Optimal convergence of the arbitrary Lagrangian–Eulerian interface tracking method for two-phase Navier–Stokes flow without surface tension

BUYANG LI†

Department of Applied Mathematics, The Hong Kong Polytechnic University, Hong Kong.
AND

SHU MA* AND WEIFENG QIU‡

Department of Mathematics, City University of Hong Kong, Kowloon, Hong Kong.

Optimal-order convergence in the H^1 norm is proved for an arbitrary Lagrangian–Eulerian interface tracking finite element method for the sharp interface model of two-phase Navier–Stokes flow without surface tension, using high-order curved evolving mesh. In this method, the interfacial mesh points move with the fluid's velocity to track the sharp interface between two phases of the fluid, and the interior mesh points move according to a harmonic extension of the interface velocity. The error of the semidiscrete arbitrary Lagrangian–Eulerian interface tracking finite element method is shown to be $O(h^k)$ in the $L^\infty(0, T; H^1(\Omega))$ norm for the Taylor-Hood finite elements of degree $k \geqslant 2$. This high-order convergence is achieved by utilizing the piecewise smoothness of the solution on each subdomain occupied by one phase of the fluid, relying on a low global regularity on the entire moving domain. Numerical experiments illustrate and complement the theoretical results.

Keywords: Two-phase Navier–Stokes flow, arbitrary Lagrangian–Eulerian, finite element method, convergence, error estimates

1. Introduction

Immiscible fluids mixture separated by a moving interface appears widely in nature and engineering applications. Current models for two-phase or multiphase flow can generally be categorized into two main types: diffuse interface models and sharp interface models. Diffuse interface models treat the interface as an extremely thin transition zone that separates the two fluids. In these models, physical quantities like density and viscosity change rapidly but smoothly within this transition zone. On the other hand, sharp interface models consider the interface as a surface without any thickness. In such models, physical quantities exhibit discontinuities across this surface.

We consider the sharp interface model of two-phase Navier–Stokes (NS) flow [25] without surface tension in a bounded domain $\Omega(t) = \Omega_+(t) \cup \Omega_-(t) \cup \Gamma(t)$ in $\mathbb{R}^d$, with $d \in \{2, 3\}$, where $\Gamma(t) = \overline{\Omega}_+(t) \cap \overline{\Omega}_-(t)$ is a sharp interface which separates the two fluids occupying two subdomains $\Omega_+(t)$ and $\Omega_-(t)$, respectively, see Figure 1. The fluid velocity u and pressure p are governed by NS equations in each subdomain $\Omega_\pm(t)$

†Email: buyang.li@polyu.edu.hk
‡Email: shuma2@cityu.edu.hk and weifeqiu@cityu.edu.hk

and satisfy certain jump conditions on the interface $\Gamma(t)$, i.e.,

$$\begin{cases} \rho_\pm(\partial_t u + u \cdot \nabla u) - \nabla \cdot \sigma = f & \text{in } \bigcup_{t \in [0,T]} \Omega_\pm(t) \times \{t\}, \\[2mm] \nabla \cdot u = 0 & \text{in } \bigcup_{t \in [0,T]} \Omega_\pm(t) \times \{t\}, \\[2mm] [u]_-^+ = 0 \quad \text{and} \quad [\sigma v]_-^+ = 0 & \text{on } \bigcup_{t \in [0,T]} \Gamma(t) \times \{t\}, \\[2mm] u = 0 & \text{in } \partial\Omega \times (0,T], \\[2mm] u = u^0 & \text{on } \Omega \times \{0\}, \end{cases} \tag{1.1}$$

where $\rho_\pm$ are the densities of the fluids in $\Omega_\pm(t)$, v is the unit normal vector on $\Gamma(t)$ pointing to $\Omega_+(t)$; $\sigma = (2\mu_\pm \mathbb{D}(u) - p\mathbb{I})$ is the stress tensor, in which $\mathbb{D}(u) = \frac{1}{2}(\nabla u + (\nabla u)^\top)$ and $\mathbb{I}$ denote the deformation matrix and identity matrix, respectively; $\mu_\pm$ are the viscosities of the two fluids in $\Omega_\pm(t)$, respectively; $[u]_-^+ = u_+ - u_-$ and $[\sigma v]_-^+ = \sigma_+ v - \sigma_- v$ denote the jumps of the quantities u and σv across the interface, respectively. The gravitational force is given by $f = (0, -g)^\top$ in two dimensions and $f = (0,0,-g)^\top$ in three dimensions with g a constant (the gravitational acceleration). The interface $\Gamma(t)$ moves along with the fluid
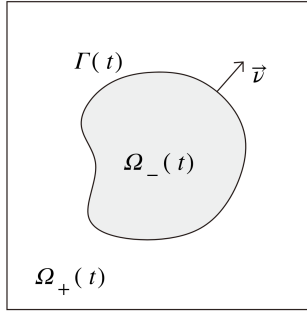

FIG. 1. Domain Ω in the case $d = 2$.

and therefore also has velocity u. If Γ^0 is the initial interface, through a flow map $X : \Gamma^0 \times [0,T] \to \mathbb{R}^d$ we can describe the evolution of interface $\Gamma(t) = \{X(y,t) : y \in \Gamma^0\}$ by the following equation:

$$\begin{cases} \partial_t X(\xi,t) = u(X(\xi,t),t) & \text{for } (\xi,t) \in \Gamma^0 \times (0,T], \\ X(\xi,0) = \xi & \text{for } \xi \in \Gamma^0. \end{cases} \tag{1.2}$$

We do not take surface tension into account on the interface $\Gamma(t)$, which differs from models that include surface tension, where the appearance of curvature leads to a coupling between the geometric properties of the interface and the fluid flow in the bulk. The fluid dynamics in (1.1)–(1.2) involve the movement of the interface $\Gamma(t)$ and two subdomains $\Omega_\pm(t)$ over time, while the overall domain $\Omega = \Omega(t)$ remains unchanged in shape.

Numerical methods for (1.1) can generally be classified into three categories, i.e., the Eulerian approach, the Lagrangian approach, and the arbitrary Lagrangian–Eulerian approach (ALE). The arbitrary Lagrangian–Eulerian approach serves as a bridge between the Lagrangian and Eulerian approaches, by enabling the frame to move with an "arbitrary" bulk velocity that fits the interface motion. At the discrete level, the mesh points can be displaced independently of the flow. Utilizing the ALE technique allows for flexible movement of the inner domain mesh, while the mesh on the interface can move alongside materials to precisely track the interfaces of a multi-material system.

In an early investigation of ALE methods, the stability analysis of the ALE finite element method (FEM) was firstly conducted by Formaggia and Nobile [20] for a linear parabolic equation. Subsequently, Gastaldi [23] developed a priori error estimates in the L^2 norm for ALE-FEM when solving parabolic equations. Many

studies on the numerical analysis of ALE-FEMs for parabolic moving boundary/interface problems can be found in works such as [4, 10, 19, 31, 32, 37]. Optimal-order error bounds of $O(h^{k+1})$ in the $L^\infty(0,T;L^2)$ norm of ALE semidiscrete FEM with curved triangles/simplices for parabolic moving boundary problems were established by Elliott and Ranner [16]. We also refer to [15] for an ALE method with harmonically evolving mesh. Optimal-order H^1 convergence of the ALE-FEM for PDEs coupling boundary evolution arising from shape optimization problems was proved in [24]. These results were established for high-order curved evolving mesh. Optimal convergence of $O(h^{k+1})$ in the $L^\infty(0,T;L^2)$ norm, with flat evolving simplices in the interior and curved simplices exclusively on the boundary, was prove in [35] for the ALE semidiscrete FEM utilizing a standard iso-parametric element of degree k in [34]. The ALE methods for PDEs in bulk domains [24] are also closely related to the evolving FEMs for PDEs on evolving surfaces, for which optimal-order convergence in the L^2 and H^1 norms were shown in [14, 17, 30]. The above-mentioned research efforts have focused on diffusion equations.

The numerical analysis of ALE methods for the Stokes and Navier–Stokes equations has also noteworthy results but remained suboptimal. Legendre and Takahashi [33] established an L^2 error estimate of $O(\tau + h^{\frac{1}{2}})$ for the method of characteristics combined with a $P_{1b} - P_1$ finite element approximation to the ALE formulation of 2D Navier–Stokes equations. In [39], by using the Taylor-Hood P_2-P_1 element, the error estimates of $O(h^2|\log h|)$ and $O(\tau + h^2 + h^2/\tau)$ were proved for the semi-discrete and fully discrete ALE-FEM for the Stokes equations in a time-dependent domain, respectively. The error of ALE finite element solutions to the Stokes equations on a time-varying domain, with BDF-r in time (for $1 \leqslant r \leqslant 5$) and the Taylor-Hood P_k–P_{k-1} elements in space (with degree $k \geqslant 2$), was proved to be $O(\tau^r + h^k)$ in the L^2 norm in [36]. Optimal-order error bounds of $O(h^{k+1})$ and $O(\tau^2 + h^{k+1})$ of the semidiscrete and fully discrete ALE-FEMs in the L^2 norm for the Stokes equations were proved recently in [38].

Many efforts have also been made in developing efficient ALE-FEMs for two-phase NS flow problems, primarily for models with surface tension. For example, Anjos et al. [3] introduced a 3D ALE-FEM scheme with adaptive mesh to achieve good control of the mesh quality by using repeated interpolations of the velocity. Barrett et al. [7] also considered surface tension and designed a stable ALE-FEM in which the mesh points used to describe the interface are not mesh points of the underlying bulk finite element mesh. An implicit tangential velocity of the interface nodes is introduced in [7] to acquire good mesh quality of the interface. Barrett, Garcke and Nürnberg introduced a novel weak formulation [9], which is referred to as the BGN formulation from now on. The BGN formulation allows for tangential degrees of freedom to improve the mesh quality. This formulation was employed for two-phase Stokes flow [6] and two-phase NS flow [8], both with surface tension and with unfitted finite element approximations, which leads to an unconditional stability bound. Moreover, the applications of BGN formulation to the two-phase NS flow with surface tension with moving fitted finite element methods were also investigated in [1] in both Eulerian and ALE approaches, but no stability bound was available. Fu [21] proposed an ALE-HDG FEM for two-phase NS flow with surface tension with an implicit tangential velocity to acquire good mesh quality of the interface, by using HDG to produce an exactly divergence-free velocity approximation. Recently, Duan et al. [12] proposed an energy-diminishing ALE scheme with surface tension by using Stokes extension of the velocity from the interface to the bulk domain, in order to produce an exactly divergence-free mesh velocity. Energy stability of the ALE-FEMs with surface tension was discussed in [6–8, 12, 21]. We also refer to [22] for the structure-preserving ALE-FEMs for the two-phase NS flow with surface tension in both the conservative and non-conservative forms. The methods introduced in [22] were shown to satisfy unconditional stability and exact volume preservation on the fully discrete level. The numerical investigation of convergence behavior (without error analysis) with surface tension, can be found in [12, 21]. As far as we know, there is no error analysis of ALE-FEMs for the sharp interface model of two-phase NS flow problems with surface tension. In this work, we focus on a simplified model of two-phase NS flow without surface tension. Two major difficulties are discussed below.

First, the error analysis of ALE-FEMs for parabolic and Stokes/NS equations in the above-mentioned articles all require the triangulation of the domain to fit the moving boundary or interface exactly, i.e., the triangulated domain $\Omega_h(t)$ for computation that interpolates the exact domain $\Omega(t)$. This is the case when the

boundary or interface of domain $\Omega(t)$ is given, determined either by a given flow map or by a given velocity field. However, this is the not the case in two-phase NS flow where the velocity of the interface is unknown (determined by the solution of the PDE problem). Thus the location of the exact interface is unknown and needs to be approximated based on the numerical solution of the two-phase NS flow. As a result, the numerically computed interface at $t > 0$ generally has a distance from (therefore is not an interpolation of) the exact interface, and the triangulated subdomains $\Omega_{h,\pm}(t)$ are not interpolations of the exact subdomains $\Omega_{\pm}(t)$ at $t > 0$. Consequently, the error analysis of ALE-FEMs in the pre-existing articles for parabolic and Stokes/NS equations with a given moving boundary/interface cannot be applied to the two-phase NS flow without surface tension.

Second, the error analysis of ALE-FEMs for parabolic and Stokes/NS equations in the above-mentioned articles all require the solution to be sufficiently smooth in the whole domain, while the solution of the two-phase NS equation in (1.1) is generally not smooth in the whole domain due to the existence of the free interface which separates the two phases of the fluid, e.g., ∇u is generally discontinuous across the interface $\Gamma(t)$. Since the numerical interface and the exact interface differ from each other by a certain distance at $t > 0$, the triangles of the mesh (which fit the numerical interface) generally do not fit the exact interface $\Gamma(t)$. Thus the exact solution u may be nonsmooth inside a triangle of the mesh. This situation renders it difficult to prove high-order convergence using standard interpolation onto the mesh.

The aim of this article is to establish an optimal-order error bound in the $L^{\infty}(0,T;H^1(\Omega))$ norm for the semidiscrete ALE-FEM for the two-phase NS flow problem without surface tension in (1.1) using high-order curved evolving mesh, where the mesh points on the numerical interface move with the fluid's velocity and is harmonically extended to the bulk subdomains. We address the two major difficulties mentioned above by using the matrix-vector formulations of the ALE-FEM and the error equations. Such techniques were originally introduced in [30] for analyzing solution-driven surface evolutions. It was later developed for analyzing the convergence of evolving FEMs for surface evolution under geometric flows [5, 26, 28, 29] and domain evolution under shape gradient flows [24]. We adapt such matrix-vector formulation techniques here to address the mismatch between the numerical interface and the exact interface, using an intermediate auxiliary triangulated domain, as well as the missing of global regularity of the solution by using lift operations which map the numerical subdomains to the exact subdomains. This approach only requires using piecewise smoothness of the solution u on the subdomains $\Omega_{\pm}(t)$ instead of its global smoothness of u on the whole domain $\Omega(t)$. For a smoothly evolving sharp interface $\Gamma(t)$ without topological changes, and a solution $(u,p) \in C([0,T];H^1(\Omega)^d) \times L^2(0,T;L^2(\Omega))$ which is piecewise smooth in the subdomains $\Omega_{\pm}(t)$ for $t \in [0,T]$, we prove optimal-order error bound of $O(h^k)$ in the $C([0,T];H^1(\Omega)^d) \times L^2(0,T;L^2(\Omega))$ norm for the semi-discrete ALE-FEMs scheme with Taylor-Hood P_k–P_{k-1} elements of degree $k \geqslant 2$; see Theorem 4.1.

The rest of this article is organized as follows. Section 2 introduces the ALE reformulation of (1.1) and the ALE weak formulation on the evolving domain. In Section 3 we introduce the ALE weak formulation and ALE evolving FEM on the interpolated evolving domain and provide a consistency analysis. In Section 3 we formulate the consistency error estimates based on interpolated evolving mesh and matrix-vector formulation, which serve not only for implementation purposes but also play a key role in the subsequent convergence analysis. In Section 4 we prove optimal-order convergence of the ALE evolving FEM considered in this article. Numerical experiments are presented in Section 5 to illustrate the convergence of the ALE evolving FEM and its performance on a benchmark example.

2. The ALE evolving finite element method

2.1 *The ALE reformulation of PDE*

Let $w(\cdot,t) : \Omega(t) \to \mathbb{R}^d$ be the harmonic extension of fluid's velocity u from the interface $\Gamma(t)$ to the bulk subdomains $\Omega_{\pm}(t)$, i.e.,

$$-\Delta w = 0 \text{ in } \Omega_{\pm}(t) \quad \text{with} \quad w = u \text{ on } \Gamma(t) \quad \text{and} \quad w = 0 \text{ on } \partial\Omega(t). \tag{2.1}$$

Harmonic extension is a common method for generating mesh velocity w in the bulk subdomains; see [10, 15, 21]. Since $w = u$ on $\Gamma(t)$, it follows that w and u generate the same evolution of interface $\Gamma(t)$ and subdomains $\Omega_\pm(t)$.

Denote the initial domain by $\Omega^0 = \Omega(0)$ and the initial subdomains by $\Omega_\pm^0 = \Omega_\pm(0)$. Let $\phi : \Omega^0 \times [0,T] \to \mathbb{R}^d$ be the flow map generated by the velocity field w. Then ϕ maps $\Omega_\pm^0$ to $\Omega_\pm(t)$ and Γ^0 to $\Gamma(t)$, respectively, with

$$\Omega(t) = \{\phi(\xi,t) : \xi \in \Omega^0\} \quad \text{with} \quad \phi(\xi,0) = \xi \quad \text{for} \quad \xi \in \Omega^0,$$

and satisfies the following ordinary differential equation (ODE):

$$\partial_t \phi(\xi,t) = w(\phi(\xi,t),t) \quad \text{with} \quad \phi(\xi,0) = \xi \quad \text{for} \quad \xi \in \Omega^0. \tag{2.2}$$

We consider an ALE moving frame associated with the flow map ϕ, with frame velocity w. Then the original problem in (1.1) can be written as

$$\begin{cases} \rho_\pm \partial_t^\bullet u + \rho_\pm((u - w) \cdot \nabla u) - \nabla \cdot \sigma = 0 & \text{in } \Omega_\pm(t), \\ \nabla \cdot u = 0 & \text{in } \Omega_\pm(t), \end{cases} \tag{2.3}$$

where $\partial_t^\bullet u = \partial_t u + w \cdot \nabla u$ is the material derivative with respect to the frame velocity w. The subdomains $\Omega_\pm(t)$ evolve with velocity w.

We assume that problem (1.1), or equivalently (2.3), has a solution

$$(u, p) \in C([0,T]; H_0^1(\Omega)^d) \times C([0,T]; L_0^2(\Omega)),$$

which is piecewise smooth in subdomains $\Omega_\pm(t)$ for $t \in (0,T]$, where

$$H_0^1(\Omega)^d = \{v \in H^1(\Omega)^d : v = 0 \text{ on } \partial\Omega\} \quad \text{and} \quad L_0^2(\Omega) = \{q \in L^2(\Omega) \text{ and } \int_\Omega q\,\mathrm{d}x = 0\}.$$

By piecewise smooth we mean that $\sum_{j=0}^2 \|(\partial_t^\bullet)^j u\|_{H^{k+1}(\Omega_\pm(t))} + \sum_{j=0}^1 \|(\partial_t^\bullet)^j p\|_{H^k(\Omega_\pm(t))}$ is bounded uniformly with respect to $t \in [0,T]$.

We assume that the flow map $\phi(\cdot,t) \in C^k(\bar{\Omega}_\pm^0)$ is piecewise smooth in the subdomains $\Omega_\pm^0$. By pulling back to the initial domain Ω^0 after composition with $\phi(\cdot,t)$ and utilizing the regularity estimate of the elliptic equation [18, Theorem 5, Section 6.3] on the fixed domain Ω^0, the frame velocity w defined in (2.1) satisfies

$$\|w(t)\|_{H^k(\Omega_\pm(t))} \leqslant C\|u(t)\|_{H^k(\Omega_\pm(t))},$$

which is also piecewise smooth. Then, by using notations

$$(u, v) = \int_{\Omega_+(t)} uv\,\mathrm{d}x + \int_{\Omega_-(t)} uv\,\mathrm{d}x, \quad \rho = \begin{cases} \rho_+ & \text{in } \Omega_+(t), \\ \rho_- & \text{in } \Omega_-(t), \end{cases} \quad \text{and} \quad \mu = \begin{cases} \mu_+ & \text{in } \Omega_+(t), \\ \mu_- & \text{in } \Omega_-(t), \end{cases}$$

testing (2.3) by v and using the interface and boundary conditions in (1.1), we obtain the following weak reformulation.

ALE weak formulation: Find $(u,p) \in C([0,T]; H_0^1(\Omega)^d) \times C([0,T]; L_0^2(\Omega))$ with $\partial_t^\bullet u \in C([0,T]; L^2(\Omega)^d)$, $w \in C([0,T]; H^1(\Omega_\pm)^d)$, and $\Omega_\pm(t) = \phi(\Omega_\pm^0, t)$, such that

$$\begin{cases} (\rho\partial_t^\bullet u, v) + (\rho(u-w)\cdot\nabla u, v) + 2(\mu\mathbb{D}(u), \mathbb{D}(v)) - (p, \nabla\cdot v) = (f, v) & \forall v \in H_0^1(\Omega)^d \\ (\nabla\cdot u, q) = 0 & \forall q \in L_0^2(\Omega), \\ (\nabla w, \nabla\chi) = 0 \quad \forall\chi \in H_0^1(\Omega_\pm)^d \quad \text{and} \quad w = u \quad \text{on } \Gamma(t), \\ \partial_t\phi = w\circ\phi \quad \text{on } \Omega^0, \end{cases} \tag{2.4}$$

with initial conditions $\phi(0) = \mathrm{id}|_{\Omega^0}$ and $u(0) = u^0$.

