ROBUST AND EXPLICIT A POSTERIORI ERROR ESTIMATION TECHNIQUES IN ADAPTIVE FINITE ELEMENT METHOD

A Dissertation

Submitted to the Faculty

of

Purdue University

by

Difeng Cai

In Partial Fulfillment of the

Requirements for the Degree

of

Doctor of Philosophy

August 2019

Purdue University

West Lafayette, Indiana

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF DISSERTATION APPROVAL

Dr. Zhiqiang Cai, Chair

Department of Mathematics

Dr. Peijun Li

Department of Mathematics

Dr. Guang Lin

Department of Mathematics

Dr. Jie Shen

Department of Mathematics

Approved by:

Dr. David Goldberg

Head for Graduate Studies, Department of Mathematics

To go

ACKNOWLEDGMENTS

First I would like to thank my advisor, Professor Zhiqiang Cai, for introducing me to the field of adaptive mesh refinement and for giving me plenty of freedom in doing research. I also thank my committee members and Professor Jianlin Xia for their support. It has always been an enjoyable experience discussing mathematics with factulty members in the department, including Professors Rodrigo Bañuelos, Steven Bell, Fabrice Baudoin, Marius Dadarlat, Bill Heinzer, Plamen Stefanov, Monica Torres, Changyou Wang, etc. I'm particularly grateful to Professor Changyou Wang for generously taking his time to discuss math problems. Thanks also go to Shannon Cassady in the graduate office, who is always there to unravel administrative work for us students.

Needless to say, I am deeply grateful to my family. My family members, e.g., my parents, grandparents, aunts and uncles, have provided the most pleasant environment since my birth, for which I will always be in debt. I feel lucky to be born and raised in a country whose culture (food, for example) is too rich for me to savour comprehensively. In short, I could never imagine a better family and a better place to be given birth to. From that perspective, I consider myself extremely lucky.

Last but not least, I have been fortunate enough to have met quite an exciting number of geniuses over the years, who belong to the union of the following (possibly overlapping) sets: high school classmates, college roommates, college classmates, graduate school roommates, graduate school classmates/colleagues. Being acquainted with those phenomenal human beings is by far my biggest achievement and my life has been complete (for multiple times already). How many people are able to meet their idols on a daily basis? How many people are able to talk to their idols on a daily basis? How many people are able to live and play with their idols all year round? From that perspective, I consider myself extremely successful.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SYMBOLS

η	an error estimator (usually denoting a residual-based one)
η_K	a local error indicator on K
ξ	an error estimator (usually denoting a hybrid or equilibrated one)
ξ_K	a local error indicator on K
$\|\| \cdot \|\|$	L^2 norm (unless otherwise stated)
$\|\| \cdot \|\|_S$	L^2 norm over a set S (unless otherwise stated)
Ω	a computational domain
$\partial\Omega$	boundary of Ω
Γ_D	Dirichlet boundary
Γ_N	Neumann boundary
∇	gradient operator
curl	curl operator
$\mathcal{T}$	a shape-regular triangulation of the computational domain (2D or 3D)
$\mathcal{N}$	the set of all vertices in $\mathcal{T}$
$\mathcal{E}$	the set of all edges($d = 2$)/faces($d = 3$) in $\mathcal{T}$
$\mathcal{E}_I$	the set of all interior edges ($d = 2$)/faces ($d = 3$) in $\mathcal{T}$
$\mathcal{E}_D$	the set of edges($d = 2$)/faces($d = 3$) on Γ_D
$\mathcal{E}_N$	the set of edges($d = 2$)/faces($d = 3$) on Γ_N
$\mathcal{E}_K$	the set of edges($d = 2$)/faces($d = 3$) in $K \in \mathcal{T}$
h_S	diameter of S
$\boldsymbol{n}_e$	a prescribed unit vector that is normal to $e \in \mathcal{E}$ and is pointing outward if e is on the boundary

s_K	a sign function defined on ∂K: $s_K\|_e = 1$ if $\boldsymbol{n}_e$ is an outward normal on $e \subset \partial K$; $s_K\|_e = -1$ if $\boldsymbol{n}_e$ is an inward normal on $e \subset \partial K$
P_k	set of polynomials with degrees less than or equal to k
$\mathrm{RT}_k(K)$	Raviart-Thomas space on K of index k
$\mathrm{BDM}_k(K)$	Brezzi-Douglas-Marini space on K of index k
$\Pi_{\mathrm{RT}_k(K)}$	the interpolation operator from $H(\mathrm{div}; K)$ to $\mathrm{RT}_k(K)$
$u_{\mathcal{T}}$	the finite element solution (unless otherwise stated)
$\boldsymbol{\sigma}_{\mathcal{T}}$	the numerical flux $\boldsymbol{\sigma}_{\mathcal{T}} := -A\nabla u_{\mathcal{T}}$
$\hat{\boldsymbol{\sigma}}_{\mathcal{T}}$	a recovered flux

ABSTRACT

Cai, Difeng PhD, Purdue University, August 2019. Robust And Explicit A Posteriori Error Estimation Techniques In Adaptive Finite Element Method . Major Professor: Zhiqiang Cai.

We present a comprehensive study of robust a posteriori error estimation for finite element approximations to elliptic partial differential equations. New results are presented for two types of estimators: the hybrid estimator and the equilibrated estimator. The framework of the hybrid estimator is introduced for two classes of problems: diffusion equations with discontinuous coefficients and stationary convection-diffusion-reaction equations with dominated convection/reaction. The hybrid estimator inherits all advantages of the residual estimator - explicit, general, robust- and is shown to be much more accurate according to extensive numerical experiments. Another kind of estimator - the flux equilibration-based estimator (or "equilibrated estimator" for simplicity) - is a second focus of the thesis. This kind of estimator is perfect for error control as it yields a guaranteed upper bound of the true error. We investiage equilibrated estimators for conforming, nonconforming, and discontinuous Galerkin discretizations of diffusion equations with discontinuous coefficients. For conforming elements, we present first a general procedure in both two and three dimensions for computing an equilibrated estimator via local minimizations and prove its robustness. We then propose an explicit equilibrated estimator for two dimensional problems and prove the robustness. This is the first robust equilibrated estimator that is explicit for conforming elements. Numerical results show that the new explicit estimator requires significantly less computational time than the state-of-the-art local minimization-based estimator. For nonconforming elements, we introduce an explicit equilibrated estimator for arbitrary order discretizations. The new estimator is robust and is independent of the odd/even order of the nonconforming finite element. This

is the first robust equilibrated estimator that enjoys the degree-independent property for nonconforming elements. Besides the design of error estimators, we discuss the quasi-monotonicity condition, a central assumption for deriving robust a posteriori error estimates for interface problems. We construct an example to show that the assumption is not only sufficient but also necessary for the robustness of equilibrated estimators. All results in the thesis hold for arbitrary order finite element approximations and for both two and three dimensional problems (unless otherwise stated).

1. INTRODUCTION

In numerical simulations, the oringal problem, governed by a differential equation, for example, is usually posed in an infinite dimensional space, which cannot be handled by a computer. Hence numerical solutions are sought after, where one discretizes the continuous model and solves the resulting finite dimensional system. However, once a numerical solution is computed, two questions arise:

1. how do we know that the computed solution is accurate enough ?
2. if the computed solution is not accurate, how to adapt the characteristics of the discrete model (e.g., mesh) to obtain a better solution ?

