Foundations of Computational Mathematics
https://doi.org/10.1007/s10208-023-09622-x

FOUNDATIONS OF COMPUTATIONAL MATHEMATICS
The Journal of the Society for the Foundations of Computational Mathematics

A New Approach to the Analysis of Parametric Finite Element Approximations to Mean Curvature Flow

Genming Bai[1] · **Buyang Li**[1]

Received: 14 February 2023 / Revised: 28 June 2023 / Accepted: 12 July 2023

Abstract
Parametric finite element methods have achieved great success in approximating the evolution of surfaces under various different geometric flows, including mean curvature flow, Willmore flow, surface diffusion, and so on. However, the convergence of Dziuk's parametric finite element method, as well as many other widely used parametric finite element methods for these geometric flows, remains open. In this article, we introduce a new approach and a corresponding new framework for the analysis of parametric finite element approximations to surface evolution under geometric flows, by estimating the projected distance from the numerically computed surface to the exact surface, rather than estimating the distance between particle trajectories of the two surfaces as in the currently available numerical analyses. The new framework can recover some hidden geometric structures in geometric flows, such as the full H^1 parabolicity in mean curvature flow, which is used to prove the convergence of Dziuk's parametric finite element method with finite elements of degree $k \geq 3$ for surfaces in the three-dimensional space. The new framework introduced in this article also provides a foundational mathematical tool for analyzing other geometric flows and other parametric finite element methods with artificial tangential motions to improve the mesh quality.

Keywords Geometric flow · Mean curvature flow · Parametric finite element method · Stability, convergence · Particle trajectory · Projected distance · H^1 parabolicity

---

Communicated by Rob Stevenson.

✉ Buyang Li
buyang.li@polyu.edu.hk

Genming Bai
genming.bai@connect.polyu.hk

[1] Department of Applied Mathematics, The Hong Kong Polytechnic University, Hung Hom, Hong Kong

Mathematics Subject Classification 65M15 · 65M60 · 49M10 · 35K65

Contents

1 Introduction

The evolution of surfaces driven by the curvature on the surfaces under geometric flows, including mean curvature flow, Willmore flow and surface diffusion, has been widely used in modelling the formation of grain boundaries in the annealing of metal, the evolution of soap films, the shape of oil drops on the surface of water, the shape evolution of cell membranes, solid-state dewetting, and so on. Among the geometric flows, mean curvature flow is most natural and intensively studied in geometric analysis in the last decades; see the review article [44] and monographs by Ecker [29] and Mantegazza [42].

In the mean curvature flow of closed surfaces in three-dimensional space, the surface evolves with velocity

$$v = -Hn = \Delta_\Gamma \mathrm{id}, \tag{1.1}$$

where H and n denote the mean curvature and the normal vector of the surface Γ, and Δ_Γ denotes the Laplace–Beltrami operator on the surface, with id denoting the identity function on $\mathbb{R}^3$ restricted to Γ. The numerical approximation of mean curvature flow of closed surfaces was first addressed by Dziuk in his paper [23] published in 1990. He proposed the first parametric finite element algorithm for approximating mean curvature flow: assuming that $\Gamma(t_m)$ is already approximated by a piecewise triangular surface Γ_h^m, find a parametrization of the surface Γ_h^{m+1} at t_{m+1}, denoted by $X_h^{m+1} : \Gamma_h^m \to \mathbb{R}^3$, such that X_h^{m+1} is in a vector-valued finite element space $S_h(\Gamma_h^m)$ and satisfies the following weak formulation:

$$\int_{\Gamma_h^m} \frac{X_h^{m+1} - \mathrm{id}}{\tau} \cdot \chi_h + \int_{\Gamma_h^m} \nabla_{\Gamma_h^m} X_h^{m+1} \cdot \nabla_{\Gamma_h^m} \chi_h = 0 \qquad \forall\, \chi_h \in S_h(\Gamma_h^m), \tag{1.2}$$

where $\tau = t_{m+1} - t_m$ is the stepsize of time discretization. The methods of parametrizing the unknown surface Γ_h^{m+1} by a finite element function on the known surface Γ_h^m are now referred to as parametric finite element methods (FEMs). In the matrix–vector form, Dziuk's parametric FEM can be written as

$$\mathbf{M}(\mathbf{x}^m)\frac{\mathbf{x}^{m+1}-\mathbf{x}^m}{\tau}+\mathbf{A}(\mathbf{x}^m)\mathbf{x}^{m+1}=\mathbf{0}, \tag{1.3}$$

where $\mathbf{x}^m$ denotes the vector which collects the node positions of the finite element surface Γ_h^m, and $\mathbf{M}(\mathbf{x}^m)$ and $\mathbf{A}(\mathbf{x}^m)$ are the mass and stiffness matrices on the surface Γ_h^m. Therefore, at every time level, Dziuk's parametric FEM only requires solving a linear system which is similar to that from the heat equation on the given surface Γ_h^m.

Since Dziuk's paper [23] was published, the parametric FEMs have been widely adopted for approximating the evolution of surfaces under other geometric flows, including surface diffusion, Willmore flow, Helfrich flow, and so on; see [2, 10, 14, 25]. In addition to the parametric FEMs, these problems have also motivated and facilitated the development of many other computational techniques and numerical analysis, including the following:

- The mesh redistribution technique proposed by Bänsch et al. [2] for improving the mesh quality and computational stability in approximating curvature-driven surface evolutions.
- The evolving surface finite element methods developed by Dziuk and Elliott in [26, 27] for solving partial differential equations on evolving surfaces.
- The matrix–vector technique for the analysis of the evolving surface FEMs approximating the surface evolutions under geometric flows; see [37, 39].
- The artificial tangential velocity methods introduced by Barrett, Garcke and Nürnberg [7–9] (the BGN methods), Elliott and Fritz [31] (the DeTurck trick) and Hu and Li[35] for improving the mesh quality without using the mesh redistribution techniques.
- The novel computational methods for simulating the evolution of solid-thin films on a substrate described by anisotropic surface diffusion flow and contact line migration; see [3, 4, 46],

The techniques have largely improved the performance of parametric FEMs in approximating geometric flows and related problems. However, the analysis of convergence of these methods remains challenging. The available numerical analyses can be divided into the following several classes:

- The convergence of parametric and evolving surface FEMs for mean curvature flow and Willmore flow of *curves*: see the work of Dziuk [24], Deckelnick and Dziuk [17, 19], Bartels [11], Elliott and Fritz [31], Li [40], Ye and Cui [45], and so on. The techniques in the proofs are not applicable to the analysis of mean curvature flow or Willmore flow of closed surfaces.
- The convergence of finite element and finite difference methods for mean curvature flow and Willmore flow of *graph surfaces* and *axisymmetric surfaces*; see [5, 16, 18, 20, 21]. The techniques in the proofs are not applicable to the analysis of mean curvature flow or Willmore flow of general closed surfaces.
- The convergence of some level set methods and diffuse-interface methods for mean curvature flow of *closed surfaces* with low-order accuracy in approximating the sharp interface evolution; see [15] and [12, 32, 33]. The techniques in the proofs are not applicable to the analysis of parametric FEMs.

- The convergence of evolving surface FEMs, with tangential velocity based on the DeTurck trick, for mean curvature flow of closed curves, graph surfaces, axisymmetric surfaces, and surfaces of torus type; see [5, 6, 20, 21, 31, 43]. The techniques in the proofs are not directly applicable to the analysis of mean curvature flow or Willmore flow of general closed surfaces.
- The convergence of Dziuk's algorithm for mean curvature flow of *closed surfaces* in the spatial semi-discretization case with sufficiently high-order finite elements of degree $k \geq 6$; see [1, 41]. The techniques are not applicable to lower-order finite elements or other algorithms with artificial tangential velocity.
- The convergence of evolving surface FEMs with finite elements of degree $k \geq 2$ based on different formulations of mean curvature flow and Willmore flow by using the evolution equations of n and H; see [13, 30, 35, 37, 38]. The techniques are not applicable to Dziuk's algorithm and the BGN type of algorithms.

To summarize, the convergence of some foundational algorithms for surface evolution under geometric flows remains open, including Dziuk's fully discrete parametric FEMs and the BGN methods for mean curvature flow, Willmore flow and surface diffusion of closed surfaces.

Currently, the analysis of parametric FEMs for the evolution of surfaces under geometric flows is largely limited by the classical approach which estimates the error along the particle trajectories of the numerically computed surface and the exact surface, i.e., estimating $\mathbf{e}^m = \mathbf{x}^m - \mathbf{x}_*^m$ by comparing equation (1.3) with the corresponding equation (up to some defect term $\mathbf{d}^m$)

$$\mathbf{M}(\mathbf{x}_*^m)\frac{\mathbf{x}_*^{m+1} - \mathbf{x}_*^m}{\tau} + \mathbf{A}(\mathbf{x}_*^m)\mathbf{x}_*^{m+1} = \mathbf{d}^m \tag{1.4}$$

satisfied by the nodal vector $\mathbf{x}_*^m$ which collects the nodes which move along the particle trajectories of the exact surface (with the initial condition $\mathbf{x}_*^0 = \mathbf{x}^0$). For the semi-discretization in space, denoting by $e_h(t)$ the finite element function associated to the nodal vector $\mathbf{e} = \mathbf{x} - \mathbf{x}_*$ on the interpolated surface $\Gamma_{h,*}(t)$, it was shown in [1, 41] that the classical approach which tracks the error along the particle trajectories can yield the following error estimate for finite elements of degree $k \geq 6$:

$$\|e_h\|_{L_t^\infty(0,T;L^2(\Gamma_{h,*}(t)))} + \|(\nabla_{\Gamma_{h,*}} e_h)n_{h,*}\|_{L_t^2(0,T;L^2(\Gamma_{h,*}(t)))} \leq Ch^{k-1}, \tag{1.5}$$

where $n_{h,*}$ denotes the unit outward normal vector on the interpolated surface $\Gamma_{h,*}(t)$. The second term in this error estimate only contains the normal component of the gradient and therefore could not be used to control the full H^1 semi-norm of the error arising from the stability estimates. The lack of control of the full H^1 semi-norm of the error makes the numerical analysis challenging and unsatisfactory, and requires finite elements of degree ≥ 6 (which is seldom used in the practical computation) for controlling the nonlinear stability terms in the error estimation through the inverse inequalities for finite element functions. This difficulty was circumvented in [35, 37, 38] by discretizing the evolution equations of H and n. It is noted that the algorithms which discretize the evolution equations of H and n suffer from a history effect in long-time simulations, i.e., the errors of n and H may accumulate in a long-time simulation and therefore certain re-initializations of n and H are needed in the computations once the accumulation errors reach a tolerance

In addition to the lack of full H^1 semi-norm estimates, the classical approach which estimates the error along the particle trajectories also could not be used to prove the convergence of BGN-type methods, as the BGN-type methods may contain implicitly determined artificial tangential velocities which cannot be trivially tracked by means of particle trajectories unless the analytic expressions of such tangential velocities are available. As a result, the convergence of BGN-type methods for mean curvature flow and other geometric flows remains open for both curves in two dimensions and surfaces in three dimensions.

In this article, we address the above-mentioned problems by introducing a new approach for analyzing parametric finite element approximations to the evolution of surfaces under geometric flows—to estimate the *projected distance* from the numerically computed surface to the exact surface, rather than the distance between particle trajectories of the two surfaces. To illustrate the basic methodology of this new approach, we define $\hat{\mathbf{x}}_*^m$ as the *distance projection* of $\mathbf{x}^m$, which is assumed to be sufficiently close to $\Gamma(t_m)$, onto the exact surface $\Gamma(t_m)$, i.e., $\hat{\mathbf{x}}_*^m = (\hat{x}_{1,*}^m, \cdots, \hat{x}_{J,*}^m)^\top$ with

$$x_j^m - \hat{x}_{j,*}^m = \pm|x_j^m - \hat{x}_{j,*}^m| n(\hat{x}_{j,*}^m), \tag{1.6}$$

where $n(\hat{x}_{j,*}^m)$ denotes the unit normal vector at point $\hat{x}_{j,*}^m$ on the exact surface $\Gamma(t_m)$, and J is the number of the nodes. Using finite element space of order k, we define $\hat{\Gamma}_{h,*}^m$ to be the piecewise curved triangular surface which interpolates the exact surface at the nodes in $\hat{\mathbf{x}}_*^m$, and define $\mathbf{x}_*^{m+1}$ to be the nodal vector consisting of the new positions of the nodes in $\hat{\mathbf{x}}_*^m$ evolving under mean curvature flow from t_m to t_{m+1}. Then we compare equation (1.3) with the following equation:

$$\mathbf{M}(\hat{\mathbf{x}}_*^m)\frac{\mathbf{x}_*^{m+1} - \hat{\mathbf{x}}_*^m}{\tau} + \mathbf{A}(\hat{\mathbf{x}}_*^m)\mathbf{x}_*^{m+1} = \hat{\mathbf{d}}^m, \tag{1.7}$$

which is satisfied for some defect $\hat{\mathbf{d}}^m$ of $O(\tau + h^k)$ measured in the discrete H^{-1} norm on $\hat{\Gamma}_{h,*}^m$.

To simplify the notations, we will always identify a finite element function with a nodal vector. Such a representation is unique once we have specified the underlying surface. For example, the two integrands of

$$\int_{\hat{\Gamma}_{h,*}^m} v_h \quad \text{and} \quad \int_{\Gamma_h^m} v_h$$

have the same nodal vector but are finite element functions defined on different surfaces. When the underlying surface is specified, the meaning of v_h has no ambiguity. Since all the calculations in this article involve either integrals or norms, the above-mentioned notations for finite element functions will always have unique and clear meanings on the corresponding surfaces specified in the notations of integrations and norms. Let e_h^{m+1} and $\hat{e}_h^m$ be the finite element functions of nodal vectors

$$\mathbf{e}^{m+1} = \mathbf{x}^{m+1} - \mathbf{x}_*^{m+1} \quad \text{and} \quad \hat{\mathbf{e}}^m = \mathbf{x}^m - \hat{\mathbf{x}}_*^m,$$

respectively, and denote by $\nabla_{\hat{\Gamma}_{h,*}^m} e_h^{m+1}$ the matrix with column vectors being the gradients of the components of e_h^{m+1}. The matrix–vector product $(\nabla_{\hat{\Gamma}_{h,*}^m} e_h^{m+1})\hat{n}_{h,*}^m$ is well defined by considering $\hat{n}_{h,*}^m$ as a column vector. Thus the errors at time levels t_{m+1} and t_m are measured in two different ways. In particular, $\hat{e}_h^m$ is the error by projecting the nodes of Γ_h^m onto $\Gamma^m := \Gamma(t_m)$ in the normal direction (orthogonal to the tangential plane at the nodes). It is this orthogonality that leads to the recovery of full H^1 parabolicity, as explained below.

One important geometric structure of mean curvature flow discovered in this article is the following estimate (see Sect. 5.2):

$$\begin{aligned}
&\int_{\Gamma_h^m} \nabla_{\Gamma_h^m} X_h^{m+1} \cdot \nabla_{\Gamma_h^m} e_h^{m+1} - \int_{\hat{\Gamma}_{h,*}^m} \nabla_{\hat{\Gamma}_{h,*}^m} X_{h,*}^{m+1} \cdot \nabla_{\hat{\Gamma}_{h,*}^m} e_h^{m+1} \\
&\quad \geq \frac{1}{2} A_{\hat{\Gamma}_{h,*}^m}^N (e_h^{m+1}, e_h^{m+1}) + \frac{1}{2} A_{\hat{\Gamma}_{h,*}^m}^T (e_h^{m+1}, e_h^{m+1}) - \frac{1}{2} A_{\hat{\Gamma}_{h,*}^m}^T (\hat{e}_h^m, \hat{e}_h^m) \\
&\qquad + B_{\Gamma^m}(\hat{e}_h^m, e_h^{m+1}) + K^m(e_h^{m+1}),
\end{aligned} \tag{1.8}$$

where $A_{\hat{\Gamma}_{h,*}^m}^N (u_h, v_h)$ and $A_{\hat{\Gamma}_{h,*}^m}^T (u_h, v_h)$ are the normal and tangential components of the H^1 bilinear form on $\hat{\Gamma}_{h,*}^m$, i.e.,

$$\begin{aligned}
A_{\hat{\Gamma}_{h,*}^m}^N (u_h, v_h) &= \int_{\hat{\Gamma}_{h,*}^m} \big[(\nabla_{\hat{\Gamma}_{h,*}^m} u_h)\hat{n}_{h,*}^m\big] \cdot \big[(\nabla_{\hat{\Gamma}_{h,*}^m} v_h)\hat{n}_{h,*}^m\big] \\
A_{\hat{\Gamma}_{h,*}^m}^T (u_h, v_h) &= \int_{\hat{\Gamma}_{h,*}^m} \mathrm{tr}\big[(\nabla_{\hat{\Gamma}_{h,*}^m} u_h)(\mathrm{I} - \hat{n}_{h,*}^m \otimes \hat{n}_{h,*}^m)(\nabla_{\hat{\Gamma}_{h,*}^m} v_h)^\top\big]
\end{aligned}$$

and the following estimates hold:

$$\begin{aligned}
|B_{\Gamma^m}(\hat{e}_h^m, e_h^{m+1})| &\leq C\|\hat{e}_h^m\|_{L^2(\hat{\Gamma}_{h,*}^m)} \|\nabla_{\hat{\Gamma}_{h,*}^m} e_h^{m+1}\|_{L^2(\hat{\Gamma}_{h,*}^m)}, \\
|K^m(e_h^{m+1})| &\leq Ch^{0.5}\big(\|\nabla_{\hat{\Gamma}_{h,*}^m} \hat{e}_h^m\|_{L^2(\hat{\Gamma}_{h,*}^m)} \\
&\quad + \|\nabla_{\hat{\Gamma}_{h,*}^m} e_h^{m+1}\|_{L^2(\hat{\Gamma}_{h,*}^m)}\big)\|\nabla_{\hat{\Gamma}_{h,*}^m} e_h^{m+1}\|_{L^2(\hat{\Gamma}_{h,*}^m)}.
\end{aligned}$$

Note that the two tangential components in (1.8), i.e.,

$$\frac{1}{2} A_{\hat{\Gamma}_{h,*}^m}^T (e_h^{m+1}, e_h^{m+1}) \quad \text{and} \quad \frac{1}{2} A_{\hat{\Gamma}_{h,*}^m}^T (\hat{e}_h^m, \hat{e}_h^m),$$

would be cancelled in the spatial semi-discretization by the classical approach by estimating the error between particle trajectories (in the classical approach there is no

$\hat{}$ on e_h^m). This is the reason that one could only get control of the normal component of the H^1 semi-norm in the classical approach. The advantage of the new approach proposed in this article is that $\hat{e}_h^m$ is the error from projecting the nodes of Γ_h^m onto Γ^m in the normal direction, and therefore $\hat{e}_h^m$ is orthogonal to the tangential plane of Γ^m at the nodes. This orthogonality relation could be used to eliminate the term $-\frac{1}{2}A^T_{\hat{\Gamma}^m_{h,*}}(\hat{e}_h^m, \hat{e}_h^m)$ with the following estimate (changing the tangential H^1 semi-norm to the weaker L^2 norm):

$$|A^T_{\hat{\Gamma}^m_{h,*}}(\hat{e}_h^m, \hat{e}_h^m)| \le C\|\hat{e}_h^m\|^2_{L^2(\hat{\Gamma}^m_{h,*})} + Ch^{1.5}\|\nabla_{\hat{\Gamma}^m_{h,*}}\hat{e}_h^m\|^2_{L^2(\hat{\Gamma}^m_{h,*})}. \tag{1.9}$$

Therefore, the tangential component $\frac{1}{2}A^T_{\hat{\Gamma}^m_{h,*}}(e_h^{m+1}, e_h^{m+1})$ could be kept and combined with the normal component $\frac{1}{2}A^N_{\hat{\Gamma}^m_{h,*}}(e_h^{m+1}, e_h^{m+1})$, leading to the following type of estimates:

$$\begin{aligned} &\frac{1}{\tau}\left(\|e_h^{m+1}\|^2_{L^2(\hat{\Gamma}^m_{h,*})} - \|\hat{e}_h^m\|^2_{L^2(\hat{\Gamma}^m_{h,*})}\right) + \|\nabla_{\hat{\Gamma}^m_{h,*}}e_h^{m+1}\|^2_{L^2(\hat{\Gamma}^m_{h,*})} \\ &\le C\|\hat{e}_h^m\|^2_{L^2(\hat{\Gamma}^m_{h,*})} + C\epsilon\|\nabla_{\hat{\Gamma}^m_{h,*}}\hat{e}_h^m\|^2_{L^2(\hat{\Gamma}^m_{h,*})} + C(\tau + h^k)^2. \end{aligned} \tag{1.10}$$

This recovers a full H^1 semi-norm in the error estimates and therefore could help us to control the nonlinear stability terms in a much better way than the classical approach.

One of the difficulties of applying this new approach is the conversion of hatted and un-hatted errors in (1.10). More specifically, the two terms

$$\|e_h^{m+1}\|^2_{L^2(\hat{\Gamma}^m_{h,*})} \quad \text{and} \quad \|\nabla_{\hat{\Gamma}^m_{h,*}}\hat{e}_h^m\|^2_{L^2(\hat{\Gamma}^m_{h,*})}$$

in (1.10) should be converted to

$$\|\hat{e}_h^{m+1}\|^2_{L^2(\hat{\Gamma}^{m+1}_{h,*})} \quad \text{and} \quad \|\nabla_{\hat{\Gamma}^m_{h,*}}e_h^{m+1}\|^2_{L^2(\hat{\Gamma}^m_{h,*})},$$

respectively, in order to apply the discrete version of Grönwall's inequality. This is addressed based on the observation of the following two geometric relations.

First, $\|\nabla_{\hat{\Gamma}^m_{h,*}}(e_h^{m+1} - \hat{e}_h^m)\|^2_{L^2(\hat{\Gamma}^m_{h,*})}$ is high-order smaller than $\|\nabla_{\hat{\Gamma}^m_{h,*}}\hat{e}_h^m\|^2_{L^2(\hat{\Gamma}^m_{h,*})}$. This could be used to convert $\epsilon\|\nabla_{\hat{\Gamma}^m_{h,*}}\hat{e}_h^m\|^2_{L^2(\hat{\Gamma}^m_{h,*})}$ to $\epsilon\|\nabla_{\hat{\Gamma}^m_{h,*}}e_h^{m+1}\|^2_{L^2(\hat{\Gamma}^m_{h,*})}$ in (1.10) (plus high-order smaller terms); see Sect. 5.3.

Second, the orthogonality of $\hat{e}_h^m$ to the tangential plane at nodes leads to the following geometric relation (see Sect. 3.4):

$$\hat{e}_h^{m+1} = I_h[(e_h^{m+1} \cdot n_*^{m+1})n_*^{m+1}] + f_h^{m+1}$$

where $f_h^{m+1} \in S_h(\Gamma_{h,*}^{m+1})$ is the higher-order corrector from the Taylor expansion, satisfying the following estimate (i.e., the remainder is *quadratically smaller*):

$$|f_h^{m+1}| \lesssim |(\mathrm{I} - n_*^{m+1} \otimes n_*^{m+1}) e_h^{m+1}|^2 \quad \text{at the nodes.}$$

This relation could be used to convert $\|e_h^{m+1}\|^2_{L^2(\hat\Gamma^m_{h,*})}$ to $\|\hat e_h^{m+1}\|^2_{L^2(\hat\Gamma^{m+1}_{h,*})}$ in (1.10) (plus high-order smaller terms); see Sect. 5.7. These two geometric relations lead to the following type of estimates:

$$\begin{aligned} &\frac{1}{\tau}\big(\|\hat e_h^{m+1}\|^2_{L^2(\hat\Gamma^{m+1}_{h,*})} - \|\hat e_h^{m}\|^2_{L^2(\hat\Gamma^{m}_{h,*})}\big) + \|\nabla_{\hat\Gamma^m_{h,*}} e_h^{m+1}\|^2_{L^2(\hat\Gamma^m_{h,*})} \\ &\le C\|\hat e_h^m\|^2_{L^2(\hat\Gamma^m_{h,*})} + C\epsilon\|\nabla_{\hat\Gamma^m_{h,*}} e_h^{m+1}\|^2_{L^2(\hat\Gamma^m_{h,*})} + C(\tau + h^k)^2. \end{aligned} \tag{1.11}$$

The H^1 semi-norm in (1.11) could also be used to obtain estimates in the following forms:

$$\|\nabla_{\hat\Gamma^{m+1}_{h,*}} e_h^{m+1}\|^2_{L^2(\hat\Gamma^{m+1}_{h,*})} \ \text{(change of underlying surface)} \quad \text{and} \quad \|\nabla_{\hat\Gamma^{m+1}_{h,*}} \hat e_h^{m+1}\|^2_{L^2(\hat\Gamma^{m+1}_{h,*})}.$$

The recovery of the full H^1 parabolicity in the new approach described above is based on utilizing the geometric structures of the curvature flow and the time discretization which allows us to introduce an intermediate local flow from $\hat{\mathbf{x}}_*^m$ to $\mathbf{x}_*^{m+1}$. This local flow could not be defined in the time-continuous case (spatial semi-discretization), nor could it be extended to the whole time interval $[0, T]$ as a continuous flow map. We present this new approach through analyzing Dziuk's fully discrete parametric FEM for mean curvature flow of closed surfaces in the three-dimensional space, and show that this new approach could yield much better results than the classical approach, i.e.,

$$\max_{1\le m\le[T/\tau]} \|\hat e_h^m\|^2_{L^2(\hat\Gamma^m_{h,*})} + \sum_{m=1}^{[T/\tau]} \tau\|\nabla_{\hat\Gamma^m_{h,*}} \hat e_h^m\|^2_{L^2(\hat\Gamma^m_{h,*})} \le C(\tau + h^k)^2. \tag{1.12}$$

In particular, higher-order convergence in space is proved in comparison with the result (1.5) given by the classical approach, and the requirement on the degree of finite elements (to control the nonlinear terms in the stability estimates) is relaxed from finite elements of degree $k \ge 6$ to $k \ge 3$ for surfaces in the three-dimensional space.

Moreover, since this new approach is proposed to estimate the projected distance instead of the error between particle trajectories, it automatically neglects the tangential motion in the numerical approximation and therefore provides a foundational mathematical tool for analyzing other parametric FEMs which contain artificial tangential motions. This will be demonstrated in the subsequent articles.

The rest of this article is organized as follows. In Sect. 2 we present the statement of the main theorem on the convergence of Dziuk's fully discrete parametric FEM for mean curvature flow using the concept of the distance projection error. In Sect. 3, we present the general settings of the new framework introduced in this article, including the approximation properties of the interpolated surface, the induction assumptions for the accuracy of approximations, and the geometric relations arising from distance projection at nodes. The proof of the main theorem is presented in Sect. 4 (consistency estimates) and Sect. 5 (stability estimates) based on the general settings of the new framework established in Sect. 3 and the optimal-order approximation properties of the interpolated surface. Finally, numerical results are presented in Sect. 6 to support the theoretical analysis, and concluding remarks are presented in Sect. 7. The rigorous proof of the optimal-order approximation properties of the interpolated surface is presented in Appendix.

2 Convergence Results for Mean Curvature Flow

We use the following notations for the initial configurations of the surfaces.

Γ^0: Exact surface at the initial time $t = 0$, i.e. $\Gamma^0 = \Gamma(0)$.
Γ_h^0: The piecewise curved triangular surface (each piece being the image of a reference triangle under a polynomial map of degree k) that approximates Γ^0.
$\Gamma_{h,\mathrm{f}}^0$: The piecewise flat triangular surface (each piece being the image of a reference triangle under a polynomial map of degree 1) whose vertices coincide with the vertices of Γ_h^0. Thus $\Gamma_{h,\mathrm{f}}^0$ is uniquely determined by Γ_h^0.

If K^0 is a curved triangle on Γ_h^0 and K_f^0 is the flat triangle on $\Gamma_{h,\mathrm{f}}^0$ with the same three vertices as K^0, then we denote by $F_{K^0} : K_\mathrm{f}^0 \to K^0$ the unique polynomial of degree k which parametrizes K^0. We assume that the initial triangulation is sufficiently good with the following property:

$$\max_{K^0 \subset \Gamma_h^0} \left(\|F_{K^0}\|_{W^{k,\infty}(K_\mathrm{f}^0)} + \|F_{K^0}^{-1}\|_{W^{1,\infty}(K^0)} \right) \le \kappa_0, \tag{2.1}$$

where κ_0 is some constant that is independent of h. This property holds for standard parametric finite elements and guarantees the following optimal-order approximation of the piecewise triangular surface Γ_h^0 to the smooth surface Γ^0 (the error of Lagrange interpolation):

$$\max_{K^0 \subset \Gamma_h^0} \left(\|a^0 \circ F_{K^0} - F_{K^0}\|_{L^\infty(K_\mathrm{f}^0)} + h\|a^0 \circ F_{K^0} - F_{K^0}\|_{W^{1,\infty}(K_\mathrm{f}^0)} \right) \le Ch^{k+1}, \tag{2.2}$$

where $a^0(x)$ denotes the projection of x onto Γ^0 such that $x - a^0(x) = \pm|x - a^0(x)|n^0(a^0(x))$ with n^0 denoting the normal vector on Γ^0. The projection $a^0(x)$ is well defined for points x in a neighborhood of Γ^0 and therefore well defined on Γ_h^0 for sufficiently small mesh size h.

Let $t_m = m\tau$, $m = 0, 1, \dots, N$, be a partition of the time interval $[0, T]$ with stepsize $\tau > 0$, and let x_j^m, $j = 1, \dots, J$, be the nodes of the approximate surface Γ_h^m given by Dziuk's parametric FEM at time level $t = t_m$ with finite elements of degree

$k \geq 1$. We denote by $\mathscr{K}_h^m$ the set of curved triangles which form the approximate surface Γ_h^m. Each curved triangle $K \in \mathscr{K}_h^m$ is the image of a curved triangle $K^0 \subset \Gamma_h^0$ under the discrete flow map X_h^m and has a parametrization $F_K : K_{\mathrm{f}}^0 \to K$, which is the unique polynomial of degree k that maps K_{f}^0 onto K, where K_{f}^0 is the unique flat triangle which has the same three vertices as K^0. The finite element space on the approximate surface Γ_h^m is defined as

$$S_h(\Gamma_h^m) = \{v_h \in H^1(\Gamma_h^m) : v_h \circ F_K \in \mathbb{P}^k(K_{\mathrm{f}}^0)^3 \text{ for all } K \in \mathscr{K}_h^m\},$$

where $\mathbb{P}^k(K_{\mathrm{f}}^0)$ denotes the space of polynomials of degree k on the flat triangle K_{f}^0.

The following notations are used in the statement of the main results for mean curvature flow (more notations can be found in Sect. 3.1).

- Γ^m: Exact surface at time level $t = t_m$.
- Γ_h^m: The numerically computed surface at time level $t = t_m$.
- $\mathbf{x}^m$: The nodal vector consisting of nodes' positions on the numerically computed surface Γ_h^m.
- $\hat{\mathbf{x}}_*^m$: The distance projection of $\mathbf{x}^m$ onto the exact surface Γ^m.
- $\hat{\Gamma}_{h,*}^m$: The piecewise triangular surface which interpolates Γ^m at the nodes in $\hat{\mathbf{x}}_*^m$.
- X_h^m: The finite element function with nodal vector $\mathbf{x}^m$. It coincides with the identity map, i.e., $\mathrm{id}(x) = x$, when it is considered as a function on Γ_h^m.

Let $\delta > 0$ be a sufficiently small constant such that every point x in the δ-neighborhood of the exact surface $\Gamma^m = \Gamma(t_m)$, denoted by $D_\delta(\Gamma^m) = \{x \in \mathbb{R}^3 : \mathrm{dist}(x, \Gamma^m) \leq \delta\}$, has a unique normal projection onto Γ^m, denoted by $a^m(x)$, satisfying the following relation:

$$x - a^m(x) = \pm|x - a^m(x)|n^m(a^m(x)),$$

i.e., $x - a^m(x)$ is orthogonal to the tangent plane of Γ^m at $a^m(x)$. Thus the distance projection $a^m : D_\delta(\Gamma^m) \to \Gamma^m$ is well defined, with a constant δ which is independent of m (but possibly dependent on T).

The interpolated surface $\hat{\Gamma}_{h,*}^m$ is determined by the nodes in $\hat{\mathbf{x}}_*^m$, which is obtained by projecting the nodes of the numerically computed surface Γ_h^m onto the exact surface Γ^m. We shall prove that the numerically computed surface Γ_h^m is in a δ-neighborhood of the exact smooth surface Γ^m so that the projection of the nodes of Γ_h^m onto Γ^m are well defined (thus the interpolated surface $\hat{\Gamma}_{h,*}^m$ is well defined).

As mentioned in the introduction section, we always identify a finite element function with a nodal vector. Correspondingly, the interpolation operator I_h should be interpreted as the determination of the nodal vector which uniquely corresponds to a finite element function after specifying the underlying surface. The lift of a finite element function v_h onto the smooth surface Γ^m is defined as

$$v_h^l = v_h \circ (a^m|_{\hat{\Gamma}_{h,*}^m})^{-1}$$

by first identifying v_h as a finite element function on the interpolated surface $\hat{\Gamma}^m_{h,*}$; see [22, Section 2.4] and [37, Section 3.4]. The inverse lift of $v \in L^2(\Gamma^m)$ onto $\hat{\Gamma}^m_{h,*}$ is defined as $v^{-l} = v \circ a^m$.

In order to measure the error between the numerically computed surface Γ^m_h and the smooth surface Γ^m, we define the lifted error function

$$\hat{e}^m = \hat{X}^{m,l}_h - \mathrm{id}_{\Gamma^m} \in H^1(\Gamma^m),$$

where $\hat{X}^{m,l}_h$ denotes the lift of X^m_h onto Γ^m through the interpolated surface $\hat{\Gamma}^m_{h,*}$.

By the new approach described in the introduction section (described more specifically in the next section), we shall prove the following theorem on the convergence of Dziuk's fully discrete parametric FEM for mean curvature flow of closed surfaces in the three-dimensional space.

Theorem 2.1 *Suppose that the flow map $\phi : \Gamma^0 \times [0, T] \to \mathbb{R}^3$ of the mean curvature flow and its inverse map $\phi(\cdot, t)^{-1} : \Gamma(t) \to \Gamma^0$ are both sufficiently smooth, uniformly with respect to $t \in [0, T]$, and the initial triangulation of the surface is sufficiently good, satisfying (2.1). Let X^m_h be the finite element solution given by Dziuk's parametric FEM in (1.2) with initial condition $X^0_h = \mathrm{id}$ on Γ^0_h. Then for any given constant c (independent of τ and h), there exists a positive constant h_0 such that for $\tau \le ch^k$ and $h \le h_0$ the following error estimate holds for finite elements of degree $k \ge 3$:*

$$\max_{1\le m\le [T/\tau]} \|\hat{e}^m\|^2_{L^2(\Gamma^m)} + \sum_{m=1}^{[T/\tau]} \tau \|\nabla_{\Gamma^m} \hat{e}^m\|^2_{L^2(\Gamma^m)} \le Ch^{2k}, \tag{2.3}$$

where the constant C is independent of τ and h (but possibly dependent on κ_0 and T).

