UNIVERSITY OF CALIFORNIA SAN DIEGO

Some superconvergence estimates of mixed and nonconforming finite element methods

A dissertation submitted in partial satisfaction of the
requirements for the degree
Doctor of Philosophy

in

Mathematics

by

Yuwen Li

Committee in charge:

Professor Randolph E. Bank, Chair
Professor Jiun-Shyan Chen
Professor Philip Gill
Professor Michael Holst
Professor Petr Krysl

2019

The dissertation of Yuwen Li is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

_____

_____

Chair

University of California San Diego

2019

DEDICATION

For my parents who helped me in all things when I was a teenager and encouraged me to pursue an academic career.

EPIGRAPH

Think globally,

Act locally.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to express my gratitude to Prof. Randy Bank who carefully guided me, encouraged and supported me, initiated my research on finite element superconvergence, and carefully checked the draft of my dissertation. Without his support, I could not finish this dissertation.

I would also like to express my sincere thanks to Prof. Michael Holst who gave me many helpful suggestions, cooperated with me in finite element exterior calculus and thus broaden my research area.

Besides Professors Randy Bank and Michael Holst, I am also grateful to other members of my doctoral committee, Prof. Jiun-Shyan Chen, Prof. Philip Gill, Prof. Petr Krysl, for their time, encouragement, and suggestions on this dissertation.

Chapter 2, in part, contains the original results in SIAM J. Numer. Anal. 56 (2018), 792–815, Li, Yu-Wen, 2018 Society for Industrial and Applied Mathematics. The dissertation author was the author of this paper.

Chapter 2, in part, has been submitted for publication of the material as it may appear in Journal of Scientific Computing, 2019, Bank, Randolph E.; Li, Yuwen, Springer, 2019. The dissertation author was the coauthor of this paper.

Chapter 3, in full, has been submitted for publication of the material as it may appear in Numerische Mathematik, 2019, Li, Yu-Wen, Springer, 2019. The dissertation author was the author of this paper.

Chapter 4, in part, has been submitted for publication of the material as it may appear in Numerische Mathematik, 2019, Li, Yu-Wen, Springer, 2019. The dissertation author was the author of this paper.

VITA

2013	B. S. in Mathematics, Nanjing University
2016	M. S. in Mathematics, Nanjing University
2019	Ph. D. in Mathematics, University of California San Diego

PUBLICATIONS

Yu-Wen Li, "Global superconvergence of the lowest-order mixed finite element on mildly structured meshes". *SIAM J. Numer. Anal.* 56 (2018), no. 2, 792–815.

Randolph Bank and Yuwen Li, " Superconvergent recovery of Raviart–Thomas mixed finite elements on triangular grids", under review.

Yu-Wen Li, "New superconvergence estimates of the Hellan–Herrmann–Johnson and Morley element methods", under review.

Yuwen Li, " Superconvergent flux recovery of Rannacher–Turek nonconforming elements", in preparation.

Yuwen Li, "Some convergence and optimality results of adaptive mixed methods in finite element exterior calculus", under review.

Michael Holst, Yuwen Li, Adam Mihalik, and Ryan Szypowski, "Convergence and optimality of adaptive mixed methods for Poisson's equation in the FEEC framework", accepted by *Journal of Computational Mathematics.*

ABSTRACT OF THE DISSERTATION

Some superconvergence estimates of mixed and nonconforming finite element methods

by

Yuwen Li

Doctor of Philosophy in Mathematics

University of California San Diego, 2019

Professor Randolph E. Bank, Chair

In this dissertation, we develop new superconvergence estimates of mixed and nonconforming finite element methods on mildly structured grids, where most pairs of adjacent triangles form approximate parallelograms. In particular, we consider the Raviart–Thomas mixed method and Crouzeix–Raviart nonconforming method for second order elliptic equations, and the Hellan–Herrmann–Johnson mixed method and Morley nonconforming method for fourth order elliptic equations. We first prove some supercloseness estimates, that is, the canonical interpolant and finite element solution are superclose. We then develop a new family of recovery operators on irregular triangular grids by using the idea of local least-squares fittings. Combining these two ingredients, we prove that the postprocessed solution superconverges to the exact solution. Compared to existing superconvergence results, our estimates are sharper and applicable to more flexible grids and partial differential equations with variable coefficients and lower order terms.

Chapter 1

Introduction

This chapter concerns with the background of finite element superconvergence analysis and motivation of the dissertation.

1.1 List of notations

For convenience, we introduce a list of notations that will appear throughout the dissertation. Let $\Omega \subset \mathbb{R}^2$ be an open polygonal domain. Let $\mathcal{T}_h = \{T_i : i \in I\}$ be a triangulation of Ω, which is a collection of triangles such that

1. $\bigcup_{i\in I} \bar{T}_i = \bar{\Omega}$;

2. For $i, j \in I$ with $i \neq j$, $\bar{T}_i \cap \bar{T}_j$ is an edge or a vertex shared by T_i and T_j.

Here $\bar{U}$ denotes the closure of U. $\mathcal{T}_h$ is assumed to be shape regular, i.e., there exists a fixed constant $\Theta > 0$, such that $\theta \geq \Theta > 0$ for the angle θ of any $T \in \mathcal{T}_h$. The shape-regularity here based on the minimum angle condition (MAC) is equivalent to the shape-regularity based on the ratio of the radii of the inscribed and circumscribed circles of $T \in \mathcal{T}_h$. Let $h_T = |T|^{\frac{1}{2}}$ be the diameter of T and $h = \max_{T\in\mathcal{T}_h} h_T$ be the mesh-size. We also assume that $\mathcal{T}_h$ is quasi-uniform, i.e., $\max_{T\in\mathcal{T}_h} h_T \leq C(\min_{T\in\mathcal{T}_h} h_T)$ for

some generic constant C, unless we clarify the mesh condition. Given a one-dimensional or two-dimensional subset $U \in \mathbb{R}^2$, the polynomial space

$$\mathcal{P}_r(U) = \{v : v \text{ is a polynomial on } U \text{ of degree } \leq r\}$$

is useful in defining shape functions of finite elements. Let $\boldsymbol{n}$ denote the outward unit normal to ∂U and $\boldsymbol{t}$ the counterclockwise unit tangent to ∂U. We do not use $\boldsymbol{n}_U$ and $\boldsymbol{t}_U$ unless confusion arises. For example, $\partial_{\boldsymbol{n}}$ can be the outward normal derivative of the domain Ω or a triangle T.

In our error analysis of finite element methods (FEMs), the set of edges, interior edges and boundary edges in $\mathcal{T}_h$ are denoted by $\mathcal{E}_h, \mathcal{E}_h^o, \mathcal{E}_h^\partial$, respectively; the set of nodes, interior nodes and boundary nodes in $\mathcal{T}_h$ are denoted by $\mathcal{N}_h, \mathcal{N}_h^o, \mathcal{N}_h^\partial$, respectively. Several kinds of local patches are useful for finite element superconvergence analysis. For $\boldsymbol{z} \in \mathcal{N}_h$, let ω_z be the union of triangles in $\mathcal{T}_h$ sharing z as a vertex. For $e \in \mathcal{E}_h$, let ω_e be the union of triangles in $\mathcal{T}_h$ sharing e as an edge. For $T \in \mathcal{T}_h$, let ω_T be the union of T and triangles in $\mathcal{T}_h$ sharing at least one vertex with T. The local nodes, edges, and triangles in U are $\mathcal{N}_h(U) = \{z \in \mathcal{N}_h : z \in \bar{U}\}$, $\mathcal{E}_h(U) = \{e \in \mathcal{E}_h : e \subset \bar{U}\}$, and $\mathcal{T}_h(U) = \{T \in \mathcal{T}_h : T \subseteq \bar{U}\}$, respectively.

For error estimates, we need several Sobolev semi-norms and norms:

$$\begin{aligned} |v|_{k,p,U} &= \big(\int_U |D^k v|^p \big)^{\frac{1}{p}}, \quad \|v\|_{k,p,U} = \big(\sum_{m=0}^{k} |v|_{m,p,U}^p \big)^{\frac{1}{p}}, \\ |v|_{k,U} &= |v|_{k,2,U}, \quad \|v\|_{k,U} = \|v\|_{k,2,U}. \end{aligned} \tag{1.1}$$

where

$$|D^k v| := \sum_{\alpha_1+\alpha_2=k} \left| \frac{\partial^{\alpha_1+\alpha_2}}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2}} v \right|.$$

The integral in (1.1) is understood as the double integral if U is a two-dimensional sub-

set or the line integral is U is a one-dimensional subset. Sobolev norms with ∞-index, norms of vector/matrix-valued functions, and fractional order norms are generalized in usual ways. The thesis considers Sobolev spaces

$$W^{k,p}(\Omega) = \{v \in L^2(\Omega) : \|v\|_{k,p,\Omega} < \infty\}, \quad H_0^1(\Omega) = \{v \in W^{1,2}(\Omega) : v = 0 \text{ on } \partial\Omega\}.$$

Let $|U|$ denote the measure (length, area etc.) of U and $\fint_U f = \int_U f/|U|$ the average of f over U. In finite element interpolation theory, the mesh-dependent semi-norm

$$|v|_{h,k,\Omega} = \big(\sum_{T \in \mathcal{T}_h} |v|_{k,T}^2 \big)^{\frac{1}{2}}$$

is quite useful. It is also convenient to introduce the notation $\lesssim$, i.e., $A \lesssim B$ if $A \leq C \cdot B$, where C is a generic constant that may change from line to line, may depend on the domain Ω, the minimum angle θ_0, datum of PDEs, but is independent of the PDE solution, the mesh-size h, and other crucial parameters. We say $A \approx B$ if $A \lesssim B$ and $B \lesssim A$.

1.2 The finite element method and superconvergent recovery

Finite element superconvergence theory of Lagrange elements applied to second order elliptic equations has been well established (cf. [67] and references therein). Generally speaking, there are two types of superconvergence phenomena. One is called natural superconvergence which is local, that is, the finite element solution superconverges to the true solution at some special points (cf. [60]). The other is recovery-type superconvergence, which is often global (cf. [1] and references therein).

Due to its simplicity, universality, and efficiency, finite element superconvergent recovery is quite popular in practice. We illustrate it by Poisson's equation

$$\begin{aligned} -\Delta u &= f \qquad \text{in } \Omega, \\ u &= g \qquad \text{on } \partial\Omega. \end{aligned} \tag{1.2}$$

Given a mesh $\mathcal{T}_h$, let ϕ_z be the nodal basis function at $z \in \mathcal{N}_h$, that is, $\phi_z(z') = \delta_{zz'}$ for any $z' \in \mathcal{N}_h$. Here $\delta_{zz'}$ is the Kronecker delta. Let $g_h = \sum_{z\in\mathcal{N}_h^\partial} g(z)\phi_z$, namely, the boundary nodal interpolant of g. For positive integer r, define

$$\mathcal{U}_h^r = \{v_h \in H_0^1(\Omega) : v_h|_T \in \mathcal{P}_r(T), \quad \forall T \in \mathcal{T}_h\}.$$

We also need the nodal finite element space without essential boundary condition

$$\mathcal{S}_h^r = \{v_h \in H_0^1(\Omega) : v_h|_T \in \mathcal{P}_r(T), \quad \forall T \in \mathcal{T}_h\}.$$

The linear FEM for solving (1.2) is to find $u_h \in \mathcal{U}_h^1 + g_h$, such that

$$(\nabla u_h, \nabla v) = (f, v), \quad \forall v \in \mathcal{U}_h^1, \tag{1.3}$$

where $(\cdot,\cdot)$ denotes the L^2-inner product on Ω. To estimate the finite element error, we introduce the linear interpolant $v_I = \sum_{z\in\mathcal{N}_h} v(z)\phi_z$ for any continuous function v. By finite element interpolation theory, we have

$$\|v - v_I\|_{p,\Omega} \lesssim h^{2-p}|v|_{2,\Omega}, \quad p = 0, 1.$$

Assume $\mathcal{T}_h$ is uniformly parallel, i.e., each pair of adjacent triangles in $\mathcal{T}_h$ forms a

parallelogram. Then the supercloseness estimate holds

$$\|\nabla(u_I - u_h)\|_{0,\Omega} \lesssim h^{\frac{3}{2}}\|u\|_{3,\Omega}. \tag{1.4}$$

In practice, u_I is not available. To obtain a more accurate solution from (1.4), one needs the so called recovery or postprocessing operator G_h acting on ∇u_h, which is often called gradient recovery. $G_h\nabla u_h$ is often H^1-conforming, i.e., $G_h\nabla u_h \in H^1(\Omega)$, and superconverges to ∇u in some situations. In addition, G_h can be used to develop a posteriori error estimators. The recovery-based a posteriori error estimators are popular for their simplicity and asymptotic exactness (cf. [72, 73, 74, 1, 9, 71]).

For example, let $G_h\nabla u_h$ be the finite element function in $\mathcal{S}_h^1 \times \mathcal{S}_h^1$ with

$$(G_h\nabla u_h)(z) = \frac{1}{|\omega_z|}\sum_{T\ni z}|T|(\nabla u_h)|_T, \quad \forall z \in \mathcal{N}_h, \tag{1.5}$$

namely, $G_h\nabla u_h = \sum_{z\in\mathcal{N}_h}(G_h\nabla u_h)(z)\phi_z$. It is not hard to show that

$$\|G_h\nabla u_h\|_{0,\Omega} \lesssim \|\nabla u_h\|_{0,\Omega}, \tag{1.6a}$$

$$\|\nabla u - G_h\nabla u_I\| \lesssim h^{\frac{3}{2}}\|u\|_{3,\Omega}. \tag{1.6b}$$

(1.6a) is the boundedness of G_h and (1.6b) is the super-approximation property of G_h. Combining (1.6) with the supercloseness estimate (1.4) and using a triangle inequality, we obtain the recovery superconvergence

$$\|\nabla u - G_h\nabla u_h\| \le \|\nabla u - G_h\nabla u_I\| + \|G_h(\nabla u_I - \nabla u_h)\| \lesssim h^{\frac{3}{2}}\|u\|_{3,\Omega}. \tag{1.7}$$

The more accurate $G_h\nabla u_h$ can be used to develop the recovery type a posteriori estimator. For each triangle $T \in \mathcal{T}_h$, let $\eta_T = \|\nabla u_h - G_h\nabla u_h\|_T$ be the computable local

error indicator, and $\eta_h = (\sum_{T\in\mathcal{T}_h} \eta_T^2)^{\frac{1}{2}} = \|\nabla u_h - G_h\nabla u_h\|_{0,\Omega}$. The triangle inequality yields

$$\|\nabla u - \nabla u_h\|_{0,\Omega} - \|\nabla u - G_h\nabla u_h\|_{0,\Omega} \leq \eta_h \leq \|\nabla u - \nabla u_h\|_{0,\Omega} + \|\nabla u - G_h\nabla u_h\|_{0,\Omega}.$$

Under mild condition, $\|\nabla(u - u_h)\|_{0,\Omega} \approx h$. It then follows from (1.7) that

$$\lim_{h\to 0} \frac{\eta_h}{\|\nabla u - \nabla u_h\|_{0,\Omega}} = 1,$$

namely, the error estimator η_h is asymptotically equivalent to the true error $\|\nabla u - \nabla u_h\|_{0,\Omega}$, which is called asymptotic exactness. Hence we have a reliable and efficient control of the finite element error. Gradient recovery method for Lagrange elements have been studied extensively by many authors, (cf.[72, 73, 74, 9, 10, 11, 71, 68]). In the aforementioned superconvergence analysis, we define the linear interpolation $I_h^1 v = v_I$. For error analysis of the quadratic FEM, we need the quadratic Lagrange interpolation $I_h^2 : C(\Omega) \to \mathcal{S}_h^2$, where $I_h^2 v(z) = v(z)$ holds for any vertex z and midpoint z of any edge in $\mathcal{T}_h$.

1.3 Motivation and originality

While superconvergence analysis of Lagrange elements reaches its maturity (cf.[60, 9, 69] and references therein), superconvergence results of many important types of finite elements are quite limited. The goal of the dissertation is to develop sophisticated superconvergence theory for mixed and nonconforming FEMs. Mixed and nonconforming FEMs are two classes of finite element methods which are seemingly very different but indeed closely related to each other in theory. They are particularly useful for dealing with systems of equations and higher order equations. The variational

formulation of mixed FEMs uses more than one finite element space. Among numerous mixed finite elements, Raviart–Thomas (RT) finite element is the first and one of most important one for discretizing space pairs $H(\mathrm{div};\Omega) \times L^2(\Omega)$ arising from the first order formulation (2.2) of second order elliptic equations. The Hellan–Herrmann-Johnson mixed method for the fourth order elliptic equation can be viewed as a generalization of RT mixed methods. Classical superconvergence results ([12, 13, 30]) for mixed methods are typically concerned with uniformly structured meshes such as uniformly parallel or rectangular meshes. Even in these cases, the error bounds are often not sharp. In Chapters 2 and 3, we shall address the originality of the part on our superconvergence analysis of mixed FEMs in detail. On the other hand, the nonconforming FEM is a family of FEMs that are seemingly different from mixed FEMs based on suitably relaxing the inter-element continuity constraint in standard FEMs. Due to the deviation from conformity, superconvergence analysis of nonconforming FEMs is very difficult. Arnold and Brezzi [3] have shown that the Crouzeix–Raviart (CR) nonconforming method and RT mixed method are equivalent for Poisson's equation. Similarly, the Morley nonconforming method and HHJ mixed method are equivalent for the biharmonic equation. Built upon this equivalence, [36] gave a superconvergence estimate for the CR and Morley element on uniformly parallel grids. However, this result are limited to uniformly structured meshes and not sharp. In addition, it only works for Poisson's equation because of the equivalence only applies to Poisson's equation. In Chapters 4, we shall address the originality of our superconvergence analysis of nonconforming FEMs in detail. Here we briefly list the originality of the dissertation.

1. Our superconvergence estimates hold on meshes satisfying the (α, β)-condition (see the next section) instead of uniform grids. As pointed out in [9, 69], the (α, β)-condition is very flexible and satisfied by many mature finite element codes.

2. The proof of superconvergence estimates are constructive. It shows how the lower order error cancels in a transparent way and characterizes the rate of superconvergence in terms of two parameters α, β.

3. Our superconvergence estimates deal with variable coefficient problems. The error analysis of mixed methods with lower order terms is often much more involved (cf.[27, 7]).

4. For the lowest order RT, HHJ mixed methods and CR, Morley nonconforming methods, classical superconvergence estimates are suboptimal. Our estimates improve existing ones by $h^{\frac{1}{2}}$.

5. We construct the recovery operator R_h based on solving local least squares problems. The super-approximation property of R_h is better than simple averaging procedures. R_h can be generalized to mixed finite elements of general order.

1.4 Mildly structured meshes

$\mathcal{T}_h$ is rarely uniformly parallel in practice. In order to generalize the superconvergent recovery framework to a more realistic setting, one needs to deal with more flexible mesh structures. In the dissertation, we focus on mildly structured grids, which were introduced in [48, 9, 69, 42, 46] and references therein. Roughly speaking, most pairs of adjacent triangles in mildly structured grids form approximate parallelograms except for a region with small measure.

We begin with geometric identities on a local element T. It has three vertices $\{z_k\}_{k=1}^3$, oriented counterclockwise, and corresponding barycentric coordinates $\{\lambda_k\}_{k=1}^3$. Let e_k denote the edge opposite to z_k, θ_k the angle opposite to e_k, ℓ_k the length of e_k, d_k the distance from z_k to e_k, $\boldsymbol{t}_k$ the unit tangent to e_k, oriented counterclockwise, $\boldsymbol{n}_k$

the unit outward normal to e_k, $\partial_{\boldsymbol{t}_k}$ the tangential derivative, $\partial_{\boldsymbol{n}_k}$ the normal derivative, and $\partial^2_{\boldsymbol{t}_k\boldsymbol{n}_k}$ the second mixed derivative, see (1.1). Corresponding quantities on triangles T' and T'' have superscripts $\prime$ and $\prime\prime$ respectively. The subscripts are equivalent mod 3, e.g., $\ell_4 = \ell_1, \theta_0 = \theta_3$.

Definition 1.4.1. For $e \in \mathcal{E}_h^o$, let $T, T' \in \mathcal{T}_h$ be the two adjacent elements sharing e. Define $e_1 = e_1' = e$. By going along ∂T and $\partial T'$ counterclockwise, we obtain other two pairs of corresponding edges e_2, e_2' and e_3, e_3'. We say $\omega_e = T \cup T'$ is an $O(h^{1+\alpha})$-approximate parallelogram if $|e_i| = |e_i'| + O(h^{1+\alpha})$, $i = 1, 2, 3$.

Definition 1.4.2. For $z \in \mathcal{N}_h^\partial$, let $e, e' \in \mathcal{E}_h^\partial$ be the two boundary edges sharing z as an endpoint. Let $\boldsymbol{t}$ and $\boldsymbol{t}'$ be the unit tangents to e and e' respectively, oriented counterclockwise. Let $T, T' \in \mathcal{T}_h$ be the two triangles having $e = e_1$ and $e' = e_1'$ as edges respectively. By going along ∂T and $\partial T'$ counterclockwise, we have other two pairs of corresponding edges e_2, e_2' and e_3, e_3'. We say that $\{T, T'\}$ is an $O(h^{1+\alpha})$-approximately parallel pair if $|\boldsymbol{t} - \boldsymbol{t}'| = O(h^\alpha)$ and $|e_i| = |e_i'| + O(h^{1+\alpha})$, $i = 1, 2, 3$.

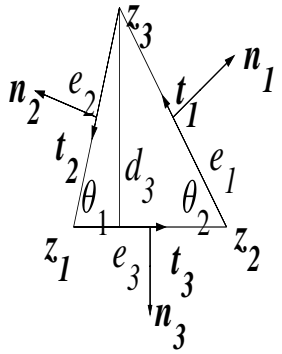

Figure 1.1: a local triangle and associated quantities

Figure 1.2 illustrates Definitions 1.4.1 and 1.4.2. For uniformly parallel grids, $\alpha = \infty$, i.e., any two corresponding edges in both Definitions 1.4.1 and 1.4.2 are parallel and of equal length. $\alpha = 1$ in [9], and $\alpha > 0$ in this dissertation.

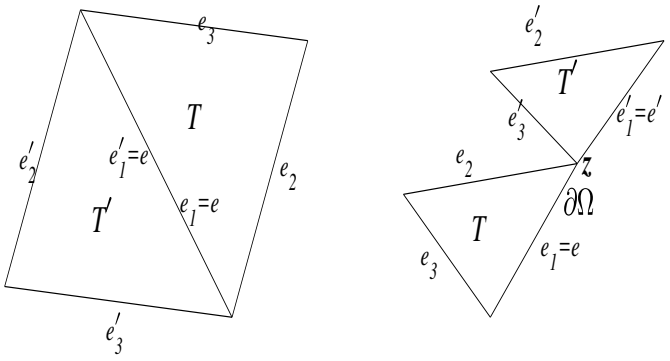

Figure 1.2: (left)approximate parallelogram (right)approximately parallel pair

Built upon approximate parallelograms, we introduce several types of mildly structured meshes characterized by two parameters α and β. $\biguplus$ stands for the disjoint union.

Definition 1.4.3. Let $\mathcal{E}_h^o = \mathcal{E}_{h,1}^o \biguplus \mathcal{E}_{h,2}^o$. We say the triangulation $\mathcal{T}_h$ satisfies the (α, β)-condition if for each $e \in \mathcal{E}_{h,1}^o$, ω_e an $O(h^{1+\alpha})$-approximate parallelogram, while $\sum_{e\in\mathcal{E}_{h,2}^o} |\omega_e| = O(h^\beta)$.

Definition 1.4.4. Assume (α, β)-condition holds and $\mathcal{N}_h^\partial = \mathcal{N}_{h,1}^\partial \biguplus \mathcal{N}_{h,2}^\partial$. We say $\mathcal{T}_h$ satisfies the strong (α, β)-condition if the adjacent boundary elements T and T' in Definition 1.4.2 associated with each $z \in \mathcal{N}_{h,1}^\partial$ form an $O(h^{1+\alpha})$-approximately parallel pair, while $\#\mathcal{N}_{h,2}^\partial$ is uniformly bounded w.r.t. h. We say $\mathcal{T}_h$ satisfies the piecewise strong (α, β)-condition if Ω can be decomposed into finitely many subdomains aligned with $\mathcal{T}_h$ and $\mathcal{T}_h$ restricted to each subdomain satisfies the strong (α, β)-condition.

For example, if Ω is a parallelogram with uniformly parallel triangulation $\mathcal{T}_h$, then $\mathcal{N}_{h,2}^\partial$ is the set of corner points, $\#\mathcal{N}_{h,2}^\partial = 4$, $\mathcal{N}_{h,1}^\partial = \mathcal{N}_h^\partial \backslash \mathcal{N}_{h,2}^\partial$ and $\mathcal{T}_h$ satisfies the strong (α, β)-condition with $\alpha = \beta = \infty$. The piecewise strong (α, β)-condition is weaker than the strong (α, β)-condition. For more examples, readers are referred to numerical experiments in the dissertation.

1.5 Integral inequalities

Inequalities are essential in finite element analysis. We list several important inequalities that will be used in following chapters. The following error estimate is the standard error bound for nodal interpoaltion (cf.[14]).

Theorem 1.5.1. *For $T \in \mathcal{T}_h$ and $v \in H^{k+1}(T)$ with $k \geq 1$,*

$$\|v - I_h^k v\|_{0,\gamma,T} \lesssim h^{k+\frac{2}{\gamma}} |v|_{k+1,T}, \quad 1 \leq \gamma \leq \infty. \tag{1.8}$$

The next inequality gives an upper bound of boundary norm in terms of volume norms.

Theorem 1.5.2 (Trace inequalities)**.** *For $T \in \mathcal{T}_h$,*

$$\begin{aligned} \|v\|_{0,\partial T} &\lesssim h_T^{-\frac{1}{2}} \|v\|_{0,T} + h_T^{\frac{1}{2}} \|\nabla v\|_{0,T}, \quad v \in H^1(T), \\ \int_{\partial T} |v| &\lesssim h_T^{-1} \int_T |v| + \int_T |\nabla v|, \quad v \in W^{1,1}(T). \end{aligned} \tag{1.9}$$

We also need the well-known inverse inequality for polynomials.

Theorem 1.5.3 (Inverse inequality)**.** *For $T \in \mathcal{T}_h$ and $v \in \mathcal{P}_r(T)$ with some integer $r > 0$,*

$$\|v\|_{k,p,T} \lesssim h_T^{-k+\min(\frac{2}{p}-\frac{2}{q},0)} \|v\|_{0,q,T}. \tag{1.10}$$

It is well known that $H^1(\Omega) \not\subseteq C(\Omega)$ in $\mathbb{R}^2$ and thus the Sobolev inequality $\|v\|_{0,\infty,\Omega} \lesssim \|v\|_{1,\Omega}$ fails for general continuous v. However, a discrete Sobolev inequality holds for $v \in \mathcal{S}_h^r$.

Theorem 1.5.4 (Discrete Sobolev inequality)**.** *For $v_h \in \mathcal{S}_h^r$ with fixed $r \geq 1$,*

$$\|v_h\|_{0,\infty,\Omega} \lesssim |\log h|^{\frac{1}{2}} \|v_h\|_{1,\Omega}. \tag{1.11}$$

Chapter 2

Superconvergence of Raviart–Thomas elements

2.1 Mixed methods and Raviart–Thomas finite elements

In this chapter, we consider the second order elliptic equation

$$-\operatorname{div}(\boldsymbol{a}_2(\boldsymbol{x})\nabla u + \boldsymbol{a}_1(\boldsymbol{x})u) + a_0(\boldsymbol{x})u = f(\boldsymbol{x}), \quad \boldsymbol{x} \in \Omega, \tag{2.1a}$$

$$u = g(\boldsymbol{x}), \quad \boldsymbol{x} \in \partial\Omega, \tag{2.1b}$$

where $\operatorname{div} = \nabla\cdot$ is the divergence operator, $\boldsymbol{a}_2 = (a_{ij})_{1\le i,j\le 2}$ is a symmetric matrix, $\boldsymbol{a}_1$ is vector-valued, $\boldsymbol{a}_2, \boldsymbol{a}_1$ and a_0 are sufficiently smooth on $\overline{\Omega}$. In addition, $\boldsymbol{a}_2$ is uniformly elliptic in the sense that $\sum_{i,j} a_{ij}(\boldsymbol{x})\xi_i\xi_j \ge \Lambda|\boldsymbol{\xi}|^2$ for all $\boldsymbol{x} \in \Omega$ and $\boldsymbol{\xi} \in \mathbb{R}^2$,

where $\Lambda > 0$ is a constant. Let

$$\begin{aligned}
&\boldsymbol{p} = \boldsymbol{a}_2 \nabla u + \boldsymbol{a}_1 u, \\
&\boldsymbol{a} = \boldsymbol{a}_2^{-1}, \quad \boldsymbol{b} = \boldsymbol{a}_2^{-1}\boldsymbol{a}_1, \quad c = a_0.
\end{aligned}$$

Problem (2.1) is equivalent to the first order system

$$\boldsymbol{a}\boldsymbol{p} - \boldsymbol{b}u - \nabla u = 0, \quad \boldsymbol{x} \in \Omega, \tag{2.2a}$$

$$-\operatorname{div} \boldsymbol{p} + cu = f, \quad \boldsymbol{x} \in \Omega, \tag{2.2b}$$

$$u = g, \quad \boldsymbol{x} \in \partial\Omega. \tag{2.2c}$$

Let $\mathcal{Q} = \{\boldsymbol{q} \in L^2(\Omega)^2 : \operatorname{div} \boldsymbol{q} \in L^2(\Omega)\}$ and $\mathcal{V} = L^2(\Omega)$. The mixed formulation for (2.2) is to find the pair $\{\boldsymbol{p}, u\} \in \mathcal{Q} \times \mathcal{V}$, such that

$$(\boldsymbol{a}\boldsymbol{p}, \boldsymbol{q}) - (\boldsymbol{q}, \boldsymbol{b}u) + (\operatorname{div} \boldsymbol{q}, u) = \langle \boldsymbol{q} \cdot \boldsymbol{n}, g \rangle, \tag{2.3a}$$

$$-(\operatorname{div} \boldsymbol{p}, v) + (cu, v) = (f, v), \tag{2.3b}$$

for each pair $\{\boldsymbol{q}, v\} \in \mathcal{Q} \times \mathcal{V}$. Here $\langle \cdot, \cdot \rangle$ denotes the L^2-inner product on $\partial\Omega$. For $r \geq 0$ and $T \in \mathcal{T}_h$, define

$$\mathcal{RT}_r(T) := \left\{ \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} + v_3 \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} : v_i \in \mathcal{P}_r(T), i = 1, 2, 3 \right\}. \tag{2.4}$$

The Raviart–Thomas finite element spaces with index $r \geq 0$ (denoted by RT_r) are

$$\begin{aligned}
&\mathcal{Q}_h^r := \{\boldsymbol{q}_h \in \mathcal{Q} : \boldsymbol{q}_h|_T \in \mathcal{RT}_r(T), \ \forall T \in \mathcal{T}_h\}, \\
&\mathcal{V}_h^r := \{v_h \in \mathcal{V} : v_h|_T \in \mathcal{P}_r(T), \ \forall T \in \mathcal{T}_h\}.
\end{aligned}$$

The mixed finite element approximation to the problem (2.3) is to find $\{\boldsymbol{p}_h^r, u_h^r\} \in \mathcal{Q}_h^r \times \mathcal{V}_h^r$, such that

$$(\boldsymbol{a}\boldsymbol{p}_h^r, \boldsymbol{q}_h) - (\boldsymbol{q}_h, \boldsymbol{b}u_h^r) + (\operatorname{div} \boldsymbol{q}_h, u_h^r) = \langle \boldsymbol{q}_h \cdot \boldsymbol{n}, g\rangle, \qquad \boldsymbol{q}_h \in \mathcal{Q}_h^r, \tag{2.5a}$$

$$-(\operatorname{div} \boldsymbol{p}_h^r, v_h) + (cu_h^r, v_h) = (f, v_h), \qquad v_h \in \mathcal{V}_h^r. \tag{2.5b}$$

Under mild assumptions, Douglas and Roberts [27] has shown the well-posedness and a priori error estimates for the method (2.5).

In this chapter, we shall prove supercloseness estimate and present superconvergent recovery procedure for the lowest order (r=0) and next lowest order (r=1) RT finite elements. In particular, we shall prove that

$$\|\Pi_h^r \boldsymbol{p} - \boldsymbol{p}_h^r\|_{0,\Omega} = O(h^{r+1+\rho_r}), \tag{2.6a}$$

$$\|\operatorname{div}(\Pi_h^r \boldsymbol{p} - \boldsymbol{p}_h^r)\|_{0,\Omega} = O(h^{r+2}), \tag{2.6b}$$

where the canonical interpolation Π_h^r is defined in (2.10), $0 < \rho_0 \leq 1, 0 < \rho_1 \leq 1/2$ are rates of superconvergence determined by the parameters in Definitions 1.4.3 and 1.4.4. (2.6b) holds on general shape regular meshes while (2.6a) holds on mildly structured quasi-uniform $\mathcal{T}_h$.

(2.6) is closely related to superconvergence of the finite element solution to the exact solution. For example, $\boldsymbol{p}_h^0$ can be postprocessed by a simple local averaging operator K_h proposed in [12]. Using the super-approximation property of K_h and (2.6) with $k = 0$, we obtain the superconvergence estimate

$$\|\boldsymbol{p} - K_h \boldsymbol{p}_h^0\|_{0,\Omega} = O(h^{1+\rho_0}). \tag{2.7}$$

The recovered flux $K_h \boldsymbol{p}_h^0$ can be used to develop a posteriori error estimates. Due to

the superconvergence (2.7), $\|K_h\boldsymbol{p}_h^0 - \boldsymbol{p}_h^0\|_{0,\Omega}$ is known to be an asymptotically exact a posteriori estimator for $\|\boldsymbol{p} - \boldsymbol{p}_h^0\|_{0,\Omega}$ (cf. [12, 10, 1]), namely,

$$\lim_{h\to 0} \frac{\|K_h\boldsymbol{p}_h^0 - \boldsymbol{p}_h^0\|_{0,\Omega}}{\|\boldsymbol{p} - \boldsymbol{p}_h^0\|_{0,\Omega}} = 1.$$

The study of supercloseness between the finite element interpolant and finite element solution has a long history. For the analogue of (2.6a) for Lagrange elements on mildly structured grids, see [9, 69, 42] and references therein. For superconvergence of the scalar variable u in mixed methods, see [3, 15, 65] and references therein. In practice, it is frequently the case that the vector variable $\boldsymbol{p}$ is more important than the scalar u. Superconvergence results of rectangular/quadrilateral mixed finite elements for the vector variable $\boldsymbol{p}$ are well established (cf. [30, 32, 31]). However, corresponding superconvergence theory of triangular mixed finite elements are much less sophisticated. To our best knowledge, the only proven superconvergence estimate of triangular mixed elements for the vector variable are in [28, 12, 13]. In [28], the authors postprocessed $\boldsymbol{p}_h$ and achieved interior superconvergence by convolution with a Bramble–Schatz kernel [18] which is constructed on uniform grids. For RT_0 element on uniformly parallel grids in the case that $\boldsymbol{b} = \boldsymbol{0}, c = 0$, Brandts [12] proved

$$\|\Pi_h^0\boldsymbol{p} - \boldsymbol{p}_h^0\|_{0,\Omega} \lesssim h^{\frac{3}{2}}(\|\boldsymbol{p}\|_{\frac{3}{2},\Omega} + h^{\frac{1}{2}}|\boldsymbol{p}|_{1,\Omega} + h^{\frac{1}{2}}|\boldsymbol{p}|_{2,\Omega}), \tag{2.8}$$

In [13], he also proved an analogue of (2.8) for RT_1 elements on uniformly parallel grids in the case that $\boldsymbol{A} = I_{2\times 2}, \boldsymbol{b} = \boldsymbol{0}, c = 0$.

Our result (2.6) improves existing results in several ways. First, our estimate holds on general mildly structured grids instead of uniform grids. As pointed out in [9, 69], the (α, β)-condition is very flexible and satisfied by many mature finite element codes. Second, for RT_0 element and $\rho_0 = 1$, (2.6a) becomes $\|\Pi_h^0\boldsymbol{p} - \boldsymbol{p}_h^0\|_{0,\Omega} =$

$O(h^2|\log h|^{\frac{1}{2}})$, which shows that the estimate (2.8) is suboptimal. This improvement results from carefully handling the boundary error, which is usually the trickiest part in global superconvergence estimates. In addition, due to the cancellation of errors on boundary elements, (2.6a) holds on not only (α, β)-grids but also piecewise (α, β)-grids. Third, our superconvergence results allow the convection term $\boldsymbol{b}(x) \cdot \nabla u$ and reaction term $c(x)u$. The error analysis of mixed methods with lower order terms is often much more involved than the case $\boldsymbol{b} = \mathbf{0}, c = 0$ (cf.[27, 7]).