2.2 *The ALE mapping and ALE evolving FEM*

Assume that the given smooth initial domain $\Omega^0 \subset \mathbb{R}^d$, $d \in \{2,3\}$, is divided into a set $\mathscr{K}_h^0$ of shape-regular and quasi-uniform simplices with mesh size h, using curved finite elements of degree k, and fitting the interface $\Gamma^0 = \Gamma(0)$. Every possibly curved closed simplex $K \in \mathscr{K}_h^0$ is the image of a unique polynomial of degree k defined on the reference simplex $\hat{K}$, denoted by $F_K : \hat{K} \to K$, called the parametrization of K; see [16, §8.6]. Every boundary simplex $K \in \mathscr{K}_h^0$ (with one face or edge attached to Γ^0) contains a possibly curved face or edge to interpolate Γ^0 with accuracy of $O(h^{k+1})$. This can be obtained by using Lenoir's isoparametric finite elements [34]. Thus the initial domain Ω^0 is approximated by the triangulated domain

$$\Omega_h^0 = \Big(\bigcup_{K \in \mathscr{K}_h^0} K \Big) \backslash \partial\Omega_h^0 = \Omega_{h,+}^0 \cup \Omega_{h,-}^0 \cup \Gamma_h^0.$$

Let $\mathbf{x}^0 = (\xi_1, \cdots, \xi_M) \in \mathbb{R}^{dM}$ be the vector that collects all the nodes $\xi_j \in \mathbb{R}^d$, $j = 1,\ldots,M$, (which define finite elements of degree k) in the triangulation $\mathscr{K}_h^0$. We evolve the vector $\mathbf{x}^0$ in time and denote its position at time t by

$$\mathbf{x}(t) = (x_1(t), \cdots, x_M(t)),$$

which determines the triangulation $\mathscr{K}_h[\mathbf{x}(t)]$ and domain $\Omega_h[\mathbf{x}(t)]$, as well as subdomains $\Omega_{h,\pm}[\mathbf{x}(t)]$ and interface $\Gamma_h[\mathbf{x}(t)]$ such that

$$\Omega_h[\mathbf{x}(t)] = \Big(\bigcup_{K \in \mathscr{K}_h[\mathbf{x}(t)]} K \Big) \backslash \partial\Omega_h[\mathbf{x}(t)] = \Omega_{h,+}[\mathbf{x}(t)] \cup \Omega_{h,-}[\mathbf{x}(t)] \cup \Gamma_h[\mathbf{x}(t)].$$

Similarly as the simplices on the initial domain, if $K \in \mathscr{K}_h[\mathbf{x}(t)]$ is a simplex on the evolving domain $\Omega_h[\mathbf{x}(t)]$, then we denote by $F_K : \hat{K} \to K$ the parametrization of K. For the simplicity of notation, we denote

$$\Omega_h(t) = \Omega_h[\mathbf{x}(t)], \quad \Omega_{h,\pm}(t) = \Omega_{h,\pm}[\mathbf{x}(t)] \quad \text{and} \quad \Gamma_h(t) = \Gamma_h[\mathbf{x}(t)].$$

Correspondingly, the scalar-valued finite element space on the evolving domain $\Omega_h(t)$ is defined as

$$S_h^k[\mathbf{x}(t)] := \{v_h \in H^1(\Omega_h(t)) : v_h \circ F_K \in P^k(\hat{K}) \text{ for all } K \in \mathscr{K}_h[\mathbf{x}(t)]\}$$

$$S_{h,\mathrm{dc}}^{k-1}[\mathbf{x}(t)] := \{v_h \in H^1(\Omega_{h,\pm}(t)) : v_h \circ F_K \in P^{k-1}(\hat{K}) \text{ for all } K \in \mathscr{K}_h[\mathbf{x}(t)]\}.$$

The ALE mapping $\phi_h(\cdot, t)$ is defined as the unique finite element function $\phi_h(\cdot, t) \in S_h^k[\mathbf{x}^0]^d$ such that

$$\phi_h(\xi_j, t) = x_j(t) \quad \text{for} \quad j = 1,\ldots,M,$$

which maps the initial domain $\Omega_{h,\pm}^0$ onto the evolving domain $\Omega_{h,\pm}(t)$, and the initial interface Γ_h^0 onto the evolving interface $\Gamma_h(t)$, i.e.,

$$\Omega_{h,\pm}(t) = \phi_h(\Omega_{h,\pm}^0, t) \quad \text{and} \quad \Gamma_h(t) = \phi_h(\Gamma_h^0, t) \quad \text{for} \quad t \in [0, T].$$

The ALE mapping $\phi_h(\cdot, t)$ is the flow map generated by a mesh velocity $w_h(\cdot, t)$ and satisfies the following relation:

$$\frac{\mathrm{d}}{\mathrm{d}t} \phi_h(\xi, t) = w_h(t) \circ \phi_h(\xi, t) \quad \text{for} \quad \xi \in \Omega_h^0. \tag{2.5}$$

Based on the ALE mapping $\phi_h(\cdot, t)$ and the mesh velocity $w_h(\cdot, t)$, if $v(\cdot, t)$ is a function defined on $\Omega_h(t)$, its discrete material derivative is defined as

$$\partial_{t,h}^\bullet v(x, t) = \frac{\mathrm{d}}{\mathrm{d}t} v(\phi_h(\xi, t), t) \quad \text{for } x = \phi_h(\xi, t) \in \Omega_h(t) \quad \text{for } \xi \in \Omega_h^0.$$

Let $X_h[\mathbf{x}(t)] \times M_h[\mathbf{x}(t)] \subset H_0^1(\Omega_h(t))^d \times L_0^2(\Omega_h(t))$ denote the Taylor–Hood finite element space of degree $k \geqslant 2$ subject to the triangulation $\mathscr{K}_h[\mathbf{x}(t)]$ of evolving domain $\Omega_h(t)$, i.e.,

$$X_h[\mathbf{x}(t)] = S_h^k[\mathbf{x}(t)]^d \cap H_0^1(\Omega_h(t))^d \quad \text{and} \quad M_h[\mathbf{x}(t)] = S_{h,\mathrm{dc}}^{k-1}[\mathbf{x}(t)] \cap L_0^2(\Omega_h(t)).$$

We also denote by $\mathring{X}_h[\mathbf{x}(t)]$ the subspace of $X_h[\mathbf{x}(t)]$ with zero interface and boundary conditions, i.e.,

$$\mathring{X}_h[\mathbf{x}(t)] = \{v_h \in X_h[\mathbf{x}(t)] : v_h = 0 \text{ on } \Gamma_h(t) = \Gamma_h[\mathbf{x}(t)] \text{ and on } \partial\Omega_h(t) = \partial\Omega_h[\mathbf{x}(t)]\}.$$

It is known that the Taylor–Hood finite element space satisfies the discrete inf-sup condition (see [11, §8.5.3] on the domain decomposition techniques), i.e., there exists positive constant κ independent of h such that

$$\kappa \|q_h\|_{L^2(\Omega_h(t))} \leqslant \sup_{0 \neq v_h \in X_h[\mathbf{x}(t)]} \frac{(q_h, \nabla \cdot v_h)_{\Omega_h}}{\|v_h\|_{H^1(\Omega_h(t))}} \quad \forall q_h \in M_h[\mathbf{x}(t)], \tag{2.6}$$

where

$$(u_h, v_h)_{\Omega_h} := \int_{\Omega_{h,+}(t)} u_h v_h \, dx + \int_{\Omega_{h,-}(t)} u_h v_h \, dx.$$

Since the mesh velocity w_h should be an approximation of w defined in Section 2.1, we define $w_h(\cdot,t)$ to be a discrete harmonic finite element function in the numerically computed subdomains $\Omega_{h,\pm}(t)$ with boundary condition matching the velocity of the fluid on the boundary $\partial \Omega_h(t)$ and the numerically computed interface $\Gamma_h(t)$. This is equivalent to finding $w_h \in u_h + \mathring{X}_h[\mathbf{x}(t)]$ such that

$$\int_{\Omega_h(t)} \nabla w_h \cdot \nabla \chi_h \, dx = 0 \quad \text{for} \quad \chi_h \in \mathring{X}_h[\mathbf{x}(t)]. \tag{2.7}$$

We denote by ρ_h and μ_h the piecewise constant functions in $\Omega_h(t)$, defined by

$$\rho_h = \begin{cases} \rho_+ & \text{in } \Omega_{h,+}(t) \\ \rho_- & \text{in } \Omega_{h,-}(t) \end{cases} \quad \text{and} \quad \mu_h = \begin{cases} \mu_+ & \text{in } \Omega_{h,+}(t) \\ \mu_- & \text{in } \Omega_{h,-}(t). \end{cases}$$

The ALE evolving FEM for (2.4) is defined as follows.

ALE evolving FEM: Find a flow map $\phi_h(\cdot,t) \in S_h^k[\mathbf{x}^0]^d$ which determines nodal vector $\mathbf{x}(t) = \phi_h(\mathbf{x}^0, t)$ and corresponding subdomains $\Omega_{h,\pm}(t) = \Omega_{h,\pm}[\mathbf{x}(t)]$, as well as $(u_h(\cdot,t), p_h(\cdot,t)) \in X_h[\mathbf{x}(t)] \times M_h[\mathbf{x}(t)]$ and $w_h(\cdot,t) \in u_h(\cdot,t) + \mathring{X}_h[\mathbf{x}(t)]$ such that

$$\begin{cases} (\rho_h \partial_{t,h}^\bullet u_h, v_h)_{\Omega_h} + 2(\mu_h \mathbb{D}(u_h), \mathbb{D}(v_h))_{\Omega_h} - (p_h, \nabla \cdot v_h)_{\Omega_h} \\ \qquad\qquad + (\rho_h(u_h - w_h) \cdot \nabla u_h, v_h)_{\Omega_h} = (f_h, v_h)_{\Omega_h} \qquad \forall v_h \in X_h[\mathbf{x}(t)], \\ (\nabla \cdot u_h, q_h)_{\Omega_h} = 0 \qquad\qquad\qquad\qquad\qquad\qquad\qquad \forall q_h \in M_h[\mathbf{x}(t)], \\ (\nabla w_h, \nabla \chi_h)_{\Omega_h} = 0 \qquad\qquad\qquad\qquad\qquad\qquad\qquad \forall \chi_h \in \mathring{X}_h[\mathbf{x}(t)], \\ \partial_t \phi_h = w_h \circ \phi_h, \end{cases} \tag{2.8}$$

with initial conditions $\phi_h(0) = \text{id}|_{\Omega_h^0}$ and $u_h(0) = u_h^0$, where u_h^0 is the Lagrange interpolation of u^0 and f_h is the Lagrange interpolation of f on $\Omega_h(t)$.

REMARK 2.1 The finite element basis functions of $S_h^k[\mathbf{x}(t)]$ are denoted by $\varphi_j[\mathbf{x}(t)] : \Omega_h(t) \to \mathbb{R}$ for $j = 1, \ldots, M$, which have the property that on every triangle their pullback to the reference triangle is polynomial of degree k, satisfying the identities

$$\varphi_j[\mathbf{x}(t)](x_i(t)) = \delta_{ij} \quad \text{for all} \quad i, j = 1, \ldots, M.$$

The pullback of $\varphi_j[\mathbf{x}(t)]$ from $\Omega_h(t)$ to Ω_h^0 is $\varphi_j[\mathbf{x}(t)] \circ \phi_h(\cdot,t) = \varphi_j[\mathbf{x}^0]$, which is simply the finite element basis functions on Ω_h^0. Therefore, the basis functions $\varphi_j[\mathbf{x}(t)]$ satisfies the transport property [14]:

$$\partial_{t,h}^\bullet \varphi_j[\mathbf{x}(t)](x) = 0 \quad \text{for } x \in \Omega_h(t), \ j = 1, \ldots, M. \tag{2.9}$$

By using the basis functions $\varphi_j[\mathbf{x}(t)]$, the finite element space $S_h^k[\mathbf{x}(t)]$ and the discrete material derivative of the finite element solutions $v_h(\cdot,t) \in S_h^k[\mathbf{x}(t)]$ can be written as

$$S_h^k[\mathbf{x}(t)] = \Big\{ \sum_{j=1}^M c_j \varphi_j[\mathbf{x}(t)] : c_j \in \mathbb{R} \Big\},$$

$$\partial_{t,h}^\bullet v_h(x,t) = \sum_{j=1}^M \dot{v}_j(t) \varphi_j[\mathbf{x}(t)](x) \quad \text{with} \quad v_h(x,t) = \sum_{j=1}^M v_j(t) \varphi_j[\mathbf{x}(t)](x) \quad \text{for } x \in \Omega_h(t),$$

where the dot denotes the time derivative with respect to t.

REMARK 2.2 Due to the use of higher-order isoparametric elements, the integrals in (2.8) cannot be evaluated exactly in practice and are instead approximated using numerical quadrature. The quadrature rules are chosen based on the geometry and polynomial order of the elements to ensure high accuracy. In our analysis, quadrature errors are not considered, as they are typically negligible compared to the discretization error.

3. Interpolated evolving mesh and consistency analysis

In the error analysis, we need to compare functions on two different domains, i.e., the exact domain $\Omega(t)$, and the triangulated domain $\Omega_h(t)$ determined by the numerical solution. The consistency error is the defect upon inserting the Lagrange interpolation of the exact solution u and the interpolated interface into the discretized equations. To this end, we define the interpolated ALE mapping and interpolated evolving mesh below.

3.1 *Interpolated ALE mapping and interpolated evolving mesh*

We evolve the vector $\mathbf{x}^0 = (\xi_1, \cdots, \xi_M)$ in time and denote $\mathbf{x}^*(t) = (x_1^*(t), \ldots, x_M^*(t))$ the nodal vector which collects the points $x_j^*(t) = \phi(\xi_j, t)$ on the exact domain $\Omega(t)$, which determines triangulation $\mathscr{K}_h[\mathbf{x}^*(t)]$ and interpolated domain $\Omega_h^*(t)$. Since the initial triangulation is shape-regular and quasi-uniform, the interpolated triangulations $\mathscr{K}_h[\mathbf{x}^*(t)]$ will keep shape-regular and quasi-uniform for $t \in [0, T]$ when the flow map $\phi(\cdot, t) : \Omega_\pm(0) \to \Omega_\pm(t)$ and its inverse are piecewise smooth on $\Omega_\pm(0)$ and $\Omega_\pm(t)$ for $t \in [0, T]$, respectively.

Subject to the triangulation $\mathscr{K}_h[\mathbf{x}^*(t)]$ of domain $\Omega_h^*(t)$, the space of scalar-valued finite element functions on $\Omega_h^*(t)$ is defined as

$$S_h^k[\mathbf{x}^*(t)] := \{v_h \in H^1(\Omega_h^*(t)) : v_h \circ F_K \in P^k(\hat{K}) \text{ for all } K \in \mathscr{K}_h[\mathbf{x}^*(t)]\}.$$

We define the interpolated ALE mapping $\phi_h^*(\cdot, t) : \Omega_h^0 \to \Omega_h^*(t)$ to be the Lagrange interpolation of flow map $\phi(\cdot, t) : \Omega(0) \to \Omega(t)$, which has the following properties:

$$\phi_h^*(\xi_j, t) = x_j^*(t) = \phi(\xi_j, t) \quad \text{for} \quad j = 1, \ldots, M,$$

and

$$\Omega_{h,\pm}^*(t) = \phi_h^*(\Omega_{h,\pm}^0, t) \quad \text{and} \quad \Gamma_h^*(t) = \phi_h^*(\Gamma_h^0, t).$$

If we denote by $w_h^*(\cdot, t) = I_h^*(t) w(\cdot, t)$ the Lagrange interpolations of the exact velocity $w(\cdot, t)$, then the following relation holds:

$$\frac{\mathrm{d}}{\mathrm{d}t} \phi_h^*(\xi, t) = w_h^*(\phi_h^*(\xi, t), t) \quad \text{for} \quad \xi \in \Omega_h^0. \tag{3.1}$$

This equality holds because both sides are finite element functions on Ω_h^0 and they are equal on the nodes.

Let u_h^* be the Lagrange interpolation of the exact solution u, and denote by $w_{h,\pm}^*$ and $u_{h,\pm}^*$ the restriction of w_h^* and u_h^* to subdomains $\Omega_{h,\pm}^*(t)$. Then the interpolated mesh velocity satisfies $w_h^* = u_h^*$ on the interpolated interface $\Gamma_h^*(t)$. Thus $w_{h,\pm}^* - u_{h,\pm}^* \in \mathring{X}_{h,\pm}[\mathbf{x}^*(t)]$.

For a function $v(\cdot, t)$ defined on $\Omega_h^*(t)$, we can define the discrete material derivative $\partial_{t,*}^\bullet$ by using the interpolated ALE mapping $\phi_h^*(\cdot, t)$ and the interpolated mesh velocity $w_h^*(\cdot, t)$ as

$$\partial_{t,*}^\bullet v(x, t) = \frac{\mathrm{d}}{\mathrm{d}t} v(\phi_h^*(\xi, t), t) \quad \text{at point} \quad x = \phi_h^*(\xi, t) \in \Omega_h^*(t). \tag{3.2}$$

3.2 *Lifting operator*

The transition between the interpolated domain $\Omega_h^*(t)$ and the exact domain $\Omega_h(t)$ is done by a lift operator. In Lenoir's isoparametric finite element approximation to Ω^0, there exists a map $\Phi : \Omega_h^0 \to \Omega^0$ such that (cf. [34]):

$$\|\Phi(\cdot) - \mathrm{id}\|_{L^\infty(\Omega_h^0)} + h\|\Phi(\cdot) - \mathrm{id}\|_{W^{1,\infty}(\Omega_h^0)} \leqslant Ch^{k+1} \quad \text{in} \quad \Omega_h^0.$$

For the triangulation $\mathscr{K}_h[\mathbf{x}^*(t)]$ which fits the interface $\Gamma(t)$, there exists a continuous mapping $\Phi_h^*(\cdot, t) : \Omega_h^*(t) \to \Omega(t)$ given by

$$\Phi_h^*(\cdot, t) = \phi(\cdot, t) \circ \Phi \circ \phi_h^*(\cdot, t)^{-1} \quad \text{in} \quad \Omega_h^*(t), \tag{3.3}$$

that maps $\Gamma_h^*(t)$ to $\Gamma(t)$ and satisfies the following estimates

$$\|\Phi_h^*(\cdot, t) - \text{id}\|_{L^\infty(\Omega_h^*(t))} + h\|\Phi_h^*(\cdot, t) - \text{id}\|_{L^\infty(\Omega_h^*(t))} \leqslant Ch^{k+1} \quad \text{in} \quad \Omega_h^*(t). \tag{3.4}$$

Correspondingly, for $v_h : \Omega_h^*(t) \to \mathbb{R}$, we denote its lift by v_h^ℓ given by

$$v_h^\ell = v_h \circ \Phi_h^*(\cdot, t)^{-1} \quad \text{in} \quad \Omega(t). \tag{3.5}$$

- (Lagrange interpolation) Define $\hat{I}_{h,\hat{K}}$ be the Lagrange interpolation on the reference simplex $\hat{K}$, then the following relation hold:

$$(I_h^*(t)v) \circ F_K = \hat{I}_{h,\hat{K}}\left(v \circ \Phi_h^*(\cdot, t) \circ F_K\right). \tag{3.6}$$

- (Commutation of material derivative and lift) (cf. [16, Lemma 3.5]):

$$\partial_t^\bullet(v_h^\ell) = \left(\partial_{t,*}^\bullet v_h\right)^\ell \qquad \text{for} \quad v_h : \Omega_h^*(t) \to \mathbb{R}. \tag{3.7}$$

- (Transport property of the lifted basis functions $\varphi_j[\mathbf{x}^*]^\ell$) (cf. [13, Lemma 4.1]):

$$\partial_t^\bullet\left(\varphi_j[\mathbf{x}^*(t)]^\ell\right) = \left(\partial_{t,*}^\bullet \varphi_j[\mathbf{x}^*(t)]\right)^\ell = 0, \quad j = 1, \ldots, M. \tag{3.8}$$

- (Commutation of material derivative of Lagrange interpolation): Utilizing the transport property of the lifted basis functions, we have

$$\partial_{t,*}^\bullet I_h^*(t)v = I_h^*(t)\partial_t^\bullet v. \tag{3.9}$$

Define $\mathscr{I}_h^* u := (I_h^* u)^\ell : \Omega(t) \to \mathbb{R}$. It follows from (3.8) and (3.7) that

$$\partial_t^\bullet \mathscr{I}_h^* u = \mathscr{I}_h^* \partial_t^\bullet u. \tag{3.10}$$

It plays a crucial role in proving the optimal-order in consistency error.

- (Approximation properties for the Lagrange interpolation) There exists a constant $C > 0$ independent of $h \leqslant h_0$, with a sufficiently small $h_0 > 0$, and t such that for $u(\cdot, t) \in H^{k+1}(\Omega_\pm(t))$, for $0 \leqslant t \leqslant T$ (cf. [30, Lemma 7.5]):

$$\|u - \mathscr{I}_h^*(t)u\|_{L^2(\Omega_\pm(t))} + h\|\nabla u - \nabla \mathscr{I}_h^*(t)u\|_{L^2(\Omega_\pm(t))} \leqslant Ch^{k+1}\|u\|_{H^{k+1}(\Omega_\pm(t))}. \tag{3.11}$$

- (Equivalence of norms) The L^p and $W^{1,p}$ norms on the discrete and continuous domains are equivalent for $1 \leqslant p \leqslant \infty$, uniformly in the mesh size $h \leqslant h_0$ (with sufficiently small $h_0 > 0$) and in $t \in [0,T]$. In particular, for $v_h : \Omega_h^*(t) \to \mathbb{R}$ with lift $v_h^\ell : \Omega(t) \to \mathbb{R}$, there is a constant C such that for $h \leqslant h_0$ and $0 \leqslant t \leqslant T$,

$$C^{-1}\|v_h\|_{L^2(\Omega_h^*(t))} \leqslant \|v_h^\ell\|_{L^2(\Omega(t))} \leqslant C\|v_h\|_{L^2(\Omega_h^*(t))}, \tag{3.12}$$

$$C^{-1}\|v_h\|_{H^1(\Omega_h^*(t))} \leqslant \|v_h^\ell\|_{H^1(\Omega(t))} \leqslant C\|v_h\|_{H^1(\Omega_h^*(t))}. \tag{3.13}$$

The lift operator ℓ maps a function on the interpolated domain $\Omega_h^*(t)$ to a function on the exact domain $\Omega(t)$, provided that $\Omega_h^*(t)$ is sufficiently close to $\Omega(t)$.

