

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

BUYANG LI AND RONG TANG<sup>†</sup>

**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 FEM, dynamic Ritz projection, convergence.

**AMS subject classifications.** 65M15, 65M60, 49M10, 35K65

**1. Introduction.** The numerical approximation of surface evolution under geometric flows, including mean curvature flow, Willmore flow and surface diffusion, have been intensively studied in the past. The most well-known example of geometric flows is mean curvature flow, which describe 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  $Hn = -\Delta_\Gamma \text{id}_\Gamma$ , where  $\Delta_\Gamma$  and  $\text{id}_\Gamma$  denote the Laplace–Beltrami operator and identity map on surface  $\Gamma$ , mean curvature flow can also be written as

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

By utilizing the formulation in (1.1), Dziuk introduced the following type of finite element method (FEM) in [22] for approximating surface evolution under mean curvature flow: assuming that  $\Gamma(t_{m-1})$  is already approximated by a piecewise triangular surface  $\Gamma_h^{m-1}$ , find a parametrization of surface  $\Gamma_h^m$  through a finite element function  $u_h^m : \Gamma_h^{m-1} \rightarrow \mathbb{R}^3$  which is determined by some weak formulation of (1.1). Such methods are referred to as parametric FEMs.

Since 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 & Nürnberg’s (BGN) [7–9], Elliott & Fritz [27, 28], Hu & Li [29], Duan & Li [21], and the structure-preserving parametric FEMs [3, 4, 43]. 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 [10, 14, 17, 23, 28, 35, 42], and for graph surfaces and axisymmetric surfaces in [5, 13, 16, 18, 19]. However, the techniques developed in the proofs are not applicable to the analysis of geometric flow of general closed

---

\*This work is supported in part by AMSS-PolyU Joint Laboratory and the Research Grants Council of Hong Kong (PolyU/GRF15303022 and PolyU/RFS2324-5S03).

<sup>†</sup>Department of Applied Mathematics, The Hong Kong Polytechnic University, Hong Kong. E-mail address: buyang.li@polyu.edu.hk and claire.tang@polyu.edu.hk

surfaces. We also refer to [5, 6, 18, 19, 28, 38] 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 & Fritz’s tangential motion [27, 28] generated by a tangential transformation on  $\Gamma^0$ .

Convergence of the parametric FEMs for mean curvature flow, Willmore flow and surface diffusion of closed surfaces was firstly proved in [12, 26, 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 [24] 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  $X(\cdot, t) : \Gamma^0 \rightarrow \mathbb{R}^3$  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.,

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

The analyses in [29, 31, 32] are restricted to finite elements of degree  $k \geq 2$ . 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  $L^p$  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 Laplacian–Beltrami operator  $\Delta_\Gamma \text{id}_\Gamma$  acting on surface  $\Gamma$ ; see the discussions in [2, 36]. In the spatially semi-discretization 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  $k \geq 6$  based on a discovery that the nonlinear Laplacian–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  $k \geq 6$  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  $k \geq 3$  based on a discovery that the nonlinear Laplacian–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 degree from  $k \geq 6$  to  $k \geq 3$ .

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  $k \geq 2$  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  $\Gamma_h^0$  be a piecewise polynomial surface that interpolates the smooth surface  $\Gamma^0$ , with each piece being the image of the reference triangle under a polynomial map of degree  $k \geq 1$ , and assume that the curved triangles are shape-regular and quasi-uniform with mesh size  $h$ ; see [20,30] and Lenoir's isoparametric approximation of a surface [34].

Let  $\mathbf{x}^0 = (p_1, \dots, p_N) \in \mathbb{R}^{3N}$  be the nodal vector that collects all the nodes  $p_j \in \mathbb{R}^3$ ,  $j = 1, \dots, N$ , in  $\Gamma_h^0$ . We evolve  $\mathbf{x}^0$  in time and denote its position at time  $t$  by  $\mathbf{x}(t) = (x_1(t), \dots, x_N(t))$ , 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

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

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

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

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

$$\partial_t X_h = v_h \circ X_h \quad \text{on } \Gamma_h[\mathbf{x}^0] \quad (2.1a)$$

$$v_h = -I_h(H_h n_h) \quad \text{on } \Gamma_h[\mathbf{x}] \quad (2.1b)$$

$$\int_{\Gamma_h[\mathbf{x}]} \partial_{t,h}^\bullet H_h \chi_H + \int_{\Gamma_h[\mathbf{x}]} \nabla_{\Gamma_h[\mathbf{x}]} H_h \cdot \nabla_{\Gamma_h[\mathbf{x}]} \chi_H = \int_{\Gamma_h[\mathbf{x}]} |\nabla_{\Gamma_h[\mathbf{x}]} n_h|^2 H_h \chi_H \quad (2.1c)$$

$$\int_{\Gamma_h[\mathbf{x}]} \partial_{t,h}^\bullet n_h \cdot \chi_n + \int_{\Gamma_h[\mathbf{x}]} \nabla_{\Gamma_h[\mathbf{x}]} n_h \cdot \nabla_{\Gamma_h[\mathbf{x}]} \chi_n = \int_{\Gamma_h[\mathbf{x}]} |\nabla_{\Gamma_h[\mathbf{x}]} n_h|^2 n_h \cdot \chi_n, \quad (2.1d)$$

where  $I_h$  denotes the Lagrange interpolation onto the finite element space  $S_h[\mathbf{x}(t)]$  and  $\partial_{t,h}^\bullet$  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:  $X_h(\cdot, 0) = \text{id}$  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  $X(\cdot, t) : \Gamma^0 \rightarrow \mathbb{R}^3$ , 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

$$\int_{\Gamma_h[\mathbf{x}^*]} (\hat{R}_h^* H \varphi_h + \nabla_{\Gamma_h[\mathbf{x}^*]} \hat{R}_h^* H \cdot \nabla_{\Gamma_h[\mathbf{x}^*]} \varphi_h) = \int_{\Gamma} (H \varphi_h^l + \nabla_{\Gamma} H \cdot \nabla_{\Gamma} \varphi_h^l)$$

$$\forall \varphi_h \in S_h(\Gamma_h[\mathbf{x}^*]), \quad (2.2a)$$

$$\int_{\Gamma_h[\mathbf{x}^*]} (\hat{R}_h^* n \cdot \phi_h + \nabla_{\Gamma_h[\mathbf{x}^*]} \hat{R}_h^* n \cdot \nabla_{\Gamma_h[\mathbf{x}^*]} \phi_h) = \int_{\Gamma} (n \cdot \phi_h^l + \nabla_{\Gamma} n \cdot \nabla_{\Gamma} \phi_h^l) \\ \forall \phi_h \in S_h(\Gamma_h[\mathbf{x}^*])^3, \quad (2.2b)$$

In (2.2),  $\varphi_h^l$  and  $\phi_h^l$  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.,

$$\partial_t \hat{I}_h^* X = \hat{I}_h^* v \circ \hat{I}_h^* X \quad \text{on } \Gamma_h[\mathbf{x}^0] \quad (2.3a)$$

$$\hat{I}_h^* v = -\hat{I}_h^* (\hat{R}_h^* H \hat{R}_h^* n) + d_v \quad \text{on } \Gamma_h[\mathbf{x}^*] \quad (2.3b)$$

$$\int_{\Gamma_h[\mathbf{x}^*]} \partial_{t,h}^* \hat{R}_h^* H \chi_H + \int_{\Gamma_h[\mathbf{x}^*]} \nabla_{\Gamma_h[\mathbf{x}^*]} \hat{R}_h^* H \cdot \nabla_{\Gamma_h[\mathbf{x}^*]} \chi_H \quad \forall \chi_H \in S_h(\Gamma_h[\mathbf{x}^*]) \\ = \int_{\Gamma_h[\mathbf{x}^*]} |\nabla_{\Gamma_h[\mathbf{x}^*]} \hat{R}_h^* n|^2 \hat{R}_h^* H \chi_H + \int_{\Gamma_h[\mathbf{x}^*]} d_H \chi_H \quad (2.3c)$$

$$\int_{\Gamma_h[\mathbf{x}^*]} \partial_{t,h}^* \hat{R}_h^* n \cdot \chi_n + \int_{\Gamma_h[\mathbf{x}^*]} \nabla_{\Gamma_h[\mathbf{x}^*]} \hat{R}_h^* n \cdot \nabla_{\Gamma_h[\mathbf{x}^*]} \chi_n \quad \forall \chi_n \in S_h(\Gamma_h[\mathbf{x}^*])^3 \\ = \int_{\Gamma_h[\mathbf{x}^*]} |\nabla_{\Gamma_h[\mathbf{x}^*]} \hat{R}_h^* n|^2 \hat{R}_h^* n \cdot \chi_n + \int_{\Gamma_h[\mathbf{x}^*]} d_n \cdot \chi_n, \quad (2.3d)$$

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

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

$$\left| \int_{\Gamma_h[\mathbf{x}^*]} d_H \chi_H \right| + \left| \int_{\Gamma_h[\mathbf{x}^*]} d_n \chi_n \right| \leq Ch^{k+1} (\|\chi_H\|_{H^1(\Gamma_h[\mathbf{x}^*])} + \|\chi_n\|_{H^1(\Gamma_h[\mathbf{x}^*])}). \quad (2.5)$$

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  $\Gamma_h[\mathbf{x}^*]$ :

$$\hat{e}_{x,h} = \hat{x}_h - \text{id}_{\Gamma_h[\mathbf{x}^*]}, \quad \hat{e}_{v,h} = \hat{v}_h - \hat{I}_h^* v, \\ \hat{e}_{H,h} = \hat{H}_h - \hat{R}_h^* H, \quad \hat{e}_{n,h} = \hat{n}_h - \hat{R}_h^* n, \quad (2.6)$$

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  $k \geq 2$ .

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 [41], 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 that 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

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

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 super-convergence rate in  $L^2(0, T; H^1)$  norm:

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

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

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

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

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

This can be used to prove the uniform  $W^{1,\infty}$  boundedness of  $\hat{e}_{x,h}$  for  $k \geq 1$ . 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  $Y_h^*(\cdot, t) : \Gamma_h^0 \rightarrow \mathbb{R}^3$  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:

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

where  $I_h^*$  denotes the Lagrange interpolation operator onto  $S_h(\Gamma_h[\mathbf{y}^*])$ , and  $\varphi_h^l$  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  $\Gamma$ , and  $\varphi_h^*$  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  $Y_h^*(\cdot, 0) = \text{id}$  on  $\Gamma_h^0$ . 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 [15, 16, 19]. 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 the flow map  $X(\cdot, t) : \Gamma^0 \rightarrow \Gamma(t)$  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  $k \geq 1$ . Then there exists a constant  $h_0 > 0$  such that the following error bound holds for mesh size  $h \leq h_0$ :*

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

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  $t \in [0, T]$  (though it 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.,

$$\partial_t Y_h^* = v_h^* \circ Y_h^* \quad \text{on } \Gamma_h^0 \quad (2.13a)$$

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

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

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

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

$$\partial_t Y_h^* = v_h^* \circ Y_h^* \quad \text{on } \Gamma_h^0 \quad (2.14a)$$

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

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

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):

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

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

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

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 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:*

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

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

The error estimates in Theorem 2.1 and Theorem 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  $k \geq 1$ . Before stating this result, we list some basic notations for finite element functions on  $\Gamma_h[\mathbf{x}]$ ,  $\Gamma_h[\mathbf{x}^*]$  and  $\Gamma$ .

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  $k \geq 1$ , satisfies the following error bound:*

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

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

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 Theorem 2.1 and Theorem 2.2. The proof of Theorem 2.3 is standard and therefore omitted (in fact, (2.17) follows from Theorem 2.1 and Theorem 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 provide essential results and detailed proofs of Lemmas 3.5–3.6, which are used for proving Theorems 2.1 and 2.2. These proofs, following similar techniques as those in Lemmas 3.3–3.4, 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  $u : \Gamma \rightarrow \mathbb{R}$  is a column vector denoted by  $\nabla_\Gamma u$ . For a vector-valued function  $u = (u_1, u_2, u_3)^\top : \Gamma \rightarrow \mathbb{R}^3$ , we define

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

We denote by  $\text{id}$  the identity function on  $\mathbb{R}^3$ , i.e.,  $\text{id}(x) = x$  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:

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

This definition of basis functions implies the following transport property (see [20]):

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

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 [20, (2.15)-(2.16)] we know that, there exists  $h_0 > 0$  such that for  $h \leq h_0$  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 relation

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

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

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

which 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 [29, 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

$$\begin{aligned} \mathbf{f}_1(\mathbf{x}, \mathbf{n})_{j+(m-1)N} &= \int_{\Gamma_h[\mathbf{x}]} |\nabla_{\Gamma_h[\mathbf{x}]} n_h|^2 (n_h)_m \phi_j \\ \mathbf{f}_2(\mathbf{x}, \mathbf{n}, \mathbf{H})_j &= \int_{\Gamma_h[\mathbf{x}]} |\nabla_{\Gamma_h[\mathbf{x}]} n_h|^2 H_h \phi_j, \end{aligned}$$

with  $j = 1, \dots, N$  and  $m = 1, 2, 3$ , where  $(n_h)_m$  denotes the  $m$ th 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:

$$\dot{\mathbf{x}} = \mathbf{v} \quad (3.4a)$$

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

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

where

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

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:

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

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

where  $\boldsymbol{\varphi}$  is the nodal vector of finite element function  $\varphi_h$ . 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:

$$\dot{\mathbf{y}}^* = \mathbf{v}^* \quad (3.7)$$

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

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

where  $\mathbf{d}_{\mathbf{u}}^*$  denotes the nodal vector of the finite element function  $d_{\mathbf{u}}^* = (d_H^*, (d_n^*)^{\top})^{\top}$ , 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  $\mathbf{e}_{\mathbf{y}} = \mathbf{y}^* - \mathbf{x}^*$ , which is the nodal vector of the finite element function  $\hat{e}_{\mathbf{y}} = \hat{y}_h^* - \text{id}$  on  $\Gamma_h[\mathbf{x}^*]$ , we consider the following intermediate surfaces

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

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

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

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

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

where  $c_p$  is a constant independent of  $\theta$  and  $h$ , and  $c_{\infty} = 2$ .

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

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

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

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

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

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

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

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

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

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

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

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

In addition to Lemma 3.1 and Lemma 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

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

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  $h \leq h_0$  there exists a unique solution  $(\mathbf{y}^*, \mathbf{H}^*, \mathbf{n}^*)$  of (3.6) satisfying the following estimates for all  $2 \leq p < \infty$ :*

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

*Proof.* Problem (3.6) is essentially a system of ordinary differential equations (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 hold for  $t \in [0, t_*]$ :

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

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

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

Under this condition, we shall prove that (3.19) holds for  $t \in [0, t_*]$  and all  $2 \leq p < \infty$  (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  $2C_4 h^{k-2/4} \leq 1/2$ , 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

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

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:

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

where  $\varphi$  denotes the inverse lift of  $\varphi^l$  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  $1 \leq p \leq \infty$  (see Appendix B in the supplementary material) and the  $L^p$  norms of  $\varphi$  and  $\varphi^l$  are equivalent for  $1 \leq p \leq \infty$ , it follows that

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

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

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

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

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

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

$$\begin{aligned} |\ell(\varphi^l)| &\leq C \|\hat{u}_h^*\|_{W^{1,\infty}(\Gamma_h[\mathbf{x}^*])} \|\text{id}_{\Gamma_h[\mathbf{x}^*]} - \hat{y}_h^*\|_{W^{1,p}(\Gamma_h[\mathbf{x}^*])} \|\varphi\|_{W^{1,p'}(\Gamma_h[\mathbf{x}^*])} \\ &\leq C \|\text{id}_{\Gamma_h[\mathbf{x}^*]} - \hat{y}_h^*\|_{W^{1,p}(\Gamma_h[\mathbf{x}^*])} \|\varphi^l\|_{W^{1,p'}(\Gamma)}. \end{aligned}$$

This means that

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

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

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

*Estimates for  $L^2$  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 [20, Corollaries 4.2 and 4.5]:

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

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

Since the complex interpolation spaces between  $H^{k+1}(\Gamma)$  and  $W^{k+1,\infty}(\Gamma)$  are  $W^{k+1,p}(\Gamma)$  for  $2 \leq p \leq \infty$  (cf. [11, Theorem 6.4.5]), the complex interpolation (cf. [11, Theorem 4.1.2]) of the above two estimates yields, for  $2 \leq p \leq \infty$ ,

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

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

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

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

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

The complex interpolation between (3.29) and (3.30) yields the following result (cf. [11, Theorem 4.1.2]) for  $2 \leq p \leq \infty$ :

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

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

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

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

*Estimates for  $L^2$  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

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

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

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

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

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

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

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

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

which can be written into the integral form:

$$X_h^*(s) - Y_h^*(s) = -\int_0^s \hat{I}_h^*(\hat{I}_h^*H\hat{I}_h^*n - \hat{H}_h^*\hat{n}_h^*) \circ X_h^* dt. \quad (3.38)$$

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

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

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

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

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:

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

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

$$\begin{aligned} \|\text{id}_{\Gamma_h[\mathbf{x}^*(s)]} - \hat{y}_h^*\|_{W^{1,p}(\Gamma[\mathbf{x}^*(s)])} &\leq C \int_0^s (\|\hat{I}_h^*H - \hat{H}_h^*\|_{W^{1,p}(\Gamma_h[\mathbf{x}^*])} + \|\hat{I}_h^*n - \hat{n}_h^*\|_{W^{1,p}(\Gamma_h[\mathbf{x}^*])}) dt \\ &\leq C \int_0^s (\|H - (\hat{H}_h^*)^l\|_{W^{1,p}(\Gamma)} + \|n - (\hat{n}_h^*)^l\|_{W^{1,p}(\Gamma)}) dt + Ch^k \end{aligned}$$

$$\leq C \int_0^s \|u - (\hat{u}_h^*)^l\|_{W^{1,p}(\Gamma)} dt + Ch^k \quad \text{for } s \in (0, t_*]. \quad (3.41)$$

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

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

Since  $\|(\hat{I}_h^* u)^l - (\hat{u}_h^*)^l\|_{W^{1,p}(\Gamma)} \leq \|(\hat{I}_h^* u)^l - u\|_{W^{1,p}(\Gamma)} + \|(\hat{u}_h^*)^l - u\|_{W^{1,p}(\Gamma)} \leq Ch^k$  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

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

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

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

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

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

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).  $\square$

### 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  $h \leq h_0$ , (3.6) has a unique solution  $(\mathbf{y}^*, \mathbf{H}^*, \mathbf{n}^*)$ , which satisfies the following estimate:*

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

*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)–(3.16), along with the results in Lemma 3.3 and the inverse inequality for finite element functions, we immediately get the following estimate:

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

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:

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

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

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

Substituting this into (3.34) yields

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

It remains to estimate  $\|w\|_{L^2(\Gamma)}$  and  $\|\text{id}_{\Gamma_h[\mathbf{x}^*]} - \hat{g}_h^*\|_{L^2(\Gamma_h[\mathbf{x}^*])}$ . 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

$$\varphi^l - \Delta_\Gamma \varphi^l = w \text{ on } \Gamma, \quad (3.47)$$

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

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

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  $\mathbf{e}_y := \mathbf{y}^* - \mathbf{x}^*$ , 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:

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

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

$$\begin{aligned} \|w\|_{L^2(\Gamma)}^2 &= \int_\Gamma (w\varphi^l + \nabla_\Gamma w \cdot \nabla_\Gamma \varphi^l) \\ &= (\mathbf{K}(\mathbf{x}^*) - \mathbf{K}(\mathbf{y}^*))\mathbf{u}^* \cdot (\mathbf{P}_h\varphi) \\ &= \left| \int_0^1 \frac{d}{d\theta} \int_{\Gamma_h[\mathbf{y}^\theta]} (\hat{u}_h^{*,\theta} \cdot \hat{\varphi}_h^\theta + \nabla_{\Gamma_h[\mathbf{y}^\theta]} \hat{u}_h^{*,\theta} \cdot \nabla_{\Gamma_h[\mathbf{y}^\theta]} \hat{\varphi}_h^\theta) \right| \\ &= \left| \int_0^1 \int_{\Gamma_h[\mathbf{y}^\theta]} (\hat{u}_h^{*,\theta} \cdot \hat{\varphi}_h^\theta \nabla_{\Gamma_h[\mathbf{y}^\theta]} \cdot \hat{e}_y^\theta + \nabla_{\Gamma_h[\mathbf{y}^\theta]} \hat{u}_h^{*,\theta} \cdot D_{\Gamma_h[\mathbf{y}^\theta]} \hat{e}_y^\theta \nabla_{\Gamma_h[\mathbf{y}^\theta]} \hat{\varphi}_h^\theta) d\theta \right| \\ &\hspace{15em} \text{(see Lemma 3.2)} \\ &= |B_0 + B_1 + B_2 + B_3|, \end{aligned} \quad (3.49)$$

where

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

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

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

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

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:

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

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

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

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. [20, Proposition 2.3]), we can derive the following result:

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

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

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

for fixed  $\theta \in [0,1]$ . As  $\alpha$  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  $\|\hat{e}_y^0\|_{W^{1,p}(\Gamma_h[\mathbf{x}^*])} \leq Ch$  from (3.19) in Lemma 3.3, we can obtain (details are omitted)

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

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

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

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

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

Then, substituting this result into (3.46), we obtain

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

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

$$\|\text{id}_{\Gamma_h[\mathbf{x}^*(s)]} - \hat{y}_h^*\|_{L^2(\Gamma[\mathbf{x}^*(s)])} \leq C \int_0^s \|(\hat{u}_h^*)^l(t) - u(t)\|_{L^2(\Gamma)} dt + Ch^{k+1}, \quad \text{for } s \in [0, T]. \quad (3.60)$$

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

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

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

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

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

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

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.  $\square$

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

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

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 equation (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  $h \leq h_0$ , the solution  $(\mathbf{y}^*, \mathbf{H}^*, \mathbf{n}^*)$  of (3.6) satisfies the following estimates:*

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

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

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  $\hat{\mathbf{H}}^*$ .