The framework of the proof of (2.6a) is as follows. We consider the orthogonal decomposition $\Pi_h^r \boldsymbol{p} - \boldsymbol{p}_h^r = \nabla^\perp r_h \oplus \operatorname{grad}_h v_h$, where $r_h \in \mathcal{S}_h^{r+1}$ and $v_h \in \mathcal{V}_h^r$. First, we develop the variational error expansion for RT_0 and RT_1 elements on a local triangle in terms of $\partial_{\boldsymbol{t}_k} r_h$, the tangential derivative of r_h on the e_k, $k = 1, 2, 3$. Due to the continuity of $\partial_{\boldsymbol{t}_k} r_h$ on e_k and the (α, β)-condition, the lower order global variational error associated with interior edges is canceled in a transparent way instead of using soft analysis tools (the Bramble–Hilbert lemma etc., cf. [12, 13]). The aforementioned basic idea is motivated by Bank and Xu [9]. But the technicality here is quite different because of the apparent difference between Lagrange elements and RT elements. Then $\|\nabla^\perp r_h\|_{0,\Omega}$ can be bounded by the supersmall global variational error while $\|\operatorname{grad}_h v_h\|_{0,\Omega}$ is estimated by (2.6b). To obtain optimal order global superconvergence, the error associated with another part occurring on triangles near the boundary is treated carefully by the discrete Sobolev inequalities.

The second major component of this chapter is the recovery operator R_h^r based on solving local least squares problems. For Raviart–Thomas mixed methods, $\boldsymbol{p}_h^r$ instead of u_h^r is the main quantity of physical interest. As far as we know, the implementation of existing recovery techniques are restricted to uniformly structured grids (cf.[28, 30, 29, 13]). In addition, most of the existing results of recovery methods focus on the lowest order case while the analysis of recovery operators for higher order

elements is indeed much more involved, especially on unstructured grids. We shall construct a new family of recovery operators R_h^r for RT_r ($r \geq 0$) elements by fitting the numerical solution $\boldsymbol{p}_h^r$ with a vector polynomial of degree $r+1$ in the least-squares (LS) sense on each local patch ω_z. Assuming each LS problem has a unique solution, we show $\|\boldsymbol{p} - R_h^r \Pi_h^r \boldsymbol{p}\| = O(h^{r+2})$. The rate of superconvergence is almost independent of the mesh structure. In the end, combining R_h^r with the estimate (2.6a), we obtain the superconvergence

$$\|\boldsymbol{p} - R_h^r \boldsymbol{p}_h^r\|_{0,\Omega} = O(h^{r+1+\rho_r})$$

by recovery.

LS-type recovery is not a new idea, e.g., the famous Zienkiewicz–Zhu(ZZ) superconvergence patch recovery (cf.[73, 74]). Zhang and Naga [71] proposed a different LS-based patch recovery operator G_h^{r+1} for Lagrange elements of degree $r+1$ by post-processing the Lagrange finite element solution u_h rather than ∇u_h. R_h^r can be viewed as a Raviart–Thomas version of it. One can also see nice relationship between G_h^{r+1} and R_h^r in Lemma 2.4.2 and Theorem 3.3.1. Despite the excellent superconvergence property of G_h^{r+1} in practice, it is difficult to prove the uniqueness of local LS solution related to G_h^{r+1} on irregular grids. In fact [57] is essentially devoted to the analysis of the uniqueness of the LS solution for G_h^1 on irregular grids and there is no proof for $G_h^{r+1}, r \geq 1$. We give a proof for G_h^2 on irregular grids, which in turn implies the boundedness and superconvergence of R_h^1, see Theorems 2.4.3, 2.4.4 and 2.4.5.

The rest of this chapter is organized as follows. Section 2.2 contains technical local error expansions. In Section 2.3, we estimate the global variational error and derive supercloseness estimates. In Section 2.4, we develop the superconvergent recovery operators. Numerical examples are presented in Section 2.6.

2.2 Local error expansions

We have the rotational gradient

$$\nabla^{\perp} v = \left(-\frac{\partial v}{\partial x_2}, \frac{\partial v}{\partial x_1} \right)^{\mathsf{T}},$$

and its adjoint

$$\nabla \times \boldsymbol{q} = \frac{\partial q_2}{\partial x_1} - \frac{\partial q_1}{\partial x_2}.$$

$\nabla^{\perp}$ and $\nabla\times$ are related by Green's formula

$$\int_T \nabla^{\perp} r \cdot \boldsymbol{q} = \int_{\partial T} r \boldsymbol{q} \cdot \boldsymbol{t} - \int_T r \nabla \times \boldsymbol{q}, \tag{2.9}$$

where $\boldsymbol{t}$ is the unit tangent to ∂T oriented counterclockwise. For $\boldsymbol{v} \in \mathbb{R}^2$, define $\boldsymbol{v}^{\perp} = (-v_2, v_1)$. Clearly, $\boldsymbol{n}_k^{\perp} = \boldsymbol{t}_k$, $\boldsymbol{t}_k^{\perp} = -\boldsymbol{n}_k$.

Now we introduce basic definitions for Raviart–Thomas elements. On a triangle $T \in \mathcal{T}_h$, the degrees of freedom of the RT_0 element are

$$\mathcal{N}_k(\boldsymbol{q}) = \fint_{e_k} \boldsymbol{q} \cdot \boldsymbol{n}_k, \quad 1 \le k \le 3.$$

The degrees of freedom of the RT_1 element are

$$\begin{aligned}
\mathcal{N}_{1,k}(\boldsymbol{q}) &= \fint_{e_k} \lambda_{k-1} \boldsymbol{q} \cdot \boldsymbol{n}_k, \quad 1 \le k \le 3, \\
\mathcal{N}_{2,k}(\boldsymbol{q}) &= \fint_{e_k} \lambda_{k+1} \boldsymbol{q} \cdot \boldsymbol{n}_k, \quad 1 \le k \le 3, \\
\mathcal{N}_{3,l}(\boldsymbol{q}) &= \fint_T q_l, \quad 1 \le l \le 2, \quad \boldsymbol{q} = (q_1, q_2)^{\mathsf{T}}.
\end{aligned}$$

For $\boldsymbol{q} \in \mathcal{Q}$, $\Pi_h^0 \boldsymbol{q} \in \mathcal{Q}_h^0$ is determined by

$$\mathcal{N}_k(\Pi_h^0 \boldsymbol{q}|_T) = \mathcal{N}_k(\boldsymbol{q}|_T), \quad 1 \le k \le 3, \quad \forall T \in \mathcal{T}_h, \tag{2.10}$$

$\Pi_h^1 \boldsymbol{q} \in \mathcal{Q}_h^1$ is determined by

$$\begin{aligned} &\mathcal{N}_{1,k}(\Pi_h^1 \boldsymbol{q}|_T) = \mathcal{N}_{1,k}(\boldsymbol{q}|_T), \\ &\mathcal{N}_{2,k}(\Pi_h^1 \boldsymbol{q}|_T) = \mathcal{N}_{2,k}(\boldsymbol{q}|_T), \quad 1 \le k \le 3, \\ &\mathcal{N}_{3,l}(\Pi_h^1 \boldsymbol{q}|_T) = \mathcal{N}_{3,l}(\boldsymbol{q}|_T) \quad l = 1, 2, \quad \forall T \in \mathcal{T}_h. \end{aligned} \tag{2.11}$$

The existence and uniqueness of $\Pi_h^r \boldsymbol{q}$ is always guaranteed. For $v \in \mathcal{V}$, $P_h^r v$ is the L^2-projection of v onto $\mathcal{V}_h^r$. There is a very nice relation among P_h^k, Π_h^r and div, namely

$$\operatorname{div}(\Pi_h^r \boldsymbol{q}) = P_h^r(\operatorname{div} \boldsymbol{q}), \quad \forall \boldsymbol{q} \in \mathcal{Q}. \tag{2.12}$$

(2.12) is crucial to the stability and error analysis of mixed methods (cf. [59, 5]). In addition, the following approximation properties hold:

$$\|\boldsymbol{q} - \Pi_h^r \boldsymbol{q}\|_{0,\Omega} \lesssim h^{r+1} |\boldsymbol{q}|_{h,r+1,\Omega}, \tag{2.13a}$$

$$\|\operatorname{div}(\boldsymbol{q} - \Pi_h^r \boldsymbol{q})\|_{0,\Omega} \lesssim h^{r+1} |\operatorname{div} \boldsymbol{q}|_{h,r+1,\Omega}, \tag{2.13b}$$

$$\|v - P_h^r v\|_{0,\Omega} \lesssim h^{r+1} |v|_{h,r+1,\Omega}. \tag{2.13c}$$

In the rest of this section, we will present technical variational error expansions for RT_0 and RT_1 elements. To shed some light on the global picture of this framework, we first expand the interpolation error of linear functions for RT_0 elements. The theory of RT_1 is similar although much more complicated. Let $\phi_k = \lambda_{k-1}\lambda_{k+1}$ be the quadratic bump function on e_k.

Lemma 2.2.1 (Lemma 2.2, [46]). *For* $\boldsymbol{p}_1 \in \mathcal{P}_1(T)^2$,

$$\boldsymbol{p}_1 - \Pi_h^0 \boldsymbol{p}_1 = \nabla^\perp w,$$

where

$$w = \sum_{k=1}^{3} \frac{\ell_k^2}{2} \boldsymbol{n}_k \cdot \partial_{\boldsymbol{t}_k} \boldsymbol{p}_1 \phi_k.$$

Proof. Using (2.12), we have

$$\operatorname{div}(\boldsymbol{p}_1 - \Pi_h^0 \boldsymbol{p}_1) = \operatorname{div} \boldsymbol{p}_1 - P_h^0 \operatorname{div} \boldsymbol{p}_1 = 0.$$

Hence there exists a $w \in \mathcal{P}_1(T)$, such that

$$\boldsymbol{p}_1 - \Pi_h^0 \boldsymbol{p}_1 = \nabla^\perp w. \tag{2.14}$$

Using the definition of Π_h^0, we know $\mathcal{N}_k(\boldsymbol{p}_1 - \Pi_h^0 \boldsymbol{p}_1) = 0$ and thus

$$w(\boldsymbol{z}_{k-1}) - w(\boldsymbol{z}_{k+1}) = \ell_k \mathcal{N}_k(\nabla^\perp w) = 0.$$

Subtracting a constant from w, we can assume that $w(\boldsymbol{z}_k) = 0$ for $k = 1, 2, 3$. Hence $w = \sum_{k=1}^{3} \alpha_k \phi_k$. It remains to verify that

$$\alpha_k = \frac{\ell_k^2}{2} \boldsymbol{n}_k \cdot \partial_{\boldsymbol{t}_k} \boldsymbol{p}_1. \tag{2.15}$$

Applying the operator $\boldsymbol{n}_k \cdot \partial_{\boldsymbol{t}_k}$ to (2.14) gives

$$\boldsymbol{n}_k \cdot \left(\partial_{\boldsymbol{t}_k} \boldsymbol{p}_1 - \partial_{\boldsymbol{t}_k} \Pi_h^0 \boldsymbol{p}_1 \right) = -\partial_{\boldsymbol{t}_k}^2 w. \tag{2.16}$$

$\Pi_h^0 \boldsymbol{p}_1 \in \mathcal{RT}_0(T)$ implies $\partial_{\boldsymbol{t}_k} \Pi_h^0 \boldsymbol{p}_1$ is proportional to $\boldsymbol{t}_k$ and the left hand side of (2.16) is merely $\boldsymbol{n}_k \cdot \partial_{\boldsymbol{t}_k} \boldsymbol{p}_1$. For the right hand side,

$$\partial_{\boldsymbol{t}_k}^2 w = 2\alpha_k \partial_{\boldsymbol{t}_k} \lambda_{k-1} \partial_{\boldsymbol{t}_k} \lambda_{k+1} = -\frac{2\alpha_k}{\ell_k^2}. \tag{2.17}$$

Combing (2.16) and (2.17), we obtain (2.15). The proof is complete. □

The next lemma is our main technical tool for estimating the global variational error of the RT_0 mixed method. We shall use elementary identities

$$\begin{aligned} \boldsymbol{n}_{k-1} &= -\sin\theta_{k+1}\boldsymbol{t}_k - \cos\theta_{k+1}\boldsymbol{n}_k, \quad \boldsymbol{t}_{k-1} = -\cos\theta_{k+1}\boldsymbol{t}_k + \sin\theta_{k+1}\boldsymbol{n}_k, \\ \boldsymbol{n}_{k+1} &= \sin\theta_{k-1}\boldsymbol{t}_k - \cos\theta_{k-1}\boldsymbol{n}_k, \quad \boldsymbol{t}_{k+1} = -\cos\theta_{k-1}\boldsymbol{t}_k - \sin\theta_{k-1}\boldsymbol{n}_k. \end{aligned} \tag{2.18}$$

Throughout this chapter, let $\bar{\boldsymbol{a}}$ be a constant symmetric matrix and $\mu_k = \boldsymbol{t}_{k-1} \cdot \bar{\boldsymbol{a}} \boldsymbol{t}_{k+1}$. For each edge e_k, we define several associated geometric quantities $\{\mu_{j,k}^i\}_{1\le i,j\le 2}$

$$\mu_{1,k}^1 = \frac{\mu_k |T|}{6\sin\theta_k}, \quad \mu_{2,k}^1 = 0, \quad \mu_{1,k}^2 = \frac{\mu_k}{12\sin\theta_k}(\ell_{k-1}^2 - \ell_{k+1}^2), \quad \mu_{2,k}^2 = -\frac{\mu_k |T|}{6\sin\theta_k},$$

and first order derivatives $\mathcal{D}_{i,k}^j$

$$\mathcal{D}_{1,k}^1 = \boldsymbol{t}_k \cdot \partial_{\boldsymbol{t}_k}, \quad \mathcal{D}_{1,k}^2 = \boldsymbol{t}_k \cdot \partial_{\boldsymbol{n}_k}, \quad \mathcal{D}_{2,k}^1 = \boldsymbol{n}_k \cdot \partial_{\boldsymbol{t}_k}, \quad \mathcal{D}_{2,k}^2 = \boldsymbol{n}_k \cdot \partial_{\boldsymbol{n}_k}.$$

Finally, $\mathcal{A}_k = \mu_{j,k}^i \mathcal{D}_{i,k}^j := \mu_{1,k}^1 \mathcal{D}_{1,k}^1 + \mu_{2,k}^1 \mathcal{D}_{1,k}^2 + \mu_{1,k}^2 \mathcal{D}_{2,k}^1 + \mu_{2,k}^2 \mathcal{D}_{2,k}^2$. In this notation, w in Lemma 2.2.1 is $w = \sum_{k=1}^3 \frac{\ell_k^2}{2} \mathcal{D}_{2,k}^1(\boldsymbol{p}_1)\phi_k$.

Lemma 2.2.2 (Lemma 3.1, [46]). *For $\boldsymbol{p}_1 \in \mathcal{P}_1(T)^2$ and $w_1 \in \mathcal{P}_1(T)$,*

$$\int_T (\boldsymbol{p}_1 - \Pi_h^0 \boldsymbol{p}_1) \cdot \bar{\boldsymbol{a}} \nabla^\perp w_1 = \sum_{k=1}^3 \int_{e_k} \mathcal{A}_k(\boldsymbol{p}_1) \partial_{\boldsymbol{t}_k} w_1. \tag{2.19}$$

Proof. Using Lemma 2.2.1, formula (2.9), and Simpson's quadrature rule,

$$\begin{aligned}\int_T (\boldsymbol{p}_1 - \Pi_h^0 \boldsymbol{p}_1) \cdot \bar{\boldsymbol{a}} \nabla^\perp w_1 &= \int_T \nabla^\perp w \cdot \bar{\boldsymbol{a}} \nabla^\perp w_1 \\ &= \sum_{k=1}^3 \int_{e_k} w (\bar{\boldsymbol{a}} \nabla^\perp w_1) \cdot \boldsymbol{t}_k = \frac{1}{12} \sum_{k=1}^3 \ell_k^3 \mathcal{D}_{2,k}^1(\boldsymbol{p}_1) \nabla^\perp w_1 \cdot \bar{\boldsymbol{a}} \boldsymbol{t}_k.\end{aligned}$$

It then follows from

$$\bar{\boldsymbol{a}} \boldsymbol{t}_k = \frac{\mu_{k-1}}{\sin\theta_k} \boldsymbol{n}_{k-1} - \frac{\mu_{k+1}}{\sin\theta_k} \boldsymbol{n}_{k+1} \tag{2.20}$$

that

$$\begin{aligned}&\int_T (\boldsymbol{p}_1 - \Pi_h^0 \boldsymbol{p}_1) \cdot \bar{\boldsymbol{a}} \nabla^\perp w_1 \\ &= \frac{1}{12} \sum_{k=1}^3 \ell_k^3 \mathcal{D}_{2,k}^1(\boldsymbol{p}_1) \left(\frac{\mu_{k+1}}{\sin\theta_k} \partial_{\boldsymbol{t}_{k+1}} w_1 - \frac{\mu_{k-1}}{\sin\theta_k} \partial_{\boldsymbol{t}_{k-1}} w_1 \right) \\ &= \sum_{k=1}^3 \frac{\mu_k}{12} \left(\ell_{k-1}^3 \mathcal{D}_{2,k-1}^1(\boldsymbol{p}_1) \frac{1}{\sin\theta_{k-1}} - \ell_{k+1}^3 \mathcal{D}_{2,k+1}^1(\boldsymbol{p}_1) \frac{1}{\sin\theta_{k+1}} \right) \partial_{\boldsymbol{t}_k} w_1 \\ &= \sum_{k=1}^3 \frac{1}{12} \frac{\mu_k \ell_k}{\sin\theta_k} \left(\ell_{k-1}^2 \mathcal{D}_{2,k-1}^1(\boldsymbol{p}_1) - \ell_{k+1}^2 \mathcal{D}_{2,k+1}^1(\boldsymbol{p}_1) \right) \partial_{\boldsymbol{t}_k} w_1.\end{aligned} \tag{2.21}$$

In the last equation, we use

$$\frac{\ell_{k+1}}{\sin\theta_{k+1}} = \frac{\ell_{k-1}}{\sin\theta_{k-1}} = \frac{\ell_k}{\sin\theta_k}.$$

Using (2.21), (2.18), and following identities

$$\begin{aligned}&\ell_{k-1} \sin\theta_{k+1} = \ell_{k+1} \sin\theta_{k-1} = d_k, \\ &\ell_{k-1}^2 \cos^2\theta_{k+1} - \ell_{k+1}^2 \cos^2\theta_{k-1} = \ell_{k-1}^2 - \ell_{k+1}^2, \\ &\ell_{k+1} \cos\theta_{k-1} + \ell_{k-1} \cos\theta_{k+1} = \ell_k,\end{aligned}$$

we obtain (2.19). □

By replacing $\boldsymbol{p}$ with its linear interpolant and vice versa, we obtain the local variational error expansion for general vector-valued function $\boldsymbol{p}$.

Theorem 2.2.3. *For $w_1 \in \mathcal{P}_1(T)$,*

$$\int_T (\boldsymbol{p} - \Pi_h^0 \boldsymbol{p}) \cdot \bar{\boldsymbol{a}} \nabla^\perp w_1 = \sum_{k=1}^{3} \int_{e_k} \mathcal{A}_k(\boldsymbol{p}) \partial_{\boldsymbol{t}_k} w_1 + O(h_T^2) |\boldsymbol{p}|_{2,T} \|\nabla^\perp w_1\|_{0,T}.$$

Proof. By Lemma 2.2.2, we have

$$\begin{aligned}
&\int_T (\boldsymbol{p} - \Pi_h^0 \boldsymbol{p}) \cdot \bar{\boldsymbol{a}} \nabla^\perp w_1 \\
&= \int_T (\boldsymbol{p} - I_h^1 \boldsymbol{p}) \cdot \bar{\boldsymbol{a}} \nabla^\perp w_1 + \int_T (I_h^1 \boldsymbol{p} - \Pi_h^0 I_h^1 \boldsymbol{p}) \cdot \bar{\boldsymbol{a}} \nabla^\perp w_1 \\
&\quad + \int_T (\Pi_h^0 I_h^1 \boldsymbol{p} - \Pi_h^0 \boldsymbol{p}) \cdot \bar{\boldsymbol{a}} \nabla^\perp w_1 \\
&= \int_T (\boldsymbol{p} - I_h^1 \boldsymbol{p}) \cdot \bar{\boldsymbol{a}} \nabla^\perp w_1 + \int_T (\Pi_h^0 I_h^1 \boldsymbol{p} - \Pi_h^0 \boldsymbol{p}) \cdot \bar{\boldsymbol{a}} \nabla^\perp w_1 \\
&\quad + \sum_{k=1}^{3} \int_{e_k} \mathcal{A}_k(I_h^1 \boldsymbol{p} - \boldsymbol{p}) \partial_{\boldsymbol{t}_k} w_1 + \sum_{k=1}^{3} \int_{e_k} \mathcal{A}_k(\boldsymbol{p}) \partial_{\boldsymbol{t}_k} w_1 \\
&= I + II + III + IV.
\end{aligned} \tag{2.22}$$

(1.8) theory gives the upper bound

$$\begin{aligned}
|I| + |II| &\lesssim \|(\mathrm{id} - \Pi_h^0)(\boldsymbol{p} - I_h^1 \boldsymbol{p})\|_{0,T} \|\nabla^\perp w_1\|_{0,T} \\
&\lesssim h_T \|(\mathrm{id} - \Pi_h^0)(\boldsymbol{p} - I_h^1 \boldsymbol{p})\|_{\infty,T} \|\nabla^\perp w_1\|_{0,T} \\
&\lesssim h_T \|\boldsymbol{p} - I_h^1 \boldsymbol{p}\|_{\infty,T} \|\nabla^\perp w_1\|_{0,T} \\
&\lesssim h_T^2 |\boldsymbol{p}|_{2,T} \|\nabla^\perp w_1\|_{0,T}.
\end{aligned} \tag{2.23}$$

Using the trace inequality (1.9) and inverse inequality (1.10),

$$
\begin{aligned}
|III| &\lesssim \sum_{k=1}^{3} \|\mathcal{A}_k(I_h^1\boldsymbol{p}-\boldsymbol{p})\|_{0,e_k}\|\partial_{\boldsymbol{t}_k} w_1\|_{0,e_k} \\
&\lesssim \sum_{k=1}^{3} (h_T^{-\frac{1}{2}}\|\mathcal{A}_k(I_h^1\boldsymbol{p}-\boldsymbol{p})\|_{0,T} + h_T^{\frac{1}{2}}|\mathcal{A}_k(I_h^1\boldsymbol{p}-\boldsymbol{p})|_{1,T}) \\
&\qquad \times h_T^{-\frac{1}{2}}\|\nabla^{\perp} w_1\|_{0,T} \\
&\lesssim \sum_{k=1}^{3} (h_T^{-1}|h_T^2(I_h^1\boldsymbol{p}-\boldsymbol{p})|_{1,T} + |h_T^2(I_h^1\boldsymbol{p}-\boldsymbol{p})|_{2,T})\|\nabla^{\perp} w_1\|_{0,T} \\
&\lesssim h_K^2|\boldsymbol{p}|_{2,T}\|\nabla^{\perp} w_1\|_{0,T}.
\end{aligned}
\tag{2.24}
$$

Combining (2.22)–(2.24) with Lemma 2.2.2 completes the proof. □

We finish deriving local variational error expansion for the RT_0 element. In general, the principle of the local expansion of RT_1 is similar to RT_0. Unfortunately, it is much more technical and the final expression is not as elegant as the RT_0 case. For this reason, we assume $\boldsymbol{a}$ in (2.2) is a scalar-valued function when dealing with RT_1 finite element. We present Lemma 2.2.5 and leave the proof in Section 2.5.

Let d be the diameter of the circumscribed circle of T. For each edge e_k, there are several associated geometric quantities

$$
\begin{aligned}
&\mu_{11,k}^1 = \frac{1}{5760}\big(3\ell_k^4 - 3(\ell_{k-1}^2-\ell_{k+1}^2)^2 - 4\ell_k^2(\ell_{k-1}^2+\ell_{k+1}^2)\big), \\
&\mu_{12,k}^1 = \mu_{21,k}^1 = \frac{1}{1440d}\ell_1\ell_2\ell_3(\ell_{k-1}^2-\ell_{k+1}^2), \quad \mu_{22,k}^1 = -\frac{1}{1440d^2}\ell_1^2\ell_2^2\ell_3^2, \\
&\mu_{11,k}^2 = \frac{1}{2880\ell_1\ell_2\ell_3}d(\ell_{k-1}^2-\ell_{k+1}^2)\big(4\ell_k^2 - (\ell_{k-1}^2-\ell_{k+1}^2)^2 - 3\ell_k^2(\ell_{k-1}^2+\ell_{k+1}^2)\big), \\
&\mu_{12,k}^2 = \mu_{21,k}^2 = -\mu_{11,k}^1, \quad \mu_{22,k}^2 = -\mu_{12,k}^1,
\end{aligned}
$$

and second order derivatives $\{\mathcal{D}_{i,k}^{jl}\}_{1\le i,j,l\le 2}$

$$\mathcal{D}_{1,k}^{11} = \boldsymbol{t}_k \cdot \partial_{\boldsymbol{t}_k}^2, \quad \mathcal{D}_{1,k}^{12} = \mathcal{D}_{1,k}^{21} = \boldsymbol{t}_k \cdot \partial_{\boldsymbol{t}_k \boldsymbol{n}_k}^2, \quad \mathcal{D}_{1,k}^{22} = \boldsymbol{t}_k \cdot \partial_{\boldsymbol{n}_k}^2,$$
$$\mathcal{D}_{2,k}^{11} = \boldsymbol{n}_k \cdot \partial_{\boldsymbol{t}_k}^2, \quad \mathcal{D}_{2,k}^{12} = \mathcal{D}_{2,k}^{21} = \boldsymbol{n}_k \cdot \partial_{\boldsymbol{t}_k \boldsymbol{n}_k}^2, \quad \mathcal{D}_{2,k}^{22} = \boldsymbol{n}_k \cdot \partial_{\boldsymbol{n}_k}^2.$$

Then we define the second order derivative $\mathcal{B}_k(\boldsymbol{q}) = \mu_{jl,k}^i \mathcal{D}_{i,k}^{jl}(\boldsymbol{q})$, where we use Einstein summation notation.

Lemma 2.2.4. *For $\boldsymbol{p}_2 \in \mathcal{P}_2(T)^2$ and $w_2 \in \mathcal{P}_2(T)$,*

$$\int_T (\boldsymbol{p}_2 - \Pi_h^1 \boldsymbol{p}_2) \cdot \nabla^\perp w_2 = \sum_{k=1}^3 \int_{e_k} \mathcal{B}_k(\boldsymbol{p}_2) \partial_{\boldsymbol{t}_k}^2 w_2.$$

Built upon Lemma 2.2.4, we derive the local error expansion for general $\boldsymbol{p}$.

Theorem 2.2.5. *For $w_2 \in \mathcal{P}_2(T)$,*

$$\int_T (\boldsymbol{p} - \Pi_h^1 \boldsymbol{p}) \cdot \nabla^\perp w_2 = \sum_{k=1}^3 \int_{e_k} \mathcal{B}_k(\boldsymbol{p}) \partial_{\boldsymbol{t}_k}^2 w_2 + O(h_T^3) |\boldsymbol{p}|_{3,T} \|\nabla^\perp w_2\|_{0,T}.$$

We skip the proof of Theorem 2.2.5 because it is the same as Theorem 2.2.3.

Remark. The operator $\mathcal{A}_k$ is simple but the expression of $\mathcal{B}_k$ is a bit complicated. However, for superconvergence analysis, it suffices to keep the following features in mind.

1. $\{\mathcal{A}_k\}_{k=1}^3$ are first order derivatives of magnitude h_T^2 and $\{\mathcal{B}_k\}_{k=1}^3$ are second order derivatives of magnitude h_T^4, namely,

$$\mathcal{A}_k(\boldsymbol{q}) = O(h_T^2) \sum_{i,j=1}^2 \partial_{x_i} q_j, \quad \mathcal{B}_k(\boldsymbol{q}) = O(h_T^4) \sum_{i,j,l=1}^2 \partial_{x_i} \partial_{x_j} q_l.$$

2. For $e \in \mathcal{E}_h^o$ or $\mathcal{E}_h^\partial$, we have $\omega_e = T \cup T'$ or $\omega_e = T$ respectively. Let $\boldsymbol{t}_e$ denote the

unit tangent and and $\boldsymbol{n}_e$ the unit normal to e whose directions are induced by T. Let $\bar{\boldsymbol{a}} = \fint_T \boldsymbol{a}$ and $\bar{\boldsymbol{a}}' = \fint_{T'} \boldsymbol{a}$. Let $\mathcal{A}_e, \mathcal{B}_e$ be the operators based on T and $\mathcal{A}'_e, \mathcal{B}'_e$ based on T'. If ω_e is an $O(h^{1+\alpha})$-approximate parallelogram, on the edge e, we have the cancellation

$$\bar{\boldsymbol{a}}\mathcal{A}_e(\boldsymbol{q}) - \bar{\boldsymbol{a}}'\mathcal{A}'_e(\boldsymbol{q}) = O(h_e^{2+\min(1,\alpha)}) \sum_{i,j=1}^{2} \partial_{x_i} q_j, \tag{2.25a}$$

$$\bar{\boldsymbol{a}}\mathcal{B}_e(\boldsymbol{q}) - \bar{\boldsymbol{a}}'\mathcal{B}'_e(\boldsymbol{q}) = O(h_e^{4+\min(1,\alpha)}) \sum_{i,j,l=1}^{2} \partial_{x_i}\partial_{x_j} q_l. \tag{2.25b}$$

Indeed, ω_e is a approximate parallelogram implies that $\ell_k = \ell'_k + O(h^{1+\alpha})$, $\boldsymbol{t}_k = \boldsymbol{t}'_k + O(h^\alpha)$, $\sin\theta_k = \sin\theta'_k + O(h^\alpha)$, $\mu_k = \mu'_k + O(h^\alpha)$, and $|T| = |T'| + O(h^{2+\alpha})$. Combining these estimates with $\bar{\boldsymbol{a}} = \bar{\boldsymbol{a}}' + O(h)$, (2.25a) then comes from the telescoping type inequality

$$\left| \prod_{i=1}^{n} a_i - \prod_{i=1}^{n} b_i \right| \le \sum_{i=1}^{n} |a_i - b_i| \prod_{j\neq i} \max(a_j, b_j). \tag{2.26}$$

(2.25b) can be derived in the same way.

2.3 Supercloseness

Lemma 2.3.1 (Lemma 3.7, [46]). *Assume that $\mathcal{T}_h$ satisfy the strong (α,β)-condition. For $w_h \in \mathcal{S}_h^1$,*

$$(\bar{\boldsymbol{a}}(\boldsymbol{p} - \Pi_h^0 \boldsymbol{p}), \nabla^\perp w_h) \lesssim h^{1+\min(1,\alpha,\beta/2)} |\log h|^{\frac{1}{2}} \|\boldsymbol{p}\|_{2,\infty,\Omega} \|\nabla^\perp w_h\|_{0,\Omega}. \tag{2.27}$$

Proof. By Theorem 2.2.3, the left hand side is

$$
\begin{aligned}
(\bar{\boldsymbol{a}}(\boldsymbol{p}-\Pi_h^0\boldsymbol{p}),\nabla^\perp w_h) &= \sum_{T\in\mathcal{T}_h}\sum_{k=1}^{3}\int_{e_k}\mathcal{A}_k(\boldsymbol{p})\partial_{\boldsymbol{t}_k}w_h + O(h^2)|\boldsymbol{p}|_{2,T}\|\nabla^\perp w_h\|_T \\
&= \Big(\sum_{e\in\mathcal{E}_{h,1}^o}+\sum_{e\in\mathcal{E}_{h,2}^o}+\sum_{e\in\mathcal{E}_h^\partial}\Big)\int_e\big(\mathcal{A}_e(\boldsymbol{p})-\mathcal{A}_e'(\boldsymbol{p})\big)\partial_{\boldsymbol{t}_e}w_h \qquad (2.28)\\
&+\sum_{T\in\mathcal{T}_h}O(h^2)|\boldsymbol{p}|_{2,T}\|\nabla^\perp w_h\|_T = I+II+III+O(h^2)|\boldsymbol{p}|_{2,\Omega}\|\nabla^\perp w_h\|_{0,\Omega}.
\end{aligned}
$$

Here the notations in Remark 2.2 are adopted. $\mathcal{A}_e'(\boldsymbol{p})=0$ if $e\in\mathcal{E}_h^\partial$. By the cancellation (2.25) and the trace inequality (1.9),

$$
\begin{aligned}
|I| &\lesssim \sum_{e\in\mathcal{E}_{h,1}^o}h_e^{2+\alpha}\int_e|D\boldsymbol{p}|\cdot|\nabla^\perp w_h| \\
&\lesssim \sum_{e\in\mathcal{E}_{h,1}^o}h_e^{2+\alpha}\Big(\int_T h_T^{-1}|D^2\boldsymbol{p}|+|D\boldsymbol{p}|\Big)|\nabla^\perp w_h| \\
&\lesssim \sum_{e\in\mathcal{E}_{h,1}^o}h_e^{1+\alpha}\|D\boldsymbol{p}\|_{1,T}\|\nabla^\perp w_h\|_{0,T} \\
&\lesssim h^{1+\alpha}\|D\boldsymbol{p}\|_{1,\Omega}\|\nabla^\perp w_h\|_{0,\Omega}.
\end{aligned} \qquad (2.29)
$$

For $e\in\mathcal{E}_{h,2}^o$, there is no cancellation. Let $\Omega_\beta=\cup_{e\in\mathcal{E}_{h,2}^o}\omega_e$. The sum over $\mathcal{E}_{h,2}^o$ is simply estimated by

$$
\begin{aligned}
|II| &\lesssim \sum_{e\in\mathcal{E}_{h,2}^o}h_e^3\|D\boldsymbol{p}\|_{0,\infty,T}\|\nabla^\perp w_h\|_{0,\infty,T} \\
&\lesssim h\|D\boldsymbol{p}\|_{0,\infty,\Omega}\sum_{e\in\mathcal{E}_{h,2}^o}\int_T|\nabla^\perp w_h| \qquad (2.30)\\
&= h|\boldsymbol{p}|_{1,\infty,\Omega}\int_{\Omega_\beta}|\nabla^\perp w_h|\lesssim h^{1+\frac{\beta}{2}}|\boldsymbol{p}|_{1,\infty,\Omega}\|\nabla^\perp w_h\|_{0,\Omega}.
\end{aligned}
$$

The trickiest part of this proof is to bound III. We can assume that $\int_\Omega w_h=0$ by

subtracting a constant from w_h. Then Poincaré inequality gives

$$\|w_h\|_{1,\Omega} \lesssim \|\nabla^\perp w_h\|_{0,\Omega}. \tag{2.31}$$

Define $\bar{\mathcal{A}}_e(\boldsymbol{p}) = \fint_e \mathcal{A}_e(\boldsymbol{p})$. III can be split as

$$\begin{aligned} III &= \sum_{e\in\mathcal{E}_h^\partial} \int_e \big(\mathcal{A}_e(\boldsymbol{p}) - \bar{\mathcal{A}}_e(\boldsymbol{p})\big)\partial_{\boldsymbol{t}_e} w_h + \sum_{e\in\mathcal{E}_h^\partial} \int_e \bar{\mathcal{A}}_e(\boldsymbol{p})\partial_{\boldsymbol{t}_e} w_h \\ &= III_1 + III_2. \end{aligned}$$

The first term can be estimated by (2.31):

$$\begin{aligned} |III_1| &\lesssim \sum_{e\in\mathcal{E}_h^\partial} h_e^3 |\boldsymbol{p}|_{2,\infty,e} \int_e |\nabla^\perp w_h| \\ &\lesssim h^2 |\boldsymbol{p}|_{2,\infty,\Omega} \sum_{e\in\mathcal{E}_h^\partial} \int_T |\nabla^\perp w_h| \\ &\lesssim h^2 |\boldsymbol{p}|_{2,\infty,\Omega} \|\nabla^\perp w_h\|_{0,\Omega}. \end{aligned} \tag{2.32}$$

For $\boldsymbol{z} \in \mathcal{N}_h^\partial$, let e, e' be the two edges on $\partial\Omega$ sharing z as an ending point. Then the second term becomes

$$III_2 = \sum_{z\in\mathcal{N}_h^\partial} \big(\bar{\mathcal{A}}_e(\boldsymbol{p}) - \bar{\mathcal{A}}_{e'}(\boldsymbol{p})\big)\, w_h(\boldsymbol{z}).$$

For $\boldsymbol{z} \in \mathcal{N}_{h,1}^\partial$, Definition 1.4.4 near the boundary implies cancellation and thus

$$\big|\bar{\mathcal{A}}_e(\boldsymbol{p}) - \bar{\mathcal{A}}_{e'}(\boldsymbol{p})\big| \lesssim h^{2+\alpha} \|D\boldsymbol{p}\|_{1,\infty,\Omega}. \tag{2.33}$$

For $\boldsymbol{z} \in \mathcal{N}_{h,2}^{\partial}$,

$$\left|\bar{\mathcal{A}}_e(\boldsymbol{p}) - \bar{\mathcal{A}}_{e'}(\boldsymbol{p})\right| \le \left|\bar{\mathcal{A}}_e(\boldsymbol{p})\right| + \left|\bar{\mathcal{A}}_{e'}(\boldsymbol{p})\right| \lesssim h^2 \|D\boldsymbol{p}\|_{0,\infty,\Omega}. \tag{2.34}$$

It follows from (1.11), $\#\mathcal{N}_{h,1}^{\partial} \lesssim h^{-1}$, $\#\mathcal{N}_{h,2}^{\partial} \lesssim 1$, (2.31), (2.33), and (2.34) that

$$\begin{aligned} |III_2| &\lesssim \left(\sum_{\boldsymbol{z} \in \mathcal{N}_{h,1}^{\partial}} h^{2+\alpha} \|D\boldsymbol{p}\|_{1,\infty,\Omega} + \sum_{\boldsymbol{z} \in \mathcal{N}_{h,2}^{\partial}} h^2 \|D\boldsymbol{p}\|_{0,\infty,\Omega} \right) \|w_h\|_{0,\infty,\partial\Omega} \\ &\lesssim h^{1+\alpha} |\log h|^{\frac{1}{2}} |\boldsymbol{p}|_{2,\infty,\Omega} \|\nabla^{\perp} w_h\|_{1,\Omega}. \end{aligned} \tag{2.35}$$

Now, combining (2.28), (2.29), (2.30), (2.32), and (2.35), we obtain (2.3.1). □

The next lemma is about the variational error expansion of RT_1 elements. The proof is the same Lemma 2.3.1.