In the following we drop the argument t when it is not essential.

3.3 *Weak formulation on the interpolated evolving mesh*

For the simplicity of notation, we denote by ρ_h^* and μ_h^* the piecewise constant functions in $\Omega_h^*(t)$ defined by

$$\rho_h^* = \begin{cases} \rho_+ & \text{in } \Omega_{h,+}^*(t), \\ \rho_- & \text{in } \Omega_{h,-}^*(t), \end{cases} \quad \text{and} \quad \mu_h^* = \begin{cases} \mu_+ & \text{in } \Omega_{h,+}^*(t), \\ \mu_- & \text{in } \Omega_{h,-}^*(t), \end{cases} \tag{3.14}$$

and define the following inner product:

$$(u_h, v_h)_{\Omega_h^*} = \int_{\Omega_{h,+}^*(t)} u_h v_h \, \mathrm{d}x + \int_{\Omega_{h,-}^*(t)} u_h v_h \, \mathrm{d}x. \tag{3.15}$$

After replacing the solution by its finite element interpolation in weak formulation (2.4), and replacing domains $\Omega(t)$ and $\Omega_\pm(t)$ by the interpolated domains $\Omega_h^*(t)$ and $\Omega_{h,\pm}^*(t)$, respectively, we obtain the following weak formulation satisfied by the interpolated functions u_h^*, p_h^*, w_h^* and ϕ_h^*:

$$\begin{cases} (\rho_h^* \partial_{t,*}^\bullet u_h^*, v_h)_{\Omega_h^*} + 2(\mu_h^* \mathbb{D}(u_h^*), \mathbb{D}(v_h))_{\Omega_h^*} - (p_h^*, \nabla \cdot v_h)_{\Omega_h^*} \\ \qquad + (\rho_h^*(u_h^* - w_h^*) \cdot \nabla u_h^*, v_h)_{\Omega_h^*} - (f_h^*, v_h)_{\Omega_h^*} = \mathscr{E}_{u1}(t, v_h) \quad \forall\, v_h \in X_h[\mathbf{x}^*(t)], \\ (\nabla \cdot u_h^*, q_h)_{\Omega_h^*} = \mathscr{E}_{u2}(t, q_h) \qquad\qquad\qquad\qquad\qquad\quad \forall\, q_h \in M_h[\mathbf{x}^*(t)], \\ (\nabla w_h^*, \nabla \chi_h)_{\Omega_h^*} = \mathscr{E}_w(t, \chi_h) \qquad\qquad\qquad\qquad\qquad\;\; \forall\, \chi_h \in \mathring{X}_h[\mathbf{x}^*(t)], \\ \partial_t \phi_h^* = w_h^* \circ \phi_h^*, \end{cases} \tag{3.16}$$

with initial conditions $\phi_h^*(0) = \mathrm{id}|_{\Omega_h^0}$ and $u_h^*(0) = u_h^0$, and f_h^* is the Lagrange interpolation of f on $\Omega_h^*(t)$. where $\mathscr{E}_{u1}(t, v_h)$, $\mathscr{E}_{u2}(t, q_h)$ and $\mathscr{E}_w(t, \chi_h)$ are consistency errors arising from finite element approximations to the solution and domain. For example,

$$\begin{aligned} \mathscr{E}_{u1}(t, v_h) := & (\rho_h^* \partial_{t,*}^\bullet u_h^*, v_h)_{\Omega_h^*} - (\rho \partial_t^\bullet u, v_h^\ell) \\ & + 2(\mu_h^* \mathbb{D}(u_h^*), \mathbb{D}(v_h))_{\Omega_h^*} - 2(\mu \mathbb{D}(u), \mathbb{D}(v_h^\ell)) \\ & + (\rho_h^*(u_h^* - w_h^*) \cdot \nabla u_h^*, v_h)_{\Omega_h^*} - (\rho(u - w) \cdot \nabla u, v_h^\ell) \\ & - (p_h^*, \nabla \cdot v_h)_{\Omega_h^*} + (p, \nabla \cdot v_h^\ell) - (f_h^*, v_h)_{\Omega_h^*} + (f, v_h^l), \end{aligned}$$

which contains errors arising from approximating domain $\Omega_\pm(t)$ by $\Omega_{h,\pm}^*(t)$. This can be controlled by using the following estimates for the error caused by domain perturbation (see [30]):

$$\left| \int_{\Omega_{h,\pm}^*(t)} g \chi_h \mathrm{d}x - \int_{\Omega_\pm(t)} g^\ell \chi_h^\ell \mathrm{d}x \right| \leqslant C h^k \|\chi_h\|_{L^2(\Omega_{h,\pm}^*(t))} \|g\|_{L^2(\Omega_{h,\pm}^*(t))},$$

$$\left| \int_{\Omega_{h,\pm}^*(t)} g \nabla \chi_h \mathrm{d}x - \int_{\Omega_\pm(t)} g^\ell \nabla \chi_h^\ell \mathrm{d}x \right| \leqslant C h^k \|\chi_h\|_{H^1(\Omega_{h,\pm}^*(t))} \|g\|_{L^2(\Omega_{h,\pm}^*(t))},$$

$$\left| \int_{\Omega_{h,\pm}^*(t)} \nabla \chi_h \cdot \nabla \psi_h \mathrm{d}x - \int_{\Omega_\pm(t)} \nabla \chi_h^\ell \cdot \nabla \psi_h^\ell \mathrm{d}x \right| \leqslant C h^k \|\chi_h\|_{H^1(\Omega_{h,\pm}^*(t))} \|\psi_h\|_{H^1(\Omega_{h,\pm}^*(t))},$$

where $\chi_h, \psi_h \in S_h^k[\mathbf{x}^*(t)]$ and $g \in L^2(\Omega_h^*(t))$. By using this result and the estimates of the interpolation error shown in (3.11), the following estimates of the consistency errors can be derived when the domain $\Omega(t)$ and interface $\Gamma(t)$ are smooth, and the solution is piecewise smooth in each subdomain $\Omega_\pm(t)$:

$$\left| \mathscr{E}_{u1}(t, v_h) \right| \leqslant C h^k \|v_h\|_{H^1(\Omega_h^*(t))} \quad \text{and} \quad \left| \mathscr{E}_{u2}(t, q_h) \right| \leqslant C h^k \|q_h\|_{L^2(\Omega_h^*(t))}, \tag{3.17}$$

$$\left| \mathscr{E}_w(t, \chi_h) \right| \leqslant C h^k \|\chi_h\|_{H^1(\Omega_h^*(t))}, \tag{3.18}$$

for $v_h \in X_h[\mathbf{x}^*(t)]$, $q_h \in M_h[\mathbf{x}^*(t)]$ and $\chi_h \in \mathring{X}_h[\mathbf{x}^*(t)]$. In the case $\partial_{t,*}^\bullet v_h = 0$ and $\partial_{t,*}^\bullet q_h = 0$ on $\Omega_h^*(t)$, the derivatives of these consistency errors also satisfy similar estimates, i.e.,

$$\left| \frac{\mathrm{d}}{\mathrm{d}t} \mathscr{E}_{u1}(t, v_h) \right| \leqslant C h^k \|v_h\|_{H^1(\Omega_h^*(t))} \quad \text{and} \quad \left| \frac{\mathrm{d}}{\mathrm{d}t} \mathscr{E}_{u2}(t, q_h) \right| \leqslant C h^k \|q_h\|_{L^2(\Omega_h^*(t))}. \tag{3.19}$$

REMARK 3.1 From the expression of $\mathscr{E}_{u1}(t, v_h)$ we can see that the first consistency estimate in (3.19) requires not only $\|u\|_{H^{k+1}(\Omega_\pm(t))}$ and $\|\partial_t^\bullet u\|_{H^{k+1}(\Omega_\pm(t))}$ but also $\|\partial_t^\bullet \partial_t^\bullet u\|_{H^{k+1}(\Omega_\pm(t))}$ to be bounded. This is the strongest regularity required for the consistency estimates, and is assumed in Proposition 3.1 for consistency estimates and Theorem 4.1 for error estimates.

3.4 *Matrix-vector form of the weak formulation*

For both computation and error analysis it is convenient to use the matrix-vector form of the finite element weak formulation in (2.8). Through the utilization of finite element basis functions $\varphi_j[\mathbf{x}(t)]$ on the domain $\Omega_h(t)$, the numerical solution $u_h(t)$ can be expressed as a column vector $\mathbf{u} = (u_1(t), \ldots, u_M(t))$, where

$$u_h(t) = \sum_{j=1}^{M} u_j(t) \varphi_j[\mathbf{x}(t)] \qquad \text{for} \qquad u_j(t) \in \mathbb{R}^d.$$

Similarly, the solutions $p_h(t)$ and $w_h(t)$ on $\Omega_h(t)$ can also be represented as column vector $\mathbf{p}$ and $\mathbf{w}$, respectively. Then $\mathbf{x}(t) = (x_1(t), \ldots, x_M(t)) = \phi_h(\mathbf{x}^0, t)$ is the vector of positions of all finite element nodes at t, satisfying

$$\frac{d\mathbf{x}(t)}{dt} = \mathbf{w}(t) \quad \text{with} \quad \mathbf{x}(0) = \mathbf{x}^0, \tag{3.20}$$

where the initial value $\mathbf{x}^0$ is given by the triangulation of the domain Ω at time $t = 0$.

We introduce domain-dependent mass matrix $\mathbf{M}(\mathbf{x})$, the stiffness matrix $\mathbf{A}(\mathbf{x})$, and the matrix $\mathbf{M}_{\boldsymbol{\rho}}(\mathbf{x})$ and $\mathbf{A}_{\boldsymbol{\mu}}(\mathbf{x})$ on the domain $\Omega_h(t)$ as follows:

$$\mathbf{v}^\top \mathbf{M}(\mathbf{x})\mathbf{u} = \sum_{K \in \mathscr{K}_h[\mathbf{x}]} \int_K u_h \cdot v_h \, dx, \qquad \mathbf{v}^\top \mathbf{M}_{\boldsymbol{\rho}}(\mathbf{x})\mathbf{u} = \sum_{K \in \mathscr{K}_h[\mathbf{x}]} \rho_K \int_K u_h \cdot v_h \, dx,$$

$$\mathbf{v}^\top \mathbf{A}(\mathbf{x})\mathbf{u} = \sum_{K \in \mathscr{K}_h[\mathbf{x}]} \int_K \nabla u_h \cdot \nabla v_h \, dx, \qquad \mathbf{v}^\top \mathbf{A}_{\boldsymbol{\mu}}(\mathbf{x})\mathbf{u} = \sum_{K \in \mathscr{K}_h[\mathbf{x}]} 2\mu_K \int_K \mathbb{D}(u_h) \cdot \mathbb{D}(v_h) \, dx.$$

With identity matrix $I_d \in \mathbb{R}^{d \times d}$, we define $\mathbf{K}(\mathbf{x}^n)^d$ as the Kronecker product of I_d and $\mathbf{K}(\mathbf{x}^n) := \mathbf{M}(\mathbf{x}^n) + \mathbf{A}(\mathbf{x}^n)$, i.e.,

$$\mathbf{K}(\mathbf{x}^n)^d := I_d \otimes (\mathbf{M}(\mathbf{x}^n) + \mathbf{A}(\mathbf{x}^n)).$$

To simplify the notation, we will use $\mathbf{K}(\mathbf{x}^n)$ to represent $\mathbf{K}(\mathbf{x}^n)^d$ when the dimension of the matrix is clear and therefore no confusion arises. With the matrices defined above, the L^2 and H^1 norm of finite element functions can be expressed as quadratic forms:

$$\|\mathbf{u}\|_{\mathbf{M}(\mathbf{x})}^2 := \mathbf{u}^\top \mathbf{M}(\mathbf{x})\mathbf{u} = \|u_h\|_{L^2(\Omega_h(t))}^2, \tag{3.21}$$

$$\|\mathbf{u}\|_{\mathbf{A}(\mathbf{x})}^2 := \mathbf{u}^\top \mathbf{A}(\mathbf{x})\mathbf{u} = \|\nabla u_h\|_{L^2(\Omega_h(t))}^2, \tag{3.22}$$

$$\|\mathbf{u}\|_{\mathbf{K}(\mathbf{x})}^2 := \mathbf{u}^\top \mathbf{K}(\mathbf{x})\mathbf{u} = \|u_h\|_{H^1(\Omega_h(t))}^2. \tag{3.23}$$

In the same way, we can define the matrices $\mathbf{B}_{\boldsymbol{\rho}}(\mathbf{x}, \mathbf{u})$ and $\mathbf{C}(\mathbf{x})$ such that

$$\mathbf{v}^\top \mathbf{B}_{\boldsymbol{\rho}}(\mathbf{x}, \mathbf{u})\boldsymbol{\chi} = \sum_{K \in \mathscr{K}_h[\mathbf{x}]} \rho_K \int_K (u_h \cdot \nabla \chi_h) \cdot v_h \, dx,$$

$$\mathbf{q}^\top \mathbf{C}(\mathbf{x})\mathbf{v} = \sum_{K \in \mathscr{K}_h[\mathbf{x}]} \int_K (\nabla \cdot v_h) q_h \, dx = \mathbf{v}^\top \mathbf{C}(\mathbf{x})^\top \mathbf{q}.$$

In situations where there is no ambiguity in the context, we omit the time dependency and rewrite the weak formulation (2.8) into the following matrix-vector form:

$$\begin{cases} \mathbf{M}_{\boldsymbol{\rho}}(\mathbf{x})\dot{\mathbf{u}} + \mathbf{B}_{\boldsymbol{\rho}}(\mathbf{x}, \mathbf{u} - \mathbf{w})\mathbf{u} + \mathbf{A}_{\boldsymbol{\mu}}(\mathbf{x})\mathbf{u} - \mathbf{C}(\mathbf{x})^\top \mathbf{p} - \mathbf{M}(\mathbf{x})\mathbf{f} = 0, & \text{(3.24a)} \\[4pt] \mathbf{C}(\mathbf{x})\mathbf{u} = 0, & \text{(3.24b)} \\[4pt] \boldsymbol{\chi}^\top \mathbf{A}(\mathbf{x})\mathbf{w} = \mathbf{0} \text{ for nodal vectors } \boldsymbol{\chi} \text{ associated to } \chi_h \in \mathring{X}_h[\mathbf{x}(t)], & \text{(3.24c)} \\[4pt] \dot{\mathbf{x}} = \mathbf{w}, & \text{(3.24d)} \end{cases}$$

where the column vector $\mathbf{f}$ collects the nodal values of f_h. In the same way, we can define the corresponding matrices and vectors over the interpolated domain $\Omega_h^*(t)$. The matrices and vectors denoted with a superscript

$*$ collect the nodal values of the interpolation of the exact solutions. For example, the vector $\mathbf{v}^*$ collects the nodal values of v_h^*. With this convention, we can rewrite the weak formulation (3.16) in the following form:

$$
\begin{cases}
\mathbf{M}_{\boldsymbol{\rho}}(\mathbf{x}^*)\dot{\mathbf{u}}^* + \mathbf{B}_{\boldsymbol{\rho}}(\mathbf{x}^*, \mathbf{u}^* - \mathbf{w}^*)\mathbf{u}^* + \mathbf{A}_{\boldsymbol{\mu}}(\mathbf{x}^*)\mathbf{u}^* - \mathbf{C}(\mathbf{x}^*)^\top \mathbf{p}^* - \mathbf{M}(\mathbf{x}^*)\mathbf{f}^* = \mathscr{E}_{\mathbf{u}1}, & (3.25a) \\[2mm]
\mathbf{C}(\mathbf{x}^*)\mathbf{u}^* = \mathscr{E}_{\mathbf{u}2}, & (3.25b) \\[2mm]
\boldsymbol{\chi}^\top \mathbf{A}(\mathbf{x}^*)\mathbf{w}^* = \boldsymbol{\chi}^\top \mathscr{E}_{\mathbf{w}} \text{ for nodal vectors } \boldsymbol{\chi} \text{ associated to } \chi_h \in \mathring{X}_h[\mathbf{x}^*(t)], & (3.25c) \\[2mm]
\dot{\mathbf{x}}^* = \mathbf{w}^*, & (3.25d)
\end{cases}
$$

where $\mathscr{E}_{\mathbf{u}1}$, $\mathscr{E}_{\mathbf{u}2}$, $\mathscr{E}_{\mathbf{w}}$ denote the vectorized forms of the defect terms $\mathscr{E}_{u1}(t,\cdot)$, $\mathscr{E}_{u2}(t,\cdot)$, and $\mathscr{E}_w(t,\cdot)$, which satisfy the following relations:

$$
\mathbf{v}^\top \mathscr{E}_{\mathbf{u}1} = \mathscr{E}_{u1}(t, v_h^*), \qquad \mathbf{q}^\top \mathscr{E}_{\mathbf{u}2} = \mathscr{E}_{u2}(t, q_h^*), \qquad \boldsymbol{\chi}^\top \mathscr{E}_{\mathbf{w}} = \mathscr{E}_w(t, \chi_h^*) \tag{3.26}
$$

for $v_h^* \in X_h[\mathbf{x}^*(t)]$, $q_h^* \in M_h[\mathbf{x}^*(t)]$, $\chi_h^* \in \mathring{X}_h[\mathbf{x}^*(t)]$ and the associated nodal vectors $\mathbf{v}$, $\mathbf{q}$, $\boldsymbol{\chi}$.

Convergence of the ALE evolving FEM is established in this article by comparing the matrix-vector formulations in (3.24) and (3.25). By subtracting (3.25) from (3.24), we obtain the following equations for the errors $\mathbf{e}_{\mathbf{w}} = \mathbf{w} - \mathbf{w}^*$, $\mathbf{e}_{\mathbf{u}} = \mathbf{u} - \mathbf{u}^*$, $\mathbf{e}_{\mathbf{p}} = \mathbf{p} - \mathbf{p}^*$, and $\mathbf{e}_{\mathbf{x}} = \mathbf{x} - \mathbf{x}^*$:

$$
\begin{cases}
\mathbf{M}_{\boldsymbol{\rho}}(\mathbf{x})\dot{\mathbf{e}}_{\mathbf{u}} + \mathbf{A}_{\boldsymbol{\mu}}(\mathbf{x}^*)\mathbf{e}_{\mathbf{u}} - \mathbf{C}(\mathbf{x})^\top \mathbf{e}_{\mathbf{p}} = -(\mathbf{M}_{\boldsymbol{\rho}}(\mathbf{x}) - \mathbf{M}_{\boldsymbol{\rho}}(\mathbf{x}^*))\dot{\mathbf{u}}^* & (3.27a) \\[1mm]
\hspace{5.5cm} - (\mathbf{A}_{\boldsymbol{\mu}}(\mathbf{x}) - \mathbf{A}_{\boldsymbol{\mu}}(\mathbf{x}^*))\mathbf{e}_{\mathbf{u}} \\[1mm]
\hspace{5.5cm} - (\mathbf{A}_{\boldsymbol{\mu}}(\mathbf{x}) - \mathbf{A}_{\boldsymbol{\mu}}(\mathbf{x}^*))\mathbf{u}^* \\[1mm]
\hspace{5.5cm} - \big(\mathbf{B}_{\boldsymbol{\rho}}(\mathbf{x}, \mathbf{u} - \mathbf{w})\mathbf{u} - \mathbf{B}_{\boldsymbol{\rho}}(\mathbf{x}^*, \mathbf{u}^* - \mathbf{w}^*)\mathbf{u}^*\big) \\[1mm]
\hspace{5.5cm} + (\mathbf{C}(\mathbf{x})^\top - \mathbf{C}(\mathbf{x}^*)^\top)\mathbf{p}^* + (\mathbf{M}(\mathbf{x})\mathbf{f} - \mathbf{M}(\mathbf{x}^*)\mathbf{f}^*) - \mathscr{E}_{\mathbf{u}1}, \\[2mm]
\mathbf{C}(\mathbf{x})\mathbf{e}_{\mathbf{u}} = -(\mathbf{C}(\mathbf{x}) - \mathbf{C}(\mathbf{x}^*))\mathbf{u}^* - \mathscr{E}_{\mathbf{u}2}, & (3.27b) \\[2mm]
\boldsymbol{\chi}^\top \mathbf{A}(\mathbf{x})\mathbf{e}_{\mathbf{w}} = -\boldsymbol{\chi}^\top [\mathbf{A}(\mathbf{x}) - \mathbf{A}(\mathbf{x}^*)]\mathbf{w}^* - \boldsymbol{\chi}^\top \mathscr{E}_{\mathbf{w}}, & (3.27c) \\[2mm]
\dot{\mathbf{e}}_{\mathbf{x}} = \mathbf{e}_{\mathbf{w}}, & (3.27d)
\end{cases}
$$

for all $\boldsymbol{\chi}$ associated to some $\chi_h \in \mathring{X}_h[\mathbf{x}^*(t)]$.

