

CONVERGENT EVOLVING FINITE ELEMENT APPROXIMATIONS OF BOUNDARY EVOLUTION UNDER SHAPE GRADIENT FLOW

WEI GONG , BUYANG LI , AND QIQI RAO

ABSTRACT. As a specific type of shape gradient descent algorithm, shape gradient flow is widely used for shape optimization problems constrained by partial differential equations. In this approach, the constraint partial differential equations could be solved by finite element methods on a domain with a solution-driven evolving boundary. Rigorous analysis for the stability and convergence of such finite element approximations is still missing from the literature due to the complex nonlinear dependence of the boundary evolution on the solution. In this article, rigorous analysis of numerical approximations to the evolution of the boundary in a prototypical shape gradient flow is addressed. First-order convergence in time and k th order convergence in space for finite elements of degree $k \geq 2$ are proved for a linearly semi-implicit evolving finite element algorithm up to a given time. The theoretical analysis is consistent with the numerical experiments, which also illustrate the effectiveness of the proposed method in simulating two- and three-dimensional boundary evolution under shape gradient flow. The extension of the formulation, algorithm and analysis to more general shape density functions and constraint partial differential equations is also discussed.

1. Introduction

Shape optimization constrained by partial differential equations (PDEs) has wide applications in modern science and engineering, such as the airfoil designs in aerodynamics [39], automotive industry [1, 27], turbomachinery, structural design [27], and so on. These applications typically concern minimizing a shape functional

$$J(\Gamma) = \int_{\Omega} j(x, u(x)) dx$$

for some shape density function $j(\cdot, u)$ subject to a constraint such that u is the solution of a PDE problem in the domain Ω with boundary Γ . Both the shape density function $j(\cdot, u)$ and the PDE problem depend on the applications. Due to their wide applications, PDE-constrained shape optimization problems have generated much interest in developing both theoretical analysis [27, 44, 28, 10] and efficient computational methods [1, 39, 26, 25].

The boundary parametrization of an elliptic shape optimization problem was considered in [17], where error estimates for a finite element method (FEM) were obtained under the assumption that the optimal domain is star-shaped and the infinite-dimensional shape optimization problem admits a stable optimizer satisfying the second-order optimality condition. A two-dimensional shape optimization problem with the portion of the boundary to be optimized being the graph of a function was studied in [31], where second-order convergence of the numerical approximations to a local solution of the optimization problem was proved under the second-order sufficient optimality condition. The approach was extended to a Stokes shape optimization problem in [20]. The analyses in these articles are based on the second-order optimality condition and the computation of the shape Hessian, and are restricted to parametrization of boundaries with special shapes. Abstract convergence of the finite element discrete optimal shape to the optimal shape in the continuous shape optimization problem was proved in [7] for an elliptic PDE-constrained shape optimization problem in two dimensions based on the compactness argument.

2010 *Mathematics Subject Classification.* 49Q10, 65K10, 65N12, 65N15, 65N30.

Key words and phrases. Evolution of boundary, shape gradient flow, nonlinear, flow map, Poisson equation, Stokes flow, evolving finite elements, stability, convergence.

An alternative way to compute critical points of shape optimization problems without requiring the second-order optimality condition is through shape gradient flow, which is a specific type of shape gradient descent algorithm (representing a method for selecting the descent velocity) and has been widely used in the shape gradient descent algorithms for PDE-constrained shape optimization; see [5, 8]. In this approach, the boundary $\Gamma(t)$ which evolves under the shape gradient flow converges to a minimizer of the shape functional. Along with the evolution of $\Gamma(t)$, the constraint PDE problem can be solved by FEMs on the evolving domain $\Omega(t)$ enclosed by $\Gamma(t)$. We refer to [11] for a variational framework of discrete shape gradient flow for shape optimization problems. Many different FEMs have been proposed for solving PDEs on evolving domains, including cut FEM [6], immersed FEM [24], isogeometric analysis [9, 46], adaptive FEM [40], coupling of FEM and BEM [18], arbitrary Lagrangian–Eulerian FEM [19] and evolving bulk FEMs with possibly curved triangles [14], just to name a few.

The convergence of finite element approximations to linear parabolic PDEs and Stokes equations on evolving domains (with evolving boundary or interface) has been studied in many articles; see [23, 37] and [38, 3, 2, 14, 13, 21, 35, 41] for unfitted FEMs and fitted arbitrary Lagrangian–Eulerian evolving bulk FEMs, respectively. In particular, optimal-order convergence of evolving bulk FEMs has been proved in [14, 13, 36] for different isoparametric finite elements. We also refer to [15, 16, 33] for evolving surface FEMs solving PDEs on evolving surfaces. However, to the best of our knowledge, the convergence of finite element approximations to the boundary evolution (a general closed smooth boundary) under shape gradient flow is still challenging and missing from the literature. This is addressed in the current article for a class of shape gradient flows formulated below.

There are many different ways to choose the descent velocity w , most of which are based on solving the following equation (cf. [1, 5])

$$\text{Find } w \in H : \quad b(w, v) = -dJ(\Gamma(t); v) \quad \forall v \in H$$

for an abstract inner product $b(\cdot, \cdot) : H \times H \rightarrow \mathbb{R}$ associated with some Hilbert space H . This is referred to as the Hilbertian extension-regularization approach in [1, Sec. 5.2]. Different choices of H correspond to different shape gradient flows. For example, $H = L^2(\Gamma)^d$ gives L^2 or Hadamard flow, $H = H^1(\Gamma)^d$ leads to Laplace–Beltrami flow (cf. [42, 43]), and $H = H^{\frac{1}{2}}(\Gamma)^d$ yields Stefan-like flow. We refer to [5] for a comprehensive review of different shape gradient flows. The L^2 flow is generally irregular and makes sense only on the boundary Γ . The most natural choice is $H = H^m(\Omega)^d$ for some $m > \frac{d}{2} + 1$ so that $H \subset W^{1,\infty}(\Omega)^d$ is well-defined for the velocity field, which is however, computationally expensive and inconvenient.

In this article, we present formulation, algorithm and convergence analysis for a shape optimization problem constrained by the Poisson equation with a given source function f , i.e.,

$$\min_{\Gamma} J(\Gamma) = \int_{\Omega} j(x, u) dx \quad \text{subject to} \quad \begin{cases} -\Delta u = f & \text{in } \Omega \subset \mathbb{R}^d, \text{ with } d \in \{2, 3\}, \\ u = 0 & \text{on } \Gamma = \partial\Omega, \end{cases} \quad (1.1)$$

with the shape density function $j(\cdot, u) = \frac{1}{2}|\nabla u|^2$ or $j(\cdot, u) = \frac{1}{2}|u - u_d|^2$, which correspond to minimal energy dissipation and optimal shape reconstruction, respectively. For the stability and convergence of the numerical approximations, we consider the H^1 shape gradient flow of the shape functional in (1.1), i.e., the evolution of boundary $\Gamma(t) = \partial\Omega(t)$, $t \in [0, T]$, with initial position $\Gamma^0 = \partial\Omega^0$, determined by the following coupled system of equations:

$$\partial_t \phi = w \circ \phi \quad \text{in } \Omega^0, \quad \phi(0) = \text{id}|_{\Omega^0} \quad \text{in } \Omega^0, \quad (1.2a)$$

$$-\Delta w + w = 0 \quad \text{in } \Omega(t), \quad \partial_\nu w = -J'(\Gamma(t))\nu \quad \text{on } \Gamma(t), \quad (1.2b)$$

$$-\Delta u = f \quad \text{in } \Omega(t), \quad u = 0 \quad \text{on } \Gamma(t), \quad (1.2c)$$

$$-\Delta p = j'_u(\cdot, u) \quad \text{in } \Omega(t), \quad p = 0 \quad \text{on } \Gamma(t), \quad (1.2d)$$

where $\phi(\cdot, t) : \Omega^0 \rightarrow \Omega(t)$ is the flow map which generates the evolution of the boundary through $\Gamma(t) = \phi(\Gamma^0)$ under the velocity field w , j'_u is the derivative of $j(\cdot, u)$ with respect to u . By using the shape gradient $J'(\Gamma) = j(\cdot, u) + \partial_\nu p \partial_\nu u$ defined in (2.4), the rate of change of the shape

functional $J(\Gamma)$ satisfies the following relation:

$$\frac{dJ(\Gamma(t))}{dt} = \int_{\Gamma(t)} J'(\Gamma(t))w(t) \cdot \nu \, d\Gamma(t). \quad (1.3)$$

Therefore, by testing (1.2b) with w and using relation (1.3), the following property can be shown:

$$\frac{dJ(\Gamma(t))}{dt} = -\|\nabla w(t)\|_{L^2(\Omega(t))}^2 - \|w(t)\|_{L^2(\Omega(t))}^2 \leq 0, \quad (1.4)$$

i.e., the shape functional decreases as time grows. Correspondingly, the H^1 shape gradient flow evolves to a critical point of the PDE-constrained shape optimization problem.

Our analysis in this article shows that, although $H = H^1(\Omega)^d$ is not a subspace of $W^{1,\infty}(\Omega)^d$, the H^1 gradient flow naturally fits the stability and convergence analysis of evolving finite element approximations for a wide class of problems. The novelty and contribution of this article include:

- An H^1 shape gradient flow of PDEs on an evolving domain $\Omega(t)$, with a solution-driven evolving boundary $\Gamma(t) = \partial\Omega(t)$, is formulated in a way that convergence of evolving finite element approximations to the boundary evolution could be proved.
- The distributed Eulerian derivative of the shape functional on the bulk domain is used for the convenience of computation and for proving the stability and convergence of numerical approximations.
- A linearly semi-implicit evolving FEM is proposed for the nonlinear PDEs and the solution-driven bulk and boundary evolution. The method only requires solving several decoupled linear systems at every time level. First-order convergence in time and k th order convergence in space for finite elements of degrees $k \geq 2$ are proved up to any given time.
- The analysis could cover a general class of shape optimization problems for a class of shape density functions including

$$j(\cdot, u) = \frac{1}{2}|\nabla u|^2 \quad \text{and} \quad j(\cdot, u) = \frac{1}{2}|u - u_d|^2,$$

for constraints which include both the Poisson equation and the Stokes equations in two and three dimensions.

- For certain shape optimization problems, the volume constraint is indispensable to ensure the existence of a solution; see [1, Sec. 5.3]. The stability and convergence analysis in this article could be naturally extended to a velocity w determined by the following weak formulation: Find $(w, q) \in H^1(\Omega) \times \mathbb{R}$ such that

$$\begin{aligned} \int_{\Omega} (\nabla w : \nabla v + wv) \, dx - \int_{\Omega} q \nabla \cdot v \, dx &= - \int_{\Gamma} J'(\Gamma)v \cdot \nu \, d\Gamma \quad \forall v \in H^1(\Omega), \\ \int_{\Omega} \nabla \cdot w \eta \, dx &= 0 \quad \forall \eta \in \mathbb{R}, \end{aligned} \quad (1.5)$$

where the second equation is equivalent to requiring the velocity field w to be volume-conserving, and q can be regarded as a Lagrange multiplier. The stability estimates for this type of equations can be done similarly.

The rest of this article is organized as follows. In Section 2, we formulate (1.2) into a computationally convenient form in terms of the distributed Eulerian derivative of the shape functional on the bulk domain, and propose a linearly semi-implicit evolving FEM for approximating the evolution of $\Omega(t)$. Then we present the main theoretical result on the convergence of the evolving finite element approximations. The proof of the main theoretical result is presented in Section 3. Numerical tests are presented in Section 4 to illustrate the convergence of the numerical approximations and the evolution of the domain under shape gradient flow. Concluding remarks and extension to other shape density functions and constraints are discussed in Section 5.

2. Notation and main results

In this section we present the notation and main results of this article, including the weak formulation of the nonlinear PDEs with solution-driven bulk and boundary evolution, the evolving finite element algorithm which tracks the bulk and boundary evolution, and the convergence of the numerical approximations.

2.1. Preliminaries of shape calculus

In this subsection, we introduce some basic ingredients of the velocity method for shape calculus. We refer the readers to [10, 44] for more details on geometry, shape calculus, and shape optimization.

Since we consider the evolution of a domain under shape gradient flow, we shall frequently calculate the rate of change of integrals over the moving domain. This can be obtained by using the following result.

Lemma 2.1 ([45, Lemma 5.7]). *If the domain Ω moves with velocity $v \in W^{1,\infty}(\Omega)$, then*

$$\frac{d}{dt} \int_{\Omega} f dx = \int_{\Omega} (\partial_t^{\bullet} f + f \nabla \cdot v) dx, \quad (2.1)$$

where $\partial_t^{\bullet} f := \partial_t f + \nabla f \cdot v$ denotes the material derivative of f .

For any smooth vector field $v : \mathbb{R}^d \rightarrow \mathbb{R}^d$ we denote by $\mathcal{F}^t[v] : \Gamma \rightarrow \mathbb{R}^d$, $t \geq 0$, the flow map determined by the velocity field v , defined by

$$\frac{d}{dt} \mathcal{F}^t[v] = v \circ \mathcal{F}^t[v] \text{ on } \Gamma \text{ with initial condition } \mathcal{F}^0[v] = \text{id}|_{\Gamma}. \quad (2.2)$$

The *Eulerian derivative* of $J(\Gamma)$ at Γ in the direction v is defined as

$$dJ(\Gamma; v) := \left. \frac{d}{dt} J(\mathcal{F}^t[v](\Gamma)) \right|_{t=0}. \quad (2.3)$$

It can be formulated as an integral on the boundary in terms of the shape gradient defined on the boundary Γ (see [10, Chap. 9, Sec. 3.4]), i.e.,

$$dJ(\Gamma; v) = \int_{\Gamma} J'(\Gamma) v \cdot \nu d\Gamma. \quad (2.4)$$

In fact, $J'(\Gamma)$ is defined as the function on Γ satisfying relation (2.4).

We shall focus on the shape gradient flow associated to the shape densify function $j(\cdot, u) = \frac{1}{2}|u - u_d|^2$, where u_d is a given smooth target function. The numerical approximation and analysis for the shape gradient flow associated to the shape densify function $j(\cdot, u) = \frac{1}{2}|\nabla u|^2$ could be analyzed similarly and therefore are only briefly discussed in the conclusion section. For the shape densify function $j(\cdot, u) = \frac{1}{2}|u - u_d|^2$, the Eulerian derivative of the functional $J(\Gamma)$ can also be written in terms of integrals over the bulk domain Ω (called the *distributed Eulerian derivative*), as shown in the following lemma.

Lemma 2.2 ([29, eq. (2.9)-(2.10)] and [22, eq. (2.6)]). *Let u and p be the solutions of the primal equation (1.2c) and the adjoint state equation (1.2d), respectively. Then the Eulerian derivative of the functional $J(\Gamma)$ with $j(\cdot, u) = \frac{1}{2}|u - u_d|^2$ defined in (1.1) has the following closed form:*

$$\begin{aligned} dJ(\Gamma; v) = dJ(\Gamma, u, p; v) := & \int_{\Omega} \nabla u \cdot (\nabla v + \nabla v^{\top}) \nabla p - f \nabla p \cdot v dx \\ & + \int_{\Omega} \left(\frac{1}{2}|u - u_d|^2 - \nabla u \cdot \nabla p \right) \nabla \cdot v - (u - u_d) \nabla u_d \cdot v dx, \end{aligned} \quad (2.5)$$

where u and p are determined by Γ through the following equations:

$$\begin{aligned} -\Delta u &= f \quad \text{in } \Omega, \quad \text{with } u = 0 \text{ on } \Gamma \\ -\Delta p &= u - u_d \quad \text{in } \Omega, \quad \text{with } p = 0 \text{ on } \Gamma. \end{aligned}$$

2.2. Weak formulation of the PDEs

For any flow map $\phi : \Omega^0 \times [0, T] \rightarrow \mathbb{R}^d$, we denote by $\Omega[\phi(\cdot, t)]$ the image of Ω^0 under the map $\phi(\cdot, t)$. Since the boundary type Eulerian derivative in (2.4) and the distributed type Eulerian derivative in (2.5) are equivalent, (1.2b) can be equivalently written into the following weak formulation by using integration by part: Find $w \in H^1(\Omega)^d$ such that

$$\begin{aligned} \int_{\Omega} (\nabla w : \nabla v + w \cdot v) dx &= \int_{\Gamma} \partial_{\nu} w \cdot \nu d\Gamma = - \int_{\Gamma} J'(\Gamma) v \cdot \nu d\Gamma \\ &= -dJ(\Gamma; v) = -dJ(\Gamma, u, p; v) \quad \forall v \in H^1(\Omega)^d, \end{aligned} \quad (2.6)$$

where the closed form of $dJ(\Gamma, u, p; v)$ is given by (2.5). Then the moving boundary problem in (1.2) can be written into the following weak formulation:

$$\partial_t \phi = w \circ \phi, \quad (2.7a)$$

$$\int_{\Omega[\phi(\cdot, t)]} \nabla w : \nabla \chi_w + w \cdot \chi_w dx = -dJ(\Gamma(t), u, p; \chi_w) \quad \forall \chi_w \in H^1(\Omega[\phi(\cdot, t)])^d, \quad (2.7b)$$

$$\int_{\Omega[\phi(\cdot, t)]} \nabla u \cdot \nabla \chi_u dx = \int_{\Omega[\phi(\cdot, t)]} f \chi_u dx \quad \forall \chi_u \in H_0^1(\Omega[\phi(\cdot, t)]), \quad (2.7c)$$

$$\int_{\Omega[\phi(\cdot, t)]} \nabla p \cdot \nabla \chi_p dx = \int_{\Omega[\phi(\cdot, t)]} j'_u(x, u) \chi_p dx \quad \forall \chi_p \in H_0^1(\Omega[\phi(\cdot, t)]), \quad (2.7d)$$

under the initial condition $\phi(\cdot, 0) = \text{id}|_{\Omega^0}$, with $\Gamma(t) = \partial\Omega(t)$.