Lemma 2.3.2. *Let $\mathcal{T}_h$ satisfy the (α, β)-condition. For $w_h \in \mathcal{S}_h^2$,*

$$(\bar{\boldsymbol{a}}(\boldsymbol{p} - \Pi_h^1 \boldsymbol{p}), \nabla^{\perp} w_h) \lesssim h^{2+\min(1/2,\alpha,\beta/2)} \big(\|\boldsymbol{p}\|_{3,\Omega} + |\boldsymbol{p}|_{2,\infty,\Omega} \big) \|\nabla^{\perp} w_h\|_{0,\Omega}.$$

Proof. By Theorem 2.2.5, the left hand side is

$$\begin{aligned} (\bar{\boldsymbol{a}}(\boldsymbol{p} - \Pi_h^1 \boldsymbol{p}), \nabla^{\perp} r_h) &= \sum_{T \in \mathcal{T}_h} \sum_{k=1}^{3} \int_{e_k} \bar{\boldsymbol{a}} \mathcal{B}_k(\boldsymbol{p}) \partial_{\boldsymbol{t}_k}^2 r_h + O(h_T^3) |\boldsymbol{p}|_{3,T} \|\nabla^{\perp} w_h\|_{0,T} \\ &= \Big(\sum_{e \in \mathcal{E}_{h,1}^o} + \sum_{e \in \mathcal{E}_{h,2}^o} + \sum_{e \in \mathcal{E}_h^{\partial}} \Big) \int_e \big(\bar{\boldsymbol{a}} \mathcal{B}_e(\boldsymbol{p}) - \bar{\boldsymbol{a}}' \mathcal{B}_e'(\boldsymbol{p}) \big) \partial_{\boldsymbol{t}_e}^2 w_h \\ &\quad + \sum_{T \in \mathcal{T}_h} O(h_T^3) |\boldsymbol{p}|_{3,T} \|\nabla^{\perp} w_h\|_{0,T} = I + II + III + O(h^3) |\boldsymbol{p}|_{3,\Omega} \|\nabla^{\perp} w_h\|_{0,\Omega}. \end{aligned} \tag{2.36}$$

It then follows from the inverse estimate $|r_h|_{2,T} \lesssim h_T^{-1} \|\nabla^{\perp} r_h\|_{0,T}$, the proof of (2.29)

and (2.30) that

$$\begin{aligned} |I| &\lesssim h^{2+\alpha}\|D^2\boldsymbol{p}\|_{1,\Omega}\|\nabla^{\perp}r_h\|_{0,\Omega}, \\ |II| &\lesssim h^{2+\frac{\beta}{2}}|\boldsymbol{p}|_{2,\infty,\Omega}\|\nabla^{\perp}r_h\|_{0,\Omega}. \end{aligned} \tag{2.37}$$

Let $\Omega_{\partial} = \cup_{e\in\mathcal{E}_h^{\partial}}\omega_e$. The area of Ω_{∂} is $|\Omega_{\partial}| = O(h)$. Then III is estimated in the same way as (2.30).

$$\begin{aligned} |III| &\lesssim h^3|\boldsymbol{p}|_{2,\infty,\partial\Omega}\sum_{e\in\mathcal{E}_h^{\partial}}\int_{\omega_e}|D^2r_h| \\ &\lesssim h^2|\boldsymbol{p}|_{2,\infty,\Omega}\int_{\Omega_{\partial}}|\nabla^{\perp}r_h| \lesssim h^{2+\frac{1}{2}}|\boldsymbol{p}|_{2,\infty,\Omega}\|\nabla^{\perp}r_h\|_{0,\Omega}. \end{aligned} \tag{2.38}$$

Combining (2.36), (2.37), and (2.38), we prove the theorem. □

Built upon Lemmas 2.3.1 and 2.3.2, we are able to prove supercloseness results, namely, to show that $\|\Pi_h^r\boldsymbol{p} - \boldsymbol{p}_h^r\|_{0,\Omega}$ is supersmall. From (2.3) and (2.5), we have the error equation

$$(\boldsymbol{a}(\boldsymbol{p}-\boldsymbol{p}_h^r),\boldsymbol{q}_h) - (\boldsymbol{q}_h,\boldsymbol{b}(u-u_h^r)) + (\operatorname{div}\boldsymbol{q}_h, u-u_h^r) = 0, \quad \boldsymbol{q}_h\in\mathcal{Q}_h^r, \tag{2.39a}$$

$$-(\operatorname{div}(\boldsymbol{p}-\boldsymbol{p}_h^r), v_h) + (c(u-u_h^r), v_h) = 0, \quad v_h\in\mathcal{V}_h^r. \tag{2.39b}$$

From now on, we introduce several error quantities

$$\boldsymbol{\xi}_h = \Pi_h^r\boldsymbol{p} - \boldsymbol{p}_h^r, \quad \tau_h = P_h^r u - u_h^r, \quad \boldsymbol{\eta} = \boldsymbol{p} - \boldsymbol{p}_h^r.$$

Douglas and Roberts [27] have shown the standard a priori error estimates:

$$\begin{aligned} &\|\boldsymbol{p}-\boldsymbol{p}_h^r\|_{0,\Omega} \lesssim h^{r+1}\|u\|_{r+2,\Omega}, \\ &\|\operatorname{div}(\boldsymbol{p}-\boldsymbol{p}_h^r)\|_{0,\Omega} \lesssim h^{r+1}\|u\|_{r+3,\Omega}, \\ &\|u-u_h^r\|_{0,\Omega} \lesssim h^{r+1}\|u\|_{r+1+\delta_{r0},\Omega}. \end{aligned} \tag{2.40}$$

In addition, [27] gives the well-known supercloseness result for the scalar unknown u

$$\|\tau_h\|_{0,\Omega} \lesssim h^{r+2}\|u\|_{r+2+\delta_{r0},\Omega}, \tag{2.41}$$

(2.41) holds on shape regular unstructured meshes. Built upon (2.41), we are able to prove that $\|\operatorname{div}(\Pi_h^r \boldsymbol{p} - \boldsymbol{p}_h^r)\|_{0,\Omega}$ is supersmall.

Theorem 2.3.3. *For general shape regular $\mathcal{T}_h$ and $r \geq 0$,*

$$\|\operatorname{div}(\Pi_h^r \boldsymbol{p} - \boldsymbol{p}_h^r)\|_{0,\Omega} \lesssim h^{r+2}\|u\|_{2+r+\delta_{r0},\Omega}.$$

Proof. Let

$$v_h = \frac{\operatorname{div} \boldsymbol{\xi}_h}{\|\operatorname{div} \boldsymbol{\xi}_h\|_{0,\Omega}} \in \mathcal{V}_h^r.$$

By (2.12) and (2.39), we have

$$\begin{aligned}\|\operatorname{div} \boldsymbol{\xi}_h\|_{0,\Omega} &= (\operatorname{div} \boldsymbol{\xi}_h, v_h) = (P_h^r \operatorname{div} \boldsymbol{p} - \operatorname{div} \boldsymbol{p}_h^r, v_h) \\ &= (\operatorname{div}(\boldsymbol{p} - \boldsymbol{p}_h^r), v_h) = (u - P_h^r u, cv_h) + (P_h^r u - u_h^r, cv_h).\end{aligned}$$

It then follows from (2.13), (2.40) and (2.41) that

$$\begin{aligned}\|\operatorname{div} \boldsymbol{\xi}_h\|_{0,\Omega} &= (u - P_h^r u, cv_h - P_h^r(cv_h)) + O(h^{r+2})\|u\|_{2+r+\delta_{r0},\Omega}\|v_h\|_{0,\Omega} \\ &= O(h^{2r+2})\|u\|_{r+1,\Omega}\|cv_h\|_{r+1,\Omega} + O(h^{r+2})\|u\|_{2+r+\delta_{r0},\Omega}\|v_h\|_{0,\Omega} \\ &= O(h^{r+2})\|u\|_{2+r+\delta_{r0},\Omega}\|v_h\|_{0,\Omega}.\end{aligned}$$

The proof is complete. □

Before proving the superconvergence result for $\|\Pi_h^r \boldsymbol{p} - \boldsymbol{p}_h^r\|_{0,\Omega}$, it is necessary to

discuss the L^2 de Rham complex in $\mathbb{R}^2$:

$$L^2(\Omega) \xrightarrow{\nabla^\perp} L^2(\Omega)^2 \xrightarrow{\operatorname{div}} L^2(\Omega) \to 0. \tag{2.42}$$

Here $L^2(\Omega) = \mathcal{V}$ is equipped with the standard inner product $(\cdot,\cdot)$. Since we are dealing with variable coefficients, $L^2(\Omega)^2$ is equipped with the weighted L^2 inner product $(\cdot,\cdot)_a$:

$$(\boldsymbol{q}_1, \boldsymbol{q}_2)_a := (\boldsymbol{a}\boldsymbol{q}_1, \boldsymbol{q}_2), \quad \boldsymbol{q}_1, \boldsymbol{q}_2 \in L^2(\Omega)^2.$$

The weighted L^2-norm is $\|\boldsymbol{q}\|_a = (\boldsymbol{a}\boldsymbol{q}, \boldsymbol{q})^{\frac{1}{2}}$. Since $\boldsymbol{a}$ is uniformly elliptic, $\|\boldsymbol{q}\|_{0,\Omega} \approx \|\boldsymbol{q}\|_a$ for all $\boldsymbol{q} \in L^2(\Omega)^2$. By the theory finite element exterior calculus (cf. [5, 6]), the discrete complex

$$\mathcal{S}_h^{r+1} \xrightarrow{\nabla^\perp} \mathcal{Q}_h^r \xrightarrow{\operatorname{div}} \mathcal{V}_h^r \to 0$$

inherits the aforementioned nice properties of (2.42). In particular, we have the discrete Helmholtz decomposition

$$\mathcal{Q}_h^r = \nabla^\perp \mathcal{S}_h^{r+1} \oplus \operatorname{grad}_h \mathcal{V}_h^r, \tag{2.43}$$

where $\operatorname{grad}_h : \mathcal{V}_h^r \to \mathcal{Q}_h^r$ is the adjoint of $-\operatorname{div} : \mathcal{Q}_h^r \to \mathcal{V}_h^r$ w.r.t. the weighted inner product $(\cdot,\cdot)_a$, namely, $(\boldsymbol{a}\operatorname{grad}_h v_h, \boldsymbol{q}_h) = -(v_h, \operatorname{div}\boldsymbol{q}_h)$ for all $\boldsymbol{q}_h \in \mathcal{Q}_h^r$.

The last ingredient for our supercloseness analysis is a discrete Poincaré inequality.

Lemma 2.3.4. *For $v \in \mathcal{V}_h^r$,*

$$\|v_h\|_{0,\Omega} \lesssim \|\operatorname{grad}_h v_h\|_a.$$

Proof. $\operatorname{div} : \mathcal{Q}_h^r \to \mathcal{V}_h^r$ is surjective and there exists $\boldsymbol{q}_h \in \mathcal{Q}_h^r$, such that $\operatorname{div}\boldsymbol{q}_h = v_h$. In

addition, Raviart and Thomas [59] have shown that

$$\|\boldsymbol{q}_h\|_a \approx \|\boldsymbol{q}_h\|_{0,\Omega} \lesssim \|v_h\|_{0,\Omega}.$$

It then follows

$$\|v_h\|_{0,\Omega}^2 = -(\boldsymbol{a}\,\mathrm{grad}_h\, v_h, \boldsymbol{q}_h) \lesssim \|\,\mathrm{grad}_h\, v_h\|_a \|v_h\|_{0,\Omega},$$

which completes the proof. □

With the above preparations, we are able to prove supercloseness estimates for the RT_0 and RT_1 mixed methods.

Theorem 2.3.5. *Assume that $\mathcal{T}_h$ satisfies the (α,β)-condition. Then*

$$\|\Pi_h^r \boldsymbol{p} - \boldsymbol{p}_h^r\|_a \lesssim h^{r+1+\min(\frac{1}{2},\alpha,\frac{\beta}{2})}\big(|\boldsymbol{p}|_{r+1,\infty,\Omega} + \|\boldsymbol{p}\|_{r+2,\Omega}\big), \quad r = 0, 1.$$

Assume that $\mathcal{T}_h$ satisfies the piecewise strong (α,β)-condition. Then for the RT_0 finite element,

$$\|\Pi_h^0 \boldsymbol{p} - \boldsymbol{p}_h^0\|_{0,\Omega} \lesssim h^{1+\min(1,\alpha,\frac{\beta}{2})} |\log h|^{\frac{1}{2}} \|u\|_{3,\infty,\Omega}.$$

Proof. Consider the discrete Helmholtz decomposition

$$\boldsymbol{\xi}_h = \nabla^\perp w_h \oplus \mathrm{grad}_h\, v_h, \tag{2.44}$$

with $\{v_h, w_h\} \in \mathcal{V}_h^r \times \mathcal{S}_h^{r+1}$. Let $\tilde{\boldsymbol{q}}_h = \mathrm{grad}_h\, v_h / \|\,\mathrm{grad}_h\, v_h\|_a$. By Lemma 2.3.4 and

Lemma 2.3.3,

$$\begin{aligned} &\|\operatorname{grad}_h v_h\|_a = (\operatorname{grad}_h v_h, \tilde{\boldsymbol{q}}_h)_a = -(v_h, \operatorname{div} \tilde{\boldsymbol{q}}_h) \\ &= -\big(v_h, \frac{\operatorname{div} \boldsymbol{\xi}_h}{\|\operatorname{grad}_h v_h\|_a}\big) \lesssim \|\operatorname{div} \boldsymbol{\xi}_h\|_{0,\Omega} \lesssim h^{r+2}\|u\|_{r+2+\delta_{r0}}. \end{aligned} \tag{2.45}$$

It remains to bound $\nabla^\perp w_h$. Let $\boldsymbol{q}_h = \nabla^\perp w_h / \|\nabla^\perp w_h\|_a$. The orthogonality implies

$$\begin{aligned} \|\nabla^\perp w_h\|_a &= (\boldsymbol{a}\boldsymbol{\xi}_h, \boldsymbol{q}_h) \\ &= -(\boldsymbol{a}(\boldsymbol{p} - \Pi_h^r \boldsymbol{p}), \boldsymbol{q}_h) + (\boldsymbol{a}(\boldsymbol{p} - \boldsymbol{p}_h^r), \boldsymbol{q}_h) \\ &= I + II. \end{aligned} \tag{2.46}$$

I is split as

$$I = ((\bar{\boldsymbol{a}} - \boldsymbol{a})(\boldsymbol{p} - \Pi_h^r \boldsymbol{p}), \boldsymbol{q}_h) - (\bar{\boldsymbol{a}}(\boldsymbol{p} - \Pi_h^r \boldsymbol{p}), \boldsymbol{q}_h) = I_1 + I_2.$$

By Lemma 2.13 and $\|\bar{\boldsymbol{a}} - \boldsymbol{a}\|_{0,\infty,\Omega} = O(h)$,

$$I_1 = O(h^{r+2})|\boldsymbol{p}|_{r+1}. \tag{2.47}$$

I_2 can be estimated by Lemmas 2.3.1 or 2.3.2.

$$|I_2| \lesssim h^{1+\min(1,\alpha,\frac{\beta}{2})} |\log h|^{\frac{1}{2}} \|u\|_{3,\infty,\Omega} \text{ or } h^{r+1+\min(\frac{1}{2},\alpha,\frac{\beta}{2})} \big(|\boldsymbol{p}|_{r+1,\infty,\Omega} + \|\boldsymbol{p}\|_{r+2,\Omega}\big). \tag{2.48}$$

Since $\operatorname{div} \boldsymbol{q}_h = 0$, (2.39)–(2.41) imply

$$\begin{aligned} II &= (\boldsymbol{q}_h, \boldsymbol{b}(u - u_h)) \\ &= (\boldsymbol{b} \cdot \boldsymbol{q}_h, u - P_h^r u + P_h^r u - u_h^r) \\ &= (\boldsymbol{b} \cdot \boldsymbol{q}_h - P_h^r(\boldsymbol{b} \cdot \boldsymbol{q}_h), u - P_h^r u) + O(h^{r+2})\|u\|_{r+2+\delta_{r0},\Omega}\|\boldsymbol{q}_h\|_{0,\Omega} \\ &= O(h^{2r+2})|\boldsymbol{b} \cdot \boldsymbol{q}_h|_{h,r+1,\Omega}|u|_{r+1,\Omega} + O(h^{r+2})\|u\|_{r+2+\delta_{r0},\Omega}\|\boldsymbol{q}_h\|_{0,\Omega}. \end{aligned} \tag{2.49}$$

Since $\boldsymbol{q}_h|_T \in \mathcal{P}_r(T)^2$, inverse estimate (1.10) implies

$$|\boldsymbol{b} \cdot \boldsymbol{q}_h|_{r+1,T} \lesssim \sum_{l=0}^{r} \|D^l \boldsymbol{q}_h\|_{0,T} \lesssim h_T^{-r}\|\boldsymbol{q}_h\|_{0,T}.$$

(2.49) then reduces to

$$II = O(h^{r+2})\|u\|_{r+2+\delta_{r0},\Omega}\|\boldsymbol{q}_h\|_{0,\Omega}. \tag{2.50}$$

Then the theorem follows from (2.45)–(2.48), and (2.50). □

2.4 Superconvergent recovery

In this section, we consider RT_r finite elements with general $r \geq 0$. For $e \in \mathcal{E}_h$, let $\{(w_j, \boldsymbol{g}_j)\}_{j=1}^{r+1}$ denote the Gaussian quadrature on e that are exact for $\mathcal{P}_{2r+1}(e)$, where $\{\boldsymbol{g}_j\}_{j=1}^{r+1}$ are quadrature points and $\{w_j\}_{j=1}^{r+1}$ the corresponding weights. Let $v_j \in \mathcal{P}_r(e)$ the polynomial that is w_j^{-1} at $\boldsymbol{g}_j$ and 0 at the rest of quadrature points. For $T \in \mathcal{T}_h$, let $\{\lambda_l\}_{l=1}^{r(r+1)/2}$ be the nodal basis function of Lagrange elements of degree $r-1$ on T ($\{\lambda_l\} = \emptyset$ if $r = 0$, $\{\lambda_l\} = \{1\}$ if $r = 1$). We can specify degrees of freedom of RT_r elements as

$$\mathcal{N}_e^j(\boldsymbol{q}) := \fint_e \boldsymbol{q} \cdot \boldsymbol{n}_e v_j, \quad \mathcal{N}_T^{lm}(\boldsymbol{q}) := \fint_T q_m \lambda_l,$$

where $e \in \mathcal{E}_h$, $T \in \mathcal{T}_h$, and $\boldsymbol{n}_e$ is a unit normal to e,

Now we introduce a new recovery operator $R_h^r : \mathcal{Q}_h^r \to \mathcal{S}_h^{r+1} \times \mathcal{S}_h^{r+1}$. For $\boldsymbol{q}_h \in \mathcal{Q}_h^r$, it suffices to specify nodal values of $R_h^r \boldsymbol{q}_h$. Here a node is the location of the degree of freedom of Lagrange elements, which can be a vertex of a triangle or an interior point of an edge/ triangle. For vertices $\boldsymbol{z}_1, \boldsymbol{z}_2, \boldsymbol{z}_3 \in \mathcal{N}_h$, let $\overline{\boldsymbol{z}_1\boldsymbol{z}_2}$ denote the edge with endpoints $\boldsymbol{z}_1, \boldsymbol{z}_2$ and $\overline{\boldsymbol{z}_1\boldsymbol{z}_2\boldsymbol{z}_3}$ the triangle with vertices $\boldsymbol{z}_1, \boldsymbol{z}_2, \boldsymbol{z}_3$. R_h^r is defined in three steps.

Step 1. For each vertex $\boldsymbol{z} \in \mathcal{N}_h$, let $R_h^r \boldsymbol{q}_h(\boldsymbol{z}) := \boldsymbol{q}_z(\boldsymbol{z})$, where $\boldsymbol{q}_z \in \mathcal{P}_{r+1}(\omega_z)^2$ minimizes the quadratic functional

$$\begin{aligned}\mathcal{F}(\boldsymbol{q}) = &\sum_{e \in \mathcal{E}_h(\omega_z)} \sum_{j=1}^{r+1} \left(\mathcal{N}_e^j(\boldsymbol{q}) - \mathcal{N}_e^j(\boldsymbol{q}_h)\right)^2 \\ &+ \sum_{T \in \mathcal{T}_h(\omega_z)} \sum_{l=1}^{r(r+1)/2} \sum_{m=1}^{2} \left(\mathcal{N}_T^{lm}(\boldsymbol{q}) - \mathcal{N}_T^{lm}(\boldsymbol{q}_h)\right)^2\end{aligned}$$

subject to $\boldsymbol{q} \in \mathcal{P}_{r+1}(\omega_z)^2$.

Step 2. For each node $\boldsymbol{z}$ in the interior of an edge $e = \overline{\boldsymbol{z}_1\boldsymbol{z}_2} \in \mathcal{E}_h$, let

$$R_h^r \boldsymbol{q}_h(\boldsymbol{z}) = (1-\alpha)\boldsymbol{q}_{z_1}(\boldsymbol{z}) + \alpha \boldsymbol{q}_{z_2}(\boldsymbol{z}), \quad \alpha = |\boldsymbol{z} - \boldsymbol{z}_1|/|e|.$$

Step 3. For each node $\boldsymbol{z}$ in the interior of the triangle $T = \overline{\boldsymbol{z}_1\boldsymbol{z}_2\boldsymbol{z}_3} \in \mathcal{T}_h$, define

$$R_h^r \boldsymbol{q}_h(\boldsymbol{z}) = \alpha_1 \boldsymbol{q}_{z_1}(\boldsymbol{z}) + \alpha_2 \boldsymbol{q}_{z_2}(\boldsymbol{z}) + \alpha_3 \boldsymbol{q}_{z_3}(\boldsymbol{z}),$$

where $\alpha_1, \alpha_2, \alpha_3$ are barycentric coordinates of $\boldsymbol{z}$ w.r.t. $\boldsymbol{z}_1, \boldsymbol{z}_2$, and $\boldsymbol{z}_3$.

In some cases, ω_z needs be enlarged to ensure that the above LS problem has a unique solution. Since R_h^r depends only on the degrees of freedom of the RT_r element, $R_h^r \boldsymbol{q}$ is well-defined for all $\boldsymbol{q} \in \mathcal{Q}$ and $R_h^r \Pi_h^r \boldsymbol{q} = R_h^r \boldsymbol{q}$. In addtion, $\mathcal{N}_e^j(\boldsymbol{q}) = \boldsymbol{q}(\boldsymbol{g}_j) \cdot \boldsymbol{n}_e$

for $\boldsymbol{q} \in \mathcal{P}_{r+1}(T)^2$ and $e \in \mathcal{E}_h(T)$.

To clarify the recovery procedure, we give details to the two important cases: RT_0 and RT_1 elements.

Example 1. *RT_0 elements on triangular meshes.* In this case, $R_h^0 \boldsymbol{q}_h$ is a continuous piecewise linear function. At step 1, let $\{e_j\}_{j=1}^J = \mathcal{E}_h(\omega_z)$. Let $\boldsymbol{m}_j = (m_{j1}, m_{j2})^\intercal$ be the midpoint of e_j and $\boldsymbol{n}_j = (n_{j1}, n_{j2})^\intercal$ be a unit normal to e_j. Then $\boldsymbol{q}_z = (c_1 + c_2 x_1 + c_3 x_2, c_4 + c_5 x_1 + c_6 x_2)^\intercal \in \mathcal{P}_1(\omega_z)^2$ is the minimizer of

$$\mathcal{F}(\boldsymbol{q}) = \sum_{j=1}^{J} (\boldsymbol{q}(\boldsymbol{m}_j) \cdot \boldsymbol{n}_j - \boldsymbol{q}_h(\boldsymbol{m}_j) \cdot \boldsymbol{n}_j)^2,$$

$$\text{subject to } \boldsymbol{q} \in \mathcal{P}_1(\omega_z)^2.$$

Equivalently, $\boldsymbol{c}_z = (c_1, \ldots, c_6)^\intercal$ satisfies the normal equation $\boldsymbol{A}_z^T \boldsymbol{A}_z \boldsymbol{c}_z = \boldsymbol{A}_z^T \boldsymbol{d}_z$, where $\boldsymbol{d}_z = (\boldsymbol{q}_h(\boldsymbol{m}_1) \cdot \boldsymbol{n}_1, \ldots, \boldsymbol{q}_h(\boldsymbol{m}_J) \cdot \boldsymbol{n}_J)^\intercal$, $\boldsymbol{A}_z = (\boldsymbol{a}_1^\intercal, \ldots, \boldsymbol{a}_J^\intercal)^\intercal$ is an $J \times 6$ matrix, $\boldsymbol{a}_j = (n_{j1}, m_{j1} n_{j1}, m_{j2} n_{j1}, n_{j2}, m_{j1} n_{j2}, m_{j2} n_{j2})$. Then $R_h \boldsymbol{q}_h(\boldsymbol{z}) = \boldsymbol{q}_z(\boldsymbol{z})$ for $\boldsymbol{z} \in \mathcal{N}_h$.

To avoid ill-conditioned $\boldsymbol{A}_z$ on graded meshes, we calculate $\boldsymbol{q}_z$ by scaling it properly. Let $h_z = |\omega_z|^{\frac{1}{2}}$ and $\hat{\boldsymbol{q}}_z(\hat{\boldsymbol{x}}) = \boldsymbol{q}_z(\boldsymbol{z} + h_z \hat{\boldsymbol{x}}) = (\hat{c}_1 + \hat{c}_2 \hat{x}_1 + \hat{c}_3 \hat{x}_2, \hat{c}_4 + \hat{c}_5 \hat{x}_1 + \hat{c}_6 \hat{x}_2)^\intercal$. $\hat{\boldsymbol{c}}_z = (\hat{c}_1, \ldots, \hat{c}_6)^\intercal$ solves $\hat{\boldsymbol{A}}_z^\intercal \hat{\boldsymbol{A}}_z \hat{\boldsymbol{c}}_z = \hat{\boldsymbol{A}}_z^\intercal \boldsymbol{d}_z$, where $\hat{\boldsymbol{A}}_z = (\hat{\boldsymbol{a}}_1^\intercal, \ldots, \hat{\boldsymbol{a}}_J^\intercal)^\intercal$, $\hat{\boldsymbol{a}}_j = (n_{j1}, \hat{m}_{j1} n_{j1}, \hat{m}_{j2} n_{j1}, n_{j2}, \hat{m}_{j1} n_{j2}, \hat{m}_{j2} n_{j2}), \hat{\boldsymbol{m}}_j = (\boldsymbol{m}_j - \boldsymbol{z})/h_z = (\hat{m}_{j1}, \hat{m}_{j2})$. Finally $R_h^0 \boldsymbol{q}_h(\boldsymbol{z}) = (\hat{c}_1, \hat{c}_4)^\intercal$.

Example 2. *RT_1 elements on triangular meshes.* In this case, $R_h^1 \boldsymbol{q}_h$ is a continuous piecewise quadratic function. At step 1, let $\{e_j\}_{j=1}^J = \mathcal{E}_h(\omega_z)$ and $\{T_j\}_{j=1}^L = \mathcal{T}_h(\omega_z)$. Let

$$\boldsymbol{q}_z = \begin{pmatrix} c_1 + c_2 x + c_3 x_2 + c_4 x^2 + c_5 x_1 x_2 + c_6 x_2^2 \\ c_7 + c_8 x + c_9 x_2 + c_{10} x_1^2 + c_{11} x_1 x_2 + c_{12} x_2^2 \end{pmatrix} \in \mathcal{P}_2(\omega_z)^2$$

minimize

$$\mathcal{F}(\boldsymbol{q}) = \sum_{j=1}^{J} (\boldsymbol{q}(\boldsymbol{x}_j)\cdot\boldsymbol{n}_j - \boldsymbol{q}_h(\boldsymbol{x}_j)\cdot\boldsymbol{n}_j)^2 + (\boldsymbol{q}(\boldsymbol{y}_j)\cdot\boldsymbol{n}_j - \boldsymbol{q}_h(\boldsymbol{y}_j)\cdot\boldsymbol{n}_j)^2 + \sum_{l=1}^{L}\sum_{m=1}^{2}\left(\fint_{T_l} q_m - \fint_{T_l} q_{h,m}\right)^2, \quad \boldsymbol{q}\in\mathcal{P}_2(\omega_z)^2,$$

where $\boldsymbol{q} = (q_1, q_2)^\intercal, \boldsymbol{q}_h = (q_{h,1}, q_{h,2})^\intercal$, $\boldsymbol{x}_j = \frac{3+\sqrt{3}}{6}\boldsymbol{a}_j + \frac{3-\sqrt{3}}{6}\boldsymbol{b}_j, \boldsymbol{y}_j = \frac{3-\sqrt{3}}{6}\boldsymbol{a}_j + \frac{3+\sqrt{3}}{6}\boldsymbol{b}_j$, and $e_j = \overline{\boldsymbol{a}_j\boldsymbol{b}_j}$. Equivalently, $\boldsymbol{c}_z = (c_1, \ldots, c_{12})^\intercal$ solves the normal equation $\boldsymbol{A}_z^\intercal\boldsymbol{A}_z\boldsymbol{c}_z = \boldsymbol{A}_z^\intercal\boldsymbol{d}_z$, where

$$\boldsymbol{d}_z = \left(\boldsymbol{q}_h(\boldsymbol{x}_1)\cdot\boldsymbol{n}_1, \boldsymbol{q}_h(\boldsymbol{y}_1)\cdot\boldsymbol{n}_1, \boldsymbol{q}_h(\boldsymbol{x}_2)\cdot\boldsymbol{n}_2, \boldsymbol{q}_h(\boldsymbol{y}_2)\cdot\boldsymbol{n}_2, \ldots, \boldsymbol{q}_h(\boldsymbol{y}_J)\cdot\boldsymbol{n}_J, \fint_{T_1} q_{h,1}, \fint_{T_1} q_{h,2}, \fint_{T_2} q_{h,1}, \fint_{T_2} q_{h,2}, \ldots, \fint_{T_L} q_{h,2}\right)^\intercal,$$

$\boldsymbol{A}_z = (\boldsymbol{a}_1^\intercal, \ldots, \boldsymbol{a}_{2J+2L}^\intercal)^\intercal$ is an $(2J+2L)\times 12$ matrix,

$$\begin{aligned}
&\boldsymbol{a}_{2j-1} = (n_{j1}\boldsymbol{\xi}_j, n_{j2}\boldsymbol{\xi}_j), \quad \boldsymbol{a}_{2j} = (n_{j1}\boldsymbol{\eta}_j, n_{j2}\boldsymbol{\eta}_j),\\
&\boldsymbol{\xi}_j = (1, x_{j1}, x_{j2}, x_{j1}^2, x_{j1}x_{j2}, x_{j2}^2),\\
&\boldsymbol{\eta}_j = (1, y_{j1}, y_{j2}, y_{j1}^2, y_{j1}y_{j2}, y_{j2}^2), \quad 1\le j\le J,\\
&\boldsymbol{a}_{2J+2l-1} = \fint_{T_l}(1, x_1, x_2, x_1^2, x_1x_2, x_2^2, 0, 0, 0, 0, 0, 0),\\
&\boldsymbol{a}_{2J+2l} = \fint_{T_l}(0, 0, 0, 0, 0, 0, 1, x_1, x_2, x^2, x_1x_2, x_2^2), \quad 1\le l\le L.
\end{aligned}$$

Then $R_h^1\boldsymbol{q}_h(\boldsymbol{z}) = \boldsymbol{q}_z(\boldsymbol{z})$ for $\boldsymbol{z}\in\mathcal{N}_h$. At step 2, for the midpoint $\boldsymbol{z}$ of the edge $e = \overline{\boldsymbol{z}_1\boldsymbol{z}_2}$, $R_h^1\boldsymbol{q}_h(\boldsymbol{z}) = (\boldsymbol{q}_{z_1}(\boldsymbol{z}) + \boldsymbol{q}_{z_2}(\boldsymbol{z}))/2$. one can again introduce the scaled polynomial $\hat{\boldsymbol{q}}_z(\hat{\boldsymbol{x}}) = \boldsymbol{q}_z(\boldsymbol{z} + h_z\hat{\boldsymbol{x}})$ in practice.

Assume that the solution of each local LS problem at each vertex $\boldsymbol{z}$ is unique. By definition R_h^r preserves $(r+1)$-degree polynomials, namely, $R_h^r\boldsymbol{q} = \boldsymbol{q}$ on T for

$\boldsymbol{q} \in \mathcal{P}_{r+1}(\omega_T)^2$. Then the polynomial preserving property (cf.[71]) leads to the super-approximation property

$$\|\boldsymbol{q} - R_h^r \boldsymbol{q}\|_{0,\Omega} = O(h^{r+2}). \tag{2.51}$$

However, it's not obvious that these local LS problems are uniquely solvable. The next obvious lemma gives several statements equivalent to uniqueness.

Lemma 2.4.1. *The following statements are equivalent:*

1. *There exists a unique* $\boldsymbol{q}_z$ *at* $\boldsymbol{z}$.
2. $A_z \boldsymbol{c} = \boldsymbol{0}$ *implies* $\boldsymbol{c} = \boldsymbol{0}$.
3. $\Pi_h^r \boldsymbol{q}_z = 0$ *on* ω_z *implies* $\boldsymbol{q}_z \equiv 0$.

Hence it suffices to study the unisolvence of Π_h^r on $\mathcal{P}_{r+1}(\omega_z)^2$. Π_h^r is moment-based interpolation while nodal interpolation is often preferred. The next lemma relates statement 3 to nodal interpolation.

Lemma 2.4.2. *Assume* $\Pi_h^r \boldsymbol{q}_z = 0$ *on* ω_z. *Then* $\boldsymbol{q}_z = \nabla^\perp w$ *for some* $w \in \mathcal{P}_{r+2}(\omega_z)$. *In addition, for any* $e \in \mathcal{E}(\omega_z)$, $w(\boldsymbol{l}) = 0$ *for any Lobatto quadrature point* $\boldsymbol{l}$ *on* e.

