SIAM J. NUMER. ANAL.
Vol. 63, No. 4, pp. 1454–1481

DYNAMIC RITZ PROJECTION OF MEAN CURVATURE FLOW AND OPTIMAL L^2 CONVERGENCE OF PARAMETRIC FEM*

BUYANG LI† AND RONG TANG‡

Abstract. A new approach is developed to study the convergence of parametric finite element approximations to the mean curvature flow of closed surfaces in three-dimensional space. In this approach, the error analysis is conducted by comparing the numerical solution to a dynamic Ritz projection of the mean curvature flow introduced in this paper rather than an interpolation of the mean curvature flow, as commonly used in the literature. The errors associated with the dynamic Ritz projection in approximating the mean curvature flow are established in the L^2 and $W^{1,p}$ norms. Leveraging these results, optimal-order convergence of parametric finite element methods for mean curvature flow of closed surfaces in the $L^\infty(0,T;L^2)$ norm is proved, including the convergence of parametric finite element methods with piecewise linear finite elements.

Key words. surface evolution, mean curvature flow, parametric finite element method, dynamic Ritz projection, convergence

MSC codes. 65M15, 65M60, 49M10, 35K65

DOI. 10.1137/24M1689053

1. Introduction.

The numerical approximation of surface evolution under geometric flows, including mean curvature flow, Willmore flow, and surface diffusion, has been intensively studied in the past. The most well-known example of geometric flows is mean curvature flow, which describes the evolution of a surface $\Gamma(t) \subset \mathbb{R}^3$, with certain initial condition $\Gamma(0) = \Gamma^0$, moving with velocity

$$v = -Hn,$$

where H and n denote the mean curvature and the normal vector of the surface. By using identity $Hn = -\Delta_\Gamma \mathrm{id}_\Gamma$, where Δ_Γ and id_Γ denote the Laplace–Beltrami operator and identity map on surface Γ, mean curvature flow can also be written as

$$v = \Delta_\Gamma \mathrm{id}_\Gamma. \tag{1.1}$$

By utilizing the formulation in (1.1), Dziuk introduced the following type of finite element method (FEM) in [21] for approximating surface evolution under mean curvature flow: Assuming that $\Gamma(t_{m-1})$ is already approximated by a piecewise triangular surface Γ_h^{m-1}, find a parametrization of surface Γ_h^m through a finite element function $u_h^m : \Gamma_h^{m-1} \to \mathbb{R}^3$, which is determined by some weak formulation of (1.1). Such methods are referred to as parametric FEMs.

*Received by the editors August 29, 2024; accepted for publication (in revised form) March 26, 2025; published electronically July 14, 2025.
https://doi.org/10.1137/24M1689053
Funding: The research was supported in part by the AMSS-PolyU Joint Laboratory; the Research Grants Council of the Hong Kong Special Administrative Region, China (projects PolyU/RFS2324-5S03 and PolyU/GRF15303022); and an internal grant of The Hong Kong Polytechnic University (project ID P0051154).
†Department of Applied Mathematics, The Hong Kong Polytechnic University, Hong Kong (buyang.li@polyu.edu.hk).
‡Corresponding author. Department of Applied Mathematics, The Hong Kong Polytechnic University, Hong Kong (claire.tang@polyu.edu.hk).

Since the 1990s, parametric FEMs have been widely used for approximating surface evolution under various geometric flows and interface evolution in various related problems. Many novel numerical methods were developed to address the challenges (such as prevention of mesh distortion and preservation of energy stability) in approximating surface evolution; see the artificial tangential motion constructed by Barrett, Garcke, and Nürnberg [6, 7, 8], Elliott and Fritz [26, 27], Hu and Li [29], and Duan and Li [20] and the structure-preserving parametric FEMs [42, 3]. These techniques have significantly improved the performance of parametric FEMs in approximating surface evolution under geometric flows. However, proving convergence of these methods remains challenging.

Convergence of parametric FEMs for geometric flows has been addressed for mean curvature flow and Willmore flow of curves in [9, 16, 15, 22, 27, 35, 41] and for graph surfaces and axisymmetric surfaces in [4, 12, 14, 17, 18]. However, the techniques developed in the proofs are not applicable to the analysis of geometric flow of general closed surfaces. We also refer the reader to [4, 5, 17, 18, 27, 39] for the proofs of convergence of parametric FEMs for mean curvature flow of closed curves, graph surfaces, axisymmetric surfaces, and surfaces of torus type, with Elliott and Fritz's tangential motion [26, 27] generated by a tangential transformation on Γ^0.

Convergence of the parametric FEMs for mean curvature flow, Willmore flow, and surface diffusion of closed surfaces was first proved in [11, 25, 29, 31, 32] for some equivalent formulations of the geometric flows which couple the velocity equation of geometric flows with the geometric evolution equations of mean curvature H and normal vector n by formulating the algorithms into evolving FEMs as in [23] and utilizing the matrix-vector formulation of evolving FEMs introduced in [33]. For example, convergence of parametric FEMs for mean curvature flow of closed surfaces is proved for the following equivalent formulation of mean curvature flow:

$$\begin{aligned} \partial_t X &= v \circ X && \text{on } \Gamma^0, \\ v &= -Hn && \text{on } \Gamma(t), \\ \partial_t^\bullet H - \Delta_{\Gamma(t)} H &= |\nabla_{\Gamma(t)} n|^2 H && \text{on } \Gamma(t), \\ \partial_t^\bullet n - \Delta_{\Gamma(t)} n &= |\nabla_{\Gamma(t)} n|^2 n && \text{on } \Gamma(t), \end{aligned} \tag{1.2}$$

where $X(\cdot,t) : \Gamma^0 \to \mathbb{R}^3$ is the flow map which determines $\Gamma(t) = \{X(p,t) : p \in \Gamma^0\}$ and $\partial_t^\bullet$ denotes material derivative along the particle trajectories of the flow map, i.e.,

$$\partial_t^\bullet u(x,t) = \frac{\mathrm{d}}{\mathrm{d}t} u(X(p,t),t) \text{ at point } x = X(p,t) \text{ on } \Gamma(t).$$

The analyses in [29, 31, 32] are restricted to finite elements of degree $k \geqslant 2$. This condition is required for proving convergence of numerical solutions in the $W^{1,\infty}$ norm in order to control the nonlinear terms and to utilize the equivalence of the L^p and $W^{1,p}$ norms of functions on the numerical-solution surface and interpolated surface.

The convergence of other parametric FEMs, designed for approximating (1.1) instead of (1.2), is more challenging due to the degeneracy of the nonlinear Laplace–Beltrami operator $\Delta_\Gamma \mathrm{id}_\Gamma$ acting on surface Γ; see the discussions in [2, 36]. In the spatially semidiscrete setting, the convergence of parametric FEMs for mean curvature flow of surfaces with formulation (1.1) was proved in [1, 36] for finite elements of degree $k \geqslant 6$ based on a discovery that the nonlinear Laplace–Beltrami operator is H^1 elliptic with respect to the normal component of the trajectory error (error between the exact and numerical flow maps). The restriction to finite elements of degree $k \geqslant 6$ is needed to control the nonlinearities in error analysis by using the very weak estimates obtained based on the partial H^1 ellipticity in the normal direction. In the fully discrete setting, the convergence of Dziuk's semi-implicit parametric FEM for mean curvature flow of surfaces was proved in [2] for finite elements of degree $k \geqslant 3$ based on a discovery that the nonlinear Laplace–Beltrami operator is H^1 elliptic with respect to the distance error (distance between the exact and numerical surfaces)—the strong H^1 ellipticity in both normal and tangential directions leads to stronger estimates of the errors for controlling the nonlinearities and therefore reduces the requirements of finite elements of degrees $k \geqslant 6$–$k \geqslant 3$.

In summary, the convergence of some fundamental algorithms for geometric flows of surfaces still remains open. The existing proofs of convergence of parametric FEMs for mean curvature flow and other geometric flows of closed surfaces are all based on optimal-order H^1-norm error estimates that require using finite elements of degree $k \geqslant 2$ to control the $W^{1,\infty}$ boundedness of numerical solutions of surface position X, mean curvature H, and normal vector n. The following two questions remain open:

- Convergence of parametric FEMs for mean curvature flow and other geometric flows of closed surfaces with piecewise linear finite elements still remains open.
- Optimal-order convergence of parametric FEMs in the $L^\infty(0,T;L^2)$ norm for these geometric flows remains open.

The two questions are addressed simultaneously in the current paper for approximating formulation (1.2) of mean curvature flow. It turns out that the two questions are closely related such that our answer to the second question (by introducing a dynamic Ritz projection which reduces the remainders in the error equations) also addresses the first question. In particular, the new framework developed in this paper, by defining and utilizing a dynamic Ritz projection of geometric flow in error analysis, is promising for proving convergence of parametric FEMs for geometric flows with optimal-order convergence and lower-degree finite elements.

2. Dynamic Ritz projection and main results.

Let Γ_h^0 be a piecewise polynomial surface that interpolates the smooth surface Γ^0, with each piece being the image of the reference triangle under a polynomial map of degree $k \geqslant 1$, and assume that the curved triangles are shape-regular and quasi-uniform with mesh size h; see [19, 30] and Lenoir's isoparametric approximation of a surface [34].

Let $\mathbf{x}^0 = (p_1, \cdots, p_N) \in \mathbb{R}^{3N}$ be the nodal vector that collects all the nodes $p_j \in \mathbb{R}^3$, $j = 1, \ldots, N$, in Γ_h^0. We evolve $\mathbf{x}^0$ in time and denote its position at time t by $\mathbf{x}(t) = (x_1(t), \cdots, x_N(t))$, which determines a piecewise (possibly curved) triangular surface $\Gamma_h[\mathbf{x}(t)]$ via piecewise polynomial interpolation on a plane reference triangle, and denote by $S_h[\mathbf{x}(t)]$ the finite element space of polynomial degree k on the piecewise triangular surface $\Gamma_h[\mathbf{x}(t)]$.

There exists a unique finite element function $X_h(\cdot, t)$ of piecewise polynomial degree k defined on $\Gamma_h[\mathbf{x}^0]$ satisfying

$$X_h(p_j, t) = x_j(t) \quad \text{for} \quad j = 1, \ldots, N.$$

This is the discrete flow map which maps $\Gamma_h[\mathbf{x}^0]$ to $\Gamma_h[\mathbf{x}(t)]$. The semidiscrete parametric FEM for (1.2) is to find

$$(X_h(\cdot,t), v_h(\cdot,t), H_h(\cdot,t), n_h(\cdot,t)) \in S_h[\mathbf{x}^0]^3 \times S_h[\mathbf{x}(t)]^3 \times S_h[\mathbf{x}(t)] \times S_h[\mathbf{x}(t)]^3$$

such that the following weak formulation holds for all $(\chi_H, \chi_n) \in S_h(\Gamma_h[\mathbf{x}]) \times S_h(\Gamma_h[\mathbf{x}])^3$:

$$\partial_t X_h = v_h \circ X_h \quad \text{on } \Gamma_h[\mathbf{x}^0], \tag{2.1a}$$
$$v_h = -I_h(H_h n_h) \quad \text{on } \Gamma_h[\mathbf{x}], \tag{2.1b}$$
$$\int_{\Gamma_h[\mathbf{x}]} \partial^\bullet_{t,h} H_h \chi_H + \int_{\Gamma_h[\mathbf{x}]} \nabla_{\Gamma_h[\mathbf{x}]} H_h \cdot \nabla_{\Gamma_h[\mathbf{x}]} \chi_H = \int_{\Gamma_h[\mathbf{x}]} |\nabla_{\Gamma_h[\mathbf{x}]} n_h|^2 H_h \chi_H, \tag{2.1c}$$
$$\int_{\Gamma_h[\mathbf{x}]} \partial^\bullet_{t,h} n_h \cdot \chi_n + \int_{\Gamma_h[\mathbf{x}]} \nabla_{\Gamma_h[\mathbf{x}]} n_h \cdot \nabla_{\Gamma_h[\mathbf{x}]} \chi_n = \int_{\Gamma_h[\mathbf{x}]} |\nabla_{\Gamma_h[\mathbf{x}]} n_h|^2 n_h \cdot \chi_n, \tag{2.1d}$$

where I_h denotes the Lagrange interpolation onto the finite element space $S_h[\mathbf{x}(t)]$ and $\partial^\bullet_{t,h}$ denotes the material derivative on $\Gamma_h[\mathbf{x}(t)]$ with respect to the discrete flow map $X_h(\cdot,t)$. The initial value for (2.1) can be chosen as follows: $X_h(\cdot,0) = \mathrm{id}$ on $\Gamma_h[\mathbf{x}^0]$; $H_h(\cdot,0)$ and $n_h(\cdot,0)$ are the Lagrange interpolations of $H(\cdot,0)$ and $n(\cdot,0)$, respectively.

Let $\mathbf{x}^* = \mathbf{x}^*(t)$ be the nodal vector which collects the nodes evolving according to the exact flow map $X(\cdot,t) : \Gamma^0 \to \mathbb{R}^3$, and denote by $\Gamma_h[\mathbf{x}^*]$ the piecewise curved triangular surface that interpolates $\Gamma(t)$ at the nodes in $\mathbf{x}^*$. The finite element space on $\Gamma_h[\mathbf{x}^*]$ is denoted by $S_h(\Gamma_h[\mathbf{x}^*])$.

The analyses in [29, 31, 32] are based on estimating the error between numerical solution (X_h, v_h, H_h, n_h) and $(\hat I_h^* X, \hat I_h^* v, \hat R_h^* H, \hat R_h^* n)$, where $\hat I_h^* X$ and $\hat I_h^* v$ are Lagrange interpolations of X and v onto $\Gamma_h[\mathbf{x}^0]$ and $\Gamma_h[\mathbf{x}^*]$, respectively, and $\hat R_h^* H, \hat R_h^* n \in S_h(\Gamma_h[\mathbf{x}^*])$ are the linear Ritz projections of H, n onto surface $\Gamma_h[\mathbf{x}^*]$, respectively, defined by

$$\int_{\Gamma_h[\mathbf{x}^*]} (\hat R_h^* H\, \varphi_h + \nabla_{\Gamma_h[\mathbf{x}^*]} \hat R_h^* H \cdot \nabla_{\Gamma_h[\mathbf{x}^*]} \varphi_h) = \int_\Gamma (H \varphi_h^l + \nabla_\Gamma H \cdot \nabla_\Gamma \varphi_h^l) \quad \forall \varphi_h \in S_h(\Gamma_h[\mathbf{x}^*]), \tag{2.2a}$$
$$\int_{\Gamma_h[\mathbf{x}^*]} (\hat R_h^* n \cdot \phi_h + \nabla_{\Gamma_h[\mathbf{x}^*]} \hat R_h^* n \cdot \nabla_{\Gamma_h[\mathbf{x}^*]} \phi_h) = \int_\Gamma (n \cdot \phi_h^l + \nabla_\Gamma n \cdot \nabla_\Gamma \phi_h^l) \quad \forall \phi_h \in S_h(\Gamma_h[\mathbf{x}^*])^3. \tag{2.2b}$$

In (2.2), φ_h^l and ϕ_h^l denote the lifts of functions $\varphi_h \in S_h(\Gamma_h[\mathbf{x}^*])$ and $\phi_h \in S_h(\Gamma_h[\mathbf{x}^*])^3$ to the exact surface $\Gamma = \Gamma(t)$, respectively; see section 3.1. By this definition of Ritz projection, $(\hat I_h^* X, \hat I_h^* v, \hat R_h^* H, \hat R_h^* n)$ satisfies numerical scheme (2.1) up to some remainders, i.e.,

$$\partial_t \hat I_h^* X = \hat I_h^* v \circ \hat I_h^* X \quad \text{on } \Gamma_h[\mathbf{x}^0], \tag{2.3a}$$
$$\hat I_h^* v = -\hat I_h^*(\hat R_h^* H\, \hat R_h^* n) + d_v \quad \text{on } \Gamma_h[\mathbf{x}^*], \tag{2.3b}$$
$$\begin{aligned}\int_{\Gamma_h[\mathbf{x}^*]} \partial^\bullet_{t,h} \hat R_h^* H \chi_H + \int_{\Gamma_h[\mathbf{x}^*]} \nabla_{\Gamma_h[\mathbf{x}^*]} \hat R_h^* H \cdot \nabla_{\Gamma_h[\mathbf{x}^*]} \chi_H \quad \forall \chi_H \in S_h(\Gamma_h[\mathbf{x}^*])\\ = \int_{\Gamma_h[\mathbf{x}^*]} |\nabla_{\Gamma_h[\mathbf{x}^*]} \hat R_h^* n|^2 \hat R_h^* H \chi_H + \int_{\Gamma_h[\mathbf{x}^*]} d_H \chi_H,\end{aligned} \tag{2.3c}$$
$$\begin{aligned}\int_{\Gamma_h[\mathbf{x}^*]} \partial^\bullet_{t,h} \hat R_h^* n \cdot \chi_n + \int_{\Gamma_h[\mathbf{x}^*]} \nabla_{\Gamma_h[\mathbf{x}^*]} \hat R_h^* n \cdot \nabla_{\Gamma_h[\mathbf{x}^*]} \chi_n \quad \forall \chi_n \in S_h(\Gamma_h[\mathbf{x}^*])^3\\ = \int_{\Gamma_h[\mathbf{x}^*]} |\nabla_{\Gamma_h[\mathbf{x}^*]} \hat R_h^* n|^2 \hat R_h^* n \cdot \chi_n + \int_{\Gamma_h[\mathbf{x}^*]} d_n \cdot \chi_n,\end{aligned} \tag{2.3d}$$

where $d_v = \hat I_h^*(\hat R_h^* H\, \hat R_h^* n) - \hat I_h^*(\hat I_h^* H\, \hat I_h^* n)$ and d_H and d_n are remainders which satisfy the following estimates (for some constant C which is independent of mesh size h):

$$\|d_v\|_{L^2(\Gamma_h[\mathbf{x}^*])} + h\|d_v\|_{H^1(\Gamma_h[\mathbf{x}^*])} \leqslant Ch^{k+1}, \tag{2.4}$$

$$\Big|\int_{\Gamma_h[\mathbf{x}^*]} d_H\chi_H\Big| + \Big|\int_{\Gamma_h[\mathbf{x}^*]} d_n\chi_n\Big| \leqslant Ch^{k+1}(\|\chi_H\|_{H^1(\Gamma_h[\mathbf{x}^*])} + \|\chi_n\|_{H^1(\Gamma_h[\mathbf{x}^*])}). \tag{2.5}$$

We denote by $\hat{x}_h$, $\hat{v}_h$, $\hat{H}_h$, and $\hat{n}_h$ the lift of X_h, v_h, H_h, and n_h to $\Gamma_h[\mathbf{x}^*]$, i.e., finite element functions on $\Gamma_h[\mathbf{x}^*]$ with the same nodal vectors as X_h, v_h, H_h, and n_h, respectively. By using this notation, we can define the following finite element error functions on $\Gamma_h[\mathbf{x}^*]$:

$$\begin{aligned} &\hat{e}_{x,h} = \hat{x}_h - \mathrm{id}_{\Gamma_h[\mathbf{x}^*]}, \quad \hat{e}_{v,h} = \hat{v}_h - \hat{I}_h^* v, \\ &\hat{e}_{H,h} = \hat{H}_h - \hat{R}_h^* H, \qquad \hat{e}_{n,h} = \hat{n}_h - \hat{R}_h^* n, \end{aligned} \tag{2.6}$$

where $\hat{e}_{x,h}$ represents the error between surfaces $\Gamma_h[\mathbf{x}]$ and $\Gamma_h[\mathbf{x}^*]$ and $\hat{e}_{v,h}$, $\hat{e}_{H,h}$, and $\hat{e}_{n,h}$ represent the errors in the numerical approximations of velocity, mean curvature, and normal vector, respectively. Such definitions of error functions are used in [29, 31, 32] in proving convergence of parametric FEMs for mean curvature flow and Willmore flow. However, the presence of remainder d_v in (2.3b) hinders people from proving optimal-order convergence in the $L^\infty(0,T;L^2)$ norm for the reason that $\|\hat{e}_{x,h}\|_{H^1(\Gamma_h[\mathbf{x}^*])}$ frequently appears due to surface location errors, and this needs to be controlled by using $\|d_v\|_{H^1(\Gamma_h[\mathbf{x}^*])}$ instead of $\|d_v\|_{L^2(\Gamma_h[\mathbf{x}^*])}$. Therefore, optimal-order convergence of parametric FEMs for mean curvature flow was only proved in the $L^\infty(0,T;H^1)$ norm in the literature, and such $L^\infty(0,T;H^1)$ error analysis typically requires $W^{1,\infty}$ boundedness of the numerical solutions in order to bound the nonlinear terms in the error analysis. This requires the convergence order to be sufficiently high in order to apply the inverse inequality of finite element functions to prove $W^{1,\infty}$ boundedness of the numerical solutions, and this limits the error analyses to high-order finite elements of degree $k \geqslant 2$.

