STABILITY AND CONVERGENCE OF FINITE ELEMENT METHODS IN COMPLEX GEOMETRIES

YUPEI XIE

PhD

The Hong Kong Polytechnic University

2025

The Hong Kong Polytechnic University

Department of Applied Mathematics

Stability and convergence of finite element methods in complex geometries

Yupei Xie

A thesis submitted in partial fulfilment of the requirements for the degree of

Doctor of Philosophy

2025 May

CERTIFICATE OF ORIGINALITY

I hereby declare that this thesis is my own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person, nor material which has been accepted for the award of any other degree or diploma at a university or other institute of higher learning, except where due acknowledgment has been made in the text.

____________________________(Signed)

Yupei Xie (Name of student)

Abstract

This thesis investigates the stability and error estimates of finite element methods (FEM) for partial differential equations (PDEs) in complex and evolving geometries. It aims to advance the mathematical understanding and numerical analysis of FEM in three challenging settings: time-dependent domains, fluid–structure interaction (FSI), and the maximum-norm stability of isoparametric FEM.

The first part addresses the Arbitrary Lagrangian–Eulerian (ALE) FEM for the Stokes equations on moving domains. By establishing optimal L^2 error bounds of order $O(h^{r+1})$ for the velocity and $O(h^r)$ for the pressure, this work closes a long-standing gap in the literature, where only sub-optimal convergence rates were previously available. A novel duality argument for H^{-1}-error estimate of pressure is developed to obtain optimal estimates for the commutator between the material derivative and the Stokes–Ritz projection.

The second part develops and analyzes a fully-discrete loosely coupled scheme for fluid thin-structure interaction problems. A key innovation is the construction and analysis of a coupled non-stationary Ritz projection that satisfies the kinematic interface condition and enables the derivation of optimal L^2 error estimates. The proposed loosely coupled scheme incorporates stabilization terms to ensure unconditional energy stability and is rigorously shown to achieve optimal convergence in the L^2 norm.

The third part focuses on maximum norm stability of isoparametric FEM in curvilinear polyhedral domains where the geometry cannot be exactly triangulated. This includes the proof of a weak discrete maximum principle and the derivation of optimal maximum-norm error estimates for elliptic equations. For parabolic problems, the thesis establishes the analyticity and maximal L^p-regularity of the semi-discrete FEM and further proves optimal maximum-norm error estimates.

Publications arising from the thesis

The following publications are based on the research presented in this thesis. All authors contributed equally to the respective works.

1. 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 Journal of Numerical Analysis*, drae097, 2025. doi: 10.1093/imanum/drae097.

2. B. Li, W. Sun, Y. Xie, and W. Yu. *Optimal L^2 error analysis of a loosely coupled finite element scheme for thin-structure interactions. SIAM Journal on Numerical Analysis*, 62(4):1782–1813, 2024. doi: 10.1137/23M1578401.

3. B. Li, W. Qiu, Y. Xie, and W. Yu. *Weak discrete maximum principle of isoparametric finite element methods in curvilinear polyhedra. Mathematics of Computation*, 93(345):1–34, 2024. doi: 10.1090/mcom/3876.

4. W. Qiu and Y. Xie. *Stability, analyticity and maximal regularity of semi-discrete isoparametric finite element solutions of parabolic equations in curvilinear polyhedra. Submitted for publication.*

Acknowledgements

First and foremost, I would like to express my deepest gratitude to my supervisor, Prof. Buyang Li, for his exceptional guidance, unwavering patience, and original perspective throughout my PhD experience. I am sincerely thankful for the opportunity he has given me to pursue research under his supervision, as well as for his kindness and tolerance during times of difficulty. His remarkable insight into research and genuine passion for mathematics have been a continual source of inspiration. I am especially grateful for his open-mindedness and the many thoughtful discussions that have shaped the direction and depth of this thesis.

I would also like to thank Prof. Weiwei Sun, Prof. Jilu Wang, and Prof. Weifeng Qiu for their valuable collaboration and insightful suggestions. Their generous support and constructive feedback have greatly enriched the quality of my research.

My heartfelt thanks go to my friends and colleagues at The Hong Kong Polytechnic University and other institutions for their warm support and stimulating discussions. Their support, encouragement, and intellectual exchange have contributed meaningfully to my academic development. Lastly, I am deeply thankful to my family and friends for their unwavering love, trust, and encouragement.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

This thesis is organized into three main parts, each addressing a fundamental challenge in the numerical analysis of finite element methods (FEM) for partial differential equations in complex geometries. The first part (Chapter 2) focuses on the optimal L^2 error analysis of ALE finite element methods for the Stokes equations on time-dependent domains. The second part (Chapter 3) presents an optimal L^2 error analysis of loosely coupled schemes for fluid–structure interaction (FSI) problems. The third part (Chapters 4 and 5) examines the maximum-norm stability and error estimates of isoparametric FEM on curvilinear polyhedral domains, including the establishment of a weak discrete maximum principle for elliptic problems and the discrete analyticity and maximal regularity for parabolic equations. Together, these contributions provide new theoretical insights and practical methodologies that enhance the accuracy and reliability of FEM-based simulations, particularly in evolving domains and coupled physical systems.

Chapter 2 of this thesis is devoted to establishing optimal L^2-error estimates for the arbitrary Lagrangian–Eulerian (ALE) finite element method for the Stokes equations on time-dependent domains. We consider the following model problem:

$$\partial_t u - \Delta u + \nabla p = f \quad \text{in } \bigcup_{t\in(0,T]} \Omega(t)\times\{t\}, \tag{1.0.1a}$$

$$\nabla\cdot u = 0 \quad \text{in } \bigcup_{t\in(0,T]} \Omega(t)\times\{t\}, \tag{1.0.1b}$$

$$u = w \quad \text{on } \bigcup_{t\in(0,T]} \partial\Omega(t)\times\{t\}, \tag{1.0.1c}$$

$$u = u_0 \quad \text{on } \Omega^0 := \Omega(0), \tag{1.0.1d}$$

where the domain $\Omega(t)$ evolves with a smooth boundary $\Gamma(t) := \partial\Omega(t)$, driven by a given smooth velocity field $w(\cdot,t)$.

Significant progress has recently been made in Eulerian FEMs for fluid equations. In particular, Lehrenfeld and Olshanskii [93] proposed a CutFEM-based Eulerian framework for parabolic problems on moving domains, which was later extended to the Stokes equations by Burman et al. [31], who proved optimal-order estimates in the L^2H^1 and L^2L^2 norms. Alternatively, the ALE method has been a widely used approach for addressing the challenges posed by domain motion, and it is the method employed in Chapter 2.

We consider the semidiscrete finite element approximation of problem (1.0.1). Let $V_h^r(\Omega_h(t))$ and $Q_h^{r-1}(\Omega_h(t))$ denote the Taylor–Hood P_r–P_{r-1} finite element spaces on the evolving computational domain $\Omega_h(t)$. We seek functions $u_h(t) \in V_h^r(\Omega_h(t))$ and $p_h(t) \in Q_h^{r-1}(\Omega_h(t))$ such that $u_h(0) = I_h u(0)$ and $u_h = I_h w$ on $\partial\Omega_h(t)$, satisfying:

$$(D_{t,h}u_h - w_h\cdot\nabla u_h, v_h)_{\Omega_h(t)} + (\nabla u_h, \nabla v_h)_{\Omega_h(t)} - (\nabla\cdot v_h, p_h)_{\Omega_h(t)} = (f, v_h)_{\Omega_h(t)}, \tag{1.0.2a}$$

$$(\nabla \cdot u_h, q_h)_{\Omega_h(t)} = 0, \tag{1.0.2b}$$

for all test functions $v_h \in \mathring{V}_h^r(\Omega_h(t))$ and $q_h \in Q_h^{r-1}(\Omega_h(t))$.

Optimal convergence of order $O(h^{r+1})$ in the $L^\infty(0,T;L^2)$ norm for the ALE semidiscrete FEM has been proved for diffusion equations in moving domains by [64] and [102], assuming polynomial degree $r \geq 1$. However, for the Stokes and Navier–Stokes equations, prior analyses have yielded only suboptimal rates. Specifically, for the ALE FEM with Taylor–Hood elements, existing results show an L^2 error of order $O(h^r)$; see [92, 122, 108].

In Chapter 2, we close this gap by proving that the semidiscrete ALE FEM achieves $O(h^{r+1})$ convergence in the L^2 norm for the velocity and $O(h^r)$ convergence for the pressure. As demonstrated in [64, 102], obtaining optimal convergence requires establishing the following estimate for the commutator between the material derivative $D_{t,h}$ and the Stokes–Ritz projection R_h:

$$\|D_{t,h}R_h u - R_h D_{t,h} u\|_{L^2} \leq C h^{r+1}. \tag{1.0.3}$$

To establish (1.0.3), the main challenge lies in the involvement of the pressure component of the Stokes–Ritz projection in the L^2 duality argument for the commutator term. This difficulty is resolved by additionally deriving and employing an optimal H^{-1}-norm error estimate for the pressure component of the Stokes–Ritz projection.

Chapter 3 of this thesis is dedicated to the optimal L^2 error analysis of finite element methods (FEM) for fluid–structure interaction (FSI) problems. This chapter addresses two primary aspects: the first is the optimal error analysis for the semidiscrete finite element approximation of the FSI system; the second concerns the stability and convergence of loosely coupled schemes for time-discretized FSI problems.

To simplify the analysis, we consider a model problem describing the interaction between a viscous incompressible fluid and a thin elastic structure. The fluid is governed by the Stokes equations:

$$\begin{cases} \rho_f \partial_t \mathbf{u} - \nabla \cdot \boldsymbol{\sigma}(\mathbf{u}, p) = 0, & \text{in } (0,T) \times \Omega, \\ \nabla \cdot \mathbf{u} = 0, & \text{in } (0,T) \times \Omega, \\ \mathbf{u}(0,\cdot) = \mathbf{u}_0(x), & \text{on } \Omega, \end{cases} \tag{1.0.4}$$

and the structure is modeled by a linear elastic wave equation:

$$\begin{cases} \rho_s \epsilon_s \partial_{tt} \boldsymbol{\eta} - \mathcal{L}_s \boldsymbol{\eta} = -\boldsymbol{\sigma}(\mathbf{u}, p)\mathbf{n}, & \text{in } (0,T) \times \Sigma, \\ \boldsymbol{\eta}(0,x) = \boldsymbol{\eta}_0(x), & \text{on } \Sigma, \\ \partial_t \boldsymbol{\eta}(0,x) = \mathbf{u}_0(x), & \text{on } \Sigma, \end{cases} \tag{1.0.5}$$

subject to the kinematic interface condition:

$$\partial_t \boldsymbol{\eta} = \mathbf{u} \qquad \text{on } (0,T) \times \Sigma, \tag{1.0.6}$$

and periodic inflow–outflow conditions at the lateral boundaries Σ_l and Σ_r (see Figure 1.1).

In this analysis, several simplifying assumptions are made: (i) the domain deformation induced by structural displacement is neglected (fixed-domain model), (ii) the domain Ω admits an exact triangulation (no domain approximation error), (iii) the fluid and structure are governed by linear equations (Stokes and elastic wave models), and (iv) the

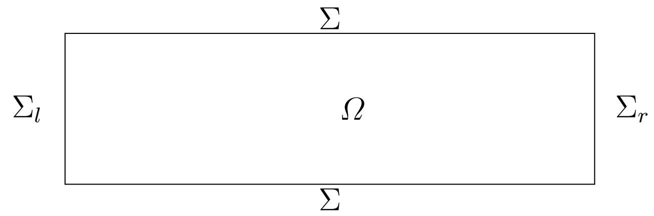


Figure 1.1: The computational domain for the thin-structure interaction problem

domain is assumed to be smooth with periodic inflow–outflow conditions (no geometric singularities).

To the best of our knowledge, even under these simplified assumptions, optimal L^2-norm error estimates for FEM applied to FSI problems have not been previously established. A primary difficulty lies in the absence of an appropriate Ritz projection that accounts for the coupling between the fluid and structure. Standard Ritz projections applied separately to the fluid and structure components fail to yield optimal L^2-error bounds for the coupled system; see, e.g., [5, 25, 54, 91, 120].

In Chapter 3, we resolve this issue by introducing a novel coupled non-stationary Ritz projection. This projection consists of a triple $(R_h\mathbf{u}, R_hp, R_h\boldsymbol{\eta})$ of finite element functions that satisfy a weak form of the coupled system, together with the time-dependent interface constraint $(R_h\mathbf{u})|_\Sigma = \partial_t R_h\boldsymbol{\eta}$ on $\Sigma \times [0,T]$. This construction is equivalent to solving an evolution problem for $R_h\boldsymbol{\eta}$ with a suitably chosen initial condition $R_h\boldsymbol{\eta}(0)$.

Moreover, the dual problem associated with the non-stationary Ritz projection is formulated as a backward-in-time initial–boundary value problem:

$$
\begin{aligned}
-\mathcal{L}_s\boldsymbol{\phi} + \boldsymbol{\phi} &= \partial_t\boldsymbol{\sigma}(\boldsymbol{\phi}, q)\mathbf{n} + \boldsymbol{f}, && \text{on } \Sigma \times [0,T), && (1.0.7a)\\
-\nabla\cdot\boldsymbol{\sigma}(\boldsymbol{\phi}, q) + \boldsymbol{\phi} &= 0, && \text{in } \Omega \times [0,T), && (1.0.7b)\\
\nabla\cdot\boldsymbol{\phi} &= 0, && \text{in } \Omega \times [0,T), && (1.0.7c)\\
\boldsymbol{\sigma}(\boldsymbol{\phi}, q)\mathbf{n} &= 0, && \text{at } t = T. && (1.0.7d)
\end{aligned}
$$

This system is equivalent to a backward evolution equation for $\boldsymbol{\xi} := \boldsymbol{\sigma}(\boldsymbol{\phi}, q)\mathbf{n}$ of the form:

$$
-\mathcal{L}_s\mathcal{N}\boldsymbol{\xi} + \mathcal{N}\boldsymbol{\xi} - \partial_t\boldsymbol{\xi} = \boldsymbol{f} \quad \text{on } \Sigma \times [0,T), \quad \boldsymbol{\xi}(T) = 0, \tag{1.0.8}
$$

where $\mathcal{N} : H^{-1/2}(\Sigma)^d \to H^{1/2}(\Sigma)^d$ denotes the Neumann-to-Dirichlet map associated with the Stokes system. By choosing a well-designed initial value $R_h\boldsymbol{\eta}(0)$ and utilizing the regularity properties of the dual problem (1.0.7), which are shown by analyzing the equivalent formulation in (1.0.8), we are able to derive optimal L^2-error estimates for the Ritz projection and, consequently, for the semidiscrete FEM approximation of the FSI problem.

For time discretization, loosely coupled schemes enable separate treatment of the fluid and structure subproblems without requiring extra iterations. However, ensuring stability—especially in the presence of strong added-mass effects, such as in hemodynamics—is a major challenge; see [33]. The design of stable loosely coupled methods has been an active area of research [29, 30, 13, 68, 72].

Among these, the kinematically coupled scheme has gained prominence due to its modularity, stability, and ease of implementation. The scheme was first analyzed in [23, 72, 25]. It proceeds in two steps: first, compute $(\boldsymbol{s}^n, \boldsymbol{\eta}^n)$ satisfying

$$
\rho_s\epsilon_s\frac{\boldsymbol{s}^n - \mathbf{u}^{n-1}}{\tau} - \mathcal{L}_s\boldsymbol{\eta}^n = -\boldsymbol{\sigma}^{n-1}\cdot\boldsymbol{n}, \qquad \text{on } \Sigma, \tag{1.0.9}
$$

$$\eta^n = \eta^{n-1} + \tau \boldsymbol{s}^n, \qquad \text{on } \Sigma,$$

then compute $(\mathbf{u}^n, p^n)$ satisfying

$$\begin{aligned} \rho_f D_\tau \mathbf{u}^n + \nabla \cdot \boldsymbol{\sigma}^n &= 0, && \text{in } \Omega, \\ \nabla \cdot \mathbf{u}^n &= 0, && \text{in } \Omega, \\ \rho_s \epsilon_s \frac{\mathbf{u}^n - \boldsymbol{s}^n}{\tau} + (\boldsymbol{\sigma}^n - \boldsymbol{\sigma}^{n-1}) \cdot \boldsymbol{n} &= 0, && \text{on } \Sigma. \end{aligned} \tag{1.0.10}$$

In Chapter 3, we propose a finite element fully discrete version of this kinematically coupled scheme. The fluid stress is evaluated explicitly as

$$\boldsymbol{\sigma}_h^n \cdot \boldsymbol{n} := \boldsymbol{\sigma}(\mathbf{u}_h^n, p_h^n) \cdot \boldsymbol{n} = (-p_h^n I + 2\mu \mathbb{D}(\mathbf{u}_h^n)) \cdot \boldsymbol{n}.$$

To ensure unconditional energy stability, additional stabilization terms are incorporated:

$$\rho_s \epsilon_s \left(\frac{\mathbf{u}_h^n - \boldsymbol{s}_h^n}{\tau}, \frac{\tau}{\rho_s \epsilon_s} \boldsymbol{\sigma}(\mathbf{v}_h, q_h) \cdot \boldsymbol{n} \right)_\Sigma, \quad \left((\boldsymbol{\sigma}_h^n - \boldsymbol{\sigma}_h^{n-1}) \cdot \boldsymbol{n}, \frac{\tau(1+\beta)}{\rho_s \epsilon_s} \boldsymbol{\sigma}(\mathbf{v}_h, q_h) \cdot \boldsymbol{n} \right)_\Sigma.$$

Although alternative unconditionally stable fully discrete kinematically coupled schemes have been proposed and studied [26, 120, 54, 5, 28, 25], their analyses have generally yielded only suboptimal L^2 error estimates, partly due to the lack of an appropriate Ritz projection for the coupled system. In contrast, by employing the coupled non-stationary Ritz projection introduced in this chapter, we are able to establish an optimal L^2 error estimate for the proposed fully discrete scheme.

The third part of this thesis concerns the maximum-norm stability and error estimates of isoparametric finite element methods (FEM). Such stability and error estimates have been established in the literature in settings where the domain Ω can be exactly triangulated by finite elements—for instance, polygonal or polyhedral domains that admit exact triangulations using linear simplicies, so that the computational domain Ω_h coincides with the exact domain Ω. However, in practical computations, curved boundaries of smooth domains—or more generally, curvilinear polygons or polyhedra that may include curved faces, edges, and corners—are typically approximated using isoparametric finite elements. In such cases, the discrepancy between the exact domain Ω and the computational domain Ω_h introduces a domain perturbation that must be carefully accounted for in both the stability analysis and the error estimates.

In Chapter 4, we investigate the weak discrete maximum principle and derive optimal maximum-norm error estimates for isoparametric FEM applied to elliptic problems. Let $\Omega \subset \mathbb{R}^N$ with $N \in \{2, 3\}$ be a (possibly concave) curvilinear polyhedral domain with edge openings smaller than π, and let $\mathcal{T}_h$ be a quasi-uniform family of meshes composed of isoparametric elements of order r, such that the Hausdorff distance between Ω and the computational domain $\Omega_h = (\bigcup_{K \in \mathcal{T}_h} K)^\circ$ is of order $O(h^{r+1})$. Denote by $S_h(\Omega_h)$ the associated isoparametric finite element space. A function $u_h \in S_h(\Omega_h)$ is said to be discrete harmonic if

$$\int_{\Omega_h} \nabla u_h \cdot \nabla \chi_h = 0 \quad \forall \chi_h \in S_h^\circ(\Omega_h). \tag{1.0.11}$$

If all such discrete harmonic functions satisfy

$$\|u_h\|_{L^\infty(\Omega_h)} \le C \|u_h\|_{L^\infty(\partial \Omega_h)}, \tag{1.0.12}$$

with a constant C independent of h, the FEM is said to satisfy the *weak discrete maximum principle*.

The result in [125] established this principle for a broad class of H^1-conforming elements on quasi-uniform meshes in polygonal domains. Moreover, the principle was used to derive maximum-norm stability and best-approximation properties of the Ritz projection:

$$\|u - R_h u\|_{L^\infty(\Omega)} \leqslant C\ell_h \inf_{v_h \in S_h^\circ(\Omega)} \|u - v_h\|_{L^\infty(\Omega)} \quad \forall\, u \in H_0^1(\Omega) \cap L^\infty(\Omega), \tag{1.0.13}$$

where $R_h : H_0^1(\Omega) \to \check{S}_h^\circ(\Omega)$ is the Ritz projection and

$$\ell_h = \begin{cases} \ln(2 + 1/h) & \text{for piecewise linear elements,} \\ 1 & \text{for higher-order finite elements.} \end{cases}$$

The method employed in [125] remains a fundamental approach for establishing maximum norm estimates in FEM. Specifically, [125] reduces the problem to an L^1-type error estimate between the regularized elliptic Green's function and its Ritz projection onto the finite element space, using a dyadic decomposition of the domain and a kick-back argument to derive the desired L^1-type error estimate. This argument was further refined in [96], extending the result to three-dimensional polyhedral domains.

Some related results have been proved in the case $\Omega_h \neq \Omega$. For general bounded smooth domains which may be concave, thus the finite element space may be non-conforming, Kashiwabara & Kemmochi [79] have obtained the following error estimate for piecewise linear finite elements for the Poisson equation under the Neumann boundary condition:

$$\|\tilde{u} - u_h\|_{L^\infty(\Omega_h)} \leqslant Ch|\log h| \inf_{v_h \in S_h} \|\tilde{u} - v_h\|_{W^{1,\infty}(\Omega_h)} + Ch^2|\log h|\|u\|_{W^{2,\infty}(\Omega)}, \tag{1.0.14}$$

where $\tilde{u}$ is any extension of u in $W^{2,\infty}(\Omega_\delta)$ and Ω_δ is a neighborhood of $\overline{\Omega}$. More recently, the $W^{1,\infty}$ stability of the Ritz projection was proved in [43] for isoparametric FEMs on $C^{r+1,1}$-smooth domains based on weighted-norm estimates, where r denotes the degree of finite elements.

For curvilinear polyhedral domains, the weak maximum principle and the best approximation results in the L^∞ norm have not been proved. In this chapter, we close the gap mentioned above by proving the weak maximum principle in (1.0.12) for isoparametric finite elements of degree $r \geqslant 1$ in a bounded smooth domain or a curvilinear polyhedron (possibly concave) with edge openings smaller than π. The weak maximum principle is proved by converting the finite element weak form on Ω_h to a weak form on Ω by using a bijective transformation $\mathbf{\Phi}_h : \Omega_h \to \Omega$ which is piecewisely defined on the triangles/tetrahedra. This yields a bilinear form with a discontinuous coefficient matrix A_h. To align the reduction step with [125, 96], we reduce the weak discrete maximum principle to a L^1-type estimate for $v - R_h v$, where v is a regularized Green function on Ω with respect to the coefficient matrix A_h (see (4.3.16) for definition of v). The difficulty arises from the limited regularity of v, as it solves an elliptic equation with discontinuous coefficients. To address this, we decomposes v into two components: v_1, a regularized Green's function for the original Laplacian equation, and v_2, which corresponds to an elliptic equation with discontinuous coefficients A_h but with a small source term arising from the domain perturbation, and then estimate the two parts separately by using the H^2 and $W^{1,p}$ regularity of the respective problems.

As an application of the weak maximum principle, we prove that the finite element solution $u_h \in S_h^{\circ}(\Omega_h)$ of the Poisson equation

$$\begin{cases} -\Delta u = f & \text{in } \Omega \\ \quad\; u = 0 & \text{on } \partial\Omega \end{cases} \tag{1.0.15}$$

using isoparametric finite elements of degree $r \geqslant 1$ has the following optimal-order error bound (for any $p > N$):

$$\|u - u_h\|_{L^\infty(\Omega)} \leqslant C\ell_h \|u - \check{I}_h u\|_{L^\infty(\Omega)} + Ch^{r+1}\ell_h \|f\|_{L^p(\Omega)}, \tag{1.0.16}$$

where u_h is extended to be zero in $\Omega\backslash\Omega_h$, and $\check{I}_h u$ denotes a Lagrange interpolation operator. The maximum-norm error estimate is established in two steps. First, we follow the approach of [125] to derive the L^∞-stability of the Ritz projection R_h from the weak discrete maximum principle, which yields an optimal maximum-norm error estimate between the finite element solution u_h and the auxiliary solution $u^{(h)}$ of the Poisson equation posed on Ω_h, where the source term f is extended by zero outside Ω. The second step is to estimate the difference between u and $u^{(h)}$, which is achieved using a maximum principle argument in an enlarged domain Ω^t that contains both Ω and Ω_h. This step leads to the error term $Ch^{r+1}\ell_h\|f\|_{L^p(\Omega)}$ in (1.0.16), capturing the effect of domain perturbation $\Omega \neq \Omega_h$.

In Chapter 5, we study the analyticity, maximal L^p-regularity, and optimal maximum-norm error estimates of isoparametric FEM for the heat equation in curvilinear polyhedral domains (possibly with non-convex corners):

$$\frac{\partial u(t,x)}{\partial t} - \Delta u(t,x) = f(t,x), \qquad \forall (t,x) \in \mathbb{R}_+ \times \Omega, \tag{1.0.17}$$

$$u(t,x) = 0, \qquad \forall (t,x) \in \mathbb{R}_+ \times \partial\Omega, \tag{1.0.18}$$

$$u(0,x) = u_0(x), \qquad \forall x \in \Omega. \tag{1.0.19}$$

The semigroup $E(t) = e^{t\Delta}$ generated by the Laplacian operator Δ satisfies the following analyticity estimates:

$$\sup_{t>0} \big(\|E(t)v\|_{L^q(\Omega)} + t\|\partial_t E(t)v\|_{L^q(\Omega)}\big) \le C\|v\|_{L^q(\Omega)}, \quad \forall v \in L^q(\Omega), \quad 1 \le q < \infty \tag{1.0.20a}$$

$$\sup_{t>0} \Big(\|E(t)v\|_{C_0(\overline{\Omega})} + t\|\partial_t E(t)v\|_{C_0(\overline{\Omega})}\Big) \le C\|v\|_{C_0(\overline{\Omega})}, \quad \forall v \in C_0(\overline{\Omega}), \tag{1.0.20b}$$

Moreover, when $u_0 = 0$, the solution of (1.0.17) exhibits maximal L^p regularity in the space $L^q(\Omega)$:

$$\|\partial_t u\|_{L^p(\mathbb{R}_+;L^q(\Omega))} + \|\Delta u\|_{L^p(\mathbb{R}_+;L^q(\Omega))} \le C_{p,q}\|f\|_{L^p(\mathbb{R}_+;L^q(\Omega))} \quad \forall 1 < p,q < \infty. \tag{1.0.21}$$

For spatially semidiscrete FEM, we study whether discrete analogues of (1.0.20)–(1.0.21) hold. Let $E_h(t) := e^{t\Delta_h}$ be the discrete semigroup. Then the corresponding estimates are:

$$\sup_{t>0} \big(\|E_h(t)v_h\|_{L^q(\Omega_h)} + t\|\partial_t E_h(t)v_h\|_{L^q(\Omega_h)}\big) \le C\|v_h\|_{L^q(\Omega_h)} \tag{1.0.22a}$$

$$\forall v_h \in S_h^{\circ}(\Omega_h), \quad 1 \le q \le \infty,$$

$$\|\partial_t u_h\|_{L^p(\mathbb{R}_+;L^q(\Omega_h))} + \|\Delta_h u_h\|_{L^p(\mathbb{R}_+;L^q(\Omega_h))} \le C_{p,q}\|f_h\|_{L^p(\mathbb{R}_+;L^q(\Omega_h))} \quad \text{if } u_{h,0}=0,\ \forall 1<p,q<\infty. \tag{1.0.22b}$$

The analyticity and maximal regularity of the finite element semi-discrete or fully-discrete problem have numerous applications and serve as important tools for the convergence analysis of numerical schemes for nonlinear parabolic equations [2, 52, 67, 104, 88, 143].

By reducing the problem to an L^1-type error estimate between the discrete Green's function and a regularized Green's function of the parabolic equation, [126, 133] established the analyticity property (5.1.8a) of the discrete semigroup E_h for smooth domains. The key estimate for the discrete Green's function in these works was subsequently employed in [66] to prove the maximal L^p-regularity (5.1.8b) of the discrete semigroup $E_h(t)$ when both the domain and the coefficients of the parabolic equation are sufficiently smooth. Later studies further relaxed the regularity requirements on the domain and the coefficients. In particular, the results in [101, 100, 104] demonstrated that both (5.1.8a) and (5.1.8b) hold when Ω is a (possibly nonconvex) polyhedral domain, provided the coefficients satisfy $a_{ij} \in W^{1,N+\varepsilon}(\Omega)$.

However, these results remain valid only under the assumption that the domain Ω is exactly triangulated. In the setting of isoparametric FEM, as discussed in Chapter 4, the discrepancy between the exact domain Ω and the computational domain Ω_h—i.e., domain perturbation—must be properly addressed. Using the extension method, [80] proved the discrete analyticity and maximal regularity properties (1.0.22a)–(1.0.22b) for finite element semi-discretizations of parabolic equations on smooth domains with Neumann boundary conditions, where Ω_h approximates Ω via a quasi-uniform triangulation $\mathcal{T}_h$ composed of linear simplices, and $S_h(\Omega_h)$ denotes the P^1 continuous finite element space on Ω_h.

The analyticity (1.0.22a) and maximal regularity (1.0.21) of isoparametric FEM on curvilinear polyhedral domains with possibly nonconvex corners had not been established. In Chapter 5, we close this gap by following the strategy of [101] to address the regularity issues posed by nonconvex corners and, in place of the extension method, adopt the transformation method to handle domain perturbation $\Omega \neq \Omega_h$. Specifically, we employ a Lipschitz homeomorphism $\mathbf{\Phi}_h : \Omega_h \to \Omega$, as introduced in Chapter 4, to transform the finite element problem on Ω_h into one on Ω, where the transformed equation involves a discontinuous coefficient matrix A_h. Let $\check{\Gamma}_h$ denote the discrete Green's function of the transformed problem and Γ the regularized Green's function for the original heat equation. Then, following [101], the estimates (1.0.22a) and (1.0.21) are reduced to an L^1-type estimate for the difference $\Gamma - \check{\Gamma}_h$, which is obtained via a kick-back argument involving parabolic dyadic decomposition, local energy error estimates, and local duality arguments. A key challenge stems from the fact that $\Gamma - \check{\Gamma}_h$ satisfies only an *almost Galerkin orthogonality*:

$$(\partial_t \Gamma - a_h(x)\partial_t \check{\Gamma}_h, \check{\chi}_h)_\Omega + (\nabla \Gamma - A_h(x)\nabla \check{\Gamma}_h, \nabla \check{\chi}_h)_\Omega = 0 \quad \forall\, \check{\chi}_h \in \check{S}_h^\circ(\Omega).$$

This complicates the derivation of local energy error estimates and the application of local duality arguments.

As shown in [101], the quasi-maximal L^∞-regularity of the FEM facilitates reducing the maximum-norm stability of finite element solutions for parabolic equations to the maximum-norm stability of the elliptic Ritz projection. Specifically, the following estimate holds:

$$\|u - u_h\|_{L^\infty(0,T;L^\infty(\Omega))} \le C\left(\ell_h \|u - R_h u\|_{L^\infty(0,T;L^\infty(\Omega))} + \|u_0 - u_{h,0}\|_{L^\infty(\Omega)}\right). \tag{1.0.23}$$

In the context of isoparametric FEM, we establish a corresponding estimate in which the additional error introduced by the domain perturbation remains of optimal order. Specifically, we show that

$$\begin{aligned} \|\widetilde{u} - u_h\|_{L^\infty(0,T;L^\infty(\Omega_h))} \le C \left(\ell_h \|\widetilde{u} - R_h \widetilde{u}\|_{L^\infty(0,T;L^\infty(\Omega_h))} + \|\widetilde{u}_0 - u_{h,0}\|_{L^\infty(\Omega_h)}\right) \\ + Ch^{r+1} \left(\|u\|_{L^\infty(0,T;W^{2,\infty}(\Omega))} + \|\partial_t u\|_{L^\infty(0,T;L^\infty(\Omega))}\right). \end{aligned} \tag{1.0.24}$$

Here, $\widetilde{u}$ denotes a Sobolev extension of the exact solution u to the larger domain $\Omega \cup \Omega_h$.

Summary. This thesis contributes to the numerical analysis of partial differential equations by developing novel stability and error estimation techniques for finite element methods in evolving and curvilinear domains. The results offer new theoretical insights and practical tools for improving the accuracy and robustness of FEM-based simulations, particularly in moving domain problems and fluid–structure interaction systems.

Chapter 2

Optimal convergence of arbitrary Lagrangian Eulerian finite element methods for the Stokes equations in an evolving domain

2.1 Introduction

The Stokes equations are 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.,

$$\partial_t u - \Delta u + \nabla p = f \quad \text{in} \quad \bigcup_{t\in(0,T]} \Omega(t) \times \{t\}, \tag{2.1.1a}$$

$$\nabla \cdot u = 0 \quad \text{in} \quad \bigcup_{t\in(0,T]} \Omega(t) \times \{t\}, \tag{2.1.1b}$$

$$u = w \quad \text{on} \quad \bigcup_{t\in(0,T]} \partial\Omega(t) \times \{t\}, \tag{2.1.1c}$$

$$u = u_0 \quad \text{on} \quad \Omega^0 = \Omega(0), \tag{2.1.1d}$$

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 (2.1.1), 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 (2.1.1) are assumed to be sufficiently smooth. To ensure uniqueness of the solutions, we assume that $p(\cdot, t) \in L^2_0(\Omega(t))$, which is the space of functions p in $L^2(\Omega(t))$ such that $\int_{\Omega(t)} p \, \mathbf{d}x = 0$.

Recent advancements have brought significant progress in the convergence analysis of finite element methods (FEMs) for fluid equations in evolving domains. The well-posedness of the Oseen equation in time-dependent domains was proved in [41] by using an evolving space framework. Lozovskiy et al. [111] introduced a quasi-Lagrangian FEM for Navier–Stokes equations in time-dependent domains, demonstrating optimal-order error estimates in the energy norm. In [113] a kth-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 [93], while Burman et al. [31] extended this framework to the Stokes equations, 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 the Stokes and Oseen equations was provided in subsequent studies of von Wahl et al. [135] and Neilan & Olshanskii [116].

Another prevalent method used to handle the complexities arising from domain evolution is the arbitrary Lagrangian-Eulerian (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 Φ_h of Φ, 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

$$D_t u(x,t) := \frac{\mathbf{d}}{\mathbf{d}t} u(\Phi(\xi,t),t) = \partial_t u + w \cdot \nabla u \text{ at } x = \Phi(\xi,t) \in \Omega(t) \text{ for } \xi \in \Omega^0. \quad (2.1.2)$$

Using this definition of material derivative, the first two equations in (2.1.1) can be rewritten as

$$D_t u - w \cdot \nabla u - \Delta u + \nabla p = f, \quad (2.1.3a)$$

$$\nabla \cdot u = 0, \quad (2.1.3b)$$

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 [118] provided stability results for two different ALE finite element schemes. Subsequently, Gastaldi [63] 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 [117], 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 [17] and [58]. 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 [10] 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 [64] for finite element schemes of degree $r \geqslant 1$. We also refer to [50] and [47] 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 [70]. 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 [106] for the ALE semidiscrete FEM utilizing the standard iso-parametric element of degree r in [94].

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 [19] and [18], respectively. The ALE methods for PDEs in bulk domains [70] 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 [45, 51, 86].

The above-mentioned research efforts have focused on diffusion equations with and without advection terms. The analysis of ALE methods for the Stokes and Navier–Stokes equations has also yielded noteworthy results but remained suboptimal, as discussed below. In [92], 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 [122], 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 [122]. 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 [108].

As far as we know, optimal-order convergence of ALE semidiscrete and fully discrete FEMs were not established for the Stokes and Navier–Stokes equations in an evolving domain. As shown in [64, 102], the optimal-order convergence of ALE semidiscrete FEM requires proving the following optimal-order approximation property for the material derivative of the Ritz projection:

$$\|D_{t,h}R_h u - R_h D_{t,h} u\|_{L^2} \leqslant Ch^{r+1}. \tag{2.1.4}$$

In line with the fixed domain case, achieving optimal consistency error in analysis of finite element approximation for the Stokes equations necessitates the use of the Stokes-Ritz projection R_h. As a result, when trying to obtain the optimal-order approximation property (2.1.4) following the duality argument as in [64, 102], 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., (2.4.43), which is used in Lemma 2.4.5. This leads to optimal-order convergence of the ALE semidiscrete FEM, as the main result of this chapter (see Theorem 2.2.1).

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 2.2 we present the basic notation of evolving mesh and ALE finite elements, as well as the semidiscrete ALE FEM for (2.1.1) and the main theorem of this chapter. In Section 2.3 we present some preliminary results for the evolving mesh, ALE finite element spaces, and boundary-skin estimates. The proof of the main theorem is presented in Section 2.4. Section 2.5 includes numerical results for the Stokes equations and Navier-Stokes equations as empirical evidence supporting our theoretical findings.

2.2 Notation and main results

2.2.1 Evolving mesh and ALE finite element spaces

Suppose that the initial smooth domain $\varOmega^0$ 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 [50]). 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 $\Gamma^0 = \partial\varOmega^0$. To achieve this interpolation, we employ iso-parametric finite elements of Lenoir's type (see [94] for further details) at time $t = 0$ based on the parametrization of the boundary which is denoted by $\varUpsilon : \partial\tilde{D} \to \Gamma^0$. Here, $\partial\tilde{D}$ represents the flat boundary face of the triangulated flat domain, which has the same vertices as the curved triangulated domain $\varOmega_h^0 = \bigcup_{K\in\mathcal{T}_h^0} K$. In practical implementations, the parametrization $\varUpsilon$ can be chosen as the normal projection onto Γ^0. In other words, it computes the unique point $\varUpsilon(x) \in \Gamma^0$ satisfying the equation:

$$x = \varUpsilon(x) + \mathrm{sign}(x, \varOmega^0)|x - \varUpsilon(x)|\mathbf{n}(\varUpsilon(x)),$$

where $\mathbf{n}(\varUpsilon(x))$ is the unit outward normal vector at point $\varUpsilon(x)$ and

$$\mathrm{sign}(x, \varOmega^0) = \begin{cases} 1 & \text{for} \quad x \in \mathbb{R}^d\backslash\overline{\varOmega^0}, \\ -1 & \text{for} \quad x \in \varOmega^0. \end{cases}$$

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 ξ_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):

$$\frac{\mathbf{d}}{\mathbf{d}t}x_j(t) = w(x_j(t), t) \quad \text{and} \quad x_j(0) = \xi_j.$$

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 ξ_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 ξ_j form the vertices of a simplex in $\mathcal{T}_h^0$. Hence, the evolving domain $\varOmega_h(t) = \bigcup_{K\in\mathcal{T}_h(t)} K$ serves as an approximation of the exact domain $\varOmega(t)$. 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 ϕ_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 Ω_h^0 and $\Omega_h(t)$ at time t, with a velocity field $w_h \in V_h^r(\Omega_h(t))$ satisfying:

$$w_h(\Phi_h(\xi, t), t) = \frac{\mathbf{d}}{\mathbf{d}t}\Phi_h(\xi, t) = \sum_{j=1}^{N} w(x_j(t), t)\phi_j^0(\xi) \text{ for } \xi \in \Omega_h^0. \tag{2.2.5}$$

This representation corresponds to the unique Lagrange interpolation of the exact velocity $w(\Phi(\cdot, t), t)$. Analogous to definition (2.1.2), 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:

$$D_{t,h}v(x, t) := \frac{\mathbf{d}}{\mathbf{d}t}v(\Phi_h(\xi, t), t) = \partial_t v + \nabla v \cdot w_h \text{ at } x = \Phi_h(\xi, t) \in \Omega_h(t) \text{ for } \xi \in \Omega_h^0. \tag{2.2.6}$$

The pullback of the finite element basis function ϕ_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 ϕ_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 ϕ_j^t, which states:

$$D_{t,h}\phi_j^t(x) = \frac{\mathbf{d}}{\mathbf{d}t}\phi_j^0(\xi) = 0 \text{ at } x = \Phi_h(\xi, t). \tag{2.2.7}$$

2.2.2 The semidiscrete finite element approximation and main results

We consider 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:

$$\mathring{V}_h^r(\Omega_h(t)) := \{u \in V_h^r(\Omega_h(t)) : u|_{\partial\Omega_h(t)} = 0\},$$

$$Q_h^{r-1}(\Omega_h(t)) := \{p \in S_h^{r-1}(\Omega_h(t)) : \int_{\Omega_h(t)} p \, \mathbf{d}x = 0\}.$$

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 $q_h \in Q_h^{r-1}(\Omega_h(t))$:

$$(D_{t,h}u_h - w_h \cdot \nabla u_h, v_h)_{\Omega_h(t)} + (\nabla u_h, \nabla v_h)_{\Omega_h(t)} - (\nabla \cdot v_h, p_h)_{\Omega_h(t)} = (f, v_h)_{\Omega_h(t)}, \quad (2.2.8a)$$

$$(\nabla \cdot u_h, q_h)_{\Omega_h(t)} = 0, \quad (2.2.8b)$$

The main result of this chapter is the following theorem.

Theorem 2.2.1 (Error estimates of the semidiscrete FEM). *Consider the semidiscrete finite element solutions* (u_h, p_h) *given by* (2.2.8). *Assuming that the exact solutions* (u, p) *to problem* (2.1.1) *are sufficiently smooth and have been extended to be defined on* $\mathbb{R}^d$ *via* (2.3.15)*, the following estimate holds under condition that* w *is sufficiently smooth:*

$$\sup_{t\in[0,T]} \|u - u_h\|_{L^2(\Omega_h(t))} \leqslant C R_{u,p} h^{r+1}, \quad (2.2.9)$$

$$\|p - p_h\|_{L^2(0,T;L^2(\Omega_h(t)))} \leqslant C R_{u,p} h^r, \quad (2.2.10)$$

where C *is a constant independent of the mesh size* h *and* $R_{u,p}$ *is a norm of* (u, p) *defined as follows:*

$$\begin{aligned} R_{u,p} :=& \|\partial_t u\|_{L^2(0,T;W^{r+1,\infty}(\mathbb{R}^d)} + \|u\|_{L^2(0,T;W^{r+2,\infty}(\mathbb{R}^d))} \\ &+ \|\partial_t p\|_{L^2(0,T;H^r(\mathbb{R}^d))} + \|p\|_{L^2(0,T;H^{r+1}(\mathbb{R}^d))} \\ &+ \|u\|_{L^\infty(0,T;W^{r+1,\infty}(\mathbb{R}^d))} + \|p\|_{L^\infty(0,T;H^r(\mathbb{R}^d))}. \end{aligned}$$

The rest of this chapter is devoted to the proof of Theorem 2.2.1.

2.3 Preliminary

The analysis of integrals over dynamically evolving domains necessitates the application of the Transport Theorem, as established in [137, 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 2.3.1 (Transport Theorem). *If the domain* Ω *undergoes motion with a velocity field* $w \in W^{1,\infty}(\Omega)$*, we have*

$$\frac{\mathbf{d}}{\mathbf{d}t} \int_\Omega f \, \mathbf{d}x = \int_\Omega D_t f + f \nabla \cdot w \, \mathbf{d}x, \quad (2.3.11)$$

where $D_t f$ *is the material derivative of* f *with respect to the velocity* w.

The interaction between the operators D_t and ∇ plays an essential role in the error analysis. Consequently, we establish the following lemma as a direct consequence of (2.1.2):

Lemma 2.3.2. *For any vector-valued function f, the material derivative of ∇f and $\nabla\cdot f$ with respect to the velocity field w can be expressed as follows:*

$$D_t\nabla f = \nabla D_t f - \nabla f \nabla w, \tag{2.3.12}$$
$$D_t\nabla\cdot f = \nabla\cdot D_t f - (\nabla f) : (\nabla w)^\top. \tag{2.3.13}$$

By employing Verfürth's trick and utilizing the macros-element criterion, as described in [16, Section 8.5 and Section 8.8], we establish the inf-sup condition for the Taylor–Hood type isoparametric elements.

Lemma 2.3.3 (Inf-sup condtion). *There exists a constant $\kappa > 0$, independent of h and $t \in [0,T]$ for $r \geqslant 2$, such that*

$$\sup_{0\neq v_h\in \mathring{V}_h^r(\Omega_h(t))} \frac{(\mathrm{div} v_h, p_h)_{\Omega_h(t)}}{\|\nabla v_h\|_{L^2(\Omega_h(t))}} \geq \kappa \|p_h\|_{L^2(\Omega_h(t))} \quad \forall p_h \in Q_h^{r-1}(\Omega_h(t)). \tag{2.3.14}$$

2.3.1 Boundedness of partial derivatives of the mesh velocity

For any function u defined on $\bigcup_{0\leq t\leq T} \Omega(t)\times\{t\}$, there is an extension function $\widetilde{u}$ defined on $\mathbb{R}^d\times[0,T]$ such that

$$\widetilde{u}(\cdot,t) := E(u(\cdot,t)\circ\Phi(\cdot,t))\circ\Phi(\cdot,t)^{-1}, \tag{2.3.15}$$

where the operator $E : L^1(\Omega(0)) \to L^1(\mathbb{R}^d)$ refers to Stein's extension operator in [130, p. 181, Theorem 5]. It holds that

$$\|\widetilde{u}(\cdot,t)\|_{W^{k,p}(\mathbb{R}^d)} \leq C\|u(\cdot,t)\|_{W^{k,p}(\Omega(t))}. \tag{2.3.16}$$

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 $I_h(t) : C(\Omega_h(t)) \to S_h^r(\Omega_h(t))$. 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 (2.2.5), 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:

$$\|w_h(\cdot,t) - w(\cdot,t)\|_{W_h^{k,\infty}(\Omega_h(t))} \leqslant Ch^{r+1-k}\|w(\cdot,t)\|_{W^{r+1,\infty}(\Omega_h(t))} \quad \forall\, 0\leq k\leq r+1, \tag{2.3.17}$$

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 ξ_j of discrete domain Ω_h^0, thus there holds $\Phi_h(\cdot,t) = I_h\Phi(\cdot,t)$ on Ω_h^0 and we can derive the error between $\Phi(t)$ and $\Phi_h(t)$ as follows:

$$\|\Phi(\cdot,t) - \Phi_h(\cdot,t)\|_{W^{1,\infty}(\Omega_h^0)} \leqslant Ch^r. \tag{2.3.18}$$

The estimates in (2.3.17) and (2.3.18) lead to the following result when h is sufficiently small

$$\|w_h(\cdot,t)\|_{W^{1,\infty}(\Omega_h(t))} + \|\Phi_h(\cdot,t)\|_{W^{1,\infty}(\Omega_h^0)} + \|\Phi_h^{-1}(\cdot,t)\|_{W^{1,\infty}(\Omega_h(t))} \leqslant C, \qquad (2.3.19)$$

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.

2.3.2 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 2.3.4. *For any finite element function $v_h \in \mathring{V}_h^r(\Omega_h(t))$, the following inequalities hold:*

$$\|v_h\|_{L^2(\Omega_h(t)\backslash\Omega(t))} \le Ch^{3(r+1)/2-d/2}\|\nabla v_h\|_{L^2(\Omega_h(t))} \le Ch^{3(r+1)/2-d/2-1}\|v_h\|_{L^2(\Omega_h(t))}.$$

Proof. Using Hölder's inequality, Newton-Leibniz formula and the fact $v_h|_{\partial\Omega_h(t)} = 0$, we have

$$\begin{aligned}\|v_h\|_{L^2(\Omega_h(t)\backslash\Omega(t))} &\le |\Omega_h(t) \setminus \Omega(t)|^{1/2}\|v_h\|_{L^\infty(\Omega_h(t)\backslash\Omega(t))} \\ &\le |\Omega_h(t) \setminus \Omega(t)|^{1/2} \sup_{x\in\Omega_h(t)\backslash\Omega(t)} \mathrm{dist}(x, \partial\Omega_h)\|\nabla v_h\|_{L^\infty(\Omega_h(t))} \\ &\le Ch^{3(r+1)/2}\|\nabla v_h\|_{L^\infty(\Omega_h(t))} \\ &\le Ch^{3(r+1)/2-d/2}\|\nabla v_h\|_{L^2(\Omega_h(t))} \le Ch^{3(r+1)/2-d/2-1}\|v_h\|_{L^2(\Omega_h(t))},\end{aligned}$$

where we used the fact that the distance from $x \in \Omega_h \setminus \Omega$ 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 [102, eq. (3.32)].

Lemma 2.3.5. *Let $g \in W^{1,1}(\mathbb{R}^d)$. Then the following inequality holds:*

$$\|g\|_{L^1(\partial\Omega_h(t))} \le C\|g\|_{L^1(\partial\Omega(t))} + C\|\nabla g\|_{L^1(\Omega(t)\cup\Omega_h(t))}, \qquad (2.3.20)$$

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 [53, Corollary 1.5].

Lemma 2.3.6. *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:*

$$\|\chi\|_{H^2(\Omega(t))} \le C\|\lambda\|_{H^1(\Omega(t))},$$

where the constant C is independent of $t \in [0,T]$.

2.4 Error estimates of the semidiscrete FEM

2.4.1 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 $\partial\Omega_h(t)$:

$$(\nabla R_h v, \nabla \chi_h)_{\Omega_h(t)} - (\nabla \cdot \chi_h, R_h q)_{\Omega_h(t)} = (\nabla v, \nabla \chi_h)_{\Omega_h(t)} - (\nabla \cdot \chi_h, q)_{\Omega_h(t)}, \quad (2.4.21a)$$

$$(\nabla \cdot R_h v, \lambda_h)_{\Omega_h(t)} = (\nabla \cdot v, \lambda_h)_{\Omega_h(t)}. \quad (2.4.21b)$$

Additionally, we define the norm $\|\cdot\|'$ over any domain $D \subset \mathbb{R}^d$ as follows:

$$\|f\|'_{L^2(D)} := \|f - \bar{f}\|_{L^2(D)}, \quad (2.4.22)$$

where $\bar{f}$ denotes the average of f over D, given by $\bar{f} := \frac{1}{|D|}\int_D f \,\mathbf{d}x$.

By utilizing the inf-sup condition (2.3.14), the Stokes–Ritz projection exhibits quasi-optimal error estimates, as stated in the following lemma:

Lemma 2.4.1. *[69, 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*

$$\|\nabla(v - R_h v)\|_{L^2(\Omega_h(t))} + \|q - R_h q\|'_{L^2(\Omega_h(t))} \leq Ch^r \Big(\|v\|_{H_h^{r+1}(\Omega_h(t))} + \|q\|_{H_h^r(\Omega_h(t))}\Big),$$

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 2.4.1. Lemma 2.4.1 is a corollary of [69, Chapter 2, Theorem 1.1] and error estimates of Lagrange interpolation:

$$\|\nabla(v - I_h v)\|_{L^2(\Omega_h(t))} \leq Ch^k \|v\|_{H_h^{k+1}(\Omega_h(t))}, \quad \|q - I_h q\|_{L^2(\Omega_h(t))} \leq Ch^k \|q\|_{H_h^k(\Omega_h(t))},$$

where k is restricted by condition that $\frac{d}{2} < k \leq r$ 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 2.4.1, it suffices to take $k = r$. Since $r \geq 2$ 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 φ of the duality problem (2.4.32), 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. [129] and [21, Section 4.8]). Though we are working with finite element space consisting of isoparametric elements, the same strategy as in [129, Theorem 3.1] still applies to prove the following first-order error estimate:

$$\|q - \mathcal{I}_h q\|_{L^2(\Omega_h(t))} \leq Ch\|q\|_{H^1(\Omega_h(t))}. \quad (2.4.23)$$

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:

$$\|q - \mathbb{P}_h q\|_{L^2(\Omega_h(t))} \leq \|q - \mathcal{I}_h q\|_{L^2(\Omega_h(t))} \leq Ch\|q\|_{H^1(\Omega_h(t))} \quad (2.4.24)$$

We shall utilize (2.4.24) in our duality argument contained in Lemma 2.4.3 and Lemma 2.4.5 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:

$$E_{t,h} := D_{t,h} R_h - R_h D_{t,h}. \tag{2.4.25}$$

We can establish the following lemma about the estimate of $\nabla E_{t,h} v$ and $E_{t,h} q$.

Lemma 2.4.2. *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:*

$$\|\nabla E_{t,h} v\|_{L^2(\Omega_h(t))} + \|E_{t,h} q\|'_{L^2(\Omega_h(t))} \le C h^r \Big(\|v\|_{H_h^{r+1}(\Omega_h(t))} + \|q\|_{H_h^r(\Omega_h(t))} \Big). \tag{2.4.26}$$

Proof. Since equations (2.4.21) 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 (2.4.21), the following equation holds for each $s \in [0, T]$:

$$(\nabla (R_h v - v)(s), \nabla \chi_h(s))_{\Omega_h(s)} - (\nabla \cdot \chi_h(s), (R_h q - q)(s))_{\Omega_h(s)} = 0, \tag{2.4.27a}$$
$$(\nabla \cdot (R_h v - v)(s), \lambda_h(s) - \overline{\lambda_h(s)})_{\Omega_h(s)} = 0, \tag{2.4.27b}$$

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 χ_h and λ_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 (2.4.27), and using Lemma 2.3.1 and Lemma 2.3.2, we obtain

$$\begin{aligned}
&(\nabla D_{t,h} R_h v, \nabla \chi_h)_{\Omega_h(t)} - (\nabla \cdot \chi_h, D_{t,h} R_h q)_{\Omega_h(t)} - (\nabla D_{t,h} v, \nabla \chi_h)_{\Omega_h(t)} + (\nabla \cdot \chi_h, D_{t,h} q)_{\Omega_h(t)} \\
&= -(\nabla (v - R_h v) \nabla w_h, \nabla \chi_h)_{\Omega_h(t)} + (\nabla (v - R_h v), \nabla \chi_h (\nabla \cdot w_h - \nabla w_h))_{\Omega_h(t)} \\
&\quad + (\nabla \chi_h : (\nabla w_h)^\top - \nabla \cdot \chi_h \nabla \cdot w_h, q - R_h q)_{\Omega_h(t)},
\end{aligned} \tag{2.4.28a}$$

$$\begin{aligned}
&(\nabla \cdot D_{t,h} R_h v, \lambda_h)_{\Omega_h(t)} - (\nabla \cdot D_{t,h} v, \lambda_h)_{\Omega_h(t)} \\
&= -(\nabla (v - R_h v) : (\nabla w_h)^\top, \lambda_h)_{\Omega_h(t)} + (\nabla \cdot (v - R_h v), \lambda_h \nabla \cdot w_h)_{\Omega_h(t)} \\
&\quad + (\nabla \cdot (R_h v - v), 1)_{\Omega_h(t)} \frac{(\lambda_h, \nabla \cdot w_h)_{\Omega_h(t)}}{|\Omega_h(t)|}.
\end{aligned} \tag{2.4.28b}$$

Similarly to the definition of (2.4.21), we can define the Stokes–Ritz projection of $(D_{t,h} v, D_{t,h} q)$, and substitute the definition into (2.4.28), we obtain

$$\begin{aligned}
&(\nabla E_{t,h} v, \nabla \chi_h)_{\Omega_h(t)} - (\nabla \cdot \chi_h, E_{t,h} q)_{\Omega_h(t)} \\
&= -(\nabla (v - R_h v) \nabla w_h, \nabla \chi_h)_{\Omega_h(t)} + (\nabla (v - R_h v), \nabla \chi_h \nabla \cdot w_h - \nabla \chi_h \nabla w_h)_{\Omega_h(t)} \\
&\quad + (\nabla \chi_h : (\nabla w_h)^\top - \nabla \cdot \chi_h \nabla \cdot w_h, q - R_h q)_{\Omega_h(t)},
\end{aligned} \tag{2.4.29a}$$

$$\begin{aligned}
&(\nabla \cdot E_{t,h} v, \lambda_h)_{\Omega_h(t)} \\
&= -(\nabla (v - R_h v) : (\nabla w_h)^\top, \lambda_h)_{\Omega_h(t)} + (\nabla \cdot (v - R_h v), \lambda_h \nabla \cdot w_h)_{\Omega_h(t)} \\
&\quad + (\nabla \cdot (R_h v - v), 1)_{\Omega_h(t)} \frac{(\lambda_h, \nabla \cdot w_h)_{\Omega_h(t)}}{|\Omega_h(t)|}.
\end{aligned} \tag{2.4.29b}$$

By the definition of the Stokes–Ritz projection, $R_h v = I_h v$ and $R_h D_{t,h} v = I_h D_{t,h} v$ on the boundary $\partial\Omega_h(t)$. Then $D_{t,h} R_h v = D_{t,h} I_h v = I_h D_{t,h} v$ 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 (2.4.29) 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

$$\begin{aligned}\|\nabla E_{t,h} v\|^2_{L^2(\Omega_h(t))} \leqslant & C\|\nabla(R_h v - v)\|^2_{L^2(\Omega_h(t))} + C\|q - R_h q\|^2_{L^2(\Omega_h(t))} \\ & + C\|E_{t,h} q - \overline{E_{t,h} q}\|_{L^2(\Omega_h(t))}\|\nabla(R_h v - v)\|_{L^2(\Omega_h(t))}. \end{aligned} \tag{2.4.30}$$

By using the inf-sup condition (2.3.14) and the equation (2.4.29a), we have

$$\begin{aligned}\|E_{t,h} q - \overline{E_{t,h} q}\|_{L^2(\Omega_h(t))} \leqslant & C \sup_{0\neq\chi_h\in\mathring{V}_h^r} \frac{(\nabla\cdot\chi_h, E_{t,h} q - \overline{E_{t,h} q})}{\|\nabla\chi_h\|_{L^2(\Omega_h(t))}} = C \sup_{0\neq\chi_h\in\mathring{V}_h^r} \frac{(\nabla\cdot\chi_h, E_{t,h} q)}{\|\nabla\chi_h\|_{L^2(\Omega_h(t))}} \\ \leqslant & C\Big(\|\nabla E_{t,h} v\|_{L^2(\Omega_h(t))} + \|\nabla(R_h v - v)\|_{L^2(\Omega_h(t))} + \|q - R_h q\|_{L^2(\Omega_h(t))}\Big). \end{aligned} \tag{2.4.31}$$

By substituting (2.4.31) into (2.4.30), and using Young's inequality and Lemma 2.4.1 under the assumption $\overline{q} = 0$, we obtain the desired result. ∎

2.4.2 The Nitsche's trick and duality argument

In order to obtain an optimal order error estimate of $R_h u - u$, we will apply Nitsche'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 $(\psi, \varphi) \in H_0^1(\Omega(t)) \times L_0^2(\Omega(t))$:

$$-\Delta\psi + \nabla\varphi = g_h \quad \text{in } \Omega(t), \tag{2.4.32a}$$

$$\nabla\cdot\psi = 0 \quad \text{in } \Omega(t), \quad \psi|_{\partial\Omega(t)} = 0. \tag{2.4.32b}$$

By applying regularity estimates for the Stokes equations in $\Omega(t)$, we obtain the following result:

$$\|\psi\|_{H^2(\Omega(t))} + \|\nabla\varphi\|_{L^2(\Omega(t))} \le C\|g_h\|_{L^2(\Omega_h(t))}. \tag{2.4.33}$$

To extend the functions ψ and φ to $\widetilde{\psi}$ and $\widetilde{\varphi}$, respectively, we employ the Stein extension operator as in (2.3.15). 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:

$$\|g_h\|^2_{L^2(\Omega_h(t))} = (\nabla\widetilde{\psi}, \nabla g_h)_{\Omega_h(t)} - (\nabla\cdot g_h, \widetilde{\varphi})_{\Omega_h(t)} - (g_h, \widetilde{\eta})_{\Omega_h(t)}. \tag{2.4.34}$$

Notably, since $\widetilde{\eta}$ vanishes in $\Omega(t)$ and we have $r \ge 2$, we can utilize Lemma 2.3.4 along with the regularity estimate (2.4.33) to obtain the following inequality:

$$|(g_h, \widetilde{\eta})_{\Omega_h(t)}| = |(g_h, \widetilde{\eta})_{\Omega_h(t)\setminus\Omega(t)}| \le \|\widetilde{\eta}\|_{L^2(\mathbb{R}^d)}\|g_h\|_{L^2(\Omega_h(t)\setminus\Omega(t))} \le Ch^2\|g_h\|^2_{L^2(\Omega_h(t))}. \tag{2.4.35}$$

Consequently, when $h > 0$ is sufficiently small, we can absorb $|(g_h, \widetilde{\eta})_{\Omega_h(t)}|$ on the right-hand side of (2.4.34) by the left-hand side. This yields the following estimate:

$$\|g_h\|^2_{L^2(\Omega_h(t))} \leqslant C\Big|(\nabla\widetilde{\psi}, \nabla g_h)_{\Omega_h(t)} - (\nabla\cdot g_h, \widetilde{\varphi})_{\Omega_h(t)}\Big|. \tag{2.4.36}$$

We choose $g_h = R_h v - I_h v \in \mathring{V}_h^r(\Omega_h(t))$ in (2.4.32). By appropriately estimating the right-hand side of (2.4.36), we can derive the following lemma.

Lemma 2.4.3. *Let (R_hv, R_hq) 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*

$$\|R_hv - v\|_{L^2(\Omega_h(t))} \leq Ch^{r+1} \left(\|v\|_{W_h^{r+1,\infty}(\Omega_h(t))} + \|q\|_{H_h^r(\Omega_h(t))} \right). \tag{2.4.37}$$

Proof. For $g_h = R_hv - I_hv$, by utilizing (2.4.36), the definition of Stokes–Ritz projection (2.4.21) with $(\chi_h, \lambda_h) = (I_h\widetilde{\psi}, \mathbb{P}_h^*\widetilde{\varphi})$ (noting that $I_h\tilde{\psi}|_{\partial\Omega_h(t)} = 0$ 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

$$\begin{aligned}
\|g_h\|^2_{L^2(\Omega_h(t))} \leqslant & C\Big|(\nabla\widetilde{\psi}, \nabla(R_hv - v))_{\Omega_h(t)} - (\nabla \cdot (R_hv - v), \widetilde{\varphi})_{\Omega_h(t)}\Big| \\
& + C\Big|(\nabla\widetilde{\psi}, \nabla(v - I_hv))_{\Omega_h(t)} - (\nabla \cdot (v - I_hv), \widetilde{\varphi})_{\Omega_h(t)}\Big| \\
\leqslant & C\Big|(\nabla(\widetilde{\psi} - I_h\widetilde{\psi}), \nabla(R_hv - v))_{\Omega_h(t)} + (\nabla \cdot I_h\widetilde{\psi}, R_hq - q - \overline{q})_{\Omega_h(t)}\Big| \\
& + C\Big|(\nabla \cdot (R_hv - v), \widetilde{\varphi} - \mathbb{P}_h^*\widetilde{\varphi})_{\Omega_h(t)}\Big| \\
& + C\Big|(-\Delta\widetilde{\psi} + \nabla\widetilde{\varphi}, v - I_hv)_{\Omega_h(t)} + (\nabla\widetilde{\psi} \cdot \mathbf{n} - \widetilde{\varphi}\mathbf{n}, v - I_hv)_{\partial\Omega_h(t)}\Big|,
\end{aligned}$$

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. (2.4.24) of Remark 2.4.1)

$$\|\mathbb{P}_h\widetilde{\varphi} - \widetilde{\varphi}\|_{L^2(\Omega_h(t))} \leq Ch\|\widetilde{\varphi}\|_{H^1(\Omega_h(t))}. \tag{2.4.38}$$

And we can deduce following estimate for $|\overline{\widetilde{\varphi}}|$ from condition $\varphi \in L_0^2(\Omega(t))$,

$$\begin{aligned}
|\overline{\widetilde{\varphi}}| \leq & C \left(\int_{\Omega_h(t)\setminus\Omega(t)} |\widetilde{\varphi}| + \int_{\Omega(t)\setminus\Omega_h(t)} |\widetilde{\varphi}| \right) \\
\leq & C\|\widetilde{\varphi}\|_{L^6(\mathbb{R}^d)}(|\Omega(t) \setminus \Omega_h(t)|^{1/3} + |\Omega_h(t) \setminus \Omega(t)|^{1/3}) \\
\leq & Ch\|\widetilde{\varphi}\|_{H^1(\mathbb{R}^d)}.
\end{aligned} \tag{2.4.39}$$

As a corollary, we have

$$\|\mathbb{P}_h^*\widetilde{\varphi} - \widetilde{\varphi}\|_{L^2(\Omega_h(t))} \leq Ch\|\widetilde{\varphi}\|_{H^1(\Omega_h(t))}. \tag{2.4.40}$$

By using the error estimate of interpolation I_h, the error estimate of $\mathbb{P}_h^*$ (2.4.40), Lemma 2.4.1, and the regularity result (2.4.33), we have

$$\begin{aligned}
& \|g_h\|^2_{L^2(\Omega_h(t))} \\
\leqslant & C\Big(\|\widetilde{\psi}\|_{H^2(\mathbb{R}^d)} + \|\widetilde{\varphi}\|_{H^1(\mathbb{R}^d)}\Big)\Big(h\|\nabla(R_hv - v)\|_{L^2(\Omega_h(t))} + \|v - I_hv\|_{L^2(\Omega_h(t))}\Big) \\
& + C\Big|(\nabla \cdot I_h\widetilde{\psi}, q - R_hq - \overline{q})_{\Omega_h(t)}\Big| + C\Big|(\mathbf{n} \cdot \nabla\widetilde{\psi} - \widetilde{\varphi}\mathbf{n}, v - I_hv)_{\partial\Omega_h(t)}\Big| \\
\leqslant & Ch^{r+1}\Big(\|v\|_{H^{r+1}(\Omega_h(t))} + \|q\|_{H^r(\Omega_h(t))}\Big)\|g_h\|_{L^2(\Omega_h(t))} + C\Big|(\nabla \cdot (\widetilde{\psi} - I_h\widetilde{\psi}), q - R_hq - \overline{q})_{\Omega_h(t)}\Big| \\
& + C\Big|(\nabla \cdot \widetilde{\psi}, q - R_hq - \overline{q})_{\Omega_h(t)}\Big| + C\Big|(\nabla\widetilde{\psi} \cdot \mathbf{n} - \widetilde{\varphi}\mathbf{n}, v - I_hv)_{\partial\Omega_h(t)}\Big|.
\end{aligned}$$

We estimate the left terms subsequently.

$$\left|(\nabla\cdot(\widetilde{\psi}-I_h\widetilde{\psi}), q-R_hq-\overline{q})_{\Omega_h(t)}\right| \leqslant Ch\|\widetilde{\psi}\|_{H^2(\Omega_h(t))}\|q-R_hq\|'_{L^2(\Omega_h(t))}.$$

It is known that $\nabla\cdot\widetilde{\psi}=0$ in $\Omega(t)$. Hence, we have

$$\begin{aligned}\left|(\nabla\cdot\widetilde{\psi}, q-R_hq-\overline{q})_{\Omega_h(t)}\right| \leqslant&\|\nabla\cdot\widetilde{\psi}\|_{L^2(\Omega_h(t)\setminus\Omega(t))}\|q-R_hq\|'_{L^2(\Omega_h(t))}\\ \leqslant&C|\Omega_h(t)\setminus\Omega(t)|^{\frac{1}{3}}\|\nabla\cdot\widetilde{\psi}\|_{L^6(\mathbb{R}^d)}\|q-R_hq\|'_{L^2(\Omega_h(t))}\\ \leqslant&Ch\|\widetilde{\psi}\|_{H^2(\mathbb{R}^d)}\|q-R_hq\|'_{L^2(\Omega_h(t))},\end{aligned} \tag{2.4.41}$$

where the last inequality follows from the fact $r\geqslant 2$, Lemma 2.4.1, and (2.4.33). By using Lemma 2.3.5 and (2.4.33), we have

$$\begin{aligned}\left|(\mathbf{n}\cdot\nabla\widetilde{\psi}-\widetilde{\varphi}\mathbf{n}, v-I_hv)_{\partial\Omega_h(t)}\right| \leqslant&C\|v-I_hv\|_{L^\infty(\partial\Omega_h(t))}\Big(\|\nabla\widetilde{\psi}\|_{L^1(\partial\Omega_h(t))}+\|\widetilde{\varphi}\|_{L^1(\partial\Omega_h(t))}\Big)\\ \leqslant&C\|v-I_hv\|_{L^\infty(\Omega_h(t))}\Big(\|\widetilde{\psi}\|_{H^2(\mathbb{R}^d)}+\|\widetilde{\varphi}\|_{H^1(\mathbb{R}^d)}\Big)\\ \leqslant&Ch^{r+1}\|v\|_{W_h^{r+1,\infty}(\Omega_h(t))}\|g_h\|_{L^2(\Omega_h(t))}.\end{aligned} \tag{2.4.42}$$

Combining the above estimates, we obtain the desired estimate (2.4.37). ∎

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 2.4.4. *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.*

$$\left|(q-R_hq-\overline{q},\lambda)_{\Omega_h(t)}\right| \leqslant Ch^{r+1}\left(\|v\|_{W_h^{r+1,\infty}(\Omega_h(t))}+\|q\|_{H_h^r(\Omega_h(t))}\right)\|\lambda\|_{H^1(\mathbb{R}^d)}. \tag{2.4.43}$$

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 2.3.6, there exists $\chi\in H^2(\Omega(t))\cap H_0^1(\Omega(t))$ such that

$$\mathrm{div}\chi=\lambda-\overline{\lambda}_* \quad \text{in } \Omega(t), \quad \|\chi\|_{H^2(\Omega(t))}\leq C\|\lambda\|_{H^1(\mathbb{R}^d)}. \tag{2.4.44}$$

We extend χ to $\widetilde{\chi}\in H^2(\mathbb{R}^d)$ as mentioned in (2.3.15). By decomposing the integral, we have

$$\begin{aligned}\left|(q-R_hq-\overline{q},\lambda)_{\Omega_h(t)}\right| =& \left|(q-R_hq-\overline{q},\lambda-\overline{\lambda}_*)_{\Omega_h(t)}\right|\\ \leqslant& \left|(q-R_hq-\overline{q},\lambda-\overline{\lambda}_*)_{\Omega_h(t)\setminus\Omega(t)}\right|+\left|(q-R_hq-\overline{q},\nabla\cdot\widetilde{\chi})_{\Omega_h(t)\cap\Omega(t)}\right|\\ \leq& \left|(q-R_hq-\overline{q},\lambda-\overline{\lambda}_*)_{\Omega_h(t)\setminus\Omega(t)}\right|+\left|(q-R_hq-\overline{q},\nabla\cdot\widetilde{\chi})_{\Omega_h(t)}\right|\\ &+\left|(q-R_hq-\overline{q},\nabla\cdot\widetilde{\chi})_{\Omega_h(t)\setminus\Omega(t)}\right|.\end{aligned}$$

To estimate the boundary-skin integral, we derive the following inequalities by using Lemma 2.4.1:

$$\begin{aligned}&\left|(q-R_hq-\overline{q},\lambda-\overline{\lambda}_*)_{\Omega_h(t)\setminus\Omega(t)}\right|+\left|(q-R_hq-\overline{q},\nabla\cdot\widetilde{\chi})_{\Omega_h(t)\setminus\Omega(t)}\right|\\ \leq&\|q-R_hq\|'_{L^2(\Omega_h(t))}\left(\|\overline{\lambda}_*\|_{L^2(\Omega_h(t)\setminus\Omega(t))}+\|\lambda\|_{L^2(\Omega_h(t)\setminus\Omega(t))}+\|\nabla\cdot\tilde{\chi}\|_{L^2(\Omega_h(t)\setminus\Omega(t))}\right)\end{aligned}$$

$$
\begin{aligned}
&\le \|q-R_hq\|'_{L^2(\Omega_h(t))}\left(|\overline{\lambda}_*||\Omega_h(t)\setminus\Omega(t)|^{1/2}+(\|\nabla\cdot\tilde{\chi}\|_{L^6(\mathbb{R}^d)}+\|\lambda\|_{L^6(\mathbb{R}^d)})|\Omega_h(t)\setminus\Omega(t)|^{1/3}\right)\\
&\le \|q-R_hq\|'_{L^2(\Omega_h(t))}\left(|\overline{\lambda}_*||\Omega_h(t)\setminus\Omega(t)|^{1/2}+(\|\tilde{\chi}\|_{H^2(\mathbb{R}^d)}+\|\lambda\|_{H^1(\mathbb{R}^d)})|\Omega_h(t)\setminus\Omega(t)|^{1/3}\right)\\
&\le Ch^{r+1}\Big(\|v\|_{H_h^{r+1}(\Omega_h(t))}+\|q\|_{H_h^r(\Omega_h(t))}\Big)\|\lambda\|_{H^1(\mathbb{R}^d)},
\end{aligned}
\tag{2.4.45}
$$

where we have used Hölder's inequality, Sobolev embedding $H^1(\mathbb{R}^d)\hookrightarrow L^6(\mathbb{R}^d)$, regularity estimate (2.4.44), and the fact $r\ge 2$.

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

$$
\begin{aligned}
&\left|(q-R_hq-\overline{q},\nabla\cdot\widetilde{\chi})_{\Omega_h(t)}\right|\\
\le&\left|(q-R_hq-\overline{q},\nabla\cdot(\widetilde{\chi}-\chi_h))_{\Omega_h(t)}\right|+\left|(q-R_hq-\overline{q},\nabla\cdot\chi_h)_{\Omega_h(t)}\right|\\
\le&\left|(q-R_hq-\overline{q},\nabla\cdot(\widetilde{\chi}-\chi_h))_{\Omega_h(t)}\right|+\left|(\nabla(R_hv-v),\nabla(\chi_h-\widetilde{\chi}))_{\Omega_h(t)}\right|+\left|(\nabla(R_hv-v),\nabla\widetilde{\chi})_{\Omega_h(t)}\right|\\
\le&Ch\|\widetilde{\chi}\|_{H^2(\mathbb{R}^d)}\Big(\|q-R_hq\|'_{L^2(\Omega_h(t))}+\|\nabla(R_hv-v)\|_{L^2(\Omega_h(t))}\Big)+\left|(\nabla(R_hv-v),\nabla\widetilde{\chi})_{\Omega_h(t)}\right|\\
\leqslant&Ch^{r+1}\Big(\|v\|_{H_h^{r+1}(\Omega_h(t))}+\|q\|_{H_h^r(\Omega_h(t))}\Big)\|\lambda\|_{H^1(\mathbb{R}^d)}+\left|(\nabla(R_hv-v),\nabla\widetilde{\chi})_{\Omega_h(t)}\right|.
\end{aligned}
$$

Integrating by parts and dealing the boundary integral term as in (2.4.42) and using regularity estimate (2.4.44) as well as Lemma 2.4.3, we obtain

$$
\begin{aligned}
\left|(\nabla(R_hv-v),\nabla\widetilde{\chi})_{L^2(\Omega_h(t))}\right|\le&C\Big(\|R_hv-v\|_{L^2(\Omega_h(t))}+\|R_hv-v\|_{L^\infty(\partial\Omega_h(t))}\Big)\|\widetilde{\chi}\|_{H^2(\mathbb{R}^d)}\\
\le&C\Big(\|R_hv-v\|_{L^2(\Omega_h(t))}+\|R_hv-v\|_{L^\infty(\partial\Omega_h(t))}\Big)\|\lambda\|_{H^1(\mathbb{R}^d)}\\
\le&C\left(\|v\|_{W_h^{r+1,\infty}(\Omega_h(t))}+\|q\|_{H_h^r(\Omega_h(t))}\right)\|\lambda\|_{H^1(\mathbb{R}^d)}
\end{aligned}
$$

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 2.4.3, leveraging the insights gained from (2.4.36) and (2.4.43).

Lemma 2.4.5. *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:*

$$
\|E_{t,h}v\|_{L^2(\Omega_h(t))}\le Ch^{r+1}\left(\|v\|_{W_h^{r+1,\infty}(\Omega_h(t))}+\|q\|_{H_h^r(\Omega_h(t))}\right). \tag{2.4.46}
$$

Proof. Similarly to the proof of Lemma 2.4.2, 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 (2.4.32). By (2.4.36) and the definition of Stokes–Ritz projection, we have

$$
\begin{aligned}
\|E_{t,h}v\|^2_{L^2(\Omega_h(t))}\leqslant&C\left|(\nabla\widetilde{\psi},\nabla g_h)_{\Omega_h(t)}-(\nabla\cdot g_h,\widetilde{\varphi})_{\Omega_h(t)}\right|\\
\le&C\left|(\nabla I_h\widetilde{\psi},\nabla E_{t,h}v)_{\Omega_h(t)}-(\nabla\cdot E_{t,h}v,\mathbb{P}_h^*\widetilde{\varphi})_{\Omega_h(t)}\right|\\
&+C\left|(\nabla(1-I_h)\widetilde{\psi},\nabla E_{t,h}v)_{\Omega_h(t)}-(\nabla\cdot E_{t,h}v,(1-\mathbb{P}_h^*)\widetilde{\varphi})_{\Omega_h(t)}\right|.
\end{aligned}
$$

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 (2.4.29) (where $\mathbb{P}_h$ is the same L^2-orthogonal projection operator as used in proof of Lemma 2.4.3), we obtain

$$
\|E_{t,h}v\|^2_{L^2(\Omega_h(t))}
$$

$$
\begin{aligned}
\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^*\tilde{\varphi})_{\Omega_h(t)} + (\nabla\cdot(v-R_hv), \mathbb{P}_h^*\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\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| \\
&+ 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 (2.3.17), applying the same routine as in the proof of Lemma 2.4.3, i.e. using the error of interpolation I_h, error of modified L^2-projection $\mathbb{P}_h^*$ (2.4.40), Lemma 2.4.2 and an analogue of estimate (2.4.41), 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 (2.4.33), Lemma 2.4.4, and integration by parts, we obtain

$$
\begin{aligned}
\|E_{t,h}v\|^2_{L^2(\Omega_h(t))} \leq & Ch^{r+1}\left(\|v\|_{W_h^{r+1,\infty}(\Omega_h(t))} + \|q\|_{H_h^r(\Omega_h(t))}\right)\|E_{t,h}v\|_{L^2(\Omega_h(t))} \\
&+ C\Big|(R_hv-v,\mathbf{n})_{\partial\Omega_h(t)}\frac{(\mathbb{P}_h^*\tilde{\varphi},\nabla\cdot w_h)_{\Omega_h(t)}}{|\Omega_h(t)|}\Big|
\end{aligned}
$$

$$
\begin{aligned}
&+C\Big| -(R_hv-v,\nabla\cdot(\nabla\widetilde{\psi}\nabla w^\top))_{\Omega_h(t)}+(R_hv-v,\mathbf{n}\cdot(\nabla\widetilde{\psi}\nabla w^\top))_{\partial\Omega_h(t)}\Big|\\
&+C\Big| -(R_hv-v,\nabla\cdot(\nabla\widetilde{\psi}(\nabla\cdot w-\nabla w)))_{\Omega_h(t)}\Big|\\
&+C\Big|(v-R_hv,\mathbf{n}\cdot(\nabla\widetilde{\psi}(\nabla\cdot w-\nabla w)))_{\partial\Omega_h(t)}\Big|\\
&+C\Big|(v-R_hv,\nabla\cdot(\tilde{\varphi}(\nabla w)^\top))_{\Omega_h(t)}-(v-R_hv,\nabla(\tilde{\varphi}\nabla\cdot w))_{\Omega_h(t)}\Big|\\
&+C\Big| -(v-R_hv,\tilde{\varphi}\mathbf{n}\cdot(\nabla w)^\top)_{\partial\Omega_h(t)}+((v-R_hv)\cdot\mathbf{n},\tilde{\varphi}\nabla\cdot w)_{\partial\Omega_h(t)}\Big|.
\end{aligned}
$$

Since $R_hv=I_hv$ on $\partial\Omega_h(t)$, by using Lemma 2.4.3, we have

$$
\begin{aligned}
\|E_{t,h}v\|^2_{L^2(\Omega_h(t))}\leqslant& Ch^{r+1}\left(\|v\|_{W_h^{r+1,\infty}(\Omega_h(t))}+\|q\|_{H_h^r(\Omega_h(t))}\right)\|E_{t,h}v\|_{L^2(\Omega_h(t))}\\
&+C\Big(\|R_hv-v\|_{L^2(\Omega_h(t))}+\|v-I_hv\|_{L^\infty(\partial\Omega_h(t))}\Big)(\|\widetilde{\psi}\|_{H^2(\mathbb{R}^d)}+\|\widetilde{\varphi}\|_{H^1(\mathbb{R}^d)})\\
\leqslant& Ch^{r+1}\left(\|v\|_{W_h^{r+1,\infty}(\Omega_h(t))}+\|q\|_{H_h^r(\Omega_h(t))}\right)\|E_{t,h}v\|_{L^2(\Omega_h(t))}. \qquad (2.4.47)
\end{aligned}
$$

By using Young's inequality, we finish the proof. ∎

With these preparations done, we can go start proving Theorem 2.2.1, which is shown in next subsection.

2.4.3 Proof of Theorem 2.2.1

Proof. We define the auxiliary function ξ in $\mathbb{R}^d$ as follows:

$$
\xi:=\partial_tu-\Delta u+\nabla p-f, \qquad (2.4.48)
$$

where u, p, f represent their extensions to $\mathbb{R}^d$. By testing the equation (2.4.48) 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 2.3.4 and the fact $r\geq 2$, we can derive the following estimate:

$$
|(\xi,v_h)_{\Omega_h(t)}|=|(\xi,v_h)_{\Omega_h(t)\setminus\Omega(t)}|\leq Ch^{r+1}\|\nabla v_h\|_{L^2(\Omega_h(t))}\|\xi\|_{L^2(\mathbb{R}^d)}. \qquad (2.4.49)
$$

It follows from (2.4.21) that the Stokes–Ritz projection (R_hu,R_hp) satisfies the following equation

$$
\begin{aligned}
(D_{t,h}R_hu-&w_h\cdot\nabla R_hu,v_h)_{\Omega_h(t)}+(\nabla R_hu,\nabla v_h)_{\Omega_h(t)}-(\nabla\cdot v_h,R_hp)_{\Omega_h(t)}\\
&=(f,v_h)_{\Omega_h(t)}+(\xi,v_h)_{\Omega_h(t)}+(\mathscr{F},v_h)_{\Omega_h(t)}\quad\forall v_h\in\mathring{V}_h^r(\Omega_h(t)), \qquad (2.4.50a)\\
(\nabla\cdot R_hu,q_h)_{\Omega_h(t)}&=0\quad\forall q_h\in Q_h^{r-1}(\Omega_h(t)), \qquad (2.4.50b)
\end{aligned}
$$

where the remainder $\mathscr{F}:=D_{t,h}(R_hu-u)-w_h\cdot\nabla(R_hu-u)$ 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 $w_h\cdot\nabla(R_hu-u)$ as follows:

$$
\big|(w_h\cdot\nabla(R_hu-u),v_h)_{\Omega_h(t)}\big|=\big|-(v_h\nabla\cdot w_h,R_hu-u)_{\Omega_h(t)}-(R_hu-u,w_h\cdot\nabla v_h)_{\Omega_h(t)}\big|
$$

$$\leq C\|R_h u - u\|_{L^2(\Omega_h(t))}\|v_h\|_{H^1(\Omega_h(t))}, \tag{2.4.51}$$

where we have used the $W^{1,\infty}$ boundedness of the mesh velocity, i.e., $\|w_h(t)\|_{W^{1,\infty}(\Omega_h(t))} \leq C$, which follows from (2.3.17) and the triangle inequality. Thus, we have

$$\begin{aligned}
\left|(\mathscr{F}, v_h)_{\Omega_h(t)}\right| \leq & C\Big(\|D_{t,h}(R_h u - u)\|_{L^2(\Omega_h(t))} + \|R_h u - u\|_{L^2(\Omega_h(t))}\Big)\|v_h\|_{H^1(\Omega_h(t))}. \\
\leqslant & C\Big(\|E_{t,h}u\|_{L^2(\Omega_h(t))} + \|R_h D_{t,h}u - D_{t,h}u\|_{L^2(\Omega_h(t))}\Big)\|v_h\|_{H^1(\Omega_h(t))} \\
& + C\|R_h u - u\|_{L^2(\Omega_h(t))}\|v_h\|_{H^1(\Omega_h(t))}
\end{aligned} \tag{2.4.52}$$

by the definition (2.4.25) of $E_{t,h}$. We can estimate $\|R_h D_{t,h}u - D_{t,h}u\|_{L^2(\Omega_h(t))}$ term in (2.4.52) by Lemma 2.4.3 as follows:

$$\begin{aligned}
& \|R_h D_{t,h}u - D_{t,h}u\|_{L^2(\Omega_h(t))} \\
\leq & Ch^{r+1}\left(\|D_{t,h}u\|_{W_h^{r+1,\infty}(\Omega_h(t))} + \|D_{t,h}p\|_{H_h^r(\Omega_h(t))}\right) \\
\leq & Ch^{r+1}\Big(\|\partial_t u\|_{W^{r+1,\infty}(\Omega_h(t))} + \|w_h\|_{W_h^{r+1,\infty}(\Omega_h(t))}\|u\|_{W^{r+2,\infty}(\Omega_h(t))} \\
& + \|\partial_t p\|_{H^r(\Omega_h(t))} + \|w_h\|_{W_h^{r,\infty}(\Omega_h(t))}\|p\|_{H^{r+1}(\Omega_h(t))}\Big) \\
\leq & Ch^{r+1}\left(\|\partial_t u\|_{W^{r+1,\infty}(\Omega_h(t))} + \|u\|_{W^{r+2,\infty}(\Omega_h(t))} + \|\partial_t p\|_{H^r(\Omega_h(t))} + \|p\|_{H^{r+1}(\Omega_h(t))}\right),
\end{aligned} \tag{2.4.53}$$

where we have employed formula (2.2.6) of material derivative and the $W_h^{r+1,\infty}$-boundedness (2.3.17) of discrete velocity w_h. Combining the estimate (2.4.53) and (2.4.52) as well as using Lemma 2.4.3 and Lemma 2.4.5 , we can derive that

$$|(\mathscr{F}, v_h)_{\Omega_h(t)}| \leqslant Ch^{r+1}A_{u,p}(t)\|v_h\|_{H^1(\Omega_h(t))}, \tag{2.4.54}$$

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 (2.2.8) from equation (2.4.50) we obtain the following equations for any $v_h \in \mathring{V}_h^r(\Omega_h(t))$ and $q_h \in Q_h^{r-1}(\Omega_h(t))$

$$(D_{t,h}e_u - w_h \cdot \nabla e_u, v_h)_{\Omega_h(t)} + (\nabla e_u, \nabla v_h)_{\Omega_h(t)} - (\nabla \cdot v_h, e_p)_{\Omega_h(t)} = (\xi + \mathscr{F}, v_h)_{\Omega_h(t)}, \tag{2.4.55a}$$

$$(\nabla \cdot e_u, q_h)_{\Omega_h(t)} = 0. \tag{2.4.55b}$$

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 (2.4.55a) with $v_h = e_u$, we obtain:

$$\begin{aligned}
\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|e_u(t)\|^2_{L^2(\Omega_h(t))} + \|\nabla e_u(t)\|^2_{L^2(\Omega_h(t))} = & (\xi, e_u)_{\Omega_h(t)} + (\mathscr{F}, e_u)_{\Omega_h(t)} \\
\leq & Ch^{r+1}A_{u,p}(t)\|e_u(t)\|_{H^1(\Omega_h(t))},
\end{aligned}$$

where the last inequality follows from (2.4.49) and (2.4.54).