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 as those for proving Lemma 3.3 and Lemma 3.4. Therefore, we omit the proof here and refer the readers to Appendix C 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 Lemma 3.3–3.5. This result is presented in the following lemma and the proof is omitted. We refer the readers to Appendix D 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  $h \leq h_0$ , the remainder  $d_u^*$  defined in (2.14) satisfies the following estimate:*

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

**4. Proof of Theorem 2.2 and Theorem 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 Lemma 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 equations (3.4) and (3.7), respectively. By subtracting (3.7) from (3.4), we find that the errors 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:

$$\dot{\mathbf{e}}_x = \mathbf{e}_v \quad (4.1a)$$

$$\mathbf{e}_v = -(\mathbf{I}_h(\mathbf{H} \bullet \mathbf{n}) - \mathbf{I}_h(\mathbf{H}^* \bullet \mathbf{n}^*)) \quad (4.1b)$$

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

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  $e_x^\theta$  as parameter  $\theta \in [0, 1]$  changes, and denote by  $e_x^\theta$ ,  $e_v^\theta$  and  $e_u^\theta$  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 as 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:

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

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

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

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

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

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  $1 \leq p \leq \infty$  (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  $h \leq h_0$ , the following estimate holds:*

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

**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) holds 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  $h \leq h_0$  the following estimate holds for  $t \in [0, t^*]$ :*

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

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

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

where  $e_u^1$  denotes the finite element function on  $\Gamma_h[\mathbf{x}^\theta]$  with  $\theta = 1$ . Additionally,  $\partial_\theta^\bullet e_x^\theta = \partial_\theta^\bullet e_u^\theta = 0$  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  $L^p$  and  $W^{1,p}$  norms on  $\Gamma_h[\mathbf{x}^\theta]$  and  $\Gamma_h[\mathbf{y}^*]$ , and Lemma 3.4–3.5, we have

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

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

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

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

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

Lemma 3.6 guarantees that

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

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:

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

where  $n_{\Gamma_h[\mathbf{x}^\theta]}$  is the unit normal vector of  $\Gamma_h[\mathbf{x}^\theta]$ , while  $n_h^\theta$  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:

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

Then, with norm equivalence of  $L^p$  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  $I_{31j}$ ,  $j = 1, \dots, 6$ , can be derived:

$$\begin{aligned} I_{311} &\leq C \|e_u\|_{L^2(\Gamma_h[\mathbf{y}^*])} \|\nabla_{\Gamma_h[\mathbf{y}^*]} e_x\|_{L^2(\Gamma_h[\mathbf{y}^*])}, \\ I_{312} &\leq C \|e_u\|_{L^2(\Gamma_h[\mathbf{y}^*])} \|\nabla_{\Gamma_h[\mathbf{y}^*]} e_x\|_{L^2(\Gamma_h[\mathbf{y}^*])} \|e_u\|_{L^\infty(\Gamma_h[\mathbf{y}^*])} \end{aligned}$$

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

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

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

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

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

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

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

These estimates lead to

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

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

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

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

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

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

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

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

**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  $h \leq h_0$ , the following estimate holds for  $t \in [0, t^*]$ :*

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

*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 Appendix A and the norm equivalence between surfaces  $\Gamma_h[\mathbf{x}^*]$  and  $\Gamma_h[\mathbf{y}^*]$ )

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

where (4.2b) and (4.6) are used in the last inequality.  $\square$

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

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

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

$$\begin{aligned} &\|e_x(t)\|_{H^1(\Gamma_h[\mathbf{y}^*(t)])}^2 \\ &= \int_0^t \frac{d}{ds} \|e_x(s)\|_{H^1(\Gamma_h[\mathbf{y}^*(s)])}^2 ds \\ &= \int_0^t \left( 2\mathbf{e}_x(s)^\top \mathbf{K}(\mathbf{y}^*(s)) \dot{\mathbf{e}}_x(s) + \mathbf{e}_x(s)^\top \frac{d}{ds} \mathbf{K}(\mathbf{y}^*(s)) \mathbf{e}_x(s) \right) ds \\ &\leq C \int_0^t \left( \|e_x(s)\|_{H^1(\Gamma_h[\mathbf{y}^*(s)])} \|e_v(s)\|_{H^1(\Gamma_h[\mathbf{y}^*(s)])} + \|e_x(s)\|_{H^1(\Gamma_h[\mathbf{y}^*(s)])}^2 \right) ds \\ &\leq C \int_0^t \|e_v(s)\|_{H^1(\Gamma_h[\mathbf{y}^*(s)])}^2 ds + C \int_0^t \|e_x(s)\|_{H^1(\Gamma_h[\mathbf{y}^*(s)])}^2 ds, \end{aligned} \quad (4.24)$$

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

$$\begin{aligned} \left| \mathbf{e}_x(s)^\top \frac{d}{ds} \mathbf{K}(\mathbf{y}^*(s)) \mathbf{e}_x(s) \right| &\leq C \|e_x(s)\|_{H^1(\Gamma_h[\mathbf{y}^*(s)])} \|v_h^*(s)\|_{W^{1,\infty}(\Gamma_h[\mathbf{y}^*(s)])} \\ &\leq C \|e_x(s)\|_{H^1(\Gamma_h[\mathbf{y}^*(s)])}^2, \end{aligned}$$

in which 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, inverse inequality and the  $W^{1,p}$  estimate in (3.19) with  $p = 2$ , i.e.,

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

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

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

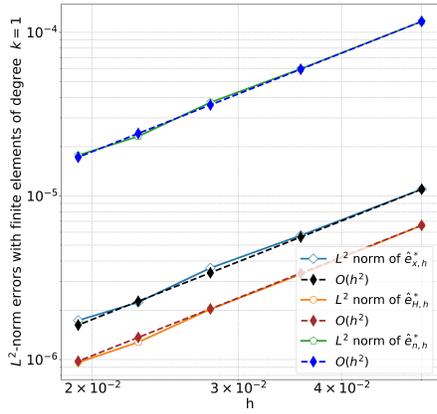
Then, summing up (4.4) and (4.25), we obtain the following result for  $t \in (0, t_*)$ :

$$\begin{aligned} & \|e_x(t)\|_{H^1(\Gamma_h[\mathbf{y}^*(t)])}^2 + \|e_u(t)\|_{L^2(\Gamma_h[\mathbf{y}^*(t)])}^2 + \int_0^t \|\nabla_{\Gamma_h[\mathbf{y}^*]} e_u(s)\|_{L^2(\Gamma_h[\mathbf{y}^*(s)])}^2 ds \\ & \leq C \int_0^t (\|e_x(s)\|_{H^1(\Gamma_h[\mathbf{y}^*(s)])}^2 + \|e_u(s)\|_{L^2(\Gamma_h[\mathbf{y}^*(s)])}^2) ds + Ch^{2k+2}. \end{aligned} \quad (4.26)$$

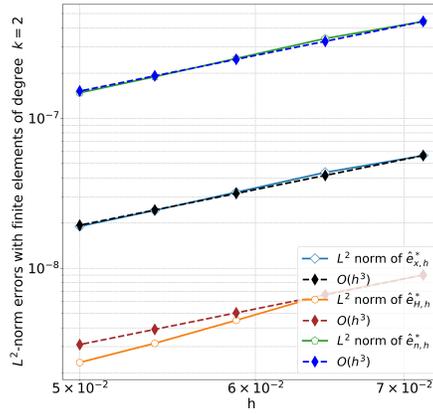
The result of Proposition 4.1 follows from applying Gronwall's inequality to (4.26). Moreover, the discussions in Remark 4.1 show that  $t^* = T$ . This completes the proof of Theorem 2.2.  $\square$

**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).

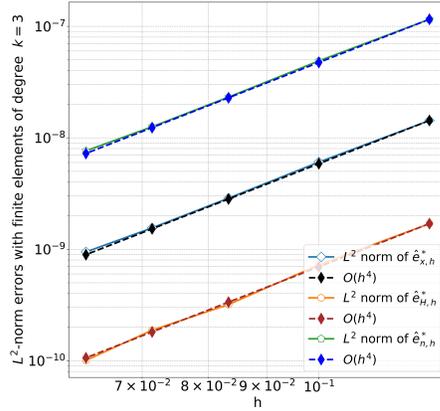
We consider the evolution of the two-dimensional sphere  $\Gamma(t)$  under mean curvature flow, which was used for testing the convergence rates of numerical methods for mean curvature flow in [31]. The exact solution of the surface at time  $t > 0$  is a sphere of radius  $R(t) = \sqrt{R(0)^2 - 4t}$  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.



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



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



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

Fig. 5.1.  $L^2$  norm for errors and convergence rates of the numerical solutions.

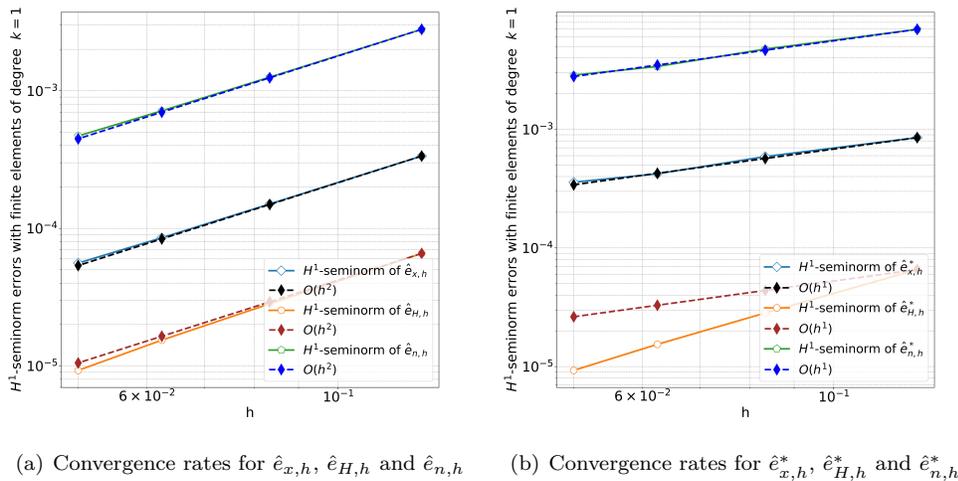


Fig. 5.2. The  $H^1$ -seminorm of the errors for finite elements of degree  $k = 1$

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 (BDF), with a sufficiently small time stepsize  $\tau = 0.001$  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  $\hat{e}_{x,h}^* = \hat{x}_h - \text{id}_{\Gamma_h[\mathbf{x}^*]}$ ,  $\hat{e}_{H,h}^* = \hat{H}_h - \hat{I}_h^* H$  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).

**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.

**Acknowledgement.** The authors would like to thank Prof. Christian Lubich for helpful discussions and valuable comments in an early stage of the research. The research was supported in part by the AMSS-PolyU Joint Laboratory, the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No. PolyU/RFS2324-5S03, PolyU/GRF15303022), and an internal grant of The Hong Kong Polytechnic University (Project ID: P0051154).

## REFERENCES

- [1] G. Bai and B. Li. Erratum: Convergence of Dziuk’s semidiscrete finite element method for mean

- curvature flow of closed surfaces with high-order finite elements. *SIAM J. Numer. Anal.*, 61(3):1609–1612, 2023.
- [2] G. Bai and B. Li. A new approach to the analysis of parametric finite element approximations to mean curvature flow. *Found. Comput. Math.*, 2023, 64 pages, DOI: <https://doi.org/10.1007/s10208-023-09622-x>.
  - [3] W. Bao, W. Jiang, and Q. Zhao. A parametric finite element method for solid-state dewetting problems in three dimensions. *SIAM J. Sci. Comput.*, 42:B327–B352, 2020.
  - [4] W. Bao and Y. Li. A structure-preserving parametric finite element method for geometric flows with anisotropic surface energy. *Numer. Math.*, 156:609–639, 2024.
  - [5] J. Barrett, K. Deckelnick, and R. Nürnberg. A finite element error analysis for axisymmetric mean curvature flow. *IMA J. Numer. Anal.*, 41(3):1641–1667, 2021.
  - [6] J. W. Barrett, K. Deckelnick, and V. Styles. Numerical analysis for a system coupling curve evolution to reaction diffusion on the curve. *SIAM J. Numer. Anal.*, 55(2):1080–1100, 2017.
  - [7] J. W. Barrett, H. Garcke, and R. Nürnberg. A parametric finite element method for fourth order geometric evolution equations. *J. Comput. Phys.*, 222:441–467, 2007.
  - [8] J. W. Barrett, H. Garcke, and R. Nürnberg. On the parametric finite element approximation of evolving hypersurfaces in  $\mathbb{R}^3$ . *J. Comput. Phys.*, 227:4281–4307, 2008.
  - [9] J. W. Barrett, H. Garcke, and R. Nürnberg. Parametric approximation of willmore flow and related geometric evolution equations. *SIAM Journal on Scientific Computing*, 31(1):225–253, 2008.
  - [10] S. Bartels. A simple scheme for the approximation of the elastic flow of inextensible curves. *IMA J. Numer. Anal.*, 33:1115–1125, 2013.
  - [11] J. Bergh and J. Löfström. *Interpolation Spaces. An Introduction*. Springer-Verlag, Berlin, Germany, 1976.
  - [12] T. Binz and B. Kovács. A convergent finite element algorithm for mean curvature flow in arbitrary codimension. *Interfaces and Free Boundaries*, 25, 08 2023.
  - [13] K. Deckelnick and G. Dziuk. Convergence of a finite element method for non-parametric mean curvature flow. *Numer. Math.*, 72:197–222, 1995.
  - [14] K. Deckelnick and G. Dziuk. On the approximation of the curve shortening flow. In *Calculus of variations, applications and computations (Pont-à-Mousson, 1994)*, volume 326 of *Pitman Res. Notes Math. Ser.*, pages 100–108. Longman Sci. Tech., Harlow, 1995.
  - [15] K. Deckelnick and G. Dziuk. Error estimates for a semi-implicit fully discrete finite element scheme for the mean curvature flow of graphs. *Interfaces and Free Boundaries*, pages 341–359, 2000.
  - [16] K. Deckelnick and G. Dziuk. Error analysis of a finite element method for the Willmore flow of graphs. *Interfaces Free Bound.*, 8:21–46, 2006.
  - [17] K. Deckelnick and G. Dziuk. Error analysis for the elastic flow of parametrized curves. *Math. Comp.*, 78:645–671, 2009.
  - [18] K. Deckelnick and R. Nürnberg. Error analysis for a finite difference scheme for axisymmetric mean curvature flow of genus-0 surfaces. *SIAM J. Numer. Anal.*, 59(5):2698–2721, 2021.
  - [19] K. Deckelnick and V. Styles. Finite element error analysis for a system coupling surface evolution to diffusion on the surface. *Interfaces Free Bound.*, 24:63–93, 2022.
  - [20] A. Demlow. Higher-order finite element methods and pointwise error estimates for elliptic problems on surfaces. *SIAM J. Numer. Anal.*, 47(2):805–827, 2009.
  - [21] B. Duan and B. Li. New artificial tangential motions for parametric finite element approximation of surface evolution. *SIAM J. Sci. Comput.*, 45:A587–A608, 2024.
  - [22] G. Dziuk. An algorithm for evolutionary surfaces. *Numer. Math.*, 58:603–611, 1990.
  - [23] G. Dziuk. Convergence of a semi-discrete scheme for the curve shortening flow. *Math. Models Methods Appl. Sci.*, 4:589–606, 1994.
  - [24] G. Dziuk and C. M. Elliott. Finite elements on evolving surfaces. *IMA J. Numer. Anal.*, 27:262–292, 2007.
  - [25] G. Dziuk, C. Lubich, and D. Mansour. Runge-Kutta time discretization of parabolic differential equations on evolving surfaces. *IMA J. Numer. Anal.*, 32:394–416, 2011.
  - [26] C. Elliott, H. Garcke, and B. Kovács. Numerical analysis for the interaction of mean curvature flow and diffusion on closed surfaces. *Numer. Math.*, 151:873–925, 2022.
  - [27] 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.*, 2:141–176, 2016.
  - [28] C. M. Elliott and H. Fritz. On approximations of the curve shortening flow and of the mean curvature flow based on the DeTurck trick. *IMA J. Numer. Anal.*, 37:543–603, 2017.
  - [29] J. Hu and B. Li. Evolving finite element methods with an artificial tangential velocity for mean curvature flow and willmore flow. *Numer. Math.*, 152:127–181, 2022.

- [30] B. Kovács. High-order evolving surface finite element method for parabolic problems on evolving surfaces. *IMA J. Numer. Anal.*, 38(1):430–459, 2018.
- [31] B. Kovács, B. Li, and C. Lubich. A convergent evolving finite element algorithm for mean curvature flow of closed surfaces. *Numer. Math.*, 143:797–853, 2019.
- [32] B. Kovács, B. Li, and C. Lubich. A convergent evolving finite element algorithm for Willmore flow of closed surfaces. *Numer. Math.*, 149:595–643, 2021.
- [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.*, 137:643–689, 2017.
- [34] M. Lenoir. Optimal iso-parametric finite elements and error estimates for domains involving curved boundaries. *SIAM J. Numer. Anal.*, 23:562–580, 1986.
- [35] B. Li. Convergence of Dziuk’s linearly implicit parametric finite element method for curve shortening flow. *SIAM J. Numer. Anal.*, 58:2315–2333, 2020.
- [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.*, 59:1592–1617, 2021.
- [37] N. G. Meyers. An  $L^p$ -estimate for the gradient of solutions of second order elliptic divergence equations. *Annali della Scuola Normale Superiore di Pisa - Classe di Scienze*, 17(3):189–206, 1963.
- [38] A. Mierswa. *Error estimates for a finite difference approximation of mean curvature flow for surfaces of torus type*. phdthesis, PhD Thesis, Otto-von-Guericke-Universität, Magdeburg, 2020.
- [39] V. Thomee. *Galerkin Finite Element Methods for Parabolic Problems*. Number 25 in Springer Series in Computational Mathematics. Springer-Verlag Berlin Heidelberg, second edition, 2006.
- [40] M. F. Wheeler. A priori  $L_2$  error estimates for Galerkin approximations to parabolic partial differential equations. *SIAM J. Numer. Anal.*, 10(4):723–759, 1973.
- [41] B. L. X. Gui and J. Wang. Convergence of renormalized finite element methods for heat flow of harmonic maps. *SIAM J. Numer. Anal.*, 60:312–338, 2022.
- [42] C. Ye and J. Cui. Convergence of Dziuk’s fully discrete linearly implicit scheme for curve shortening flow. *SIAM J. Numer. Anal.*, 59:2823–2842, 2021.
- [43] Q. Zhao, W. Jiang, and W. Bao. A parametric finite element method for solid-state dewetting problems in three dimensions. *SIAM J. Sci. Comput.*, 42:B327–B352, 2020.

## Supplementary material

This supplementary material complements the paper by providing some fundamental results that will be used in this paper and the detailed proofs of Lemmas 3.5–3.6, which are omitted in the paper as they are based on the same techniques used in the proofs of Lemmas 3.3–3.4.

### Appendix A. $L^p$ and $W^{1,p}$ stability of interpolation operator on finite element functions of higher degree.

Let  $\hat{I}_h^* : C(\Gamma_h[\mathbf{x}^*]) \rightarrow S_h(\Gamma_h[\mathbf{x}^*])$  be the Lagrange interpolation operator onto the finite element space of polynomial degree  $k \geq 1$ , and let  $S_h^r[\mathbf{x}^*]$  be the space of finite element functions of polynomial degree  $r > k$ . In the proof of Lemma 3.3 and 3.4 we have used the following stability results:

$$\begin{aligned} \|\hat{I}_h^* \varphi_h\|_{L^p(\Gamma_h[\mathbf{x}^*])} &\leq C \|\varphi_h\|_{L^p(\Gamma_h[\mathbf{x}^*])} && \text{for } \varphi_h \in S_h^r[\mathbf{x}^*], \quad 1 \leq p \leq \infty, \\ \|\hat{I}_h^* \varphi_h\|_{W^{1,p}(\Gamma_h[\mathbf{x}^*])} &\leq C \|\varphi_h\|_{W^{1,p}(\Gamma_h[\mathbf{x}^*])} && \text{for } \varphi_h \in S_h^r[\mathbf{x}^*], \quad 1 \leq p \leq \infty, \end{aligned} \quad (\text{A.1})$$

which can be proved as follows: First,

$$\|\hat{I}_h^* \varphi_h - \varphi_h\|_{L^p(\Gamma_h[\mathbf{x}^*])} \leq Ch^{k+1} \|\nabla_{\Gamma_h[\mathbf{x}^*]}^{k+1} \varphi_h\|_{L^p(\Gamma_h[\mathbf{x}^*])} \leq C \|\varphi_h\|_{L^p(\Gamma_h[\mathbf{x}^*])},$$

where the second to last inequality follows from the standard error estimate of Lagrange interpolation, and the last inequality follows from the inverse inequality of finite element functions in  $S_h^r[\mathbf{x}^*]$ . Then, by using the triangle inequality, we have

$$\|\hat{I}_h^* \varphi_h\|_{L^p(\Gamma_h[\mathbf{x}^*])} \leq \|\hat{I}_h^* \varphi_h - \varphi_h\|_{L^p(\Gamma_h[\mathbf{x}^*])} + \|\varphi_h\|_{L^p(\Gamma_h[\mathbf{x}^*])} \leq C \|\varphi_h\|_{L^p(\Gamma_h[\mathbf{x}^*])}.$$

The  $W^{1,p}$  stability of  $\hat{I}_h^*$  can be proved similarly.