2.3. The evolving finite elements

Assume that the given smooth initial domain $\Omega^0 \subset \mathbb{R}^d$, $d \in \{1, 2, 3\}$, is divided into a set \mathcal{K}_h^0 of shape-regular and quasi-uniform simplices with mesh size h , using curved finite elements of degree $k \geq 2$ to approximate the boundary $\Gamma^0 = \partial\Omega^0$. In particular, every curved simplex $K \in \mathcal{K}_h^0$ is the image of a unique polynomial of degree k defined on the reference simplex \hat{K} , denoted by $F_K : \hat{K} \rightarrow K$, called the parametrization of K ; see [14, §8.6]. Moreover, every boundary simplex $K \in \mathcal{K}_h^0$ (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 [30] based on some parametrization of the boundary $\Upsilon : \partial\tilde{D} \rightarrow \partial\Omega^0$, where $\partial\tilde{D}$ denotes the flat boundary face in the flat simplex with the same vertices as the curved boundary simplex K . In practice one can choose the parametrization Υ to be the normal projection onto $\partial\Omega^0$, i.e., the unique point $\Upsilon(x) \in \Gamma$ such that (cf. [12, Section 2.1])

$$x = \Upsilon(x) + \text{sign}(x, \Omega) |x - \Upsilon(x)| n(\Upsilon(x)),$$

where $n(\Upsilon(x))$ is the unit outward normal vector and

$$\text{sign}(x, \Omega) = \begin{cases} 1 & \text{for } x \in \mathbb{R}^d \setminus \bar{\Omega}(t), \\ -1 & \text{for } x \in \Omega(t). \end{cases}$$

Thus the initial domain Ω^0 is approximated by the triangulated domain $\Omega_h^0 = \bigcup_{K \in \mathcal{K}_h^0} K$.

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

$$\mathbf{x}(t) = (x_1(t), \dots, x_N(t)),$$

which determines the triangulation $\mathcal{K}_h[\mathbf{x}(t)]$ and the domain $\Omega_h[\mathbf{x}(t)] = \bigcup_{K \in \mathcal{K}_h[\mathbf{x}(t)]} K$ by piecewise polynomial interpolation on the reference simplex. Thus the edges or surfaces of the simplices on both interior and boundary are curved, as in references [13, 14].

Similarly as the simplices on the initial domain, if $K \in \mathcal{K}_h[\mathbf{x}(t)]$ is a simplex on the evolving domain $\Omega_h[\mathbf{x}(t)]$ then we denote by $F_K : \hat{K} \rightarrow K$ the parametrization of K , i.e., the unique polynomial of degree k that maps the reference flat simplex \hat{K} onto the possibly curved simplex

K . Correspondingly, the finite element space on the evolving domain $\Omega_h[\mathbf{x}(t)]$ is defined as

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

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

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

It maps the initial domain Ω_h^0 onto the evolving domain $\Omega_h[\mathbf{x}(t)]$. The velocity of the mesh movement is the unique function $w_h(t) \in S_h^k[\mathbf{x}(t)]^d$ satisfying the following relation:

$$w_h(t) \circ \phi_h(\xi, t) = \frac{d}{dt} \phi_h(\xi, t) \quad \forall \xi \in \Omega_h^0.$$

If $v(\cdot, t)$ is a function defined on $\Omega_h[\mathbf{x}(t)]$ for $t \in [0, T]$, then its material derivative with respect to the velocity field w_h is defined as

$$\partial_{t,h}^\bullet v(x, t) = \frac{d}{dt} v(\phi_h(\xi, t), t) \quad \text{at } x = \phi_h(\xi, t) \in \Omega_h[\mathbf{x}(t)].$$

The finite element basis functions of $S_h^k[\mathbf{x}(t)]$ are denoted by $\phi_j[\mathbf{x}(t)]$, $j = 1, \dots, N$, satisfying the identities:

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

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

$$\partial_{t,h}^\bullet \phi_j[\mathbf{x}(t)] = 0 \quad \text{on } \Omega_h[\mathbf{x}(t)], \quad j = 1, \dots, N. \quad (2.8)$$

By using the basis functions $\phi_j[\mathbf{x}(t)]$, the finite element space on the evolving domain $\Omega_h[\mathbf{x}(t)]$ can be written as

$$S_h^k[\mathbf{x}(t)] = \left\{ \sum_{j=1}^N c_j \phi_j[\mathbf{x}(t)] : c_j \in \mathbb{R} \right\},$$

$$\hat{S}_h^k[\mathbf{x}(t)] = \left\{ v \in S_h^k[\mathbf{x}(t)] : v = 0 \text{ on } \partial\Omega_h[\mathbf{x}(t)] \right\}.$$

2.4. The numerical method and its convergence

The semidiscrete evolving FEM for (2.7) reads: Find $\phi_h(\cdot, t) \in S_h^k[\mathbf{x}^0]^d \subset H^1(\Omega_h^0)^d$ and $\mathbf{x}(t) = \phi_h(\mathbf{x}^0, t)$, along with $w_h(t) \in S_h^k[\mathbf{x}(t)]^d \subset H^1(\Omega_h[\mathbf{x}(t)])^d$ and $u_h(t), p_h(t) \in \hat{S}_h^k[\mathbf{x}(t)] \subset H_0^1(\Omega_h[\mathbf{x}(t)])$, satisfying the following weak formulation:

$$\partial_t \phi_h = w_h \circ \phi_h \quad (2.9a)$$

$$\int_{\Omega_h[\mathbf{x}(t)]} \nabla w_h : \nabla \chi_w + w_h \cdot \chi_w dx = -dJ(\Gamma_h[\mathbf{x}(t)], u_h, p_h; \chi_w) \quad \forall \chi_w \in S_h^k[\mathbf{x}(t)]^d \quad (2.9b)$$

$$\int_{\Omega_h[\mathbf{x}(t)]} \nabla u_h \cdot \nabla \chi_u dx = \int_{\Omega_h[\mathbf{x}(t)]} f \chi_u dx \quad \forall \chi_u \in \hat{S}_h^k[\mathbf{x}(t)] \quad (2.9c)$$

$$\int_{\Omega_h[\mathbf{x}(t)]} \nabla p_h \cdot \nabla \chi_p dx = \int_{\Omega_h[\mathbf{x}(t)]} j'_u(x, u_h) \chi_p dx \quad \forall \chi_p \in \hat{S}_h^k[\mathbf{x}(t)] \quad (2.9d)$$

under the initial condition $\phi_h(0) = \phi_h^0 := \text{id}|_{\Omega_h^0}$.

We consider the following time discretization of (2.9) with a linearly semi-implicit Euler method: For given $\phi_h^n \in S_h^k[\mathbf{x}^0]^d$ and $\mathbf{x}^n = \phi_h^n(\mathbf{x}^0)$, find $\phi_h^{n+1} \in S_h^k[\mathbf{x}^0]^d$, $w_h^n \in S_h^k[\mathbf{x}^n]^d$ and $u_h^n, p_h^n \in \hat{S}_h^k[\mathbf{x}^n]$ such that

$$\frac{\phi_h^{n+1} - \phi_h^n}{\tau} = w_h^n \circ \phi_h^n \quad (2.10a)$$

$$\int_{\Omega_h[\mathbf{x}^n]} \nabla w_h^n : \nabla \chi_w + w_h^n \cdot \chi_w dx = -dJ(\Gamma_h[\mathbf{x}^n], u_h^n, p_h^n; \chi_w) \quad \forall \chi_w \in S_h^k[\mathbf{x}^n]^d \quad (2.10b)$$

$$\int_{\Omega_h[\mathbf{x}^n]} \nabla u_h^n \cdot \nabla \chi_u dx = \int_{\Omega_h[\mathbf{x}^n]} f \chi_u dx \quad \forall \chi_u \in \mathring{S}_h^k[\mathbf{x}^n] \quad (2.10c)$$

$$\int_{\Omega_h[\mathbf{x}^n]} \nabla p_h^n \cdot \nabla \chi_p dx = \int_{\Omega_h[\mathbf{x}^n]} j'_u(x, u_h^n) \chi_p dx \quad \forall \chi_p \in \mathring{S}_h^k[\mathbf{x}^n] \quad (2.10d)$$

and then set $\mathbf{x}^{n+1} = \phi_h^{n+1}(\mathbf{x}^0)$. The algorithm only requires solving several decoupled linear systems at every time level. In practice, u_h^n and p_h^n can be first solved from (2.10c)–(2.10d) and substituted into (2.10b) to yield w_h^n . The latter is used to defined ϕ_h^{n+1} through (2.10a).

For the practical computation and numerical analysis it is convenient to use matrix-vector formulation to represent the system of equations in (2.10). For this purpose, we express the unknown solutions in terms of the finite element basis functions $\phi_j[\mathbf{x}^n]$, $j = 1, \dots, N$ on domain $\Omega_h[\mathbf{x}^n]$, i.e.,

$$\begin{aligned} u_h^n &= \sum_{j=1}^N u_j^n \phi_j[\mathbf{x}^n], & \text{with } u_j^n &= u_h^n(x_j^n) \in \mathbb{R}, \\ p_h^n &= \sum_{j=1}^N p_j^n \phi_j[\mathbf{x}^n], & \text{with } p_j^n &= p_h^n(x_j^n) \in \mathbb{R}, \\ w_h^n &= \sum_{j=1}^N w_j^n \phi_j[\mathbf{x}^n], & \text{with } w_j^n &= w_h^n(x_j^n) \in \mathbb{R}^d, \end{aligned}$$

and collect the nodal values in column vectors

$$\mathbf{u}^n = (u_j^n) \in \mathbb{R}^N, \quad \mathbf{p}^n = (p_j^n) \in \mathbb{R}^N \quad \text{and} \quad \mathbf{w}^n = (w_j^n) \in \mathbb{R}^{dN}.$$

We define the domain-dependent mass matrix $\mathbf{M}(\mathbf{x}^n)$ and stiffness matrix $\mathbf{A}(\mathbf{x}^n)$ on the domain $\Omega_h[\mathbf{x}^n]$ determined by the nodal vector \mathbf{x}^n , i.e.,

$$\begin{aligned} \mathbf{M}(\mathbf{x}^n)|_{jk} &= \int_{\Omega_h[\mathbf{x}^n]} \phi_j[\mathbf{x}^n] \phi_k[\mathbf{x}^n] dx \\ \mathbf{A}(\mathbf{x}^n)|_{jk} &= \int_{\Omega_h[\mathbf{x}^n]} \nabla \phi_j[\mathbf{x}^n] \cdot \nabla \phi_k[\mathbf{x}^n] dx \quad \text{for } j, k = 1, \dots, N. \end{aligned}$$

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 $\mathbf{K}(\mathbf{x}^n) := \mathbf{M}(\mathbf{x}^n) + \mathbf{A}(\mathbf{x}^n)$,

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

To simplify the notation, we will use $\mathbf{K}(x^n)$ 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:

$$\begin{aligned} \|\mathbf{u}^n\|_{\mathbf{M}(\mathbf{x}^n)}^2 &:= (\mathbf{u}^n)^\top \mathbf{M}(\mathbf{x}^n) \mathbf{u}^n = \|u_h^n\|_{L^2(\Omega_h[\mathbf{x}^n])}^2, \\ \|\mathbf{u}^n\|_{\mathbf{A}(\mathbf{x}^n)}^2 &:= (\mathbf{u}^n)^\top \mathbf{A}(\mathbf{x}^n) \mathbf{u}^n = \|\nabla u_h^n\|_{L^2(\Omega_h[\mathbf{x}^n])}^2, \\ \|\mathbf{w}^n\|_{\mathbf{K}(\mathbf{x}^n)}^2 &:= (\mathbf{w}^n)^\top \mathbf{K}(\mathbf{x}^n) \mathbf{w}^n = \|w_h^n\|_{H^1(\Omega_h[\mathbf{x}^n])}^2. \end{aligned}$$

The right-hand side vectors $\mathbf{f}(\mathbf{x}^n)$, $\mathbf{J}'_u(\mathbf{x}^n, \mathbf{u}^n) \in \mathbb{R}^N$ and $d\mathbf{J}(\mathbf{x}^n, \mathbf{u}^n, \mathbf{p}^n) \in \mathbb{R}^{dN}$ are given by

$$\begin{aligned} \mathbf{f}(\mathbf{x}^n)|_j &= \int_{\Omega_h[\mathbf{x}^n]} f \phi_j[\mathbf{x}^n] dx, \\ \mathbf{J}'_u(\mathbf{x}^n, \mathbf{u}^n)|_j &= \int_{\Omega_h[\mathbf{x}^n]} j'_u(x, u_h^n) \phi_j[\mathbf{x}^n] dx, \\ d\mathbf{J}(\mathbf{x}^n, \mathbf{u}^n, \mathbf{p}^n)|_{d(j-1)+l} &= \int_{\Omega_h[\mathbf{x}^n]} (\nabla \phi_j[\mathbf{x}^n] \cdot \nabla u_h^n - f \phi_j[\mathbf{x}^n]) (\nabla p_h^n)_l + \nabla \phi_j[\mathbf{x}^n] \cdot \nabla p_h^n (\nabla u_h^n)_l \\ &\quad + (j'_u(x, u_h^n) - \nabla u_h^n \cdot \nabla p_h^n) (\nabla \phi_j[\mathbf{x}^n])_l - j'_u(x, u_h^n) (\nabla u_d)_l \phi_j[\mathbf{x}^n] dx, \end{aligned}$$

for $j = 1, \dots, N$ and $1 \leq l \leq d$.

Using the nodal values vectors \mathbf{x} , \mathbf{w} , \mathbf{u} and \mathbf{p} and the matrices and right-hand side vectors defined above, the fully discrete algorithm in (2.10) can be written into the following matrix-vector form:

$$\mathbf{x}^{n+1} - \mathbf{x}^n = \tau \mathbf{w}^n, \quad (2.11a)$$

$$\mathbf{K}(\mathbf{x}^n) \mathbf{w}^n = -d\mathbf{J}(\mathbf{x}^n, \mathbf{u}^n, \mathbf{p}^n), \quad (2.11b)$$

$$\mathbf{A}(\mathbf{x}^n) \mathbf{u}^n = \mathbf{f}(\mathbf{x}^n), \quad (2.11c)$$

$$\mathbf{A}(\mathbf{x}^n) \mathbf{p}^n = \mathbf{J}'_u(\mathbf{x}^n, \mathbf{u}^n). \quad (2.11d)$$

We are now in the position to state the main result of this article, i.e., the convergence of the fully discrete evolving FEM in (2.10). To this end, we denote by $\mathbf{x}_*^n = \mathbf{x}_*(t_n) = \phi(\mathbf{x}^0, t_n)$ the image of the nodes of Ω_h^0 under the exact flow map. Correspondingly, $\mathcal{K}[\mathbf{x}_*^n]$ is a triangulation of the domain $\Omega[\phi(\cdot, t_n)]$ based on interpolation at the nodes in the vector \mathbf{x}_*^n , and $S_h^k[\mathbf{x}_*^n]$ is the finite element space on the triangulated domain $\Omega_h[\mathbf{x}_*^n] = \bigcup_{K \in \mathcal{K}[\mathbf{x}_*^n]} K$. At the initial moment, we have $\Omega_h[\mathbf{x}_*^0] = \Omega_h[\mathbf{x}^0] = \Omega_h^0$. The Lagrangian interpolation of the exact solution, denoted by $\hat{\phi}_h^n \in S_h^k[\mathbf{x}^0]^d$ and $\hat{w}_h^n \in S_h^k[\mathbf{x}_*^n]^d$ and $\hat{u}_h^n, \hat{p}_h^n \in \hat{S}_h^k[\mathbf{x}_*^n]$, can be compared with the numerical solution after both being pulled back to the initial domain Ω_h^0 . This is presented in the following theorem.