Proof. $\Pi_h^r \boldsymbol{q}_z = 0$ and (2.12) imply

$$\operatorname{div} \boldsymbol{q}_z = \operatorname{div}(\boldsymbol{q}_z - \Pi_h^r \boldsymbol{q}_z) = \operatorname{div} \boldsymbol{q}_z - \Pi_h^r \operatorname{div} \boldsymbol{q}_z = 0.$$

Hence $\boldsymbol{q}_z = \nabla^\perp w$ for some $w \in \mathcal{P}_{r+2}(\omega_z)$. Given $e = \overline{\boldsymbol{ab}} \in \mathcal{E}_h(\omega_z)$,

$$w(\boldsymbol{b}) - w(\boldsymbol{a}) = \int_e \partial_{\boldsymbol{t}_e} w = \int_e \boldsymbol{q}_z \cdot \boldsymbol{n}_e = \int_e \Pi_h^r \boldsymbol{q}_z \cdot \boldsymbol{n}_e = 0.$$

Hence $w(\boldsymbol{z}) \equiv c$ at all vertices $\boldsymbol{z}$ in ω_z. By subtracting c from w, we can assume that

w vanishes at all vertices. For $v \in \mathcal{P}_r(e)$, it then follows

$$\int_e r\partial_{\boldsymbol{t}_e} v = -\int_e v\partial_{\boldsymbol{t}_e} r = -\int_e \boldsymbol{q}_z \cdot \boldsymbol{n}_e v = -\int_e \Pi_h^r \boldsymbol{q}_z \cdot \boldsymbol{n}_e v = 0,$$

and thus

$$\int_e rw = 0 \quad \text{for all } w \in \mathcal{P}_{r-1}(e). \tag{2.52}$$

Notice that on e, the Lobatto quadrature $\int_e f = \sum_{j=1}^{r+2} \mu_j f(\boldsymbol{l}_j)$ is exact for $f \in \mathcal{P}_{2r+1}(e)$, where $\boldsymbol{l}_j = \boldsymbol{a} + (\boldsymbol{b} - \boldsymbol{a})\hat{l}_j$, $\{\hat{l}_j\}_{j=1}^{r+2}$ are zeros of $\frac{d^r}{ds^r}(s^{r+1}(1-s)^{r+1})$, and $\{\mu_j\}_{j=1}^{r+2}$ are corresponding weights. Let w be the polynomial which is μ_j^{-1} at $\boldsymbol{l}_j$ and 0 at rest of the $(r-1)$ interior quadrature points $\{\boldsymbol{l}_i\}_{i=2,i\neq j}^{r+1}$ in (2.52). Then $r(\boldsymbol{l}_j) = \int_e rw = 0$. The proof is complete. □

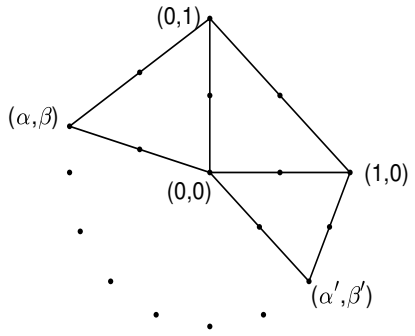

Figure 2.1: a local patch containing the reference triangle

The next theorem gives practical criteria for checking the well-posedness of R_h^0 and R_h^1.

Theorem 2.4.3. *Let $\boldsymbol{z}$ be a vertex in $\mathcal{T}_h$. If $\#\mathcal{T}_h(\omega_z) \geq 5$ and the sum of each pair of adjacent angles in ω_z is $\leq \pi$, then there exists a unique $\boldsymbol{q}_z$ at $\boldsymbol{z}$ for R_h^0. If $\#\mathcal{T}_h(\omega_z) \geq 4$, then there exists a unique $\boldsymbol{q}_z$ at $\boldsymbol{z}$ for R_h^1.*

Proof. Assume $\Pi_h^r \boldsymbol{q}_z = 0$ on ω_z. By Lemma 2.4.2, $\boldsymbol{q}_z = \nabla^\perp w$ for some $w \in \mathcal{P}_{r+2}(\omega_z)$. If $r = 0$, then w vanishes on all vertices in ω_z and thus $\boldsymbol{q}_z = 0$ by Theorem 2.3 in [57].

If $r = 1$, $w \in \mathcal{P}_3(\omega_z)$ vanishes at all vertices and midpoints of edges in ω_z. Without loss of generality, we can assume that $\boldsymbol{z} = (0,0)$ and the reference triangle $\hat{T}$ spanned by $(0,0),(0,1),(1,0)$ is in $\mathcal{T}(\omega_z)$.

If w is reducible, then the zero set $w^{-1}(0)$ is the union of three straight lines (counting multiplicity) or the union of a straight line and a conic. Clearly three lines cannot pass all vertices and midpoints in ω_z provided $\#\mathcal{T}_h(\omega_z) \geq 4$. If $w^{-1}(0)$ contains a conic branch C, clearly C must contain at least two vertices $\boldsymbol{a}, \boldsymbol{b}$ in ω_z. However, C cannot pass through $\frac{\boldsymbol{a}+\boldsymbol{b}}{2}$ by elementary geometry.

Hence a reducible w cannot vanish at all nodes in ω_z and we can assume

$$w(x_1,x_2) = c_1x_1^3 + c_2x_1^2x_2 + c_3x_1x_2^2 + c_4x_2^3 + c_5x_2^2 + c_6x_1x_2 + c_7x_2^2 + c_8x_1 + c_9x_2$$

is irreducible. Furthermore, we can assume one of the coefficients of highest order terms is 1, say $c_1 = 1$(similar argument for c_2, c_3 or $c_4 = 1$). Let (α,β) be the vertex outside $\hat{T}$ next to $(0,1)$, see Figure 2.1. Solving the linear system of equations

$$\begin{aligned} &w(1,0) = w(0,1) = w(1/2,0) = w(0,1/2) = w(1/2,1/2) \\ &= w(\alpha,\beta) = w\left(\frac{\alpha}{2}, \frac{\beta+1}{2}\right) = w\left(\frac{\alpha}{2}, \frac{\beta}{2}\right) = 0, \end{aligned}$$

we have

$$c_1 = \frac{3-3\alpha}{1+\beta}, \quad c_2 = \frac{3\alpha(\alpha-1)}{\beta(1+\beta)}. \tag{2.53}$$

Note that $\beta \neq 0, \beta \neq -1$ in (2.53), otherwise the irreducible cubic curve $w^{-1}(0)$ intersects with a line at five distinct points, which is impossible by Bézout's theorem (cf.[61]). Also $\alpha \neq 1$ otherwise it violates the topology of the patch ω_i. Hence $\alpha/\beta =$

$-c_2/c_1$. Let (α', β') be the vertex outside $\hat{T}$ next to $(1, 0)$. Similarly we have $\alpha'/\beta' = -c_2/c_1$. Then it forces that $(\alpha, \beta) = (\alpha', \beta')$, which contradicts $\#\mathcal{T}_h(\omega_z) \geq 4$. Hence $w \equiv 0$ and $\boldsymbol{q}_z \equiv 0$.

Therefore by Lemma 2.4.1, there exists a unique $\boldsymbol{q}_z$ for $r = 0, 1$. □

We say a vertex z is good if the condition in Theorem 2.4.3 holds at z, otherwise it is a bad vertex. In practice, $\mathcal{T}_h$ typically has a few bad vertices, e.g., boundary vertices. There are several ways of dealing with a bad vertex z. If z is directly connected to a good vertex z', one can define $\omega_z := \omega_{z'}$ and thus $\boldsymbol{A}_z$ is of full column rank. More conveniently, one can empirically add some extra elements to the patch ω_z in practice, e.g., enlarge ω_z by one layer. Alternatively, one can solve a rank-deficient local least squares problem, which might reduce the super-approximation property of R_h:

In the rest of this chapter, we assume that

At each vertex $\boldsymbol{z}$, there exists a unique $\boldsymbol{q}_z$.

Using the uniqueness of the local LS solution, we can obtain boundedness of R_h^r.

Theorem 2.4.4. *For $\boldsymbol{q}_h \in \mathcal{Q}_h^r$ and $T \in \mathcal{T}_h$,*

$$\|R_h^r \boldsymbol{q}\|_{0,T} \lesssim \|\boldsymbol{q}\|_{0,\omega_T}, \quad r = 0, 1.$$

Proof. For $\boldsymbol{z} \in \mathcal{N}_h$, Let $\sigma_{\min}$ and $\sigma_{\max}$ be the minimum and maximum singular values of $\hat{\boldsymbol{A}}_z$ respectively. The goal is to show that $\sigma_{\min}$ is uniformly bounded away from 0. MAC implies $\#\mathcal{T}_h(\omega_z) \leq N_{\max} = 2\pi/\Theta$. Hence it suffices to consider the case $\#\mathcal{T}_h(\omega_z) = N$ for some fixed $N \leq N_{\max}$. In this case, $\#\mathcal{E}_h(\omega_z) = 2N$. Let $N_1 = 2N, N_2 = 6$ provided $k = 0$ and $N_1 = 6N, N_2 = 12$ provided $k = 1$. Let $M_{N_1 \times N_2}$ and $S_{N_1 \times N_2}$ be the set of $N_1 \times N_2$ matrices and $N_1 \times N_2$ rank-deficient matrices, respectively. It is well known

that $\sigma_{\min} = \mathrm{dist}(\hat{\boldsymbol{A}}_z, S_{N_1\times N_2})$, the distance (measured by matrix 2-norm) from $\hat{\boldsymbol{A}}_z$ to rank-deficient matrices. $\mathrm{dist}(\cdot, S_{N_1\times N_2})$ is continuous on $M_{N_1\times N_2}$. Recall that $\hat{\boldsymbol{A}}_z$ is the scaled LS coefficient matrix determined by ω_z. Consider all possible ω_z and define

$$\mathcal{A}_z = \{\hat{\boldsymbol{A}}_z \in M_{N_1\times N_2} : \#\mathcal{T}_h(\omega_z) = N, \omega_z \text{ satisfies MAC}\}.$$

Clearly $\mathcal{A}_z$ is a compact set in $M_{N_1\times N_2}$ and any $\hat{\boldsymbol{A}}_z \in \mathcal{A}_z$ is of full rank by the uniqueness assumption, which imply that $\sigma_{\min} = \mathrm{dist}(\hat{\boldsymbol{A}}_z, S_{N_1\times N_2}) \geq C_1 > 0$, where C_1 depends only on the minimum angle Θ. The maximum singular value $\sigma_{\max} \leq C_2$, where C_2 only depends on Ω. For $\boldsymbol{q}_h \in \mathcal{Q}_h^k$,

$$\begin{aligned} |\hat{\boldsymbol{c}}_z| &\leq \|(\hat{\boldsymbol{A}}_z^\intercal \hat{\boldsymbol{A}}_z)^{-1}\|_2 |\hat{\boldsymbol{A}}_z^\intercal \boldsymbol{d}_z| \leq \sigma_{\min}^{-2}\sigma_{\max}|\boldsymbol{d}_z| \\ &\leq C_1^{-2}C_2\|\boldsymbol{q}_h\|_{0,\infty,\omega_z} \lesssim h_z^{-1}\|\boldsymbol{q}_h\|_{0,\omega_z}, \end{aligned} \tag{2.54}$$

where $|\cdot|$ is the Euclidean norm. Finally by (2.54), we have

$$\|R_h^r \boldsymbol{q}_h\|_{0,T} \lesssim h\|R_h^r \boldsymbol{q}_h\|_{0,\infty,T} \lesssim h|\hat{\boldsymbol{c}}_z| \lesssim \|\boldsymbol{q}_h\|_{0,\omega_T},$$

which completes the proof. □

The super-approximation property of R_h follows from the uniqueness and boundedness results.

Theorem 2.4.5. *For $\boldsymbol{q} \in H^{r+2}(\Omega)$,*

$$\|\boldsymbol{q} - R_h^r \boldsymbol{q}\|_{0,\Omega} \lesssim h^{r+2}|\boldsymbol{q}|_{r+2,\Omega}, \quad r = 0, 1.$$

Proof. Let $T = \overline{\boldsymbol{z}_1\boldsymbol{z}_2\boldsymbol{z}_3} \in \mathcal{T}_h$ and $T_1 \subset \overline{\Omega}$ be a smallest local triangle containing ω_T. Let $\boldsymbol{q}_{r+1} \in \mathcal{P}_{r+1}(T_1)^2$ be the degree-$(r+1)$ local Lagrange interpolant of $\boldsymbol{q}$ using based

on T_1. By the uniqueness assumption, $R_h^r \boldsymbol{q}_{r+1} = \boldsymbol{q}_{r+1}$ on T. It then follows from $R_h^r \Pi_h^r = R_h^r$ that

$$\|\boldsymbol{q} - R_h^r \boldsymbol{q}\|_{0,T} \leq \|\boldsymbol{q} - \boldsymbol{q}_{r+1}\|_{0,T} + \|R_h^r \Pi_h^r (\boldsymbol{q}_{k+1} - \boldsymbol{q})\|_{0,T}. \tag{2.55}$$

Using the boundedness from Theorem 2.4.4 and (1.8),

$$\begin{aligned} \|R_h^r \Pi_h^r (\boldsymbol{q}_{r+1} - \boldsymbol{q})\|_{0,T} &\lesssim \|\Pi_h^r (\boldsymbol{q}_{r+1} - \boldsymbol{q})\|_{0,\omega_T} \\ &\lesssim h \|\Pi_h^r (\boldsymbol{q}_{r+1} - \boldsymbol{q})\|_{0,\infty,\omega_T} \lesssim h \|\boldsymbol{q}_{r+1} - \boldsymbol{q}\|_{0,\infty,\omega_T} \lesssim h^{r+2} |\boldsymbol{q}|_{r+2,T_1}. \end{aligned} \tag{2.56}$$

Combining (2.55), (2.56) and the shape regularity $\mathcal{T}_h$ completes the proof. □

Combining Theorems 2.3.5 and 2.4.5, we obtain the superconvergent recovery estimate.

Theorem 2.4.6. *Assume that $\mathcal{T}_h$ satisfies the (α, β)-condition. Then*

$$\|\boldsymbol{p} - R_h^r \boldsymbol{p}_h^r\|_a \lesssim h^{r+1+\min(\frac{1}{2}, \alpha, \frac{\beta}{2})} \big(|\boldsymbol{p}|_{r+1,\infty,\Omega} + \|\boldsymbol{p}\|_{r+2,\Omega} \big), \quad r = 0, 1.$$

Assume that $\mathcal{T}_h$ satisfies the piecewise strong (α, β)-condition. Then for the RT_0 finite element,

$$\|\boldsymbol{p} - R_h^0 \boldsymbol{p}_h^0\|_{0,\Omega} \lesssim h^{1+\min(1, \alpha, \frac{\beta}{2})} |\log h|^{\frac{1}{2}} \|u\|_{3,\infty,\Omega}.$$

Proof. The theorem follows from

$$\|\boldsymbol{p} - R_h^r \boldsymbol{p}_h^r\|_{0,\Omega} \leq \|\boldsymbol{p} - R_h^r \boldsymbol{p}\|_{0,\Omega} + \|R_h^r (\Pi_h^r \boldsymbol{p} - \boldsymbol{p}_h^r)\|_{0,\Omega},$$

and Theorems 2.3.5 and 2.4.5. □

2.5 Proofs of Lemma 2.2.4

To prove Lemma 2.2.4, we introduce cubic bubble functions

$$\psi_0 = \lambda_1\lambda_2\lambda_3, \quad \psi_k = \lambda_{k-1}\lambda_{k+1}(\lambda_{k-1} - \lambda_{k+1}), \quad 1 \le k \le 3.$$

By counting the dimension, it is clear that $\{\psi_k\}_{k=0}^3$ can span polynomials in $\mathcal{P}_3(T)$ that vanish at $\{z_k\}_{k=1}^3$ and midpoints of $\{e_k\}_{k=1}^3$. In fact, $\{\psi_k\}_{k=0}^3$ has been used to derive superconvergence of quadratic Lagrange elements (cf. [42]) and a posteriori error estimators (cf. [11]). In addition, we have the several elementary identities:

$$\begin{aligned}
&\cos\theta_k = (\ell_{k-1}^2 + \ell_{k+1}^2 - \ell_k^2)/(2\ell_{k-1}\ell_{k+1}), \quad \sin\theta_k = \ell_k/d, \quad d_k = \ell_{k-1}\ell_{k+1}/d,\\
&\boldsymbol{n}_{k-1} = -\sin\theta_{k+1}\boldsymbol{t}_k - \cos\theta_{k+1}\boldsymbol{n}_k, \quad \boldsymbol{n}_{k+1} = \sin\theta_{k-1}\boldsymbol{t}_k - \cos\theta_{k-1}\boldsymbol{n}_k,\\
&\partial_{\boldsymbol{t}_{k-1}}^2 = \cos^2\theta_{k+1}\partial_{\boldsymbol{t}_k}^2 - 2\cos\theta_{k+1}\sin\theta_{k+1}\partial_{\boldsymbol{t}_k\boldsymbol{n}_k}^2 + \sin^2\theta_{k+1}\partial_{\boldsymbol{n}_k}^2,\\
&\partial_{\boldsymbol{t}_{k+1}}^2 = \cos^2\theta_{k-1}\partial_{\boldsymbol{t}_k}^2 + 2\cos\theta_{k-1}\sin\theta_{k-1}\partial_{\boldsymbol{t}_k\boldsymbol{n}_k}^2 + \sin^2\theta_{k-1}\partial_{\boldsymbol{n}_k}^2.
\end{aligned} \tag{2.57}$$

For each edge e_k, we define several associated geometric quantities $\{\alpha_{jl,k}^i\}_{1\le i,j,l\le 2}$

$$\begin{aligned}
&\alpha_{11,k}^1 = \frac{1}{24d\ell_k^2}\ell_{k-1}\ell_{k+1}\big(3\ell_k^4 - (\ell_{k-1}^2 - \ell_{k+1}^2)^2\big),\\
&\alpha_{12,k}^1 = \alpha_{21,k}^1 = \frac{1}{12d^2\ell_k}\ell_{k-1}^2\ell_{k+1}^2(\ell_{k-1}^2 - \ell_{k+1}^2), \quad \alpha_{22,k}^1 = -\frac{1}{6d^3}\ell_{k+1}^3\ell_{k-1}^3,\\
&\alpha_{11,k}^2 = \frac{1}{48\ell_k^3}(\ell_{k-1}^2 - \ell_{k+1}^2)\big(9\ell_k^4 - (\ell_{k-1}^2 - \ell_{k+1}^2)^2\big),\\
&\alpha_{12,k}^2 = \alpha_{21,k}^2 = -\alpha_{11,k}^1, \quad \alpha_{22,k}^2 = -\alpha_{12,k}^1.
\end{aligned}$$

With this preparation, we present an analogue of Lemma 2.2.1.

Lemma 2.5.1. *For $\boldsymbol{p}_2 \in \mathcal{P}_2(T)^2$,*

$$\boldsymbol{p}_2 - \Pi_h^1\boldsymbol{p}_2 = \nabla^\perp w,$$

where

$$w = \alpha^i_{jl,k}\mathcal{D}^{jl}_{i,k}(\boldsymbol{p}_2)\psi_0 + \sum_{k=1}^{3}\frac{\ell_k^3}{12}\mathcal{D}^{11}_{2,k}(\boldsymbol{p}_2)\psi_k.$$

Proof. Notice that $\mathcal{N}_{1,k}(\boldsymbol{p}_2 - \Pi_h^1\boldsymbol{p}_2) = \mathcal{N}_{2,k}(\boldsymbol{p}_2 - \Pi_h^1\boldsymbol{p}_2) = \mathcal{N}_{3,l}(\boldsymbol{p}_2 - \Pi_h^1\boldsymbol{p}_2) = 0$ for $k = 1, 2, 3$ and $l = 1, 2$. Then Lemma 2.4.2 leads to

$$\boldsymbol{p}_2 - \Pi_h^1\boldsymbol{p}_2 = \nabla^\perp\big(\sum_{k=0}^{3} c_k\psi_k\big). \tag{2.58}$$

For a unit vector $\boldsymbol{d}$ and the directional derivative $\partial_{\boldsymbol{d}}$, the definition of $\mathcal{RT}_1(T)$ implies that $\partial^2_{\boldsymbol{d}}\Pi_h^1\boldsymbol{p}_2$ is proportional to $\boldsymbol{d}$. Then applying $\boldsymbol{d}^\perp\cdot\partial^2_{\boldsymbol{d}}$ to (2.58) gives

$$\boldsymbol{d}^\perp\cdot\partial^2_{\boldsymbol{d}}\boldsymbol{p}_2 = \sum_{k=0}^{3} c_k\partial^3_{\boldsymbol{d}}\psi_k. \tag{2.59}$$

By direct calculation,

$$\partial^3_{\boldsymbol{d}}\psi_0 = 6\partial_{\boldsymbol{d}}\lambda_1\partial_{\boldsymbol{d}}\lambda_2\partial_{\boldsymbol{d}}\lambda_3, \tag{2.60a}$$

$$\partial^3_{\boldsymbol{d}}\psi_k = 6\partial_{\boldsymbol{d}}\lambda_{k-1}\partial_{\boldsymbol{d}}\lambda_{k+1}(\partial_{\boldsymbol{d}}\lambda_{k-1} - \partial_{\boldsymbol{d}}\lambda_{k+1}), \quad 1 \le k \le 3. \tag{2.60b}$$

In particular, $\partial^3_{\boldsymbol{t}_k}\psi_0 = 0$ and $\partial^3_{\boldsymbol{t}_k}\psi_j = -12\delta_{jk}/\ell_k^3$. By (2.60) and (2.59) with $\boldsymbol{d} = \boldsymbol{t}_k$, we have

$$c_k = \frac{\ell_k^3}{12}\boldsymbol{n}_k\cdot\partial^2_{\boldsymbol{t}_k}\boldsymbol{p}_2 = \frac{\ell_k^3}{12}\mathcal{D}^{11}_{2,k}(\boldsymbol{p}_2), \quad 1 \le k \le 3. \tag{2.61}$$

It remains to determine c_0. It follows from (2.59) with $\boldsymbol{d} = \boldsymbol{n}_k$ that

$$\mathcal{D}^{22}_{1,k}(\boldsymbol{p}_2) = c_0\partial^3_{\boldsymbol{n}_k}\psi_0 + c_k\partial^3_{\boldsymbol{n}_k}\psi_k + c_{k-1}\partial^3_{\boldsymbol{n}_k}\psi_{k-1} + c_{k+1}\partial^3_{\boldsymbol{n}_k}\psi_{k+1}, \tag{2.62}$$

Using $\partial_{\boldsymbol{n}_k}\lambda_k = -1/d_k, \partial_{\boldsymbol{n}_k}\lambda_{k+1} = \cos\theta_{k-1}/d_{k+1}, \partial_{\boldsymbol{n}_k}\lambda_{k-1} = \cos\theta_{k+1}/d_{k-1}$, (2.60) with

$\boldsymbol{d} = \boldsymbol{n}_k$, (2.61), and (2.62), we obtain

$$\begin{aligned} c_0 = &-\frac{d_{k-1}d_k d_{k+1}}{6\cos\theta_{k-1}\cos\theta_{k+1}}\mathcal{D}^{22}_{1,k}(\boldsymbol{p}_2) \\ &+ \frac{\ell_k^3}{12} d_k \left(\frac{\cos\theta_{k+1}}{d_{k-1}} - \frac{\cos\theta_{k-1}}{d_{k+1}}\right)\mathcal{D}^{11}_{2,k}(\boldsymbol{p}_2) \\ &- \frac{\ell_{k-1}^3}{12}\frac{d_{k-1}}{\cos\theta_{k+1}}\left(\frac{1}{d_k} + \frac{\cos\theta_{k-1}}{d_{k+1}}\right)\mathcal{D}^{11}_{2,k-1}(\boldsymbol{p}_2) \\ &+ \frac{\ell_{k+1}^3}{12}\frac{d_{k+1}}{\cos\theta_{k-1}}\left(\frac{1}{d_k} + \frac{\cos\theta_{k+1}}{d_{k-1}}\right)\mathcal{D}^{11}_{2,k+1}(\boldsymbol{p}_2). \end{aligned} \tag{2.63}$$

Plugging in (2.57), we obtain $c_0 = \alpha^i_{jl,k}\mathcal{D}^{jl}_{i,k}(\boldsymbol{p}_2), 1 \le k \le 3$. □

Now we can prove Lemma 2.2.4. In the proof, we use the integral formulas

$$\begin{aligned} \int_T \lambda_1^{m_1}\lambda_2^{m_2}\lambda_3^{m_3} &= \frac{2|T|m_1!m_2!m_3!}{(m_1+m_2+m_3+2)!}, \\ \int_e \lambda_1^{m_1}\lambda_2^{m_2} &= \frac{|e|m_1!m_2!}{(m_1+m_2+1)!}, \end{aligned} \tag{2.64}$$

where $\{m_i\}_{i=1}^3$ are non-negative integers and λ_1, λ_2 are barycentric coordinates w.r.t. the edge e.

Proof of Lemma 2.2.4. Using (2.9) and Lemma 2.5.1, we have

$$\begin{aligned} \int_T (\boldsymbol{p}_2 - \Pi_h^1\boldsymbol{p}_2)\cdot\nabla^\perp w_2 &= \sum_{k=1}^3 \int_{e_k} w\nabla^\perp r_2\cdot\boldsymbol{t}_k - \int_T r\Delta w_2 \\ &= I + II. \end{aligned} \tag{2.65}$$

Recall that $\phi_k = \lambda_{k-1}\lambda_{k+1}$. Then using the hierarchical representation

$$w_2 - I_h^1 w_2 = -\frac{1}{2}\sum_{k=1}^3 \ell_k^2\phi_k\partial^2_{\boldsymbol{t}_k} w_2, \tag{2.66}$$

and $\Delta\phi_k = 2\nabla\lambda_{k-1}\cdot\nabla\lambda_{k+1} = -2\cos\theta_k/(d_{k-1}d_{k+1})$, we obtain

$$\Delta w_2 = \frac{1}{4|T|^2}\sum_{k=1}^{3}\ell_k^2\ell_{k-1}\ell_{k+1}\cos\theta_k\partial^2_{\boldsymbol{t}_k}w_2. \tag{2.67}$$

It then follows from Lemma 2.2.4, (2.67), and $\int_T\psi_0 = |T|/60$, $\int_T\psi_k = 0, 1\le k\le 3$, that

$$\begin{aligned} II &= -\frac{|T|}{60}c_0\Delta w_2 \\ &= -\frac{1}{240|T|}\sum_{k=1}^{3}c_0\ell_k^2\ell_{k-1}\ell_{k+1}\cos\theta_k\partial^2_{\boldsymbol{t}_k}w_2 \\ &= -\frac{1}{120}\sum_{k=1}^{3}\int_{e_k}\alpha^i_{jl,k}\mathcal{D}^{jl}_{i,k}(\boldsymbol{p}_2)\ell_k\cot\theta_k\partial^2_{\boldsymbol{t}_k}w_2. \end{aligned} \tag{2.68}$$

By Lemma 2.2.4, (2.20), and $\psi_k = -\ell_k\partial_{\boldsymbol{t}_k}\phi_k^2/2$, we have

$$\begin{aligned} I &= -\sum_{k=1}^{3}\frac{1}{12}\int_{e_k}\ell_k^3\mathcal{D}^{11}_{2,k}(\boldsymbol{p}_2)\psi_k\nabla^\perp w_2\cdot\left(\frac{\cos\theta_{k-1}}{\sin\theta_k}\boldsymbol{n}_{k-1} - \frac{\cos\theta_{k+1}}{\sin\theta_k}\boldsymbol{n}_{k+1}\right) \\ &= \sum_{k=1}^{3}\frac{1}{24}\int_{e_k}\ell_k^4\mathcal{D}^{11}_{2,k}(\boldsymbol{p}_2)\phi_k^2\left(\frac{\cos\theta_{k-1}}{\sin\theta_k}\partial^2_{\boldsymbol{t}_k\boldsymbol{t}_{k-1}}w_2 - \frac{\cos\theta_{k+1}}{\sin\theta_k}\partial^2_{\boldsymbol{t}_k\boldsymbol{t}_{k+1}}w_2\right). \end{aligned} \tag{2.69}$$

Then using the quadrature rule (2.64),

$$I = \frac{1}{720}\sum_{k=1}^{3}\ell_k^5\mathcal{D}^{11}_{2,k}(\boldsymbol{p}_2)\left(\frac{\cos\theta_{k-1}}{\sin\theta_k}\partial^2_{\boldsymbol{t}_k\boldsymbol{t}_{k-1}}w_2 - \frac{\cos\theta_{k+1}}{\sin\theta_k}\partial^2_{\boldsymbol{t}_k\boldsymbol{t}_{k+1}}w_2\right).$$

In addition, (2.66) gives

$$\begin{aligned} \partial^2_{\boldsymbol{t}_k\boldsymbol{t}_{k-1}}w_2 &= -\frac{\ell_k}{2\ell_{k-1}}\partial^2_{\boldsymbol{t}_k}w_2 + \frac{\ell_{k+1}^2}{2\ell_{k-1}\ell_k}\partial^2_{\boldsymbol{t}_{k+1}}w_2 - \frac{\ell_{k-1}}{2\ell_k}\partial^2_{\boldsymbol{t}_{k-1}}w_2, \\ \partial^2_{\boldsymbol{t}_k\boldsymbol{t}_{k+1}}w_2 &= -\frac{\ell_k}{2\ell_{k+1}}\partial^2_{\boldsymbol{t}_k}w_2 - \frac{\ell_{k+1}}{2\ell_k}\partial^2_{\boldsymbol{t}_{k+1}}w_2 + \frac{\ell_{k-1}^2}{2\ell_k\ell_{k+1}}\partial^2_{\boldsymbol{t}_{k-1}}w_2. \end{aligned}$$

Therefore,

$$
\begin{aligned}
I = \frac{1}{1440}\sum_{k=1}^{3}\int_{e_k}\Bigg\{ & \frac{\ell_k^5}{\sin\theta_k}\mathcal{D}_{2,k}^{11}(\boldsymbol{p}_2)\left(\frac{\cos\theta_{k+1}}{\ell_{k+1}}-\frac{\cos\theta_{k-1}}{\ell_{k-1}}\right) \\
& + \frac{\ell_{k-1}^4}{\sin\theta_{k-1}}\mathcal{D}_{2,k-1}^{11}(\boldsymbol{p}_2)\left(\cos\theta_k+\frac{\ell_k}{\ell_{k+1}}\cos\theta_{k+1}\right) \\
& - \frac{\ell_{k+1}^4}{\sin\theta_{k+1}}\mathcal{D}_{2,k+1}^{11}(\boldsymbol{p}_2)\left(\frac{\ell_k}{\ell_{k-1}}\cos\theta_{k-1}+\cos\theta_k\right)\Bigg\}\partial_{\boldsymbol{t}_k}^2 w_2
\end{aligned}
\tag{2.70}
$$

Combining (2.68), (2.70) and using (2.57), we obtain Lemma 2.2.4. □

2.6 Numerical experiments

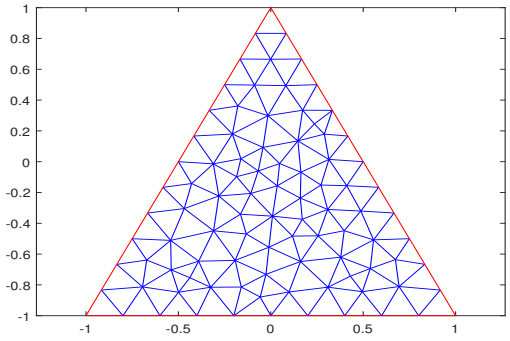

Figure 2.2: initial Delaunay triangulation

Problem1: We first test our recovery operators R_h^0 and R_h^1 on

$$-\Delta u + u = f, \quad \boldsymbol{x} \in \Omega,$$

where Ω is the triangle spanned by $(0,1)$, $(-1,-1)$ and $(1,-1)$. Let $u = \exp(x_1+x_2)$, $g = u|_{\partial\Omega}$ and f be the corresponding source term. In tables, 'nt' denotes the number of triangles. For RT_0 elements, we also test the postprocessing operator K_h in [12]. The experiments are performed using the PDE toolbox in Matlab 2016b. In this section,

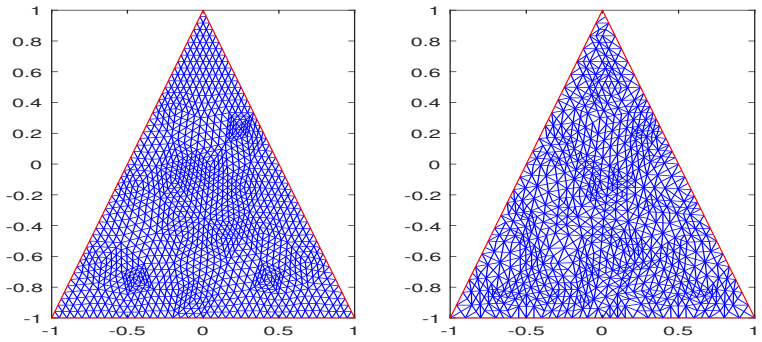

Figure 2.3: (left)regular refinement (right)longest edge bisection

the norm $\|\cdot\|$ is short for $\|\cdot\|_{0,\Omega}$. The order of convergence is p such that error$\approx$ dof$^{-\frac{p}{2}}$, where dof is the number of degrees of freedom. p is calculated by least squares fitting using data in tables. The first row is dropped in Tables 2.1–2.4 because of abnormal error reduction in coarse grids.

We start with the Delaunay triangulation in Figure 2.2, and computed a sequence of meshes by regular refinement, i.e., dividing an element into four similar subelements by connecting the midpoints of each edge, see Tables 2.1 and 2.2. We also computed a sequence of meshes by longest edge bisection, see Figure 2.3, Tables 2.3 and 2.4.

For regular refinement, the grids satisfy (α,β)-condition with $(\alpha,\beta)=(\infty,1)$ as well as piecewise strong (α,σ)-condition with $(\alpha,\beta)=(\infty,\infty)$ (i.e., piecewise uniformly parallel grids). For RT_0 elements, $\|\Pi_h^0\boldsymbol{p}-\boldsymbol{p}_h^0\|=O(h^2)$ and $\|\boldsymbol{p}-R_h^0\boldsymbol{p}_h^0\|=O(h^2)$ by Theorem 2.4.6, which are verified by Table 2.1. For RT_1 elements, Theorem 2.3.5 predicts that $\|\Pi_h^1\boldsymbol{p}-\boldsymbol{p}_h^1\|=O(h^{2.5})$, which is confirmed by Table 2.2. However, $\|\boldsymbol{p}-R_h^1\boldsymbol{p}_h^1\|=O(h^3)$ according to Table 2.2. It shows that the estimate $\|\boldsymbol{p}-R_h^1\boldsymbol{p}_h^1\|=O(h^{2.5})$ for RT_1 elements by Theorem 2.4.6 may be suboptimal.

For longest edge bisection, the resulting sequence of grids is completely unstructured, i.e., almost no pair of adjacent triangles forms an $O(h^{1+\alpha})$ approximate paral-

lelogram. Hence there is no supercloseness and superconvergence of K_h, see Tables 2.3 and 2.4. Surprisingly, there is still apparent superconvergence for $\|\boldsymbol{p} - R_h^k \boldsymbol{p}_h^k\|, k = 0, 1$.

Problem2: In the second experiment, we report a negative result for the lowest order Brezzi–Douglas–Marini (BDM_1) elements. We test the BDM_1 element for Poisson's equation $-\Delta u = f$ with exact solution $u = \exp(x_1 + x_2)$. Let $\tilde{\boldsymbol{p}}_h$ denote the numerical solution, $\tilde{\boldsymbol{p}}_I$ the canonical interpolant and $\tilde{R}_h$ the LS-based recovery operator fitting degrees of freedom of BDM_1 elements, respectively. We start with the uniformly parallel mesh, see Figure 2.4(left), and regularly refine it. Even though $\tilde{R}_h \tilde{\boldsymbol{p}}_I$ has excellent superapproximation property, there is no supercloseness and superconvergence of postprocessed solution to exact solution for BDM_1 elements, see Table 2.5.

Problem3: In the end, we test R_h^0 and R_h^1 by Poisson's equation $-\Delta u = f$ on the square $[-1, 1] \times [-1, 1]$ with a notch whose angle is $\omega = \pi/24$. We choose

$$u(r, \theta) = r^{\frac{\pi}{2\pi - \omega}} \sin\left(\frac{\pi}{2\pi - \omega}\theta\right) - \frac{r^2}{4},$$

where (r, θ) is the polar coordinate. The corresponding $f = 1$. We use $\eta_T = \|R_h^k \boldsymbol{p}_h^k - \boldsymbol{p}_h^k\|_{0,T}$ as a posteriori error estimator and start from the initial grid in Figure 2.4(right). The adaptive feedback loop is the classical "Solve $\rightarrow$ Estimate $\rightarrow$ Mark $\rightarrow$ Refine" loop (cf. [26, 55]). The simple recovery-based error indicator η_T generates correct girds for the solution with a point singularity, see Figure 2.7. In addition, there is apparent superconvergence under adaptively refined meshes, see Figure 2.5 and 2.6.

Chapter 2, in part, contains the original results in SIAM J. Numer. Anal. 56 (2018), 792–815, Li, Yu-Wen, 2018 Society for Industrial and Applied Mathematics. The dissertation author was the author of this paper.

Chapter 2, in part, has been submitted for publication of the material as it may appear in Journal of Scientific Computing, 2019, Bank, Randolph E.; Li, Yuwen, Springer, 2019. The dissertation author was the coauthor of this paper.