The first question is the very motivation for the so-called **a posteriori error estimation**, where one tries to estimate the error once a numerical solution is obtained. The second question can be tackled by using **Adaptive Mesh Refinement(AMR)**, identified as one of the two indispensable tools for computationally grand challenges by the US National Research Council [1]. AMR repeats the loop below until the desired accuracy is reached:

$$\text{SOLVE} \rightarrow \text{ESTIMATE} \rightarrow \text{MARK} \rightarrow \text{REFINE}.$$

The key step in AMR is ESTIMATE: the a posteriori error estimation. To see how a posteriori error estimation addresses the two questions above, we provide an example below. Consider solving an equation (with certain boundary conditions if needed)

$$\mathcal{L}u = f,$$

where $\mathcal{L}$ denotes a partial differential operator and f is given. Suppose an approximate solution $u_{\mathcal{T}}$ is computed based on certain numerical scheme (finite element method,

for example). The a posteriori error estimation aims to find a *computable* quantity η, called **a posteriori error estimator** or simply estimator, such that

$$||u - u_\tau|| \leq C_1 \eta, \tag{1.0.1}$$

$$\eta_K \leq C_2 ||u - u_\tau||_{\omega_K}, \tag{1.0.2}$$

where $\|\cdot\|$ denotes certain norm, K is an element in the mesh (a triangle or tetrahedron in 2D or 3D), ω_K denotes a local region around K, and η_K denotes the local error indicator on K (i.e., $\eta^2 = \sum_{K \in \mathcal{T}} \eta_K^2$). Here C_1 and C_2 are constants that are independent of u, u_τ, mesh-size but may depend on the shape parameter of the mesh, PDE coefficients, the polynomial degree of the approximation space, etc. The estimates in (1.0.1) and (1.0.2) imply that, up to certain constants, the estimator provides global upper bound and local lower bound of the true error. The upper bound (1.0.1) is sometimes referred as the *reliability* bound and the lower bound (1.0.2) the *efficiency* bound. In view of the two questions at the beginning, we see that they correspond to the reliability bound and efficiency bound in the a posteriori error estimation. That is, if one finds an error estimator η, then η (value over entire domain) indicates the global approximation accuracy (hence resolves the first question) and η_K (value on element K) indicates the local error on K, according to which one can then specify a set of elements (on which η_K is relatively large) for further refinement (hence resolves the second question). It is not hard to construct an estimator "intuitively". For example, one may define η_K as $\eta_K := \|f - \mathcal{L}u_\tau\|_K$, which measures the "residual" of the approximation. It is a computable quantity and can be used to indicate the error to some degree. However, it fails to yield a reliability bound (1.0.1) in general because it may happen that $\eta = 0$ while the true error is not (consider $f = 0$, $\mathcal{L} = \Delta$ and $u = 0$ while u_τ is a piecewise linear nonzero function over $\mathcal{T}$). Therefore, we see that an error estimator has to be judiciously chosen in order to be both reliable and efficient.

Generally speaking, the a posteriori error estimation is extremely challenging because the true solution is unknown and to find a computable quantity η that yields

the estimates (1.0.1) and (1.0.2) is highly nontrivial. More challenges arise when the underlying problem is not "nice": the operator $\mathcal{L}$ contains high-constrast coefficients; the solution u displays boundary or interior layers; the solution u contains singularities, etc. So how to evaluate the quality of an estimator or what are the desired properties of a good estimator?

1. For computational efficiency, an estimator should be easy to compute (much less costly than solving the original PDE), preferably with an explicit formula.

2. An estimator should be as general as possible. That is, it should be valid for arbitrary order finite element approximations, for both two and three dimensional problems, for large classes of equations, etc.

3. For challenging problems, the performance of an estimator is significantly affected by its *robustness*. Consider the example above where the operator $\mathcal{L}$ contains discontinuous coefficients. An estimator η is called *robust* (with respect to coefficient jump) if the constants C_1, C_2 in (1.0.1), (1.0.2) are independent of the coefficient jump.

In this thesis, we discuss two types of problems: diffusion problems with discontinuous coefficients and stationary convection-diffusion-reaction equations (with dominated convection/reaction). In terms of a posteriori error estimation, we focus on two types of estimators: hybrid estimators and equilibrated estimators. Though efforts have been made over the years and various estimators are proposed, for challenging problems, few estimators satisfy *all* the desired properties mentioned above, i.e., explicit, general, and most importantly, robust. In Tables 1.1 — 1.4, we summarize existing developments and new results in the thesis on the robust a posteriori error estimation. Table 1.1 is for diffusion equations with discontinuous coefficients, where *robust* means provable estimates as in (1.0.1), (1.0.2) with constants *independent* of the coefficient jump. Table 1.2 is for stationary convection-diffusion-reaction equations, where *robust* means provable estimates as in (1.0.1), (1.0.2) with constants *independent* of the sizes of convection and reaction as well as mesh-size (true error is

measured in the norm introduced in [2]). Table 1.3 presents error estimators based on flux equilibration, which will be termed as "equilibrated estimators" in this thesis. Table 1.4 summarizes error estimators for nonconforming finite element discretizations.

The major contributions of the thesis include the following:

1. set up the framework of the *hybrid estimator*, which is as general as the standard residual estimator. In fact, apart from the residual estimator, the hybrid estimator is the only explicit estimator with provable robustness for both diffusion equations with discontinuous coefficients and convection-diffusion-reaction equations with dominated convection/reaction (cf. Chapter 4 and Chapter 5). See also Table 1.1 and Table 1.2. Numerical experiments in Section 4.5 and Section 5.6 imply that the hybrid estimator is more accurate than residual estimator and for convection-diffusion-reaction equations, the hybrid estimator is much less sensitive to the size of convection/reaction.

2. for arbitrary order conforming finite element discretizations, give the first algorithm to compute an equilibrated flux in both 2D and 3D such that the resulting estimator is proved to be robust for interface problems (cf. Chapter 6);

3. for arbitrary order conforming finite element discretizations, propose for the first time an *explicit* equilibrated estimator that is robust with respect to coefficient jump in two dimensional interface problems (cf. Chapter 7);

4. for arbitrary order nonconforming finite element discretizations, propose the first explicit and robust equilibrated estimator that is *independent* of the odd/even order of the nonconforming element (cf. Chapter 8);

5. give a counter example to demonstrate for the first time that the quasi-monotonicity condition on the diffusion coefficient is not only sufficient but also *necessary* for the robustness of equilibrated estimators for interface problems (cf. 10).