Theorem 2.3 (Convergence of the evolving finite element approximations). *Suppose that the solution of (2.7), the flow map $\phi : \Omega^0 \times [0, T] \rightarrow \mathbb{R}^d$ and its inverse $\phi(\cdot, t)^{-1} : \Omega[\phi(\cdot, t)] \rightarrow \Omega^0$, and the domain $\Omega[\phi(\cdot, t)]$ are all sufficiently smooth, and assume that the triangulations $\mathcal{K}[x_*(t)]$ keep shape-regular and quasi-uniform (see Remark 2.5). Then there exist positive constants τ_0 and h_0 such that for $\tau \leq \tau_0$ and $h \leq h_0$ together with the restriction $\tau = o(h^{\frac{d}{2}})$, the following error estimates hold:*

$$\begin{aligned} \|\phi_h^n - \hat{\phi}_h^n\|_{H^1(\Omega_h^0)} + \|w_h^n \circ \phi_h^n - \hat{w}_h^n \circ \hat{\phi}_h^n\|_{H^1(\Omega_h^0)} + \|u_h^n \circ \phi_h^n - \hat{u}_h^n \circ \hat{\phi}_h^n\|_{H^1(\Omega_h^0)} \\ + \|p_h^n \circ \phi_h^n - \hat{p}_h^n \circ \hat{\phi}_h^n\|_{H^1(\Omega_h^0)} \leq C(\tau + h^k), \end{aligned} \quad (2.12)$$

where $\hat{\phi}_h^n \in S_h^k[\mathbf{x}^0]^d$ and $\hat{w}_h^n \in S_h^k[\mathbf{x}_*^n]^d$ and $\hat{u}_h^n, \hat{p}_h^n \in \hat{S}_h^k[\mathbf{x}_*^n]$ are the Lagrangian interpolations of the exact solutions.

Remark 2.4. If the domain $\Omega[\phi^n]$ is smooth and the triangulation $\mathcal{K}[\mathbf{x}_*^n]$ is shape-regular and quasi-uniform, and the exact solutions $w(\cdot, t_n)$, $u(\cdot, t_n)$ and $p(\cdot, t_n)$ are sufficiently smooth, then the following errors estimates of the Lagrangian interpolation are known (see [4]):

$$\|\tilde{w}^n - \hat{w}_h^n\|_{W^{1,\infty}(\Omega_h[\mathbf{x}_*^n])} + \|\tilde{u}^n - \hat{u}_h^n\|_{W^{1,\infty}(\Omega_h[\mathbf{x}_*^n])} + \|\tilde{p}^n - \hat{p}_h^n\|_{W^{1,\infty}(\Omega_h[\mathbf{x}_*^n])} \leq Ch^k, \quad (2.13a)$$

$$\|\tilde{w}^n - \hat{w}_h^n\|_{L^\infty(\Omega_h[\mathbf{x}_*^n])} + \|\tilde{u}^n - \hat{u}_h^n\|_{L^\infty(\Omega_h[\mathbf{x}_*^n])} + \|\tilde{p}^n - \hat{p}_h^n\|_{L^\infty(\Omega_h[\mathbf{x}_*^n])} \leq Ch^{k+1}, \quad (2.13b)$$

where $\tilde{w}^n = \tilde{w}(\cdot, t_n)$, $\tilde{u}^n = \tilde{u}(\cdot, t_n)$ and $\tilde{p}^n = \tilde{p}(\cdot, t_n)$ are the smooth extensions of $w(\cdot, t_n)$, $u(\cdot, t_n)$ and $p(\cdot, t_n)$ onto \mathbb{R}^d .

Remark 2.5. The constants τ_0 , h_0 and C depend on $\|\nabla\phi\|_{L^\infty(\Omega)}$, which represents the deformation of the domain. Therefore, the larger the deformation, the smaller stepsize and mesh size are required. Moreover, if the initial triangulation is shape-regular and quasi-uniform, then the triangulations $\mathcal{K}[\mathbf{x}_*(t)]$ will keep shape-regular and quasi-uniform when $\|\nabla\phi\|_{L^\infty(\Omega)}$ and $\|\nabla\phi^{-1}\|_{L^\infty(\Omega(t))}$ are bounded. In practice, we could divide the time interval $[0, T]$ into several sufficiently small subintervals $[T_{j-1}, T_j]$, $j = 1, \dots, m$, with $T_0 = 0$, $T_m = T$ and $T_j - T_{j-1} = O(1)$, such that on each subinterval $[T_{j-1}, T_j]$ the deformation is not large, and then re-initialize the mesh at the time levels T_j , $j = 1, \dots, m-1$. This would keep the triangulations shape-regular and quasi-uniform and avoid requiring too small stepsize and mesh size, compared to evolving the mesh over the entire time interval $[0, T]$.

3. Convergence of the numerical approximations

The proof of convergence consists of consistency and stability analysis. In the stability analysis we need to compare two different triangulated domains, i.e, the triangulated domain $\Omega_h[\mathbf{x}_*^n]$ obtained from interpolating the exact domain $\Omega[\phi(\cdot, t_n)]$, and the triangulated domain $\Omega_h[\mathbf{x}^n]$ determined by the numerical solution. In the spirit of the techniques in [32, 13], using bulk domains instead of surfaces, we can obtain a similar sequence of results employing shape derivatives by constructing a homotopy map between the $\Omega_h[\mathbf{x}_*^n]$ and $\Omega_h[\mathbf{x}^n]$. The corresponding results are listed in the following subsection and will be used in the stability analysis in Section 3.3.

3.1. Comparison of norms and integrals on two different domains

Let $\mathbf{y}, \mathbf{z} \in \mathbb{R}^{dN}$ be the two nodal vectors which define the discrete finite element domains $\Omega_h[\mathbf{y}]$ and $\Omega_h[\mathbf{z}]$, respectively. Let $\mathbf{e} = (e_j) := \mathbf{y} - \mathbf{z} \in \mathbb{R}^{dN}$. By means of a linear homology, the intermediate domain $\Omega_h^\theta := \Omega_h[\mathbf{z} + \theta\mathbf{e}]$ changes continuously from $\Omega_h[\mathbf{z}]$ to $\Omega_h[\mathbf{y}]$ when the parameter θ takes values in $[0, 1]$. For a vector $\mathbf{u} = (u_j) \in \mathbb{R}^N$ we denote by $u_h^\theta \in S_h^k[\mathbf{z} + \theta\mathbf{e}]$ the finite element function on $\Omega_h[\mathbf{z} + \theta\mathbf{e}]$ defined by

$$u_h^\theta = \sum_{j=1}^N u_j \phi_j[\mathbf{z} + \theta\mathbf{e}].$$

Similarly to the definition of the scalar-valued function u_h^θ by using a N -dimensional vector \mathbf{u} , we can define a d -dimensional vector-valued function e_h^θ by using the dN -dimensional vector \mathbf{e} .

In combination with the fundamental theorem of calculus, Lemma 2.1 and the transport property (2.8), the following lemma was proved in [13, Lemma 5.1].

Lemma 3.1. *In the above setting the following identities hold:*

$$\mathbf{u}^\top (\mathbf{M}(\mathbf{y}) - \mathbf{M}(\mathbf{z})) \mathbf{v} = \int_0^1 \int_{\Omega_h^\theta} u_h^\theta (\nabla \cdot e_h^\theta) v_h^\theta dx d\theta, \quad (3.1)$$

$$\mathbf{u}^\top (\mathbf{A}(\mathbf{y}) - \mathbf{A}(\mathbf{z})) \mathbf{v} = \int_0^1 \int_{\Omega_h^\theta} \nabla u_h^\theta \cdot (D_{\Omega_h^\theta} e_h^\theta) \nabla v_h^\theta dx d\theta, \quad (3.2)$$

with $D_{\Omega_h^\theta} e_h^\theta = \text{trace}(E)I_d - (E + E^\top)$ for $E = \nabla e_h^\theta \in L^2(\Omega_h^\theta)^{d \times d}$.

The two formulae in above lemma directly show that if ∇e_h^θ is small, the norms of finite element functions on the two finite element domains $\Omega_h[\mathbf{z}]$ and $\Omega_h[\mathbf{y}]$ with same nodal vectors are equivalent. The following lemma was proved in [13, Lemma 5.2].

Lemma 3.2. *If $\|\nabla \cdot e_h^\theta\|_{L^\infty(\Omega_h^\theta)} \leq \mu$ for $0 \leq \theta \leq 1$, then*

$$\|\mathbf{v}\|_{\mathbf{M}(\mathbf{z}+\theta\mathbf{e})} \leq e^{\mu\theta/2} \|\mathbf{v}\|_{\mathbf{M}(\mathbf{z})}.$$

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

$$\|\mathbf{v}\|_{\mathbf{A}(\mathbf{z}+\theta\mathbf{e})} \leq e^{\eta\theta/2} \|\mathbf{v}\|_{\mathbf{A}(\mathbf{z})}.$$

The following lemma was proved in [13, Lemma 5.3], which says that the condition in Lemma 3.2 can be reduced to $\theta = 0$.

Lemma 3.3. *If $\|\nabla e_h^0\|_{L^\infty(\Omega_h^0)} \leq \frac{1}{2}$ then the finite element function v_h^θ on Ω_h^θ , with $0 \leq \theta \leq 1$, satisfies the following estimate:*

$$\|\nabla_{\Omega_h^\theta} v_h^\theta\|_{L^p(\Omega_h^\theta)} \leq c_p \|\nabla_{\Omega_h^0} v_h^0\|_{L^p(\Omega_h^0)} \text{ for } 1 \leq p \leq \infty, \quad (3.3)$$

where c_p depends only on p .

In Lenoir's isoparametric finite element approximation to Ω^0 , there exists a lift map $\Psi : \Omega_h^0 \rightarrow \Omega^0$ satisfying the following estimates (cf. [30]):

$$|\Psi(x) - x| \leq Ch^{k+1} \text{ for } x \in \Omega_h^0 \quad \text{and} \quad \|\nabla \Psi - I\|_{L^\infty(\Omega_h^0)} \leq Ch^k.$$

Let $\phi : \Omega^0 \times [0, T] \rightarrow \mathbb{R}^d$ be the flow map which determines the domain $\Omega(t) = \phi(\Omega^0, t)$, and let $\phi_h^* = I_h \phi : \Omega_h^0 \times [0, T] \rightarrow \mathbb{R}^d$ be the Lagrangian interpolation of the flow map. Let $\mathbf{x}_*(t)$ be the image of \mathbf{x}^0 under the flow map ϕ . Thus $\Omega_h[\mathbf{x}_*(t)]$ is the triangulated domain which approximates $\Omega(t)$ based on the nodes in $\mathbf{x}_*(t)$. Then $\Phi(\cdot, t) = \phi(\cdot, t) \circ \Psi \circ \phi_h^*(\cdot, t)^{-1} : \Omega_h[\mathbf{x}_*(t)] \rightarrow \Omega(t)$ is a lift map at time t such that

$$|\Phi(x, t) - x| \leq Ch^{k+1} \text{ for } x \in \Omega_h[\mathbf{x}_*(t)] \quad \text{and} \quad \|\nabla \Phi(\cdot, t) - I\|_{L^\infty(\Omega_h[\mathbf{x}_*(t)])} \leq Ch^k.$$

Correspondingly, for a finite element function $\chi_h \in S_h^k[\mathbf{x}_*(t)]$ defined on $\Omega_h[\mathbf{x}_*(t)]$, we can define its lift onto $\Omega(t)$ by

$$\tilde{\chi}_h = \chi_h \circ \Phi(\cdot, t)^{-1}.$$

Then the following lemma is a generalization of the geometric estimates in [32] and is used in the proof of the consistency estimates in Lemma 3.5.

Lemma 3.4. *The following estimates hold for $\chi, \psi \in S_h^k[\mathbf{x}_*(t)]$ and $g \in W^{1,\infty}(\mathbb{R}^d)$:*

$$\begin{aligned} \left| \int_{\Omega_h[\mathbf{x}_*(t)]} \chi \psi dx - \int_{\Omega(t)} \tilde{\chi} \tilde{\psi} dx \right| &\leq Ch^k \|\chi\|_{L^2(\Omega_h[\mathbf{x}_*(t)])} \|\psi\|_{L^2(\Omega_h[\mathbf{x}_*(t)])}, \\ \left| \int_{\Omega_h[\mathbf{x}_*(t)]} \nabla \chi \cdot \nabla \psi dx - \int_{\Omega(t)} \nabla \tilde{\chi} \cdot \nabla \tilde{\psi} dx \right| &\leq Ch^k \|\chi\|_{H^1(\Omega_h[\mathbf{x}_*(t)])} \|\psi\|_{H^1(\Omega_h[\mathbf{x}_*(t)])}, \\ \left| \int_{\Omega_h[\mathbf{x}_*(t)]} g \chi dx - \int_{\Omega(t)} g \tilde{\chi} dx \right| &\leq Ch^k \|\chi\|_{L^2(\Omega_h[\mathbf{x}_*(t)])} \|g\|_{W^{1,\infty}(\mathbb{R}^d)}, \\ \left| \int_{\Omega_h[\mathbf{x}_*(t)]} g \nabla \chi dx - \int_{\Omega(t)} g \nabla \tilde{\chi} dx \right| &\leq Ch^k \|\chi\|_{H^1(\Omega_h[\mathbf{x}_*(t)])} \|g\|_{W^{1,\infty}(\mathbb{R}^d)}. \end{aligned}$$

3.2. Error equations and consistency estimates

We compare the Lagrangian interpolations of the exact solution, denoted by $\hat{\phi}_h^n \in S_h^k[\mathbf{x}^0]^d$ and $\hat{w}_h^n \in S_h^k[\mathbf{x}_*^n]^d$ and $\hat{u}_h^n, \hat{p}_h^n \in \hat{S}_h^k[\mathbf{x}_*^n]$, with the numerical solutions ϕ_h^n, w_h^n, u_h^n and p_h^n , after pulling these functions back to the initial domain Ω_h^0 .

The finite element functions $\hat{\phi}_h^n \in S_h^k[\mathbf{x}^0]^d$, $\hat{w}_h^n \in S_h^k[\mathbf{x}_*^n]^d$ and $\hat{u}_h^n, \hat{p}_h^n \in \hat{S}_h^k[\mathbf{x}_*^n]$ satisfy the weak formulations up to some defects:

$$\frac{\hat{\phi}_h^{n+1} - \hat{\phi}_h^n}{\tau} = \hat{w}_h^n \circ \hat{\phi}_h^n + d_\phi^n, \quad (3.4a)$$

$$\int_{\Omega_h[\mathbf{x}_*^n]} \nabla \hat{w}_h^n : \nabla \chi_w + \hat{w}_h^n \cdot \chi_w dx = -dJ(\Gamma_h[\mathbf{x}_*^n], \hat{u}_h^n, \hat{p}_h^n; \chi_w) + \int_{\Omega_h[\mathbf{x}_*^n]} d_w^m \cdot \chi_w dx, \quad (3.4b)$$

$$\int_{\Omega_h[\mathbf{x}_*^n]} \nabla \hat{u}_h^n \cdot \nabla \chi_u dx = \int_{\Omega_h[\mathbf{x}_*^n]} f \chi_u dx + \int_{\Omega_h[\mathbf{x}_*^n]} d_u^m \chi_u dx, \quad (3.4c)$$

$$\int_{\Omega_h[\mathbf{x}_*^n]} \nabla \hat{p}_h^n \cdot \nabla \chi_p dx = \int_{\Omega_h[\mathbf{x}_*^n]} j'_u(x, \hat{u}_h^n) \chi_p dx + \int_{\Omega_h[\mathbf{x}_*^n]} d_p^m \chi_p dx, \quad (3.4d)$$

for test functions $\chi_w \in S_h^k[\mathbf{x}_*^n]^d$ and $\chi_u, \chi_p \in \hat{S}_h^k[\mathbf{x}_*^n]$, where $d_\phi^n \in S_h^k[\mathbf{x}^0]^d$, $d_w^m \in S_h^k[\mathbf{x}_*^n]^d$ and $d_u^m, d_p^m \in S_h^k[\mathbf{x}_*^n]$ are defects (consistency errors).