Remark 2.1 The condition $\tau \le ch^k$ is required in proving the shape regularity of triangulations and the optimal-order approximation to Γ^m of the curved finite element interpolated surface $\hat{\Gamma}^m_{h,*}$. The restriction to finite elements of degree $k \ge 3$ is due to several technical difficulties, including

(1) The application of the inverse inequality on the two-dimensional surface $\hat{\Gamma}^m_{h,*}$, i.e.,

$$\|\hat{e}^m_h\|_{W^{1,\infty}(\hat{\Gamma}^m_{h,*})} \le Ch^{-2}\|\hat{e}^m_h\|_{L^2(\hat{\Gamma}^m_{h,*})} \le Ch^{-2}(\tau + h^k), \tag{2.4}$$

which is used to guarantee the equivalence of norms between the approximate surfaces Γ^m_h, $\hat{\Gamma}^m_{h,*}$ and $\Gamma^{m+1}_{h,*}$. The convergence of the numerical approximations in the $W^{1,\infty}$ norm based on inequality (2.4) requires $\tau = o(h^2)$ and $k \ge 3$.

(2) The proof of shape regularity of triangulations and the optimal-order approximation to Γ^m of the curved finite element interpolated surface $\hat{\Gamma}^m_{h,*}$ also requires $k \ge 3$.

Remark 2.2 The proof of Theorem 2.1 could be trivially extended to any partition $0 = t_0 < t_1 < \cdots < t_N = T$ with variable time stepsizes $\tau_m = t_m - t_{m-1}$ under the condition $\max_{1\le m\le N} \tau_m = o(h^{2.5})$. Although the stepsize condition is required in the proof of convergence for the parametric FEM, it is not observed in the numerical experiments.

The rigorous proof of Theorem 2.1 via the new approach discussed in the introduction section is presented in the following sections.

3 General Settings of the New Framework

In this section, we present the general settings of the new framework introduced in this article for estimating the projected distance from the numerically computed surface to the exact surface. These general settings, including the properties of the interpolated surface, the induction assumptions for the accuracy of approximations, and the geometric relations arising from distance projection at nodes, are also applicable to the analysis of other geometric flows and parametric finite element algorithms. The analysis of Dziuk's parametric FEM for mean curvature flow will be presented in the next section based on the general settings established in this section.

3.1 Notations

For the simplicity of notations, we identify a finite element function with a nodal vector. For example, $\|v_h\|_{\hat\Gamma^m_{h,*}}$ and $\|v_h\|_{\Gamma^m_h}$ denote the norms of a finite element function (a nodal vector) on the two different surfaces $\hat\Gamma^m_{h,*}$ and Γ^m_h, respectively. The following notations for different surfaces and flow maps will be frequently used in the analysis of parametric finite element approximations to geometric flows. These notations are defined in the text and summarized below for the convenience of the readers.

Γ^m: Exact surface at time level $t = t_m$.
Γ^m_h: The numerically computed surface at time level $t = t_m$.
$\mathbf{x}^m$: The nodal vector consisting of nodes on the numerically computed surface Γ^m_h.
$\hat{\mathbf{x}}^m_*$:The distance projection of $\mathbf{x}^m$ onto the exact surface Γ^m.
$\mathbf{x}^{m+1}_*$: The new position of $\hat{\mathbf{x}}^m_*$ evolving under mean curvature flow from t_m to t_{m+1}.
$\hat\Gamma^m_{h,*}$: The piecewise triangular surface which interpolates Γ^m at the nodes in $\hat{\mathbf{x}}^m_*$.
$\Gamma^{m+1}_{h,*}$: The piecewise triangular surface which interpolates Γ^{m+1} at the nodes in $\mathbf{x}^{m+1}_*$.
X^m_h: The finite element function with nodal vector $\mathbf{x}^m$. It coincides with the identity map, i.e., $\mathrm{id}(x) = x$, when it is considered as a function on Γ^m_h.
X^{m+1}_h: The finite element function with nodal vector $\mathbf{x}^{m+1}$. When it is considered as a function on Γ^m_h, it represents the local flow map from Γ^m_h to Γ^{m+1}_h.
$\hat X^m_{h,*}$: The finite element function with nodal vector $\hat{\mathbf{x}}^m_*$. It coincides with the identity map, i.e., $\mathrm{id}(x) = x$, when it is considered as a function on $\hat\Gamma^m_{h,*}$.

$X_{h,*}^{m+1}$: The finite element function with nodal vector $\mathbf{x}_*^{m+1}$. When it is considered as a function on $\hat{\Gamma}_{h,*}^m$, it represents the local flow map from $\hat{\Gamma}_{h,*}^m$ to $\Gamma_{h,*}^{m+1}$.
X^{m+1}: The local flow map from Γ^m to Γ^{m+1} under mean curvature flow.
$\hat{e}_h^m$: The finite element error function with nodal vector $\hat{\mathbf{e}}^m = \mathbf{x}^m - \hat{\mathbf{x}}_*^m$.
e_h^{m+1}: The auxiliary error function with nodal vector $\mathbf{e}^{m+1} = \mathbf{x}^{m+1} - \mathbf{x}_*^{m+1}$.
ϕ_h^m: The flow map from Γ_h^0 to Γ_h^m.
$\hat{\phi}_{h,*}^m$: The flow map from Γ_h^0 to $\hat{\Gamma}_{h,*}^m$.
H^m: The mean curvature on Γ^m.
n^m: The normal vector on Γ^m.
$a^m(x)$: The distance projection of x onto Γ^m.
n_*^m: The extension of n^m to a neighborhood of Γ^m by $n_*^m = n^m \circ a^m$. It can also be viewed as the inversely lift of n^m onto $\hat{\Gamma}_{h,*}^m$.
$\hat{n}_{h,*}^m$: The normal vector on $\hat{\Gamma}_{h,*}^m$.
n_h^m: The normal vector on Γ_h^m.
N_*^m: The normal projection operator $N_*^m = n_*^m (n_*^m)^\top$ on $\hat{\Gamma}_{h,*}^m$.
N^m: The normal projection operator $N^m = n^m (n^m)^\top$ on Γ^m. Thus N^m is the lift of N_*^m onto Γ^m.
$\hat{T}_{h,*}^m$: The tangential projection operator $T_*^m = I - \hat{n}_{h,*}^m (\hat{n}_{h,*}^m)^\top$ on $\hat{\Gamma}_{h,*}^m$.
T_*^m: The tangential projection operator $T_*^m = I - n_*^m (n_*^m)^\top$ on $\hat{\Gamma}_{h,*}^m$.
T^m: The tangential projection operator $T^m = I - n^m (n^m)^\top$ on Γ^m. Thus T^m is the lift of T_*^m onto Γ^m.

3.2 Approximation Properties of the Interpolated Surface $\hat{\Gamma}_{h,*}^m$

If K is a curved triangle on $\hat{\Gamma}_{h,*}^m$ then we denote by K^0 the curved triangle on Γ_h^0 which is mapped to K by the discrete flow map $\hat{X}_{h,*}^m$, and denote by $F_{K^0} : K_{\mathrm{f}}^0 \to K^0$ the parametrization of the curved triangle $K^0 \subset \Gamma_h^0$ (as in the beginning of Sect. 2), where K_{f}^0 is the flat triangle which has the same three vertices as K^0. The flat triangles K_{f}^0 form a piecewise flat triangular surface

$$\Gamma_{h,\mathrm{f}}^0 = \bigcup_{K^0 \in \Gamma_h^0} K_{\mathrm{f}}^0.$$

We still denote by $\hat{X}_{h,*}^m : \Gamma_{h,\mathrm{f}}^0 \to \hat{\Gamma}_{h,*}^m$ the unique piecewise polynomial of degree k (with the nodal vector $\hat{\mathbf{x}}_*^m$ as before) which parametrizes $\hat{\Gamma}_{h,*}^m$, and denote by $\|\hat{X}_{h,*}^m\|_{H_h^j(\Gamma_{h,\mathrm{f}}^0)}$ and $\|\hat{X}_{h,*}^m\|_{W_h^{j,\infty}(\Gamma_{h,\mathrm{f}}^0)}$ the piecewise Sobolev norms on $\Gamma_{h,\mathrm{f}}^0$, i.e.,

$$\|\hat{X}_{h,*}^m\|_{H_h^j(\Gamma_{h,\mathrm{f}}^0)} := \Big(\sum_{K_{\mathrm{f}}^0 \subset \Gamma_{h,\mathrm{f}}^0} \|\hat{X}_{h,*}^m\|_{H^j(K_{\mathrm{f}}^0)}^2 \Big)^{\frac{1}{2}} \quad \text{and}$$

$$\|\hat{X}_{h,*}^m\|_{W_h^{j,\infty}(\Gamma_{h,\mathrm{f}}^0)} := \max_{K_{\mathrm{f}}^0 \subset \Gamma_{h,\mathrm{f}}^0} \|\hat{X}_{h,*}^m\|_{W^{j,\infty}(K_{\mathrm{f}}^0)}.$$

For the discrete flow maps $\hat{X}^m_{h,*} : \Gamma^0_{h,\mathrm{f}} \to \hat{\Gamma}^m_{h,*}$, $m = 0, 1, \ldots$, we denote

$$\begin{aligned} \kappa_l &:= \max_{0\le m\le l} (\|\hat{X}^m_{h,*}\|_{H^{k-1}_h(\Gamma^0_{h,\mathrm{f}})} + \|\hat{X}^m_{h,*}\|_{W^{k-2,\infty}_h(\Gamma^0_{h,\mathrm{f}})} + \|(\hat{X}^m_{h,*})^{-1}\|_{W^{1,\infty}_h(\hat{\Gamma}^m_{h,*})}), \\ \kappa_{*,l} &:= \max_{0\le m\le l} (\|\hat{X}^m_{h,*}\|_{H^{k}_h(\Gamma^0_{h,\mathrm{f}})} + \|\hat{X}^m_{h,*}\|_{W^{k-1,\infty}_h(\Gamma^0_{h,\mathrm{f}})}). \end{aligned} \tag{3.1}$$

By pulling functions on $\hat{\Gamma}^m_{h,*}$ back to $\Gamma^0_{h,\mathrm{f}}$ via the map $\hat{X}^m_{h,*} : \Gamma^0_{h,\mathrm{f}} \to \hat{\Gamma}^m_{h,*}$ (and vice visa), one can see that the $W^{1,p}$, $p \in [1, \infty]$, norms of a finite element function (with a fixed nodal vector) on $\Gamma^0_{h,\mathrm{f}}$ and $\hat{\Gamma}^m_{h,*}$ are equivalent up to constants which depend on κ_l, i.e.,

$$C^{-1}_{\kappa_l}\|v_h\|_{W^{1,p}(\hat{\Gamma}^m_{h,*})} \le \|v_h\|_{W^{1,p}(\Gamma^0_{h,\mathrm{f}})} \le C_{\kappa_l}\|v_h\|_{W^{1,p}(\hat{\Gamma}^m_{h,*})} \quad \text{for } \ 0 \le m \le l.$$

Accordingly, the interpolated surface $\hat{\Gamma}^m_{h,*}$ approximates the smooth surface Γ^m to the optimal order, as shown below.

In fact, for a curved triangle $K \subset \hat{\Gamma}^m_{h,*}$, its parametrization $F_K = \hat{X}^m_{h,*}|_{K^0_\mathrm{f}} : K^0_\mathrm{f} \to K$ (a polynomial of degree k) satisfies the following estimates as a result of (3.1): For $m = 0, \ldots, l$,

$$\begin{aligned} \Big(\sum_{K\subset\hat{\Gamma}^m_{h,*}} \|F_K\|^2_{H^{k-1}(K^0_\mathrm{f})}\Big)^{\frac12} + \max_{K\subset\hat{\Gamma}^m_{h,*}} \|F_K\|_{W^{k-2,\infty}(K^0_\mathrm{f})} + \max_{K\subset\hat{\Gamma}^m_{h,*}} \|F_K\|_{W^{1,\infty}(K^0_\mathrm{f})} \le \kappa_l, \\ \Big(\sum_{K\subset\hat{\Gamma}^m_{h,*}} \|F_K\|^2_{H^{k}(K^0_\mathrm{f})}\Big)^{\frac12} + \max_{K\subset\hat{\Gamma}^m_{h,*}} \|F_K\|_{W^{k-1,\infty}(K^0_\mathrm{f})} \le \kappa_{*,l}. \end{aligned} \tag{3.2}$$

Then, in terms of the normal projection $a^m : D_\delta(\Gamma^m) \to \Gamma^m$, we obtain a map $a^m \circ F_K : K^0_\mathrm{f} \to \Gamma^m$, and F_K is the unique polynomial of degree k which interpolates the function $a^m \circ F_K$ at the nodes of the flat triangle K^0_f. Therefore, the following polynomial approximation property holds:

$$\begin{aligned} &\sum_{K\subset\hat{\Gamma}^m_{h,*}} \big(\|a^m \circ F_K - F_K\|^2_{L^2(K^0_\mathrm{f})} + h^2\|a^m \circ F_K - F_K\|^2_{H^1(K^0_\mathrm{f})}\big) \\ &\le Ch^{2k+2} \sum_{K\subset\hat{\Gamma}^m_{h,*}} \|\nabla^{k+1}_{K^0_\mathrm{f}}(a^m \circ F_K)\|^2_{L^2(K^0_\mathrm{f})} \\ &\le Ch^{2k+2} \sum_{K\subset\hat{\Gamma}^m_{h,*}} \sum_{\substack{j_1+\cdots+j_l=k+1\\ \max(j_1,\ldots,j_l)\le k}} \|\nabla^{j_1}_{K^0_\mathrm{f}} F_K \cdots \nabla^{j_l}_{K^0_\mathrm{f}} F_K\|^2_{L^2(K^0_\mathrm{f})} \\ &\le C_{\kappa_l}(1 + \kappa^2_{*,l})h^{2k+2}, \quad \text{for } m = 0, \ldots, l, \end{aligned} \tag{3.3}$$

where we have used (3.2) in the last inequality. The right-hand sides of (3.3) may depend nonlinearly on κ_l but only have quadratic growth with respect to $\kappa_{*,l}$. This quadratic growth with respect to $\kappa_{*,l}$ is crucial for us to prove that $\kappa_{*,l}$ is independent of τ, h and l (possibly depending on T) by using Grönwall's inequality on $\|\hat{X}^m_{h,*}\|_{H^k_h(\Gamma^0_h)}$; see Appendix.

We denote by I_K the interpolation operator on the flat triangle K^0_{f}. Since $F_K = a^m \circ F_K$ at the nodes of K^0_{f}, it follows that $I_K[a^m \circ F_K] = F_K$. The interpolation of the distance projection $a^m : \hat{\Gamma}^m_{h,*} \to \Gamma^m$ onto the curved surface $\hat{\Gamma}^m_{h,*}$ is denoted by

$$I_h a^m = I_K[a^m \circ F_K] \circ F_K^{-1} = \mathrm{id} \quad \text{on a curved triangle } K \subset \hat{\Gamma}^m_{h,*}.$$

Therefore, the parametrization $a^m : \hat{\Gamma}^m_{h,*} \to \Gamma^m$ of the smooth surface Γ^m satisfies, for $m = 0, \dots, l$,

$$\|a^m - I_h a^m\|_{L^2(\hat{\Gamma}^m_{h,*})} + h\|a^m - I_h a^m\|_{H^1(\hat{\Gamma}^m_{h,*})} \le C_{\kappa_l}(1 + \kappa_{*,l})h^{k+1}, \tag{3.4}$$

which can be obtained by pushing forward the estimate in (3.3) from K^0_{f} to K.

For a smooth function f on the smooth surface Γ^m, we denote by $I_h f$ the interpolation of the inversely lifted function $f^{-l} = f \circ a^m$ onto $\hat{\Gamma}^m_{h,*}$, and denote by $(I_h f)^l$ the lift of $I_h f$ onto Γ^m. Then the following approximation estimates hold for $m = 0, \dots, l$:

$$\begin{aligned} \|f^{-l} - I_h f\|_{L^2(\hat{\Gamma}^m_{h,*})} + h\|f^{-l} - I_h f\|_{H^1(\hat{\Gamma}^m_{h,*})} &\le C_{\kappa_l}(1 + \kappa_{*,l})h^{k+1}, \\ \|f - (I_h f)^l\|_{L^2(\Gamma^m)} + h\|f - (I_h f)^l\|_{H^1(\Gamma^m)} &\le C_{\kappa_l}(1 + \kappa_{*,l})h^{k+1}, \end{aligned} \tag{3.5}$$

which can be proved similarly as (3.4), i.e., replacing a^m by $f \circ a^m$ in (3.3).

We denote by n^m the unit normal vector on Γ^m and denote by $n^m_* = n^m \circ a^m$ a smooth extension of n^m to a neighborhood of Γ^m. In particular, n^m_* is well defined on $\hat{\Gamma}^m_{h,*}$ as the inverse lift of n^m via the distance projection a, and $\|n^m_*\|_{W^{1,\infty}(\hat{\Gamma}^m_{h,*})} \le C$.

Let $K \subset \hat{\Gamma}^m_{h,*}$ be a curved triangle and let $\widetilde{K} \subset \Gamma^m$ be the image of K under the distance projection $a^m : \hat{\Gamma}^m_{h,*} \to \Gamma^m$. In the parametrization $F_K : K^0_{\mathrm{f}} \to K$ of the curved triangle $K \subset \hat{\Gamma}^m_{h,*}$, via a rotation we can assume that the flat triangle K^0_{f} is in $\mathbb{R}^2$ with coordinates u and v. Then the normal vectors on $K \subset \hat{\Gamma}^m_{h,*}$ and $\widetilde{K} \subset \Gamma^m$, both being pulled back to K^0_{f}, are given by

$$\hat{n}^m_{h,*} \circ F_K = \frac{\partial_u F_K \times \partial_v F_K}{|\partial_u F_K \times \partial_v F_K|} \quad \text{and} \quad n^m_* \circ F_K = \frac{\partial_u(a^m \circ F_K) \times \partial_v(a^m \circ F_K)}{|\partial_u(a^m \circ F_K) \times \partial_v(a^m \circ F_K)|},$$

respectively. The first inequality in (3.2) implies $|\hat{n}^m_{h,*} \circ F_K - n^m_* \circ F_K| \le C_{\kappa_l}|\nabla_{K^0_{\mathrm{f}}} F_K|$ and therefore, from the L^∞ version of (3.3) we obtain

$$\|\hat{n}^m_{h,*} - n^m_*\|_{L^\infty(\hat{\Gamma}^m_{h,*})}$$

$$
\begin{aligned}
&= \max_{K\subset\hat{\Gamma}^m_{h,*}} \|\hat{n}^m_{h,*}\circ F_K - n^m_*\circ F_K\|_{L^\infty(K^0_{\rm f})}\\
&\le C_{\kappa_l} \max_{K\subset\hat{\Gamma}^m_{h,*}} \|\nabla_{K^0_{\rm f}}(F_K - a\circ F_K)\|_{L^\infty(K^0_{\rm f})}\\
&\le C_{\kappa_l} \max_{K\subset\hat{\Gamma}^m_{h,*}} h^k\|\nabla^{k+1}_{K^0_{\rm f}}(a\circ F_K)\|_{L^\infty(K^0_{\rm f})}\\
&\le C_{\kappa_l} \max_{K\subset\hat{\Gamma}^m_{h,*}} h^k \sum_{\substack{j_1+\cdots+j_l=k+1\\ \max(j_1,\dots,j_l)\le k-2}} \|\nabla^{j_1}_{K^0_{\rm f}}F_K\cdots\nabla^{j_l}_{K^0_{\rm f}}F_K\|_{L^\infty(K^0_{\rm f})}\\
&\quad + C_{\kappa_l} \max_{K\subset\hat{\Gamma}^m_{h,*}} h^k\big(\|\nabla_{K^0_{\rm f}}F_K\|^2_{L^\infty(K^0_{\rm f})}\|\nabla^{k-1}_{K^0_{\rm f}}F_K\|_{L^\infty(K^0_{\rm f})}\\
&\quad + \|\nabla^2_{K^0_{\rm f}}F_K\|_{L^\infty(K^0_{\rm f})}\|\nabla^{k-1}_{K^0_{\rm f}}F_K\|_{L^\infty(K^0_{\rm f})}\big)\\
&\quad + C_{\kappa_l} \max_{K\subset\hat{\Gamma}^m_{h,*}} h^k\|\nabla_{K^0_{\rm f}}F_K\|_{L^\infty(K^0_{\rm f})}\|\nabla^{k}_{K^0_{\rm f}}F_K\|_{L^\infty(K^0_{\rm f})} \quad \text{(chain rule)}\\
&\le C_{\kappa_l}(1+\kappa_{*,l})h^{k-1} \quad \text{for } k\ge 3, \qquad (3.6)
\end{aligned}
$$

where we have used the following inverse inequalities:

$$
\begin{aligned}
&\|\nabla^2_{K^0_{\rm f}}F_K\|_{L^\infty(K^0_{\rm f})} \le C_{\kappa_l}h^{-1}\|\nabla_{K^0_{\rm f}}F_K\|_{L^\infty(K^0_{\rm f})} \le C_{\kappa_l}h^{-1}\kappa_l,\\
&\|\nabla^k_{K^0_{\rm f}}F_K\|_{L^\infty(K^0_{\rm f})} \le C_{\kappa_l}h^{-1}\|\nabla^{k-1}_{K^0_{\rm f}}F_K\|_{L^\infty(K^0_{\rm f})} \le C_{\kappa_l}h^{-1}\kappa_{*,l}.
\end{aligned}
$$

Following the same proof, we also have the L^2-version of the above estimate:

$$
\|\hat{n}^m_{h,*} - n^m_*\|_{L^2(\hat{\Gamma}^m_{h,*})} \le C_{\kappa_l}(1+\kappa_{*,l})h^k. \qquad (3.7)
$$

From (3.6) we see that, if the mesh size h is sufficiently small such that

$$
(1+\kappa_{*,l})h^{k-2.5} \le C_{\kappa_l}^{-1}, \qquad (3.8)
$$

then the following inequality holds:

$$
\|\hat{n}^m_{h,*} - n^m_*\|_{L^\infty(\hat{\Gamma}^m_{h,*})} \le h^{1.5}. \qquad (3.9)
$$

This estimate will be used in the consistency and stability estimates in the next two sections. The existence of a constant $h_{\kappa_l,\kappa_{*,l}}$ such that condition (3.8) can be satisfied for $h \le h_{\kappa_l,\kappa_{*,l}}$ will be guaranteed by induction assumption (3) in the next subsection.

In the rest of this article, we denote by C a generic positive constant which may be different at different occurrences, possibly dependent on κ_l and T, but is independent of τ, h, m and $\kappa_{*,l}$. We denote by C_0 generic positive constant which is independent of κ_l. For the simplicity of notation, we denote by $A \lesssim B$ the statement "$A \le CB$ for some constant C".

3.3 Induction Assumptions

The error estimates for the parametric finite element approximations to evolving surfaces generally require some mathematical induction on the accuracy of numerical approximations at previous time levels. For the analysis of Dziuk's parametric FEM for mean curvature flow, we assume that the following conditions hold for $m = 0, \dots, l$ (and then prove that these conditions could be recovered for $m = l + 1$):

(1) The numerically computed surface Γ_h^m is in a δ-neighborhood of the exact surface Γ^m. Therefore, the distance projection of the nodes of Γ_h^m onto Γ^m are well defined (thus the interpolated surface $\hat{\Gamma}_{h,*}^m$ is well defined).

(2) The error $\hat{e}_h^m = X_h^m - \hat{X}_{h,*}^m$ satisfies the following estimates:

$$\|\hat{e}_h^m\|_{L^2(\hat{\Gamma}_{h,*}^m)} + h\|\hat{e}_h^m\|_{H^1(\hat{\Gamma}_{h,*}^m)} \le h^{2.5}. \tag{3.10}$$

(3) The constants κ_l and $\kappa_{*,l}$ are independent of τ and h. Accordingly, for any constant C_{κ_l} appearing in the following analysis, condition (3.8) can be satisfied for sufficiently small h.

For the other methods and other geometric flows, these mathematical induction assumptions could be adjusted according to the numerical analysis.

The following results can be obtained from (3.10) by applying the inverse inequality of finite element functions:

$$\|\nabla_{\hat{\Gamma}_{h,*}^m} \hat{e}_h^m\|_{L^2(\hat{\Gamma}_{h,*}^m)} \lesssim h^{1.5}, \quad \|\hat{e}_h^m\|_{L^\infty(\hat{\Gamma}_{h,*}^m)} \lesssim h^{1.5} \quad \text{and} \quad \|\nabla_{\hat{\Gamma}_{h,*}^m} \hat{e}_h^m\|_{L^\infty(\hat{\Gamma}_{h,*}^m)} \lesssim h^{0.5}, \tag{3.11}$$

which guarantee the equivalence of L^p and $W^{1,p}$ norms, $1 \le p \le \infty$, of a finite element function v_h (with a fixed nodal vector) on the family of surfaces

$$\hat{\Gamma}_{h,\theta}^m = (1-\theta)\hat{\Gamma}_{h,*}^m + \theta\Gamma_h^m, \quad \theta \in [0, 1],$$

which are intermediate between the interpolated surface $\hat{\Gamma}_{h,*}^m$ and the numerically computed surface Γ_h^m; see [39, Lemma 4.3]. In particular, the L^p and $W^{1,p}$ norms of a finite element function (with a fixed nodal vector) on $\hat{\Gamma}_{h,*}^m$ and Γ_h^m are equivalent.

3.4 Geometric Relations

In this subsection, we present two types of geometric relations which will be frequently used in the analysis of parametric FEMs by the new approach proposed in this article.

Firstly, if the nodes of Γ_h^{m+1} are in a δ-neighborhood of the smooth surface Γ^{m+1} (thus the projections of these nodes onto Γ^{m+1} are well defined) then, since the hatted error is defined through distance projection, we have the following formula at the finite

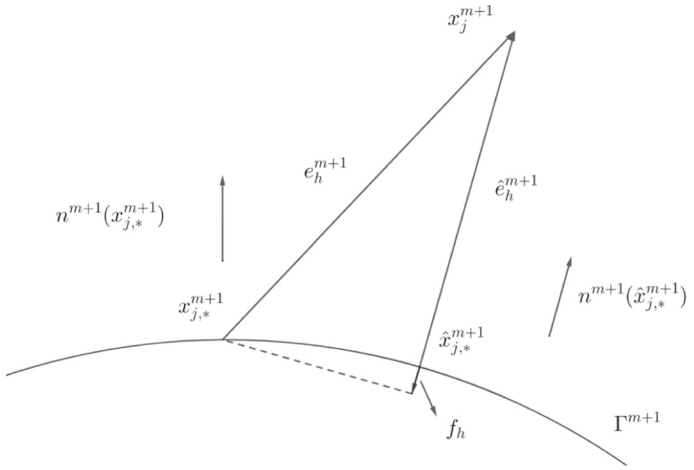


Fig. 1 The geometric relation at the j-th node

element nodes:

$$\hat{e}_h^{m+1} = I_h\big[(e_h^{m+1} \cdot n_*^{m+1})n_*^{m+1}\big] + f_h \tag{3.12}$$

where f_h is the higher order corrector from the Taylor expansion by considering the discrepancy between the distance projection from X_h^{m+1} onto Γ^{m+1} and the orthogonal projection of $e_h^{m+1} = X_{h,*}^{m+1} - X_h^{m+1}$ along n_*^{m+1}; see Fig. 1. From the orthogonality relation, it follows that the amplitude of f_h is no greater than the square of the tangential projection of e_h^{m+1} at the nodes, i.e.,

$$|f_h| \lesssim |[I - n_*^{m+1}(n_*^{m+1})^\top]e_h^{m+1}|^2 \quad \text{at the finite element nodes.} \tag{3.13}$$

This geometric relation plays an important role in the following analysis.

Secondly, we denote by $X_{h,*}^{m+1} : \hat{\Gamma}_{h,*}^m \to \Gamma_{h,*}^{m+1}$ the local flow map under which the nodes of $\hat{\Gamma}_{h,*}^m$ move exactly according to mean curvature flow, and denote by $X^{m+1} : \Gamma^m \to \Gamma^{m+1}$ the local flow map of mean curvature flow. Since $X_{h,*}^{m+1} - \hat{X}_{h,*}^m = X^{m+1} - \mathrm{id}$ at the finite element nodes, it follows that

$$\begin{aligned} X_{h,*}^{m+1} - \hat{X}_{h,*}^m &= I_h(X^{m+1} - \mathrm{id}) \quad &&\text{on } \hat{\Gamma}_{h,*}^m, \\ X^{m+1} - \mathrm{id} &= (v^m + g^m)\tau &&\text{on } \Gamma^m, \end{aligned}$$

where v^m is the exact velocity of the geometric flow at time level $t = t_m$ (for mean curvature flow we have $v^m = -H^m n^m$ with H^m and n^m being the mean curvature and normal vector on Γ^m), and g^m is some smooth correction from the Taylor expansion, satisfying the following estimate:

$$\|g^m\|_{W^{1,\infty}(\Gamma^m)} \le C\tau. \tag{3.14}$$

Therefore, we obtain

$$\begin{aligned} X_h^{m+1} - X_h^m &= e_h^{m+1} - \hat{e}_h^m + X_{h,*}^{m+1} - \hat{X}_{h,*}^m \\ &= e_h^{m+1} - \hat{e}_h^m + \tau I_h(v^m + g^m). \end{aligned} \tag{3.15}$$

This relation plays an important role in estimating the numerical displacement $X_h^{m+1} - X_h^m$.

3.5 Recovery of Full H^1 Norm by the Normal Component's H^1 Norm

Since $(I - n_*^m (n_*^m)^\top)\hat{e}_h^m = 0$ at all the nodes of $\hat{\Gamma}_{h,*}^m$, its interpolation on $\hat{\Gamma}_{h,*}^m$ vanishes, i.e. $I_h[(I - n_*^m (n_*^m)^\top)\hat{e}_h^m] = 0$. Therefore, with the help of norm equivalence relations on different surfaces in [22, Section 2.4] and the super-convergence type arguments, on a curved triangle $K \subset \hat{\Gamma}_{h,*}^m$ (with parametrization $F_K : K_{\rm f}^0 \to K$ defined in Sect. 3.2) we have

$$\begin{aligned} &\|(I - n_*^m (n_*^m)^\top)\hat{e}_h^m\|_{L^2(K)} \\ &\quad = \|(I - n_*^m (n_*^m)^\top)\hat{e}_h^m - I_h[(I - n_*^m (n_*^m)^\top)\hat{e}_h^m]\|_{L^2(K)} \\ &\quad \sim \|[(I - n_*^m (n_*^m)^\top)\hat{e}_h^m] \circ F_K - I_h[(I - n_*^m (n_*^m)^\top)\hat{e}_h^m] \circ F_K\|_{L^2(K_{\rm f}^0)} \\ &\quad \text{(here the norm equivalence[22, } Eq. \text{ (2.18)] is used)} \\ &\quad \lesssim h^{k+1} |[(I - n_*^m (n_*^m)^\top)\hat{e}_h^m] \circ F_K|_{H^{k+1}(K_{\rm f}^0)} \\ &\quad \lesssim h^{k+1} \sum_{i=0}^{k} \Big(\|\hat{e}_h^m \circ F_K\|_{H^{k-i}(K_{\rm f}^0)} \sum_{j_1+\cdots+j_l=i+1} \|\nabla_{K_{\rm f}^0}^{j_1} F_K\|_{L^\infty(K_{\rm f}^0)} \cdots \\ &\qquad \times \|\nabla_{K_{\rm f}^0}^{j_l} F_K\|_{L^\infty(K_{\rm f}^0)} \Big) \\ &\quad \lesssim h^{k+1} \sum_{i=0}^{k} \Big(h^{-k+i} \|\hat{e}_h^m \circ F_K\|_{L^2(K_{\rm f}^0)} \sum_{l=1}^{i+1} h^{-i-1+l} \|F_K\|_{W^{1,\infty}(K_{\rm f}^0)}^l \Big) \\ &\quad \lesssim h \|\hat{e}_h^m\|_{L^2(K)} \qquad \text{(the first inequality of (3.2) is used)} \\ &\quad \lesssim h \|(\hat{e}_h^m \cdot n_*^m) n_*^m\|_{L^2(K)} + h \|(I - n_*^m (n_*^m)^\top)\hat{e}_h^m\|_{L^2(K)}, \end{aligned} \tag{3.16}$$

where $\sim$ denotes the norm equivalence relation. For sufficiently small h, the last term on the right-hand side of the inequality above can be absorbed by the left-hand side, and therefore

$$\|(I - n_*^m (n_*^m)^\top)\hat{e}_h^m\|_{L^2(\hat{\Gamma}_{h,*}^m)} \lesssim h \|(\hat{e}_h^m \cdot n_*^m) n_*^m\|_{L^2(\hat{\Gamma}_{h,*}^m)}. \tag{3.17}$$

The corresponding H^1-norm estimate can be proved similarly:

$$\|(I - n_*^m (n_*^m)^\top)\hat{e}_h^m\|_{H^1(\hat{\Gamma}_{h,*}^m)} \lesssim h \|(\hat{e}_h^m \cdot n_*^m) n_*^m\|_{H^1(\hat{\Gamma}_{h,*}^m)}. \tag{3.18}$$

This implies that the tangential component of $\hat{e}_h^m$ is much smaller than its normal component, and the full L^2 and H^1 norms can be bounded by their normal components (for sufficiently small h), i.e.,

$$\|\hat{e}_h^m\|_{L^2(\hat{\Gamma}_{h,*}^m)} \leq 2\|(\hat{e}_h^m \cdot n_*^m)n_*^m\|_{L^2(\hat{\Gamma}_{h,*}^m)}, \tag{3.19}$$

$$\|\hat{e}_h^m\|_{H^1(\hat{\Gamma}_{h,*}^m)} \leq 2\|(\hat{e}_h^m \cdot n_*^m)n_*^m\|_{H^1(\hat{\Gamma}_{h,*}^m)}. \tag{3.20}$$

4 Consistency Estimates

Under the induction assumptions in Sect. 3.3, we present estimates for the consistency error of Dziuk's method for mean curvature flow for $m = 0, \dots, l$. The following lemma, proved in [39, Lemma 4.3], states that the norms of the finite element functions with same nodal vectors on the family of surfaces

$$\hat{\Gamma}_{h,\theta}^m = (1-\theta)\hat{\Gamma}_{h,*}^m + \theta\Gamma_h^m, \quad \theta \in [0,1],$$

are equivalent.