The estimates of the consistency errors in (3.17)–(3.19) can also be written into the matrix-vector form, as shown in the following proposition. This proposition generalizes the consistency estimates found in [30, Lemma 8.1], as it directly follows from the estimates for the error caused by domain perturbation and the approximation properties of Lagrangian interpolation.

PROPOSITION 3.1 (Consistency estimates) We assume that the flow map $\phi : \Omega_\pm(0) \times [0,T] \to \mathbb{R}^d$ and its inverse $\phi(\cdot, t)^{-1} : \Omega_\pm(t) \to \Omega_\pm(0)$ are both sufficiently smooth so that the subdomains $\Omega_\pm(t)$ are sufficiently smooth and the interpolated triangulations $\mathscr{K}_h[\mathbf{x}^*(t)]$ are shape-regular and quasi-uniform for $t \in [0,T]$, and assume that the solution of (1.1) is piecewise smooth on these subdomains in the sense that

$$
\max_{t \in [0,T]} \sum_{j=0}^{2} \|(\partial_t^\bullet)^j u\|_{H^{k+1}(\Omega_\pm(t))} \leqslant C_0 \quad \text{for some constant } C_0 > 0.
$$

Then there exist positive constants h_0 such that for $h \leqslant h_0$, the following consistency estimates hold:

$$
\left| \mathbf{v}^\top \mathscr{E}_{\mathbf{u}1} \right| \leqslant Ch^k \|\mathbf{v}\|_{\mathbf{K}(\mathbf{x}^*)} \quad \text{and} \quad \left| \mathbf{q}^\top \mathscr{E}_{\mathbf{u}2} \right| \leqslant Ch^k \|\mathbf{q}\|_{\mathbf{M}(\mathbf{x}^*)}, \tag{3.28}
$$

$$
\left| \boldsymbol{\chi}^\top \mathscr{E}_{\mathbf{w}} \right| \leqslant Ch^k \|\boldsymbol{\chi}\|_{\mathbf{K}(\mathbf{x}^*)} \quad \text{for } \boldsymbol{\chi} \text{ associated to some } \chi_h \in \mathring{X}_h[\mathbf{x}^*(t)]. \tag{3.29}
$$

Moreover, the derivatives of these defect terms satisfy similar estimates:

$$
\left| \mathbf{v}^\top \frac{\mathrm{d}}{\mathrm{d}t} \mathscr{E}_{\mathbf{u}1} \right| \leqslant Ch^k \|\mathbf{v}\|_{\mathbf{K}(\mathbf{x}^*)} \quad \text{and} \quad \left| \mathbf{q}^\top \frac{\mathrm{d}}{\mathrm{d}t} \mathscr{E}_{\mathbf{u}2} \right| \leqslant Ch^k \|\mathbf{q}\|_{\mathbf{M}(\mathbf{x}^*)}. \tag{3.30}
$$

4. Convergence of the ALE evolving FEM

We are now in the position to formulate the main result of this paper, which yields optimal-order convergence in the H^1 norm for the Taylor–Hood finite element semidiscretization of problem (3.24).

Theorem 4.1 (Convergence of the ALE evolving FEM) We assume that the solution of (1.1) satisfies the following global regularity

$$(u,p) \in C([0,T];H^1(\Omega)^d) \times L^2(0,T;L^2(\Omega)),$$

and is piecewise smooth in the sense that

$$\max_{t \in [0,T]} \left(\|u\|_{H^{k+1}(\Omega_\pm(t))} + \|\partial_t^\bullet u\|_{H^{k+1}(\Omega_\pm(t))} + \|\partial_t^\bullet \partial_t^\bullet u\|_{H^{k+1}(\Omega_\pm(t))} \right) \leqslant C_0 \tag{4.1}$$

$$\max_{t \in [0,T]} \left(\|p\|_{H^k(\Omega_\pm(t))} + \|\partial_t^\bullet p\|_{H^k(\Omega_\pm(t))} \right) \leqslant C_0 \quad \text{for some constant } C_0 > 0, \tag{4.2}$$

and assume that flow map $\phi : \Omega_\pm(0) \times [0,T] \to \mathbb{R}^d$ and its inverse $\phi(\cdot,t)^{-1} : \Omega_\pm(t) \to \Omega_\pm(0)$ are both sufficiently smooth so that the subdomains $\Omega_\pm(t)$ are sufficiently smooth and the interpolated triangulations $\mathscr{K}_h[\mathbf{x}^*(t)]$ are shape-regular and quasi-uniform for $t \in [0,T]$. Then there exists a positive constant h_0 such that for $h \leqslant h_0$, the finite element method in (2.8) with Taylor–Hood finite elements of degree $k \geqslant 2$ has following error bound:

$$\max_{t \in [0,T]} \|\phi_h(\cdot,t) - \phi_h^*(\cdot,t)\|_{H^1(\Omega_h^0)} + \max_{t \in [0,T]} \|u_h \circ \phi_h(\cdot,t) - u_h^* \circ \phi_h^*(\cdot,t)\|_{H^1(\Omega_h^0)}$$

$$+ \max_{t \in [0,T]} \|w_h \circ \phi_h(\cdot,t) - w_h^* \circ \phi_h^*(\cdot,t)\|_{H^1(\Omega_h^0)}$$

$$+ \|p_h \circ \phi_h(\cdot,s) - p_h^* \circ \phi_h^*(\cdot,s)\|_{L^2(0,T;L^2(\Omega_h^0))} \leqslant Ch^k, \tag{4.3}$$

where $\phi_h^* \in S_h^k[\mathbf{x}^0]^d$, $(u_h^*,p_h^*) \in X_h[\mathbf{x}^*(t)] \times M_h[\mathbf{x}^*(t)]$ and $w_h^* \in u_h^* + \mathring{X}_h[\mathbf{x}^*(t)]$ are the Lagrange interpolations of the exact solutions.

The proof of Theorem 4.1 will be presented in the next two subsections based on the technique of a homotopy map between two different domains $\Omega_h^*(t)$ and $\Omega_h(t)$.

4.1 *Comparison of integrals on two different domains*

To establish the convergence of the numerical schemes, it is necessary to compare the integrals over two different domains, i.e, the triangulated domain $\Omega_h^*(t)$ obtained from interpolating the exact domain $\Omega(t)$, and the triangulated domain $\Omega_h(t)$ determined by the numerical solution. In the spirit of the techniques in [15,30], we can obtain a similar sequence of results employing shape derivatives by constructing a homotopy map between $\Omega_h^*(t)$ and $\Omega_h(t)$.

To handle the rate of change of an integral over a moving domain, we will frequently make use of the following lemma (cf. [40, Lemma 5.7]):

Lemma 4.1 If the domain $\Omega(t)$ moves with velocity $w \in W^{1,\infty}(\Omega(t))$, then

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_{\Omega(t)} f \mathrm{d}x = \int_{\Omega(t)} (\partial_t^\bullet f + f \nabla \cdot w) \mathrm{d}x, \tag{4.4}$$

where the material derivative $\partial_t^\bullet f = \partial_t f + \nabla f \cdot w$.

REMARK 4.1 The commutation of the material derivative and gradient is crucial in the error analysis. Through direct computation, we obtain the following identities:

$$\partial_t^\bullet(\nabla v) = \nabla(\partial_t^\bullet v) - \nabla v \nabla w \quad \text{and} \quad \partial_t^\bullet(\nabla \cdot v) = \nabla \cdot (\partial_t^\bullet v) - \mathrm{tr}[\nabla v \nabla w]. \tag{4.5}$$

We establish a linear continuous deformation from $\Omega_h^*(t)$ to $\Omega_h(t)$. Using the basis functions $\varphi_j[\mathbf{x}(0)]$ of

$S_h^k(\Omega_h^0)^d$, we uniquely determine the domains $\Omega_h^*(t)$ and $\Omega_h(t)$ by vectors $\mathbf{x}^*$ and $\mathbf{x}$, respectively. The vector

$$\mathbf{x}^\theta = (1-\theta)\mathbf{x}^* + \theta\mathbf{x} = \mathbf{x}^* + \theta\mathbf{e_x}, \quad \text{for} \quad 0 \leqslant \theta \leqslant 1, \tag{4.6}$$

then defines an intermediate domain $\Omega_h^\theta(t)$ that changes continuously from $\Omega_h^*(t)$ to $\Omega_h(t)$ when the parameter θ varies over the interval $[0,1]$. For the vector $\mathbf{u}^* = (u_1^*(t),\ldots,u_M^*(t))$, we define the finite element function on $\Omega_h^\theta(t)$ as

$$u_h^{*,\theta} = \sum_{j=1}^M u_j^*(t)\varphi_j[\mathbf{x}^\theta(t)] \quad \text{for} \quad u_j^*(t) \in \mathbb{R}^d \quad \text{and} \quad 0 \leqslant \theta \leqslant 1. \tag{4.7}$$

Similarly, we can define functions $e_x^\theta,\ e_u^\theta, e_w^\theta,\ w_h^{*,\theta}$ and $(\partial_{t,*}^\bullet u_h^*)^\theta$, which are defined on intermediate domain $\Omega_h^\theta(t)$ and share the nodal vectors $\mathbf{e_x}, \mathbf{e_u}, \mathbf{e_w}, \mathbf{w}^*$ and $\dot{\mathbf{u}}^*$, respectively. These notations are consistently applied to other functions without redundant introductions.

Let $\partial_\theta^\bullet$ denote the material derivative with respect to the velocity field e_x^θ. By using the transport property, we have

$$\partial_\theta^\bullet u_h^{*,\theta} = 0 \qquad \text{and} \qquad \partial_\theta^\bullet e_u^\theta = 0. \tag{4.8}$$

Then the following relations follow from (4.8) and (4.5):

$$\partial_\theta^\bullet(\nabla e_u^\theta) = -\nabla e_u^\theta \nabla e_x^\theta \qquad \text{and} \qquad \partial_\theta^\bullet(\nabla \cdot e_u^\theta) = -\mathrm{tr}[\nabla e_u^\theta \nabla e_x^\theta], \tag{4.9}$$

$$\partial_\theta^\bullet(\nabla u_h^{*,\theta}) = -\nabla u_h^{*,\theta} \nabla e_x^\theta \quad \text{and} \quad \partial_\theta^\bullet(\nabla \cdot u_h^{*,\theta}) = -\mathrm{tr}[\nabla u_h^{*,\theta} \nabla e_x^\theta]. \tag{4.10}$$

We denote by ρ_h^θ and μ_h^θ the piecewise constant functions in $\Omega_h^\theta(t)$ defined by

$$\rho_h^\theta = \begin{cases} \rho_+ & \text{in } \Omega_{h,+}^\theta(t), \\ \rho_- & \text{in } \Omega_{h,-}^\theta(t), \end{cases} \quad \text{and} \quad \mu_h^\theta = \begin{cases} \mu_+ & \text{in } \Omega_{h,+}^\theta(t), \\ \mu_- & \text{in } \Omega_{h,-}^\theta(t). \end{cases} \tag{4.11}$$

In combination with the fundamental theorem of calculus, Lemma 4.1 and the transport property (2.9), the following lemma was proved in [15, Lemma 5.1].

Lemma 4.2 In the above setting the following identities hold (for the corresponding finite element functions and the associated nodal vectors):

$$(\mathbf{u}^*)^\top(\mathbf{M}(\mathbf{x}) - \mathbf{M}(\mathbf{x}^*))\mathbf{v}^* = \int_0^1 \int_{\Omega_h^\theta(t)} u_h^{*,\theta}(\nabla \cdot e_x^\theta)v_h^{*,\theta}\,\mathrm{d}x\mathrm{d}\theta, \tag{4.12}$$

$$(\mathbf{u}^*)^\top(\mathbf{A}(\mathbf{x}) - \mathbf{A}(\mathbf{x}^*))\mathbf{v}^* = \int_0^1 \int_{\Omega_h^\theta(t)} \nabla u_h^{*,\theta}(D_{\Omega_h^\theta}e_x^\theta)\cdot\nabla v_h^{*,\theta}\,\mathrm{d}x\mathrm{d}\theta, \tag{4.13}$$

$$(\mathbf{u}^*)^\top(\mathbf{M}_{\boldsymbol{\rho}}(\mathbf{x}) - \mathbf{M}_{\boldsymbol{\rho}}(\mathbf{x}^*))\mathbf{v}^* = \int_0^1 \int_{\Omega_h^\theta(t)} \rho_h^\theta u_h^{*,\theta}(\nabla \cdot e_x^\theta)v_h^{*,\theta}\,\mathrm{d}x\mathrm{d}\theta, \tag{4.14}$$

$$(\mathbf{u}^*)^\top(\mathbf{A}_{\boldsymbol{\mu}}(\mathbf{x}) - \mathbf{A}_{\boldsymbol{\mu}}(\mathbf{x}^*))\mathbf{v}^* = 2\int_0^1 \int_{\Omega_h^\theta(t)} \mu_h^\theta\Big[\mathbb{D}(u_h^{*,\theta})(\nabla \cdot e_x^\theta)\cdot\mathbb{D}(v_h^{*,\theta}) \tag{4.15}$$

$$- \mathbb{S}(\nabla u_h^{*,\theta}\nabla e_x^\theta)\cdot\mathbb{D}(v_h^{*,\theta}) - \mathbb{D}(u_h^{*,\theta})\cdot\mathbb{S}(\nabla v_h^{*,\theta}\nabla e_x^\theta)\Big]\,\mathrm{d}x\mathrm{d}\theta,$$

$$(\mathbf{u}^*)^\top(\mathbf{C}(\mathbf{x}) - \mathbf{C}(\mathbf{x}^*))^\top\mathbf{p}^* = \int_0^1 \int_{\Omega_h^\theta(t)} \Big[(\nabla \cdot u_h^{*,\theta})(\nabla \cdot e_x^\theta) - \mathrm{tr}(\nabla u_h^{*,\theta}\nabla e_x^\theta)\Big]p_h^{*,\theta}\,\mathrm{d}x\mathrm{d}\theta, \tag{4.16}$$

where $D_{\Omega_h^\theta}e_x^\theta = \nabla \cdot e_x^\theta - 2\mathbb{S}(\nabla e_x^\theta)$ with $\mathbb{S}(E) = \frac{1}{2}(E + E^\top)$.

A direct consequence of Lemma (4.2) is the following conditional equivalence of norms, i.e., if ∇e_x^θ is small, the norms of finite element functions on the two domains $\Omega_h^*(t)$ and $\Omega_h(t)$ with same nodal vectors are equivalent. This lemma was proved in [15, Lemma 5.2] by using the Gronwall's inequality.

Lemma 4.3 If $\|\nabla \cdot e_x^\theta\|_{L^\infty(\Omega_h^\theta(t))} \leqslant \alpha$ for $0 \leqslant \theta \leqslant 1$, then

$$\|\mathbf{v}\|_{\mathbf{M}(\mathbf{x}^\theta)} \leqslant e^{\alpha/2}\|\mathbf{v}\|_{\mathbf{M}(\mathbf{x}^*)}.$$

If $\|D_{\Omega_h^\theta} e_x^\theta\|_{L^\infty(\Omega_h^\theta(t))} \leqslant \eta$ for $0 \leqslant \theta \leqslant 1$, then

$$\|\mathbf{v}\|_{\mathbf{A}(\mathbf{x}^\theta)} \leqslant e^{\eta/2}\|\mathbf{v}\|_{\mathbf{A}(\mathbf{x}^*)}.$$

REMARK 4.2 (Norm equivalence) By using the Poincaré's inequality, the norms $\|\cdot\|_{\mathbf{A}(\mathbf{x}^*)}$ and $\|\cdot\|_{\mathbf{K}(\mathbf{x}^*)}$ are equivalent for finite element functions in $H_0^1(\Omega_h^*(t))$, and the equivalence is independent of h since there is a one-to-one $W^{1,\infty}$-uniformly bounded lift map from $\Omega_h^*(t)$ onto $\Omega_h(t)$.

In combination with the above Lemmas, the definition of the norms in (3.21)–(3.23) implies the following conclusion under the condition in Lemma 4.3:

$$\begin{aligned} &\text{the norms } \|\cdot\|_{\mathbf{M}(\mathbf{x}^\theta)} \text{ are } h\text{-uniformly equivalent for } 0 \leqslant \theta \leqslant 1, \\ &\text{and so are the norms } \|\cdot\|_{\mathbf{A}(\mathbf{x}^\theta)} \text{ and } \|\cdot\|_{\mathbf{K}(\mathbf{x}^\theta)}. \end{aligned} \tag{4.17}$$

Then by using Korn's inequality, the estimate (3.4), Lemma 4.3 and the equivalence (4.17), the following norms are h-uniformly equivalent when h is sufficiently small for $0 \leqslant \theta \leqslant 1$ under the condition in Lemma 4.3:

$$\|\cdot\|_{\mathbf{M}(\mathbf{x}^\theta)} \sim \|\cdot\|_{\mathbf{M}_{\boldsymbol{\rho}}(\mathbf{x}^\theta)} \qquad \text{and} \qquad \|\cdot\|_{\mathbf{K}(\mathbf{x}^\theta)} \sim \|\cdot\|_{\mathbf{A}(\mathbf{x}^\theta)} \sim \|\cdot\|_{\mathbf{A}_{\boldsymbol{\mu}}(\mathbf{x}^\theta)}. \tag{4.18}$$

The following lemma was proved in [15, Lemma 5.3], which says that the condition for e_x^θ in Lemma 4.3 can be reduced to $\theta = 0$.

Lemma 4.4 Let $e_x^* = e_x^\theta|_{\theta=0}$ be the finite element error function on $\Omega_h^*(t)$ with nodal vectors $\mathbf{e_x}$. If $\|\nabla e_x^*\|_{L^\infty(\Omega_h^*(t))} \leqslant \frac{1}{2}$, then the finite element function $v_h^{*,\theta}(t)$ defined on $\Omega_h^\theta(t)$, with $0 \leqslant \theta \leqslant 1$, satisfies the following estimate:

$$\|\nabla v_h^{*,\theta}\|_{L^p(\Omega_h^\theta(t))} \leqslant C_p \|\nabla v_h^*\|_{L^p(\Omega_h^*(t))} \quad \text{for} \quad 1 \leqslant p \leqslant \infty,$$

where C_p depends only on p.

We also need results that bound the time derivatives of the mass and stiffness matrices. The following results can be shown in a similar way as the analogous results on a surface [30, Lemma 4.6].

Lemma 4.5 For the nodal vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^{dM}$, there holds

$$\mathbf{u}^\top \left(\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{M}(\mathbf{x}^*)\right)\mathbf{v} \leqslant C\|\mathbf{u}\|_{\mathbf{M}(\mathbf{x}^*)}\|\mathbf{v}\|_{\mathbf{M}(\mathbf{x}^*)}, \tag{4.19}$$

$$\mathbf{u}^\top \left(\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{A}(\mathbf{x}^*)\right)\mathbf{v} \leqslant C\|\mathbf{u}\|_{\mathbf{A}(\mathbf{x}^*)}\|\mathbf{v}\|_{\mathbf{A}(\mathbf{x}^*)}, \tag{4.20}$$

$$\mathbf{u}^\top \left(\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{A}_{\boldsymbol{\mu}}(\mathbf{x}^*)\right)\mathbf{v} \leqslant C\|\mathbf{u}\|_{\mathbf{A}(\mathbf{x}^*)}\|\mathbf{v}\|_{\mathbf{A}(\mathbf{x}^*)}, \tag{4.21}$$

where C depends only on a bound of the $W^{1,\infty}$ norm of the domain velocity w_h^*.

4.2 *The proof of Theorem 4.1*

Let $e_x^* \in S_h^k[\mathbf{x}^*(t)]^d$, $e_u^* \in X_h[\mathbf{x}^*(t)]$, $e_p^* \in M_h[\mathbf{x}^*(t)]$ and $e_w^* \in X_{h,\pm}[\mathbf{x}^*(t)](e_u^*)$ be the finite element error functions on $\Omega_h^*(t)$ with nodal vectors $\mathbf{e_x}$, $\mathbf{e_u}$, $\mathbf{e_p}$ and $\mathbf{e_w}$, respectively.

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

$$\|e_x^*(t)\|_{W^{1,\infty}(\Omega_h^*(t))} \leqslant h^{-\frac{d}{4}} h^{\frac{k}{2}}, \tag{4.22a}$$

$$\|e_u^*(t)\|_{W^{1,\infty}(\Omega_h^*(t))} \leqslant 1, \tag{4.22b}$$

$$\|e_w^*(t)\|_{W^{1,\infty}(\Omega_h^*(t))} \leqslant 1. \tag{4.22c}$$

A key issue in the proof is to ensure that the $W^{1,\infty}$ norm of the position error $e_x^*(t)$ remains small. Since (4.22a) is small only for $k \geqslant 2$, the estimates in this proof cannot be derived for the linear case $k = 1$. In fact, t^* is positive because the inequalities above hold for $t = 0$ when h is smaller than some sufficiently small constant.

