# NUMERICAL ANALYSIS OF SUBDIFFUSION, NAVIER-STOKES EQUATIONS, AND FLUID-STRUCTURE INTERACTION

QIQI RAO

PhD

The Hong Kong Polytechnic University

2025

The Hong Kong Polytechnic University

Department of Applied Mathematics

# Numerical Analysis of Subdiffusion, Navier-Stokes Equations, and Fluid-Structure Interaction

Qiqi RAO

A thesis submitted in partial fulfilment of the requirements

for the degree of Doctor of Philosophy

April 2025

# Certificate of Originality

I hereby declare that this thesis is my own work and that, to the best of my knowledge and belief, it reproduces no material previously published or written, nor material that has been accepted for the award of any other degree or diploma, except where due acknowledgement has been made in the text.

_____ (Signed)

Qiqi RAO (Name of student)

# Abstract

This thesis is devoted to high-order convergent numerical methods for some nonlinear parabolic equations with rough initial data and evolving domains.

Chapter 2 and Chapter 3 are devoted to the analysis of numerical methods for subdiffusion and Navier–Stokes (NS) equations with rough data. In Chapter 2, a new spectral method is constructed for the linear and semilinear subdiffusion equations with possibly discontinuous rough initial data. The new method effectively combines several computational techniques, including the contour integral representation of the solutions, the quadrature approximation of contour integrals, the exponential integrator using the de la Vallée Poussin (VP) means of the source function, and a decomposition of the time interval geometrically refined towards the singularity of the solution and the source function. Rigorous error analysis shows that the proposed method has spectral convergence for the linear and semilinear subdiffusion equations with bounded measurable initial data and possibly singular source functions under the natural regularity of the solutions.

Chapter 3 concerns the numerical solution of the two-dimensional NS equations with nonsmooth initial data in the $L^2$ space, which is the critical space for the two-dimensional NS equations to be well-posed. In this case, the solutions of the NS equations exhibit certain singularities at $t = 0$, e.g., the $H^s$ norm of the solution blows up as $t \to 0$ when $s > 0$. To date, the best convergence result proved in the literature are first-order accuracy in both time and space for the semi-implicit Euler time-stepping scheme and divergence-free finite elements (even high-order finite elements are used), while numerical results demonstrate that second-order convergence in time and space may be achieved. Therefore, there is still a gap between numerical analysis and numerical computation for the NS equations with $L^2$ initial data. The primary challenge to realizing high-order convergence is the insufficient regularity in the solutions due to the rough initial condition and the nonlinearity of the equations. In this work, we propose a fully discrete numerical scheme that utilizes the Taylor–Hood or Stokes-MINI finite element method (FEM) for spatial discretization and an implicit–explicit Runge–Kutta (RK) time-stepping method in conjunction with

graded stepsizes. By employing discrete semigroup techniques, sharp regularity estimates, negative norm estimates and the $L^2$ projection onto the divergence-free Raviart–Thomas (RT) element space, we prove that the proposed scheme attains second-order convergence in both space and time. Numerical examples are presented to support the theoretical analysis. In particular, the convergence in space is at most second order even higher-order finite elements are used. This shows the sharpness of the convergence order proved in this chapter.

Chapter 4, Chapter 5 and Chapter 6 are devoted to the high-order convergent FEM for parabolic equations in evolving domains.

As a specific type of shape gradient descent algorithm, shape gradient flow is widely used for shape optimization problems constrained by partial differential equations (PDEs). In this approach, the constraint PDEs 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 Chapter 4, 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 $k \geqslant 2$ 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 PDEs is also discussed.

In Chapter 5, the numerical solution of the Stokes equations on an evolving domain with a moving boundary is studied based on the arbitrary Lagrangian–Eulerian (ALE) finite element method along the trajectories of the evolving mesh. The error of the semidiscrete ALE method is shown to be $O(h^{r+1})$ for velocity in $L^\infty(0,T;L^2)$ norm and $O(h^r)$ for pressure in $L^2(0,T;L^2)$ norm by employing the Taylor–Hood finite elements of degree $r \geqslant 2$, using Nitsche's duality argument adapted to an evolving mesh, by proving that the material derivative and the Stokes–Ritz projection commute up to terms which have optimal-order convergence in the $L^2$ norm. Numerical examples are provided to support the theoretical analysis.

In Chapter 6, the convergence of ALE–FEMs for fluid-structure interaction (FSI) problems with a solution-driven moving interface is studied. In addition to the finite element

approximation errors, the geometric approximation errors due to the unknown interface motion and mesh evolvement are considered as well within the ALE frame. By adding an initial correction term to the numerical scheme, the optimal convergence rates of fluid velocity, structural displacement and ALE mesh motion in $L^\infty(0,T;H^1)$ norm, as well as of pressure in $L^\infty(0,T;L^2)$ norm, are proved for a fully discrete, monolithic FEM which tracks the unknown moving interface using the ALE approach. Numerical experiments in both two and three dimensions are presented to support the theoretical error analysis.

# Publications arising from the thesis

1. B. Li, Y. Lin, S. Ma, and Q. Rao. An Exponential Spectral Method Using VP Means for Semilinear Subdiffusion Equations with Rough Data, *SIAM J. Numer. Anal.*, 61(5):2305–2326, 2023.

2. W. Gong, B. Li, and Q. Rao. Convergent Evolving Finite Element Approximations of Boundary Evolution under Shape Gradient Flow, *IMA J. Numer. Anal.*, 44(5):2667–2697, 2024.

3. Q. Rao, J. Wang, and Y. Xie. Optimal Convergence of Arbitrary Lagrangian–Eulerian Finite Element Methods for the Stokes Equation in an Evolving Domain, *IMA J. Numer. Anal.*, drae097, 2025.

4. B. Li, Q. Rao, H. Zhang, and Z. Zhou. Numerical Analysis of the 2D Navier–Stokes Equations with nonsmooth initial value in the Critical Space, submitted for publication.

5. B. Li, Q. Rao, and P. Sun. Optimal Convergence of an Arbitrary Lagrangian–Eulerian Finite Element Method for Fluid-Structure Interactions, submitted for publication.

# Acknowledgements

First and foremost, I wish to express my deepest gratitude to my esteemed supervisors, Professor Yanping Lin and Professor Buyang Li, whose extraordinary mentorship has profoundly shaped my academic journey. Their intellectual generosity and innovative thinking have been instrumental in cultivating my identity as an open-minded researcher. Their unwavering passion for academic inquiry and steadfast encouragement have constantly inspired me to pursue research with substantive societal impacts. Possessing exceptional academic vision, they have consistently provided discerning critiques that elevated the quality of my work, while their talent for identifying research potential has been pivotal in guiding my scholarly development.

I am particularly indebted to both supervisors for their unparalleled support in facilitating my research endeavors. Through their extensive academic networks, they have consistently connected me with valuable opportunities and resources. Whenever encountering challenges, I could invariably rely on their wholehearted commitment to ensure the continuity of my research progress. The successful completion of this work owes much to their expert guidance and the intellectually stimulating discussions we shared.

My sincere appreciation extends to Professor Jilu Wang, Professor Wei Gong, and Professor Zhi Zhou for their insightful suggestions and constructive feedback throughout this journey. Their exemplary dedication to academic excellence has established an inspiring model of rigorous scholarship, challenging me to uphold the highest standards in my research pursuits.

To my cherished colleagues and friends at PolyU, I extend heartfelt thanks for your invaluable assistance, stimulating academic exchanges, and the camaraderie that transformed challenges into rewarding experiences. Our collaborative spirit and shared moments of discovery have significantly enriched this endeavor.

Finally, I must reserve my most profound gratitude for my family. Your unconditional love, unwavering trust, and patient understanding have been my enduring foundation. The silent strength drawn from your support has sustained me through every phase of this

intellectual odyssey.

# Contents

**Certificate of Originality** ii

**Abstract** iii

**Publications arising from the thesis** vi

**Acknowledgements** vii

**List of Figures** xii

**1 Introduction** 1

**Part I: High-order Methods for Parabolic Equations with Rough Data** 8

**2 An Exponential Spectral Method Using VP Means for Semilinear Subdiffusion Equations with Rough Data** 10

2.1 Introduction . . . . . 10

2.2 Construction of the spectral method . . . . . 15

2.2.1 Bounded mild solutions to the subdiffusion equation . . . . . 16

2.2.2 Quadrature approximation of the mild solution . . . . . 17

2.2.3 Uniform polynomial approximation by VP means . . . . . 18

2.2.4 The exponential integrator using VP means . . . . . 20

2.3 The linear problem with a singular source . . . . . 20

2.4 The semilinear problem with rough initial data . . . . . 25

2.4.1 The spectral collocation method . . . . . 25

2.4.2 Existence, uniqueness and boundedness of the numerical solution . . 28

2.4.3 Error estimation . . . . . 29

2.5 Numerical tests . . . . . 32

2.5.1 The linear subdiffusion equation . . . . . 32

2.5.2 The semilinear subdiffusion equation . . . . . 34

**3 Numerical Analysis of the 2D Navier–Stokes Equations with nonsmooth initial value in the Critical Space 35**

3.1 Introduction . . . 35

3.2 Spatial semi-discretization by finite element method . . . 38

3.2.1 Weak solution . . . 38

3.2.2 Spatial semi-discretization . . . 40

3.3 Fully discretization . . . 48

3.3.1 Runge-Kutta method and error equations . . . 48

3.3.2 Regularities of numerical solutions and estimates for operators . . . 51

3.3.3 Error analysis . . . 56

3.4 Numerical examples . . . 64

**Part II: High-order Methods for Parabolic Equations in Evolving Domains 69**

**4 Convergent Evolving Finite Element Approximations of Boundary Evolution under Shape Gradient Flow 72**

4.1 Introduction . . . 72

4.2 Notation and main results . . . 76

4.2.1 Preliminaries of shape calculus . . . 76

4.2.2 Weak formulation of the PDEs . . . 77

4.2.3 The evolving finite elements . . . 78

4.2.4 The numerical method and its convergence . . . 80

4.3 Convergence of the numerical approximations . . . 83

4.3.1 Comparison of norms and integrals on two different domains . . . 84

4.3.2 Error equations and consistency estimates . . . 85

4.3.3 Stability estimates . . . 88

4.3.4 Proof of Theorem 4.1 . . . 95

4.4 Numerical examples . . . 95

**5 Optimal Convergence of Arbitrary Lagrangian–Eulerian Finite Element Methods for the Stokes Equation in an Evolving Domain 101**

5.1 Introduction . . . 101

5.2 Preliminary . . . 105

5.2.1 Evolving mesh and ALE finite element spaces . . . 105

5.2.2 Boundedness of partial derivatives of the mesh velocity . . . 108

5.2.3 Error of domain approximation . . . 109
5.3 The semidiscrete finite element approximation . . . 110
5.3.1 The Stokes–Ritz projection . . . 111
5.3.2 The Nitche's trick and duality argument . . . 114
5.3.3 Proof of Theorem 5.1 . . . 120
5.4 Numerical experiments . . . 123

**6 Optimal Convergence of an Arbitrary Lagrangian–Eulerian Finite Element Method for Fluid-Structure Interactions 129**

6.1 Introduction . . . 129
6.2 Model description . . . 132
6.3 ALE mapping and ALE weak formulation for fully-discrete scheme . . . 134
6.3.1 Interpolated ALE mapping and interpolated evolving mesh . . . 135
6.3.2 Weak formulation on the interpolated evolving mesh . . . 137
6.3.3 ALE mapping and weak formulation . . . 139
6.3.4 Fully-discrete scheme . . . 143
6.4 Error estimates . . . 144
6.4.1 Matrix-vector form of the weak formulations . . . 144
6.4.2 Comparison of integrals over two different domains . . . 148
6.4.3 Convergence of evolving finite element approximations . . . 150
6.5 Numerical Examples . . . 167

**7 Conclusion 172**

**Bibliography 177**

# List of Figures


2.1 Errors of the numerical solutions for the linear subdiffusion equation . . . . 33

2.2 Errors of the numerical solutions for the semilinear subdiffusion equation . . 33

3.1 $L^2$ errors of $u$ . . . . . 65

3.2 Isocontours of the velocity $u$ . . . . . 66

3.3 $L^2$ errors of $u$ . . . . . 67

3.4 Isocontours of the velocity $u$ . . . . . 68

4.1 Errors of the numerical solutions at time $T = 10$ (Example 4.1) . . . . . 97

4.2 Evolution of the 2D domain in Example 4.1. . . . . 98

4.3 Evolution of the 3D domain in Example 4.2. . . . . 98

4.4 Evolution of the shape of the obstacle in Example 4.3. . . . . 100

5.1 Meshes of $P_1$ and $P_2$ elements at time $T = 0$ and $T = 1$. . . . . 126

5.2 Errors from spatial discretization for $T = 1$. . . . . 127

5.3 Errors from temporal discretization for velocity $u$ at time $T = 1$. . . . . 127

5.4 Errors from spatial discretization for $T = 1$. . . . . 127

5.5 Errors from temporal discretization for velocity $u$ at time $T = 1$. . . . . 128

6.1 Two schematic domain decompositions divided by the interface $\Gamma(t)$: (1) the immersed case (left); (2) the back-to-back case (right). . . . . 132

6.2 Errors of the numerical solutions at time $T = 0.1$ . . . . . 169

6.3 Evolution of the domain and mesh . . . . . 169

6.4 Errors of the numerical solutions at time $T = 0.1$ . . . . . 171

6.5 Evolution of the domain and mesh . . . . . 171

# Chapter 1

## Introduction

This thesis is devoted to the development of high-order convergent numerical methods for nonlinear equations. It consists of two parts:

1. High-order methods for parabolic equations with rough data.
2. High-order methods for parabolic equations in evolving domains.

The first part of this thesis focuses on the numerical analysis of semilinear subdiffusion equations and Navier-Stokes (NS) equations under rough data conditions. Such problems arise pervasively in modeling physical phenomena characterized by singular perturbations or discontinuous inputs. For instance, localized disturbances in fluid dynamics or abrupt changes in material properties can induce solutions with severely limited regularity. Developing high-order convergent numerical methods for these scenarios has emerged as a critical research direction in recent years. The primary challenge stems from the low regularity of solutions caused by rough data, which often renders conventional numerical schemes, such as standard finite element or finite difference methods with uniform meshsizes, incapable of achieving optimal convergence rates. This necessitates the design of specialized algorithms that rigorously account for the solution's nonsmooth features while preserving accuracy and stability.

The construction of numerical methods fundamentally hinges on the structural characteristics of solutions and the nature of nonlinear terms. In scenarios where solutions exhibit singularities at isolated points while maintaining smoothness elsewhere, a mathematically rigorous approach involves developing high-order methods through graded meshes localized near singular regions. In Chapter 2, we use this method to construct a spectral convergent method for the following subequations:

$$\begin{cases} \partial_t^\alpha u(x,t) - \Delta u(x,t) = f(u(x,t),x,t) & \text{for } (x,t) \in \Omega \times (0,T], \\ u(x,t) = 0 & \text{for } (x,t) \in \partial\Omega \times (0,T], \\ u(x,0) = u_0(x) & \text{for } x \in \Omega, \end{cases} \tag{1.1}$$

where $f : \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}_+ \to \mathbb{R}$ and $u_0$ are the given nonlinear function and initial value, and $\partial_t^\alpha u$ denotes the Caputo fractional derivative of order $\alpha \in (0, 1)$. The solutions of equation (1.1) may have singularity at $t = 0$.

In general, for the linear subdiffusion equation with initial value $u_0 \in H_0^1(\Omega) \cap H^2(\Omega)$ and a temporally smooth source function $f(x, t)$ (independent of $u$), the solution generally exhibits the following type of weak singularity at $t = 0$ (see [85, Theorem 1]):

$$\|\partial_t^m(u(\cdot, t) - u_0)\|_{L^2} \leqslant C_m t^{\alpha-m} \quad \text{for } m \geqslant 0. \tag{1.2}$$

The extension of the above-mentioned results to rough initial data in $L^p(\Omega)$ with $1 \leqslant p \leqslant \infty$ (without any differentiability), using graded stepsizes to improve the convergence orders, is still challenging due to the stronger singularity in this case (see [85, Theorem 1]), i.e.,

$$\|\partial_t^m u(\cdot, t)\|_{L^p} \leqslant C_m t^{-m} \quad \text{for } m \geqslant 0. \tag{1.3}$$

Lots of high-order numerical methods are designed based on the estimates (1.2) and (1.3).

For the semilinear subdiffusion equation with nonsmooth initial data, the problem becomes more difficult since the source function $f(u(x, t), x, t)$ also becomes singular at $t = 0$ due to the singularity of $u(x, t)$. In this case, the regularity of the solution and the source function is

$$\|\partial_t^m(u(\cdot, t) - u_0)\|_{L^\infty} + \|\partial_t^m f(u(\cdot, t), \cdot, t)\|_{L^\infty} \leqslant C_m t^{\alpha-m} \quad \text{for } m \geqslant 0 \tag{1.4}$$

and

$$\|\partial_t^m u(\cdot, t)\|_{L^\infty} + \|\partial_t^m f(u(\cdot, t), \cdot, t)\|_{L^\infty} \leqslant C_m t^{-m} \quad \text{for } m \geqslant 0 \tag{1.5}$$

for initial data in $H_0^1(\Omega) \cap C^2(\overline{\Omega})$ and $L^\infty(\Omega)$, respectively; see [106, 173]. For fixed convergence order numerical methods, there are already some results, see [84, 2, 173]. But for the development of spectral methods, which converge faster than $O(N^{-k})$ for any fixed positive integer $k$ (where $N$ denotes the degrees of freedom in the time discretization), is still challenging for the subdiffusion equation with singular solutions and source functions. The objective of Chapter 2 is to fill in this gap.

When the nonlinear term exhibits high complexity, specialized techniques are required to achieve high-order convergence rates. In Chapter 3, we show how to construct high-order convergent numerical method for NS equations with $L^2$ initial value under no-slip boundary condition.

Optimal error estimates for high-order methods can be proved when the solutions to the Navier-Stokes equations are sufficiently regular, meaning they are sufficiently smooth

and adhere to the compatibility conditions. For example, if the initial values are sufficiently smooth, i.e. $u_0 \in \dot{H}_0^1(\Omega) \cap H^2(\Omega)^2$ or above, then optimal-order convergence of temporal and spatial discretizations of the NS equations have been proved for various methods in [7, 10, 52, 70, 97, 80, 148, 166, 67, 157, 158, 77], where the finite element and spectral Galerkin methods were usually used for spatial discretization, and the time-stepping schemes include varies of the Crank–Nicolson method, Euler method, two-step backward difference formula, projection methods, fractional step methods and so on. However, the error estimates discussed in the aforementioned articles are not applicable to nonsmooth initial data.

Discussions concerning the case that $u_0 \in \dot{L}^2(\Omega)$ are less prevalent in the literature. It has been known that $\dot{L}^2(\Omega)$ is a critical space for the well-posedness of the two-dimensional NS equations [51]. The error analysis in this case turns out to be significantly more challenging than for cases with smoother initial data, and the literature offers only a limited number of relevant results. Under the CFL condition, $\tau \leqslant C\lambda_m^{-1}$, it was shown in [65] that the implicit-explicit Euler spectral Galerkin method has an error bound of $O(\lambda_m^{-1/2} + \tau^{1/2})$ over a bounded time interval. For the implicit-explicit Euler scheme with finite element spatial discretization, several stability results were proved in [68] without error estimates. In more recent developments, first-order convergence in both time and space was shown in [110] for high-order divergence-free finite elements. To our knowledge, this represents the most advanced convergence result obtained to date. However, there is still a gap between the numerical analysis and the numerical results, which demonstrate the possibility of achieving second-order convergence in space by using the Taylor–Hood finite elements. Proving second-order convergence of any numerical method for the NS equations remains an open and challenging task.

In Chapter 3, we consider a fully discrete implicit-explicit Runge–Kutta finite element scheme for the NS equations with $L^2$ initial data by utilizing an $L^2$ projection $P_h^{\mathrm{RT}}$ onto the divergence-free subspace of the Raviart–Thomas element space in the numerical scheme. The linear term is discretized using the Runge–Kutta Lobatto IIIC scheme, while the nonlinear term is handled through an extrapolation approximation. To address the solution's singularity near $t = 0$, we employ graded stepsizes that provide enhanced resolution where needed. We prove the a nearly optimal error estimate. More specifically, let $u_h^{n+1}$ be the numerical solution of the fully discrete scheme at time level $t = t_{n+1}$. Theorems 3.1 and 3.2 show that, for arbitrarily small $\varepsilon > 0$,

$$\|u(t_{n+1}) - u_h^{n+1}\|_{L^2(\Omega)} \leqslant C_\varepsilon(h^{2-2\varepsilon} t_{n+1}^{\varepsilon-1} + t_{n+1}^{-2-\varepsilon}\tau_{n+1}^2),$$

where $\tau_{n+1}$ and $h$ denote the temporal stepsize of the $(n+1)$th step and spatial mesh size, respectively.

The second part of this thesis investigates the optimal-order convergence of evolving finite element methods (FEM) for shape gradient flows, Stokes equations and fluid–structure interaction (FSI) problem under $H^1$- and $L^2$-norm error metrics. These problems, characterized by solution-driven domain boundaries or interfaces, represent a class of nonlinear PDEs with evolving geometries. A critical challenge arises from the time-dependent domain deformation, which introduces two fundamental difficulties:

1. Dynamic Mesh Adaptation: The evolving domain necessitates distinct finite element spaces at different temporal discretization levels. To ensure robust approximation properties, we construct hierarchically compatible finite element spaces that preserve geometric consistency across deformations.

2. Nonlinear Coupling Analysis: The interplay between domain evolution and PDE solutions complicates traditional error estimation frameworks. To address this, we develop matrix-vector formulations that decouple geometric motion from physical field variations, enabling rigorous convergence analysis through discrete energy arguments.

In Chapter 4, we will prove the optimal-order convergence of $H^1$-norm for shape optimization problem constrained by PDEs. It typically concerns minimizing a shape functional

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

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 $\Omega$ with boundary $\Gamma$. A popular way to compute critical points of shape optimziation problems without requiring the second-order optimality is through shape gradient flow, which is a specific type of shape gradient descent algorithm. 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 $\Gamma(t)$.

There are many different ways to choose the descent velocity $w$, most of which are based on solving the following equation (cf. [3, 16])

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

for an abstract inner product $b(\cdot, \cdot) : H \times H \to \mathbb{R}$ associated with some Hilbert space $H$. In Chapter 4, 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.,

$$\min_{\Gamma} J(\Gamma) = \int_{\Omega} j(x,u)\mathrm{d}x \quad \text{subject to} \quad \begin{cases} -\Delta u = f & \text{in } \Omega \subset \mathbb{R}^d, \text{ with } d \in \{2,3\}, \\ u = 0 & \text{on } \Gamma = \partial\Omega, \end{cases} \tag{1.6}$$

with the shape densify 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 (4.1), 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:

$$\partial_t \phi = w \circ \phi \quad \text{in } \Omega^0, \qquad \phi(0) = \mathrm{id}|_{\Omega^0} \quad \text{in } \Omega^0, \tag{1.7a}$$

$$-\Delta w + w = 0 \quad \text{in } \Omega(t), \qquad \partial_\nu w = -J'(\Gamma(t))\nu \quad \text{on } \Gamma(t), \tag{1.7b}$$

$$-\Delta u = f \quad \text{in } \Omega(t), \qquad u = 0 \quad \text{on } \Gamma(t), \tag{1.7c}$$

$$-\Delta p = j_u'(\cdot,u) \quad \text{in } \Omega(t), \qquad p = 0 \quad \text{on } \Gamma(t), \tag{1.7d}$$

where $\phi(\cdot,t) : \Omega^0 \to \Omega(t)$ 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 $u$. We proved the optimal-order convergence for this problem by using evolving FEM, which is still absent in literatures.

In Chapter 5, we further prove the optimal-order convergence of $L^2$ norm for Stokes and Navier–Stokes equations in an evolving domain. For simplicity, we consider the Stokes equations in an evolving domain with given velocity. For NS equations, the analysis is similar. Let $\Omega(t) \subset \mathbb{R}^d$, $d \in \{2,3\}$, be a time-dependent domain with a smooth boundary $\Gamma(t) = \partial\Omega(t)$ moving under a given smooth vector field $w(\cdot,t)$.

The analysis of ALE methods for the Stokes and NS equations has yielded noteworthy results but remained suboptimal, as discussed below. In [101], Legendre & Takahashi introduced a novel approach that combines the method of characteristics with finite element approximation to the ALE formulation of the Navier–Stokes equations in two dimensions, and established an $L^2$ error estimate of $O(\tau + h^{1/2})$ for the $P_{1b}$–$P_1$ elements under certain restrictions on the time step size. In a related work [151], an error estimate of $O(h^2|\log h|)$ was obtained for the ALE semidiscrete FEM with the Taylor–Hood $P_2$–$P_1$ elements for the Stokes equations in a time-dependent domain. Moreover, for a fully discrete ALE method with the implicit Euler scheme in time, convergence of $O(\tau + h^2 + h^2/\tau)$ was proved in [151]. The errors of ALE finite element solutions to the Stokes equations on a time-varying

domain, with BDF-$k$ in time (for $1 \leqslant k \leqslant 5$) and the Taylor–Hood $P_r$–$P_{r-1}$ elements in space (with degree $r \geqslant 2$), were shown to be $O(\tau^k + h^r)$ in the $L^2$ norm in [119].

As far as we know, optimal-order convergence of ALE semidiscrete and fully discrete FEMs were not established for the Stokes and NS equations in an evolving domain. As shown in [54, 113], the optimal-order convergence of ALE semidiscrete FEM requires proving the following optimal-order approximation property for the material derivative of the Ritz projection:

$$\|D_{t,h} R_h u - R_h D_{t,h} u\|_{L^2} \leqslant C h^{r+1}. \tag{1.8}$$

In line with the fixed domain case, achieving optimal consistency error in analysis of finite element approximation for Stokes equation necessitates the use of the Stokes-Ritz projection $R_h$. As a result, when trying to obtain the optimal-order approximation property (5.4) following the duality argument as in [54, 113], a problem occurs that the error estimate of Stokes-Ritz projection of pressure is involved in the analysis. This problem was addressed by additionally establishing and utilizing an optimal $H^{-1}$ error estimate for the Stokes-Ritz projection of pressure, i.e., (5.43), which is used in Lemma 5.11. This leads to optimal-order convergence of the ALE semidiscrete FEM, as the main result of Chapter 5 (see Theorem 5.1).

Chapter 6 focuses on optimal-order convergence analysis for fluid–structure interaction (FSI) problems. The bidirectional coupling between fluid and solid phases induces solution nonsmoothness across the dynamic interface due to discontinuities in material properties and kinematic constraints. However, solutions often exhibit piecewise smoothness within individual subdomains (fluid and solid regions), which motivates the design of high-order convergent evolving FEM tailored for FSI systems.

Recently, the errors of semi- or fully discrete ALE-FEMs for FSI problems were analyzed in [99, 100] without considering the errors in approximating the interface motion. To the best of our knowledge, the only proof of convergence of ALE-FEMs for FSI problems with a moving interface was given in [101] for rigid structures which can be modelled by ordinary differential equations instead of PDEs. Consequently, a comprehensive error analysis of ALE-FEMs for FSI problems that addresses the errors in approximating the interface motion, with elastic structures which need to be modelled by PDEs instead of ODEs, is still lacking.

The objective of Chapter 6 is to establish the optimal-order convergence in $L^\infty(0, T; H^1)$ of a monolithic fully discrete ALE-FEM with evolving mesh for a dynamic FSI problem with an unknown interface evolution driven by the solution of FSI problem. In particular,

the errors of fluid velocity and structural displacement in $L^\infty(0,T;H^1)$ norm and the error of pressure in $L^\infty(0,T;L^2)$ norm are proved to be $O(\tau+h^k)$ for a semi-implicit Euler scheme with a high-order evolving mesh ALE-FEM described in [41], using a mixed finite element method with $P_k$–$P_{k-1}$–$P_k$ elements for velocity, pressure and structural displacement, respectively. This is achieved by designing an initial correction that improves the order of convergence in the error analysis of the developed ALE-FEM for the presented FSI problem, utilizing the matrix-vector formulation techniques developed in [95] for analyzing the errors in approximating solution-driven interface evolution in the FSI problem, as well as a mathematical induction which guarantees that the errors in approximating the evolving domain (including the unknown interface) will be uniformly bounded and sufficiently small when $\tau$ and $h$ are small enough and $\tau = o(h^{\frac{d}{2}})$.

The content of this thesis consists of the research in the following papers of the author.

### Papers published

1. B. Li, Y. Lin, S. Ma, and Q. Rao. An Exponential Spectral Method Using VP Means for Semilinear Subdiffusion Equations with Rough Data, *SIAM J. Numer. Anal.*, 61(5):2305–2326, 2023.

2. W. Gong, B. Li, and Q. Rao. Convergent Evolving Finite Element Approximations of Boundary Evolution under Shape Gradient Flow, *IMA J. Numer. Anal.*, 44(5):2667–2697, 2024.

3. Q. Rao, J. Wang, and Y. Xie. Optimal Convergence of Arbitrary Lagrangian–Eulerian Finite Element Methods for the Stokes Equation in an Evolving Domain, *IMA J. Numer. Anal.*, drae097, 2025.

### Papers submitted

4. B. Li, Q. Rao, H. Zhang, and Z. Zhou. Numerical Analysis of the 2D Navier–Stokes Equations with nonsmooth initial value in the Critical Space, submitted for publication.

5. B. Li, Q. Rao, and P. Sun. Optimal Convergence of an Arbitrary Lagrangian–Eulerian Finite Element Method for Fluid-Structure Interactions, submitted for publication.

# Part I: High-order Methods for Parabolic Equations with Rough Data

## Brief Introduction

The first part of the thesis is concerned with the semilinear subdiffusion equation and NS equation with rough data.

In Chapter 2 we propose a new spectral method for the linear and semilinear subdiffusion equations with possibly discontinuous rough initial data. The new method effectively combines several computational techniques, including the contour integral representation of the solutions, the quadrature approximation of contour integrals, the exponential integrator using the de la Vallée Poussin (VP) means of the source function, and a decomposition of the time interval geometrically refined towards the singularity of the solution and the source function. Rigorous error analysis shows that the proposed method has spectral convergence for the linear and semilinear subdiffusion equations with bounded measurable initial data and possibly singular source functions under the natural regularity of the solutions.

In Chapter 3, we propose a fully discrete numerical scheme that utilizes the Taylor–Hood or Stokes-MINI finite element method (FEM) for spatial discretization and an implicit–explicit Runge–Kutta (RK) time-stepping method in conjunction with graded stepsizes. By employing discrete semigroup techniques, sharp regularity estimates, negative norm estimates and the $L^2$ projection onto the divergence-free Raviart–Thomas (RT) element space, we prove that the proposed scheme attains second-order convergence in both space and time. Numerical examples are presented to support the theoretical analysis. In particular, the convergence in space is at most second order even higher-order finite elements are used. This shows the sharpness of the convergence order proved in this chapter.

## Papers published & submitted

1. B. Li, Y. Lin, S. Ma, and Q. Rao. An Exponential Spectral Method Using VP Means for Semilinear Subdiffusion Equations with Rough Data, *SIAM J. Numer. Anal.*, 61(5):2305–2326, 2023.

2. B. Li, Q. Rao, H. Zhang, and Z. Zhou. Numerical Analysis of the 2D Navier–Stokes Equations with nonsmooth initial value in the Critical Space, submitted.

# Chapter 2

## An Exponential Spectral Method Using VP Means for Semilinear Subdiffusion Equations with Rough Data

The content of this chapter has been published in "*B. Li, Y. Lin, S. Ma, and Q. Rao. An Exponential Spectral Method Using VP Means for Semilinear Subdiffusion Equations with Rough Data, SIAM J. Numer. Anal., 61(5):2305–2326, 2023.*"

**2.1 Introduction.** We consider the semilinear subdiffusion equation in a bounded Lipschitz domain $\Omega \subset \mathbb{R}^d$ up to a given time $T > 0$, under the Dirichlet boundary condition, i.e.,

$$\begin{cases} \partial_t^\alpha u(x,t) - \Delta u(x,t) = f(u(x,t),x,t) & \text{for } (x,t) \in \Omega \times (0,T], \\ u(x,t) = 0 & \text{for } (x,t) \in \partial\Omega \times (0,T], \\ u(x,0) = u_0(x) & \text{for } x \in \Omega, \end{cases} \tag{2.1}$$

where $f : \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}_+ \to \mathbb{R}$ and $u_0$ are the given nonlinear function and initial value, and $\partial_t^\alpha u$ denotes the Caputo fractional derivative of order $\alpha \in (0,1)$. The subdiffusion equations which can model the sublinear growth of mean squared particle displacement have generated much interests from physicists, engineers and applied mathematicians in developing new computational methods and rigorous numerical analyses because of their excellent capability in modelling the anomalous transport processes. The construction and analysis of high-order computational methods for the subdiffusion equations, especially for the semilinear subdiffusion equation, have been challenging due to the possible singularity of the solution and the source function at $t = 0$.

In general, for the linear subdiffusion equation with initial value $u_0 \in H_0^1(\Omega) \cap H^2(\Omega)$ and a temporally smooth source function $f(x,t)$ (independent of $u$), the solution generally exhibits the following type of weak singularity at $t = 0$ (see [85, Theorem 1]):

$$\|\partial_t^m (u(\cdot,t) - u_0)\|_{L^2} \leqslant C_m t^{\alpha - m} \quad \text{for } m \geqslant 0. \tag{2.2}$$

Under this limited regularity condition, the classical L1, L2, dG and convolution quadrature (CQ) with a uniform stepsize generally have first-order convergence; for example, see [82, 134, 83, 175]. The analyses in [90, 92, 138, 162] show that the L1 and L2 methods, and a low-order dG time-stepping method, can have the desired optimal-order convergence by using graded stepsizes locally refined towards $t = 0$. The extension of these low-order convergence results to the semilinear subdiffusion equation may be established by using the fractional version of discrete Gronwall's inequalities in [84, 117, 118]. The sharp pointwise-in-time error bounds on quasi-graded temporal meshes with arbitrary degree of grading are obtained in [89] using method of upper and lower solutions for L1 scheme. Higher-order sub-optimal convergence in time was proved for the dG time-stepping method in [137] under condition (2.2) and some additional regularity assumptions such as $\partial_t u \in L^2(0,T;H^2(\Omega))$, which generally requires the initial value to satisfy $u_0 \in H^{5/2}(\Omega) \cap H_0^1(\Omega)$ plus a compatibility condition $\Delta u_0 = 0$ on $\partial\Omega$. In the case $u_0 = 0$, high-order convergence of the Runge–Kutta convolution quadrature was proved in [8].

The extension of the above-mentioned results to rough initial data in $L^p(\Omega)$ with $1 \leqslant p \leqslant \infty$ (without any differentiability), using graded stepsizes to improve the convergence orders, is still challenging due to the stronger singularity in this case (see [85, Theorem 1]), i.e.,

$$\|\partial_t^m u(\cdot,t)\|_{L^p} \leqslant C_m t^{-m} \quad \text{for } m \geqslant 0. \tag{2.3}$$

Under this regularity condition, the L1, L2 and CQ schemes with a uniform stepsize and appropriate initial corrections can have high-order convergence at time levels far away from $t = 0$, i.e.,

$$\|u(\cdot,t_n) - u_n\|_{L^p} \leqslant C t_n^{-k} \tau^k, \tag{2.4}$$

where $u_n$ denotes the numerical solution using a uniform stepsize $\tau$. The results were established in [83, 128, 175] in the $L^2$-norm framework, i.e., with error estimates in the $L^2$ norm and initial data in $L^2(\Omega)$, by comparing the numerical solution with the solution through their Laplace transform representations, a framework developed by Lubich for the analysis of CQ for convolution integrals; see [123, 124, 125]. Nevertheless, the analyses can be naturally extended to the $L^p$-norm framework by using the corresponding resolvent estimate in $L^p(\Omega)$, i.e.,

$$\|(z - \Delta)^{-1}\|_{L^p(\Omega)\to L^p(\Omega)} \leqslant C|z|^{-1}, \quad 1 \leqslant p \leqslant \infty,$$

which holds for both the Laplacian $\Delta$ and the finite element discrete Laplacian $\Delta_h$; see [104,

105]. However, these high-order convergence results only hold for the linear subdiffusion equation with a given temporally smooth source function.

The extension of the high-order methods to the semilinear subdiffusion equation with nonsmooth initial data is still challenging. The main difficulty in the numerical solution of the semilinear subdiffusion equation is that the source function $f(u(x,t),x,t)$ also becomes singular at $t=0$ due to the singularity of $u(x,t)$. In this case, the regularity of the solution and the source function is

$$\|\partial_t^m(u(\cdot,t)-u_0)\|_{L^\infty}+\|\partial_t^m f(u(\cdot,t),\cdot,t)\|_{L^\infty}\leqslant C_m t^{\alpha-m}\quad \text{for } m\geqslant 0 \tag{2.5}$$

and

$$\|\partial_t^m u(\cdot,t)\|_{L^\infty}+\|\partial_t^m f(u(\cdot,t),\cdot,t)\|_{L^\infty}\leqslant C_m t^{-m}\quad \text{for } m\geqslant 0 \tag{2.6}$$

for initial data in $H_0^1(\Omega)\cap C^2(\overline{\Omega})$ and $L^\infty(\Omega)$, respectively; see [106, 173]. The convergence of the backward Euler CQ was not affected by the singularity of the source function, as shown in [84] and [2] for initial data in $H_0^1(\Omega)\cap H^2(\Omega)$ and $H^s(\Omega)$, $s>0$, respectively. However, the convergence of higher-order CQs can be affected by the singularity of the source function at $t=0$. In particular, the analysis and numerical experiments in [173] show that the convergence order of high-order BDF-CQs, using a uniform stepsize with appropriate initial corrections, is at most $1+2\alpha$ for general initial data in $H_0^1(\Omega)\cap C^2(\overline{\Omega})$. More recently, a $k$-step exponential CQ was proposed in [106] to achieve $k$th-order convergence for semilinear subdiffusion equation with initial data in $L^\infty(\Omega)$, where $k$ is an arbitrary prescribed positive integer which determines the numerical scheme and the convergence order.

The above-mentioned methods all have fixed convergence orders. The development of spectral methods, which converge faster than $O(N^{-k})$ for any fixed positive integer $k$ (where $N$ denotes the degrees of freedom in the time discretization), is still challenging for the subdiffusion equation with singular solutions and source functions. Many efficient spectral methods for the fractional ordinary and partial differential equations have been proposed and analyzed:

- A new multi-domain spectral method for high-order time discretizations of fractional ordinary differential equations (ODEs) and the subdiffusion equation was proposed in [21]. The stability of the method was studied by identifying the method as a generalized linear method. The rigorous analysis of its spectral convergence for the subdiffusion equation with rough initial data still remains challenging.

- A class of spectral collocation methods based on the generalized Jacobi functions was proposed in [179, 180] for some fractional differential equations. The approximation errors of the generalized Jacobi polynomials in weighted Sobolev spaces, as well as the convergence of the spectral Petrov–Galerkin method for fractional ordinary differential equations (ODEs), were presented in [24]. The current analysis requires some fractional derivatives of the solution to be smooth, which is true for some fractional ODEs but cannot be readily extended to the subdiffusion equation. The rigorous analysis of this class of methods for the subdiffusion equation with rough initial data still remains challenging.

- The generalied Jacobi polynomials were also used for calculating the eigenvalues of the space-fractional differential equations in [22] and for determining the superconvergence points of fractional spectral interpolations in [181].

- A class of Müntz spectral Galerkin methods were proposed and analyzed in [75, 74] based on the Müntz polynomials and weighted Sobolev spaces. The methods have spectral convergence for solutions with the following regularity:

$$t^{-1+m}\partial_t^m[u(x,t^{1/\lambda})] \in L^2(0,T;H_0^1(\Omega)) \quad \forall\, m \geqslant 0. \tag{2.7}$$

  This covers a wide class of solutions, including solutions in the form of

$$u(x,t) = \sum_{j=1}^{\infty} t^{j\alpha}\phi_j(x),$$

  which can be approximated with spectral convergence by choosing $\lambda = 1/\alpha$. For more general solutions of the subdiffusion equation with initial data in $H_0^1(\Omega) \cap H^2(\Omega)$, when the solutions have the regularity in (2.2) but may not satisfy condition (2.7), the Müntz spectral Galerkin methods can still have a fixed high order of convergence by choosing a sufficiently small parameter $\lambda$.

- A class of log-orthogonal spectral methods for the subdiffusion equation was proposed and analyzed in [25] by using the log-orthogonal polynomials introduced in [23]. It is shown that the method can have spectral convergence if the initial value satisfies $u_0 \in \dot{H}^3(\Omega)$, which is equivalent to requiring $u_0 \in H_0^1(\Omega) \cap H^3(\Omega)$ plus one compatibility condition $\Delta u_0 = 0$ on $\partial\Omega$.

As far as we know, the following questions are still unsolved in the development of spectral methods for the subdiffusion equations.

- The current error analysis of spectral methods for the subdiffusion equation requires the initial data to be smoother than $H_0^1(\Omega)\cap H^2(\Omega)$ and satisfying some compatibility conditions, or satisfying a regularity condition in the form of (2.5). The construction of spectral methods for the semilinear subdiffusion equation with rough initial data in $L^\infty(\Omega)$ with strong singularity in the form of (2.6) is still challenging.

- The existing error estimates for the spectral methods are all for the linear subdiffusion equation based on the Hilbert space framework. The extension of the error analysis for the spectral methods to the semilinear subdiffusion equation, with nonlinear source functions which may not be globally Lipschitz continuous, requires error estimates in the $L^\infty$-norm based Banach space framework and therefore still remains open. There are few analysis of spectral methods in the $L^\infty$-norm framework. We are only aware of the $L^\infty$-norm analysis of spectral methods in [116] by the effective maximum principle of spatial discretizations. The techniques cannot be applied to the time discretization of the subdiffusion equation in the presence of singularity.

These questions are addressed in this chapter.

We propose a new spectral method for the semilinear subdiffusion equation with rough initial data in $L^\infty(\Omega)$ under the natural regularity condition (2.6), based on quadrature approximations to the contour integral representation of the solution, the exponential integrator using the de la Vallée Poussin means (i.e., VP means, see [165]), and error estimates via a fixed-point argument in an $L^\infty$-norm framework. The contour integral approximation techniques were used in [121] and [58] for evaluating exponential-type of operators and for solving linear convection–diffusion equations, respectively. The techniques were also used in the construction of a high-order backward extrapolated multi-step exponential convolution quadrature for the semilinear subdiffusion equation in [106]. However, the method in [106] only has a fixed order of convergence and cannot be extended to a spectral method directly due to the following three challenges: (1) The temporal Lagrange interpolation used in [106] is not stable in the $L^\infty$ norm as the degrees of freedom tend to infinity; (2) The integrals of the exponential function times the Lagrange basis is difficult to be evaluated analytically when the degree of the polynomial is large; (3) The source function is singular at $t = 0$ and therefore the Lagrange interpolation on the whole time interval does not have uniform high-order convergence.

We overcome these difficulties by utilizing an exponential type of spectral method in terms of the interpolation polynomial VP means on subintervals that are geometrically

refined towards $t = 0$ according to the singularity of the solution and the source function. We derive analytical formulas for the exponential integrals arising from the proposed method by utilizing the differentiation properties of the Jacobi polynomials, and prove the spectral convergence of the proposed method based on the natural regularity condition in (2.6) using the $L^\infty$ stability of the polynomial VP means and their approximation properties on the geometrically refined subintervals. The spectral convergence in the $L^\infty$ norm is established, which allows us to handle nonlinear source functions which are only locally Lipschitz continuous (rather than globally Lipschitz continuous). These results make the proposed method practically computable and spectrally convergent for rough solutions of the semilinear subdiffusion equation in the most general setting.

In view of the equivalence between CQs and their Laplace transform representations, e.g., the Runge–Kutta (or BDF) CQ for a parabolic or subdiffusion equation is equivalent to applying the Runge–Kutta method (or BDF) method to the ODE arising from the Laplace transform representation of the solution (as shown in [126]), the exponential spectral method proposed in this chapter can also be viewed as a contour integral approximation to a spectrally convergent CQ constructed by utilizing an exponential integrator for the ODE arising from the Laplace transform representation of the solution, and by utilizing VP means of the source function on geometrically refined subintervals adapted to the singularity at $t = 0$.

The rest of this chapter is organized as follows. The quadrature approximation of the mild solution and the uniform polynomial approximation of functions by VP means are presented in Section 2.2. The new spectral method and its error analysis for the linear subdiffusion equation with a possibly singular source function are presented in Section 2.3. The new spectral method and its error analysis for the semilinear subdiffusion equation are presented in Section 2.4. Numerical experiments are presented in Section 2.5 to illustrate the spectral convergence of the proposed method for linear and semilinear subdiffusion problems with rough initial data in $L^\infty(\Omega)$.

**2.2 Construction of the spectral method.** In this section, we introduce the several ingredients we use to construct the spectral method, including the contour integral representation of the solution, quadrature approximation to the contour integrals, polynomial approximation by the VP means, and the exponential integrator using VP means. The construction of the spectral method is discussed at the end of this section, while the complete algorithms are presented in the next two sections for the linear and semilinear

subdiffusion equations, respectively.

**2.2.1 Bounded mild solutions to the subdiffusion equation.** For the simplicity of notations, we denote $u(t) = u(\cdot, t)$ for any function $u$ defined on $\Omega \times (0, T]$. A function $u \in L^\infty(0, T; L^\infty(\Omega)) \cap C([0, T]; L^2(\Omega))$ is called a bounded mild solution of (2.1) if it satisfies the following equation

$$u(t) = F(t)u_0 + \int_0^t E(t-s)f(u(s), \cdot, t)\mathrm{d}s \quad \forall\, t \in (0, T], \tag{2.8}$$

where $F(t)$ and $E(t)$ are the solution operators associated to the subdiffusion equation, defined as the inverse Laplace transform of the operators $z^{\alpha-1}(z^\alpha - \Delta)^{-1}$ and $(z^\alpha - \Delta)^{-1}$, respectively, i.e.,

$$F(t) = \frac{1}{2\pi i}\int_{\mathrm{Re}(z)=\sigma} z^{\alpha-1}(z^\alpha - \Delta)^{-1}e^{zt}\mathrm{d}z$$
$$E(t) = \frac{1}{2\pi i}\int_{\mathrm{Re}(z)=\sigma} (z^\alpha - \Delta)^{-1}e^{zt}\mathrm{d}z,$$

with an arbitrary parameter $\sigma > 0$. The paths of integration in the expressions of $F(t)$ and $E(t)$ can be respectively deformed to the following two contours on the complex plane:

$$\Gamma_\lambda = \{\lambda(1 - \sin(\beta + is)) : s \in \mathbb{R}\} \quad \text{and} \quad \Gamma_{\tilde\lambda} = \{\tilde\lambda(1 - \sin(\beta + is)) : s \in \mathbb{R}\}, \tag{2.9}$$

where $\beta \in (0, \varphi - \frac{\pi}{2})$, $\varphi \in (\frac{\pi}{2}, \pi)$ and $\lambda, \tilde\lambda > 0$ can be arbitrary fixed parameters. These contours are contained in the region between the two sectors $\Sigma_\varphi = \{z \in \mathbb{C} : |\arg(z)| \leqslant \varphi\}$ and $\lambda + \Sigma_{\beta+\frac{\pi}{2}}$; see [106, Figure 2.1]. Therefore,

$$\mathrm{Re}(z) \sim -|z| \quad \text{and} \quad |\mathrm{Im}(z)| \sim |\mathrm{Re}(z)| \quad \text{for } z \in \Gamma_\lambda \cup \Gamma_{\tilde\lambda}.$$

The representation of the solutions to the subdiffusion equation in (2.8) has been used in analyzing the regularity of solutions and the error of numerical approximations in [84, 106, 128, 173, 175]. The integration contours in the form of (2.9) have been used in approximating exponential-type functions of elliptic operators in [58, 121].

It is known that the resolvent operators $(z^\alpha - \Delta)^{-1}$ of the Dirichlet Laplacian satisfy the following estimate for some constant $C > 0$ (for example, see [106, Appendix A]):

$$\|(z^\alpha - \Delta)^{-1}\|_{L^\infty \to L^\infty} \leqslant C(1 + |z|)^{-\alpha} \text{ for } z \in \Gamma_\lambda. \tag{2.10}$$

As a result of the resolvent estimate in (2.10), the solution operators $F(t)$ and $E(t)$ were proved to satisfy the following estimates (cf. [106, Lemma 3.1]):

$$\|F(t)\|_{L^\infty \to L^\infty} \leqslant C \quad \text{and} \quad \|E(t)\|_{L^\infty \to L^\infty} \leqslant Ct^{\alpha-1}. \tag{2.11}$$

By using these properties of the solution operators, it can be shown that equation (2.1) indeed has a bounded mild solution if the one of the following two conditions is satisfied:

(1) The nonlinear function $f(u)$ is Lipschitz continuous with respect to $u$. In this case, the proof of well-posedness in [84] is still valid if the underlying space $L^2(\Omega)$ is replaced by $L^\infty(\Omega)$, as the proof only requires using the resolvent estimate in (2.10).

(2) The nonlinear function $f(u) = -F'(u)$ is the derivative of a double-well potential $F(u)$ with two wells at $\pm\alpha$, and the initial value satisfies that $|u_0| \leqslant \alpha$. For example, $f(u) = (u-u^3)/\varepsilon^2 = -F'(u)$ with $F(u) = (1-u^2)^2/(4\varepsilon^2)$ being the Ginzburg–Landau potential. In this case, the maximum principle of the subdiffusion equation (cf. [178, Theorems 3.1–3.2], [129, Theorem 2.1] and [91]) guarantees the boundedness of the solution, i.e., $|u(x,t)| \leqslant \alpha$.

In view of these results, we simply assume that the semilinear subdiffusion equation has a bounded mild solution and propose a class of spectral methods for the linear and semilinear problems, respectively, with rigorous analysis for the existence, uniqueness and spectral convergence of the numerical approximations.

Throughout this chapter, we denote by $C$ a generic positive constant that may be different at different occurrences but is independent of the number $N_m$, $N_1(m)$ of subintervals, the time level $t_n$, and the number $M$ of quadrature points for approximating the contour integrals.

**2.2.2 Quadrature approximation of the mild solution.** By substituting the contour integral expressions of $F(t)$ and $E(t)$ into (2.8), and make a change of variable in the first integral of (2.8), we can express the mild solution as

$$\begin{aligned} u(t) =& \frac{1}{2\pi i}\int_{\Gamma_\lambda} e^z z^{\alpha-1}(z^\alpha - t^\alpha \Delta)^{-1} u_0 \mathrm{d}z \\ &+ \frac{1}{2\pi i}\int_{\Gamma_{\tilde{\lambda}}} (z^\alpha - \Delta)^{-1} \int_0^t e^{z(t-s)} f(u(s),\cdot,s)\mathrm{d}s\mathrm{d}z, \end{aligned} \tag{2.12}$$

It is also known that the two contour integrals on the right-hand side of (2.12) can be approximated by a quadrature which has spectral convergence with respect to the number

of quadrature points, i.e.,

$$\begin{aligned} u(t) = &\sum_{j=-M}^{M} \omega_j e^{z_j} z_j^{\alpha-1} (z_j^\alpha - t^\alpha \Delta)^{-1} u_0 + \mathcal{E}_{1,q}(t) \\ &+ \sum_{j=-M}^{M} \tilde{\omega}_j (\tilde{z}_j^\alpha - \Delta)^{-1} y(\tilde{z}_j, t) + \mathcal{E}_{2,q}(t), \end{aligned} \tag{2.13}$$

where $2M+1$ quadrature points are used, $(\omega_j, z_j)$ and $(\tilde{\omega}_j, \tilde{z}_j)$ are two pairs of quadrature weights and nodes, with $\mathcal{E}_{1,q}(t)$ and $\mathcal{E}_{2,q}(t)$ denoting the remainders of the quadrature approximations, and $y(z,t) = \int_0^t e^{z(t-s)} f(u(s), \cdot, s) \mathrm{d}s$ is the solution of the following ordinary differential equation (ODE):

$$\frac{\mathrm{d}}{\mathrm{d}t} y(z,t) = z y(z,t) + f(u(t), \cdot, t). \tag{2.14}$$

The explicit expressions of the quadrature weights and nodes can be found in [106]. Moreover, the remainders satisfy the estimates in the following lemma.

**Lemma 2.1** (cf. [106, Lemma 3.3]). *If* $\|f(u,\cdot,\cdot)\|_{L^\infty(0,t;L^\infty(\Omega))} \leqslant C$ *then*

$$\|\mathcal{E}_{1,q}(t)\|_{L^\infty(\Omega)} + \|\mathcal{E}_{2,q}(t)\|_{L^\infty(\Omega)} \leqslant C e^{-\sqrt{M}/C}. \tag{2.15}$$

The representation of the mild solution by the discrete contour integrals in (2.13)–(2.14) was proposed in [106] for solving the semilinear subdiffusion equation by an exponential CQ with $k$-step extrapolation, which can only have $k$th-order convergence with a fixed $k \geqslant 1$. In the next subsection, we propose a spectrally convergent method based on the expression in (2.13)–(2.14) and a uniform polynomial interpolation technique using the VP means.

**2.2.3 Uniform polynomial approximation by VP means.** We denote the space of polynomials of degree $\leqslant m$ on the interval $[-1,1]$ by $P_m([-1,1])$, and define the $L^\infty$-norm polynomial approximation error as

$$E_m(g)(x) := \inf_{p \in P_m} \|g(x,\cdot) - p(\cdot)\|_{L^\infty(-1,1)} \quad \text{for } g \in C(\overline{\Omega} \times [-1,1]). \tag{2.16}$$

In the approximation theory, the following results are known (see [19, (5.4.16)]).

**Lemma 2.2.** *For* $0 \leqslant k \leqslant m$ *and* $g \in C^k([-1,1]; L^\infty(\Omega))$ *the following estimates hold:*

$$\begin{aligned} |E_m(g)(x)| &\leqslant C_k m^{-k} \|g(x,\cdot)\|_{C^k([-1,1])}, \\ \|E_m(g)\|_{L^\infty(\Omega)} &\leqslant C_k m^{-k} \|\partial_t^k g\|_{C([-1,1];L^\infty(\Omega))}. \end{aligned} \tag{2.17}$$

It is known that the Lagrange interpolation operator $I_m : C([-1,1]) \to P_m([-1,1])$ is not stable with respect to $m$, i.e.,

$$\|I_m g\|_{C([-1,1])} \leqslant C \log(m) \|g\|_{C([-1,1])}, \tag{2.18}$$

where the logarithmic factor in $m$ cannot be removed. This significantly affects the stability of the spectral methods based on the Lagrange interpolation for nonlinear problem, for which the error estimates require using Gronwall's inequality. We shall use a different interpolation technique in our spectral method, called the polynomial VP means, which turns out to be not only stable in $C([-1,1]; L^\infty(\Omega))$ but also convenient for the practical computations in combination with exponential integrators.

Let $t_{m,1} < t_{m,2} < \cdots < t_{m,m}$ be the $m$ zeros of the normalized Jacobi polynomial $J_m^{\alpha,\beta}(t)$ of degree $m$ on the interval $[-1,1]$, with the Jacobi weight $\omega^{\alpha,\beta} = (1-t)^\alpha (1+t)^\beta$. The following polynomial VP means of a function $g \in C([-1,1])$ will be used:

$$V_m^r g(t) = \sum_{i=1}^{m} g(t_{m,i}) \Phi_{m,i}^r(t) \quad \text{for } g \in C([-1,1]), \tag{2.19}$$

where

$$\Phi_{m,i}^r(t) = \frac{\sum_{j=0}^{m+r-1} \mu_{m,j}^r J_j^{\alpha,\beta}(t_{m,i}) J_j^{\alpha,\beta}(t)}{\sum_{j=0}^{m-1} [J_j^{\alpha,\beta}(t_{m,i})]^2}, \ \ j = 1, \dots m, \tag{2.20}$$

with

$$\mu_{m,j}^r := \begin{cases} 1 & \text{if} \quad 0 \leqslant j \leqslant m-r, \\ \frac{m+r-j}{2r} & \text{if} \quad m-r < j < m+r, \\ 0 & \text{if} \quad j \geqslant m+r. \end{cases} \tag{2.21}$$

For $|\alpha| = |\beta| = \frac{1}{2}$, the VP means enjoy the interpolation properties, i.e., $V_m^r g(t_{m,i}) = g(t_{m,i})$ for $i = 1, \dots, m$.

In this chapter, we simply choose $r \leqslant \theta m$ for some fixed $\theta \in (0,1)$. Then the approximation errors of the polynomial VP means are given in the following lemma, which shows that the $L^\infty$-stability constant of the polynomial VP means is independent of the polynomial degree.

**Lemma 2.3** (165, Theorem 3.2]). *For sufficiently large positive integers $m$ and $r$ satisfying $r \leqslant \theta m$, the following estimates hold:*

$$\|V_m^r g\|_{C([-1,1];L^\infty(\Omega))} \leqslant C \sup_{1 \leqslant i \leqslant m} \|g(t_{m,i})\|_{L^\infty(\Omega)} \qquad \text{for } g \in C([-1,1]; L^\infty(\Omega)), \tag{2.22}$$

$$\|g - V_m^r g\|_{C([-1,1];L^\infty(\Omega))} \leqslant C \|E_{m-r}(g)\|_{L^\infty(\Omega)} \qquad \text{for } g \in C([-1,1]; L^\infty(\Omega)), \tag{2.23}$$

*for some constant $C$ which does not depend on $m$.*

**2.2.4 The exponential integrator using VP means.** The only unknown on the right-hand side of (2.13) is $y(\tilde{z}_j, t)$, which we shall approximate by the exponential integrator

$$y(\tilde{z}_j, t_{ni}) = e^{\tilde{z}_j(t_{ni}-t_{n-1})} y(\tilde{z}_j, t_{n-1}) + \int_{t_{n-1}}^{t_{ni}} e^{\tilde{z}_j(t_{ni}-s)} V_m^r f(u(s), \cdot, s) \mathrm{d}s \tag{2.24}$$

at the finitely many internal nodes $t_{ni} \in (t_{n-1}, t_n]$, $i = 1, \ldots, m$, where the polynomial VP mean $V_m^r f(u(s), \cdot, s)$ only depends on the finitely many values $f(u(t_{ni}), \cdot, t_{ni})$, $i = 1, \ldots, m$.

Since the VP means can be expressed as linear combinations of the Jacobi polynomials, the integral in (2.24) can be evaluated analytically by using the following formula:

$$\int_{-1}^{t} e^{z(t-s)} J_m^{\alpha,\beta}(s) \mathrm{d}s = \sum_{k=0}^{m} \frac{d_{m,k}^{\alpha,\beta}}{z^{k+1}} \sqrt{\frac{\gamma_{m-k}^{\alpha+k,\beta+k}}{\gamma_m^{\alpha,\beta}}} (-J_{m-k}^{\alpha+k,\beta+k}(t) + e^{z(t+1)} J_{m-k}^{\alpha+k,\beta+k}(-1)). \tag{2.25}$$

This formula can be derived by using integration by parts and the following differentiation property of the Jacobi polynomials ([159, (3.101)]):

$$\partial_t^k J_m^{\alpha,\beta}(t) = d_{m,k}^{\alpha,\beta} \sqrt{\frac{\gamma_{m-k}^{\alpha+k,\beta+k}}{\gamma_m^{\alpha,\beta}}} J_{m-k}^{\alpha+k,\beta+k}(t),$$

where

$$d_{m,k}^{\alpha,\beta} = \frac{\Gamma(m+k+\alpha+\beta+1)}{2^k \Gamma(m+\alpha+\beta+1)},$$

$$\gamma_m^{\alpha,\beta} = \frac{2^{\alpha+\beta+1}\Gamma(m+\alpha+1)\Gamma(m+\beta+1)}{(2m+\alpha+\beta+1)m!\Gamma(m+\alpha+\beta+1)}.$$

By a simple scaling transformation, we can use formula (2.25) to compute the integral in (2.24) in terms of the values $f(u(t_{ni}), \cdot, t_{ni})$, $i = 1, \ldots, m$. Then, by substituting the computed values $y(\tilde{z}_j, t_{ni})$ into (2.13), we can obtain the desired spectral method. The complete algorithms are presented in the next two sections for the linear and semilinear subdiffusion equations, respectively.

**2.3 The linear problem with a singular source.** In this section, we present the spectral algorithm for the linear subdiffusion equation with a given source function $f(x, t)$ which may be singular at $t = 0$ (but is independent of $u$).

Since the solution representation in (2.13)–(2.14) does not depend on the history values of the solution, i.e., it only depends on the value of $y(z, t)$ which satisfies the ODE problem

in (2.14), which does not contain history integrals, the case $T > 1$ can be reduced to the case $T = 1$ by dividing the time interval $[0, T]$ into several parts. The case $T < 1$ can be converted to the case $T = 1$ via a temporal scaling transformation. Therefore, without loss of generality, we focus on the case $T = 1$ and consider the subdiffusion equation on the unit time interval $[0, 1]$.

If the source function is smooth, i.e., $f \in C^\infty([0,1]; L^\infty(\Omega))$, then we can approximate the source function $f$ by its VP mean $V_m^r f$ on the whole interval $[0, 1]$. In particular, let $t_{m,1} < t_{m,2} < \cdots < t_{m,m}$ be the zeros of the Jacobi polynomial $J_m^{\alpha,\beta}(t)$ of degree $m$ on the time interval $[0, 1]$. For any fixed $t \in [0, 1]$ we can approximate $y(z, t)$ by

$$y_m(z,t) = \int_0^t e^{z(t-s)} V_m^r f(s) \mathrm{d}s, \tag{2.26}$$

which can be evaluated exactly by using formula (2.25). By substituting (2.26) into (2.13) and dropping the two remainders $\mathcal{E}_{1,q}(t)$ and $\mathcal{E}_{2,q}(t)$, we obtain the following algorithm for any $t \in [0, 1]$:

$$u_m(t) = \sum_{j=-M}^{M} \omega_j e^{z_j} z_j^{\alpha-1} (z_j^\alpha - t^\alpha \Delta)^{-1} u_0 + \sum_{j=-M}^{M} \tilde{\omega}_j (\tilde{z}_j^\alpha - \Delta)^{-1} y_m(\tilde{z}_j, t). \tag{2.27}$$

The error bound of this method is an immediate consequence of Lemma 2.1, Lemma 2.2 and Lemma 2.3 (with $r \leqslant \theta m$ for some fixed parameter $0 < \theta < 1$). We present the result in the following theorem and omit the proof.

**Theorem 2.1.** *Let $u_0 \in L^\infty(\Omega)$ and $f \in C^\infty([0,1]; L^\infty(\Omega))$. Then the numerical solution defined in (2.26)–(2.27) has the following error bound for the linear subdiffusion equation:*

$$\max_{t\in[0,1]} \|u(t) - u_m(t)\|_{L^\infty(\Omega)} \leqslant C_k m^{-k} + C e^{-\sqrt{M}/C}, \tag{2.28}$$

*which holds for all fixed integer $k \geqslant 1$, all $M$ and all sufficiently large $m$.*

We are more interested in the development of high-order methods for the subdiffusion equation with a source function singular at $t = 0$. This is often the case when the source function is related to the solution of a subdiffusion equation. The strength of such a singularity at $t = 0$ can be characterized by the following condition with a parameter $\gamma \in (0, 1]$:

$$\|(t^\gamma \partial_t)^k f(\cdot, t)\|_{L^\infty(\Omega)} \leqslant C_k \ \text{ for } k \geqslant 0 \text{ and } t \in (0, 1]. \tag{2.29}$$

The larger value of $\gamma$, the stronger of the singularity.

For a source function which exhibits a singularity at $t = 0$ in the form of (2.29), we choose a fixed parameter $\lambda > 1$ and divide the time interval $[0, 1]$ into $N$ subintervals, i.e.,

$$I_1 = [t_0, t_1] := [0, \lambda^{1-N}] \quad \text{and} \quad I_n = (t_{n-1}, t_n] := (\lambda^{n-1-N}, \lambda^{n-N}], \quad 2 \leqslant n \leqslant N,$$

where $N = N_m$ can be any integer satisfying $\lim_{m\to\infty} N_m/\log m = \infty$ (this requirement will become clear in the error analysis). On each subinterval $I_n$, we approximate the source function $f(\cdot, t)$ by its VP means $V_m^r f(\cdot, t)$ and substitute the value

$$y_n(z) = e^{z(t_n - t_{n-1})} y_{n-1}(z) + \int_{t_{n-1}}^{t_n} e^{z(t_n - s)} V_m^r f(s) \mathrm{d}s \tag{2.30}$$

into (2.13). Then, by dropping the two remainders $\mathcal{E}_{1,q}(t)$ and $\mathcal{E}_{2,q}(t)$ in (2.13), we obtain the following algorithm:

$$u_n = \sum_{j=-M}^{M} \omega_j e^{z_j} z_j^{\alpha-1} (z_j^\alpha - t_n^\alpha \Delta)^{-1} u_0 + \sum_{j=-M}^{M} \tilde{\omega}_j (\tilde{z}_j^\alpha - \Delta)^{-1} y_n(\tilde{z}_j). \tag{2.31}$$

The algorithm in (2.30)–(2.31) only requires polynomial interpolation based on the VP means, the evaluation of the exponential integrals in (2.30), and the solution of the linear systems associated to the operators $z_j^\alpha - \Delta$ and $\tilde{z}_j^\alpha - \Delta$. This is different from the spectral method for the semilinear problem to be presented in the next section, which requires solving certain nonlinear systems by fixed-point iterations or the Newton iterations.

For the linear problem with singular source function, the accuracy of the numerical approximation by (2.31) is guaranteed by the following theorem.

**Theorem 2.2.** *Let $u_0 \in L^\infty(\Omega)$ and assume that the source function $f(x, t)$ satisfies the regularity condition in* (2.29)*. Let $u_n$, $n = 1, \ldots, N$ be the numerical solutions given by* (2.30)–(2.31) *with $N = N_m$ satisfying*

$$\lim_{m\to\infty} \frac{N_m}{\log(m)} = \infty. \tag{2.32}$$

*Then the following error bound holds (for all integer $k \geqslant 1$, all $M$ and all sufficiently large $m$.):*

$$\max_{1\leqslant n\leqslant N} \|u(t_n) - u_n\|_{L^\infty(\Omega)} \leqslant C_k m^{-k} + C e^{-\sqrt{M}/C}. \tag{2.33}$$

**Remark 2.1.** The total number of degrees of freedom in the time discretization is $mN$ with $N = N_m$ satisfying condition (2.32). By choosing a moderate growing $N_m$, such as

$N_m = m/2$, the total number of degrees of freedom is $O(m^2)$ while the error of the numerical approximation is $O(m^{-k})$ for arbitrarily large $k$. This means that the proposed method has spectral convergence (i.e., arbitrarily large convergence orders) with respect to the total number of degrees of freedom.

**Remark 2.2.** Theorems 2.1 and 2.2 are still valid if the $L^\infty$ norms are changed to $L^2$ norms. Namely, if $u_0 \in L^2(\Omega)$ and $\|(t^\gamma \partial_t)^k f(\cdot,t)\|_{L^2(\Omega)} \leqslant C_k$ for $k \geqslant 0$ and $t \in (0,1]$, then the numerical solution defined in (2.30)–(2.31) has the following error bound for the linear subdiffusion equation:

$$\max_{1\leqslant n\leqslant N} \|u(t_n) - u_n\|_{L^2(\Omega)} \leqslant C_k m^{-k} + Ce^{-\sqrt{M}/C},$$

which holds for all fixed integer $k \geqslant 1$ and sufficiently large $m$ and $M$. The proof of this result is the same as the proof of Theorem 2.2 below by changing all the $L^\infty$ norms to $L^2$ norms.

*Proof.* We define an auxiliary function $u^*$, which is the solution of the following linear subdiffusion problem (with the source function replaced by $V_m^r f$):

$$\begin{cases} \partial_t^\alpha u^* - \Delta u^* = V_m^r f & \text{in } \Omega \times (0,1], \\ u^* = 0 & \text{on } \partial\Omega \times (0,1], \\ u^*(0) = u_0 & \text{in } \Omega. \end{cases} \tag{2.34}$$

Then the error $e_n = u(t_n) - u_n$ can be decomposed into $e_n = \tilde{e}_n + e_n^*$, with

$$\tilde{e}_n = u(t_n) - u^*(t_n) \quad \text{and} \quad e_n^* = u^*(t_n) - u_n.$$

Since $u_n$ is the quadrature approximation of the contour integral representation of $u^*(t_n)$, i.e., the value after dropping the two remainders in (2.13), the estimates in Lemma 2.1 and the $L^\infty$-stability of the VP means imply that $\|V_m^r f\|_{L^\infty(0,1;L^\infty(\Omega))} \leqslant C\|f\|_{L^\infty(0,1;L^\infty(\Omega))} \leqslant C$ and therefore

$$\|e_n^*\|_{L^\infty(\Omega)} \leqslant Ce^{-\sqrt{M}/C}. \tag{2.35}$$

Since $\tilde{e}_n$ represents the error between the exact solution and the auxiliary function $u^*$ due to the change of the source function from $f$ to $V_m^r f$, by using formula (2.8) we can express $\tilde{e}_n$ as follows:

$$\begin{aligned} \tilde{e}_n &= \int_0^{t_1} E(t_n - s)[f(s) - V_m^r f(s)]\mathrm{d}s + \sum_{j=2}^{n} \int_{t_{j-1}}^{t_j} E(t_n - s)[f(s) - V_m^r f(s)]\mathrm{d}s \\ &=: \mathcal{E}_{n,1} + \mathcal{E}_{n,2}. \end{aligned} \tag{2.36}$$

The first term on the right-hand side of (2.36) can be estimated by using the boundedness of $f$ and $V_m^r f$, and the estimate $\|E(t-s)\|_{L^\infty(\Omega)\to L^\infty(\Omega)} \leqslant C(t-s)^{\alpha-1}$ as shown in (2.11), i.e.,

$$\|\mathcal{E}_{n,1}\|_{L^\infty(\Omega)} \leqslant C\int_0^{t_1}(t_1-s)^{\alpha-1}\|f\|_{L^\infty(I_1;L^\infty(\Omega))}\mathrm{d}s \leqslant Ct_1^\alpha\|f\|_{L^\infty([0,1];L^\infty(\Omega))}.$$

Since $t_1 = \lambda^{1-N}$ and $N = N_m$ satisfies condition (2.32), it follows that $t_1^\alpha \leqslant m^{-k}$ as $m \to \infty$ for any $k \geqslant 1$. This proves that

$$\|\mathcal{E}_{n,1}\|_{L^\infty(\Omega)} \leqslant C_k m^{-k} \quad \text{as } m \to \infty. \tag{2.37}$$

The second term on the right-hand side of (2.36) can be estimated by using the following approximation error estimate of the VP means, i.e.,

$$\begin{aligned}\|f - V_m^r f\|_{C([t_{j-1},t_j];L^\infty(\Omega))} &\leqslant C_k(m-r)^{-k}\Big(\frac{t_j-t_{j-1}}{2}\Big)^k\|\partial_t^k f\|_{C([t_{j-1},t_j];L^\infty(\Omega))}\\ &\leqslant C_k m^{-k}\Big(\frac{t_j-t_{j-1}}{2}\Big)^k\|\partial_t^k f\|_{C([t_{j-1},t_j];L^\infty(\Omega))},\end{aligned} \tag{2.38}$$

which is an immediate consequence of Lemma 2.2 and Lemma 2.3, and a scaling transformation from the standard interval $[-1,1]$ to the current interval $[t_{j-1},t_j]$. The last inequality is due to the fact that $r \leqslant \theta m$ for some fixed $\theta \in (0,1)$. Therefore,

$$\begin{aligned}\|\mathcal{E}_{n,2}\|_{L^\infty(\Omega)} &\leqslant \sum_{j=2}^n\int_{t_{j-1}}^{t_j}(t_n-s)^{\alpha-1}\|f - V_m^r f\|_{L^\infty(I_j;L^\infty(\Omega))}\mathrm{d}s\\ &\leqslant C_k\sum_{j=2}^n m^{-k}\int_{t_{j-1}}^{t_j}(t_n-s)^{\alpha-1}\Big(\frac{t_j-t_{j-1}}{2}\Big)^k\|\partial_s^k f\|_{L^\infty(I_j;L^\infty(\Omega))}\mathrm{d}s.\end{aligned} \tag{2.39}$$

The condition in (2.29) implies that $\|\partial_s^k f\|_{L^\infty(I_j;L^\infty(\Omega))} \leqslant Ct_{j-1}^{-\gamma k}$ and therefore

$$\begin{aligned}\|\mathcal{E}_{n,2}\|_{L^\infty(\Omega)} &\leqslant C_k\sum_{j=2}^n m^{-k}\Big(\frac{t_j-t_{j-1}}{2t_{j-1}^\gamma}\Big)^k\int_{t_{j-1}}^{t_j}(t_n-s)^{\alpha-1}\mathrm{d}s\\ &\leqslant C_k m^{-k}\sum_{j=2}^n\lambda^{[(1-\gamma)+\alpha](j-N)},\end{aligned}$$

where have substituted $t_j = \lambda^{j-N}$ into the last inequality. Since $\lambda > 1$ and $n \leqslant N$, the summation of $\lambda^{[(1-\gamma)+\alpha](j-N)}$ for $j = 2,\ldots,n$ is finite and independent of $m$. This proves that

$$\|\mathcal{E}_{n,2}\|_{L^\infty(\Omega)} \leqslant C_k m^{-k}. \tag{2.40}$$

Then, substituting the estimates of $\|\mathcal{E}_{n,1}\|_{L^\infty(\Omega)}$ and $\|\mathcal{E}_{n,2}\|_{L^\infty(\Omega)}$ into (2.36), we obtain

$$\|\tilde{e}_n\|_{L^\infty(\Omega)} \leqslant C_k m^{-k} \quad \text{for} \quad 1 \leqslant n \leqslant N. \tag{2.41}$$

The error bound in Theorem 2.2 follows from the two estimates in (2.35) and (2.41). □

**2.4 The semilinear problem with rough initial data.** In this section, we present the spectral method for the semilinear subdiffusion equation with a rough initial value $u_0 \in L^\infty(\Omega)$. Similarly to the linear problem (as explained at the beginning of Section 2.3), we can focus on the case $T = 1$ and consider the semilinear subdiffusion equation on the unit time interval $[0, 1]$. For the simplicity of presentation, we focus on the case $f(u, x, t) = f(u)$ without loss of generality.

**2.4.1 The spectral collocation method.** Differently from the linear problem, we first divide the whole interval $[0, 1]$ uniformly into $N$ smaller subintervals $I_n = ((n-1)/N, n/N] = (t_{n-1}, t_n]$, $1 \leqslant n \leqslant N$, and then refine the first subinterval $I_1 = [0, t_1]$ into $N_1$ smaller subintervals. In particular, for a constant $\lambda > 1$ we define

$$\begin{aligned} I_{1,1} &= [t_0, t_{1,1}] := [0, \lambda^{1-N_1}/N], \\ I_{1,j} &= (t_{1,j-1}, t_{1,j}] := (\lambda^{j-1-N_1}/N, \lambda^{j-N_1}/N] \quad \text{for } j = 2, \ldots, N_1. \end{aligned}$$

The division of the whole interval $[0, 1]$ uniformly into $N$ smaller subintervals $I_n = ((n-1)/N, n/N]$, $n = 1, \ldots, N$, is to guarantee the stability of the spectral method on each subinterval with respect to the polynomial interpolation of the nonlinear source function. This will become clear in the error analysis, i.e., the $L^\infty$-norm stability with respect to the polynomial interpolation requires the length of the interval to be sufficiently small. The division of the first subinterval $I_1$ into $N_1$ smaller subintervals $I_{1,j}$, $j = 1, \ldots, N_1$, with graded stepsizes locally refined towards $t = 0$ is to resolve the singularity of the nonlinear source function at $t = 0$ (similarly to the linear problem).

On each subinterval $I_{1,n}$, we approximate $f(u)$ by its VP means $V_m^r f(u_m)$, where $u_m$ denotes the numerical solution obtained by using polynomial VP means of degree $m$ on each subinterval. In particular, we denote by $t_{1,n}^{m,1} < t_{1,n}^{m,2} < \cdots < t_{1,n}^{m,m}$ the $m$ zeros of the Jacobi polynomial $J_m^{\alpha,\beta}(t)$ of degree $m$ on the interval $I_{1,n} = (t_{1,n-1}, t_{1,n}]$, $1 \leqslant n \leqslant N_1$, and denote by

$$U_{1,n} = (u_{1,n}^{m,1}, u_{1,n}^{m,2}, \ldots, u_{1,n}^{m,m})^\top$$

the vector which contains the numerical solutions at the discrete time levels $t_{1,n}^{m,i}$, $i = 1, \ldots, m$, determined by the following nonlinear system of equations:

$$U_{1,n} = G_{1,n}(U_{1,n}), \tag{2.42}$$

where $G_{1,n} : L^\infty(\Omega)^m \to L^\infty(\Omega)^m$ is a nonlinear mapping defined by

$$G_{1,n}(U_{1,n}) := (v_{1,n}^{m,1}, v_{1,n}^{m,2}, \ldots, v_{1,n}^{m,m})^\top,$$

with

$$\begin{aligned} v_{1,n}^{m,i} = & \sum_{j=-M}^{M} \omega_j e^{z_j} z_j^{\alpha-1} (z_j^\alpha - (t_{1,n}^{m,i})^\alpha \Delta)^{-1} u_0 \\ & + \sum_{j=-M}^{M} \tilde{\omega}_j (\tilde{z}_j^\alpha - \Delta)^{-1} y_{1,n}^{m,i}(\tilde{z}_j) \quad \text{for } i = 1, \ldots, m, \end{aligned} \tag{2.43}$$

and

$$y_{1,n}^{m,i}(\tilde{z}_j) = e^{\tilde{z}_j (t_{1,n}^{m,i} - t_{1,n-1})} y_{1,n-1}(\tilde{z}_j) + \int_{t_{1,n-1}}^{t_{1,n}^{m,i}} e^{\tilde{z}_j (t_{1,n}^{m,i} - s)} [V_m^r f(u_m)](s) \mathrm{d}s. \tag{2.44}$$

If $y_{1,n-1}(\tilde{z}_j)$ is known then we can compute $u_{1,n}^{m,i}$, $i = 1, \ldots, m$ by solving the collocation system (2.42) and then compute $y_{1,n}(z)$ by

$$y_{1,n}(\tilde{z}_j) = e^{\tilde{z}_j (t_{1,n} - t_{1,n-1})} y_{1,n-1}(\tilde{z}_j) + \int_{t_{1,n-1}}^{t_{1,n}} e^{\tilde{z}_j (t_{1,n} - s)} [V_m^r f(u_m)](s) \mathrm{d}s. \tag{2.45}$$

Similarly, for each subinterval $I_n$, $2 \leqslant n \leqslant N$, we denote by $t_n^{m,1} < \cdots < t_n^{m,m}$ the $m$ zeros of Jacobi polynomial $J_m^{\alpha,\beta}(x)$ on $I_n$, and denote by

$$U_n = (u_n^{m,1}, u_n^{m,2}, \ldots, u_n^{m,m})^\top,$$

the vector which contains the numerical solutions at time levels $t_n^{m,i}$, $i = 1, \ldots, m$, determined by the following nonlinear system of equations:

$$U_n = G_n(U_n), \tag{2.46}$$

where $G_n(U_n) := (v_n^{m,1}, v_n^{m,2}, \ldots, v_n^{m,m})^\top$, with

$$\begin{aligned} v_n^{m,i} = & \sum_{j=-M}^{M} \omega_j e^{z_j} z_j^{\alpha-1} (z_j^\alpha - (t_n^{m,i})^\alpha \Delta)^{-1} u_0 \\ & + \sum_{j=-M}^{M} \tilde{\omega}_j (\tilde{z}_j^\alpha - \Delta)^{-1} y_n^{m,i}(\tilde{z}_j) \quad \text{for } i = 1, \ldots, m, \end{aligned} \tag{2.47}$$

and

$$y_n^{m,i}(\tilde{z}_j) = e^{\tilde{z}_j(t_n^{m,i}-t_{n-1})}y_{n-1}(\tilde{z}_j) + \int_{t_{n-1}}^{t_n^{m,i}} e^{\tilde{z}_j(t_n^{m,i}-s)}[V_m^r f(u_m)](s)\mathrm{d}s. \tag{2.48}$$

If $y_{n-1}(\tilde{z}_j)$ is known then we can compute $u_n^{m,i}$, $i = 1, \ldots, m$ by solving collocation system (2.46) and then compute $y_n(z)$ by

$$y_n(\tilde{z}_j) = e^{\tilde{z}_j(t_n-t_{n-1})}y_{1,n-1}(\tilde{z}_j) + \int_{t_{n-1}}^{t_n} e^{\tilde{z}_j(t_n-s)}[V_m^r f(u_m)](s)\mathrm{d}s. \tag{2.49}$$

The existence, uniqueness and spectral convergence of the numerical solutions defined by (2.42) and (2.46) are guaranteed by the following theorem.

**Theorem 2.3.** *Let $u \in C([0,1]; L^2(\Omega)) \cap L^\infty(0,1; L^\infty(\Omega))$ be a bounded mild solution of (2.1) with initial value $u_0 \in L^\infty(\Omega)$ and nonlinear source function $f \in C^\infty(\mathbb{R})$. Let $u_{1,n}^{m,i}$ and $u_n^{m,i}$ be the numerical solutions given by (2.43) and (2.47), respectively. Then there exists a positive constant $N_*$ such that for $N \geqslant N_*$ and $N_1 = N_1(m)$ satisfying*

$$\lim_{m\to\infty} \frac{N_1(m)}{\log(m)} = \infty, \tag{2.50}$$

*the nonlinear systems (2.42) and (2.46) have unique solutions in an $L^\infty$-neighborhood of the mild solution, with the following error bounds:*

$$\max_{1\leqslant n\leqslant N_1} \max_{1\leqslant i\leqslant m} \|u(t_{1,n}^{m,i}) - u_{1,n}^{m,i}\|_{L^\infty(\Omega)} \leqslant C_k(m^{-k} + e^{-\sqrt{M}/C}), \tag{2.51}$$

$$\max_{2\leqslant n\leqslant N} \max_{1\leqslant i\leqslant m} \|u(t_n^{m,i}) - u_n^{m,i}\|_{L^\infty(\Omega)} \leqslant C_k(m^{-k} + e^{-\sqrt{M}/C}), \tag{2.52}$$

*which hold for all integer $k \geqslant 1$ and sufficiently large $m$ and $M$ (larger than some constants which are independent of $m$).*

**Remark 2.3.** Since all the results are based on the properties of the resolvent operator $(z - \Delta)^{-1}$, which has similar properties under the Dirichlet and periodic boundary conditions, the results in this chapter can be extended to the periodic bounadry condition.

**Remark 2.4.** It is mentioned at the beginning of Section 3 that the case $T > 1$ can be solved by dividing the time interval into a number of subintervals. Using the same method together with discrete Gronwall's inequality, the error bound will be multiplied by $e^{CT}$ for long-term computation. The factor $e^{CT}$ usually appears in the error estimates for semilinear parabolic equations and subdiffusion equations. The factor $e^{CT}$ may be removed

for the linear subdiffusion equation but requires a closer look at the error analysis by taking account of the regularity behaviour of the mild solution as $t \to 0$. This is not studied in the current thesis.

**Remark 2.5.** For the convenience of illustration, we have focused on the semilinear subdiffusion equation with a Laplacian operator in space. However, the results can be extended to more general elliptic partial differential operators which satisfy the resolved estimate in (2.10), such as second-order elliptic partial differential operators with variable coefficients.

*Proof.* The roughness of the initial value will generate singularity in the solution and the nonlinear source function at $t = 0$. For a bounded mild solution of the semilinear subdiffusion equation with initial value $u_0 \in L^\infty(\Omega)$, it is shown in [106, inequality (3.8)] that both the solution and the source function exhibit singularities in the form of (2.6). In the next two subsections, we present the proof of Theorem 2.3 based on the regularity estimates in (2.6). For the simplicity of notations, we omit the dependence of the constants $C$ on $k$ in the proof.

**2.4.2 Existence, uniqueness and boundedness of the numerical solution.** We denote $L = \|u\|_{L^\infty(0,1;L^\infty(\Omega))}$ and modify the definition of the nonlinear function $f : \mathbb{R} \to \mathbb{R}$ in the region $|\sigma| \geqslant L+1$ so that $f(\sigma) = 0$ for $|\sigma| \geqslant L+2$. For the simplicity of the notation, we still use $f$ to denote the modified function. This modification make the source function $f : \mathbb{R} \to \mathbb{R}$ globally bounded and Lipschitz continuous, but would not have influence on the numerical solution if it satisfies the following condition:

$$\|u_{1,n}^{m,i}\|_{L^\infty(\Omega)} \leqslant L + 1 \quad \text{and} \quad \|u_n^{m,i}\|_{L^\infty(\Omega)} \leqslant L + 1. \tag{2.53}$$

We shall prove the existence and uniqueness of a numerical solution satisfying this condition (for sufficiently large $N$ and $N_1$).

The existence and uniqueness of numerical solutions would follow from the contractivity of the map $G_{1,n} : L^\infty(\Omega)^m \to L^\infty(\Omega)^m$. For $V = (v_1, \ldots, v_m) \in L^\infty(\Omega)^m$, we denote $f(V) = (f(v_1), \ldots, f(v_m))$ and $V_m^r f(V)$ the polynomial on the interval $I_{1,n}$ based on the VP mean of the nodal values $f(v_1), \ldots, f(v_m)$. Then

$$\begin{aligned}
&\|G_{1,n}(U) - G_{1,n}(V)\|_{L^\infty(\Omega)^m} \\
&\leqslant \Big\| \sum_{j=-M}^{M} \tilde{\omega}_j (\tilde{z}_j^\alpha - \Delta)^{-1} \int_{t_{1,n-1}}^{t} e^{\tilde{z}_j(t-s)} V_m^r [f(U) - f(V)] \mathrm{d}s \Big\|_{L^\infty(I_{1,n};L^\infty(\Omega))} \\
&\leqslant \Big\| \sum_{j=-M}^{M} \tilde{\omega}_j (\tilde{z}_j^\alpha - \Delta)^{-1} \int_{t_{1,n-1}}^{t} e^{\tilde{z}_j(t-s)} V_m^r [f(U) - f(V)] \mathrm{d}s
\end{aligned}$$

$$
\begin{aligned}
&- \int_{\Gamma_{\tilde{\lambda}}} (z^\alpha - \Delta)^{-1} \int_{t_{1,n-1}}^{t} e^{z(t-s)} V_m^r[f(U) - f(V)] \mathrm{d}s\mathrm{d}z \Big\|_{L^\infty(I_{1,n};L^\infty(\Omega))} \\
&+ \Big\| \int_{\Gamma_{\tilde{\lambda}}} (z^\alpha - \Delta)^{-1} \int_{t_{1,n-1}}^{t} e^{z(t-s)} V_m^r[f(U) - f(V)] \mathrm{d}s\mathrm{d}z \Big\|_{L^\infty(I_{1,n};L^\infty(\Omega))} \\
&=: \mathcal{F}_1 + \mathcal{F}_2.
\end{aligned}
$$

$\mathcal{F}_1$ can be estimated by using the $L^\infty$-stability of the VP means and the Lipschitz continuity of the modified function $f$, i.e., $\|V_m^r[f(U) - f(V)]\|_{L^\infty(I_{1,n};L^\infty(\Omega))} \leqslant C\|U - V\|_{L^\infty(\Omega)^m}$, and Lemma 2.1, which together imply that

$$
\mathcal{F}_1 \leqslant Ce^{-\sqrt{M}/C} \|U - V\|_{L^\infty(\Omega)^m}. \tag{2.54}
$$

$\mathcal{F}_2$ can be converted into the following form:

$$
\begin{aligned}
\mathcal{F}_2 &= \Big\| \int_{t_{n-1}}^{t} E(t-s) V_m^r[f(U) - f(V)] \mathrm{d}s \Big\|_{L^\infty(I_{1,n};L^\infty(\Omega))} \\
&\leqslant C\|U - V\|_{L^\infty(\Omega)^m} \int_{t_{n-1}}^{t_n} (t_n - s)^{\alpha-1} \mathrm{d}s \\
&= C|I_{1,n}|^\alpha \|U - V\|_{L^\infty(\Omega)^m}.
\end{aligned} \tag{2.55}
$$

Since $|I_{1,n}| \leqslant 1/N$, it follows that

$$
\|G_{1,n}(U) - G_{1,n}(V)\|_{L^\infty(\Omega)^m} \leqslant C(e^{-\sqrt{M}/C} + N^{-\alpha}) \|U - V\|_{L^\infty(\Omega)^m}. \tag{2.56}
$$

For sufficiently large $M$ and $N$ (bigger than some constants), (2.56) implies that $G_{1,n} : L^\infty(\Omega)^m \to L^\infty(\Omega)^m$ is a contraction and therefore has a unique fixed point, i.e., a numerical solution of (2.42). The existence and uniqueness of a fixed point for the map $G_n : L^\infty(\Omega)^m \to L^\infty(\Omega)^m$, i.e., the existence and uniqueness of a numerical solution of (2.46), can be proved in the same way and therefore omitted.

