_{\Gamma_h[\mathbf{x}]} V_h\varphi_h^V + \int_{\Gamma_h[\mathbf{x}]} \nabla_{\Gamma_h[\mathbf{x}]}H_h \cdot \nabla_{\Gamma_h[\mathbf{x}]}\varphi_h^V = \int_{\Gamma_h[\mathbf{x}]} Q_h\varphi_h^V; \qquad (3.2b)$$

$$\int_{\Gamma_h[\mathbf{x}]} \partial_h^\bullet \nu_h \cdot \varphi_h^\nu - \int_{\Gamma_h[\mathbf{x}]} \nabla_{\Gamma_h[\mathbf{x}]}z_h \cdot \nabla_{\Gamma_h[\mathbf{x}]}\varphi_h^\nu = \int_{\Gamma_h[\mathbf{x}]} (H_h A_h - A_h^2)z_h \cdot \varphi_h^\nu$$
$$+ \int_{\Gamma_h[\mathbf{x}]} \left(|\nabla_{\Gamma_h[\mathbf{x}]}H_h|^2\nu_h + A_h^2 \nabla_{\Gamma_h[\mathbf{x}]}H_h\right) \cdot \varphi_h^\nu$$
$$+ 2\int_{\Gamma_h[\mathbf{x}]} (A_h\nabla_{\Gamma_h[\mathbf{x}]}H_h) \cdot (\nabla_{\Gamma_h[\mathbf{x}]}\varphi_h^\nu\, \nu_h)$$
$$+ \int_{\Gamma_h[\mathbf{x}]} Q_h\, \nabla_{\Gamma_h[\mathbf{x}]} \cdot \varphi_h^\nu - \int_{\Gamma_h[\mathbf{x}]} Q_h H_h\, \nu_h \cdot \varphi_h^\nu, \qquad (3.2c)$$

$$\int_{\Gamma_h[\mathbf{x}]} z_h \cdot \varphi_h^z + \int_{\Gamma_h[\mathbf{x}]} \nabla_{\Gamma_h[\mathbf{x}]}\nu_h \cdot \nabla_{\Gamma_h[\mathbf{x}]}\varphi_h^z = \int_{\Gamma_h[\mathbf{x}]} |A_h|^2\nu_h \cdot \varphi_h^z, \qquad (3.2d)$$

for all $\varphi_h^v \in S_h[\mathbf{x}(t)]^3$, $\varphi_h^H \in S_h[\mathbf{x}(t)]$, $\varphi_h^V \in S_h[\mathbf{x}(t)]$, $\varphi_h^\nu \in S_h[\mathbf{x}(t)]^3$, and $\varphi_h^z \in S_h[\mathbf{x}(t)]^3$.

The initial values for the nodal vector $\mathbf{x}$ are taken as the positions of the nodes of the triangulation of the given initial surface Γ^0. The initial data for H_h and ν_h are determined by Lagrange interpolation of H^0 and ν^0, respectively.

We note that the finite element functions ν_h and H_h are *not* the normal vector and mean curvature of the discrete surface $\Gamma_h[\mathbf{x}]$.

Our error analysis indicates that the above semi-discretization is not yet convergent of optimal order. To achieve this, we need to add correction terms to (3.2b) and (3.2d). These will be constructed in Section 3.4, using the matrix–vector formulation of the next subsection.

Remark 3.1 In the evolving finite element algorithm for mean curvature flow of [KLL19], the velocity approximation v_h was chosen as the Ritz projection of $V_h\nu_h$ (with $V_h = -H_h$ for mean curvature flow), as opposed to the simpler nodal interpolation of $V_h\nu_h$ proposed here. Meanwhile, we have realized that the same H^1 error bounds as in [KLL19] can be proved for mean curvature flow also with interpolation instead of the Ritz projection. Here we use interpolation for the velocity approximation from the outset, and in our proof of stability we will provide the argument that could have been used for interpolation also in the case of mean curvature flow.

3.3 Matrix–vector formulation

We collect the nodal values in column vectors $\mathbf{v} = (v_j) \in \mathbb{R}^{3N}$, $\mathbf{H} = (H_j) \in \mathbb{R}^N$, $\mathbf{n} = (\nu_j) \in \mathbb{R}^{3N}$ and $\boldsymbol{V} = (V_j) \in \mathbb{R}^N$, $\mathbf{z} = (z_j) \in \mathbb{R}^{3N}$, and denote

$$\mathbf{u} = \begin{pmatrix} \mathbf{H} \\ \mathbf{n} \end{pmatrix} \in \mathbb{R}^{4N} \quad \text{and} \quad \mathbf{w} = \begin{pmatrix} \boldsymbol{V} \\ \mathbf{z} \end{pmatrix} \in \mathbb{R}^{4N}.$$

We define the surface-dependent mass matrix $\mathbf{M}(\mathbf{x}) \in \mathbb{R}^{N \times N}$ and stiffness matrix $\mathbf{A}(\mathbf{x}) \in \mathbb{R}^{N \times N}$ on the surface determined by the nodal vector $\mathbf{x}$:

$$\mathbf{M}(\mathbf{x})|_{ij} = \int_{\Gamma_h[\mathbf{x}]} \phi_i[\mathbf{x}]\phi_j[\mathbf{x}],$$
$$\mathbf{A}(\mathbf{x})|_{ij} = \int_{\Gamma_h[\mathbf{x}]} \nabla_{\Gamma_h[\mathbf{x}]}\phi_i[\mathbf{x}] \cdot \nabla_{\Gamma_h[\mathbf{x}]}\phi_j[\mathbf{x}], \qquad i, j = 1, \dots, N,$$

with the finite element nodal basis functions $\phi_j[\mathbf{x}] \in S_h[\mathbf{x}]$. We further let, for $d = 3$ or 4 (with the identity matrices $I_d \in \mathbb{R}^{d \times d}$)

$$\mathbf{M}^{[d]}(\mathbf{x}) = I_d \otimes \mathbf{M}(\mathbf{x}), \qquad \mathbf{A}^{[d]}(\mathbf{x}) = I_d \otimes \mathbf{A}(\mathbf{x}).$$

When no confusion can arise, we write $\mathbf{M}(\mathbf{x})$ for $\mathbf{M}^{[d]}(\mathbf{x})$ and $\mathbf{A}(\mathbf{x})$ for $\mathbf{A}^{[d]}(\mathbf{x})$.

We recall that we denote by $|A_h|^2$ the squared Frobenius norm of $A_h = \frac{1}{2}\big(\nabla_{\Gamma_h[\mathbf{x}]}\nu_h + (\nabla_{\Gamma_h[\mathbf{x}]}\nu_h)^T\big)$. We define the surface- and $|A_h|^2$-dependent matrix $\mathbf{F}_1(\mathbf{x}, \mathbf{u}) \in \mathbb{R}^{N \times N}$ by

$$\mathbf{F}_1(\mathbf{x}, \mathbf{u})|_{ij} = -\int_{\Gamma_h[\mathbf{x}]} |A_h|^2 \, \phi_i[\mathbf{x}]\phi_j[\mathbf{x}], \qquad i, j = 1, \dots, N, \qquad (3.3a)$$

and similarly the matrix $\mathbf{F}_2(\mathbf{x}, \mathbf{u}) \in \mathbb{R}^{3N \times 3N}$ by

$$\mathbf{F}_2(\mathbf{x}, \mathbf{u})|_{ij} = \int_{\Gamma_h[\mathbf{x}]} (H_h A_h - A_h^2) \, \phi_i[\mathbf{x}] \cdot \phi_j[\mathbf{x}], \qquad i, j = 1, \dots, 3N. \quad (3.3b)$$

We then define the block diagonal matrix

$$\mathbf{F}(\mathbf{x}, \mathbf{u}) = \begin{pmatrix} \mathbf{F}_1(\mathbf{x}, \mathbf{u}) & 0 \\ 0 & \mathbf{F}_2(\mathbf{x}, \mathbf{u}) \end{pmatrix} \in \mathbb{R}^{4N \times 4N}.$$

We define the non-linear functions $\mathbf{f}(\mathbf{x}, \mathbf{u}), \mathbf{g}(\mathbf{x}, \mathbf{u}) \in \mathbb{R}^{4N}$ by

$$\mathbf{f}(\mathbf{x}, \mathbf{u}) = \begin{pmatrix} 0 \\ \mathbf{f}_2(\mathbf{x}, \mathbf{u}) \end{pmatrix} \quad \text{and} \quad \mathbf{g}(\mathbf{x}, \mathbf{u}) = \begin{pmatrix} \mathbf{g}_1(\mathbf{x}, \mathbf{u}) \\ \mathbf{g}_2(\mathbf{x}, \mathbf{u}) \end{pmatrix},$$

with $\mathbf{f}_2(\mathbf{x}, \mathbf{u}) \in \mathbb{R}^{3N}$, $\mathbf{g}_1(\mathbf{x}, \mathbf{u}) \in \mathbb{R}^{N}$ and $\mathbf{g}_2(\mathbf{x}, \mathbf{u}) \in \mathbb{R}^{3N}$, given by

$$
\begin{aligned}
\mathbf{f}_2(\mathbf{x}, \mathbf{u})|_{j+(\ell-1)N} &= \int_{\Gamma_h[\mathbf{x}]} \left(|\nabla_{\Gamma_h[\mathbf{x}]} H_h|^2 \nu_h + A_h^2 \nabla_{\Gamma_h[\mathbf{x}]} H_h \right)_\ell \phi_j[\mathbf{x}] \\
&\quad + 2 \int_{\Gamma_h[\mathbf{x}]} \left(A_h \nabla_{\Gamma_h[\mathbf{x}]} H_h \right) \cdot \left(\nabla_{\Gamma_h[\mathbf{x}]} \phi_j[\mathbf{x}] (\nu_h)_\ell \right) \\
&\quad + \int_{\Gamma_h[\mathbf{x}]} Q_h \left(\nabla_{\Gamma_h[\mathbf{x}]} \right)_\ell \phi_j[\mathbf{x}] - \int_{\Gamma_h[\mathbf{x}]} Q_h H_h (\nu_h)_\ell \, \phi_j[\mathbf{x}], \\
\mathbf{g}_1(\mathbf{x}, \mathbf{u})|_{j} &= \int_{\Gamma_h[\mathbf{x}]} Q_h \, \phi_j[\mathbf{x}], \\
\mathbf{g}_2(\mathbf{x}, \mathbf{u})|_{j+(\ell-1)N} &= \int_{\Gamma_h[\mathbf{x}]} |A_h|^2 (\nu_h)_\ell \, \phi_j[\mathbf{x}],
\end{aligned}
$$

for $j = 1, \dots, N$ and $\ell = 1, 2, 3$.

The position and velocity equations (3.1) are equivalent to

$$
\begin{aligned}
\dot{\mathbf{x}} &= \mathbf{v}, &\text{(3.4a)} \\
\mathbf{v} &= \mathbf{V} \bullet \mathbf{n}, &\text{(3.4b)}
\end{aligned}
$$

where $\bullet$ denotes the componentwise product of vectors, $(\mathbf{V} \bullet \mathbf{n})_j = (\mathbf{V}_j \mathbf{n}_j)$. With the notation introduced above, the equations (3.2) can be written in the following matrix–vector form, where we recall that $\mathbf{u} = (\mathbf{H}; \mathbf{n})$ and $\mathbf{w} = (\mathbf{V}; \mathbf{z})$:

$$
\begin{aligned}
\mathbf{M}^{[4]}(\mathbf{x})\dot{\mathbf{u}} - \mathbf{A}^{[4]}(\mathbf{x})\mathbf{w} &= \mathbf{F}(\mathbf{x}, \mathbf{u})\mathbf{w} + \mathbf{f}(\mathbf{x}, \mathbf{u}), &\text{(3.5a)} \\
\mathbf{M}^{[4]}(\mathbf{x})\mathbf{w} + \mathbf{A}^{[4]}(\mathbf{x})\mathbf{u} &= \mathbf{g}(\mathbf{x}, \mathbf{u}). &\text{(3.5b)}
\end{aligned}
$$

3.4 Modified ESFEM spatial semi-discretization

The initial values of V_h and z_h in (3.2b) and (3.2d), or equivalently of $\mathbf{w}$ in the algebraic equations (3.5b), are not at our disposal, but are determined by the equations. On the other hand, our error analysis yields convergence in the H^1-norm of optimal order k only if these initial values are $O(h^k)$ close to the Ritz projection of the exact initial values in the H^1-norm. We can verify this only in the weaker H^{-1}-norm but not in the H^1-norm. For the method as stated (3.4)–(3.5), we can then show convergence in the H^1-norm only of reduced order $k-2$ (and only for polynomial degree $k \geq 4$).

To obtain the full order k (for polynomial degree $k \geq 2$), we modify equation (3.5b) by adding a time-independent correction term such that the initial value of the nodal vector $\mathbf{w}$ equals the nodal vector $\bar{\mathbf{w}}^*(0)$ of exact initial values: Let $\bar{\mathbf{w}}(0)$ be the vector obtained from (3.5b) at time $t = 0$, and set

$$
\boldsymbol{\vartheta} = \mathbf{M}^{[4]}(\mathbf{x}^0)(\bar{\mathbf{w}}^*(0) - \bar{\mathbf{w}}(0)). \tag{3.6}
$$

We then modify (3.5b) to

$$\mathbf{M}^{[4]}(\mathbf{x})\mathbf{w} + \mathbf{A}^{[4]}(\mathbf{x})\mathbf{u} = \mathbf{g}(\mathbf{x}, \mathbf{u}) + \boldsymbol{\vartheta}, \qquad (3.7)$$

so that the initial value becomes $\mathbf{w}(0) = \bar{\mathbf{w}}^*(0)$. Note that the time-differentiated equation (3.7) is still the same as the time-differentiated equation (3.5b), since $\boldsymbol{\vartheta}$ does not depend on time.

As an aside, to interpret this modification in the weak form (3.2) (which is not necessary for the algorithm or its error analysis), we introduce

$$\boldsymbol{\eta}(t) = (\eta_j(t)) = \mathbf{M}^{[4]}(\mathbf{x}(t))^{-1}\boldsymbol{\vartheta}$$

with $\eta_j = (\eta_j^V; \eta_j^z) \in \mathbb{R}^{1+3}$ and define the finite element functions

$$\eta_h^V(x,t) = \sum_{j=1}^{N} \eta_j^V(t)\varphi_j[\mathbf{x}(t)](x) \qquad \text{for} \quad x \in \Gamma_h[\mathbf{x}(t)],$$

$$\eta_h^z(x,t) = \sum_{j=1}^{N} \eta_j^z(t)\,\varphi_j[\mathbf{x}(t)](x) \qquad \text{for} \quad x \in \Gamma_h[\mathbf{x}(t)].$$

Then, the modification (3.7) corresponds to modifying (3.2b) and (3.2d) to

$$\int_{\Gamma_h[\mathbf{x}]} V_h \varphi_h^V + \int_{\Gamma_h[\mathbf{x}]} \nabla_{\Gamma_h[\mathbf{x}]} H_h \cdot \nabla_{\Gamma_h[\mathbf{x}]} \varphi_h^V = \int_{\Gamma_h[\mathbf{x}]} Q_h \varphi_h^V + \int_{\Gamma_h[\mathbf{x}]} \eta_h^V \varphi_h^V, \quad (3.8a)$$

$$\int_{\Gamma_h[\mathbf{x}]} z_h \cdot \varphi_h^z + \int_{\Gamma_h[\mathbf{x}]} \nabla_{\Gamma_h[\mathbf{x}]} \nu_h \cdot \nabla_{\Gamma_h[\mathbf{x}]} \varphi_h^z = \int_{\Gamma_h[\mathbf{x}]} |A_h|^2 \nu_h \cdot \varphi_h^z + \int_{\Gamma_h[\mathbf{x}]} \eta_h^z \, \varphi_h^z. \quad (3.8b)$$

3.5 Lifts

(The text of this preparatory subsection is taken verbatim from [KLL19, Section 3.4].) We need to compare functions on the *exact surface* $\Gamma(t) = \Gamma[X(\cdot,t)]$ with functions on the *discrete surface* $\Gamma_h(t) = \Gamma_h[\mathbf{x}(t)]$. To this end, we further work with functions on the *interpolated surface* $\Gamma_h^*(t) = \Gamma_h[\mathbf{x}^*(t)]$, where $\mathbf{x}^*(t)$ denotes the nodal vector collecting the grid points $x_j^*(t) = X(p_j, t)$ on the exact surface.

Any finite element function $w_h : \Gamma_h(t) \to \mathbb{R}^m$ ($m = 1$ or 3) on the discrete surface, with nodal values w_j, is associated with a finite element function $\widehat{w}_h$ on the interpolated surface $\Gamma_h^*(t)$ that is defined by

$$\widehat{w}_h = \sum_{j=1}^{N} w_j \phi_j[\mathbf{x}^*(t)].$$

This can be further lifted to a function on the exact surface by using the *lift operator l*, which was introduced for linear and higher-order surface approximations in [Dzi88] and [Dem09], respectively. The lift operator l maps a

function on the interpolated surface Γ_h^* to a function on the exact surface Γ, provided that Γ_h^* is sufficiently close to Γ. The exact regular surface Γ can be represented, in some neighbourhood of the surface, by a smooth signed distance function $d : \mathbb{R}^3 \times [0, T] \to \mathbb{R}$, cf. [DE07, Section 2.1], such that Γ is the zero level set of d (i.e., $x \in \Gamma$ if and only if $d(x) = 0$), with negative values of d in the interior. Using this distance function, the lift of a continuous function $\eta_h : \Gamma_h^* \to \mathbb{R}^m$ is defined as $\eta_h^l(y) := \eta_h(x)$ for $x \in \Gamma_h^*$, where for every $x \in \Gamma_h^*$ the point $y = y(x) \in \Gamma$ is uniquely defined via $y = x - \nu(y)d(x)$.

The composed lift L from finite element functions on the discrete surface $\Gamma_h(t)$ to functions on the exact surface $\Gamma(t)$ via the interpolated surface $\Gamma_h^*(t)$ is denoted by

$$w_h^L = (\widehat{w}_h)^l.$$

4 Convergence of the modified ESFEM semi-discretization

We are now in the position to formulate the first main result of this paper, which yields optimal-order error bounds for the finite element semi-discretization (3.1)–(3.2) with the modification (3.8) using finite elements of polynomial degree $k \geq 2$. We denote by $\Gamma(t) = \Gamma[X(\cdot, t)]$ the exact surface and by $\Gamma_h(t) = \Gamma[X_h(\cdot, t)] = \Gamma_h[\mathbf{x}(t)]$ the discrete surface at time t, and introduce the notation

$$x_h^L(x, t) = X_h^L(p, t) \in \Gamma_h(t) \qquad \text{for} \quad x = X(p, t) \in \Gamma(t).$$

Theorem 4.1 *Consider the ESFEM semi-discretization* (3.1)–(3.2) *with the modification* (3.8) *(i.e.,* (3.4)–(3.5) *with the modification* (3.7) *in matrix–vector form) of the Willmore flow problem* (2.11)–(2.12)*, using evolving surface finite elements of polynomial degree $k \geq 2$. Suppose that the problem admits an exact solution (X, v, ν, H, z, V) that is sufficiently differentiable on the time interval $t \in [0, T]$, and that the flow map $X(\cdot, t) : \Gamma^0 \to \Gamma(t) \subset \mathbb{R}^3$ is non-degenerate so that $\Gamma(t)$ is a regular surface on the time interval $t \in [0, T]$.*

Then, there exists a constant $h_0 > 0$ such that for all mesh sizes $h \leq h_0$ the semi-discrete equations have a unique solution on $[0, T]$, and the following error bounds for the lifts of the discrete position, velocity, normal vector, mean curvature and further geometric quantities hold over the exact surface $\Gamma(t) = \Gamma[X(\cdot, t)]$ for $0 \leq t \leq T$:

$$\|x_h^L(\cdot, t) - \mathrm{id}_{\Gamma(t)}\|_{H^1(\Gamma(t))^3} \leq Ch^k, \qquad \|v_h^L(\cdot, t) - v(\cdot, t)\|_{H^1(\Gamma(t))^3} \leq Ch^k,$$

$$\|H_h^L(\cdot, t) - H(\cdot, t)\|_{H^1(\Gamma(t))} \leq Ch^k, \qquad \|\nu_h^L(\cdot, t) - \nu(\cdot, t)\|_{H^1(\Gamma(t))^3} \leq Ch^k,$$

$$\|V_h^L(\cdot, t) - V(\cdot, t)\|_{H^1(\Gamma(t))} \leq Ch^k, \qquad \|z_h^L(\cdot, t) - \nabla_{\Gamma(t)} H(\cdot, t)\|_{H^1(\Gamma(t))^3} \leq Ch^k,$$

and also

$$\|X_h^l(\cdot, t) - X(\cdot, t)\|_{H^1(\Gamma^0)^3} \leq Ch^k,$$

where the constant C is independent of h and t, but depends on bounds of higher derivatives of the solution (X, v, ν, H, z, V) of Willmore flow and on the length T of the time interval.

Sufficient regularity assumptions are the following: with bounds that are uniform in $t \in [0, T]$, we assume $X(\cdot, t) \in H^{k+1}(\Gamma^0)$, $v(\cdot, t) \in H^{k+1}(\Gamma(X(\cdot, t)))^3$, and for $u = (\nu, H)$ we have $u(\cdot, t), \partial^\bullet u(\cdot, t), \partial^{(2)} u \in W^{k+1,\infty}(\Gamma(X(\cdot, t)))^4$, for $w = (V, z)$ we have $w, \partial^\bullet w \in W^{k+1,\infty}(\Gamma(X(\cdot, t)))^4$.

We note that the remarks made after the convergence result in [KLL19] apply also here, in particular on the preservation of admissibility of the triangulation over the whole time interval $[0, T]$ for sufficiently small h.