(1) Since $\mathbf{x}^*(0) = \mathbf{x}^0$ and $\mathbf{u}^*(0) = \mathbf{u}^0$, it follows that $\mathbf{e_x}(0) = 0$ and $\mathbf{e_u}(0) = 0$, which gives $e_x^*(0) = 0$ and $e_u^*(0) = 0$, respectively.

(2) Testing (3.27c) with $\mathbf{e_w} - \mathbf{e_u}$ at $t = 0$ and using the fact that $\mathbf{e_x}(0) = \mathbf{e_u}(0) = 0$ yields the following relation:

$$\|\mathbf{e_w}(0)\|_{\mathbf{A}(\mathbf{x}^*(0))}^2 = -\mathbf{e_w}(0)^\top \mathscr{E}_\mathbf{w}(0).$$

The relation above, together with Lemma 3.1, implies that $\|\mathbf{e_w}(0)\|_{\mathbf{K}(\mathbf{x}^*)} \leqslant Ch^k$. By the inverse inequality of finite element function, we have

$$\|e_w^*(0)\|_{W^{1,\infty}(\Omega_h^*(t))} \leqslant Ch^{-\frac{d}{2}} \|e_w^*(0)\|_{H^1(\Omega_h^*(t))} \leqslant Ch^{-\frac{d}{2}} \|\mathbf{e_w}(0)\|_{\mathbf{K}(\mathbf{x}^*)} \leqslant Ch^{k-\frac{d}{2}} \leqslant 1.$$

Since $k \geqslant 2 > d/2$, it follows that (4.22b)–(4.22c) hold for $t = 0$ when h is sufficiently small.

We first prove that error estimate (4.3) holds for $t \in [0, t^*]$ under condition (4.22). Then we complete the proof by showing that $t^* = T$.

(A) *Estimates for* $\mathbf{e_x}$: Multiplying (3.27d) by matrix $\mathbf{A}(\mathbf{x}^*)$, testing with $\mathbf{e_x}$, and dropping the omnipresent argument $t \in [0, t^*]$, we have

$$(\mathbf{e_x})^\top \mathbf{A}(\mathbf{x}^*)\dot{\mathbf{e}}_\mathbf{x} = (\mathbf{e_x})^\top \mathbf{A}(\mathbf{x}^*)\mathbf{e_w}. \tag{4.23}$$

In order to apply Gronwall's inequality, we relate $\frac{d}{dt}\|\mathbf{e_x}\|_{\mathbf{A}(\mathbf{x}^*)}^2$ to $(\mathbf{e_x})^\top \mathbf{A}(\mathbf{x}^*)\dot{\mathbf{e}}_\mathbf{x}$ as follows

$$\frac{1}{2}\frac{d}{dt}\|\mathbf{e_x}\|_{\mathbf{A}(\mathbf{x}^*)}^2 = \frac{1}{2}\frac{d}{dt}\big((\mathbf{e_x})^\top \mathbf{A}(\mathbf{x}^*)\mathbf{e_x}\big) = (\mathbf{e_x})^\top \mathbf{A}(\mathbf{x}^*)\dot{\mathbf{e}}_\mathbf{x} + \frac{1}{2}(\mathbf{e_x})^\top \Big(\frac{d}{dt}\mathbf{A}(\mathbf{x}^*)\Big)\mathbf{e_x}. \tag{4.24}$$

By using Lemma 4.5, we have

$$\frac{1}{2}\frac{d}{dt}\|\mathbf{e_x}\|_{\mathbf{A}(\mathbf{x}^*)}^2 = (\mathbf{e_x})^\top \mathbf{A}(\mathbf{x}^*)\mathbf{e_w} + \frac{1}{2}(\mathbf{e_x})^\top \Big(\frac{d}{dt}\mathbf{A}(\mathbf{x}^*)\Big)\mathbf{e_x} \tag{4.25}$$

$$\leqslant \int_{\Omega_h^*(t)} \nabla e_w^* \cdot \nabla e_x^* \, dx + C\|\mathbf{e_x}\|_{\mathbf{A}(\mathbf{x}^*)}^2$$

$$\leqslant C\|\nabla e_w^*\|_{L^2(\Omega_h^*(t))}\|\nabla e_x^*\|_{L^2(\Omega_h^*(t))} + C\|\mathbf{e_x}\|_{\mathbf{A}(\mathbf{x}^*)}^2$$

$$\leqslant C\|\mathbf{e_w}\|_{\mathbf{A}(\mathbf{x}^*)}^2 + C\|\mathbf{e_x}\|_{\mathbf{A}(\mathbf{x}^*)}^2.$$

(B) *Estimates for* $\mathbf{e_w}$ *under condition* (4.22): Since $e_w^* - e_u^* = 0$ on $\Gamma_h^*(t)$, i.e., $e_w^* - e_u^* \in \mathring{X}_h[\mathbf{x}^*(t)]$, we can choose $\boldsymbol{\chi} = \mathbf{e_w} - \mathbf{e_u}$ in (3.27c). This yields the following result:

$$\|\mathbf{e_w}\|_{\mathbf{A}(\mathbf{x}^*)}^2 = -(\mathbf{e_w} - \mathbf{e_u})^\top (\mathbf{A}(\mathbf{x}) - \mathbf{A}(\mathbf{x}^*))\mathbf{w} + (\mathbf{e_u})^\top \mathbf{A}(\mathbf{x}^*)\mathbf{e_w} \tag{4.26}$$

$$- (\mathbf{e_w} - \mathbf{e_u})^\top \mathscr{E}_\mathbf{w}$$

$$=: \mathscr{H}_1 + \mathscr{H}_2 + \mathscr{H}_3.$$

The first term on the right hand side of (4.26) can be estimated by using Lemma 4.2 and Hölder's inequality, i.e.,

$$\mathscr{H}_1 \leqslant C\|\nabla w_h^\theta\|_{L^\infty(\Omega_h^\theta(t))}\|\nabla e_x^\theta\|_{L^2(\Omega_h^\theta(t))}\|\nabla e_w^\theta - \nabla e_u^\theta\|_{L^2(\Omega_h^\theta(t))}$$

$$\leqslant C\big(1 + \|\nabla e_w^*\|_{L^\infty(\Omega_h^*(t))}\big)\|\nabla e_x^*\|_{L^2(\Omega_h^*(t))}\big(\|\nabla e_w^*\|_{L^2(\Omega_h^*(t))} + \|\nabla e_u^*\|_{L^2(\Omega_h^*(t))}\big)$$

$$\leqslant C\|\mathbf{e_x}\|_{\mathbf{A}(\mathbf{x}^*)}\big(\|\mathbf{e_w}\|_{\mathbf{A}(\mathbf{x}^*)} + \|\mathbf{e_u}\|_{\mathbf{A}(\mathbf{x}^*)}\big)$$

$$\leqslant C\|\mathbf{e_x}\|_{\mathbf{A}(\mathbf{x}^*)}^2 + C\|\mathbf{e_u}\|_{\mathbf{A}(\mathbf{x}^*)}^2 + \frac{1}{8}\|\mathbf{e_w}\|_{\mathbf{A}(\mathbf{x}^*)}^2,$$

where we have used $\|\nabla w_h^\theta\|_{L^\infty(\Omega_h^\theta(t))} = \|\nabla e_w^\theta - \nabla w_h^{*,\theta}\|_{L^\infty(\Omega_h^\theta(t))} \leqslant C\big(1 + \|\nabla e_w^*\|_{L^\infty(\Omega_h^*(t))}\big).$

The second term on the right hand side of (4.26) can be bounded by

$$\mathscr{H}_2 \leqslant C\|\mathbf{e_u}\|_{\mathbf{A(x^*)}}\|\mathbf{e_w}\|_{\mathbf{A(x^*)}} \leqslant C\|\mathbf{e_u}\|_{\mathbf{A(x^*)}}^2 + \frac{1}{8}\|\mathbf{e_w}\|_{\mathbf{A(x^*)}}^2.$$

The last term on the right hand side of (4.26) can be estimated similarly, i.e.,

$$\mathscr{H}_3 \leqslant Ch^k\|\mathbf{e_w} - \mathbf{e_u}\|_{\mathbf{K(x^*)}} \leqslant Ch^k\|\mathbf{e_w} - \mathbf{e_u}\|_{\mathbf{A(x^*)}} \tag{4.27}$$

$$\leqslant Ch^{2k} + \frac{1}{8}\|\mathbf{e_u}\|_{\mathbf{A(x^*)}}^2 + \frac{1}{8}\|\mathbf{e_w}\|_{\mathbf{A(x^*)}}^2,$$

where we have converted $\|\mathbf{e_w} - \mathbf{e_u}\|_{\mathbf{K(x^*)}}$ (i.e., the full H^1 norm) to $\|\mathbf{e_w} - \mathbf{e_u}\|_{\mathbf{A(x^*)}}$ (i.e., the H^1 semi-norm) because the corresponding function $e_w^* - e_u^*$ satisfies the zero boundary condition in $\Omega_{h,\pm}^*(t)$. Altogether we obtain the bound as

$$\|\mathbf{e_w}\|_{\mathbf{A(x^*)}}^2 \leqslant Ch^{2k} + C\|\mathbf{e_x}\|_{\mathbf{A(x^*)}}^2 + C\|\mathbf{e_u}\|_{\mathbf{A(x^*)}}^2. \tag{4.28}$$

Then, substituting (4.28) into (4.25), we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\|\mathbf{e_x}\|_{\mathbf{A(x^*)}}^2 \leqslant Ch^{2k} + C\|\mathbf{e_u}\|_{\mathbf{A(x^*)}}^2 + C\|\mathbf{e_x}\|_{\mathbf{A(x^*)}}^2. \tag{4.29}$$

By integrating (4.29) on both sides with respect to time from 0 to t, we obtain

$$\|\mathbf{e_x}\|_{\mathbf{A(x^*)}}^2 \leqslant Ch^{2k} + \int_0^t \Big(\|\mathbf{e_x}(s)\|_{\mathbf{A(x^*)}}^2 + \|\mathbf{e_u}(s)\|_{\mathbf{A(x^*)}}^2\Big)\,\mathrm{d}s. \tag{4.30}$$

(C) *Estimates for* $\mathbf{e_u}$ *under condition* (4.22): Testing (3.27a) by $\dot{\mathbf{e}}_\mathbf{u}$, we obtain

$$(\dot{\mathbf{e}}_\mathbf{u})^\top \mathbf{M_\rho(x)}\dot{\mathbf{e}}_\mathbf{u} + \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\Big((\mathbf{e_u})^\top \mathbf{A_\mu(x^*)}\mathbf{e_u}\Big) - \frac{1}{2}(\mathbf{e_u})^\top \Big(\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{A_\mu(x^*)}\Big)\mathbf{e_u} \tag{4.31}$$

$$= -(\dot{\mathbf{e}}_\mathbf{u})^\top (\mathbf{M_\rho(x)} - \mathbf{M_\rho(x^*)})\dot{\mathbf{u}}^*$$

$$\quad -(\dot{\mathbf{e}}_\mathbf{u})^\top (\mathbf{A_\mu(x)} - \mathbf{A_\mu(x^*)})\mathbf{u}^*$$

$$\quad -(\dot{\mathbf{e}}_\mathbf{u})^\top (\mathbf{A_\mu(x)} - \mathbf{A_\mu(x^*)})\mathbf{e_u}$$

$$\quad -(\dot{\mathbf{e}}_\mathbf{u})^\top \Big(\mathbf{B_\rho(x, u - w)u} - \mathbf{B_\rho(x^*, u^* - w^*)u^*}\Big)$$

$$\quad +(\dot{\mathbf{e}}_\mathbf{u})^\top (\mathbf{C(x)}^\top - \mathbf{C(x^*)}^\top)\mathbf{p}^* + (\dot{\mathbf{e}}_\mathbf{u})^\top \mathbf{C(x)}^\top \mathbf{e_p}$$

$$\quad +(\dot{\mathbf{e}}_\mathbf{u})^\top (\mathbf{M(x)f} - \mathbf{M(x^*)f^*}) - (\dot{\mathbf{e}}_\mathbf{u})^\top \mathscr{E}_{\mathbf{u1}}$$

$$=: \sum_{j=1}^{7} S_j - (\dot{\mathbf{e}}_\mathbf{u})^\top \mathscr{E}_{\mathbf{u1}}.$$

The last term on the left-hand side of (4.31) can be estimated by using Lemma 4.5, i.e.,

$$\left|\frac{1}{2}(\mathbf{e_u})^\top \Big(\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{A_\mu(x^*)}\Big)\mathbf{e_u}\right| \leqslant C\|\mathbf{e_u}\|_{\mathbf{A(x^*)}}^2. \tag{4.32}$$

In the following, we present estimates for each term on the right-hand side of (4.31).

The first term on the right hand side of (4.31) can be estimated by using Lemma 4.2 and Hölder's inequality, and the norm equivalence in Lemma 4.3, i.e.,

$$S_1 = (\dot{\mathbf{e}}_\mathbf{u})^\top (\mathbf{M_\rho(x)} - \mathbf{M_\rho(x^*)})\dot{\mathbf{u}}^* \tag{4.33}$$

$$= \int_0^1 \int_{\Omega_h^\theta(t)} \rho_h^\theta (\partial_{t,*}^\bullet e_u^*)^\theta (\nabla \cdot e_x^\theta)(\partial_{t,*}^\bullet u_h^*)^\theta \,\mathrm{d}x\mathrm{d}\theta$$

$$\leqslant C\|\nabla \cdot e_x^\theta\|_{L^2(\Omega_h^\theta(t))}\|(\partial_{t,*}^\bullet u_h^*)^\theta\|_{L^\infty(\Omega_h^\theta(t))}\|(\partial_{t,*}^\bullet e_u^*)^\theta\|_{L^2(\Omega_h^\theta(t))}$$

$$\leqslant C\|\nabla e_x^*\|_{L^2(\Omega_h^*(t))}\|\partial_{t,*}^\bullet u_h^*\|_{L^\infty(\Omega_h^*(t))}\|\partial_{t,*}^\bullet e_u^*\|_{L^2(\Omega_h^*(t))}$$

$$\leqslant C\|\mathbf{e_x}\|_{\mathbf{A(x^*)}}\|\dot{\mathbf{e}}_\mathbf{u}\|_{\mathbf{M(x)}}$$

$$\leqslant \varepsilon \|\dot{\mathbf{e}}_\mathbf{u}\|^2_{\mathbf{M}_{\boldsymbol{\rho}}(\mathbf{x})} + C\varepsilon^{-1}\|\mathbf{e}_\mathbf{x}\|^2_{\mathbf{A}(\mathbf{x}^*)},$$

where the positive constant ε can be arbitrarily small and we have used $\|\partial^\bullet_{t,*} u^*_h\|_{L^\infty(\Omega^*_h(t))} \leqslant C$.

The second term on the right hand side of (4.31) can be estimated by using integration by parts in time, i.e.,

$$
\begin{aligned}
S_2 &= -(\dot{\mathbf{e}}_\mathbf{u})^\top (\mathbf{A}_{\boldsymbol{\mu}}(\mathbf{x}) - \mathbf{A}_{\boldsymbol{\mu}}(\mathbf{x}^*))\mathbf{u}^* \\
&= -\frac{\mathrm{d}}{\mathrm{d}t}\Big((\mathbf{e}_\mathbf{u})^\top (\mathbf{A}_{\boldsymbol{\mu}}(\mathbf{x}) - \mathbf{A}_{\boldsymbol{\mu}}(\mathbf{x}^*))\mathbf{u}^* \Big) + (\mathbf{e}_\mathbf{u})^\top \Big(\frac{\mathrm{d}}{\mathrm{d}t}(\mathbf{A}_{\boldsymbol{\mu}}(\mathbf{x}) - \mathbf{A}_{\boldsymbol{\mu}}(\mathbf{x}^*))\Big)\mathbf{u}^* \\
&\quad + (\mathbf{e}_\mathbf{u})^\top (\mathbf{A}_{\boldsymbol{\mu}}(\mathbf{x}) - \mathbf{A}_{\boldsymbol{\mu}}(\mathbf{x}^*))\dot{\mathbf{u}}^* \\
&=: S_{21} + S_{22} + S_{23},
\end{aligned}
\tag{4.34}
$$

where the first term on the right-hand side of (4.34) can be bounded after integrating it over $[0,t]$ for $0 \leqslant t \leqslant t^*$, i.e.,

$$
\begin{aligned}
\int_0^t S_{21}(s)\,\mathrm{d}s &= -(\mathbf{e}_\mathbf{u})^\top (\mathbf{A}_{\boldsymbol{\mu}}(\mathbf{x}) - \mathbf{A}_{\boldsymbol{\mu}}(\mathbf{x}^*))\mathbf{u}^* \\
&= -2\int_0^1 \int_{\Omega^\theta_h(t)} \mu^\theta_h \Big[\mathbb{D}(u^{*,\theta}_h)(\nabla \cdot e^\theta_x)\cdot \mathbb{D}(e^\theta_u) \quad \text{(Lemma 4.2 is used)} \\
&\qquad\qquad\qquad\qquad - \mathbb{S}(\nabla u^{*,\theta}_h \nabla e^\theta_x)\cdot \mathbb{D}(e^\theta_u) + \mathbb{D}(u^{*,\theta}_h)\cdot \mathbb{S}(\nabla e^\theta_u \nabla e^\theta_x) \Big]\,\mathrm{d}x\mathrm{d}\theta \\
&\leqslant C\|\nabla u^*_h(t)\|_{L^\infty(\Omega^*_h(t))} \|\nabla e^*_x(t)\|_{L^2(\Omega^*_h(t))}\|\nabla e^*_u(t)\|_{L^2(\Omega^*_h(t))} \quad \text{(Lemma 4.3 is used)} \\
&\leqslant C\|\mathbf{e}_\mathbf{x}(t)\|_{\mathbf{A}(\mathbf{x}^*)}\|\mathbf{e}_\mathbf{u}(t)\|_{\mathbf{A}(\mathbf{x}^*)} \\
&\leqslant \varepsilon\|\mathbf{e}_\mathbf{u}(t)\|^2_{\mathbf{A}_{\boldsymbol{\mu}}(\mathbf{x}^*)} + C\varepsilon^{-1}\|\mathbf{e}_\mathbf{x}(t)\|^2_{\mathbf{A}(\mathbf{x}^*)}.
\end{aligned}
\tag{4.35}
$$

In order to estimate S_{22}, we define $w^\theta_h = w^{*,\theta}_h + \theta e^\theta_w$ to be the finite element function on $\Omega^\theta_h(t)$ with nodal vector $\mathbf{w}^\theta = \mathbf{w}^* + \theta \mathbf{e}_\mathbf{w}$. Note that $\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x}^\theta = \dot{\mathbf{x}} = \mathbf{w}^\theta$ and $\frac{\mathrm{d}}{\mathrm{d}\theta}\mathbf{x}^\theta = \mathbf{e}_\mathbf{x}$. By using the transport formula, Lemma 4.1 and (4.5), we have

$$
\frac{\mathrm{d}}{\mathrm{d}t}\int_{\Omega^\theta_h(t)} f\mathrm{d}x = \int_{\Omega^\theta_h(t)} (\partial^\bullet_{t,\theta} f + f\nabla \cdot w^\theta_h)\mathrm{d}x,
$$

$$
\frac{\mathrm{d}}{\mathrm{d}\theta}\int_{\Omega^\theta_h(t)} f\mathrm{d}x = \int_{\Omega^\theta_h(t)} (\partial^\bullet_\theta f + f\nabla \cdot e^\theta_x)\mathrm{d}x,
$$

where $\partial^\bullet_\theta$ denotes the material derivative with respect to the velocity field e^θ_x with

$$
\partial^\bullet_\theta (\nabla w^\theta_h) = \nabla e^\theta_w - \nabla w^\theta_h \nabla e^\theta_x \quad \text{and} \quad \partial^\bullet_\theta (\nabla \cdot w^\theta_h) = \nabla \cdot e^\theta_w - \mathrm{tr}[\nabla w^\theta_h \nabla e^\theta_x],
\tag{4.36}
$$

and $\partial^\bullet_{t,\theta}$ denotes the material derivative with respect to the velocity field w^θ_h with

$$
\partial^\bullet_{t,\theta} \mathbb{D}(\eta^\theta_h) = -\mathbb{S}(\nabla \eta^\theta_h \nabla e^\theta_x) \qquad \text{for} \qquad \partial^\bullet_{t,\theta} \eta^\theta_h = 0,
\tag{4.37}
$$

$$
\partial^\bullet_{t,\theta} \mathbb{S}(\nabla \eta^\theta_h \nabla e^\theta_x) = \mathbb{S}(\nabla \eta^\theta_h \nabla e^\theta_w) - \mathbb{S}(\nabla \eta^\theta_h (\nabla w^\theta_h \nabla e^\theta_x + \nabla e^\theta_x \nabla w^\theta_h)),
\tag{4.38}
$$

where we note $\partial^\bullet_{t,\theta} e^\theta_x = e^\theta_w$ and η^θ_h is defined from (4.7), which is the finite element function on $\Omega^\theta_h(t)$ and shares the nodal vector $\boldsymbol{\eta} \in \mathbb{R}^{dM}$ with $\dot{\boldsymbol{\eta}} = 0$.