In the next subsection, we prove that for sufficiently large $m$, $M$ and $N$, the numerical solutions $U_{1,n}$ and $U_n$ with the modified source function actually satisfies condition (2.53) and therefore are also the numerical solutions with the original source function. This would prove the existence and uniqueness of numerical solutions with the original source function.

**2.4.3 Error estimation.** We define an auxiliary function $u^*$, which is the solution of the following subdiffusion problem (with the modified source function):

$$
\begin{cases}
\partial_t^\alpha u^* - \Delta u^* = V_m^r f(u_m) & \text{in } \Omega \times (0,1], \\
u^* = 0 & \text{on } \partial\Omega \times (0,1], \\
u^*(0) = u_0 & \text{in } \Omega.
\end{cases} \tag{2.57}
$$

We also know that the exact solution $u$ satisfies the following equation:

$$\begin{cases} \partial_t^\alpha u - \Delta u = V_m^r f(u) + \mathcal{E}_f & \text{in } \Omega \times (0,1], \\ u = 0 & \text{on } \partial\Omega \times (0,1], \\ u(0) = u_0 & \text{in } \Omega, \end{cases} \tag{2.58}$$

where the remainder $\mathcal{E}_f = f(u) - V_m^r f(u)$ satisfies the following estimate for any subinterval $[a,b] \subset (0,1]$:

$$\|\mathcal{E}_f\|_{C([a,b];L^\infty(\Omega))} \leqslant Cm^{-k}\Big(\frac{b-a}{2}\Big)^k \|\partial_t^k f(u)\|_{C([a,b];L^\infty(\Omega))}. \tag{2.59}$$

This is similar to the approximation error bound of VP means for the linear problem; see (2.38). By using expression (2.8) of the solution, the function $\tilde{e} = u - u^*$ can be written as

$$\tilde{e}(t) = \int_0^t E(t-s)(V_m^r f(u) - V_m^r f(u_m))\mathrm{d}s + \int_0^t E(t-s)\mathcal{E}_f(s)\mathrm{d}s. \tag{2.60}$$

Since the modified source function $f : \mathbb{R} \to \mathbb{R}$ is globally bounded and Lipschitz continuous, and the VP means are uniform bounded, it follows that

$$\|V_m^r f(u_m)\|_{L^\infty(I_{1,n};L^\infty(\Omega))} \leqslant C\|f(u_m)\|_{L^\infty(I_{1,n};L^\infty(\Omega))} \leqslant C.$$

Since the numerical solution $u_m$ is the quadrature approximation of the contour integral representation of $u^*$, it follows from Lemma 2.1 that

$$\|u^* - u_m\|_{L^\infty(I_{1,n};L^\infty(\Omega))} \leqslant Ce^{-\sqrt{M}/C} \text{ for } 1 \leqslant n \leqslant N_1. \tag{2.61}$$

For $\tilde{e}_{1,n}^{m,i} = u(t_{1,n}^{m,i}) - u^*(t_{1,n}^{m,i})$, with $1 \leqslant i \leqslant m$, $1 \leqslant n \leqslant N_1$, setting $t = t_{1,n}^{m,i}$ in (2.60) yields the following estimate:

$$\begin{aligned} \|\tilde{e}_{1,n}^{m,i}\|_{L^\infty(\Omega)} \leqslant{}& C\int_0^{t_{1,1}} (t_{1,n}^{m,i} - s)^{\alpha-1}\|f(u(s)) - f(u_m)\|_{L^\infty(I_{1,1};L^\infty(\Omega))}\mathrm{d}s \\ &+ C\sum_{j=2}^{n-1}\int_{t_{1,j-1}}^{t_{1,j}} (t_{1,n}^{m,i} - s)^{\alpha-1} \max_{1\leqslant i\leqslant m}\|e_{1,j}^{m,i}\|_{L^\infty(\Omega)}\mathrm{d}s \\ &+ C\int_{t_{1,n-1}}^{t_{1,n}^{m,i}} (t_{1,n}^{m,i} - s)^{\alpha-1} \max_{1\leqslant i\leqslant m}\|e_{1,n}^{m,i}\|_{L^\infty(\Omega)}\mathrm{d}s + Cm^{-k}, \end{aligned}$$

where we have used the result $\|\int_0^{t_{1,n}^{m,i}} E(t_{1,n}^{m,i} - s)\mathcal{E}_f(s)\mathrm{d}s\|_{L^\infty(\Omega)} \leqslant Cm^{-k}$, which follows from the same argument as the proof in (2.39)–(2.41) (with $\gamma = 1$ therein). Since $\|f(u(s)) - f(u_0)\|_{L^\infty(I_{1,1};L^\infty(\Omega))} \leqslant C$, from the inequality above we obtain

$$\max_{1\leqslant n\leqslant N_1}\max_{1\leqslant i\leqslant m}\|\tilde{e}_{1,n}^{m,i}\|_{L^\infty(\Omega)} \leqslant Ct_{1,1}^\alpha + Ct_{1,N}^\alpha \max_{1\leqslant n\leqslant N_1}\max_{1\leqslant i\leqslant m}\|e_{1,n}^{m,i}\|_{L^\infty(\Omega)} + Cm^{-k}$$

$$= C\lambda^{(1-N_1)\alpha} N^{-\alpha} + CN^{-\alpha} \max_{1\leqslant n\leqslant N_1} \max_{1\leqslant i\leqslant m} \|e_{1,n}^{m,i}\|_{L^\infty(\Omega)} + Cm^{-k}.$$

Similarly to the proof of (2.37), for sufficiently large $N_1$ satisfying condition (2.50), we have

$$\max_{1\leqslant n\leqslant N_1} \max_{1\leqslant i\leqslant m} \|\tilde{e}_{1,n}^{m,i}\|_{L^\infty(\Omega)} \leqslant CN^{-\alpha} \max_{1\leqslant n\leqslant N_1} \max_{1\leqslant i\leqslant m} \|e_{1,n}^{m,i}\|_{L^\infty(\Omega)} + Cm^{-k}.$$

Then, substituting (2.61) into the inequality above, we can convert $\tilde{e}_{1,j}^{m,i}$ to $e_{1,j}^{m,i}$ on the left-hand side, i.e.,

$$\begin{aligned}
&\max_{1\leqslant n\leqslant N_1} \max_{1\leqslant i\leqslant m} \|e_{1,n}^{m,i}\|_{L^\infty(\Omega)} \\
&\leqslant CN^{-\alpha} \max_{1\leqslant n\leqslant N_1} \max_{1\leqslant i\leqslant m} \|e_{1,n}^{m,i}\|_{L^\infty(\Omega)} + Cm^{-k} + Ce^{-\sqrt{M}/C}.
\end{aligned}$$

For sufficiently large $N$ (larger than some constant which is independent of $m$), the first term on the right-hand side above can be absorbed by the left-hand side. This yields the following error esetimate:

$$\max_{1\leqslant n\leqslant N_1} \max_{1\leqslant i\leqslant m} \|e_{1,n}^{m,i}\|_{L^\infty(\Omega)} \leqslant Cm^{-k} + Ce^{-\sqrt{M}/C}. \tag{2.62}$$

By choosing $m$ and $M$ large enough, we have

$$\max_{1\leqslant n\leqslant N_1} \max_{1\leqslant i\leqslant m} \|e_{1,n}^{m,i}\|_{L^\infty(\Omega)} \leqslant 1$$

and therefore

$$\max_{1\leqslant n\leqslant N_1} \|U_{1,n}\|_{L^\infty(\Omega)^m} \leqslant L + 1. \tag{2.63}$$

The same argument can be used to prove that (for sufficiently large $m$ and $M$)

$$\max_{2\leqslant n\leqslant N} \max_{1\leqslant i\leqslant m} \|e_n^{m,i}\|_{L^\infty(\Omega)} \leqslant Cm^{-k} + Ce^{-\sqrt{M}/C}, \tag{2.64}$$

$$\max_{2\leqslant n\leqslant N} \|U_n\|_{L^\infty(\Omega)^m} \leqslant L + 1. \tag{2.65}$$

This proves that the numerical solution satisfies the constraint in (2.53). Therefore, as we have discussed at the end of Section 2.4.2, $U_{1,n}$ and $U_n$ are the numerical solutions with the original source function $f$ (which is possibly not globally Lipschitz continuous), satisfying the error bounds in (2.62) and (2.64). This completes the proof of Theorem 2.3. □

**2.5 Numerical tests.** In this section, we present numerical tests to illustrate spectral convergence of the proposed time discretizations for both linear and semilinear subdiffusion equations with rough initial data. The piecewise linear Galerkin finite element method in space is used with a sufficiently small mesh size that does not affect the observation of the time discretization errors. All the computations are performed by MATLAB R2020b on a personal laptop.

We consider the subdiffusion equation in (2.1) in the domain $\Omega = (0,1)$ up to time $T = 1$, with a discontinuous initial value $u_0 = \chi_{[1/2,1)} \in L^\infty(\Omega)$, where $\chi_{[1/2,1)}$ denotes the characteristic function of the subinterval $[\frac{1}{2}, 1)$. The parameter $\lambda$ in the algorithm is chosen to be 2. The number of quadrature points are $2M + 1$ with $M = O(m \log^3 m)$, which satisfies the conditions in Theorems 2.1, 2.2 and 2.3. Since the closed form of the exact solution is not known, we compute a reference solution $u_{m_{\rm ref}}$ with $m_{\rm ref} = 24$, and compute the errors for $m \leqslant 16$.

The principle of choosing $M$ is to make $e^{-C\sqrt{M}}$ smaller than $m^{-k}$ for all $k$ as $m \to \infty$. Therefore, $M = O(m \log^2 m)$ and $M = O(m \log^3 m)$ both satisfy the requirement theoretically. In the numerical tests we observe that the choice of $M = O(m \log^2 m)$ performs well when $m$ is large, while the choice of $M = O(m \log^3 m)$ performs well for both large $m$ and small $m$. Therefore, we choose $M = O(m \log^3 m)$ in our numerical experiments below.

**2.5.1 The linear subdiffusion equation.** We solve the linear subdiffusion equation with a given source function $f(x,t)$ by the proposed algorithm in Section 2.3 with $N = \frac{m}{2}$ subintervals. Then the total number of degrees of freedom is $\frac{m^2}{2}$. The errors of the numerical solutions for the smooth source function

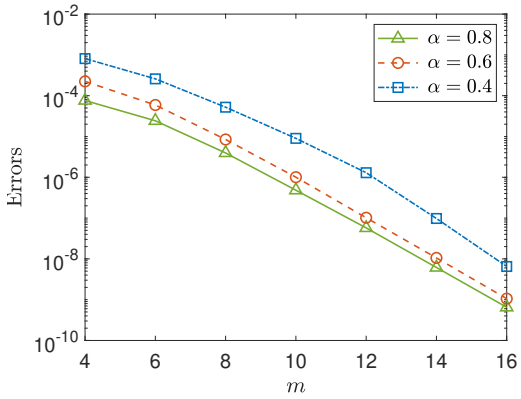
$$f(x,t) = (\sin t) \cos \pi x$$

and the singular source functions

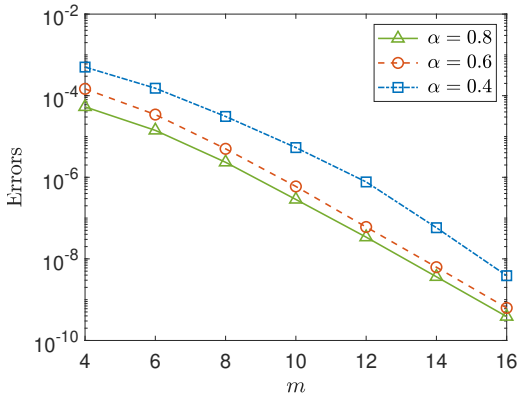
$$f(x,t) = t^\sigma \cos \pi x \quad \text{with } \sigma = 0.75,\ 0.5 \text{ and } 0.25,$$

are presented in Figure 2.1. In particular, the smooth and singular functions satisfy the conditions of Theorems 2.1 and 2.2, respectively.

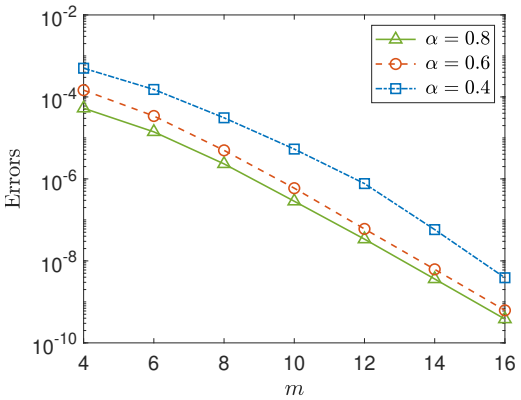
The numerical results in Figure 2.1 indicate that the proposed method has spectral convergence for the linear subdiffusion equation with both smooth and singular source functions. This is consistent with the theoretical results proved in Theorems 2.1 and 2.2.



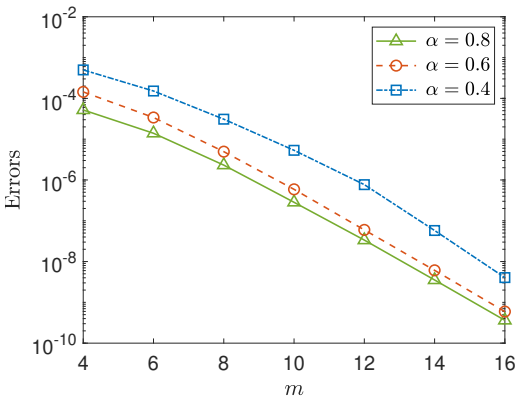

(a) $f(x,t) = \sin t \cos \pi x$




(b) $f(x,t) = t^{0.75} \cos \pi x$




(c) $f(x,t) = t^{0.5} \cos \pi x$




(d) $f(x,t) = t^{0.25} \cos \pi x$

*Figure 2.1. Errors of the numerical solutions for the linear subdiffusion equation*

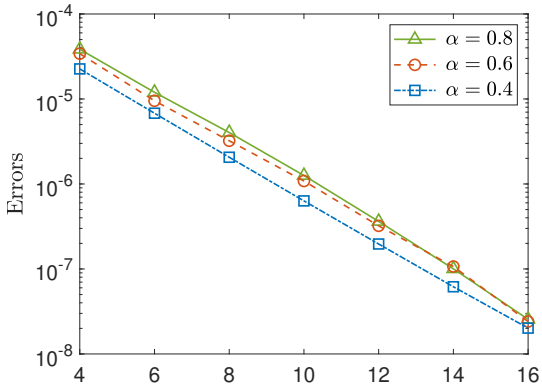



*Figure 2.2. Errors of the numerical solutions for the semilinear subdiffusion equation*

**2.5.2 The semilinear subdiffusion equation.** We consider the semilinear subdiffusion equation in (2.1) with the following nonlinear source function:

$$f(u) = \sin u,$$

which satisfies the condition of Theorem 2.3 and guarantees that the semilinear subdiffusion equation has a unique bounded mild solution. We divide the interval $[0, 1]$ into several subintervals with parameters $N = 1$, $N_1 = \frac{m}{2}$, and total number of degrees of freedom $\frac{m^2}{2}$, and present the errors of the numerical solutions in Figure 2.2 for several different values of $\alpha \in (0, 1)$. The numerical results in Figure 2.2 indicate that the proposed method has spectral convergence for the seimlinear subdiffusion equation with the discontinuous initial value $u_0 = \chi_{[1/2,1)} \in L^\infty(\Omega)$. This is consistent with the theoretical results proved in Theorem 2.3.

# Chapter 3

## Numerical Analysis of the 2D Navier–Stokes Equations with nonsmooth initial value in the Critical Space

The content of this chapter has been published in "*B. Li, Q. Rao, H. Zhang, and Z. Zhou. Numerical Analysis of the 2D Navier–Stokes Equations with nonsmooth initial value in the Critical Space, submitted.*"

**3.1 Introduction.** We denote by $\Omega$ a convex polygonal domain in $\mathbb{R}^2$ and consider the NS equations on $\Omega$ with the no-slip boundary condition up to a given time $T > 0$, i.e.,

$$\begin{cases} \partial_t u + u \cdot \nabla u - \Delta u + \nabla p = 0 & \text{in} \quad \Omega \times (0, T], \\ \nabla \cdot u = 0 & \text{in} \quad \Omega \times (0, T], \\ u = 0 & \text{on} \quad \partial\Omega \times (0, T], \\ u = u_0 & \text{on} \quad \Omega \times \{0\}, \end{cases} \tag{3.1}$$

where $\partial\Omega$ denotes the boundary of domain $\Omega$. In particular, we assume that the initial value $u_0$ belongs to $\dot{L}^2(\Omega)$, which is defined as

$$\dot{L}^2(\Omega) = \{v \in L^2(\Omega)^2 : \nabla \cdot v = 0 \text{ in } \Omega, v \cdot \nu = 0 \text{ on } \partial\Omega\}, \tag{3.2}$$

where $\nu$ denotes the unit outward normal vector on $\partial\Omega$. It is known that problem (3.1) has a unique weak solution $u \in L^2(0, T; \dot{H}_0^1(\Omega)) \cap H^1(0, T; \dot{H}^{-1}(\Omega)) \hookrightarrow C([0, T]; \dot{L}^2(\Omega))$, where $\dot{H}_0^1(\Omega) = \{v \in H_0^1(\Omega)^2 : \nabla \cdot v = 0\}$ and $\dot{H}^{-1}(\Omega)$ is the dual space of $\dot{H}_0^1(\Omega)$; see [164] for a rigorous proof of this result. The uniqueness of solution $p$ can be guaranteed by requiring $p \in L_0^2(\Omega) := \{v \in L^2(\Omega) : \int_\Omega v \, \mathrm{d}x = 0\}$.

The NS equations are the fundamental partial differential equations describing the motion of incompressible viscous fluids. They are widely used in fluid dynamics to model water and blood flows, air flow around a wing, and ocean currents. As exact solutions are unknown for most practical applications, numerical solutions of the NS equations are of

paramount importance. Error estimates can be obtained based on the regularity assumptions of the solution and the initial data. Optimal error estimates for high-order methods can be proved when the solutions to the Navier-Stokes equations are sufficiently regular, meaning they are sufficiently smooth and adhere to the compatibility conditions. For example, if the initial values are sufficiently smooth, i.e. $u_0 \in \dot{H}_0^1(\Omega) \cap H^2(\Omega)^2$ or above, then optimal-order convergence of temporal and spatial discretizations of the NS equations have been proved for various methods in [7, 10, 52, 70, 97, 80, 148, 166, 67, 157, 158, 77], where the finite element and spectral Galerkin methods were usually used for spatial discretization, and the time-stepping schemes include varies of the Crank–Nicolson method, Euler method, two-step backward difference formula, projection methods, fractional step methods and so on. However, the error estimates discussed in the aforementioned articles are not applicable to nonsmooth initial data.

When the initial value $u_0$ belongs to the space $\dot{H}_0^1(\Omega)$, a number of numerical analyses for the Navier-Stokes equations are available. The analysis in [127] essentially proves almost first-order convergence in time of the Runge–Kutta method for the two-dimensional NS equations when the initial value is in $\dot{H}_0^1(\Omega)$. In [71], Hill and Süli proved second-order convergence of the semidiscrete finite element method. For the implicit-explicit finite element method, first-order convergence in time and second-order convergence in space were proved under condition $\tau|\log h| \leqslant C$ in [64], where $\tau$ and $h$ are the temporal stepsize and spatial mesh size, respectively. Additionally, the error of semi-discretization in time by the Crank–Nicolson/Adams–Bashforth implicit-explicit scheme with a uniform stepsize was shown to be $O(\tau^{\frac{3}{2}})$ in [66]. This convergence rate is sharp with respect to the empirical numerical results. Second-order convergence in time and space was proved for a linearly extrapolated Crank–Nicolson scheme and a two-step backward differentiation formula by utilizing graded stepsizes locally refined towards $t = 0$; see [27, 111].

Discussions concerning the case that $u_0 \in \dot{L}^2(\Omega)$ are less prevalent in the literature. It has been known that $\dot{L}^2(\Omega)$ is a critical space for the well-posedness of the two-dimensional NS equations [51]. The error analysis in this case turns out to be significantly more challenging than for cases with smoother initial data, and the literature offers only a limited number of relevant results. Under the CFL condition, $\tau \leqslant C\lambda_m^{-1}$, it was shown in [65] that the implicit-explicit Euler spectral Galerkin method has an error bound of $O(\lambda_m^{-1/2} + \tau^{1/2})$ over a bounded time interval. For the implicit-explicit Euler scheme with finite element spatial discretization, several stability results were proved in [68] without error estimates. In more recent developments, first-order convergence in both time and space was shown

in [110] for high-order divergence-free finite elements. To our knowledge, this represents the most advanced convergence result obtained to date. However, there is still a gap between the numerical analysis and the numerical results, which demonstrate the possibility of achieving second-order convergence in space by using the Taylor–Hood finite elements. Proving second-order convergence of any numerical method for the NS equations remains an open and challenging task. Furthermore, the employed time-stepping scheme is of low order. developing higher-order schemes (with rigorous proof of the convergence rates) presents additional challenges due to limited smoothness of the solution and the nonlinearity of the NS equations. Recently, the construction and analysis of low-regularity integrators for nonlinear dispersive equations and NS equations based on energy techniques as well as harmonic analysis techniques become an active research area; see [108, 145, 152, 174]. The analyses in these articles generally require discovering and utilizing certain cancellation structures in the equations. An application of the general framework in [150] to the NS equations was shown in [108]. Since this approach does not use the smoothing property of the NS equations (thus the results are independent of the viscosity of fluid), it requires the initial value to be in $\dot{H}_0^1(\Omega) \cap H^2(\Omega)$ for the numerical solution to have first-order convergence in time.

In this chapter, we consider a fully discrete implicit-explicit Runge–Kutta finite element scheme for the NS equations with $L^2$ initial data by utilizing an $L^2$ projection $P_h^{\mathrm{RT}}$ onto the divergence-free subspace of the Raviart–Thomas element space in the numerical scheme. The linear term is discretized using the Runge–Kutta Lobatto IIIC scheme, while the nonlinear term is handled through an extrapolation approximation. To address the solution's singularity near $t = 0$, we employ graded stepsizes that provide enhanced resolution where needed. We prove the a nearly optimal error estimate. More specifically, let $u_h^{n+1}$ be the numerical solution of the fully discrete scheme at time level $t = t_{n+1}$. Theorems 3.1 and 3.2 show that, for arbitrarily small $\varepsilon > 0$,

$$\|u(t_{n+1}) - u_h^{n+1}\|_{L^2(\Omega)} \leqslant C_\varepsilon(h^{2-2\varepsilon} t_{n+1}^{\varepsilon-1} + t_{n+1}^{-2-\varepsilon}\tau_{n+1}^2),$$

where $\tau_{n+1}$ and $h$ denote the temporal stepsize of the $(n+1)$th step and spatial mesh size, respectively. A crucial element in our error analysis is the utilization of the $L^2$ projection $P_h^{\mathrm{RT}}$, which plays a key role in achieving second-order convergence in space and in deriving discrete energy decay, as detailed in Lemma 3.5. Our analysis also employs the discrete semigroup technique and the estimate of numerical solution in $H^1$ norm (Lemma 3.6), as well as some negative norm error estimates (Lemma 3.7). The choice of the Lobatto IIIC scheme is also critical for our analysis due to its distinctive property that the second internal

stage coincides with the endpoint of the time interval. This property is extensively used in the stability estimates, e.g., Lemma 3.6. Numerical examples are provided to support the theoretical analysis, which show that the numerical solutions of the NS equations with $L^2$ initial data achieve second-order convergence in both time and space. This is consistent with our theoretical analysis. Moreover, the convergence in space is at most second order even higher-order finite elements are used. This shows the sharpness of the convergence order proved in this chapter.

The rest of this chapter is organized as follows. In Section 3.2, we describe the finite element method for the spatial discretization using Taylor–Hood element, and present the error analysis of the semi-discrete scheme. The fully discrete scheme is developed and analyzed in Section 3.3. Some numerical experiments are shown in Section 3.4 to support and complement our theoretical analysis.

**3.2 Spatial semi-discretization by finite element method.** For $s \geqslant 0$ and $1 \leqslant p \leqslant \infty$, we denote by $W^{s,p}(\Omega)$ the conventional Sobolev spaces of functions defined on $\Omega$, with abbreviation $H^s(\Omega) = W^{s,2}(\Omega)$ and $L^p(\Omega) = W^{0,p}(\Omega)$. For the simplicity of notation, we denote by $\|\cdot\|_{W^{s,p}(\Omega)}$ the norm of the spaces $W^{s,p}(\Omega)$, $W^{s,p}(\Omega)^2$ and $W^{s,p}(\Omega)^{2\times 2}$, omitting the dependence on dimension.

Let $\dot{H}^1_0(\Omega) = \{v \in H^1_0(\Omega)^2 : \nabla \cdot v = 0\}$ and let $\dot{H}^s_0(\Omega) = (\dot{L}^2(\Omega), \dot{H}^1_0(\Omega))_{[s]}$ be the complex interpolation space between $\dot{L}^2(\Omega)$ and $\dot{H}^1_0(\Omega)$. The dual space of $\dot{H}^s_0(\Omega)$ is denoted by $\dot{H}^{-s}(\Omega)$.

**3.2.1 Weak solution.** Let $P_X$ be the $L^2$-orthogonal projection from $L^2(\Omega)^2$ to $\dot{L}^2(\Omega)$. Then any function $v \in L^2(\Omega)^2$ has a decomposition

$$v = P_X v + \nabla \eta, \tag{3.3}$$

where $\eta \in H^1(\Omega) \cap L^2_0(\Omega)$ satisfies the following elliptic equation with Neumann boundary condition

$$\begin{cases} \Delta \eta = \nabla \cdot v & \text{in} \quad \Omega, \\ \frac{\partial \eta}{\partial \nu} = v \cdot \nu & \text{on} \quad \partial\Omega. \end{cases}$$

Since $\nabla p$ is orthogonal to $\dot{L}^2(\Omega)$ for any function $p \in H^1(\Omega)$, it follows that $P_X \nabla p = 0$.

We denote by $A := P_X \Delta$ the Stokes operator on $\dot{L}^2(\Omega)$ with domain $D(A) = \dot{H}^1_0(\Omega) \cap H^2(\Omega)^2$, which is a self-adjoint operator and negative-definite. The Stokes operator has an extension to a bounded operator $A : \dot{H}^1_0(\Omega) \to \dot{H}^{-1}(\Omega)$ defined by

$$(Au, v) = -\int_\Omega \nabla u \cdot \nabla v \mathrm{d}x \quad \forall u, v \in \dot{H}^1_0(\Omega). \tag{3.4}$$

By applying $P_X$ to the first equation in (3.1), we obtain the following abstract parabolic equation in terms of the Stokes operator $A$:

$$\partial_t u - Au = -P_X(u \cdot \nabla u) \text{ in } \Omega \times (0, T]. \tag{3.5}$$

The the weak solution of (3.5) can be expressed as

$$u(\cdot, t) = e^{tA}u_0 - \int_0^t e^{(t-s)A} P_X(u(s) \cdot \nabla u(s))\mathrm{d}s. \tag{3.6}$$

The properties of operator $A$ are similar to the Laplacian operator $\Delta$. For example, for any functions $v, w \in \dot{H}_0^1(\Omega)$, $(Av, w) = -(\nabla v, \nabla w)$.

We recall the following regularity estimate of the solution proved in [110, Lemma 3.2].

**Lemma 3.1.** *For any given initial value $u_0 \in \dot{L}^2(\Omega)$, the exact solution $u$ of problem (3.1) satisfy the following regularity result.*

$$\|\partial_t^m u(\cdot, t)\|_{H^s(\Omega)} \leqslant Ct^{-\frac{s}{2}-m} \text{ for } 0 \leqslant s \leqslant 2, \ m = 0, 1, 2, \ldots \tag{3.7}$$

The exponential operator $e^{tA}$ plays a crucial role in the error analysis. By taking Laplace transform and inverse Laplace transform, we have

$$e^{tA} = \int_{|z|=\sigma} e^{zt}(z - A)^{-1}\mathrm{d}z,$$

for some constant $\sigma > 0$. Due to the analyticity of $e^{zt}(z - A)^{-1}$ in the sector $\{z \in \mathbf{C} : |\arg(z)| \leqslant \pi\}$, the straight line $|z| = \sigma$ in the complex plane can be deformed to a contour $\Gamma_{\delta,\kappa}$

$$\Gamma_{\delta,\kappa} = \{\kappa e^{i\theta} : -\delta \leqslant \theta \leqslant \delta\} \cup \{\rho e^{\pm i\delta} : \kappa \leqslant \rho < \infty\}.$$

Hence, the operator $e^{tA}$ has the form

$$e^{tA} = \int_{\Gamma_{\delta,\kappa}} e^{zt}(z - A)^{-1}\mathrm{d}z. \tag{3.8}$$

The stability estimate of the operator $e^{tA}$ then follows from the estimate of the resolvent operator $(z - A)^{-1}$.

**Lemma 3.1.** *The operator $e^{tA}P_X$ satisfies the following stability estimates.*

$$\|e^{tA}P_X f\|_{L^2(\Omega)} \leqslant \|f\|_{L^2(\Omega)}, \tag{3.9}$$

$$\|e^{tA}P_X f\|_{L^2(\Omega)} \leqslant Ct^{-\frac{s}{2}}\|f\|_{H^{-s}(\Omega)} \text{ for } 0 \leqslant s \leqslant 2, \tag{3.10}$$

$$\|e^{tA}P_X f\|_{L^2(\Omega)} \leqslant t^{-\frac{1}{r}}\|f\|_{W^{-1,r}(\Omega)} \text{ for } 1 < r \leqslant 2. \tag{3.11}$$

*Proof.* The first inequality follows from the relation (3.8) and the standard resolvent estimate (see [4, Theorem 3.7.11])

$$\|(z-A)^{-1}P_X f\|_{L^2} \leqslant C|z|^{-1}\|f\|_{L^2(\Omega)} \text{ for } z \in \Gamma_{\delta,\kappa}. \tag{3.12}$$

To show the second estimate, we let $w = (z-A)^{-1}P_X f$, then according to 3.12 we have $\|w\|_{L^2(\Omega)} \leqslant C|z|^{-1}\|f\|_{L^2(\Omega)}$. This together with the elliptic regularity estimate implies

$$\|w\|_{H^2(\Omega)} \leqslant \|Aw\|_{L^2(\Omega)} \leqslant \|zw - P_X f\|_{L^2(\Omega)} \leqslant C\|f\|_{L^2(\Omega)},$$

and hence

$$\|(z-A)^{-1}P_X f\|_{H^2(\Omega)} \leqslant C\|f\|_{L^2(\Omega)}.$$

Then by means of interpolation there holds

$$\|(z-A)^{-1}P_X f\|_{H^s(\Omega)} \leqslant C|z|^{-1+\frac{s}{2}}\|f\|_{L^2(\Omega)} \text{ for } 0 \leqslant s \leqslant 2. \tag{3.13}$$

Since the resolvent operator $(z-A)^{-1}P_X : L^2 \to \dot{H}_0^s$ is self-adjoint, we have

$$\|(z-A)^{-1}P_X f\|_{L^2} \leqslant C|z|^{-1+\frac{s}{2}}\|f\|_{H^{-s}(\Omega)} \text{ for } 0 \leqslant s \leqslant 2. \tag{3.14}$$

Then Substituting (3.14) into (3.8) and evaluating the integral leads to (3.10).

To prove (3.11), we apply the following embedding estimate in two dimension that

$$W^{-1,r}(\Omega) \hookrightarrow H^{-2/r}(\Omega) \text{ for } 1 < r \leqslant 2. \tag{3.15}$$

This, together with (3.10) with $s = 2/r$, leads to the estimate (3.11). □

**3.2.2 Spatial semi-discretization.** Let $\mathcal{T}_h$ denote a shape-regular and quasi-uniform triangulation of mesh size $h$. We define $\mathrm{RT}^1(\mathcal{T}_h)$ to be the $\mathrm{H}(\mathrm{div},\Omega)$-conforming Raviart-Thomas finite element space:

$$\mathrm{RT}^1(\mathcal{T}_h) := \{w \in \mathrm{H}(\mathrm{div},\Omega) : w|_K \in P_1(K)^2 + xP_1(K),\ \forall\ K \in \mathcal{T}_h\}.$$

Furthermore, we let $\mathrm{RT}_0^1(\Omega)$ be a subspace of $\mathrm{RT}^1(\Omega)$ such that

$$\mathrm{RT}_0^1(\mathcal{T}_h) := \{v_h \in \mathrm{RT}^1(\mathcal{T}_h) : \nabla \cdot v_h = 0 \text{ in } \Omega \text{ and } v_h \cdot \nu = 0 \text{ on } \partial\Omega\}.$$

Define the $L^2$ projection $P_h^{\mathrm{RT}}$ from $\dot{L}^2$ to $\mathrm{RT}_0^1$, that satisfies

$$(v - P_h^{\mathrm{RT}}v, \chi_h) = 0 \text{ for any } v \in \dot{L}^2(\Omega) \text{ and } \chi_h \in \mathrm{RT}_0^1(\mathcal{T}_h). \tag{3.16}$$

The projection $P_h^{\mathrm{RT}}$ satisfies the following estimate for $v \in X$ (cf. [112, Eq. (3.4)]):

$$\|P_h^{\mathrm{RT}}v - v\|_{L^2(\Omega)} \leqslant Ch^l\|v\|_{H^l(\Omega)},\ l = 1, 2. \tag{3.17}$$

Let the pair $(V_h, Q_h) \subset (H_0^1(\Omega), L_0^2(\Omega))$ denote the Taylor–Hood element spaces or Stokes-MINI element space, which have the following approximation properties (see [5, 169, 55]):

$$\inf_{v_h \in V_h} \|v - v_h\|_{H^s(\Omega)} + \inf_{q_h \in Q_h} \|q - q_h\|_{H^{s-1}(\Omega)} \leqslant C h^{m-s} \|v\|_{H^m(\Omega)}, \quad 0 \leqslant s \leqslant 1, \quad 1 \leqslant m \leqslant 2. \tag{3.18}$$

Both the Taylor–Hood and Stokes-MINI finite element spaces satisfy the discrete inf-sup condition, i.e., there is a generic constant $\kappa > 0$ such that

$$\sup_{v_h \in V_h, \nabla v_h \neq 0} \frac{(q_h, \nabla v_h)}{\|\nabla v_h\|_{L^2(\Omega)}} \geqslant \kappa \|q_h\|_{L^2(\Omega)} \quad \forall\, q_h \in Q_h. \tag{3.19}$$

We denote by $X_h := \{v_h \in V_h : (\nabla \cdot v_h, q_h) = 0 \ \forall\, q_h \in Q_h\}$ the discrete divergence-free subspace of $V_h$, and define the $L^2$ projection $P_{X_h}$ from $\dot{L}^2(\Omega)$ onto $X_h$ by the following relation:

$$(v - P_{X_h} v, w_h) = 0 \quad \forall\, w_h \in X_h. \tag{3.20}$$

The semi-discrete scheme for the NS equations in (3.1) reads: Find $(u_h, p_h) \in (V_h, Q_h)$ such that

$$(\partial_t u_h, v_h) + (P_h^{\mathrm{RT}} u_h \cdot \nabla u_h, v_h) + (\nabla u_h, \nabla v_h) - (p_h, \nabla \cdot v_h) = 0 \quad \forall\, v_h \in V_h, \tag{3.21a}$$

$$(\nabla \cdot u_h, q_h) = 0 \quad \forall\, q_h \in Q_h. \tag{3.21b}$$

Let $A_h : X_h \to X_h$ be the discrete Stokes operator defined by

$$(A_h v_h, w_h) = -(\nabla v_h, \nabla w_h) \quad \forall\, v_h, w_h \in X_h.$$

Then, by applying projection operator $P_{X_h}$ to (3.21), the semi-discrete scheme in (3.21) can be rewritten as

$$\partial_t u_h(\cdot, t) - A_h u_h(\cdot, t) = -P_{X_h}(P_h^{\mathrm{RT}} u_h(s) \cdot \nabla u_h(s)), \tag{3.22}$$

with initial value $u_h(\cdot, 0) = u_h^0 := P_{X_h} u_0$. By using Duhamel's formula, the solution to the semidiscrete problem (3.22) can be written as

$$u_h(\cdot, t) = e^{t A_h} u_h^0 - \int_0^t e^{(t-s) A_h} P_{X_h}(P_h^{\mathrm{RT}} u_h(s) \cdot \nabla u_h(s)) \mathrm{d}s. \tag{3.23}$$

**Remark 3.1.** If $\varphi_h \in X_h$ and $\varphi \in \dot{H}_0^1(\Omega)^2 \cap H^2(\Omega)^2$ satisfies the following relation:

$$A\varphi = A_h \varphi_h. \tag{3.24}$$

then there exist $q \in L^2(\Omega)$ and $q_h \in Q_h$ such that $(\varphi_h, q_h)$ is the Stokes-Ritz projection of $(\varphi, q)$, i.e., Ritz projection associated to the linear Stokes equations. This can be shown as follows: Let $q \in L^2_0(\Omega)$ and $q_h \in Q_h$ be the unique functions (determined via the continuous and discrete inf-sup conditions) such that

$$\begin{aligned} -(A\varphi, v) &= (\nabla\varphi, \nabla v) - (q, \nabla \cdot v) \qquad \forall\, v \in H^1_0(\Omega)^2, \\ -(A_h\varphi_h, v_h) &= (\nabla\varphi_h, \nabla v_h) - (q_h, \nabla \cdot v_h) \quad \forall\, v_h \in V_h. \end{aligned}$$

Then testing equation $-A\varphi = -A_h\varphi_h$ by $v_h \in V_h$ yields

$$(\nabla\varphi, \nabla v_h) - (q, \nabla \cdot v_h) = (\nabla\varphi_h, \nabla v_h) - (q_h, \nabla \cdot v_h) \quad \forall\, v_h \in V_h.$$

This shows that $(\varphi_h, q_h)$ is the Ritz projection of $(\varphi, q)$ associated to the linear Stokes equations. Moreover, via integration by parts we derive $\nabla q = \Delta\varphi - A\varphi$, which implies that

$$\|q\|_{H^{l-1}(\Omega)} \leqslant C\|\varphi\|_{H^l(\Omega)} \quad \text{for } l = 1, 2.$$

Therefore, the standard $L^2$ and $H^1$ error estimates for the Stokes-Ritz projection (see [55]) imply the following result:

$$\begin{aligned} \|\varphi_h - \varphi\|_{L^2(\Omega)} + h\|\varphi_h - \varphi\|_{H^1(\Omega)} &\leqslant Ch^l(\|\varphi\|_{H^l(\Omega)} + \|q\|_{H^{l-1}(\Omega)}) \\ &\leqslant Ch^l\|\varphi\|_{H^l(\Omega)} \quad \text{for} \quad l = 1, 2. \end{aligned} \tag{3.25}$$

Let $v \in \dot{H}^1_0(\Omega)^2 \cap H^2(\Omega)$ be the solution of the PDE problem $Av = \varphi$, and let $v_h \in X_h$ be the Stokes-Ritz projection of $v$ defined by $A_h v_h = Av = \varphi$. Then testing equation $-A\varphi = -A_h\varphi_h$ yields

$$\begin{aligned} \|\varphi\|^2_{L^2(\Omega)} &= (-A_h\varphi_h, v - v_h) - (\varphi_h, A_h v_h) \\ &\leqslant C\|A_h\varphi_h\|_{L^2(\Omega)}\|v - v_h\|_{L^2(\Omega)} + C\|\varphi_h\|_{L^2(\Omega)}\|A_h v_h\|_{L^2(\Omega)} \\ &\leqslant \|A_h\varphi_h\|_{L^2(\Omega)}Ch^2\|v\|_{H^2(\Omega)} + C\|\varphi_h\|_{L^2(\Omega)}\|\varphi\|_{L^2(\Omega)} \\ &\leqslant C\|\varphi_h\|_{L^2(\Omega)}\|\varphi\|_{L^2(\Omega)} + C\|\varphi_h\|_{L^2(\Omega)}\|\varphi\|_{L^2(\Omega)}, \end{aligned}$$

which implies the following $L^2$ stability result:

$$\|\varphi\|_{L^2(\Omega)} \leqslant C\|\varphi_h\|_{L^2(\Omega)}. \tag{3.26}$$

By testing equation $-A\varphi = -A_h\varphi_h$ with $\varphi$ we also obtain the following $H^1$ stability result:

$$\|\varphi\|_{H^1(\Omega)} \leqslant C\|\varphi_h\|_{H^1(\Omega)}. \tag{3.27}$$

The $L^p$ stability of the projection operator $P_h^{\mathrm{RT}}$ plays a pivotal role in the ensuing error analysis. The following lemma presents a fundamental result crucial for our investigations:

**Lemma 3.2.** *Let $\varphi_h \in X_h$, and $2 \leqslant p \leqslant \infty$, the following inequality holds:*

$$\|P_h^{RT}\varphi_h\|_{L^p(\Omega)} \leqslant \|\varphi_h\|_{L^p(\Omega)} + C\|\varphi_h\|_{L^2(\Omega)}^{\frac{2}{p}}\|\varphi_h\|_{H^1(\Omega)}^{1-\frac{2}{p}}. \tag{3.28}$$

*Proof.* For a function $\varphi_h \in X_h$, we let $\varphi$ be the solution to the elliptic PDE problem in (3.24). Thus $\varphi_h$ is the Stokes-Ritz projection of $\varphi$, satisfying the estimates in (3.25)–(3.27). Next, we proceed to estimate the $L^p$ norm of $P_h^{\mathrm{RT}}\varphi_h$ as follows:

$$\begin{aligned}
\|P_h^{\mathrm{RT}}\varphi_h\|_{L^p(\Omega)} \leqslant & \|\varphi_h\|_{L^p(\Omega)} + \|P_h^{\mathrm{RT}}\varphi_h - \varphi_h\|_{L^p(\Omega)} \\
\leqslant & \|\varphi_h\|_{L^p(\Omega)} + Ch^{\frac{2}{p}-1}\|P_h^{\mathrm{RT}}\varphi_h - \varphi_h\|_{L^2(\Omega)} \\
\leqslant & \|\varphi_h\|_{L^p(\Omega)} + Ch^{\frac{2}{p}-1}\big(\|P_h^{\mathrm{RT}}(\varphi_h - \varphi)\|_{L^2(\Omega)} + \|P_h^{\mathrm{RT}}\varphi - \varphi\|_{L^2(\Omega)} + \|\varphi - \varphi_h\|_{L^2(\Omega)}\big).
\end{aligned}$$

By incorporating the error estimates (3.17), (3.25), the stability estimate in (3.27), and the $L^2$ stability of $P_h^{\mathrm{RT}}$, we obtain

$$\|P_h^{\mathrm{RT}}\varphi_h\|_{L^p(\Omega)} \leqslant \|\varphi_h\|_{L^p(\Omega)} + Ch^{\frac{2}{p}}\|\varphi_h\|_{H^1(\Omega)} \leqslant \|\varphi_h\|_{L^p(\Omega)} + C\|\varphi_h\|_{L^2(\Omega)}^{\frac{2}{p}}\|\varphi_h\|_{H^1(\Omega)}^{1-\frac{2}{p}},$$

where we have used the inverse inequality of finite element functions. This proves the result in (3.28). □

When $p < \infty$, leveraging the interpolation inequality allows us to eliminate the first term on the right-hand side of (3.28). However, in the case when $p = \infty$, we encounter the task of estimating the $L^\infty$ norm of a finite element function in $X_h$. To address this, we present the following lemma.

**Lemma 3.3.** *The following inequality holds:*

$$\|\varphi_h\|_{L^\infty(\Omega)} \leqslant C\|\varphi_h\|_{L^2(\Omega)}^{\frac{1}{2}}\|A_h\varphi_h\|_{L^2(\Omega)}^{\frac{1}{2}}, \quad \forall \varphi_h \in X_h. \tag{3.29}$$

*Proof.* Let $\varphi$ be the solution of equation (3.24). Then the following standard regularity result hold:

$$\|\varphi\|_{H^2(\Omega)} \leqslant C\|A_h\varphi_h\|_{L^2(\Omega)}. \tag{3.30}$$

Therefore, the Sobolev interpolation inequality in [1, Theorem 5.9] implies that

$$\|\varphi\|_{L^\infty(\Omega)} \leqslant C\|\varphi\|_{L^2(\Omega)}^{\frac{1}{2}}\|\varphi\|_{H^2(\Omega)}^{\frac{1}{2}} \leqslant C\|\varphi_h\|_{L^2(\Omega)}^{\frac{1}{2}}\|A_h\varphi_h\|_{L^2(\Omega)}^{\frac{1}{2}}. \tag{3.31}$$

Using the inverse inequality and the error estimate (3.25), we have

$$\|I_h\varphi - \varphi_h\|_{L^\infty(\Omega)} \leqslant Ch^{-1}\|I_h\varphi - \varphi_h\|_{L^2(\Omega)} \leqslant Ch\|\varphi\|_{H^2(\Omega)}. \tag{3.32}$$

Using this result and the triangle inequality, we can bound $\|\varphi_h\|_{L^\infty(\Omega)}$ by

$$\begin{aligned}
\|\varphi_h\|_{L^\infty(\Omega)} \leqslant & \|I_h\varphi\|_{L^\infty(\Omega)} + \|I_h\varphi - \varphi_h\|_{L^\infty(\Omega)} \\
\leqslant & C\|\varphi\|_{L^\infty(\Omega)} + Ch\|\varphi\|_{H^2(\Omega)} && (L^\infty\text{-stability of } I_h) \\
\leqslant & C\|\varphi_h\|_{L^2(\Omega)}^{\frac{1}{2}}\|A_h\varphi_h\|_{L^2(\Omega)}^{\frac{1}{2}} + Ch\|A_h\varphi_h\|_{L^2(\Omega)} && (\text{here (3.30) and (3.31) are used}) \\
\leqslant & C\|\varphi_h\|_{L^2(\Omega)}^{\frac{1}{2}}\|A_h\varphi_h\|_{L^2(\Omega)}^{\frac{1}{2}} && (\text{inverse inequality}). \qquad (3.33)
\end{aligned}$$

This proves the result of Lemma 3.3. □

The discrete operator $A_h$ has similar property to $A$, we can obtain the regularity result for the semi-discrete numerical solution $u_h$ in the following lemma. The proof is similar to that of Lemma 3.1.

**Lemma 3.4.** *The semi-discrete solution $u_h$ to problem (3.22) is a function of $L^2(0,T;\dot{H}_0^1(\Omega))$ and satisfies*

$$\|\partial_t^m u_h(\cdot,t)\|_{H^s(\Omega)} \leqslant Ct^{-\frac{s}{2}-m} \text{ for } 0 \leqslant s \leqslant 1,\ m = 0,1,2,\ldots \tag{3.34}$$

According to [110, Eq. (3.5)], the projection operator $P_{X_h}$ is $H_0^1$ stable. By using a duality argument, we can derive that $P_{X_h}$ is $H^{-1}$ stable. The following corollary present some a priori estimates for the semi-discrete solution $u_h$ in negative norms.

**Corollary 3.1.** *This is the extension of Lemma 3.4. The semi-discrete numerical solution $u_h$ is a function of $L^2(0,T;\dot{H}_0^1(\Omega))$ satisfying that*

$$\|\partial_t^m u_h(\cdot,t)\|_{H^{-s}(\Omega)} \leqslant Ct^{-m+\frac{s}{2}} \text{ for } 0 \leqslant s \leqslant 1, \quad m = 1,2,\ldots \tag{3.35}$$

*Proof.* By the equation (3.22), the $H^{-1}$ stability of $P_{X_h}$, and the inequality (3.28), we have

$$\begin{aligned}
\|\partial_t u_h(\cdot,t)\|_{H^{-1}(\Omega)} \leqslant & C\|u_h(\cdot,t)\|_{H^1(\Omega)} + C\|[P_h^{\mathrm{RT}}u_h \cdot \nabla u_h](\cdot,t)\|_{H^{-1}(\Omega)} \\
\leqslant & C\|u_h(\cdot,t)\|_{H^1(\Omega)} + C\|[P_h^{\mathrm{RT}}u_h \otimes u_h](\cdot,t)\|_{L^2(\Omega)} \\
\leqslant & C\|u_h(\cdot,t)\|_{H^1(\Omega)} + C\|P_h^{\mathrm{RT}}u_h(\cdot,t)\|_{L^4(\Omega)}\|u_h(\cdot,t)\|_{L^4(\Omega)} \\
\leqslant & C\|u_h(\cdot,t)\|_{H^1(\Omega)} + C\|u_h(\cdot,t)\|_{H^1(\Omega)}\|u_h(\cdot,t)\|_{L^2(\Omega)} \leqslant Ct^{-\frac{1}{2}}.
\end{aligned}$$

We denote $u_h^{(m-1)} = \partial_t^{m-1}u_h$, $m \geqslant 2$, and differentiate (3.22) $m-1$ times, we obtain

$$\partial_t u_h^{(m-1)} - A_h u_h^{(m-1)} = -P_{X_h}\sum_{j=0}^{m-1}\binom{m-1}{j}(P_h^{\mathrm{RT}}u_h^{(j)} \cdot \nabla u_h^{(m-1-j)}).$$

Similar to the above process, we derive that

$$\|\partial_t u_h^{(m-1)}\|_{H^{-1}(\Omega)} \leqslant Ct^{-m+\frac{1}{2}}.$$

Using the interpolation inequality, (3.35) is verified. □

The next lemma provides error bounds between $e^{tA}P_X$ and $e^{tA_h}P_{X_h}$. [109, Lemma 4.5] The error between exact operator $e^{tA}P_X$ and $e^{tA_h}P_{X_h}$ is presented as follows

$$\|e^{tA}P_X - e^{tA_h}P_{X_h}\|_{L^2\to L^2} \leqslant Ct^{-1}h^2, \tag{3.36}$$

$$\|e^{tA}P_X - e^{tA_h}P_{X_h}\|_{L^2\to L^2} \leqslant Ct^{-\frac{1}{2}}h, \tag{3.37}$$

$$\|e^{tA}P_X - e^{tA_h}P_{X_h}\|_{H^{-1}\to L^2} \leqslant Ct^{-\frac{1}{2}}. \tag{3.38}$$

Next, we present an optimal error estimate for the semi-discrete scheme (3.22). Here we only consider the short-time error estimate, i.e., $T \leqslant T_0$ with $T_0$ sufficiently small. This case is more tricky since the $H^2$ norm of the solution exhibits singularity near $t = 0$. For large time estimate with $t > T_0$, the standard argument for the case that $u_0 \in \dot{H}_0^1(\Omega) \cap [H^2(\Omega)]^2$ works directly.

**Theorem 3.1.** *Suppose $u$ is the mild solution of (3.1) defined by (3.6), $u_h$ is the numerical solution defined by (3.23). Then the error $e(t) := u(t) - u_h(t)$ satisfies*

$$\|e(t)\|_{L^2(\Omega)} \leqslant Ct^{-1+\varepsilon}h^{2-2\varepsilon} \quad \forall\ t \in (0, T] \tag{3.39}$$

*for arbitrarily small $\varepsilon > 0$ and sufficiently small $T$.*

*Proof.* By using the equations (3.6) and (3.23), the error $\|e(t)\|_{L^2(\Omega)}$ can be decomposed as:

$$\begin{aligned}\|e(t)\|_{L^2(\Omega)} \leqslant& \Big\|(e^{tA}P_X - e^{tA_h}P_{X_h})u_0\Big\|_{L^2(\Omega)}\\ &+ \Big\| \int_0^t e^{(t-s)A}P_X\Big[u(s)\cdot\nabla u(s) - P_h^{\mathrm{RT}}u_h(s)\cdot\nabla u_h(s)\Big]\mathrm{d}s\Big\|_{L^2(\Omega)}\\ &+ \Big\| \int_0^t \Big[e^{(t-s)A}P_X - e^{(t-s)A_h}P_{X_h}\Big](P_h^{\mathrm{RT}}u_h(s)\cdot\nabla u_h(s))\mathrm{d}s\Big\|_{L^2(\Omega)}\\ =:& \mathcal{E}_1(t) + \mathcal{E}_2(t) + \mathcal{E}_3(t).\end{aligned}$$

The error $\mathcal{E}_1(t)$ follows from (3.36) and the $L^2$ stability of $e^{tA}$ and $e^{tA_h}$ such that

$$\mathcal{E}_1(t) \leqslant Ct^{-1+\varepsilon}h^{2-2\varepsilon}\|u_0\|_{L^2(\Omega)}. \tag{3.40}$$

For the estimate of $\mathcal{E}_2(t)$, since $u$ and $P_h^{\mathrm{RT}}u_h$ are both divergence free, by using (3.11) and choosing $r = 1/(1-\frac{\varepsilon}{2})$, we have

$$\mathcal{E}_2(t) \leqslant C\int_0^t (t-s)^{-1+\frac{\varepsilon}{2}}\|u(s)\otimes u(s) - P_h^{\mathrm{RT}}u_h(s)\otimes u_h(s)\|_{L^{1/(1-\varepsilon/2)}(\Omega)}\mathrm{d}s$$

$$
\begin{aligned}
\leqslant & C\int_0^t (t-s)^{-1+\frac{\varepsilon}{2}} \|P_h^{\mathrm{RT}} e(s)\otimes u(s) + P_h^{\mathrm{RT}} u_h(s)\otimes e(s)\|_{L^{1/(1-\varepsilon/2)}(\Omega)} \mathrm{d}s \\
& + C\int_0^t (t-s)^{-1+\frac{\varepsilon}{2}} \|(u(s)-P_h^{\mathrm{RT}} u(s))\otimes u(s)\|_{L^{1/(1-\varepsilon/2)}(\Omega)} \mathrm{d}s \\
\leqslant & C\int_0^t (t-s)^{-1+\frac{\varepsilon}{2}} \Big( \|P_h^{\mathrm{RT}} e(t)\|_{L^2(\Omega)} \|u(s)\|_{L^{\frac{2}{1-\varepsilon}}(\Omega)} + \|e(s)\|_{L^2(\Omega)} \|P_h^{\mathrm{RT}} u_h(s)\|_{L^{\frac{2}{1-\varepsilon}}(\Omega)} \Big) \mathrm{d}s \\
& + C\int_0^t (t-s)^{-1+\frac{\varepsilon}{2}} \|u(s)-P_h^{\mathrm{RT}} u(s)\|_{L^2(\Omega)} \|u(s)\|_{L^{\frac{2}{1-\varepsilon}}(\Omega)} \mathrm{d}s
\end{aligned}
$$

By using Lemma 3.1, Lemma 3.4, the error estimate (3.17), the $L^2(\Omega)$ stability of $P_h^{\mathrm{RT}}$, the estimate (3.28) for $p = 2/(1-\varepsilon)$, and the interpolation inequality, we have

$$
\begin{aligned}
\mathcal{E}_2(t) \leqslant & C\int_0^t (t-s)^{-1+\frac{\varepsilon}{2}} s^{-\frac{\varepsilon}{4}} \Big( \|u(s)\|_{H^1(\Omega)}^{\frac{\varepsilon}{2}} + \|u_h(s)\|_{H^1(\Omega)}^{\frac{\varepsilon}{2}} \Big) \|e(t)\|_{L^2(\Omega)} \mathrm{d}s \\
& + Ch^{2-2\varepsilon} \int_0^t (t-s)^{-1+\frac{\varepsilon}{2}} \|u(s)\|_{H^2(\Omega)}^{1-\varepsilon} \|u(s)\|_{H^1(\Omega)}^{\varepsilon} \mathrm{d}s \\
\leqslant & Ct^{-1+\varepsilon} h^{2-2\varepsilon} + C\int_0^t (t-s)^{-1+\frac{\varepsilon}{2}} s^{-\frac{\varepsilon}{4}} \Big( \|u(s)\|_{H^1(\Omega)}^{\frac{\varepsilon}{2}} + \|u_h(s)\|_{H^1(\Omega)}^{\frac{\varepsilon}{2}} \Big) \|e(t)\|_{L^2(\Omega)} \mathrm{d}s.
\end{aligned} \tag{3.41}
$$

For the estimate of $\mathcal{E}_3(t)$, by using Lemma 3.2.2, we have

$$
\begin{aligned}
\mathcal{E}_3(t) \leqslant & Ch^{2-2\varepsilon} \int_0^t (t-s)^{-1+\frac{\varepsilon}{2}} \|P_h^{\mathrm{RT}} u_h(s)\cdot\nabla u_h(s)\|_{L^2(\Omega)}^{1-\varepsilon} \|P_h^{\mathrm{RT}} u_h(s)\cdot\nabla u_h(s)\|_{H^{-1}(\Omega)}^{\varepsilon} \mathrm{d}s \\
\leqslant & Ch^{2-2\varepsilon} \int_0^t (t-s)^{-1+\frac{\varepsilon}{2}} \|P_h^{\mathrm{RT}} u_h(s)\|_{L^\infty(\Omega)}^{1-\varepsilon} \|\nabla u_h(s)\|_{L^2(\Omega)}^{1-\varepsilon} \|P_h^{\mathrm{RT}} u_h(s)\otimes u_h(s)\|_{L^2(\Omega)}^{\varepsilon} \mathrm{d}s \\
\leqslant & Ch^{2-2\varepsilon} \int_0^t (t-s)^{-1+\frac{\varepsilon}{2}} \|P_h^{\mathrm{RT}} u_h(s)\|_{L^\infty(\Omega)} \|\nabla u_h(s)\|_{L^2(\Omega)}^{1-\varepsilon} \|u_h(s)\|_{L^2(\Omega)}^{\varepsilon} \mathrm{d}s \\
\leqslant & Ch^{2-2\varepsilon} \int_0^t (t-s)^{-1+\frac{\varepsilon}{2}} \Big( \|u_h(s)\|_{L^\infty(\Omega)} + \|u_h(s)\|_{H^1(\Omega)} \Big) \|\nabla u_h(s)\|_{L^2(\Omega)}^{1-\varepsilon} \mathrm{d}s,
\end{aligned} \tag{3.42}
$$

where the last inequality follows from (3.28). By using Lemma 3.3, we have

$$
\|u_h(s)\|_{L^\infty(\Omega)} \leqslant C\|u_h(s)\|_{L^2(\Omega)}^{\frac{1}{2}} \|A_h u_h(s)\|_{L^2(\Omega)}^{\frac{1}{2}}. \tag{3.43}
$$

From the equation (3.22), we can estimate $\|A_h u_h(s)\|_{L^2(\Omega)}$ as follows by using the $L^2$ stability of $P_{X_h}$, (3.28) and Lemma 3.4

$$
\begin{aligned}
\|A_h u_h(s)\|_{L^2(\Omega)} \leqslant & \|\partial_t u_h(s)\|_{L^2(\Omega)} + \|P_h^{\mathrm{RT}} u_h(s)\cdot\nabla u_h(s)\|_{L^2(\Omega)} \\
\leqslant & Cs^{-1} + C\|P_h^{\mathrm{RT}} u_h(s)\|_{L^\infty(\Omega)} \|\nabla u_h(s)\|_{L^2(\Omega)}
\end{aligned}
$$

$$\begin{aligned}\leqslant & Cs^{-1}+C\Big(\|u_h(s)\|_{L^\infty(\Omega)}+\|u_h\|_{H^1(\Omega)}\Big)\|\nabla u_h(s)\|_{L^2(\Omega)}\\ \leqslant & Cs^{-1}+Cs^{-\frac{1}{2}}\|u_h(s)\|_{L^\infty(\Omega)}.\end{aligned}\tag{3.44}$$

Substituting (3.44) into (3.43) and using Young's inequality, we obtain

$$\|u_h(s)\|_{L^\infty(\Omega)}\leqslant Cs^{-\frac{1}{2}}.\tag{3.45}$$

Substituting (3.45) into (3.42) and using Lemma 3.4, we have

$$\mathcal{E}_3(t)\leqslant Ch^{2-2\varepsilon}\int_0^t(t-s)^{-1+\frac{\varepsilon}{2}}s^{-1+\frac{\varepsilon}{2}}\mathrm{d}s\leqslant Ct^{-1+\varepsilon}h^{2-2\varepsilon}.\tag{3.46}$$

Combining the estimates (3.40), (3.41) and (3.46), we obtain the estimate for $e(t)$

$$\begin{aligned}\|e(t)\|_{L^2(\Omega)}\leqslant & C\int_0^t(t-s)^{-1+\frac{\varepsilon}{2}}s^{-\frac{\varepsilon}{4}}\Big(\|u_h(s)\|_{H^1(\Omega)}^{\frac{\varepsilon}{2}}+\|u(s)\|_{H^1(\Omega)}^{\frac{\varepsilon}{2}}\Big)\|e(s)\|_{L^2(\Omega)}\mathrm{d}s\\ & +Ct^{-1+\varepsilon}h^{2-2\varepsilon}.\end{aligned}$$

Multiplying $t^{1-\varepsilon}$ on both sides derives that

$$\begin{aligned}& t^{1-\varepsilon}\|e(t)\|_{L^2(\Omega)}\\ \leqslant & Ct^{1-\varepsilon}\int_0^t(t-s)^{-1+\frac{\varepsilon}{2}}s^{-1+\frac{3\varepsilon}{4}}(\|u_h(s)\|_{H^1(\Omega)}^{\frac{\varepsilon}{2}}+\|u(s)\|_{H^1(\Omega)}^{\frac{\varepsilon}{2}})s^{1-\varepsilon}\|e(s)\|_{L^2(\Omega)}\mathrm{d}s+Ch^{2-2\varepsilon}.\end{aligned}$$

By Hölder's inequality, we have

$$\begin{aligned}\int_0^t(t-s)^{-1+\frac{\varepsilon}{2}}s^{-1+\frac{3\varepsilon}{4}}\|u(s)\|_{H^1(\Omega)}^{\frac{\varepsilon}{2}}\mathrm{d}s\leqslant & \|u\|_{L^2(0,t;\dot{H}_0^1(\Omega))}^{\frac{\varepsilon}{2}}\Big(\int_0^t\Big[(t-s)^{-1+\frac{\varepsilon}{2}}s^{-1+\frac{3\varepsilon}{4}}\Big]^{\frac{4}{4-\varepsilon}}\mathrm{d}s\Big)^{\frac{4-\varepsilon}{4}}\\ \leqslant & Ct^{-1+\varepsilon}\|u\|_{L^2(0,t;\dot{H}_0^1(\Omega))}^{\frac{\varepsilon}{2}}.\end{aligned}$$

Combining the above inequalities above, we have

$$t^{1-\varepsilon}\|e(t)\|_{L^2(\Omega)}\leqslant Ch^{2-2\varepsilon}+C\Big(\|u\|_{L^2(0,t;\dot{H}_0^1(\Omega))}^{\frac{\varepsilon}{2}}+\|u_h\|_{L^2(0,t;\dot{H}_0^1(\Omega))}^{\frac{\varepsilon}{2}}\Big)\sup_{0<s\leqslant t}s^{1-\varepsilon}\|e(s)\|_{L^2(\Omega)}.$$

Taking the supremum with respect to $t$ on both sides deduce that

$$\begin{aligned}\sup_{0<t\leqslant T}t^{1-\varepsilon}\|e(t)\|_{L^2(\Omega)}\leqslant & Ch^{2-2\varepsilon}\\ & +C\Big(\|u\|_{L^2(0,T;\dot{H}_0^1(\Omega))}^{\frac{\varepsilon}{2}}+\|u_h\|_{L^2(0,T;\dot{H}_0^1(\Omega))}^{\frac{\varepsilon}{2}}\Big)\sup_{0<t\leqslant T}t^{1-\varepsilon}\|e(t)\|_{L^2(\Omega)}.\end{aligned}$$

According to [110, Lemma 3.5], for any small $\sigma>0$, there exists $T_\sigma>0$ such that

$$\|u\|_{L^2(0,T;\dot{H}_0^1(\Omega))}+\|u_h\|_{L^2(0,T;\dot{H}_0^1(\Omega))}\leqslant\sigma\quad\forall T\in(0,T_\sigma].$$

If $T$ satisfies $C\Big(\|u\|^{\varepsilon}_{L^2(0,T;\dot{H}^1_0(\Omega))} + \|u_h\|^{\varepsilon}_{L^2(0,T;\dot{H}^1_0(\Omega))}\Big) < 1$, then we have

$$\sup_{0<t\leqslant T} t^{1-\varepsilon}\|e(t)\|_{L^2(\Omega)} \leqslant Ch^{2-2\varepsilon},$$

and complete the proof of theorem.

□

**3.3 Fully discretization.** In this section, we propose and analyze a fully discrete scheme by using a second-order implicit-explicit Runge–Kutta method.

**3.3.1 Runge-Kutta method and error equations.** Let $0 = t_0 < t_1 < ... < t_N = T$ be a partition of the time interval $[0, T]$ with stepsize

$$\tau_1 = \tau^{\frac{1}{1-\alpha}} \text{ and } \tau_n = t_n - t_{n-1} \sim (t_{n-1}/T)^{\alpha}\tau \text{ for } 2 \leqslant n \leqslant N, \tag{3.47}$$

where $\tau$ is the maximal stepsize, and " $\sim$ " means equivalent magnitude (up to a constant multiple). The parameter $\alpha \in (0, 1)$ determines how fast the temporal grids are refined towards $t = 0$. The stepsizes defined in this way have the following properties:

1. $\tau_n \sim \tau_{n-1}$ for two consecutive stepsizes.

2. For any fixed integer $M_0$, $\tau_1 \sim \tau_2 \sim \cdots \sim \tau_{M_0} \sim \tau^{\frac{1}{1-\alpha}}$, the equivalence depends on $M_0$, but is independent on $\tau$ and $n$. Hence, the starting stepsize is much smaller than the maximal stepsize. This resolves the solution's singularity near $t = 0$.

3. The total number of time levels is $O(T/\tau)$. Therefore, the total computational cost is equivalent to using a uniform stepsize $\tau$.

Next, we introduce an implicit Runge–Kutta method with $q$ stages for the time discretization of the evolution equation (3.5). The coefficients of the method are given by the Butcher tableau

$$\begin{array}{c|ccc} a_{11} & \cdots & a_{1q} & c_1 \\ \vdots & & \vdots & \vdots \\ a_{q1} & \cdots & a_{qq} & c_q \\ \hline b_1 & \cdots & b_q & \end{array}$$

with $c_1, \ldots, c_q \in (0, 1]$. Here the quadrature points $c_i, 1 \leqslant i \leqslant q$, are distinct numbers in $[0, 1]$ and the coefficients $a_{ij}$ and $b_j$ are associated with the quadrature formulas

$$\int_0^1 \varphi \mathrm{d}t \approx \sum_{j=1}^{q} b_j\varphi(c_j), \qquad \int_0^{c_i} \varphi \mathrm{d}t \approx \sum_{j=1}^{q} a_{ij}\varphi(c_j),\ i = 1, \ldots, q. \tag{3.48}$$

We assume that (3.48) are exact for polynomials of degree $p-1$ and $p-2$, respectively. It implies that the method is accurate of order $p$. Now we introduce error functionals for the quadrature formulae (3.48) for the interval $(t_n, t_{n+1})$ as

$$\begin{aligned} Q_{n,i}(\varphi) &= \int_{t_n}^{t_{n,i}} \varphi \mathrm{d}s - \tau_{n+1} \sum_{j=1}^{q} a_{ij} \varphi(t_{n,j}), \quad i = 1, \cdots, q, \\ Q_{n+1}(\varphi) &= \int_{t_n}^{t_{n+1}} \varphi \mathrm{d}s - \tau_{n+1} \sum_{i=1}^{q} b_l \varphi(t_{n,i}). \end{aligned} \tag{3.49}$$

Recall the assumption that the quadrature formulae (3.48) are exact for polynomials of degree $p-1$ and $p-2$, respectively (this means that the time discretization scheme is strictly accurate of $p$). As a result, we have (see [29])

$$\begin{aligned} &\|Q_{n,i}(\varphi)\| \leqslant C\tau_{n+1}^{l+1} \sup_{t_n<s<t_{n+1}} \|\varphi^{(l)}(s)\| \quad \text{for } l \leqslant p-1,\ i = 1, 2. \\ &\|Q_{n+1}(\varphi)\| \leqslant C\tau_{n+1}^{l+1} \sup_{t_n<s<t_{n+1}} \|\varphi^{(l)}(s)\| \quad \text{for } l \leqslant p. \end{aligned} \tag{3.50}$$

where $\|\cdot\|$ can be $L^2(\Omega)$ norm or $H^1(\Omega)$ norm.

Taking $\mathcal{O} = (a_{ij})$, the vectors $b = (b_j)$ and $c = (c_i)^T$ for $i, j = 1, \cdots, q$. Here we use the two-stage Lobatto IIIC scheme, with $p = q = 2$, namely

$$\mathcal{O} = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}, \quad b = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \end{pmatrix}, \quad c = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

It is well-known that the method is implicit and algebraically stable [61]. In the numerical scheme, we linearize the nonlinear term in the Navier–Stokes equation. For a sequence of finite element functions $\{v_h^{n,i}\}$ for $n = 0, 1, \cdots$ and $i = 1, 2$, we define the extrapolation operator $\hat{I}_h$ as follows:

$$\hat{I}_h v_h^{n,i} = \begin{cases} v_h^0, & n = 0, \\ v_h^n + c_i \frac{\tau_{n+1}}{\tau_n} (v_h^n - v_h^{n-1}), & n \geqslant 1. \end{cases} \tag{3.51}$$

Then for a function $f$, we have the following error estimate for the extrapolation operator $\hat{I}_h$ for $n \geqslant 1$:

$$\|\hat{I}_h f(t_{n,i}) - f(t_{n,i})\| \leqslant C\tau_{n+1}^2 \sup_{t_{n-1}<s<t_{n+1}} \|\partial_t^2 f(\cdot, s)\| \quad \text{for} \quad i = 1, 2, \tag{3.52}$$

where $\|\cdot\|$ can be $L^2(\Omega)$ norm or $H^1(\Omega)$ norm.

For given numerical solutions $u_h^{n-1}, u_h^n \in X_h$, we compute $u_h^{n+1} \in X_h$ by

$$u_h^{n,i} = u_h^n + \tau_{n+1} \sum_{j=1}^{2} a_{ij} \left[ A_h u_h^{n,j} - P_{X_h}(P_h^{\mathrm{RT}} \hat{I}_h u_h^{n,j} \cdot \nabla u_h^{n,j}) \right], \ i = 1, 2, \tag{3.53a}$$

$$u_h^{n+1} = u_h^n + \tau_{n+1} \sum_{i=1}^{2} b_i \left[ A_h u_h^{n,i} - P_{X_h}(P_h^{\mathrm{RT}} \hat{I}_h u_h^{n,i} \cdot \nabla u_h^{n,i}) \right], \tag{3.53b}$$

Here $u_h^{n,i}$ are approximations to $u_h(t_{n,i})$ for $i = 1, 2$, with $t_{n,i} = t_n + c_i \tau_{n+1}$ being the internal Runge-Kutta nodes.

Recalling the truncation errors $Q_{n,i}(\partial_t u_h)$ and $Q_n(\partial_t u_h)$, we write the the semi-discrete solution $u_h$ as

$$\begin{aligned} u_h(t_{n,i}) =& u_h(t_n) + \tau_{n+1} \sum_{j=1}^{2} a_{ij} \left[ A_h u_h(t_{n,j}) - P_{X_h}(P_h^{\mathrm{RT}} \hat{I}_h u_h(t_{n,j}) \cdot \nabla u_h(t_{n,j})) \right] \\ &+ \tau_{n+1} \sum_{j=1}^{2} a_{ij} \left[ P_{X_h}(P_h^{\mathrm{RT}} \hat{I}_h u_h(t_{n,j}) \cdot \nabla u_h(t_{n,j})) - P_{X_h}(P_h^{\mathrm{RT}} u_h(t_{n,j}) \cdot \nabla u_h(t_{n,j})) \right] \\ &+ Q_{n,i}(\partial_t u_h), \quad i = 1, 2, \end{aligned} \tag{3.54}$$

$$\begin{aligned} u_h(t_{n+1}) =& u_h(t_n) + \tau_{n+1} \sum_{i=1}^{2} b_i \left[ A_h u_h(t_{n,i}) - P_{X_h}(P_h^{\mathrm{RT}} \hat{I}_h u_h(t_{n,i}) \cdot \nabla u_h(t_{n,i})) \right] \\ &+ \tau_{n+1} \sum_{i=1}^{2} b_i \left[ P_{X_h}(P_h^{\mathrm{RT}} \hat{I}_h u_h(t_{n,i}) \cdot \nabla u_h(t_{n,i})) - P_{X_h}(P_h^{\mathrm{RT}} u_h(t_{n,i}) \cdot \nabla u_h(t_{n,i})) \right] \\ &+ Q_{n+1}(\partial_t u_h). \end{aligned} \tag{3.55}$$

Now we define

$$\begin{aligned} \mathcal{G}^{n,i} &= P_{X_h}(P_h^{\mathrm{RT}} \hat{I}_h u_h(t_{n,i}) \cdot \nabla u_h(t_{n,i})) - P_{X_h}(P_h^{\mathrm{RT}} u_h(t_{n,i}) \cdot \nabla u_h(t_{n,i})), \\ T^{n,i} &= -P_{X_h}(P_h^{\mathrm{RT}} \hat{I}_h u_h^{n,i} \cdot \nabla u_h^{n,i}) + P_{X_h}(P_h^{\mathrm{RT}} \hat{I}_h u_h(t_{n,i}) \cdot \nabla u_h(t_{n,i})). \end{aligned}$$

Then the errors $\eta^{n+1} = u_h^{n+1} - u_h(t_{n+1})$ and $\eta^{n,i} = u_h^{n,i} - u_h(t_{n,i})$ satisfy

$$\begin{aligned} &\dot{\eta}^{n,i} = A_h \eta^{n,i} + T^{n,i}, \\ &\eta^{n,i} = \eta^n + \tau_{n+1} \sum_{j=1}^{2} a_{ij} \dot{\eta}^{n,j} - \tau_{n+1} \sum_{j=1}^{2} a_{ij} \mathcal{G}^{n,j} - Q_{n,i}(\partial_t u_h) \quad i = 1, 2, \\ &\eta^{n+1} = \eta^n + \tau_{n+1} \sum_{i=1}^{2} b_i \dot{\eta}^{n,i} - \tau_{n+1} \sum_{i=1}^{2} b_i \mathcal{G}^{n,i} - Q_{n+1}(\partial_t u_h). \end{aligned} \tag{3.56}$$

In order to estimate the extrapolation error $\mathcal{G}^{n,i}$, we first derive an a priori estimate for $A_h u_h(t_{n,i})$. In combination with (3.44) and (3.45), we have

$$\|A_h u_h(t_{n,i})\|_{L^2(\Omega)} \leqslant C t_{n,i}^{-1}. \tag{3.57}$$

According to (3.52) and (3.57), $\mathcal{G}^{n,i}$ satisfies

$$\begin{aligned}\|\mathcal{G}^{n,i}\|_{L^2(\Omega)} \leqslant & C\|\hat{I}_h u_h(t_{n,i}) - u_h(t_{n,i})\|_{L^4(\Omega)}\|\nabla u_h(t_{n,i})\|_{L^4(\Omega)} \\ \leqslant & C\|\hat{I}_h u_h(t_{n,i}) - u_h(t_{n,i})\|_{L^2(\Omega)}^{1/2}\|\hat{I}_h u_h(t_{n,i}) - u_h(t_{n,i})\|_{H^1(\Omega)}^{1/2} \\ & \cdot \|u_h(t_{n,i})\|_{H^1(\Omega)}^{1/2}\|A_h u_h(t_{n,i})\|_{L^2(\Omega)}^{1/2} \\ \leqslant & \tau_{n+1}^2 t_{n+1}^{-3}.\end{aligned} \tag{3.58}$$

Similarly, we have the estimate in $H^{-1}$ norm

$$\|\mathcal{G}^{n,i}\|_{H^{-1}(\Omega)} \leqslant C\|P_h^{\mathrm{RT}}(\hat{I}_h u_h(t_{n,i}) - u_h(t_{n,i})) \otimes u_h(t_{n,i})\|_{L^2(\Omega)} \leqslant C t_{n+1}^{-5/2}\tau_{n+1}^2. \tag{3.59}$$

**3.3.2 Regularities of numerical solutions and estimates for operators.** In this subsection, we prove $L^2(\Omega)^2$ boundedness, $L^2(0,T;H_0^1(\Omega)^2)$ boundedness and $H^1(\Omega)^2$ estimate of the fully discrete solution in (3.53) by using energy estimates.

The $L^2(\Omega)^2$ and $L^2(0,T;H_0^1(\Omega)^2)$ boundedness of the solution of the fully discrete scheme (3.53) is presented in the following lemma.

**Lemma 3.5.** *(Discrete energy decay for the NS equation) Assume that $u_h^n \in X_h$ is given. Then, the solutions $u_h^{n,i} \in X_h$, $i = 1, 2$ and $u_h^{n+1} \in X_h$ of fully discrete scheme* (3.53) *satisfy the following estimate:*

$$\|u_h^{n+1}\|_{L^2(\Omega)}^2 \leqslant \|u_h^n\|_{L^2(\Omega)}^2 - 2\tau_{n+1}\sum_{i=1}^{2} b_i\|\nabla u_h^{n,i}\|_{L^2(\Omega)}^2, \quad \textit{for} \quad n \geqslant 0, \tag{3.60}$$

$$\begin{aligned}\sum_{i=1}^{2}\|u_h^{n,i}\|_{L^2(\Omega)}^2 \leqslant & C\|u_h^n\|_{L^2(\Omega)}^2 + C\tau_{n+1}\sum_{i=1}^{2}\|\nabla u_h^{n,i}\|_{L^2(\Omega)}^2 \\ & + C\tau_{n+1}^2\sum_{i=1}^{2}\|\hat{I}_h u_h^{n,i}\|_{L^2(\Omega)}^2\|\hat{I}_h u_h^{n,i}\|_{H^1(\Omega)}^2\|u_h^{n,i}\|_{H^1(\Omega)}^2, \quad \textit{for} \quad n \geqslant 0.\end{aligned} \tag{3.61}$$

*Proof.* First, we rewrite the numerical scheme (3.53) as

$$\dot{u}_h^{n,i} = A_h u_h^{n,i} - P_{X_h}(P_h^{\mathrm{RT}}\hat{I}_h u_h^{n,i} \cdot \nabla u_h^{n,i}),\ i = 1, 2, \tag{3.62a}$$

$$u_h^{n,i} = u_h^n + \tau_{n+1}\sum_{j=1}^{2} a_{ij}\dot{u}_h^{n,j},\ i = 1, 2, \tag{3.62b}$$

$$u_h^{n+1} = u_h^n + \tau_{n+1}\sum_{i=1}^{2} b_i \dot{u}_h^{n,i}. \tag{3.62c}$$

According to the (3.62c), we conclude

$$\begin{aligned}\|u_h^{n+1}\|_{L^2(\Omega)}^2 =& \left(u_h^n + \tau_{n+1}\sum_{i=1}^{2} b_i \dot{u}_h^{n,i}, u_h^n + \tau_{n+1}\sum_{i=1}^{2} b_i \dot{u}_h^{n,i}\right)\\ =&\|u_h^n\|_{L^2(\Omega)}^2 + 2\tau_{n+1}\sum_{i=1}^{2} b_i(\dot{u}_h^{n,i}, u_h^n) + \tau_{n+1}^2\sum_{i,j=1}^{2} b_i b_j(\dot{u}_h^{n,i}, \dot{u}_h^{n,j}).\end{aligned} \tag{3.63}$$

Substituting (3.62b) into the second term on the right-hand side of (3.63), we obtain

$$\begin{aligned}\|u_h^{n+1}\|_{L^2(\Omega)}^2 =&\|u_h^n\|_{L^2(\Omega)}^2 + 2\tau_{n+1}\sum_{i=1}^{2} b_i\Big(\dot{u}_h^{n,i}, u_h^{n,i} - \tau_{n+1}\sum_{j=1}^{2} a_{ij}\dot{u}_h^{n,j}\Big)\\ &+ \tau_{n+1}^2\sum_{i,j=1}^{2} b_i b_j(\dot{u}_h^{n,i}, \dot{u}_h^{n,j}).\end{aligned}$$

Hence

$$\|u_h^{n+1}\|_{L^2(\Omega)}^2 = \|u_h^n\|_{L^2(\Omega)}^2 + 2\tau_{n+1}\sum_{i=1}^{2} b_i(\dot{u}_h^{n,i}, u_h^{n,i}) - \tau_{n+1}^2\sum_{i,j=1}^{2} d_{ij}(\dot{u}_h^{n,i}, \dot{u}_h^{n,j}),$$

with $d_{ij} = b_i a_{ij} + b_j a_{ji} - b_i b_j$, $i, j = 1, 2$. The scheme is algebraic stable, i.e. the symmetric matrix $(d_{ij})$ is positive semidefinite. Therefore,

$$\|u_h^{n+1}\|_{L^2(\Omega)}^2 \leqslant \|u_h^n\|_{L^2(\Omega)}^2 + 2\tau_{n+1}\sum_{i=1}^{2} b_i(\dot{u}_h^{n,i}, u_h^{n,i}). \tag{3.64}$$

Testing (3.62a) with $u_h^{n,i}$ yields

$$(\dot{u}_h^{n,i}, u_h^{n,i}) = -\|\nabla u_h^{n,i}\|_{L^2(\Omega)}^2 - (P_{X_h}(P_h^{\mathrm{RT}}\hat{I}_h u_h^{n,i}\cdot\nabla u_h^{n,i}), u_h^{n,i}),\ i = 1, 2. \tag{3.65}$$

Note that $P_h^{\mathrm{RT}}\hat{I}_h u_h^{n,i}$ is divergence free. Then we have

$$(P_h^{\mathrm{RT}}\hat{I}_h u_h^{n,i}\cdot\nabla u_h^{n,i}, u_h^{n,i}) = \left(P_h^{\mathrm{RT}}\hat{I}_h u_h^{n,i}, \nabla\frac{1}{2}|u_h^{n,i}|^2\right) = -\left(\nabla\cdot P_h^{\mathrm{RT}}\hat{I}_h u_h^{n,i}, \frac{1}{2}|u_h^{n,i}|^2\right) = 0.$$

As a result, we obtain the inequality (3.60) by substituting (3.65) into (3.64).

To prove the $L^2$ boundedness of $u_h^{n,i}$, we test the equation (3.53a) with $u_h^{n,i}$ and obtain

$$\|u_h^{n,i}\|_{L^2(\Omega)}^2 =(u_h^n, u_h^{n,i}) - \tau_{n+1}\sum_{j=1}^{2} a_{ij}\Big[(\nabla u_h^{n,j}, \nabla u_h^{n,i}) + (P_h^{\mathrm{RT}}\hat{I}_h u_h^{n,j}\cdot\nabla u_h^{n,j}, u_h^{n,i})\Big]$$

$$\begin{aligned}
\leqslant & \frac{1}{2}\|u_h^{n,i}\|_{L^2(\Omega)}^2 + \frac{1}{2}\|u_h^n\|_{L^2(\Omega)}^2 + C\tau_{n+1}\sum_{j=1}^{2}\|\nabla u_h^{n,j}\|_{L^2(\Omega)}^2 \\
& + \tau_{n+1}\|u_h^{n,i}\|_{H^1(\Omega)}\sum_{j=1}^{2}\|P_h^{\mathrm{RT}}\hat{I}_h u_h^{n,j}\cdot\nabla u_h^{n,j}\|_{H^{-1}(\Omega)} \qquad (3.66) \\
\leqslant & \frac{1}{2}\|u_h^{n,i}\|_{L^2(\Omega)}^2 + \frac{1}{2}\|u_h^n\|_{L^2(\Omega)}^2 + C\tau_{n+1}\sum_{j=1}^{2}\|\nabla u_h^{n,j}\|_{L^2(\Omega)}^2 \\
& + C\tau_{n+1}\sum_{j=1}^{2}\|P_h^{\mathrm{RT}}\hat{I}_h u_h^{n,j}\otimes u_h^{n,j}\|_{L^2(\Omega)}^2. \qquad (3.67)
\end{aligned}$$

By using Hölder's inequality, the estimate (3.28), we have

$$\begin{aligned}
\|P_h^{\mathrm{RT}}\hat{I}_h u_h^{n,j}\otimes u_h^{n,j}\|_{L^2(\Omega)}^2 \leqslant & \|P_h^{\mathrm{RT}}\hat{I}_h u_h^{n,j}\|_{L^4(\Omega)}^2\|u_h^{n,j}\|_{L^4(\Omega)}^2 \\
\leqslant & C\|\hat{I}_h u_h^{n,j}\|_{L^2(\Omega)}\|\hat{I}_h u_h^{n,j}\|_{H^1(\Omega)}\|u_h^{n,j}\|_{L^2(\Omega)}\|u_h^{n,j}\|_{H^1(\Omega)} \qquad (3.68)
\end{aligned}$$

Substituting (3.68) into (3.67), summing up the obtained inequality with respect to $i$ from $i=1$ to $i=2$, and using Young's inequality, we obtain the desired result (3.61).

□

Then next lemma gives an a priori estimate for $\|\nabla u_h^{n,i}\|_{L^2(\Omega)}$.