Our solution to the above-mentioned two questions, i.e., optimal-order convergence in the $L^\infty(0,T;L^2)$ norm and convergence of parametric FEM for mean curvature flow with piecewise linear finite elements, is based on the following two observations:

1. The evolution equations of H and n in (1.2) have similar nonlinear structures as the harmonic map heat flow studied in [28], where the proof of optimal-order convergence of FEMs in the $L^\infty(0,T;L^2)$ norm only requires utilizing $W^{1,4}$ boundedness of $e_{H,h}$ and $e_{n,h}$ to bound the nonlinear terms appearing in the error analysis. This motivates us to consider $L^\infty(0,T;L^2)$ error estimates of parametric FEMs for (1.2) based on boundedness of $e_{H,h}$ and $e_{n,h}$ in a norm weaker than the $W^{1,\infty}$ norm instead of the $L^\infty(0,T;H^1)$ error estimates considered in [29, 31, 32]. The latter approach requires $W^{1,\infty}$ boundedness of $e_{H,h}$ and $e_{n,h}$ (this cannot be proved for piecewise linear finite elements so far) to bound the nonlinear terms in error analysis.
2. However, convergence of parametric FEMs for mean curvature flow of closed surfaces requires $W^{1,\infty}$ boundedness of $\hat{e}_{x,h}$ to guarantee the norm equivalence of finite element functions on $\Gamma_h[\mathbf{x}]$ and $\Gamma_h[\mathbf{x}^*]$ associated to a common nodal vector. Such norm equivalence is frequently used and can hardly be relaxed in analyzing the error of parametric FEMs for surface evolution under geometric flows. However, an optimal-order error estimate such as

$$\|\hat{e}_{x,h}\|_{L^\infty(0,T;L^2(\Gamma_h[\mathbf{x}^*]))} \leqslant Ch^{k+1} \tag{2.7}$$

is not enough to guarantee the $W^{1,\infty}$ boundedness of $\hat{e}_{x,h}$ in the case $k = 1$ (i.e., for piecewise linear finite elements)—we actually need $k > 1$

(quadratic or higher-order finite elements) to have some extra convergence rate in order to prove the $W^{1,\infty}$ boundedness of $\hat{e}_{x,h}$ from (2.7); see the arguments in [29, 31, 32].

This difficulty can be overcome if we can prove the following superconvergence rate in the $L^2(0,T;H^1)$ norm:

$$\|\hat{e}_{v,h}\|_{L^2(0,T;H^1(\Gamma_h[\mathbf{x}^*]))} \leqslant Ch^{k+1}. \tag{2.8}$$

This would imply the following result (position is the time integral of velocity):

$$\|\hat{e}_{x,h}\|_{L^\infty(0,T;H^1(\Gamma_h[\mathbf{x}^*]))} \leqslant Ch^{k+1}. \tag{2.9}$$

This would further imply (via the inverse inequality in two dimensions) that

$$\|\hat{e}_{x,h}\|_{L^\infty(0,T;W^{1,\infty}(\Gamma_h[\mathbf{x}^*]))} \leqslant Ch^{k}. \tag{2.10}$$

This can be used to prove the uniform $W^{1,\infty}$ boundedness of $\hat{e}_{x,h}$ for $k \geqslant 1$. However, the main difficulty of this approach is that (2.8) cannot be shown with the presence of remainder d_v in (2.3b). To overcome this difficulty, we *redefine* the error functions using a modified Ritz projection of the mean curvature flow that could exclude the remainder term d_v in (2.3b).

Our idea is to define a dynamic Ritz projection of (X, v, H, n) as the finite element solution $(Y_h^*, v_h^*, H_h^*, n_h^*)$ of a nonlinearly coupled surface evolution problem in (2.11), with $Y_h^*(\cdot,t) : \Gamma_h^0 \to \mathbb{R}^3$ being a finite element flow map with nodal vector $\mathbf{y}^* = \mathbf{y}^*(t)$ and v_h^*, H_h^*, and n_h^* being functions on $\Gamma_h[\mathbf{y}^*]$ determined by the equations

$$\begin{aligned} &\partial_t Y_h^* = v_h^* \circ Y_h^* && \text{on } \Gamma_h^0, \\ &v_h^* = -I_h^*(H_h^* n_h^*) && \text{on } \Gamma_h[\mathbf{y}^*], \\ &\int_{\Gamma_h[\mathbf{y}^*]} (H_h^* \varphi_h^* + \nabla_{\Gamma_h[\mathbf{y}^*]} H_h^* \cdot \nabla_{\Gamma_h[\mathbf{y}^*]} \varphi_h^*) = \int_\Gamma (H\varphi_h^l + \nabla_\Gamma H \cdot \nabla_\Gamma \varphi_h^l) && \forall\, \varphi_h \in S_h(\Gamma_h[\mathbf{x}^*]), \\ &\int_{\Gamma_h[\mathbf{y}^*]} (n_h^* \phi_h^* + \nabla_{\Gamma_h[\mathbf{y}^*]} n_h^* \cdot \nabla_{\Gamma_h[\mathbf{y}^*]} \phi_h^*) = \int_\Gamma (n\phi_h^l + \nabla_\Gamma n \cdot \nabla_\Gamma \phi_h^l) && \forall\, \phi_h \in S_h(\Gamma_h[\mathbf{x}^*])^3, \end{aligned} \tag{2.11}$$

where I_h^* denotes the Lagrange interpolation operator onto $S_h(\Gamma_h[\mathbf{y}^*])$; φ_h^l denotes the lift of a function $\varphi_h \in S_h(\Gamma_h[\mathbf{x}^*])$ to surface $\Gamma = \Gamma(t)$, i.e., $\varphi_h^l(x^l) = \varphi_h(x)$ for $x \in \Gamma_h[\mathbf{x}^*]$, with x^l denoting the lift of point x from $\Gamma_h[\mathbf{x}^*]$ to Γ; and φ_h^* denotes the finite element function on $\Gamma_h[\mathbf{y}^*]$ with the same nodal vector as $\varphi_h \in S_h(\Gamma_h[\mathbf{x}^*])$. The initial value for the system (2.11) is given by $Y_h^*(\cdot,0) = \mathrm{id}$ on Γ_h^0. In this definition, H_h^* and n_h^* are Ritz projections of H and n onto an unknown surface $\Gamma_h[\mathbf{y}^*]$ which evolves with velocity $v_h^* = -I_h^*(H_h^* n_h^*)$ determined by this Ritz projection.

Note that the idea of using Ritz projection to achieve H^1 superconvergence and subsequently obtain optimal L^2 error estimates was first introduced in [40] for a class of nonlinear parabolic equations. Nonlinear types of Ritz projections were employed to ensure uniform control over the gradient of the height function for the mean curvature flow, Willmore flow, and surface diffusion of graphs; see [13, 14, 18]. The dynamic Ritz projection introduced in this paper distinguishes itself from the classical Ritz projection primarily through modifications to the first two equations in (2.11). These alterations enable the surface $\Gamma_h[\,\mathbf{y}^*]$ to evolve according to an evolution equation (thus earning the name "dynamic"), thereby differentiating it from the interpolated surface $\Gamma_h[\mathbf{x}^*]$.

The first main result of this paper is the following theorem about optimal-order approximation properties for this dynamic Ritz projection of mean curvature flow.

THEOREM 2.1. *Assume that the exact solution (X, v, H, n) of the mean curvature flow is sufficiently smooth and that the flow map $X(\cdot, t) : \Gamma^0 \to \Gamma(t)$ is a diffeomorphism for $t \in [0, T]$. Let $(\hat{y}_h^*, \hat{v}_h^*, \hat{H}_h^*, \hat{n}_h^*)$ be the lift of the dynamic Ritz projection $(Y_h^*, v_h^*, H_h^*, n_h^*)$ defined in (2.11), with finite elements of degree $k \geqslant 1$. Then there exists a constant $h_0 > 0$ such that the following error bound holds for mesh size $h \leqslant h_0$:*

$$
\begin{aligned}
&\|\hat{y}_h^* - \mathrm{id}_{\Gamma_h[\mathbf{x}^*]}\|_{L^\infty(0,T;L^2(\Gamma_h[\mathbf{x}^*]))} + \|\hat{v}_h^* - \hat{I}_h^* v\|_{L^\infty(0,T;L^2(\Gamma_h[\mathbf{x}^*]))} \\
&\quad + \|\hat{H}_h^* - \hat{I}_h^* H\|_{L^\infty(0,T;L^2(\Gamma_h[\mathbf{x}^*]))} + \|\hat{n}_h^* - \hat{I}_h^* n\|_{L^\infty(0,T;L^2(\Gamma_h[\mathbf{x}^*]))} \\
&\quad + \|\partial_{t,h}^\bullet(\hat{H}_h^* - \hat{I}_h^* H)\|_{L^\infty(0,T;L^2(\Gamma_h[\mathbf{x}^*]))} + \|\partial_{t,h}^\bullet(\hat{n}_h^* - \hat{I}_h^* n)\|_{L^\infty(0,T;L^2(\Gamma_h[\mathbf{x}^*]))} \leqslant C h^{k+1},
\end{aligned}
\tag{2.12}
$$

where $\hat{I}_h^* X$ and $\hat{I}_h^* v$ denote the Lagrange interpolations of X and v onto $\Gamma_h[\mathbf{x}^0]$ and $\Gamma_h[\mathbf{x}^*]$, respectively. The constant C is independent of h and $t \in [0, T]$ (but may depend on T).

In view of the results in Theorem 2.1, we compare the numerical solution (X_h, v_h, H_h, n_h) with the dynamic Ritz projection $(Y_h^*, v_h^*, H_h^*, n_h^*)$, which satisfies (2.1) up to some remainders $d_H^* \in S_h[\mathbf{y}^*]$ and $d_n^* \in S_h^3[\mathbf{y}^*]$, i.e.,

$$
\partial_t Y_h^* = v_h^* \circ Y_h^* \qquad \text{on } \Gamma_h^0, \tag{2.13a}
$$

$$
v_h^* = -I_h^*(H_h^* n_h^*) \qquad \text{on } \Gamma_h[\mathbf{y}^*], \tag{2.13b}
$$

$$
\begin{aligned}
&\int_{\Gamma_h[\mathbf{y}^*]} \partial_{t,h}^\bullet H_h^* \chi_H + \int_{\Gamma_h[\mathbf{y}^*]} \nabla_{\Gamma_h[\mathbf{y}^*]} H_h^* \cdot \nabla_{\Gamma_h[\mathbf{y}^*]} \chi_H \qquad \forall \chi_H \in S_h(\Gamma_h[\mathbf{y}^*]) \\
&\qquad = \int_{\Gamma_h[\mathbf{y}^*]} |\nabla_{\Gamma_h[\mathbf{y}^*]} n_h^*|^2 H_h^* \chi_H + \int_{\Gamma_h[\mathbf{y}^*]} d_H^* \chi_H,
\end{aligned}
\tag{2.13c}
$$

$$
\begin{aligned}
&\int_{\Gamma_h[\mathbf{y}^*]} \partial_{t,h}^\bullet n_h^* \cdot \chi_n + \int_{\Gamma_h[\mathbf{y}^*]} \nabla_{\Gamma_h[\mathbf{y}^*]} n_h^* \cdot \nabla_{\Gamma_h[\mathbf{y}^*]} \chi_n \qquad \forall \chi_n \in S_h^3(\Gamma_h[\mathbf{y}^*]) \\
&\qquad = \int_{\Gamma_h[\mathbf{y}^*]} |\nabla_{\Gamma_h[\mathbf{y}^*]} n_h^*|^2 n_h^* \cdot \chi_n + \int_{\Gamma_h[\mathbf{y}^*]} d_n^* \cdot \chi_n.
\end{aligned}
\tag{2.13d}
$$

By introducing $u_h^* := (n_h^*, H_h^*)$, the weak formulation in (2.13) can be rewritten as follows:

$$
\partial_t Y_h^* = v_h^* \circ Y_h^* \qquad \text{on } \Gamma_h^0, \tag{2.14a}
$$

$$
v_h^* = -I_h^*(H_h^* n_h^*) \qquad \text{on } \Gamma_h[\mathbf{y}^*], \tag{2.14b}
$$

$$
\begin{aligned}
&\int_{\Gamma_h[\mathbf{y}^*]} \partial_{t,h}^\bullet u_h^* \cdot \chi_u + \int_{\Gamma_h[\mathbf{y}^*]} \nabla_{\Gamma_h[\mathbf{y}^*]} u_h^* \cdot \nabla_{\Gamma_h[\mathbf{y}^*]} \chi_u \qquad \forall \chi_u \in S_h^4(\Gamma_h[\mathbf{y}^*]) \\
&\qquad = \int_{\Gamma_h[\mathbf{y}^*]} |\nabla_{\Gamma_h[\mathbf{y}^*]} n_h^*|^2 u_h^* \cdot \chi_u + \int_{\Gamma_h[\mathbf{y}^*]} d_u^* \cdot \chi_u,
\end{aligned}
\tag{2.14c}
$$

where d_u^* is remainder which satisfies the following estimate (this can be shown by using the result of Theorem 2.1; see Lemma 3.6):

$$
\Big| \int_{\Gamma_h[\mathbf{y}^*]} d_u^* \cdot \chi_u \Big| \leqslant C h^{k+1} \|\chi_u\|_{H^1(\Gamma_h[\mathbf{y}^*])}.
$$

In particular, compared with (2.3b), no remainder appears in (2.14b). This makes it possible to prove the estimates in (2.8) and (2.9) by redefining the error functions as follows:

$$
\hat{e}_{x,h} := \hat{x}_h - \hat{y}_h^*, \quad \hat{e}_{v,h} := \hat{v}_h - \hat{v}_h^*, \quad \hat{e}_{H,h} := \hat{H}_h - \hat{H}_h^*, \quad \text{and} \quad \hat{e}_{n,h} := \hat{n}_h - \hat{n}_h^*,
\tag{2.15}
$$

where $(\hat{x}_h, \hat{v}_h, \hat{H}_h, \hat{n}_h)$ and $(\hat{y}_h^*, \hat{v}_h^*, \hat{H}_h^*, \hat{n}_h^*)$ are the lifts of (X_h, v_h, H_h, n_h) and $(Y_h^*, v_h^*, H_h^*, n_h^*)$, respectively, i.e., the numerical solution of mean curvature flow defined in (2.1) and the dynamic Ritz projection of mean curvature flow defined in (2.11), respectively. This leads to the second main result of this paper, which is presented in the following theorem.

THEOREM 2.2. *Under the assumptions of Theorem* 2.1*, the following error bound holds:*

$$\begin{aligned}&\|\hat{e}_{x,h}\|_{L^\infty(0,T;H^1(\Gamma_h[\mathbf{x}^*]))} + \|\hat{e}_{H,h}\|_{L^\infty(0,T;L^2(\Gamma_h[\mathbf{x}^*]))} + \|\hat{e}_{n,h}\|_{L^\infty(0,T;L^2(\Gamma_h[\mathbf{x}^*]))}\\ &\quad + \|\hat{e}_{v,h}\|_{L^2(0,T;H^1(\Gamma_h[\mathbf{x}^*]))} + \|\hat{e}_{H,h}\|_{L^2(0,T;H^1(\Gamma_h[\mathbf{x}^*]))} + \|\hat{e}_{n,h}\|_{L^2(0,T;H^1(\Gamma_h[\mathbf{x}^*]))}\\ &\leqslant Ch^{k+1},\end{aligned} \tag{2.16}$$

where C is a constant independent of h and $t \in [0,T]$ (but may depend on T).

The error estimates in Theorems 2.1 and 2.2 and the estimates of the Lagrange interpolation error lead to the third main result of this paper, i.e., optimal-order convergence of FEMs for mean curvature flow in the $L^\infty(0,T;L^2)$ norm with finite elements of degree $k \geqslant 1$. Before stating this result, we list some basic notations for finite element functions on $\Gamma_h[\mathbf{x}]$, $\Gamma_h[\mathbf{x}^*]$, and Γ.

For any given finite element function $w_h \in S_h(\Gamma_h[\mathbf{x}])$ on the approximate surface $\Gamma_h[\mathbf{x}]$, we denote its nodal vector by $\mathbf{w}$, which collects all the values of w_h at the nodes of $\Gamma_h[\mathbf{x}]$. The finite element function on the interpolated surface $\Gamma_h[\mathbf{x}^*]$ with the same nodal vector $\mathbf{w}$ is denoted by $\hat{w}_h$. The function $\hat{w}_h$ can be further lifted to $\Gamma[X]$ as $(\hat{w}_h)^l$; see details in section 3.1. The lift from $S_h(\Gamma_h[\mathbf{x}])$ to $\Gamma[X]$ is denoted by $w_h^L = (\hat{w}_h)^l$.

THEOREM 2.3. *Under the assumptions of Theorem* 2.1*, the numerical solution of mean curvature flow defined in* (2.1)*, with finite elements of degree $k \geqslant 1$, satisfies the following error bound:*

$$\begin{aligned}&\|\hat{x}_h - \mathrm{id}_{\Gamma_h[\mathbf{x}^*]}\|_{L^\infty(0,T;L^2(\Gamma_h[\mathbf{x}^*]))} + \|\hat{v}_h - \hat{I}_h^* v\|_{L^\infty(0,T;L^2(\Gamma_h[\mathbf{x}^*]))}\\ &\quad + \|\hat{H}_h - \hat{I}_h^* H\|_{L^\infty(0,T;L^2(\Gamma_h[\mathbf{x}^*]))} + \|\hat{n}_h - \hat{I}_h^* n\|_{L^\infty(0,T;L^2(\Gamma_h[\mathbf{x}^*]))} \leqslant Ch^{k+1},\end{aligned} \tag{2.17}$$

$$\begin{aligned}&\|X_h^L - \mathrm{id}\|_{L^\infty(0,T;L^2(\Gamma[X]))} + \|H_h^L - H\|_{L^\infty(0,T;L^2(\Gamma[X]))}\\ &\quad + \|n_h^L - n\|_{L^\infty(0,T;L^2(\Gamma[X]))} + \|v_h^L - v\|_{L^2(0,T;L^2(\Gamma[X]))} \leqslant Ch^{k+1},\end{aligned} \tag{2.18}$$

where C is a constant independent of h and $t \in [0,T]$ (but may depend on T).