In the computation and analysis it is more convenient to write the above linear systems into the matrix-vector form. To this end, we denote by $\mathbf{w}_*^n, \mathbf{u}_*^n, \mathbf{p}_*^n, \mathbf{d}_x^n, \mathbf{d}_w^n, \mathbf{d}_u^n$ and \mathbf{d}_p^n the column vectors that collect the nodal values of $\hat{w}_h^n, \hat{u}_h^n, \hat{p}_h^n, d_\phi^n, d_w^m, d_u^m$ and d_p^m , respectively. The right-hand side vectors $\mathbf{f}(\mathbf{x}_*^n)$, $\mathbf{J}'_u(\mathbf{x}_*^n, \mathbf{u}_*^n)$ and $-\mathbf{dJ}(\mathbf{x}_*^n, \mathbf{u}_*^n, \mathbf{p}_*^n)$ are defined by

$$\begin{aligned} \mathbf{f}(\mathbf{x}_*^n)|_j &= \int_{\Omega_h[\mathbf{x}_*^n]} f \phi_j[\mathbf{x}_*^n] dx, \\ \mathbf{J}'_u(\mathbf{x}_*^n, \mathbf{u}_*^n) &= \int_{\Omega_h[\mathbf{x}_*^n]} j'_u(x, \hat{u}_h^n) \phi_j[\mathbf{x}_*^n] dx, \end{aligned}$$

$$\begin{aligned} \mathbf{dJ}(\mathbf{x}_*^n, \mathbf{u}_*^n, \mathbf{p}_*^n)|_{j+N(l-1)} &= \int_{\Omega_h[\mathbf{x}_*^n]} (\nabla\phi_j[\mathbf{x}_*^n] \cdot \nabla\hat{u}_h^n - f\phi_j[\mathbf{x}_*^n]) (\nabla\hat{p}_h^n)_l + \nabla\phi_j[\mathbf{x}_*^n] \cdot \nabla\hat{p}_h^n (\nabla\hat{u}_h^n)_l \\ &\quad + (j(x, \hat{u}_h^n) - \nabla\hat{u}_h^n \cdot \nabla\hat{p}_h^n) (\nabla\phi_j[\mathbf{x}_*^n])_l - j'_u(x, \hat{u}_h^n) (\nabla u_d)_l \phi_j[\mathbf{x}_*^n] dx, \end{aligned}$$

for $j = 1, \dots, N$ and $1 \leq l \leq d$. Then (3.4) can be written into the following matrix-vector form:

$$\mathbf{x}_*^{n+1} - \mathbf{x}_*^n = \tau \mathbf{w}_*^n + \tau \mathbf{d}_x^n, \quad (3.5a)$$

$$\mathbf{K}(\mathbf{x}_*^n) \mathbf{w}_*^n = -\mathbf{dJ}(\mathbf{x}_*^n, \mathbf{u}_*^n, \mathbf{p}_*^n) + \mathbf{M}^{[d]}(\mathbf{x}_*^n) \mathbf{d}_w^n, \quad (3.5b)$$

$$\mathbf{A}(\mathbf{x}_*^n) \mathbf{u}_*^n = \mathbf{f}(\mathbf{x}_*^n) + \mathbf{M}(\mathbf{x}_*^n) \mathbf{d}_u^n, \quad (3.5c)$$

$$\mathbf{A}(\mathbf{x}_*^n) \mathbf{p}_*^n = \mathbf{J}'_u(\mathbf{x}_*^n, \mathbf{u}_*^n) + \mathbf{M}(\mathbf{x}_*^n) \mathbf{d}_p^n, \quad (3.5d)$$

with $\mathbf{M}^{[d]}(\mathbf{x}_*^n) = I_d \otimes \mathbf{M}(\mathbf{x}_*^n)$. When no confusion arises, we simply write $\mathbf{M}(\mathbf{x}_*^n)$ for $\mathbf{M}^{[d]}(\mathbf{x}_*^n)$ and $\|\cdot\|_{H^1(\Omega)}$ for $\|\cdot\|_{H^1(\Omega)^d}$ throughout.

The H^{-1} norm of the defect d_w^n will be used in stability analysis. It has the following expression:

$$\begin{aligned} \|d_w^n\|_{H_h^{-1}(\Omega_h[\mathbf{x}_*^n])} &:= \sup_{0 \neq \psi_h \in S_h^k[\mathbf{x}_*^n]^d} \|\psi_h\|_{H^1(\Omega_h[\mathbf{x}_*^n])}^{-1} \int_{\Omega_h[\mathbf{x}_*^n]} d_w^n \cdot \psi_h dx \\ &= \sup_{0 \neq \mathbf{z} \in \mathbb{R}^{dN}} \frac{(\mathbf{d}_w^n)^\top \mathbf{M}(\mathbf{x}_*^n) \mathbf{z}}{(\mathbf{z}^\top \mathbf{K}(\mathbf{x}_*^n) \mathbf{z})^{\frac{1}{2}}} = \sup_{0 \neq \mathbf{z} \in \mathbb{R}^{dN}} \frac{(\mathbf{d}_w^n)^\top \mathbf{M}(\mathbf{x}_*^n) \mathbf{K}(\mathbf{x}_*^n)^{-\frac{1}{2}} \mathbf{K}(\mathbf{x}_*^n)^{\frac{1}{2}} \mathbf{z}}{(\mathbf{z}^\top \mathbf{K}(\mathbf{x}_*^n) \mathbf{z})^{\frac{1}{2}}} \\ &= \|\mathbf{K}(\mathbf{x}_*^n)^{-\frac{1}{2}} \mathbf{M}(\mathbf{x}_*^n) \mathbf{d}_w^n\|_2 = \left((\mathbf{d}_w^n)^\top \mathbf{M}(\mathbf{x}_*^n) \mathbf{K}(\mathbf{x}_*^n)^{-1} \mathbf{M}(\mathbf{x}_*^n) \mathbf{d}_w^n \right)^{\frac{1}{2}}. \end{aligned}$$

Correspondingly, in the matrix-vector notation, we denote the discrete dual norm of \mathbf{d}_w^n by

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

Similarly, we denote the discrete dual norms of $\mathbf{d}_u^n, \mathbf{d}_p^n \in \mathbb{R}^N$ by

$$\|\mathbf{d}_u^n\|_{*, \mathbf{x}_*^n}^2 := (\mathbf{d}_u^n)^\top \mathbf{M}(\mathbf{x}_*^n) \mathbf{K}(\mathbf{x}_*^n)^{-1} \mathbf{M}(\mathbf{x}_*^n) \mathbf{d}_u^n = \|d_u^n\|_{H_h^{-1}(\Omega_h[\mathbf{x}_*^n])}^2,$$

$$\|\mathbf{d}_p^n\|_{*, \mathbf{x}_*^n}^2 := (\mathbf{d}_p^n)^\top \mathbf{M}(\mathbf{x}_*^n) \mathbf{K}(\mathbf{x}_*^n)^{-1} \mathbf{M}(\mathbf{x}_*^n) \mathbf{d}_p^n = \|d_p^n\|_{H_h^{-1}(\Omega_h[\mathbf{x}_*^n])}^2.$$

The stability estimates will be established by comparing the matrix-vector formulations (2.11) and (3.5). By subtracting (3.5) from (2.11), we obtain the following equations for the errors $\mathbf{e}_x^n = \mathbf{x}^n - \mathbf{x}_*^n$, $\mathbf{e}_w^n = \mathbf{w}^n - \mathbf{w}_*^n$, $\mathbf{e}_u^n = \mathbf{u}^n - \mathbf{u}_*^n$ and $\mathbf{e}_p^n = \mathbf{p}^n - \mathbf{p}_*^n$:

$$\mathbf{e}_x^{n+1} = \mathbf{e}_x^n + \tau \mathbf{e}_w^n - \tau \mathbf{d}_x^n, \quad (3.6a)$$

$$\begin{aligned} \mathbf{K}(\mathbf{x}_*^n) \mathbf{e}_w^n &= -(\mathbf{K}(\mathbf{x}^n) - \mathbf{K}(\mathbf{x}_*^n)) \mathbf{e}_w^n - (\mathbf{K}(\mathbf{x}^n) - \mathbf{K}(\mathbf{x}_*^n)) \mathbf{w}_*^n \\ &\quad - (\mathbf{dJ}(\mathbf{x}^n, \mathbf{u}^n, \mathbf{p}^n) - \mathbf{dJ}(\mathbf{x}_*^n, \mathbf{u}_*^n, \mathbf{p}_*^n)) - \mathbf{M}(\mathbf{x}_*^n) \mathbf{d}_w^n, \end{aligned} \quad (3.6b)$$

$$\begin{aligned} \mathbf{A}(\mathbf{x}_*^n) \mathbf{e}_u^n &= -(\mathbf{A}(\mathbf{x}^n) - \mathbf{A}(\mathbf{x}_*^n)) \mathbf{e}_u^n - (\mathbf{A}(\mathbf{x}^n) - \mathbf{A}(\mathbf{x}_*^n)) \mathbf{u}_*^n \\ &\quad + (\mathbf{f}(\mathbf{x}^n) - \mathbf{f}(\mathbf{x}_*^n)) - \mathbf{M}(\mathbf{x}_*^n) \mathbf{d}_u^n, \end{aligned} \quad (3.6c)$$

$$\begin{aligned} \mathbf{A}(\mathbf{x}_*^n) \mathbf{e}_p^n &= -(\mathbf{A}(\mathbf{x}^n) - \mathbf{A}(\mathbf{x}_*^n)) \mathbf{e}_p^n - (\mathbf{A}(\mathbf{x}^n) - \mathbf{A}(\mathbf{x}_*^n)) \mathbf{p}_*^n \\ &\quad + (\mathbf{J}'_u(\mathbf{x}^n, \mathbf{u}^n) - \mathbf{J}'_u(\mathbf{x}_*^n, \mathbf{u}_*^n)) - \mathbf{M}(\mathbf{x}_*^n) \mathbf{d}_p^n. \end{aligned} \quad (3.6d)$$

The error estimates depend on the estimates for the defect terms \mathbf{d}_x^n , \mathbf{d}_w^n , \mathbf{d}_u^n and \mathbf{d}_p^n (the consistency errors), which are presented in the following lemma. The proof of this lemma is omitted as it basically follows from the approximation properties of the Lagrangian interpolation, Taylor's formula and Lemma 3.4.

Lemma 3.5 (Consistency estimates). *Under the assumptions of Theorem 2.3, there exist positive constants τ_0 and h_0 such that for $\tau \leq \tau_0$ and $h \leq h_0$, the following consistency error estimates hold:*

$$\sup_{1 \leq n \leq [T/\tau]} \|\mathbf{d}_x^n\|_{\mathbf{K}(\mathbf{x}_*^n)} \leq C\tau, \quad (3.7a)$$

$$\sup_{1 \leq n \leq [T/\tau]} (\|\mathbf{d}_w^n\|_{*,\mathbf{x}_*^n} + \|\mathbf{d}_u^n\|_{*,\mathbf{x}_*^n} + \|\mathbf{d}_p^n\|_{*,\mathbf{x}_*^n}) \leq Ch^k. \quad (3.7b)$$

3.3. Stability estimates

From the error equations in (3.6) and the consistency estimates in Lemma 3.5, we can derive the following stability estimates.

Proposition 3.6 (Stability estimates). *Under the assumptions of Theorem 2.3, there exists a positive constant h_0 such that for $\tau = o(h^{\frac{d}{2}})$ and $h \leq h_0$ the following stability estimate holds:*

$$\begin{aligned} & \sup_{0 \leq n \leq [T/\tau]} (\|\mathbf{e}_x^n\|_{\mathbf{K}(\mathbf{x}_*^n)} + \|\mathbf{e}_w^n\|_{\mathbf{K}(\mathbf{x}_*^n)} + \|\mathbf{e}_u^n\|_{\mathbf{K}(\mathbf{x}_*^n)} + \|\mathbf{e}_p^n\|_{\mathbf{K}(\mathbf{x}_*^n)}) \\ & \leq C \sup_{0 \leq n \leq [T/\tau]} (\|\mathbf{d}_x^n\|_{\mathbf{K}(\mathbf{x}_*^n)} + \|\mathbf{d}_w^n\|_{*,\mathbf{x}_*^n} + \|\mathbf{d}_u^n\|_{*,\mathbf{x}_*^n} + \|\mathbf{d}_p^n\|_{*,\mathbf{x}_*^n}), \end{aligned} \quad (3.8)$$

where C is independent of τ , h and n (but may depend on T).

Proof. Let $e_x^n \in S_h^k[\mathbf{x}_*^n]^d$, $e_w^n \in S_h^k[\mathbf{x}_*^n]^d$ and $e_u^n, e_p^n \in \tilde{S}_h^k[\mathbf{x}_*^n]$ be the finite element error functions on $\Omega_h[\mathbf{x}_*^n]$ with nodal vectors \mathbf{e}_x^n , \mathbf{e}_w^n , \mathbf{e}_u^n and \mathbf{e}_p^n , respectively. We make the following mathematical induction on the boundedness of the errors: We assume that there exists an integer $1 \leq m \leq [T/\tau]$ such that the following inequalities hold for all $0 \leq n \leq m-1$,

$$\|e_x^n\|_{W^{1,\infty}(\Omega_h[\mathbf{x}_*^n])} \leq h^{-\frac{d}{4}}(\tau^{\frac{1}{2}} + h^{\frac{k}{2}}), \quad (3.9a)$$

$$\|e_u^n\|_{W^{1,\infty}(\Omega_h[\mathbf{x}_*^n])} \leq 1, \quad (3.9b)$$

$$\|e_p^n\|_{W^{1,\infty}(\Omega_h[\mathbf{x}_*^n])} \leq 1. \quad (3.9c)$$

In fact, the inequalities above hold at least for $m=1$ because of the following two reasons:

- (i) Since $\mathbf{x}_*^0 = \mathbf{x}^0$, it follows that $e_x^0 = 0$.
- (ii) Testing (3.6c) and (3.6d) with \mathbf{e}_u^0 and \mathbf{e}_p^0 , respectively, yields the following relations:

$$\begin{aligned} & (\mathbf{e}_u^0)^\top \mathbf{A}(\mathbf{x}_*^0) \mathbf{e}_u^0 = -(\mathbf{e}_u^0)^\top \mathbf{M}(\mathbf{x}_*^0) \mathbf{d}_u^0, \\ & (\mathbf{e}_p^0)^\top \mathbf{A}(\mathbf{x}_*^0) \mathbf{e}_p^0 = (\mathbf{e}_p^0)^\top (\mathbf{J}'_u(\mathbf{x}^0, \mathbf{u}^0) - \mathbf{J}'_u(\mathbf{x}_*^0, \mathbf{u}_*^0)) - (\mathbf{e}_p^0)^\top \mathbf{M}(\mathbf{x}_*^0) \mathbf{d}_p^0. \end{aligned}$$

By Poincaré's inequality, $\|\cdot\|_{\mathbf{A}(\mathbf{x}_*^n)}$ and $\|\cdot\|_{\mathbf{K}(\mathbf{x}_*^n)}$ are equivalent for functions in $H_0^1(\Omega_h[\mathbf{x}_*^n])$, and the equivalence is independent of h since there is a one-to-one $W^{1,\infty}$ -uniformly bounded lift map from $\Omega_h[\mathbf{x}_*^n]$ onto $\Omega[\mathbf{x}(t_n)]$. Therefore, the relations above together with Cauchy-Schwarz inequality imply that

$$\begin{aligned} & \|\mathbf{e}_u^0\|_{\mathbf{K}(\mathbf{x}_*^0)} \leq C \|\mathbf{d}_u^0\|_{*,\mathbf{x}_*^0} \leq Ch^k \\ & (\mathbf{e}_p^0)^\top (\mathbf{J}'_u(\mathbf{x}^0, \mathbf{u}^0) - \mathbf{J}'_u(\mathbf{x}_*^0, \mathbf{u}_*^0)) = \int_{\Omega_h^0} e_p^0 (u_h^0 - \hat{u}_h^0) dx \leq Ch^{k+1} \|\mathbf{e}_p^0\|_{\mathbf{M}(\mathbf{x}_*^0)}, \end{aligned}$$

and therefore

$$\|\mathbf{e}_p^0\|_{\mathbf{K}(\mathbf{x}_*^0)} \leq Ch^k.$$

By the inverse inequality of finite element functions, we have

$$\begin{aligned} & \|e_u^0\|_{W^{1,\infty}(\Omega_h[\mathbf{x}_*^0])} \leq Ch^{-\frac{d}{2}} \|\mathbf{e}_u^0\|_{\mathbf{K}(\mathbf{x}_*^0)} \leq Ch^{k-\frac{d}{2}}, \\ & \|e_p^0\|_{W^{1,\infty}(\Omega_h[\mathbf{x}_*^0])} \leq Ch^{-\frac{d}{2}} \|\mathbf{e}_p^0\|_{\mathbf{K}(\mathbf{x}_*^0)} \leq Ch^{k-\frac{d}{2}}. \end{aligned}$$

Therefore, in the case $k \geq 2 > d/2$, assumptions (3.9b)–(3.9c) hold for $n=0$ when h is sufficiently small.

Now we prove that the stated error bounds (3.8) hold for all time levels $0 \leq n \leq m$ under the induction assumption. The mathematical induction will be completed by proving that (3.9) also holds for $n=m$.