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:

$$\|e_u\|^2_{L^2(\Omega_h(t))} + \int_0^t \|\nabla e_u\|^2_{L^2(\Omega_h(s))}\mathbf{d}s \leq \|e_h(0)\|^2_{L^2(\Omega_h(0))} + CR^2_{u,p}h^{2(r+1)} \leqslant CR^2_{u,p}h^{2(r+1)}.$$

Combining this result with Lemma 2.4.3, 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 (2.4.55a), we have

$$\begin{aligned}\|e_p\|_{L^2(\Omega_h(t))} \leqslant & C \sup_{0\neq v_h \in \mathring{V}^r_h(\Omega_h(t))} \frac{(\nabla\cdot v_h, e_p)}{\|\nabla v_h\|_{L^2(\Omega_h(t))}} \\ \leqslant & C\left(h^{r+1}A_{u,p}(t) + \|\nabla e_u\|_{L^2(\Omega_h(t))} + \|D_{t,h}e_u\|_{L^2(\Omega_h(t))}\right). \end{aligned} \tag{2.4.56}$$

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 (2.4.55a) and obtain

$$\begin{aligned}&\|D_{t,h}e_u\|^2_{L^2(\Omega_h(t))} - (w_h\cdot\nabla e_u, D_{t,h}e_u)_{\Omega_h(t)} + (\nabla e_u, \nabla D_{t,h}e_u)_{\Omega_h(t)} - (\nabla\cdot D_{t,h}e_u, e_p)_{\Omega_h(t)} \\ &= (\xi + \mathscr{F}, D_{t,h}e_u)_{\Omega_h(t)}.\end{aligned} \tag{2.4.57}$$

From Lemma 2.3.1 and Lemma 2.3.2, it is known that

$$\begin{aligned}(\nabla e_u, \nabla D_{t,h}e_u)_{\Omega_h(t)} =& \frac{1}{2}\frac{\mathbf{d}}{\mathbf{d}t}\|\nabla e_u\|^2_{L^2(\Omega_h(t))} - \frac{1}{2}(|\nabla e_u|^2, \nabla\cdot w_h)_{\Omega_h(t)} \\ &+ \frac{1}{2}(\nabla e_u(\nabla w_h + (\nabla w_h)^\top), \nabla e_u)_{\Omega_h(t)}.\end{aligned} \tag{2.4.58}$$

By taking derivative to (2.4.55b) with respect time, we obtain that

$$(\nabla\cdot D_{t,h}e_u, q_h)_{\Omega_h(t)} + (\nabla\cdot e_u \nabla\cdot w_h - \nabla e_u : (\nabla w_h)^\top, q_h)_{\Omega_h(t)} = 0. \tag{2.4.59}$$

Let $q_h = e_p$ in (2.4.59), we have

$$(\nabla\cdot D_{t,h}e_u, e_p)_{\Omega_h(t)} + (\nabla\cdot e_u \nabla\cdot w_h - \nabla e_u : (\nabla w_h)^\top, e_p)_{\Omega_h(t)} = 0 \tag{2.4.60}$$

Substituting (2.4.49), (2.4.54), (2.4.58) and (2.4.60) into (2.4.57), and using inverse estimate, we can obtain that

$$\begin{aligned}&\|D_{t,h}e_u\|^2_{L^2(\Omega_h(t))} + \frac{1}{2}\frac{\mathbf{d}}{\mathbf{d}t}\|\nabla e_u\|^2_{L^2(\Omega_h(t))} \\ \leqslant & Ch^{r+1}A_{u,p}(t)\|D_{t,h}e_u\|_{H^1(\Omega_h(t))} + C\|D_{t,h}e_u\|_{L^2(\Omega_h(t))}\|\nabla e_u\|_{L^2(\Omega_h(t))} \\ &+ C\|\nabla e_u\|_{L^2(\Omega_h(t))}\|e_p\|_{L^2(\Omega_h(t))} + C\|\nabla e_u\|^2_{L^2(\Omega_h(t))} \\ \leq & Ch^r A_{u,p}(t)\|D_{t,h}e_u\|_{L^2(\Omega_h(t))} + C\|D_{t,h}e_u\|_{L^2(\Omega_h(t))}\|\nabla e_u\|_{L^2(\Omega_h(t))} \\ &+ C\|\nabla e_u\|_{L^2(\Omega_h(t))}\|e_p\|_{L^2(\Omega_h(t))} + C\|\nabla e_u\|^2_{L^2(\Omega_h(t))}\end{aligned} \tag{2.4.61}$$

By substituting (2.4.56) into (2.4.61), using Young's inequality, and integrating both sides from 0 to t, we have

$$\int_0^t \|D_{t,h}e_u\|^2_{L^2(\Omega_h(s))}ds + \|\nabla e_u\|^2_{L^2(\Omega_h(t))}$$

$$\leqslant Ch^{2r}R_{u,p}^2 + C\int_0^t \|\nabla e_u\|_{L^2(\Omega_h(s))}^2 ds + \|\nabla e_u(\cdot,0)\|_{L^2(\Omega_h^0)}^2 \leqslant Ch^{2r}R_{u,p}^2. \tag{2.4.62}$$

Combining (2.4.62) and (2.4.56), we obtain that

$$\|e_p\|_{L^2(0,T;L^2(\Omega_h(t)))} \leqslant CR_{u,p}h^r.$$

By using Lemma 2.4.1, we can deduce that

$$\|p - p_h\|_{L^2(0,T;L^2(\Omega_h(t)))} \leqslant CR_{u,p}h^r + C\|\overline{p}\|_{L^2(0,T)}, \tag{2.4.63}$$

where $\bar{p}(t)$ is the average of $p(t)$ on $\Omega_h(t)$. Since $p \in L_0^2(\Omega(t))$, we have that

$$\begin{aligned}
|\overline{p}(t)| \leqslant& C\left(\Big|\int_{\Omega_h(t)\backslash\Omega(t)} pdx\Big| + \Big|\int_{\Omega(t)\backslash\Omega_h(t)} pdx\Big|\right)\\
\leqslant& C\|p(\cdot,t)\|_{L^\infty(\mathbb{R}^d)}(|\Omega_h(t)\backslash\Omega(t)| + |\Omega(t)\backslash\Omega_h(t)|)\\
\leq& CA_{u,p}(t)h^{r+1} \quad \text{(Sobolev embedding } H^{r+1}(\mathbb{R}^d)\hookrightarrow L^\infty(\mathbb{R}^d) \text{ used)}
\end{aligned} \tag{2.4.64}$$

Combining (2.4.63) and (2.4.64), we complete the proof. ∎

2.5 Numerical experiments

In this section, we provide numerical tests for problem (2.1.1) to support the theoretical result proved in Theorem 2.2.1. 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

$$\begin{aligned}
&\frac{1}{2\tau}\Big[\big(u_h^{n+1} - P_h^{n,n+1}u_h^n, v_h\big)_{\Omega_h^{n+1}} + \big(P_h^{n+1,n}u_h^{n+1} - u_h^n, P_h^{n+1,n}v_h\big)_{\Omega_h^n}\Big]\\
&- \frac{1}{8}\Big((w_h^{n+1} + P_h^{n,n+1}w_h^n)\cdot\nabla(u_h^{n+1} + P_h^{n,n+1}u_h^n), v_h\Big)_{\Omega_h^{n+1}}\\
&- \frac{1}{8}\Big((P_h^{n+1,n}w_h^{n+1} + w_h^n)\cdot\nabla(P_h^{n+1,n}u_h^{n+1} + u_h^n), P_h^{n+1,n}v_h\Big)_{\Omega_h^n}\\
&+ \frac{1}{4}\Big[\Big(\nabla(u_h^{n+1} + P_h^{n,n+1}u_h^n), \nabla v_h\Big)_{\Omega_h^{n+1}} + \Big(\nabla(P_h^{n+1,n}u_h^{n+1} + u_h^n), \nabla P_h^{n+1,n}v_h\Big)_{\Omega_h^n}\Big]\\
&- \frac{1}{2}\Big[\big(\nabla\cdot v_h, p_h^n\circ\Phi_h^n\circ(\Phi_h^{n+1})^{-1}\big)_{\Omega_h^{n+1}} + \big(\nabla\cdot P_h^{n+1,n}v_h, p_h^n\big)_{\Omega_h^n}\Big]\\
=&\frac{1}{2}\Big[(f(t_{n+1}), v_h)_{\Omega_h^{n+1}} + (f(t_n), P_h^{n+1,n}v_h)_{\Omega_h^n}\Big] \quad \forall v_h \in \mathring{V}_h^r(\Omega_h^{n+1}),
\end{aligned} \tag{2.5.65a}$$

$$(\nabla\cdot u_h^{n+1}, q_h)_{\Omega_h^{n+1}} + \beta\tau\big(\nabla(p_h^{n+1} - p_h^n\circ\Phi_h^n\circ(\Phi_h^{n+1})^{-1}), \nabla q_h\big)_{\Omega_h^{n+1}} = 0 \quad \forall q_h \in Q_h^{r-1}(\Omega_h^{n+1}), \tag{2.5.65b}$$

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 (2.5.65a), and subsequently, p_h^{n+1} is computed using equation (2.5.65b) and u_h^{n+1}.

Example 2.5.1. 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 τ 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 2.1. 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 2.2 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 2.2.1 for $r = 2$. Notably, for the $P_{1b} - P_1$ element, we verified that the inf-sup condition (2.3.14) 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 2.3 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 2.5.2. 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

$$\Omega(0) = \{(x, y) : \frac{25}{16}x^2 + \frac{25}{9}y^2 \leqslant 1\}.$$

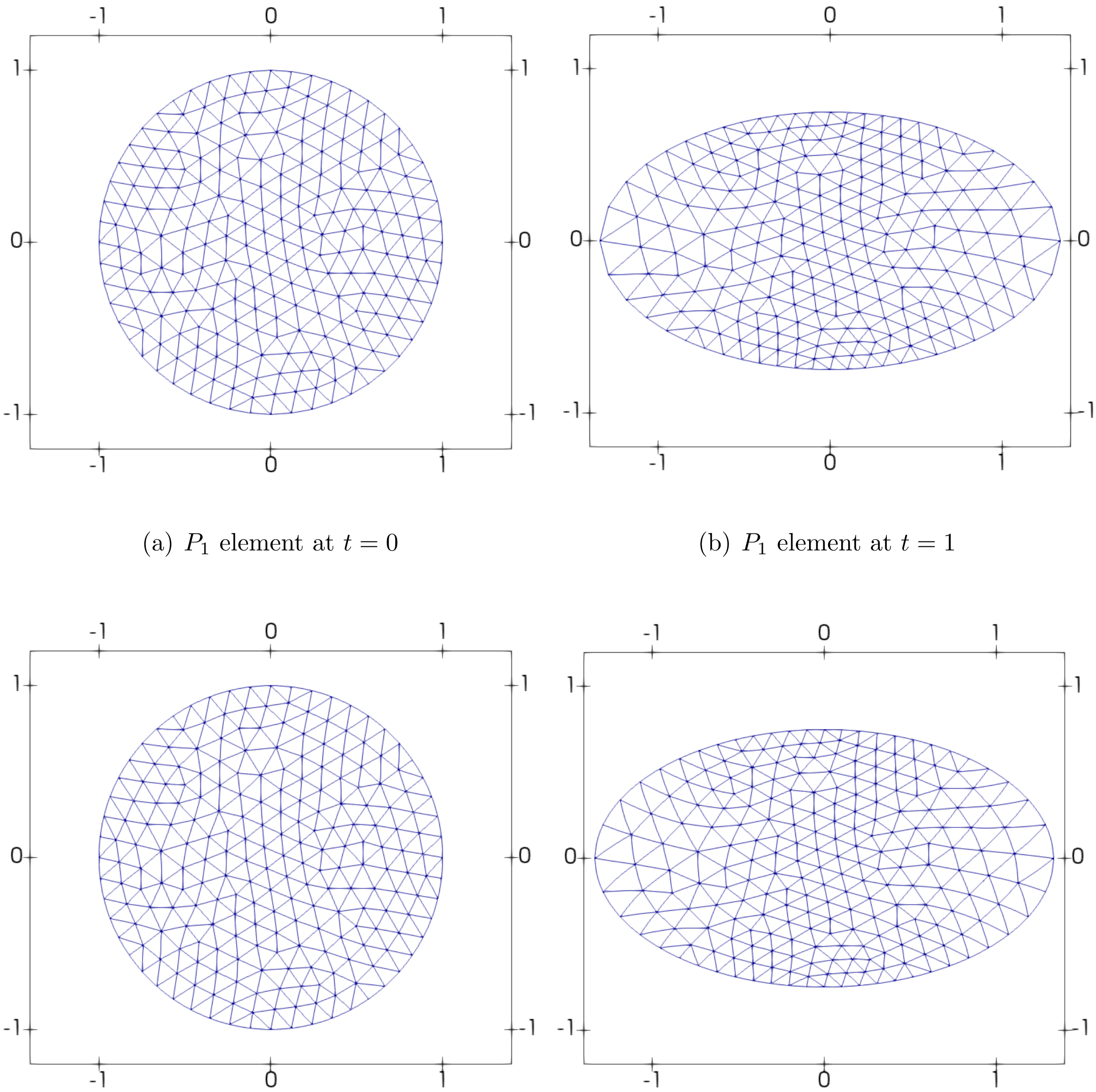
The domain $\Omega(t)$ is generated by the rotating mesh velocity field $w(x, y, t)$, which is given by

$$w(x, y, t) = (-y \sin t, x \cos t).$$

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 (2.1.1a).

Similarly to Example 2.5.1, 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 2.4 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 2.2.1.

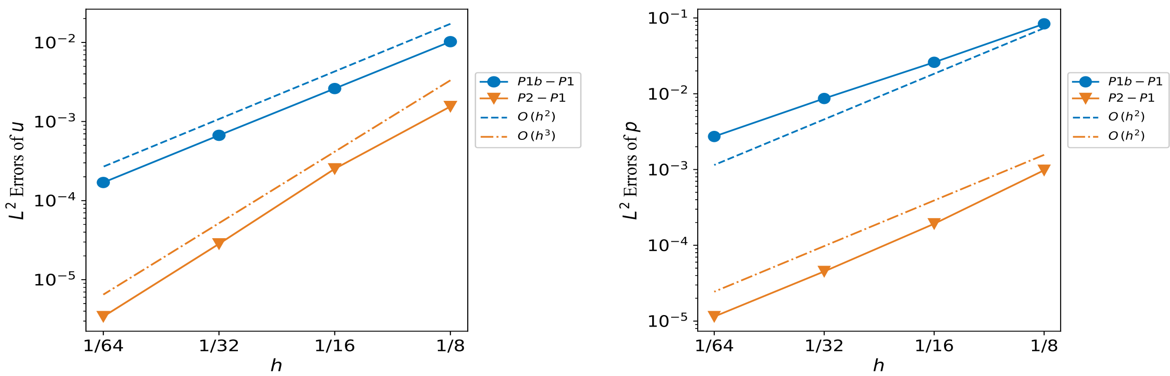
In addition, we examine the convergence of the velocity u in time at $T = 1$ using the $P_2 - P_1$ element, with a sufficiently small mesh size that ensures the errors from



(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 2.1: 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 2.2: Errors from spatial discretization for $T = 1$.

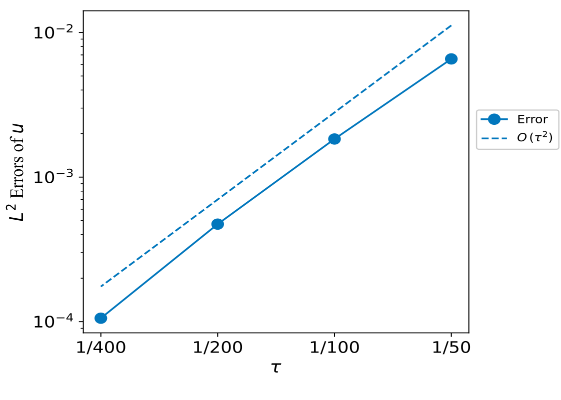
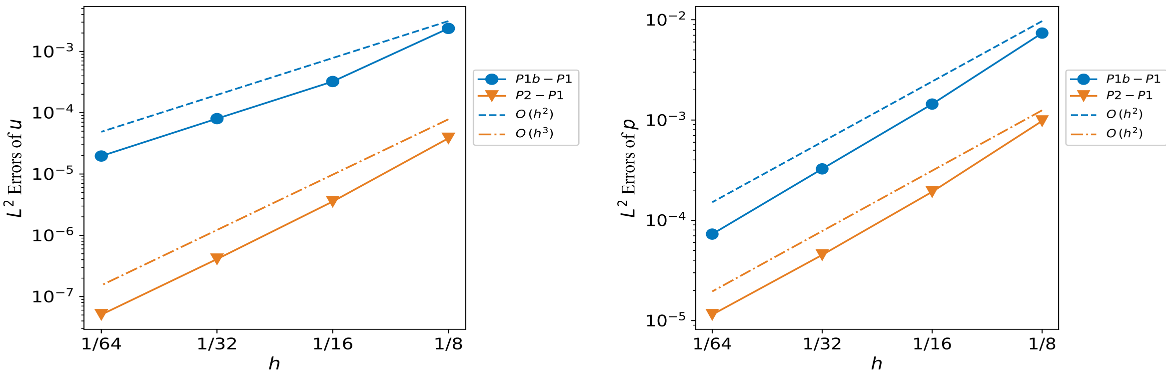


Figure 2.3: Errors from temporal discretization for velocity u at time $T = 1$.

spatial discretization are negligible. The errors of the numerical solutions are presented in Figure 2.5 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) L^2 error of u from spatial discretization at $T = 1$ (b) L^2L^2 error of p from spatial discretization

Figure 2.4: Errors from spatial discretization for $T = 1$.

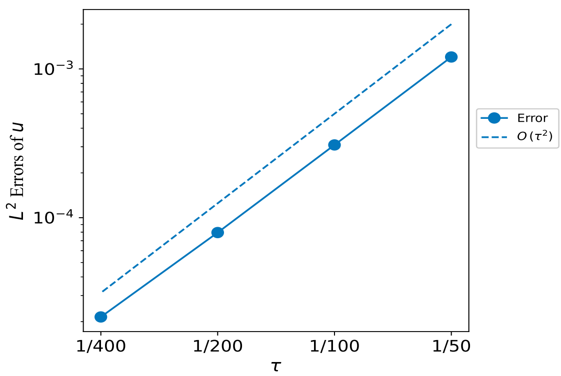


Figure 2.5: Errors from temporal discretization for velocity u at time $T = 1$.

Example 2.5.3 (Navier–Stokes flow in a domain with rotating propeller). In this example, we investigate the fluid motion surrounding a rotating propeller, governed by the Navier–Stokes equation with slip boundary conditions, i.e.

$$\begin{cases} \partial_t u + u \cdot \nabla u - \nabla \cdot (2\mu \mathbb{D}u - \frac{1}{\rho} pI) = 0 & \text{in } \bigcup_{t \in (0,T]} \Omega(t) \times \{t\}, \\ \nabla \cdot u = 0, & \text{in } \bigcup_{t \in (0,T]} \Omega(t) \times \{t\}, \\ u \cdot \mathbf{n} = w \cdot \mathbf{n} & \text{on } \bigcup_{t \in (0,T]} \partial\Omega(t) \times \{t\}, \\ ((2\mu \mathbb{D}u - \frac{1}{\rho} pI) \cdot n)_{\tan} + k u_{\tan} = 0 & \text{on } \bigcup_{t \in (0,T]} \partial\Omega(t) \times \{t\}, \\ u = u_0 & \text{on } \Omega(0), \end{cases} \tag{2.5.66}$$

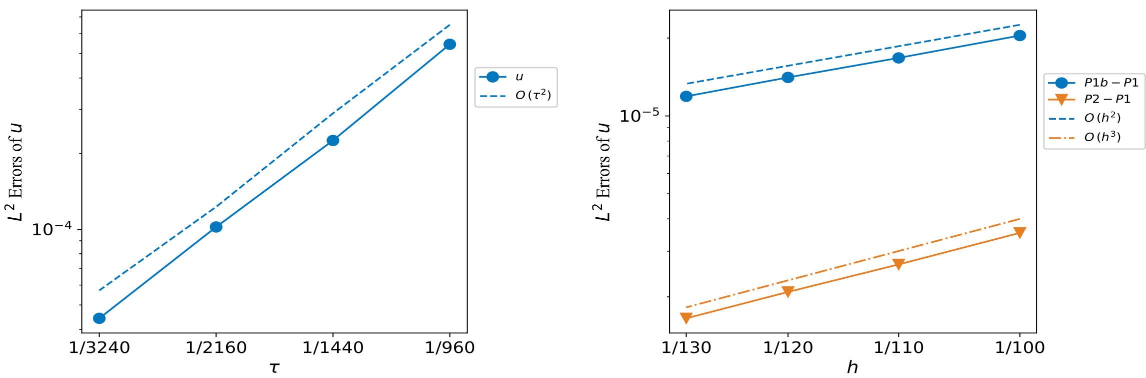
where the subindex tan stands for the tangential component of a vector, w is the velocity of the propeller defined on the boundary and has a natural extension to the whole domain $\Omega(t)$, $\mu = 0.001$ denotes the viscosity, $\rho = 1000$ is the fluid density, and $\mathbf{n}$ is the outward normal vector on the boundary. The initial domain $\Omega(0)$ corresponds to a unit sphere with an ellipse removed, defined as:

$$\Omega(0) = \{(x, y) : x^2 + y^2 \leqslant 1, \text{ and } 2x^2 + 4y^2 \geqslant 1\}.$$

The propeller, depicted as the middle ellipse in Figure 2.7, has a prescribed velocity profile. Specifically, when $0 \leq t \leq 2$, the propeller velocity is given by $w(x, y; t) := (-2ty, 2tx)$, and for $t > 2$, it is defined as $w(x, y; t) := (-2y, 2x)$. This velocity naturally extends to domain Ω.

The numerical method studied in this chapter can be extended to the Navier–Stokes equations with slip boundary conditions. The scheme needs to be modified to suit the slip boundary conditions, and optimal-order convergence in space and second-order convergence in time can be established similarly.

We perform convergence tests for the accuracy of the numerical scheme. To investigate the convergence in time, we select the $P_2 - P_1$ element with a sufficiently small mesh size, ensuring that the errors from space discretization are negligible. The results, presented in Figure 2.6 (a), demonstrate the errors of the numerical solutions for various time step sizes: $\tau = 1/960, 1/1440, 1/2160, 1/3240$, and indicating that the numerical solutions exhibit second-order convergence in time.



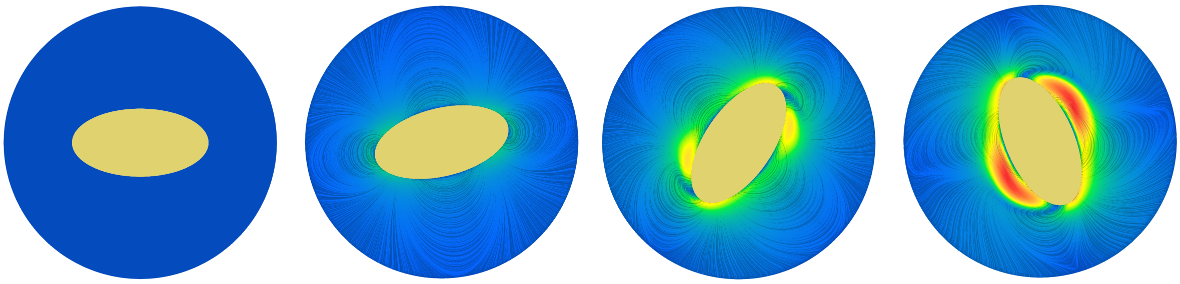
(a) L^2 error of u from time discretization

(b) L^2 error of u from space discretization

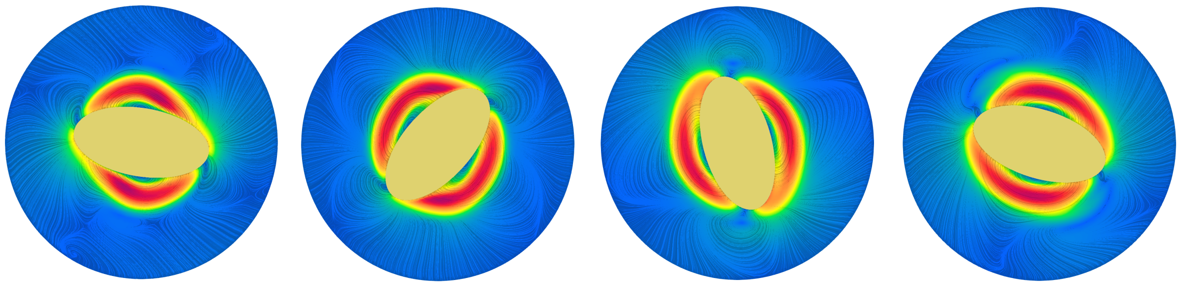
Figure 2.6: Errors from time and space discretization at time $T = 1$.

In addition to time discretization, we illustrate the convergence of spatial discretization using both the $P_{1b} - P_1$ and $P_2 - P_1$ elements, with sufficiently small time step sizes to ensure negligible errors from time discretization. Figure 2.6 (b) presents the errors of the numerical solutions for various mesh sizes: $h = 1/100, 1/110, 1/120, 1/130$. The results demonstrate that the numerical solutions exhibit $(r + 1)$th-order convergence in space for finite elements of degree r. This aligns with the theoretical results presented in Theorem 2.2.1 in the case $r = 2$.

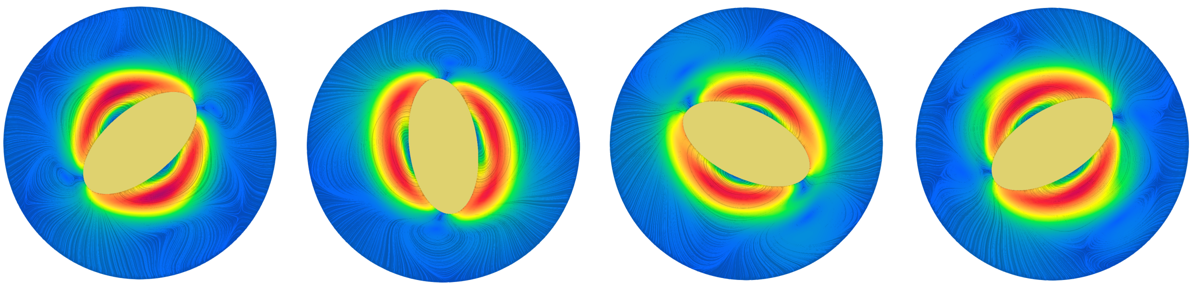
To illustrate the propeller rotation, we conduct simulations with a mesh size of $h = 0.01$ and a time step size of $\tau = 0.001$. Figure 2.7 depicts the process of the propeller rotation and displays the magnitude of the velocity field $|u|$. The figure portrays the flow of the fluid driven by the yellow elliptic propeller, offering insights into propeller-driven flows.



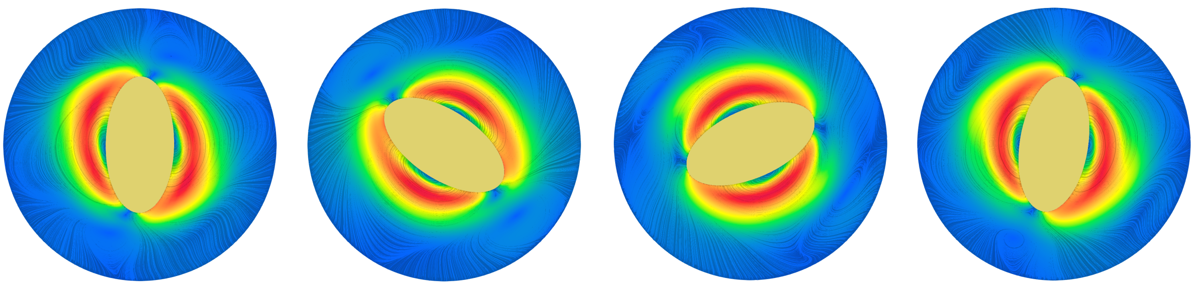
(a) $t = 0$ (b) $t = 0.5$ (c) $t = 1$ (d) $t = 1.5$



(e) $t = 2$ (f) $t = 2.5$ (g) $t = 3$ (h) $t = 3.5$



(i) $t = 4$ (j) $t = 4.5$ (k) $t = 5$ (l) $t = 5.5$



(m) $t = 6$ (n) $t = 6.5$ (o) $t = 7$ (p) $t = 7.5$

(q) The range of $|u|$

Figure 2.7: Flow of the fluid driven by propeller rotation.

Chapter 3

Optimal L^2 error analysis of a loosely coupled finite element scheme for thin-structure interactions

3.1 Introduction

There has been increasing interest in studying fluid-structure interaction due to its diverse applications in many areas [44, 59, 77, 110, 117]. Numerical simulations are crucial in this field, and over the past two decades, numerous efforts have been devoted to developing efficient numerical algorithms and analysis methods.

This chapter focus on a commonly-used academic model problem, where an incompressible fluid interacts with thin structure described by some lower-dimensional, linearly elastic model (such as membranes in 3D, strings in 2D). This thin-structure interaction model is described by the following equations

$$\begin{cases} \rho_f \partial_t \mathbf{u} - \operatorname{div} \boldsymbol{\sigma}(\mathbf{u}, p) = 0, & \text{in } (0,T) \times \Omega, \\ \operatorname{div} \mathbf{u} = 0, & \text{in } (0,T) \times \Omega, \\ \mathbf{u}(0, \cdot) = \mathbf{u}_0(x), & \text{on } \Omega, \end{cases} \tag{3.1.1}$$

$$\begin{cases} \rho_s \epsilon_s \partial_{tt} \eta - \mathcal{L}_s \eta = -\boldsymbol{\sigma}(\mathbf{u}, p)\mathbf{n}, & \text{in } (0,T) \times \Sigma, \\ \eta(0, x) = \eta_0(x), & \text{on } \Sigma, \\ \partial_t \eta(0, x) = \mathbf{u}_0(x), & \text{on } \Sigma \end{cases} \tag{3.1.2}$$

with the kinematic interface condition

$$\partial_t \eta = \mathbf{u} \qquad \text{on } (0,T) \times \Sigma \tag{3.1.3}$$

and certain inflow and outflow conditions at Σ_l and Σ_r; see Figure 3.1. The unknown solutions in (3.1.1) –(3.1.3) are fluid velocity $\mathbf{u}$, fluid pressure p and structure displacement η. The following notations are also used in the model:

ϵ_s:	The thickness of the structure.
μ:	The fluid viscosity.
ρ_f:	The fluid density.
ρ_s:	The structure density.
$\mathbf{n}$:	The outward normal vector on $\partial\Omega$.

$\boldsymbol{\sigma}(\mathbf{u}, p) = -pI + 2\mu\mathbf{D}(\mathbf{u})$: The fluid stress tensor.
$\mathbf{D}(\mathbf{u}) = \frac{1}{2}(\nabla\mathbf{u} + (\nabla\mathbf{u})^T)$: The strain-rate tensor.
$\mathcal{L}_s$: An elliptic differential operator on Σ, such as $\mathcal{L}_s = -I + \Delta_s$, where Δ_s is the Laplace-Beltrami operator on Σ.

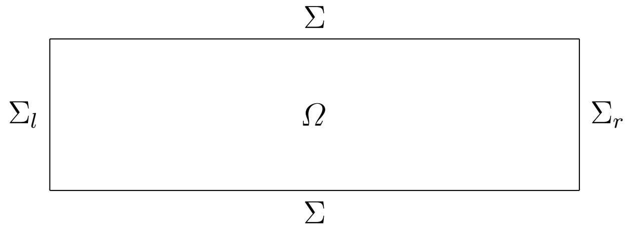


Figure 3.1: The computational domain in the thin-structure interaction problem

In general, two strategies can be employed to construct numerical schemes for solving fluid-structure interaction problems. Monolithic algorithms solve a fully coupled system, which can be expensive for complex fluid-structure problems. Various studies have focused on the numerical simulation and analysis of monolithic algorithms, as can be found in [76, 78, 65, 91, 90, 109, 117]. Alternatively, the fluid and structure sub-problems can be solved separately by partitioned type schemes. A strongly-coupled partitioned scheme often requires extra iterations for the sub-problems at each time step to obtain the solution which at convergence coincides with the monolithic one [117, 55], while the extra iterations are not needed in loosely-coupled partitioned schemes. However, the stability is a key issue for loosely-coupled partitioned schemes, which may be hard to be ensured for highly added mass effect problems such as hemodynamics (e.g.[33]). The development and study of stable loosely-coupled partitioned schemes have been an active area of research (e.g. [29, 30, 13, 68, 72]).

Among those loosely-coupled partitioned schemes, the kinematically coupled scheme is the most popular one due to its modularity, stability, and ease of implementation. The scheme was first studied in [72] for the fluid-structure interaction problems and subsequently by numerous researchers [23, 25, 26, 112, 120]. However, the analysis of kinematically coupled schemes has been challenging due to the specific coupling of two distinct physical phenomena. In [54], Fernandez proposed an incremental displacement-correction scheme, which proved to be stable, and the following energy-norm error estimate was established using piecewise polynomials of degree k for both $\mathbf{u}_h^n$ and $\boldsymbol{\eta}_h^n$ in (3.1.4), i.e.,

$$\|\mathbf{u}^n - \mathbf{u}_h^n\|_{L^2(\Omega)} + \Big(\sum_{m=1}^{n} \tau\|\mathbf{u}^m - \mathbf{u}_h^m\|_f^2\Big)^{\frac{1}{2}} + \|\mathbf{u}^n - \mathbf{u}_h^n\|_{L^2(\Sigma)} + \|\boldsymbol{\eta}^n - \boldsymbol{\eta}_h^n\|_s \leqslant C(\tau + h^k). \tag{3.1.4}$$

The above estimate is optimal only for the velocity in the weak H^1-norm (more precisely, $L^2(H^1)$-norm) and not optimal in L^2-norm. Several different schemes were investigated, and similar error estimates, such as those given in [25, 120], were provided. The kinematic coupling has been extended to other applications, such as composite structures and non-Newtonian flow [24, 112], by many researchers. Additionally, a fully discrete loosely coupled Robin-Robin scheme for thick structures was proposed in [28], where they showed that the error estimate in the same energy norm as in (3.1.4) is in the order of $O(\sqrt{\tau}+h)$ for $k=1$. Recently, a splitting scheme was proposed in [5] for the fluid-structure interaction problem with immersed thin-walled structures. The scheme was proved to be unconditionally stable, and a suboptimal L^2-norm error estimate was presented.

Optimal L^2-norm error estimates play a crucial role in both theoretical analysis of algorithms and development of novel algorithms for practical applications. However, to the best of our knowledge, such results have not been established due to the lack of properly defined Ritz projections for fluid-structure interaction problems. This is in contrast to the error analysis of finite element methods for parabolic equations, where the Ritz projections have been well defined since the early work of Wheeler [140]. For instance, for the heat equation $\partial_t u - \Delta u = f$, the Ritz projection is a finite element function $R_h u$ that satisfies the weak formulation:

$$\int_{\Omega} \nabla(u - R_h u)\cdot\nabla v_h dx = 0 \quad \text{for all finite element functions } v_h. \tag{3.1.5}$$

With this projection R_h, the error of the finite element solution can be decomposed into two parts:

$$u - u_h = (u - R_h u) + (R_h u - u_h).$$

In the analysis of the second part, the pollution from the approximation of the diffusion term is not involved, thus enabling the establishment of an optimal-order error estimate for $\|R_h u - u_h\|_{L^2(\Omega)}$. The optimal estimate for $\|u-u_h\|_{L^2(\Omega)}$ can be derived from the fact that the projection error $\|u - R_h u\|_{L^2(\Omega)}$ is also of optimal order. However, formulating and determining optimal L^2-norm error estimates for a suitably defined Ritz projection in fluid-structure interaction systems remains a challenge. The standard elliptic Ritz projection for the Stokes equations, while widely employed for obtaining error estimates in the energy norm, no longer produces optimal L^2-norm error estimates for such fluid-structure interaction systems; see [5, 25, 54, 91, 120].

In this chapter, we propose a new kinematically coupled scheme which decouples $(\mathbf{u}, p)$ and η for solving the thin-structure interaction problem, and demonstrate its unconditional stability for long-time computation. More importantly, we establish an optimal L^2-norm error estimate for the proposed method, i.e.,

$$\|\mathbf{u}^n - \mathbf{u}_h^n\|_{L^2(\Omega)} + \|\mathbf{u}^n - \mathbf{u}_h^n\|_{L^2(\Sigma)} + \|\boldsymbol{\eta}^n - \boldsymbol{\eta}_h^n\|_{L^2(\Sigma)} \leqslant C(\tau + h^{k+1})\,, \tag{3.1.6}$$

by developing a new framework for the numerical analysis of fluid-structure interaction problems in terms of a newly introduced coupled non-stationary Ritz projection, which is defined as a triple of finite element functions $(R_h\mathbf{u}, R_h p, R_h\boldsymbol{\eta})$ satisfying a weak formulation plus a constraint condition $(R_h\mathbf{u})|_\Sigma = \partial_t R_h\boldsymbol{\eta}$ on $\Sigma\times[0,T]$. This is equivalent to solving an evolution equation of $R_h\boldsymbol{\eta}$ under some initial condition $R_h\boldsymbol{\eta}(0)$. Moreover, the dual problem of the non-stationary Ritz projection, required in the optimal L^2-norm error estimates for the fluid-structure interaction problem, is a backward initial-boundary value problem

$$\begin{aligned}
-\mathcal{L}_s\boldsymbol{\phi} + \boldsymbol{\phi} &= \partial_t\boldsymbol{\sigma}(\boldsymbol{\phi}, q)\mathbf{n} + \mathbf{f} && \text{on } \Sigma\times[0,T) \quad \text{(the boundary condition)} && (3.1.7a)\\
-\nabla\cdot\sigma(\boldsymbol{\phi}, q) + \boldsymbol{\phi} &= 0 && \text{in } \Omega\times[0,T) && (3.1.7b)\\
\nabla\cdot\boldsymbol{\phi} &= 0 && \text{in } \Omega\times[0,T) && (3.1.7c)\\
\boldsymbol{\sigma}(\boldsymbol{\phi}, q)\mathbf{n} &= 0 && \text{at } t = T \quad \text{(the initial condition).} && (3.1.7d)
\end{aligned}$$

which turns out to be equivalent to a backward evolution equation of $\boldsymbol{\xi} = \boldsymbol{\sigma}(\boldsymbol{\phi}, q)\mathbf{n}$, i.e.,

$$-\mathcal{L}_s\mathcal{N}\boldsymbol{\xi} + \mathcal{N}\boldsymbol{\xi} - \partial_t\boldsymbol{\xi} = \mathbf{f} \ \text{ on } \ \Sigma \times [0,T), \ \text{ with initial condition } \ \boldsymbol{\xi}(T) = 0, \qquad (3.1.8)$$

where $\mathcal{N} : H^{-\frac{1}{2}}(\Sigma)^d \to H^{\frac{1}{2}}(\Sigma)^d$ is the Neumann-to-Dirichlet map associated to the Stokes equations. By choosing a well-designed initial value $R_h\boldsymbol{\eta}(0)$ and utilizing the regularity properties of the dual problem (3.1.7), which are shown by analyzing the equivalent formulation in (3.1.8), we are able to establish optimal L^2 error estimates for the non-stationary Ritz projection and, subsequently, optimal L^2-norm error estimates for the finite element solutions of the thin-structure interaction problem.

The rest of this chapter is organized as follows. In Section 2, we introduce a kinematically coupled scheme and present our main theoretical results on the unconditional stability and optimal L^2-norm error estimates of the scheme. We focus on a first-order kinematically coupled time-stepping method and the class of H^1-conforming inf-sup stable finite element spaces, including the classical Taylor–Hood and MINI elements. In Section 3, we introduce a new non-stationary coupled Ritz projection and present the corresponding projection error estimates (with its proof deferred to Section 4). Then we establish unconditionally stability and optimal L^2-norm error estimates for the fully discrete finite element solutions by utilizing the error estimates for the non-stationary coupled Ritz projection. Section 4 is devoted to the proof of the error estimates of the non-stationary coupled Ritz projection. We present a well-designed initial value of the projection and the corresponding error estimates based on duality arguments on the thin solid structure. In Section 5, we provide three numerical examples to support the theoretical analysis presented in this chapter. The first example illustrates the optimal L^2-norm convergence of the proposed fully-discrete kinematically coupled scheme. The second example demonstrates the simulation of certain physical features, which are consistent with previous works. The third example is the 3D simulation of common cardiac arteries in hemodynamics.

3.2 Notations, assumptions and main results

In this section, we propose a stable fully-discrete kinematically coupled FEM for the FSI problem (3.1.1)–(3.1.3). Then, we present main theoretical results in this work.

3.2.1 Notation and weak formulation

Some standard notations and operators are defined below. For any two function u, $v \in L^2(\Omega)$, we denote the inner products and norms of $L^2(\Omega)$ and $L^2(\Sigma)$ by

$$(u, v) = \int_\Omega u(\mathbf{x})v(\mathbf{x})\,\mathrm{d}\mathbf{x}, \qquad \|u\|^2 := (u, u),$$
$$(w, \xi)_\Sigma = \int_\Sigma w(\mathbf{x})\xi(\mathbf{x})\,\mathrm{d}\mathbf{x}, \qquad \|w\|_\Sigma^2 := (w, w)_\Sigma.$$

We assume that $\Omega \subset \mathbb{R}^d$ $(d = 2, 3)$ is a bounded domain with $\partial\Omega = \Sigma_l \cup \Sigma_r \cup \Sigma$, where Σ denotes the fluid-structure interface, Σ_l and Σ_r are two disks (or lines in 2-dimensional case) denoting the inflow and outflow boundary. Moreover, $\Sigma_r = \{(x, y, z+L) : (x, y, z) \in \Sigma_l \ \text{ for some } L > 0\}$.

For the simplicity of analysis, we consider the problem with the periodic boundary condition on Σ_l and Σ_r. Assume that the extended domains Ω_∞ and Σ_∞ are smooth, where

$$\begin{aligned}\Omega_\infty &:= \{(x,y,z) : \exists k \in \mathbb{Z} \text{ such that } (x,y,z+Lk) \in \Omega \cup \Sigma_l\},\\ \Sigma_\infty &:= \{(x,y,z) : \exists k \in \mathbb{Z} \text{ such that } (x,y,z+Lk) \in \overline{\Sigma}\}.\end{aligned}$$

We say a function f defined in Ω_∞ is periodic if

$$f(x,y,z) = f(x,y,z+kL) \quad \forall (x,y,z) \in \Omega \cup \Sigma_l \quad \forall\, k \in \mathbb{Z}.$$

The space of periodic smooth functions on Ω_∞ is denoted as $C^\infty(\Omega_\infty)$. The periodic Sobolev spaces $H^s(\Omega)$ and $H^s(\Sigma)$, with $s \geqslant 0$, are defined as

$$\begin{aligned}H^s(\Omega) &:= \text{The closure of } C^\infty(\Omega_\infty) \text{ under the conventional norm of } H^s(\Omega),\\ H^s(\Sigma) &:= \text{The closure of } C^\infty(\Sigma_\infty) \text{ under the conventional norm of } H^s(\Sigma),\end{aligned}$$

which are equivalent to the Sobolev spaces by considering Ω and Σ as tori in the z direction. The dual spaces of $H^s(\Omega)$ and $H^s(\Sigma)$ are denoted by $H^{-s}(\Omega)$ and $H^{-s}(\Sigma)$, respectively.

We define the following function spaces associated to velocity, pressure and thin structure, respectively:

$$\mathbf{X}(\Omega) := H^1(\Omega)^d, \quad Q(\Omega) := L^2(\Omega), \quad \mathbf{S}(\Sigma) := H^1(\Sigma)^d.$$

Correspondingly, we define the following bilinear forms:

$$a_f(\mathbf{u},\mathbf{v}) := 2\mu(\mathbf{D}(\mathbf{u}), \mathbf{D}(\mathbf{v})) \qquad \text{for } \mathbf{u},\mathbf{v} \in \mathbf{X}(\Omega), \tag{3.2.1}$$

$$b(p,\mathbf{v}) := (p, \nabla\cdot\mathbf{v}) \qquad \text{for } \mathbf{v} \in \mathbf{X}(\Omega) \text{ and } p \in Q(\Omega), \tag{3.2.2}$$

$$a_s(\eta,\mathbf{w}) := (-\mathcal{L}_s\eta, \mathbf{w})_\Sigma \qquad \text{for } \eta,\mathbf{w} \in \mathbf{S}(\Sigma).$$

We assume that $\mathcal{L}_s$ is a second-order differential operator on Σ satisfying the following conditions:

$$\|\mathcal{L}_s\mathbf{w}\|_{H^k(\Sigma)} \le C\|\mathbf{w}\|_{H^{k+2}(\Sigma)} \qquad \forall\,\mathbf{w} \in H^k(\Sigma)^d,\ \forall k \ge -1,\ k \in \mathbb{R}, \tag{3.2.3}$$

$$a_s(\eta,\mathbf{w}) = a_s(\mathbf{w},\eta) \quad \text{and} \quad a_s(\eta,\eta) \ge 0 \qquad \forall \eta \in H^1(\Sigma)^d, \tag{3.2.4}$$

$$\|\eta\|_s + \|\eta\|_\Sigma \sim \|\eta\|_{H^1(\Sigma)} \qquad \text{for } \|\eta\|_s := \sqrt{a_s(\eta,\eta)}. \tag{3.2.5}$$

In addition, we denote $\|\mathbf{u}\|_f := \sqrt{(\mathbf{D}(\mathbf{u}),\mathbf{D}(\mathbf{u}))}$ and mention that the following norm equivalence holds (according to Korn's inequality):

$$\|\mathbf{u}\|_f + \|\mathbf{u}\| \sim \|\mathbf{u}\|_{H^1(\Omega)}.$$

For the simplicity of notations, we denote by $\|\mathbf{v}\|_{L^pX}$ the Bochner norm (or semi-norm) defined by

$$\|\mathbf{v}\|_{L^pX} := \begin{cases} \left(\int_{t=0}^{t=T} \|\mathbf{v}(t,\cdot)\|_X^p dt\right)^{1/p} & 1 \le p < \infty \\ \sup_{t\in[0,T]} \|\mathbf{v}(t,\cdot)\|_X & p = \infty, \end{cases}$$

where $\|\cdot\|_X$ is any norm or semi-norm in space, such as $\|\cdot\|_f$, $\|\cdot\|_s$ or $\|\cdot\|_{L^2(\Sigma)}$. The following conventional notations will be used: $\|\cdot\|_X := \|\cdot\|_{X(\Omega)}$, $\|\cdot\| := \|\cdot\|_{L^2(\Omega)}$, $\|\cdot\|_\Sigma := \|\cdot\|_{L^2(\Sigma)}$ and $\|\cdot\|_f := \|\cdot\|_{H_f}$, $\|\cdot\|_s := \|\cdot\|_{H_s}$.

For smooth solutions of (3.1.1)–(3.1.3), one can verify that (via integration by parts) the following equations hold for all test functions $(\mathbf{v}, q, \mathbf{w}) \in \mathbf{X} \times Q \times \mathbf{S}$ with $\mathbf{v}|_\Sigma = \mathbf{w}$:

$$\begin{aligned} &\partial_t \eta = \mathbf{u} \qquad \text{on } \Sigma, \\ &\rho_f(\partial_t \mathbf{u}, \mathbf{v}) + a_f(\mathbf{u}, \mathbf{v}) - b(p, \mathbf{v}) + b(q, \mathbf{u}) + \rho_s \epsilon_s (\partial_{tt}\eta, \mathbf{w})_\Sigma + a_s(\eta, \mathbf{w}) = 0. \end{aligned} \tag{3.2.6}$$

3.2.2 Regularity assumptions

To establish the optimal error estimates for the finite element solutions to the thin-structure interaction problem, we need to use the following regularity results.

- We assume that the domain Ω is smooth so that the the solution $(\mathbf{u}, p, \eta)$ of the fluid-structure interaction problem (3.1.1)–(3.1.3) is sufficiently smooth.
- The weak solution $(\omega, \lambda) \in H^1(\Omega)^d \times L^2(\Omega)$ of the Stokes equations

$$\begin{aligned} -\nabla \cdot \sigma(\omega, \lambda) + \omega &= \mathbf{f} \\ \nabla \cdot \omega &= 0 \end{aligned}$$

  has the following regularity estimates:

$$\|\omega\|_{H^{k+3/2}} + \|\lambda\|_{H^{k+1/2}} \le C\|\mathbf{f}\|_{H^{k-1/2}} + \|\sigma(\omega, \lambda)\cdot \mathbf{n}\|_{H^k(\Sigma)} \quad \text{for } k \ge -1/2, \quad k \in \mathbb{R}, \tag{3.2.7}$$

$$\|\omega\|_{H^{k+1/2}} + \|\lambda - \bar{\lambda}\|_{H^{k-1/2}} \le C\|\mathbf{f}\|_{H^{k-3/2}} + \|\omega\|_{H^k(\Sigma)} \quad \text{for } k \ge 1/2, \quad k \in \mathbb{R}, \tag{3.2.8}$$

  where $\bar{\lambda} := \frac{1}{|\Omega|}\int_\Omega \lambda$ is the mean value of λ over Ω. The estimates in (3.2.7) and (3.2.8) correspond to the Neumann and Dirichlet boundary conditions, respectively; see [61, Theorem IV.6.1] for a proof of (3.2.8) in smooth domains, with a similar approach as in [61, Chapter IV] one can prove (3.2.7). We also refer to [74, Theorem 4.15] for a proof of (3.2.7) in the case of polygonal domain.

- We assume that operator $\mathcal{L}_s$ possesses the following elliptic regularity: The weak solution $\boldsymbol{\xi} \in H^1(\Sigma)^d$ of the equation (in the weak formulation)

$$a_s(\boldsymbol{\xi}, \mathbf{w}) + (\boldsymbol{\xi}, \mathbf{w})_\Sigma = (\mathbf{g}, \mathbf{w})_\Sigma \quad \forall \mathbf{w} \in H^1(\Sigma)^d,$$

  has the following regularity estimate:

$$\|\boldsymbol{\xi}\|_{H^{2+k}(\Sigma)} \le C\|\mathbf{g}\|_{H^k(\Sigma)} \quad \text{for } k \ge -1, \ k \in \mathbb{R}. \tag{3.2.9}$$

3.2.3 Assumptions on the finite element spaces

Let $\mathcal{T}_h$ denote a quasi-uniform partition on Ω with $\overline{\Omega} = \bigcup_{K\in\mathcal{T}_h} K$. Each K is a curvilinear polyhedron/polygon with $\mathrm{diam}(K) \le h$. All boundary faces of $\mathcal{T}_h$ on Σ form a partition $\mathcal{T}_h(\Sigma)$, $\Sigma = \bigcup_{D\in\mathcal{T}_h(\Sigma)} D$. All boundary faces of $\mathcal{T}_h$ on Σ_l or Σ_r form a partition for Σ_l or Σ_r, respectively, and these two partitions coincide after shifting L in z-direction. To approximate the weak form (3.2.6) by finite element method, we assume that there are finite element spaces $(\mathbf{X}_h^r, \mathbf{S}_h^r, Q_h^{r-1})$ on $\mathcal{T}_h$ (where $r \ge 1$) with the following properties.

- **(A1)** $\mathbf{X}_h^r \subseteq \mathbf{X}$, $\mathbf{S}_h^r \subseteq \mathbf{S}$ and $\mathbb{R} \subseteq Q_h^{r-1} \subseteq Q$, with $\mathbf{S}_h^r = \{\mathbf{v}_h|_\Sigma : \mathbf{v}_h \in \mathbf{X}_h^r\}$.
- **(A2)** For $\mathbf{X}_h^r$ and Q_h^{r-1}, the following local inverse estimate holds on each $K \in \mathcal{T}_h$ for $0 \le l \le k, 1 \le p, q \le \infty$:

$$\|\mathbf{v}_h\|_{W^{k,p}(K)} \le Ch^{-(k-l)+(d/p-d/q)}\|\mathbf{v}_h\|_{W^{l,q}(K)} \quad \forall \mathbf{v}_h \in \mathbf{X}_h^r \text{ or } Q_h^{r-1}, \tag{3.2.10}$$

  For $\mathbf{S}_h^r$, the following global inverse estimate holds:

$$\|\mathbf{w}_h\|_{H^s(\Sigma)} \le Ch^{k-s}\|\mathbf{w}_h\|_{H^k(\Sigma)} \quad \forall \mathbf{w}_h \in \mathbf{S}_h^r;\ \forall\, k, s \in \mathbb{R} \text{ with } 0 \le k \le s \le 1. \tag{3.2.11}$$

- **(A3)** There are interpolation/projection operators $I_h^X : \mathbf{X} \to \mathbf{X}_h^r$ and $I_h^Q : Q \to Q_h^{r-1}$ which have the following local L^p approximation properties on each $K \in \mathcal{T}_h$, for all $1 \le p \le \infty$:

$$\|I_h^X\mathbf{u} - \mathbf{u}\|_{L^p(K)} + h\|I_h^X\mathbf{u} - \mathbf{u}\|_{W^{1,p}(K)} \le Ch^{k+1}\|\mathbf{u}\|_{W^{k+1,p}(\Delta_K)} \quad \forall\, 0 \le k \le r, \tag{3.2.12a}$$

$$\|I_h^Q p - p\|_{L^p(K)} \le Ch^{k+1}\|p\|_{W^{k+1,p}(\Delta_K)} \quad \forall\, 0 \le k \le r-1, \tag{3.2.12b}$$

  where Δ_K is the macro element including all the elements which have a common vertex with K. And there is an interpolation/projection operator $I_h^S : \mathbf{S} \to \mathbf{S}_h^r$ satisfying $(I_h^X\mathbf{u})|_\Sigma = I_h^S(\mathbf{u}|_\Sigma)$ for all $\mathbf{u} \in \mathbf{X}$ with $\mathbf{u}|_\Sigma \in \mathbf{S}$. Moreover, we require the following optimal order error estimate

$$\|I_h^S\mathbf{w} - \mathbf{w}\|_\Sigma + h\|I_h^S\mathbf{w} - \mathbf{w}\|_{H^1(\Sigma)} \le Ch^{k+1}\|\mathbf{w}\|_{H_h^{k+1}(\Sigma)} \quad \forall 0 \le k \le r, \tag{3.2.13}$$

  where $\|\cdot\|_{H_h^{k+1}(\Sigma)}$ is the piecewise H^{k+1}-norm associated with partition $\mathcal{T}_h(\Sigma)$. We will use I_h to denote one of the operators I_h^X, I_h^S and I_h^Q when there is no confusion.

- **(A4)** Let $\mathring{\mathbf{X}}_h^r := \{\mathbf{v}_h \in \mathbf{X}_h^r : \mathbf{v}_h|_\Sigma = 0\}$ and $Q_{h,0}^{r-1} := \{q_h \in Q_h^{r-1} : q_h \in L_0^2(\Omega)\}$. The following inf-sup condition holds:

$$\|q_h\| \le C \sup_{0 \ne \mathbf{v}_h \in \mathring{\mathbf{X}}_h^r} \frac{(\operatorname{div}\mathbf{v}_h, q_h)}{\|\mathbf{v}_h\|_{H^1}} \quad \forall q_h \in Q_{h,0}^{r-1} \tag{3.2.14}$$

Remark 3.2.1. Examples of finite element spaces which satisfy Assumptions (A1)–(A4) include the Taylor–Hood finite element space with I_h^X, I_h^Q and I_h^S being the Scott–Zhang interpolation operators onto $\mathbf{X}_h^r$, Q_h^{r-1} and $\mathbf{S}_h^r$ respectively. We refer to [21, Section 4.8] and the references therein for the details on construction and properties of Scott-Zhang interpolation, and refer to [16, Section 8.8] for a proof of (3.2.14) for the Taylor-Hood finite element spaces. The following properties are consequences of the assumptions **(A1)**–**(A4)**.

1. From **(A2)** and **(A3)** we can derive the following estimate for $\mathbf{v}_h \in \mathbf{X}_h^r$:

$$\|\mathbf{D}(\mathbf{v}_h)\mathbf{n}\|_\Sigma = \Big(\sum_{D \in \mathcal{T}_h(\Sigma)} \|\mathbf{D}(\mathbf{v}_h)\mathbf{n}\|_{L^2(D)}^2\Big)^{1/2}$$

$$\le C\Big(\sum_{D\in\mathcal{T}_h(\Sigma)} h^{d-1}\|\mathbf{v}_h\|^2_{W^{1,\infty}(K)}\Big)^{1/2} \quad (K\in\mathcal{T}_h \text{ contains } D)$$
$$\le C\Big(\sum_{D\in\mathcal{T}_h(\Sigma)} h^{-1}\|\mathbf{v}_h\|^2_{H^1(K)}\Big)^{1/2} \le Ch^{-1/2}\|\mathbf{v}_h\|_{H^1}.$$

Therefore, we can obtain the following inverse estimate for the boundary term $\boldsymbol{\sigma}(\mathbf{v}_h, q_h)\mathbf{n}$:

$$\|\boldsymbol{\sigma}(\mathbf{v}_h, q_h)\mathbf{n}\|_\Sigma \le Ch^{-1/2}(\|\mathbf{v}_h\|_{H^1} + \|q_h\|). \tag{3.2.15}$$

2. From **(A3)** and **(A4)** we can see that when $r\ge 2$, the mixed finite element space $(\mathbf{X}_h^r, Q_h^{r-1})$ can be realized by the $(r, r-1)$ Taylor-Hood finite element space. When $r=1$, $(\mathbf{X}_h^1, Q_h^0)$ can be realized by the MINI element space.

3. From inf-sup condition (3.2.14), we can deduce the following alternative version of inf-sup condition (involving $H^1(\Sigma)$-norm in the denominator)

$$\|q_h\|\le C \sup_{0\ne\mathbf{v}_h\in\mathbf{X}_h^r} \frac{(\operatorname{div}\mathbf{v}_h, q_h)}{\|\mathbf{v}_h\|_{H^1} + \|\mathbf{v}_h\|_{H^1(\Sigma)}} \quad \forall q_h\in Q_h^{r-1}. \tag{3.2.16}$$

An inf-sup condition similar to (3.2.16) was proved in [141, Lemma 2], though thick structure problem is considered there. For the reader's convenience, we present a proof of (3.2.16) in Section 3.9.

4. For each $\mathbf{w}_h\in\mathbf{S}_h^r$, we denote by $E_h\mathbf{w}_h\in\mathbf{X}_h^r$ an extension such that $E_h\mathbf{w}_h := I_h^X\mathbf{v}$, where $\mathbf{v}\in H^1(\Omega)^d$ is the extension of $\mathbf{w}_h$ by trace theorem, satisfying $\|\mathbf{v}\|_{H^1}\le C\|\mathbf{w}_h\|_{H^{1/2}(\Sigma)}$ and $\mathbf{v}|_\Sigma = \mathbf{w}_h$. Combining (3.2.12) with (3.2.11) we see that

$$\|E_h\mathbf{w}_h\|_{H^1} \le Ch^{-1/2}\|\mathbf{w}_h\|_\Sigma. \tag{3.2.17}$$

5. Combining (3.2.12) with (3.2.15) we have for any $\mathbf{u}_h\in\mathbf{X}_h^r$, $p_h\in Q_h^{r-1}$

$$\begin{aligned}
&\|\boldsymbol{\sigma}(\mathbf{u}-\mathbf{u}_h, p-p_h)\mathbf{n}\|_\Sigma\\
&\le \|\boldsymbol{\sigma}(\mathbf{u}-I_h\mathbf{u}, p-I_hp)\mathbf{n}\|_\Sigma + \|\boldsymbol{\sigma}(I_h\mathbf{u}-\mathbf{u}_h, I_hp-p_h)\mathbf{n}\|_\Sigma\\
&\le C(\|\mathbf{u}-I_h\mathbf{u}\|_{W^{1,\infty}} + \|p-I_hp\|_{L^\infty}) + \|\boldsymbol{\sigma}(I_h\mathbf{u}-\mathbf{u}_h, I_hp-p_h)\mathbf{n}\|_\Sigma\\
&\le Ch^r + Ch^{-1/2}(\|I_h\mathbf{u}-\mathbf{u}_h\|_{H^1} + \|I_hp-p_h\|)\\
&\le Ch^{r-1/2} + Ch^{-1/2}(\|\mathbf{u}-\mathbf{u}_h\|_{H^1} + \|p-p_h\|),
\end{aligned} \tag{3.2.18}$$

where we have used (3.2.12) with $p=\infty$ and (3.2.15) in the second to last inequality.

3.2.4 A new kinematically coupled scheme and main theoretical results

Let $\{t_n\}_{n=0}^N$ be a uniform partition of the time interval $[0,T]$ with stepsize $\tau = T/N$. For a sequence of functions $\{\mathbf{u}^n\}_{n=0}^N$ we denote

$$D_\tau\mathbf{u}^n = \frac{\mathbf{u}^n - \mathbf{u}^{n-1}}{\tau}, \quad \text{for } n = 1,\, 2,\, \ldots,\, N.$$

With the above notations, we present a fully discrete kinematically coupled algorithm.

Step 1: For given $\mathbf{u}_h^{n-1}, p_h^{n-1}, \eta_h^{n-1}$, find η_h^n and $\mathbf{s}_h^n \in \mathbf{S}_h^r$ such that

$$\rho_s\epsilon_s\left(\frac{\mathbf{s}_h^n - \mathbf{u}_h^{n-1}}{\tau}, \mathbf{w}_h\right)_\Sigma + a_s(\eta_h^n, \mathbf{w}_h) = -(\boldsymbol{\sigma}_h^{n-1}\cdot\mathbf{n}, \mathbf{w}_h)_\Sigma, \qquad \forall \mathbf{w}_h \in \mathbf{S}_h^r \tag{3.2.19}$$
$$\eta_h^n = \eta_h^{n-1} + \tau\mathbf{s}_h^n.$$

Step 2: Then find $(\mathbf{u}_h^n, p_h^n) \in \mathbf{X}_h^r \times Q_h^{r-1}$ satisfying

$$\begin{aligned}&\rho_f(D_\tau\mathbf{u}_h^n, \mathbf{v}_h) + a_f(\mathbf{u}_h^n, \mathbf{v}_h) - b(p_h^n, \mathbf{v}_h) + b(q_h, \mathbf{u}_h^n) - (\boldsymbol{\sigma}_h^n\cdot\mathbf{n}, \mathbf{v}_h)_\Sigma \\ &+ \rho_s\epsilon_s\left(\frac{\mathbf{u}_h^n - \mathbf{s}_h^n}{\tau}, \mathbf{v}_h + \frac{\tau}{\rho_s\epsilon_s}\boldsymbol{\sigma}(\mathbf{v}_h, q_h)\cdot\mathbf{n}\right)_\Sigma \\ &+ \left((\boldsymbol{\sigma}_h^n - \boldsymbol{\sigma}_h^{n-1})\cdot\mathbf{n}, \mathbf{v}_h + \frac{\tau(1+\beta)}{\rho_s\epsilon_s}\boldsymbol{\sigma}(\mathbf{v}_h, q_h)\cdot\mathbf{n}\right)_\Sigma = 0\end{aligned} \tag{3.2.20}$$

for all $(\mathbf{v}_h, q_h) \in \mathbf{X}_h^r \times Q_h^{r-1}$, where $\boldsymbol{\sigma}_h^n = \boldsymbol{\sigma}(\mathbf{u}_h^n, p_h^n)$ and $\beta \geq 0$ denotes a stabilization parameter.

Initial values: Since $\boldsymbol{\sigma}_h^{n-1}$ depends on both $\mathbf{u}_h^{n-1}$ and p_h^{n-1}, the numerical scheme in (3.2.19)–(3.2.20) requires the initial value $(\mathbf{u}_h^0, p_h^0, \eta_h^0)$ to be given. We simply assume that the initial value $(\mathbf{u}_h^0, p_h^0, \eta_h^0)$ are given sufficiently accurately, satisfying the following conditions:

$$\begin{aligned}\|\mathbf{u}_h^0 - R_h\mathbf{u}^0\| + \|\mathbf{u}_h^0 - R_h\mathbf{u}^0\|_\Sigma + \|\eta_h^0 - R_h\eta^0\|_{H^1(\Sigma)} &\leqslant Ch^{r+1}, \\ \|p_h^0 - R_hp^0\|_\Sigma &\leqslant C,\end{aligned} \tag{3.2.21}$$

where $(R_h\mathbf{u}^0, R_hp^0, R_h\eta^0)$ satisfies a coupled non-stationary Ritz projection defined in Section 3.3.2.

Remark 3.2.2. Kinematically coupled schemes were firstly proposed in [23, 72, 25] with the following time discretization: Find $(\mathbf{s}^n, \eta^n)$ such that

$$\begin{aligned}&\rho_s\epsilon_s\frac{\mathbf{s}^n - \mathbf{u}^{n-1}}{\tau} - \mathcal{L}_s(\eta^n) = -\boldsymbol{\sigma}^{n-1}\cdot\mathbf{n} && \text{on } \Sigma \\ &\eta^n = \eta^{n-1} + \tau\mathbf{s}^n && \text{on } \Sigma\end{aligned} \tag{3.2.22}$$

and then find $(\mathbf{u}^n, p^n)$ satisfying

$$\begin{aligned}&\rho_f D_\tau\mathbf{u}^n + \nabla\cdot\boldsymbol{\sigma}^n = 0 \quad \text{and} \quad \nabla\cdot\mathbf{u}^n = 0 && \text{in } \Omega, \\ &\rho_s\epsilon_s\frac{\mathbf{u}^n - \mathbf{s}^n}{\tau} + (\boldsymbol{\sigma}^n - \boldsymbol{\sigma}^{n-1})\cdot\mathbf{n} = 0 && \text{on } \Sigma.\end{aligned} \tag{3.2.23}$$

The extension to full discretization was considered by several authors [25, 120], while the analysis for full discretization is incomplete and the energy stability is proved only for time-discrete schemes.

Remark 3.2.3. Our scheme in (3.2.19)–(3.2.20) is designed with two new ingredients. First, we have added two stabilization terms

$$\rho_s\epsilon_s\left(\frac{\mathbf{u}_h^n - \mathbf{s}_h^n}{\tau}, \frac{\tau}{\rho_s\epsilon_s}\boldsymbol{\sigma}(\mathbf{v}_h, q_h)\cdot\mathbf{n}\right)_\Sigma \quad \text{and} \quad \left((\boldsymbol{\sigma}_h^n - \boldsymbol{\sigma}_h^{n-1})\cdot\mathbf{n}, \frac{\tau(1+\beta)}{\rho_s\epsilon_s}\boldsymbol{\sigma}(\mathbf{v}_h, q_h)\cdot\mathbf{n}\right)_\Sigma,$$

which guarantee unconditional energy stability of the scheme in (3.2.19)–(3.2.20). Otherwise the unconditional energy stability cannot be proved in the fully discrete finite element setting. Second, we have introduced an additional parameter $\beta > 0$ to the scheme, and this additional parameter allows us to prove optimal-order convergence in the L^2 norm (especially optimal order in space). More specifically, this parameter $\beta > 0$ leads to the following term in the E_1 of (3.2.26) :

$$\beta_0 \frac{\rho_s \epsilon_s}{2\tau} \|\mathbf{s}_h^n - \mathbf{u}_h^n\|_\Sigma^2 \quad \text{with } \beta_0 = 1 - (\sqrt{4+\beta^2} - \beta)/2,$$

which is used to absorb other undesired terms on the right-hand side of the inequalities in our error estimation. Therefore, the optimal-order L^2 error estimate does benefits from our scheme (with the parameter $\beta > 0$).

Remark 3.2.4. For the Taylor–Hood finite element spaces, the conditions in (3.2.21) on the initial values can be satisfied if one chooses $\mathbf{u}_h^0$ and p_h^0 to be the Lagrange interpolations of $\mathbf{u}^0$ and p^0, respectively, and chooses $\boldsymbol{\eta}_h^0 = R_{sh}\boldsymbol{\eta}(0)$, where $R_{sh}\boldsymbol{\eta}(0)$ is defined in Section 3.4; see Definition 3.4.4 and estimate (5.3.44).

The main theoretical results of this chapter are the following two theorems.

Theorem 3.2.1. *Under the assumptions in Section 3.2.3 (on the finite element spaces), the finite element system in (3.2.19)–(3.2.20) is uniquely solvable, and the following inequality holds:*

$$E_0(\mathbf{u}_h^n, p_h^n, \boldsymbol{\eta}_h^n) + \sum_{m=1}^{n} \tau E_1(\mathbf{u}_h^m, \mathbf{s}_h^m, \boldsymbol{\eta}_h^m) \leqslant E_0(\mathbf{u}_h^0, p_h^0, \boldsymbol{\eta}_h^0), \quad n = 1, 2, ..., N, \tag{3.2.24}$$

where

$$E_0(\mathbf{u}_h^n, p_h^n, \boldsymbol{\eta}_h^n) = \frac{\rho_f}{2}\|\mathbf{u}_h^n\|^2 + \frac{1}{2}\|\boldsymbol{\eta}_h^n\|_s^2 + \frac{\tau^2(1+\beta)}{2\rho_s\epsilon_s}\|\boldsymbol{\sigma}_h^n \cdot \mathbf{n}\|_\Sigma^2 + \frac{\rho_s\epsilon_s}{2}\|\mathbf{u}_h^n\|_\Sigma^2, \tag{3.2.25}$$

$$\begin{aligned} E_1(\mathbf{u}_h^n, \mathbf{s}_h^n, \boldsymbol{\eta}_h^n) = {} & 2\mu\|\mathbf{u}_h^n\|_f^2 + \frac{\rho_f}{2\tau}\|\mathbf{u}_h^n - \mathbf{u}_h^{n-1}\|^2 + \frac{\rho_s\epsilon_s}{2\tau}\|\mathbf{s}_h^n - \mathbf{u}_h^{n-1}\|_\Sigma^2 + \frac{\rho_s\epsilon_s\beta_0}{2\tau}\|\mathbf{s}_h^n - \mathbf{u}_h^n\|_\Sigma^2 \\ & + \frac{\tau\beta_0}{2\rho_s\epsilon_s}\|(\boldsymbol{\sigma}_h^n - \boldsymbol{\sigma}_h^{n-1})\cdot\mathbf{n}\|_\Sigma^2 + \frac{\tau}{2}\|D_\tau\boldsymbol{\eta}_h^n\|_s^2, \end{aligned} \tag{3.2.26}$$

with $\beta_0 = 1 - (\sqrt{4+\beta^2} - \beta)/2$ *and* $\beta \geqslant 0$.

Theorem 3.2.2. *For finite elements of degree $r \geqslant 2$, under the assumptions in Sections 3.2.2–3.2.3 (on the regularity of solutions and finite element spaces), there exist positive constants τ_0 and h_0 such that, for sufficiently small stepsize and mesh size $\tau \leqslant \tau_0$ and $h \leqslant h_0$, the finite element solutions given by (3.2.19)–(3.2.20) with initial values satisfying* (3.2.21) *and $\beta > 0$ has the following error bound:*

$$\max_{1\leqslant n\leqslant N}\left(\|\mathbf{u}(t_n,\cdot) - \mathbf{u}_h^n\| + \|\boldsymbol{\eta}(t_n,\cdot) - \boldsymbol{\eta}_h^n\|_\Sigma + \|\mathbf{u}(t_n,\cdot) - \mathbf{u}_h^n\|_\Sigma\right) \leqslant C(\tau + h^{r+1}), \tag{3.2.27}$$

where C is some positive constant independent of n, h and τ.

The proofs of Theorem 3.2.1 and Theorem 3.2.2 are presented in the next section.

3.3 Analysis of the proposed algorithm

This section is devoted to the proof of Theorems 3.2.1 and 3.2.2. For the simplicity of notation, we denote by C a generic positive constant, which is independent of n, h and τ but may depend on the physical parameters $\rho_s, \epsilon, \mu, \rho_f$ and the exact solution $(\mathbf{u}, p, \boldsymbol{\eta})$.

3.3.1 Proof of Theorem 3.2.1

We rewrite (3.2.20) into

$$\rho_f(D_\tau \mathbf{u}_h^n,\, \mathbf{v}_h) + a_f(\mathbf{u}_h^n,\, \mathbf{v}_h) - b(p_h^n,\, \mathbf{v}_h^n) + b(q_h,\, \mathbf{u}_h^n) + \rho_s\epsilon_s \left(\frac{\mathbf{u}_h^n - \mathbf{s}_h^n}{\tau},\, \mathbf{v}_h\right)_\Sigma \tag{3.3.1}$$
$$= (\boldsymbol{\sigma}_h^{n-1}\cdot \mathbf{n},\, \mathbf{v}_h)_\Sigma - (\mathbf{u}_h^n - \mathbf{s}_h^n,\, \boldsymbol{\sigma}(\mathbf{v}_h, q_h)\cdot \mathbf{n})_\Sigma - \frac{\tau(1+\beta)}{\rho_s\epsilon_s}((\boldsymbol{\sigma}_h^n - \boldsymbol{\sigma}_h^{n-1})\cdot \mathbf{n},\, \boldsymbol{\sigma}(\mathbf{v}_h, q_h)\cdot\mathbf{n})_\Sigma\,.$$

Taking $\mathbf{v}_h = \mathbf{u}_h^n, q_h = p_h^n$ in (3.3.1) and $\mathbf{w}_h = \mathbf{s}_h^n = D_\tau \eta_h^n$ in (3.2.19), respectively, gives the following relations:

$$\frac{\rho_f}{2\tau}\left(\|\mathbf{u}_h^n\|^2 - \|\mathbf{u}_h^{n-1}\|^2 + \|\mathbf{u}_h^n - \mathbf{u}_h^{n-1}\|^2\right) + 2\mu\|\mathbf{u}_h^n\|_f^2 + \rho_s\epsilon_s\left(\frac{\mathbf{u}_h^n - \mathbf{s}_h^n}{\tau},\, \mathbf{u}_h^n\right)_\Sigma$$
$$= (\boldsymbol{\sigma}_h^{n-1}\cdot\mathbf{n},\, \mathbf{u}_h^n)_\Sigma - (\mathbf{u}_h^n - \mathbf{s}_h^n,\, \boldsymbol{\sigma}_h^n\cdot\mathbf{n})_\Sigma - \frac{\tau(1+\beta)}{\rho_s\epsilon_s}((\boldsymbol{\sigma}_h^n - \boldsymbol{\sigma}_h^{n-1})\cdot\mathbf{n},\, \boldsymbol{\sigma}_h^n\cdot\mathbf{n})_\Sigma$$

and

$$\frac{1}{2\tau}\left(a_s(\eta_h^n, \eta_h^n) - a_s(\eta_h^{n-1}, \eta_h^{n-1}) + \tau^2 a_s(\mathbf{s}_h^n, \mathbf{s}_h^n)\right) + \rho_s\epsilon_s\left(\frac{\mathbf{s}_h^n - \mathbf{u}_h^{n-1}}{\tau},\, \mathbf{s}_h^n\right)_\Sigma = -(\boldsymbol{\sigma}_h^{n-1}\cdot\mathbf{n},\, \mathbf{s}_h^n)_\Sigma\,.$$

By summing up the last two equations, we have

$$\frac{\rho_f}{2}\left(\|\mathbf{u}_h^n\|^2 - \|\mathbf{u}_h^{n-1}\|^2 + \|\mathbf{u}_h^n - \mathbf{u}_h^{n-1}\|^2\right) + 2\mu\tau\|\mathbf{u}_h^n\|_f^2 + \frac{\rho_s\epsilon_s}{2}\left(\|\mathbf{s}_h^n - \mathbf{u}_h^{n-1}\|_\Sigma^2 + \|\mathbf{u}_h^n - \mathbf{s}_h^n\|_\Sigma^2\right)$$
$$+ \frac{1}{2}\left(a_s(\eta_h^n, \eta_h^n) - a_s(\eta_h^{n-1}, \eta_h^{n-1}) + \tau^2 a_s(\mathbf{s}_h^n, \mathbf{s}_h^n)\right) + \frac{\rho_s\epsilon_s}{2}\left(\|\mathbf{u}_h^n\|_\Sigma^2 - \|\mathbf{u}_h^{n-1}\|_\Sigma^2\right)$$
$$= \tau((\boldsymbol{\sigma}_h^{n-1} - \boldsymbol{\sigma}_h^n)\cdot\mathbf{n}, \mathbf{u}_h^n - \mathbf{s}_h^n)_\Sigma - \frac{\tau^2(1+\beta)}{\rho_s\epsilon_s}((\boldsymbol{\sigma}_h^n - \boldsymbol{\sigma}_h^{n-1})\cdot\mathbf{n},\, \boldsymbol{\sigma}_h^n\cdot\mathbf{n})_\Sigma$$
$$\leqslant \frac{\tau^2(1+\beta-\beta_0)}{2\rho_s\epsilon_s}\|(\boldsymbol{\sigma}_h^n - \boldsymbol{\sigma}_h^{n-1})\cdot\mathbf{n}\|_\Sigma^2 + \frac{\rho_s\epsilon_s}{2(1+\beta-\beta_0)}\|\mathbf{u}_h^n - \mathbf{s}_h^n\|_\Sigma^2$$
$$- \frac{\tau^2(1+\beta)}{2\rho_s\epsilon_s}\left(\|\boldsymbol{\sigma}_h^n\cdot\mathbf{n}\|_\Sigma^2 - \|\boldsymbol{\sigma}_h^{n-1}\cdot\mathbf{n}\|_\Sigma^2 + \|(\boldsymbol{\sigma}_h^n - \boldsymbol{\sigma}_h^{n-1})\cdot\mathbf{n}\|_\Sigma^2\right)$$
$$\leqslant \frac{\rho_s\epsilon_s(1-\beta_0)}{2}\|\mathbf{u}_h^n - \mathbf{s}_h^n\|_\Sigma^2 - \frac{\tau^2(1+\beta)}{2\rho_s\epsilon_s}\left(\|\boldsymbol{\sigma}_h^n\cdot\mathbf{n}\|_\Sigma^2 - \|\boldsymbol{\sigma}_h^{n-1}\cdot\mathbf{n}\|_\Sigma^2\right) - \frac{\tau^2\beta_0}{2\rho_s\epsilon_s}\|(\boldsymbol{\sigma}_h^n - \boldsymbol{\sigma}_h^{n-1})\cdot\mathbf{n}\|_\Sigma^2,$$

which leads to the following energy inequality:

$$E_0(\mathbf{u}_h^n, p_h^n, \eta_h^n) - E_0(\mathbf{u}_h^{n-1}, p_h^{n-1}, \eta_h^{n-1}) + E_1(\mathbf{u}_h^n, p_h^n, \eta_h^n)\tau \leqslant 0\,. \tag{3.3.2}$$

This implies (3.2.24) and completes the proof of Theorem 3.2.1. ∎

3.3.2 A coupled non-stationary Ritz projection

To establish L^2-norm optimal error estimate as given in Theorem 5.2.2, we need to introduce a new coupled Ritz projection. Since the FSI model is governed by the Stokes type equation for fluid coupled with the hyperbolic type equation for solid, the coupled projection, which is non-stationary and much more complicated than the standard Ritz projections, plays a key role in proving the optimal-order convergence of finite element solutions to the FSI model.

Definition 3.3.1 (Coupled non-stationary Ritz projection). Let $(\mathbf{u}, p, \eta) \in \mathbf{X} \times Q \times \mathbf{S}$ be a triple of functions smoothly depending on $t \in [0, T]$ and satisfying the condition $\mathbf{u}|_\Sigma = \partial_t \eta$. For a given initial value $R_h\eta(0)$, the coupled Stokes–Ritz projection $R_h(\mathbf{u}, p, \eta)$ is defined as a triple of functions $(R_h\mathbf{u}, R_hp, R_h\eta) \in \mathbf{X}_h^r \times Q_h^{r-1} \times \mathbf{S}_h^r$ satisfying $(R_h\mathbf{u})|_\Sigma = \partial_t R_h\eta$ and the following weak formulation for every $t \in [0, T]$:

$$
\begin{aligned}
a_f(\mathbf{u} - R_h\mathbf{u}, \mathbf{v}_h) - b(p - R_hp, \mathbf{v}_h) + b(q_h, \mathbf{u} - R_h\mathbf{u}) + (\mathbf{u} - R_h\mathbf{u}, \mathbf{v}_h) \\
+ a_s(\eta - R_h\eta, \mathbf{v}_h) + (\eta - R_h\eta, \mathbf{v}_h)_\Sigma = 0, \qquad \forall(\mathbf{v}_h, q_h) \in \mathbf{X}_h^r \times Q_h^{r-1}.
\end{aligned} \tag{3.3.3}
$$

Remark 3.3.1. Given an initial value $R_h\eta(0)$, there exists a unique solution $(R_h\mathbf{u}, R_hp, R_h\eta_h)$ for the finite element semi-discrete problem (3.3.3). To see this, we firstly introduce a linear operator $\mathcal{S}_h : (\mathbf{X}_h^r)^* \times (Q_h^{r-1})^* \to \mathbf{X}_h^r \times Q_h^{r-1}$, where $(\mathbf{X}_h^r)^*$ and $(Q_h^{r-1})^*$ denote the dual space of $\mathbf{X}_h^r$ and Q_h^{r-1}, respectively. For a given $(\phi, \ell) \in (\mathbf{X}_h^r)^* \times (Q_h^{r-1})^*$, denote by $(\mathbf{u}_h, p_h) \in \mathbf{X}_h^r \times Q_h^{r-1}$ the solution of the following Neumann-type discrete Stokes equation

$$
\begin{aligned}
&a_f(\mathbf{u}_h, \mathbf{v}_h) - b(p_h, \mathbf{v}_h) + (\mathbf{u}_h, \mathbf{v}_h) = \phi(\mathbf{v}_h) \quad \forall \mathbf{v}_h \in \mathbf{X}_h^r, \\
&b(q_h, \mathbf{u}_h) = \ell(q_h) \quad \forall q_h \in Q_h^{r-1},
\end{aligned}
$$

and define $\mathcal{S}_h(\phi, \ell) = (\mathcal{S}_h^v(\phi, \ell), \mathcal{S}_h^p(\phi, \ell)) := (\mathbf{u}_h, p_h)$. The well-posedness of the above equation follows the inf-sup condition (3.2.16).