5 Stability

5.1 Preparation: Estimates relating different finite element surfaces

In our previous work [KLLP17, Section 4] and [KLL19, Section 7.1] we proved technical results relating different finite element surfaces, which we recapitulate here, *taking verbatim the text of [KLL19, Section 7.1] in this preparatory subsection.*

The finite element matrices of Section 3.3 induce discrete versions of Sobolev norms. Let $\mathbf{x} \in \mathbb{R}^{3N}$ be a nodal vector defining the discrete surface $\Gamma_h[\mathbf{x}]$. For any nodal vector $\mathbf{w} = (w_j) \in \mathbb{R}^N$, with the corresponding finite element function $w_h = \sum_{j=1}^{N} w_j \phi_j[\mathbf{x}] \in S_h[\mathbf{x}]$, we define the following norms, where $\mathbf{K}(\mathbf{x}) = \mathbf{M}(\mathbf{x}) + \mathbf{A}(\mathbf{x})$ in the third line:

$$\|\mathbf{w}\|_{\mathbf{M}(\mathbf{x})}^2 = \mathbf{w}^T \mathbf{M}(\mathbf{x}) \mathbf{w} = \|w_h\|_{L^2(\Gamma_h[\mathbf{x}])}^2, \tag{5.1}$$

$$\|\mathbf{w}\|_{\mathbf{A}(\mathbf{x})}^2 = \mathbf{w}^T \mathbf{A}(\mathbf{x}) \mathbf{w} = \|\nabla_{\Gamma_h[\mathbf{x}]} w_h\|_{L^2(\Gamma_h[\mathbf{x}])}^2, \tag{5.2}$$

$$\|\mathbf{w}\|_{\mathbf{K}(\mathbf{x})}^2 = \mathbf{w}^T \mathbf{K}(\mathbf{x}) \mathbf{w} = \|w_h\|_{H^1(\Gamma_h[\mathbf{x}])}^2. \tag{5.3}$$

When $\mathbf{w} \in \mathbb{R}^{dN}$, so that the corresponding finite element function w_h maps into $\mathbb{R}^d$, we write in the following $\|w_h\|_{L^2(\Gamma)}$ for $\|w_h\|_{L^2(\Gamma)^d}$ and $\|w_h\|_{H^1(\Gamma)}$ for $\|w_h\|_{H^1(\Gamma)^d}$.

Let now $\mathbf{x}, \mathbf{y} \in \mathbb{R}^{3N}$ be two nodal vectors defining discrete surfaces $\Gamma_h[\mathbf{x}]$ and $\Gamma_h[\mathbf{y}]$, respectively. We denote the difference by $\mathbf{e} = (e_j) = \mathbf{x} - \mathbf{y} \in \mathbb{R}^{3N}$. For $\theta \in [0, 1]$, we consider the intermediate surface $\Gamma_h^\theta = \Gamma_h[\mathbf{y} + \theta \mathbf{e}]$ and the corresponding finite element functions given as

$$e_h^\theta = \sum_{j=1}^{N} e_j \phi_j[\mathbf{y} + \theta \mathbf{e}],$$

and in the same way, for any vectors $\mathbf{w}, \mathbf{z} \in \mathbb{R}^N$,

$$w_h^\theta = \sum_{j=1}^{N} w_j \phi_j[\mathbf{y} + \theta \mathbf{e}] \quad \text{and} \quad z_h^\theta = \sum_{j=1}^{N} z_j \phi_j[\mathbf{y} + \theta \mathbf{e}].$$
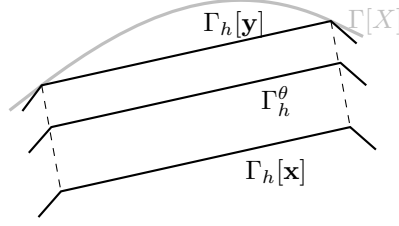
Fig. 5.1 The construction of the intermediate surfaces Γ_h^θ

Figure 5.1 illustrates the described construction.

The following formulae relate the mass and stiffness matrices for the discrete surfaces $\Gamma_h[\mathbf{x}]$ and $\Gamma_h[\mathbf{y}]$. They result from the Leibniz rule and are given in Lemma 4.1 of [KLLP17].

Lemma 5.1 *In the above setting, the following identities hold true:*

$$\mathbf{w}^T(\mathbf{M}(\mathbf{x}) - \mathbf{M}(\mathbf{y}))\mathbf{z} = \int_0^1 \int_{\Gamma_h^\theta} w_h^\theta (\nabla_{\Gamma_h^\theta} \cdot e_h^\theta) z_h^\theta \, \mathrm{d}\theta, \qquad (5.4)$$

$$\mathbf{w}^T(\mathbf{A}(\mathbf{x}) - \mathbf{A}(\mathbf{y}))\mathbf{z} = \int_0^1 \int_{\Gamma_h^\theta} \nabla_{\Gamma_h^\theta} w_h^\theta \cdot (D_{\Gamma_h^\theta} e_h^\theta) \nabla_{\Gamma_h^\theta} z_h^\theta \, \mathrm{d}\theta, \qquad (5.5)$$

where $D_{\Gamma_h^\theta} e_h^\theta = \mathrm{tr}(E^\theta) I_3 - (E^\theta + (E^\theta)^T)$ with $E^\theta = \nabla_{\Gamma_h^\theta} e_h^\theta \in \mathbb{R}^{3\times 3}$.

The following lemma combines Lemmas 4.2 and 4.3 of [KLLP17].

Lemma 5.2 *In the above setting, if*

$$\|\nabla_{\Gamma_h[\mathbf{y}]} e_h^0\|_{L^\infty(\Gamma_h[\mathbf{y}])} \leq \tfrac{1}{2},$$

then, for $0 \leq \theta \leq 1$ and $1 \leq p \leq \infty$, the finite element function $w_h^\theta = \sum_{j=1}^N w_j \phi_j[\mathbf{y} + \theta\mathbf{e}]$ on $\Gamma_h^\theta = \Gamma_h[\mathbf{y} + \theta\mathbf{e}]$ is bounded by

$$\|w_h^\theta\|_{L^p(\Gamma_h^\theta)} \leq c_p \|w_h^0\|_{L^p(\Gamma_h^0)},$$
$$\|\nabla_{\Gamma_h^\theta} w_h^\theta\|_{L^p(\Gamma_h^\theta)} \leq c_p \|\nabla_{\Gamma_h^0} w_h^0\|_{L^p(\Gamma_h^0)},$$

where c_p is an absolute constant (in particular, independent of $0 \leq \theta \leq 1$ and h). Moreover, $c_\infty = 2$.

The first estimate is not stated explicitly in [KLLP17], but follows immediately with the proof of Lemma 4.3 in [KLLP17].

If $\|\nabla_{\Gamma_h[\mathbf{y}]} e_h^0\|_{L^\infty(\Gamma_h[\mathbf{y}])} \leq \tfrac{1}{4}$, using the lemma for $w_h^\theta = e_h^\theta$ shows that

$$\|\nabla_{\Gamma_h^\theta} e_h^\theta\|_{L^\infty(\Gamma_h^\theta)} \leq \tfrac{1}{2}, \qquad 0 \leq \theta \leq 1, \qquad (5.6)$$

and then the lemma with $p = 2$ and the definition of the norms (5.1) and (5.2) (and interchanging the roles of $\mathbf{y}$ and $\mathbf{y} + \theta\mathbf{e}$) show that

> the norms $\|\cdot\|_{\mathbf{M}(\mathbf{y}+\theta\mathbf{e})}$ are h-uniformly equivalent for $0 \leq \theta \leq 1$, and so are the norms $\|\cdot\|_{\mathbf{A}(\mathbf{y}+\theta\mathbf{e})}$. (5.7)

Under the condition that $\varepsilon := \|\nabla_{\Gamma_h[\mathbf{y}]} e_h^0\|_{L^\infty(\Gamma_h[\mathbf{y}])} \leq \frac{1}{4}$, using (5.6) in Lemma 5.1 and applying the Cauchy-Schwarz inequality yields the bounds, with $c = c_\infty c_2^2$,

$$\mathbf{w}^T(\mathbf{M}(\mathbf{x}) - \mathbf{M}(\mathbf{y}))\mathbf{z} \leq c\varepsilon \|\mathbf{w}\|_{\mathbf{M}(\mathbf{y})}\|\mathbf{z}\|_{\mathbf{M}(\mathbf{y})},$$
$$\mathbf{w}^T(\mathbf{A}(\mathbf{x}) - \mathbf{A}(\mathbf{y}))\mathbf{z} \leq c\varepsilon \|\mathbf{w}\|_{\mathbf{A}(\mathbf{y})}\|\mathbf{z}\|_{\mathbf{A}(\mathbf{y})}. \tag{5.8}$$

We will also use similar bounds where we use the L^∞ norm of z_h or its gradient and the L^2 norm of the gradient of e_h:

$$\mathbf{w}^T(\mathbf{M}(\mathbf{x}) - \mathbf{M}(\mathbf{y}))\mathbf{z} \leq c \|\mathbf{w}\|_{\mathbf{M}(\mathbf{y})}\|\mathbf{e}\|_{\mathbf{A}(\mathbf{y})},$$
$$\mathbf{w}^T(\mathbf{A}(\mathbf{x}) - \mathbf{A}(\mathbf{y}))\mathbf{z} \leq c \|\mathbf{w}\|_{\mathbf{A}(\mathbf{y})}\|\mathbf{e}\|_{\mathbf{A}(\mathbf{y})}. \tag{5.9}$$

Consider now a continuously differentiable function $\mathbf{x} : [0, T] \to \mathbb{R}^{3N}$ that defines a finite element surface $\Gamma_h[\mathbf{x}(t)]$ for every $t \in [0, T]$, and assume that its time derivative $\mathbf{v}(t) = \dot{\mathbf{x}}(t)$ is the nodal vector of a finite element function $v_h(\cdot, t)$ that satisfies

$$\|\nabla_{\Gamma_h[\mathbf{x}(t)]} v_h(\cdot, t)\|_{L^\infty(\Gamma_h[\mathbf{x}(t)])} \leq K, \qquad 0 \leq t \leq T. \tag{5.10}$$

With $\mathbf{e} = \mathbf{x}(t) - \mathbf{x}(s) = \int_s^t \mathbf{v}(r)\,dr$, the bounds (5.8) then yield the following bounds, which were first shown in Lemma 4.1 of [DLM12]:
for $0 \leq s, t \leq T$ with $K|t - s| \leq \frac{1}{4}$, we have with $C = cK$

$$\mathbf{w}^T\big(\mathbf{M}(\mathbf{x}(t)) - \mathbf{M}(\mathbf{x}(s))\big)\mathbf{z} \leq C\,|t - s|\,\|\mathbf{w}\|_{\mathbf{M}(\mathbf{x}(t))}\|\mathbf{z}\|_{\mathbf{M}(\mathbf{x}(t))},$$
$$\mathbf{w}^T\big(\mathbf{A}(\mathbf{x}(t)) - \mathbf{A}(\mathbf{x}(s))\big)\mathbf{z} \leq C\,|t - s|\,\|\mathbf{w}\|_{\mathbf{A}(\mathbf{x}(t))}\|\mathbf{z}\|_{\mathbf{A}(\mathbf{x}(t))}. \tag{5.11}$$

Letting $s \to t$, this implies the bounds stated in Lemma 4.6 of [KLLP17]:

$$\mathbf{w}^T \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{M}(\mathbf{x}(t))\mathbf{z} \leq C \|\mathbf{w}\|_{\mathbf{M}(\mathbf{x}(t))}\|\mathbf{z}\|_{\mathbf{M}(\mathbf{x}(t))},$$
$$\mathbf{w}^T \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{A}(\mathbf{x}(t))\mathbf{z} \leq C \|\mathbf{w}\|_{\mathbf{A}(\mathbf{x}(t))}\|\mathbf{z}\|_{\mathbf{A}(\mathbf{x}(t))}. \tag{5.12}$$

Moreover, by patching together finitely many intervals over which $K|t-s| \leq \frac{1}{4}$, we obtain that

> the norms $\|\cdot\|_{\mathbf{M}(\mathbf{x}(t))}$ are h-uniformly equivalent for $0 \leq t \leq T$, and so are the norms $\|\cdot\|_{\mathbf{A}(\mathbf{x}(t))}$. (5.13)

5.2 Preparation: Interpolation of products of finite element functions

For the stability of the velocity approximation we will need the following result.

Lemma 5.3 *For an admissible triangulation of a smooth surface Γ, let Γ_h^* be the interpolated surface with finite elements of polynomial degree $k \geq 1$. Let $\widetilde{I}_h^* : C(\Gamma_h^*) \to S_h(\Gamma_h^*)$ denote the finite element interpolation operator on Γ_h^*. Then, the interpolation of the product of two finite element functions a_h, b_h on Γ_h^* is bounded by*

$$\|\widetilde{I}_h^*(a_h b_h)\|_{H^1(\Gamma_h^*)} \leq C \|a_h\|_{H^1(\Gamma_h^*)} \|b_h\|_{W^{1,\infty}(\Gamma_h^*)},$$

where C depends only on Γ (more precisely, on bounds of higher derivatives of a parametrization of Γ), on shape-regularity and quasi-uniformity of the triangulation, and on the degree k.

Proof Let K be a curved triangle of the triangulation Γ_h^*. Using the interpolation error estimate of [Dem09, Proposition 2.7], on the element K we obtain (with different constants c)

$$\|\widetilde{I}_h^*(a_h b_h) - a_h b_h\|_{H^1(K)} \leq ch^k\|a_h b_h\|_{H^{k+1}(K)} \leq ch^k \sum_{j=0}^{k+1} |a_h b_h|_{H^j(K)}$$

$$\leq ch^k \sum_{j=0}^{k+1} \sum_{i=0}^{j} \|\nabla_{\Gamma_h}^i a_h\|_{L^2(K)} \|\nabla_{\Gamma_h}^{j-i} b_h\|_{L^\infty(K)}$$

$$= c \sum_{j=0}^{k} h^{k-j} \sum_{i=0}^{j} h^i \|\nabla_{\Gamma_h}^i a_h\|_{L^2(K)} \, h^{j-i} \|\nabla_{\Gamma_h}^{j-i} b_h\|_{L^\infty(K)}$$

$$+ c \sum_{i=0}^{k} h^i \|\nabla_{\Gamma_h}^i a_h\|_{L^2(K)} \, h^{k-i} \|\nabla_{\Gamma_h}^{k+1-i} b_h\|_{L^\infty(K)} + ch^k \|\nabla_{\Gamma_h}^{k+1} a_h\|_{L^2(K)} \|b_h\|_{L^\infty(K)}.$$

By inverse estimates, this is further bounded by

$$\begin{aligned}
\|\widetilde{I}_h^*(a_h b_h) - a_h b_h\|_{H^1(K)} &\leq c\|a_h\|_{L^2(K)} \|b_h\|_{L^\infty(K)} \\
&\quad + c\|a_h\|_{L^2(K)} \|\nabla_{\Gamma_h} b_h\|_{L^\infty(K)} + c\|\nabla_{\Gamma_h} a_h\|_{L^2(K)} \|b_h\|_{L^\infty(K)} \\
&\leq c \, \|a_h\|_{H^1(K)} \|b_h\|_{W^{1,\infty}(K)}
\end{aligned}$$

Squaring and summing up over the triangles then shows that

$$\|\widetilde{I}_h^*(a_h b_h) - a_h b_h\|_{H^1(\Gamma_h^*)} \leq c\|a_h\|_{H^1(\Gamma_h^*)} \|b_h\|_{W^{1,\infty}(\Gamma_h^*)}.$$

The stated result then follows with the triangle inequality. $\qquad \square$

5.3 Defects and errors

We choose reference finite element functions $x_h^*(\cdot, t)$, $v_h^*(\cdot, t)$, $u_h^*(\cdot, t)$, $w_h^*(\cdot, t)$ on the interpolated surface $\Gamma_h[\mathbf{x}^*(t)]$ with nodal vectors

$$\mathbf{x}^*(t) \in \mathbb{R}^{3N}, \quad \mathbf{v}^*(t) \in \mathbb{R}^{3N},$$

$$\mathbf{u}^*(t) = \begin{pmatrix} \mathbf{H}^*(t) \\ \mathbf{n}^*(t) \end{pmatrix} \in \mathbb{R}^{4N}, \quad \mathbf{w}^*(t) = \begin{pmatrix} \mathbf{V}^*(t) \\ \mathbf{z}^*(t) \end{pmatrix} \in \mathbb{R}^{4N},$$

which are related to the exact solution X, v and $u = (H, \nu)$, $w = (V, z)$ as follows: $\mathbf{x}^*(t)$ and $\mathbf{v}^*(t)$ collect the values at the finite element nodes of $X(\cdot, t)$ and $v(\cdot, t)$, respectively. The vectors $\mathbf{u}^*(t)$ and $\mathbf{w}^*(t)$ need to be chosen in a different way. The vector $\mathbf{u}^*(t)$ contains the nodal values of the finite element function $u_h^*(\cdot, t) \in S_h[\mathbf{x}^*(t)]^4$ that is determined on the interpolated surface $\Gamma_h[\mathbf{x}^*(t)]$ by a *modified* Ritz map of the exact solution component u that will be defined in Section 6. Similarly, the vector $\mathbf{w}^*(t)$ contains the nodal values of the finite element function $w_h^*(\cdot, t) \in S_h[\mathbf{x}^*(t)]^4$ that is defined on the interpolated surface $\Gamma_h[\mathbf{x}^*(t)]$ by the Ritz map of the solution component w.

The nodal vectors $\mathbf{x}^*(t)$, $\mathbf{v}^*(t)$ then satisfy the equations (3.4) up to some defect $\mathbf{d_v}$ that will be studied in Section 6,

$$\dot{\mathbf{x}}^* = \mathbf{v}^*, \tag{5.14a}$$

$$\mathbf{v}^* = \mathbf{V}^* \bullet \mathbf{n}^* + \mathbf{d_v}, \tag{5.14b}$$

and $\mathbf{u}^*(t)$, $\mathbf{w}^*(t)$ satisfy the equations (3.5) up to some defects $\mathbf{d_u}$ and $\mathbf{d_w}$ that will also be studied in Section 6:

$$\mathbf{M}^{[4]}(\mathbf{x}^*)\dot{\mathbf{u}}^* - \mathbf{A}^{[4]}(\mathbf{x}^*)\mathbf{w}^* = \mathbf{F}(\mathbf{x}^*, \mathbf{u}^*)\mathbf{w}^* + \mathbf{f}(\mathbf{x}^*, \mathbf{u}^*) + \mathbf{M}^{[4]}(\mathbf{x}^*)\mathbf{d_u}, \tag{5.15a}$$

$$\mathbf{M}^{[4]}(\mathbf{x}^*)\mathbf{w}^* + \mathbf{A}^{[4]}(\mathbf{x}^*)\mathbf{u}^* = \mathbf{g}(\mathbf{x}^*, \mathbf{u}^*) + \mathbf{M}^{[4]}(\mathbf{x}^*)\mathbf{d_w}. \tag{5.15b}$$

In the following we omit the superscript [4] on $\mathbf{M}$ and $\mathbf{A}$. Furthermore, we simplify the notation and abbreviate $\mathbf{M}(\mathbf{x}(t))$ and $\mathbf{M}(\mathbf{x}^*(t))$ to $\mathbf{M}(t)$ and $\mathbf{M}^*(t)$, respectively. Similarly we write $\mathbf{A}$ and $\mathbf{A}^*$.

The errors between the nodal values of the numerical solutions and the nodal values of the interpolated exact values are denoted by $\mathbf{e_x} = \mathbf{x} - \mathbf{x}^*$, $\mathbf{e_v} = \mathbf{v} - \mathbf{v}^*$, $\mathbf{e_u} = \mathbf{u} - \mathbf{u}^*$ and $\mathbf{e_w} = \mathbf{w} - \mathbf{w}^*$ and their corresponding finite element functions on the interpolated surface $\Gamma_h[\mathbf{x}^*]$ are denoted by e_x, e_v, e_u, and e_w, respectively.

We obtain the error equations by subtracting (5.15) from (3.5) with (3.7) instead of (3.5b), and (5.14) from (3.4):

$$\dot{\mathbf{e}}_{\mathbf{x}} = \mathbf{e}_{\mathbf{v}}, \tag{5.16a}$$

$$\mathbf{e}_{\mathbf{v}} = \mathbf{V} \bullet \mathbf{n} - \mathbf{V}^* \bullet \mathbf{n}^* - \mathbf{d}_{\mathbf{v}}, \tag{5.16b}$$

$$\begin{aligned}
\mathbf{M}\dot{\mathbf{e}}_{\mathbf{u}} - \mathbf{A}\mathbf{e}_{\mathbf{w}} = & - \big(\mathbf{M} - \mathbf{M}^*\big)\dot{\mathbf{u}}^* + \big(\mathbf{A} - \mathbf{A}^*\big)\mathbf{w}^* \\
& + \big(\mathbf{F}(\mathbf{x}, \mathbf{u})\mathbf{w} - \mathbf{F}(\mathbf{x}^*, \mathbf{u}^*)\mathbf{w}^*\big) \\
& + \big(\mathbf{f}(\mathbf{x}, \mathbf{u}) - \mathbf{f}(\mathbf{x}^*, \mathbf{u}^*)\big) - \mathbf{M}^*\mathbf{d}_{\mathbf{u}},
\end{aligned} \tag{5.16c}$$

$$\begin{aligned}
\mathbf{M}\mathbf{e}_{\mathbf{w}} + \mathbf{A}\mathbf{e}_{\mathbf{u}} = & - \big(\mathbf{M} - \mathbf{M}^*\big)\mathbf{w}^* - \big(\mathbf{A} - \mathbf{A}^*\big)\mathbf{u}^* \\
& + \big(\mathbf{g}(\mathbf{x}, \mathbf{u}) - \mathbf{g}(\mathbf{x}^*, \mathbf{u}^*)\big) - \mathbf{M}^*\mathbf{d}_{\mathbf{w}} + \boldsymbol{\vartheta}.
\end{aligned} \tag{5.16d}$$

We note further that $\mathbf{e}_{\mathbf{x}}(0) = 0$, but in general $\mathbf{e}_{\mathbf{v}}(0)$, $\mathbf{e}_{\mathbf{u}}(0)$ and $\mathbf{e}_{\mathbf{w}}(0)$ are different from 0. We recall from (3.6) that

$$\boldsymbol{\vartheta} = \mathbf{M}(\mathbf{x}(0))\,\bar{\mathbf{e}}_{\mathbf{w}}(0) \quad \text{with} \quad \bar{\mathbf{e}}_{\mathbf{w}}(0) = \bar{\mathbf{w}}^*(0) - \bar{\mathbf{w}}(0).$$

This is not to be confounded with $\mathbf{e}_{\mathbf{w}}(0) = \bar{\mathbf{w}}^*(0) - \mathbf{w}^*(0)$, where we remind that $\bar{\mathbf{w}}^*(0)$ contains the values of the exact solution w in the nodes, whereas $\mathbf{w}^*(0)$ is constructed by a Ritz map from the exact solution w at the initial time 0.