The second term on the right hand side of (4.34) can be estimated by using the above-mentioned setting and Lemma 4.2, we have

$$
\begin{aligned}
|S_{22}| &= \Big| (\mathbf{e}_\mathbf{u})^\top \Big(\frac{\mathrm{d}}{\mathrm{d}t}(\mathbf{A}_{\boldsymbol{\mu}}(\mathbf{x}) - \mathbf{A}_{\boldsymbol{\mu}}(\mathbf{x}^*))\Big)\mathbf{u}^* \Big| \\
&= \Big| \frac{\mathrm{d}}{\mathrm{d}t}\int_0^1 \int_{\Omega^\theta_h(t)} 2\mu^\theta_h \Big[\mathbb{D}(u^{*,\theta}_h)(\nabla \cdot e^\theta_x)\cdot \mathbb{D}(e^\theta_u) \\
&\qquad\qquad\qquad - \mathbb{S}(\nabla u^{*,\theta}_h \nabla e^\theta_x)\cdot \mathbb{D}(e^\theta_u) - \mathbb{D}(u^{*,\theta}_h)\cdot \mathbb{S}(\nabla e^\theta_u \nabla e^\theta_x) \Big]\mathrm{d}x\mathrm{d}\theta \Big|
\end{aligned}
\tag{4.39}
$$

$$= \Big| \int_0^1 \int_{\Omega_h^\theta(t)} 2\mu_h^\theta \Big[-\mathbb{S}(u_h^{*,\theta} w_h^\theta)(\nabla \cdot e_x^\theta) \cdot \mathbb{D}(e_u^\theta) - \mathbb{D}(u_h^{*,\theta})(\nabla \cdot e_x^\theta) \cdot \mathbb{S}(e_u^\theta w_h^\theta)$$

$$+ \mathbb{D}(u_h^{*,\theta})\big(\nabla \cdot e_w^\theta - \text{tr}[\nabla e_x^\theta \nabla w_h^\theta]\big) \cdot \mathbb{D}(e_u^\theta) \Big] \mathrm{d}x \mathrm{d}\theta$$

$$+ \int_0^1 \int_{\Omega_h^\theta(t)} 2\mu_h^\theta \Big[\big(\mathbb{S}(\nabla u_h^{*,\theta} \nabla e_w^\theta) - \mathbb{S}(\nabla u_h^{*,\theta}(\nabla e_x^\theta \nabla w_h^\theta + \nabla w_h^\theta \nabla e_x^\theta))\big) \cdot \mathbb{D}(e_u^\theta)$$

$$+ \mathbb{D}(u_h^{*,\theta}) \cdot \big(\mathbb{S}(\nabla e_u^\theta \nabla e_w^\theta) - \mathbb{S}(\nabla e_u^\theta(\nabla e_x^\theta \nabla w_h^\theta + \nabla w_h^\theta \nabla e_x^\theta))\big)$$

$$- \mathbb{S}(\nabla u_h^{*,\theta} \nabla e_x^\theta) \cdot \mathbb{S}(\nabla e_u^\theta \nabla w_h^\theta) - \mathbb{S}(u_h^{*,\theta} \nabla w_h^\theta) \cdot \mathbb{S}(\nabla e_u^\theta \nabla e_x^\theta) \Big] \mathrm{d}x \mathrm{d}\theta$$

$$+ \int_0^1 \int_{\Omega_h^\theta(t)} 2\mu_h^\theta \Big[\mathbb{D}(u_h^{*,\theta})(\nabla \cdot e_x^\theta) \cdot \mathbb{D}(e_u^\theta)$$

$$- \mathbb{S}(\nabla u_h^{*,\theta} \nabla e_x^\theta) \cdot \mathbb{D}(e_u^\theta) - \mathbb{D}(u_h^{*,\theta}) \cdot \mathbb{S}(\nabla e_u^\theta \nabla e_x^\theta) \Big] (\nabla \cdot w_h^\theta) \, \mathrm{d}x \mathrm{d}\theta \Big|$$

$$\leqslant C \|\nabla u_h^*\|_{L^\infty(\Omega_h^*(t))} \big(1 + \|\nabla e_w^*\|_{L^\infty(\Omega_h^*(t))}\big) \|\nabla e_x^*\|_{L^2(\Omega_h^*(t))} \|\nabla e_u^*\|_{L^2(\Omega_h^*(t))}$$

$$+ C \|\nabla u_h^*\|_{L^\infty(\Omega_h^*(t))} \big(1 + \|\nabla e_w^*\|_{L^\infty(\Omega_h^*(t))}\big) \|\nabla e_w^*\|_{L^2(\Omega_h^*(t))} \|\nabla e_u^*\|_{L^2(\Omega_h^*(t))}$$

$$\leqslant C \big(\|\nabla e_w^*\|_{L^2(\Omega_h^*(t))} + \|\nabla e_x^*\|_{L^2(\Omega_h^*(t))}\big) \|\nabla e_u^*\|_{L^2(\Omega_h^*(t))}$$

$$\leqslant C \|\mathbf{e_u}\|_{\mathbf{A(x^*)}}^2 + C \|\mathbf{e_x}\|_{\mathbf{A(x^*)}}^2 + C \|\mathbf{e_w}\|_{\mathbf{A(x^*)}}^2.$$

Similarly as the estimates of (4.35), the third term on the right hand side of (4.34) can be estimated as

$$|S_{23}| = \big|(\mathbf{e_u})^\top(\mathbf{A_\mu(x)} - \mathbf{A_\mu(x^*)})\dot{\mathbf{u}}^*\big| \tag{4.40}$$

$$\leqslant C \|\nabla \partial_{t,*}^\bullet u_h^*\|_{L^\infty(\Omega_h^*(t))} \|\nabla e_x^*\|_{L^2(\Omega_h^*(t))} \|\nabla e_u^*\|_{L^2(\Omega_h^*(t))}$$

$$\leqslant C \|\mathbf{e_u}\|_{\mathbf{A(x^*)}}^2 + C \|\mathbf{e_x}\|_{\mathbf{A(x^*)}}^2.$$

The estimates of $\int_0^t S_{21}(s)\mathrm{d}s$, S_{22} and S_{23} show that

$$\Big| \int_0^t S_2(s) \, \mathrm{d}s \Big| \tag{4.41}$$

$$\leqslant \varepsilon \|\mathbf{e_u}(t)\|_{\mathbf{A_\mu(x^*)}}^2 + C\varepsilon^{-1} \|\mathbf{e_x}(t)\|_{\mathbf{A(x^*)}}^2 + C \int_0^t \Big(\|\mathbf{e_u}(s)\|_{\mathbf{A(x^*)}}^2 + \|\mathbf{e_x}(s)\|_{\mathbf{A(x^*)}}^2 + \|\mathbf{e_w}(s)\|_{\mathbf{A(x^*)}}^2 \Big) \mathrm{d}s.$$

The third term on the right hand side of (4.31) can be written as

$$S_3 = -(\dot{\mathbf{e}}_\mathbf{u})^\top(\mathbf{A_\mu(x)} - \mathbf{A_\mu(x^*)})\mathbf{e_u} \tag{4.42}$$

$$= -\frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{1}{2}(\mathbf{e_u})^\top(\mathbf{A_\mu(x)} - \mathbf{A_\mu(x^*)})\mathbf{e_u}\Big) + \frac{1}{2}(\mathbf{e_u})^\top\Big(\frac{\mathrm{d}}{\mathrm{d}t}(\mathbf{A_\mu(x)} - \mathbf{A_\mu(x^*)})\Big)\mathbf{e_u}$$

$$=: S_{31} + S_{32},$$

where S_{31} and S_{32} can be treated similarly as S_{21} and S_{22}, with the following estimates:

$$\Big| \int_0^t S_{31}(s)\mathrm{d}s \Big| = \Big|\frac{1}{2}(\mathbf{e_u})^\top(\mathbf{A_\mu(x)} - \mathbf{A_\mu(x^*)})\mathbf{e_u}\Big| \tag{4.43}$$

$$\leqslant C \|e_x^*\|_{W^{1,\infty}(\Omega_h^*(t))} \|\mathbf{e_u}(t)\|_{\mathbf{A_\mu(x^*)}}^2$$

$$\leqslant C h^{\frac{k}{2}-\frac{d}{4}} \|\mathbf{e_u}(t)\|_{\mathbf{A_\mu(x^*)}}^2$$

$$|S_{32}| = \Big|(\mathbf{e_u})^\top\Big(\frac{\mathrm{d}}{\mathrm{d}t}(\mathbf{A_\mu(x)} - \mathbf{A_\mu(x^*)})\Big)\mathbf{e_u}\Big| \tag{4.44}$$

$$\leqslant C \|e_u^*\|_{W^{1,\infty}(\Omega_h^*(t))} \big(\|\nabla e_w^*\|_{L^2(\Omega_h^*(t))} + \|\nabla e_x^*\|_{L^2(\Omega_h^*(t))}\big) \|\nabla e_u^*\|_{L^2(\Omega_h^*(t))}$$

$$\leqslant C \|\mathbf{e_u}\|_{\mathbf{A(x^*)}}^2 + C \|\mathbf{e_x}\|_{\mathbf{A(x^*)}}^2 + C \|\mathbf{e_w}\|_{\mathbf{A(x^*)}}^2,$$

where we have used the results $\|e_x^*\|_{W^{1,\infty}(\Omega_h^*(t))} \leqslant h^{-\frac{d}{4}} h^{\frac{k}{2}}$ and $\|e_u^*\|_{W^{1,\infty}(\Omega_h^*(t))} \leqslant 1$ in (4.22). Therefore,

$$\left| \int_0^t S_3(s)\,\mathrm{d}s \right| \leqslant Ch^{\frac{k}{2}-\frac{d}{4}} \|\mathbf{e_u}(t)\|_{\mathbf{A_\mu}(\mathbf{x}^*)}^2 + C \int_0^t \left(\|\mathbf{e_u}(s)\|_{\mathbf{A}(\mathbf{x}^*)}^2 + \|\mathbf{e_x}(s)\|_{\mathbf{A}(\mathbf{x}^*)}^2 + \|\mathbf{e_w}(s)\|_{\mathbf{A}(\mathbf{x}^*)}^2 \right) \mathrm{d}s. \tag{4.45}$$

The fourth term on the right hand side of (4.31) can be estimated similarly, i.e.,

$$S_4 = - (\dot{\mathbf{e}}_\mathbf{u})^\top \left(\mathbf{B_\rho}(\mathbf{x}, \mathbf{u} - \mathbf{w})\mathbf{u} - \mathbf{B_\rho}(\mathbf{x}, \mathbf{u}^* - \mathbf{w}^*)\mathbf{u}^* \right) \tag{4.46}$$

$$- (\dot{\mathbf{e}}_\mathbf{u})^\top \left(\mathbf{B_\rho}(\mathbf{x}, \mathbf{u}^* - \mathbf{w}^*)\mathbf{u}^* - \mathbf{B_\rho}(\mathbf{x}^*, \mathbf{u}^* - \mathbf{w}^*)\mathbf{u}^* \right)$$

$$= - \int_{\Omega_h(t)} \rho_h \left[(e_u^\theta - e_w^\theta) \cdot \nabla u_h^{*,\theta} + (u_h^\theta - w_h^\theta) \cdot \nabla e_u^\theta \right] \cdot (\partial_{t,*}^\bullet e_u^*)^\theta \Big|_{\theta=1} \mathrm{d}x$$

$$+ \int_0^1 \int_{\Omega_h^\theta(t)} \rho_h^\theta \left[(u_h^{*,\theta} - w_h^{*,\theta}) \cdot (\nabla u_h^{*,\theta} \nabla e_x^\theta - \nabla u_h^{*,\theta}(\nabla \cdot e_x^\theta)) \right] \cdot (\partial_{t,*}^\bullet e_u^*)^\theta \mathrm{d}x\mathrm{d}\theta$$

$$\leqslant C \big(\|e_u^*\|_{L^2(\Omega_h^*(t))} + \|e_w^*\|_{L^2(\Omega_h^*(t))} \big) \big(\|\nabla u_h^*\|_{L^\infty(\Omega_h^*(t))} + \|\nabla e_u^*\|_{L^\infty(\Omega_h^*(t))} \big) \|\partial_{t,*}^\bullet e_u^*\|_{L^2(\Omega_h^*(t))}$$

$$+ C \|u_h^* - w_h^*\|_{L^\infty(\Omega_h^*(t))} \|\nabla u_h^*\|_{L^\infty(\Omega_h^*(t))} \|\nabla e_x^*\|_{L^2(\Omega_h^*(t))} \|\partial_{t,*}^\bullet e_u^*\|_{L^2(\Omega_h^*(t))}$$

$$\leqslant C \|\dot{\mathbf{e}}_\mathbf{u}\|_{\mathbf{M}(\mathbf{x}^*)} \big(\|\mathbf{e_u}\|_{\mathbf{A}(\mathbf{x}^*)} + \|\mathbf{e_w}\|_{\mathbf{A}(\mathbf{x}^*)} + \|\mathbf{e_x}\|_{\mathbf{A}(\mathbf{x}^*)} \big)$$

$$\leqslant \varepsilon \|\dot{\mathbf{e}}_\mathbf{u}\|_{\mathbf{M_\rho}(\mathbf{x})}^2 + C\varepsilon^{-1} \big(\|\mathbf{e_u}\|_{\mathbf{A}(\mathbf{x}^*)}^2 + \|\mathbf{e_w}\|_{\mathbf{A}(\mathbf{x}^*)}^2 + \|\mathbf{e_x}\|_{\mathbf{A}(\mathbf{x}^*)}^2 \big).$$

The fifth term on the right hand side of (4.31) can be written as

$$S_5 = (\dot{\mathbf{e}}_\mathbf{u})^\top (\mathbf{C}(\mathbf{x})^\top - \mathbf{C}(\mathbf{x}^*)^\top)\mathbf{p}^* \tag{4.47}$$

$$= \frac{\mathrm{d}}{\mathrm{d}t} \left((\mathbf{e_u})^\top (\mathbf{C}(\mathbf{x})^\top - \mathbf{C}(\mathbf{x}^*)^\top)\mathbf{p}^* \right) - (\mathbf{e_u})^\top \left(\frac{\mathrm{d}}{\mathrm{d}t}(\mathbf{C}(\mathbf{x})^\top - \mathbf{C}(\mathbf{x}^*)^\top) \right)\mathbf{p}^*$$

$$- (\mathbf{e_u})^\top (\mathbf{C}(\mathbf{x})^\top - \mathbf{C}(\mathbf{x}^*)^\top)\dot{\mathbf{p}}^*$$

$$=: S_{51} + S_{52} + S_{53},$$

where the first term on the right-hand side of (4.47) can be bounded after integrating it over $[0,t]$ for $0 \leqslant t \leqslant t^*$, i.e.,

$$\int_0^t S_{51}(s)\,\mathrm{d}s = (\mathbf{e_u})^\top (\mathbf{C}(\mathbf{x})^\top - \mathbf{C}(\mathbf{x}^*)^\top)\mathbf{p}^* \tag{4.48}$$

$$= \int_0^1 \int_{\Omega_h^\theta(t)} \left((\nabla \cdot e_x^\theta)(\nabla \cdot e_u^\theta) - \mathrm{tr}[\nabla e_u^\theta \nabla e_x^\theta] \right) p_h^{*,\theta} \mathrm{d}x\mathrm{d}\theta$$

$$\leqslant C \|p_h^*\|_{L^\infty(\Omega_h^*(t))} \|\nabla e_x^*\|_{L^2(\Omega_h^*(t))} \|\nabla e_u^*\|_{L^2(\Omega_h^*(t))}$$

$$\leqslant \varepsilon \|\mathbf{e_u}(t)\|_{\mathbf{A_\mu}(\mathbf{x}^*)}^2 + C\varepsilon^{-1} \|\mathbf{e_x}(t)\|_{\mathbf{A}(\mathbf{x}^*)}^2.$$

The second and third terms on the right-hand side of (4.47) can be treated similarly as S_{22} and S_{23}, with the following estimates:

$$|S_{52}| \leqslant C \|\mathbf{e_u}\|_{\mathbf{A}(\mathbf{x}^*)}^2 + C \|\mathbf{e_w}\|_{\mathbf{A}(\mathbf{x}^*)}^2 + C \|\mathbf{e_x}\|_{\mathbf{A}(\mathbf{x}^*)}^2, \tag{4.49}$$

$$|S_{53}| \leqslant C \|\mathbf{e_u}\|_{\mathbf{A}(\mathbf{x}^*)}^2 + C \|\mathbf{e_x}\|_{\mathbf{A}(\mathbf{x}^*)}^2. \tag{4.50}$$

The sixth term on the right-hand side of (4.31) can be estimated by differentiating (3.27b) in time. This yields the following relation:

$$\mathbf{C}(\mathbf{x})\dot{\mathbf{e}}_\mathbf{u} = - \left(\frac{\mathrm{d}}{\mathrm{d}t}(\mathbf{C}(\mathbf{x}) - \mathbf{C}(\mathbf{x}^*)) \right)\mathbf{u}^* - (\mathbf{C}(\mathbf{x}) - \mathbf{C}(\mathbf{x}^*))\dot{\mathbf{u}}^* - \left(\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{C}(\mathbf{x}) \right)\mathbf{e_u} - \frac{\mathrm{d}}{\mathrm{d}t}\mathscr{E}_{\mathbf{u}2}. \tag{4.51}$$

Testing this equation with $\mathbf{e_p}$ and using relation $\mathbf{q}^\top \mathbf{C}(\mathbf{x})\mathbf{v} = (\mathbf{C}(\mathbf{x})\mathbf{v})^\top \mathbf{q}$, we have

$$S_6 = (\mathbf{C}(\mathbf{x})\dot{\mathbf{e}}_\mathbf{u})^\top \mathbf{e_p} = -(\mathbf{u}^*)^\top \left(\frac{\mathrm{d}}{\mathrm{d}t}(\mathbf{C}(\mathbf{x})^\top - \mathbf{C}(\mathbf{x}^*)^\top) \right)\mathbf{e_p} - (\dot{\mathbf{u}}^*)^\top (\mathbf{C}(\mathbf{x})^\top - \mathbf{C}(\mathbf{x}^*)^\top)\mathbf{e_p} \tag{4.52}$$

$$- (\mathbf{e_u})^\top \left(\frac{\mathrm{d}}{\mathrm{d}t} \mathbf{C}(\mathbf{x})^\top \right) \mathbf{e_p} - (\mathbf{e_p})^\top \left(\frac{\mathrm{d}}{\mathrm{d}t} \mathscr{E}_{\mathbf{u}2} \right),$$

which can be estimated similarly as S_{22}, S_{23}, and so on. In particular, we have

$$S_6 \leqslant \varepsilon \|\mathbf{e_p}\|^2_{\mathbf{M}(\mathbf{x}^*)} + C\varepsilon^{-1} \left(\|\mathbf{e_u}\|^2_{\mathbf{A}(\mathbf{x}^*)} + \|\mathbf{e_w}\|^2_{\mathbf{A}(\mathbf{x}^*)} + \|\mathbf{e_x}\|^2_{\mathbf{A}(\mathbf{x}^*)} + h^{2k} \right). \tag{4.53}$$

The seventh term on the right hand side of (4.31) can be written as

$$S_7 = (\dot{\mathbf{e}}_\mathbf{u})^\top (\mathbf{M}(\mathbf{x})\mathbf{f} - \mathbf{M}(\mathbf{x}^*)\mathbf{f}^*) \tag{4.54}$$

$$= (f_h, (\partial^\bullet_{t,*} e^*_u)^\theta|_{\theta=1})_{\Omega_h(t)} - (f^*_h, \partial^\bullet_{t,*} e^*_u)_{\Omega^*_h(t)} \qquad (\text{here we use } (\partial^\bullet_{t,*} e^*_u)^\theta|_{\theta=0} = \partial^\bullet_{t,*} e^*_u)$$

$$= (f_h - f, (\partial^\bullet_{t,*} e^*_u)^\theta|_{\theta=1})_{\Omega_h(t)} - (f^*_h - f, \partial^\bullet_{t,*} e^*_u)_{\Omega^*_h(t)} + \left((f, (\partial^\bullet_{t,*} e^*_u)^\theta|_{\theta=1})_{\Omega_h(t)} - (f, \partial^\bullet_{t,*} e^*_u)_{\Omega^*_h(t)} \right)$$

$$=: S_{71} + S_{72} + S_{73}.$$

By using the approximation properties of the Lagrange interpolations f_h and f^*_h for f on the domains $\Omega_h(t)$ and $\Omega^*_h(t)$, respectively, the terms S_{71} and S_{72} can be estimates as

$$|S_{71}| + |S_{72}| \leqslant Ch^k \|f\|_{H^{k-1}(\Omega(t))} \|\partial^\bullet_{t,*} e^*_u\|_{L^2(\Omega^*_h(t))} \leqslant Ch^k \|\dot{\mathbf{e}}_\mathbf{u}\|_{\mathbf{M}(\mathbf{x})} \leqslant \varepsilon \|\dot{\mathbf{e}}_\mathbf{u}\|^2_{\mathbf{M}_\rho(\mathbf{x})} + C\varepsilon^{-1} h^{2k}. \tag{4.55}$$