Lemma 4.1 *If* $\|\nabla_{\hat{\Gamma}_{h,*}^m} \hat{e}_h^m\|_{L^\infty(\hat{\Gamma}_{h,*}^m)} \leq \frac{1}{2}$, *then the following equivalence of norms hold for* $1 \leq p \leq \infty$:

$$\|v_h\|_{L^p(\hat{\Gamma}_{h,*}^m)} \lesssim \|v_h\|_{L^p(\hat{\Gamma}_{h,\theta}^m)} \lesssim \|v_h\|_{L^p(\hat{\Gamma}_{h,*}^m)},$$

$$\|\nabla_{\hat{\Gamma}_{h,*}^m} v_h\|_{L^p(\hat{\Gamma}_{h,*}^m)} \lesssim \|\nabla_{\hat{\Gamma}_{h,\theta}^m} v_h\|_{L^p(\hat{\Gamma}_{h,\theta}^m)} \lesssim \|\nabla_{\hat{\Gamma}_{h,*}^m} v_h\|_{L^p(\hat{\Gamma}_{h,*}^m)}.$$

The following lemma states that the error between two integrals due to the perturbation of the surface in the normal direction is $O(h^{k+1})$.

Lemma 4.2 *The following estimates hold for* $f_1, f_2 \in W^{1,\infty}(\hat{\Gamma}_{h,*}^m)$ *and their lifts* $f_1^l, f_2^l \in W^{1,\infty}(\Gamma^m)$:

$$\left|\int_{\hat{\Gamma}_{h,*}^m} f_1 f_2 - \int_{\Gamma^m} f_1^l f_2^l\right| \lesssim (1+\kappa_{*,l})h^{k+1}\|f_1\|_{L^\infty(\hat{\Gamma}_{h,*}^m)}\|f_2\|_{L^2(\hat{\Gamma}_{h,*}^m)}$$

and

$$\left|\int_{\hat{\Gamma}_{h,*}^m} \nabla_{\hat{\Gamma}_{h,*}^m} f_1 \cdot \nabla_{\hat{\Gamma}_{h,*}^m} f_2 - \int_{\Gamma^m} \nabla_{\Gamma^m} f_1^l \cdot \nabla_{\Gamma^m} f_2^l\right| \\ \lesssim (1+\kappa_{*,l})h^{k+1}\|\nabla_{\hat{\Gamma}_{h,*}^m} f_1\|_{L^\infty(\hat{\Gamma}_{h,*}^m)}\|\nabla_{\hat{\Gamma}_{h,*}^m} f_2\|_{L^2(\hat{\Gamma}_{h,*}^m)},$$

for sufficiently small h.

Proof We defined the distortion factor as $\hat{\delta}^m_{h,*}(x) := \mathrm{d}\sigma^m(a^m(x))/\mathrm{d}\hat{\sigma}^m_{h,*}(x)$ for all $x \in \hat{\Gamma}^m_{h,*}$, where σ^m and $\hat{\sigma}^m_{h,*}$ are the measures on Γ^m and $\hat{\Gamma}^m_{h,*}$ respectively. From [22, Proposition 2.5], we have the formula

$$\hat{\delta}^m_{h,*}(x) = n^m(a^m(x)) \cdot \hat{n}^m_{h,*}(x) \prod_{i=1}^{2} \Big(1 - \frac{q^m(x)\kappa^m_i(a^m(x))}{1 + q^m(x)\kappa^m_i(a^m(x))}\Big) \quad \forall x \in \hat{\Gamma}^m_{h,*},$$

where $q^m(x) = |a^m(x) - x| = |a^m(x) - I_h a^m(x)|$ is the distance from $x \in \hat{\Gamma}^m_{h,*}$ to Γ^m, and κ^m_i, $i = 1, 2$ are the principle curvatures of Γ^m. The L^∞ analouge of (3.3) (which follows from (3.3) and the inverse inequality) implies that $\|q^m(x)\|_{L^\infty} \le C_{\kappa_l}(1 + \kappa_{*,l})h^k$. Therefore, based on induction assumption (3) in Sect. 3.3, for sufficiently small h satisfying $(1 + \kappa_{*,l})h^k \le C^{-1}_{\kappa_l}$, $q^m(x)$ is a small quantity such that $1 + q^m(x)\kappa^m_i(a^m(x)) \ge 1/2$. Therefore, using the notation $n^m_*(x) = n^m(a^m(x))$, we have

$$\begin{aligned}
|1 - \hat{\delta}^m_{h,*}(x)| &\le |1 - n^m_*(x) \cdot \hat{n}^m_{h,*}(x)| \\
&\quad + \Big| n^m_*(x) \cdot \hat{n}^m_{h,*}(x) \Big(1 - \prod_{i=1}^{2} \big(1 - \frac{q^m(x)\kappa^m_i(a^m(x))}{1 + q^m(x)\kappa^m_i(a^m(x))}\big)\Big)\Big| \\
&\lesssim |n^m_*(x) \cdot (n^m_*(x) - \hat{n}^m_{h,*}(x))| + |q^m(x)| \\
&\lesssim |n^m_*(x) - \hat{n}^m_{h,*}(x)|^2 + |a^m(x) - I_h a^m(x)| \quad \forall x \in \hat{\Gamma}^m_{h,*},
\end{aligned}$$

where the last inequality follows from the almost orthogonality between $n^m_*(x)$ and $n^m_*(x) - \hat{n}^m_{h,*}(x)$ leading to a squared small term. As a result, using a change of variables and (3.6)–(3.7), we get

$$\begin{aligned}
&\Big| \int_{\hat{\Gamma}^m_{h,*}} f_1 f_2 - \int_{\Gamma^m} f^l_1 f^l_2 \Big| = \Big| \int_{\hat{\Gamma}^m_{h,*}} (1 - \hat{\delta}^m_{h,*}) f_1 f_2 \Big| \\
&\lesssim \|1 - \hat{\delta}^m_{h,*}\|_{L^2(\hat{\Gamma}^m_{h,*})} \|f_1\|_{L^\infty(\hat{\Gamma}^m_{h,*})} \|f_2\|_{L^2(\hat{\Gamma}^m_{h,*})} \\
&\lesssim \big(\|a^m - I_h a^m\|_{L^2(\hat{\Gamma}^m_{h,*})} + \|n^m_* - \hat{n}^m_{h,*}\|_{L^2(\hat{\Gamma}^m_{h,*})} \|n^m_* - \hat{n}^m_{h,*}\|_{L^\infty(\hat{\Gamma}^m_{h,*})}\big) \\
&\quad \|f_1\|_{L^\infty(\hat{\Gamma}^m_{h,*})} \|f_2\|_{L^2(\hat{\Gamma}^m_{h,*})} \\
&\lesssim \big((1 + \kappa_{*,l})h^{k+1} + (1 + \kappa_{*,l})^2 h^{2k-1}\big) \|f_1\|_{L^\infty(\hat{\Gamma}^m_{h,*})} \|f_2\|_{L^2(\hat{\Gamma}^m_{h,*})} \\
&\lesssim (1 + \kappa_{*,l})h^{k+1} \|f_1\|_{L^\infty(\hat{\Gamma}^m_{h,*})} \|f_2\|_{L^2(\hat{\Gamma}^m_{h,*})},
\end{aligned}$$

where the last inequality requires h to satisfy the mesh size condition in (3.8). This proves the first result of Lemma 4.2. The proof of the second result is similar and omitted.

The consistency error of Dziuk's method for mean curvature flow is defined as the following linear functional on $\phi_h \in S_h(\hat{\Gamma}^m_{h,*})$:

$$\begin{aligned} d^m(\phi_h) &:= \int_{\hat{\Gamma}^m_{h,*}} \frac{X^{m+1}_{h,*} - \mathrm{id}}{\tau} \cdot \phi_h + \int_{\hat{\Gamma}^m_{h,*}} \nabla_{\hat{\Gamma}^m_{h,*}} X^{m+1}_{h,*} \cdot \nabla_{\hat{\Gamma}^m_{h,*}} \phi_h \\ &= \int_{\hat{\Gamma}^m_{h,*}} \frac{X^{m+1}_{h,*} - \mathrm{id}}{\tau} \cdot \phi_h + \int_{\Gamma^m} H^m n^m \cdot \phi_h^l \\ &\quad - \int_{\Gamma^m} \nabla_{\Gamma^m} \mathrm{id} \cdot \nabla_{\Gamma^m} \phi_h^l + \int_{\hat{\Gamma}^m_{h,*}} \nabla_{\hat{\Gamma}^m_{h,*}} X^{m+1}_{h,*} \cdot \nabla_{\hat{\Gamma}^m_{h,*}} \phi_h \\ &\quad (\text{since } -\Delta_{\Gamma^m} \mathrm{id} = H^m n^m \text{ on } \Gamma^m) \\ &=: d_1^m(\phi_h) + d_2^m(\phi_h), \end{aligned} \tag{4.1}$$

which is estimated in the following lemma.

Lemma 4.3 *Under the conditions of Theorem 2.1, the following estimate holds for the consistency error*:

$$\begin{aligned} |d^m(\phi_h)| &\lesssim (1 + \kappa_{*,l} h^{k-1}) \tau \|\phi_h\|_{L^2(\hat{\Gamma}^m_{h,*})} \\ &\quad + (1 + \kappa_{*,l}) h^k \|\phi_h\|_{H^1(\hat{\Gamma}^m_{h,*})} \quad \forall \phi_h \in S_h(\hat{\Gamma}^m_{h,*}). \end{aligned}$$

Proof Since $X^{m+1}_{h,*} - \mathrm{id} = I_h(X^{m+1} - \mathrm{id}_{\Gamma^m})$ on $\hat{\Gamma}^m_{h,*}$, where X^{m+1} denotes the local flow map from Γ^m to Γ^{m+1} under mean curvature flow, it follows that $d_1^m(\phi_h)$ can be decomposed into the following parts:

$$\begin{aligned} d_1^m(\phi_h) &= \int_{\hat{\Gamma}^m_{h,*}} \frac{X^{m+1}_{h,*} - \mathrm{id}}{\tau} \cdot \phi_h + \int_{\Gamma^m} H^m n^m \cdot \phi_h^l \\ &= \int_{\hat{\Gamma}^m_{h,*}} \Big(I_h \frac{X^{m+1} - \mathrm{id}_{\Gamma^m}}{\tau} + I_h(H^m n^m) \Big) \cdot \phi_h \\ &\quad - \int_{\hat{\Gamma}^m_{h,*}} (I_h(H^m n^m) - H^{m,-l} n^{m,-l}) \cdot \phi_h \\ &\quad - \int_{\hat{\Gamma}^m_{h,*}} H^{m,-l} n^{m,-l} \cdot \phi_h + \int_{\Gamma^m} H^m n^m \cdot \phi_h^l \\ &=: d_{11}^m(\phi_h) + d_{12}^m(\phi_h) + d_{13}^m(\phi_h). \end{aligned} \tag{4.2}$$

The first term on the right-hand side of (4.2) can be estimated by using Taylor's expansion of equation $\partial_t X = -Hn$ at time level $t = t_m$, which implies that

$$|d_{11}^m(\phi_h)| \lesssim \Big\| \frac{X^{m+1} - \mathrm{id}_{\Gamma^m}}{\tau} + H^m n^m \Big\|_{L^\infty(\Gamma^m)} \|\phi_h\|_{L^2(\hat{\Gamma}^m_{h,*})} \lesssim \tau \|\phi_h\|_{L^2(\hat{\Gamma}^m_{h,*})}.$$

The second and third terms on the right-hand side of (4.2) can be estimated by using the approximation property of the Lagrange interpolation in (3.5) and the geometric perturbation estimates in Lemma 4.2, i.e.,

$$|d_{12}^m(\phi_h)| + |d_{13}^m(\phi_h)| \lesssim (1+\kappa_{*,l})h^{k+1}\|\phi_h\|_{L^2(\hat{\Gamma}_{h,*}^m)}.$$

Similarly, since $X_{h,*}^{m+1} = I_h X^{m+1}$ it follows that $d_2^m(\phi_h)$ can be decomposed into the following parts:

$$\begin{aligned}
d_2^m(\phi_h) &= \int_{\hat{\Gamma}_{h,*}^m} \nabla_{\hat{\Gamma}_{h,*}^m} X_{h,*}^{m+1} \cdot \nabla_{\hat{\Gamma}_{h,*}^m} \phi_h - \int_{\Gamma^m} \nabla_{\Gamma^m} \mathrm{id} \cdot \nabla_{\Gamma^m} \phi_h^l \\
&= \int_{\hat{\Gamma}_{h,*}^m} \nabla_{\hat{\Gamma}_{h,*}^m} (X_{h,*}^{m+1} - X_{h,*}^m) \cdot \nabla_{\hat{\Gamma}_{h,*}^m} \phi_h \\
&\quad + \int_{\hat{\Gamma}_{h,*}^m} \nabla_{\hat{\Gamma}_{h,*}^m} X_{h,*}^m \cdot \nabla_{\hat{\Gamma}_{h,*}^m} \phi_h - \int_{\Gamma^m} \nabla_{\Gamma^m} (X_{h,*}^m)^l \cdot \nabla_{\Gamma^m} \phi_h^l \\
&\quad + \int_{\Gamma^m} \nabla_{\Gamma^m} [(X_{h,*}^m)^l - \mathrm{id}] \cdot \nabla_{\Gamma^m} \phi_h^l \\
&= d_{21}^m(\phi_h) + d_{22}^m(\phi_h) + d_{23}^m(\phi_h).
\end{aligned} \tag{4.3}$$

Since $X_{h,*}^{m+1} - X_{h,*}^m = I_h(X^{m+1} - \mathrm{id}_{\Gamma^m})$ on $\hat{\Gamma}_{h,*}^m$, we furthermore decompose as follows

$$\begin{aligned}
d_{21}^m(\phi_h) &= \int_{\hat{\Gamma}_{h,*}^m} \nabla_{\hat{\Gamma}_{h,*}^m} I_h(X^{m+1} - \mathrm{id}_{\Gamma^m}) \cdot \nabla_{\hat{\Gamma}_{h,*}^m} \phi_h \\
&\quad - \int_{\Gamma^m} \nabla_{\Gamma^m} [I_h(X^{m+1} - \mathrm{id}_{\Gamma^m})]^l \cdot \nabla_{\Gamma^m} \phi_h^l \\
&\quad + \int_{\Gamma^m} \nabla_{\Gamma^m} \Big([I_h(X^{m+1} - \mathrm{id}_{\Gamma^m})]^l - (X^{m+1} - \mathrm{id}_{\Gamma^m}) \Big) \cdot \nabla_{\Gamma^m} \phi_h^l \\
&\quad + \int_{\Gamma^m} \nabla_{\Gamma^m} (X^{m+1} - \mathrm{id}_{\Gamma^m}) \cdot \nabla_{\Gamma^m} \phi_h^l \\
&=: d_{211}^m(\phi_h) + d_{212}^m(\phi_h) + d_{213}^m(\phi_h).
\end{aligned}$$

The first term, $d_{211}^m(\phi_h)$, can be estimated by the geometric perturbation estimates in Lemma 4.2, i.e.,

$$\begin{aligned}
|d_{211}^m(\phi_h)| &\lesssim (1+\kappa_{*,l})h^{k+1}\|I_h(X^{m+1} - \mathrm{id}_{\Gamma^m})\|_{W^{1,\infty}(\hat{\Gamma}_{h,*}^m)} \|\phi_h\|_{H^1(\hat{\Gamma}_{h,*}^m)} \\
&\lesssim (1+\kappa_{*,l})h^{k+1}\tau\|\phi_h\|_{H^1(\hat{\Gamma}_{h,*}^m)} \lesssim (1+\kappa_{*,l})h^k\tau\|\phi_h\|_{L^2(\hat{\Gamma}_{h,*}^m)}.
\end{aligned}$$

The second term, $d_{212}^m(\phi_h)$, can be estimated by using the approximation property of the Lagrange interpolation, i.e.,

$$
\begin{aligned}
|d_{212}^m(\phi_h)| &\lesssim (1+\kappa_{*,l})h^k\|X^{m+1}-\mathrm{id}_{\Gamma^m}\|_{H^{k+1}(\Gamma^m)}\|\phi_h\|_{H^1(\hat{\Gamma}_{h,*}^m)} \\
&\lesssim (1+\kappa_{*,l})h^k\tau\|\phi_h\|_{H^1(\hat{\Gamma}_{h,*}^m)} \lesssim (1+\kappa_{*,l})h^{k-1}\tau\|\phi_h\|_{L^2(\hat{\Gamma}_{h,*}^m)}.
\end{aligned}
$$

The third term, $d_{213}^m(\phi_h)$, can be estimated by using integration by parts (see Lemma 5.1, item 3), i.e.,

$$
|d_{213}^m(\phi_h)| = \left|\int_{\Gamma^m} \Delta_{\Gamma^m}(X^{m+1}-\mathrm{id}_{\Gamma^m})\phi_h^l\right| \lesssim \tau\|\phi_h\|_{L^2(\hat{\Gamma}_{h,*}^m)}.
$$

This proves that

$$
|d_{21}^m(\phi_h)| \lesssim \tau\|\phi_h\|_{L^2(\hat{\Gamma}_{h,*}^m)}.
$$

The second term on the right-hand side of (4.3) can be estimated by using the geometric perturbation estimate (Lemma 4.2), with

$$
|d_{22}^m(\phi_h)| \lesssim (1+\kappa_{*,l})h^{k+1}\|\phi_h\|_{H^1(\hat{\Gamma}_{h,*}^m)}.
$$

Since $(X_{h,*}^m)^l - \mathrm{id} = (I_h\mathrm{id})^l - \mathrm{id}$ on Γ^m, the last term on the right-hand side of (4.3) can be estimated by using the approximation property of the Lagrange interpolation in (3.5), which implies that

$$
|d_{23}^m(\phi_h)| \lesssim (1+\kappa_{*,l})h^k\|\phi_h\|_{H^1(\hat{\Gamma}_{h,*}^m)}.
$$

Combining the estimates above, we obtain the result of Lemma 4.3.

Lemma 4.3 shows that the consistency error is $O(\tau\|\phi_h\|_{L^2(\hat{\Gamma}_{h,*}^m)} + h^k\|\phi_h\|_{H^1(\hat{\Gamma}_{h,*}^m)})$. In Lemma 4.5, we show that the consistency error can be improved to $O(\tau\|\phi_h\|_{L^2(\hat{\Gamma}_{h,*}^m)} + h^{k+1}\|\phi_h\|_{H^1(\hat{\Gamma}_{h,*}^m)})$ if the test function ϕ_h is approximately in the tangential plane. The proof relies on a geometric structure of mean curvature flow and the super-approximation estimates in Lemma 4.4. The proof of the super-approximation estimates in Lemma 4.4 is the same as that of [35, Lemma A], also similar to the proof of (3.16), and therefore omitted.

Lemma 4.4 *Let $v_h, w_h \in S_h(\hat{\Gamma}_{h,*}^m)$ be two finite element functions, and let $T_*^m = I - n_*^m(n_*^m)^\top$ be the tangential projection matrix. Then the following estimates hold:*

$$
\begin{aligned}
\|(1-I_h)T_*^m v_h\|_{L^2(\hat{\Gamma}_{h,*}^m)} &\lesssim h\|v_h\|_{L^2(\hat{\Gamma}_{h,*}^m)}, \\
\|\nabla_{\hat{\Gamma}_{h,*}^m}[(1-I_h)T_*^m v_h]\|_{L^2(\hat{\Gamma}_{h,*}^m)} &\lesssim h\|v_h\|_{H^1(\hat{\Gamma}_{h,*}^m)}, \\
\|(1-I_h)(v_h w_h)\|_{L^1(\hat{\Gamma}_{h,*}^m)} &\lesssim h^2\|v_h\|_{H^1(\hat{\Gamma}_{h,*}^m)}\|w_h\|_{H^1(\hat{\Gamma}_{h,*}^m)}.
\end{aligned}
$$

Now we are in a position to establish the following improved estimate for the consistency error, which will be used for estimating the tangential motion given by Dziuk's parametric FEM.

Lemma 4.5 *Under the conditions of Theorem 2.1, the following estimate holds for* $\phi_h \in S_h(\hat\Gamma^m_{h,*})$:

$$|d^m(I_hT^m_*\phi_h)| \lesssim (1+\kappa_{*,l}h^{k-1})\tau\|I_hT^m_*\phi_h\|_{L^2(\hat\Gamma^m_{h,*})} + (1+\kappa_{*,l})h^{k+1}\|I_hT^m_*\phi_h\|_{H^1(\hat\Gamma^m_{h,*})}. \tag{4.4}$$

Proof From the proof of Lemma 4.3 we see that all the terms in $d^m_{1j}(\phi_h)$ and $d^m_{2j}(\phi_h)$, $j=1,2,3$, are bounded by

$$(1+\kappa_{*,l}h^{k-1})\tau\|\phi_h\|_{L^2(\hat\Gamma^m_{h,*})} + (1+\kappa_{*,l})h^{k+1}\|\phi_h\|_{H^1(\hat\Gamma^m_{h,*})}$$

except $d^m_{23}(\phi_h)$, which is only $O(h^k\|\phi_h\|_{H^1(\hat\Gamma^m_{h,*})})$. Since $I_hT^m_*\phi_h = I_hT^m_*I_hT^m_*\phi_h$, it suffices to prove the following result:

$$|d^m_{23}(I_hT^m_*\phi_h)| \lesssim (1+\kappa_{*,l})h^{k+1}\|\phi_h\|_{H^1(\hat\Gamma^m_{h,*})}.$$

Then replacing ϕ_h by $I_hT^m_*\phi_h$ yields the desired estimate in Lemma 4.5.

Since $X^m_{h,*} = \mathrm{id}$ on $\hat\Gamma^m_{h,*}$, it follows that $(X^m_{h,*})^l = \mathrm{id}^l$ on Γ^m. Therefore, we can rewrite $d^m_{23}(I_hT^m_*\phi_h)$ into the following form:

$$\begin{aligned} d^m_{23}(I_hT^m_*\phi_h) &= \int_{\Gamma^m} \nabla_{\Gamma^m}(\mathrm{id}^l-\mathrm{id})\cdot\nabla_{\Gamma^m}(I_hT^m_*\phi_h)^l \\ &= \int_{\Gamma^m} \nabla_{\Gamma^m}(\mathrm{id}^l-\mathrm{id})\cdot\nabla_{\Gamma^m}T^m\phi_h^l \\ &\quad + \int_{\Gamma^m} \nabla_{\Gamma^m}(\mathrm{id}^l-\mathrm{id})\cdot\nabla_{\Gamma^m}[(I_hT^m_*\phi_h)^l - T^m\phi_h^l] \\ &=: d^m_{231}(\phi_h) + d^m_{232}(\phi_h). \end{aligned} \tag{4.5}$$

Since $T^m = I - n^m(n^m)^\top$ and T^m_* is the inverse lift of T^m onto $\hat\Gamma^m_{h,*}$, it follows that

$$\begin{aligned} d^m_{231}(\phi_h) &= \int_{\Gamma^m} \nabla_{\Gamma^m}(\mathrm{id}^l-\mathrm{id})\cdot[(\nabla_{\Gamma^m}\phi_h^l)(I-n^m(n^m)^\top)] \\ &\quad + \int_{\Gamma^m} \nabla_{\Gamma^m}(\mathrm{id}^l-\mathrm{id})\cdot[\nabla_{\Gamma^m}(I-n^m(n^m)^\top)]\phi_h^l \\ &= \int_{\Gamma^m} \nabla_{\Gamma^m}[(I-n^m(n^m)^\top)(\mathrm{id}^l-\mathrm{id})]\nabla_{\Gamma^m}\phi_h^l \\ &\quad - \int_{\Gamma^m} [\nabla_{\Gamma^m}(I-n^m(n^m)^\top)](\mathrm{id}^l-\mathrm{id})\cdot\nabla_{\Gamma^m}\phi_h^l \end{aligned}$$

$$+\int_{\Gamma^m} \nabla_{\Gamma^m}(\mathrm{id}^l - \mathrm{id}) \cdot [\nabla_{\Gamma^m}(I - n^m (n^m)^\top)]\phi_h^l. \qquad (4.6)$$

The orthogonality of $\mathrm{id}^l - \mathrm{id}$ to the tangent plane of Γ^m, i.e.,

$$(I - n^m (n^m)^\top)(\mathrm{id}^l - \mathrm{id}) = 0 \quad \text{on} \quad \Gamma^m,$$

is a geometric structure which can help to improve the consistency error by one order. Therefore, only the second and third terms on the right-hand side of (4.6) are not zero. Since the third term on the right-hand side of (4.6) can be estimated by using integration by parts, which removes the partial derivative from $(\mathrm{id}^l - \mathrm{id})$, it follows that

$$|d_{231}^m(\phi_h)| \lesssim (1 + \kappa_{*,l})h^{k+1}\|\phi_h^l\|_{H^1(\Gamma^m)} \lesssim (1 + \kappa_{*,l})h^{k+1}\|\phi_h\|_{H^1(\hat{\Gamma}_{h,*}^m)}. \qquad (4.7)$$

Since $I_h T_*^m \phi_h = I_h(T^m \phi_h^l)$, the last term on the right-hand side of (4.5) can be estimated by using the super-approximation estimates in Lemma 4.4, which implies that

$$|d_{232}^m(\phi_h)| \lesssim \|\mathrm{id}^l - \mathrm{id}\|_{H^1(\Gamma^m)} h \|\phi_h\|_{H^1(\hat{\Gamma}_{h,*}^m)} \lesssim (1 + \kappa_{*,l})h^{k+1}\|\phi_h\|_{H^1(\hat{\Gamma}_{h,*}^m)}.$$

This proves Lemma 4.5.

5 Stability Estimates

In this section, we present the stability estimates for Dziuk's parametric FEM for mean curvature flow by utilizing the new approach outlined in the introduction section and the general settings in Sect. 3. Under the induction assumptions in Sect. 3.3, the stability estimates in this section hold for $m = 0, \ldots, l$.

To simplify the computations, we first introduce the partial derivatives on surfaces. The i-th component of the surface derivative, denoted by $\underline{D}_i$, is defined as the i-th component of the surface gradient, i.e. $\underline{D}_i u := (\nabla_\Gamma u)_i$ for $i = 1, ..., 3$. Given the local parametrization $(\{\frac{\partial}{\partial \theta^i}\}_{i=1}^2, F)$, we define the push-forward of the i-th coordinate vector field as $A_i := F_*(\frac{\partial}{\partial \theta^i})$. Then we have the following local formula

$$\begin{aligned}(\underline{D}_i u) \circ F = \left(\nabla_\Gamma u \cdot \frac{\partial}{\partial x^i}\right) \circ F &= g^{jk}\left(A_j u \left(A_k \cdot \frac{\partial}{\partial x^i}\right)\right) \circ F \\ &= g^{jk} \frac{\partial (u \circ F)}{\partial \theta^j} \cdot \frac{\partial F_i}{\partial \theta^k},\end{aligned} \qquad (5.1)$$

where $\{g^{jk}\}_{j,k=1}^2$ is the local representation of the metric tensor of Γ with respect to the local coordinate frame $\{\frac{\partial}{\partial \theta^i}\}_{i=1}^2$. Upon the local formula (5.1) of $\underline{D}$, the correspondent Leibniz rule, chain rule, integration-by-parts and commutators formulas can be established.

Lemma 5.1 *Given two smooth surfaces* Γ, Γ' *and functions* $f, h \in C^\infty(\Gamma)$, $g \in C^\infty(\Gamma'; \Gamma)$. *The following identities hold*

1. $\underline{D}_i(fh) = \underline{D}_i fh + f\underline{D}_i h$.
2. $\underline{D}_i(f \circ g) = \underline{D}_j f \circ g \cdot \underline{D}_i g_j$.
3. *If* Γ *is closed then* $\int_\Gamma \underline{D}_i f = \int_\Gamma fHn_i$, *where* n *is the normal direction and* $H := \underline{D}_i n_i$ *is the mean curvature.*
4. $\underline{D}_i\underline{D}_j f = \underline{D}_j\underline{D}_i f + n_i H_{jl}\underline{D}_l f - n_j H_{il}\underline{D}_l f$ *where the symmetric matrix* $H_{ij} := \underline{D}_i n_j = \underline{D}_j n_i$ *is the matrix representation of the shape operator* $\mathrm{d}n : T\Gamma \to T\mathbb{S}^2$.
5. *If* Γ *evolves under the velocity field* v *whose graph we denote by* $G_T := \cup_{t\in[0,T]}\Gamma(t) \times t$. *For* $f \in C^2(G_T)$ *we have*
$$\partial_t^\bullet(\underline{D}_i f) = \underline{D}_i(\partial_t^\bullet f) - (\underline{D}_i v_j - n_i n_l \underline{D}_j v_l)\underline{D}_j f$$
where $\partial_t^\bullet$ *to denote the material derivative with respect to* v.
6. *If* $f, h \in C^2(G_T)$, *then it holds that*
$$\frac{\mathrm{d}}{\mathrm{d}t}\int_\Gamma fh = \int_\Gamma \partial_t^\bullet fh + \int_\Gamma f\partial_t^\bullet h + \int_\Gamma fh(\nabla_\Gamma \cdot v).$$
The divergence is defined as $\nabla_\Gamma \cdot v := \underline{D}_i v_i$ *and coincides with the surface divergence if* v *is a vector field on* Γ.

Proof The first two relations are obvious from the local formula of $\underline{D}$; see (5.1). The third relation is shown in [34, Lemma 16.1]. The fourth and fifth equalities are proved in [28, Lemma 2.4 and 2.6], and the proof of the last formula can be found in [26, Appendix A].

5.1 The Error Equation

The error equation follows from subtracting the consistency equation (4.1) from the scheme (1.2), i.e.,

$$\begin{aligned}
&\int_{\Gamma_h^m} \frac{X_h^{m+1} - X_h^m}{\tau} \cdot \phi_h - \int_{\hat{\Gamma}_{h,*}^m} \frac{X_{h,*}^{m+1} - \hat{X}_{h,*}^m}{\tau} \cdot \phi_h \\
&+ \int_{\Gamma_h^m} \nabla_{\Gamma_h^m} X_h^{m+1} \cdot \nabla_{\Gamma_h^m}\phi_h - \int_{\hat{\Gamma}_{h,*}^m} \nabla_{\hat{\Gamma}_{h,*}^m} X_{h,*}^{m+1} \cdot \nabla_{\hat{\Gamma}_{h,*}^m}\phi_h \\
&= -d^m(\phi_h),
\end{aligned} \tag{5.2}$$

where

$$\begin{aligned}
&\int_{\Gamma_h^m} \frac{X_h^{m+1} - X_h^m}{\tau} \cdot \phi_h - \int_{\hat{\Gamma}_{h,*}^m} \frac{X_{h,*}^{m+1} - \hat{X}_{h,*}^m}{\tau} \cdot \phi_h \\
&= \int_{\hat{\Gamma}_{h,*}^m} \frac{e_h^{m+1} - \hat{e}_h^m}{\tau} \cdot \phi_h + \left(\int_{\Gamma_h^m} - \int_{\hat{\Gamma}_{h,*}^m}\right)\frac{X_h^{m+1} - X_h^m}{\tau} \cdot \phi_h
\end{aligned}$$

$$=! : \int_{\hat{\Gamma}^m_{h,*}} \frac{e_h^{m+1} - \hat{e}_h^m}{\tau} \cdot \phi_h + J^m(\phi_h). \tag{5.3}$$

Since X_h^{m+1}, X_h^m and ϕ_h are viewed as finite element functions on the family of intermediate surfaces $\hat{\Gamma}^m_{h,\theta} = (1-\theta)\hat{\Gamma}^m_{h,*} + \theta\Gamma^m_h$, $\theta \in [0, 1]$, with fixed nodal values, they have the following transport property: $\partial^\bullet_\theta X_h^{m+1} = \partial^\bullet_\theta X_h^m = \partial^\bullet_\theta \phi_h = 0$. Therefore,

$$\begin{aligned} J^m(\phi_h) &= \int_{\hat{\Gamma}^m_{h,\theta}} \frac{X_h^{m+1} - X_h^m}{\tau} \cdot \phi_h \bigg|_{\theta=0}^{\theta=1} \\ &= \int_0^1 \frac{\mathrm{d}}{\mathrm{d}\theta} \int_{\hat{\Gamma}^m_{h,\theta}} \frac{X_h^{m+1} - X_h^m}{\tau} \cdot \phi_h \mathrm{d}\theta \\ &= \int_0^1 \int_{\hat{\Gamma}^m_{h,\theta}} \frac{X_h^{m+1} - X_h^m}{\tau} \cdot \phi_h (\nabla_{\hat{\Gamma}^m_{h,\theta}} \cdot \hat{e}_h^m) \mathrm{d}\theta \\ &= \int_0^1 \int_{\hat{\Gamma}^m_{h,\theta}} \Big(\frac{e_h^{m+1} - \hat{e}_h^m}{\tau} - I_h(H^m n^m - g^m) \Big) \cdot \phi_h (\nabla_{\hat{\Gamma}^m_{h,\theta}} \cdot \hat{e}_h^m) \mathrm{d}\theta, \end{aligned} \tag{5.4}$$

where we have used (3.15) in the last equality. The remainder $J^m(\phi_h)$ can be estimated by using the norm equivalence of the surfaces $\hat{\Gamma}^m_{h,*}$ and $\hat{\Gamma}^m_{h,\theta}$ (cf. Lemma 4.1), i.e.,

$$\begin{aligned} |J^m(\phi_h)| &\lesssim \|\nabla_{\hat{\Gamma}^m_{h,*}} \hat{e}_h^m\|_{L^2(\hat{\Gamma}^m_{h,*})} \|\phi_h\|_{L^2(\hat{\Gamma}^m_{h,*})} \\ &\quad + \Big\| \frac{e_h^{m+1} - \hat{e}_h^m}{\tau} \Big\|_{L^2(\hat{\Gamma}^m_{h,*})} \|\nabla_{\hat{\Gamma}^m_{h,*}} \hat{e}_h^m\|_{L^\infty(\hat{\Gamma}^m_{h,*})} \|\phi_h\|_{L^2(\hat{\Gamma}^m_{h,*})}. \end{aligned} \tag{5.5}$$

5.2 Recovery of Full H^1 Parabolicity

The following formula accounts for the error of H^1 bilinear forms from surface discrepancy; see [37, Lemma 7.1].

Lemma 5.2

$$\begin{aligned} &\int_{\Gamma^m_h} \nabla_{\Gamma^m_h} w_h \cdot \nabla_{\Gamma^m_h} z_h - \int_{\hat{\Gamma}^m_{h,*}} \nabla_{\hat{\Gamma}^m_{h,*}} w_h \cdot \nabla_{\hat{\Gamma}^m_{h,*}} z_h \\ &\quad = \int_0^1 \int_{\hat{\Gamma}^m_{h,\theta}} \nabla_{\hat{\Gamma}^m_{h,\theta}} w_h \cdot (D_{\hat{\Gamma}^m_{h,\theta}} \hat{e}_h^m) \nabla_{\hat{\Gamma}^m_{h,\theta}} z_h \mathrm{d}\theta \end{aligned} \tag{5.6}$$

where $(D_{\hat{\Gamma}^m_{h,\theta}} v)_{rl} := -\underline{D}_l v_r - \underline{D}_r v_l + \delta_{rl} \underline{D}_m v_m$.

For $z_h = e_h^{m+1}$, the right hand side of (5.6) is a dominant error which cannot be trivially written into some positive-definite bilinear form, and this is the main difficulty in the numerical analysis. In [1, 41], a geometric structure was discovered and used to cancel out the tangential part of the full stiffness matrix, leading to an H^1 parabolic structure of the normal component. However, the H^1 parabolic structure of the tangential component is still missing.