### Appendix B. $L^p$ and $W^{1,p}$ stability of $L^2$ projection onto $\Gamma_h[\mathbf{x}^*]$ .

The  $L^2$  projection operator  $P_h : L^2(\Gamma_h[\mathbf{x}^*]) \rightarrow S_h(\Gamma_h[\mathbf{x}^*])$  is defined as

$$\int_{\Gamma_h[\mathbf{x}^*]} (\varphi - P_h \varphi) \chi_h = 0 \quad \forall \varphi \in L^2(\Gamma_h[\mathbf{x}^*]) \text{ and } \chi_h \in S_h(\Gamma_h[\mathbf{x}^*]).$$

It is naturally bounded in the  $L^2$  norm, i.e.,  $\|P_h \varphi\|_{L^2(\Gamma_h[\mathbf{x}^*])} \leq C \|\varphi\|_{L^2(\Gamma_h[\mathbf{x}^*])}$  for  $\varphi \in L^2(\Gamma_h[\mathbf{x}^*])$ . Additionally, by the same proof of [39, Lemma 6.1], it is shown that the  $L^2$  projection  $P_h$  is bounded in the  $L^\infty$  norm, i.e.,  $\|P_h \varphi\|_{L^\infty(\Gamma_h[\mathbf{x}^*])} \leq C \|\varphi\|_{L^\infty(\Gamma_h[\mathbf{x}^*])}$ . Then, by the real interpolation between the  $L^2$  and  $L^\infty$  stability estimates, we derive the  $L^p$  stability of  $P_h$  for  $2 \leq p \leq \infty$ , i.e.,

$$\|P_h \varphi\|_{L^p(\Gamma_h[\mathbf{x}^*])} \leq C \|\varphi\|_{L^p(\Gamma_h[\mathbf{x}^*])} \quad \forall \varphi \in L^p(\Gamma_h[\mathbf{x}^*]) \text{ and } 2 \leq p \leq \infty.$$

The  $W^{1,\infty}$  stability of  $P_h$  can be shown by using the inverse inequality of finite element functions, the  $L^\infty$  stability of  $P_h$ , and the approximation property of the Lagrange interpolation, i.e.,

$$\begin{aligned} \|P_h \varphi - I_h \varphi\|_{W^{1,\infty}(\Gamma_h[\mathbf{x}^*])} &\leq Ch^{-1} \|P_h \varphi - I_h \varphi\|_{L^\infty(\Gamma_h[\mathbf{x}^*])} \\ &= Ch^{-1} \|P_h(\varphi - I_h \varphi)\|_{L^\infty(\Gamma_h[\mathbf{x}^*])} \\ &\leq Ch^{-1} \|\varphi - I_h \varphi\|_{L^\infty(\Gamma_h[\mathbf{x}^*])} \\ &\leq C \|\varphi\|_{W^{1,\infty}(\Gamma_h[\mathbf{x}^*])}, \end{aligned} \quad (\text{B.1})$$

which implies that

$$\|P_h \varphi\|_{W^{1,\infty}(\Gamma_h[\mathbf{x}^*])} \leq \|P_h \varphi - I_h \varphi\|_{W^{1,\infty}(\Gamma_h[\mathbf{x}^*])} + \|I_h \varphi\|_{W^{1,\infty}(\Gamma_h[\mathbf{x}^*])} \leq C \|\varphi\|_{W^{1,\infty}(\Gamma_h[\mathbf{x}^*])},$$

where the last inequality uses the  $W^{1,\infty}$  stability of the Lagrange interpolation. The  $H^1$  stability of  $P_h$  can be shown similarly by comparing  $P_h$  with the standard  $H^1$  projection

$R_h : H^1(\Gamma_h[\mathbf{x}^*]) \rightarrow S_h(\Gamma_h[\mathbf{x}^*])$ , i.e.,

$$\begin{aligned}
\|P_h\varphi - R_h\varphi\|_{H^1(\Gamma_h[\mathbf{x}^*])} &\leq Ch^{-1}\|P_h\varphi - R_h\varphi\|_{L^2(\Gamma_h[\mathbf{x}^*])} \\
&= Ch^{-1}\|P_h(\varphi - R_h\varphi)\|_{L^2(\Gamma_h[\mathbf{x}^*])} \\
&\leq Ch^{-1}\|\varphi - R_h\varphi\|_{L^2(\Gamma_h[\mathbf{x}^*])} \\
&\leq C\|\varphi\|_{H^1(\Gamma_h[\mathbf{x}^*])},
\end{aligned} \tag{B.2}$$

which implies that

$$\|P_h\varphi\|_{H^1(\Gamma_h[\mathbf{x}^*])} \leq \|P_h\varphi - R_h\varphi\|_{H^1(\Gamma_h[\mathbf{x}^*])} + \|R_h\varphi\|_{H^1(\Gamma_h[\mathbf{x}^*])} \leq C\|\varphi\|_{H^1(\Gamma_h[\mathbf{x}^*])},$$

Then, by the complex interpolation between the  $H^1$  and  $W^{1,\infty}$  stability results, we have

$$\|P_h\varphi\|_{W^{1,p}(\Gamma_h[\mathbf{x}^*])} \leq C\|\varphi\|_{W^{1,p}(\Gamma_h[\mathbf{x}^*])} \quad \text{for } 2 \leq p \leq \infty.$$

### Appendix C. Proof of Lemma 3.5.

The proof of Lemma 3.5 is similar as the proofs for Lemma 3.3 and Lemma 3.4. For the readers' convenience, we present a complete proof and divide it into several subsections.

**C.1. Basic setting.** We define a discrete velocity  $\tilde{v}_h^*$  on the smooth surface  $\Gamma$  in the same way as [30, Section 5.3] and [33, Section 7.3]. In particular, the surface  $\Gamma$  is decomposed into curved elements which move with velocity  $\tilde{v}_h^*$ , which is defined by

$$\tilde{v}_h^*(X_h^*(\cdot, t)^l, t) = \frac{d}{dt}(X_h^*(\cdot, t)^l).$$

Then the discrete material derivatives on the smooth surface  $\Gamma$  and its interpolated surface  $\Gamma_h[\mathbf{x}^*]$  are given by

$$\begin{aligned}
\partial_{t,h}^\bullet f &= \partial_t f + \tilde{v}_h^* \cdot \nabla f && \text{for } f \text{ defined on } \Gamma, \\
\partial_{t,h}^\bullet \varphi_h &= \partial_t \varphi_h + \hat{I}_h^* v \cdot \nabla \varphi_h && \text{for } \varphi_h \text{ defined on } \Gamma_h[\mathbf{x}^*],
\end{aligned}$$

where  $\hat{I}_h^* v \in S_h[\mathbf{x}^*]$  is the interpolated velocity (i.e., velocity of the interpolated surface  $\Gamma_h[\mathbf{x}^*]$ ). The two material derivatives above are related by

$$\partial_{t,h}^\bullet \varphi_h^l = (\partial_{t,h}^\bullet \varphi_h)^l \quad \text{for } \varphi_h \in S_h(\Gamma_h[\mathbf{x}^*]). \tag{C.1}$$

It is known that (cf. [30, Lemma 5.4])

$$\|\tilde{v}_h^* - v\|_{L^\infty(\Gamma)} + h\|\nabla_\Gamma(\tilde{v}_h^* - v)\|_{L^\infty(\Gamma)} \leq Ch^{k+1}. \tag{C.2}$$

Moreover, the surface  $\Gamma$  can be considered as moving with velocity  $\tilde{v}_h^*$ , and the following identities hold:

$$\frac{d}{dt} \int_\Gamma f = \int_\Gamma (\partial_{t,h}^\bullet f + f \nabla_\Gamma \cdot \tilde{v}_h^*), \tag{C.3}$$

$$\frac{d}{dt} \int_\Gamma \nabla_\Gamma f \cdot \nabla_\Gamma g = \int_\Gamma (\nabla_\Gamma \partial_{t,h}^\bullet f \cdot \nabla_\Gamma g + \nabla_\Gamma f \cdot \nabla_\Gamma \partial_{t,h}^\bullet g + \nabla_\Gamma f \cdot D_\Gamma \tilde{v}_h^* \nabla_\Gamma g), \tag{C.4}$$

where  $D_\Gamma \tilde{v}_h^* = (\nabla_\Gamma \cdot \tilde{v}_h^*)I_3 - (\nabla_\Gamma \tilde{v}_h^* + (\nabla_\Gamma \tilde{v}_h^*)^\top)$ ; see Lemma 3.2 and [25, Lemma 3.1]. Since the two functions  $u_h^*$  (defined on surface  $\Gamma_h[\mathbf{y}^*]$ ) and  $\hat{u}_h^*$  (defined on surface  $\Gamma_h[\mathbf{x}^*]$ ) have the same nodal vector, it follows that

$$(\partial_{t,h}^\bullet u_h^*)^\wedge = \partial_{t,h}^\bullet \hat{u}_h^* \quad (\text{both with nodal vector } \hat{\mathbf{u}}^*). \tag{C.5}$$

Since problem (3.6) is essentially an ODE system, we can assume that  $t_* \in (0, T]$  is the maximal time such that the solution of problem (3.6) satisfies the following estimate for  $t \in [0, t_*]$ :

$$\|\partial_{t,h}^\bullet \hat{u}_h^* - \hat{I}_h^* \partial_t^* u\|_{W^{1,6}(\Gamma_h[\mathbf{x}^*])} \leq 1. \tag{C.6}$$

Under this condition, we shall prove that (3.64)–(3.65) hold for  $t \in [0, t_*]$  (for some constants  $h_0$  and  $C$  which are independent of  $t_*$ ). Then, since problem (3.6) is essentially an ODE system, the local well-posedness of the ODE system and the continuity of  $\partial_{t,h}^\bullet \hat{u}_h^*$  in time imply that (3.64)–(3.65) hold for  $t \in [0, t_* + \delta_h]$  with some  $\delta_h > 0$  and with  $C$  replaced by  $2C$ . This implies that

$$\|\partial_{t,h}^\bullet \hat{u}_h^* - \hat{I}_h^* \partial_t^\bullet u\|_{W^{1,6}(\Gamma_h[\mathbf{x}^*])} \leq Ch^{-\frac{2}{3}} \|\partial_{t,h}^\bullet \hat{u}_h^* - \hat{I}_h^* \partial_t^\bullet u\|_{H^1(\Gamma_h[\mathbf{x}^*])} \leq Ch^{k-\frac{2}{3}} \quad \text{for } t \in [0, t_* + \delta].$$

The above inequality implies that (C.6) holds for  $t \in [0, t_* + \delta_h]$  (when  $k \geq 1$  and  $h$  is smaller than some constant which is independent of  $t_*$ ). This would imply that  $t_* = T$  (otherwise  $t_*$  is not the maximal time for (C.6) to hold) and therefore completes the proof of Lemma 3.5.

**C.2. Reduction to estimates of  $\partial_{t,h}^\bullet w$  and  $g$ .** For  $\varphi_h \in S_h[\mathbf{x}^*]^4$  satisfying  $\partial_{t,h}^\bullet \varphi_h = 0$  on  $\Gamma_h[\mathbf{x}^*]$  and accordingly  $\partial_{t,h}^\bullet \varphi_h^l = 0$  on  $\Gamma$  as a result of (C.1), we differentiate (3.23) in time and using (C.3)–(C.4). This yields that

$$\begin{aligned} & \mathbf{K}(\mathbf{x}^*) \dot{\mathbf{u}}^* \cdot \varphi \\ &= \int_{\Gamma} ((\partial_{t,h}^\bullet w + \partial_{t,h}^\bullet u) \cdot \varphi_h^l + \nabla_{\Gamma}(\partial_{t,h}^\bullet w + \partial_{t,h}^\bullet u) \cdot \nabla_{\Gamma} \varphi_h^l) \\ & \quad + \int_{\Gamma} ((w + u) \cdot \varphi_h^l \nabla_{\Gamma} \cdot \tilde{v}_h^* + \nabla_{\Gamma}(w + u) \cdot D_{\Gamma} \tilde{v}_h^* \nabla_{\Gamma} \varphi_h^l) - \left( \frac{d}{dt} \mathbf{K}(\mathbf{x}^*) \right) \mathbf{u}^* \cdot \varphi, \end{aligned} \quad (\text{C.7})$$

where

$$\left( \frac{d}{dt} \mathbf{K}(\mathbf{x}^*) \right) \mathbf{u}^* \cdot \varphi = \int_{\Gamma_h[\mathbf{x}^*]} (\hat{u}_h^* \cdot \varphi_h \nabla_{\Gamma_h[\mathbf{x}^*]} \cdot \hat{I}_h^* v + \nabla_{\Gamma_h[\mathbf{x}^*]} \hat{u}_h^* \cdot D_{\Gamma_h[\mathbf{x}^*]} \hat{I}_h^* v \nabla_{\Gamma_h[\mathbf{x}^*]} \varphi_h),$$

which follows from the identities in Lemma 3.2. Let  $g \in H^1(\Gamma)^4$  be the solution of the following weak formulation:

$$\begin{aligned} \int_{\Gamma} (g \cdot \varphi^l + \nabla_{\Gamma} g \cdot \nabla_{\Gamma} \varphi^l) &= \int_{\Gamma} (w \cdot \varphi^l \nabla_{\Gamma} \cdot \tilde{v}_h^* + \nabla_{\Gamma} w \cdot D_{\Gamma} \tilde{v}_h^* \nabla_{\Gamma} \varphi^l) \\ & \quad + \int_{\Gamma} u \cdot \varphi^l \nabla_{\Gamma} \cdot \tilde{v}_h^* - \int_{\Gamma_h[\mathbf{x}^*]} \hat{u}_h^* \cdot \varphi \nabla_{\Gamma_h[\mathbf{x}^*]} \cdot \hat{I}_h^* v \\ & \quad + \int_{\Gamma} \nabla_{\Gamma} u \cdot D_{\Gamma} \tilde{v}_h^* \nabla_{\Gamma} \varphi^l - \int_{\Gamma_h[\mathbf{x}^*]} \nabla_{\Gamma_h[\mathbf{x}^*]} \hat{u}_h^* \cdot D_{\Gamma_h[\mathbf{x}^*]} \hat{I}_h^* v \nabla_{\Gamma_h[\mathbf{x}^*]} \varphi \\ & =: L_1(\varphi^l) + L_2(\varphi^l) + L_3(\varphi^l) \quad \forall \varphi^l \in H^1(\Gamma)^4, \end{aligned} \quad (\text{C.8})$$

where  $\varphi$  is the inverse lift of  $\varphi^l$  from  $\Gamma$  onto  $\Gamma_h[\mathbf{x}^*]$ . Then, substituting (C.8) into (C.7), we have

$$\mathbf{K}(\mathbf{x}^*) \dot{\mathbf{u}}^* \cdot \varphi = \int_{\Gamma} ((\partial_{t,h}^\bullet w + \partial_{t,h}^\bullet u + g) \varphi_h^l + \nabla_{\Gamma}(\partial_{t,h}^\bullet w + \partial_{t,h}^\bullet u + g) \cdot \nabla_{\Gamma} \varphi_h^l) \quad \forall \varphi_h \in S_h(\Gamma_h[\mathbf{x}^*])^4,$$

which implies that  $\partial_{t,h}^\bullet \hat{u}_h^*$  (with nodal vector  $\dot{\mathbf{u}}^*$ ) is the linear Ritz projection of  $\partial_{t,h}^\bullet w + \partial_{t,h}^\bullet u + g$  onto  $S_h(\Gamma_h[\mathbf{x}^*])^4$ . As a result, we have (see [20, Proof of Corollary 4.2], inserting (4.4) into (3.3) and (3.5) therein)

$$\begin{aligned} \|(\partial_{t,h}^\bullet \hat{u}_h^*)^l - \partial_{t,h}^\bullet w - \partial_{t,h}^\bullet u - g\|_{H^1(\Gamma)} &\leq Ch^{k+1} \|\partial_{t,h}^\bullet w + \partial_{t,h}^\bullet u + g\|_{H^1(\Gamma)} \\ & \quad + C \inf_{\chi_h \in S_h(\Gamma_h[\mathbf{x}^*])} \|\partial_{t,h}^\bullet w + \partial_{t,h}^\bullet u + g - \chi_h^l\|_{H^1(\Gamma)}. \end{aligned} \quad (\text{C.9})$$

and

$$\begin{aligned} \|(\partial_{t,h}^\bullet \hat{u}_h^*)^l - \partial_{t,h}^\bullet w - \partial_{t,h}^\bullet u - g\|_{L^2(\Gamma)} &\leq Ch^{k+1} \|\partial_{t,h}^\bullet w + \partial_{t,h}^\bullet u + g\|_{H^1(\Gamma)} \\ & \quad + Ch \inf_{\chi_h \in S_h(\Gamma_h[\mathbf{x}^*])} \|\partial_{t,h}^\bullet w + \partial_{t,h}^\bullet u + g - \chi_h^l\|_{H^1(\Gamma)}. \end{aligned} \quad (\text{C.10})$$

Hence, by using the triangle inequality, we have

$$\begin{aligned} \|(\partial_{t,h}^{\bullet} \hat{u}_h^*)^l - \partial_{t,h}^{\bullet} u\|_{H^1(\Gamma)} &\leq C(\|\partial_{t,h}^{\bullet} w\|_{H^1(\Gamma)} + \|g\|_{H^1(\Gamma)}) + Ch^{k+1} \|\partial_{t,h}^{\bullet} u\|_{H^1(\Gamma)} \\ &\quad + C \inf_{\chi_h \in S_h(\Gamma_h[\mathbf{x}^*])} \|\partial_{t,h}^{\bullet} u - \chi_h^l\|_{H^1(\Gamma)}. \end{aligned} \quad (\text{C.11})$$

and

$$\begin{aligned} \|(\partial_{t,h}^{\bullet} \hat{u}_h^*)^l - \partial_{t,h}^{\bullet} u\|_{L^2(\Gamma)} &\leq \|\partial_{t,h}^{\bullet} w\|_{L^2(\Gamma)} + \|g\|_{L^2(\Gamma)} + Ch(\|\partial_{t,h}^{\bullet} w\|_{H^1(\Gamma)} + \|g\|_{H^1(\Gamma)}) \\ &\quad + Ch^{k+1} \|\partial_{t,h}^{\bullet} u\|_{H^1(\Gamma)} + Ch \inf_{\chi_h \in S_h(\Gamma_h[\mathbf{x}^*])} \|\partial_{t,h}^{\bullet} u - \chi_h^l\|_{H^1(\Gamma)}. \end{aligned} \quad (\text{C.12})$$

It remains to estimate the right-hand sides of (C.11) and (C.12).

**C.3. Estimate for  $\|g\|_{H^1(\Gamma)}$ .** We estimate  $\|g\|_{H^1(\Gamma)}$  from the expression in (C.8), where  $L_1(\varphi^l)$  can be estimated by using boundedness of  $\|\nabla_{\Gamma} \tilde{v}_h^*\|_{L^\infty(\Gamma)}$  as shown in (C.2) and the estimate of  $\|w\|_{H^1(\Gamma)}$  in (3.63), i.e.,

$$|L_1(\varphi^l)| \leq C\|w\|_{H^1(\Gamma)} \|\nabla_{\Gamma} \tilde{v}_h^*\|_{L^\infty(\Gamma)} \|\varphi^l\|_{H^1(\Gamma)} \leq Ch^k \|\varphi^l\|_{H^1(\Gamma)}.$$

Given the following established estimate(which follows from [20, Proposition 2.7 and relations (2.1)-(2.12)]):

$$\begin{aligned} &\|(\nabla_{\Gamma} \cdot \tilde{v}_h^*)^{-l} - \nabla_{\Gamma_h[\mathbf{x}^*]} \cdot \hat{I}_h^* v\|_{L^p(\Gamma_h[\mathbf{x}^*])} \\ &\sim \|\nabla_{\Gamma} \cdot \tilde{v}_h^* - (\nabla_{\Gamma_h[\mathbf{x}^*]} \cdot \hat{I}_h^* v)^l\|_{L^p(\Gamma)} \\ &\leq \|\nabla_{\Gamma} \cdot (\tilde{v}_h^* - v)\|_{L^p(\Gamma)} + \|\nabla_{\Gamma} \cdot v - (\nabla_{\Gamma_h[\mathbf{x}^*]} \cdot \hat{I}_h^* v)^l\|_{L^p(\Gamma)} \\ &\leq Ch^k, \end{aligned} \quad (\text{C.13})$$

we apply Lemma 3.4 and (3.37), along with considering the estimates for surface perturbation via the lift map. Consequently, we have

$$\begin{aligned} |L_2(\varphi^l)| &\leq \left| \int_{\Gamma} u \cdot \varphi^l \nabla_{\Gamma} \cdot \tilde{v}_h^* - \int_{\Gamma_h[\mathbf{x}^*]} u^{-l} \cdot \varphi(\nabla_{\Gamma} \cdot \tilde{v}_h^*)^{-l} \right| \\ &\quad + \left| \int_{\Gamma_h[\mathbf{x}^*]} u^{-l} \cdot \varphi(\nabla_{\Gamma} \cdot \tilde{v}_h^*)^{-l} - \int_{\Gamma_h[\mathbf{x}^*]} u^{-l} \cdot \varphi(\nabla_{\Gamma_h[\mathbf{x}^*]} \cdot \hat{I}_h^* v) \right| \\ &\quad + \left| \int_{\Gamma_h[\mathbf{x}^*]} u^{-l} \cdot \varphi(\nabla_{\Gamma_h[\mathbf{x}^*]} \cdot \hat{I}_h^* v) - \int_{\Gamma_h[\mathbf{x}^*]} \hat{u}_h^* \cdot \varphi(\nabla_{\Gamma_h[\mathbf{x}^*]} \cdot \hat{I}_h^* v) \right| \\ &\leq Ch^k \|\varphi^l\|_{L^2(\Gamma)}. \end{aligned}$$

Similarly, by employing the result of Lemma 3.3 and following a process analogous to the one described above, we obtain

$$|L_3(\varphi^l)| \leq Ch^k \|\varphi^l\|_{H^1(\Gamma)}$$

Therefore, by substituting  $\varphi^l = g$  into (C.8) and utilizing the above-established estimates of  $L_1(\varphi^l)$ ,  $L_2(\varphi^l)$  and  $L_3(\varphi^l)$ , we obtain

$$\|g\|_{H^1(\Gamma)} \leq Ch^k. \quad (\text{C.14})$$

**C.4. Estimate for  $\|g\|_{L^2(\Gamma)}$ .** Next, we estimate  $\|g\|_{L^2(\Gamma)}$  via a duality argument. Specifically, we define  $\varphi^l \in H^2(\Gamma)^4$  to be the solution of

$$\varphi^l - \Delta_{\Gamma} \varphi^l = g \text{ on } \Gamma, \quad (\text{C.15})$$

which satisfies the following standard  $H^2$  regularity estimate:

$$\|\varphi^l\|_{H^2(\Gamma)} \leq C\|g\|_{L^2(\Gamma)}. \quad (\text{C.16})$$

By testing (C.15) with  $g$  and estimating the right-hand side of (C.8), we shall prove the following result:

$$\|g\|_{L^2(\Gamma)}^2 = \int_{\Gamma} (g\varphi^l + \nabla_{\Gamma} g \cdot \nabla_{\Gamma} \varphi^l) \leq Ch^{k+1} \|\varphi^l\|_{H^2(\Gamma)}. \quad (\text{C.17})$$

This, together with (C.16), would imply the following estimate:

$$\|g\|_{L^2(\Gamma)} \leq Ch^{k+1}. \quad (\text{C.18})$$

It remains to prove the inequality in (C.17). This is achieved by estimating the right-hand side of (C.8) as follows.