Table 2.1: RT_0 with regular refinement

nt	$\|\boldsymbol{p}-\boldsymbol{p}_h^0\|$	$\|\Pi_h^0\boldsymbol{p}-\boldsymbol{p}_h^0\|$	$\|\boldsymbol{p}-R_h^0\boldsymbol{p}_h^0\|$	$\|\boldsymbol{p}-K_h\boldsymbol{p}_h^0\|$
150	1.011e-1	2.118e-02	3.542e-02	2.948e-2
600	5.124e-2	6.060e-03	8.251e-03	9.782e-3
2400	2.574e-2	1.684e-03	2.114e-03	2.296e-3
9600	1.289e-2	4.595e-04	5.492e-04	1.121e-3
38400	6.447e-3	1.237e-04	1.434e-04	3.856e-4
order	1.002	1.880	1.957	1.562

Table 2.2: RT_1 with regular refinement

nt	$\|\boldsymbol{p}-\boldsymbol{p}_h^1\|$	$\|\Pi_h^1\boldsymbol{p}-\boldsymbol{p}_h^1\|$	$\|\boldsymbol{p}-R_h^1\boldsymbol{p}_h^1\|$
150	2.161e-3	4.740e-4	9.774e-4
600	5.471e-4	8.444e-5	9.061e-5
2400	1.376e-4	1.493e-5	1.058e-5
9600	3.451e-5	2.640e-6	1.460e-6
38400	8.641e-6	4.660e-7	2.271e-7
order	2.001	2.510	3.023

Table 2.3: RT_0 with bisection refinement

nt	$\|\boldsymbol{p}-\boldsymbol{p}_h^0\|$	$\|\Pi_h^0\boldsymbol{p}-\boldsymbol{p}_h^0\|$	$\|\boldsymbol{p}-R_h^0\boldsymbol{p}_h^0\|$	$\|\boldsymbol{p}-K_h\boldsymbol{p}_h^0\|$
150	1.011e-1	2.118e-2	3.542e-2	2.948e-2
378	7.812e-2	2.440e-2	2.110e-2	5.388e-2
865	5.099e-1	1.357e-2	9.894-3	3.392e-2
1889	3.567e-2	9.904e-3	5.728e-3	2.252e-2
4031	2.457e-2	6.179e-3	2.677e-3	1.533e-2
8476	1.682e-2	4.460e-3	1.647e-3	9.946e-3
order	0.986	1.087	1.663	1.083

Table 2.4: RT_1 with bisection refinement

nt	$\|\boldsymbol{p}-\boldsymbol{p}_h^1\|$	$\|\Pi_h^1\boldsymbol{p}-\boldsymbol{p}_h^1\|$	$\|\boldsymbol{p}-R_h^1\boldsymbol{p}_h^1\|$
150	2.161e-3	4.740e-4	9.774e-4
378	1.252e-3	4.881e-4	5.140e-4
865	5.440e-4	2.347e-4	1.688e-4
1889	2.694e-4	9.748e-5	6.048e-5
4031	1.252e-4	4.857e-5	1.889e-5
8476	6.123e-5	2.110e-5	7.793e-6
order	1.942	2.033	2.735

Table 2.5: BDM_1 on uniformly parallel grids with regular refinement

nt	$\|\boldsymbol{p}-\tilde{\boldsymbol{p}}_h\|$	$\|\tilde{\boldsymbol{p}}_I-\tilde{\boldsymbol{p}}_h\|$	$\|\boldsymbol{p}-\tilde{R}_h\tilde{\boldsymbol{p}}_I\|$	$\|\boldsymbol{p}-\tilde{R}_h\tilde{\boldsymbol{p}}_h\|$
64	8.541e-03	1.195e-02	2.492e-03	1.042e-02
256	2.171e-03	2.958e-03	2.422e-04	2.779e-03
1024	5.465e-04	7.350e-04	2.240e-05	7.140e-04
4096	1.370e-04	1.832e-04	2.053e-06	1.807e-04
16384	3.430e-05	4.570e-05	1.910e-07	4.540e-05
order	2.020	2.037	3.474	1.992

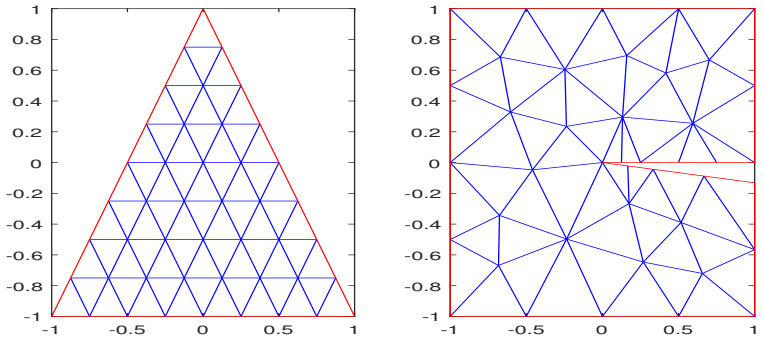

Figure 2.4: (left)uniformly parallel triangulation (right)adaptive initial grid

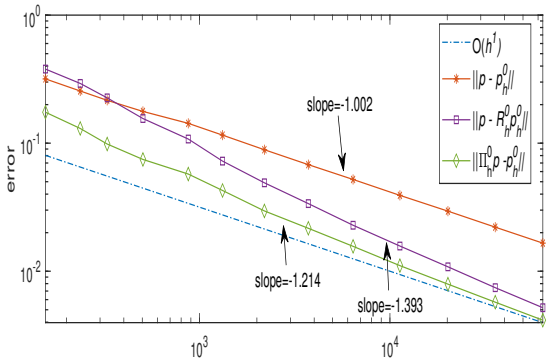

Figure 2.5: error curves for RT_0

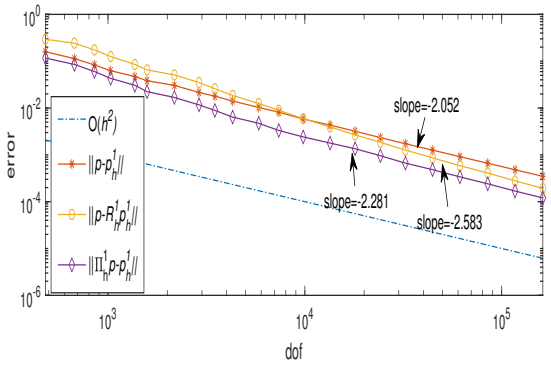

Figure 2.6: error curves for RT_1

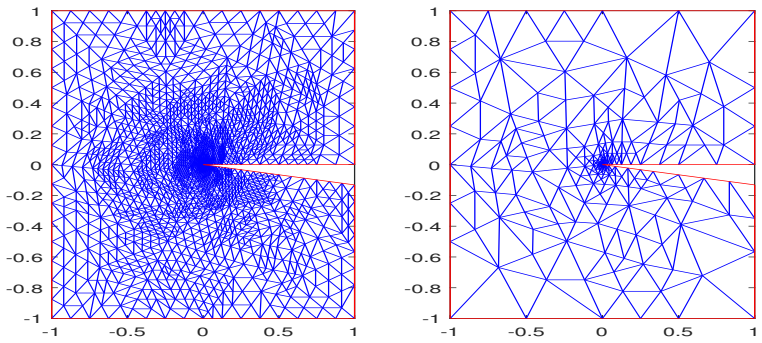

Figure 2.7: (left)graded grid for RT_0 (right)graded grid for RT_1

Chapter 3

Superconvergence of the Hellan–Herrmann-Johnson mixed method

3.1 Introduction

Since there is no simple finite element subspace of $H_0^2(\Omega)$ in the primal formulation (3.2), the fourth order elliptic equation (3.1) is often discretized by the nonconforming Morley element method, or the mixed Hellan–Herrmann–Johnson(HHJ) method. In the HHJ method, the stress tensor $\boldsymbol{\sigma} = \mathcal{A}\nabla^2 u$ and the displacement u in (3.4) are approximated by the finite element solutions $\boldsymbol{\sigma}_h$ and u_h, respectively. It can be shown using a duality argument that $I_h u$ and u_h are superclose in the H^1-norm on unstructured grids and thus a postprocessing technique can be applied to achieve superconvergence for the displacement variable, cf. [3, 23]. It is often the case that the stress tensor $\boldsymbol{\sigma}$, or simply the Hessian of u, is of more practical interest than u itself. For example, the adaptive version of Morley and HHJ methods are based on a

posteriori error estimation of $\boldsymbol{\sigma}$ instead of the H^1- or L^2-norm of $u - u_h$, cf. [41, 39]. However, superconvergence analysis of the tensor variable is mesh-dependent and more involved. Up to now, there has been little work for superconvergence of stress tensors in mixed methods, except for [36], which presents recovery superconvergence results on $\boldsymbol{\sigma} = \nabla^2 u$ for the lowest order HHJ mixed method on uniformly parallel grids.

In this chapter, we shall prove Theorem 3.2.5, a supercloseness estimate for $\boldsymbol{\sigma}$ in the lowest order HHJ mixed method under mildly structured grids. The proof is constructive and different from the proof in [36]. Our supercloseness result shows that the rate of superconvergence proved in [36] on uniform grids is suboptimal (an extra $h^{\frac{1}{2}}$ is obtained in Theorem 3.2.5). The main tool for superconvergence analysis is Lemma 3.2.1 about the local interpolation error expansion. Global supercloseness then follows from the framework for superconvergence analysis of RT mixed methods in Chapter 2. The additional difficulties arise from the symmetry of the stress tensor in the HHJ method and the double divergence operator div Div.

The second major component of this chapter is the new recovery operator R_h based on superconvergent patch recovery technique, compared with R_h^r for the RT_r element in Chapter 2. We shall show that R_h has a nice super-approximation property and thus achieves recovery superconvergence using the aforementioned supercloseness result. The analysis is still based on Lemma 3.2.1. Meanwhile, $\|\boldsymbol{\sigma}_h - R_h\boldsymbol{\sigma}_h\|_{0,T}$ can serve as a posteriori error estimators. In numerical experiments, it is shown to be asymptotically exact. The asymptotic exactness of recovery-based error estimators is often attributed to superconvergence.

Preliminaries are introduced in the rest of this section. Throughout this chapter, variables in boldface stand for vectors or matrices. A vector is viewed as a column by default unless confusion arises. For a scalar-valued function v and vector/matrix-valued

functions

$$\boldsymbol{\phi} = \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix}, \quad \boldsymbol{\tau} = \begin{pmatrix} \boldsymbol{\tau}_1 \\ \boldsymbol{\tau}_2 \end{pmatrix},$$

several differential operators are defined as

$$\begin{aligned} \mathrm{Curl}\boldsymbol{\phi} &= (\nabla^\perp \phi_1, \nabla^\perp \phi_2)^\intercal, \quad \mathrm{Rot}\boldsymbol{\tau} = (\nabla \times \boldsymbol{\tau}_1, \nabla \times \boldsymbol{\tau}_2)^\intercal, \\ \mathrm{Grad}\boldsymbol{\phi} &= (\nabla \phi_1, \nabla \phi_2)^\intercal, \quad \mathrm{Div}\boldsymbol{\tau} = (\operatorname{div} \boldsymbol{\tau}_1, \operatorname{div} \boldsymbol{\tau}_2)^\intercal. \end{aligned}$$

Symmetrized gradient and curl are defined as

$$\mathrm{Grad}^s(\boldsymbol{\phi}) = \frac{1}{2}(\mathrm{Grad}\boldsymbol{\phi} + (\mathrm{Grad}\boldsymbol{\phi})^\intercal), \quad \mathrm{Curl}^s\boldsymbol{\phi} = \frac{1}{2}(\mathrm{Curl}\boldsymbol{\phi} + (\mathrm{Curl}\boldsymbol{\phi})^\intercal),$$

respectively. The linear operator $\mathcal{A}$ is defined to be

$$\mathcal{A}\boldsymbol{\tau} := \frac{E}{12(1-\nu^2)}((1-\nu)\boldsymbol{\tau} + \nu \operatorname{tr}(\boldsymbol{\tau})\boldsymbol{I}),$$

where $\boldsymbol{\tau}$ a symmetric 2×2 matrix, $\boldsymbol{I}$ is the 2×2 identity matrix, $\operatorname{tr}(\boldsymbol{\tau})$ is the trace of $\boldsymbol{\tau}$, E is the Young's modulus and $0 \le \nu < 1$ is a constant. Without loss of generality, we set $E/12 = 1$. Let ∇^2 denote Hessian. We consider the boundary value problem

$$\operatorname{div} \mathrm{Div}\, \mathcal{A}\nabla^2 u = f \quad \text{in } \Omega, \tag{3.1a}$$

$$u = \partial_{\boldsymbol{n}} u = 0 \quad \text{on } \partial\Omega. \tag{3.1b}$$

Let : denote the component wise product for matrices, namely, $\boldsymbol{A} : \boldsymbol{B} = \sum_{i=1}^m \sum_{j=1}^n a_{ij} b_{ij}$, where $\boldsymbol{A} = (a_{ij})_{1 \le i \le m, 1 \le j \le n}, \boldsymbol{B} = (b_{ij})_{1 \le i \le m, 1 \le j \le n}$. Let $(\cdot, \cdot)$ denote the $L^2(\Omega)$-inner product: $(\boldsymbol{A}, \boldsymbol{B}) = \int_\Omega \boldsymbol{A} : \boldsymbol{B}$. The primal formulation for (3.1) is to find $u \in H_0^2(\Omega)$, such that

$$(\mathcal{A}\nabla^2 u, \nabla^2 v) = (f, v), \quad \text{for all } v \in H_0^2(\Omega). \tag{3.2}$$

Here $H_0^2(\Omega) = \{v \in H^2(\Omega) : v = \partial_{\boldsymbol{n}} v = 0 \text{ on } \partial\Omega\}$. (3.2) is the model variational formulation for Kirchhoff plate theory, cf. [16, 41, 39].

For a symmetric 2×2 matrix $\boldsymbol{\tau}$, two scalars $\boldsymbol{\tau}_{nn} := \boldsymbol{n}^\intercal \boldsymbol{\tau} \boldsymbol{n}, \boldsymbol{\tau}_{nt} := \boldsymbol{n}^\intercal \boldsymbol{\tau} \boldsymbol{t}$ are defined on ∂T, where $\boldsymbol{n}$ and $\boldsymbol{t}$ are unit normal and tangent to ∂T, respectively. For a space V,

$$
\begin{aligned}
[V]_s^4 &:= \left\{ v = \begin{pmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{pmatrix} : v_{12} = v_{21}, v_{ij} \in V,\ i,j = 1,2 \right\}, \\
[V]^n &:= \{v = (v_1, \ldots, v_n)^\intercal : v_i \in V, 1 \le i \le n\}.
\end{aligned}
\tag{3.3}
$$

Recall that $\mathcal{U}_h = \mathcal{S}_h^1 \cap H_0^1(\Omega)$ and define

$$
\begin{aligned}
\boldsymbol{\Sigma}_h = \{ &\boldsymbol{\tau}_h \in [L^2(\Omega)]_s^4 : \boldsymbol{\tau}_h|_T \in [\mathcal{P}_0(T)]_s^4,\ \forall\, T \in \mathcal{T}_h, \\
&(\boldsymbol{\tau}_h)_{nn} \text{ is single-valued on each } e \in \mathcal{E}_h^o\}.
\end{aligned}
$$

Equation (3.1) is equivalent to the second order system

$$
\boldsymbol{\sigma} = \mathcal{A} \nabla^2 u, \tag{3.4a}
$$

$$
\operatorname{div} \operatorname{Div} \boldsymbol{\sigma} = f. \tag{3.4b}
$$

It follows from direct calculation that

$$
\mathcal{C}\boldsymbol{\tau} := \mathcal{A}^{-1}\boldsymbol{\tau} = (1+\nu)\boldsymbol{\tau} - \nu \operatorname{tr}(\boldsymbol{\tau}) \boldsymbol{I}.
$$

By (3.4) and integrating by parts element by element, we have

$$
\begin{aligned}
(\mathcal{C}\boldsymbol{\sigma}, \boldsymbol{\tau}_h) + b_h(\boldsymbol{\tau}_h, u) &= 0, \\
b_h(\boldsymbol{\sigma}, v_h) &= -(f, v_h),
\end{aligned}
\tag{3.5}
$$

for all $(\boldsymbol{\tau}_h, v_h) \in \boldsymbol{\Sigma}_h \times \mathcal{U}_h$, where

$$\begin{aligned} b_h(\boldsymbol{\tau}, v) :&= \sum_{T\in\mathcal{T}_h} \int_T \operatorname{Div}\boldsymbol{\tau}\cdot\nabla v - \int_{\partial T} \boldsymbol{\tau}_{nt}\partial_{\boldsymbol{t}} v \\ &= \sum_{T\in\mathcal{T}_h} \int_T -\boldsymbol{\tau} : \nabla^2 v + \int_{\partial T} \boldsymbol{\tau}_{nn}\partial_{\boldsymbol{n}} v. \end{aligned} \tag{3.6}$$

Here we use the tensor version integration by parts

$$\int_T \operatorname{Div}\boldsymbol{\tau}\cdot\nabla v = \int_{\partial T} (\boldsymbol{\tau}\boldsymbol{n})\cdot\nabla v - \int_T \boldsymbol{\tau} : \nabla^2 v, \tag{3.7}$$

where $v \in H^2(T)$ and each component of $\boldsymbol{\tau}$ is in $H^1(T)$. Another useful formula is

$$\int_T \operatorname{Curl}\boldsymbol{\varphi} : \boldsymbol{\tau} = \int_{\partial T} \boldsymbol{\varphi}\cdot\boldsymbol{\tau}\boldsymbol{t} - \int_T \boldsymbol{\varphi}\cdot\operatorname{Rot}\boldsymbol{\tau}, \tag{3.8}$$

for the same $\boldsymbol{\tau}$ and $\boldsymbol{\varphi} \in [H^1(T)]^2$. The HHJ mixed method for solving (3.4) is to find $(\boldsymbol{\sigma}_h, u_h) \in \boldsymbol{\Sigma}_h \times \mathcal{U}_h$, such that

$$(\mathcal{C}\boldsymbol{\sigma}_h, \boldsymbol{\tau}_h) + b_h(\boldsymbol{\tau}_h, u_h) = 0, \tag{3.9a}$$

$$b_h(\boldsymbol{\sigma}_h, v_h) = -(f, v_h), \tag{3.9b}$$

for all $(\boldsymbol{\tau}_h, v_h) \in \boldsymbol{\Sigma}_h \times \mathcal{U}_h$. For $\boldsymbol{\tau} \in [H^1(\Omega)]^4_s$, $\pi_h\boldsymbol{\tau}$ is the unique element in $\boldsymbol{\Sigma}_h$ satisfying

$$\int_{\partial T} (\pi_h\boldsymbol{\tau})_{nn} = \int_{\partial T} \boldsymbol{\tau}_{nn} \text{ for all } T \in \mathcal{T}_h. \tag{3.10}$$

We also need the linear Lagrange interpolation I_h. It is readily checked that π_h, I_h and

b_h satisfy

$$b_h(\pi_h\boldsymbol{\tau}, v_h) = b_h(\boldsymbol{\tau}, v_h), \quad \text{for all } v_h \in \mathcal{U}_h, \tag{3.11a}$$

$$b_h(\boldsymbol{\tau}_h, I_h v) = b_h(\boldsymbol{\tau}_h, v), \quad \text{for all } \boldsymbol{\tau}_h \in \boldsymbol{\Sigma}_h. \tag{3.11b}$$

The rest of this chapter is organized as follows. Section 3.2 is devoted to estimating the global variational error and developing a supercloseness estimate, i.e., Theorem 3.2.5. In Section 3.3, we construct the recovery operator and prove Theorem 3.3.3, a recovery-type superconvergence estimate. Numerical experiments including both smooth problems under quasi-uniform grids and singular problems under graded meshes are reported in Section 3.4.

3.2 Supercloseness

First we expand the local interpolation error for linear tensor-valued functions. Let $\boldsymbol{Q}$ be the matrix representing rotation by $\frac{\pi}{2}$ counterclockwise and $\boldsymbol{d}^{\perp} = \boldsymbol{Q}\boldsymbol{d}$ for $\boldsymbol{d} \in \mathbb{R}^2$.

Lemma 3.2.1. *For* $\boldsymbol{\sigma}_1 \in [\mathcal{P}_1(T)]^4_s$,

$$\boldsymbol{\sigma}_1 - \pi_h\boldsymbol{\sigma}_1 = \mathrm{Curl}^{\mathrm{s}}\, \boldsymbol{r}^{\boldsymbol{\sigma}_1},$$

where

$$\boldsymbol{r}^{\boldsymbol{\sigma}_1} = \sum_{k=1}^{3} \phi_k \boldsymbol{\gamma}_k^{\boldsymbol{\sigma}_1},$$

with

$$\boldsymbol{\gamma}_k^{\boldsymbol{\sigma_1}} \cdot \boldsymbol{n}_k = \frac{\ell_k^2}{2} \boldsymbol{n}_k^\intercal \partial_{\boldsymbol{t}_k} \boldsymbol{\sigma}_1 \boldsymbol{n}_k, \tag{3.12a}$$

$$\boldsymbol{\gamma}_k^{\boldsymbol{\sigma_1}} \cdot \boldsymbol{t}_k = \frac{\ell_k^2}{2} \boldsymbol{n}_k^\intercal \partial_{\boldsymbol{n}_k} \boldsymbol{\sigma}_1 \boldsymbol{n}_k + \ell_k^2 \boldsymbol{t}_k^\intercal \partial_{\boldsymbol{t}_k} \boldsymbol{\sigma}_1 \boldsymbol{n}_k. \tag{3.12b}$$

Proof. First note that

$$\begin{aligned} &\int_{e_k} \boldsymbol{n}_k^\intercal \left(\operatorname{Curl}^{\mathrm{s}} \begin{pmatrix} \phi_k \\ 0 \end{pmatrix} \right) \boldsymbol{n}_k \\ &= \int_{e_k} \boldsymbol{n}_k^\intercal \left(\operatorname{Curl}^{\mathrm{s}} \begin{pmatrix} 0 \\ \phi_k \end{pmatrix} \right) \boldsymbol{n}_k = 0, \quad 1 \le k \le 3. \end{aligned} \tag{3.13}$$

Now assume that

$$\left\{ \operatorname{Curl}^{\mathrm{s}} \begin{pmatrix} \phi_i \\ 0 \end{pmatrix} \right\}_{i=1}^{3} \bigcup \left\{ \operatorname{Curl}^{\mathrm{s}} \begin{pmatrix} 0 \\ \phi_i \end{pmatrix} \right\}_{i=1}^{3} \tag{3.14}$$

are linearly independent. The reason is given at the end of the proof. Since $\int_{e_k} (\boldsymbol{\sigma}_1 - \pi_h \boldsymbol{\sigma}_1)_{nn} = 0$ for $1 \le k \le 3$, by counting the dimension, we have

$$\boldsymbol{\sigma}_1 - \pi_h \boldsymbol{\sigma}_1 = \sum_{i=1}^{3} \alpha_i \operatorname{Curl}^{\mathrm{s}} \begin{pmatrix} \phi_i \\ 0 \end{pmatrix} + \sum_{i=1}^{3} \beta_i \operatorname{Curl}^{\mathrm{s}} \begin{pmatrix} 0 \\ \phi_i \end{pmatrix} = \operatorname{Curl}^{\mathrm{s}} \boldsymbol{r}^{\boldsymbol{\sigma}_1}, \tag{3.15}$$

where $\boldsymbol{r}^{\boldsymbol{\sigma}_1} = \sum_{i=1}^{3} \phi_i \boldsymbol{\gamma}_i^{\boldsymbol{\sigma_1}}$ with $\boldsymbol{\gamma}_i^{\boldsymbol{\sigma_1}} := (\alpha_i, \beta_i)^\intercal$, and α_i, β_i are some undetermined constants. Given two unit vectors $\boldsymbol{d}_1$ and $\boldsymbol{d}_2$, it follows from (3.15) that

$$\boldsymbol{d}_1^\intercal (\boldsymbol{\sigma}_1 - \pi_h \boldsymbol{\sigma}_1) \boldsymbol{d}_2 = -\frac{1}{2} \sum_{i=1}^{3} \left(\boldsymbol{d}_1^\intercal \boldsymbol{\gamma}_i^{\boldsymbol{\sigma_1}} \frac{\partial \phi_i}{\partial \boldsymbol{d}_2^\perp} + \frac{\partial \phi_i}{\partial \boldsymbol{d}_1^\perp} \boldsymbol{\gamma}_i^{\boldsymbol{\sigma_1}\intercal} \boldsymbol{d}_2 \right). \tag{3.16}$$

First, applying $\partial_{\boldsymbol{t}_k}$ to (3.16) with $\boldsymbol{d}_1 = \boldsymbol{d}_2 = \boldsymbol{n}_k$, and using

$$\partial^2_{\boldsymbol{t}_k}\phi_i = -2\delta_{ki}/\ell_k^2, \tag{3.17}$$

where δ_{ki} is the Kronecker delta, we obtain (3.12a), the normal component of $\boldsymbol{\gamma}_i^{\boldsymbol{\sigma_1}}$. Second, by applying $\partial_{\boldsymbol{t}_k}$ to (3.16) with $\boldsymbol{d}_1 = \boldsymbol{t}_k, \boldsymbol{d}_2 = \boldsymbol{n}_k$ and using (3.17), we have

$$\boldsymbol{t}_k^\intercal \partial_{\boldsymbol{t}_k}\boldsymbol{\sigma}_1\boldsymbol{n}_k = \frac{1}{\ell_k^2}\boldsymbol{t}_k^\intercal\boldsymbol{\gamma}_k^{\boldsymbol{\sigma_1}} + \frac{1}{2}\sum_{i=1}^{3}\partial^2_{\boldsymbol{t}_k\boldsymbol{n}_k}\phi_i\boldsymbol{\gamma}_i^{\boldsymbol{\sigma_1}\intercal}\boldsymbol{n}_k. \tag{3.18}$$

Third, applying $\partial_{\boldsymbol{n}_k}$ to (3.16) with $\boldsymbol{d}_1 = \boldsymbol{d}_2 = \boldsymbol{n}_k$ leads to

$$\boldsymbol{n}_k^\intercal \partial_{\boldsymbol{n}_k}\boldsymbol{\sigma}_1\boldsymbol{n}_k = -\sum_{i=1}^{3}\boldsymbol{\gamma}_i^{\boldsymbol{\sigma_1}\intercal}\boldsymbol{n}_k\partial^2_{\boldsymbol{t}_k\boldsymbol{n}_k}\phi_i. \tag{3.19}$$

Comparing (3.19) with (3.18), we obtain (3.12b), the tangential component of $\boldsymbol{\gamma}_i^{\boldsymbol{\sigma_1}}$. As for the linear independence of (3.14), let

$$\sum_{i=1}^{3}\alpha_i'\,\mathrm{Curl}^{\mathrm{s}}\begin{pmatrix}\phi_i\\ 0\end{pmatrix} + \sum_{i=1}^{3}\beta_i'\,\mathrm{Curl}^{\mathrm{s}}\begin{pmatrix}0\\ \phi_i\end{pmatrix} = 0.$$

Then by using the same argument below (3.15), we obtain that both the normal and tangential components of (α_i', β_i') are zero, i.e., $\alpha_i' = \beta_i' = 0$. □

Any vector $\boldsymbol{\xi}_k \in \mathbb{R}^2$ is a linear combination of $\boldsymbol{t}_{k-1}$ and $\boldsymbol{t}_{k+1}$. Similar to (2.20),

$$\boldsymbol{\xi}_k = -\frac{\boldsymbol{\xi}_k\cdot\boldsymbol{n}_{k+1}}{\sin\theta_k}\boldsymbol{t}_{k-1} + \frac{\boldsymbol{\xi}_k\cdot\boldsymbol{n}_{k-1}}{\sin\theta_k}\boldsymbol{t}_{k+1}, \quad 1 \le k \le 3.$$

Combining it with $\ell_1/\sin\theta_1 = \ell_2/\sin\theta_2 = \ell_3/\sin\theta_3$, we obtain

$$\begin{aligned}\sum_{k=1}^{3}\ell_k\nabla v\cdot\boldsymbol{\xi}_k &= \sum_{k=1}^{3}\frac{\ell_k}{\sin\theta_k}\boldsymbol{\xi}_k\cdot\boldsymbol{n}_{k-1}\partial_{\boldsymbol{t}_{k+1}}v - \sum_{k=1}^{3}\frac{\ell_k}{\sin\theta_k}\boldsymbol{\xi}_k\cdot\boldsymbol{n}_{k+1}\partial_{\boldsymbol{t}_{k-1}}v\\ &= \sum_{k=1}^{3}\frac{\ell_k}{\sin\theta_k}\big(\boldsymbol{\xi}_{k-1}\cdot\boldsymbol{n}_{k+1} - \boldsymbol{\xi}_{k+1}\cdot\boldsymbol{n}_{k-1}\big)\partial_{\boldsymbol{t}_k}v.\end{aligned} \tag{3.20}$$

The identity (3.20) is essentially first proved in [9] but the proof here is simplified, compared also with the proof of Lemma 2.2.2. Setting $\boldsymbol{\xi}_k = \delta_k\boldsymbol{n}_k$ for $1\le k\le 3$ in (3.20) and using $\boldsymbol{n}_{k-1}\cdot\boldsymbol{n}_{k+1} = -\cos\theta_k$, we have

$$\sum_{k=1}^{3}\ell_k\delta_k\nabla^{\perp}v\cdot\boldsymbol{t}_k = \sum_{k=1}^{3}\ell_k\cot\theta_k\left(\delta_{k+1}-\delta_{k-1}\right)\partial_{\boldsymbol{t}_k}v.$$

Let $\boldsymbol{\varphi} = (\varphi_1,\varphi_2)^{\intercal}$ be a $\mathbb{R}^2$-valued function and $\boldsymbol{\delta}_k = (\delta_{k,1},\delta_{k,2})^{\intercal}\in\mathbb{R}^2$. Setting $v=\varphi_1$, $\delta_k = \delta_{k,1}$ or $v=\varphi_2$, $\delta_k=\delta_{k,2}$ for $1\le k\le 3$ in the above identity, we obtain

$$\sum_{k=1}^{3}\ell_k\boldsymbol{\delta}_k^{\intercal}(\operatorname{Curl}\boldsymbol{\varphi})\boldsymbol{t}_k = \sum_{k=1}^{3}\ell_k\cot\theta_k(\boldsymbol{\delta}_{k+1}-\boldsymbol{\delta}_{k-1})^{\intercal}\partial_{\boldsymbol{t}_k}\boldsymbol{\varphi}. \tag{3.21}$$

The next lemma is our main technical tool for estimating the global variational error in the HHJ method.

Lemma 3.2.2. *For $\boldsymbol{\sigma}_1\in[\mathcal{P}_1(T)]^4_s$ and $\boldsymbol{\varphi}\in[\mathcal{P}_1(T)]^2$,*

$$\int_T\mathcal{C}(\boldsymbol{\sigma}_1-\pi_h\boldsymbol{\sigma}_1):\operatorname{Curl}^{\mathrm{s}}\boldsymbol{\varphi} = \sum_{k=1}^{3}\int_{e_k}\boldsymbol{g}_k^{\boldsymbol{\sigma}_1}\partial_{\boldsymbol{t}_k}\boldsymbol{\varphi}, \tag{3.22}$$

where

$$\begin{aligned}\boldsymbol{g}_k^{\boldsymbol{\sigma}_1} = &-\frac{1}{12}(1+\nu)\{(\boldsymbol{\gamma}_{k-1}^{\boldsymbol{\sigma}_1}-\boldsymbol{\gamma}_{k+1}^{\boldsymbol{\sigma}_1})^{\intercal}\cot\theta_k + \boldsymbol{\gamma}_k^{\boldsymbol{\sigma}_1\intercal}\boldsymbol{n}_k\boldsymbol{t}_k^{\intercal}\\ &+(\boldsymbol{\gamma}_{k-1}^{\boldsymbol{\sigma}_1\intercal}\boldsymbol{t}_{k-1}\boldsymbol{t}_{k-1}^{\intercal}-\boldsymbol{\gamma}_{k+1}^{\boldsymbol{\sigma}_1\intercal}\boldsymbol{t}_{k+1}\boldsymbol{t}_{k+1}^{\intercal})\cot\theta_k\}\\ &+\frac{\nu}{6}\frac{1}{\sin\theta_k}(\boldsymbol{\gamma}_{k-1}^{\boldsymbol{\sigma}_1\intercal}\boldsymbol{t}_{k-1}\boldsymbol{n}_{k+1}^{\intercal}-\boldsymbol{\gamma}_{k+1}^{\boldsymbol{\sigma}_1\intercal}\boldsymbol{t}_{k+1}\boldsymbol{n}_{k-1}^{\intercal})\boldsymbol{Q},\end{aligned} \tag{3.23}$$

and $\{\boldsymbol{\gamma}_k^{\boldsymbol{\sigma}_1}\}_{k=1}^3$ *are defined in* (3.12).