To describe the normal and tangential components of the stiffness matrix in a clearer way, we define the following symmetric bilinear forms for any two $\mathbb{R}^3$-valued functions u and v on Γ:

$$\begin{aligned} A_\Gamma(u,v) &:= \int_\Gamma \nabla_\Gamma u \cdot \nabla_\Gamma v, \\ A_\Gamma^N(u,v) &:= \int_\Gamma [(\nabla_\Gamma u)n]\cdot[(\nabla_\Gamma v)n], \\ A_\Gamma^T(u,v) &:= \int_\Gamma \mathrm{tr}\big[(\nabla_\Gamma u)(I-nn^\top)(\nabla_\Gamma v)^T\big], \\ B_\Gamma(u,v) &:= \int_\Gamma (\nabla_\Gamma\cdot u)(\nabla_\Gamma\cdot v) - \mathrm{tr}(\nabla_\Gamma u \nabla_\Gamma v). \end{aligned} \tag{5.7}$$

Thus $A_\Gamma(u,v) = A_\Gamma^N(u,v) + A_\Gamma^T(u,v)$. These bilinear forms can be defined similarly on the approximate surfaces, e.g., $\hat\Gamma_{h,*}^m$ and $\hat\Gamma_{h,\theta}^m$. By using the identity $\nabla_\Gamma \mathrm{id} = I - nn^\top =: P$, we find the following relation:

$$\begin{aligned} \int_\Gamma \nabla_\Gamma \mathrm{id}\cdot(D_\Gamma u)\nabla_\Gamma v &= \int_\Gamma P_{ri}(-\underline{D}_l u_r - \underline{D}_r u_l + \delta_{rl}\underline{D}_m u_m)\underline{D}_l v_i \\ &= -\int_\Gamma P_{ri}\underline{D}_l u_r \underline{D}_l v_i - \int_\Gamma P_{ri}\underline{D}_r u_l \underline{D}_l v_i + \int_\Gamma P_{ri}\underline{D}_m u_m \underline{D}_r v_i \\ &= -\int_\Gamma \mathrm{tr}\big[(\nabla_\Gamma u)P(\nabla_\Gamma v)^T\big] - \int_\Gamma \underline{D}_r u_l \underline{D}_l v_r + \int_\Gamma \underline{D}_m u_m \underline{D}_r v_r \\ &= -A_\Gamma^T(u,v) + B_\Gamma(u,v). \end{aligned} \tag{5.8}$$

This formula also holds for the approximate surfaces $\hat\Gamma_{h,*}^m$ and $\hat\Gamma_{h,\theta}^m$.

According to [1, Eq. (2.1)] and the surface calculus in Lemma 5.1, if the underlying surface is sufficiently smooth, then the symmetric bilinear form $B_\Gamma(u,v)$ can be furthermore written into the following form via integration by parts:

$$\begin{aligned} B_\Gamma(u,v) = &\int_\Gamma u_j \underline{D}_i v_i H n_j - \int_\Gamma u_j \underline{D}_j v_i H n_i \\ &+ \int_\Gamma u_j \underline{D}_k v_i n_i H_{jk} - \int_\Gamma u_j \underline{D}_k v_i H_{ik} n_j \quad \text{for } u,v\in H^1(\Gamma). \end{aligned} \tag{5.9}$$

Now we are in a good position to estimate the error coming from the stiffness matrix. If we define $\hat X_{h,\theta}^m := (1-\theta)\hat X_{h,*}^m + \theta X_h^m$ and $X_{h,\theta}^{m+1} := (1-\theta)X_{h,*}^{m+1} + \theta X_h^{m+1}$ in the sense of nodal vectors, then from the fundamental theorem of calculus we obtain

$$\begin{aligned} &\int_{\Gamma_h^m} \nabla_{\Gamma_h^m} X_h^{m+1}\cdot\nabla_{\Gamma_h^m}\phi_h - \int_{\hat\Gamma_{h,*}^m} \nabla_{\hat\Gamma_{h,*}^m} X_{h,*}^{m+1}\cdot\nabla_{\hat\Gamma_{h,*}^m}\phi_h \\ &= \int_{\hat\Gamma_{h,\theta}^m} \nabla_{\hat\Gamma_{h,\theta}^m} X_{h,\theta}^{m+1}\cdot\nabla_{\hat\Gamma_{h,\theta}^m}\phi_h \Big|_{\theta=0}^{\theta=1} \end{aligned}$$

$$
\begin{aligned}
&= \int_0^1 \frac{\mathrm{d}}{\mathrm{d}\theta} \int_{\hat{\Gamma}^m_{h,\theta}} \nabla_{\hat{\Gamma}^m_{h,\theta}} X^{m+1}_{h,\theta} \cdot \nabla_{\hat{\Gamma}^m_{h,\theta}} \phi_h \mathrm{d}\theta \\
&= \int_0^1 \int_{\hat{\Gamma}^m_{h,\theta}} \nabla_{\hat{\Gamma}^m_{h,\theta}} e^{m+1}_h \cdot \nabla_{\hat{\Gamma}^m_{h,\theta}} \phi_h \mathrm{d}\theta + \int_0^1 \int_{\hat{\Gamma}^m_{h,\theta}} \nabla_{\hat{\Gamma}^m_{h,\theta}} X^{m+1}_{h,\theta} \cdot D_{\hat{\Gamma}^m_{h,\theta}} \hat{e}^m_h \nabla_{\hat{\Gamma}^m_{h,\theta}} \phi_h \mathrm{d}\theta
\end{aligned}
$$

(Lemma 5.1 (item 5) and Lemma 5.2 are used)

$$
\begin{aligned}
&= \int_0^1 \int_{\hat{\Gamma}^m_{h,\theta}} \Big(\nabla_{\hat{\Gamma}^m_{h,\theta}} e^{m+1}_h \cdot \nabla_{\hat{\Gamma}^m_{h,\theta}} \phi_h + \nabla_{\hat{\Gamma}^m_{h,\theta}} \hat{X}^m_{h,\theta} \cdot D_{\hat{\Gamma}^m_{h,\theta}} e^{m+1}_h \nabla_{\hat{\Gamma}^m_{h,\theta}} \phi_h \Big) \mathrm{d}\theta \\
&\quad - \int_0^1 \int_{\hat{\Gamma}^m_{h,\theta}} \nabla_{\hat{\Gamma}^m_{h,\theta}} \hat{X}^m_{h,\theta} \cdot D_{\hat{\Gamma}^m_{h,\theta}} (e^{m+1}_h - \hat{e}^m_h) \nabla_{\hat{\Gamma}^m_{h,\theta}} \phi_h \mathrm{d}\theta \\
&\quad + \int_0^1 \int_{\hat{\Gamma}^m_{h,\theta}} \nabla_{\hat{\Gamma}^m_{h,\theta}} (X^{m+1}_{h,\theta} - \hat{X}^m_{h,\theta}) \cdot D_{\hat{\Gamma}^m_{h,\theta}} \hat{e}^m_h \nabla_{\hat{\Gamma}^m_{h,\theta}} \phi_h \mathrm{d}\theta \\
&= \int_0^1 [A_{\hat{\Gamma}^m_{h,\theta}} (e^{m+1}_h, \phi_h) - A^T_{\hat{\Gamma}^m_{h,\theta}} (e^{m+1}_h, \phi_h) + B_{\hat{\Gamma}^m_{h,\theta}} (e^{m+1}_h, \phi_h)] \mathrm{d}\theta \\
&\quad + \int_0^1 [A^T_{\hat{\Gamma}^m_{h,\theta}} (e^{m+1}_h - \hat{e}^m_h, \phi_h) - B_{\hat{\Gamma}^m_{h,\theta}} (e^{m+1}_h - \hat{e}^m_h, \phi_h)] \mathrm{d}\theta \\
&\qquad\qquad\qquad (\text{The relations } \hat{X}^m_{h,\theta} = \text{id on } \hat{\Gamma}^m_{h,\theta} \text{ and (5.8) is used}) \\
&\quad + \int_0^1 \int_{\hat{\Gamma}^m_{h,\theta}} \nabla_{\hat{\Gamma}^m_{h,\theta}} (X^{m+1}_{h,\theta} - \hat{X}^m_{h,\theta}) \cdot D_{\hat{\Gamma}^m_{h,\theta}} \hat{e}^m_h \nabla_{\hat{\Gamma}^m_{h,\theta}} \phi_h \mathrm{d}\theta \\
&= A^N_{h,*}(e^{m+1}_h, \phi_h) + A^T_{h,*}(e^{m+1}_h - \hat{e}^m_h, \phi_h) + B^m(\hat{e}^m_h, \phi_h) + K^m(\phi_h)
\end{aligned} \tag{5.10}
$$

where we have used the following notations for simplicity:

$$
A^N_{h,*}(u_h, v_h) := A^N_{\hat{\Gamma}^m_{h,*}} (u_h, v_h) \quad \text{and} \quad A^T_{h,*}(u_h, v_h) := A^T_{\hat{\Gamma}^m_{h,*}} (u_h, v_h), \tag{5.11}
$$

$$
B^m(u_h, v_h) := B_{\Gamma^m}(u^l_h, v^l_h) \tag{5.12}
$$

$$
\begin{aligned}
K^m(\phi_h) &= \int_0^1 \big[A^N_{\hat{\Gamma}^m_{h,\theta}} (e^{m+1}_h, \phi_h) - A^N_{\hat{\Gamma}^m_{h,*}} (e^{m+1}_h, \phi_h) \big] \mathrm{d}\theta \\
&\quad + \int_0^1 \big[A^T_{\hat{\Gamma}^m_{h,\theta}} (e^{m+1}_h - \hat{e}^m_h, \phi_h) - A^T_{\hat{\Gamma}^m_{h,*}} (e^{m+1}_h - \hat{e}^m_h, \phi_h) \big] \mathrm{d}\theta \\
&\quad + \int_0^1 \big[B_{\hat{\Gamma}^m_{h,\theta}} (\hat{e}^m_h, \phi_h) - B_{\hat{\Gamma}^m_{h,*}} (\hat{e}^m_h, \phi_h) \big] \mathrm{d}\theta \\
&\quad + B_{\hat{\Gamma}^m_{h,*}} (\hat{e}^m_h, \phi_h) - B_{\Gamma^m} (\hat{e}^{m,l}_h, \phi^l_h) \\
&\quad + \int_0^1 \int_{\hat{\Gamma}^m_{h,\theta}} \nabla_{\hat{\Gamma}^m_{h,\theta}} (X^{m+1}_{h,\theta} - \hat{X}^m_{h,\theta}) \cdot D_{\hat{\Gamma}^m_{h,\theta}} \hat{e}^m_h \nabla_{\hat{\Gamma}^m_{h,\theta}} \phi_h \mathrm{d}\theta \\
&=: K^m_1(\phi_h) + K^m_2(\phi_h) + K^m_3(\phi_h) + K^m_4(\phi_h) + K^m_5(\phi_h).
\end{aligned} \tag{5.13}
$$

The error equation (5.2) can be written into the following form by using the expressions in (5.3) and (5.10):

$$\int_{\hat{\Gamma}^m_{h,*}} \frac{e_h^{m+1}-\hat{e}_h^m}{\tau}\cdot\phi_h + A^N_{h,*}(e_h^{m+1},\phi_h) + A^T_{h,*}(e_h^{m+1}-\hat{e}_h^m,\phi_h) + B^m(\hat{e}_h^m,\phi_h)$$
$$= -J^m(\phi_h) - K^m(\phi_h) - d^m(\phi_h). \tag{5.14}$$

The recovery of full H^1 parabolicity becomes clear after we test (5.10) with $\phi_h = e_h^{m+1}$:

$$\begin{aligned}
&\int_{\Gamma_h^m} \nabla_{\Gamma_h^m} X_h^{m+1}\cdot\nabla_{\Gamma_h^m} e_h^{m+1} - \int_{\hat{\Gamma}^m_{h,*}} \nabla_{\hat{\Gamma}^m_{h,*}} X_{h,*}^{m+1}\cdot\nabla_{\hat{\Gamma}^m_{h,*}} e_h^{m+1}\\
&\quad = A^N_{h,*}(e_h^{m+1},e_h^{m+1}) + A^T_{h,*}(e_h^{m+1}-\hat{e}_h^m,e_h^{m+1}) + B^m(\hat{e}_h^m,e_h^{m+1}) + K^m(e_h^{m+1})\\
&\quad \ge A^N_{h,*}(e_h^{m+1},e_h^{m+1}) + \frac{1}{2}A^T_{h,*}(e_h^{m+1},e_h^{m+1}) - \frac{1}{2}A^T_{h,*}(\hat{e}_h^m,\hat{e}_h^m)\\
&\quad + B^m(\hat{e}_h^m,e_h^{m+1}) + K^m(e_h^{m+1})\\
&\quad \ge \frac{1}{2}A_{h,*}(e_h^{m+1},e_h^{m+1}) - \frac{1}{2}A^T_{h,*}(\hat{e}_h^m,\hat{e}_h^m) + B^m(\hat{e}_h^m,e_h^{m+1}) + K^m(e_h^{m+1}).
\end{aligned}\tag{5.15}$$

The full H^1 parabolicity stems from the dominant term $\frac{1}{2}A_{h,*}(e_h^{m+1},e_h^{m+1})$ and the fact that $\frac{1}{2}A^T_{h,*}(\hat{e}_h^m,\hat{e}_h^m)$ is much smaller than $\frac{1}{2}A_{h,*}(e_h^{m+1},e_h^{m+1})$ due to the orthogonality between $\hat{e}_h^m$ and the tangent plane of Γ^m at the nodes. This can be seen from the following calculations:

$$\begin{aligned}
&|A^T_{h,*}(\hat{e}_h^m,\hat{e}_h^m)|\\
&= \Big|\int_{\hat{\Gamma}^m_{h,*}} \mathrm{tr}\Big((\nabla_{\hat{\Gamma}^m_{h,*}}\hat{e}_h^m)[I-n^m_{h,*}(n^m_{h,*})^\top]^2(\nabla_{\hat{\Gamma}^m_{h,*}}\hat{e}_h^m)^\top\Big)\Big|\\
&\lesssim \Big|\int_{\hat{\Gamma}^m_{h,*}} \mathrm{tr}\Big((\nabla_{\hat{\Gamma}^m_{h,*}}\hat{e}_h^m)[I-n^m_{*}(n^m_{*})^\top]^2(\nabla_{\hat{\Gamma}^m_{h,*}}\hat{e}_h^m)^\top\Big)\Big|\\
&\quad + \|n^m_{h,*}-n^m_*\|_{L^\infty(\hat{\Gamma}^m_{h,*})}\|\nabla_{\hat{\Gamma}^m_{h,*}}\hat{e}_h^m\|^2_{L^2(\hat{\Gamma}^m_{h,*})}\\
&\lesssim \|\nabla_{\hat{\Gamma}^m_{h,*}}\big([I-n^m_*(n^m_*)^\top]\hat{e}_h^m\big)\|^2_{L^2(\hat{\Gamma}^m_{h,*})} + \|\nabla_{\hat{\Gamma}^m_{h,*}}[I-n^m_*(n^m_*)^\top]\hat{e}_h^m\|^2_{L^2(\hat{\Gamma}^m_{h,*})}\\
&\quad + \|n^m_{h,*}-n^m_*\|_{L^\infty(\hat{\Gamma}^m_{h,*})}\|\nabla_{\hat{\Gamma}^m_{h,*}}\hat{e}_h^m\|^2_{L^2(\hat{\Gamma}^m_{h,*})} \quad \text{(Leibniz rule and Young's inequality)}\\
&\lesssim h^2\|\hat{e}_h^m\|^2_{H^1(\hat{\Gamma}^m_{h,*})} + \|\hat{e}_h^m\|^2_{L^2(\hat{\Gamma}^m_{h,*})} + \|n^m_{h,*}-n^m_*\|_{L^\infty(\hat{\Gamma}^m_{h,*})}\|\nabla_{\hat{\Gamma}^m_{h,*}}\hat{e}_h^m\|^2_{L^2(\hat{\Gamma}^m_{h,*})}\\
&\qquad \text{(Inequality (3.18) is used)}\\
&\lesssim \|\hat{e}_h^m\|^2_{L^2(\hat{\Gamma}^m_{h,*})} + h^{1.5}\|\nabla_{\hat{\Gamma}^m_{h,*}}\hat{e}_h^m\|^2_{L^2(\hat{\Gamma}^m_{h,*})},
\end{aligned}\tag{5.16}$$

where we have used (3.9) in the last inequality. From (5.9), Young's inequality and norm equivalence between the surfaces Γ^m and $\hat{\Gamma}^m_{h,*}$, we also obtain the following result:

$$
\begin{aligned}
|B^m(\hat{e}_h^m, e_h^{m+1})| &= \Big| \int_{\Gamma^m} \hat{e}_{h,j}^{m,l} \underline{D}_i e_{h,i}^{m+1,l} H^m n_j^m - \int_{\Gamma^m} \hat{e}_{h,j}^{m,l} \underline{D}_j e_{h,i}^{m+1,l} H^m n_i^m \\
&\quad + \int_{\Gamma^m} \hat{e}_{h,j}^{m,l} \underline{D}_k e_{h,i}^{m+1,l} n_i^m H_{jk}^m - \int_{\Gamma^m} \hat{e}_{h,j}^{m,l} \underline{D}_k e_{h,i}^{m+1,l} n_j^m H_{ik}^m \Big| \\
&\lesssim \epsilon^{-1} \|\hat{e}_h^m\|_{L^2(\hat{\Gamma}_{h,*}^m)}^2 + \epsilon \|\nabla_{\hat{\Gamma}_{h,*}^m} e_h^{m+1}\|_{L^2(\hat{\Gamma}_{h,*}^m)}^2,
\end{aligned}
\tag{5.17}
$$

where ϵ can be arbitrarily small.

Therefore, both $A_{h,*}^T(\hat{e}_h^m, \hat{e}_h^m)$ and $B^m(\hat{e}_h^m, e_h^{m+1})$ are much smaller than the term $\frac{1}{2}A_{h,*}(e_h^{m+1}, e_h^{m+1})$ in (5.15). This recovery of the full H^1 parabolicity heavily depends on the orthogonality between $\hat{e}_h^m$ and the tangent plane of Γ^m. This orthogonality arises from our definition of the error $\hat{e}_h^m$ in terms of the distance projection onto Γ^m.

5.3 Boundedness of Velocity

In this subsection, we present the estimates for the velocity of the numerical solution by testing the error equation in (5.14) with the error of the velocity, i.e. $\phi_h = (e_h^{m+1} - \hat{e}_h^m)/\tau$. Since $A_{h,*}^N\big(e_h^{m+1} - \hat{e}_h^m, (e_h^{m+1} - \hat{e}_h^m)/\tau\big) \geq 0$ and $A_{h,*}^T\big(e_h^{m+1} - \hat{e}_h^m, (e_h^{m+1} - \hat{e}_h^m)/\tau\big) \geq 0$, we obtain the following relation:

$$
\begin{aligned}
&\int_{\hat{\Gamma}_{h,*}^m} \frac{e_h^{m+1} - \hat{e}_h^m}{\tau} \cdot \frac{e_h^{m+1} - \hat{e}_h^m}{\tau} \\
&\leq -A_{h,*}^N\Big(e_h^{m+1}, \frac{e_h^{m+1} - \hat{e}_h^m}{\tau}\Big) - A_{h,*}^T\Big(e_h^{m+1} - \hat{e}_h^m, \frac{e_h^{m+1} - \hat{e}_h^m}{\tau}\Big) \\
&\quad - B^m\Big(\hat{e}_h^m, \frac{e_h^{m+1} - \hat{e}_h^m}{\tau}\Big) \\
&\quad - J^m\Big(\frac{e_h^{m+1} - \hat{e}_h^m}{\tau}\Big) - K^m\Big(\frac{e_h^{m+1} - \hat{e}_h^m}{\tau}\Big) - d^m\Big(\frac{e_h^{m+1} - \hat{e}_h^m}{\tau}\Big) \\
&\leq -A_{h,*}^N\Big(\hat{e}_h^m, \frac{e_h^{m+1} - \hat{e}_h^m}{\tau}\Big) - B^m\Big(\hat{e}_h^m, \frac{e_h^{m+1} - \hat{e}_h^m}{\tau}\Big) \\
&\quad - J^m\Big(\frac{e_h^{m+1} - \hat{e}_h^m}{\tau}\Big) - K^m\Big(\frac{e_h^{m+1} - \hat{e}_h^m}{\tau}\Big) - d^m\Big(\frac{e_h^{m+1} - \hat{e}_h^m}{\tau}\Big).
\end{aligned}
\tag{5.18}
$$

By using the estimate (5.5) with $\phi_h = (e_h^{m+1} - \hat{e}_h^m)/\tau$, the following estimate can be derived:

$$
\begin{aligned}
\Big| J^m\Big(\frac{e_h^{m+1} - \hat{e}_h^m}{\tau}\Big)\Big| &\lesssim \|\nabla_{\hat{\Gamma}_{h,*}^m} \hat{e}_h^m\|_{L^2(\hat{\Gamma}_{h,*}^m)} \Big\| \frac{e_h^{m+1} - \hat{e}_h^m}{\tau} \Big\|_{L^2(\hat{\Gamma}_{h,*}^m)} \\
&\quad + \|\nabla_{\hat{\Gamma}_{h,*}^m} \hat{e}_h^m\|_{L^\infty(\hat{\Gamma}_{h,*}^m)} \Big\| \frac{e_h^{m+1} - \hat{e}_h^m}{\tau} \Big\|_{L^2(\hat{\Gamma}_{h,*}^m)}^2.
\end{aligned}
\tag{5.19}
$$

By using the expression of $K^m(\phi_h)$ in (5.13), the following estimates can be derived for $j = 1, 2, 3$ from the fundamental theorem of calculus (analogous to the formula in Lemma 5.2):

$$
\begin{aligned}
&|K_j^m(\phi_h)| \\
&\lesssim \|\nabla_{\hat\Gamma^m_{h,*}} \hat e^m_h\|_{L^\infty(\hat\Gamma^m_{h,*})}\Big(\|\nabla_{\hat\Gamma^m_{h,*}} \hat e^m_h\|_{L^2(\hat\Gamma^m_{h,*})} + \|\nabla_{\hat\Gamma^m_{h,*}} e^{m+1}_h\|_{L^2(\hat\Gamma^m_{h,*})}\Big)\|\nabla_{\hat\Gamma^m_{h,*}} \phi_h\|_{L^2(\hat\Gamma^m_{h,*})},
\end{aligned}
$$

and the following estimate can be obtained by using the geometric perturbation estimate (cf. [36, Lemma 5.6]):

$$
\begin{aligned}
|K_4^m(\phi_h)| &\lesssim (1+\kappa_{*,l})h^k \|\nabla_{\hat\Gamma^m_{h,*}} \hat e^m_h\|_{L^\infty(\hat\Gamma^m_{h,*})}\|\nabla_{\hat\Gamma^m_{h,*}} \phi_h\|_{L^2(\hat\Gamma^m_{h,*})} \\
&\lesssim (1+\kappa_{*,l})h^{k-1} \|\nabla_{\hat\Gamma^m_{h,*}} \hat e^m_h\|_{L^2(\hat\Gamma^m_{h,*})}\|\nabla_{\hat\Gamma^m_{h,*}} \phi_h\|_{L^2(\hat\Gamma^m_{h,*})}. \qquad (5.20)
\end{aligned}
$$

By using the relation

$$
\begin{aligned}
X^{m+1}_{h,\theta} - \hat X^m_{h,\theta} &= X^{m+1}_h - X^m_h - (1-\theta)(e^{m+1}_h - \hat e^m_h) \\
&= \theta(e^{m+1}_h - \hat e^m_h) + \tau I_h(v^m + g^m), \quad \text{(Relation (3.15) is used)},
\end{aligned}
$$

we have

$$
\begin{aligned}
|K_5^m(\phi_h)| &= \Big|\int_0^1 \int_{\hat\Gamma^m_{h,\theta}} \nabla_{\hat\Gamma^m_{h,\theta}}(X^{m+1}_{h,\theta} - \hat X^m_{h,\theta}) \cdot D_{\hat\Gamma^m_{h,\theta}} \hat e^m_h \nabla_{\hat\Gamma^m_{h,\theta}} \phi_h \mathrm{d}\theta\Big| \\
&= \Big|\int_0^1 \int_{\hat\Gamma^m_{h,\theta}} \nabla_{\hat\Gamma^m_{h,\theta}}\Big(\theta(e^{m+1}_h - \hat e^m_h) - \tau I_h(H^m n^m - g^m)\Big) \cdot D_{\hat\Gamma^m_{h,\theta}} \hat e^m_h \nabla_{\hat\Gamma^m_{h,\theta}} \phi_h \mathrm{d}\theta\Big| \\
&\lesssim \|\nabla_{\hat\Gamma^m_{h,*}} \hat e^m_h\|_{L^\infty(\hat\Gamma^m_{h,*})}\Big(\|\nabla_{\hat\Gamma^m_{h,*}} \hat e^m_h\|_{L^2(\hat\Gamma^m_{h,*})} + \|\nabla_{\hat\Gamma^m_{h,*}} e^{m+1}_h\|_{L^2(\hat\Gamma^m_{h,*})}\Big)\|\nabla_{\hat\Gamma^m_{h,*}} \phi_h\|_{L^2(\hat\Gamma^m_{h,*})} \\
&\quad + \tau \|\nabla_{\hat\Gamma^m_{h,*}} \hat e^m_h\|_{L^2(\hat\Gamma^m_{h,*})}\|\nabla_{\hat\Gamma^m_{h,*}} \phi_h\|_{L^2(\hat\Gamma^m_{h,*})}. \qquad (5.21)
\end{aligned}
$$

Therefore, by collecting the above estimates, we obtain the following result:

$$
\begin{aligned}
|K^m(\phi_h)| &\lesssim \|\nabla_{\hat\Gamma^m_{h,*}} \hat e^m_h\|_{L^\infty(\hat\Gamma^m_{h,*})}\Big(\|\nabla_{\hat\Gamma^m_{h,*}} \hat e^m_h\|_{L^2(\hat\Gamma^m_{h,*})} + \|\nabla_{\hat\Gamma^m_{h,*}} e^{m+1}_h\|_{L^2(\hat\Gamma^m_{h,*})}\Big)\|\nabla_{\hat\Gamma^m_{h,*}} \phi_h\|_{L^2(\hat\Gamma^m_{h,*})} \\
&\quad + [\tau + (1+\kappa_{*,l})h^{k-1}]\|\nabla_{\hat\Gamma^m_{h,*}} \hat e^m_h\|_{L^2(\hat\Gamma^m_{h,*})}\|\nabla_{\hat\Gamma^m_{h,*}} \phi_h\|_{L^2(\hat\Gamma^m_{h,*})}. \qquad (5.22)
\end{aligned}
$$

By decomposing e^{m+1}_h into $(e^{m+1}_h - \hat e^m_h) + \hat e^m_h$ on the right-hand side of (5.22), we obtain the following result:

$$
\begin{aligned}
\Big|K^m(\phi_h)\Big| &\lesssim \|\nabla_{\hat\Gamma^m_{h,*}} \hat e^m_h\|_{L^\infty(\hat\Gamma^m_{h,*})}\|\nabla_{\hat\Gamma^m_{h,*}} \hat e^m_h\|_{L^2(\hat\Gamma^m_{h,*})}\|\nabla_{\hat\Gamma^m_{h,*}} \phi_h\|_{L^2(\hat\Gamma^m_{h,*})} \\
&\quad + \|\nabla_{\hat\Gamma^m_{h,*}} \hat e^m_h\|_{L^\infty(\hat\Gamma^m_{h,*})}\|\nabla_{\hat\Gamma^m_{h,*}} (e^{m+1}_h - \hat e^m_h)\|_{L^2(\hat\Gamma^m_{h,*})}\|\nabla_{\hat\Gamma^m_{h,*}} \phi_h\|_{L^2(\hat\Gamma^m_{h,*})} \\
&\quad + [\tau + (1+\kappa_{*,l})h^{k-1}]\|\nabla_{\hat\Gamma^m_{h,*}} \hat e^m_h\|_{L^2(\hat\Gamma^m_{h,*})}\|\nabla_{\hat\Gamma^m_{h,*}} \phi_h\|_{L^2(\hat\Gamma^m_{h,*})} \\
&\lesssim h^{-2}\|\nabla_{\hat\Gamma^m_{h,*}} \hat e^m_h\|^2_{L^2(\hat\Gamma^m_{h,*})}\|\phi_h\|_{L^2(\hat\Gamma^m_{h,*})}
\end{aligned}
$$

$$+\tau h^{-2}\|\nabla_{\hat{\Gamma}^m_{h,*}}\hat{e}^m_h\|_{L^\infty(\hat{\Gamma}^m_{h,*})}\left\|\frac{e^{m+1}_h-\hat{e}^m_h}{\tau}\right\|_{L^2(\hat{\Gamma}^m_{h,*})}$$

$$\|\phi_h\|_{L^2(\hat{\Gamma}^m_{h,*})}\quad\text{(inverse inequality)}$$

$$+h^{-1}[\tau+(1+\kappa_{*,l})h^{k-1}]\|\nabla_{\hat{\Gamma}^m_{h,*}}\hat{e}^m_h\|_{L^2(\hat{\Gamma}^m_{h,*})}\|\phi_h\|_{L^2(\hat{\Gamma}^m_{h,*})}. \tag{5.23}$$

Under the stepsize condition $\tau = o(h^{2.5})$ and the mesh size condition in (3.8), the following result can be derived from (5.23) by using the estimates in (3.11):

$$\left|K^m\Big(\frac{e^{m+1}_h-\hat{e}^m_h}{\tau}\Big)\right|$$

$$\lesssim h^{-0.5}\|\nabla_{\hat{\Gamma}^m_{h,*}}\hat{e}^m_h\|_{L^2(\hat{\Gamma}^m_{h,*})}\left\|\frac{e^{m+1}_h-\hat{e}^m_h}{\tau}\right\|_{L^2(\hat{\Gamma}^m_{h,*})}+h\left\|\frac{e^{m+1}_h-\hat{e}^m_h}{\tau}\right\|^2_{L^2(\hat{\Gamma}^m_{h,*})}. \tag{5.24}$$

Furthermore, by applying (5.9) and the inverse inequality, it is straightforward to show that

$$\left|B^m\Big(\hat{e}^m_h,\frac{e^{m+1}_h-\hat{e}^m_h}{\tau}\Big)\right|\lesssim\|\nabla_{\hat{\Gamma}^m_{h,*}}\hat{e}^m_h\|_{L^2(\hat{\Gamma}^m_{h,*})}\left\|\frac{e^{m+1}_h-\hat{e}^m_h}{\tau}\right\|_{L^2(\hat{\Gamma}^m_{h,*})}, \tag{5.25}$$

$$\left|A^N_{h,*}\Big(\hat{e}^m_h,\frac{e^{m+1}_h-\hat{e}^m_h}{\tau}\Big)\right|\lesssim h^{-1}\|\nabla_{\hat{\Gamma}^m_{h,*}}\hat{e}^m_h\|_{L^2(\hat{\Gamma}^m_{h,*})}\left\|\frac{e^{m+1}_h-\hat{e}^m_h}{\tau}\right\|_{L^2(\hat{\Gamma}^m_{h,*})}. \tag{5.26}$$

The estimate for the last term on the right-hand side of (5.18) follows from Lemma 4.3, i.e.,

$$\left|d^m\Big(\frac{e^{m+1}_h-\hat{e}^m_h}{\tau}\Big)\right|\lesssim[(1+\kappa_{*,l}h^{k-1})\tau+(1+\kappa_{*,l})h^{k-1}]\left\|\frac{e^{m+1}_h-\hat{e}^m_h}{\tau}\right\|_{L^2(\hat{\Gamma}^m_{h,*})}, \tag{5.27}$$

Then, substituting these estimates into (5.18), we obtain the following result under the mesh size condition in (3.8):

$$\left\|\frac{e^{m+1}_h-\hat{e}^m_h}{\tau}\right\|_{L^2(\hat{\Gamma}^m_{h,*})}\lesssim h^{-1}\|\nabla_{\hat{\Gamma}^m_{h,*}}\hat{e}^m_h\|_{L^2(\hat{\Gamma}^m_{h,*})}+\tau+(1+\kappa_{*,l})h^{k-1}. \tag{5.28}$$

5.4 Estimates of $\hat{e}^{m+1}_h$ in Terms of e^{m+1}_h

In this subsection, we prove the existence of the interpolated surface $\hat{\Gamma}^{m+1}_{h,*}$ and present estimates for $\|\nabla_{\hat{\Gamma}^m_{h,*}}\hat{e}^{m+1}_h\|_{L^2(\hat{\Gamma}^m_{h,*})}$ in terms of $\|\nabla_{\hat{\Gamma}^m_{h,*}}e^{m+1}_h\|_{L^2(\hat{\Gamma}^m_{h,*})}$ and $\|e^{m+1}_h\|_{L^2(\hat{\Gamma}^m_{h,*})}$ by using the estimate of the velocity in (5.28).

From (5.28) we see that, under the mesh size condition in (3.8),

$$\|e_h^{m+1} - \hat{e}_h^m\|_{L^\infty(\hat{\Gamma}_{h,*}^m)} \lesssim h^{-1}\|e_h^{m+1} - \hat{e}_h^m\|_{L^2(\hat{\Gamma}_{h,*}^m)}$$
$$\lesssim \tau h^{-1}\big[h^{-1}\|\nabla_{\hat{\Gamma}_{h,*}^m} \hat{e}_h^m\|_{L^2(\hat{\Gamma}_{h,*}^m)} + \tau + (1+\kappa_{*,l})h^{k-1}\big] \lesssim h^{-0.5}\tau \tag{5.29}$$

and therefore

$$\|e_h^{m+1}\|_{L^\infty(\hat{\Gamma}_{h,*}^m)} \le \|e_h^{m+1} - \hat{e}_h^m\|_{L^\infty(\hat{\Gamma}_{h,*}^m)} + \|\hat{e}_h^m\|_{L^\infty(\hat{\Gamma}_{h,*}^m)} \lesssim h^{-0.5}\tau + h^{1.5} \lesssim h^{1.5}, \tag{5.30}$$

where the last inequality uses the stepsize condition $\tau = o(h^{2.5})$. Since the estimates above hold for $m = 0, \ldots, l$, it follows that, for sufficiently small mesh size h, the nodes of the numerically computed surface Γ_h^{l+1} are in a δ-neighborhood of the exact surface Γ^{l+1} and correspondingly, the distance projection of the nodes of Γ_h^{l+1} onto Γ^{l+1} are well defined (thus the interpolated surface $\hat{\Gamma}_{h,*}^{l+1}$ is well defined). This recovers the first induction assumption in Sect. 3.3 at time level t_{l+1}.