Table 1.1.
Comparison of existing error estimators for diffusion equations with *discontinuous* coefficients (*conforming* finite element approximation)

estimator	robust	explicit	higher order elements
ZZ [3, 4]	no	yes	no
ZZ-variants [5–7]	no	no	yes
residual [8, 9]	yes	yes	yes
equilibrated [10]	yes	no	yes
$H(\mathrm{div})$ projection [11]	unknown	no	yes
hybrid [12](Chapter 4)	yes	yes	yes

Table 1.2.
Comparison of existing error estimators for stationary convection-diffusion-reaction equations with dominated convection/reaction (conforming finite element approximation)

estimator	robust	explicit	higher order elements
ZZ [3, 4]	no	yes	no
ZZ-variants [5–7]	no	no	yes
residual [2]	yes	yes	yes
$H(\mathrm{div})$ projection [11]	unknown	no	yes
flux recovery [13]	special case: zero reaction	yes	no
hybrid [14] (Chapter 5)	yes	yes	yes

Table 1.3.
Comparison of *equilibrated* estimators for diffusion equations with discontinuous coefficients (*conforming* finite element approximation)

estimator	robust	explicit	higher order elements	3D
[15]	no	no	no	no
[16]	no	no	yes	yes
[17, 18]	no	yes	no	no
[19]	no	no	yes	no
[10] or [20] (Chapter 6)	yes	no	yes	yes
[21] (Chapter 7)	yes	yes	yes	no

Table 1.4.
Comparison of *equilibrated* estimators for diffusion equations with discontinuous coefficients (*nonconforming* finite element approximation)

estimator	robust	explicit	higher order elements	even/odd degree-independent
[22]	yes	yes	yes	no
[23]	no	yes	no	no
[24] (Chapter 8)	yes	yes	yes	yes

The rest of the thesis is organized as follows. In Chapter 2, we introduce function spaces, operators and prove some results on the decomposition of vector fields that will be used to derive guaranteed reliability of equilibrated estimators. Chapter 3 specifies the notation concerning the adaptive mesh refinement. In Chapter 4 and Chapter 5, we present the framework of the hybrid estimator for diffusion equations with discontinuous coefficients and stationary convection-diffusion-reaction equations, respectively. Robust estimates are derived and various numerical examples are presented to demonstrate the advantages over residual estimators. In Chapter 6 and Chapter 7, we study the design of equilibrated estimators for conforming finite element discretizations of interface problems. Chapter 6 presents a general procedure in both 2D and 3D to compute equilibrated estimator based on local minimizations and proves the robustness with respect to coefficient jump. Chapter 7 proposes an *explicit* equilibrated estimator for two dimensional interface problems and gives a straightforward proof of the robustness, which is entirely different from Chapter 6. In Chapter 8, we study nonconforming finite element discretizations of interface problems and propose an explicit, robust, equilibrated estimator that is independent of the even/odd polynomial degree in the nonconforming finite element space. In Chapter 9, we present an explicit, robust, equilibrated estimator for discontinuous Galerkin discretizations of the interface problems. In Chapter 10, we investigate the quasi-monotonicity condition, which is used in existing literatures to derive robust a posteriori error estimates for interface problems. A counter example is constructed to show that the condition is not only sufficient but also necessary for the robustness of equilibrated estimators. In Chapter 11, we discuss possible misconceptions in PDEs, numerical solutions, and a posteriori error estimations. Figure 1.1 gives a schematic illustration of the major part of the thesis.

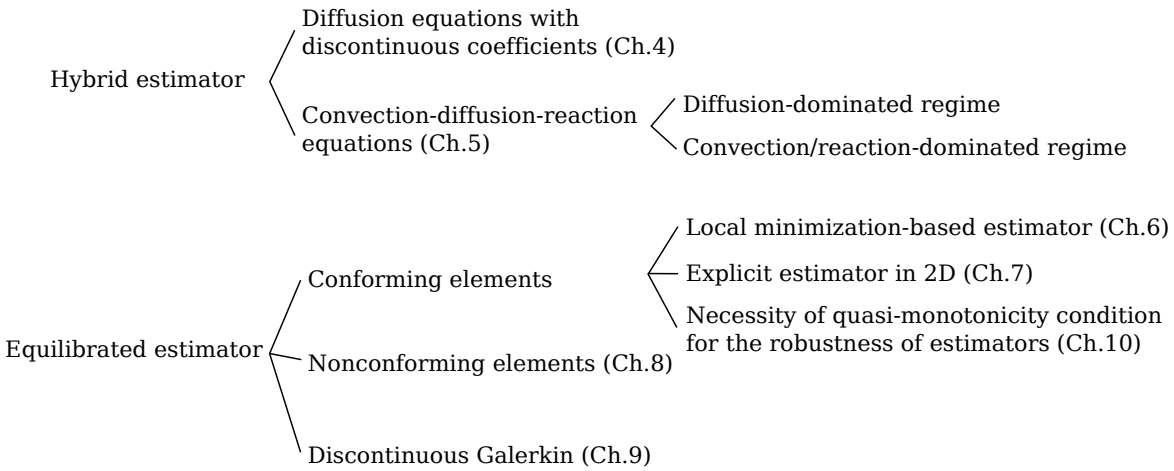

Figure 1.1. Structural overview of the topics discussed in the thesis

2. FUNCTION SPACES, OPERATORS AND DECOMPOSITIONS

2.1 Function Spaces

Assume in this section that Ω is a bounded Lipschitz domain in $\mathbb{R}^d (d = 2, 3)$. If necessary, the boundary of Ω, denoted by $\partial\Omega$, will be further assumed to be piecewise smooth.

Let $\boldsymbol{n}$ denote the unit outward normal on the boundary of a region of interest. In two dimensions, if $\boldsymbol{n} = (n_1, n_2)$ is the unit outward normal, then denote by $\boldsymbol{t}$ the unit tangent given by $\boldsymbol{t} = (-n_2, n_1)$.

Let $\mathcal{D}(\Omega)$ denote the space of infinitely differentiable functions with compact support on Ω. The dual space of $\mathcal{D}(\Omega)$ is denoted by $\mathcal{D}'(\Omega)$, often called the space of distributions. Derivatives of a distribution can be defined via *integration by parts* (cf. [25]). For $s > 0$, define

$$H^s(\mathbb{R}^d) = \left\{ v \in L^2(\mathbb{R}^d) : (1 + |\xi|^2)^{s/2} \hat{v}(\xi) \in L^2(\mathbb{R}^d) \right\}$$

with norm

$$\|v\|_{H^s(\mathbb{R}^d)} = \left(\|v\|^2_{L^2(\mathbb{R}^d)} + \|(1 + |\xi|^2)^{s/2} \hat{v}(\xi)\|^2_{L^2(\mathbb{R}^d)} \right)^{1/2},$$

where $\hat{v}$ denotes the Fourier transform of v. The space $H^s(\Omega)$ is given by restriction:

$$H^s(\Omega) = \left\{ v \in L^2(\Omega) : \exists \tilde{v} \in H^s(\mathbb{R}^d) \text{ such that } v = \tilde{v}|_\Omega \right\},$$

where the norm is

$$\|v\|_{H^s(\Omega)} = \inf_{\tilde{v} \in H^s(\mathbb{R}^d), \tilde{v}|_\Omega = v} \|\tilde{v}\|_{H^s(\mathbb{R}^d)}.$$

Under the above assumptions on Ω, it is well-known that the Bessel potential space $H^s(\Omega)$ is identitcal to the Sobolev space $W^{s,2}(\Omega)$ defined by Sobolev (if s is an integer) or Sobolev-Slobodeckij (if s is non-integer) norms ([26, Ch.16]). Define

$$H_0^s(\Omega) = \overline{\mathcal{D}(\Omega)}^{H^s(\Omega)},$$

i.e., the closure of $\mathcal{D}(\Omega)$ in $H^s(\Omega)$. Let $H^{-s}(\Omega)$ be the dual space of $H_0^s(\Omega)$. If $s = m$ is a positive integer, then $H^{-m}(\Omega)$ can be characterized below (cf. [25, Lemma 1.2]).