The rest of this paper is devoted to the proofs of Theorems 2.1 and 2.2. The proof of Theorem 2.3 is standard and therefore omitted (in fact, (2.17) follows from Theorems 2.1 and 2.2 with the application of the triangle inequality, and (2.18) follows from an additional estimate for the interpolation errors). The appendices in the supplementary material (Supp_Materials.pdf [local/web 379KB])provide essential results and detailed proofs of Lemmas 3.5 and 3.6, which are used for proving Theorems 2.1 and 2.2. These proofs, following techniques similar to those in Lemmas 3.3 and 3.4, have been omitted from the main paper.

3. Proof of Theorem 2.1.

3.1. Lifts. Throughout this article, we denote by C and h_0 two generic positive constants which are different at different occurrences, possibly depending on the norms of the exact solution and T, but are independent of mesh size h.

Given a smooth surface $\Gamma \subset \mathbb{R}^3$, the surface tangential gradient of a scalar function $u : \Gamma \to \mathbb{R}$ is a column vector denoted by $\nabla_\Gamma u$. For a vector-valued function $u = (u_1, u_2, u_3)^\top : \Gamma \to \mathbb{R}^3$, we define

$$\nabla_\Gamma u := (\nabla_\Gamma u_1, \nabla_\Gamma u_2, \nabla_\Gamma u_3).$$

We denote by id the identity function on $\mathbb{R}^3$, i.e., $\mathrm{id}(x) = x$ for $x \in \mathbb{R}^3$. Its domain of definition can be restricted to any surface in $\mathbb{R}^3$.

The finite element basis functions on $\Gamma_h[\mathbf{x}]$ are denoted by $\phi_j[\mathbf{x}]$, $j = 1, \dots, N$, which are polynomials of degree k after being pulled back to the reference plane triangle, and satisfy the following identities:

$$\phi_j[\mathbf{x}](x_i) = \delta_{ij}, \qquad i, j = 1, \dots, N.$$

This definition of basis functions implies the following transport property (see [19]):

$$\partial^\bullet_{t,h} \phi_j[\mathbf{x}(t)] = 0 \quad \text{on } \Gamma_h[\mathbf{x}(t)],\ j = 1, \dots, N. \tag{3.1}$$

The finite element space on $\Gamma_h[\mathbf{x}]$ is defined as $S_h(\Gamma_h[\mathbf{x}]) := \{ \sum_{j=1}^N c_j \phi_j[\mathbf{x}] : c_j \in \mathbb{R} \}$.

From [33, Lemma 7.1] or [19, equations (2.15) and (2.16)], we know that there exists $h_0 > 0$ such that for $h \leqslant h_0$ and $t \in [0, T]$, any point $x \in \Gamma_h[\mathbf{x}^*(t)]$ can be lifted to a unique point $x^l \in \Gamma(t)$ satisfying the relation

$$x^l - x = \pm |x^l - x| n(x^l).$$

Then any function φ on $\Gamma_h[\mathbf{x}^*(t)]$ can be lifted to a function φ^l on $\Gamma(t)$, defined as

$$\varphi^l(x^l) = \varphi(x) \quad \forall\, x \in \Gamma_h[\mathbf{x}^*(t)].$$

The lifted functions satisfy the following estimates uniformly for h and t:

$$\begin{aligned} C^{-1} \|\phi\|_{L^2(\Gamma_h[\mathbf{x}^*])} &\le \|\phi^l\|_{L^2(\Gamma[X])} \leqslant C \|\phi\|_{L^2(\Gamma_h[\mathbf{x}^*])}, \\ C^{-1} \|\nabla_{\Gamma_h[\mathbf{x}^*]} \varphi\|_{L^2(\Gamma_h[\mathbf{x}^*])} &\le \|\nabla_{\Gamma[X]} \varphi^l\|_{L^2(\Gamma[X])} \leqslant C \|\nabla_{\Gamma_h[\mathbf{x}^*]} \varphi\|_{L^2(\Gamma_h[\mathbf{x}^*])}. \end{aligned} \tag{3.2}$$

These hold for all $\phi \in L^2(\Gamma_h[\mathbf{x}^*])$ and $\varphi \in H^1(\Gamma_h[\mathbf{x}^*])$, respectively.

3.2. Matrix-vector formulation. The matrix-vector notations of [26, 31, 32] will be used in this paper. In particular, we define $\mathbf{K}(\mathbf{x}) = \mathbf{M}(\mathbf{x}) + \mathbf{A}(\mathbf{x})$, with $\mathbf{M}(\mathbf{x}) \in \mathbb{R}^{N\times N}$ and $\mathbf{A}(\mathbf{x}) \in \mathbb{R}^{N\times N}$ denoting the mass matrix and stiffness matrix associated to finite element space $S_h(\Gamma_h[\mathbf{x}])$ on surface $\Gamma_h[\mathbf{x}]$, respectively, and define

$$\mathbf{M}^{[d]}(\mathbf{x}) = I_d \otimes \mathbf{M}(\mathbf{x}) \quad \text{and} \quad \mathbf{A}^{[d]}(\mathbf{x}) = I_d \otimes \mathbf{A}(\mathbf{x}),$$

where I_d is the $d \times d$ identity matrix. We denote by $\mathbf{v}$, $\mathbf{n}$, and $\mathbf{H}$ the nodal vectors of v_h, n_h, and H_h, respectively, and denote by $\mathbf{f}_1(\mathbf{x}, \mathbf{n}, \mathbf{v}) \in \mathbb{R}^{3N}$ and $\mathbf{f}_2(\mathbf{x}, \mathbf{n}, \mathbf{v}, \mathbf{H}) \in \mathbb{R}^N$ the nonlinear terms associated to the right-hand side of (2.1c) and (2.1d), respectively, defined by

$$\mathbf{f}_1(\mathbf{x}, \mathbf{n})_{j+(m-1)N} = \int_{\Gamma_h[\mathbf{x}]} |\nabla_{\Gamma_h[\mathbf{x}]} n_h|^2 (n_h)_m \phi_j, \tag{3.3a}$$

$$\mathbf{f}_2(\mathbf{x}, \mathbf{n}, \mathbf{H})_j = \int_{\Gamma_h[\mathbf{x}]} |\nabla_{\Gamma_h[\mathbf{x}]} n_h|^2 H_h \phi_j, \tag{3.3b}$$

with $j = 1, \dots, N$ and $m = 1, 2, 3$, where $(n_h)_m$ denotes the mth component of $n_h \in \mathbb{R}^3$.

By introducing $\mathbf{u} := (\mathbf{n}, \mathbf{H})^\top$, the spatially semidiscrete parametric surface FEM in (2.1) can be rewritten into the matrix-vector form

$$\dot{\mathbf{x}} = \mathbf{v}, \tag{3.4a}$$

$$\mathbf{v} = -\mathbf{I}_h(\mathbf{H} \bullet \mathbf{n}), \tag{3.4b}$$

$$\mathbf{M}^{[4]}(\mathbf{x})\dot{\mathbf{u}} + \mathbf{A}^{[4]}(\mathbf{x})\mathbf{u} = \mathbf{f}(\mathbf{x}, \mathbf{u}), \tag{3.4c}$$

where

$$\mathbf{f}(\mathbf{x}, \mathbf{u}) = \begin{pmatrix} \mathbf{f}_1(\mathbf{x}, \mathbf{n}) \\ \mathbf{f}_2(\mathbf{x}, \mathbf{n}, \mathbf{H}) \end{pmatrix} \in \mathbb{R}^{3N+N} \tag{3.5}$$

and $\mathbf{I}_h(\mathbf{H} \bullet \mathbf{n})$ denotes the nodal vector of the finite element function $I_h(H_h n_h)$.

We denote by $\mathbf{y}^*$, $\mathbf{v}^*$, and $\mathbf{u}^* = ((\mathbf{H}^*)^\top, (\mathbf{n}^*)^\top)^\top$ the nodal vectors of Y_h^*, v_h^*, and $u_h^* = (H_h^*, (n_h^*)^\top)^\top$, respectively. The latter are defined in (2.11), which can be written into the following matrix-vector form:

$$\dot{\mathbf{y}}^* = -\mathbf{I}_h(\mathbf{H}^* \bullet \mathbf{n}^*), \tag{3.6a}$$

$$\mathbf{K}^{[4]}(\mathbf{y}^*)\mathbf{u}^* \cdot \boldsymbol{\varphi} = \int_\Gamma (u \cdot \varphi_h^l + \nabla_\Gamma u \cdot \nabla_\Gamma \varphi_h^l) \qquad \forall \varphi_h \in S_h(\Gamma_h[\mathbf{x}^*])^4, \tag{3.6b}$$

where $\boldsymbol{\varphi}$ is the nodal vector of finite element function φ_h. The existence of $\mathbf{y}^*$ as a sufficiently good approximation to $\mathbf{x}^*$ will be proved. Then (2.13) can be written into the following matrix-vector form:

$$\dot{\mathbf{y}}^* = \mathbf{v}^*, \tag{3.7}$$

$$\mathbf{v}^* = -\mathbf{I}_h(\mathbf{H}^* \bullet \mathbf{n}^*), \tag{3.8}$$

$$\mathbf{M}^{[4]}(\mathbf{y}^*)\dot{\mathbf{u}}^* + \mathbf{A}^{[4]}(\mathbf{y}^*)\mathbf{u}^* = \mathbf{f}(\mathbf{y}^*, \mathbf{u}^*) + \ \mathbf{M}^{[4]}(\mathbf{y}^*)\mathbf{d}_{\mathbf{u}}, \tag{3.9}$$

where $\mathbf{d}_{\mathbf{u}}^*$ denotes the nodal vector of the finite element function $d_u^* = (d_H^*, (d_n^*)^\top)^\top$, with d_H^* and d_n^* being the remainders defined in (2.13). In the rest of this paper, we omit the superscripts in $\mathbf{M}^{[4]}(\mathbf{x})$, $\mathbf{A}^{[4]}(\mathbf{x})$, and $\mathbf{K}^{[4]}(\mathbf{x})$ for the simplicity of notations.

3.3. Perturbation of mass matrix and stiffness matrix. For $\mathbf{e_y} = \mathbf{y}^* - \mathbf{x}^*$, which is the nodal vector of the finite element function $\hat{e}_y = \hat{y}_h^* - \mathrm{id}$ on $\Gamma_h[\mathbf{x}^*]$, we consider the following intermediate surfaces:

$$\Gamma_h[\mathbf{y}^\theta] \quad \text{with} \quad \mathbf{y}^\theta = (1-\theta)\mathbf{x}^* + \theta\mathbf{y}^* = \mathbf{x}^* + \theta\mathbf{e_y} \quad \text{for } \theta \in [0,1].$$

The finite element functions on $\Gamma_h[\mathbf{y}^\theta]$ with nodal vectors $\mathbf{e_y}$, $\mathbf{z}$, and $\mathbf{w}$ are denoted by $\hat{e}_y^\theta$, $\hat{z}_h^\theta$, and $\hat{w}_h^\theta$, respectively (thus, $\hat{e}_y^0 = \hat{e}_y$). The following result was proved in [33, Lemma 4.3] and [31, Lemma 7.2].

Lemma 3.1. *If $\|\nabla_{\Gamma_h[\mathbf{x}^*]}\hat{e}_y\|_{L^\infty(\Gamma_h[\mathbf{x}^*])} \leqslant 1/2$, then the following inequalities hold for $\theta \in [0,1]$ and $1 \leqslant p \leqslant \infty$:*

$$\|\hat{w}_h^\theta\|_{L^p(\Gamma_h[\mathbf{y}^\theta])} \leqslant c_p \|\hat{w}_h^0\|_{L^p(\Gamma_h[\mathbf{x}^*])}, \tag{3.10}$$

$$\|\nabla_{\Gamma_h[\mathbf{y}^\theta]}\hat{w}_h^\theta\|_{L^p(\Gamma_h[\mathbf{y}^\theta])} \leqslant c_p \|\nabla_{\Gamma_h[\mathbf{x}^*]}\hat{w}_h^0\|_{L^p(\Gamma_h[\mathbf{x}^*])}, \tag{3.11}$$

where c_p is a constant independent of θ and h and $c_\infty = 2$.

The following lemma was proved in [31, Lemma 7.1].

LEMMA 3.2. *If* $\|\nabla_{\Gamma_h[\mathbf{x}^*]}\hat{e}_y\|_{L^\infty(\Gamma_h[\mathbf{x}^*])} \leqslant 1/2$*, then the following relations hold:*

$$(\mathbf{M}(\mathbf{y}^*) - \mathbf{M}(\mathbf{x}^*))\mathbf{z}\cdot\mathbf{w} = \int_0^1 \int_{\Gamma_h[\mathbf{y}^\theta]} \hat{w}_h^\theta (\nabla_{\Gamma_h[\mathbf{y}^\theta]} \cdot \hat{e}_y^\theta)\hat{z}_h^\theta, \tag{3.12}$$

$$(\mathbf{A}(\mathbf{y}^*) - \mathbf{A}(\mathbf{x}^*))\mathbf{z}\cdot\mathbf{w} = \int_0^1 \int_{\Gamma_h[\mathbf{y}^\theta]} \nabla_{\Gamma_h[\mathbf{y}^\theta]}\hat{w}_h^\theta \cdot \big(D_{\Gamma_h[\mathbf{y}^\theta]}\hat{e}_y^\theta\big)\nabla_{\Gamma_h[\mathbf{y}^\theta]}\hat{z}_h^\theta, \tag{3.13}$$

where $D_{\Gamma_h[\mathbf{y}^\theta]}\hat{e}_y^\theta = \mathrm{tr}\,\big(E^\theta\big)I_3 - \big(E^\theta + (E^\theta)^\top\big)$ *and* $E^\theta = \nabla_{\Gamma_h[\mathbf{y}^\theta]}\hat{e}_y^\theta \in \mathbb{R}^{3\times 3}$.

Lemmas 3.1 and 3.2 imply the following result: If $\|\nabla_{\Gamma_h[\mathbf{x}^*]}\hat{e}_y\|_{L^\infty(\Gamma_h[\mathbf{x}^*])} \leqslant 1/2$, then $\|\nabla_{\Gamma_h[\mathbf{y}^\theta]}\hat{e}_y^\theta\|_{L^\infty(\Gamma_h[\mathbf{y}^\theta])} \leqslant 1$ for $\theta \in [0,1]$, and therefore

(3.14)
the norms $\|\cdot\|_{\mathbf{M}(\mathbf{x}^*+\theta\mathbf{e_y})}$ are h-uniformly equivalent for $\theta \in [0,1]$,
the norms $\|\cdot\|_{\mathbf{A}(\mathbf{x}^*+\theta\mathbf{e_y})}$ are h-uniformly equivalent for $\theta \in [0,1]$,

(3.15)
$$(\mathbf{M}(\mathbf{y}^*) - \mathbf{M}(\mathbf{x}^*))\mathbf{z}\cdot\mathbf{w} \leqslant C\|\hat{w}_h^0\|_{L^p(\Gamma_h[\mathbf{x}^*])}\|\nabla_{\Gamma_h[\mathbf{x}^*]}\hat{e}_y^0\|_{L^{p'}(\Gamma_h[\mathbf{x}^*])}\|\hat{z}_h^0\|_{L^\infty(\Gamma_h[\mathbf{x}^*])},$$

(3.16)
$$(\mathbf{A}(\mathbf{y}^*) - \mathbf{A}(\mathbf{x}^*))\mathbf{z}\cdot\mathbf{w} \leqslant C\|\hat{w}_h^0\|_{W^{1,p}(\Gamma_h[\mathbf{x}^*])}\|\hat{e}_y^0\|_{W^{1,p'}(\Gamma_h[\mathbf{x}^*])}\|\hat{z}_h^0\|_{W^{1,\infty}(\Gamma_h[\mathbf{x}^*])},$$

(3.17)
$$(\mathbf{A}(\mathbf{y}^*) - \mathbf{A}(\mathbf{x}^*))\mathbf{z}\cdot\mathbf{w} \leqslant C\|\hat{e}_y^0\|_{W^{1,\infty}(\Gamma_h[\mathbf{x}^*])}\|\hat{z}_h^0\|_{W^{1,p}(\Gamma_h[\mathbf{x}^*])}\|\hat{w}_h^0\|_{W^{1,p'}(\Gamma_h[\mathbf{x}^*])},$$

given that p and p' satisfy the relation $\frac{1}{p} + \frac{1}{p'} = 1$.

In addition to Lemmas 3.1 and 3.2, the following relation will be used: If $K \subset \Gamma$ or $K \subset \Gamma_h[\,\mathbf{x}^*]$ is a smooth piece of surface which evolves under the velocity field w and $\partial_t^\bullet$ denotes the material derivative with respect to w, then

$$\partial_t^\bullet \nabla_K f = \nabla_K \partial_t^\bullet f - (\nabla_K w - n_K n_K^\top (\nabla_K w)^\top)\nabla_K f, \tag{3.18}$$

where n_K denotes the unit normal vector of K.

3.4. $W^{1,p}$ error estimates for the dynamic Ritz projection.

Let $\hat{y}_h^*$, $\hat{H}_h^*$, and $\hat{n}_h^*$ be the finite element functions in $S_h(\Gamma_h[\mathbf{x}^*])$ with the nodal vectors $\mathbf{y}^*$, $\mathbf{H}^*$, and $\mathbf{n}^*$ defined in (3.6), respectively. In this subsection, we prove the following lemma.