Proof. Let $\boldsymbol{\tau} = \mathrm{Curl}^{\mathrm{s}}\boldsymbol{\varphi}$. By the self-adjointness of $\mathcal{C}$, the symmetry of $\mathcal{C}\boldsymbol{\tau}$, the integration-by-parts formula (3.8) and Lemma 3.2.1, we have

$$\begin{aligned}\int_T \mathcal{C}(\boldsymbol{\sigma}_1 - \pi_h\boldsymbol{\sigma}_1) : \boldsymbol{\tau} &= \int_T (\boldsymbol{\sigma}_1 - \pi_h\boldsymbol{\sigma}_1) : \mathcal{C}\boldsymbol{\tau} = \int_T \mathrm{Curl}^{\mathrm{s}}\boldsymbol{r}^{\boldsymbol{\sigma}_1} : \mathcal{C}\boldsymbol{\tau} \\ &= \int_T \mathrm{Curl}\,\boldsymbol{r}^{\boldsymbol{\sigma}_1} : \mathcal{C}\boldsymbol{\tau} = \sum_{k=1}^3 \int_{e_k} \phi_k \boldsymbol{\gamma}_k^{\boldsymbol{\sigma}_1\intercal}(\mathcal{C}\boldsymbol{\tau})\boldsymbol{t}_k = \frac{1}{6}\sum_{k=1}^3 \ell_k \boldsymbol{\gamma}_k^{\boldsymbol{\sigma}_1\intercal}(\mathcal{C}\boldsymbol{\tau})\boldsymbol{t}_k.\end{aligned}$$

In the last step, we use Simpson's quadrature rule. It follows that

$$\int_T \mathcal{C}(\boldsymbol{\sigma}_1 - \pi_h\boldsymbol{\sigma}_1) : \boldsymbol{\tau} = \frac{1}{6}(1+\nu)I - \frac{1}{6}\nu II, \tag{3.24}$$

where

$$I = \sum_{k=1}^3 \ell_k \boldsymbol{\gamma}_k^{\boldsymbol{\sigma}_1\intercal}\boldsymbol{\tau}\boldsymbol{t}_k, \quad II = \sum_{k=1}^3 \ell_k \boldsymbol{\gamma}_k^{\boldsymbol{\sigma}_1\intercal}\boldsymbol{t}_k \,\mathrm{tr}(\boldsymbol{\tau}).$$

By $\boldsymbol{\tau} = \{\mathrm{Curl}\,\boldsymbol{\varphi} + (\mathrm{Curl}\,\boldsymbol{\varphi})^\intercal\}/2$ and the identity (3.21),

$$\begin{aligned}I &= \frac{1}{2}\sum_{k=1}^3 \ell_k \cot\theta_k (\boldsymbol{\gamma}_{k+1}^{\boldsymbol{\sigma}_1} - \boldsymbol{\gamma}_{k-1}^{\boldsymbol{\sigma}_1})^\intercal \partial_{\boldsymbol{t}_k}\boldsymbol{\varphi} \\ &\quad + \frac{1}{2}\sum_{k=1}^3 \ell_k \boldsymbol{\gamma}_k^{\boldsymbol{\sigma}_1\intercal}(\mathrm{Curl}\,\boldsymbol{\varphi})^\intercal \boldsymbol{t}_k = I_1 + I_2.\end{aligned} \tag{3.25}$$

By writing $\boldsymbol{\gamma}_k^{\boldsymbol{\sigma}_1\intercal} = \boldsymbol{\gamma}_k^{\boldsymbol{\sigma}_1\intercal}\boldsymbol{t}_k\boldsymbol{t}_k^\intercal + \boldsymbol{\gamma}_k^{\boldsymbol{\sigma}_1\intercal}\boldsymbol{n}_k\boldsymbol{n}_k^\intercal$ and using (3.21),

$$\begin{aligned}I_2 &= \frac{1}{2}\sum_{k=1}^3 \ell_k \boldsymbol{\gamma}_k^{\boldsymbol{\sigma}_1\intercal}\boldsymbol{t}_k\boldsymbol{t}_k^\intercal(\mathrm{Curl}\,\boldsymbol{\varphi})\boldsymbol{t}_k + \frac{1}{2}\sum_{k=1}^3 \ell_k \boldsymbol{\gamma}_k^{\boldsymbol{\sigma}_1\intercal}\boldsymbol{n}_k\boldsymbol{t}_k^\intercal(\mathrm{Curl}\,\boldsymbol{\varphi})\boldsymbol{n}_k \\ &= \frac{1}{2}\sum_{k=1}^3 \ell_k \cot\theta_k (\boldsymbol{\gamma}_{k+1}^{\boldsymbol{\sigma}_1\intercal}\boldsymbol{t}_{k+1}\boldsymbol{t}_{k+1}^\intercal - \boldsymbol{\gamma}_{k-1}^{\boldsymbol{\sigma}_1\intercal}\boldsymbol{t}_{k-1}\boldsymbol{t}_{k-1}^\intercal)\partial_{\boldsymbol{t}_k}\boldsymbol{\varphi} \\ &\quad - \frac{1}{2}\sum_{k=1}^3 \ell_k \boldsymbol{\gamma}_k^{\boldsymbol{\sigma}_1\intercal}\boldsymbol{n}_k\boldsymbol{t}_k^\intercal\partial_{\boldsymbol{t}_k}\boldsymbol{\varphi}.\end{aligned} \tag{3.26}$$

Let $\boldsymbol{\varphi} = (\varphi_1, \varphi_2)^\intercal$. By the identity (3.20) with $\boldsymbol{\xi}_k = (0,-1)^\intercal(\boldsymbol{\gamma}_k^{\boldsymbol{\sigma}_1\intercal}\boldsymbol{t}_k), v = \varphi_1$ and $\boldsymbol{\xi}_k = (1,0)^\intercal(\boldsymbol{\gamma}_k^{\boldsymbol{\sigma}_1\intercal}\boldsymbol{t}_k), v = \varphi_2$, respectively, we have

$$\begin{aligned} II &= \sum_{k=1}^{3} \ell_k \boldsymbol{\gamma}_k^{\boldsymbol{\sigma}_1\intercal}\boldsymbol{t}_k(-\partial_{x_2}\varphi_1 + \partial_{x_1}\varphi_2) \\ &= \sum_{k=1}^{3} \ell_k (p_k^{\boldsymbol{\sigma}_1}\partial_{\boldsymbol{t}_k}\varphi_1 + q_k^{\boldsymbol{\sigma}_1}\partial_{\boldsymbol{t}_k}\varphi_2), \end{aligned} \tag{3.27}$$

where

$$\begin{aligned} p_k^{\boldsymbol{\sigma}_1} &= \frac{1}{\sin\theta_k}\{\boldsymbol{n}_{k+1}^\intercal(0,-1)^\intercal(\boldsymbol{\gamma}_{k-1}^{\boldsymbol{\sigma}_1\intercal}\boldsymbol{t}_{k-1}) - \boldsymbol{n}_{k-1}^\intercal(0,-1)^\intercal(\boldsymbol{\gamma}_{k+1}^{\boldsymbol{\sigma}_1\intercal}\boldsymbol{t}_{k+1})\}, \\ q_k^{\boldsymbol{\sigma}_1} &= \frac{1}{\sin\theta_k}\{\boldsymbol{n}_{k+1}^\intercal(1,0)^\intercal(\boldsymbol{\gamma}_{k-1}^{\boldsymbol{\sigma}_1\intercal}\boldsymbol{t}_{k-1}) - \boldsymbol{n}_{k-1}^\intercal(1,0)^\intercal(\boldsymbol{\gamma}_{k+1}^{\boldsymbol{\sigma}_1\intercal}\boldsymbol{t}_{k+1})\}. \end{aligned}$$

Combining (3.24)–(3.27), we obtain (3.21). □

We expand the left hand side of (3.22) in terms of the tangential derivative $\partial_{\boldsymbol{t}_k}\boldsymbol{\varphi}$ which is single-valued over the inter-element boundary provided $\boldsymbol{\varphi}_h \in [\mathcal{S}_h^1]^2$. This fact is crucial for error cancellation, see the proof of Theorem 3.2.5. Note that $\boldsymbol{g}_k^{\boldsymbol{\sigma}_1}$ is linear in $\boldsymbol{\sigma}_1$. By passing $\boldsymbol{\sigma}$ to $I_h\boldsymbol{\sigma}$ and vice versa, we obtain a corollary for general $\boldsymbol{\sigma}$. The proof is skipped because it is the same as Lemma 2.2.3.

Lemma 3.2.3. *For $\boldsymbol{\varphi} \in [\mathcal{P}_1(T)]^2$,*

$$\int_T \mathcal{C}(\boldsymbol{\sigma} - \pi_T\boldsymbol{\sigma}) : \operatorname{Curl}^{\mathrm{s}}\boldsymbol{\varphi} = \sum_{k=1}^{3}\int_{e_k} \boldsymbol{g}_k^{\boldsymbol{\sigma}}\partial_{\boldsymbol{t}_k}\boldsymbol{\varphi} + O(h^2)|\boldsymbol{\sigma}|_{2,T}|\boldsymbol{\varphi}|_{1,T}.$$

Although $\boldsymbol{g}_k^{\boldsymbol{\sigma}}$ in Corollary 3.2.3 is complicated, $\boldsymbol{g}_k^{\boldsymbol{\sigma}}$ is clearly of magnitude $O(h^2)$. In addition, $|\boldsymbol{g}_k^{\boldsymbol{\sigma}} - \boldsymbol{g}_k^{\boldsymbol{\sigma}'}| \lesssim h^{2+\alpha}|D\boldsymbol{\sigma}|$ on e_k, where e_k is shared by T and T' which form an $O(h^{1+\alpha})$-approximate parallelogram, and $\boldsymbol{g}_k^{\boldsymbol{\sigma}'}$ is defined by the quantities on T', see the proof of Lemma 3.2.4. Now we present our main lemma.

Lemma 3.2.4. *For any $\boldsymbol{\varphi}_h \in [\mathcal{S}_h]^2$, let $\boldsymbol{\tau}_h = \operatorname{Curl}^{\mathrm{s}}\boldsymbol{\varphi}_h$. If $\mathcal{T}_h$ satisfies the (α,β)-*

condition, then

$$(\mathcal{C}(\boldsymbol{\sigma}-\pi_h\boldsymbol{\sigma}),\boldsymbol{\tau}_h)\lesssim h^{1+\min(\frac{1}{2},\alpha,\frac{\beta}{2})}(|\boldsymbol{\sigma}|_{1,\infty,\Omega}+\|\boldsymbol{\sigma}\|_{2,\Omega})\|\boldsymbol{\tau}_h\|_{0,\Omega}. \tag{3.28}$$

If $\mathcal{T}_h$ satisfies the piecewise strong (α,β)-condition, then

$$(\mathcal{C}(\boldsymbol{\sigma}-\pi_h\boldsymbol{\sigma}),\boldsymbol{\tau}_h)\lesssim h^{1+\min(1,\alpha,\frac{\beta}{2})}|\log h|^{\frac{1}{2}}\|\boldsymbol{\sigma}\|_{2,\infty,\Omega}\|\boldsymbol{\tau}_h\|_{0,\Omega}. \tag{3.29}$$

Proof. First assume that $\mathcal{T}_h$ satisfies the strong (α,β)-condition. By a Korn's inequality in quotient space (cf. Theorems 2.3 and 2.2 in [20]),

$$\|\boldsymbol{\varphi}_h^{\perp}+\boldsymbol{r}_1\|_{1,\Omega}\lesssim\|\mathrm{Grad}^s(\boldsymbol{\varphi}_h^{\perp})\|_{0,\Omega} \tag{3.30}$$

for some $\boldsymbol{r}_1=a\boldsymbol{x}^{\perp}+\boldsymbol{b}\in\ker\mathrm{Grad}^s$, where $a\in\mathbb{R},\boldsymbol{b}\in\mathbb{R}^2$. Then using (3.30), $\mathrm{Grad}^s(\boldsymbol{\varphi}_h^{\perp})=\boldsymbol{Q}(\mathrm{Curl}^{\mathrm{s}}\,\boldsymbol{\varphi}_h)\boldsymbol{Q}^{\intercal}$, and setting $\hat{\boldsymbol{\varphi}}_h=\boldsymbol{\varphi}_h-\boldsymbol{r}_1^{\perp}$, we obtain $\mathrm{Curl}^{\mathrm{s}}\,\hat{\boldsymbol{\varphi}}_h=\mathrm{Curl}^{\mathrm{s}}\,\boldsymbol{\varphi}_h$ and

$$\|\hat{\boldsymbol{\varphi}}_h\|_{1,\Omega}\lesssim\|\,\mathrm{Curl}^{\mathrm{s}}\,\hat{\boldsymbol{\varphi}}_h\|_{0,\Omega}=\|\boldsymbol{\tau}_h\|_{0,\Omega}. \tag{3.31}$$

Let $g_{e_k}^{\boldsymbol{\sigma}}=g_k^{\boldsymbol{\sigma}}$ in Corollary 3.2.3. Note that $\hat{\boldsymbol{\varphi}}_h\in[\mathcal{S}_h]^2$ and thus $\partial_{\boldsymbol{t}_k}\hat{\boldsymbol{\varphi}}_h$ is single-valued on e_k. Hence by Corollary 3.2.3 and rearranging the sum, we have

$$\begin{aligned}(\mathcal{C}(\boldsymbol{\sigma}-\pi_h\boldsymbol{\sigma}),\boldsymbol{\tau}_h)&=(\mathcal{C}(\boldsymbol{\sigma}-\pi_h\boldsymbol{\sigma}),\mathrm{Curl}^{\mathrm{s}}\,\hat{\boldsymbol{\varphi}}_h)\\&=\sum_{T\in\mathcal{T}_h}\left\{\sum_{k=1}^{3}\int_{e_k}\boldsymbol{g}_{e_k}^{\boldsymbol{\sigma}}\partial_{\boldsymbol{t}_k}\hat{\boldsymbol{\varphi}}_h+O(h^2)|\boldsymbol{\sigma}|_{2,T}|\hat{\boldsymbol{\varphi}}_h|_{1,T}\right\}\\&=\sum_{i=1}^{4}I_i,\end{aligned} \tag{3.32}$$

where

$$I_j = \sum_{e\in\mathcal{E}_{h,j}^o}\int_e [\![\boldsymbol{g}_e^{\boldsymbol{\sigma}}]\!]\partial_{\boldsymbol{t}_e}\hat{\boldsymbol{\varphi}}_h, \quad j=1,2, \quad I_3 = \sum_{e\in\mathcal{E}_h^\partial}\int_e \boldsymbol{g}_e^{\boldsymbol{\sigma}}\partial_{\boldsymbol{t}_e}\hat{\boldsymbol{\varphi}}_h,$$

$$I_4 = \sum_{T\in\mathcal{T}_h} O(h^2)|\boldsymbol{\sigma}|_{2,T}|\hat{\boldsymbol{\varphi}}_h|_{1,T},$$

and $[\![\boldsymbol{g}_e^{\boldsymbol{\sigma}}]\!] := \boldsymbol{g}_e^{\boldsymbol{\sigma}} - \boldsymbol{g}_e^{\boldsymbol{\sigma}'}$ if $e \in \mathcal{E}_{h,j}^o$ is shared by T and T'. In this case, $\boldsymbol{t}_e$ is the unit tangent induced by T. The quantities associated with T' have a superscript $'$ and $\boldsymbol{g}_e^{\boldsymbol{\sigma}'}$ is computed based on the quantities in T'. In the rest of the proof, the telescoping inequality (2.26) for products is frequently used. For $e \in \mathcal{E}_{h,1}^o$, we claim that

$$|[\![\boldsymbol{g}_e^{\boldsymbol{\sigma}}]\!]| \lesssim h^{2+\alpha}|D\boldsymbol{\sigma}| \quad \text{on } e. \tag{3.33}$$

We can view e as e_1 and e_1' in T and T', respectively. Then $\boldsymbol{t}_1 = -\boldsymbol{t}_1'$, $\boldsymbol{n}_1 = -\boldsymbol{n}_1'$ etc. (3.33) can be seen by three steps. First since T and T' form an $O(h^{1+\alpha})$-approximate parallelogram,

$$\sum_{k=1}^{3}|\boldsymbol{t}_k + \boldsymbol{t}_k'| + |\boldsymbol{n}_k + \boldsymbol{n}_k'| = O(h^\alpha). \tag{3.34}$$

Second, by the definition (3.12) and using (2.26) and (3.34),

$$|\boldsymbol{\gamma}_k^{\boldsymbol{\sigma}} - \boldsymbol{\gamma}_k^{\boldsymbol{\sigma}'}| \lesssim h^{2+\alpha}|D\boldsymbol{\sigma}| \text{ on } e, \quad 1\le k\le 3. \tag{3.35}$$

Finally, (3.33) follows from (2.26), (3.34), (3.35), and the definition of $g_e^{\boldsymbol{\sigma}}$ in (3.23). Hence by (3.33), the trace inequality (1.9) and the Cauchy–Schwarz inequality, we have

$$\begin{aligned}
|I_1| &\lesssim \sum_{e\in\mathcal{E}_{h,1}^o} h^{2+\alpha}|\boldsymbol{\sigma}|_{1,e}|\hat{\boldsymbol{\varphi}}_h|_{1,e} \\
&\lesssim \sum_{e\in\mathcal{E}_{h,1}^o} h^{2+\alpha}\left(h^{-\frac{1}{2}}|\boldsymbol{\sigma}|_{1,T} + h^{\frac{1}{2}}|\boldsymbol{\sigma}|_{2,T}\right)\cdot h^{-\frac{1}{2}}|\hat{\boldsymbol{\varphi}}_h|_{1,T} \\
&\lesssim h^{1+\alpha}\|\boldsymbol{\sigma}\|_{2,\Omega}|\hat{\boldsymbol{\varphi}}_h|_{1,\Omega}.
\end{aligned} \tag{3.36}$$

Following the same argument in the proof of (2.30), (2.32), and (2.35), we obtain

$$|I_2| \lesssim h^{1+\frac{\beta}{2}} |\boldsymbol{\sigma}|_{1,\infty,\Omega} |\hat{\boldsymbol{\varphi}}_h|_{1,\Omega}, \tag{3.37}$$

and

$$|I_3| \lesssim h^{1+\alpha} |\log h|^{\frac{1}{2}} \|\boldsymbol{\sigma}\|_{2,\infty,\Omega} \|\hat{\boldsymbol{\varphi}}_h\|_{1,\Omega}. \tag{3.38}$$

In the end, the Cauchy–Schwarz inequality implies

$$|I_4| \lesssim h^2 |\boldsymbol{\sigma}|_{2,\Omega} |\hat{\boldsymbol{\varphi}}_h|_{1,\Omega}. \tag{3.39}$$

Combining (3.32), (3.36)–(3.39) and using (3.31) we obtain the estimate (3.29). Since the above argument is completely local, the estimate (3.29) also holds on $\mathcal{T}_h$ satisfying the piecewise strong (α, β)-condition. If I_3 is trivially estimated by

$$|I_3| \lesssim h^{1+\frac{1}{2}} |\boldsymbol{\sigma}|_{1,\infty,\Omega} |\hat{\boldsymbol{\varphi}}_h|_{1,\Omega}$$

as I_2, we obtain the estimate (3.28) on $\mathcal{T}_h$ satisfying the (α, β)-condition. □

Subtracting (3.9) from (3.5) gives the error equation

$$(\mathcal{C}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h), \boldsymbol{\tau}_h) + b_h(\boldsymbol{\tau}_h, u - u_h) = 0, \tag{3.40a}$$

$$b_h(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, v_h) = 0. \tag{3.40b}$$

With the help of Lemma 3.2.4 and (3.40), we can derive supercloseness estimates.

Theorem 3.2.5. *If $\mathcal{T}_h$ satisfies the (α, β)-condition, then*

$$\|\pi_h \boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{0,\Omega} \lesssim (1-\nu)^{-1} h^{1+\min(\frac{1}{2}, \alpha, \frac{\beta}{2})} (|\boldsymbol{\sigma}|_{1,\infty,\Omega} + \|\boldsymbol{\sigma}\|_{2,\Omega}).$$

If $\mathcal{T}_h$ satisfies the piecewise strong (α, β)-condition, then

$$\|\pi_h \boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{0,\Omega} \lesssim (1-\nu)^{-1} h^{1+\min(1,\alpha,\frac{\beta}{2})} |\log h|^{\frac{1}{2}} \|\boldsymbol{\sigma}\|_{2,\infty,\Omega}.$$

Proof. Let $\boldsymbol{\tau}_h = \pi_h \boldsymbol{\sigma} - \boldsymbol{\sigma}_h$. $(\operatorname{div}\operatorname{Div})_h \boldsymbol{\tau}_h$ is defined to be the function in $\mathcal{U}_h$ determined by

$$(v_h, (\operatorname{div}\operatorname{Div})_h \boldsymbol{\tau}_h) = b_h(\boldsymbol{\tau}_h, v_h) \quad \text{for all } v_h \in \mathcal{U}_h.$$

By (3.11a) and (3.40b),

$$b_h(\boldsymbol{\tau}_h, v_h) = b_h(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, v_h) = 0, \tag{3.41}$$

i.e., $(\operatorname{div}\operatorname{Div})_h \boldsymbol{\tau}_h = 0$. From [21], we have the exact sequence

$$[\mathcal{S}_h]^2 \xrightarrow{\operatorname{Curl}^{\mathrm{s}}} \boldsymbol{\Sigma}_h \xrightarrow{(\operatorname{div}\operatorname{Div})_h} \mathcal{U}_h \to 0.$$

Hence $\boldsymbol{\tau}_h = \operatorname{Curl}^{\mathrm{s}} \boldsymbol{\varphi}_h$ for some $\boldsymbol{\varphi}_h \in [\mathcal{S}_h^1]^2$. By (3.40a), (3.11b) and (3.41),

$$\begin{aligned}
(1-\nu)\|\boldsymbol{\tau}_h\|_{0,\Omega}^2 &\le \langle \mathcal{C}\boldsymbol{\tau}_h, \boldsymbol{\tau}_h \rangle \\
&= (\mathcal{C}(\pi_h \boldsymbol{\sigma} - \boldsymbol{\sigma}), \boldsymbol{\tau}_h) - b_h(\boldsymbol{\tau}_h, u - u_h) \\
&= (\mathcal{C}(\pi_h \boldsymbol{\sigma} - \boldsymbol{\sigma}), \boldsymbol{\tau}_h) - b_h(\boldsymbol{\tau}_h, I_h u - u_h) \\
&= (\mathcal{C}(\pi_h \boldsymbol{\sigma} - \boldsymbol{\sigma}), \boldsymbol{\tau}_h).
\end{aligned}$$

Then Theorem 3.2.5 follows from Lemma 3.2.4 and the above inequality. □

On uniformly parallel grids, Theorem 3.2.5 gives $\|\pi_h \boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{0,\Omega} = O(h^2 |\log h|^{\frac{1}{2}})$. If $\mathcal{C} = \mathcal{A}$ is the identity operator, Theorem 5.3 in [36] says that $\|\pi_h \boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{0,\Omega} = O(h^{1.5})$ which is suboptimal according to Theorem 3.2.5.

3.3 Superconvergent recovery

In this section, we introduce the recovery operator R_h based on the local least squares fitting similar to R_h^r in Chapter 2.

Definition 3.3.1. The operator $R_h : \boldsymbol{\Sigma}_h \to [\mathcal{S}_h^1]_s^4$ is defined as follows. For $\boldsymbol{z} \in \mathcal{N}_h$, let $\boldsymbol{\tau}_z \in [\mathcal{P}_1(\omega_z)]_s^4$ minimize the functional

$$\mathcal{F}(\boldsymbol{\tau}) = \sum_{e \in \mathcal{E}_h(\omega_z)} ([\boldsymbol{\tau}(\boldsymbol{m}_e)]_{nn} - [\boldsymbol{\tau}_h(\boldsymbol{m}_e)]_{nn})^2$$

subject to $\boldsymbol{\tau} \in [\mathcal{P}_1(\omega_z)]_s^4$, where $\boldsymbol{m}_e$ is the midpoint of e and $\boldsymbol{n}$ is a unit normal to e. Then $R_h\boldsymbol{\tau}_h(\boldsymbol{z}) := \boldsymbol{\tau}_z(\boldsymbol{z})$.

Clearly, R_h can also be applied to $\boldsymbol{\sigma}$ and $R_h\boldsymbol{\sigma} = R_h\pi_h\boldsymbol{\sigma}$. To clarify the post-processing procedure, we rewrite R_h in linear algebra language. For an internal vertex $\boldsymbol{z}$, let $\{e_j\}_{j=1}^N$ be the set of edges in the local patch ω_z. Assume

$$\boldsymbol{\tau}_z = \begin{pmatrix} c_1 + c_2x_1 + c_3x_2, & c_4 + c_5x_1 + c_6x_2 \\ c_4 + c_5x_1 + c_6x_2, & c_7 + c_8x_1 + c_9x_2 \end{pmatrix}.$$

Let $\boldsymbol{m}_j = (m_{j1}, m_{j2})^\intercal$ and $\boldsymbol{n}_j = (n_{j1}, n_{j2})^\intercal$ be the midpoint and unit normal to e_j respectively. Let $\boldsymbol{d}_z = (\boldsymbol{n}_1^\intercal\boldsymbol{\tau}_h(\boldsymbol{m}_1)\boldsymbol{n}_1, \ldots, \boldsymbol{n}_N^\intercal\boldsymbol{\tau}_h(\boldsymbol{m}_N)\boldsymbol{n}_N)^\intercal$, and $\boldsymbol{A}_z = (\boldsymbol{a}_1^\intercal, \ldots, \boldsymbol{a}_N^\intercal)^\intercal$ be an $N \times 9$ matrix with

$$\boldsymbol{a}_j = (n_{j1}^2, n_{j1}^2m_{j1}, n_{j1}^2m_{j2}, 2n_{j1}n_{j2}, 2n_{j1}n_{j2}m_{j1}, 2n_{j1}n_{j2}m_{j2}, n_{j2}^2, n_{j2}^2m_{j1}, n_{j2}^2m_{j2}).$$

Then $\boldsymbol{c} = (c_1, \ldots, c_9)^\intercal$ solves $\min_{\hat{\boldsymbol{c}} \in \mathbb{R}^9} |\boldsymbol{A}_z\hat{\boldsymbol{c}} - \boldsymbol{d}_z|^2$. In other words, the normal equation $\boldsymbol{A}_z^\intercal\boldsymbol{A}_z\boldsymbol{c} = \boldsymbol{A}_z^\intercal\boldsymbol{d}_z$ holds.

In some cases, the local least squares problem may not have a unique solution.

One can add some extra elements to the local patch ω_z, e.g. enlarge ω_z by one layer. The following theorem shows that there is almost no need to enlarge the local patch under a good mesh $\mathcal{T}_h$.

Lemma 3.3.1. *Let $\boldsymbol{z}$ be an internal vertex. Assume $\#\mathcal{T}_h(\omega_z) \geq 5$ and the sum of each pair of adjacent angles in ω_z is $\leq \pi$. Then there exists a unique $\boldsymbol{\tau}_z$ at $\boldsymbol{z}$.*

Proof. Assume $\pi_h \boldsymbol{\tau}_z = \boldsymbol{O}_{2\times 2}$ on ω_z. Then by Lemma 3.2.1,

$$\boldsymbol{\tau}_z|_T = \boldsymbol{\tau}_z|_T - \pi_T(\boldsymbol{\tau}_z|_T) = \operatorname{Curl}^{\mathrm{s}} \boldsymbol{r}_T \tag{3.42}$$

for each $T \in \mathcal{T}_h(\omega_z)$ with $\boldsymbol{r}_T = \sum_{k=1}^{3} \phi_k \boldsymbol{\gamma}^{\boldsymbol{\tau}_z|_T} \in [\mathcal{P}_2(K)]^2$, where $\boldsymbol{r}_T$ vanishes on three vertices of T. Let $\boldsymbol{r}_z$ be the piecewise quadratic polynomial on ω_z whose restriction to K is $\boldsymbol{r}_T$. We claim that $\boldsymbol{r}_z$ is indeed a quadratic polynomial on ω_z. Consider an edge $e \in \mathcal{E}_h(\omega_z)$ shared by $T, T' \in \mathcal{T}_h(\omega_z)$ with unit tangent $\boldsymbol{t}_e$ and normal $\boldsymbol{n}_e = \boldsymbol{Q}\boldsymbol{t}_e$. It suffices to show that $\partial_{\boldsymbol{t}_e}^{\alpha_1} \partial_{\boldsymbol{n}_e}^{\alpha_2} \boldsymbol{r}_T = \partial_{\boldsymbol{t}_e}^{\alpha_1} \partial_{\boldsymbol{n}_e}^{\alpha_2} \boldsymbol{r}_{T'}$ on e for $\alpha_1 + \alpha_2 \leq 2$. By the formula for $\boldsymbol{\gamma}^{\boldsymbol{\tau}_z|_T}$ and $\boldsymbol{\gamma}^{\boldsymbol{\tau}_z|_{T'}}$ in Lemma 3.2.1, $\boldsymbol{r}_T = \boldsymbol{r}_{T'}$ on e. Hence

$$\partial_{\boldsymbol{t}_e} \boldsymbol{r}_T = \partial_{\boldsymbol{t}_e} \boldsymbol{r}_{T'}, \quad \partial_{\boldsymbol{t}_e}^2 \boldsymbol{r}_T = \partial_{\boldsymbol{t}_e}^2 \boldsymbol{r}_{T'} \quad \text{on } e. \tag{3.43}$$

By (3.42), we have

$$\begin{aligned}
&\boldsymbol{t}_e^{\intercal} \partial_{\boldsymbol{n}_e} \boldsymbol{r}_K = \boldsymbol{t}_e^{\intercal} \boldsymbol{\tau}_z \boldsymbol{t}_e \text{ on } T, \quad \boldsymbol{t}_e^{\intercal} \partial_{\boldsymbol{n}_e} \boldsymbol{r}_{T'} = \boldsymbol{t}_e^{\intercal} \boldsymbol{\tau}_z \boldsymbol{t}_e \quad \text{on } T', \\
&\frac{1}{2} \left(\boldsymbol{n}_e^{\intercal} \partial_{\boldsymbol{n}_e} \boldsymbol{r}_T - \boldsymbol{t}_e^{\intercal} \partial_{\boldsymbol{t}_e} \boldsymbol{r}_T \right) = \boldsymbol{t}_e^{\intercal} \boldsymbol{\tau}_z \boldsymbol{n}_e \text{ on } T, \\
&\frac{1}{2} \left(\boldsymbol{n}_e^{\intercal} \partial_{\boldsymbol{n}_e} \boldsymbol{r}_{T'} - \boldsymbol{t}_e^{\intercal} \partial_{\boldsymbol{t}_e} \boldsymbol{r}_{T'} \right) = \boldsymbol{t}_e^{\intercal} \boldsymbol{\tau}_z \boldsymbol{n}_e \text{ on } T'.
\end{aligned} \tag{3.44}$$

(3.44) with (3.43) imply

$$\partial_{\boldsymbol{n}_e} \boldsymbol{r}_T = \partial_{\boldsymbol{n}_e} \boldsymbol{r}_{T'}, \quad \partial_{\boldsymbol{t}_e} \partial_{\boldsymbol{n}_e} \boldsymbol{r}_T = \partial_{\boldsymbol{t}_e} \partial_{\boldsymbol{n}_e} \boldsymbol{r}_{T'} \quad \text{on } e. \tag{3.45}$$

Applying $\partial_{\boldsymbol{n}_e}$ to (3.44) and using (3.45), we obtain $\partial^2_{\boldsymbol{n}_e}\boldsymbol{r}_T = \partial^2_{\boldsymbol{n}_e}\boldsymbol{r}_{T'}$ on e. Therefore the claim is confirmed and $\boldsymbol{\tau}_z = \mathrm{Curl}^{\mathrm{s}}\, \boldsymbol{r}_z$, where $\boldsymbol{r}_z \in [\mathcal{P}_2(\omega_z)]^2$ vanishing at all vertices in ω_z. By Theorem 2.3 in [57], $\boldsymbol{r}_z = \boldsymbol{0}$ and thus $\boldsymbol{\tau}_z = \boldsymbol{O}_{2\times 2}$. Hence $\pi_h\boldsymbol{\tau}_z = \boldsymbol{O}_{2\times 2}$ on ω_z implies $\boldsymbol{\tau}_z = \boldsymbol{O}_{2\times 2}$. In other words, $\boldsymbol{A}_z\boldsymbol{c} = \boldsymbol{0}$ implies $\boldsymbol{c} = 0$, which implies the uniqueness of the least squares solution $\boldsymbol{\tau}_z$ at $\boldsymbol{z}$. □

If the condition in Lemma 3.3.1 is violated, one can add extra elements to ω_z, see the remark below Lemma 2.4.3. In the rest of this chapter, we assume that there exists a unique $\boldsymbol{\tau}_z$ at each $\boldsymbol{z} \in \mathcal{N}_h$.

The uniqueness of the least squares solution implies $R_h\boldsymbol{\tau} = \boldsymbol{\tau}$ on T for $\boldsymbol{\tau} \in [\mathcal{P}_1(\omega_T)]^4_s$, which is called the polynomial preserving property, see Chapter 2. The super-approximation property then follows from a homogeneity or compactness argument.

Theorem 3.3.2. *R_h is locally stable, namely, $|(R_h\boldsymbol{\tau}_h)(\boldsymbol{z})| \lesssim \|\boldsymbol{\tau}_h\|_{0,\infty,\omega_z}$, $\|R_h\boldsymbol{\tau}_h\|_{0,T} \lesssim \|\boldsymbol{\tau}_h\|_{0,\omega_T}$ for all $\boldsymbol{\tau}_h \in \boldsymbol{\Sigma}_h$. In addition,*

$$\|\boldsymbol{\sigma} - R_h\boldsymbol{\sigma}\|_{0,\Omega} \lesssim h^2|\boldsymbol{\sigma}|_{2,\Omega}.$$

Combining Theorems 3.2.5, and 3.3.2 and splitting $\boldsymbol{\sigma} - R_h\boldsymbol{\sigma}_h = \boldsymbol{\sigma} - R_h\boldsymbol{\sigma} + R_h(\pi_h\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)$, we obtain the following superconvergent recovery theorem.

Theorem 3.3.3. *If $\mathcal{T}_h$ satisfies the (α,β)-condition, then*

$$\|\boldsymbol{\sigma} - R_h\boldsymbol{\sigma}_h\|_{0,\Omega} \lesssim (1-\nu)^{-1}h^{1+\min(\frac{1}{2},\alpha,\frac{\beta}{2})}\big(|\boldsymbol{\sigma}|_{1,\infty,\Omega} + \|\boldsymbol{\sigma}\|_{2,\Omega}\big).$$

If $\mathcal{T}_h$ satisfies the piecewise strong (α,β)-condition, then

$$\|\boldsymbol{\sigma} - R_h\boldsymbol{\sigma}_h\|_{0,\Omega} \lesssim (1-\nu)^{-1}h^{1+\min(1,\alpha,\frac{\beta}{2})}|\log h|^{\frac{1}{2}}\|\boldsymbol{\sigma}\|_{2,\infty,\Omega}.$$

Remark. The piecewise strong (α,β)-condition with $(\alpha,\beta)=(\infty,\infty)$ holds on any sequence of grids obtained from uniformly refining an arbitrary initial grid. Hence $\|\boldsymbol{\sigma}-R_h\boldsymbol{\sigma}_h\|_{0,\Omega}=O(h^2|\log h|^{\frac{1}{2}})$ provided $\boldsymbol{\sigma}$ is smooth enough.

Since $\boldsymbol{\sigma}_h$ is piecewise constant, the gradient recovery operator G_h defined by (1.5) can be applied to it. As mentioned before, G_h has boundedness and super-approximation property. Hence $\|\boldsymbol{\sigma}-G_h\boldsymbol{\sigma}_h\|_{0,\Omega}=O(h^{1+\min(\frac{1}{2},\alpha,\frac{\beta}{2})})$.

3.4 Numerical experiments

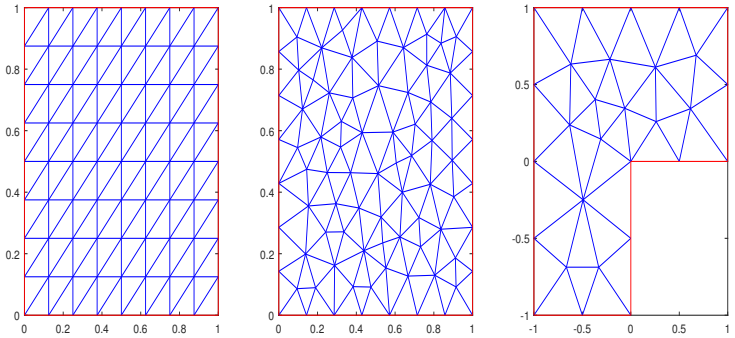

Figure 3.1: (left)uniform grid (middle)Delaunay grid (right)adaptive initial grid

Table 3.1: uniform initial grid with regular refinement

nt	$\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|$	$\|\boldsymbol{\sigma}-G_h\boldsymbol{\sigma}_h\|$	$\|\boldsymbol{\sigma}-R_h\boldsymbol{\sigma}_h\|$	$\|\pi_h\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|$
128	1.642e+00	5.792e-01	8.585e-01	4.889e-01
512	8.506e-01	1.924e-01	2.501e-01	1.391e-01
2048	4.294e-01	6.071e-02	6.490e-02	3.619e-02
8192	2.152e-01	1.951e-02	1.624e-02	9.145e-03
32768	1.077e-01	6.484e-03	4.037e-03	2.29e-03
order	0.994	1.631	1.986	1.976

In this section, we test the recovery operators G_h and R_h. The experiments are performed using the PDE toolbox in Matlab 2016b. In tables, 'nt' denotes the number of triangles. The norm $\|\cdot\|$ is short for $\|\cdot\|_{0,\Omega}$. The order of convergence is p such that

Table 3.2: Delaunay initial grid with regular refinement

nt	$\|\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|\|$	$\|\|\boldsymbol{\sigma} - G_h\boldsymbol{\sigma}_h\|\|$	$\|\|\boldsymbol{\sigma} - R_h\boldsymbol{\sigma}_h\|\|$	$\|\|\pi_h\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|\|$
148	1.412e+00	5.452e-01	8.009e-01	3.714e-01
592	7.238e-01	1.881e-01	2.113e-01	1.079e-01
2368	3.646e-01	6.325e-02	5.519e-02	2.929e-02
9472	1.827e-01	2.144e-02	1.416e-02	7.757e-03
37888	9.141e-02	7.371e-03	3.602e-03	2.04e-03
order	0.996	1.558	1.959	1.911

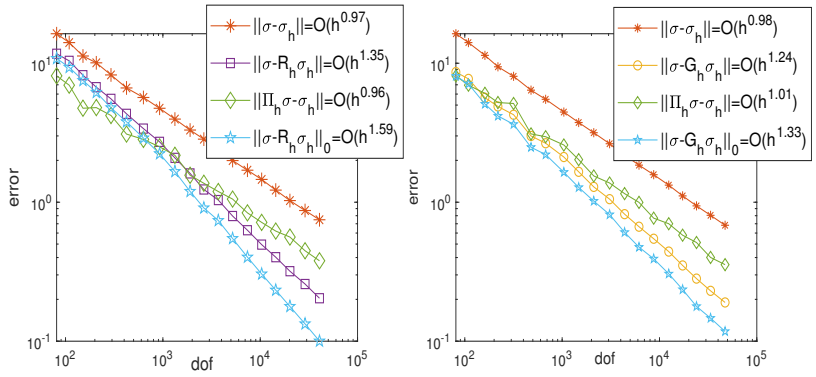

Figure 3.2: (left)error curve from R_h (right)error curve from G_h

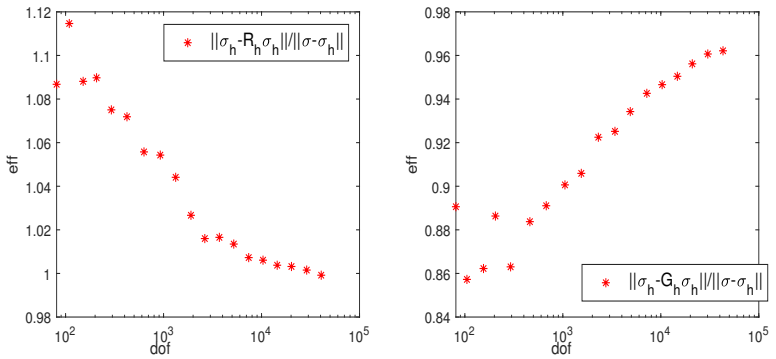

Figure 3.3: (left)efficiency curve from R_h (right)efficiency curve from G_h

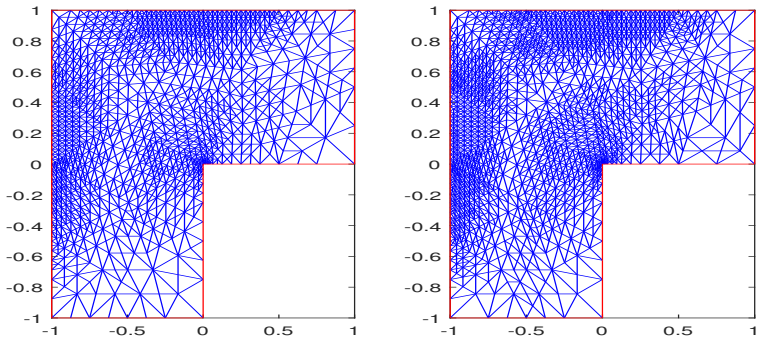

Figure 3.4: (left)graded grid from R_h (right)grade grid from G_h

error/dof$^{\frac{p}{2}}$ is approximately constant, where dof is the number of degrees of freedom. p is calculated by least squares using data in tables.