**Lemma 3.6.** *If $u_0 \in \dot{L}^2(\Omega)$, then the fully discrete scheme* (3.53) *satisfy*

$$\sum_{i=1}^{2}\|\nabla u_h^{n,i}\|_{L^2(\Omega)} \leqslant Ct_{n+1}^{-1/2}, \quad \text{for} \quad n \geqslant 0. \qquad (3.69)$$

*Proof.* First of all, we note that the inequality (3.69) holds when $n = 0, 1, 2$ according to Lemma 3.5. Then for $n \geqslant 3$, taking gradient on both sides of (3.62c) and squaring, we have

$$\|\nabla u_h^{n+1}\|_{L^2(\Omega)}^2 = \|\nabla u_h^n\|_{L^2(\Omega)}^2 + 2\tau_{n+1}\sum_{i=1}^{2} b_i(\nabla \dot{u}_h^{n,i}, \nabla u_h^n) + \tau_{n+1}^2\sum_{i,j=1}^{2} b_i b_j(\nabla \dot{u}_h^{n,i}, \nabla \dot{u}_h^{n,j}).$$

Meanwhile, we recall (3.62b) and obtain

$$(\nabla u_h^n, \nabla \dot{u}_h^{n,i}) = (\nabla u_h^{n,i}, \nabla \dot{u}_h^{n,i}) - \tau_{n+1}\sum_{j=1}^{2} a_{ij}(\nabla \dot{u}_h^{n,j}, \nabla \dot{u}_h^{n,i})$$

Therefore,

$$\|\nabla u_h^{n+1}\|_{L^2(\Omega)}^2 = \|\nabla u_h^n\|_{L^2(\Omega)}^2 + 2\tau_{n+1}\sum_{i=1}^{2} b_i(\nabla u_h^{n,i}, \nabla \dot{u}_h^{n,i}) - 2\tau_{n+1}^2\sum_{i,j=1}^{2} a_{ij} b_i(\nabla \dot{u}_h^{n,i}, \nabla \dot{u}_h^{n,j})$$

$$+\tau_{n+1}^2\sum_{i,j=1}^{2} b_i b_j(\nabla \dot{u}_h^{n,i},\nabla \dot{u}_h^{n,j})$$

Then we apply the algebraical stability of the scheme to obtain

$$\|\nabla u_h^{n+1}\|_{L^2(\Omega)}^2 \leqslant \|\nabla u_h^{n}\|_{L^2(\Omega)}^2 + 2\tau_{n+1}\sum_{i=1}^{2} b_i(\nabla u_h^{n,i},\nabla \dot{u}_h^{n,i}). \tag{3.70}$$

Testing (3.62a) with $\dot{u}_h^{n,i}$, we have

$$(\nabla u_h^{n,i},\nabla \dot{u}_h^{n,i}) = -\|\dot{u}_h^{n,i}\|_{L^2(\Omega)}^2 - (P_h^{\mathrm{RT}}\hat{I}_h u_h^{n,i}\cdot\nabla u_h^{n,i},\dot{u}_h^{n,i}). \tag{3.71}$$

Substituting (3.71) into (3.70), we have

$$\|\nabla u_h^{n+1}\|_{L^2(\Omega)}^2 \leqslant \|\nabla u_h^{n}\|_{L^2(\Omega)}^2 - 2\tau_{n+1}\sum_{i=1}^{2} b_i\|\dot{u}_h^{n,i}\|_{L^2(\Omega)}^2 - 2\tau_{n+1}\sum_{i=1}^{2} b_i(P_h^{\mathrm{RT}}\hat{I}_h u_h^{n,i}\cdot\nabla u_h^{n,i},\dot{u}_h^{n,i}).$$

Since $b_i > 0$ for each $i = 1, 2$, by Hölder's inequality, we have

$$\begin{aligned}\|\nabla u_h^{n+1}\|_{L^2(\Omega)}^2 &+\tau_{n+1}\sum_{i=1}^{2} b_i\|\dot{u}_h^{n,i}\|_{L^2(\Omega)}^2\\ &\leqslant \|\nabla u_h^{n}\|_{L^2(\Omega)}^2 + C\tau_{n+1}\sum_{i=1}^{2}\|P_h^{\mathrm{RT}}\hat{I}_h u_h^{n,i}\cdot\nabla u_h^{n,i}\|_{L^2(\Omega)}^2.\end{aligned} \tag{3.72}$$

According to [110, Lemma 3.1] and estimate (3.28), we have

$$\|P_h^{\mathrm{RT}}\hat{I}_h u_h^{n,i}\cdot\nabla u_h^{n,i}\|_{L^2(\Omega)} \leqslant \|\hat{I}_h u_h^{n,i}\|_{L^2(\Omega)}^{1/2}\|\hat{I}_h u_h^{n,i}\|_{H^1(\Omega)}^{1/2}\|u_h^{n,i}\|_{H^1(\Omega)}^{1/2}\|A_h u_h^{n,i}\|_{L^2(\Omega)}^{1/2}, \tag{3.73}$$

where we have used interpolation inequality. By (3.62a), we have

$$\|A_h u_h^{n,i}\|_{L^2(\Omega)} \leqslant \|\dot{u}_h^{n,i}\|_{L^2(\Omega)} + C\|P_h^{\mathrm{RT}}\hat{I}_h u_h^{n,i}\cdot\nabla u_h^{n,i}\|_{L^2(\Omega)}. \tag{3.74}$$

Substituting (3.74) into (3.73) and using Young's inequality, we have

$$\begin{aligned}\|P_h^{\mathrm{RT}}\hat{I}_h u_h^{n,i}\cdot\nabla u_h^{n,i}\|_{L^2(\Omega)} \leqslant& C\|\hat{I}_h u_h^{n,i}\|_{L^2(\Omega)}^{1/2}\|\hat{I}_h u_h^{n,i}\|_{H^1(\Omega)}^{1/2}\|u_h^{n,i}\|_{H^1(\Omega)}^{1/2}\|\dot{u}_h^{n,i}\|_{L^2(\Omega)}^{1/2}\\ &+ C\|\hat{I}_h u_h^{n,i}\|_{L^2(\Omega)}\|\hat{I}_h u_h^{n,i}\|_{H^1(\Omega)}\|u_h^{n,i}\|_{H^1(\Omega)}.\end{aligned} \tag{3.75}$$

Now we substitute this estimate into (3.72) and absorb $\|\dot{u}_h^{n,i}\|_{L^2(\Omega)}$ on the right-hand side by using Young's inequality. Following from these steps and the definition of the extrapolation operator $\hat{I}_h$, we obtain for $n \geqslant 3$

$$\begin{aligned}&\frac{1}{2}\tau_{n+1}\sum_{i=1}^{2} b_i\|\dot{u}_h^{n,i}\|_{L^2(\Omega)}^2 + \|\nabla u_h^{n+1}\|_{L^2(\Omega)}^2 - \|\nabla u_h^{n}\|_{L^2(\Omega)}^2\\ \leqslant& C\tau_{n+1}\sum_{i=1}^{2}\|\hat{I}_h u_h^{n,i}\|_{L^2(\Omega)}^2\left(\|u_h^{n-1}\|_{H^1(\Omega)}^2 + \|u_h^{n}\|_{H^1(\Omega)}^2\right)\|u_h^{n,i}\|_{H^1(\Omega)}^2.\end{aligned}$$

Due to the $L^2$ boundedness of $\hat{I}_h u_h^{n,i}$, we can Multiply $t_{n+1}$ on both sides of the above estimate and obtain

$$\begin{aligned}
&\frac{1}{2}t_{n+1}\tau_{n+1}\sum_{i=1}^{2} b_i\|\dot{u}_h^{n,i}\|_{L^2(\Omega)}^2 + t_{n+1}\|\nabla u_h^{n+1}\|_{L^2(\Omega)}^2 - t_n\|\nabla u_h^n\|_{L^2(\Omega)}^2 \\
\leqslant&\tau_{n+1}\|\nabla u_h^n\|_{L^2(\Omega)}^2 + C\Big(t_{n-1}\|u_h^{n-1}\|_{H^1(\Omega)}^2 + t_n\|u_h^n\|_{H^1(\Omega)}^2\Big)\tau_{n+1}\sum_{i=1}^{2}\|u_h^{n,i}\|_{H^1(\Omega)}^2 \\
&+ C\Big((\tau_n+\tau_{n+1})\|u_h^{n-1}\|_{H^1(\Omega)}^2 + \tau_{n+1}\|u_h^n\|_{H^1(\Omega)}^2\Big)\tau_{n+1}\sum_{i=1}^{2}\|u_h^{n,i}\|_{H^1(\Omega)}^2. \qquad (3.76)
\end{aligned}$$

From (3.60), we have that

$$2\sum_{n=0}^{m}\tau_{n+1}\sum_{i=1}^{2} b_i\|\nabla u_h^{n,i}\|_{L^2(\Omega)}^2 \leqslant \|u_h^0\|_{L^2(\Omega)}^2. \qquad (3.77)$$

Since $\tau_{n-1}\sim\tau_n\sim\tau_{n+1}$ and $t_3\sim\tau_3$, we can sum up (3.76) with respect to $n$ from 3 to $m$ and obtain the following inequality in combination with (3.77)

$$\begin{aligned}
&\frac{1}{2}\sum_{n=3}^{m} t_{n+1}\tau_{n+1}\sum_{i=1}^{2} b_i\|\dot{u}_h^{n,i}\|_{L^2(\Omega)}^2 + t_{m+1}\|\nabla u_h^{m+1}\|_{L^2(\Omega)}^2 \\
\leqslant&C + C\sum_{n=3}^{m}\Big(t_{n-1}\|u_h^{n-1}\|_{H^1(\Omega)}^2 + t_n\|u_h^n\|_{H^1(\Omega)}^2\Big)\tau_{n+1}\sum_{i=1}^{2}\|u_h^{n,i}\|_{H^1(\Omega)}^2
\end{aligned}$$

By using discrete Gronwall's inequality and (3.77), we then obtain that

$$\frac{1}{2}\sum_{n=3}^{m} t_{n+1}\tau_{n+1}\sum_{i=1}^{2} b_i\|\dot{u}_h^{n,i}\|_{L^2(\Omega)}^2 + t_{m+1}\|\nabla u_h^{m+1}\|_{L^2(\Omega)}^2 \leqslant C. \qquad (3.78)$$

Based on (3.65), we have

$$\begin{aligned}
\|\nabla u_h^{n,i}\|_{L^2(\Omega)}^2 =& -(\dot{u}_h^{n,i}, u_h^{n,i}) = (\dot{u}_h^{n,i}, u_h^n + \tau_{n+1}\sum_{j=1}^{2} a_{ij}\dot{u}_h^{n,j}) \\
\leqslant&\|\dot{u}_h^{n,i}\|_{H^{-1}(\Omega)}\|u_h^n\|_{H^1(\Omega)} + C\tau_{n+1}\|\dot{u}_h^{n,i}\|_{L^2(\Omega)}\sum_{j=1}^{2}\|\dot{u}_h^{n,j}\|_{L^2(\Omega)}. \qquad (3.79)
\end{aligned}$$

It follows from (3.62a) and (3.28) that

$$\begin{aligned}
\|\dot{u}_h^{n,i}\|_{H^{-1}(\Omega)} \leqslant& C\|u_h^{n,i}\|_{H^1(\Omega)} + C\|P_h^{\mathrm{RT}}\hat{I}_h u_h^{n,i}\otimes u_h^{n,i}\|_{L^2(\Omega)} \\
\leqslant& C\|u_h^{n,i}\|_{H^1(\Omega)} + C\|\hat{I}_h u_h^{n,i}\|_{L^2(\Omega)}^{1/2}\|\hat{I}_h u_h^{n,i}\|_{H^1(\Omega)}^{1/2}\|u_h^{n,i}\|_{L^2(\Omega)}^{1/2}\|u_h^{n,i}\|_{H^1(\Omega)}^{1/2}. \qquad (3.80)
\end{aligned}$$

Substituting (3.80) into (3.79), it gives

$$\begin{aligned}
&\|\nabla u_h^{n,i}\|_{L^2(\Omega)}^2\\
\leqslant& C\|u_h^n\|_{H^1(\Omega)}\|u_h^{n,i}\|_{H^1(\Omega)} + C\|u_h^n\|_{H^1(\Omega)}\|\hat{I}_h u_h^{n,i}\|_{H^1(\Omega)}^{1/2}\|u_h^{n,i}\|_{H^1(\Omega)}^{1/2} + C\tau_{n+1}\sum_{j=1}^{2}\|\dot{u}_h^{n,j}\|_{L^2(\Omega)}^2\\
\leqslant& Ct_n^{-1/2}\|u_h^{n,i}\|_{H^1(\Omega)} + Ct_n^{-3/4}\|u_h^{n,i}\|_{H^1(\Omega)}^{1/2} + C\tau_{n+1}\sum_{j=1}^{2}\|\dot{u}_h^{n,j}\|_{L^2(\Omega)}^2\\
\leqslant& \delta\|u_h^{n,i}\|_{H^1(\Omega)}^2 + Ct_{n+1}^{-1} + C\tau_{n+1}\sum_{j=1}^{2}\|\dot{u}_h^{n,j}\|_{L^2(\Omega)}^{2,}
\end{aligned}$$

where $\delta > 0$ is a sufficiently small number so that the first term can be absorbed by the left-hand side. Combining (3.78), we obtain the desired estimate. □

**3.3.3 Error analysis.** The following lemma gives an a priori error bound for the time discretization.

**Lemma 3.7.** *Let $u_h(t_{n+1})$ be the solution to the semidiscrete scheme* (3.22)*. If the time stepsizes satisfy* (3.47) *with a fixed constant $\alpha \in \left(\frac{3}{4}, 1\right)$ and $u_h^{n+1}$ is the solution to the fully discrete scheme* (3.53)*, the error $\eta^{n+1} := u_h^{n+1} - u_h(t_{n+1})$ satisfies the following error bound*

$$\Big(\sum_{l=0}^{n}\tau_{l+1}\sum_{i=1}^{2}b_i\|\eta^{l,i}\|_{L^2(\Omega)}^2\Big)^{1/2} + \|\eta^{n+1}\|_{H^{-1}(\Omega)} \leqslant Ct_{n+1}^{-3/2-\varepsilon}\tau_{n+1}^2. \tag{3.81}$$

*Here $\varepsilon \in (0, 2\alpha - 3/2)$ could be arbitrarily small.*

*Proof.* Multiplying $-A_h^{-1}$ on both sides of the third relation in (3.56) and testing with $\eta^{n+1}$, we obtain

$$\begin{aligned}
&\|\eta^{n+1}\|_{H^{-1}(\Omega)}^2 - \|\eta^n\|_{H^{-1}(\Omega)}^2\\
=&2\tau_{n+1}\sum_{i=1}^{2}b_i(-A_h^{-1}\dot{\eta}^{n,i}, \eta^n) + 2\Big(A_h^{-1}Q_{n+1}(\partial_t u_h), \eta^n\Big)\\
&+ 2\tau_{n+1}\sum_{i=1}^{2}b_i(A_h^{-1}\mathcal{G}^{n,i}, \eta^n) - \tau_{n+1}^2\sum_{i,j=1}^{2}b_i b_j(A_h^{-1}\dot{\eta}^{n,i}, \dot{\eta}^{n,j}) + \|Q_{n+1}(\partial_t u_h)\|_{H^{-1}(\Omega)}^2\\
&- 2\tau_{n+1}^2\sum_{i,j=1}^{2}b_i b_j(-A_h^{-1}\dot{\eta}^{n,i}, \mathcal{G}^{n,i}) - 2\tau_{n+1}\sum_{i=1}^{2}b_i\Big(-A_h^{-1}\dot{\eta}^{n,i}, Q_{n+1}(\partial_t u_h)\Big)\\
&+ \tau_{n+1}^2\sum_{i,j=1}^{2}b_i b_j(-A_h^{-1}\mathcal{G}^{n,i}, \mathcal{G}^{n,j}) + 2\tau_{n+1}\sum_{i=1}^{2}b_i\Big(-A_h^{-1}\mathcal{G}^{n,i}, Q_{n+1}(\partial_t u_h)\Big).
\end{aligned}$$

Next the second relation in (3.56) and the algebraical stability of the Gauss–Lobatto IIIC lead to

$$\begin{aligned}
&\|\eta^{n+1}\|^2_{H^{-1}(\Omega)} - \|\eta^{n}\|^2_{H^{-1}(\Omega)}\\
\leqslant& 2\tau_{n+1}\sum_{i=1}^{2} b_i\Big( - A_h^{-1}\dot\eta^{n,i}, \eta^{n,i} + \tau_{n+1}\sum_{j=1}^{2} a_{ij}\mathcal{G}^{n,j} + Q_{n,i}(\partial_t u_h)\Big)\\
&- 2\tau_{n+1}\sum_{i=1}^{2} b_i(-A_h^{-1}\mathcal{G}^{n,i}, \eta^n) + 2\Big(A_h^{-1}Q_{n+1}(\partial_t u_h), \eta^n\Big) + \|Q_{n+1}(\partial_t u_h)\|^2_{H^{-1}(\Omega)}\\
&- 2\tau_{n+1}^2\sum_{i,j=1}^{2} b_i b_j(-A_h^{-1}\dot\eta^{n,i}, \mathcal{G}^{n,j}) - 2\tau_{n+1}\sum_{i=1}^{2} b_i\Big( - A_h^{-1}\dot\eta^{n,i}, Q_{n+1}(\partial_t u_h)\Big)\\
&+ \tau_{n+1}^2\sum_{i,j=1}^{2} b_i b_j(-A_h^{-1}\mathcal{G}^{n,i}, \mathcal{G}^{n,j}) + 2\tau_{n+1}\sum_{i=1}^{2} b_i\Big( - A_h^{-1}\mathcal{G}^{n,i}, Q_{n+1}(\partial_t u_h)\Big).
\end{aligned}\tag{3.82}$$

It follows from the first relation of (3.56) that $-A_h^{-1}\dot\eta^{n,i} = -\eta^{n,i} - A_h^{-1}T^{n,i}$. Then we obtain

$$\begin{aligned}
&2\tau_{n+1}\sum_{i=1}^{2} b_i\|\eta^{n,i}\|^2_{L^2(\Omega)} + \|\eta^{n+1}\|^2_{H^{-1}(\Omega)} - \|\eta^{n}\|^2_{H^{-1}(\Omega)}\\
\leqslant& 2\tau_{n+1}\sum_{i=1}^{2} b_i(-A_h^{-1}T^{n,i}, \eta^{n,i}) - 2\tau_{n+1}\sum_{i=1}^{2} b_i\Big(\eta^{n,i} + A_h^{-1}T^{n,i}, \tau_{n+1}\sum_{j=1}^{2} a_{ij}\mathcal{G}^{n,j} + Q_{n,i}(\partial_t u_h)\Big)\\
&- 2\tau_{n+1}\sum_{i=1}^{2} b_i(-A_h^{-1}\mathcal{G}^{n,i}, \eta^n) + 2\Big(A_h^{-1}Q_{n+1}(\partial_t u_h), \eta^n\Big) + \|Q_{n+1}(\partial_t u_h)\|^2_{H^{-1}(\Omega)}\\
&+ 2\tau_{n+1}^2\sum_{i,j=1}^{2} b_i b_j(\eta^{n,i} + A_h^{-1}T^{n,i}, \mathcal{G}^{n,j}) + 2\tau_{n+1}\sum_{i=1}^{2} b_i\Big(\eta^{n,i} + A_h^{-1}T^{n,i}, Q_{n+1}(\partial_t u_h)\Big)\\
&+ \tau_{n+1}^2\sum_{i,j=1}^{2} b_i b_j(-A_h^{-1}\mathcal{G}^{n,i}, \mathcal{G}^{n,j}) + 2\tau_{n+1}\sum_{i=1}^{2} b_i\Big( - A_h^{-1}\mathcal{G}^{n,i}, Q_{n+1}(\partial_t u_h)\Big).
\end{aligned}\tag{3.83}$$

We estimate the terms on the right-hand side of (3.83) subsequently. For $n \geqslant 2$, the first term can be bounded by

$$\begin{aligned}
&2\tau_{n+1}\sum_{i=1}^{2} b_i(-A_h^{-1}T^{n,i}, \eta^{n,i})\\
\leqslant& C\tau_{n+1}\sum_{i=1}^{2}\|\eta^{n,i}\|_{W^{-1,4}(\Omega)}\|T^{n,i}\|_{W^{-1,4/3}(\Omega)}\\
\leqslant& C\tau_{n+1}\sum_{i=1}^{2}\|\eta^{n,i}\|_{W^{-1,4}(\Omega)}\Big(\|P_h^{\mathrm{RT}}\hat I_h\eta^{n,i}\otimes u_h^{n,i}\|_{L^{4/3}(\Omega)} + \|P_h^{\mathrm{RT}}\hat I_h u_h(t_{n,i})\otimes\eta^{n,i}\|_{L^{4/3}(\Omega)}\Big)
\end{aligned}$$

$$\leqslant C\tau_{n+1}\sum_{i=1}^{2}\|\eta^{n,i}\|_{H^{-1}(\Omega)}^{1/2}\|\eta^{n,i}\|_{L^2(\Omega)}^{1/2}\Big(\|\hat{I}_h\eta^{n,i}\|_{L^2(\Omega)}\|u_h^{n,i}\|_{L^4(\Omega)} + \|P_h^{\mathrm{RT}}\hat{I}_h u_h(t_{n,i})\|_{L^4(\Omega)}\|\eta^{n,i}\|_{L^2(\Omega)}\Big). \tag{3.84}$$

By testing the second relation in (3.56) with $-A_h^{-1}\eta^{n,i}$ and using the first relation in (3.56), we obtain

$$\begin{aligned}\|\eta^{n,i}\|_{H^{-1}(\Omega)}^2 \leqslant& C\|\eta^n\|_{H^{-1}(\Omega)}^2 + C\tau_{n+1}^2\sum_{j=1}^{2}\Big(\|T^{n,j}\|_{H^{-1}(\Omega)}^2 + \|\mathcal{G}^{n,j}\|_{H^{-1}(\Omega)}^2\Big)\\ &+ C\tau_{n+1}\sum_{i=1}^{2}\|\eta^{n,j}\|_{L^2(\Omega)}^2 + C\|Q_{n,i}(\partial_t u_h)\|_{H^{-1}(\Omega)}^2.\end{aligned} \tag{3.85}$$

By substituting (3.85) into (3.84) and utilizing Hölder's inequality together the imbedding $H^{1/2}(\Omega)\hookrightarrow L^4(\Omega)$, we obtain

$$\begin{aligned}&2\tau_{n+1}\sum_{i=1}^{2}b_i(-A_h^{-1}T^{n,i},\eta^{n,i})\\ \leqslant&\delta\tau_{n+1}\sum_{i=1}^{2}b_i\Big(\|\eta^{n,i}\|_{L^2(\Omega)}^2 + \|\hat{I}_h\eta^{n,i}\|_{L^2(\Omega)}^2\Big) + C\tau_{n+1}^{3/2}\sum_{i=1}^{2}\|\eta^{n,i}\|_{L^2(\Omega)}^2\|u_h^{n,i}\|_{L^4(\Omega)}^2\\ &+ C_\delta\tau_{n+1}\sum_{i=1}^{2}\|\eta^n\|_{H^{-1}(\Omega)}^2\Big(\|u_h^{n,i}\|_{L^4(\Omega)}^4 + \|P_h^{\mathrm{RT}}\hat{I}_h u_h(t_{n,i})\|_{L^4(\Omega)}^4\Big)\\ &+ C_\delta\tau_{n+1}^{5/4}\sum_{i=1}^{2}\|\eta^{n,i}\|_{L^2(\Omega)}^2\|P_h^{\mathrm{RT}}\hat{I}_h u_h(t_{n,i})\|_{L^4(\Omega)} + C_\delta\tau_{n+1}\sum_{i=1}^{2}\|Q_{n,i}(\partial_t u_h)\|_{H^{-1}(\Omega)}^2\\ &+ C_\delta\tau_{n+1}^{3/2}\sum_{j=1}^{2}\Big(\|T^{n,j}\|_{H^{-1}(\Omega)}^{1/2} + \|\mathcal{G}^{n,j}\|_{H^{-1}(\Omega)}^{1/2}\Big)\\ &\cdot\sum_{i=1}^{2}\|\eta^{n,i}\|_{L^2(\Omega)}^{1/2}\Big(\|\hat{I}_h\eta^{n,i}\|_{L^2(\Omega)}\|u_h^{n,i}\|_{L^4(\Omega)} + \|P_h^{\mathrm{RT}}\hat{I}_h u_h(t_{n,i})\|_{L^4(\Omega)}\|\eta^{n,i}\|_{L^2(\Omega)}\Big).\end{aligned} \tag{3.86}$$

By using Lemma 3.4, Lemma 3.6, the $L^\infty$ estimate (3.45) of $u_h$, the $L^\infty$ stability estimate (3.28) and the interpolation inequality, we have

$$\begin{aligned}\|T^{n,j}\|_{H^{-1}(\Omega)} \leqslant&\|P_h^{\mathrm{RT}}\hat{I}_h\eta^{n,j}\otimes(\eta^{n,j}+u_h(t_{n,j}))\|_{L^2(\Omega)} + \|P_h^{\mathrm{RT}}\hat{I}_h u_h(t_{n,j})\otimes\eta^{n,j}\|_{L^2(\Omega)}\\ \leqslant&C\|\hat{I}_h\eta^{n,j}\|_{L^2(\Omega)}^{1/2}\|\hat{I}_h\eta^{n,j}\|_{H^1(\Omega)}^{1/2}\|\eta^{n,j}\|_{L^2(\Omega)}^{1/2}\|\eta^{n,j}\|_{H^1(\Omega)}^{1/2}\\ &+\|\hat{I}_h\eta^{n,j}\|_{L^2(\Omega)}\|u_h(t_{n,j})\|_{L^\infty(\Omega)}\end{aligned}$$

$$
\begin{aligned}
&+C\|\eta^{n,j}\|_{L^2(\Omega)}\Big(\|\hat{I}_h u_h(t_{n,j})\|_{L^\infty(\Omega)}+\|\hat{I}_h u_h(t_{n,j})\|_{H^1(\Omega)}\Big)\\
\leqslant& Ct_{n+1}^{-1/2}\|\hat{I}_h\eta^{n,j}\|_{L^2(\Omega)}^{1/2}\|\eta^{n,j}\|_{L^2(\Omega)}^{1/2}+Ct_{n+1}^{-1/2}\Big(\|\hat{I}_h\eta^{n,j}\|_{L^2(\Omega)}+\|\eta^{n,j}\|_{L^2(\Omega)}\Big),
\end{aligned}
\tag{3.87}
$$

By substituting (3.87) and (3.59) into (3.86), and using Corollary 3.1 and estimate (3.50), we obtain

$$
\begin{aligned}
&2\tau_{n+1}\sum_{i=1}^{2}b_i(-A_h^{-1}T^{n,i},\eta^{n,i})\\
\leqslant&\delta\tau_{n+1}\sum_{i=1}^{2}b_i\Big(\|\eta^{n,i}\|_{L^2(\Omega)}^2+\|\hat{I}_h\eta^{n,i}\|_{L^2(\Omega)}^2\Big)+C_\delta\tau_{n+1}^{3/2}t_{n+1}^{-1/2}\sum_{i=1}^{2}\Big(\|\eta^{n,i}\|_{L^2(\Omega)}^2+\|\hat{I}_h\eta^{n,i}\|_{L^2(\Omega)}^2\Big)\\
&+C_\delta\tau_{n+1}\sum_{i=1}^{2}\|\eta^{n}\|_{H^{-1}(\Omega)}^2\Big(\|u_h^{n,i}\|_{H^1(\Omega)}^2+\|\hat{I}_h u_h(t_{n,i})\|_{H^1(\Omega)}^2\Big)\\
&+C_\delta\tau_{n+1}^{5/4}t_{n+1}^{-1/4}\sum_{i=1}^{2}\|\eta^{n,i}\|_{L^2(\Omega)}^2+C_\delta\tau_{n+1}^{5}t_{n+1}^{-3}+C_\delta\tau_{n+1}^{7}t_{n+1}^{-6}.
\end{aligned}
\tag{3.88}
$$

Substituting (3.88) into (3.83) together with estimates (3.50), (3.50), (3.59), (3.58) and (3.87), we obtain

$$
\begin{aligned}
&2\tau_{n+1}\sum_{i=1}^{2}b_i\|\eta^{n,i}\|_{L^2(\Omega)}^2+\|\eta^{n+1}\|_{H^{-1}(\Omega)}^2-\|\eta^{n}\|_{H^{-1}(\Omega)}^2\\
\leqslant&\delta\tau_{n+1}\sum_{i=1}^{2}b_i\Big(\|\eta^{n,i}\|_{L^2(\Omega)}^2+\|\hat{I}_h\eta^{n,i}\|_{L^2(\Omega)}^2\Big)+C_\delta\tau_{n+1}^{3/2}t_{n+1}^{-1/2}\sum_{i=1}^{2}\Big(\|\eta^{n,i}\|_{L^2(\Omega)}^2+\|\hat{I}_h\eta^{n,i}\|_{L^2(\Omega)}^2\Big)\\
&+C_\delta\tau_{n+1}\sum_{i=1}^{2}\|\eta^{n}\|_{H^{-1}(\Omega)}^2\Big(\|u_h^{n,i}\|_{H^1(\Omega)}^2+\|\hat{I}_h u_h(t_{n,i})\|_{H^1(\Omega)}^2\Big)+C_\delta\tau_{n+1}^{5/4}t_{n+1}^{-1/4}\sum_{i=1}^{2}\|\eta^{n,i}\|_{L^2(\Omega)}^2\\
&+C_\delta\tau_{n+1}^{3}t_{n+1}^{-5/2}\|\eta^{n}\|_{H^{-1}(\Omega)}+C\tau_{n+1}^{5}t_{n+1}^{-3}+C_\delta\tau_{n+1}^{6}t_{n+1}^{-5}+C\tau_{n+1}^{7}t_{n+1}^{-6}
\end{aligned}
\tag{3.89}
$$

$$
\begin{aligned}
\leqslant&\delta\tau_{n+1}\sum_{i=1}^{2}b_i\Big(\|\eta^{n,i}\|_{L^2(\Omega)}^2+\|\hat{I}_h\eta^{n,i}\|_{L^2(\Omega)}^2\Big)+C_\delta\tau_{n+1}^{3/2}t_{n+1}^{-1/2}\sum_{i=1}^{2}\Big(\|\eta^{n,i}\|_{L^2(\Omega)}^2+\|\hat{I}_h\eta^{n,i}\|_{L^2(\Omega)}^2\Big)\\
&+C_\delta\tau_{n+1}\sum_{i=1}^{2}\|\eta^{n}\|_{H^{-1}(\Omega)}^2\Big(\|u_h^{n,i}\|_{H^1(\Omega)}^2+\|\hat{I}_h u_h(t_{n,i})\|_{H^1(\Omega)}^2\Big)+C_\delta\tau_{n+1}^{5/4}t_{n+1}^{-1/4}\sum_{i=1}^{2}\|\eta^{n,i}\|_{L^2(\Omega)}^2\\
&+C_\delta\tau_{n+1}t_{n+1}^{-1+2\varepsilon}\|\eta^{n}\|_{H^{-1}(\Omega)}^2+C_\delta\tau_{n+1}^{5}t_{n+1}^{-4-2\varepsilon},
\end{aligned}
\tag{3.90}
$$

where $\varepsilon\in(0,2\alpha-3/2)$. When $\delta$ is sufficiently small, $n$ is larger than a fixed integer $N^*$ such that $\tau_{n+1}^{1/4}t_{n+1}^{-1/4}$ and $\tau_{n+1}^{1/2}t_{n+1}^{-1/2}$ are sufficiently small. Then the corresponding terms on

the right-hand side can be absorbed by the left-hand side. Summing up (3.90) for $n \geqslant N^*$ and utilizing Gronwall's inequality give that

$$\begin{aligned}&\sum_{l=N^*}^{n} \tau_{l+1} \sum_{i=1}^{2} b_i \|\eta^{l,i}\|^2_{L^2(\Omega)} + \|\eta^{n+1}\|^2_{H^{-1}(\Omega)} \\ \leqslant & C\|\eta^{N^*}\|^2_{H^{-1}(\Omega)} + C\tau_{N^*} \sum_{i=1}^{2} \|\hat{I}_h \eta^{N^*,i}\|^2_{L^2(\Omega)} + C \sum_{l=N^*}^{n} \tau_{l+1}^5 t_{l+1}^{-4-2\varepsilon}. \end{aligned} \tag{3.91}$$

Since $\alpha \in (3/4, 1)$, we can deduce the following inequality by utilizing (3.47)

$$\sum_{l=0}^{n} \tau_{l+1}^5 t_{l+1}^{-4-2\varepsilon} \leqslant C\tau^4 \sum_{l=N^*}^{n} \tau_{l+1} t_{l+1}^{4\alpha-4-2\varepsilon} \leqslant C\tau^4 t_{n+1}^{4\alpha-3-2\varepsilon} = C\tau_{n+1}^4 t_{n+1}^{-3-2\varepsilon}. \tag{3.92}$$

Substituting (3.92) into (3.91), we obtain

$$\begin{aligned}&\sum_{l=N^*}^{n} \tau_{l+1} \sum_{i=1}^{2} b_i \|\eta^{l,i}\|^2_{L^2(\Omega)} + \|\eta^{n+1}\|^2_{H^{-1}(\Omega)} \\ \leqslant & C\|\eta^{N^*}\|^2_{H^{-1}(\Omega)} + C\tau_{N^*} \sum_{i=1}^{2} \|\hat{I}_h \eta^{N^*,i}\|^2_{L^2(\Omega)} + C\tau_{n+1}^4 t_{n+1}^{-3-2\varepsilon}. \end{aligned} \tag{3.93}$$

To finalize the proof, we need to estimate the error $\|\eta^n\|_{H^{-1}(\Omega)}$ for $1 \leqslant n < N^*$. The inequality (3.89) is valid for $0 \leqslant n < N^*$. Summing up (3.89) with respect to $0 \leqslant n < N^*$, the following inequality is then followed from the $L^2(\Omega)$ boundedness of $\eta^{n,i}$ and $\eta^n$

$$\begin{aligned}&2\sum_{l=0}^{n} \tau_{l+1} \sum_{i=1}^{2} b_i \|\eta^{l,i}\|^2_{L^2(\Omega)} + \|\eta^{n+1}\|^2_{H^{-1}(\Omega)} \\ \leqslant & CN^*\tau_{n+1} + C\sum_{l=0}^{n} \tau_{l+1} \sum_{i=1}^{2} \|\eta^l\|^2_{H^{-1}(\Omega)} \Big( \|u_h^{l,i}\|^2_{H^1(\Omega)} + \|\hat{I}_h u_h(t_{l,i})\|^2_{H^1(\Omega)} \Big) \\ & + C\sum_{l=1}^{n} \|\eta^l\|_{H^{-1}(\Omega)} \quad \text{for} \quad 1 \leqslant n < N^*. \end{aligned} \tag{3.94}$$

By using (3.92) and applying discrete Gronwall inequality, we obtain the result for $1 \leqslant n < N^*$

$$2\sum_{l=0}^{n} \tau_{l+1} \sum_{i=1}^{2} b_i \|\eta^{l,i}\|^2_{L^2(\Omega)} + \|\eta^{n+1}\|^2_{H^{-1}(\Omega)} \leqslant C\tau_{n+1}, \tag{3.95}$$

where the constant $C$ is dependent on the fixed constant $N^*$, but is independent on $n$. By using the property (2) specified at the beginning of section 3.3.1 and choosing $M_0 = N^*$, we have that

$$\tau_1 \sim \tau_2 \sim \cdots \sim \tau_{N^*} \sim \tau^{\frac{1}{1-\alpha}}, \tag{3.96}$$

when $\alpha \in (\frac{3}{4}, 1)$ and the equivalence depends on $N^*$. Then for $1 \leqslant n \leqslant N^*$, $\tau_n \leqslant t_n \leqslant N^*\tau_n$, which implies $t_n \sim \tau_n$. Therefore, we have

$$\tau_{n+1} \sim \tau_{n+1}^4 t_{n+1}^{-3} \leqslant C\tau^4 \tau_{n+1}^{-3-2\varepsilon} \quad \text{for} \quad 0 \leqslant n \leqslant N^* - 1. \tag{3.97}$$

Substituting (3.97) into (3.95), we obtain the desired result for $1 \leqslant n < N^*$ such that

$$2\sum_{l=0}^{n} \tau_{l+1} \sum_{i=1}^{2} b_i \|\eta^{l,i}\|_{L^2(\Omega)}^2 + \|\eta^{n+1}\|_{H^{-1}(\Omega)}^2 \leqslant C\tau_{n+1}^4 t_{n+1}^{-3-2\varepsilon}. \tag{3.98}$$

Then we can substitute (3.95) into (3.93) and obtain the following inequality for $n \geqslant N^*$ by using the $L^2$ boundedness of $\eta^{n,i}$ and $\eta^n$

$$\sum_{l=N^*}^{n} \tau_{l+1} \sum_{i=1}^{2} b_i \|\eta^{l,i}\|_{L^2(\Omega)}^2 + \|\eta^{n+1}\|_{H^{-1}(\Omega)}^2 \leqslant C\tau_{N^*} + C\tau_{n+1}^4 t_{n+1}^{-3-2\varepsilon}. \tag{3.99}$$

By adding up (3.99) and the inequality (3.95), we have

$$\sum_{l=0}^{n} \tau_{l+1} \sum_{i=1}^{2} b_i \|\eta^{l,i}\|_{L^2(\Omega)}^2 + \|\eta^{n+1}\|_{H^{-1}(\Omega)}^2 \leqslant C\tau_{N^*} + C\tau_{n+1}^4 t_{n+1}^{-3-2\varepsilon}, \quad \text{for} \quad n \geqslant N^*. \tag{3.100}$$

Again, by using the equivalence (3.96), we can derive that

$$\tau_{N^*} \sim \tau^{\frac{1}{1-\alpha}} = \tau^4 \tau^{\frac{4\alpha-3}{1-\alpha}} \sim \tau^4 \tau_1^{4\alpha-3} \leqslant \tau^4 t_{n+1}^{4\alpha-3} \sim \tau_{n+1}^4 t_{n+1}^{-3} \leqslant C\tau_{n+1}^4 t_{n+1}^{-3-2\varepsilon}. \tag{3.101}$$

Combining (3.98) and (3.100), (3.101), we complete the proof.

□

Using the $H^{-1}(\Omega)$ estimate of the errors proved in Lemma 3.7, we can derive the $L^2(\Omega)$ error bound, which is present in the following theorem.

**Theorem 3.2.** *Under the same conditions of $\alpha$ and $\varepsilon$ as Lemma 3.7, the error $\eta^{n+1} := u_h^{n+1} - u_h(t_{n+1})$ satisfies the following error bound*

$$\|\eta^{n+1}\|_{L^2(\Omega)} \leqslant Ct_{n+1}^{-2-\varepsilon}\tau_{n+1}^2. \tag{3.102}$$

*Proof.* Let $N^*$ be the fixed integer defined in Lemma 3.7. Then for $n < N^*$, (3.102) follows directly from the $L^2(\Omega)$ boundedness of $u_h^{n+1}$ and $u_h(t_{n+1})$. And it suffices to prove the desired result for $n \geqslant N^*$.

Squaring the third relation (3.56) on both sides, we obtain

$$\|\eta^{n+1}\|_{L^2(\Omega)}^2 = \|\eta^n\|_{L^2(\Omega)}^2 + 2\tau_{n+1}\sum_{i=1}^{2} b_i(\dot{\eta}^{n,i}, \eta^n) - 2\tau_{n+1}\sum_{i=1}^{2} b_i(\eta^n, \mathcal{G}^{n,i}) - 2\Big(\eta^n, Q_{n+1}(\partial_t u_h)\Big)$$

$$
\begin{aligned}
&+\tau_{n+1}^2\sum_{i,j=1}^{2} b_i b_j(\dot\eta^{n,i},\dot\eta^{n,j}) - 2\tau_{n+1}^2\sum_{i,j=1}^{2} b_i b_j(\dot\eta^{n,i},\mathcal{G}^{n,j})\\
&-2\tau_{n+1}\sum_{i=1}^{2} b_i\Big(\dot\eta^{n,i},Q_{n+1}(\partial_t u_h)\Big)+\tau_{n+1}^2\sum_{i,j=1}^{2} b_i b_j(\mathcal{G}^{n,i},\mathcal{G}^{n,j})\\
&+2\tau_{n+1}\sum_{i=1}^{2} b_i\Big(\mathcal{G}^{n,i},Q_{n+1}(\partial_t u_h)\Big)+\|Q_{n+1}(\partial_t u_h)\|_{L^2(\Omega)}^2. \qquad (3.103)
\end{aligned}
$$

Similarly to the deduction of (3.82), by representing $\eta^n$ using the second relation in (3.56) and utilizing the algebraical stability of Gauss–Lobatto IIIC, we obtain

$$
\begin{aligned}
\|\eta^{n+1}\|_{L^2(\Omega)}^2 \leqslant&\|\eta^n\|_{L^2(\Omega)}^2+2\tau_{n+1}\sum_{i=1}^{2} b_i\Big(\dot\eta^{n,i},\eta^{n,i}+\tau_{n+1}\sum_{j=1}^{2} a_{ij}\mathcal{G}^{n,j}+Q_{n,i}(\partial_t u_h)\Big)\\
&-2\tau_{n+1}\sum_{i=1}^{2} b_i(\eta^n,\mathcal{G}^{n,i})-2\Big(\eta^n,Q_{n+1}(\partial_t u_h)\Big)+\tau_{n+1}^2\sum_{i,j=1}^{2} b_i b_j(\mathcal{G}^{n,i},\mathcal{G}^{n,j})\\
&-2\tau_{n+1}\sum_{i=1}^{2} b_i\Big(\dot\eta^{n,i},Q_{n+1}(\partial_t u_h)\Big)+\|Q_{n+1}(\partial_t u_h)\|_{L^2(\Omega)}^2\\
&+2\tau_{n+1}\sum_{i=1}^{2} b_i\Big(\mathcal{G}^{n,i},Q_{n+1}(\partial_t u_h)\Big)-2\tau_{n+1}^2\sum_{i,j=1}^{2} b_i b_j(\dot\eta^{n,i},\mathcal{G}^{n,j}). \qquad (3.104)
\end{aligned}
$$

By the first relation in (3.56), we have

$$
\begin{aligned}
&2\tau_{n+1}\sum_{i=1}^{2} b_i\|\nabla\eta^{n,i}\|_{L^2(\Omega)}^2+\|\eta^{n+1}\|_{L^2(\Omega)}^2-\|\eta^n\|_{L^2(\Omega)}^2\\
\leqslant&C\tau_{n+1}^2\sum_{i,j=1}^{2}\|\dot\eta^{n,i}\|_{L^2(\Omega)}\|\mathcal{G}^{n,j}\|_{L^2(\Omega)}+\|Q_{n+1}(\partial_t u_h)\|_{L^2(\Omega)}^2\\
&+C\tau_{n+1}\sum_{i=1}^{2}\|T^{n,i}\|_{H^{-1}(\Omega)}\|\eta^{n,i}\|_{H^1(\Omega)}+C\tau_{n+1}\sum_{i=1}^{2}\|\eta^n\|_{L^2(\Omega)}\|\mathcal{G}^{n,i}\|_{L^2(\Omega)}\\
&+C\|\eta^n\|_{L^2(\Omega)}\|Q_{n+1}(\partial_t u_h)\|_{L^2(\Omega)}+C\tau_{n+1}^2\sum_{i=1}^{2}\|\mathcal{G}^{n,i}\|_{L^2(\Omega)}^2\\
&+C\tau_{n+1}\sum_{i=1}^{2}\Big(\|\eta^{n,i}\|_{H^1(\Omega)}+\|T^{n,i}\|_{H^{-1}(\Omega)}\Big)\Big(\|Q_{n,i}(\partial_t u_h)\|_{H^1(\Omega)}+\|Q_{n+1}(\partial_t u_h)\|_{H^1(\Omega)}\Big). \qquad (3.105)
\end{aligned}
$$

Suppose $\mathcal{O}^{-1} = (r_{ij})$. By the second relation in (3.56), we have

$$\dot{\eta}^{n,i} = \tau_{n+1}^{-1} \sum_{j=1}^{2} r_{ij}(\eta^{n,j} - \eta^n) + \mathcal{G}^{n,i} + \tau_{n+1}^{-1} \sum_{j=1}^{2} r_{ij} Q_{n,j}(\partial_t u_h).$$

Hence, we can derive the estimate of $\|\dot{\eta}^{n,i}\|_{L^2(\Omega)}$ such that

$$\|\dot{\eta}^{n,i}\|_{L^2(\Omega)} \leqslant C\tau_{n+1}^{-1} \sum_{j=1}^{2} \Big( \|\eta^{n,j}\|_{L^2(\Omega)} + \|\eta^n\|_{L^2(\Omega)} + \|Q_{n,j}(\partial_t u_h)\|_{L^2(\Omega)} \Big) + \|\mathcal{G}^{n,i}\|_{L^2(\Omega)}.$$

Substituting this estimate into (3.105), we have □

$$\begin{aligned}
&2\tau_{n+1} \sum_{i=1}^{2} b_i \|\nabla \eta^{n,i}\|_{L^2(\Omega)}^2 + \|\eta^{n+1}\|_{L^2(\Omega)}^2 - \|\eta^n\|_{L^2(\Omega)}^2 \\
\leqslant &\delta\tau_{n+1} \sum_{i=1}^{2} b_i \|\eta^{n,i}\|_{H^1(\Omega)}^2 + C\tau_{n+1} t_{n+1}^{-1} \Big( \|\eta^n\|_{L^2(\Omega)}^2 + \sum_{i=1}^{2} \|\eta^{n,i}\|_{L^2(\Omega)}^2 \Big) \\
&+ C\|Q_{n+1}(\partial_t u_h)\|_{L^2(\Omega)}^2 + C\delta^{-1}\tau_{n+1} \sum_{i=1}^{2} \Big( \|Q_{n,i}(\partial_t u_h)\|_{H^1(\Omega)} + \|Q_{n+1}(\partial_t u_h)\|_{H^1(\Omega)} \Big)^2 \\
&+ C\tau_{n+1} t_{n+1} \sum_{i=1}^{2} \|\mathcal{G}^{n,i}\|_{L^2(\Omega)}^2 + C\delta^{-1}\tau_{n+1} \sum_{i=1}^{2} \|T^{n,i}\|_{H^{-1}(\Omega)}^2. \qquad (3.106)
\end{aligned}$$

By using estimates (3.87), (3.50), (3.59) and (3.58), when $\delta$ is sufficiently small, we obtain

$$\begin{aligned}
&\tau_{n+1} \sum_{i=1}^{2} b_i \|\nabla \eta^{n,i}\|_{L^2(\Omega)}^2 + \|\eta^{n+1}\|_{L^2(\Omega)}^2 - \|\eta^n\|_{L^2(\Omega)}^2 \\
\leqslant &C\tau_{n+1} t_{n+1}^{-1} \Big( \|\eta^n\|_{L^2(\Omega)}^2 + \sum_{i=1}^{2} \|\eta^{n,i}\|_{L^2(\Omega)}^2 + \sum_{i=1}^{2} \|\hat{I}_h \eta^{n,i}\|_{L^2(\Omega)}^2 \Big) + C\tau_{n+1}^5 t_{n+1}^{-5}. \qquad (3.107)
\end{aligned}$$

Multiplying $t_{n+1}$ on both sides of (3.107), we have

$$\begin{aligned}
&t_{n+1}\tau_{n+1} \sum_{i=1}^{2} b_i \|\nabla \eta^{n,i}\|_{L^2(\Omega)}^2 + t_{n+1}\|\eta^{n+1}\|_{L^2(\Omega)}^2 - t_n \|\eta^n\|_{L^2(\Omega)}^2 \\
\leqslant &\tau_{n+1}\|\eta^n\|_{L^2(\Omega)}^2 + C\tau_{n+1} \Big( \|\eta^n\|_{L^2(\Omega)}^2 + \sum_{i=1}^{2} \|\eta^{n,i}\|_{L^2(\Omega)}^2 + \sum_{i=1}^{2} \|\hat{I}_h \eta^{n,i}\|_{L^2(\Omega)}^2 \Big) + C\tau_{n+1}^5 t_{n+1}^{-4}.
\end{aligned}$$

By summing up the above inequality with respect to $n$ from 0 to $m$ and following the proof similarly to (3.92), we have

$$\sum_{n=0}^{m} t_{n+1}\tau_{n+1} \sum_{i=1}^{2} b_i \|\nabla \eta^{n,i}\|_{L^2(\Omega)}^2 + t_{m+1}\|\eta^{m+1}\|_{L^2(\Omega)}^2$$

$$\leqslant C\sum_{n=0}^{m}\tau_{n+1}\Big(\|\eta^n\|^2_{L^2(\Omega)}+\sum_{i=1}^{2}\|\eta^{n,i}\|^2_{L^2(\Omega)}+\sum_{i=1}^{2}\|\hat{I}_h\eta^{n,i}\|^2_{L^2(\Omega)}\Big)+C\tau_{m+1}^4 t_{m+1}^{-3}.$$

Then the desired result follows from Lemma 3.7.

**3.4 Numerical examples.** In this section we present numerical examples to support the theoretical results in Theorem 3.1 and Theorem 3.2. All examples concern the incompressible NS problem in (3.1).

**Example 3.1** (The merging of two co-rotating vortices)**.** In this example, we investigate the simulation of the merging of two co-rotating Lamb-Oseen vortices within a two-dimensional domain $\Omega = (-\pi,\pi)\times(-\pi,\pi)$. The initial value of the standard Lamb-Oseen vortex [47, 86, 132, 144] is inherently a function in $L^p(\Omega)^2$ for $p<2$. To ensure that the initial value belongs to $L^2(\Omega)^2$ but not to $H^\varepsilon(\Omega)^2$, we make a slight modification to the data by selecting the initial value $u_0 = u_1 + u_2$ and

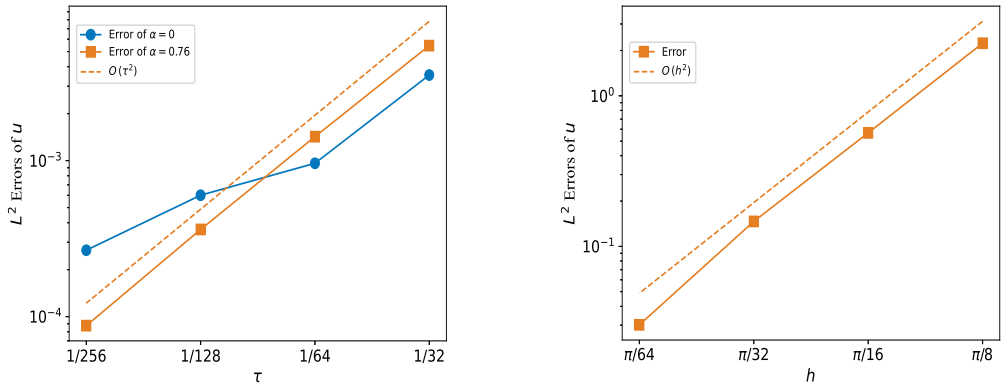
$$u_1 = \Big(-\frac{y\Gamma}{2\pi r_1^{2-\varepsilon}}, \frac{(x+0.5)\Gamma}{2\pi r_1^{2-\varepsilon}}\Big), \quad u_2 = \Big(-\frac{y\Gamma}{2\pi r_2^{2-\varepsilon}}, \frac{(x-0.5)\Gamma}{2\pi r_2^{2-\varepsilon}}\Big),$$

with $r_1 = \sqrt{(x+0.5)^2+y^2}$, $r_2 = \sqrt{(x-0.5)^2+y^2}$, and $\varepsilon = 0.1$. Here, $\Gamma$ denotes the circulation, set to $2\pi$ for this test. The viscosity $\nu$ is chosen as 0.1. We choose the domain $\Omega$ so large that we may assume that $u$ satisfies 0 Dirichlet boundary condition.

We test temporal convergence at $T = 0.1$ using graded stepsizes (3.47) with $\alpha = 0$ (uniform) and $\alpha = 0.76$. The reference solution $u^N_{h,\mathrm{ref}}$ is computed with $\tau = 1/1024$. Temporal errors $\|u^N_h - u^N_{h,\mathrm{ref}}\|_{L^2(\Omega)}$ in Figure 3.1 (a) for $\tau = 1/32, 1/64, 1/128, 1/256$ (spatial errors negligible for sufficiently small $h$) show second-order convergence for $\alpha = 0.76$ (consistent with Theorem 3.2) but irregular convergence for $\alpha = 0$, justifying the necessity of graded stepsizes in (3.47).

In Figure 3.1 (b), we present the spatial discretization errors $\|u^N_h - u^N_{h,\mathrm{ref}}\|_{L^2(\Omega)}$ and convergence rates for mesh sizes $h = \pi/8, \pi/16, \pi/32, \pi/64$ with a sufficiently small temporal stepsize that ensures the errors from temporal discretization are negligible. The reference solution $u^N_{h,\mathrm{ref}}$ is chosen to be the numerical solution with mesh size $h = \pi/128$. We use $P_2$–$P_1$ Taylor–Hood elements and observe that the convergence in space is second order. This aligns with the theoretical result proved in Theorem 3.1 and shows the sharpness of the convergence rate in space.

The evolution of the velocity $u$ for the co-rotating vortices is illustrated at various time instances, specifically at $t = 0.1, 0.3, 0.5, 0.7, 1.0, 2.0$. These visualizations are depicted in




(a) $L^2$ error of $u$ from temporal discretization

(b) $L^2$ error of $u$ from spatial discretization

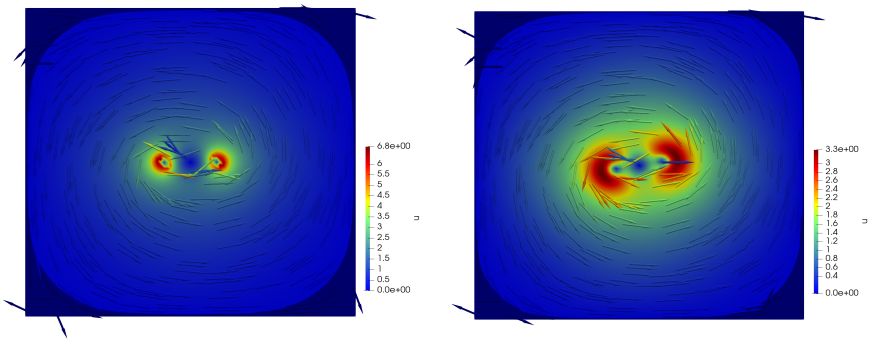
*Figure 3.1. $L^2$ errors of $u$*

Figure 3.2 (a)–(f) with mesh size $h = \pi/50$ and time stepsize $\tau = 0.005$. The parameter $\alpha$ is chosen as 0.76. The numerical simulation indicates a gradual merging of the two co-rotating vortices over time. Notably, at $t = 2.0$, the vortices have completely merged into a single vortex, as shown in Figure 3.2 (f).

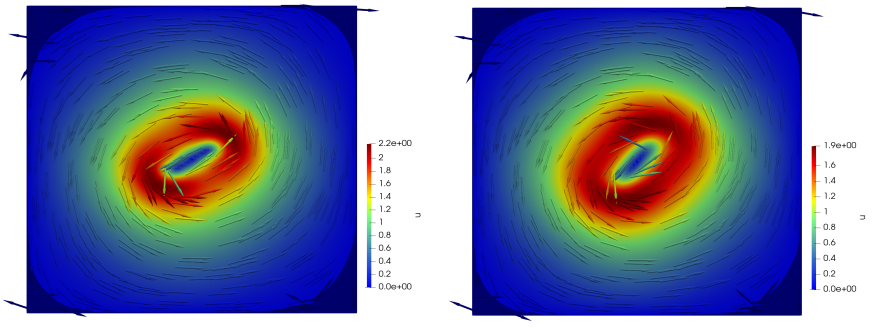
**Example 3.2** (Piecewise constant initial value)**.** In this example, we present numerical simulation of the Navier–Stokes equations with a piecewise constant initial value in $\Omega = (-\pi, \pi) \times (-\pi, \pi)$. The viscosity $\nu$ is chosen to be 0.1. The initial value $u_0$ takes value $(10, 0)$ when $y > 0$, and $(-10, 0)$ when $y < 0$. This initial value is in $\dot{L}^2(\Omega) \cap H^{\frac{1}{2}-\varepsilon}(\Omega)^2$ for any $\varepsilon \in (0, \frac{1}{2})$ but not in $H^{\frac{1}{2}}(\Omega)^2$.

We test temporal convergence at $T = 1$ using graded stepsizes (3.47) with $\alpha = 0.76$. The reference solution $u^N_{h,\mathrm{ref}}$ is computed with $\tau = 1/1024$. Temporal errors $\|u^N_h - u^N_{h,\mathrm{ref}}\|_{L^2(\Omega)}$ in Figure 3.3 (a) for $\tau = 1/32, 1/64, 1/128, 1/256$ (spatial errors negligible for sufficiently small $h$) show second-order convergence for $\alpha = 0.76$, which is consistent with Theorem 3.2.

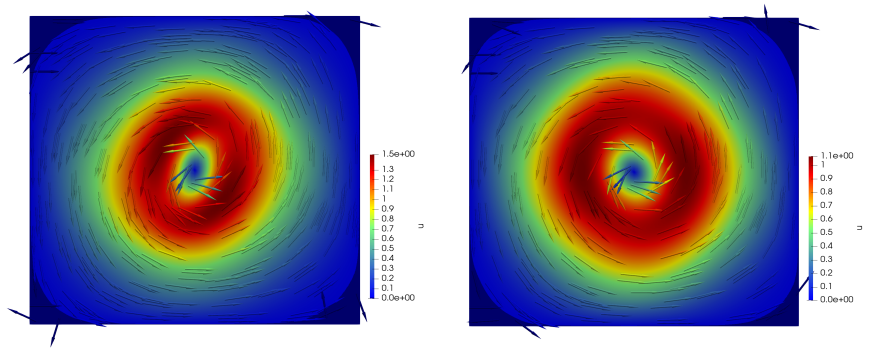
In Figure 3.3 (b), we present the spatial discretization errors $\|u^N_h - u^N_{h,\mathrm{ref}}\|_{L^2(\Omega)}$ and convergence rates for mesh sizes $h = 2\pi/15, 2\pi/30, 2\pi/60, 2\pi/120$ with a sufficiently small temporal stepsize that ensures the errors from temporal discretization are negligible. The reference solution $u^N_{h,\mathrm{ref}}$ is chosen to be the numerical solution with mesh size $h = 2\pi/240$. We use $P_2$–$P_1$ Taylor–Hood elements and observe that the convergence in space is second order. This aligns with the theoretical result proved in Theorem 3.1 and shows the sharpness




(a) The vorticity at time $t = 0.01$ (b) The vorticity at time $t = 0.1$

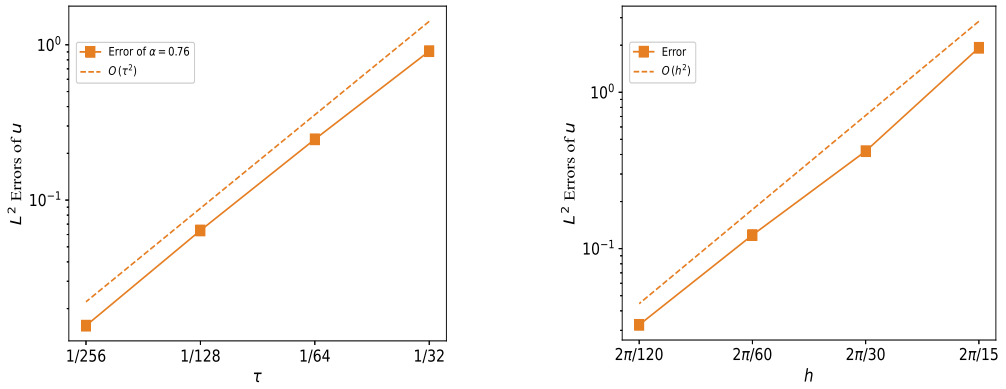



(c) The vorticity at time $t = 0.3$ (d) The vorticity at time $t = 0.5$




(e) The vorticity at time $t = 1.0$ (f) The vorticity at time $t = 2.0$

*Figure 3.2. Isocontours of the velocity u*



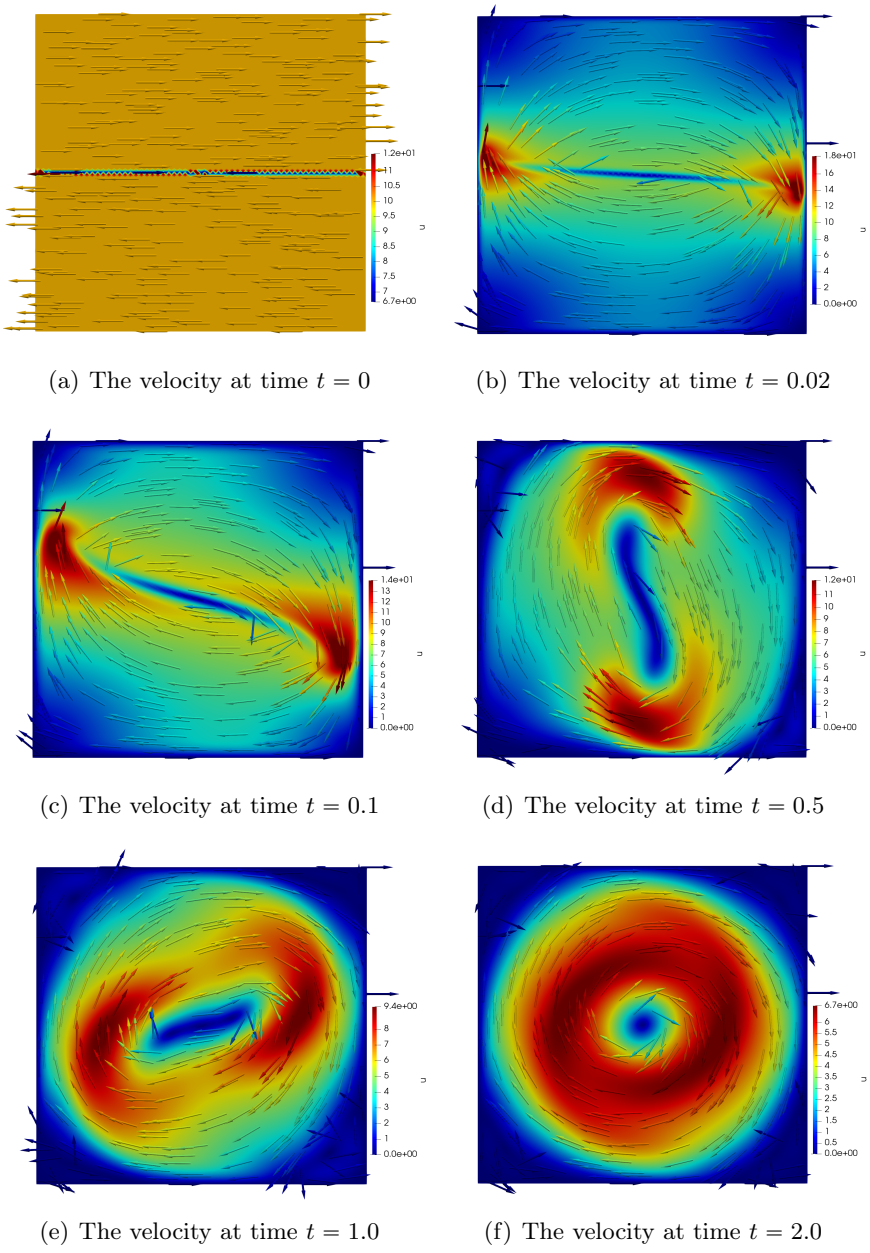

(a) $L^2$ error of $u$ from temporal discretization

(b) $L^2$ error of $u$ from spatial discretization

*Figure 3.3. $L^2$ errors of $u$*

of the convergence rate in space.

The evolution of the velocity field $u$ computed by the proposed method is illustrated at various time instances, specifically at $t = 0, 0.02, 0.1, 0.5, 1.0, 2.0$. These visualizations are depicted in Figure 3.4 with mesh size $h = 0.06$ and time stepsize $\tau = 0.01$. The parameter $\alpha$ is chosen to be 0.76. Notably, the discontinuous initial velocity field gradually becomes smooth as time evolves.




(a) The velocity at time $t = 0$

(b) The velocity at time $t = 0.02$

(c) The velocity at time $t = 0.1$

(d) The velocity at time $t = 0.5$

(e) The velocity at time $t = 1.0$

(f) The velocity at time $t = 2.0$

*Figure 3.4. Isocontours of the velocity $u$*

# Part II: High-order Methods for Parabolic Equations in Evolving Domains

## Brief Introduction

The second part of the thesis is concerned with the optimal-order convergence of the evolving FEM for shape gradient flow, Stokes equations and fluid–structure interaction (FSI) problems under $H^1$-norm and $L^2$-norm.

In Chapter 4, 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 $k \geqslant 2$ 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 PDEs is also discussed.

In Chapter 5, the numerical solution of the Stokes equations on an evolving domain with a moving boundary is studied based on the arbitrary Lagrangian–Eulerian (ALE) finite element method along the trajectories of the evolving mesh. The error of the semidiscrete ALE method is shown to be $O(h^{r+1})$ for velocity in $L^\infty(0,T;L^2)$ norm and $O(h^r)$ for pressure in $L^2(0,T;L^2)$ norm by employing the Taylor–Hood finite elements of degree $r \geqslant 2$, using Nitsche's duality argument adapted to an evolving mesh, by proving that the material derivative and the Stokes–Ritz projection commute up to terms which have optimal-order convergence in the $L^2$ norm. Numerical examples are provided to support the theoretical analysis.

In Chapter 6, the convergence of ALE–FEMs for fluid-structure interaction (FSI) problems with a solution-driven moving interface is studied. In addition to the finite element approximation errors, the geometric approximation errors due to the unknown interface motion and mesh evolvement are considered as well within the ALE frame. By adding an initial correction term to the numerical scheme, the optimal convergence rates of fluid velocity, structural displacement and ALE mesh motion in $L^\infty(0,T;H^1)$ norm, as well as of pressure in $L^\infty(0,T;L^2)$ norm, are proved for a fully discrete, monolithic FEM which tracks the unknown moving interface using the ALE approach. Numerical experiments in both two and three dimensions are presented to support the theoretical error analysis.

The content of this part consists of the research in the following papers of the author.

## Papers published & submitted

1. W. Gong, B. Li, and Q. Rao. Convergent Evolving Finite Element Approximations of Boundary Evolution under Shape Gradient Flow, *IMA J. Numer. Anal.*, 44(5):2667–

2697, 2024.

2. Q. Rao, J. Wang, and Y. Xie. Optimal Convergence of Arbitrary Lagrangian–Eulerian Finite Element Methods for the Stokes Equation in an Evolving Domain, *IMA J. Numer. Anal.*, drae097, 2025.

3. B. Li, Q. Rao, and P. Sun. Optimal Convergence of an Arbitrary Lagrangian–Eulerian Finite Element Method for Fluid-Structure Interactions, submitted for publication.

# Chapter 4

## Convergent Evolving Finite Element Approximations of Boundary Evolution under Shape Gradient Flow

The content of this chapter has been published in "*W. Gong, B. Li, and Q. Rao. Convergent Evolving Finite Element Approximations of Boundary Evolution under Shape Gradient Flow, IMA J. Numer. Anal., 44(5):2667–2697, 2024.*"

**4.1 Introduction.** Shape optimization constrained by partial differential equations (PDEs) has wide applications in modern science and engineering, such as the airfoil designs in aerodynamics [135], automative industry [3, 88], turbomachinery, structural design [88], and so on. These applications typically concern minimizing a shape functional

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

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 $\Omega$ with boundary $\Gamma$. 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 [88, 81, 69, 32] and efficient computational methods [3, 135, 63, 62].

The boundary parametrization of an elliptic shape optimization problem was considered in [45], 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 [87], 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 [50]. 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 [26] for an elliptic PDE-constrained shape optimization problem in two dimensions based on the compactness argument.

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 [16, 30]. 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 [34] 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 [17], immersed FEM [60], isogeometric analysis [31, 172], adaptive FEM [136], coupling of FEM and BEM [44], ALE FEM [143] and evolving bulk FEMs with possibly curved triangles [41], 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 [59, 131] and [151, 12, 6, 41, 38, 53, 98, 141] 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 [41, 38, 113] for different isoparametric finite elements. We also refer to [42, 43, 96] 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 chapter 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. [3, 16])

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

for an abstract inner product $b(\cdot, \cdot) : H \times H \to \mathbb{R}$ associated with some Hilbert space $H$. This

is referred to as the Hilbertian extension-regularization approach in [3, 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. [153, 154]), and $H = H^{\frac{1}{2}}(\Gamma)^d$ yields Stefan-like flow. We refer to [16] for a comprehensive review of different shape gradient flows. The $L^2$ flow is generally irregular and makes sense only on the boundary $\Gamma$. 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 chapter, 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.,

$$\min_{\Gamma} J(\Gamma) = \int_{\Omega} j(x,u)\mathrm{d}x \quad \text{subject to} \quad \begin{cases} -\Delta u = f & \text{in } \Omega \subset \mathbb{R}^d, \text{ with } d \in \{2,3\}, \\ u = 0 & \text{on } \Gamma = \partial\Omega, \end{cases} \tag{4.1}$$

with the shape densify 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 (4.1), 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:

$$\partial_t \phi = w \circ \phi \quad \text{in } \Omega^0, \qquad \phi(0) = \mathrm{id}|_{\Omega^0} \quad \text{in } \Omega^0, \tag{4.2a}$$

$$-\Delta w + w = 0 \quad \text{in } \Omega(t), \qquad \partial_\nu w = -J'(\Gamma(t))\nu \quad \text{on } \Gamma(t), \tag{4.2b}$$

$$-\Delta u = f \quad \text{in } \Omega(t), \qquad u = 0 \quad \text{on } \Gamma(t), \tag{4.2c}$$

$$-\Delta p = j'_u(\cdot,u) \quad \text{in } \Omega(t), \qquad p = 0 \quad \text{on } \Gamma(t), \tag{4.2d}$$

where $\phi(\cdot,t) : \Omega^0 \to \Omega(t)$ 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 (4.9), the rate of change of the shape functional $J(\Gamma)$ satisfies the following relation:

$$\frac{\mathrm{d}J(\Gamma(t))}{\mathrm{d}t} = \int_{\Gamma(t)} J'(\Gamma(t))w(t)\cdot\nu \,\mathrm{d}\Gamma(t). \tag{4.3}$$

Therefore, by testing (4.2b) with $w$ and using relation (4.3), the following property can be shown:

$$\frac{\mathrm{d}J(\Gamma(t))}{\mathrm{d}t} = -\|\nabla w(t)\|^2_{L^2(\Omega(t))} - \|w(t)\|^2_{L^2(\Omega(t))} \leqslant 0, \tag{4.4}$$

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 chapter 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 chapter 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 $k \geqslant 2$ 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 [3, Sec. 5.3]. The stability and convergence analysis in this chapter 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

$$\begin{aligned} &\int_\Omega (\nabla w : \nabla v + wv)\,\mathrm{d}x - \int_\Omega q\,\nabla\cdot v\,\mathrm{d}x = -\int_\Gamma J'(\Gamma)v\cdot\nu\mathrm{d}\Gamma \quad \forall\, v \in H^1(\Omega), \\ &\int_\Omega \nabla\cdot w\,\eta\,\mathrm{d}x = 0 \qquad\qquad \forall\, \eta \in \mathbb{R}, \end{aligned} \tag{4.5}$$

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 chapter is organized as follows. In Section 4.2, we formulate (4.2) 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 4.3. Numerical tests are presented in Section 4.4 to illustrate the convergence of the numerical approximations and the evolution of the domain under shape gradient flow.

## 4.2 Notation and main results.

In this section we present the notation and main results of this chapter, 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.

### 4.2.1 Preliminaries of shape calculus.

In this subsection, we introduce some basic ingredients of the velocity method for shape calculus. We refer the readers to [32, 81] 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 4.1** ([171, Lemma 5.7]). *If the domain $\Omega$ moves with velocity $v \in W^{1,\infty}(\Omega)$, then*

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega f \mathrm{d}x = \int_\Omega (\partial_t^\bullet f + f\nabla\cdot v)\mathrm{d}x, \tag{4.6}$$

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

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

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{F}^t[v] = v \circ \mathcal{F}^t[v] \text{ on } \Gamma \text{ with initial condition } \mathcal{F}^0[v] = \mathrm{id}|_\Gamma. \tag{4.7}$$

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

$$\mathrm{d}J(\Gamma; v) := \frac{\mathrm{d}}{\mathrm{d}t} J(\mathcal{F}^t[v](\Gamma))\Big|_{t=0}. \tag{4.8}$$

It can be formulated as an integral on the boundary in terms of the shape gradient defined on the boundary $\Gamma$ (see [32, Chap. 9, Sec. 3.4]), i.e.,

$$\mathrm{d}J(\Gamma; v) = \int_\Gamma J'(\Gamma) v \cdot \nu \mathrm{d}\Gamma. \tag{4.9}$$

In fact, $J'(\Gamma)$ is defined as the function on $\Gamma$ satisfying relation (4.9).

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 $\Omega$ (called the *distributed Eulerian derivative*), as shown in the following lemma.

**Lemma 4.2** ([72, eq. (2.9)-(2.10)] and [57, eq. (2.6)]). *Let $u$ and $p$ be the solutions of the primal equation (4.2c) and the adjoint state equation (4.2d), respectively. Then the Eulerian derivative of the functional $J(\Gamma)$ with $j(\cdot, u) = \frac{1}{2}|u - u_d|^2$ defined in (4.1) has the following closed form:*

$$\begin{aligned}\mathrm{d}J(\Gamma; v) = \mathrm{d}J(\Gamma, u, p; v) := &\int_\Omega \nabla u \cdot (\nabla v + \nabla v^\top)\nabla p - f \nabla p \cdot v \mathrm{d}x \\ &+ \int_\Omega (\frac{1}{2}|u - u_d|^2 - \nabla u \cdot \nabla p)\nabla \cdot v - (u - u_d)\nabla u_d \cdot v \mathrm{d}x, \end{aligned} \tag{4.10}$$

*where $u$ and $p$ are determined by $\Gamma$ through the following equations:*

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

**4.2.2 Weak formulation of the PDEs.** For any flow map $\phi : \Omega^0 \times [0, T] \to \mathbb{R}^d$, we denote by $\Omega[\phi(\cdot, t)]$ the image of $\Omega^0$ under the map $\phi(\cdot, t)$. Since the boundary type Eulerian derivative in (4.9) and the distributed type Eulerian derivative in (4.10) are equivalent, (4.2b) can be equivalently written into the following weak formulation by using integration by part: Find $w \in H^1(\Omega)^d$ such that

$$\begin{aligned}\int_\Omega \big(\nabla w : \nabla v + w \cdot v\big) dx = -\int_\Gamma \partial_\nu w \cdot v \mathrm{d}\Gamma \; &= \int_\Gamma J'(\Gamma) v \cdot \nu \mathrm{d}\Gamma \\ &= \mathrm{d}J(\Gamma; v) = -\mathrm{d}J(\Gamma, u, p; v) \quad \forall\, v \in H^1(\Omega)^d, \end{aligned} \tag{4.11}$$

where the closed form of $\mathrm{d}J(\Gamma, u, p; v)$ is given by (4.10). Then the moving boundary problem in (4.2) can be written into the following weak formulation:

$$\partial_t \phi = w \circ \phi, \tag{4.12a}$$

$$\int_{\Omega[\phi(\cdot,t)]} \nabla w : \nabla \chi_w + w \cdot \chi_w \mathrm{d}x = -\mathrm{d}J(\Gamma(t), u, p; \chi_w) \quad \forall\, \chi_w \in H^1(\Omega[\phi(\cdot,t)])^d, \tag{4.12b}$$

$$\int_{\Omega[\phi(\cdot,t)]} \nabla u \cdot \nabla \chi_u \mathrm{d}x = \int_{\Omega[\phi(\cdot,t)]} f \chi_u \mathrm{d}x \qquad \forall\, \chi_u \in H_0^1(\Omega[\phi(\cdot,t)]), \tag{4.12c}$$

$$\int_{\Omega[\phi(\cdot,t)]} \nabla p \cdot \nabla \chi_p \mathrm{d}x = \int_{\Omega[\phi(\cdot,t)]} j'_u(x,u) \chi_p \mathrm{d}x \qquad \forall\, \chi_p \in H_0^1(\Omega[\phi(\cdot,t)]), \tag{4.12d}$$

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

**4.2.3 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 $k \geqslant 2$ 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 $F_K : \hat{K} \to K$, called the parametrization of $K$; see [41, §8.6]. Moreover, every boundary simplex $K \in \mathcal{K}_h^0$ (with one face or edge attached to $\Gamma^0$) contains a possibly curved face or edge to interpolate $\Gamma^0$ with accuracy of $O(h^{k+1})$. This can be obtained by using Lenoir's isoparametric finite elements [103] based on some parametrization of the boundary $\Upsilon : \partial\widetilde{D} \to \partial\Omega^0$, where $\partial\widetilde{D}$ 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 $\Upsilon$ to be the normal projection onto $\partial\Omega^0$, i.e., the unique point $\Upsilon(x) \in \Gamma$ such that (cf. [37, Section 2.1])

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

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

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

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

Let $\mathbf{x}^0 = (\xi_1, \cdots, \xi_N) \in \mathbb{R}^{dN}$ be the vector that collects all the nodes $\xi_j \in \mathbb{R}^d$, $j = 1, \ldots, N$, (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

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

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 of the triangles on both interior and boundary are curved, as in references [38, 41].

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 $F_K : \hat{K} \to K$ 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

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

It maps the initial domain $\Omega_h^0$ 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:

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

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

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

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:

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

The pullback of $\phi_j[\mathbf{x}(t)]$ from $\Omega_h[\mathbf{x}(t)]$ to $\Omega_h^0$ 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:

$$\partial_{t,h}^{\bullet}\phi_j[\mathbf{x}(t)] = 0 \quad \text{on } \ \Omega_h[\mathbf{x}(t)], \ \ j = 1,\dots,N. \tag{4.13}$$

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

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

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

**4.2.4 The numerical method and its convergence.** The semidiscrete evolving FEM for (4.12) 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 $u_h(t), p_h(t) \in \mathring{S}_h^k[\mathbf{x}(t)] \subset H_0^1(\Omega_h[\mathbf{x}(t)])$, satisfying the following weak formulation:

$$\partial_t \phi_h = w_h \circ \phi_h \tag{4.14a}$$

$$\int_{\Omega_h[\mathbf{x}(t)]} \nabla w_h : \nabla \chi_w + w_h \cdot \chi_w \mathrm{d}x = -\mathrm{d}J(\Gamma_h[\mathbf{x}(t)], u_h, p_h; \chi_w) \quad \forall\, \chi_w \in S_h^k[\mathbf{x}(t)]^d \tag{4.14b}$$

$$\int_{\Omega_h[\mathbf{x}(t)]} \nabla u_h \cdot \nabla \chi_u \mathrm{d}x = \int_{\Omega_h[\mathbf{x}(t)]} f \chi_u \mathrm{d}x \qquad \forall\, \chi_u \in \mathring{S}_h^k[\mathbf{x}(t)] \tag{4.14c}$$

$$\int_{\Omega_h[\mathbf{x}(t)]} \nabla p_h \cdot \nabla \chi_p \mathrm{d}x = \int_{\Omega_h[\mathbf{x}(t)]} j_u'(x, u_h) \chi_p \mathrm{d}x \qquad \forall\, \chi_p \in \mathring{S}_h^k[\mathbf{x}(t)] \tag{4.14d}$$

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

We consider the following time discretization of (4.14) 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 $u_h^n, p_h^n \in \mathring{S}_h^k[\mathbf{x}^n]$ such that

$$\frac{\phi_h^{n+1} - \phi_h^n}{\tau} = w_h^n \circ \phi_h^n \tag{4.15a}$$

$$\int_{\Omega_h[\mathbf{x}^n]} \nabla w_h^n : \nabla \chi_w + w_h^n \cdot \chi_w \mathrm{d}x = -\mathrm{d}J(\Gamma_h[\mathbf{x}^n], u_h^n, p_h^n; \chi_w) \quad \forall\, \chi_w \in S_h^k[\mathbf{x}^n]^d \tag{4.15b}$$

$$\int_{\Omega_h[\mathbf{x}^n]} \nabla u_h^n \cdot \nabla \chi_u \mathrm{d}x = \int_{\Omega_h[\mathbf{x}^n]} f \chi_u \mathrm{d}x \qquad \forall\, \chi_u \in \mathring{S}_h^k[\mathbf{x}^n] \tag{4.15c}$$

$$\int_{\Omega_h[\mathbf{x}^n]} \nabla p_h^n \cdot \nabla \chi_p \mathrm{d}x = \int_{\Omega_h[\mathbf{x}^n]} j_u'(x, u_h^n) \chi_p \mathrm{d}x \qquad \forall\, \chi_p \in \mathring{S}_h^k[\mathbf{x}^n] \tag{4.15d}$$

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 (4.15c)–(4.15d) and substituted into (4.15b) to yield $w_h^n$. The latter is used to defined $\phi_h^{n+1}$ through (4.15a).

**Remark 4.1.** We can also use BDF method to discrete (4.14a), i.e., the scheme (4.15a) can be generalized to BDF method. The temporal error analysis in this chapter can be extended to BDF method directly by using Taylor's expansion. By using BDF method, the equations (4.15a) and (4.15b) will become a nonlinear coupled system. In this case, we can use iteration method such as fixed point iteration to solve equations (4.15a) and (4.15b).

For the practical computation and numerical analysis it is convenient to use matrix-vector formulation to represent the system of equations in (4.15). For this purpose, we

express the unknown solutions in terms of the finite element basis functions $\phi_j[\mathbf{x}^n]$, $j = 1, \ldots, N$ on domain $\Omega_h[\mathbf{x}^n]$, i.e.,

$$u_h^n = \sum_{j=1}^{N} u_j^n \phi_j[\mathbf{x}^n], \qquad \text{with } u_j^n = u_h^n(x_j^n) \in \mathbb{R},$$

$$p_h^n = \sum_{j=1}^{N} p_j^n \phi_j[\mathbf{x}^n], \qquad \text{with } p_j^n = p_h^n(x_j^n) \in \mathbb{R},$$

$$w_h^n = \sum_{j=1}^{N} w_j^n \phi_j[\mathbf{x}^n], \qquad \text{with } w_j^n = w_h^n(x_j^n) \in \mathbb{R}^d,$$

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.,

$$\mathbf{M}(\mathbf{x}^n)|_{jk} = \int_{\Omega_h[\mathbf{x}^n]} \phi_j[\mathbf{x}^n]\phi_k[\mathbf{x}^n] \mathrm{d}x$$

$$\mathbf{A}(\mathbf{x}^n)|_{jk} = \int_{\Omega_h[\mathbf{x}^n]} \nabla\phi_j[\mathbf{x}^n] \cdot \nabla\phi_k[\mathbf{x}^n] \mathrm{d}x \quad \text{for } j,k = 1, \cdots, N.$$

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:

$$\|\mathbf{u}^n\|^2_{\mathbf{M}(\mathbf{x}^n)} := (\mathbf{u}^n)^\top \mathbf{M}(\mathbf{x}^n)\mathbf{u}^n = \|u_h^n\|^2_{L^2(\Omega_h[\mathbf{x}^n])},$$

$$\|\mathbf{u}^n\|^2_{\mathbf{A}(\mathbf{x}^n)} := (\mathbf{u}^n)^\top \mathbf{A}(\mathbf{x}^n)\mathbf{u}^n = \|\nabla u_h^n\|^2_{L^2(\Omega_h[\mathbf{x}^n])},$$

$$\|\mathbf{w}^n\|^2_{\mathbf{K}(\mathbf{x}^n)} := (\mathbf{w}^n)^\top \mathbf{K}(\mathbf{x}^n)\mathbf{w}^n = \|w_h^n\|^2_{H^1(\Omega_h[\mathbf{x}^n])}.$$

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

$$\mathbf{f}(\mathbf{x}^n)|_j = \int_{\Omega_h[\mathbf{x}^n]} f\phi_j[\mathbf{x}^n] \mathrm{d}x,$$

$$\mathbf{J}'_u(\mathbf{x}^n,\mathbf{u}^n)|_j = \int_{\Omega_h[\mathbf{x}^n]} j'_u(x,u_h^n)\phi_j[\mathbf{x}^n]\mathrm{d}x,$$

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

for $j = 1,\cdots,N$ and $1\leqslant l\leqslant d$.

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 (4.15) can be written into the following matrix-vector form:

$$\mathbf{x}^{n+1} - \mathbf{x}^n = \tau\mathbf{w}^n, \tag{4.16a}$$

$$\mathbf{K}(\mathbf{x}^n)\mathbf{w}^n = -\mathrm{d}\mathbf{J}(\mathbf{x}^n,\mathbf{u}^n,\mathbf{p}^n), \tag{4.16b}$$

$$\mathbf{A}(\mathbf{x}^n)\mathbf{u}^n = \mathbf{f}(\mathbf{x}^n), \tag{4.16c}$$

$$\mathbf{A}(\mathbf{x}^n)\mathbf{p}^n = \mathbf{J}'_u(\mathbf{x}^n,\mathbf{u}^n). \tag{4.16d}$$

We are now in the position to state the main result of this chapter, i.e., the convergence of the fully discrete evolving FEM in (4.15). 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 $\Omega_h^0$ 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 $\hat u_h^n, \hat p_h^n \in \mathring{S}_h^k[\mathbf{x}_*^n]$, can be compared with the numerical solution after both being pulled back to the initial domain $\Omega_h^0$. This is presented in the following theorem.

**Theorem 4.1** (Convergence of the evolving finite element approximations)**.** *Suppose that the solution of* (4.12)*, the flow map $\phi : \Omega^0\times[0,T]\to\mathbb{R}^d$ and its inverse $\phi(\cdot,t)^{-1} : \Omega[\phi(\cdot,t)]\to\Omega^0$, 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 4.3). Then there exist positive constants $\tau_0$ and $h_0$ such that for $\tau\leqslant\tau_0$ and $h\leqslant h_0$ together with the restriction $\tau = o(h^{\frac{d}{2}})$, the following error estimates hold:*

$$\begin{aligned}\|\phi_h^n - \hat\phi_h^n\|_{H^1(\Omega_h^0)} + \|w_h^n\circ\phi_h^n - \hat w_h^n\circ\hat\phi_h^n\|_{H^1(\Omega_h^0)} + \|u_h^n\circ\phi_h^n - \hat u_h^n\circ\hat\phi_h^n\|_{H^1(\Omega_h^0)}&\\ + \|p_h^n\circ\phi_h^n - \hat p_h^n\circ\hat\phi_h^n\|_{H^1(\Omega_h^0)} \leqslant C(\tau + h^k),&\end{aligned} \tag{4.17}$$

*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* $\hat{u}_h^n, \hat{p}_h^n \in \mathring{S}_h^k[\mathbf{x}_*^n]$ *are the Lagrangian interpolations of the exact solutions.*

**Remark 4.2.** 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 [15]):

$$\|\tilde{w}^n - \hat{w}_h^n\|_{W^{1,\infty}(\Omega_h[\mathbf{x}_*^n])} + \|\tilde{u}^n - \hat{u}_h^n\|_{W^{1,\infty}(\Omega_h[\mathbf{x}_*^n])} + \|\tilde{p}^n - \hat{p}_h^n\|_{W^{1,\infty}(\Omega_h[\mathbf{x}_*^n])} \leqslant Ch^k, \quad (4.18a)$$

$$\|\tilde{w}^n - \hat{w}_h^n\|_{L^{\infty}(\Omega_h[\mathbf{x}_*^n])} + \|\tilde{u}^n - \hat{u}_h^n\|_{L^{\infty}(\Omega_h[\mathbf{x}_*^n])} + \|\tilde{p}^n - \hat{p}_h^n\|_{L^{\infty}(\Omega_h[\mathbf{x}_*^n])} \leqslant Ch^{k+1}, \quad (4.18b)$$

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 4.3.** The constants $\tau_0$, $h_0$ 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 $[T_{j-1},T_j]$, $j=1,\ldots,m$, 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 $T_j$, $j=1,\ldots,m-1$. 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 $[0,T]$.

**4.3 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 [95, 38], 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 4.3.3.

**4.3.1 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 homolopy, 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 $\theta$ 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 $u_h^\theta$ by using a $N$-dimensional vector $\mathbf{u}$, we can define a $d$-dimensional vector-valued function $e_h^\theta$ by using the $dN$-dimensional vector $\mathbf{e}$.

In combination with the fundamental theorem of calculus, Lemma 4.1 and the transport property (4.13), the following lemma was proved in [38, Lemma 5.1].

**Lemma 4.3.** *In the above setting the following identities hold:*

$$\mathbf{u}^\top(\mathbf{M}(\mathbf{y}) - \mathbf{M}(\mathbf{z}))\mathbf{v} = \int_0^1 \int_{\Omega_h^\theta} u_h^\theta (\nabla \cdot e_h^\theta) v_h^\theta \mathrm{d}x\mathrm{d}\theta, \tag{4.19}$$

$$\mathbf{u}^\top(\mathbf{A}(\mathbf{y}) - \mathbf{A}(\mathbf{z}))\mathbf{v} = \int_0^1 \int_{\Omega_h^\theta} \nabla u_h^\theta \cdot (D_{\Omega_h^\theta} e_h^\theta) \nabla v_h^\theta \mathrm{d}x\mathrm{d}\theta, \tag{4.20}$$

*with* $D_{\Omega_h^\theta} e_h^\theta = trace(E) I_d - (E + E^\top)$ *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 $\nabla e_h^\theta$ 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 [38, Lemma 5.2].

**Lemma 4.4.** *If* $\|\nabla \cdot e_h^\theta\|_{L^\infty(\Omega_h^\theta)} \leqslant \mu$ *for* $0 \leqslant \theta \leqslant 1$, *then*

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

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

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

The following lemma was proved in [38, Lemma 5.3], which says that the condition in Lemma 4.4 can be reduced to $\theta = 0$.

**Lemma 4.5.** *If* $\|\nabla e_h^0\|_{L^\infty(\Omega_h^0)} \leqslant \frac{1}{2}$ *then the finite element function* $v_h^\theta$ *on* $\Omega_h^\theta$*, with* $0 \leqslant \theta \leqslant 1$*, satisfies the following estimate:*

$$\|\nabla_{\Omega_h^\theta} v_h^\theta\|_{L^p(\Omega_h^\theta)} \leqslant c_p \|\nabla_{\Omega_h^0} v_h^0\|_{L^p(\Omega_h^0)} \text{ for } 1 \leqslant p \leqslant \infty, \tag{4.21}$$

*where* $c_p$ *depends only on* $p$*.*

In Lenoir's isoparametric finite element approximation to $\Omega^0$, there exists a lift map $\Psi : \Omega_h^0 \to \Omega^0$ satisfying the following estimates (cf. [103]):

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

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

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

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 [95] and is used in the proof of the consistency estimates in Lemma 4.7.