Next, we denote

$$
\begin{aligned}
\phi_{(u,p,\eta)}(\mathbf{v}_h) &:= a_f(\mathbf{u}, \mathbf{v}_h) - b(p, \mathbf{v}_h) + (\mathbf{u}, \mathbf{v}_h) + a_s(\eta, \mathbf{v}_h) + (\eta, \mathbf{v}_h)_\Sigma, \\
\phi_{R_h\eta}(\mathbf{v}_h) &:= a_s(R_h\eta, \mathbf{v}_h) + (R_h\eta, \mathbf{v}_h)_\Sigma, \\
\ell_u(q_h) &:= b(q_h, \mathbf{u}).
\end{aligned}
$$

Then $(R_h\mathbf{u}, R_hp, R_h\eta)$ is a solution to (3.3.3) if and only if the following equations are satisfied:

$$
\partial_t R_h\eta = \mathcal{S}_h^v(\phi_{(u,p,\eta)} - \phi_{R_h\eta}, \ell_u)|_\Sigma, \tag{3.3.4a}
$$

$$
R_h\mathbf{u} = \mathcal{S}_h^v(\phi_{(u,p,\eta)} - \phi_{R_h\eta}, \ell_u), \quad R_hp = \mathcal{S}_h^p(\phi_{(u,p,\eta)} - \phi_{R_h\eta}, \ell_u). \tag{3.3.4b}
$$

Therefore, the uniqueness and existence of solution to (3.3.3) follows the uniqueness and existence of solution to (3.3.4a). Since $\mathcal{S}_h^v$ is a linear operator on $(\mathbf{X}_h^r)^* \times (Q_h^{r-1})^*$ and $\phi_{R_h\eta}$ is linear with respect to $R_h\eta$, (3.3.4a) is an in-homogeneous linear ordinary differential equation for $R_h\eta$ and thus admits a unique solution for a given initial value $R_h\eta(0)$. Next, we can obtain $R_h\mathbf{u}$ and R_hp from (3.3.4b).

In order to guarantee that the coupled non-stationary Ritz projection R_h possesses optimal-order approximation properties, we need to define $R_h\eta(0)$ in a rather technical way. Therefore, we present error estimates for this projection in Theorem 3.3.1 and postpone the definition of $R_h\eta(0)$ and the proof of Theorem 3.3.1 to Section 3.4.

Theorem 3.3.1 (Error estimates for the coupled non-stationary Ritz projection). *For sufficiently smooth functions $(\mathbf{u}, p, \eta)$ satisfying $\mathbf{u}|_\Sigma = \partial_t\eta$, there exists $\mathbf{w}_h \in \mathbf{S}_h^r$ such that when $R_h\eta(0) = \mathbf{w}_h$, the following estimates hold uniformly for $t \in [0, T]$:*

$$
\max_{t\in[0,T]} \big(\|\eta - R_h\eta\|_\Sigma + \|\mathbf{u} - R_h\mathbf{u}\| + \|\mathbf{u} - R_h\mathbf{u}\|_\Sigma + h\|p - R_hp\| \big) \leqslant Ch^{r+1}, \tag{3.3.5}
$$

$$
\max_{t\in[0,T]} \big(\|\partial_t(\mathbf{u} - R_h\mathbf{u})\|_{H^1} + \|\partial_t(\mathbf{u} - R_h\mathbf{u})\|_{H^1(\Sigma)} + \|\partial_t(p - R_hp)\| \big) \leqslant Ch^r, \tag{3.3.6}
$$

$$
\|\partial_t(\mathbf{u} - R_h\mathbf{u})\|_{L^2L^2(\Sigma)} + \|\partial_t(\mathbf{u} - R_h\mathbf{u})\|_{L^2L^2} \leqslant Ch^{r+1}. \tag{3.3.7}
$$

3.3.3 Proof of Theorem 5.2.2

For the solution $(\mathbf{u}, p, \eta)$ of the problem (3.1.1)–(3.1.3), we define the notations:

$$\mathbf{u}^n = \mathbf{u}(t_n, \cdot), \quad \eta^n = \eta(t_n, \cdot), \quad p^n = p(t_n, \cdot). \tag{3.3.8}$$

For the analysis of the kinematically coupled scheme, we introduce $\mathbf{s}^n \in H^1(\Sigma)$ and $R_h\mathbf{s}^n \in \mathbf{S}_h^r$ by

$$\mathbf{s}^n = \partial_t\eta(t_n, \cdot) = \mathbf{u}(t_n, \cdot) \quad \text{and} \quad R_h\mathbf{s}^n := (R_h\mathbf{u})(t_n) = \partial_t R_h\eta(t_n) \quad \text{on } \Sigma,$$

which satisfy the estimate:

$$\|\mathbf{s}^n - R_h\mathbf{s}^n\|_\Sigma \leqslant Ch^{r+1} \tag{3.3.9}$$

according to the estimates in Theorem 3.3.1.

By Taylor's expansion, we have $\eta^n = \eta^{n-1} + \tau\mathbf{s}^n + \mathcal{T}_0^n$, with a truncation error $\mathcal{T}_0^n$ which has the following bound:

$$\|\mathcal{T}_0^n\|_{H^1(\Sigma)} \leq C\tau^2 \quad \forall\, n \geq 1\,. \tag{3.3.10}$$

By (3.1.1)–(3.1.3), we can see that the sequence $(\mathbf{u}^n, p^n, \eta^n, \mathbf{s}^n)$ satisfies the following weak formulations

$$\rho_s\epsilon_s\left(\frac{\mathbf{s}^n - \mathbf{u}^{n-1}}{\tau}, \mathbf{w}_h\right)_\Sigma + a_s(\eta^n, \mathbf{w}_h) + (\sigma^{n-1}\cdot\mathbf{n}, \mathbf{w}_h)_\Sigma = \mathcal{E}_s^n(\mathbf{w}_h), \qquad \forall\mathbf{w}_h \in \mathbf{S}_h^r \tag{3.3.11}$$

and

$$\begin{aligned}
&\rho_f(D_\tau\mathbf{u}^n, \mathbf{v}_h) + a_f(\mathbf{u}^n, \mathbf{v}_h) - b(p^n, \mathbf{v}_h) + b(q_h, \mathbf{u}^n) + \rho_s\epsilon_s\left(\frac{\mathbf{u}^n - \mathbf{s}^n}{\tau}, \mathbf{v}_h\right)_\Sigma \\
&= (\sigma^{n-1}\cdot\mathbf{n}, \mathbf{v}_h)_\Sigma - (\mathbf{u}^n - \mathbf{s}^n, \sigma(\mathbf{v}_h, q_h)\cdot\mathbf{n})_\Sigma - \frac{\tau(1+\beta)}{\rho_s\epsilon_s}((\sigma^n - \sigma^{n-1})\cdot\mathbf{n}, \sigma(\mathbf{v}_h, q_h)\cdot\mathbf{n})_\Sigma \\
&\quad + \mathcal{E}_f^n(\mathbf{v}_h, q_h), \qquad \forall(\mathbf{v}_h, q_h) \in \mathbf{X}_h^r \times Q_h^{r-1}
\end{aligned} \tag{3.3.12}$$

where $\sigma^n = \sigma(\mathbf{u}^n, p^n)$ and the truncation error functions satisfy the following estimates:

$$\begin{aligned}
&|\mathcal{E}_s^n(\mathbf{w}_h)| \leqslant C\tau\|\mathbf{w}_h\|_\Sigma, \\
&|\mathcal{E}_f^n(\mathbf{v}_h, \mathbf{q}_h)| \leqslant C\tau(\|\mathbf{v}_h\|_\Sigma + \|\mathbf{v}_h\|) + C\tau^2\|\sigma(\mathbf{v}_h, \mathbf{q}_h)\cdot\mathbf{n}\|_\Sigma\,.
\end{aligned} \tag{3.3.13}$$

For given $(\mathbf{u}^n, p^n, \eta^n, \mathbf{s}^n)$, we denote by $(R_h\mathbf{u}^n, R_hp^n, R_h\eta^n, R_h\mathbf{s}^n)$ the corresponding coupled non-stationary Ritz projection and define $R_h\mathcal{T}_0^n$ satisfying

$$R_h\eta^n = R_h\eta^{n-1} + \tau R_h\mathbf{s}^n + R_h\mathcal{T}_0^n \qquad \forall n \geq 1.$$

Then we introduce the following error decomposition:

$$\begin{aligned}
&e_u^n := \mathbf{u}^n - \mathbf{u}_h^n = \mathbf{u}^n - R_h\mathbf{u}^n + R_h\mathbf{u}^n - \mathbf{u}_h^n := \theta_u^n + \delta_u^n, && \text{in } \Omega. \\
&e_p^n := p^n - p_h^n = p^n - R_hp^n + R_hp^n - p_h^n := \theta_p^n + \delta_p^n, && \text{in } \Omega. \\
&e_\sigma^n := \sigma(\mathbf{u}^n, p^n) - \sigma(\mathbf{u}_h^n, p_h^n) = \sigma(\theta_u^n, \theta_p^n) + \sigma(\delta_u^n, \delta_p^n) := \theta_\sigma^n + \delta_\sigma^n, && \text{in } \Omega. \\
&e_s^n := \mathbf{s}^n - \mathbf{s}_h^n = \mathbf{s}^n - R_h\mathbf{s}^n + R_h\mathbf{s}^n - \mathbf{s}_h^n := \theta_s^n + \delta_s^n, && \text{on } \Sigma\,.
\end{aligned}$$

$$e_\eta^n := \boldsymbol{\eta}^n - \boldsymbol{\eta}_h^n = \boldsymbol{\eta}^n - R_h\boldsymbol{\eta}^n + R_h\boldsymbol{\eta}^n - \boldsymbol{\eta}_h^n := \theta_\eta^n + \delta_\eta^n, \qquad \text{on } \Sigma\,.$$

Since $\mathbf{u}^n|_\Sigma = \mathbf{s}^n$, it follows that $\theta_u^n|_\Sigma = \theta_s^n$. Moreover, the following relations hold:

$$(\mathbf{u}^n - \mathbf{u}^{n-1}) - (\mathbf{s}_h^n - \mathbf{u}_h^{n-1}) = \theta_u^n + \delta_s^n - \theta_u^{n-1} - \delta_u^{n-1},$$
$$(\mathbf{u}^n - \mathbf{u}^n) - (\mathbf{u}_h^n - \mathbf{s}_h^n) = \theta_u^n + \delta_u^n - \theta_u^n - \delta_s^n = \delta_u^n - \delta_s^n \quad \text{on } \Sigma.$$

By using (3.2.19)–(3.2.20) and (3.3.11)–(3.3.12), we can write down the following error equations:

$$\rho_s\epsilon_s\left(\frac{\delta_s^n - \delta_u^{n-1}}{\tau},\, \mathbf{w}_h\right)_\Sigma + a_s(\delta_\eta^n,\, \mathbf{w}_h) + (\delta_\sigma^{n-1}\cdot\mathbf{n},\, \mathbf{w}_h)_\Sigma = \mathcal{E}_s^n(\mathbf{w}_h) - F_s^n(\mathbf{w}_h), \quad \forall \mathbf{w}_h \in \mathbf{S}_h^r \tag{3.3.14}$$

$$\delta_\eta^n = \delta_\eta^{n-1} + \tau\delta_s^n + R_h\mathcal{T}_0^n, \qquad \text{on } \Sigma \tag{3.3.15}$$

$$\begin{aligned}
&\rho_f\left(\frac{\delta_u^n - \delta_s^n}{\tau},\, \mathbf{v}_h\right) + a_f(\delta_u^n,\, \mathbf{v}_h) - b(\delta_p^n,\, \mathbf{v}_h^n) + b(q_h,\, \delta_u^n) + \rho_s\epsilon_s\left(\frac{\delta_u^n - \delta_s^n}{\tau},\, \mathbf{v}_h\right)_\Sigma \\
&= (\delta_\sigma^{n-1}\cdot\mathbf{n},\, \mathbf{v}_h)_\Sigma - (\delta_u^n - \delta_s^n,\, \boldsymbol{\sigma}(\mathbf{v}_h,\, q_h))_\Sigma - \frac{\tau(1+\beta)}{\rho_s\epsilon_s}((\delta_\sigma^n - \delta_\sigma^{n-1})\cdot\mathbf{n},\, \boldsymbol{\sigma}(\mathbf{v}_h, q_h)\cdot\mathbf{n})_\Sigma \\
&\quad + \mathcal{E}_f^n(\mathbf{v}_h, q_h) - F_f^n(\mathbf{v}_h, q_h), \qquad \forall(\mathbf{v}_h, q_h) \in \mathbf{X}_h^r \times Q_h^{r-1}
\end{aligned} \tag{3.3.16}$$

where

$$F_s^n(\mathbf{w}_h) = \rho_s\epsilon_s(D_\tau\theta_u^n,\, \mathbf{w}_h)_\Sigma + a_s(\theta_\eta^n,\, \mathbf{w}_h) + (\theta_\sigma^{n-1}\cdot\mathbf{n},\, \mathbf{w}_h)_\Sigma \tag{3.3.17}$$

$$\begin{aligned}
F_f^n(\mathbf{v}_h, q_h) &= \rho_f(D_\tau\theta_u^n, \mathbf{v}_h) + a_f(\theta_u^n,\, \mathbf{v}_h) - b(\theta_p^n,\, \mathbf{v}_h) \\
&\quad - (\theta_\sigma^{n-1}\cdot\mathbf{n},\, \mathbf{v}_h)_\Sigma + \frac{\tau(1+\beta)}{\rho_s\epsilon_s}((\theta_\sigma^n - \theta_\sigma^{n-1})\cdot\mathbf{n}, \boldsymbol{\sigma}(\mathbf{v}_h,\, q_h)\cdot\mathbf{n})_\Sigma
\end{aligned} \tag{3.3.18}$$

Moreover, we have the following result:

$$\theta_\eta^n = \theta_\eta^{n-1} + \tau\theta_s^n + (\mathcal{T}_0^n - R_h\mathcal{T}_0^n),$$

where the last term can be estimated by using (3.3.6), i.e.,

$$\|\mathcal{T}_0^n - R_h\mathcal{T}_0^n\|_{H^1(\Sigma)} \le C\tau^2\|\partial_t(R_h\mathbf{u} - \mathbf{u})\|_{L^\infty H^1(\Sigma)} \le C\tau^2h^r. \tag{3.3.19}$$

Therefore, by the triangle inequality with estimates (3.3.10) and (3.3.19), we have

$$\|R_h\mathcal{T}_0^n\|_{H^1(\Sigma)} \leqslant \|\mathcal{T}_0^n\|_{H^1(\Sigma)} + \|\mathcal{T}_0^n - R_h\mathcal{T}_0^n\|_{H^1(\Sigma)} \le C\tau^2 \quad \forall n \ge 1 \tag{3.3.20}$$

We take $(\mathbf{v}_h, q_h) = (\delta_u^n, \delta_p^n) \in \mathbf{X}_h^r \times Q_h^{r-1}$ in (3.3.16) and $\mathbf{w}_h = \delta_s^n \in \mathbf{S}_h^r$ in (3.3.14), respectively, and then sum up the two results. Using the stability analysis in (3.3.2) and the relation

$$\delta_s^n = D_\tau\delta_\eta^n - \tau^{-1}R_h\mathcal{T}_0^n,$$

we obtain

$$\begin{aligned}
&D_\tau E_0(\delta_u^n, \delta_p^n, \delta_\eta^n) + E_1(\delta_u^n, \delta_s^n, \delta_\eta^n) \\
&\leqslant \mathcal{E}_s^n(\delta_s^n) - F_s^n(\delta_s^n) + \mathcal{E}_f^n(\delta_u^n, \delta_p^n) - F_f^n(\delta_u^n, \delta_p^n) + \tau^{-1}a_s(\delta_\eta^n, R_h\mathcal{T}_0^n)\,.
\end{aligned} \tag{3.3.21}$$

To establish the error estimate, we need to estimate each term on the right-hand side of (3.3.21). From (3.3.13) and (3.3.20) we can see that

$$\begin{aligned}
&|\mathcal{E}_s^n(\delta_s^n)| \leqslant C\tau\|\delta_s^n\|_\Sigma \\
&|\mathcal{E}_f^n(\delta_u^n,\, \delta_p^n)| \leqslant C\tau(\|\delta_u^n\|_\Sigma + \|\delta_u^n\|) + \tau^2\|\delta_\sigma^n\cdot\mathbf{n}\|_\Sigma \\
&|\tau^{-1}a_s(\delta_\eta^n,\, R_h\mathcal{T}_0^n)| \leq C\tau\|\delta_\eta^n\|_s
\end{aligned} \tag{3.3.22}$$

It remains to estimate $F_s^n(\delta_s) + F_f^n(\delta_u, \delta_p)$ from the right hand side of (3.3.21).

1. The second term in (3.3.17) plus the second and third terms in (3.3.18) can be estimated as follows. Let $\xi_h^n := \delta_u^n - E_h(\delta_u^n - \delta_s^n)$, where $E_h(\delta_u^n - \delta_s^n)$ is an extension of $\delta_u^n - \delta_s^n$ to Ω satisfying estimate (3.2.17) and $\xi_h^n|_\Sigma = \delta_s^n$. By choosing $v_h = \xi_h^n$ and $q_h = 0$ in (3.3.3) (definition of the coupled Ritz projection), we obtain the following relation:

$$\begin{aligned}
&a_f(\theta_u^n, \delta_u^n) - b(\theta_p^n, \delta_u^n) + a_s(\theta_\eta^n, \delta_s^n) \\
&= a_f(\theta_u^n, E_h(\delta_u^n - \delta_s^n)) - b(\theta_p^n, E_h(\delta_u^n - \delta_s^n)) - (\theta_u^n, \xi_h^n) - (\theta_\eta^n, \delta_s^n)_\Sigma \\
&\leq Ch^r\|E_h(\delta_u^n - \delta_s^n)\|_f + Ch^{r+1}(\|\xi_h^n\| + \|\delta_s^n\|_\Sigma) \\
&\leq Ch^{r-1/2}\|\delta_u^n - \delta_s^n\|_\Sigma + Ch^{r+1}(\|\delta_u^n\| + \|\delta_s^n\|_\Sigma),
\end{aligned} \tag{3.3.23}$$

   where we have used estimate (3.3.5)–(3.3.6).

2. The third term in (3.3.17) plus the fourth term in (3.3.18) can be estimated as follows:

$$\begin{aligned}
&(\theta_\sigma^{n-1}\cdot\mathbf{n}, \delta_s^n)_\Sigma - (\theta_\sigma^{n-1}\cdot\mathbf{n}, \delta_u^n)_\Sigma \\
\leq&\|\theta_\sigma^{n-1}\cdot\mathbf{n}\|_\Sigma\|\delta_s^n - \delta_u^n\|_\Sigma \\
\leq&C(h^{r-1/2} + h^{-1/2}(\|\theta_u^{n-1}\|_{H^1} + \|\theta_p^{n-1}\|))\|\delta_s^n - \delta_u^n\|_\Sigma \\
\leq&Ch^{r-1/2}\|\delta_s^n - \delta_u^n\|_\Sigma,
\end{aligned} \tag{3.3.24}$$

   where we used (3.2.18) in the second inequality and (3.3.5) in the last inequality.

3. For the first term in (3.3.17) and (3.3.18), respectively, we have

$$\rho_s\epsilon_s(D_\tau\theta_u^n,\, \delta_s^n)_\Sigma \leq \frac{C}{\tau}\|\delta_s^n\|_\Sigma\int_{t_{n-1}}^{t_n}\|\partial_t\theta_u(t)\|_\Sigma dt, \tag{3.3.25}$$

$$\rho_f(D_\tau\theta_u^n, \delta_u^n) \leq \frac{C}{\tau}\|\delta_u^n\|\int_{t_{n-1}}^{t_n}\|\partial_t\theta_u(t)\|dt. \tag{3.3.26}$$

4. The last term in (3.3.18) can be estimated by using (3.3.6) and (3.2.18), i.e.,

$$\begin{aligned}
&\frac{\tau}{\rho_s\epsilon}((\theta_\sigma^n - \theta_\sigma^{n-1})\cdot\mathbf{n}, \boldsymbol{\sigma}(\delta_u^n,\, \delta_p^n)\cdot\mathbf{n})_\Sigma \\
\leqslant&C\tau\left(\int_{t_{n-1}}^{t_n}\|\boldsymbol{\sigma}(\partial_t\theta_u,\, \partial_t\theta_p)(t)\cdot\mathbf{n}\|_\Sigma dt\right)\|\boldsymbol{\sigma}(\delta_u^n,\, \delta_p^n)\cdot\mathbf{n}\|_\Sigma \\
\leq&C\tau^2h^{r-1/2}\|\boldsymbol{\sigma}(\delta_u^n,\, \delta_p^n)\cdot\mathbf{n}\|_\Sigma.
\end{aligned} \tag{3.3.27}$$

Now we can substitute estimates (3.3.22)–(3.3.27) into the energy inequality in (3.3.21). This yields the following result:

$$\begin{aligned} &D_\tau E_0(\delta_u^n, \delta_p^n, \delta_\eta^n) + E_1(\delta_u^n, \delta_s^n, \delta_\eta^n) \\ &\leqslant C\tau(\|\delta_s^n\|_\Sigma + \|\delta_u^n\|_\Sigma + \|\delta_u^n\| + \|\delta_\eta^n\|_s) + Ch^{r-1/2}\|\delta_u^n - \delta_s^n\|_\Sigma + Ch^{r+1}(\|\delta_u^n\| + \|\delta_s^n\|_\Sigma) \\ &\quad + \frac{C}{\tau}\|\delta_s^n\|_\Sigma \int_{t_{n-1}}^{t_n} \|\partial_t\theta_u(t)\|_\Sigma dt + \frac{C}{\tau}\|\delta_u^n\| \int_{t_{n-1}}^{t_n} \|\partial_t\theta_u(t)\| dt + C\tau^2\|\delta_\sigma^n \cdot \mathbf{n}\|_\Sigma . \qquad (3.3.28) \end{aligned}$$

Since $\|\delta_s^n\|_\Sigma \leqslant \|\delta_s^n - \delta_u^n\|_\Sigma + \|\delta_u^n\|_\Sigma$, by using Young's inequality, we can re-arrange the right hand side of (3.3.28) to obtain

$$\begin{aligned} &D_\tau E_0(\delta_u^n, \delta_p^n, \delta_\eta^n) + E_1(\delta_u^n, \delta_s^n, \delta_\eta^n) \\ &\leqslant C\varepsilon^{-1}(\tau^2 + Ch^{2(r+1)} + \tau h^{2r-1}) + C\varepsilon(\|\delta_u^n\|_\Sigma^2 + \|\delta_u^n\|^2 + \|\delta_\eta^n\|_s^2) + \frac{C\varepsilon}{\tau}\|\delta_u^n - \delta_s^n\|_\Sigma^2 \\ &\quad + \frac{C\varepsilon^{-1}}{\tau}\left(\int_{t_{n-1}}^{t_n} \|\partial_t\theta_u(t)\|_\Sigma^2 dt + \int_{t_{n-1}}^{t_n} \|\partial_t\theta_u(t)\|^2 dt\right) + C\tau^2\|\delta_\sigma^n \cdot \mathbf{n}\|_\Sigma^2, \qquad (3.3.29) \end{aligned}$$

where $0 < \varepsilon < 1$ is an arbitrary constant.

We can choose a sufficiently small ε so that the term $\frac{C\varepsilon}{\tau}\|\delta_u^n - \delta_s^n\|_\Sigma^2$ can be absorbed by $E_1(\delta_u^n, \delta_s^n, \delta_\eta^n)$ on the left-hand side. Then, using the discrete Gronwall's inequality and the estimates of θ_u in (3.3.7), as well as the definition of E_0 and E_1 in (3.2.25)–(3.2.26), we obtain

$$E_0(\delta_u^n, \delta_p^n, \delta_\eta^n) + \sum_{m=1}^{n} \tau E_1(\delta_u^m, \delta_s^m, \delta_\eta^m) \leq CE_0(\delta_u^0, \delta_p^0, \delta_\eta^0) + C(\tau^2 + Ch^{2(r+1)} + \tau h^{2r-1}). \qquad (3.3.30)$$

Since the initial values satisfy the estimates in (3.2.21), the term $E_0(\delta_u^0, \delta_p^0, \delta_\eta^0)$ can be estimated to the optimal order. Thus inequality (3.3.30) reduces to

$$\|\delta_u^n\| + \|\delta_u^n\|_\Sigma + \|\delta_\eta^n\|_s + \|\delta_u^n - \delta_s^n\|_\Sigma \leqslant C(h^{r-1/2}\tau^{1/2} + \tau + h^{r+1}). \qquad (3.3.31)$$

It follows from the relation $\delta_\eta^n = \delta_\eta^{n-1} + \tau\delta_s^n + R_h\mathcal{T}_0^n$, $n \geqslant 1$, that

$$\|\delta_\eta^n\|_\Sigma \leq \|\delta_\eta^0\|_\Sigma + \sum_{m=1}^{n} \tau\|\delta_s^m\|_\Sigma + \sum_{m=1}^{n} \|R_h\mathcal{T}_0^m\|_\Sigma \leq C(h^{r-1/2}\tau^{1/2} + \tau + h^{r+1}), \qquad (3.3.32)$$

where we have used (3.3.31) and (3.3.20). Then, combining the two estimates above with the following estimate for the projection error:

$$\|\theta_u^n\| + \|\theta_u^n\|_\Sigma + \|\theta_\eta^n\|_\Sigma \leq Ch^{r+1} \quad \forall n \geq 0,$$

we obtain the following error bound:

$$\|e_u^n\| + \|e_u^n\|_\Sigma + \|e_\eta^n\|_\Sigma \leqslant C(h^{r-1/2}\tau^{1/2} + \tau + h^{r+1}) \leqslant C(\tau + h^{r+1}),$$

where the last inequality uses $h^{r-1/2}\tau^{1/2} \leqslant \tau + h^{2r-1}$ and $r \geqslant 2$. This completes the proof of Theorem 5.2.2. ∎

3.4 The proof of Theorem 3.3.1

We present the proof of the Theorem 3.3.1 step-by-step in the next three subsections.

3.4.1 The definition of $R_h\boldsymbol{\eta}(0)$ in the coupled Ritz projection

In this subsection, we focus on designing the initial value $R_h\boldsymbol{\eta}(0)$ for our coupled non-stationary Ritz projection.

We first present two auxiliary Ritz projections R_h^S and R_h^D associated to the structure model and the fluid model in Definitions 3.4.1-3.4.2, respectively. Next, in terms of these two auxiliary Ritz projections, we define the initial value $R_h\boldsymbol{\eta}(0)$ in Definition 3.4.3 which is only for our theoretical purpose. Finally, an alternative definition of $R_h\boldsymbol{\eta}(0)$ for practical computation is given in Definition 3.4.4.

Definition 3.4.1 (Structure–Ritz projection R_h^S)**.** We define an auxiliary Ritz projection $R_h^S : \mathbf{S} \to \mathbf{S}_h^r$ for the elastic structure problem by

$$a_s(R_h^S\mathbf{s} - \mathbf{s}, \mathbf{w}_h) + (R_h^S\mathbf{s} - \mathbf{s}, \mathbf{w}_h)_\Sigma = 0 \quad \forall \mathbf{w}_h \in \mathbf{S}_h^r. \tag{3.4.33}$$

This is the standard Ritz projection on Σ, which satisfies the estimate $\|R_h^S\mathbf{s}-\mathbf{s}\|_\Sigma \le Ch^{r+1}$ when $\mathbf{s}$ is sufficiently smooth. Moreover when $r \ge 2$, there holds the negative norm estimate:

$$\|R_h^S\mathbf{s} - \mathbf{s}\|_{H^{-1}(\Sigma)} \le Ch^{r+2}. \tag{3.4.34}$$

Let $\mathring{\mathbf{X}}_h^r := \{\mathbf{v}_h \in \mathbf{X}_h^r : \mathbf{v}_h|_\Sigma = 0\}$ and $Q_{h,0}^{r-1} := \{q_h \in Q_h^{r-1} : q_h \in L_0^2(\Omega)\}$. We denote $\widetilde{\mathbf{S}}_h^r := \{\mathbf{v}_h \in \mathbf{S}_h^r : (\mathbf{v}_h, \mathbf{n})_\Sigma = 0\}$ and by $\widetilde{P}$ the $L^2(\Sigma)$-orthogonal projection from $\mathbf{S}_h^r$ to $\widetilde{\mathbf{S}}_h^r$.

Definition 3.4.2 (Dirichlet Stokes–Ritz projection R_h^D)**.** Let $\hat{\mathbf{X}} := \{\mathbf{u} \in \mathbf{X} : \mathbf{u}|_\Sigma \in \mathbf{S}\}$. We define an auxiliary Dirichlet Stokes–Ritz projection $R_h^D : \hat{\mathbf{X}} \times Q \to \mathbf{X}_h^r \times Q_h^{r-1}$ by

$$a_f(\mathbf{u} - R_h^D\mathbf{u}, \mathbf{v}_h) - b(p - R_h^D p, \mathbf{v}_h) + (\mathbf{u} - R_h^D\mathbf{u}, \mathbf{v}_h) = 0 \quad \forall \mathbf{v}_h \in \mathring{\mathbf{X}}_h^r, \tag{3.4.35a}$$

$$b(q_h, \mathbf{u} - R_h^D\mathbf{u}) = 0 \quad \forall q_h \in Q_{h,0}^{r-1}; \quad \text{with } R_h^D\mathbf{u} = \widetilde{P}R_h^S(\mathbf{u}|_\Sigma) \quad \text{on } \Sigma, \tag{3.4.35b}$$

In addition, we choose $R_h^D p$ to satisfy $R_h^D p - p \in L_0^2(\Omega)$. This uniquely determines a solution $(R_h^D u, R_h^D p) \in \mathbf{X}_h^r \times Q_h^{r-1}$, as explained in the following Remark.

Remark 3.4.1. In order to see the existence and uniqueness of solution $(R_h^D u, R_h^D p)$ defined by (3.4.35), we let $\hat{\mathbf{u}}_h \in \mathbf{X}_h^r$ be an extension of $\widetilde{P}R_h^S\mathbf{u}$ to the bulk domain Ω and let $\hat{p}_h$ be the $L^2(\Omega)$-orthogonal projection of p onto Q_h^{r-1}. Then $\hat{\mathbf{u}}_h - R_h^D\mathbf{u} \in \mathring{\mathbf{X}}_h^r$ and $\hat{p}_h - R_h^D p \in Q_{h,0}^{r-1}$. Replacing $(\mathbf{u}, p)$ and $(R_h^D\mathbf{u}, R_h^D p)$ by $(\mathbf{u} - \hat{\mathbf{u}}_h, p - \hat{p}_h)$ and $(R_h^D\mathbf{u} - \hat{\mathbf{u}}_h, R_h^D p - \hat{p}_h)$ in (3.4.35a)-(3.4.35b) respectively, we obtain a standard Stokes FE system with a homogeneous Dirichlet boundary condition for $(R_h^D\mathbf{u} - \hat{\mathbf{u}}_h, R_h^D p - \hat{p}_h)$. The well-posedness directly follows the inf-sup condition (3.2.14).

Remark 3.4.2. The projection $\widetilde{P}$ in (3.4.35b) is introduced to guarantees that the $b(q_h, \mathbf{u} - R_h^D\mathbf{u}) = 0$ holds not only for $q_h \in Q_{h,0}^{r-1}$ but also for $q_h \in Q_h^{r-1}$. That is,

$$b(q_h, \mathbf{u} - R_h^D\mathbf{u}) = 0 \quad \forall q_h \in Q_h^{r-1}. \tag{3.4.36}$$

Since $Q_h^{r-1} = \{1\} \oplus Q_{h,0}^{r-1}$, this follows from the first relation in (3.4.35b) and the following relation:

$$b(1, \mathbf{u} - R_h^D\mathbf{u}) = (R_h^D\mathbf{u}, \mathbf{n})_\Sigma = (\widetilde{P}R_h^S\mathbf{u}, \mathbf{n})_\Sigma = 0\,,$$

where $b(1, \mathbf{u}) = 0$ for the exact solution $\mathbf{u}$ which satisfies $\nabla \cdot \mathbf{u} = 0$. Especially, when $\mathbf{u}$ is replaced with $\partial_t \mathbf{u}(0)$, we have

$$b(q_h, (\partial_t \mathbf{u} - R_h^D \partial_t \mathbf{u})(0)) = 0 \quad \forall q_h \in Q_h^{r-1}. \tag{3.4.37}$$

The relation (3.4.37) is needed in error estimates between $(\partial_t R_h \mathbf{u}(0), \partial_t R_h p(0))$ and $(\partial_t \mathbf{u}(0), \partial_t p(0))$ in the Lemma 3.4.4 below. Furthermore, in the Definition 3.4.3, we defined $(R_h \mathbf{u}(0), R_h p(0))$ via a Dirichlet-type Stokes-Ritz projection with the boundary condition $R_h \mathbf{u}(0)|_\Sigma = \widetilde{P} R_{sh} \mathbf{u}(0)$.

To facilitate further use of $\widetilde{P}$ in the following analysis, here we derive an explicit formula for $\widetilde{P}$. We denote by $\mathbf{n}_h \in \mathbf{S}_h^r$ the $L^2(\Sigma)$-orthogonal projection of unit normal vector field $\mathbf{n}$ of Σ to $\mathbf{S}_h^r$, i.e.,

$$(\mathbf{n}, \mathbf{w}_h)_\Sigma = (\mathbf{n}_h, \mathbf{w}_h)_\Sigma \quad \forall \mathbf{w}_h \in \mathbf{S}_h^r. \tag{3.4.38}$$

Then for any $\mathbf{w}_h \in \mathbf{S}_h^r$, we have

$$\widetilde{P} \mathbf{w}_h = \mathbf{w}_h - \lambda(\mathbf{w}_h) \mathbf{n}_h \in \tilde{\mathbf{S}}_h^r \quad \text{with } \lambda(\mathbf{w}_h) := \frac{(\mathbf{w}_h, \mathbf{n})_\Sigma}{\|\mathbf{n}_h\|_\Sigma^2}. \tag{3.4.39}$$

From $\|\mathbf{n} - \mathbf{n}_h\|_\Sigma \le \|\mathbf{n} - I_h \mathbf{n}\|_\Sigma \le C h^{r+1}$ (since $\mathbf{n}$ is smooth on Σ), especially we have $\|\mathbf{n}_h\|_\Sigma \sim C$ and

$$|\lambda(R_h^S \mathbf{u})| = \frac{|(R_h^S \mathbf{u} - \mathbf{u}, \mathbf{n})_\Sigma|}{\|\mathbf{n}_h\|_\Sigma^2} \le C h^{r+1} \text{ and } \|\widetilde{P} R_h^S \mathbf{u} - R_h^S \mathbf{u}\| \le C h^{r+1}. \tag{3.4.40}$$

Therefore we obtain the estimate $\|R_h^D \mathbf{u} - \mathbf{u}\|_\Sigma \le C h^{r+1}$.

The following lemma on the error estimates of the Dirichlet Stokes–Ritz projection is standard. We refer to [73, Proposition 8, Proposition 9] for the proof of (3.4.41). The negative norm estimate of pressure in (3.4.42) requires a further duality argument, which is presented in the proof of Lemma 3.8.3 of Section 3.8. We omit the details here.

Lemma 3.4.1. *Under the regularity assumptions in Section 3.2.2, the Dirichlet Stokes–Ritz projection R_h^D defined in* (3.4.35) *satisfies the following estimates:*

$$\|\mathbf{u} - R_h^D \mathbf{u}\|_\Sigma + \|\mathbf{u} - R_h^D \mathbf{u}\| + h \left(\|\mathbf{u} - R_h^D \mathbf{u}\|_{H^1} + \|p - R_h^D p\| \right) \leqslant C h^{r+1}, \tag{3.4.41}$$

$$\|R_h^D p - p\|_{H^{-1}} \le C h^{r+1}. \tag{3.4.42}$$

We define an initial value $R_h \boldsymbol{\eta}(0)$ as follows in terms of the Dirichlet Ritz projection R_h^D.

Definition 3.4.3 (Initial value $R_h \boldsymbol{\eta}(0)$)**.** Firstly, assuming that the function $R_h^D \partial_t \mathbf{u}(0)$ and $R_h^D \partial_t p(0)$ are known with operator R_h^D defined by (3.4.35), we define $R_{sh} \mathbf{u}(0) \in \mathbf{S}_h^r$ to be the solution of the following weak formulation:

$$\begin{aligned} &a_s((\mathbf{u} - R_{sh} \mathbf{u})(0), \mathbf{w}_h) + ((\mathbf{u} - R_{sh} \mathbf{u})(0), \mathbf{w}_h)_\Sigma + a_f((\partial_t \mathbf{u} - R_h^D \partial_t \mathbf{u})(0), E_h \mathbf{w}_h) \\ &\quad - b((\partial_t p - R_h^D \partial_t p)(0), E_h \mathbf{w}_h) + ((\partial_t \mathbf{u} - R_h^D \partial_t \mathbf{u})(0), E_h \mathbf{w}_h) = 0 \quad \forall \mathbf{w}_h \in \mathbf{S}_h^r, \end{aligned} \tag{3.4.43}$$

where $E_h \mathbf{w}_h$ denotes an extension of $\mathbf{w}_h$ to the bulk domain Ω. From the definition of R_h^D in (3.4.35) we can conclude that this definition is independent of the specific extension. Therefore, (3.4.43) still holds when replacing both $\mathbf{w}_h$ and $E \mathbf{w}_h$ with $\mathbf{v}_h \in \mathbf{X}_h^r$.

Secondly, we denote by $(R_h\mathbf{u}(0), R_hp(0)) \in \mathbf{X}_h^r \times Q_h^{r-1}$ a Dirichlet-type Stokes–Ritz projection satisfying

$$a_f(\mathbf{u}(0)-R_h\mathbf{u}(0), \mathbf{v}_h) - b(p(0) - R_hp(0), \mathbf{v}_h) + (\mathbf{u}(0) - R_h\mathbf{u}(0), \mathbf{v}_h) = 0 \quad \forall \mathbf{v}_h \in \mathring{\mathbf{X}}_h^r, \tag{3.4.44a}$$

$$b(q_h, \mathbf{u}(0) - R_h\mathbf{u}(0)) = 0 \quad \forall q_h \in Q_{h,0}^{r-1}; \quad R_h\mathbf{u}(0) = \widetilde{P}R_{sh}\mathbf{u}(0) \quad \text{on } \Sigma, \tag{3.4.44b}$$

where we require $p(0) - R_hp(0) \in L_0^2(\Omega)$.

Finally, with the $R_h\mathbf{u}(0)$ and $R_hp(0)$ defined above, we define $R_h\boldsymbol{\eta}(0) \in \mathbf{S}_h^r$ to be the solution of the following weak formulation on Σ:

$$\begin{aligned} &a_f(\mathbf{u}(0) - R_h\mathbf{u}(0), E_h\mathbf{w}_h) - b(p(0) - R_hp(0), E_h\mathbf{w}_h) + (\mathbf{u}(0) - R_h\mathbf{u}(0), E_h\mathbf{w}_h) \\ &\qquad + a_s(\boldsymbol{\eta}(0) - R_h\boldsymbol{\eta}(0), \mathbf{w}_h) + (\boldsymbol{\eta}(0) - R_h\boldsymbol{\eta}(0), \mathbf{w}_h)_\Sigma = 0 \quad \forall \mathbf{w}_h \in \mathbf{S}_h^r. \end{aligned} \tag{3.4.45}$$

Again (3.4.45) also holds when replacing $\mathbf{w}_h$ and $E_h\mathbf{w}_h$ with $\mathbf{v}_h \in \mathbf{X}_h^r$.

For the computation with the numerical scheme (3.2.19)–(3.2.20), we can define the initial value $\boldsymbol{\eta}_h^0 = R_{sh}\boldsymbol{\eta}(0) \in \mathbf{S}_h^r$ in an alternative way below.

Definition 3.4.4 (Ritz projection $R_{sh}\boldsymbol{\eta}(0)$)**.** We define $\boldsymbol{\eta}_h^0 = R_{sh}\boldsymbol{\eta}(0) \in \mathbf{S}_h^r$ as the solution of the following weak formulation:

$$\begin{aligned} &a_s((R_{sh}\boldsymbol{\eta} - \boldsymbol{\eta})(0), \mathbf{w}_h) + ((R_{sh}\boldsymbol{\eta} - \boldsymbol{\eta})(0), \mathbf{w}_h)_\Sigma \qquad \forall \mathbf{w}_h \in \mathbf{S}_h^r \\ &= -a_f((R_h^D\mathbf{u} - \mathbf{u})(0), E_h\mathbf{w}_h) + b((R_h^Dp - p)(0), E_h\mathbf{w}_h) - ((R_h^D\mathbf{u} - \mathbf{u})(0), E_h\mathbf{w}_h)\,, \end{aligned} \tag{3.4.46}$$

which does not require knowledge of $\partial_t\mathbf{u}(0)$ or $\partial_tp(0)$. Again, $E_h\mathbf{w}_h$ denotes an extension of $\mathbf{w}_h$ to the bulk domain Ω, and this definition is independent of the specific extension. Therefore, (3.4.46) holds for all $\mathbf{v}_h \in \mathbf{X}_h^r$ with $\mathbf{w}_h$ and $E_h\mathbf{w}_h$ replaced by $\mathbf{v}_h$ in the equation. For $r \geq 2$, the following result can be proved in Section 3.8:

$$\|R_{sh}\boldsymbol{\eta}(0) - R_h\boldsymbol{\eta}(0)\|_{H^1(\Sigma)} \leq Ch^{r+1}. \tag{3.4.47}$$

In addition, by differentiating (3.3.3) with respect to time, we have the following evolution equations:

$$\begin{aligned} &a_s(\mathbf{u} - R_h\mathbf{u}, \mathbf{v}_h) + (\mathbf{u} - R_h\mathbf{u}, \mathbf{v}_h)_\Sigma + a_f(\partial_t(\mathbf{u} - R_h\mathbf{u}), \mathbf{v}_h) \\ &\qquad -b(\partial_t(p - R_hp), \mathbf{v}_h) + (\partial_t(\mathbf{u} - R_h\mathbf{u}), \mathbf{v}_h) = 0 \quad \forall \mathbf{v}_h \in \mathbf{X}_h^r, \end{aligned} \tag{3.4.48a}$$

$$b(q_h, \partial_t(\mathbf{u} - R_h\mathbf{u})) = 0 \quad \forall q_h \in Q_h^{r-1}\,, \tag{3.4.48b}$$

which are used not only to design the above $R_h\boldsymbol{\eta}(0)$, but also to estimate errors in the following subsections.

3.4.2 Error estimates for the coupled Ritz projection at $t = 0$

Firstly, we consider the estimation of $R_{sh}\mathbf{u}(0)$ which occurs as an auxiliary function in the definition of $R_h\boldsymbol{\eta}(0)$ in Lemma 3.4.2. Secondly, we present estimates for $\mathbf{u}(0) - R_h\mathbf{u}(0)$, $\boldsymbol{\eta}(0) - R_h\boldsymbol{\eta}(0)$ and $p(0) - R_hp(0)$ in Lemma 3.4.3. Finally, we present estimates for the time derivatives $\partial_t(\mathbf{u} - R_h\mathbf{u})(0)$ and $\partial_t(p - R_hp)(0)$ in Lemma 3.4.4.

Lemma 3.4.2. *Under the assumptions in Sections 3.2.2 and 3.2.3, the following error estimate holds for the* $R_{sh}\mathbf{u}(0)$ *defined in* (3.4.43):

$$\|R_{sh}\mathbf{u}(0) - \mathbf{u}(0)\|_\Sigma + h\|R_{sh}\mathbf{u}(0) - \mathbf{u}(0)\|_s \le Ch^{r+1}. \qquad (3.4.49)$$

Proof. Since we can choose an extension $E_h\boldsymbol{\xi}_h$ of $\boldsymbol{\xi}_h \in \mathbf{S}_h^r$ to satisfy that $\|E_h\boldsymbol{\xi}_h\|_{H^1(\Omega)} \le C\|\boldsymbol{\xi}_h\|_{H^1(\Sigma)}$, equation (3.4.43) implies that

$$a_s(\mathbf{u}(0) - R_{sh}\mathbf{u}(0), \boldsymbol{\xi}_h) + (\mathbf{u}(0) - R_{sh}\mathbf{u}(0), \boldsymbol{\xi}_h)_\Sigma \leqslant Ch^r\|\boldsymbol{\xi}_h\|_{H^1(\Sigma)}\,.$$

This leads to the following standard H^1-norm estimate:

$$\|\mathbf{u}(0) - R_{sh}\mathbf{u}(0)\|_s + \|\mathbf{u}(0) - R_{sh}\mathbf{u}(0)\|_\Sigma \le Ch^r.$$

In order to obtain an optimal-order L^2-norm estimate for $\mathbf{u}(0) - R_{sh}\mathbf{u}(0)$, we introduce the following dual problem:

$$-\mathcal{L}_s\psi + \psi = R_{sh}\mathbf{u}(0) - \mathbf{u}(0), \quad \psi \text{ has periodic boundary condition on } \Sigma. \qquad (3.4.50)$$

The regularity assumption in (3.2.9) implies that

$$a_s(\psi, \boldsymbol{\xi}) + (\psi, \boldsymbol{\xi})_\Sigma = (\mathbf{u}(0) - R_{sh}\mathbf{u}(0), \boldsymbol{\xi})_\Sigma \quad \forall \boldsymbol{\xi} \in \mathbf{S} \quad \text{and } \|\psi\|_{H^2(\Sigma)} \le C\|\mathbf{u}(0) - R_{sh}\mathbf{u}(0)\|_\Sigma\,.$$

We can extend ψ to be a function on Ω, still denoted by ψ, satisfying the periodic boundary condition and $\|\psi\|_{H^2(\Omega)} \le C\|\psi\|_{H^2(\Sigma)}$. Therefore, choosing $\boldsymbol{\xi} = \mathbf{u}(0) - R_{sh}\mathbf{u}(0)$ in the equation above leads to

$$\begin{aligned}
\|\mathbf{u}(0) - R_{sh}\mathbf{u}(0)\|_\Sigma^2 =\,& a_s(\mathbf{u}(0) - R_{sh}\mathbf{u}(0), \psi) + (\mathbf{u}(0) - R_{sh}\mathbf{u}(0), \psi)_\Sigma \\
=\,& a_s(\mathbf{u}(0) - R_{sh}\mathbf{u}(0), \psi - I_h\psi) + (\mathbf{u}(0) - R_{sh}\mathbf{u}(0), \psi - I_h\psi)_\Sigma \\
& - a_f(\partial_t\mathbf{u}(0) - R_h^D\partial_t\mathbf{u}(0), I_h\psi) + b(\partial_t p(0) - R_h^D\partial_t p(0), I_h\psi) \\
& - (\partial_t\mathbf{u}(0) - R_h^D\partial_t\mathbf{u}(0), I_h\psi) \quad \text{(relation (3.4.43) is used)} \\
\le\,& Ch^{r+1}\|\psi\|_{H^2(\Sigma)} + |a_f(\partial_t\mathbf{u}(0) - R_h^D\partial_t\mathbf{u}(0), \psi)| \\
& + |b(\partial_t p(0) - R_h^D\partial_t p(0), \psi)| + |(\partial_t\mathbf{u}(0) - R_h^D\partial_t\mathbf{u}(0), \psi)|\,.
\end{aligned}$$

Since

$$\begin{aligned}
&|(\mathbf{D}(\partial_t\mathbf{u}(0) - R_h^D\partial_t\mathbf{u}(0)), \mathbf{D}\psi)| \\
&= |-(\partial_t\mathbf{u}(0) - R_h^D\partial_t\mathbf{u}(0), \nabla\cdot\mathbf{D}\psi) + (\partial_t\mathbf{u}(0) - R_h^D\partial_t\mathbf{u}(0), \mathbf{D}\psi\cdot\mathbf{n})_\Sigma| \\
&\le Ch^{r+1}\|\psi\|_{H^2(\Sigma)},
\end{aligned}$$

where the last inequality uses the estimate $\|\psi\|_{H^2(\Omega)} \le C\|\psi\|_{H^2(\Sigma)}$ as well as the estimates of $\|\partial_t\mathbf{u}(0) - R_h^D\partial_t\mathbf{u}(0)\|$ and $\|\partial_t\mathbf{u}(0) - R_h^D\partial_t\mathbf{u}(0)\|_\Sigma$ in (3.4.41) (with $\mathbf{u}(0)$ replaced by $\partial_t\mathbf{u}(0)$ therein). Furthermore, using the H^{-1} estimate in (3.4.42), we have

$$|b(\partial_t p(0) - R_h^D\partial_t p(0), \psi)| \le C\|\partial_t p(0) - R_h^D\partial_t p(0)\|_{H^{-1}}\|\psi\|_{H^2} \leqslant Ch^{r+1}\|\psi\|_{H^2(\Sigma)}.$$

Then, summing up the estimates above, we obtain

$$\|\mathbf{u}(0) - R_{sh}\mathbf{u}(0)\|_\Sigma \le Ch^{r+1}\,.$$

The proof of Lemma 3.4.2 is complete. ∎

Lemma 3.4.3. *Under the assumptions in Sections 3.2.2 and 3.2.3, the following error estimates hold (for the coupled Ritz projection in Definition 3.4.3)*:

$$\|\boldsymbol{\eta}(0) - R_h\boldsymbol{\eta}(0)\|_\Sigma + h\|\boldsymbol{\eta}(0) - R_h\boldsymbol{\eta}(0)\|_s + \|\mathbf{u}(0) - R_h\mathbf{u}(0)\|_\Sigma \le Ch^{r+1}, \tag{3.4.51}$$

$$\|\mathbf{u}(0) - R_h\mathbf{u}(0)\| + h\|p(0) - R_hp(0)\| \le Ch^{r+1}. \tag{3.4.52}$$

Proof. From (3.4.39) we know that $R_h\mathbf{u}(0) = \widetilde{P}R_{sh}\mathbf{u}(0) = R_{sh}\mathbf{u}(0) - \lambda(R_{sh}\mathbf{u}(0))\mathbf{n}_h$ on Σ, with

$$|\lambda(R_{sh}\mathbf{u}(0))| = \frac{|(R_{sh}\mathbf{u}(0), \mathbf{n})_\Sigma|}{\|\mathbf{n}_h\|_\Sigma^2} = \frac{|(R_{sh}\mathbf{u}(0) - \mathbf{u}(0), \mathbf{n})_\Sigma|}{\|\mathbf{n}_h\|_\Sigma^2} \leqslant C\|R_{sh}\mathbf{u}(0) - \mathbf{u}(0)\|_\Sigma \leqslant Ch^{r+1}.$$

Therefore, using the triangle inequality, we have

$$\|\mathbf{u}(0) - R_h\mathbf{u}(0)\|_\Sigma \leqslant \|\mathbf{u}(0) - R_{sh}\mathbf{u}(0)\|_\Sigma + |\lambda(R_{sh}\mathbf{u}(0))|\|\mathbf{n}_h\|_\Sigma \leqslant Ch^{r+1},$$

where the estimate (3.4.49) is used.

Since $(R_h\mathbf{u}(0), R_hp(0))$ is essentially a Dirichlet Ritz projection with a different boundary value, i.e., $\widetilde{P}R_{sh}\mathbf{u}(0)$, the error estimates for $\|\mathbf{u}(0) - R_h\mathbf{u}(0)\|$ and $\|p(0) - R_hp(0)\|$ are the same as those in Lemma 3.4.1. With the optimal-order estimates of $\|\mathbf{u}(0) - R_h\mathbf{u}(0)\|_\Sigma$, $\|\mathbf{u}(0) - R_h\mathbf{u}(0)\|$ and $\|p(0) - R_hp(0)\|$, the estimation of $\|\boldsymbol{\eta}(0) - R_h\boldsymbol{\eta}(0)\|_\Sigma$ and $\|\boldsymbol{\eta}(0) - R_h\boldsymbol{\eta}(0)\|_s$ would be the same as the proof of Lemma 3.4.2. ∎

Next, we present estimates for the time derivatives $\partial_t(\mathbf{u} - R_h\mathbf{u})(0)$ and $\partial_t(p - R_hp)(0)$. To this end, we use the following relation:

$$(\mathbf{u} - R_h\mathbf{u})(0) = (\mathbf{u} - R_{sh}\mathbf{u})(0) + \lambda(R_{sh}\mathbf{u}(0))\mathbf{n}_h \quad \text{on } \Sigma. \tag{3.4.53}$$

Replacing $(\mathbf{u} - R_{sh}\mathbf{u})(0)$ by $(\mathbf{u} - R_h\mathbf{u})(0) - \lambda(R_{sh}\mathbf{u}(0))\mathbf{n}_h$ in (3.4.43), we have

$$\begin{aligned} &a_s((\mathbf{u} - R_h\mathbf{u})(0), \mathbf{v}_h) + ((\mathbf{u} - R_h\mathbf{u})(0), \mathbf{v}_h)_\Sigma + a_f((\partial_t\mathbf{u} - R_h^D\partial_t\mathbf{u})(0), \mathbf{v}_h) \\ &\quad - b((\partial_tp - R_h^D\partial_tp)(0), \mathbf{v}_h) + ((\partial_t\mathbf{u} - R_h^D\partial_t\mathbf{u})(0), \mathbf{v}_h) \\ &= \lambda(R_{sh}\mathbf{u}(0))(a_s(\mathbf{n}_h, \mathbf{v}_h) + (\mathbf{n}_h, \mathbf{v}_h)_\Sigma) \quad \forall \mathbf{v}_h \in \mathbf{X}_h^r. \end{aligned} \tag{3.4.54}$$

Let $(\mathbf{u}^\#, p^\#) \in \mathbf{X} \times Q$ be the weak solution of

$$a_f(\mathbf{u}^\#, \mathbf{v}) - b(p^\#, \mathbf{v}) + (\mathbf{u}^\#, \mathbf{v}) = a_s(\mathbf{n}, \mathbf{v}) + (\mathbf{n}, \mathbf{v})_\Sigma \qquad \forall \mathbf{v} \in \mathbf{X}, \tag{3.4.55a}$$

$$b(q, \mathbf{u}^\#) = 0 \qquad \forall q \in Q. \tag{3.4.55b}$$

Denote by $(\mathbf{u}_h^\#, p_h^\#) \in (\mathbf{X}_h^r, Q_h^{r-1})$ the corresponding FE solution satisfying

$$a_f(\mathbf{u}_h^\#, \mathbf{v}_h) - b(p_h^\#, \mathbf{v}_h) + (\mathbf{u}_h^\#, \mathbf{v}_h) = a_s(\mathbf{n}_h, \mathbf{v}_h) + (\mathbf{n}_h, \mathbf{v}_h)_\Sigma \quad \forall \mathbf{v}_h \in \mathbf{X}_h^r, \tag{3.4.56a}$$

$$b(q_h, \mathbf{u}_h^\#) = 0 \qquad \forall q_h \in Q_h^{r-1}, \tag{3.4.56b}$$

where $\mathbf{n}_h$ is defined in (3.4.38).

Note that (3.4.55) is equivalent to the weak solution of

$$\begin{aligned} &-\nabla \cdot \boldsymbol{\sigma}(\mathbf{u}^\#, p^\#) + \mathbf{u}^\# = \mathbf{0} \text{ in } \Omega \quad \text{with } \boldsymbol{\sigma}(\mathbf{u}^\#, p^\#)\mathbf{n} = -\mathcal{L}_s\mathbf{n} + \mathbf{n} \text{ on } \Sigma \\ &\nabla \cdot \mathbf{u}^\# = 0 \text{ in } \Omega. \end{aligned}$$

Therefore, from the regularity estimate in (3.2.7) (with $k = r - 1/2$ therein) and assumption (3.2.3) on $\mathcal{L}_s$, we obtain the following regularity estimate for the solutions of (3.4.55):

$$\|\mathbf{u}^{\#}\|_{H^{r+1}} + \|p^{\#}\|_{H^r} \le C\|\mathbf{n}\|_{H^{r+3/2}(\Sigma)} \le C.$$

By considering the difference between (3.4.55) and (3.4.56), the following estimates of $\mathbf{e}_h^{\#} := I_h\mathbf{u}^{\#} - \mathbf{u}_h^{\#}$ and $m_h^{\#} := I_h p^{\#} - p_h^{\#}$ can be derived for all $\mathbf{v}_h \in \mathbf{X}_h^r$ and $q_h \in Q_h^{r-1}$:

$$\begin{aligned} a_f(\mathbf{e}_h^{\#}, \mathbf{v}_h) - b(m_h^{\#}, \mathbf{v}_h) + (\mathbf{e}_h^{\#}, \mathbf{v}_h) &\le Ch^r\|\mathbf{v}_h\|_{H^1(\Sigma)} + Ch^r\|\mathbf{v}_h\|_{H^1} \le Ch^{r-1/2}\|\mathbf{v}_h\|_{H^1} \\ b(q_h, \mathbf{e}_h^{\#}) &\le Ch^r\|q_h\|, \end{aligned}$$

where we have used the inverse estimate in (3.2.11) and the following trace inequality:

$$\|\mathbf{v}_h\|_{H^1(\Sigma)} \le Ch^{-1/2}\|\mathbf{v}_h\|_{H^{1/2}(\Sigma)} \le Ch^{-1/2}\|\mathbf{v}_h\|_{H^1}.$$

From Korn's inequality and inf-sup condition (3.2.16), choosing $\mathbf{v}_h = \mathbf{e}_h^{\#}$ yields the following result:

$$\|\mathbf{e}_h^{\#}\|_{H^1} + \|m_h^{\#}\| \le Ch^{r-1/2},$$

which also implies the following boundedness through the application of the triangle inequality:

$$\|\mathbf{u}_h^{\#}\|_{H^1} + \|p_h^{\#}\| \le C.$$

By using the boundedness of $H^1(\Omega)$-norm of $\mathbf{u}_h^{\#}$ and $L^2(\Omega)$-norm of $p_h^{\#}$, we can estimate $\partial_t(\mathbf{u} - R_h\mathbf{u})(0)$ and $\partial_t(p - R_h p)(0)$ as follows.

Lemma 3.4.4. *Under the assumptions in Sections 3.2.2 and 3.2.3, the following error estimates hold (for the time derivative of the coupled Ritz projection in Definition 3.4.3)*:

$$\|\partial_t(\mathbf{u} - R_h\mathbf{u})(0)\| + \|\partial_t(\mathbf{u} - R_h\mathbf{u})(0)\|_\Sigma + h\|\partial_t(p - R_h p)(0)\| \le Ch^{r+1}. \tag{3.4.57}$$

Proof. By comparing (3.4.54) with (3.4.56a), and comparing (3.4.37) with (3.4.56b), we obtain

$$\begin{aligned} & a_s((\mathbf{u} - R_h\mathbf{u})(0), \mathbf{v}_h) + ((\mathbf{u} - R_h\mathbf{u})(0), \mathbf{v}_h)_\Sigma \\ & \quad + a_f((\partial_t\mathbf{u} - R_h^D\partial_t\mathbf{u})(0) - \lambda(R_{sh}\mathbf{u}(0))\mathbf{u}_h^{\#}, \mathbf{v}_h) - b((\partial_t p - R_h^D\partial_t p)(0) - \lambda(R_{sh}\mathbf{u}(0))p_h^{\#}, \mathbf{v}_h) \\ & \quad + ((\partial_t\mathbf{u} - R_h^D\partial_t\mathbf{u})(0) - \lambda(R_{sh}\mathbf{u}(0))\mathbf{u}_h^{\#}, \mathbf{v}_h) = 0 \quad \forall \mathbf{v}_h \in \mathbf{X}_h^r \qquad (3.4.58) \\ & b\left(q_h, (\partial_t\mathbf{u} - R_h^D\partial_t\mathbf{u})(0) - \lambda(R_{sh}u(0))\mathbf{u}_h^{\#}\right) = 0 \qquad \forall q_h \in Q_h^{r-1} \qquad (3.4.59) \end{aligned}$$

Then, by comparing (3.4.58)-(3.4.59) with (3.4.48a)-(3.4.48b), we find the following relations:

$$\begin{aligned} \partial_t(\mathbf{u} - R_h\mathbf{u})(0) &= \left(\partial_t\mathbf{u} - R_h^D\partial_t\mathbf{u}\right)(0) - \lambda(R_{sh}\mathbf{u}(0))\mathbf{u}_h^{\#}, \\ \partial_t(p - R_h p)(0) &= (\partial_t p - R_h^D\partial_t p)(0) - \lambda(R_{sh}\mathbf{u}(0))p_h^{\#}. \end{aligned}$$

Since $|\lambda(R_{sh}\mathbf{u}(0))| \le Ch^{r+1}$ and $\|\mathbf{u}_h^{\#}\| + \|\mathbf{u}_h^{\#}\|_\Sigma + \|p_h^{\#}\| \le C$, the result of this lemma follows from the estimates of the Dirichlet Stokes–Ritz projection in Lemma 3.4.1 (with $\mathbf{u}$ and p replaced by $\partial_t\mathbf{u}$ and $\partial_t p$ therein). ∎

3.4.3 Error estimates of the coupled Ritz projection for $t > 0$

In this subsection, using the results in the subsection 3.4.2, we present the proof of the H^1-error estimates and L^2-error estimates results in Theorem 3.3.1.

We first present H^1-norm error estimates for the coupled Ritz projection by employing the auxiliary Ritz projections R_h^S and R_h^D defined in (3.4.33) and (3.4.35), respectively. From (3.4.35b) we see that

$$R_h^D\mathbf{u} - R_h^S\mathbf{u} = \widetilde{P}R_h^S\mathbf{u} - R_h^S\mathbf{u} = -\lambda(R_h^S\mathbf{u})\mathbf{n}_h \quad \text{with } \lambda(R_h^S\mathbf{u}) \in \mathbb{R},$$

where the last equality follows from relation (3.4.39). Therefore, with the relation above we have

$$\begin{aligned} & a_s(\mathbf{u} - R_h^D\mathbf{u}, \mathbf{v}_h) + (\mathbf{u} - R_h^D\mathbf{u}, \mathbf{v}_h)_\Sigma \\ &= a_s(\mathbf{u} - R_h^S\mathbf{u}, \mathbf{v}_h) + (\mathbf{u} - R_h^S\mathbf{u}, \mathbf{v}_h)_\Sigma + \lambda(R_h^S\mathbf{u})\,(a_s(\mathbf{n}_h, \mathbf{v}_h) + (\mathbf{n}_h, \mathbf{v}_h)_\Sigma) \\ &\le Ch^{r+1}\|\mathbf{v}_h\|_{H^1(\Sigma)} \le Ch^{r+1/2}\|\mathbf{v}_h\|_{H^{1/2}(\Sigma)} \le Ch^{r+1/2}\|\mathbf{v}_h\|_{H^1} \quad \forall \mathbf{v}_h \in \mathbf{X}_h^r, \end{aligned} \tag{3.4.60}$$

where we have used the inverse inequality in (3.2.11) and the trace inequality in the derivation of the last two inequalities. Moreover, since the auxiliary Ritz projection R_h^D defined in (3.4.35) is time-independent, it follows that $(\partial_t R_h^D u, \partial_t R_h^D p) = (R_h^D \partial_t u, R_h^D \partial_t p)$. Therefore, in view of estimate (3.4.41) for the Dirichlet Stokes–Ritz projection, the following estimate can be found:

$$\begin{aligned} & a_s(\mathbf{u} - R_h^D\mathbf{u}, \mathbf{v}_h) + (\mathbf{u} - R_h^D\mathbf{u}, \mathbf{v}_h)_\Sigma + a_f(\partial_t(\mathbf{u} - R_h^D\mathbf{u}), \mathbf{v}_h) \\ & - b(\partial_t(p - R_h^D p), \mathbf{v}_h) + (\partial_t(\mathbf{u} - R_h^D\mathbf{u}), \mathbf{v}_h) \le Ch^r\|\mathbf{v}_h\|_{H^1} \quad \forall \mathbf{v}_h \in \mathbf{X}_h^r. \end{aligned} \tag{3.4.61}$$

By considering the difference between (3.4.48a) and (3.4.61), we can derive the following inequality:

$$\begin{aligned} & a_s(R_h\mathbf{u} - R_h^D\mathbf{u}, \mathbf{v}_h) + (R_h\mathbf{u} - R_h^D\mathbf{u}, \mathbf{v}_h)_\Sigma + a_f(\partial_t(R_h\mathbf{u} - R_h^D\mathbf{u}), \mathbf{v}_h) \\ & - b(\partial_t(R_h p - R_h^D p), \mathbf{v}_h) + (\partial_t(R_h\mathbf{u} - R_h^D\mathbf{u}), \mathbf{v}_h) \le Ch^r\|\mathbf{v}_h\|_{H^1} \quad \forall \mathbf{v}_h \in \mathbf{X}_h^r. \end{aligned} \tag{3.4.62}$$

Then, choosing $\mathbf{v}_h = \partial_t(R_h\mathbf{u} - R_h^D\mathbf{u})$ in (3.4.62) and using relation $b(\partial_t(R_h p - R_h^D p), \partial_t(R_h\mathbf{u} - R_h^D\mathbf{u})) = 0$ (which follows from (3.4.37) and (3.4.48b)), using Young's inequality

$$Ch^r\|\partial_t(R_h\mathbf{u} - R_h^D\mathbf{u})\|_{H^1} \leqslant C\varepsilon^{-1}h^{2r} + \varepsilon\|\partial_t(R_h\mathbf{u} - R_h^D\mathbf{u})\|_{H^1}^2$$

with a small constant ε so that $\varepsilon\|\partial_t(R_h\mathbf{u} - R_h^D\mathbf{u})\|_{H^1}^2$ can be absorbed by the left hand side of (3.4.62), we obtain

$$\begin{aligned} & \|R_h\mathbf{u} - R_h^D\mathbf{u}\|_{L^\infty H^1(\Sigma)} + \|\partial_t(R_h\mathbf{u} - R_h^D\mathbf{u})\|_{L^2H^1} \\ & \le Ch^r + C\|(R_h\mathbf{u} - R_h^D\mathbf{u})(0)\|_s + C\|(R_h\mathbf{u} - R_h^D\mathbf{u})(0)\|_\Sigma \le Ch^r, \end{aligned} \tag{3.4.63}$$

where the last inequality uses the estimates in Lemma 3.4.3 and Lemma 3.4.1. Then, by applying the inf-sup condition in (3.2.16) (which involves $\|\mathbf{v}_h\|_{H^1(\Sigma)}$ in the denominator), we can obtain the following estimate from (3.4.62):

$$\|\partial_t(R_h p - R_h^D p)\| \le C\|R_h\mathbf{u} - R_h^D\mathbf{u}\|_{H^1(\Sigma)} + C\|\partial_t(R_h\mathbf{u} - R_h^D\mathbf{u})\|_{H^1} + Ch^r, \tag{3.4.64}$$

which combined with the estimate in (3.4.63), leads to the following estimate:

$$\|\partial_t(R_h p - R_h^D p)\|_{L^2L^2} \le Ch^r. \tag{3.4.65}$$

Therefore, using an additional triangle inequality, the estimates in (3.4.63)–(3.4.65) can be written as follows:

$$\|\partial_t(R_h\mathbf{u}-\mathbf{u})\|_{L^2H^1} + \|R_h\mathbf{u}-\mathbf{u}\|_{L^\infty H^1(\Sigma)} + \|\partial_t(R_hp-p)\|_{L^2L^2} \le Ch^r. \tag{3.4.66}$$

With the initial estimates in Lemma 3.4.3, the estimate of $\|\partial_t(R_h\mathbf{u}-\mathbf{u})\|_{L^2H^1}$ above further implies that

$$\|R_h\mathbf{u}-\mathbf{u}\|_{L^\infty H^1} \le \|(R_h\mathbf{u}-\mathbf{u})(0)\|_{H^1} + C\|\partial_t(R_h\mathbf{u}-\mathbf{u})\|_{L^2H^1} \le Ch^r\,. \tag{3.4.67}$$

Since $\partial_t(R_h\eta-\eta) = R_h\mathbf{u}-\mathbf{u}$ on the boundary Σ, by using the Newton–Leibniz formula with respect to $t\in[0,T]$, the estimate in (3.4.66) and initial estimates in Lemma 3.4.3, we have

$$\begin{aligned}\|R_h\eta-\eta\|_{L^\infty H^1(\Sigma)} &\le \|(R_h\eta-\eta)(0)\|_{H^1(\Sigma)} + C\|\partial_t(R_h\eta-\eta)\|_{L^2H^1(\Sigma)}\\ &\le \|(R_h\eta-\eta)(0)\|_{H^1(\Sigma)} + C\|R_h\mathbf{u}-\mathbf{u}\|_{L^2H^1(\Sigma)} \le Ch^r.\end{aligned} \tag{3.4.68}$$

In the same way, from (3.4.66) and initial estimates in Lemma 3.4.3 we have

$$\|R_hp-p\|_{L^\infty L^2} \le C\|(R_hp-p)(0)\| + C\|R_hp-\mathbf{u}\|_{L^2L^2} \le Ch^r. \tag{3.4.69}$$

Thus we can summarize what we have proved as follows:

$$\begin{aligned}&\|R_h\mathbf{u}-\mathbf{u}\|_{L^\infty H^1} + \|R_h\mathbf{u}-\mathbf{u}\|_{L^\infty H^1(\Sigma)} + \|R_hp-p\|_{L^\infty L^2}\\ &+ \|R_h\eta-\eta\|_{L^\infty H^1(\Sigma)} + \|\partial_t(R_h\mathbf{u}-\mathbf{u})\|_{L^2H^1} + \|\partial_t(R_hp-p)\|_{L^2L^2} \le Ch^r.\end{aligned} \tag{3.4.70}$$

Moreover, by differentiating (3.4.48) with respect to time, we have

$$\begin{aligned}&a_s(\partial_t(R_h\mathbf{u}-\mathbf{u}),\mathbf{v}_h) + (\partial_t(R_h\mathbf{u}-\mathbf{u}),\mathbf{v}_h)_\Sigma + a_f(\partial_t^2(R_h\mathbf{u}-\mathbf{u}),\mathbf{v}_h)\\ &\quad - b(\partial_t^2(R_hp-p),\mathbf{v}_h) + (\partial_t^2(R_h\mathbf{u}-\mathbf{u}),\mathbf{v}_h) = 0 \qquad \forall \mathbf{v}_h\in\mathbf{X}_h^r, \quad (3.4.71\text{a})\\ &b(q_h,\partial_t^2(R_h\mathbf{u}-\mathbf{u})) = 0 \qquad \forall q_h\in Q_h^{r-1}. \quad (3.4.71\text{b})\end{aligned}$$

Similarly, by choosing $\mathbf{v}_h = \partial_t^2(R_h\mathbf{u}-R_h^D\mathbf{u})$ in (3.4.71a) and using the same approach as above with the initial value estimates in (3.4.57), we can obtain the following estimate (the details are omitted):

$$\begin{aligned}&\|\partial_t(R_h\mathbf{u}-\mathbf{u})\|_{L^\infty H^1} + \|\partial_t(R_h\mathbf{u}-\mathbf{u})\|_{L^\infty H^1(\Sigma)} + \|\partial_t(R_hp-p)\|_{L^\infty L^2}\\ &\qquad + \|\partial_t^2(R_h\mathbf{u}-\mathbf{u})\|_{L^2H^1} + \|\partial_t^2(R_hp-p)\|_{L^2L^2} \le Ch^r\,.\end{aligned} \tag{3.4.72}$$

(3.4.70) and (3.4.72) establish the H^1-norm error estimates for the coupled non-stationary Ritz projection defined in (3.3.3).

We then present L^2-norm error estimates for the coupled non-stationary Ritz projection. To this end, we introduce the following dual problem:

$$\begin{aligned}-\mathcal{L}_s\boldsymbol{\phi} + \boldsymbol{\phi} &= \partial_t\boldsymbol{\sigma}(\boldsymbol{\phi},q)\mathbf{n} + \mathbf{f} && \text{in } \Sigma && (3.4.73\text{a})\\ -\nabla\cdot\sigma(\boldsymbol{\phi},q) + \boldsymbol{\phi} &= 0 && \text{in } \Omega && (3.4.73\text{b})\\ \nabla\cdot\boldsymbol{\phi} &= 0 && \text{in } \Omega, && (3.4.73\text{c})\end{aligned}$$

with the initial condition $\boldsymbol{\sigma}(\boldsymbol{\phi},q)\mathbf{n} = 0$ at $t = T$. Problem (3.4.73) can be equivalently written as a backward evolution equation of $\boldsymbol{\xi} = \boldsymbol{\sigma}(\boldsymbol{\phi},q)\mathbf{n}$, i.e.,

$$-\mathcal{L}_s\mathcal{N}\boldsymbol{\xi} + \mathcal{N}\boldsymbol{\xi} - \partial_t\boldsymbol{\xi} = \mathbf{f} \text{ on } \Sigma\times[0,T), \text{ with initial condition } \boldsymbol{\xi}(T) = 0, \tag{3.4.74}$$

where $\mathcal{N} : H^{-\frac{1}{2}}(\Sigma)^d \to H^{\frac{1}{2}}(\Sigma)^d$ is the Neumann-to-Dirichlet map associated to the Stokes equations. The existence, uniqueness and regularity of solutions to (3.4.73) are presented in the following lemma, for which the proof is given in Section 3.7 by utilizing and analyzing (3.4.74).

Lemma 3.4.5. *Problem* (3.4.73) *has a unique solution which satisfies the following estimate:*

$$\|\boldsymbol{\phi}\|_{L^2H^2} + \|\boldsymbol{\phi}\|_{L^2H^2(\Sigma)} + \|q\|_{L^2H^1} + \|\boldsymbol{\sigma}(\boldsymbol{\phi}, q)(0)\mathbf{n}\|_\Sigma \le C\|\mathbf{f}\|_{L^2L^2(\Sigma)}. \tag{3.4.75}$$

By choosing $\mathbf{f} = R_h\boldsymbol{\eta} - \boldsymbol{\eta}$ and, testing equations (3.4.73a) and (3.4.73b) with $R_h\boldsymbol{\eta} - \boldsymbol{\eta}$ and $R_h\mathbf{u} - \mathbf{u}$, respectively, and using relation $\partial_t(R_h\boldsymbol{\eta} - \boldsymbol{\eta}) = R_h\mathbf{u} - \mathbf{u}$ on Σ, we have

$$\begin{aligned}
&a_s(\boldsymbol{\phi}, R_h\boldsymbol{\eta} - \boldsymbol{\eta}) + (\boldsymbol{\phi}, R_h\boldsymbol{\eta} - \boldsymbol{\eta})_\Sigma + a_f(\boldsymbol{\phi}, R_h\mathbf{u} - \mathbf{u}) - b(q, R_h\mathbf{u} - \mathbf{u}) + (\boldsymbol{\phi}, R_h\mathbf{u} - \mathbf{u}) \\
&= \frac{d}{dt}(\boldsymbol{\sigma}(\boldsymbol{\phi}, q)\cdot\mathbf{n}, R_h\boldsymbol{\eta} - \boldsymbol{\eta})_\Sigma + \|R_h\boldsymbol{\eta} - \boldsymbol{\eta}\|_\Sigma^2.
\end{aligned}$$

In view of the definition of the non-stationary Ritz projection in (3.3.3), we can subtract $I_h\boldsymbol{\phi}$ from $\boldsymbol{\phi}$ in the inequality above by generating an additional remainder $b(R_hp - p, \boldsymbol{\phi} - I_h\boldsymbol{\phi})$. This leads to the following result in view of the estimate in (3.4.66):

$$\begin{aligned}
&\frac{d}{dt}(\boldsymbol{\sigma}(\boldsymbol{\phi}, q)\mathbf{n}, R_h\boldsymbol{\eta} - \boldsymbol{\eta})_\Sigma + \|R_h\boldsymbol{\eta} - \boldsymbol{\eta}\|_\Sigma^2 = a_s(\boldsymbol{\phi} - I_h\boldsymbol{\phi}, R_h\boldsymbol{\eta} - \boldsymbol{\eta}) + (\boldsymbol{\phi} - I_h\boldsymbol{\phi}, R_h\boldsymbol{\eta} - \boldsymbol{\eta})_\Sigma \\
&\quad + a_f(\boldsymbol{\phi} - I_h\boldsymbol{\phi}, R_h\mathbf{u} - \mathbf{u}) - b(q - I_hq, R_h\mathbf{u} - \mathbf{u}) + (\boldsymbol{\phi} - I_h\boldsymbol{\phi}, R_h\mathbf{u} - \mathbf{u}) - b(R_hp - p, \boldsymbol{\phi} - I_h\boldsymbol{\phi}) \\
&\leqslant Ch^{r+1}(\|\boldsymbol{\phi}\|_{H^2} + \|\boldsymbol{\phi}\|_{H^2(\Sigma)} + \|q\|_{H^1}).
\end{aligned}$$

Since $\|(R_h\boldsymbol{\eta} - \boldsymbol{\eta})(0)\|_\Sigma \leqslant Ch^{r+1}$ (see Lemma 3.4.3), the inequality above leads to the following result:

$$\begin{aligned}
&\|R_h\boldsymbol{\eta} - \boldsymbol{\eta}\|_{L^2L^2(\Sigma)}^2 \\
&\le Ch^{r+1}\|R_h\boldsymbol{\eta} - \boldsymbol{\eta}\|_{L^2L^2(\Sigma)} + \|R_h\boldsymbol{\eta}(0) - \boldsymbol{\eta}(0)\|_{L^2(\Sigma)}\|(\boldsymbol{\sigma}(\boldsymbol{\phi}, q)\mathbf{n})(0)\|_{L^2(\Sigma)} \\
&\le Ch^{r+1}\|R_h\boldsymbol{\eta} - \boldsymbol{\eta}\|_{L^2L^2(\Sigma)} + Ch^{r+1}\|R_h\boldsymbol{\eta} - \boldsymbol{\eta}\|_{L^2L^2(\Sigma)},
\end{aligned}$$

and therefore

$$\|R_h\boldsymbol{\eta} - \boldsymbol{\eta}\|_{L^2L^2(\Sigma)} \le Ch^{r+1}. \tag{3.4.76}$$

By using the same approach, choosing $\mathbf{f} = R_h\mathbf{u} - \mathbf{u}$ and $\mathbf{f} = \partial_t(R_h\mathbf{u} - \mathbf{u})$ in (3.4.73a), respectively, the following result can be shown (the details are omitted):

$$\|R_h\mathbf{u} - \mathbf{u}\|_{L^2L^2(\Sigma)} + \|\partial_t(R_h\mathbf{u} - \mathbf{u})\|_{L^2L^2(\Sigma)} \le Ch^{r+1}. \tag{3.4.77}$$

This also implies, via the Newton–Leibniz formula in time,

$$\|R_h\boldsymbol{\eta} - \boldsymbol{\eta}\|_{L^\infty L^2(\Sigma)} + \|R_h\mathbf{u} - \mathbf{u}\|_{L^\infty L^2(\Sigma)} \le Ch^{r+1}\,. \tag{3.4.78}$$

Furthermore, we consider a dual problem defined by

$$\begin{cases}
-\nabla\cdot\sigma(\boldsymbol{\phi}, q) + \boldsymbol{\phi} = R_h\mathbf{u} - \mathbf{u} & \text{in } \mathit{\Omega} \\
\nabla\cdot\boldsymbol{\phi} = 0 & \text{in } \mathit{\Omega} \\
\boldsymbol{\phi}|_\Sigma = 0, \quad q \in L_0^2(\mathit{\Omega}),
\end{cases} \tag{3.4.79}$$

which satisfies the following standard H^2 regularity estimate

$$\|\phi\|_{H^2} + \|q\|_{H^1} + \|\sigma(\phi, q)\mathbf{n}\|_{L^2(\Sigma)} \leq C\|R_h\mathbf{u} - \mathbf{u}\|,$$

where the term $\|\sigma(\phi, q)\mathbf{n}\|_{L^2(\Sigma)}$ is included on the left-hand side because it is actually bounded by $\|\phi\|_{H^2} + \|q\|_{H^1}$. Then, testing (3.4.79) with $R_h\mathbf{u} - \mathbf{u}$, we have

$$\begin{aligned}
&\|R_h\mathbf{u} - \mathbf{u}\|^2 \\
&= a_f(\phi, R_h\mathbf{u} - \mathbf{u}) - b(q, R_h\mathbf{u} - \mathbf{u}) + (\phi, R_h\mathbf{u} - \mathbf{u}) - (\boldsymbol{\sigma}(\phi, q)\mathbf{n}, R_h\mathbf{u} - \mathbf{u})_\Sigma \\
&= a_f(\phi - I_h\phi, R_h\mathbf{u} - \mathbf{u}) - b(q - I_hq, R_h\mathbf{u} - \mathbf{u}) - (\boldsymbol{\sigma}(\phi, q)\mathbf{n}, R_h\mathbf{u} - \mathbf{u})_\Sigma \\
&\quad + (\phi - I_h\phi, R_h\mathbf{u} - \mathbf{u}) - b(R_hp - p, \phi - I_h\phi) \quad \text{(as a result of (3.3.3) with } \mathbf{v}_h = I_h\phi, q_h = I_hq) \\
&\leqslant Ch(\|\phi\|_{H^2} + \|q\|_{H^1})(\|R_h\mathbf{u} - \mathbf{u}\|_{H^1} + \|R_hp - p\|) \\
&\quad + \|\boldsymbol{\sigma}(\phi, q) \cdot \mathbf{n}\|_\Sigma \|R_h\mathbf{u} - \mathbf{u}\|_\Sigma \\
&\leqslant Ch^{r+1}\|R_h\mathbf{u} - \mathbf{u}\| + C\|R_h\mathbf{u} - \mathbf{u}\|\|R_h\mathbf{u} - \mathbf{u}\|_\Sigma.
\end{aligned}$$

The last inequality implies, in combination with (3.4.78), the following result:

$$\|R_h\mathbf{u} - \mathbf{u}\| \leq Ch^{r+1}. \tag{3.4.80}$$

By using the same approach, replacing $R_h\mathbf{u} - \mathbf{u}$ by $\partial_t(R_h\mathbf{u} - \mathbf{u})$ in (3.4.79), the following estimate can be shown (the details are omitted):

$$\|\partial_t(R_h\mathbf{u} - \mathbf{u})\|_{L^2L^2} \leq Ch^{r+1}. \tag{3.4.81}$$

The proof of Theorem 3.3.1 is complete. ∎

3.5 Numerical examples

In this section, we present numerical tests to support the theoretical analysis in this chapter and to show the efficiency of the proposed algorithm. For 2D numerical examples, the operator $\mathcal{L}_s\eta = C_0\partial_{xx}\eta - C_1\eta$ on the interface Σ is considered. All computations are performed by the finite element package NGSolve; see [128].