Proposition 2.1.1. *A distribution f belongs to $H^{-m}(\Omega)$ if and only if there exist functions $f_\alpha \in L^2(\Omega)$ for $|\alpha| \le m$, such that*

$$f = \sum_{|\alpha| \le m} D^\alpha f_\alpha,$$

where $\alpha = (\alpha_1, \ldots, \alpha_d)$ denotes a multi-index and $D^\alpha f_\alpha = \frac{\partial^{|\alpha|} f_\alpha}{\partial x_1^{\alpha 1} \ldots \partial x_d^{\alpha d}}$.

Now we define spaces on the boundary $\partial\Omega$. Define

$$L^2(\partial\Omega) = \left\{ v : \int_{\partial\Omega} |v|^2 ds < \infty \right\}$$

with norm

$$\|v\|_{L^2(\partial\Omega)} = \left(\int_{\partial\Omega} |v|^2 ds \right)^{1/2}.$$

Define

$$H^{1/2}(\partial\Omega) = \left\{ v \in L^2(\partial\Omega) : \exists \tilde{v} \in H^1(\Omega) \text{ such that } v = \tilde{v}|_{\partial\Omega} \right\}.$$

In general, $H^s(\partial\Omega)$ can be defined using an atlas of $\partial\Omega$ assuming the boundary $\partial\Omega$ is smooth enough (in terms of s). See, for example, [25, 26] for details. For $s \in (\frac{1}{2}, \frac{3}{2})$, the trace operator $\gamma : H^s(\Omega) \to H^{s-1/2}(\partial\Omega)$ is bounded (cf. [27]).

2.2 Differential Operators and Vector Fields

The operator ∇ is defined by:

$$\nabla u = (\frac{\partial u}{\partial x_1}, \ldots, \frac{\partial u}{\partial x_d}).$$

The operator div is defined by:

$$\operatorname{div} \boldsymbol{v} = \sum_{i=1}^{d} \frac{\partial v_i}{\partial x_i},$$

where $\boldsymbol{v} = (v_1, \ldots, v_d)$.

The operator curl is defined below:

- in $\mathbb{R}^2$,
$$\begin{cases} \operatorname{curl} \phi = (\dfrac{\partial \phi}{\partial x_2}, -\dfrac{\partial \phi}{\partial x_1}), & \phi \in H^1(\Omega), \\ \operatorname{curl} \boldsymbol{v} = \dfrac{\partial v_2}{\partial x_1} - \dfrac{\partial v_1}{\partial x_2}, & \boldsymbol{v} = (v_1, v_2). \end{cases}$$
- in $\mathbb{R}^3$,
$$\operatorname{curl} \boldsymbol{v} = \left(\frac{\partial v_3}{\partial x_2} - \frac{\partial v_2}{\partial x_3}, \ \frac{\partial v_1}{\partial x_3} - \frac{\partial v_3}{\partial x_1}, \ \frac{\partial v_2}{\partial x_1} - \frac{\partial v_1}{\partial x_2} \right).$$

Simply speaking, for a vector field $\boldsymbol{v} = (v_1, \ldots, v_d)$, $\operatorname{curl} \boldsymbol{v}$ can be obtained by computing the differential form $d(\sum_{i=1}^{d} v_i dx_i)$.

In two dimensions, if we define the rotation operator rot by

$$\operatorname{rot} : \mathbb{R}^2 \to \mathbb{R}^2, \quad (v_1, v_2) \to (-v_2, v_1),$$

then the following relations hold

$$\boldsymbol{t} = \operatorname{rot} \boldsymbol{n}, \quad \operatorname{curl} = \operatorname{rot} \circ \nabla \quad \text{and} \quad \operatorname{curl} = -\operatorname{div} \circ \operatorname{rot}, \tag{2.2.1}$$

where the first curl acts on a scalar function and the second curl acts on a vector field.

To specify spaces with boundary conditions, assume that $\partial\Omega$ is splitted into the Dirichlet boundary Γ_D and the Neumann boundary Γ_N, where Γ_D is assumed to be a closed non-empty subset of $\partial\Omega$ and $\Gamma_N = \partial\Omega \backslash \Gamma_D$. Define the subspaces of $H^1(\Omega)$:

$$\begin{aligned} H^1_D(\Omega) &= \left\{ v \in H^1(\Omega) : v|_{\Gamma_D} = 0 \right\}, \\ H^1_{t,N}(\Omega) &= \left\{ v \in H^1(\Omega) : \nabla v \cdot \boldsymbol{t}|_{\Gamma_N} = 0 \right\}, \quad \text{if } d = 2. \end{aligned} \tag{2.2.2}$$

Define $H(\mathrm{div};\Omega)$ and its subspaces by:

$$\begin{aligned}
H(\mathrm{div};\Omega) &= \left\{\boldsymbol{v} \in L^2(\Omega)^d : \mathrm{div}\,\boldsymbol{v} \in L^2(\Omega)\right\},\\
H_0(\mathrm{div};\Omega) &= \left\{\boldsymbol{v} \in H(\mathrm{div};\Omega) : \boldsymbol{v}\cdot\boldsymbol{n}|_{\partial\Omega} = 0\right\},\\
\mathrm{Ker}(\mathrm{div}) &= \left\{\boldsymbol{v} \in H(\mathrm{div};\Omega) : \mathrm{div}\,\boldsymbol{v} = 0\right\},\\
\mathrm{Ker}_0(\mathrm{div}) &= \left\{\boldsymbol{v} \in H(\mathrm{div};\Omega) : \mathrm{div}\,\boldsymbol{v} = 0, \boldsymbol{v}\cdot\boldsymbol{n}|_{\partial\Omega} = 0\right\},\\
\mathrm{Ker}_D(\mathrm{div}) &= \left\{\boldsymbol{v} \in H(\mathrm{div};\Omega) : \mathrm{div}\,\boldsymbol{v} = 0, \boldsymbol{v}\cdot\boldsymbol{n}|_{\Gamma_D} = 0\right\},\\
\mathrm{Ker}_N(\mathrm{div}) &= \left\{\boldsymbol{v} \in H(\mathrm{div};\Omega) : \mathrm{div}\,\boldsymbol{v} = 0, \boldsymbol{v}\cdot\boldsymbol{n}|_{\Gamma_N} = 0\right\}.
\end{aligned} \tag{2.2.3}$$