Firstly,  $|L_1(\varphi^l)|$  can be estimated by using integration by parts (after approximating  $\tilde{v}_h^*$  by  $v$  in the integrand), i.e.,

$$\begin{aligned} |L_1(\varphi^l)| &= \int_{\Gamma} (w \cdot \varphi^l \nabla_{\Gamma} \cdot \tilde{v}_h^* + \nabla_{\Gamma} w \cdot D_{\Gamma} \tilde{v}_h^* \nabla_{\Gamma} \varphi^l) \\ &= \int_{\Gamma} (w \cdot \varphi^l \nabla_{\Gamma} \cdot \tilde{v}_h^* + \nabla_{\Gamma} w \cdot D_{\Gamma} v \nabla_{\Gamma} \varphi^l) + \int_{\Gamma} \nabla_{\Gamma} w \cdot D_{\Gamma} (\tilde{v}_h^* - v) \nabla_{\Gamma} \varphi^l \\ &= \int_{\Gamma} [w \cdot \varphi^l \nabla_{\Gamma} \cdot \tilde{v}_h^* - w \cdot (\nabla_{\Gamma} \cdot (D_{\Gamma} v \nabla_{\Gamma} \varphi^l))] + \int_{\Gamma} \nabla_{\Gamma} w \cdot D_{\Gamma} (\tilde{v}_h^* - v) \nabla_{\Gamma} \varphi^l \\ &\leq C \|w\|_{L^2(\Gamma)} \|\nabla_{\Gamma} \cdot \tilde{v}_h^*\|_{L^{\infty}(\Gamma)} \|\varphi^l\|_{H^2(\Gamma)} + C \|w\|_{L^2(\Gamma)} \|\varphi^l\|_{H^2(\Gamma)} \\ &\quad + C \|w\|_{H^1(\Gamma)} \|D_{\Gamma} (\tilde{v}_h^* - v)\|_{L^{\infty}} \|\varphi^l\|_{H^1(\Gamma)} \\ &\leq C \|w\|_{L^2(\Gamma)} \|\varphi^l\|_{H^2(\Gamma)} + Ch \|w\|_{H^1(\Gamma)} \|\varphi^l\|_{H^2(\Gamma)} \quad (\text{here (C.2) is used}) \\ &\leq Ch^{k+1} \|\varphi^l\|_{H^2(\Gamma)}, \end{aligned}$$

where we have used the estimate of  $\|w\|_{L^2(\Gamma)}$  and  $\|w\|_{H^1(\Gamma)}$  in (3.63).

Secondly,  $|L_2(\varphi^l)|$  can be estimated by defining and utilizing intermediate surfaces between  $\Gamma_h[\mathbf{x}^*]$  to  $\Gamma$ . Namely, we define

$$x^{\theta} := (1 - \theta)x^l + \theta x \text{ for } x \in \Gamma_h[\mathbf{x}^*] \quad \text{and} \quad \Gamma^{\theta} := (1 - \theta)\Gamma + \theta\Gamma_h[\mathbf{x}^*] = \{x^{\theta} : x \in \Gamma_h[\mathbf{x}^*]\}.$$

Let  $u^{\theta}$ ,  $\varphi^{\theta}$ ,  $\tilde{v}_h^{\theta}$  and  $e^{\theta}$  be functions on  $\Gamma^{\theta}$  defined by

$$\begin{aligned} u^{\theta}(x^{\theta}) &= (1 - \theta)u(x^l) + \theta \hat{u}_h^*(x), \\ \varphi^{\theta}(x^{\theta}) &= \varphi(x) = \varphi^l(x^l), \\ \tilde{v}_h^{\theta}(x^{\theta}) &= (1 - \theta)\tilde{v}_h^*(x^l) + \theta \hat{I}_h^* v(x), \\ e^{\theta}(x^{\theta}) &= x - x^l \quad \text{for } x \in \Gamma_h[\mathbf{x}^*]. \end{aligned}$$

Then the surface  $\Gamma^{\theta}$  moves with velocity  $e^{\theta}$  (as  $\theta$  changes) and

$$\partial_{\theta}^{\bullet} u^{\theta}(x^{\theta}) = \hat{u}_h^*(x) - u(x^l), \quad \partial_{\theta}^{\bullet} \varphi^{\theta} = 0 \quad \text{and} \quad \partial_{\theta}^{\bullet} \tilde{v}_h^{\theta}(x^{\theta}) = \hat{I}_h^* v(x) - \tilde{v}_h^*(x^l).$$

From (C.2) we conclude that

$$|\partial_{\theta}^{\bullet} \tilde{v}_h^{\theta}(x^{\theta})| = |\hat{I}_h^* v(x) - \tilde{v}_h^*(x^l)| \leq |\hat{I}_h^* v(x) - v(x^l)| + |v(x^l) - \tilde{v}_h^*(x^l)| \leq Ch^{k+1}. \quad (\text{C.19})$$

Since

$$\|\nabla_{\Gamma_h[\mathbf{x}^*]} x^{\theta} - \nabla_{\Gamma_h[\mathbf{x}^*]} \text{id}\|_{L^p(\Gamma_h[\mathbf{x}^*])} = (1 - \theta) \|\nabla_{\Gamma_h[\mathbf{x}^*]} e^l\|_{L^p(\Gamma_h[\mathbf{x}^*])} \leq Ch^k \text{ for } 2 \leq p \leq \infty,$$

it follows that the  $L^p(\Gamma^{\theta})$  norm of  $f$  is equivalent to the  $L^p(\Gamma_h[\mathbf{x}^*])$  norm of  $f \circ X^{\theta}$ . The result of Lemma 3.4 and the norm equivalence between  $\Gamma^{\theta}$  and  $\Gamma_h[\mathbf{x}^*]$  imply that

$$\|\partial_{\theta}^{\bullet} u^{\theta}(x^{\theta})\|_{L^2(\Gamma^{\theta})} \leq C \|\hat{u}_h^* - u^{-l}\|_{L^2(\Gamma_h[\mathbf{x}^*])} \leq Ch^{k+1}. \quad (\text{C.20})$$

Then  $|L_2(\varphi^l)|$  can be rewritten as follows:

$$\begin{aligned}
|L_2(\varphi^l)| &= \left| \int_{\Gamma} u \cdot \varphi^l \nabla_{\Gamma} \cdot \tilde{v}_h^* - \int_{\Gamma_h[\mathbf{x}^*]} \hat{u}_h^* \cdot \varphi \nabla_{\Gamma_h[\mathbf{x}^*]} \cdot \hat{I}_h^* v \right| \\
&= \left| \int_0^1 \frac{d}{d\theta} \int_{\Gamma^\theta} u^\theta \cdot \varphi^\theta \nabla_{\Gamma^\theta} \cdot \tilde{v}_h^\theta d\theta \right| \\
&= \left| \int_0^1 \int_{\Gamma^\theta} [u^\theta \cdot \varphi^\theta (\nabla_{\Gamma^\theta} \cdot \tilde{v}_h^\theta) (\nabla_{\Gamma^\theta} \cdot e^\theta) + \partial_\theta^\bullet u^\theta \cdot \varphi^\theta (\nabla_{\Gamma^\theta} \cdot \tilde{v}_h^\theta)] d\theta \right. \\
&\quad \left. + \int_0^1 \int_{\Gamma^\theta} u^\theta \cdot \varphi^\theta (\nabla_{\Gamma^\theta} \cdot \partial_\theta^\bullet \tilde{v}_h^\theta - \text{tr}(\nabla_{\Gamma^\theta} e^\theta \nabla_{\Gamma^\theta} \tilde{v}_h^\theta) + \text{tr}[n^\theta (n^\theta)^\top (\nabla_{\Gamma^\theta} e^\theta)^\top \nabla_{\Gamma^\theta} \tilde{v}_h^\theta]) d\theta \right| \\
&=: |L_{21}(\varphi^l) + L_{22}(\varphi^l) + L_{23}(\varphi^l) + L_{24}(\varphi^l) + L_{25}(\varphi^l)|.
\end{aligned}$$

The term  $L_{22}(\varphi^l)$  can be estimated by utilizing (C.20) and  $\|\nabla_{\Gamma^\theta} \cdot \tilde{v}_h^\theta\|_{L^\infty(\Gamma^\theta)} \leq C$ . The latter result follows from (C.2) and  $\|\nabla_{\Gamma_h[\mathbf{x}^*]} \cdot \hat{I}_h v\|_{L^\infty(\Gamma_h[\mathbf{x}^*])} \leq C$ , and the equivalence of norms on  $\Gamma^\theta$ ,  $\theta \in [0, 1]$ . These results, and the equivalence between  $\|\varphi^\theta\|_{L^2(\Gamma^\theta)}$  and  $\|\varphi^l\|_{L^2(\Gamma)}$ , lead to the following estimate of  $L_{22}(\varphi^l)$ :

$$|L_{22}(\varphi^l)| \leq Ch^{k+1} \|\varphi^l\|_{L^2(\Gamma)}.$$

The term  $L_{21}(\varphi^l)$  can be rewritten as follows:

$$\begin{aligned}
|L_{21}(\varphi^l)| &= \left| \int_0^1 \left[ \int_{\Gamma^\theta} u^\theta \cdot \varphi^\theta (\nabla_{\Gamma^\theta} \cdot \tilde{v}_h^\theta) (\nabla_{\Gamma^\theta} \cdot e^\theta) - \int_{\Gamma} u \cdot \varphi^l (\nabla_{\Gamma} \cdot \tilde{v}_h^*) (\nabla_{\Gamma} \cdot e^1) \right] d\theta \right. \\
&\quad \left. + \int_{\Gamma} u \cdot \varphi^l (\nabla_{\Gamma} \cdot \tilde{v}_h^* - \nabla_{\Gamma} \cdot v) (\nabla_{\Gamma} e^1) + \int_{\Gamma} u \cdot \varphi^l (\nabla_{\Gamma} \cdot v) (\nabla_{\Gamma} e^1) \right| \\
&=: |L_{211}(\varphi^l) + L_{212}(\varphi^l) + L_{213}(\varphi^l)|.
\end{aligned}$$

Since

$$\begin{aligned}
L_{211}(\varphi^l) &= \int_0^\theta \frac{d}{d\alpha} \int_{\Gamma^\alpha} u^\alpha \varphi^\alpha (\nabla_{\Gamma^\alpha} \cdot \tilde{v}_h^\alpha) (\nabla_{\Gamma^\alpha} \cdot e^\alpha) d\alpha \\
&= \int_0^\theta \int_{\Gamma^\alpha} [u^\alpha \varphi^\alpha (\nabla_{\Gamma^\alpha} \cdot \tilde{v}_h^\alpha) (\nabla_{\Gamma^\alpha} \cdot e^\alpha) (\nabla_{\Gamma^\alpha} \cdot e^\alpha) + \partial_\alpha^\bullet u^\alpha \varphi^\alpha (\nabla_{\Gamma^\alpha} \cdot \tilde{v}_h^\alpha) (\nabla_{\Gamma^\alpha} \cdot e^\alpha) \\
&\quad + u^\alpha \varphi^\alpha \partial_\alpha^\bullet (\nabla_{\Gamma^\alpha} \cdot \tilde{v}_h^\alpha) (\nabla_{\Gamma^\alpha} \cdot e^\alpha) + u^\alpha \varphi^\alpha (\nabla_{\Gamma^\alpha} \cdot \tilde{v}_h^\alpha) \partial_\alpha^\bullet (\nabla_{\Gamma^\alpha} \cdot e^\alpha)] d\alpha,
\end{aligned}$$

where  $\partial_\alpha^\bullet (\nabla_{\Gamma^\alpha} \cdot \tilde{v}_h^\alpha)$  and  $\partial_\alpha^\bullet (\nabla_{\Gamma^\alpha} \cdot e^\alpha)$  can be further expressed as follows using relation (3.18):

$$\begin{aligned}
\partial_\alpha^\bullet (\nabla_{\Gamma^\alpha} \cdot \tilde{v}_h^\alpha) &= (\nabla_{\Gamma^\alpha} \cdot \partial_\alpha^\bullet \tilde{v}_h^\alpha) - (\nabla_{\Gamma^\alpha} e^\alpha - n_{\Gamma^\alpha} n_{\Gamma^\alpha}^\top (\nabla_{\Gamma^\alpha} e^\alpha)^\top) \nabla_{\Gamma^\alpha} \tilde{v}_h^\alpha \\
\partial_\alpha^\bullet (\nabla_{\Gamma^\alpha} \cdot e^\alpha) &= (\nabla_{\Gamma^\alpha} \cdot \partial_\alpha^\bullet e^\alpha) - (\nabla_{\Gamma^\alpha} e^\alpha - n_{\Gamma^\alpha} n_{\Gamma^\alpha}^\top (\nabla_{\Gamma^\alpha} e^\alpha)^\top) \nabla_{\Gamma^\alpha} e^\alpha.
\end{aligned}$$

By using the norm equivalence between  $\Gamma$  and  $\Gamma^\alpha$ , and the interpolation error estimates

$$\begin{aligned}
\|e^\alpha\|_{L^\infty(\Gamma^\alpha)} &\sim \|\text{id}_{\Gamma_h[\mathbf{x}^*]}^l - \text{id}_{\Gamma_h[\mathbf{x}^*]}\|_{L^\infty(\Gamma_h[\mathbf{x}^*])} \leq Ch^{k+1} \\
\|\nabla_{\Gamma^\alpha} e^\alpha\|_{L^\infty(\Gamma^\alpha)} &\sim \|\nabla_{\Gamma_h[\mathbf{x}^*]} (\text{id}_{\Gamma_h[\mathbf{x}^*]}^l - \text{id}_{\Gamma_h[\mathbf{x}^*]})\|_{L^\infty(\Gamma_h[\mathbf{x}^*])} \leq Ch^k,
\end{aligned} \tag{C.21}$$

as well as the estimates in (C.19)–(C.20), we can derive the following estimate of  $L_{211}(\varphi^l)$ :

$$|L_{211}(\varphi^l)| \leq Ch^{2k} \|\varphi^l\|_{L^2(\Gamma)} \leq Ch^{k+1} \|\varphi^l\|_{L^2(\Gamma)}.$$

The term  $L_{212}(\varphi^l)$  can be estimated by employing inequalities (C.2) and (C.21), i.e.,

$$|L_{212}(\varphi^l)| \leq Ch^{2k} \|\varphi^l\|_{L^2(\Gamma)} \leq Ch^{k+1} \|\varphi^l\|_{L^2(\Gamma)}.$$

The term  $L_{213}(\varphi^l)$  can be estimated with integration by parts, which allows us to transfer the gradient from  $e^1$  to  $H\varphi^l(\nabla_{\Gamma} \cdot v)$ . This, along with (C.21), leads to the following result:

$$|L_{213}(\varphi^l)| \leq Ch^{k+1} \|\varphi^l\|_{H^1(\Gamma)}.$$

By employing a similar method as that for estimating  $L_{21}(\varphi^l)$ , we can derive the following result:

$$|L_{23}(\varphi^l) + L_{24}(\varphi^l) + L_{25}(\varphi^l)| \leq Ch^{k+1} \|\varphi^l\|_{H^1(\Gamma)}.$$

Combining the estimates for  $|L_{2i}(\varphi^l)|$  for  $i = 1, \dots, 5$ , we have the following result:

$$|L_2(\varphi^l)| \leq Ch^{k+1} \|\varphi^l\|_{H^1(\Gamma)}.$$

Thirdly, we can estimate  $|L_3(\varphi^l)|$  similarly as  $|L_2(\varphi^l)|$ , by utilizing the intermediate surface  $\Gamma^\theta$  between  $\Gamma_h[\mathbf{x}^*]$  and  $\Gamma$ . This leads to the following result:

$$|L_3(\varphi^l)| \leq Ch^{k+1} \|\varphi^l\|_{H^2(\Gamma)}.$$

Substituting these estimates for  $L_i(\varphi^l)$  for  $i = 1, 2, 3$  into (C.8), we obtain

$$\left| \int_{\Gamma} (g\varphi^l + \nabla_{\Gamma} g \cdot \nabla_{\Gamma} \varphi^l) \right| \leq Ch^{k+1} \|\varphi^l\|_{H^2(\Gamma)}. \quad (\text{C.22})$$

This proves (C.17) and therefore completes the proof of (C.18).

**C.5. Estimate for  $\|\partial_{t,h}^{\bullet} w\|_{H^1(\Gamma)}$ .** To estimate  $\|\partial_{t,h}^{\bullet} w\|_{H^1(\Gamma)}$ , we need to consider the time derivative of (3.22) by extending  $\varphi^l \in H^1(\Gamma)^4$  in time in such a way that

$$\partial_{t,h}^{\bullet} \varphi^l = (\partial_{t,h}^{\bullet} \varphi^l)^l = 0, \quad (\text{C.23})$$

where  $\partial_{t,h}^{\bullet} \varphi^l$  denotes the material derivative with respect to velocity  $\tilde{v}_h^*$ . 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 the intermediate surface  $\Gamma_h[\mathbf{y}^\theta]$  with nodal vectors  $\mathbf{u}^*$ ,  $\mathbf{P}_h \varphi$  and  $\mathbf{e}_y := \mathbf{y}^* - \mathbf{x}^*$ , respectively. Then the surface  $\Gamma_h[\mathbf{y}^\theta]$  moves with velocity  $\hat{e}_y^\theta$  as  $\theta \in [0, 1]$  changes, and

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

When  $\theta$  is fixed and  $t$  increases, the surface  $\Gamma_h[\mathbf{y}^\theta]$  moves with velocity  $v_h^\theta$ , which is a finite element function on  $\Gamma_h[\mathbf{y}^\theta]$  with nodal vector

$$\mathbf{v}^\theta = (1-\theta)\mathbf{I}_h v + \theta\mathbf{v}^*.$$

Let  $e_v^\theta$  be the finite element function on  $\Gamma_h[\mathbf{y}^\theta]$  with nodal vector  $\mathbf{v}^* - \mathbf{I}_h v$ , which is the same as the nodal vector of  $-\hat{I}_h^* (\hat{H}_h^* \hat{n}_h^* - \hat{I}_h^* H \hat{I}_h^* n)$ . Hence, by using the  $L^p$  and  $W^{1,p}$  stability of  $\hat{I}_h^*$  (see Appendix A), and the  $W^{1,p}$  and  $L^2$  estimates of the dynamic Ritz projection in Lemma 3.3 and Lemma 3.4, we can derive the following estimates:

$$\|e_v^\theta\|_{L^2(\Gamma_h[\mathbf{y}^\theta])} \leq Ch^{k+1} \quad \text{and} \quad \|e_v^\theta\|_{W^{1,p}(\Gamma_h[\mathbf{y}^\theta])} \leq Ch^k \quad \text{for } 2 \leq p < \infty. \quad (\text{C.24})$$

For a function  $f$  defined on  $\Gamma_h[\mathbf{y}^\theta]$  (such as  $f = \hat{u}_h^\theta$  and  $f = \hat{\varphi}_h^\theta$ ), we denote by  $\partial_{t,h}^{\bullet} f$  the material derivative with respect to the velocity  $v_h^\theta$ . By differentiating (3.22) in time we obtain

$$\begin{aligned} & \int_{\Gamma} (\partial_{t,h}^{\bullet} w \cdot \varphi^l + \nabla_{\Gamma} \partial_{t,h}^{\bullet} w \cdot \nabla_{\Gamma} \varphi^l) \\ &= \frac{d}{dt} [(\mathbf{K}(\mathbf{x}^*) - \mathbf{K}(\mathbf{y}^*))\mathbf{u}^* \cdot (\mathbf{P}_h \varphi)] - \int_{\Gamma} (w \cdot \varphi^l \nabla_{\Gamma} \cdot \tilde{v}_h^* + \nabla_{\Gamma} w \cdot D_{\Gamma} \tilde{v}_h^* \nabla_{\Gamma} \varphi^l) \\ &= -\frac{d}{dt} \int_0^1 \frac{d}{d\theta} \int_{\Gamma_h[\mathbf{y}^\theta]} (\hat{u}_h^\theta \cdot \hat{\varphi}_h^\theta + \nabla_{\Gamma_h[\mathbf{y}^\theta]} \hat{u}_h^\theta \cdot \nabla_{\Gamma_h[\mathbf{y}^\theta]} \hat{\varphi}_h^\theta) d\theta - \int_{\Gamma} (w \cdot \varphi^l \nabla_{\Gamma} \cdot \tilde{v}_h^* + \nabla_{\Gamma} w \cdot D_{\Gamma} \tilde{v}_h^* \nabla_{\Gamma} \varphi^l) \\ &= -\frac{d}{dt} \int_0^1 \int_{\Gamma_h[\mathbf{y}^\theta]} (\hat{u}_h^\theta \cdot \hat{\varphi}_h^\theta \nabla_{\Gamma_h[\mathbf{y}^\theta]} \cdot \hat{e}_y^\theta + \nabla_{\Gamma_h[\mathbf{y}^\theta]} \hat{u}_h^\theta \cdot D_{\Gamma_h[\mathbf{y}^\theta]} \hat{e}_y^\theta \nabla_{\Gamma_h[\mathbf{y}^\theta]} \hat{\varphi}_h^\theta) d\theta \\ & \quad - \int_{\Gamma} (w \cdot \varphi^l \nabla_{\Gamma} \cdot \tilde{v}_h^* + \nabla_{\Gamma} w \cdot D_{\Gamma} \tilde{v}_h^* \nabla_{\Gamma} \varphi^l) \end{aligned}$$

$$\begin{aligned}
&= - \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]} \cdot v_h^\theta d\theta \\
&\quad - \int_0^1 \int_{\Gamma_h[\mathbf{y}^\theta]} \partial_{t,h}^\bullet \hat{u}_h^\theta \cdot \hat{\varphi}_h^\theta \nabla_{\Gamma_h[\mathbf{y}^\theta]} \cdot \hat{e}_y^\theta d\theta \\
&\quad - \int_0^1 \int_{\Gamma_h[\mathbf{y}^\theta]} \hat{u}_h^\theta \cdot \partial_{t,h}^\bullet \hat{\varphi}_h^\theta \nabla_{\Gamma_h[\mathbf{y}^\theta]} \cdot \hat{e}_y^\theta d\theta \\
&\quad - \int_0^1 \int_{\Gamma_h[\mathbf{y}^\theta]} \hat{u}_h^\theta \cdot \hat{\varphi}_h^\theta \left( \nabla_{\Gamma_h[\mathbf{y}^\theta]} \cdot \partial_{t,h}^\bullet \hat{e}_y^\theta - \text{tr}(\nabla_{\Gamma_h[\mathbf{y}^\theta]} v_h^\theta \nabla_{\Gamma_h[\mathbf{y}^\theta]} \hat{e}_y^\theta) \right. \\
&\quad \quad \left. + \text{tr}[n_{\Gamma_h[\mathbf{y}^\theta]} n_{\Gamma_h[\mathbf{y}^\theta]}^\top (\nabla_{\Gamma_h[\mathbf{y}^\theta]} v_h^\theta)^\top \nabla_{\Gamma_h[\mathbf{y}^\theta]} \hat{e}_y^\theta] \right) d\theta \\
&\quad - \int_0^1 \int_{\Gamma_h[\mathbf{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 \nabla_{\Gamma_h[\mathbf{y}^\theta]} \cdot v_h^\theta d\theta \\
&\quad - \int_0^1 \int_{\Gamma_h[\mathbf{y}^\theta]} \partial_{t,h}^\bullet \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 d\theta \\
&\quad - \int_0^1 \int_{\Gamma_h[\mathbf{y}^\theta]} \nabla_{\Gamma_h[\mathbf{y}^\theta]} \hat{u}_h^\theta \cdot \partial_{t,h}^\bullet D_{\Gamma_h[\mathbf{y}^\theta]} \hat{e}_y^\theta \nabla_{\Gamma_h[\mathbf{y}^\theta]} \hat{\varphi}_h^\theta d\theta \\
&\quad - \int_0^1 \int_{\Gamma_h[\mathbf{y}^\theta]} \nabla_{\Gamma_h[\mathbf{y}^\theta]} \hat{u}_h^\theta \cdot D_{\Gamma_h[\mathbf{y}^\theta]} \hat{e}_y^\theta \partial_{t,h}^\bullet \nabla_{\Gamma_h[\mathbf{y}^\theta]} \hat{\varphi}_h^\theta d\theta \\
&\quad - \int_{\Gamma} (w \cdot \varphi^l \nabla_{\Gamma} \cdot \tilde{v}_h^* + \nabla_{\Gamma} w \cdot D_{\Gamma} \tilde{v}_h^* \nabla_{\Gamma} \varphi^l) \\
&=: \sum_{j=1}^9 W_j. \tag{C.25}
\end{aligned}$$