Example 3.5.1. To test the convergence rate of the algorithm, we consider an artificial example of two-dimensional thin structure models given in (3.1.1)–(3.1.3) with extra source terms such that the exact solution is given by

$$\begin{aligned}
u_1 &= 4\sin(2\pi x)\sin(2\pi y)\sin(t), \\
u_2 &= 4(\cos(2\pi x)\cos(2\pi y))\sin(t), \\
p &= 8(\cos(4\pi x) - \cos(4\pi y))\sin(t), \\
\eta_1 &= 0, \quad \eta_2 = -4\cos(2\pi x)\cos(t).
\end{aligned}$$

First, we examine this problem involving left/right-side periodic boundary conditions and top/bottom interfaces in the domain $\bar{\Omega} = [0, 2] \times [0, 1]$. A uniform triangular partition is employed, featuring $M + 1$ vertices in the y-direction and $2M + 1$ vertices in the x-direction, where $h = 1/M$. The classical lowest-order Taylor–Hood element is utilized for spatial discretization. For simplicity, we set all involved parameters to 1. Our algorithm is applied to solve the system with $M = 8, 16, 32$, $\tau = h^3$, and the terminal time $T = 0.1$. The numerical results are presented in the Table 3.1, which shows that the algorithm has

Table 3.1: The convergence order of the algorithm under periodic boundary conditions

Taylor–Hood elements ($\tau = h^3$)	$\|\mathbf{u}^N - \mathbf{u}_h^N\|$	$\|p^N - p_h^N\|$	$\|\boldsymbol{\eta}^N - \boldsymbol{\eta}_h^N\|_\Sigma$	$\|\boldsymbol{\eta}^N - \boldsymbol{\eta}_h^N\|_s$
$h = 1/8$	6.852e-3	1.403e-1	1.324e-2	8.075e-1
$h = 1/16$	6.848e-4	2.691e-2	1.644e-3	2.029e-1
$h = 1/32$	7.937e-5	6.297e-3	2.052e-4	5.079e-2
order	3.10	2.10	3.00	2.00

the third-order accuracy for the velocity and the displacement in the L^2-norm, as well as the second-order accuracy for the pressure in the L^2-norm and the displacement in the energy-norm. These numerical results align with our theoretical analysis.

Next, we test our algorithm for the case of the left/right-side Dirichlet boundary conditions, using the same configuration as previously described. Both the lowest-order Taylor-Hood element and the MINI element are employed for spatial discretization. We set $\tau = h^3$ and $\tau = h^2$ for the Taylor-Hood element and the MINI element, respectively. The numerical results are displayed in the Table 3.2. As observed in the Table 3.2, the algorithm, when paired with both the Taylor–Hood element and the MINI element, yields numerical results exhibiting optimal convergence orders for $\mathbf{u}$ and η.

Table 3.2: The convergence order of the algorithm under Dirichlet boundary conditions

Taylor–Hood elements ($\tau = h^3$)	$\|u^N - u_h^N\|$	$\|p^N - p_h^N\|$	$\|\boldsymbol{\eta}^N - \boldsymbol{\eta}_h^N\|_\Sigma$	$\|\boldsymbol{\eta}^N - \boldsymbol{\eta}_h^N\|_s$
$h = 1/8$	4.553e-3	1.354e-1	1.313e-2	8.069e-1
$h = 1/16$	6.009e-4	2.775e-2	1.645e-3	2.029e-1
$h = 1/32$	7.693e-5	6.470e-3	2.055e-4	5.079e-2
order	2.97	2.10	3.00	2.00
MINI elements ($\tau = h^2$)	$\|u^N - u_h^N\|$	$\|p^N - p_h^N\|$	$\|\boldsymbol{\eta}^N - \boldsymbol{\eta}_h^N\|_\Sigma$	$\|\boldsymbol{\eta}^N - \boldsymbol{\eta}_h^N\|_s$
$h = 1/16$	1.324e-2	3.186e-1	7.971e-2	4.001e0
$h = 1/32$	3.349e-3	1.192e-1	1.999e-2	2.003e0
$h = 1/64$	8.327e-4	4.641e-2	5.001e-3	1.002e0
order	2.00	1.36	2.00	1.00

Example 3.5.2. We consider a benchmark model which was studied by many researchers [25, 26, 54, 57, 72, 109, 120]. All the quantities will be given in the CGS system of units [54]. The model is described by (3.1.1)–(3.1.3) in $\bar{\Omega} = [0, 5] \times [0, 0.5]$ with the physical parameters: fluid density $\rho_f = 1$, fluid viscosity $\mu = 0.035$, solid density $\rho_s = 1.1$, the thickness of wall $\epsilon_s = 0.1$, Young's modulus $E = 0.75 \times 10^6$, Poisson's ratio $\sigma = 0.5$ and

$$C_0 = \frac{E\epsilon_s}{2(1+\sigma)}, \quad C_1 = \frac{E\epsilon_s}{R^2(1-\sigma^2)},$$

where $R = 0.5$ is the width of the domain Ω. The boundary conditions on the in/out-flow sides ($x = 0, x = 5$) are defined by $\sigma(\mathbf{u}, p)\mathbf{n} = -p_{\text{in/out}}\mathbf{n}$ where

$$p_{\text{in}}(t) = \begin{cases} \frac{p_{\max}}{2}\left[1 - \cos\left(\frac{2\pi t}{t_{\max}}\right)\right] & \text{if } t \le t_{\max} \\ 0 & \text{if } t > t_{\max} \end{cases}, \quad p_{\text{out}}(t) = 0 \ \ \forall t \in (0, T].$$

with $p_{\max} = 1.3333 \times 10^4$ and $t_{\max} = 0.003$. The top and bottom sides of Ω are thin structures, and the fluid is initially at rest. We take a uniform triangular partition with $M+1$ vertices in y-direction and $10M+1$ vertices in x-direction ($h = 1/M$), and solve the system by our algorithm where the lowest-order Taylor–Hood finite element approximation is used with the spatial mesh size $h = 1/64$ ($M = 64$), the temporal step size $\tau = h^3$ and the parameter $\beta = 0.5$. We present the contour of pressure p in the Figure 3.2 at $t = 0.003, 0.009, 0.016, 0.026$ (from top to bottom). We can see a forward moving pressure wave(red), which reaches the right-end of the domain and gets reflected. The reflected wave is characterized by the different color(blue), which was also observed in [54, 57, 72].

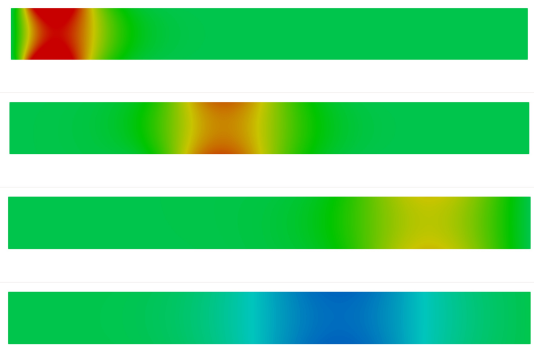


Figure 3.2: The contour of the pressure when $t = 0.003, 0.009, 0.016, 0.026$ (from top to bottom)

Example 3.5.3. We consider an example of 3D blood flow simulation in common carotid arteries studied in [120]. The blood flow is modeled by the Navier-Stokes equation, while our analysis was presented only for the model with the Stokes equation. The weak form of the arterial wall model is:

$$\rho_s\epsilon_s(\eta_{tt}, \mathbf{w})_\Sigma + D_1(\eta, \mathbf{w})_\Sigma + D_2(\eta_t, \mathbf{w})_\Sigma + \epsilon_s(\mathbf{\Pi}_s(\eta), \nabla_s \mathbf{w})_\Sigma = (-\sigma(\mathbf{u}, p)\mathbf{n}, \mathbf{w})_\Sigma$$

for any $\mathbf{w} \in \mathbf{S}$, where ∇_s denote the surface gradient on the interface Σ and

$$\mathbf{\Pi}_s(\eta) = \frac{E}{1+\sigma^2}\frac{\nabla_s\eta + \nabla_s^T\eta}{2} + \frac{E\sigma}{1-\sigma^2}\nabla_s \cdot \eta\mathbf{I}$$

for a linearly elastic isotropic structure. The geometrical domain is a straight cylinder of length 4 cm and radius 0.3 cm, see the Fugure 3.3. The hemodynamical parameters used in this model are given in the Table 3.3. For the inlet and outlet boundary conditions, we set

$$\mathbf{u} = (u_D(t)\frac{R^2 - r^2}{R^2}, 0, 0) \quad \text{on } \Sigma_{in} \qquad \text{and} \quad \sigma(\mathbf{u}, p)\mathbf{n} = -p_{out}(t)\mathbf{n} \quad \text{on } \Sigma_{out}.$$

The given data for $u_D(t)$ and $p_{out}(t)$, as shown in the Figure 5.2, are taken from [120]. Mmore realistic and delicate treatment of boundary conditions can be found in [56].

The fluid mesh used in this example consists of 11745 tetrahedra, and the structure mesh consists of 3786 triangles. We utilize the $P2-P1$ finite element approximation for the velocity and pressure of the fluid, the P2 finite element approximation for the displacement of the structure. For comparison, both classical monolithic scheme and the proposed partitioned scheme are implemented to solve this example, where the parameter $\beta = 0.5$. The initial velocity/pressure is the smooth constant extension of the inlet/outlet

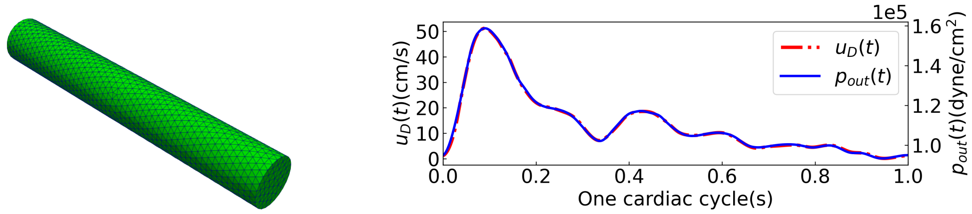


Figure 3.3: The geometrical domain(left) and the given data for $u_D(t)$ and $p_{out}(t)$(right)

Table 3.3: The hemodynamical parameters in the PDE model

Parameter	Value	Parameter	Value
Wall thickness ϵ_s(cm)	0.06	Poisson's ratio σ	0.5
Fluid viscosity μ(g/cm s)	0.04	Young's modulo E(dyne/cm^2)	$2.6 \cdot 10^6$
Fluid density ρ_f(g/cm^3)	1	Coefficient D_1(dyne/cm^3)	$6 \cdot 10^5$
Wall density ρ_s(g/cm^3)	1.1	Coefficient D_2(dyne s/cm^3)	$2 \cdot 10^5$

boundary data at $t = 0$ for both schemes. The terminal time $T = 3$ s which corresponds to 3 cardiac cycles. We have observed that the periodic pattern was established after 1 cardiac cycle. Some comparison between monolithic and partitioned schemes is done. In the Figure 3.4, the magnitude of the radial displacement for the artery wall is shown at the interface point $(2, 0.3, 0)$ in the whole 3 cardiac cycles. In the Figures 3.5 and 3.6, the axial velocity and the pressure are presented at the center point $(2, 0, 0)$ in the third cardiac cycle, respectively. The waveforms of velocity and pressure are generally not be the same. The difference waveforms between velocity and pressure can be observed in the numerical results by comparing Figure 5.4 and Figure 5.5.

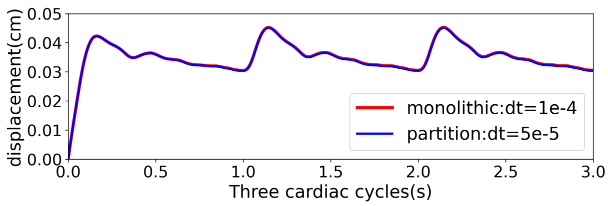


Figure 3.4: Comparison of the radial displacement

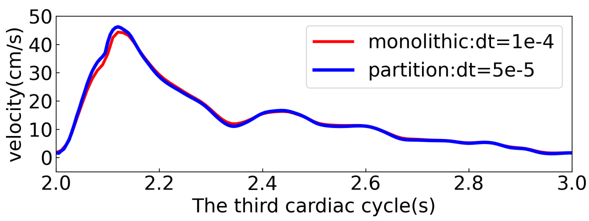


Figure 3.5: Comparison of the axial velocity

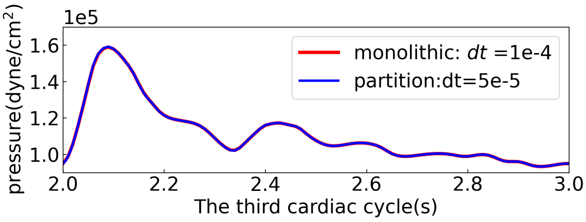


Figure 3.6: Comparison of the pressure

3.6 Conclusion

We have proposed a new stable fully-discrete kinematically coupled scheme which decouples fluid velocity from the structure displacement for solving a thin-structure interaction problem described by (3.1.1)–(3.1.3). To the best of our knowledge, the optimal-order convergence in L^2 norm of spatially finite element methods for such problems has not been established in the previous works. Our scheme in (3.2.19)–(3.2.20) contains two stabilization terms

$$\rho_s\epsilon_s\left(\frac{\mathbf{u}_h^n-\mathbf{s}_h^n}{\tau},\ \frac{\tau}{\rho_s\epsilon_s}\boldsymbol{\sigma}(\mathbf{v}_h,q_h)\cdot\mathbf{n}\right)_\Sigma \quad \text{and} \quad \left((\boldsymbol{\sigma}_h^n-\boldsymbol{\sigma}_h^{n-1})\cdot\mathbf{n},\ \frac{\tau(1+\beta)}{\rho_s\epsilon_s}\boldsymbol{\sigma}(\mathbf{v}_h,q_h)\cdot\mathbf{n}\right)_\Sigma$$

which guarantee the unconditional stability of the method, and an additional parameter $\beta > 0$ which is helpful for us to prove optimal-order convergence in the L^2 norm for the fully discrete finite element scheme. Moreover, we have developed a new approach for the numerical analysis of such thin-structure interaction problems in terms of a newly introduced coupled non-stationary Ritz projection, with rigorous analysis for its approximation properties through analyzing its dual problem, which turns out to be equivalent to a backward evolution equation on the boundary Σ, i.e.,

$$-\mathcal{L}_s\mathcal{N}\boldsymbol{\xi}+\mathcal{N}\boldsymbol{\xi}-\partial_t\boldsymbol{\xi}=\mathbf{f}\ \text{ on }\ \Sigma\times[0,T),\ \ \text{with initial condition }\ \boldsymbol{\xi}(T)=0,$$

in terms of the Neumann-to-Dirichlet map $\mathcal{N}: H^{-\frac{1}{2}}(\Sigma)^d \to H^{\frac{1}{2}}(\Sigma)^d$ associated to the Stokes equations. Although we have focused on the analysis for the specific kinematically coupled scheme proposed in this chapter for a thin-structure interaction problem, the new approach developed in this chapter, including the non-stationary Ritz projection and its approximation properties, may be extended to many other fully-discrete monolithic and partitioned coupled algorithms and to more general fluid-structure interaction models.

3.7 Appendix A: Proof of Lemma 3.4.5

In this appendix, we prove Lemma 3.4.5 via the following proposition, where equation (3.7.1) differs from (3.4.73) via a change of variable $t \to T-t$ in time.

Proposition 3.7.1. *The initial-boundary value problem*

$$\begin{aligned}
-\mathcal{L}_s\boldsymbol{\phi}+\boldsymbol{\phi}&=-\partial_t\boldsymbol{\sigma}(\boldsymbol{\phi},q)\mathbf{n}+\mathbf{f} && \textit{on } \Sigma\times(0,T] \quad \text{(the boundary condition)} && (3.7.1a)\\
-\nabla\cdot\sigma(\boldsymbol{\phi},q)+\boldsymbol{\phi}&=0 && \textit{in } \Omega\times(0,T] && (3.7.1b)\\
\nabla\cdot\boldsymbol{\phi}&=0 && \textit{in } \Omega\times(0,T] && (3.7.1c)\\
\boldsymbol{\sigma}(\boldsymbol{\phi},q)\mathbf{n}&=0 && \textit{at } t=0 \quad \text{(the initial condition)}, && (3.7.1d)
\end{aligned}$$

has a unique solution (ϕ, q) *which satisfies the following regularity estimate:*

$$\|\phi\|_{L^2H^2} + \|\phi\|_{L^2H^2(\Sigma)} + \|q\|_{L^2H^1} + \|\boldsymbol{\sigma}(\phi, q)\mathbf{n}\|_{L^\infty L^2(\Sigma)} \le C\|\mathbf{f}\|_{L^2L^2(\Sigma)} \tag{3.7.2}$$

Proof. We divide the proof into three parts. In the first part, we introduce the Neumann-to-Dirichlet operator and reformulate (3.7.1) into an evolution equation (3.7.4) on the boundary Σ with the aid of Neumann-to-Dirichlet operator, and then establish some mapping properties of the Neumann-to-Dirichlet operator to be used in the proof of Proposition 3.7.1. In the second part, we establish the existence, uniqueness and regularity of solutions to an equivalent formulation of (3.7.1), i.e., equation (3.7.4) below. Finally, in the third part, we establish regularity estimates for the solutions to (3.7.1).

Part 1. We can define the Neumann-to-Dirichlet operator $\mathcal{N} : H^{-1/2}(\Sigma)^d \to H^{1/2}(\Sigma)^d$ as $\boldsymbol{\zeta} \mapsto (\mathcal{N}^v\boldsymbol{\zeta})|_\Sigma$, with $(\mathcal{N}^v\boldsymbol{\zeta}, \mathcal{N}^p\boldsymbol{\zeta})$ being the solution of the following Stokes equation:

$$a_f(\mathcal{N}^v\boldsymbol{\zeta}, \mathbf{v}) - b(\mathcal{N}^p\boldsymbol{\zeta}, \mathbf{v}) + (\mathcal{N}^v\boldsymbol{\zeta}, \mathbf{v}) = (\boldsymbol{\zeta}, \mathbf{v})_\Sigma \qquad \forall \mathbf{v} \in H^1(\Omega)^d \tag{3.7.3a}$$

$$b(q, \mathcal{N}^v\boldsymbol{\zeta}) = 0 \qquad \forall q \in L^2(\Omega). \tag{3.7.3b}$$

Therefore,

$$-\nabla \cdot \sigma(\mathcal{N}^v\boldsymbol{\zeta}, \mathcal{N}^p\boldsymbol{\zeta}) + \mathcal{N}^v\boldsymbol{\zeta} = 0 \text{ in } \Omega \quad \text{and} \quad \sigma(\mathcal{N}^v\boldsymbol{\zeta}, \mathcal{N}^p\boldsymbol{\zeta})\mathbf{n} = \boldsymbol{\zeta} \text{ on } \Sigma.$$

Let $\boldsymbol{\xi} = \boldsymbol{\sigma}(\phi, q)\mathbf{n}$. Then it is easy to see that problem (3.7.1) can be equivalently formulated as follows: Find $\boldsymbol{\xi}(t) \in H^1(\Sigma)^d$ for $t \in [0, T]$ satisfying the following evolution equation:

$$-\mathcal{L}_s\mathcal{N}\boldsymbol{\xi} + \mathcal{N}\boldsymbol{\xi} + \partial_t\boldsymbol{\xi} = \mathbf{f} \text{ on } \Sigma \times (0, T], \text{ with initial condition } \boldsymbol{\xi}(0) = 0. \tag{3.7.4}$$

By choosing $\mathbf{v} = \mathcal{N}^v\boldsymbol{\varphi}$ in (3.7.3) and using relation $b(\mathcal{N}^p\boldsymbol{\zeta}, \mathcal{N}^v\boldsymbol{\varphi}) = 0$ (due to the definition of $\mathcal{N}^v\boldsymbol{\varphi}$), we obtain

$$(\boldsymbol{\zeta}, \mathcal{N}\boldsymbol{\varphi})_\Sigma = a_f(\mathcal{N}^v\boldsymbol{\zeta}, \mathcal{N}^v\boldsymbol{\varphi}) + (\mathcal{N}^v\boldsymbol{\zeta}, \mathcal{N}^v\boldsymbol{\varphi}) \quad \forall \boldsymbol{\zeta}, \boldsymbol{\varphi} \in H^{-1/2}(\Sigma)^d. \tag{3.7.5}$$

Especially, this implies that

$$(\boldsymbol{\zeta}, \mathcal{N}\boldsymbol{\zeta})_\Sigma = 2\mu\|\mathcal{N}^v\boldsymbol{\zeta}\|_f^2 + \|\mathcal{N}^v\boldsymbol{\zeta}\|^2 \sim \|\mathcal{N}^v\boldsymbol{\zeta}\|_{H^1}^2 \sim \|\mathcal{N}\boldsymbol{\zeta}\|_{H^{1/2}(\Sigma)}^2 \quad \forall \boldsymbol{\zeta} \in H^{-1/2}(\Sigma)^d. \tag{3.7.6}$$

By choosing $k = s$ in the regularity result in (3.2.7) with $s \ge -1/2, s \in \mathbb{R}$ and noting the trace inequality, we can establish the following mapping property of the Neumann-to-Dirichlet operator:

$$\|\mathcal{N}\boldsymbol{\zeta}\|_{H^{s+1}(\Sigma)} \le C\|\mathcal{N}^v\boldsymbol{\zeta}\|_{H^{s+3/2}(\Omega)} \le C\|\boldsymbol{\zeta}\|_{H^s(\Sigma)} \quad \forall s \ge -1/2, s \in \mathbb{R}, \tag{3.7.7}$$

Note that $\mathcal{N}\boldsymbol{\zeta} = 0$ if and only if $\boldsymbol{\zeta} = \lambda\mathbf{n}$ for some scalar constant $\lambda \in \mathbb{R}$. This motivates us to define the following subspace of $H^s(\Sigma)^d$ for $s \in \mathbb{R}$:

$$\widetilde{H}^s(\Sigma)^d := \{\boldsymbol{\zeta} \in H^s(\Sigma)^d : (\boldsymbol{\zeta}, \mathbf{n})_\Sigma = 0\}.$$

Then we define the Dirichlet-to-Neumann operator $\mathcal{D} : \widetilde{H}^{1/2}(\Sigma)^d \to \widetilde{H}^{-1/2}(\Sigma)^d$ as follows: For $\boldsymbol{\zeta} \in \widetilde{H}^{1/2}(\Sigma)^d$, let $(\mathcal{D}^v\boldsymbol{\zeta}, \mathcal{D}^p\boldsymbol{\zeta})$ be the weak solution of

$$\begin{aligned} -\nabla \cdot \boldsymbol{\sigma}(\mathcal{D}^v\boldsymbol{\zeta}, \mathcal{D}^p\boldsymbol{\zeta}) + \mathcal{D}^v\boldsymbol{\zeta} &= 0 \quad \text{in } \Omega \\ \nabla \cdot \mathcal{D}^v\boldsymbol{\zeta} &= 0 \quad \text{in } \Omega \\ (\mathcal{D}^v\boldsymbol{\zeta})|_\Sigma &= \boldsymbol{\zeta} \quad \text{on } \Sigma, \end{aligned} \tag{3.7.8}$$

and then define $\mathcal{D}\boldsymbol{\zeta} \in \widetilde{H}^{-1/2}(\Sigma)^d$ by the following equation

$$a_f(\mathcal{D}^v\boldsymbol{\zeta}, \mathbf{v}) - b(\mathcal{D}^p\boldsymbol{\zeta}, \mathbf{v}) + (\mathcal{D}^v\boldsymbol{\zeta}, \mathbf{v}) = (\mathcal{D}\boldsymbol{\zeta}, \mathbf{v})_\Sigma \quad \forall \mathbf{v} \in H^1(\Omega)^d. \tag{3.7.9}$$

Since the function $\mathcal{D}^p\boldsymbol{\zeta}$ in equation (3.7.8) is only determined up to a constant, we can choose this constant in such a way that the function $\mathcal{D}\boldsymbol{\zeta}$ defined by (3.7.9) lies in $\widetilde{H}^{-1/2}(\Sigma)^d$. Using trace theorem and Bogovoski's map (cf. [53, Corollary 1.5]) there exists $\mathbf{v} \in H^1(\Omega)^d$ such that $\mathbf{v}|_\Sigma = \mathbf{n}$, $\nabla \cdot \mathbf{v} = \frac{\|\mathbf{n}\|_\Sigma^2}{|\Omega|}$ with $\|\mathbf{v}\|_{H^1} \le C$, testing (3.7.9) with such $\mathbf{v}$, noting the assumption that $(\mathcal{D}\boldsymbol{\zeta}, \mathbf{n})_\Sigma = 0$, we obtain

$$|\overline{\mathcal{D}^p\boldsymbol{\zeta}}| \le C\|\mathcal{D}^v\boldsymbol{\zeta}\|_{H^1}, \tag{3.7.10}$$

where $\overline{\mathcal{D}^p\boldsymbol{\zeta}}$ is the mean value of $\mathcal{D}^p\boldsymbol{\zeta}$ over Ω. Therefore, choosing $k = s$ with $s \ge 1/2, s \in \mathbb{R}$ in (3.2.8) and combining (3.7.10) leads to the following estimates

$$\|\mathcal{D}^v\boldsymbol{\zeta}\|_{H^{s+1/2}} + \|\mathcal{D}^p\boldsymbol{\zeta}\|_{H^{s-1/2}} \le C\|\boldsymbol{\zeta}\|_{H^s(\Sigma)} \quad \forall s \ge 1/2, s \in \mathbb{R}. \tag{3.7.11}$$

From the weak form (3.7.9), it follows that

$$\|\mathcal{D}\boldsymbol{\zeta}\|_{H^{-1/2}(\Sigma)} \le C\left(\|\mathcal{D}^v\boldsymbol{\zeta}\|_{H^1} + \|\mathcal{D}^p\boldsymbol{\zeta}\|\right). \tag{3.7.12}$$

Meanwhile when $s \ge 3/2$, by trace inequality we have

$$\|\mathcal{D}\boldsymbol{\zeta}\|_{H^{s-1}(\Sigma)} \le C\left(\|\mathcal{D}^v\boldsymbol{\zeta}\|_{H^{s+1/2}} + \|\mathcal{D}^p\boldsymbol{\zeta}\|_{s-1/2}\right) \quad \forall s \ge 3/2, s \in \mathbb{R}. \tag{3.7.13}$$

Combining (3.7.12), (3.7.13) and (3.7.11) leads to the following estimates of the Neumann value $\mathcal{D}\boldsymbol{\zeta}$ in terms of the Dirichlet value $\boldsymbol{\zeta}$:

$$\begin{aligned} \|\mathcal{D}\boldsymbol{\zeta}\|_{H^{-1/2}(\Sigma)} &\le C\|\boldsymbol{\zeta}\|_{H^{1/2}(\Sigma)} \quad \forall \boldsymbol{\zeta} \in \widetilde{H}^{1/2}(\Sigma)^d, \\ \|\mathcal{D}\boldsymbol{\zeta}\|_{H^{s-1}(\Sigma)} &\le C\|\boldsymbol{\zeta}\|_{H^s(\Sigma)} \quad \forall \boldsymbol{\zeta} \in \widetilde{H}^s(\Sigma)^d. \quad (\text{whenever } s \ge 3/2, s \in \mathbb{R}) \end{aligned} \tag{3.7.14}$$

The following complex interpolation of Sobolev spaces hold:

$$[H^k(\Sigma)^d, H^s(\Sigma)^d]_\theta = H^{\theta s + (1-\theta)k}(\Sigma)^d \quad \forall k, s \in \mathbb{R}, \theta \in [0,1]; \tag{3.7.15a}$$

$$[\widetilde{H}^k(\Sigma)^d, \widetilde{H}^s(\Sigma)^d]_\theta = \widetilde{H}^{\theta s + (1-\theta)k}(\Sigma)^d \quad \forall k, s \in \mathbb{R}, \theta \in [0,1]; \tag{3.7.15b}$$

where (3.7.15a) follows from [131, Proposition 3.1-3.2 of Chapter 4] and (3.7.15b) follows from (3.7.15a) because $\widetilde{H}^s(\Sigma)^d$ is a retract of $H^s(\Sigma)^d$ for $s \in \mathbb{R}$ via projection $\pi : H^s(\Sigma)^d \to \widetilde{H}^s(\Sigma)^d$, with

$$\pi(\boldsymbol{\zeta}) := \boldsymbol{\zeta} - \frac{(\boldsymbol{\zeta}, \mathbf{n})_\Sigma}{\|\mathbf{n}\|_\Sigma^2}\mathbf{n}. \tag{3.7.16}$$

Therefore, the following result follows from the complex interpolation between the two estimates in (3.7.14):

$$\|\mathcal{D}\boldsymbol{\zeta}\|_{H^{s-1}(\Sigma)} \le C\|\boldsymbol{\zeta}\|_{H^s(\Sigma)} \quad \forall \boldsymbol{\zeta} \in \widetilde{H}^s(\Sigma)^d \quad \forall s \ge 1/2, s \in \mathbb{R}. \tag{3.7.17}$$

If we restrict the domain of $\mathcal{N}$ to $\widetilde{H}^{-1/2}(\Sigma)^d$, then $\mathcal{N} : \widetilde{H}^{-1/2}(\Sigma)^d \to \widetilde{H}^{1/2}(\Sigma)^d$ and $\mathcal{D} : \widetilde{H}^{1/2}(\Sigma)^d \to \widetilde{H}^{-1/2}(\Sigma)^d$ are inverse maps of each other. This leads to the following norm equivalence:

$$\|\boldsymbol{\zeta}\|_{H^{-1/2}(\Sigma)} \sim \|\mathcal{N}\boldsymbol{\zeta}\|_{H^{1/2}(\Sigma)} \quad \forall \boldsymbol{\zeta} \in \widetilde{H}^{-1/2}(\Sigma)^d.$$

Similarly, from identity $\mathcal{D}\mathcal{N}\boldsymbol{\zeta} = \mathcal{N}\mathcal{D}\boldsymbol{\zeta} = \boldsymbol{\zeta}$ for $\boldsymbol{\zeta} \in \widetilde{H}^{1/2}(\Sigma)^d$ and the mapping property in (3.7.17) and (3.7.7), we conclude that the maps $\mathcal{N} : \widetilde{H}^s(\Sigma)^d \to \widetilde{H}^{s+1}(\Sigma)^d$ and $\mathcal{D} : \widetilde{H}^{s+1}(\Sigma)^d \to \widetilde{H}^s(\Sigma)^d$ are also inverse to each other for all $s \geq -1/2, s \in \mathbb{R}$. This implies the following norm equivalence for $s \geq -1/2, s \in \mathbb{R}$:

$$\|\boldsymbol{\zeta}\|_{H^s(\Sigma)} \sim \|\mathcal{N}\boldsymbol{\zeta}\|_{H^{s+1}(\Sigma)}; \quad \|\boldsymbol{\zeta}\|_{H^{s+1}(\Sigma)} \sim \|\mathcal{D}\boldsymbol{\zeta}\|_{H^s(\Sigma)} \quad \forall \boldsymbol{\zeta} \in \widetilde{H}^{-1/2}(\Sigma)^d \tag{3.7.18}$$

To facilitate further use, we summarize the properties of the NtD (Neumann to Dirichlet) operator and DtN (Dirichlet to Neumann) operator in the following lemma:

Lemma 3.7.2.

1. *For $s \geq -1/2, s \in \mathbb{R}$, the NtD operator $\mathcal{N} : \widetilde{H}^s(\Sigma)^d \to \widetilde{H}^{s+1}(\Sigma)^d$ and DtN operator $\mathcal{D} : \widetilde{H}^{s+1}(\Sigma)^d \to \widetilde{H}^s(\Sigma)^d$ are bounded and inverse to each other.*

2. *With domain* $\mathrm{dom}(\mathcal{D}) := \widetilde{H}^1(\Sigma)^d \subseteq \widetilde{L}^2(\Sigma)^d$*, the DtN operator $\mathcal{D}$ is a self-adjoint positive-definite operator on $\widetilde{L}^2(\Sigma)^d$. The NtD operator $\mathcal{N} : \widetilde{L}^2(\Sigma)^d \to \widetilde{L}^2(\Sigma)^d$ is a compact self-adjoint positive-definite operator on $\widetilde{L}^2(\Sigma)^d$.*

3. *The square root operators $\mathcal{D}^{1/2}$ and $\mathcal{N}^{1/2}$ are well defined. Moreover, for $s \geq -1/2, s \in \mathbb{R}$, operators $\mathcal{N}^{1/2} : \widetilde{H}^s(\Sigma)^d \to \widetilde{H}^{s+1/2}(\Sigma)^d$ and $\mathcal{D}^{1/2} : \widetilde{H}^{s+1/2}(\Sigma)^d \to \widetilde{H}^s(\Sigma)^d$ are bounded and inverse to each other.*

Proof. The three statements are proved as follows.

1. The first statement has been proved in (3.7.18).

2. From (3.7.5) and (3.7.6) it follows that $\mathcal{N}$ is self-adjoint positive-definite operator on $\widetilde{L}^2(\Sigma)^d$. Since $\widetilde{H}^1(\Sigma)^d \to \widetilde{L}^2(\Sigma)^d$ is a compact embedding by Rellich-Kondrachov theorem (cf. [131, Proposition 3.4 of Chapter 4]), from mapping property (3.7.7) of $\mathcal{N}$ it follows that $\mathcal{N}$ is a compact operator. To verify $\mathcal{D} : \mathrm{dom}(\mathcal{D}) \to \widetilde{L}^2(\Sigma)^d$ is self-adjoint, it suffices to show that if $\boldsymbol{\zeta}$ satisfies

$$|(\boldsymbol{\zeta}, \mathcal{D}\boldsymbol{\varphi})_\Sigma| \leq C\|\boldsymbol{\varphi}\|_\Sigma \quad \forall \boldsymbol{\varphi} \in \widetilde{H}^1(\Sigma)^d, \tag{3.7.19}$$

then $\boldsymbol{\zeta} \in \widetilde{H}^1(\Sigma)^d$. From (3.7.19), by Riesz representation theorem there exists $\mathbf{g} \in \widetilde{L}^2(\Sigma)^d$ such that

$$(\boldsymbol{\zeta}, \mathcal{D}\boldsymbol{\varphi})_\Sigma = (\mathbf{g}, \boldsymbol{\varphi})_\Sigma \quad \forall \boldsymbol{\varphi} \in \widetilde{H}^1(\Sigma)^d.$$

Especially, taking $\boldsymbol{\varphi} = \mathcal{N}\boldsymbol{\xi}$, it follows that

$$(\boldsymbol{\zeta}, \boldsymbol{\xi})_\Sigma = (\mathbf{g}, \mathcal{N}\boldsymbol{\xi})_\Sigma = (\mathcal{N}\mathbf{g}, \boldsymbol{\xi})_\Sigma \quad \forall \boldsymbol{\xi} \in \widetilde{L}^2(\Sigma)^d.$$

Therefore $\boldsymbol{\zeta} = \mathcal{N}\mathbf{g} \in \widetilde{H}^1(\Sigma)^d$, proof of the second statement is complete.

3. By the spectrum theory of compact self-adjoint operator (cf. [22, Theorem 5.3.16]), $\widetilde{L}^2(\Sigma)^d$ admits an orthonormal basis of eigenvectors $\{\boldsymbol{\omega}_i\}_{i\in\mathbb{N}}$ of $\mathcal{N}$ and $\mathcal{N}$ has the following expression

$$\mathcal{N}\boldsymbol{\zeta} = \sum_{i=1}^{\infty} (\mathcal{N}\boldsymbol{\zeta}, \boldsymbol{\omega}_i)_\Sigma \boldsymbol{\omega}_i = \sum_{i=1}^{\infty} \lambda_i (\boldsymbol{\zeta}, \boldsymbol{\omega}_i)_\Sigma \boldsymbol{\omega}_i \quad \forall \boldsymbol{\zeta} \in \widetilde{H}^{-1/2}(\Sigma)^d,$$

where $\lambda_i > 0$ is the eigenvalue associated with ω_i. From norm equivalence (3.7.18), we can deduce that for $s \in \mathbb{N}$ there holds

$$\|\zeta\|_{H^s(\Sigma)} \sim \|\mathcal{D}^s \zeta\|_\Sigma = \left(\sum_{i=1}^{\infty} \lambda_i^{-2s} |(\zeta, \omega_i)_\Sigma|^2 \right)^{1/2} \qquad \forall s \in \mathbb{N}. \tag{3.7.20}$$

In view of complex interpolation result of weighted ℓ^2-sequence spaces (cf. [15, Theorem 5.5.3]), in fact (3.7.20) is valid for all $s \geq 0, s \in \mathbb{R}$ by complex interpolation method. Moreover for $-1/2 \leq s < 0$, using norm equivalence (3.7.18) and (3.7.20) (which is valid for $s \geq 0, s \in \mathbb{R}$) we have

$$\|\zeta\|_{H^s(\Sigma)} \sim \|\mathcal{N}\zeta\|_{H^{s+1}(\Sigma)} \sim \left(\sum_{i=1}^{\infty} \lambda_i^{-2s} |(\zeta, \omega_i)_\Sigma|^2 \right)^{1/2} \qquad \forall s \in \mathbb{R}, -1/2 \leq s < 0. \tag{3.7.21}$$

Combining (3.7.20) and (3.7.21), we arrive at

$$\|\zeta\|_{H^s(\Sigma)} \sim \left(\sum_{i=1}^{\infty} \lambda_i^{-2s} |(\zeta, \omega_i)_\Sigma|^2 \right)^{1/2} \qquad \forall s \geq -1/2, s \in \mathbb{R}. \tag{3.7.22}$$

We can define square root operators $\mathcal{N}^{1/2}$ and $\mathcal{D}^{1/2}$ by formula

$$\mathcal{N}^{1/2}\zeta := \sum_{i=1}^{\infty} \lambda_i^{1/2} (\zeta, \omega_i)_\Sigma \omega_i \qquad \forall \zeta \in \widetilde{H}^{-1/2}(\Sigma)^d \quad \text{(this series converges in } \widetilde{L}^2(\Sigma)^d\text{)}$$

$$\mathcal{D}^{1/2}\zeta := \sum_{i=1}^{\infty} \lambda_i^{-1/2} (\zeta, \omega_i)_\Sigma \omega_i \qquad \forall \zeta \in \widetilde{L}^2(\Sigma)^d \quad \text{(this series converges in } \widetilde{H}^{-1/2}(\Sigma)^d\text{)},$$

from the norm equivalence in (3.7.22), it is direct to verify that operators $\mathcal{N}^{1/2}$ and $\mathcal{D}^{1/2}$ are inverse to each other and satisfy the following mapping property for $s \geq -1/2, s \in \mathbb{R}$

$$\|\zeta\|_{H^s(\Sigma)} \sim \|\mathcal{N}^{1/2}\zeta\|_{H^{s+1/2}(\Sigma)}; \quad \|\zeta\|_{H^{s+1/2}(\Sigma)} \sim \|\mathcal{D}^{1/2}\zeta\|_{H^s(\Sigma)} \quad \forall \zeta \in \widetilde{H}^{-1/2}(\Sigma)^d. \tag{3.7.23}$$

The proof of third statement is complete.

∎

Part 2. Taking into account of the fact that $\mathcal{N}$ is not injective on $L^2(\Sigma)^d$, for convenience of our further construction we first take the L^2-orthogonal projection $\pi : L^2(\Sigma)^d \to \widetilde{L}^2(\Sigma)^d$ defined as in (3.7.16) on the both side of (3.7.4), and obtain the following equation with solution space contained in $\widetilde{L}^2(\Sigma)^d$: seek $\widetilde{\boldsymbol{\xi}} \in L^2\widetilde{H}^1(\Sigma)^d$ with $\partial_t\widetilde{\boldsymbol{\xi}} \in L^2\widetilde{L}^2(\Sigma)^d$ satisfying

$$\partial_t\widetilde{\boldsymbol{\xi}} + \mathcal{A}\widetilde{\boldsymbol{\xi}} = \widetilde{\mathbf{f}}; \quad \widetilde{\boldsymbol{\xi}}(0) = 0, \tag{3.7.24}$$

where $\mathcal{A} = \pi(I - \mathcal{L}_s)\mathcal{N}$ and $\widetilde{\mathbf{f}} = \pi\mathbf{f}$. One difficulty in proving existence and regularity of solution to (3.7.24) is that the operator $\mathcal{A} : \widetilde{H}^1(\Sigma)^d \to \widetilde{L}^2(\Sigma)^d$ is not a self-adjoint operator in $\widetilde{L}^2(\Sigma)^d$. To overcome this difficulty, we consider the following change of variable $\omega = \mathcal{N}^{1/2}\widetilde{\boldsymbol{\xi}}$, and reformulate (3.7.24) into an abstract Cauchy problem on ω: seek $\omega \in L^2\widetilde{H}^1(\Sigma)^d$ with $\partial_t\omega \in L^2\widetilde{L}^2(\Sigma)^d$ satisfying

$$\partial_t\omega + \mathcal{B}\omega = \mathbf{g}; \quad \omega(0) = 0, \tag{3.7.25}$$

where $\mathcal{B} := \mathcal{N}^{1/2}\pi(I - \mathcal{L}_s)\mathcal{N}^{1/2}$ and $\mathbf{g} := \mathcal{N}^{1/2}\widetilde{\mathbf{f}}$. We summarize some useful properties on the operators $\mathcal{A}$ and $\mathcal{B}$ in the following lemma:

Lemma 3.7.3.

1. *There holds norm equivalence for* $0 \le s \le 1, s \in \mathbb{R}$

$$\|\mathcal{A}\boldsymbol{\zeta}\|_{H^s(\Sigma)} \sim \|\boldsymbol{\zeta}\|_{H^{s+1}(\Sigma)}; \quad \|\mathcal{B}\boldsymbol{\zeta}\|_{H^s(\Sigma)} \sim \|\boldsymbol{\zeta}\|_{H^{s+1}(\Sigma)} \quad \forall \boldsymbol{\zeta} \in \widetilde{H}^{-1/2}(\Sigma)^d. \qquad (3.7.26)$$

2. $\mathcal{B}$ *is a self-adjoint positive-definite operator on* $\widetilde{L}^2(\Sigma)^d$ *with domain* $\mathrm{dom}(\mathcal{B}) := \widetilde{H}^1(\Sigma)^d$.

Proof. The two statements are proved as follows.

1. In view of the norm equivalence relations in (3.7.18) and (3.7.23), it suffices to show the following norm equivalence for $-1 \le s \le 1, s \in \mathbb{R}$

$$\|\pi(I - \mathcal{L}_s)\boldsymbol{\zeta}\|_{H^s(\Sigma)} \sim \|\boldsymbol{\zeta}\|_{H^{s+2}(\Sigma)} \quad \forall \boldsymbol{\zeta} \in \widetilde{H}^{-1/2}(\Sigma)^d. \qquad (3.7.27)$$

Note that one direction of the norm equivalence in (3.7.27) is given by assumption (3.2.3). To prove the opposite direction, observe first that

$$\|\pi(I - \mathcal{L}_s)\boldsymbol{\zeta}\|_{H^{-1}(\Sigma)}\|\boldsymbol{\zeta}\|_{H^1(\Sigma)} \ge (\pi(I - \mathcal{L}_s)\boldsymbol{\zeta}, \boldsymbol{\zeta})_\Sigma = ((I - \mathcal{L}_s)\boldsymbol{\zeta}, \boldsymbol{\zeta})_\Sigma \ge C\|\boldsymbol{\zeta}\|^2_{H^1(\Sigma)}.$$

It follows that (3.7.27) is valid for $s = -1$. Next we note that, by definition (3.7.16) of projection $\pi : H^s(\Sigma)^d \to \widetilde{H}^s(\Sigma)^d$,

$$\|\pi(I - \mathcal{L}_s)\boldsymbol{\zeta} - (I - \mathcal{L}_s)\boldsymbol{\zeta}\|_{H^s(\Sigma)} \le C|(\boldsymbol{\zeta}, (I - \mathcal{L}_s)\mathbf{n})_\Sigma| \le C\|\boldsymbol{\zeta}\|_{H^1(\Sigma)} \qquad (3.7.28)$$

For $-1 \le s \le 1, s \in \mathbb{R}$, in view of regularity assumption (3.2.9) and the estimate (3.7.28) above, we have

$$\begin{aligned}\|\boldsymbol{\zeta}\|_{H^{s+2}(\Sigma)} &\le C\|(I - \mathcal{L}_s)\boldsymbol{\zeta}\|_{H^s(\Sigma)}\\ &\le C\|\pi(I - \mathcal{L}_s)\boldsymbol{\zeta}\|_{H^s(\Sigma)} + C\|\boldsymbol{\zeta}\|_{H^1(\Sigma)}\\ &\le C\|\pi(I - \mathcal{L}_s)\boldsymbol{\zeta}\|_{H^s(\Sigma)} \quad \text{(this is because (3.7.27) is valid for } s = -1).\end{aligned}$$

Thus (3.7.27) is proved and the first statement follows directly.

2. Since $\mathcal{B}$ is obviously symmetric and positive definite on its domain $\mathrm{dom}(B) = \widetilde{H}^1(\Sigma)^d$. To prove that $\mathcal{B}$ is self-adjoint, it remains to show that the domain of the dual operator $\mathcal{B}'$ defined by

$$\mathrm{dom}(\mathcal{B}') = \{\mathbf{w} \in \widetilde{L}^2(\Sigma)^d : \exists\ \mathbf{g} \in \widetilde{L}^2(\Sigma)^d \text{ such that } (\boldsymbol{\omega}, \mathcal{B}\boldsymbol{\zeta})_\Sigma = (\mathbf{g}, \boldsymbol{\zeta})_\Sigma \quad \forall \boldsymbol{\zeta} \in \widetilde{H}^1(\Sigma)^d\},$$

coincides with the domain of $\mathcal{B}$. Therefore, we need to prove that if $\boldsymbol{\omega} \in \widetilde{L}^2(\Sigma)^d$ satisfies

$$(\boldsymbol{\omega}, \mathcal{B}\boldsymbol{\zeta})_\Sigma = (\mathbf{g}, \boldsymbol{\zeta})_\Sigma \quad \forall \boldsymbol{\zeta} \in \widetilde{H}^1(\Sigma)^d, \qquad (3.7.29)$$

for some $\mathbf{g} \in \widetilde{L}^2(\Sigma)^d$, then $\boldsymbol{\omega} \in \widetilde{H}^1(\Sigma)^d$. To this end, we define $\boldsymbol{\varphi} \in \widetilde{H}^1(\Sigma)^d$ to be the weak solution of equation

$$a_s(\boldsymbol{\varphi}, \boldsymbol{\zeta}) + (\boldsymbol{\varphi}, \boldsymbol{\zeta})_\Sigma = (\mathcal{D}^{1/2}\mathbf{g}, \boldsymbol{\zeta})_\Sigma \quad \forall \boldsymbol{\zeta} \in \widetilde{H}^1(\Sigma)^d, \qquad (3.7.30)$$

where the existence and uniqueness of solution to (3.7.30) is due to coercive property: $\|\zeta\|_s^2 + \|\zeta\|_\Sigma^2 \sim \|\zeta\|_{H^1(\Sigma)}^2, \forall \zeta \in \widetilde{H}^1(\Sigma)^d$. Equation (3.7.30) means

$$\pi(I - \mathcal{L}_s)\varphi = \mathcal{D}^{1/2}\mathbf{g} \in \widetilde{H}^{-1/2}(\Sigma)^d,$$

thus by norm equivalence (3.7.27) we have $\varphi \in \widetilde{H}^{3/2}(\Sigma)^d$. Now we observe

$$(\mathcal{D}^{1/2}\varphi, \mathcal{B}\zeta)_\Sigma = \left(\pi(I - \mathcal{L}_s)\varphi, \mathcal{N}^{1/2}\zeta\right)_\Sigma = (\mathcal{D}^{1/2}\mathbf{g}, \mathcal{N}^{1/2}\zeta)_\Sigma = (\mathbf{g}, \zeta)_\Sigma \quad \forall \zeta \in \widetilde{H}^1(\Sigma)^d \tag{3.7.31}$$

By comparing (3.7.29) with (3.7.31) we obtain $\omega = \mathcal{D}^{1/2}\varphi \in \widetilde{H}^1(\Sigma)^d$. This completes the proof.

∎

Especially, since $\mathcal{B}$ is a self-adjoint positive-definite operator on $\widetilde{L}^2(\Sigma)^d$ with domain $\mathrm{dom}(\mathcal{B}) := \widetilde{H}^1(\Sigma)^d$, $-\mathcal{B}$ generates an analytic semigroup $E(t) : \widetilde{L}^2(\Sigma)^d \to \widetilde{L}^2(\Sigma)^d$ for $t \geq 0$ (cf. [22, Example 7.4.5]), and the unique solution to (3.7.25) is given by

$$\omega(t) = \int_0^t E(t-s)\mathbf{g}(s)ds.$$

Moreover, for self-adjoint semigroup on a Hilbert space, the following L^2-maximal regularity estimate holds (cf. [22, Theorem 7.6.11]):

$$\|\partial_t \omega\|_{L^2L^2(\Sigma)} + \|\mathcal{B}\omega\|_{L^2L^2(\Sigma)} \leq C\|\mathbf{g}\|_{L^2L^2(\Sigma)}, \tag{3.7.32}$$

which can be obtained by testing (3.7.25) with $\partial_t\omega$. If the source term $\mathbf{g}$ in (3.7.25) possesses higher spacial regularity, the solution ω also inherits higher spacial regularity. To see this, assume $\mathbf{g} \in L^2\widetilde{H}^1(\Sigma)^d$, then since

$$\mathcal{B}\omega(t) = \int_0^t E(t-s)\mathcal{B}\mathbf{g}(s)ds$$

is the solution to (3.7.25) with the source term replaced by $\mathcal{B}\mathbf{g}$. Thus again by maximal L^2-regularity estimate, we have

$$\|\mathcal{B}\partial_t \omega\|_{L^2L^2(\Sigma)} + \|\mathcal{B}^2\omega\|_{L^2L^2(\Sigma)} \leq C\|\mathcal{B}\mathbf{g}\|_{L^2L^2(\Sigma)}. \tag{3.7.33}$$

By norm equivalence in (3.7.26), it follows that

$$\|\partial_t \omega\|_{L^2H^1(\Sigma)} + \|\mathcal{B}\omega\|_{L^2H^1(\Sigma)} \leq C\|\mathbf{g}\|_{L^2H^1(\Sigma)}. \tag{3.7.34}$$

Complex interpolation of (3.7.32) and 3.7.34 gives

$$\|\partial_t \omega\|_{L^2H^{1/2}(\Sigma)} + \|\mathcal{B}\omega\|_{L^2H^{1/2}(\Sigma)} \leq C\|\mathbf{g}\|_{L^2H^{1/2}(\Sigma)}. \tag{3.7.35}$$

Now we take $\mathbf{g} = \mathcal{N}^{1/2}\widetilde{\mathbf{f}}$, then it is direct to verify that $\widetilde{\boldsymbol{\xi}} := \mathcal{D}^{1/2}\omega$ is the solution to (3.7.24) and satisfies estimate

$$\|\partial_t\widetilde{\boldsymbol{\xi}}\|_{L^2L^2(\Sigma)} + \|\widetilde{\boldsymbol{\xi}}\|_{L^2H^1(\Sigma)} \leq C\|\widetilde{\mathbf{f}}\|_{L^2L^2(\Sigma)}, \tag{3.7.36}$$

where we have used norm equivalences in (3.7.26) and (3.7.23). Having obtained the solution $\widetilde{\boldsymbol{\xi}}$ to equation (3.7.24), if we write $\mathbf{f}(t) = \widetilde{\mathbf{f}}(t) + c(t)\mathbf{n}$ then it is direct to verify that $\boldsymbol{\xi}(t) = \widetilde{\boldsymbol{\xi}}(t) + k(t)\mathbf{n}$ is the solution to (3.7.4), where $k(t)$ is given by

$$\partial_t k = c - r(\widetilde{\boldsymbol{\xi}}); \quad k(0) = 0$$
$$r(\widetilde{\boldsymbol{\xi}}) := \frac{\big((I - \mathcal{L}_s)\mathcal{N}\widetilde{\boldsymbol{\xi}}, \mathbf{n}\big)_\Sigma}{\|\mathbf{n}\|_\Sigma^2},$$

it follows that

$$\|\partial_t k\|_{L^2(0,T)} \le C(\|\mathbf{f}\|_{L^2L^2(\Sigma)} + \|\widetilde{\boldsymbol{\xi}}\|_{L^2H^1(\Sigma)}) \le C\|\mathbf{f}\|_{L^2L^2(\Sigma)}. \tag{3.7.37}$$

Therefore, combining (3.7.37) and (3.7.36) we obtain

$$\|\partial_t \boldsymbol{\xi}\|_{L^2L^2(\Sigma)} + \|\boldsymbol{\xi}\|_{L^2H^1(\Sigma)} \le C\|\mathbf{f}\|_{L^2L^2(\Sigma)} \tag{3.7.38}$$

Part 3. Given the solution $\boldsymbol{\xi}$ to equation (3.7.4), we define $(\boldsymbol{\phi}, q) = (\mathcal{N}^v\boldsymbol{\xi}, \mathcal{N}^p\boldsymbol{\xi})$. Then $\boldsymbol{\xi} = \boldsymbol{\sigma}(\boldsymbol{\phi}, q)\mathbf{n}$ and $\mathcal{N}\boldsymbol{\xi} = \boldsymbol{\phi}|_\Sigma$. Therefore, equation (3.7.4) can be written as

$$-\mathcal{L}_s\boldsymbol{\phi} + \boldsymbol{\phi} = -\partial_t\boldsymbol{\sigma}(\boldsymbol{\phi}, q)\mathbf{n} + \mathbf{f} \quad \text{on } \Sigma \times (0, T], \quad \text{with initial condition } (\boldsymbol{\phi}, q)(0) = (0, 0).$$

Thus $(\boldsymbol{\phi}, q)$ is a solution of equation (3.7.1). Since $\boldsymbol{\sigma}(\boldsymbol{\phi}, q)\mathbf{n}(0) = 0$, it follows that

$$\|\boldsymbol{\sigma}(\boldsymbol{\phi}, q)\mathbf{n}\|_{L^\infty L^2(\Sigma)} \le C\|\partial_t\boldsymbol{\sigma}(\boldsymbol{\phi}, q)\mathbf{n}\|_{L^2L^2(\Sigma)}.$$

Therefore, $(\boldsymbol{\phi}, q)$ satsifies the following estimate according to (3.7.38) and the inequality above:

$$\|\boldsymbol{\sigma}(\boldsymbol{\phi}, q)\mathbf{n}\|_{L^\infty L^2(\Sigma)} + \|\boldsymbol{\phi}\|_{L^2H^2(\Sigma)} \le C\|\mathbf{f}\|_{L^2L^2(\Sigma)}. \tag{3.7.39}$$

Moreover, since (ϕ, q) is the solution of the homogeneous Stokes equation (with boundary value $\phi|_\Sigma = \mathcal{N}\boldsymbol{\xi}$), the following two estimates follow from the regularity results of the Stokes equation in (3.2.7)–(3.2.8):

$$\|\boldsymbol{\phi}\|_{H^2} + \|\nabla q\| \le C\|\boldsymbol{\phi}|_\Sigma\|_{H^2(\Sigma)}; \quad \|\boldsymbol{\phi}\|_{H^1} + \|q\| \le C\|\boldsymbol{\sigma}(\boldsymbol{\phi}, q)\mathbf{n}\|_{H^{-1/2}(\Sigma)} \tag{3.7.40}$$

Combining the estimates in (3.7.39) and (3.7.40), we obtain the result of Proposition 3.7.1. ∎

Proof of (3.4.75). Let $\widehat{\boldsymbol{\phi}} = \boldsymbol{\phi}(T-t)$ and $\widehat{q} = q(T-t)$, then $\widehat{\boldsymbol{\phi}}$ and $\widehat{q}$ is a solution of (3.7.1) in Proposition 3.7.1, with source term $\widehat{\mathbf{f}}(t) = \mathbf{f}(T-t)$, thus we have

$$\|\boldsymbol{\phi}\|_{L^2H^2(\Omega)} + \|\boldsymbol{\phi}\|_{L^2H^2(\Sigma)} + \|q\|_{L^2H^1(\Omega)} + \|\boldsymbol{\sigma}(\boldsymbol{\phi}, q)\mathbf{n}\|_{L^\infty L^2(\Sigma)} \le C\|\mathbf{f}\|_{L^2L^2(\Sigma)}.$$

The proof of Lemma 3.4.5 is complete. ∎

3.8 Appendix B: Proof of (3.4.47)

In this subsection, we assume that $r \ge 2$. Under this assumption, we establish a negative-norm estimate for the Dirichlet Stokes–Ritz projection R_h^D in the following lemma.

Lemma 3.8.1. *For the Dirichlet Stokes–Ritz projection R_h^D defined in* (3.4.35)*, the following error estimate holds:*

$$\|\mathbf{u} - R_h^D \mathbf{u}\|_{H^{-1}} + \|\mathbf{u} - R_h^D \mathbf{u}\|_{H^{-1}(\Sigma)} + h\|p - R_h^D p\|_{H^{-1}} \leqslant Ch^{r+2}. \tag{3.8.1}$$

Skectch of Proof. From the definition of $R_h^D \mathbf{u}$ in (3.4.35) we can see that the following relation holds on the boundary Σ:

$$R_h^S \mathbf{u} - R_h^D \mathbf{u} = \frac{(R_h^S \mathbf{u}, \mathbf{n})_\Sigma}{\|\mathbf{n}_h\|^2} \mathbf{n}_h = \frac{(R_h^S \mathbf{u} - \mathbf{u}, \mathbf{n})_\Sigma}{\|\mathbf{n}_h\|^2} \mathbf{n}_h.$$

Since

$$\frac{(R_h^S \mathbf{u} - \mathbf{u}, \mathbf{n})_\Sigma}{\|\mathbf{n}_h\|^2} \lesssim C\|R_h^S \mathbf{u} - \mathbf{u}\|_{H^{-1}(\Sigma)} \leq Ch^{r+2},$$

it follows that $\|\mathbf{u} - R_h^D \mathbf{u}\|_{H^{-1}(\Sigma)} \leq Ch^{r+2}$. Then (3.8.1) follows from the same routine of duality argument for the Dirichlet Stokes–Ritz projection. ∎

Next, we note that $\|(R_{sh}\mathbf{u} - \mathbf{u})(0)\|$ also satisfies negative norm estimate below.

Lemma 3.8.2. *For the $R_{sh}\mathbf{u}(0)$ defined in* (3.4.43)*, the following negative-norm estimate holds:*

$$\|(R_{sh}\mathbf{u} - \mathbf{u})(0)\|_{H^{-1}(\Sigma)} \leq Ch^{r+2}. \tag{3.8.2}$$

Proof. We introduce a dual equation

$$-\mathcal{L}_s \psi + \psi = \varphi \quad \psi \text{ has periodic boundary condition on } \Sigma. \tag{3.8.3}$$

The regularity assumption in (3.2.9) implies that $\|\psi\|_{H^3(\Sigma)} \leq \|\varphi\|_{H^1(\Sigma)}$. We can extend ψ to be a function (still denoted by ψ) which is defined in Ω with periodic boundary condition and satisfies $\|\psi\|_{H^3} \leq C\|\psi\|_{H^3(\Sigma)}$. Then the following relation can be derived:

$$\begin{aligned}
((R_{sh}\mathbf{u} - \mathbf{u})(0), \varphi)_\Sigma =& a_s((R_{sh}\mathbf{u} - \mathbf{u})(0), \psi) + ((R_{sh}\mathbf{u} - \mathbf{u})(0), \psi)_\Sigma \\
=& a_s((R_{sh}\mathbf{u} - \mathbf{u})(0), \psi - I_h\psi)_\Sigma + ((R_{sh}\mathbf{u} - \mathbf{u})(0), (\psi - I_h\psi))_\Sigma \\
& - a_f((R_h^D \partial_t \mathbf{u} - \partial_t \mathbf{u})(0), I_h\psi) + b((R_h^D \partial_t p - \partial_t p)(0), I_h\psi) \\
& - ((R_h^D \partial_t \mathbf{u} - \partial_t \mathbf{u})(0), I_h\psi) \\
\leq& Ch^{r+2}\|\psi\|_{H^3(\Sigma)} + |a_f((R_h^D \partial_t \mathbf{u} - \partial_t \mathbf{u})(0), \psi)| + |b((R_h^D \partial_t p \partial_t - p)(0), \psi)| \\
& + |((R_h^D \partial_t \mathbf{u} - \partial_t \mathbf{u})(0), \psi)|.
\end{aligned}$$

Since

$$\begin{aligned}
(\mathbf{D}(R_h^D \partial_t \mathbf{u} - \partial_t \mathbf{u})(0), \mathbf{D}\psi) &= -((R_h^D \partial_t \mathbf{u} - \partial_t \mathbf{u})(0), \nabla \cdot \mathbf{D}\psi) + ((R_h^D \partial_t \mathbf{u} - \partial_t \mathbf{u})(0), \mathbf{n} \cdot \mathbf{D}\psi)_\Sigma \\
&\lesssim \left(\|(R_h^D \partial_t \mathbf{u} - \partial_t \mathbf{u})(0)\|_{H^{-1}(\Omega)} + \|(R_h^D \partial_t \mathbf{u} - \partial_t \mathbf{u})(0)\|_{H^{-1}(\Sigma)}\right) \|\psi\|_{H^3(\Omega)} \\
&\leq Ch^{r+2}\|\varphi\|_{H^1(\Sigma)}
\end{aligned}$$

and

$$\begin{aligned}
b((R_h^D \partial_t p - \partial_t p)(0), \psi) &\lesssim C\|(R_h^D \partial_t p - \partial_t p)(0)\|_{H^{-2}}\|\psi\|_{H^3} \leq Ch^{r+2}\|\varphi\|_{H^1(\Sigma)} \\
((R_h^D \partial_t \mathbf{u} - \partial_t \mathbf{u})(0), \psi) &\lesssim \|(R_h^D \partial_t \mathbf{u} - \partial_t \mathbf{u})(0)\|_{H^{-1}}\|\psi\|_{H^1} \leq Ch^{r+2}\|\varphi\|_{H^1(\Sigma)},
\end{aligned}$$

summing up the estimates above yields the result in (3.8.2). The proof is complete. ∎

Lemma 3.8.3. *Let* $e_h^u := (R_h\mathbf{u} - R_h^D\mathbf{u})(0)$ *and* $e_h^p := (R_h p - R_h^D p)(0)$*. Then the following estimates hold:*

$$\|e_h^u\|_{H^1} + \|e_h^p\| \le Ch^{r+1/2}, \tag{3.8.4}$$
$$\|e_h^u\|_{H^{-1}(\Sigma)} + \|e_h^u\|_{H^{-1/2}} + \|e_h^p\|_{H^{-3/2}} \le Ch^{r+2}. \tag{3.8.5}$$

Proof. To prove the first inequality in Lemma 3.8.3, we note that

$$a_f(e_h^u, \mathbf{v}_h) + (e_h^u, \mathbf{v}_h) - b(e_h^p, \mathbf{v}_h) = 0 \quad \forall \mathbf{v}_h \in \mathring{\mathbf{X}}_h^r. \tag{3.8.6}$$

Let $\mathbf{u}_h = E_h(e_h^u|_\Sigma)$, where E_h is an extension operator as in (3.2.17). Then $e_h^u - \mathbf{u}_h \in \mathring{\mathbf{X}}_h^r$ and $\|\mathbf{u}_h\|_{H^1} \le Ch^{-1/2}\|e_h^u\|_\Sigma \le Ch^{r+1/2}$. This estimate of $\|\mathbf{u}_h\|_{H^1}$ and relation (3.8.6) imply that

$$a_f(e_h^u - \mathbf{u}_h, \mathbf{v}_h) + (e_h^u - \mathbf{u}_h, \mathbf{v}_h) - b(e_h^p, \mathbf{v}_h) \lesssim Ch^{r+1/2}\|\mathbf{v}_h\|_{H^1} \quad \forall \mathbf{v}_h \in \mathring{\mathbf{X}}_h^r.$$

Now, choosing $\mathbf{v}_h = e_h^u - \mathbf{u}_h$ in the inequality above, we obtain (3.8.4)

$$\|e_h^u\|_{H^1(\Omega)} + \|e_h^p\| \le Ch^{r+1/2}.$$

We prove the second inequality in Lemma 3.8.3 now. On the boundary Σ, relations (3.4.44b) and (3.4.39) imply that $(R_h\mathbf{u} - \mathbf{u})(0) = (R_{sh}\mathbf{u} - \mathbf{u})(0) - \lambda(R_{sh}\mathbf{u}(0))\mathbf{n}_h$. Since

$$\lambda(R_{sh}\mathbf{u}(0)) = \frac{(R_{sh}\mathbf{u}(0), \mathbf{n})_\Sigma}{\|\mathbf{n}_h\|_\Sigma^2} = \frac{(R_{sh}\mathbf{u}(0) - \mathbf{u}(0), \mathbf{n})_\Sigma}{\|\mathbf{n}_h\|_\Sigma^2} \lesssim C\|R_{sh}\mathbf{u}(0) - \mathbf{u}(0)\|_{H^{-1}(\Sigma)},$$

it follows from (3.8.2) in Lemma 3.8.2 that $\|(R_h\mathbf{u} - \mathbf{u})(0)\|_{H^{-1}(\Sigma)} \le Ch^{r+2}$. Using inequality $\|(R_h^D\mathbf{u} - \mathbf{u})(0)\|_{H^{-1}(\Sigma)} \le Ch^{r+2}$ in (3.8.1) of Lemma 3.8.1, we obtain

$$\|(R_h\mathbf{u} - R_h^D\mathbf{u})(0)\|_{H^{-1}(\Sigma)} \le Ch^{r+2}. \tag{3.8.7}$$

Next, we consider a dual problem: For given $\mathbf{f} \in H^{1/2}(\Omega)^d$, we construct $(\boldsymbol{\phi}, q)$ to be the solution of

$$-\nabla \cdot \sigma(\boldsymbol{\phi}, q) + \boldsymbol{\phi} = \mathbf{f}; \quad \nabla \cdot \boldsymbol{\phi} = 0; \quad \boldsymbol{\phi}|_\Sigma = 0; \; q \in L_0^2(\Omega).$$

By the regularity assumptions in (3.2.8), the following estimate of ϕ and q can be written down:

$$\|\boldsymbol{\phi}\|_{H^{5/2}} + \|p\|_{H^{3/2}} \le C\|\mathbf{f}\|_{H^{1/2}}.$$

From equation (3.8.6) one can see that

$$\begin{aligned}(e_h^u, \mathbf{f}) =& a_f(e_h^u, \boldsymbol{\phi} - I_h\boldsymbol{\phi}) + (e_h^u, \boldsymbol{\phi} - I_h\boldsymbol{\phi}) - (\boldsymbol{\sigma}(\boldsymbol{\phi}, q)\mathbf{n}, e_h^u)_\Sigma \\ &+ b(e_h^p, I_h\boldsymbol{\phi} - \boldsymbol{\phi}) \\ \lesssim& Ch^{r+2}\|\mathbf{f}\|_{H^{1/2}} + \|e_h^u\|_{H_p^{-1}(\Sigma)}\|\mathbf{f}\|_{H^{1/2}} \le Ch^{r+2}\|\mathbf{f}\|_{H^{1/2}}.\end{aligned}$$

Therefore, the following result is proved:

$$|(e_h^u, \mathbf{f})| \le Ch^{r+2}\|\mathbf{f}\|_{H^{1/2}}; \quad \|e_h^u\|_{H^{-1/2}} \le Ch^{r+2}.$$

We move on to consider the dual problem for pressure: For given $f \in H^{3/2}(\Omega)$, since $e_h^p \in L_0^2(\Omega)$ it follows that

$$(e_h^p, f) = (e_h^p, f - \overline{f}).$$

Thus it suffices to assume that $\int_\Omega f = 0$. Then, using Bogovoski's map (see details in [53, Corollary 1.5] and [4, Theorem 4]), there exists $\mathbf{v} \in H^{5/2}(\Omega)$ such that

$$\nabla \cdot \mathbf{v} = f, \quad \|\mathbf{v}\|_{H^{5/2}} \le C\|f\|_{H^{3/2}}, \quad \mathbf{v}|_\Sigma = 0.$$

From equation (3.8.6), we can find that

$$(e_h^p, f) = b(e_h^p, \mathbf{v}) = b(e_h^p, \mathbf{v} - I_h\mathbf{v}) + a_f(e_h^u, I_h\mathbf{v} - \mathbf{v}) + (e_h^u, I_h\mathbf{v} - \mathbf{v}) + a_f(e_h^u, \mathbf{v}) + (e_h^u, \mathbf{v}).$$

Thus, combining the known estimate $\|e_h^u\|_{H^{-1/2}} \le Ch^{r+2}$ and $\|e_h^u\|_{H^1} + \|e_h^p\| \le Ch^r$, we have

$$|(e_h^p, f)| \le Ch^{r+2}\|f\|_{H^{3/2}} + |a_f(e_h^u, \mathbf{v})|.$$

Using integration by parts, we derive that

$$a_f(e_h^u, \mathbf{v}) = 2\mu(\mathbf{D}e_h^u, \mathbf{D}\mathbf{v}) = -2\mu(e_h^u, \nabla \cdot \mathbf{D}\mathbf{v}) + 2\mu(e_h^u, \mathbf{D}\mathbf{v} \cdot \mathbf{n})_\Sigma \lesssim Ch^{r+2}\|f\|_{H^{3/2}}.$$

Therefore, we have proved the following result:

$$|(e_h^p, f)| \le Ch^{r+2}\|f\|_{H^{3/2}}; \quad \|e_h^p\|_{H^{-3/2}} \le Ch^{r+2}.$$

This completes the proof of Lemma 3.8.3. ∎

Lemma 3.8.4. *For the projection operators R_h and R_{sh} defined in* (3.4.3) *and* (3.4.4)*, respectively, the following estimate holds:*

$$\|R_{sh}\boldsymbol{\eta}(0) - R_h\boldsymbol{\eta}(0)\|_{H^1(\Sigma)} \le Ch^{r+1}.$$

Proof. By denoting $\delta_h := R_h\boldsymbol{\eta}(0) - R_{\mathrm{sh}}\boldsymbol{\eta}(0)$, $e_h^u := (R_h\mathbf{u} - R_h^D\mathbf{u})(0)$ and $e_h^p := (R_hp - R_h^Dp)(0)$, we can write down the following equation according to the definitions of the two projection operators:

$$a_s(\delta_h, \mathbf{v}_h) + (\delta_h, \mathbf{v}_h)_\Sigma + a_f(e_h^u, \mathbf{v}_h) - b(e_h^p, \mathbf{v}_h) + (e_h^u, \mathbf{v}_h) = 0 \quad \forall \mathbf{v}_h \in \mathbf{X}_h^r.$$

Then, choosing $\mathbf{v}_h = E_h\delta_h \in \mathbf{X}_h^r$ in the relation above and note that $\|\mathbf{v}_h\|_{H^1} \le Ch^{-1/2}\|\delta_h\|_\Sigma$, we derive that

$$\|\delta_h\|_{H^1(\Sigma)}^2 \le Ch^{-1/2}\|\delta_h\|_\Sigma(\|e_h^u\|_{H^1} + \|e_h^p\|) \le Ch^r\|\delta_h\|_\Sigma. \tag{3.8.8}$$

Next, we consider the following dual problem: Let ψ be the solution of

$$-\mathcal{L}_s\psi + \psi = \delta_h \quad \psi \text{ has periodic boundary condition on } \Sigma.$$

Then

$$a_s(\psi, \boldsymbol{\xi}) + (\psi, \boldsymbol{\xi})_\Sigma = (\delta_h, \boldsymbol{\xi})_\Sigma \quad \forall \boldsymbol{\xi} \in \mathbf{S} \quad \text{ and } \|\psi\|_{H^2(\Sigma)} \le \|\delta_h\|_\Sigma.$$

We can extend ψ to a function (still denoted by ψ) defined on Ω with periodic boundary condition and $\|\psi\|_{H^{5/2}} \le C\|\psi\|_{H^2}$. Therefore

$$\begin{aligned}\|\delta_h\|_\Sigma^2 =& a_s(\delta_h, \psi - I_h\psi)_\Sigma + (\delta_h, (\psi - I_h\psi))_\Sigma \\ & - a_f(e_h^u, I_h\psi) + b(e_h^p, I_h\psi) - (e_h^u, I_h\psi) \\ \le& Ch\|\delta_h\|_{H^1(\Sigma)}\|\psi\|_{H^2(\Sigma)} + Ch^{3/2}(\|e_h^u\|_{H^1} + \|e_h^p\|)\|\psi\|_{H^{5/2}} \\ & + |a_f(e_h^u, \psi) - b(e_h^p, \psi) + (e_h^u, \psi)|.\end{aligned} \tag{3.8.9}$$

Using integrating by parts and negative-norm estimates in Lemma 3.8.3, the following estimates can obtain:

$$b(e_h^p, \psi) \lesssim C\|\psi\|_{H^{5/2}}\|e_h^p\|_{H^{-3/2}} \le Ch^{r+2}\|\delta_h\|_\Sigma, \tag{3.8.10}$$

$$a_f(e_h^u, \psi) + (e_h^u, \psi) \lesssim C\|e_h^u\|_{H^{-1/2}}\|\psi\|_{H^{5/2}} + C\|e_h^u\|_{H^{-1}(\Sigma)}\|\psi\|_{H^{5/2}} \le Ch^{r+2}\|\delta_h\|_\Sigma. \tag{3.8.11}$$

Therefore, combining the estimates in (3.8.8)–(3.8.11), we obtain the following error estimate for δ_h:

$$\|\delta_h\|_\Sigma \le Ch^{r/2+1}\|\delta_h\|_\Sigma^{1/2} + Ch^{r+2} \Rightarrow \|\delta_h\|_\Sigma \le Ch^{r+2}.$$

The inverse inequality implies $\|\delta_h\|_{H^1(\Sigma)} \le Ch^{r+1}$. This completes the proof of Lemma 3.8.4. ∎

3.9 Appendix C: Proof of (3.2.16)

In this appendix, we prove (3.2.16) in the following lemma.

Lemma 3.9.1. *Under assumptions* **(A1)**–**(A4)** *on the finite element spaces, the following type of inf-sup condition holds (the $H^1(\Sigma)$-norm is involved on the right-hand side of the inequality)*:

$$\|p_h\| \le C \sup_{0\ne \mathbf{v}_h \in \mathbf{X}_h^r} \frac{(\mathrm{div}\mathbf{v}_h, p_h)}{\|\mathbf{v}_h\|_{H^1} + \|\mathbf{v}\|_{H^1(\Sigma)}} \quad \forall p_h \in Q_h^{r-1},$$

where $C > 0$ is a constant independent of p_h and the mesh size h.

Proof. Each $p_h \in Q_h^{r-1}$ can be decomposed into $p_h = \widetilde{p}_h + \bar{p}_h$, with $\widetilde{p}_h \in Q_{h,0}^{r-1}$ and $\bar{p}_h = \frac{1}{|\Omega|}\int_\Omega p_h \mathrm{d}x$. Since we have assumed that inf-sup condition (3.2.14) holds, there exists $\widetilde{\mathbf{v}}_h \in \mathring{\mathbf{X}}_h^r$ such that

$$\|\widetilde{\mathbf{v}}_h\|_{H^1} \le \|\widetilde{p}_h\| \quad \text{and} \quad b(\widetilde{p}_h, \widetilde{\mathbf{v}}_h) \ge C_1\|\widetilde{p}_h\|^2 \quad \text{for some constant } C_1 > 0. \tag{3.9.1}$$

For the constant $\bar{p}_h \in \mathbb{R}$, we note that

$$b(\bar{p}_h, \mathbf{v}_h) = \bar{p}_h b(1, \mathbf{v}_h) = \bar{p}_h(\mathbf{v}_h, \mathbf{n})_\Sigma.$$

Let $\mathbf{v}_h^* \in \mathbf{X}_h^r$ be defined as $\mathbf{v}_h^* = E_h(\mathbf{n}_h)$, where $\mathbf{n}_h$ is defined in (3.4.38) and E_h is the extension operator defined in item 4 of Remark 3.2.1, i.e., $\mathbf{v}_h^* = I_h^X\mathbf{v} \in \mathbf{X}_h^r$ with $\mathbf{v} \in H^1(\Omega)^d$ being an extension of $\mathbf{n}_h$ such that $\mathbf{v}|_\Sigma = \mathbf{n}_h$. By the definition of $\mathbf{v}_h^*$, we have

$$\|\mathbf{v}_h^*\|_{H^1} = \|I_h^X\mathbf{v}\|_{H^1} \le C\|\mathbf{v}\|_{H^1} \le C\|\mathbf{n}_h\|_{H^{1/2}(\Sigma)} \le C$$
$$\|\mathbf{v}_h^*\|_{H^1(\Sigma)} = \|\mathbf{n}_h\|_{H^1(\Sigma)} \le C.$$

Moreover, the following relation holds:

$$b(1, \mathbf{v}_h^*) = (\mathbf{v}_h^*, \mathbf{n})_\Sigma = (\mathbf{n}_h, \mathbf{n})_\Sigma = \|\mathbf{n}_h\|_\Sigma^2 \ge C > 0.$$

Therefore, the function $\mathbf{v}_h^1 := \bar{p}_h \mathbf{v}_h^*$ has the following property:

$$\|\mathbf{v}_h^1\|_{H^1} + \|\mathbf{v}_h^1\|_{H^1(\Sigma)} \le C|\bar{p}_h| \le C_0\|\bar{p}_h\| \quad \text{for some constant } C_0 > 0.$$

We can re-scale $\mathbf{v}_h^1$ to $\mathbf{v}_h^e = \frac{1}{C_0}\mathbf{v}_h^1$ so that the following inequalities hold for some constant $C_2 > 0$:

$$\|\mathbf{v}_h^e\|_{H^1} + \|\mathbf{v}_h^e\|_{H^1(\Sigma)} \le \|\bar{p}_h\| \quad \text{and} \quad b(\bar{p}_h, \mathbf{v}_h^e) = \frac{|\bar{p}_h|^2}{C_0} b(1, \mathbf{v}_h^*) \ge C_2\|\bar{p}_h\|^2. \tag{3.9.2}$$

By considering $\mathbf{v}_h = \widetilde{\mathbf{v}}_h + \epsilon \mathbf{v}_h^e$, with a parameter $\epsilon > 0$ to be determined later, and using the relation $b(\bar{p}_h, \widetilde{\mathbf{v}}_h) = \bar{p}_h(\widetilde{\mathbf{v}}_h, \mathbf{n}_h)_\Sigma = 0$, we have

$$\begin{aligned}
b(p_h, \mathbf{v}_h) &= b(\widetilde{p}_h + \bar{p}_h, \widetilde{\mathbf{v}}_h + \epsilon \mathbf{v}_h^e) \\
&= b(\widetilde{p}_h, \widetilde{\mathbf{v}}_h) + \epsilon b(\widetilde{p}_h, \mathbf{v}_h^e) + \epsilon b(\bar{p}_h, \mathbf{v}_h^e) \\
&\ge C_1\|\widetilde{p}_h\|^2 + \epsilon b(\widetilde{p}_h, \mathbf{v}_h^e) + \epsilon C_2\|\bar{p}_h\|^2 \qquad \text{(here (3.9.1) and (3.9.2) are used)} \\
&\ge C_1\|\widetilde{p}_h\|^2 - C\epsilon\|\widetilde{p}_h\|\|\bar{p}_h\| + \epsilon C_2\|\bar{p}_h\|^2 \quad \text{(the first inequality in (3.9.2) is used).}
\end{aligned}$$

By using Young's inequality, we can reduce the last inequality to the following one:

$$b(p_h, \mathbf{v}_h) \ge C_1\|\widetilde{p}_h\|^2 + \epsilon C_2\|\bar{p}_h\|^2 - \left(\frac{C_1}{2}\|\widetilde{p}_h\|^2 + \frac{C^2\epsilon^2}{2C_1}\|\bar{p}_h\|^2\right).$$

Then, choosing $\epsilon = \frac{C_1 C_2}{C^2}$, we derive that

$$b(p_h, \mathbf{v}_h) \ge \frac{C_1}{2}\|\widetilde{p}_h\|^2 + \frac{C_1 C_2^2}{2C^2}\|\bar{p}_h\|^2 \ge C_3\|p_h\|^2 \quad \text{with } C_3 := \min\{\tfrac{C_1}{2}, \tfrac{C_1 C_2^2}{2C^2}\}. \tag{3.9.3}$$

Since $\mathbf{v}_h = \widetilde{\mathbf{v}}_h + \epsilon \mathbf{v}_h^e$ with $\widetilde{\mathbf{v}}_h = 0$ on Σ, it follows from the triangle inequality and (3.9.1)–(3.9.2) that

$$\|\mathbf{v}_h\|_{H^1} + \|\mathbf{v}_h\|_{H^1(\Sigma)} \le \|\widetilde{\mathbf{v}}_h\|_{H^1} + \epsilon(\|\mathbf{v}_h^e\|_{H^1} + \|\mathbf{v}_h^e\|_{H^1(\Sigma)}) \tag{3.9.4}$$

$$\le \|\widetilde{p}_h\| + \frac{C_1 C_2}{C^2}\|\bar{p}_h\| \le \Big(1 + \frac{C_1 C_2}{C^2}\Big)\|p_h\|. \tag{3.9.5}$$

Therefore, (3.9.3) and (3.9.4) imply that

$$\frac{b(p_h, \mathbf{v}_h)}{\|\mathbf{v}_h\|_{H^1} + \|\mathbf{v}_h\|_{H^1(\Sigma)}} \geqslant \frac{C_3\|p_h\|^2}{\big(1 + \frac{C_1 C_2}{C^2}\big)\|p_h\|} = \frac{C_3\|p_h\|}{\big(1 + \frac{C_1 C_2}{C^2}\big)}.$$

This proves that

$$\|p_h\| \le \frac{1}{C_3}\Big(1 + \frac{C_1 C_2}{C^2}\Big)\frac{b(p_h, \mathbf{v}_h)}{\|\mathbf{v}_h\|_{H^1} + \|\mathbf{v}_h\|_{H^1(\Sigma)}}.$$

and therefore completes the proof of (3.2.16). ∎

Chapter 4

Weak Discrete Maximum Principle of Isoparametric Finite Element Methods in Curvilinear Polyhedra

4.1 Introduction

Let Ω be a bounded domain in $\mathbb{R}^N$ with $N \in \{2,3\}$ and consider a quasi-uniform triangulation of the domain Ω with mesh size h, denoted by $\mathcal{T}_h$. Hence, $\Omega_h = (\bigcup_{K\in\mathcal{T}_h} K)^\circ$ is an approximation of Ω. Let $S_h(\Omega_h)$ be a finite element space subject to the triangulation $\mathcal{T}_h$, and denote by $S_h^\circ(\Omega_h) = \{v_h \in S_h(\Omega_h) : v_h = 0 \text{ on } \partial\Omega_h\}$ the finite element subspace under the homogeneous boundary condition. A function $u_h \in S_h(\Omega_h)$ is called discrete harmonic if it satisfies

$$\int_{\Omega_h} \nabla u_h \cdot \nabla \chi_h = 0 \quad \forall\, \chi_h \in S_h^\circ(\Omega_h). \tag{4.1.1}$$

For a given mesh and finite element space, if all the discrete harmonic functions satisfy the following inequality:

$$\|u_h\|_{L^\infty(\Omega_h)} \leqslant \|u_h\|_{L^\infty(\partial\Omega_h)}, \tag{4.1.2}$$

then it is said that the *discrete maximum principle* holds.

The discrete maximum principle of finite element methods (FEMs) has attracted much attention from numerical analysts due to its importance for the stability and accuracy of numerical solutions; for example, see [34, 36, 121, 134, 138]. However, strong restrictions on the geometry of the mesh are required for the discrete maximum principle to hold. For example, for piecewise linear finite elements on a two-dimensional triangular mesh, the discrete maximum principle generally requires the angles of the triangles to be less than $\pi/2$; see [138, §5]. In three dimensions, it is hard to have such discrete maximum principe even for piecewise linear finite elements; see [20, 83, 84, 142].

Schatz considered a different approach in [125] by proving the weak maximum principle (also called the Agmon–Miranda maximum principle) ,

$$\|u_h\|_{L^\infty(\Omega_h)} \leqslant C\, \|u_h\|_{L^\infty(\partial\Omega_h)}, \tag{4.1.3}$$

for some constant C which is independent of u_h and h, for a wide class of H^1-conforming finite elements on a general quasi-uniform triangulation of a two-dimensional polygonal domain. It was shown in [125] that the weak maximum principle can be used to prove the maximum-norm stability and best approximation results of FEMs in a plane polygonal domain, i.e.,

$$\|u - R_h u\|_{L^\infty(\Omega)} \leqslant C\ell_h \inf_{v_h \in S_h^\circ} \|u - v_h\|_{L^\infty(\Omega)} \quad \forall\, u \in H_0^1(\Omega) \cap L^\infty(\Omega), \tag{4.1.4}$$

where $R_h : H_0^1(\Omega) \to S_h^\circ$ is the Ritz projection operator, and

$$\ell_h = \begin{cases} \ln(2 + 1/h) & \text{for piecewise linear elements,} \\ 1 & \text{for higher-order finite elements.} \end{cases}$$

Such maximum-norm stability and best approximation results have a number of applications in the error estimates of finite element solutions for parabolic problems [101, 102, 80, 98], Stokes systems [14], nonlinear problems [60, 40, 115], optimal control problems [6, 7], and so on.