LEMMA 3.3. *Under the assumptions of Theorem* 2.1*, there exists* $h_0 > 0$ *such that for mesh size* $h \leqslant h_0$*, there exists a unique solution* $(\mathbf{y}^*, \mathbf{H}^*, \mathbf{n}^*)$ *of* (3.6) *satisfying the following estimates for all* $2 \leqslant p < \infty$*:*

$$\begin{aligned}&\|\hat{y}_h^* - \mathrm{id}_{\Gamma_h[\mathbf{x}^*]}\|_{W^{1,p}(\Gamma_h[\mathbf{x}^*])} + \|\hat{v}_h^* - \hat{I}_h^* v\|_{W^{1,p}(\Gamma_h[\mathbf{x}^*])}\\ &\quad + \|\hat{H}_h^* - \hat{I}_h^* H\|_{W^{1,p}(\Gamma_h[\mathbf{x}^*])} + \|\hat{n}_h^* - \hat{I}_h^* n\|_{W^{1,p}(\Gamma_h[\mathbf{x}^*])} \leqslant C_p h^k.\end{aligned} \tag{3.19}$$

Proof. Problem (3.6) is essentially a system of ODEs. We assume that $t_* \in (0,T]$ is the maximal time such that the solution of problem (3.6) exists and satisfies the following estimates for $t \in [0, t_*]$:

$$\|\hat{y}_h^* - \mathrm{id}_{\Gamma_h[\mathbf{x}^*]}\|_{W^{1,\infty}(\Gamma_h[\mathbf{x}^*])} \leqslant 1/2, \tag{3.20a}$$

$$\|\hat{H}_h^* - \hat{I}_h^* H\|_{W^{1,\infty}(\Gamma_h[\mathbf{x}^*])} \leqslant 1/2, \tag{3.20b}$$

$$\|\hat{n}_h^* - \hat{I}_h^* n\|_{W^{1,\infty}(\Gamma_h[\mathbf{x}^*])} \leqslant 1/2. \tag{3.20c}$$

Under this condition, we shall prove that (3.19) holds for $t\in[0,t_*]$ and all $2\leqslant p<\infty$ (with some constants h_0 and C_p that are independent of t_*). In particular, for $p=4$, the local-in-time existence and uniqueness of solutions to ODE system (2.11) (and the continuity in time of solutions to the ODE system) guarantee that its solution extends to $t\in[0,t_*+\delta_h]$ for some $\delta_h>0$ and satisfies (3.19) for $t\in[0,t_*+\delta_h]$ with C_4 replaced by $2C_4$. For sufficiently small h (smaller than some constant independent of t_*) such that $2C_4h^{k-2/4}\leqslant 1/2$, this implies that (3.20) holds for $t\in[0,t_*+\delta_h]$. This will prove that $t_*=T$ (otherwise, $t_*\in(0,T]$ is not the maximal time for (3.20) to hold).

Since u_h^* is defined on the surface $\Gamma_h[\mathbf{y}^*]$ via (3.6b), we must bridge the gap between the discrete surfaces $\Gamma_h[\mathbf{y}^*]$ and $\Gamma_h[\mathbf{x}^*]$ to estimate $\|\hat u_h^*-\hat I_h^*u\|_{W^{1,p}(\Gamma_h[\mathbf{x}^*])}$. Under condition (3.20), we can rewrite (3.6b) as

$$\mathbf{K}(\mathbf{x}^*)\mathbf{u}^*\cdot\boldsymbol{\varphi}=(\mathbf{K}(\mathbf{x}^*)-\mathbf{K}(\mathbf{y}^*))\ \mathbf{u}^*\cdot\boldsymbol{\varphi}+\int_\Gamma\Big(u\,\varphi_h^l+\nabla_\Gamma u\cdot\nabla_\Gamma\varphi_h^l\Big). \tag{3.21}$$

To characterize the gap term $(\mathbf{K}(\mathbf{x}^*)-\mathbf{K}(\mathbf{y}^*))\mathbf{u}^*\cdot\boldsymbol{\varphi}$, we define $w\in H^1(\Gamma)^4$ as the solution of the following weak formulation:

$$\int_\Gamma\Big(w\cdot\varphi^l+\nabla_\Gamma w\cdot\nabla_\Gamma\varphi^l\Big)=(\mathbf{K}(\mathbf{x}^*)-\mathbf{K}(\mathbf{y}^*))\mathbf{u}^*\cdot(\mathbf{P}_h\varphi)\quad\forall\varphi^l\in H^1(\Gamma)^4, \tag{3.22}$$

where φ denotes the inverse lift of φ^l onto $\Gamma_h[\mathbf{x}^*]$ and $\mathbf{P}_h\varphi$ is the nodal vector of $P_h\varphi$, i.e., the L^2 projection of $\varphi\in H^1(\Gamma_h[\mathbf{x}^*])^4$ onto $S_h(\Gamma_h[\ \mathbf{x}^*])^4$. Since the L^2 projection operator P_h is bounded in the L^p norm for $1\leqslant p\leqslant\infty$ (see section SM2 in the supplementary material) and the L^p norms of φ and φ^l are equivalent for $1\leqslant p\leqslant\infty$, it follows that

$$\|P_h\varphi\|_{L^p(\Gamma_h[\mathbf{x}^*])}\leqslant C\|\varphi\|_{L^p(\Gamma_h[\mathbf{x}^*])}\leqslant C\|\varphi^l\|_{L^p(\Gamma)}\quad\forall 1\leqslant p\leqslant\infty.$$

Since $\mathbf{P}_h\varphi_h=\boldsymbol{\varphi}$ for all $\varphi_h\in S_h(\Gamma_h[\mathbf{x}^*])$, it follows from (3.21) and (3.22) that

$$\mathbf{K}(\mathbf{x}^*)\mathbf{u}^*\cdot\boldsymbol{\varphi}=\int_\Gamma((w+u)\cdot\varphi_h^l+\nabla_\Gamma(w+u)\cdot\nabla_\Gamma\varphi_h^l)\quad\forall\varphi_h\in S_h(\Gamma_h[\mathbf{x}^*]). \tag{3.23}$$

This means that $\hat u_h^*$ is the linear Ritz projection of $w+u$ onto $S_h(\Gamma_h[\mathbf{x}^*])$. If we further define $\hat w_h^*$ as the linear Ritz projection of w onto $S_h(\Gamma_h[\mathbf{x}^*])$, then $\hat u_h^*-\hat w_h^*$ is the linear Ritz projection of u onto $S_h(\Gamma_h[\mathbf{x}^*])$.

Estimate for $\|w\|_{W^{1,p}(\Gamma)}$. To derive an estimate for $\|w\|_{W^{1,p}(\Gamma)}$, we consider the PDE problem (3.22) on the continuous surface, which can be reformulated as

$$\int_\Gamma(w\varphi^l+\nabla_\Gamma w\cdot\nabla_\Gamma\varphi^l)=\ell(\varphi^l), \tag{3.24}$$

where $\ell(\varphi^l)=(\mathbf{K}(\mathbf{x}^*)-\mathbf{K}(\mathbf{y}^*))\mathbf{u}^*\cdot\mathbf{P}_h\varphi$ is a linear functional on φ^l. Under condition (3.20), the following estimate follows from inequalities (3.15) and (3.16):

$$\begin{aligned}|\ell(\varphi^l)|&\leqslant C\|\hat u_h^*\|_{W^{1,\infty}(\Gamma_h[\mathbf{x}^*])}\|\mathrm{id}_{\Gamma_h[\mathbf{x}^*]}-\hat y_h^*\|_{W^{1,p}(\Gamma_h[\mathbf{x}^*])}\|\varphi\|_{W^{1,p'}(\Gamma_h[\mathbf{x}^*])}\\&\leqslant C\|\mathrm{id}_{\Gamma_h[\mathbf{x}^*]}-\hat y_h^*\|_{W^{1,p}(\Gamma_h[\mathbf{x}^*])}\|\varphi^l\|_{W^{1,p'}(\Gamma)}.\end{aligned}$$

This means that

$$\|\ell\|_{W^{-1,p}(\Gamma)}:=\|\ell\|_{(W^{1,p'}(\Gamma))'}\leqslant C\|\mathrm{id}_{\Gamma_h[\mathbf{x}^*]}-\hat y_h^*\|_{W^{1,p}(\Gamma_h[\mathbf{x}^*])}.$$

Then the standard $W^{1,p}$ estimates for the elliptic PDE problem in (3.24) imply (cf. [37, Theorem 1], which extends to PDEs on surfaces via estimates on local coordinate charts) that

$$\|w\|_{W^{1,p}(\Gamma)} \leqslant C\|\ell\|_{W^{-1,p}(\Gamma)} \leqslant C\|\mathrm{id}_{\Gamma_h[\mathbf{x}^*]} - \hat{y}_h^*\|_{W^{1,p}(\Gamma_h[\mathbf{x}^*])} \quad \text{for } 2 \leqslant p < \infty. \tag{3.25}$$

Estimates for the L^2 and $W^{1,p}$ norms of $(\hat{u}_h^*)^l - (\hat{w}_h^*)^l - u$. The following $W^{1,p}$ and H^1 estimates for the linear Ritz projection onto the interpolated surface $\Gamma_h[\mathbf{x}^*]$ were established in [19, Corollaries 4.2 and 4.5]:

$$\|(\hat{w}_h^*)^l - w\|_{L^2(\Gamma)} + h\|(\hat{w}_h^*)^l - w\|_{H^1(\Gamma)} \leqslant Ch^{k+1}\|w\|_{H^{k+1}(\Gamma)}, \tag{3.26}$$

$$\begin{aligned}\|(\hat{w}_h^*)^l - w\|_{W^{1,\infty}(\Gamma)} &\leqslant C\|w - (\hat{I}_h^* w)^l\|_{W^{1,\infty}(\Gamma)} + Ch^{k+1}|\ln h|\|w\|_{W^{1,\infty}(\Gamma)} \\ &\leqslant Ch^k\|w\|_{W^{k+1,\infty}(\Gamma)}.\end{aligned} \tag{3.27}$$

Since the complex interpolation spaces between $H^{k+1}(\Gamma)$ and $W^{k+1,\infty}(\Gamma)$ are $W^{k+1,p}(\Gamma)$ for $2 \leqslant p \leqslant \infty$ (cf. [10, Theorem 6.4.5]), the complex interpolation (cf. [10, Theorem 4.1.2]) of the above two estimates yields, for $2 \leqslant p \leqslant \infty$,

$$\|(w_h^*)^l - w\|_{W^{1,p}(\Gamma)} \leqslant Ch^k\|w\|_{W^{k+1,p}(\Gamma)}. \tag{3.28}$$

Moreover, the linear Ritz projection onto the interpolated surface is naturally stable in the H^1 norm, i.e.,

$$\|(\hat{w}_h^*)^l\|_{H^1(\Gamma)} \leqslant C\|w\|_{H^1(\Gamma)}. \tag{3.29}$$

By utilizing (3.27) and the $W^{1,\infty}$ stability of Lagrange interpolation, we derive the following result:

$$\begin{aligned}\|(\hat{w}_h^*)^l\|_{W^{1,\infty}(\Gamma)} &\leqslant \|(\hat{w}_h^*)^l - w\|_{W^{1,\infty}(\Gamma)} + \|w\|_{W^{1,\infty}(\Gamma)} \\ &\leqslant C\|w - (\hat{I}_h^* w)^l\|_{W^{1,\infty}(\Gamma)} + C\|w\|_{W^{1,\infty}(\Gamma)} \leqslant C\|w\|_{W^{1,\infty}(\Gamma)}.\end{aligned} \tag{3.30}$$

The complex interpolation between (3.29) and (3.30) yields the following result (cf. [10, Theorem 4.1.2]) for $2 \leqslant p \leqslant \infty$:

$$\|(\hat{w}_h^*)^l\|_{W^{1,p}(\Gamma)} \leqslant C\|w\|_{W^{1,p}(\Gamma)}. \tag{3.31}$$

Since $\hat{u}_h^* - \hat{w}_h^*$ is the linear Ritz projection of u onto $S_h(\Gamma_h[\mathbf{x}^*])$, replacing w by u in (3.26) and (3.28) yields that

$$\|(\hat{u}_h^*)^l - (\hat{w}_h^*)^l - u\|_{L^2(\Gamma)} + h\|(\hat{u}_h^*)^l - (\hat{w}_h^*)^l - u\|_{H^1(\Gamma)} \leqslant C\|u\|_{H^{k+1}(\Gamma)}h^{k+1}, \tag{3.32}$$

$$\|(\hat{u}_h^*)^l - (\hat{w}_h^*)^l - u\|_{W^{1,p}(\Gamma)} \leqslant C\|u\|_{W^{k+1,p}(\Gamma)}h^k \quad \text{for } 2 \leqslant p \leqslant \infty. \tag{3.33}$$

Estimates for the L^2 and $W^{1,p}$ norms of $(\hat{u}_h^*)^l - u$. As a result of (3.26) and (3.32), by using the triangle inequality, we have

$$\begin{aligned}\|(\hat{u}_h^*)^l - u\|_{L^2(\Gamma)} &\leqslant \|(\hat{w}_h^*)^l\|_{L^2(\Gamma)} + C\|u\|_{H^{k+1}(\Gamma)}h^{k+1} \\ &\leqslant \|w\|_{L^2(\Gamma)} + \|(\hat{w}_h^*)^l - w\|_{L^2(\Gamma)} + C\|u\|_{H^{k+1}(\Gamma)}h^{k+1} \\ &\leqslant \|w\|_{L^2(\Gamma)} + Ch^2\|w\|_{H^2(\Gamma)} + C\|u\|_{H^{k+1}(\Gamma)}h^{k+1},\end{aligned} \tag{3.34}$$

$$\|(\hat{u}_h^*)^l - u\|_{W^{1,p}(\Gamma)} \leqslant \|(\hat{w}_h^*)^l\|_{W^{1,p}(\Gamma)} + C\|u\|_{W^{k+1,p}(\Gamma)}h^k. \tag{3.35}$$

Then substituting inequality (3.31) into (3.35), we obtain

$$\|(\hat{u}_h^*)^l - u\|_{W^{1,p}(\Gamma)} \leqslant C\|w\|_{W^{1,p}(\Gamma)} + C\|u\|_{W^{k+1,p}(\Gamma)}h^k. \tag{3.36}$$

From (3.36) and (3.25), we obtain the following result for $2 \leqslant p < \infty$:

$$\|(\hat{u}_h^*)^l - u\|_{W^{1,p}(\Gamma)} \leqslant C\|\mathrm{id}_{\Gamma_h[\mathbf{x}^*]} - \hat{y}_h^*\|_{W^{1,p}(\Gamma_h[\mathbf{x}^*])} + C\|u\|_{W^{k+1,p}(\Gamma)}h^k. \tag{3.37}$$

In order to establish an estimate for $\|\mathrm{id}_{\Gamma_h[\mathbf{x}^*]} - \hat{y}_h^*\|_{W^{1,p}(\Gamma_h[\mathbf{x}^*])}$ on the right-hand side of (3.37), we consider the flow maps $X_h^* : \Gamma_h[\mathbf{x}^0] \to \Gamma_h[\mathbf{x}^*]$ and $Y_h^* = \hat{y}_h^* \circ X_h^*$, which satisfy the relation

$$\frac{\mathrm{d}}{\mathrm{d}t}(X_h^* - Y_h^*) = -\hat{I}_h^*(\hat{I}_h^* H \hat{I}_h^* n - \hat{H}_h^* \hat{n}_h^*) \circ X_h^*,$$

which can be written into the integral form:

$$X_h^*(s) - Y_h^*(s) = -\int_0^s \hat{I}_h^*(\hat{I}_h^* H \hat{I}_h^* n - \hat{H}_h^* \hat{n}_h^*) \circ X_h^* \,\mathrm{d}t. \tag{3.38}$$

Then applying gradient to (3.38) and using the chain rule of partial differentiation, we have

$$\nabla_{\Gamma_h^0}(X_h^*(s) - Y_h^*(s)) = -\int_0^s \nabla_{\Gamma_h^0} X_h^* \,[\nabla_{\Gamma_h[\mathbf{x}^*]} \hat{I}_h^*(\hat{I}_h^* H \hat{I}_h^* n - \hat{H}_h^* \hat{n}_h^*)] \circ X_h^* \,\mathrm{d}t.$$

Then by considering the L^p norm of both sides of this relation, we obtain the following result for $s \in (0, t_*]$:

$$\begin{aligned}
&\|\nabla_{\Gamma_h^0}(X_h^*(s) - Y_h^*(s))\|_{L^p(\Gamma_h^0)} \\
&\quad\leqslant C\int_0^s \|[\nabla_{\Gamma_h[\mathbf{x}^*]} \hat{I}_h^*(\hat{I}_h^* H \hat{I}_h^* n - \hat{H}_h^* \hat{n}_h^*)] \circ X_h^*\|_{L^p(\Gamma_h^0)} \mathrm{d}t \quad (\text{since } \|X_h^*\|_{W^{1,\infty}(\Gamma_h^0)} \text{ is bounded}) \\
&\quad\leqslant C\int_0^s (\|\hat{I}_h^* H - \hat{H}_h^*\|_{W^{1,p}(\Gamma_h[\mathbf{x}^*])} \|\hat{I}_h^* n\|_{W^{1,\infty}(\Gamma_h[\mathbf{x}^*])} + \|\hat{H}_h^*\|_{W^{1,\infty}(\Gamma_h[\mathbf{x}^*])} \|\hat{I}_h^* n - \hat{n}_h^*\|_{W^{1,p}(\Gamma_h[\mathbf{x}^*])}) \mathrm{d}t \\
&\qquad\qquad (W^{1,p} \text{ stability of } \hat{I}_h^* \text{ is used; see Appendix A in the supplementary material}) \\
&\quad\leqslant C\int_0^s (\|\hat{I}_h^* H - \hat{H}_h^*\|_{W^{1,p}(\Gamma_h[\mathbf{x}^*])} + \|\hat{I}_h^* n - \hat{n}_h^*\|_{W^{1,p}(\Gamma_h[\mathbf{x}^*])}) \mathrm{d}t,
\end{aligned} \tag{3.39}$$

where the last inequality uses the boundedness of $\|\hat{H}_h^*\|_{W^{1,\infty}(\Gamma_h[\mathbf{x}^*])}$, which follows from (3.20). Similarly, by considering the L^p norm of (3.38) directly, we can obtain the following result:

$$\|X_h^*(s) - Y_h^*(s)\|_{L^p(\Gamma_h^0)} \leqslant C\int_0^s (\|\hat{I}_h^* H - \hat{H}_h^*\|_{L^p(\Gamma_h[\mathbf{x}^*])} + \|\hat{I}_h^* n - \hat{n}_h^*\|_{L^p(\Gamma_h[\mathbf{x}^*])}) \mathrm{d}t. \tag{3.40}$$

Since $\mathrm{id}_{\Gamma_h[\mathbf{x}^*]} - \hat{y}_h^* = (X_h^* - Y_h^*) \circ (X_h^*)^{-1}$ and therefore $\|\mathrm{id}_{\Gamma_h[\mathbf{x}^*]} - \hat{y}_h^*\|_{W^{1,p}(\Gamma[\mathbf{x}^*(s)])} \sim \|X_h^*(s) - Y_h^*(s)\|_{W^{1,p}(\Gamma_h^0)}$, the last two estimates and Lagrange interpolation error estimates imply that

$$\begin{aligned}
&\|\mathrm{id}_{\Gamma_h[\mathbf{x}^*(s)]} - \hat{y}_h^*\|_{W^{1,p}(\Gamma[\mathbf{x}^*(s)])} \\
&\quad\leqslant C\int_0^s (\|\hat{I}_h^* H - \hat{H}_h^*\|_{W^{1,p}(\Gamma_h[\mathbf{x}^*])} + \|\hat{I}_h^* n - \hat{n}_h^*\|_{W^{1,p}(\Gamma_h[\mathbf{x}^*])}) \mathrm{d}t \\
&\quad\leqslant C\int_0^s (\|H - (\hat{H}_h^*)^l\|_{W^{1,p}(\Gamma)} + \|n - (\hat{n}_h^*)^l\|_{W^{1,p}(\Gamma)}) \mathrm{d}t + Ch^k \\
&\quad\leqslant C\int_0^s \|u - (\hat{u}_h^*)^l\|_{W^{1,p}(\Gamma)} \mathrm{d}t + Ch^k \quad \text{for } s \in (0, t_*].
\end{aligned} \tag{3.41}$$