Define $H(\mathrm{curl};\Omega)$ and its subspaces by:

$$\begin{aligned}
H(\mathrm{curl};\Omega) &= \begin{cases} \left\{\boldsymbol{v} \in L^2(\Omega)^d : \mathrm{curl}\,\boldsymbol{v} \in L^2(\Omega)\right\}, & \text{if } d = 2\\ \left\{\boldsymbol{v} \in L^2(\Omega)^d : \mathrm{curl}\,\boldsymbol{v} \in L^2(\Omega)^d\right\}, & \text{if } d = 3, \end{cases}\\
H_0(\mathrm{curl};\Omega) &= \left\{\boldsymbol{v} \in H(\mathrm{curl};\Omega) : \gamma_t\boldsymbol{v}|_{\partial\Omega} = 0\right\},\\
H_N(\mathrm{curl};\Omega) &= \left\{\boldsymbol{v} \in H(\mathrm{curl};\Omega) : (\mathrm{curl}\,\boldsymbol{v})\cdot\boldsymbol{n}|_{\Gamma_N} = 0\right\}, \quad \text{if } d = 3,\\
\mathrm{Ker}(\mathrm{curl}) &= \left\{\boldsymbol{v} \in H(\mathrm{curl};\Omega) : \mathrm{curl}\,\boldsymbol{v} = 0\right\},\\
\mathrm{Ker}_D(\mathrm{curl}) &= \left\{\boldsymbol{v} \in \mathrm{Ker}(\mathrm{curl}) : \gamma_t\boldsymbol{v}|_{\Gamma_D} = 0\right\},\\
\mathrm{Ker}_N(\mathrm{curl}) &= \left\{\boldsymbol{v} \in \mathrm{Ker}(\mathrm{curl}) : \gamma_t\boldsymbol{v}|_{\Gamma_N} = 0\right\},
\end{aligned} \tag{2.2.4}$$

where, as in [25, Theorem 2.11], the tangential component operator γ_t is defined by

$$\begin{aligned}
\gamma_t &: \boldsymbol{v} \to \boldsymbol{v}\cdot\boldsymbol{t}|_{\partial\Omega}, \quad \text{if } d = 2,\\
\gamma_t &: \boldsymbol{v} \to \boldsymbol{v}\times\boldsymbol{n}|_{\partial\Omega}, \quad \text{if } d = 3.
\end{aligned} \tag{2.2.5}$$

In two dimensions, $H(\mathrm{curl};\Omega)$ is isomorphic to $H(\mathrm{div};\Omega)$ due to (2.2.1) [28, Chapter 2].

For notational convenience, if $W \subseteq H^1(\Omega)$, then we define

$$\nabla W := \left\{\nabla u : u \in W\right\}.$$

Similar definitions apply to curl, div, etc.

2.3 Decompositions of Vector Fields

Assume in this section that Ω is a bounded simply-connected Lipschitz domain in $\mathbb{R}^d$. The following orthogonal decomposition of $L^2(\Omega)^d$ is well known (cf., [25, Theorem 2.7]):

Theorem 2.3.1.

$$L^2(\Omega)^d = \nabla H^1(\Omega) \oplus \mathrm{Ker}_0(\mathrm{div}). \tag{2.3.1}$$

Moreover, the space $\nabla H^1(\Omega)$ is known to admit the following characterization (cf., [25, Theorem 2.9]):

Theorem 2.3.2. *The mapping* $\nabla : H^1(\Omega)/\mathbb{R} \to \mathrm{Ker}(\mathrm{curl})$ *defines a bijection. Consequently,*

$$\nabla H^1(\Omega) = \mathrm{Ker}(\mathrm{curl}). \tag{2.3.2}$$

If boundary conditions are imposed on Γ_D and Γ_D is assumed to be connected, then it is easy to obtain the following based on Theorem 2.3.2.

Corollary 2.3.1. *If* Γ_D *is a connected subset of* $\partial\Omega$*, then the mapping* $\nabla : H^1_D(\Omega)/\mathbb{R} \to \mathrm{Ker}_D(\mathrm{curl})$ *defines a bijection. Consequently,*

$$\nabla H^1_D(\Omega) = \mathrm{Ker}_D(\mathrm{curl}).$$

Note that if Γ_D is not connected, then in general

$$\nabla H^1_D(\Omega) \subsetneqq \mathrm{Ker}_D(\mathrm{curl}).$$

A example can be easily constructed as follows. Let Ω be the unit square with edges $e_1, \ldots, e_4$, where e_1, e_2 are vertical edges(closed) and e_3, e_4 are horizontal edges. Define $\Gamma_D = e_1 \cup e_2$ and $\Gamma_N = \partial\Omega \backslash \Gamma_D$. Choose a function $u \in H^1(\Omega)$ such that

$$u|_{e_1} = 0, \quad u|_{e_2} = 1, \quad \text{and} \quad \nabla u \cdot \boldsymbol{n}|_{\Gamma_N} = 0.$$

It is easy to see that $\nabla u \in \mathrm{Ker}_D(\mathrm{curl})$ and $\nabla u \notin \nabla H^1_D(\Omega)$.

In view of the decomposition theorem 2.3.1, apart from the characterization of $\nabla H^1(\Omega)$ in Theorem 2.3.2, it is known that $\mathrm{Ker}_0(\mathrm{div})$ also has alternative characterizations, which depend on dimension d.

Since Ω is simply-connected, $\partial\Omega$ is connected. The following characterization holds for two dimensional vector fields (cf., [25, Corollary 3.1]).

Theorem 2.3.3. *If $d = 2$, then*

$$\mathrm{Ker}_0(\mathrm{div}) = \mathrm{curl}\, H_0^1(\Omega) \tag{2.3.3}$$

and

$$\mathrm{curl} : H_0^1(\Omega) \to \mathrm{Ker}_0(\mathrm{div})$$

is an isomorphism, where for $\phi \in H_0^1(\Omega)$, $\mathrm{curl}\,\phi = \boldsymbol{v} \in \mathrm{Ker}_0(\mathrm{div})$ if and only if

$$(\mathrm{curl}\,\phi, \mathrm{curl}\, q) = (\boldsymbol{v}, \mathrm{curl}\, q), \quad \forall\, q \in H_0^1(\Omega). \tag{2.3.4}$$

The characterization for three dimensional vector fields is given below (cf., [25, Theorem 3.6]).

Theorem 2.3.4. *If $d = 3$, then*

$$\mathrm{Ker}_0(\mathrm{div}) = \mathrm{curl}\,\Phi, \tag{2.3.5}$$

where

$$\Phi = \{\boldsymbol{\phi} \in H_0(\mathrm{curl}; \Omega) : \mathrm{div}\,\boldsymbol{\phi} = 0\}. \tag{2.3.6}$$

As Ω is simply-connected, it holds that (cf., [25, Theorem 3.1, Theorem 3.4]).

Theorem 2.3.5.

$$\mathrm{Ker}(\mathrm{div}) = \mathrm{curl}\,(H^1(\Omega)), \quad \textit{if } d = 2,$$

and

$$\mathrm{Ker}(\mathrm{div}) = \mathrm{curl}\,(H^1(\Omega)^3) = \mathrm{curl}\,(H(\mathrm{curl}; \Omega)), \quad \textit{if } d = 3.$$

Theorem 2.3.2, Theorem 2.3.5 and the surjectivity of div from $H(\mathrm{div}; \Omega)$ to $L^2(\Omega)$ as well as the isomorphism in (2.2.1) between curl and div (in two dimensions) yield the following well-known result (cf., [28, Chapter 11.1]).

Theorem 2.3.6. *The following two sequences*

$$\mathbb{R} \xrightarrow{\text{id}} H^1(\Omega) \xrightarrow{\nabla} H(\text{curl};\Omega) \xrightarrow{\text{curl}} L^2(\Omega) \xrightarrow{0} 0 \quad (d=2)$$

and

$$\mathbb{R} \xrightarrow{\text{id}} H^1(\Omega) \xrightarrow{\nabla} H(\text{curl};\Omega) \xrightarrow{\text{curl}} H(\text{div};\Omega) \xrightarrow{\text{div}} L^2(\Omega) \xrightarrow{0} 0 \quad (d=3)$$

are exact. Here ∇ *in the first exact sequence can be replaced by* curl *as a result of the isomorphism in* (2.2.1).