In three dimensions, the weak maximum principle was extended to convex polyhedral domains in [96] and used to prove the L^∞-norm stability and best approximation results of FEMs on convex polyhedral domains, removing an extra logarithmic factor $\ln(2 + 1/h)$ in the stability constant for quadratic and higher-order elements obtained in other approaches (for example, see [97]). When Ω is a smooth domain and $\Omega_h = \Omega$ (the triangulation is assumed to match the curved boundary exactly), the weak maximum principle of quadratic or higher-order FEMs is a result of the maximum-norm stability result in [127, 124], and the weak maximum principle of linear finite elements can be proved similarly as in [96]. In all these articles, the triangulation is assumed to match the boundary of the domain exactly, with $\Omega_h = \Omega$.

In the practical computation, the curved boundary of a bounded smooth domain, or more generally a curvilinear polygon or polyhedron (which may contain both curved faces, curved edges, and corners), is generally approximated by isoparametric finite elements instead of being matched exactly by the triangulation. In this case, the weak maximum principle of FEMs has not been proved yet. Correspondingly, the best approximation results such as (4.1.4) are not known for isoparametric FEMs in a curved domain.

Some related results have been proved in the case $\Omega_h \neq \Omega$. For the Poisson equation with Dirichlet boundary conditions in convex smooth domains, the piecewise linear finite element space with a zero extension in $\Omega\backslash\Omega_h$ is conforming, i.e., $S_h(\Omega_h) \subset H_0^1(\Omega)$. In this case, pointwise error estimates of FEMs have been established in [12, 124]. For general bounded smooth domains which may be concave, thus the finite element space may be non-conforming, Kashiwabara & Kemmochi [79] have obtained the following error estimate for piecewise linear finite elements for the Poisson equation under the Neumann boundary condition:

$$\|\tilde{u} - u_h\|_{L^\infty(\Omega_h)} \leqslant Ch|\log h| \inf_{v_h \in S_h} \|\tilde{u} - v_h\|_{W^{1,\infty}(\Omega_h)} + Ch^2|\log h|\|u\|_{W^{2,\infty}(\Omega)}, \tag{4.1.5}$$

where $\tilde{u}$ is any extension of u in $W^{2,\infty}(\Omega_\delta)$ and Ω_δ is a neighborhood of $\overline{\Omega}$. In the case $u \in W^{2,\infty}(\Omega)$, this error estimate is a consequence of the best approximation result in (4.1.4). More recently, the $W^{1,\infty}$ stability of the Ritz projection was proved in [43] for isoparametric FEMs on $C^{r+1,1}$-smooth domains based on weighted-norm estimates, where r denotes the degree of finite elements. For curvilinear polyhedra or smooth domains which may be concave, the weak maximum principle and the best approximation results in the L^∞ norm have not been proved.

In this chapter, we close the gap mentioned above by proving the weak maximum principle in (4.1.3) for isoparametric finite elements of degree $r \geqslant 1$ in a bounded smooth domain or a curvilinear polyhedron (possibly concave) with edge openings smaller than π. As an application of the weak maximum principle, we prove that the finite element solution $u_h \in S_h^{\circ}(\Omega_h)$ of the Poisson equation

$$\begin{cases} -\Delta u = f & \text{in } \Omega \\ \quad\;\; u = 0 & \text{on } \partial\Omega \end{cases} \tag{4.1.6}$$

using isoparametric finite elements of degree $r \geqslant 1$ has the following optimal-order error bound (for any $p > N$):

$$\|u - u_h\|_{L^\infty(\Omega)} \leqslant C\ell_h \|u - \check{I}_h u\|_{L^\infty(\Omega)} + Ch^{r+1}\ell_h \|f\|_{L^p(\Omega)}, \tag{4.1.7}$$

where u_h is extended to be zero in $\Omega\backslash\Omega_h$, and $\check{I}_h u$ denotes a Lagrange interpolation operator (which will be defined in the next section). Inequality (4.1.7) can be viewed as a variant of the best approximation result in (4.1.4) by taking account of the geometry change of the domain, which produces an additional optimal-order term $Ch^{r+1}\|f\|_{L^p(\Omega)}$ independent of the higher regularity of f. In particular, inequality (4.1.7) implies the following error estimate:

$$\|u - u_h\|_{L^\infty(\Omega)} \leqslant C\ell_h h^s \|u\|_{C^s(\overline{\Omega})} + Ch^{r+1}\ell_h \|f\|_{L^p(\Omega)} \quad \text{for } u \in C^s(\overline{\Omega}),\;\; 0 \leqslant s \leqslant r+1, \tag{4.1.8}$$

which adapts to the regularity of u.

The weak maximum principle is proved by converting the finite element weak form on Ω_h to a weak form on Ω by using a bijective transformation $\Phi_h : \Omega_h \to \Omega$ which is piecewisely defined on the triangles/tetrahedra. This yields a bilinear form with a discontinuous coefficient matrix. The main technical difficulty is that the elliptic partial differential equation associated to this coefficient matrix does not have the H^2 regularity estimate, which is required in the proof of weak maximum principle in the literature; see [96]. We overcome this difficulty by decomposing the solution v (of a duality problem) into two parts, $v = v_1 + v_2$, with v_1 corresponding to the Poisson equation with H^2 regularity, and v_2 corresponding to an elliptic equation with discontinuous coefficients but with a small source term arising from the geometry perturbation, and then estimate the two parts separately by using the H^2 and $W^{1,p}$ regularity of the respective problems.

The maximum-norm error estimate is proved by using Schatz argument through estimating the difference between the solutions of the Poisson equations in Ω_h and Ω. However, in order to avoid using the partial derivatives of f in the proof of (4.1.7), we have to estimate the error between the solutions of the Poisson equation in the two domains Ω_h and Ω under the Dirichlet boundary conditions, respectively. This is accomplished by perturbing the curvilinear polyhedron through a globally smooth flow map pointing outward the domain and establishing the $W^{1,\infty}$ regularity estimate of the Poisson equation in a slightly larger perturbed domain Ω^t (uniformly with respect to the perturbation), which contains both Ω_h and Ω and satisfies that $\mathrm{dist}(x, \partial\Omega) \sim h^{r+1}$ for $x \in \partial\Omega^t$.

The rest of this chapter is organized as follows. In Section 4.2, we present the main results to be proved in this chapter, including the weak maximum principle of the isoparametric FEM in a curvilinear polyhedron, and the best approximation result of finite element solutions in the maximum norm. The proofs of the two main results are presented in Sections 4.3 and 4.4, respectively. The conclusions are presented in Section 4.5.

4.2 Main results

In this chapter, we assume that $\Omega \subset \mathbb{R}^N$, with $N \in \{2, 3\}$, is either a bounded smooth domain or a curvilinear polyhedron (possibly concave) with edge openings smaller than π. More specifically, in the three-dimensional space, this means that for every $x \in \partial\Omega$ there is a neighborhood U_x and a smooth diffeomorphism $\varphi_x : U_x \to B_0(\varepsilon_x)$ mapping x to 0 such that one of the following three conditions holds:

1. x is a smooth point, i.e., $\varphi_x(U_x \cap \Omega) = B_0(\varepsilon_x) \cap \mathbb{R}^3_+$, where $\mathbb{R}^3_+ = \{x \in \mathbb{R}^3 : x_3 > 0\}$ is a half space in $\mathbb{R}^3$.
2. x is an edge point, i.e., $\varphi_x(U_x \cap \Omega) = B_0(\varepsilon_x) \cap K_x$, where $K_x = \mathbb{R} \times \Sigma$, where $\Sigma \subseteq \mathbb{R}^2$ is a sector with angle less than π.
3. x is a vertex point, i.e., $\varphi_x(U_x \cap \Omega) = B_0(\varepsilon_x) \cap K_x$, where K_x is a convex polyhedral cone with a vertex at 0. Therefore, the boundary of K_x consists of several smooth faces intersecting at some edges which pass through the vertex 0.

We refer to [9, Definition 2.1] for the definition of general curvilinear polyhedron.

Let $\mathcal{T}_h$ be the set of closed simplices in a quasi-uniform triangulation of the domain Ω with isoparametric finite elements of degree $r \geqslant 1$ approximating the boundary $\partial\Omega$, as described in [94], with flat interior simplices which have at most one vertex on $\partial\Omega$ and possibly curved boundary simplices. Each boundary simplex contains a possibly curved face or edge interpolating $\partial\Omega$ with an accuracy of $O(h^{r+1})$, where h denotes the mesh size of the triangulation. Hence, $\Omega_h = (\bigcup_{K\in\mathcal{T}} K)^\circ$ is an approximation to Ω such that $\mathrm{dist}(x, \partial\Omega) = O(h^{r+1})$ for $x \in \partial\Omega_h$.

We prove the following weak maximum principle of the isoparametric FEM.

Theorem 4.2.1. *For the isoparametric FEM of degree $r \geqslant 1$ on a quasi-uniform triangulation of Ω, all the discrete harmonic functions $u_h \in S_h(\Omega_h)$ satisfying* (4.1.1) *have the following estimate:*

$$\|u_h\|_{L^\infty(\Omega_h)} \leqslant C \|u_h\|_{L^\infty(\partial\Omega_h)}, \tag{4.2.1}$$

where the constant C is independent of u_h and the mesh size h.

In the isoparametric finite elements described in [94], each curved simplex $K \in \mathcal{T}_h$ is the image of a map $F_K : \hat{K} \to K$ defined on the reference simplex $\hat{K}$, which is a polynomial of degree no larger than r and transforms the finite element structure of $\hat{K}$ to K. There is a homeomorphism $\Phi_h : \Omega_h \to \Omega$, which is piecewise smooth on each simplex and globally Lipschitz continuous. If we denote $\Phi_{h,K} := \Phi_h|_K$ and $\check{K} := \Phi_h(K)$, then $\Phi_{h,K} : K \to \check{K}$ is a diffeomorphism which transforms the finite element structure of K to $\check{K}$. Therefore, $\check{\mathcal{T}}_h = \{\check{K} : K \in \mathcal{T}\}$ is a triangulation of the curved domain Ω, with

$$\Omega_h = \bigcup_{K\in\mathcal{T}_h} K \quad \text{and} \quad \Omega = \bigcup_{K\in\mathcal{T}_h} \check{K}.$$

One can define isoparametric finite element space $S_h(\Omega_h)$ as

$$S_h(\Omega_h) = \{v_h \in H^1(\Omega_h) : v_h|_K \circ F_K \text{ is a polynomial on } \hat{K} \text{ of degree} \leqslant r \text{ for } K \in \mathscr{K}\}. \tag{4.2.2}$$

The finite element spaces on Ω can be defined as

$$\check{S}_h(\Omega) = \{\check{v}_h \in H^1(\Omega) : \check{v}_h \circ \Phi_h \in S_h(\Omega_h)\} \quad \text{and} \quad \check{S}_h^\circ(\Omega) = \{\check{v}_h \in \check{S}_h(\Omega) : \check{v}_h = 0 \text{ on } \partial\Omega\}. \tag{4.2.3}$$

For a finite element function $v_h \in S_h(\Omega_h)$, we can associate it with a finite element function $\check{v}_h \in \check{S}_h(\Omega)$ defined by $v_h \circ \Phi_h^{-1} := \check{v}_h$.

Remark 4.2.1. *By using the notation which link $v_h \in S_h(\Omega_h)$ and $\check{v}_h \in \check{S}_h(\Omega)$, the weak maximum principle in* (4.2.1) *can be equivalently written as*

$$\|\check{u}_h\|_{L^\infty(\Omega)} \leqslant C \|\check{u}_h\|_{L^\infty(\partial\Omega)} . \tag{4.2.4}$$

For a function $f \in C^0(\overline{\Omega}_h)$, one can define its local interpolation $I_{h,K}f$ on a simplex $K \in \mathcal{T}$ as the function satisfying

$$(I_{h,K}f) \circ F_K := I_{\hat{K}}(f \circ F_K),$$

where $I_{\hat{K}}$ is the standard Lagrange interpolation on the reference simplex $\hat{K}$ (onto the space of polynomials of degree $\leqslant r$). The global interpolation $I_h f \in S_h(\Omega_h)$ is defined as

$$I_h f|_K := I_{h,K} f \quad \forall\, K \in \mathcal{T}.$$

For the analysis of the isoparametric FEM, we also define an interpolation operator $\check{I}_h : C(\overline{\Omega}) \to \check{S}_h(\Omega)$ by

$$(\check{I}_h v) \circ \Phi_h = I_h(v \circ \Phi_h) \quad \forall\, v \in C(\overline{\Omega}).$$

As an application of the weak maximum principle, we establish an L^∞-norm best approximation result of isoparametric FEM for the Poisson equation in a curvilinear polyehdron. We assume that the triangulation can be extended to a bigger domain which contains $\overline{\Omega}$, as stated below.

Assumption 4.2.1. The curvilinear polyhedral domain Ω can be extended to a larger convex polyhedron Ω_* with piecewise flat boundaries such that $\overline{\Omega} \subset \Omega_*$ and the triangulation $\mathcal{T}_h$ can be extended to a quasi-uniform triangulation $\mathcal{T}_{*,h}$ on Ω_* (thus the triangulation in $\Omega_* \backslash \overline{\Omega}$ is also isoparametric on its boundary $\partial\Omega$).

Remark 4.2.2. Here Ω_* can be chosen as a large cube whose interior contains $\overline{\Omega}$. Note that the triangulation $\mathcal{T}_h$ is obtained from some triangulation $\widetilde{\mathcal{T}}_h$ consisting of flat simplexes by the method in Lenoir's paper [94]. We can first extend $\widetilde{K}$ to a quasi-uniform flat triangulation $\widetilde{\mathcal{T}}_{*,h}$ of Ω_*, and then modify those flat simplexes $\widetilde{\mathcal{T}}_h$ with one of whose edges/faces attaches to the boundary $\partial\Omega$, to isoparametric elements by the method in Lenoir's paper [94]. This leads to a quasi-uniform triangulation $\mathcal{T}_{*,h}$ on Ω_* which extends $\mathcal{T}_h$. By our construction, the triangulation on $\Omega_* \backslash \overline{\Omega}$ is also isoparametric on its boundary $\partial\Omega$.

Theorem 4.2.2. *For $f \in L^p(\Omega)$ with some $p > N$, we consider the Poisson equation*

$$\begin{cases} -\Delta u = f & \text{in } \Omega \\ \quad u = 0 & \text{on } \partial\Omega \end{cases} \tag{4.2.5}$$

and the isoparametric FEM of degree $r \geqslant 1$ for (4.2.5)*: Find $u_h \in S_h^\circ(\Omega_h)$ such that*

$$\int_{\Omega_h} \nabla u_h \cdot \nabla \chi_h \mathbf{d}x = \int_{\Omega_h} \tilde{f} \chi_h \mathbf{d}x \quad \forall\, \chi_h \in S_h^\circ(\Omega_h), \tag{4.2.6}$$

where $\tilde{f} \in L^p(\Omega \cup \Omega_h)$ is any extension of $f \in L^p(\Omega)$ satisfying $\|\tilde{f}\|_{L^p(\Omega\cup\Omega_h)} \leqslant C\|f\|_{L^p(\Omega)}$. Assuming that the triangulation satisfies Assumption 4.2.1*, there exist positive constants h_0 and C (independent of f, u and h) such that the solutions of* (4.2.5) *and* (4.2.6) *satisfy the following inequality for $h \leq h_0$:*

$$\|u - u_h\|_{L^\infty(\Omega)} \leq C\ell_h \|u - \check{I}_h u\|_{L^\infty(\Omega)} + Ch^{r+1}\ell_h\|f\|_{L^p(\Omega)}, \tag{4.2.7}$$

where u_h is extended to be zero on $\Omega \backslash \Omega_h$, and ℓ_h is defined as

$$\ell_h = \begin{cases} \ln(2 + 1/h) & \text{for piecewise linear elements,} \\ 1 & \text{for higher-order finite elements.} \end{cases}$$

The proofs of Theorems 4.2.1 and 4.2.2 are presented in the next two sections, respectively. For the simplicity of notation, we denote by C a generic positive constant which may be different at different occurrences, possibly depending on the specific domain Ω and the shape-regularity and quasi-uniformity of the triangulation, and the polynomial degree $r \geqslant 1$, but is independent of the mesh size h.

4.3 Proof of Theorem 4.2.1

The proof of Theorem 4.2.1 is divided into six parts, presented in the following six subsections.

4.3.1 Properties of the isomparametric FEM

In this subsection, we summarize the basic properties of the isoparametric FEM to be used in the proof of Theorem 4.2.1.

Lemma 4.3.1 ([94, Theorem 1, Theorem 2, Proposition 2, Proposition 3, Proposition 4])**.** *Let $\check{\mathcal{T}}_h$ be the triangulation of Ω by isoparametric finite elements of degree $r \geqslant 1$, with the maps $F_K : \hat{K} \to K$ and $\Phi_{h,K} : K \to \check{K}$ described in Section* 4.2*. Let D^s denote the Fréchet derivative of order s. Then the following results hold:*

1. *$F_K : \hat{K} \to K$ is a diffeomorphism such that*

$$\begin{aligned} \|D^s F_K\|_{L^\infty(\hat{K})} &\leq Ch^s && \forall s \in [1, r+1] \\ \|D^s F_K^{-1}\|_{L^\infty(K)} &\leq Ch^{-s} && \forall s \in [1, r+1] \end{aligned} \tag{4.3.1}$$

2. *$\Phi_{h,K} : K \to \check{K}$ is a diffeomorphism such that*

$$\begin{aligned} \|D^s(\Phi_{h,K} - \mathrm{Id})\|_{L^\infty(K)} &\leq Ch^{r+1-s} && \forall s \in [1, r+1] \\ \|D^s(\Phi_{h,K}^{-1} - \mathrm{Id})\|_{L^\infty(\check{K})} &\leq Ch^{r+1-s} && \forall s \in [1, r+1] \end{aligned} \tag{4.3.2}$$

3. *For $v \in H^m(K)$ and integer $m \in [0, r+1]$, the norms $\|v\|_{H^m(K)}$ and $\|v \circ \Phi_{h,K}^{-1}\|_{H^m(\check{K})}$ are uniformly equivalent with respect to h.*

4. *Each curved simplex $K \in \mathcal{T}_h$ corresponds to a flat simplex $\widetilde{K}$ (which has the same vertices as K), and there is a unique linear bijection $F_{\widetilde{K}} : \hat{K} \to \widetilde{K}$ which maps the reference simplex $\hat{K}$ onto $\widetilde{K}$. The map $\widetilde{\Psi}_K := F_K \circ F_{\widetilde{K}}^{-1} : \widetilde{K} \to K$ is a diffeomorphism satisfying the following estimates:*

$$\begin{aligned} &\|D(\widetilde{\Psi}_K - \mathrm{Id})\|_{L^\infty(\widetilde{K})} \le Ch, \quad \|D(\widetilde{\Psi}_K^{-1} - \mathrm{Id})\|_{L^\infty(K)} \le Ch \\ &\|D^s\widetilde{\Psi}_K\|_{L^\infty(\widetilde{K})} \le C, \quad \|D^s\widetilde{\Psi}_K^{-1}\|_{L^\infty(K)} \le C \quad \forall s \in [1, r+1]. \end{aligned} \tag{4.3.3}$$

5. *For $v \in H^m(K)$ and integer $m \in [0, r+1]$, the norms $\|v\|_{H^m(K)}$ and $\|v \circ \widetilde{\Psi}_K\|_{H^m(\widetilde{K})}$ are uniformly equivalent with respect to h.*

Let $W_h^{k,p}(\Omega)$ be the space of functions on Ω whose restriction on each $\check{K} \in \check{\mathcal{T}}_h$ lies in $W^{k,p}(\check{K})$, equipped with the following norm:

$$\|v\|_{W_h^{k,p}(\Omega)} := \begin{cases} \left(\sum_{K \in \mathcal{T}} \|v\|^p_{W^{k,p}(K)} \right)^{\frac{1}{p}} & \text{for } 1 \leqslant p < \infty, \\ \sup_{K \in \mathcal{T}} \|v\|_{W^{k,p}(K)} & \text{for } p = \infty. \end{cases}$$

In the case $p = 2$ we write $H^{l,h}(\Omega) = W_h^{l,2}(\Omega)$. The following local interpolation error estimate was proved in [94, Lemma 7]; also see [35, Theorem 4.3.4]. Although it was proved only for $p = 2$ in [94, Lemma 7], the proof can be extended to $1 \leqslant p \leqslant \infty$ straightforwardly.

Lemma 4.3.2 (Lagrange interpolation)**.** *Let $\check{I}_{h,K} : C(\overline{\Omega}) \to \check{S}_h(\Omega)$ be the interpolation operator defined by*

$$\check{I}_{h,K} f \circ \Phi_h := I_{h,K} f \quad \forall f \in C(\overline{\Omega}).$$

Then, for $1 \le k \le r+1$ and $1 \leqslant p \leqslant \infty$ such that $W_h^{k,p}(\Omega) \hookrightarrow C(\overline{\Omega})$ (e.g., $kp > N$ when $p > 1$ or $k \ge N$ when $p = 1$), the following error estimate holds:

$$|u - \check{I}_{h,K} u|_{W^{i,p}(\check{K})} \le Ch^{k-i} \|u\|_{W^{k,p}(\check{K})} \quad \forall\, 0 \le i \le k,\ \forall\, \check{K} \in \check{\mathcal{T}}_h,\ \forall\, u \in C(\bar{\Omega}) \cap W_h^{k,p}(\Omega).$$

Since the Lagrange interpolation is defined by using the pointwise values of a function at the Lagrange nodes, its stability in the $W^{k,p}$ norm is valid only when $W^{k,p}(\Omega) \hookrightarrow C(\overline{\Omega})$, i.e., in the case "$kp > N$ and $p > 1$" or "$k \ge N$ and $p = 1$". One can remove this restriction by using the Scott–Zhang interpolation, which can be constructed first in the flat triangulation $\widetilde{\mathcal{T}}_h = \{\widetilde{K} : K \in \mathcal{T}_h\}$ as in [21, Section 4.8] and then be transformed to $\mathcal{T}_h$ via the maps $\widetilde{\Psi}_K$. Namely, by denoting $\widetilde{\Omega}_h = \bigcup_{K \in \mathcal{T}_h} \widetilde{K}$ and $\widetilde{\Psi}_h : \widetilde{\Omega}_h \to \Omega_h$, we can define

$$(\mathcal{I}^h v) \circ \widetilde{\Psi}_h := \widetilde{\mathcal{I}}^h (v \circ \widetilde{\Psi}_h) \quad \forall\, v \in L^1(\Omega_h),$$

where $\widetilde{\mathcal{I}}^h$ denotes the Scott–Zhang interpolation on the flat triangulation $\widetilde{\mathcal{T}}_h$. Since the maps $\widetilde{\Psi}_h$ induces norm equivalence on every simplex, as a result of (4.3.3), we have the following result.

Lemma 4.3.3 (Scott–Zhang interpolation)**.** *There is a global interpolation operator*

$$\mathcal{I}^h : L^1(\Omega_h) \to S_h(\Omega_h)$$

such that

$$|u - \mathcal{I}^h u|_{W_h^{i,p}(\Omega_h)} \le Ch^{k-i} \|u\|_{W_h^{k,p}(\Omega_h)} \quad \forall\, 0 \le i \le k,\ \forall\, 1 \le k \le r+1,\ \forall\, u \in W_h^{k,p}(\Omega_h).$$

The inverse estimate for isoparametric finite elements follows from Lemma 4.3.1, Part 1. This is presented in the following lemma.

Lemma 4.3.4 (Inverse estimate)**.** *For* $1 \le k \le l \le r+1$ *and* $1 \le p, q \le \infty$ *the following estimate holds:*

$$\|\check{u}_h\|_{W^{l,p}(\check{K})} \le Ch^{k-l+N/p-N/q}\|\check{u}_h\|_{W^{k,q}(\check{K})} \quad \forall\, \check{u}_h \in \check{S}_h(\Omega),\ \forall\, \check{K} \in \check{\mathcal{T}}_h. \tag{4.3.4}$$

The following lemma says that the $(r+1)$th-order derivative of a finite element function in $\check{S}_h(\Omega)$ can be bounded by its lower-order derivatives. This result is often used to prove a super-approximation property which is stated in Lemma 4.3.6 for iso-parametric finite elements.

Lemma 4.3.5. *The following result holds for iso-parametric finite element functions in* $\check{S}_h(\Omega)$:

$$|D^{r+1}\check{v}_h|(x) \le C\sum_{i=1}^{r}|D^i\check{v}_h|(x) \quad \forall\, x \in \check{K},\ \forall\, \check{K} \in \check{\mathcal{T}}_h,\ \forall\, \check{v}_h \in \check{S}_h(\Omega). \tag{4.3.5}$$

Proof. Let $M_K := \Phi_{h,K} \circ \widetilde{\Psi}_K$, which is a diffeomorpshism between the flat simplex $\widetilde{K}$ and the curved simplex $\check{K}$ (according to Lemma 4.3.1), satisfying the following estimates:

$$\|D^sM_K\|_{L^\infty(\widetilde{K})} \le C \quad \text{and} \quad \|D^sM_K^{-1}\|_{L^\infty(\check{K})} \le C \quad \forall\, 1 \le s \le r+1.$$

According to the definition of $\check{S}_h(\Omega)$, a function $\check{v}_h$ is in $\check{S}_h(\Omega)$ if and only if the pull-back function $\check{v}_h \circ M_K$ is a polynomial degree $\leqslant r$ on the flat simplex $\widetilde{K}$. Therefore, from the estimate on higher order derivatives of composed functions (see [36, Lemma 3]), we have

$$\begin{aligned}
&|D^{r+1}\check{v}_h|(x)\\
&= |D^{r+1}((\check{v}_h \circ M_K)\circ M_K^{-1})|(x)\\
&\le C\sum_{l=1}^{r+1}|D^l(\check{v}_h \circ M_K)(M_K^{-1}(x))| \sum_{i\in I(l,r+1)}|DM_K^{-1}(x)|^{i_1}|D^2M_K^{-1}(x)|^{i_2}...|D^{r+1}M_K^{-1}(x)|^{i_{r+1}}\\
&\le C\sum_{l=1}^{r+1}|D^l(\check{v}_h \circ M_K)|(M_K^{-1}(x))\\
&= C\sum_{l=1}^{r}|D^l(\check{v}_h \circ M_K)|(M_K^{-1}(x)),
\end{aligned}$$

where

$$I(l, r+1) := \{i = (i_1, i_2, ..., i_{r+1}) \in \mathbb{Z}^{r+1} : i_k \ge 0, \sum_{k=1}^{r+1} i_k = l; \sum_{k=1}^{r+1} ki_k = r+1\}.$$

We can estimate $|D^l(\check{v}_h \circ M_K)|(M_K^{-1}(x))$ using the same estimate on higher order derivatives of composed function

$$\begin{aligned}
&|D^l(\check{v}_h \circ M_K)|(M_K^{-1}(x))\\
\le& C\sum_{k=1}^{l}|D^k\check{v}_h|(x)\sum_{i\in I(k,l)}|DM_K(M_K^{-1}(x))|^{i_1}|D^2M_K(M_K^{-1}(x))|^{i_2}...|D^lM_K(M_K^{-1}(x))|^{i_l}
\end{aligned}$$

$$\leq C\sum_{k=1}^{l}|D^k\check{v}_h|(x)$$

The result of Lemma 4.3.5 is obtained by combining the two estimates above. ∎ The result above is the key to the superapproximation results for the isoparametric case. For the standard elements the $r+1$ derivative just vanishes.

Lemma 4.3.6 (Super-approximation). *Let $\omega \in C_0^\infty(\mathbb{R}^N)$ be a smooth cut-off function such that $0 \leq \omega \leq 1$ and $\operatorname{supp}(\omega)\cap\Omega \subset \Omega_0 \subset \Omega$, with $\Omega_0(d) := \{x\in\Omega : \operatorname{dist}(x,\Omega_0)\leq d\} \subset \Omega_1$ for some $d > h$. Then the following estimate holds for $\check{v}_h \in \check{S}_h^\circ(\Omega)$:*

$$\|\omega\check{v}_h - \check{I}_h(\omega\check{v}_h)\|_{H^1(\Omega_1)} \leq Ch\Big(\sum_{j=1}^{r} h^{j-1}\|\omega\|_{W^{j,\infty}(\mathbb{R}^N)}\Big)\|\check{v}_h\|_{H^1(\Omega_1)} + Ch^r\|\omega\|_{r+1,\infty}\|\check{v}_h\|_{L^2(\Omega_1)},$$

$$\|\omega\check{v}_h - \check{I}_h(\omega\check{v}_h)\|_{H^1(\Omega_1)} \leq C\Big(\sum_{j=1}^{r+1} h^{j-1}\|\omega\|_{W^{j,\infty}(\mathbb{R}^N)}\Big)\|\check{v}_h\|_{L^2(\Omega_1)}.$$

Proof. Since $\operatorname{supp}(\omega\check{v}_h)\subset\Omega_0$, it follows that $\check{I}_h(\omega\check{v}_h)$ vanishes on all $\check{K}$ such that $\check{K}\cap\Omega_0 = \emptyset$. Since $\Omega_0(d)\subset\Omega_1$, all the simplices $\check{K}$ such that $\check{K}\cap\Omega_0 \neq \emptyset$ are contained in Ω_1. Therefore, we have

$$\begin{aligned}\|\omega\check{v}_h - \check{I}_h(\omega\check{v}_h)\|^2_{H^1(\Omega_1)} &= \sum_{\check{K}\cap\Omega_0\neq\emptyset}\|\omega\check{v}_h - \check{I}_h(\omega\check{v}_h)\|^2_{H^1(\check{K})}\\ &\leqslant \sum_{\check{K}\cap\Omega_0\neq\emptyset} Ch^{2r}\|\omega\check{v}_h\|^2_{H^{r+1}(\check{K})}\\ &\leqslant \sum_{\check{K}\cap\Omega_0\neq\emptyset} Ch^{2r}\Big(|\check{v}_h|^2_{H^{r+1}(\check{K})} + \sum_{i=0}^{r}\|\omega\|^2_{W^{r+1-i,\infty}(\mathbb{R}^N)}\|\check{v}_h\|^2_{H^i(\check{K})}\Big).\end{aligned} \tag{4.3.6}$$

The term $|\check{v}_h|^2_{H^{r+1}(\check{K})}$ can be estimated by using Lemma 4.3.5, i.e.,

$$|\check{v}_h|^2_{H^{r+1}(\check{K})} \leq C\sum_{i=1}^{r}|\check{v}_h|^2_{H^i(\check{K})}. \tag{4.3.7}$$

For $0\leqslant i\leqslant r$, the term $|\check{v}_h|_{H^i(\check{K})}$ can be estimated by using the inverse estimate for isoparametric finite element functions (see Lemma 4.3.4). This yields the first result of Lemma 4.3.6. The second result can be proved similarly. ∎

4.3.2 The perturbed bilinear form associated to the isoparametric FEM

By using the notation $\check{u}_h\circ\Phi_h = u_h$ and $\check{v}_h\circ\Phi_h = v_h$ for $u_h, v_h \in S_h(\Omega_h)$, the following identity holds:

$$\begin{aligned}\int_{\Omega_h}\nabla u_h\cdot\nabla v_h\mathbf{d}x &= \int_{\Omega_h}\nabla(\check{u}_h\circ\Phi_h)\cdot\nabla(\check{u}_h\circ\Phi_h)\mathbf{d}x\\ &= \int_{\Omega}A_h\nabla\check{u}_h\cdot\nabla\check{v}_h\mathbf{d}x \quad \forall\,\check{v}_h\in\check{S}_h^\circ(\Omega),\end{aligned} \tag{4.3.8}$$

where

$$A_h = (\nabla\Phi_h(\nabla\Phi_h)^\top J^{-1}) \circ \Phi_h^{-1}$$

is a piecewise smooth (globally discontinuous) and symmetric matrix-valued function, and $J = \det(\nabla\Phi_h) \in L^\infty(\Omega_h)$ is the Jacobian of the mapping $\Phi_h : \Omega_h \to \Omega$, piecewisely defined on every simplex $K \in \mathcal{T}$. Therefore, a function $u_h \in S_h(\Omega_h)$ is discrete harmonic if and only if

$$\int_\Omega A_h \nabla \check{u}_h \cdot \nabla \check{v}_h \mathbf{d}x = 0 \quad \forall\, \check{v}_h \in \check{S}_h^\circ(\Omega), \tag{4.3.9}$$

Identity (4.3.9) will be used frequently in the following proof.

Since the map $\Phi_h : \Omega_h \to \Omega$ is close to the identity map $\mathrm{Id} : \mathbb{R}^N \to \mathbb{R}^N$ (which satisfies $\mathrm{Id}(x) \equiv x$), it follows that the matrix A_h is close to the identity matrix. In particular, the following results are corollaries of the second statement of Lemma 4.3.1:

$$\|\nabla^j(\Phi_h - \mathrm{Id})\|_{L^\infty(\Omega_h)} \le Ch^{r+1-j} \quad \text{and} \quad \|A_h - I\|_{L^\infty(\Omega)} \le Ch^r, \quad \text{for } j = 0, 1. \tag{4.3.10}$$

Therefore, for sufficiently small mesh size h, the perturbed bilinear form $\check{B}_h : H^1(\Omega) \times H^1(\Omega) \to \mathbb{R}$ defined by

$$\check{B}_h(v, \chi) = \int_\Omega A_h \nabla v \cdot \nabla \chi \mathbf{d}x \tag{4.3.11}$$

is continuous and coercive on $H_0^1(\Omega)$, i.e.,

$$\begin{aligned} &\check{B}_h(v, \chi) \leqslant C\|\nabla v\|_{L^2(\Omega)}\|\nabla\chi\|_{L^2(\Omega)} && \forall\, v, \chi \in H^1(\Omega), \\ &\check{B}_h(v, v) \geqslant C^{-1}\|\nabla v\|^2_{L^2(\Omega)} \sim \|v\|^2_{H^1(\Omega)} && \forall\, v \in H_0^1(\Omega). \end{aligned} \tag{4.3.12}$$

More precisely, the difference between $\check{B}_h(u, v)$ and $B(u, v)$ is estimated in the following lemma.

Lemma 4.3.7. *There exists a positive constant $h_1 > 0$ such that for $h \leqslant h_1$ the following result holds: If $1 \le p, q \le \infty$, $\frac{1}{p} + \frac{1}{q} = 1$, and $u \in W^{1,p}(\Omega)$, $v \in W^{1,q}(\Omega)$, then*

$$\left|\check{B}_h(u, v) - B(u, v)\right| \le Ch^r \|\nabla u\|_{L^p(\Lambda_h)} \|\nabla v\|_{L^q(\Lambda_h)}$$

where $\Lambda_h := \{x \in \Omega : \mathrm{dist}(x, \partial\Omega) \le 2h\}$.

Proof. Since $\Phi_h = \mathrm{Id}$ at all interior simplices, it follows that $A_h \circ \Phi_h = I$ outside the subdomain $D_h = \{x \in \Omega_h : \mathrm{dist}(x, \partial\Omega_h) \le h\}$. Correspondingly, $A_h = I$ outside the subdomain $\Phi_h(D_h)$ and therefore,

$$\begin{aligned} \left|\check{B}_h(u, v) - B(u, v)\right| &\le \|A_h - I\|_{L^\infty(\Phi_h(D_h))} \|\nabla u\|_{L^p(\Phi_h(D_h))} \|\nabla v\|_{L^q(\Phi_h(D_h))} \\ &\le Ch^r \|\nabla u\|_{L^p(\Phi_h(D_h))} \|\nabla v\|_{L^q(\Phi_h(D_h))}. \end{aligned}$$

If $x \in D_h$, then there exists $x' \in \partial\Omega_h$ such that $|x - x'| = \mathrm{dist}(x, \partial\Omega_h) \leqslant h$ and

$$|\Phi_h(x) - \Phi_h(x')| \leqslant |\Phi_h(x) - x| + |x - x'| + |x' - \Phi_h(x')| \leqslant Ch^{r+1} + h + Ch^{r+1},$$

which implies that

$$\mathrm{dist}(\Phi_h(x), \partial\Omega) \leqslant Ch^{r+1} + h.$$

For sufficiently small h we obtain $\mathrm{dist}(\Phi_h(x), \partial\Omega) \leqslant 2h$ and therefore $\Phi_h(D_h) \subset \Lambda_h$. ∎

4.3.3 Reduction of the problem

Let $x_0 \in \overline{\Omega}$ be a point satisfying

$$|\check{u}_h(x_0)| = \|\check{u}_h\|_{L^\infty(\Omega)} \quad \text{with} \quad d = \operatorname{dist}(x_0, \partial\Omega).$$

If $d \geqslant 2kh$ for some fixed $k \geqslant 1$, i.e., x_0 is relatively far away from the boundary $\partial\Omega$, then we can choose $\Omega_1 = \{x_0\}$ and $\Omega_2 = S_{d/2}(x_0)$ and use the interior L^∞ estimate established in [127, Corollary 5.1]. This yields the following result:

$$|\check{u}_h(x_0)| \leqslant C d^{-\frac{N}{2}} \|\check{u}_h\|_{L^2(S_d(x_0))}.$$

Otherwise, we have $d < 2kh$. In this case, assuming that $x_0 \in \check{K}$ for some curved simplex $\check{K} \in \check{\mathcal{T}}$, by the inverse estimate in Lemma 4.3.4 we have

$$|\check{u}_h(x_0)| = \|\check{u}_h\|_{L^\infty(\check{K})} \leqslant C h^{-\frac{N}{2}} \|\check{u}_h\|_{L^2(\check{K})} \leq C h^{-\frac{N}{2}} \|\check{u}_h\|_{L^2(S_{2kh}(x_0))}.$$

Overall, for either $d \geqslant 2kh$ or $d < 2kh$, the following estimate holds:

$$|\check{u}_h(x_0)| \leqslant C \rho^{-\frac{N}{2}} \|\check{u}_h\|_{L^2(S_\rho(x_0))}, \quad \text{with} \quad \rho = d + 2kh. \tag{4.3.13}$$

To estimate the term $\|\check{u}_h\|_{L^2(S_\rho(x_0))}$ on the right-hand side of (4.3.13), we use the following duality property:

$$\|\check{u}_h\|_{L^2(S_\rho(x_0))} = \sup_{\substack{\operatorname{supp}(\varphi) \subset S_\rho(x_0) \\ \|\varphi\|_{L^2(S_\rho(x_0))} \leqslant 1}} |(\check{u}_h, \varphi)|,$$

where $(\cdot, \cdot)$ denotes the inner product of $L^2(\Omega)$ (or $L^2(\Omega)^N$ for vector-valued functions), i.e.,

$$(u, v) := \int_\Omega u \cdot v \mathbf{d}x.$$

Hence, there exists a function $\varphi \in C_0^\infty(\Omega)$ with the following properties:

$$\operatorname{supp}(\varphi) \subset S_\rho(x_0), \quad \|\varphi\|_{L^2(S_\rho(x_0))} \leqslant 1, \tag{4.3.14}$$

$$\|\check{u}_h\|_{L^2(S_\rho(x_0))} \leqslant 2 |(\check{u}_h, \varphi)|. \tag{4.3.15}$$

For this function φ, we define $v \in H_0^1(\Omega)$ and $u \in H^1(\Omega)$ to be the solutions of the following elliptic equations (in the weak form):

$$\begin{cases} (A_h \nabla v, \nabla \chi) = (\varphi, \chi) & \forall \chi \in H_0^1(\Omega), \\ \qquad v = 0 & \text{on } \partial\Omega, \end{cases} \tag{4.3.16}$$

and

$$\begin{cases} (A_h \nabla u, \nabla \chi) = 0 & \forall \chi \in H_0^1(\Omega), \\ \qquad u = \check{u}_h & \text{on } \partial\Omega, \end{cases} \tag{4.3.17}$$

respectively. The maximum principle of the continuous problem (4.3.17) implies that

$$\|u\|_{L^\infty(\Omega)} \leqslant \|\check{u}_h\|_{L^\infty(\partial\Omega)}. \tag{4.3.18}$$

Therefore, we have

$$
\begin{aligned}
\|\breve{u}_h\|_{L^2(S_\rho(x_0))} &\leqslant 2|(\breve{u}_h,\varphi)| && \text{(here we have used (4.3.15))}\\
&= 2|(\breve{u}_h-u,\varphi)+(u,\varphi)| \\
&= 2|(A_h\nabla(\breve{u}_h-u),\nabla v)+(u,\varphi)| && \text{(here we have used (4.3.16))}\\
&= 2|(A_h\nabla\breve{u}_h,\nabla v)+(u,\varphi)| && \text{(here we have used (4.3.17))}\\
&\leqslant 2|(A_h\nabla\breve{u}_h,\nabla v)|+2\|u\|_{L^\infty(\Omega)}\|\varphi\|_{L^1(S_\rho(x_0))} && \text{(since } \mathrm{supp}(\varphi)\subset S_\rho(x_0))\\
&\leqslant 2|(A_h\nabla\breve{u}_h,\nabla v)|+C\rho^{\frac{N}{2}}\|\breve{u}_h\|_{L^\infty(\partial\Omega)}\|\varphi\|_{L^2(S_\rho(x_0))}, && (4.3.19)
\end{aligned}
$$

where we have used (4.3.18) and the Hölder inequality in deriving the last inequality. Combing inequalities (4.3.13) and (4.3.19), we have

$$\|\breve{u}_h\|_{L^\infty(\Omega)} = |\breve{u}_h(x_0)| \le C\rho^{-\frac{N}{2}}|(A_h\nabla\breve{u}_h,\nabla v)| + C\|\breve{u}_h\|_{L^\infty(\partial\Omega)} \tag{4.3.20}$$

where we have used the fact that $\|\varphi\|_{L^2(S_\rho(x_0))} \le 1$.

It remains to estimate $\rho^{-\frac{N}{2}}|(A_h\nabla\breve{u}_h,\nabla v)|$. To this end, we define $R_h : H_0^1(\Omega) \to \breve{S}_h^\circ(\Omega)$ to be the Ritz projection associated with the perturbed bilinear form defined in (4.3.11), i.e.,

$$(A_h\nabla(v-R_hv),\nabla\breve{\chi}_h) = 0 \quad \forall\, \breve{\chi}_h \in \breve{S}_h^\circ(\Omega), \tag{4.3.21}$$

which is well defined in view of the coercivity of the bilinear form; see (4.3.12). By using identity (4.3.9) for the discrete harmonic function u_h and the definition of the Ritz projection R_h in (4.3.21), we have

$$
\begin{aligned}
(A_h\nabla\breve{u}_h,\nabla v) &= (A_h\nabla\breve{u}_h,\nabla(v-R_hv)) \\
&= (A_h\nabla(\breve{u}_h-\breve{\chi}_h),\nabla(v-R_hv)) \quad \forall\breve{\chi}_h \in \breve{S}_h^\circ(\Omega).
\end{aligned} \tag{4.3.22}
$$

In particular, we can choose $\breve{\chi}_h = \chi_h \circ \Phi_h^{-1} \in \breve{S}_h^\circ(\Omega)$ to satisfy $\chi_h = u_h$ on all interior Lagrange nodes while $\chi_h = 0$ on all the boundary nodes (which implies $\chi_h = 0$ on $\partial\Omega_h$ and therefore $\breve{\chi}_h \equiv 0$ on $\partial\Omega$). Then

$$\|\breve{\chi}_h - \breve{u}_h\|_{L^\infty(\Omega)} \le C\|\breve{u}_h\|_{L^\infty(\partial\Omega)}. \tag{4.3.23}$$

Let $\Lambda_h = \{x \in \Omega : \mathrm{dist}(x,\partial\Omega) \leqslant 2h\}$ be a neighborhood of the boundary $\partial\Omega$, when h sufficiently small, $\breve{u}_h - \breve{\chi}_h = 0$ outside Λ_h. Then

$$
\begin{aligned}
|(A_h\nabla(\breve{u}_h-\breve{\chi}_h),\nabla(v-R_hv))| &\le C\|\nabla(\breve{\chi}_h-\breve{u}_h)\|_{L^\infty(\Omega)}\|\nabla(v-R_hv)\|_{L^1(\Lambda_h)} \\
&\le Ch^{-1}\|\breve{u}_h\|_{L^\infty(\partial\Omega)}\|\nabla(v-R_hv)\|_{L^1(\Lambda_h)},
\end{aligned} \tag{4.3.24}
$$

where we have used (4.3.23) and the inverse estimate for finite element functions. Substituting (4.3.22) and (4.3.24) into (4.3.20), we obtain

$$\|\breve{u}_h\|_{L^\infty(\Omega)} \leqslant C\big(\rho^{-\frac{N}{2}}h^{-1}\|\nabla(v-R_hv)\|_{L^1(\Lambda_h)}+1\big)\|u_h\|_{L^\infty(\partial\Omega)}. \tag{4.3.25}$$

The proof of Theorem 4.2.1 will be completed if the following result holds:

$$\rho^{-\frac{N}{2}}h^{-1}\|\nabla(v-R_hv)\|_{L^1(\Lambda_h)} \leqslant C, \tag{4.3.26}$$

which will be proved in the following subsections.

4.3.4 Regularity decomposition

In order to estimate the left-hand side of (4.3.26), we need to use a local energy estimate and a duality argument, which is based on the regularity result of the following elliptic equation (in the weak form): Find $v \in H_0^1(\Omega)$ such that

$$(A_h \nabla v, \nabla \chi) = (f, \chi) \quad \forall\, \chi \in H_0^1(\Omega), \tag{4.3.27}$$

where A_h is a globally discontinuous matrix-valued function defined in Section 4.3.2.

Due to the discontinuity of the coefficient matrix A_h, the standard H^2 regularity does not hold for the elliptic equation (4.3.27). We decompose the solution $v \in H_0^1(\Omega)$ of equation (4.3.27) into the following two parts:

$$v = v_1 + v_2, \tag{4.3.28}$$

where $v_1 \in H_0^1(\Omega)$ and $v_2 \in H_0^1(\Omega)$ are the weak solutions of the equations

$$(\nabla v_1, \nabla \chi) = (f, \chi) \qquad \forall \chi \in H_0^1(\Omega), \tag{4.3.29}$$
$$(A_h \nabla v_2, \nabla \chi) = ((I - A_h)\nabla v_1, \nabla \chi) \qquad \forall \chi \in H_0^1(\Omega). \tag{4.3.30}$$

Equation (4.3.29) has a constant coefficient and therefore the classical $W^{2,q}$ regularity estimate holds for $1 < q < 2 + \varepsilon$, for some $\varepsilon > 0$ which depends on the interior angles at the edges and corners of the domain Ω (see [39, Corollaries 3.7, 3.9 and 3.12]), i.e.,

$$\|v_1\|_{W^{2,q}(\Omega)} \le C_q \|f\|_{L^q(\Omega)} \quad \forall\, 1 < q < 2 + \varepsilon. \tag{4.3.31}$$

Since equation (4.3.30) has discontinuous coefficients, the $W^{2,q}$ regularity estimate does not hold. We have to estimate v_2 by using the $W^{1,p}$ estimate in the following lemma.

Lemma 4.3.8. *For every $1 < p < \infty$ there exists $h_p > 0$ (which depends on p), such that for $h \le h_p$, the solution $w \in H_0^1(\Omega)$ of the equation*

$$(A_h \nabla w, \nabla \chi) = (\vec{g}, \nabla \chi) \quad \forall \chi \in H_0^1(\Omega) \quad \textit{with} \quad \vec{g} \in L^p(\Omega)^N \cap L^2(\Omega)^N, \tag{4.3.32}$$

satisfies $w \in W^{1,p}(\Omega)$ and

$$\|w\|_{W^{1,p}(\Omega)} \le C_p \|\vec{g}\|_{L^p(\Omega)}, \tag{4.3.33}$$

where C_p is a constant which is independent of h (possibly depending on p).

Proof. We can rewrite equation (4.3.32) into the following form:

$$(\nabla w, \nabla \chi) = (\vec{g}, \nabla \chi) + ((I - A_h)\nabla w, \nabla \chi) \quad \forall \chi \in H_0^1(\Omega),$$

and apply the $W^{1,p}$ regularity estimate for the Poisson equation (which holds in a smooth domain or curvilinear polyhedron with edge openings smaller than π; see [39, Corollaries 3.7, 3.9 and 3.12]). This yields the following inequality:

$$\|w\|_{W^{1,p}(\Omega)} \le C_p \|\vec{g}\|_{L^p(\Omega)} + C_p \|I - A_h\|_{L^\infty(\Omega)} \|w\|_{W^{1,p}(\Omega)}.$$

Since $\|A_h - I\|_{L^\infty} \le Ch$, for sufficiently small h (depending on p) the last term on the right-hand side can be absorbed by the left-hand side. This yields the result of Lemma 4.3.8. ∎

By combining the $W^{2,q}$ regularity estimate in (4.3.31) and the $W^{1,p}$ regularity estimate in Lemma 4.3.8, we can prove the following result.

Lemma 4.3.9. *Let $1 < p, q < \infty$ be numbers such that $1/q \leq 1/n + 1/p$, and assume that $h \leq h_p$, where h_p is given in Lemma 4.3.8. Let $w \in H_0^1(\Omega)$ be the weak solution of the equation*

$$(A_h \nabla w, \nabla \chi) = (f, \chi) + (\vec{g}, \nabla \chi) \quad \forall \chi \in H_0^1(\Omega) \tag{4.3.34}$$

for some $f \in L^q(\Omega) \cap L^2(\Omega)$ and $\vec{g} \in L^p(\Omega)^N \cap L^2(\Omega)^N$. Then $w \in W^{1,p}(\Omega)$ and

$$\|w\|_{W^{1,p}(\Omega)} \leq C_p \|f\|_{L^q(\Omega)} + C_p \|\vec{g}\|_{L^p(\Omega)}. \tag{4.3.35}$$

Proof. We consider the decomposition $w = w_1 + w_2$ with $w_1, w_2 \in H_0^1(\Omega)$ weakly solving

$$\begin{aligned} &(\nabla w_1, \nabla \chi) = (f, \chi) && \forall \chi \in H_0^1(\Omega), \\ &(A_h \nabla w_2, \chi) = ((I - A_h)\nabla w_1 + \vec{g}, \nabla \chi) && \forall \chi \in H_0^1(\Omega). \end{aligned}$$

Note that for $\chi \in W_0^{1,p'}(\Omega)$ where $1/p + 1/p' = 1$

$$\begin{aligned} |(f, \chi)| \leq & \|f\|_{L^q(\Omega)} \|\chi\|_{L^{q'}(\Omega)} \quad (1/q + 1/q' = 1) \\ \leq & C\|f\|_{L^q(\Omega)} \|\chi\|_{W^{1,p'}(\Omega)} \quad (\text{embedding } W^{1,p'} \hookrightarrow L^{q'} \text{ used}), \end{aligned}$$

therefore we have $\|f\|_{W^{-1,p}(\Omega)} \leq C\|f\|_{L^q(\Omega)}$. By the $W^{1,p}$ regularity estimate for the Poisson equation on curvilinear polyhedron (see [39, Corollaries 3.7, 3.9 and 3.12]), there holds

$$\|w_1\|_{W^{1,p}(\Omega)} \leq C_p \|f\|_{W^{-1,p}(\Omega)} \leq C_p \|f\|_{L^q(\Omega)}$$

Then we apply the $W^{1,p}$ estimate in Lemma 4.3.8 to the equation of w_2. This yields the following result:

$$\|w_2\|_{W^{1,p}(\Omega)} \leq C_p \|\vec{g} + (I - A_h)\nabla w_1\|_{L^p(\Omega)} \leq C_p \|\vec{g}\|_{L^p(\Omega)} + C_p \|f\|_{L^q(\Omega)}.$$

The result of Lemma 4.3.9 follows from combining the estimates for w_1 and w_2. ∎

The following lemma was proved in [96, Lemma 2.2] for polyhedral domains. The proof of this result for smooth domains and curvilinear polyhedron is the same.

Lemma 4.3.10. *If $\chi \in W_0^{1,p}(\Omega)$ for some $1 < p < \infty$ and $x^* \in \partial\Omega$, then*

$$\|\chi\|_{L^p(S_{d_*}(x^*))} \leqslant C d_* \|\nabla \chi\|_{L^p(\Omega)},$$

where $S_{d_*}(x^*) := \{x \in \Omega : |x - x^*| < d_*\}$.

Lemma 4.3.11. *Let $1 < p < \infty$ and $h \leq h_p$, where h_p is given in Lemma 4.3.8. For*

$$f \in L^p(\Omega) \cap L^2(\Omega) \text{ with } \operatorname{supp}(f) \subset S_{d_*}(x_0), \text{ where } x_0 \in \overline{\Omega} \text{ and } \operatorname{dist}(x_0, \partial\Omega) \leq d_*,$$

the solution $v \in H_0^1(\Omega)$ of equation (4.3.27) satisfies

$$\|v\|_{W^{1,p}(\Omega)} \leq C_p d_* \|f\|_{L^p(\Omega)} \tag{4.3.36}$$

Proof. We consider the decomposition $v = v_1 + v_2$ in (4.3.28)–(4.3.30). If $\operatorname{dist}(x_0, \partial\Omega) \leq d_*$, then $S_{d_*}(x_0) \subset S_{2d_*}(\bar{x}_0)$ for some $\bar{x}_0 \in \partial\Omega$. Note that for $\chi \in W_0^{1,p'}(\Omega)$ where $1/p + 1/p' = 1$, we have

$$\begin{aligned}|(f,\chi)| \leq& \|f\|_{L^p(S_{d_*}(x_0))}\|\chi\|_{L^{p'}(S_{d_*}(x_0))}\\ \leq& \|f\|_{L^p(S_{d_*}(x_0))}\|\chi\|_{L^{p'}(S_{2d_*}(\bar{x}_0))}\\ \leq& C d_* \|f\|_{L^p(\Omega)}\|\nabla\chi\|_{L^{p'}(\Omega)}, \quad \text{(Lemma 4.3.10 used)}\end{aligned}$$

which implies that $\|f\|_{W^{-1,p}(\Omega)} \leq C d_* \|f\|_{L^p(\Omega)}$. Thus by the $W^{1,p}$ regularity estimate for the Poisson equation on curvilinear polyhedron (see [39, Corollaries 3.7, 3.9 and 3.12]), there holds:

$$\|v_1\|_{W^{1,p}(\Omega)} \leq C_p \|f\|_{W^{-1,p}(\Omega)} \leq C_p d_* \|f\|_{L^p(\Omega)}.$$

By applying Lemma 4.3.8 to equation (4.3.30), we obtain

$$\|v_2\|_{W^{1,p}(\Omega)} \leq C_p \|(I - A_h)\nabla v_1\|_{L^p(\Omega)} \leq C_p h \|v_1\|_{W^{1,p}(\Omega)} \leq C_p h d_* \|f\|_{L^p(\Omega)}.$$

The last two inequalities imply the result of Lemma 4.3.11. ∎

The next lemma is about the Cacciopoli inequality for harmonic functions which is the same as in [123, Lemma 8.3]. The result holds for smooth domains and curvilinear polyhedra on which the elliptic H^2 regularity result holds for the Poisson equation.

Lemma 4.3.12. *Let D and D_d be two subdomains of Ω satisfying $D \subset D_d \subset \Omega$, with*

$$D_d = \{x \in \Omega : dist(x, D) \leqslant d\},$$

where d is a positive constant. If $v \in H_0^1(\Omega)$ and v is harmonic on D_d, i.e.

$$(\nabla v, \nabla w) = 0 \quad \forall w \in H_0^1(D_d),$$

then the following estimates hold:

$$|v|_{H^2(D)} \leqslant C d^{-1} \|v\|_{H^1(D_d)}, \tag{4.3.37a}$$

$$\|v\|_{H^1(D)} \leqslant C d^{-1} \|v\|_{L_2(D_d)}. \tag{4.3.37b}$$

We also need the following interior estimate in the estimation of v_2.

Lemma 4.3.13. *Let $1 < p, q < \infty$ be numbers such that $1/q \leq 1/n + 1/p$ and assume that $h \leq h_p$, where h_p is given in Lemma 4.3.8. Let $D \subset D_d \subset \Omega$ be subdomains, with $D_d = \{x \in \Omega : dist(x, D) \leqslant d\}$. If $v \in W_0^{1,p}(\Omega) \cap H_0^1(\Omega)$ satisfies equation*

$$(A_h \nabla v, \nabla \chi) = 0 \quad \forall \chi \in H_0^1(D_d), \tag{4.3.38}$$

or

$$(\nabla v, \nabla \chi) = 0 \quad \forall \chi \in H_0^1(D_d). \tag{4.3.39}$$

Then

$$\|v\|_{W^{1,p}(D)} \leq \frac{C_p}{d}(\|v\|_{L^p(D_d)} + \|v\|_{W^{1,q}(D_d)}). \tag{4.3.40}$$

Proof. We focus on the first case: v satisfies equation (4.3.38). The proof for the second case is the same and therefore omitted.

First, we choose a cut-off function $\omega \in C_0^\infty(\mathbb{R}^N)$, $\omega \equiv 1$ on D, $\mathrm{supp}(\omega) \cap \Omega \subset D_d$, with $\|\omega\|_{W^{1,\infty}(\mathbb{R}^N)} \leq Cd^{-1}$. Then $\omega v \in H_0^1(\Omega)$ satisfies the following equation:

$$\begin{aligned}(A_h\nabla(\omega v), \nabla\chi) &= (\omega A_h \nabla v, \nabla\chi) + (A_h\nabla\omega, v\nabla\chi)\\ &= (A_h\nabla v, \nabla(\omega\chi)) - (A_h\nabla v, \chi\nabla\omega) + (A_h\nabla\omega, v\nabla\chi)\\ &= (A_h v\nabla\omega, \nabla\chi) - (A_h\nabla v \cdot \nabla\omega, \chi) \quad \forall\chi \in H_0^1(\Omega)\end{aligned}$$

where we have used the identity $(A_h\nabla v, \nabla(\omega\chi)) = 0$ in the derivation of the last equality, which is a consequence of (4.3.38) and $\omega\chi \in H_0^1(D_d)$. Then we can apply Lemma 4.3.9 to the above equation satisfied by ωv. This yields the following result:

$$\begin{aligned}\|\omega v\|_{W^{1,p}(\Omega)} &\leq C_p\|A_h v\nabla\omega\|_{L^p(\Omega)} + C_p\|A_h\nabla v \cdot \nabla\omega\|_{L^q(\Omega)}\\ &\leq \frac{C_p}{d}\|v\|_{L^p(D_d)} + \frac{C_p}{d}\|v\|_{W^{1,q}(D_d)}.\end{aligned}$$

Since $\omega = 1$ on D, the last inequality implies the result of Lemma 4.3.13. ∎

Lemma 4.3.14. *Let $1 < p, q < \infty$ be numbers such that $1/q \leq 1/n + 1/p$ and assume that $h \leq \min\{h_p, h_q\}$, where h_p, h_q are given in Lemma 4.3.8. Let $D \subset D_d \subset \Omega$ be subdomains, with $D_d = \{x \in \Omega : dist(x, D) \leqslant d\}$. If the source function f has* $\mathrm{supp}(f) \cap D_d = \emptyset$*, then the solution v_2 of equation* (4.3.30) *satisfies the following estimate:*

$$\|v_2\|_{W^{1,p}(D)} \leq \frac{C_{p,q}}{d} h\|v_1\|_{W^{1,q}(\Omega)}. \tag{4.3.41}$$

Proof. We consider a cut-off function ω such that $\omega \equiv 1$ in D and $\mathrm{supp}(\omega) \subset D_{d/2}$, with $\|\omega\|_{W^{1,\infty}(\mathbb{R}^N)} \leq Cd^{-1}$. Then the following equation can be written down similarly as in the proof of Lemma 4.3.13:

$$\begin{aligned}(A_h\nabla(\omega v_2), \nabla\chi) = &(\omega(I - A_h)\nabla v_1, \nabla\chi) + ((I - A_h)\nabla v_1 \cdot \nabla\omega, \chi)\\ &+ (v_2 A_h\nabla\omega, \nabla\chi) - (A_h\nabla v_2 \cdot \nabla\omega, \chi) \quad \forall\chi \in H_0^1(\Omega).\end{aligned}$$

By applying Lemma 4.3.9 to the equation above, we obtain

$$\begin{aligned}\|v_2\|_{W^{1,p}(D)} &\leq C_p h\|v_1\|_{W^{1,p}(D_{d/2})} + \frac{C_p h}{d}\|v_1\|_{W^{1,q}(\Omega)} + \frac{C_p}{d}\|v_2\|_{L^p(\Omega)} + \frac{C_p}{d}\|v_2\|_{W^{1,q}(\Omega)}\\ &\leq C_p h\|v_1\|_{W^{1,p}(D_{d/2})} + \frac{C_p h}{d}\|v_1\|_{W^{1,q}(\Omega)} + \frac{C_p}{d}\|v_2\|_{W^{1,q}(\Omega)},\end{aligned} \tag{4.3.42}$$

where we have used Sobolev embedding $W^{1,q}(\Omega) \hookrightarrow L^p(\Omega)$.

Since $\mathrm{supp}(f) \cap D_d = \emptyset$, it follows that the solution v_1 of (4.3.29) satisfies equation (4.3.39). Therefore, Lemma 4.3.13 implies that

$$\|v_1\|_{W^{1,p}(D_{d/2})} \leq \frac{C_p}{d}(\|v\|_{L^p(D_d)} + \|v\|_{W^{1,q}(D_d)}) \leq \frac{C_p}{d}\|v_1\|_{W^{1,q}(\Omega)}.$$

By applying Lemma 4.3.8 to equation (4.3.30), we also obtain

$$\|v_2\|_{W^{1,q}(\Omega)} \leq C_q\|I - A_h\|_{L^\infty(\Omega)}\|v_1\|_{W^{1,q}(\Omega)} \leq C_q h\|v_1\|_{W^{1,q}(\Omega)}.$$

Then, substituting the last two inequalities into (4.3.42), we obtain the result of Lemma 4.3.14. ∎

4.3.5 $W^{1,p}$ stablity of the Ritz projection (with discontinuous coefficients)

In [75] the $W^{1,\infty}$ stability of the Ritz projection is proved for the Poisson equation in convex polyhedral domains. The proof is based on the following properties of the domain and finite elements:

(P1) Hölder estimates of the Green function for the Poisson equation, i.e.,

$$\begin{aligned} \frac{|\partial_{x_i} G(x,\xi) - \partial_{y_i} G(y,\xi)|}{|x-y|^\sigma} &\le C\left(|x-\xi|^{-2-\sigma} + |y-\xi|^{-2-\sigma}\right) \\ \frac{\left|\partial_{x_i}\partial_{\xi_j} G(x,\xi) - \partial_{y_i}\partial_{\xi_j} G(y,\xi)\right|}{|x-y|^\sigma} &\le C\left(|x-\xi|^{-3-\sigma} + |y-\xi|^{-3-\sigma}\right) \end{aligned} \tag{4.3.43}$$

for $i,j = 1,2,3$.

(P2) Elliptic H^2 regularity result for the Poisson equation.

(P3) Exact triangulation which matches the boundary $\partial\Omega$.

(P4) Error estimates for the Lagrange interpolation holds as in Lemma 4.3.2.

Note that the Hölder estimates for the Green function in (4.3.43) was proved in [75] for general curvilinear polyhedral domains with edge opening smaller than π, instead of merely classical polyhedral domains. If we define a modified Ritz projection R_h^* associated to the Poisson equation (without the discontinuous coefficient A_h), i.e.,

$$\int_\Omega \nabla(v - R_h^* v)\cdot\nabla\check{\chi}_h \mathbf{d}x = 0 \quad \forall \check{\chi}_h \in \check{S}_h^\circ(\Omega), \tag{4.3.44}$$

then all the properties in (P1)–(P4) are possessed by the curvilinear polyhedral domain Ω and the finite element space $\check{S}_h^\circ(\Omega)$. The latter is based on the triangulation $\check{K}$ which matches the boundary $\partial\Omega$ exactly. Therefore, the $W^{1,\infty}$ stability still holds for the modified Ritz projection defined in (4.3.44). The result is stated in the following lemma.

Lemma 4.3.15.

$$\|R_h^* v\|_{W^{1,\infty}(\Omega)} \le C\|v\|_{W^{1,\infty}(\Omega)} \quad \forall\, v \in H_0^1(\Omega)\cap W^{1,\infty}(\Omega). \tag{4.3.45}$$

By real interpolation between the H^1 and $W^{1,\infty}$ stability estimates (see [32, result in (5.1)]), we obtain the $W^{1,p}$ stability of the modified Ritz projection for $2 \le p \le \infty$. The result can also be extended to $1 < p \le 2$ by a duality argument as in [21, Section 8.5], which requires Poisson equation to have the $W^{1,p'}$ regularity (this is true for a curvilinear polyhedron with edge opening smaller than π). The result is summarized below.

Lemma 4.3.16 ($W^{1,p}$ stability of the modified Ritz projection R_h^*)**.** *For any $1 < p \leqslant \infty$, there exists a positive constant h_p such that for $h \leqslant h_p$ the following result holds:*

$$\|R_h^* u\|_{W^{1,p}(\Omega)} \le C_p\|u\|_{W^{1,p}(\Omega)} \quad \forall u \in W^{1,p}(\Omega)\cap H_0^1(\Omega). \tag{4.3.46}$$

By a "perturbation" argument, similar as [21, Section 8.6], one can obtain the $W^{1,p}$ stability of the Ritz projection R_h. This is stated in the following proposition.

Proposition 4.3.17 ($W^{1,p}$ stability of the Ritz projection R_h)**.** *For any $1 < p < \infty$, there exists a positive constant h_p such that for $h \leqslant h_p$ the following result holds:*

$$\|R_h u\|_{W^{1,p}(\Omega)} \leq C_p \|u\|_{W^{1,p}(\Omega)} \quad \forall u \in W^{1,p}(\Omega) \cap H_0^1(\Omega). \tag{4.3.47}$$

Proof. For $v \in H_0^1(\Omega)$, its Ritz projection $R_h v \in \check{S}_h^\circ(\Omega)$ satisfies the following equation:

$$\int_\Omega \nabla(v - R_h v) \cdot \nabla \check{\chi}_h \mathbf{d}x = \int_\Omega (I - A_h)\nabla(v - R_h v) \cdot \nabla \check{\chi}_h \mathbf{d}x \quad \forall \check{\chi}_h \in \check{S}_h^\circ(\Omega).$$

If we define w to be the solution of the following elliptic equation (in the weak form):

$$\int_\Omega \nabla w \cdot \nabla \check{\chi} \mathbf{d}x = -\int_\Omega (I - A_h)\nabla(v - R_h v) \cdot \nabla \check{\chi} \mathbf{d}x \quad \forall \check{\chi} \in H_0^1(\Omega),$$

then

$$\int_\Omega \nabla(w + v - R_h v) \cdot \nabla \check{\chi}_h \mathbf{d}x = 0 \quad \forall \check{\chi}_h \in \check{S}_h^\circ(\Omega),$$

which means that $R_h v = R_h^*(w + v)$. Lemma 4.3.16 implies that

$$\begin{aligned}\|R_h v\|_{W^{1,p}(\Omega)} = \|R_h^*(w+v)\|_{W^{1,p}(\Omega)} &\leqslant C_p \|w + v\|_{W^{1,p}(\Omega)} \\ &\leqslant C_p \|I - A_h\|_{L^\infty(\Omega)} \|v - R_h v\|_{W^{1,p}(\Omega)} + C_p \|v\|_{W^{1,p}(\Omega)} \\ &\leqslant C_p h \|R_h v\|_{W^{1,p}(\Omega)} + C_p \|v\|_{W^{1,p}(\Omega)}.\end{aligned}$$

There exists a constant h_p such that for $h \leqslant h_p$ the first term on the right-hand side can be absorbed by the left-hand side. In this case we obtain the result of Proposition 4.3.17. ∎

As a result of Proposition 4.3.17, we obtain the following $W^{1,p}$ error estimate for the Ritz projection.

Lemma 4.3.18. *For any $1 < q < 2 + \varepsilon$, there exists a positive constant h_q such that for $h \leqslant h_q$ the solution of equation* (4.3.27) *has the following error bound*:

$$\|v - R_h v\|_{W^{1,q}(\Omega)} \leqslant C_q h \|f\|_{L^q(\Omega)} \quad \forall f \in L^q(\Omega) \cap L^2(\Omega).$$

Proof. We consider the decomposition $v = v_1 + v_2$ in (4.3.28)–(4.3.30). The $W^{2,q}$ estimate in (4.3.31) and the $W^{1,p}$ estimate in Lemma 4.3.8 imply that v_1 and v_2 satisfy the following estimates:

$$\begin{aligned}&\|v_1\|_{W^{2,q}(\Omega)} \leq C_q \|f\|_{L^q(\Omega)} \quad \forall 1 < q < 2 + \varepsilon, \\ &\|v_2\|_{W^{1,q}(\Omega)} \leq C_q h \|v_1\|_{W^{1,q}(\Omega)} \leq C_q h \|f\|_{L^q(\Omega)}.\end{aligned}$$

Applying the $W^{1,q}$ stability of the Ritz projection, we obtain the following estimates:

$$\|v_1 - R_h v_1\|_{W^{1,q}(\Omega)} \leq C_q \inf_{\check{\chi}_h \in \check{S}_h^\circ(\Omega)} \|v_1 - \check{\chi}_h\|_{W^{1,q}(\Omega)} \leq C_q h \|v_1\|_{W^{2,q}(\Omega)} \leq C_q h \|f\|_{L^q(\Omega)},$$

$$\|v_2 - R_h v_2\|_{W^{1,q}(\Omega)} \leq C_q \|v_2\|_{W^{1,q}(\Omega)} \leq C_q h \|f\|_{L^q(\Omega)}.$$

The result of Lemma 4.3.18 is obtained by combining the two estimates above. ∎

Finally, the L^p error estimate for the Ritz projection follows from a standard duality argument, again by using the regularity decomposition as in (4.3.28)–(4.3.30) for the dual problem.

Lemma 4.3.19. *For any $1 < q < 2 + \varepsilon$, there exists a positive constant h_q such that for $h \leqslant h_q$ the following error estimate holds:*

$$\|u - R_h u\|_{L^{q'}(\Omega)} \leq C_q h \|u - R_h u\|_{W^{1,q'}(\Omega)} \quad \forall\, u \in H_0^1(\Omega) \cap W^{1,q'}(\Omega), \tag{4.3.48}$$

where $1/q + 1/q' = 1$.

Proof. By using the duality between $L^q(\Omega)$ and $L^{q'}(\Omega)$, we can express the $L^{q'}$ error of the Ritz projection as

$$\|R_h u - u\|_{L^{q'}(\Omega)} = \sup_{\substack{\varphi \in C_0^\infty(\Omega) \\ \|\varphi\|_{L^q(\Omega)} \leq 1}} (R_h u - u, \varphi),$$

In particular, there exists $\varphi \in C_0^\infty(\Omega)$ with $\|\varphi\|_{L^q(\Omega)} \leq 1$ such that

$$\|R_h u - u\|_{L^{q'}(\Omega)} \leq 2(R_h u - u, \varphi).$$

Let $v \in H_0^1(\Omega)$ be the weak solution of the following elliptic equation (in the weak form):

$$(A_h \nabla v, \nabla \chi) = (\varphi, \chi) \quad \forall\, \chi \in H_0^1(\Omega).$$

Then

$$\begin{aligned}
(R_h u - u, \varphi) &= (A_h \nabla v, \nabla (R_h u - u)) \\
&= (A_h \nabla (R_h u - u), \nabla v) \\
&= (A_h \nabla (R_h u - u), \nabla (v - R_h v)) \\
&\leq C \|R_h u - u\|_{W^{1,q'}(\Omega)} \|R_h v - v\|_{W^{1,q}(\Omega)} \\
&\leq C_q h \|R_h u - u\|_{W^{1,q'}(\Omega)} \|\varphi\|_{L^q(\Omega)} \quad \text{(Lemma 4.3.18 is used here)} \\
&\leq C_q h \|R_h u - u\|_{W^{1,q'}(\Omega)}.
\end{aligned}$$

This proves the result of Lemma 4.3.19 . ∎

4.3.6 Estimation of $\rho^{-\frac{N}{2}} h^{-1} \|\nabla (v - R_h v)\|_{L^1(\Lambda_h)}$

In this subsection, we prove (4.3.26) by utilizing the results established in Sections 4.3.4–4.3.5, where v is the solution of (4.3.16). This would complete the proof of Theorem 4.2.1. To this end, we consider a dyadic decomposition of the domain as in the literature; see [75, 96, 125].

Let $R_0 = \mathrm{diam}(\Omega)$ and $d_j = R_0 2^{-j}$ for $j \geqslant 0$. We define a sequence of subdomains

$$D_j = \{x \in \Omega : d_{j+1} \leqslant |x - x_0| \leqslant d_j\} \quad \text{for } j \geqslant 0.$$

For each j we denote by D_j^l a subdomain slightly larger than D_j, defined by

$$D_j^l = D_{j-l} \cup \cdots \cup D_j \cup D_{j+1} \cup \cdots \cup D_{j+l} \quad (D_i := \emptyset \text{ for } i < 0.)$$

Let $J = [\ln_2(R_0/2\kappa\rho)] + 1$, where $[\ln_2(R_0/2\kappa\rho)]$ denotes the biggest integer not exceeding $\ln_2(R_0/2\kappa\rho)$. The constant $\kappa > 32$ will be determined below, and the generic constant C will be independent on κ until it is determined (unless it contains a subscript κ). The definition above implies that

$$\frac{1}{2}\kappa\rho \leqslant d_{J+1} \leqslant \kappa\rho$$

and

$$\text{measure}(D_j \cap \Lambda_h) \leq Chd_j^{N-1}. \tag{4.3.49}$$

Note that v is the solution of (4.3.16), where $\varphi = 0$ outside $S_\rho(x_0)$. Therefore, $\varphi = 0$ in D_j^3 for $1 \leqslant j \leqslant J$. This result will be used below.

By using the subdomains defined above, we have

$$\begin{aligned}
&\rho^{-\frac{N}{2}} h^{-1} \|\nabla(v - R_h v)\|_{L^1(\Lambda_h)} \\
&\leqslant \rho^{-\frac{N}{2}} h^{-1} \left(\sum_{j=0}^{J} \|\nabla(v - R_h v)\|_{L^1(\Lambda_h \cap D_j)} + \|\nabla(v - R_h v)\|_{L^1(\Lambda_h \cap S_{\kappa\rho}(x_0))} \right) \\
&\leqslant C\rho^{-\frac{N}{2}} h^{-1} \sum_{j=0}^{J} h^{\frac{1}{2}} d_j^{\frac{N-1}{2}} \|\nabla(v - R_h v)\|_{L^2(\Lambda_h \cap D_j)} \\
&\quad + C\kappa^{\frac{N-1}{2}} \rho^{-\frac{1}{2}} h^{-\frac{1}{2}} \|\nabla(v - R_h v)\|_{L^2(\Lambda_h \cap S_{\kappa\rho}(x_0))},
\end{aligned} \tag{4.3.50}$$

where the Hölder inequality and (4.3.49) are used in the derivation of the last inequality. By choosing $q = 2$ in Lemma 4.3.18 we have

$$\|\nabla(v - R_h v)\|_{L^2(\Omega)} \leq Ch\|\varphi\|_{L^2(\Omega)} \leq Ch. \tag{4.3.51}$$

Then, substituting (4.3.51) into the last term on the right-hand side of (4.3.50) and using the fact that $\rho \geqslant h$ (which follows from the definition of ρ in (4.3.13)), we obtain

$$\rho^{-\frac{N}{2}} h^{-1} \|\nabla(v - R_h v)\|_{L^1(\Lambda_h)} \leq C\rho^{-\frac{N}{2}} h^{-\frac{1}{2}} \sum_{j=0}^{J} d_j^{\frac{N-1}{2}} \|\nabla(v - R_h v)\|_{L^2(D_j)} + C_\kappa, \tag{4.3.52}$$

where C_κ denotes a constant which depends on the parameter κ.

It remains to estimate $\|\nabla(v - R_h v)\|_{L^2(D_j)}$. To this end, we use the following interior energy estimate for the solution of (4.3.16):

$$\|v - R_h v\|_{H^1(D_j)} \leq C\|v - \check{I}_h v\|_{H^1(D_j^1)} + Cd_j^{-1}\|v - \check{I}_h v\|_{L^2(D_j^1)} + Cd_j^{-1}\|v - R_h v\|_{L^2(D_j^1)}. \tag{4.3.53}$$

The proof of such interior energy estimate is omitted as it only requires the coefficient matrix A_h to be L^∞ in the perturbed bilinear form in (4.3.11), without additional smoothness, and therefore is the same as the proof for standard finite elements for the Poisson equation.

We use the decomposition $v = v_1 + v_2$ in (4.3.28)–(4.3.30) with $f = \varphi$ supported in $S_\rho(x_0)$, and consider interpolation error of v_1 and v_2, respectively. First, by applying the result of Lemma 4.3.2 and using the fact that $d_j > h$, we have

$$\|v_1 - \check{I}_h v_1\|_{H^1(D_j^1)} + d_j^{-1}\|v_1 - \check{I}_h v_1\|_{L^2(D_j^1)} \leqslant Ch\|v_1\|_{H^2(D_j^2)} \leqslant Chd_j^{-1+\frac{N}{2}-\frac{N}{p}}\|v_1\|_{W^{1,p}(\Omega)} \quad \text{for } \tfrac{2N}{N+2} < p < 2, \tag{4.3.54}$$

where we have used the following inequality in deriving the last inequality:

$$\|v_1\|_{H^2(D_j^2)} \leqslant Cd_j^{\frac{1}{2}-\frac{3}{p}}\|v_1\|_{W^{1,p}(\Omega)} \quad \text{for } \tfrac{2N}{N+2} < p < 2. \tag{4.3.55}$$

The inequality above follows from Lemma 4.3.12 (because v_1 is the solution of (4.3.29) with $f = \varphi = 0$ in D_j^3), the Hölder inequality and the Sobolev embedding inequality, i.e.,

$$\begin{aligned}\|v_1\|_{H^2(D_j^2)} &\leqslant Cd_j^{-2}\|v_1\|_{L^2(D_j^3)}\\ &\leqslant Cd_j^{-2+\frac{N}{2}-\frac{N}{p_*}}\|v_1\|_{L^{p_*}(D_j^3)} \quad \text{if } p_*>2\\ &\leqslant Cd_j^{-1+\frac{N}{2}-\frac{N}{p}}\|v_1\|_{W^{1,p}(\Omega)} \quad \text{for } \tfrac{N}{p_*}=\tfrac{N}{p}-1 \text{ and } \tfrac{2N}{N+2}<p<2\\ &\qquad\qquad \text{so that } p_*>2 \text{ and } W^{1,p}(\Omega)\hookrightarrow L^{p^*}(\Omega).\end{aligned}$$

Here we require $\kappa > 32$ to guarantee that $d_{J+5} > \rho$, which is required in the use Lemma 4.3.12. This proves the last inequality in (4.3.54).

Next, we consider the interpolation error of v_2 by using Lemma 4.3.2 and Hölder inequality, i.e.,

$$\begin{aligned}\|v_2-\check{I}_hv_2\|_{H^1(D_j^1)}+d_j^{-1}\|v_2-\check{I}_hv_2\|_{L^2(D_j^1)} &\le Cd_j^{\frac{N}{2}-\frac{N}{p_1}}\|v_2\|_{W^{1,p_1}(D_j^2)}\\ &\le Cd_j^{\frac{N}{2}-\frac{N}{q_1}}h\|v_1\|_{W^{1,q_1}(\Omega)}\\ &\text{for some } p_1>N \text{ and } \frac{N}{q_1}=\frac{N}{p_1}+1, \quad (4.3.56)\end{aligned}$$

where we have applied Corollary 4.3.14 in deriving the last inequality. (Here we only need p_1 to be slightly bigger than N, and therefore the corresponding q_1 here can be smaller than 2, so that we can use Hölder inequality to estimate $\|\varphi\|_{L^{q_1}(S_\rho(x_0))}$ below.)

By combining (4.3.54) and (4.3.56), we obtain

$$\begin{aligned}&C\|v-\check{I}_hv\|_{H^1(D_j^1)}+Cd_j^{-1}\|v-\check{I}_hv\|_{L^2(D_j^1)}\\ &\le Chd_j^{-1+\frac{N}{2}-\frac{N}{p}}\|v_1\|_{W^{1,p}(\Omega)}+Cd_j^{\frac{N}{2}-\frac{N}{q_1}}h\|v_1\|_{W^{1,q_1}(\Omega)}\\ &\le Chd_j^{-1+\frac{N}{2}-\frac{N}{p}}\rho\|\varphi\|_{L^p(S_\rho(x_0))}+Chd_j^{\frac{N}{2}-\frac{N}{q_1}}\rho\|\varphi\|_{L^{q_1}(S_\rho(x_0))}\\ &\le Chd_j^{-1+\frac{N}{2}-\frac{N}{p}}\rho^{1-\frac{N}{2}+\frac{N}{p}}+Chd_j^{\frac{N}{2}-\frac{N}{q_1}}\rho^{1-\frac{N}{2}+\frac{N}{q_1}}, \qquad (4.3.57)\end{aligned}$$

where we have applied Lemma 4.3.11 to equation (4.3.29) in the derivation of the second inequality, and used Hölder inequality in the derivation of the last inequality.

Finally, substituting (4.3.57) into (4.3.53), we obtain

$$\begin{aligned}&d_j^{\frac{N-1}{2}}\|\nabla(v-R_hv)\|_{L^2(D_j)}\\ &\le Chd_j^{N-\frac{3}{2}-\frac{N}{p}}\rho^{1-\frac{N}{2}+\frac{N}{p}}+Chd_j^{N-\frac{1}{2}-\frac{N}{q_1}}\rho^{1-\frac{N}{2}+\frac{N}{q_1}}+Cd_j^{\frac{N-3}{2}}\|v-R_hv\|_{L^2(D_j^1)}\\ &\le Chd_j^{N-\frac{3}{2}-\frac{N}{p}}\rho^{1-\frac{N}{2}+\frac{N}{p}}+Cd_j^{\frac{N-3}{2}}\|v-R_hv\|_{L^2(D_j^1)}, \qquad (4.3.58)\end{aligned}$$

where we have chosen $p = q_1 < 2$ and used $d_j \le C$ in the derivation of the last inequality. Here we can make p as close to 2 as possible so that $p = q'$ satisfies the condition in Lemma 4.3.19 (which will be used in the subsequent analysis).