Since  $\partial_{t,h}^\bullet \hat{u}_h^\theta$  has nodal vector  $\hat{\mathbf{u}}^*$ , it follows from (C.6) that  $\|\partial_{t,h}^\bullet \hat{u}_h^\theta\|_{W^{1,6}(\Gamma_h[\mathbf{y}^\theta])} \leq C$ . By using the estimate of  $\|\nabla_{\Gamma_h[\mathbf{y}^\theta]} \hat{e}_y^\theta\|_{L^2(\Gamma_h[\mathbf{y}^\theta])} \sim \|\nabla_{\Gamma_h[\mathbf{x}^*]}(\text{id}_{\Gamma_h[\mathbf{x}^*]} - \hat{y}_h^*)\|_{L^2(\Gamma_h[\mathbf{x}^*])}$  in Lemma 3.3, as well as inequalities (3.63), (C.2) and (C.6), we can derive the following estimate:

$$|W_1| + |W_2| + |W_5| + |W_9| \leq Ch^k \|\varphi^l\|_{H^1(\Gamma)},$$

The term  $|W_6|$  can be estimated by using relation (3.18), which allows us to write  $\partial_{t,h}^\bullet \nabla_{\Gamma_h[\mathbf{y}^\theta]} \hat{u}_h^\theta$  as

$$\partial_{t,h}^\bullet \nabla_{\Gamma_h[\mathbf{y}^\theta]} \hat{u}_h^\theta = \nabla_{\Gamma_h[\mathbf{y}^\theta]} \partial_{t,h}^\bullet \hat{u}_h^\theta - (\nabla_{\Gamma_h[\mathbf{y}^\theta]} v_h^\theta - n_{\Gamma_h[\mathbf{y}^\theta]} n_{\Gamma_h[\mathbf{y}^\theta]}^\top (\nabla_{\Gamma_h[\mathbf{y}^\theta]} v_h^\theta)^\top) \nabla_{\Gamma_h[\mathbf{y}^\theta]} \hat{u}_h^\theta.$$

This, together with Lemma 3.3 and (C.6), implies  $\|\partial_{t,h}^\bullet \nabla_{\Gamma_h[\mathbf{y}^\theta]} \hat{u}_h^\theta\|_{L^6(\Gamma_h[\mathbf{y}^\theta])} \leq C$  and therefore

$$|W_6| \leq C \|\partial_{t,h}^\bullet \nabla_{\Gamma_h[\mathbf{y}^\theta]} \hat{u}_h^\theta\|_{L^6(\Gamma_h[\mathbf{y}^\theta])} \|D_{\Gamma_h[\mathbf{y}^\theta]} \hat{e}_y^\theta\|_{L^3(\Gamma_h[\mathbf{y}^\theta])} \|\nabla_{\Gamma_h[\mathbf{y}^\theta]} \hat{\varphi}_h^\theta\|_{L^2(\Gamma_h[\mathbf{y}^\theta])} \leq Ch^k \|\varphi^l\|_{H^1(\Gamma)}.$$

The estimate of  $|W_3|$  depends on the estimate of  $\partial_{t,h}^\bullet \hat{\varphi}_h^\theta$ . Since  $(\partial_{t,h}^\bullet \varphi)^l = (\partial_{t,h}^\bullet \varphi^l) = 0$ , as shown in (C.23), where  $\varphi = (\varphi^l)^{-l}$  and  $\partial_{t,h}^\bullet \varphi^l$  denotes the material derivative with respect to velocity  $\tilde{v}_h^*$ , differentiating equation  $0 = \int_{\Gamma_h[\mathbf{x}^*]} (P_h \varphi - \varphi) \chi_h$  with respect to time, we obtain the following relation for all functions  $\chi_h \in S_h(\Gamma_h[\mathbf{x}^*])$  such that  $\partial_{t,h}^\bullet \chi_h = 0$ :

$$\begin{aligned}
0 &= \frac{d}{dt} \int_{\Gamma_h[\mathbf{x}^*]} (P_h \varphi - \varphi) \chi_h \\
&= \int_{\Gamma_h[\mathbf{x}^*]} \partial_{t,h}^\bullet (P_h \varphi - \varphi) \chi_h + \int_{\Gamma_h[\mathbf{x}^*]} (P_h \varphi - \varphi) \chi_h \nabla_{\Gamma_h[\mathbf{x}^*]} \cdot v_h^* \\
&= \int_{\Gamma_h[\mathbf{x}^*]} \partial_{t,h}^\bullet P_h \varphi \chi_h + \int_{\Gamma_h[\mathbf{x}^*]} (P_h \varphi - \varphi) \chi_h \nabla_{\Gamma_h[\mathbf{x}^*]} \cdot v_h^*. \tag{C.26}
\end{aligned}$$

This implies the following estimate by the duality argument:

$$\|\partial_{t,h}^\bullet P_h \varphi\|_{L^2(\Gamma_h[\mathbf{x}^*])} \leq C \|P_h \varphi - \varphi\|_{L^2(\Gamma_h[\mathbf{x}^*])} \leq Ch \|\varphi^l\|_{H^1(\Gamma)}. \tag{C.27}$$

From the definition of  $\partial_{t,h}^\bullet \hat{\varphi}_h^\theta$  and inequality (C.27), along with the norm equivalence on  $\Gamma_h[\mathbf{y}^\theta]$  and  $\Gamma_h[\mathbf{x}^*]$ , we obtain

$$\|\partial_{t,h}^\bullet \hat{\varphi}_h^\theta\|_{L^2(\Gamma_h[\mathbf{y}^\theta])} = \|(\partial_{t,h}^\bullet P_h \varphi) \circ X_h^* \circ (Y_h^\theta)^{-1}\|_{L^2(\Gamma_h[\mathbf{y}^\theta])} \leq Ch \|\varphi^l\|_{H^1(\Gamma)}, \quad (\text{C.28})$$

where  $Y_h^\theta : \Gamma_h^0 \rightarrow \Gamma_h[\mathbf{y}^\theta]$  is the finite element flow map with nodal vector  $\mathbf{y}^\theta$ . By utilizing (C.28) and the estimate of  $\|\nabla_{\Gamma_h[\mathbf{y}^\theta]} \hat{e}_y^\theta\|_{L^2(\Gamma_h[\mathbf{y}^\theta])} \sim \|\nabla_{\Gamma_h[\mathbf{x}^*]}(\text{id}_{\Gamma_h[\mathbf{x}^*]} - \hat{y}_h^*)\|_{L^2(\Gamma_h[\mathbf{x}^*])}$  in Lemma 3.3, we obtain

$$|W_3| \leq Ch^{k+1} \|\varphi^l\|_{H^1(\Gamma)}.$$

In order to estimate  $|W_8|$ , we use relation (3.18) again to write  $\partial_{t,h}^\bullet \nabla_{\Gamma_h[\mathbf{y}^\theta]} \hat{\varphi}_h^\theta$  as

$$\partial_{t,h}^\bullet \nabla_{\Gamma_h[\mathbf{y}^\theta]} \hat{\varphi}_h^\theta = \nabla_{\Gamma_h[\mathbf{y}^\theta]} \partial_{t,h}^\bullet \hat{\varphi}_h^\theta - (\nabla_{\Gamma_h[\mathbf{y}^\theta]} v_h^\theta - n_{\Gamma_h[\mathbf{y}^\theta]} n_{\Gamma_h[\mathbf{y}^\theta]}^\top (\nabla_{\Gamma_h[\mathbf{y}^\theta]} v_h^\theta)^\top) \nabla_{\Gamma_h[\mathbf{y}^\theta]} \hat{\varphi}_h^\theta.$$

where  $\nabla_{\Gamma_h[\mathbf{y}^\theta]} \partial_{t,h}^\bullet \hat{\varphi}_h^\theta$  can be estimated by combining the inverse inequality with (C.28), i.e.,

$$\|\nabla_{\Gamma_h[\mathbf{y}^\theta]} \partial_{t,h}^\bullet \hat{\varphi}_h^\theta\|_{L^2(\Gamma_h[\mathbf{y}^\theta])} \leq Ch^{-1} \|\partial_{t,h}^\bullet \hat{\varphi}_h^\theta\|_{L^2(\Gamma_h[\mathbf{y}^\theta])} \leq C \|\varphi^l\|_{H^1(\Gamma)}. \quad (\text{C.29})$$

This, together with the estimate of  $\|\nabla_{\Gamma_h[\mathbf{y}^\theta]} \hat{e}_y^\theta\|_{L^2(\Gamma_h[\mathbf{y}^\theta])} \sim \|\nabla_{\Gamma_h[\mathbf{x}^*]}(\text{id}_{\Gamma_h[\mathbf{x}^*]} - \hat{y}_h^*)\|_{L^2(\Gamma_h[\mathbf{x}^*])}$  in Lemma 3.3, leads to the following result:

$$|W_8| \leq Ch^k \|\varphi^l\|_{H^1(\Gamma)}.$$

Since  $\partial_{t,h}^\bullet \text{id}_{\Gamma_h[\mathbf{x}^*]} = \hat{I}_h^* v = -\hat{I}_h^*(Hn) = -\hat{I}_h^*(\hat{I}_h^* H \hat{I}_h^* n)$  and  $\partial_{t,h}^\bullet \hat{y}_h^* = -\hat{I}_h^*(\hat{H}^* \hat{n}^*)$  on  $\Gamma_h[\mathbf{x}^*]$ . The latter follows from relation  $\dot{\mathbf{y}}^* = -\mathbf{I}_h(\mathbf{H}^* \bullet \mathbf{n}^*)$  in (3.6), the following identity holds:

$$\begin{aligned} \partial_{t,h}^\bullet \hat{e}_y^\theta &= \left[ \frac{d}{dt} ((\hat{y}_h^* - \text{id}_{\Gamma_h[\mathbf{x}^*]}) \circ X_h^*) \right] \circ (Y_h^\theta)^{-1} = \partial_{t,h}^\bullet (\hat{y}_h^* - \text{id}_{\Gamma_h[\mathbf{x}^*]}) \circ X_h^* \circ (Y_h^\theta)^{-1} \\ &= -(\hat{I}_h^*(\hat{H}^* \hat{n}^*) - \hat{I}_h^*(\hat{I}_h^* H \hat{I}_h^* n)) \circ X_h^* \circ (Y_h^\theta)^{-1}. \end{aligned}$$

By using the stability of  $\hat{I}_h^*$  in the  $H^1(\Gamma_h[\mathbf{x}^*])$  norm for products of finite element functions (see [32, Lemma 5.3]) and Lemma 3.3, we obtain:

$$\begin{aligned} \|\partial_{t,h}^\bullet \hat{e}_y^\theta\|_{H^1(\Gamma_h[\mathbf{y}^\theta])} &\leq C \|\hat{I}_h^*(\hat{H}^* \hat{n}^*) - \hat{I}_h^*(\hat{I}_h^* H \hat{I}_h^* n)\|_{H^1(\Gamma_h[\mathbf{x}^*])} \\ &\leq C \|\hat{H}^* - \hat{I}_h^* H\|_{H^1(\Gamma_h[\mathbf{x}^*])} + C \|\hat{n}^* - \hat{I}_h^* n\|_{H^1(\Gamma_h[\mathbf{x}^*])} \leq Ch^k. \end{aligned} \quad (\text{C.30})$$

Using the inequality (C.30) and the estimate of  $\|\nabla_{\Gamma_h[\mathbf{y}^\theta]} \hat{e}_y^\theta\|_{L^2(\Gamma_h[\mathbf{y}^\theta])} \sim \|\nabla_{\Gamma_h[\mathbf{x}^*]}(\text{id}_{\Gamma_h[\mathbf{x}^*]} - \hat{y}_h^*)\|_{L^2(\Gamma_h[\mathbf{x}^*])}$  in Lemma 3.3, we derive that

$$|W_4| + |W_7| \leq Ch^k \|\varphi^l\|_{H^1(\Gamma)}.$$

As a result, substituting  $\varphi^l = \partial_{t,h}^\bullet w$  into (C.25) yields

$$\|\partial_{t,h}^\bullet w\|_{H^1(\Gamma)} \leq Ch^k. \quad (\text{C.31})$$

**C.6. Proof of the  $H^1$  estimates in Lemma 3.5.** Now, substituting (C.14) and (C.31) into (C.11), along with assumption (C.6), we obtain

$$\begin{aligned} \|(\partial_{t,h}^\bullet \hat{u}_h^*)^l - \partial_{t,h}^\bullet u\|_{H^1(\Gamma)} &\leq Ch^k + C \inf_{\chi_h \in S_h(\Gamma_h[\mathbf{x}^*])} \|\partial_{t,h}^\bullet u - \chi_h^l\|_{H^1(\Gamma)} \\ &= Ch^k + C \inf_{\chi_h \in S_h(\Gamma_h[\mathbf{x}^*])} \|\partial_{t,h}^\bullet u + (\tilde{v}_h^* - v) \cdot \nabla \Gamma u - \chi_h^l\|_{H^1(\Gamma)} \\ &\leq Ch^k + C \|\partial_{t,h}^\bullet u - (I_h^* \partial_{t,h}^\bullet u)^l\|_{H^1(\Gamma)}, \end{aligned}$$

where the last inequality follows from choosing  $\chi_h = I_h^* \partial_{t,h}^\bullet u$  and utilizing  $\|\tilde{v}_h^* - v\|_{H^1(\Gamma)} \leq Ch^k$ . The latter has been shown in (C.2). Since  $\partial_{t,h}^\bullet u$  is a smooth function on  $\Gamma$ , it follows that  $\|\partial_{t,h}^\bullet u - (I_h^* \partial_{t,h}^\bullet u)^l\|_{H^1(\Gamma)} \leq Ch^k$  and therefore

$$\|(\partial_{t,h}^\bullet \hat{u}_h^*)^l - \partial_{t,h}^\bullet u\|_{H^1(\Gamma)} \leq Ch^k \quad \text{for } t \in [0, t_* + \delta]. \quad (\text{C.32})$$

**C.7. Estimate for  $\|\partial_{t,h}^\bullet w\|_{L^2(\Gamma)}$ .** We estimate  $\|\partial_{t,h}^\bullet w\|_{L^2(\Gamma)}$  via a duality argument. Specifically, we define  $\varphi^l \in H^2(\Gamma)^4$  to be the solution of

$$\varphi^l - \Delta_\Gamma \varphi^l = \partial_{t,h}^\bullet w \text{ on } \Gamma, \quad (\text{C.33})$$

which satisfies the following standard  $H^2$  regularity estimate:

$$\|\varphi^l\|_{H^2(\Gamma)} \leq C \|\partial_{t,h}^\bullet w\|_{L^2(\Gamma)}. \quad (\text{C.34})$$

By testing (C.33) with  $\partial_{t,h}^\bullet w$  and estimating the right-hand side of (C.25), we shall prove the following result:

$$\|\partial_{t,h}^\bullet w\|_{L^2(\Gamma)}^2 = \int_\Gamma (\partial_{t,h}^\bullet w \cdot \varphi^l + \nabla_\Gamma \partial_{t,h}^\bullet w \cdot \nabla_\Gamma \varphi^l) \leq Ch^{k+1} \|\varphi^l\|_{H^2(\Gamma)}. \quad (\text{C.35})$$

This, together with (C.34), would imply the following estimate:

$$\|\partial_{t,h}^\bullet w\|_{L^2(\Gamma)} \leq Ch^{k+1}. \quad (\text{C.36})$$

It remains to prove the inequality in (C.35). This is achieved by estimating  $W_j$ ,  $j = 1, \dots, 9$ , on the right-hand side of (C.25) as follows.

The same as before (see Section C.5), we denote by  $\hat{u}_h^\theta$ ,  $\hat{\varphi}_h^\theta$  and  $\hat{e}_y^\theta$  the finite element functions on surface  $\Gamma_h[\mathbf{y}^\theta]$  with nodal vectors  $\mathbf{u}^*$ ,  $\mathbf{P}_h \varphi$  and  $\mathbf{e}_y := \mathbf{y}^* - \mathbf{x}^*$ , respectively, and denote  $\hat{\varphi}_h^* = \hat{\varphi}_h^0 = P_h \varphi \in S_h[\mathbf{x}^*] = S_h[\mathbf{y}^0]$ . Then the surface  $\Gamma_h[\mathbf{y}^\theta]$  moves with velocity  $\hat{e}_y^\theta$  as  $\theta \in [0, 1]$  changes, and

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

When  $\theta$  is fixed and  $t$  increases, the surface  $\Gamma_h[\mathbf{y}^\theta]$  moves with velocity  $v_h^\theta$ , which is a finite element function on  $\Gamma_h[\mathbf{y}^\theta]$  with nodal vector

$$\mathbf{v}^\theta = (1 - \theta)\mathbf{I}_h v + \theta \mathbf{v}^*.$$

By comparing  $\int_{\Gamma_h[\mathbf{y}^\theta]} \hat{u}_h^\theta \cdot \hat{\varphi}_h^\theta \nabla_{\Gamma_h[\mathbf{y}^\theta]} \cdot \hat{e}_y^\theta \nabla_{\Gamma_h[\mathbf{y}^\theta]} \cdot v_h^\theta d\theta$  with its value at  $\theta = 0$ , and using the Newton–Leibniz rule, we have

$$\begin{aligned} W_1 &= \int_0^1 \int_{\Gamma_h[\mathbf{y}^\theta]} \hat{u}_h^\theta \cdot \hat{\varphi}_h^\theta \nabla_{\Gamma_h[\mathbf{y}^\theta]} \cdot \hat{e}_y^\theta \nabla_{\Gamma_h[\mathbf{y}^\theta]} \cdot v_h^\theta d\theta \\ &= \int_0^1 \int_0^\theta \frac{d}{d\alpha} \int_{\Gamma_h[\mathbf{y}^\alpha]} \hat{u}_h^\alpha \cdot \hat{\varphi}_h^\alpha \nabla_{\Gamma_h[\mathbf{y}^\alpha]} \cdot \hat{e}_y^\alpha \nabla_{\Gamma_h[\mathbf{y}^\alpha]} \cdot v_h^\alpha d\alpha d\theta \quad (\text{Newton–Leibniz rule}) \\ &\quad + \int_{\Gamma_h[\mathbf{x}^*]} \hat{u}_h^* \cdot \hat{\varphi}_h^* \nabla_{\Gamma_h[\mathbf{x}^*]} \cdot \hat{e}_y \nabla_{\Gamma_h[\mathbf{x}^*]} \cdot \hat{I}_h^* v \quad (\text{value at } \theta = 0) \\ &=: W_{11} + W_{12}. \end{aligned}$$

Since  $\partial_\alpha \hat{u}_h^\alpha = \partial_\alpha \hat{e}_y^\alpha = 0$ ,  $\partial_\alpha \hat{\varphi}_h^\alpha = 0$  and  $\partial_\alpha v_h^\alpha = e_v^\alpha$ , it follows that, by using formula (3.18),

$$\begin{aligned} |W_{11}| &\leq C (\|\hat{e}_y^\alpha\|_{W^{1,6}(\Gamma_h[\mathbf{y}^\alpha])}^2 + \|\hat{e}_y^\alpha\|_{W^{1,6}(\Gamma_h[\mathbf{y}^\alpha])} \|\hat{e}_v^\alpha\|_{W^{1,6}(\Gamma_h[\mathbf{y}^\alpha])}) \|\varphi^l\|_{H^1(\Gamma)} \\ &\leq Ch^{2k} \|\varphi^l\|_{H^1(\Gamma)}, \end{aligned}$$

where we have used Lemma 3.3 and (C.24) in estimating  $\|\hat{e}_y^\alpha\|_{W^{1,6}(\Gamma_h[\mathbf{y}^\alpha])}$  and  $\|\hat{e}_v^\alpha\|_{W^{1,6}(\Gamma_h[\mathbf{y}^\alpha])}$  in the last inequality, respectively. By estimating the surface perturbation error between  $\Gamma_h[\mathbf{x}^*]$  and  $\Gamma$ , and using integration by parts for the integral on  $\Gamma$  to remove the gradient operator  $\nabla_\Gamma$  from  $(\hat{e}_y)^l$ , along with error estimates of Lagrange interpolation, we have

$$\begin{aligned} |W_{12}| &\leq \left| \int_{\Gamma_h[\mathbf{x}^*]} \hat{u}_h^* \cdot \hat{\varphi}_h^* \nabla_{\Gamma_h[\mathbf{x}^*]} \cdot \hat{e}_y \nabla_{\Gamma_h[\mathbf{x}^*]} \cdot \hat{I}_h^* v - \int_\Gamma (\hat{u}_h^*)^l \cdot (\hat{\varphi}_h^*)^l \nabla_\Gamma \cdot (\hat{e}_y)^l \nabla_\Gamma \cdot (\hat{I}_h^* v)^l \right| \\ &\quad + \left| \int_\Gamma (\hat{u}_h^*)^l \cdot (\hat{\varphi}_h^*)^l \nabla_\Gamma \cdot (\hat{e}_y)^l (\nabla_\Gamma \cdot (\hat{I}_h^* v)^l - \nabla_\Gamma \cdot v) \right| + \left| \int_\Gamma (\hat{u}_h^*)^l \cdot (\hat{\varphi}_h^*)^l \nabla_\Gamma \cdot (\hat{e}_y)^l \nabla_\Gamma \cdot v \right| \\ &\leq Ch^{2k} \|\varphi^l\|_{H^1(\Gamma)} + Ch^{2k} \|\varphi^l\|_{H^1(\Gamma)} + Ch^{k+1} \|\varphi^l\|_{H^1(\Gamma)} \leq Ch^{k+1} \|\varphi^l\|_{H^1(\Gamma)}. \end{aligned}$$

This proves that

$$|W_1| \leq Ch^{k+1} \|\varphi^l\|_{H^1(\Gamma)}. \quad (\text{C.37})$$

Similarly, by converting the integral from  $\Gamma_h[\mathbf{y}^\theta]$  to  $\Gamma$  and then applying integration by parts along with the induction assumption (C.6), we can derive the following estimate:

$$|W_2| := \left| \int_0^1 \int_{\Gamma_h[\mathbf{y}^\theta]} \partial_{t,h}^\bullet \hat{u}_h^\theta \cdot \hat{\varphi}_h^\theta \nabla_{\Gamma_h[\mathbf{y}^\theta]} \cdot \hat{e}_y^\theta d\theta \right| \leq Ch^{k+1} \|\varphi^l\|_{H^1(\Gamma)}. \quad (\text{C.38})$$