5.4 Stability estimate

We need to bound the errors at time t in terms of the defects up to time t and the errors in the initial values. The errors will be estimated in the H^1 norm on the interpolated surface $\Gamma_h[\mathbf{x}^*]$: for a nodal vector $\mathbf{e}$ corresponding to a finite element function $e \in S_h(\mathbf{x}^*)$, we have with the matrix $\mathbf{K}(\mathbf{x}^*) = \mathbf{M}(\mathbf{x}^*) + \mathbf{A}(\mathbf{x}^*)$ that

$$\|\mathbf{e}\|_{\mathbf{K}(\mathbf{x}^*)}^2 = \mathbf{e}^T \mathbf{K}(\mathbf{x}^*)\mathbf{e} = \|e\|_{H^1(\Gamma_h[\mathbf{x}^*])}^2.$$

The defect $\mathbf{d}_{\mathbf{v}}$ then needs to be sufficiently small in the H^1 norm, and the other defects will be required to be small in the L^2 norm

$$\|\mathbf{d}\|_{\mathbf{M}(\mathbf{x}^*)}^2 = \mathbf{d}^T \mathbf{M}(\mathbf{x}^*)\mathbf{d} = \|d\|_{L^2(\Gamma_h[\mathbf{x}^*])}^2,$$

and their time derivatives as well as $\bar{\mathbf{e}}_{\mathbf{w}}(0)$ will be required to be small in the norm given by

$$\|\mathbf{d}\|_{\star, \mathbf{x}^*}^2 := \mathbf{d}^T \mathbf{M}(\mathbf{x}^*)\mathbf{K}(\mathbf{x}^*)^{-1}\mathbf{M}(\mathbf{x}^*)\mathbf{d}.$$

By [KLLP17, Eq. (5.5)], this equals the following dual norm for the corresponding finite element function $d \in S_h[\mathbf{x}^*]$, which has the vector of nodal values $\mathbf{d}$,

$$\|\mathbf{d}\|_{\star, \mathbf{x}^*} = \|d\|_{H_h^{-1}(\Gamma_h[\mathbf{x}^*])} := \sup_{0 \neq \varphi_h \in S_h[\mathbf{x}^*]} \frac{\int_{\Gamma_h[\mathbf{x}^*]} d \cdot \varphi_h}{\|\varphi_h\|_{H^1(\Gamma_h[\mathbf{x}^*])}}.$$

The following result provides the key stability estimate, which bounds the errors in terms of the defects and the initial errors.

Proposition 5.1 *Assume that the reference finite element functions $x_h^*(\cdot, t)$, $v_h^*(\cdot, t)$, $u_h^*(\cdot, t)$, $w_h^*(\cdot, t)$ on the interpolated surface $\Gamma_h[\mathbf{x}^*(t)]$ have $W^{1,\infty}$ norms that are bounded independently of h, for all $t \in [0, T]$. Assume that, for some κ with $1 < \kappa \le k$, the defects are bounded by*

$$\|\mathbf{d_v}(t)\|_{\mathbf{K}(\mathbf{x}^*(t))} \le ch^\kappa,$$

$$\|\mathbf{d_u}(t)\|_{\mathbf{M}(\mathbf{x}^*(t))} + \|\dot{\mathbf{d}}_{\mathbf{u}}(t)\|_{\star, \mathbf{x}^*(t)} \le ch^\kappa, \qquad for \ \ 0 \le t \le T, \qquad (5.17)$$

$$\|\mathbf{d_w}(t)\|_{\mathbf{M}(\mathbf{x}^*(t))} + \|\dot{\mathbf{d}}_{\mathbf{w}}(t)\|_{\star, \mathbf{x}^*(t)} \le ch^\kappa,$$

and that the errors of the initial values are bounded by

$$\|\mathbf{e_u}(0)\|_{\mathbf{K}(\mathbf{x}^0)} + \|\mathbf{e_w}(0)\|_{\mathbf{K}(\mathbf{x}^0)} \le ch^\kappa,$$
$$\|\bar{\mathbf{e}}_{\mathbf{w}}(0)\|_{\star, \mathbf{x}^0} \quad \le ch^\kappa. \qquad (5.18)$$

Then, there exists $h_0 > 0$ such that the following stability estimate holds for all $h \le h_0$ and $0 \le t \le T$:

$$\|\mathbf{e_x}(t)\|_{\mathbf{K}(\mathbf{x}^*(t))}^2 + \|\mathbf{e_v}(t)\|_{\mathbf{K}(\mathbf{x}^*(t))}^2 + \|\mathbf{e_u}(t)\|_{\mathbf{K}(\mathbf{x}^*(t))}^2 + \|\mathbf{e_w}(t)\|_{\mathbf{K}(\mathbf{x}^*(t))}^2$$

$$\le C \max_{0 \le s \le t} \left(\|\mathbf{d_v}(s)\|_{\mathbf{K}(\mathbf{x}^*(s))}^2 + \|\mathbf{d_u}(s)\|_{\star, \mathbf{x}^*(s)}^2 \right)$$

$$+ C \int_0^t \left(\|\mathbf{d_u}(s)\|_{\mathbf{M}(\mathbf{x}^*(s))}^2 + \|\dot{\mathbf{d}}_{\mathbf{u}}(s)\|_{\star, \mathbf{x}^*(s)}^2 \right. \qquad (5.19)$$

$$\left. + \|\mathbf{d_w}(s)\|_{\mathbf{M}(\mathbf{x}^*(s))}^2 + \|\dot{\mathbf{d}}_{\mathbf{w}}(s)\|_{\star, \mathbf{x}^*(s)}^2 \right) \mathrm{d}s$$

$$+ C \left(\|\mathbf{e_u}(0)\|_{\mathbf{K}(\mathbf{x}^0)}^2 + \|\mathbf{e_w}(0)\|_{\mathbf{K}(\mathbf{x}^0)}^2 \right) + C \|\bar{\mathbf{e}}_{\mathbf{w}}(0)\|_{\star, \mathbf{x}^0}^2,$$

where C is independent of h and t, but depends on the final time T.

Proof The proof uses energy estimates for the error equations in the matrix–vector formulation (5.16) and relies on the lemmas of Section 5.1, which relate the different finite element surfaces. While this basic procedure of the proof looks similar to that of [KLL19] and [KLLP17], there are substantial differences and technical difficulties that are peculiar to the fourth-order system.

We exploit the skew-symmetric structure of the system of error equations (5.16c)–(5.16d). The uniform-in-time stability estimate follows from the combination of four different sets of auxiliary energy estimates (denoted by (i)–(iv)), divided into two major parts, Parts (A.1) and (A.2), and finally combining them in Part (A.3), which gives the uniform-in-time stability bound for the errors $\mathbf{e_u}$ and $\mathbf{e_w}$, as illustrated in Figure 5.2. Part (B) of the proof contains the estimates for the velocity error equation (5.16b), and finally the two are combined in Part (C), to show the stability bound (5.19).

Let $t^* \in (0, T]$ be the maximal time such that the following inequalities hold:

$$\|e_x(\cdot, t)\|_{W^{1,\infty}(\Gamma_h[\mathbf{x}^*(t)])} \le h^{(\kappa-1)/2},$$
$$\|e_v(\cdot, t)\|_{W^{1,\infty}(\Gamma_h[\mathbf{x}^*(t)])} \le h^{(\kappa-1)/2},$$
$$\|e_u(\cdot, t)\|_{W^{1,\infty}(\Gamma_h[\mathbf{x}^*(t)])} \le h^{(\kappa-1)/2}, \qquad for \quad t \in [0, t^*]. \qquad (5.20)$$
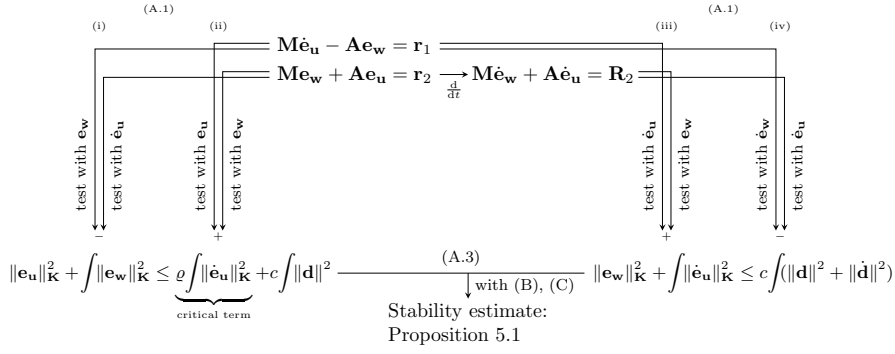$$\|e_w(\cdot, t)\|_{W^{1,\infty}(\Gamma_h[\mathbf{x}^*(t)])} \le h^{(\kappa-1)/2},$$

Fig. 5.2 Sketch of the structure of the energy estimates of Part (A) for the stability proof. In the diagram, $\mathbf{r}_1$ and $\mathbf{r}_2$ denote the right-hand sides of (5.16c) and (5.16d), respectively. (Note that, after time differentiation, $\mathbf{R}_2$ contains terms other than only the time derivative of $\mathbf{r}_2$.)

Note that $t^* > 0$ since initially $e_x(\cdot, 0) = 0$ and, by an inverse inequality, we have

$$\|e_u(\cdot, 0)\|_{W^{1,\infty}(\Gamma_h[\mathbf{x}^*(0)])} \leq ch^{-1}\|e_u(\cdot, 0)\|_{H^1(\Gamma_h[\mathbf{x}^*(0)])}$$
$$= ch^{-1}\|\mathbf{e_u}(0)\|_{\mathbf{K}(\mathbf{x}^*(t))} \leq Ch^{\kappa-1},$$

where the last inequality holds by assumption (5.18). By the same argument, also $\|e_w(\cdot, 0)\|_{W^{1,\infty}(\Gamma_h[\mathbf{x}^*(0)])} \leq Ch^{\kappa-1}$ and, as a result of $v_h = \widetilde{I}_h(V_h \nu_h)$,

$$\|e_v(\cdot, 0)\|_{W^{1,\infty}(\Gamma_h[\mathbf{x}^*(0)])}$$
$$\leq c\|e_u(\cdot, 0)\|_{W^{1,\infty}(\Gamma_h[\mathbf{x}^*(0)])} + c\|e_w(\cdot, 0)\|_{W^{1,\infty}(\Gamma_h[\mathbf{x}^*(0)])} \leq Ch^{\kappa-1}.$$

By continuity, we then obtain for sufficiently small h that $t^* > 0$ (which *a priori* might depend on h). We first prove the stated error bounds for $0 \leq t \leq t^*$. At the end of the proof we will show that in fact t^* coincides with T.

Since the reference finite element functions $x_h^*(\cdot, t), v_h^*(\cdot, t), u_h^*(\cdot, t), w_h^*(\cdot, t)$ on the interpolated surface $\Gamma_h[\mathbf{x}^*(t)]$ have $W^{1,\infty}$ norms that are bounded independently of h for all $t \in [0, T]$, the bounds (5.20) together with Lemma 5.2 imply that the $W^{1,\infty}$ norms of the ESFEM functions $x_h(\cdot, t), v_h(\cdot, t), u_h(\cdot, t), w_h(\cdot, t)$ on the discrete surface $\Gamma_h[\mathbf{x}(t)]$ are also bounded independently of h and $t \in [0, t^*]$, and so are their lifts to the interpolated surface $\Gamma_h[\mathbf{x}^*(t)]$. In particular, it will be important that the discrete velocity v_h with nodal vector $\mathbf{v}$ satisfies (5.10).

The estimate on the position errors e_x in (5.20) and the $W^{1,\infty}$ bound on v_h immediately imply that the results of Section 5.1 apply with $\mathbf{x}$ and $\mathbf{x}^*$ in the roles of $\mathbf{x}$ and $\mathbf{y}$, respectively. In particular, due to the bounds in (5.20) (for a sufficiently small $h \leq h_0$), the main condition of Lemma 5.2 and also (5.6) is satisfied (with $e_h^\theta = e_x^\theta$), hence the h-uniform norm equivalences in (5.7) and the estimates in (5.8) and (5.9) hold between the surfaces defined by $\mathbf{x}$

and $\mathbf{x}^*$. Similarly, again due to (5.20) the bound (5.10) also holds, and hence the estimates in (5.11), (5.12) and the h-uniform norm equivalences in time (5.13) also hold. When referring to these results from Section 5.1 ((5.7)–(5.9), (5.11)–(5.13)) within the stability proof below, following the above argument, we always mean that their respective smallness assumptions are satisfied via (5.20), but we will not repeat this argument at each instance.

In the following c and C will denote generic constants that might take different values on different occurrences. In contrast, constants with a subscript (such as c_0) will play a distinctive role in the proof, and will not change their value between appearances.

(A) *Estimates for the surface PDEs:*

(A.1) We start by showing an error estimate for $\mathbf{e_u}$ in the $\mathbf{K}$-norm uniformly in time for $0 \leq t \leq t^*$. A critical term will appear, which will be controlled in Part (A.2) and eliminated in (A.3).

Energy estimate (i): We test (5.16c) with $\mathbf{e_w}$, and test (5.16d) with $\dot{\mathbf{e}}_\mathbf{u}$, so that we obtain the two equations

$$
\begin{aligned}
\mathbf{e_w}^T \mathbf{M} \dot{\mathbf{e}}_\mathbf{u} - \mathbf{e_w}^T \mathbf{A} \mathbf{e_w} = \; & -\mathbf{e_w}^T (\mathbf{M} - \mathbf{M}^*) \dot{\mathbf{u}}^* + \mathbf{e_w}^T (\mathbf{A} - \mathbf{A}^*) \mathbf{w}^* \\
& + \mathbf{e_w}^T \big(\mathbf{F}(\mathbf{x}, \mathbf{u}) \mathbf{w} - \mathbf{F}(\mathbf{x}^*, \mathbf{u}^*) \mathbf{w}^* \big) \\
& + \mathbf{e_w}^T \big(\mathbf{f}(\mathbf{x}, \mathbf{u}) - \mathbf{f}(\mathbf{x}^*, \mathbf{u}^*) \big) - \mathbf{e_w}^T \mathbf{M}^* \mathbf{d_u}, \\
\dot{\mathbf{e}}_\mathbf{u}^T \mathbf{M} \mathbf{e_w} + \dot{\mathbf{e}}_\mathbf{u}^T \mathbf{A} \mathbf{e_u} = \; & -\dot{\mathbf{e}}_\mathbf{u}^T (\mathbf{M} - \mathbf{M}^*) \mathbf{w}^* - \dot{\mathbf{e}}_\mathbf{u}^T (\mathbf{A} - \mathbf{A}^*) \mathbf{u}^* \\
& + \dot{\mathbf{e}}_\mathbf{u}^T \big(\mathbf{g}(\mathbf{x}, \mathbf{u}) - \mathbf{g}(\mathbf{x}^*, \mathbf{u}^*) \big) - \dot{\mathbf{e}}_\mathbf{u}^T \mathbf{M}^* \mathbf{d_w} + \dot{\mathbf{e}}_\mathbf{u}^T \boldsymbol{\vartheta},
\end{aligned}
$$

In order to eliminate the mixed term $\mathbf{e_w}^T \mathbf{M} \dot{\mathbf{e}}_\mathbf{u}$ (using the symmetry of the mass matrix $\mathbf{M}$), we then subtract the former equation from the latter, and obtain

$$
\begin{aligned}
\|\mathbf{e_w}\|_\mathbf{A}^2 + \dot{\mathbf{e}}_\mathbf{u}^T \mathbf{A} \mathbf{e_u} = \; & + \mathbf{e_w}^T (\mathbf{M} - \mathbf{M}^*) \dot{\mathbf{u}}^* - \mathbf{e_w}^T (\mathbf{A} - \mathbf{A}^*) \mathbf{w}^* \\
& - \mathbf{e_w}^T \big(\mathbf{F}(\mathbf{x}, \mathbf{u}) \mathbf{w} - \mathbf{F}(\mathbf{x}^*, \mathbf{u}^*) \mathbf{w}^* \big) \\
& - \mathbf{e_w}^T \big(\mathbf{f}(\mathbf{x}, \mathbf{u}) - \mathbf{f}(\mathbf{x}^*, \mathbf{u}^*) \big) + \mathbf{e_w}^T \mathbf{M}^* \mathbf{d_u} \\
& - \dot{\mathbf{e}}_\mathbf{u}^T (\mathbf{M} - \mathbf{M}^*) \mathbf{w}^* - \dot{\mathbf{e}}_\mathbf{u}^T (\mathbf{A} - \mathbf{A}^*) \mathbf{u}^* \\
& + \dot{\mathbf{e}}_\mathbf{u}^T \big(\mathbf{g}(\mathbf{x}, \mathbf{u}) - \mathbf{g}(\mathbf{x}^*, \mathbf{u}^*) \big) - \dot{\mathbf{e}}_\mathbf{u}^T \mathbf{M}^* \mathbf{d_w} + \dot{\mathbf{e}}_\mathbf{u}^T \boldsymbol{\vartheta}.
\end{aligned} \tag{5.21}
$$

The terms in equation (5.21) will now be bounded separately, using the results of Section 5.1.

For the second term on the left-hand side of (5.21), using the product rule, the symmetry of $\mathbf{A}$ and the bound (5.12), we have

$$
\begin{aligned}
\dot{\mathbf{e}}_\mathbf{u}^T \mathbf{A} \mathbf{e_u} &= \frac{1}{2} \frac{\mathrm{d}}{\mathrm{d}t} \big(\mathbf{e_u}^T \mathbf{A} \mathbf{e_u} \big) - \frac{1}{2} \mathbf{e_u}^T \dot{\mathbf{A}} \mathbf{e_u} \\
&\geq \frac{1}{2} \frac{\mathrm{d}}{\mathrm{d}t} \|\mathbf{e_u}\|_\mathbf{A}^2 - c \|\mathbf{e_u}\|_\mathbf{A}^2.
\end{aligned} \tag{5.22}
$$

On the right-hand side, for the matrix difference terms in the first and fourth line of (5.21) we use (5.9), and, recalling that $\mathbf{M}^* + \mathbf{A}^* = \mathbf{K}^*$, we

obtain

$$
\begin{aligned}
\mathbf{e_w}^T\big(\mathbf{M} - \mathbf{M}^*\big)\dot{\mathbf{u}}^* - \mathbf{e_w}^T\big(\mathbf{A} - \mathbf{A}^*\big)\mathbf{w}^* &\leq c\|\mathbf{e_w}\|_{\mathbf{K}}\|\mathbf{e_x}\|_{\mathbf{K}}, \\
-\dot{\mathbf{e}}_{\mathbf{u}}^T\big(\mathbf{M} - \mathbf{M}^*\big)\mathbf{w}^* - \dot{\mathbf{e}}_{\mathbf{u}}^T\big(\mathbf{A} - \mathbf{A}^*\big)\mathbf{u}^* &\leq c\|\dot{\mathbf{e}}_{\mathbf{u}}\|_{\mathbf{K}}\|\mathbf{e_x}\|_{\mathbf{K}}.
\end{aligned}
\tag{5.23}
$$

For the defect terms in (5.21), we obtain

$$
\begin{aligned}
\mathbf{e_w}^T\mathbf{M}^*\mathbf{d_u} &= \int_{\Gamma_h[\mathbf{x}^*]} e_w\, d_u \\
&\leq \|e_w\|_{H^1(\Gamma_h[\mathbf{x}^*])}\|d_u\|_{H_h^{-1}(\Gamma_h[\mathbf{x}^*])} = \|\mathbf{e_w}\|_{\mathbf{K}^*}\|\mathbf{d_u}\|_{\star,\mathbf{x}^*},
\end{aligned}
$$

and in the same way

$$
-\dot{\mathbf{e}}_{\mathbf{u}}^T\mathbf{M}^*\mathbf{d_w} \leq \|\dot{\mathbf{e}}_{\mathbf{u}}\|_{\mathbf{K}^*}\|\mathbf{d_w}\|_{\star,\mathbf{x}^*},
$$

and similarly, using the norm equivalence (5.13),

$$
\begin{aligned}
\dot{\mathbf{e}}_{\mathbf{u}}(t)^T\boldsymbol{\vartheta} &= \dot{\mathbf{e}}_{\mathbf{u}}(t)^T\mathbf{M}(\mathbf{x}^0)\bar{\mathbf{e}}_{\mathbf{w}}^0 \\
&\leq \|\dot{\mathbf{e}}_{\mathbf{u}}(t)\|_{\mathbf{K}(\mathbf{x}^0)}\|\bar{\mathbf{e}}_{\mathbf{w}}^0\|_{\star,\mathbf{x}^0} \leq c\|\dot{\mathbf{e}}_{\mathbf{u}}(t)\|_{\mathbf{K}(\mathbf{x}(t))}\|\bar{\mathbf{e}}_{\mathbf{w}}^0\|_{\star,\mathbf{x}^0}.
\end{aligned}
$$

We thus obtain the following bounds for the defect terms:

$$
\begin{aligned}
\mathbf{e_w}^T\mathbf{M}^*\mathbf{d_u} &\leq c\|\mathbf{e_w}\|_{\mathbf{K}}\|\mathbf{d_u}\|_{\star,\mathbf{x}^*}, \\
-\dot{\mathbf{e}}_{\mathbf{u}}^T\mathbf{M}^*\mathbf{d_w} &\leq c\|\dot{\mathbf{e}}_{\mathbf{u}}\|_{\mathbf{K}}\|\mathbf{d_w}\|_{\star,\mathbf{x}^*}, \\
\dot{\mathbf{e}}_{\mathbf{u}}^T\boldsymbol{\vartheta} &\leq c\|\dot{\mathbf{e}}_{\mathbf{u}}\|_{\mathbf{K}}\|\bar{\mathbf{e}}_{\mathbf{w}}^0\|_{\star,\mathbf{x}^0}.
\end{aligned}
\tag{5.24}
$$

For the non-linear terms we use the following bounds: The two terms involving $\mathbf{f}$ and $\mathbf{g}$ are bounded exactly as the (general, locally Lipschitz) non-linear term in Part (A.v) in the proof of Proposition 7.1 in [KLL19] (using the fact that by the $W^{1,\infty}$ bound for the exact solution u_h^* and the error e_u in (5.20), the numerical solution u_h is also bounded in the $W^{1,\infty}$ norm) and the norm equivalence (5.7). We altogether have

$$
\begin{aligned}
\mathbf{e_w}^T\big(\mathbf{f}(\mathbf{x},\mathbf{u}) - \mathbf{f}(\mathbf{x}^*,\mathbf{u}^*)\big) &\leq c\|\mathbf{e_w}\|_{\mathbf{K}}\big(\|\mathbf{e_x}\|_{\mathbf{K}} + \|\mathbf{e_u}\|_{\mathbf{K}}\big), \\
\dot{\mathbf{e}}_{\mathbf{u}}^T\big(\mathbf{g}(\mathbf{x},\mathbf{u}) - \mathbf{g}(\mathbf{x}^*,\mathbf{u}^*)\big) &\leq c\|\dot{\mathbf{e}}_{\mathbf{u}}\|_{\mathbf{M}}\big(\|\mathbf{e_x}\|_{\mathbf{K}} + \|\mathbf{e_u}\|_{\mathbf{K}}\big).
\end{aligned}
\tag{5.25}
$$

For the term involving the state-dependent mass matrix $\mathbf{F}(\mathbf{x},\mathbf{u})$ we estimate by

$$
\begin{aligned}
&- \mathbf{e_w}^T\big(\mathbf{F}(\mathbf{x},\mathbf{u})\mathbf{w} - \mathbf{F}(\mathbf{x}^*,\mathbf{u}^*)\mathbf{w}^*\big) \\
&= - \mathbf{e_w}^T\mathbf{F}(\mathbf{x},\mathbf{u})\mathbf{e_w} - \mathbf{e_w}^T\big(\mathbf{F}(\mathbf{x},\mathbf{u}) - \mathbf{F}(\mathbf{x}^*,\mathbf{u}^*)\big)\mathbf{w}^* \\
&\leq c_0\|\mathbf{e_w}\|_{\mathbf{M}}^2 + c\|\mathbf{e_w}\|_{\mathbf{M}}\big(\|\mathbf{e_x}\|_{\mathbf{K}} + \|\mathbf{e_u}\|_{\mathbf{K}}\big),
\end{aligned}
\tag{5.26}
$$

where the first term in the middle line is estimated using (3.3) together with the $W^{1,\infty}$ boundedness of the numerical solutions u_h (due to (5.20)), while the second term is again estimated using the above argument from [KLL19, Proposition 7.1, Part (A.v)], and using the norm equivalence (5.7). Here, c_0 is a positive constant.