Substituting this into (3.37) and using Gronwall's inequality, we obtain

$$\|\mathrm{id}_{\Gamma_h[\mathbf{x}^*]} - \hat{y}_h^*\|_{W^{1,p}(\Gamma_h[\mathbf{x}^*])} + \|(\hat{u}_h^*)^l - u\|_{W^{1,p}(\Gamma)} \leqslant Ch^k \quad \text{for } 2 \leqslant p < \infty.$$

Since $\|(\hat{I}_h^* u)^l - (\hat{u}_h^*)^l\|_{W^{1,p}(\Gamma)} \leqslant \|(\hat{I}_h^* u)^l - u\|_{W^{1,p}(\Gamma)} + \|(\hat{u}_h^*)^l - u\|_{W^{1,p}(\Gamma)} \leqslant Ch^k$ and $\|(\hat{I}_h^* u)^l - (\hat{u}_h^*)^l\|_{W^{1,p}(\Gamma)} \sim \|\hat{I}_h^* u - \hat{u}_h^*\|_{W^{1,p}(\Gamma_h[\mathbf{x}^*])}$, it follows that

$$\|\mathrm{id}_{\Gamma_h[\mathbf{x}^*]} - \hat{y}_h^*\|_{W^{1,p}(\Gamma_h[\mathbf{x}^*])} + \|\hat{I}_h^* u - \hat{u}_h^*\|_{W^{1,p}(\Gamma_h[\mathbf{x}^*])} \leqslant Ch^k \quad \text{for } 2 \leqslant p < \infty. \tag{3.42}$$

Moreover, we can express $\hat{I}_h^* v - \hat{v}_h^*$ as

$$\begin{aligned} \hat{I}_h^* v - \hat{v}_h^* &= -\hat{I}_h^*(\hat{I}_h^* H \hat{I}_h^* n) - \hat{I}_h^*(\hat{H}_h^* \hat{n}_h^*) \\ &= -\hat{I}_h^*[(\hat{I}_h^* H - \hat{H}_h^*)\hat{I}_h^* n] - \hat{I}_h^*[\hat{H}_h^*(\hat{I}_h^* n - \hat{n}_h^*)] \end{aligned} \tag{3.43}$$

and apply the $W^{1,p}$ stability of Lagrange interpolation (as shown in section SM1 in the supplementary material). This leads to the following result for $t \in [0, t_*]$:

$$\begin{aligned} \|\hat{I}_h^* v - \hat{v}_h^*\|_{W^{1,p}(\Gamma_h[\mathbf{x}^*])} &\leqslant C\|\hat{I}_h^* H - \hat{H}_h^*\|_{W^{1,p}(\Gamma_h[\mathbf{x}^*])}\|\hat{I}_h^* n\|_{W^{1,\infty}(\Gamma_h[\mathbf{x}^*])} \\ &\quad + C\|\hat{H}_h^*\|_{W^{1,\infty}(\Gamma_h[\mathbf{x}^*])}\|\hat{I}_h^* n - \hat{n}_h^*\|_{W^{1,p}(\Gamma_h[\mathbf{x}^*])} \\ &\leqslant C\|\hat{I}_h^* u - \hat{u}_h^*\|_{W^{1,p}(\Gamma_h[\mathbf{x}^*])} \leqslant Ch^k, \end{aligned} \tag{3.44}$$

where the boundedness of $\|\hat{H}_h^*\|_{W^{1,\infty}(\Gamma_h[\mathbf{x}^*])}$ follows from (3.20) and the last inequality follows from (3.42). This proves Lemma 3.3 according to the discussions in the text below (3.20). ∎

3.5. L^2 error estimates for the dynamic Ritz projection.

LEMMA 3.4. *Under the assumptions of Theorem* 2.1*, there exists* $h_0 > 0$ *such that for mesh size* $h \leqslant h_0$*,* (3.6) *has a unique solution* $(\mathbf{y}^*, \mathbf{H}^*, \mathbf{n}^*)$*, which satisfies the following estimate:*

$$\|\hat{y}_h^* - \mathrm{id}_{\Gamma_h[\mathbf{x}^*]}\|_{L^2(\Gamma_h[\mathbf{x}^*])} + \|\hat{H}_h^* - \hat{I}_h^* H\|_{L^2(\Gamma_h[\mathbf{x}^*])} + \|\hat{n}_h^* - \hat{I}_h^* n\|_{L^2(\Gamma_h[\mathbf{x}^*])} \leqslant Ch^{k+1}. \tag{3.45}$$

Proof. From (3.34), we see that in order to estimate $\|(\hat{u}_h^*)^l - u\|_{L^2(\Gamma)}$, it suffices to estimate $\|w\|_{L^2(\Gamma)}$ and $\|w\|_{H^2(\Gamma)}$. By estimating the right-hand side of (3.22) using (3.15) and (3.16) along with the results in Lemma 3.3 and the inverse inequality for finite element functions, we immediately get the estimate

$$\begin{aligned} \left| \int_\Gamma (w\varphi^l + \nabla_\Gamma w \cdot \nabla_\Gamma \varphi^l) \right| &\leqslant C\|\mathrm{id}_{\Gamma_h[\mathbf{x}^*]} - \hat{y}_h^*\|_{H^1(\Gamma_h[\mathbf{x}^*])}\|\hat{u}_h^*\|_{W^{1,\infty}(\Gamma_h[\mathbf{x}^*])}\|P_h\varphi\|_{H^1(\Gamma_h[\mathbf{x}^*])} \\ &\leqslant Ch^{-2}\|\mathrm{id}_{\Gamma_h[\mathbf{x}^*]} - \hat{y}_h^*\|_{L^2(\Gamma_h[\mathbf{x}^*])}\|\hat{u}_h^*\|_{W^{1,\infty}(\Gamma_h[\mathbf{x}^*])}\|\varphi\|_{L^2(\Gamma_h[\mathbf{x}^*])} \\ &\leqslant Ch^{-2}\|\mathrm{id}_{\Gamma_h[\mathbf{x}^*]} - \hat{y}_h^*\|_{L^2(\Gamma_h[\mathbf{x}^*])}\|\varphi^l\|_{L^2(\Gamma)}, \end{aligned}$$

which implies that (via duality) $\int_\Gamma (w\varphi^l + \nabla_\Gamma w \cdot \nabla_\Gamma \varphi^l) = \int_\Gamma f\varphi^l$ for some function f satisfying the following inequality:

$$\|f\|_{L^2(\Gamma)} \leqslant Ch^{-2}\|\mathrm{id}_{\Gamma_h[\mathbf{x}^*]} - \hat{y}_h^*\|_{L^2(\Gamma_h[\mathbf{x}^*])}.$$

The standard H^2 regularity estimate of elliptic equations says that

$$\|w\|_{H^2(\Gamma)} \leqslant C\|f\|_{L^2(\Gamma)} \leqslant Ch^{-2}\|\mathrm{id}_{\Gamma_h[\mathbf{x}^*]} - \hat{y}_h^*\|_{L^2(\Gamma_h[\mathbf{x}^*])}.$$

Substituting this into (3.34) yields

$$\|(\hat{u}_h^*)^l - u\|_{L^2(\Gamma)} \leqslant \|w\|_{L^2(\Gamma)} + C\|\mathrm{id}_{\Gamma_h[\mathbf{x}^*]} - \hat{y}_h^*\|_{L^2(\Gamma_h[\mathbf{x}^*])} + Ch^{k+1}. \tag{3.46}$$

It remains to estimate $\|w\|_{L^2(\Gamma)}$ and $\|\mathrm{id}_{\Gamma_h[\mathbf{x}^*]} - \hat{y}_h^*\|_{L^2(\Gamma_h[\mathbf{x}^*])}$. This can be done as follows by using (3.22) with a duality argument. We define $\varphi^l \in H^2(\Gamma)$ to be the solution of

$$\varphi^l - \Delta_\Gamma \varphi^l = w \text{ on } \Gamma \tag{3.47}$$

and let φ be the inverse lift of φ^l onto $\Gamma_h[\mathbf{x}^*]$. Then

$$\|\varphi^l\|_{H^2(\Gamma)} \leqslant C\|w\|_{L^2(\Gamma)}. \tag{3.48}$$

For $\mathbf{y}^\theta := (1-\theta)\mathbf{x}^* + \theta\mathbf{y}^*$, let $\hat{u}_h^{*,\theta}$, $\hat{\varphi}_h^\theta$, and $\hat{e}_y^\theta$ be the finite element functions on $\Gamma_h[\mathbf{y}^\theta]$ with nodal vectors $\mathbf{u}^*$, $\mathbf{P}_h\varphi$, and $\mathbf{e}_\mathbf{y} := \mathbf{y}^* - \mathbf{x}^*$, respectively. In particular, $\hat{u}_h^{*,0} = \hat{u}_h^*$. Then surface $\Gamma_h[\mathbf{y}^\theta]$ moves with velocity $\hat{e}_y^\theta$ as parameter $\theta \in [0,1]$ changes, and the following relation holds:

$$\partial_\theta^\bullet \hat{u}_h^{*,\theta} = \partial_\theta^\bullet \hat{\varphi}_h^\theta = 0.$$

Testing (3.47) by w and using (3.22), we have

$$\begin{aligned}
\|w\|_{L^2(\Gamma)}^2 &= \int_\Gamma (w\varphi^l + \nabla_\Gamma w \cdot \nabla_\Gamma \varphi^l) \\
&= (\mathbf{K}(\mathbf{x}^*) - \mathbf{K}(\mathbf{y}^*))\mathbf{u}^* \cdot (\mathbf{P}_h\varphi) \\
&= \left| \int_0^1 \frac{\mathrm{d}}{\mathrm{d}\theta} \int_{\Gamma_h[\mathbf{y}^\theta]} (\hat{u}_h^{*,\theta} \cdot \hat{\varphi}_h^\theta + \nabla_{\Gamma_h[\mathbf{y}^\theta]} \hat{u}_h^{*,\theta} \cdot \nabla_{\Gamma_h[\mathbf{y}^\theta]} \hat{\varphi}_h^\theta) \right| \\
&= \left| \int_0^1 \int_{\Gamma_h[\mathbf{y}^\theta]} (\hat{u}_h^{*,\theta} \cdot \hat{\varphi}_h^\theta \nabla_{\Gamma_h[\mathbf{y}^\theta]} \cdot \hat{e}_y^\theta + \nabla_{\Gamma_h[\mathbf{y}^\theta]} \hat{u}_h^{*,\theta} \cdot D_{\Gamma_h[\mathbf{y}^\theta]} \hat{e}_y^\theta \nabla_{\Gamma_h[\mathbf{y}^\theta]} \hat{\varphi}_h^\theta) \,\mathrm{d}\theta \right| \\
&\qquad \text{(see Lemma 3.2)} \\
&= \left| B_0 + B_1 + B_2 + B_3 \right|,
\end{aligned} \tag{3.49}$$

where

$$B_0 := \int_\Gamma (\hat{u}_h^{*,l} \cdot \hat{\varphi}_h^{0,l} \nabla_\Gamma \cdot \hat{e}_y^{0,l} + \nabla_\Gamma u \cdot D_\Gamma \hat{e}_y^{0,l} \nabla_\Gamma \varphi^l), \tag{3.50}$$

$$B_1 := \int_\Gamma (\nabla_\Gamma \hat{u}_h^{*,l} \cdot D_\Gamma \hat{e}_y^{0,l} \nabla_\Gamma \hat{\varphi}_h^{0,l} - \nabla_\Gamma u \cdot D_\Gamma \hat{e}_y^{0,l} \nabla_\Gamma \varphi^l), \tag{3.51}$$

$$\begin{aligned}
B_2 := &\int_{\Gamma_h[\mathbf{x}^*]} (\hat{u}_h^* \cdot \hat{\varphi}_h^0 \nabla_{\Gamma_h[\mathbf{x}^*]} \cdot \hat{e}_y^0 + \nabla_{\Gamma_h[\mathbf{x}^*]} \hat{u}_h^* \cdot D_{\Gamma_h[\mathbf{x}^*]} \hat{e}_y^0 \nabla_{\Gamma_h[\mathbf{x}^*]} \hat{\varphi}_h^0) \\
&- \int_\Gamma (\hat{u}_h^{*,l} \cdot \hat{\varphi}_h^{0,l} \nabla_\Gamma \cdot \hat{e}_y^{0,l} + \nabla_\Gamma \hat{u}_h^{*,l} \cdot D_\Gamma \hat{e}_y^{0,l} \nabla_\Gamma \hat{\varphi}_h^{0,l}),
\end{aligned} \tag{3.52}$$

$$\begin{aligned}
B_3 := &\int_0^1 \int_{\Gamma_h[\mathbf{y}^\theta]} (\hat{u}_h^{*,\theta} \cdot \hat{\varphi}_h^\theta \nabla_{\Gamma_h[\mathbf{y}^\theta]} \cdot \hat{e}_y^\theta + \nabla_{\Gamma_h[\mathbf{y}^\theta]} \hat{u}_h^{*,\theta} \cdot D_{\Gamma_h[\mathbf{y}^\theta]} \hat{e}_y^\theta \nabla_{\Gamma_h[\mathbf{y}^\theta]} \hat{\varphi}_h^\theta) \,\mathrm{d}\theta \\
&- \int_0^1 \int_{\Gamma_h[\mathbf{x}^*]} (\hat{u}_h^* \cdot \hat{\varphi}_h^0 \nabla_{\Gamma_h[\mathbf{x}^*]} \cdot \hat{e}_y^0 + \nabla_{\Gamma_h[\mathbf{x}^*]} \hat{u}_h^* \cdot D_{\Gamma_h[\mathbf{x}^*]} \hat{e}_y^0 \nabla_{\Gamma_h[\mathbf{x}^*]} \hat{\varphi}_h^0) \,\mathrm{d}\theta.
\end{aligned} \tag{3.53}$$

In the expression of B_0, we can remove the gradient acting on $\hat{e}_y^{0,l}$ by utilizing integration by parts. This will yield the following result:

$$(3.54)\qquad |B_0| \leqslant C\|\hat{e}_y^{0,l}\|_{L^2(\Gamma)}\|\varphi^l\|_{H^2(\Gamma)} \leqslant C\|\hat{e}_y^0\|_{L^2(\Gamma_h[\mathbf{x}^*])}\|\varphi^l\|_{H^2(\Gamma)}.$$

Since $\hat{\varphi}_h^{0,l} = (P_h\varphi)^l$, it follows that $\|\nabla_\Gamma(\varphi^l - \hat{\varphi}_h^{0,l})\|_{L^2(\Gamma)} \leqslant Ch\|\varphi^l\|_{H^2(\Gamma)}$. This implies that

$$\begin{aligned}|B_1| \leqslant& C\|(\hat{u}_h^*)^l - u\|_{W^{1,p}(\Gamma)}\|\hat{e}_y^{0,l}\|_{H^1(\Gamma)}\|\varphi^l\|_{W^{1,\frac{2p}{p-2}}(\Gamma)} \\ &+ \|(\hat{u}_h^*)^l\|_{W^{1,\infty}(\Gamma_h[\mathbf{x}^*])}\|\hat{e}_y^{0,l}\|_{H^1(\Gamma)}Ch\|\varphi^l\|_{H^2(\Gamma)} \quad (\text{for some } p>2) \\ (3.55)\qquad \leqslant& Ch\|\hat{e}_y^{0,l}\|_{H^1(\Gamma)}\|\varphi^l\|_{H^2(\Gamma)},\end{aligned}$$

where the last inequality follows from the $W^{1,p}$ error estimate in Lemma 3.3. Similarly, by estimating B_2 with the error between the interpolated surface $\Gamma_h[\mathbf{x}^*]$ and the exact surface $\Gamma[X]$ (cf. [19, Proposition 2.3]), we can derive the following result:

$$(3.56)\qquad |B_2| \leqslant Ch\|\hat{e}_y^{0,l}\|_{H^1(\Gamma)}\|\varphi^l\|_{H^2(\Gamma)}.$$

We consider the intermediate surfaces between $\Gamma_h[\mathbf{y}^\theta]$ and $\Gamma_h[\mathbf{x}^*]$ defined by

$$\Gamma_h[\mathbf{y}^{\theta,\alpha}], \quad \text{with} \quad \mathbf{y}^{\theta,\alpha} = (1-\alpha)\mathbf{x}^* + \alpha\mathbf{y}^\theta = \mathbf{x}^* + \alpha\theta\mathbf{e_y}, \quad \alpha\in[0,1],$$

for fixed $\theta\in[0,1]$. As α varies, the intermediate surface $\Gamma_h[\mathbf{y}^{\theta,\alpha}]$ moves with velocity $\theta\hat{e}_y^\theta$. By employing these intermediate surfaces along with the estimates in (3.15)–(3.17) and the bound $\|\hat{e}_y^0\|_{W^{1,p}(\Gamma_h[\mathbf{x}^*])} \leqslant Ch$ from (3.19) in Lemma 3.3, we can obtain (details are omitted)

$$(3.57)\qquad |B_3| \leqslant Ch\|\hat{e}_y^{0,l}\|_{H^1(\Gamma)}\|\varphi^l\|_{H^2(\Gamma)}.$$

Therefore, by substituting the estimates of B_0, B_1, B_2, and B_3 into (3.49) and using the estimate of $\|\varphi^l\|_{H^2(\Gamma)}$ in (3.48), we obtain

$$\begin{aligned}\|w\|_{L^2(\Gamma)}^2 &\leqslant C\|\hat{e}_y^0\|_{L^2(\Gamma_h[\mathbf{x}^*])}\|w\|_{L^2(\Gamma)} + Ch\|\hat{e}_y^0\|_{H^1(\Gamma_h[\mathbf{x}^*])}\|w\|_{L^2(\Gamma)} \\ &\leqslant C\|\hat{e}_y^0\|_{L^2(\Gamma_h[\mathbf{x}^*])}\|w\|_{L^2(\Gamma)} + Ch^{k+1}\|w\|_{L^2(\Gamma)},\end{aligned}$$

where the last inequality follows from the $W^{1,p}$ estimate of $\hat{e}_y^0 = \mathrm{id}_{\Gamma_h[\mathbf{x}^*]} - \hat{y}_h^*$ in Lemma 3.3 with $p=2$. This implies that

$$(3.58)\qquad \|w\|_{L^2(\Gamma)} \leqslant C\|\mathrm{id}_{\Gamma_h[\mathbf{x}^*]} - \hat{y}_h^*\|_{L^2(\Gamma_h[\mathbf{x}^*])} + Ch^{k+1}.$$

Then substituting this result into (3.46), we obtain

$$(3.59)\qquad \|(\hat{u}_h^*)^l - u\|_{L^2(\Gamma)} \leqslant C\|\mathrm{id}_{\Gamma_h[\mathbf{x}^*]} - \hat{y}_h^*\|_{L^2(\Gamma_h[\mathbf{x}^*])} + Ch^{k+1}.$$

The first term on the right-hand side of (3.59) can be estimated similarly to (3.41) by choosing $p=2$ in (3.40) and rewriting it equivalently as follows:

(3.60)

$$\|\mathrm{id}_{\Gamma_h[\mathbf{x}^*(s)]} - \hat{y}_h^*\|_{L^2(\Gamma[\mathbf{x}^*(s)])} \leqslant C\int_0^s \|(\hat{u}_h^*)^l(t) - u(t)\|_{L^2(\Gamma)}\mathrm{d}t + Ch^{k+1} \quad \text{for } s\in[0,T].$$

Then substituting (3.59) into (3.60) and using Gronwall's inequality, we obtain an optimal-order estimate of $\|\mathrm{id}_{\Gamma_h[\mathbf{x}^*]} - \hat{y}_h^*\|_{L^2(\Gamma_h[\mathbf{x}^*])}$, i.e.,

$$(3.61)\qquad \|\mathrm{id}_{\Gamma_h[\mathbf{x}^*]} - \hat{y}_h^*\|_{L^2(\Gamma_h[\mathbf{x}^*])} \leqslant Ch^{k+1} \quad \text{for } s\in[0,T].$$

Substituting this estimate back into (3.59) yields the following result:

$$\|(\hat{u}_h^*)^l - u\|_{L^2(\Gamma)} \leqslant Ch^{k+1} \quad \text{for } s \in [0,T]. \tag{3.62}$$

Since $\|(\hat{I}_h^* u)^l - u\|_{L^2(\Gamma)} \leqslant Ch^{k+1}$, by using (3.62) and the triangle inequality

$$\|(\hat{u}_h^*)^l - (\hat{I}_h^* u)^l\|_{L^2(\Gamma)} \leqslant \|(\hat{u}_h^*)^l - u\|_{L^2(\Gamma)} + \|(\hat{I}_h^* u)^l - u\|_{L^2(\Gamma)}$$

as well as the norm equivalence $\|(\hat{u}_h^*)^l - (\hat{I}_h^* u)^l\|_{L^2(\Gamma)} \sim \|\hat{u}_h^* - \hat{I}_h^* u\|_{L^2(\Gamma_h[\mathbf{x}^*])}$, we obtain the L^2 error estimate of the dynamic Ritz projection in Lemma 3.4. ∎

Substituting Lemma 3.3 into (3.25) with $p = 2$ and substituting Lemma 3.4 into (3.58), we obtain

$$\|w\|_{L^2(\Gamma)} + h\|w\|_{H^1(\Gamma)} \leqslant Ch^{k+1}. \tag{3.63}$$

This result will be used in the next subsection.