2.3.1 Results for mixed boundary conditions

This subsection presents decompositions of vector fields when mixed boundary conditions are imposed, which will be used later for the a posteriori error estimation for nonconforming elements (and also conforming elements). The main result is stated in Theorem 2.3.8. First we present a decomposition of $L^2(\Omega)^d$ similar to Theorem 2.3.1.

Theorem 2.3.7.

$$L^2(\Omega)^d = \nabla H_0^1(\Omega) \oplus \text{Ker}(\text{div}). \tag{2.3.7}$$

Proof. Let $X = \nabla H_0^1(\Omega)$. Then

$$L^2(\Omega)^d = X \oplus X^\perp,$$

where $X^\perp$ denotes the orthogonal complement of X in $L^2(\Omega)^d$. It follows from integration by parts that $\text{Ker}(\text{div}) \subseteq X^\perp$. We then complete the proof by showing that $X^\perp \subseteq \text{Ker}(\text{div})$. In fact, for any $\boldsymbol{q} \in X^\perp$, there holds

$$(\boldsymbol{q}, \nabla u) = 0, \quad \forall\, u \in H_0^1(\Omega).$$

Note that $\text{div}\,\boldsymbol{q} \in H^{-1}(\Omega)$ and it then follows

$$(\text{div}\,\boldsymbol{q}, u) = -(\boldsymbol{q}, \nabla u) = 0, \quad \forall\, u \in H_0^1(\Omega),$$

which implies $\text{div}\,\boldsymbol{q} = 0$. □

The main result of this section concerns a decomposition of $L^2(\Omega)^d$ when mixed boundary conditions are taken into account.

Theorem 2.3.8.

$$L^2(\Omega)^d = \nabla H^1_D(\Omega) \oplus \mathrm{Ker}_N(\mathrm{div}). \tag{2.3.8}$$

Proof. Let $X = \nabla H^1_D(\Omega)$. The fact that $\mathrm{Ker}_N(\mathrm{div}) \subseteq X^\perp$ follows from integration by parts. For an arbitrary $\boldsymbol{q} \in X^\perp$, as in the proof of Theorem 2.3.7, we deduce that $\mathrm{div}\,\boldsymbol{q} = 0$. Therefore,

$$(\boldsymbol{q}\cdot\boldsymbol{n}, v)_{\Gamma_N} = (\boldsymbol{q}, \nabla v) + (\mathrm{div}\,\boldsymbol{q}, v) = 0, \quad \forall\, v \in H^1_D(\Omega).$$

Hence $\boldsymbol{q}\cdot\boldsymbol{n}|_{\Gamma_N} = 0$ and $\boldsymbol{q} \in \mathrm{Ker}_N(\mathrm{div})$. □

The so-called Prager-Synge Theorem (cf. [29] or [17, Theorem 5.1]) is an immediate consequence of Theorem 2.3.8.

Corollary 2.3.2 (Prager-Synge Theorem)**.** *Let u be the solution to the boundary value problem:*

$$\begin{cases} -div(A\nabla u) = f, & in\ \ \Omega, \\ u = g_{_D}, & on\ \ \Gamma_D, \\ -A\nabla u\cdot\boldsymbol{n} = g_{_N}, & on\ \ \Gamma_N. \end{cases} \tag{2.3.9}$$

If $u_{_\mathcal{T}} \in H^1(\Omega)$ with $u_{_\mathcal{T}} = g_{_D}$ on Γ_D and $\boldsymbol{\sigma} \in H(\mathrm{div};\Omega)$ satisfies

$$\mathrm{div}\,\boldsymbol{\sigma} = f \ \ in\ \Omega \quad and \quad \boldsymbol{\sigma}\cdot\boldsymbol{n} = g_{_N} \ \ on\ \Gamma_N,$$

then

$$\|A^{1/2}\nabla(u - u_{_\mathcal{T}})\|^2 + \|A^{1/2}\nabla u + A^{-1/2}\boldsymbol{\sigma}\|^2 = \|A^{1/2}\nabla u_{_\mathcal{T}} + A^{-1/2}\boldsymbol{\sigma}\|^2.$$

In particular,

$$\|A^{1/2}\nabla(u - u_{_\mathcal{T}})\| \le \|A^{1/2}\nabla u_{_\mathcal{T}} + A^{-1/2}\boldsymbol{\sigma}\|. \tag{2.3.10}$$

Proof. Since $\nabla(u-u_{\mathcal{T}}) \in \nabla H^1_D(\Omega)$ and $A\nabla u+\boldsymbol{\sigma} \in \mathrm{Ker}_N(\mathrm{div})$, Theorem 2.3.8 implies that they are orthogonal (with respect to the L^2 inner product). Hence $A^{1/2}\nabla(u-u_{\mathcal{T}})$ and $A^{-1/2}(A\nabla u+\boldsymbol{\sigma})$ are orthogonal. Consequently,

$$\|A^{1/2}\nabla(u-u_{\mathcal{T}})\|^2+\|A^{-1/2}(A\nabla u+\boldsymbol{\sigma})\|^2=\|A^{1/2}\nabla u_{\mathcal{T}}+A^{-1/2}\boldsymbol{\sigma}\|^2,$$

which completes the proof. □

The inequality (2.3.10) is used extensively in a posteriori error estimation, where $u_{\mathcal{T}}$ denotes a finite element solution, $\boldsymbol{\sigma}$ is an equilibrated flux computed from the reconstruction or postprocessing, and $\|A^{1/2}\nabla u_{\mathcal{T}}+A^{-1/2}\boldsymbol{\sigma}\|$ is defined as the estimator. Then (2.3.10) shows that the estimator is always an upper bound of the true error (in energy norm) with constant being exactly 1, often termed as the *guaranteed* reliability bound. This kind of estimator, i.e., based on equilibrated flux reconstruction, will be elaborated in Chapter 6 — Chapter 9.

Below we present some auxiliary results. Using Theorem 2.3.5, another characterization of $\mathrm{Ker}_N(\mathrm{div})$ can be obtained below.