Now we substitute (4.3.58) into (4.3.52) and use the result $\sum_{j=0}^{J} d_j^{N-\frac{3}{2}-\frac{N}{p}} \le C_\kappa\rho^{N-\frac{3}{2}-\frac{N}{p}}$, we obtain

$$\sum_{j=0}^{J} d_j^{\frac{N-1}{2}}\|\nabla(v-R_hv)\|_{L^2(D_j)} \leqslant C_\kappa h\rho^{\frac{N-1}{2}}+\sum_{j=0}^{J} Cd_j^{\frac{N-3}{2}}\|v-R_hv\|_{L^2(D_j^1)}, \qquad (4.3.59)$$

and therefore

$$\rho^{-\frac{N}{2}}h^{-1}\|\nabla(v-R_hv)\|_{L^1(\Lambda_h)} \le C\rho^{-\frac{N}{2}}h^{-\frac{1}{2}}\sum_{j=0}^{J} d_j^{\frac{N-1}{2}}\|\nabla(v-R_hv)\|_{L^2(D_j)} + C_\kappa$$
$$\le C_\kappa + C\rho^{-\frac{N}{2}}h^{-\frac{1}{2}}\sum_{j=0}^{J} d_j^{\frac{N-3}{2}}\|v-R_hv\|_{L^2(D_j^1)}. \tag{4.3.60}$$

It remains to estimate $\sum_{j=0}^{J} d_j^{\frac{N-3}{2}}\|v-R_hv\|_{L^2(D_j^1)}$. To this end, we let χ be a smooth cut-off function satisfying

$$\chi = 1 \text{ on } D_j^1, \quad \chi = 0 \text{ outside } D_j^2 \quad \text{and} \quad |\nabla\chi| \leqslant Cd_j^{-1}.$$

For $N = 2, 3$ the following Sobolev interpolation inequality holds:

$$\|\chi(v-R_hv)\|_{L^2(\Omega)} \leqslant \|\chi(v-R_hv)\|_{L^p(\Omega)}^{1-\theta}\|\chi(v-R_hv)\|_{H^1(\Omega)}^{\theta} \quad \text{with } \frac{1}{2} = \frac{1-\theta}{p} + \frac{\theta}{p_*}, \tag{4.3.61}$$

where $p_* = \infty$ for $N = 2$ and $p_* = 6$ for $N = 3$. For both $N = 2$ and $N = 3$, the parameter θ determined by (4.3.61) satisfies the following relation:

$$\frac{N}{p} - \frac{N}{2} = \frac{\theta}{1-\theta}. \tag{4.3.62}$$

We can choose p sufficiently close to 2 as mentioned below (4.3.58). Since

$$C\|\chi(v-R_hv)\|_{H^1(\Omega)} \leqslant C\|\nabla(v-R_hv)\|_{L^2(D_j^2)} + Cd_j^{-1}\|v-R_hv\|_{L^2(D_j^2)} \tag{4.3.63}$$

it follows that

$$\|v-R_hv\|_{L^2(D_j^1)}$$
$$\leqslant \|v-R_hv\|_{L^p(D_j^2)}^{1-\theta}\big(C\|\nabla(v-R_hv)\|_{L^2(D_j^2)} + Cd_j^{-1}\|v-R_hv\|_{L^2(D_j^2)}\big)^{\theta}$$
$$= (\epsilon^{-\frac{\theta}{1-\theta}}\|v-R_hv\|_{L^p(D_j^2)})^{1-\theta}\big(C\epsilon\|\nabla(v-R_hv)\|_{L^2(D_j^2)} + C\epsilon d_j^{-1}\|v-R_hv\|_{L^2(D_j^2)}\big)^{\theta}$$
$$\leqslant C\epsilon^{-\frac{\theta}{1-\theta}}\|v-R_hv\|_{L^p(D_j^2)} + C\epsilon\|\nabla(v-R_hv)\|_{L^2(D_j^2)} + C\epsilon d_j^{-1}\|v-R_hv\|_{L^2(D_j^2)},$$

where ϵ can be an arbitrary positive number.

By choosing $\epsilon = d_j(\rho/d_j)^\sigma$ with a fixed $\sigma \in (0,1)$, we obtain

$$\|v-R_hv\|_{L^2(D_j^1)} \leqslant C\left(\frac{\rho}{d_j}\right)^{-\frac{\theta\sigma}{1-\theta}} d_j^{-\frac{\theta}{1-\theta}}\|v-R_hv\|_{L^p(D_j^1)} \tag{4.3.64}$$
$$+\left(\frac{\rho}{d_j}\right)^{\sigma}\big(Cd_j\|\nabla(v-R_hv)\|_{L^2(D_j^2)} + C\|v-R_hv\|_{L^2(D_j^2)}\big).$$

Hence,

$$\rho^{-\frac{N}{2}}h^{-\frac{1}{2}}\sum_{j=0}^{J} d_j^{\frac{N-3}{2}}\|v-R_hv\|_{L^2(D_j^1)}$$

$$
\begin{aligned}
&\leqslant C\rho^{-\frac{N}{2}}h^{-\frac{1}{2}}\sum_{j=0}^{J}\left(\frac{\rho}{d_j}\right)^{-\frac{\theta\sigma}{1-\theta}} d_j^{-\frac{\theta}{1-\theta}+\frac{N-3}{2}}\|v-R_hv\|_{L^p(D_j^2)}\\
&\quad + C\rho^{-\frac{N}{2}}h^{-\frac{1}{2}}\sum_{j=0}^{J}\left(\frac{\rho}{d_j}\right)^{\sigma}\big(d_j^{\frac{N-1}{2}}\|\nabla(v-R_hv)\|_{L^2(D_j^2)}+Cd_j^{\frac{N-3}{2}}\|v-R_hv\|_{L^2(D_j^2)}\big)\\
&\leqslant C\rho^{-\frac{N}{2}}h^{-\frac{1}{2}}\sum_{j=0}^{J}\left(\frac{\rho}{d_j}\right)^{-\frac{\theta\sigma}{1-\theta}} d_j^{-\frac{\theta}{1-\theta}+\frac{N-3}{2}}\|v-R_hv\|_{L^p(D_j^2)}\\
&\quad + C_\kappa + C\rho^{-\frac{N}{2}}h^{-\frac{1}{2}}\left(\frac{\rho}{d_J}\right)^{\sigma}\sum_{j=0}^{J} d_j^{\frac{N-3}{2}}\|v-R_hv\|_{L^2(D_j^3)},
\end{aligned}
\tag{4.3.65}
$$

where we have used (4.3.59) and the fact $\frac{\rho}{d_j}\le\frac{\rho}{d_J}$ in deriving the last inequality. Note that

$$
\sum_{j=0}^{J} d_j^{\frac{N-3}{2}}\|v-R_hv\|_{L^2(D_j^3)}\leqslant Cd_J^{\frac{N-3}{2}}\|v-R_hv\|_{L^2(S_{\kappa\rho}(x_0))}+3\sum_{j=0}^{J} d_j^{\frac{N-3}{2}}\|v-R_hv\|_{L^2(D_j^1)}.
$$

Combining the last two estimates, we obtain

$$
\begin{aligned}
\rho^{-\frac{N}{2}}h^{-\frac{1}{2}}\sum_{j=0}^{J} d_j^{\frac{N-3}{2}}\|v-R_hv\|_{L^2(D_j^1)} &\leqslant C\rho^{-\frac{N}{2}}h^{-\frac{1}{2}}\sum_{j=0}^{J}\left(\frac{\rho}{d_j}\right)^{-\frac{\theta\sigma}{1-\theta}} d_j^{-\frac{\theta}{1-\theta}+\frac{N-3}{2}}\|v-R_hv\|_{L^p(D_j^2)}\\
&\quad + C_\kappa + C\rho^{-\frac{N}{2}}h^{-\frac{1}{2}}\left(\frac{\rho}{d_J}\right)^{\sigma} d_J^{\frac{N-3}{2}}\|v-R_hv\|_{L^2(S_{\kappa\rho}(x_0))}\\
&\quad + C\rho^{-\frac{N}{2}}h^{-\frac{1}{2}}\left(\frac{\rho}{d_J}\right)^{\sigma}\sum_{j=0}^{J} d_j^{\frac{N-3}{2}}\|v-R_hv\|_{L^2(D_j^1)}.
\end{aligned}
$$

For the fixed $\sigma\in(0,1)$, by choosing a sufficiently large parameter κ we have $\left(\frac{\rho}{d_J}\right)^{\sigma}\le\frac{C}{\kappa^\sigma}$, and therefore the last term of the inequality above can be absorbed by the left-hand side. From now on we fix the parameter κ. Then we have

$$
\begin{aligned}
\sum_{j=0}^{J}\rho^{-\frac{N}{2}}h^{-\frac{1}{2}}d_j^{\frac{N-3}{2}}\|v-R_hv\|_{L^2(D_j^1)} &\leqslant \sum_{j=0}^{J} C\rho^{-\frac{N}{2}}h^{-\frac{1}{2}}\left(\frac{\rho}{d_j}\right)^{-\frac{\theta\sigma}{1-\theta}} d_j^{-\frac{\theta}{1-\theta}+\frac{N-3}{2}}\|v-R_hv\|_{L^p(D_j^2)}\\
&\quad + C_\kappa + C\rho^{-\frac{N}{2}}h^{-\frac{1}{2}}\left(\frac{\rho}{d_J}\right)^{\sigma} d_J^{\frac{N-3}{2}}\|v-R_hv\|_{L^2(S_{\kappa\rho}(x_0))}.
\end{aligned}
\tag{4.3.66}
$$

It remains to estimate $\|v-R_hv\|_{L^p(D_j^1)}$ and $\|v-R_hv\|_{L^2(S_{\kappa\rho}(x_0))}$. This is done by applying Lemma 4.3.19 (with $q'=p$ therein), Lemma 4.3.18 (with $q=p$ therein) and Hölder's inequality, i.e.,

$$
\|v-R_hv\|_{L^p(\Omega)}\leqslant Ch^2\|\varphi\|_{L^p(\Omega)}\leqslant Ch^2\rho^{\frac{N}{p}-\frac{N}{2}}, \tag{4.3.67}
$$

$$
\|v-R_hv\|_{L^2(\Omega)}\leqslant Ch^2 \quad \text{(setting } q'=q=2 \text{ in Lemma 4.3.19 and Lemma 4.3.18).} \tag{4.3.68}
$$

Then, substituting these estimates into (4.3.66), we obtain

$$\sum_{j=0}^{J}\rho^{-\frac{N}{2}}h^{-\frac{1}{2}}d_j^{\frac{N-3}{2}}\|v-R_hv\|_{L^2(D_j^1)} \leqslant \sum_{j=0}^{J}C\left(\frac{h}{\rho}\right)^{\frac{N}{2}}\left(\frac{h}{d_j}\right)^{\frac{3-N}{2}}\left(\frac{\rho}{d_j}\right)^{\frac{N}{p}-\frac{N}{2}-\frac{\theta\sigma}{1-\theta}}d_j^{\frac{N}{p}-\frac{N}{2}-\frac{\theta}{1-\theta}}$$
$$+C_\kappa + C\left(\frac{h}{\rho}\right)^{\frac{N}{2}}\left(\frac{h}{d_J}\right)^{\frac{3-N}{2}}\left(\frac{\rho}{d_J}\right)^{\sigma} \qquad (4.3.69)$$

By choosing $p<2$ to be sufficiently close to 2 (so that $q'=p$ satisfies the condition of Lemma 4.3.19) and using the relation $\frac{N}{p}-\frac{N}{2}=\frac{\theta}{1-\theta}$ as shown in (4.3.62), we obtain

$$\sum_{j=0}^{J}\rho^{-\frac{N}{2}}h^{-\frac{1}{2}}d_j^{\frac{N-3}{2}}\|v-R_hv\|_{L^2(D_j^1)} \leqslant C. \qquad (4.3.70)$$

Then, substituting the last inequality into the right-hand side of (4.3.60), we obtain

$$\rho^{-\frac{N}{2}}h^{-1}\|\nabla(v-R_hv)\|_{L^1(\Lambda_h)} \leqslant C.$$

This proves (4.3.26) for sufficiently small mesh size, say $h\leqslant h_0$. This condition is required when we use Corollary 4.3.14, Lemma 4.3.18 and Lemma 4.3.19 in this subsection.

In the case $h\geqslant h_0$, we denote by $\tilde g_h\in S_h(\Omega_h)$ the isoparametric finite element function satisfying $\tilde g_h=u_h$ on $\partial\Omega_h$ and $\tilde g_h=0$ at the interior nodes of the domain Ω_h. Then the following estimate holds:

$$\|\tilde g_h\|_{L^\infty(\Omega_h)} \leqslant C\|u_h\|_{L^\infty(\partial\Omega_h)}.$$

Since $\chi_h=u_h-\tilde g_h\in S_h^\circ(\Omega_h)$, it follows from (4.1.1) that

$$0=\int_{\Omega_h}\nabla u_h\cdot\nabla(u_h-\tilde g_h)=\|\nabla(u_h-\tilde g_h)\|_{L^2(\Omega_h)}^2+\int_{\Omega_h}\nabla\tilde g_h\cdot\nabla(u_h-\tilde g_h),$$

and therefore

$$\|\nabla(u_h-\tilde g_h)\|_{L^2(\Omega_h)}^2=-\int_{\Omega_h}\nabla\tilde g_h\cdot\nabla(u_h-\tilde g_h)\leqslant C\|\nabla\tilde g_h\|_{L^2(\Omega_h)}\|\nabla(u_h-\tilde g_h)\|_{L^2(\Omega_h)}.$$

Thus, by using the inverse inequality and the condition $h\geqslant h_0$, we have

$$\|\nabla(u_h-\tilde g_h)\|_{L^2(\Omega_h)}\leqslant C\|\nabla\tilde g_h\|_{L^2(\Omega_h)}\leqslant Ch^{-1}\|\tilde g_h\|_{L^2(\Omega_h)}\leqslant Ch_0^{-1}\|\tilde g_h\|_{L^\infty(\Omega_h)}$$
$$\leqslant Ch_0^{-1}\|u_h\|_{L^\infty(\partial\Omega_h)}.$$

By using the inverse inequality again, we obtain

$$\begin{aligned}\|u_h-\tilde g_h\|_{L^\infty(\Omega_h)}&\leqslant Ch^{-\frac{N}{2}}\|u_h-\tilde g_h\|_{L^2(\Omega_h)}\\&\leqslant Ch^{-\frac{N}{2}}\|\nabla(u_h-\tilde g_h)\|_{L^2(\Omega_h)}\\&\leqslant Ch_0^{-\frac{N}{2}-1}\|u_h\|_{L^\infty(\partial\Omega_h)}.\end{aligned}$$

By the triangle inequality, this proves

$$\|u_h\|_{L^\infty(\Omega_h)}\leqslant\|\tilde g_h\|_{L^\infty(\Omega_h)}+\|u_h-\tilde g_h\|_{L^\infty(\Omega_h)}\leqslant C\|u_h\|_{L^\infty(\partial\Omega_h)}$$

for $h\geqslant h_0$.

Combining the two cases $h\leqslant h_0$ and $h\geqslant h_0$, we obtain the result of Theorem 4.2.1. □

4.4 Proof of Theorem 4.2.2

In this section, we adapt Schatz's argument in [125] to the proof of maximum-norm stability of isoparametric finite element solutions of the Poisson equation in the curvilinear polyhedron considered here. The argument is based on the weak maximum principle established in Theorem 4.2.1 and the following technical result, which asserts that the $W^{1,\infty}$ regularity estimate of the Poisson equation can hold in a family of larger perturbed domains Ω^t, $t \in [0, \delta]$, such that $\mathrm{dist}(\partial\Omega^t, \partial\Omega) \sim t$ and the $W^{1,\infty}$ estimate is uniformly with respect to $t \in [0, \delta]$.

Remark 4.4.1. Here we make a remark on the idea of our proof. To prove Theorem 4.2.2, we observe that the numerical solution u_h is in fact the Ritz projection of $u^{(h)} \in H_0^1(\Omega_h)$ which is the exact solution of the Poisson equation on Ω_h:

$$-\Delta u^{(h)} = f \quad \text{in } \Omega_h \quad (f \text{ is extended by zero outside } \Omega),$$

in the sense that

$$R_h(u^{(h)} \circ \Phi_h^{-1}) = u_h \circ \Phi_h^{-1}.$$

Using the weak maximum principle established in Theorem 4.2.1, one can imitate the proof of [96, Theorem 5.1] to show that there holds L^∞ stability for our Ritz projection R_h. It follows that

$$\|u^{(h)} - u_h\|_{L^\infty(\Omega_h)} \le C\|u^{(h)} - I_h u^{(h)}\|_{L^\infty(\Omega_h)}.$$

Now we can obtain the result of Theorem 4.2.2 as long as we establish the estimate

$$\|u - u^{(h)}\|_{L^\infty(\Omega_h)} \le Ch^{r+1}\|f\|_{L^p(\Omega)} \quad (p > N),$$

where we have extended u by zero outside Ω. To this end, we consider employing the maximum principle of harmonic functions since $\Delta(u^{(h)} - u) = 0$ in $\Omega \cap \Omega_h$. Here technically we introduce larger perturbed domain Ω^t and solution u^t

$$-\Delta u^t = f \quad \text{in } \Omega^t,$$

in the larger perturbed domain Ω^t. Then using maximum principle, we compare u and $u^{(h)}$ with u^t respectively, for example we have

$$\|u - u^t\|_{L^\infty(\Omega)} \le \|u^t\|_{L^\infty(\partial\Omega)} \le Ch^{r+1}\|u^t\|_{W^{1,\infty}(\Omega^t)}.$$

This explains the motivation of establishing Proposition 4.4.1.

Proposition 4.4.1. *Let Ω be a curvilinear polyhedron with edge openings smaller than π, and define*

$$\Omega(\varepsilon) := \{x \in \mathbb{R}^N \, : \, \mathrm{dist}(x, \Omega) < \varepsilon\},$$

which is an ε neighborhood of Ω. Then there exist constants $\delta > 0$ and $\lambda > 0$ and a family of larger bounded domains Ω^t satisfying

$$\Omega(\lambda t) \subseteq \Omega^t \subseteq \Omega(\lambda^{-1} t) \quad \forall t \in [0, \delta],$$

such that the weak solution $u^t \in H_0^1(\Omega^t)$ of the Poisson equation

$$-\Delta u^t = f \quad \text{in } \Omega^t, \quad \text{with } f \in L^p(\Omega^t) \text{ for some } p > N, \tag{4.4.1}$$

satisfies the following estimate:

$$\|u^t\|_{W^{1,\infty}(\Omega^t)} \leq C_p \|f\|_{L^p(\Omega^t)} \quad \text{for } t \in [0, \delta], \tag{4.4.2}$$

where C_p is some constant which is independent of $t \in [0, \delta]$.

Proof. In a standard convex polyhedron $\hat{\Omega}$, the following estimate holds for $p > N$ (cf. [104, Lemma 2.1]):

$$\|\nabla w\|_{L^\infty(\hat{\Omega})} \leq C_p \|\nabla \cdot (a\nabla w)\|_{L^p(\hat{\Omega})} \quad \forall\, w \in H_0^1(\hat{\Omega}) \text{ such that } \nabla \cdot (a\nabla w) \in L^2(\hat{\Omega}). \tag{4.4.3}$$

where $a = (a_{ij})$ is any symmetric positive definite matrix in $W^{1,q}(\hat{\Omega})$ with $q > N$, satisfying the following estimate:

$$C^{-1}|\xi|^2 \leqslant a\xi \cdot \xi \leqslant C|\xi|^2. \tag{4.4.4}$$

On the curvilinear polyhedron Ω considered in this chapter, by using a partition of unity we can reduce the problem to an open subset of Ω which is diffeomorphic to a convex polyhedral cone. Therefore, the following result still holds for $p > N$:

$$\|\nabla w\|_{L^\infty(\Omega)} \leq C_p \|\nabla \cdot (a\nabla w)\|_{L^p(\Omega)} \quad \forall\, w \in H_0^1(\Omega) \text{ such that } \nabla \cdot (a\nabla w) \in L^2(\Omega). \tag{4.4.5}$$

If there exists a smooth diffeomorphism $\Psi_t : \Omega \to \Omega^t$ (smooth uniformly with respect to $t \in [0, \delta]$), then we can pull the Poisson equation on $\Omega^t = \Psi_t(\Omega)$ back to the curvilinear polyhedron Ω as an elliptic equation with some coefficient matrix a satisfying (4.4.4), and then use the result in (4.4.5). This would prove (4.4.2). If the partial derivatives of the diffeomorphism from Ω to Ω^t can be uniformly bounded with respect to $t \in [0, \delta]$, then the constant in (4.4.2) is independent of $t \in [0, \delta]$.

It remains to prove the existence of a smooth diffeomorphism $\Psi_t : \Omega \to \Omega^t = \Psi_t(\Omega)$. This is presented in the following lemma. ∎

Lemma 4.4.2. *Let Ω be a curvilinear polyherdon. Then there exist constants $\delta > 0$ and $\lambda > 0$ (which only depend on Ω), and a family of diffeomorphisms $\Psi_t : \mathbb{R}^N \to \mathbb{R}^N$ for $t \in [0, \delta]$, such that*

1. *$\Omega(\lambda t) \subseteq \Psi_t(\Omega) \subseteq \Omega(\lambda^{-1} t)$ for $t \in [0, \delta]$ and some constant $\lambda > 0$.*
2. *The partial derivatives of Ψ_t are bounded uniformly with respect to $t \in [0, \delta]$, i.e.,*

$$|\nabla^k \Psi_t(x)| \leqslant C_k \quad \forall x \in \mathbb{R}^N, \ \forall k \geqslant 1, \ \text{where } C_k \text{ is independent of } t \in [0, \delta].$$

Proof. It is known that any given smooth and compactly supported vector field X on $\mathbb{R}$ induces a flow map

$$\Psi : \mathbb{R} \times \mathbb{R}^N \to \mathbb{R}^N \quad (t, x) \mapsto \Phi(t, x),$$

such that each $\Psi_t = \Psi(t, \cdot) : \mathbb{R}^N \to \mathbb{R}^N$ is a diffeomorphism of $\mathbb{R}^N$ for sufficiently small t, say $|t| \leqslant \delta$. Moreover, $\Psi_0 = \mathrm{Id}$, $\Psi_{t+s} = \Psi_t \circ \Psi_s$ for $t, s \in \mathbb{R}$, and the partial derivatives of Ψ_t are uniformly bounded by constants which only depend on X and δ (independent of t).

Therefore, in order to prove Lemma 4.4.2, it suffices to construct a compactly supported smooth vector field X, such that the flow map induced by X satisfies $\Omega(\lambda t) \subseteq \Psi_t(\Omega) \subseteq \Omega(\lambda^{-1}t)$ for $t \in [0, \delta]$ (with some constants $\lambda > 0$ and $\delta > 0$). This can be proved by utilizing the following result, which provides a criteria for the construction of such a vector field.

Lemma 4.4.3. *Let Ω be a curvilinear polyhedron, and let X be a smooth and compactly supported vector field on $\mathbb{R}^N$ satisfying the following conditions:*

1. $X|_{\Omega'} \equiv 0$ *for some nonempty open subset* $\Omega' \subset\subset \Omega$.
2. $\langle X(x), N_x \rangle \geqslant c$ *at all smooth points* $x \in \partial\Omega$, *where* N_x *denotes the unit outward normal vector at* $x \in \partial\Omega$ *and* $c > 0$ *is some constant.*
3. $|X(x)| \leqslant 1 \quad \forall x \in \mathbb{R}^N$

Then there are constants $\lambda > 0$ and $\delta > 0$, which only depend on X and Ω, such that the flow map Ψ_t induced by the vector field X has the following property:

$$\Omega(\lambda t) \subseteq \Psi_t(\Omega) \subseteq \Omega(\lambda^{-1}t) \quad \textit{for } t \in [0, \delta].$$

Let us temporarily assume that Lemma 4.4.3 holds, and use it to prove Lemma 4.4.2. To this end, it suffices to construct a vector field which satisfies the conditions in Lemma 4.4.3.

From the definition of the curvilinear polyhedron we know that for every $x \in \partial\Omega$ there exists a map $\varphi_x : U_x \to B_{\varepsilon_x}(0)$ which is a diffeomorphism from a neighborhood U_x of x in $\mathbb{R}^N$ to a ball centered at 0 with radius ε_x, such that $\varphi_x(x) = 0$ and $\varphi_x(U_x \cap \Omega) = K_x \cap B_0(\varepsilon_x)$, where $K_x = \{y \in \mathbb{R}^3 : y/|y| \in \Theta\}$ is a cone corresponding to a spherical region $\Theta \subset \mathbb{S}^2$ which is contained in an open half sphere, say $\mathbb{S}^2_+ = \{x \in \mathbb{R}^3 : |x| = 1,\ x_3 > 0\}$. We shall use the following terminology:

1. By composing φ_x with an additional linear transformation if necessary, we can assume that $\nabla\varphi_x(x) = I$ (which holds only at the point x in U_x).
2. If p is a smooth point on ∂K_x (not on the edges or vertex of ∂K_x), then we denote by $N_{x,p}$ the unit outward normal vector of ∂K_x at p, and define
$$\hat{N}_x = \big\{N_{x,p} : p \text{ in some smooth piece of } \partial K_x\big\}$$
to be the set of all outward unit normal vectors on the smooth faces of ∂K_x. When x is a smooth point of $\partial\Omega$, $\hat{N}_x$ consists of only one vector, i.e., the usual unit normal vector N_x. Therefore, the set $\hat{N}_x$ can be viewed as generalization of normal vector at x when x is not a smooth point.
3. Let y be an interior point in the polyhedral cone K_x. Then the unit vector $V_x = -y/|y|$ satisfies that $\langle V_x, N_{x,p} \rangle > 0$ for all $N_{x,p} \in \hat{N}_x$.

We will construct a smooth vector field X on $\mathbb{R}^N$ as follows, by using a partition of unity. By the three properties above and the compactness of $\partial\Omega$, there is constant $c > 0$ only dependent on Ω such that for each $x \in \partial\Omega$, there is a unit vector $V_x \in \mathbb{R}^N$ such that

$$\langle V_x, N_{x,p} \rangle \geqslant 2c \quad \forall N_{x,p} \in \hat{N}_x.$$

Since the normal vector at a smooth point of $\partial\Omega$ changes continuously in a smooth piece of $\partial\Omega$, one can shrink the neighborhood U_x of $x \in \partial\Omega$ so that

$$\langle V_x, N_y\rangle \geqslant c \text{ for all smooth points } y \in \partial\Omega \cap U_x,$$

where N_y denotes the unit outward normal vector at $y \in \partial\Omega \cap U_x$. We define a smooth vector field X_x on U_x by

$$X_x(y) = V_x \quad \forall y \in U_x,$$

and choose a finite covering $\{U_{x_\ell}\}_{1\leqslant \ell\leqslant L}$ of $\partial\Omega$ from these U_x, $x \in \partial\Omega$, and a family of smooth cut-off functions $\{\chi_\ell\}_{1\leqslant\ell\leqslant L}$ such that $0 \le \chi_\ell \le 1$ and

$$\operatorname{supp}(\chi_\ell) \subseteq U_{x_\ell} \text{ and } \sum_{1\le\ell\le L} \chi_\ell(x) = 1, \quad \forall x \in \partial\Omega.$$

Then we denote by X_{x_ℓ} the above-mentioned vector field defined on U_{x_ℓ}, and define

$$X = \sum_{\ell=1}^{L} \chi_\ell X_{x_\ell},$$

so that X is a compactly supported smooth vector field such that

$$\langle X(y), N_y\rangle = \sum_{\chi_\ell(y)\neq 0} \chi_\ell(y)\langle X_{x_\ell}, N_y\rangle \ge c, \quad \text{for all smooth point } y \in \partial\Omega.$$

and clearly $|X(x)| \le 1$, $\forall x \in \mathbb{R}^N$. This proves the existence of a desired vector field X, and therefore completes the proof of Proposition 4.4.1. ∎

Proof of Lemma 4.4.3. For each $x \in \partial\Omega$, let $\varphi_x : U_x \to B_{\varepsilon_x}(0)$ be the map as in the definition of the curvilinear polyhedron. Here we do not require $\varphi_x(U_x)$ to be a ball so that we can assume U_x to be convex.

By composing φ_x with an additional linear transformation if necessary, we can assume that $\nabla\varphi_x(x) = I$ (as in the proof of Lemma 4.4.2). Since $c \leqslant \langle X(x), N_x\rangle \leqslant 1$ (as a the condition in Lemma 4.4.3), we can shrink the neighborhood U_x small enough so that

$$\frac{c}{2} \leqslant \langle(\nabla\varphi_x(y))^\top X(y), N_{x,p}\rangle \leqslant 2 \quad \forall y \in U_x,\ p \in \varphi_x(U_x \cap \partial\Omega) = \varphi_x(U_x) \cap \partial K_x,$$
$$p \text{ is a smooth point.} \quad (4.4.6)$$

Moreover, since $(\nabla\varphi_x)^\top = I$ at x, we can shrink U_x so that the following equivalence relation holds:

$$d(y_1, y_2) \sim d(\varphi_x(y_1), \varphi_x(y_2)) \quad \forall y_1, y_2 \in U_x,$$

where $d(\cdot,\cdot)$ denotes the Euclidean distance in $\mathbb{R}^N$. As a result,

$$d(y, U_x \cap \Omega) \sim d(\varphi_x(y), \varphi_x(U_x \cap \Omega)) \quad \forall y \in U_x.$$

We can choose a finite covering $\{U_{x_\ell}\}_{1\leqslant\ell\leqslant L}$ of $\partial\Omega$ from these U_x. Then there exists a sufficiently small $\delta > 0$ such that for any $x \in \partial\Omega$ there exists $1 \leqslant \ell \leqslant L$ such that for all $t \in [0, \delta]$,

$$\Psi_t(x) \in U_{x_\ell} \text{ for some } 1 \leqslant \ell \leqslant L.$$

Moreover,

$$d(\Psi_t(x), \Omega) = d(\Psi_t(x), U_\ell \cap \Omega) \tag{4.4.7}$$

and

$$d(\Psi_t(x), \Omega) \sim d(\varphi_{x_\ell}(\Psi_t(x)), \varphi_{x_\ell}(U_\ell \cap \Omega)). \tag{4.4.8}$$

Let $Y_\ell = (\nabla \varphi_{x_\ell})^\top X|_{U_\ell}$ be the pushforward vector field under φ_{x_ℓ}, then $\varphi_{x_\ell}(\Psi_t(x))$ is the integral curve of vector field Y_ℓ, with initial value point $\varphi_{x_\ell}(x)$. From (4.4.6) we know that

$$\frac{c}{2} \leqslant \langle Y_\ell(z), N_{x_\ell,p} \rangle \leqslant 2 \quad \forall z \in \varphi_{x_\ell}(U_{x_\ell}),\ \ \forall p \in \varphi_{x_\ell}(U_{x_\ell} \cap \partial\Omega) = \varphi_{x_\ell}(U_{x_\ell}) \cap \partial K_{x_\ell},$$

which implies that the integral curve $\varphi_{x_\ell}(\Psi_t(x))$ is flowing outside $\varphi_{x_\ell}(U_{x_\ell} \cap \Omega)$, i.e.,

$$\frac{ct}{2} \leqslant d(\varphi_{x_\ell}(\Psi_t(x)), \varphi_{x_\ell}(U_{x_\ell} \cap \Omega)) \leqslant 2t.$$

Then, from the equivalence of distance as shown in (4.4.7)–(4.4.8), we conclude that there exists a constant $\lambda > 0$ such that

$$2\lambda t \leqslant d(\Psi_t(x), \Omega) \leqslant \frac{1}{2}\lambda^{-1} t \quad \forall t \in [0, \delta],\ \ \forall x \in \partial\Omega.$$

We consider the domain $\Omega(\lambda t) := \{x \in \mathbb{R}^N \ : \ \mathrm{dist}(x, \Omega) < \lambda t\} \supset \Omega$. On the one hand, since $X|_{\Omega'} = 0$ for some subdomain $\Omega' \subset\subset \Omega$ it follows that $\Psi_t(\Omega) \cap \Omega(\lambda t) \neq \emptyset$. On the other hand, since $d(\Psi_t(x), \Omega) > \lambda t$ for all $x \in \partial\Omega$, the boundaries of $\Psi_t(\Omega)$ and $\Omega(\lambda t)$ are disjoint. It follows that $\Omega(\lambda t) \subseteq \Psi_t(\Omega)$ for $t \in [0, \delta]$. Similarly, one can prove that $\Omega(\lambda^{-1} t) \supset \Psi_t(\Omega)$. This completes the proof of Lemma 4.4.3. ∎

Lemma 4.4.4. *Let Ω^t be the domain in Proposition 4.4.1, satisfying $\Omega(\lambda t) \subseteq \Omega^t \subseteq \Omega(\lambda^{-1} t)$ for $t \in [0, \delta]$, with $\Omega(\lambda t) = \{x \in \mathbb{R}^N \ : \ \mathrm{dist}(x, \Omega) < \lambda t\}$. Suppose that $f \in L^p(\Omega^t)$ for some $p > N$, and $\Omega_h \subset \Omega^t$ for some $t = O(h^{r+1})$ and $h \leqslant h_1$, where $h_1 > 0$ is some constant. Let $u \in H_0^1(\Omega)$ and $u^{(h)} \in H_0^1(\Omega_h)$ be the weak solutions of the following PDE problems*:

$$\begin{aligned} -\Delta u &= f \quad \text{in } \Omega, \\ -\Delta u^{(h)} &= f \quad \text{in } \Omega_h, \end{aligned}$$

and extend u and $u^{(h)}$ by zero to the larger domain Ω^t. Then there exists $h_2 > 0$ such that for $h \leqslant h_2$ the following estimate holds:

$$\|u - u^{(h)}\|_{L^\infty(\Omega^t)} \leq C h^{r+1} \|f\|_{L^p(\Omega^t)} \tag{4.4.9}$$

Proof. Since $\max_{x \in \Omega_h} |\Phi_h(x) - x| \leq C_0 h^{r+1}$ for some constant C_0, it follows that $\Omega_h \subset \Omega(C_0 h^{r+1}) \subset \Omega^t$ for $t = C_0 \lambda^{-1} h^{r+1}$. When h is sufficiently small we have $t = C_0 \lambda^{-1} h^{r+1} \leqslant \delta$ and therefore Ω^t is well defined. Let $u^t \in H_0^1(\Omega^t)$ be a weak solution of the Poisson equation

$$-\Delta u^t = f \quad \text{in } \Omega^t.$$

Proposition 4.4.1 implies that

$$\|u^t\|_{W^{1,\infty}(\Omega^t)} \leq C\|f\|_{L^p(\Omega^t)}. \tag{4.4.10}$$

Since $u^t - u$ is harmonic in $\Omega \subset \Omega^t$ and $u^t - u^{(h)}$ is harmonic in $\Omega_h \subset \Omega^t$, the maximum principle of the continuous problem implies that

$$\begin{aligned}
\|u^t - u^{(h)}\|_{L^\infty(\Omega_h)} &\leq \|u^t - u^{(h)}\|_{L^\infty(\partial\Omega_h)} \\
&= \|u^t\|_{L^\infty(\partial\Omega_h)} \qquad\qquad (\text{since } u^{(h)} = 0 \text{ on } \partial\Omega_h) \\
&\leq Ch^{r+1}\|u^t\|_{W^{1,\infty}(\Omega^t)} \\
&\leq Ch^{r+1}\|f\|_{L^p(\Omega^t)},
\end{aligned} \tag{4.4.11}$$

where we have used the fact that $\mathrm{dist}(x, \partial\Omega^t) \leq 2C_0h^{r+1}$ for $x \in \partial\Omega_h$. Therefore,

$$\begin{aligned}
\|u^t - u^{(h)}\|_{L^\infty(\Omega^t)} &\leq \|u^t - u^{(h)}\|_{L^\infty(\Omega_h)} + \|u^t\|_{L^\infty(\Omega^t\backslash\Omega_h)} \\
&\leq Ch^{r+1}\|f\|_{L^p(\Omega^t)} + Ch^{r+1}\|u^t\|_{W^{1,\infty}(\Omega^t)} \\
&\leq Ch^{r+1}\|f\|_{L^p(\Omega^t)}.
\end{aligned} \tag{4.4.12}$$

The following result can be proved in the same way:

$$\|u^t - u\|_{L^\infty(\Omega^t)} \leq Ch^{r+1}\|f\|_{L^p(\Omega^t)}. \tag{4.4.13}$$

The result of Lemma 4.4.4 follows from (4.4.12)–(4.4.13) and the triangle inequality. ∎

In the following, we prove Theorem 4.2.2 by using the technical result in Proposition 4.4.1.

Let Ω^t be the domain in Proposition 4.4.1, satisfying $\Omega(\lambda t) \subseteq \Omega^t \subseteq \Omega(\lambda^{-1}t)$ for $t \in [0, \delta]$, with $\Omega(\lambda t) = \{x \in \mathbb{R}^N \ : \ \mathrm{dist}(x, \Omega) < \lambda t\}$. For the simplicity of notation, we still denote by $f \in L^p(\Omega^t)$ an extension of $\tilde{f} \in L^p(\Omega \cup \Omega_h)$ satisfying $\|f\|_{L^p(\Omega^t)} \leqslant C\|\tilde{f}\|_{L^p(\Omega\cup\Omega_h)} \leqslant C\|f\|_{L^p(\Omega)}$.

Under assumption 4.2.1, the curvilinear polyhedral domain Ω can be extended to a larger convex polyhedron Ω_* with a piecewise flat boundary such that $\overline{\Omega} \subset \Omega_*$ and the triangulation $\mathcal{T}_h$ can be extended to a quasi-uniform triangulation $\mathcal{T}_{*,h}$ on Ω_* (thus the triangulation in $\Omega_*\backslash\overline{\Omega}$ is also isoparametric on its boundary $\partial\Omega$).

Let $\tilde{u}$ be an extension of $u^{(h)}$ such that $\tilde{u} = u^{(h)}$ on Ω_h and $\tilde{u} = 0$ in $\Omega_*\backslash\Omega_h$. Let $S_h^\circ(\Omega_*) \subset H_0^1(\Omega_*)$ be the H^1-conforming isoparametric finite element space on Ω_* with triangulation $\mathcal{T}_{*,h}$. Let $\tilde{u}_h \in S_h^\circ(\Omega_*)$ be the Ritz projection of $\tilde{u}$ defined by

$$\int_{\Omega_*} \nabla(\tilde{u} - \tilde{u}_h) \cdot \nabla\chi_h = 0 \quad \forall\, \chi_h \in S_h^\circ(\Omega_*).$$

Then

$$\begin{aligned}
\|u^{(h)} - u_h\|_{L^\infty(\Omega_h)} &= \|\tilde{u} - u_h\|_{L^\infty(\Omega_h)} \\
&\leqslant \|\tilde{u} - \tilde{u}_h\|_{L^\infty(\Omega_h)} + \|\tilde{u}_h - u_h\|_{L^\infty(\Omega_h)} \\
&\leqslant \|\tilde{u} - \tilde{u}_h\|_{L^\infty(\Omega_*)} + \|\tilde{u}_h - u_h\|_{L^\infty(\Omega_h)},
\end{aligned} \tag{4.4.14}$$

where $\|\tilde{u} - \tilde{u}_h\|_{L^\infty(\Omega_*)}$ is the error of the Ritz projection of an H^1-conforming FEM in a standard convex polyhedron and therefore can be estimated by using the result on a standard convex polyhedron (or using the interior maximum-norm estimate as in [127, Theorem 5.1] and [96, Proof of Theorem 5.1]), i.e.,

$$
\begin{aligned}
\|\tilde{u}-\tilde{u}_h\|_{L^\infty(\Omega_*)} &\leqslant C\ell_h\|\tilde{u}-I_h\tilde{u}\|_{L^\infty(\Omega_*)}\\
&\leqslant C\ell_h\|u^{(h)}-I_hu^{(h)}\|_{L^\infty(\Omega_h)}\\
&\leqslant C\ell_h\|u-I_hu\|_{L^\infty(\Omega_h)}+C\ell_hh^{r+1}\|f\|_{L^p(\Omega^t)}, \qquad (4.4.15)
\end{aligned}
$$

where the last inequality uses the triangle inequality and (4.4.9), and $I_h\widetilde{u}$ is the interpolation operator associated with the larger triangulation $\mathcal{T}_{*,h}$ which extends the interpolation operator $I_h : C(\bar{\Omega}_h) \to S_h(\Omega_h)$ associated with $\mathcal{T}_h$. Since $\tilde{u}_h - u_h$ is discrete harmonic in Ω_h, i.e.,

$$
\int_{\Omega_h}\nabla(\tilde{u}_h-u_h)\cdot\nabla\chi_h\mathbf{d}x=\int_{\Omega_h}\nabla(\tilde{u}-u^{(h)})\cdot\nabla\chi_h\mathbf{d}x=0\quad\forall\,\chi_h\in S_h^\circ(\Omega_h),
$$

it follows from Theorem 4.2.1 that $\tilde{u}_h - u_h$ satisfies the discrete maximum principle, i.e.,

$$
\begin{aligned}
\|\tilde{u}_h-u_h\|_{L^\infty(\Omega_h)} &\leqslant C\|\tilde{u}_h-u_h\|_{L^\infty(\partial\Omega_h)}\\
&=C\|\tilde{u}_h\|_{L^\infty(\partial\Omega_h)}\\
&=C\|\tilde{u}_h-\tilde{u}\|_{L^\infty(\partial\Omega_h)}\quad(\text{since } \tilde{u}|_{\partial\Omega_h}=0)\\
&\leqslant C\|\widetilde{u}_h-\widetilde{u}\|_{L^\infty(\Omega_*)}. \qquad (4.4.16)
\end{aligned}
$$

Substituting (4.4.15) and (4.4.16) into (4.4.14) yields

$$
\|u^{(h)}-u_h\|_{L^\infty(\Omega_h)}\leqslant C\ell_h\|u-I_hu\|_{L^\infty(\Omega_h)}+C\ell_hh^{r+1}\|f\|_{L^p(\Omega^t)}.
$$

Since $u^{(h)} = u_h = 0$ in $\Omega\backslash\Omega_h$, it follows that

$$
\|u^{(h)}-u_h\|_{L^\infty(\Omega)}=\|u^{(h)}-u_h\|_{L^\infty(\Omega\cap\Omega_h)}\leqslant C\ell_h\|u-I_hu\|_{L^\infty(\Omega_h)}+C\ell_hh^{r+1}\|f\|_{L^p(\Omega^t)}.
$$

Then, combining this with (4.4.9), we obtain the following error bound:

$$
\|u-u_h\|_{L^\infty(\Omega)}\leqslant C\ell_h\|u-I_hu\|_{L^\infty(\Omega_h)}+C\ell_hh^{r+1}\|f\|_{L^p(\Omega^t)}.
$$

Finally, we note that

$$
\begin{aligned}
\|u-\check{I}_hu\|_{L^\infty(\Omega)} =&\|u\circ\Phi_h-I_h(u\circ\Phi_h)\|_{L^\infty(\Omega_h)}\\
\geq&\|u-I_hu\|_{L^\infty(\Omega_h)}-C\|u-u\circ\Phi_h\|_{L^\infty(\Omega_h)}\\
\geq&\|u-I_hu\|_{L^\infty(\Omega_h)}-C\|u\|_{W^{1,\infty}(\mathbb{R}^d)}\|\Phi_h-\mathrm{Id}\|_{L^\infty(\Omega_h)}\\
\geq&\|u-I_hu\|_{L^\infty(\Omega_h)}-Ch^{r+1}\|u\|_{W^{1,\infty}(\mathbb{R}^d)}\\
\geq&\|u-I_hu\|_{L^\infty(\Omega_h)}-Ch^{r+1}\|f\|_{L^p(\Omega^t)}.
\end{aligned}
$$

This proves the result of Theorem 4.2.2. □

4.5 Conclusion

We have proved the weak maximum principle of the isoparametric FEM for the Poisson equation in curvilinear polyhedral domains with edge openings smaller than π, which include smooth domains and smooth deformations of convex polyhedra. The proof requires using a duality argument for an elliptic equation with some discontinuous coefficients arising from the use of isoparametric finite elements. Hence, the standard H^2 elliptic regularity does not hold for the solution of the corresponding dual problem. We have overcome the difficulty by decomposing the solution into a smooth H^2 part and a nonsmooth $W^{1,p}$ part, separately, and replaced the H^2 regularity required in a standard duality argument by some $W^{1,p}$ estimates for the nonsmooth part of the solution.

As an application of the weak maximum principle, we have proved an L^∞-norm best approximation property of the isoparametric FEM for the Poisson equation. All the analysis for the Poisson equation in this chapter can be extended to elliptic equations with $W^{1,\infty}$ coefficients. However, the current analysis does not allow us to extend the results to curvilinear polyhedral domains with edge openings bigger than π (smooth deformations of nonconvex polyhedra) or graded mesh in three dimensions. These would be the subject of future research.

There are other approaches to the maximum principle of finite element methods for elliptic equations using non-obtuse meshes, which is restricted to piecewise linear finite elements and Poisson equation with constant coefficients; see [62]. The approach in the current manuscript is applicable to elliptic equations with $W^{1,\infty}$ coefficients, general quasi-uniform meshes, and high-order finite elements, and therefore requires completely different analysis from the approaches using non-obtuse meshes.

Chapter 5

Stability, analyticity and maximal regularity of semi-discrete isoparametric finite element solutions of parabolic equations in curvilinear polyhedra

5.1 Introduction

Let Ω be a curvilinear polygonal (in 2D) or polyhedral (in 3D) domain in $\mathbb{R}^N$ (where $N \in \{2,3\}$) with edge openings possibly larger than π, and consider the heat equation

$$\frac{\partial u(t,x)}{\partial t} - \Delta u(t,x) = f(t,x), \qquad \forall (t,x) \in \mathbb{R}_+ \times \Omega, \tag{5.1.1}$$

$$u(t,x) = 0, \qquad \forall (t,x) \in \mathbb{R}_+ \times \partial\Omega, \tag{5.1.2}$$

$$u(0,x) = u_0(x), \qquad \forall x \in \Omega. \tag{5.1.3}$$

In the case of $f = 0$ it is well-known that the solution of (5.1.1) is given by $u(t,x) = e^{t\Delta}u(x)$, where $E(t) = e^{t\Delta}$ extends to an analytic semigroup on $C_0(\overline{\Omega})$ and $L^q(\Omega)$ for any $1 \le q < \infty$ (cf. [119]), and satisfies the following analytic estimates:

$$\sup_{t>0} \left(\|E(t)v\|_{L^q(\Omega)} + t\|\partial_t E(t)v\|_{L^q(\Omega)} \right) \le C\|v\|_{L^q(\Omega)}, \quad \forall v \in L^q(\Omega), \quad 1 \le q < \infty \tag{5.1.4a}$$

$$\sup_{t>0} \left(\|E(t)v\|_{C_0(\overline{\Omega})} + t\|\partial_t E(t)v\|_{C_0(\overline{\Omega})} \right) \le C\|v\|_{C_0(\overline{\Omega})}, \quad \forall v \in C_0(\overline{\Omega}). \tag{5.1.4b}$$

When $u_0 = 0$, the solution of (5.1.1) exhibits maximal L^p regularity in the space $L^q(\Omega)$. Specifically, for all $f \in L^p(\mathbb{R}_+; L^q(\Omega))$, the solution satisfies:

$$\|\partial_t u\|_{L^p(\mathbb{R}_+;L^q(\Omega))} + \|\Delta u\|_{L^p(\mathbb{R}_+;L^q(\Omega))} \le C_{p,q}\|f\|_{L^p(\mathbb{R}_+;L^q(\Omega))} \quad \forall 1 < p,q < \infty. \tag{5.1.5}$$

Maximal L_p-regularity, as described in (5.1.5), plays a crucial role in the analysis of nonlinear partial differential equations (PDEs) [3, 37, 38, 107] and has been extensively studied in the literature; see [89, 114, 139] and the references therein.

This chapter addresses the heat equation on a curvilinear polyhedral domain Ω, which cannot be exactly triangulated by linear simplices. To achieve high-order finite element methods (FEM) in such cases, an effective approach is to use isoparametric elements. The work of [94] provides a systematic way to construct a family $\mathcal{T}_h$ of isoparametric elements of order r for each $h > 0$. Each boundary simplex $K \in \mathcal{T}_h$ contains a curved face or edge interpolating $\partial\Omega$ with an accuracy of $O(h^{r+1})$. The approximate domain Ω_h = interior of $\left(\bigcup_{K\in\mathcal{T}_h} K\right)$ satisfies $\mathrm{dist}(x,\Omega) = O(h^{r+1})$ for $x \in \Omega_h$ and $\mathrm{dist}(x,\Omega_h) = O(h^{r+1})$ for $x \in \Omega$. Based on these isoparametric elements, we define the finite element space $S_h(\Omega_h) \subseteq H^1(\Omega_h) \cap C(\overline{\Omega}_h)$. The semi-discrete isoparametric FEM approximation for the heat equation (5.1.1) then involves finding $u_h(t) \in S_h^\circ(\Omega_h)$ that satisfies:

$$\begin{cases} (\partial_t u_h, \chi_h)_{\Omega_h} + (\nabla u_h, \nabla \chi_h)_{\Omega_h} = (f_h, v_h)_{\Omega_h}, \quad \forall \chi_h \in S_h^\circ(\Omega_h), \forall t > 0, \\ u_h(0) = u_{h,0} \in S_h^\circ(\Omega_h). \end{cases} \tag{5.1.6}$$

where

$$S_h^\circ(\Omega_h) := \{\chi_h \in S_h(\Omega_h) : \chi_h|_{\partial\Omega_h} = 0\} \tag{5.1.7}$$

and $f_h(t) \in S_h^\circ(\Omega_h)$ is some source term function. Let $E_h(t) = e^{t\Delta_h}$ denote the discrete semigroup on $S_h^\circ(\Omega_h)$ generated by the operator Δ_h. Then $u_h(t) = E_h(t)v_h$ gives the solution of equation (5.1.6) when $u_{h,0} = v_h$ and $f_h = 0$. The aim of this chapter is to prove the following analogues of (5.1.4) and (5.1.5) for the semi-discrete problem (5.1.6):

$$\sup_{t>0} \left(\|E_h(t)v_h\|_{L^q(\Omega_h)} + t\|\partial_t E_h(t)v_h\|_{L^q(\Omega_h)}\right) \le C\|v_h\|_{L^q(\Omega_h)} \tag{5.1.8a}$$
$$\forall v_h \in S_h^\circ(\Omega_h), \quad 1 \le q \le \infty,$$

$$\|\partial_t u_h\|_{L^p(\mathbb{R}_+;L^q(\Omega_h))} + \|\Delta_h u_h\|_{L^p(\mathbb{R}_+;L^q(\Omega_h))} \le C_{p,q}\|f_h\|_{L^p(\mathbb{R}_+;L^q(\Omega_h))} \tag{5.1.8b}$$
$$\text{if } u_{h,0} = 0, \ \forall 1 < p, q < \infty.$$

The analyticity and maximal regularity of the finite element semi-discrete or fully-discrete problem have numerous applications and serve as important tools for the convergence analysis of numerical schemes for nonlinear parabolic equations [2, 52, 67, 104, 88, 143].

Historically, there has been extensive literature examining the analyticity (5.1.8a) and maximal regularity (5.1.8b) of the finite element discrete semigroup. By analyzing the discrete Green's function, [126, 133] established the analyticity (5.1.8a) of the discrete semigroup E_h when the domain is smooth. The key estimate for the discrete Green's function discussed in [126, 133] was subsequently utilized in [66] to demonstrate the maximal L^p regularity (5.1.8b) of the discrete semigroup $E_h(t)$ when the domain and coefficients of parabolic equation are sufficiently smooth. The extension of semidiscrete maximal L^p-regularity (5.1.8b) to fully discrete finite element methods has been established for various time discretization methods, including the backward Euler method [8, 105], discontinuous Galerkin method [99], θ-schemes [82] and A-stable multistep and Runge–Kutta methods [85].

Subsequent studies have relaxed the requirements on the smoothness of the domain and coefficients necessary to obtain analyticity and maximal regularity estimates. The results in [101, 100, 104] have shown that (5.1.8a) and (5.1.8b) hold when Ω is a polyhedral domain (possibly nonconvex) and the coefficients of the parabolic equation satisfy $a_{ij} \in W^{1,N+\varepsilon}(\Omega)$. Furthermore, the discrete maximal regularity of fully-discrete k-step BDF methods for parabolic equations in polyhedral (possibly nonconvex) domain was established in [102].

The above results regarding (5.1.8a)-(5.1.8b) are valid only when the domain Ω is assumed to be exactly triangulated. When $\Omega \neq \Omega_h$, it becomes necessary to address the domain perturbation effect. As far as we know, there are in general two approaches to deal with the domain perturbation effect. One approach is the extension method, where the exact solution u is extended over a neighborhood of Ω. The extended solution $\widetilde{u}$ satisfies the original equation in Ω_h except along the boundary skin $\Omega\Delta\Omega_h$. Analyzing the consistency error terms associated with the boundary skin effect is crucial in this method; see the analysis using the extension methods in [80, 79, 81, 71]. Specifically, using the extension method, [80] proved (5.1.8a)-(5.1.8b) for finite element semi-discretization of parabolic problems on a smooth domain Ω with Neumann boundary conditions, where Ω_h approximates the original Ω through a quasi-uniform triangulation $\mathcal{T}_h$ consisting of linear simplicies, and $S_h(\Omega_h)$ is the continuous P^1 element space based on $\mathcal{T}_h$.

In this chapter, we adopt an alternative approach—the transformation method—to address the domain perturbation $\Omega \neq \Omega_h$. Specifically, we utilize the Lipschitz homeomorphism $\mathbf{\Phi}_h : \Omega_h \to \Omega$ constructed in [94], to transform the equation (5.1.6) into an equation defined on domain Ω:

$$\begin{cases} (a_h(x)\partial_t \check{u}_h, \check{v}_h)_\Omega + (A_h(x)\nabla \check{u}_h, \nabla \check{v}_h)_\Omega = (a_h(x)\check{f}_h, \check{v}_h)_\Omega, \quad \forall \check{v}_h \in \check{S}_h^\circ(\Omega), \forall t > 0, \\ \check{u}_h(0) = \check{u}_{h,0} := u_{h,0} \circ \mathbf{\Phi}_h^{-1}. \end{cases} \tag{5.1.9}$$

Then (5.1.8a)-(5.1.8b) would become equivalent to the analyticity and maximal regularity estimate of equation (5.1.9) (see Section 5.3.2). Let $\check{\Gamma}_h$ be the discrete Green's function for equation (5.1.9) and Γ be a regularized Green's function for the original equation (5.1.1). Both $\check{\Gamma}_h$ and Γ are defined on Ω; however, it is important to note that $\check{\Gamma}_h - \Gamma$ does not satisfy Galerkin orthogonality due to $a_h(x) \neq 1$ and $A_h(x) \neq I_N$ for $x \in \Omega$ with $\mathrm{dist}(x, \partial\Omega) \le ch$.

In general, the transformation method involves using a map like $\mathbf{\Phi}_h : \Omega_h \to \Omega$ to lift finite element functions to Ω by $u_h^l := u_h \circ \mathbf{\Phi}_h^{-1}$. The domain perturbation $\Omega \neq \Omega_h$ is then captured in the discrepancies of the mass bilinear forms and stiffness bilinear forms:

$$\begin{aligned} m_h(u_h, v_h) - m(u_h, v_h) &:= (u_h, v_h)_{\Omega_h} - (u_h^l, v_h^l)_\Omega \\ A_h(u_h, v_h) - A(u_h, v_h) &:= (\nabla u_h, \nabla v_h)_{\Omega_h} - (\nabla u_h^l, \nabla v_h^l)_\Omega. \end{aligned}$$

This approach is widely used in analyzing domain perturbation effects [46, 49, 27] and can be adapted for problems involving moving domains or surfaces [11, 87, 48, 50]. Specifically, optimal H^1 and L^2 error estimates for isoparametric FEM applied to the heat equation and the Cahn–Hilliard equation with dynamic boundary conditions are provided in [87] and [27], respectively. Evolving bulk and surface isoparametric finite element spaces on evolving triangulations are defined and developed in [49, 50] for coupled bulk–surface system. [48] examines an evolving bulk–surface model in which a Poisson equation with a generalized Robin boundary condition on the domain is coupled to a forced mean curvature flow of the free boundary, proving an optimal H^1 error estimate for the spatial semi-discretization using bulk–surface finite elements. Lastly, [11] establishes maximal regularity for evolving surface FEM applied to parabolic equations on moving surfaces, utilizing a temporal perturbation argument to extend results from stationary to evolving surfaces.

Our proof of (5.1.8a)-(5.1.8b) essentially follows the same strategy in [101]. We reduce (5.1.8a)-(5.1.8b) to an L^1-type error estimate between $\check{\Gamma}_h$ and Γ (Lemma 5.4.3). The estimates utilize a dyadic decomposition of the domain $(0,1) \times \Omega = \cup_{*,j} Q_j$ and the kick-back argument from [126]. To address the singularity arising from non-convex corners, we analyze the local $L^2H^{1+\alpha}(Q_j)$ and $L^\infty H^{1+\alpha}(Q_j)$ estimates of the Green's function (Lemma 5.4.2), as in [101]. A local energy error estimate for finite element solutions of parabolic equations (Lemma 5.5.1) and a local duality argument (see (5.5.111)–(5.5.120)) are key components of the kick-back argument. The main challenge in our proof arises during the local energy error estimate and the duality argument, similar to [80]. Since only perturbed Galerkin orthogonality holds for $\check{\Gamma}_h - \Gamma$, extra terms arising from the domain perturbation effect (terms involving $a_h - 1$ or $A_h - I_N$) must be handled carefully.

As shown in [101], the quasi-maximal L^∞-regularity facilitates reducing the maximum-norm stability of finite element solutions for parabolic equations to the maximum-norm stability of the elliptic Ritz projection. Specifically, the following estimate holds:

$$\|u - u_h\|_{L^\infty(0,T;L^\infty(\Omega))} \le C\left(\ell_h\|u - R_h u\|_{L^\infty(0,T;L^\infty(\Omega))} + \|u_0 - u_{h,0}\|_{L^\infty(\Omega)}\right). \quad (5.1.10)$$

In our context, since u_h solves the equation in a perturbed domain Ω_h, the error introduced by domain perturbation is nontrivial to eliminate. Nevertheless, this chapter establishes that the additional error caused by the domain perturbation remains of optimal order. Specifically, we show:

$$\begin{aligned}\|\widetilde{u} - u_h\|_{L^\infty(0,T;L^\infty(\Omega_h))} \le & C\left(\ell_h\|\widetilde{u} - R_h\widetilde{u}\|_{L^\infty(0,T;L^\infty(\Omega_h))} + \|\widetilde{u}_0 - u_{h,0}\|_{L^\infty(\Omega_h)}\right) \\ & + Ch^{r+1}\left(\|u\|_{L^\infty(0,T;W^{2,\infty}(\Omega))} + \|\partial_t u\|_{L^\infty(0,T;L^\infty(\Omega))}\right), \quad (5.1.11)\end{aligned}$$

where $\widetilde{u}$ is a Sobolev extension of the exact solution u over $\Omega \cup \Omega_h$.

Our current work presents both connections and distinctions with the previous study [103], which investigated the weak maximal principle for isoparametric FEM in the context of elliptic equations. In both studies, the primary result is reduced to obtaining an L^1-type error estimate of a regularized Green's function. Specifically, the L^1-type error estimate pertains to the function v defined in [103, Eq. (2.16)] and the function Γ defined in (5.3.53), respectively.

However, in [103], to align with the reduction process of [96], v is defined as the solution of an elliptic equation with discontinuous coefficients (cf. [103, Eq. (2.16)]). The main result is reduced to proving an L^1-type error estimate (cf. [103, Eq. (2.26)]) for $v - R_h v$, where R_h denotes the Ritz projection for elliptic equations with the same discontinuous coefficients (cf. [103, Eq. (2.21)]). Although $v - R_h v$ satisfies Galerkin orthogonality, the difficulty arises from the limited regularity of v, as it solves an elliptic equation with discontinuous coefficients. To address this, [103] decomposes v into two components: v_1, a regularized Green's function for the original Laplacian equation, and v_2, which accounts for the effects of domain perturbation (cf. [103, Section 2.4]). Consequently, the analysis in [103] does not require addressing local energy error estimates or duality arguments under almost Galerkin orthogonality.

In contrast, in the current work, the function Γ retains the same regularity as the Green's function for the heat equation in the curvilinear polyhedron. However, the error $\Gamma - \check{\Gamma}_h$ satisfies only an almost Galerkin orthogonality:

$$(\partial_t\Gamma - a_h(x)\partial_t\check{\Gamma}_h, \check{\chi}_h)_\Omega + (\nabla\Gamma - A_h(x)\nabla\check{\Gamma}_h, \nabla\check{\chi}_h)_\Omega = 0 \quad \forall \check{\chi}_h \in \check{S}_h^\circ(\Omega).$$

The primary difficulty lies in handling the local energy error estimates and the local duality arguments in the context of almost Galerkin orthogonality. We address the local energy error estimate by decomposing the error $\phi - \check{\phi}_h$ into two terms: $\phi - \check{\theta}_h$ and $\check{\eta}_h$. The term $\phi - \check{\theta}_h$ satisfies a local Galerkin orthogonality and can be analyzed using [101, Lemma

5.1], while the term $\check{\eta}_h$, which captures the domain perturbation effects ($A_h - I \neq 0$ and $a_h - 1 \neq 0$), satisfies the following equation:

$$(a_h(x)\partial_t\check{\eta}_h, \check{\chi}_h) + (A_h(x)\nabla\check{\eta}_h, \nabla\check{\chi}_h) = ((1 - a_h)\partial_t\widetilde{\phi}, \check{\chi}_h) + ((I - A_h)\nabla\widetilde{\phi}, \nabla\check{\chi}_h) \quad (5.1.12)$$
$$\forall\check{\chi}_h \in \check{S}_h^\circ(\Omega), t \in (0, 1) \text{ and } \check{\eta}_h(0) = 0.$$

The additional term $\check{\eta}_h$ is bounded using a global parabolic energy estimate, introducing the extra term $Y_j(\phi)$ in the local energy error estimate (cf. Lemma 5.5.1). While the local duality argument largely follows the structure in [101], additional terms, such as $\mathcal{I}_2$, $\mathcal{I}_3$, and $\mathcal{I}_4$ in (5.5.113), arise due to the domain perturbation effects. For example, $\mathcal{I}_2$ is given by:

$$\mathcal{J}_2 = [(1 - a_h(x))\partial_t\Gamma, w]_{\mathcal{Q}} + [(I_N - A_h(x))\nabla\Gamma, \nabla w]_{\mathcal{Q}}$$

As in [101], these additional terms are controlled using local $H^{1+\alpha}$ estimates (5.5.119) for w (the solution of the parabolic duality problem) and local energy estimates (5.4.65c) for Γ, based on the parabolic dyadic decomposition.

The rest of this chapter is organized as follows. In Section 5.2, we present the main results concerning the analyticity and maximal regularity of the discrete semigroup, as well as the quasi-optimal maximum norm error estimate for isoparametric FEM. Section 5.3 introduces some preliminary results on isoparametric FEM and the Green's functions, and reformulates (5.1.8a)-(5.1.8b) using the transformation method. In Section 5.4, we prove the main results with the assistance of Lemma 5.4.3, which provides an L^1-type error estimate between $\check{\Gamma}_h$ and Γ. Section 5.5 contains the proof of the key Lemma 5.4.3 through a local energy error estimate and a local duality argument. Finally, in Section 5.6, we establish Lemma 5.5.1, which is utilized in the local energy error estimate.

5.2 Main results

The discrete semigroup E_h has an associated kernel $\Gamma_h(t, x, x_0) := (E_h(t)\delta_{h,x_0})(x)$ such that

$$(E_h(t)v_h)(x_0) = \int_{\Omega_h} \Gamma_h(t, x, x_0)v_h(x)dx \quad \forall v_h \in S_h^\circ(\Omega_h), \quad (5.2.13)$$

where $\delta_{h,x_0} \in S_h^\circ(\Omega_h)$ is the discrete delta function satisfying $(\delta_{h,x_0}, v_h)_{\Omega_h} = v_h(x_0)$ for all $v_h \in S_h^\circ(\Omega_h)$. We can define $|E_h(t)|$ as a linear operator on $L^q(\Omega_h)$ with kernel $|\Gamma_h(t, x, x_0)|$,

$$(|E_h(t)|v)(x_0) := \int_{\Omega_h} |\Gamma_h(t, x, x_0)|v(x)dx \quad \forall v \in L^q(\Omega_h). \quad (5.2.14)$$

The main result of this chapter is the following theorem.

Theorem 5.2.1. *Let Ω be a curvilinear polyhedral domain in $\mathbb{R}^N$ (with edge openings possibly larger than π), and let $S_h^\circ(\Omega_h)$, $0 < h < h_0$ be the finite element spaces based on the family $\mathcal{T}_h$ of isoparametric elements. Then, for the semi-discrete equation* (5.1.6)*, we have the following analytic semigroup estimate:*

$$\sup_{t>0} \left(\|E_h(t)v_h\|_{L^q(\Omega_h)} + t\|\partial_t(E_h(t)v_h)\|_{L^q(\Omega_h)}\right) \leq C\|v_h\|_{L^q(\Omega_h)} \quad \forall v_h \in S_h^\circ(\Omega_h), \forall 1 \leq q \leq \infty, \quad (5.2.15a)$$

$$\|\sup_{t>0}|E_h(t)||v|\|_{L^q(\Omega_h)} \le C_q\|v\|_{L^q(\Omega_h)}, \quad \forall v \in L^q(\Omega_h),\ \forall 1 < q \le \infty. \tag{5.2.15b}$$

Further, if $u_{h,0} = 0$ and $f_h \in L^p(0,T;L^q(\Omega_h))$, then the solution $u_h(t)$ of equation (5.1.6) *possesses the following maximal L^p-regularity:*

$$\|\partial_t u_h\|_{L^p(0,T;L^q(\Omega_h))} + \|\Delta_h u_h\|_{L^p(0,T;L^q(\Omega_h))} \le \max(p,(p-1)^{-1})C_q\|f_h\|_{L^p(0,T;L^q(\Omega_h))} \tag{5.2.16a}$$

$$\forall 1 < p,q < \infty$$

$$\|\partial_t u_h\|_{L^\infty(0,T;L^q(\Omega_h))} + \|\Delta_h u_h\|_{L^\infty(0,T;L^q(\Omega_h))} \le C\ell_h\|f_h\|_{L^\infty(0,T;L^q(\Omega_h))}, \tag{5.2.16b}$$

$$\forall 1 \le q \le \infty,$$

where $\ell_h := \log_2(2+1/h)$. The constant C in (5.2.15a) *and* (5.2.16b) *is independent of f_h, h, p, q and T. The constant C_q in* (5.2.15b) *and* (5.2.16a) *is independent of f_h, h, p and T.*

To analyze the error between exact and numerical solutions, let $E : L^1(\Omega) \to L^1(\mathbb{R}^N)$ denote the Stein extension operator (cf. [130, page 181, Theorem 5]) which continuously maps the Sobolev spaces $W^{k,p}(\Omega)$ into $W^{k,p}(\mathbb{R}^N)$ for each $1 \le p \le \infty$ and $k \ge 0$. We use abbreviation $\widetilde{\phi} := E\phi$ for a function ϕ defined on Ω.

To solve the semi-discrete equation (5.1.6) as an approximation of the original heat equation (5.1.1), one choice is to set $f_h(t) = P_h\widetilde{f}(t)$, where $\widetilde{f}$ is the extension of f discussed as above and P_h is the $L^2(\Omega_h)$-orthogonal projection onto $S_h^\circ(\Omega_h)$ i.e.,

$$(P_hf, v_h)_{\Omega_h} = (f, v_h)_{\Omega_h} \quad \forall v_h \in S_h^\circ(\Omega_h).$$

We define the elliptic Ritz projection $R_h : H^1(\Omega_h) \to S_h^\circ$ as:

$$(\nabla R_h\phi, \nabla v_h)_{\Omega_h} = (\nabla\phi, \nabla v_h)_{\Omega_h} \quad \forall v_h \in S_h^\circ(\Omega_h). \tag{5.2.17}$$

Then, as an application of Theorem 5.2.1, we can prove the following L^∞-norm error estimate which is analogous to [101, Corollary 2.2] in the context of isoparametric FEM.

Theorem 5.2.2. *Let Ω be a curvilinear polyhedral domain with edge openings possibly larger than π. Let u be the solution of equation* (5.1.1)*, and let u_h be the solution of equation* (5.1.6) *with $f_h = P_h\widetilde{f}$. When h is sufficiently small, the following holds:*

$$\begin{aligned}\|\widetilde{u} - u_h\|_{L^\infty(0,T;L^\infty(\Omega_h))} \le & C\ell_h\|\widetilde{u} - R_h\widetilde{u}\|_{L^\infty(0,T;L^\infty(\Omega_h))} + C\|u_{h,0} - \widetilde{u}_0\|_{L^\infty(\Omega_h)} \\ & + Ch^{r+1}\left(\|u\|_{L^\infty(0,T;W^{2,\infty}(\Omega))} + \|\partial_t u\|_{L^\infty(0,T;L^\infty(\Omega))}\right),\end{aligned} \tag{5.2.18}$$

where $\widetilde{u} = Eu$ and $\widetilde{f} = Ef$ denote the extensions of the exact solution u and the source term f, respectively, using the Stein extension operator $E : L^1(\Omega) \to L^1(\mathbb{R}^N)$. The constant C is independent of h, T, f and u, and r is the order of the isoparametric elements

Remark 5.2.1. When Ω is a curvilinear polyhedral domain with edge openings smaller than π, we apply [103, Theorem 1.2] to derive a maximum norm estimate for the projection error $\widetilde{u}(t) - R_h\widetilde{u}(t)$. Specifically, [103, Theorem 1.2] establishes the following: Let $g \in L^\infty(\Omega)$, and v solves the Poisson equation

$$-\Delta v = g \text{ in } \Omega, \quad v = 0 \text{ on } \partial\Omega. \tag{5.2.19}$$

Let $v_h \in S_h^\circ(\Omega_h)$ be the finite element approximation solving

$$(\nabla v_h, \nabla \chi_h)_{\Omega_h} = (\widetilde{g}, \chi_h)_{\Omega_h} \quad \forall \chi_h \in S_h^\circ(\Omega_h), \tag{5.2.20}$$

where $\widetilde{g}$ is an extension of g to $\mathbb{R}^N$. For sufficiently small h, if v and v_h are extended by zero to $\Omega \cup \Omega_h$, the following holds:

$$\|v - v_h\|_{L^\infty(\Omega\cup\Omega_h)} \leq C\ell_h h^{r+1} \left(\|v\|_{W^{r+1,\infty}(\Omega)} + \|\widetilde{g}\|_{L^\infty(\mathbb{R}^N)}\right). \tag{5.2.21}$$

We use the result of (5.2.21) to estimate $\widetilde{u}(t) - R_h\widetilde{u}(t)$ by setting $v = u(t)$, $g = -\Delta u(t)$ in the Poisson equation (5.2.19), and $v_h = R_h\widetilde{u}(t)$, $\widetilde{g} = -\Delta\widetilde{u}(t)$ in its finite element counterpart (5.2.20). Consequently, we obtain

$$\begin{aligned}
&\|u(t) - R_h\widetilde{u}(t)\|_{L^\infty(\Omega_h\cap\Omega)} + \|u(t)\|_{L^\infty(\Omega\backslash\Omega_h)} + \|R_h\widetilde{u}(t)\|_{L^\infty(\Omega_h\backslash\Omega)} \\
\leq & C\ell_h h^{r+1} \left(\|u\|_{W^{r+1,\infty}(\Omega)} + \|\widetilde{u}(t)\|_{W^{2,\infty}(\mathbb{R}^N)}\right).
\end{aligned}$$

Since $\widetilde{u}(t)|_{\partial\Omega} = 0$ and $\mathrm{dist}(x, \partial\Omega) \leq Ch^{r+1}$ for $x \in \Omega_h \setminus \Omega$, we further deduce

$$\|\widetilde{u}(t)\|_{L^\infty(\Omega_h\backslash\Omega)} \leq Ch^{r+1}\|\widetilde{u}(t)\|_{W^{1,\infty}(\mathbb{R}^N)}.$$

Combining these results and utilizing the $W^{k,p}$-boundedness of the Stein extension operator, we obtain

$$\|\widetilde{u}(t) - R_h\widetilde{u}(t)\|_{L^\infty(\Omega_h)} \leq C\ell_h h^{r+1}\|u(t)\|_{W^{r+1,\infty}(\Omega)} \quad \forall r \geq 1.$$

Thus, the following quasi-optimal maximum norm error estimate holds for the parabolic problem:

$$\begin{aligned}
\|\widetilde{u} - u_h\|_{L^\infty(0,T;L^\infty(\Omega_h))} \leq & C\|u_{h,0} - \widetilde{u}_0\|_{L^\infty(\Omega_h)} \\
& + Ch^{r+1}\ell_h^2 \left(\|u\|_{L^\infty(0,T;W^{r+1,\infty}(\Omega))} + \|\partial_t u\|_{L^\infty(0,T;L^\infty(\Omega))}\right).
\end{aligned} \tag{5.2.22}$$

These two theorems are demonstrated in Section 5.4.

5.3 Preliminary

5.3.1 Notations of function spaces

Let $\Omega \subseteq \mathbb{R}^N$ be the curvilinear polyhedral domain in (5.1.1). We use the conventional notations of Sobolev spaces $W^{s,q}(\Omega)$, $s \geq 0, 1 \leq q \leq \infty$ (c.f. [1]) and Hölder spaces $C^\gamma(\overline{\Omega})$, $0 < \gamma < 1$, with the abbreviations $L^q = W^{0,q}(\Omega)$, $W^{s,q} = W^{s,q}(\Omega)$, $C^\gamma = C^\gamma(\overline{\Omega})$ and $H^s := W^{s,2}(\Omega)$. The notation $H^{-s}(\Omega)$ denotes the dual space of $H_0^s(\Omega)$, which is the closure of $C_0^\infty(\Omega)$ in $H^s(\Omega)$. The Bochner norm of a function $f : (0,T) \to W^{s,p}$ is defined as

$$\|f\|_{L^p(0,T;W^{s,q})} := \|\|f(\cdot)\|_{W^{s,q}}\|_{L^p(0,T)} \quad \forall 1 \leq p, q \leq \infty, s \in \mathbb{R},$$

For any subdomain $D \subseteq \Omega$, we make the following convention concerning the Sobolev spaces on D:

$$\|f\|_{W^{s,q}(D)} := \inf_{\widetilde{f}|_D = f} \|\widetilde{f}\|_{W^{s,q}(\Omega)} \quad \forall 1 \leq p, q \leq \infty, s \in \mathbb{R}, \tag{5.3.23}$$

where the infimum extends over all possible $\widetilde{f}$ defined on Ω with $\widetilde{f}|_D = f$. Similarly, for any time-space subdomain $Q \subseteq \mathcal{Q} = (0,1) \times \Omega$, we define

$$\|f\|_{L^pW^{s,q}(Q)} := \inf_{\widetilde{f}|_Q = f} \|\widetilde{f}\|_{L^p(0,1;W^{s,q})} \quad \forall 1 \le p, q \le \infty, s \in \mathbb{R}, \tag{5.3.24}$$

where the infimum extends over all possible $\widetilde{f}$ defined on $\mathcal{Q}$ such that $\widetilde{f}|_Q = f$ in Q. We define the Bochner norms for the Hölder spaces in the same way:

$$\|f\|_{L^pC^\gamma(Q)} := \inf_{\widetilde{f}|_Q = f} \|\widetilde{f}\|_{L^p(0,1;C^\gamma)} \quad \forall 1 \le p, q \le \infty, s \in \mathbb{R}.$$

One advantage of our convention for Sobolev and Hölder norms on $Q \subseteq \mathcal{Q}$ is the following: if there holds Sobolev embedding $W^{s,q}(\Omega) \hookrightarrow C^\gamma(\overline{\Omega})$, then

$$\|f\|_{L^pC^\gamma(Q)} \le C\|f\|_{L^pW^{s,q}(Q)},$$

where the constant C is independent of the subdomain $Q \subseteq \mathcal{Q}$.

Finally, we use the abbreviations

$$(\phi, \varphi) := \int_\Omega \phi(x)\varphi(x)dx \qquad [u, v] := \int_0^1 \int_\Omega u(t,x)v(t,x)dxdt, \tag{5.3.25}$$

$$(\phi, \varphi)_{\Omega_h} := \int_{\Omega_h} \phi(x)\varphi(x)dx \qquad [u, v]_{\Omega_h} := \int_0^1 \int_{\Omega_h} u(t,x)v(t,x)dxdt, \tag{5.3.26}$$

and denote $\omega(t) := \omega(t, \cdot)$ for the slice at time t of any function ω defined on $\mathcal{Q}$.

5.3.2 Preliminary of the isoparametric FEM

Definition of isoparametric FEM

We denote by $\widetilde{\mathcal{T}}_h$ a quasi-uniform triangulation of the curvilinear polygonal or polyhedral domain $\Omega \subseteq \mathbb{R}^N$, using triangles in 2D or tetrahedra in 3D. For each simplex $\widetilde{K} \in \widetilde{\mathcal{T}}_h$, there is a linear parametric map $\boldsymbol{F}_{\widetilde{K}} : \hat{K} \to \widetilde{K}$ from the reference simplex $\hat{K}$ to $\widetilde{K}$. For a boundary simplexes $\widetilde{K}$, we denote by $\widetilde{D}$ the face or edge of $\widetilde{K}$ attaching to the boundary $\partial\Omega$, and let $\hat{D} \subseteq \hat{K}$ be corresponding face or edge of the reference simplex such that $\boldsymbol{F}_{\widetilde{K}}(\hat{D}) = \widetilde{D}$. The work in [94] provides a systematic way to modify the linear parametric map $\boldsymbol{F}_{\widetilde{K}} : \hat{K} \to \widetilde{K}$ of a boundary simplex into $\boldsymbol{F}_K : \hat{K} \to \mathbb{R}^N$ so that $\boldsymbol{F}_K$ is a vector-valued polynomial on $\hat{K}$ with degree no greater than a given integer $r \ge 1$. Moreover $\boldsymbol{F}_K|_{\hat{D}}$ interpolates the boundary $\partial\Omega$ at the Lagrangian nodes of degree r on the face or edge $\hat{D}$. For interior simplexes, the parametric maps remain unchanged, i.e., $\boldsymbol{F}_K = \boldsymbol{F}_{\widetilde{K}}$. Let K be the image of the parametric map $\boldsymbol{F}_K$; in this way we obtain a family $\mathcal{T}_h$ consisting of these possibly curved simplexes K. The family of parametric maps F_K constructed in [94] satisfies the following mesh regularity condition:

$$\|D^s\boldsymbol{F}_K\|_{L^\infty(\hat{K})} \le C_s h^s, \ \|D^s(\boldsymbol{F}_K^{-1})\|_{L^\infty(K)} \le C_s h^{-s} \quad \forall K \in \mathcal{T}_h, \forall s \ge 1, \tag{5.3.27}$$

where C_s is a positive constant independent of h. In addition, the parametric maps $\boldsymbol{F}_K$ are arranged in a mutually consistent way so that each $\mathcal{T}_h$ is still a triangulation with the same structure as $\widetilde{\mathcal{T}}_h$, i.e., $K, K' \in \mathcal{T}_h$ share a common face if and only if the corresponding pair $\widetilde{K}, \widetilde{K'} \in \widetilde{\mathcal{T}}_h$ share a common face. We define the approximate domain as $\Omega_h :=$ interior of $\bigcup_{K\in\mathcal{T}_h} K$, and the isoparametric finite element space $S_h(\Omega_h)$ of order r is defined as

$$S_h(\Omega_h) := \{\chi_h \in C(\overline{\Omega}_h) : \chi_h|_K \circ \boldsymbol{F}_K \in \mathcal{P}^r(\hat{K}) \ \ \forall K \in \mathcal{T}_h\},$$

where $\mathcal{P}^r(\hat{K})$ denotes the space of polynomials on $\hat{K}$ with degree no greater than r. The mesh regularity condition (5.3.27) guarantees that the finite element space $S_h(\Omega_h)$ satisfies the same local interpolation error estimate and inverse estimate as the usual Lagrangian finite element space based on a quasi-uniform triangle/tetrahedron mesh (cf. [94]).

Furthermore, similar as the construction of $\boldsymbol{F}_K$, for every given $m \geq r$, [94] associates each $K \in \mathcal{T}_h$ with a map $\boldsymbol{\Psi}_K^m : K \to \mathbb{R}^N$, which is a C^{m+1}-diffeomorphism from K to $\check{K}^m := \boldsymbol{\Psi}_K^m(K)$. In this chapter, we just choose $m = r$ and omit the superscript m; specifically the transformation $\boldsymbol{\Psi}_K$ is C^2-diffeomorphism when $r = 1$. For interior simplex K, $\boldsymbol{\Psi}_K$ equals to identity map, and for boundary simplex K, $\boldsymbol{\Psi}_K$ maps the curved boundary face/edge of the simplex K onto the exact domain boundary $\partial\Omega$. The map $\boldsymbol{\Phi}_h : \Omega_h \to \Omega$ defined by $\boldsymbol{\Phi}_h|_K = \boldsymbol{\Psi}_K$ gives a globally Lipschitz homeomorphism from the approximate domain Ω_h to the exact domain Ω. Let $\check{\mathcal{T}}_h$ be the family of $\{\check{K} = \boldsymbol{\Psi}_K(K) : K \in \mathcal{T}_h\}$; it follows that $\check{\mathcal{T}}_h$ gives a triangulation on Ω. The finite element space $\check{S}_h(\Omega) \subseteq C(\bar{\Omega})$ associated to $\check{\mathcal{T}}_h$ is defined via the Lipschitz homeomorphism $\boldsymbol{\Phi}_h$:

$$\check{S}_h(\Omega) := \{\check{\chi}_h \in C(\bar{\Omega}) : \check{\chi}_h \circ \boldsymbol{\Phi}_h \in S_h(\Omega_h)\}$$

Alternative formulation of Theorem 5.2.1

For any element $K \in \mathcal{T}_h$, we have estimates (cf. [94, Proposition 2 & Proposition 3 of Section 5])

$$\begin{aligned} &\|D^s(\boldsymbol{\Psi}_K - Id)\|_{L^\infty(K)} \leq Ch^{r+1-s} \quad \forall s \in [1, r+1], \\ &\|D^s(\boldsymbol{\Psi}_K^{-1} - Id)\|_{L^\infty(\check{K})} \leq Ch^{r+1-s} \quad \forall s \in [1, r+1]. \end{aligned} \tag{5.3.28}$$

As a corollary, let $\boldsymbol{F}_{\check{K}} := \boldsymbol{\Psi}_K \circ \boldsymbol{F}_K : \hat{K} \to \check{K}$ be the parametric map of $\check{K} \in \check{\mathcal{T}}_h$. Then $\boldsymbol{F}_{\check{K}}$ is a C^{r+1}-diffeomorphism from $\hat{K}$ to $\check{K}$ with derivatives satisfying estimate:

$$\|D^s \boldsymbol{F}_{\check{K}}\|_{L^\infty(\hat{K})} \leq Ch^s, \ \|D^s(\boldsymbol{F}_{\check{K}}^{-1})\|_{L^\infty(\check{K})} \leq Ch^{-s} \quad \forall K \in \mathcal{T}_h, \forall 1 \leq s \leq r+1 \tag{5.3.29}$$

From estimate (5.3.28), when h is sufficiently small, $\boldsymbol{\Phi}_h$ induces the following L^p and $W^{1,p}$ norm equivalence for each $1 \leq p \leq \infty$:

$$C^{-1}\|v \circ \boldsymbol{\Phi}_h\|_{L^p(\Omega_h)} \leq \|v\|_{L^p(\Omega)} \leq C\|v \circ \boldsymbol{\Phi}_h\|_{L^p(\Omega_h)} \quad \forall v \in L^p(\Omega) \tag{5.3.30}$$

$$C^{-1}\|\nabla(v \circ \boldsymbol{\Phi}_h)\|_{L^p(\Omega_h)} \leq \|\nabla v\|_{L^p(\Omega)} \leq C\|\nabla(v \circ \boldsymbol{\Phi}_h)\|_{L^p(\Omega_h)} \quad \forall v \in W^{1,p}(\Omega). \tag{5.3.31}$$

We define $a_h(x) := |\det\left(D\boldsymbol{\Psi}_K^{-1}(x)\right)|$ and $A_h(x) := a_h(x)\left(D\boldsymbol{\Psi}_K^{-1}(x)\right)^{-1}\left(D\boldsymbol{\Psi}_K^{-1}(x)\right)^{-\top}$ for any $x \in \Omega$. By (5.3.28), we have

$$\|a_h - 1\|_{L^\infty(\Omega)} + \|A_h - I_N\|_{L^\infty(\Omega)} \leq Ch^r, \tag{5.3.32}$$

where I_N denotes the $N \times N$ identity matrix. Observe that the isoparametric finite element method (5.1.6) is equivalent to

$$\begin{cases} (a_h(x)\partial_t \check{u}_h, \check{v}_h) + (A_h(x)\nabla\check{u}_h, \nabla\check{v}_h) = (a_h(x)\check{f}_h, \check{v}_h), \quad \forall \check{v}_h \in \check{S}_h^\circ(\Omega), \forall t > 0, \\ \check{u}_h(0) = \check{u}_{h,0} := u_{h,0} \circ \boldsymbol{\Phi}_h^{-1}, \end{cases} \tag{5.3.33}$$

where for elements $v_h, u_h, f_h \in S_h(\Omega_h)$ we use the convention $\check{v}_h := v_h \circ \mathbf{\Phi}_h^{-1}, \check{u}_h := u_h \circ \mathbf{\Phi}_h^{-1}, \check{f}_h := f_h \circ \mathbf{\Phi}_h^{-1}$, and we define $\check{S}_h^\circ(\Omega)$ as

$$\check{S}_h^\circ(\Omega) := \{\check{\chi}_h \in \check{S}_h(\Omega) : \check{\chi}_h|_{\partial\Omega} = 0\}. \tag{5.3.34}$$

We can define the corresponding operators associated with finite element space $\check{S}_h^\circ(\Omega)$, which have natural relations with those associated with the finite element space $S_h^\circ(\Omega_h)$:

The L^2-projection We denote by $\check{P}_h$ the weighted $L^2(\Omega)$-orthogonal projection onto $\check{S}_h^\circ(\Omega)$ defined as follows:

$$(a_h(x)\check{P}_h v, \check{\chi}_h) = (a_h(x)v, \check{\chi}_h) \quad \forall \check{\chi}_h \in \check{S}_h^\circ(\Omega). \tag{5.3.35}$$

Clearly, there holds following relation between P_h and $\check{P}_h$:

$$P_h(u \circ \mathbf{\Phi}_h) = \check{P}_h u \circ \mathbf{\Phi}_h.$$

The $L^2(\Omega_h)$-orthogonal projection P_h can be extended to a bounded operator on $L^q(\Omega_h)$, $1 \le q \le \infty$, i.e.,

$$\|P_h f\|_{L^q(\Omega_h)} \le C\|f\|_{L^q(\Omega_h)} \quad \forall f \in L^q(\Omega_h),\ \forall 1 \le q \le \infty, \tag{5.3.36}$$

where the constant C is independent of h and q. The estimate above is a consequence of [132, Lemma 6.1] and the self-adjointness of P_h. In view of the norm equivalence (5.3.30), it follows that $\check{P}_h$ also possesses L^q-stability:

$$\|\check{P}_h f\|_{L^q(\Omega)} \le C\|f\|_{L^q(\Omega)} \quad \forall f \in L^q(\Omega),\ \forall 1 \le q \le \infty. \tag{5.3.37}$$

The discrete Laplacian In the same way, one can transform the discrete Laplacian Δ_h to the corresponding operator $\check{\Delta}_h : \check{S}_h^\circ(\Omega) \to \check{S}_h^\circ(\Omega)$ defined by,

$$-(A_h(x)\nabla\check{u}_h, \nabla\check{\chi}_h)_\Omega = (a_h(x)\check{\Delta}_h\check{u}_h, \check{\chi}_h)_\Omega \quad \forall \check{\chi}_h \in \check{S}_h^\circ(\Omega), \tag{5.3.38}$$

with the relation

$$\Delta_h(\check{u}_h \circ \mathbf{\Phi}_h) = \check{\Delta}_h\check{u}_h \circ \mathbf{\Phi}_h.$$

The discrete semigroup With the operator $\check{\Delta}_h$, we can rewrite (5.3.33) equivalently as,

$$\begin{cases} \partial_t\check{u}_h - \check{\Delta}_h\check{u}_h = \check{f}_h \quad \forall t > 0, \\ \check{u}_h(0) = \check{u}_{h,0} := u_{h,0} \circ \mathbf{\Phi}_h^{-1}. \end{cases} \tag{5.3.39}$$

In the case of $\check{f}_h = 0$, there is a discrete semigroup $\check{E}_h(t)$ on finite element space $\check{S}_h^\circ(\Omega)$, generated by $\check{\Delta}_h$, such that $\check{E}_h(t)(\check{v}_h)$ is the solution of (5.3.33) when $\check{u}_{h,0} = \check{v}_h$. The following relation holds between $E_h(t)$ and $\check{E}_h(t)$:

$$(\check{E}_h(t)\check{v}_h) \circ \mathbf{\Phi}_h = E_h(t)v_h \quad \text{(convention } v_h = \check{v}_h \circ \mathbf{\Phi}_h \text{ used)},$$

which implies that

$$(\check{E}_h(t)\check{v}_h)(x_0) = \int_\Omega a_h(x)\Gamma_h(\mathbf{\Phi}_h^{-1}x, \mathbf{\Phi}_h^{-1}x_0, t)\check{v}_h(x)dx \quad \forall x_0 \in \Omega, \forall \check{v}_h \in \check{S}_h^\circ(\Omega).$$

We can define $|\check{E}_h(t)|$ to be the linear operator on $L^q(\Omega)$ with the following representation

$$(|\check{E}_h(t)|v)(x_0) = \int_\Omega a_h(x)|\Gamma_h(\mathbf{\Phi}_h^{-1}x, \mathbf{\Phi}_h^{-1}x_0, t)|v(x)dx \quad \forall x_0 \in \Omega, \forall v \in L^q(\Omega),$$

then we have the relation

$$(|\check{E}_h(t)|v) \circ \mathbf{\Phi}_h = |E_h(t)|(v \circ \mathbf{\Phi}_h) \quad \forall v \in L^q(\Omega).$$

From the discussion above, the main Theorem 5.2.1 can be equivalently expressed in the following form.