Altogether, the combination of the above bounds yields the first energy estimate

$$\|\mathbf{e_w}\|_{\mathbf{A}}^2 + \frac{\mathrm{d}}{\mathrm{d}t}\|\mathbf{e_u}\|_{\mathbf{A}}^2 \leq c\|\mathbf{e_u}\|_{\mathbf{A}}^2 + 2c_0\|\mathbf{e_w}\|_{\mathbf{M}}^2 \tag{5.27}$$
$$+ c\|\mathbf{e_w}\|_{\mathbf{K}}\big(\|\mathbf{e_x}\|_{\mathbf{K}} + \|\mathbf{e_u}\|_{\mathbf{K}} + \|\mathbf{d_u}\|_{\star,\mathbf{x}^*}\big)$$
$$+ c\|\dot{\mathbf{e}}_\mathbf{u}\|_{\mathbf{K}}\big(\|\mathbf{e_x}\|_{\mathbf{K}} + \|\mathbf{e_u}\|_{\mathbf{K}} + \|\mathbf{d_w}\|_{\star,\mathbf{x}^*} + \|\bar{\mathbf{e}}_\mathbf{w}^0\|_{\star,\mathbf{x}^0}\big).$$

Energy estimate (ii): We test (5.16c) with $\mathbf{e_u}$ and (5.16d) with $\mathbf{e_w}$, then sum up to cancel the mixed terms $\mathbf{e_u}^T \mathbf{A} \mathbf{e_w}$, to obtain

$$\mathbf{e_u}^T \mathbf{M}\dot{\mathbf{e}}_\mathbf{u} + \|\mathbf{e_w}\|_{\mathbf{M}}^2 = -\mathbf{e_u}^T(\mathbf{M} - \mathbf{M}^*)\dot{\mathbf{u}}^* + \mathbf{e_u}^T(\mathbf{A} - \mathbf{A}^*)\mathbf{w}^*$$
$$+ \mathbf{e_u}^T\big(\mathbf{F}(\mathbf{x}, \mathbf{u})\mathbf{w} - \mathbf{F}(\mathbf{x}^*, \mathbf{u}^*)\mathbf{w}^*\big)$$
$$+ \mathbf{e_u}^T\big(\mathbf{f}(\mathbf{x}, \mathbf{u}) - \mathbf{f}(\mathbf{x}^*, \mathbf{u}^*)\big) - \mathbf{e_u}^T\mathbf{M}^*\mathbf{d_u}$$
$$- \mathbf{e_w}^T(\mathbf{M} - \mathbf{M}^*)\mathbf{w}^* - \mathbf{e_w}^T(\mathbf{A} - \mathbf{A}^*)\mathbf{u}^*$$
$$+ \mathbf{e_w}^T\big(\mathbf{g}(\mathbf{x}, \mathbf{u}) - \mathbf{g}(\mathbf{x}^*, \mathbf{u}^*)\big) - \mathbf{e_w}^T\mathbf{M}^*\mathbf{d_w} + \mathbf{e_w}^T\boldsymbol{\vartheta}.$$

We estimate these terms by applying the same techniques as in (i): using (5.22) on the left-hand side, and on the right-hand side using (5.23), (5.24), (5.25), and (5.26).

We thus obtain the second energy estimate

$$\frac{\mathrm{d}}{\mathrm{d}t}\|\mathbf{e_u}\|_{\mathbf{M}}^2 + \|\mathbf{e_w}\|_{\mathbf{M}}^2 \leq c\|\mathbf{e_u}\|_{\mathbf{K}}\big(\|\mathbf{e_x}\|_{\mathbf{K}} + \|\mathbf{e_u}\|_{\mathbf{K}} + \|\mathbf{d_u}\|_{\star,\mathbf{x}^*}\big)$$
$$+ c\|\mathbf{e_w}\|_{\mathbf{K}}\big(\|\mathbf{e_x}\|_{\mathbf{K}} + \|\mathbf{e_u}\|_{\mathbf{K}} + \|\mathbf{d_w}\|_{\star,\mathbf{x}^*} + \|\bar{\mathbf{e}}_\mathbf{w}^0\|_{\star,\mathbf{x}^0}\big).$$
$$\tag{5.28}$$

We now take the weighted linear combination of the two energy estimates (5.27) and (5.28), with weights 1 and $4c_0$, respectively, to obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\|\mathbf{e_u}\|_{\mathbf{A}}^2 + 4c_0\frac{\mathrm{d}}{\mathrm{d}t}\|\mathbf{e_u}\|_{\mathbf{M}}^2 + \|\mathbf{e_w}\|_{\mathbf{A}}^2 + 4c_0\|\mathbf{e_w}\|_{\mathbf{M}}^2$$
$$\leq c\|\mathbf{e_u}\|_{\mathbf{A}}^2 + 2c_0\|\mathbf{e_w}\|_{\mathbf{M}}^2$$
$$+ c\|\mathbf{e_w}\|_{\mathbf{K}}\big(\|\mathbf{e_x}\|_{\mathbf{K}} + \|\mathbf{e_u}\|_{\mathbf{K}} + \|\mathbf{d_u}\|_{\star,\mathbf{x}^*} + \|\mathbf{d_w}\|_{\star,\mathbf{x}^*} + \|\bar{\mathbf{e}}_\mathbf{w}^0\|_{\star,\mathbf{x}^0}\big)$$
$$+ c\|\dot{\mathbf{e}}_\mathbf{u}\|_{\mathbf{K}}\big(\|\mathbf{e_x}\|_{\mathbf{K}} + \|\mathbf{e_u}\|_{\mathbf{K}} + \|\mathbf{d_w}\|_{\star,\mathbf{x}^*} + \|\bar{\mathbf{e}}_\mathbf{w}^0\|_{\star,\mathbf{x}^0}\big)$$
$$+ c\|\mathbf{e_u}\|_{\mathbf{K}}\big(\|\mathbf{e_x}\|_{\mathbf{K}} + \|\mathbf{e_u}\|_{\mathbf{K}} + \|\mathbf{d_u}\|_{\star,\mathbf{x}^*}\big).$$

The term $\|\mathbf{e_w}\|_{\mathbf{M}}^2$ is absorbed to the left-hand side. Then, using Young's inequality (often weighted with a small constant $\varrho > 0$ that can be chosen arbitrarily), by further absorptions to the left-hand side, and by collecting the terms, we obtain the estimate (with constants c that depend on the choice of ϱ)

$$\frac{\mathrm{d}}{\mathrm{d}t}\|\mathbf{e_u}\|_{\mathbf{A}}^2 + 8c_0\frac{\mathrm{d}}{\mathrm{d}t}\|\mathbf{e_u}\|_{\mathbf{M}}^2 + c_0\|\mathbf{e_w}\|_{\mathbf{K}}^2$$
$$\leq \varrho\|\dot{\mathbf{e}}_\mathbf{u}\|_{\mathbf{K}}^2 + c\big(\|\mathbf{e_x}\|_{\mathbf{K}}^2 + \|\mathbf{e_u}\|_{\mathbf{K}}^2\big) + c\big(\|\mathbf{d_u}\|_{\star,\mathbf{x}^*}^2 + \|\mathbf{d_w}\|_{\star,\mathbf{x}^*}^2 + \|\bar{\mathbf{e}}_\mathbf{w}^0\|_{\star,\mathbf{x}^0}^2\big).$$
$$\tag{5.29}$$

We now integrate the inequality (5.29) in time, divide by $\min\{1, c_0\}$, and use the norm equivalence (5.7), which altogether yields

$$
\begin{aligned}
\|\mathbf{e_u}(t)\|^2_{\mathbf{K}(\mathbf{x}^*(t))} + & \int_0^t \|\mathbf{e_w}(s)\|^2_{\mathbf{K}(\mathbf{x}^*(s))}\mathrm{d}s \\
& \leq \|\mathbf{e_u}(0)\|^2_{\mathbf{K}(\mathbf{x}^*(0))} + \varrho \int_0^t \|\dot{\mathbf{e}}_{\mathbf{u}}(s)\|^2_{\mathbf{K}(\mathbf{x}^*(s))}\mathrm{d}s \\
& + c \int_0^t \big(\|\mathbf{e_u}(s)\|^2_{\mathbf{K}(\mathbf{x}^*(s))} + \|\mathbf{e_x}(s)\|^2_{\mathbf{K}(\mathbf{x}^*(s))}\big)\mathrm{d}s \\
& + c \int_0^t \big(\|\mathbf{d_u}(s)\|^2_{\star,\mathbf{x}^*(s)} + \|\mathbf{d_w}(s)\|^2_{\star,\mathbf{x}^*(s)} + \|\bar{\mathbf{e}}_{\mathbf{w}}^0\|^2_{\star,\mathbf{x}^0}\big)\mathrm{d}s.
\end{aligned}
\tag{5.30}
$$

(A.2) To establish a bound for the critical term involving $\|\dot{\mathbf{e}}_{\mathbf{u}}\|_{\mathbf{K}^*}$ in (5.30) and to show a uniform-in-time bound for $\mathbf{e_w}$ in the $\mathbf{K}$-norm, we perform a second pair of energy estimates. We take the time derivative of equation (5.16d):

$$
\begin{aligned}
\mathbf{M}\dot{\mathbf{e}}_{\mathbf{w}} + \mathbf{A}\dot{\mathbf{e}}_{\mathbf{u}} = & -\dot{\mathbf{M}}\mathbf{e_w} - \dot{\mathbf{A}}\mathbf{e_u} \\
& - (\mathbf{M} - \mathbf{M}^*)\dot{\mathbf{w}}^* - \frac{\mathrm{d}}{\mathrm{d}t}(\mathbf{M} - \mathbf{M}^*)\mathbf{w}^* \\
& - (\mathbf{A} - \mathbf{A}^*)\dot{\mathbf{u}}^* - \frac{\mathrm{d}}{\mathrm{d}t}(\mathbf{A} - \mathbf{A}^*)\mathbf{u}^* \\
& + \frac{\mathrm{d}}{\mathrm{d}t}\big(\mathbf{g}(\mathbf{x}, \mathbf{u}) - \mathbf{g}(\mathbf{x}^*, \mathbf{u}^*)\big) \\
& - \mathbf{M}^*\dot{\mathbf{d}}_{\mathbf{w}} - \dot{\mathbf{M}}^*\mathbf{d_w}.
\end{aligned}
\tag{5.31}
$$

Energy estimate (iii): We test (5.16c) with $\dot{\mathbf{e}}_{\mathbf{u}}$ and (5.31) with $\mathbf{e_w}$, then sum up to cancel the terms $\mathbf{e_w}^T \mathbf{A}\dot{\mathbf{e}}_{\mathbf{u}}$, and obtain

$$
\begin{aligned}
\mathbf{e_w}^T \mathbf{M}\dot{\mathbf{e}}_{\mathbf{w}} + \|\dot{\mathbf{e}}_{\mathbf{u}}\|^2_{\mathbf{M}} = & -\dot{\mathbf{e}}_{\mathbf{u}}^T(\mathbf{M} - \mathbf{M}^*)\dot{\mathbf{u}}^* + \dot{\mathbf{e}}_{\mathbf{u}}^T(\mathbf{A} - \mathbf{A}^*)\mathbf{w}^* \\
& + \dot{\mathbf{e}}_{\mathbf{u}}^T\big(\mathbf{F}(\mathbf{x}, \mathbf{u})\mathbf{w} - \mathbf{F}(\mathbf{x}^*, \mathbf{u}^*)\mathbf{w}^*\big) \\
& + \dot{\mathbf{e}}_{\mathbf{u}}^T\big(\mathbf{f}(\mathbf{x}, \mathbf{u}) - \mathbf{f}(\mathbf{x}^*, \mathbf{u}^*)\big) - \dot{\mathbf{e}}_{\mathbf{u}}^T\mathbf{M}^*\mathbf{d_u} \\
& - \mathbf{e_w}^T\dot{\mathbf{M}}\mathbf{e_w} - \mathbf{e_w}^T\dot{\mathbf{A}}\mathbf{e_u} \\
& - \mathbf{e_w}^T(\mathbf{M} - \mathbf{M}^*)\dot{\mathbf{w}}^* - \mathbf{e_w}^T\frac{\mathrm{d}}{\mathrm{d}t}(\mathbf{M} - \mathbf{M}^*)\mathbf{w}^* \\
& - \mathbf{e_w}^T(\mathbf{A} - \mathbf{A}^*)\dot{\mathbf{u}}^* - \mathbf{e_w}^T\frac{\mathrm{d}}{\mathrm{d}t}(\mathbf{A} - \mathbf{A}^*)\mathbf{u}^* \\
& + \mathbf{e_w}^T\frac{\mathrm{d}}{\mathrm{d}t}\big(\mathbf{g}(\mathbf{x}, \mathbf{u}) - \mathbf{g}(\mathbf{x}^*, \mathbf{u}^*)\big) \\
& - \mathbf{e_w}^T\mathbf{M}^*\dot{\mathbf{d}}_{\mathbf{w}} - \mathbf{e_w}^T\dot{\mathbf{M}}^*\mathbf{d_w}.
\end{aligned}
$$

We estimate these terms by applying the same techniques as in (i) and (ii): Using (5.22) on the left-hand side, and on the right-hand side using (5.23), (5.24), and (5.25) and (5.26).

The new terms involving time derivatives of the mass or stiffness matrix or their differences are bounded as follows.

Using (5.12) (via the bounds (5.20)) we obtain

$$
\begin{aligned}
-\mathbf{e}_{\mathbf{w}}^T \dot{\mathbf{M}} \mathbf{e}_{\mathbf{w}} &\leq c\|\mathbf{e}_{\mathbf{w}}\|_{\mathbf{M}}^2 \leq c\|\mathbf{e}_{\mathbf{w}}\|_{\mathbf{K}}^2, \\
-\mathbf{e}_{\mathbf{w}}^T \dot{\mathbf{A}} \mathbf{e}_{\mathbf{u}} &\leq c\|\mathbf{e}_{\mathbf{w}}\|_{\mathbf{A}}\|\mathbf{e}_{\mathbf{u}}\|_{\mathbf{A}} \leq c\|\mathbf{e}_{\mathbf{w}}\|_{\mathbf{K}}\|\mathbf{e}_{\mathbf{u}}\|_{\mathbf{K}}.
\end{aligned}
\tag{5.32}
$$

Time derivatives of the differences of mass or stiffness matrices are bounded, exactly as in Part (A.iv) in the proof of Proposition 7.1 in [KLL19], by

$$
\begin{aligned}
-\mathbf{e}_{\mathbf{w}}^T \frac{\mathrm{d}}{\mathrm{d}t}(\mathbf{M} - \mathbf{M}^*)\mathbf{w}^* &\leq c\|\mathbf{e}_{\mathbf{w}}\|_{\mathbf{K}}\big(\|\mathbf{e}_{\mathbf{v}}\|_{\mathbf{K}} + \|\mathbf{e}_{\mathbf{x}}\|_{\mathbf{K}}\big), \\
-\mathbf{e}_{\mathbf{w}}^T \frac{\mathrm{d}}{\mathrm{d}t}(\mathbf{A} - \mathbf{A}^*)\mathbf{u}^* &\leq c\|\mathbf{e}_{\mathbf{w}}\|_{\mathbf{K}}\big(\|\mathbf{e}_{\mathbf{v}}\|_{\mathbf{K}} + \|\mathbf{e}_{\mathbf{x}}\|_{\mathbf{K}}\big).
\end{aligned}
\tag{5.33}
$$

The defect term with a time derivative of the mass matrix is bounded using (5.12), by

$$
\mathbf{e}_{\mathbf{w}}^T \dot{\mathbf{M}}^* \mathbf{d}_{\mathbf{w}} \leq c\|\mathbf{e}_{\mathbf{w}}\|_{\mathbf{M}^*}\|\mathbf{d}_{\mathbf{w}}\|_{\mathbf{M}^*} \leq c\|\mathbf{e}_{\mathbf{w}}\|_{\mathbf{M}}\|\mathbf{d}_{\mathbf{w}}\|_{\mathbf{M}^*}.
\tag{5.34}
$$

The term containing the derivative of the nonlinearity $\mathbf{g}$ is estimated by techniques used in Part (A.v) of the proof of Proposition 7.1 in [KLL19]. A lengthy calculation yields

$$
\mathbf{e}_{\mathbf{w}}^T \frac{\mathrm{d}}{\mathrm{d}t}\big(\mathbf{g}(\mathbf{x}, \mathbf{u}) - \mathbf{g}(\mathbf{x}^*, \mathbf{u}^*)\big) \leq c\|\mathbf{e}_{\mathbf{w}}\|_{\mathbf{M}}\Big(\|\mathbf{e}_{\mathbf{x}}\|_{\mathbf{K}} + \|\mathbf{e}_{\mathbf{v}}\|_{\mathbf{K}} + \|\mathbf{e}_{\mathbf{u}}\|_{\mathbf{K}} + \|\dot{\mathbf{e}}_{\mathbf{u}}\|_{\mathbf{K}}\Big).
\tag{5.35}
$$

By combining the above estimates we obtain the third energy estimate

$$
\begin{aligned}
\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}&\|\mathbf{e}_{\mathbf{w}}\|_{\mathbf{M}}^2 + \frac{1}{2}\|\dot{\mathbf{e}}_{\mathbf{u}}\|_{\mathbf{M}}^2 \\
\leq\ & c\|\dot{\mathbf{e}}_{\mathbf{u}}\|_{\mathbf{K}}\big(\|\mathbf{e}_{\mathbf{x}}\|_{\mathbf{K}} + \|\mathbf{e}_{\mathbf{u}}\|_{\mathbf{K}} + \|\mathbf{e}_{\mathbf{w}}\|_{\mathbf{K}}\big) \\
& + c\|\mathbf{e}_{\mathbf{w}}\|_{\mathbf{K}}\big(\|\mathbf{e}_{\mathbf{x}}\|_{\mathbf{K}} + \|\mathbf{e}_{\mathbf{v}}\|_{\mathbf{K}} + \|\mathbf{e}_{\mathbf{u}}\|_{\mathbf{K}} + \|\dot{\mathbf{e}}_{\mathbf{u}}\|_{\mathbf{K}} + \|\mathbf{e}_{\mathbf{w}}\|_{\mathbf{K}}\big) \\
& + c\|\dot{\mathbf{e}}_{\mathbf{u}}\|_{\mathbf{K}}\,\|\mathbf{d}_{\mathbf{u}}\|_{\star,\mathbf{x}^*} + c\|\mathbf{e}_{\mathbf{w}}\|_{\mathbf{K}}\big(\|\dot{\mathbf{d}}_{\mathbf{w}}\|_{\star,\mathbf{x}^*} + \|\mathbf{d}_{\mathbf{w}}\|_{\mathbf{M}^*}\big).
\end{aligned}
\tag{5.36}
$$

Integrating this inequality in time, we obtain with a constant ϱ that is chosen suitably small, and with constants c depending on the choice of ϱ,

$$
\begin{aligned}
\frac{1}{2}&\|\mathbf{e}_{\mathbf{w}}(t)\|_{\mathbf{M}(\mathbf{x}(t))}^2 + \frac{1}{2}\int_0^t \|\dot{\mathbf{e}}_{\mathbf{u}}(s)\|_{\mathbf{M}(\mathbf{x}(s))}^2 \mathrm{d}s \\
\leq\ & \frac{1}{2}\|\mathbf{e}_{\mathbf{w}}(0)\|_{\mathbf{M}(\mathbf{x}(0))}^2 + \varrho\int_0^t \|\dot{\mathbf{e}}_{\mathbf{u}}(s)\|_{\mathbf{K}(\mathbf{x}(s))}^2 \mathrm{d}s \\
& + c\int_0^t \big(\|\mathbf{e}_{\mathbf{x}}(s)\|_{\mathbf{K}(\mathbf{x}(s))}^2 + \|\mathbf{e}_{\mathbf{v}}(s)\|_{\mathbf{K}(\mathbf{x}(s))}^2\big)\mathrm{d}s \\
& + c\int_0^t \big(\|\mathbf{e}_{\mathbf{u}}(s)\|_{\mathbf{K}(\mathbf{x}(s))}^2 + \|\mathbf{e}_{\mathbf{w}}(s)\|_{\mathbf{K}(\mathbf{x}(s))}^2\big)\mathrm{d}s \\
& + c\int_0^t \big(\|\mathbf{d}_{\mathbf{u}}(s)\|_{\star,\mathbf{x}^*}^2 + \|\mathbf{d}_{\mathbf{w}}(s)\|_{\mathbf{M}^*}^2 + \|\dot{\mathbf{d}}_{\mathbf{w}}(s)\|_{\star,\mathbf{x}^*}^2\big)\mathrm{d}s.
\end{aligned}
\tag{5.37}
$$