Note that $\hat{e}_h^{m+1}$ is also small in view of (3.12)–(3.13). Since e_h^{m+1} is small, the value of $\hat{X}_{h,*}^{m+1} - X_{h,*}^{m+1}$ at the nodes is approximately the tangential projection of e_h^{m+1}. Besides, $T^{m+1} \circ \hat{X}_{h,*}^{m+1}$ is well defined at the nodes since $\hat{X}_{h,*}^{m+1}$ takes value on Γ^{m+1} there. Hence the following inequality holds at the nodes:

$$\begin{aligned}
&|\hat{X}_{h,*}^{m+1} - X_{h,*}^{m+1}| \\
&\lesssim |(T^{m+1} \circ \hat{X}_{h,*}^{m+1})e_h^{m+1}| \\
&\lesssim |[T^{m+1} \circ \hat{X}_{h,*}^{m+1} - T^{m+1} \circ X_{h,*}^{m+1}]e_h^{m+1}| + |[T^{m+1} \circ X_{h,*}^{m+1} - T^m \circ \hat{X}_{h,*}^m]e_h^{m+1}| \\
&\quad + |(T^m \circ \hat{X}_{h,*}^m)(e_h^{m+1} - \hat{e}_h^m)| \qquad ((T^m \circ \hat{X}_{h,*}^m)\hat{e}_h^m = 0 \text{ due to orthogonality}) \\
&\lesssim |\hat{X}_{h,*}^{m+1} - X_{h,*}^{m+1}||e_h^{m+1}| + |X_{h,*}^{m+1} - \hat{X}_{h,*}^m||e_h^{m+1}| + |(T^m \circ \hat{X}_{h,*}^m)(e_h^{m+1} - \hat{e}_h^m)|.
\end{aligned}$$

Since $|e_h^{m+1}| \lesssim h^{1.5}$, the first term on the right-hand side can be absorbed by the left-hand side. Since $|X_{h,*}^{m+1} - \hat{X}_{h,*}^m| = |X^{m+1} - \mathrm{id}| = O(\tau)$ at the nodes on Γ^m, it follows that

$$|\hat{X}_{h,*}^{m+1} - X_{h,*}^{m+1}| \lesssim \tau|e_h^{m+1}| + |(T^m \circ \hat{X}_{h,*}^m)(e_h^{m+1} - \hat{e}_h^m)| \quad \text{at the nodes.}$$

Therefore,

$$\begin{aligned}
|n^{m+1} \circ \hat{X}_{h,*}^{m+1} - n^{m+1} \circ X_{h,*}^{m+1}| &\lesssim |\hat{X}_{h,*}^{m+1} - X_{h,*}^{m+1}| \\
&\lesssim \tau|e_h^{m+1}| + |T_*^m(e_h^{m+1} - \hat{e}_h^m)| \quad \text{at the nodes.}
\end{aligned}$$

For the simplicity of notation, we use the same notation n_*^{m+1} to denote the pull-back function $n_*^{m+1} \circ \hat{X}_{h,*}^{m+1}$ on $\hat{\Gamma}_{h,*}^m$. Then the following relation holds at the nodes (in view of the last inequality above):

$$
\begin{aligned}
n_*^{m+1} - n_*^m &= n^{m+1} \circ \hat{X}_{h,*}^{m+1} - n^m \circ \hat{X}_{h,*}^m \\
&= (n^{m+1} \circ \hat{X}_{h,*}^{m+1} - n^{m+1} \circ X_{h,*}^{m+1}) + (n^{m+1} \circ X_{h,*}^{m+1} - n^m \circ \hat{X}_{h,*}^m) \\
&\lesssim \tau |e_h^{m+1}| + |T_*^m (e_h^{m+1} - \hat{e}_h^m)| + |X_{h,*}^{m+1} - \hat{X}_{h,*}^m| \\
&\lesssim \tau |e_h^{m+1}| + |T_*^m (e_h^{m+1} - \hat{e}_h^m)| + \tau \\
&\lesssim \tau + |T_*^m (e_h^{m+1} - \hat{e}_h^m)| \quad \text{at the nodes.} \qquad (5.31)
\end{aligned}
$$

By using this result, we have

$$
\begin{aligned}
&\| I_h([I - n_*^{m+1}(n_*^{m+1})^\top] e_h^{m+1}) \|_{L^\infty(\hat{\Gamma}_{h,*}^m)} \\
&\lesssim \max_{\text{at nodes}} |[I - n_*^{m+1}(n_*^{m+1})^\top] e_h^{m+1}| \\
&\leq \max_{\text{at nodes}} |[I - n_*^{m+1}(n_*^{m+1})^\top](e_h^{m+1} - \hat{e}_h^m)| \\
&\quad + \max_{\text{at nodes}} |[n_*^m (n_*^m)^\top - n_*^{m+1}(n_*^{m+1})^\top] \hat{e}_h^m| \quad ([I - n_*^m (n_*^m)^\top] \hat{e}_h^m = 0 \text{ at nodes}) \\
&\lesssim \| e_h^{m+1} - \hat{e}_h^m \|_{L^\infty(\hat{\Gamma}_{h,*}^m)} + (\tau + \| e_h^{m+1} - \hat{e}_h^m \|_{L^\infty(\hat{\Gamma}_{h,*}^m)}) \| \hat{e}_h^m \|_{L^\infty(\hat{\Gamma}_{h,*}^m)} \\
&\lesssim \tau h^{-1} \Big\| \frac{e_h^{m+1} - \hat{e}_h^m}{\tau} \Big\|_{L^2(\hat{\Gamma}_{h,*}^m)} \\
&\quad + \tau h^{-1} \| \hat{e}_h^m \|_{L^2(\hat{\Gamma}_{h,*}^m)} \quad (\text{inverse inequality, } \| \hat{e}_h^m \|_{L^\infty(\hat{\Gamma}_{h,*}^m)} \lesssim 1) \\
&\lesssim \tau h^{-2} \| \hat{e}_h^m \|_{H^1(\hat{\Gamma}_{h,*}^m)} + \tau h^{-1} [\tau + (1 + \kappa_{*,l}) h^{k-1}] \qquad (\text{here (5.28) is used}). \quad (5.32)
\end{aligned}
$$

Since "$|f_h| \lesssim |g_h|$ at nodes" implies "$\| f_h \|_{L^2(\hat{\Gamma}_{h,*}^m)} \lesssim \| g_h \|_{L^2(\hat{\Gamma}_{h,*}^m)}$" for any two finite element functions f_h and g_h, it follows that

$$
\begin{aligned}
&\| \nabla_{\hat{\Gamma}_{h,*}^m} f_h \|_{L^2(\hat{\Gamma}_{h,*}^m)} \\
&\lesssim h^{-1} \| f_h \|_{L^2(\hat{\Gamma}_{h,*}^m)} \quad (\text{inverse inequality is used}) \\
&\lesssim h^{-1} \big\| |I_h([I - n_*^{m+1}(n_*^{m+1})^\top] e_h^{m+1})|^2 \big\|_{L^2(\hat{\Gamma}_{h,*}^m)} \quad (\text{inequality (3.13) is used}) \\
&\lesssim h^{-1} \| I_h([I - n_*^{m+1}(n_*^{m+1})^\top] e_h^{m+1}) \|_{L^\infty(\hat{\Gamma}_{h,*}^m)} \| e_h^{m+1} \|_{L^2(\hat{\Gamma}_{h,*}^m)} \\
&\lesssim \tau h^{-3} (\| \hat{e}_h^m \|_{H^1(\hat{\Gamma}_{h,*}^m)} + \tau h + (1 + \kappa_{*,l}) h^k) \| e_h^{m+1} \|_{L^2(\hat{\Gamma}_{h,*}^m)} \quad (\text{here (5.32) is used}). \\
&\qquad (5.33)
\end{aligned}
$$

Therefore, using relation $\hat{e}_h^{m+1} = I_h\big[(e_h^{m+1} \cdot n_*^{m+1}) n_*^{m+1}\big] + f_h$ in (3.12), we have

$$
\begin{aligned}
&\| \nabla_{\hat{\Gamma}_{h,*}^m} \hat{e}_h^{m+1} \|_{L^2(\hat{\Gamma}_{h,*}^m)} \\
&\lesssim \| \nabla_{\hat{\Gamma}_{h,*}^m} I_h [n_*^{m+1} (n_*^{m+1})^\top e_h^{m+1}] \|_{L^2(\hat{\Gamma}_{h,*}^m)} + \| \nabla_{\hat{\Gamma}_{h,*}^m} f_h \|_{L^2(\hat{\Gamma}_{h,*}^m)} \\
&\lesssim \| \nabla_{\hat{\Gamma}_{h,*}^m} I_h [(n_*^{m+1} (n_*^{m+1})^\top - n_*^m (n_*^m)^\top) e_h^{m+1}] \|_{L^2(\hat{\Gamma}_{h,*}^m)}
\end{aligned}
$$

$$
\begin{aligned}
&+\|\nabla_{\hat\Gamma^m_{h,*}} I_h[n^m_*(n^m_*)^\top e^{m+1}_h]\|_{L^2(\hat\Gamma^m_{h,*})} + \|\nabla_{\hat\Gamma^m_{h,*}} f_h\|_{L^2(\hat\Gamma^m_{h,*})} \\
&\lesssim h^{-1}\|I_h[(n^{m+1}_*(n^{m+1}_*)^\top - n^m_*(n^m_*)^\top)e^{m+1}_h]\|_{L^\infty(\hat\Gamma^m_{h,*})} \\
&\quad + \|\nabla_{\hat\Gamma^m_{h,*}}[n^m_*(n^m_*)^\top e^{m+1}_h]\|_{L^2(\hat\Gamma^m_{h,*})} + \|\nabla_{\hat\Gamma^m_{h,*}}(1-I_h)[n^m_*(n^m_*)^\top e^{m+1}_h]\|_{L^2(\hat\Gamma^m_{h,*})} \\
&\quad + \tau h^{-3}(\|\hat e^m_h\|_{H^1(\hat\Gamma^m_{h,*})} + \tau h + (1+\kappa_{*,l})h^k)\|e^{m+1}_h\|_{L^2(\hat\Gamma^m_{h,*})} \quad \text{(here (5.33) are used)} \\
&\lesssim \|e^{m+1}_h\|_{L^2(\hat\Gamma^m_{h,*})} + \|\nabla_{\hat\Gamma^m_{h,*}} e^{m+1}_h\|_{L^2(\hat\Gamma^m_{h,*})},
\end{aligned}
\tag{5.34}
$$

where have used $|n^{m+1}_*(n^{m+1}_*)^\top - n^m_*(n^m_*)^\top| \lesssim \tau + |e^{m+1}_h - \hat e^m_h| \lesssim h^2$ at the nodes, which follows from (5.31) and the stepsize condition $\tau = o(h^{2.5})$. In addition, we have used the inverse inequality and (3.10), as well as the mesh size condition in (3.8), in the derivation of the last inequality.

5.5 Norm Equivalence on the Surfaces Γ^m_h, Γ^{m+1}_h, $\hat\Gamma^m_{h,*}$, $\hat\Gamma^{m+1}_{h,*}$ and $\Gamma^{m+1}_{h,*}$

By using the results in Sects. 5.3 and 5.4, and using the inverse inequality of finite element functions, we obtain the following results from (5.28) and (3.10) under the mesh size condition in (3.8):

$$
\|e^{m+1}_h - \hat e^m_h\|_{H^1(\hat\Gamma^m_{h,*})} \lesssim \tau h^{-2}(\|\nabla_{\hat\Gamma^m_{h,*}} \hat e^m_h\|_{L^2(\hat\Gamma^m_{h,*})} + \tau h + (1+\kappa_{*,l})h^k) \le o(h^2),
\tag{5.35}
$$

$$
\|e^{m+1}_h - \hat e^m_h\|_{W^{1,\infty}(\hat\Gamma^m_{h,*})} \lesssim h^{-1}\|e^{m+1}_h - \hat e^m_h\|_{H^1(\hat\Gamma^m_{h,*})} \le o(h).
\tag{5.36}
$$

From (3.10) we also obtain $\|\hat e^m_h\|_{H^1(\hat\Gamma^m_{h,*})} \lesssim h^{1.5}$ and $\|\hat e^m_h\|_{W^{1,\infty}(\hat\Gamma^m_{h,*})} \lesssim h^{0.5}$ (as a result of the inverse inequality). These results imply that

$$
\|e^{m+1}_h\|_{H^1(\hat\Gamma^m_{h,*})} \lesssim h^{1.5} \quad \text{and} \quad \|e^{m+1}_h\|_{W^{1,\infty}(\hat\Gamma^m_{h,*})} \lesssim h^{0.5}.
\tag{5.37}
$$

Substituting these results into (5.34) and using the inverse inequality (the L^2 norm of $\hat e^{m+1}_h$ is similar as (5.34) and omitted), we obtain

$$
\|\hat e^{m+1}_h\|_{H^1(\hat\Gamma^m_{h,*})} \lesssim h^{1.5} \quad \text{and} \quad \|\hat e^{m+1}_h\|_{W^{1,\infty}(\hat\Gamma^m_{h,*})} \lesssim h^{0.5}.
\tag{5.38}
$$

For the smooth function $X^{m+1} - \mathrm{id}$ defined on Γ^m, we have

$$
\begin{aligned}
\|X^{m+1}_{h,*} - \hat X^m_{h,*}\|_{W^{1,\infty}(\hat\Gamma^m_{h,*})} &= \|I_h[X^{m+1} - \mathrm{id}]\|_{W^{1,\infty}(\hat\Gamma^m_{h,*})} \\
&\lesssim \|X^{m+1} - \mathrm{id}\|_{W^{1,\infty}(\Gamma^m)} \lesssim \tau,
\end{aligned}
$$

and therefore

$$
\|X^{m+1}_h - X^m_h\|_{W^{1,\infty}(\hat\Gamma^m_{h,*})}
$$

$$= \|e_h^{m+1} - \hat{e}_h^m + \tau I_h(v^m + g^m)\|_{W^{1,\infty}(\hat{\Gamma}_{h,*}^m)} \lesssim h + \tau \quad \text{(relation (3.15) is used).}$$

This implies that

$$\begin{aligned}
&\|\hat{X}_{h,*}^{m+1} - \hat{X}_{h,*}^m\|_{W^{1,\infty}(\hat{\Gamma}_{h,*}^m)} \\
&\le \|\hat{X}_{h,*}^{m+1} - X_h^{m+1}\|_{W^{1,\infty}(\hat{\Gamma}_{h,*}^m)} + \|X_h^{m+1} - X_h^m\|_{W^{1,\infty}(\hat{\Gamma}_{h,*}^m)} + \|X_h^m - \hat{X}_{h,*}^m\|_{W^{1,\infty}(\hat{\Gamma}_{h,*}^m)} \\
&\le \|\hat{e}_h^{m+1}\|_{W^{1,\infty}(\hat{\Gamma}_{h,*}^m)} + \|X_h^{m+1} - X_h^m\|_{W^{1,\infty}(\hat{\Gamma}_{h,*}^m)} + \|\hat{e}_h^m\|_{W^{1,\infty}(\hat{\Gamma}_{h,*}^m)} \\
&\lesssim h^{0.5}, \qquad (5.39)
\end{aligned}$$

where we have used (5.38) in the last inequality. By the norm equivalence in Lemma 4.1, for $\tau = o(h^{2.5})$ and sufficiently small h satisfying the mesh size condition in (3.8), the L^p and $W^{1,p}$ norms of a finite element function v_h (with fixed nodal vector) on the surfaces Γ_h^m, Γ_h^{m+1}, $\hat{\Gamma}_{h,*}^m$, $\hat{\Gamma}_{h,*}^{m+1}$, $\Gamma_{h,*}^{m+1}$ are all equivalent for $p \in [1, \infty]$.

5.6 Improved Estimates for the Tangential Velocity

We show that the tangential component of the velocity has better estimates than the general result proved in (5.28). The improved estimates for the tangential velocity established in this subsection could help us to convert $\|e_h^{m+1}\|^2_{L^2(\hat{\Gamma}_{h,*}^m)}$ to $\|\hat{e}_h^{m+1}\|^2_{L^2(\hat{\Gamma}_{h,*}^{m+1})}$ with coefficient 1 (up to some $O(\tau)$ remainder).

In order to estimate the tangential component of the velocity, we test the error equation in (5.14) with

$$\phi_h = \frac{1}{\tau} I_h T_*^m (e_h^{m+1} - \hat{e}_h^m) = \frac{1}{\tau} I_h\big([I - n_*^m (n_*^m)^\top](e_h^{m+1} - \hat{e}_h^m)\big).$$

This yields the following relation:

$$\begin{aligned}
&\int_{\hat{\Gamma}_{h,*}^m} \frac{e_h^{m+1} - \hat{e}_h^m}{\tau} \cdot I_h T_*^m \frac{e_h^{m+1} - \hat{e}_h^m}{\tau} \\
&= -A_{h,*}^N\Big(e_h^{m+1}, T_*^m I_h T_*^m \frac{e_h^{m+1} - \hat{e}_h^m}{\tau}\Big) - A_{h,*}^T\Big(e_h^{m+1} - \hat{e}_h^m, T_*^m I_h T_*^m \frac{e_h^{m+1} - \hat{e}_h^m}{\tau}\Big) \\
&\quad - A_{h,*}^N\Big(e_h^{m+1}, (I_h T_*^m - T_*^m) I_h T_*^m \frac{e_h^{m+1} - \hat{e}_h^m}{\tau}\Big) \\
&\quad - A_{h,*}^T\Big(e_h^{m+1} - \hat{e}_h^m, (I_h T_*^m - T_*^m) I_h T_*^m \frac{e_h^{m+1} - \hat{e}_h^m}{\tau}\Big) \\
&\quad - B^m\Big(\hat{e}_h^m, I_h T_*^m \frac{e_h^{m+1} - \hat{e}_h^m}{\tau}\Big) - J^m\Big(I_h T_*^m \frac{e_h^{m+1} - \hat{e}_h^m}{\tau}\Big) \\
&\quad - K^m\Big(I_h T_*^m \frac{e_h^{m+1} - \hat{e}_h^m}{\tau}\Big) - d^m\Big(I_h T_*^m \frac{e_h^{m+1} - \hat{e}_h^m}{\tau}\Big). \qquad (5.40)
\end{aligned}$$

We present estimates for each terms on the right-hand side of (5.40) separately.

First, the third and fourth terms can be estimated by the super-approximation estimates in Lemma 4.4 and the inverse inequality:

$$\begin{aligned}
&\Big|A_{h,*}^N\Big(e_h^{m+1},(I_hT_*^m-T_*^m)I_hT_*^m\frac{e_h^{m+1}-\hat{e}_h^m}{\tau}\Big)\Big|\\
&\lesssim \|\nabla_{\hat{\Gamma}_{h,*}^m}e_h^{m+1}\|_{L^2(\hat{\Gamma}_{h,*}^m)}\Big\|I_hT_*^m\frac{e_h^{m+1}-\hat{e}_h^m}{\tau}\Big\|_{L^2(\hat{\Gamma}_{h,*}^m)}, \qquad (5.41)\\
&\Big|A_{h,*}^T\Big(e_h^{m+1}-\hat{e}_h^m,(I_hT_*^m-T_*^m)I_hT_*^m\frac{e_h^{m+1}-\hat{e}_h^m}{\tau}\Big)\Big|\\
&\lesssim \|\nabla_{\hat{\Gamma}_{h,*}^m}(e_h^{m+1}-\hat{e}_h^m)\|_{L^2(\hat{\Gamma}_{h,*}^m)}\Big\|I_hT_*^m\frac{e_h^{m+1}-\hat{e}_h^m}{\tau}\Big\|_{L^2(\hat{\Gamma}_{h,*}^m)}. \qquad (5.42)
\end{aligned}$$

Second, by considering the definition of B^m in (5.12) and (5.9), it is straightforward to show that

$$\Big|B^m\Big(\hat{e}_h^m,I_hT_*^m\frac{e_h^{m+1}-\hat{e}_h^m}{\tau}\Big)\Big|\lesssim\|\nabla_{\hat{\Gamma}_{h,*}^m}\hat{e}_h^m\|_{L^2(\hat{\Gamma}_{h,*}^m)}\Big\|I_hT_*^m\frac{e_h^{m+1}-\hat{e}_h^m}{\tau}\Big\|_{L^2(\hat{\Gamma}_{h,*}^m)}. \qquad (5.43)$$

Next, from the expression of J^m in (5.4) we see that

$$\begin{aligned}
J^m(I_hT_*^m\phi_h)&=\int_0^1\int_{\hat{\Gamma}_{h,\theta}^m}\frac{e_h^{m+1}-\hat{e}_h^m}{\tau}\cdot I_hT_*^m\phi_h(\nabla_{\hat{\Gamma}_{h,\theta}^m}\cdot\hat{e}_h^m)\mathrm{d}\theta\\
&\quad-\int_0^1\int_{\hat{\Gamma}_{h,\theta}^m}I_h(H^mn^m+g^m)\cdot I_hT_*^m\phi_h(\nabla_{\hat{\Gamma}_{h,\theta}^m}\cdot\hat{e}_h^m)\mathrm{d}\theta\\
&=:J_1^m(I_hT_*^m\phi_h)+J_2^m(I_hT_*^m\phi_h), \qquad (5.44)
\end{aligned}$$

where $J_1^m(I_hT_*^m\phi_h)$ can be estimated by using the norm equivalence relations in (4.1) and the super-approximation estimates in Lemma 4.4, i.e.,

$$\begin{aligned}
|J_1^m(I_hT_*^m\phi_h)|&=\Big|\int_0^1\int_{\hat{\Gamma}_{h,\theta}^m}\frac{e_h^{m+1}-\hat{e}_h^m}{\tau}\cdot T_*^mI_hT_*^m\phi_h(\nabla_{\hat{\Gamma}_{h,\theta}^m}\cdot\hat{e}_h^m)\mathrm{d}\theta\\
&\quad+\int_0^1\int_{\hat{\Gamma}_{h,\theta}^m}\frac{e_h^{m+1}-\hat{e}_h^m}{\tau}\cdot(I_hT_*^m-T_*^m)I_hT_*^m\phi_h(\nabla_{\hat{\Gamma}_{h,\theta}^m}\cdot\hat{e}_h^m)\mathrm{d}\theta\Big|\\
&\lesssim\Big\|T_*^m\frac{e_h^{m+1}-\hat{e}_h^m}{\tau}\Big\|_{L^2(\hat{\Gamma}_{h,*}^m)}\|I_hT_*^m\phi_h\|_{L^2(\hat{\Gamma}_{h,*}^m)}\|\nabla_{\hat{\Gamma}_{h,*}^m}\hat{e}_h^m\|_{L^\infty(\hat{\Gamma}_{h,*}^m)}\\
&\quad+h\Big\|\frac{e_h^{m+1}-\hat{e}_h^m}{\tau}\Big\|_{L^2(\hat{\Gamma}_{h,*}^m)}\|I_hT_*^m\phi_h\|_{L^2(\hat{\Gamma}_{h,*}^m)}\|\nabla_{\hat{\Gamma}_{h,*}^m}\hat{e}_h^m\|_{L^\infty(\hat{\Gamma}_{h,*}^m)}.
\end{aligned}$$

Then, using the super-approximation estimates again, we can replace $\|T_*^m\phi_h\|_{L^2(\hat\Gamma_{h,\theta}^m)}$ by $\|I_hT_*^m\phi_h\|_{L^2(\hat\Gamma_{h,\theta}^m)}$ plus a higher-order smaller term, i.e.,

$$\begin{aligned}\|T_*^m\phi_h\|_{L^2(\hat\Gamma_{h,*}^m)} &\le \|I_hT_*^m\phi_h\|_{L^2(\hat\Gamma_{h,*}^m)} + \|(1-I_h)T_*^m\phi_h\|_{L^2(\hat\Gamma_{h,*}^m)} \\ &\lesssim \|I_hT_*^m\phi_h\|_{L^2(\hat\Gamma_{h,*}^m)} + h\|\phi_h\|_{L^2(\hat\Gamma_{h,*}^m)} \quad \forall\phi_h \in S_h(\hat\Gamma_{h,*}^m).\end{aligned}$$

Thus, substituting this inequality into the estimate of $J_1^m(I_hT_*^m\phi_h)$ with $\phi_h = (e_h^{m+1} - \hat e_h^m)/\tau$, we obtain

$$\begin{aligned}\Big|J_1^m\Big(I_hT_*^m\frac{e_h^{m+1}-\hat e_h^m}{\tau}\Big)\Big| &\lesssim \Big\|I_hT_*^m\frac{e_h^{m+1}-\hat e_h^m}{\tau}\Big\|_{L^2(\hat\Gamma_{h,*}^m)}^2 \|\nabla_{\hat\Gamma_{h,*}^m}\hat e_h^m\|_{L^\infty(\hat\Gamma_{h,*}^m)} \\ &\quad + h\Big\|\frac{e_h^{m+1}-\hat e_h^m}{\tau}\Big\|_{L^2(\hat\Gamma_{h,*}^m)} \\ &\quad \Big\|I_hT_*^m\frac{e_h^{m+1}-\hat e_h^m}{\tau}\Big\|_{L^2(\hat\Gamma_{h,*}^m)} \|\nabla_{\hat\Gamma_{h,*}^m}\hat e_h^m\|_{L^\infty(\hat\Gamma_{h,*}^m)}.\end{aligned}$$

The second term on the right-hand side of (5.44) can be estimated in the usual way, by using the Cauchy–Schwartz inequality. This leads to the following result:

$$\begin{aligned}\Big|J^m\Big(I_hT_*^m\frac{e_h^{m+1}-\hat e_h^m}{\tau}\Big)\Big| &\lesssim \Big\|I_hT_*^m\frac{e_h^{m+1}-\hat e_h^m}{\tau}\Big\|_{L^2(\hat\Gamma_{h,*}^m)}^2 \|\nabla_{\hat\Gamma_{h,*}^m}\hat e_h^m\|_{L^\infty(\hat\Gamma_{h,*}^m)} \\ &\quad + h\Big\|\frac{e_h^{m+1}-\hat e_h^m}{\tau}\Big\|_{L^2(\hat\Gamma_{h,*}^m)} \Big\|I_hT_*^m\frac{e_h^{m+1}-\hat e_h^m}{\tau}\Big\|_{L^2(\hat\Gamma_{h,*}^m)} \\ &\quad \|\nabla_{\hat\Gamma_{h,*}^m}\hat e_h^m\|_{L^\infty(\hat\Gamma_{h,*}^m)} \\ &\quad + \Big\|I_hT_*^m\frac{e_h^{m+1}-\hat e_h^m}{\tau}\Big\|_{L^2(\hat\Gamma_{h,*}^m)} \|\nabla_{\hat\Gamma_{h,*}^m}\hat e_h^m\|_{L^2(\hat\Gamma_{h,*}^m)}. \qquad (5.45)\end{aligned}$$

Then, under the stepsize condition $\tau = o(h^{2.5})$ and the induction assumption which implies that $\|\nabla_{\hat\Gamma_{h,*}^m}\hat e_h^m\|_{L^\infty(\hat\Gamma_{h,*}^m)} \lesssim h^{0.5}$, inequality (5.23) reduces to

$$\begin{aligned}&\Big|K^m\Big(I_hT_*^m\frac{e_h^{m+1}-\hat e_h^m}{\tau}\Big)\Big| \\ &\lesssim h^{-2}\|\nabla_{\hat\Gamma_{h,*}^m}\hat e_h^m\|_{L^2(\hat\Gamma_{h,*}^m)}^2 \Big\|I_hT_*^m\frac{e_h^{m+1}-\hat e_h^m}{\tau}\Big\|_{L^2(\hat\Gamma_{h,*}^m)} \\ &\quad + h\Big\|\frac{e_h^{m+1}-\hat e_h^m}{\tau}\Big\|_{L^2(\hat\Gamma_{h,*}^m)} \Big\|I_hT_*^m\frac{e_h^{m+1}-\hat e_h^m}{\tau}\Big\|_{L^2(\hat\Gamma_{h,*}^m)} \\ &\quad + h^{-1}(\tau + (1+\kappa_{*,l})h^{k-1})\|\nabla_{\hat\Gamma_{h,*}^m}\hat e_h^m\|_{L^2(\hat\Gamma_{h,*}^m)} \Big\|I_hT_*^m\frac{e_h^{m+1}-\hat e_h^m}{\tau}\Big\|_{L^2(\hat\Gamma_{h,*}^m)}. \qquad (5.46)\end{aligned}$$

Finally, the last term on the right-hand side of (5.40) has been estimated in Lemma 4.5, i.e.,

$$\left|d^m\Big(I_h T_*^m \frac{e_h^{m+1}-\hat e_h^m}{\tau}\Big)\right| \lesssim (\tau + (1+\kappa_{*,l})h^k)\left\|I_h T_*^m \frac{e_h^{m+1}-\hat e_h^m}{\tau}\right\|_{L^2(\hat\Gamma^m_{h,*})}. \tag{5.47}$$

Substituting estimates (5.41)–(5.47) into (5.40), we obtain the following result:

$$\begin{aligned}
&\int_{\hat\Gamma^m_{h,*}} \frac{e_h^{m+1}-\hat e_h^m}{\tau}\cdot I_h T_*^m \frac{e_h^{m+1}-\hat e_h^m}{\tau}\\
&\lesssim -A^N_{h,*}\Big(e_h^{m+1}, T_*^m I_h T_*^m \frac{e_h^{m+1}-\hat e_h^m}{\tau}\Big) - A^T_{h,*}\Big(e_h^{m+1}-\hat e_h^m, T_*^m I_h T_*^m \frac{e_h^{m+1}-\hat e_h^m}{\tau}\Big)\\
&\quad + \big(\|\nabla_{\hat\Gamma^m_{h,*}} \hat e_h^m\|_{L^2(\hat\Gamma^m_{h,*})} + \|\nabla_{\hat\Gamma^m_{h,*}} e_h^{m+1}\|_{L^2(\hat\Gamma^m_{h,*})}\big)\left\|I_h T_*^m \frac{e_h^{m+1}-\hat e_h^m}{\tau}\right\|_{L^2(\hat\Gamma^m_{h,*})}\\
&\quad + h^{0.5}\left\|I_h T_*^m \frac{e_h^{m+1}-\hat e_h^m}{\tau}\right\|^2_{L^2(\hat\Gamma^m_{h,*})} + h\left\|\frac{e_h^{m+1}-\hat e_h^m}{\tau}\right\|_{L^2(\hat\Gamma^m_{h,*})}\left\|I_h T_*^m \frac{e_h^{m+1}-\hat e_h^m}{\tau}\right\|_{L^2(\hat\Gamma^m_{h,*})}\\
&\quad + h^{-2}\|\nabla_{\hat\Gamma^m_{h,*}} \hat e_h^m\|^2_{L^2(\hat\Gamma^m_{h,*})}\left\|I_h T_*^m \frac{e_h^{m+1}-\hat e_h^m}{\tau}\right\|_{L^2(\hat\Gamma^m_{h,*})}\\
&\quad + (\tau + (1+\kappa_{*,l})h^k)\left\|I_h T_*^m \frac{e_h^{m+1}-\hat e_h^m}{\tau}\right\|_{L^2(\hat\Gamma^m_{h,*})}.
\end{aligned} \tag{5.48}$$

The first two terms on the right hand side of (5.48) have better upper bounds than their straightforward estimates due to the orthogonality relation. For example, the first term can be estimated by

$$\begin{aligned}
&-A^N_{h,*}\left(e_h^{m+1}, T_*^m I_h T_*^m \frac{e_h^{m+1}-\hat e_h^m}{\tau}\right)\\
&\quad = -\int_{\hat\Gamma^m_{h,*}} \big[(\nabla_{\hat\Gamma^m_{h,*}} e_h^{m+1}) n^m_{h,*}\big]\cdot\Big[\Big(\nabla_{\hat\Gamma^m_{h,*}} T_*^m I_h T_*^m \frac{e_h^{m+1}-\hat e_h^m}{\tau}\Big)n^m_{h,*}\Big]\\
&\quad = -\int_{\hat\Gamma^m_{h,*}} \big[(\nabla_{\hat\Gamma^m_{h,*}} e_h^{m+1}) n^m_{h,*}\big]\cdot\nabla_{\hat\Gamma^m_{h,*}}\Big([I - n_*^m (n_*^m)^\top] I_h T_*^m \frac{e_h^{m+1}-\hat e_h^m}{\tau}\cdot n^m_{h,*}\Big)\\
&\qquad + \int_{\hat\Gamma^m_{h,*}} \big[(\nabla_{\hat\Gamma^m_{h,*}} e_h^{m+1}) n^m_{h,*}\big]\cdot\Big[(\nabla_{\hat\Gamma^m_{h,*}} n^m_{h,*})[I - n_*^m (n_*^m)^\top] I_h T_*^m \frac{e_h^{m+1}-\hat e_h^m}{\tau}\Big]\\
&\quad \lesssim \|\nabla_{\hat\Gamma^m_{h,*}} e_h^{m+1}\|_{L^2(\hat\Gamma^m_{h,*})}\left\|I_h T_*^m \frac{e_h^{m+1}-\hat e_h^m}{\tau}\right\|_{L^2(\hat\Gamma^m_{h,*})}\\
&\qquad + \|\nabla_{\hat\Gamma^m_{h,*}} e_h^{m+1}\|_{L^2(\hat\Gamma^m_{h,*})}\|n^m_{h,*} - n^m_*\|_{L^\infty(\hat\Gamma^m_{h,*})}\left\|\nabla_{\hat\Gamma^m_{h,*}} I_h T_*^m \frac{e_h^{m+1}-\hat e_h^m}{\tau}\right\|_{L^2(\hat\Gamma^m_{h,*})}\\
&\quad \lesssim \|\nabla_{\hat\Gamma^m_{h,*}} e_h^{m+1}\|_{L^2(\hat\Gamma^m_{h,*})}\left\|I_h T_*^m \frac{e_h^{m+1}-\hat e_h^m}{\tau}\right\|_{L^2(\hat\Gamma^m_{h,*})}\\
&\quad \text{((3.9) and inverse inequality are used).}
\end{aligned} \tag{5.49}$$