By using (C.28), we have

$$|W_3| \leq Ch^{k+1} \|\varphi^l\|_{H^1(\Gamma)}. \quad (\text{C.39})$$

By using (B.2), the first inequality in (C.27), and the first equality in (C.28), we have

$$\begin{aligned} \|\partial_{t,h}^\bullet \hat{\varphi}_h^\theta\|_{H^1(\Gamma_h[\mathbf{x}^*])} &\leq Ch^{-1} \|\partial_{t,h}^\bullet \hat{\varphi}_h^\theta\|_{L^2(\Gamma_h[\mathbf{x}^*])} \\ &\leq Ch^{-1} \|P_h \varphi - \varphi\|_{L^2(\Gamma_h[\mathbf{x}^*])} \leq Ch \|\varphi^l\|_{H^2(\Gamma)}, \end{aligned}$$

which implies that

$$|W_8| \leq Ch^{k+1} \|\varphi^l\|_{H^2(\Gamma)}. \quad (\text{C.40})$$

Since  $\partial_{t,h}^\bullet \hat{e}_y^\theta$  has the same nodal vector as  $\hat{v}_h^* - \hat{I}_h^* v$ , by the similar method along with Lemma 3.4 and inequality (3.58), we can obtain the following estimates for  $W_j$ ,  $j = 4, 5, 7, 9$ , i.e.,

$$|W_4| + |W_5| + |W_7| + |W_9| \leq Ch^{k+1} \|\varphi^l\|_{H^2(\Gamma)}. \quad (\text{C.41})$$

It remains to prove the following result:

$$|W_6| \leq Ch^{k+1} \|\varphi^l\|_{H^2(\Gamma)}. \quad (\text{C.42})$$

This can be done similarly by converting the integral in the expression of  $W_6$  from  $\Gamma_h[\mathbf{y}^\theta]$  to  $\Gamma_h[\mathbf{x}^*]$ , and using the Newton–Leibniz formula, we have

$$\begin{aligned} W_6 &= \int_0^1 \int_{\Gamma_h[\mathbf{y}^\theta]} \partial_{t,h}^\bullet \nabla_{\Gamma_h[\mathbf{y}^\theta]} \hat{u}_h^\theta \cdot D_{\Gamma_h[\mathbf{y}^\theta]} \hat{e}_y^\theta \nabla_{\Gamma_h[\mathbf{y}^\theta]} \hat{\varphi}_h^\theta d\theta \\ &= \int_0^1 \int_0^\theta \frac{d}{d\alpha} \int_{\Gamma_h[\mathbf{y}^\alpha]} \partial_{t,h}^\bullet \nabla_{\Gamma_h[\mathbf{y}^\alpha]} \hat{u}_h^\alpha \cdot D_{\Gamma_h[\mathbf{y}^\alpha]} \hat{e}_y^\alpha \nabla_{\Gamma_h[\mathbf{y}^\alpha]} \hat{\varphi}_h^\alpha d\alpha d\theta \\ &\quad + \int_{\Gamma_h[\mathbf{x}^*]} \partial_{t,h}^\bullet \nabla_{\Gamma_h[\mathbf{x}^*]} \hat{u}_h^* \cdot D_{\Gamma_h[\mathbf{x}^*]} \hat{e}_y \nabla_{\Gamma_h[\mathbf{x}^*]} \hat{\varphi}_h^* d\theta \\ &=: W_{61} + W_{62}, \end{aligned} \quad (\text{C.43})$$

where  $W_{61}$  can be rewritten as follows:

$$\begin{aligned} W_{61} &= \int_0^1 \int_0^\theta \int_{\Gamma_h^\alpha} \partial_{t,h}^\bullet \nabla_{\Gamma_h[\mathbf{y}^\alpha]} \hat{u}_h^\alpha \cdot D_{\Gamma_h[\mathbf{y}^\alpha]} \hat{e}_y^\alpha \nabla_{\Gamma_h[\mathbf{y}^\alpha]} \hat{\varphi}_h^\alpha \nabla_{\Gamma_h[\mathbf{y}^\alpha]} \cdot \hat{e}_y^\alpha d\alpha d\theta \\ &\quad + \int_0^1 \int_0^\theta \int_{\Gamma_h[\mathbf{y}^\alpha]} \partial_\alpha^\bullet (\partial_{t,h}^\bullet \nabla_{\Gamma_h[\mathbf{y}^\alpha]} \hat{u}_h^\alpha) \cdot D_{\Gamma_h[\mathbf{y}^\alpha]} \hat{e}_y^\alpha \nabla_{\Gamma_h[\mathbf{y}^\alpha]} \hat{\varphi}_h^\alpha d\alpha d\theta \\ &\quad + \int_0^1 \int_0^\theta \int_{\Gamma_h[\mathbf{y}^\alpha]} \partial_{t,h}^\bullet \nabla_{\Gamma_h[\mathbf{y}^\alpha]} \hat{u}_h^\alpha \cdot \partial_\alpha^\bullet (D_{\Gamma_h[\mathbf{y}^\alpha]} \hat{e}_y^\alpha) \nabla_{\Gamma_h[\mathbf{y}^\alpha]} \hat{\varphi}_h^\alpha d\alpha d\theta \\ &\quad + \int_0^1 \int_0^\theta \int_{\Gamma_h[\mathbf{y}^\alpha]} \partial_{t,h}^\bullet \nabla_{\Gamma_h[\mathbf{y}^\alpha]} \hat{u}_h^\alpha \cdot D_{\Gamma_h[\mathbf{y}^\alpha]} \hat{e}_y^\alpha \partial_\alpha^\bullet (\nabla_{\Gamma_h[\mathbf{y}^\alpha]} \hat{\varphi}_h^\alpha) d\alpha d\theta \\ &=: W_{611} + W_{612} + W_{613} + W_{614}. \end{aligned}$$

Since the nodal vector of  $\partial_{t,h}^{\bullet} \hat{u}_h^\alpha$  is  $\hat{\mathbf{u}}$ , which is independent of  $\alpha$ , it follows that  $\partial_\alpha^\bullet \partial_{t,h}^{\bullet} \hat{u}_h^\alpha = 0$ . Hence, by applying  $\partial_{t,h}^{\bullet}$  and  $\partial_\alpha^\bullet$  to  $\nabla_{\Gamma_h[\mathbf{y}^\alpha]} \hat{u}_h^\alpha$  sequentially and using relation (3.18), we obtain

$$\begin{aligned} \partial_{t,h}^{\bullet} \nabla_{\Gamma_h[\mathbf{y}^\alpha]} \hat{u}_h^\alpha &= \nabla_{\Gamma_h[\mathbf{y}^\alpha]} \partial_{t,h}^{\bullet} \hat{u}_h^\alpha - (\nabla_{\Gamma_h[\mathbf{y}^\alpha]} v_h^\alpha - n_{\Gamma_h[\mathbf{y}^\alpha]} n_{\Gamma_h[\mathbf{y}^\alpha]}^\top (\nabla_{\Gamma_h[\mathbf{y}^\alpha]} v_h^\alpha)^\top) \nabla_{\Gamma_h[\mathbf{y}^\alpha]} \hat{u}_h^\alpha \\ \partial_\alpha^\bullet \partial_{t,h}^{\bullet} \nabla_{\Gamma_h[\mathbf{y}^\alpha]} \hat{u}_h^\alpha &= -(\nabla_{\Gamma_h[\mathbf{y}^\alpha]} \hat{e}_y^\alpha - n_{\Gamma_h[\mathbf{y}^\alpha]} n_{\Gamma_h[\mathbf{y}^\alpha]}^\top (\nabla_{\Gamma_h[\mathbf{y}^\alpha]} \hat{e}_y^\alpha)^\top) \nabla_{\Gamma_h[\mathbf{y}^\alpha]} \partial_{t,h}^{\bullet} \hat{u}_h^\alpha \\ &\quad - \partial_\alpha^\bullet (\nabla_{\Gamma_h[\mathbf{y}^\alpha]} v_h^\alpha - n_{\Gamma_h[\mathbf{y}^\alpha]} n_{\Gamma_h[\mathbf{y}^\alpha]}^\top (\nabla_{\Gamma_h[\mathbf{y}^\alpha]} v_h^\alpha)^\top) \nabla_{\Gamma_h[\mathbf{y}^\alpha]} \hat{u}_h^\alpha \\ &\quad + (\nabla_{\Gamma_h[\mathbf{y}^\alpha]} v_h^\alpha - n_{\Gamma_h[\mathbf{y}^\alpha]} n_{\Gamma_h[\mathbf{y}^\alpha]}^\top (\nabla_{\Gamma_h[\mathbf{y}^\alpha]} v_h^\alpha)^\top) \\ &\quad \cdot (\nabla_{\Gamma_h[\mathbf{y}^\alpha]} \hat{e}_y^\alpha - n_{\Gamma_h[\mathbf{y}^\alpha]} n_{\Gamma_h[\mathbf{y}^\alpha]}^\top (\nabla_{\Gamma_h[\mathbf{y}^\alpha]} \hat{e}_y^\alpha)^\top) \nabla_{\Gamma_h[\mathbf{y}^\alpha]} \hat{u}_h^\alpha \end{aligned}$$

where

$$\begin{aligned} \partial_\alpha^\bullet (\nabla_{\Gamma_h[\mathbf{y}^\alpha]} v_h^\alpha) &= \nabla_{\Gamma_h[\mathbf{y}^\alpha]} \partial_\alpha^\bullet v_h^\alpha - (\nabla_{\Gamma_h[\mathbf{y}^\alpha]} \hat{e}_y^\alpha - n_{\Gamma_h[\mathbf{y}^\alpha]} n_{\Gamma_h[\mathbf{y}^\alpha]}^\top (\nabla_{\Gamma_h[\mathbf{y}^\alpha]} \hat{e}_y^\alpha)^\top) \nabla_{\Gamma_h[\mathbf{y}^\alpha]} v_h^\alpha \\ \partial_\alpha^\bullet n_{\Gamma_h[\mathbf{y}^\alpha]} &= -(\nabla_{\Gamma_h[\mathbf{y}^\alpha]} \hat{e}_y^\alpha) n_{\Gamma_h[\mathbf{y}^\alpha]}. \end{aligned}$$

Since  $\partial_\alpha^\bullet v_h^\alpha = e_v^\alpha$ , which has been estimated in (C.24), it follows that

$$\begin{aligned} &\|\partial_\alpha^\bullet (\nabla_{\Gamma_h[\mathbf{y}^\alpha]} v_h^\alpha)\|_{L^6(\Gamma_h[\mathbf{y}^\alpha])} + \|\partial_\alpha^\bullet n_{\Gamma_h[\mathbf{y}^\alpha]}\|_{L^6(\Gamma_h[\mathbf{y}^\alpha])} \\ &\leq C \|e_v^\alpha\|_{W^{1,6}(\Gamma_h[\mathbf{y}^\alpha])} + C \|\hat{e}_y^\alpha\|_{W^{1,6}(\Gamma_h[\mathbf{y}^\alpha])} \leq Ch^k. \end{aligned}$$

By using this result and the estimate of  $\|\nabla_{\Gamma_h[\mathbf{y}^\alpha]} \hat{e}_y^\alpha\|_{L^p(\Gamma_h[\mathbf{y}^\alpha])} \sim \|\nabla_{\Gamma_h[\mathbf{x}^*]} (\text{id}_{\Gamma_h[\mathbf{x}^*]} - \hat{y}_h^*)\|_{L^p(\Gamma_h[\mathbf{x}^*])}$  in Lemma 3.3, as well as the estimate of  $\|\nabla_{\Gamma_h[\mathbf{y}^\alpha]} \partial_{t,h}^{\bullet} \hat{u}_h^\alpha\|_{L^6(\Gamma_h[\mathbf{y}^\alpha])}$  in (C.6), we have

$$\|\partial_{t,h}^{\bullet} \nabla_{\Gamma_h[\mathbf{y}^\alpha]} \hat{u}_h^\alpha\|_{L^6(\Gamma_h[\mathbf{y}^\alpha])} \leq C$$

and

$$\begin{aligned} \|\partial_\alpha^\bullet \partial_{t,h}^{\bullet} \nabla_{\Gamma_h[\mathbf{y}^\alpha]} \hat{u}_h^\alpha\|_{L^3(\Gamma_h[\mathbf{y}^\alpha])} &\leq C \|\hat{e}_y^\alpha\|_{W^{1,6}(\Gamma_h[\mathbf{y}^\alpha])} \|\partial_{t,h}^{\bullet} \hat{u}_h^\alpha\|_{W^{1,6}(\Gamma_h[\mathbf{y}^\alpha])} \\ &\quad + C \|\partial_\alpha^\bullet (\nabla_{\Gamma_h[\mathbf{y}^\alpha]} v_h^\alpha)\|_{L^6(\Gamma_h[\mathbf{y}^\alpha])} + C \|\partial_\alpha^\bullet n_{\Gamma_h[\mathbf{y}^\alpha]}\|_{L^6(\Gamma_h[\mathbf{y}^\alpha])} \\ &\quad + C \|\nabla_{\Gamma_h[\mathbf{y}^\alpha]} \hat{e}_y^\alpha\|_{L^6(\Gamma_h[\mathbf{y}^\alpha])} \\ &\leq Ch^k. \end{aligned}$$

This implies that

$$\begin{aligned} |W_{611}| &\leq C \|\partial_{t,h}^{\bullet} \nabla_{\Gamma_h[\mathbf{y}^\alpha]} \hat{u}_h^\alpha\|_{L^6(\Gamma_h[\mathbf{y}^\alpha])} \|\hat{e}_y^\alpha\|_{W^{1,6}(\Gamma_h[\mathbf{y}^\alpha])}^2 \|\hat{\varphi}_h^\alpha\|_{H^1(\Gamma_h[\mathbf{y}^\alpha])} \leq Ch^{2k} \|\varphi^l\|_{H^1(\Gamma)} \\ |W_{612}| &\leq C \|\partial_\alpha^\bullet \partial_{t,h}^{\bullet} \nabla_{\Gamma_h[\mathbf{y}^\alpha]} \hat{u}_h^\alpha\|_{L^3(\Gamma_h[\mathbf{y}^\alpha])} \|\hat{e}_y^\alpha\|_{W^{1,6}(\Gamma_h[\mathbf{y}^\alpha])} \|\hat{\varphi}_h^\alpha\|_{H^1(\Gamma_h[\mathbf{y}^\alpha])} \leq Ch^{2k} \|\varphi^l\|_{H^1(\Gamma)}. \end{aligned}$$

The estimates of  $|W_{613}|$  and  $|W_{614}|$  can be done similarly by using relation (3.18), i.e.,

$$\begin{aligned} &|W_{613}| + |W_{614}| \\ &\leq C \|\partial_{t,h}^{\bullet} \nabla_{\Gamma_h[\mathbf{y}^\alpha]} \hat{u}_h^\alpha\|_{L^6(\Gamma_h[\mathbf{y}^\alpha])} \|\hat{e}_y^\alpha\|_{W^{1,6}(\Gamma_h[\mathbf{y}^\alpha])}^2 \|\hat{\varphi}_h^\alpha\|_{H^1(\Gamma_h[\mathbf{y}^\alpha])} \leq Ch^{2k} \|\varphi^l\|_{H^1(\Gamma)}. \end{aligned}$$

Therefore, we have

$$|W_{61}| \leq Ch^{k+1} \|\varphi^l\|_{H^1(\Gamma)}. \quad (\text{C.44})$$

The estimate of  $|W_{62}|$  can be obtained by changing the underlying surface from  $\Gamma_h[\mathbf{x}^*]$  to  $\Gamma$  and then performing integration by parts on  $\Gamma$ . This can be done by defining a family of intermediate surfaces  $\Gamma^\theta$ ,  $\theta \in [0, 1]$ , between smooth surface  $\Gamma$  and interpolated surface  $\Gamma_h[\mathbf{x}^*]$ , i.e.,

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

Let  $u^\theta$ ,  $\varphi^\theta$ ,  $\tilde{v}_h^\theta$ ,  $\hat{e}_y^\theta$  and  $e^\theta$  be functions on  $\Gamma^\theta$  defined by

$$\begin{aligned} \hat{u}^\theta(x^\theta) &= (1 - \theta)u(x^l) + \theta \hat{u}_h^*(x), \\ \hat{\varphi}^\theta(x^\theta) &= (1 - \theta)\varphi^l(x^l) + \theta \hat{\varphi}_h^*(x), \end{aligned}$$

$$\begin{aligned}\hat{e}_y^\theta(x^\theta) &= \hat{e}_y(x) = \hat{y}_h^* - x \quad (\text{see the definition in Section 3.3}), \\ e^\theta(x^\theta) &= x - x^l \quad \text{for } x \in \Gamma_h[\mathbf{x}^*].\end{aligned}$$

By this definition,  $\varphi^0 = \varphi^l$  and  $\hat{e}_y^0 = \hat{e}_y^l$ , and surface  $\Gamma^\theta$  moves with velocity  $e^\theta$  as  $\theta \in [0, 1]$  changes. Moreover, the following relations hold:

$$\partial_\theta^\bullet \hat{u}^\theta(x^\theta) = \hat{u}_h^*(x) - u(x^l), \quad \partial_\theta^\bullet \varphi^\theta(x^\theta) = \hat{\varphi}_h^*(x) - \varphi^l(x^l), \quad \text{and} \quad \partial_\theta^\bullet \hat{e}_y^\theta(x^\theta) = 0.$$

Then we can rewrite  $W_{62}$  as follows by using formula (3.18):

$$\begin{aligned}W_{62} &= \int_{\Gamma_h[\mathbf{x}^*]} \nabla_{\Gamma_h[\mathbf{x}^*]} \partial_{t,h}^\bullet \hat{u}_h^* \cdot D_{\Gamma_h[\mathbf{x}^*]} \hat{e}_y \nabla_{\Gamma_h[\mathbf{x}^*]} \hat{\varphi}_h^* \\ &\quad - \int_{\Gamma_h[\mathbf{x}^*]} [\nabla_{\Gamma_h[\mathbf{x}^*]} \hat{I}_h^* v - n_{\Gamma_h[\mathbf{x}^*]}^* (n_{\Gamma_h[\mathbf{x}^*]}^*)^\top (\nabla_{\Gamma_h[\mathbf{x}^*]} \hat{I}_h^* v)^\top] \nabla_{\Gamma_h[\mathbf{x}^*]} \hat{u}_h^* \cdot D_{\Gamma_h[\mathbf{x}^*]} \hat{e}_y \nabla_{\Gamma_h[\mathbf{x}^*]} \hat{\varphi}_h^* \\ &=: W_{621} + W_{622}.\end{aligned}$$

We can further approximate  $W_{621}$  by  $\int_\Gamma \nabla_\Gamma \partial_{t,h}^\bullet u \cdot D_\Gamma \hat{e}_y^l \nabla_\Gamma \varphi^l$  and use Newton–Leibniz rule, i.e.,

$$\begin{aligned}W_{621} &= \int_0^1 \frac{d}{d\theta} \int_{\Gamma^\theta} \nabla_{\Gamma^\theta} \partial_{t,h}^\bullet \hat{u}^\theta \cdot D_{\Gamma^\theta} \hat{e}_y^\theta \nabla_{\Gamma^\theta} \hat{\varphi}^\theta \, d\theta + \int_\Gamma \nabla_\Gamma \partial_t^\bullet u \cdot D_\Gamma \hat{e}_y^l \nabla_\Gamma \varphi^l \\ &= \int_0^1 \int_{\Gamma^\theta} \nabla_{\Gamma^\theta} \partial_{t,h}^\bullet \hat{u}^\theta \cdot D_{\Gamma^\theta} \hat{e}_y^\theta \nabla_{\Gamma^\theta} \hat{\varphi}^\theta \nabla_{\Gamma^\theta} \cdot e^\theta \, d\theta + \int_0^1 \int_{\Gamma^\theta} \partial_\theta^\bullet \nabla_{\Gamma^\theta} \partial_{t,h}^\bullet \hat{u}^\theta \cdot D_{\Gamma^\theta} \hat{e}_y^\theta \nabla_{\Gamma^\theta} \hat{\varphi}^\theta \, d\theta \\ &\quad + \int_0^1 \int_{\Gamma^\theta} \nabla_{\Gamma^\theta} \partial_{t,h}^\bullet \hat{u}^\theta \cdot \partial_\theta^\bullet D_{\Gamma^\theta} \hat{e}_y^\theta \nabla_{\Gamma^\theta} \hat{\varphi}^\theta \, d\theta + \int_0^1 \int_{\Gamma^\theta} \nabla_{\Gamma^\theta} \partial_{t,h}^\bullet \hat{u}^\theta \cdot D_{\Gamma^\theta} \hat{e}_y^\theta \partial_\theta^\bullet \nabla_{\Gamma^\theta} \hat{\varphi}^\theta \, d\theta \\ &\quad + \int_\Gamma \nabla_\Gamma \partial_t^\bullet u \cdot D_\Gamma \hat{e}_y \nabla_\Gamma \varphi^l =: W_{6211} + W_{6212} + W_{6213} + W_{6214} + W_{6215}.\end{aligned}$$

By using the following estimates

$$\begin{aligned}\|\partial_{t,h}^\bullet \hat{u}^\theta\|_{W^{1,6}(\Gamma^\theta)} &\leq C \quad (\text{as a result of (C.6)}) \\ \|\hat{e}_y^\theta\|_{W^{1,6}(\Gamma^\theta)} &\sim \|\hat{y}_h^* - \text{id}_{\Gamma_h[\mathbf{x}^*]}\|_{W^{1,6}(\Gamma_h[\mathbf{x}^*])} \leq Ch^k \quad (\text{shown in in Lemma 3.3}) \\ \|e^\theta\|_{W^{1,6}(\Gamma^\theta)} &\leq Ch^k \quad (\text{error of interpolating surface } \Gamma).\end{aligned}$$

we have

$$|W_{6211}| \leq \|\partial_{t,h}^\bullet \hat{u}^\theta\|_{W^{1,6}(\Gamma^\theta)} \|\hat{e}_y^\theta\|_{W^{1,6}(\Gamma^\theta)} \|\hat{\varphi}^\theta\|_{H^1(\Gamma^\theta)} \|e^\theta\|_{W^{1,6}(\Gamma^\theta)} \leq Ch^{2k} \|\varphi^l\|_{H^1(\Gamma)}.$$