(A) *Estimates for \mathbf{e}_x^n* : First, multiplying (3.6a) by matrix $\mathbf{K}(\mathbf{x}_*^n)$, we have

$$\mathbf{K}(\mathbf{x}_*^n) \mathbf{e}_x^n = \mathbf{K}(\mathbf{x}_*^n) \mathbf{e}_x^{n-1} + \tau \mathbf{K}(\mathbf{x}_*^n) \mathbf{e}_w^{n-1} - \tau \mathbf{K}(\mathbf{x}_*^n) \mathbf{d}_x^{n-1}, \text{ for } 1 \leq n \leq m.$$

Then, testing the equation above by \mathbf{e}_x^n , the following relation is derived:

$$\begin{aligned} \|\mathbf{e}_x^n\|_{\mathbf{K}(\mathbf{x}_*^n)}^2 &= (\mathbf{e}_x^n)^\top \mathbf{K}(\mathbf{x}_*^n) \mathbf{e}_x^{n-1} + \tau (\mathbf{e}_x^n)^\top \mathbf{K}(\mathbf{x}_*^n) \mathbf{e}_w^{n-1} - \tau (\mathbf{e}_x^n)^\top \mathbf{K}(\mathbf{x}_*^n) \mathbf{d}_x^{n-1} \\ &\leq \|\mathbf{e}_x^n\|_{\mathbf{K}(\mathbf{x}_*^n)} \|\mathbf{e}_x^{n-1}\|_{\mathbf{K}(\mathbf{x}_*^n)} + \tau \|\mathbf{e}_x^n\|_{\mathbf{K}(\mathbf{x}_*^n)} \left(\|\mathbf{e}_w^{n-1}\|_{\mathbf{K}(\mathbf{x}_*^n)} + \|\mathbf{d}_x^{n-1}\|_{\mathbf{K}(\mathbf{x}_*^n)} \right). \end{aligned}$$

Next, by dividing both sides by $\|\mathbf{e}_x^n\|_{\mathbf{K}(\mathbf{x}_*^n)}$, we obtain

$$\|\mathbf{e}_x^n\|_{\mathbf{K}(\mathbf{x}_*^n)} \leq \|\mathbf{e}_x^{n-1}\|_{\mathbf{K}(\mathbf{x}_*^n)} + \tau \|\mathbf{e}_w^{n-1}\|_{\mathbf{K}(\mathbf{x}_*^n)} + \tau \|\mathbf{d}_x^{n-1}\|_{\mathbf{K}(\mathbf{x}_*^n)}. \quad (3.10)$$

In order to iterate the inequality above with respect to n , we need to convert $\|\mathbf{e}_x^{n-1}\|_{\mathbf{K}(\mathbf{x}_*^n)}$ to $\|\mathbf{e}_x^{n-1}\|_{\mathbf{K}(\mathbf{x}_*^{n-1})}$ on the right-hand side by utilizing Lemmas 3.1–3.3, which imply the following result:

$$\|\mathbf{e}_x^{n-1}\|_{\mathbf{K}(\mathbf{x}_*^n)}^2 - \|\mathbf{e}_x^{n-1}\|_{\mathbf{K}(\mathbf{x}_*^{n-1})}^2 \leq C\tau \|\mathbf{e}_x^{n-1}\|_{\mathbf{K}(\mathbf{x}_*^{n-1})}^2. \quad (3.11)$$

Now, we can substitute (3.11) into (3.10) and sum up the result with respect to n . This yields the following estimate:

$$\|\mathbf{e}_x^n\|_{\mathbf{K}(\mathbf{x}_*^n)} \leq C\tau \sum_{l=1}^n \left(\|\mathbf{e}_w^{l-1}\|_{\mathbf{K}(\mathbf{x}_*^{l-1})} + \|\mathbf{d}_x^{l-1}\|_{\mathbf{K}(\mathbf{x}_*^{l-1})} \right) + C\tau \sum_{l=1}^n \|\mathbf{e}_x^{l-1}\|_{\mathbf{K}(\mathbf{x}_*^{l-1})} \quad (3.12)$$

for $1 \leq n \leq m$.

Similarly, testing equations (3.6b), (3.6c) and (3.6d) with \mathbf{e}_w^n , \mathbf{e}_u^n and \mathbf{e}_p^n , respectively, we obtain

$$\|\mathbf{e}_w^n\|_{\mathbf{K}(\mathbf{x}_*^n)}^2 = -(\mathbf{e}_w^n)^\top (\mathbf{K}(\mathbf{x}^n) - \mathbf{K}(\mathbf{x}_*^n)) \mathbf{e}_w^n - (\mathbf{e}_w^n)^\top (\mathbf{K}(\mathbf{x}^n) - \mathbf{K}(\mathbf{x}_*^n)) \mathbf{w}_*^n \quad (3.13a)$$

$$- (\mathbf{e}_w^n)^\top (\mathbf{d}\mathbf{J}(\mathbf{x}^n, \mathbf{u}^n, \mathbf{p}^n) - \mathbf{d}\mathbf{J}(\mathbf{x}_*^n, \mathbf{u}_*^n, \mathbf{p}_*^n)) - (\mathbf{e}_w^n)^\top \mathbf{M}(\mathbf{x}_*^n) \mathbf{d}_w^n,$$

$$\|\mathbf{e}_u^n\|_{\mathbf{A}(\mathbf{x}_*^n)}^2 = -(\mathbf{e}_u^n)^\top (\mathbf{A}(\mathbf{x}^n) - \mathbf{A}(\mathbf{x}_*^n)) \mathbf{e}_u^n - (\mathbf{e}_u^n)^\top (\mathbf{A}(\mathbf{x}^n) - \mathbf{A}(\mathbf{x}_*^n)) \mathbf{u}_*^n \quad (3.13b)$$

$$+ (\mathbf{e}_u^n)^\top (\mathbf{f}(\mathbf{x}^n) - \mathbf{f}(\mathbf{x}_*^n)) - (\mathbf{e}_u^n)^\top \mathbf{M}(\mathbf{x}_*^n) \mathbf{d}_u^n,$$

$$\|\mathbf{e}_p^n\|_{\mathbf{A}(\mathbf{x}_*^n)}^2 = -(\mathbf{e}_p^n)^\top (\mathbf{A}(\mathbf{x}^n) - \mathbf{A}(\mathbf{x}_*^n)) \mathbf{e}_p^n - (\mathbf{e}_p^n)^\top (\mathbf{A}(\mathbf{x}^n) - \mathbf{A}(\mathbf{x}_*^n)) \mathbf{p}_*^n \quad (3.13c)$$

$$+ (\mathbf{e}_p^n)^\top (\mathbf{J}'_u(\mathbf{x}^n, \mathbf{u}^n) - \mathbf{J}'_u(\mathbf{x}_*^n, \mathbf{u}_*^n)) - (\mathbf{e}_p^n)^\top \mathbf{M}(\mathbf{x}_*^n) \mathbf{d}_p^n.$$

In the following, we present estimates for \mathbf{e}_u^n , \mathbf{e}_p^n and \mathbf{e}_w^n by using the equations in (3.13).

(B) *Estimates for \mathbf{e}_u^n under assumption (3.9a)*: We estimate the four terms on the right-hand side of (3.13b) separately by using Lemmas 3.1–3.3. For $\theta \in [0, 1]$, we denote by $e_{x,\theta}^n, e_{u,\theta}^n, \hat{u}_{h,\theta}^n$ the finite element functions on the intermediate domain $\Omega_h[\mathbf{x}_\theta^n]$ for $\mathbf{x}_\theta^n = \mathbf{x}_*^n + \theta \mathbf{e}_x^n$ with nodal vectors $\mathbf{e}_x^n, \mathbf{e}_u^n, \mathbf{u}_*^n$, respectively.

The first term on the right hand side of (3.13b) can be estimated by using Lemma 3.1 and Hölder's inequality, i.e.,

$$\begin{aligned} (\mathbf{e}_u^n)^\top (\mathbf{A}(\mathbf{x}^n) - \mathbf{A}(\mathbf{x}_*^n)) \mathbf{e}_u^n &= \int_0^1 \int_{\Omega_h[\mathbf{x}_\theta^n]} \nabla e_{u,\theta}^n \cdot (D_{\Omega_h[\mathbf{x}_\theta^n]} e_{x,\theta}^n) \nabla e_{u,\theta}^n dx d\theta \\ &\leq \int_0^1 \|\nabla e_{u,\theta}^n\|_{L^2(\Omega_h[\mathbf{x}_\theta^n])}^2 \|D_{\Omega_h[\mathbf{x}_\theta^n]} e_{x,\theta}^n\|_{L^\infty(\Omega_h[\mathbf{x}_\theta^n])} d\theta. \end{aligned}$$

Under the induction assumption in (3.9a) we obtain from Lemma 3.3 that

$$(\mathbf{e}_u^n)^\top (\mathbf{A}(\mathbf{x}^n) - \mathbf{A}(\mathbf{x}_*^n)) \mathbf{e}_u^n \leq Ch^{-\frac{d}{4}} (\tau^{\frac{1}{2}} + h^{\frac{k}{2}}) \|\mathbf{e}_u^n\|_{\mathbf{A}(\mathbf{x}_*^n)}^2 \text{ for } 0 \leq n \leq m-1.$$

The second term on the right hand side of (3.13b) can be estimated similarly, i.e.,

$$\begin{aligned} (\mathbf{e}_u^n)^\top (\mathbf{A}(\mathbf{x}^n) - \mathbf{A}(\mathbf{x}_*^n)) \mathbf{u}_*^n &\leq \int_0^1 \|\nabla e_{u,\theta}^n\|_{L^2(\Omega_h[\mathbf{x}_\theta^n])} \|\nabla e_{x,\theta}^n\|_{L^2(\Omega_h[\mathbf{x}_\theta^n])} \|\nabla \hat{u}_{h,\theta}^n\|_{L^\infty(\Omega_h[\mathbf{x}_\theta^n])} d\theta \\ &\leq C \|\nabla e_u^n\|_{L^2(\Omega_h[\mathbf{x}_*^n])} \|\nabla e_x^n\|_{L^2(\Omega_h[\mathbf{x}_*^n])} \|\hat{u}_h^n\|_{W^{1,\infty}(\Omega_h[\mathbf{x}_*^n])} \\ &\leq C \|\mathbf{e}_u^n\|_{\mathbf{A}(\mathbf{x}_*^n)} \|\mathbf{e}_x^n\|_{\mathbf{A}(\mathbf{x}_*^n)}, \end{aligned}$$

where the second to last inequality follows from Lemma 3.3.

The third term on the right hand side of (3.13b) can be written as

$$\begin{aligned} (\mathbf{e}_{\mathbf{u}}^n)^\top (\mathbf{f}(\mathbf{x}^n) - \mathbf{f}(\mathbf{x}_*^n)) &= \int_{\Omega_h[\mathbf{x}_1^n]} f e_{u,1}^n dx - \int_{\Omega_h[\mathbf{x}_0^n]} f e_{u,0}^n dx \\ &= \int_0^1 \frac{d}{d\theta} \int_{\Omega_h[\mathbf{x}_\theta^n]} f e_{u,\theta}^n dx d\theta \\ &= \int_0^1 \int_{\Omega_h[\mathbf{x}_\theta^n]} (\partial_\theta^\bullet (f e_{u,\theta}^n) + f e_{u,\theta}^n \nabla \cdot e_{x,\theta}^n) dx d\theta, \end{aligned}$$

where ∂_θ^\bullet denotes the material derivative with respect to θ . By using the transport property $\partial_\theta^\bullet e_{u,\theta}^n = 0$ we have $\partial_\theta^\bullet (f e_{u,\theta}^n) = e_{u,\theta}^n \partial_\theta^\bullet f$. Since f is a function of only x , it follows that $\partial_\theta^\bullet f = \nabla f \cdot e_{x,\theta}^n$. This gives the expression $\partial_\theta^\bullet (f e_{u,\theta}^n) = e_{u,\theta}^n \nabla f \cdot e_{x,\theta}^n$. Therefore, by using Hölder's inequality, we obtain

$$\begin{aligned} (\mathbf{e}_{\mathbf{u}}^n)^\top (\mathbf{f}(\mathbf{x}^n) - \mathbf{f}(\mathbf{x}_*^n)) &= \int_0^1 \int_{\Omega_h[\mathbf{x}_\theta^n]} (e_{u,\theta}^n \nabla f \cdot e_{x,\theta}^n + f e_{u,\theta}^n \nabla \cdot e_{x,\theta}^n) dx d\theta \\ &\leq \int_0^1 \|e_{u,\theta}^n\|_{L^2(\Omega_h[\mathbf{x}_\theta^n])} \|\nabla f\|_{L^\infty(\Omega_h[\mathbf{x}_\theta^n])} \|e_{x,\theta}^n\|_{L^2(\Omega_h[\mathbf{x}_\theta^n])} d\theta \\ &\quad + \int_0^1 \|f\|_{L^\infty(\Omega_h[\mathbf{x}_\theta^n])} \|e_{u,\theta}^n\|_{L^2(\Omega_h[\mathbf{x}_\theta^n])} \|\nabla \cdot e_{x,\theta}^n\|_{L^2(\Omega_h[\mathbf{x}_\theta^n])} d\theta. \end{aligned}$$

In view of Lemmas 3.2 and 3.3, we have

$$(\mathbf{e}_{\mathbf{u}}^n)^\top (\mathbf{f}(\mathbf{x}^n) - \mathbf{f}(\mathbf{x}_*^n)) \leq C \|\mathbf{e}_{\mathbf{u}}^n\|_{\mathbf{M}(\mathbf{x}_*^n)} \|\mathbf{e}_{\mathbf{x}}^n\|_{\mathbf{K}(\mathbf{x}_*^n)}.$$

The fourth term on the right hand side of (3.13b) can be estimated directly as follows:

$$(\mathbf{e}_{\mathbf{u}}^n)^\top \mathbf{M}(\mathbf{x}_*^n) \mathbf{d}_{\mathbf{u}}^n \leq \|\mathbf{e}_{\mathbf{u}}^n\|_{\mathbf{K}(\mathbf{x}_*^n)} \|\mathbf{d}_{\mathbf{u}}^n\|_{*,\mathbf{x}_*^n}.$$

Combining the four estimates above, we derive that

$$\|\mathbf{e}_{\mathbf{u}}^n\|_{\mathbf{A}(\mathbf{x}_*^n)}^2 \leq C h^{-\frac{d}{4}} (\tau^{\frac{1}{2}} + h^{\frac{k}{2}}) \|\mathbf{e}_{\mathbf{u}}^n\|_{\mathbf{A}(\mathbf{x}_*^n)}^2 + C \|\mathbf{e}_{\mathbf{u}}^n\|_{\mathbf{K}(\mathbf{x}_*^n)} \|\mathbf{e}_{\mathbf{x}}^n\|_{\mathbf{K}(\mathbf{x}_*^n)} + \|\mathbf{e}_{\mathbf{u}}^n\|_{\mathbf{K}(\mathbf{x}_*^n)} \|\mathbf{d}_{\mathbf{u}}^n\|_{*,\mathbf{x}_*^n}.$$

Since $\|\cdot\|_{\mathbf{A}(\mathbf{x}_*^n)}$ and $\|\cdot\|_{\mathbf{K}(\mathbf{x}_*^n)}$ are equivalent for functions in $\mathring{S}_h^k[\mathbf{x}_*^n]$, when $\tau = o(h^{\frac{d}{2}})$ and h is sufficiently small so that $C h^{-\frac{d}{2}} (\tau + h^k) \leq \frac{1}{4}$, we have

$$\|\mathbf{e}_{\mathbf{u}}^n\|_{\mathbf{K}(\mathbf{x}_*^n)} \leq C \|\mathbf{e}_{\mathbf{x}}^n\|_{\mathbf{K}(\mathbf{x}_*^n)} + C \|\mathbf{d}_{\mathbf{u}}^n\|_{*,\mathbf{x}_*^n} \quad \text{for } 0 \leq n \leq m-1. \quad (3.14)$$

(C) *Estimates for $\mathbf{e}_{\mathbf{p}}^n$ under assumption (3.9a)*: The estimation of the first, second and fourth term on the right-hand side of (3.13c) are similar as that in part (B) of this proof and therefore omitted. We focus on the estimation of $(\mathbf{e}_{\mathbf{p}}^n)^\top (\mathbf{J}'_u(\mathbf{x}^n, \mathbf{u}^n) - \mathbf{J}'_u(\mathbf{x}_*^n, \mathbf{u}_*^n))$. Let $e_{p,\theta}^n$ and $u_{h,\theta}^n$ be the finite element functions in $\mathring{S}_h^k[\mathbf{x}_\theta^n]$ with nodal vectors $\mathbf{e}_{\mathbf{p}}^n$ and $\mathbf{u}_*^n + \theta \mathbf{e}_{\mathbf{u}}^n$, i.e., $u_{h,\theta}^n = \hat{u}_{h,\theta}^n + \theta e_{u,\theta}^n$. Then

$$\begin{aligned} &(\mathbf{e}_{\mathbf{p}}^n)^\top (\mathbf{J}'_u(\mathbf{x}^n, \mathbf{u}^n) - \mathbf{J}'_u(\mathbf{x}_*^n, \mathbf{u}_*^n)) \\ &= \int_0^1 \frac{d}{d\theta} \int_{\Omega_h[\mathbf{x}_\theta^n]} j'_u(x, u_{h,\theta}^n) e_{p,\theta}^n dx d\theta \\ &= \int_0^1 \int_{\Omega_h[\mathbf{x}_\theta^n]} (e_{u,\theta}^n - \nabla u_d \cdot e_{x,\theta}^n) e_{p,\theta}^n + (u_{h,\theta}^n - u_d) e_{p,\theta}^n \nabla \cdot e_{x,\theta}^n dx d\theta. \end{aligned}$$

By Lemma 3.3 and Hölder's inequality, and induction assumption (3.9a), we have

$$\begin{aligned} &(\mathbf{e}_{\mathbf{p}}^n)^\top (\mathbf{J}'_u(\mathbf{x}^n, \mathbf{u}^n) - \mathbf{J}'_u(\mathbf{x}_*^n, \mathbf{u}_*^n)) \\ &\leq C (\|\mathbf{e}_{\mathbf{u}}^n\|_{\mathbf{M}(\mathbf{x}_*^n)} + \|\mathbf{e}_{\mathbf{x}}^n\|_{\mathbf{M}(\mathbf{x}_*^n)}) \|\mathbf{e}_{\mathbf{p}}^n\|_{\mathbf{M}(\mathbf{x}_*^n)} + C \|\mathbf{e}_{\mathbf{p}}^n\|_{\mathbf{M}(\mathbf{x}_*^n)} \|\mathbf{e}_{\mathbf{x}}^n\|_{\mathbf{A}(\mathbf{x}_*^n)} \\ &\quad + C \|\mathbf{e}_{\mathbf{u}}^n\|_{\mathbf{M}(\mathbf{x}_*^n)} \|\mathbf{e}_{\mathbf{p}}^n\|_{\mathbf{M}(\mathbf{x}_*^n)} \|e_{x,\theta}^n\|_{W^{1,\infty}(\Omega_h[\mathbf{x}_\theta^n])}. \end{aligned}$$