Analogously, the second term on the right hand side of (5.48) can be estimated by using the definition of $A^T_{h,*}(\cdot,\cdot)$ in (5.7) and (5.11), as well as the decomposition $\hat{T}^m_{h,*} = T^m_* + (\hat{T}^m_{h,*} - T^m_*) = T^m_* + O(h)$, i.e.,

$$
\begin{aligned}
&- A^T_{h,*}\left(e^{m+1}_h - \hat{e}^m_h, T^m_* I_h T^m_* \frac{e^{m+1}_h - \hat{e}^m_h}{\tau}\right)\\
&= - \int_{\hat{\Gamma}^m_{h,*}} \mathrm{tr}\Big[(\nabla_{\hat{\Gamma}^m_{h,*}}(e^{m+1}_h - \hat{e}^m_h))\hat{T}^m_{h,*}\Big(\nabla_{\hat{\Gamma}^m_{h,*}} T^m_* I_h T^m_* \frac{e^{m+1}_h - \hat{e}^m_h}{\tau}\Big)^\top\Big]\\
&\lesssim - \int_{\hat{\Gamma}^m_{h,*}} \mathrm{tr}\Big[(\nabla_{\hat{\Gamma}^m_{h,*}}(e^{m+1}_h - \hat{e}^m_h))T^m_*\Big(\nabla_{\hat{\Gamma}^m_{h,*}} T^m_* I_h T^m_* \frac{e^{m+1}_h - \hat{e}^m_h}{\tau}\Big)^\top\Big]\\
&\quad + h^{0.5}\|\nabla_{\hat{\Gamma}^m_{h,*}}(e^{m+1}_h - \hat{e}^m_h)\|_{L^2(\hat{\Gamma}^m_{h,*})}\Big\| I_h T^m_* \frac{e^{m+1}_h - \hat{e}^m_h}{\tau}\Big\|_{L^2(\hat{\Gamma}^m_{h,*})}\\
&\text{((3.9) and inverse inequality are used)}\\
&\lesssim - \int_{\hat{\Gamma}^m_{h,*}} \mathrm{tr}\Big[(\nabla_{\hat{\Gamma}^m_{h,*}}(e^{m+1}_h - \hat{e}^m_h))T^m_*\Big(\nabla_{\hat{\Gamma}^m_{h,*}} I_h T^m_* \frac{e^{m+1}_h - \hat{e}^m_h}{\tau}\Big)^\top\Big]\\
&\quad + \|\nabla_{\hat{\Gamma}^m_{h,*}}(e^{m+1}_h - \hat{e}^m_h)\|_{L^2(\hat{\Gamma}^m_{h,*})}\Big\| I_h T^m_* \frac{e^{m+1}_h - \hat{e}^m_h}{\tau}\Big\|_{L^2(\hat{\Gamma}^m_{h,*})} \quad \text{(Leibniz rule)}\\
&\lesssim - \int_{\hat{\Gamma}^m_{h,*}} \nabla_{\hat{\Gamma}^m_{h,*}}[T^m_*(e^{m+1}_h - \hat{e}^m_h)] \cdot \nabla_{\hat{\Gamma}^m_{h,*}} I_h T^m_* \frac{e^{m+1}_h - \hat{e}^m_h}{\tau}\\
&\quad + h^{-1}\|e^{m+1}_h - \hat{e}^m_h\|_{L^2(\hat{\Gamma}^m_{h,*})}\Big\| I_h T^m_* \frac{e^{m+1}_h - \hat{e}^m_h}{\tau}\Big\|_{L^2(\hat{\Gamma}^m_{h,*})}\\
&\text{(Leibniz rule and inverse inequality)}\\
&\lesssim -\tau \int_{\hat{\Gamma}^m_{h,*}} \nabla_{\hat{\Gamma}^m_{h,*}} I_h T^m_* \frac{e^{m+1}_h - \hat{e}^m_h}{\tau} \cdot \nabla_{\hat{\Gamma}^m_{h,*}} I_h T^m_* \frac{e^{m+1}_h - \hat{e}^m_h}{\tau}\\
&\quad + h^{-1}\|e^{m+1}_h - \hat{e}^m_h\|_{L^2(\hat{\Gamma}^m_{h,*})}\Big\| I_h T^m_* \frac{e^{m+1}_h - \hat{e}^m_h}{\tau}\Big\|_{L^2(\hat{\Gamma}^m_{h,*})} \text{(Lemma 4.4 is used)}\\
&\lesssim \tau h^{-1}\Big\| \frac{e^{m+1}_h - \hat{e}^m_h}{\tau}\Big\|_{L^2(\hat{\Gamma}^m_{h,*})}\Big\| I_h T^m_* \frac{e^{m+1}_h - \hat{e}^m_h}{\tau}\Big\|_{L^2(\hat{\Gamma}^m_{h,*})}.
\end{aligned}
\tag{5.50}
$$

By utilizing the estimates in (5.48)–(5.50) and the stepsize condition $\tau = o(h^{2.5})$, we can estimate the tangential component of the numerical velocity as follows:

$$
\begin{aligned}
&\int_{\hat{\Gamma}^m_{h,*}} I_h T^m_* \frac{e^{m+1}_h - \hat{e}^m_h}{\tau} \cdot I_h T^m_* \frac{e^{m+1}_h - \hat{e}^m_h}{\tau}\\
&\quad = \int_{\hat{\Gamma}^m_{h,*}} \frac{e^{m+1}_h - \hat{e}^m_h}{\tau} \cdot I_h T^m_* \frac{e^{m+1}_h - \hat{e}^m_h}{\tau}
\end{aligned}
$$

$$
\begin{aligned}
&+\int_{\hat\Gamma^m_{h,*}}(1-I_h)N^m_*\frac{e^{m+1}_h-\hat e^m_h}{\tau}\cdot I_hT^m_*\frac{e^{m+1}_h-\hat e^m_h}{\tau}\\
&-\int_{\hat\Gamma^m_{h,*}}N^m_*\frac{e^{m+1}_h-\hat e^m_h}{\tau}\cdot I_hT^m_*I_hT^m_*\frac{e^{m+1}_h-\hat e^m_h}{\tau}\\
=&\int_{\hat\Gamma^m_{h,*}}\frac{e^{m+1}_h-\hat e^m_h}{\tau}\cdot I_hT^m_*\frac{e^{m+1}_h-\hat e^m_h}{\tau}\\
&+\int_{\hat\Gamma^m_{h,*}}(1-I_h)N^m_*\frac{e^{m+1}_h-\hat e^m_h}{\tau}\cdot I_hT^m_*\frac{e^{m+1}_h-\hat e^m_h}{\tau}\\
&+\int_{\hat\Gamma^m_{h,*}}\frac{e^{m+1}_h-\hat e^m_h}{\tau}\cdot N^m_*(1-I_h)T^m_*I_hT^m_*\frac{e^{m+1}_h-\hat e^m_h}{\tau}\\
\lesssim&\left(\|\nabla_{\hat\Gamma^m_{h,*}}\hat e^m_h\|_{L^2(\hat\Gamma^m_{h,*})}+\|\nabla_{\hat\Gamma^m_{h,*}}e^{m+1}_h\|_{L^2(\hat\Gamma^m_{h,*})}\right)\left\|I_hT^m_*\frac{e^{m+1}_h-\hat e^m_h}{\tau}\right\|_{L^2(\hat\Gamma^m_{h,*})}\\
&+h^{0.5}\left\|I_hT^m_*\frac{e^{m+1}_h-\hat e^m_h}{\tau}\right\|^2_{L^2(\hat\Gamma^m_{h,*})}\\
&+h\left\|\frac{e^{m+1}_h-\hat e^m_h}{\tau}\right\|_{L^2(\hat\Gamma^m_{h,*})}\left\|I_hT^m_*\frac{e^{m+1}_h-\hat e^m_h}{\tau}\right\|_{L^2(\hat\Gamma^m_{h,*})}\\
&+h^{-2}\|\nabla_{\hat\Gamma^m_{h,*}}\hat e^m_h\|^2_{L^2(\hat\Gamma^m_{h,*})}\left\|I_hT^m_*\frac{e^{m+1}_h-\hat e^m_h}{\tau}\right\|_{L^2(\hat\Gamma^m_{h,*})}\\
&+(\tau+(1+\kappa_{*,l})h^k)\left\|I_hT^m_*\frac{e^{m+1}_h-\hat e^m_h}{\tau}\right\|_{L^2(\hat\Gamma^m_{h,*})}. \qquad (5.51)
\end{aligned}
$$

Substituting (5.28) and the inequality

$$
\left\|T^m_*\frac{e^{m+1}_h-\hat e^m_h}{\tau}\right\|_{L^2(\hat\Gamma^m_{h,*})}\lesssim\left\|I_hT^m_*\frac{e^{m+1}_h-\hat e^m_h}{\tau}\right\|_{L^2(\hat\Gamma^m_{h,*})}+h\left\|\frac{e^{m+1}_h-\hat e^m_h}{\tau}\right\|_{L^2(\hat\Gamma^m_{h,*})}
$$

into (5.51) and using (5.28) to eliminate $h\|\frac{e^{m+1}_h-\hat e^m_h}{\tau}\|_{L^2(\hat\Gamma^m_{h,*})}$, we obtain

$$
\begin{aligned}
\left\|T^m_*\frac{e^{m+1}_h-\hat e^m_h}{\tau}\right\|_{L^2(\hat\Gamma^m_{h,*})}\lesssim&\|\nabla_{\hat\Gamma^m_{h,*}}\hat e^m_h\|_{L^2(\hat\Gamma^m_{h,*})}+\|\nabla_{\hat\Gamma^m_{h,*}}e^{m+1}_h\|_{L^2(\hat\Gamma^m_{h,*})}\\
&+h^{-2}\|\nabla_{\hat\Gamma^m_{h,*}}\hat e^m_h\|^2_{L^2(\hat\Gamma^m_{h,*})}\\
&+\tau+(1+\kappa_{*,l})h^k.
\end{aligned}
$$

By decomposing e_h^{m+1} into $e_h^{m+1} - \hat{e}_h^m$ and $\hat{e}_h^m$, and using (5.28) again with the stepsize condition $\tau = o(h^{2.5})$, we obtain the following result:

$$\left\| T_*^m \frac{e_h^{m+1} - \hat{e}_h^m}{\tau} \right\|_{L^2(\hat{\Gamma}_{h,*}^m)} \lesssim \|\nabla_{\hat{\Gamma}_{h,*}^m} \hat{e}_h^m\|_{L^2(\hat{\Gamma}_{h,*}^m)} + h^{-2}\|\nabla_{\hat{\Gamma}_{h,*}^m} \hat{e}_h^m\|_{L^2(\hat{\Gamma}_{h,*}^m)}^2 + \tau + (1+\kappa_{*,l})h^k. \quad (5.52)$$

5.7 Stability of Orthogonal Projection on the Error

Since the value of $\hat{e}_h^{m+1}$ at the nodes is the distance between the numerically computed surface and the exact surface, it follows that $\hat{e}_h^{m+1}$ can be approximately viewed as the orthogonal projection of e_h^{m+1} to the normal direction. In order to apply Grönwall's inequality, we need to estimate $\|\hat{e}_h^{m+1}\|_{L^2(\hat{\Gamma}_{h,*}^{m+1})}^2 - \|e_h^{m+1}\|_{L^2(\hat{\Gamma}_{h,*}^m)}^2$, which accounts for the stability of the projection on the error and stability of changing the underlying surface. To this end, we denote by

$$\begin{aligned} T_*^m &= I - n_*^m (n_*^m)^\top \quad \text{and} \quad N_*^m = n_*^m (n_*^m)^\top \\ T_*^{m+1} &= I - n_*^{m+1} (n_*^{m+1})^\top \quad \text{and} \quad N_*^{m+1} = n_*^{m+1} (n_*^{m+1})^\top \end{aligned}$$

the tangential projection matrices and normal projection matrices, all pulled back to the surface $\hat{\Gamma}_{h,*}^m$, where n_*^{m+1} actually denotes the pulled-back function $n_*^{m+1} \circ \hat{X}_{h,*}^{m+1}$ on $\hat{\Gamma}_{h,*}^m$ (with abbreviation). Then consider the following decomposition:

$$\begin{aligned} &\|\hat{e}_h^{m+1}\|_{L^2(\hat{\Gamma}_{h,*}^{m+1})}^2 - \|e_h^{m+1}\|_{L^2(\hat{\Gamma}_{h,*}^m)}^2 \\ &= \|\hat{e}_h^{m+1}\|_{L^2(\hat{\Gamma}_{h,*}^{m+1})}^2 - \|\hat{e}_h^{m+1}\|_{L^2(\hat{\Gamma}_{h,*}^m)}^2 \\ &\quad + \|\hat{e}_h^{m+1}\|_{L^2(\hat{\Gamma}_{h,*}^m)}^2 - \|e_h^{m+1}\|_{L^2(\hat{\Gamma}_{h,*}^m)}^2 \\ &= \|\hat{e}_h^{m+1}\|_{L^2(\hat{\Gamma}_{h,*}^{m+1})}^2 - \|\hat{e}_h^{m+1}\|_{L^2(\hat{\Gamma}_{h,*}^m)}^2 \\ &\quad + \|\hat{e}_h^{m+1}\|_{L^2(\hat{\Gamma}_{h,*}^m)}^2 - \|I_h N_*^{m+1} e_h^{m+1}\|_{L^2(\hat{\Gamma}_{h,*}^m)}^2 - \|I_h T_*^m e_h^{m+1}\|_{L^2(\hat{\Gamma}_{h,*}^m)}^2 \\ &\quad + \|I_h N_*^{m+1} e_h^{m+1}\|_{L^2(\hat{\Gamma}_{h,*}^m)}^2 - \|I_h N_*^m e_h^{m+1}\|_{L^2(\hat{\Gamma}_{h,*}^m)}^2 \\ &\quad - 2\int_{\hat{\Gamma}_{h,*}^m} I_h N_*^m e_h^{m+1} \cdot I_h T_*^m e_h^{m+1} \\ &=: L_1 + L_2 + L_3 + L_4. \end{aligned} \quad (5.53)$$

We denote by $\hat{\Gamma}_{h,*}^{m+\theta} = (1-\theta)\hat{\Gamma}_{h,*}^m + \theta\hat{\Gamma}_{h,*}^{m+1}$ and consider the following decomposition of L_1:

$$
\begin{aligned}
L_1 &= \|\hat{e}_h^{m+1}\|^2_{L^2(\hat{\Gamma}^{m+1}_{h,*})} - \|\hat{e}_h^{m+1}\|^2_{L^2(\hat{\Gamma}^{m}_{h,*})} \\
&= \int_0^1 \frac{\mathrm{d}}{\mathrm{d}\theta} \int_{\hat{\Gamma}^{m+\theta}_{h,*}} \hat{e}_h^{m+1} \cdot \hat{e}_h^{m+1} \mathrm{d}\theta \\
&= \int_0^1 \int_{\hat{\Gamma}^{m+\theta}_{h,*}} \hat{e}_h^{m+1} \cdot \hat{e}_h^{m+1} \nabla_{\hat{\Gamma}^{m+\theta}_{h,*}} \cdot (\hat{X}^{m+1}_{h,*} - \hat{X}^{m}_{h,*}) \mathrm{d}\theta \\
&= \int_{\Gamma^m} \hat{e}_h^{m+1,l} \cdot \hat{e}_h^{m+1,l} [\nabla_{\Gamma^m} \cdot (\hat{X}^{m+1}_{h,*} - \hat{X}^{m}_{h,*})^l] \mathrm{d}\theta \\
&\quad + \Big[\int_{\hat{\Gamma}^{m}_{h,*}} \hat{e}_h^{m+1} \cdot \hat{e}_h^{m+1} \nabla_{\hat{\Gamma}^{m}_{h,*}} \cdot (\hat{X}^{m+1}_{h,*} - \hat{X}^{m}_{h,*}) \\
&\qquad - \int_{\Gamma^m} \hat{e}_h^{m+1,l} \cdot \hat{e}_h^{m+1,l} [\nabla_{\Gamma^m} \cdot (\hat{X}^{m+1}_{h,*} - \hat{X}^{m}_{h,*})^l] \Big] \\
&\quad + \int_0^1 \Big[\int_{\hat{\Gamma}^{m+\theta}_{h,*}} \hat{e}_h^{m+1} \cdot \hat{e}_h^{m+1} \nabla_{\hat{\Gamma}^{m+\theta}_{h,*}} \cdot (\hat{X}^{m+1}_{h,*} - \hat{X}^{m}_{h,*}) \\
&\qquad - \int_{\hat{\Gamma}^{m}_{h,*}} \hat{e}_h^{m+1} \cdot \hat{e}_h^{m+1} \nabla_{\hat{\Gamma}^{m}_{h,*}} \cdot (\hat{X}^{m+1}_{h,*} - \hat{X}^{m}_{h,*}) \Big] \mathrm{d}\theta \\
&=: L_{11} + L_{12} + L_{13}.
\end{aligned} \tag{5.54}
$$

The first term can be estimated by using integration by parts (Lemma 5.1, (item 3)), i.e.,

$$
\begin{aligned}
L_{11} &= \int_{\Gamma^m} \hat{e}_h^{m+1,l} \cdot \hat{e}_h^{m+1,l} [\nabla_{\Gamma^m} \cdot (\hat{X}^{m+1}_{h,*} - \hat{X}^{m}_{h,*})^l] \mathrm{d}\theta \\
&= -2 \int_{\Gamma^m} (\nabla_{\Gamma^m} \hat{e}_h^{m+1,l} \cdot \hat{e}_h^{m+1,l}) \cdot (\hat{X}^{m+1}_{h,*} - \hat{X}^{m}_{h,*})^l \mathrm{d}\theta \\
&\quad + \int_{\Gamma^m} (\hat{e}_h^{m+1,l} \cdot \hat{e}_h^{m+1,l}) H^m n^m \cdot (\hat{X}^{m+1}_{h,*} - \hat{X}^{m}_{h,*})^l \mathrm{d}\theta \\
&\lesssim \|\hat{X}^{m+1}_{h,*} - \hat{X}^{m}_{h,*}\|_{L^3(\hat{\Gamma}^m_{h,*})} \|\hat{e}_h^{m+1}\|_{L^6(\hat{\Gamma}^m_{h,*})} \|\hat{e}_h^{m+1}\|_{H^1(\hat{\Gamma}^m_{h,*})}.
\end{aligned} \tag{5.55}
$$

We express $\hat{X}^{m+1}_{h,*} - \hat{X}^{m}_{h,*}$ into the following form:

$$
\begin{aligned}
\hat{X}^{m+1}_{h,*} - \hat{X}^{m}_{h,*} &= X^{m+1}_{h,*} - \hat{X}^{m}_{h,*} + \hat{X}^{m+1}_{h,*} - X^{m+1}_{h,*} \\
&= (X^{m+1}_{h,*} - \hat{X}^{m}_{h,*}) + (e_h^{m+1} - \hat{e}_h^{m+1}),
\end{aligned} \tag{5.56}
$$

where

$$
\|X^{m+1}_{h,*} - \hat{X}^{m}_{h,*}\|_{L^3(\hat{\Gamma}^m_{h,*})} = \|I_h(X^{m+1} - \mathrm{id})\|_{L^3(\hat{\Gamma}^m_{h,*})} \lesssim \tau. \tag{5.57}
$$

By using relation $\hat{e}_h^{m+1} = I_h\big[(e_h^{m+1} \cdot n_*^{m+1}) n_*^{m+1}\big] + f_h$ in (3.12), we have

$$
\begin{aligned}
e_h^{m+1} - \hat{e}_h^{m+1} &= I_h T_*^{m+1} e_h^{m+1} - f_h \\
&= I_h T_*^m (e_h^{m+1} - \hat{e}_h^m) - f_h + I_h (T_*^{m+1} - T_*^m) e_h^{m+1},
\end{aligned}
$$

where we have used the orthogonality $I_h T_*^m \hat{e}_h^m = 0$. Since the continuous L^p norm and the discrete L^p norm at nodes are equivalent for $p \in [1, \infty]$, it follows that

$$
\begin{aligned}
\|e_h^{m+1} - \hat{e}_h^{m+1}\|_{L^3(\hat{\Gamma}_{h,*}^m)} &\lesssim \|e_h^{m+1} - \hat{e}_h^m\|_{L^3(\hat{\Gamma}_{h,*}^m)} + \tau \|e_h^{m+1}\|_{L^3(\hat{\Gamma}_{h,*}^m)} + \|f_h\|_{L^3(\hat{\Gamma}_{h,*}^m)} \\
&\lesssim \|e_h^{m+1} - \hat{e}_h^m\|_{L^3(\hat{\Gamma}_{h,*}^m)} + \tau \|e_h^{m+1}\|_{L^3(\hat{\Gamma}_{h,*}^m)} \\
&\quad + \|I_h T_*^{m+1} e_h^{m+1}\|_{L^\infty(\hat{\Gamma}_{h,*}^m)} \|I_h T_*^{m+1} e_h^{m+1}\|_{L^3(\hat{\Gamma}_{h,*}^m)}
\end{aligned}
$$

where we have used the estimate of $|f_h| \lesssim |I_h T_*^{m+1} e_h^{m+1}|^2$ at nodes in (3.13). Since $I_h T_*^m \hat{e}_h^m = 0$, it follows that

$$
\begin{aligned}
\|I_h T_*^{m+1} e_h^{m+1}\|_{L^p(\hat{\Gamma}_{h,*}^m)} &= \|I_h T_*^m (e_h^{m+1} - \hat{e}_h^m) + I_h (T_*^{m+1} - T_*^m) e_h^{m+1}\|_{L^p(\hat{\Gamma}_{h,*}^m)} \\
&\lesssim \|I_h T_*^m (e_h^{m+1} - \hat{e}_h^m)\|_{L^p(\hat{\Gamma}_{h,*}^m)} + \tau \|e_h^{m+1}\|_{L^p(\hat{\Gamma}_{h,*}^m)}
\end{aligned}
$$

Then, using the relations $\|e_h^{m+1}\|_{L^3(\hat{\Gamma}_{h,*}^m)} \le \|e_h^{m+1} - \hat{e}_h^m\|_{L^3(\hat{\Gamma}_{h,*}^m)} + \|\hat{e}_h^m\|_{L^3(\hat{\Gamma}_{h,*}^m)}$, we have

$$
\|e_h^{m+1} - \hat{e}_h^{m+1}\|_{L^3(\hat{\Gamma}_{h,*}^m)} \lesssim \|e_h^{m+1} - \hat{e}_h^m\|_{L^3(\hat{\Gamma}_{h,*}^m)} + \tau \|\hat{e}_h^m\|_{L^3(\hat{\Gamma}_{h,*}^m)}. \tag{5.58}
$$

By substituting (5.57) and (5.58) into (5.56), we obtain

$$
\begin{aligned}
\|\hat{X}_{h,*}^{m+1} - \hat{X}_{h,*}^m\|_{L^3(\hat{\Gamma}_{h,*}^m)} &\lesssim \tau + \|e_h^{m+1} - \hat{e}_h^m\|_{L^3(\hat{\Gamma}_{h,*}^m)} + \tau \|\hat{e}_h^m\|_{L^3(\hat{\Gamma}_{h,*}^m)} \\
&\lesssim \tau + \tau h^{-\frac{4}{3}} \|\nabla_{\hat{\Gamma}_{h,*}^m} \hat{e}_h^m\|_{L^2(\hat{\Gamma}_{h,*}^m)} \\
&\quad + \tau h^{-\frac{1}{3}} [\tau + (1 + \kappa_{*,l}) h^{k-1}] + \tau h^{-\frac{1}{3}} \|\hat{e}_h^m\|_{L^2(\hat{\Gamma}_{h,*}^m)} \\
&\lesssim \tau,
\end{aligned} \tag{5.59}
$$

where we have used (5.28) and the inverse inequality in the derivation of the second to last inequality, and have used (3.10) as well as condition (3.8) in the derivation of the last inequality. Therefore, substituting this into (5.55), we have

$$
|L_{11}| \lesssim \tau \|\hat{e}_h^{m+1}\|_{L^6(\hat{\Gamma}_{h,*}^m)} \|\hat{e}_h^{m+1}\|_{H^1(\hat{\Gamma}_{h,*}^m)} \lesssim \epsilon \tau \|\hat{e}_h^{m+1}\|_{H^1(\hat{\Gamma}_{h,*}^m)}^2 + \epsilon^{-1} \tau \|\hat{e}_h^{m+1}\|_{L^2(\hat{\Gamma}_{h,*}^m)}^2, \tag{5.60}
$$

where the last inequality uses the interpolation inequality on Γ^m (i.e., the L^6 norm is intermediate between the L^2 and H^1 norms) and the equivalence of L^p and $W^{1,p}$ norms on $\hat{\Gamma}^m_{h,*}$ and Γ^m (for $\hat{e}_h^{m+1}$ and its lift $(\hat{e}_h^{m+1})^l$).

The estimate of L_{12} comes from the geometric perturbation estimate (cf. [36, Lemma 5.6]), i.e.,

$$
\begin{aligned}
|L_{12}| &\lesssim \|\nabla_{\hat{\Gamma}^m_{h,*}}(a-\mathrm{id})\|_{L^2(\hat{\Gamma}^m_{h,*})}\|\nabla_{\hat{\Gamma}^m_{h,*}}(\hat{X}^{m+1}_{h,*}-\hat{X}^m_{h,*})\|_{L^3(\hat{\Gamma}^m_{h,*})}\|\hat{e}_h^{m+1}\|^2_{L^{12}(\hat{\Gamma}^m_{h,*})}\\
&\lesssim h^k\|\nabla_{\hat{\Gamma}^m_{h,*}}(\hat{X}^{m+1}_{h,*}-\hat{X}^m_{h,*})\|_{L^3(\hat{\Gamma}^m_{h,*})}\|\hat{e}_h^{m+1}\|^2_{H^1(\hat{\Gamma}^m_{h,*})}\\
&\lesssim h^{k-\frac{5}{6}}\tau\|\hat{e}_h^{m+1}\|^2_{H^1(\hat{\Gamma}^m_{h,*})} \quad \text{(inverse inequality and (5.59) are used).} \qquad (5.61)
\end{aligned}
$$

Analogously,

$$
\begin{aligned}
|L_{13}| &\lesssim \|\nabla_{\hat{\Gamma}^m_{h,*}}(\hat{X}^{m+1}_{h,*}-\hat{X}^m_{h,*})\|^2_{L^3(\hat{\Gamma}^m_{h,*})}\|\hat{e}_h^{m+1}\|^2_{L^6(\hat{\Gamma}^m_{h,*})}\\
&\lesssim h^{-\frac{5}{3}}\tau^2\|\hat{e}_h^{m+1}\|^2_{H^1(\hat{\Gamma}^m_{h,*})} \quad \text{(inverse inequality and (5.59) are used).} \qquad (5.62)
\end{aligned}
$$

In summary we obtain the following estimate for L_1:

$$
|L_1| \lesssim \epsilon\tau\|\hat{e}_h^{m+1}\|^2_{H^1(\hat{\Gamma}^m_{h,*})}+\epsilon^{-1}\tau\|\hat{e}_h^{m+1}\|^2_{L^2(\hat{\Gamma}^m_{h,*})}, \qquad (5.63)
$$

where we have used the stepsize condition $\tau = o(h^{2.5})$ and $k \geq 3$.

The term L_2 defined in (5.53) can be estimated by using the almost orthogonality relation in (3.12)–(3.13), i.e.,

$$
\begin{aligned}
L_2 &= \|\hat{e}_h^{m+1}\|^2_{L^2(\hat{\Gamma}^m_{h,*})}-\|I_hN^{m+1}_*e_h^{m+1}\|^2_{L^2(\hat{\Gamma}^m_{h,*})}-\|I_hT^m_*e_h^{m+1}\|^2_{L^2(\hat{\Gamma}^m_{h,*})}\\
&\leq \|\hat{e}_h^{m+1}\|^2_{L^2(\hat{\Gamma}^m_{h,*})}-\|I_hN^{m+1}_*e_h^{m+1}\|^2_{L^2(\hat{\Gamma}^m_{h,*})}\\
&\lesssim \|\hat{e}_h^{m+1}-I_h\big[(e_h^{m+1}\cdot n_*^{m+1})n_*^{m+1}\big]\|_{L^2(\hat{\Gamma}^m_{h,*})}(\|\hat{e}_h^{m+1}\|_{L^2(\hat{\Gamma}^m_{h,*})}+\|e_h^{m+1}\|_{L^2(\hat{\Gamma}^m_{h,*})})\\
&= \|f_h\|_{L^2(\hat{\Gamma}^m_{h,*})}(\|\hat{e}_h^{m+1}\|_{L^2(\hat{\Gamma}^m_{h,*})}+\|e_h^{m+1}\|_{L^2(\hat{\Gamma}^m_{h,*})}) \quad \text{(relation (3.12)is used)}\\
&\lesssim \|[I-n_*^{m+1}(n_*^{m+1})^\top]e_h^{m+1}\|^2_{L^\infty(\hat{\Gamma}^m_{h,*})}(\|\hat{e}_h^{m+1}\|_{L^2(\hat{\Gamma}^m_{h,*})}+\|e_h^{m+1}\|_{L^2(\hat{\Gamma}^m_{h,*})})\\
&\qquad \text{(relation (3.13) is used)}\\
&\lesssim [\tau h^{-2}\|\hat{e}_h^m\|_{H^1(\hat{\Gamma}^m_{h,*})}+\tau h^{-1}(\tau+(1+\kappa_{*,l})h^{k-1})]^2(\|\hat{e}_h^{m+1}\|_{L^2(\hat{\Gamma}^m_{h,*})}\\
&\quad +\|e_h^{m+1}\|_{L^2(\hat{\Gamma}^m_{h,*})})\\
&\qquad \text{(relation (5.32) is used)}\\
&\lesssim o(1)\tau\|\hat{e}_h^m\|_{H^1(\hat{\Gamma}^m_{h,*})}(\|\hat{e}_h^{m+1}\|_{L^2(\hat{\Gamma}^m_{h,*})}+\|e_h^{m+1}\|_{L^2(\hat{\Gamma}^m_{h,*})})\\
&\quad +\tau(\tau+(1+\kappa_{*,l})h^k)(\|\hat{e}_h^{m+1}\|_{L^2(\hat{\Gamma}^m_{h,*})}+\|e_h^{m+1}\|_{L^2(\hat{\Gamma}^m_{h,*})}), \qquad (5.64)
\end{aligned}
$$

where we have used the stepsize condition $\tau = o(h^{2.5})$ and the induction assumption in (3.10) which implies that $\|\hat{e}_h^m\|_{H^1(\hat{\Gamma}_{h,*}^m)} \lesssim h^{1.5}$.

The term L_3 defined in (5.53) can be estimated by using relation (5.31), i.e., $|n_*^{m+1} - n_*^m| \lesssim \tau + |T_*^m(e_h^{m+1} - \hat{e}_h^m)|$ at the nodes:

$$\begin{aligned} L_3 &= \|I_h N_*^{m+1} e_h^{m+1}\|^2_{L^2(\hat{\Gamma}_{h,*}^m)} - \|I_h N_*^m e_h^{m+1}\|^2_{L^2(\hat{\Gamma}_{h,*}^m)} \\ &\lesssim \|I_h([n_*^{m+1}(n_*^{m+1})^\top - n_*^m(n_*^m)^\top]e_h^{m+1})\|_{L^2(\hat{\Gamma}_{h,*}^m)} \\ &\quad \times \|I_h([n_*^{m+1}(n_*^{m+1})^\top + n_*^m(n_*^m)^\top]e_h^{m+1})\|_{L^2(\hat{\Gamma}_{h,*}^m)} \\ &\lesssim (\tau + \|I_h T_*^m(e_h^{m+1} - \hat{e}_h^m)\|_{L^\infty(\hat{\Gamma}_{h,*}^m)})\|e_h^{m+1}\|^2_{L^2(\hat{\Gamma}_{h,*}^m)} \\ &\lesssim \tau\|e_h^{m+1}\|^2_{L^2(\hat{\Gamma}_{h,*}^m)}, \end{aligned} \tag{5.65}$$

where we have used inequality (5.52) with induction assumption $\|\nabla_{\hat{\Gamma}_{h,*}^m} \hat{e}_h^m\|_{L^2(\hat{\Gamma}_{h,*}^m)} \lesssim h^{1.5}$, the inverse inequality and the mesh size condition in (3.8).