**Lemma 4.6.** *The following estimates hold for* $\chi, \psi \in S_h^k[\mathbf{x}_*(t)]$ *and* $g \in W^{1,\infty}(\mathbb{R}^d)$*:*

$$\Big|\int_{\Omega_h[\mathbf{x}_*(t)]} \chi\psi \mathrm{d}x - \int_{\Omega(t)} \tilde{\chi}\tilde{\psi}\mathrm{d}x\Big| \leqslant Ch^k \|\chi\|_{L^2(\Omega_h[\mathbf{x}_*(t)])} \|\psi\|_{L^2(\Omega_h[\mathbf{x}_*(t)])},$$

$$\Big|\int_{\Omega_h[\mathbf{x}_*(t)]} \nabla\chi\cdot\nabla\psi \mathrm{d}x - \int_{\Omega(t)} \nabla\tilde{\chi}\cdot\nabla\tilde{\psi}\mathrm{d}x\Big| \leqslant Ch^k \|\chi\|_{H^1(\Omega_h[\mathbf{x}_*(t)])} \|\psi\|_{H^1(\Omega_h[\mathbf{x}_*(t)])},$$

$$\Big|\int_{\Omega_h[\mathbf{x}_*(t)]} g\chi \mathrm{d}x - \int_{\Omega(t)} g\tilde{\chi}\mathrm{d}x\Big| \leqslant Ch^k \|\chi\|_{L^2(\Omega_h[\mathbf{x}_*(t)])} \|g\|_{W^{1,\infty}(\mathbb{R}^d)},$$

$$\Big|\int_{\Omega_h[\mathbf{x}_*(t)]} g\nabla\chi \mathrm{d}x - \int_{\Omega(t)} g\nabla\tilde{\chi}\mathrm{d}x\Big| \leqslant Ch^k \|\chi\|_{H^1(\Omega_h[\mathbf{x}_*(t)])} \|g\|_{W^{1,\infty}(\mathbb{R}^d)}.$$

**4.3.2 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 $\hat{u}_h^n, \hat{p}_h^n \in$

$\mathring{S}_h^k[\mathbf{x}_*^n]$, with the numerical solutions $\phi_h^n$, $w_h^n$, $u_h^n$ and $p_h^n$, after pulling these functions back to the initial domain $\Omega_h^0$.

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 $\hat{u}_h^n, \hat{p}_h^n \in \mathring{S}_h^k[\mathbf{x}_*^n]$ satisfy the weak formulations up to some defects:

$$\frac{\hat{\phi}_h^{n+1} - \hat{\phi}_h^n}{\tau} = \hat{w}_h^n \circ \hat{\phi}_h^n + d_\phi^n, \tag{4.22a}$$

$$\int_{\Omega_h[\mathbf{x}_*^n]} \nabla \hat{w}_h^n : \nabla \chi_w + \hat{w}_h^n \cdot \chi_w \mathrm{d}x = -\mathrm{d}J(\Gamma_h[\mathbf{x}_*^n], \hat{u}_h^n, \hat{p}_h^n; \chi_w) + \int_{\Omega_h[\mathbf{x}_*^n]} d_w^n \cdot \chi_w \mathrm{d}x, \tag{4.22b}$$

$$\int_{\Omega_h[\mathbf{x}_*^n]} \nabla \hat{u}_h^n \cdot \nabla \chi_u \mathrm{d}x = \int_{\Omega_h[\mathbf{x}_*^n]} f \chi_u \mathrm{d}x + \int_{\Omega_h[\mathbf{x}_*^n]} d_u^n \chi_u \mathrm{d}x, \tag{4.22c}$$

$$\int_{\Omega_h[\mathbf{x}_*^n]} \nabla \hat{p}_h^n \cdot \nabla \chi_p \mathrm{d}x = \int_{\Omega_h[\mathbf{x}_*^n]} j_u'(x, \hat{u}_h^n) \chi_p \mathrm{d}x + \int_{\Omega_h[\mathbf{x}_*^n]} d_p^n \chi_p \mathrm{d}x, \tag{4.22d}$$

for test functions $\chi_w \in S_h^k[\mathbf{x}_*^n]^d$ and $\chi_u, \chi_p \in \mathring{S}_h^k[\mathbf{x}_*^n]$, where $d_\phi^n \in S_h^k[\mathbf{x}_*^0]^d$, $d_w^n \in S_h^k[\mathbf{x}_*^n]^d$ and $d_u^n, d_p^n \in S_h^k[\mathbf{x}_*^n]$ 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 $\mathbf{w}_*^n$, $\mathbf{u}_*^n$, $\mathbf{p}_*^n$ $\mathbf{d}_\mathbf{x}^n$, $\mathbf{d}_\mathbf{w}^n$ $\mathbf{d}_\mathbf{u}^n$ and $\mathbf{d}_\mathbf{p}^n$ the column vectors that collect the nodal values of $\hat{w}_h^n$, $\hat{u}_h^n$, $\hat{p}_h^n$ $d_\phi^n$, $d_w^n$, $d_u^n$ and $d_p^n$, respectively. The right-hand side vectors $\mathbf{f}(\mathbf{x}_*^n)$, $\mathbf{J}_u'(\mathbf{x}_*^n, \mathbf{u}_*^n)$ and $-\mathrm{d}\mathbf{J}(\mathbf{x}_*^n, \mathbf{u}_*^n, \mathbf{p}_*^n)$ are defined by

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

for $j = 1, \cdots, N$ and $1 \leqslant l \leqslant d$. Then (4.22) can be written into the following matrix-vector form:

$$\mathbf{x}_*^{n+1} - \mathbf{x}_*^n = \tau \mathbf{w}_*^n + \tau \mathbf{d}_\mathbf{x}^n, \tag{4.23a}$$

$$\mathbf{K}(\mathbf{x}_*^n)\mathbf{w}_*^n = -\mathrm{d}\mathbf{J}(\mathbf{x}_*^n, \mathbf{u}_*^n, \mathbf{p}_*^n) + \mathbf{M}^{[d]}(\mathbf{x}_*^n)\mathbf{d}_\mathbf{w}^n, \tag{4.23b}$$

$$\mathbf{A}(\mathbf{x}_*^n)\mathbf{u}_*^n = \mathbf{f}(\mathbf{x}_*^n) + \mathbf{M}(\mathbf{x}_*^n)\mathbf{d}_\mathbf{u}^n, \tag{4.23c}$$

$$\mathbf{A}(\mathbf{x}_*^n)\mathbf{p}_*^n = \mathbf{J}_\mathbf{u}'(\mathbf{x}_*^n, \mathbf{u}_*^n) + \mathbf{M}(\mathbf{x}_*^n)\mathbf{d}_\mathbf{p}^n, \tag{4.23d}$$

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:

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

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

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

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

$$\begin{aligned}\|\mathbf{d}_\mathbf{u}^n\|_{\star,\mathbf{x}_*^n}^2 &:= (\mathbf{d}_\mathbf{u}^n)^\top \mathbf{M}(\mathbf{x}_*^n)\mathbf{K}(\mathbf{x}_*^n)^{-1}\mathbf{M}(\mathbf{x}_*^n)\mathbf{d}_\mathbf{u}^n = \|d_u^n\|_{H_h^{-1}(\Omega_h[\mathbf{x}_*^n])}^2, \\ \|\mathbf{d}_\mathbf{p}^n\|_{\star,\mathbf{x}_*^n}^2 &:= (\mathbf{d}_\mathbf{p}^n)^\top \mathbf{M}(\mathbf{x}_*^n)\mathbf{K}(\mathbf{x}_*^n)^{-1}\mathbf{M}(\mathbf{x}_*^n)\mathbf{d}_\mathbf{p}^n = \|d_p^n\|_{H_h^{-1}(\Omega_h[\mathbf{x}_*^n])}^2.\end{aligned}$$

The stability estimates will be established by comparing the matrix-vector formulations (4.16) and (4.23). By subtracting (4.23) from (4.16), we obtain the following equations for the errors $\mathbf{e}_\mathbf{x}^n = \mathbf{x}^n - \mathbf{x}_*^n$, $\mathbf{e}_\mathbf{w}^n = \mathbf{w}^n - \mathbf{w}_*^n$, $\mathbf{e}_\mathbf{u}^n = \mathbf{u}^n - \mathbf{u}_*^n$ and $\mathbf{e}_\mathbf{p}^n = \mathbf{p}^n - \mathbf{p}_*^n$:

$$\mathbf{e}_\mathbf{x}^{n+1} = \mathbf{e}_\mathbf{x}^n + \tau\mathbf{e}_\mathbf{w}^n - \tau\mathbf{d}_\mathbf{x}^n, \tag{4.24a}$$

$$\begin{aligned}\mathbf{K}(\mathbf{x}_*^n)\mathbf{e}_\mathbf{w}^n =& -(\mathbf{K}(\mathbf{x}^n) - \mathbf{K}(\mathbf{x}_*^n))\mathbf{e}_\mathbf{w}^n - (\mathbf{K}(\mathbf{x}^n) - \mathbf{K}(\mathbf{x}_*^n))\mathbf{w}_*^n \\ &- (\mathrm{d}\mathbf{J}(\mathbf{x}^n, \mathbf{u}^n, \mathbf{p}^n) - \mathrm{d}\mathbf{J}(\mathbf{x}_*^n, \mathbf{u}_*^n, \mathbf{p}_*^n)) - \mathbf{M}(\mathbf{x}_*^n)\mathbf{d}_\mathbf{w}^n,\end{aligned} \tag{4.24b}$$

$$\begin{aligned}\mathbf{A}(\mathbf{x}_*^n)\mathbf{e}_\mathbf{u}^n =& -(\mathbf{A}(\mathbf{x}^n) - \mathbf{A}(\mathbf{x}_*^n))\mathbf{e}_\mathbf{u}^n - (\mathbf{A}(\mathbf{x}^n) - \mathbf{A}(\mathbf{x}_*^n))\mathbf{u}_*^n \\ &+ (\mathbf{f}(\mathbf{x}^n) - \mathbf{f}(\mathbf{x}_*^n)) - \mathbf{M}(\mathbf{x}_*^n)\mathbf{d}_\mathbf{u}^n,\end{aligned} \tag{4.24c}$$

$$\begin{aligned}\mathbf{A}(\mathbf{x}_*^n)\mathbf{e}_\mathbf{p}^n =& -(\mathbf{A}(\mathbf{x}^n) - \mathbf{A}(\mathbf{x}_*^n))\mathbf{e}_\mathbf{p}^n - (\mathbf{A}(\mathbf{x}^n) - \mathbf{A}(\mathbf{x}_*^n))\mathbf{p}_*^n \\ &+ (\mathbf{J}_u'(\mathbf{x}^n, \mathbf{u}^n) - \mathbf{J}_u'(\mathbf{x}_*^n, \mathbf{u}_*^n)) - \mathbf{M}(\mathbf{x}_*^n)\mathbf{d}_\mathbf{p}^n.\end{aligned} \tag{4.24d}$$

The error estimates depend on the estimates for the defect terms $\mathbf{d}_\mathbf{x}^n$, $\mathbf{d}_\mathbf{w}^n$, $\mathbf{d}_\mathbf{u}^n$ and $\mathbf{d}_\mathbf{p}^n$ (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 4.6.

**Lemma 4.7** (Consistency estimates)**.** *Under the assumptions of Theorem 4.1, there exist positive constants $\tau_0$ and $h_0$ such that for $\tau \leqslant \tau_0$ and $h \leqslant h_0$, the following consistency error estimates hold:*

$$\sup_{1\leqslant n\leqslant[T/\tau]} \|\mathbf{d}_{\mathbf{x}}^n\|_{\mathbf{K}(\mathbf{x}_*^n)} \leqslant C\tau, \tag{4.25a}$$

$$\sup_{1\leqslant n\leqslant[T/\tau]} \left(\|\mathbf{d}_{\mathbf{w}}^n\|_{\star,\mathbf{x}_*^n} + \|\mathbf{d}_{\mathbf{u}}^n\|_{\star,\mathbf{x}_*^n} + \|\mathbf{d}_{\mathbf{p}}^n\|_{\star,\mathbf{x}_*^n}\right) \leqslant Ch^k. \tag{4.25b}$$

**4.3.3 Stability estimates.** From the error equations in (4.24) and the consistency estimates in Lemma 4.7, we can derive the following stability estimates.

**Proposition 4.1** (Stability estimates)**.** *Under the assumptions of Theorem 4.1, there exists a positive constant $h_0$ such that for $\tau = o(h^{\frac{d}{2}})$ and $h \leqslant h_0$ the following stability estimate holds:*

$$\begin{aligned}&\sup_{0\leqslant n\leqslant[T/\tau]} \left(\|\mathbf{e}_{\mathbf{x}}^n\|_{\mathbf{K}(\mathbf{x}_*^n)} + \|\mathbf{e}_{\mathbf{w}}^n\|_{\mathbf{K}(\mathbf{x}_*^n)} + \|\mathbf{e}_{\mathbf{u}}^n\|_{\mathbf{K}(\mathbf{x}_*^n)} + \|\mathbf{e}_{\mathbf{p}}^n\|_{\mathbf{K}(\mathbf{x}_*^n)}\right)\\ &\leqslant C \sup_{0\leqslant n\leqslant[T/\tau]} (\|\mathbf{d}_{\mathbf{x}}^n\|_{\mathbf{K}(\mathbf{x}_*^n)} + \|\mathbf{d}_{\mathbf{w}}^n\|_{\star,\mathbf{x}_*^n} + \|\mathbf{d}_{\mathbf{u}}^n\|_{\star,\mathbf{x}_*^n} + \|\mathbf{d}_{\mathbf{p}}^n\|_{\star,\mathbf{x}_*^n}),\end{aligned} \tag{4.26}$$

*where $C$ is independent of $\tau$, $h$ 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 $e_u^n, e_p^n \in \mathring{S}_h^k[\mathbf{x}_*^n]$ be the finite element error functions on $\Omega_h[\mathbf{x}_*^n]$ with nodal vectors $\mathbf{e}_{\mathbf{x}}^n$, $\mathbf{e}_{\mathbf{w}}^n$, $\mathbf{e}_{\mathbf{u}}^n$ and $\mathbf{e}_{\mathbf{p}}^n$, respectively. We make the following mathematical induction on the boundedness of the errors: We assume that there exists an integer $1 \leqslant m \leqslant [T/\tau]$ such that the following inequalities hold for all $0 \leqslant n \leqslant m-1$,

$$\|e_x^n\|_{W^{1,\infty}(\Omega_h[\mathbf{x}_*^n])} \leqslant h^{-\frac{d}{4}}(\tau^{\frac{1}{2}} + h^{\frac{k}{2}}), \tag{4.27a}$$

$$\|e_u^n\|_{W^{1,\infty}(\Omega_h[\mathbf{x}_*^n])} \leqslant 1, \tag{4.27b}$$

$$\|e_p^n\|_{W^{1,\infty}(\Omega_h[\mathbf{x}_*^n])} \leqslant 1. \tag{4.27c}$$

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 (4.24c) and (4.24d) with $\mathbf{e}_{\mathbf{u}}^0$ and $\mathbf{e}_{\mathbf{p}}^0$, respectively, yields the following relations:

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

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

$$\|\mathbf{e}_{\mathbf{u}}^0\|_{\mathbf{K}(\mathbf{x}_*^0)} \leqslant C\|\mathbf{d}_{\mathbf{u}}^0\|_{\star,\mathbf{x}_*^0} \leqslant Ch^k$$

$$(\mathbf{e}_{\mathbf{p}}^0)^\top(\mathbf{J}_u'(\mathbf{x}^0,\mathbf{u}^0) - \mathbf{J}_u'(\mathbf{x}_*^0,\mathbf{u}_*^0)) = \int_{\Omega_h^0} e_p^0(u_h^0 - \hat{u}_h^0)\mathrm{d}x \leqslant Ch^{k+1}\|\mathbf{e}_{\mathbf{p}}^0\|_{\mathbf{M}(\mathbf{x}_*^0)},$$

and therefore

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

By the inverse inequality of finite element functions, we have

$$\|e_u^0\|_{W^{1,\infty}(\Omega_h[\mathbf{x}_*^0])} \leqslant Ch^{-\frac{d}{2}}\|\mathbf{e}_{\mathbf{u}}^0\|_{\mathbf{K}(\mathbf{x}_*^0)} \leqslant Ch^{k-\frac{d}{2}},$$

$$\|e_p^0\|_{W^{1,\infty}(\Omega_h[\mathbf{x}_*^0])} \leqslant Ch^{-\frac{d}{2}}\|\mathbf{e}_{\mathbf{p}}^0\|_{\mathbf{K}(\mathbf{x}_*^0)} \leqslant Ch^{k-\frac{d}{2}}.$$

Therefore, in the case $k \geqslant 2 > d/2$, assumptions (4.27b)–(4.27c) hold for $n = 0$ when $h$ is sufficiently small.

Now we prove that the stated error bounds (4.26) hold for all time levels $0 \leqslant n \leqslant m$ under the induction assumption. The mathematical induction will be completed by proving that (4.27) also holds for $n = m$.

(A) *Estimates for* $\mathbf{e}_{\mathbf{x}}^n$ : First, multiplying (4.24a) by matrix $\mathbf{K}(\mathbf{x}_*^n)$, we have

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

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

$$\begin{aligned}\|\mathbf{e}_{\mathbf{x}}^n\|_{\mathbf{K}(\mathbf{x}_*^n)}^2 &= (\mathbf{e}_{\mathbf{x}}^n)^\top\mathbf{K}(\mathbf{x}_*^n)\mathbf{e}_{\mathbf{x}}^{n-1} + \tau(\mathbf{e}_{\mathbf{x}}^n)^\top\mathbf{K}(\mathbf{x}_*^n)\mathbf{e}_{\mathbf{w}}^{n-1} - \tau(\mathbf{e}_{\mathbf{x}}^n)^\top\mathbf{K}(\mathbf{x}_*^n)\mathbf{d}_{\mathbf{x}}^{n-1}\\ &\leqslant \|\mathbf{e}_{\mathbf{x}}^n\|_{\mathbf{K}(\mathbf{x}_*^n)}\|\mathbf{e}_{\mathbf{x}}^{n-1}\|_{\mathbf{K}(\mathbf{x}_*^n)} + \tau\|\mathbf{e}_{\mathbf{x}}^n\|_{\mathbf{K}(\mathbf{x}_*^n)}\Big(\|\mathbf{e}_{\mathbf{w}}^{n-1}\|_{\mathbf{K}(\mathbf{x}_*^n)} + \|\mathbf{d}_{\mathbf{x}}^{n-1}\|_{\mathbf{K}(\mathbf{x}_*^n)}\Big).\end{aligned}$$

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

$$\|\mathbf{e}_{\mathbf{x}}^n\|_{\mathbf{K}(\mathbf{x}_*^n)} \leqslant \|\mathbf{e}_{\mathbf{x}}^{n-1}\|_{\mathbf{K}(\mathbf{x}_*^n)} + \tau\|\mathbf{e}_{\mathbf{w}}^{n-1}\|_{\mathbf{K}(\mathbf{x}_*^n)} + \tau\|\mathbf{d}_{\mathbf{x}}^{n-1}\|_{\mathbf{K}(\mathbf{x}_*^n)}. \tag{4.28}$$

In order to iterate the inequality above with respect to $n$, we need to convert $\|\mathbf{e}_{\mathbf{x}}^{n-1}\|_{\mathbf{K}(\mathbf{x}_*^n)}$ to $\|\mathbf{e}_{\mathbf{x}}^{n-1}\|_{\mathbf{K}(\mathbf{x}_*^{n-1})}$ on the right-hand side by utilizing Lemmas 4.3–4.5, which imply the following result:

$$\|\mathbf{e}_{\mathbf{x}}^{n-1}\|_{\mathbf{K}(\mathbf{x}_*^n)}^2 - \|\mathbf{e}_{\mathbf{x}}^{n-1}\|_{\mathbf{K}(\mathbf{x}_*^{n-1})}^2 \leqslant C\tau\|\mathbf{e}_{\mathbf{x}}^{n-1}\|_{\mathbf{K}(\mathbf{x}_*^{n-1})}^2. \tag{4.29}$$

Now, we can substitute (4.29) into (4.28) and sum up the result with respect to $n$. This yields the following estimate:

$$\|\mathbf{e}_{\mathbf{x}}^n\|_{\mathbf{K}(\mathbf{x}_*^n)} \leqslant C\tau \sum_{l=1}^{n} \left( \|\mathbf{e}_{\mathbf{w}}^{l-1}\|_{\mathbf{K}(\mathbf{x}_*^{l-1})} + \|\mathbf{d}_{\mathbf{x}}^{l-1}\|_{\mathbf{K}(\mathbf{x}_*^{l-1})} \right) + C\tau \sum_{l=1}^{n} \|\mathbf{e}_{\mathbf{x}}^{l-1}\|_{\mathbf{K}(\mathbf{x}_*^{l-1})} \quad \text{for } 1 \leqslant n \leqslant m. \tag{4.30}$$

Similarly, testing equations (4.24b), (4.24c) and (4.24d) with $\mathbf{e}_{\mathbf{w}}^n$, $\mathbf{e}_{\mathbf{u}}^n$ and $\mathbf{e}_{\mathbf{p}}^n$, respectively, we obtain

$$\begin{aligned}
\|\mathbf{e}_{\mathbf{w}}^n\|_{\mathbf{K}(\mathbf{x}_*^n)}^2 =& -(\mathbf{e}_{\mathbf{w}}^n)^\top(\mathbf{K}(\mathbf{x}^n)-\mathbf{K}(\mathbf{x}_*^n))\mathbf{e}_{\mathbf{w}}^n - (\mathbf{e}_{\mathbf{w}}^n)^\top(\mathbf{K}(\mathbf{x}^n)-\mathbf{K}(\mathbf{x}_*^n))\mathbf{w}_*^n \qquad (4.31a)\\
&-(\mathbf{e}_{\mathbf{w}}^n)^\top(\mathrm{d}\mathbf{J}(\mathbf{x}^n,\mathbf{u}^n,\mathbf{p}^n)-\mathrm{d}\mathbf{J}(\mathbf{x}_*^n,\mathbf{u}_*^n,\mathbf{p}_*^n)) - (\mathbf{e}_{\mathbf{w}}^n)^\top\mathbf{M}(\mathbf{x}_*^n)\mathbf{d}_{\mathbf{w}}^n,\\
\|\mathbf{e}_{\mathbf{u}}^n\|_{\mathbf{A}(\mathbf{x}_*^n)}^2 =& -(\mathbf{e}_{\mathbf{u}}^n)^\top(\mathbf{A}(\mathbf{x}^n)-\mathbf{A}(\mathbf{x}_*^n))\mathbf{e}_{\mathbf{u}}^n - (\mathbf{e}_{\mathbf{u}}^n)^\top(\mathbf{A}(\mathbf{x}^n)-\mathbf{A}(\mathbf{x}_*^n))\mathbf{u}_*^n \qquad (4.31b)\\
&+(\mathbf{e}_{\mathbf{u}}^n)^\top(\mathbf{f}(\mathbf{x}^n)-\mathbf{f}(\mathbf{x}_*^n)) - (\mathbf{e}_{\mathbf{u}}^n)^\top\mathbf{M}(\mathbf{x}_*^n)\mathbf{d}_{\mathbf{u}}^n,\\
\|\mathbf{e}_{\mathbf{p}}^n\|_{\mathbf{A}(\mathbf{x}_*^n)}^2 =& -(\mathbf{e}_{\mathbf{p}}^n)^\top(\mathbf{A}(\mathbf{x}^n)-\mathbf{A}(\mathbf{x}_*^n))\mathbf{e}_{\mathbf{p}}^n - (\mathbf{e}_{\mathbf{p}}^n)^\top(\mathbf{A}(\mathbf{x}^n)-\mathbf{A}(\mathbf{x}_*^n))\mathbf{p}_*^n \qquad (4.31c)\\
&+(\mathbf{e}_{\mathbf{p}}^n)^\top(\mathbf{J}_u'(\mathbf{x}^n,\mathbf{u}^n)-\mathbf{J}_u'(\mathbf{x}_*^n,\mathbf{u}_*^n)) - (\mathbf{e}_{\mathbf{p}}^n)^\top\mathbf{M}(\mathbf{x}_*^n)\mathbf{d}_{\mathbf{p}}^n.
\end{aligned}$$

In the following, we present estimates for $\mathbf{e}_{\mathbf{u}}^n$, $\mathbf{e}_{\mathbf{p}}^n$ and $\mathbf{e}_{\mathbf{w}}^n$ by using the equations in (4.31).

(B) *Estimates for* $\mathbf{e}_{\mathbf{u}}^n$ *under assumption* (4.27a): We estimate the four terms on the right-hand side of (4.31b) separately by using Lemmas 4.3–4.5. 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 $\mathbf{x}_\theta^n = \mathbf{x}_*^n + \theta\mathbf{e}_{\mathbf{x}}^n$ with nodal vectors $\mathbf{e}_{\mathbf{x}}^n, \mathbf{e}_{\mathbf{u}}^n, \mathbf{u}_*^n$, respectively.

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

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

Under the induction assumption in (4.27a) we obtain from Lemma 4.5 that

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

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

$$(\mathbf{e}_{\mathbf{u}}^n)^\top(\mathbf{A}(\mathbf{x}^n)-\mathbf{A}(\mathbf{x}_*^n))\mathbf{u}_*^n \leqslant \int_0^1 \|\nabla e_{u,\theta}^n\|_{L^2(\Omega_h[\mathbf{x}_\theta^n])} \|\nabla e_{x,\theta}^n\|_{L^2(\Omega_h[\mathbf{x}_\theta^n])} \|\nabla \hat{u}_{h,\theta}^n\|_{L^\infty(\Omega_h[\mathbf{x}_\theta^n])} \mathrm{d}\theta$$

$$\leqslant C\|\nabla e_u^n\|_{L^2(\Omega_h[\mathbf{x}_*^n])}\|\nabla e_x^n\|_{L^2(\Omega_h[\mathbf{x}_*^n])}\|\hat{u}_h^n\|_{W^{1,\infty}(\Omega_h[\mathbf{x}_*^n])}$$
$$\leqslant C\|\mathbf{e}_{\mathbf{u}}^n\|_{\mathbf{A}(\mathbf{x}_*^n)}\|\mathbf{e}_{\mathbf{x}}^n\|_{\mathbf{A}(\mathbf{x}_*^n)},$$

where the second to last inequality follows from Lemma 4.5.

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

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

where $\partial_\theta^\bullet$ denotes the material derivative with respect to $\theta$. 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

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

In view of Lemmas 4.4 and 4.5, we have

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

The fourth term on the right hand side of (4.31b) can be estimated directly as follows:

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

Combining the four estimates above, we derive that

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

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 $Ch^{-\frac{d}{2}}(\tau+h^k) \leqslant \frac{1}{4}$, we have

$$\|\mathbf{e}_{\mathbf{u}}^n\|_{\mathbf{K}(\mathbf{x}_*^n)} \leqslant C\|\mathbf{e}_{\mathbf{x}}^n\|_{\mathbf{K}(\mathbf{x}_*^n)} + C\|\mathbf{d}_{\mathbf{u}}^n\|_{\star,\mathbf{x}_*^n} \quad \text{for } 0\leqslant n\leqslant m-1. \tag{4.32}$$

(C) *Estimates for* $\mathbf{e}_{\mathbf{p}}^n$ *under assumption* (4.27a): The estimation of the first, second and fourth term on the right-hand side of (4.31c) 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

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

By Lemma 4.5 and Hölder's inequality, and induction assumption (4.27a), we have

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

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

$$\|\mathbf{e}_{\mathbf{p}}^n\|_{\mathbf{K}(\mathbf{x}_*^n)}\leqslant C\|\mathbf{e}_{\mathbf{u}}^n\|_{\mathbf{K}(\mathbf{x}_*^n)}+C\|\mathbf{e}_{\mathbf{x}}^n\|_{\mathbf{K}(\mathbf{x}_*^n)}+C\|\mathbf{d}_{\mathbf{p}}^n\|_{\star,\mathbf{x}_*^n}\quad\text{for }0\leqslant n\leqslant m-1.\tag{4.33}$$

(D) *Estimates for* $\mathbf{e}_{\mathbf{w}}^n$ *under assumption* (4.27): The first, second and fourth terms on the right-hand side of (4.31a) can be estimated similarly. The third term need to be estimated by using the expressions of $\mathrm{d}\mathbf{J}(\mathbf{x}^n,\mathbf{u}^n,\mathbf{p}^n)$ and $\mathrm{d}\mathbf{J}(\mathbf{x}_*^n,\mathbf{u}_*^n,\mathbf{p}_*^n)$ 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 4.2, which gives us the following expression:

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

Then

$$\partial_\theta^\bullet\big(\nabla u_{h,\theta}^n\cdot(\nabla e_{w,\theta}^n+(\nabla e_{w,\theta}^n)^\top)\nabla p_{h,\theta}^n\big)$$

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

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

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

Therefore,

$$
\begin{aligned}
&(\mathbf{e}_{\mathbf{w}}^n)^\top(\mathrm{d}\mathbf{J}(\mathbf{x}^n,\mathbf{u}^n,\mathbf{p}^n)-\mathrm{d}\mathbf{J}(\mathbf{x}_*^n,\mathbf{u}_*^n,\mathbf{p}_*^n))\\
=&\int_0^1\int_{\Omega_h[\mathbf{x}_\theta^n]}\Big[\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\Big]\nabla\cdot e_{x,\theta}^n\mathrm{d}x\mathrm{d}\theta\\
&+\int_0^1\int_{\Omega_h[\mathbf{x}_\theta^n]}\Big[\Big(j(x,u_{h,\theta}^n)-(\nabla u_{h,\theta}^n)\cdot(\nabla p_{h,\theta}^n)\Big)\nabla\cdot e_{w,\theta}^n\Big]\nabla\cdot e_{x,\theta}^n\mathrm{d}x\mathrm{d}\theta\\
&-\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\mathrm{d}x\mathrm{d}\theta\\
&+\int_0^1\int_{\Omega_h[\mathbf{x}_\theta^n]}\Big(\nabla e_{u,\theta}^n-(\nabla e_{x,\theta}^n)^\top(\nabla u_{h,\theta}^n)\Big)\cdot(\nabla e_{w,\theta}^n+(\nabla e_{w,\theta}^n)^\top)(\nabla p_{h,\theta}^n)\mathrm{d}x\mathrm{d}\theta\\
&+\int_0^1\int_{\Omega_h[\mathbf{x}_\theta^n]}\Big(\nabla e_{p,\theta}^n-(\nabla e_{x,\theta}^n)^\top(\nabla p_{h,\theta}^n)\Big)\cdot(\nabla e_{w,\theta}^n+(\nabla e_{w,\theta}^n)^\top)(\nabla u_{h,\theta}^n)\mathrm{d}x\mathrm{d}\theta\\
&-\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)\mathrm{d}x\mathrm{d}\theta\\
&-\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)+fe_{w,\theta}^n\cdot\Big(\nabla e_{p,\theta}^n-(\nabla e_{x,\theta}^n)^\top(\nabla p_{h,\theta}^n)\Big)\mathrm{d}x\mathrm{d}\theta\\
&-\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^2u_d)^\top e_{w,\theta}^n\mathrm{d}x\mathrm{d}\theta\\
&+\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)\mathrm{d}x\mathrm{d}\theta\\
&-\int_0^1\int_{\Omega_h[\mathbf{x}_\theta^n]}(\nabla e_{u,\theta}^n-(\nabla e_{x,\theta})^\top\nabla u_{h,\theta}^n)\cdot\nabla p_{h,\theta}^n(\nabla\cdot e_{w,\theta}^n)\mathrm{d}x\mathrm{d}\theta\\
&-\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)\mathrm{d}x\mathrm{d}\theta
\end{aligned}
$$

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

By Lemmas 4.4 and 4.5 we obtain the following estimates of this term under the induction assumption in (4.27):

$$\begin{aligned}
&(\mathbf{e}_{\mathbf{w}}^n)^\top (\mathrm{d}\mathbf{J}(\mathbf{x}^n, \mathbf{u}^n, \mathbf{p}^n) - \mathrm{d}\mathbf{J}(\mathbf{x}_*^n, \mathbf{u}_*^n, \mathbf{p}_*^n)) \\
&\leqslant C\|\mathbf{e}_{\mathbf{w}}^n\|_{\mathbf{K}(\mathbf{x}_*^n)} (\|\mathbf{e}_{\mathbf{p}}^n\|_{\mathbf{K}(\mathbf{x}_*^n)} + \|\mathbf{e}_{\mathbf{x}}^n\|_{\mathbf{K}(\mathbf{x}_*^n)} + \|\mathbf{e}_{\mathbf{u}}^n\|_{\mathbf{K}(\mathbf{x}_*^n)})(1 + \|e_p^n\|_{W^{1,\infty}})(1 + \|e_x^n\|_{W^{1,\infty}}) \\
&\quad + C\|\mathbf{e}_{\mathbf{w}}^n\|_{\mathbf{K}(\mathbf{x}_*^n)} (\|\mathbf{e}_{\mathbf{x}}^n\|_{\mathbf{K}(\mathbf{x}_*^n)} + \|\mathbf{e}_{\mathbf{u}}^n\|_{\mathbf{K}(\mathbf{x}_*^n)} + \|\mathbf{e}_{\mathbf{p}}^n\|_{\mathbf{K}(\mathbf{x}_*^n)})(1 + \|e_u^n\|_{W^{1,\infty}})(1 + \|e_x^n\|_{W^{1,\infty}}) \\
&\leqslant C\|\mathbf{e}_{\mathbf{w}}^n\|_{\mathbf{K}(\mathbf{x}_*^n)} (\|\mathbf{e}_{\mathbf{p}}^n\|_{\mathbf{K}(\mathbf{x}_*^n)} + \|\mathbf{e}_{\mathbf{x}}^n\|_{\mathbf{K}(\mathbf{x}_*^n)} + \|\mathbf{e}_{\mathbf{u}}^n\|_{\mathbf{K}(\mathbf{x}_*^n)}),
\end{aligned}$$

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 (4.31a), for which the estimates are omitted due to the similarity as Part (B) of the proof, we have

$$\|\mathbf{e}_{\mathbf{w}}^n\|_{\mathbf{K}(\mathbf{x}_*^n)} \leqslant C(\|\mathbf{e}_{\mathbf{x}}^n\|_{\mathbf{K}(\mathbf{x}_*^n)} + \|\mathbf{e}_{\mathbf{u}}^n\|_{\mathbf{K}(\mathbf{x}_*^n)} + \|\mathbf{e}_{\mathbf{p}}^n\|_{\mathbf{K}(\mathbf{x}_*^n)} + \|\mathbf{d}_{\mathbf{w}}^n\|_{\star,\mathbf{x}_*^n}). \tag{4.34}$$

(E) *Combination of estimates for* $\|\mathbf{e}_{\mathbf{x}}^n\|_{\mathbf{K}(\mathbf{x}_*^n)}$, $\|\mathbf{e}_{\mathbf{u}}^n\|_{\mathbf{K}(\mathbf{x}_*^n)}$, $\|\mathbf{e}_{\mathbf{p}}^n\|_{\mathbf{K}(\mathbf{x}_*^n)}$ *and* $\|\mathbf{e}_{\mathbf{w}}^n\|_{\mathbf{K}(\mathbf{x}_*^n)}$: Finally, combining the estimates in (4.30), (4.32), (4.33) and (4.34), we have

$$\begin{aligned}
\|\mathbf{e}_{\mathbf{x}}^n\|_{\mathbf{K}(\mathbf{x}_*^n)} \leqslant & C\tau \sum_{l=1}^{n} \left( \|\mathbf{d}_{\mathbf{x}}^{l-1}\|_{\mathbf{K}(\mathbf{x}_*^{l-1})} + \|\mathbf{d}_{\mathbf{u}}^{l-1}\|_{\star,\mathbf{x}_*^{l-1}} + \|\mathbf{d}_{\mathbf{p}}^{l-1}\|_{\star,\mathbf{x}_*^{l-1}} + \|\mathbf{d}_{\mathbf{w}}^{l-1}\|_{\star,\mathbf{x}_*^{l-1}} \right) \\
& + C\tau \sum_{l=1}^{n} \|\mathbf{e}_{\mathbf{x}}^{l-1}\|_{\mathbf{K}(\mathbf{x}_*^{l-1})} \qquad \text{for } 1 \leqslant n \leqslant m.
\end{aligned}$$

By applying the discrete Gronwall's inequality we obtain

$$\begin{aligned}
\|\mathbf{e}_{\mathbf{x}}^n\|_{\mathbf{K}(\mathbf{x}_*^n)} \leqslant C \sup_{1 \leqslant l \leqslant n} & \left( \|\mathbf{d}_{\mathbf{x}}^{l-1}\|_{\mathbf{K}(\mathbf{x}_*^{l-1})} + \|\mathbf{d}_{\mathbf{u}}^{l-1}\|_{\star,\mathbf{x}_*^{l-1}} + \|\mathbf{d}_{\mathbf{p}}^{l-1}\|_{\star,\mathbf{x}_*^{l-1}} + \|\mathbf{d}_{\mathbf{w}}^{l-1}\|_{\star,\mathbf{x}_*^{l-1}} \right) \\
& \text{for } 1 \leqslant n \leqslant m.
\end{aligned} \tag{4.35}$$

Furthermore, under the assumptions of Proposition 4.1, the inequality above implies the following result:

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

With this result, estimates (4.32) and (4.33) also hold for $1 \leqslant n \leqslant m$. This leads to

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

which further imply (4.27) for $1 \leqslant n \leqslant m$ when $\tau = o(h^{\frac{d}{2}})$ and $h$ is sufficiently small. This completes the mathematical induction.

Accordingly, estimates (4.32), (4.33), (4.34) and (4.35) hold for all $1 \leqslant n \leqslant [T/\tau]$. Substituting (4.35) into the other three estimates yields the desired result in Proposition 4.1. □

**4.3.4 Proof of Theorem 4.1.** Theorem 4.1 is an immediate consequence of the consistency estimates in Lemma 4.7 and the stability estimates in Proposition 4.1. □

**4.4 Numerical examples.** In this section, we present numerical examples to support the theoretical analysis in this chapter 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:

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

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

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

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

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

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

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

As $t \to \infty$, the boundary $\Gamma(t)$ converges to the optimal boundary $\Gamma_\infty$ which minimizes the energy functional $J(\Gamma) = \int_\Omega j(x, u)\mathrm{d}x$ 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 4.1 (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 4.1 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 difference formula of order $1, 2, 3$, which we refer to as BDF-1, BDF-2 and BDF-3, respectively. For the extension to BDF-$k$, we need to replace $w_h^n$ by a $k$th-order approximation of $w_h^{n+1}$ in (4.15a) using the $k$-step backward extrapolation. This will maintain the $k$th-order accuracy while decoupling equations (4.15a) and (4.15b). We have commented on this in the revised manuscript. The errors of the time discretizations are presented in Figure 4.1 (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 4.1. The analysis of stability and convergence of numerical approximations by BDF-k, $k = 2, \ldots, 5$, should be similar as that for BDF-1 by using the multiplier and energy techniques in [93].

The shape and mesh of the evolving domain at time $t = 0$, $t = 5$ and $t = 30$ are presented in Figure 4.2 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 $\Gamma_\infty$.

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

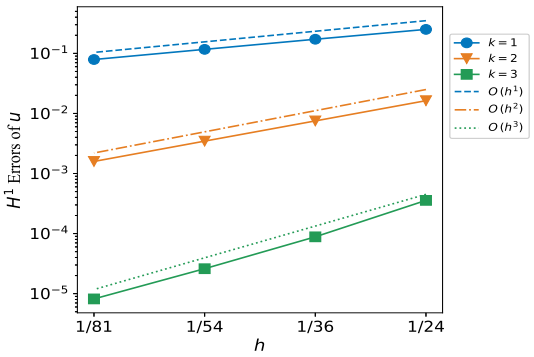
$$\Omega^0 = \{(x_1, x_2, x_3) : \frac{x_1^2}{0.85^2} + \frac{x_2^2}{0.45^2} + \frac{x_3^2}{0.45^2} \leqslant 1\}$$

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

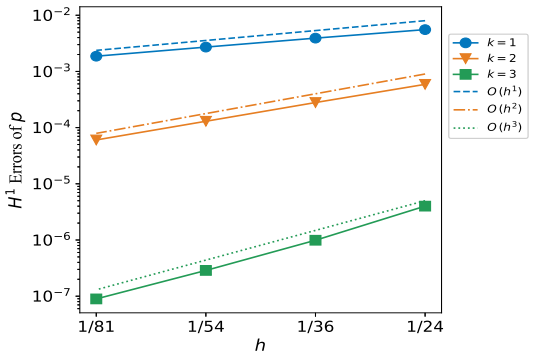
$$f = -\Delta u_d + u_d.$$

As $t \to \infty$, the boundary $\Gamma(t)$ converges to a three-dimensional dumbbell-shape optimal boundary $\Gamma_\infty$, as shown in Figure 4.3 (d). The shape and mesh of the evolving 3D domain at time $t = 0$, $t = 10$ and $t = 100$ are presented in Figure 4.3 (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 $\Gamma_\infty$.

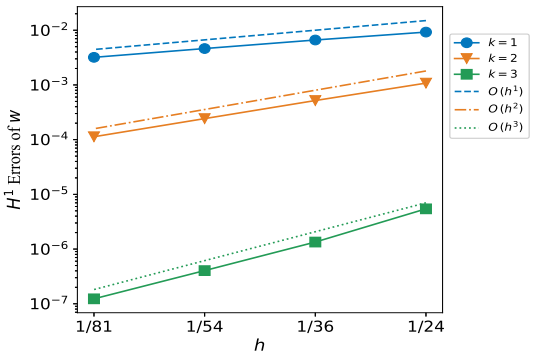
**Example 4.3** (Drag minimization under Stokes flow in three dimensions)**.**
We consider an example from drag minimization on the exterior of a three-dimensional



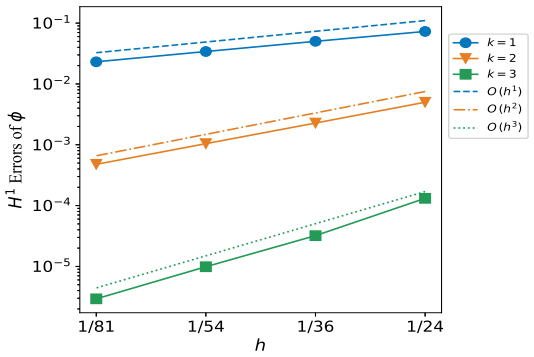

(a) $H^1$ error of $u$ from space discretization



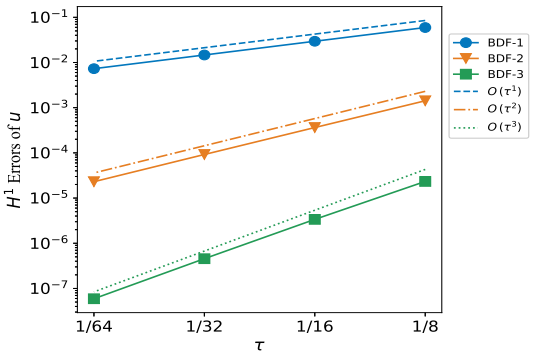

(b) $H^1$ error of $p$ from space discretization



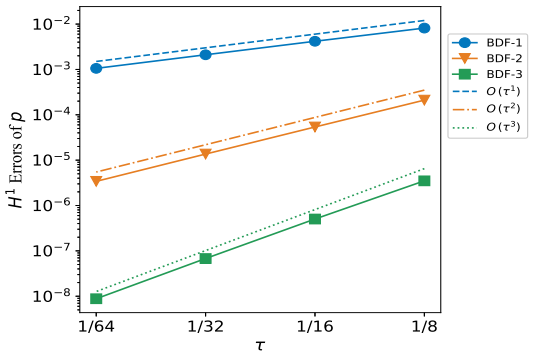

(c) $H^1$ error of $w$ from space discretization



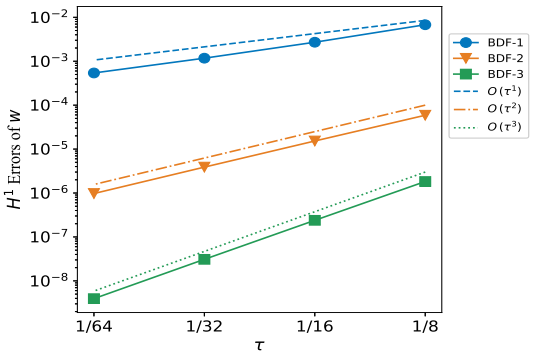

(d) $H^1$ error of $\phi$ from space discretization



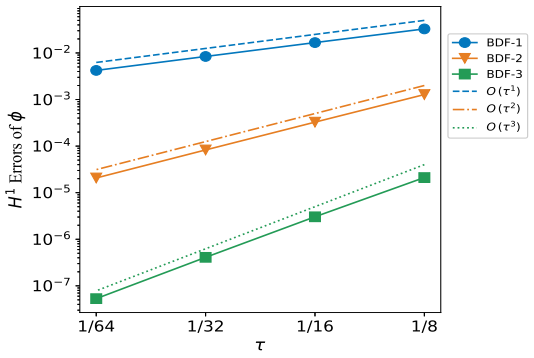

(e) $H^1$ error of $u$ from time discretization




(f) $H^1$ error of $p$ from time discretization

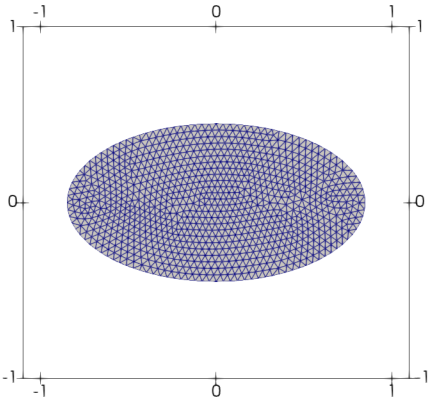



(g) $H^1$ error of $w$ from time discretization

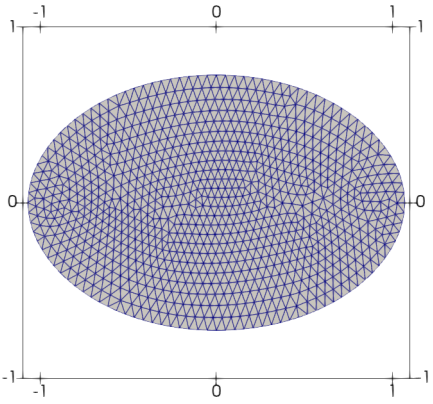



(h) $H^1$ error of $\phi$ from time discretization

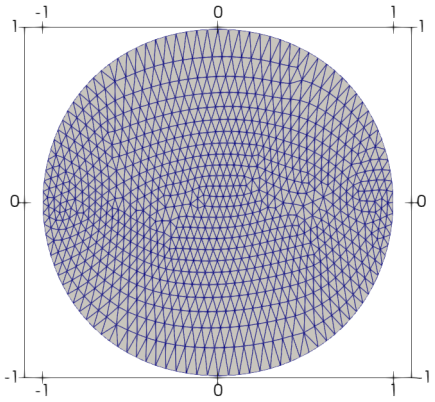
*Figure 4.1. Errors of the numerical solutions at time $T = 10$ (Example 4.1)*



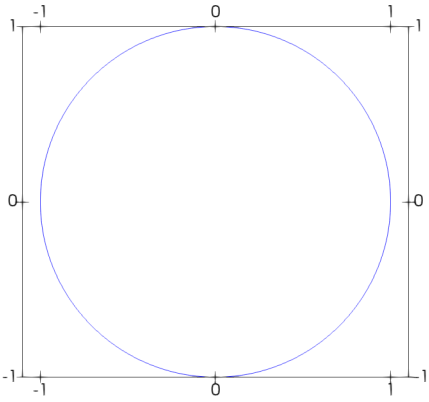
(a) The domain and mesh at $t = 0$



(b) The domain and mesh at $t = 5$

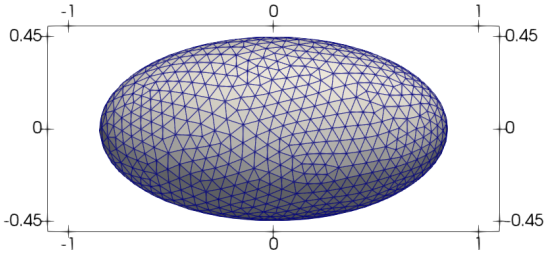


(c) The domain and mesh at $t = 30$

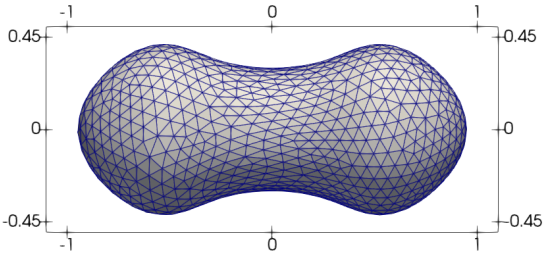


(d) The optimal boundary $\Gamma_\infty$

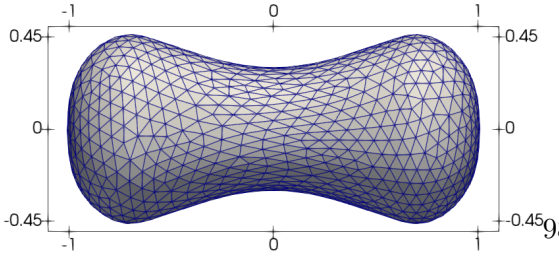
*Figure 4.2. Evolution of the 2D domain in Example 4.1.*



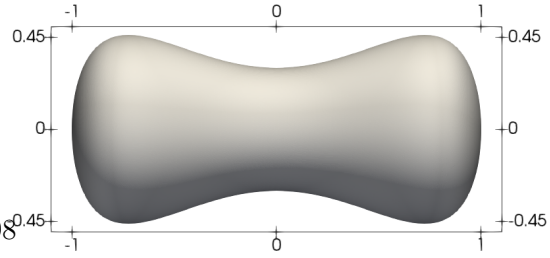
(a) The domain and mesh at $t = 0$



(b) The domain and mesh at $t = 10$



(c) The domain and mesh at $t = 100$



(d) The optimal boundary $\Gamma_\infty$

*Figure 4.3. Evolution of the 3D domain in Example 4.2.*

obstacle surrounded by viscous incompressible Stokes flow (see [57, Fig. 3.2]), which has many applications in airfoil design and obstacle problems. The analysis in this chapter could be extended to this example.

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

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

where $\mu = 0.1$ denotes the viscosity of the fluid and $\nu$ 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 $\Omega$, 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 $\Gamma_{\mathrm{in}}$ and $\Gamma_{\mathrm{out}}$, which are the left and right sides of the cube, respectively. The no-slip boundary condition is imposed on the other parts $\Gamma_{\mathrm{w}}$ of the boundary and the boundary $\Gamma$ of the obstacle. In our numerical test we set $f = 0$ and $u_{\mathrm{in}} = (1, 0, 0)^\top$.

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

$$\min_{\substack{\partial\Omega=\Gamma \\ \mathrm{Vol}(\Omega)=\mathrm{const}}} J(\Gamma) = \mu \int_{D\backslash\Omega} |\mathbb{D}u|^2 \mathrm{d}x.$$

The distributed Eulerian derivative of the energy functional $J(\Gamma)$ has the following expression (see [182, (36)]):

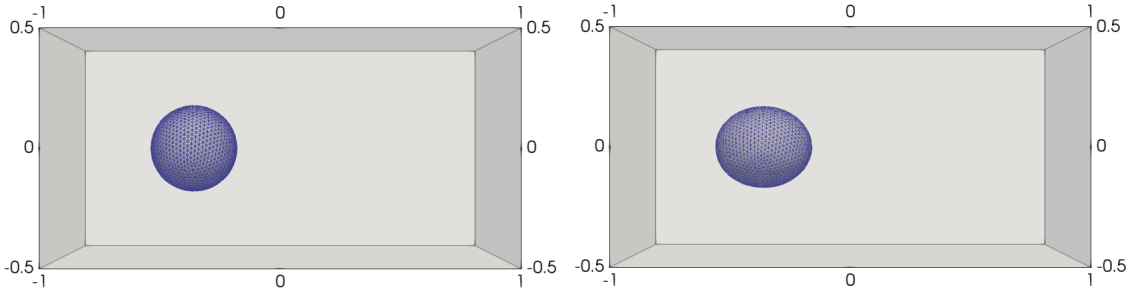
$$\mathrm{d}J(\Gamma; v) = \int_{D\backslash\Omega} \Big[ \mu \Big( |\mathbb{D}u|^2 \nabla \cdot v - \mathbb{D}u : (\nabla u \nabla v^\top + \nabla v \nabla u^\top) \Big) + p \nabla u : \nabla v^\top \Big] \mathrm{d}x.$$

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 $\Omega$ in the evolution, we choose a divergence-free descent velocity $w$ from the solution of the following equation (in the weak formulation):

$$\begin{aligned} &\int_{D\backslash\Omega} \nabla w : \nabla v \, \mathrm{d}x - \int_{D\backslash\Omega} q \, \nabla \cdot v \, \mathrm{d}x = -\mathrm{d}J(\Gamma; v), \\ &\int_{D\backslash\Omega} \nabla \cdot w \, \eta \, \mathrm{d}x = 0, \end{aligned} \tag{4.3}$$

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

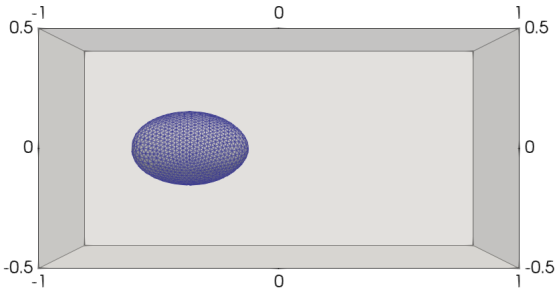
The evolution of shape and mesh of the three-dimensional domain at time $t = 0$, $t = 0.01$, $t = 0.04$ and $t = 0.05$ are presented in Figure 4.4 (a)-(d) with time stepsize $\tau = 0.001$, maximal mesh size $h = 0.06$ in $\Omega$, and local mesh size $h = 0.02$ near the obstacle. Our numerical simulation 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 4.4(d).



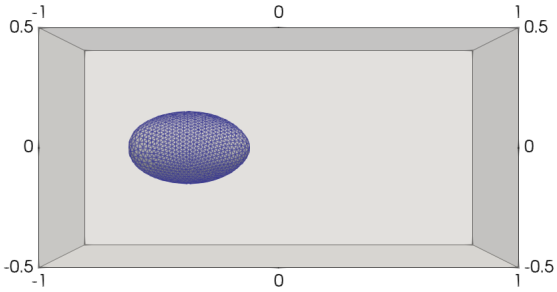
(a) The domain and mesh at $t = 0$



(b) The domain and mesh at $t = 0.01$



(c) The domain and mesh at $t = 0.04$



(d) The domain and mesh at $t = 0.05$

*Figure 4.4. Evolution of the shape of the obstacle in Example 4.3.*

# Chapter 5

## Optimal Convergence of Arbitrary Lagrangian–Eulerian Finite Element Methods for the Stokes Equation in an Evolving Domain

The content of this chapter has been published in "*Q. Rao, J. Wang, and Y. Xie. Optimal Convergence of Arbitrary Lagrangian–Eulerian Finite Element Methods for the Stokes Equation in an Evolving Domain, IMA J. Numer. Anal., drae097, 2025.*" This study was conducted in collaboration with another Ph.D, Y. Xie, and we contributed equally to the work.

**5.1 Introduction.** The Stokes equations are PDEs widely used to describe the motion of viscous fluids such as water and air. Solving the Stokes equations is a critical area of research in fluid dynamics, particularly when the domain is not fixed, such as in moving boundary/interface or fluid-structure interaction problems. The inclusion of such a dynamic domain introduces an additional layer of intricacy to the problem.

This chapter concerns the numerical solution of the Stokes equations in a time-dependent domain $\Omega(t) \subset \mathbb{R}^d$ with $d \in \{2, 3\}$, i.e.,

$$\partial_t u - \Delta u + \nabla p = f \quad \text{in} \quad \bigcup_{t\in(0,T]} \Omega(t) \times \{t\}, \tag{5.1a}$$

$$\nabla \cdot u = 0 \quad \text{in} \quad \bigcup_{t\in(0,T]} \Omega(t) \times \{t\}, \tag{5.1b}$$

$$u = w \quad \text{on} \quad \bigcup_{t\in(0,T]} \partial\Omega(t) \times \{t\}, \tag{5.1c}$$

$$u = u_0 \quad \text{on} \quad \Omega^0 = \Omega(0), \tag{5.1d}$$

where the domain $\Omega(t)$ has a smooth boundary $\Gamma(t) = \partial\Omega(t)$ which moves under a given smooth vector field $w(\cdot, t)$. For well-posedness of system (5.1), velocity field $w$ should satisfy condition $\int_{\partial\Omega(t)} w(\cdot, t) \cdot \mathbf{n} = 0$ for each $t \in [0, T]$, where $\mathbf{n}$ denotes the outward unit

normal vector of $\partial\Omega(t)$. For simplicity, we assume that the vector field $w$ has a smooth extension (which we do not need to know explicitly) to the entire space $\mathbb{R}^d$ and generates a smooth flow map $\Phi(\cdot, t)$ defined on the entire space $\mathbb{R}^d$. The equation also includes a source term $f$, a given smooth function that depends on both space and time variables. In our analysis, the solutions $(u, p)$ of equation (5.1) are assumed to be sufficiently smooth. To ensure uniqueness of the solutions, we assume that $p(\cdot, t) \in L_0^2(\Omega(t))$, which is the space of functions $p$ in $L^2(\Omega(t))$ such that $\int_{\Omega(t)} p \, \mathrm{d}x = 0$.

Recent advancements have brought significant progress in the convergence analysis of FEMs for fluid equations in evolving domains. The well-posedness of the Oseen equation in time-dependent domains was proved in [33] by using an evolving space framework. Lozovskiy et al. [122] introduced a quasi-Lagrangian FEM for NS equations in time-dependent domains, demonstrating optimal-order error estimates in the energy norm. In [130] a $k$th-order unfitted characteristic finite element method(UCFEM) was studied for the time-varying interface problem of two-dimensional Oseen equations. Moreover, Eulerian FEMs for fluid equations have made significant progress. An Eulerian coordinate framework using CutFEM for parabolic equations on moving domains was proposed by Lehrenfeld & Olshanskii [102], while Burman et al. [18] extended this framework to the Stokes equation, proving optimal-order error estimates for the velocity in $L^2H^1$-norm and $L^2L^2$-norm. Further enhanced analysis of related CutFEMs for Stokes and Oseen equation was provided in subsequent studies of von Wahl et al. [170] and Neilan & Olshanskii [139].

Another prevalent method used to handle the complexities arising from domain evolution is the ALE method, which will be employed in this chapter. The ALE method allows the mesh to move according to an ALE mapping, such as the interpolation $\Phi_h$ of $\Phi$, to fit the evolving domain. To employ the ALE formulation, one can define the material derivative of the solution $u$ with respect to the velocity field $w$ as

$$D_t u(x,t) := \frac{\mathrm{d}}{\mathrm{d}t} u(\Phi(\xi,t),t) = \partial_t u + w \cdot \nabla u \text{ at } x = \Phi(\xi,t) \in \Omega(t) \text{ for } \xi \in \Omega^0. \tag{5.2}$$

Using this definition of material derivative, the first two equations in (5.1) can be rewritten as

$$D_t u - w \cdot \nabla u - \Delta u + \nabla p = f, \tag{5.3a}$$

$$\nabla \cdot u = 0, \tag{5.3b}$$

and the ALE method can be employed to discretize the material derivative $D_t u$ along the characteristic lines of the evolving mesh.

In an early investigation of ALE methods, Formaggia & Nobile [143] provided stability results for two different ALE finite element schemes. Subsequently, Gastaldi [53] established a priori error estimates of ALE FEMs for parabolic equations, illustrating that a piecewise linear element can yield $L^2$ error of order $O(h)$ when the mesh size $h$ is sufficiently small. In a related study [142], Nobile obtained an error estimate of $O(h^k)$ in the $L^2$ norm for spatially semidiscrete ALE finite element schemes, with $k$ denoting the degree of the piecewise polynomials utilized. The stability of time-stepping schemes in the context of ALE formulations, such as implicit Euler, Crank–Nicolson, and backward differentiation formulae (BDF), were proved in [12] and [49]. Under specific generalized compatibility conditions and step size restrictions, these investigations yielded $L^2$ error estimates of $O(\tau^s + h^k)$, where $s = 1, 2$ corresponds to the order of the time schemes and $k$ denotes the degree of the finite element space employed. Moreover, Badia & Codina [6] obtained $L^2$ error bounds of $O(\tau^s + \tau^{-1/2}h^{k+1})$ for $s = 1, 2$ for fully discretized ALE methods that employ BDF in time and FEM in space. These sub-optimal error bounds were obtained when the mesh dependent stabilization parameter appearing in fully discrete scheme is as small as the time step size.

Optimal convergence of $O(h^{r+1})$ in the $L^\infty(0, T; L^2)$ norm of ALE semidiscrete FEM for diffusion equations in a bulk domain with a moving boundary was established by Gawlik & Lew in [54] for finite element schemes of degree $r \geqslant 1$. We also refer to [41] and [38] for a unified framework of ALE evolving FEMs and an ALE method with harmonically evolving mesh, respectively. Optimal-order $H^1$ convergence of the ALE FEM for PDEs coupling boundary evolution arising from shape optimziation problems was proved in [56]. These results were established for high-order curved evolving mesh. Optimal convergence of $O(h^{r+1})$ in the $L^\infty(0, T; L^2)$ norm, with flat evolving simplices in the interior and curved simplices exclusively on the boundary, was proved in [113] for the ALE semidiscrete FEM utilizing the standard iso-parametric element of degree $r$ in [103].

In addition to the ALE spatial discretizations mentioned above, the stability and error estimates of discontinuous Galerkin (dG) semi-discretizations in time for diffusion equations in a moving domain using ALE formulations were established in [14] and [13], respectively. The ALE methods for PDEs in bulk domains [56] are also closely related to the evolving FEMs for PDEs on evolving surfaces. Optimal-order convergence in the $L^2$ and $H^1$ norms of evolving FEMs for linear and nonlinear PDEs on evolving surfaces has been shown in [37, 43, 95].

The above-mentioned research efforts have focused on diffusion equations with and with-

out advection terms. The analysis of ALE methods for the Stokes and NS equations has also yielded noteworthy results but remained suboptimal, as discussed below. In [101], Legendre & Takahashi introduced a novel approach that combines the method of characteristics with finite element approximation to the ALE formulation of the Navier–Stokes equations in two dimensions, and established an $L^2$ error estimate of $O(\tau + h^{1/2})$ for the $P_{1b}$–$P_1$ elements under certain restrictions on the time step size. In a related work [151], an error estimate of $O(h^2|\log h|)$ was obtained for the ALE semidiscrete FEM with the Taylor–Hood $P_2$–$P_1$ elements for the Stokes equations in a time-dependent domain. Moreover, for a fully discrete ALE method with the implicit Euler scheme in time, convergence of $O(\tau + h^2 + h^2/\tau)$ was proved in [151]. The errors of ALE finite element solutions to the Stokes equations on a time-varying domain, with BDF-$k$ in time (for $1 \leqslant k \leqslant 5$) and the Taylor–Hood $P_r$–$P_{r-1}$ elements in space (with degree $r \geqslant 2$), were shown to be $O(\tau^k + h^r)$ in the $L^2$ norm in [119].

As far as we know, optimal-order convergence of ALE semidiscrete and fully discrete FEMs were not established for the Stokes and NS equations in an evolving domain. As shown in [54, 113], the optimal-order convergence of ALE semidiscrete FEM requires proving the following optimal-order approximation property for the material derivative of the Ritz projection:

$$\|D_{t,h}R_h u - R_h D_{t,h} u\|_{L^2} \leqslant Ch^{r+1}. \tag{5.4}$$

In line with the fixed domain case, achieving optimal consistency error in analysis of finite element approximation for Stokes equation necessitates the use of the Stokes-Ritz projection $R_h$. As a result, when trying to obtain the optimal-order approximation property (5.4) following the duality argument as in [54, 113], a problem occurs that the error estimate of Stokes-Ritz projection of pressure is involved in the analysis. This problem was addressed by additionally establishing and utilizing an optimal $H^{-1}$ error estimate for the Stokes-Ritz projection of pressure, i.e., (5.43), which is used in Lemma 5.11. This leads to optimal-order convergence of the ALE semidiscrete FEM, as the main result of this chapter (see Theorem 5.1).

A fully discrete second-order projection method along the trajectories of the evolving mesh for decoupling the unknown solutions of velocity and pressure is proposed to compute the numerical solutions in the section of numerical examples.

For simplicity, we focus on the analysis of ALE semidiscrete method for the Stokes equations. However, the numerical scheme and analysis presented in this chapter can be readily extended to the Navier–Stokes equations. The methodologies employed can be effectively utilized to tackle the nonlinear terms as well.

The rest of this chapter is organized as follows. In Section 5.2 we present preliminary results for the evolving mesh, ALE finite element spaces, and boundary-skin estimates. In Section 5.3, we present the formula for the semidiscrete finite element approximation and prove an optimal-order error estimate for it. To reinforce the theoretical analysis, Section 5.4 includes numerical results for the Stokes equation as empirical evidence supporting our theoretical findings.

## 5.2 Preliminary.

**5.2.1 Evolving mesh and ALE finite element spaces.** Suppose that the initial smooth domain $\Omega^0$ is divided into a set $\mathcal{T}_h^0$ of shape-regular and quasi-uniform curved simplices with maximal mesh size $h$. Each curved simplex $K$ is associated with a unique polynomial $F_K$ of degree $r$, referred to as the parametrization of $K$ (as described in [41]). This parametrization maps the reference simplex $\hat{K}$ onto the curved simplex $K$. Additionally, each boundary simplex $K$ (with one face or edge attached to the boundary) may contain a curved face or edge that needs to interpolate the boundary $\Gamma^0 = \partial\Omega^0$. To achieve this interpolation, we employ iso-parametric finite elements of Lenoir's type (see [103] for further details) at time $t = 0$ based on the parametrization of the boundary which is denoted by $\Upsilon : \partial\tilde{D} \to \Gamma^0$. Here, $\partial\tilde{D}$ represents the flat boundary face of the triangulated flat domain, which has the same vertices as the curved triangulated domain $\Omega_h^0 = \bigcup_{K\in\mathcal{T}_h^0} K$. In practical implementations, the parametrization $\Upsilon$ can be chosen as the normal projection onto $\Gamma^0$. In other words, it computes the unique point $\Upsilon(x) \in \Gamma^0$ satisfying the equation:

$$x = \Upsilon(x) + \operatorname{sign}(x, \Omega^0)|x - \Upsilon(x)|\mathbf{n}(\Upsilon(x)),$$

where $\mathbf{n}(\Upsilon(x))$ is the unit outward normal vector at point $\Upsilon(x)$ and

$$\operatorname{sign}(x, \Omega^0) = \begin{cases} 1 & \text{for} \quad x \in \mathbb{R}^d \backslash \overline{\Omega^0}, \\ -1 & \text{for} \quad x \in \Omega^0. \end{cases}$$

Let us denote the nodes of the triangulation $\mathcal{T}_h^0$ as $\xi_j \in \mathbb{R}^d$, where $j = 1, \ldots, N$. Each node $\xi_j$ undergoes a time evolution with velocity $w$, resulting in the movement of the node to a point $x_j(t) \in \mathbb{R}^d$ at time $t$. This evolution is governed by an ordinary differential equation (ODE):

$$\frac{\mathrm{d}}{\mathrm{d}t} x_j(t) = w(x_j(t), t) \quad \text{and} \quad x_j(0) = \xi_j.$$

Consequently, the points $x_j(t)$, where $j = 1, \ldots, N$, constitute the nodes of a time-dependent triangulation denoted as $\mathcal{T}_h(t)$. The relations among these points mirror those among the

original nodes $\xi_j$, namely, a set of nodes $x_j(t)$ form the vertices of a simplex in $\mathcal{T}_h(t)$ if and only if the corresponding nodes $\xi_j$ form the vertices of a simplex in $\mathcal{T}_h^0$. Hence, the evolving domain $\Omega_h(t) = \bigcup_{K\in\mathcal{T}_h(t)} K$ serves as an approximation of the exact domain $\Omega(t)$. This approximation is achieved by employing piecewise polynomial interpolation of degree $r$ on the reference simplex, with an associated interpolation error of $O(h^{r+1})$. Note that the approximation to $\Omega(t)$ by $\Omega_h(t)$ may not be Lenoir's type for $t > 0$.

In a manner similar to the initial triangulation $\mathcal{T}_h^0$, each simplex $K \in \mathcal{T}_h(t)$ is associated with a unique polynomial of degree $r$, denoted as $F_K^t : \hat{K} \to K$, which serves as a parametrization of $K$ over time. Therefore, the finite element space defined on the evolving discrete domain $\Omega_h(t)$ is given by:

$$S_h^r(\Omega_h(t)) := \{v_h \in C(\Omega_h(t)) : v_h \circ F_K^t \in P^r(\hat{K}) \text{ for all } K \in \mathcal{T}_h(t)\},$$

where $P^r(\hat{K})$ represents the set of polynomials on $\hat{K}$ with degree less than or equal to $r$. We denote $V_h^r(\Omega_h(t)) := S_h^r(\Omega_h(t))^d$ as the corresponding vector-valued finite element spaces. The finite element basis functions of $S_h^r(\Omega_h(t))$ are denoted as $\phi_j^t$, where $j = 1, \ldots, N$. These basis functions satisfy the property:

$$\phi_j^t(x_i(t)) = \delta_{ij},\ i, j = 1, \ldots, N.$$

In terms of these basis functions, the approximated flow map $\Phi_h(\cdot, t) \in V_h^r(\Omega_h^0)$ can be expressed as

$$\Phi_h(\xi, t) = \sum_{j=1}^{N} x_j(t)\phi_j^0(\xi) \text{ for } \xi \in \Omega_h^0.$$

The flow map $\Phi_h(\cdot, t)$ establishes a one to one correspondence between $\Omega_h^0$ and $\Omega_h(t)$ at time $t$, with a velocity field $w_h \in V_h^r(\Omega_h(t))$ satisfying:

$$w_h(\Phi_h(\xi, t), t) = \frac{\mathrm{d}}{\mathrm{d}t}\Phi_h(\xi, t) = \sum_{j=1}^{N} w(x_j(t), t)\phi_j^0(\xi) \text{ for } \xi \in \Omega_h^0. \tag{5.5}$$

This representation corresponds to the unique Lagrange interpolation of the exact velocity $w(\Phi(\cdot, t), t)$. Analogous to definition (5.2), we can define the material derivative of any vector or scalar valued function $v$ with respect to the discrete velocity field $w_h$ as follows:

$$D_{t,h}v(x, t) := \frac{\mathrm{d}}{\mathrm{d}t}v(\Phi_h(\xi, t), t) = \partial_t v + \nabla v \cdot w_h \text{ at } x = \Phi_h(\xi, t) \in \Omega_h(t) \text{ for } \xi \in \Omega_h^0. \tag{5.6}$$

The pullback of the finite element basis function $\phi_j^t$ from the domain $\Omega_h(t)$ to $\Omega_h(s)$, i.e., $\phi_j^t \circ \Phi_h(\cdot, t) \circ \Phi_h(\cdot, s)^{-1}$, gives rise to a finite element function defined on $\Omega_h(s)$. Remarkably,

the nodal values of this function coincide with those of $\phi_j^s$. As a result, we establish the equality $\phi_j^t \circ \Phi_h(\cdot,t) \circ \Phi_h(\cdot,s)^{-1} = \phi_j^s$. Exploiting this relationship, we can derive the well-known transport property of the basis function $\phi_j^t$, which states:

$$D_{t,h}\phi_j^t(x) = \frac{\mathrm{d}}{\mathrm{d}t}\phi_j^0(\xi) = 0 \text{ at } x = \Phi_h(\xi,t). \tag{5.7}$$

The analysis of integrals over dynamically evolving domains necessitates the application of the Transport Theorem, as established in [171, Lemma 5.7]. This pivotal theorem provides a concise and indispensable description of the intrinsic relationship between the time derivative of an integral over a domain that evolves with time and the derivatives of the integrated function and domain velocity.

**Lemma 5.1** (Transport Theorem)**.** *If the domain $\Omega$ undergoes motion with a velocity field $w \in W^{1,\infty}(\Omega)$, we have*

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega f \,\mathrm{d}x = \int_\Omega D_t f + f\nabla\cdot w \,\mathrm{d}x, \tag{5.8}$$

*where $D_t f$ is the material derivative of $f$ with respect to the velocity $w$.*

The interaction between the operators $D_t$ and $\nabla$ plays an essential role in the error analysis. Consequently, we establish the following lemma as a direct consequence of (5.2):

**Lemma 5.2.** *For any vector-valued function $f$, the material derivative of $\nabla f$ and $\nabla\cdot f$ with respect to the velocity field $w$ can be expressed as follows:*

$$D_t\nabla f = \nabla D_t f - \nabla f\nabla w, \tag{5.9}$$

$$D_t\nabla\cdot f = \nabla\cdot D_t f - (\nabla f) : (\nabla w)^\top. \tag{5.10}$$

We adopt the Taylor–Hood type finite element spaces on the evolving domain $\Omega(t)$, which allow for a continuous approximation of the pressure. Specifically, we define the following spaces:

$$\mathring{V}_h^r(\Omega_h(t)) := \{u \in V_h^r(\Omega_h(t)) : u|_{\partial\Omega_h(t)} = 0\},$$
$$Q_h^{r-1}(\Omega_h(t)) := \{p \in S_h^{r-1}(\Omega_h(t)) : \int_{\Omega_h(t)} p \,\mathrm{d}x = 0\}.$$

By employing Verfürth's trick and utilizing the macros-element criterion, as described in [11, Section 8.5 and Section 8.8], we establish the inf-sup condition for the Taylor–Hood type isoparametric elements.

**Lemma 5.3** (Inf-sup condtion)**.** *There exists a constant $\kappa > 0$, independent of $h$ and $t \in [0,T]$ for $r \geqslant 2$, such that*

$$\sup_{0\neq v_h\in \mathring{V}_h^r(\Omega_h(t))} \frac{(\mathrm{div} v_h, p_h)_{\Omega_h(t)}}{\|\nabla v_h\|_{L^2(\Omega_h(t))}} \geqslant \kappa \|p_h\|_{L^2(\Omega_h(t))} \quad \forall p_h \in Q_h^{r-1}(\Omega_h(t)). \tag{5.11}$$

**5.2.2 Boundedness of partial derivatives of the mesh velocity.** For any function $u$ defined on $\bigcup_{0\leqslant t\leqslant T} \Omega(t)\times\{t\}$, there is an extension function $\widetilde{u}$ defined on $\mathbb{R}^d\times[0,T]$ such that

$$\widetilde{u}(\cdot,t) := E(u(\cdot,t)\circ\Phi(\cdot,t))\circ\Phi(\cdot,t)^{-1}, \tag{5.12}$$

where the operator $E: L^1(\Omega(0)) \to L^1(\mathbb{R}^d)$ refers to Stein's extension operator in [161, p. 181, Theorem 5]. It holds that

$$\|\widetilde{u}(\cdot,t)\|_{W^{k,p}(\mathbb{R}^d)} \leqslant C\|u(\cdot,t)\|_{W^{k,p}(\Omega(t))}. \tag{5.13}$$

Similarly, we can define the function $\widetilde{p}$ as the extension of $p$ to the whole space $\mathbb{R}^d$. To simplify the notation, we will just use $(u,p)$ to represent $(\widetilde{u},\widetilde{p})$ if there is no confusion arisen within the context.

We denote the interpolation operators as $I_h : C(\Omega_h(t)) \to S_h^r(\Omega_h(t))$. Throughout this discussion, the explicit time dependency $t$ is often omitted, and we will use $I_h$ instead.

In certain cases, we come across vector-valued spaces such as $V_h^r(\Omega_h(t)) = S_h^r(\Omega_h(t))^d$ and the corresponding vector-valued interpolation operators such as $I_h^d$. To streamline the notation, we will use $I_h$ when referring to vector-valued objects, provided there is no ambiguity within the context. In the same spirit, we use notation like $\|\cdot\|_{H^1(\Omega(t))}$ instead of $\|\cdot\|_{H^1(\Omega(t))^d}$ when referring to norms of vector-valued objects.

By (5.5), the interpolation $w_h = I_h w$ serves as an approximation of $w$. Consequently, we can establish an error estimate for $w_h$ in the piece-wise Sobolev norm $W_h^{k,\infty}(\Omega_h(t))$ with respect to triangulation $\mathcal{T}_h(t)$ as follows:

$$\|w_h(\cdot,t) - w(\cdot,t)\|_{W_h^{k,\infty}(\Omega_h(t))} \leqslant Ch^{r+1-k}\|w(\cdot,t)\|_{W^{r+1,\infty}(\Omega_h(t))} \quad \forall\, 0\leqslant k\leqslant r+1, \tag{5.14}$$

which especially implies the $W^{1,\infty}$-boundedness of the discrete velocity $w_h$. Observe that $\Phi(\xi_j,t) = \Phi_h(\xi_j,t)$ for each nodes $\xi_j$ of discrete domain $\Omega_h^0$, thus there holds $\Phi_h(\cdot,t) = I_h\Phi(\cdot,t)$ on $\Omega_h^0$ and we can derive the error between $\Phi(t)$ and $\Phi_h(t)$ as follows:

$$\|\Phi(\cdot,t) - \Phi_h(\cdot,t)\|_{W^{1,\infty}(\Omega_h^0)} \leqslant Ch^r. \tag{5.15}$$

The estimates in (5.14) and (5.15) lead to the following result when $h$ is sufficiently small

$$\|w_h(\cdot,t)\|_{W^{1,\infty}(\Omega_h(t))} + \|\Phi_h(\cdot,t)\|_{W^{1,\infty}(\Omega_h^0)} + \|\Phi_h^{-1}(\cdot,t)\|_{W^{1,\infty}(\Omega_h(t))} \leqslant C, \tag{5.16}$$

where $C$ is a constant independent of the mesh size $h$ and time $t$. This serves as a basic condition on the mesh velocity in the subsequent analysis.

**5.2.3 Error of domain approximation.** To address the discrepancy between $\Omega(t)$ and its finite element approximation $\Omega_h(t)$, we utilize the boundary-skin estimate. This estimate is essential for effectively managing errors that arise from the finite element approximation of the domain.

**Lemma 5.4.** *For any finite element function $v_h \in \mathring{V}_h^r(\Omega_h(t))$, the following inequalities hold:*

$$\|v_h\|_{L^2(\Omega_h(t)\backslash\Omega(t))} \leqslant Ch^{3(r+1)/2-d/2}\|\nabla v_h\|_{L^2(\Omega_h(t))} \leqslant Ch^{3(r+1)/2-d/2-1}\|v_h\|_{L^2(\Omega_h(t))}.$$

*Proof.* Using Hölder's inequality, Newton-Leibniz formula and the fact $v_h|_{\partial\Omega_h(t)} = 0$, we have

$$\begin{aligned}
\|v_h\|_{L^2(\Omega_h(t)\backslash\Omega(t))} &\leqslant |\Omega_h(t)\backslash\Omega(t)|^{1/2}\|v_h\|_{L^\infty(\Omega_h(t)\backslash\Omega(t))} \\
&\leqslant |\Omega_h(t)\backslash\Omega(t)|^{1/2} \sup_{x\in\Omega_h(t)\backslash\Omega(t)} \operatorname{dist}(x,\partial\Omega_h)\|\nabla v_h\|_{L^\infty(\Omega_h(t))} \\
&\leqslant Ch^{3(r+1)/2}\|\nabla v_h\|_{L^\infty(\Omega_h(t))} \\
&\leqslant Ch^{3(r+1)/2-d/2}\|\nabla v_h\|_{L^2(\Omega_h(t))} \leqslant Ch^{3(r+1)/2-d/2-1}\|v_h\|_{L^2(\Omega_h(t))},
\end{aligned}$$

where we used the fact that the distance from $x \in \Omega_h\backslash\Omega$ to $\partial\Omega_h(t)$ is no greater than $Ch^{r+1}$ and the inverse estimate of finite element functions in the last two inequalities. □

Due to the inherent discrepancy between the finite element domain $\Omega_h(t)$ and the exact domain $\Omega(t)$, the exact solution $u$ does not vanish on $\partial\Omega_h(t)$. To handle this situation, we rely on the following lemma to derive an estimate for the integral over the boundary $\partial\Omega_h(t)$. A proof of this lemma can be found in [113, eq. (3.32)].

**Lemma 5.5.** *Let $g \in W^{1,1}(\mathbb{R}^d)$. Then the following inequality holds:*

$$\|g\|_{L^1(\partial\Omega_h(t))} \leqslant C\|g\|_{L^1(\partial\Omega(t))} + C\|\nabla g\|_{L^1(\Omega(t)\cup\Omega_h(t))}, \tag{5.17}$$

*where $C$ is a constant independent of the mesh size $h$ and time $t$.*

The significance of the ensuing lemma lies in its pivotal role in acquiring optimal $H^{-1}$-norm estimates for pressure through implementation of a duality argument. A rigorous proof of this lemma can be found in [46, Corollary 1.5].

**Lemma 5.6.** *For each $\lambda \in H^1(\Omega(t)) \cap L_0^2(\Omega(t))$, there is a function $\chi \in H^2(\Omega(t))^d \cap H_0^1(\Omega(t))^d$ such that $\mathrm{div}\chi = \lambda$, and the following inequality holds:*

$$\|\chi\|_{H^2(\Omega(t))} \leqslant C\|\lambda\|_{H^1(\Omega(t))},$$

*where the constant $C$ is independent of $t \in [0,T]$.*

**5.3 The semidiscrete finite element approximation.** In this section, we present a semidiscrete FEM for solving problem (5.1) using the ALE formulation (5.3). We also provide an error estimate for this method.

The semidiscrete finite element problem can be formulated as follows: Seek solutions $u_h(t) \in V_h^r(\Omega_h(t))$ with initial value $u_h(0) = I_h u(0)$ and the boundary condition $u_h = w_h$ on $\partial\Omega_h(t)$, and $p_h(t) \in Q_h^{r-1}(\Omega_h(t))$ that satisfy the following equations for all test functions $v_h \in \mathring{V}_h^r(\Omega_h(t))$ and $q_h \in Q_h^{r-1}(\Omega_h(t))$:

$$(D_{t,h}u_h - w_h \cdot \nabla u_h, v_h)_{\Omega_h(t)} + (\nabla u_h, \nabla v_h)_{\Omega_h(t)} - (\nabla \cdot v_h, p_h)_{\Omega_h(t)} = (f, v_h)_{\Omega_h(t)}, \quad (5.18a)$$

$$(\nabla \cdot u_h, q_h)_{\Omega_h(t)} = 0, \quad (5.18b)$$

The main result of this section is the following theorem.

**Theorem 5.1** (Error estimate of the semidiscrete FEM)**.** *Consider the semidiscrete finite element solutions $(u_h, p_h)$ given by (5.18). Assuming that the exact solutions $(u,p)$ to problem (5.1) are sufficiently smooth and have been extended to be defined on $\mathbb{R}^d$ via (5.12), the following estimate holds under condition that $w$ is sufficiently smooth:*

$$\sup_{t\in[0,T]} \|u - u_h\|_{L^2(\Omega_h(t))} \leqslant C R_{u,p} h^{r+1}, \quad (5.19)$$

$$\|p - p_h\|_{L^2(0,T;L^2(\Omega_h(t)))} \leqslant C R_{u,p} h^{r}. \quad (5.20)$$

*where $C$ is a constant independent of the mesh size $h$ and $R_{u,p}$ is a norm of $(u,p)$ defined as follows:*

$$\begin{aligned} R_{u,p} :=& \|\partial_t u\|_{L^2(0,T;W^{r+1,\infty}(\mathbb{R}^d)} + \|u\|_{L^2(0,T;W^{r+2,\infty}(\mathbb{R}^d))} \\ &+ \|\partial_t p\|_{L^2(0,T;H^r(\mathbb{R}^d))} + \|p\|_{L^2(0,T;H^{r+1}(\mathbb{R}^d))} \\ &+ \|u\|_{L^\infty(0,T;W^{r+1,\infty}(\mathbb{R}^d))} + \|p\|_{L^\infty(0,T;H^r(\mathbb{R}^d))}. \end{aligned}$$

**5.3.1 The Stokes–Ritz projection.** Analogous to the Stokes–Ritz projection in a fixed domain, we introduce the concept of the Stokes–Ritz projection for the pair $(v(\cdot,t),q(\cdot,t)) \in H^1(\Omega_h(t)) \times L^2(\Omega_h(t))$ for $t \in [0,T]$ over a time-dependent finite element domain $\Omega_h(t)$, denoted as $(R_h v, R_h q) \in V_h^r(\Omega_h(t)) \times Q_h^{r-1}(\Omega_h(t))$. The Stokes–Ritz projection satisfies the following equations for all test functions $\chi_h \in \mathring{V}_h^r(\Omega_h(t))$ and $\lambda_h \in Q_h^{r-1}(\Omega_h(t))$ under the boundary condition $R_h v = I_h v$ on $\partial\Omega_h(t)$:

$$(\nabla R_h v, \nabla \chi_h)_{\Omega_h(t)} - (\nabla \cdot \chi_h, R_h q)_{\Omega_h(t)} = (\nabla v, \nabla \chi_h)_{\Omega_h(t)} - (\nabla \cdot \chi_h, q)_{\Omega_h(t)}, \tag{5.21a}$$

$$(\nabla \cdot R_h v, \lambda_h)_{\Omega_h(t)} = (\nabla \cdot v, \lambda_h)_{\Omega_h(t)}. \tag{5.21b}$$

Additionally, we define the norm $\|\cdot\|'$ over any domain $D \subset \mathbb{R}^d$ as follows:

$$\|f\|'_{L^2(D)} := \|f - \bar{f}\|_{L^2(D)}, \tag{5.22}$$

where $\bar{f}$ denotes the average of $f$ over $D$, given by $\bar{f} := \frac{1}{|D|}\int_D f \,\mathrm{d}x$.

By utilizing the inf-sup condition (5.11), the Stokes–Ritz projection exhibits quasi-optimal error estimates, as stated in the following lemma:

**Lemma 5.7.** *[55, Chapter 2, Theorem 1.1] Let $(R_h v, R_h q)$ denote the Stokes–Ritz projections of $(v,q)$. Suppose that $(v,q)$ are sufficiently smooth. Then the following estimate holds*

$$\|\nabla(v - R_h v)\|_{L^2(\Omega_h(t))} + \|q - R_h q\|'_{L^2(\Omega_h(t))} \leqslant C h^r \Big( \|v\|_{H_h^{r+1}(\Omega_h(t))} + \|q\|_{H_h^r(\Omega_h(t))} \Big),$$

*where $H_h^r(\Omega_h(t))$ means the piece-wise Sobolev norm with respect to the mesh $\mathcal{T}_h(t)$. The constant $C$ is independent of $h, t$ and the function $(v,q)$.*

**Remark 5.1.** Lemma 5.7 is a corollary of [55, Chapter 2, Theorem 1.1] and error estimates of Lagrange interpolation:

$$\|\nabla(v - I_h v)\|_{L^2(\Omega_h(t))} \leqslant C h^k \|v\|_{H_h^{k+1}(\Omega_h(t))}, \quad \|q - I_h q\|_{L^2(\Omega_h(t))} \leqslant C h^k \|q\|_{H_h^k(\Omega_h(t))},$$

where $k$ is restricted by condition that $\frac{d}{2} < k \leqslant r$ due to the requirement of Sobolev embedding $H^k(\mathbb{R}^d) \hookrightarrow C^0(\mathbb{R}^d)$ for the stability of Lagrange interpolation. To obtain Lemma 5.7, it suffices to take $k = r$. Since $r \geqslant 2$ and $d \in \{2,3\}$ by our assumption, the restriction $k = r > \frac{d}{2}$ is satisfied. However, if $q$ only possesses $H^1$-regularity, as is the case for the solution $\varphi$ of the duality problem (5.32), there is no desired estimate of Lagrange interpolation error of $q$. To overcome this problem, we can consider the Scott-Zhang interpolation $\mathcal{I}_h$ (cf. [155] and [15, Section 4.8]). Though we are working with finite element space consisting

of isoparametric elements, the same strategy as in [155, Theorem 3.1] still applies to prove the following first-order error estimate:

$$\|q - \mathcal{I}_h q\|_{L^2(\Omega_h(t))} \leqslant Ch\|q\|_{H^1(\Omega_h(t))}. \tag{5.23}$$

As a corollary, let $\mathbb{P}_h : L^2(\Omega_h(t)) \to S_h^r(\Omega_h(t))$ be the $L^2(\Omega_h(t))$-orthogonal projection onto the finite element space $S_h^r(\Omega_h(t))$. Then, there holds:

$$\|q - \mathbb{P}_h q\|_{L^2(\Omega_h(t))} \leqslant \|q - \mathcal{I}_h q\|_{L^2(\Omega_h(t))} \leqslant Ch\|q\|_{H^1(\Omega_h(t))} \tag{5.24}$$

We shall utilize (5.24) in our duality argument contained in Lemma 5.9 and Lemma 5.11 below.

In order to prove the optimal-order estimate of the error between exact solutions and numerical solutions, we need to facilitate the estimation of errors such as $D_{t,h}(v - R_h v)$ and $D_{t,h}(q - R_h q)$. It is convenient to introduce the operator $E_{t,h}$ defined as:

$$E_{t,h} := D_{t,h}R_h - R_h D_{t,h}. \tag{5.25}$$

We can establish the following lemma about the estimate of $\nabla E_{t,h} v$ and $E_{t,h} q$.