Problem1: We first test G_h and R_h with

$$u = \exp(x_1 + x_2)x_1^2(x_1 - 1)^2 \sin^2(\pi x_2), \quad \Omega = [0, 1] \times [0, 1]$$

with $\nu = 0.3$ in (3.4). We start with the uniform and Delaunay triangulations in Figure 3.1 respectively, and then computed a sequence of meshes by regular refinement, i.e., dividing an element into four similar subelements by connecting the midpoints of each edge. Numerical results are presented in Tables 3.1 and 3.2. The first row in tables are not used in least squares fitting for evaluating order of convergence.

Grids starting from the uniform initial grid satisfy the strong (α, β)-condition with $(\alpha, \beta) = (\infty, \infty)$. Grids starting from the Delaunay initial grid satisfy the strong (α, β)-condition with $(\alpha, \beta) = (\infty, 1/2)$ as well as the piecewise strong (α, β)-condition with $(\alpha, \beta) = (\infty, \infty)$. Numerical results confirm Theorem 3.2.5 and 3.3.3. It should be noted that R_h has better superconvergent property than G_h, since $\|\boldsymbol{\sigma} - R_h\boldsymbol{\sigma}_h\|_{0,\Omega} = O(h^2)$, while $\|\boldsymbol{\sigma} - G_h\boldsymbol{\sigma}_h\|_{0,\Omega} = O(h^{1.5})$ in both cases.

Problem2: In the second experiment, we test the performance of the a pos-

teriori error estimator $\eta_T = \|\boldsymbol{\sigma}_h - A_h\boldsymbol{\sigma}_h\|_{0,T}$ with $A_h = R_h$ or G_h on (3.4), where $\nu = 0.3$,

$$
\begin{aligned}
&u(r,\theta) = (r^2\cos^2\theta - 1)^2(r^2\sin^2\theta - 1)^2 r^{1+\gamma} g(\theta), \quad \gamma = 0.544483736782464,\\
&g(\theta) = \left(\frac{1}{\gamma-1}\sin((\gamma-1)\omega) - \frac{1}{\gamma+1}\sin((\gamma+1)\omega)\right)(\cos((\gamma-1)\theta) - \cos((\gamma+1)\theta))\\
&\quad - (\frac{1}{\gamma-1}\sin((\gamma-1)\theta) - \frac{1}{\gamma+1}\sin((\gamma+1)\theta))(\cos((\gamma-1)\omega) - \cos((\gamma+1)\omega)),\\
&\Omega = ([-1,1]\times[-1,1])\backslash([0,1]\times[-1,0]), \quad \omega = \frac{3\pi}{2},
\end{aligned}
$$

where (r,θ) is the polar coordinate, see [34] for the construction of u. We start from the initial grid in Figure 3.1. The adaptive feedback loop is the classical

$$\mathsf{Solve} \rightarrow \mathsf{Estimate} \rightarrow \mathsf{Mark} \rightarrow \mathsf{Refine}$$

loop (cf.[26]). In each marking step, we mark triangles in a set $\mathcal{M}$ such that $\sum_{T\in\mathcal{M}}\eta_T^2 \geq 0.49\sum_{T\in\mathcal{T}_h}\eta_T^2$. Marked triangles are refined by longest edge bisection. The efficiency index

$$\text{eff} = \frac{\|\boldsymbol{\sigma}_h - A_h\boldsymbol{\sigma}_h\|}{\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|}$$

is computed. We also compute the interior error measured by $\|\cdot\|_0 := \|\cdot\|_{0,\Omega_0}$, where $\Omega_0 = \Omega\backslash\{x_1^2 + x_2^2 \leq 0.2^2\}$.

The simple recovery-based error indicator η_T generates correct girds for the solution with a point singularity, see Figure 3.4. It is not hard to see that bisection tends to destroy the approximate parallelogram structure and thus there is no supercloseness as shown in Figure 3.2. Although the supercloseness estimate in section 3.2 fails in this case, there is still apparent global and interior superconvergence under adaptive

meshes, see Figure 3.2. Use triangle

$$\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\| - \text{h.o.t.} \leq \|\boldsymbol{\sigma}_h - A_h\boldsymbol{\sigma}_h\| \leq \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\| + \text{h.o.t.},$$

where h.o.t.$= \|\boldsymbol{\sigma} - A_h\boldsymbol{\sigma}_h\|$. Hence $\|\boldsymbol{\sigma}_h - A_h\boldsymbol{\sigma}_h\|$ is asymptotically exact, i.e., eff goes to 1, see Figure 3.3. We conjecture that this type of superconvergence is due to a large number of locally symmetric patches with respect to a point in grids, cf. [60] regarding second order elliptic equations.

Chapter 3, in full, has been submitted for publication of the material as it may appear in Numerische Mathematik, 2019, Li, Yu-Wen, Springer, 2019. The dissertation author was the author of this paper.

Chapter 4

Superconvergence of nonconforming methods

In this chapter, we develop superconvergence estimates of nonconforming FEMs for second and fourth order elliptic equations. As mentioned before, nonconforming FEMs are particularly suitable for discretizing systems of PDEs and higher order PDEs by relaxing the interelement continuity constraint in their finite element spaces. Due to this deviation from conformity in standard FEMs, it is very difficult to prove superconvergence of nonconforming methods directly (cf. [36] and references therein).

4.1 Superconvergence of Crouzeix–Raviart and Rannacher–Turek elements

For example, the standard lowest order triangular nonconforming element is the Crouzeix–Raviart(CR) element (cf.[25]). It can be observed in numerical experiments that the canonical interpolant u_I and the finite element solution u_h of the CR method for Poisson's equation are not superclose in the energy norm and thus the

aforementioned recovery framework does not work. In [19], Chen proved a superconvergence estimate by using a corrected interpolation instead of the canonical interpolation for CR elements. In [70], Ye develop superconvergence estimates of CR elements based on the framework of least-squares surface fitting. Based on an equivalence between the CR method and the lowest order Raviart–Thomas (RT) method for Poisson's equation (cf. [53, 3]), the authors in [36] first prove a superconvergent gradient recovery estimate for CR elements by using superconvergence estimates of RT elements in [12]. Readers are also referred to [24, 52, 51, 38] and references therein for superconvergence of other nonconforming elements.

The nonconforming Rannacher–Turek (NCRT) element (cf.[58]) is a generalization of CR elements on quadrilateral meshes which is very popular for numerically solving the Stokes problem. We mention that there is a superconvergence estimate of NCRT at some special points under some mildly distorted square meshes, see [54]. Considering Poisson's equation, the authors in [49] show that various rectangular nonconforming methods fail to have natural supercloseness estimates. In particular for NCRT elements, u_I and u_h are superclose in the energy norm only under square meshes. To overcome this barrier, they enrich the NCRT element by one degree of freedom at the centroid of each element and achieve superconvergent gradient recovery for the modified nonconforming element.

We shall consider the standard NCRT method (4.3) for solving the elliptic equation (4.1) with variable coefficients and lower order terms. First we compute a new numerical flux $\boldsymbol{p}_h$ from the NCRT finite element solution, see Theorem 4.1.3 for details. We shall show that $\boldsymbol{p}_h$ is superclose to $\Pi_h(a\nabla u)$ by comparing it with an auxiliary $H(\mathrm{div})$-conforming flux $\bar{\boldsymbol{p}}_h$ and using well-established superconvergence tools and techniques for RT elements in [30, 46], compared also with Chapter 2. Here Π_h is the canonical interpolation of the lowest order rectangular RT element. We then

construct a local edge-based weighted averaging operator A_h, which makes $\|a\nabla u - A_h\Pi_h(a\nabla u)\|_{0,\Omega}$ supersmall on any rectangular mesh. Hence $A_h\boldsymbol{p}_h$ superconverges to $a\nabla u$ on any rectangular mesh by a triangle-inequality argument. As far as we know, even for Poisson's equation, this is the first global superconvergent recovery method for the NCRT element on any rectangular mesh. Our supercloseness estimate also directly extends to $\mathbb{R}^d$ with $d \geq 3$.

For elliptic equations with variable coefficients and lower order terms, Arbogast and Chen in [2] can reformulate various mixed methods into modified nonconforming methods to obtain multigrid algorithms for mixed methods. However, the general equivalence expression is complicated and it is unclear how far the standard nonconforming finite element solution is from the modified one. On the other hand, superconvergence analysis of $H(\mathrm{div})$-conforming mixed finite elements is often more tractable than nonconforming methods and many sophisticated superconvergence estimates of mixed methods are available (cf.[30, 46]). Hence in this paper, we aim to push nonconforming methods to their mixed side just as in [36]. But we will not try to write the NCRT method (4.3) for general elliptic equations as an equivalent mixed method. All we need is the equivalence Lemma 3.2.4 for Poisson's equation. Moreover, our recovery scheme can be directly generalized to the Courzeix–Raviart method on triangular meshes. To our best knowledge, it is the first superconvergence estimate of nonconforming methods for elliptic equations with variable coefficients, lower order terms. It is surprising that everything matches so well.

Let $\Omega = [a, b] \times [c, d] \subset \mathbb{R}^2$ be a rectangle. Consider the second order elliptic equation

$$-\nabla \cdot (a\nabla u) + \boldsymbol{b} \cdot \nabla u + cu = f \quad \text{in } \Omega, \tag{4.1a}$$

$$u = g \quad \text{on } \partial\Omega, \tag{4.1b}$$

where $a(\boldsymbol{x}) \geq a_0 > 0$ for all $\boldsymbol{x} \in \Omega$, $a, \boldsymbol{b}, c$, and f are smooth functions on $\bar{\Omega}$.

Let $\mathcal{T}_h$ denote a partition of Ω by rectangles. Given a rectangle $T \in \mathcal{T}_h$, let $\ell_{T,1}$ and $\ell_{T,2}$ denote the width and height of T, respectively and $h = \max_{T\in\mathcal{T}_h} \max(\ell_{T,1}, \ell_{T,2})$ the mesh size. We assume that $h < 1$ and $\mathcal{T}_h$ is nondegenerate, i.e.

$$\max_{T\in\mathcal{T}_h} \max\left\{\frac{\ell_{T,1}}{\ell_{T,2}}, \frac{\ell_{T,2}}{\ell_{T,1}}\right\} \leq C < \infty,$$

where C is a constant independent of h. Let $\mathcal{E}_h$, $\mathcal{E}_h^o$, and $\mathcal{E}_h^\partial$ denote the set of edges, interior edges, and boundary edges, respectively. Let

$$\begin{aligned}\mathcal{V}_{g,h} :=& \{v_h \in L^2(\Omega) : v_h|_T \in \operatorname{span}\{1, x_1, x_2, x_1^2 - x_2^2\} \text{ for all } T \in \mathcal{T}_h,\\ & \fint_e v_h \text{ is single-valued for all } e \in \mathcal{E}_h^o, \fint_e v_h = \fint_e g \text{ for all } e \in \mathcal{E}_h^\partial\}.\end{aligned} \tag{4.2}$$

The Rannacher–Turek nonconforming method for (4.1) is to find $u_h \in \mathcal{V}_{g,h}$, such that

$$(a\nabla_h u_h, \nabla_h v) + (\boldsymbol{b}\cdot\nabla_h u_h, v) + (cu_h, v) = (f, v), \quad \forall v \in \mathcal{V}_{0,h}, \tag{4.3}$$

where ∇_h denotes the piecewise gradient w.r.t. $\mathcal{T}_h$. We assume that the standard a priori error estimate for the NCRT method holds:

$$\|u - u_h\|_{0,\Omega} + h\|\nabla_h(u - u_h)\|_{0,\Omega} \lesssim h^2\|u\|_{2,\Omega}. \tag{4.4}$$

Readers are referred to [14] for the analogue of (4.4) for the CR method.

The NCRT space $\tilde{\mathcal{V}}_h$ using pointwise function evaluation will be used to define the averaging operator A_h.

$$\begin{aligned}\tilde{\mathcal{V}}_h :=& \{v_h \in L^2(\Omega) : v_h|_T \in \operatorname{span}\{1, x_1, x_2, x_1^2 - x_2^2\} \text{ for all } T \in \mathcal{T}_h,\\ & v_h \text{ is continuous at the midpoint of each } E \in \mathcal{E}_h^o\}.\end{aligned} \tag{4.5}$$

Let $Q_{k,l}(T)$ denote the set of polynomials of degree $\leq k$ in x_1 and of degree $\leq l$ in x_2 on the element T. Let $H(\mathrm{div},\Omega) := \{\boldsymbol{q} \in [L^2(\Omega)]^2 : \mathrm{div}\,\boldsymbol{q} \in L^2(\Omega)\}$. The lowest order rectangular Raviart–Thomas finite element space is

$$\mathcal{RT}_h := \{\boldsymbol{q}_h \in H(\mathrm{div},\Omega) : \boldsymbol{q}_h|_T \in Q_{1,0}(T) \times Q_{0,1}(T),\ \forall T \in \mathcal{T}_h\}.$$

The broken Raviart–Thomas space is

$$\mathcal{RT}_h^{-1} := \{\boldsymbol{q}_h \in [L^2(\Omega)]^2 : \boldsymbol{q}_h|_T \in Q_{1,0}(T) \times Q_{0,1}(T),\ \forall T \in \mathcal{T}_h\}.$$

Given $\boldsymbol{q} \in H(\mathrm{div},\Omega)$, $\Pi_h\boldsymbol{q}$ is the unique element in $\mathcal{RT}_h$ satisfying $\int_e (\Pi_h\boldsymbol{q})\cdot\boldsymbol{n} = \int_e \boldsymbol{q}\cdot\boldsymbol{n}$, for all $e \in \mathcal{E}_h$, where $\boldsymbol{n}$ is a unit normal to e. Let P_h be the $L^2(\Omega)$-projection onto the space of piecewise constant functions. The commutativity (2.12) still holds.

Let $e \in \mathcal{E}_h^o$ and T, T' be the two rectangles sharing e. Let $\boldsymbol{n}$ and $\boldsymbol{n}'$ denote the outward unit normal induced by T and T' respectively. In the analysis of nonconforming methods, it is convenient to introduce notations for jumps and averages on e:

$$\begin{aligned}
[\![\boldsymbol{q}]\!] &:= \boldsymbol{q}|_T \cdot \boldsymbol{n} + \boldsymbol{q}|_{T'} \cdot \boldsymbol{n}',\\
\{\boldsymbol{q}\} &:= (\boldsymbol{q}|_T + \boldsymbol{q}|_{T'})/2,\\
[\![v]\!] &:= (v|_T\boldsymbol{n} + v|_{T'}\boldsymbol{n}')/2,\\
\{v\} &:= (v|_T + v|_{T'})/2,
\end{aligned}$$

where q is a vector and v is a scalar. For $e \in \mathcal{E}_h^\partial$, $[\![\boldsymbol{q}]\!] := \boldsymbol{q}\cdot\boldsymbol{n}$, $\{v\} := v$, $[\![v]\!] := \boldsymbol{0}$. It is readily to check that

$$[\![\boldsymbol{q}v]\!] = [\![\boldsymbol{q}]\!]\{v\} + [\![v]\!]\cdot\{\boldsymbol{q}\}. \tag{4.6}$$

With these notations, a useful fact is that

$$\boldsymbol{q}_h \in \mathcal{RT}_h \text{ if and only if } \boldsymbol{q}_h \in \mathcal{RT}_h^{-1} \text{ and } [\![\boldsymbol{q}_h]\!] = 0, \forall e \in \mathcal{E}_h^o. \tag{4.7}$$

4.1.1 Supercloseness

In this section, we derive a supercloseness estimate for NCRT elements, which is essential to develop superconvergent flux recovery. First we need a lemma in the spirit of Marini (cf. [53]). Let div_h denote the piecewise divergence and define the piecewise function $\boldsymbol{r}_h$ as

$$\boldsymbol{r}_h|_T := \left(\frac{\ell_{T,2}^2}{\ell_{T,1}^2 + \ell_{T,2}^2}(x_1 - x_{T,1}), \quad \frac{\ell_{T,1}^2}{\ell_{T,1}^2 + \ell_{T,2}^2}(x_2 - x_{T,2}) \right)^{\mathsf{T}}, \ T \in \mathcal{T}_h,$$

where $T = [x_{1,i}, x_{1,i+1}] \times [x_{2,j}, x_{2,j+1}]$, $\ell_{T,1} = x_{1,i+1} - x_{1,i}$, $\ell_{T,2} = x_{2,j+1} - x_{2,j}$, and $(x_{T,1}, x_{T,2})^{\mathsf{T}}$ is the centroid of T.

Lemma 4.1.1. *Let $\bar{f}$ be a piecewise constant, $\boldsymbol{q}_h|_T \in Q_{1,0}(T) \times Q_{0,1}(T)$ and $\mathrm{div}_h \boldsymbol{q}_h = 0$. Assume that*

$$(\boldsymbol{q}_h, \nabla_h v) = (\bar{f}, v) \tag{4.8}$$

for all $v \in \mathcal{V}_{0,h}$. Then $\boldsymbol{q}_h - \bar{f}\boldsymbol{r}_h \in \mathcal{RT}_h$.

Proof. Consider the vertical edge $e \in \mathcal{E}_h^o$ and the two rectangles

$$T = [x_{1,i}, x_{1,i+1}] \times [x_{2,j}, x_{2,j+1}], \quad T' = [x_{1,i+1}, x_{1,i+2}] \times [x_{2,j}, x_{2,j+1}]$$

sharing it. Let $v_e \in \mathcal{V}_{0,h}$ be the basis function such that $\fint_e v_e = 1$ and $\fint_{e'} v_e = 0$ for $e' \neq e$ in $\mathcal{E}_h$. Note that $(\boldsymbol{q}_h \cdot \boldsymbol{n})|_{e'}$ is constant for all $e' \in \mathcal{E}_h$. It then follows from (4.8),

$\operatorname{div} \boldsymbol{q}_h|_T = \operatorname{div} \boldsymbol{q}_h|_{T'} = 0$ and integration by parts that

$$\int_e [\![\boldsymbol{q}_h]\!] = \int_{T'\cup T} \bar{f} v_e. \tag{4.9}$$

By direct calculation, the right hand side of (4.9) is

$$\int_T v_e = \frac{|T|\ell_{T,2}^2}{2(\ell_{T,1}^2+\ell_{T,2}^2)}, \quad \int_{T'} v_e = \frac{|T'|\ell_{T',2}^2}{2(\ell_{T',1}^2+\ell_{T',2}^2)},$$

and thus

$$[\![\boldsymbol{q}_h - \bar{f}\boldsymbol{r}_h]\!] = 0 \text{ on } e. \tag{4.10}$$

Similarly, (4.10) also holds for horizontal edges. Combining (4.10) with the fact $(\boldsymbol{q}_h - \bar{f}\boldsymbol{r}_h)|_T \in Q_{1,0}(T) \times Q_{0,1}(T)$, we conclude that $\boldsymbol{q}_h - \bar{f}\boldsymbol{r}_h \in \mathcal{RT}_h$. □

Remark. It seems that the NCRT method using degrees of freedom based on pointwise function evaluation does not have a similar equivalence.

To apply Lemma 4.1.1, we then introduce the auxiliary nonconforming method: Find $\bar{u}_h \in \mathcal{V}_{g,h}$, such that

$$(a\nabla_h \bar{u}_h, \nabla_h v) = (P_h(f - cu - \boldsymbol{b}\cdot\nabla u), v), \quad \forall v \in \mathcal{V}_{0,h}. \tag{4.11}$$

The following lemma shows that u_h and $\bar{u}_h$ are superclose in the H^1-norm.

Lemma 4.1.2. *Let u_h and $\bar{u}_h$ solve* (4.3) *and* (4.11)*, respectively. Then*

$$\|\nabla_h(u_h - \bar{u}_h)\|_{0,\Omega} \lesssim h^2 \|u\|_{2,\Omega}.$$

Proof. Subtracting (4.11) from (4.3) yields

$$(a\nabla_h(u_h-\bar{u}_h),\nabla_h v)=(f-cu_h-\boldsymbol{b}\cdot\nabla_h u_h-P_h(f-cu-\boldsymbol{b}\cdot\nabla u),v).$$

It then follows from (4.4), and the discrete Poincaré inequality for nonconforming finite elements (cf. Theorem 10.6.12. in [14])

$$\|v\|_{h,1,\Omega}\lesssim\|\nabla_h v\|_{0,\Omega},\quad\forall v\in\mathcal{V}_{0,h},\tag{4.12}$$

that

$$\begin{aligned}&(a\nabla_h(u_h-\bar{u}_h),\nabla_h v)\\&=(f-cu-\boldsymbol{b}\cdot\nabla u-P_h(f-cu-\boldsymbol{b}\cdot\nabla u),v-P_h v)\\&\qquad+(c(u-u_h),v)+(\boldsymbol{b}\cdot\nabla_h(u-u_h),v)\\&=O(h^2)(|f|_{1,\Omega}+\|u\|_{2,\Omega})\|\nabla_h v\|_{0,\Omega}+(\boldsymbol{b}\cdot\nabla_h(u-u_h),v).\end{aligned}\tag{4.13}$$

It remains to show that $(\boldsymbol{b}\cdot\nabla_h(u-u_h),v)$ is supersmall. Using integrating by parts, (4.6), and that $⨍_e[\![u-u_h]\!]=\mathbf{0}$, we have

$$\begin{aligned}&(\boldsymbol{b}\cdot\nabla_h(u-u_h),v)\\&=\sum_{T\in\mathcal{T}_h}\int_{\partial T}(u-u_h)v\boldsymbol{b}\cdot\boldsymbol{n}-\int_T(u-u_h)\nabla\cdot(\boldsymbol{b}v)\\&=\sum_{e\in\mathcal{E}_h}\int_e\{u-u_h\}[\![v\boldsymbol{b}-\boldsymbol{c}_e]\!]+[\![u-u_h]\!]\cdot\{v\boldsymbol{b}-\boldsymbol{c}_e\}-\int_\Omega(u-u_h)\nabla_h\cdot(\boldsymbol{b}v)\end{aligned}\tag{4.14}$$

for any constants $\boldsymbol{c}_e\in\mathbb{R}^2$. In particular, we choose $\boldsymbol{c}_e=\boldsymbol{b}(m_e)⨍_e v$. By the trace inequality (1.9), we have

$$\begin{aligned}&\|\{u-u_h\}\|_{0,e}+\|[\![u-u_h]\!]\|_{0,e}\\&\quad\lesssim h^{-\frac{1}{2}}\|u-u_h\|_{0,\omega_e}+h^{\frac{1}{2}}\|\nabla_h(u-u_h)\|_{0,\omega_e}\end{aligned}\tag{4.15}$$

and

$$\|[\![v\boldsymbol{b}-\boldsymbol{c}_e]\!]\|_{0,e}+\|\{v\boldsymbol{b}-\boldsymbol{c}_e\}\|_{0,e}\lesssim h^{\frac{1}{2}}\|v\|_{h,1,\omega_e}. \tag{4.16}$$

Combining (4.14) with Hölder's inequality, (4.12), (4.15), (4.16) and (4.4), we have

$$\begin{aligned}
&|(\boldsymbol{b}\cdot\nabla_h(u-u_h),v)|\\
&\lesssim\sum_{e\in\mathcal{E}_h}\|\{u-u_h\}\|_{0,e}\|[\![v\boldsymbol{b}-\boldsymbol{c}_e]\!]\|_{0,e}\\
&\qquad+\|[\![u-u_h]\!]\|_{0,e}\|\{v\boldsymbol{b}-\boldsymbol{d}_e\}\|_{0,e}+\|u-u_h\|\|v\|_{h,1,\Omega}\\
&\lesssim\sum_{e\in\mathcal{E}_h}(\|u-u_h\|_{0,\omega_e}+h\|\nabla_h(u-u_h)\|_{0,\omega_e})\|v\|_{h,1,\omega_e}\\
&\qquad+\|u-u_h\|_{0,\Omega}\|v\|_{h,1,\Omega}\\
&\lesssim h^2\|u\|_{2,\Omega}\|\nabla_h v\|_{0,\Omega}.
\end{aligned} \tag{4.17}$$

Combining (4.17) with (4.13), we obtain Lemma 4.1.2. □

Now we are in a position to present supercloseness results.

Theorem 4.1.3. *Let $\boldsymbol{p}=a\nabla u$. Let Q_h be the L^2-projection onto $\nabla_h\mathcal{V}_{0,h}$ and $\boldsymbol{p}_h:=Q_h(a\nabla_h u_h)-\boldsymbol{r}_hP_h(f-cu_h-\boldsymbol{b}\cdot\nabla_h u_h)$, where $\boldsymbol{r}_h$ is defined in Lemma 4.1.1. Then*

$$\|\Pi_h\boldsymbol{p}-\boldsymbol{p}_h\|_{0,\Omega}\lesssim h^2\|u\|_{3,\Omega}.$$

Proof. It follows from Lemma 4.1.1 and $\operatorname{div}_h\circ Q_h=0$ that $\bar{\boldsymbol{p}}_h:=Q_h(a\nabla_h\bar{u}_h)-\boldsymbol{r}_hP_h(f-cu-\boldsymbol{b}\cdot\nabla u)\in\mathcal{RT}_h\subset H(\operatorname{div},\Omega)$. Let $\boldsymbol{q}_h=\Pi_h\boldsymbol{p}-\bar{\boldsymbol{p}}_h$. Using (2.12) and $\operatorname{div}_h\boldsymbol{r}_h=1$,

$$\operatorname{div}\boldsymbol{q}_h=P_h\operatorname{div}(a\nabla u)-P_h(f-cu-\boldsymbol{b}\cdot\nabla u)=0.$$

and thus $\boldsymbol{q}_h|_T=(c_1x_1+c_2,-c_1x_2+c_3)^\intercal$ for some $c_i\in\mathbb{R}$ on an element $T\in\mathcal{T}_h$.

Simpson's quadrature rule then yields

$$\begin{aligned}\int_T \boldsymbol{r}_h \cdot \boldsymbol{q}_h &= \int_T \boldsymbol{r}_h \cdot (\boldsymbol{q}_h - (c_2 + c_1 x_{T,1}, c_3 - c_1 x_{T,2})^\intercal) \\ &= \frac{c_1}{\ell_{T,1}^2 + \ell_{T,2}^2} \int_T \ell_{T,2}^2 (x_1 - x_{T,1})^2 - \ell_{T,1}^2 (x_2 - x_{T,2})^2 = 0.\end{aligned}$$

Hence

$$\begin{aligned}\|\Pi_h \boldsymbol{p} - \bar{\boldsymbol{p}}_h\|^2 &= (\Pi_h \boldsymbol{p} - \boldsymbol{p}, \boldsymbol{q}_h) + (\boldsymbol{p} - \bar{\boldsymbol{p}}_h, \boldsymbol{q}_h) \\ &= (\Pi_h \boldsymbol{p} - \boldsymbol{p}, \boldsymbol{q}_h) + (a \nabla_h (u - u_h), \boldsymbol{q}_h) \\ &:= I + II.\end{aligned} \tag{4.18}$$

By Lemma 3.1 with $k = 0$ in [30] and the Bramble–Hilbert lemma,

$$|I| \lesssim |\boldsymbol{p}|_{2,\Omega} \|\boldsymbol{q}_h\|_{0,\Omega}. \tag{4.19}$$

For part II, since $\operatorname{div}(\boldsymbol{q}_h|_T) = 0$, we have

$$\begin{aligned}II &= \sum_{T \in \mathcal{T}_h} \int_T a \nabla (u - \bar{u}_h) \cdot \boldsymbol{q}_h \\ &= \sum_{T \in \mathcal{T}_h} \int_T (\nabla (a(u - \bar{u}_h)) - (u - \bar{u}_h) \nabla a) \cdot \boldsymbol{q}_h \\ &= \sum_{T \in \mathcal{T}_h} \int_{\partial T} a(u - \bar{u}_h) \boldsymbol{q}_h \cdot \boldsymbol{n} - ((u - \bar{u}_h) \nabla a, \boldsymbol{q}_h) \\ &:= II_1 + II_2.\end{aligned} \tag{4.20}$$

II_2 is estimated by Lemma 4.1.2 and the apriori error estimate (4.4):

$$|II_2| \lesssim h^2 \|u\|_{2,\Omega} \|\boldsymbol{q}_h\|_{0,\Omega}. \tag{4.21}$$

Note that the normal component of $\{\boldsymbol{q}_h\}$ is constant on e and $[\![\boldsymbol{q}_h]\!] = 0$ by (4.7). It then follows from $\fint_e [\![\bar{u}_h]\!] = \boldsymbol{0}$, (4.6) , the trace inequality (1.9), an inverse inequality,

(4.4), and Lemma 4.1.2, that

$$
\begin{aligned}
II_1 &= \sum_{e\in\mathcal{E}_h}\int_e [\![a(u-\bar{u}_h)\boldsymbol{q}_h]\!] \\
&= \sum_{e\in\mathcal{E}_h}\int_e [\![(a-\fint_e a)(u-\bar{u}_h)]\!]\cdot\{\boldsymbol{q}_h\} \\
&\lesssim h\sum_{e\in\mathcal{E}_h}\|[\![u-\bar{u}_h]\!]\|_{0,e}\|\{\boldsymbol{q}_h\}\|_{0,e} \qquad (4.22)\\
&\lesssim h^{\frac{1}{2}}\sum_{e\in\mathcal{E}_h}\big(h^{-\frac{1}{2}}\|u-\bar{u}_h\|_{2,\omega_e}+h^{\frac{1}{2}}\|\nabla_h(u-\bar{u}_h)\|_{0,\omega_e}\big)\|\boldsymbol{q}_h\|_{0,\omega_e} \\
&\lesssim (\|u-\bar{u}_h\|_{0,\Omega}+h\|\nabla_h(u-\bar{u}_h)\|_{0,\Omega})\|\boldsymbol{q}_h\|_{0,\Omega}\lesssim h^2\|u\|_{2,\Omega}\|\boldsymbol{q}_h\|_{0,\Omega}.
\end{aligned}
$$

Combining (4.18)–(4.22), we obtain

$$\|\Pi_h\boldsymbol{p}-\bar{\boldsymbol{p}}_h\|_{0,\Omega}\lesssim h^2\|u\|_{3,\Omega}. \qquad (4.23)$$

Lemma 4.1.2 implies

$$\|\boldsymbol{p}_h-\bar{\boldsymbol{p}}_h\|_{0,\Omega}\lesssim h^2\|u\|_{2,\Omega}. \qquad (4.24)$$

The theorem follows from (4.23) and (4.24). □

Note that Q_h is in fact an element-by-element projection and $Q_h(a\nabla_h u_h) = a\nabla_h u_h$ if a is a piecewise constant.

4.1.2 Superconvergent recovery

Considering rectangular Raviart–Thomas elements, Durán in [30] proposed a postprocessing operator K_h^D satisfying:

$$\|K_h^D\boldsymbol{q}_h\|_{0,\Omega}\lesssim\|\boldsymbol{q}_h\|_{0,\Omega},\quad \forall\boldsymbol{q}_h\in\mathcal{RT}_h, \qquad (4.25a)$$

$$\|\boldsymbol{p}-K_h^D\Pi_h\boldsymbol{p}\|_{0,\Omega}\lesssim h^2|\boldsymbol{p}|_{2,\Omega}. \qquad (4.25b)$$

Here the input for K_h^D needs to be $H(\mathrm{div})$-conforming. Assume $\boldsymbol{p}_h \in \mathcal{RT}_h$, e.g., f is piecewise constant, $\boldsymbol{b} = \boldsymbol{0}$, and $c = 0$. Using (4.25), Theorem 4.1.3, and the triangle inequality $\|a\nabla u - K_h^D \boldsymbol{p}_h\|_{0,\Omega} \leq \|a\nabla u - K_h^D \Pi_h \boldsymbol{p}\|_{0,\Omega} + \|K_h^D(\Pi_h \boldsymbol{p} - \boldsymbol{p}_h)\|_{0,\Omega}$, we obtain $\|a\nabla u - K_h^D \boldsymbol{p}_h\|_{0,\Omega} \lesssim h^2 \|u\|_{3,\Omega}$. However, $\boldsymbol{p}_h \notin \mathcal{RT}_h$ in general and thus K_h^D cannot be directly applied to $\boldsymbol{p}_h$. In addition, the implementation of K_h^D requires exactly rectangular mesh structure. In this section, we introduce a simple recovery operator A_h by local weighted averaging. It also applies to quadrilateral meshes and any piecewise polynomial function.

Definition 4.1.1. The operator $A_h : \mathcal{RT}_h^{-1} \to \tilde{\mathcal{V}}_h$ is defined as follows.

1. For each $e \in \mathcal{E}_h^o$, let m be the midpoint of e. Let T and T' be the two rectangles sharing e as an edge. Define

$$(A_h \boldsymbol{q}_h)(m) := \frac{|T|}{|T| + |T'|} \boldsymbol{q}_h|_{T'}(m) + \frac{|T'|}{|T| + |T'|} \boldsymbol{q}_h|_T(m).$$

2. For each $e \in \mathcal{E}_h^\partial$, let m denote the midpoint of e and T the element having e as an edge. Let e' be the edge of T opposite to e with midpoint m'. Let T' be the other element having e' as an edge and m'' the midpoint of the edge of T' opposite to e'. Define

$$(A_h \boldsymbol{q}_h)(m) := ((A_h \boldsymbol{q}_h)(m') - w'(A_h \boldsymbol{q}_h)(m''))/w,$$

where

$$w = \frac{|T'|}{|T| + |T'|}, \quad w' = \frac{|T|}{|T| + |T'|}.$$

Then $A_h \boldsymbol{q}_h$ is the unique element in $\tilde{\mathcal{V}}_h$ with midpoint values are given in the above two steps.

Note that the weight constants in Definition 4.1.1 are not chosen in a standard way. We show that A_h has a super-approximation property on any nondegenerate rectangular meshes.

Theorem 4.1.4. *For $\boldsymbol{q}_h \in \mathcal{RT}_h^{-1}$ and $\boldsymbol{q} \in H(\mathrm{div};\Omega)$,*

$$\|A_h \boldsymbol{q}_h\|_{0,\Omega} \lesssim \|\boldsymbol{q}_h\|_{0,\Omega}, \tag{4.26a}$$

$$\|\boldsymbol{q} - A_h \Pi_h \boldsymbol{q}\|_{0,\Omega} \lesssim h^2 |\boldsymbol{q}|_{2,\Omega}. \tag{4.26b}$$

Proof. Consider $T \in \mathcal{T}_h$ and $\tilde{T} := \bigcup_{e \subset \partial T} \omega_e$. Using the stability of A_h in the L^∞-norm and the inverse inequality, we prove the stability of A_h in the L^2-norm:

$$\|A_h \boldsymbol{q}_h\|_{0,T} \lesssim h \|A_h \boldsymbol{q}_h\|_{0,\infty,T} \lesssim h \|\boldsymbol{q}_h\|_{0,\infty,\tilde{T}} \lesssim \|\boldsymbol{q}_h\|_{0,\tilde{T}}.$$

(4.26a) then follows from the above estimate and sum of squares.

Let $e \in \mathcal{E}_h^o$ with midpoint m and two adjacent elements T, T' sharing e. For $\boldsymbol{q}_1 \in Q_{1,1}(\omega_e) \times Q_{1,1}(\omega_e)$, we first want to show $(\boldsymbol{q}_1 - A_h \Pi_h \boldsymbol{q}_1)(m) = \mathbf{0}$. Since Π_h preserves functions in $Q_{1,0}(\omega_e) \times Q_{0,1}(\omega_e)$, it suffices to check when $\boldsymbol{q}_1 = (x_2, 0)^\intercal$ or $(0, x_1)^\intercal$. By linearity we can assume $m = \mathbf{0}$ without loss of generality. If e is a horizontal interior edge, let $T = [-\ell_1/2, \ell_1/2] \times [0, \ell_2]$, $T' = [-\ell_1/2, \ell_1/2] \times [-\ell_2', 0]$. Then,

$$\Pi_h \begin{pmatrix} x_2 \\ 0 \end{pmatrix} = \begin{cases} (\ell_2/2, 0)^\intercal & \text{if } x_2 > 0 \\ (-\ell_2'/2, 0)^\intercal & \text{if } x_2 < 0 \end{cases}, \quad \Pi_h \begin{pmatrix} 0 \\ x_1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

In either case, $(\boldsymbol{q}_1 - A_h \Pi_h \boldsymbol{q}_1)(m) = 0$. The same argument works for vertical interior edges.