Similarly as the estimation of $\mathbf{e}_{\mathbf{u}}^n$, by Cauchy-Schwartz inequality, we obtain that

$$\|\mathbf{e}_{\mathbf{p}}^n\|_{\mathbf{K}(\mathbf{x}_*^n)} \leq C \|\mathbf{e}_{\mathbf{u}}^n\|_{\mathbf{K}(\mathbf{x}_*^n)} + C \|\mathbf{e}_{\mathbf{x}}^n\|_{\mathbf{K}(\mathbf{x}_*^n)} + C \|\mathbf{d}_{\mathbf{p}}^n\|_{*,\mathbf{x}_*^n} \quad \text{for } 0 \leq n \leq m-1. \quad (3.15)$$

(D) *Estimates for \mathbf{e}_w^n under assumption (3.9)*: The first, second and fourth terms on the right-hand side of (3.13a) can be estimated similarly. The third term need to be estimated by using the expressions of $d\mathbf{J}(\mathbf{x}^n, \mathbf{u}^n, \mathbf{p}^n)$ and $d\mathbf{J}(\mathbf{x}_*^n, \mathbf{u}_*^n, \mathbf{p}_*^n)$ and therefore more complicated. To this end, we denote $p_{h,\theta}^n = \hat{p}_{h,\theta}^n + \theta e_{p,\theta}^n$ and use Lemma 2.2, which gives us the following expression:

$$\begin{aligned} & (\mathbf{e}_w^n)^\top (d\mathbf{J}(\mathbf{x}^n, \mathbf{u}^n, \mathbf{p}^n) - d\mathbf{J}(\mathbf{x}_*^n, \mathbf{u}_*^n, \mathbf{p}_*^n)) \\ &= \int_0^1 \frac{d}{d\theta} \int_{\Omega_h[\mathbf{x}_\theta^n]} \nabla u_{h,\theta}^n \cdot (\nabla e_{w,\theta}^n + (\nabla e_{w,\theta}^n)^\top) \nabla p_{h,\theta}^n - f e_{w,\theta}^n \cdot \nabla p_{h,\theta}^n - j'_u(x, u_{h,\theta}^n) \nabla u_d \cdot e_{w,\theta}^n \\ & \quad + (j(x, u_{h,\theta}^n) - \nabla u_{h,\theta}^n \cdot \nabla p_{h,\theta}^n) \nabla \cdot e_{w,\theta}^n dx d\theta. \end{aligned}$$

Then

$$\begin{aligned} & \partial_\theta^\bullet (\nabla u_{h,\theta}^n \cdot (\nabla e_{w,\theta}^n + (\nabla e_{w,\theta}^n)^\top) \nabla p_{h,\theta}^n) \\ &= \partial_\theta^\bullet (\nabla u_{h,\theta}^n) \cdot (\nabla e_{w,\theta}^n + (\nabla e_{w,\theta}^n)^\top) \nabla p_{h,\theta}^n + \nabla u_{h,\theta}^n \cdot \partial_\theta^\bullet (\nabla e_{w,\theta}^n + (\nabla e_{w,\theta}^n)^\top) \nabla p_{h,\theta}^n \\ & \quad + \nabla u_{h,\theta}^n \cdot (\nabla e_{w,\theta}^n + (\nabla e_{w,\theta}^n)^\top) \partial_\theta^\bullet (\nabla p_{h,\theta}^n). \end{aligned}$$

The material derivatives of $\nabla u_{h,\theta}^n$, $\nabla p_{h,\theta}^n$, $\nabla e_{w,\theta}^n$ and $\nabla \cdot e_{w,\theta}^n$ are given by

$$\begin{aligned} \partial_\theta^\bullet (\nabla u_{h,\theta}^n) &= \nabla e_{u,\theta}^n - (\nabla e_{x,\theta}^n)^\top \nabla u_{h,\theta}^n, \\ \partial_\theta^\bullet (\nabla p_{h,\theta}^n) &= \nabla e_{p,\theta}^n - (\nabla e_{x,\theta}^n)^\top \nabla p_{h,\theta}^n, \\ \partial_\theta^\bullet (\nabla e_{w,\theta}^n) &= -(\nabla e_{x,\theta}^n)^\top \nabla e_{w,\theta}^n, \\ \partial_\theta^\bullet (\nabla \cdot e_{w,\theta}^n) &= -\text{tr}[(\nabla e_{x,\theta}^n)^\top \nabla e_{w,\theta}^n]. \end{aligned}$$

Therefore,

$$\begin{aligned} & (\mathbf{e}_w^n)^\top (d\mathbf{J}(\mathbf{x}^n, \mathbf{u}^n, \mathbf{p}^n) - d\mathbf{J}(\mathbf{x}_*^n, \mathbf{u}_*^n, \mathbf{p}_*^n)) \\ &= \int_0^1 \int_{\Omega_h[\mathbf{x}_\theta^n]} \left[\nabla u_{h,\theta}^n \cdot (\nabla e_{w,\theta}^n + (\nabla e_{w,\theta}^n)^\top) (\nabla p_{h,\theta}^n) - e_{w,\theta}^n \cdot (\nabla p_{h,\theta}^n) f \right] \nabla \cdot e_{x,\theta}^n dx d\theta \\ & \quad + \int_0^1 \int_{\Omega_h[\mathbf{x}_\theta^n]} \left[(j(x, u_{h,\theta}^n) - (\nabla u_{h,\theta}^n) \cdot (\nabla p_{h,\theta}^n)) \nabla \cdot e_{w,\theta}^n \right] \nabla \cdot e_{x,\theta}^n dx d\theta \\ & \quad - \int_0^1 \int_{\Omega_h[\mathbf{x}_\theta^n]} j'_u(x, u_{h,\theta}^n) (\nabla u_d \cdot e_{w,\theta}^n) \nabla \cdot e_{x,\theta}^n dx d\theta \\ & \quad + \int_0^1 \int_{\Omega_h[\mathbf{x}_\theta^n]} \left(\nabla e_{u,\theta}^n - (\nabla e_{x,\theta}^n)^\top (\nabla u_{h,\theta}^n) \right) \cdot (\nabla e_{w,\theta}^n + (\nabla e_{w,\theta}^n)^\top) (\nabla p_{h,\theta}^n) dx d\theta \\ & \quad + \int_0^1 \int_{\Omega_h[\mathbf{x}_\theta^n]} \left(\nabla e_{p,\theta}^n - (\nabla e_{x,\theta}^n)^\top (\nabla p_{h,\theta}^n) \right) \cdot (\nabla e_{w,\theta}^n + (\nabla e_{w,\theta}^n)^\top) (\nabla u_{h,\theta}^n) dx d\theta \\ & \quad - \int_0^1 \int_{\Omega_h[\mathbf{x}_\theta^n]} (\nabla u_{h,\theta}^n) \cdot (\nabla e_{w,\theta}^n \nabla e_{x,\theta}^n + (\nabla e_{w,\theta}^n \nabla e_{x,\theta}^n)^\top) (\nabla p_{h,\theta}^n) dx d\theta \\ & \quad - \int_0^1 \int_{\Omega_h[\mathbf{x}_\theta^n]} (\nabla f \cdot e_{x,\theta}^n) (e_{w,\theta}^n \cdot \nabla p_{h,\theta}^n) + f e_{w,\theta}^n \cdot \left(\nabla e_{p,\theta}^n - (\nabla e_{x,\theta}^n)^\top (\nabla p_{h,\theta}^n) \right) dx d\theta \\ & \quad - \int_0^1 \int_{\Omega_h[\mathbf{x}_\theta^n]} (e_{u,\theta}^n - \nabla u_d \cdot e_{x,\theta}^n) \nabla u_d \cdot e_{w,\theta}^n + (u_{h,\theta}^n - u_d) (e_{x,\theta}^n)^\top (\nabla^2 u_d)^\top e_{w,\theta}^n dx d\theta \\ & \quad + \int_0^1 \int_{\Omega_h[\mathbf{x}_\theta^n]} (u_{h,\theta}^n - u_d) (e_{u,\theta}^n - \nabla u_d \cdot e_{x,\theta}^n) (\nabla \cdot e_{w,\theta}^n) dx d\theta \\ & \quad - \int_0^1 \int_{\Omega_h[\mathbf{x}_\theta^n]} (\nabla e_{u,\theta}^n - (\nabla e_{x,\theta}^n)^\top \nabla u_{h,\theta}^n) \cdot \nabla p_{h,\theta}^n (\nabla \cdot e_{w,\theta}^n) dx d\theta \end{aligned}$$

$$\begin{aligned}
& - \int_0^1 \int_{\Omega_h[\mathbf{x}_\theta^n]} \nabla u_{h,\theta}^n \cdot (\nabla e_{p,\theta}^n - (\nabla e_{x,\theta}^n)^\top \nabla p_{h,\theta}^n) (\nabla \cdot e_{w,\theta}^n) dx d\theta \\
& - \int_0^1 \int_{\Omega_h[\mathbf{x}_\theta^n]} (j(x, u_{h,\theta}^n) - \nabla u_{h,\theta}^n \cdot \nabla p_{h,\theta}^n) \sum_{l=1}^d \nabla [(e_{w,\theta}^n)_l] \cdot \partial_{x_l} e_{x,\theta}^n dx d\theta.
\end{aligned}$$

By Lemmas 3.2 and 3.3 we obtain the following estimates of this term under the induction assumption in (3.9):

$$\begin{aligned}
& (\mathbf{e}_w^n)^\top (\mathbf{dJ}(\mathbf{x}^n, \mathbf{u}^n, \mathbf{p}^n) - \mathbf{dJ}(\mathbf{x}_*^n, \mathbf{u}_*^n, \mathbf{p}_*^n)) \\
& \leq C \|\mathbf{e}_w^n\|_{\mathbf{K}(\mathbf{x}_*^n)} (\|\mathbf{e}_p^n\|_{\mathbf{K}(\mathbf{x}_*^n)} + \|\mathbf{e}_x^n\|_{\mathbf{K}(\mathbf{x}_*^n)} + \|\mathbf{e}_u^n\|_{\mathbf{K}(\mathbf{x}_*^n)}) (1 + \|e_p^n\|_{W^{1,\infty}}) (1 + \|e_x^n\|_{W^{1,\infty}}) \\
& \quad + C \|\mathbf{e}_w^n\|_{\mathbf{K}(\mathbf{x}_*^n)} (\|\mathbf{e}_x^n\|_{\mathbf{K}(\mathbf{x}_*^n)} + \|\mathbf{e}_u^n\|_{\mathbf{K}(\mathbf{x}_*^n)} + \|\mathbf{e}_p^n\|_{\mathbf{K}(\mathbf{x}_*^n)}) (1 + \|e_u^n\|_{W^{1,\infty}}) (1 + \|e_x^n\|_{W^{1,\infty}}) \\
& \leq C \|\mathbf{e}_w^n\|_{\mathbf{K}(\mathbf{x}_*^n)} (\|\mathbf{e}_p^n\|_{\mathbf{K}(\mathbf{x}_*^n)} + \|\mathbf{e}_x^n\|_{\mathbf{K}(\mathbf{x}_*^n)} + \|\mathbf{e}_u^n\|_{\mathbf{K}(\mathbf{x}_*^n)}),
\end{aligned}$$

where the $W^{1,\infty}$ norm is taken on the domain $\Omega_h[\mathbf{x}_*^n]$. Combining this estimate with the other three terms on the right-hand side of (3.13a), for which the estimates are omitted due to the similarity as Part (B) of the proof, we have

$$\|\mathbf{e}_w^n\|_{\mathbf{K}(\mathbf{x}_*^n)} \leq C (\|\mathbf{e}_x^n\|_{\mathbf{K}(\mathbf{x}_*^n)} + \|\mathbf{e}_u^n\|_{\mathbf{K}(\mathbf{x}_*^n)} + \|\mathbf{e}_p^n\|_{\mathbf{K}(\mathbf{x}_*^n)} + \|\mathbf{d}_w^n\|_{*,\mathbf{x}_*^n}). \quad (3.16)$$

(E) *Combination of estimates for $\|\mathbf{e}_x^n\|_{\mathbf{K}(\mathbf{x}_*^n)}$, $\|\mathbf{e}_u^n\|_{\mathbf{K}(\mathbf{x}_*^n)}$, $\|\mathbf{e}_p^n\|_{\mathbf{K}(\mathbf{x}_*^n)}$ and $\|\mathbf{e}_w^n\|_{\mathbf{K}(\mathbf{x}_*^n)}$* : Finally, combining the estimates in (3.12), (3.14), (3.15) and (3.16), we have

$$\begin{aligned}
\|\mathbf{e}_x^n\|_{\mathbf{K}(\mathbf{x}_*^n)} & \leq C\tau \sum_{l=1}^n \left(\|\mathbf{d}_x^{l-1}\|_{\mathbf{K}(\mathbf{x}_*^{l-1})} + \|\mathbf{d}_u^{l-1}\|_{*,\mathbf{x}_*^{l-1}} + \|\mathbf{d}_p^{l-1}\|_{*,\mathbf{x}_*^{l-1}} + \|\mathbf{d}_w^{l-1}\|_{*,\mathbf{x}_*^{l-1}} \right) \\
& \quad + C\tau \sum_{l=1}^n \|\mathbf{e}_x^{l-1}\|_{\mathbf{K}(\mathbf{x}_*^{l-1})} \quad \text{for } 1 \leq n \leq m.
\end{aligned}$$

By applying the discrete Grönwall's inequality we obtain

$$\|\mathbf{e}_x^n\|_{\mathbf{K}(\mathbf{x}_*^n)} \leq C \sup_{1 \leq l \leq n} \left(\|\mathbf{d}_x^{l-1}\|_{\mathbf{K}(\mathbf{x}_*^{l-1})} + \|\mathbf{d}_u^{l-1}\|_{*,\mathbf{x}_*^{l-1}} + \|\mathbf{d}_p^{l-1}\|_{*,\mathbf{x}_*^{l-1}} + \|\mathbf{d}_w^{l-1}\|_{*,\mathbf{x}_*^{l-1}} \right) \quad \text{for } 1 \leq n \leq m. \quad (3.17)$$

Furthermore, under the assumptions of Proposition 3.6, the inequality above implies the following result:

$$\|\mathbf{e}_x^n\|_{\mathbf{K}(\mathbf{x}_*^n)} \leq C(\tau + h^k) \quad \text{for } 1 \leq n \leq m.$$

With this result, estimates (3.14) and (3.15) also hold for $1 \leq n \leq m$. This leads to

$$\|\mathbf{e}_u^n\|_{\mathbf{K}(\mathbf{x}_*^n)} \leq C(\tau + h^k) \quad \text{and} \quad \|\mathbf{e}_p^n\|_{\mathbf{K}(\mathbf{x}_*^n)} \leq C(\tau + h^k) \quad \text{for } 1 \leq n \leq m,$$

which further imply (3.9) for $1 \leq n \leq m$ when $\tau = o(h^{\frac{d}{2}})$ and h is sufficiently small. This completes the mathematical induction.

Accordingly, estimates (3.14), (3.15), (3.16) and (3.17) hold for all $1 \leq n \leq [T/\tau]$. Substituting (3.17) into the other three estimates yields the desired result in Proposition 3.6. \square

3.4. Proof of Theorem 2.3

Theorem 2.3 is an immediate consequence of the consistency estimates in Lemma 3.5 and the stability estimates in Proposition 3.6. \square

4. Numerical examples

In this section, we present numerical examples to support the theoretical analysis in this article by illustrating the convergence of the proposed method and the effectiveness of the method in simulating boundary evolution under shape gradient flow in three-dimensional space.

The computations are performed by the finite element software package NGSolve; see <https://ngsolve.org/>.

Example 4.1 (Convergence of the evolving FEM). We consider the shape gradient flow governed by the following moving-boundary PDEs:

$$\partial_t \phi = w \circ \phi \quad \text{in } \Omega^0, \quad \phi(0) = \text{id}|_{\Omega^0} \quad \text{in } \Omega^0, \quad (4.1a)$$

$$-\Delta w + w = 0 \quad \text{in } \Omega(t), \quad \partial_\nu w = -J'(\Gamma(t))\nu \quad \text{on } \Gamma(t), \quad (4.1b)$$

$$-\Delta u + u = f \quad \text{in } \Omega(t), \quad u = 0 \quad \text{on } \Gamma(t), \quad (4.1c)$$

$$-\Delta p + p = j'_u(\cdot, u) \quad \text{in } \Omega(t), \quad p = 0 \quad \text{on } \Gamma(t), \quad (4.1d)$$

with the following initial condition, source functions and shape density:

$$\Omega^0 = \{(x_1, x_2) : \frac{x_1^2}{0.85^2} + \frac{x_2^2}{0.45^2} \leq 1\}, \quad f = 5 - x_1^2 - x_2^2,$$

$$j(\cdot, u) = \frac{1}{2}|u - u_d|^2 \quad \text{with } u_d = 1 - x_1^2 - x_2^2.$$

As $t \rightarrow \infty$, the boundary $\Gamma(t)$ converges to the optimal boundary Γ_∞ which minimizes the energy functional $J(\Gamma) = \int_\Omega j(x, u) dx$ under the constraint $-\Delta u + u = f$, i.e., the unit circle $\Gamma_\infty = \{(x_1, x_2) : x_1^2 + x_2^2 = 1\}$.