**Lemma 5.8.** *Let $(R_h v, R_h q)$ denote the Stokes–Ritz projections of $(v, q)$. Suppose that $(v, q)$ are sufficiently smooth. There is a constant $C$ independent of $h, t$ and the function $(v, q)$ so that the following estimate holds:*

$$\|\nabla E_{t,h} v\|_{L^2(\Omega_h(t))} + \|E_{t,h} q\|'_{L^2(\Omega_h(t))} \leqslant Ch^r \left( \|v\|_{H_h^{r+1}(\Omega_h(t))} + \|q\|_{H_h^r(\Omega_h(t))} \right). \tag{5.26}$$

*Proof.* Since equations (5.21) are invariant under the substitution $q$ to $q - \overline{q}$ with $\overline{q}$ being the average of $q$ over $\Omega_h(t)$, it suffices to assume that $\overline{q} = 0$. Now, we fix a time $t \in [0, T]$ and a pair of testing functions $\chi_h \in \mathring{V}_h^r(\Omega_h(t))$ and $\lambda_h \in Q_h^{r-1}(\Omega_h(t))$. From (5.21), the following equation holds for each $s \in [0, T]$:

$$(\nabla(R_h v - v)(s), \nabla\chi_h(s))_{\Omega_h(s)} - (\nabla \cdot \chi_h(s), (R_h q - q)(s))_{\Omega_h(s)} = 0, \tag{5.27a}$$

$$(\nabla \cdot (R_h v - v)(s), \lambda_h(s) - \overline{\lambda_h(s)})_{\Omega_h(s)} = 0, \tag{5.27b}$$

where $\chi_h(s) \in \mathring{V}_h^r(\Omega_h(s))$ and $\lambda_h(s) \in S_h^{r-1}(\Omega_h(s))$ are defined by $\chi_h(s) := \chi_h(t) \circ \phi_h(t) \circ (\phi_h(s))^{-1}$ and $\lambda_h(s) := \lambda_h(t) \circ \phi_h(t) \circ (\phi_h(s))^{-1}$, i.e. the finite element functions on $\Omega_h(s)$ with the same nodal values as $\chi_h$ and $\lambda_h$ respectively. Note that by definition $D_{t,h}\chi_h(s) = D_{t,h}\lambda_h(s) = 0$ for all $s \in [0, T]$ and $\overline{\lambda_h(t)} = \overline{\lambda_h} = 0$ but in general $\overline{\lambda_h(s)} \neq 0$. By taking

derivative with respective to time $s$ at $s = t$ on both sides of (5.27), and using Lemma 5.1 and Lemma 5.2, we obtain

$$\begin{aligned}
&(\nabla D_{t,h}R_hv, \nabla\chi_h)_{\Omega_h(t)} - (\nabla\cdot\chi_h, D_{t,h}R_hq)_{\Omega_h(t)} - (\nabla D_{t,h}v, \nabla\chi_h)_{\Omega_h(t)} + (\nabla\cdot\chi_h, D_{t,h}q)_{\Omega_h(t)}\\
&= -(\nabla(v - R_hv)\nabla w_h, \nabla\chi_h)_{\Omega_h(t)} + (\nabla(v - R_hv), \nabla\chi_h(\nabla\cdot w_h - \nabla w_h))_{\Omega_h(t)}\\
&\quad + (\nabla\chi_h : (\nabla w_h)^\top - \nabla\cdot\chi_h\nabla\cdot w_h, q - R_hq)_{\Omega_h(t)}, \qquad (5.28a)\\
&(\nabla\cdot D_{t,h}R_hv, \lambda_h)_{\Omega_h(t)} - (\nabla\cdot D_{t,h}v, \lambda_h)_{\Omega_h(t)}\\
&= -(\nabla(v - R_hv) : (\nabla w_h)^\top, \lambda_h)_{\Omega_h(t)} + (\nabla\cdot(v - R_hv), \lambda_h\nabla\cdot w_h)_{\Omega_h(t)}\\
&\quad + (\nabla\cdot(R_hv - v), 1)_{\Omega_h(t)}\frac{(\lambda_h, \nabla\cdot w_h)_{\Omega_h(t)}}{|\Omega_h(t)|}. \qquad (5.28b)
\end{aligned}$$

Similarly to the definition of (5.21), we can define the Stokes–Ritz projection of $(D_{t,h}v, D_{t,h}q)$, and substitute the definition into (5.28), we obtain

$$\begin{aligned}
&(\nabla E_{t,h}v, \nabla\chi_h)_{\Omega_h(t)} - (\nabla\cdot\chi_h, E_{t,h}q)_{\Omega_h(t)}\\
&= -(\nabla(v - R_hv)\nabla w_h, \nabla\chi_h)_{\Omega_h(t)} + (\nabla(v - R_hv), \nabla\chi_h\nabla\cdot w_h - \nabla\chi_h\nabla w_h)_{\Omega_h(t)}\\
&\quad + (\nabla\chi_h : (\nabla w_h)^\top - \nabla\cdot\chi_h\nabla\cdot w_h, q - R_hq)_{\Omega_h(t)}, \qquad (5.29a)\\
&(\nabla\cdot E_{t,h}v, \lambda_h)_{\Omega_h(t)}\\
&= -(\nabla(v - R_hv) : (\nabla w_h)^\top, \lambda_h)_{\Omega_h(t)} + (\nabla\cdot(v - R_hv), \lambda_h\nabla\cdot w_h)_{\Omega_h(t)}\\
&\quad + (\nabla\cdot(R_hv - v), 1)_{\Omega_h(t)}\frac{(\lambda_h, \nabla\cdot w_h)_{\Omega_h(t)}}{|\Omega_h(t)|}. \qquad (5.29b)
\end{aligned}$$

By the definition of the Stokes–Ritz projection, $R_hv = I_hv$ and $R_hD_{t,h}v = I_hD_{t,h}v$ on the boundary $\partial\Omega_h(t)$. Then $D_{t,h}R_hv = D_{t,h}I_hv = I_hD_{t,h}v$ on $\partial\Omega_h(t)$, which means $E_{t,h}v = 0$ on $\partial\Omega_h(t)$. Hence, we can choose $\chi_h = E_{t,h}v$ and $\lambda_h = E_{t,h}q - \overline{E_{t,h}q}$ in equation (5.29) with $\overline{E_{t,h}q}$ being the average of $E_{t,h}q$ over $\Omega_h(t)$, and obtain the following estimate by using the $W^{1,\infty}$ boundedness of $w_h$

$$\begin{aligned}
\|\nabla E_{t,h}v\|^2_{L^2(\Omega_h(t))} \leqslant& C\|\nabla(R_hv - v)\|^2_{L^2(\Omega_h(t))} + C\|q - R_hq\|^2_{L^2(\Omega_h(t))}\\
&+ C\|E_{t,h}q - \overline{E_{t,h}q}\|_{L^2(\Omega_h(t))}\|\nabla(R_hv - v)\|_{L^2(\Omega_h(t))}. \qquad (5.30)
\end{aligned}$$

By using the inf-sup condition (5.11) and the equation (5.29a), we have

$$\begin{aligned}
\|E_{t,h}q - \overline{E_{t,h}q}\|_{L^2(\Omega_h(t))} \leqslant& C\sup_{0\neq\chi_h\in\mathring{V}_h^r}\frac{(\nabla\cdot\chi_h, E_{t,h}q - \overline{E_{t,h}q})}{\|\nabla\chi_h\|_{L^2(\Omega_h(t))}} = C\sup_{0\neq\chi_h\in\mathring{V}_h^r}\frac{(\nabla\cdot\chi_h, E_{t,h}q)}{\|\nabla\chi_h\|_{L^2(\Omega_h(t))}}\\
\leqslant& C\Big(\|\nabla E_{t,h}v\|_{L^2(\Omega_h(t))} + \|\nabla(R_hv - v)\|_{L^2(\Omega_h(t))} + \|q - R_hq\|_{L^2(\Omega_h(t))}\Big). \qquad (5.31)
\end{aligned}$$

By substituting (5.31) into (5.30), and using Young's inequality and Lemma 5.7 under the assumption $\overline{q} = 0$, we obtain the desired result. □

**5.3.2 The Nitche's trick and duality argument.** In order to obtain an optimal order error estimate of $R_h u - u$, we will apply Nitche's trick. Let $g_h$ be a function in $\mathring{V}_h^r(\Omega_h(t))$ that we can extend outside of $\Omega_h(t)$ by setting it to zero. We solve the following equations in $\Omega(t)$ for $(\psi, \varphi) \in H_0^1(\Omega(t)) \times L_0^2(\Omega(t))$:

$$-\Delta\psi + \nabla\varphi = g_h \quad \text{in } \Omega(t), \tag{5.32a}$$

$$\nabla \cdot \psi = 0 \quad \text{in } \Omega(t), \quad \psi|_{\partial\Omega(t)} = 0. \tag{5.32b}$$

By applying regularity estimates for the Stokes equation in $\Omega(t)$, we obtain the following result:

$$\|\psi\|_{H^2(\Omega(t))} + \|\nabla\varphi\|_{L^2(\Omega(t))} \leqslant C\|g_h\|_{L^2(\Omega_h(t))}. \tag{5.33}$$

To extend the functions $\psi$ and $\varphi$ to $\widetilde{\psi}$ and $\widetilde{\varphi}$, respectively, we employ the Stein extension operator as in (5.12). By applying this operator, we can define $\widetilde{\eta}$ as $\widetilde{\eta} := -\Delta\widetilde{\psi} + \nabla\widetilde{\varphi} - g_h$ and arrive at the following expression:

$$\|g_h\|_{L^2(\Omega_h(t))}^2 = (\nabla\widetilde{\psi}, \nabla g_h)_{\Omega_h(t)} - (\nabla \cdot g_h, \widetilde{\varphi})_{\Omega_h(t)} - (g_h, \widetilde{\eta})_{\Omega_h(t)}. \tag{5.34}$$

Notably, since $\widetilde{\eta}$ vanishes in $\Omega(t)$ and we have $r \geqslant 2$, we can utilize Lemma 5.4 along with the regularity estimate (5.33) to obtain the following inequality:

$$|(g_h, \widetilde{\eta})_{\Omega_h(t)}| = |(g_h, \widetilde{\eta})_{\Omega_h(t)\setminus\Omega(t)}| \leqslant \|\widetilde{\eta}\|_{L^2(\mathbb{R}^d)}\|g_h\|_{L^2(\Omega_h(t)\setminus\Omega(t))} \leqslant Ch^2\|g_h\|_{L^2(\Omega_h(t))}^2. \tag{5.35}$$

Consequently, when $h > 0$ is sufficiently small, we can absorb $|(g_h, \widetilde{\eta})_{\Omega_h(t)}|$ on the right-hand side of (5.34) by the left-hand side. This yields the following estimate:

$$\|g_h\|_{L^2(\Omega_h(t))}^2 \leqslant C\Big|(\nabla\widetilde{\psi}, \nabla g_h)_{\Omega_h(t)} - (\nabla \cdot g_h, \widetilde{\varphi})_{\Omega_h(t)}\Big|. \tag{5.36}$$

We choose $g_h = R_h v - I_h v \in \mathring{V}_h^r(\Omega_h(t))$ in (5.32). By appropriately estimating the right-hand side of (5.36), we can derive the following lemma.

**Lemma 5.9.** *Let $(R_h v, R_h q)$ be the Stokes–Ritz projection of $(v, q)$. Suppose that $(v, q)$ are sufficiently smooth. Then there exists a constant $C$ independent of $h$, $t$ and $(v, q)$ such that*

$$\|R_h v - v\|_{L^2(\Omega_h(t))} \leqslant Ch^{r+1}\left(\|v\|_{W_h^{r+1,\infty}(\Omega_h(t))} + \|q\|_{H_h^r(\Omega_h(t))}\right). \tag{5.37}$$

*Proof.* For $g_h = R_h v - I_h v$, by utilizing (5.36), the definition of Stokes–Ritz projection (5.21) with $(\chi_h, \lambda_h) = (I_h\widetilde{\psi}, \mathbb{P}_h^*\widetilde{\varphi})$ (noting that $I_h\widetilde{\psi}|_{\partial\Omega_h(t)} = 0$ since $\widetilde{\psi}|_{\partial\Omega(t)} = 0$ and thus $\tilde{\psi}$ vanishes on all the boundary nodes of $\Omega_h(t)$), and integration by parts, we have

$$\begin{aligned}
\|g_h\|^2_{L^2(\Omega_h(t))} \leqslant & C\Big|(\nabla\widetilde{\psi}, \nabla(R_h v - v))_{\Omega_h(t)} - (\nabla\cdot(R_h v - v), \widetilde{\varphi})_{\Omega_h(t)}\Big| \\
& + C\Big|(\nabla\widetilde{\psi}, \nabla(v - I_h v))_{\Omega_h(t)} - (\nabla\cdot(v - I_h v), \widetilde{\varphi})_{\Omega_h(t)}\Big| \\
\leqslant & C\Big|(\nabla(\widetilde{\psi} - I_h\widetilde{\psi}), \nabla(R_h v - v))_{\Omega_h(t)} + (\nabla\cdot I_h\widetilde{\psi}, R_h q - q - \overline{q})_{\Omega_h(t)}\Big| \\
& + C\Big|(\nabla\cdot(R_h v - v), \widetilde{\varphi} - \mathbb{P}_h^*\widetilde{\varphi})_{\Omega_h(t)}\Big| \\
& + C\Big|(-\Delta\widetilde{\psi} + \nabla\widetilde{\varphi}, v - I_h v)_{\Omega_h(t)} + (\nabla\widetilde{\psi}\cdot\mathbf{n} - \widetilde{\varphi}\mathbf{n}, v - I_h v)_{\partial\Omega_h(t)}\Big|,
\end{aligned}$$

where we used following notations: $\mathbb{P}_h\widetilde{\varphi}$ is the $L^2(\Omega_h(t))$-orthogonal projection of $\widetilde{\varphi}$ onto the space $S_h^{r-1}(\Omega_h(t))$ and $\overline{\mathbb{P}_h\widetilde{\varphi}} = \overline{\widetilde{\varphi}}$ is the average of $\mathbb{P}_h\widetilde{\varphi}$ on $\Omega_h(t)$ so that $\mathbb{P}_h^*\widetilde{\varphi} := \mathbb{P}_h\widetilde{\varphi} - \overline{\mathbb{P}_h\widetilde{\varphi}}$ belongs to $Q_h^{r-1}(\Omega_h(t))$. For $L^2$-orthogonal projection $\mathbb{P}_h$, there holds error estimate (cf. (5.24) of Remark 5.1)

$$\|\mathbb{P}_h\widetilde{\varphi} - \widetilde{\varphi}\|_{L^2(\Omega_h(t))} \leqslant Ch\|\widetilde{\varphi}\|_{H^1(\Omega_h(t))}. \tag{5.38}$$

And we can deduce following estimate for $|\overline{\widetilde{\varphi}}|$ from condition $\varphi \in L_0^2(\Omega(t))$,

$$\begin{aligned}
|\overline{\widetilde{\varphi}}| \leqslant & C\left(\int_{\Omega_h(t)\backslash\Omega(t)} |\widetilde{\varphi}| + \int_{\Omega(t)\backslash\Omega_h(t)} |\widetilde{\varphi}|\right) \\
\leqslant & C\|\widetilde{\varphi}\|_{L^6(\mathbb{R}^d)}(|\Omega(t)\backslash\Omega_h(t)|^{1/3} + |\Omega_h(t)\backslash\Omega(t)|^{1/3}) \\
\leqslant & Ch\|\widetilde{\varphi}\|_{H^1(\mathbb{R}^d)}.
\end{aligned} \tag{5.39}$$

As a corollary, we have

$$\|\mathbb{P}_h^*\widetilde{\varphi} - \widetilde{\varphi}\|_{L^2(\Omega_h(t))} \leqslant Ch\|\widetilde{\varphi}\|_{H^1(\Omega_h(t))}. \tag{5.40}$$

By using the error estimate of interpolation $I_h$, the error estimate of $\mathbb{P}_h^*$ (5.40), Lemma 5.7, and the regularity result (5.33), we have

$$\begin{aligned}
& \|g_h\|^2_{L^2(\Omega_h(t))} \\
\leqslant & C\Big(\|\widetilde{\psi}\|_{H^2(\mathbb{R}^d)} + \|\widetilde{\varphi}\|_{H^1(\mathbb{R}^d)}\Big)\Big(h\|\nabla(R_h v - v)\|_{L^2(\Omega_h(t))} + \|v - I_h v\|_{L^2(\Omega_h(t))}\Big) \\
& + C\Big|(\nabla\cdot I_h\widetilde{\psi}, q - R_h q - \overline{q})_{\Omega_h(t)}\Big| + C\Big|(\mathbf{n}\cdot\nabla\widetilde{\psi} - \widetilde{\varphi}\mathbf{n}, v - I_h v)_{\partial\Omega_h(t)}\Big| \\
\leqslant & Ch^{r+1}\Big(\|v\|_{H^{r+1}(\Omega_h(t))} + \|q\|_{H^r(\Omega_h(t))}\Big)\|g_h\|_{L^2(\Omega_h(t))} + C\Big|(\nabla\cdot(\widetilde{\psi} - I_h\widetilde{\psi}), q - R_h q - \overline{q})_{\Omega_h(t)}\Big|
\end{aligned}$$

$$+C\Big|(\nabla\cdot\widetilde{\psi},q-R_hq-\overline{q})_{\Omega_h(t)}\Big|+C\Big|(\nabla\widetilde{\psi}\cdot\mathbf{n}-\widetilde{\varphi}\mathbf{n},v-I_hv)_{\partial\Omega_h(t)}\Big|.$$

We estimate the left terms subsequently.

$$\Big|(\nabla\cdot(\widetilde{\psi}-I_h\widetilde{\psi}),q-R_hq-\overline{q})_{\Omega_h(t)}\Big|\leqslant Ch\|\widetilde{\psi}\|_{H^2(\Omega_h(t))}\|q-R_hq\|'_{L^2(\Omega_h(t))}.$$

It is known that $\nabla\cdot\widetilde{\psi}=0$ in $\Omega(t)$. Hence, we have

$$\begin{aligned}\Big|(\nabla\cdot\widetilde{\psi},q-R_hq-\overline{q})_{\Omega_h(t)}\Big|\leqslant&\|\nabla\cdot\widetilde{\psi}\|_{L^2(\Omega_h(t)\backslash\Omega(t))}\|q-R_hq\|'_{L^2(\Omega_h(t))}\\ \leqslant&C|\Omega_h(t)\backslash\Omega(t)|^{\frac{1}{3}}\|\nabla\cdot\widetilde{\psi}\|_{L^6(\mathbb{R}^d)}\|q-R_hq\|'_{L^2(\Omega_h(t))}\\ \leqslant&Ch\|\widetilde{\psi}\|_{H^2(\mathbb{R}^d)}\|q-R_hq\|'_{L^2(\Omega_h(t))},\end{aligned}\tag{5.41}$$

where the last inequality follows from the fact $r\geqslant 2$, Lemma 5.7, and (5.33). By using Lemma 5.5 and (5.33), we have

$$\begin{aligned}\Big|(\mathbf{n}\cdot\nabla\widetilde{\psi}-\widetilde{\varphi}\mathbf{n},v-I_hv)_{\partial\Omega_h(t)}\Big|\leqslant&C\|v-I_hv\|_{L^\infty(\partial\Omega_h(t))}\Big(\|\nabla\widetilde{\psi}\|_{L^1(\partial\Omega_h(t))}+\|\widetilde{\varphi}\|_{L^1(\partial\Omega_h(t))}\Big)\\ \leqslant&C\|v-I_hv\|_{L^\infty(\Omega_h(t))}\Big(\|\widetilde{\psi}\|_{H^2(\mathbb{R}^d)}+\|\widetilde{\varphi}\|_{H^1(\mathbb{R}^d)}\Big)\\ \leqslant&Ch^{r+1}\|v\|_{W_h^{r+1,\infty}(\Omega_h(t))}\|g_h\|_{L^2(\Omega_h(t))}.\end{aligned}\tag{5.42}$$

Combining the above estimates, we obtain the desired estimate (5.37). □

To obtain the optimal order estimate of $E_{t,h}v$, we rely on the negative norm estimate of $R_hq-q$, which is shown in the following lemma.

**Lemma 5.10.** *Let $(R_hv,R_hq)$ be the Stokes–Ritz projection of $(v,q)$. Let $\overline{q}$ denote the average of $q$ over $\Omega_h(t)$. Suppose that $(v,q)$ are sufficiently smooth. Then for each $\lambda\in H^1(\mathbb{R}^d)$, the following inequality holds.*

$$\big|(q-R_hq-\overline{q},\lambda)_{\Omega_h(t)}\big|\leqslant Ch^{r+1}\left(\|v\|_{W_h^{r+1,\infty}(\Omega_h(t))}+\|q\|_{H_h^r(\Omega_h(t))}\right)\|\lambda\|_{H^1(\mathbb{R}^d)}.\tag{5.43}$$

*The constant $C$ is independent of $h,t$ and $(v,q)$.*

*Proof.* For each $\lambda\in H^1(\mathbb{R}^d)$, let $\overline{\lambda}_*:=\frac{1}{|\Omega(t)|}\int_{\Omega(t)}\lambda\,\mathrm{d}x$ denote its average over $\Omega(t)$. By Lemma 5.6, there exists $\chi\in H^2(\Omega(t))\cap H_0^1(\Omega(t))$ such that

$$\mathrm{div}\chi=\lambda-\overline{\lambda}_*\quad\text{ in }\Omega(t),\quad\|\chi\|_{H^2(\Omega(t))}\leqslant C\|\lambda\|_{H^1(\mathbb{R}^d)}.\tag{5.44}$$

We extend $\chi$ to $\widetilde{\chi}\in H^2(\mathbb{R}^d)$ as mentioned in (5.12). By decomposing the integral, we have

$$\big|(q-R_hq-\overline{q},\lambda)_{\Omega_h(t)}\big|=\big|(q-R_hq-\overline{q},\lambda-\overline{\lambda}_*)_{\Omega_h(t)}\big|$$

$$
\begin{aligned}
&\leqslant \big|(q - R_h q - \overline{q}, \lambda - \overline{\lambda}_*)_{\Omega_h(t)\setminus\Omega(t)}\big| + \big|(q - R_h q - \overline{q}, \nabla\cdot\widetilde{\chi})_{\Omega_h(t)\cap\Omega(t)}\big| \\
&\leqslant \big|(q - R_h q - \overline{q}, \lambda - \overline{\lambda}_*)_{\Omega_h(t)\setminus\Omega(t)}\big| + \big|(q - R_h q - \overline{q}, \nabla\cdot\widetilde{\chi})_{\Omega_h(t)}\big| \\
&\quad + \big|(q - R_h q - \overline{q}, \nabla\cdot\widetilde{\chi})_{\Omega_h(t)\setminus\Omega(t)}\big|.
\end{aligned}
$$

To estimate the boundary-skin integral, we derive the following inequalities by using Lemma 5.7:

$$
\begin{aligned}
&\big|(q - R_h q - \overline{q}, \lambda - \overline{\lambda}_*)_{\Omega_h(t)\setminus\Omega(t)}\big| + \big|(q - R_h q - \overline{q}, \nabla\cdot\widetilde{\chi})_{\Omega_h(t)\setminus\Omega(t)}\big| \\
\leqslant &\|q - R_h q\|'_{L^2(\Omega_h(t))}\left(\|\overline{\lambda}_*\|_{L^2(\Omega_h(t)\setminus\Omega(t))} + \|\lambda\|_{L^2(\Omega_h(t)\setminus\Omega(t))} + \|\nabla\cdot\tilde{\chi}\|_{L^2(\Omega_h(t)\setminus\Omega(t))}\right) \\
\leqslant &\|q - R_h q\|'_{L^2(\Omega_h(t))}\left(|\overline{\lambda}_*||\Omega_h(t)\setminus\Omega(t)|^{1/2} + (\|\nabla\cdot\tilde{\chi}\|_{L^6(\mathbb{R}^d)} + \|\lambda\|_{L^6(\mathbb{R}^d)})|\Omega_h(t)\setminus\Omega(t)|^{1/3}\right) \\
\leqslant &\|q - R_h q\|'_{L^2(\Omega_h(t))}\left(|\overline{\lambda}_*||\Omega_h(t)\setminus\Omega(t)|^{1/2} + (\|\tilde{\chi}\|_{H^2(\mathbb{R}^d)} + \|\lambda\|_{H^1(\mathbb{R}^d)})|\Omega_h(t)\setminus\Omega(t)|^{1/3}\right) \\
\leqslant &Ch^{r+1}\Big(\|v\|_{H_h^{r+1}(\Omega_h(t))} + \|q\|_{H_h^r(\Omega_h(t))}\Big)\|\lambda\|_{H^1(\mathbb{R}^d)},
\end{aligned}
\tag{5.45}
$$

where we have used Hölder's inequality, Sobolev embedding $H^1(\mathbb{R}^d) \hookrightarrow L^6(\mathbb{R}^d)$, regularity estimate (5.44), and the fact $r \geqslant 2$.

Since $\widetilde{\chi}|_{\partial\Omega(t)} = 0$, we can interpolate $\widetilde{\chi}$ to a function $\chi_h \in \mathring{V}_h^r(\Omega_h(t))$, i.e., $\chi_h = I_h\widetilde{\chi}$. Then we have

$$
\begin{aligned}
&\big|(q - R_h q - \overline{q}, \nabla\cdot\widetilde{\chi})_{\Omega_h(t)}\big| \\
\leqslant &\big|(q - R_h q - \overline{q}, \nabla\cdot(\widetilde{\chi} - \chi_h))_{\Omega_h(t)}\big| + \big|(q - R_h q - \overline{q}, \nabla\cdot\chi_h)_{\Omega_h(t)}\big| \\
\leqslant &\big|(q - R_h q - \overline{q}, \nabla\cdot(\widetilde{\chi} - \chi_h))_{\Omega_h(t)}\big| + \big|(\nabla(R_h v - v), \nabla(\chi_h - \widetilde{\chi}))_{\Omega_h(t)}\big| + \big|(\nabla(R_h v - v), \nabla\widetilde{\chi})_{\Omega_h(t)}\big| \\
\leqslant &Ch\|\widetilde{\chi}\|_{H^2(\mathbb{R}^d)}\Big(\|q - R_h q\|'_{L^2(\Omega_h(t))} + \|\nabla(R_h v - v)\|_{L^2(\Omega_h(t))}\Big) + \big|(\nabla(R_h v - v), \nabla\widetilde{\chi})_{\Omega_h(t)}\big| \\
\leqslant &Ch^{r+1}\Big(\|v\|_{H_h^{r+1}(\Omega_h(t))} + \|q\|_{H_h^r(\Omega_h(t))}\Big)\|\lambda\|_{H^1(\mathbb{R}^d)} + \big|(\nabla(R_h v - v), \nabla\widetilde{\chi})_{\Omega_h(t)}\big|.
\end{aligned}
$$

Integrating by parts and dealing the boundary integral term as in (5.42) and using regularity estimate (5.44) as well as Lemma 5.9, we obtain

$$
\begin{aligned}
\big|(\nabla(R_h v - v), \nabla\widetilde{\chi})_{L^2(\Omega_h(t))}\big| &\leqslant C\Big(\|R_h v - v\|_{L^2(\Omega_h(t))} + \|R_h v - v\|_{L^\infty(\partial\Omega_h(t))}\Big)\|\widetilde{\chi}\|_{H^2(\mathbb{R}^d)} \\
&\leqslant C\Big(\|R_h v - v\|_{L^2(\Omega_h(t))} + \|R_h v - v\|_{L^\infty(\partial\Omega_h(t))}\Big)\|\lambda\|_{H^1(\mathbb{R}^d)} \\
&\leqslant C\left(\|v\|_{W_h^{r+1,\infty}(\Omega_h(t))} + \|q\|_{H_h^r(\Omega_h(t))}\right)\|\lambda\|_{H^1(\mathbb{R}^d)}
\end{aligned}
$$

Combining the above estimates completes this proof. □

Having completed the necessary preparations, we are now poised to establish the $L^2$-estimate of $E_{t,h}v$. To achieve this, we adopt a proof technique akin to that used in Lemma 5.9, leveraging the insights gained from (5.36) and (5.43).

**Lemma 5.11.** *Let $(R_hv, R_hq)$ be the Stokes–Ritz projection of $(v, q)$. Suppose that $(v, q)$ are sufficiently smooth. Then there is a constant $C$ independent on $h, t$ and $(v, q)$ such that the following estimate holds:*

$$\|E_{t,h}v\|_{L^2(\Omega_h(t))} \leqslant Ch^{r+1}\left(\|v\|_{W_h^{r+1,\infty}(\Omega_h(t))} + \|q\|_{H_h^r(\Omega_h(t))}\right). \tag{5.46}$$

*Proof.* Similarly to the proof of Lemma 5.8, we may assume that $\overline{q} = 0$, where $\overline{q}$ is the average of $q$ over $\Omega_h(t)$. Since $E_{t,h}v = 0$ on the boundary $\partial\Omega_h(t)$, we can choose $g_h = E_{t,h}v$ in (5.32). By (5.36) and the definition of Stokes–Ritz projection, we have

$$\begin{aligned}
\|E_{t,h}v\|^2_{L^2(\Omega_h(t))} \leqslant& C\Big|(\nabla\widetilde{\psi}, \nabla g_h)_{\Omega_h(t)} - (\nabla\cdot g_h, \widetilde{\varphi})_{\Omega_h(t)}\Big| \\
\leqslant& C\Big|(\nabla I_h\widetilde{\psi}, \nabla E_{t,h}v)_{\Omega_h(t)} - (\nabla\cdot E_{t,h}v, \mathbb{P}_h^*\widetilde{\varphi})_{\Omega_h(t)}\Big| \\
&+ C\Big|(\nabla(1-I_h)\widetilde{\psi}, \nabla E_{t,h}v)_{\Omega_h(t)} - (\nabla\cdot E_{t,h}v, (1-\mathbb{P}_h^*)\widetilde{\varphi})_{\Omega_h(t)}\Big|.
\end{aligned}$$

Let $\chi_h = I_h\widetilde{\psi}$, $\lambda_h = \mathbb{P}_h^*\widetilde{\varphi} := \mathbb{P}_h\widetilde{\varphi} - \overline{\mathbb{P}_h\widetilde{\varphi}}$ in (5.29) (where $\mathbb{P}_h$ is the same $L^2$-orthogonal projection operator as used in proof of Lemma 5.9), we obtain

$$\begin{aligned}
&\|E_{t,h}v\|^2_{L^2(\Omega_h(t))} \\
\leqslant& C\Big| - (\nabla(v-R_hv)\nabla w_h, \nabla I_h\widetilde{\psi})_{\Omega_h(t)} + (\nabla(v-R_hv), \nabla I_h\widetilde{\psi}(\nabla\cdot w_h - \nabla w_h))_{\Omega_h(t)}\Big| \\
&+ C\Big|(\nabla I_h\widetilde{\psi} : (\nabla w_h)^\top - \nabla\cdot I_h\widetilde{\psi}\nabla\cdot w_h, q - R_hq)_{\Omega_h(t)}\Big| \\
&+ C\Big| - (\nabla(v-R_hv) : (\nabla w_h)^\top, \mathbb{P}_h^*\widetilde{\varphi})_{\Omega_h(t)} + (\nabla\cdot(v-R_hv), \mathbb{P}_h^*\widetilde{\varphi}\nabla\cdot w_h)_{\Omega_h(t)}\Big| \\
&+ C\Big|(\nabla\cdot(R_hv - v), 1)_{\Omega_h(t)}\frac{(\mathbb{P}_h^*\widetilde{\varphi}, \nabla\cdot w_h)_{\Omega_h(t)}}{|\Omega_h(t)|}\Big| \\
&+ C\Big|(\nabla\cdot I_h\widetilde{\psi}, E_{t,h}q - \overline{E_{t,h}q})_{\Omega_h(t)}\Big| \\
&+ C\Big|(\nabla(1-I_h)\widetilde{\psi}, \nabla E_{t,h}v)_{\Omega_h(t)} - (\nabla\cdot E_{t,h}v, (1-\mathbb{P}_h^*)\widetilde{\varphi})_{\Omega_h(t)}\Big| \\
\leqslant& C\Big| - (\nabla(v-R_hv)\nabla w_h, \nabla(I_h\widetilde{\psi} - \widetilde{\psi}))_{\Omega_h(t)} + (\nabla(v-R_hv), \nabla(I_h\widetilde{\psi} - \widetilde{\psi})(\nabla\cdot w_h - \nabla w_h))_{\Omega_h(t)}\Big| \\
&+ C\Big|(\nabla(I_h\widetilde{\psi} - \widetilde{\psi}) : (\nabla w_h)^\top - \nabla\cdot(I_h\widetilde{\psi} - \widetilde{\psi})\nabla\cdot w_h, q - R_hq)_{\Omega_h(t)}\Big| \\
&+ C\Big| - (\nabla(v-R_hv) : (\nabla w_h)^\top, \mathbb{P}_h^*\tilde{\varphi} - \tilde{\varphi})_{\Omega_h(t)} + (\nabla\cdot(v-R_hv), (\mathbb{P}_h^*\tilde{\varphi} - \tilde{\varphi})\nabla\cdot w_h)_{\Omega_h(t)}\Big|
\end{aligned}$$

$$
\begin{aligned}
&+C\Big|(\nabla(R_hv-v)\nabla w_h,\nabla\widetilde{\psi})_{\Omega_h(t)}+(\nabla(v-R_hv),\nabla\widetilde{\psi}(\nabla\cdot w_h-\nabla w_h))_{\Omega_h(t)}\Big|\\
&+C\Big|(\nabla\widetilde{\psi}:(\nabla w_h)^\top-\nabla\cdot\widetilde{\psi}\nabla\cdot w_h,q-R_hq)_{\Omega_h(t)}\Big|\\
&+C\Big|-(\nabla(v-R_hv):(\nabla w_h)^\top,\tilde{\varphi})_{\Omega_h(t)}+(\nabla\cdot(v-R_hv),\tilde{\varphi}\nabla\cdot w_h)_{\Omega_h(t)}\Big|\\
&+C\Big|(\nabla\cdot(R_hv-v),1)_{\Omega_h(t)}\frac{(\mathbb{P}_h^*\tilde{\varphi},\nabla\cdot w_h)_{\Omega_h(t)}}{|\Omega_h(t)|}\Big|\\
&+C\Big|(\nabla\cdot(I_h-1)\widetilde{\psi},E_{t,h}q-\overline{E_{t,h}q})_{\Omega_h(t)}\Big|+C\Big|(\nabla\cdot\widetilde{\psi},E_{t,h}q-\overline{E_{t,h}q})_{\Omega_h(t)}\Big|\\
&+C\Big|(\nabla(1-I_h)\widetilde{\psi},\nabla E_{t,h}v)_{\Omega_h(t)}-(\nabla\cdot E_{t,h}v,(1-\mathbb{P}_h^*)\widetilde{\varphi})_{\Omega_h(t)}\Big|.
\end{aligned}
$$

Furthermore, $w_h$ can be replaced by $(w_h-w)+w$. In view of the error estimate (5.14), applying the same routine as in the proof of Lemma 5.9, i.e. using the error of interpolation $I_h$, error of modified $L^2$-projection $\mathbb{P}_h^*$ (5.40), Lemma 5.8 and an analogue of estimate (5.41), we have

$$
\begin{aligned}
\|E_{t,h}v\|^2_{L^2(\Omega_h(t))}\leqslant&Ch^{r+1}\left(\|v\|_{H_h^{r+1}(\Omega_h(t))}+\|q\|_{H_h^r(\Omega_h(t))}\right)\left(\|\widetilde{\psi}\|_{H^2(\Omega_h(t))}+\|\widetilde{\varphi}\|_{H^1(\Omega_h(t))}\right)\\
&+C\Big|(\nabla\cdot(R_hv-v),1)_{\Omega_h(t)}\frac{(\mathbb{P}_h^*\tilde{\varphi},\nabla\cdot w_h)_{\Omega_h(t)}}{|\Omega_h(t)|}\Big|\\
&+C\Big|(\nabla(R_hv-v)\nabla w,\nabla\widetilde{\psi})_{\Omega_h(t)}+(\nabla(v-R_hv),\nabla\widetilde{\psi}(\nabla\cdot w-\nabla w))_{\Omega_h(t)}\Big|\\
&+C\Big|(\nabla\widetilde{\psi}:(\nabla w)^\top-\nabla\cdot\widetilde{\psi}\nabla\cdot w,q-R_hq)_{\Omega_h(t)}\Big|\\
&+C\Big|-(\nabla(v-R_hv):(\nabla w)^\top,\tilde{\varphi})_{\Omega_h(t)}+(\nabla\cdot(v-R_hv),\tilde{\varphi}\nabla\cdot w)_{\Omega_h(t)}\Big|.
\end{aligned}
$$

By using the regularity result (5.33), Lemma 5.10, and integration by parts, we obtain

$$
\begin{aligned}
\|E_{t,h}v\|^2_{L^2(\Omega_h(t))}\leqslant&Ch^{r+1}\left(\|v\|_{W_h^{r+1,\infty}(\Omega_h(t))}+\|q\|_{H_h^r(\Omega_h(t))}\right)\|E_{t,h}v\|_{L^2(\Omega_h(t))}\\
&+C\Big|(R_hv-v,\mathbf{n})_{\partial\Omega_h(t)}\frac{(\mathbb{P}_h^*\tilde{\varphi},\nabla\cdot w_h)_{\Omega_h(t)}}{|\Omega_h(t)|}\Big|\\
&+C\Big|-(R_hv-v,\nabla\cdot(\nabla\widetilde{\psi}\nabla w^\top))_{\Omega_h(t)}+(R_hv-v,\mathbf{n}\cdot(\nabla\widetilde{\psi}\nabla w^\top))_{\partial\Omega_h(t)}\Big|\\
&+C\Big|-(R_hv-v,\nabla\cdot(\nabla\widetilde{\psi}(\nabla\cdot w-\nabla w)))_{\Omega_h(t)}\Big|\\
&+C\Big|(v-R_hv,\mathbf{n}\cdot(\nabla\widetilde{\psi}(\nabla\cdot w-\nabla w)))_{\partial\Omega_h(t)}\Big|\\
&+C\Big|(v-R_hv,\nabla\cdot(\tilde{\varphi}(\nabla w)^\top))_{\Omega_h(t)}-(v-R_hv,\nabla(\tilde{\varphi}\nabla\cdot w))_{\Omega_h(t)}\Big|\\
&+C\Big|-(v-R_hv,\tilde{\varphi}\mathbf{n}\cdot(\nabla w)^\top)_{\partial\Omega_h(t)}+((v-R_hv)\cdot\mathbf{n},\tilde{\varphi}\nabla\cdot w)_{\partial\Omega_h(t)}\Big|.
\end{aligned}
$$

Since $R_h v = I_h v$ on $\partial\Omega_h(t)$, by using Lemma 5.9, we have

$$\begin{aligned}\|E_{t,h}v\|^2_{L^2(\Omega_h(t))} \leqslant & Ch^{r+1}\left(\|v\|_{W_h^{r+1,\infty}(\Omega_h(t))} + \|q\|_{H_h^r(\Omega_h(t))}\right)\|E_{t,h}v\|_{L^2(\Omega_h(t))} \\ & + C\Big(\|R_h v - v\|_{L^2(\Omega_h(t))} + \|v - I_h v\|_{L^\infty(\partial\Omega_h(t))}\Big)(\|\widetilde{\psi}\|_{H^2(\mathbb{R}^d)} + \|\widetilde{\varphi}\|_{H^1(\mathbb{R}^d)}) \\ \leqslant & Ch^{r+1}\left(\|v\|_{W_h^{r+1,\infty}(\Omega_h(t))} + \|q\|_{H_h^r(\Omega_h(t))}\right)\|E_{t,h}v\|_{L^2(\Omega_h(t))}. \end{aligned} \tag{5.47}$$

By using Young's inequality, we finish the proof. □

With these preparations done, we can go start proving Theorem 5.1, which is shown in next subsection.

### 5.3.3 Proof of Theorem 5.1.

*Proof.* We define the auxiliary function $\xi$ in $\mathbb{R}^d$ as follows:

$$\xi := \partial_t u - \Delta u + \nabla p - f, \tag{5.48}$$

where $u$, $p$, $f$ represent their extensions to $\mathbb{R}^d$. By testing the equation (5.48) with $v_h \in \mathring{V}_h^r(\Omega_h(t))$, we obtain:

$$(D_{t,h}u - w_h \cdot \nabla u, v_h)_{\Omega_h(t)} + (\nabla u, \nabla v_h)_{\Omega_h(t)} - (\nabla \cdot v_h, p)_{\Omega_h(t)} = (f, v_h)_{\Omega_h(t)} + (\xi, v_h)_{\Omega_h(t)}.$$

Applying Hölder's inequality, Lemma 5.4 and the fact $r \geqslant 2$, we can derive the following estimate:

$$|(\xi, v_h)_{\Omega_h(t)}| = |(\xi, v_h)_{\Omega_h(t)\setminus\Omega(t)}| \leqslant Ch^{r+1}\|\nabla v_h\|_{L^2(\Omega_h(t))}\|\xi\|_{L^2(\mathbb{R}^d)}. \tag{5.49}$$

It follows from (5.21) that the Stokes–Ritz projection $(R_h u, R_h p)$ satisfies the following equation

$$\begin{aligned}(D_{t,h}R_h u - w_h \cdot \nabla R_h u, v_h)_{\Omega_h(t)} + (\nabla R_h u, \nabla v_h)_{\Omega_h(t)} - (\nabla \cdot v_h, R_h p)_{\Omega_h(t)} & \\ = (f, v_h)_{\Omega_h(t)} + (\xi, v_h)_{\Omega_h(t)} + (\mathscr{F}, v_h)_{\Omega_h(t)} \quad \forall v_h \in \mathring{V}_h^r(\Omega_h(t)), & \text{(5.50a)} \\ (\nabla \cdot R_h u, q_h)_{\Omega_h(t)} = 0 \quad \forall q_h \in Q_h^{r-1}(\Omega_h(t)), & \text{(5.50b)}\end{aligned}$$

where the remainder $\mathscr{F} := D_{t,h}(R_h u - u) - w_h \cdot \nabla(R_h u - u)$ represents the consistency error of the spatial discretization.

Since $v_h \in \mathring{V}_h^r(\Omega_h(t))$, via integration by parts, we can estimate $w_h \cdot \nabla(R_h u - u)$ as follows:

$$\big|(w_h \cdot \nabla(R_h u - u), v_h)_{\Omega_h(t)}\big| = \big| - (v_h \nabla \cdot w_h, R_h u - u)_{\Omega_h(t)} - (R_h u - u, w_h \cdot \nabla v_h)_{\Omega_h(t)}\big|$$

$$\leqslant C\|R_h u - u\|_{L^2(\Omega_h(t))}\|v_h\|_{H^1(\Omega_h(t))}, \tag{5.51}$$

where we have used the $W^{1,\infty}$ boundedness of the mesh velocity, i.e., $\|w_h(t)\|_{W^{1,\infty}(\Omega_h(t))} \leqslant C$, which follows from (5.14) and the triangle inequality. Thus, we have

$$\begin{aligned}\left|(\mathscr{F}, v_h)_{\Omega_h(t)}\right| \leqslant & C\Big(\|D_{t,h}(R_h u - u)\|_{L^2(\Omega_h(t))} + \|R_h u - u\|_{L^2(\Omega_h(t))}\Big)\|v_h\|_{H^1(\Omega_h(t))}. \\ \leqslant & C\Big(\|E_{t,h}u\|_{L^2(\Omega_h(t))} + \|R_h D_{t,h}u - D_{t,h}u\|_{L^2(\Omega_h(t))}\Big)\|v_h\|_{H^1(\Omega_h(t))} \\ & + C\|R_h u - u\|_{L^2(\Omega_h(t))}\|v_h\|_{H^1(\Omega_h(t))}\end{aligned} \tag{5.52}$$

by the definition (5.25) of $E_{t,h}$. We can estimate $\|R_h D_{t,h}u - D_{t,h}u\|_{L^2(\Omega_h(t))}$ term in (5.52) by Lemma 5.9 as follows:

$$\begin{aligned}&\|R_h D_{t,h}u - D_{t,h}u\|_{L^2(\Omega_h(t))} \\ \leqslant & Ch^{r+1}\left(\|D_{t,h}u\|_{W_h^{r+1,\infty}(\Omega_h(t))} + \|D_{t,h}p\|_{H_h^r(\Omega_h(t))}\right) \\ \leqslant & Ch^{r+1}\Big(\|\partial_t u\|_{W^{r+1,\infty}(\Omega_h(t))} + \|w_h\|_{W_h^{r+1,\infty}(\Omega_h(t))}\|u\|_{W^{r+2,\infty}(\Omega_h(t))} \\ & + \|\partial_t p\|_{H^r(\Omega_h(t))} + \|w_h\|_{W_h^{r,\infty}(\Omega_h(t))}\|p\|_{H^{r+1}(\Omega_h(t))}\Big) \\ \leqslant & Ch^{r+1}\left(\|\partial_t u\|_{W^{r+1,\infty}(\Omega_h(t))} + \|u\|_{W^{r+2,\infty}(\Omega_h(t))} + \|\partial_t p\|_{H^r(\Omega_h(t))} + \|p\|_{H^{r+1}(\Omega_h(t))}\right),\end{aligned} \tag{5.53}$$

where we have employed formula (5.6) of material derivative and the $W_h^{r+1,\infty}$-boundedness (5.14) of discrete velocity $w_h$. Combining the estimate (5.53) and (5.52) as well as using Lemma 5.9 and Lemma 5.11 , we can derive that

$$|(\mathscr{F}, v_h)_{\Omega_h(t)}| \leqslant Ch^{r+1}A_{u,p}(t)\|v_h\|_{H^1(\Omega_h(t))}, \tag{5.54}$$

where we used notation $A_{u,p}(t)$ which is an abbreviation defined as follows

$$A_{u,p}(t) := \|\partial_t u(\cdot,t)\|_{W^{r+1,\infty}(\mathbb{R}^d)} + \|u(\cdot,t)\|_{W^{r+2,\infty}(\mathbb{R}^d)} + \|\partial_t p(\cdot,t)\|_{H^r(\mathbb{R}^d)} + \|p(\cdot,t)\|_{H^{r+1}(\mathbb{R}^d)}.$$

Let us define $e_u := R_h u - u_h$ and $e_p := R_h p - p_h$. By subtracting equation (5.18) from equation (5.50) we obtain the following equations for any $v_h \in \mathring{V}_h^r(\Omega_h(t))$ and $q_h \in Q_h^{r-1}(\Omega_h(t))$

$$(D_{t,h}e_u - w_h \cdot \nabla e_u, v_h)_{\Omega_h(t)} + (\nabla e_u, \nabla v_h)_{\Omega_h(t)} - (\nabla \cdot v_h, e_p)_{\Omega_h(t)} = (\xi + \mathscr{F}, v_h)_{\Omega_h(t)}, \tag{5.55a}$$

$$(\nabla \cdot e_u, q_h)_{\Omega_h(t)} = 0. \tag{5.55b}$$

Since $R_h u = I_h u = w_h = u_h$ on the boundary $\partial\Omega_h(t)$, we have $e_u \in \mathring{V}_h^r(\Omega_h(t))$. By testing (5.55a) with $v_h = e_u$, we obtain:

$$\frac{1}{2}\partial_t\|e_u(t)\|^2_{L^2(\Omega_h(t))} + \|\nabla e_u(t)\|^2_{L^2(\Omega_h(t))} = (\xi, e_u)_{\Omega_h(t)} + (\mathscr{F}, e_u)_{\Omega_h(t)}$$

$$\leqslant Ch^{r+1}A_{u,p}(t)\|e_u(t)\|_{H^1(\Omega_h(t))},$$

where the last inequality follows from (5.49) and (5.54).

By applying Young's inequality and absorbing $\|e_u\|^2_{H^1(\Omega_h(t))}$ on the right-hand side, we can integrate the inequality from 0 to $t$ to obtain:

$$\|e_u\|^2_{L^2(\Omega_h(t))} + \int_0^t \|\nabla e_u\|^2_{L^2(\Omega_h(s))}\mathrm{d}s \leqslant \|e_h(0)\|^2_{L^2(\Omega_h(0))} + CR^2_{u,p}h^{2(r+1)} \leqslant CR^2_{u,p}h^{2(r+1)}.$$

Combining this result with Lemma 5.9, we derive the estimate for $u - u_h$.

For the estimate of $p - p_h$, by using inf-sup condition, the $W^{1,\infty}$-boundedness of $w_h$, and equation (5.55a), we have

$$\begin{aligned}\|e_p\|_{L^2(\Omega_h(t))} \leqslant & C \sup_{0\neq v_h\in \mathring{V}_h^r(\Omega_h(t))} \frac{(\nabla\cdot v_h, e_p)}{\|\nabla v_h\|_{L^2(\Omega_h(t))}} \\ \leqslant & C\left(h^{r+1}A_{u,p}(t) + \|\nabla e_u\|_{L^2(\Omega_h(t))} + \|D_{t,h}e_u\|_{L^2(\Omega_h(t))}\right). \end{aligned} \tag{5.56}$$

Since $D_{t,h}e_u = 0$ on the boundary $\partial\Omega_h(t)$, we can chooose $v_h = D_{t,h}e_u$ in (5.55a) and obtain

$$\begin{aligned}&\|D_{t,h}e_u\|^2_{L^2(\Omega_h(t))} - (w_h\cdot\nabla e_u, D_{t,h}e_u)_{\Omega_h(t)} + (\nabla e_u, \nabla D_{t,h}e_u)_{\Omega_h(t)} - (\nabla\cdot D_{t,h}e_u, e_p)_{\Omega_h(t)} \\ &= (\xi + \mathscr{F}, D_{t,h}e_u)_{\Omega_h(t)}. \end{aligned} \tag{5.57}$$

From Lemma 5.1 and Lemma 5.2, it is known that

$$\begin{aligned}(\nabla e_u, \nabla D_{t,h}e_u)_{\Omega_h(t)} =& \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\nabla e_u\|^2_{L^2(\Omega_h(t))} - \frac{1}{2}(|\nabla e_u|^2, \nabla\cdot w_h)_{\Omega_h(t)} \\ &+ \frac{1}{2}(\nabla e_u(\nabla w_h + (\nabla w_h)^\top), \nabla e_u)_{\Omega_h(t)}. \end{aligned} \tag{5.58}$$

By taking derivative to (5.55b) with respect time, we obtain that

$$(\nabla\cdot D_{t,h}e_u, q_h)_{\Omega_h(t)} + (\nabla\cdot e_u\nabla\cdot w_h - \nabla e_u : (\nabla w_h)^\top, q_h)_{\Omega_h(t)} = 0. \tag{5.59}$$

Let $q_h = e_p$ in (5.59), we have

$$(\nabla\cdot D_{t,h}e_u, e_p)_{\Omega_h(t)} + (\nabla\cdot e_u\nabla\cdot w_h - \nabla e_u : (\nabla w_h)^\top, e_p)_{\Omega_h(t)} = 0 \tag{5.60}$$

Substituting (5.49), (5.54), (5.58) and (5.60) into (5.57), and using inverse estimate, we can obtain that

$$\begin{aligned}&\|D_{t,h}e_u\|^2_{L^2(\Omega_h(t))} + \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\nabla e_u\|^2_{L^2(\Omega_h(t))} \\ \leqslant & Ch^{r+1}A_{u,p}(t)\|D_{t,h}e_u\|_{H^1(\Omega_h(t))} + C\|D_{t,h}e_u\|_{L^2(\Omega_h(t))}\|\nabla e_u\|_{L^2(\Omega_h(t))}\end{aligned}$$

$$
\begin{aligned}
&+C\|\nabla e_u\|_{L^2(\Omega_h(t))}\|e_p\|_{L^2(\Omega_h(t))}+C\|\nabla e_u\|^2_{L^2(\Omega_h(t))}\\
\leqslant& Ch^r A_{u,p}(t)\|D_{t,h}e_u\|_{L^2(\Omega_h(t))}+C\|D_{t,h}e_u\|_{L^2(\Omega_h(t))}\|\nabla e_u\|_{L^2(\Omega_h(t))}\\
&+C\|\nabla e_u\|_{L^2(\Omega_h(t))}\|e_p\|_{L^2(\Omega_h(t))}+C\|\nabla e_u\|^2_{L^2(\Omega_h(t))} \qquad (5.61)
\end{aligned}
$$

By substituting (5.56) into (5.61), using Young's inequality, and integrating both sides from 0 to $t$, we have

$$
\begin{aligned}
&\int_0^t \|D_{t,h}e_u\|^2_{L^2(\Omega_h(s))}ds+\|\nabla e_u\|^2_{L^2(\Omega_h(t))}\\
\leqslant& Ch^{2r}R^2_{u,p}+C\int_0^t \|\nabla e_u\|^2_{L^2(\Omega_h(s))}\mathrm{d}s+\|\nabla e_u(\cdot,0)\|^2_{L^2(\Omega_h^0)}\leqslant Ch^{2r}R^2_{u,p}. \qquad (5.62)
\end{aligned}
$$

Combining (5.62) and (5.56), we obtain that

$$
\|e_p\|_{L^2(0,T;L^2(\Omega_h(t)))}\leqslant CR_{u,p}h^r.
$$

By using Lemma 5.7, we can deduce that

$$
\|p-p_h\|_{L^2(0,T;L^2(\Omega_h(t)))}\leqslant CR_{u,p}h^r+C\|\overline{p}\|_{L^2(0,T)}, \qquad (5.63)
$$

where $\bar{p}(t)$ is the average of $p(t)$ on $\Omega_h(t)$. Since $p\in L^2_0(\Omega(t))$, we have that

$$
\begin{aligned}
|\overline{p}(t)|\leqslant& C\left(\Big|\int_{\Omega_h(t)\backslash\Omega(t)}p\mathrm{d}x\Big|+\Big|\int_{\Omega(t)\backslash\Omega_h(t)}p\mathrm{d}x\Big|\right)\\
\leqslant& C\|p(\cdot,t)\|_{L^\infty(\mathbb{R}^d)}(|\Omega_h(t)\backslash\Omega(t)|+|\Omega(t)\backslash\Omega_h(t)|)\\
\leqslant& CA_{u,p}(t)h^{r+1}\quad\text{(Sobolev embedding } H^{r+1}(\mathbb{R}^d)\hookrightarrow L^\infty(\mathbb{R}^d)\text{ used)} \qquad (5.64)
\end{aligned}
$$

Combining (5.63) and (5.64), we complete the proof. □

**5.4 Numerical experiments.** In this section, we provide numerical tests for problem (5.1) to support the theoretical result proved in Theorem 5.1. For temporal discretization, We use the second-order projection method. If we define the pull back operator $P_h^{n,m}$ : $V_h^r(\Omega_h^n)\to V_h^r(\Omega_h^m)$ as $P_h^{n,m}v_h=v_h\circ\Phi_h^n\circ(\Phi_h^m)^{-1}$ for any $v_h\in V_h^r(\Omega_h^n)$, then the fully discrete scheme is shown as follows: Find $u_h^{n+1}\in V_h^r(\Omega_h^{n+1})$ and $p_h^{n+1}\in Q_h^{r-1}(\Omega_h^{n+1})$ at step $n+1$ such that

$$
\begin{aligned}
&\frac{1}{2\tau}\Big[\Big(u_h^{n+1}-P_h^{n,n+1}u_h^n,v_h\Big)_{\Omega_h^{n+1}}+\Big(P_h^{n+1,n}u_h^{n+1}-u_h^n,P_h^{n+1,n}v_h\Big)_{\Omega_h^n}\Big]\\
&-\frac{1}{8}\Big((w_h^{n+1}+P_h^{n,n+1}w_h^n)\cdot\nabla(u_h^{n+1}+P_h^{n,n+1}u_h^n),v_h\Big)_{\Omega_h^{n+1}}
\end{aligned}
$$

$$
\begin{aligned}
&-\frac{1}{8}\Big((P_h^{n+1,n}w_h^{n+1}+w_h^n)\cdot\nabla(P_h^{n+1,n}u_h^{n+1}+u_h^n),P_h^{n+1,n}v_h\Big)_{\Omega_h^n}\\
&+\frac{1}{4}\Big[\Big(\nabla(u_h^{n+1}+P_h^{n,n+1}u_h^n),\nabla v_h\Big)_{\Omega_h^{n+1}}+\Big(\nabla(P_h^{n+1,n}u_h^{n+1}+u_h^n),\nabla P_h^{n+1,n}v_h\Big)_{\Omega_h^n}\Big]\\
&-\frac{1}{2}\Big[\big(\nabla\cdot v_h,p_h^n\circ\Phi_h^n\circ(\Phi_h^{n+1})^{-1}\big)_{\Omega_h^{n+1}}+\Big(\nabla\cdot P_h^{n+1,n}v_h,p_h^n\Big)_{\Omega_h^n}\Big]\\
=&\frac{1}{2}\Big[(f(t_{n+1}),v_h)_{\Omega_h^{n+1}}+(f(t_n),P_h^{n+1,n}v_h)_{\Omega_h^n}\Big]\quad\forall v_h\in\mathring{V}_h^r(\Omega_h^{n+1}),
\end{aligned}
\tag{5.65a}
$$

$$
(\nabla\cdot u_h^{n+1},q_h)_{\Omega_h^{n+1}}+\beta\tau\left(\nabla(p_h^{n+1}-p_h^n\circ\Phi_h^n\circ(\Phi_h^{n+1})^{-1}),\nabla q_h\right)_{\Omega_h^{n+1}}=0\quad\forall q_h\in Q_h^{r-1}(\Omega_h^{n+1}),
\tag{5.65b}
$$

where $\beta>1$ is a constant. In the numerical tests, we choose $\beta=2$. The solution $u_h^{n+1}$ is obtained by solving equation (5.65a), and subsequently, $p_h^{n+1}$ is computed using equation (5.65b) and $u_h^{n+1}$.

**Example 5.1.** Let $\Omega(t)$ be an ellipse given by:

$$
\Omega(t)=\{(x,y):F(x,y)\leqslant 0\}\ \text{ for }\ F(x,y)=(1-\frac{t}{4})^2x^2+(1-\frac{t}{4})^{-2}y^2-1.
$$

Then for $t\geqslant 0$, the domain $\Omega(t)$ evolves with volume invariant. We select the velocity function $w$ to be

$$
w(x,y;t)=-\frac{\partial_t F\nabla F}{\nabla F^\top\nabla F}\ \text{ on }\ \partial\Omega(t),\quad\text{and}\quad -\Delta w=0\ \text{ in }\ \Omega(t).
$$

The initial value $u_0$ is chosen to be $w(\cdot,0)$ and $f=0$. Since the exact solution is not known, we compute a numerical solution for sufficiently small $\tau$ and $h$ as reference solution.

The initial and final discretized domains, denoted as $\Omega_h(0)$ and $\Omega_h(1)$ respectively, are illustrated in Figure 5.1. These domains are obtained by employing the $P_1$ element and $P_2$ element, representing the piecewise linear and quadratic finite elements, respectively.

To assess the convergence properties of the numerical scheme, we conducted a convergence test at time $T=1$ to assess the spatial discretization. For this purpose, we employed two different sets of finite elements: $P_{1b}-P_1$, and $P_2-P_1$, while keeping the time step sizes sufficiently small to ensure minimal errors from the time discretization. The errors of the numerical solutions are presented in Figure 5.2 for varying mesh sizes: $h=1/8,1/16,1/32,1/64$. The results demonstrate that the numerical solutions exhibit $r+1$-th order convergence in space, where $r$ corresponds to the order of the FEM. This finding aligns with the theoretical results established in Theorem 5.1 for $r=2$. Notably, for

the $P_{1b} - P_1$ element, we verified that the inf-sup condition (5.11) is satisfied. Therefore, we can attain second-order convergence using the same approach presented in this chapter.

In addition to investigating the convergence in space, we also conducted a temporal convergence test at $T = 1$ using the $P_2 - P_1$ element and a suitably small mesh size that ensures negligible errors from the space discretization. The resulting errors of the numerical solutions are depicted in Figure 5.3 for different time step sizes: $\tau = 1/50, 1/100, 1/200, 1/400$. The observed errors demonstrate second-order convergence of velocity $u$ in time.

**Example 5.2.** In this example, we investigate the convergence order of numerical solutions in a rotating domain. Let the initial $\Omega(0)$ be an ellipse given by

$$\Omega(0) = \{(x, y) : \frac{25}{16}x^2 + \frac{25}{9}y^2 \leqslant 1\}.$$

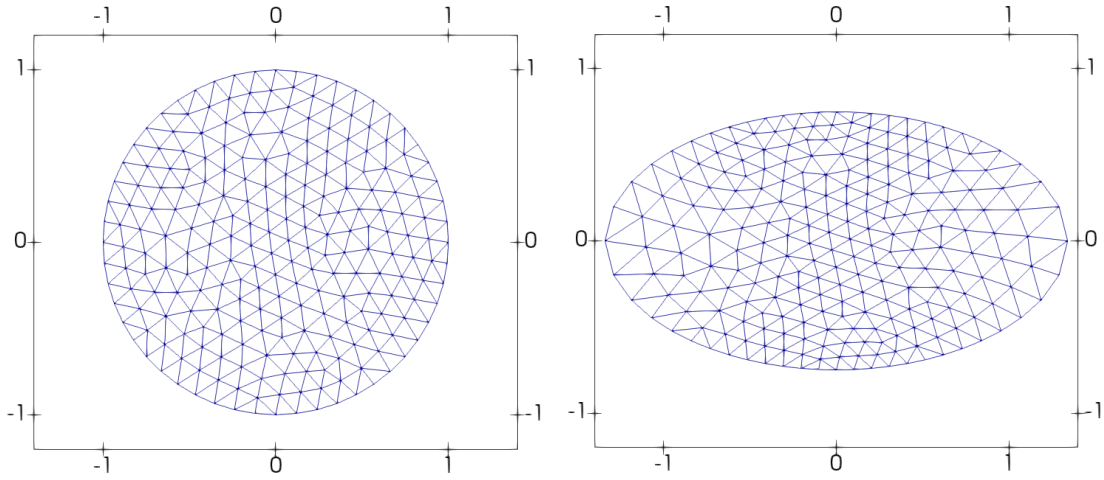
The domain $\Omega(t)$ is generated by the rotating mesh velocity field $w(x, y, t)$, which is given by

$$w(x, y, t) = (-y \sin t, x \cos t).$$

The exact solutions $(u, p)$ are chosen to be $u(x, y, t) = w(x, y, t)$ and $p(x, y, t) = x + y$. The source function $f$ is chosen to be consistent with the equation (5.1a).

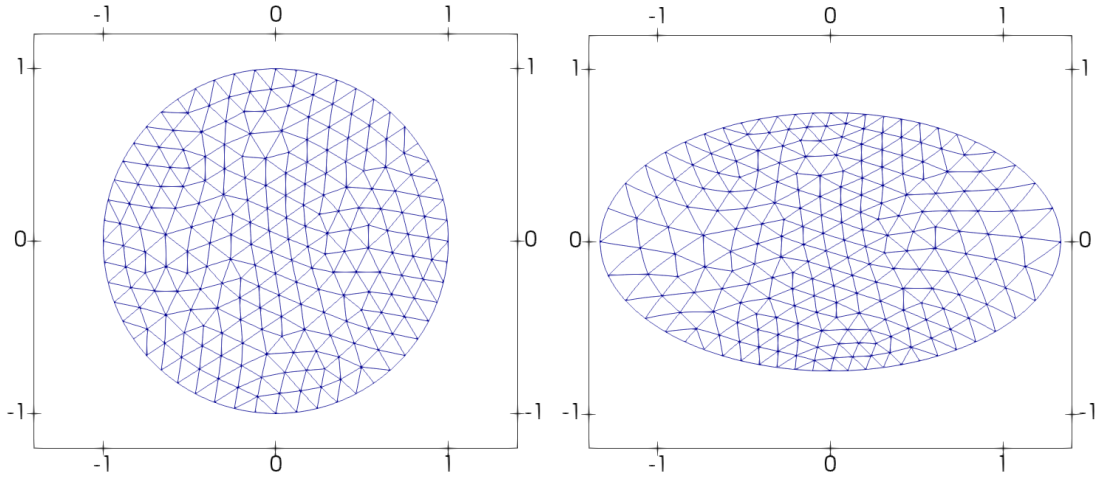
Similarly to Example 5.1, we assess the convergence behavior of the numerical solutions. Specifically, we investigate the convergence of spatial discretization using the $P_{1b} - P_1$, and $P_2 - P_1$ elements, considering sufficiently small time step sizes that ensure the errors from time discretization are negligible. Figure 5.4 illustrates the errors of the numerical solutions for different mesh sizes: $h = 1/8, 1/16, 1/32, 1/64$. The results indicate that the numerical solutions exhibit $r+1$-th order convergence in space for $r$-th order FEMs. This convergence behavior aligns with the Theorem 5.1.

In addition, we examine the convergence of the velocity $u$ in time at $T = 1$ using the $P_2 - P_1$ element, with a sufficiently small mesh size that ensures the errors from spatial discretization are negligible. The errors of the numerical solutions are presented in Figure 5.5 for various time step sizes: $\tau = 1/50, 1/100, 1/200, 1/400$. The numerical results demonstrate a second-order convergence in time.




(a) $P_1$ element at $t = 0$

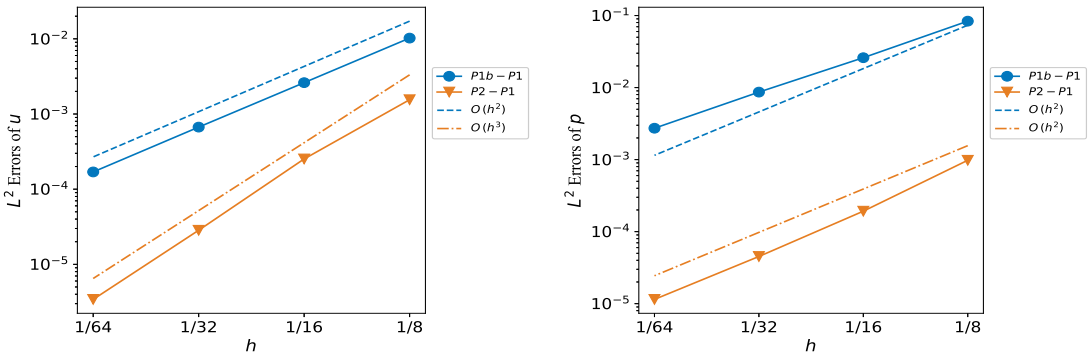
(b) $P_1$ element at $t = 1$




(c) $P_2$ element at $t = 0$

(d) $P_2$ element at $t = 1$

*Figure 5.1. Meshes of $P_1$ and $P_2$ elements at time $T = 0$ and $T = 1$.*




(a) $L^2$ error of $u$ from spatial discretization at $T = 1$ (b) $L^2L^2$ error of $p$ from spatial discretization

*Figure 5.2. Errors from spatial discretization for $T = 1$.*

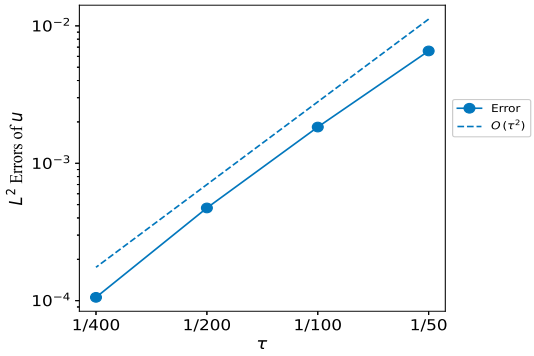
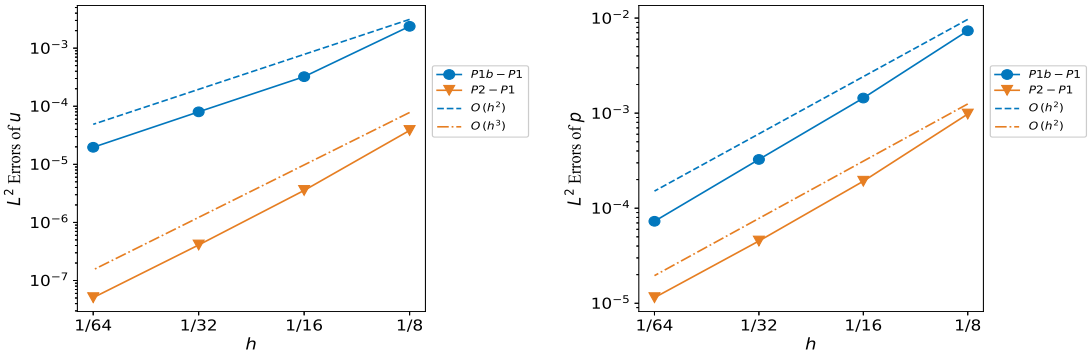



*Figure 5.3. Errors from temporal discretization for velocity $u$ at time $T = 1$.*




(a) $L^2$ error of $u$ from spatial discretization at $T = 1$ (b) $L^2L^2$ error of $p$ from spatial discretization

*Figure 5.4. Errors from spatial discretization for $T = 1$.*

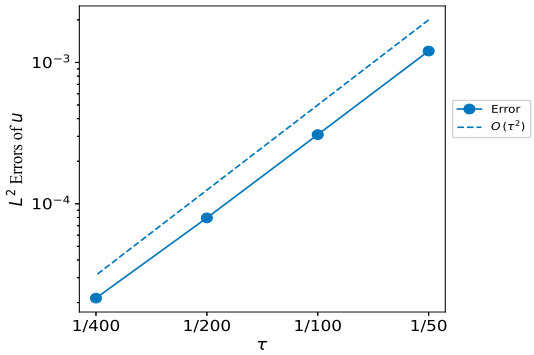



*Figure 5.5. Errors from temporal discretization for velocity $u$ at time $T = 1$.*

# Chapter 6

## Optimal Convergence of an Arbitrary Lagrangian–Eulerian Finite Element Method for Fluid-Structure Interactions

The content of this chapter has been published in "*B. Li, Q. Rao, and P. Sun. Optimal Convergence of an Arbitrary Lagrangian–Eulerian Finite Element Method for Fluid-Structure Interactions, submitted.*"

**6.1 Introduction.** In this chapter, we consider the problem of interactions between the free viscous fluid flow and a deformable elastic structure, which gives rise to a rich variety of physical phenomena with applications in many fields of science and engineering, such as the vibration of rotating turbine blades impacted by the fluid flow, the response of bridges and tall buildings to winds, the floating parachute wafted by the air current, the blood flow through arteries that are made up of a deformable hyperelastic medium, and etc. These problems are often referred to as fluid-structure interactions (FSI); see [176, 163, 20, 35, 156, 120, 168, 167]. The incompressible Navier–Stokes equations, defined in Eulerian coordinates, are frequently used to describe fluid motion. To define the structural motion, the linear elasticity model or neo-Hookean-type material models in Lagrangian coordinates are commonly employed; see [115]. Generally, FSI problems necessitate that both the fluid and structural fields at the common interface not only share the same velocity but also exhibit consistent normal stresses. Current major numerical approaches to solve FSI problems are distinguished by their methods of handling the interface conditions along the moving interface and by their strategies for managing the mesh in the fluid region.

Due to their high-order accuracy, body-fitted mesh methods have become a reliable numerical approach for solving unsteady moving domain/interface problems including FSI, provided that the moving mesh – which adapts to the moving boundary/interface over time – can be efficiently generated. In this context, the arbitrary Lagrangian–Eulerian

(ALE) method [73, 79, 78, 140, 160] has emerged as the most popular one among body-fitted mesh methods due to its high feasibility and fidelity, where the mesh at the interface is continuously adjusted to be shared by both the fluid and structure, thereby naturally satisfying the interface conditions of FSI within a location-dependent finite element space tailored to the interface.

The convergence of ALE-based finite element methods (ALE-FEM) for time-dependent partial differential equations (PDEs) in evolving domains has been studied in many articles. When the velocity of the moving boundary/interface was prescribed, stability analysis of ALE-FEM for parabolic equations was provided in [48], optimal-order error estimates in $L^2(0,T;H^1)$ norm were given in [53], and optimal-order error estimates in $L^\infty(0,T;L^2)$ norm were established in [54, 114]. We also refer to [41] and [38] for a unified framework of ALE evolving FEMs and an ALE method with harmonically evolving mesh, respectively. For the Stokes equations, suboptimal convergence rates in $L^\infty(0,T;H^1)$ and $L^\infty(0,T;L^2)$ were given in [133], and optimal convergence rate in $L^\infty(0,T;L^2)$ was established in [147] recently.

However, very few work has been done when the moving boundary/interface is driven by the solutions of PDEs in the domain. This is the case of FSI problems and many other related free boundary/interface problems, as well as shape optimization problems and surface evolution under geometric curvature flows. Optimal-order convergence in $L^\infty(0,T;H^1)$ of evolving FEM for surface evolution under geometric curvature flows were established in [93, 94, 76, 40] by utilizing a matrix-vector framework introduced in [95] for analyzing the errors in approximating surface-driven surface evolution problems. For discussions on the convergence of finite element methods for curve evolution, readers are referred to articles [36, 39, 146, 9, 177]. Besides, optimal-order convergence in $L^\infty(0,T;H^1)$ of an ALE-FEM for boundary evolution under a gradient flow arising from shape optimization problems was obtained in [56], and optimal-order convergence in $L^\infty(0,T;H^1)$ of an ALE-FEM for the sharp interface model of two-phase Navier-Stokes equations was established in [107].

Recently, the errors of semi- or fully discrete ALE-FEMs for FSI problems were analyzed in [99, 100] without considering the errors in approximating the interface motion. To the best of our knowledge, the only proof of convergence of ALE-FEMs for FSI problems with a moving interface was given in [101] for rigid structures which can be modelled by ordinary differential equations instead of PDEs. Consequently, a comprehensive error analysis of ALE-FEMs for FSI problems that addresses the errors in approximating the interface motion, with elastic structures which need to be modelled by PDEs instead of

ODEs, is still lacking.

The objective of this chapter is to establish the optimal-order convergence in $L^\infty(0,T;H^1)$ of a monolithic fully discrete ALE-FEM with evolving mesh for a dynamic FSI problem with an unknown interface evolution driven by the solution of FSI problem. In particular, the errors of fluid velocity and structural displacement in $L^\infty(0,T;H^1)$ norm and the error of pressure in $L^\infty(0,T;L^2)$ norm are proved to be $O(\tau+h^k)$ for a semi-implicit Euler scheme with a high-order evolving mesh ALE-FEM described in [41], using a mixed finite element method with $P_k$–$P_{k-1}$–$P_k$ elements for velocity, pressure and structural displacement, respectively. This is achieved by designing an initial correction that improves the order of convergence in the error analysis of the developed ALE-FEM for the presented FSI problem, utilizing the matrix-vector formulation techniques developed in [95] for analyzing the errors in approximating solution-driven interface evolution in the FSI problem, as well as a mathematical induction which guarantees that the errors in approximating the evolving domain (including the unknown interface) will be uniformly bounded and sufficiently small when $\tau$ and $h$ are small enough and $\tau = o(h^{\frac{d}{2}})$.

Note that the coupling between the Navier-Stokes equations and the second-order hyperbolic system in the FSI problem, along with the continuity conditions (6.5) across the interface (which requires the velocity of fluid to match the velocity of structure on the interface), presents challenges in deriving stability estimates in the $L^\infty(0,T;H^1)$ norm for the ALE-FEM. In general, to derive $L^\infty(0,T;H^1)$ stability for individual Navier-Stokes equations or the hyperbolic system, one can choose the test function to be the time derivative of the solution. However, in the coupled system (6.7), this approach encounters difficulty when choosing the test function $\eta = \partial_t v$ to obtain $L^\infty(0,T;H^1)$ stability of the fluid velocity $v$. Therefore, we are compelled to choose $\xi = \partial_{tt} u$ due to the continuity condition (6.5a) across the interface, which induces the term $(\epsilon(u), \epsilon(\partial_{tt} u))_{\Omega_2}$. Integration by parts necessitates estimating the $H^1$ norm of $\partial_t u$, which cannot be directly controlled via energy estimates. To overcome this challenge, we can take an additional time derivative to the equation (6.7), and select the test function $(\eta, \xi) = (\partial_t v, \partial_{tt} u)$ in the newly obtained equation.

In our error analysis, we use this idea to establish the $H^1$ stability estimates of the fully discrete ALE-FEM for the coupled Navier-Stokes equations and second-order hyperbolic (structural) equation with some consistency error terms. In this case, we need to estimate the $L^2$ errors of $\partial_t v$ and $\partial_{tt} u$ at initial time. Then the standard geometric estimates using interpolation are inadequate, indicating that the discrete $H^{-1}$ norm of the consistency

error is $O(h^k)$. Utilizing inverse estimates, this implies that the discrete $L^2$ norm of the consistency error is of order $O(h^{k-1})$. To address this limitation, we introduce an initial correction specifically designed to enhance the convergence order, thereby allowing the discrete $L^2$ norm of the consistency error to achieve $O(h^k)$ at the initial time. Furthermore, we construct two analytic solutions for the FSI problem in both 2D and 3D for testing the convergence order of the numerical solutions.

The structure of this chapter is organized as follows: we present a realistic FSI model and its original weak form in Section 6.2. Then, we define the ALE mapping, ALE-based weak formulation of the FSI problem, as well as the fully discrete ALE-FEM in Section 6.3. We conduct a rigorous optimal error analysis for the developed fully discrete scheme in Section 6.4. Numerical experiments are carried out in Section 6.5 to validate the theoretical results.

**6.2 Model description.** We consider the motion of a viscous fluid and an elastic structure in a bounded domain $\Omega(t) \subset \mathbb{R}^d$ $(d = 2, 3)$ separated by a smooth interface $\Gamma(t)$ which is the common boundary of the fluid domain $\Omega_1(t)$ and structural domain $\Omega_2(t)$ at every time $t \in [0, T]$, i.e., $\Omega(t) = \Omega_1(t) \cup \Gamma(t) \cup \Omega_2(t)$, $\Omega_1(t) \cap \Omega_2(t) = \varnothing$, $\overline{\Omega}_1(t) \cap \overline{\Omega}_2(t) = \Gamma(t)$. Hence, the interface $\Gamma(t)$ moves along with the fluid and structure, as shown in Figure 6.1. For example, we are interested in a pressure-driven flow through the deformable channel

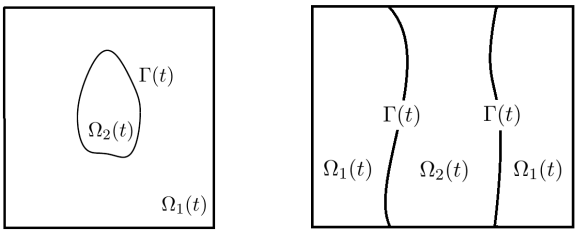



*Figure 6.1. Two schematic domain decompositions divided by the interface $\Gamma(t)$: (1) the immersed case (left); (2) the back-to-back case (right).*

with a two-way coupling between the incompressible fluid and the elastic material, where the fluid motion is defined in the current/Eulerian description, while the structure motion is defined in the initial/Lagrangian description.

The flow map $X(x, t) : \Omega_2(0) \to \Omega_2(t)$ of the structural domain is generated by the structural displacement $u$, and is defined as

$$X(x, t) = x + u(x, t), \quad \text{for } x \in \Omega_2(0).$$

The corresponding deformation gradient tensor and its Jacobian are defined as $F(x,t) = \nabla X(x,t)$ and $J(x,t) = \det(F(x,t))$, respectively, for $x \in \Omega_2(0)$. The fluid and structure motions are governed by a coupled system of partial differential equations (PDEs) as shown below.

The fluid motion is described by the following incompressible Navier–Stokes equations with respect to its velocity $v : \Omega_1(t) \to \mathbb{R}^d$ and pressure $p : \Omega_1(t) \to \mathbb{R}$ for $t \in [0,T]$,

$$\rho_1\Big(\frac{\partial v}{\partial t} + v \cdot \nabla v\Big) - \nabla \cdot \sigma_1 = \rho_1 f_1, \qquad \text{in } \bigcup_{t\in(0,T]} \Omega_1(t) \times \{t\}, \tag{6.1a}$$

$$\nabla \cdot v = 0, \qquad \text{in } \bigcup_{t\in(0,T]} \Omega_1(t) \times \{t\}, \tag{6.1b}$$

$$\sigma_1 n = 0, \qquad \text{on } \bigcup_{t\in(0,T]} [\partial\Omega_1(t)\backslash\Gamma(t)]_N \times \{t\}, \tag{6.1c}$$

$$v = v_0, \qquad \text{in } \Omega_1(0), \tag{6.1d}$$

where $\sigma_1$ is the stress tensor of fluid, defined by

$$\sigma_1 = -pI + 2\mu_1\mathbb{D}(v) \quad \text{with} \quad \mathbb{D}(v) := \frac{1}{2}(\nabla v + (\nabla v)^\top), \tag{6.2a}$$

with $\rho_1$ and $\mu_1$ being the density and viscosity of fluid, respectively.

The motion of elastic structure is described by the following linear elasticity equations with respect to its displacement $u : \Omega_2(0) \to \mathbb{R}^d$,

$$\rho_2\frac{\partial^2 u}{\partial t^2} - \nabla \cdot \sigma_2 = \rho_2 f_2, \qquad \text{in } \bigcup_{t\in(0,T]} \Omega_2(0) \times \{t\}, \tag{6.3a}$$

$$\sigma_2\hat{n} = 0, \qquad \text{on } \bigcup_{t\in(0,T]} [\partial\Omega_2(0)\backslash\Gamma(0)]_N \times \{t\}, \tag{6.3b}$$

$$u = u_0, \qquad \text{in } \Omega_2(0), \tag{6.3c}$$

$$\frac{\partial u}{\partial t} = w_0, \qquad \text{in } \Omega_2(0), \tag{6.3d}$$

where $w_0$ agrees with $v_0$ on the initial interface $\Gamma(0)$, $\sigma_2$ is the stress tensor of elastic structure, defined by

$$\sigma_2 = 2\mu_2\epsilon(u) + \lambda_2\nabla \cdot uI \quad \text{with} \quad \epsilon(u) := \frac{1}{2}(\nabla u + (\nabla u)^\top), \tag{6.4}$$

with $\rho_2$, $\mu_2$ and $\lambda_2$ being the density, shear modulus, and Lamé constant of the elastic structure, respectively, $\hat{n}$ denotes the unit normal vector on $\partial\Omega_2(0)$ including $\Gamma(0)$.

The fluid and elastic structure are coupled through the interface, satisfying the following interface conditions:

$$v \circ X = \partial_t u \qquad \text{on } \Gamma(0) \times (0,T], \tag{6.5a}$$

$$J(\sigma_1 \circ X) F^{-\top} \hat{n} = \sigma_2 \hat{n} \qquad \text{on } \Gamma(0) \times (0,T]. \tag{6.5b}$$

Due to the Piola transformation [149, §2.3, Definition 2 and Lemma 2], the following relations hold:

$$\begin{aligned}\int_{\Gamma(t)} \sigma_1 n \cdot \eta = \int_{\Gamma(t)} n \cdot \sigma_1 \eta = \int_{\Gamma(0)} \hat{n} \cdot J F^{-1} (\sigma_1 \circ X)(\eta \circ X) = \int_{\Gamma(0)} J(\sigma_1 \circ X) F^{-\top} \hat{n} \cdot (\eta \circ X) \\ = \int_{\Gamma(0)} \sigma_2 \hat{n} \cdot (\eta \circ X). \end{aligned} \tag{6.6}$$

This guarantees that, by testing (6.1a) with $\eta \in H^1(\Omega_1(t))^d$ and testing (6.3a) with $\xi \in H^1(\Omega_2(0))^d$, the interface terms from integration by parts are cancelled under the compatibility condition $\xi = \eta \circ X$ on $\Gamma(0)$.

Accordingly, the solutions of (6.1)–(6.5) satisfy the following weak formulation for $t \in (0,T]$: Find $v \in H^1(0,T; H^1(\Omega_1(t))^d)$, $u \in H^2(0,T; H^1(\Omega_2(0))^d)$ and $p \in L^2(0,T; L^2(\Omega_1(t)))$ satisfying (6.5a) such that

$$\begin{aligned}&\int_{\Omega_1(t)} \rho_1 \Big( \frac{\partial v}{\partial t} + v \cdot \nabla v \Big) \cdot \eta + \int_{\Omega_1(t)} \Big( 2\mu_1 \mathbb{D}(v) : \mathbb{D}(\eta) - p \nabla \cdot \eta \Big) + \int_{\Omega_1(t)} \nabla \cdot v q \\ &\quad + \int_{\Omega_2(0)} \rho_2 \frac{\partial^2 u}{\partial t^2} \cdot \xi + \int_{\Omega_2(0)} 2\mu_2 \epsilon(u) : \epsilon(\xi) + \lambda_2 (\nabla \cdot u)(\nabla \cdot \xi) \\ &= \int_{\Omega_1(t)} \rho_1 f_1 \cdot \eta + \int_{\Omega_2(0)} \rho_2 f_2 \cdot \xi, \\ &\forall\, \eta \in H^1(\Omega_1(t))^d,\ \xi \in H^1(\Omega_2(0))^d,\ q \in L^2(\Omega_1(t)),\ \text{where } \xi = \eta \circ X \text{ on } \Gamma(0). \end{aligned} \tag{6.7}$$

Note that the first interface condition, (6.5a), says that the fluid and structure moves with the same velocity on the interface. Thus the interface moves with the same velocity as well. If we use the same notation of the flow map for the elastic structure to denote the flow map for the interface by $X(\cdot,t) : \Gamma(0) \to \Gamma(t)$, then

$$\partial_t X = v \circ X \quad \text{on } \Gamma(0) \times (0,T]. \tag{6.8}$$

### 6.3 ALE mapping and ALE weak formulation for fully-discrete scheme.

Assume that the solution $u$ is smooth. We introduce a velocity field denoted by $w(X(x,t),t)$ that serves as a piecewise smooth extension of the structure's velocity, $\partial_t u \circ X^{-1}$, from $\Omega_2(t)$

to $\Omega(t)$. This extension is achieved by defining $w$ as a harmonic function within $\Omega_1(t)$ while $w \circ X = \partial_t u$ in $\Omega_2(0)$, and subject to the following boundary conditions:

$$-\Delta w = 0 \ \text{ in } \ \Omega_1(t), \quad w = 0 \ \text{ on } \ \partial\Omega_1(t)\backslash\Gamma(t), \ \ w = v \ \text{ on } \ \Gamma(t). \tag{6.9}$$

Thus $w$ is smooth in both subdomains $\Omega_1(t)$ and $\Omega_2(t)$, and $w_0$ denotes the initial velocity of $\Omega_1(0)$ and $\Omega_2(0)$ according to the condition (6.3d). We denote the piecewise smooth diffeomorphic flow map, generated by the piecewise smooth vector field $w$, as $X(\cdot,t)$ : $\Omega(0) \to \Omega(t)$ that is defined by

$$\partial_t X(x,t) = w(X(x,t),t) \quad \text{for } \ x \in \Omega(0), \tag{6.10}$$

with $\Omega_i(t) = X(\Omega_i(0),t)$ for $i = 1,2$ and $\Gamma(t) = X(\Gamma(0))$.

Let $\Omega_h(0)$ be the triangulated high-order isoparametric finite element domain with curved boundary and mesh fitting the interface $\Gamma(0)$ that serves as an approximation of $\Omega(0)$, where the subscript $h$ denotes the mesh size. Therefore, $\Omega_h(0) = \Omega_{1,h}(0) \cup \Gamma_h(0) \cup \Omega_{2,h}(0)$, where each $\Omega_{i,h}(0)$ is a triangulated high-order isoparametric finite element domain that serves as an approximation of $\Omega_i(0)$ for $i = 1,2$. The set of simplices in the triangulation/partition at $t = 0$ is denoted by $\mathscr{K}(0) = \mathscr{K}_1(0) \cup \mathscr{K}_2(0)$, where $\mathscr{K}_i(0)$ is the triangulation/partition of $\Omega_{i,h}(0)$ for $i = 1,2$. Since we do not need to use $\Omega_{2,h}(t)$, we will just use $\Omega_{2,h}$ to denote $\Omega_{2,h}(0)$.

**6.3.1 Interpolated ALE mapping and interpolated evolving mesh.** We define the interpolated ALE mapping $X^*_{1,h}(\cdot,t) = I^*_h X(\cdot,t)$ to be the piecewise polynomial interpolation of the piecewise smooth mapping $X(\cdot,t)$ : $\Omega_1(0) \to \Omega_1(t)$, where $I^*_h$ denotes the interpolation operator onto the finite element space $S^k_h(\Omega_{1,h}(0))$ of piecewise polynomial degree $k \geqslant 2$. Then we define the interpolated interface and subdomains as,

$$\Gamma^*_h(t) = X^*_{1,h}(\Gamma_h(0)) \ \text{ and } \ \Omega^*_{1,h}(t) = X^*_{1,h}(\Omega_{1,h}(0)) \quad \text{for } \ t \in [0,T].$$

Since the fluidic and structural velocities $v$ and $w$ match on the interface $\Gamma(t)$, it follows that their interpolations $I^*_h v$ and $I^*_h w$ match on the interpolated interface $\Gamma^*_h(t)$.