The term L_4 can be estimated by using the orthogonality relation $I_h(I_h N_*^m e_h^{m+1} \cdot I_h T_*^m e_h^{m+1}) = 0$ and the super-approximation estimates in Lemma 4.4, i.e.,

$$\begin{aligned} L_4 &= -2\int_{\hat{\Gamma}_{h,*}^m} I_h N_*^m e_h^{m+1} \cdot I_h T_*^m e_h^{m+1} \\ &= -2\int_{\hat{\Gamma}_{h,*}^m} \left[I_h N_*^m e_h^{m+1} \cdot I_h T_*^m e_h^{m+1} - I_h(I_h N_*^m e_h^{m+1} \cdot I_h T_*^m e_h^{m+1})\right] \\ &\lesssim h^2\|I_h N_*^m e_h^{m+1}\|_{H^1(\hat{\Gamma}_{h,*}^m)}\|I_h T_*^m e_h^{m+1}\|_{H^1(\hat{\Gamma}_{h,*}^m)} \qquad \text{(Lemma 4.4 is used)} \\ &\lesssim h\|I_h N_*^m e_h^{m+1}\|_{H^1(\hat{\Gamma}_{h,*}^m)}\|I_h T_*^m(e_h^{m+1} - \hat{e}_h^m)\|_{L^2(\hat{\Gamma}_{h,*}^m)} \\ &\qquad \text{(inverse inequality and } I_h T_*^m \hat{e}_h^m = 0 \text{ are used)} \\ &\lesssim h\|I_h N_*^m e_h^{m+1}\|_{H^1(\hat{\Gamma}_{h,*}^m)}\|T_*^m(e_h^{m+1} - \hat{e}_h^m)\|_{L^2(\hat{\Gamma}_{h,*}^m)} \\ &\quad + h^2\|I_h N_*^m e_h^{m+1}\|_{H^1(\hat{\Gamma}_{h,*}^m)}\|e_h^{m+1} - \hat{e}_h^m\|_{L^2(\hat{\Gamma}_{h,*}^m)} \qquad \text{(Lemma 4.4 is used)} \\ &\lesssim h\tau\|I_h N_*^m e_h^{m+1}\|_{H^1(\hat{\Gamma}_{h,*}^m)}\big(\|\nabla_{\hat{\Gamma}_{h,*}^m} \hat{e}_h^m\|_{L^2(\hat{\Gamma}_{h,*}^m)} \\ &\quad + h^{-2}\|\nabla_{\hat{\Gamma}_{h,*}^m} \hat{e}_h^m\|^2_{L^2(\hat{\Gamma}_{h,*}^m)} + \tau + (1+\kappa_{*,l})h^k\big) \\ &\qquad \text{((5.28) and (5.52) are used)} \\ &\lesssim h\tau\|N_*^m e_h^{m+1}\|_{H^1(\hat{\Gamma}_{h,*}^m)}\big(\|\nabla_{\hat{\Gamma}_{h,*}^m} \hat{e}_h^m\|_{L^2(\hat{\Gamma}_{h,*}^m)} \\ &\quad + h^{-2}\|\nabla_{\hat{\Gamma}_{h,*}^m} \hat{e}_h^m\|^2_{L^2(\hat{\Gamma}_{h,*}^m)} + \tau + (1+\kappa_{*,l})h^k\big) \\ &\quad + h\tau\|(1-I_h)N_*^m e_h^{m+1}\|_{H^1(\hat{\Gamma}_{h,*}^m)}\big(\|\nabla_{\hat{\Gamma}_{h,*}^m} \hat{e}_h^m\|_{L^2(\hat{\Gamma}_{h,*}^m)} \\ &\quad + h^{-2}\|\nabla_{\hat{\Gamma}_{h,*}^m} \hat{e}_h^m\|^2_{L^2(\hat{\Gamma}_{h,*}^m)} + \tau + (1+\kappa_{*,l})h^k\big) \end{aligned}$$

$$\lesssim h\tau\|e_h^{m+1}\|_{H^1(\hat\Gamma_{h,*}^m)}\big(\|\nabla_{\hat\Gamma_{h,*}^m}\hat e_h^m\|_{L^2(\hat\Gamma_{h,*}^m)}+h^{-2}\|\nabla_{\hat\Gamma_{h,*}^m}\hat e_h^m\|_{L^2(\hat\Gamma_{h,*}^m)}^2$$
$$+\tau+(1+\kappa_{*,l})h^k\big)$$
(Lemma 4.4 is used)
$$\lesssim \tau\|e_h^{m+1}\|_{H^1(\hat\Gamma_{h,*}^m)}\big(\|\hat e_h^m\|_{L^2(\hat\Gamma_{h,*}^m)}+h^{-2}\|\hat e_h^m\|_{L^2(\hat\Gamma_{h,*}^m)}\|\nabla_{\hat\Gamma_{h,*}^m}\hat e_h^m\|_{L^2(\hat\Gamma_{h,*}^m)}$$
$$+\tau h+(1+\kappa_{*,l})h^{k+1}\big)$$
$$\lesssim \epsilon\tau\|e_h^{m+1}\|_{H^1(\hat\Gamma_{h,*}^m)}^2+h^{0.5}\tau\|\nabla_{\hat\Gamma_{h,*}^m}\hat e_h^m\|_{L^2(\hat\Gamma_{h,*}^m)}^2+\epsilon^{-1}\tau\|\hat e_h^m\|_{L^2(\hat\Gamma_{h,*}^m)}^2$$
$$+\epsilon^{-1}\tau(\tau h+(1+\kappa_{*,l})h^{k+1})^2. \tag{5.66}$$

In the derivation of the last inequality we have used the induction assumption in (3.10), which implies

$$h^{-2}\|\hat e_h^m\|_{L^2(\hat\Gamma_{h,*}^m)}\|\nabla_{\hat\Gamma_{h,*}^m}\hat e_h^m\|_{L^2(\hat\Gamma_{h,*}^m)}\lesssim h^{0.5}\|\nabla_{\hat\Gamma_{h,*}^m}\hat e_h^m\|_{L^2(\hat\Gamma_{h,*}^m)}.$$

In the case $m\geq 1$, inequality (5.34) implies that $\|\hat e_h^m\|_{H^1(\hat\Gamma_{h,*}^m)}\lesssim\|e_h^m\|_{H^1(\hat\Gamma_{h,*}^m)}$. In the case $m=0$ we simply use $\|\hat e_h^0\|_{H^1(\hat\Gamma_{h,*}^m)}\lesssim h^k$, which holds for the initial triangulation at $t=0$. Therefore, in either case, we have

$$\|\hat e_h^m\|_{H^1(\hat\Gamma_{h,*}^m)}\lesssim\|e_h^m\|_{H^1(\hat\Gamma_{h,*}^m)}+h^k. \tag{5.67}$$

From (5.28) we see that, under the stepsize condition $\tau=o(h^{2.5})$ and the mesh size condition in (3.8),

$$\|e_h^{m+1}-\hat e_h^m\|_{L^2(\hat\Gamma_{h,*}^m)}\lesssim\tau h^{-1}\|\nabla_{\hat\Gamma_{h,*}^m}\hat e_h^m\|_{L^2(\hat\Gamma_{h,*}^m)}+\tau(\tau+(1+\kappa_{*,l})h^{k-1})$$
$$\lesssim\|\hat e_h^m\|_{L^2(\hat\Gamma_{h,*}^m)}+\tau,$$

and

$$\|e_h^{m+1}-\hat e_h^m\|_{H^1(\hat\Gamma_{h,*}^m)}\lesssim\tau h^{-2}\|\nabla_{\hat\Gamma_{h,*}^m}\hat e_h^m\|_{L^2(\hat\Gamma_{h,*}^m)}+\tau h^{-1}(\tau+(1+\kappa_{*,l})h^{k-1})$$
$$\lesssim\|\nabla_{\hat\Gamma_{h,*}^m}\hat e_h^m\|_{L^2(\hat\Gamma_{h,*}^m)}+\tau,$$

which imply that

$$\|e_h^{m+1}\|_{L^2(\hat\Gamma_{h,*}^m)}\lesssim\|\hat e_h^m\|_{L^2(\hat\Gamma_{h,*}^m)}+\tau, \tag{5.68}$$
$$\|e_h^{m+1}\|_{H^1(\hat\Gamma_{h,*}^m)}\lesssim\|\hat e_h^m\|_{H^1(\hat\Gamma_{h,*}^m)}+\tau. \tag{5.69}$$

Then, substituting (5.63)–(5.66) into (5.53) and use relations (5.68)–(5.67), we obtain

$$
\begin{aligned}
&\|\hat{e}_h^{m+1}\|^2_{L^2(\hat{\Gamma}^{m+1}_{h,*})} - \|e_h^{m+1}\|^2_{L^2(\hat{\Gamma}^m_{h,*})} \\
&= L_1 + L_2 + L_3 + L_4 \\
&\lesssim \epsilon\tau\|\nabla_{\hat{\Gamma}^m_{h,*}}\hat{e}_h^m\|^2_{L^2(\hat{\Gamma}^m_{h,*})} + C_\epsilon\tau\|\hat{e}_h^m\|^2_{L^2(\hat{\Gamma}^m_{h,*})} + C_\epsilon\tau(\tau + (1+\kappa_{*,l})h^k)^2. \qquad (5.70)
\end{aligned}
$$

5.8 Convergence of Numerical Solutions

By choosing $\phi_h = e_h^{m+1}$ in (5.5) and estimating $\|\frac{e_h^{m+1}-\hat{e}_h^m}{\tau}\|_{L^2(\hat{\Gamma}^m_{h,*})}$ by (5.28), and then replacing $\|e_h^{m+1}\|_{L^2(\hat{\Gamma}^m_{h,*})}$ by $\|\hat{e}_h^m\|_{L^2(\hat{\Gamma}^m_{h,*})}$ using (5.68), the following result can be derived:

$$
\begin{aligned}
|J^m(e_h^{m+1})| &\lesssim \|\nabla_{\hat{\Gamma}^m_{h,*}}\hat{e}_h^m\|_{L^2(\hat{\Gamma}^m_{h,*})}\|e_h^{m+1}\|_{L^2(\hat{\Gamma}^m_{h,*})} \\
&\quad + \left[h^{-1}\|\nabla_{\hat{\Gamma}^m_{h,*}}\hat{e}_h^m\|_{L^2(\hat{\Gamma}^m_{h,*})} + (\tau + (1+\kappa_{*,l})h^{k-1})\right] \\
&\quad \|\nabla_{\hat{\Gamma}^m_{h,*}}\hat{e}_h^m\|_{L^\infty(\hat{\Gamma}^m_{h,*})}\|e_h^{m+1}\|_{L^2(\hat{\Gamma}^m_{h,*})} \\
&\lesssim \|\nabla_{\hat{\Gamma}^m_{h,*}}\hat{e}_h^m\|_{L^2(\hat{\Gamma}^m_{h,*})}(\|\hat{e}_h^m\|_{L^2(\hat{\Gamma}^m_{h,*})} + \tau) \\
&\quad + \left[h^{-1}\|\nabla_{\hat{\Gamma}^m_{h,*}}\hat{e}_h^m\|_{L^2(\hat{\Gamma}^m_{h,*})} + (\tau + (1+\kappa_{*,l})h^{k-1})\right] \\
&\quad \|\nabla_{\hat{\Gamma}^m_{h,*}}\hat{e}_h^m\|_{L^\infty(\hat{\Gamma}^m_{h,*})}(\|\hat{e}_h^m\|_{L^2(\hat{\Gamma}^m_{h,*})} + \tau) \\
&\lesssim \epsilon\|\nabla_{\hat{\Gamma}^m_{h,*}}\hat{e}_h^m\|^2_{L^2(\hat{\Gamma}^m_{h,*})} + \epsilon^{-1}\|\hat{e}_h^m\|^2_{L^2(\hat{\Gamma}^m_{h,*})} + \epsilon^{-1}(\tau + (1+\kappa_{*,l})h^k)^2,
\end{aligned}
\qquad (5.71)
$$

where the last inequality follows from using the stepsize condition $\tau = o(h^{2.5})$ and Young's inequality.

Testing (5.14) with $\phi_h = e_h^{m+1}$ and using relation (5.15), we obtain

$$
\begin{aligned}
&\int_{\hat{\Gamma}^m_{h,*}} \frac{e_h^{m+1} - \hat{e}_h^m}{\tau} \cdot e_h^{m+1} + \frac{1}{2}A_{h,*}(e_h^{m+1}, e_h^{m+1}) \\
&\le \frac{1}{2}A^T_{h,*}(\hat{e}_h^m, \hat{e}_h^m) - B^m(\hat{e}_h^m, e_h^{m+1}) - J^m(e_h^{m+1}) - K^m(e_h^{m+1}) - d^m(e_h^{m+1}) \\
&\lesssim \epsilon\|\nabla_{\hat{\Gamma}^m_{h,*}}\hat{e}_h^m\|^2_{L^2(\hat{\Gamma}^m_{h,*})} + \epsilon^{-1}\|\hat{e}_h^m\|^2_{L^2(\hat{\Gamma}^m_{h,*})} + \epsilon^{-1}(\tau + (1+\kappa_{*,l})h^k)^2, \qquad (5.72)
\end{aligned}
$$

where we have estimated $J^m(e_h^{m+1})$ with (5.71) and used the estimates of $A^T_{h,*}(\hat{e}_h^m, \hat{e}_h^m)$, $B^m(\hat{e}_h^m, e_h^{m+1})$, $J^m(e_h^{m+1})$, $K^m(e_h^{m+1})$ and $d^m(e_h^{m+1})$ in (5.16)–(5.17), (5.22) and Lemma 4.3, respectively, and we have replaced e_h^{m+1} by $\hat{e}_h^m$ using the relations in (5.68)–(5.67) in these estimates. Then, using relations (5.34) and (5.70), we can further reduce (5.72) to the following result:

$$\frac{\|\hat e_h^{m+1}\|^2_{L^2(\hat\Gamma^{m+1}_{h,*})}-\|\hat e_h^{m}\|^2_{L^2(\hat\Gamma^{m}_{h,*})}}{2\tau}+C^{-1}\|\nabla_{\hat\Gamma^m_{h,*}}\hat e_h^{m+1}\|^2_{L^2(\hat\Gamma^{m}_{h,*})}$$
$$\lesssim \epsilon\|\nabla_{\hat\Gamma^m_{h,*}}\hat e_h^{m}\|^2_{L^2(\hat\Gamma^{m}_{h,*})}+C_\epsilon\|\hat e_h^{m}\|^2_{L^2(\hat\Gamma^{m}_{h,*})}+C_\epsilon(\tau+(1+\kappa_{*,l})h^k)^2, \tag{5.73}$$

where ϵ is an arbitrary small constant. Note that the constant $\kappa_{*,l}$ on the right-hand side of (5.73) can be replaced by $\kappa_{*,m}$ because all the analysis above relies on the surface $\hat\Gamma^m_{h,*}$ instead of $\hat\Gamma^l_{h,*}$. Therefore, by applying Grönwall's inequality and the norm equivalence, we obtain the following error estimate for some constant C_{κ_l} which may depend on the κ_l defined in (3.1):

$$\max_{0\le m\le l}\|\hat e_h^{m+1}\|^2_{L^2(\hat\Gamma^{m}_{h,*})}+\sum_{m=0}^{l}\tau\|\nabla_{\hat\Gamma^m_{h,*}}\hat e_h^{m+1}\|^2_{L^2(\hat\Gamma^{m}_{h,*})}$$
$$\le C_{\kappa_l}\Big(\tau+\Big(1+\sum_{m=0}^{l}\tau\kappa_{*,m}^2\Big)^{\frac12}h^k\Big)^2. \tag{5.74}$$

In view of (5.68)–(5.67), we also obtain the following result:

$$\max_{0\le m\le l}\|e_h^{m+1}\|^2_{L^2(\hat\Gamma^{m}_{h,*})}+\sum_{m=0}^{l}\tau\|\nabla_{\hat\Gamma^m_{h,*}}e_h^{m+1}\|^2_{L^2(\hat\Gamma^{m}_{h,*})}$$
$$\le C_{\kappa_l}\Big(\tau+\Big(1+\sum_{m=0}^{l}\tau\kappa_{*,m}^2\Big)^{\frac12}h^k\Big)^2. \tag{5.75}$$

Note that the lifted error $\hat e^m=\hat X_h^{m,l}-\mathrm{id}_{\Gamma^m}$ can be written as

$$\hat e^m=\hat X_h^{m,l}-\hat X_{h,*}^{m,l}+\hat X_{h,*}^{m,l}-\mathrm{id}_{\Gamma^m}=(e_h^m)^l-(a^m-I_ha^m)^l,$$

with e_h^m and $a^m-I_ha^m$ being function defined on $\hat\Gamma^m_{h,*}$, where we have used the relation $\hat X^m_{h,*}=\mathrm{id}=I_ha^m$ on $\hat\Gamma^m_{h,*}$. Since $\|a^m-I_ha^m\|_{L^2(\hat\Gamma^m_{h,*})}\le C_{\kappa_l}(1+\kappa_{*,l})h^{k+1}$, as shown in (3.4), the error estimate (5.74) also implies the following result for the lifted error $\hat e^m=\hat X_h^{m,l}-\mathrm{id}_{\Gamma^m}$:

$$\max_{0\le m\le l}\|\hat e^{m+1}\|^2_{L^2(\Gamma^m)}+\sum_{m=0}^{l}\tau\|\nabla_{\Gamma^m}\hat e^{m+1}\|^2_{L^2(\Gamma^m)}\le C_{\kappa_l}(\tau+(1+\kappa_{*,l})h^k)^2. \tag{5.76}$$

It remains to show that the constants κ_l and $\kappa_{*,l}$ defined in (3.1) are bounded uniformly with respect to τ, h and l (thus κ_{l+1} and $\kappa_{*,l+1}$ are also bounded uniformly with respect to τ and h). This would recover induction hypotheses (3) in Sect. 3.3 at time level t_{l+1}. Then, under the stepsize condition $\tau=o(h^{2.5})$, for sufficiently small h satisfying (3.8), the error estimate in (5.74) can be used to recover the induction hypotheses (1)–(2) in Sect. 3.3 at time level t_{l+1}. In fact, induction hypothesis (1) follows from the L^∞ error estimate using (5.74) and the inverse inequality, for sufficiently small h satisfying (3.8); induction hypothesis (2) follows from (5.74) directly for sufficiently small h satisfying (3.8).

The boundedness of the constants κ_l and $\kappa_{*,l}$ uniformly with respect to τ, h and l is proved under a stronger stepsize condition $\tau \le ch^k$ in Appendix based on the error estimates in (5.74)–(5.75). This would complete the proof of Theorem 2.1. □

6 Numerical Experiments

In this section, we present numerical experiments to support the theoretical analysis by testing the convergence rate of Dziuk's fully discrete parametric FEM for the evolution of surface $\Gamma^0 = \mathbb{S}^2$ under mean curvature flow. Since the unit sphere $\mathbb{S}^2$ is a self-shrinker under mean curvature flow, the exact solution at time t is a sphere of radius

$$R(t) = \sqrt{1-4t} \quad \text{for} \quad t \in \Big[0, \frac{1}{4}\Big). \tag{6.1}$$

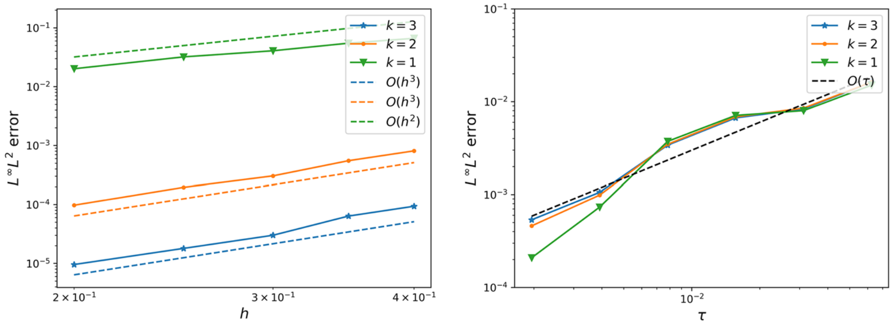
The surface shrinks to a point singularity at $t = \frac{1}{4}$.

We test the convergence of algorithm for the evolution of the surface under mean curvature flow up to $T = 0.1$. Thus the surface keeps to be smooth with bounded curvature for $t \in [0, T]$. Although we have only proved the convergence of Dziuk's fully discrete parametric FEM for finite elements of degree $k \ge 3$ in Theorem 2.1, we test all the cases of $k = 1, 2, 3$ for the $L^\infty L^2$ norm and $L^2 H^1$ seminorm of the error, i.e. $\max\limits_{0\le m\le[T/\tau]} \|\hat e_h^m\|_{L^2(\hat\Gamma_{h,*}^m)}$ and $\sum_{m=0}^{[T/\tau]} \tau \|\nabla_{\hat\Gamma_{h,*}^m} \hat e_h^m\|^2_{L^2(\hat\Gamma_{h,*}^m)}$. The $L^\infty L^2$ and $L^2 H^1$ spatial discretization errors at $T = 0.1$ (with a sufficiently small stepsize $\tau_{\rm ref} = 2^{-13}, 2^{-14}, 2^{-15}$ for $k = 1, 2, 3$, respectively) are presented in Figs. 2A and 3A. The $L^\infty L^2$ and $L^2 H^1$ temporal discretization errors at $T = 0.1$ (measured with a sufficiently small mesh size $h_{\rm ref} = 0.025, 0.2, 0.2$ for $k = 1, 2, 3$, respectively) are presented in Figs. 2B and 3B.

From the numerical results in Figs. 2 and 3 we observe $O(\tau + h^k)$ rate of convergence for $k = 3$, and $O(\tau + h^{k+1})$ rate of convergence for $k = 1, 2$. For finite elements of degree $k = 3$, the convergence order observed in the numerical results is consistent with the theoretical analysis in Theorem 2.1. However, the necessity of the stepsize condition $\tau \le ch^k$ is not observed. Thus the stepsize condition $\tau \le ch^k$ may be a technical condition that could be removed from the convergence analysis. The rigorous analysis of stability and convergence of Dziuk's parametric FEMs with the low-order finite elements of degree $k = 1, 2$ still remains open.

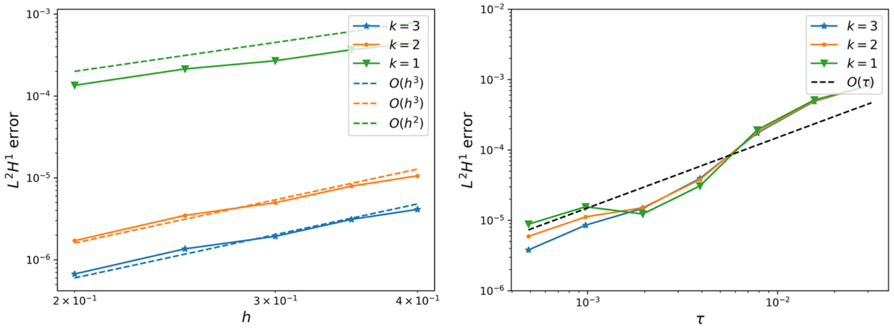
7 Conclusions

We have introduced a new approach for analyzing the errors of parametric finite element approximations to surface evolution under geometric flows—to estimate the



(A) Spatial discretization errors (B) Temporal discretization errors

Fig. 2 $L^\infty L^2$ errors of numerical approximations to mean curvature flow



(A) Spatial discretization errors (B) Temporal discretization errors

Fig. 3 L^2H^1 errors of numerical approximations to mean curvature flow

projected distance from the numerically computed surface to the exact surface, rather than the distance between particle trajectories of the two surfaces as in the literature. We have established a corresponding new framework in Sect. 3 which includes the analysis of the approximation properties of the interpolated surface, the induction assumptions for the accuracy of approximations, and the geometric relations arising from distance projection at nodes, which are not only used in this article to prove the convergence of Dziuk's parametric FEM for mean curvature flow, but also applicable to the analysis of other geometric flows and parametric finite element algorithms. Based on the new approach introduced in this article, we have recovered the full H^1 parabolicity of mean curvature flow and correspondingly established improved convergence order for Dziuk's fully discrete parametric FEM (for which the convergence has remained open in the last three decades) for finite elements of degree $k \geq 3$.

In addition to the proof of convergence of Dziuk's fully discrete parametric FEM for mean curvature flow, since this new approach is proposed to estimate the projected distance instead of the error between particle trajectories, it automatically neglects the tangential motion in the numerical approximation and therefore provides a foundational mathematical tool for analyzing other parametric FEMs which contain artificial tangential motions. This will be demonstrated in some subsequent articles through the analysis of stability and convergence of algorithms which contain artificial tangential velocities to improve the mesh quality of the numerically computed surfaces, such as the BGN type of methods for mean curvature flow and surface diffusion.

Acknowledgements The work in this paper was partially supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No. PolyU15303022), and an internal grant of The Hong Kong Polytechnic University (Project ID: P0038843).

Appendix: Optimal Approximation Properties of the Interpolated Surfaces

The quantities κ_l and $\kappa_{*,l}$ defined in (3.1) characterize the shape regularity, quasi-uniformity and optimal approximation properties of the interpolated surface $\hat\Gamma^m_{h,*}$. In this appendix, we show that κ_l and $\kappa_{*,l}$ have an upper bound which may depend on the exact solution and T, but is independent of τ, h and l. In order to make the argument clear, we denote by C_{κ_l} and C_0 some generic constants which are dependent and independent of κ_l, respectively.

A.1. Boundedness of Discrete Flow Maps in the $W^{k-1,\infty}$ and H^k Norms: Part I

In terms of the notation in Sect. 3.2, for a curved triangle $K^0 \subset \Gamma^0_h$ we denote by $K^0_{\rm f}$ the unique flat triangle with the same three vertices as K^0, and consider the piecewise flat triangular surface

$$\Gamma^0_{h,{\rm f}} = \bigcup_{K^0\subset\Gamma^0_h} K^0_{\rm f}.$$

We still denote by $\hat X^m_{h,*} : \Gamma^0_{h,{\rm f}} \to \hat\Gamma^m_{h,*}$ the unique piecewise polynomial of degree k (with the nodal vector $\hat{\mathbf{x}}^m_*$ as before) which parametrizes $\hat\Gamma^m_{h,*}$, and consider the following decomposition:

$$\hat X^{q+1}_{h,*} = \hat X^0_{h,*} + \sum_{m=0}^{q} (\hat X^{m+1}_{h,*} - \hat X^m_{h,*}).$$

Using the triangle inequality and the good quality of initial triangulation at $t = 0$, as shown in (2.1), we have

$$\|\hat X^{q+1}_{h,*}\|_{W^{j,\infty}_h(\Gamma^0_{h,{\rm f}})} \le C_0 + \sum_{m=0}^{q} \|\hat X^{m+1}_{h,*} - \hat X^m_{h,*}\|_{W^{j,\infty}_h(\Gamma^0_{h,{\rm f}})} \quad \text{for } 1 \le j \le k-1, \tag{A.1}$$

$$\|\hat X^{q+1}_{h,*}\|_{H^j_h(\Gamma^0_{h,{\rm f}})} \le C_0 + \sum_{m=0}^{q} \|\hat X^{m+1}_{h,*} - \hat X^m_{h,*}\|_{H^j_h(\Gamma^0_{h,{\rm f}})} \quad \text{for } 1 \le j \le k, \tag{A.2}$$

where $\|\cdot\|_{W_h^{j,\infty}(\Gamma_{h,\mathrm{f}}^0)}$ and $\|\cdot\|_{H_h^j(\Gamma_{h,\mathrm{f}}^0)}$ denote the piecewise $W^{j,\infty}$ norm and piecewise H^j norm, respectively, on the piecewise flat triangular surface $\Gamma_{h,\mathrm{f}}^0$; see the definition of these piecewise Sobolev norms in Sect. 3.2.

In the next section, we shall prove the following two results under the condition $\tau \le ch^k$ and $h \le h_{\kappa_m}$ (where h_{κ_m} is some constant depending on κ_m):

$$\begin{aligned}
&\|\hat{X}_{h,*}^{m+1} - \hat{X}_{h,*}^m\|_{W_h^{j,\infty}(\Gamma_{h,\mathrm{f}}^0)} \\
&\le C_0\tau[1 + (j-1)\|\hat{X}_{h,*}^m\|_{W_h^{j-1,\infty}(\Gamma_{h,\mathrm{f}}^0)}^j] + C_0\tau\|\hat{X}_{h,*}^m\|_{W_h^{j,\infty}(\Gamma_{h,\mathrm{f}}^0)} \\
&\quad + C_0h^{k-j}\tau + C_{\kappa_m}(1+\kappa_{*,m})h^{k-j-1}\tau \\
&\quad + C_{\kappa_m}h^{-j-1}\big(\|T_*^m(e_h^{m+1} - \hat{e}_h^m)\|_{L^2(\hat{\Gamma}_{h,*}^m)} + h\|e_h^{m+1} - \hat{e}_h^m\|_{L^2(\hat{\Gamma}_{h,*}^m)}\big) \qquad \text{(A.3)}
\end{aligned}$$

and

$$\begin{aligned}
&\|\hat{X}_{h,*}^{m+1} - \hat{X}_{h,*}^m\|_{H_h^j(\Gamma_{h,\mathrm{f}}^0)} \\
&\le C_0\tau[1 + (j-1)(j-2)\|\hat{X}_{h,*}^m\|_{W_h^{j-2,\infty}(\Gamma_{h,\mathrm{f}}^0)}^j \\
&\quad + (j-1)\|\hat{X}_{h,*}^m\|_{W^{1,\infty}(\Gamma_{h,\mathrm{f}}^0)}\|\hat{X}_{h,*}^m\|_{H_h^{j-1,\infty}(\Gamma_{h,\mathrm{f}}^0)}^j] \\
&\quad + C_0j\tau\|\hat{X}_{h,*}^m\|_{H_h^j(\Gamma_{h,\mathrm{f}}^0)} + C_0h^{k-j}\tau + C_{\kappa_m}(1+\kappa_{*,m})h^{k-j}\tau \\
&\quad + C_{\kappa_m}h^{-j}\big(\|T_*^m(e_h^{m+1} - \hat{e}_h^m)\|_{L^2(\hat{\Gamma}_{h,*}^m)} + h\|e_h^{m+1} - \hat{e}_h^m\|_{L^2(\hat{\Gamma}_{h,*}^m)}\big). \qquad \text{(A.4)}
\end{aligned}$$

From (5.28), (5.52) and (5.74) we see that, by applying the inverse inequality,

$$\begin{aligned}
&h^{-j-1}\|T_*^m(e_h^{m+1} - \hat{e}_h^m)\|_{L^2(\hat{\Gamma}_{h,*}^m)} + h^{-j}\|e_h^{m+1} - \hat{e}_h^m\|_{L^2(\hat{\Gamma}_{h,*}^m)} \\
&\le C_{\kappa_m}h^{-j-1}\tau\|\nabla_{\hat{\Gamma}_{h,*}^m}\hat{e}_h^m\|_{L^2(\hat{\Gamma}_{h,*}^m)} + C_{\kappa_m}h^{-j-3}\tau\|\nabla_{\hat{\Gamma}_{h,*}^m}\hat{e}_h^m\|_{L^2(\hat{\Gamma}_{h,*}^m)}^2 \\
&\quad + C_{\kappa_m}h^{-j-1}\tau(1+\kappa_{*,m})h^k. \qquad \text{(A.5)}
\end{aligned}$$

Then, substituting this result into (A.3) and choosing $j = 1$, we derive the following estimate for $0 \le q \le l$:

$$\begin{aligned}
&\sum_{m=0}^{q}\|\hat{X}_{h,*}^{m+1} - \hat{X}_{h,*}^m\|_{W^{1,\infty}(\Gamma_{h,\mathrm{f}}^0)} \\
&\le C_0 + \sum_{m=0}^{q}C_0\tau\|\hat{X}_{h,*}^m\|_{W^{1,\infty}(\Gamma_{h,\mathrm{f}}^0)} + C_0h^{k-1} + \sum_{m=0}^{q}C_{\kappa_m}\tau(1+\kappa_{*,m})h^{k-2}
\end{aligned}$$

$$+\sum_{m=0}^{q} C_{\kappa_m}\Big[h^{-2}\tau\|\nabla_{\hat\Gamma^m_{h,*}}\hat e^m_h\|_{L^2(\hat\Gamma^m_{h,*})}+h^{-4}\tau\|\nabla_{\hat\Gamma^m_{h,*}}\hat e^m_h\|^2_{L^2(\hat\Gamma^m_{h,*})}\Big]$$

$$\le C_0+C_{\kappa_l}(1+\kappa_{*,l})h^{k-2}+C_{\kappa_l}(1+\kappa_{*,l})^2h^{2k-4}+\sum_{m=0}^{q}C_0\tau\|\hat X^m_{h,*}\|_{W^{1,\infty}(\Gamma^0_{h,\mathrm{f}})}.$$

Here we have used the error estimate in (5.74) with $\tau \le ch^k$ in the last inequality. Since $k \ge 3$, for sufficiently small mesh size $h \le h_{\kappa_l,\kappa_{*,l}}$ (with some constant which depends on κ_l and $\kappa_{*,l}$), substituting the last inequality into (A.1) and taking the square yield the following result:

$$\|\hat X^{q+1}_{h,*}\|^2_{W^{1,\infty}(\Gamma^0_{h,\mathrm{f}})}\le C_0+\sum_{m=0}^{q}C_0\tau\|\hat X^m_{h,*}\|^2_{W^{1,\infty}(\Gamma^0_{h,\mathrm{f}})}\quad\text{for }\ 0\le q\le l. \tag{A.6}$$

Then, by applying the discrete Grönwall's inequality and taking the square root, we obtain

$$\max_{0\le q\le l}\|\hat X^{q+1}_{h,*}\|_{W^{1,\infty}(\Gamma^0_{h,\mathrm{f}})}\le C_0. \tag{A.7}$$

Now, by using mathematical induction, we shall prove that if $j \le k-2$ and $\max\limits_{0\le q\le l}\|\hat X^{q+1}_{h,*}\|_{W_h^{j-1,\infty}(\Gamma^0_{h,\mathrm{f}})}\le C_0$ then

$$\max_{0\le q\le l}\|\hat X^{q+1}_{h,*}\|_{W_h^{j,\infty}(\Gamma^0_{h,\mathrm{f}})}\le C_0.$$

In fact, if $\max\limits_{0\le q\le l}\|\hat X^{q+1}_{h,*}\|_{W^{j-1,\infty}(\Gamma^0_{h,\mathrm{f}})}\le C_0$ then summing up (A.3) for $m=0,\dots,q$ and using (A.5) yield the following result:

$$\sum_{m=0}^{q}\|\hat X^{m+1}_{h,*}-\hat X^m_{h,*}\|_{W_h^{j,\infty}(\Gamma^0_h)}$$

$$\le C_0(1+h^{k-j})+\sum_{m=0}^{q}C_0\tau\|\hat X^m_{h,*}\|_{W_h^{j,\infty}(\Gamma^0_{h,\mathrm{f}})}+\sum_{m=0}^{q}C_{\kappa_m}h^{k-j-1}\tau(1+\kappa_{*,m})$$

$$+\sum_{m=0}^{q}C_{\kappa_m}\Big[h^{-j-1}\tau\|\nabla_{\hat\Gamma^m_{h,*}}\hat e^m_h\|_{L^2(\hat\Gamma^m_{h,*})}+h^{-j-3}\tau\|\nabla_{\hat\Gamma^m_{h,*}}\hat e^m_h\|^2_{L^2(\hat\Gamma^m_{h,*})}\Big]$$

$$\le C_0(1+h^{k-j})+\sum_{m=0}^{q}C_0\tau\|\hat X^m_{h,*}\|_{W_h^{j,\infty}(\Gamma^0_{h,\mathrm{f}})}$$

$$+C_{\kappa_l}\Big(1+\sum_{m=0}^{q}\tau\kappa^2_{*,m}\Big)^{\frac12}h^{k-j-1}+C_{\kappa_l}(1+\kappa_{*,l})^2h^{2k-j-3},$$

where we have used the error estimate in (5.74) in the last inequality. Since $k \geq 3$, for sufficiently small mesh size $h \leq h_{\kappa_l,\kappa_{*,l}}$ (with some constant which depends on κ_l and $\kappa_{*,l}$), substituting the last inequality into (A.1) and taking the square yield the following result for $j \leq k-2$:

$$\|\hat{X}_{h,*}^{q+1}\|^2_{W_h^{j,\infty}(\Gamma_{h,\mathrm{f}}^0)} \leq C_0 + \sum_{m=0}^{q} C_0\tau\|\hat{X}_{h,*}^m\|^2_{W_h^{j,\infty}(\Gamma_{h,\mathrm{f}}^0)} \quad \text{for } 0 \leq q \leq l.$$

Then, by applying the discrete Grönwall's inequality and taking the square root, we obtain

$$\max_{0\leq q\leq l} \|\hat{X}_{h,*}^{q+1}\|_{W_h^{j,\infty}(\Gamma_{h,\mathrm{f}}^0)} \leq C_0. \tag{A.8}$$

This proves (A.8) for $1 \leq j \leq k-2$. Analogously, by using (A.4), we can prove the following result for $1 \leq j \leq k-1$:

$$\max_{0\leq q\leq l} \|\hat{X}_{h,*}^{q+1}\|_{H_h^{j}(\Gamma_{h,\mathrm{f}}^0)} \leq C_0. \tag{A.9}$$

Similar estimate for $\max_{0\leq q\leq l} \|(\hat{X}_{h,*}^{q+1})^{-1}\|_{W_h^{1,\infty}(\hat{\Gamma}_{h,*}^j)}$ can also be proved and omitted here. This proves that if $\tau \leq ch^k$ and $h \leq h_{\kappa_l,\kappa_{*,l}}$ then $\kappa_{l+1} \leq C_0$ in view of the definition in (3.1).