We test the convergence in space at time $T = 10$ by choosing a sufficiently small time stepsize such that the errors from the time discretization are negligibly small. The errors of the numerical solutions are presented in Figure 1 (a)–(d) for several different spatial mesh sizes $h = 1/24, 1/36, 1/54, 1/81$, for finite elements of degree $k = 1, 2, 3$. The numerical results indicate that the numerical solutions have k th-order convergence in space. This is consistent with the theoretical result proved in Theorem 2.3 in the case $k = 2, 3$. The analysis of stability and convergence of numerical approximations in the case $k = 1$ is still challenging.

In addition to spatial discretization, we also test the convergence of time discretizations by backward differentiation formula (BDF) of order 1, 2, 3, which we refer to as BDF-1, BDF-2 and BDF-3, respectively. In the BDF methods, we replace w_h^n by w_h^{n+1} in equation (2.10a). Accordingly, (2.10) becomes a nonlinearly coupled system which needs to be solved by using fixed point iterations. The errors of the time discretizations are presented in Figure 1 (e)–(h), which indicate that the numerical solutions given by BDF- k have k th-order convergence in time. For BDF-1, this is consistent with the theoretical result proved in Theorem 2.3.

The shape and mesh of the evolving domain at time $t = 0$, $t = 5$ and $t = 30$ are presented in Figure 2 with meshsize $h = 0.06$ and time stepsize $\tau = 0.1$, where we can observe that the shape of the boundary at $t = 30$ is already almost the same as the optimal boundary Γ_∞ .

Example 4.2 (Dumbbell shape in three dimensions). We consider problem (4.1) with the following initial condition, shape density and source functions:

$$\Omega^0 = \{(x_1, x_2, x_3) : \frac{x_1^2}{0.85^2} + \frac{x_2^2}{0.45^2} + \frac{x_3^2}{0.45^2} \leq 1\}$$

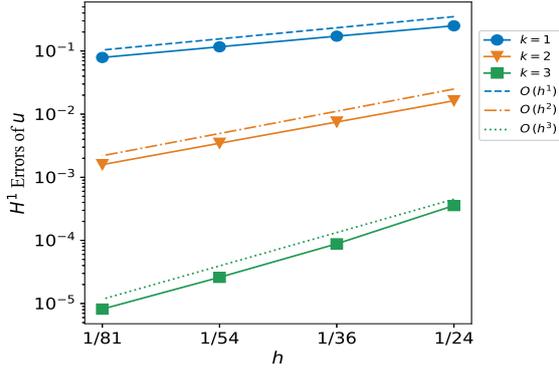
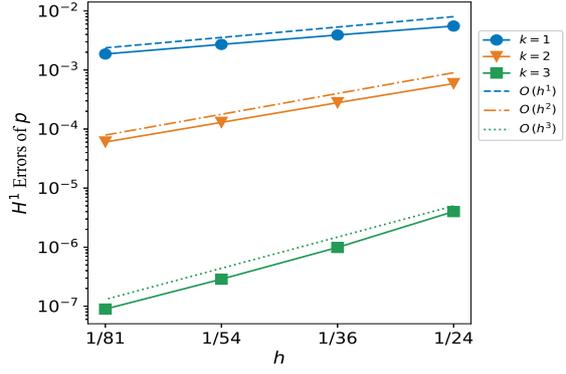
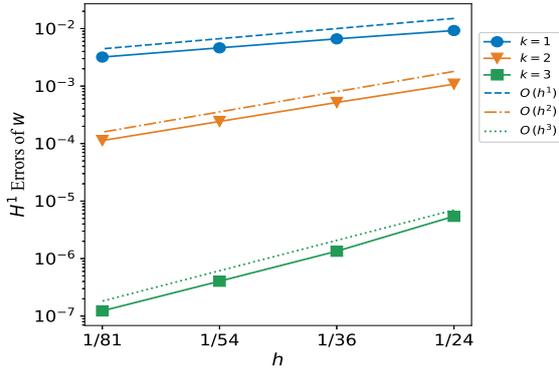
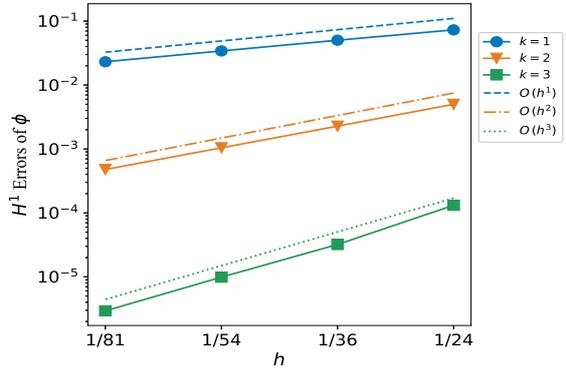
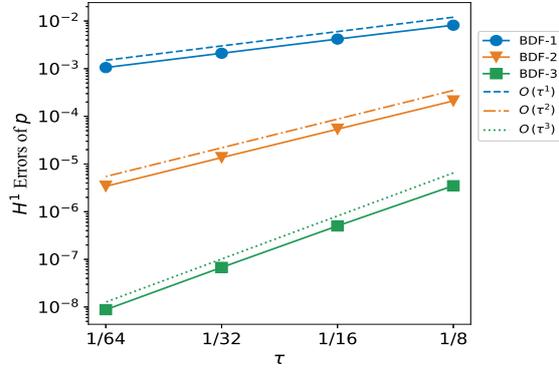
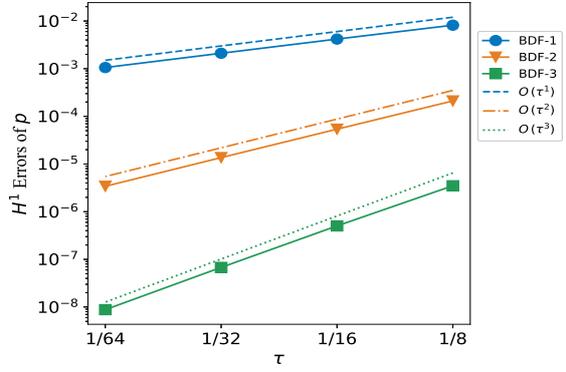
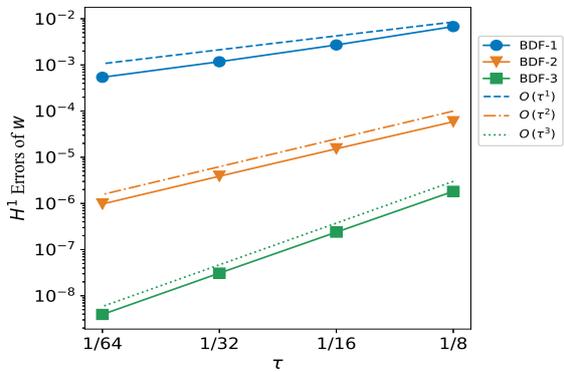
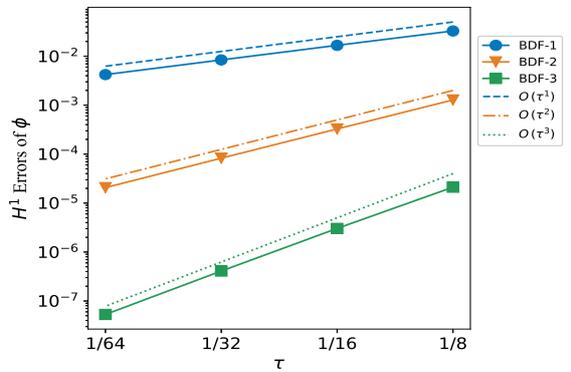
$$j(\cdot, u) = \frac{1}{2}|u - u_d|^2 \quad \text{with } u_d = x_1^2 + \frac{x_2^2 + x_3^2}{(0.7x_1^2 + 0.3)^2} - 1$$

$$f = -\Delta u_d + u_d.$$

As $t \rightarrow \infty$, the boundary $\Gamma(t)$ converges to a three-dimensional dumbbell-shape optimal boundary Γ_∞ , as shown in Figure 3 (d). The shape and mesh of the evolving 3D domain at time $t = 0$, $t = 10$ and $t = 100$ are presented in Figure 3 (a)–(c) with meshsize $h = 0.06$ and time stepsize $\tau = 0.1$, where we can observe that the shape of the boundary at $t = 100$ is almost the same as the optimal boundary Γ_∞ .

Example 4.3 (Drag minimization under Stokes flow in three dimensions).

We consider an example from drag minimization on the exterior of a three-dimensional obstacle surrounded by viscous incompressible Stokes flow (see [22, Fig. 3.2]), which has many applications in airfoil design and obstacle problems. The analysis in this article could be extended to this example; see the discussions in Section 5.

(a) H^1 error of u from space discretization(b) H^1 error of p from space discretization(c) H^1 error of w from space discretization(d) H^1 error of ϕ from space discretization(e) H^1 error of u from time discretization(f) H^1 error of p from time discretization(g) H^1 error of w from time discretization(h) H^1 error of ϕ from time discretizationFIGURE 1. Errors of the numerical solutions at time $T = 10$ (Example 4.1)

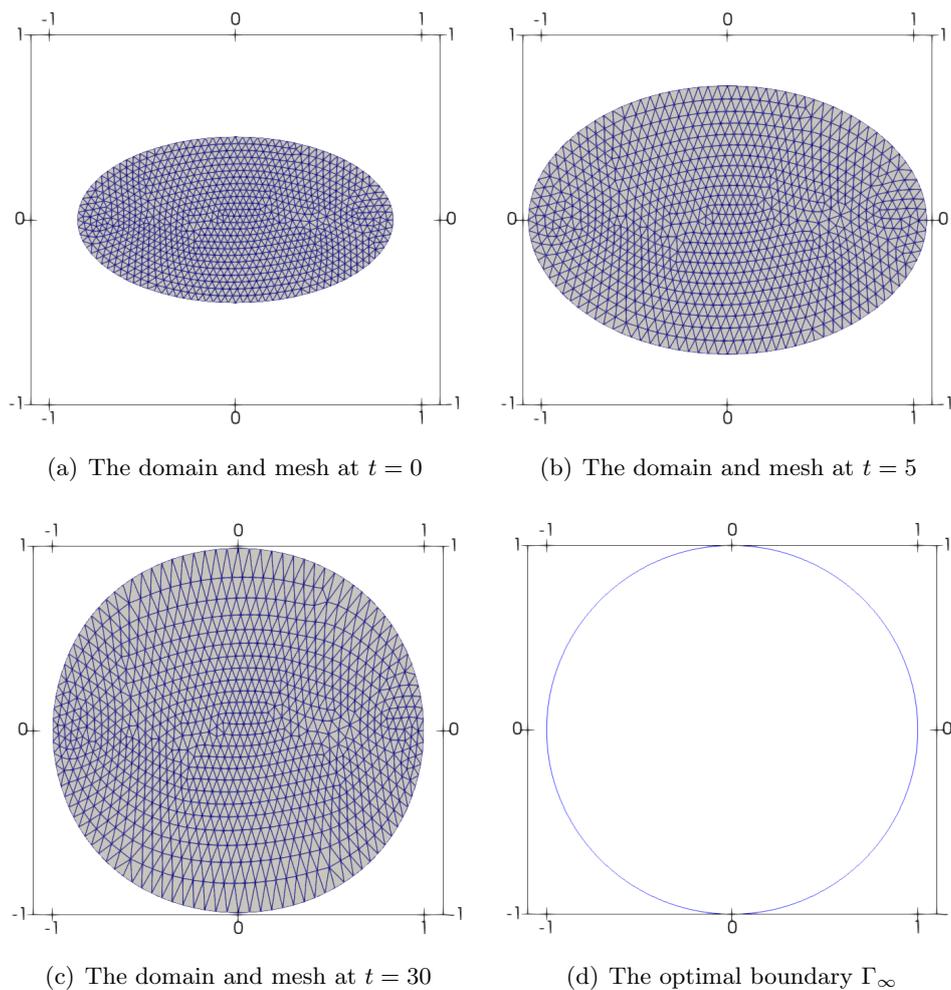


FIGURE 2. Evolution of the 2D domain in Example 4.1.

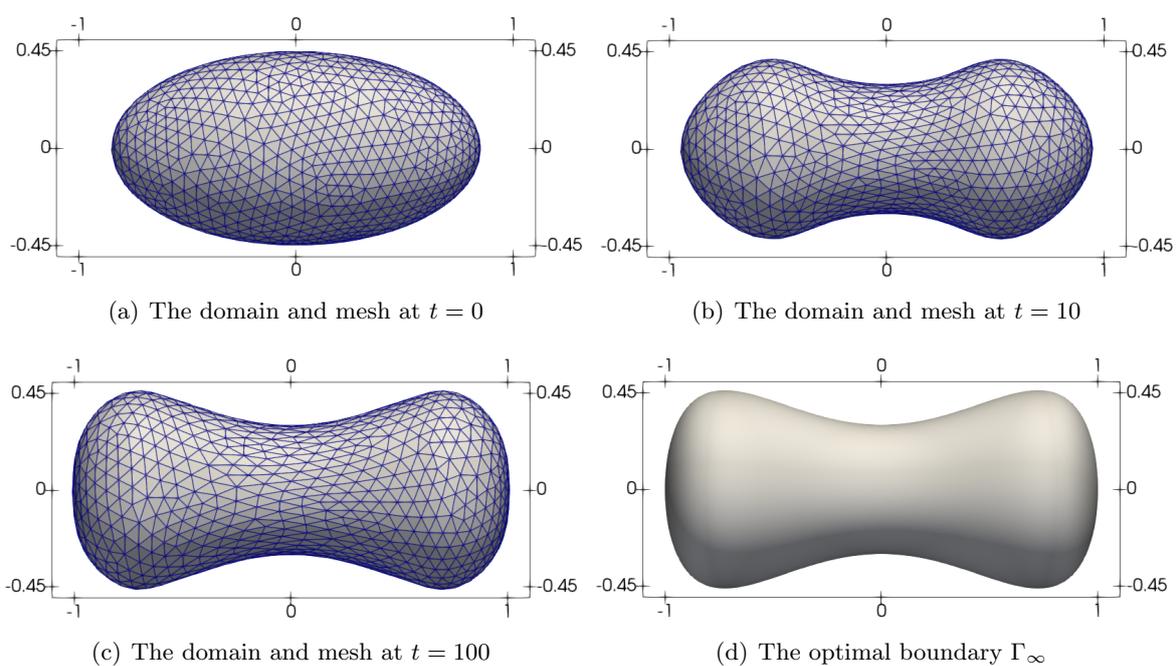


FIGURE 3. Evolution of the 3D domain in Example 4.2.

The velocity of the fluid which surrounds the obstacle is governed by the following Stokes equation:

$$\begin{cases} -\nabla \cdot (2\mu\mathbb{D}(u) - pI) = f & \text{in } D \setminus \Omega, \\ \nabla \cdot u = 0 & \text{in } D \setminus \Omega, \\ u = 0 & \text{on } \Gamma \cup \Gamma_w, \\ u = u_{\text{in}} & \text{on } \Gamma_{\text{in}}, \\ (2\mu\mathbb{D}(u) - pI)\nu = 0 & \text{on } \Gamma_{\text{out}}, \end{cases} \quad (4.2)$$

where $\mu = 0.1$ denotes the viscosity of the fluid and ν is the unit outward normal vector on the boundary of domain $D = (-1, 1) \times (-0.5, 0.5) \times (-0.5, 0.5)$ excluding the obstacle Ω , and $\mathbb{D}(u) = \frac{1}{2}(\nabla u + \nabla u^\top)$ is the deformation tensor. The inflow and outflow parts of the boundary are denoted by Γ_{in} and Γ_{out} , which are the left and right sides of the cube, respectively. The no-slip boundary condition is imposed on the other parts Γ_w of the boundary and the boundary Γ of the obstacle. In our numerical test we set $f = 0$ and $u_{\text{in}} = (1, 0, 0)^\top$.

The drag minimization problem seeks a boundary Γ which minimizes the energy dissipation of fluids, i.e.,

$$\min_{\substack{\partial\Omega=\Gamma \\ \text{Vol}(\Omega)=\text{const}}} J(\Gamma) = \mu \int_{D \setminus \Omega} |\mathbb{D}u|^2 dx.$$

The distributed Eulerian derivative of the energy functional $J(\Gamma)$ has the following expression (see [47, (36)]):

$$dJ(\Gamma; v) = \int_{D \setminus \Omega} \left[\mu \left(|\mathbb{D}u|^2 \nabla \cdot v - \mathbb{D}u : (\nabla u \nabla v^\top + \nabla v \nabla u^\top) \right) + p \nabla u : \nabla v^\top \right] dx.$$

By applying our method to the drag minimization under Stokes flow, we evolve the domain under the H^1 shape gradient flow associated to the energy functional $J(\Gamma)$, where the initial shape of the domain is a sphere centered at $(-0.4, 0, 0)$ with radius 0.2. In order to maintain the volume constraint of Ω in the evolution, we choose a divergence-free descent velocity w from the solution of the following equation (in the weak formulation):

$$\begin{aligned} \int_{D \setminus \Omega} \nabla w : \nabla v \, dx - \int_{D \setminus \Omega} q \nabla \cdot v \, dx &= -dJ(\Gamma; v), \\ \int_{D \setminus \Omega} \nabla \cdot w \, \eta \, dx &= 0, \end{aligned} \quad (4.3)$$

for any test functions $\eta \in \mathbb{R}$ and $v \in H^1(D \setminus \Omega)$ with no-slip boundary condition on $\Gamma_w \cup \Gamma_{\text{in}} \cup \Gamma_{\text{out}}$ in the finite element discretization.