Correspondingly, we denote by $\mathscr{K}_1{}^*(t)$ the set of possibly curved simplices at time $t$ as the image of the initial simplices under the interpolated ALE mapping, i.e., $\mathscr{K}_1{}^*(t) = \{X^*_{1,h}(K,t) : K \in \mathscr{K}_1(0)\}$, and denote by $\Omega^*_{1,h}(t)$ the triangulated high-order isoparametric finite element domain with mesh determined by $\mathscr{K}_1{}^*(t)$. Let $\hat{K}$ denote the reference simplex. For each $K \in \mathscr{K}_1{}^*(t)$, there is a $k$-th order polynomial $F_K : \hat{K} \to K$ that maps $\hat{K}$ onto $K$.

The space of vector-valued finite element functions in $\Omega_{1,h}^*(t)$ is denoted by:

$$S_h^k(\Omega_{1,h}^*(t)) = \{\eta_h \in C^0(\Omega_{1,h}^*(t))^d : \eta_h \circ F_K \text{ is a polynomial of degree} \leqslant k \text{ for } \forall\ K \in \mathscr{K}_1^*(t)\},$$

and its subspace with zero boundary condition is denoted by $\mathring{S}_h^k(\Omega_{1,h}^*(t))$. When utilizing the Taylor–Hood finite element method to solve fluid equations, we establish the following scalar-valued finite element space for pressure

$$Q_h^{k-1}(\Omega_{1,h}^*(t)) = \{q_h \in C^0(\Omega_{1,h}^*(t)) : q_h \circ F_K \text{ is a polynomial of degree} \leqslant k-1 \text{ for } \forall\ K \in \mathscr{K}_1^*(t)\}.$$

If we denote by $w_h^*(\cdot,t) = I_h^* w(\cdot,t)$ the interpolated mesh velocity in $\Omega_{1,h}^*(t)$ which interpolates $\Omega(t)$, then the following relation holds:

$$\partial_t X_{1,h}^*(x,t) = w_h^*(X_{1,h}^*(x,t),t) \quad \text{for } x \in \Omega_{1,h}(0) \supset \Gamma_h(0).$$

The interpolated functions $X_{1,h}^*$ and $w_h^*$ are sufficiently good approximations to the smooth functions $X$ and $w$, satisfying the following error estimates:

$$\begin{aligned}
&\max_{t\in[0,T]} \Big( \|X_{1,h}^*(\cdot,t) - X(\cdot,t)\|_{L^\infty(\Omega_{1,h}(0))} + \|w_h^*(\cdot,t) - w(\cdot,t)\|_{L^\infty(\Omega_{1,h}^*(t))} \Big) \leqslant C h^{k+1}, \\
&\max_{t\in[0,T]} \Big( \|X_{1,h}^*(\cdot,t) - X(\cdot,t)\|_{W^{1,\infty}(\Omega_{1,h}(0))} + \|w_h^*(\cdot,t) - w(\cdot,t)\|_{W^{1,\infty}(\Omega_{1,h}^*(t))} \Big) \leqslant C h^{k}, \quad (6.11)
\end{aligned}$$

where the values of both $X$ and $w$ in $\Omega_{1,h}(0)$ and $\Omega_{1,h}^*(t)$ correspond to the values obtained by extending $X$ and $w$ from $\Omega_1(0)$ and $\Omega_1(t)$ to $\mathbb{R}^d$ smoothly. Hence, for sufficiently small mesh size $h$, the interpolated ALE mapping is invertible and

$$\max_{t\in[0,T]} \Big( \|X_{1,h}^*(\cdot,t)\|_{W^{1,\infty}(\Omega_{1,h}(0))} + \|(X_{1,h}^*)^{-1}(\cdot,t)\|_{W^{1,\infty}(\Omega_{1,h}^*(t))} \Big) \leqslant C. \quad (6.12)$$

Based on the interpolated ALE mapping $X_{1,h}^*(\cdot,t)$ and the interpolated mesh velocity $w_h^*$, we define the ALE time derivative of a function $\psi : \bigcup_{t\in[0,T]} \Omega_{1,h}^*(t) \times \{t\} \to \mathbb{R}$ as

$$\partial_{t,*}^\bullet \psi(x^*,t) = \frac{\mathrm{d}}{\mathrm{d}t} \psi(X_{1,h}^*(\xi,t),t) \quad \text{at } x^* = X_{1,h}^*(\xi,t) \in \Omega_{1,h}^*(t) \text{ for } \xi \in \Omega_{1,h}(0). \quad (6.13)$$

Hence, by applying the chain rule of partial differentiation, one obtains

$$\partial_{t,*}^\bullet \psi(x^*,t) = \partial_t \psi(x^*,t) + w_h^*(x^*,t) \cdot \nabla \psi(x^*,t) \quad \text{for } x^* \in \Omega_{1,h}^*(t).$$

The finite element nodes of $\Omega_{1,h}^*(t)$ are denoted as $\{\zeta_{1,i}^*(t)\}_{i=1}^{N_1}$, and we define the vector $\mathbf{x}_1^*(t)$ to collect all these nodes as follows:

$$\mathbf{x}_1^*(t) = (\zeta_{1,1}^*(t), \cdots, \zeta_{1,N_1}^*(t)).$$

In the vector-valued space $S_h^k(\Omega_{1,h}^*(t))$, the corresponding finite element basis functions are represented as $\phi_{1,i}[\mathbf{x}_1^*(t)]$ for $1 \leqslant i \leqslant dN_1$. Since $\phi_{1,i}[\mathbf{x}_1^*(t)](X_h^*(\cdot,t)) = \phi_{1,i}[\mathbf{x}_1^*(0)](\cdot)$, these basis functions exhibit a transport property, which can be expressed as:

$$\partial_{t,*}^\bullet \phi_{1,i}[\mathbf{x}_1^*(t)] = 0, \quad \text{for} \quad 1 \leqslant i \leqslant dN_1. \tag{6.14}$$

There exist smooth extensions of $v$ and $u$ from $\Omega_1(t)$ and $\Omega_2(0)$ to $\mathbb{R}^d$, respectively. To simplify our notations, we shall continue to refer to these extended functions as $v$ and $u$ in $\mathbb{R}^d$. We establish the following standard approximation properties for the Lagrangian interpolation:

$$\|v - I_h^* v\|_{L^\infty(\Omega_{1,h}^*(t))} + h\|v - I_h^* v\|_{W^{1,\infty}(\Omega_{1,h}^*(t))} \leqslant Ch^{k+1}, \tag{6.15a}$$

$$\|u - I_h^* u\|_{L^\infty(\Omega_{2,h})} + h\|u - I_h^* u\|_{W^{1,\infty}(\Omega_{2,h})} \leqslant Ch^{k+1}. \tag{6.15b}$$

**6.3.2 Weak formulation on the interpolated evolving mesh.** For the triangulation which fits the interface $\Gamma(t)$, there exists a Lipschitz continuous mapping $\Phi_h^0(\cdot) : \Omega_h(0) \to \Omega(0)$ that maps $\Gamma_h(0)$ to $\Gamma(0)$. The mapping $\Phi_h^0$ induces a Lipschitz continuous mapping $\Phi_{1,h}^*(\cdot,t) : \Omega_{1,h}^*(t) \to \Omega_1(t)$ that maps $\Gamma_h^*(t)$ to $\Gamma(t)$ and satisfies the relationship given by

$$\Phi_{1,h}^*(\cdot,t) = X(\cdot,t) \circ \Phi_h^0(\cdot) \circ X_{1,h}^*(\cdot,t)^{-1} \quad \text{in} \quad \Omega_{1,h}^*(t). \tag{6.16}$$

Furthermore, this mapping adheres to the following estimates:

$$\|\Phi_{1,h}^*(\cdot,t) - \mathrm{id}\|_{L^\infty(\Omega_{1,h}^*(t))} + h\|\Phi_{1,h}^*(\cdot,t) - \mathrm{id}\|_{W^{1,\infty}(\Omega_{1,h}^*(t))} \leqslant Ch^{k+1}, \tag{6.17a}$$

$$\|\Phi_h^0 - \mathrm{id}\|_{L^\infty(\Omega_{2,h})} + h\|\Phi_h^0 - \mathrm{id}\|_{W^{1,\infty}(\Omega_{2,h})} \leqslant Ch^{k+1}. \tag{6.17b}$$

Utilizing the relation (6.16), we can demonstrate that for $1 \leqslant i \leqslant dN_1$, it holds that:

$$\partial_t^\bullet \big(\phi_{1,i}^*[\mathbf{x}_1(t)] \circ (\Phi_{1,h}^*)^{-1}\big) = 0. \tag{6.18}$$

Consider the finite element functions $\eta_h \in S_h^k(\Omega_{1,h}^*(t))$ and $\xi_h \in S_h^k(\Omega_{2,h})$, which share agreement along the interface $\Gamma_h^*(t)$ in the sense that as $\xi_h = \eta_h \circ X_{1,h}^*$ on $\Gamma_h(0)$. We define $\eta = \eta_h \circ (\Phi_{1,h}^*)^{-1}$ and $\xi = \xi_h \circ (\Phi_h^0)^{-1}$ as functions in the respective smooth subdomains, $\Omega_1(t)$ and $\Omega_2(0)$. These functions maintain agreement along the interface $\Gamma(t)$. Specifically, we have:

$$\xi = \eta_h \circ X_{1,h}^* \circ (\Phi_h^0)^{-1} = \eta \circ \Phi_{1,h}^* \circ X_{1,h}^* \circ (\Phi_h^0)^{-1} = \eta \circ X, \quad \text{on } \Gamma(0).$$

As a result, the weak formulation in (6.7) remains valid when employing $\eta$ and $\xi$ as test functions. Similarly, for $q_h \in Q_h^{k-1}(\Omega_{1,h}^*(t))$, we define $q = q_h \circ (\Phi_{1,h}^*)^{-1}$ as a function in $\Omega_1(t)$.

By using the properties (6.14) and (6.18), we can deduce that the functions $\eta_h$, $q_h$, $\eta$ and $q$ satisfy the similar properties:

$$\partial^\bullet_{t,*}\eta_h = 0, \quad \partial^\bullet_{t,*}q_h = 0, \quad \partial^\bullet_t \eta = 0, \quad \partial^\bullet_t q = 0. \tag{6.19}$$

In order to deform the integrations in the weak formulation from $\Omega_1(t)$ and $\Omega_2(0)$ to $\Omega^*_{1,h}(t)$ and $\Omega_{2,h}$, respectively, we utilize the following type of estimates for the error caused by domain perturbation (see a proof in [95, Lemma 8.1]):

$$\Big| \int_{\Omega^*_{1,h}(t)} \rho_1 \Big( \frac{\partial v}{\partial t} + v \cdot \nabla v \Big) \circ \Phi^*_{1,h} \cdot \eta \circ \Phi^*_{1,h} - \int_{\Omega_1(t)} \rho_1 \Big( \frac{\partial v}{\partial t} + v \cdot \nabla v \Big) \cdot \eta \Big| \leqslant C h^k \|\eta_h\|_{L^2(\Omega^*_{1,h}(t))}.$$

As a result, by leveraging the triangle inequality, interpolation error estimates, and the commutative property of the interpolation operator $I^*_h$ and the material derivative $\partial^\bullet_{t,*}$, we can further replace $v$ by its interpolation $v^*_h = I^*_h v$ with an additional remainder of $O(h^k)$, i.e.,

$$\Big| \int_{\Omega^*_{1,h}(t)} \rho_1 \Big( \frac{\partial v^*_h}{\partial t} + v^*_h \cdot \nabla v^*_h \Big) \cdot \eta_h - \int_{\Omega_1(t)} \rho_1 \Big( \frac{\partial v}{\partial t} + v \cdot \nabla v \Big) \cdot \eta \Big| \leqslant C h^k \|\eta_h\|_{L^2(\Omega^*_{1,h}(t))}.$$

Similarly,

$$\Big| \int_{\Omega^*_{1,h}(t)} \Big( 2\mu_1 \mathbb{D}(v^*_h) : \mathbb{D}(\eta_h) - p^*_h \nabla \cdot \eta_h \Big) - \int_{\Omega_1(t)} \Big( 2\mu_1 \mathbb{D}(v) : \mathbb{D}(\eta) - p \nabla \cdot \eta \Big) \Big| \leqslant C h^k \|\eta_h\|_{H^1(\Omega^*_{1,h}(t))},$$

$$\Big| \int_{\Omega^*_{1,h}(t)} q_h \nabla \cdot v^*_h - \int_{\Omega_1(t)} q \nabla \cdot v \Big| \leqslant C h^k \|q_h\|_{L^2(\Omega^*_{1,h}(t))},$$

$$\Big| \int_{\Omega_{2,h}} \rho_2 \frac{\partial^2 u^*_h}{\partial t^2} \cdot \xi_h - \int_{\Omega_2(0)} \rho_2 \frac{\partial^2 u}{\partial t^2} \cdot \xi \Big| \leqslant C h^k \|\xi_h\|_{L^2(\Omega_{2,h})},$$

$$\Big| \int_{\Omega_{2,h}} 2\mu_2 \epsilon(u^*_h) : \epsilon(\xi_h) + \lambda_2 (\nabla \cdot u^*_h)(\nabla \cdot \xi_h) - \int_{\Omega_2(0)} 2\mu_2 \epsilon(u) : \epsilon(\xi) + \lambda_2 (\nabla \cdot u)(\nabla \cdot \xi) \Big|$$
$$\leqslant C h^k \|\xi_h\|_{H^1(\Omega_{2,h})}.$$

Therefore, the weak formulation in (6.7) can be reduced to

$$\begin{aligned}
&\int_{\Omega^*_{1,h}(t)} \rho_1 \Big( \partial^\bullet_{t,*} v^*_h + (v^*_h - w^*_h) \cdot \nabla v^*_h \Big) \cdot \eta_h + \int_{\Omega^*_{1,h}(t)} \Big( 2\mu_1 \mathbb{D}(v^*_h) : \mathbb{D}(\eta_h) - p^*_h \nabla \cdot \eta_h \Big) \\
&+ \int_{\Omega^*_{1,h}(t)} q_h \nabla \cdot v^*_h + \int_{\Omega_{2,h}(0)} \rho_2 \frac{\partial^2 u^*_h}{\partial t^2} \cdot \xi_h + \int_{\Omega_{2,h}} 2\mu_2 \epsilon(u^*_h) : \epsilon(\xi_h) + \lambda_2 (\nabla \cdot u^*_h)(\nabla \cdot \xi_h) \\
&= \int_{\Omega^*_{1,h}(t)} \rho_1 f_1 \cdot \eta_h + \int_{\Omega_{2,h}} \rho_2 f_2 \cdot \xi_h + \mathcal{E}_1(t, \eta_h) + \mathcal{E}_2(t, \xi_h) + \mathcal{E}_3(t, q_h), \qquad \forall\, t \in (0, T],
\end{aligned} \tag{6.20}$$

where $\mathcal{E}_1(t,\eta_h)$, $\mathcal{E}_2(t,\xi_h)$ and $\mathcal{E}_3(t,q_h)$ are some linear functionals acting on $\eta_h$, $\xi_h$ and $q_h$, satisfying the following estimates:

$$|\mathcal{E}_1(t,\eta_h)| \leqslant Ch^k\|\eta_h\|_{H^1(\Omega^*_{1,h}(t))} \quad \text{and} \quad |\mathcal{E}_2(t,\xi_h)| \leqslant Ch^k\|\xi_h\|_{H^1(\Omega_{2,h})}, \tag{6.21a}$$

$$|\mathcal{E}_3(t,q_h)| \leqslant Ch^k\|q_h\|_{L^2(\Omega^*_{1,h}(t))}, \tag{6.21b}$$

which are later considered as the consistency errors of the numerical scheme. Using the same method, we can prove that temporal derivatives of these consistency errors satisfy similar estimates:

$$\left|\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}_1(t,\eta_h)\right| \leqslant Ch^k\|\eta_h\|_{H^1(\Omega^*_{1,h}(t))} \quad \text{and} \quad \left|\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}_2(t,\xi_h)\right| \leqslant Ch^k\|\xi_h\|_{H^1(\Omega_{2,h})}, \tag{6.22a}$$

$$\left|\frac{\mathrm{d}^2}{\mathrm{d}t^2}\mathcal{E}_1(t,\eta_h)\right| \leqslant Ch^k\|\eta_h\|_{H^1(\Omega^*_{1,h}(t))} \quad \text{and} \quad \left|\frac{\mathrm{d}^2}{\mathrm{d}t^2}\mathcal{E}_2(t,\xi_h)\right| \leqslant Ch^k\|\xi_h\|_{H^1(\Omega_{2,h})}, \tag{6.22b}$$

$$\left|\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}_3(t,q_h)\right| \leqslant Ch^k\|q_h\|_{L^2(\Omega^*_{1,h}(t))}. \tag{6.22c}$$

**6.3.3 ALE mapping and weak formulation.** The above weak formulation is based on the interpolated ALE mapping and interpolated mesh velocity. However, these interpolated ALE mapping and mesh velocity are unknown in practice. To address this limitation in practical computations, we introduce an alternative weak formulation based on numerically computed ALE mapping $X_{1,h}(\cdot,t) \in S_h^k(\Omega_{1,h}(0))$ and ALE mesh velocity $w_h(\cdot,t)$.

We denote by $\mathscr{K}_1(t)$ the set of potentially curved simplices at time $t$, which results from the initial simplices under the ALE mapping, i.e., $\mathscr{K}_1(t) = \{X_{1,h}(K,t) : K \in \mathscr{K}_1(0)\}$. Furthermore, we designate $\Omega_{1,h}(t)$ as the finite element domain determined by $\mathscr{K}_1(t)$. The spaces of vector-valued and scalar-valued finite element functions in $\Omega_{1,h}(t)$ are represented as:

$$S_h^k(\Omega_{1,h}(t)) := \{\eta_h \in C^0(\Omega_{1,h}(t))^d : \eta_h \circ F_K \text{ is a polynomial of degree} \leqslant k \text{ for } \forall\ K \in \mathscr{K}_1(t)\},$$
$$Q_h^{k-1}(\Omega_{1,h}(t)) := \{q_h \in C^0(\Omega_{1,h}(t)) : q_h \circ F_K \text{ is a polynomial of degree} \leqslant k-1 \text{ for } \forall\ K \in \mathscr{K}_1(t)\}.$$

The subspace of $S_h^k(\Omega_{1,h}(t))$ incorporating zero boundary condition is denoted as $\mathring{S}_h^k(\Omega_{1,h}(t))$. Similarly, the space of vector-valued finite element functions in $\Omega_{2,h}$ is denoted as $S_h^k(\Omega_{2,h})$, while its subspace with zero boundary condition is designated as $\mathring{S}_h^k(\Omega_{2,h})$. Finally, approximate interface and subdomain can be expressed as follows:

$$\Gamma_h(t) = X_{1,h}(\Gamma_h(0)), \quad \Omega_{1,h}(t) = X_{1,h}(\Omega_{1,h}(0)) \quad \text{for } t \in [0,T].$$

The ALE mapping $X_{1,h}(\cdot,t)$ is chosen to be the flow map generated by a mesh velocity $w_h(\cdot,t)$:

$$\partial_t X_{1,h} = w_h \circ X_{1,h} \quad \text{in } \Omega_{1,h}(0)\times(0,T], \tag{6.23}$$

where $w_h$ is a discrete harmonic finite element function in the numerically computed subdomain $\Omega_{1,h}(t)$ with boundary condition matching the discrete velocity of the fluid/structure, $v_h = \partial_t u_h \circ X_{1,h}^{-1}$, on the numerically computed interface $\Gamma_h(t)$, i.e.,

$$\begin{aligned} \int_{\Omega_{1,h}(t)} \nabla w_h : \nabla \chi_h = 0, \quad & \forall\, \chi_h \in \mathring{S}_h^k(\Omega_{1,h}(t)), \\ w_h = v_h, \quad & \text{on } \Gamma_h(t), \\ w_h = 0, \quad & \text{on } \partial\Omega_{1,h}(t)\backslash\Gamma_h(t). \end{aligned} \tag{6.24}$$

Similarly to the definition in (6.13), we can define the discrete material derivative $\partial_{t,h}^{\bullet}$ by using the numerical ALE mapping $X_{1,h}(\cdot,t)$. And similarly to (6.20), we can replace $w_h$ by the interpolation $w_h^* = I_h^* w$ with a reminder $O(h^k)$, where $w$ is defined in (6.9). Thus,

$$\int_{\Omega_{1,h}^*(t)} \nabla w_h^* : \nabla \chi_h = \mathcal{E}_4(t,\chi_h), \quad \forall \chi_h \in \mathring{S}_h^k(\Omega_{1,h}^*(t)). \tag{6.25}$$

with a consitency error $\mathcal{E}_4(t,\chi_h)$ satisfying the following estimate:

$$\Big|\mathcal{E}_4(t,\chi_h)\Big| \leqslant C h^k \|\chi_h\|_{H^1(\Omega_{1,h}^*(t))}. \tag{6.26}$$

By dropping the remainders in (6.20), we define the FEM weak formulation based on the ALE mapping and the corresponding evolving mesh as follows: Find $v_h \in S_h^k(\Omega_{1,h}(t))$, $p_h \in Q_h^{k-1}(\Omega_{1,h}(t))$ and $u_h \in S_h^k(\Omega_{2,h})$ such that

$$\begin{aligned} &\int_{\Omega_{1,h}(t)} \rho_1\Big(\partial_{t,h}^{\bullet} v_h + (v_h - w_h)\cdot\nabla v_h\Big)\cdot\eta_h + \int_{\Omega_{1,h}(t)} \Big(2\mu_1 \mathbb{D}(v_h) : \mathbb{D}(\eta_h) - p_h \nabla\cdot\eta_h\Big) \\ &+ \int_{\Omega_{1,h}(t)} q_h \nabla\cdot v_h + \int_{\Omega_{2,h}} \rho_2 \frac{\partial^2 u_h}{\partial t^2}\cdot\xi_h + \int_{\Omega_{2,h}} 2\mu_2 \epsilon(u_h) : \epsilon(\xi_h) + \lambda_2(\nabla\cdot u_h)(\nabla\cdot\xi_h) \\ &= \int_{\Omega_{1,h}(t)} \rho_1 f_1\cdot\eta_h + \int_{\Omega_{2,h}} \rho_2 f_2\cdot\xi_h, \qquad \forall\, t\in(0,T], \end{aligned} \tag{6.27}$$

for test functions $\eta_h \in S_h^k(\Omega_{1,h}(t))$, $q_h \in Q_h^{k-1}(\Omega_{1,h}(t))$ and $\xi_h \in S_h^k(\Omega_{2,h})$ satisfying the condition $\eta_h \circ X_{1,h}(\cdot,t) = \xi_h$ on $\Gamma_h(0)$.

Taking the divergence of equation (6.1a), we derive the pressure equation at time $t = 0$, which is expressed as:

$$-\Delta p = \rho_1 \nabla v_0 : \nabla v_0^{\top} - \rho_1 \nabla\cdot f_1(\cdot,0), \quad \text{in } \Omega_1(0),$$

$$p = 2\mu_1 n^\top \mathbb{D}(v_0)n \ \text{ on } \ \partial\Omega_1(0)\backslash\Gamma(0), \quad p = 2\mu_1 n^\top \mathbb{D}(v_0)n - J^{-1}n^\top \left(\sigma_2(0)\circ X^{-1}\right) F^\top n \ \text{ on } \ \Gamma(0). \tag{6.28}$$

To solve this elliptic equation, we can employ standard finite element methods of order $(k-1)$ and $k$. The solutions obtained through these methods are represented as $\bar{p}_h \in Q_h^{k-1}(\Omega_{1,h}(0))$ and $\hat{p}_h \in Q_h^k(\Omega_{1,h}(0))$, respectively. Utilizing the standard error estimates for the elliptic equation, we obtain the following inequalities:

$$\|\bar{p}_h - p\|_{L^2(\Omega_{1,h}(0))} + h\|\nabla(\bar{p}_h - p)\|_{L^2(\Omega_{1,h}(0))} \leqslant Ch^k, \tag{6.29a}$$

$$\|\hat{p}_h - p\|_{L^2(\Omega_{1,h}(0))} + h\|\nabla(\hat{p}_h - p)\|_{L^2(\Omega_{1,h}(0))} \leqslant Ch^{k+1}. \tag{6.29b}$$

Similarly to [94] and [76], to obtain the optimal order of convergence, we modify the equation (6.27) by adding two time-independent correction terms and obtain the following spatially semi-discrete scheme:

$$\begin{aligned}
&\int_{\Omega_{1,h}(t)} \rho_1\Big(\partial_{t,h}^\bullet v_h + (v_h - w_h)\cdot\nabla v_h\Big)\cdot\eta_h + \int_{\Omega_{1,h}(t)} \Big(2\mu_1\mathbb{D}(v_h) : \mathbb{D}(\eta_h) - p_h\nabla\cdot\eta_h\Big) \\
&+ \int_{\Omega_{1,h}(t)} q_h\nabla\cdot v_h + \int_{\Omega_{2,h}} \rho_2\frac{\partial^2 u_h}{\partial t^2}\cdot\xi_h + \int_{\Omega_{2,h}} 2\mu_2\epsilon(u_h) : \epsilon(\xi_h) + \lambda_2(\nabla\cdot u_h)(\nabla\cdot\xi_h) \\
&= \int_{\Omega_{1,h}(t)} \rho_1 f_1\cdot\eta_h + \int_{\Omega_{2,h}} \rho_2 f_2\cdot\xi_h + \vartheta_1(\eta_h) + \vartheta_2(\xi_h) \qquad \forall\, t\in(0,T],
\end{aligned} \tag{6.30}$$

for test functions $\eta_h \in S_h^k(\Omega_{1,h}(t))$, $q_h \in Q_h^{k-1}(\Omega_{1,h}(t))$ and $\xi_h \in S_h^k(\Omega_{2,h})$ satisfying the condition $\eta_h \circ X_{1,h}(\cdot,t) = \xi_h$ on $\Gamma_h(0)$, where the correction terms $\vartheta_1(\eta_h)$ and $\vartheta_2(\xi_h)$ are defined as

$$\begin{aligned}
\vartheta_1(\eta_h) :=& \int_{\Omega_{1,h}(0)} \Big(2\mu_1\mathbb{D}(I_h^* v_0) : \mathbb{D}(\eta_h\circ X_{1,h}(\cdot,t)) - \bar{p}_h\nabla\cdot(\eta_h\circ X_{1,h}(\cdot,t))\Big) \\
&+ \int_{\Omega_{1,h}(0)} (\mu_1 I_h^*\Delta v_0 - \nabla\hat{p}_h)\cdot(\eta_h\circ X_{1,h}(\cdot,t)),
\end{aligned} \tag{6.31a}$$

$$\vartheta_2(\xi_h) := \int_{\Omega_{2,h}} \Big(2\mu_2\epsilon(I_h^* u_0) : \epsilon(\xi_h) + \lambda_2(\nabla\cdot I_h^* u_0)(\nabla\cdot\xi_h) + (I_h^*\nabla\cdot\sigma_2(\cdot,0))\cdot\xi_h\Big). \tag{6.31b}$$

Note that $\vartheta_1(\eta_h)$ and $\vartheta_2(\xi_h)$ solely rely on the nodal values of $\eta_h$ and $\xi_h$, respectively, without additional dependence on time $t$. Consider $\vartheta_1$ and $\vartheta_2$ as the operators acting on $\eta_h$ and $\xi_h$, respectively. we can establish the estimate for the discrete $H^{-1}$-norm of $\vartheta_1 + \vartheta_2$. Let $\hat{\eta}_h = \eta_h \circ X_{1,h}$ and $\hat{\eta} = \eta\circ X$, then $\hat{\eta}_h \in S_h^k(\Omega_{1,h}(0))$. Through a method analogous to the proof of (6.21), along with the estimates provided in (6.29), we obtain:

$\vartheta_1(\eta_h) + \vartheta_2(\xi_h)$

$$
\begin{aligned}
&= \int_{\Omega_{1,h}(0)} 2\mu_1 \Big( \mathbb{D}(I_h^* v_0) - \mathbb{D}(v_0) \Big) : \mathbb{D}(\hat{\eta}_h) + (p(0) - \bar{p}_h) \nabla \cdot \hat{\eta}_h + \int_{\Omega_{1,h}(0)} \mu_1 (I_h^* \Delta v_0 - \Delta v_0) \cdot \hat{\eta}_h \\
&\quad - \int_{\Omega_{1,h}(0)} (\nabla \hat{p}_h - \nabla p(0)) \cdot \hat{\eta}_h + \int_{\Omega_{1,h}(0)} \Big( 2\mu_1 \mathbb{D}(v_0) : \mathbb{D}(\hat{\eta}_h) - p(0) \nabla \cdot \hat{\eta}_h + (\nabla \cdot \sigma_1(\cdot, 0)) \cdot \hat{\eta}_h \Big) \\
&\quad - \int_{\Omega_1(0)} \Big( 2\mu_1 \mathbb{D}(v_0) : \mathbb{D}(\hat{\eta}) - p(0) \nabla \cdot \hat{\eta} + (\nabla \cdot \sigma_1(\cdot, 0)) \cdot \hat{\eta} \Big) + \int_{\Omega_{2,h}} 2\mu_2 \Big( \epsilon(I_h^* u_0) - \epsilon(u_0) \Big) : \mathbb{D}(\xi_h) \\
&\quad + \int_{\Omega_{2,h}} \lambda_2 (\nabla \cdot I_h^* u_0 - \nabla \cdot u_0)(\nabla \cdot \xi_h) + (I_h^* \nabla \cdot \sigma_2(\cdot, 0) - \nabla \cdot \sigma_2(\cdot, 0)) \cdot \xi_h \\
&\quad + \int_{\Omega_{2,h}} \Big( 2\mu_2 \epsilon(u_0) : \epsilon(\xi_h) + \lambda_2 (\nabla \cdot u_0)(\nabla \cdot \xi_h) + (\nabla \cdot \sigma_2(\cdot, 0)) \cdot \xi_h \Big) \\
&\quad - \int_{\Omega_2(0)} \Big( 2\mu_2 \epsilon(u_0) : \epsilon(\xi) + \lambda_2 (\nabla \cdot u_0)(\nabla \cdot \xi) + (\nabla \cdot \sigma_2(\cdot, 0)) \cdot \xi \Big) \\
&\leqslant C h^k \Big( \|\hat{\eta}_h\|_{H^1(\Omega_{1,h}(0))} + \|\xi_h\|_{H^1(\Omega_{2,h}(0))} \Big) \leqslant C h^k \Big( \|\eta_h\|_{H^1(\Omega_{1,h}(t))} + \|\xi_h\|_{H^1(\Omega_{2,h}(0))} \Big). \quad (6.32)
\end{aligned}
$$

Furthermore, if we estimate $\vartheta_1$, $\vartheta_2$, $\mathcal{E}_1$ and $\mathcal{E}_2$ together, we can improve the result given in (6.32) by establishing the estimate for the discrete $L^2$-norm of $\vartheta_1 - \mathcal{E}_1(\cdot, 0) + \vartheta_2 - \mathcal{E}_2(\cdot, 0)$. For $\eta_h \in S_h^k(\Omega_{1,h}(0))$, by utilizing (6.20), (6.31) and the original equations (6.1) and (6.3), we obtain:

$$
\begin{aligned}
&\vartheta_1(\eta_h) - \mathcal{E}_1(\eta_h, 0) + \vartheta_2(\xi_h) - \mathcal{E}_2(\xi_h, 0) \\
&= \int_{\Omega_1(0)} \rho_1 \Big( \partial_t^\bullet v + (v_0 - w) \cdot \nabla v_0 - f_1 \Big) \cdot \eta - \int_{\Omega_{1,h}(0)} \rho_1 \Big( \partial_{t,*}^\bullet v_h^* + (v_h^* - w_h^*) \cdot \nabla v_h^* - f_1 \Big) \cdot \eta_h \\
&\quad - \int_{\Omega_{1,h}(0)} (\bar{p}_h - p_h^*) \nabla \cdot \eta_h + \int_{\Omega_{1,h}(0)} (\mu_1 I_h^* \Delta v_0 - \nabla \hat{p}_h) \cdot \eta_h - \int_{\Omega_1(0)} (\mu_1 \Delta v_0 - \nabla p) \cdot \eta \\
&\quad + \int_{\Omega_2(0)} \Big( \rho_2 \frac{\partial^2 u}{\partial t^2} - \rho_2 f_2 - \nabla \cdot \sigma_2(\cdot, 0) \Big) \cdot \xi - \int_{\Omega_{2,h}} \Big( \rho_2 \frac{\partial^2 u_h^*}{\partial t^2} - \rho_2 f_2 - I_h^* \nabla \cdot \sigma_2(\cdot, 0) \Big) \cdot \xi_h.
\end{aligned}
$$

Notice that there are no derivatives for $\eta_h$, $\eta$, $\xi_h$ and $\xi$ except for the term $(\bar{p_h} - p_h^*, \nabla \cdot \eta_h)_{\Omega_{1,h}(0)}$ on the right-hand side. Therefore, similarly to the deduction of the consistency error bounds (6.21), and by considering the interpolation errors (6.15a) along with the error estimates in (6.29), we obtain:

$$
\begin{aligned}
&\vartheta_1(\eta_h) - \mathcal{E}_1(\eta_h, 0) + \vartheta_2(\xi_h) - \mathcal{E}_2(\xi_h, 0) \\
&\leqslant C h^k \Big( \|\eta_h\|_{L^2(\Omega_{1,h}(0))}^2 + \|\xi_h\|_{L^2(\Omega_{2,h})}^2 \Big) - \int_{\Omega_{1,h}(0)} (\bar{p}_h - p_h^*) \nabla \cdot \eta_h. \quad (6.33)
\end{aligned}
$$

If $\eta_h$ is chosen to be discrete divergence free or with small remainders, then the result (6.33) is expected to be better than (6.32).

It's worth emphasizing that the ALE mapping $X_{1,h}(\cdot,t)$ is contingent upon the velocity field $w_h$, and this velocity field is dependent on the numerical solution $v_h$ on $\Gamma_h(t)$. Consequently, the dynamic evolution of the mesh, as determined by equations (6.23)–(6.24), is intricately coupled with the finite element solutions of (6.30).

**6.3.4 Fully-discrete scheme.** We proceed to fully discretize the spatially semi-discrete weak formulation (6.30) by employing the semi-implicit Euler scheme for temporal discretization. Throughout this study, the superscript $n$ denote the time level $t_n$.

We define the pull back operator $P_h^{m,n}: C^0(\Omega_{1,h}^m) \to C^0(\Omega_{1,h}^n)$ as follows:

$$P_h^{m,n}\chi_h^m = \chi_h^m \circ X_{1,h}^m \circ (X_{1,h}^n)^{-1} \quad \text{for} \quad \forall \chi_h^m \in C^0(\Omega_{1,h}^m).$$

Let $\dot{u}_h^n$ represent the numerical approximation of $\partial_t u(\cdot,t_n)$. The fully-discrete scheme is formulated as follows: Given $X_{1,h}^{n-1} \in S_h^k(\Omega_{1,h}^0)$, $v_h^{n-1} \in S_h^k(\Omega_{1,h}^{n-1})$, $\dot{u}_h^{n-1} \in S_h^k(\Omega_{2,h})$, and $u_h^{n-1} \in S_h^k(\Omega_{2,h})$, seek $w_h^{n-1} \in S_h^k(\Omega_{1,h}^{n-1})$, $X_{1,h}^n \in S_h^k(\Omega_{1,h}^0)$, $v_h^n \in S_h^k(\Omega_{1,h}^n)$, $p_h^n \in Q_h^{k-1}(\Omega_{1,h}^n)$, $\dot{u}_h^n \in S_h^k(\Omega_{2,h})$, and $u_h^n \in S_h^k(\Omega_{2,h})$ satisfying $\dot{u}_h^n = v_h^n \circ X_{1,h}^n$ on $\Gamma_h^n$ and the following equations:

$$\int_{\Omega_{1,h}^{n-1}} \nabla w_h^{n-1} : \nabla \chi_h = 0, \;\; w_h^{n-1} = v_h^{n-1} \;\text{ on }\; \Gamma_h^{n-1}, \;\; w_h^{n-1} = 0 \;\text{ on }\; \partial\Omega_{1,h}^{n-1}\backslash\Gamma_h^{n-1}, \tag{6.34a}$$

$$\frac{X_{1,h}^n - X_{1,h}^{n-1}}{\tau} = w_h^{n-1} \circ X_{1,h}^{n-1}, \tag{6.34b}$$

$$\begin{aligned}
&\int_{\Omega_{1,h}^n} \rho_1\Big(\frac{v_h^n - P_h^{n-1,n}v_h^{n-1}}{\tau} + P_h^{n-1,n}(v_h^{n-1} - w_h^{n-1})\cdot\nabla v_h^n\Big)\cdot\eta_h + 2\mu_1\mathbb{D}(v_h^n):\mathbb{D}(\eta_h) \\
&+ \int_{\Omega_{1,h}^n} q_h\nabla\cdot v_h^n - p_h^n\nabla\cdot\eta_h + \int_{\Omega_{2,h}} \rho_2\frac{\dot{u}_h^n - \dot{u}_h^{n-1}}{\tau}\cdot\xi_h + 2\mu_2\epsilon(u_h^n):\epsilon(\xi_h) + \lambda_2(\nabla\cdot u_h^n)(\nabla\cdot\xi_h) \\
&= \int_{\Omega_{1,h}^n} \rho_1 f_1^n\cdot\eta_h + \int_{\Omega_{2,h}} \rho_2 f_2^n\cdot\xi_h + \vartheta_1(\eta_h) + \vartheta_2(\xi_h),
\end{aligned} \tag{6.34c}$$

$$\frac{u_h^n - u_h^{n-1}}{\tau} = \dot{u}_h^n, \tag{6.34d}$$

for any test functions $\chi_h \in \mathring{S}_h^k(\Omega_{1,h}^{n-1}), q_h \in Q_h^{k-1}(\Omega_{1,h}^n), \eta_h \in S_h^k(\Omega_{1,h}^n)$ and $\xi_h \in S_h^k(\Omega_{2,h})$ that satisfy condition $\eta_h \circ X_{1,h}^n = \xi_h$ on the interface $\Gamma_h^0$. The flow map $X_{1,h}^n$ can be solved from the first two equations (6.34a) and (6.34b). Then the domain $\Omega_{1,h}^n$ is defined

as $X^n_{1,h}(\Omega^0_{1,h})$. And from equations (6.34c) and (6.34d), we can finally obtain $v^n_h, p^n_h, \dot{u}^n_h$ and $u^n_h$.

We assume that the inf-sup condition is satisfied, as discussed in [15, Chapter 12]. For any $q_h \in Q^{k-1}_h(\Omega_{1,h}(t))$, there is a positive constant $\kappa$ independent of $h$ and $t$, such that

$$\kappa \|q_h\|_{L^2(\Omega^n_{1,h})} \leqslant \sup_{0\neq\eta_h\in S^k_h(\Omega_{1,h}(t))} \frac{(\nabla\cdot\eta_h, q_h)_{\Omega^n_{1,h}}}{\|\eta_h\|_{H^1(\Omega^n_{1,h})}}. \tag{6.35}$$

**6.4 Error estimates.** The error estimates involve a comparison between (6.34) and (6.20), which are weak formulations defined in different domains. Consequently, the error analysis generally fits within the framework of evolving finite element methods for nonlinear PDEs on evolving meshes. In this context, the techniques and methods developed in references such as [56, 41, 93, 95] can be effectively employed for establishing error estimates.

Throughout this chapter, we use the symbol $C$ to represent a generic positive constant. It depends solely on the dimension $d$, the smooth domain $\Omega$, and the solution $(v, p, u)$, but remains independent of the mesh size $h$ and time step size $\tau$.

**6.4.1 Matrix-vector form of the weak formulations.** It is convenient to represent the weak formulations (6.20), (6.34) in matrix-vector form. Let $\{\zeta^n_{1,i}\}_{i=1}^{N_1}$ represent the finite element nodes of $\Omega^n_{1,h}$. Define the vector $\mathbf{x}^n_1 := (\zeta^n_{1,1}, \cdots, \zeta^n_{1,N_1})$ as the collection of all finite element nodes within $\Omega^n_{1,h}$. The finite element basis functions of the vector-valued space $S^k_h(\Omega^n_{1,h})$ are denoted as $\phi_{1,i}[\mathbf{x}^n_1]$ for $1 \leqslant i \leqslant dN_1$.

By utilizing basis functions, the numerical solution $v^n_h$ can be represented as a vector $\mathbf{v}^n = (v^n_1, \cdots, v^n_{dN_1})$, where each component $v^n_i \in \mathbb{R}$ is defined by:

$$v^n_h = \sum_{i=1}^{dN_1} v^n_i \phi_{1,i}[\mathbf{x}^n_1].$$

Similarly, the solutions $p^n_h$ and $w^n_h$ within $\Omega^n_{1,h}$ can also be represented as vectors $\mathbf{p}^n$ and $\mathbf{w}^n_1$, respectively. The vector representation of $X^n_{1,h}$ corresponds to $\mathbf{x}^n_1$. We introduce domain-dependent mass matrix $\mathbf{M}_f(\mathbf{x}^n_1)$, the stiffness matrix $\mathbf{A}_f(\mathbf{x}^n_1)$, and the matrix $\mathbf{C}_f(\mathbf{x}^n_1)$ in $\Omega^n_{1,h}$ as follows:

$$\mathbf{M}_f(\mathbf{x}^n_1)|_{ij} := \rho_1 \int_{\Omega^n_{1,h}} \phi_{1,i}[\mathbf{x}^n_1]\cdot\phi_{1,j}[\mathbf{x}^n_1] \ \text{ for } \ 1 \leqslant i,j \leqslant dN_1,$$

$$\mathbf{A}_f(\mathbf{x}^n_1)|_{ij} := \int_{\Omega^n_{1,h}} \nabla\phi_{1,i}[\mathbf{x}^n_1] : \nabla\phi_{1,j}[\mathbf{x}^n_1] \ \text{ for } \ 1 \leqslant i,j \leqslant dN_1,$$

$$\mathbf{C}_f(\mathbf{x}_1^n)|_{ij} := 2\mu_1 \int_{\Omega_{1,h}^n} \mathbb{D}(\phi_{1,i}[\mathbf{x}_1^n]) : \mathbb{D}(\phi_{1,j}[\mathbf{x}_1^n]) \text{ for } 1 \leqslant i,j \leqslant dN_1.$$

The mass and stiffness matrices associated with the finite element spaces $S_h^k(\Omega_{1,h}^n)$ and $Q_h^{k-1}(\Omega_{1,h}^n)$ are distinct. However, for the sake of simplicity, we will use the same notations unless ambiguity arises within the context. By utilizing the matrix-vector form, the norms of the finite element functions over the domain $\Omega_{1,h}^n$ can be expressed as quadratic forms:

$$\begin{aligned} \rho_1 \|v_h^n\|^2_{L^2(\Omega_{1,h}^n)} &= \mathbf{v}^n \cdot \mathbf{M}_f(\mathbf{x}_1^n)\mathbf{v}^n =: \|\mathbf{v}^n\|^2_{\mathbf{M}_f(\mathbf{x}_1^n)}, \\ \|\nabla w_h^n\|^2_{L^2(\Omega_{1,h}^n)} &= \mathbf{w}_1^n \cdot \mathbf{A}_f(\mathbf{x}_1^n)\mathbf{w}_1^n =: \|\mathbf{w}_1^n\|^2_{\mathbf{A}_f(\mathbf{x}_1^n)}, \\ 2\mu_1 \|\mathbb{D}(v_h^n)\|^2_{L^2(\Omega_{1,h}^n)} &= \mathbf{v}^n \cdot \mathbf{C}_f(\mathbf{x}_1^n)\mathbf{v}^n =: \|\mathbf{v}^n\|^2_{\mathbf{C}_f(\mathbf{x}_1^n)}. \end{aligned} \tag{6.36}$$

In the same way, we can define the matrices $\mathbf{D}_f(\mathbf{x}_1^n)$ and $\mathbf{B}_f(\mathbf{x}_1^n, \mathbf{w}_1^n)$ and the vector $\mathbf{f}_1^n$ such that

$$\mathbf{p}^n \cdot \mathbf{D}_f(\mathbf{x}_1^n)\mathbf{v}^n := \int_{\Omega_{1,h}^n} p_h^n \nabla \cdot v_h^n, \quad \text{for} \quad \forall\, p_h^n \in Q_h^{k-1}(\Omega_{1,h}^n),\ v_h^n \in S_h^k(\Omega_{1,h}^n),$$

$$\boldsymbol{\eta} \cdot \mathbf{B}_f(\mathbf{x}_1^n, \mathbf{w}_1^n)\mathbf{v}^n := \rho_1 \int_{\Omega_{1,h}^n} (w_h^n \cdot \nabla v_h^n) \cdot \eta_h, \quad \text{for} \quad \forall\, w_h^n, v_h^n, \eta_h \in S_h^k(\Omega_{1,h}^n),$$

$$\mathbf{f}_1^n|_i = \rho_1 \int_{\Omega_{1,h}^n} \phi_{1,i}[\mathbf{x}_1^n] \cdot f_1^n \quad \text{for} \quad 1 \leqslant i \leqslant dN_1,$$

where $\boldsymbol{\eta}$ is the vector representation of $\eta_h$.

In the domain $\Omega_{2,h}$, similar definitions apply. Let $\mathbf{x}_2$ denote the vector that collects all finite element nodes within $\Omega_{2,h}$. Let $\dot{\mathbf{u}}^n$ and $\mathbf{u}^n$ represent the vector forms of the functions $\dot{u}_h^n$ and $u_h^n$ in $\Omega_{2,h}$. We define $\mathbf{f}_2^n$ as the vector that comprises the integrals of $\rho_2 f_2^n$ and the basis functions of $S_h^k(\Omega_{2,h})$, mirroring the definition of $\mathbf{f}_1^n$. For any function $\xi_h \in S_h^k(\Omega_{2,h})$, having a vector representation denoted as $\boldsymbol{\xi}$, the matrices $\mathbf{M}_s(\mathbf{x}_2)$ and $\mathbf{E}_s(\mathbf{x}_2)$ are defined as follows:

$$\begin{aligned} \boldsymbol{\xi} \cdot \mathbf{M}_s(\mathbf{x}_2)\mathbf{u}^n &= \int_{\Omega_{2,h}} \rho_2 u_h^n \cdot \xi_h, \\ \boldsymbol{\xi} \cdot \mathbf{E}_s(\mathbf{x}_2)\mathbf{u}^n &= \int_{\Omega_{2,h}} 2\mu_2 \epsilon(u_h^n) : \epsilon(\xi_h) + \lambda_2 (\nabla \cdot u_h^n)(\nabla \cdot \xi_h). \end{aligned}$$

We define the difference operator $\delta$ between two time levels $n$ and $n-1$ such that:

$$\delta \mathbf{v}^n := \mathbf{v}^n - \mathbf{v}^{n-1}.$$

By using the matrix-vector form, the fully-discrete scheme (6.34) can be expressed as follows:

$$\mathbf{A}_f(\mathbf{x}_1^{n-1})\mathbf{w}_1^{n-1}=0,\quad \mathbf{w}_1^{n-1}|_{\Gamma_h^{n-1}}=\mathbf{v}^{n-1}|_{\Gamma_h^{n-1}},\quad \mathbf{w}_1^{n-1}|_{\partial\Omega_{1,h}^{n-1}\setminus\Gamma_h^{n-1}}=0, \tag{6.37a}$$

$$\frac{\mathbf{x}_1^n-\mathbf{x}_1^{n-1}}{\tau}=\mathbf{w}_1^{n-1}, \tag{6.37b}$$

$$\boldsymbol{\eta}\cdot\Big(\mathbf{M}_f(\mathbf{x}_1^n)\frac{\mathbf{v}^n-\mathbf{v}^{n-1}}{\tau}+\mathbf{B}_f(\mathbf{x}_1^n,\mathbf{v}^{n-1}-\mathbf{w}_1^{n-1})\mathbf{v}^n\Big)+\boldsymbol{\eta}\cdot\mathbf{C}_f(\mathbf{x}_1^n)\mathbf{v}^n-\boldsymbol{\eta}\cdot\mathbf{D}_f(\mathbf{x}_1^n)^\top\mathbf{p}^n \tag{6.37c}$$

$$+\boldsymbol{\xi}\cdot\mathbf{M}_s(\mathbf{x}_2)\frac{\dot{\mathbf{u}}^n-\dot{\mathbf{u}}^{n-1}}{\tau}+\boldsymbol{\xi}\cdot\mathbf{E}_s(\mathbf{x}_2)\mathbf{u}^n=\boldsymbol{\eta}\cdot\mathbf{f}_1^n+\boldsymbol{\xi}\cdot\mathbf{f}_2^n+\boldsymbol{\eta}\cdot\boldsymbol{\vartheta}_1+\boldsymbol{\xi}\cdot\boldsymbol{\vartheta}_2,$$

$$\mathbf{D}_f(\mathbf{x}_1^n)\mathbf{v}^n=0, \tag{6.37d}$$

$$\frac{\mathbf{u}^n-\mathbf{u}^{n-1}}{\tau}=\dot{\mathbf{u}}^n, \tag{6.37e}$$

where $\mathbf{v}^n$ and $\dot{\mathbf{u}}^n$ are vectors that should agree on the interface $\Gamma_h^n$ in the sense that $(v_h^n\circ X_{1,h}^n)|_{\Gamma_h^0}=\dot{u}_h^n|_{\Gamma_h^0}$, $\boldsymbol{\eta}$ and $\boldsymbol{\xi}$ are vectors that should agree on the interface $\Gamma_h^n$ in the sense that $(\eta_h\circ X_{1,h}^n)|_{\Gamma_h^0}=\xi_h|_{\Gamma_h^0}$. The vectors $\boldsymbol{\vartheta}_1$ and $\boldsymbol{\vartheta}_2$ represent the vector forms of $\vartheta_1(\cdot)$ and $\vartheta_2(\cdot)$, respectively, i.e., $\boldsymbol{\eta}\cdot\boldsymbol{\vartheta}_1=\vartheta_1(\eta_h)$, $\boldsymbol{\xi}\cdot\boldsymbol{\vartheta}_2=\vartheta_2(\xi_h)$. Notably, the vectors $\boldsymbol{\vartheta}_1$ and $\boldsymbol{\vartheta}_2$ are independent of $n$.

Similarly, we extend the definitions of matrices and vectors to the interpolated domain $\Omega_h^{*,n}$. The objects denoted with a superscript $*$ represent the interpolation of exact solutions. Specifically, the vectors $\mathbf{v}^{*,n}$, $\mathbf{p}^{*,n}$, $\mathbf{w}_1^{*,n}$, $\mathbf{u}^{*,n}$, $\mathbf{x}_1^{*,n}$, $\boldsymbol{\eta}^*$ collect the nodal values of $v_h^*(\cdot,t_n)$, $p_h^*(\cdot,t_n)$, $w_h^*(\cdot,t_n)$, $u_h^*(\cdot,t_n)$, $X_{1,h}^*(\cdot,t_n)$, and $\eta_h^*$. Additionally, we define the following derivatives:

$$\dot{\mathbf{x}}_1^{*,n}=\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x}_1^*(t_n),\quad \dot{\mathbf{u}}^{*,n}=\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^*(t_n),\quad \ddot{\mathbf{u}}^{*,n}=\frac{\mathrm{d}^2}{\mathrm{d}t^2}\mathbf{u}^*(t_n).$$

The weak formulation (6.20) can be rewritten in the following form:

$$\mathbf{A}_f(\mathbf{x}_1^{*,n-1})\mathbf{w}_1^{*,n-1}=\boldsymbol{\mathcal{E}}_4^{n-1},\quad \mathbf{w}_1^{*,n-1}|_{\Gamma_h^{*,n-1}}=\mathbf{v}^{*,n-1}|_{\Gamma_h^{*,n-1}},\quad \mathbf{w}_1^{*,n-1}|_{\partial\Omega_{1,h}^{*,n-1}\setminus\Gamma_h^{*,n-1}}=0, \tag{6.38a}$$

$$\frac{\mathbf{x}_1^{*,n}-\mathbf{x}_1^{*,n-1}}{\tau}=\mathbf{w}_1^{*,n-1}+\boldsymbol{\mathcal{F}}_1^{n-1}, \tag{6.38b}$$

$$\boldsymbol{\eta}^*\cdot\Big(\mathbf{M}_f(\mathbf{x}_1^{*,n})\frac{\mathbf{v}^{*,n}-\mathbf{v}^{*,n-1}}{\tau}+\mathbf{B}_f(\mathbf{x}_1^{*,n},\mathbf{v}^{*,n-1}-\mathbf{w}_1^{*,n-1})\mathbf{v}^{*,n}\Big)+\boldsymbol{\eta}^*\cdot\mathbf{C}_f(\mathbf{x}_1^{*,n})\mathbf{v}^{*,n}$$

$$
\begin{aligned}
&-\boldsymbol{\eta}^*\cdot\mathbf{D}_f(\mathbf{x}_1^{*,n})^\top\mathbf{p}^{*,n}+\boldsymbol{\xi}\cdot\mathbf{M}_s(\mathbf{x}_2)\frac{\dot{\mathbf{u}}^{*,n}-\dot{\mathbf{u}}^{*,n-1}}{\tau}+\boldsymbol{\xi}\cdot\mathbf{E}_s(\mathbf{x}_2)\mathbf{u}^{*,n}\\
=&\boldsymbol{\eta}^*\cdot\mathbf{f}_1^{*,n}+\boldsymbol{\xi}\cdot\mathbf{f}_2^{*,n}+\boldsymbol{\eta}^*\cdot\boldsymbol{\mathcal{E}}_1^n+\boldsymbol{\xi}\cdot\boldsymbol{\mathcal{E}}_2^n+\boldsymbol{\eta}^*\cdot\boldsymbol{\mathcal{F}}_2^n+\boldsymbol{\xi}\cdot\boldsymbol{\mathcal{F}}_3^n, &(6.38c)\\
&\mathbf{D}_f(\mathbf{x}_1^{*,n})\mathbf{v}^{*,n}=\boldsymbol{\mathcal{E}}_3^n, &(6.38d)\\
&\frac{\mathbf{u}^{*,n}-\mathbf{u}^{*,n-1}}{\tau}=\dot{\mathbf{u}}^{*,n}+\boldsymbol{\mathcal{F}}_4^n, &(6.38e)
\end{aligned}
$$

where $\mathbf{v}^{*,n}$ and $\dot{\mathbf{u}}^{*,n}$ are vectors that should agree on the interface $\Gamma_h^{*,n}$ in the sense that $(v_h^{*,n}\circ X_{1,h}^{*,n})|_{\Gamma_h^0}=\dot{u}_h^{*,n}|_{\Gamma_h^0}$, $\boldsymbol{\eta}^*$ and $\boldsymbol{\xi}$ are vectors that should agree on the interface $\Gamma_h^{*,n}$ in the sense that $(\eta_h^*\circ X_{1,h}^{*,n})|_{\Gamma_h^0}=\xi_h|_{\Gamma_h^0}$. The vectors $\boldsymbol{\mathcal{E}}_1^n$, $\boldsymbol{\mathcal{E}}_2^n$, $\boldsymbol{\mathcal{E}}_3^n$, and $\boldsymbol{\mathcal{E}}_4^{n-1}$ represent the vectorized forms of the defect terms $\mathcal{E}_1(t_n,\cdot)$, $\mathcal{E}_2(t_n,\cdot)$, $\mathcal{E}_3(t_n,\cdot)$, and $\mathcal{E}_4(t_{n-1},\cdot)$ defined in (6.20) and (6.25). When nodal values are considered, we can choose $\boldsymbol{\eta}=\boldsymbol{\eta}^*$. By Taylor's expansion, it is easy to derive that the consistency errors $\boldsymbol{\mathcal{F}}_1^{n-1}$, $\boldsymbol{\mathcal{F}}_2^n$, $\boldsymbol{\mathcal{F}}_3^n$ and $\boldsymbol{\mathcal{F}}_4^n$ satisfy the following estimates:

$$
\begin{aligned}
&\|\boldsymbol{\mathcal{F}}_1^{n-1}\|_{\mathbf{M}_f(\mathbf{x}_1^0)}+\|\boldsymbol{\mathcal{F}}_1^{n-1}\|_{\mathbf{A}_f(\mathbf{x}_1^0)}\leqslant C\tau, \qquad &|\boldsymbol{\eta}\cdot\boldsymbol{\mathcal{F}}_2^n|\leqslant C\tau\|\boldsymbol{\eta}\|_{\mathbf{M}_f(\mathbf{x}_1^{*,n})} &(6.39a)\\
&|\boldsymbol{\xi}\cdot\boldsymbol{\mathcal{F}}_3^n|\leqslant C\tau\|\boldsymbol{\xi}\|_{\mathbf{M}_s(\mathbf{x}_2)}, &\|\boldsymbol{\mathcal{F}}_4^n\|_{\mathbf{M}_s(\mathbf{x}_2)}+\|\boldsymbol{\mathcal{F}}_4^n\|_{\mathbf{E}_s(\mathbf{x}_2)}\leqslant C\tau, &(6.39b)\\
&|\boldsymbol{\eta}\cdot(\delta\boldsymbol{\mathcal{F}}_2^n)|\leqslant C\tau^2\|\eta\|_{\mathbf{M}_f(\mathbf{x}_1^{*,n})}, &|\boldsymbol{\xi}\cdot(\delta\boldsymbol{\mathcal{F}}_3^n)|\leqslant C\tau^2\|\xi\|_{\mathbf{M}_s(\mathbf{x}_2)}, &(6.39c)\\
&|\boldsymbol{\eta}\cdot(\delta^2\boldsymbol{\mathcal{F}}_2^n)|\leqslant C\tau^3\|\eta\|_{\mathbf{M}_f(\mathbf{x}_1^{*,n})}, &|\boldsymbol{\xi}\cdot(\delta^2\boldsymbol{\mathcal{F}}_3^n)|\leqslant C\tau^3\|\xi\|_{\mathbf{M}_s(\mathbf{x}_2)}, &(6.39d)\\
&\|\delta\boldsymbol{\mathcal{F}}_4^n\|_{\mathbf{M}_s(\mathbf{x}_2)}+\|\delta\boldsymbol{\mathcal{F}}_4^n\|_{\mathbf{E}_s(\mathbf{x}_2)}\leqslant C\tau^2. &&(6.39e)
\end{aligned}
$$

Subtracting (6.38) from (6.37), we obtain equations describing the errors: $\mathbf{e}_{\mathbf{x}_1}^n=\mathbf{x}_1^n-\mathbf{x}_1^{*,n}$, $\mathbf{e}_{\mathbf{w}_1}^n=\mathbf{w}_1^n-\mathbf{w}_1^{*,n}$, $\mathbf{e}_{\mathbf{v}}^n=\mathbf{v}^n-\mathbf{v}^{*,n}$, $\mathbf{e}_{\mathbf{p}}^n=\mathbf{p}^n-\mathbf{p}^{*,n}$, $\mathbf{e}_{\mathbf{u}}^n=\mathbf{u}^n-\mathbf{u}^{*,n}$, and $\dot{\mathbf{e}}_{\mathbf{u}}^n=\dot{\mathbf{u}}^n-\dot{\mathbf{u}}^{*,n}$.

$$
\begin{aligned}
&\mathbf{A}_f(\mathbf{x}_1^{n-1})\mathbf{e}_{\mathbf{w}_1}^{n-1}=-(\mathbf{A}_f(\mathbf{x}_1^{n-1})-\mathbf{A}_f(\mathbf{x}_1^{*,n-1}))\mathbf{w}_1^{*,n-1}-\boldsymbol{\mathcal{E}}_4^{n-1}, &(6.40a)\\
&\mathbf{e}_{\mathbf{w}_1}^{n-1}|_{\Gamma_h^{*,n-1}}=\mathbf{e}_{\mathbf{v}}^{n-1}|_{\Gamma_h^{*,n-1}},\quad \mathbf{e}_{\mathbf{w}_1}^{n-1}|_{\partial\Omega_{1,h}^{*,n-1}\setminus\Gamma_h^{*,n-1}}=0,
\end{aligned}
$$

$$
\frac{\mathbf{e}_{\mathbf{x}_1}^n-\mathbf{e}_{\mathbf{x}_1}^{n-1}}{\tau}=\mathbf{e}_{\mathbf{w}_1}^{n-1}-\boldsymbol{\mathcal{F}}_1^{n-1}, \tag{6.40b}
$$

$$
\boldsymbol{\eta}\cdot\Big(\mathbf{M}_f(\mathbf{x}_1^n)\frac{\mathbf{e}_{\mathbf{v}}^n-\mathbf{e}_{\mathbf{v}}^{n-1}}{\tau}+\mathbf{C}_f(\mathbf{x}_1^n)\mathbf{e}_{\mathbf{v}}^n-\mathbf{D}_f(\mathbf{x}_1^n)^\top\mathbf{e}_{\mathbf{p}}^n\Big)+\boldsymbol{\xi}\cdot\Big(\mathbf{M}_s(\mathbf{x}_2)\frac{\dot{\mathbf{e}}_{\mathbf{u}}^n-\dot{\mathbf{e}}_{\mathbf{u}}^{n-1}}{\tau}+\mathbf{E}_s(\mathbf{x}_2)\mathbf{e}_{\mathbf{u}}^n\Big) \tag{6.40c}
$$

$$
\begin{aligned}
=&-\boldsymbol{\eta}\cdot(\mathbf{M}_f(\mathbf{x}_1^n)-\mathbf{M}_f(\mathbf{x}_1^{*,n}))\frac{\mathbf{v}^{*,n}-\mathbf{v}^{*,n-1}}{\tau}-\boldsymbol{\eta}\cdot\mathbf{B}_f(\mathbf{x}_1^n,\mathbf{v}^{n-1}-\mathbf{w}_1^{n-1})\mathbf{v}^n\\
&+\boldsymbol{\eta}\cdot\mathbf{B}_f(\mathbf{x}_1^{*,n},\mathbf{v}^{*,n-1}-\mathbf{w}_1^{*,n-1})\mathbf{v}^{*,n}-\boldsymbol{\eta}\cdot(\mathbf{C}_f(\mathbf{x}_1^n)-\mathbf{C}_f(\mathbf{x}_1^{*,n}))\mathbf{v}^{*,n}
\end{aligned}
$$

$$+\boldsymbol{\eta}\cdot(\mathbf{D}_f(\mathbf{x}_1^n)^\top-\mathbf{D}_f(\mathbf{x}_1^{*,n})^\top)\mathbf{p}^{*,n}+\boldsymbol{\eta}\cdot(\mathbf{f}_1^n-\mathbf{f}_1^{*,n})$$
$$+\boldsymbol{\eta}\cdot(\boldsymbol{\vartheta}_1-\boldsymbol{\mathcal{E}}_1^n-\boldsymbol{\mathcal{F}}_2^n)+\boldsymbol{\xi}\cdot(\boldsymbol{\vartheta}_2-\boldsymbol{\mathcal{E}}_2^n-\boldsymbol{\mathcal{F}}_3^n),\quad \mathbf{e}_{\mathbf{v}}^n|_{\Gamma_h^n}=\dot{\mathbf{e}}_{\mathbf{u}}^n|_{\Gamma_h^0},$$

$$\mathbf{D}_f(\mathbf{x}_1^n)\mathbf{e}_{\mathbf{v}}^n=-(\mathbf{D}_f(\mathbf{x}_1^n)-\mathbf{D}_f(\mathbf{x}_1^{*,n}))\mathbf{v}^{*,n}-\boldsymbol{\mathcal{E}}_3^n, \tag{6.40d}$$

$$\frac{\mathbf{e}_{\mathbf{u}}^n-\mathbf{e}_{\mathbf{u}}^{n-1}}{\tau}=\dot{\mathbf{e}}_{\mathbf{u}}^n-\boldsymbol{\mathcal{F}}_4^n. \tag{6.40e}$$

**6.4.2 Comparison of integrals over two different domains.** To prove the convergence of the numerical schemes, it is necessary to compare the integrals over two domains: $\Omega_{1,h}^{*,n}$ and $\Omega_{1,h}^n$, $\Omega_{1,h}^{n-1}$ and $\Omega_{1,h}^n$. In this subsection, we will establish a sequence of Lemmas that are used to compare the norms of finite element functions with same nodal values over two different finite element domains by constructing a continuous deformation between them.

To handle the rate of change of an integral over a moving domain, we will frequently make use of the following lemma:

**Lemma 6.1.** *If the domain $\Omega(t)$ moves with velocity $w\in W^{1,\infty}(\Omega(t))$, then we have*

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

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

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

$$\partial_t^\bullet\nabla v=\nabla\partial_t^\bullet v-\nabla v\nabla w. \tag{6.41}$$

Let $\mathbf{y},\mathbf{z}\in\mathbb{R}^{dN_1}$ be two nodal vectors which define the discrete finite element domains $\Omega_{1,h}[\mathbf{y}]$ and $\Omega_{1,h}[\mathbf{z}]$, respectively. Let $\mathbf{e}=(e_j):=\mathbf{y}-\mathbf{z}\in\mathbb{R}^{dN_1}$. By means of a linear homolopy, the intermediate domain $\Omega_{1,h}^\theta:=\Omega_{1,h}[\mathbf{z}+\theta\mathbf{e}]$ changes continuously from $\Omega_{1,h}[\mathbf{z}]$ to $\Omega_{1,h}[\mathbf{y}]$ when the parameter $\theta$ takes values in $[0,1]$.

For a vector $\mathbf{v}=(v_1,\cdots,v_{dN_1})\in\mathbb{R}^{dN_1}$, we define the function $v_h^\theta\in S_h^k(\Omega_{1,h}^\theta)$ as

$$v_h^\theta:=\sum_{i=1}^{dN_1}v_i\phi_{1,i}[\mathbf{z}+\theta\mathbf{e}]. \tag{6.42}$$

(6.42) implies that there is a one to one correspondence between $\mathbb{R}^{dN_1}$ and $S_h^k(\Omega_{1,h}^\theta)$. Hence, we define the functions $e_h^\theta$, $\eta_h^\theta$, $p_h^\theta$ defined on $\Omega_{1,h}^\theta$ that correspond to the vectors $\mathbf{e}$, $\boldsymbol{\eta}$, $\mathbf{p}$.

The following lemma is a direct result of Lemma 6.1.

**Lemma 6.2.** *[38, Lemma 5.1] In the aforementioned setup, let $\partial_\theta^\bullet$ denote the material derivative with respect to the velocity field $e_h^\theta$. Similarly to the transport property* (6.14)*, we have $\partial_\theta^\bullet v_h^\theta = 0$. Subsequently, employing* (6.41)*, we obtain the following equalities:*

$$\begin{aligned}
\boldsymbol{\eta}\cdot(\mathbf{M}_f(\mathbf{y})-\mathbf{M}_f(\mathbf{z}))\mathbf{v} &= \int_0^1\int_{\Omega_{1,h}^\theta} \eta_h^\theta\cdot v_h^\theta\nabla\cdot e_h^\theta \mathrm{d}x\mathrm{d}\theta,\\
\boldsymbol{\eta}\cdot(\mathbf{A}_f(\mathbf{y})-\mathbf{A}_f(\mathbf{z}))\mathbf{v} &= \int_0^1\int_{\Omega_{1,h}^\theta} \Big[\nabla\eta_h^\theta(\nabla\cdot e_h^\theta I_d-\nabla e_h^\theta-(\nabla e_h^\theta)^\top)\Big]:\nabla v_h^\theta \mathrm{d}x\mathrm{d}\theta,\\
\boldsymbol{\eta}\cdot(\mathbf{C}_f(\mathbf{y})-\mathbf{C}_f(\mathbf{z}))\mathbf{v} &= \int_0^1\int_{\Omega_{1,h}^\theta} \mathbb{D}(\eta_h^\theta):\mathbb{D}(v_h^\theta)\nabla\cdot e_h^{n,\theta}-\mathbb{D}(\eta_h^\theta):\big(\nabla v_h^\theta\nabla e_h^\theta\big)\\
&\quad -\big(\nabla\eta_h^\theta\nabla e_h^\theta\big):\mathbb{D}(v_h^\theta)\mathrm{d}x\mathrm{d}\theta,\\
\boldsymbol{\eta}\cdot(\mathbf{D}_f(\mathbf{y})-\mathbf{D}_f(\mathbf{z}))^\top\mathbf{p} &= \int_0^1\int_{\Omega_{1,h}^\theta} -p_h^\theta\nabla\eta_h^\theta:(\nabla e_h^\theta)^\top+p_h^\theta\nabla\cdot\eta_h^\theta\nabla\cdot e_h^\theta\mathrm{d}x\mathrm{d}\theta.
\end{aligned}$$

Based on Lemma 6.2, we can infer that when the $W^{1,\infty}$ semi-norm of $e_h^\theta$ is bounded, it enables a comparison of norms over the two distinct domains, namely $\Omega_{1,h}[\mathbf{y}]$ and $\Omega_{1,h}[\mathbf{z}]$. Furthermore, the application of Gronwall's inequality allows us to establish the following lemma.

**Lemma 6.3.** *[38, Lemma 5.2] If $\|\nabla e_h^\theta\|_{L^\infty(\Omega_{1,h}^\theta)}\leqslant\alpha$ for $0\leqslant\theta\leqslant 1$, then the following inequalities hold for $0\leqslant\theta\leqslant 1$:*

$$\begin{aligned}
\|\mathbf{v}\|_{\mathbf{M}_f(\mathbf{z}+\theta\mathbf{e})} &\leqslant Ce^{C\alpha}\|\mathbf{v}\|_{\mathbf{M}_f(\mathbf{z})},\\
\|\mathbf{v}\|_{\mathbf{A}_f(\mathbf{z}+\theta\mathbf{e})} &\leqslant Ce^{C\alpha}\|\mathbf{v}\|_{\mathbf{A}_f(\mathbf{z})},\\
\|\mathbf{v}\|_{\mathbf{C}_f(\mathbf{z}+\theta\mathbf{e})} &\leqslant C_{\alpha,\varepsilon}\|\mathbf{v}\|_{\mathbf{C}_f(\mathbf{z})}+\varepsilon\|\mathbf{v}\|_{\mathbf{A}_f(\mathbf{z})},
\end{aligned}$$

*where $\varepsilon$ is an arbitrary real number in $(0,1)$, $C_{\alpha,\varepsilon}$ is a constant that depends on $\alpha$ and $\varepsilon$.*

According to the following lemma, the condition in Lemma 6.3 for $\nabla e_h^\theta$ can be relaxed to $\theta=0$.

**Lemma 6.4.** *[38, Lemma 5.3] If $\|\nabla e_h^0\|_{L^\infty(\Omega_{1,h}^0)}\leqslant\frac{1}{2}$, then the finite element function $v_h^\theta\in S_h^k(\Omega_{1,h}^\theta)$, with $0\leqslant\theta\leqslant 1$, satisfies the following estimate:*

$$\|\nabla v_h^\theta\|_{L^p(\Omega_{1,h}^\theta)}\leqslant C_p\|\nabla v_h^0\|_{L^p(\Omega_{1,h}^0)}\quad \textit{for}\quad 1\leqslant p\leqslant\infty,$$

*where $C_p$ is a constant that only depends on $p$.*

In combination with the above lemmas, the definition of the norms in (6.36) implies the following conclusion under the condition in Lemma 6.4:

$$\begin{aligned} &\text{the norms } \|\cdot\|_{\mathbf{M}_f(\mathbf{z}+\theta\mathbf{e})} \text{ are } h\text{-uniformly equivalent for } 0 \leqslant \theta \leqslant 1, \\ &\text{and so are the norms } \|\cdot\|_{\mathbf{A}_f(\mathbf{z}+\theta\mathbf{e})}. \end{aligned} \tag{6.43}$$

Then by using Korn's inequality (see [28, Theorem 1.1-2]), the estimate (6.17a), Lemma 6.3 and the equivalence (6.43), the following norms are $h$-uniformly equivalent when $h$ is sufficiently small for $0 \leqslant \theta \leqslant 1$ under the condition in Lemma 6.4.

$$\|\cdot\|_{\mathbf{M}_f(\mathbf{z}+\theta\mathbf{e})} + \|\cdot\|_{\mathbf{A}_f(\mathbf{z}+\theta\mathbf{e})} \sim \|\cdot\|_{\mathbf{M}_f(\mathbf{z}+\theta\mathbf{e})} + \|\cdot\|_{\mathbf{C}_f(\mathbf{z}+\theta\mathbf{e})}. \tag{6.44}$$

**6.4.3 Convergence of evolving finite element approximations.** We are now ready to present the main result of this chapter, which demonstrates the convergence of the evolving finite element methods to the problems (6.1)–(6.8). This convergence is outlined in Theorem 6.1.

**Theorem 6.1.** *We assume that the solutions of the fluid-structure interaction problems* (6.1)–(6.8)*, the flow map* $X : \Omega(0) \times [0,T] \to \mathbb{R}^d$*, and the inverse flow map* $X^{-1}$ *are sufficiently smooth in* $\Omega_i(t)$ *for* $i = 1, 2$*. This guarantees that the interpolated triangulations* $\mathscr{K}_1^*(t)$ *keep shape-regular and quasi-uniform for* $t \in [0,T]$*, as well as that the boundary* $\partial\Omega(t)$ *and the interface* $\Gamma(t)$ *are sufficient smooth. Then the following error estimates hold for all* $n \geqslant 1$ *if* $\tau, h$ *are sufficiently small and* $\tau = o(h^{\frac{d}{2}})$

$$\begin{aligned} &\|v_h^n \circ X_{1,h}^n - v_h^{*,n} \circ X_{1,h}^{*,n}\|_{H^1(\Omega_{1,h}^0)} + \|p_h^n \circ X_{1,h}^n - p_h^{*,n} \circ X_{1,h}^{*,n}\|_{L^2(\Omega_{1,h}^0)} + \|X_{1,h}^n - X_{1,h}^{*,n}\|_{H^1(\Omega_{1,h}^0)} \\ &+ \|w_h^n \circ X_{1,h}^n - w_h^{*,n} \circ X_{1,h}^{*,n}\|_{H^1(\Omega_{1,h}^0)} + \|u_h^n - u_h^{*,n}\|_{H^1(\Omega_{2,h})} \leqslant C(\tau + h^k). \end{aligned} \tag{6.45}$$

*Proof.* Let $\Omega_{1,h}^{n,\theta}$ denote the intermediate domain generated by the vector $\mathbf{x}_1^{*,n} + \theta\mathbf{e}_{\mathbf{x}_1}^n$ for $0 \leqslant \theta \leqslant 1$. The finite element functions $\eta_h^\theta$, $e_{x_1}^{n,\theta}$, $v_h^{n,\theta}$, $p_h^{n,\theta}$, $\hat{v}_h^{n,\theta}$, $\hat{w}_h^{n,\theta}$, $v_h^{*,n,\theta}$ are defined on $\Omega_{1,h}^{n,\theta}$ with nodal vectors $\boldsymbol{\eta}$, $\mathbf{e}_{\mathbf{x}_1}^n$, $\mathbf{v}^n$, $\mathbf{p}^n$, $\mathbf{v}^{*,n} + \theta\mathbf{e}_{\mathbf{v}}^n$, $\mathbf{w}_1^{*,n} + \theta\mathbf{e}_{\mathbf{w}_1}^n$, $\mathbf{v}^{*,n}$. The finite element functions $e_{x_1}^{*,n}$, $e_v^{*,n}$, $e_{w_1}^{*,n}$ are defined on $\Omega_{1,h}^{*,n}$ with nodal vectors $\mathbf{e}_{\mathbf{x}_1}^n$, $\mathbf{e}_{\mathbf{v}}^n$ and $\mathbf{e}_{\mathbf{w}_1}^n$.

We will utilize mathematical induction to establish the results. Suppose that the following inequalities are satisfied for $0 \leqslant n \leqslant m-1$:

$$\|\nabla e_{x_1}^{*,n}\|_{L^\infty(\Omega_{1,h}^{*,n})} \leqslant h^{-\frac{d}{4}}(\tau^{\frac{1}{2}} + h^{\frac{k}{2}}), \tag{6.46a}$$

$$\|e_v^{*,n}\|_{W^{1,\infty}(\Omega_{1,h}^{*,n})} + \|e_{w_1}^{*,n}\|_{W^{1,\infty}(\Omega_{1,h}^{*,n})} \leqslant 1. \tag{6.46b}$$

It is evident that $m \geqslant 1$ since $e_{x_1}^{*,0} = 0$, $e_v^{*,0} = 0$, and $e_{w_1}^{*,0}$ represents the error of solving a single Laplace equation. Our approach is to establish the validity of the error estimates (6.45) for $0 \leqslant n \leqslant m$ and subsequently demonstrate that the induction assumptions (6.46) also hold for $n = m$.

Let $\mathbf{y} = \mathbf{x}_1^n$ and $\mathbf{z} = \mathbf{x}_1^{*,n}$ in Lemma 6.4, where the domain $\Omega_{1,h}^{n,\theta}$ is the intermediate domain defined by the vector $\hat{\mathbf{x}}_1^{n,\theta} = \mathbf{x}_1^{*,n} + \theta \mathbf{e}_{\mathbf{x}_1}^n$. When $\tau$ and $h$ are sufficiently small, the conditions of Lemma 6.4 are satisfied under the induction assumption (6.46). Therefore, for $0 \leqslant n \leqslant m-1$, we establish the following equivalence relation using Lemma 6.3 and Lemma 6.4:

$$\begin{aligned} &\text{the norms } \|\cdot\|_{\mathbf{M}_f(\hat{\mathbf{x}}_1^{n,\theta})} \text{ are } h\text{-uniformly equivalent for } 0 \leqslant \theta \leqslant 1, \\ &\text{and so are the norms } \|\cdot\|_{\mathbf{A}_f(\hat{\mathbf{x}}_1^{n,\theta})} \text{ and } \|\cdot\|_{\mathbf{M}_f(\hat{\mathbf{x}}_1^{n,\theta})} + \|\cdot\|_{\mathbf{C}_f(\hat{\mathbf{x}}_1^{n,\theta})}. \end{aligned} \tag{6.47}$$

Let $\mathbf{y} = \mathbf{x}_1^n$ and $\mathbf{z} = \mathbf{x}_1^{n-1}$ in Lemma 6.4. From equation (6.37b), we have $\mathbf{x}_1^n - \mathbf{x}_1^{n-1} = \tau \mathbf{w}_1^{n-1}$. Under the induction assumption (6.46), the $W^{1,\infty}$ boundedness of $w_h^{*,n-1}$ ensures that $\|\nabla w_h^{n-1}\|_{L^\infty(\Omega_{1,h}^{n-1})}$ is bounded. Consequently, when $\tau$ is sufficiently small, the conditions of Lemma 6.4 are satisfied. By applying Lemma 6.2–6.4, for $0 \leqslant n \leqslant m$, we establish that

$$\left| \|\cdot\|_{\mathbf{M}_f(\mathbf{x}_1^n)}^2 - \|\cdot\|_{\mathbf{M}_f(\mathbf{x}_1^{n-1})}^2 \right| \leqslant C\tau \|\cdot\|_{\mathbf{M}_f(\mathbf{x}_1^{n-1})}^2, \tag{6.48a}$$

$$\left| \|\cdot\|_{\mathbf{A}_f(\mathbf{x}_1^n)}^2 - \|\cdot\|_{\mathbf{A}_f(\mathbf{x}_1^{n-1})}^2 \right| \leqslant C\tau \|\cdot\|_{\mathbf{A}_f(\mathbf{x}_1^{n-1})}^2. \tag{6.48b}$$

By applying discrete Gronwall's inequality, we then obtain that the following norms are equivalent for $0 \leqslant n \leqslant m$

$$\|\cdot\|_{\mathbf{M}_f(\mathbf{x}_1^n)}^2 \sim \|\cdot\|_{\mathbf{M}_f(\mathbf{x}_1^0)}^2, \qquad \|\cdot\|_{\mathbf{A}_f(\mathbf{x}_1^n)}^2 \sim \|\cdot\|_{\mathbf{A}_f(\mathbf{x}_1^0)}^2. \tag{6.49}$$

Note that (6.48) and (6.49) are still valid when $\mathbf{x}_1^n$ and $\mathbf{x}_1^{n-1}$ are replaced by $\mathbf{x}_1^{*,n}$ and $\mathbf{x}_1^{*,n-1}$.

(A) *Estimates for* $\mathbf{e}_{\mathbf{x}_1}^n$: Multiplying (6.40b) by $\mathbf{e}_{\mathbf{x}_1}^n \cdot \mathbf{M}_f(\mathbf{x}_1^n)$ results in

$$\frac{1}{2\tau}\Big((\mathbf{e}_{\mathbf{x}_1}^n + \mathbf{e}_{\mathbf{x}_1}^{n-1}) + (\mathbf{e}_{\mathbf{x}_1}^n - \mathbf{e}_{\mathbf{x}_1}^{n-1})\Big) \cdot \mathbf{M}_f(\mathbf{x}_1^n)(\mathbf{e}_{\mathbf{x}_1}^n - \mathbf{e}_{\mathbf{x}_1}^{n-1}) = \mathbf{e}_{\mathbf{x}_1}^n \cdot \mathbf{M}_f(\mathbf{x}_1^n)(\mathbf{e}_{\mathbf{w}_1}^{n-1} - \boldsymbol{\mathcal{F}}_1^{n-1}),$$

which implies the following inequality by using Cauchy-Schwarz inequality

$$\frac{1}{2\tau}\Big(\|\mathbf{e}_{\mathbf{x}_1}^n\|_{\mathbf{M}_f(\mathbf{x}_1^n)}^2 - \|\mathbf{e}_{\mathbf{x}_1}^{n-1}\|_{\mathbf{M}_f(\mathbf{x}_1^n)}^2 + \|\delta \mathbf{e}_{\mathbf{x}_1}^n\|_{\mathbf{M}_f(\mathbf{x}_1^n)}^2\Big) \leqslant \|\mathbf{e}_{\mathbf{x}_1}^n\|_{\mathbf{M}_f(\mathbf{x}_1^n)}\Big(\|\mathbf{e}_{\mathbf{w}_1}^{n-1}\|_{\mathbf{M}_f(\mathbf{x}_1^n)} + \|\boldsymbol{\mathcal{F}}_1^{n-1}\|_{\mathbf{M}_f(\mathbf{x}_1^n)}\Big).$$

On the left-hand side, the term $\|\mathbf{e}_{\mathbf{x}_1}^{n-1}\|_{\mathbf{M}_f(\mathbf{x}_1^n)}$ can be estimated by using (6.48a). On the right-hand side, by using (6.49) and applying Hölder's inequality, we can further establish

$$\begin{aligned}&\frac{1}{2\tau}\Big(\|\mathbf{e}_{\mathbf{x}_1}^{n}\|^2_{\mathbf{M}_f(\mathbf{x}_1^n)}-\|\mathbf{e}_{\mathbf{x}_1}^{n-1}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{n-1})}+\|\delta\mathbf{e}_{\mathbf{x}_1}^{n}\|^2_{\mathbf{M}_f(\mathbf{x}_1^n)}\Big)\\ \leqslant& C\|\mathbf{e}_{\mathbf{x}_1}^{n-1}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{n-1})}+C\|\mathbf{e}_{\mathbf{w}_1}^{n-1}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{n-1})}+\|\mathbf{e}_{\mathbf{x}_1}^{n}\|^2_{\mathbf{M}_f(\mathbf{x}_1^n)}+C\|\boldsymbol{\mathcal{F}}_1^{n-1}\|^2_{\mathbf{M}_f(\mathbf{x}_1^0)}.\end{aligned}\tag{6.50}$$

In a similar manner, by multiplying (6.40b) by $\mathbf{e}_{\mathbf{x}_1}^n\cdot\mathbf{A}_f(\mathbf{x}_1^n)$, we obtain the following inequality:

$$\begin{aligned}&\frac{1}{2\tau}\Big(\|\mathbf{e}_{\mathbf{x}_1}^{n}\|^2_{\mathbf{A}_f(\mathbf{x}_1^n)}-\|\mathbf{e}_{\mathbf{x}_1}^{n-1}\|^2_{\mathbf{A}_f(\mathbf{x}_1^{n-1})}+\|\delta\mathbf{e}_{\mathbf{x}_1}^{n}\|^2_{\mathbf{A}_f(\mathbf{x}_1^n)}\Big)\\ \leqslant& C\|\mathbf{e}_{\mathbf{x}_1}^{n-1}\|^2_{\mathbf{A}_f(\mathbf{x}_1^{n-1})}+C\|\mathbf{e}_{\mathbf{w}_1}^{n-1}\|^2_{\mathbf{A}_f(\mathbf{x}_1^{n-1})}+\|\mathbf{e}_{\mathbf{x}_1}^{n}\|^2_{\mathbf{A}_f(\mathbf{x}_1^n)}+C\|\boldsymbol{\mathcal{F}}_1^{n-1}\|^2_{\mathbf{A}_f(\mathbf{x}_1^0)}.\end{aligned}\tag{6.51}$$

(B) *Estimates for* $\mathbf{e}_{\mathbf{w}_1}^{n-1}$: Since (6.40a) is a single Poisson equation with Dirichlet boundary condition and the boundary value of $e_{w_1}^{*,n-1}$ is only a part of $e_v^{*,n-1}$, we can extend the trace of $e_{w_1}^{*,n-1}$ to a function $\varphi_h^*\in S_h^k(\Omega_{1,h}^{*,n-1})$ such that

$$\|\varphi_h^*\|_{H^1(\Omega_{1,h}^{*,n-1})}\leqslant C\|e_v^{*,n-1}\|_{H^1(\Omega_{1,h}^{*,n-1})},\quad \varphi_h^*=e_{w_1}^{*,n-1}\ \text{ on }\ \partial\Omega_{1,h}^{*,n-1}.\tag{6.52}$$

We can construct the function $\varphi_h^*\in S_h^k(\Omega_{1,h}^{*,n-1})$ in the following way. By using the lift map specified in (6.16), the function $e_{w_1}^{*,n-1}\circ(\Phi_{1,h}^*)^{-1}$ belongs to $H^1(\Omega_1(t_{n-1}))$. Then by Trace Theorem (see [1, Lemma 7.41]), there is a function $\varphi\in H^1(\Omega_1(t_{n-1}))$ such that the trace of $\varphi$ matches $e_{w_1}^{*,n-1}\circ(\Phi_{1,h}^*)^{-1}$ on the boundary $\partial\Omega_1(t_{n-1})$. Furthermore, the function $\varphi$ satisfies the following estimate:

$$\|\varphi\|_{H^1(\Omega_1(t_{n-1}))}\leqslant C\|e_{w_1}^{*,n-1}\circ(\Phi_{1,h}^*)^{-1}\|_{H^{\frac{1}{2}}(\partial\Omega_1(t_{n-1}))},\tag{6.53}$$

where $C$ is a constant independent of $h$, $e_{w_1}^{*,n-1}$ and $\varphi$. Note that $|e_{w_1}^{*,n-1}|\leqslant|e_v^{*,n-1}|$ pointwisely on $\partial\Omega_{1,h}^{*,n-1}$ since $e_{w_1}^{*,n-1}=0$ on $\partial\Omega_{1,h}^{*,n-1}\backslash\Gamma_h^{*,n-1}$ and $e_{w_1}^{*,n-1}=e_v^{*,n-1}$ on $\Gamma_h^{*,n-1}$. Hence, we have that

$$\|e_{w_1}^{*,n-1}\circ(\Phi_{1,h}^*)^{-1}\|_{H^{\frac{1}{2}}(\partial(\Omega_1(t_{n-1})))}\leqslant\|e_v^{*,n-1}\circ(\Phi_{1,h}^*)^{-1}\|_{H^{\frac{1}{2}}(\partial(\Omega_1(t_{n-1})))}.\tag{6.54}$$

Combining the inequalities (6.53) and (6.54), and applying Trace Theorem (see [1, Lemma 7.40]) to function $e_v^{*,n-1}\circ(\Phi_{1,h}^*)^{-1}$, we have

$$\|\varphi\|_{H^1(\Omega_1(t_{n-1}))}\leqslant C\|e_v^{*,n-1}\circ(\Phi_{1,h}^*)^{-1}\|_{H^{\frac{1}{2}}(\partial(\Omega_1(t_{n-1})))}\leqslant C\|e_v^{*,n-1}\circ(\Phi_{1,h}^*)^{-1}\|_{H^1((\Omega_1(t_{n-1})))}.\tag{6.55}$$

Let $\varphi_h^* \in S_h^k(\Omega_{1,h}^{*,n-1})$ denote the Scott-Zhang interpolation (see [155]) of $\varphi \circ \Phi_{1,h}^*$ over $\Omega_{1,h}^{*,n-1}$. Due to the $H^1$-stability of the Scott-Zhang interpolation (see [155, Corollary 4.1]), we can conclude:

$$\|\varphi_h^*\|_{H^1(\Omega_{1,h}^{*,n-1})} \leqslant C\|\varphi \circ \Phi_{1,h}^*\|_{H^1(\Omega_{1,h}^{*,n-1})} \leqslant C\|e_v^{*,n-1}\|_{H^1(\Omega_{1,h}^{*,n-1})}, \tag{6.56}$$

where the last inequality follows from (6.55) by making a change of variable and using the $W^{1,\infty}$ boundedness of $\Phi_{1,h}^*$ and $(\Phi_{1,h}^*)^{-1}$ given in (6.12). The Scott-Zhang interpolation preserves boundary values if the boundary values are the trace of some finite element functions. Hence, we have that $\varphi_h^* = \varphi \circ \Phi_{1,h}^* = e_{w_1}^{*,n-1}$ on $\partial\Omega_{1,h}^{*,n-1}$. This proves the result (6.52).