In addition to the $W^{1,p}$ and L^2 error estimates of the dynamic Ritz projection, we can also differentiate (3.23) in time and, by using the resulting derivative equation, prove the following L^2 error estimates for the material derivative of the dynamic Ritz projection.

LEMMA 3.5. *Under the assumptions of Theorem* 2.1*, there exists* $h_0 > 0$ *such that for mesh size* $h \leqslant h_0$*, the solution* $(\mathbf{y}^*, \mathbf{H}^*, \mathbf{n}^*)$ *of* (3.6) *satisfies the following estimates:*

$$\|(\partial_{t,h}^\bullet H_h^*)^\wedge - \hat{I}_h^* \partial_t^\bullet H\|_{L^2(\Gamma_h[\mathbf{x}^*])} + h\|(\partial_{t,h}^\bullet H_h^*)^\wedge - \hat{I}_h^* \partial_t^\bullet H\|_{H^1(\Gamma_h[\mathbf{x}^*])} \leqslant Ch^{k+1}, \tag{3.64}$$

$$\|(\partial_{t,h}^\bullet n_h^*)^\wedge - \hat{I}_h^* \partial_t^\bullet n\|_{L^2(\Gamma_h[\mathbf{x}^*])} + h\|(\partial_{t,h}^\bullet n_h^*)^\wedge - \hat{I}_h^* \partial_t^\bullet n\|_{H^1(\Gamma_h[\mathbf{x}^*])} \leqslant Ch^{k+1}, \tag{3.65}$$

where $\partial_{t,h}^\bullet H_h^*$ *and* $(\partial_{t,h}^\bullet H_h^*)^\wedge$ *are finite element functions on* $\Gamma_h[\mathbf{y}^*]$ *and* $\Gamma_h[\mathbf{x}^*]$*, respectively, with a common nodal vector* $\dot{\mathbf{H}}^*$*.*

The proof of Lemma 3.5 is based on differentiating (3.23) in time, which leads to complicated expressions. However, the techniques for proving Lemma 3.5 are similar to those for proving Lemmas 3.3 and 3.4. Therefore, we omit the proof here and refer the reader to section SM3 in the supplementary material.

3.6. Estimates of the remainders. The remainders d_H^* and d_n^* defined in (2.13) can be estimated by using the approximation properties of the dynamic Ritz projection in Lemmas 3.3–3.5. This result is presented in the following lemma, and the proof is omitted. We refer the reader to section SM4 in the supplementary material for more details.

LEMMA 3.6. *Under the assumptions of Theorem* 2.1*, there exists* $h_0 > 0$ *such that for mesh size* $h \leqslant h_0$*, the remainder* d_u^* *defined in* (2.14) *satisfes the following estimate:*

$$\left| \int_{\Gamma_h[\mathbf{y}^*]} d_u^* \cdot \chi_u \right| \leqslant Ch^{k+1} \|\chi_u\|_{H^1(\Gamma_h[\mathbf{y}^*])^4}.$$

4. Proof of Theorems 2.2 and 2.3. In this section, we prove Theorem 2.2 on the optimal-order convergence of the parametric FEM for mean curvature flow in the $L^\infty(0,T;L^2)$ norm by utilizing the estimates of the dynamic Ritz projection in Lemmas 3.3–3.5 and the estimates of the remainders in Lemma 3.6.

4.1. Basic settings. The numerical solution $(\mathbf{x},\mathbf{v},\mathbf{u})$ and the dynamic Ritz projection $(\mathbf{x}^*,\mathbf{v}^*,\mathbf{u}^*)$ satisfy (3.4) and (3.7), respectively. By subtracting (3.7) from (3.4), we find that the error functions

$$\mathbf{e_x}=\mathbf{x}-\mathbf{y}^*,\quad \mathbf{e_v}=\mathbf{v}-\mathbf{v}^*,\quad \text{and}\quad \mathbf{e_u}=\mathbf{u}-\mathbf{u}^*$$

satisfy the following equations:

$$\dot{\mathbf{e}}_\mathbf{x}=\mathbf{e_v}, \tag{4.1a}$$

$$\mathbf{e_v}=-(\mathbf{I}_h(\mathbf{H}\bullet\mathbf{n})-\mathbf{I}_h(\mathbf{H}^*\bullet\mathbf{n}^*)), \tag{4.1b}$$

$$\begin{aligned}\mathbf{M}(\mathbf{x})\dot{\mathbf{e}}_\mathbf{u}+\mathbf{A}(\mathbf{x})\mathbf{e_u}&=-(\mathbf{M}(x)-\mathbf{M}(\mathbf{y}^*))\dot{\mathbf{u}}^*-(\mathbf{A}(x)-\mathbf{A}(\mathbf{y}^*))\mathbf{u}^*\\&\quad+(\mathbf{f}(\mathbf{x},\mathbf{u})-\mathbf{f}(\mathbf{y},\mathbf{u}^*))-\mathbf{M}(\mathbf{y}^*)\mathbf{d_u}.\end{aligned} \tag{4.1c}$$

Let $\mathbf{x}^\theta=(1-\theta)\mathbf{y}^*+\theta\mathbf{x}$ for $\theta\in[0,1]$, which defines an intermediate surface $\Gamma_h[\mathbf{x}^\theta]$ moving with the velocity e_x^θ as parameter $\theta\in[0,1]$ changes, and denote by e_x^θ, e_v^θ, and e_u^θ the finite element functions on $\Gamma_h[\mathbf{x}^\theta]$ with nodal vectors $\mathbf{e_x}$, $\mathbf{e_v}$, and $\mathbf{e_u}$, respectively. In particular, we denote $e_x=e_x^0$, $e_v=e_v^0$, and $e_u=e_u^0$, which are finite element functions on $\Gamma_h[\mathbf{y}^*]$. On the intermediate surface $\Gamma_h[\mathbf{x}^\theta]$, we also define finite element functions

$$v_h^\theta\quad\text{and}\quad u_h^\theta=(H_h^\theta,(n_h^\theta)^\top)^\top$$

with nodal vectors $\mathbf{v}^\theta=(1-\theta)\mathbf{v}^*+\theta\mathbf{v}$ and $\mathbf{u}^\theta=(1-\theta)\mathbf{u}^*+\theta\mathbf{u}$, respectively. We also denote by $u_h^{*,\theta}=(H_h^{*,\theta},(n_h^{*,\theta})^\top)^\top$ the finite element function on $\Gamma_h[\mathbf{x}^\theta]$ with nodal vector $\mathbf{u}^*$.

Similar to the proof of Theorem 2.1, we define $t^*\in[0,T]$ as the maximal time such that the numerical solution exists and the following inequalities are satisfied:

$$\|e_x(\cdot,t)\|_{W^{1,\infty}(\Gamma_h[\mathbf{y}^*])}\le h^{k-0.1}, \tag{4.2a}$$

$$\|e_u(\cdot,t)\|_{L^\infty(\Gamma_h[\mathbf{y}^*])}\le h^{k-0.1}\quad\text{for}\quad t\in[0,t^*]. \tag{4.2b}$$

At time $t=0$, we have $e_x(\cdot,0)=0$ and $e_u=\hat{I}_h^*u-\hat{u}_h^*$ on $\Gamma_h[\mathbf{x}^0]$. Therefore, by using the inverse inequality of finite element functions and the L^2 estimates of $\hat{I}_h^*u-\hat{u}_h^*$ in Theorem 2.1, we have

$$\|e_u(\cdot,0)\|_{L^\infty(\Gamma_h[\mathbf{y}^*])}\le Ch^{-1}\|e_u(\cdot,0)\|_{L^2(\Gamma_h[\mathbf{y}^*])}\leqslant Ch^k.$$

For sufficiently small h such that $Ch^k\leqslant h^{k-0.1}$, the inequality above implies that $t^*>0$.

Note that under condition (4.2), the L^p and $W^{1,p}$ norms on surfaces $\Gamma_h[\mathbf{x}]$ and $\Gamma_h[\mathbf{y}^*]$ are equivalent for $1\le p\le\infty$ (as shown in Lemma 3.1). This norm equivalence will be used frequently in the following subsections. In particular, under condition (4.2), we shall prove the following proposition (with some constants h_0 and C that are independent of t_*).

PROPOSITION 4.1. *Under the assumptions in Theorem* 2.1 *and* (4.2)*, there exists* $h_0>0$ *such that when mesh size* $h\leqslant h_0$*, the following estimate holds:*

$$\|e_x\|_{L^\infty(0,t^*;H^1(\Gamma_h[\mathbf{y}^*]))}+\|e_u\|_{L^\infty(0,t^*;L^2(\Gamma_h[\mathbf{y}^*]))}+\|e_u\|_{L^2(0,t^*;H^1(\Gamma_h[\mathbf{y}^*]))}\leqslant Ch^{k+1}. \tag{4.3}$$

Remark 4.1. By the local-in-time existence and uniqueness as well as the continuity of solutions to the ODE system (3.4), there exists $\delta_h>0$ such that the numerical solution and the error estimate in (4.3) hold for $t \in [0, t_* + \delta_h]$, with C replaced by $2C$ therein. This would imply that when h is smaller than some constant (which is independent of t_*), (4.2) holds for $t \in [0, t_* + \delta_h]$. This would imply that $t_* = T$ (otherwise, $t_* \in (0, T]$ would not be the maximal time satisfying the condition), and therefore the error estimate in (4.3) holds for $t \in [0, T]$. Then by the norm equivalence between $\Gamma_h[\mathbf{y}^*]$ and $\Gamma_h[\mathbf{x}^*]$, the error estimate in (4.3) can be equivalently written into (2.16). This would complete the proof of Theorem 2.2.

4.2. Estimates of $\|e_u\|_{L^\infty(0,t;L^2(\Gamma_h[\mathbf{y}^*]))}$ and $\|e_u\|_{L^2(0,t;H^1(\Gamma_h[\mathbf{y}^*]))}$. In this subsection, we establish an estimate of $\|e_u\|_{L^\infty(0,t;L^2(\Gamma_h[\mathbf{y}^*]))}$ and $\|e_u\|_{L^2(0,t;H^1(\Gamma_h[\mathbf{y}^*]))}$ in terms of $\|e_x\|_{L^2(0,t;H^1(\Gamma_h[\mathbf{y}^*]))}$ and $\|e_u\|_{L^2(0,t;L^2(\Gamma_h[\mathbf{y}^*]))}$.

LEMMA 4.2. *Under the assumptions in Theorem* 2.1 *and* (4.2), *there exists* $h_0 > 0$ *such that when mesh size* $h \leqslant h_0$, *the following estimate holds for* $t \in [0, t^*]$*:*

$$\begin{aligned}
&\|e_u(t)\|^2_{L^2(\Gamma_h[\mathbf{y}^*(t)])} + \int_0^t \|\nabla_{\Gamma_h[\mathbf{y}^*]} e_u(s)\|^2_{L^2(\Gamma_h[\mathbf{y}^*(s)])}\,\mathrm{d}s \\
&\quad \le C\int_0^t \left(\|e_x(s)\|^2_{H^1(\Gamma_h[\mathbf{y}^*(s)])} + \|e_u(s)\|^2_{L^2(\Gamma_h[\mathbf{y}^*(s)])}\right)\mathrm{d}s + Ch^{2k+2}.
\end{aligned} \tag{4.4}$$

Proof. Testing (4.1c) with $\mathbf{e_u}$, we obtain the following relation:

$$\begin{aligned}
&\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|e_u^1\|^2_{L^2(\Gamma_h[\mathbf{x}])} + \|\nabla_{\Gamma_h[\mathbf{x}]} e_u^1\|^2_{L^2(\Gamma_h[\mathbf{x}])} \\
&\quad = -\mathbf{e_u}^\top(\mathbf{M}(\mathbf{x}) - \mathbf{M}(\mathbf{y}^*))\dot{\mathbf{u}}^* - \mathbf{e_u}^\top(\mathbf{A}(\mathbf{x}) - \mathbf{A}(\mathbf{y}^*))\dot{\mathbf{u}}^* \\
&\qquad + \mathbf{e_u}^\top(\mathbf{f}(\mathbf{x},\mathbf{u}) - \mathbf{f}(\mathbf{y}^*,\mathbf{u}^*)) - \mathbf{e_u}^\top\mathbf{M}(\mathbf{y}^*)\mathbf{d_u} \\
&\quad =: I_1 + I_2 + I_3 + I_4,
\end{aligned} \tag{4.5}$$

where e_u^1 denotes the finite element function on $\Gamma_h[\mathbf{x}^\theta]$ with $\theta = 1$. Additionally, $\partial_\theta^\bullet e_x^\theta = \partial_\theta^\bullet e_v^\theta = 0$ and $\partial_\theta^\bullet e_u^\theta = 0$. Since $u_h^* \in \Gamma_h[\mathbf{y}^*]$ and $\hat{u}_h^* \in \Gamma_h[\mathbf{x}^*]$ have the same nodal vectors, by using the equivalence of the L^p and $W^{1,p}$ norms on $\Gamma_h[\mathbf{x}^\theta]$ and $\Gamma_h[\mathbf{y}^*]$ and Lemmas 3.4 and 3.5, we have

$$\begin{aligned}
\|u_h^*\|_{W^{1,\infty}(\Gamma_h[\mathbf{y}^*])} &\le C\|\hat{u}_h^*\|_{W^{1,\infty}(\Gamma_h[\mathbf{x}^*])} \\
&\le C\|\hat{u}_h^* - \hat{I}_h^* u\|_{W^{1,\infty}(\Gamma_h[\mathbf{x}^*])} + C\|\hat{I}_h^* u\|_{W^{1,\infty}(\Gamma_h[\mathbf{x}^*])} \\
&\le Ch^{-2}h^{k+1} + C \le C,
\end{aligned} \tag{4.6}$$

$$\begin{aligned}
\|\partial_{t,h}^\bullet u_h^*\|_{W^{1,\infty}(\Gamma_h[\mathbf{y}^*])} &\le C\|(\partial_t^\bullet u_h^*)^\wedge\|_{W^{1,\infty}(\Gamma_h[\mathbf{x}^*])} \\
&\le C\|(\partial_t^\bullet u_h^*)^\wedge - \hat{I}_h^*\partial_t^\bullet u\|_{W^{1,\infty}(\Gamma_h[\mathbf{x}^*])} + C\|\hat{I}_h^*\partial_t^\bullet u\|_{W^{1,\infty}(\Gamma_h[\mathbf{x}^*])} \\
&\le Ch^{-2}h^{k+1} + C \le C,
\end{aligned} \tag{4.7}$$

where Lemma 3.5 and the inverse inequality are used in the last inequalities of (4.6) and (4.7). Then with inequalities (4.6) and 4.7) and (3.15)–(3.17), we can estimate $|I_1|$ and $|I_2|$ as follows:

$$|I_1| \leqslant C\|\nabla_{\Gamma_h[\mathbf{y}^*]} e_x\|_{L^2(\Gamma_h[\mathbf{y}^*])}\|e_u\|_{L^2(\Gamma_h[\mathbf{y}^*])} \tag{4.8}$$

$$|I_2| \leqslant C\|\nabla_{\Gamma_h[\mathbf{y}^*]} e_x\|_{L^2(\Gamma_h[\mathbf{y}^*])}\|\nabla_{\Gamma_h[\mathbf{y}^*]} e_u\|_{L^2(\Gamma_h[\mathbf{y}^*])}. \tag{4.9}$$