The evolution of shape and mesh of the three-dimensional domain at time $t = 0$, $t = 0.01$, $t = 0.04$ and $t = 0.05$ are presented in Figure 4 (a)-(d) with time stepsize $\tau = 0.001$, maximal mesh size $h = 0.06$ in Ω , and local mesh size $h = 0.02$ near the obstacle. Our numerical simulation shows that the shape of the domain is close to stationary at $t = 0.05$. The corresponding optimal shape of the obstacle under volume constraint dragged by the viscous incompressible Stokes flow is presented in Figure 4(d).

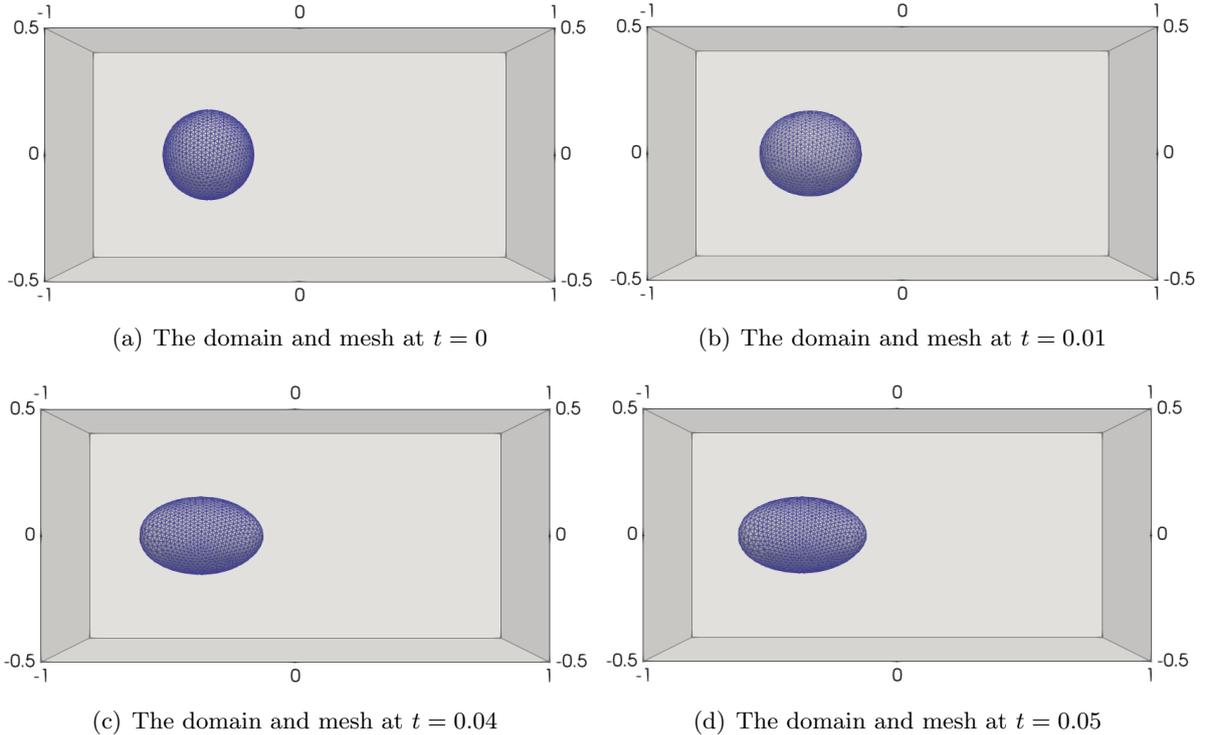


FIGURE 4. Evolution of the shape of the obstacle in Example 4.3.

5. Conclusions and extensions

We have formulated the shape optimization problem with shape density function $j(\cdot, u) = \frac{1}{2}|u - u_d|^2$, constrained by the Poisson equation, into a gradient flow system of nonlinear PDEs on an evolving domain with a solution-driven evolving boundary that tends to the optimal shape. The formulation is intended to make the evolving finite element approximations to the boundary evolution have stability and convergence with optimal-order accuracy up to a given time through utilizing the H^1 shape gradient flow and distributed Eulerian derivative. The main advantage of using the distributed Eulerian shape derivative lies in the convenience it offers for proving the stability estimates. If we were to use formula (2.4), it would necessitate computing the material derivative ∂_θ^\bullet of the normal vector of the boundary Γ_h^θ . However, this approach would lead to several difficulties. First, we would need to estimate the errors associated with the normal vector, which can be a complex task. Additionally, the usage of Sobolev spaces on the boundary becomes necessary, introducing further intricacies in the analysis. Moreover, the computation of the material derivative of the normal vector is not as clear-cut as with the distributed Eulerian derivative. The lack of clarity in this process adds to the overall complexity and makes the stability analysis more challenging.

We have proved the stability and convergence of the numerical approximations to the boundary evolution under shape gradient flow by using the evolving bulk finite elements systematically studied in [14] and the geometric estimates developed in [32, 13]. The latter is used to compare the error between two functions defined on different domains. We have illustrated the convergence and performance of the proposed method through numerical examples in both two and three dimensions.

The formulation, algorithm and analysis in this article could be extended to other shape density functions, such as $j(\cdot, u) = \frac{1}{2}|\nabla u|^2$. For example, for the shape density function $j(\cdot, u) = \frac{1}{2}|\nabla u|^2$ we have $p = u$ and the following expression of the distributed Eulerian derivative:

$$dJ(\Gamma, u; v) = \frac{1}{2} \int_{\Omega} \left(\nabla u \cdot (\nabla v + \nabla v^\top) \nabla u - |\nabla u|^2 \nabla \cdot v \right) dx - \int_{\Omega} f v \cdot \nabla u dx.$$

The nonlinear structure in this expression is similar as the distributed Eulerian derivative for $j(\cdot, u) = \frac{1}{2}|u - u_d|^2$, and therefore the analysis in this article could be trivially extended to the case $j(\cdot, u) = \frac{1}{2}|\nabla u|^2$. Similarly, the analysis in this article could be equally extended to constraints with the Stokes equations and additional constraint on the volume conservation enclosed by the boundary. Therefore, the numerical analysis is also applicable to the drag minimization problem in Example 4.3.

The improvement of the formulation, algorithm and analysis to contain artificial tangential motions which can improve the mesh quality in approximating the evolving boundary of general closed surfaces in the three-dimensional space is interesting and nontrivial.

Acknowledgment

The work of the first author was partially supported by the National Key Research and Development Program of China (project no. 2022YFA1004402), the Strategic Priority Research Program of Chinese Academy of Sciences (project no. XDB 41000000) and the National Natural Science Foundation of China (project no. 12071468). The work of the second and third authors was partially supported by the Research Grants Council of the Hong Kong (GRF project no. 15300920).

References

- [1] G. Allaire, C. Dapogny and F. Jouve. *Shape and topology optimization, in Geometric Partial Differential Equations, Part II*. A. Bonito and R. Nochetto, eds., Handb. Numer. Anal. 22, Elsevier, Amsterdam, 2021, pp. 1–132.
- [2] S. Badia and R. Codina. Analysis of a stabilized finite element approximation of the transient convection-diffusion equation using an arbitrary Lagrangian–Eulerian framework. *SIAM J. Numer. Anal.*, 44(2006), pp. 2159–2197.
- [3] D. Boffi and L. Gastaldi. Stability and geometric conservation laws for arbitrary Lagrangian–Eulerian formulations. *Comput. Methods Appl. Mech. Engrg.*, 193(2004), pp. 4717–4739.
- [4] S. C. Brenner and L. R. Scott. *The Mathematical Theory of FEMs*. 3rd edn., Springer, New York, 2008.
- [5] M. Burger. A framework for the construction of level set methods for shape optimization and reconstruction. *Interfaces and Free boundaries*, 5(2003), pp. 301–329.
- [6] E. Burman, D. Elfverson, P. Hansbo, M. G. Larson and K. Larsson. A cut finite element method for the Bernoulli free boundary value problem. *Comput. Methods Appl. Mech. Engrg.*, 317(2017), pp. 598–618.
- [7] D. Chenais and E. Zuazua. Finite-element approximation of 2D elliptic optimal design. *J. Math. Pures Appl.*, 85(2006), pp. 225–249.
- [8] F. de Gournay. Velocity extension for the level-set method and multiple eigenvalues in shape optimization. *SIAM J. Control Optim.*, 45(2006), pp. 343–367.
- [9] L. Dedè, M.J. Borden and T.J.R. Hughes. Isogeometric analysis for topology optimization with a phase field model. *Arch. Comput. Methods Engrg.*, 19(2012), pp. 427–465.
- [10] M. C. Delfour and J. P. Zolésio. *Shapes and Geometries: Metrics, Analysis, Differential Calculus, and Optimization*. 2nd edn., SIAM, Philadelphia, 2011.
- [11] G. Dogan, P. Morin, R. H. Nochetto and M. Verani. Discrete gradient flows for shape optimization and applications. *Comput. Methods Appl. Mech. Engrg.*, 196(2007), pp. 3898–3914.
- [12] G. Dziuk and C. M. Elliott. Finite elements on evolving surfaces. *IMA J. Numer. Anal.*, 27(2007), pp. 262–292.
- [13] D. Edelmann. Finite element analysis for a diffusion equation on a harmonically evolving domain. *IMA J. Numer. Anal.*, 42(2022), pp. 1866–1901.
- [14] 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(2021), no. 3, pp. 1696–1845.
- [15] C. M. Elliott and V. Styles. An arbitrary Lagrangian–Eulerian ESFEM for solving PDEs on evolving surfaces. *Milan J. Math.*, 80(2012), pp. 469–501.
- [16] C. M. Elliott and C. Venkataraman. Error analysis for an arbitrary Lagrangian–Eulerian evolving surface FEM. *Numer. Meth. Part. Diff. Eq.*, 31(2015), pp. 459–499.
- [17] K. Eppler, H. Harbrecht and R. Schneider. On convergence in elliptic shape optimization. *SIAM J. Control Optim.*, 46(2007), no. 1, pp. 61–83.
- [18] K. Eppler and H. Harbrecht. Coupling of FEM and BEM in shape optimization. *Numer. Math.*, 104(2006), no. 1, pp. 47–68.

- [19] L. Formaggia and F. Nobile. A stability analysis for the arbitrary Lagrangian Eulerian formulation with finite elements. *East-West J. Numer. Math.*, 7(1999), pp. 105–132.
- [20] I. Fumagalli, N. Parolini and M. Verani. Shape optimization for Stokes flows: a finite element convergence analysis. *ESAIM Math. Model. Numer. Anal.*, 49(2015), no. 4, pp. 921–951.
- [21] L. Gastaldi. A priori error estimates for the arbitrary Lagrangian Eulerian formulation with finite elements. *East-West J. Numer. Math.*, 9(2001), pp. 123–156.
- [22] W. Gong, J. Li and S. Zhu. Improved discrete boundary type shape gradients for PDE-constrained shape optimization. *SIAM J. Sci. Comput.*, 44(2022), pp. A2464–A2505.
- [23] R. Guo. Solving parabolic moving interface problems with dynamical immersed spaces on unfitted meshes: fully discrete analysis. *SIAM J. Numer. Anal.*, 59(2021), pp. 797–828.
- [24] R. Guo, T. Lin and Y. Lin. Recovering elastic inclusions by shape optimization methods with immersed finite elements. *J. Comput. Phys.*, 404(2020), 109123, pp. 24.
- [25] J. Haslinger and R. Mäkinen. *Introduction to Shape Optimization: Theory, Approximation, and Computation*. SIAM, 2003.
- [26] J. Haslinger and P. Neittaanmaki. *Finite Element Approximation for Optimal Shape, Material and Topology Design*. 2nd edn., J. Wiley & Sons: Chichester, 1996.
- [27] E. J. Haug, K. K. Choi and V. Komkov. *Design Sensitivity Analysis of Structural Systems*. Academic Press, Orlando, FL, 1986.
- [28] A. Henrot and M. Pierre. *Shape Variation and Optimization*. EMS Tracts in Mathematics, Vol. 28, 2018.
- [29] R. Hiptmair, A. Paganini and S. Sargheini. Comparison of approximate shape gradients. *BIT Numer. Math.*, 55(2015), pp. 459–485.
- [30] M. Lenoir. Optimal iso-parametric finite elements and error estimates for domains involving curved boundaries. *SIAM J. Numer. Anal.*, 23(1986), pp. 562–580.
- [31] B. Kiniger and B. Vexler. A priori error estimates for finite element discretizations of a shape optimization problem. *ESAIM Math. Model. Numer. Anal.*, 47(2013), pp. 1733–1763.
- [32] B. Kovács, B. Li and C. Lubich. Convergence of finite elements on an evolving surface driven by diffusion on the surface. *Numer. Math.*, 137(2017), pp. 643–689.
- [33] B. Kovács and C. A. Power Guerra. Higher order time discretizations with arbitrary Lagrangian–Eulerian finite elements for parabolic problems on evolving surfaces. *IMA J. Numer. Anal.*, 38(2018), pp. 460–494.
- [34] B. Kovács, B. Li and C. Lubich. A convergent evolving finite element algorithm for mean curvature flow of closed surfaces. *Numer. Math.*, 143(2019), pp. 797–853.
- [35] R. Lan and P. Sun. A novel arbitrary Lagrangian–Eulerian FEM for a parabolic/mixed parabolic moving interface problem. *J. Comput. Appl. Math.*, 383(2021), 113125.
- [36] 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.*, 2022, DOI: 10.1093/imanum/drab099.
- [37] C. Ma, Q. Zhang and W. Zheng. A fourth-order unfitted characteristic FEM for solving the advection-diffusion equation on time-varying domains. *SIAM J. Numer. Anal.*, 60(2022), pp. 2203–2224.
- [38] J. S. Martín, L. Smaranda and T. Takahashi. Convergence of a finite element/arbitrary Lagrangian–Eulerian method for the Stokes equations in a domain depending on time. *J. Comput. Appl. Math.*, 230(2009), pp. 521–545.
- [39] B. Mohammadi and O. Pironneau. *Applied Shape Optimization for Fluids*. 2nd edn., Oxford University Press, Oxford, 2010.
- [40] P. Morin, R. H. Nochetto, M. S. Pauletti and M. Verani. Adaptive FEM for shape optimization. *ESAIM: COCV*, 18(2012), pp. 1122–1149.
- [41] F. Nobile. *Numerical approximation of fluid-structure interaction problems with application to haemodynamics*. Ph.D. thesis, Department of Mathematics, Ecole Polytechnique Fdrale de Lausanne, Switzerland, 2001.
- [42] V. Schulz and M. Siebenborn. Computational comparison of surface metrics for PDE constrained shape optimization. *Comput. Methods Appl. Math.*, 16(2016), pp. 485–496.
- [43] V. Schulz, M. Siebenborn and K. Welker. Efficient PDE constrained shape optimization based on Steklov-Poincaré-type metrics. *SIAM J. Optim.*, 26(2016), pp. 2800–2819.
- [44] J. Sokolowski and J.P. Zolésio. *Introduction to Shape Optimization: Shape Sensitivity Analysis*. Springer, Heidelberg, 1992.
- [45] Shawn W. Walker. *The Shapes of Things: A Practical Guide to Differential Geometry and the Shape Derivative*. SIAM, 2015.
- [46] C. F. Wang, S. T. Xia, X .L. Wang and X. P. Qian. Isogeometric shape optimization on triangulations. *Comput. Methods Appl. Mech. Engrg.*, 331(2018), pp. 585–622.
- [47] S. Zhu and Z. Gao. Convergence analysis of mixed finite element approximations to shape gradients in the Stokes equation. *Comput. Methods Appl. Mech. Engrg.*, 343(2019), pp. 127–150.

WEI GONG: THE STATE KEY LABORATORY OF SCIENTIFIC AND ENGINEERING COMPUTING, INSTITUTE OF COMPUTATIONAL MATHEMATICS & NATIONAL CENTER FOR MATHEMATICS AND INTERDISCIPLINARY SCIENCES, ACADEMY OF MATHEMATICS AND SYSTEMS SCIENCE, CHINESE ACADEMY OF SCIENCES, 100190 BEIJING, CHINA. E-mail address: wgong@lsec.cc.ac.cn

BUYANG LI: DEPARTMENT OF APPLIED MATHEMATICS, THE HONG KONG POLYTECHNIC UNIVERSITY, HUNG HOM, HONG KONG. E-mail address: buyang.li@polyu.edu.hk

QIQI RAO: DEPARTMENT OF APPLIED MATHEMATICS, THE HONG KONG POLYTECHNIC UNIVERSITY, HUNG HOM, HONG KONG. E-mail address: qi-qi.rao@connect.polyu.hk