Energy estimate (iv): Finally, we test (5.16c) with $\dot{\mathbf{e}}_{\mathbf{w}}$ and (5.31) with $\dot{\mathbf{e}}_{\mathbf{u}}$, then subtract the former from the latter to obtain

$$
\begin{aligned}
\dot{\mathbf{e}}_{\mathbf{w}}^T \mathbf{A} \mathbf{e}_{\mathbf{w}} + \|\dot{\mathbf{e}}_{\mathbf{u}}\|_{\mathbf{A}}^2 = {}& + \dot{\mathbf{e}}_{\mathbf{w}}^T (\mathbf{M} - \mathbf{M}^*) \dot{\mathbf{u}}^* - \dot{\mathbf{e}}_{\mathbf{w}}^T (\mathbf{A} - \mathbf{A}^*) \mathbf{w}^* \\
& - \dot{\mathbf{e}}_{\mathbf{w}}^T \big(\mathbf{F}(\mathbf{x}, \mathbf{u}) \mathbf{w} - \mathbf{F}(\mathbf{x}^*, \mathbf{u}^*) \mathbf{w}^* \big) \\
& - \dot{\mathbf{e}}_{\mathbf{w}}^T \big(\mathbf{f}(\mathbf{x}, \mathbf{u}) - \mathbf{f}(\mathbf{x}^*, \mathbf{u}^*) \big) + \dot{\mathbf{e}}_{\mathbf{w}}^T \mathbf{M}^* \mathbf{d}_{\mathbf{u}} \\
& - \dot{\mathbf{e}}_{\mathbf{u}}^T \dot{\mathbf{M}} \mathbf{e}_{\mathbf{w}} - \dot{\mathbf{e}}_{\mathbf{u}}^T \dot{\mathbf{A}} \mathbf{e}_{\mathbf{u}} \\
& - \dot{\mathbf{e}}_{\mathbf{u}}^T (\mathbf{M} - \mathbf{M}^*) \dot{\mathbf{w}}^* - \dot{\mathbf{e}}_{\mathbf{u}}^T \frac{\mathrm{d}}{\mathrm{d}t} (\mathbf{M} - \mathbf{M}^*) \mathbf{w}^* \\
& - \dot{\mathbf{e}}_{\mathbf{u}}^T (\mathbf{A} - \mathbf{A}^*) \dot{\mathbf{u}}^* - \dot{\mathbf{e}}_{\mathbf{u}}^T \frac{\mathrm{d}}{\mathrm{d}t} (\mathbf{A} - \mathbf{A}^*) \mathbf{u}^* \\
& + \dot{\mathbf{e}}_{\mathbf{u}}^T \frac{\mathrm{d}}{\mathrm{d}t} \big(\mathbf{g}(\mathbf{x}, \mathbf{u}) - \mathbf{g}(\mathbf{x}^*, \mathbf{u}^*) \big) \\
& - \dot{\mathbf{e}}_{\mathbf{u}}^T \mathbf{M}^* \dot{\mathbf{d}}_{\mathbf{w}} - \dot{\mathbf{e}}_{\mathbf{u}}^T \dot{\mathbf{M}}^* \mathbf{d}_{\mathbf{w}}.
\end{aligned}
$$

We bound these terms by applying the same techniques as in (iii) and (i)–(ii): We bound the terms using (5.22) on the left-hand side, and on the right-hand side all the terms that do not involve $\dot{\mathbf{e}}_{\mathbf{w}}$, using (5.23), (5.24), and (5.25) and (5.26). The terms involving time derivatives of $\mathbf{M}$ and $\mathbf{A}$ are bounded using (5.32)– (5.34). As in (5.35) we have

$$
\dot{\mathbf{e}}_{\mathbf{u}}^T \frac{\mathrm{d}}{\mathrm{d}t} \big(\mathbf{g}(\mathbf{x}, \mathbf{u}) - \mathbf{g}(\mathbf{x}^*, \mathbf{u}^*) \big) \leq c \|\dot{\mathbf{e}}_{\mathbf{u}}\|_{\mathbf{M}} \Big(\|\mathbf{e}_{\mathbf{x}}\|_{\mathbf{K}} + \|\mathbf{e}_{\mathbf{v}}\|_{\mathbf{K}} + \|\mathbf{e}_{\mathbf{u}}\|_{\mathbf{K}} + \|\dot{\mathbf{e}}_{\mathbf{u}}\|_{\mathbf{K}} \Big).
\tag{5.38}
$$

The critical term $c \|\dot{\mathbf{e}}_{\mathbf{u}}\|_{\mathbf{M}} \|\dot{\mathbf{e}}_{\mathbf{u}}\|_{\mathbf{K}}$ is bounded, using Young's inequality and recalling that $\mathbf{K} = \mathbf{M} + \mathbf{A}$, as follows:

$$
\begin{aligned}
c \|\dot{\mathbf{e}}_{\mathbf{u}}\|_{\mathbf{M}} \|\dot{\mathbf{e}}_{\mathbf{u}}\|_{\mathbf{K}} &= c \|\dot{\mathbf{e}}_{\mathbf{u}}\|_{\mathbf{M}} \|\dot{\mathbf{e}}_{\mathbf{u}}\|_{\mathbf{M}+\mathbf{A}} \\
&\leq c \|\dot{\mathbf{e}}_{\mathbf{u}}\|_{\mathbf{M}}^2 + c \|\dot{\mathbf{e}}_{\mathbf{u}}\|_{\mathbf{M}} \|\dot{\mathbf{e}}_{\mathbf{u}}\|_{\mathbf{A}} \\
&\leq c_1 \|\dot{\mathbf{e}}_{\mathbf{u}}\|_{\mathbf{M}}^2 + \tfrac{1}{2} \|\dot{\mathbf{e}}_{\mathbf{u}}\|_{\mathbf{A}}^2, \qquad \text{with a constant } c_1 > 0.
\end{aligned}
\tag{5.39}
$$

The terms involving $\dot{\mathbf{e}}_{\mathbf{w}}$ are rewritten and bounded by the following trick, cf. [KLL19, equation (7.23)], and then bounded using (5.9), see the arguments

to (5.23) and (5.33) (cf. Part (A.iv) in [KLL19]):

$$
\begin{aligned}
\dot{\mathbf{e}}_{\mathbf{w}}^T(\mathbf{M} - \mathbf{M}^*)\dot{\mathbf{u}}^* &= \frac{\mathrm{d}}{\mathrm{d}t}\Big(\mathbf{e}_{\mathbf{w}}^T(\mathbf{M} - \mathbf{M}^*)\dot{\mathbf{u}}^*\Big) \\
&\quad - \mathbf{e}_{\mathbf{w}}^T\frac{\mathrm{d}}{\mathrm{d}t}(\mathbf{M} - \mathbf{M}^*)\dot{\mathbf{u}}^* - \mathbf{e}_{\mathbf{w}}^T(\mathbf{M} - \mathbf{M}^*)\ddot{\mathbf{u}}^* \\
&\leq \frac{\mathrm{d}}{\mathrm{d}t}\Big(\mathbf{e}_{\mathbf{w}}^T(\mathbf{M} - \mathbf{M}^*)\dot{\mathbf{u}}^*\Big) + c\|\mathbf{e}_{\mathbf{w}}\|_{\mathbf{K}}\big(\|\mathbf{e}_{\mathbf{x}}\|_{\mathbf{K}} + \|\mathbf{e}_{\mathbf{v}}\|_{\mathbf{K}}\big), \\
-\dot{\mathbf{e}}_{\mathbf{w}}^T(\mathbf{A} - \mathbf{A}^*)\mathbf{w}^* &= -\frac{\mathrm{d}}{\mathrm{d}t}\Big(\mathbf{e}_{\mathbf{w}}^T(\mathbf{A} - \mathbf{A}^*)\mathbf{w}^*\Big) \\
&\quad + \mathbf{e}_{\mathbf{w}}^T\frac{\mathrm{d}}{\mathrm{d}t}(\mathbf{A} - \mathbf{A}^*)\mathbf{w}^* + \mathbf{e}_{\mathbf{w}}^T(\mathbf{A} - \mathbf{A}^*)\dot{\mathbf{w}}^* \\
&\leq -\frac{\mathrm{d}}{\mathrm{d}t}\Big(\mathbf{e}_{\mathbf{w}}^T(\mathbf{A} - \mathbf{A}^*)\mathbf{w}^*\Big) + c\|\mathbf{e}_{\mathbf{w}}\|_{\mathbf{K}}\big(\|\mathbf{e}_{\mathbf{x}}\|_{\mathbf{K}} + \|\mathbf{e}_{\mathbf{v}}\|_{\mathbf{K}}\big).
\end{aligned}
\tag{5.40}
$$

Analogously, for the non-linear terms we obtain

$$
\begin{aligned}
&-\dot{\mathbf{e}}_{\mathbf{w}}^T\big(\mathbf{F}(\mathbf{x}, \mathbf{u})\mathbf{w} - \mathbf{F}(\mathbf{x}^*, \mathbf{u}^*)\mathbf{w}^*\big) \\
&= -\dot{\mathbf{e}}_{\mathbf{w}}^T\mathbf{F}(\mathbf{x}, \mathbf{u})\mathbf{e}_{\mathbf{w}} - \dot{\mathbf{e}}_{\mathbf{w}}^T\big(\mathbf{F}(\mathbf{x}, \mathbf{u}) - \mathbf{F}(\mathbf{x}^*, \mathbf{u}^*)\big)\mathbf{w}^* \\
&= -\frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{1}{2}\mathbf{e}_{\mathbf{w}}^T\mathbf{F}(\mathbf{x}, \mathbf{u})\mathbf{e}_{\mathbf{w}} + \mathbf{e}_{\mathbf{w}}^T\big(\mathbf{F}(\mathbf{x}, \mathbf{u}) - \mathbf{F}(\mathbf{x}^*, \mathbf{u}^*)\big)\mathbf{w}^*\Big) \\
&\quad + \frac{1}{2}\mathbf{e}_{\mathbf{w}}^T\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{F}(\mathbf{x}, \mathbf{u})\mathbf{e}_{\mathbf{w}} + \mathbf{e}_{\mathbf{w}}^T\frac{\mathrm{d}}{\mathrm{d}t}\big(\mathbf{F}(\mathbf{x}, \mathbf{u}) - \mathbf{F}(\mathbf{x}^*, \mathbf{u}^*)\big)\mathbf{w}^* \\
&\leq -\frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{1}{2}\mathbf{e}_{\mathbf{w}}^T\mathbf{F}(\mathbf{x}, \mathbf{u})\mathbf{e}_{\mathbf{w}} + \mathbf{e}_{\mathbf{w}}^T\big(\mathbf{F}(\mathbf{x}, \mathbf{u}) - \mathbf{F}(\mathbf{x}^*, \mathbf{u}^*)\big)\mathbf{w}^*\Big) \\
&\quad + c\|\mathbf{e}_{\mathbf{w}}\|_{\mathbf{M}}\Big(\|\mathbf{e}_{\mathbf{x}}\|_{\mathbf{K}} + \|\mathbf{e}_{\mathbf{v}}\|_{\mathbf{K}} + \|\mathbf{e}_{\mathbf{u}}\|_{\mathbf{K}} + \|\dot{\mathbf{e}}_{\mathbf{u}}\|_{\mathbf{K}} + \|\mathbf{e}_{\mathbf{w}}\|_{\mathbf{M}}\Big),
\end{aligned}
$$

where the inequality for the third term is shown as in (5.35). Similarly, we have for the other term

$$
\begin{aligned}
-\dot{\mathbf{e}}_{\mathbf{w}}^T\big(\mathbf{f}(\mathbf{x}, \mathbf{u}) - \mathbf{f}(\mathbf{x}^*, \mathbf{u}^*)\big) &\leq -\frac{\mathrm{d}}{\mathrm{d}t}\Big(\mathbf{e}_{\mathbf{w}}^T\big(\mathbf{f}(\mathbf{x}, \mathbf{u}) - \mathbf{f}(\mathbf{x}^*, \mathbf{u}^*)\big)\Big) \\
&\quad + c\|\mathbf{e}_{\mathbf{w}}\|_{\mathbf{K}}\Big(\|\mathbf{e}_{\mathbf{x}}\|_{\mathbf{K}} + \|\mathbf{e}_{\mathbf{v}}\|_{\mathbf{K}} + \|\mathbf{e}_{\mathbf{u}}\|_{\mathbf{K}} + \|\dot{\mathbf{e}}_{\mathbf{u}}\|_{\mathbf{K}}\Big).
\end{aligned}
$$

Finally, the defect term is bounded using (5.34),

$$
\begin{aligned}
\dot{\mathbf{e}}_{\mathbf{w}}^T\mathbf{M}^*\mathbf{d}_{\mathbf{u}} &= \frac{\mathrm{d}}{\mathrm{d}t}\Big(\mathbf{e}_{\mathbf{w}}^T\mathbf{M}^*\mathbf{d}_{\mathbf{u}}\Big) - \mathbf{e}_{\mathbf{w}}^T\mathbf{M}^*\dot{\mathbf{d}}_{\mathbf{u}} - \mathbf{e}_{\mathbf{w}}^T\dot{\mathbf{M}}^*\mathbf{d}_{\mathbf{u}} \\
&\leq \frac{\mathrm{d}}{\mathrm{d}t}\Big(\mathbf{e}_{\mathbf{w}}^T\mathbf{M}^*\mathbf{d}_{\mathbf{u}}\Big) + c\|\mathbf{e}_{\mathbf{w}}\|_{\mathbf{K}}\big(\|\dot{\mathbf{d}}_{\mathbf{u}}\|_{\star, \mathbf{x}^*} + \|\mathbf{d}_{\mathbf{u}}\|_{\mathbf{M}^*}\big).
\end{aligned}
$$

Then the combination of the bounds above, and an absorption to the left-hand side of the last term in (5.39), yields the fourth energy estimate

$$
\begin{aligned}
&\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\mathbf{e_w}\|_{\mathbf{A}}^2 + \frac{1}{2}\|\dot{\mathbf{e}}_{\mathbf{u}}\|_{\mathbf{A}}^2 \\
&\leq c_1\|\dot{\mathbf{e}}_{\mathbf{u}}\|_{\mathbf{M}}^2 \\
&\quad + c\|\mathbf{e_w}\|_{\mathbf{K}}\Big(\|\mathbf{e_x}\|_{\mathbf{K}} + \|\mathbf{e_v}\|_{\mathbf{K}} + \|\mathbf{e_u}\|_{\mathbf{K}} + \|\dot{\mathbf{e}}_{\mathbf{u}}\|_{\mathbf{K}} + \|\mathbf{e_w}\|_{\mathbf{K}}\Big) \\
&\quad + c\|\dot{\mathbf{e}}_{\mathbf{u}}\|_{\mathbf{K}}\Big(\|\mathbf{e_x}\|_{\mathbf{K}} + \|\mathbf{e_v}\|_{\mathbf{K}} + \|\mathbf{e_u}\|_{\mathbf{K}} + \|\mathbf{e_w}\|_{\mathbf{K}}\Big) \\
&\quad + c\|\mathbf{e_w}\|_{\mathbf{K}}\big(\|\mathbf{d_u}\|_{\mathbf{M}^*} + \|\dot{\mathbf{d}}_{\mathbf{u}}\|_{\star,\mathbf{x}^*}\big) + c\|\dot{\mathbf{e}}_{\mathbf{u}}\|_{\mathbf{K}}\big(\|\mathbf{d_w}\|_{\mathbf{M}^*} + \|\dot{\mathbf{d}}_{\mathbf{w}}\|_{\star,\mathbf{x}^*}\big) \\
&\quad + \frac{\mathrm{d}}{\mathrm{d}t}\Big(\mathbf{e_w}^T(\mathbf{M}-\mathbf{M}^*)\dot{\mathbf{u}}^*\Big) - \frac{\mathrm{d}}{\mathrm{d}t}\Big(\mathbf{e_w}^T(\mathbf{M}-\mathbf{M}^*)\mathbf{w}^*\Big) + \frac{\mathrm{d}}{\mathrm{d}t}\Big(\mathbf{e_w}^T\mathbf{M}^*\mathbf{d_u}\Big) \\
&\quad - \frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{1}{2}\mathbf{e_w}^T\mathbf{F}(\mathbf{x},\mathbf{u})\mathbf{e_w} + \mathbf{e_w}^T\big(\mathbf{F}(\mathbf{x},\mathbf{u})-\mathbf{F}(\mathbf{x}^*,\mathbf{u}^*)\big)\mathbf{w}^*\Big) \\
&\quad - \frac{\mathrm{d}}{\mathrm{d}t}\Big(\mathbf{e_w}^T\big(\mathbf{f}(\mathbf{x},\mathbf{u})-\mathbf{f}(\mathbf{x}^*,\mathbf{u}^*)\big)\Big).
\end{aligned}
$$

(5.41)

Now we take the weighted sum of the the two energy estimates from (iii) and (iv), i.e. we sum $4c_1$ times (5.36) and (5.41). Then we absorb the term $c_1\|\dot{\mathbf{e}}_{\mathbf{u}}\|_{\mathbf{M}}^2$ to the left-hand side. Using Young's inequality (often with a small parameter), absorptions to the left-hand side, and collecting the terms, we obtain

$$
\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}\Big(2c_1\|\mathbf{e_w}\|_{\mathbf{M}}^2 + \frac{1}{2}\|\mathbf{e_w}\|_{\mathbf{A}}^2\Big) + c_1\|\dot{\mathbf{e}}_{\mathbf{u}}\|_{\mathbf{M}}^2 + \frac{1}{2}\|\dot{\mathbf{e}}_{\mathbf{u}}\|_{\mathbf{A}}^2 \\
&\leq c\|\mathbf{e_u}\|_{\mathbf{K}}^2 + c\|\mathbf{e_x}\|_{\mathbf{K}}^2 + c\|\mathbf{e_v}\|_{\mathbf{K}}^2 + c\|\mathbf{e_w}\|_{\mathbf{K}}^2 \\
&\quad + c\|\mathbf{d_u}\|_{\mathbf{M}^*}^2 + c\|\dot{\mathbf{d}}_{\mathbf{u}}\|_{\star,\mathbf{x}^*}^2 + c\|\mathbf{d_w}\|_{\mathbf{M}^*}^2 + c\|\dot{\mathbf{d}}_{\mathbf{w}}\|_{\star,\mathbf{x}^*}^2 \\
&\quad + c\frac{\mathrm{d}}{\mathrm{d}t}\Big(\mathbf{e_w}^T(\mathbf{M}-\mathbf{M}^*)\dot{\mathbf{u}}^*\Big) - c\frac{\mathrm{d}}{\mathrm{d}t}\Big(\mathbf{e_w}^T(\mathbf{M}-\mathbf{M}^*)\dot{\mathbf{w}}^*\Big) + c\frac{\mathrm{d}}{\mathrm{d}t}\Big(\mathbf{e_w}^T\mathbf{M}^*\mathbf{d_u}\Big) \\
&\quad - c\frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{1}{2}\mathbf{e_w}^T\mathbf{F}(\mathbf{x},\mathbf{u})\mathbf{e_w} + \mathbf{e_w}^T\big(\mathbf{F}(\mathbf{x},\mathbf{u})-\mathbf{F}(\mathbf{x}^*,\mathbf{u}^*)\big)\mathbf{w}^*\Big) \\
&\quad - c\frac{\mathrm{d}}{\mathrm{d}t}\Big(\mathbf{e_w}^T\big(\mathbf{f}(\mathbf{x},\mathbf{u})-\mathbf{f}(\mathbf{x}^*,\mathbf{u}^*)\big)\Big).
\end{aligned}
$$

We integrate this inequality in time from 0 to t, to obtain

$$
\|\mathbf{e_w}(t)\|^2_{\mathbf{K}(\mathbf{x}(t))} + \int_0^t \|\dot{\mathbf{e}}_{\mathbf{u}}(s)\|^2_{\mathbf{K}(\mathbf{x}(s))}\mathrm{d}s
$$
$$
\leq c\,\|\mathbf{e_w}(0)\|^2_{\mathbf{K}(\mathbf{x}(0))}
$$
$$
+ c\int_0^t \big(\|\mathbf{e_x}(s)\|^2_{\mathbf{K}(\mathbf{x}(s))} + \|\mathbf{e_v}(s)\|^2_{\mathbf{K}(\mathbf{x}(s))} + \|\mathbf{e_u}(s)\|^2_{\mathbf{K}(\mathbf{x}(s))} + \|\mathbf{e_w}(s)\|^2_{\mathbf{K}(\mathbf{x}(s))}\big)\mathrm{d}s
$$
$$
+ c\int_0^t \big(\|\mathbf{d_u}(s)\|^2_{\mathbf{M}(\mathbf{x}(s))} + \|\dot{\mathbf{d}}_{\mathbf{u}}(s)\|^2_{\star,\mathbf{x}(s)} + \|\mathbf{d_w}(s)\|^2_{\mathbf{M}(\mathbf{x}(s))} + c\|\dot{\mathbf{d}}_{\mathbf{w}}(s)\|^2_{\star,\mathbf{x}(s)}\big)\mathrm{d}s
$$
$$
+ c\,\mathbf{e_w}(t)^T\Big(\mathbf{M}(\mathbf{x}(t)) - \mathbf{M}(\mathbf{x}^*(t))\Big)\dot{\mathbf{u}}^*(t)
$$
$$
- c\,\mathbf{e_w}(t)^T\Big(\mathbf{M}(\mathbf{x}(t)) - \mathbf{M}(\mathbf{x}^*(t))\Big)\dot{\mathbf{w}}^*(t) + c\,\mathbf{e_w}(t)^T\mathbf{M}(\mathbf{x}(t))\mathbf{d_u}(t)
$$
$$
- c\,\mathbf{e_w}(t)^T\mathbf{F}(\mathbf{x}(t),\mathbf{u}(t))\mathbf{e_w}(t) - c\,\mathbf{e_w}(t)^T\big(\mathbf{F}(\mathbf{x}(t),\mathbf{u}(t)) - \mathbf{F}(\mathbf{x}^*(t),\mathbf{u}^*(t))\big)\mathbf{w}^*(t)
$$
$$
+ c\,\mathbf{e_w}(t)^T\Big(\mathbf{f}(\mathbf{x}(t),\mathbf{u}(t)) - \mathbf{f}(\mathbf{x}(t),\mathbf{u}^*(t))\Big).
$$