By using formula (3.18), we see that  $\partial_\theta^\bullet \nabla_{\Gamma^\theta} \partial_{t,h}^\bullet \hat{u}^\theta$  is equal to  $\nabla_{\Gamma^\theta} \partial_\theta^\bullet \partial_{t,h}^\bullet \hat{u}^\theta$  plus a term which is bounded by  $\nabla_{\Gamma^\theta} e^\theta$  times  $\nabla_{\Gamma^\theta} \partial_{t,h}^\bullet \hat{u}^\theta$ . Therefore,

$$\begin{aligned}|W_{6212}| &\leq (\|\nabla_{\Gamma^\theta} \partial_\theta^\bullet \partial_{t,h}^\bullet \hat{u}^\theta\|_{L^2(\Gamma^\theta)} + \|e^\theta\|_{W^{1,3}(\Gamma^\theta)} \|\partial_{t,h}^\bullet \hat{u}^\theta\|_{W^{1,6}(\Gamma^\theta)}) \|\hat{e}_y^\theta\|_{W^{1,3}(\Gamma^\theta)} \|\hat{\varphi}^\theta\|_{W^{1,6}(\Gamma^\theta)} \\ &\leq Ch^{2k} \|\varphi^l\|_{H^2(\Gamma)},\end{aligned}$$

where the last inequality uses the Sobolev embedding  $H^2(\Gamma) \hookrightarrow W^{1,6}(\Gamma)$ , the following result that follows from (C.32):

$$\|\nabla_{\Gamma^\theta} \partial_\theta^\bullet \partial_{t,h}^\bullet \hat{u}^\theta\|_{L^2(\Gamma^\theta)} \leq C \|(\partial_\theta^\bullet \partial_{t,h}^\bullet \hat{u}^\theta)^l\|_{H^1(\Gamma)} = C \|(\partial_{t,h}^* \hat{u}_h^*)^l - \partial_t^\bullet u\|_{H^1(\Gamma)} \leq Ch^k,$$

and the following result pertains to the interpolation error caused by the perturbation of the surface:

$$\|\nabla_{\Gamma^\theta} e^\theta\|_{L^\infty(\Gamma^\theta)} \leq Ch^k.$$

Moreover,  $W_{6213}$  and  $W_{6214}$  can be estimated similarly by using the following results:

$$\|\nabla_{\Gamma^\theta} \partial_\theta^\bullet \hat{\varphi}_h^\theta\|_{L^2(\Gamma^\theta)} \leq \|\varphi^l - (\hat{\varphi}_h^*)^l\|_{H^1(\Gamma)} = \|\varphi^l - (P_h \varphi)^l\|_{H^1(\Gamma)} \leq Ch \|\varphi^l\|_{H^2(\Gamma)}$$

This would yield the following results:

$$|W_{6213}| + |W_{6214}| \leq Ch^{2k} \|\varphi^l\|_{H^2(\Gamma)}.$$

Then,  $W_{6215}$  can be estimated with integration by parts and the estimate of  $\|\hat{e}_y\|_{L^2(\Gamma)} \sim \|\hat{y}_h^* - \text{id}_{\Gamma_h[\mathbf{x}^*]}\|_{L^2(\Gamma_h[\mathbf{x}^*])}$  in Lemma 3.4, with  $|W_{6215}| \leq Ch^{k+1} \|\varphi^l\|_{H^2(\Gamma)}$ . This proves the following result (by summing up the estimates of  $|W_{621j}|$ ,  $j = 1, \dots, 5$ ):

$$|W_{621}| \leq Ch^{k+1} \|\varphi^l\|_{H^2(\Gamma)}.$$

Note that  $|W_{622}|$  can be estimated similarly as  $|W_{621}|$ , with  $|W_{622}| \leq Ch^{k+1} \|\varphi^l\|_{H^2(\Gamma)}$ . By summing up the estimates of  $|W_{621}|$  and  $|W_{622}|$  we obtain

$$|W_{62}| \leq Ch^{k+1} \|\varphi^l\|_{H^2(\Gamma)}. \quad (\text{C.45})$$

Now, substituting (C.44) and (C.45) into (C.43) yields (C.42). Then, substituting the estimates in (C.37)–(C.42) into the right-hand side of (C.25) yields (C.35) and therefore completes the proof of (C.36).

**C.8. Proof of the  $L^2$  estimates in Lemma 3.5.** By substituting estimates (C.14), (C.18), (C.31) and (C.36) into (C.12), we obtain the following result:

$$\begin{aligned} \|(\partial_{t,h}^{\bullet} \hat{u}_h^*)^l - \partial_{t,h}^{\bullet} u\|_{L^2(\Gamma)} &\leq Ch^{k+1} + Ch \inf_{\chi_h \in S_h(\Gamma_h[\mathbf{x}^*])} \|\partial_{t,h}^{\bullet} u - \chi_h^l\|_{H^1(\Gamma)} \\ &\leq Ch^{k+1} + Ch \|\partial_{t,h}^{\bullet} u - (I_h^* \partial_t^{\bullet} u)^l\|_{H^1(\Gamma)} \\ &= Ch^{k+1} + Ch \|\partial_t^{\bullet} u + (\tilde{v}_h^* - v) \cdot \nabla u - (I_h^* \partial_t^{\bullet} u)^l\|_{H^1(\Gamma)} \\ &\leq Ch^{k+1} + Ch \|\partial_t^{\bullet} u - (I_h^* \partial_t^{\bullet} u)^l\|_{H^1(\Gamma)}, \end{aligned}$$

where the last inequality follows from the application of the triangle inequality and the estimate of  $\|\tilde{v}_h^* - v\|_{H^1(\Gamma)}$  in (C.2). The last inequality, together with a standard estimate of interpolation error  $\|\partial_t^{\bullet} u - (I_h^* \partial_t^{\bullet} u)^l\|_{H^1(\Gamma)}$ , implies the following  $L^2$  estimate:

$$\|(\partial_{t,h}^{\bullet} \hat{u}_h^*)^l - \partial_t^{\bullet} u\|_{L^2(\Gamma)} \leq Ch^{k+1} \quad \text{for } t \in [0, t_* + \delta]. \quad (\text{C.46})$$

This proves (3.64)–(3.65) for  $t \in [0, t_*]$  and therefore completes the proof of Lemma 3.5, as discussed at the end of Section C.1.  $\square$

#### Appendix D. Proof of Lemma 3.6.

The proof of Lemma 3.6 is similar as the proof for Lemma 3.5. For the readers' convenience, we present a complete proof and divide into several sections.

For a finite element function  $\chi_u \in S_h[\mathbf{y}^*]^4$  we denote by  $\hat{\chi}_u$  the finite element function on  $\Gamma_h[\mathbf{x}^*]$  with the same nodal vector as  $\chi_u$  and denote by  $\hat{\chi}_u^l$  the lift of  $\hat{\chi}_u$  from  $\Gamma_h[\mathbf{x}^*]$  to  $\Gamma$ . Then equation (1.2) implies that

$$\int_{\Gamma} \partial_t^{\bullet} u \cdot \hat{\chi}_u^l + \int_{\Gamma} \nabla_{\Gamma} u \cdot \nabla_{\Gamma} \hat{\chi}_u^l = \int_{\Gamma} |\nabla_{\Gamma} n|^2 u \cdot \hat{\chi}_u^l \quad \forall \chi_u \in S_h[\mathbf{y}^*]^4. \quad (\text{D.1})$$

The definition of the dynamic Ritz projection in (2.11) implies that

$$\int_{\Gamma_h[\mathbf{y}^*]} \nabla_{\Gamma_h[\mathbf{y}^*]} u_h^* \cdot \nabla_{\Gamma_h[\mathbf{y}^*]} \chi_u - \int_{\Gamma} \nabla_{\Gamma} u \cdot \nabla_{\Gamma} \hat{\chi}_u^l = \int_{\Gamma} u \cdot \hat{\chi}_u^l - \int_{\Gamma_h[\mathbf{y}^*]} u_h^* \cdot \chi_u \quad (\text{D.2})$$

By subtracting (D.1) from (2.14) and utilizing (D.2), we have

$$\begin{aligned}
& \int_{\Gamma_h[\mathbf{y}^*]} d_u^* \cdot \chi_u \\
&= \int_{\Gamma_h[\mathbf{y}^*]} \partial_{t,h}^* u_h^* \cdot \chi_u - \int_{\Gamma} \partial_t^* u \cdot \hat{\chi}_u^l + \int_{\Gamma_h[\mathbf{y}^*]} \nabla_{\Gamma_h[\mathbf{y}^*]} u_h^* \cdot \nabla_{\Gamma_h[\mathbf{y}^*]} \chi_u - \int_{\Gamma} \nabla_{\Gamma} u \cdot \nabla_{\Gamma} \hat{\chi}_u^l \\
&\quad + \int_{\Gamma} |\nabla_{\Gamma} n|^2 u \cdot \hat{\chi}_u^l - \int_{\Gamma_h[\mathbf{y}^*]} |\nabla_{\Gamma_h[\mathbf{y}^*]} n_h^*|^2 u_h^* \cdot \chi_u \\
&= \int_{\Gamma_h[\mathbf{y}^*]} \partial_{t,h}^* u_h^* \cdot \chi_u - \int_{\Gamma} \partial_t^* u \cdot \hat{\chi}_u^l + \int_{\Gamma} u \cdot \hat{\chi}_u^l - \int_{\Gamma_h[\mathbf{y}^*]} u_h^* \cdot \chi_u \\
&\quad + \int_{\Gamma} |\nabla_{\Gamma} n|^2 u \cdot \hat{\chi}_u^l - \int_{\Gamma_h[\mathbf{y}^*]} |\nabla_{\Gamma_h[\mathbf{y}^*]} n_h^*|^2 u_h^* \cdot \chi_u \\
&=: \int_{\Gamma_h[\mathbf{y}^*]} d_{u,1}^* \cdot \chi_u + \int_{\Gamma_h[\mathbf{y}^*]} d_{u,2}^* \cdot \chi_u + \int_{\Gamma_h[\mathbf{y}^*]} d_{u,3}^* \cdot \chi_u, \quad \forall \chi_u \in S_h[\mathbf{y}^*]^4.
\end{aligned} \tag{D.3}$$

The three terms on the right-hand side of (D.3) are estimated in the following lemmas.

**Lemma D.1.** *Under the assumptions of Theorem 2.1, the following estimates hold:*

$$\begin{aligned}
\left| \int_{\Gamma_h[\mathbf{y}^*]} d_{u,1}^* \cdot \chi_u \right| &\leq Ch^{k+1} \|\chi_u\|_{H^1(\Gamma_h[\mathbf{y}^*])}, \\
\left| \int_{\Gamma_h[\mathbf{y}^*]} d_{u,2}^* \cdot \chi_u \right| &\leq Ch^{k+1} \|\chi_u\|_{H^1(\Gamma_h[\mathbf{y}^*])}.
\end{aligned}$$

*Proof.* By using the triangle inequality, we have

$$\begin{aligned}
& \left| \int_{\Gamma_h[\mathbf{y}^*]} d_{u,2}^* \cdot \chi_u \right| \\
&= \left| \int_{\Gamma} u \cdot \hat{\chi}_u^l - \int_{\Gamma_h[\mathbf{y}^*]} u_h^* \cdot \chi_u \right| \\
&\leq \left| \int_{\Gamma} u \cdot \hat{\chi}_u^l - \int_{\Gamma_h[\mathbf{x}^*]} u^{-l} \cdot \hat{\chi}_u \right| + \left| \int_{\Gamma_h[\mathbf{x}^*]} u^{-l} \cdot \hat{\chi}_u - \int_{\Gamma_h[\mathbf{x}^*]} \hat{I}_h^* u \cdot \hat{\chi}_u \right| \\
&\quad + \left| \int_{\Gamma_h[\mathbf{x}^*]} \hat{I}_h^* u \cdot \hat{\chi}_u - \int_{\Gamma_h[\mathbf{y}^*]} u_h^* \cdot \chi_u \right| \\
&\leq Ch^{k+1} \|\chi_u\|_{L^2(\Gamma_h[\mathbf{y}^*])} + Ch^{k+1} \|\chi_u\|_{L^2(\Gamma_h[\mathbf{y}^*])} + \left| \int_{\Gamma_h[\mathbf{x}^*]} \hat{I}_h^* u \cdot \hat{\chi}_u - \int_{\Gamma_h[\mathbf{y}^*]} u_h^* \cdot \chi_u \right|,
\end{aligned} \tag{D.4}$$

where the first term on the right-hand side is obtained by using the standard estimates for the geometric perturbation errors in [30, Lemma 5.6] (also see [33, Lemma 7.4]), and the second term on the right-hand side of the inequality above is obtained by using error estimates for the Lagrange interpolation, i.e.,  $\|u - (\hat{I}_h^* u)^l\|_{L^\infty(\Gamma)} \leq Ch^{k+1}$ . In both terms we have converted  $\|\chi_u\|_{L^2(\Gamma_h[\mathbf{x}^*])}$  to  $\|\chi_u\|_{L^2(\Gamma_h[\mathbf{y}^*])}$  by using the norm equivalence between  $\Gamma_h[\mathbf{x}^*]$  and  $\Gamma_h[\mathbf{y}^*]$ . The third term on the right-hand side of the inequality above can be estimated as follows.

Let  $\mathbf{y}^\theta = (1 - \theta)\mathbf{x}^* + \theta\mathbf{y}^*$ , let  $\chi_u^\theta$ ,  $u_h^\theta$  and  $e_y^\theta$  be the finite element functions on  $\Gamma_h[\mathbf{y}^\theta]$  with nodal vectors  $\chi_u$ ,  $(1 - \theta)\mathbf{I}_h u + \theta\mathbf{u}^*$  and  $\mathbf{e}_y := \mathbf{y}^* - \mathbf{x}^*$ , respectively. Then

$$\partial_\theta^* u_h^\theta = e_u^\theta \quad \text{and} \quad \partial_\theta^* \hat{\chi}_u^\theta = 0,$$

where  $e_u^\theta$  is the finite element function on  $\Gamma_h[\mathbf{y}^\theta]$  with nodal vector  $\mathbf{u}^* - \mathbf{I}_h u$ . The surface  $\Gamma_h[\mathbf{y}^\theta]$  is moving with velocity  $e_y^\theta$  as  $\theta$  changes from 0 to 1. Then, Lemma 3.4 and the equivalence of norms on  $\Gamma_h[\mathbf{y}^\theta]$ ,  $\theta \in [0, 1]$ , imply that

$$\|e_y^\theta\|_{L^2(\Gamma_h[\mathbf{y}^\theta])} + \|e_u^\theta\|_{L^2(\Gamma_h[\mathbf{y}^\theta])} \leq Ch^{k+1}. \tag{D.5}$$

By using (D.5), along with the equivalence of  $L^p$  norm on  $\Gamma_h[\mathbf{y}^\theta]$  and  $\Gamma_h[\mathbf{y}^*]$ , we have

$$\begin{aligned}
& \left| \int_{\Gamma_h[\mathbf{x}^*]} \hat{I}_h^* u \cdot \hat{\chi}_u - \int_{\Gamma_h[\mathbf{y}^*]} u_h^* \cdot \chi_u \right| \\
&= \left| \int_0^1 \frac{d}{d\theta} \int_{\Gamma_h[\mathbf{y}^\theta]} u_h^\theta \cdot \chi_u^\theta d\theta \right| \\
&= \left| \int_0^1 \int_{\Gamma_h[\mathbf{y}^\theta]} (e_u^\theta \cdot \chi_u^\theta + u_h^\theta \cdot \chi_u^\theta \nabla \cdot e_y^\theta) d\theta \right| \\
&\leq C \int_0^1 \|e_u^\theta\|_{L^2(\Gamma_h[\mathbf{y}^\theta])} \|\chi_u^\theta\|_{L^2(\Gamma_h[\mathbf{y}^\theta])} d\theta + \left| \int_0^1 \int_{\Gamma_h[\mathbf{y}^\theta]} u_h^\theta \cdot \chi_u^\theta \nabla_{\Gamma_h[\mathbf{y}^\theta]} \cdot e_y^\theta d\theta \right| \\
&\leq Ch^{k+1} \|\chi_u\|_{L^2(\Gamma_h[\mathbf{y}^*])} + \left| \int_0^1 \int_{\Gamma_h[\mathbf{y}^\theta]} u_h^\theta \cdot \chi_u^\theta \nabla_{\Gamma_h[\mathbf{y}^\theta]} \cdot e_y^\theta d\theta \right|.
\end{aligned}$$

The second term on the right-hand side of the last inequality can be estimated as follows (by using the triangle inequality):

$$\begin{aligned}
\left| \int_0^1 \int_{\Gamma_h[\mathbf{y}^\theta]} u_h^\theta \cdot \chi_u^\theta \nabla_{\Gamma_h[\mathbf{y}^\theta]} \cdot e_y^\theta d\theta \right| &\leq \left| \int_{\Gamma_h[\mathbf{x}^*]} \hat{u}_h^* \cdot \hat{\chi}_u^* \nabla_{\Gamma_h[\mathbf{x}^*]} \cdot e_y^0 - \int_{\Gamma} (\hat{u}_h^*)^l \cdot (\hat{\chi}_u^*)^l (\nabla_{\Gamma_h[\mathbf{x}^*]} \cdot e_y^0)^l \right| \\
&\quad + \left| \int_{\Gamma} (\hat{u}_h^*)^l \cdot (\hat{\chi}_u^*)^l [(\nabla_{\Gamma_h[\mathbf{x}^*]} \cdot e_y^0)^l - \nabla_{\Gamma} \cdot (e_y^0)^l] \right| \\
&\quad + \left| \int_0^1 \int_0^\theta \frac{d}{d\alpha} \int_{\Gamma_h[\mathbf{y}^\alpha]} u_h^\alpha \cdot \chi_h^\alpha \nabla_{\Gamma_h[\mathbf{y}^\alpha]} \cdot e_y^\alpha d\alpha d\theta \right| \\
&\quad + \left| \int_{\Gamma} (\hat{u}_h^*)^l \cdot (\hat{\chi}_u^*)^l \nabla_{\Gamma} \cdot (e_y)^l \right| \\
&=: J_1 + J_2 + J_3 + J_4 \leq Ch^{k+1} \|\chi_u\|_{H^1(\Gamma_h[\mathbf{y}^*])},
\end{aligned}$$

where  $J_1$  is estimated by applying the geometric perturbation errors (see [30, Lemma 5.6] and [33, Lemma 7.4]);  $J_2$  is estimated by applying the chain rule of partial differentiation and Lemma 3.19, which imply that

$$J_2 \leq C \|(\hat{\chi}_u^*)^l\|_{L^2(\Gamma)} h^k \|\nabla_{\Gamma_h[\mathbf{x}^*]} e_y^0\|_{L^2(\Gamma_h[\mathbf{x}^*])} \leq Ch^{2k} \|\hat{\chi}_u^*\|_{L^2(\Gamma_h[\mathbf{y}^*])};$$

$J_3$  is estimated by utilizing identities (3.12) and (3.18);  $J_4$  is estimated by applying integration by parts which transfers the gradient operator  $\nabla_{\Gamma}$  from  $(e_y)^l$  to  $(\hat{u}_h^*)^l \cdot (\hat{\chi}_u^*)^l$ .

This proves the desired upper bound for the third term on the right-hand side of (D.4), and therefore completes the proof for the second result of Lemma D.1. The first result of Lemma D.1 can be proved in the same way by using Lemma 3.5.  $\square$

**Lemma D.2.** *Under the assumptions of Theorem 2.1, the following estimate holds:*

$$\left| \int_{\Gamma_h[\mathbf{y}^*]} d_{u,3}^* \cdot \chi_u \right| \leq Ch^{k+1} \|\chi_u\|_{H^1(\Gamma_h[\mathbf{y}^*])}.$$

*Proof.* Observe that

$$\begin{aligned}
\left| \int_{\Gamma_h[\mathbf{y}^*]} d_{u,3}^* \cdot \chi_u \right| &= \left| \int_{\Gamma} |\nabla_{\Gamma} n|^2 u \cdot \hat{\chi}_u^l - \int_{\Gamma_h[\mathbf{y}^*]} |\nabla_{\Gamma_h[\mathbf{y}^*]} n_h^*|^2 u_h^* \cdot \chi_u \right| \\
&\leq \left| \int_{\Gamma} |\nabla_{\Gamma} n|^2 u \cdot \hat{\chi}_u^l - \int_{\Gamma_h[\mathbf{x}^*]} |\nabla_{\Gamma_h[\mathbf{x}^*]} \hat{I}_h^* n|^2 \hat{I}_h^* u \cdot \hat{\chi}_u \right| \\
&\quad + \left| \int_{\Gamma_h[\mathbf{x}^*]} |\nabla_{\Gamma_h[\mathbf{x}^*]} \hat{I}_h^* n|^2 \hat{I}_h^* u \cdot \hat{\chi}_u - \int_{\Gamma_h[\mathbf{y}^*]} |\nabla_{\Gamma_h[\mathbf{y}^*]} n_h^*|^2 u_h^* \cdot \chi_u \right|. \quad (\text{D.6})
\end{aligned}$$

The first term on the right-hand side of (D.6) can be estimated as follows:

$$\left| \int_{\Gamma} |\nabla_{\Gamma} n|^2 u \cdot \hat{\chi}_u^l - \int_{\Gamma_h[\mathbf{x}^*]} |\nabla_{\Gamma_h[\mathbf{x}^*]} \hat{I}_h^* n|^2 \hat{I}_h^* u \cdot \hat{\chi}_u \right|$$

$$\begin{aligned}
&\leq \left| \int_{\Gamma} |\nabla_{\Gamma} n|^2 u \cdot \hat{\chi}_u^l - \int_{\Gamma_h[\mathbf{x}^*]} |\nabla_{\Gamma_h[\mathbf{x}^*]} n^{-l}|^2 u^{-l} \cdot \hat{\chi}_u \right| \\
&\quad + \left| \int_{\Gamma_h[\mathbf{x}^*]} |\nabla_{\Gamma_h[\mathbf{x}^*]} n^{-l}|^2 u^{-l} \cdot \hat{\chi}_u - \int_{\Gamma_h[\mathbf{x}^*]} |\nabla_{\Gamma_h[\mathbf{x}^*]} \hat{I}_h^* n|^2 \hat{I}_h^* u \cdot \hat{\chi}_u \right| \\
&=: F_1 + F_2.
\end{aligned} \tag{D.7}$$

To estimate  $F_1$ , we denote

$$x^\theta := (1 - \theta)x^l + \theta x \text{ for } x \in \Gamma_h[\mathbf{x}^*] \quad \text{and} \quad \Gamma^\theta := (1 - \theta)\Gamma + \theta\Gamma_h[\mathbf{x}^*] = \{x^\theta : x \in \Gamma_h[\mathbf{x}^*]\}.$$