Theorem 2.3.9.

$$\mathrm{Ker}_N(\mathrm{div})=\begin{cases}\mathrm{curl}\,(H^1_{t,N}(\Omega)), & \text{if } d=2,\\ \mathrm{curl}\,(H_N(\mathrm{curl};\Omega)), & \text{if } d=3,\end{cases}$$

where the spaces $H^1_{t,N}(\Omega)$ and $H_N(\mathrm{curl};\Omega)$ are defined in (2.2.2) *and* (2.2.4)*, respectively.*

Theorem 2.3.8 and Theorem 2.3.9 yield the following decomposition of $L^2(\Omega)^d$:

Corollary 2.3.3.

$$L^2(\Omega)^d=\begin{cases}\nabla H^1_D(\Omega)\oplus\mathrm{curl}\,(H^1_{t,N}(\Omega)), & \text{if } d=2,\\ \nabla H^1_D(\Omega)\oplus\mathrm{curl}\,(H_N(\mathrm{curl};\Omega)), & \text{if } d=3.\end{cases}\tag{2.3.11}$$

2.3.2 Decomposition of error in nonconforming finite element

Let $u \in H^1_D(\Omega)$ and u_τ be piecewise H^1 with respect to mesh $\mathcal{T}$. Denote by ∇_h the gradient operator defined piecewisely with respect to mesh $\mathcal{T}$. Let u_1 be the projection of u_τ into $H^1_D(\Omega)$ given by

$$(A\nabla u_1, \nabla v) = (A\nabla_h u_\tau, \nabla v), \quad \forall\, v \in H^1_D(\Omega). \tag{2.3.12}$$

Then u_τ can be written as

$$u_\tau = u_1 + u_0 \tag{2.3.13}$$

with

$$(A\nabla_h u_0, \nabla v) = 0, \quad \forall\, v \in H^1_D(\Omega). \tag{2.3.14}$$

Thus we know from Theorem 2.3.8 that

$$A\nabla_h u_0 \in \mathrm{Ker}_N(\mathrm{div}).$$

Writing

$$A\nabla_h(u_\tau - u) = A\nabla(u_1 - u) + A\nabla_h u_0, \tag{2.3.15}$$

we are able to obtain the following result with the help of Theorem 2.3.8, which will be the key to the guaranteed a posteriori error estimation for nonconforming elements in Chapter 8. Theorem 2.3.10 can be viewed as a generalization of the Prager-Synge Theorem in Corollary 2.3.2 for *piecewise* H^1 functions.

Theorem 2.3.10. *Let u be the solution of (4.1.1) and u_τ be a function that is piecewise H^1 with respect to mesh $\mathcal{T}$. Let u_τ be decomposed as in (2.3.13) with $u_1 \in H^1_D(\Omega)$ in (2.3.12). Then for any vector field $\boldsymbol{\tau}$ with*

$$\boldsymbol{\tau} \in \Sigma_f := \{\boldsymbol{\tau} \in H(\mathrm{div};\Omega) : \mathrm{div}\,\boldsymbol{\tau} = f,\ \boldsymbol{\tau}\cdot\boldsymbol{n}|_{\Gamma_N} = g_{_N}\},$$

the following identities hold:

$$\begin{aligned} \|A^{1/2}\nabla(u_1 - u)\|^2 &= \left(A\nabla_h u_\tau + \boldsymbol{\tau}, \nabla(u_1 - u)\right), \\ \|A^{1/2}\nabla_h u_0\| &= \inf_{v\in H^1_D(\Omega)} \|A^{1/2}\nabla_h(u_\tau - v)\|. \end{aligned} \tag{2.3.16}$$

In particular,

$$\begin{aligned}\|A^{1/2}\nabla_h(u_{\mathcal{T}}-u)\|^2 &= \|A^{1/2}\nabla(u_1-u)\|^2+\|A^{1/2}\nabla_h u_0\|^2\\ &\le \inf_{\boldsymbol{\tau}\in\Sigma_f}\|A^{1/2}\nabla_h u_{\mathcal{T}}+A^{-1/2}\boldsymbol{\tau}\|^2+\inf_{v\in H^1_D(\Omega)}\|A^{1/2}\nabla_h(u_{\mathcal{T}}-v)\|^2.\end{aligned} \tag{2.3.17}$$

Proof. The second identity in (2.3.16) follows from (2.3.14). We next prove the first identity in (2.3.16). Since $\boldsymbol{\tau}+A\nabla u\in \mathrm{Ker}_N(\mathrm{div})$ and $A\nabla_h u_0\in \mathrm{Ker}_N(\mathrm{div})$, we deduce from Theorem 2.3.8 that

$$\begin{aligned}(A\nabla(u_1-u),\nabla(u_1-u)) &= (A\nabla u_1+\boldsymbol{\tau},\nabla(u_1-u))\\ &= (A\nabla_h u_{\mathcal{T}}-A\nabla_h u_0+\boldsymbol{\tau},\nabla(u_1-u))\\ &= (A\nabla_h u_{\mathcal{T}}+\boldsymbol{\tau},\nabla(u_1-u)).\end{aligned}$$

(2.3.17) follows immediately from (2.3.14) and (2.3.16). □

Let $\mathcal{S}$ be the interpolation operator defined in [22] so that $\mathcal{S}(u_{\mathcal{T}})\in H^1_D(\Omega)$. Then the so-called nonconforming error on K can be estimated by the quantity

$$\xi_{nc,K}=\|A^{1/2}\nabla_h(u_{\mathcal{T}}-\mathcal{S}(u_{\mathcal{T}}))\|_K.$$

It is known that (cf. [22, Lemma 4.2]), under the monotonicity condition of diffusion coefficient,

$$\xi_{nc,K}\le C\|A^{1/2}\nabla_h(u_{\mathcal{T}}-u)\|_{\hat{\omega}_K},$$

where $\hat{\omega}_K$ denotes the union of elements sharing a common vertex with K and C is a constant independent of the jump of the diffusion coefficient.

Remark 2.3.1. *Particularly, in two dimensions, combining* (2.3.14) *with Theorem 2.3.8 and Theorem 2.3.9, we see that*

$$A\nabla_h u_0=\mathrm{curl}\,\phi,\quad \textit{for some } \phi\in H^1_{t,N}(\Omega).$$

Hence the decomposition in (2.3.15) *becomes*

$$A\nabla_h(u_{\mathcal{T}}-u)=A\nabla(u_1-u)+\mathrm{curl}\,\phi.$$

2.3.3 Remarks on Helmholtz decompositions

The so-called Helmholtz decomposition of vector fields in $L^2(\Omega)^d$ is used extensively in the a posteriori error estimation for nonconforming elements. Below we give some comments on the proofs concerning Helmholtz decompositions.

- In [30, (3.6)], the proof of the Helmholtz decomposition concerns a boundary value problem for w, where the Neumann boundary data is $a\nabla_j e \cdot n$. (∇_j denotes the piecewise gradient in [30].) Note that in general the vector field $a\nabla_j e \notin H(\mathrm{div};\Omega)$, so the trace $a\nabla_j e$ on Γ_2 is *not* well-defined. As a result, the meaning of the normal component $a\nabla_j e \cdot n$ on Γ_2 is questionable. (Recall that (cf. [26, Lemma 20.2]) for $\boldsymbol{\tau} \in H(\mathrm{div};\Omega)$, $\boldsymbol{\tau}\cdot\boldsymbol{n} \in H^{-1/2}(\partial\Omega)$ and the mapping is surjective.) The result in [30, (3.6)] is cited in a series of papers, including [22, 31, 32].

- In [31, Lemma 3.1], the proof made use of the Green's formula for a vector field $\boldsymbol{w} \in L^2(\Omega)^2$. However, since $\boldsymbol{w} \notin H(\mathrm{div};\Omega)$ in general, it is non-trivial to interpret the two integrals $\int_\Omega v\mathrm{div}\,\boldsymbol{w}dx$ and $\int_\Gamma v\boldsymbol{w}\cdot\boldsymbol{n}ds$ in the "Green's formula" ($v \in H^1(\Omega)$ vanishes on the Dirichlet boundary Γ_D). So far, we can only find Green's formula for vector fields in $H(\mathrm{div};\Omega)$ (see, e.g., [25, 2.17]). Same to the issue discussed above, the normal component $\boldsymbol{w}\cdot\boldsymbol{n}$ is not well-defined for $\boldsymbol{w} \in L^2(\Omega)^2$.