By applying $\partial^\bullet_\theta f = \nabla f \cdot e^\theta_x$ on $\Omega^\theta_h(t)$, the term S_{73} can be bounded similarly to the estimates for (4.33) as

$$|S_{73}| = \left| \int_0^1 \int_{\Omega^\theta_h(t)} (\nabla f \cdot e^\theta_x)(\partial^\bullet_{t,*} e^*_u)^\theta \, \mathrm{d}x \mathrm{d}\theta + \int_0^1 \int_{\Omega^\theta_h(t)} f(\nabla \cdot e^\theta_x)(\partial^\bullet_{t,*} e^*_u)^\theta \, \mathrm{d}x \mathrm{d}\theta \right| \tag{4.56}$$

$$\leqslant C \left(\|e^*_x\|_{L^2(\Omega^*_h(t))} \|\nabla f\|_{L^\infty(\Omega^*_h(t))} + \|\nabla e^*_x\|_{L^2(\Omega^*_h(t))} \|f\|_{L^\infty(\Omega^*_h(t))} \right) \|\partial^\bullet_{t,*} e^*_u\|_{L^2(\Omega^*_h(t))}$$

$$\leqslant C\|\mathbf{e_x}\|_{\mathbf{A}(\mathbf{x}^*)} \|\dot{\mathbf{e}}_\mathbf{u}\|_{\mathbf{M}(\mathbf{x})}$$

$$\leqslant \varepsilon \|\dot{\mathbf{e}}_\mathbf{u}\|^2_{\mathbf{M}_\rho(\mathbf{x})} + C\varepsilon^{-1} \|\mathbf{e_x}\|^2_{\mathbf{A}(\mathbf{x}^*)}.$$

The last term on the right hand side of (4.31) can also be estimated by using integration by parts in time, i.e.,

$$-(\dot{\mathbf{e}}_\mathbf{u})^\top \mathscr{E}_{\mathbf{u}1} = -\frac{\mathrm{d}}{\mathrm{d}t} \left((\mathbf{e_u})^\top \mathscr{E}_{\mathbf{u}1} \right) + (\mathbf{e_u})^\top \left(\frac{\mathrm{d}}{\mathrm{d}t} \mathscr{E}_{\mathbf{u}1} \right) =: S_{81} + S_{82}. \tag{4.57}$$

Then, using the consistency estimates in Proposition 3.1, we have

$$\left| \int_0^t S_{81}(s) \, \mathrm{d}s \right| \leqslant Ch^k \|\mathbf{e_u}(t)\|_{\mathbf{K}(\mathbf{x}^*)} \leqslant \varepsilon \|\mathbf{e_u}(t)\|^2_{\mathbf{A}_\mu(\mathbf{x}^*)} + C\varepsilon^{-1} h^{2k}, \tag{4.58a}$$

$$|S_{82}| \leqslant Ch^k \|\mathbf{e_u}\|_{\mathbf{K}(\mathbf{x}^*)} \leqslant C\|\mathbf{e_u}\|^2_{\mathbf{A}(\mathbf{x}^*)} + Ch^{2k}. \tag{4.58b}$$

Integrating the inequality (4.31) in time from 0 to t and using the estimates of S_j, $j = 1, \ldots, 6$, as well as the estimates in (4.58), with a sufficiently small ε in these estimates, the terms $\varepsilon \|\mathbf{e_u}(t)\|^2_{\mathbf{A}_\mu(\mathbf{x}^*)}$ and $\varepsilon \int_0^t \|\dot{\mathbf{e}}_\mathbf{u}(s)\|^2_{\mathbf{M}_\rho(\mathbf{x})} \mathrm{d}s$ on the right-hand side can be absorbed by the left-hand side. For $k \geqslant 2$ and sufficiently small h, the term $Ch^{\frac{k}{2}-\frac{d}{4}} \|\mathbf{e_u}(t)\|^2_{\mathbf{A}_\mu(\mathbf{x}^*)}$ on the right-hand side can also be absorbed by the left-hand side. Then we obtain the following result:

$$\|\mathbf{e_u}(t)\|^2_{\mathbf{A}(\mathbf{x}^*)} + \int_0^t \|\dot{\mathbf{e}}_\mathbf{u}(s)\|^2_{\mathbf{M}(\mathbf{x}^*)} \, \mathrm{d}s \tag{4.59}$$

$$\leqslant \varepsilon \int_0^t \|\mathbf{e_p}(s)\|^2_{\mathbf{M}(\mathbf{x}^*)} \, \mathrm{d}s + C\varepsilon^{-1} h^{2k}$$

$$+ C\varepsilon^{-1} \int_0^t \left(\|\mathbf{e_x}(s)\|^2_{\mathbf{A}(\mathbf{x}^*)} + \|\mathbf{e_u}(s)\|^2_{\mathbf{A}(\mathbf{x}^*)} + \|\mathbf{e_w}(s)\|^2_{\mathbf{A}(\mathbf{x}^*)} \right) \mathrm{d}s.$$

The first term on the right-hand side of (4.59) is estimated in the next part of the proof.
(D) *Estimates for* $\mathbf{e_p}$ *under condition* (4.22): To estimate $\|\mathbf{e_p}\|^2_{\mathbf{M}(\mathbf{x}^*)}$, we test (3.27a) by a vector $\mathbf{v}$ associated to a finite element function v_h on $\Omega_h(t)$ to get

$$(\mathbf{v})^\top \mathbf{C}(\mathbf{x})^\top \mathbf{e_p} = (\mathbf{v})^\top \mathbf{M}_\rho(\mathbf{x})\dot{\mathbf{e}}_\mathbf{u} + (\mathbf{v})^\top (\mathbf{M}_\rho(\mathbf{x}) - \mathbf{M}_\rho(\mathbf{x}^*))\dot{\mathbf{u}}^* \tag{4.60}$$

$$+ (\mathbf{v})^{\top} \mathbf{A}_{\boldsymbol{\mu}}(\mathbf{x}) \mathbf{e_u} + (\mathbf{v})^{\top} (\mathbf{A}_{\boldsymbol{\mu}}(\mathbf{x}) - \mathbf{A}_{\boldsymbol{\mu}}(\mathbf{x}^*)) \mathbf{u}^*$$

$$+ (\mathbf{v})^{\top} (\mathbf{B}_{\boldsymbol{\rho}}(\mathbf{x}, \mathbf{u} - \mathbf{w}) \mathbf{u} - \mathbf{B}_{\boldsymbol{\rho}}(\mathbf{x}^*, \mathbf{u}^* - \mathbf{w}^*) \mathbf{u}^*)$$

$$- (\mathbf{v})^{\top} (\mathbf{C}(\mathbf{x})^{\top} - \mathbf{C}(\mathbf{x}^*)^{\top}) \mathbf{p}^* - (\mathbf{v})^{\top} (\mathbf{M}(\mathbf{x}) \mathbf{f} - \mathbf{M}(\mathbf{x}^*) \mathbf{f}^*) + (\mathbf{v})^{\top} \mathscr{E}_{\mathbf{u}1}.$$

The estimates of the terms on the right-hand side of (4.60) are similar as that in part (C) of this proof and therefore omitted. Specifically, we have

$$|(\mathbf{v})^{\top} \mathbf{C}(\mathbf{x})^{\top} \mathbf{e_p}| \leqslant C (\|\dot{\mathbf{e}}_{\mathbf{u}}\|_{\mathbf{M}(\mathbf{x})} + \|\mathbf{e_x}\|_{\mathbf{A}(\mathbf{x}^*)} + \|\mathbf{e_u}\|_{\mathbf{A}(\mathbf{x}^*)} + \|\mathbf{e_w}\|_{\mathbf{A}(\mathbf{x}^*)} + h^k) \|\mathbf{v}\|_{\mathbf{K}(\mathbf{x})}. \tag{4.61}$$

By using the inf-sup condition (2.6) in the expression for $\|\mathbf{e_p}\|_{\mathbf{M}(\mathbf{x})}$, we can estimate $\|\mathbf{e_p}\|_{\mathbf{M}(\mathbf{x})}$ as follows:

$$\|\mathbf{e_p}\|_{\mathbf{M}(\mathbf{x})} \leqslant \sup_{\mathbf{v} \neq 0} \frac{(\mathbf{v})^{\top} \mathbf{C}(\mathbf{x})^{\top} \mathbf{e_p}}{\|\mathbf{v}\|_{\mathbf{A}(\mathbf{x})}} \tag{4.62}$$

$$\leqslant C (\|\dot{\mathbf{e}}_{\mathbf{u}}\|_{\mathbf{M}(\mathbf{x})} + \|\mathbf{e_x}\|_{\mathbf{A}(\mathbf{x}^*)} + \|\mathbf{e_u}\|_{\mathbf{A}(\mathbf{x}^*)} + \|\mathbf{e_w}\|_{\mathbf{A}(\mathbf{x}^*)} + h^k),$$

which together with the norm equivalence in Remark 4.2 gives

$$\|\mathbf{e_p}\|^2_{\mathbf{M}(\mathbf{x}^*)} \leqslant C (\|\dot{\mathbf{e}}_{\mathbf{u}}\|^2_{\mathbf{M}(\mathbf{x}^*)} + \|\mathbf{e_x}\|^2_{\mathbf{A}(\mathbf{x}^*)} + \|\mathbf{e_u}\|^2_{\mathbf{A}(\mathbf{x}^*)} + \|\mathbf{e_w}\|^2_{\mathbf{A}(\mathbf{x}^*)} + h^{2k}). \tag{4.63}$$

By integrating the above inequality with respect to time from 0 to t, we have

$$\int_0^t \|\mathbf{e_p}(s)\|^2_{\mathbf{M}(\mathbf{x}^*)} \, \mathrm{d}s \leqslant C h^{2k} + C \int_0^t \|\dot{\mathbf{e}}_{\mathbf{u}}(s)\|^2_{\mathbf{M}(\mathbf{x}^*)} \, \mathrm{d}s \tag{4.64}$$

$$+ C \int_0^t \left(\|\mathbf{e_x}(s)\|^2_{\mathbf{A}(\mathbf{x}^*)} + \|\mathbf{e_u}(s)\|^2_{\mathbf{A}(\mathbf{x}^*)} + \|\mathbf{e_w}(s)\|^2_{\mathbf{A}(\mathbf{x}^*)} \right) \mathrm{d}s.$$

(E) *Combination of the estimates*: Summing up inequalities (4.30), (4.59) and $\varepsilon \times (4.64)$, substituting estimate (4.28) into the resulted inequality and choosing a sufficiently small ε, we obtain the following combined estimate:

$$\|\mathbf{e_u}(t)\|^2_{\mathbf{A}(\mathbf{x}^*)} + \|\mathbf{e_x}(t)\|^2_{\mathbf{A}(\mathbf{x}^*)} + \int_0^t \|\dot{\mathbf{e}}_{\mathbf{u}}(s)\|^2_{\mathbf{M}(\mathbf{x}^*)} \, \mathrm{d}s \tag{4.65}$$

$$\leqslant C h^{2k} + C \int_0^t \left(\|\mathbf{e_x}(s)\|^2_{\mathbf{A}(\mathbf{x}^*)} + \|\mathbf{e_u}(s)\|^2_{\mathbf{A}(\mathbf{x}^*)} \right) \mathrm{d}s.$$

By applying Gronwall's inequality, we obtain

$$\sup_{0 \leqslant t \leqslant t^*} \left(\|\mathbf{e_u}(t)\|^2_{\mathbf{K}(\mathbf{x}^*)} + \|\mathbf{e_x}(t)\|^2_{\mathbf{K}(\mathbf{x}^*)} + \int_0^t \|\dot{\mathbf{e}}_{\mathbf{u}}(s)\|^2_{\mathbf{M}(\mathbf{x}^*)} \, \mathrm{d}s \right) \leqslant C h^{2k}. \tag{4.66}$$

This result can be substituted into (4.28) and (4.64) to yield an error estimate for $\mathbf{e_w}$ and $\mathbf{e_p}$, i.e.,

$$\sup_{0 \leqslant t \leqslant t^*} \left(\|\mathbf{e_w}(t)\|^2_{\mathbf{K}(\mathbf{x}^*)} + \int_0^t \|\mathbf{e_p}(s)\|^2_{\mathbf{M}(\mathbf{x}^*)} \, \mathrm{d}s \right) \leqslant C h^{2k}. \tag{4.67}$$

It remains to show that $t^* = T$ when h is sufficiently small (smaller than some constant). In fact, for $0 \leqslant t \leqslant t^*$ we can use the inverse inequality and (4.66) to bound the left-hand side in (4.22a) as

$$\|\nabla e_x^*(t)\|_{L^{\infty}(\Omega_h^*(t))} \leqslant C h^{-\frac{d}{2}} \|\mathbf{e_x}(t)\|_{\mathbf{K}(\mathbf{x}^*)} \leqslant C h^{k - \frac{d}{2}}, \tag{4.68}$$

where C is a constant dependent on T, but independent of t^*. Since $k \geqslant 2$, we have the following inequality for $0 \leqslant t \leqslant t^*$ when h is sufficiently small:

$$\|\nabla e_x^*(t)\|_{L^{\infty}(\Omega_h^*(t))} \leqslant \frac{1}{2} h^{-\frac{d}{4}} h^{\frac{k}{2}}. \tag{4.69}$$

For the same reason, for $k \geqslant 2$ and sufficiently small h, we have

$$\|\nabla e_u(t)\|_{L^{\infty}(\Omega_h^*(t))} \leqslant \frac{1}{2} \quad \text{and} \quad \|\nabla e_w(t)\|_{L^{\infty}(\Omega_h^*(t))} \leqslant \frac{1}{2} \quad \text{for} \quad 0 \leqslant t \leqslant t^*. \tag{4.70}$$

Hence, we can extend the bound (4.22a)–(4.22c) beyond t^*, which contradicts the maximality of t^* unless $t^* = T$. This proves that the error estimates in (4.66)–(4.67) actually hold for $t^* = T$. Since the flow map

$\phi_h^*(\cdot,t)$ and its inverse are bounded in the $W^{1,\infty}$ norm, it follows that error estimates can be pulled back to the initial domain Ω_h^0 after composition with $\phi_h^*(\cdot,t)$. This proves the desired error estimates in Theorem 4.1.

5. Numerical results

In this section we present numerical results for a benchmark example without surface tension. One of the most well-known benchmark examples for the two-phase NS flow problem was proposed in [27, Table I] for simulating a moving bubble in fluid. This example includes surface tension and has been widely adopted to test the convergence and robustness of numerical methods for two-phase NS flows; see [2, 8, 21]. In our study, we modify this model by considering a simplified version without surface tension and with no-slip boundary conditions. All the computations are performed using the software package NGSolve (https://ngsolve.org), and the mesh is generated by Gmsh (https://gmsh.info/).

The computational domain is $\Omega = (0,1) \times (0,2)$ and the no-slip pieces of the boundary are given by

$$\Gamma_1 = [0,1] \times \{0,2\} \quad \text{and} \quad \Gamma_2 = \{0,1\} \times [0,2].$$

The initial interface is a circle with radius 0.25, i.e.,

$$\Gamma(0) = \{(x,y) : |(x,y) - (0.5,0.5)| = 0.25\}.$$

The gravitational force is $f = (0,-0.98)^\top$, and the physical parameters are shown in Table 1.

Table 1. *Physical parameters for the benchmark problem*

	ρ_+	ρ_-	μ_+	μ_-
BP-1	10^3	10^2	10	1

We first present numerical results to illustrate the convergence rate of the proposed method (2.8) for BP-1 in Table 1. We apply the time discretization of the proposed method (2.8) with a linearly semi-implicit Euler method, which only requires solving several decoupled linear systems at every time level. Since the exact solution of the problem is unknown, we evaluate the convergence rate in space by the formula

$$\text{conv. rate in space} = \log\left(\frac{\|u_{\tau,h} - u_{\tau,h/m}\|}{\|u_{\tau,h/m} - u_{\tau,h/m^2}\|}\right)/\log(m),$$

based on the finest three meshes, where $u_{\tau,h}$ denotes the numerical solution at $t_N = T$ computed by using a stepsize τ and mesh size h.

The spatial discretization errors of the numerical solutions at $T = 0.1$ and $T = 0.5$ are presented in Figure 2 and Figure 3, respectively, where we have used the following finite elements

$$k = 1: \text{(Mini element } P_{1b}\text{--}P_1)$$
$$k = 2: \text{(Taylor–Hood } P_2\text{--}P_1) \quad \text{and} \quad k = 3: \text{(Taylor–Hood } P_3\text{--}P_2)$$

with sufficiently small time step size $\tau = 10^{-3}$ so that the temporal discretization errors are negligibly small compared to the spatial errors. From Figures 2–3, we see that the error of space discretization is $O(h^k)$, which is consistent with the theoretical results proved in Theorem 4.1.

Next, we present simulations of the bubble's circularity, center of mass, rise velocity, and the total energy of the two-phase NS system with parameters BP-1 in Table 1. We focus on the following benchmark quantities in [27] with respect to time.

(a) *Center of mass*:

$$X_d = \frac{1}{|\Omega_-|}\int_{\Omega_-} x_d \, dx.$$

(b) *Circularity* (referred as *sphericity* in 3-D): In two dimensions it is defined as the quotient between *perimeter of area-equivalent circle* and *perimeter of the bubble*. In three dimensions it is defined as
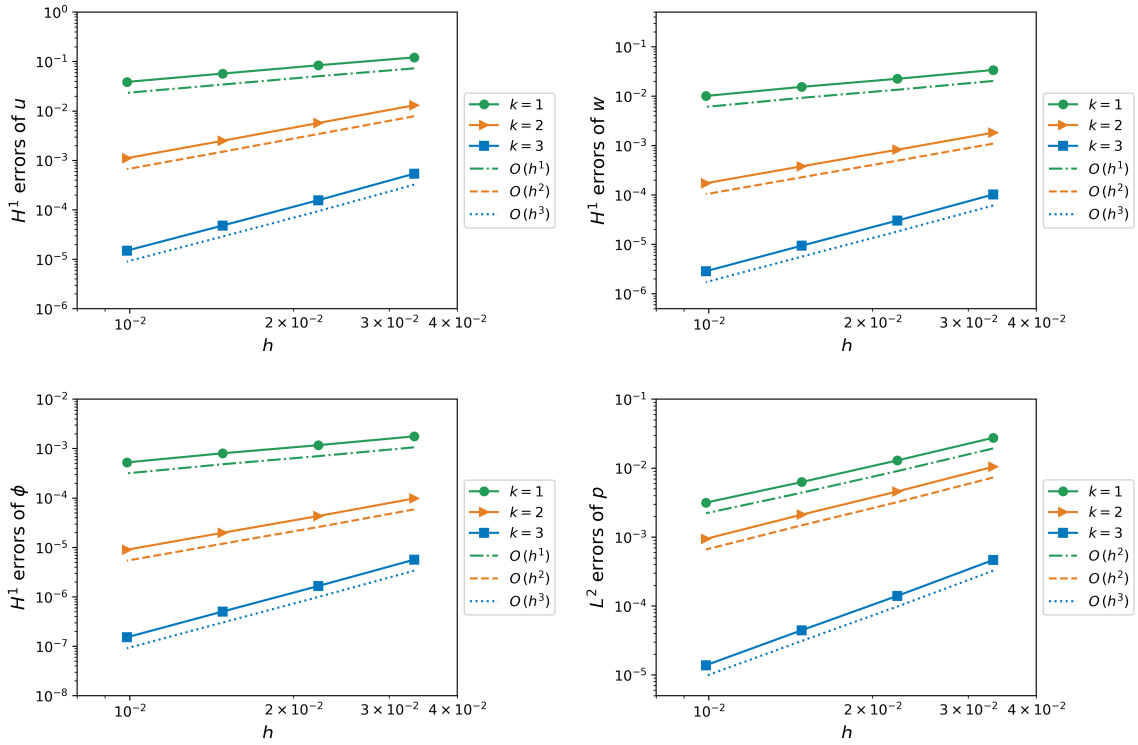
B. LI, S. MA & W. QIU

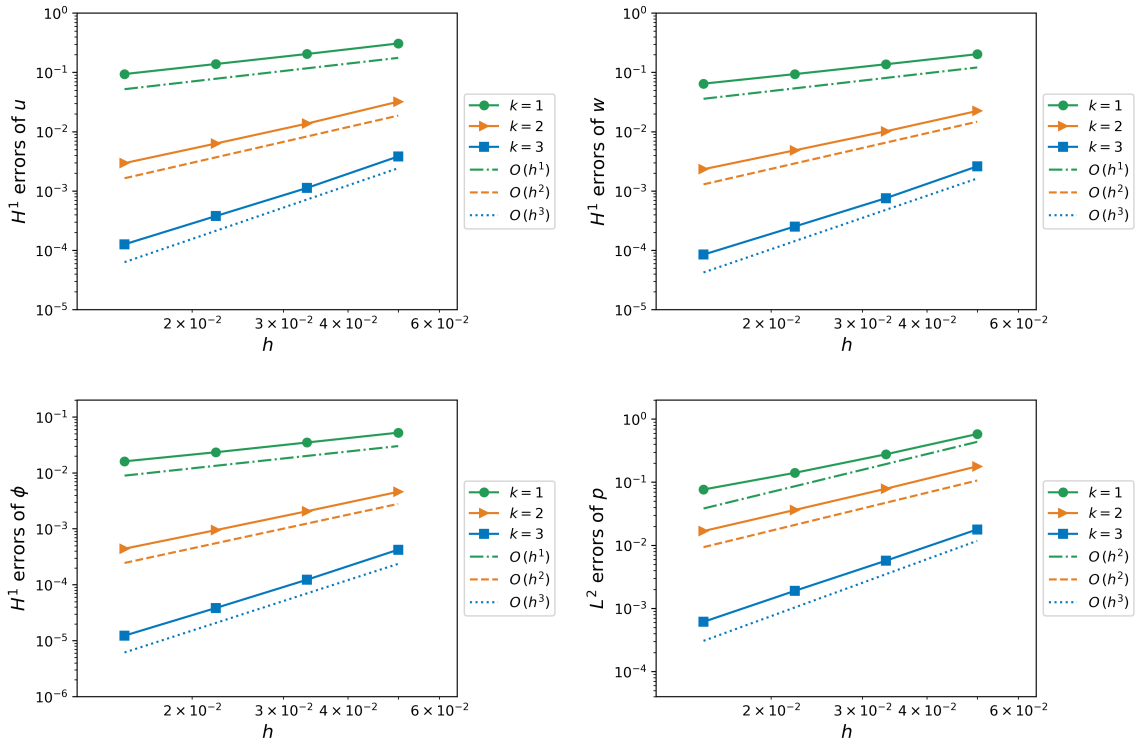FIG. 2. *Spatial discretization errors of the numerical solutions at $T = 0.1$.*

FIG. 3. *Spatial discretization errors of the numerical solutions at $T = 0.5$.*

the quotient between *surface area of volume-equivalent ball* and *surface area of the bubble*, i.e.,

$$Cir = \frac{2\sqrt{|\Omega_-|\pi}}{|\Gamma|} \quad \text{(in 2-D)} \qquad \text{and} \qquad Cir = \frac{\pi^{1/3}|\Omega_-|^{2/3}}{|\Gamma|} \quad \text{(in 3-D)}.$$

(c) *Rise velocity*:

$$V_d = \frac{1}{|\Omega_-|} \int_{\Omega_-} u_h \cdot e_d \, dx,$$

where the integrand $u_h \cdot e_d$ refers to the vertical component of the velocity u_h.