Lemma 3.6 guarantees that

$$|I_4| \leqslant Ch^{k+1}\|e_u\|_{H^1(\Gamma_h[\mathbf{y}^*])}. \tag{4.10}$$

It remains to estimate $|I_3|$. By employing the identity (3.18), we can bound $|I_3|$ by the sum of five terms as follows:

$$
\begin{aligned}
|I_3| &= |\mathbf{e}_\mathbf{u}^\top(\mathbf{f}(\mathbf{x},\mathbf{u})-\mathbf{f}(\mathbf{y}^*,\mathbf{u}^*))| \\
&= \Big|\int_{\Gamma_h[\mathbf{x}]} \big|\nabla_{\Gamma_h[\mathbf{x}]} n_h\big|^2 u_h\cdot e_u^1 - \int_{\Gamma_h[\mathbf{y}^*]} \big|\nabla_{\Gamma_h[\mathbf{y}^*]} n_h^*\big|^2 u_h^*\cdot e_u\Big| \\
&= \Big|\int_0^1 \frac{\mathrm{d}}{\mathrm{d}\theta}\int_{\Gamma_h[\mathbf{x}^\theta]} \big|\nabla_{\Gamma_h[\mathbf{x}^\theta]} n_h^\theta\big|^2 u_h^\theta\cdot e_u^\theta \mathrm{d}\theta\Big| \\
&\leqslant \Big|\int_0^1 \int_{\Gamma_h[\mathbf{x}^\theta]} \big|\nabla_{\Gamma_h[\mathbf{x}^\theta]} n_h^\theta\big|^2 u_h^\theta\cdot e_u^\theta \nabla_{\Gamma_h[\mathbf{x}^\theta]}\cdot e_x^\theta \,\mathrm{d}\theta\Big| \\
&\quad + \Big|\int_0^1 \int_{\Gamma_h[\mathbf{x}^\theta]} \big|\nabla_{\Gamma_h[\mathbf{x}^\theta]} n_h^\theta\big|^2 e_u^\theta\cdot e_u^\theta \,\mathrm{d}\theta\Big| \\
&\quad + \Big|\int_0^1 \int_{\Gamma_h[\mathbf{x}^\theta]} 2\nabla_{\Gamma_h[\mathbf{x}^\theta]} n_h^\theta \Big(\nabla_{\Gamma_h[\mathbf{x}^\theta]}\partial_\theta^\bullet n_h^\theta - \nabla_{\Gamma_h[\mathbf{x}^\theta]} e_x^\theta \nabla_{\Gamma_h[\mathbf{x}^\theta]} n_h^\theta \\
&\quad + n_{\Gamma_h[\mathbf{x}^\theta]} n_{\Gamma_h[\mathbf{x}^\theta]}^\top (\nabla_{\Gamma_h[\mathbf{x}^\theta]} e_x^\theta)^\top \nabla_{\Gamma_h[\mathbf{x}^\theta]} n_h^\theta\Big) u_h^\theta\cdot e_u^\theta \,\mathrm{d}\theta\Big| \\
&=: I_{31}+I_{32}+I_{33}+I_{34}+I_{35},
\end{aligned} \tag{4.11}
$$

where $n_{\Gamma_h[\mathbf{x}^\theta]}$ is the unit normal vector of $\Gamma_h[\mathbf{x}^\theta]$, while n_h^θ is the finite element function with nodal vector $\mathbf{n}^\theta = (1-\theta)\mathbf{n}^* + \theta\mathbf{n}$, with $\mathbf{n}^*$ and $\mathbf{n}$ being the nodal vectors of the dynamic Ritz projection $n_h^* \in S_h(\Gamma_h[\mathbf{y}^*])$ and numerical solution $n_h \in S_h(\Gamma_h[\mathbf{x}])$, respectively. Since $n_h^\theta = n_h^{*,\theta} + \theta e_n^\theta$ and $u_h^\theta = u_h^{*,\theta} + \theta e_u^\theta$, I_{31} can be bounded as follows:

$$
\begin{aligned}
I_{31} &= \Big|\int_0^1 \int_{\Gamma_h[\mathbf{x}^\theta]} \big|\nabla_{\Gamma_h[\mathbf{x}^\theta]}(n_h^{*,\theta}+\theta e_n^\theta)\big|^2 (u_h^{*,\theta}+\theta e_u^\theta)\cdot e_u^\theta \nabla_{\Gamma_h[\mathbf{x}^\theta]}\cdot e_x^\theta \,\mathrm{d}\theta\Big| \\
&\leqslant \Big|\int_0^1 \int_{\Gamma_h[\mathbf{x}^\theta]} \big|\nabla_{\Gamma_h[\mathbf{x}^\theta]} n_h^{*,\theta}\big|^2 (u_h^{*,\theta}\cdot e_u^\theta)\nabla_{\Gamma_h[\mathbf{x}^\theta]}\cdot e_x^\theta \,\mathrm{d}\theta\Big| \\
&\quad + \Big|\int_0^1 \int_{\Gamma_h[\mathbf{x}^\theta]} \theta\big|\nabla_{\Gamma_h[\mathbf{x}^\theta]} n_h^{*,\theta}\big|^2 \big|e_u^\theta\big|^2 \nabla_{\Gamma_h[\mathbf{x}^\theta]}\cdot e_x^\theta \,\mathrm{d}\theta\Big| \\
&\quad + \Big|\int_0^1 \int_{\Gamma_h[\mathbf{x}^\theta]} \theta^2\big|\nabla_{\Gamma_h[\mathbf{x}^\theta]} e_n^\theta\big|^2 (u_h^{*,\theta}\cdot e_u^\theta)\nabla_{\Gamma_h[\mathbf{x}^\theta]}\cdot e_x^\theta \,\mathrm{d}\theta\Big| \\
&\quad + \Big|\int_0^1 \int_{\Gamma_h[\mathbf{x}^\theta]} \theta^3\big|\nabla_{\Gamma_h[\mathbf{x}^\theta]} e_n^\theta\big|^2 \big|e_u^\theta\big|^2 \nabla_{\Gamma_h[\mathbf{x}^\theta]}\cdot e_x^\theta \,\mathrm{d}\theta\Big| \\
&\quad + \Big|\int_0^1 \int_{\Gamma_h[\mathbf{x}^\theta]} 2\theta(\nabla_{\Gamma_h[\mathbf{x}^\theta]} n_h^{*,\theta}\cdot\nabla_{\Gamma_h[\mathbf{x}^\theta]} e_n^\theta)(u_h^{*,\theta}\cdot e_u^\theta)\nabla_{\Gamma_h[\mathbf{x}^\theta]}\cdot e_x^\theta \,\mathrm{d}\theta\Big| \\
&\quad + \Big|\int_0^1 \int_{\Gamma_h[\mathbf{x}^\theta]} 2\theta^2\nabla_{\Gamma_h[\mathbf{x}^\theta]} n_h^{*,\theta}\cdot\nabla_{\Gamma_h[\mathbf{x}^\theta]} e_n^\theta \,|e_u^\theta|^2\, \nabla_{\Gamma_h[\mathbf{x}^\theta]}\cdot e_x^\theta \,\mathrm{d}\theta\Big| \\
&=: I_{311}+I_{312}+I_{313}+I_{314}+I_{315}+I_{316}.
\end{aligned} \tag{4.12}
$$

Then with the norm equivalence of the L^p and $W^{1,p}$ norms on $\Gamma_h[\mathbf{x}^\theta]$ and $\Gamma_h[\mathbf{y}^*]$ as well as estimates (4.2) and (4.6), the following estimates of I_{31j}, $j=1,\ldots,6$, can be derived:

$$
\begin{aligned}
I_{311} &\leqslant C\|e_u\|_{L^2(\Gamma_h[\mathbf{y}^*])}\|\nabla_{\Gamma_h[\mathbf{y}^*]}e_x\|_{L^2(\Gamma_h[\mathbf{y}^*])},\\
I_{312} &\leqslant C\|e_u\|_{L^2(\Gamma_h[\mathbf{y}^*])}\|\nabla_{\Gamma_h[\mathbf{y}^*]}e_x\|_{L^2(\Gamma_h[\mathbf{y}^*])}\|e_u\|_{L^\infty(\Gamma_h[\mathbf{y}^*])}\\
&\leqslant C\|e_u\|_{L^2(\Gamma_h[\mathbf{y}^*])}\|\nabla_{\Gamma_h[\mathbf{y}^*]}e_x\|_{L^2(\Gamma_h[\mathbf{y}^*])},\\
I_{313} &\leqslant C\|\nabla_{\Gamma_h[\mathbf{y}^*]}e_n\|^2_{L^2(\Gamma_h[\mathbf{y}^*])}\|\nabla_{\Gamma_h[\mathbf{y}^*]}e_x\|_{L^\infty(\Gamma_h[\mathbf{y}^*])}\|e_u\|_{L^\infty(\Gamma_h[\mathbf{y}^*])}\\
&\leqslant Ch^{2k-0.2}\|\nabla_{\Gamma_h[\mathbf{y}^*]}e_u\|^2_{L^2(\Gamma_h[\mathbf{y}^*])},\\
I_{314} &\leqslant C\|\nabla_{\Gamma_h[\mathbf{y}^*]}e_n\|^2_{L^2(\Gamma_h[\mathbf{y}^*])}\|\nabla_{\Gamma_h[\mathbf{y}^*]}e_x\|_{L^\infty(\Gamma_h[\mathbf{y}^*])}\|e_u\|^2_{L^\infty(\Gamma_h[\mathbf{y}^*])}\\
&\leqslant Ch^{3k-0.3}\|\nabla_{\Gamma_h[\mathbf{y}^*]}e_u\|^2_{L^2(\Gamma_h[\mathbf{y}^*])},\\
I_{315} &\leqslant C\|\nabla_{\Gamma_h[\mathbf{y}^*]}e_n\|_{L^2(\Gamma_h[\mathbf{y}^*])}\|e_u\|_{L^2(\Gamma_h[\mathbf{y}^*])}\|\nabla_{\Gamma_h[\mathbf{y}^*]}e_x\|_{L^\infty(\Gamma_h[\mathbf{y}^*])}\\
&\leqslant Ch^{k-0.1}\|e_u\|_{L^2(\Gamma_h[\mathbf{y}^*])}\|\nabla_{\Gamma_h[\mathbf{y}^*]}e_u\|_{L^2(\Gamma_h[\mathbf{y}^*])},\\
I_{316} &\leqslant C\|\nabla_{\Gamma_h[\mathbf{y}^*]}e_n\|_{L^2(\Gamma_h[\mathbf{y}^*])}\|e_u\|_{L^2(\Gamma_h[\mathbf{y}^*])}\|e_u\|_{L^\infty(\Gamma_h[\mathbf{y}^*])}\|\nabla_{\Gamma_h[\mathbf{y}^*]}e_x\|_{L^\infty(\Gamma_h[\mathbf{y}^*])}\\
&\leqslant Ch^{2k-0.2}\|e_u\|_{L^2(\Gamma_h[\mathbf{y}^*])}\|\nabla_{\Gamma_h[\mathbf{y}^*]}e_u\|_{L^2(\Gamma_h[\mathbf{y}^*])}.
\end{aligned}
\tag{4.13}
$$

By summing up the estimates of I_{31j}, $j=1,\dots,6$, in (4.13), we obtain

$$
I_{31} \leqslant C\|e_u\|^2_{L^2(\Gamma_h[\mathbf{y}^*])} + C\|e_x\|^2_{H^1(\Gamma_h[\mathbf{y}^*])} + Ch^{2k-0.2}\|\nabla_{\Gamma_h[\mathbf{y}^*]}e_u\|^2_{L^2(\Gamma_h[\mathbf{y}^*])}.
\tag{4.14}
$$

In the same way, the following estimates of I_{3j}, $j=2,\dots,5$, can be obtained:

$$
\begin{aligned}
I_{32}+I_{34}+I_{35} &\leqslant C\|e_u\|^2_{L^2(\Gamma_h[\mathbf{y}^*])} + C\|\nabla_{\Gamma_h[\mathbf{y}^*]}e_x\|^2_{L^2(\Gamma_h[\mathbf{y}^*])}\\
&\quad + Ch^{2k-0.2}\|\nabla_{\Gamma_h[\mathbf{y}^*]}e_u\|^2_{L^2(\Gamma_h[\mathbf{y}^*])}.
\end{aligned}
\tag{4.15}
$$

Furthermore, by applying Young's inequality, the term I_{33} can be bounded as follows:

$$
I_{33} \leqslant \epsilon\|\nabla_{\Gamma_h[\mathbf{y}^*]}e_u\|^2_{L^2(\Gamma_h[\mathbf{y}^*])} + C(\epsilon)\|e_u\|^2_{L^2(\Gamma_h[\mathbf{y}^*])}.
\tag{4.16}
$$

These estimates lead to

$$
|I_3| \leqslant C\|e_u\|^2_{L^2(\Gamma_h[\mathbf{y}^*])} + C\|e_x\|^2_{H^1(\Gamma_h[\mathbf{y}^*])} + (Ch^{2k-0.2}+\epsilon)\|\nabla_{\Gamma_h[\mathbf{y}^*]}e_u\|^2_{L^2(\Gamma_h[\mathbf{y}^*])}.
\tag{4.17}
$$

Then substituting estimates (4.8)–(4.10) and (4.17) into (4.5) yields the following result:

$$
\begin{aligned}
&\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|e_u^1\|^2_{L^2(\Gamma_h[\mathbf{x}])} + \|\nabla_{\Gamma_h[\mathbf{x}]}e_u^1\|^2_{L^2(\Gamma_h[\mathbf{x}])}\\
&\leqslant C\|e_x\|^2_{H^1(\Gamma_h[\mathbf{y}^*])} + C\|e_u\|^2_{L^2(\Gamma_h[\mathbf{y}^*])} + (Ch^{2k-0.2}+\epsilon)\|\nabla_{\Gamma_h[\mathbf{y}^*]}e_u\|^2_{L^2(\Gamma_h[\mathbf{y}^*])} + Ch^{2k+2}.
\end{aligned}
\tag{4.18}
$$

By employing the H^1 seminorm equivalence between the surfaces $\Gamma_h[\mathbf{y}^*]$ and $\Gamma_h[\mathbf{x}]$ and by choosing h and ϵ sufficiently small, the term $(Ch^{2k-0.2}+\epsilon)\|\nabla_{\Gamma_h[\mathbf{y}^*]}e_u\|^2_{L^2(\Gamma_h[\mathbf{y}^*])}$ can be absorbed into the left-hand side of the above inequality. This reduces (4.18) to the following result:

$$
\begin{aligned}
&\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|e_u^1\|^2_{L^2(\Gamma_h[\mathbf{x}])} + \|\nabla_{\Gamma_h[\mathbf{x}]}e_u^1\|^2_{L^2(\Gamma_h[\mathbf{x}])}\\
&\quad\leqslant C\|e_x\|^2_{H^1(\Gamma_h[\mathbf{y}^*])} + C\|e_u\|^2_{L^2(\Gamma_h[\mathbf{y}^*])} + Ch^{2k+2}.
\end{aligned}
\tag{4.19}
$$

Integrating the above inequality from 0 to t along with the norm equivalence between surfaces $\Gamma_h[\mathbf{x}]$ and $\Gamma_h[\mathbf{y}^*(t)]$, we have

$$\begin{aligned}&\|e_u(t)\|^2_{L^2(\Gamma_h[\mathbf{y}^*(t)])}+\int_0^t\|\nabla_{\Gamma_h[\mathbf{y}^*(s)]}e_u(s)\|^2_{L^2(\Gamma_h[\mathbf{y}^*(s)])}\,\mathrm{d}s\\ &\leqslant C\int_0^t\big(\|e_x(s)\|^2_{H^1(\Gamma_h[\mathbf{y}^*(s)])}+\|e_u(s)\|^2_{L^2(\Gamma_h[\mathbf{y}^*(s)])}\big)\,\mathrm{d}s+Ch^{2k+2},\end{aligned}\tag{4.20}$$

where $\|e_u(0)\|_{L^2(\Gamma_h[\mathbf{y}^*(0)])}\leqslant Ch^{k+1}$ is used. This proves the result of Lemma 4.2. ∎

4.3. Estimates of $\|e_v(t)\|_{H^1(\Gamma_h[\mathbf{y}^*(t)])}$. In this subsection, we establish an estimate of $\|e_v(t)\|_{H^1(\Gamma_h[\mathbf{y}^*(t)])}$ in terms of $\|e_u(t)\|_{H^1(\Gamma_h[\ \mathbf{y}^*(t)])}$.

LEMMA 4.3. *Under the assumptions in Theorem* 2.1 *and* (4.2)*, there exists $h_0>0$ such that when mesh size $h\leqslant h_0$, the following estimate holds for $t\in[0,t^*]$:*

$$\|e_v(t)\|_{H^1(\Gamma_h[\mathbf{y}^*(t)])}\leqslant C\|e_u(t)\|_{H^1(\Gamma_h[\mathbf{y}^*(t)])}.\tag{4.21}$$

Proof. Equation (4.1b) can be written as $\mathbf{e_v}=-\mathbf{I}_h[(\mathbf{H}-\mathbf{H}^*)\bullet\mathbf{n}]-\mathbf{I}_h[\mathbf{H}^*\bullet(\mathbf{n}-\mathbf{n}^*)]$, which implies (using the H^1 stability of Lagrange interpolation in section SM1 and the norm equivalence between surfaces $\Gamma_h[\mathbf{x}^*]$ and $\Gamma_h[\mathbf{y}^*]$) that

$$\begin{aligned}\|e_v\|_{H^1(\Gamma_h[\mathbf{y}^*])}&\leqslant C\|e_v\|_{H^1(\Gamma_h[\mathbf{x}^*])}\\ &\leqslant C\|e_H\|_{H^1(\Gamma_h[\mathbf{x}^*])}\|n_h^*\|_{W^{1,\infty}(\Gamma_h[\mathbf{x}^*])}+C\|e_n\|_{H^1(\Gamma_h[\mathbf{x}^*])}\|H_h^*\|_{W^{1,\infty}(\Gamma_h[\mathbf{x}^*])}\\ &\quad+C\|e_H\|_{H^1(\Gamma_h[\mathbf{x}^*])}\|e_n\|_{L^\infty(\Gamma_h[\mathbf{x}^*])}+C\|e_n\|_{H^1(\Gamma_h[\ \mathbf{x}^*])}\|e_H\|_{L^\infty(\Gamma_h[\mathbf{x}^*])}\\ &\leqslant C\|e_H\|_{H^1(\Gamma_h[\mathbf{y}^*])}\|n_h^*\|_{W^{1,\infty}(\Gamma_h[\mathbf{y}^*])}+C\|e_n\|_{H^1(\Gamma_h[\mathbf{y}^*])}\|H_h^*\|_{W^{1,\infty}(\Gamma_h[\mathbf{y}^*])}\\ &\quad+C\|e_H\|_{H^1(\Gamma_h[\mathbf{y}^*])}\|e_n\|_{L^\infty(\Gamma_h[\mathbf{y}^*])}+C\|e_n\|_{H^1(\Gamma_h[\ \mathbf{y}^*])}\|e_H\|_{L^\infty(\Gamma_h[\mathbf{y}^*])}\\ &\leqslant C\|e_u\|_{H^1(\Gamma_h[\mathbf{y}^*])},\end{aligned}\tag{4.22}$$

where (4.2b) and (4.6) are used in the last inequality. ∎

We can substitute the estimate in Lemma 4.2 into the estimate in Lemma 4.3. This yields the following inequality:

$$\begin{aligned}\|e_v\|_{L^2(0,t;H^1(\Gamma_h[\mathbf{y}^*]))}&\leqslant C\|e_u\|_{L^2(0,t;H^1(\Gamma_h[\mathbf{y}^*]))}\\ &\leqslant C(\|e_x\|_{L^2(0,t;H^1(\Gamma_h[\mathbf{y}^*]))}+\|e_u\|_{L^2(0,t;L^2(\Gamma_h[\mathbf{y}^*]))})+Ch^{k+1}.\end{aligned}\tag{4.23}$$