Let us denote the vector that collect the nodal values of $\varphi_h^*$ as $\boldsymbol{\varphi}$. Then we can test (6.40a) with $\mathbf{e}_{\mathbf{w}_1}^{n-1} - \boldsymbol{\varphi}$. Using Lemma 6.2 and Lemma 6.4, and the $W^{1,\infty}$ boundedness of the mesh velocity $w_h^*$, and the consistency error estimate (6.21) we obtain

$$\begin{aligned}
&\|\mathbf{e}_{\mathbf{w}_1}^{n-1}\|_{\mathbf{A}_f(\mathbf{x}_1^{n-1})}^2 \\
=&\boldsymbol{\varphi} \cdot \mathbf{A}_f(\mathbf{x}_1^{n-1})\mathbf{e}_{\mathbf{w}_1}^{n-1} - (\mathbf{e}_{\mathbf{w}_1}^{n-1} - \boldsymbol{\varphi}) \cdot (\mathbf{A}_f(\mathbf{x}_1^{n-1}) - \mathbf{A}_f(\mathbf{x}_1^{*,n-1}))\mathbf{w}_1^{*,n-1} - \boldsymbol{\mathcal{E}}_4^{n-1} \cdot (\mathbf{e}_{\mathbf{w}_1}^{n-1} - \boldsymbol{\varphi}) \\
\leqslant&\|\boldsymbol{\varphi}\|_{\mathbf{A}_f(\mathbf{x}_1^{n-1})}\|\mathbf{e}_{\mathbf{w}_1}^{n-1}\|_{\mathbf{A}_f(\mathbf{x}_1^{n-1})} + C\Big(\|\mathbf{e}_{\mathbf{w}_1}^{n-1}\|_{\mathbf{A}_f(\mathbf{x}_1^{*,n-1})} + \|\boldsymbol{\varphi}\|_{\mathbf{A}_f(\mathbf{x}_1^{*,n-1})}\Big)\|\mathbf{e}_{\mathbf{x}_1}^{n-1}\|_{\mathbf{A}_f(\mathbf{x}_1^{*,n-1})} \\
&+ Ch^k\Big(\|\mathbf{e}_{\mathbf{w}_1}^{n-1}\|_{\mathbf{M}_f(\mathbf{x}_1^{*,n-1})} + \|\mathbf{e}_{\mathbf{w}_1}^{n-1}\|_{\mathbf{A}_f(\mathbf{x}_1^{*,n-1})} + \|\boldsymbol{\varphi}\|_{\mathbf{M}_f(\mathbf{x}_1^{*,n-1})} + \|\boldsymbol{\varphi}\|_{\mathbf{A}_f(\mathbf{x}_1^{*,n-1})}\Big).
\end{aligned}$$

By applying Young's inequality, the term $\|\mathbf{e}_{\mathbf{w}_1}^{n-1}\|_{\mathbf{A}_f(\mathbf{x}_1^{n-1})}^2$ on the right-hand side can be absorbed by the left-hand side. Then by using (6.56) and the equivalence of the norms (6.47), we have

$$\begin{aligned}
&\|\mathbf{e}_{\mathbf{w}_1}^{n-1}\|_{\mathbf{A}_f(\mathbf{x}_1^{n-1})}^2 \\
\leqslant&C\Big(h^{2k} + \|\boldsymbol{\varphi}\|_{\mathbf{M}_f(\mathbf{x}_1^{n-1})}^2 + \|\boldsymbol{\varphi}\|_{\mathbf{A}_f(\mathbf{x}_1^{n-1})}^2 + \|\mathbf{e}_{\mathbf{x}_1}^{n-1}\|_{\mathbf{A}_f(\mathbf{x}_1^{n-1})}^2 + h^k\|\mathbf{e}_{\mathbf{w}_1}^{n-1}\|_{\mathbf{M}_f(\mathbf{x}_1^{n-1})}\Big) \\
\leqslant&C\Big(h^{2k} + \|\mathbf{e}_{\mathbf{v}}^{n-1}\|_{\mathbf{M}_f(\mathbf{x}_1^{n-1})}^2 + \|\mathbf{e}_{\mathbf{v}}^{n-1}\|_{\mathbf{A}_f(\mathbf{x}_1^{n-1})}^2 + \|\mathbf{e}_{\mathbf{x}_1}^{n-1}\|_{\mathbf{A}_f(\mathbf{x}_1^{n-1})}^2 + h^k\|\mathbf{e}_{\mathbf{w}_1}^{n-1}\|_{\mathbf{M}_f(\mathbf{x}_1^{n-1})}\Big).
\end{aligned}$$

Since $\mathbf{e}_{\mathbf{w}_1}^{n-1} = 0$ on $\partial\Omega_{1,h}^{n-1}\backslash\Gamma_h^{n-1}$, by applying Poincaré's inequality and Young's inequality, the term $\|\mathbf{e}_{\mathbf{w}_1}^{n-1}\|_{\mathbf{M}_f(\mathbf{x}_1^{n-1})}$ on the right-hand side can be absorbed by the left-hand side, which gives the following result (6.57).

$$\|\mathbf{e}_{\mathbf{w}_1}^{n-1}\|_{\mathbf{A}_f(\mathbf{x}_1^{n-1})}^2 \leqslant Ch^{2k} + C\|\mathbf{e}_{\mathbf{v}}^{n-1}\|_{\mathbf{M}_f(\mathbf{x}_1^{n-1})}^2 + C\|\mathbf{e}_{\mathbf{v}}^{n-1}\|_{\mathbf{A}_f(\mathbf{x}_1^{n-1})}^2 + C\|\mathbf{e}_{\mathbf{x}_1}^{n-1}\|_{\mathbf{A}_f(\mathbf{x}_1^{n-1})}^2. \tag{6.57}$$

By applying Poincaré's inequality to $\mathbf{e}_{\mathbf{w}_1}^{n-1}$, we can deduce that the inequality (6.57) is also satisfied for $\|\mathbf{e}_{\mathbf{w}_1}^{n-1}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{n-1})}$. By substituting these results into the inequalities (6.50) and (6.51), and using the estimate (6.39a), we obtain that

$$\begin{aligned}
&\frac{1}{2\tau}\Big(\|\mathbf{e}_{\mathbf{x}_1}^{n}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{n})}+\|\mathbf{e}_{\mathbf{x}_1}^{n}\|^2_{\mathbf{A}_f(\mathbf{x}_1^{n})}-\|\mathbf{e}_{\mathbf{x}_1}^{n-1}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{n-1})}-\|\mathbf{e}_{\mathbf{x}_1}^{n-1}\|^2_{\mathbf{A}_f(\mathbf{x}_1^{n-1})}\Big)\\
\leqslant& C\Big(\tau^2+h^{2k}+\|\mathbf{e}_{\mathbf{x}_1}^{n-1}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{n-1})}+\|\mathbf{e}_{\mathbf{x}_1}^{n-1}\|^2_{\mathbf{A}_f(\mathbf{x}_1^{n-1})}+\|\mathbf{e}_{\mathbf{x}_1}^{n}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{n})}+\|\mathbf{e}_{\mathbf{x}_1}^{n}\|^2_{\mathbf{A}_f(\mathbf{x}_1^{n})}\Big)\\
&+C\Big(\|\mathbf{e}_{\mathbf{v}}^{n-1}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{n-1})}+\|\mathbf{e}_{\mathbf{v}}^{n-1}\|^2_{\mathbf{A}_f(\mathbf{x}_1^{n-1})}\Big).
\end{aligned}\tag{6.58}$$

By summing up (6.58) from $n=1$ to $n=\ell$ for $1\leqslant\ell\leqslant m$, we obtain the following inequality when $\tau$ is sufficiently small

$$\begin{aligned}
\|\mathbf{e}_{\mathbf{x}_1}^{\ell}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{\ell})}+\|\mathbf{e}_{\mathbf{x}_1}^{\ell}\|^2_{\mathbf{A}_f(\mathbf{x}_1^{\ell})}\leqslant& C(\tau^2+h^{2k})+C\tau\sum_{n=1}^{\ell-1}\|\mathbf{e}_{\mathbf{x}_1}^{n}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{n})}\\
&+C\tau\sum_{n=1}^{\ell-1}\Big(\|\mathbf{e}_{\mathbf{x}_1}^{n}\|^2_{\mathbf{A}_f(\mathbf{x}_1^{n})}+\|\mathbf{e}_{\mathbf{v}}^{n}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{n})}+\|\mathbf{e}_{\mathbf{v}}^{n}\|^2_{\mathbf{A}_f(\mathbf{x}_1^{n})}\Big).
\end{aligned}\tag{6.59}$$

(C) *Estimates for* $\|\mathbf{e}_{\mathbf{v}}^{n}\|_{\mathbf{M}_f(\mathbf{x}_1^{n})}$ *and* $\|\mathbf{e}_{\mathbf{u}}^{n}\|_{\mathbf{E}_s(\mathbf{x}_2)}$: In this part, we want to obtain the estimates for $\|\mathbf{e}_{\mathbf{v}}^{n}\|_{\mathbf{M}_f(\mathbf{x}_1^{n})}$ and $\|\mathbf{e}_{\mathbf{u}}^{n}\|_{\mathbf{E}_s(\mathbf{x}_2)}$ as a preliminarty result, which will be used to estimate the $H^1$-norm of $\mathbf{e}_{\mathbf{v}}^{n}$ later.

By choosing $\boldsymbol{\eta}=\mathbf{e}_{\mathbf{v}}^{n}$ and $\boldsymbol{\xi}=\dot{\mathbf{e}}_{\mathbf{u}}^{n}$ in (6.40c), since $\partial_\theta^\bullet\hat{v}_h^{n,\theta}=e_v^{n,\theta}$ and $\partial_\theta^\bullet\hat{w}_h^{n,\theta}=e_{w_1}^{n,\theta}$, we have:

$$\begin{aligned}
&\frac{1}{2\tau}\Big(\|\mathbf{e}_{\mathbf{v}}^{n}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{n})}-\|\mathbf{e}_{\mathbf{v}}^{n-1}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{n})}+\|\delta\mathbf{e}_{\mathbf{v}}^{n}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{n})}\Big)+\|\mathbf{e}_{\mathbf{v}}^{n}\|^2_{\mathbf{C}_f(\mathbf{x}_1^{n})}-\mathbf{e}_{\mathbf{v}}^{n}\cdot\mathbf{D}_f(\mathbf{x}_1^{n})^\top\mathbf{e}_{\mathbf{p}}^{n}\\
&+\frac{1}{2\tau}\Big(\|\dot{\mathbf{e}}_{\mathbf{u}}^{n}\|^2_{\mathbf{M}_s(\mathbf{x}_2)}-\|\dot{\mathbf{e}}_{\mathbf{u}}^{n-1}\|^2_{\mathbf{M}_s(\mathbf{x}_2)}+\|\delta\dot{\mathbf{e}}_{\mathbf{u}}^{n}\|^2_{\mathbf{M}_s(\mathbf{x}_2)}+\|\mathbf{e}_{\mathbf{u}}^{n}\|^2_{\mathbf{E}_s(\mathbf{x}_2)}-\|\mathbf{e}_{\mathbf{u}}^{n-1}\|^2_{\mathbf{E}_s(\mathbf{x}_2)}+\|\delta\mathbf{e}_{\mathbf{u}}^{n}\|^2_{\mathbf{E}_s(\mathbf{x}_2)}\Big)\\
=&-\rho_1\int_0^1\int_{\Omega_{1,h}^{n,\theta}}e_v^{n,\theta}\frac{v_h^{*,n,\theta}-v_h^{*,n-1,\theta}}{\tau}\nabla\cdot e_{x_1}^{n,\theta}+(\hat{v}_h^{n-1,\theta}-\hat{w}_h^{n-1,\theta})\cdot\nabla\hat{v}_h^{n,\theta}e_v^{n,\theta}\nabla\cdot e_{x_1}^{n,\theta}\mathrm{d}x\mathrm{d}\theta\\
&-\rho_1\int_0^1\int_{\Omega_{1,h}^{n,\theta}}(e_v^{n-1,\theta}-e_{w_1}^{n-1,\theta})\cdot\nabla\hat{v}_h^{n,\theta}e_v^{n,\theta}+(\hat{v}_h^{n-1,\theta}-\hat{w}_h^{n-1,\theta})\cdot(\nabla e_v^{n,\theta}-\nabla\hat{v}_h^{n,\theta}\nabla e_{x_1}^{n,\theta})e_v^{n,\theta}\mathrm{d}x\mathrm{d}\theta\\
&+2\mu_1\int_0^1\int_{\Omega_{1,h}^{n,\theta}}(\nabla e_v^{n,\theta}\nabla e_{x_1}^{n,\theta}):\mathbb{D}(v_h^{*,n,\theta})+\mathbb{D}(e_v^{n,\theta}):(\nabla v_h^{*,n,\theta}\nabla e_{x_1}^{n,\theta})\mathrm{d}x\mathrm{d}\theta\\
&-2\mu_1\int_0^1\int_{\Omega_{1,h}^{n,\theta}}\mathbb{D}(e_v^{n,\theta}):\mathbb{D}(v_h^{*,n,\theta})\nabla\cdot e_{x_1}^{n,\theta}+p_h^{*,n,\theta}\nabla\cdot e_v^{n,\theta}\nabla\cdot e_{x_1}^{n,\theta}-p_h^{*,n,\theta}\nabla e_v^{n,\theta}:(\nabla e_{x_1}^{n,\theta})^\top\mathrm{d}x\mathrm{d}\theta\\
&+\rho_1\int_0^1\int_{\Omega_{1,h}^{n,\theta}}e_v^{n,\theta}\cdot(\nabla f_1(\cdot,t_n)e_{x_1}^{n,\theta})+e_v^{n,\theta}\cdot f_1(\cdot,t_n)\nabla\cdot e_{x_1}^{n,\theta}\mathrm{d}x\mathrm{d}\theta+\mathbf{e}_{\mathbf{v}}^{n}\cdot(\boldsymbol{\vartheta}_1-\boldsymbol{\mathcal{E}}_1^n-\boldsymbol{\mathcal{F}}_2^n)
\end{aligned}$$

$$+\dot{\mathbf{e}}_{\mathbf{u}}^n\cdot(\boldsymbol{\vartheta}_2^n-\boldsymbol{\mathcal{E}}_2^n-\boldsymbol{\mathcal{F}}_3^n)-\boldsymbol{\mathcal{F}}_4^n\cdot\mathbf{E}_s(\mathbf{x}_2)\mathbf{e}_{\mathbf{u}}^n.$$

When $1\leqslant n\leqslant m-1$, by using the equivalence (6.47), (6.49), the estimate (6.48), the consistency error estimates (6.21), (6.32), (6.39a), (6.39b) and Hölder's inequality, we can deduce that

$$\begin{aligned}
&\frac{1}{2\tau}\Big(\|\mathbf{e}_{\mathbf{v}}^n\|^2_{\mathbf{M}_f(\mathbf{x}_1^n)}-\|\mathbf{e}_{\mathbf{v}}^{n-1}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{n-1})}\Big)+\|\mathbf{e}_{\mathbf{v}}^n\|^2_{\mathbf{C}_f(\mathbf{x}_1^n)}+\frac{1}{2\tau}\Big(\|\dot{\mathbf{e}}_{\mathbf{u}}^n\|^2_{\mathbf{M}_s(\mathbf{x}_2)}-\|\dot{\mathbf{e}}_{\mathbf{u}}^{n-1}\|^2_{\mathbf{M}_s(\mathbf{x}_2)}\Big)\\
&+\frac{1}{2\tau}\Big(\|\mathbf{e}_{\mathbf{u}}^n\|^2_{\mathbf{E}_s(\mathbf{x}_2)}-\|\mathbf{e}_{\mathbf{u}}^{n-1}\|^2_{\mathbf{E}_s(\mathbf{x}_2)}\Big)\\
\leqslant&C\|\mathbf{e}_{\mathbf{v}}^{n-1}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{n-1})}+Ch^k\Big(\|\mathbf{e}_{\mathbf{v}}^n\|_{\mathbf{M}_f(\mathbf{x}_1^n)}+\|\mathbf{e}_{\mathbf{v}}^n\|_{\mathbf{A}_f(\mathbf{x}_1^n)}+\|\dot{\mathbf{e}}_{\mathbf{u}}^n\|_{\mathbf{M}_s(\mathbf{x}_2)}+\|\dot{\mathbf{e}}_{\mathbf{u}}^n\|_{\mathbf{E}_s(\mathbf{x}_2)}\Big)\\
&+C\tau\Big(\|\mathbf{e}_{\mathbf{v}}^n\|_{\mathbf{M}_f(\mathbf{x}_1^n)}+\|\dot{\mathbf{e}}_{\mathbf{u}}^n\|_{\mathbf{M}_s(\mathbf{x}_2)}+\|\mathbf{e}_{\mathbf{u}}^n\|_{\mathbf{E}_s(\mathbf{x}_2)}\Big)\\
&+\frac{C}{\tau}\|\mathbf{e}_{\mathbf{v}}^n\|_{\mathbf{M}_f(\mathbf{x}_1^n)}\|\mathbf{e}_{\mathbf{x}_1}^n\|_{\mathbf{A}_f(\mathbf{x}_1^n)}\|v_h^{*,n}-v_h^{*,n-1}\|_{L^\infty(\Omega_{1,h}^{*,n})}+\mathbf{e}_{\mathbf{v}}^n\cdot\mathbf{D}_f(\mathbf{x}_1^n)^\top\mathbf{e}_{\mathbf{p}}^n\\
&+C\|\mathbf{e}_{\mathbf{v}}^n\|_{\mathbf{M}_f(\mathbf{x}_1^n)}\|\mathbf{e}_{\mathbf{x}_1}^n\|_{\mathbf{A}_f(\mathbf{x}_1^n)}\|\nabla\hat{v}_h^{n,0}\|_{L^\infty(\Omega_{1,h}^{*,n})}\Big(\|\hat{v}_h^{n-1,0}\|_{L^\infty(\Omega_{1,h}^{*,n})}+\|\hat{w}_h^{n-1,0}\|_{L^\infty(\Omega_{1,h}^{*,n})}\Big)\\
&+C\Big(\|\mathbf{e}_{\mathbf{v}}^{n-1}\|_{\mathbf{M}_f(\mathbf{x}_1^n)}+\|\mathbf{e}_{\mathbf{w}_1}^{n-1}\|_{\mathbf{M}_f(\mathbf{x}_1^n)}\Big)\|\mathbf{e}_{\mathbf{v}}^n\|_{\mathbf{M}_f(\mathbf{x}_1^n)}\|\nabla\hat{v}_h^{n,0}\|_{L^\infty(\Omega_{1,h}^{*,n})}\\
&+C\|\mathbf{e}_{\mathbf{v}}^n\|_{\mathbf{M}_f(\mathbf{x}_1^n)}\|\mathbf{e}_{\mathbf{v}}^n\|_{\mathbf{A}_f(\mathbf{x}_1^n)}\Big(\|\hat{v}_h^{n-1,0}\|_{L^\infty(\Omega_{1,h}^{*,n})}+\|\hat{w}_h^{n-1,0}\|_{L^\infty(\Omega_{1,h}^{*,n})}\Big)\\
&+C\|\mathbf{e}_{\mathbf{v}}^n\|_{\mathbf{A}_f(\mathbf{x}_1^n)}\|\mathbf{e}_{\mathbf{x}_1}^n\|_{\mathbf{A}_f(\mathbf{x}_1^n)}\Big(\|\nabla v_h^{*,n}\|_{L^\infty(\Omega_{1,h}^{*,n})}+\|p_h^{*,n}\|_{L^\infty(\Omega_{1,h}^{*,n})}\Big)\\
&+C\|\mathbf{e}_{\mathbf{v}}^n\|_{\mathbf{M}_f(\mathbf{x}_1^n)}\Big(\|\mathbf{e}_{\mathbf{x}_1}^n\|_{\mathbf{M}_f(\mathbf{x}_1^n)}+\|\mathbf{e}_{\mathbf{x}_1}^n\|_{\mathbf{A}_f(\mathbf{x}_1^n)}\Big)\|f_1(\cdot,t_n)\|_{W^{1,\infty}(\Omega_{1,h}^{*,n})}.
\end{aligned}$$

By Taylor's expansion, we can derive that $\frac{1}{\tau}\|v_h^{*,n}-v_h^{*,n-1}\|_{L^\infty(\Omega_{1,h}^{*,n})}$ is bounded. By using the induction assumption (6.46) and the assumption that the exact solutions are sufficiently smooth, we can infer that all the $L^\infty$ and $W^{1,\infty}$ norms in the above inequality are bounded. To handle the term $\mathbf{e}_{\mathbf{v}}^n\cdot\mathbf{D}_f(\mathbf{x}_1^n)^\top\mathbf{e}_{\mathbf{p}}^n$, we can test equation (6.40d) with $\mathbf{e}_{\mathbf{p}}^n$. By applying Young's inequality, Korn's inequality and inequality (6.48), we arrive at the following inequality when $h$ is sufficiently small:

$$\begin{aligned}
&\frac{1}{2\tau}\Big(\|\mathbf{e}_{\mathbf{v}}^n\|^2_{\mathbf{M}_f(\mathbf{x}_1^n)}-\|\mathbf{e}_{\mathbf{v}}^{n-1}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{n-1})}\Big)+\|\mathbf{e}_{\mathbf{v}}^n\|^2_{\mathbf{C}_f(\mathbf{x}_1^n)}+\frac{1}{2\tau}\Big(\|\dot{\mathbf{e}}_{\mathbf{u}}^n\|^2_{\mathbf{M}_s(\mathbf{x}_2)}-\|\dot{\mathbf{e}}_{\mathbf{u}}^{n-1}\|^2_{\mathbf{M}_s(\mathbf{x}_2)}\Big)\\
&+\frac{1}{2\tau}\Big(\|\mathbf{e}_{\mathbf{u}}^n\|^2_{\mathbf{E}_s(\mathbf{x}_2)}-\|\mathbf{e}_{\mathbf{u}}^{n-1}\|^2_{\mathbf{E}_s(\mathbf{x}_2)}\Big)\\
\leqslant&C_1(\tau^2+h^{2k})+C_1\Big(\|\mathbf{e}_{\mathbf{x}_1}^n\|^2_{\mathbf{A}_f(\mathbf{x}_1^n)}+\|\mathbf{e}_{\mathbf{v}}^{n-1}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{n-1})}+\|\mathbf{e}_{\mathbf{v}}^n\|^2_{\mathbf{M}_f(\mathbf{x}_1^n)}+\|\mathbf{e}_{\mathbf{u}}^n\|^2_{\mathbf{E}_s(\mathbf{x}_2)}\Big)\\
&+\varepsilon_1\Big(\|\mathbf{e}_{\mathbf{p}}^n\|^2_{\mathbf{M}_f(\mathbf{x}_1^n)}+\|\mathbf{e}_{\mathbf{x}_1}^n\|^2_{\mathbf{M}_f(\mathbf{x}_1^n)}+\|\mathbf{e}_{\mathbf{w}_1}^{n-1}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{n-1})}+\|\dot{\mathbf{e}}_{\mathbf{u}}^n\|^2_{\mathbf{M}_s(\mathbf{x}_2)}+\|\dot{\mathbf{e}}_{\mathbf{u}}^n\|^2_{\mathbf{E}_s(\mathbf{x}_2)}\Big),\quad(6.60)
\end{aligned}$$

where $\varepsilon_1$ is a small positive real number, and $C_1$ is a constant that depends on $1/\varepsilon_1$.

(D) *Estimates for* $\|\mathbf{e}_{\mathbf{p}}^n\|_{\mathbf{M}_f(\mathbf{x}_1^n)}$: Starting from the expression of $\boldsymbol{\eta}\cdot\mathbf{D}_f(\mathbf{x}_1^n)^\top\mathbf{e}_{\mathbf{p}}^n$ obtained in (6.40c), by using inf-sup condition (6.35) and incorporating the consistency error estimates (6.21), (6.32), (6.39a), and (6.39b), we can derive the estimate for $\|\mathbf{e}_{\mathbf{p}}^n\|_{\mathbf{M}_f(\mathbf{x}_1^n)}$ as follows:

$$\begin{aligned}\|\mathbf{e}_{\mathbf{p}}^n\|_{\mathbf{M}_f(\mathbf{x}_1^n)}\leqslant& C\Big(\tau+h^k+\frac{1}{\tau}\|\delta\mathbf{e}_{\mathbf{v}}^n\|_{\mathbf{M}_f(\mathbf{x}_1^n)}+\|\mathbf{e}_{\mathbf{v}}^{n-1}\|_{\mathbf{M}_f(\mathbf{x}_1^{n-1})}+\|\mathbf{e}_{\mathbf{v}}^n\|_{\mathbf{C}_f(\mathbf{x}_1^n)}\Big)\\&+C\Big(\|\mathbf{e}_{\mathbf{w}_1}^{n-1}\|_{\mathbf{M}_f(\mathbf{x}_1^{n-1})}+\|\mathbf{e}_{\mathbf{x}_1}^n\|_{\mathbf{M}_f(\mathbf{x}_1^n)}+\|\mathbf{e}_{\mathbf{x}_1}^n\|_{\mathbf{A}_f(\mathbf{x}_1^n)}+\|\mathbf{e}_{\mathbf{u}}^n\|_{\mathbf{E}_s(\mathbf{x}_2)}+\frac{1}{\tau}\|\delta\dot{\mathbf{e}}_{\mathbf{u}}^n\|_{\mathbf{M}_s(\mathbf{x}_2)}\Big).\end{aligned}\tag{6.61}$$

Substituting the estimates (6.57) and (6.61) into (6.60), we obtain:

$$\begin{aligned}&\frac{1}{2\tau}\Big(\|\mathbf{e}_{\mathbf{v}}^n\|^2_{\mathbf{M}_f(\mathbf{x}_1^n)}-\|\mathbf{e}_{\mathbf{v}}^{n-1}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{n-1})}\Big)+\frac{1}{2}\|\mathbf{e}_{\mathbf{v}}^n\|^2_{\mathbf{C}_f(\mathbf{x}_1^n)}\\&+\frac{1}{2\tau}\Big(\|\dot{\mathbf{e}}_{\mathbf{u}}^n\|^2_{\mathbf{M}_s(\mathbf{x}_2)}-\|\dot{\mathbf{e}}_{\mathbf{u}}^{n-1}\|^2_{\mathbf{M}_s(\mathbf{x}_2)}\Big)+\frac{1}{2\tau}\Big(\|\mathbf{e}_{\mathbf{u}}^n\|^2_{\mathbf{E}_s(\mathbf{x}_2)}-\|\mathbf{e}_{\mathbf{u}}^{n-1}\|^2_{\mathbf{E}_s(\mathbf{x}_2)}\Big)\\\leqslant&C_1\Big(\tau^2+h^{2k}+\|\mathbf{e}_{\mathbf{v}}^{n-1}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{n-1})}+\|\mathbf{e}_{\mathbf{v}}^n\|^2_{\mathbf{M}_f(\mathbf{x}_1^n)}+\|\mathbf{e}_{\mathbf{x}_1}^n\|^2_{\mathbf{A}_f(\mathbf{x}_1^n)}+\|\mathbf{e}_{\mathbf{u}}^n\|^2_{\mathbf{E}_s(\mathbf{x}_2)}\Big)\\&+C\varepsilon_1\Big(\frac{1}{\tau^2}\|\delta\mathbf{e}_{\mathbf{v}}^n\|^2_{\mathbf{M}_f(\mathbf{x}_1^n)}+\|\mathbf{e}_{\mathbf{v}}^{n-1}\|^2_{\mathbf{C}_f(\mathbf{x}_1^{n-1})}+\|\dot{\mathbf{e}}_{\mathbf{u}}^n\|^2_{\mathbf{M}_s(\mathbf{x}_2)}+\|\dot{\mathbf{e}}_{\mathbf{u}}^n\|^2_{\mathbf{E}_s(\mathbf{x}_2)}\Big)\\&+C\varepsilon_1\Big(\|\mathbf{e}_{\mathbf{x}_1}^n\|^2_{\mathbf{M}_f(\mathbf{x}_1^n)}+\|\mathbf{e}_{\mathbf{x}_1}^{n-1}\|^2_{\mathbf{A}_f(\mathbf{x}_1^{n-1})}+\frac{1}{\tau^2}\|\delta\dot{\mathbf{e}}_{\mathbf{u}}^n\|^2_{\mathbf{M}_s(\mathbf{x}_2)}\Big),\end{aligned}\tag{6.62}$$

where $C$ is a constant independent of $\varepsilon_1$. By summing up the inequality (6.62) from $n=1$ to $n=\ell-1$ for $2\leqslant\ell\leqslant m$, we obtain:

$$\begin{aligned}&\frac{1}{2}\|\mathbf{e}_{\mathbf{v}}^{\ell-1}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{\ell-1})}+\frac{1}{4}\sum_{n=1}^{\ell-1}\tau\|\mathbf{e}_{\mathbf{v}}^n\|^2_{\mathbf{C}_f(\mathbf{x}_1^n)}+\frac{1}{2}\|\dot{\mathbf{e}}_{\mathbf{u}}^{\ell-1}\|^2_{\mathbf{M}_s(\mathbf{x}_2)}+\frac{1}{2}\|\mathbf{e}_{\mathbf{u}}^{\ell-1}\|^2_{\mathbf{E}_s(\mathbf{x}_2)}\\\leqslant&C_1(\tau^2+h^{2k})+C_1\sum_{n=1}^{\ell-1}\tau\Big(\|\mathbf{e}_{\mathbf{v}}^n\|^2_{\mathbf{M}_f(\mathbf{x}_1^n)}+\|\mathbf{e}_{\mathbf{x}_1}^n\|^2_{\mathbf{A}_f(\mathbf{x}_1^n)}+\|\mathbf{e}_{\mathbf{u}}^n\|^2_{\mathbf{E}_s(\mathbf{x}_2)}+\varepsilon_1\|\mathbf{e}_{\mathbf{x}_1}^n\|^2_{\mathbf{M}_f(\mathbf{x}_1^n)}\Big)\\&+C\varepsilon_1\sum_{n=1}^{\ell-1}\tau\Big(\frac{1}{\tau^2}\|\delta\mathbf{e}_{\mathbf{v}}^n\|^2_{\mathbf{M}_f(\mathbf{x}_1^n)}+\|\dot{\mathbf{e}}_{\mathbf{u}}^n\|^2_{\mathbf{M}_s(\mathbf{x}_2)}+\|\dot{\mathbf{e}}_{\mathbf{u}}^n\|^2_{\mathbf{E}_s(\mathbf{x}_2)}+\frac{1}{\tau^2}\|\delta\dot{\mathbf{e}}_{\mathbf{u}}^n\|^2_{\mathbf{M}_s(\mathbf{x}_2)}\Big).\end{aligned}\tag{6.63}$$

(E) *Estimates for* $\|\mathbf{e}_{\mathbf{v}}^n\|_{\mathbf{C}_f(\mathbf{x}_1^n)}$: The error equation (6.40c) is a parabolic system about the variable $\mathbf{e}_{\mathbf{v}}^n$. To derive the $H^1$ estimate for $\mathbf{e}_{\mathbf{v}}^n$, we can test (6.40c) with $\boldsymbol{\eta}=\delta\mathbf{e}_{\mathbf{v}}^n$ and $\boldsymbol{\xi}=\delta\dot{\mathbf{e}}_{\mathbf{u}}^n$ for $n\geqslant 1$, and obtain:

$$\begin{aligned}&\frac{1}{\tau}\|\delta\mathbf{e}_{\mathbf{v}}^n\|^2_{\mathbf{M}_f(\mathbf{x}_1^n)}+\frac{1}{2}\Big(\|\mathbf{e}_{\mathbf{v}}^n\|^2_{\mathbf{C}_f(\mathbf{x}_1^n)}-\|\mathbf{e}_{\mathbf{v}}^{n-1}\|^2_{\mathbf{C}_f(\mathbf{x}_1^{n-1})}+\|\delta\mathbf{e}_{\mathbf{v}}^n\|^2_{\mathbf{C}_f(\mathbf{x}_1^n)}\Big)\\&+\frac{1}{\tau}\|\delta\dot{\mathbf{e}}_{\mathbf{u}}^n\|^2_{\mathbf{M}_s(\mathbf{x}_2)}+(\delta\dot{\mathbf{e}}_{\mathbf{u}}^n)\cdot\mathbf{E}_s(\mathbf{x}_2)\mathbf{e}_{\mathbf{u}}^n-(\delta\mathbf{e}_{\mathbf{v}}^n)\cdot\mathbf{D}_f(\mathbf{x}_1^n)^\top\mathbf{e}_{\mathbf{p}}^n\end{aligned}$$

$$
\begin{aligned}
=&-\frac{1}{\tau}(\delta\mathbf{e}_{\mathbf{v}}^n)\cdot\Big(\mathbf{M}_f(\mathbf{x}_1^n)-\mathbf{M}_f(\mathbf{x}_1^{*,n})\Big)(\delta\mathbf{v}^{*,n})-(\delta\mathbf{e}_{\mathbf{v}}^n)\cdot\mathbf{B}_f(\mathbf{x}_1^n,\mathbf{v}^{n-1}-\mathbf{w}_1^{n-1})\mathbf{v}^n\\
&+(\delta\mathbf{e}_{\mathbf{v}}^n)\cdot\mathbf{B}_f(\mathbf{x}_1^{*,n},\mathbf{v}^{*,n-1}-\mathbf{w}_1^{*,n-1})\mathbf{v}^{*,n}-(\delta\mathbf{e}_{\mathbf{v}}^n)\cdot\Big(\mathbf{C}_f(\mathbf{x}_1^n)-\mathbf{C}_f(\mathbf{x}_1^{*,n})\Big)\mathbf{v}^{*,n}\\
&+(\delta\mathbf{e}_{\mathbf{v}}^n)\cdot\Big(\mathbf{D}_f(\mathbf{x}_1^n)^\top-\mathbf{D}_f(\mathbf{x}_1^{*,n})^\top\Big)\mathbf{p}^{*,n}+(\delta\mathbf{e}_{\mathbf{v}}^n)\cdot(\mathbf{f}_1^n-\mathbf{f}_1^{*,n})\\
&+(\delta\mathbf{e}_{\mathbf{v}}^n)\cdot(\boldsymbol{\vartheta}_1-\boldsymbol{\mathcal{E}}_1^n-\boldsymbol{\mathcal{F}}_2^n)+(\delta\dot{\mathbf{e}}_{\mathbf{u}}^n)\cdot(\boldsymbol{\vartheta}_2-\boldsymbol{\mathcal{E}}_2^n-\boldsymbol{\mathcal{F}}_3^n).
\end{aligned}\tag{6.64}
$$

For the estimation of the nonlinear terms $-(\delta\mathbf{e}_{\mathbf{v}}^n)\cdot\mathbf{B}_f(\mathbf{x}_1^n,\mathbf{v}^{n-1}-\mathbf{w}_1^{n-1})\mathbf{v}^n+(\delta\mathbf{e}_{\mathbf{v}}^n)\cdot\mathbf{B}_f(\mathbf{x}_1^{*,n},\mathbf{v}^{*,n-1}-\mathbf{w}_1^{*,n-1})\mathbf{v}^{*,n}$, we only encounter the $L^2$-norm of $\delta\mathbf{e}_{\mathbf{v}}^n$, i.e., $\|\delta\mathbf{e}_{\mathbf{v}}^n\|_{\mathbf{M}_f(\mathbf{x}_1^n)}$. Following a similar proof in the derivation of (6.60), $\|\delta\mathbf{e}_{\mathbf{v}}^n\|_{\mathbf{M}_f(\mathbf{x}_1^n)}$ can be absorbed by the left-hand side of (6.64). The terms $-\frac{1}{\tau}(\delta\mathbf{e}_{\mathbf{v}}^n)\cdot\Big(\mathbf{M}_f(\mathbf{x}_1^n)-\mathbf{M}_f(\mathbf{x}_1^{*,n})\Big)(\delta\mathbf{v}^{*,n})$ and $(\delta\mathbf{e}_{\mathbf{v}}^n)\cdot(\mathbf{f}_1^n-\mathbf{f}_1^{*,n})$ can be estimated in the same way. For the other terms, we will use the following equalities:

$$
\begin{aligned}
&(\delta\dot{\mathbf{e}}_{\mathbf{u}}^n)\cdot\mathbf{E}_s(\mathbf{x}_2)\mathbf{e}_{\mathbf{u}}^n=\Big(\dot{\mathbf{e}}_{\mathbf{u}}^n\cdot\mathbf{E}_s(\mathbf{x}_2)\mathbf{e}_{\mathbf{u}}^n-\dot{\mathbf{e}}_{\mathbf{u}}^{n-1}\cdot\mathbf{E}_s(\mathbf{x}_2)\mathbf{e}_{\mathbf{u}}^{n-1}\Big)-\tau\dot{\mathbf{e}}_{\mathbf{u}}^{n-1}\cdot\mathbf{E}_s(\mathbf{x}_2)\dot{\mathbf{e}}_{\mathbf{u}}^n+\tau\dot{\mathbf{e}}_{\mathbf{u}}^{n-1}\cdot\mathbf{E}_s(\mathbf{x}_2)\boldsymbol{\mathcal{F}}_4^n,\\
&(\delta\mathbf{e}_{\mathbf{v}}^n)\cdot\Big(\mathbf{C}_f(\mathbf{x}_1^n)-\mathbf{C}_f(\mathbf{x}_1^{*,n})\Big)\mathbf{v}^{*,n}=-\mathbf{e}_{\mathbf{v}}^{n-1}\cdot(\mathbf{C}_f(\mathbf{x}_1^n)-\mathbf{C}_f(\mathbf{x}_1^{*,n}))(\delta\mathbf{v}^{*,n})\\
&\quad+\Big(\mathbf{e}_{\mathbf{v}}^n\cdot(\mathbf{C}_f(\mathbf{x}_1^n)-\mathbf{C}_f(\mathbf{x}_1^{*,n}))\mathbf{v}^{*,n}-\mathbf{e}_{\mathbf{v}}^{n-1}\cdot(\mathbf{C}_f(\mathbf{x}_1^{n-1})-\mathbf{C}_f(\mathbf{x}_1^{*,n-1}))\mathbf{v}^{*,n-1}\Big)\\
&\quad-\mathbf{e}_{\mathbf{v}}^{n-1}\cdot\Big((\mathbf{C}_f(\mathbf{x}_1^n)-\mathbf{C}_f(\mathbf{x}_1^{*,n}))-(\mathbf{C}_f(\mathbf{x}_1^{n-1})-\mathbf{C}_f(\mathbf{x}_1^{*,n-1}))\Big)\mathbf{v}^{*,n-1},\\
&(\delta\mathbf{e}_{\mathbf{v}}^n)\cdot\Big(\mathbf{D}_f(\mathbf{x}_1^n)^\top-\mathbf{D}_f(\mathbf{x}_1^{*,n})^\top\Big)\mathbf{p}^{*,n}=-\mathbf{e}_{\mathbf{v}}^{n-1}\cdot(\mathbf{D}_f(\mathbf{x}_1^n)^\top-\mathbf{D}_f(\mathbf{x}_1^{*,n})^\top)(\delta\mathbf{p}^{*,n})\\
&\quad+\Big(\mathbf{e}_{\mathbf{v}}^n\cdot(\mathbf{D}_f(\mathbf{x}_1^n)^\top-\mathbf{D}_f(\mathbf{x}_1^{*,n})^\top)\mathbf{p}^{*,n}-\mathbf{e}_{\mathbf{v}}^{n-1}\cdot(\mathbf{D}_f(\mathbf{x}_1^{n-1})^\top-\mathbf{D}_f(\mathbf{x}_1^{*,n-1})^\top)\mathbf{p}^{*,n-1}\Big)\\
&\quad-\mathbf{e}_{\mathbf{v}}^{n-1}\cdot\Big((\mathbf{D}_f(\mathbf{x}_1^n)^\top-\mathbf{D}_f(\mathbf{x}_1^{*,n})^\top)-(\mathbf{D}_f(\mathbf{x}_1^{n-1})^\top-\mathbf{D}_f(\mathbf{x}_1^{*,n-1})^\top)\Big)\mathbf{p}^{*,n-1},\\
&(\delta\mathbf{e}_{\mathbf{v}}^n)\cdot(\boldsymbol{\mathcal{E}}_1^n+\boldsymbol{\mathcal{F}}_2^n)=\Big(\mathbf{e}_{\mathbf{v}}^n\cdot(\boldsymbol{\mathcal{E}}_1^n+\boldsymbol{\mathcal{F}}_2^n)-\mathbf{e}_{\mathbf{v}}^{n-1}\cdot(\boldsymbol{\mathcal{E}}_1^{n-1}+\boldsymbol{\mathcal{F}}_2^{n-1})\Big)-\mathbf{e}_{\mathbf{v}}^{n-1}\cdot(\delta\boldsymbol{\mathcal{E}}_1^n+\delta\boldsymbol{\mathcal{F}}_2^n),\\
&(\delta\dot{\mathbf{e}}_{\mathbf{u}}^n)\cdot(\boldsymbol{\mathcal{E}}_2^n+\boldsymbol{\mathcal{F}}_3^n)=\Big(\dot{\mathbf{e}}_{\mathbf{u}}^n\cdot(\boldsymbol{\mathcal{E}}_2^n+\boldsymbol{\mathcal{F}}_3^n)-\dot{\mathbf{e}}_{\mathbf{u}}^{n-1}\cdot(\boldsymbol{\mathcal{E}}_2^{n-1}+\boldsymbol{\mathcal{F}}_3^{n-1})\Big)-\dot{\mathbf{e}}_{\mathbf{u}}^{n-1}\cdot(\delta\boldsymbol{\mathcal{E}}_2^n+\delta\boldsymbol{\mathcal{F}}_3^n).
\end{aligned}
$$

Note that $\delta\mathbf{v}^{*,n}$, $\delta\mathbf{p}^{*,n}$, $\delta\boldsymbol{\mathcal{E}}_1^n$, $\delta\boldsymbol{\mathcal{F}}_2^n$, $\delta\boldsymbol{\mathcal{E}}_2^n$ and $\delta\boldsymbol{\mathcal{F}}_3^n$ have an additional $\tau$, as shown in (6.22) and (6.39). Substituting the above equalities into (6.64) and following a similar procedure in the derivation of (6.60) again, we obtain the following inequality:

$$
\begin{aligned}
&\frac{1}{2\tau}\|\delta\mathbf{e}_{\mathbf{v}}^n\|^2_{\mathbf{M}_f(\mathbf{x}_1^n)}+\frac{1}{2}\Big(\|\mathbf{e}_{\mathbf{v}}^n\|^2_{\mathbf{C}_f(\mathbf{x}_1^n)}-\|\mathbf{e}_{\mathbf{v}}^{n-1}\|^2_{\mathbf{C}_f(\mathbf{x}_1^{n-1})}\Big)+\frac{1}{\tau}\|\delta\dot{\mathbf{e}}_{\mathbf{u}}^n\|^2_{\mathbf{M}_s(\mathbf{x}_2)}\\
\leqslant&C\tau(\tau^2+h^{2k})+C\tau\Big(\|\mathbf{e}_{\mathbf{x}_1}^{n-1}\|^2_{\mathbf{A}_f(\mathbf{x}_1^{n-1})}+\|\mathbf{e}_{\mathbf{x}_1}^{n}\|^2_{\mathbf{A}_f(\mathbf{x}_1^{n})}+\|\mathbf{e}_{\mathbf{w}_1}^{n-1}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{n-1})}+\|\mathbf{e}_{\mathbf{w}_1}^{n-1}\|^2_{\mathbf{A}_f(\mathbf{x}_1^{n-1})}\Big)\\
&+C\tau\Big(\|\mathbf{e}_{\mathbf{x}_1}^{n}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{n})}+\|\mathbf{e}_{\mathbf{v}}^{n-1}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{n-1})}+\|\mathbf{e}_{\mathbf{v}}^{n-1}\|^2_{\mathbf{C}_f(\mathbf{x}_1^{n-1})}+\|\mathbf{e}_{\mathbf{v}}^{n}\|^2_{\mathbf{A}_f(\mathbf{x}_1^{n})}\Big)
\end{aligned}
$$

$$
\begin{aligned}
&+ C\tau\Big(\|\dot{\mathbf{e}}_{\mathbf{u}}^{n-1}\|^2_{\mathbf{M}_s(\mathbf{x}_2)} + \|\dot{\mathbf{e}}_{\mathbf{u}}^{n-1}\|^2_{\mathbf{E}_s(\mathbf{x}_2)} + \|\dot{\mathbf{e}}_{\mathbf{u}}^{n}\|^2_{\mathbf{E}_s(\mathbf{x}_2)}\Big) + (\delta\mathbf{e}_{\mathbf{v}}^n)\cdot\mathbf{D}_f(\mathbf{x}_1^n)^\top\mathbf{e}_{\mathbf{p}}^n\\
&- \Big(\mathbf{e}_{\mathbf{v}}^n\cdot(\mathbf{C}_f(\mathbf{x}_1^n)-\mathbf{C}_f(\mathbf{x}_1^{*,n}))\mathbf{v}^{*,n} - \mathbf{e}_{\mathbf{v}}^{n-1}(\mathbf{C}_f(\mathbf{x}_1^{n-1})-\mathbf{C}_f(\mathbf{x}_1^{*,n-1}))\mathbf{v}^{*,n-1}\Big)\\
&+ \mathbf{e}_{\mathbf{v}}^{n-1}\cdot\Big((\mathbf{C}_f(\mathbf{x}_1^n)-\mathbf{C}_f(\mathbf{x}_1^{*,n})) - (\mathbf{C}_f(\mathbf{x}_1^{n-1})-\mathbf{C}_f(\mathbf{x}_1^{*,n-1}))\Big)\mathbf{v}^{*,n-1}\\
&+ \Big(\mathbf{e}_{\mathbf{v}}^n(\mathbf{D}_f(\mathbf{x}_1^n)-\mathbf{D}_f(\mathbf{x}_1^{*,n}))^\top\mathbf{p}^{*,n} - \mathbf{e}_{\mathbf{v}}^n(\mathbf{D}_f(\mathbf{x}_1^{n-1})-\mathbf{D}_f(\mathbf{x}_1^{*,n-1}))^\top\mathbf{p}^{*,n-1}\Big)\\
&- \mathbf{e}_{\mathbf{v}}^{n-1}\cdot\Big((\mathbf{D}_f(\mathbf{x}_1^n)^\top-\mathbf{D}_f(\mathbf{x}_1^{*,n})^\top) - (\mathbf{D}_f(\mathbf{x}_1^{n-1})^\top-\mathbf{D}_f(\mathbf{x}_1^{*,n-1})^\top)\Big)\mathbf{p}^{*,n-1}\\
&- \Big(\dot{\mathbf{e}}_{\mathbf{u}}^n\cdot\mathbf{E}_s(\mathbf{x}_2)\mathbf{e}_{\mathbf{u}}^n - \dot{\mathbf{e}}_{\mathbf{u}}^{n-1}\cdot\mathbf{E}_s(\mathbf{x}_2)\mathbf{e}_{\mathbf{u}}^{n-1}\Big)\\
&+ (\delta\mathbf{e}_{\mathbf{v}}^n)\cdot\boldsymbol{\vartheta}_1 - \Big(\mathbf{e}_{\mathbf{v}}^n\cdot(\boldsymbol{\mathcal{E}}_1^n+\boldsymbol{\mathcal{F}}_2^n) - \mathbf{e}_{\mathbf{v}}^{n-1}\cdot(\boldsymbol{\mathcal{E}}_1^{n-1}+\boldsymbol{\mathcal{F}}_2^{n-1})\Big)\\
&+ (\delta\dot{\mathbf{e}}_{\mathbf{u}}^n)\cdot\boldsymbol{\vartheta}_2 - \Big(\dot{\mathbf{e}}_{\mathbf{u}}^n\cdot(\boldsymbol{\mathcal{E}}_2^n+\boldsymbol{\mathcal{F}}_3^n) - \dot{\mathbf{e}}_{\mathbf{u}}^{n-1}\cdot(\boldsymbol{\mathcal{E}}_2^{n-1}+\boldsymbol{\mathcal{F}}_3^{n-1})\Big).
\end{aligned}
\tag{6.65}
$$

By taking the difference of the equation (6.40d) on time levels $n$ and $n-1$, we obtain

$$
\begin{aligned}
\mathbf{D}_f(\mathbf{x}_1^n)(\delta\mathbf{e}_{\mathbf{v}}^n) = &-\Big(\mathbf{D}_f(\mathbf{x}_1^n)-\mathbf{D}_f(\mathbf{x}_1^{n-1})\Big)\mathbf{e}_{\mathbf{v}}^{n-1} - \Big(\mathbf{D}_f(\mathbf{x}_1^n)-\mathbf{D}_f(\mathbf{x}_1^{*,n})\Big)(\delta\mathbf{v}^{*,n})\\
&- \Big((\mathbf{D}_f(\mathbf{x}_1^n)-\mathbf{D}_f(\mathbf{x}_1^{*,n})) - (\mathbf{D}_f(\mathbf{x}_1^{n-1})-\mathbf{D}_f(\mathbf{x}_1^{*,n-1}))\Big)\mathbf{v}^{*,n-1} - \delta\boldsymbol{\mathcal{E}}_3^n.
\end{aligned}
\tag{6.66}
$$

We can define a continuous deformation from the intermediate domain $\Omega_{1,h}^{n-1,\theta}$ to $\Omega_{1,h}^{n,\theta}$ by introducing the domain $\Omega_{1,h}^{n-1,\theta,\iota}$ which is determined by the vector $\mathbf{x}_1^{*,n-1}+\theta\mathbf{e}_{\mathbf{x}_1}^{n-1}+\iota\tau(\mathbf{w}_1^{*,n-1}+\theta\mathbf{e}_{\mathbf{w}_1}^{n-1})$. Let $p_h^{\theta,\iota}$ and $v_h^{\theta,\iota}$ be the finite element functions defined on domain $\Omega_{1,h}^{n-1,\theta,\iota}$ with nodal vectors $\mathbf{p}$ and $\mathbf{v}$, respectively. Then we obtain that

$$
\begin{aligned}
&\mathbf{p}\cdot\Big((\mathbf{D}_f(\mathbf{x}_1^n)-\mathbf{D}_f(\mathbf{x}_1^{*,n})) - (\mathbf{D}_f(\mathbf{x}_1^{n-1})-\mathbf{D}_f(\mathbf{x}_1^{*,n-1}))\Big)\mathbf{v}\\
&= \int_0^1\int_0^1 \frac{\mathrm{d}}{\mathrm{d}\iota}\frac{\mathrm{d}}{\mathrm{d}\theta}\int_{\Omega_{1,h}^{n-1,\theta,\iota}} p_h^{\theta,\iota}\nabla\cdot v_h^{\theta,\iota}\mathrm{d}x\mathrm{d}\theta\mathrm{d}\iota.
\end{aligned}
\tag{6.67}
$$

By testing equation (6.66) with $\mathbf{e}_{\mathbf{p}}^n$, applying Lemma 6.1 to (6.67), and using Lemma 6.3–6.4 we obtain the following estimate under the induction assumption (6.46) when $\tau$ is sufficiently small.

$$
\begin{aligned}
\mathbf{e}_{\mathbf{p}}^n\cdot\mathbf{D}_f(\mathbf{x}_1^n)(\delta\mathbf{e}_{\mathbf{v}}^n) \leqslant C\tau\|\mathbf{e}_{\mathbf{p}}^n\|_{\mathbf{M}_f(\mathbf{x}_1^n)}\Big(&h^k + \|\mathbf{e}_{\mathbf{v}}^{n-1}\|_{\mathbf{A}_f(\mathbf{x}_1^{n-1})} + \|\mathbf{e}_{\mathbf{x}_1}^{n-1}\|_{\mathbf{A}_f(\mathbf{x}_1^{n-1})}\\
&+ \|\mathbf{e}_{\mathbf{x}_1}^{n}\|_{\mathbf{A}_f(\mathbf{x}_1^{n})} + \|\mathbf{e}_{\mathbf{w}_1}^{n-1}\|_{\mathbf{A}_f(\mathbf{x}_1^{n-1})}\Big).
\end{aligned}
\tag{6.68}
$$

The terms $\mathbf{e}_{\mathbf{v}}^{n-1}\cdot\Big((\mathbf{C}_f(\mathbf{x}_1^n)-\mathbf{C}_f(\mathbf{x}_1^{*,n}))-(\mathbf{C}_f(\mathbf{x}_1^{n-1})-\mathbf{C}_f(\mathbf{x}_1^{*,n-1}))\Big)\mathbf{v}^{*,n-1}$ and $\mathbf{e}_{\mathbf{v}}^{n-1}\cdot\Big((\mathbf{D}_f(\mathbf{x}_1^n)^\top-\mathbf{D}_f(\mathbf{x}_1^{*,n})^\top)-(\mathbf{D}_f(\mathbf{x}_1^{n-1})^\top-\mathbf{D}_f(\mathbf{x}_1^{*,n-1})^\top)\Big)\mathbf{p}^{*,n-1}$ can be estimated in a similar way. Substituting the obtained results, (6.61) and (6.68) into (6.65), we obtain (6.69).

$$
\begin{aligned}
&\frac{1}{4\tau}\|\delta\mathbf{e}_{\mathbf{v}}^n\|^2_{\mathbf{M}_f(\mathbf{x}_1^n)}+\frac{1}{2}\Big(\|\mathbf{e}_{\mathbf{v}}^n\|^2_{\mathbf{C}_f(\mathbf{x}_1^n)}-\|\mathbf{e}_{\mathbf{v}}^{n-1}\|^2_{\mathbf{C}_f(\mathbf{x}_1^{n-1})}\Big)+\frac{1}{2\tau}\|\delta\dot{\mathbf{e}}_{\mathbf{u}}^n\|^2_{\mathbf{M}_s(\mathbf{x}_2)}\\
\leqslant& C_2\tau(\tau^2+h^{2k})+C\tau\Big(\|\mathbf{e}_{\mathbf{x}_1}^n\|^2_{\mathbf{M}_f(\mathbf{x}_1^n)}+\|\mathbf{e}_{\mathbf{x}_1}^{n-1}\|^2_{\mathbf{A}_f(\mathbf{x}_1^{n-1})}+\|\mathbf{e}_{\mathbf{x}_1}^n\|^2_{\mathbf{A}_f(\mathbf{x}_1^n)}\Big)\\
&+C_2\tau\Big(\|\mathbf{e}_{\mathbf{v}}^{n-1}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{n-1})}+\|\mathbf{e}_{\mathbf{v}}^{n-1}\|^2_{\mathbf{C}_f(\mathbf{x}_1^{n-1})}+\|\mathbf{e}_{\mathbf{v}}^n\|^2_{\mathbf{A}_f(\mathbf{x}_1^n)}+\|\dot{\mathbf{e}}_{\mathbf{u}}^n\|^2_{\mathbf{E}_s(\mathbf{x}_2)}+\|\dot{\mathbf{e}}_{\mathbf{u}}^{n-1}\|^2_{\mathbf{E}_s(\mathbf{x}_2)}\Big)\\
&+\varepsilon_2\tau\Big(\|\dot{\mathbf{e}}_{\mathbf{u}}^{n-1}\|^2_{\mathbf{M}_s(\mathbf{x}_2)}+\|\mathbf{e}_{\mathbf{u}}^n\|^2_{\mathbf{E}_s(\mathbf{x}_2)}\Big)-\Big(\dot{\mathbf{e}}_{\mathbf{u}}^n\cdot\mathbf{E}_s(\mathbf{x}_2)\mathbf{e}_{\mathbf{u}}^n-\dot{\mathbf{e}}_{\mathbf{u}}^{n-1}\cdot\mathbf{E}_s(\mathbf{x}_2)\mathbf{e}_{\mathbf{u}}^{n-1}\Big)\\
&-\Big(\mathbf{e}_{\mathbf{v}}^n\cdot(\mathbf{C}_f(\mathbf{x}_1^n)-\mathbf{C}_f(\mathbf{x}_1^{*,n}))\mathbf{v}^{*,n}-\mathbf{e}_{\mathbf{v}}^{n-1}(\mathbf{C}_f(\mathbf{x}_1^{n-1})-\mathbf{C}_f(\mathbf{x}_1^{*,n-1}))\mathbf{v}^{*,n-1}\Big)\\
&+\Big(\mathbf{e}_{\mathbf{v}}^n(\mathbf{D}_f(\mathbf{x}_1^n)-\mathbf{D}_f(\mathbf{x}_1^{*,n}))^\top\mathbf{p}^{*,n}-\mathbf{e}_{\mathbf{v}}^n(\mathbf{D}_f(\mathbf{x}_1^{n-1})-\mathbf{D}_f(\mathbf{x}_1^{*,n-1}))^\top\mathbf{p}^{*,n-1}\Big)\\
&+(\delta\mathbf{e}_{\mathbf{v}}^n)\cdot\boldsymbol{\vartheta}_1-\Big(\mathbf{e}_{\mathbf{v}}^n\cdot(\boldsymbol{\mathcal{E}}_1^n+\boldsymbol{\mathcal{F}}_2^n)-\mathbf{e}_{\mathbf{v}}^{n-1}\cdot(\boldsymbol{\mathcal{E}}_1^{n-1}+\boldsymbol{\mathcal{F}}_2^{n-1})\Big)\\
&+(\delta\dot{\mathbf{e}}_{\mathbf{u}}^n)\cdot\boldsymbol{\vartheta}_2-\Big(\dot{\mathbf{e}}_{\mathbf{u}}^n\cdot(\boldsymbol{\mathcal{E}}_2^n+\boldsymbol{\mathcal{F}}_3^n)-\dot{\mathbf{e}}_{\mathbf{u}}^{n-1}\cdot(\boldsymbol{\mathcal{E}}_2^{n-1}+\boldsymbol{\mathcal{F}}_3^{n-1})\Big).
\end{aligned}
\tag{6.69}
$$

where $\varepsilon_2$ is a small positive real number, $C_2$ is a constant that depends on $1/\varepsilon_2$.

Since $\mathbf{e}_{\mathbf{v}}^0=0$ and $\dot{\mathbf{e}}_{\mathbf{u}}^0=0$, by summing up (6.69) from $n=1$ to $n=\ell-1$ for $2\leqslant\ell\leqslant m$, and using Lemma 6.2–6.4, the induction assumption (6.46) and Young's inequality, we obtain:

$$
\begin{aligned}
&\frac{1}{4\tau}\sum_{n=1}^{\ell-1}\|\delta\mathbf{e}_{\mathbf{v}}^n\|^2_{\mathbf{M}_f(\mathbf{x}_1^n)}+\frac{1}{2}\|\mathbf{e}_{\mathbf{v}}^{\ell-1}\|^2_{\mathbf{C}_f(\mathbf{x}_1^{\ell-1})}+\frac{1}{2\tau}\sum_{n=1}^{\ell-1}\|\delta\dot{\mathbf{e}}_{\mathbf{u}}^n\|^2_{\mathbf{M}_s(\mathbf{x}_2)}\\
\leqslant& C_2(\tau^2+h^{2k})+C_2\sum_{n=1}^{\ell-1}\tau\Big(\|\mathbf{e}_{\mathbf{x}_1}^n\|^2_{\mathbf{M}_f(\mathbf{x}_1^n)}+\|\mathbf{e}_{\mathbf{x}_1}^n\|^2_{\mathbf{A}_f(\mathbf{x}_1^n)}+\|\mathbf{e}_{\mathbf{v}}^{n-1}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{n-1})}+\|\mathbf{e}_{\mathbf{v}}^n\|^2_{\mathbf{A}_f(\mathbf{x}_1^n)}\Big)\\
&+C_2\Big(\|\mathbf{e}_{\mathbf{x}_1}^{\ell-1}\|^2_{\mathbf{A}_f(\mathbf{x}_1^{\ell-1})}+\|\mathbf{e}_{\mathbf{u}}^{\ell-1}\|^2_{\mathbf{E}_s(\mathbf{x}_2)}\Big)+\varepsilon_2\sum_{n=1}^{\ell-1}\tau\Big(\|\dot{\mathbf{e}}_{\mathbf{u}}^{n-1}\|^2_{\mathbf{M}_s(\mathbf{x}_2)}+\|\mathbf{e}_{\mathbf{u}}^n\|^2_{\mathbf{E}_s(\mathbf{x}_2)}\Big)\\
&+\varepsilon_2\Big(\|\dot{\mathbf{e}}_{\mathbf{u}}^{\ell-1}\|^2_{\mathbf{M}_s(\mathbf{x}_2)}+\|\dot{\mathbf{e}}_{\mathbf{u}}^{\ell-1}\|^2_{\mathbf{E}_s(\mathbf{x}_2)}+\|\mathbf{e}_{\mathbf{v}}^{\ell-1}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{\ell-1})}+\|\mathbf{e}_{\mathbf{v}}^{\ell-1}\|^2_{\mathbf{C}_f(\mathbf{x}_1^{\ell-1})}\Big)+C_2\sum_{n=1}^{\ell-1}\tau\|\dot{\mathbf{e}}_{\mathbf{u}}^n\|^2_{\mathbf{E}_s(\mathbf{x}_2)}.
\end{aligned}
\tag{6.70}
$$

(F) *Estimates for* $\|\dot{\mathbf{e}}_{\mathbf{u}}^n\|_{\mathbf{E}_s(\mathbf{x}_2)}$: To obtain the error estimates from inequalities (6.60) and (6.70), we need to further estimate $\|\dot{\mathbf{e}}_{\mathbf{u}}^n\|_{\mathbf{C}(\mathbf{x}_2)}$. By taking the difference between the equations of (6.40c) on time levels $n$ and $n-1$, we obtain the following equality:

$$
\begin{aligned}
&\frac{1}{\tau}\boldsymbol{\eta}\cdot\Big(\mathbf{M}_f(\mathbf{x}_1^n)(\delta\mathbf{e}_{\mathbf{v}}^n)-\mathbf{M}_f(\mathbf{x}_1^{n-1})(\delta\mathbf{e}_{\mathbf{v}}^{n-1})\Big)+\boldsymbol{\eta}\cdot\Big(\mathbf{C}_f(\mathbf{x}_1^n)\mathbf{e}_{\mathbf{v}}^n-\mathbf{C}_f(\mathbf{x}_1^{n-1})\mathbf{e}_{\mathbf{v}}^{n-1}\Big)\\
&-\boldsymbol{\eta}\cdot\Big(\mathbf{D}_f(\mathbf{x}_1^n)^\top\mathbf{e}_{\mathbf{p}}^n-\mathbf{D}_f(\mathbf{x}_1^{n-1})^\top\mathbf{e}_{\mathbf{p}}^{n-1}\Big)+\frac{1}{\tau}\boldsymbol{\xi}\cdot\mathbf{M}_s(\mathbf{x}_2)(\delta^2\dot{\mathbf{e}}_{\mathbf{u}}^n)+\boldsymbol{\xi}\cdot\mathbf{E}_s(\mathbf{x}_2)(\delta\mathbf{e}_{\mathbf{u}}^n)\\
=&-\frac{1}{\tau}\boldsymbol{\eta}\cdot\Big((\mathbf{M}_f(\mathbf{x}_1^n)-\mathbf{M}_f(\mathbf{x}_1^{*,n}))\cdot(\delta\mathbf{v}^{*,n})-(\mathbf{M}_f(\mathbf{x}_1^{n-1})-\mathbf{M}_f(\mathbf{x}_1^{*,n-1}))\cdot(\delta\mathbf{v}^{*,n-1})\Big)\\
&-\boldsymbol{\eta}\cdot\Big(\mathbf{B}_f(\mathbf{x}_1^n,\mathbf{v}^{n-1}-\mathbf{w}_1^{n-1})\mathbf{v}^n-\mathbf{B}_f(\mathbf{x}_1^{n-1},\mathbf{v}^{n-2}-\mathbf{w}_1^{n-2})\mathbf{v}^{n-1}\Big)+\boldsymbol{\eta}\cdot\Big((\delta\mathbf{f}_1^n)-(\delta\mathbf{f}_1^{*,n})\Big)\\
&+\boldsymbol{\eta}\cdot\Big(\mathbf{B}_f(\mathbf{x}_1^{*,n},\mathbf{v}^{*,n-1}-\mathbf{w}_1^{*,n-1})\mathbf{v}^{*,n}-\mathbf{B}_f(\mathbf{x}_1^{*,n-1},\mathbf{v}^{*,n-2}-\mathbf{w}_1^{*,n-2})\mathbf{v}^{*,n-1}\Big)\\
&-\boldsymbol{\eta}\cdot\Big((\mathbf{C}_f(\mathbf{x}_1^n)-\mathbf{C}_f(\mathbf{x}_1^{*,n}))\mathbf{v}^{*,n}-(\mathbf{C}_f(\mathbf{x}_1^{n-1})-\mathbf{C}_f(\mathbf{x}_1^{*,n-1}))\mathbf{v}^{*,n-1}\Big)\\
&+\boldsymbol{\eta}\cdot\Big((\mathbf{D}_f(\mathbf{x}_1^n)-\mathbf{D}_f(\mathbf{x}_1^{*,n}))^\top\mathbf{p}^{*,n}-(\mathbf{D}_f(\mathbf{x}_1^{n-1})-\mathbf{D}_f(\mathbf{x}_1^{*,n-1}))^\top\mathbf{p}^{*,n-1}\Big)\\
&-\boldsymbol{\eta}\cdot\Big((\delta\boldsymbol{\mathcal{E}}_1^n)+(\delta\boldsymbol{\mathcal{F}}_2^n)\Big)-\boldsymbol{\xi}\cdot\Big((\delta\boldsymbol{\mathcal{E}}_2^n)+(\delta\boldsymbol{\mathcal{F}}_3^n)\Big).
\end{aligned}
\tag{6.71}
$$

By choosing $\boldsymbol{\eta}=\delta\mathbf{e}_{\mathbf{v}}^n$ and $\boldsymbol{\xi}=\delta\dot{\mathbf{e}}_{\mathbf{u}}^n$ in (6.71), we have that

$$
\begin{aligned}
(\delta\mathbf{e}_{\mathbf{v}}^n)\cdot\Big(\mathbf{M}_f(\mathbf{x}_1^n)(\delta\mathbf{e}_{\mathbf{v}}^n)-\mathbf{M}_f(\mathbf{x}_1^{n-1})(\delta\mathbf{e}_{\mathbf{v}}^{n-1})\Big)=&\frac{1}{2}\Big(\|\delta\mathbf{e}_{\mathbf{v}}^n\|^2_{\mathbf{M}_f(\mathbf{x}_1^n)}-\|\delta\mathbf{e}_{\mathbf{v}}^{n-1}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{n-1})}\Big)\\
&+\frac{1}{2}(\delta\mathbf{e}_{\mathbf{v}}^n)\cdot\Big(\mathbf{M}_f(\mathbf{x}_1^n)-\mathbf{M}_f(\mathbf{x}_1^{n-1})\Big)(\delta\mathbf{e}_{\mathbf{v}}^n)+\frac{1}{2}\|\delta^2\mathbf{e}_{\mathbf{v}}^n\|^2_{\mathbf{M}_f(\mathbf{x}_1^{n-1})}.
\end{aligned}
$$

For the other terms, the process is similar. Then by using the same method as in the derivation of (6.60), and (6.67)–(6.68), we obtain:

$$
\begin{aligned}
&\frac{1}{2\tau}\Big(\|\delta\mathbf{e}_{\mathbf{v}}^n\|^2_{\mathbf{M}_f(\mathbf{x}_1^n)}-\|\delta\mathbf{e}_{\mathbf{v}}^{n-1}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{n-1})}\Big)+\frac{1}{2}\|\delta\mathbf{e}_{\mathbf{v}}^n\|^2_{\mathbf{C}_f(\mathbf{x}_1^n)}+\frac{1}{2\tau}\Big(\|\delta\dot{\mathbf{e}}_{\mathbf{u}}^n\|^2_{\mathbf{M}_f(\mathbf{x}_2)}-\|\delta\dot{\mathbf{e}}_{\mathbf{u}}^{n-1}\|^2_{\mathbf{M}_f(\mathbf{x}_2)}\Big)\\
&+\frac{1}{2\tau}\Big(\|\delta\mathbf{e}_{\mathbf{u}}^n\|^2_{\mathbf{E}_s(\mathbf{x}_2)}-\|\delta\mathbf{e}_{\mathbf{u}}^{n-1}\|^2_{\mathbf{E}_s(\mathbf{x}_2)}\Big)\\
\leqslant&C_3(\tau^2h^{2k}+\tau^4)+\varepsilon_3\|\delta\mathbf{e}_{\mathbf{w}_1}^{n-1}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{n-1})}+C_3\tau^2\Big(\|\mathbf{e}_{\mathbf{w}_1}^{n-1}\|^2_{\mathbf{A}_f(\mathbf{x}_1^{n-1})}+\|\mathbf{e}_{\mathbf{x}_1}^{n-1}\|^2_{\mathbf{A}_f(\mathbf{x}_1^{n-1})}\Big)\\
&+C_3\Big(\|\delta\mathbf{e}_{\mathbf{v}}^n\|^2_{\mathbf{M}_f(\mathbf{x}_1^n)}+\|\delta\mathbf{e}_{\mathbf{v}}^{n-1}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{n-1})}\Big)+C_3\tau^2\Big(\|\mathbf{e}_{\mathbf{v}}^{n-1}\|^2_{\mathbf{A}_f(\mathbf{x}_1^{n-1})}+\|\mathbf{e}_{\mathbf{v}}^{n-1}\|^2_{\mathbf{C}_f(\mathbf{x}_1^{n-1})}\Big)\\
&+\varepsilon_3\tau^2\Big(\|\mathbf{e}_{\mathbf{v}}^{n-2}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{n-2})}+\|\mathbf{e}_{\mathbf{w}_1}^{n-2}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{n-2})}+\|\mathbf{e}_{\mathbf{x}_1}^{n}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{n})}+\|\mathbf{e}_{\mathbf{x}_1}^{n}\|^2_{\mathbf{A}_f(\mathbf{x}_1^{n})}\Big)\\
&+(\delta\mathbf{e}_{\mathbf{v}}^n)\cdot\Big(\mathbf{D}_f(\mathbf{x}_1^n)^\top\mathbf{e}_{\mathbf{p}}^n-\mathbf{D}_f(\mathbf{x}_1^{n-1})^\top\mathbf{e}_{\mathbf{p}}^{n-1}\Big)-(\delta\boldsymbol{\mathcal{F}}_4^n)\cdot\mathbf{E}_s(\mathbf{x}_2)(\delta\mathbf{e}_{\mathbf{u}}^n)\\
&-\Big(\dot{\mathbf{e}}_{\mathbf{u}}^n(\delta\boldsymbol{\mathcal{E}}_2^n)-\dot{\mathbf{e}}_{\mathbf{u}}^{n-1}(\delta\boldsymbol{\mathcal{E}}_2^{n-1})\Big)+\dot{\mathbf{e}}_{\mathbf{u}}^{n-1}(\delta^2\boldsymbol{\mathcal{E}}_2^n)-\Big(\dot{\mathbf{e}}_{\mathbf{u}}^n(\delta\boldsymbol{\mathcal{F}}_2^n)-\dot{\mathbf{e}}_{\mathbf{u}}^{n-1}(\delta\boldsymbol{\mathcal{F}}_2^{n-1})\Big)+\dot{\mathbf{e}}_{\mathbf{u}}^{n-1}(\delta^2\boldsymbol{\mathcal{F}}_2^n).
\end{aligned}
\tag{6.72}
$$

To estimate $(\delta\mathbf{e}_{\mathbf{v}}^n)\cdot\Big(\mathbf{D}_f(\mathbf{x}_1^n)^\top\mathbf{e}_{\mathbf{p}}^n-\mathbf{D}_f(\mathbf{x}_1^{n-1})^\top\mathbf{e}_{\mathbf{p}}^{n-1}\Big)$ while avoiding the norm of $\delta\mathbf{e}_{\mathbf{p}}^n$, we use the following equality:

$$\begin{aligned}(\delta\mathbf{e}_{\mathbf{v}}^n)\cdot\Big(\mathbf{D}_f(\mathbf{x}_1^n)^\top\mathbf{e}_{\mathbf{p}}^n-\mathbf{D}_f(\mathbf{x}_1^{n-1})^\top\mathbf{e}_{\mathbf{p}}^{n-1}\Big)&=\Big((\delta\mathbf{e}_{\mathbf{v}}^n)\mathbf{D}_f(\mathbf{x}_1^n)^\top\mathbf{e}_{\mathbf{p}}^n-(\delta\mathbf{e}_{\mathbf{v}}^{n-1})\mathbf{D}_f(\mathbf{x}_1^{n-1})^\top\mathbf{e}_{\mathbf{p}}^{n-1}\Big)\\&\quad-\Big((\delta\mathbf{e}_{\mathbf{v}}^n)\mathbf{D}_f(\mathbf{x}_1^n)^\top-(\delta\mathbf{e}_{\mathbf{v}}^{n-1})\mathbf{D}_f(\mathbf{x}_1^{n-1})^\top\Big)\mathbf{e}_{\mathbf{p}}^{n-1}+(\delta\mathbf{e}_{\mathbf{v}}^n)(\mathbf{D}_f(\mathbf{x}_1^n)^\top-\mathbf{D}_f(\mathbf{x}_1^{n-1})^\top)\mathbf{e}_{\mathbf{p}}^{n-1}.\end{aligned}\tag{6.73}$$

Taking the difference of the equation (6.66) on time levels $n$ and $n-1$, and testing the obtained equation with $\mathbf{e}_{\mathbf{p}}^{n-1}$, we have:

$$\begin{aligned}&-\Big((\delta\mathbf{e}_{\mathbf{v}}^n)\cdot\mathbf{D}_f(\mathbf{x}_1^n)^\top-(\delta\mathbf{e}_{\mathbf{v}}^{n-1})\cdot\mathbf{D}_f(\mathbf{x}_1^{n-1})^\top\Big)\mathbf{e}_{\mathbf{p}}^{n-1}\\=&\mathbf{e}_{\mathbf{v}}^{n-1}\Big(\mathbf{D}_f(\mathbf{x}_1^n)-2\mathbf{D}_f(\mathbf{x}_1^{n-1})+\mathbf{D}_f(\mathbf{x}_1^{n-2})\Big)^\top\mathbf{e}_{\mathbf{p}}^{n-1}+(\delta\mathbf{e}_{\mathbf{v}}^{n-1})\Big(\mathbf{D}_f(\mathbf{x}_1^{n-1})-\mathbf{D}_f(\mathbf{x}_1^{n-2})\Big)^\top\mathbf{e}_{\mathbf{p}}^{n-1}\\&+(\delta\mathbf{v}^{*,n})\Big((\mathbf{D}_f(\mathbf{x}_1^n)-\mathbf{D}_f(\mathbf{x}_1^{*,n}))-(\mathbf{D}_f(\mathbf{x}_1^{n-1})-\mathbf{D}_f(\mathbf{x}_1^{*,n-1}))\Big)^\top\mathbf{e}_{\mathbf{p}}^{n-1}\\&+(\delta^2\mathbf{v}^{*,n})\Big(\mathbf{D}_f(\mathbf{x}_1^{n-1})-\mathbf{D}_f(\mathbf{x}_1^{*,n-1})\Big)^\top\mathbf{e}_{\mathbf{p}}^{n-1}+(\delta^2\boldsymbol{\mathcal{E}}_3^n)\cdot\mathbf{e}_{\mathbf{p}}^{n-1}\\&+\mathbf{v}^{*,n-1}\Big((\mathbf{D}_f(\mathbf{x}_1^n)-2\mathbf{D}_f(\mathbf{x}_1^{n-1})+\mathbf{D}_f(\mathbf{x}_1^{n-2}))-(\mathbf{D}_f(\mathbf{x}_1^{*,n})-2\mathbf{D}_f(\mathbf{x}_1^{*,n-1})+\mathbf{D}_f(\mathbf{x}_1^{*,n-2}))\Big)^\top\mathbf{e}_{\mathbf{p}}^{n-1}\\&+(\delta\mathbf{v}^{*,n-1})\Big((\mathbf{D}_f(\mathbf{x}_1^{n-1})-\mathbf{D}_f(\mathbf{x}_1^{*,n-1}))-(\mathbf{D}_f(\mathbf{x}_1^{n-2})-\mathbf{D}_f(\mathbf{x}_1^{*,n-2}))\Big)^\top\mathbf{e}_{\mathbf{p}}^{n-1}\\\leqslant&C\tau\|\mathbf{e}_{\mathbf{p}}^{n-1}\|_{\mathbf{M}_f(\mathbf{x}_1^{n-1})}\Big(\|\delta\mathbf{e}_{\mathbf{w}_1}^{n-1}\|_{\mathbf{A}_f(\mathbf{x}_1^{n-1})}+\|\delta\mathbf{e}_{\mathbf{v}}^{n-1}\|_{\mathbf{A}_f(\mathbf{x}_1^{n-1})}+\tau\|\mathbf{e}_{\mathbf{v}}^{n-1}\|_{\mathbf{A}_f(\mathbf{x}_1^{n-1})}+\tau h^k\Big)\\&+C\tau^2\|\mathbf{e}_{\mathbf{p}}^{n-1}\|_{\mathbf{M}_f(\mathbf{x}_1^{n-1})}\Big(\|\mathbf{e}_{\mathbf{w}_1}^{n-2}\|_{\mathbf{A}_f(\mathbf{x}_1^{n-2})}+\|\mathbf{e}_{\mathbf{x}_1}^{n-2}\|_{\mathbf{A}_f(\mathbf{x}_1^{n-2})}+\|\mathbf{e}_{\mathbf{x}_1}^{n-1}\|_{\mathbf{A}_f(\mathbf{x}_1^{n-1})}+\|\mathbf{e}_{\mathbf{w}_1}^{n-1}\|_{\mathbf{A}_f(\mathbf{x}_1^{n-1})}\Big).\end{aligned}\tag{6.74}$$

To estimate $\|\delta\mathbf{e}_{\mathbf{w}_1}^{n-1}\|_{\mathbf{A}_f(\mathbf{x}_1^{n-1})}$, we take the difference of the equation (6.40a) on the time steps $n$ and $n-1$, and obtain:

$$\begin{aligned}\mathbf{A}_f(\mathbf{x}_1^{n-1})(\delta\mathbf{e}_{\mathbf{w}_1}^{n-1})=&-(\mathbf{A}_f(\mathbf{x}_1^{n-1})-\mathbf{A}_f(\mathbf{x}_1^{n-2}))\mathbf{e}_{\mathbf{w}_1}^{n-2}-(\mathbf{A}_f(\mathbf{x}_1^{n-2})-\mathbf{A}_f(\mathbf{x}_1^{*,n-2}))(\delta\mathbf{w}_1^{*,n-2})\\&-\Big((\mathbf{A}_f(\mathbf{x}_1^{n-1})-\mathbf{A}_f(\mathbf{x}_1^{*,n-1}))-(\mathbf{A}_f(\mathbf{x}_1^{n-2})-\mathbf{A}_f(\mathbf{x}_1^{*,n-2}))\Big)\mathbf{w}_1^{*,n-2}-\delta\boldsymbol{\mathcal{E}}_4^{n-1}.\end{aligned}\tag{6.75}$$

Similarly to the proof of (6.57), and by using the estimate (6.57), we have

$$\begin{aligned}\|\delta\mathbf{e}_{\mathbf{w}_1}^{n-1}\|^2_{\mathbf{A}_f(\mathbf{x}_1^{n-1})}\leqslant&C\Big(\|\delta\mathbf{e}_{\mathbf{v}}^{n-1}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{n-1})}+\|\delta\mathbf{e}_{\mathbf{v}}^{n-1}\|^2_{\mathbf{C}_f(\mathbf{x}_1^{n-1})}\Big)\\&+C\tau^2\Big(h^{2k}+\|\mathbf{e}_{\mathbf{x}_1}^{n-1}\|^2_{\mathbf{A}_f(\mathbf{x}_1^{n-1})}+\|\mathbf{e}_{\mathbf{x}_1}^{n-2}\|^2_{\mathbf{A}_f(\mathbf{x}_1^{n-2})}+\|\mathbf{e}_{\mathbf{v}}^{n-2}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{n-2})}+\|\mathbf{e}_{\mathbf{v}}^{n-2}\|^2_{\mathbf{C}_f(\mathbf{x}_1^{n-2})}\Big).\end{aligned}\tag{6.76}$$

Substituting (6.57), (6.61), (6.74)–, (6.74), and (6.76) into (6.72), we obtain:

$$
\begin{aligned}
&\frac{1}{2\tau}\Big(\|\delta\mathbf{e}_{\mathbf{v}}^{n}\|^2_{\mathbf{M}_f(\mathbf{x}_1^n)}-\|\delta\mathbf{e}_{\mathbf{v}}^{n-1}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{n-1})}\Big)+\frac{1}{2}\|\delta\mathbf{e}_{\mathbf{v}}^{n}\|^2_{\mathbf{C}_f(\mathbf{x}_1^n)}+\frac{1}{2\tau}\Big(\|\delta\dot{\mathbf{e}}_{\mathbf{u}}^{n}\|^2_{\mathbf{M}_s(\mathbf{x}_2)}-\|\delta\dot{\mathbf{e}}_{\mathbf{u}}^{n-1}\|^2_{\mathbf{M}_s(\mathbf{x}_2)}\Big)\\
&+\frac{1}{2\tau}\Big(\|\delta\mathbf{e}_{\mathbf{u}}^{n}\|^2_{\mathbf{E}_s(\mathbf{x}_2)}-\|\delta\mathbf{e}_{\mathbf{u}}^{n-1}\|^2_{\mathbf{E}_s(\mathbf{x}_2)}\Big)\\
\leqslant&C_3(\tau^2h^{2k}+\tau^4)+C\varepsilon_3\|\delta\mathbf{e}_{\mathbf{v}}^{n-1}\|^2_{\mathbf{C}_f(\mathbf{x}_1^{n-1})}+C_3\tau^2\Big(\|\mathbf{e}_{\mathbf{x}_1}^{n-1}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{n-1})}+\|\mathbf{e}_{\mathbf{x}_1}^{n-1}\|^2_{\mathbf{A}_f(\mathbf{x}_1^{n-1})}\Big)\\
&+C_3\Big(\|\delta\mathbf{e}_{\mathbf{v}}^{n}\|^2_{\mathbf{M}_f(\mathbf{x}_1^n)}+\|\delta\mathbf{e}_{\mathbf{v}}^{n-1}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{n-1})}\Big)+C_3\tau^2\Big(\|\mathbf{e}_{\mathbf{v}}^{n-1}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{n-1})}+\|\mathbf{e}_{\mathbf{v}}^{n-1}\|^2_{\mathbf{C}_f(\mathbf{x}_1^{n-1})}\Big)\\
&+C\varepsilon_3\tau^2\Big(\|\mathbf{e}_{\mathbf{v}}^{n-2}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{n-2})}+\|\mathbf{e}_{\mathbf{v}}^{n-2}\|^2_{\mathbf{C}_f(\mathbf{x}_1^{n-2})}+\|\mathbf{e}_{\mathbf{x}_1}^{n}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{n})}+\|\mathbf{e}_{\mathbf{x}_1}^{n}\|^2_{\mathbf{A}_f(\mathbf{x}_1^{n})}+\|\mathbf{e}_{\mathbf{x}_1}^{n-2}\|^2_{\mathbf{A}_f(\mathbf{x}_1^{n-2})}\Big)\\
&+\varepsilon_3\|\delta\mathbf{e}_{\mathbf{u}}^{n}\|^2_{\mathbf{E}_s(\mathbf{x}_2)}+\Big((\delta\mathbf{e}_{\mathbf{v}}^{n})\mathbf{D}_f(\mathbf{x}_1^n)^{\top}\mathbf{e}_{\mathbf{p}}^{n}-(\delta\mathbf{e}_{\mathbf{v}}^{n-1})\mathbf{D}_f(\mathbf{x}_1^{n-1})^{\top}\mathbf{e}_{\mathbf{p}}^{n-1}\Big)\\
&+\varepsilon_3\tau^2\Big(\|\dot{\mathbf{e}}_{\mathbf{u}}^{n-1}\|^2_{\mathbf{M}_s(\mathbf{x}_2)}+\|\dot{\mathbf{e}}_{\mathbf{u}}^{n-1}\|^2_{\mathbf{E}_s(\mathbf{x}_2)}\Big)+C_3\tau^2\|\mathbf{e}_{\mathbf{u}}^{n-1}\|^2_{\mathbf{E}_s(\mathbf{x}_2)}\\
&+C_3\|\delta\dot{\mathbf{e}}_{\mathbf{u}}^{n-1}\|^2_{\mathbf{M}_s(\mathbf{x}_2)}-\Big(\dot{\mathbf{e}}_{\mathbf{u}}^{n}(\delta\boldsymbol{\mathcal{E}}_2^{n})-\dot{\mathbf{e}}_{\mathbf{u}}^{n-1}(\delta\boldsymbol{\mathcal{E}}_2^{n-1})\Big)-\Big(\dot{\mathbf{e}}_{\mathbf{u}}^{n}(\delta\boldsymbol{\mathcal{F}}_2^{n})-\dot{\mathbf{e}}_{\mathbf{u}}^{n-1}(\delta\boldsymbol{\mathcal{F}}_2^{n-1})\Big).
\end{aligned}
\tag{6.77}
$$

If $m\geqslant 3$, we can sum up the inequality (6.77) from $n=2$ to $n=\ell-1$ for $3\leqslant\ell\leqslant m$, and obtain the following inequality by using the estimates (6.68) and (6.61), (6.57).

$$
\begin{aligned}
&\frac{1}{2\tau^2}\|\delta\mathbf{e}_{\mathbf{v}}^{\ell-1}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{\ell-1})}+\frac{1}{4\tau}\sum_{n=2}^{\ell-1}\|\delta\mathbf{e}_{\mathbf{v}}^{n}\|^2_{\mathbf{C}_f(\mathbf{x}_1^{n})}+\frac{1}{2\tau^2}\|\delta\dot{\mathbf{e}}_{\mathbf{u}}^{\ell-1}\|^2_{\mathbf{M}_s(\mathbf{x}_2)}+\frac{1}{2\tau^2}\|\delta\mathbf{e}_{\mathbf{u}}^{\ell-1}\|^2_{\mathbf{E}_s(\mathbf{x}_2)}\\
\leqslant&\frac{1}{2\tau^2}\Big(\|\delta\mathbf{e}_{\mathbf{v}}^{1}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{1})}+\|\delta\dot{\mathbf{e}}_{\mathbf{u}}^{1}\|^2_{\mathbf{M}_f(\mathbf{x}_2)}+\|\delta\mathbf{e}_{\mathbf{u}}^{1}\|^2_{\mathbf{E}_s(\mathbf{x}_2)}\Big)+C_3(\tau^2+h^{2k})+\frac{\varepsilon_3}{\tau}\sum_{n=2}^{\ell-1}\|\delta\mathbf{e}_{\mathbf{u}}^{n}\|^2_{\mathbf{E}_s(\mathbf{x}_2)}\\
&+2C_3\tau\sum_{n=1}^{\ell-1}\Big(\|\mathbf{e}_{\mathbf{x}_1}^{n}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{n})}+\|\mathbf{e}_{\mathbf{x}_1}^{n}\|^2_{\mathbf{A}_f(\mathbf{x}_1^{n})}\Big)+\varepsilon_3\tau\sum_{n=2}^{\ell-1}\Big(\|\dot{\mathbf{e}}_{\mathbf{u}}^{n-1}\|^2_{\mathbf{M}_s(\mathbf{x}_2)}+\|\dot{\mathbf{e}}_{\mathbf{u}}^{n-1}\|^2_{\mathbf{E}_s(\mathbf{x}_2)}\Big)\\
&+2C_3\tau\sum_{n=2}^{\ell-1}\Big(\|\mathbf{e}_{\mathbf{v}}^{n-1}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{n-1})}+\|\mathbf{e}_{\mathbf{v}}^{n-1}\|^2_{\mathbf{C}_f(\mathbf{x}_1^{n-1})}\Big)+C_3\tau\Big(\|\dot{\mathbf{e}}_{\mathbf{u}}^{\ell-1}\|_{\mathbf{M}_s(\mathbf{x}_2)}+\|\dot{\mathbf{e}}_{\mathbf{u}}^{1}\|_{\mathbf{M}_s(\mathbf{x}_2)}\Big)\\
&+\varepsilon_3\Big(\|\mathbf{e}_{\mathbf{v}}^{\ell-1}\|^2_{\mathbf{C}_f(\mathbf{x}_1^{\ell-1})}+\|\mathbf{e}_{\mathbf{x}_1}^{\ell-1}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{\ell-1})}+\|\mathbf{e}_{\mathbf{u}}^{\ell-1}\|^2_{\mathbf{E}_s(\mathbf{x}_2)}\Big)+\frac{2C_3}{\tau}\sum_{n=1}^{\ell-1}\|\delta\mathbf{e}_{\mathbf{v}}^{n}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{n})}\\
&+C_3\tau\sum_{n=2}^{\ell-1}\Big(\|\mathbf{e}_{\mathbf{u}}^{n-1}\|^2_{\mathbf{E}_s(\mathbf{x}_2)}\Big)+\frac{C_3}{\tau}\sum_{n=2}^{\ell-1}\|\delta\dot{\mathbf{e}}_{\mathbf{u}}^{n-1}\|^2_{\mathbf{M}_s(\mathbf{x}_2)}+\frac{1}{\tau}\|\delta\mathbf{e}_{\mathbf{v}}^{1}\|_{\mathbf{A}_f(\mathbf{x}_1^{1})}\|\mathbf{e}_{\mathbf{p}}^{1}\|_{\mathbf{M}_f(\mathbf{x}_1^{1})}\\
&+C_3\Big(\|\mathbf{e}_{\mathbf{v}}^{\ell-2}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{\ell-2})}+\|\mathbf{e}_{\mathbf{v}}^{\ell-2}\|^2_{\mathbf{C}_f(\mathbf{x}_1^{\ell-2})}+\|\mathbf{e}_{\mathbf{x}_1}^{\ell-2}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{\ell-2})}+\|\mathbf{e}_{\mathbf{x}_1}^{\ell-2}\|^2_{\mathbf{A}_f(\mathbf{x}_1^{\ell-2})}\Big)\\
&+\varepsilon_3\Big(\|\dot{\mathbf{e}}_{\mathbf{u}}^{\ell-1}\|^2_{\mathbf{M}_s(\mathbf{x}_2)}+\|\dot{\mathbf{e}}_{\mathbf{u}}^{\ell-1}\|^2_{\mathbf{E}_s(\mathbf{x}_2)}+\|\dot{\mathbf{e}}_{\mathbf{u}}^{1}\|^2_{\mathbf{M}_s(\mathbf{x}_2)}+\|\dot{\mathbf{e}}_{\mathbf{u}}^{1}\|^2_{\mathbf{E}_s(\mathbf{x}_2)}\Big).
\end{aligned}
\tag{6.78}
$$

(G) *Error estimates for* $m \leqslant 3$: The inequality (6.78) holds only when $m \geqslant 3$. To make use of (6.78), we must first prove that the induction assumption (6.46) holds for $m \leqslant 3$. Consequently, we need to prove the case when $m = 2, 3$ separately.