Theorem 5.3.1. *Let Ω be a curvilinear polyhedral domain in $\mathbb{R}^3$ (possibly with edge openings larger than π), and let $\check{S}_h^\circ(\Omega) \subseteq H_0^1(\Omega)$ be the finite element spaces defined in* (5.3.34)*. Assume that h is sufficiently small. Then for the semi-discrete equation* (5.3.33)*, we have the following analytic semigroup estimate:*

$$\sup_{t>0} \left(\|\check{E}_h(t)\check{v}_h\|_{L^q(\Omega)} + t\|\partial_t(\check{E}_h(t)\check{v}_h)\|_{L^q(\Omega)} \right) \leq C\|\check{v}_h\|_{L^q(\Omega)} \quad \forall \check{v}_h \in \check{S}_h^\circ(\Omega_h), \forall 1 \leq q \leq \infty, \tag{5.3.40a}$$

$$\| \sup_{t>0} |\check{E}_h(t)||v|\|_{L^q(\Omega)} \leq C_q\|v\|_{L^q(\Omega)}, \quad \forall v \in L^q(\Omega), \ \forall 1 < q \leq \infty. \tag{5.3.40b}$$

Further, if $\check{u}_{h,0} = 0$ and $\check{f}_h \in L^p(0,T; L^q(\Omega))$, then the solution $\check{u}_h(t)$ of equation (5.3.33) *possesses the following maximal L^p-regularity estimate:*

$$\|\partial_t\check{u}_h\|_{L^p(0,T;L^q(\Omega))} + \|\check{\Delta}_h u_h\|_{L^p(0,T;L^q(\Omega))} \leq \max(p,(p-1)^{-1})C_q\|\check{f}_h\|_{L^p(0,T;L^q(\Omega))} \tag{5.3.41a}$$
$$\forall 1 < p, q < \infty,$$

$$\|\partial_t\check{u}_h\|_{L^\infty(0,T;L^q(\Omega))} + \|\check{\Delta}_h\check{u}_h\|_{L^\infty(0,T;L^q(\Omega))} \leq C\ell_h\|\check{f}_h\|_{L^\infty(0,T;L^q(\Omega))}, \tag{5.3.41b}$$
$$\forall 1 \leq q \leq \infty,$$

where $\ell_h := \log_2(2+1/h)$. The constant C in (5.3.40a) *and* (5.3.41b) *is independent of $\check{f}_h, p, q, h$ and T, and the constant C_q in* (5.3.40b) *and* (5.3.41a) *is independent of $\check{f}_h, h, p$ and T.*

Theorem 5.3.1 is proved in Section 5.4.2 and Section 5.4.3.

Properties of the isoparametric finite element space

For any subregion $D \subseteq \Omega$, we define $\check{S}_h^\circ(D)$ as the subspace of $\check{S}_h^\circ(\Omega)$ consisting of functions that equal zero outside of D. For a given subset $D \subseteq \Omega$, denote $B_d(D) := \{x \in \Omega : \mathrm{dist}(x, D) \leq d\}$ for $d \geq 0$. There exists positive constants C and c_0 such that the finite element space $\check{S}_h(\Omega)$ possesses the following properties, independent of the subset D and mesh size h:

(P1) Quasi-uniformity For all $\check{K} \in \check{\mathcal{T}}_h$ (we make the convention that an element K denotes a closed domain), the following analogue properties of quasi-uniform triangle/tetrahedron meshes hold:

$$\mathrm{diam}(\check{K}) \leq h \text{ and } |\check{K}| \geq c_0^{-1}h^N,$$
$$\#\{\check{K}' \in \check{\mathcal{T}}_h : \check{K}' \cap \check{K} \neq \emptyset\} \leq c_0.$$

(P2) Inverse property If D is a union of elements in partition $\check{\mathcal{T}}_h$, then

$$\|\check{\chi}_h\|_{W^{l,p}(D)} \leq Ch^{-(l-k)-(N/q-N/p)}\|\check{\chi}_h\|_{W^{k,q}(D)} \quad \forall \check{\chi}_h \in \check{S}_h(\Omega), \tag{5.3.42}$$

for $0 \leq k \leq l \leq 1$ and $1 \leq q \leq p \leq \infty$.

(P3) Local appoximation and superapproximation There exists a quasi-interpolation operator $\check{I}_h : H_0^1(\Omega) \to \check{S}_h^\circ(\Omega)$ with the following properties:

1. for $\forall v \in H^{1+\alpha}(\Omega) \cap H_0^1(\Omega),\ \alpha \in [0,1]$, the following error estimate holds:

$$\|v - \check{I}_h v\|_{L^2(\Omega)} + h\|\nabla(v - \check{I}_h v)\|_{L^2(\Omega)} \leq Ch^{1+\alpha}\|v\|_{H^{1+\alpha}(\Omega)}, \tag{5.3.43}$$

2. If $d \geq 2h$, then the value of $\check{I}_h v$ in D depends only on the value of v in $B_d(D)$. If $d \geq 2h$ and $\mathrm{supp}(v) \subset \bar{D}$, then $\check{I}_h v \in \check{S}_h^\circ(B_d(D))$.
3. If $d \geq 2h$, $\omega = 0$ outside D and $|\partial^\beta \omega| \leq Cd^{-|\beta|}$ for all multi-index β, then for any $\check{\psi}_h \in \check{S}_h^\circ(\Omega)$, $\check{I}_h(\omega\check{\psi}_h) \in \check{S}_h^\circ(B_d(D))$ and

$$\|\omega\check{\psi}_h - \check{I}_h(\omega\check{\psi}_h)\|_{L^2(\Omega)} + h\|\omega\check{\psi}_h - \check{I}_h(\omega\check{\psi}_h)\|_{H^1(\Omega)} \leq Chd^{-1}\|\check{\psi}_h\|_{L^2(B_d(D))}. \tag{5.3.44}$$

By (5.3.43), Property (P3)-(2), and our definition of Sobolev spaces (5.3.23), we have the following estimate for $\alpha \in [0,1]$:

$$\begin{aligned} \|v - \check{I}_h v\|_{L^2(D)} + h\|v - \check{I}_h v\|_{H^1(D)} &\leq Ch^{1+\alpha}\|v\|_{H^{1+\alpha}(B_d(D))} \\ &\forall v \in H^{1+\alpha}(B_d(D)) \cap H_0^1(\Omega). \end{aligned} \tag{5.3.45}$$

The properties (P1)-(P2) can be directly verified via employing the mesh-regularity condition (5.3.27) and the norm equivalence property (5.3.30)-(5.3.31) of the Lipschitz homeomorphism $\mathbf{\Phi}_h$. The operator $\check{I}_h$ in (P3) can be constructed by the same method as in [101, Appendix B], for the reader's convenience we demonstrate (P3) in Section 5.7 of our manuscript.

5.3.3 Green's functions

For any $x_0 \in \check{K} \in \check{\mathcal{T}}_h$, using the mesh regularity estimate (5.3.29), we can mimic the proof of [133, Lemma 2.2] to obtain a function $\check{\delta}_{x_0} \in C^{r+1}(\bar{\Omega})$ with $\mathrm{supp}\check{\delta}_{x_0} \subseteq \check{K}$ and $\mathrm{dist}(\mathrm{supp}\check{\delta}_{x_0}, \partial\check{K}) \geq k_0 h$ (k_0 is a positive constant independent of h) such that

$$\check{\chi}_h(x_0) = \int_\Omega a_h(x)\check{\delta}_{x_0}\check{\chi}_h dx \quad \forall \check{\chi}_h \in \check{S}_h(\Omega), \tag{5.3.46}$$

and

$$\|\check{\delta}_{x_0}\|_{W^{l,p}(\Omega)} \leq Ch^{-l-N(1-1/p)} \quad \forall 1 \leq p \leq \infty, 0 \leq l \leq r+1, \tag{5.3.47a}$$

$$\sup_{y\in\Omega}\int_\Omega |\check{\delta}_y(x)|dx + \sup_{x\in\Omega}\int_\Omega |\check{\delta}_y(x)|dx \leq C, \tag{5.3.47b}$$

$$\int_\Omega a_h(x)\check{\delta}_{x_0}dx = 1. \tag{5.3.47c}$$

Let $\check\delta_{h,x_0} := \check P_h \check\delta_{x_0} \in \check S_h^\circ(\Omega)$ be the weighted $L^2(\Omega)$ projection of $\check\delta_{x_0}$. In view of (5.3.35), (5.3.47a) and properties (P1)-(P3) of $\check S_h^\circ(\Omega)$, the same proof as in [136, Lemma 7.2] shows that there exists constant $C > 0$ independent of h such that

$$|\check\delta_{h,x_0}(x)| \le Ch^{-N} e^{-\frac{|x-x_0|}{Ch}} \quad \forall x, x_0 \in \Omega. \tag{5.3.48}$$

Since $\check\Delta_h \check P_h \check\delta_{x_0} \in \check S_h^\circ(\Omega)$, (5.3.46) and (5.3.35) imply that for any $y \in \Omega$,

$$\left(\check\Delta_h \check P_h \check\delta_{x_0}\right)(y) = (a_h(x)\check\Delta_h \check P_h \check\delta_{x_0}, \check P_h \check\delta_y) = (A_h(x)\nabla\check\delta_{h,x_0}, \nabla\check\delta_{h,y}) \tag{5.3.49}$$

From the exponential decay property (5.3.48), using inverse estimate (5.3.42), one can show that $\nabla\check\delta_{h,x_0}$ also possesses exponential decay property i.e

$$|\nabla\check\delta_{h,x_0}|(x) \le Ch^{-N-1} e^{-\frac{|x-x_0|}{Ch}}.$$

Thus by formula (5.3.49) above, one can deduce that the discrete Laplacian of $\check\delta_{h,x_0}$ also possesses exponential decay property:

$$|\check\Delta_h \check\delta_{h,x_0}|(y) \le Ch^{-2N-2} \int_\Omega e^{-\frac{|x-x_0|+|x-y|}{Ch}} dx \le Ch^{-N-2} e^{-\frac{|y-x_0|}{Ch}} \quad \forall y, x_0 \in \Omega. \tag{5.3.50}$$

Let $G(t, x, x_0)$ denote the Green's function of the parabolic equation,

$$\partial_t G(\cdot,\cdot,x_0) - \Delta G(\cdot,\cdot,x_0) = 0 \quad \text{in } (0,\infty)\times\Omega, \tag{5.3.51a}$$
$$G(\cdot,\cdot,x_0) = 0 \quad \text{on } (0,\infty)\times\partial\Omega, \tag{5.3.51b}$$
$$G(0,\cdot,x_0) = \delta_{x_0} \quad \text{in } \Omega, \tag{5.3.51c}$$

where δ_{x_0} is the Dirac delta function centered at x_0. The Green's function $G(t, x, y)$ is symmetric with respect to x and y and satisfies the following Gaussian pointwise estimate for the time derivatives (cf. [101, (3.12)])

$$|\partial_t^k G(t,x,x_0)| \le \frac{C_k}{t^{k+N/2}} e^{-\frac{|x-x_0|^2}{C_k t}} \quad \forall x, x_0 \in \Omega \quad \forall t > 0, \quad k = 0, 1, 2, \cdots. \tag{5.3.52}$$

Let $\Gamma = \Gamma(t, x, x_0)$ be the regularized Green's function of the parabolic equation, defined by

$$\partial_t \Gamma(\cdot,\cdot,x_0) - \Delta\Gamma(\cdot,\cdot,x_0) = 0 \quad \text{in } (0,\infty)\times\Omega, \tag{5.3.53a}$$
$$\Gamma(\cdot,\cdot,x_0) = 0 \quad \text{on } (0,\infty)\times\partial\Omega, \tag{5.3.53b}$$
$$\Gamma(0,\cdot,x_0) = \check\delta_{x_0} \quad \text{in } \Omega. \tag{5.3.53c}$$

Let $\check\Gamma_h = \check\Gamma_h(t, x, x_0)$ be the finite element approximation of Γ, defined by

$$(a_h(\cdot)\partial_t\check\Gamma_h(t,\cdot,x_0), \check v_h)_\Omega + (A_h(\cdot)\nabla\check\Gamma_h(t,\cdot,x_0), \nabla\check v_h)_\Omega = 0 \quad \forall \check v_h \in \check S_h^0(\Omega), t \in (0,\infty), \tag{5.3.54a}$$
$$\check\Gamma_h(0,\cdot,x_0) = \check\delta_{h,x_0}. \tag{5.3.54b}$$

By using the Green's function, the solution $u(x,t)$ of the heat equation (5.1.1) has the formula:

$$u(t,x_0) = \int_\Omega G(t,x,x_0)u_0(x)dx + \int_0^t \int_\Omega G(t-s,x,x_0)f(s,x)dxds. \tag{5.3.55}$$

Notice that the discrete Laplacian operator $\check{\Delta}_h$ is self-adjoint with respect to the weighted $L^2(\Omega)$ inner product on $\check{S}_h^\circ(\Omega)$, i.e.,

$$(a_h(x)\check{\Delta}_h\check{u}_h, \check{v}_h)_\Omega = (\check{u}_h, a_h(x)\check{\Delta}_h v_h)_\Omega.$$

It follows that $\check{E}_h(t)$ is self-adjoint with respect to the weighted $L^2(\Omega)$ inner product on $\check{S}_h^\circ(\Omega)$. Since $\check{\Gamma}_h(t,\cdot,x_0) = \check{E}_h(t)\check{\delta}_{h,x_0}$ by the definition of $\check{E}_h(t)$, we have

$$\begin{aligned}\check{\Gamma}_h(t,x,x_0) =&(\check{E}_h(t)\check{\delta}_{h,x_0})(x) = (a_h\check{E}_h(t)\check{\delta}_{h,x_0}, \check{\delta}_{h,x})\\ =&(a_h\check{\delta}_{h,x_0}, \check{E}_h(t)\check{\delta}_{h,x}) = (\check{E}_h(t)\check{\delta}_{h,x})(x_0) = \check{\Gamma}_h(t,x_0,x),\end{aligned} \tag{5.3.56}$$

where we used the fact that $\check{v}_h(x_0) = (a_h\check{v}_h, \check{\delta}_{x_0}) = (a_h\check{v}_h, \check{\delta}_{h,x_0})$ for all $v_h \in \check{S}_h^\circ(\Omega)$, which follows from (5.3.46) and (5.3.35). Therefore, $\check{\Gamma}_h(t,x,y)$ is symmetric with respect to x and y. Moreover, $\check{\Gamma}_h$ gives the kernel of the discrete semigroup $\check{E}_h(t)$:

$$\begin{aligned}(\check{E}_h(t)\check{v}_h)(x_0) &= (a_h\check{E}_h(t)\check{v}_h, \check{\delta}_{h,x_0})_\Omega = (a_h\check{v}_h, \check{E}_h(t)\check{\delta}_{h,x_0})_\Omega\\ &= \int_\Omega a_h(x)\check{\Gamma}_h(t,x,x_0)\check{v}_h(x)dx \quad \forall \check{v}_h \in \check{S}_h^\circ(\Omega),\end{aligned} \tag{5.3.57}$$

and the solution $\check{u}_h(t,x_0)$ of (5.3.39) can be represented by

$$\check{u}_h(t,x_0) = \int_\Omega a_h(x)\check{\Gamma}_h(t,x,x_0)\check{u}_{h,0}(x)dx + \int_0^t\int_\Omega a_h(x)\check{\Gamma}_h(t-s,x,x_0)\check{f}_h(s,x)dxds, \tag{5.3.58}$$

and we have

$$(|\check{E}_h(t)|v)(x_0) = \int_\Omega a_h(x)|\check{\Gamma}_h(t,x,x_0)|v(x)dx \quad \forall v \in L^q(\Omega). \tag{5.3.59}$$

The regularized Green's function can be represented as follows:

$$\Gamma(t,x,x_0) = \int_\Omega G(t,y,x)\check{\delta}_{x_0}(y)dy = \int_\Omega G(t,x,y)\check{\delta}_{x_0}(y)dy. \tag{5.3.60}$$

From the representation (5.3.60) and estimate (5.3.52), the regularized Green's function Γ also satisfies the Gaussian pointwise estimate:

$$|\partial_t^k\Gamma(t,x,x_0)| \le \frac{C_k}{t^{k+N/2}}e^{-\frac{|x-x_0|^2}{C_k t}} \quad \forall x,x_0 \in \Omega \quad \forall t > 0 \text{ such that } \max(|x-x_0|,\sqrt{t}) \ge 2h, \tag{5.3.61}$$

with $k = 0,1,2,\cdots$.

5.3.4 Dyadic decomposition of the domain $\mathcal{Q} = (0,1) \times \Omega$

We will employ the same dyadic decomposition method as used in [101] to prove Theorem 5.3.1. Readers who are already familiar with this method may choose to skip this subsection. The dyadic decomposition method, originally introduced in [126], has been widely utilized by various authors [104, 66, 95, 100, 133].

For any positive integer j, we define $d_j = 2^{-j}$. For a given point $x_0 \in \Omega$, we set $J_1 = 1$, $J_0 = 0$, and let J_* be the integer satisfying $2^{-J_*} = C_*h$ where C_* is a constant with $C_* \ge 16$, to be determined later. If the condition

$$h < \frac{1}{4C_*} \tag{5.3.62}$$

is satisfied, then

$$2 \le J_* \le \log_2(2 + 1/h) = \ell_h.$$

In the following manuscript, for conclusions related to the dyadic decomposition, we will assume that h is sufficiently small to satisfy condition (5.3.62). Let

$$\begin{aligned} Q_*(x_0) &= \{(t,x) \in \mathcal{Q} : \max(|x - x_0|, t^{1/2}) \le d_{J_*}\}, \\ \Omega_*(x_0) &= \{x \in \Omega : |x - x_0| \le d_{J_*}\}. \end{aligned}$$

we define

$$\begin{aligned} Q_j(x_0) &= \{(t,x) \in \mathcal{Q} : d_j \le \max(|x - x_0|, t^{1/2}) \le 2d_j\} \quad \text{for } j \ge 1, \\ \Omega_j(x_0) &= \{x \in \Omega : d_j \le |x - x_0| \le 2d_j\} \quad \text{for } j \ge 1, \\ D_j(x_0) &= \{x \in \Omega : |x - x_0| \le 2d_j\} \quad \text{for } j \ge 1, \end{aligned}$$

and

$$\begin{aligned} Q_0(x_0) &= \mathcal{Q} \setminus \left(\bigcup_{j=1}^{J_*} Q_j(x_0) \cup Q_*(x_0) \right) \\ \Omega_0(x_0) &= \Omega \setminus \left(\bigcup_{j=1}^{J_*} \Omega_j(x_0) \cup \Omega_*(x_0) \right). \end{aligned}$$

For $j < 0$, we define $Q_j(x_0) = \Omega_j(x_0) = \emptyset$. For all integer $j \ge 0$, we define:

$$\begin{aligned} \Omega'_j(x_0) &= \Omega_{j-1}(x_0) \cup \Omega_j(x_0) \cup \Omega_{j+1}(x_0) \\ Q'_j(x_0) &= Q_{j-1}(x_0) \cup Q_j(x_0) \cup Q_{j+1}(x_0) \\ \Omega''_j(x_0) &= \Omega_{j-2}(x_0) \cup \Omega'_j(x_0) \cup \Omega_{j+2}(x_0) \\ Q''_j(x_0) &= Q_{j-2}(x_0) \cup Q'_j(x_0) \cup Q_{j+2}(x_0) \\ D'_j(x_0) &= D_j(x_0) \cup \Omega_{j-1}(x_0), \quad Q'_*(x_0) = Q_*(x_0) \cup Q_{J_*}(x_0) \\ D''_j(x_0) &= D'_j(x_0) \cup \Omega_{j-2}(x_0), \quad Q''_*(x_0) = Q'_*(x_0) \cup Q_{J_*-1}(x_0) \end{aligned}$$

Then we have

$$\mathcal{Q} = \bigcup_{j=0}^{J_*} Q_j(x_0) \cup Q_*(x_0) \quad \Omega = \bigcup_{j=0}^{J_*} \Omega_j(x_0) \cup \Omega_*(x_0)$$

We refer to $Q_*(x_0)$ as the "innermost" set. We will use the notation $\sum_{*,j}$ to indicate that the innermost set is included, and $\sum_j$ when it is not. When x_0 is fixed and there is no ambiguity, we will simplify the notation by writing $Q_j = Q_j(x_0)$, $\Omega_j = \Omega_j(x_0)$, $Q'_j = Q'_j(x_0)$, and $\Omega'_j = \Omega'_j(x_0)$.

We will use the following notations

$$\|v\|_{k,D} = \left(\int_D \sum_{|\alpha|=k} |\partial^\alpha v|^2 dx \right)^{1/2}, \quad |||v|||_{k,D} = \left(\int_Q \sum_{|\alpha|=k} |\partial^\alpha v|^2 dx dt \right)^{1/2} \tag{5.3.63}$$

for any subdomains $D \subseteq \Omega$ and $Q \subseteq \mathcal{Q}$, where ∂^α denotes the derivative in x with respect to the multi-index α. Throughout this manuscript, we use C to represent a generic positive constant that is independent of h, x_0 and C_* (until C^* is determined in Section 5.5). To simplify notations, we also denote $d_* = d_{J_*}$.

5.4 Proofs of the main results

5.4.1 Estimates of Green's function

In the proof below, we will apply the following lemma (cf. [101, Lemma 4.3]) which is a consequence of the elliptic regularity estimate in Ω.

Lemma 5.4.1. *Let Ω be a curvilinear polyhedral domain in $\mathbb{R}^N$ (with edge openings possibly larger than π). Then there exists $\alpha \in (1/2, 1]$ and constant C such that*

$$\|u\|_{H^{1+\alpha}} \leq C\|\nabla u\|_{L^2}^{1-\alpha}\|\Delta u\|_{L^2}^{\alpha} \quad \forall u \in H_0^1(\Omega) \text{ with } \Delta u \in L^2(\Omega). \tag{5.4.64}$$

Similar to [101, Lemma 4.1], we have the following estimates on the (regularized) Green's functions.

Lemma 5.4.2. *Let $\alpha \in (\frac{1}{2}, 1]$ be as in Lemma 5.4.1 and assume that condition* (5.3.62) *holds. There exists $C > 0$, independent of h and x_0, such that the Green's function G defined in (5.3.51) and the regularized Green's function Gamma defined in (5.3.53), satisfy the following estimates:*

$$\begin{aligned} &d_j^{-4-\alpha+N/2}\|\Gamma(\cdot,\cdot,x_0)\|_{L^\infty(Q_j(x_0))} + d_j^{-4-\alpha}\|\nabla\Gamma(\cdot,\cdot,x_0)\|_{L^2(Q_j(x_0))} \\ &\quad + d_j^{-4}\|\Gamma(\cdot,\cdot,x_0)\|_{L^2H^{1+\alpha}(Q_j(x_0))} + d_j^{-2}\|\partial_t\Gamma(\cdot,\cdot,x_0)\|_{L^2H^{1+\alpha}(Q_j(x_0))} \\ &\quad + \|\partial_{tt}\Gamma(\cdot,\cdot,x_0)\|_{L^2H^{1+\alpha}(Q_j(x_0))} \leq Cd_j^{-N/2-4-\alpha}, \end{aligned} \tag{5.4.65a}$$

$$\begin{aligned} &\|G(\cdot,\cdot,x_0)\|_{L^\infty H^{1+\alpha}(\cup_{k\leq j}Q_k(x_0))} \\ &\quad + d_j^2\|\partial_t G(\cdot,\cdot,x_0)\|_{L^\infty H^{1+\alpha}(\cup_{k\leq j}Q_k(x_0))} \leq Cd_j^{-N/2-1-\alpha}, \end{aligned} \tag{5.4.65b}$$

$$\begin{aligned} &d_j^4|||\partial_{ttt}\Gamma(\cdot,\cdot,x_0)|||_{Q_j(x_0)} + d_j^3|||\partial_{tt}\Gamma(\cdot,\cdot,x_0)|||_{1,Q_j(x_0)} + d_j^2|||\partial_{tt}\Gamma(\cdot,\cdot,x_0)|||_{Q_j(x_0)} \\ &\quad + d_j|||\partial_t\Gamma(\cdot,\cdot,x_0)|||_{1,Q_j(x_0)} + |||\partial_t\Gamma(\cdot,\cdot,x_0)|||_{Q_j(x_0)} + d_j^{-1}|||\Gamma(\cdot,\cdot,x_0)|||_{1,Q_j(x_0)} \\ &\quad + d_j^{-2}|||\Gamma(\cdot,\cdot,x_0)|||_{Q_j(x_0)} \leq Cd_j^{-1-N/2}. \end{aligned} \tag{5.4.65c}$$

Proof. Due to [101, (4.2)], (5.4.65b) is true. We note that there exists a $\check{K} \in \check{\mathcal{T}}_h$ such that $x_0 \in \check{K}$, and supp$\check{\delta}_{x_0}$ is contained in $\check{K}$. Therefore (5.3.60) implies

$$\Gamma(t,x,x_0) = \int_{\check{K}} G(t,y,x)\check{\delta}_{x_0}(y)dy = \int_{\check{K}} G(t,x,y)\check{\delta}_{x_0}(y)dy.$$

For (5.4.65a) and (5.4.65c), we can proceed exactly as in the proof of [101, Lemma 4.1]. Namely, we first establish the corresponding estimate for Green's function $G(\cdot,\cdot,x_0)$ using the local energy estimate and Gaussian pointwise estimate (5.3.52). Then, by applying (5.3.47b) and the identity above, we conclude that (5.4.65a) and (5.4.65c) also hold for $\Gamma(\cdot,\cdot,x_0)$. ∎

In addition to Lemma 5.4.2, we require the following critical lemma for the proof of Theorem 5.3.1. The proof of this lemma is deferred to Section 5.5.

Lemma 5.4.3. *There exists $h_0 > 0$ such that for any $0 < h < h_0$, the functions $\check{\Gamma}_h(t,x,x_0)$, $\Gamma(t,x,x_0)$, and $F(t,x,x_0) := \check{\Gamma}_h(t,x,x_0) - \Gamma(t,x,x_0)$ satisfy*

$$\sup_{t\in(0,\infty)} \left(\|\check{\Gamma}_h(t,\cdot,x_0)\|_{L^1(\Omega)} + t\|\partial_t\check{\Gamma}_h(t,\cdot,x_0)\|_{L^1(\Omega)}\right) \leq C, \tag{5.4.66a}$$

$$\sup_{t\in(0,\infty)} \left(\|\Gamma(t,\cdot,x_0)\|_{L^1(\Omega)} + t\|\partial_t\Gamma(t,\cdot,x_0)\|_{L^1(\Omega)}\right) \le C, \tag{5.4.66b}$$

$$\|\partial_t F(\cdot,\cdot,x_0)\|_{L^1((0,\infty)\times\Omega)} + \|t\partial_{tt}F(\cdot,\cdot,x_0)\|_{L^1((0,\infty)\times\Omega)} \le C, \tag{5.4.66c}$$

$$\|\partial_t\check{\Gamma}_h(t,\cdot,x_0)\|_{L^1(\Omega)} \le Ce^{-\lambda_0 t} \quad \forall t\ge 1, \tag{5.4.66d}$$

where the constants C and λ_0 are independent of h.

5.4.2 Proof of (5.3.40) in Theorem 5.3.1

According to (5.3.57) and (5.4.66a), we have

$$\begin{aligned}
&|(\check{E}_h(t)\check{v}_h)(x_0)| + |(t\partial_t\check{E}_h(t)\check{v}_h)(x_0)| \\
\le&\|a_h\|_{L^\infty(\Omega)}\left(\|\check{\Gamma}_h(t,\cdot,x_0)\|_{L^1(\Omega)} + t\|\partial_t\check{\Gamma}_h(t,\cdot,x_0)\|_{L^1(\Omega)}\right)\|\check{v}_h\|_{L^\infty(\Omega)} \\
\le&C\|\check{v}_h\|_{L^\infty(\Omega)} \quad \forall t>0, \forall \check{v}_h\in\check{S}_h^\circ(\Omega).
\end{aligned}$$

Therefore, (5.3.40a) is proven for $q=\infty$, if $h>0$ is small enough. The case $q=2$ follows from energy estimate. Thus, the general case $1\le q\le\infty$ follows from interpolation and duality (the operators $\check{E}_h(t)$ and $\partial_t\check{E}_h(t)$ are self-adjoint w.r.t the weighted L^2 inner product). This completes the proof of (5.3.40a).

In order to prove (5.3.40b), we need a symmetrically truncated Green's function $G^*_{\rm tr}$ as used in [101, Section 4.2] (see also [100, 104]). $G^*_{\rm tr}$ satisfies the following conditions:

$$G^*_{\rm tr}(t,x,y) \text{ is symmetric with respect to } x \text{ and } y, \text{ namely, } G^*_{\rm tr}(t,x,y) = G^*_{\rm tr}(t,y,x). \tag{5.4.67a}$$

$$G^*_{\rm tr}(\cdot,\cdot,y) = 0 \text{ in } Q_*(y) := \{(t,x)\in\mathcal{Q} : \max(|x-y|,\sqrt{t})\le d_*\}, \tag{5.4.67b}$$
$$\text{and } G^*_{\rm tr}(0,\cdot,y)\equiv 0 \text{ in } \Omega.$$

$$0\le G^*_{\rm tr}(t,x,y)\le G(t,x,y) \text{ and } G^*_{\rm tr}(t,x,y) = G(t,x,y) \tag{5.4.67c}$$
$$\text{when } \max(|x-y|,\sqrt{t}) > 2d_*.$$

$$|\partial_t G^*_{\rm tr}(t,x,y)| \le Cd_*^{-N-2} \text{ when } \max(|x-y|,\sqrt{t})\le 2d_*. \tag{5.4.67d}$$

Using the same reasoning as in [101, (4.32)—(4.37)] and in view of (5.4.66c), to establish (5.3.40b), it suffices to prove

$$\int\!\!\int_{(0,\infty)\times\Omega} |\partial_t\Gamma(t,x,x_0) - \partial_t G^*_{\rm tr}(t,x,x_0)|dxdt \le C. \tag{5.4.68}$$

Let $\check{K}\in\check{\mathcal{T}}_h$ such that $x_0\in\check{K}$ and supp$\check{\delta}_{x_0}$ is contained in $\check{K}$, and we denote $Q_{2*}(x_0) := \{(t,x)\in\mathcal{Q} : \max(|x-x_0|,\sqrt{t}) < 2d_*\}$. Then, by (5.4.67c), (5.3.47c) and (5.3.60), we have

$$\begin{aligned}
&\int\!\!\int_{[(0,\infty)\times\Omega]\backslash Q_{2*}(x_0)} |\partial_t\Gamma(t,x,x_0) - \partial_t G^*_{\rm tr}(t,x,x_0)|dxdt \\
=&\int\!\!\int_{[(0,\infty)\times\Omega]\backslash Q_{2*}(x_0)} |\partial_t\Gamma(t,x,x_0) - \partial_t G(t,x,x_0)|dxdt \\
\le&\int\!\!\int_{[(0,1)\times\Omega]\backslash Q_{2*}(x_0)} \int_{\check{K}} |\partial_t G(t,x,y) - \partial_t G(t,x,x_0)|\,|\check{\delta}_{x_0}(y)|dydxdt \\
&+\int\!\!\int_{[(0,1)\times\Omega]\backslash Q_{2*}(x_0)} \int_{\check{K}} \left|(a_h(y)-1)\partial_t G(t,x,x_0)\check{\delta}_{x_0}(y)\right| dydxdt
\end{aligned}$$

$$
\begin{aligned}
&+\int\int_{(1,\infty)\times\Omega} |\partial_t\Gamma(t,x,x_0)-\partial_t G(t,x,x_0)|dxdt\\
\le& C\int\int_{[(0,1)\times\Omega]\backslash Q_{2*}(x_0)} h^{\alpha-(N-2)/2}|\partial_t G(t,x,\cdot)|_{C^{\alpha-(N-2)/2}(\check{K})}dxdt\\
&+C\int\int_{[(0,1)\times\Omega]\backslash Q_{2*}(x_0)} h^{r}\|\partial_t G(t,x,\cdot)\|_{L^\infty(\check{K})}dxdt\\
&+\int\int_{(1,\infty)\times\Omega} |\partial_t\Gamma(t,x,x_0)-\partial_t G(t,x,x_0)|dxdt\\
\le& C\int\int_{[(0,1)\times\Omega]\backslash Q_{2*}(x_0)} h^{\alpha-(N-2)/2}\|\partial_t G(t,x,\cdot)\|_{C^{\alpha-(N-2)/2}(\check{K})}dxdt\\
&\quad+\int\int_{(1,\infty)\times\Omega} |\partial_t\Gamma(t,x,x_0)-\partial_t G(t,x,x_0)|dxdt\\
=:&\mathcal{I}_1+\mathcal{I}_2,
\end{aligned}
$$

where we have used (5.3.47b) and (5.3.32) in deriving the second inequality and used the fact $\alpha-(N-2)/2\le 1$ in deriving the third inequality. In view of (5.3.61) and (5.3.52), we can conclude that $\mathcal{I}_2\le C$. The inequality $\mathcal{I}_1\le C$ was demonstrated in the proof of [101, estimate (4.30)]. Furthermore, by applying the basic energy estimate as in [101, estimate (4.31)] and considering (5.4.67d), we have

$$
\int\int_{Q_{2*}(x_0)} |\partial_t\Gamma(t,x,x_0)-\partial_t G^*_{\mathrm{tr}}(t,x,x_0)|dxdt\le C. \tag{5.4.69}
$$

This establishes (5.4.68) and completes the proof of (5.3.40b).

5.4.3 Proof of (5.3.41) in Theorem 5.3.1

Since we have established the analytic estimate (5.3.40a) of $E_h(t)$, by applying the general theory of maximal regularity (cf. [42, Theorem 4.2], see also the proof in [101, Section 4.4]), it suffices to show (5.3.41a) for the case $p=q$. Let $\mathcal{E}_h$ denote the linear operator on $L^q(Q_T)$ defined by

$$
\mathcal{E}_h f(t):=\int_0^t \partial_t\check{E}_h(t-s)\check{P}_h f(s)ds. \tag{5.4.70}
$$

Thus, when $\check{u}_{h,0}=0$, we have $\check{\Delta}_h\check{u}_h=\mathcal{E}_h\check{f}_h$, which means that the maximal regularity estimate (5.3.41a) for the case $1<p=q<\infty$ is equivalent to the $L^q(Q_T)$-boundedness of the operator $\mathcal{E}_h$. Furthermore, since the discrete semigroup $\check{E}_h(t)$ is self-adjoint w.r.t. the weighted $L^2(\Omega)$ inner product, there holds:

$$
\begin{aligned}
\int_0^T (a_h\mathcal{E}_h f(t),g(t))dt&=\int_0^T\int_0^t (a_h\partial_t\check{E}_h(t-s)\check{P}_h f(s),g(t))dsdt\\
&=\int_0^T\int_0^t (a_h f(s),\partial_t\check{E}_h(t-s)\check{P}_h g(t))dsdt\\
&=\int_0^T\left(a_h f(s),\int_s^T \partial_t\check{E}_h(t-s)\check{P}_h g(t)dt\right)ds\\
&=\int_0^T (a_h f(s),\mathcal{E}_h g^*(T-s))\,ds
\end{aligned} \tag{5.4.71}
$$

where $g^*(t) := g(T-t)$. In view of (5.4.71), via duality, it suffices to prove that $\mathcal{E}_h$ is a bounded operator on $L^q(Q_T)$ for $2 \le q < \infty$. We decompose operator $\mathcal{E}_h$ as follows

$$
\begin{aligned}
\mathcal{E}_h f(t,x_0) &= \int_0^t \left(\partial_t \check{E}_h(t-s)\check{P}_h f(s)\right)(x_0) ds \\
&= \int_0^t \int_\Omega a_h(x)\partial_t \check{\Gamma}_h(t-s,x,x_0) f(s,x) dx ds \\
&= \int_0^t \int_\Omega a_h(x)\partial_t F(t-s,x,x_0) f(s,x) dx ds \\
&\quad + \int_0^t \int_\Omega a_h(x)\partial_t \Gamma(t-s,x,x_0) f(s,x) dx ds \\
&= \mathcal{M}_h f(t,x_0) + \mathcal{K}_h(a_h f)(t,x_0)
\end{aligned}
$$

where we denote by $\mathcal{M}_h$ and $\mathcal{K}_h$ the following operators:

$$
\begin{aligned}
\mathcal{K}_h g(t,x_0) &:= \int_0^t \int_\Omega \partial_t \Gamma(t-s,x,x_0) g(s,x) dx ds \\
\mathcal{M}_h g(t,x_0) &:= \int_0^t \int_\Omega a_h(x)\partial_t F(t-s,x,x_0) g(s,x) dx ds
\end{aligned}
$$

The same proof as in [101, (4.43)—(4.46)] yields that

$$
\|\mathcal{K}_h f\|_{L^q(Q_T)} \le C_q \|f\|_{L^q(Q_T)} \quad \forall 1 < q < \infty, \tag{5.4.72}
$$

where the constant C_q is independent of h and T. By the classical energy estimate, the result (5.3.41a) is true for $p = q = 2$. Combining this with the $L^2(Q_T)$-boundedness (5.4.72) of the operator $\mathcal{K}_h$, it follows that

$$
\|\mathcal{M}_h f\|_{L^2(Q_T)} \le C\|f\|_{L^2(Q_T)}. \tag{5.4.73}
$$

By (5.4.66c) of Lemma 5.4.3 we have

$$
\int_0^t \int_\Omega a_h(x)|\partial_t F(t-s,x,x_0)| dx ds \le C \int_0^\infty \int_\Omega |\partial_t F(s,x,x_0)| dx ds \le C,
$$

which implies

$$
\|\mathcal{M}_h \check{f}_h\|_{L^\infty(Q_T)} \le C\|\check{f}_h\|_{L^\infty(Q_T)}, \tag{5.4.74}
$$

and the interpolation of (5.4.73) and (5.4.74) yields

$$
\|\mathcal{M}_h \check{f}_h\|_{L^q(Q_T)} \le C\|\check{f}_h\|_{L^q(Q_T)} \quad \forall 2 \le q \le \infty.
$$

The estimate for $\mathcal{M}_h$ above, combined with the estimate for $\mathcal{K}_h$, establishes the boundedness of $\mathcal{E}_h$ on $L^q(Q_T)$ for $2 \le q < \infty$. Thus, the proof of (5.3.41a) is complete.

The proof of (5.3.41b) is exactly as in [101, Section 4.5]. We reduce to prove that the operator $\mathcal{E}_h$ satisfies the following estimate

$$
\|\mathcal{E}_h f\|_{L^\infty(0,T;L^q)} \le C\ell_h \|f\|_{L^\infty(0,T;L^q)} \quad \forall 1 \le q \le \infty. \tag{5.4.75}
$$

By the same deduction as in [101, (4.61),(4.62)], we can show that

$$
\|\mathcal{E}_h f\|_{L^\infty(0,T;L^q)} \le C\left(\int_0^\infty \sup_{x_0\in\Omega} \int_\Omega |\partial_t \check{\Gamma}_h(t,x,x_0)| dx dt\right) \|f\|_{L^\infty(0,T;L^q)} \quad \forall 1 \le q \le \infty. \tag{5.4.76}
$$

It remains to prove

$$\int_0^\infty \sup_{x_0\in\Omega}\int_\Omega |\partial_t\check{\Gamma}_h(t,x,x_0)|dxdt \le C\log_2(2+1/h). \tag{5.4.77}$$

To this end, note that $\partial_t\check{\Gamma}_h(t,\cdot,x_0) = \check{\Delta}_h\check{\Gamma}_h(t,\cdot,x_0) = \check{E}_h(t)\check{\Delta}_h\check{P}_h\check{\delta}_{x_0}$. By using (5.3.40a) of Theorem 5.3.1 (proved in Section 5.4.2) and (5.4.66a) of Lemma (5.4.3), we have

$$\|\partial_t\check{\Gamma}_h(t,\cdot,x_0)\|_{L^1} \le Ct^{-1}, \tag{5.4.78}$$

$$\|\partial_t\check{\Gamma}_h(t,\cdot,x_0)\|_{L^1} \le C\|\check{\Delta}_h\check{P}_h\check{\delta}_{x_0}\|_{L^1} \le Ch^{-2}\|\check{P}_h\check{\delta}_{x_0}\|_{L^1} \le Ch^{-2}, \tag{5.4.79}$$

where we have used inverse estimate, L^1-stability of $\check{P}_h$ and (5.3.47b). The interpolation of the last two inequalities gives (see [101, estimate (4.67)])

$$\int_0^1 \sup_{x_0\in\Omega}\int_\Omega |\partial_t\check{\Gamma}_h(t,x,x_0)|dxdt \le C\log_2(2+1/h). \tag{5.4.80}$$

While estimate (5.4.66d) implies

$$\int_1^\infty \sup_{x_0\in\Omega}\int_\Omega |\partial_t\check{\Gamma}_h(t,x,x_0)|dxdt \le C. \tag{5.4.81}$$

The last two inequalities combine to give (5.4.77), completing the proof of (5.3.41b).

5.4.4 Proof of Theorem 5.2.2

For a solution u_h with general initial value $u_{h,0}$, we denote u_h^* as the solution of (5.1.6) with $u_h^*(0) = P_h\widetilde{u}_0$ and $f_h = P_h\widetilde{f}$. Then, by the maximum norm stability (5.2.15a) of discrete semigroup $E_h(t)$, we have

$$\|u_h(t) - u_h^*(t)\|_{L^\infty(\Omega_h)} = \|E_h(t)(u_{h,0} - P_h\widetilde{u}_0)\|_{L^\infty(\Omega_h)} \le C\|u_{h,0} - P_h\widetilde{u}_0\|_{L^\infty(\Omega_h)}.$$

Combined with the L^∞-stability (5.3.36) of $L^2(\Omega_h)$-orthogonal projection P_h, this implies:

$$\|u_h - u_h^*\|_{L^\infty(0,T;L^\infty(\Omega_h))} \le C\|u_{h,0} - \widetilde{u}_0\|_{L^\infty(\Omega_h)}$$

Therefore, it suffices to prove the error estimate in the case where $u_{h,0} = P_h\widetilde{u}_0$. Denote $e_h := P_h\widetilde{u} - u_h$. We first derive the error equation satisfied by e_h. For each $v_h \in S_h^\circ(\Omega_h)$, we have:

$$\begin{aligned}
&(\partial_t P_h\widetilde{u}, v_h)_{\Omega_h} + (\nabla P_h\widetilde{u}, \nabla v_h)_{\Omega_h}\\
=&(\partial_t\widetilde{u}, v_h)_{\Omega_h} + (\nabla(P_h\widetilde{u} - R_h\widetilde{u}), \nabla v_h)_{\Omega_h} + (\nabla R_h\widetilde{u}, \nabla v_h)_{\Omega_h}\\
=&(\partial_t\widetilde{u}, v_h)_{\Omega_h} + (\nabla\widetilde{u}, \nabla v_h)_{\Omega_h} + (\nabla(P_h\widetilde{u} - R_h\widetilde{u}), \nabla v_h)_{\Omega_h} \quad \text{(definition (5.2.17) of } R_h \text{ used)}\\
=&(\partial_t\widetilde{u} - \Delta\widetilde{u} - \widetilde{f}, v_h)_{\Omega_h} + (\widetilde{f}, v_h)_{\Omega_h} + (\nabla(P_h\widetilde{u} - R_h\widetilde{u}), \nabla v_h)_{\Omega_h}\\
=&(\widetilde{\psi}, v_h)_{\Omega_h\setminus\Omega} + (\widetilde{f}, v_h)_{\Omega_h} + (\Delta_h(R_h\widetilde{u} - P_h\widetilde{u}), v_h)_{\Omega_h},
\end{aligned}$$

where we have denoted $\widetilde{\psi} := \partial_t\widetilde{u} - \Delta\widetilde{u} - \widetilde{f}$ and observed that $\widetilde{\psi}|_\Omega \equiv 0$. Thus, the error equation satisfied by e_h reads: $e_h(0) = 0$ and

$$(\partial_t e_h, v_h)_{\Omega_h} + (\nabla e_h, \nabla v_h)_{\Omega_h} = (\widetilde{\psi}, v_h)_{\Omega_h\setminus\Omega} + (\Delta_h(R_h\widetilde{u} - P_h\widetilde{u}), v_h)_{\Omega_h} \quad \forall v_h \in S_h^\circ(\Omega_h).$$

We split the error e_h as $e_h = \delta_h + \theta_h$ where θ_h is defined by equation: $\theta_h(0) = 0$ and

$$(\partial_t \theta_h, v_h)_{\Omega_h} + (\nabla \theta_h, \nabla v_h)_{\Omega_h} = (P_h \widetilde{\psi}, v_h)_{\Omega_h} \quad \forall v_h \in S_h^{\circ}(\Omega_h).$$

Then δ_h satisfies $\delta_h(0) = 0$ and

$$\partial_t \delta_h - \Delta_h \delta_h = \Delta_h (R_h \widetilde{u} - P_h \widetilde{u})$$

Using the quasi-maximal L^∞-L^∞ regularity estimate in (5.2.16b), and following the same argument as in [101, Corollary 2.2], we obtain the following result:

$$\|\delta_h\|_{L^\infty(0,T;L^\infty(\Omega_h))} = \|\Delta_h(\Delta_h^{-1}\delta_h)\|_{L^\infty(0,T;L^\infty(\Omega_h))} \le C\ell_h \|R_h \widetilde{u} - P_h \widetilde{u}\|_{L^\infty(0,T;L^\infty(\Omega_h))} \tag{5.4.82}$$

Next, we represent θ_h via the discrete Green's function Γ_h:

$$\begin{aligned} \theta_h(t,x) &= \int_0^t \int_{\Omega_h} \Gamma_h(t-s,y,x) P_h \widetilde{\psi}(s,y) dy ds \\ &= \int_0^t \int_{\Omega_h \setminus \Omega} \Gamma_h(t-s,y,x) \widetilde{\psi}(s,y) dy ds. \end{aligned}$$

Therefore, for each $0 < t < T$ and $x \in \Omega_h$,

$$|\theta_h(t,x)| \le \|\widetilde{\psi}\|_{L^\infty(0,T;L^\infty(\Omega_h))} \int_0^T \|\Gamma_h(t,y,x)\|_{L^1_y(\Omega_h \setminus \Omega)} dt.$$

From the boundary skin estimate [80, (3.3) of Lemma 3.2] and the fact $\Gamma_h|_{\partial\Omega_h} = 0$, we have:

$$\|\Gamma_h(t,\cdot,x)\|_{L^1(\Omega_h \setminus \Omega)} \le C h^{r+1} \|\nabla \Gamma_h(t,\cdot,x)\|_{L^1(\Omega_h)}$$

It follows that

$$\|\theta_h\|_{L^\infty(0,T;L^\infty(\Omega_h))} \le C h^{r+1} \|\widetilde{\psi}\|_{L^\infty(0,T;L^\infty(\Omega_h))} \sup_{x \in \Omega_h} \|\nabla \Gamma_h(\cdot,\cdot,x)\|_{L^1(0,T;L^1(\Omega_h))}. \tag{5.4.83}$$

It remains to prove the boundedness of $\|\nabla \Gamma_h(\cdot,\cdot,x)\|_{L^1(0,\infty;L^1(\Omega_h))}$. Since

$$\Gamma_h(t,x,y) = \check{\Gamma}_h(t, \mathbf{\Phi}_h(x), \mathbf{\Phi}_h(y))$$

and $\mathbf{\Phi}_h$ induces norm equivalence (5.3.31) in $W^{1,1}$-norm, it suffices to prove such boundedness result for $\check{\Gamma}_h$. We split the $L^1(0,\infty;W^{1,1})$-norm of $\check{\Gamma}_h$ as follows

$$\begin{aligned} \|\nabla \check{\Gamma}_h(\cdot,\cdot,x)\|_{L^1(0,\infty;L^1(\Omega))} \le & \|\nabla F(\cdot,\cdot,x)\|_{L^1(0,1;L^1(\Omega))} + \|\nabla \Gamma(\cdot,\cdot,x)\|_{L^1(0,1;L^1(\Omega))} \\ & + \|\nabla \check{\Gamma}_h(\cdot,\cdot,x)\|_{L^1(1,\infty;L^1(\Omega))} \end{aligned} \tag{5.4.84}$$

We can convert the L^1-norm into a summation of L^2-norms by the dyadic decomposition introduced in Section 5.3.4 and employ the local estimate (5.4.65a) and global estimate (5.5.92c) of Γ to obtain

$$\|\nabla \Gamma(\cdot,\cdot,x)\|_{L^1(\mathcal{Q})} \le C \sum_{j,*} d_j^{N/2+1} \|\nabla \Gamma\|_{L^2(Q_j(x))}$$

$$\leq C \sum_j d_j^{N/2+1} d_j^{-N/2} + C d_*^{N/2+1} h^{-N/2} \leq C. \tag{5.4.85}$$

Similarly, from (5.5.99) and (5.5.123) established in the next section, specifically there holds $\|\nabla F\|_{L^2(Q_j(x))} \leq C d_j^{-N/2}$. Combining with the global estimate (5.5.92c), we have:

$$\begin{aligned}\|\nabla F(\cdot,\cdot,x)\|_{L^1(\mathcal{Q})} \leq & C \sum_{j,*} d_j^{N/2+1} \|\nabla F\|_{L^2(Q_j(x))} \\ \leq & C \sum_j d_j^{N/2+1} d_j^{-N/2} + C d_*^{N/2+1} h^{-N/2} \leq C. \end{aligned} \tag{5.4.86}$$

Finally, from the exponential decay estimate (5.5.125), we can deduce by elliptic energy estimate that

$$\|\nabla \check{\Gamma}_h(t,\cdot,x)\|_{L^2(\Omega)} \leq C e^{-\lambda_0(t-1)} \quad \forall t \geq 1, \tag{5.4.87}$$

which yields the boundedness of $\|\nabla \check{\Gamma}_h(\cdot,\cdot,x)\|_{L^1(1,\infty;L^1(\Omega))}$. Combining estimates (5.4.85), (5.4.86) and (5.4.87), we proved that

$$\sup_{x\in\Omega_h} \|\nabla \Gamma_h(\cdot,\cdot,x)\|_{L^1(0,\infty;L^1(\Omega_h))} \leq C. \tag{5.4.88}$$

Therefore, summarizing the estimates (5.4.82, 5.4.83, 5.4.88), we obtain

$$\begin{aligned}\|P_h\widetilde{u} - u_h\|_{L^\infty(0,T;L^\infty(\Omega_h))} \leq & C h^{r+1} \|\widetilde{\psi}\|_{L^\infty(0,T;L^\infty(\Omega_h))} + C\ell_h \|R_h\widetilde{u} - P_h\widetilde{u}\|_{L^\infty(0,T;L^\infty(\Omega_h))} \\ \leq & C h^{r+1} \left(\|u\|_{L^\infty(0,T;W^{2,\infty}(\Omega))} + \|\partial_t u\|_{L^\infty(0,T;L^\infty(\Omega))}\right) \\ & + C\ell_h \|R_h\widetilde{u} - \widetilde{u}\|_{L^\infty(0,T;L^\infty(\Omega_h))}, \end{aligned} \tag{5.4.89}$$

where we have used the stability in Sobolev norms of the Stein extension operator and the L^∞-stability (5.3.36) of P_h in deducing the last inequality. The proof of Theorem 5.2.2 is complete.

5.5 Proof of Lemma 5.4.3

We use the following local energy error estimate for finite element solutions of parabolic equations.

Lemma 5.5.1. *Suppose that $\phi \in L^2(0,1;H_0^1(\Omega)) \cap H^1(0,1;L^2(\Omega))$ and $\check{\phi}_h \in H^1(0,1;\check{S}_h^\circ(\Omega))$ satisfy the equation*

$$(\partial_t\phi - a_h(x)\partial_t\check{\phi}_h, \check{\chi}_h) + (\nabla\phi - A_h(x)\nabla\check{\phi}_h, \nabla\check{\chi}_h) = 0, \tag{5.5.90}$$

for any $\check{\chi}_h \in \check{S}_h^\circ(\Omega)$, and $0 < t < 1$, with $\phi(0) = 0$ in Ω_j''. Then there exists $h_0 > 0$ such that for any $0 < h < h_0$,

$$\begin{aligned} & |||\partial_t(\phi - \check{\phi}_h)|||_{Q_j} + d_j^{-1} |||\phi - \check{\phi}_h|||_{1,Q_j} \\ \leq & C\epsilon^{-3} \left(I_j(\check{\phi}_h(0)) + X_j(\check{I}_h\phi - \phi) + Y_j(\phi) + d_j^{-2} |||\phi - \check{\phi}_h|||_{Q_j'} \right) \\ & + \left(C h^{1/2} d_j^{-1/2} + C\epsilon^{-1} h d_j^{-1} + \epsilon \right) \left(|||\partial_t(\phi - \check{\phi}_h)|||_{Q_j'} + d_j^{-1} |||\phi - \check{\phi}_h|||_{1,Q_j'} \right), \end{aligned} \tag{5.5.91}$$

where

$$\begin{aligned}
I_j(\check{\phi}_h(0)) =&\|\check{\phi}_h(0)\|_{1,\Omega_j'} + d_j^{-1}\|\check{\phi}_h(0)\|_{\Omega_j'},\\
X_j(\check{I}_h\phi-\phi) =&d_j\|\nabla\partial_t(\check{I}_h\phi-\phi)\|_{L^2(Q_j')} + \|\partial_t(\check{I}_h\phi-\phi)\|_{L^2(Q_j')}\\
&+ d_j^{-1}\|\nabla(\check{I}_h\phi-\phi)\|_{L^2(Q_j')} + d_j^{-2}\|\check{I}_h\phi-\phi\|_{L^2(Q_j')}\\
Y_j(\phi) =&h^r\left(d_j\|\nabla(\partial_t\phi)\|_{L^2(Q_j')} + \|\partial_t\phi\|_{L^2(Q_j')} + d_j^{-1}\|\nabla\phi\|_{L^2(Q_j')} + d_j^{-2}\|\phi\|_{L^2(Q_j')}\right).
\end{aligned}$$

Here $\epsilon \in (0,1)$ *is an arbitrary parameter, and the positive constant* C *is independent of* h*,* j*, and* C_**.*

The proof of Lemma 5.5.1 is presented in next Section. In the rest of this section, we apply Lemma 5.5.1 to prove Lemma 5.4.3 by setting $\alpha \in (\frac{1}{2}, 1]$ a fixed constant satisfying Lemma 5.4.2. The proof consists of three parts. The first part is concerned with estimates for $t \in (0,1)$, where we convert the L^1 estimates on $\mathcal{Q} = (0,1)\times\Omega = Q_* \cup (\cup_{j=0}^{J} Q_j)$ into weighted L^2 estimates on the subdomains Q_* and Q_j, $j = 0, 1, \cdots, J_*$. The second part is concerned with estimates for $t \geq 1$, which is a simple consequence of the parabolic regularity. The third part is concerned with the proof of (5.4.66a, 5.4.66b), which are simple consequences of the results proved in the first two parts.

Proof. Part I. First, we present estimates in the domain $\mathcal{Q} = (0,1)\times\Omega$ with the restriction $h < 1/(4C_*)$. In this case, the basic energy estimate gives

$$\|\partial_t\Gamma\|_{L^2(\mathcal{Q})} + \|\partial_t\check{\Gamma}_h\|_{L^2(\mathcal{Q})} \leq C\left(\|\Gamma(0)\|_{H^1(\Omega)} + \|\check{\Gamma}_h(0)\|_{H^1(\Omega)}\right) \leq Ch^{-1-N/2}, \tag{5.5.92a}$$

$$\|\Gamma\|_{L^\infty L^2(\mathcal{Q})} + \|\check{\Gamma}_h\|_{L^\infty L^2(\mathcal{Q})} \leq C\left(\|\Gamma(0)\|_{L^2(\Omega)} + \|\check{\Gamma}_h(0)\|_{L^2(\Omega)}\right) \leq Ch^{-N/2}, \tag{5.5.92b}$$

$$\|\nabla\Gamma\|_{L^2(\mathcal{Q})} + \|\nabla\check{\Gamma}_h\|_{L^2(\mathcal{Q})} \leq C\left(\|\Gamma(0)\|_{L^2(\Omega)} + \|\check{\Gamma}_h(0)\|_{L^2(\Omega)}\right) \leq Ch^{-N/2}, \tag{5.5.92c}$$

$$\|\partial_{tt}\check{\Gamma}_h\|_{L^2(\mathcal{Q})} \leq C\|\check{\Delta}_h\check{\Gamma}_h(0)\|_{H^1(\Omega)} \leq Ch^{-3-N/2}, \tag{5.5.92d}$$

$$\|\nabla\partial_t\Gamma\|_{L^2(\mathcal{Q})} + \|\nabla\partial_t\check{\Gamma}_h\|_{L^2(\mathcal{Q})} \leq C\left(\|\Delta\Gamma(0)\|_{L^2(\Omega)} + \|\check{\Delta}_h\check{\Gamma}_h(0)\|_{L^2(\Omega)}\right) \leq Ch^{-2-N/2}. \tag{5.5.92e}$$

In the estimate above, we have employed (5.3.47a), (5.3.48) and inverse properties (5.3.42) of finite element functions to deal with norms of $\Gamma(0) = \check{\delta}_{x_0}$ and $\check{\Gamma}_h(0) = \check{P}_h\check{\delta}_{x_0}$. We can decompose $\|\partial_t F\|_{L^1(\mathcal{Q})} + \|t\partial_{tt}F\|_{L^1(\mathcal{Q})}$ as follows:

$$\begin{aligned}
&\|\partial_t F\|_{L^1(\mathcal{Q})} + \|t\partial_{tt}F\|_{L^1(\mathcal{Q})} \qquad (5.5.93)\\
\leq&\|\partial_t F\|_{L^1(Q_*)} + \|t\partial_{tt}F\|_{L^1(Q_*)} + \sum_{j=0}^{J_*}\left(\|\partial_t F\|_{L^1(Q_j)} + \|t\partial_{tt}F\|_{L^1(Q_j)}\right).
\end{aligned}$$

We will bound the innermost part $\|\partial_t F\|_{L^1(Q_*)} + \|t\partial_{tt}F\|_{L^1(Q_*)}$ by separately bounding $\|\partial_t\check{\Gamma}_h\|_{L^1(Q_*)} + \|t\partial_{tt}\check{\Gamma}_h\|_{L^1(Q_*)}$ and $\|\partial_t\Gamma\|_{L^1(Q_*)} + \|t\partial_{tt}\Gamma\|_{L^1(Q_*)}$. By Hölder's inequality (noting the volume of $Q_j = Q_j(x_0)$ is Cd_j^{2+N}) and the global energy estimate (5.5.92a) and (5.5.92d), we have:

$$\begin{aligned}
\|\partial_t\check{\Gamma}_h\|_{L^1(Q_*)} + \|t\partial_{tt}\check{\Gamma}_h\|_{L^1(Q_*)} \leq&Cd_*^{N/2+1}\left(\|\partial_t\check{\Gamma}_h\|_{L^2(Q_*)} + d_*^2\|\partial_{tt}\check{\Gamma}_h\|_{L^2(Q_*)}\right)\\
\leq&Cd_*^{N/2+1}\left(\|\partial_t\check{\Gamma}_h\|_{L^2(\mathcal{Q})} + d_*^2\|\partial_{tt}\check{\Gamma}_h\|_{L^2(\mathcal{Q})}\right)\\
\leq&CC_*^{N/2+1} + CC_*^{N/2+3} \leq CC_*^{N/2+3}. \qquad (5.5.94)
\end{aligned}$$

For the term $\|t\partial_{tt}\Gamma\|_{L^1(Q_*)}$, since $\check{\delta}_{x_0}$ only belongs to $C^2(\overline{\Omega})$ when $r = 1$ (cf. (5.3.47a)), we utilize the analyticity of the parabolic semigroup $e^{t\Delta}$ to obtain

$$\|t\partial_{tt}\Gamma\|_{L^2} = \|t\partial_t e^{t\Delta}(\partial_t\Gamma(0))\|_{L^2} \le C\|\partial_t\Gamma(0)\|_{L^2} = C\|\Delta\check{\delta}_{x_0}\|_{L^2} \le Ch^{-N/2-2}. \tag{5.5.95}$$

Thus, by Hölder's inequality, (5.5.92a) and the last estimate (5.5.95), we have

$$\begin{aligned}\|\partial_t\Gamma\|_{L^1(Q_*)} + \|t\partial_{tt}\Gamma\|_{L^1(Q_*)} \le& Cd_*^{N/2+1}\|\partial_t\Gamma\|_{L^2(Q_*)} + Cd_*^{2+N/2}\|t\partial_{tt}\Gamma\|_{L^\infty L^2(Q_*)}\\ \le& CC_*^{N/2+1} + CC_*^{N/2+2} \le CC_*^{N/2+2}.\end{aligned} \tag{5.5.96}$$

It follows that

$$\begin{aligned}&\|\partial_t F\|_{L^1(\mathcal{Q})} + \|t\partial_{tt}F\|_{L^1(\mathcal{Q})} && (5.5.97)\\ \le& CC_*^{N/2+3} + \sum_{j=0}^{J_*}\left(\|\partial_t F\|_{L^1(Q_j)} + \|t\partial_{tt}F\|_{L^1(Q_j)}\right)\\ \le& CC_*^{N/2+3} + \sum_{j=0}^{J_*} Cd_j^{N/2+1}\left(\|\partial_t F\|_{L^2(Q_j)} + d_j^2\|\partial_{tt}F\|_{L^2(Q_j)}\right)\\ \le& CC_*^{N/2+3} + C\mathcal{K} && (5.5.98)\end{aligned}$$

where we have used Hölder's inequality to convert the sum of L^1-norms to a weighted sum of L^2-norms and introduced the notation:

$$\mathcal{K} := \sum_{j=0}^{J_*} d_j^{1+N/2}\left(d_j^{-1}|||F|||_{1,Q_j} + |||\partial_t F|||_{Q_j} + d_j|||\partial_t F|||_{1,Q_j} + d_j^2|||\partial_{tt}F|||_{Q_j}\right). \tag{5.5.99}$$

It remains to estimate $\mathcal{K}$. To this end, we set "$\check{\phi}_h = \check{\Gamma}_h, \phi = \Gamma, \check{\phi}_h(0) = \check{P}_h\check{\delta}_{x_0}$ and $\phi(0) = \check{\delta}_{x_0}$" and "$\check{\phi}_h = \partial_t\check{\Gamma}_h$, $\phi = \partial_t\Gamma$, $\check{\phi}_h(0) = \check{\Delta}_h\check{P}_h\check{\delta}_{x_0}$ and $\phi(0) = \Delta\check{\delta}_{x_0}$" in Lemma 5.5.1 respectively. The we obtain

$$\begin{aligned}&d_j^{-1}|||F|||_{1,Q_j'^{,1/2}} + |||\partial_t F|||_{Q_j'^{,1/2}} && (5.5.100)\\ \le& C\epsilon_1^{-3}\left(\widehat{I}_j + \widehat{X}_j + \widehat{Y}_j + d_j^{-2}|||F|||_{Q_j'}\right)\\ &+\left(Ch^{1/2}d_j^{-1/2} + C\epsilon_1^{-1}hd_j^{-1} + \epsilon_1\right)\left(d_j^{-1}|||F|||_{1,Q_j'} + |||\partial_t F|||_{Q_j'}\right)\end{aligned}$$

and

$$\begin{aligned}&d_j|||\partial_t F|||_{1,Q_j} + d_j^2|||\partial_{tt}F|||_{Q_j} && (5.5.101)\\ \le& C\epsilon_2^{-3}\left(\overline{I}_j + \overline{X}_j + \overline{Y}_j + |||\partial_t F|||_{Q_j'^{,1/2}}\right)\\ &+\left(Ch^{1/2}d_j^{-1/2} + C\epsilon_2^{-1}hd_j^{-1} + \epsilon_2\right)\left(d_j|||\partial_t F|||_{1,Q_j'^{,1/2}} + d_j^2|||\partial_{tt}F|||_{Q_j'^{,1/2}}\right),\end{aligned}$$

respectively, where $\epsilon_1, \epsilon_2 \in (0,1)$ are arbitrary constants and $Q_j'^{,1/2}$ is an intermediate subset between Q_j and Q_j':

$$Q_j'^{,1/2} := \{(t,x) \in \mathcal{Q} : \frac{3}{4}d_j \le \max(|x - x_0|, t^{1/2}) \le 3d_j\}.$$

Note that from the proof of Lemma 5.5.1, the pair (Q_j, Q'_j) in the statement of Lemma 5.5.1 could be replaced by the pair $(Q'^{,1/2}_j, Q'_j)$ or the pair $(Q_j, Q'^{,1/2}_j)$.

By using local interpolation error estimate (5.3.45), exponential decay property (5.3.48) and (5.3.50) of $\check{P}_h\check{\delta}_{x_0}$ and $\check{\Delta}_h\check{P}_h\check{\delta}_{x_0}$ respectively, and local estimates (5.4.65a) and (5.4.65c) of the regularized Green's function Γ, we have:

$$\widehat{I}_j = \|\check{P}_h\check{\delta}_{x_0}\|_{1,\Omega'_j} + d_j^{-1}\|\check{P}_h\check{\delta}_{x_0}\|_{\Omega'_j} \le Ch^2 d_j^{-3-N/2}, \tag{5.5.102}$$

$$\begin{aligned}\widehat{X}_j =& d_j\|\nabla(1-\check{I}_h)\partial_t\Gamma\|_{L^2(Q'_j)} + \|(1-\check{I}_h)\partial_t\Gamma\|_{L^2(Q'_j)} \\ &+ d_j^{-1}\|\nabla(1-\check{I}_h)\Gamma\|_{L^2(Q'_j)} + d_j^{-2}\|(1-\check{I}_h)\Gamma\|_{L^2(Q'_j)} \\ \le& C(d_jh^\alpha + h^{1+\alpha})\|\partial_t\Gamma\|_{L^2H^{1+\alpha}(Q''_j)} + C(d_j^{-1}h^\alpha + d_j^{-2}h^{1+\alpha})\|\Gamma\|_{L^2H^{1+\alpha}(Q''_j)} \\ \le& C(h^\alpha d_j^{-1-N/2-\alpha} + h^{1+\alpha}d_j^{-2-N/2-\alpha}) \le Ch^\alpha d_j^{-1-N/2-\alpha}\end{aligned} \tag{5.5.103}$$

$$\widehat{Y}_j = h^r\left(d_j|||\nabla(\partial_t\Gamma)|||_{Q'_j} + |||\partial_t\Gamma|||_{Q'_j} + d_j^{-1}|||\nabla\Gamma|||_{Q'_j} + d_j^{-2}|||\Gamma|||_{Q'_j}\right) \le Ch^r d_j^{-1-N/2}. \tag{5.5.104}$$

and

$$\overline{I}_j = d_j^2\|\check{\Delta}_h\check{P}_h\check{\delta}_{x_0}\|_{1,\Omega'^{,1/2}_j} + d_j\|\check{\Delta}_h\check{P}_h\check{\delta}_{x_0}\|_{\Omega'^{,1/2}_j} \le Ch^2 d_j^{-3-N/2}, \tag{5.5.105}$$

$$\begin{aligned}\overline{X}_j =& d_j^3\|\nabla(1-\check{I}_h)\partial_{tt}\Gamma\|_{L^2(Q'^{,1/2}_j)} + d_j^2\|(1-\check{I}_h)\partial_{tt}\Gamma\|_{L^2(Q'^{,1/2}_j)} \\ &+ d_j\|\nabla(1-\check{I}_h)\partial_t\Gamma\|_{L^2(Q'^{,1/2}_j)} + \|(1-\check{I}_h)\partial_t\Gamma\|_{L^2(Q'^{,1/2}_j)} \\ \le& C(d_j^3h^\alpha + d_j^2h^{1+\alpha})\|\partial_{tt}\Gamma\|_{L^2H^{1+\alpha}(Q'_j)} + C(d_jh^\alpha + h^{1+\alpha})\|\partial_t\Gamma\|_{L^2H^{1+\alpha}(Q'_j)} \\ \le& Ch^\alpha d_j^{-1-N/2-\alpha},\end{aligned} \tag{5.5.106}$$

$$\begin{aligned}\overline{Y}_j =& h^r d_j^2\left(d_j|||\nabla(\partial_{tt}\Gamma)|||_{Q'^{,1/2}_j} + |||\partial_{tt}\Gamma|||_{Q'^{,1/2}_j} + d_j^{-1}|||\nabla\partial_t\Gamma|||_{Q'^{,1/2}_j} + d_j^{-2}|||\partial_t\Gamma|||_{Q'^{,1/2}_j}\right) \\ \le& Ch^r d_j^{-1-N/2}.\end{aligned} \tag{5.5.107}$$

By choosing $\epsilon_1 = \epsilon^4$ in (5.5.100) and $\epsilon_2 = \epsilon$ in (5.5.101), substituting (5.5.100—5.5.107) into the expression of $\mathcal{K}$ in (5.5.99), and dealing with the term $C\epsilon^{-3}|||\partial_t F|||_{Q'^{,1/2}_j}$ appearing on the right side again by using (5.5.100), we have:

$$\begin{aligned}\mathcal{K} =& \sum_{j=0}^{J_*} d_j^{1+N/2}\left(d_j^{-1}|||F|||_{1,Q_j} + |||\partial_t F|||_{Q_j} + d_j|||\partial_t F|||_{1,Q_j} + d_j^2|||\partial_{tt}F|||_{Q_j}\right) \\ \le& C_\epsilon\sum_{j=0}^{J_*} d_j^{1+N/2}\left(h^2d_j^{-3-N/2} + hd_j^{-1-N/2} + h^\alpha d_j^{-1-\alpha-N/2} + d_j^{-2}|||F|||_{Q'_j}\right) \\ &+ \sum_{j=0}^{J_*}\left(C_\epsilon h^{1/2}d_j^{-1/2} + C_\epsilon hd_j^{-1} + \epsilon\right)d_j^{1+N/2}\left(d_j^{-1}|||F|||_{1,Q'_j} + |||\partial_t F|||_{Q'_j}\right) \\ &+ \sum_{j=0}^{J_*}\left(C_\epsilon h^{1/2}d_j^{-1/2} + C_\epsilon hd_j^{-1} + \epsilon\right)d_j^{1+N/2}\left(d_j|||\partial_t F|||_{1,Q'_j} + d_j^2|||\partial_{tt}F|||_{Q'_j}\right) \\ \le& C_\epsilon + C_\epsilon\sum_{j=0}^{J_*} d_j^{-1+N/2}|||F|||_{Q'_j}\end{aligned}$$

$$+\sum_{j=0}^{J_*}\left(C_\epsilon h^{1/2}d_j^{-1/2}+C_\epsilon h d_j^{-1}+\epsilon\right)d_j^{1+N/2}\left(d_j^{-1}|||F|||_{1,Q_j'}+|||\partial_t F|||_{Q_j'}\right)$$

$$+\sum_{j=0}^{J_*}\left(C_\epsilon h^{1/2}d_j^{-1/2}+C_\epsilon h d_j^{-1}+\epsilon\right)d_j^{1+N/2}\left(d_j|||\partial_t F|||_{1,Q_j'}+d_j^2|||\partial_{tt}F|||_{Q_j'}\right). \quad (5.5.108)$$

Since $|||F|||_{Q_j'}\le C\left(|||F|||_{Q_{j-1}}+|||F|||_{Q_j}+|||F|||_{Q_{j+1}}\right)$, we can convert the Q_j'-norm in the inequality (5.5.108) to the Q_j-norm:

$$\begin{aligned}\mathcal{K}\le & C_\epsilon+C_\epsilon\sum_{j=0}^{J_*}d_j^{-1+N/2}|||F|||_{Q_j}+C_\epsilon d_*^{-1+N/2}|||F|||_{Q_*}\\ &+\sum_{j=0}^{J_*}\left(C_\epsilon h^{1/2}d_j^{-1/2}+C_\epsilon h d_j^{-1}+\epsilon\right)d_j^{1+N/2}\left(d_j^{-1}|||F|||_{1,Q_j}+|||\partial_t F|||_{Q_j}\right)\\ &+\sum_{j=0}^{J_*}\left(C_\epsilon h^{1/2}d_j^{-1/2}+C_\epsilon h d_j^{-1}+\epsilon\right)d_j^{1+N/2}\left(d_j|||\partial_t F|||_{1,Q_j}+d_j^2|||\partial_{tt}F|||_{Q_j}\right)\\ &+\left(C_\epsilon h^{1/2}d_*^{-1/2}+C_\epsilon h d_*^{-1}+\epsilon\right)d_*^{1+N/2}\left(d_*^{-1}|||F|||_{1,Q_*}+|||\partial_t F|||_{Q_*}\right)\\ &+\left(C_\epsilon h^{1/2}d_*^{-1/2}+C_\epsilon h d_*^{-1}+\epsilon\right)d_*^{1+N/2}\left(d_*|||\partial_t F|||_{1,Q_*}+d_*^2|||\partial_{tt}F|||_{Q_{J_*+1}}\right).\end{aligned}$$

We can use global estimates (5.5.92) to bound

$$|||F|||_{Q_*},|||F|||_{1,Q_*},|||\partial_t F|||_{Q_*},|||\partial_t F|||_{1,Q_*}.$$

For the term $d_*^{N/2+3}|||\partial_{tt}F|||_{Q_{J_*+1}}$, note that by (5.4.65c), $\|\partial_{tt}\Gamma\|_{L^2(Q_{J_*+1})}\le Cd_*^{-3-N/2}$ and by (5.5.92d), $\|\partial_{tt}\check{\Gamma}_h\|_{L^2(Q_{J_*+1})}\le\|\partial_{tt}\check{\Gamma}_h\|_{L^2(Q_*)}\le Ch^{-3-N/2}$. Therefore, we have

$$\begin{aligned}\mathcal{K}\le & C_\epsilon+C_\epsilon C_*^{3+N/2}+C_\epsilon\sum_{j=0}^{J_*}d_j^{-1+N/2}|||F|||_{Q_j} \quad (5.5.109)\\ &+\sum_{j=0}^{J_*}\left(C_\epsilon h^{1/2}d_j^{-1/2}+C_\epsilon h d_j^{-1}+\epsilon\right)d_j^{1+N/2}\left(d_j^{-1}|||F|||_{1,Q_j}+|||\partial_t F|||_{Q_j}\right)\\ &+\sum_{j=0}^{J_*}\left(C_\epsilon h^{1/2}d_j^{-1/2}+C_\epsilon h d_j^{-1}+\epsilon\right)d_j^{1+N/2}\left(d_j|||\partial_t F|||_{1,Q_j}+d_j^2|||\partial_{tt}F|||_{Q_j}\right)\\ \le & C_\epsilon+C_\epsilon C_*^{3+N/2}+C_\epsilon\sum_{j=0}^{J_*}d_j^{-1+N/2}|||F|||_{Q_j}+C(C_\epsilon C_*^{-1/2}+C_\epsilon C_*^{-1}+\epsilon)\mathcal{K}.\end{aligned}$$

We have used $d_j\ge C_*h$ and the expression of $\mathcal{K}$ in (5.5.99) to obtain the last inequality in (5.5.109). By choosing ϵ small enough and then choosing C_* large enough (C_* is still to be determined later), the term $C(C_\epsilon C_*^{-1/2}+C_\epsilon C_*^{-1}+\epsilon)\mathcal{K}$ in (5.5.109) will be absorbed by the left hand side term $\mathcal{K}$. Hence, we obtain

$$\mathcal{K}\le CC_*^{3+N/2}+C\sum_{j=0}^{J_*}d_j^{-1+N/2}|||F|||_{Q_j}. \quad (5.5.110)$$

It remains to estimate $|||F|||_{Q_j}$. To this end, we apply a duality argument below. Let w be the solution of the backward parabolic equation

$$-\partial_t w - \Delta w = v \text{ with } w(1) = 0,\ \omega|_{\partial\Omega} = 0,$$

where v is a function supported on Q_j with $|||v|||_{Q_j} = 1$. The auxiliary backward parabolic equation above has been introduced in [101, Section 5]; for brevity we will directly use the estimates on ω (cf. [101, (5.24), (5.31)]) proved there.