We estimate those terms that are not integrated in time using (5.8) (with (5.20)), and for the non-linear terms we use (5.26) and (5.25). Then using Young's inequality (possibly with a sufficiently small constant $\varrho > 0$) and absorbing terms to the left-hand side, we obtain

$$
\|\mathbf{e_w}(t)\|^2_{\mathbf{K}(\mathbf{x}(t))} + \int_0^t \|\dot{\mathbf{e}}_{\mathbf{u}}(s)\|^2_{\mathbf{K}(\mathbf{x}(s))}\mathrm{d}s
$$
$$
\leq c_2\|\mathbf{e_u}(t)\|^2_{\mathbf{K}(\mathbf{x}(t))} + c\|\mathbf{e_x}(t)\|^2_{\mathbf{K}(\mathbf{x}(t))} + c\|\mathbf{e_w}(t)\|^2_{\mathbf{M}(\mathbf{x}(t))}
$$
$$
+ c\int_0^t \big(\|\mathbf{e_x}(s)\|^2_{\mathbf{K}(\mathbf{x}(s))} + \|\mathbf{e_v}(s)\|^2_{\mathbf{K}(\mathbf{x}(s))} + \|\mathbf{e_u}(s)\|^2_{\mathbf{K}(\mathbf{x}(s))} + \|\mathbf{e_w}(s)\|^2_{\mathbf{K}(\mathbf{x}(s))}\big)\mathrm{d}s
$$
$$
+ c\int_0^t \big(\|\mathbf{d_u}(s)\|^2_{\mathbf{M}(\mathbf{x}(s))} + \|\dot{\mathbf{d}}_{\mathbf{u}}(s)\|^2_{\star,\mathbf{x}(s)} + \|\mathbf{d_w}(s)\|^2_{\mathbf{M}(\mathbf{x}(s))} + c\|\dot{\mathbf{d}}_{\mathbf{w}}(s)\|^2_{\star,\mathbf{x}(s)}\big)\mathrm{d}s
$$
$$
+ c\|\mathbf{d_u}(t)\|^2_{\star,\mathbf{x}(t)} + c\|\mathbf{e_w}(0)\|^2_{\mathbf{K}(\mathbf{x}(0))}, \qquad \text{with a constant } c_2 > 0.
$$

The term $\|\mathbf{e_w}(t)\|^2_{\mathbf{M}(\mathbf{x}(t))}$ can be estimated by using (5.37). Then we have

$$
\|\mathbf{e_w}(t)\|^2_{\mathbf{K}(\mathbf{x}(t))} + \frac{1}{2}\int_0^t \|\dot{\mathbf{e}}_{\mathbf{u}}(s)\|^2_{\mathbf{K}(\mathbf{x}(s))}\mathrm{d}s
$$
$$
\leq c_2\|\mathbf{e_u}(t)\|^2_{\mathbf{K}(\mathbf{x}(t))} + c\|\mathbf{e_x}(t)\|^2_{\mathbf{K}(\mathbf{x}(t))}
$$
$$
+ c\int_0^t \big(\|\mathbf{e_x}(s)\|^2_{\mathbf{K}(\mathbf{x}(s))} + \|\mathbf{e_v}(s)\|^2_{\mathbf{K}(\mathbf{x}(s))} + \|\mathbf{e_u}(s)\|^2_{\mathbf{K}(\mathbf{x}(s))} + \|\mathbf{e_w}(s)\|^2_{\mathbf{K}(\mathbf{x}(s))}\big)\mathrm{d}s
$$
$$
+ c\int_0^t \big(\|\mathbf{d_u}(s)\|^2_{\mathbf{M}(\mathbf{x}(s))} + \|\dot{\mathbf{d}}_{\mathbf{u}}(s)\|^2_{\star,\mathbf{x}(s)} + \|\mathbf{d_w}(s)\|^2_{\mathbf{M}(\mathbf{x}(s))} + c\|\dot{\mathbf{d}}_{\mathbf{w}}(s)\|^2_{\star,\mathbf{x}(s)}\big)\mathrm{d}s
$$
$$
+ c\|\mathbf{d_u}(t)\|^2_{\star,\mathbf{x}(t)} + c\|\mathbf{e_w}(0)\|^2_{\mathbf{K}(\mathbf{x}(0))}.
$$

$$(5.42)$$

(A.3) *Combining the energy estimates (i)–(iv):* We multiply (5.30) with $2c_2$ and sum with (5.42), absorb the terms (by choosing $\varrho > 0$ in (5.30) sufficiently small) with $\|\dot{e}_\mathbf{u}\|_\mathbf{K}$ and $\|e_\mathbf{u}\|_\mathbf{K}$ to the left-hand side, to obtain

$$
\begin{aligned}
&\|e_\mathbf{u}(t)\|^2_{\mathbf{K}(\mathbf{x}(t))} + \|e_\mathbf{w}(t)\|^2_{\mathbf{K}(\mathbf{x}(t))} \\
&\leq c\|e_\mathbf{x}(t)\|^2_{\mathbf{K}(\mathbf{x}(t))} \\
&\quad + c\int_0^t \big(\|e_\mathbf{x}(s)\|^2_\mathbf{K} + \|e_\mathbf{v}(s)\|^2_{\mathbf{K}(\mathbf{x}(s))} + \|e_\mathbf{u}(s)\|^2_{\mathbf{K}(\mathbf{x}(s))} + \|e_\mathbf{w}(s)\|^2_\mathbf{K}\big)\mathrm{d}s \\
&\quad + c\int_0^t \big(\|d_\mathbf{u}(s)\|^2_{\mathbf{M}(\mathbf{x}(s))} + \|\dot{d}_\mathbf{u}(s)\|^2_{\star,\mathbf{x}(s)} + \|d_\mathbf{w}(s)\|^2_{\mathbf{M}(\mathbf{x}(s))} + c\|\dot{d}_\mathbf{w}(s)\|^2_{\star,\mathbf{x}(s)}\big)\mathrm{d}s \\
&\quad + c\|d_\mathbf{u}(t)\|^2_{\star,\mathbf{x}(t)} + c\|e_\mathbf{u}(0)\|^2_{\mathbf{K}(\mathbf{x}(0))} + c\|e_\mathbf{w}(0)\|^2_{\mathbf{K}(\mathbf{x}(0))} + c\|\bar{e}_\mathbf{w}(0)\|^2_{\star,\mathbf{x}(0)}.
\end{aligned}
\tag{5.43}
$$

(B) *Estimates for the velocity equation:* We write $e_\mathbf{v}$ as the nodal vector of the finite element function $e_v \in S_h[\mathbf{x}^*]$. To obtain an expression for the function e_v, we denote by $\widetilde{I}_h^*$ the finite element interpolation operator on $\Gamma_h[\mathbf{x}^*]$, and we set

$$
V_h^\theta = \sum_{j=1}^N V_j\, \phi_j[\mathbf{x} + \theta(\mathbf{x}^* - \mathbf{x})], \quad \nu_h^\theta = \sum_{j=1}^N \nu_j\, \phi_j[\mathbf{x} + \theta(\mathbf{x}^* - \mathbf{x})].
$$

In view of (5.16b), we then have

$$
\begin{aligned}
e_v &= \widetilde{I}_h^*(V_h^1 \nu_h^1) - \widetilde{I}_h^*(V_h^* \nu_h^*) - d_v \\
&= \widetilde{I}_h^*((V_h^1 - V_h^*)\nu_h^1) + \widetilde{I}_h^*(V_h^*(\nu_h^1 - \nu_h^*)) - d_v.
\end{aligned}
$$

Lemma 5.3 gives us the bound

$$
\begin{aligned}
\|e_v\|_{H^1(\Gamma_h[\mathbf{x}^*])} &\leq C\, \|V_h^1 - V_h^*\|_{H^1(\Gamma_h[\mathbf{x}^*])}\|\nu_h^1\|_{W^{1,\infty}(\Gamma_h[\mathbf{x}^*])} \\
&\quad + C\, \|\nu_h^1 - \nu_h^*\|_{H^1(\Gamma_h[\mathbf{x}^*])}\|V_h^*\|_{W^{1,\infty}(\Gamma_h[\mathbf{x}^*])} + \|d_v\|_{H^1(\Gamma_h[\mathbf{x}^*])}.
\end{aligned}
$$

The $W^{1,\infty}$ norms of V_h^* and ν_h^* are bounded independently of h by assumption, and that of $\nu_h^1 = \nu_h^* + e_\nu$ by (5.20). Since the nodal vector of $V_h^1 - V_h^*$ is the subvector of the first N components of $e_\mathbf{w}$ and the nodal vector of $\nu_h^1 - \nu_h^*$ is the subvector of the last $3N$ components of $e_\mathbf{u}$, the above bound yields

$$
\|e_\mathbf{v}\|_{\mathbf{K}(\mathbf{x}^*)} \leq c\left(\|e_\mathbf{u}\|_{\mathbf{K}(\mathbf{x}^*)} + \|e_\mathbf{w}\|_{\mathbf{K}(\mathbf{x}^*)}\right) + \|d_\mathbf{v}\|_{\mathbf{K}(\mathbf{x}^*)}.
\tag{5.44}
$$

(C) *Combination:* We use the differential equation $\dot{e}_\mathbf{x} = e_\mathbf{v}$ (5.16a) to show the bound

$$
\begin{aligned}
\|e_\mathbf{x}(t)\|^2_{\mathbf{K}(\mathbf{x}^*(t))} &= \int_0^t \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}s}\|e_\mathbf{x}(s)\|^2_{\mathbf{K}(\mathbf{x}^*(s))}\mathrm{d}s \\
&\leq c\int_0^t \|e_\mathbf{v}(s)\|^2_{\mathbf{K}(\mathbf{x}^*(s))}\mathrm{d}s + c\int_0^t \|e_\mathbf{x}(s)\|^2_{\mathbf{K}(\mathbf{x}^*(s))}\mathrm{d}s,
\end{aligned}
$$

which is substituted into (5.43). This combined estimate is plugged into (5.44), using the equivalence (5.7) between the norms $\|\cdot\|_{\mathbf{K}(\mathbf{x}(t))}$ and $\|\cdot\|_{\mathbf{K}(\mathbf{x}^*(t))}$. Then we sum the three inequalities and obtain

$$
\begin{aligned}
&\|\mathbf{e_x}(t)\|^2_{\mathbf{K}(\mathbf{x}^*(t))} + \|\mathbf{e_v}(t)\|^2_{\mathbf{K}(\mathbf{x}^*(t))} + \|\mathbf{e_u}(t)\|^2_{\mathbf{K}(\mathbf{x}^*(t))} + \|\mathbf{e_w}(t)\|^2_{\mathbf{K}(\mathbf{x}^*(t))} \\
\leq\ &+ c\int_0^t \big(\|\mathbf{e_x}(s)\|^2_{\mathbf{K}(\mathbf{x}^*(s))} + \|\mathbf{e_v}(s)\|^2_{\mathbf{K}(\mathbf{x}^*(s))} + \|\mathbf{e_u}(s)\|^2_{\mathbf{K}(\mathbf{x}^*(s))} + \|\mathbf{e_w}(s)\|^2_{\mathbf{K}(\mathbf{x}^*(s))}\big)\mathrm{d}s \\
&+ c\int_0^t \big(\|\mathbf{d_u}(s)\|^2_{\mathbf{M}(\mathbf{x}^*(s))} + \|\dot{\mathbf{d}}_{\mathbf{u}}(s)\|^2_{\star,\mathbf{x}^*(s)} + \|\mathbf{d_w}(s)\|^2_{\mathbf{M}(\mathbf{x}^*(s))} + c\|\dot{\mathbf{d}}_{\mathbf{w}}(s)\|^2_{\star,\mathbf{x}^*(s)}\big)\mathrm{d}s \\
&+ c\|\mathbf{d_v}(t)\|^2_{\mathbf{K}(\mathbf{x}^*(t))} + c\|\mathbf{d_u}(t)\|^2_{\star,\mathbf{x}^*(t)} \\
&+ c\|\mathbf{e_u}(0)\|^2_{\mathbf{K}(\mathbf{x}^*(0))} + c\|\mathbf{e_w}(0)\|^2_{\mathbf{K}(\mathbf{x}^*(0))} + c\|\bar{\mathbf{e}}_{\mathbf{w}}(0)\|^2_{\star,\mathbf{x}^*(0)}.
\end{aligned}
$$

By Gronwall's inequality we obtain the stability bound (5.19) for $t \in [0, t^*]$.

Finally it remains to show that $t^* = T$ for h sufficiently small. To this end we use the assumed defect bounds to obtain the error estimates of order κ:

$$
\|\mathbf{e_x}(t)\|_{\mathbf{K}(\mathbf{x}^*)} + \|\mathbf{e_v}(t)\|_{\mathbf{K}(\mathbf{x}^*)} + \|\mathbf{e_u}(t)\|_{\mathbf{K}(\mathbf{x}^*)} + \|\mathbf{e_w}(t)\|_{\mathbf{K}(\mathbf{x}^*)} \leq Ch^\kappa.
$$

Then, by the inverse inequality, we have for $t \in [0, t^*]$

$$
\begin{aligned}
&\|e_x(\cdot, t)\|_{W^{1,\infty}(\Gamma_h[\mathbf{x}^*(t)])} + \|e_v(\cdot, t)\|_{W^{1,\infty}(\Gamma_h[\mathbf{x}^*(t)])} \\
&\quad + \|e_u(\cdot, t)\|_{W^{1,\infty}(\Gamma_h[\mathbf{x}^*(t)])} + \|e_w(\cdot, t)\|_{W^{1,\infty}(\Gamma_h[\mathbf{x}^*(t)])} \\
&\leq ch^{-1}\big(\|\mathbf{e_x}(t)\|_{\mathbf{K}(\mathbf{x}^*)} + \|\mathbf{e_v}(t)\|_{\mathbf{K}(\mathbf{x}^*)} + \|\mathbf{e_u}(t)\|_{\mathbf{K}(\mathbf{x}^*)} + \|\mathbf{e_w}(t)\|_{\mathbf{K}(\mathbf{x}^*)}\big) \\
&\leq cCh^{\kappa-1} \leq \tfrac{1}{2}h^{(\kappa-1)/2},
\end{aligned}
$$

for sufficiently small h. Hence we can extend the bounds (5.20) beyond t^*, which contradicts the maximality of t^* unless $t^* = T$. Therefore we have the stability bound (5.19) for $t \in [0, T]$. □

6 Consistency

In this section we show that the defects and the errors in the initial values can be bounded by $O(h^k)$ in the appropriate norms that appear in the stability bounds of Proposition 5.1. Together with the argument of [KLL19, Section 9], this completes the proof of Theorem 4.1. We first state the bounds in Subsection 6.1 and then turn to the essentials of their proof in the following subsections.

6.1 Defect bounds and initial value error bounds

As before, we let the vectors $\mathbf{x}^*(t) \in \mathbb{R}^{3N}$ and $\mathbf{v}^*(t) \in \mathbb{R}^{3N}$ collect the positions $X(p_i, t)$ of the moving finite element nodes on the exact surface $\Gamma(t) = \Gamma[X(\cdot, t)]$ and their velocities $v(X(p_i, t), t)$, respectively. They are the

nodal vectors of finite element functions $X_h^*(\cdot, t) \in S_h[\Gamma^0]^3$ and $v_h^*(\cdot, t) \in S_h(\Gamma_h[\mathbf{x}^*(t)])^3$ on the interpolated surface $\Gamma_h^*(t) = \Gamma_h[\mathbf{x}^*(t)] = \Gamma[X_h^*(\cdot, t)]$, which moves with the velocity $v_h^*(x, t) = \frac{\mathrm{d}}{\mathrm{d}t} X_h^*(p, t)$ for $x = X_h^*(p, t)$. We write

$$\partial_h^\bullet f(x, t) = \partial_x f(x, t) v_h^*(x, t) + \partial_t f(x, t), \qquad x \in \Gamma_h^*(t),$$

for the material derivative on $\Gamma_h^*(t)$ corresponding to the velocity v_h^*. (This is not the same as the material derivative on the discrete surface $\Gamma_h[\mathbf{x}(t)]$ with velocity v_h, although we use the same symbol $\partial_h^\bullet$.)

We further need reference finite element functions

$$u_h^*(\cdot, t) = \begin{pmatrix} H_h^*(t) \\ \nu_h^*(t) \end{pmatrix} \in S_h[\mathbf{x}^*(t)]^{1+3}, \quad w_h^*(\cdot, t) = \begin{pmatrix} V_h^*(t) \\ z_h^*(t) \end{pmatrix} \in S_h[\mathbf{x}^*(t)]^{1+3}$$

with their nodal vectors

$$\mathbf{u}^*(t) = \begin{pmatrix} \mathbf{H}^*(t) \\ \mathbf{n}^*(t) \end{pmatrix} \in \mathbb{R}^{N+3N}, \quad \mathbf{w}^*(t) = \begin{pmatrix} \mathbf{V}^*(t) \\ \mathbf{z}^*(t) \end{pmatrix} \in \mathbb{R}^{N+3N}.$$

Proposition 6.1 *Under the assumptions of Theorem 4.1, there exist, for $t \in [0, T]$, finite element functions $u_h^*(\cdot, t), w_h^*(\cdot, t) \in S_h[\mathbf{x}^*(t)]^4$ such that the following holds true:*

(a) The lift to the exact surface $\Gamma(t)$ of these functions is $O(h^k)$ close to the exact solution in the H^1-norm:

$$\begin{aligned} \|(u_h^*)^l(\cdot, t) - u(\cdot, t)\|_{H^1(\Gamma(t))} &\le C h^k, \\ \|(w_h^*)^l(\cdot, t) - w(\cdot, t)\|_{H^1(\Gamma(t))} &\le C h^k. \end{aligned} \tag{6.1}$$

(b) The defect functions $d_v(\cdot, t) \in S_h[\mathbf{x}^*(t)]^3$ and $d_u(\cdot, t), d_w(\cdot, t) \in S_h[\mathbf{x}^*(t)]^4$ with nodal vectors $\mathbf{d_v}(t)$ and $\mathbf{d_u}(t)$, $\mathbf{d_w}(t)$, respectively, which are defined by (5.14)–(5.15), are bounded by

$$\begin{aligned} \|d_v(\cdot, t)\|_{H^1(\Gamma_h^*(t))} &= \|\mathbf{d_v}(t)\|_{\mathbf{K}(\mathbf{x}^*(t))} \le C h^k, \\ \|d_u(\cdot, t)\|_{L^2(\Gamma_h^*(t))} &= \|\mathbf{d_u}(t)\|_{\mathbf{M}(\mathbf{x}^*(t))} \le C h^k, \\ \|d_w(\cdot, t)\|_{L^2(\Gamma_h^*(t))} &= \|\mathbf{d_w}(t)\|_{\mathbf{M}(\mathbf{x}^*(t))} \le C h^k, \\ \|\partial_h^\bullet d_u(\cdot, t)\|_{H_h^{-1}(\Gamma_h^*(t))} &= \|\dot{\mathbf{d_u}}(t)\|_{\star, \mathbf{x}^*(t)} \le C h^k, \\ \|\partial_h^\bullet d_w(\cdot, t)\|_{H_h^{-1}(\Gamma_h^*(t))} &= \|\dot{\mathbf{d_w}}(t)\|_{\star, \mathbf{x}^*(t)} \le C h^k. \end{aligned} \tag{6.2}$$

The constants C are independent of h and $t \in [0, T]$.

Because of the large number of terms, we will not give the full proof of this proposition. We will instead consider the construction and properties of z_h^* as an exemplary and actually the most demanding case among the reference finite element functions. This function will be constructed by a modified Ritz map in the next subsection, whereas H_h^*, ν_h^* and V_h^* can be constructed with the required approximation properties via the standard Ritz map. The construction of z_h^* by a modified Ritz projection is required in order to obtain

the $O(h^k)$ estimate for the defect d_ν (where $d_u = (d_H; d_\nu) \in S_h[\mathbf{x}^*]^{1+3}$). We will therefore prove the approximation estimate for z_h^* according to part (a) and the defect estimates for d_ν and its material derivative according to part (b). All the other bounds of Proposition 6.1 are obtained by similar or simpler arguments.

Remark 6.1 The error bounds (6.1) imply that $u_h^*(\cdot, t)$ and $w_h^*(\cdot, t)$ are bounded in the $W^{1,\infty}$ norm uniformly in h and $t \in [0, T]$. This follows from the equivalent error bound $\|u_h^*(\cdot, t) - \widehat{I}_h^* u(\cdot, t)\|_{H^1(\Gamma_h^*(t))} \leq C h^k$ and the inverse estimate $\|\varphi_h\|_{W^{1,\infty}(\Gamma_h^*(t))} \leq c h^{-1} \|\varphi_h\|_{H^1(\Gamma_h^*(t))}$ for $\varphi_h \in S_h(\Gamma_h^*(t))$, where c depends only on shape regularity and quasi-uniformity of the triangulation and on the regularity and size of $\Gamma(t)$.

Proposition 6.2 *Under the assumptions of Theorem 4.1, the initial errors are bounded by*

$$\|\widetilde{I}_h u(\cdot, 0) - u_h^*(\cdot, 0)\|_{H^1(\Gamma_h^*(0))} = \|\mathbf{e_u}(0)\|_{\mathbf{K}(\mathbf{x}^0)} \leq C h^k,$$
$$\|\widetilde{I}_h w(\cdot, 0) - w_h^*(\cdot, 0)\|_{H^1(\Gamma_h^*(0))} = \|\mathbf{e_w}(0)\|_{\mathbf{K}(\mathbf{x}^0)} \leq C h^k, \qquad (6.3)$$
$$\|\widetilde{I}_h w(\cdot, 0) - \bar{w}_h(\cdot, 0)\|_{H_h^{-1}(\Gamma_h^*(0))} = \|\bar{\mathbf{e}}_\mathbf{w}(0)\|_{\star, \mathbf{x}^0} \leq C h^k,$$

where $\widetilde{I}_h : C(\Gamma^0) \to S_h[\mathbf{x}^*(0)]$ is the finite element interpolation operator and $\bar{w}_h(\cdot, 0) = (\bar{V}_h(\cdot, 0); \bar{z}_h(\cdot, 0)) \in S_h[\mathbf{x}^*(t)]^4$ is the finite element function defined by (3.2b) and (3.2d) at $t = 0$, that is, it has the nodal vector $\bar{\mathbf{w}}(0)$ defined in Section 3.4.

Here we note that the first two bounds (for $\mathbf{e_u}(0)$ and $\mathbf{e_w}(0)$) follow directly from the error bounds (6.1) at $t = 0$, the known error bounds for interpolation and the equivalence of the H^1-norms of a function on the interpolated surface and its lift on the exact surface [DE07]. The third bound will be proved for the z-component of $\bar{w}_h = (\bar{V}_h, \bar{z}_h)$ in Section 6.5. The proof for the V-component is completely analogous and is therefore omitted.

After combining Propositions 6.1 and 6.2 with the stability estimates of Proposition 5.1, our main result Theorem 4.1 is proved by the same short argument as in [KLL19, Section 9].