4.4. Proof of Proposition 4.1. Since $e_x(\cdot,0)=0$, it follows that

$$\begin{aligned}&\|e_x(t)\|^2_{H^1(\Gamma_h[\mathbf{y}^*(t)])}\\ &=\int_0^t\frac{\mathrm{d}}{\mathrm{d}s}\|e_x(s)\|^2_{H^1(\Gamma_h[\mathbf{y}^*])}\,\mathrm{d}s\\ &=\int_0^t\Big(2\mathbf{e_x}(s)^\top\mathbf{K}(\mathbf{y}^*(s))\dot{\mathbf{e}}_\mathbf{x}(s)+\mathbf{e_x}(s)^\top\frac{\mathrm{d}}{\mathrm{d}s}\mathbf{K}(\mathbf{y}^*(s))\mathbf{e_x}(s)\Big)\,\mathrm{d}s\\ &\leqslant C\int_0^t\Big(\|e_x(s)\|_{H^1(\Gamma_h[\mathbf{y}^*(s)])}\|e_v(s)\|_{H^1(\Gamma_h[\mathbf{y}^*(s)])}+\|e_x(s)\|^2_{H^1(\Gamma_h[\mathbf{y}^*(s)])}\Big)\,\mathrm{d}s\\ &\leqslant C\int_0^t\|e_v(s)\|^2_{H^1(\Gamma_h[\mathbf{y}^*(s)])}\,\mathrm{d}s+C\int_0^t\|e_x(s)\|^2_{H^1(\Gamma_h[\ \mathbf{y}^*(s)])}\,\mathrm{d}s,\end{aligned}\tag{4.24}$$

where the second-to-last inequality follows from the following estimate (which can be derived from the expressions in Lemma 3.2):

$$\begin{aligned}\Big|\mathbf{e_x}(s)^\top\frac{\mathrm{d}}{\mathrm{d}s}\mathbf{K}(\mathbf{y}^*(s))\mathbf{e_x}(s)\Big|&\leqslant C\|e_x(s)\|^2_{H^1(\Gamma_h[\mathbf{y}^*(s)])}\|v_h^*(s)\|_{W^{1,\infty}(\Gamma_h[\mathbf{y}^*(s)])}\\ &\leqslant C\|e_x(s)\|^2_{H^1(\Gamma_h[\mathbf{y}^*(s)])}.\end{aligned}$$

Here the boundedness of $\|v_h^*(s)\|_{W^{1,\infty}(\Gamma_h[\mathbf{y}^*(s)])}$ follows from applying the norm equivalence between $\Gamma_h[\mathbf{y}^*]$ and $\Gamma_h[\mathbf{x}^*]$, the triangle inequality, the inverse inequality, and the $W^{1,p}$ estimate in (3.19) with $p=2$, i.e.,

$$\begin{aligned}\|v_h^*\|_{W^{1,\infty}(\Gamma_h[\mathbf{y}^*])} \leqslant C\|\hat{v}_h^*\|_{W^{1,\infty}(\Gamma_h[\mathbf{x}^*])} &\leqslant C\|\hat{v}_h^* - \hat{I}_h^* v\|_{W^{1,\infty}(\Gamma_h[\mathbf{x}^*])} + \|\hat{I}_h^* v\|_{W^{1,\infty}(\Gamma_h[\mathbf{x}^*])}\\ &\leqslant Ch^{-1}\|\hat{v}_h^* - \hat{I}_h^* v\|_{H^1(\Gamma_h[\mathbf{x}^*])} + \|\hat{I}_h^* v\|_{W^{1,\infty}(\Gamma_h[\mathbf{x}^*])}\\ &\leqslant C.\end{aligned}$$

The right-hand side of (4.24) can be estimated with (4.23). This leads to the following result:

(4.25)
$$\|e_x(t)\|^2_{H^1(\Gamma_h[\mathbf{y}^*(t)])} \leqslant C\int_0^t \left(\|e_x(s)\|^2_{H^1(\Gamma_h[\mathbf{y}^*(s)])} + \|e_u(s)\|^2_{L^2(\Gamma_h[\mathbf{y}^*(s)])}\right) \mathrm{d}s + Ch^{2k+2}.$$

Then summing up (4.4) and (4.25), we obtain the following result for $t \in (0, t_*]$:

$$\begin{aligned}&\|e_x(t)\|^2_{H^1(\Gamma_h[\mathbf{y}^*(t)])} + \|e_u(t)\|^2_{L^2(\Gamma_h[\mathbf{y}^*(t)])} + \int_0^t \|\nabla_{\Gamma_h[\mathbf{y}^*]} e_u(s)\|^2_{L^2(\Gamma_h[\mathbf{y}^*(s)])}\,\mathrm{d}s\\ &\qquad \leqslant C\int_0^t \left(\|e_x(s)\|^2_{H^1(\Gamma_h[\mathbf{y}^*(s)])} + \|e_u(s)\|^2_{L^2(\Gamma_h[\mathbf{y}^*(s)])}\right) \mathrm{d}s + Ch^{2k+2}.\end{aligned} \tag{4.26}$$

The result of Proposition 4.1 follows from applying Gronwall's inequality to (4.26). Moreover, the discussions in Remark 4.1 show that $t^* = T$. This completes the proof of Theorem 2.2.

5. Numerical tests.

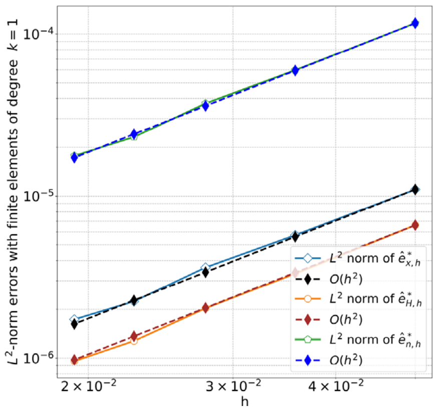
In this section, we present numerical experiments to support the theoretical analysis for the convergence rate of the semidiscrete parametric FEM in (2.1).

We consider the evolution of the two-dimensional sphere $\Gamma(t)$ under mean curvature flow, which was used for testing the convergence rates of numerical methods for mean curvature flow in [31]. The exact solution of the surface at time $t > 0$ is a sphere of radius $R(t) = \sqrt{R(0)^2 - 4t}$ with $R(0) = 2$, which reaches zero at time $t = 1$. The mean curvature H and normal vector n of the evolving sphere $\Gamma(t)$ can also be calculated explicitly.

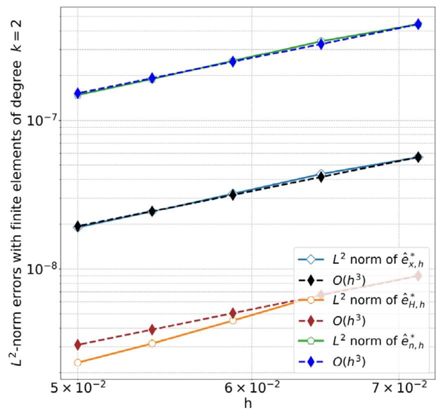
We solve the problem by the algorithm in (2.1) up to time $T = 0.125$ with finite elements of degrees 1, 2, and 3, respectively, using a four-step backward differentiation formula, with a sufficiently small time step size $\tau = 0.001$ such that the errors from temporal discretizations can be neglected in testing the convergence rates of spatial discretizations.

The L^2 norms of the error functions $\hat{e}_{x,h}^* = \hat{x}_h - \mathrm{id}_{\Gamma_h[\mathbf{x}^*]}$, $\hat{e}_{H,h}^* = \hat{H}_h - \hat{I}_h^* H$, and $\hat{e}_{n,h}^* = \hat{n}_h - \hat{I}_h^* n$ at time T are presented in Figure 5.1, which indicates that the errors of the numerical solutions are about $O(h^{k+1})$ for finite elements of degree $k = 1, 2, 3$. This is consistent with the error estimates established in Theorem 2.3.

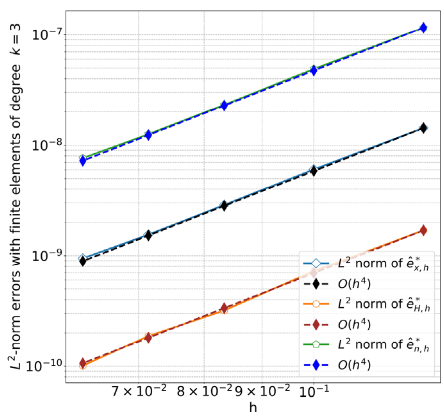
Figure 5.2(a) shows the H^1-seminorms of the error functions $\hat{e}_{x,h}$, $\hat{e}_{H,h}$, and $\hat{e}_{n,h}$ (between the numerical solution and the dynamic Ritz projection) at time T, demonstrating second-order convergence for finite elements of degree $k = 1$, which agrees with the error estimate in Theorem 2.2. In contrast, the H^1-seminorms of $\hat{e}_{x,h}^*$, $\hat{e}_{H,h}^*$, and $\hat{e}_{n,h}^*$ (between the numerical solution and the interpolated solution) is only $O(h)$, as shown in Figure 5.2(b).



(a) Finite elements of degree $k = 1$

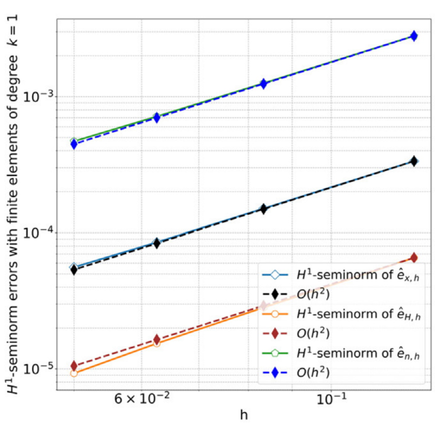


(b) Finite elements of degree $k = 2$

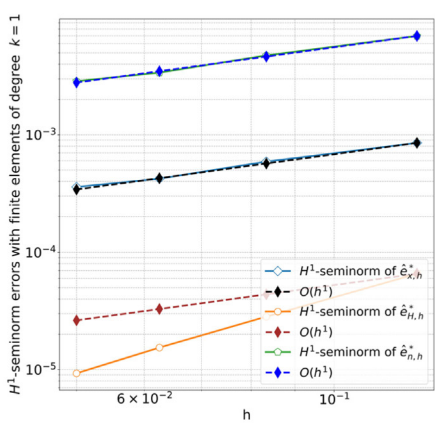


(c) Finite elements of degree $k = 3$

FIG. 5.1. *The L^2 norm for errors and convergence rates of the numerical solutions.*



(a) Convergence rates for $\hat{e}_{x,h}$, $\hat{e}_{H,h}$ and $\hat{e}_{n,h}$



(b) Convergence rates for $\hat{e}^*_{x,h}$, $\hat{e}^*_{H,h}$ and $\hat{e}^*_{n,h}$

FIG. 5.2. *The H^1-seminorm of the errors for finite elements of degree $k = 1$.*

6. Conclusion. We have defined a dynamic Ritz projection of mean curvature flow of closed surfaces in the three-dimensional space and have proved optimal-order error bounds in the L^2 and $W^{1,p}$ norms for the dynamic Ritz projection in approximating the solution of mean curvature flow. By utilizing these approximation results, we have proved optimal-order convergence of parametric FEMs for formulation (1.2) of mean curvature flow in the $L^\infty(0,T;L^2)$ norm as well as convergence of parametric FEMs for mean curvature flow with piecewise linear finite elements. The new approach developed in this paper—analyzing the error of numerical approximation through a dynamic Ritz projection of the mean curvature flow—can serve as a foundational framework for studying the convergence of numerical approximations for other geometric flows and parametric FEMs.

Acknowledgments. The authors would like to thank Prof. Christian Lubich for helpful discussions and valuable comments in an early stage of the research.

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 (2023), pp. 1609–1612, https://doi.org/10.1137/22M1521791.

[2] G. Bai and B. Li, *A new approach to the analysis of parametric finite element approximations to mean curvature flow*, Found. Comput. Math., 24 (2023), 64, https://doi.org/10.1007/s10208-023-09622-x.

[3] W. Bao and Y. Li, *A structure-preserving parametric finite element method for geometric flows with anisotropic surface energy*, Numer. Math., 156 (2024), pp. 609–639, https://doi.org/10.1007/s00211-024-01398-8.

[4] J. W. Barrett, K. Deckelnick, and R. Nürnberg, *A finite element error analysis for axisymmetric mean curvature flow*, IMA J. Numer. Anal., 41 (2021), pp. 1641–1667, https://doi.org/10.1093/imanum/draa020.

[5] 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 (2017), pp. 1080–1100, https://doi.org/10.1137/16M1083682.

[6] J. W. Barrett, H. Garcke, and R. Nürnberg, *A parametric finite element method for fourth order geometric evolution equations*, J. Comput. Phys., 222 (2007), pp. 441–467, https://doi.org/10.1016/j.jcp.2006.07.026.

[7] 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 (2008), pp. 4281–4307, https://doi.org/10.1016/j.jcp.2007.11.023.

[8] J. W. Barrett, H. Garcke, and R. Nürnberg, *Parametric approximation of Willmore flow and related geometric evolution equations*, SIAM J. Sci. Comput., 31 (2008), pp. 225–253, https://doi.org/10.1137/070700231.

[9] S. Bartels, *A simple scheme for the approximation of the elastic flow of inextensible curves*, IMA J. Numer. Anal., 33 (2013), pp. 1115–1125, https://doi.org/10.1093/imanum/drs041.

[10] J. Bergh and J. Löfström, *Interpolation Spaces: An Introduction*, Springer-Verlag, Berlin, 1976.

[11] T. Binz and B. Kovács, *A convergent finite element algorithm for mean curvature flow in arbitrary codimension*, Interfaces Free Bound., 25 (2023), 400, https://doi.org/10.4171/ifb/493.

[12] K. Deckelnick and G. Dziuk, *Convergence of a finite element method for non-parametric mean curvature flow*, Numer. Math., 72 (1995), pp. 197–222, https://doi.org/10.1007/s002110050166.

[13] K. Deckelnick and G. Dziuk, *Error estimates for a semi-implicit fully discrete finite element scheme for the mean curvature flow of graphs*, Interfaces Free Bound., 2 (2000), pp. 341–359, https://doi.org/10.4171/ifb/24.

[14] K. Deckelnick and G. Dziuk, *Error analysis of a finite element method for the Willmore flow of graphs*, Interfaces Free Bound., 8 (2006), pp. 21–46, https://doi.org/10.4171/ifb/134.
[15] K. Deckelnick and G. Dziuk, *Error analysis for the elastic flow of parametrized curves*, Math. Comp., 78 (2009), pp. 645–671, https://doi.org/10.1090/S0025-5718-08-02176-5.
[16] K. Deckelnick and G. Dziuk, *On the approximation of the curve shortening flow*, in Calculus of Variations, Applications and Computations (Pont-à-Mousson, 1994), Pitman Research Notes in Mathematics Series 326, Longman Scientific and Technical, Harlow, UK, 1995, pp. 100–108.
[17] 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 (2021), pp. 2698–2721, https://doi.org/10.1137/20M1374584.
[18] K. Deckelnick and V. Styles, *Finite element error analysis for a system coupling surface evolution to diffusion on the surface*, Interfaces Free Bound., 24 (2022), pp. 63–93, https://doi.org/10.4171/ifb/467.
[19] A. Demlow, *Higher-order finite element methods and pointwise error estimates for elliptic problems on surfaces*, SIAM J. Numer. Anal., 47 (2009), pp. 805–827, https://doi.org/10.1137/070708135.
[20] B. Duan and B. Li, *New artificial tangential motions for parametric finite element approximation of surface evolution*, SIAM J. Sci. Comput., 46 (2024), pp. A587–A608, https://doi.org/10.1137/23M1551857.
[21] G. Dziuk, *An algorithm for evolutionary surfaces*, Numer. Math., 58 (1990), pp. 603–611, https://doi.org/10.1007/BF01385643.
[22] G. Dziuk, *Convergence of a semi-discrete scheme for the curve shortening flow*, Math. Models Methods Appl. Sci., 04 (1994), pp. 589–606, https://doi.org/10.1142/S0218202594000339.
[23] G. Dziuk and C. M. Elliott, *Finite elements on evolving surfaces*, IMA J. Numer. Anal., 27 (2007), pp. 262–292, https://doi.org/10.1093/imanum/drl023.
[24] G. Dziuk, C. Lubich, and D. Mansour, *Runge-Kutta time discretization of parabolic differential equations on evolving surfaces*, IMA J. Numer. Anal., 32 (2011), pp. 394–416, https://doi.org/10.1093/imanum/drr017.
[25] C. Elliott, H. Garcke, and B. Kovács, *Numerical analysis for the interaction of mean curvature flow and diffusion on closed surfaces*, Numer. Math., 151 (2022), pp. 873–925, https://doi.org/10.1007/s00211-022-01301-3.
[26] C. M. Elliott and H. Fritz, *On algorithms with good mesh properties for problems with moving boundaries based on the harmonic map heat flow and the DeTurck trick*, SMAI J. Comput. Math., 2 (2016), pp. 141–176, https://doi.org/10.5802/smai-jcm.12.
[27] 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 (2017), pp. 543–603, https://doi.org/10.1093/imanum/drw020.
[28] X. Gui, B. Li and J. Wang, *Convergence of renormalized finite element methods for heat flow of harmonic maps*, SIAM J. Numer. Anal., 60 (2022), pp. 312–338, https://doi.org/10.1137/21M1402212.
[29] J. Hu and B. Li, *Evolving finite element methods with an artificial tangential velocity for mean curvature flow and Willmore flow*, Numer. Math., 152 (2022), pp. 127–181, https://doi.org/10.1007/s00211-022-01309-9.
[30] B. Kovács, *High-order evolving surface finite element method for parabolic problems on evolving surfaces*, IMA J. Numer. Anal., 38 (2018), pp. 430–459, https://doi.org/10.1093/imanum/drx013.
[31] B. Kovács, B. Li, and C. Lubich, *A convergent evolving finite element algorithm for mean curvature flow of closed surfaces*, Numer. Math., 143 (2019), pp. 797–853, https://doi.org/10.1007/s00211-019-01074-2.
[32] B. Kovács, B. Li, and C. Lubich, *A convergent evolving finite element algorithm for Willmore flow of closed surfaces*, Numer. Math., 149 (2021), pp. 595–643, https://doi.org/10.1007/s00211-021-01238-z.
[33] 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 (2017), pp. 643–689, https://doi.org/10.1007/s00211-017-0888-4.
[34] M. Lenoir, *Optimal iso-parametric finite elements and error estimates for domains involving curved boundaries*, SIAM J. Numer. Anal., 23 (1986), pp. 562–580, https://doi.org/10.1137/0723036.
[35] B. Li, *Convergence of Dziuk's linearly implicit parametric finite element method for curve shortening flow*, SIAM J. Numer. Anal., 58 (2020), pp. 2315–2333, https://doi.org/10.1137/19M1305483.

[36] 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 (2021), pp. 1592–1617, https://doi.org/10.1137/20M136935X.

[37] N. G. Meyers, *An L^p-estimate for the gradient of solutions of second order elliptic divergence equations*, Ann. Sc. Norm. Super. Pisa Cl. Sci., 17 (1963), pp. 189–206, http://www.numdam.org/item?id=ASNSP_1963_3_17_3_189_0.

[38] A. Mierswa, *Error Estimates for a Finite Difference Approximation of Mean Curvature Flow for Surfaces of Torus Type*, Ph.D. thesis, Otto-von-Guericke-Universität, Magdeburg, Germany, 2020.

[39] V. Thomee, *Galerkin Finite Element Methods for Parabolic Problems*, 2nd ed., Springer Series in Computational Mathematics 25, Springer-Verlag, Berlin, 2006.

[40] M. F. Wheeler, *A priori L_2 error estimates for Galerkin approximations to parabolic partial differential equations*, SIAM J. Numer. Anal., 10 (1973), pp. 723–759, https://doi.org/10.1137/0710062.

[41] C. Ye and J. Cui, *Convergence of Dziuk's fully discrete linearly implicit scheme for curve shortening flow*, SIAM J. Numer. Anal., 59 (2021), pp. 2823–2842, https://doi.org/10.1137/21M1391626.

[42] 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 (2020), pp. B327–B352, https://doi.org/10.1137/19M1281666.