Therefore, we can replace C_{κ_m} by C_0 in (A.3)–(A.4) and obtain the following results for $0 \leq q \leq l$ in the same way as above, under the conditions $\tau \leq ch^k$ and $h \leq h_{\kappa_l,\kappa_{*,l}}$:

$$\|\hat{X}_{h,*}^{q+1}\|^2_{W_h^{k-1,\infty}(\Gamma_{h,\mathrm{f}}^0)} \leq C_0 + \sum_{m=0}^{q} C_0\tau\kappa_{*,m}^2 + \sum_{m=0}^{q} C_0\tau\|\hat{X}_{h,*}^m\|^2_{W_h^{k-1,\infty}(\Gamma_{h,\mathrm{f}}^0)}, \tag{A.10}$$

$$\|\hat{X}_{h,*}^{l+1}\|^2_{H_h^{k}(\Gamma_{h,\mathrm{f}}^0)} \leq C_0 + \sum_{m=0}^{q} C_0\tau\kappa_{*,m}^2 + \sum_{m=0}^{q} C_0\tau\|\hat{X}_{h,*}^m\|^2_{H_h^{k}(\Gamma_{h,\mathrm{f}}^0)}. \tag{A.11}$$

In regard to the definition of $\kappa_{*,m}$ in (3.1), we can replace $\kappa_{*,m}$ by $\|\hat{X}_{h,*}^m\|_{W_h^{k-1,\infty}(\Gamma_{h,\mathrm{f}}^0)} + \|\hat{X}_{h,*}^m\|_{H_h^{k}(\Gamma_{h,\mathrm{f}}^0)}$ and then sum up the two inequalities above. This yields that

$$\begin{aligned}
&\|\hat{X}_{h,*}^{q+1}\|^2_{H_h^{k}(\Gamma_{h,\mathrm{f}}^0)} + \|\hat{X}_{h,*}^{q+1}\|^2_{W_h^{k-1,\infty}(\Gamma_{h,\mathrm{f}}^0)} \\
&\leq C_0 + \sum_{m=0}^{q} C_0\tau(\|\hat{X}_{h,*}^m\|^2_{H_h^{k}(\Gamma_{h,\mathrm{f}}^0)} + \|\hat{X}_{h,*}^m\|^2_{W_h^{k-1,\infty}(\Gamma_{h,\mathrm{f}}^0)}).
\end{aligned} \tag{A.12}$$

By applying Grönwall's inequality and taking the square root, we obtain

$$\max_{0\leq q\leq l}(\|\hat{X}_{h,*}^{q+1}\|_{H_h^{k}(\Gamma_{h,\mathrm{f}}^0)} + \|\hat{X}_{h,*}^{q+1}\|_{W_h^{k-1,\infty}(\Gamma_{h,\mathrm{f}}^0)}) \leq C_0. \tag{A.13}$$

This proves that $\kappa_{*,l+1} \le C_0$ in view of its definition in (3.1). For this constant C_0 (which is independent of l) we get the following result: If $\kappa_l \le C_0$ and $\kappa_{*,l} \le C_0$, and $\tau \le ch^k$ with $h \le h_{C_0,C_0}$, then

$$\kappa_{l+1} \le C_0 \quad \text{and} \quad \kappa_{*,l+1} \le C_0. \tag{A.14}$$

This proves (A.14) by mathematical induction under the conditions $\tau \le ch^k$ and $h \le h_{C_0,C_0}$. As a result, the quantities κ_l and $\kappa_{*,l}$ defined in (3.1) are uniformly bounded with respect to τ, h and l under the required conditions on the stepsize and mesh size.

A.2. Boundedness of Discrete Flow Maps in the $W^{k-1,\infty}$ and H^k Norms: Part II

In this appendix we prove (A.3)–(A.4), which are used in Appendix A.1 to prove (A.14).

Note that the nodal vectors $\hat{\mathbf{x}}_*^m$ and $\hat{\mathbf{x}}_*^{m+1}$ are defined as the distance projection of $\mathbf{x}^m$ and $\mathbf{x}^{m+1}$ onto the smooth surfaces Γ^m and Γ^{m+1}. Therefore, $\hat{e}_h^m = X_h^m - \hat{X}_{h,*}^m$ and $\hat{e}_h^{m+1} = X_h^{m+1} - \hat{X}_{h,*}^{m+1}$ are in the directions of n_*^m and n_*^{m+1} at the nodes, respectively. From the geometric relation in Fig. 4 we observe the following vector decomposition at the j-th node:

$$N_*^m(\hat{x}_{j,*}^{m+1} - \hat{x}_{j,*}^m) = x_{j,*}^{m+1} - \hat{x}_{j,*}^m + \rho_h|_{j\text{-th node}},$$

and passing to finite element functions, it holds that

$$N_*^m(\hat{X}_{h,*}^{m+1} - \hat{X}_{h,*}^m) = (X^{m+1} - \mathrm{id}) \circ a^m + \rho_h \quad \text{at the nodes of } \hat{\Gamma}_{h,*}^m \tag{A.15}$$

for some finite element function ρ_h such that by the triangle inequality

$$|\rho_h| \le C_0\tau^2 + C_0|T_*^m(\hat{X}_{h,*}^{m+1} - \hat{X}_{h,*}^m)|^2 \quad \text{at the nodes,} \tag{A.16}$$

where $C_0\tau^2$ arises from the quadratic term in the Taylor expansion of the exact flow, which measures the deviation of $X_{h,*}^{m+1}$ away from the normal direction, while $|T_*^m(\hat{X}_{h,*}^{m+1} - \hat{X}_{h,*}^m)|^2$ measures the difference of lengths in the normal direction, as shown in Fig. 4. The latter is essentially the product of $|T_*^m(\hat{X}_{h,*}^{m+1} - \hat{X}_{h,*}^m)|$ (the length of one side of a right triangle) and the tangent of an angle whose amplitude is of order $O(|T_*^m(\hat{X}_{h,*}^{m+1} - \hat{X}_{h,*}^m)|)$.

Moreover, since $T_*^m(\hat{X}_{h,*}^m - X_h^m) = 0$ at the nodes and $T_*^m N_*^m = 0$, the following relation holds:

$$\begin{aligned} T_*^m(\hat{X}_{h,*}^{m+1} - \hat{X}_{h,*}^m) &= T_*^m(X_h^{m+1} - X_h^m) + T_*^m(\hat{X}_{h,*}^{m+1} - X_h^{m+1}) - T_*^m(\hat{X}_{h,*}^m - X_h^m) \\ &= T_*^m(X_h^{m+1} - X_h^m) + T_*^m N_*^{m+1}(\hat{X}_{h,*}^{m+1} - X_h^{m+1}) \\ &= T_*^m(X_h^{m+1} - X_h^m) + T_*^m(N_*^{m+1} - N_*^m)\hat{e}_h^{m+1} \quad \text{at the nodes.} \end{aligned} \tag{A.17}$$

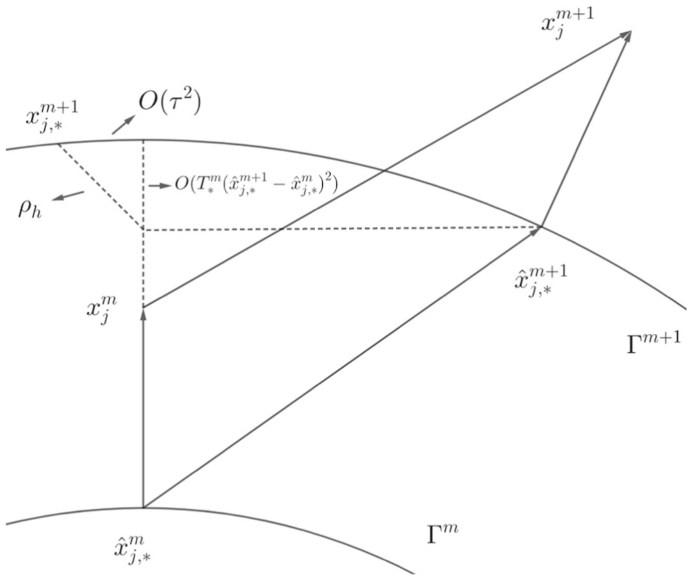


Fig. 4 The geometric relation at the j-th node

In the last equality we have used $\hat{e}_h^{m+1} = \hat{X}_{h,*}^{m+1} - X_h^{m+1}$. Therefore, by decomposing $\hat{X}_{h,*}^{m+1} - \hat{X}_{h,*}^m$ into the normal and tangential components and applying the triangle inequality, we have

$$
\begin{aligned}
&\|\hat{X}_{h,*}^{m+1} - \hat{X}_{h,*}^m\|_{W_h^{j,\infty}(\Gamma_{h,\mathrm{f}}^0)} \\
&\le \|I_h N_*^m(\hat{X}_{h,*}^{m+1} - \hat{X}_{h,*}^m)\|_{W_h^{j,\infty}(\Gamma_{h,\mathrm{f}}^0)} + \|I_h T_*^m(\hat{X}_{h,*}^{m+1} - \hat{X}_{h,*}^m)\|_{W_h^{j,\infty}(\Gamma_{h,\mathrm{f}}^0)} \\
&\le \|I_h[(X^{m+1} - \mathrm{id}) \circ a^m \circ \hat{X}_{h,*}^m] + \rho_h\|_{W_h^{j,\infty}(\Gamma_{h,\mathrm{f}}^0)} \\
&\quad \text{(relation (A.15) is pulled back to } \Gamma_{h,\mathrm{f}}^0) \\
&\quad + C_0 h^{-j} \|I_h T_*^m(\hat{X}_{h,*}^{m+1} - \hat{X}_{h,*}^m)\|_{L^\infty(\Gamma_{h,\mathrm{f}}^0)}. \qquad \text{(A.18)}
\end{aligned}
$$

The first term on the right-hand side of (A.18) can be estimated by using (A.16) as follows:

$$
\begin{aligned}
&\|I_h[(X^{m+1} - \mathrm{id}) \circ a^m \circ \hat{X}_{h,*}^m] + \rho_h\|_{W_h^{j,\infty}(\Gamma_{h,\mathrm{f}}^0)} \\
&\le C_0 \|X^{m+1} - \mathrm{id}\|_{W^{j,\infty}(\Gamma^m)} \Big(1 + \sum_{\substack{j_1+\cdots+j_i \le j \\ j_1,\ldots,j_i \ge 1}} \|\hat{X}_{h,*}^m\|_{W_h^{j_1,\infty}(\Gamma_{h,\mathrm{f}}^0)} \cdots \|\hat{X}_{h,*}^m\|_{W_h^{j_i,\infty}(\Gamma_{h,\mathrm{f}}^0)}\Big) \\
&\quad + C_0 h^{-j} \|\rho_h\|_{L^\infty(\Gamma_{h,\mathrm{f}}^0)}
\end{aligned}
$$

$$
\begin{aligned}
&\leq C_0\tau[1+j(j-1)\|\hat{X}^m_{h,*}\|^j_{W_h^{j-1,\infty}(\Gamma^0_{h,\mathrm{f}})}]+C_0 j\tau\|\hat{X}^m_{h,*}\|_{W_h^{j,\infty}(\Gamma^0_{h,\mathrm{f}})}\\
&\quad+C_0h^{-j}(\tau^2+\|I_hT^m_*(\hat{X}^{m+1}_{h,*}-\hat{X}^m_{h,*})\|^2_{L^\infty(\Gamma^0_{h,\mathrm{f}})}).
\end{aligned}\tag{A.19}
$$

Here we have added a factor $j(j-1)$ in front of $\|\hat{X}^m_{h,*}\|_{W_h^{j-1,\infty}(\Gamma^0_{h,\mathrm{f}})}$ to indicate that this term should disappear in the case $j=1$, and we have added a factor j in front of $\|\hat{X}^m_{h,*}\|_{W_h^{j,\infty}(\Gamma^0_{h,\mathrm{f}})}$ to indicate that this term should disappear in the case $j=0$.

The second term on the right-hand side of (A.18), as well as the last term on the right-hand side of (A.19), can be estimated by using relation (A.17), i.e.,

$$
\begin{aligned}
&\|I_hT^m_*(\hat{X}^{m+1}_{h,*}-\hat{X}^m_{h,*})\|_{L^\infty(\Gamma^0_{h,\mathrm{f}})}\\
&\leq \|I_hT^m_*(X^{m+1}_h-X^m_h)\|_{L^\infty(\Gamma^0_{h,\mathrm{f}})}+\|I_hT^m_*(N^{m+1}_*-N^m_*)\hat{e}^{m+1}_h\|_{L^\infty(\Gamma^0_{h,\mathrm{f}})}.
\end{aligned}\tag{A.20}
$$

In the case $j=0$ we get from (A.18) and (A.19) the following result:

$$
\begin{aligned}
&\|\hat{X}^{m+1}_{h,*}-\hat{X}^m_{h,*}\|_{L^\infty(\Gamma^0_{h,\mathrm{f}})}\\
&\leq C_0\tau+C_0\|I_hT^m_*(\hat{X}^{m+1}_{h,*}-\hat{X}^m_{h,*})\|^2_{L^\infty(\Gamma^0_{h,\mathrm{f}})}+\|I_hT^m_*(\hat{X}^{m+1}_{h,*}-\hat{X}^m_{h,*})\|_{L^\infty(\Gamma^0_{h,\mathrm{f}})}\\
&\leq C_0\tau+C_0\|I_hT^m_*(\hat{X}^{m+1}_{h,*}-\hat{X}^m_{h,*})\|_{L^\infty(\Gamma^0_{h,\mathrm{f}})},
\end{aligned}\tag{A.21}
$$

where the last inequality follows from the estimate in (5.39), which implies that

$$
\|\hat{X}^{m+1}_{h,*}-\hat{X}^m_{h,*}\|_{L^\infty(\Gamma^0_{h,\mathrm{f}})}+\|I_hT^m_*(\hat{X}^{m+1}_{h,*}-\hat{X}^m_{h,*})\|_{L^\infty(\Gamma^0_{h,\mathrm{f}})}\leq C_{\kappa_l}h^{0.5}\leq 1, \tag{A.22}
$$

when $h\leq h_{\kappa_l}$ (for some constant h_{κ_l} which may depend on κ_l).

In the case $j\geq 1$ we obtain from (A.18)–(A.19) and (A.22) the following result:

$$
\begin{aligned}
&\|\hat{X}^{m+1}_{h,*}-\hat{X}^m_{h,*}\|_{W_h^{j,\infty}(\Gamma^0_{h,\mathrm{f}})}\\
&\leq C_0\tau[1+j(j-1)\|\hat{X}^m_{h,*}\|^j_{W_h^{j-1,\infty}(\Gamma^0_{h,\mathrm{f}})}]+C_0j\tau\|\hat{X}^m_{h,*}\|_{W_h^{j,\infty}(\Gamma^0_{h,\mathrm{f}})}\\
&\quad+C_0h^{-j}\tau^2+C_0h^{-j}\|I_hT^m_*(\hat{X}^{m+1}_{h,*}-\hat{X}^m_{h,*})\|_{L^\infty(\Gamma^0_{h,\mathrm{f}})}.
\end{aligned}\tag{A.23}
$$

The first term on the right-hand side of (A.20) can be estimated by using the geometric relation in (3.15), which implies that

$$
\begin{aligned}
&\|I_hT^m_*(X^{m+1}_h-X^m_h)\|_{L^\infty(\Gamma^0_{h,\mathrm{f}})}\\
&\leq \|I_hT^m_*(e^{m+1}_h-\hat{e}^m_h)\|_{L^\infty(\Gamma^0_{h,\mathrm{f}})}+\tau\|I_hT^m_*I_h(H^mn^m-g)\|_{L^\infty(\Gamma^0_{h,\mathrm{f}})}\\
&=\|I_hT^m_*(e^{m+1}_h-\hat{e}^m_h)\|_{L^\infty(\Gamma^0_{h,\mathrm{f}})}+\tau\|I_hT^m_*I_hg\|_{L^\infty(\Gamma^0_{h,\mathrm{f}})}\\
&\quad(\text{as } T^m_*I_h(H^mn^m)=0 \text{ at the nodes})\\
&\leq \|I_hT^m_*(e^{m+1}_h-\hat{e}^m_h)\|_{L^\infty(\hat{\Gamma}^m_{h,*})}+C_0\tau\|g\|_{L^\infty(\Gamma^m)},
\end{aligned}
$$

where the last inequality follows from the L^∞ stability of the Lagrange interpolation operator (with respect to the nodal values) on the initial triangulated surface $\Gamma^0_{h,\mathrm{f}}$, and the fact that the nodal values of $I_h T^m_* I_h g$ is bounded by $\|g\|_{L^\infty(\Gamma^m)}$. Then, by applying the inverse inequality to convert the $L^\infty(\hat{\Gamma}^m_{h,*})$ norm to the $L^2(\hat{\Gamma}^m_{h,*})$ norm (with a constant depending on κ_l and independent of $\kappa_{*,l}$), we obtain

$$
\begin{aligned}
&\|I_h T^m_*(X^{m+1}_h - X^m_h)\|_{L^\infty(\Gamma^0_{h,\mathrm{f}})} \\
&\le C_{\kappa_m} h^{-1}\|I_h T^m_*(e^{m+1}_h - \hat{e}^m_h)\|_{L^2(\hat{\Gamma}^m_{h,*})} + C_0\tau^2 \qquad \text{(here (3.14) is used)} \\
&\le C_{\kappa_m} h^{-1}\big(\|T^m_*(e^{m+1}_h - \hat{e}^m_h)\|_{L^2(\hat{\Gamma}^m_{h,*})} + h\|e^{m+1}_h - \hat{e}^m_h\|_{L^2(\hat{\Gamma}^m_{h,*})}\big) + C_0\tau^2. \qquad \text{(A.24)}
\end{aligned}
$$

In the last inequality, we have used the super-approximation estimates in Lemma 4.4.

The second term on the right-hand side of (A.20) can be estimated by using the inverse inequality and the expression $N^m_* = (n^m_* \circ \hat{X}^m_{h,*})(n^m_* \circ \hat{X}^m_{h,*})^\top$ at the nodes, i.e.,

$$
\begin{aligned}
&\|I_h T^m_*(N^{m+1}_* - N^m_*)\hat{e}^{m+1}_h\|_{L^\infty(\hat{\Gamma}^m_{h,*})} \\
&\le C_{\kappa_m}\|I_h[(N^{m+1}_* - N^m_*)\hat{e}^{m+1}_h]\|_{L^\infty(\hat{\Gamma}^m_{h,*})} \\
&\le C_{\kappa_m}\|I_h[n^{m+1}_* \circ \hat{X}^{m+1}_{h,*} - n^m_* \circ \hat{X}^m_{h,*}]\|_{L^\infty(\hat{\Gamma}^m_{h,*})}\|\hat{e}^{m+1}_h\|_{L^\infty(\hat{\Gamma}^m_{h,*})} \\
&\le C_{\kappa_m}\|I_h[n^{m+1}_* \circ \hat{X}^{m+1}_{h,*} - n^{m+1}_* \circ X^{m+1}_{h,*}]\|_{L^\infty(\hat{\Gamma}^m_{h,*})}\|\hat{e}^{m+1}_h\|_{L^\infty(\hat{\Gamma}^m_{h,*})} \\
&\quad + C_{\kappa_m}\|I_h[n^{m+1}_* \circ X^{m+1}_{h,*} - n^m_* \circ \hat{X}^m_{h,*}]\|_{L^\infty(\hat{\Gamma}^m_{h,*})}\|\hat{e}^{m+1}_h\|_{L^\infty(\hat{\Gamma}^m_{h,*})} \\
&\le C_{\kappa_m}(\|I_h T^{m+1}_*(\hat{X}^{m+1}_{h,*} - X^{m+1}_{h,*})\|_{L^\infty(\hat{\Gamma}^m_{h,*})} + \|\hat{X}^{m+1}_{h,*} - \\
&\quad X^{m+1}_{h,*}\|^2_{L^\infty(\hat{\Gamma}^m_{h,*})} + \tau)\|\hat{e}^{m+1}_h\|_{L^\infty(\hat{\Gamma}^m_{h,*})} \\
&\le C_{\kappa_m}(\|I_h T^{m+1}_*(\hat{X}^{m+1}_{h,*} - \hat{X}^m_{h,*})\|_{L^\infty(\hat{\Gamma}^m_{h,*})} + \|\hat{X}^{m+1}_{h,*} \\
&\quad - \hat{X}^m_{h,*}\|^2_{L^\infty(\hat{\Gamma}^m_{h,*})} + \tau)\|\hat{e}^{m+1}_h\|_{L^\infty(\hat{\Gamma}^m_{h,*})}, \qquad \text{(A.25)}
\end{aligned}
$$

where the derivation of the second to last inequality of (A.25) uses the following two arguments:

(i) We have used the Taylor expansion of $n^{m+1}_* \circ \hat{X}^{m+1}_{h,*} - n^{m+1}_* \circ X^{m+1}_{h,*}$ at $\hat{X}^{m+1}_{h,*}$ up to the quadratic term, with the following observation: Since both $\hat{X}^{m+1}_{h,*}$ and $X^{m+1}_{h,*}$ take values on Γ^{m+1}, and the value of $[(\nabla_{\Gamma^{m+1}} n^{m+1}_*) \circ \hat{X}^{m+1}_{h,*}](\hat{X}^{m+1}_{h,*} - \hat{X}^m_{h,*})$ at a node only depends on the value of $T^{m+1}_*(\hat{X}^{m+1}_{h,*} - \hat{X}^m_{h,*})$ at the node, it follows that

$$
|I_h(\nabla_{\Gamma^{m+1}} n^{m+1}_*)(\hat{X}^{m+1}_{h,*} - \hat{X}^m_{h,*})| \le C_0|I_h T^{m+1}_*(\hat{X}^{m+1}_{h,*} - \hat{X}^m_{h,*})| \quad \text{at the nodes.}
$$

(ii) The value of $n^{m+1}_* \circ X^{m+1}_{h,*} - n^m_* \circ \hat{X}^m_{h,*}$ at a node is the change of the normal vector along a particle trajectory of length $O(\tau)$.

The last inequality of (A.25) follows from the triangle inequality and the property $|X_{h,*}^{m+1} - \hat{X}_{h,*}^m| = O(\tau)$ at the nodes, because the value of $X_{h,*}^{m+1} - \hat{X}_{h,*}^m$ at a node is the distance a particle moves within time period τ.

Then, substituting (A.21) into (A.25) and using the result $\|\hat{X}_{h,*}^{m+1} - \hat{X}_{h,*}^m\|_{L^\infty(\hat{\Gamma}_{h,*}^m)} \le 1$ shown in (A.22), we get

$$
\begin{aligned}
&\|I_h T_*^m (N_*^{m+1} - N_*^m)\hat{e}_h^{m+1}\|_{L^\infty(\hat{\Gamma}_{h,*}^m)} \\
&\le C_{\kappa_m}(\|I_h T_*^{m+1}(\hat{X}_{h,*}^{m+1} - \hat{X}_{h,*}^m)\|_{L^\infty(\hat{\Gamma}_{h,*}^m)} + \tau)\|\hat{e}_h^{m+1}\|_{L^\infty(\hat{\Gamma}_{h,*}^m)} \\
&\le C_{\kappa_m} h^{0.5}\|I_h T_*^{m+1}(\hat{X}_{h,*}^{m+1} - \hat{X}_{h,*}^m)\|_{L^\infty(\hat{\Gamma}_{h,*}^m)} + C_{\kappa_m}\tau h^{-1}\|\hat{e}_h^{m+1}\|_{L^2(\hat{\Gamma}_{h,*}^m)} \\
&\le C_{\kappa_m} h^{0.5}\|I_h T_*^m(\hat{X}_{h,*}^{m+1} - \hat{X}_{h,*}^m)\|_{L^\infty(\hat{\Gamma}_{h,*}^m)} + C_{\kappa_m}(1+\kappa_{*,m})\tau h^{-1}(\tau + h^k),
\end{aligned}
\tag{A.26}
$$

where the second to last inequality follows from the estimate $\|\hat{e}_h^{m+1}\|_{L^\infty(\hat{\Gamma}_{h,*}^m)} \le C_{\kappa_m} h^{0.5}$ in (5.38) and the inverse inequality, and the last inequality follows from the error estimate of $\|\hat{e}_h^{m+1}\|_{L^2(\hat{\Gamma}_{h,*}^m)}$ in (5.74) and from replacing T_*^{m+1} by T_*^m with an error of $O(\tau)$ at the nodes. Note that we have replaced κ_l and $\kappa_{*,l}$ by κ_m and $\kappa_{*,m}$, respectively, when we estimate $\|\hat{e}_h^{m+1}\|_{L^\infty(\hat{\Gamma}_{h,*}^m)}$ and $\|\hat{e}_h^{m+1}\|_{L^2(\hat{\Gamma}_{h,*}^m)}$ using (5.38) and (5.74). This is correct as the estimation of $\hat{e}_h^{m+1}$ only requires using κ_m and $\kappa_{*,m}$ instead of κ_l and $\kappa_{*,l}$ (unless we want to consider the maximum error among $m = 0, \dots, l$ as in (5.74)). We also note that the error estimate of $\|\hat{e}_h^{m+1}\|_{L^2(\hat{\Gamma}_{h,*}^m)}$ using (5.74) requires the mesh size to satisfy $h \le h_{\kappa_m,\kappa_{*,m}}$ for some constant $h_{\kappa_m,\kappa_{*,m}}$ which may depend on κ_m and $\kappa_{*,m}$. Now we can substitute (A.24) and (A.26) into (A.20). This yields the following estimate:

$$
\begin{aligned}
&\|I_h T_*^m(\hat{X}_{h,*}^{m+1} - \hat{X}_{h,*}^m)\|_{L^\infty(\Gamma_{h,\mathrm{f}}^0)} \\
&\le \|I_h T_*^m(X_h^{m+1} - X_h^m)\|_{L^\infty(\Gamma_{h,\mathrm{f}}^0)} + \|I_h T_*^m(N_*^{m+1} - N_*^m)\hat{e}_h^{m+1}\|_{L^\infty(\Gamma_{h,\mathrm{f}}^0)} \\
&\le C_{\kappa_m} h^{-1}\big(\|T_*^m(e_h^{m+1} - \hat{e}_h^m)\|_{L^2(\hat{\Gamma}_{h,*}^m)} + h\|e_h^{m+1} - \hat{e}_h^m\|_{L^2(\hat{\Gamma}_{h,*}^m)}\big) + C_0\tau^2 \\
&\quad + C_{\kappa_m} h^{0.5}\|I_h T_*^m(\hat{X}_{h,*}^{m+1} - \hat{X}_{h,*}^m)\|_{L^\infty(\hat{\Gamma}_{h,*}^m)} + C_{\kappa_m}(1+\kappa_{*,m})\tau h^{-1}(\tau + h^k).
\end{aligned}
\tag{A.27}
$$

The second to last term on the right-hand side of (A.27) can be absorbed by the left-hand side by choosing sufficiently small h, say $h \le h_{\kappa_m,\kappa_{*,m}}$ for some constant $h_{\kappa_m,\kappa_{*,m}}$ which may depend on κ_m and $\kappa_{*,m}$. Then it holds that

$$
\begin{aligned}
&\|I_h T_*^m(\hat{X}_{h,*}^{m+1} - \hat{X}_{h,*}^m)\|_{L^\infty(\Gamma_{h,\mathrm{f}}^0)} \\
&\le C_{\kappa_m} h^{-1}\big(\|T_*^m(e_h^{m+1} - \hat{e}_h^m)\|_{L^2(\hat{\Gamma}_{h,*}^m)} + h\|e_h^{m+1} - \hat{e}_h^m\|_{L^2(\hat{\Gamma}_{h,*}^m)}\big) \\
&\quad + C_0\tau^2 + C_{\kappa_m}(1+\kappa_{*,m})\tau h^{-1}(\tau + h^k).
\end{aligned}
\tag{A.28}
$$

By substituting (A.28) into (A.23) we obtain the following result for $j \geq 1$:

$$\begin{aligned}
&\|\hat{X}_{h,*}^{m+1} - \hat{X}_{h,*}^{m}\|_{W_h^{j,\infty}(\Gamma_{h,\mathrm{f}}^0)} \\
&\leq C_0\tau[1 + (j-1)\|\hat{X}_{h,*}^{m}\|^{j}_{W_h^{j-1,\infty}(\Gamma_{h,\mathrm{f}}^0)}] + C_0\tau\|\hat{X}_{h,*}^{m}\|_{W_h^{j,\infty}(\Gamma_{h,\mathrm{f}}^0)} \\
&\quad + C_0 h^{-j}\tau^2 + C_{\kappa_m}(1+\kappa_{*,m})\tau h^{-j-1}(\tau + h^k) \\
&\quad + C_{\kappa_m} h^{-j-1}\big(\|T_*^m(e_h^{m+1} - \hat{e}_h^m)\|_{L^2(\hat{\Gamma}_{h,*}^m)} + h\|e_h^{m+1} - \hat{e}_h^m\|_{L^2(\hat{\Gamma}_{h,*}^m)}\big).
\end{aligned} \tag{A.29}$$

This proves the relation in (A.3) under the stepsize condition $\tau \leq ch^k$. The proof of (A.4) is similar (only the norm is changed) and omitted.

References

1. G. Bai and B. Li. Erratum: Convergence of Dziuk's semidiscrete finite element method for mean curvature flow of closed surfaces with high-order finite elements. *SIAM J. Numer. Anal.*, 61(3):1609–1612, 2023.
2. E. Bänsch, P. Morin, and R. H. Nochetto. A finite element method for surface diffusion: The parametric case. *J. Comput. Phys.*, 203:321–343, 2005.
3. 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.
4. W. Bao, W. Jiang, and Q. Zhao. A parametric finite element method for solid-state dewetting problems in three dimensions. *SIAM J. Sci. Comput.*, 42:B327–B352, 2020.
5. J. Barrett, K. Deckelnick, and R. Nürnberg. A finite element error analysis for axisymmetric mean curvature flow. *IMA J. Numer. Anal.*, 41(3):1641–1667, 2021.
6. J. W. Barrett, K. Deckelnick, and V. Styles. Numerical analysis for a system coupling curve evolution to reaction diffusion on the curve. *SIAM J. Numer. Anal.*, 55(2):1080–1100, 2017.
7. J. W. Barrett, H. Garcke, and R. Nürnberg. A parametric finite element method for fourth order geometric evolution equations. *J. Comput. Phys.*, 222:441–467, 2007.
8. J. W. Barrett, H. Garcke, and R. Nürnberg. On the parametric finite element approximation of evolving hypersurfaces in $\mathbb{R}^3$. *J. Comput. Phys.*, 227:4281–4307, 2008.
9. J. W. Barrett, H. Garcke, and R. Nürnberg. Parametric approximation of willmore flow and related geometric evolution equations. *SIAM Journal on Scientific Computing*, 31(1):225–253, 2008.
10. 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. Elsevier, 2020.
11. S. Bartels. A simple scheme for the approximation of the elastic flow of inextensible curves. *IMA J. Numer. Anal.*, 33:1115–1125, 2013.
12. S. Bartels, R. Müller, and C. Ortner. Robust a priori and a posteriori error analysis for the approximation of Allen–Cahn and Ginzburg–Landau equations past topological changes. *SIAM J. Numer. Anal.*, 49:110–134, 2011.
13. T. Binz and B. Kovács. A convergent finite element algorithm for generalized mean curvature flows of closed surfaces. *IMA J. Numer. Anal.*, 42(3):2545–2588, 2021.
14. A. Bonito, R. H. Nochetto, and M. S. Pauletti. Parametric FEM for geometric biomembranes. *J. Comput. Phys.*, 229:3171–3188, 2010.
15. K. Deckelnick. Error bounds for a difference scheme approximating viscosity solutions of mean curvature flow. *Interfaces Free Bound.*, 2:117–142, 2000.
16. K. Deckelnick and G. Dziuk. Convergence of a finite element method for non-parametric mean curvature flow. *Numer. Math.*, 72:197–222, 1995.

17. K. Deckelnick and G. Dziuk. On the approximation of the curve shortening flow. In *Calculus of variations, applications and computations (Pont-à-Mousson, 1994)*, volume 326 of *Pitman Res. Notes Math. Ser.*, pages 100–108. Longman Sci. Tech., Harlow, 1995.
18. K. Deckelnick and G. Dziuk. Error analysis of a finite element method for the Willmore flow of graphs. *Interfaces Free Bound.*, 8:21–46, 2006.
19. K. Deckelnick and G. Dziuk. Error analysis for the elastic flow of parametrized curves. *Math. Comp.*, 78:645–671, 2009.
20. K. Deckelnick and R. Nürnberg. Error analysis for a finite difference scheme for axisymmetric mean curvature flow of genus-0 surfaces. *SIAM J. Numer. Anal.*, 59(5):2698–2721, 2021.
21. K. Deckelnick and V. Styles. Finite element error analysis for a system coupling surface evolution to diffusion on the surface. *Interfaces Free Bound.*, 24:63–93, 2022.
22. A. Demlow. Higher-order finite element methods and pointwise error estimates for elliptic problems on surfaces. *SIAM J. Numer. Anal.*, 47(2):805–827, 2009.
23. G. Dziuk. An algorithm for evolutionary surfaces. *Numer. Math.*, 58:603–611, 1990.
24. G. Dziuk. Convergence of a semi-discrete scheme for the curve shortening flow. *Math. Models Methods Appl. Sci.*, 4:589–606, 1994.
25. G. Dziuk. Computational parametric Willmore flow. *Numer. Math.*, 111:55–80, 2008.
26. G. Dziuk and C. M. Elliott. Finite elements on evolving surfaces. *IMA J. Numer. Anal.*, 27:262–292, 2007.
27. G. Dziuk and C. M. Elliott. A fully discrete evolving surface finite element method. *SIAM J. Numer. Anal.*, 50:2677–2694, 2012.
28. G. Dziuk, D. Kröner, and T. Müller. Scalar conservation laws on moving hypersurfaces. *Interfaces Free Bound.*, 15(2):203–236, 2013.
29. K. Ecker. *Regularity theory for mean curvature flow*. Springer, 2012.
30. C. Elliott, H. Garcke, and B. Kovács. Numerical analysis for the interaction of mean curvature flow and diffusion on closed surfaces. *Numer. Math.*, 151:873–925, 2022.
31. C. M. Elliott and H. Fritz. On approximations of the curve shortening flow and of the mean curvature flow based on the DeTurck trick. *IMA J. Numer. Anal.*, 37:543–603, 2017.
32. X. Feng and Y. Li. Analysis of symmetric interior penalty discontinuous Galerkin methods for the Allen–Cahn equation and the mean curvature flow. *IMA J. Numer. Anal.*, 35:1622–1651, 2015.
33. X. Feng and A. Prohl. Numerical analysis of the Allen–Cahn equation and approximation for mean curvature flows. *Numer. Math.*, 94:33–65, 2003.
34. D. Gilbarg and N. S. Trudinger. *Elliptic partial differential equations of second order*. Springer, Berlin, Germany, 2001.
35. J. Hu and B. Li. Evolving finite element methods with an artificial tangential velocity for mean curvature flow and willmore flow. *Numer. Math.*, 152:127–181, 2022.
36. B. Kovács. High-order evolving surface finite element method for parabolic problems on evolving surfaces. *IMA J. Numer. Anal.*, 38(1):430–459, 2018.
37. B. Kovács, B. Li, and C. Lubich. A convergent evolving finite element algorithm for mean curvature flow of closed surfaces. *Numer. Math.*, 143:797–853, 2019.
38. B. Kovács, B. Li, and C. Lubich. A convergent evolving finite element algorithm for Willmore flow of closed surfaces. *Numer. Math.*, 149:595–643, 2021.
39. B. Kovács, B. Li, C. Lubich, and C. A. P. Guerra. Convergence of finite elements on an evolving surface driven by diffusion on the surface. *Numer. Math.*, 137:643–689, 2017.
40. B. Li. Convergence of Dziuk's linearly implicit parametric finite element method for curve shortening flow. *SIAM J. Numer. Anal.*, 58:2315–2333, 2020.
41. B. Li. Convergence of Dziuk's semidiscrete finite element method for mean curvature flow of closed surfaces with high-order finite elements. *SIAM J. Numer. Anal.*, 59:1592–1617, 2021.
42. C. Mantegazza. *Lecture Notes on Mean Curvature Flow*. . Basel AG, 2012.
43. A. Mierswa. Error estimates for a finite difference approximation of mean curvature flow for surfaces of torus type, PhD Thesis, Otto-von-Guericke-Universität, Magdeburg, 2020.
44. B. White. Evolution of curves and surfaces by mean curvature. *Proceedings of the International Congress of Mathematicians*, 1:525–538, 2002.
45. C. Ye and J. Cui. Convergence of Dziuk's fully discrete linearly implicit scheme for curve shortening flow. *SIAM J. Numer. Anal.*, 59:2823–2842, 2021.
46. 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:B327–B352, 2020.