6.2 Construction of z_h^* by a modified Ritz map

The defect $d_\nu(\cdot, t) \in S_h(\Gamma_h^*(t))^3$ is defined as the finite element function on $\Gamma_h^*(t)$ with nodal vector $\mathbf{d}_\nu(t)$, which comprises the last $3N$ components of the defect vector $\mathbf{d_u}(t) = (\mathbf{d_H}(t); \mathbf{d}_\nu(t)) \in \mathbb{R}^{4N}$ defined in (5.15a). Translated back into a function setting, d_ν is determined as the defect in (3.2c) on inserting reference finite element functions H_h^*, ν_h^* and V_h^*, z_h^* on $\Gamma_h^*(t) = \Gamma_h[\mathbf{x}^*(t)]$ in place of the numerical solution H_h, ν_h and V_h, z_h on $\Gamma_h(t) = \Gamma_h[\mathbf{x}(t)]$: with

$A_h^* = \frac{1}{2}(\nabla_{\Gamma_h^*}\nu_h^* + (\nabla_{\Gamma_h^*}\nu_h^*)^T)$ and $Q_h^* = -\frac{1}{2}(H_h^*)^3 + |A_h^*|^2 H_h^*$ (and omitting the omnipresent argument t),

$$\int_{\Gamma_h[\mathbf{x}^*]} d_\nu \varphi_h = \int_{\Gamma_h[\mathbf{x}^*]} \partial_h^\bullet \nu_h^* \cdot \varphi_h \tag{6.4}$$

$$- \int_{\Gamma_h[\mathbf{x}^*]} \nabla_{\Gamma_h[\mathbf{x}^*]} z_h^* \cdot \nabla_{\Gamma_h[\mathbf{x}^*]} \varphi_h$$

$$- 2 \int_{\Gamma_h[\mathbf{x}^*]} (A_h^* \nabla_{\Gamma_h[\mathbf{x}^*]} H_h^*) \cdot (\nabla_{\Gamma_h[\mathbf{x}^*]} \varphi_h\, \nu_h^*) - \int_{\Gamma_h[\mathbf{x}^*]} Q_h^* (\nabla_{\Gamma_h[\mathbf{x}^*]} \cdot \varphi_h)$$

$$- \int_{\Gamma_h[\mathbf{x}^*]} \left(|\nabla_{\Gamma_h[\mathbf{x}^*]} H_h^*|^2 \nu_h^* + (A_h^*)^2 \nabla_{\Gamma_h[\mathbf{x}^*]} H_h^*\right) \cdot \varphi_h$$

$$+ \int_{\Gamma_h[\mathbf{x}^*]} Q_h^* H_h^* \nu_h^* \cdot \varphi_h - \int_{\Gamma_h[\mathbf{x}^*]} (H_h^* A_h^* - (A_h^*)^2)\, z_h^* \cdot \varphi_h$$

for all $\varphi_h \in S_h[\mathbf{x}^*]^3$. As we want to bound the L^2 norm of d_ν, we can expect difficulties from the terms in the second and third line of (6.4), which contain the surface gradient and divergence of the test function φ_h. We subtract equation (2.12c) for the exact solution ν from (6.4):

$$\int_{\Gamma_h[\mathbf{x}^*]} d_\nu \varphi_h = \int_{\Gamma_h[\mathbf{x}^*]} \partial_h^\bullet \nu_h^* \cdot \varphi_h - \int_{\Gamma[X]} \partial^\bullet \nu \cdot \varphi_h^l \tag{6.5}$$

$$+ \left(- \int_{\Gamma_h[\mathbf{x}^*]} \nabla_{\Gamma_h[\mathbf{x}^*]} z_h^* \cdot \nabla_{\Gamma_h[\mathbf{x}^*]} \varphi_h - 2 \int_{\Gamma_h[\mathbf{x}^*]} (A_h^* \nabla_{\Gamma_h[\mathbf{x}^*]} H_h^*) \cdot (\nabla_{\Gamma_h[\mathbf{x}^*]} \varphi_h\, \nu_h^*) \right.$$

$$- \int_{\Gamma_h[\mathbf{x}^*]} Q_h^* (\nabla_{\Gamma_h[\mathbf{x}^*]} \cdot \varphi_h)$$

$$+ \int_{\Gamma[X]} \nabla_{\Gamma[X]} z \cdot \nabla_{\Gamma[X]} \varphi_h^l + 2 \int_{\Gamma[X]} (A \nabla_{\Gamma[X]} H) \cdot (\nabla_{\Gamma[X]} \varphi_h^l\, \nu)$$

$$\left. + \int_{\Gamma[X]} Q\, \nabla_{\Gamma[X]} \cdot \varphi_h^l \right)$$

$$- \int_{\Gamma_h[\mathbf{x}^*]} \left(|\nabla_{\Gamma_h[\mathbf{x}^*]} H_h^*|^2 \nu_h^* + (A_h^*)^2 \nabla_{\Gamma_h[\mathbf{x}^*]} H_h^*\right) \cdot \varphi_h$$

$$+ \int_{\Gamma[X]} \left(|\nabla_{\Gamma[X]} H|^2 \nu + A^2 \nabla_{\Gamma[X]} H\right) \cdot \varphi_h^l$$

$$+ \int_{\Gamma_h[\mathbf{x}^*]} Q_h^* H_h^* \nu_h^* \cdot \varphi_h - \int_{\Gamma[X]} Q H \nu \cdot \varphi_h^l$$

$$- \int_{\Gamma_h[\mathbf{x}^*]} (H_h^* A_h^* - (A_h^*)^2)\, z_h^* \cdot \varphi_h + \int_{\Gamma[X]} (HA - A^2)\, z \cdot \varphi_h^l$$

for all $\varphi_h \in S_h[\mathbf{x}^*]^3$. The critical terms in the big bracket from the second to the fifth line are replaced by harmless terms, which do not contain a gradient or divergence of the test function, when we define $z_h^* \in S_h[\mathbf{x}^*]^3$ by a *modified*

Ritz map: z_h^* is determined from z (and Q_h^* and Q) by

$$\int_{\Gamma_h[\mathbf{x}^*]} \big(\nabla_{\Gamma_h[\mathbf{x}^*]} z_h^* \cdot \nabla_{\Gamma_h[\mathbf{x}^*]} \varphi_h + z_h^* \varphi_h\big) = \int_{\Gamma[X]} \big(\nabla_{\Gamma[X]} z \cdot \nabla_{\Gamma[X]} \varphi_h^l + z \varphi_h^l\big)$$

$$+ \left(2 \int_{\Gamma_h[\mathbf{x}^*]} (A_h^* \nabla_{\Gamma_h[\mathbf{x}^*]} H_h^*) \cdot (\nabla_{\Gamma_h[\mathbf{x}^*]} \varphi_h \, \nu_h^*) - 2 \int_{\Gamma[X]} (A \nabla_{\Gamma[X]} H) \cdot (\nabla_{\Gamma[X]} \varphi_h^l \nu) \right)$$

$$+ \left(\int_{\Gamma_h[\mathbf{x}^*]} Q_h^* (\nabla_{\Gamma_h[\mathbf{x}^*]} \cdot \varphi_h) - \int_{\Gamma[X]} Q \, \nabla_{\Gamma[X]} \cdot \varphi_h^l \right) \tag{6.6}$$

for all $\varphi_h \in S_h[\mathbf{x}^*]^3$. A similar modified Ritz map, also motivated by a perturbation term containing the gradient of the test function, was previously used and analysed in [LM15]. For later use we can now restate the simplified version of (6.5),

$$\int_{\Gamma_h[\mathbf{x}^*]} d_\nu \varphi_h = \int_{\Gamma_h[\mathbf{x}^*]} \partial_h^\bullet \nu_h^* \cdot \varphi_h - \int_{\Gamma[X]} \partial^\bullet \nu \cdot \varphi_h^l \tag{6.7}$$

$$+ \int_{\Gamma_h[\mathbf{x}^*]} z_h^* \cdot \varphi_h - \int_{\Gamma[X]} z \cdot \varphi_h^l$$

$$- \int_{\Gamma_h[\mathbf{x}^*]} \big(|\nabla_{\Gamma_h[\mathbf{x}^*]} H_h^*|^2 \nu_h^* + (A_h^*)^2 \nabla_{\Gamma_h[\mathbf{x}^*]} H_h^*\big) \cdot \varphi_h$$

$$+ \int_{\Gamma[X]} \big(|\nabla_{\Gamma[X]} H|^2 \nu + A^2 \nabla_{\Gamma[X]} H\big) \cdot \varphi_h^l$$

$$+ \int_{\Gamma_h[\mathbf{x}^*]} Q_h^* H_h^* \nu_h^* \cdot \varphi_h - \int_{\Gamma[X]} Q H \, \nu \cdot \varphi_h^l$$

$$- \int_{\Gamma_h[\mathbf{x}^*]} (H_h^* A_h^* - (A_h^*)^2) \, z_h^* \cdot \varphi_h + \int_{\Gamma[X]} (H A - A^2) \, z \cdot \varphi_h^l$$

for all $\varphi_h \in S_h[\mathbf{x}^*]^3$, where, thanks to the modified Ritz map (6.6), no gradients of the test function appear.

We note that the equation for ν is the only equation in (2.12) for which the right-hand side contains the gradient or the divergence of the test function. The other reference finite element functions H_h^*, ν_h^*, V_h^* are therefore defined by the standard Ritz map (which only contains the first line in (6.6)). We further note that Q_h^* is thus defined independently of z_h^*, and so all the reference finite element functions $H_h^*, \nu_h^*, V_h^*, z_h^*$ are well defined.

6.3 Error bounds of the reference finite element functions

For the functions defined by the standard Ritz map, the optimal-order H^1 error bounds of [Kov18, Theorems 6.3–6.4] yield, under the regularity assumptions

of Theorem 4.1,

$$\|(\nu_h^*)^l(\cdot,t) - \nu(\cdot,t)\|_{H^1(\Gamma(t))} \le C\,h^k$$
$$\|(\partial_h^\bullet \nu_h^*)^l(\cdot,t) - \partial^\bullet \nu(\cdot,t)\|_{H^1(\Gamma(t))} \le C\,h^k \qquad (6.8)$$
$$\|(\partial_h^\bullet \partial_h^\bullet \nu_h^*)^l(\cdot,t) - \partial^\bullet \partial^\bullet \nu(\cdot,t)\|_{H^1(\Gamma(t))} \le C\,h^k$$

and analogously for H_h^* and V_h^*. These error bounds also yield, by an inverse estimate, that these reference finite element functions functions have $W^{1,\infty}$ norms bounded independently of h. These facts, together with the equivalence of norms on $\Gamma(t)$ and $\Gamma_h^*(t)$, give us an L^2 error bound for Q_h^*:

$$\|(Q_h^*)^l(\cdot,t) - Q(\cdot,t)\|_{L^2(\Gamma(t))} \le C\,h^k.$$

We then rewrite the last term in (6.6) as

$$\int_{\Gamma_h[\mathbf{x}^*]} Q_h^*(\nabla_{\Gamma_h[\mathbf{x}^*]} \cdot \varphi_h) - \int_{\Gamma[X]} (Q_h^*)^l \nabla_{\Gamma[X]} \cdot \varphi_h^l$$
$$+ \int_{\Gamma[X]} (Q_h^*)^l \nabla_{\Gamma[X]} \cdot \varphi_h^l - \int_{\Gamma[X]} Q \nabla_{\Gamma[X]} \cdot \varphi_h^l,$$

where the difference in the second line is bounded by $C\,h^k \|\nabla_{\Gamma[X]} \cdot \varphi_h^l\|_{L^2(\Gamma[X])}$ in view of the L^2 error bound of Q_h^*. The first term has a bound of the same type by the higher-order version of [DE13b, Lemma 5.5] that follows with the geometric error bound of [Kov18, Lemma 5.2].

The difference of the integrals in the second line of (6.6) is bounded by $C\,h^k \|\nabla_{\Gamma[X]} \varphi_h^l\|_{L^2(\Gamma[X])}$ by the same arguments.

With these error bounds for the extra terms in the modified Ritz map, the proof of [LM15, Theorems 8.2 and 8.3] together with the higher-order geometric perturbation error bounds of [Kov18, Lemma 5.6] yields the error bounds

$$\|(z_h^*)^l(\cdot,t) - z(\cdot,t)\|_{H^1(\Gamma(t))} \le C\,h^k$$
$$\|\partial_h^\bullet (z_h^*)^l(\cdot,t) - \partial^\bullet z(\cdot,t)\|_{H^1(\Gamma(t))} \le C\,h^k. \qquad (6.9)$$

6.4 Bounds for the defect d_ν and its material derivative

The techniques of the proof of the defect bounds have already been used for the defect bounds in [KLLP17] and [KLL19], but we now have the additional difficulty to prove estimates for the time-differentiated defects.

The terms in (6.7) do not contain the gradient of the test function, and this allows us obtain estimates of d_ν in the L^2 norm. All the pairs of terms in (6.7) can be bounded by the same arguments as for the corresponding terms in [KLLP17, Section 8], using the Ritz map error bounds of the previous subsection instead of the bound for the interpolation error where required. These arguments are based on geometric estimates that were previously proved in [Dem09, Dzi88, DE07, DE13b, Kov18]. We refer to [KLLP17] for the details.

To bound the material derivative $\partial_h^\bullet d_\nu$ in the H_h^{-1} norm, we differentiate (6.5) in time for time-dependent test functions $\varphi_h(\cdot, t) \in S_h[\mathbf{x}^*(t)]^3$ with time-independent nodal vector and use the Leibniz formula to obtain

$$\int_{\Gamma_h^*} \partial_h^\bullet d_\nu \, \varphi_h = \frac{\mathrm{d}}{\mathrm{d}t} \int_{\Gamma_h^*} d_\nu \, \varphi_h - \int_{\Gamma_h^*} d_\nu \, (\nabla_{\Gamma_h[\mathbf{x}^*]} \cdot v_h^*) \, \varphi_h,$$

where we used that $\partial_h^\bullet \varphi_h = 0$. The last term is readily estimated using the L^2 bound of d_ν. For the first term on the right-hand side, we insert (6.7) and differentiate each pair of terms using the Leibniz rule. With the Ritz map error bounds (6.8) for the material derivatives of ν_h^* and H_h^* and again with the arguments of [KLLP17] we then obtain the stated error bound for $\partial_h^\bullet d_\nu$.

6.5 Error bound of $\bar{z}_h(\cdot, 0)$

By the construction of $\bar{z}_h(\cdot, 0)$ from (3.2d), we have (omitting in the following the argument $t = 0$)

$$\int_{\Gamma_h^0} \bar{z}_h \cdot \varphi_h = -\int_{\Gamma_h^0} \nabla_{\Gamma_h^0} \nu_h \cdot \nabla_{\Gamma_h^0} \varphi_h + \int_{\Gamma_h^0} |A_h|^2 \nu_h \cdot \varphi_h$$

for all $\varphi_h \in S_h(\Gamma_h^0)^3$. Here ν_h is the finite element interpolation of ν on Γ_h^0 and $A_h = \frac{1}{2}(\nabla_{\Gamma_h^0} \nu_h + (\nabla_{\Gamma_h^0} \nu_h)^T)$. This is to be compared with the interpolation of $z(0)$, which satisfies (2.12d), that is,

$$\int_{\Gamma^0} z \cdot \varphi = -\int_{\Gamma^0} \nabla_{\Gamma[X]} \nu \cdot \nabla_{\Gamma^0} \varphi + \int_{\Gamma^0} |A|^2 \nu \cdot \varphi$$

for all $\varphi_h \in H^1(\Gamma^0)$. So we obtain

$$\int_{\Gamma_h^0} (\bar{z}_h - \widetilde{I}_h z) \cdot \varphi_h = -\int_{\Gamma_h^0} \nabla_{\Gamma_h^0} \nu_h \cdot \nabla_{\Gamma_h^0} \varphi_h + \int_{\Gamma^0} \nabla_{\Gamma[X]} \nu \cdot \nabla_{\Gamma^0} \varphi_h^l$$
$$+ \int_{\Gamma_h^0} |A_h|^2 \nu_h \cdot \varphi_h - \int_{\Gamma^0} |A|^2 \nu \cdot \varphi_h^l - \int_{\Gamma_h^0} \widetilde{I}_h z \cdot \varphi_h + \int_{\Gamma^0} z \cdot \varphi_h^l$$

for all $\varphi_h \in S_h(\Gamma_h^0)^3$. The pairs of terms on the right-hand side can be estimated in the same way as similar pairs in the defects, and so we obtain the last bound of (6.3) for $\bar{z}_h(\cdot, 0)$ in the $H_h^{-1}(\Gamma_h^0)$ norm.

7 Numerical experiments

We performed the following numerical experiments for Willmore flow:

- Convergence tests using stationary solutions of Willmore flow, i.e. a sphere and a Clifford torus.

- We report on the Willmore energy of a few surfaces from the literature, e.g. [Dzi08] and [BGN08], as they evolve towards the stationary solution, minimising the energy.

All our numerical experiments use quadratic evolving surface finite elements. The numerical computations were carried out in Matlab. The initial meshes for all surfaces were generated using DistMesh [PS04], without taking advantage of the symmetries of the surfaces. For time discretization we use the linearly implicit backward difference formula (BDF) of order 2 applied to the system (3.4)–(3.5) in matrix-vector formulation, in the same way as described in [KLL19].

7.1 Convergence tests

We computed approximations to Willmore flow for two surfaces that are *stationary* along Willmore flow (though not along the discretized flow): a sphere of radius $R = 1$ ($W(\Gamma) = 8\pi \approx 25.1327$), and a Clifford torus ($W(\Gamma) = 4\pi^2 \approx 39.4784$), i.e. a torus with the quotient of radii $R/r = \sqrt{2}$, in our case with major radius $R = 1$ and minor radius $r = 1/\sqrt{2}$. The considered time interval is $[0, T] = [0, 1]$ in both cases.

For our convergence tests we used a sequence of time step sizes $\tau_k = \tau_{k-1}/2$ with $\tau_0 = 0.2$, and a sequence of meshes with mesh widths $h_k \approx 2^{-1/2}h_{k-1}$. For the Clifford torus, in order to reasonably resolve the surface, we used graded meshes that are more refined around the hole.

We started the time integration from the nodal interpolations of the exact initial values $\nu(\cdot, 0)$, $H(\cdot, 0)$, and $V(\cdot, 0) = 0$, $z(\cdot, 0) = \nabla_{\Gamma^0} H(\cdot, 0)$, which were computed analytically.

In Figure 7.1 we report on the errors between the numerical and exact solutions for the Willmore flow of a sphere until the final time $T = 1$ (the first three plots from left to right: the surface error and the errors of the dynamic variables ν and H), and on the Willmore energy of the discrete surfaces at the final time $T = 1$ (rightmost plot). The logarithmic error plots show the $L^\infty(H^1)$ norm errors, and the Willmore energy (the plot is semi-logarithmic) against the mesh width h. The lines marked with different symbols correspond to different time step sizes.

In order to preserve important properties of the functions in the approximation, it is recommended to *project* the computed surface normal $\nu_h \approx \nu$ onto the unit sphere and the function $z_h \approx \nabla_\Gamma H$ onto the tangent space. This is done after every time step. The advantages of a normalising projection for ν_h were already explored in numerical experiments for mean curvature flow in [KLL19].

Figure 7.2 reports on the effect of the projections using a sphere of radius 1: the left plots show the minimum ($\circ$) and maximum ($*$) lengths of the normal vectors, the middle plots the roundness of the surface (minimum ($\circ$) and maximum ($*$) distance from origin), and the right plots show the Willmore energy along the flows.
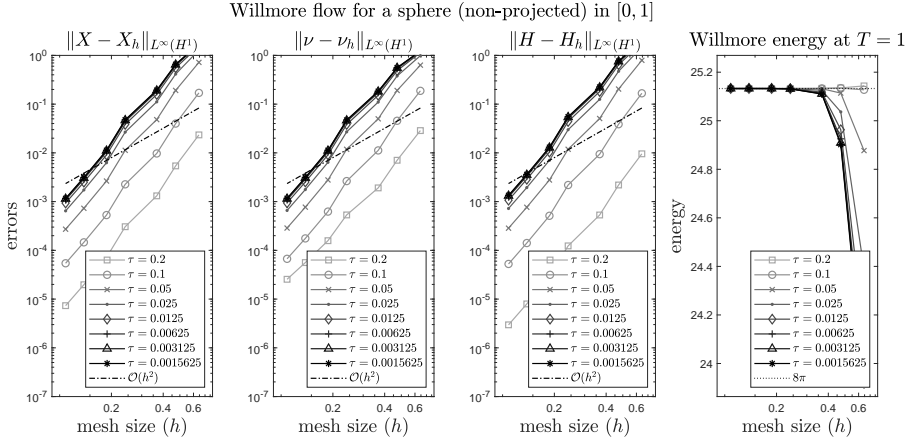
Fig. 7.1 Spatial convergence of the BDF2 / quadratic ESFEM discretization (without projecting nu_h and z_h) for the Willmore flow of a sphere for $T = 1$.
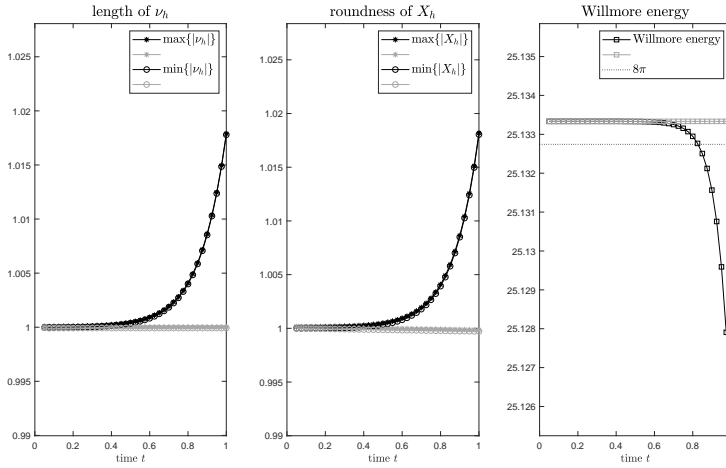

Fig. 7.2 Comparing the effect of projecting ν_h and z_h onto the unit sphere and the tangent plane, respectively, (black without projections, grey with projections) using a sphere (dof = 1914, $\tau = 0.025$) reporting on length of normal ν_h, surface roundness (i.e. radius), and Willmore energy along the flow in $[0, 1]$.