We consider the proposed method (2.8) with polynomial degree $k = 2$ and quadratic isoparametric triangular elements on the two different meshes as depicted in Figure 4, where the mesh M_1 has mesh size $h_1 = 0.04$ and the mesh M_2 has mesh size $h_2 = 0.02$. Zoom-in of the deformed mesh around $\Omega_-(t)$ at $t = 1$, 1.5 and 2.5 are shown in Figure 5.

In the computation the mesh may degenerate and in this case one has to regenerate the mesh. In our computation we regenerate the mesh when the following angle criterion is satisfied:

$$\min_{K \in \mathcal{K}_h[\mathbf{x}(t)]} \min_{\alpha \in \angle K} \leqslant \frac{\pi}{18}, \tag{5.1}$$

where $\angle K$ denotes the set of interior angles of a triangle K. The remeshing is performed using the software Gmsh. If the mesh is regenerated then we take the new mesh as the initial mesh to compute the above benchmark quantities. Figure 6 shows the number of remeshing steps over time for two different meshes, using a fixed stepsize $\tau = 1/200$. The meshes before and after the 3rd and the 6th remeshing steps are shown in Figures 7-8. In Figure 9, we present the numerical results of circularity, center of mass, rise velocity and energy on the two different meshes with $\tau = 1/200$. From Figures 4–9 we see that the method with the coarse mesh captures the shape of the interface very well. The geometry of the interface at $t = 1$, 1.5 and 2.5, as well as the distribution of interface-velocity with the mesh M_2, are presented in Figure 10.

In summary, with the numerical comparison of results between coarse and fine meshes in Figures 4–9, we have good reasons to trust the numerical simulation results for this benchmark example.

REFERENCES

[1] M. Agnese and R. Nürnberg. Fitted front tracking methods for two-phase incompressible Navier–Stokes flow: Eulerian and ALE finite element discretizations. *International Journal of Numerical Analysis & Modeling*, 17:613–642, 2020.

[2] S. Aland and A. Voigt. Benchmark computations of diffuse interface models for two-dimensional bubble dynamics. *Int. J. Numer. Methods Fluids*, 69(3):747–761, 2012.

[3] G. Anjos, N. Borhani, N. Mangiavacchi, and J. R. Thome. A 3D moving mesh finite element method for two-phase flows. *J. Comput. Phys.*, 270:366–377, 2014.

[4] S. Badia and R. Codina. Analysis of a stabilized finite element approximation of the transient convection-diffusion equation using an ALE framework. *SIAM J. Numer. Anal.*, 44(5):2159–2197, 2006.
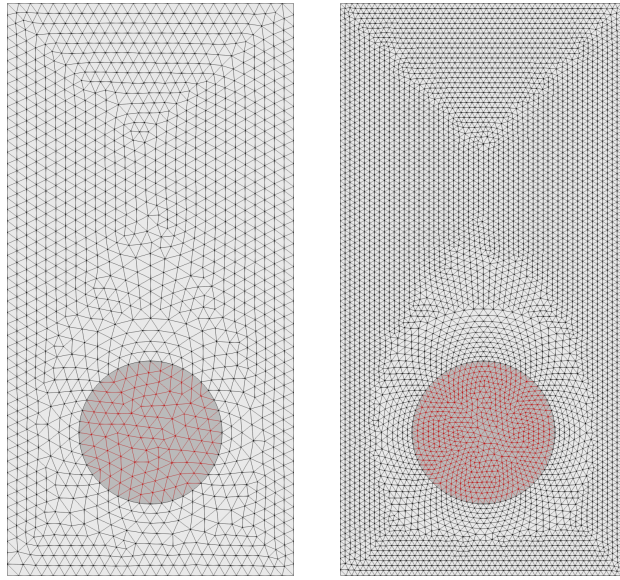
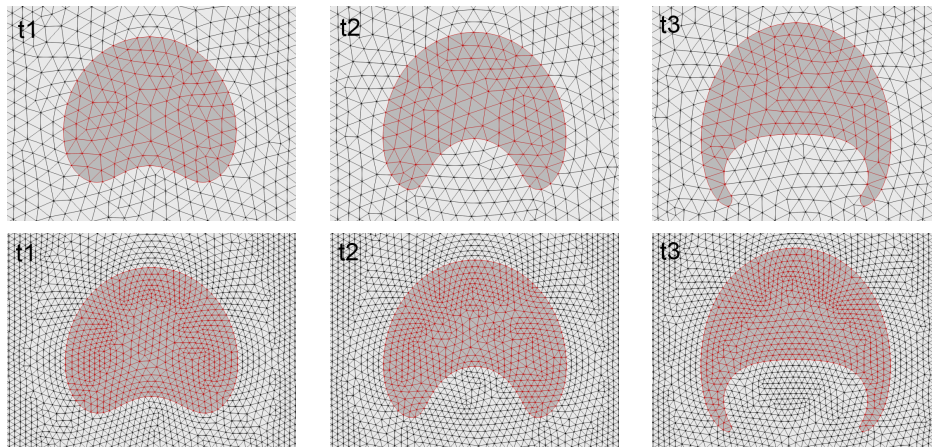FIG. 4. *Left: mesh M_1. Right: mesh M_2. Dark color: domain $\Omega_-(0)$.*

FIG. 5. *Zoom-in of the deformed mesh around $\Omega_-(t)$ with $k = 2$. Top: M_1. Bottom: M_2. Left: $t_1 = 1$. Middle: $t_2 = 1.5$. Right: $t_3 = 2.5$.*
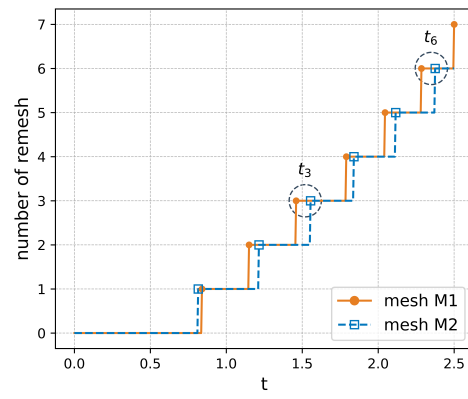
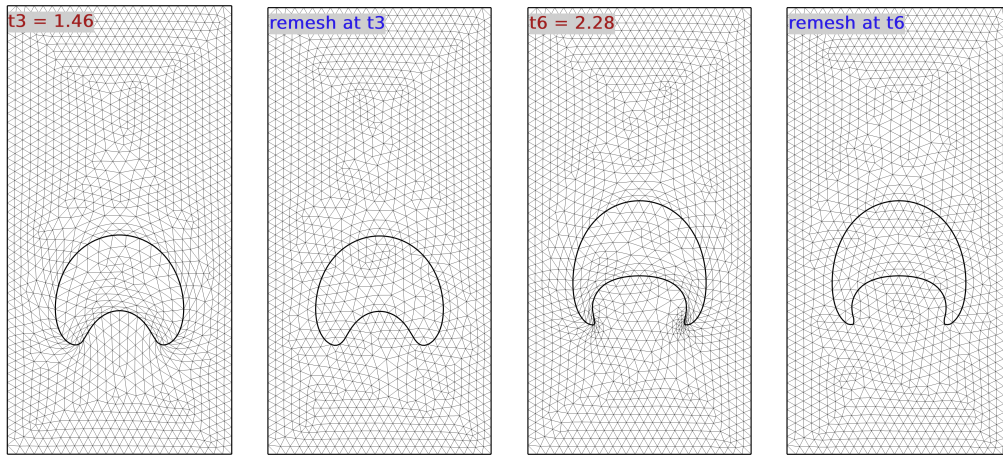FIG. 6. *The number of remeshing steps with respect to time.*

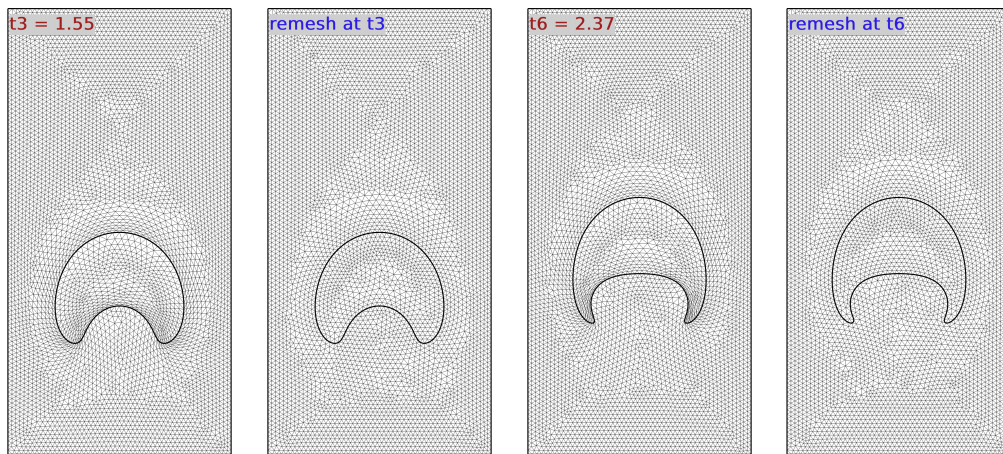FIG. 7. *Meshes before and after a remeshing step for mesh M_1 with $k = 2$.*

FIG. 8. *Meshes before and after a remeshing step for mesh M_2 with $k = 2$.*
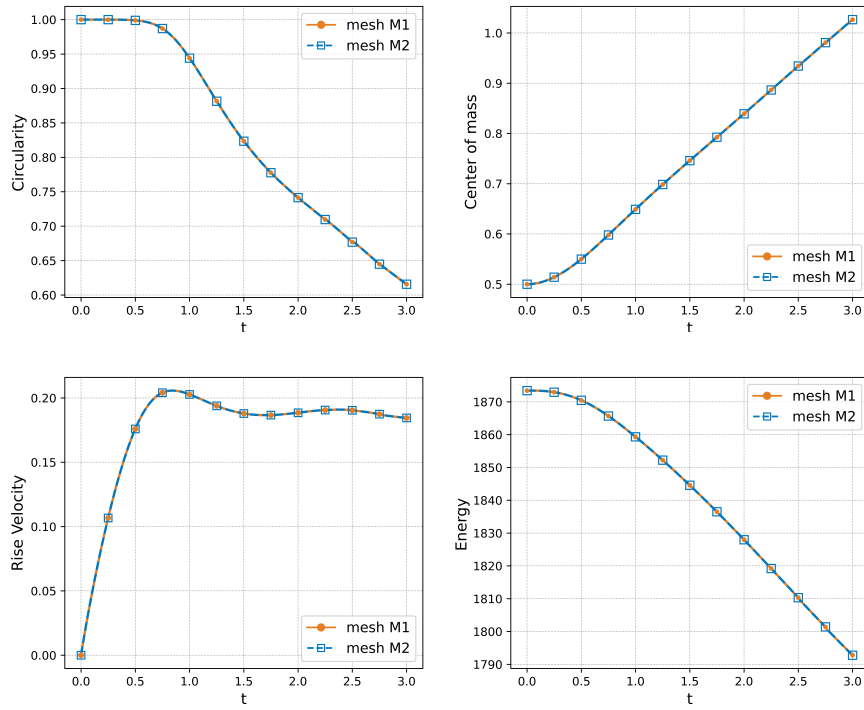
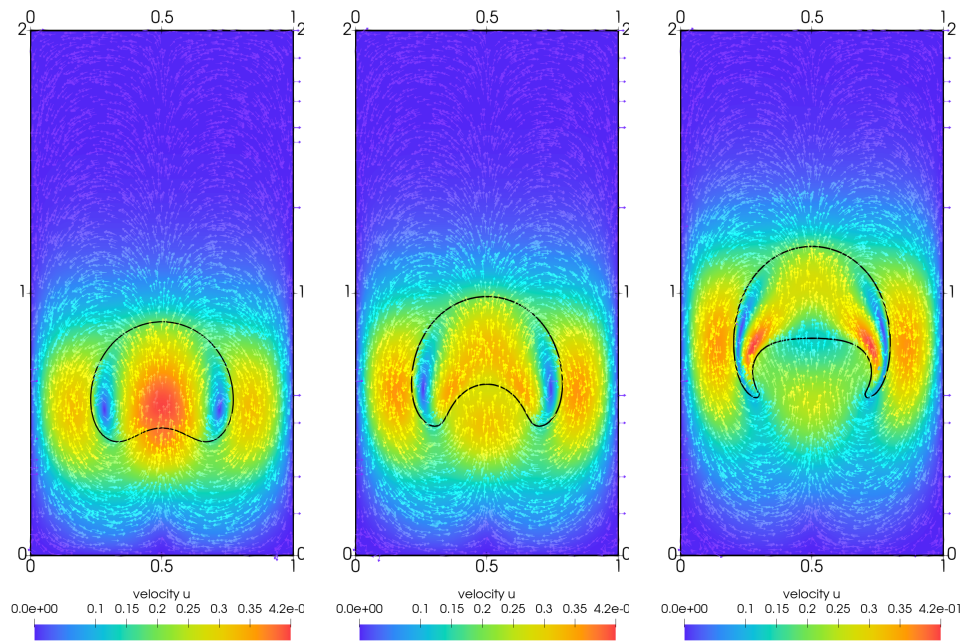FIG. 9. *Benchmark quantities with respect to time*

FIG. 10. *Velocity. Left: $t_1 = 1$. Middle: $t_2 = 1.5$. Right: $t_3 = 2.5$.*

[5] G. Bai and B. Li. A new approach to the analysis of parametric finite element approximations to mean curvature flow. *Found. Comput. Math.*, 2024, DOI: 10.1007/s10208-023-09622-x.

[6] J. W. Barrett, H. Garcke, and R. Nürnberg. Eliminating spurious velocities with a stable approximation of viscous incompressible two-phase Stokes flow. *Comput. Meth. Appl. Mech. Eng.*, 267:511–530, 2013.

[7] J. W. Barrett, H. Garcke, and R. Nürnberg. On the stable numerical approximation of two-phase flow with insoluble surfactant. *ESAIM-Math. Model. Numer. Anal.-Model. Math. Anal. Numer.*, 49(2):421–458, 2015.

[8] J. W. Barrett, H. Garcke, and R. Nürnberg. A stable parametric finite element discretization of two-phase Navier–Stokes flow. *J. Sci. Comput.*, 63:78–117, 2015.

[9] J. W. Barrett, H. Garcke, and R. Nürnberg. Parametric finite element approximations of curvature-driven interface evolutions. In *Handbook of numerical analysis*, volume 21, pages 275–423. Elsevier, 2020.

[10] D. Boffi and L. Gastaldi. Stability and geometric conservation laws for ALE formulations. *Comput. Meth. Appl. Mech. Eng.*, 193(42-44):4717–4739, 2004.

[11] F. B. D. Boffi and M. Fortin. *Mixed Finite Element Methods and Applications*. Springer-Verlag Berlin Heidelberg, 2013.

[12] B. Duan, B. Li, and Z. Yang. An energy diminishing arbitrary Lagrangian–Eulerian finite element method for two-phase Navier–Stokes flow. *J. Comput. Phys.*, 461:111215, 2022.

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

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

[15] D. Edelmann. Finite element analysis for a diffusion equation on a harmonically evolving domain. *IMA J. Numer. Anal.*, 42(2):1866–1901, 2022.

[16] C. M. Elliott and T. Ranner. A unified theory for continuous-in-time evolving finite element space approximations to partial differential equations in evolving domains. *IMA J. Numer. Anal.*, 41(3):1696–1845, 2021.

[17] C. M. Elliott and C. Venkataraman. Error analysis for an ALE evolving surface finite element method. *Numer. Meth. Part Differ. Equ.*, 31(2):459–499, 2015.

[18] L. C. Evans. *Partial differential equations*, volume 19. American Mathematical Society, 2022.

[19] L. Formaggia and F. Nobile. Stability analysis of second-order time accurate schemes for ALE–FEM. *Comput. Meth. Appl. Mech. Eng.*, 193(39-41):4097–4116, 2004.

[20] L. Formaggia, F. Nobile, et al. A stability analysis for the arbitrary Lagrangian Eulerian formulation with finite elements. *East-West Journal of Mathematics*, 7:105–132, 1999.

[21] G. Fu. Arbitrary Lagrangian–Eulerian hybridizable discontinuous Galerkin methods for incompressible flow with moving boundaries and interfaces. *Comput. Meth. Appl. Mech. Eng.*, 367:113–158, 2020.

[22] H. Garcke, R. Nürnberg, and Q. Zhao. Structure-preserving discretizations of two-phase Navier–Stokes flow using fitted and unfitted approaches. *J. Comput. Phys.*, page 112276, 2023.

[23] L. Gastaldi. A priori error estimates for the arbitrary Lagrangian Eulerian formulation with finite elements. 2001.

[24] W. Gong, B. Li, and Q. Rao. Convergent evolving finite element approximations of boundary evolution under shape gradient flow. *IMA J. Numer. Anal.*, 2023, DOI: 10.1093/imanum/drad080.

[25] S. Gross and A. Reusken. *Numerical methods for two-phase incompressible flows*, volume 40. Springer Science & Business Media, 2011.

[26] 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.

[27] S. Hysing, S. Turek, D. Kuzmin, N. Parolini, E. Burman, S. Ganesan, and L. Tobiska. Quantitative benchmark computations of two-dimensional bubble dynamics. *Int. J. Numer. Methods Fluids*, 60(11):1259–1288, 2009.

[28] 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.

[29] 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.

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

[31] B. Kovács and C. A. Power Guerra. Higher order time discretizations with ALE finite elements for parabolic problems on evolving surfaces. *IMA J. Numer. Anal.*, 38(1):460–494, 2018.

[32] R. Lan and P. Sun. A novel arbitrary Lagrangian–Eulerian finite element method for a parabolic/mixed parabolic moving interface problem. *J. Comput. Appl. Math.*, 383:1, 2020.

[33] G. Legendre and T. Takahashi. Convergence of a Lagrange-Galerkin method for a fluid-rigid body system in ALE formulation. *ESAIM-Math. Model. Numer. Anal.*, 42(4):609–644, 2008.

[34] M. Lenoir. Optimal isoparametric finite elements and error estimates for domains involving curved boundaries. *SIAM J. Numer. Anal.*, 23(3):562–580, 1986.

[35] B. Li, Y. Xia, and Z. Yang. Optimal convergence of arbitrary Lagrangian–Eulerian iso-parametric finite element methods for parabolic equations in an evolving domain. *IMA J. Numer. Anal.*, 43(1):501–534, 2023.

[36] J. Liu. Simple and efficient ALE methods with provable temporal accuracy up to fifth order for the Stokes equations on time varying domains. *SIAM J. Numer. Anal.*, 51(2):743–772, 2013.

[37] F. Nobile. Numerical approximation of fluid-structure interaction problems with application to haemodynamics. Technical report, EPFL, 2001.

[38] Q. Rao, J. Wang, and Y. Xie. Convergence of arbitrary Lagrangian-Eulerian second-order projection method for the Stokes equations on an evolving domain. *arXiv preprint*, arXiv:2310.08218.

[39] J. San Martín, L. Smaranda, and T. Takahashi. Convergence of a finite element/ALE method for the Stokes equations in a domain depending on time. *J. Comput. Appl. Math.*, 230(2):521–545, 2009.

[40] S. W. Walker. *The shapes of things: a practical guide to differential geometry and the shape derivative*. SIAM, 2015.