Multiplying the above equation by F yields (notice that $\mathcal{Q} = (0,1) \times \Omega$)

$$[F, v] = (F(0), w(0)) + [F_t, w] + [\nabla F, \nabla w]. \tag{5.5.111}$$

Here

$$\begin{aligned}
&(F(0), w(0)) \\
=&(\check{P}_h\check{\delta}_{x_0} - \check{\delta}_{x_0}, w(0))_\Omega = (\check{P}_h\check{\delta}_{x_0} - \check{\delta}_{x_0}, w(0) - a_h(x)\check{I}_h w(0))_\Omega \quad \text{(by (5.3.35))} \\
=&(\check{P}_h\check{\delta}_{x_0} - \check{\delta}_{x_0}, a_h(x)(w(0) - \check{I}_h w(0)))_\Omega + (\check{P}_h\check{\delta}_{x_0} - \check{\delta}_{x_0}, (1 - a_h(x))w(0))_\Omega \\
=&(a_h(x)\check{P}_h\check{\delta}_{x_0}, w(0) - \check{I}_h w(0))_{\Omega_j''} + (a_h(x)(\check{P}_h\check{\delta}_{x_0} - \check{\delta}_{x_0}), w(0) - \check{I}_h w(0))_{(\Omega_j'')^c} \\
&\quad + (\check{P}_h\check{\delta}_{x_0} - \check{\delta}_{x_0}, (1 - a_h(x))w(0))_\Omega \\
=:&\mathcal{I}_1 + \mathcal{I}_2 + \mathcal{I}_3.
\end{aligned}$$

By the same arguments in [101, (5.21, 5.22, 5.24)], we have

$$\mathcal{I}_1 + \mathcal{I}_2 \le Ch^{1+\alpha-N/2} d_j^{-\alpha}.$$

By (5.3.32), (5.3.47a), L^q stability (5.3.37) of $\check{P}_h$ and global energy estimate for w, we have

$$\begin{aligned}
\mathcal{I}_3 =&(\check{P}_h\check{\delta}_{x_0} - \check{\delta}_{x_0}, (1 - a_h(x))w(0))_\Omega \\
\le&Ch^r \left(\|\check{P}_h\check{\delta}_{x_0}\|_{L^{\frac{6}{5}}(\Omega)} + \|\check{\delta}_{x_0}\|_{L^{\frac{6}{5}}(\Omega)} \right) \|w(0)\|_{H^1(\Omega)} \\
\le&Ch^{r-N/6} |||v|||_{Q_j} \le Ch^{1/2}.
\end{aligned}$$

Therefore we have

$$|(F(0), w(0))| \le Ch^{1+\alpha-N/2} d_j^{-\alpha} + Ch^{1/2}. \tag{5.5.112}$$

Since $F = \check{\Gamma}_h - \Gamma$, from (5.3.54a) and (5.3.53a) we have

$$\begin{aligned}
&[F_t, w] + [\nabla F, \nabla w] \\
=&[a_h(x)F_t, w - \check{I}_h w] + [A_h(x)\nabla F, \nabla(w - \check{I}_h w)] \\
&+ [(1 - a_h(x))\partial_t \Gamma, w] + [(I_N - A_h(x))\nabla\Gamma, \nabla w] \\
&+ [(1 - a_h(x))\partial_t \Gamma, \check{I}_h w - w] + [(I_N - A_h(x))\nabla\Gamma, \nabla(\check{I}_h w - w)] \\
&+ [(1 - a_h(x))\partial_t F, w] + [(I_N - A_h(x))\nabla F, \nabla w] \\
=:&\mathcal{J}_1 + \mathcal{J}_2 + \mathcal{J}_3 + \mathcal{J}_4.
\end{aligned} \tag{5.5.113}$$

By local interpolation error estimate (5.3.45), we have

$$\mathcal{J}_1 = [a_h(x)F_t, w - \check{I}_h w] + [A_h(x)\nabla F, \nabla(w - \check{I}_h w)]$$

$$\leq C \sum_{*,i} \left(|||F_t|||_{Q_i} \left|\left|\left| w - \check{I}_h w \right|\right|\right|_{Q_i} + |||F|||_{1,Q_i} \left|\left|\left| w - \check{I}_h w \right|\right|\right|_{1,Q_i} \right)$$

$$\leq C \sum_{*,i} \left(h^{1+\alpha} |||F_t|||_{Q_i} + h^{\alpha} |||F|||_{1,Q_i} \right) \|w\|_{L^2 H^{1+\alpha}(Q_i')}. \tag{5.5.114}$$

By Hölder's inequality and Sobolev embedding $H^{1+\alpha}(\Omega) \hookrightarrow L^\infty(\Omega)$, $H^{1+\alpha}(\Omega) \hookrightarrow W^{1,3}(\Omega)$ for $\alpha > 1/2$, we have

$$\begin{aligned}
\mathcal{J}_2 =& [(1 - a_h(x))\partial_t \Gamma, w] + [(I_N - A_h(x))\nabla \Gamma, \nabla w] \\
\leq& C h^r \sum_{*,i} \left(\|\partial_t \Gamma\|_{L^2 L^1(Q_i)} \|w\|_{L^2 L^\infty(Q_i)} + \|\nabla \Gamma\|_{L^2 L^{3/2}(Q_i)} \|\nabla w\|_{L^2 L^3(Q_i)} \right) \\
\leq& C h^r \sum_{*,i} \left(\|\partial_t \Gamma\|_{L^2 L^1(Q_i)} + \|\nabla \Gamma\|_{L^2 L^{3/2}(Q_i)} \right) \|w\|_{L^2 H^{1+\alpha}(Q_i)} \\
\leq& C h^r \sum_{*,i} \left(d_i^{N/2} \|\partial_t \Gamma\|_{L^2(Q_i)} + d_i^{N/6} \|\nabla \Gamma\|_{L^2(Q_i)} \right) \|w\|_{L^2 H^{1+\alpha}(Q_i)}.
\end{aligned}$$

Then by (5.4.65c) and (5.5.92a)-(5.5.92c), we have

$$\mathcal{J}_2 \leq C h^r \sum_{*,i} \left(C_*^{N/2+1} d_i^{-1} + C_*^{N/2} d_i^{-N/3} \right) \|w\|_{L^2 H^{1+\alpha}(Q_i)}, \tag{5.5.115}$$

where we used estimates

$\|\partial_t \Gamma\|_{L^2(Q_*)} \leq \|\partial_t \Gamma\|_{L^2(\mathcal{Q})} \leq C C_*^{1+N/2} d_*^{-1-N/2}$ and $\|\nabla \Gamma\|_{L^2(Q_*)} \leq \|\nabla \Gamma\|_{L^2(\mathcal{Q})} \leq C C_*^{N/2} d_*^{-N/2}$

to deal with the innermost term. Similarly, by Hölder's inequality and Sobolev embedding $H^{1+\alpha}(\Omega) \hookrightarrow L^\infty(\Omega)$, $H^{1+\alpha}(\Omega) \hookrightarrow W^{1,3}(\Omega)$ for $\alpha > 1/2$, we have

$$\begin{aligned}
\mathcal{J}_4 =& [(1 - a_h(x))\partial_t F, w] + [(I_N - A_h(x))\nabla F, \nabla w] \\
\leq& C h^r \sum_{*,i} \left(d_i^{N/2} \|\partial_t F\|_{L^2(Q_i)} + d_i^{N/6} \|\nabla F\|_{L^2(Q_i)} \right) \|w\|_{L^2 H^{1+\alpha}(Q_i)}.
\end{aligned} \tag{5.5.116}$$

By (5.3.45, 5.4.65c, 5.5.92a, 5.5.92c), we have

$$\begin{aligned}
\mathcal{J}_3 =& [(1 - a_h(x))\partial_t \Gamma, \check{I}_h w - w]_{\mathcal{Q}} + [(I_N - A_h(x))\nabla \Gamma, \nabla(\check{I}_h w - w)]_{\mathcal{Q}} \\
\leq& C h^r \sum_{*,i} \left(|||\partial_t \Gamma|||_{Q_i} \left|\left|\left| \check{I}_h w - w \right|\right|\right|_{Q_i} + |||\nabla \Gamma|||_{Q_i} \left|\left|\left| \check{I}_h w - w \right|\right|\right|_{1,Q_i} \right) \\
\leq& C h^r \sum_{*,i} \left(h^{1+\alpha} \|\partial_t \Gamma\|_{L^2(Q_i)} + h^{\alpha} \|\nabla \Gamma\|_{L^2(Q_i)} \right) \|w\|_{L^2 H^{1+\alpha}(Q_i')} \\
\leq& C h^r \sum_{*,i} \left(h^{1+\alpha} C_*^{1+N/2} d_i^{-1-N/2} + h^{\alpha} C_*^{N/2} d_i^{-N/2} \right) \|w\|_{L^2 H^{1+\alpha}(Q_i')}.
\end{aligned} \tag{5.5.117}$$

By (5.5.113) and the estimates (5.5.114, 5.5.115, 5.5.116, 5.5.117)of $\mathcal{J}_1, \cdots, \mathcal{J}_4$, we have

$$\begin{aligned}
& [F_t, w]_{\mathcal{Q}} + [\nabla F, \nabla w]_{\mathcal{Q}} \\
\leq& C \sum_{*,i} \left((h^{1+\alpha} + h^r d_i^{N/2}) |||F_t|||_{Q_i} + (h^{\alpha} + h^r d_i^{N/6}) |||F|||_{1,Q_i} \right) \|w\|_{L^2 H^{1+\alpha}(Q_i')}
\end{aligned} \tag{5.5.118}$$

$$+ CC_*^{N/2+1} \sum_{*,i} h^r \left(d_i^{-1} + h^\alpha d_i^{-N/2} \right) \|w\|_{L^2 H^{1+\alpha}(Q_i')},$$

where we have used $d_i \le 1$ and $d_i \ge h$ to simplify the result obtained. We note that the following local $H^{1+\alpha}$-estimate of w has been proved by [101, (5.31)]

$$\|w\|_{L^2 H^{1+\alpha}(Q_i')} \le C d_i^{1-\alpha} \left(\frac{\min(d_i, d_j)}{\max(d_i, d_j)} \right)^\alpha. \tag{5.5.119}$$

Hence, substituting (5.5.112, 5.5.118) and (5.5.119) into (5.5.111) yields

$$\begin{aligned} &|||F|||_{Q_j} \qquad (5.5.120) \\ \le& C(h^{1+\alpha-N/2} d_j^{-\alpha} + h^{1/2}) + CC_*^{1+N/2} \sum_{*,i} h^r \left(d_i^{-1} + h^\alpha d_i^{-N/2} \right) d_i^{1-\alpha} \left(\frac{\min(d_i, d_j)}{\max(d_i, d_j)} \right)^\alpha \\ &+ C \sum_{*,i} \left((h^{1+\alpha} + h^r d_i^{N/2}) |||F_t|||_{Q_i} + (h^\alpha + h^r d_i^{N/6}) |||F|||_{1,Q_i} \right) d_i^{1-\alpha} \left(\frac{\min(d_i, d_j)}{\max(d_i, d_j)} \right)^\alpha \end{aligned}$$

Since $\alpha > 1/2$, it follows that for any $i \in *, 0, 1, \cdots, J_*$, we have:

$$\sum_{j=0}^{J_*} d_j^{N/2-1} \left(\frac{\min(d_i, d_j)}{\max(d_i, d_j)} \right)^\alpha \le C d_i^{N/2-1}. \tag{5.5.121}$$

By (5.5.110) and (5.5.120), we have:

$$\begin{aligned} &\mathcal{K} \\ \le& CC_*^{3+N/2} + C \sum_{j=0}^{J_*} d_j^{-1+N/2} |||F|||_{Q_j} \\ \le& CC_*^{3+N/2} + \sum_{j=0}^{j_*} h^{1/2} d_j^{-1+N/2} + \sum_{j=0}^{J_*} \left(\frac{h}{d_j} \right)^{1+\alpha-N/2} \\ &+ CC_*^{N/2+1} \sum_{j=0}^{J_*} d_j^{-1+N/2} \sum_{*,i} h^r \left(d_i^{-1} + h^\alpha d_i^{-N/2} \right) d_i^{1-\alpha} \left(\frac{\min(d_i, d_j)}{\max(d_i, d_j)} \right)^\alpha \\ &+ C \sum_{j=0}^{J_*} d_j^{-1+N/2} \sum_{*,i} \left((h^{1+\alpha} + h^r d_i^{N/2}) |||F_t|||_{Q_i} + (h^\alpha + h^r d_i^{N/6}) |||F|||_{1,Q_i} \right) d_i^{1-\alpha} \left(\frac{\min(d_i, d_j)}{\max(d_i, d_j)} \right)^\alpha \end{aligned}$$

It is easy to check that:

$$\sum_{j=0}^{J_*} h^{1/2} d_j^{-1+N/2} + \sum_{j=0}^{J_*} \left(\frac{h}{d_j} \right)^{1+\alpha-N/2} \le C. \tag{5.5.122}$$

By (5.5.121), we have:

$$\begin{aligned} &\sum_{j=0}^{J_*} d_j^{-1+N/2} \sum_{*,i} h^r \left(d_i^{-1} + h^\alpha d_i^{-N/2} \right) d_i^{1-\alpha} \left(\frac{\min(d_i, d_j)}{\max(d_i, d_j)} \right)^\alpha \\ =& \sum_{*,i} \left(h^r d_i^{-\alpha} + h^{r+\alpha} d_i^{1-\alpha-N/2} \right) \sum_{j=0}^{J_*} d_j^{-1+N/2} \left(\frac{\min(d_i, d_j)}{\max(d_i, d_j)} \right)^\alpha \end{aligned}$$

$$\leq C \sum_{*,i} \left(h^r d_i^{N/2-1-\alpha} + h^{r+\alpha} d_i^{-\alpha} \right)$$
$$\leq C \sum_{*,i} \left(h^{1-\alpha} d_i^{N/2-1} (\frac{h}{d_i})^\alpha + h(\frac{h}{d_i})^\alpha \right) \leq C \quad (\text{used } N = 2, 3,\ r \geq 1,\ d_i \leq 1).$$

By (5.5.121) again, using the facts $r \geq 1$ and $\alpha > 1/2$, and (5.5.92a, 5.5.92c), we have:

$$\sum_{j=0}^{J_*} d_j^{-1+N/2} \sum_{*,i} \left((h^{1+\alpha} + h^r d_i^{N/2}) |||F_t|||_{Q_i} + (h^\alpha + h^r d_i^{N/6}) |||F|||_{1,Q_i} \right) d_i^{1-\alpha} \left(\frac{\min(d_i, d_j)}{\max(d_i, d_j)} \right)^\alpha$$
$$= \sum_{*,i} \left((h^{1+\alpha} + h^r d_i^{N/2}) |||F_t|||_{Q_i} + (h^\alpha + h^r d_i^{N/6}) |||F|||_{1,Q_i} \right) d_i^{1-\alpha} \sum_{j=0}^{J_*} d_j^{-1+N/2} \left(\frac{\min(d_i, d_j)}{\max(d_i, d_j)} \right)^\alpha$$
$$\leq C \sum_{*,i} \left((h^{1+\alpha} + h^r d_i^{N/2}) |||F_t|||_{Q_i} + (h^\alpha + h^r d_i^{N/6}) |||F|||_{1,Q_i} \right) d_i^{N/2-\alpha}$$
$$\leq C \sum_{*,i} d_i^{1+N/2} \left(|||F_t|||_{Q_i} \left(\frac{h}{d_i} \right)^{1+\alpha} + d_i^{-1} |||F|||_{1,Q_i} \left(\frac{h}{d_j} \right)^\alpha \right)$$
$$+ C \sum_{*,i} d_i^{1+N/2} \left(|||F_t|||_{Q_i} d_i^{N/2-\alpha} \left(\frac{h}{d_i} \right) + d_i^{-1} |||F|||_{1,Q_i} h^{1-\alpha} d_i^{N/6} \left(\frac{h}{d_i} \right)^\alpha \right)$$
$$\leq C \sum_{*,i} d_i^{1+N/2} \left(|||F_t|||_{Q_i} + d_i^{-1} |||F|||_{1,Q_i} \right) \left(\frac{h}{d_i} \right)^\alpha \quad (\text{used } h \leq d_i \leq 1)$$
$$\leq C d_*^{1+N/2} \left(|||F_t|||_{Q_*} + d_i^{-1} |||F|||_{1,Q_*} \right) + C \sum_{i=0}^{J_*} d_i^{1+N/2} \left(|||F_t|||_{Q_i} + d_i^{-1} |||F|||_{1,Q_i} \right) \left(\frac{h}{d_i} \right)^\alpha$$
$$\leq C C_*^{1+N/2} + \frac{C\mathcal{K}}{C_*^\alpha}.$$

So we obtain

$$\mathcal{K} \leq C C_*^{3+N/2} + \frac{C\mathcal{K}}{C_*^\alpha}.$$

By choosing C_* to be large enough (C_* determined now), the term $\frac{C\mathcal{K}}{C_*^\alpha}$ will be absorbed by the left-hand side of the inequality above. In the case, the inequality above implies

$$\mathcal{K} \leq C. \tag{5.5.123}$$

Substituing (5.5.123) into (5.5.93) yields

$$\|\partial_t F\|_{L^1(\mathcal{Q})} + \|t \partial_{tt} F\|_{L^1(\mathcal{Q})} \leq C. \tag{5.5.124}$$

Part II. Next, we present estimates for $(t, x) \in (1, \infty) \times \Omega$. For $t > 1$, we differentiate (5.3.54) with respect to t and integrate the resulting equation against $\partial_t \check{\Gamma}_h$. Then we get:

$$\frac{d}{dt} \| \sqrt{a_h(x)} \partial_t \check{\Gamma}_h(t, \cdot, x_0) \|^2_{L^2(\Omega)} + (A_h(x) \nabla \partial_t \check{\Gamma}_h(t, \cdot, x_0), \nabla \partial_t \check{\Gamma}_h(t, \cdot, x_0))_\Omega = 0$$

for $t \geq 1$. Owing to (5.3.32), when h is sufficiently small, there exists $\lambda_0 > 0$ only dependent on Ω such that:

$$\lambda_0 \|\partial_t \check{\Gamma}_h(t, \cdot, x_0)\|^2_{L^2(\Omega)} \leq (A_h(x)\nabla\partial_t\check{\Gamma}_h(t, \cdot, x_0), \nabla\partial_t\check{\Gamma}_h(t, \cdot, x_0))_\Omega.$$

So we have:

$$\frac{d}{dt}\|\sqrt{a_h(x)}\partial_t\check{\Gamma}_h(t, \cdot, x_0)\|^2_{L^2(\Omega)} + \lambda_0\|\partial_t\check{\Gamma}_h(t, \cdot, x_0)\|^2_{L^2(\Omega)} \leq 0,$$

which implies:

$$\|\partial_t\check{\Gamma}_h(t, \cdot, x_0)\|^2_{L^2(\Omega)} \leq e^{-\lambda_0(t-1)}\|\partial_t\check{\Gamma}_h(1, \cdot, x_0)\|^2_{L^2(\Omega)}.$$

By a standard energy estimate, we have $\|\partial_t\check{\Gamma}_h(1, \cdot, x_0)\|^2_{L^2(\Omega)} \leq C$. So we have:

$$\|\partial_t\check{\Gamma}_h(t, \cdot, x_0)\|^2_{L^2(\Omega)} \leq Ce^{-\lambda_0(t-1)} \quad \forall t \geq 1. \tag{5.5.125}$$

Similarly, we also have:

$$\|\partial_{tt}\check{\Gamma}_h(t, \cdot, x_0)\|^2_{L^2(\Omega)} + \|\partial_t\Gamma(t, \cdot, x_0)\|^2_{L^2(\Omega)} + \|\partial_{tt}\Gamma(t, \cdot, x_0)\|^2_{L^2(\Omega)} \leq Ce^{-\lambda_0(t-1)} \quad \forall t \geq 1. \tag{5.5.126}$$

The estimate (5.5.124) and the last two inequalities imply (5.4.66c, 5.4.66d) in the case h is small enough.

Part III. Finally, we notice that (5.4.66b) is a simple consequence of (5.3.47b) and analyticity estimate (5.1.4) of semigroup $e^{t\Delta}$. While (5.4.66a) is a consequence of (5.4.66b), (5.4.66c) and the following inequalities:

$$\begin{aligned}\|F(t, \cdot, x_0)\|_{L^1(\Omega)} &\leq \|F(0, \cdot, x_0)\|_{L^1(\Omega)} + \int_0^t \|\partial_t F(s, \cdot, x_0)\|_{L^1(\Omega)}ds \\ &\leq \|\check{\delta}_{x_0} - \check{P}_h\check{\delta}_{x_0}\|_{L^1(\Omega)} + \|\partial_t F\|_{L^1((0,\infty)\times\Omega)} \leq C, \quad \forall t \in (0, 1)\end{aligned} \tag{5.5.127}$$

$$\begin{aligned}\|t\partial_t F(t, \cdot, x_0)\|_{L^1(\Omega)} &\leq \int_0^t \|\partial_s F(s, \cdot, x_0) + s\partial_{ss}F(s, \cdot, x_0)\|_{L^1(\Omega)}ds \\ &\leq \|\partial_t F\|_{L^1((0,\infty)\times\Omega)} + \|t\partial_{tt}F\|_{L^1((0,\infty)\times\Omega)} \leq C, \quad \forall t \in (0, 1)\end{aligned} \tag{5.5.128}$$

The proof of Lemma 5.4.3 is complete. ∎

5.6 Proof of Lemma 5.5.1

In this section, we prove Lemma 5.5.1. First, we observe that the same proof as in [101, Lemma 5.1] leads to the following lemma, which provides the local energy error estimate under the assumption of local Galerkin orthogonality.

Lemma 5.6.1. *Suppose that $\phi \in L^2(0, 1; H^1_0(\Omega)) \cap H^1(0, 1; L^2(\Omega))$ and $\check{\phi}_h \in H^1(0, 1; \check{S}^\circ_h(\Omega))$ satisfy the equation*

$$(a_h(x)\partial_t(\phi - \check{\phi}_h), \check{\chi}_h) + (A_h(x)\nabla(\phi - \check{\phi}_h), \nabla\check{\chi}_h) = 0 \quad \forall\check{\chi}_h \in \check{S}^\circ_h(\Omega'_j), t \in (0, d^2_j),$$

$$(a_h(x)\partial_t(\phi-\check{\phi}_h),\check{\chi}_h)+(A_h(x)\nabla(\phi-\check{\phi}_h),\nabla\check{\chi}_h)=0\quad \forall\check{\chi}_h\in\check{S}_h^\circ(D_j'),t\in(\frac{1}{4}d_j^2,4d_j^2)$$

with $\phi(0)=0$ *in* Ω_j'. *Then the following holds:*

$$\begin{aligned}&\left|\!\left|\!\left|\partial_t(\phi-\check{\phi}_h)\right|\!\right|\!\right|_{Q_j}+d_j^{-1}\left|\!\left|\!\left|\phi-\check{\phi}_h\right|\!\right|\!\right|_{1,Q_j} \\ \leq & C\epsilon^{-3}\left(I_j(\check{\phi}_h(0))+X_j(\check{I}_h\phi-\phi)+d_j^{-2}\left|\!\left|\!\left|\phi-\check{\phi}_h\right|\!\right|\!\right|_{Q_j'}\right) \\ &+\left(Ch^{1/2}d_j^{-1/2}+C\epsilon^{-1}hd_j^{-1}+\epsilon\right)\left(\left|\!\left|\!\left|\partial_t(\phi-\check{\phi}_h)\right|\!\right|\!\right|_{Q_j'}+d_j^{-1}\left|\!\left|\!\left|\phi-\check{\phi}_h\right|\!\right|\!\right|_{1,Q_j'}\right),\end{aligned} \tag{5.6.129}$$

where

$$\begin{aligned}I_j(\check{\phi}_h(0))=&\|\check{\phi}_h(0)\|_{1,\Omega_j'}+d_j^{-1}\|\check{\phi}_h(0)\|_{\Omega_j'},\\ X_j(\check{I}_h\phi-\phi)=&d_j\|\nabla\partial_t(\check{I}_h\phi-\phi)\|_{L^2(Q_j')}+\|\partial_t(\check{I}_h\phi-\phi)\|_{L^2(Q_j')}\\ &+d_j^{-1}\|\nabla(\check{I}_h\phi-\phi)\|_{L^2(Q_j')}+d_j^{-2}\|\check{I}_h\phi-\phi\|_{L^2(Q_j')}.\end{aligned}$$

Here $\epsilon\in(0,1)$ *is an arbitrary parameter, and the positive constant* C *is independent of* h,j *and* C_*.

The distinction between Lemma 5.5.1 and Lemma 5.6.1 is that in the condition of Lemma 5.5.1, there only holds perturbed Galerkin orthogonality (5.5.90). In fact, Lemma 5.5.1 is derived from Lemma 5.6.1 by additionally accounting for the domain perturbation term, which introduces the extra term $Y_j(\phi)$ in the error estimate of Lemma 5.5.1.

Proof of Lemma 5.5.1. Let $0\leq\widetilde{\omega}\leq 1$ be a smooth cut-off function which equals to 1 in Q_j' and $\text{supp}(\widetilde{\omega})\subset Q_j''$, with an estimate of derivatives $|\partial_t^k\nabla^l\widetilde{\omega}|\leq Cd_j^{-2k-l}$ for any non-negative integers k,l. Let $\widetilde{\phi}=\widetilde{\omega}\phi$, and let $\check{\eta}_h(t)\in\check{S}_h^\circ(\Omega)$ be the solution of

$$(a_h(x)\partial_t\check{\eta}_h,\check{\chi}_h)+(A_h(x)\nabla\check{\eta}_h,\nabla\check{\chi}_h)=((1-a_h)\partial_t\widetilde{\phi},\check{\chi}_h)+((I-A_h)\nabla\widetilde{\phi},\nabla\check{\chi}_h) \tag{5.6.130}$$
$$\forall\check{\chi}_h\in\check{S}_h^\circ(\Omega),t\in(0,1)\text{ and }\check{\eta}_h(0)=0.$$

Since $\widetilde{\phi}=\phi$ in Q_j', it follows that

$$(a_h(x)\partial_t\check{\eta}_h,\check{\chi}_h)+(A_h(x)\nabla\check{\eta}_h,\nabla\check{\chi}_h)=((1-a_h)\partial_t\phi,\check{\chi}_h)+((I-A_h)\nabla\phi,\nabla\check{\chi}_h) \tag{5.6.131}$$
$$\forall\check{\chi}_h\in\check{S}_h^\circ(\Omega_j'),t\in(0,d_j^2)$$

and

$$(a_h(x)\partial_t\check{\eta}_h,\check{\chi}_h)+(A_h(x)\nabla\check{\eta}_h,\nabla\check{\chi}_h)=((1-a_h)\partial_t\phi,\check{\chi}_h)+((I-A_h)\nabla\phi,\nabla\check{\chi}_h) \tag{5.6.132}$$
$$\forall\check{\chi}_h\in\check{S}_h^\circ(D_j'),t\in(\frac{1}{4}d_j^2,4d_j^2).$$

Now we split $\check{\phi}_h$ into $\check{\phi}_h=\check{\theta}_h+\check{\eta}_h$. Then $\check{\theta}_h$ satisfies the condition of Lemma 5.6.1, i.e.,

$$\begin{aligned}&(a_h(x)\partial_t(\phi-\check{\theta}_h),\check{\chi}_h)+(A_h(x)\nabla(\phi-\check{\theta}_h),\nabla\check{\chi}_h)=0\quad\forall\check{\chi}_h\in\check{S}_h^\circ(\Omega_j'),t\in(0,d_j^2),\\ &(a_h(x)\partial_t(\phi-\check{\theta}_h),\check{\chi}_h)+(A_h(x)\nabla(\phi-\check{\theta}_h),\nabla\check{\chi}_h)=0\quad\forall\check{\chi}_h\in\check{S}_h^\circ(D_j'),t\in(\frac{1}{4}d_j^2,4d_j^2).\end{aligned}$$

Therefore, we can apply Lemma 5.6.1 to obtain

$$\left|\!\left|\!\left|\partial_t(\phi-\check{\theta}_h)\right|\!\right|\!\right|_{Q_j}+d_j^{-1}\left|\!\left|\!\left|\phi-\check{\theta}_h\right|\!\right|\!\right|_{1,Q_j}$$

$$\leq C\epsilon^{-3}\left(I_j(\check{\theta}_h(0)) + X_j(I_h\phi - \phi) + d_j^{-2}|||\phi - \check{\theta}_h|||_{Q_j'}\right) + \left(Ch^{1/2}d_j^{-1/2} + C\epsilon^{-1}hd_j^{-1} + \epsilon\right)\left(|||\partial_t(\phi - \check{\theta}_h)|||_{Q_j'} + d_j^{-1}|||\phi - \check{\theta}_h|||_{1,Q_j'}\right).$$

Again we split $\phi - \check{\theta}_h$ into $\phi - \check{\theta}_h = (\phi - \check{\phi}_h) + \check{\eta}_h$ and note that $\check{\theta}_h(0) = \check{\phi}_h(0)$. It follows that

$$\begin{aligned}
&|||\partial_t(\phi - \check{\phi}_h)|||_{Q_j} + d_j^{-1}|||\phi - \check{\phi}_h|||_{1,Q_j} \qquad (5.6.133)\\
\leq &C\epsilon^{-3}\left(I_j(\check{\phi}_h(0)) + X_j(I_h\phi - \phi) + d_j^{-2}|||\phi - \check{\phi}_h|||_{Q_j'}\right)\\
&+ \left(Ch^{1/2}d_j^{-1/2} + C\epsilon^{-1}hd_j^{-1} + \epsilon\right)\left(|||\partial_t(\phi - \check{\phi}_h)|||_{Q_j'} + d_j^{-1}|||\phi - \check{\phi}_h|||_{1,Q_j'}\right)\\
&+ C\epsilon^{-3}\left(d_j^{-2}\|\check{\eta}_h\|_{L^2(Q_j')} + d_j^{-1}\|\nabla\check{\eta}_h\|_{L^2(Q_j')} + \|\partial_t\check{\eta}_h\|_{L^2(Q_j')}\right).
\end{aligned}$$

By Hölder's inequality, we have

$$\begin{aligned}
&d_j^{-2}\|\check{\eta}_h\|_{L^2(Q_j')} + d_j^{-1}\|\nabla\check{\eta}_h\|_{L^2(Q_j')} + \|\partial_t\check{\eta}_h\|_{L^2(Q_j')}\\
\leq &C\left(d_j^{-1}\|\check{\eta}_h\|_{L^\infty L^2(\mathcal{Q})} + d_j^{-1}\|\nabla\check{\eta}_h\|_{L^2(\mathcal{Q})} + \|\partial_t\check{\eta}_h\|_{L^2(\mathcal{Q})}\right) \qquad (5.6.134)
\end{aligned}$$

It remains to establish a global energy estimate for $\check{\eta}_h$ using the following lemma:

Lemma 5.6.2. *Assume $\check{\phi}_h \in \check{S}_h^\circ(\Omega)$ satisfies*

$$(a_h(x)\partial_t\check{\phi}_h, \check{\chi}_h) + (A_h(x)\nabla\check{\phi}_h, \nabla\check{\chi}_h) = (f, \check{\chi}_h) + (\boldsymbol{g}, \nabla\check{\chi}_h)\forall\check{\chi}_h \in \check{S}_h^\circ(\Omega), t \in (0,1),$$

and $\check{\phi}_h(0) = 0$. Then we have

$$\|\check{\phi}_h\|_{L^\infty L^2(\mathcal{Q})} + \|\nabla\check{\phi}_h\|_{L^2(\mathcal{Q})} \leq C(\|f\|_{L^1L^2(\mathcal{Q})} + \|\boldsymbol{g}\|_{L^2(\mathcal{Q})}) \qquad (5.6.135)$$

$$\|\partial_t\check{\phi}_h\|_{L^2(\mathcal{Q})} \leq C\left(\|f\|_{L^2(\mathcal{Q})} + \|\boldsymbol{g}(0)\|_{L^2} + \|\boldsymbol{g}\|^{\frac{1}{2}}_{L^2(\mathcal{Q})}\|\partial_t\boldsymbol{g}\|^{\frac{1}{2}}_{L^2(\mathcal{Q})} + \|\partial_t\boldsymbol{g}\|^{\frac{1}{2}}_{L^2(\mathcal{Q})}\|f\|^{\frac{1}{2}}_{L^1L^2(\mathcal{Q})}\right) \qquad (5.6.136)$$

Proof. Let $\check{\chi}_h = \check{\phi}_h$ and integrate in time, we have

$$\|\check{\phi}_h\|^2_{L^\infty(0,1;L^2)} + \|\nabla\check{\phi}_h\|^2_{L^2(\mathcal{Q})} \leq C\left(\int_0^1 \|f(t)\|_{L^2}\|\check{\phi}_h(t)\|_{L^2}dt + \int_0^1 \|\boldsymbol{g}(t)\|_{L^2}\|\nabla\check{\phi}_h(t)\|_{L^2}dt\right),$$

and (5.6.135) is obtained from the estimate above and Hölder's inequality. Let $\check{\chi}_h = \partial_t\check{\phi}_h$, we have

$$\begin{aligned}
\|\sqrt{a_h}\partial_t\check{\phi}_h\|^2_{L^2} + \frac{1}{2}\frac{d}{dt}(A_h\nabla\check{\phi}_h, \nabla\check{\phi}_h) &= (f, \partial_t\check{\phi}_h) + (\boldsymbol{g}, \nabla\partial_t\check{\phi}_h)\\
&= (f, \partial_t\check{\phi}_h) + \frac{d}{dt}(\boldsymbol{g}, \nabla\check{\phi}_h) - (\partial_t\boldsymbol{g}, \nabla\check{\phi}_h).
\end{aligned}$$

Integrate in time from $s = 0$ to $s = t$, we obtain

$$\begin{aligned}
&\|\partial_t\check{\phi}_h\|^2_{L^2(0,t;L^2)} + \|\nabla\check{\phi}_h(t)\|^2_{L^2}\\
\leq &C\left(\int_0^t (\|f\|_{L^2}\|\partial_t\check{\phi}_h\|_{L^2} + \|\partial_t\boldsymbol{g}\|_{L^2}\|\nabla\check{\phi}_h\|_{L^2})ds + \|\boldsymbol{g}(t)\|_{L^2}\|\|\nabla\check{\phi}_h(t)\|_{L^2}\right)\\
\leq &C\left(\|f\|_{L^2(0,t;L^2)}\|\partial_t\check{\phi}_h\|_{L^2(0,t;L^2)} + \|\partial_t\boldsymbol{g}\|_{L^2(0,t;L^2)}\|\nabla\check{\phi}_h\|_{L^2(0,t;L^2)} + \|\boldsymbol{g}(t)\|_{L^2}\|\|\nabla\check{\phi}_h(t)\|_{L^2}\right)
\end{aligned}$$

Using Young's inequality to absorb $\|\nabla\check{\phi}_h(t)\|_{L^2}$ and $\|\partial_t\check{\phi}_h\|_{L^2(0,t;L^2)}$ on the right side, we have

$$\|\partial_t\check{\phi}_h\|_{L^2(\mathcal{Q})} + \|\nabla\check{\phi}_h\|_{L^\infty L^2(\mathcal{Q})} \le C\left(\|f\|_{L^2(\mathcal{Q})} + \|\boldsymbol{g}\|_{L^\infty L^2(\mathcal{Q})} + \|\partial_t\boldsymbol{g}\|^{\frac{1}{2}}_{L^2(\mathcal{Q})}\|\nabla\check{\phi}_h\|^{\frac{1}{2}}_{L^2(\mathcal{Q})}\right). \tag{5.6.137}$$

Substitute the interpolation inequality

$$\|\boldsymbol{g}\|_{L^\infty(0,1;L^2)} \le C\|\boldsymbol{g}\|^{\frac{1}{2}}_{L^2(0,1;L^2)}\|\partial_t\boldsymbol{g}\|_{L^2(0,1;L^2)} + \|\boldsymbol{g}(0)\|_{L^2},$$

and the estimate (5.6.135) of $\|\nabla\check{\phi}_h\|_{L^2(\mathcal{Q})}$ into (5.6.137), we obtain (5.6.136). ∎

Take $\check{\phi}_h = \check{\eta}_h$, $f = (1-a_h)\partial_t\widetilde{\phi}$ and $\boldsymbol{g} = (I - A_h)\nabla\widetilde{\phi}$ in the last lemma and note (5.3.32), we have

$$\begin{aligned}\|\check{\eta}_h\|_{L^\infty L^2(\mathcal{Q})} + \|\nabla\check{\eta}_h\|_{L^2(\mathcal{Q})} &\le Ch^r(\|\partial_t\widetilde{\phi}\|_{L^1L^2(\mathcal{Q})} + \|\nabla\widetilde{\phi}\|_{L^2(\mathcal{Q})})\\ &\le Ch^r(d_j\|\partial_t\widetilde{\phi}\|_{L^2(\mathcal{Q})} + \|\nabla\widetilde{\phi}\|_{L^2(\mathcal{Q})}), \text{(used Hölder's inequality)}\end{aligned}$$

and

$$\begin{aligned}\|\partial_t\check{\eta}_h\|_{L^2(\mathcal{Q})} \le& Ch^r\left(\|\partial_t\widetilde{\phi}\|_{L^2(\mathcal{Q})} + \|\nabla\widetilde{\phi}\|^{\frac{1}{2}}_{L^2(\mathcal{Q})}\|\nabla\partial_t\widetilde{\phi}\|^{\frac{1}{2}}_{L^2(\mathcal{Q})} + \|\nabla\partial_t\widetilde{\phi}\|^{\frac{1}{2}}_{L^2(\mathcal{Q})}\|\partial_t\widetilde{\phi}\|_{L^1L^2(\mathcal{Q})}\right)\\ \le& Ch^r\left(\|\partial_t\widetilde{\phi}\|_{L^2(\mathcal{Q})} + d_j^{-1}\|\nabla\widetilde{\phi}\|_{L^2(\mathcal{Q})} + d_j\|\nabla\partial_t\widetilde{\phi}\|_{L^2(\mathcal{Q})} + d_j^{-1}\|\partial_t\widetilde{\phi}\|_{L^1L^2(\mathcal{Q})}\right)\\ \le& Ch^r\left(\|\partial_t\widetilde{\phi}\|_{L^2(\mathcal{Q})} + d_j^{-1}\|\nabla\widetilde{\phi}\|_{L^2(\mathcal{Q})} + d_j\|\nabla\partial_t\widetilde{\phi}\|_{L^2(\mathcal{Q})}\right),\end{aligned}$$

where we have used the fact that $\mathrm{supp}\widetilde{\phi} \subseteq Q_j''$ and $\widetilde{\phi}(0) = 0$, which follows from $\mathrm{supp}\widetilde{\omega} \subseteq Q_j''$ and the assumption $\phi(0) = 0$ in Ω_j''. As a result,

$$\begin{aligned}& d_j^{-1}\|\check{\eta}_h\|_{L^\infty L^2(\mathcal{Q})} + d_j^{-1}\|\nabla\check{\eta}_h\|_{L^2(\mathcal{Q})} + \|\partial_t\check{\eta}_h\|_{L^2(\mathcal{Q})} \qquad (5.6.138)\\ \le& Ch^r\left(\|\partial_t\widetilde{\phi}\|_{L^2(\mathcal{Q})} + d_j^{-1}\|\nabla\widetilde{\phi}\|_{L^2(\mathcal{Q})} + d_j\|\nabla\partial_t\widetilde{\phi}\|_{L^2(\mathcal{Q})}\right)\\ \le& Ch^r\left(\|\partial_t\phi\|_{L^2(Q_j'')} + d_j^{-1}\|\nabla\phi\|_{L^2(Q_j'')} + d_j\|\nabla\partial_t\phi\|_{L^2(Q_j'')} + d_j^{-2}\|\phi\|_{L^2(Q_j'')}\right).\end{aligned}$$

In conclusion, combining (5.6.133), (5.6.134) and (5.6.138) we have shown

$$\begin{aligned}& \left|\!\left|\!\left|\partial_t(\phi - \check{\phi}_h)\right|\!\right|\!\right|_{Q_j} + d_j^{-1}\left|\!\left|\!\left|\phi - \check{\phi}_h\right|\!\right|\!\right|_{1,Q_j} \qquad (5.6.139)\\ \le& C\epsilon^{-3}\left(I_j(\check{\phi}_h(0)) + X_j(I_h\phi - \phi) + \widetilde{Y}_j(\phi) + d_j^{-2}\left|\!\left|\!\left|\phi - \check{\phi}_h\right|\!\right|\!\right|_{Q_j'}\right)\\ &+ \left(Ch^{1/2}d_j^{-1/2} + C\epsilon^{-1}hd_j^{-1} + \epsilon\right)\left(\left|\!\left|\!\left|\partial_t(\phi - \check{\phi}_h)\right|\!\right|\!\right|_{Q_j'} + d_j^{-1}\left|\!\left|\!\left|\phi - \check{\phi}_h\right|\!\right|\!\right|_{1,Q_j'}\right),\end{aligned}$$

where in the last inequality (5.6.139), the $\widetilde{Y}_j(\phi)$ denotes

$$\widetilde{Y}_j(\phi) = h^r\left(d_j\|\nabla(\partial_t\phi)\|_{L^2(Q_j'')} + \|\partial_t\phi\|_{L^2(Q_j'')} + d_j^{-1}\|\nabla\phi\|_{L^2(Q_j'')} + d_j^{-2}\|\phi\|_{L^2(Q_j'')}\right).$$

Finally, in order to replace Q_j'' with Q_j' in the expression of $\widetilde{Y}_j(\phi)$ above, we observe that Lemma 5.6.1 remains valid if in the statement Q_j' is substituted with $Q_j^{',\frac{1}{2}}$, where $Q_j^{',\frac{1}{2}}$ denotes the following intermediate set between Q_j and Q_j':

$$Q_j^{',\frac{1}{2}} := \{(t,x) \in \mathcal{Q} : \frac{3}{4}d_j \le \max(|x - x_0|, t^{1/2}) \le 3d_j\}.$$

Therefore, we can substitute Q_j' with $Q_j^{',\frac{1}{2}}$ and Q_j'' with Q_j' throughout the proof above, leading to the statement of Lemma (5.5.1). ∎

5.7 Appendix: Property (P3) and operator $\check{I}_h$

We take the same approach as in [101, Appendix B], constructing a modified Clément's interpolation operator (which is similar to the Scott-Zhang interpolation operator cf. [21, Section 4.8]) $\widetilde{I}_h : H^1(\Omega) \to \check{S}_h(\Omega)$ which preserves homogeneous Dirichlet boundary condition i.e. if $u \in H_0^1(\Omega)$ then $\widetilde{I}_h u \in \check{S}_h^\circ(\Omega)$. We denote by $x_i \in \Omega, i = 1, \ldots, M$ the interior finite element nodes and denote by $x'_j \in \partial\Omega, j = 1, \ldots, m$ the boundary finite element nodes of the finite element space $\check{S}_h(\Omega)$. We denote by $\check{\Phi}_i \in \check{S}_h^\circ(\Omega)$ the basis function corresponding to $x_i \in \Omega$ and denote by $\check{\Phi}'_j \in \check{S}_h(\Omega)$ the basis function corresponding to $x'_j \in \partial\Omega$. In other words, we have relation

$$\check{\Phi}_i(x_j) = \delta_{ij}, \check{\Phi}'_i(x'_j) = \delta_{ij}, \check{\Phi}_i(x'_j) = 0 \text{ and } \check{\Phi}'_j(x_i) = 0.$$

Let $\tau_i = \bigcup\{\check{K} \in \check{\mathcal{T}}_h : x_i \in \check{K}\}$ and $\tau'_j = \bigcup\{\check{K} \in \check{\mathcal{T}}_h : x'_j \in \check{K}\}$. For each interior node $x_i \in \Omega$, we define $P_h^{(i)} : L^2(\Omega) \to \check{S}_h(\tau_i)$ as the local L^2-projection onto $\check{S}_h(\tau_i)$, i.e.,

$$(P_h^{(i)} v, \check{\chi}_h)_{\tau_i} = (v, \check{\chi}_h)_{\tau_i} \quad \forall \check{\chi}_h \in \check{S}_h(\tau_i).$$

For each boundary node x'_j, we define $\overline{P}_h^{(j)} : H^1(\Omega) \to \check{S}_h(\partial\Omega \cap \tau'_j)$ where

$$\check{S}_h(\partial\Omega \cap \tau'_j) = \{\chi \in C^0(\partial\Omega \cap \tau'_j) : \exists \check{\chi}_h \in \check{S}_h(\Omega) \text{ s.t. } \check{\chi}_h|_{\partial\Omega\cap\tau'_j} = \chi\},$$

and $\overline{P}_h^{(j)} v$ is the local L^2 projection of $v|_{\partial\Omega}$ (trace of v on the boundary):

$$(\overline{P}_h^{(j)} v, \check{\chi}_h)_{\partial\Omega\cap\tau'_j} = (v|_{\partial\Omega}, \check{\chi}_h)_{\partial\Omega\cap\tau'_j} \quad \forall \check{\chi}_h \in \check{S}_h(\partial\Omega \cap \tau'_j).$$

We define operator $\widetilde{I}_h : H^1(\Omega) \to \check{S}_h(\Omega)$ by setting

$$\widetilde{I}_h v = \sum_{i=1}^{M} (P_h^{(i)} v)(x_i)\check{\Phi}_i + \sum_{j=1}^{m} (\overline{P}_h^{(j)} v)(x'_j)\check{\Phi}'_j. \tag{5.7.140}$$

It follows from the definition of $\widetilde{I}_h$ that $\widetilde{I}_h \check{\chi}_h = \check{\chi}_h$ for $\check{\chi}_h \in \check{S}_h(\Omega)$ and moreover $\widetilde{I}_h v \in \check{S}_h^\circ(\Omega)$ when $v \in H_0^1(\Omega)$. Therefore, the restriction of $\widetilde{I}_h$ to $H_0^1(\Omega)$ gives a projection operator $\check{I}_h : H_0^1(\Omega) \to \check{S}_h^\circ(\Omega)$ onto the finite element space $\check{S}_h^\circ(\Omega)$. To verify (P3) for $\check{I}_h$, it suffices to prove the same statements for $\widetilde{I}_h$.

Using the mesh regularity condition (5.3.29), we can establish the following inverse estimates by pulling back to reference element:

$$|\check{\chi}_h|(x'_j) \le Ch^{-N/2+1/2} \|\check{\chi}_h\|_{L^2(\partial\Omega\cap\tau'_j)}; \quad |\check{\chi}_h|(x_i) \le Ch^{-N/2} \|\check{\chi}_h\|_{L^2(\tau_i)} \quad \forall \check{\chi}_h \in \check{S}_h(\Omega) \tag{5.7.141}$$

where x'_j can be any boundary node and x_i any interior node. Thus, from the definition of $\widetilde{I}_h$, it is straightforward to verify the following local stability:

$$\|\widetilde{I}_h v\|_{L^2(\tau_i)} + h\|\nabla \widetilde{I}_h v\|_{L^2(\tau_i)} \le C\left(\|v\|_{L^2(\widetilde{\tau}_i)} + h^{1/2}\|v\|_{L^2(\partial\Omega\cap\widetilde{\tau}_i)}\right)$$
$$\|\widetilde{I}_h v\|_{L^2(\tau'_j)} + h\|\nabla \widetilde{I}_h v\|_{L^2(\tau'_j)} \le C\left(\|v\|_{L^2(\widetilde{\tau}'_j)} + h^{1/2}\|v\|_{L^2(\partial\Omega\cap\widetilde{\tau}'_j)}\right),$$

where $\widetilde{\tau}_i = \bigcup\{\check{K} \in \check{\mathcal{T}}_h : \check{K} \cap \tau_i \neq \emptyset\}$ and $\widetilde{\tau}_j' = \bigcup\{\check{K} \in \check{\mathcal{T}}_h : \check{K} \cap \tau_j' \neq \emptyset\}$. Summing up the two inequalities above for $i = 1, \dots, M$ and $j = 1, \dots, m$ and taking into account of the quasi-uniformity (P1), we obtain the global stability:

$$\|\widetilde{I}_h v\|_{L^2(\Omega)} + h\|\nabla \widetilde{I}_h v\|_{L^2(\Omega)} \le C(\|v\|_{L^2(\Omega)} + h^{1/2}\|v\|_{L^2(\partial\Omega)}).$$

By trace inequality $\|v\|_{L^2(\partial\Omega)} \le C\|v\|_{L^2(\Omega)}^{1/2}\|v\|_{H^1(\Omega)}^{1/2}$ and Young's inequality, we have:

$$\|\widetilde{I}_h v\|_{L^2(\Omega)} + h\|\nabla \widetilde{I}_h v\|_{L^2(\Omega)} \le C(\|v\|_{L^2(\Omega)} + h\|\nabla v\|_{L^2(\Omega)}),$$

which implies the following quasi-optimal approximation property since $\widetilde{I}_h$ is a projection onto $\check{S}_h(\Omega)$:

$$\|v - \widetilde{I}_h v\|_{L^2(\Omega)} + h\|\nabla(v - \widetilde{I}_h v)\|_{L^2(\Omega)} \le C(\|v - \check{\chi}_h\|_{L^2(\Omega)} + h\|\nabla(v - \check{\chi}_h)\|_{L^2(\Omega)}) \quad \forall \check{\chi}_h \in \check{S}_h(\Omega). \tag{5.7.142}$$

When $v \in H^2(\Omega)$, by simply taking $\check{\chi}_h = \check{I}_h^L v$, i.e., the Lagrangian interpolation of v (cf. [94, Lemma 7] for error estimate of Lagrangian interpolation operator of $\check{S}_h(\Omega)$), we have

$$\|v - \widetilde{I}_h v\|_{L^2(\Omega)} + h\|\nabla(v - \widetilde{I}_h v)\|_{L^2(\Omega)} \le Ch^2\|v\|_{H^2(\Omega)} \quad \forall v \in H^2(\Omega). \tag{5.7.143}$$

To consider the case where $v \in H^1(\Omega)$, we can construct another Lipschitz homeomorphism $\widetilde{\boldsymbol{\Psi}}_h : \widetilde{\Omega}_h \to \Omega_h$ by

$$\widetilde{\boldsymbol{\Psi}}_h|_{\widetilde{K}} = \boldsymbol{F}_K \circ \boldsymbol{F}_{\widetilde{K}}^{-1},$$

where $\widetilde{\Omega}_h$ = interior of $\bigcup_{\widetilde{K} \in \widetilde{\mathcal{T}}_h} \widetilde{K}$ is the domain consisting of the triangles/tetrahedrons in the initial flat triangulation of Ω described in Section 5.3.2. The mesh regularity condition (5.3.27) guarantees that

$$\|\nabla \widetilde{\boldsymbol{\Psi}}_h\|_{L^\infty(\widetilde{\Omega}_h)} + \|\nabla \widetilde{\boldsymbol{\Psi}}_h^{-1}\|_{L^\infty(\Omega_h)} \le C. \tag{5.7.144}$$

Let $S_h(\widetilde{\Omega}_h)$ be the finite element space based on $\widetilde{\mathcal{T}}_h$ and let $\widetilde{P}_h' : L^2(\widetilde{\Omega}_h) \to S_h(\widetilde{\Omega}_h)$ be the L^2 projection onto $S_h(\widetilde{\Omega}_h)$. Since $S_h(\widetilde{\Omega}_h)$ is the usual Lagrangian finite element space based on a quasi-uniform triangle/tetrahedron mesh, there holds the following basic estimate of the L^2 projection $\widetilde{P}_h'$:

$$\|(1 - \widetilde{P}_h')f\|_{L^2(\widetilde{\Omega}_h)} + h\|\nabla(1 - \widetilde{P}_h')f\|_{L^2(\widetilde{\Omega}_h)} \le Ch\|f\|_{H^1(\widetilde{\Omega}_h)}.$$

From the definition of finite element space $\check{S}_h(\Omega)$, $\check{\chi}_h = \widetilde{P}_h'(v \circ \boldsymbol{\Phi}_h \circ \widetilde{\boldsymbol{\Psi}}_h) \circ \widetilde{\boldsymbol{\Psi}}_h^{-1} \circ \boldsymbol{\Phi}_h^{-1}$ belongs to $\check{S}_h(\Omega_h)$ and satisfies

$$\begin{aligned}
&\|v - \check{\chi}_h\|_{L^2(\Omega)} + h\|\nabla(v - \check{\chi}_h)\|_{L^2(\Omega)} \\
\le& C\|(1 - \widetilde{P}_h')(v \circ \boldsymbol{\Phi}_h \circ \widetilde{\boldsymbol{\Psi}}_h)\|_{L^2(\widetilde{\Omega}_h)} + Ch\|\nabla(1 - \widetilde{P}_h')(v \circ \boldsymbol{\Phi}_h \circ \widetilde{\boldsymbol{\Psi}}_h)\|_{L^2(\widetilde{\Omega}_h)} \\
\le& Ch\|v \circ \boldsymbol{\Phi}_h \circ \widetilde{\boldsymbol{\Psi}}_h\|_{L^2(\widetilde{\Omega}_h)} \le Ch\|v\|_{H^1(\Omega)},
\end{aligned} \tag{5.7.145}$$

where we have used the fact that $\boldsymbol{\Phi}_h$ and $\widetilde{\boldsymbol{\Psi}}_h$ induce norm equivalence in view of (5.7.144) and (5.3.30)-(5.3.31). Therefore, by (5.7.142) we also established:

$$\|v - \widetilde{I}_h v\|_{L^2(\Omega)} + h\|\nabla(v - \widetilde{I}_h v)\|_{L^2(\Omega)} \le Ch\|v\|_{H^1(\Omega)} \quad \forall v \in H^1(\Omega). \tag{5.7.146}$$

By complex interpolation method, we can deduce from (5.7.146) and (5.7.143) that

$$\|v - \widetilde{I}_h v\|_{L^2(\Omega)} + \|\nabla(v - \widetilde{I}_h v)\|_{L^2(\Omega)} \leq Ch^{1+\alpha}\|v\|_{H^{1+\alpha}(\Omega)} \quad \forall v \in H^{1+\alpha}(\Omega). \tag{5.7.147}$$

This proves (P3)-(1). From the construction of $\widetilde{I}_h v$, it is direct to verify that if $\mathrm{diam}(\check{K}) \leq h$ for all $\check{K} \in \check{\mathcal{T}}_h$, then the value of $\widetilde{I}_h v$ in $D \subseteq \Omega$ only depends on the value of v in $B_d(D)$ for $d \geq 2h$. This proves (P3)-(2). Finally, it remains to prove the super-approximation estimate

$$\|\omega\check{\psi}_h - \widetilde{I}_h(\omega\check{\psi}_h)\|_{L^2(\Omega)} + h\|\omega\check{\psi}_h - \widetilde{I}_h(\omega\check{\psi}_h)\|_{H^1(\Omega)} \leq Chd^{-1}\|\check{\psi}_h\|_{L^2(B_d(D))}.$$

From the quasi-optimal approximation estimate (5.7.142), there holds

$$\begin{aligned} &\|\omega\check{\psi}_h - \widetilde{I}_h(\omega\check{\psi}_h)\|_{L^2(\Omega)} + h\|\omega\check{\psi}_h - \widetilde{I}_h(\omega\check{\psi}_h)\|_{H^1(\Omega)} \\ \leq &\|\omega\check{\psi}_h - \check{I}_h^L(\omega\check{\psi}_h)\|_{L^2(\Omega)} + h\|\omega\check{\psi}_h - \check{I}_h^L(\omega\check{\psi}_h)\|_{H^1(\Omega)}, \end{aligned}$$

while for the Lagrangian interpolation operator $\check{I}_h^L$ we refer to [103, Lemma 2.6] for a proof of its super-approximation property.

Chapter 6

Conclusion

This thesis investigates finite element methods (FEM) for solving complex problems in fluid–structure interaction (FSI) and partial differential equations (PDEs) posed on curvilinear domains. The work is divided into three main parts, each addressing fundamental challenges in computational mathematics and numerical analysis.

The first part (Chapter 2) focuses on the numerical solution of the Stokes equations in evolving domains with moving boundaries, using the arbitrary Lagrangian–Eulerian (ALE) finite element method. For Taylor–Hood elements of degree $r \geq 2$, we establish optimal L^2 error estimates of order $O(h^{r+1})$ for the velocity and $O(h^r)$ for the pressure. The analysis employs Nitsche's duality argument, adapted to an evolving mesh, and hinges on proving that the material derivative and the Stokes–Ritz projection commute up to terms of optimal L^2 convergence order. A novel duality argument is developed to derive an optimal H^{-1} error estimate for the pressure component of the Stokes–Ritz projection, which is essential in the L^2-error analysis of the commutator. Numerical experiments support the theoretical results and demonstrate the method's effectiveness in simulating Navier–Stokes flow in domains with rotating boundaries, such as a propeller.

The second part (Chapter 3) addresses FSI problems involving the coupling of the Stokes equations in a fluid domain with an elastic wave equation on the boundary through kinematic and dynamic interface conditions. Such coupled systems are analytically and numerically challenging. On one hand, optimal L^2 error estimates for FEM in FSI problems had not been previously established, primarily due to the lack of an appropriate Ritz projection to control the consistency error. To overcome this, we introduce a coupled non-stationary Ritz projection and analyze its approximation properties, thereby providing the first optimal L^2 error estimate for the semi-discrete FEM approximation of FSI problems. On the other hand, the presence of both parabolic and hyperbolic dynamics complicates the construction of stable, decoupled numerical schemes. To address this, we propose a fully discrete, kinematically coupled scheme that decouples the fluid and structure subproblems while ensuring unconditional energy stability. Using the newly constructed non-stationary Ritz projection, we further prove optimal L^2-norm convergence of the fully discrete scheme.

The third part (Chapters 4–5) investigates the maximum-norm stability and error estimates of isoparametric FEM for elliptic and parabolic equations posed on curvilinear polyhedral domains. Since such domains cannot, in general, be exactly triangulated using linear simplices, isoparametric finite element methods are employed to better approximate curved boundaries. This, however, introduces domain perturbation errors, as the computational domain Ω_h no longer coincides with the exact domain Ω, necessitating a careful analysis. We address this by adopting a transformation technique that maps the problem from Ω_h to Ω via a bi-Lipschitz transformation $\Phi_h : \Omega_h \to \Omega$. The resulting formulation leads to a finite element discretization on Ω for an equivalent problem with perturbed coefficients that encode the effects of domain perturbation.

In Chapter 4, we establish the weak discrete maximum principle for isoparametric FEM applied to the Poisson equation with Dirichlet boundary conditions on (possibly nonconvex) curvilinear polyhedral domains with edge openings smaller than π. These domains include smooth regions and smooth deformations of convex polyhedra. The proof relies on the analysis of a dual elliptic problem with a discontinuous coefficient matrix arising from the isoparametric mapping. As standard H^2-regularity does not hold in this context, we decompose the dual solution into a smooth component and a singular component, estimating them via H^2 and $W^{1,p}$ regularity, respectively. As an application, we establish a maximum-norm best approximation result for the isoparametric FEM. To address the domain perturbation $\Omega \neq \Omega_h$, we construct a globally smooth flow map from Ω to a larger perturbed domain containing Ω_h, enabling a uniform $W^{1,\infty}$ estimate of the continuous solution with respect to the perturbation.

In Chapter 5, we investigate the isoparametric finite element semi-discretization for a parabolic problem on a curvilinear polyhedral domain $\Omega \subseteq \mathbb{R}^N$ with homogeneous Dirichlet boundary condition. The domain Ω may include non-convex corners, i.e., with edge openings possibly greater than π. We establish the analyticity and maximal regularity of the discrete semigroup by employing a transformation method to address the domain perturbation effect $\Omega \neq \Omega_h$. The analysis hinges on estimating the error $\Gamma - \check{\Gamma}_h$ between the regularized Green function of the Laplace equation in Ω and the discrete Green function $\check{\Gamma}_h$ of the transformed finite element problem in Ω. Since only a perturbed Galerkin orthogonality holds for $\Gamma - \check{\Gamma}_h$ due to the domain perturbation effect, the primary difficulty in the proof lies in handling the local energy error estimates and the local duality arguments in the context of perturbed Galerkin orthogonality. As an application of the logarithmically quasi-maximal L^∞-regularity of the discrete semigroup, we reduce the maximum norm error estimate of the parabolic problem to that of the elliptic Ritz projection, incorporating an additional optimal order term resulting from the domain perturbation effect.

Several promising directions for future research emerge from this thesis, particularly in advancing the numerical analysis of fluid–structure interaction (FSI) problems and finite element methods under domain perturbation. For FSI, a deeper investigation of the coupled non-stationary Ritz projection is warranted, particularly to extend the current framework to more complex configurations involving thick-structure interactions, interface deformation, and general inflow/outflow boundary conditions. Understanding the stability properties of the Ritz projection in these extended settings remains a key open question. In addition, the analyticity and maximal regularity of the evolution equation associated with the coupled Ritz projection deserve further study. It is also of interest to explore the removal of the current theoretical restrictions on initial data by examining more closely the stability properties of the coupled parabolic–hyperbolic system.

In the context of domain perturbation, a systematic comparison between the extension method and the transformation method raises rich and subtle theoretical challenges. While both approaches have demonstrated effectiveness in certain settings, there exist intricate scenarios in which one method proves successful while the other does not. This highlights the need for a unified analytical framework capable of capturing the effects of domain perturbation in FEM analysis across a broader class of geometries and PDE models.

Bibliography

[1] R. A. Adams and J. J. Fournier. *Sobolev spaces.* Elsevier, 2003.

[2] G. Akrivis, B. Li, and C. Lubich. Combining maximal regularity and energy estimates for time discretizations of quasilinear parabolic equations. *Mathematics of Computation*, 86(306):1527–1552, 2017.

[3] H. Amann. *Linear and Quasilinear Parabolic Problems: Volume I: Abstract Linear Theory*, volume 1. Springer Science & Business Media, 1995.

[4] C. Amrouche and V. Girault. On the existence and regularity of the solution of Stokes problem in arbitrary dimension. *Proceedings of the Japan Academy, Series A, Mathematical Sciences*, 67(5):171–175, 1991.

[5] M. Annese, M. A. Fernández, and L. Gastaldi. Splitting schemes for a Lagrange multiplier formulation of FSI with immersed thin-walled structure: stability and convergence analysis. *IMA Journal of Numerical Analysis*, 43(2):881–919, 2023.

[6] T. Apel, A. Rösch, and D. Sirch. L^∞-error estimates on graded meshes with application to optimal control. *SIAM journal on control and optimization*, 48(3):1771–1796, 2009.

[7] T. Apel, M. Winkler, and J. Pfefferer. Error estimates for the postprocessing approach applied to Neumann boundary control problems in polyhedral domains. *IMA Journal of Numerical Analysis*, 38(4):1984–2025, 2018.

[8] A. Ashyralyev, S. Piskarev, and L. Weis. On well-posedness of difference schemes for abstract parabolic equations in $L^p([0,T];E)$ spaces. *Numerical Functional Analysis and Optimization*, 23(7-8):669–693, 2002.

[9] C. Bacuta, A. L. Mazzucato, V. Nistor, and L. Zikatanov. Interface and mixed boundary value problems on n-dimensional polyhedral domains. *Documenta Mathematica*, 15:687–745, 2010.

[10] S. Badia and R. Codina. Analysis of a stabilized finite element approximation of the transient convection-diffusion equation using an ALE framework. *SIAM Journal on Numerical Analysis*, 44(5):2159–2197, 2006.

[11] G. Bai, B. Kovács, and B. Li. Maximal regularity of evolving FEMs for parabolic equations on an evolving surface. *arXiv preprint arXiv:2408.14096*, 2024.

[12] N. Bakaev, V. Thomée, and L. Wahlbin. Maximum-norm estimates for resolvents of elliptic finite element operators. *Mathematics of computation*, 72(244):1597–1610, 2003.

[13] J. W. Banks, W. D. Henshaw, and D. W. Schwendeman. An analysis of a new stable partitioned algorithm for FSI problems. Part I: Incompressible flow and elastic solids. *Journal of Computational Physics*, 269:108–137, 2014.

[14] N. Behringer, D. Leykekhman, and B. Vexler. Global and local pointwise error estimates for finite element approximations to the Stokes problem on convex polyhedra. *SIAM Journal on Numerical Analysis*, 58(3):1531–1555, 2020.

[15] J. Bergh and J. Löfström. *Interpolation spaces: an introduction*, volume 223. Springer Science & Business Media, 2012.

[16] D. Boffi, F. Brezzi, M. Fortin, et al. *Mixed finite element methods and applications*, volume 44. Springer, 2013.

[17] D. Boffi and L. Gastaldi. Stability and geometric conservation laws for ALE formulations. *Computer methods in applied mechanics and engineering*, 193(42-44):4717–4739, 2004.

[18] A. Bonito, I. Kyza, and R. H. Nochetto. Time-discrete higher order ALE formulations: A priori error analysis. *Numerische Mathematik*, 125(2):225–257, 2013.

[19] A. Bonito, I. Kyza, and R. H. Nochetto. Time-discrete higher-order ALE formulations: stability. *SIAM Journal on Numerical Analysis*, 51(1):577–604, 2013.

[20] J. Brandts, S. Korotov, M. Křížek, and J. Šolc. On nonobtuse simplicial partitions. *SIAM review*, 51(2):317–335, 2009.

[21] S. C. Brenner. *The mathematical theory of finite element methods*. Springer, 2008.

[22] T. Bühler and D. A. Salamon. *Functional analysis*, volume 191. American Mathematical Soc., 2018.

[23] M. Bukač, S. Čanić, R. Glowinski, J. Tambača, and A. Quaini. Fluid–structure interaction in blood flow capturing non-zero longitudinal structure displacement. *Journal of Computational Physics*, 235:515–541, 2013.

[24] M. Bukač, S. Čanić, and B. Muha. A partitioned scheme for fluid–composite structure interaction problems. *Journal of Computational Physics*, 281:493–517, 2015.

[25] M. Bukac and B. Muha. Stability and convergence analysis of the extensions of the kinematically coupled scheme for the fluid-structure interaction. *SIAM Journal on Numerical Analysis*, 54(5):3032–3061, 2016.

[26] M. Bukač and C. Trenchea. Adaptive, second-order, unconditionally stable partitioned method for fluid–structure interaction. *Computer Methods in Applied Mechanics and Engineering*, 393:114847, 2022.

[27] N. Bullerjahn and B. Kovács. Error estimates for full discretization of Cahn–Hilliard equation with dynamic boundary conditions. *arXiv preprint arXiv:2407.20698*, 2024.

[28] E. Burman, R. Durst, M. A. Fernández, and J. Guzmán. Fully discrete loosely coupled robin-robin scheme for incompressible fluid–structure interaction: stability and error analysis. *Numerische Mathematik*, 151(4):807–840, 2022.

[29] E. Burman and M. A. Fernández. Stabilization of explicit coupling in fluid–structure interaction involving fluid incompressibility. *Computer Methods in Applied Mechanics and Engineering*, 198(5-8):766–784, 2009.

[30] E. Burman and M. A. Fernández. Explicit strategies for incompressible fluid-structure interaction problems: Nitsche type mortaring versus robin–robin coupling. *International Journal for Numerical Methods in Engineering*, 97(10):739–758, 2014.

[31] E. Burman, S. Frei, and A. Massing. Eulerian time-stepping schemes for the non-stationary Stokes equations on time-dependent domains. *Numerische Mathematik*, pages 1–56, 2022.

[32] C. P. Calderón and M. Milman. Interpolation of sobolev spaces. the real method. *Indiana University Mathematics Journal*, 32(6):801–808, 1983.

[33] P. Causin, J.-F. Gerbeau, and F. Nobile. Added-mass effect in the design of partitioned algorithms for fluid–structure problems. *Computer methods in applied mechanics and engineering*, 194(42-44):4506–4527, 2005.

[34] P. G. Ciarlet. Discrete maximum principle for finite-difference operators. *Aequationes mathematicae*, 4(3):338–352, 1970.

[35] P. G. Ciarlet. *The finite element method for elliptic problems.* SIAM, 2002.

[36] P. G. Ciarlet and P.-A. Raviart. Maximum principle and uniform convergence for the finite element method. *Computer methods in applied mechanics and engineering*, 2(1):17–31, 1973.

[37] P. Clément. *Semigroup Theory and Evolution Equations: The Second International Conference.* CRC Press, 2023.

[38] P. Clément and J. Prüss. Global existence for a semilinear parabolic volterra equation. *Mathematische Zeitschrift*, 209:17–26, 1992.

[39] M. Dauge. Neumann and mixed problems on curvilinear polyhedra. *Integral equations and operator theory*, 15:227–261, 1992.

[40] A. Demlow. Localized pointwise a posteriori error estimates for gradients of piecewise linear finite element approximations to second-order quasilinear elliptic problems. *SIAM Journal on Numerical Analysis*, 44(2):494–514, 2006.

[41] A. Djurdjevac, C. Gräser, and P. J. Herbert. An evolving space framework for oseen equations on a moving domain. *ESAIM: Mathematical Modelling and Numerical Analysis*, 57(5):3113–3138, 2023.

[42] G. Dore. L p regularity for abstract differential equations. In *Functional Analysis and Related Topics, 1991: Proceedings of the International Conference in Memory of Professor Kôsaku Yosida held at RIMS, Kyoto University, Japan, July 29–Aug. 2, 1991*, pages 25–38. Springer, 2006.

[43] B. Dörich, J. Leibold, and B. Maier. Optimal $W^{1,\infty}$-estimates for an isoparametric finite element discretization of elliptic boundary value problems. *ETNA - Electronic Transactions on Numerical Analysis*, 2022.

[44] E. H. Dowell and K. C. Hall. Modeling of fluid-structure interaction. *Annual review of fluid mechanics*, 33(1):445–490, 2001.

[45] G. Dziuk and C. M. Elliott. Finite elements on evolving surfaces. *IMA journal of numerical analysis*, 27(2):262–292, 2007.

[46] D. Edelmann. Isoparametric finite element analysis of a generalized robin boundary value problem on curved domains. *The SMAI journal of computational mathematics*, 7:57–73, 2021.

[47] D. Edelmann. Finite element analysis for a diffusion equation on a harmonically evolving domain. *IMA Journal of Numerical Analysis*, 42(2):1866–1901, 2022.

[48] D. Edelmann, B. Kovács, and C. Lubich. Numerical analysis of an evolving bulk–surface model of tumour growth. *IMA Journal of Numerical Analysis*, page drae077, 10 2024.

[49] C. M. Elliott and T. Ranner. Finite element analysis for a coupled bulk–surface partial differential equation. *IMA Journal of Numerical Analysis*, 33(2):377–402, 2013.

[50] 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 Journal of Numerical Analysis*, 41(3):1696–1845, 2021.

[51] C. M. Elliott and C. Venkataraman. Error analysis for an ALE evolving surface finite element method. *Numerical Methods for Partial Differential Equations*, 31(2):459–499, 2015.

[52] K. Eriksson, C. Johnson, and S. Larsson. Adaptive finite element methods for parabolic problems vi: Analytic semigroups. *SIAM journal on numerical analysis*, 35(4):1315–1325, 1998.

[53] R. Farwig and H. Sohr. Generalized resolvent estimates for the Stokes system in bounded and unbounded domains. *Journal of the Mathematical Society of Japan*, 46(4):607–643, 1994.

[54] M. A. Fernández. Incremental displacement-correction schemes for incompressible fluid-structure interaction: stability and convergence analysis. *Numerische Mathematik*, 123(1):21–65, 2013.

[55] M. A. Fernández, J.-F. Gerbeau, and C. Grandmont. A projection semi-implicit scheme for the coupling of an elastic structure with an incompressible fluid. *International journal for numerical methods in engineering*, 69(4):794–821, 2007.

[56] C. A. Figueroa, I. E. Vignon-Clementel, K. E. Jansen, T. J. Hughes, and C. A. Taylor. A coupled momentum method for modeling blood flow in three-dimensional deformable arteries. *Computer methods in applied mechanics and engineering*, 195(41-43):5685–5706, 2006.

[57] L. Formaggia, J.-F. Gerbeau, F. Nobile, and A. Quarteroni. On the coupling of 3d and 1d Navier–Stokes equations for flow problems in compliant vessels. *Computer methods in applied mechanics and engineering*, 191(6-7):561–582, 2001.

[58] L. Formaggia and F. Nobile. Stability analysis of second-order time accurate schemes for ALE–FEM. *Computer methods in applied mechanics and engineering*, 193(39-41):4097–4116, 2004.

[59] L. Formaggia, A. Quarteroni, and A. Veneziani. *Cardiovascular Mathematics: Modeling and simulation of the circulatory system*, volume 1. Springer Science & Business Media, 2010.

[60] J. Frehse and R. Rannacher. Asymptotic L^∞-error estimates for linear finite element approximations of quasilinear boundary value problems. *SIAM Journal on Numerical Analysis*, 15:418–431, 1978.

[61] G. Galdi. *An introduction to the mathematical theory of the Navier-Stokes equations: Steady-state problems*. Springer Science & Business Media, 2011.

[62] H. Gao and W. Qiu. The pointwise stabilities of piecewise linear finite element method on non-obtuse tetrahedral meshes of nonconvex polyhedra. *Journal of Scientific Computing*, 87:1–5, 2021.

[63] L. Gastaldi. A priori error estimates for the Arbitrary Lagrangian Eulerian formulation with finite elements. *Journal of Numerical Mathematics*, 9(2):123–156, 2001.

[64] E. S. Gawlik and A. J. Lew. Unified analysis of finite element methods for problems with moving boundaries. *SIAM Journal on Numerical Analysis*, 53(6):2822–2846, 2015.

[65] M. W. Gee, U. Küttler, and W. Wall. Truly monolithic algebraic multigrid for fluid–structure interaction. *International Journal for Numerical Methods in Engineering*, 85(8):987–1016, 2011.

[66] M. Geissert. Discrete maximal L^p regularity for finite element operators. *SIAM journal on numerical analysis*, 44(2):677–698, 2006.

[67] M. Geissert. Applications of discrete maximal L^p regularity for finite element operators. *Numerische Mathematik*, 108:121–149, 2007.

[68] G. Gigante and C. Vergara. On the stability of a loosely-coupled scheme based on a robin interface condition for fluid-structure interaction. *Computers & Mathematics with Applications*, 96:109–119, 2021.

[69] V. Girault and P.-A. Raviart. *Finite element methods for Navier-Stokes equations: theory and algorithms*, volume 5. Springer Science & Business Media, 2012.

[70] W. Gong, B. Li, and Q. Rao. Convergent evolving finite element approximations of boundary evolution under shape gradient flow. *IMA Journal of Numerical Analysis*, 44(5):2667–2697, 2024.

[71] W. Gong, D. Liang, and X. Xie. Pointwise error estimates for linear finite element approximation to elliptic Dirichlet problems in smooth domains. *Advances in Computational Mathematics*, 49(2):15, 2023.

[72] G. Guidoboni, R. Glowinski, N. Cavallini, and S. Canic. Stable loosely-coupled-type algorithm for fluid–structure interaction in blood flow. *Journal of Computational Physics*, 228(18):6916–6937, 2009.

[73] M. D. Gunzburger and S. L. Hou. Treating inhomogeneous essential boundary conditions in finite element methods and the calculation of boundary stresses. *SIAM journal on numerical analysis*, 29(2):390–424, 1992.

[74] B. Guo and C. Schwab. Analytic regularity of Stokes flow on polygonal domains in countably weighted sobolev spaces. *Journal of Computational and Applied Mathematics*, 190(1-2):487–519, 2006.

[75] J. Guzmán, D. Leykekhman, J. Rossmann, and A. H. Schatz. Hölder estimates for Green's functions on convex polyhedral domains and their applications to finite element methods. *Numerische Mathematik*, 112(2):221–243, 2009.

[76] W. Hao, P. Sun, J. Xu, and L. Zhang. Multiscale and monolithic arbitrary Lagrangian–Eulerian finite element method for a hemodynamic fluid-structure interaction problem involving aneurysms. *Journal of Computational Physics*, 433:110181, 2021.

[77] G. Hou, J. Wang, and A. Layton. Numerical methods for fluid-structure interaction—a review. *Communications in Computational Physics*, 12(2):337–377, 2012.

[78] J. Hron and S. Turek. A monolithic FEM multigrid solver for an ALE formulation of fluid-structure interaction with applications in biomechanics. In *Fluid-Structure Interaction: Modelling, Simulation, Optimisation*, pages 146–170. Springer, 2006.

[79] T. Kashiwabara and T. Kemmochi. Pointwise error estimates of linear finite element method for Neumann boundary value problems in a smooth domain. *Numerische Mathematik*, 144(3):553–584, 2020.

[80] T. Kashiwabara and T. Kemmochi. Stability, analyticity, and maximal regularity for parabolic finite element problems on smooth domains. *Mathematics of Computation*, 89(324):1647–1679, 2020.

[81] T. Kashiwabara, I. Oikawa, and G. Zhou. Penalty method with P1/P1 finite element approximation for the Stokes equations under the slip boundary condition. *Numerische Mathematik*, 134:705–740, 2016.

[82] T. Kemmochi and N. Saito. Discrete maximal regularity and the finite element method for parabolic equations. *Numerische Mathematik*, 138:905–937, 2018.

[83] S. Korotov and M. Krízek. Acute type refinements of tetrahedral partitions of polyhedral domains. *SIAM journal on numerical analysis*, 39(2):724–733, 2001.

[84] S. Korotov, M. Křížek, and P. Neittaanmäki. Weakened acute type condition for tetrahedral triangulations and the discrete maximum principle. *Mathematics of Computation*, 70(233):107–119, 2001.

[85] B. Kovács, B. Li, and C. Lubich. A-stable time discretizations preserve maximal parabolic regularity. *SIAM Journal on Numerical Analysis*, 54(6):3600–3624, 2016.

[86] 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. *Numerische Mathematik*, 137:643–689, 2017.

[87] B. Kovács and C. Lubich. Numerical analysis of parabolic problems with dynamic boundary conditions. *IMA Journal of Numerical Analysis*, 37(1):1–39, 2017.

[88] P. C. Kunstmann, B. Li, and C. Lubich. Runge–kutta time discretization of nonlinear parabolic equations studied via discrete maximal parabolic regularity. *Foundations of Computational Mathematics*, 18:1109–1130, 2018.

[89] P. C. Kunstmann and L. Weis. Maximal l^p-regularity for parabolic equations, fourier multiplier theorems and h^∞-functional calculus. In *Functional analytic methods for evolution equations*, pages 65–311. Springer, 2004.

[90] I. S. Lan, J. Liu, W. Yang, and A. L. Marsden. A reduced unified continuum formulation for vascular fluid–structure interaction. *Computer Methods in Applied Mechanics and Engineering*, 394:114852, 2022.

[91] R. Lan and P. Sun. A monolithic arbitrary Lagrangian–Eulerian finite element analysis for a Stokes/parabolic moving interface problem. *Journal of Scientific Computing*, 82(3):59, 2020.

[92] G. Legendre and T. Takahashi. Convergence of a Lagrange-Galerkin method for a fluid-rigid body system in ALE formulation. *ESAIM: Mathematical Modelling and Numerical Analysis*, 42(4):609–644, 2008.

[93] C. Lehrenfeld and M. Olshanskii. An eulerian finite element method for pdes in time-dependent domains. *ESAIM: Mathematical Modelling and Numerical Analysis*, 53(2):585–614, 2019.

[94] M. Lenoir. Optimal isoparametric finite elements and error estimates for domains involving curved boundaries. *SIAM journal on numerical analysis*, 23(3):562–580, 1986.

[95] D. Leykekhman. Pointwise localized error estimates for parabolic finite element equations. *Numerische Mathematik*, 96:583–600, 2004.

[96] D. Leykekhman and B. Li. Weak discrete maximum principle of finite element methods in convex polyhedra. *Mathematics of Computation*, 90(327):1–18, 2021.

[97] D. Leykekhman and B. Vexler. Finite element pointwise results on convex polyhedral domains. *SIAM Journal on Numerical Analysis*, 54(2):561–587, 2016.

[98] D. Leykekhman and B. Vexler. Pointwise best approximation results for Galerkin finite element solutions of parabolic problems. *SIAM Journal on Numerical Analysis*, 54(3):1365–1384, 2016.

[99] D. Leykekhman and B. Vexler. Discrete maximal parabolic regularity for Galerkin finite element methods. *Numerische Mathematik*, 135(3):923–952, 2017.

[100] B. Li. Maximum-norm stability and maximal L^p regularity of FEMs for parabolic equations with lipschitz continuous coefficients. *Numerische Mathematik*, 131:489–516, 2015.

[101] B. Li. Analyticity, maximal regularity and maximum-norm stability of semi-discrete finite element solutions of parabolic equations in nonconvex polyhedra. *Mathematics of Computation*, 88(315):1–44, 2019.

[102] B. Li. Maximal regularity of multistep fully discrete finite element methods for parabolic equations. *IMA Journal of Numerical Analysis*, 42(2):1700–1734, 2022.

[103] B. Li, W. Qiu, Y. Xie, and W. Yu. Weak discrete maximum principle of isoparametric finite element methods in curvilinear polyhedra. *Mathematics of Computation*, 93(345):1–34, 2024.

[104] B. Li and W. Sun. Maximal L^p analysis of finite element solutions for parabolic equations with nonsmooth coefficients in convex polyhedra. *Mathematics of Computation*, 86(305):1071–1102, 2017.

[105] B. Li and W. Sun. Maximal regularity of fully discrete finite element solutions of parabolic equations. *SIAM Journal on Numerical Analysis*, 55(2):521–542, 2017.

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

[107] P.-L. Lions. *Mathematical Topics in Fluid Mechanics: Volume 1: Incompressible Models*. Mathematical Topics in Fluid Mechanics. Oxford University Press, Incorporated, 1996.

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

[109] J. Liu, R. K. Jaiman, and P. S. Gurugubelli. A stable second-order scheme for fluid–structure interaction with strong added-mass effects. *Journal of Computational Physics*, 270:687–710, 2014.

[110] J. Liu, I. S. Lan, and A. L. Marsden. Mathematical modeling of the vascular system. *Notices of the American Mathematical Society*, 2021.

[111] 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. *Computer Methods in Applied Mechanics and Engineering*, 333:55–73, 2018.

[112] M. Lukáčová-Medvid'ová, G. Rusnáková, and A. Hundertmark-Zaušková. Kinematic splitting algorithm for fluid–structure interaction in hemodynamics. *Computer Methods in Applied Mechanics and Engineering*, 265:83–106, 2013.

[113] C. Ma, T. Tian, and W. Zheng. High-order unfitted characteristic finite element methods for moving interface problem of oseen equations. *Journal of Computational and Applied Mathematics*, 425:115028, 2023.

[114] H. Matthias and P. Jan. Heat kernels and maximal $L^p - L^q$ estimates for parabolic evolution equations. *Communications in Partial Differential Equations*, 22(9-10):1647–1669, 1997.

[115] D. Meidner and B. Vexler. Optimal error estimates for fully discrete Galerkin approximations of semilinear parabolic equations. *ESAIM: Mathematical Modelling and Numerical Analysis*, 52(6):2307–2325, 2018.

[116] M. Neilan and M. Olshanskii. An eulerian finite element method for the linearized Navier–Stokes problem in an evolving domain. *IMA Journal of Numerical Analysis*, 44(6):3234–3258, 2024.

[117] F. Nobile. *Numerical Approximation of Fluid–Structure Interaction Problems with Application to Haemodynamics*. PhD thesis, EPFL, 2001. DOI: 10.5075/EPFL-THESIS-2458.

[118] F. Nobile and L. Formaggia. A stability analysis for the arbitrary Lagrangian Eulerian formulation with finite elements. *East-West Journal of Numerical Mathematics*, 7(2):105–132, 1999.

[119] E.-M. Ouhabaz. Gaussian estimates and holomorphy of semigroups. *Proceedings of the American Mathematical Society*, 123(5):1465–1474, 1995.

[120] O. Oyekole, C. Trenchea, and M. Bukac. A second-order in time approximation of fluid-structure interaction problem. *SIAM Journal on Numerical Analysis*, 56(1):590–613, 2018.

[121] V. Ruas. *On the strong maximum principle for some piecewise linear finite element approximative problems of non-positive type*. PhD thesis, INRIA, 1980.

[122] 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. *Journal of computational and applied mathematics*, 230(2):521–545, 2009.

[123] A. Schatz and L. Wahlbin. Maximum norm estimates in the finite element method on plane polygonal domains. I. *Mathematics of Computation*, 32(141):73–109, 1978.

[124] A. Schatz and L. B. Wahlbin. On the quasi-optimality in L^∞ of the H^1-projection into finite element spaces. *Mathematics of Computation*, 38(157):1–22, 1982.

[125] A. H. Schatz. A weak discrete maximum principle and stability of the finite element method in L^∞ on plane polygonal domains. I. *Mathematics of Computation*, 34(149):77–91, 1980.

[126] A. H. Schatz, V. Thomée, and L. B. Wahlbin. Stability, analyticity, and almost best approximation in maximum norm for parabolic finite element equations. *Communications on pure and applied mathematics*, 51(11-12):1349–1385, 1998.

[127] A. H. Schatz and L. B. Wahlbin. Interior maximum norm estimates for finite element methods. *Mathematics of Computation*, 31(138):414–442, 1977.

[128] J. Schöberl. C++ 11 implementation of finite elements in ngsolve. *Institute for analysis and scientific computing, Vienna University of Technology*, 30, 2014.

[129] L. R. Scott and S. Zhang. Finite element interpolation of nonsmooth functions satisfying boundary conditions. *Mathematics of computation*, 54(190):483–493, 1990.

[130] E. M. Stein. *Singular integrals and differentiability properties of functions*. Princeton university press, 1970.

[131] M. E. Taylor. *Partial Differential Equations I, Basic Theory*. Springer Cham, 2023.

[132] V. Thomée. *Galerkin finite element methods for parabolic problems*, volume 25. Springer Science & Business Media, 2007.

[133] V. Thomée and L. B. Wahlbin. Stability and analyticity in maximum-norm for simplicial Lagrange finite element semidiscretizations of parabolic equations with dirichlet boundary conditions. *Numerische Mathematik*, 87:373–389, 2000.

[134] R. Vanselow. About Delaunay triangulations and discrete maximum principles for the linear conforming FEM applied to the Poisson equation. *Applications of Mathematics*, 46:13–28, 2001.

[135] H. von Wahl, T. Richter, and C. Lehrenfeld. An unfitted eulerian finite element method for the time-dependent Stokes problem on moving domains. *IMA Journal of Numerical Analysis*, 42(3):2505–2544, 2022.

[136] L. B. Wahlbin. Local behavior in finite element methods. *Handbook of Numerical Analysis*, 2:353–522, 1991.

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

[138] J. Wang and R. Zhang. Maximum principles for $P1$-conforming finite element approximations of quasi-linear second order elliptic equations. *SIAM Journal on Numerical Analysis*, 50(2):626–642, 2012.

[139] L. Weis. A new approach to maximal lp-regularity. In *Evolution equations and their applications in physical and life sciences*, pages 195–214. CRC Press, 2019.

[140] M. F. Wheeler. A priori L^2 error estimates for Galerkin approximations to parabolic partial differential equations. *SIAM Journal on Numerical Analysis*, 10(4):723–759, 1973.

[141] J. Xu and K. Yang. Well-posedness and robust preconditioners for discretized fluid–structure interaction systems. *Computer Methods in Applied Mechanics and Engineering*, 292:69–91, 2015.

[142] J. Xu and L. Zikatanov. A monotone finite element scheme for convection-diffusion equations. *Mathematics of Computation*, 68(228):1429–1446, 1999.

[143] G. Zhou and N. Saito. Finite volume methods for a keller–segel system: discrete energy, error estimates and numerical blow-up analysis. *Numerische Mathematik*, 135(1):265–311, 2017.

‘