Let $e \in \mathcal{E}_h^\partial$ and T the element having e as an edge. Let e' be the edge of T opposite to e and T' be the element sharing the edge e' with T. Let e'' be the edge of

T' opposite to e' and T'' be the element sharing e'' with T'. Let $\omega_e = T \cup T' \cup T''$. By similar argument, we have $(\boldsymbol{q}_1 - A_h \Pi_h \boldsymbol{q}_1)(m) = 0$ when $\boldsymbol{q}_1 \in Q_{1,1}(\omega_e) \times Q_{1,1}(\omega_e)$.

By the property derived in the above three paragraphs, for $\boldsymbol{q}_1 \in Q_{1,1}(\tilde{T}) \times Q_{1,1}(\tilde{T})$,

$$\begin{aligned}\|\boldsymbol{q} - A_h \Pi_h \boldsymbol{q}\|_{0,T} &\lesssim h \|\boldsymbol{q} - A_h \Pi_h \boldsymbol{q}\|_{0,\infty,T} \\ &\lesssim h \|(\mathrm{id} - A_h \Pi_h)(\boldsymbol{q} - \boldsymbol{q}_1)\|_{0,\infty,T} \lesssim h \|\boldsymbol{q} - \boldsymbol{q}_1\|_{0,\infty,\tilde{T}}.\end{aligned}$$

Then by standard finite element approximation theory,

$$\inf_{\boldsymbol{q}_1 \in Q_{1,1}(\tilde{T}) \times Q_{1,1}(\tilde{T})} \|\boldsymbol{q} - \boldsymbol{q}_1\|_{0,\infty,\tilde{T}} \lesssim h |\boldsymbol{q}|_{2,\tilde{T}}, \tag{4.27}$$

and thus

$$\|\boldsymbol{q} - A_h \Pi_h \boldsymbol{q}\|_{0,T} \lesssim h^2 |\boldsymbol{q}|_{2,\tilde{T}}. \tag{4.28}$$

Then (4.26b) follows from (4.28) and sum of squares. □

Combining Theorems 4.1.3 and 4.1.4, we obtain the superconvergent flux recovery estimate.

Theorem 4.1.5.

$$\|a \nabla u - A_h \boldsymbol{p}_h\|_{0,\Omega} \lesssim h^2 \|u\|_{3,\Omega},$$

where $\boldsymbol{p}_h$ is defined in Theorem 4.1.3.

Proof. This theorem readily follows from Theorems 4.1.3, 4.1.4 and the triangle inequality $\|a \nabla u - A_h \boldsymbol{p}_h\| \leq \|a \nabla u - A_h \Pi_h \boldsymbol{p}\| + \|A_h (\Pi_h \boldsymbol{p} - \boldsymbol{p}_h)\|$. □

Consider $\tilde{\boldsymbol{p}}_h \in \mathcal{RT}_h^{-1}$, where

$$\tilde{\boldsymbol{p}}_h|_K = Q_h(a \nabla_h u_h) - \boldsymbol{r}_h(f - \boldsymbol{b} \cdot \nabla_h u_h - c u_h)(\boldsymbol{x}_T).$$

Since $\boldsymbol{r}_h = O(h)$, we have $\|\tilde{\boldsymbol{p}}_h - \boldsymbol{p}_h\|_{0,\Omega} \lesssim h^2\|u\|_{2,\Omega}$. and thus

$$\|a\nabla u - A_h\tilde{\boldsymbol{p}}_h\|_{0,\Omega} \lesssim h^2\|u\|_{3,\Omega}.$$

$\tilde{\boldsymbol{p}}_h$ is favorable because of lower computational cost than $\boldsymbol{p}_h$.

4.1.3 Extension to the Crouzeix–Raviart element

In this subsection, we extend superconvergence analysis to triangular CR elements. Based on the equivalence between mixed and nonconforming methods for Poisson's equation, a superconvergent recovery for CR elements applied to Poisson's equation has been developed in [36]. Our technique here can be directly extended to CR elements for elliptic equations with lower order terms and variable coefficients. The CR finite element space is

$$\begin{aligned}\mathcal{V}_{g,h}^{CR} :=& \{v_h \in L^2(\Omega) : v_h|_T \in \text{span}\{1, x_1, x_2\} \text{ for all } T \in \mathcal{T}_h,\\ & v_h \text{ is continuous at the midpoint of each } e \in \mathcal{E}_h^o,\\ & \fint_e v_h = \fint_e g \text{ for all } e \in \mathcal{E}_h^\partial\}.\end{aligned}$$

The CR method for (4.1) is to find $u_h^{CR} \in \mathcal{V}_{g,h}^{CR}$, such that

$$(a\nabla_h u_h^{CR}, \nabla_h v) + (\boldsymbol{b}\cdot\nabla_h u_h^{CR}, v) + (cu_h^{CR}, v) = (f, v), \quad \forall v \in \mathcal{V}_{0,h}^{CR}.$$

Recall that $\mathcal{Q}_h^0$ is the lowest order triangular RT finite element space. The paper [53] (essentially) shows that CR and RT finite element spaces are closely related by the following lemma.

Lemma 4.1.6. *Let $\bar{f}$ and $\boldsymbol{q}_h$ be piecewise constant functions with respect to $\mathcal{T}_h$. Assume*

that $(\boldsymbol{q}_h, \nabla_h v) = (\bar{f}, v)$ *for all* $v \in \mathcal{V}_{0,h}^{CR}$. *Then* $\boldsymbol{\tau}_h - \bar{f}\boldsymbol{r}_h^{CR} \in \mathcal{Q}_h^0$, *with*

$$\boldsymbol{r}_h^{CR}|_T := \frac{1}{2}\left(x_1 - x_{T,1}, x_2 - x_{T,2}\right)^{\intercal}, \quad T \in \mathcal{T}_h,$$

where $(x_{T,1}, x_{T,2})$ *is the centroid of* T.

The next theorem follows from the supercloseness estimate in Chapter 2, and the same procedure in section 2.3.

Theorem 4.1.7. *Let* $\mathcal{T}_h$ *be a uniformly parallel mesh. Let* $\boldsymbol{p}_h^{CR} := \bar{a}\nabla_h u_h^{CR} - \boldsymbol{r}_h^{CR} P_h(f - cu_h^{CR} - \boldsymbol{b}\cdot\nabla_h u_h^{CR})$, *where* $\boldsymbol{r}_h^{CR}$ *is defined in Lemma 4.1.6 and* $\bar{a}|_T = \fint_T a$. *Then*

$$\|\Pi_h^0 \boldsymbol{p} - \boldsymbol{p}_h^{CR}\|_{0,\Omega} \lesssim h^2 |\log h|^{\frac{1}{2}} \|u\|_{3,\infty,\Omega}.$$

Proof. We use the same notation in Theorem 4.1.3. Let $\boldsymbol{q}_h = \Pi_h^0 \boldsymbol{p} - \boldsymbol{p}_h^{CR}$ and thus $\operatorname{div} \boldsymbol{q}_h = 0$. Hence $\boldsymbol{q}_h = \nabla^{\perp} w_h$ for some continuous piecewsie linear function w_h. The bound (4.19) for part I is replaced by

$$|(\boldsymbol{p} - \Pi_h^0 \boldsymbol{p}, \nabla^{\perp} w_h)| \lesssim h^2 |\log h|^{\frac{1}{2}} \|\boldsymbol{p}\|_{2,\infty,\Omega} \|\nabla^{\perp} w_h\|_{0,\Omega},$$

which is proved in Theorem 2.3.1. The rest of the proof is the same as Theorem 4.1.3. □

For the purpose of recovery, let

$$\begin{aligned}\mathcal{V}_h^{CR} :=& \{v_h \in L^2(\Omega) : v_h|_T \in \operatorname{span}\{1, x_1, x_2\} \text{ for all } T \in \mathcal{T}_h,\\ & v_h \text{ is continuous at the midpoint of each } e \in \mathcal{E}_h^o\}.\end{aligned}$$

Then we consider the postprocessing operator K_h defined in [12], see also [30].

Definition 4.1.2. Let $\boldsymbol{\tau}_h$ be a piecewise constant function.

1. For each $e \in \mathcal{E}_h^o$, let m be the midpoint of e. Let T and T' be the two rectangles sharing e as an edge. Define

$$(K_h \boldsymbol{q}_h)(m) := \frac{1}{2}\boldsymbol{q}_h|_T(m) + \frac{1}{2}\boldsymbol{q}_h|_{T'}(m).$$

2. For each $e \in \mathcal{E}_h^\partial$, let m denote the midpoint of e and T the element having e as an edge. Let e' be another edge of T with midpoint m'. Let T' be the other element having e' as an edge and m'' the midpoint of the edge of T' that is parallel to e. Define

$$(K_h \boldsymbol{q}_h)(m) := 2(K_h \boldsymbol{q}_h)(m') - (K_h \boldsymbol{q}_h)(m'').$$

Then $K_h \boldsymbol{\tau}_h$ is the unique element in $\mathcal{V}_h^{CR}$ with midpoint values given in the above two steps.

Based on Theorem 4.1.7, we obtain the superconvergent recovery for CR elements. For simplicity of the proof, we only consider uniformly parallel grids here.

Theorem 4.1.8. *Assume that $\mathcal{T}_h$ is uniformly parallel. Then*

$$\|a\nabla u - K_h(\bar{a}\nabla_h u_h^{CR})\|_{0,\Omega} \lesssim h^2 |\log h|^{\frac{1}{2}} \|u\|_{3,\infty,\Omega}.$$

Proof. The operator K_h is known to satisfy Theorem 4.1.4 with K_h replacing A_h, see [12]. It then follows from Theorem 4.1.7 and the same argument in the proof of Theorem 4.1.5 that

$$\|a\nabla u - K_h \boldsymbol{p}_h^{CR}\|_{0,\Omega} \lesssim h^2 |\log h|^{\frac{1}{2}} \|u\|_{3,\infty,\Omega}. \tag{4.29}$$

Let $p = f - cu - \boldsymbol{b} \cdot \nabla u$ and $\bar{\boldsymbol{p}}_h^{CR} := \bar{a}\nabla_h u_h^{CR} - \boldsymbol{r}_h^{CR} P_h p$. Clearly,

$$\|\boldsymbol{p}_h^{CR} - \bar{\boldsymbol{p}}_h^{CR}\|_{0,\Omega} \lesssim h^2 \|u\|_{2,\Omega}. \tag{4.30}$$

Let $e \in \mathcal{E}_h^o$ with midpoint m. By $(K_h \boldsymbol{r}_h^{CR})(m) = 0$, we have

$$\begin{aligned}
&(K_h \boldsymbol{r}_h^{CR} P_h p)(m) \\
&= (K_h \boldsymbol{r}_h^{CR} p)(m) + (K_h \boldsymbol{r}_h^{CR}(P_h p - p))(m) \\
&= (K_h \boldsymbol{r}_h^{CR})(m) p(m) + O(h^2)\|u\|_{2,\infty,\Omega} = O(h^2)\|u\|_{2,\infty,\Omega}.
\end{aligned}$$

A similar argument works for $e \in \mathcal{E}_h^\partial$. Hence

$$\|K_h \boldsymbol{r}_h^{CR} P_h p\|_{0,\Omega} \lesssim \|K_h \boldsymbol{r}_h^{CR} P_h p\|_{0,\infty,\Omega} \lesssim h^2 \|u\|_{2,\infty,\Omega}. \tag{4.31}$$

The theorem then follows from (4.29)–(4.31) and the triangle inequality

$$\begin{aligned}
\|a\nabla u - K_h(\bar{a}\nabla u_h^{CR})\|_{0,\Omega} &\le \|a\nabla u - K_h \boldsymbol{p}_h^{CR}\|_{0,\Omega} \\
&+ \|K_h(\boldsymbol{p}_h^{CR} - \bar{\boldsymbol{p}}_h^{CR})\|_{0,\Omega} + \|K_h \boldsymbol{r}_h^{CR} P_h p\|_{0,\Omega}.
\end{aligned}$$

□

4.2 Superconvergence of Morley elements

For the popular nonconforming method using the primal formulation (3.2) and Morley elements, superconvergence estimates are limited due to the very weak inter-element continuity. In fact, one purpose of superconvergence analysis of nonconforming or mixed methods for second order elliptic equations is to shed some light on superconvergence of nonconforming or mixed methods for fourth order elliptic problems, e.g.,

see [24, 62, 36] for superconvergence results of Wilson and Courzeix–Raviart nonconforming methods for second order elliptic equations. See also [12, 30, 46] for superconvergence results of Raviat–Thomas mixed methods for second order elliptic equations. Up to now, there have been several superconvergence results for the H^2-error of the displacement variable u using rectangular Morley elements and the formulation

$$(\nabla^2 u, \nabla^2 v) = (f, v), \quad \text{for all } v \in H_0^2(\Omega), \tag{4.32}$$

i.e., $\nu = 0$ in (3.2), cf. [47, 52, 40].

The HHJ mixed method is important partly because it is closely related to the popular Morley element, see the equivalence between the HHJ and Morley element methods in [3]. Using this equivalence, [36] then gives a superconvergent recovery estimate for the Morley element method using the formulation (4.32) under uniformly parallel grids.

To apply our superconvergence estimates of the HHJ mixed method in Chapter 3 to the Morley element for (3.2), we need to generalize the equivalence between Morley and HHJ methods based on (4.32) in [3] to the more general formulation (3.2). This interesting equivalence was originally proved in [3] for (4.32) by relaxing the inter-element continuity constraint in the HHJ method and introducing Lagrange multipliers for compensation. We extend it to (3.2) by giving a direct proof, see Theorem 4.2.1. Using the generalized equivalence, superconvergence of the Morley element method for the primal formulation (3.2) are direct corollaries of our results for the lowest order HHJ mixed method (3.9). Similar to the mixed case, our superconvergence estimate for Morley elements improves the results in [36] even under uniform grids. Theorem 4.2.1 is of independent interest. For example, [41] and [39] give convergence and optimality proofs for the adaptive HHJ and Morley element methods for (3.2),

respectively. By Theorem 4.2.1, it is possible to derive the proof for one method from the other. Combining the equivalence Theorem 4.2.1 and the multigrid algorithm for the HHJ method in [21], one can also obtain a multigrid algorithm for the Morley element method.

Let

$$\begin{aligned}\mathcal{M}_h := \{v :\ & v|_T \in \mathcal{P}_2(T), v \text{ is continuous at } \boldsymbol{z} \in \mathcal{N}_h^o \\ & \text{and vanishes at } \boldsymbol{z} \in \mathcal{N}_h^\partial, \partial_{\boldsymbol{n}} v \text{ is continuous at the} \\ & \text{midpoint of } e \in \mathcal{E}_h^o, \text{ and vanishes at the midpoint of } e \in \mathcal{E}_h^\partial\}.\end{aligned}$$

The Morley element method for solving (3.2) is to find $u_h^M \in \mathcal{M}_h$ such that

$$(\mathcal{A}\nabla_h^2 u_h^M, \nabla_h^2 v_h) = (f, v_h), \quad \forall v_h \in \mathcal{M}_h, \tag{4.33}$$

where ∇_h^2 is the piecewise Hessian w.r.t. $\mathcal{T}_h$. Arnold and Brezzi [3] proved that the lowest order HHJ method is equivalent to a modified Morley element method for the formulation $(\nabla^2 u, \nabla^2 v) = (f, v)$. In this section, we extend it to the general formulation (3.2). Recall that I_h is the nodal interpolation for linear elements.

Theorem 4.2.1. *Let $\bar{u}_h^M \in \mathcal{M}_h$ solve*

$$(\mathcal{A}\nabla_h^2 \bar{u}_h^M, \nabla_h^2 v_h) = (f, I_h v_h), \quad \forall v_h \in \mathcal{M}_h. \tag{4.34}$$

Then $(\mathcal{A}\nabla_h^2 \bar{u}_h^M, I_h \bar{u}_h^M)$ is the solution of the HHJ method (3.9).

Proof. Let $\bar{\boldsymbol{\sigma}}_h = \mathcal{A}\nabla_h^2 \bar{u}_h^M$. Integrating by parts element by element in (4.34), we have

$$\sum_{T \in \mathcal{T}_h} \int_{\partial T} \{(\bar{\boldsymbol{\sigma}}_h)_{nn} \partial_{\boldsymbol{n}} v_h + (\bar{\boldsymbol{\sigma}}_h)_{nt} \partial_{\boldsymbol{t}} v_h\} = (f, I_h v_h), \quad \forall v_h \in \mathcal{M}_h. \tag{4.35}$$

For $e \in \mathcal{E}_h^o$, let $v_e \in \mathcal{M}_h$ be the test function such that $\partial_{\boldsymbol{n}} v_e = 1$ at the midpoint of e, $\partial_{\boldsymbol{n}} v_e = 0$ at the midpoint of $e' \neq e$, and $v_e(\boldsymbol{z}) = 0$ for all $\boldsymbol{z} \in \mathcal{N}_h$. Then

$$\sum_{T \in \mathcal{T}_h} \int_{\partial T} (\bar{\boldsymbol{\sigma}}_h)_{nt} \partial_{\boldsymbol{t}} v_e = \sum_{K \in \mathcal{T}_h} (\bar{\boldsymbol{\sigma}}_h)_{nt} \int_{\partial T} \partial_{\boldsymbol{t}} v_e = 0,$$

$I_h v_e = 0$, and thus (4.35) with $v_h = v_e$ gives

$$\int_e \{(\bar{\boldsymbol{\sigma}}_h)_{nn}|_T - (\bar{\boldsymbol{\sigma}}_h)_{nn}|_{T'}\} = 0, \quad e \in \partial T \cap \partial T',$$

for all $e \in \mathcal{E}_h^o$, i.e., $\bar{\boldsymbol{\sigma}}_h \in \boldsymbol{\Sigma}_h$. Hence by the continuity of $\partial_{\boldsymbol{n}} v_h$ and $(\bar{\boldsymbol{\sigma}}_h)_{nn}$ at the midpoint of each edge and the midpoint quadrature rule, (4.35) implies

$$b_h(\bar{\boldsymbol{\sigma}}_h, v_h) := -\sum_{T \in \mathcal{T}_h} \int_{\partial T} (\bar{\boldsymbol{\sigma}}_h)_{nt} \partial_{\boldsymbol{t}} v_h = -(f, I_h v_h), \quad \forall v_h \in \mathcal{M}_h. \tag{4.36}$$

By $I_h \mathcal{M}_h = \mathcal{U}_h$, (4.36) and (3.11b), we have

$$b_h(\bar{\boldsymbol{\sigma}}_h, v_h) = -(f, v_h), \quad \forall v_h \in \mathcal{U}_h.$$

Using the formula (3.7) and the continuity of $\partial_{\boldsymbol{n}} \bar{u}_h^M$ at the midpoint of $e \in \mathcal{E}_h$,

$$\begin{aligned}(\mathcal{C}\bar{\boldsymbol{\sigma}}_h, \boldsymbol{\tau}_h) &= \sum_{K \in \mathcal{T}_h} \int_T \nabla^2 \bar{u}_h^M : \boldsymbol{\tau}_h = \sum_{T \in \mathcal{T}_h} \int_{\partial T} \operatorname{grad} \bar{u}_h^M \cdot \boldsymbol{\tau}_h \boldsymbol{n} \\ &= \sum_{T \in \mathcal{T}_h} \int_{\partial T} \partial_{\boldsymbol{t}} \bar{u}_h^M (\boldsymbol{\tau}_h)_{nt} = -b_h(\boldsymbol{\tau}_h, I_h \bar{u}_h^M), \quad \text{for all } \boldsymbol{\tau}_h \in \boldsymbol{\Sigma}_h,\end{aligned}$$

which completes the proof. □

Now we need a recovery operator to achieve superconvergence. Notice that $\mathcal{A}\nabla_h^2 u_h^M \notin \boldsymbol{\Sigma}_h$. Hence we cannot postprocess $\mathcal{A}\nabla_h^2 u_h^M$ directly by R_h in Chapter 3. However, we can simply apply the gradient recovery operator G_h defined by (1.5)

in Chapter 1 to it. The analysis of super-approximation property of postprocessing operators is local and independent of PDEs. Then following the analysis in [69], we obtain superconvergent recovery for Morley elements.

Consider the difference between (4.33) and (4.34), it's readily to check that

$$\|\nabla_h^2(\bar{u}_h^M - u_h^M)\|_{0,\Omega} \lesssim h^2\|f\|_{0,\Omega}.$$

Combining it with the triangle inequality

$$\|\boldsymbol{\sigma} - G_h\mathcal{A}\nabla_h^2 u_h^M\|_{0,\Omega} \leq \|\boldsymbol{\sigma} - G_h\mathcal{A}\nabla_h^2\bar{u}_h^M\|_{0,\Omega} + \|G_h\mathcal{A}\nabla_h^2(\bar{u}_h^M - u_h^M)\|_{0,\Omega},$$

and using Theorems 3.3.3 and 4.2.1, we obtain superconvergent recovery of the Morley element method (4.33).

Corollary 4.2.1.1. *Assume that $\mathcal{T}_h$ satisfies the (α,β)-condition. Then*

$$\|\mathcal{A}\nabla^2 u - G_h\mathcal{A}\nabla_h^2 u_h^M\|_{0,\Omega} \lesssim (1-\nu)^{-1}\rho_1(h,\alpha,\beta)(\|u\|_{3,\infty,\Omega} + \|u\|_{4,\Omega}).$$

The operator K_h proposed in [36] can also be used to postprocess $\mathcal{A}\nabla_h^2 u_h^M$. Replacing Q_h with K_h in Corollary 4.2.1.1, one obtain $\|\mathcal{A}\nabla^2 u - K_h\mathcal{A}\nabla_h^2 u_h^M\|_{L^2(\Omega)} = O(h^2|\log h|^{\frac{1}{2}})$ under uniformly parallel grids. If $\mathcal{C} = \mathcal{A}$ is the identity operator, Theorem 5.6 in [36] says that $\|\nabla^2 u - K_h\nabla_h^2 u_h^M\|_{2,\Omega} = O(h^{1.5})$ on uniformly parallel grids, which is suboptimal according to the aforementioned discussion.

Chapter 4, in part, has been submitted for publication of the material as it may appear in Numerische Mathematik, 2019, Li, Yu-Wen, Springer, 2019. The dissertation author was the author of this paper.

Bibliography

[1] Ainsworth, Mark; Oden, J. Tinsley A posteriori error estimation in finite element analysis. Pure and Applied Mathematics (New York). *John Wiley&Sons, New York*, 2000.

[2] Arbogast, Todd; Chen, Zhangxin On the implementation of mixed methods as nonconforming methods for second-order elliptic problems. *Math. Comp.* 64 (1995), no. 211, 943–972.

[3] Arnold, D. N.; Brezzi, F. Mixed and nonconforming finite element methods: implementation, postprocessing and error estimates. *RAIRO Modél. Math. Anal. Numér.* 19 (1985), no. 1, 7–32.

[4] Arnold, Douglas N.; Falk, Richard S.; Winther, Ragnar Multigrid in H(div) and H(curl). *Numer. Math.* 85 (2000), no. 2, 197–217.

[5] Arnold, Douglas N.; Falk, Richard S.; Winther, Ragnar Finite element exterior calculus, homological techniques, and applications. *Acta Numer.* 15 (2006), 1–155.

[6] Arnold, Douglas N.; Falk, Richard S.; Winther, Ragnar Finite element exterior calculus: from Hodge theory to numerical stability. *Bull. Amer. Math. Soc.* 47 (2010), no. 2, 281–354.

[7] Arnold, Douglas N.; Li, Lizao Finite element exterior calculus with lower-order terms. *Math. Comp.* 86 (2017), no. 307, 2193–2212.

[8] Babuka, I.; Osborn, J.; Pitkranta, J. Analysis of mixed methods using mesh dependent norms. *Math. Comp.* 35 (1980), no. 152, 1039–1062.

[9] Bank, Randolph E.; Xu, Jinchao Asymptotically exact a posteriori error estimators. I. Grids with superconvergence. *SIAM J. Numer. Anal.* 41 (2003), no. 6, 2294–2312.

[10] Bank, Randolph E.; Xu, Jinchao Asymptotically exact a posteriori error estimators. II. General unstructured grids. *SIAM J. Numer. Anal.* 41 (2003), no. 6, 2313–2332.

[11] Bank, Randolph E.; Xu, Jinchao; Zheng, Bin Superconvergent derivative recovery for Lagrange triangular elements of degree p on unstructured grids. *SIAM J. Numer. Anal.* 45 (2007), no. 5, 2032–2046.

[12] Brandts, Jan H. Superconvergence and a posteriori error estimation for triangular mixed finite elements. *Numer. Math.* 68 (1994), no. 3, 311–324.

[13] Brandts, Jan H. Superconvergence for triangular order k=1 Raviart-Thomas mixed finite elements and for triangular standard quadratic finite element methods. *Appl. Numer. Math.* 34 (2000), no. 1, 39–58.

[14] Brenner, Susanne C.; Scott, L. Ridgway The mathematical theory of finite element methods. Third edition. Texts in Applied Mathematics, 15. *Springer, New York,* 2008.

[15] Brezzi, Franco; Douglas, Jim, Jr.; Marini, L. D. Two families of mixed finite elements for second order elliptic problems. *Numer. Math.* 47 (1985), no. 2, 217–235.

[16] Beirao da Veiga, L.; Niiranen, J.; Stenberg, R. A posteriori error estimates for the Morley plate bending element. *Numer. Math.* 106 (2007), no. 2, 165–179.

[17] Brezzi, Franco; Fortin, Michel Mixed and hybrid finite element methods. Springer Series in Computational Mathematics, 15. *Springer-Verlag, New York,* 1991.

[18] Bramble, J. H.; Schatz, A. H. Higher order local accuracy by averaging in the finite element method. *Math. Comp.* 31 (1977), no. 137, 94–111.

[19] Chen Chuanmiao Structure theory of superconvergence of finite elements. *Hunan Science and Technology Press, Changsha,* 2002.

[20] Ciarlet, Philippe G. On Korn's inequality. *Chin. Ann. Math. Ser. B* 31 (2010), no. 5, 607–618.

[21] Chen, Long; Hu, Jun; Huang, Xuehai Multigrid methods for Hellan-Herrmann-Johnson mixed method of Kirchhoff plate bending problems. *J. Sci. Comput.* 76 (2018), no. 2, 673–696.

[22] Chen, Long; Holst, Michael; Xu, Jinchao Convergence and optimality of adaptive mixed finite element methods. *Math. Comp.* 78 (2009), no. 265, 35–53.

[23] Comodi M. I. The Hellan-Herrmann-Johnson method: some new error estimates and postprocessing. *Math. Comp.* 52 (1989), no. 185, 17–29.

[24] Chen, Hongsen; Li, Bo Superconvergence analysis and error expansion for the Wilson nonconforming finite element. *Numer. Math.* 69 (1994), no. 2, 125–140.

[25] Crouzeix, M.; Raviart, P.-A. Conforming and nonconforming finite element methods for solving the stationary Stokes equations. *RAIRO Anal. Numér.* 7 (1973), no. R-3, 33–75.

[26] Dörfler, Willy A convergent adaptive algorithm for Poisson's equation. *SIAM J. Numer. Anal.* 33 (1996), no. 3, 1106–1124.

[27] Douglas, Jim, Jr.; Roberts, Jean E. Global estimates for mixed methods for second order elliptic equations. *Math. Comp.* 44 (1985), no. 169, 39–52.

[28] Douglas, J., Jr.; Milner, F. A. Interior and superconvergence estimates for mixed methods for second order elliptic problems. *RAIRO Modél. Math. Anal. Numér.* 19 (1985), no. 3, 397–428.

[29] Dupont, Todd F.; Keenan, Philip T. Superconvergence and postprocessing of fluxes from lowest-order mixed methods on triangles and tetrahedra. *SIAM J. Sci. Comput.* 19 (1998), no. 4, 1322–1332.

[30] Durán, Ricardo Superconvergence for rectangular mixed finite elements. *Numer. Math.* 58 (1990), no. 3, 287–298.

[31] Ewing, Richard E.; Liu, Michael M.; Wang, Junping Superconvergence of mixed finite element approximations over quadrilaterals. *SIAM J. Numer. Anal.* 36 (1999), no. 3, 772–787.

[32] Ewing, R. E.; Lazarov, R. D.; Wang, J. Superconvergence of the velocity along the Gauss lines in mixed finite element methods. *SIAM J. Numer. Anal.* 28 (1991), no. 4, 1015–1029.

[33] Girault, Vivette; Raviart, Pierre-Arnaud Finite element methods for Navier-Stokes equations. Theory and algorithms. Springer Series in Computational Mathematics, 5. *Springer-Verlag, Berlin*, 1986.

[34] Grisvard, P. Singularities in boundary value problems. Research in Applied Mathematics, 22. Masson, Paris; *Springer-Verlag, Berlin*, 1992.

[35] Grisvard, Pierre Elliptic problems in nonsmooth domains. Classics in Applied Mathematics, 69. *SIAM, Philadelphia, PA*, 2011.

[36] Hu, Jun; Ma, Rui Superconvergence of both the Crouzeix-Raviart and Morley elements. *Numer. Math.* 132 (2016), no. 3, 491–509.

[37] Huang, Yunqing; Li, Jichun; Wu, Chao Averaging for superconvergence: verification and application of 2D edge elements to Maxwell's equations in metamaterials. *Comput. Methods Appl. Mech. Engrg.* 255 (2013), 121–132.

[38] Hu, Jun; Shi, Zhongci Constrained quadrilateral nonconforming rotated Q_1 element. *J. Comput. Math.*, 23 (2005), 561–586.

[39] Hu, Jun; Shi, Zhongci; Xu, Jinchao Convergence and optimality of the adaptive Morley element method. *Numer. Math.* 121 (2012), no. 4, 731–752.

[40] Hu, Jun; Shi, Zhongci; Yang, Xueqin Superconvergence of both two and three dimensional rectangular Morley elements for biharmonic equations. arXiv:1501.02424v1[math.NA].

[41] Huang, Jianguo; Huang, Xuehai; Xu, Yifeng Convergence of an adaptive mixed finite element method for Kirchhoff plate bending problems. *SIAM J. Numer. Anal.* 49 (2011), no. 2, 574–607.

[42] Huang, Yunqing; Xu, Jinchao Superconvergence of quadratic finite elements on mildly structured grids. *Math. Comp.* 77 (2008), no. 263, 1253–1268.

[43] Johnson, Claes On the convergence of a mixed finite-element method for plate bending problems. *Numer. Math.* 21 (1973), 43–62.

[44] Li, Bo Lagrange interpolation and finite element superconvergence. *Numer. Methods Partial Differential Equations* 20 (2004), no. 1, 33–59.

[45] Li, Bo; Zhang, Zhimin Analysis of a class of superconvergence patch recovery techniques for linear and bilinear finite elements. *Numer. Methods Partial Differential Equations* 15 (1999), no. 2, 151–167.

[46] Li, Yu-Wen Global superconvergence of the lowest-order mixed finite element on mildly structured meshes. *SIAM J. Numer. Anal.* 56 (2018), no. 2, 792–815.

[47] Lin, Jiafu; Lin, Qun Superconvergence of a finite element method for the biharmonic equation. *Numer. Methods Partial Differential Equations* 18 (2002), no. 3, 420–427.

[48] Lin, Qun; Xu, Jinchao Linear finite elements with high accuracy. *J. Comput. Math.* 3 (1985), no. 2, 115–133.

[49] Lin, Qun; Tobiska, Lutz; Zhou, Aihui Superconvergence and extrapolation of nonconforming low order finite elements applied to the Poisson equation. *IMA J. Numer. Anal.* 25 (2005), no. 1, 160–181.

[50] Lin, Qun; Yan, Ningning Global superconvergence for Maxwell's equations. *Math. Comp.* 69 (2000), no. 229, 159–176.

[51] Liu, Huipo; Yan, Ningning Superconvergence analysis of the nonconforming quadrilateral linear-constant scheme for Stokes equations. *Adv. Comput. Math.* 29 (2008), no. 4, 375–392.

[52] Mao, Shipeng; Shi, Zhong-ci High accuracy analysis of two nonconforming plate elements. *Numer. Math.* 111 (2009), no. 3, 407–443.

[53] Marini, Luisa Donatella An inexpensive method for the evaluation of the solution of the lowest order Raviart-Thomas mixed method. *SIAM J. Numer. Anal.* 22 (1985), no. 3, 493–496.

[54] Huang, Jianguo; Shi, Zhongci; Xu, Yifeng Finite element analysis for general elastic multi-structures. *Sci. China Ser. A* 49 (2006), no. 1, 109–129.

[55] Mekchay, Khamron; Nochetto, Ricardo H. Convergence of adaptive finite element methods for general second order linear elliptic PDEs. *SIAM J. Numer. Anal.* 43 (2005), no. 5, 1803–1827.

[56] Monk, Peter Finite element methods for Maxwell's equations. Numerical Mathematics and Scientific Computation. *Oxford University Press, New York*, 2003.

[57] Naga, Ahmed; Zhang, Zhimin A posteriori error estimates based on the polynomial preserving recovery. *SIAM J. Numer. Anal.* 42 (2004), no. 4, 1780–1800.

[58] Rannacher, R.; Turek, S. Simple nonconforming quadrilateral Stokes element. *Numer. Methods Partial Differential Equations* 8 (1992), no. 2, 97–111.

[59] Raviart, P.-A.; Thomas, J. M. A mixed finite element method for 2nd order elliptic problems. Mathematical aspects of finite element methods (Proc. Conf., Consiglio Naz. delle Ricerche (C.N.R.), Rome, 1975), pp. 292–315. Lecture Notes in Math., Vol. 606, *Springer, Berlin*, 1977.

[60] Schatz, A. H.; Sloan, I. H.; Wahlbin, L. B. Superconvergence in finite element methods and meshes that are locally symmetric with respect to a point. *SIAM J. Numer. Anal.* 33 (1996), no. 2, 505–521.

[61] Shafarevich, Igor R. Basic algebraic geometry. 1. Varieties in projective space. Third edition. *Springer, Heidelberg*, 2013.

[62] Shi, Zhong-Ci; Jiang, Bin; Xue, Weimin A new superconvergence property of Wilson nonconforming finite element. *Numer. Math.* 78 (1997), no. 2, 259–268.

[63] Thome, Vidar Galerkin finite element methods for parabolic problems. Second edition. Springer Series in Computational Mathematics, 25. *Springer-Verlag, Berlin,* 2006.

[64] Verfürth, Rüdiger A review of a posteriori error estimation and adaptive mesh-refinement techniques. Wiley-Teubner series, advances in numerical mathematics. *Wiley-Teubner*, 1996.

[65] Wang, Jun Ping Asymptotic expansions and L^∞-error estimates for mixed finite element methods for second order elliptic problems. *Numer. Math.* 55 (1989), no. 4, 401–430.

[66] Schatz, A. H.; Wahlbin, L. B. Interior maximum-norm estimates for finite element methods. II. *Math. Comp.* 64 (1995), no. 211, 907–928.

[67] Wahlbin, Lars B. Superconvergence in Galerkin finite element methods. Lecture Notes in Mathematics, 1605. *Springer-Verlag, Berlin*, 1995.

[68] Wu, Haijun; Zhang, Zhimin Can we have superconvergent gradient recovery under adaptive meshes? *SIAM J. Numer. Anal.* 45 (2007), no. 4, 1701–1722.

[69] Xu, Jinchao; Zhang, Zhimin Analysis of recovery type a posteriori error estimators for mildly structured grids. *Math. Comp.* 73 (2004), no. 247, 1139–1152.

[70] Ye, Xiu Superconvergence of nonconforming finite element method for the Stokes equations. *Numer. Methods Partial Differential Equations* 18 (2002), no. 2, 143–154.

[71] Zhang, Zhimin; Naga, Ahmed A new finite element gradient recovery method: superconvergence property. *SIAM J. Sci. Comput.* 26 (2005), no. 4, 1192–1213.

[72] Zienkiewicz, O. C.; Zhu, J. Z. A simple error estimator and adaptive procedure for practical engineering analysis. *Internat. J. Numer. Methods Engrg.* 24 (1987), no. 2, 337–357.

[73] Zienkiewicz, O. C.; Zhu, J. Z. The superconvergent patch recovery and a posteriori error estimates. I. The recovery technique. *Internat. J. Numer. Methods Engrg.* 33 (1992), no. 7, 1331–1364.

[74] Zienkiewicz, O. C.; Zhu, J. Z. The superconvergent patch recovery and a posteriori error estimates. II. Error estimates and adaptivity. Internat. *J. Numer. Methods Engrg.* 33 (1992), no. 7, 1365–1382.