Since we have proved that $m \geqslant 1$, we can choose $m = 1$ in (6.59), and obtain that

$$\|\mathbf{e}_{\mathbf{x}_1}^1\|_{\mathbf{M}_f(\mathbf{x}_1^1)} + \|\mathbf{e}_{\mathbf{x}_1}^1\|_{\mathbf{A}_f(\mathbf{x}_1^1)} \leqslant C(\tau + h^k). \tag{6.79}$$

By inverse inequality, we have that

$$\|\nabla e_{x1}^{1,1}\|_{L^\infty(\Omega_{1,h}^1)} \leqslant Ch^{-\frac{d}{2}}(\tau + h^k).$$

When $h$ is sufficiently small, the equivalence (6.47) holds for $n = 1$ under the mesh size restriction $\tau = o(h^{\frac{d}{2}})$. Hence, we have that

$$\|\mathbf{e}_{\mathbf{x}_1}^1\|_{\mathbf{M}_f(\mathbf{x}_1^{*,1})} + \|\mathbf{e}_{\mathbf{x}_1}^1\|_{\mathbf{A}_f(\mathbf{x}_1^{*,1})} \leqslant C(\tau + h^k).$$

Again, by using inverse inequality, we can prove that the induction assumption (6.46a) holds when $m \leqslant 2$ for sufficiently small $h$ and $\tau = o(h^{\frac{d}{2}})$. Since (6.46b) holds for $m = 1$, we can choose $n = 1$ in (6.64) and obtain that

$$\begin{aligned}
&\frac{1}{2\tau}\|\delta\mathbf{e}_{\mathbf{v}}^1\|_{\mathbf{M}_f(\mathbf{x}_1^1)}^2 + \frac{1}{2}\|\mathbf{e}_{\mathbf{v}}^1\|_{\mathbf{C}_f(\mathbf{x}_1^1)}^2 + \frac{1}{2\tau}\|\delta\dot{\mathbf{e}}_{\mathbf{u}}^1\|_{\mathbf{M}_s(\mathbf{x}_2)}^2 + (\delta\dot{\mathbf{e}}_{\mathbf{u}}^1)\cdot\mathbf{E}_s(\mathbf{x}_2)\mathbf{e}_{\mathbf{u}}^1 \\
\leqslant& C\tau(\tau^2 + h^{2k}) + C\tau\|\mathbf{e}_{\mathbf{v}}^1\|_{\mathbf{A}_f(\mathbf{x}_1^1)}^2 + (\delta\mathbf{e}_{\mathbf{v}}^1)\cdot(\boldsymbol{\vartheta}_1 - \boldsymbol{\mathcal{E}}_1^0) + (\delta\dot{\mathbf{e}}_{\mathbf{u}}^1)\cdot(\boldsymbol{\vartheta}_2 - \boldsymbol{\mathcal{E}}_2^0) \\
&- (\delta\mathbf{e}_{\mathbf{v}}^1)\cdot(\delta\boldsymbol{\mathcal{E}}_1^1) - (\delta\dot{\mathbf{e}}_{\mathbf{u}}^1)\cdot(\delta\boldsymbol{\mathcal{E}}_2^1).
\end{aligned} \tag{6.80}$$

Since $\mathbf{e}_{\mathbf{u}}^0 = 0$ and $\dot{\mathbf{e}}_{\mathbf{u}}^0 = 0$, by using the equation (6.40e), we can derive that

$$(\delta\dot{\mathbf{e}}_{\mathbf{u}}^1)\cdot\mathbf{E}_s(\mathbf{x}_2)\mathbf{e}_{\mathbf{u}}^1 = \tau\|\dot{\mathbf{e}}_{\mathbf{u}}^1\|_{\mathbf{E}_s(\mathbf{x}_2)}^2 - \tau\dot{\mathbf{e}}_{\mathbf{u}}^1\cdot\mathbf{E}_s(\mathbf{x}_2)\boldsymbol{\mathcal{F}}_4^1. \tag{6.81}$$

$\delta\boldsymbol{\mathcal{E}}_1^1$ and $\delta\boldsymbol{\mathcal{E}}_2^1$ will have an additional $\tau$. Since $\mathbf{e}_{\mathbf{v}}^0 = 0$, the last two terms on the right-hand side of (6.80) can be absorbed by the left-hand side when $\tau$ is sufficiently small. Combining (6.80), (6.81) and (6.33), we obtain

$$\begin{aligned}
&\frac{1}{4\tau}\|\delta\mathbf{e}_{\mathbf{v}}^1\|_{\mathbf{M}_f(\mathbf{x}_1^1)}^2 + \frac{1}{4}\|\mathbf{e}_{\mathbf{v}}^1\|_{\mathbf{C}_f(\mathbf{x}_1^1)}^2 + \frac{1}{4\tau}\|\delta\dot{\mathbf{e}}_{\mathbf{u}}^1\|_{\mathbf{M}_s(\mathbf{x}_2)}^2 + \frac{1}{2}\tau\|\dot{\mathbf{e}}_{\mathbf{u}}^1\|_{\mathbf{E}_s(\mathbf{x}_2)}^2 \\
\leqslant& C\tau(\tau^2 + h^{2k}) + C\tau\|\mathbf{e}_{\mathbf{v}}^1\|_{\mathbf{A}_f(\mathbf{x}_1^1)}^2 + (\delta\mathbf{e}_{\mathbf{v}}^1)\cdot\mathbf{D}_f(\mathbf{x}_1^0)^\top(\mathbf{p}^{*,0} - \bar{\mathbf{p}}).
\end{aligned} \tag{6.82}$$

We can further bound the term $(\delta\mathbf{e}_{\mathbf{v}}^1)\cdot\mathbf{D}_f(\mathbf{x}_1^0)^\top(\mathbf{p}^{*,0} - \bar{\mathbf{p}})$ by using (6.66) and the estimates (6.29) and (6.57), which gives

$$(\delta\mathbf{e}_{\mathbf{v}}^1)\cdot\mathbf{D}_f(\mathbf{x}_1^0)^\top(\mathbf{p}^{*,0} - \bar{\mathbf{p}}) \leqslant C\tau h^{2k}. \tag{6.83}$$

Substituting (6.83) into (6.82), and applying Young's inequality and Korn's inequality, we can derive the following estimate:

$$\begin{aligned}&\frac{1}{4\tau}\|\delta\mathbf{e}_{\mathbf{v}}^1\|^2_{\mathbf{M}_f(\mathbf{x}_1^1)}+\frac{1}{8}\|\mathbf{e}_{\mathbf{v}}^1\|^2_{\mathbf{C}_f(\mathbf{x}_1^1)}+\frac{1}{4\tau}\|\delta\dot{\mathbf{e}}_{\mathbf{u}}^1\|^2_{\mathbf{M}_s(\mathbf{x}_2)}+\frac{\tau}{2}\|\dot{\mathbf{e}}_{\mathbf{u}}^1\|^2_{\mathbf{E}_s(\mathbf{x}_2)}\\ \leqslant& C\tau(\tau^2+h^{2k})+C\tau\|\mathbf{e}_{\mathbf{v}}^1\|^2_{\mathbf{M}_f(\mathbf{x}_1^1)}.\end{aligned}\tag{6.84}$$

For the same reason, we can choose $\ell=2$ in (6.63) and add up the obtained inequality and (6.84). We then deduce the following inequality when $\varepsilon_1$ and $\tau$ are sufficiently small

$$\begin{aligned}&\frac{1}{8}\|\mathbf{e}_{\mathbf{v}}^1\|^2_{\mathbf{M}_f(\mathbf{x}_1^1)}+\frac{1}{8}\tau\|\mathbf{e}_{\mathbf{v}}^1\|^2_{\mathbf{C}_f(\mathbf{x}_1^1)}+\frac{1}{2}\|\dot{\mathbf{e}}_{\mathbf{u}}^1\|^2_{\mathbf{M}_s(\mathbf{x}_2)}+\frac{1}{4}\|\mathbf{e}_{\mathbf{u}}^1\|^2_{\mathbf{E}_s(\mathbf{x}_2)}\\ &+\frac{1}{8\tau}\|\delta\mathbf{e}_{\mathbf{v}}^1\|^2_{\mathbf{M}_f(\mathbf{x}_1^1)}+\frac{1}{8}\|\mathbf{e}_{\mathbf{v}}^1\|^2_{\mathbf{C}_f(\mathbf{x}_1^1)}+\frac{1}{8\tau}\|\delta\dot{\mathbf{e}}_{\mathbf{u}}^1\|^2_{\mathbf{M}_s(\mathbf{x}_2)}+\frac{\tau}{4}\|\dot{\mathbf{e}}_{\mathbf{u}}^1\|^2_{\mathbf{E}_s(\mathbf{x}_2)}\leqslant C(\tau^2+h^{2k}).\end{aligned}\tag{6.85}$$

Substituting (6.85) into (6.84), we have that

$$\|\delta\mathbf{e}_{\mathbf{v}}^1\|^2_{\mathbf{M}_f(\mathbf{x}_1^1)}+\tau\|\mathbf{e}_{\mathbf{v}}^1\|^2_{\mathbf{C}_f(\mathbf{x}_1^1)}+\|\delta\dot{\mathbf{e}}_{\mathbf{u}}^1\|^2_{\mathbf{M}_s(\mathbf{x}_2)}+\tau^2\|\dot{\mathbf{e}}_{\mathbf{u}}^1\|^2_{\mathbf{E}_s(\mathbf{x}_2)}\leqslant C\tau^2(\tau^2+h^{2k}).\tag{6.86}$$

(6.85) together with (6.86), (6.61), (6.57), (6.79), (6.49) implies the error estimates (6.45) for $n=1$. The induction assumption (6.46) for $m=2$ then follows from inverse inequality.

Following a similar proof, we can derive that the error estimates (6.45) hold for $n=2$ and the induction assumption (6.46) holds for $m=3$.

(H) *Combination of the above estimates*: Adding up the inequalities (6.59), (6.63), (6.70) and (6.78), we obtain the following combined estimate:

$$\begin{aligned}&\|\mathbf{e}_{\mathbf{x}_1}^{\ell}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{\ell})}+\|\mathbf{e}_{\mathbf{x}_1}^{\ell}\|^2_{\mathbf{A}_f(\mathbf{x}_1^{\ell})}+\frac{1}{2}\|\mathbf{e}_{\mathbf{v}}^{\ell-1}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{\ell-1})}+\frac{1}{8}\sum_{n=1}^{\ell-1}\tau\|\mathbf{e}_{\mathbf{v}}^n\|^2_{\mathbf{C}_f(\mathbf{x}_1^n)}+\frac{1}{2}\|\dot{\mathbf{e}}_{\mathbf{u}}^{\ell-1}\|^2_{\mathbf{M}_f(\mathbf{x}_2)}\\ &+\frac{1}{2}\|\mathbf{e}_{\mathbf{u}}^{\ell-1}\|^2_{\mathbf{E}_s(\mathbf{x}_2)}+\frac{1}{4\tau}\sum_{n=1}^{\ell-1}\|\delta\mathbf{e}_{\mathbf{v}}^n\|^2_{\mathbf{M}_f(\mathbf{x}_1^n)}+\frac{1}{2\tau}\sum_{n=1}^{\ell-1}\|\delta\dot{\mathbf{e}}_{\mathbf{u}}^n\|^2_{\mathbf{M}_s(\mathbf{x}_2)}+\frac{1}{2\tau^2}\|\delta\mathbf{e}_{\mathbf{v}}^{\ell-1}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{\ell-1})}\\ &+\frac{1}{4\tau}\sum_{n=2}^{\ell-1}\|\delta\mathbf{e}_{\mathbf{v}}^n\|^2_{\mathbf{C}_f(\mathbf{x}_1^n)}+\frac{1}{2\tau^2}\|\delta\dot{\mathbf{e}}_{\mathbf{u}}^{\ell-1}\|^2_{\mathbf{M}_s(\mathbf{x}_2)}+\frac{1}{2\tau^2}\|\delta\mathbf{e}_{\mathbf{u}}^{\ell-1}\|^2_{\mathbf{E}_s(\mathbf{x}_2)}\\ \leqslant&C(\tau^2+h^{2k})+C\sum_{n=1}^{\ell-1}\tau\Big(\|\mathbf{e}_{\mathbf{u}}^n\|^2_{\mathbf{E}_s(\mathbf{x}_2)}+\|\dot{\mathbf{e}}_{\mathbf{u}}^n\|^2_{\mathbf{E}_s(\mathbf{x}_2)}\Big)+\frac{\varepsilon_3}{\tau}\sum_{n=2}^{\ell-1}\|\delta\mathbf{e}_{\mathbf{u}}^n\|^2_{\mathbf{E}_s(\mathbf{x}_2)}\\ &+\varepsilon_1\sum_{n=1}^{\ell-1}\tau\Big(\frac{1}{\tau^2}\|\delta\mathbf{e}_{\mathbf{v}}^n\|^2_{\mathbf{M}_f(\mathbf{x}_1^n)}+\|\dot{\mathbf{e}}_{\mathbf{u}}^n\|^2_{\mathbf{M}_s(\mathbf{x}_2)}+\frac{1}{\tau^2}\|\delta\dot{\mathbf{e}}_{\mathbf{u}}^n\|^2_{\mathbf{M}_s(\mathbf{x}_2)}\Big)+\varepsilon_2\sum_{n=1}^{\ell-1}\tau\|\dot{\mathbf{e}}_{\mathbf{u}}^{n-1}\|^2_{\mathbf{M}_s(\mathbf{x}_2)}\\ &+C\tau\sum_{n=1}^{\ell-1}\Big(\|\mathbf{e}_{\mathbf{x}_1}^n\|^2_{\mathbf{M}_f(\mathbf{x}_1^n)}+\|\mathbf{e}_{\mathbf{x}_1}^n\|^2_{\mathbf{A}_f(\mathbf{x}_1^n)}+\|\mathbf{e}_{\mathbf{v}}^n\|^2_{\mathbf{M}_f(\mathbf{x}_1^n)}+\|\mathbf{e}_{\mathbf{v}}^n\|^2_{\mathbf{C}_f(\mathbf{x}_1^n)}\Big)\end{aligned}$$

$$
\begin{aligned}
&+\varepsilon_2\Big(\|\mathbf{e}_{\mathbf{v}}^{\ell-1}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{\ell-1})}+\|\mathbf{e}_{\mathbf{v}}^{\ell-1}\|^2_{\mathbf{C}_f(\mathbf{x}_1^{\ell-1})}+\|\dot{\mathbf{e}}_{\mathbf{u}}^{\ell-1}\|^2_{\mathbf{M}_s(\mathbf{x}_2)}+\|\dot{\mathbf{e}}_{\mathbf{u}}^{\ell-1}\|^2_{\mathbf{E}_s(\mathbf{x}_2)}\Big)+\varepsilon_3\|\mathbf{e}_{\mathbf{v}}^{\ell-1}\|^2_{\mathbf{C}_f(\mathbf{x}_1^{\ell-1})}\\
&+C\Big(\|\mathbf{e}_{\mathbf{x}_1}^{\ell-1}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{\ell-1})}+\|\mathbf{e}_{\mathbf{x}_1}^{\ell-1}\|^2_{\mathbf{A}_f(\mathbf{x}_1^{\ell-1})}+\|\mathbf{e}_{\mathbf{u}}^{\ell-1}\|^2_{\mathbf{E}_s(\mathbf{x}_2)}\Big)+\varepsilon_3\tau\sum_{n=2}^{\ell-1}\Big(\|\dot{\mathbf{e}}_{\mathbf{u}}^{n-1}\|^2_{\mathbf{M}_s(\mathbf{x}_2)}+\|\dot{\mathbf{e}}_{\mathbf{u}}^{n-1}\|^2_{\mathbf{E}_s(\mathbf{x}_2)}\Big)\\
&+\frac{1}{2\tau^2}\|\delta\mathbf{e}_{\mathbf{v}}^{1}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{1})}+\frac{1}{2\tau^2}\|\delta\dot{\mathbf{e}}_{\mathbf{u}}^{1}\|^2_{\mathbf{M}_s(\mathbf{x}_2)}+\frac{1}{2\tau^2}\|\delta\mathbf{e}_{\mathbf{u}}^{1}\|^2_{\mathbf{E}_s(\mathbf{x}_2)}\\
&+\frac{C_3}{\tau}\sum_{n=2}^{\ell-1}\|\delta\dot{\mathbf{e}}_{\mathbf{u}}^{n-1}\|^2_{\mathbf{M}_s(\mathbf{x}_2)}+\frac{1}{\tau}\|\delta\mathbf{e}_{\mathbf{v}}^{1}\|_{\mathbf{A}_f(\mathbf{x}_1^{1})}\|\mathbf{e}_{\mathbf{p}}^{1}\|_{\mathbf{M}_f(\mathbf{x}_1^{1})}+C_3\tau\Big(\|\dot{\mathbf{e}}_{\mathbf{u}}^{\ell-1}\|_{\mathbf{M}_s(\mathbf{x}_2)}+\|\dot{\mathbf{e}}_{\mathbf{u}}^{1}\|_{\mathbf{M}_s(\mathbf{x}_2)}\Big)\\
&+C_3\Big(\|\mathbf{e}_{\mathbf{v}}^{\ell-2}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{\ell-2})}+\|\mathbf{e}_{\mathbf{v}}^{\ell-2}\|^2_{\mathbf{C}_f(\mathbf{x}_1^{\ell-2})}+\|\mathbf{e}_{\mathbf{x}_1}^{\ell-2}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{\ell-2})}+\|\mathbf{e}_{\mathbf{x}_1}^{\ell-2}\|^2_{\mathbf{A}_f(\mathbf{x}_1^{\ell-2})}\Big)\\
&+\varepsilon_3\Big(\|\dot{\mathbf{e}}_{\mathbf{u}}^{\ell-1}\|^2_{\mathbf{M}_s(\mathbf{x}_2)}+\|\dot{\mathbf{e}}_{\mathbf{u}}^{\ell-1}\|^2_{\mathbf{E}_s(\mathbf{x}_2)}+\|\dot{\mathbf{e}}_{\mathbf{u}}^{1}\|^2_{\mathbf{M}_s(\mathbf{x}_2)}+\|\dot{\mathbf{e}}_{\mathbf{u}}^{1}\|^2_{\mathbf{E}_s(\mathbf{x}_2)}\Big)+\frac{2C_3}{\tau}\sum_{n=1}^{\ell-1}\|\delta\mathbf{e}_{\mathbf{v}}^{n}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{n})},
\end{aligned}
$$

where $C$ is a constant that depends on $1/\varepsilon_1$, $1/\varepsilon_2$ and $1/\varepsilon_3$. The estimates for $\|\mathbf{e}_{\mathbf{x}_1}^{\ell-2}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{\ell-2})}$, $\|\mathbf{e}_{\mathbf{x}_1}^{\ell-1}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{\ell-1})}$, $\|\mathbf{e}_{\mathbf{x}_1}^{\ell-2}\|^2_{\mathbf{A}_f(\mathbf{x}_1^{\ell-2})}$, $\|\mathbf{e}_{\mathbf{x}_1}^{\ell-1}\|^2_{\mathbf{A}_f(\mathbf{x}_1^{\ell-1})}$, $\|\mathbf{e}_{\mathbf{v}}^{\ell-2}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{\ell-2})}$, $\|\mathbf{e}_{\mathbf{v}}^{\ell-2}\|^2_{\mathbf{C}_f(\mathbf{x}_1^{\ell-2})}$, $\|\mathbf{e}_{\mathbf{u}}^{\ell-1}\|^2_{\mathbf{E}_s(\mathbf{x}_2)}$ follow from (6.59), (6.63), and (6.70). By substituting these estimates into the inequality above and using the estimates (6.85), (6.86), (6.61), and the equation (6.40e), we obtain

$$
\begin{aligned}
&\|\mathbf{e}_{\mathbf{x}_1}^{\ell}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{\ell})}+\|\mathbf{e}_{\mathbf{x}_1}^{\ell}\|^2_{\mathbf{A}_f(\mathbf{x}_1^{\ell})}+\frac{1}{2}\|\mathbf{e}_{\mathbf{v}}^{\ell-1}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{\ell-1})}+\frac{1}{8}\sum_{n=1}^{\ell-1}\tau\|\mathbf{e}_{\mathbf{v}}^{n}\|^2_{\mathbf{C}_f(\mathbf{x}_1^{n})}+\frac{1}{2}\|\dot{\mathbf{e}}_{\mathbf{u}}^{\ell-1}\|^2_{\mathbf{M}_s(\mathbf{x}_2)}\\
&+\frac{1}{2}\|\mathbf{e}_{\mathbf{u}}^{\ell-1}\|^2_{\mathbf{E}_s(\mathbf{x}_2)}+\frac{1}{4\tau}\sum_{n=1}^{\ell-1}\|\delta\mathbf{e}_{\mathbf{v}}^{\ell-1}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{\ell-1})}+\frac{1}{4}\|\mathbf{e}_{\mathbf{v}}^{\ell-1}\|^2_{\mathbf{C}_f(\mathbf{x}_1^{\ell-1})}+\frac{1}{2\tau}\sum_{n=1}^{\ell-1}\|\delta\dot{\mathbf{e}}_{\mathbf{u}}^{\ell-1}\|^2_{\mathbf{M}_s(\mathbf{x}_2)}\\
&+\frac{1}{2\tau^2}\|\delta\mathbf{e}_{\mathbf{v}}^{\ell-1}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{\ell-1})}+\frac{1}{4\tau}\sum_{n=2}^{\ell-1}\|\delta\mathbf{e}_{\mathbf{v}}^{n}\|^2_{\mathbf{C}_f(\mathbf{x}_1^{n})}+\frac{1}{2\tau^2}\|\delta\dot{\mathbf{e}}_{\mathbf{u}}^{\ell-1}\|^2_{\mathbf{M}_s(\mathbf{x}_2)}+\frac{1}{2}\|\dot{\mathbf{e}}_{\mathbf{u}}^{\ell-1}\|^2_{\mathbf{E}_s(\mathbf{x}_2)}\\
\leqslant&C(\tau^2+h^{2k})+C\tau\sum_{n=1}^{\ell-1}\Big(\|\dot{\mathbf{e}}_{\mathbf{u}}^{n}\|^2_{\mathbf{M}_s(\mathbf{x}_2)}+\|\mathbf{e}_{\mathbf{u}}^{n}\|^2_{\mathbf{E}_s(\mathbf{x}_2)}+\|\dot{\mathbf{e}}_{\mathbf{u}}^{n}\|^2_{\mathbf{E}_s(\mathbf{x}_2)}+\|\mathbf{e}_{\mathbf{x}_1}^{n}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{n})}+\|\mathbf{e}_{\mathbf{x}_1}^{n}\|^2_{\mathbf{A}_f(\mathbf{x}_1^{n})}\Big)\\
&+C\tau\sum_{n=1}^{\ell-1}\Big(\|\mathbf{e}_{\mathbf{v}}^{n}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{n})}+\|\mathbf{e}_{\mathbf{v}}^{n}\|^2_{\mathbf{C}_f(\mathbf{x}_1^{n})}\Big)+C\frac{1}{\tau}\sum_{n=1}^{\ell-1}\Big(\|\delta\mathbf{e}_{\mathbf{v}}^{n}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{n})}+\|\delta\dot{\mathbf{e}}_{\mathbf{u}}^{n}\|^2_{\mathbf{M}_s(\mathbf{x}_2)}\Big)\\
&+(\varepsilon_2+\varepsilon_3)\Big(\|\dot{\mathbf{e}}_{\mathbf{u}}^{\ell-1}\|^2_{\mathbf{M}_s(\mathbf{x}_2)}+\|\dot{\mathbf{e}}_{\mathbf{u}}^{\ell-1}\|^2_{\mathbf{E}_s(\mathbf{x}_2)}+\|\mathbf{e}_{\mathbf{v}}^{\ell-1}\|^2_{\mathbf{C}_f(\mathbf{x}_1^{\ell-1})}\Big)+C_3\tau\|\dot{\mathbf{e}}_{\mathbf{u}}^{\ell-1}\|_{\mathbf{M}_s(\mathbf{x}_2)}.
\end{aligned}
$$

We can choose $\tau$, $\varepsilon_2$ and $\varepsilon_3$ to be sufficiently small such that the corresponding terms on the right-hand side can be absorbed by the terms on the left-hand side. Consequently, for $2\leqslant\ell\leqslant m$, we have:

$$
\|\mathbf{e}_{\mathbf{x}_1}^{\ell}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{\ell})}+\|\mathbf{e}_{\mathbf{x}_1}^{\ell}\|^2_{\mathbf{A}_f(\mathbf{x}_1^{\ell})}+\frac{1}{2}\|\mathbf{e}_{\mathbf{v}}^{\ell-1}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{\ell-1})}+\frac{1}{8}\sum_{n=1}^{\ell-1}\tau\|\mathbf{e}_{\mathbf{v}}^{n}\|^2_{\mathbf{C}_f(\mathbf{x}_1^{n})}+\frac{1}{4}\|\dot{\mathbf{e}}_{\mathbf{u}}^{\ell-1}\|^2_{\mathbf{M}_s(\mathbf{x}_2)}
$$

$$
\begin{aligned}
&+\frac{1}{2}\|\mathbf{e}_{\mathbf{u}}^{\ell-1}\|^2_{\mathbf{E}_s(\mathbf{x}_2)}+\frac{1}{4\tau}\sum_{n=1}^{\ell-1}\|\delta\mathbf{e}_{\mathbf{v}}^{n}\|^2_{\mathbf{M}_f(\mathbf{x}_1^n)}+\frac{1}{8}\|\mathbf{e}_{\mathbf{v}}^{\ell-1}\|^2_{\mathbf{C}_f(\mathbf{x}_1^{\ell-1})}+\frac{1}{2\tau}\sum_{n=1}^{\ell-1}\|\delta\dot{\mathbf{e}}_{\mathbf{u}}^{\ell-1}\|^2_{\mathbf{M}_s(\mathbf{x}_2)}\\
&+\frac{1}{2\tau^2}\|\delta\mathbf{e}_{\mathbf{v}}^{\ell-1}\|^2_{\mathbf{M}_f(\mathbf{x}_1^{\ell-1})}+\frac{1}{4\tau}\sum_{n=2}^{\ell-1}\|\delta\mathbf{e}_{\mathbf{v}}^{n}\|^2_{\mathbf{C}_f(\mathbf{x}_1^n)}+\frac{1}{2\tau^2}\|\delta\dot{\mathbf{e}}_{\mathbf{u}}^{\ell-1}\|^2_{\mathbf{M}_s(\mathbf{x}_2)}+\frac{1}{4}\|\dot{\mathbf{e}}_{\mathbf{u}}^{\ell-1}\|^2_{\mathbf{E}_s(\mathbf{x}_2)}\\
\leqslant& C(\tau^2+h^{2k})+C\sum_{n=1}^{\ell-1}\tau\Big(\|\mathbf{e}_{\mathbf{u}}^{n}\|^2_{\mathbf{E}_s(\mathbf{x}_2)}+\|\dot{\mathbf{e}}_{\mathbf{u}}^{n}\|^2_{\mathbf{E}_s(\mathbf{x}_2)}+\frac{1}{\tau^2}\|\delta\mathbf{e}_{\mathbf{v}}^{n}\|^2_{\mathbf{M}_f(\mathbf{x}_1^n)}+\|\dot{\mathbf{e}}_{\mathbf{u}}^{n}\|^2_{\mathbf{M}_s(\mathbf{x}_2)}\Big)\\
&+C\tau\sum_{n=1}^{\ell-1}\Big(\|\mathbf{e}_{\mathbf{x}_1}^{n}\|^2_{\mathbf{M}_f(\mathbf{x}_1^n)}+\|\mathbf{e}_{\mathbf{x}_1}^{n}\|^2_{\mathbf{A}_f(\mathbf{x}_1^n)}+\|\mathbf{e}_{\mathbf{v}}^{n}\|^2_{\mathbf{M}_f(\mathbf{x}_1^n)}+\|\mathbf{e}_{\mathbf{v}}^{n}\|^2_{\mathbf{C}_f(\mathbf{x}_1^n)}+\frac{1}{\tau^2}\|\delta\dot{\mathbf{e}}_{\mathbf{u}}^{n}\|^2_{\mathbf{M}_s(\mathbf{x}_2)}\Big).
\end{aligned}
\tag{6.87}
$$

By applying Gronwall's inequality to (6.87), we obtain

$$
\begin{aligned}
&\sup_{0\leqslant n\leqslant m}\Big\{\|\mathbf{e}_{\mathbf{x}_1}^{n}\|_{\mathbf{M}_f(\mathbf{x}_1^n)}+\|\mathbf{e}_{\mathbf{x}_1}^{n}\|_{\mathbf{A}_f(\mathbf{x}_1^n)}\Big\}+\sup_{0\leqslant n\leqslant m-1}\Big\{\|\mathbf{e}_{\mathbf{v}}^{n}\|_{\mathbf{M}_f(\mathbf{x}_1^n)}+\|\mathbf{e}_{\mathbf{v}}^{n}\|_{\mathbf{C}_f(\mathbf{x}_1^n)}\Big\}\\
&+\sup_{0\leqslant n\leqslant m-1}\Big\{\|\dot{\mathbf{e}}_{\mathbf{u}}^{n}\|_{\mathbf{M}_s(\mathbf{x}_2)}+\|\mathbf{e}_{\mathbf{u}}^{n}\|_{\mathbf{E}_s(\mathbf{x}_2)}+\frac{1}{\tau}\|\delta\dot{\mathbf{e}}_{\mathbf{u}}^{n}\|_{\mathbf{M}_s(\mathbf{x}_2)}+\frac{1}{\tau}\|\delta\mathbf{e}_{\mathbf{v}}^{n}\|_{\mathbf{M}_f(\mathbf{x}_1^n)}\Big\}\leqslant C(\tau+h^k).
\end{aligned}
\tag{6.88}
$$

The estimate for $\mathbf{e}_{\mathbf{p}}^{n}$ follows from (6.61) and (6.88), which implies

$$
\sup_{0\leqslant n\leqslant m-1}\|\mathbf{e}_{\mathbf{p}}^{n}\|_{\mathbf{M}_f(\mathbf{x}_1^n)}\leqslant C(\tau+h^k).
\tag{6.89}
$$

Additionally, since $\mathbf{e}_{\mathbf{u}}^{n}=\sum_{i=1}^{n}\delta\mathbf{e}_{\mathbf{u}}^{i}$, we have

$$
\sup_{0\leqslant n\leqslant m-1}\|\mathbf{e}_{\mathbf{u}}^{n}\|_{\mathbf{M}_s(\mathbf{x}_2)}\leqslant\sup_{0\leqslant n\leqslant m-1}\sum_{i=1}^{n}\|\delta\mathbf{e}_{\mathbf{u}}^{i}\|_{\mathbf{M}_s(\mathbf{x}_2)}\leqslant C\tau+T\sup_{0\leqslant n\leqslant m-1}\|\dot{\mathbf{e}}_{\mathbf{u}}^{n}\|_{\mathbf{M}_s(\mathbf{x}_2)}\leqslant C(\tau+h^k).
\tag{6.90}
$$

The estimate for $\mathbf{e}_{\mathbf{w}_1}^{n}$ follows from (6.57) such that

$$
\sup_{0\leqslant n\leqslant m-1}\|\mathbf{e}_{\mathbf{w}_1}^{n}\|_{\mathbf{A}_f(\mathbf{x}_1^n)}\leqslant C(\tau+h^k).
\tag{6.91}
$$

By using inverse inequality, we can derive the induction assumption (6.46a) for $n=m$ when $\tau$ and $h$ are sufficiently small. Then following a similar process, we can deduce that the inequality (6.87) holds for $2\leqslant\ell\leqslant m+1$. Following from the deduction of (6.88)–(6.91), we have that

$$
\begin{aligned}
&\sup_{0\leqslant n\leqslant m}\Big\{\|\mathbf{e}_{\mathbf{v}}^{n}\|_{\mathbf{M}(\mathbf{x}_1^n)}+\|\mathbf{e}_{\mathbf{v}}^{n}\|_{\mathbf{C}(\mathbf{x}_1^n)}+\|\mathbf{e}_{\mathbf{u}}^{n}\|_{\mathbf{M}(\mathbf{x}_1^n)}+\|\mathbf{e}_{\mathbf{u}}^{n}\|_{\mathbf{C}(\mathbf{x}_1^n)}+\|\mathbf{e}_{\mathbf{p}}^{n}\|_{\mathbf{M}(\mathbf{x}_1^n)}+\|\mathbf{e}_{\mathbf{w}_1}^{n}\|_{\mathbf{A}_f(\mathbf{x}_1^n)}\Big\}\\
&+\sup_{0\leqslant n\leqslant m+1}\Big\{\|\mathbf{e}_{\mathbf{x}_1}^{n}\|_{\mathbf{M}(\mathbf{x}_1^n)}+\|\mathbf{e}_{\mathbf{x}_1}^{n}\|_{\mathbf{A}(\mathbf{x}_1^n)}\Big\}\leqslant C(\tau+h^k).
\end{aligned}
\tag{6.92}
$$

Again by using inverse inequality, we prove the induction assumption (6.46b) for $n = m$, which complete the induction. The desired result (6.45) then follows from the equivalence (6.49). □

**6.5 Numerical Examples.** In this section, we provide numerical examples to substantiate the theoretical results presented in this chapter, thereby demonstrating the convergence of the proposed method. The computational analyses were conducted using the finite element software package NGSolve, available at https://ngsolve.org/.

**Example 6.1** (2D vessel wall deformations under blood flow)**.** In this example, we construct a 2D analytical solution in a situation where blood flows through a blood vessel. For given positive real numbers $R$ and $L$, the initial fluid domain is $\Omega_1(0) = [0, L] \times [-R, R]$. The reference domain of solid is $\Omega_2 = \Big([0, L] \times [-1.5R, R]\Big) \bigcup \Big([0, L] \times [R, 1.5R]\Big)$. Let $g(t, x) = 1/10 \sin^2 x \sin(L - x) \sin t$. The exact solution of fluid velocity is given by

$$v_1 = R^2 - (y - g)^2, \quad v_2 = v_1 \partial_x g + \partial_t g.$$

The velocity $v = (v_1, v_2)^\top$ is then divergence free. The pressure $p = -\mu_1 \partial_x g(\partial_y v_1 + \partial_x v_2) + 2\mu_1 \partial_y v_2 + 10\mu_1 v_1$. Next, to define the displacement $u$, we first define two auxliary functions:

$$\begin{aligned} a_1(t, x) =& \frac{1}{\mu_2}\Big(p(t, x, R)\partial_x g + \mu_1\big(\partial_y v_1(t, x, R) + \partial_x v_2(t, x, R)\big) - 2\mu_1 \partial_x v_1(t, x, R)\partial_x g\Big) - \partial_x g, \\ a_2(t, x) =& \partial_x g - \frac{1}{\mu_2}\Big( - 2\mu_1 \partial_x v_1(t, x, -R)\partial_x g + \mu_1\big(\partial_y v_1(t, x - R) + \partial_x v_2(t, x, -R)\big) \\ & + p(t, x, -R)\partial_x g\Big). \end{aligned}$$

The displacement on the upper solid reference domain $[0, L] \times [R, 1.5R]$ is defined as

$$u^{\text{upper}} = (\sin(y - R)a_1, \cos(y - R)g)^\top,$$

and the displacement on the lower solid reference domain $[0, L] \times [-1.5R, -R]$ is defined as

$$u^{\text{lower}} = (\sin(-R - y)a_2, \cos(-R - y)g)^\top.$$

The source function $f_1$ and $f_2$ are tailored to align with the characteristics of the analytical solutions. To make the test more realistic, we employ dirichlet boundary condition on the left fluid boundary $\{0\} \times [-R, R]$. On the other boundaries, we employ the same type of boundary conditions as those specified in problems (6.1) and (6.3).

In this test, we choose the parameters as follows. $R = 1, L = 2\pi, \rho_1 = 1.06, \mu_1 = 0.04, \rho_2 = 1.2, \mu_2 = E/(2(1+\nu_2)), \lambda_2 = E\nu_2/((1+\nu_2)(1-2\nu_2))$, where $E = 2$ and $\nu_2 = 0.45$.

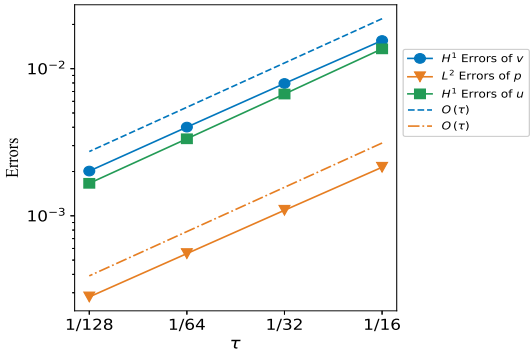
To assess the convergence properties of the numerical scheme, we conducted a convergence test at $T = 1$ using the $P_2$–$P_1$–$P_2$ element and a suitably small mesh size that ensures negligible errors from the spatial discretization. The errors of the numerical solutions are depicted in Figure 6.2 (a) for time step sizes: $\tau = 1/16, 1/32, 1/64, 1/128$. The observed errors demonstrate first-order convergence in time, which aligns with the theoretical findings established in Theorem 6.1.

In addition to investigating the convergence in time, we also conducted a convergence test at time $T = 0.1$ to assess the spatial discretization by utilizing sufficiently small time step sizes to ensure negligible errors from the temporal discretization. Numerical solutions for $v, p, u$ are evaluated, and the corresponding errors are presented in Figure 6.2 (b)–(d) across various mesh sizes $h = 1/4, 1/8, 1/16, 1/32$. The mixed element methods $P_k$–$P_{k-1}$–$P_k$ of degree $k = 2, 3$ are employed. The numerical findings reveal that the $H^1$ error of the velocity $v$ and the displacement $u$, along with the $L^2$ error of the pressure $p$ exhibit $k$-th order convergence in space. This observation aligns with the theoretical result established in Theorem 6.1, specifically for the cases where $k = 2, 3$.

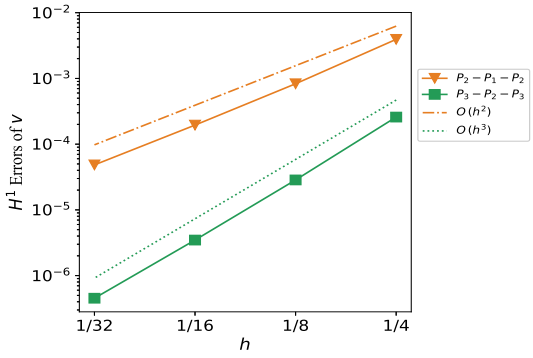
The configurations and meshes of the evolving domain at times $t = 0$ and $t = 2$ are depicted in Figure 6.3, employing a mesh size of $h = 1/8$ and a time step size of $1/50$. The solutions of the equations are obtained using the $P_2$–$P_1$ element, with meshes generated by the $P_2$ element. The fluid and solid phases are highlighted by red and blue meshes, respectively, in Figure 6.3 (a)–(b). In these representations, the interfaces are the intersections of the red and blue domains.

**Example 6.2** (3D vessel wall deformations under blood flow)**.** In this example, we construct a 3D analytical solution in a situation where blood flows through a blood vessel. For given positive real numbers $R$ and $L$, the initial fluid domain is $\Omega_1(0) = \{(y, z) : y^2 + z^2 \leqslant R^2\} \times [0, L]$. The reference domain of solid is $\Omega_2 = \{(y, z) : R^2 \leqslant y^2 + z^2 \leqslant 2.25R^2\} \times [0, L]$. Let $g(t, x) = 1/100 \sin^2 x \sin(L - x) \sin t$. The exact solution of fluid velocity is given by

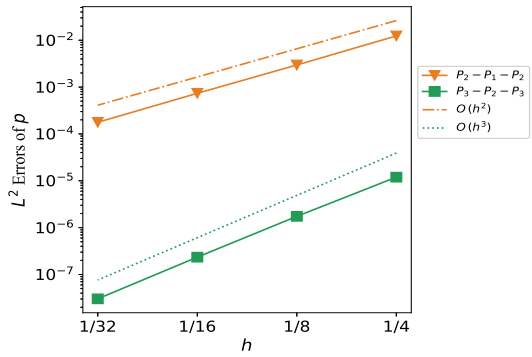
$$
\begin{aligned}
&v_1 = \frac{1}{100}\Big(R^2 - (y-g)^2 - (z-g)^2\Big)^2, \quad v_2 = v_3 = v_1\partial_x g + \partial_t g,\\
&p = -\mu_1\partial_{tx} g\partial_x g - 4(\mu_2 + \lambda_2)yzb(t, x) + \rho_1 v_1,\\
&u_1 = \sin(y^2 + z^2 - R^2)a(t, x, y, z), \quad u_2 = \sin(y^2 + z^2 - R^2)zb(x, t) + g,\\
&u_3 = \sin(y^2 + z^2 - R^2)yb(t, x) + g,
\end{aligned}
$$



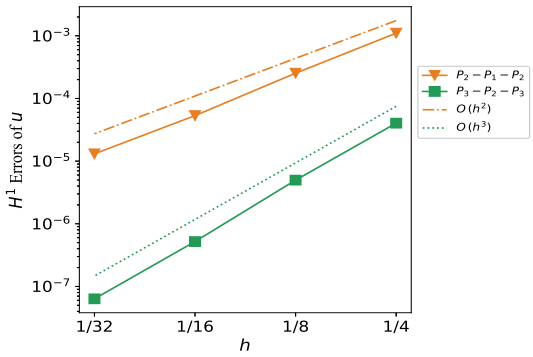

(a) Errors from temporal discretization




(b) $H^1$ error of $v$ from spatial discretization

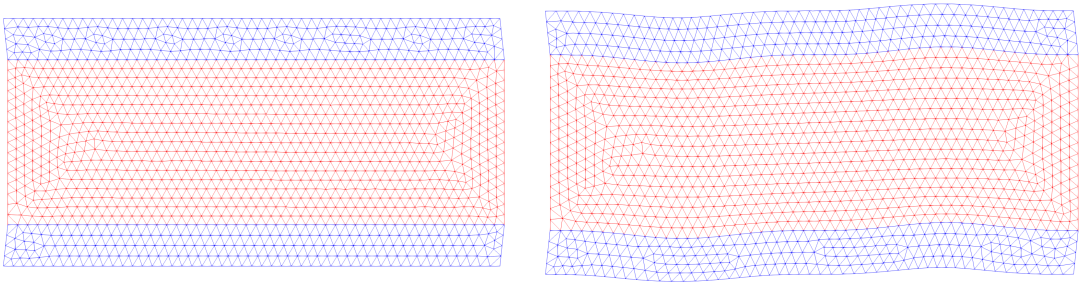



(c) $L^2$ error of $p$ from spatial discretization




(d) $H^1$ error of $u$ from spatial discretization

*Figure 6.2. Errors of the numerical solutions at time $T = 0.1$*



(a) The domain and mesh at time $t = 0$

(b) The domain and mesh at time $t = 2$

*Figure 6.3. Evolution of the domain and mesh*

where $b(t,x)$ and $a(t,x,y,z)$ are defined as

$$b(t,x) := -\frac{\mu_1}{2\mu_2 R^2}\partial_{tx} g \partial_x g,$$

$$a(t,x,y,z) := \frac{1}{2\mu_2 R^2}\Big(\mu_1(y+z)\partial_{tx} g + (y+z)p\partial_x g - \mu_2(y+z)\partial_x g\Big).$$

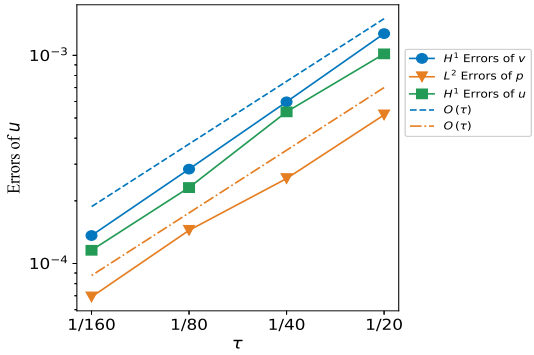
The source function $f_1$ and $f_2$ are tailored to align with the characteristics of the analytical solutions. To make the test more realistic, we employ dirichlet boundary condition on the left fluid boundary $\{(0,y,z) : y^2 + z^2 \leqslant R^2\}$. On the other boundaries, we employ the same type of boundary conditions as those specified in problems (6.1) and (6.3).

In this test, we choose the parameters as follows. $R = 1, L = 1, \rho_1 = 1.06, \mu_1 = 0.04, \rho_2 = 1.2, \mu_2 = E/(2(1+\nu_2)), \lambda_2 = E\nu_2/((1+\nu_2)(1-2\nu_2))$, where $E = 2$ and $\nu_2 = 0.45$.

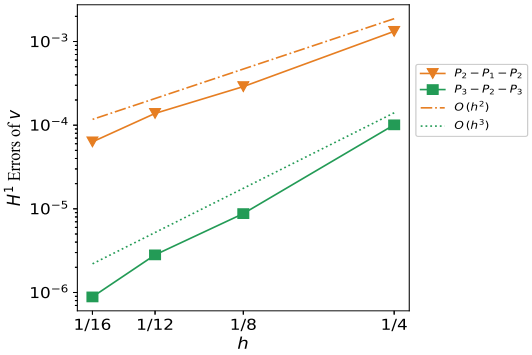
To assess the convergence properties of the numerical scheme, we conducted a convergence test at $T = 1$ using the $P_2$–$P_1$–$P_2$ element and a suitably small mesh size that ensures negligible errors from the spatial discretization. The errors of the numerical solutions are depicted in Figure 6.4 (a) for time step sizes: $\tau = 1/20, 1/40, 1/80, 1/160$. The observed errors demonstrate first-order convergence in time, which aligns with the theoretical findings established in Theorem 6.1.

In addition to investigating the convergence in time, we also conducted a convergence test at time $T = 1$ to assess the spatial discretization by utilizing sufficiently small time step sizes to ensure negligible errors from the temporal discretization. Numerical solutions for $v, p, u$ are evaluated, and the corresponding errors are presented in Figure 6.4 (b)–(d) across various mesh sizes $h = 1/4, 1/8, 1/12, 1/16$. The mixed element methods $P_k$–$P_{k-1}$–$P_k$ of degree $k = 2, 3$ are employed. The numerical findings reveal that the $H^1$ error of the velocity $v$ and the displacement $u$, along with the $L^2$ error of the pressure $p$ exhibit $k$-th order convergence in space. This observation aligns with the theoretical result established in Theorem 6.1, specifically for the cases where $k = 2, 3$.

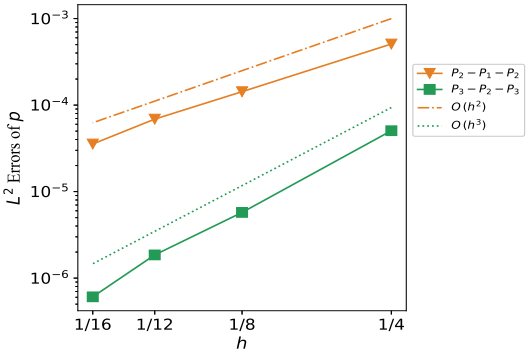
The configurations and meshes of the evolving domain for $R = 1$ and $L = 2\pi$ at times $t = 0$ and $t = 1$ are depicted in Figure 6.5, employing a mesh size of $h = 1/4$ and a time step size of 1/100. The solutions of the equations are obtained using the $P_2$–$P_1$ element, with meshes generated by the $P_2$ element. The fluid and solid phases are highlighted by red and blue meshes, respectively, in Figure 6.5. In these representations, the interfaces are the intersections of the red and blue domains.



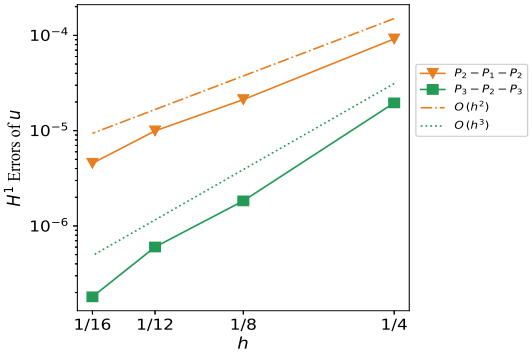

(a) Errors from temporal discretization




(b) $H^1$ error of $v$ from spatial discretization

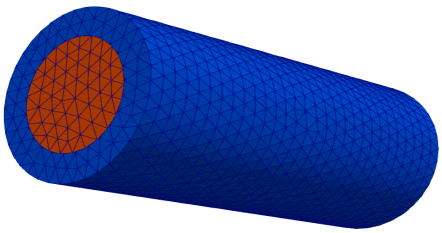



(c) $L^2$ error of $p$ from spatial discretization

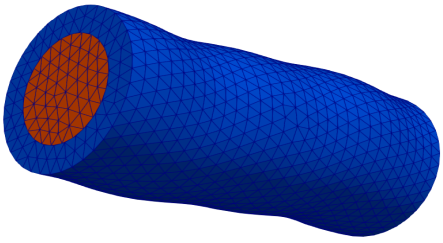



(d) $H^1$ error of $u$ from spatial discretization

*Figure 6.4. Errors of the numerical solutions at time $T = 0.1$*



(a) The domain and mesh at time $t = 0$



(b) The domain and mesh at time $t = 1$

*Figure 6.5. Evolution of the domain and mesh*

# Chapter 7

# Conclusion

In this thesis, several new high-order numerical methods and analyses are developed for some nonlinear parabolic equations with rough data and evolving domains, including

- high-order convergent numerical methods for the parabolic–type equations with rough data (in Chapter 2–3),
- high-order convergent numerical methods for the nonlinear equations with evolving domains (in Chapters 4–6).

In this chapter we draw conclusions and point out some possible research directions related to the work done in this thesis.

In Chapter 2, we have proposed a new spectral method for the linear and semilinear subdiffusion equations in a bounded domain $\Omega \subset \mathbb{R}^d$ under the Dirichlet boundary condition with rough initial data in $L^\infty(\Omega)$ and possibly singular source function by effectively combining several computational techniques, including the contour integral representation of the mild solutions, the quadrature approximation of the contour integrals, the exponential integrator using VP means, and a decomposition of the time interval geometrically refined towards $t = 0$ according to the singularity of the solution and the source function. We have proved the spectral convergence in $L^\infty(0, T; L^\infty)$ norm of the proposed method for both linear and semilinear subdiffusion equations with an arbitrary rough initial value $u_0 \in L^\infty(\Omega)$ under the natural regularity of the solutions with strong singularities at $t = 0$ in the form of (2.6).

In Chapter 3, we have studied numerical treatment for the two-dimensional Navier-Stokes equations with $L^2$ initial data. To date, the best convergence results obtained for fully discrete schemes are limited to first-order accuracy in both time and space, which are suboptimal in space and considered low-order in time. We have proposed a fully discrete

scheme that utilizes the finite element method for spatial discretization and a implicit-explicit Runge–Kutta method in conjunction with graded time meshes. By employing discrete semigroup techniques, sharp regularity estimates, negative norm estimates and the $L^2$ projection onto the divergence-free Raviart–Thomas element space, we have demonstrated that the proposed scheme attains second-order convergence in both space and time. The argument presented in Chapter 3 could be further extended to higher-order implicit-explicit Runge–Kutta schemes. The numerical results are consistent with the theoretical analysis and demonstrate the sharpness of convergence order.

In the future, we may focus on two directions: (i) problems with initial data in negative Sobolev spaces, and (ii) fully nonlinear subdiffusion equations. For the former, it will be essential to develop an efficient framework to characterize functionals in negative Sobolev spaces, and to establish regularity estimates that can guide the construction of effective numerical methods. As for the latter, further investigation is needed, since the validity of the Lipschitz condition plays a crucial role in our current numerical analysis. If this condition fails, entirely different analytical techniques will have to be employed.

In Chapter 4, 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 $\partial_\theta^\bullet$ of the normal vector of the boundary $\Gamma_h^\theta$. 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 in $L^\infty(0, T; H^1)$ norm of the numerical approximations to the boundary evolution under shape gradient flow by using the evolving bulk finite elements systematically studied in [41] and the geometric estimates developed in

[95, 38]. 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 chapter 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:

$$\mathrm{d}J(\Gamma, u; v) = \frac{1}{2}\int_\Omega \Big(\nabla u \cdot (\nabla v + \nabla v^\top)\nabla u - |\nabla u|^2 \nabla \cdot v\Big)\mathrm{d}x - \int_\Omega f v \cdot \nabla u \mathrm{d}x.$$

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 chapter could be trivially extended to the case $j(\cdot, u) = \frac{1}{2}|\nabla u|^2$. Similarly, the analysis in this chapter 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.

In Chapter 5, we proved the optimal-order convergence of the Stokes equations on an evolving domain with a moving boundary by employing the ALE-FEM along the trajectories of the evolving mesh. The error of the semidiscrete ALE method is shown to be $O(h^{r+1})$ for velocity in $L^\infty(0, T; L^2)$ norm and $O(h^r)$ for pressure in $L^2(0, T; L^2)$ norm by employing the Taylor–Hood finite elements of degree $r \geqslant 2$, using Nitsche's duality argument adapted to an evolving mesh, by proving that the material derivative and the Stokes–Ritz projection commute up to terms which have optimal-order convergence in the $L^2$ norm.

In Chapter 6, we have established the optimal convergence rates of fluid velocity, structural displacement and ALE mesh motion in $L^\infty(0, T; H^1)$ norm, as well as of pressure in $L^\infty(0, T; L^2)$ norm, for an ALE-based interface tracking, monolithic mixed FEM for a dynamic fluid-structure interaction problem with solution-driven moving interfaces, where we have developed a fully discrete scheme by means of the backward Euler scheme and initial correction terms within the frames of Eulerian-Lagrangian description and monolithic ALE-FEM.

Without adding the initial correction term, the convergence of fully discrete ALE interface tracking FEMs with the semi-implicit Euler scheme can still be proved for the fluid ve-

locity and structural displacement with suboptimal-order convergence in the $L^\infty(0,T;H^1)$ norm in the framework of this chapter. The proof of optimal-order convergence in the $L^\infty(0,T;L^2)$ norm and the numerical analysis of the ALE-based interface tracking for affine finite element approximations still remain open and challenging problems.

In the future, we plan to focus on the two aforementioned open problems. For the first problem, a promising direction is to design a specialized Ritz projection, potentially with time-dependent basis functions, and to rigorously establish its stability properties. For the second problem, a key challenge lies in achieving a sufficiently small $W^{1,\infty}$ error estimate for the displacement. Currently, when using standard affine finite elements, we can only prove boundedness due to the limited convergence rate. Addressing this issue may require the development of an entirely new analytical framework.

# Bibliography


[1] R. A. Adams and J. J. F. Fournier. *Sobolev spaces*, volume 140 of *Pure and Applied Mathematics (Amsterdam)*. Elsevier/Academic Press, Amsterdam, second edition, 2003.

[2] M. Al-Maskari and S. Karaa. Numerical approximation of semilinear subdiffusion equations with nonsmooth initial data. *SIAM J. Numer. Anal.*, 57:1524–1544, 2019.

[3] G. Allaire, C. Dapogny, and F. Jouve. *Shape and topology optimization, in Geometric Partial Differential Equations, Part II.*, volume 22 of *Handb. Numer. Anal.* Elsevier, Amsterdam, 2021.

[4] W. Arendt, C. J. Batty, M. Hieber, and F. Neubrander. *Vector-valued Laplace Transforms and Cauchy Problems*, volume 96. Springer Science & Business Media, 2011.

[5] D. Arnold, F. Brezzi, and M. Fortin. A stable finite element for the stokes equations. *Calcolo*, 21:337–344, 1984.

[6] S. Badia and R. Codina. Analysis of a stabilized finite element approximation of the transient convection-diffusion equation using an ale framework. *SIAM J. Numer. Anal.*, 44(5):2159–2197, 2006.

[7] S. Badia and R. Codina. Convergence analysis of the FEM approximation of the first order projection method for incompressible flows with and without the inf-sup condition. *Numer. Math.*, 107(4):533–557, 2007.

[8] L. Banjai and M. López-Fernández. Efficient high order algorithms for fractional integrals and fractional differential equations. *Numer. Math.*, 141:289–317, 2019.

[9] J. W. Barrett, K. Deckelnick, and V. Styles. Numerical analysis for a system coupling curve evolution to reaction diffusion on the curve. *SIAM J. Numer. Anal.*, 55(2):1080–1100, 2017.

[10] R. Bermejo, P. Galán del Sastre, and L. Saavedra. A second order in time modified Lagrange-Galerkin finite element method for the incompressible Navier-Stokes equations. *SIAM J. Numer. Anal.*, 50(6):3084–3109, 2012.

[11] D. Boffi, F. Brezzi, M. Fortin, et al. *Mixed finite element methods and applications*, volume 44. Springer, 2013.

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

[13] A. Bonito, I. Kyza, and R. H. Nochetto. Time-discrete higher order ALE formulations: A priori error analysis. *Numer. Math.*, 125(2):225–257, 2013.

[14] A. Bonito, I. Kyza, and R. H. Nochetto. Time-discrete higher-order ALE formulations: stability. *SIAM J. Numer. Anal.*, 51(1):577–604, 2013.

[15] S. Brenner and R. Scott. *The mathematical theory of finite element methods*, volume 15. Springer Science & Business Media, 2007.

[16] M. Burger. A framework for the construction of level set methods for shape optimization and reconstruction. *Interfaces Free boundaries*, 5(3):301–329, 2003.

[17] 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. Eng.*, 317:598–618, 2017.

[18] E. Burman, S. Frei, and A. Massing. Eulerian time-stepping schemes for the non-stationary stokes equations on time-dependent domains. *Numer. Math.*, pages 1–56, 2022.

[19] C. Canuto, M. Hussaini, A. Quarteroni, and T. Zang. *Spectral Methods: Fundamentals in Single Domains*, volume 23. Springer, 01 2006.

[20] S. K. Chakrabarti, editor. *Numerical Models in Fluid Structure Interaction, Advances in Fluid Mechanics*, volume 42. WIT Press, 2005.

[21] F. Chen, Q. Xu, and J. S. Hesthaven. A multi-domain spectral method for time-fractional differential equations. *J. Comput. Phys.*, 293:157–172, 2015.

[22] L. Chen, Z. Mao, and H. Li. Jacobi-Galerkin spectral method for eigenvalue problems of Riesz fractional differential equations. *arXiv preprint arXiv:1803.03556*, 2018.

[23] S. Chen and J. Shen. Log orthogonal functions: approximation properties and applications. *IMA J. Numer. Anal.*, 42(1):712–743, 2022.

[24] S. Chen, J. Shen, and L. Wang. Generalized Jacobi functions and their applications to fractional differential equations. *Math. Comp.*, 85:1603–1638, 2016.

[25] S. Chen, J. Shen, Z. Zhang, and Z. Zhou. A spectrally accurate approximation to subdiffusion equations using the log orthogonal functions. *SIAM J. Sci. Comput.*, 42:A849–A877, 2020.

[26] D. Chenais and E. Zuazua. Finite-element approximation of 2D elliptic optimal design. *J. Math. Pures Appl.*, 85(2):225–249, 2006.

[27] T. Chu, J. Wang, N. Wang, and Z. Zhang. Optimal-Order Convergence of a Two-Step BDF Method for the Navier-Stokes Equations with $H^1$ Initial Data. *J. Sci. Comput.*, 96(2):Paper No. 62, 2023.

[28] P. G. Ciarlet. Mathematical elasticity. volume ii: Theory of plates, volume 27 of. *Studies in Mathematics and its Applications*, 1997.

[29] M. Crouzeix and V. Thomée. On the discretization in time of semilinear parabolic equations with nonsmooth initial data. *Math. Comp.*, 49(180):359–377, 1987.

[30] F. De Gournay. Velocity extension for the level-set method and multiple eigenvalues in shape optimization. *SIAM J. Control Optim.*, 45(1):343–367, 2006.

[31] L. Dedè, M. J. Borden, and T. J. Hughes. Isogeometric analysis for topology optimization with a phase field model. *Arch. Comput. Methods Eng.*, 19:427–465, 2012.

[32] M. C. Delfour and J.-P. Zolésio. *Shapes and geometries: metrics, analysis, differential calculus, and optimization*. SIAM, 2011.

[33] A. Djurdjevac, C. Gräser, and P. J. Herbert. An evolving space framework for oseen equations on a moving domain. *ESAIM: Math. Model. Numer. Anal.*, 57(5):3113–3138, 2023.

[34] G. Dogan, P. Morin, R. H. Nochetto, and M. Verani. Discrete gradient flows for shape optimization and applications. *Comput. Methods Appl. Mech. Eng.*, 196(37-40):3898–3914, 2007.

[35] E. H. Dowell and K. C. Hall. Modeling of fluid-structure interaction. *Annu. Rev. Fluid Mech.*, 33:445–490, 2001.

[36] G. Dziuk. Convergence of a semi-discrete scheme for the curve shortening flow. *Math. Models Methods Appl. Sci.*, 4(04):589–606, 1994.

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

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

[39] C. M. Elliott and H. Fritz. On approximations of the curve shortening flow and of the mean curvature flow based on the Deturck trick. *IMA J. Numer. Anal.*, 37(2):543–603, 2017.

[40] C. M. Elliott, H. Garcke, and B. Kovács. Numerical analysis for the interaction of mean curvature flow and diffusion on closed surfaces. *Numer. Math.*, 151(4):873–925, 2022.

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

[42] C. M. Elliott and V. Styles. An ALE ESFEM for solving PDEs on evolving surfaces. *Milan J. Math.*, 80(2):469–501, 2012.

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

[44] K. Eppler and H. Harbrecht. Coupling of FEM and BEM in shape optimization. *Numer. Math.*, 104:47–68, 2006.

[45] K. Eppler, H. Harbrecht, and R. Schneider. On convergence in elliptic shape optimization. *SIAM J. Control Optim.*, 46(1):61–83, 2007.

[46] R. Farwig and H. Sohr. Generalized resolvent estimates for the Stokes system in bounded and unbounded domains. *J. Math. Soc. Japan*, 46(4):607–643, 1994.

[47] P. J. Ferreira de Sousa and J. C. Pereira. Reynolds number dependence of two-dimensional laminar co-rotating vortex merging. *Theor. Comput. Fluid Dyn.*, 19(1):65–75, 2005.

[48] L. Formaggia and F. Nobile. A stability analysis for the arbitrary Lagrangian–Eulerian formulation with finite elements. *East-West J. Numer. Math.*, 7:105–131, 1999.

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

[50] I. Fumagalli, N. Parolini, and M. Verani. Shape optimization for Stokes flows: a finite element convergence analysis. *ESAIM: Math. Model. Numer. Anal.*, 49(4):921–951, 2015.

[51] I. Gallagher. Critical function spaces for the well-posedness of the Navier-Stokes initial value problem. In *Handbook of mathematical analysis in mechanics of viscous fluids*, pages 647–685. Springer, Cham, 2018.

[52] B. García-Archilla, V. John, and J. Novo. Symmetric pressure stabilization for equal-order finite element approximations to the time-dependent Navier-Stokes equations. *IMA J. Numer. Anal.*, 41(2):1093–1129, 2021.

[53] L. Gastaldi. A priori error estimates for the arbitrary Lagrangian–Eulerian formulation with finite elements. *East-West J. Numer. Math.*, 9:123–156, 2001.

[54] E. S. Gawlik and A. J. Lew. Unified analysis of finite element methods for problems with moving boundaries. *SIAM J. Numer. Anal.*, 53(6):2822–2846, 2015.

[55] V. Girault and P.-A. Raviart. *Finite element methods for Navier-Stokes equations: theory and algorithms*, volume 5. Springer Science & Business Media, 2012.

[56] W. Gong, B. Li, and Q. Rao. Convergent evolving finite element approximations of boundary evolution under shape gradient flow. *IMA J. Numer. Anal.*, 44(5):2667–2697, 2024.

[57] W. Gong, J. Li, and S. Zhu. Improved discrete boundary type shape gradients for pde-constrained shape optimization. *SIAM J. Sci. Comput.*, 44(4):A2464–A2505, 2022.

[58] N. Guglielmi, M. López-Fernández, and M. Manucci. Pseudospectral roaming contour integral methods for convection-diffusion equations. *J. Sci. Comput.*, 89:22, Sep 2021.

[59] R. Guo. Solving parabolic moving interface problems with dynamical immersed spaces on unfitted meshes: fully discrete analysis. *SIAM J. Numer. Anal.*, 59(2):797–828, 2021.

[60] R. Guo, T. Lin, and Y. Lin. Recovering elastic inclusions by shape optimization methods with immersed finite elements. *J. Comput. Phys.*, 404:109123, 2020.

[61] E. Hairer and G. Wanner. *Solving ordinary differential equations. II*, volume 14 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, 2010. Stiff and differential-algebraic problems, Second revised edition, paperback.

[62] J. Haslinger and R. A. Mäkinen. *Introduction to shape optimization: theory, approximation, and computation*. SIAM, 2003.

[63] J. Haslinger and P. Neittaanmäki. Finite element approximation for optimal shape, material and topology design. *(No Title)*, 1996.

[64] Y. He. The Euler implicit/explicit scheme for the 2D time-dependent Navier-Stokes equations with smooth or non-smooth initial data. *Math. Comp.*, 77(264):2097–2124, 2008.

[65] Y. He. Stability and error analysis for spectral Galerkin method for the Navier-Stokes equations with $L^2$ initial data. *Numer. Methods Partial Differ. Equ.*, 24(1):79–103, 2008.

[66] Y. He. The Crank-Nicolson/Adams-Bashforth scheme for the time-dependent Navier-Stokes equations with nonsmooth initial data. *Numer. Methods Partial Differ. Equ.*, 28(1):155–187, 2012.

[67] Y. He and W. Sun. Stability and convergence of the Crank-Nicolson/Adams-Bashforth scheme for the time-dependent Navier-Stokes equations. *SIAM J. Numer. Anal.*, 45(2):837–869, 2007.

[68] Y. He and W. Sun. Stabilized finite element method based on the Crank-Nicolson extrapolation scheme for the time-dependent Navier-Stokes equations. *Math. Comp.*, 76(257):115–136, 2007.

[69] A. Henrot and M. Pierre. *Shape variation and optimization*. 2018.

[70] J. G. Heywood and R. Rannacher. Finite-element approximation of the nonstationary Navier-Stokes problem. IV. Error analysis for second-order time discretization. *SIAM J. Numer. Anal.*, 27(2):353–384, 1990.

[71] A. T. Hill and E. Süli. Approximation of the global attractor for the incompressible Navier-Stokes equations. *IMA J. Numer. Anal.*, 20(4):633–667, 2000.

[72] R. Hiptmair, A. Paganini, and S. Sargheini. Comparison of approximate shape gradients. *BIT Numer. Math.*, 55:459–485, 2015.

[73] C. Hirth, A. Amsden, and J. Cook. An arbitrary Lagrangian-Eulerian computing method for all flow speeds. *J. Comput. Phys.*, 14(3):227 – 253, 1974.

[74] D. Hou, M. T. Hasan, and C. Xu. Müntz spectral methods for the time-fractional diffusion equation. *Comput. Meth. Appl. Math.*, 18(1):43–62, 2018.

[75] D. Hou and C. Xu. A fractional spectral method with applications to some singular problems. *Adv. Comput. Math.*, 43(5):911–944, 2017.

[76] J. Hu and B. Li. Evolving finite element methods with an artificial tangential velocity for mean curvature flow and willmore flow. *Numer. Math.*, 152(1):127–181, 2022.

[77] F. Huang and J. Shen. Stability and error analysis of a second-order consistent splitting scheme for the Navier-Stokes equations. *SIAM J. Numer. Anal.*, 61(5):2408–2433, 2023.

[78] A. Huerta and W. Liu. Viscous flow structure interaction. *J. Pressure Vessel Technol.*, 110(1):15 – 21, 1988.

[79] T. Hughes, W. Liu, and T. Zimmermann. Lagrangian-Eulerian finite element formulation for incompressible viscous flows. *Comput. Methods Appl. Mech. Eng.*, 29(3):329 – 349, 1981.

[80] R. Ingram. A new linearly extrapolated Crank-Nicolson time-stepping scheme for the Navier-Stokes equations. *Math. Comp.*, 82(284):1953–1973, 2013.

[81] S. Jan and J. Zolesio. Introduction to shape optimization: shape sensitivity analysis, 1992.

[82] B. Jin, R. Lazarov, and Z. Zhou. An analysis of the L1 scheme for the subdiffusion equation with nonsmooth data. *IMA J. Numer. Anal.*, 36:197–221, 2015.

[83] B. Jin, B. Li, and Z. Zhou. Correction of high-order BDF convolution quadrature for fractional evolution equations. *SIAM J. Sci. Comput.*, 39:A3129–A3152, 2017.

[84] B. Jin, B. Li, and Z. Zhou. Numerical analysis of nonlinear subdiffusion equations. *SIAM J. Numer. Anal.*, 56:1–23, 2018.

[85] B. Jin, B. Li, and Z. Zhou. Subdiffusion with time-dependent coefficients: improved regularity and second-order time stepping. *Numer. Math.*, 145(4):883–913, 2020.

[86] C. Josserand and M. Rossi. The merging of two co-rotating vortices: a numerical study. *Eur. J. Mech. B Fluids*, 26(6):779–794, 2007.

[87] B. Kiniger and B. Vexler. A priori error estimates for finite elementdiscretizations of a shape optimization problem. *ESAIM: Math. Model. Numer. Anal.*, 47(6):1733–1763, 2013.

[88] V. Komkov, K. K. Choi, and E. J. Haug. *Design sensitivity analysis of structural systems*, volume 177. Academic press, 1986.

[89] N. Kopteva. Error analysis for time-fractional semilinear parabolic equations using upper and lower solutions. *SIAM J. Numer. Anal.*, 58(4):2212–2234, 2020.

[90] N. Kopteva. Error analysis of an L2-type method on graded meshes for a fractional-order parabolic problem. *Math. Comp.*, 90:19–40, 2021.

[91] N. Kopteva. Maximum principle for time-fractional parabolic equations with a reaction coefficient of arbitrary sign. *Appl. Math. Lett.*, 132:108209, 2022.

[92] N. Kopteva and X. Meng. Error analysis for a fractional-derivative parabolic problem on quasi-graded meshes using barrier functions. *SIAM J. Numer. Anal.*, 58:1217–1238, 2020.

[93] B. Kovács, B. Li, and C. Lubich. A convergent evolving finite element algorithm for mean curvature flow of closed surfaces. *Numer. Math.*, 143:797–853, 2019.

[94] B. Kovács, B. Li, and C. Lubich. A convergent evolving finite element algorithm for willmore flow of closed surfaces. *Numer. Math.*, 149(3):595–643, 2021.

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

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

[97] A. Labovsky, W. J. Layton, C. C. Manica, M. Neda, and L. G. Rebholz. The stabilized extrapolated trapezoidal finite-element method for the Navier-Stokes equations. *Comput. Methods Appl. Mech. Eng.*, 198(9-12):958–974, 2009.

[98] R. Lan and P. Sun. A novel arbitrary lagrangian–eulerian finite element method for a parabolic/mixed parabolic moving interface problem. *J. Comput. Appl. Math.*, 383:113125, 2021.

[99] H. Lee and S. Xu. Finite element error estimation for quasi-Newtonian fluid-structure interaction problems. *Appl. Math. Comput.*, 274:93–105, 2016.

[100] H. Lee and S. Xu. Fully discrete error estimation for a quasi-Newtonian fluid-structure interaction problem. *Comput. Math. Appl.*, 71:2373–2388, 2016.

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

[102] C. Lehrenfeld and M. Olshanskii. An eulerian finite element method for pdes in time-dependent domains. *ESAIM: Math. Model. Numer. Anal.*, 53(2):585–614, 2019.

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

[104] B. Li. Analyticity, maximal regularity and maximum-norm stability of semi-discrete finite element solutions of parabolic equations in nonconvex polyhedra. *Math. Comp.*, 88:1–44, 2019.

[105] B. Li. Maximal regularity of multistep fully discrete finite element methods for parabolic equations. *IMA J. Numer. Anal.*, 42(2):1700–1734, 2022.

[106] B. Li and S. Ma. Exponential convolution quadrature for nonlinear subdiffusion equations with nonsmooth initial data. *SIAM J. Numer. Anal.*, 60(2):503–528, 2022.

[107] B. Li, S. Ma, and W. Qiu. Optimal convergence of the arbitrary Lagrangian–Eulerian interface tracking method for two-phase Navier–Stokes flow without surface tension. *IMA J. Numer. Anal.*, page draf003, 03 2025.

[108] B. Li, S. Ma, and K. Schratz. A semi-implicit exponential low-regularity integrator for the Navier-Stokes equations. *SIAM J. Numer. Anal.*, 60(4):2273–2292, 2022.

[109] B. Li, S. Ma, and W. Sun. Optimal analysis of finite element methods for the stochastic Stokes equations. *Math. Comp.*, 94(352):551–583, 2025.

[110] B. Li, S. Ma, and Y. Ueda. Analysis of fully discrete finite element methods for 2D Navier-Stokes equations with critical initial data. *ESAIM: Math. Model. Numer. Anal.*, 56(6):2105–2139, 2022.

[111] B. Li, S. Ma, and N. Wang. Second-order convergence of the linearly extrapolated Crank-Nicolson method for the Navier-Stokes equations with $H^1$ initial data. *J. Sci. Comput.*, 88(3):Paper No. 70, 20, 2021.

[112] B. Li, W. Qiu, and Z. Yang. A convergent post-processed discontinuous galerkin method for incompressible flow with variable density. *J. Sci. Comput.*, 91(1):2, 2022.

[113] B. Li, Y. Xia, and Z. Yang. Optimal convergence of arbitrary lagrangian–eulerian isoparametric finite element methods for parabolic equations in an evolving domain. *IMA J. Numer. Anal.*, 43(1):501–534, 2023.

[114] B. Li, Y. Xia, and Z. Yang. Optimal convergence of arbitrary lagrangian–eulerian isoparametric finite element methods for parabolic equations in an evolving domain. *IMA J. Numer. Anal.*, 43:501–534, 2023.

[115] C. Li, R. I. Borja, and R. A. Regueiro. Dynamics of porous media at finite strain. *Comput. Methods Appl. Mech. Eng.*, 193:3837–3870, 2004.

[116] D. Li. Effective maximum principles for spectral methods. *Ann. Appl. Math.*, 2021:131–290, 2021.

[117] D. Li, H.-L. Liao, W. Sun, J. Wang, and J. Zhang. Analysis of L1-Galerkin FEMs for time-fractional nonlinear parabolic problems. *Comm. Comput. Phys.*, 24:86–103, 2018.

[118] H.-l. Liao, W. McLean, and J. Zhang. A discrete Grönwall inequality with applications to numerical schemes for subdiffusion problems. *SIAM J. Numer. Anal.*, 57:218–237, 2019.

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

[120] V. R. Loon. *A 3D method for modelling the fluid-structure interaction of heart valves*. Eindhoven: Technische Universiteit Eindhoven, DOI: 10.6100/IR594604, 2005.

[121] M. López-Fernández. A quadrature based method for evaluating exponential-type functions for exponential methods. *BIT Numer. Math.*, 50:631–655, 2010.

[122] A. Lozovskiy, M. A. Olshanskii, and Y. V. Vassilevski. A quasi-Lagrangian finite element method for the Navier–Stokes equations in a time-dependent domain. *Comput. Methods Appl. Mech. Eng.*, 333:55–73, 2018.

[123] C. Lubich. Convolution quadrature and discretized operational calculus. I. *Numer. Math.*, 52:129–145, 1988.

[124] C. Lubich. Convolution quadrature and discretized operational calculus. II. *Numer. Math.*, 52:413–425, 1988.

[125] C. Lubich. Convolution quadrature revisited. *BIT*, 44:503–514, 2004.

[126] C. Lubich and A. Ostermann. Runge-Kutta methods for parabolic equations and convolution quadrature. *Math. Comp.*, 60:105–131, 1993.

[127] C. Lubich and A. Ostermann. Runge-Kutta time discretization of reaction-diffusion and Navier-Stokes equations: nonsmooth-data error estimates and applications to long-time behaviour. volume 22, pages 279–292. 1996. Special issue celebrating the centenary of Runge-Kutta methods.

[128] C. Lubich, I. H. Sloan, and V. Thomée. Nonsmooth data error estimates for approximations of an evolution equation with a positive-type memory term. *Math. Comp.*, 65:1–17, 1996.

[129] Y. Luchko and M. Yamamoto. On the maximum principle for a time-fractional diffusion equation. *Fract Calc Appl Anal.*, 20:1131–1145, 2017.

[130] C. Ma, T. Tian, and W. Zheng. High-order unfitted characteristic finite element methods for moving interface problem of oseen equations. *J. Comput. Appl. Math.*, 425:115028, 2023.

[131] C. Ma, Q. Zhang, and W. Zheng. A fourth-order unfitted characteristic finite element method for solving the advection-diffusion equation on time-varying domains. *SIAM J. Numer. Anal.*, 60(4):2203–2224, 2022.

[132] X. Mao, S. J. Sherwin, and H. M. Blackburn. Non-normal dynamics of time-evolving co-rotating vortex pairs. *J. Fluid Mech.*, 701:430–459, 2012.

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

[134] W. McLean and K. Mustapha. Time-stepping error bounds for fractional diffusion problems with non-smooth initial data. *J. Comput. Phys.*, 293:201–217, 2015.

[135] B. Mohammadi and O. Pironneau. *Applied shape optimization for fluids*. OUP Oxford, 2009.

[136] P. Morin, R. H. Nochetto, M. S. Pauletti, and M. Verani. Adaptive finite element method for shape optimization. *ESAIM: COCV*, 18(4):1122–1149, 2012.

[137] K. Mustapha, B. Abdallah, and K. M. Furati. A discontinuous Petrov–Galerkin method for time-fractional diffusion equations. *SIAM J. Numer. Anal.*, 52:2512–2529, 2014.

[138] K. Mustapha and W. McLean. Uniform convergence for a discontinuous Galerkin, time-stepping method applied to a fractional diffusion equation. *IMA J. Numer. Anal.*, 32:906–925, 2011.

[139] M. Neilan and M. Olshanskii. An Eulerian finite element method for the linearized Navier–Stokes problem in an evolving domain. *IMA J. Numer. Anal.*, 44(6):3234–3258, 2024.

[140] C. Nitikitpaiboon and K. Bathe. An arbitrary Lagrangian-Eulerian velocity potential formulation for fluid-structure interaction. *Comput. Struct.*, 47(4):871 – 891, 1993.

[141] F. Nobile. *Numerical approximation of fluid-structure interaction problems with application to haemodynamics*. PhD thesis, EPFL, 2001.

[142] F. Nobile. *Numerical approximation of fluid-structure interaction problems with application to haemodynamics*. PhD thesis, EPFL, 2001.

[143] F. Nobile and L. Formaggia. A stability analysis for the arbitrary Lagrangian-Eulerian formulation with finite elements. *East-West J. Numer. Math.*, 7(2):105–132, 1999.

[144] P. Orlandi. Two-dimensional and three-dimensional direct numerical simulation of co-rotating vortices. *Phys. Fluids*, 19(1), 2007.

[145] A. Ostermann, F. Rousset, and K. Schratz. Error estimates of a Fourier integrator for the cubic Schrödinger equation at low regularity. *Found. Comput. Math.*, 21(3):725–765, 2021.

[146] P. Pozzi and B. Stinner. Convergence of a scheme for an elastic flow with tangential mesh movement. *ESAIM: Math. Model. Numer. Anal.*, 57(2):445–466, 2023.

[147] Q. Rao, J. Wang, and Y. Xie. Optimal convergence of arbitrary lagrangian–eulerian finite element methods for the stokes equations in an evolving domain. *IMA J. Numer. Anal.*, page drae097, 2025.

[148] L. G. Rebholz. An energy- and helicity-conserving finite element scheme for the Navier-Stokes equations. *SIAM J. Numer. Anal.*, 45(4):1622–1638, 2007.

[149] T. Richter and T. Wick. Finite elements for fluid–structure interaction in ALE and fully Eulerian coordinates. *Comput. Methods Appl. Mech. Engrg.*, 199(41):2633–2642, 2010.

[150] F. Rousset and K. Schratz. A general framework of low regularity integrators. *SIAM J. Numer. Anal.*, 59(3):1735–1768, 2021.

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

[152] K. Schratz, Y. Wang, and X. Zhao. Low-regularity integrators for nonlinear Dirac equations. *Math. Comp.*, 90(327):189–214, 2021.

[153] V. Schulz and M. Siebenborn. Computational comparison of surface metrics for pde constrained shape optimization. *Comput. Methods Appl. Math.*, 16(3):485–496, 2016.

[154] V. H. Schulz, M. Siebenborn, and K. Welker. Efficient pde constrained shape optimization based on steklov–poincaré-type metrics. *SIAM J. Optim.*, 26(4):2800–2819, 2016.

[155] L. R. Scott and S. Zhang. Finite element interpolation of nonsmooth functions satisfying boundary conditions. *Math. Comp.*, 54(190):483–493, 1990.

[156] J. H. Seo, V. Vedula, T. Abraham, and R. Mittal. Multiphysics computational models for cardiac flow and virtual cardiography. *Int. J. Numer. Meth. Biomed. Engng.*, 29:850–869, 2013.

[157] J. Shen. On error estimates of projection methods for Navier-Stokes equations: first-order schemes. *SIAM J. Numer. Anal.*, 29(1):57–77, 1992.

[158] J. Shen. On error estimates of some higher order projection and penalty-projection methods for Navier-Stokes equations. *Numer. Math.*, 62(1):49–73, 1992.

[159] J. Shen, T. Tang, and L.-L. Wang. *Spectral Methods: Algorithms, Analysis and Applications*, volume 41. Springer Science & Business Media, 2011.

[160] M. Souli and D. Benson, editors. *Arbitrary Lagrangian Eulerian and Fluid-Structure Interaction: Numerical Simulation.* John Wiley & Sons, 2010.

[161] E. M. Stein. *Singular integrals and differentiability properties of functions.* Number 30. Princeton university press, 1970.

[162] M. Stynes, E. O'Riordan, and J. L. Gracia. Error analysis of a finite difference method on graded meshes for a time-fractional diffusion equation. *SIAM J. Numer. Anal.*, 55:1057–1079, 2017.

[163] K. Takizawa, B. Henicke, T. Tezduyar, M. Hsu, and Y. Bazilevs. Stabilized space-time computation of wind-turbine rotor aerodynamics. *Comput. Mech.*, 48:333–344, 2011.

[164] R. Temam. *Navier–Stokes equations: theory and numerical analysis*, volume 343. American Mathematical Society, 2024.

[165] W. Themistoclakis. Uniform approximation on [-1, 1] via discrete de la Vallée Poussin means. *Numer. Algorithms*, 60, 08 2012.

[166] F. Tone. Error analysis for a second order scheme for the Navier-Stokes equations. *Appl. Numer. Math.*, 50(1):93–119, 2004.

[167] S. Turek and J. Hron. *Proposal for numerical benchmarking of fluid-structure interaction between an elastic object and laminar incompressible flow.* Springer, 2006.

[168] S. Turek, J. Hron, M. Mádlík, M. Razzaq, H. Wobker, and J. F. Acker. Numerical simulation and benchmarking of a monolithic multigrid solver for fluid-structure interaction problems with application to hemodynamics. In H.-J. Bungartz, M. Mehl, and M. Schäfer, editors, *Fluid Structure Interaction II*, volume 73 of *Lecture Notes in Computational Science and Engineering*, pages 193 – 220. Springer Berlin Heidelberg, Berlin, Heidelberg, 2010.

[169] R. Verfürth. Error estimates for a mixed finite element approximation of the Stokes equations. *RAIRO Anal. Numer.*, 18(2):175–182, 1984.

[170] H. von Wahl, T. Richter, and C. Lehrenfeld. An unfitted Eulerian finite element method for the time-dependent Stokes problem on moving domains. *IMA J. Numer. Anal.*, 42(3):2505–2544, 2022.

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

[172] C. Wang, S. Xia, X. Wang, and X. Qian. Isogeometric shape optimization on triangulations. *Comput. Methods Appl. Mech. Eng.*, 331:585–622, 2018.

[173] K. Wang and Z. Zhou. High-order time stepping schemes for semilinear subdiffusion equations. *SIAM J. Numer. Anal.*, 58:3226–3250, 2020.

[174] Y. Wu and X. Zhao. Optimal convergence of a second-order low-regularity integrator for the KdV equation. *IMA J. Numer. Anal.*, 42(4):3499–3528, 2022.

[175] Y. Xing and Y. Yan. A higher order numerical method for time fractional partial differential equations with nonsmooth data. *J. Comput. Phys.*, 357:305–323, 2018.

[176] K. Yang, P. Sun, L. Wang, J. Xu, and L. Zhang. Modeling and simulation for fluid-rotating structure interaction. *Comput. Methods Appl. Mech. Eng.*, 311:788–814, 2016.

[177] C. Ye and J. Cui. Convergence of Dziuk's fully discrete linearly implicit scheme for curve shortening flow. *SIAM J. Numer. Anal.*, 59(6):2823–2842, 2021.

[178] R. Zacher. Boundedness of weak solutions to evolutionary partial integro-differential equations with discontinuous coefficients. *J. Math. Anal. Appl.*, 348:137–149, 2008.

[179] M. Zayernouri and G. E. Karniadakis. Fractional Sturm-Liouville eigen-problems: Theory and numerical approximation. *J. Comput. Phys.*, 252:495–517, 2013.

[180] M. Zayernouri and G. E. Karniadakis. Fractional spectral collocation method. *SIAM J. Sci.Comput.*, 36:A40–A62, 2014.

[181] X. Zhao and Z. Zhang. Superconvergence points of fractional spectral interpolation. *SIAM J. Sci. Comput.*, 38:A598–A613, 2016.

[182] S. Zhu and Z. Gao. Convergence analysis of mixed finite element approximations to shape gradients in the stokes equation. *Comput. Methods Appl. Mech. Eng.*, 343:127–150, 2019.