Figure 7.3 and 7.4 report on the same experiments for the algorithm *with projections* for the sphere and the Clifford torus, respectively. For the Clifford torus we used the linearly implicit Euler method (BDF1), whose damping property proved to be advantageous for this delicate experiment.

In Figures 7.1, 7.3 and 7.4 the spatial discretization error dominates, matching (exceeding even) the $O(h^2)$ order of convergence (note the reference lines), in agreement with our theoretical results in Theorem 4.1.
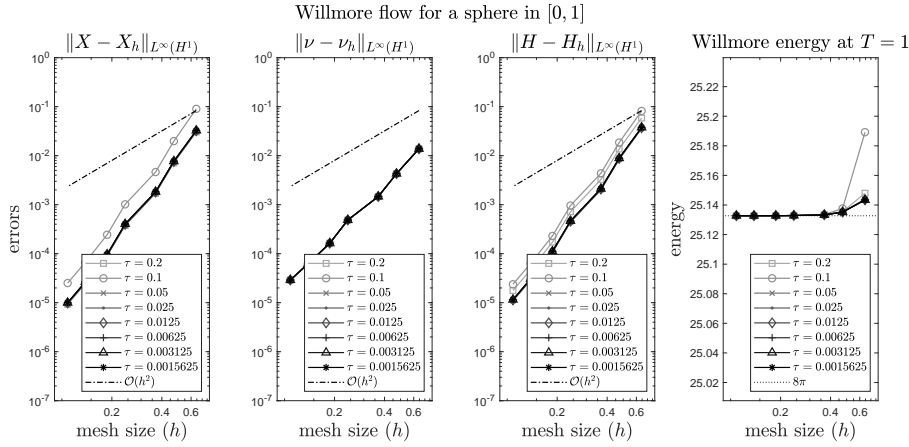
Fig. 7.3 Spatial convergence of the BDF2 / quadratic ESFEM discretization (with projection of ν_h and z_h onto the unit sphere and the tangent plane, respectively) for the Willmore flow of a sphere.
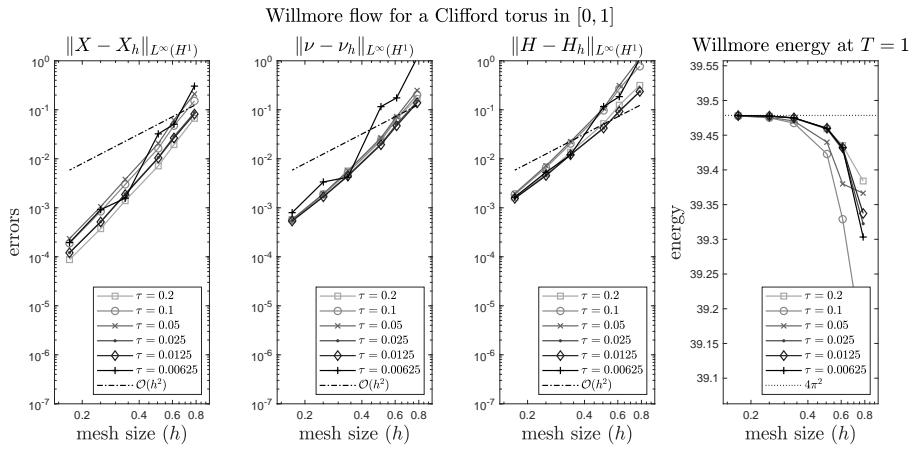

Fig. 7.4 Spatial convergence of the BDF1 / quadratic ESFEM discretization (with projection of ν_h and z_h onto the unit sphere and the tangent plane, respectively) for the Willmore flow of a Clifford torus.

7.2 Willmore flow towards stationary solutions

In Figure 7.5 we report on experiments for the surface evolution and Willmore energy for three surfaces: an ellipsoid, a surface of the shape of a red blood cell (cf. [Dzi08] and [BGN08]), and a periodically perturbed torus (cf. [Dzi08]). The meshes have 4978, 8978, and 12000 nodes, respectively, and were integrated in $[0, 4]$ with a time step size $\tau = 0.2 \cdot 2^{-8} \approx 7.8 \cdot 10^{-4}$. In the last two rows, times $t = 3$ and 4, it can be observed that the algorithm leaves the surface stationary.
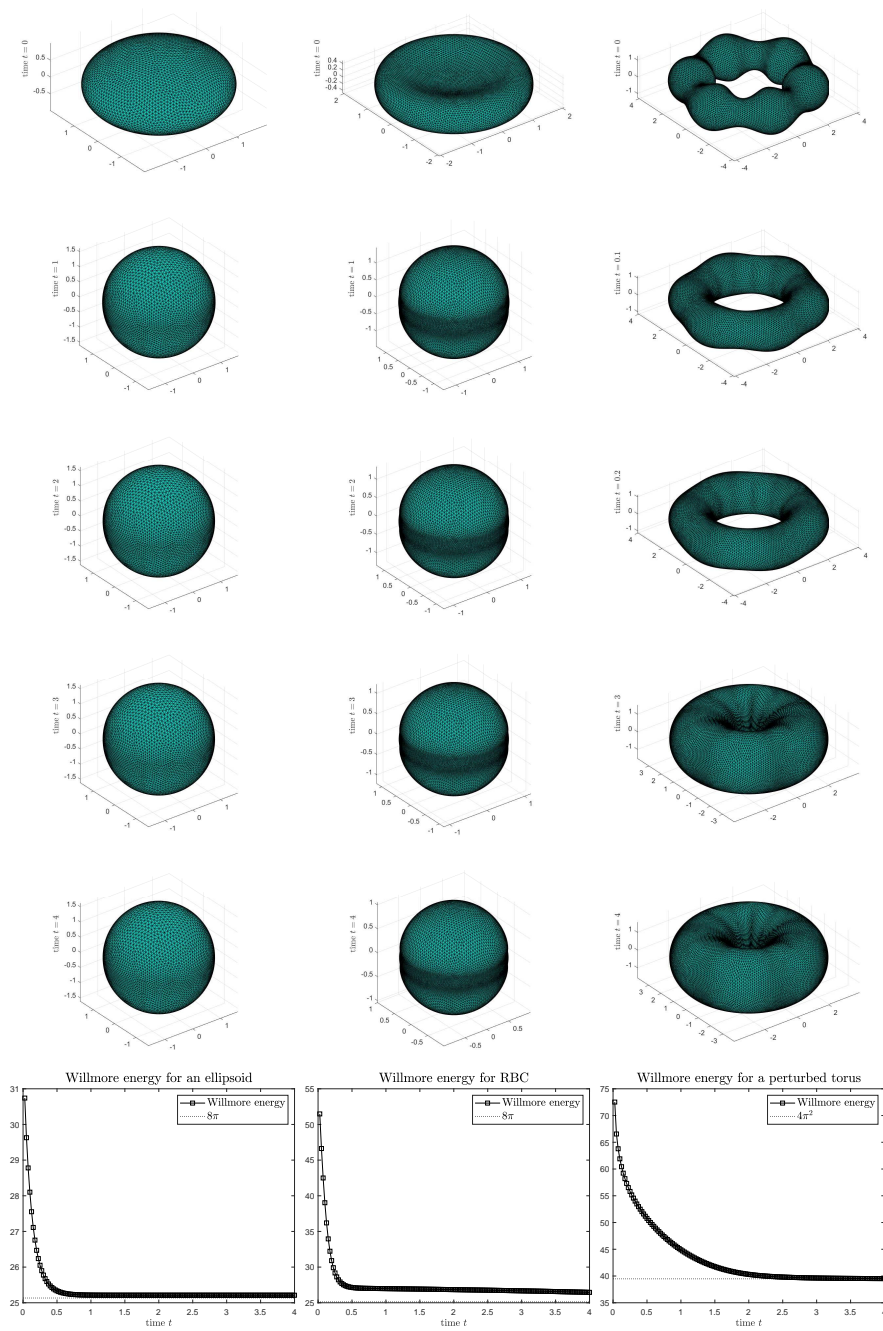
Fig. 7.5 Surface evolutions at different times and corresponding Willmore energy

Appendix: Proof of Lemma 2.2

We came up with two independent proofs of Lemma 2.2, one based on local coordinates and the other one based on the formalism of Dziuk and Elliott in [DE13a]. As should be, both approaches yield the same result as stated in Lemma 2.2. Here we present the more straightforward proof based on [DE13a].

The following commutator formula for tangential differential operators $(D_i = (\nabla_\Gamma)_i)$ is shown in [DE13a, Lemma 2.6]:

$$D_i D_j u - D_j D_i u = A_{jk} D_k u \nu_i - A_{ik} D_k u \nu_j \qquad (i, j = 1, 2, 3), \qquad (7.1)$$

where we use the convention to sum over repeated indices. We also recall that $A_{ij} = D_i \nu_j = A_{ji}$ and $H = \mathrm{tr}(\nabla_\Gamma \nu) = D_i \nu_i$.

Using (7.1) we obtain (cf. the proof of Lemma 3.2 in [DE13a])

$$\begin{aligned}
D_i D_i D_j u &= D_i\big(D_j D_i u + A_{jk} D_k u \nu_i - A_{ik} D_k u \nu_j\big) \\
&= D_i D_j D_i u + D_i\big(D_k u (A_{jk} \nu_i - A_{ik} \nu_j)\big) \\
&= D_i D_j D_i u + D_i D_k u\big(A_{jk} \nu_i - A_{ik} \nu_j\big) \\
&\qquad + D_k u D_i\big(A_{jk} \nu_i - A_{ik} \nu_j\big).
\end{aligned}$$

The second term is then rewritten as

$$\begin{aligned}
D_i D_k u\big(A_{jk} \nu_i - A_{ik} \nu_j\big) &= A_{jk} D_i D_k u \nu_i - A_{ik} D_i D_k u \nu_j \\
&= A_{jk}\big(D_k D_i u + A_{kl} D_l u \nu_i - A_{il} D_l u \nu_k\big)\nu_i - A_{ik} D_i D_k u \nu_j \\
&= A_{jk}(D_k D_i u \nu_i) + A_{jk} A_{kl} D_l u \nu_i \nu_i - A_{jk} A_{il} D_l u \nu_k \nu_i - A_{ik} D_i D_k u \nu_j \\
&= -A_{jk} A_{ki} D_i u + A_{jk} A_{kl} D_l u - A_{jk} A_{il} D_l u \nu_k \nu_i - A_{ik} D_i D_k u \nu_j \\
&= -(A^2)_{ji} D_i u + (A^2)_{jl} D_l u - A_{jk} A_{il} D_l u \nu_k \nu_i - A_{ik} D_i D_k u \nu_j \\
&= -A_{jk} A_{il} D_l u \nu_k \nu_i - A_{ik} D_i D_k u \nu_j \\
&= -A_{ik} D_i D_k u \nu_j.
\end{aligned}$$

In particular for the last term we have:

$$\begin{aligned}
D_i\big(A_{jk} \nu_i - A_{ik} \nu_j\big) &= D_i A_{jk} \nu_i - D_i A_{ik} \nu_j + A_{jk} D_i \nu_i - A_{ik} D_i \nu_j \\
&= D_i A_{jk} \nu_i - D_i A_{ik} \nu_j + A_{jk} H - A_{ki} A_{ij} \\
&= -\Delta_\Gamma \nu_k \nu_j + (HA - A^2)_{jk}.
\end{aligned}$$

The third-order term is again rewritten using (7.1) as

$$D_i D_j D_i u = D_j D_i D_i u + A_{jk} D_k D_i u \nu_i - A_{ik} D_k D_i u \nu_j,$$

which is further rewritten using

$$\begin{aligned}
D_k D_i u \nu_i &= D_k(D_i u \nu_i) - D_k \nu_i D_i u = -A_{ki} D_i u, \\
D_k D_i u \nu_j &= D_k(D_i u \nu_j) - D_k \nu_j D_i u.
\end{aligned}$$

Therefore,

$$D_i D_j D_i u = D_j D_i D_i u - A_{jk} A_{ki} D_i u - A_{ik} D_k D_i u \nu_j$$
$$= D_j D_i D_i u - (A^2)_{ji} D_i u - A_{ik} D_k D_i u \nu_j.$$

Altogether, the above calculations yield

$D_i D_i D_j u$
$= D_j D_i D_i u - (A^2)_{ji} D_i u - 2 A_{ik} D_k D_i u \nu_j - \Delta_\Gamma \nu_k \nu_j D_k u + (HA - A^2)_{jk} D_k u$
$= D_j D_i D_i u - (A^2)_{ji} D_i u - 2 D_k(A_{ik} D_i u) \nu_j + 2 D_k A_{ik} D_i u \nu_j$
$\quad - \Delta_\Gamma \nu_k \nu_j D_k u + (HA - A^2)_{jk} D_k u$
$= D_j D_i D_i u - (A^2)_{ji} D_i u - 2 D_k(A_{ik} D_i u) \nu_j + \Delta_\Gamma \nu_k \nu_j D_k u + (HA - A^2)_{jk} D_k u.$

This implies

$$\Delta_\Gamma \nabla_\Gamma u = \nabla_\Gamma \Delta_\Gamma u - A^2 \nabla_\Gamma u - 2 \nabla_\Gamma \cdot (A \nabla_\Gamma u) \nu$$
$$+ (\Delta_\Gamma \nu \cdot \nabla_\Gamma u) \nu + (HA - A^2) \nabla_\Gamma u,$$

which is the same as (2.6). □

Acknowledgement

References

Bar13.   S. Bartels. A simple scheme for the approximation of the elastic flow of inexten-
            sible curves. *IMA J. Numer. Anal.*, 33(4):1115–1125, 2013.
BGN07a.   J.W. Barrett, H. Garcke, and R. Nürnberg. A parametric finite element method
            for fourth order geometric evolution equations. *J. Comput. Phys.*, 222(1):441–
            467, 2007.
BGN07b.   J.W. Barrett, H. Garcke, and R. Nürnberg. On the variational approximation of
            combined second and fourth order geometric evolution equations. *SIAM J. Sci.
            Comput.*, 29(3):1006–1041, 2007.

BGN08.   J.W. Barrett, H. Garcke, and R. Nürnberg. Parametric approximation of Will-more flow and related geometric evolution equations. *SIAM J. on Sci. Comput.*, 31(1):225–253, 2008.

BGN15.   J.W. Barrett, H. Garcke, and R. Nürnberg. Numerical computations of the dynamics of fluidic membranes and vesicles. *Phys. Rev. E*, 92(5):052704, 2015.

BGN20.   J.W. Barrett, H. Garcke, and R. Nürnberg. Parametric finite element approximations of curvature driven interface evolutions. In *Handbook of Numerical Analysis*, volume 21, pages 275–423. 2020.

BJSW17.   W. Bao, W. Jiang, D. J. Srolovitz, and Y. Wang. Stable equilibria of anisotropic particles on substrates: a generalized Winterbottom construction. *SIAM J. Appl. Math.*, 77(6):2093–2118, 2017.

BJWZ17.   W. Bao, W. Jiang, Y. Wang, and Q. Zhao. A parametric finite element method for solid-state dewetting problems with anisotropic surface energies. *J. Comput. Phys.*, 330:380–400, 2017.

Bla29.   W. Blaschke. *Vorlesungen über Differentialgeometrie III*. Grundlehren der mathematischen Wissenschaften. Springer-Verlag, Berlin, 1929.

BMN04.   E. Bänsch, P. Morin, and R. H. Nochetto. Surface diffusion of graphs: variational formulation, error analysis, and simulation. *SIAM J. Numer. Anal.*, 42(2):773–799, 2004.

BMN05.   E. Bänsch, P. Morin, and R. H. Nochetto. A finite element method for surface diffusion: the parametric case. *J. Comput. Phys.*, 203(1):321–343, 2005.

BNP10.   A. Bonito, R. H. Nochetto, and M. S. Pauletti. Parametric FEM for geometric biomembranes. *J. Comput. Phys.*, 229(9):3171–3188, 2010.

BZ21.   Weizhu Bao and Quan Zhao. A structure-preserving parametric finite element method for surface diffusion. *arXiv:2104.01432*, 2021.

CENC96.   J.W. Cahn, C.M. Elliott, and A. Novick-Cohen. The Cahn-Hilliard equation with a concentration dependent mobility: motion by minus the Laplacian of the mean curvature. *European J. Appl. Math.*, 7(3):287–301, 1996.

CLS$^+$18.   Y. Chen, J. Lowengrub, J. Shen, C. Wang, and S. Wise. Efficient energy stable schemes for isotropic and strongly anisotropic Cahn–Hilliard systems with the Willmore regularization. *J. Comput. Phys.*, 365:56–73, 2018.

DD06.   K. Deckelnick and G. Dziuk. Error analysis of a finite element method for the Willmore flow of graphs. *Interfaces Free Bound.*, 8(1):21–46, 2006.

DD09.   K. Deckelnick and G. Dziuk. Error analysis for the elastic flow of parametrized curves. *Math. Comp.*, 78(266):645–671, 2009.

DDE03.   K. Deckelnick, G. Dziuk, and C.M. Elliott. Error analysis of a semidiscrete numerical scheme for diffusion in axially symmetric surfaces. *SIAM J. Numer. Anal.*, 41(6):2161–2179, 2003.

DDE05a.   K. Deckelnick, G. Dziuk, and C. M. Elliott. Computation of geometric partial differential equations and mean curvature flow. *Acta Numerica*, 14:139–232, 2005.

DDE05b.   K. Deckelnick, G. Dziuk, and C.M. Elliott. Fully discrete finite element approximation for anisotropic surface diffusion of graphs. *SIAM J. Numer. Anal.*, 43(3):1112–1138, 2005.

DE07.   G. Dziuk and C.M. Elliott. Finite elements on evolving surfaces. *IMA J. Numer. Anal.*, 27(2):262–292, 2007.

DE13a.   G. Dziuk and C.M. Elliott. Finite element methods for surface PDEs. *Acta Numerica*, 22:289–396, 2013.

DE13b.   G. Dziuk and C.M. Elliott. L^2–estimates for the evolving surface finite element method. *Math. Comp.*, 82(281):1–24, 2013.

Dem09.   A. Demlow. Higher–order finite element methods and pointwise error estimates for elliptic problems on surfaces. *SIAM J. Numer. Anal.*, 47(2):805–807, 2009.

DKS15.   K. Deckelnick, J. Katz, and F. Schieweck. A C^1-finite element method for the Willmore flow of two-dimensional graphs. *Math. Comp.*, 84(296):2617–2643, 2015.

DLM12.   G. Dziuk, C. Lubich, and D.E. Mansour. Runge–Kutta time discretization of parabolic differential equations on evolving surfaces. *IMA J. Numer. Anal.*, 32(2):394–416, 2012.

Dzi88. G. Dziuk. Finite elements for the Beltrami operator on arbitrary surfaces. *Partial differential equations and calculus of variations, Lecture Notes in Math., 1357, Springer, Berlin*, pages 142–155, 1988.

Dzi08. G. Dziuk. Computational parametric Willmore flow. *Numer. Math.*, 111(1):55–80, 2008.

Eck12. K. Ecker. *Regularity theory for mean curvature flow*. Birkhäuser, Boston, 2012.

ES10. C.M. Elliott and B. Stinner. Modeling and computation of two phase geometric biomembranes using surface finite elements. *J. Comput. Phys.*, 229(18):6585–6612, 2010.

Hel73. W. Helfrich. Elastic properties of lipid bilayers: theory and possible experiments. *Zeitschrift für Naturforschung C*, 28(11-12):693–703, 1973.

Hui84. G. Huisken. Flow by mean curvature of convex surfaces into spheres. *J. Differential Geometry*, 20(1):237–266, 1984.

JL21. Wei Jiang and Buyang Li. A perimeter-decreasing and area-conserving algorithm for surface diffusion flow of curves. *arXiv:2102.00374*, 2021.

KLL19. B. Kovács, B. Li, and C. Lubich. A convergent evolving finite element algorithm for mean curvature flow of closed surfaces. *Numer. Math.*, 143(4):797–853, 2019.

KLLP17. B. Kovács, B. Li, C. Lubich, and C.A. Power Guerra. Convergence of finite elements on an evolving surface driven by diffusion on the surface. *Numer. Math.*, 137(3):643–689, 2017.

Kov18. B. Kovács. High-order evolving surface finite element method for parabolic problems on evolving surfaces. *IMA J. Numer. Anal.*, 38(1):430–459, 2018.

KS01. E. Kuwert and R. Schätzle. The Willmore flow with small initial energy. *J. Differential Geom.*, 57(3):409–441, 2001.

KS02. E. Kuwert and R. Schätzle. Gradient flow for the Willmore functional. *Comm. Anal. Geom.*, 10(2):307–339, 2002.

LM15. C. Lubich and D.E. Mansour. Variational discretization of wave equations on evolving surfaces. *Math. Comp.*, 84(292):513–542, 2015.

Man12. C. Mantegazza. *Lecture Notes on Mean Curvature Flow*. Progress in Mathematics, Volume 290. Birkhäuser, Corrected Printing 2012.

MN14. F. C. Marques and A. Neves. Min-Max theory and the Willmore conjecture. *Annals Math.*, 179(2):683–782, 2014.

Mul57. W. W. Mullins. Theory of thermal grooving. *J. Appl. Phys.*, 28:333–339, 1957.

Poz15. P. Pozzi. Computational anisotropic Willmore flow. *Interfaces Free Bound.*, 17(2):189–232, 2015.

PS04. P.-O. Persson and G. Strang. A simple mesh generator in MATLAB. *SIAM Review*, 46(2):329–345, 2004.

PS18. P. Pozzi and B. Stinner. Elastic flow interacting with a lateral diffusion process: the one-dimensional graph case. *IMA J. Numer. Anal.*, 39(1):201–234, 03 2018.

Rus05. R.E. Rusu. An algorithm for the elastic flow of surfaces. *Interfaces and Free Boundaries*, 7(3):229–239, 2005.

Tho24. G. Thomsen. Grundlagen der konformen flächentheorie. *Abh. Math. Seminar Univ. Hamburg*, 3(1):31–56, 1924.

Wal15. S.W. Walker. *The Shape of Things: a Practical Guide to Differential Geometry and the Shape Derivative*. SIAM, Philadelphia, 2015.

Wil65. T.J. Willmore. Note on embedded surfaces. *An. Sti. Univ. "Al. I. Cuza" Iasi Sect. I a Mat.(NS) B*, 11:493–496, 1965.

Wil93. T. J. Willmore. *Riemannian Geometry*. Oxford Science Publications. The Clarendon Press, Oxford University Press, New York, 1993.

ZJB20. Q. Zhao, W. Jiang, and W. Bao. A parametric finite element method for solid-state dewetting problems in three dimensions. *SIAM J. Sci. Comput.*, 42(1):B327–B352, 2020.