Let  $u^\theta, n^\theta, e^\theta$  and  $\chi_u^\theta$  be functions on  $\Gamma^\theta$  defined by

$$u^\theta(x^\theta) = u(x^l), \quad n^\theta(x^\theta) = n(x^l), \quad e^\theta(x^\theta) = x - x^l \quad \text{and} \quad \chi_u^\theta(x^\theta) = \hat{\chi}_u(x) \quad \text{for } x \in \Gamma_h[\mathbf{x}^*].$$

Then

$$\partial_\theta^\bullet u^\theta = \partial_\theta^\bullet \chi_u^\theta = 0 \quad \text{and} \quad \partial_\theta^\bullet n^\theta = 0,$$

and

$$\begin{aligned}
F_1 &= \left| \int_0^1 \frac{d}{d\theta} \int_{\Gamma^\theta} |\nabla_{\Gamma^\theta} n^\theta|^2 u^\theta \cdot \chi_u^\theta d\theta \right| \\
&= \left| \int_0^1 \int_{\Gamma^\theta} |\nabla_{\Gamma^\theta} n^\theta|^2 (\nabla_{\Gamma^\theta} \cdot e^\theta) u^\theta \cdot \chi_u^\theta d\theta \right. \\
&\quad \left. + \int_0^1 \int_{\Gamma^\theta} 2\nabla_{\Gamma^\theta} n^\theta \cdot (\nabla_{\Gamma^\theta} \partial_\theta^\bullet n^\theta - \nabla_{\Gamma^\theta} e^\theta \nabla_{\Gamma^\theta} n^\theta + \nu^\theta (\nu^\theta)^\top (\nabla_{\Gamma^\theta} e^\theta)^\top \nabla_{\Gamma^\theta} n^\theta) u^\theta \cdot \chi_u^\theta d\theta \right| \\
&=: |F_{11} + F_{12}|,
\end{aligned}$$

where  $\nu^\theta$  is the unit normal vector on  $\Gamma^\theta$ . Notice that the following identity holds:

$$\begin{aligned}
F_{11} &= \int_0^1 \int_{\Gamma^\theta} |\nabla_{\Gamma^\theta} n^\theta|^2 (\nabla_{\Gamma^\theta} \cdot e^\theta) u^\theta \cdot \chi_u^\theta d\theta \\
&= \int_0^1 \left[ \int_{\Gamma^\theta} |\nabla_{\Gamma^\theta} n^\theta|^2 (\nabla_{\Gamma^\theta} \cdot e^\theta) u^\theta \cdot \chi_u^\theta - \int_{\Gamma} |\nabla_{\Gamma} n|^2 (\nabla_{\Gamma} \cdot e^l) u \cdot \hat{\chi}_u^l \right] d\theta \\
&\quad + \int_{\Gamma} |\nabla_{\Gamma} n|^2 (\nabla_{\Gamma} \cdot e^l) u \cdot \hat{\chi}_u^l =: F_{111} + F_{112}.
\end{aligned}$$

By employing the intermediate surface again and referencing the inequality  $\|\nabla_{\Gamma^\theta} \cdot e^\theta\|_{L^\infty(\Gamma^\theta)} \leq Ch^k$  induced from Lagrange interpolation error estimates, we obtain

$$|F_{111}| \leq C \|\nabla_{\Gamma^\theta} \cdot e^\theta\|_{L^\infty(\Gamma^\theta)}^2 \|\hat{\chi}_u^l\|_{L^2(\Gamma)} \leq Ch^{2k} \|\hat{\chi}_u\|_{H^1(\Gamma_h[\mathbf{x}^*])}.$$

By integration by parts and referencing the inequality  $\|e^\theta\|_{L^\infty(\Gamma^\theta)} \leq Ch^{k+1}$  from Lagrange interpolation error estimates, we have

$$|F_{112}| \leq C \|e^\theta\|_{L^\infty(\Gamma^\theta)} \|\nabla_{\Gamma} \hat{\chi}_u^l\|_{L^2(\Gamma)} \leq Ch^{k+1} \|\hat{\chi}_u\|_{H^1(\Gamma_h[\mathbf{x}^*])}.$$

The term  $F_{12}$  can be estimated in the same way as  $F_{11}$  (split it into two parts, separately, and using integration by parts for the second part). Hence,

$$F_1 \leq Ch^{k+1} \|\hat{\chi}_u\|_{H^1(\Gamma_h[\mathbf{x}^*])}. \tag{D.8}$$

By using the triangle inequality, we have

$$\begin{aligned}
F_2 &= \left| \int_{\Gamma_h[\mathbf{x}^*]} |\nabla_{\Gamma_h[\mathbf{x}^*]} n^{-l}|^2 u^{-l} \cdot \hat{\chi}_u - \int_{\Gamma_h[\mathbf{x}^*]} |\nabla_{\Gamma_h[\mathbf{x}^*]} \hat{I}_h^* n|^2 \hat{I}_h^* u \cdot \hat{\chi}_u \right| \\
&\leq \left| \int_{\Gamma_h[\mathbf{x}^*]} |\nabla_{\Gamma_h[\mathbf{x}^*]} n^{-l}|^2 (u^{-l} - \hat{I}_h^* u) \cdot \hat{\chi}_u \right|
\end{aligned}$$

$$\begin{aligned}
& + \left| \int_{\Gamma_h[\mathbf{x}^*]} (|\nabla_{\Gamma_h[\mathbf{x}^*]} n^{-l}|^2 - |\nabla_{\Gamma_h[\mathbf{x}^*]} \hat{I}_h^* n|^2) \hat{I}_h^* u \cdot \hat{\chi}_u \right| \\
& =: F_{21} + F_{22}.
\end{aligned}$$

By using error estimates of Lagrange interpolation, we have

$$F_{21} \leq C \|u^{-l} - \hat{I}_h^* u\|_{L^2(\Gamma_h[\mathbf{x}^*])} \|\hat{\chi}_u\|_{L^2(\Gamma_h[\mathbf{x}^*])} \leq Ch^{k+1} \|\hat{\chi}_u\|_{L^2(\Gamma_h[\mathbf{x}^*])} \leq Ch^{k+1} \|\hat{\chi}_u\|_{L^2(\Gamma_h[\mathbf{x}^*])}.$$

Let  $\eta = n^{-l} - \hat{I}_h^* n$ , then

$$\begin{aligned}
F_{22} & = \left| \int_{\Gamma_h[\mathbf{x}^*]} (\nabla_{\Gamma_h[\mathbf{x}^*]} n^{-l} + \nabla_{\Gamma_h[\mathbf{x}^*]} \hat{I}_h^* n) \cdot \nabla_{\Gamma_h[\mathbf{x}^*]} \eta (\hat{I}_h^* u \cdot \hat{\chi}_u) \right| \\
& \leq \left| \int_{\Gamma_h[\mathbf{x}^*]} (\nabla_{\Gamma_h[\mathbf{x}^*]} n^{-l} + \nabla_{\Gamma_h[\mathbf{x}^*]} \hat{I}_h^* n - 2(\nabla_{\Gamma} n)^{-l}) \cdot \nabla_{\Gamma_h[\mathbf{x}^*]} \eta (\hat{I}_h^* u \cdot \hat{\chi}_u) \right| \\
& + \left| \int_{\Gamma_h[\mathbf{x}^*]} 2(\nabla_{\Gamma} n)^{-l} \cdot \nabla_{\Gamma_h[\mathbf{x}^*]} \eta (\hat{I}_h^* u \cdot \hat{\chi}_u) - 2 \int_{\Gamma} \nabla_{\Gamma} n \cdot \nabla_{\Gamma} \eta^l (\hat{I}_h^* u \cdot \hat{\chi}_u)^l \right| \\
& + \left| 2 \int_{\Gamma} \nabla_{\Gamma} n \cdot \nabla_{\Gamma} \eta^l (\hat{I}_h^* u \cdot \hat{\chi}_u)^l \right| =: F_{221} + F_{222} + F_{223}.
\end{aligned}$$

Since the following inequality holds (Lagrange interpolation error estimates)

$$\|\eta\|_{L^2(\Gamma_h[\mathbf{x}^*])} + h \|\nabla_{\Gamma_h[\mathbf{x}^*]} \eta\|_{L^2(\Gamma_h[\mathbf{x}^*])} \leq Ch^{k+1}, \quad (\text{D.9})$$

along with error estimates from perturbation of surfaces, we have

$$F_{221} + F_{222} \leq Ch^{2k} \|\hat{\chi}_u\|_{L^2(\Gamma_h[\mathbf{x}^*])}.$$

By utilizing integration by parts to remove the gradient  $\nabla_{\Gamma}$  from  $\eta^l$ , along with inequality (D.9), we have

$$F_{223} \leq C \|\eta^l\|_{L^2(\Gamma)} \|\hat{\chi}_u\|_{H^1(\Gamma_h[\mathbf{x}^*])} \leq Ch^{k+1} \|\hat{\chi}_u\|_{H^1(\Gamma_h[\mathbf{x}^*])}.$$

Then we obtain

$$F_2 \leq F_{21} + F_{221} + F_{222} + F_{223} \leq Ch^{k+1} \|\hat{\chi}_u\|_{H^1(\Gamma_h[\mathbf{x}^*])}. \quad (\text{D.10})$$

Substituting the estimates of  $F_1$  and  $F_2$  into (D.7), we obtain the following result for the first term on the right-hand side of (D.6):

$$\begin{aligned}
& \left| \int_{\Gamma} |\nabla_{\Gamma} n|^2 u \cdot \hat{\chi}_u^l - \int_{\Gamma_h[\mathbf{x}^*]} |\nabla_{\Gamma_h[\mathbf{x}^*]} \hat{I}_h^* n|^2 \hat{I}_h^* u \cdot \hat{\chi}_u \right| \leq Ch^{k+1} \|\hat{\chi}_u\|_{H^1(\Gamma_h[\mathbf{x}^*])} \\
& \leq Ch^{k+1} \|\hat{\chi}_u\|_{H^1(\Gamma_h[\mathbf{y}^*])}, \quad (\text{D.11})
\end{aligned}$$

where norm equivalence of  $W^{1,p}$  norm on  $\Gamma_h[\mathbf{x}^*]$  and  $\Gamma_h[\mathbf{y}^*]$  has been used in the last inequality.

The second term on the right-hand side of (D.6) can be estimated by employing the intermediate surface again between  $\Gamma_h[\mathbf{x}^*]$  and  $\Gamma_h[\mathbf{y}^*]$ . Recall that  $\mathbf{y}^\theta = (1 - \theta)\mathbf{x}^* + \theta\mathbf{y}^*$ . Let  $\chi_u^\theta$ ,  $u_h^\theta$ ,  $n_h^\theta$  and  $e_y^\theta$  be the finite element functions on  $\Gamma_h[\mathbf{y}^\theta]$  with nodal vectors  $\chi_u$ ,  $(1 - \theta)\mathbf{I}_h u + \theta\mathbf{u}^*$ ,  $(1 - \theta)\mathbf{I}_h n + \theta\mathbf{n}^*$  and  $\mathbf{e}_y := \mathbf{y} - \mathbf{x}$ , respectively. Then

$$\partial_\theta^\bullet u_h^\theta = e_u^\theta, \quad \partial_\theta^\bullet n_h^\theta = e_n^\theta \quad \text{and} \quad \partial_\theta^\bullet \chi_u^\theta = 0,$$

where  $e_u^\theta$  and  $e_n^\theta$  are the finite element functions on  $\Gamma_h[\mathbf{y}^\theta]$  with nodal vectors  $\mathbf{u}^* - \mathbf{I}_h u$  and  $\mathbf{n}^* - \mathbf{I}_h n$ , respectively. The surface  $\Gamma_h[\mathbf{y}^\theta]$  is moving with velocity  $e_y^\theta$  as  $\theta$  changes from 0 to 1. Then, Lemmas 3.3 and 3.4 imply that

$$\|e_y^\theta\|_{L^2(\Gamma_h[\mathbf{y}^\theta])} + h \|e_y^\theta\|_{H^1(\Gamma_h[\mathbf{y}^\theta])} \leq Ch^{k+1}, \quad (\text{D.12})$$

$$\|e_n^\theta\|_{L^2(\Gamma_h[\mathbf{y}^\theta])} + h \|e_n^\theta\|_{H^1(\Gamma_h[\mathbf{y}^\theta])} \leq Ch^{k+1}, \quad (\text{D.13})$$

$$\|e_u^\theta\|_{L^2(\Gamma_h[\mathbf{y}^\theta])} + h\|e_u^\theta\|_{H^1(\Gamma_h[\mathbf{y}^\theta])} \leq Ch^{k+1}. \quad (\text{D.14})$$

Then the second term on the right-hand side of (D.6) can be estimated as follows:

$$\begin{aligned} & \left| \int_{\Gamma_h[\mathbf{x}^*]} |\nabla_{\Gamma_h[\mathbf{x}^*]} \hat{I}_h^* n|^2 \hat{I}_h^* u \cdot \hat{\chi}_u - \int_{\Gamma_h[\mathbf{y}^*]} |\nabla_{\Gamma_h[\mathbf{y}^*]} n_h^*|^2 u_h^* \cdot \chi_u \right| \\ &= \left| \int_0^1 \frac{d}{d\theta} \int_{\Gamma_h[\mathbf{y}^\theta]} |\nabla_{\Gamma_h[\mathbf{y}^\theta]} n_h^\theta|^2 u_h^\theta \cdot \chi_u^\theta d\theta \right| \\ &= \left| \int_0^1 \int_{\Gamma_h[\mathbf{y}^\theta]} (|\nabla_{\Gamma_h[\mathbf{y}^\theta]} n_h^\theta|^2 e_u^\theta \cdot \chi_u^\theta + |\nabla_{\Gamma_h[\mathbf{y}^\theta]} n_h^\theta|^2 (u_h^\theta \cdot \chi_u^\theta) \nabla_{\Gamma_h[\mathbf{y}^\theta]} \cdot e_y^\theta) d\theta \right. \\ & \quad \left. + \int_0^1 \int_{\Gamma_h[\mathbf{y}^\theta]} 2 \nabla_{\Gamma_h[\mathbf{y}^\theta]} n_h^\theta \cdot (\nabla_{\Gamma_h[\mathbf{y}^\theta]} \partial_\theta^\bullet n_h^\theta - \nabla_{\Gamma_h[\mathbf{y}^\theta]} e_y^\theta \nabla_{\Gamma_h[\mathbf{y}^\theta]} n_h^\theta + \nu^\theta (\nu^\theta)^\top (\nabla_{\Gamma_h[\mathbf{y}^\theta]} e_y^\theta)^\top \nabla_{\Gamma_h[\mathbf{y}^\theta]} n_h^\theta) (u_h^\theta \cdot \chi_u^\theta) d\theta \right| \\ &\leq \left| \int_0^1 \int_{\Gamma_h[\mathbf{y}^\theta]} |\nabla_{\Gamma_h[\mathbf{y}^\theta]} n_h^\theta|^2 e_u^\theta \cdot \chi_u^\theta d\theta \right| \\ & \quad + \left| \int_0^1 \int_{\Gamma_h[\mathbf{y}^\theta]} |\nabla_{\Gamma_h[\mathbf{y}^\theta]} n_h^\theta|^2 (u_h^\theta \cdot \chi_u^\theta) \nabla_{\Gamma_h[\mathbf{y}^\theta]} \cdot e_y^\theta d\theta \right| \\ & \quad + \left| \int_0^1 \int_{\Gamma_h[\mathbf{y}^\theta]} 2 \nabla_{\Gamma_h[\mathbf{y}^\theta]} n_h^\theta \cdot \nabla_{\Gamma_h[\mathbf{y}^\theta]} e_y^\theta (u_h^\theta \cdot \chi_u^\theta) d\theta \right| \\ & \quad + \left| \int_0^1 \int_{\Gamma_h[\mathbf{y}^\theta]} 2 (\nabla_{\Gamma_h[\mathbf{y}^\theta]} n_h^\theta \cdot \nabla_{\Gamma_h[\mathbf{y}^\theta]} e_y^\theta \nabla_{\Gamma_h[\mathbf{y}^\theta]} n_h^\theta) (u_h^\theta \cdot \chi_u^\theta) d\theta \right| \\ & \quad + \left| \int_0^1 \int_{\Gamma_h[\mathbf{y}^\theta]} 2 (\nabla_{\Gamma_h[\mathbf{y}^\theta]} n_h^\theta \cdot \nu_h^\theta (\nu_h^\theta)^\top (\nabla_{\Gamma_h[\mathbf{y}^\theta]} e_y^\theta)^\top (\nabla_{\Gamma_h[\mathbf{y}^\theta]} n_h^\theta)) (u_h^\theta \cdot \chi_u^\theta) d\theta \right| \\ &=: \sum_{j=1}^5 E_j, \end{aligned} \quad (\text{D.15})$$

where  $\nu^\theta$  is the unit normal vector on  $\Gamma_h[\mathbf{y}^\theta]$ . By using inequality (D.14), we have

$$\begin{aligned} E_1 &= \left| \int_0^1 \int_{\Gamma_h[\mathbf{y}^\theta]} |\nabla_{\Gamma_h[\mathbf{y}^\theta]} n_h^\theta|^2 (e_u^\theta \cdot \chi_u^\theta) d\theta \right| \\ &\leq C \|e_u^\theta\|_{L^2(\Gamma_h[\mathbf{y}^\theta])} \|\nabla_{\Gamma_h[\mathbf{y}^\theta]} n_h^\theta\|_{L^6(\Gamma_h[\mathbf{y}^\theta])} \|\chi_u^\theta\|_{L^6(\Gamma_h[\mathbf{y}^\theta])} \\ &\leq Ch^{k+1} \|\chi_u\|_{H^1(\Gamma_h[\mathbf{y}^*])}, \end{aligned}$$

where norm equivalence of  $L^p$  and  $W^{1,p}$  norms on  $\Gamma_h[\mathbf{y}^\theta]$  and  $\Gamma_h[\mathbf{y}^*]$  are used in the last inequality. Observe that  $E_2$  can be rewritten as follows:

$$\begin{aligned} E_2 &= \left| \int_0^1 \int_{\Gamma_h[\mathbf{y}^\theta]} |\nabla_{\Gamma_h[\mathbf{y}^\theta]} n_h^\theta|^2 (u_h^\theta \cdot \chi_u^\theta) \nabla_{\Gamma_h[\mathbf{y}^\theta]} \cdot e_y^\theta d\theta \right| \\ &\leq \left| \int_0^1 \left[ \int_{\Gamma_h[\mathbf{y}^\theta]} |\nabla_{\Gamma_h[\mathbf{y}^\theta]} n_h^\theta|^2 (u_h^\theta \cdot \chi_u^\theta) \nabla_{\Gamma_h[\mathbf{y}^\theta]} \cdot e_y^\theta - \int_{\Gamma_h[\mathbf{x}^*]} |\nabla_{\Gamma_h[\mathbf{x}^*]} \hat{I}_h^* n|^2 (\hat{I}_h^* u \cdot \hat{\chi}_u) \nabla_{\Gamma_h[\mathbf{x}^*]} \cdot e_y^0 \right] d\theta \right| \\ & \quad + \left| \int_{\Gamma_h[\mathbf{x}^*]} |\nabla_{\Gamma_h[\mathbf{x}^*]} \hat{I}_h^* n|^2 (\hat{I}_h^* u \cdot \hat{\chi}_u) \nabla_{\Gamma_h[\mathbf{x}^*]} \cdot e_y^0 - \int_{\Gamma} |\nabla_{\Gamma} n|^2 (u \cdot \hat{\chi}_u^l) \nabla_{\Gamma} \cdot (e_y^0)^l \right| \\ & \quad + \left| \int_{\Gamma} |\nabla_{\Gamma} n|^2 (u \cdot \hat{\chi}_u^l) \nabla_{\Gamma} \cdot (e_y^0)^l \right| =: E_{21} + E_{22} + E_{23}. \end{aligned} \quad (\text{D.16})$$

By employing the intermediate surface between  $\Gamma_h[\mathbf{y}^\theta]$  and  $\Gamma_h[\mathbf{x}^*]$  again and referencing the inequalities (D.12), (D.13) and (D.14), we obtain

$$E_{21} \leq Ch^{2k} \|\chi_u\|_{L^2(\Gamma_h[\mathbf{y}^*])}.$$

By employing the intermediate surface between  $\Gamma_h[\mathbf{x}^*]$  and  $\Gamma$  throughout the lift map and utilizing the inequality  $\|\nabla_{\Gamma_h[\mathbf{x}^*]} \cdot (x - x^l)\|_{L^p(\Gamma_h[\mathbf{x}^*])} \leq Ch^k$  induced from Lagrange interpolation

error estimates, we obtain

$$E_{22} \leq Ch^{2k} \|\chi_u\|_{L^2(\Gamma_h[\mathbf{y}^*])}.$$

By using integration by parts along with inequality (D.12), we derive

$$E_{23} \leq C \|e_y^0\|_{L^2(\Gamma_h[\mathbf{x}^*])} \|\chi_u\|_{H^1(\Gamma_h[\mathbf{y}^*])} \leq Ch^{k+1} \|\chi_u\|_{H^1(\Gamma_h[\mathbf{y}^*])}.$$

Substituting the estimates of  $E_{2j}$  for  $j = 1, 2, 3$  into (D.16), we obtain

$$E_2 \leq Ch^{k+1} \|\chi_u\|_{H^1(\Gamma_h[\mathbf{y}^*])}.$$

The terms  $E_3$ ,  $E_4$  and  $E_5$  can be estimated similarly as  $E_2$ , then we have

$$E_3 + E_4 + E_5 \leq Ch^{k+1} \|\chi_u\|_{H^1(\Gamma_h[\mathbf{y}^*])}.$$

Substituting the estimates of  $E_j$  for  $j = 1, 2, 3, 4, 5$  into (D.15), we obtain the following result for the second term on the right-hand side of (D.6):

$$\left| \int_{\Gamma_h[\mathbf{x}^*]} |\nabla_{\Gamma_h[\mathbf{x}^*]} \hat{I}_h^* n|^2 \hat{I}_h^* u \cdot \hat{\chi}_u - \int_{\Gamma_h[\mathbf{y}^*]} |\nabla_{\Gamma_h[\mathbf{y}^*]} n_h^*|^2 u_h^* \cdot \chi_u \right| \leq Ch^{k+1} \|\chi_u\|_{H^1(\Gamma_h[\mathbf{y}^*])}. \quad (\text{D.17})$$

Then substituting inequalities (D.11) and (D.17) into inequality (D.6), we obtain

$$\left| \int_{\Gamma_h[\mathbf{y}^*]} d_{u,3}^* \cdot \chi_u \right| \leq Ch^{k+1} \|\chi_u\|_{H^1(\Gamma_h[\mathbf{y}^*])},$$

which proves Lemma D.2.  $\square$

Substituting results from Lemma D.1 and Lemma D.2 into (D.3), we obtain

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

which proves Lemma 3.6.