Note that the Helmholtz decompositions in [30, (3.6)], [31, (3.1)], [32, (2.19)], [22, (2.35)], [33, Theorem 3.15] are for two dimensional problems only. For the general case, i.e., in d dimensions with mixed boundary conditions, the Helmholtz decomposition is stated in Corollary 2.3.3, similar to [34, Lemma 2.1].

When deriving error estimates for nonconforming elements in Section 2.3.10, we did not use the Helmholtz decomposition in Corollary 2.3.3. Instead, we utilized the more general result in Theorem 2.3.8, which holds true in arbitrary dimensions, while the Helmholtz decomposition is stated in at most three dimensions.

3. ADAPTIVE MESH REFINEMENT

3.1 Mesh

Let Ω be a bounded polygonal domain in $\mathbb{R}^d (d = 2, 3)$ whose boundary $\partial\Omega$ is composed of two disjoint parts Γ_D and Γ_N with $\text{meas}(\Gamma_D) > 0$. Let $\mathcal{T}$ be a shape-regular triangulation of Ω (see, e.g., [35]), where $K \in \mathcal{T}$ is either a triangle ($d = 2$) or a tetrahedron ($d = 3$). Define the following sets associated with $\mathcal{T}$:

$$
\begin{aligned}
\mathcal{N}: &\ \text{the set of all vertices,} \\
\mathcal{E}: &\ \text{the set of all edges } (d = 2)/\text{faces } (d = 3), \\
\mathcal{E}_I: &\ \text{the set of all interior edges } (d = 2)/\text{faces } (d = 3), \\
\mathcal{E}_D: &\ \text{the set of edges } (d = 2)/\text{faces } (d = 3) \text{ on } \Gamma_D, \\
\mathcal{E}_N: &\ \text{the set of edges } (d = 2)/\text{faces } (d = 3) \text{ on } \Gamma_N, \\
\mathcal{E}_K: &\ \text{the set of edges } (d = 2)/\text{faces } (d = 3) \text{ in } K \in \mathcal{T}.
\end{aligned}
$$

- ω_z: the union of elements that share the vertex z;
- ω_e: the union of elements adjacent to $e \in \mathcal{E}$;
- ω_K: the union of elements that share at least one edge ($d = 2$) or one face ($d = 3$) with K;
- $\hat{\omega}_K$: the union of elements that share at least one vertex with K.

For each $e \in \mathcal{E}$, associate a unit normal vector $\boldsymbol{n}_e$, where $\boldsymbol{n}_e$ is chosen as the unit outward normal if $e \in \mathcal{E}_N$. For an interior edge ($d = 2$) /face ($d = 3$) $e \in \mathcal{E}_I$, denote by K_e^+ and K_e^- the two elements sharing e such that the unit outward normal of K_e^+ on e coincides with $\boldsymbol{n}_e$. For $e \in \mathcal{E}_I \cap \mathcal{E}_K$, let K_e' denote the neighbor of K that shares e.

Let h_S denote the diameter of $S \in \mathcal{T} \cup \mathcal{E}$. $|K|$ and $|e|$ denote the area ($d = 2$) / volume ($d = 3$) of an element K and length of an edge e ($d = 2$) / area of a face e ($d = 3$), respectively.

Assuming the context is clear, $\boldsymbol{n}$ always denotes the unit outward normal vector on ∂K. For $e \in \mathcal{E}_K$, in order to determine if the unit normal $\boldsymbol{n}_e = \boldsymbol{n}$, we define below a sign function in $L^2(\partial K)$:

$$s_{_K} : \partial K \to \{-1, 1\}, \quad s_{_K}|_e = \begin{cases} 1, & \text{if } \boldsymbol{n}_e \text{ is an outward normal on } \partial K, \\ -1, & \text{if } \boldsymbol{n}_e \text{ is an inward normal on } \partial K. \end{cases}$$

That is to say, $\boldsymbol{n} = s_{_K}|_e \boldsymbol{n}_e$.

3.2 Adaptivity

To solve a problem numerically, the adaptive mesh refinement(AMR) algorithms loop in the following way until the prescribed accuracy is reached:

$$\text{Solve} \to \text{Estimate} \to \text{Mark} \to \text{Refine}.$$

In the context of numerically solving partial differential equations(PDEs),

- the *Solve* stage involves the discretization of the underlying PDE and the solution of the resulting matrix problem;
- the *Estimate* stage is the key in the AMR loop, where one estimates (both globally and locally) the error between the exact PDE solution and the numerical solution computed from the *Solve* stage;
- the *Mark* stage consists in marking elements or subregions where the local error is relatively large and this stage is performed only when the global error estimate from *Estimate* stage is larger than the prescribed tolerance;
- the *Refine* stage performs refinement of the elements or subregions marked from the *Mark* stage.

The *Estimate* stage performs what is called the *a posteriori error estimation*. Since the exact solution is in general unknown a priori, it is a very challenging task to obtain an accurate a posteriori error estimate of the true error .

3.3 Spaces

For $S \in \mathcal{T} \cup \mathcal{E}$, $P_k(S)$ denotes the set of polynomials over S with degrees less than or equal to k.

On a simplex $K \in \mathcal{T}$, let $\mathrm{RT}_{k-1}(K)$ and $\mathrm{BDM}_k(K)$ be the Raviart-Thomas space of index $k-1$ ($k \geq 1$) and the Brezzi-Douglas-Marini space of index k, respectively. Namely,

$$\mathrm{RT}_{k-1}(K) := P_{k-1}(K)^d + \boldsymbol{x} P_{k-1}(K) \quad \text{and} \quad \mathrm{BDM}_k(K) := P_k(K)^d. \tag{3.3.1}$$

Let $\mathcal{T}$ be a regular triangulation of Ω. The $H(\mathrm{div};\Omega)$-conforming and the broken Raviart-Thomas spaces of index $k-1$ are then given by

$$\mathrm{RT}_{k-1} := \{\boldsymbol{\tau} \in H(\mathrm{div};\Omega) : \boldsymbol{\tau}|_K \in \mathrm{RT}_{k-1}(K),\ \forall\, K \in \mathcal{T}\} \tag{3.3.2}$$

and

$$\mathrm{RT}_{k-1}^{-1} := \{\boldsymbol{\tau} \in L^2(\Omega)^d : \boldsymbol{\tau}|_K \in \mathrm{RT}_{k-1}(K),\ \forall\, K \in \mathcal{T}\}, \tag{3.3.3}$$

respectively. Analogously, the counterparts for the Brezzi-Douglas-Marini spaces of index k are given by

$$\mathrm{BDM}_k := \{\boldsymbol{\tau} \in H(\mathrm{div};\Omega) : \boldsymbol{\tau}|_K \in \mathrm{BDM}_k(K),\ \forall\, K \in \mathcal{T}\}$$

and

$$\mathrm{BDM}_k^{-1} := \{\boldsymbol{\tau} \in L^2(\Omega)^d : \boldsymbol{\tau}|_K \in \mathrm{BDM}_k(K),\ \forall\, K \in \mathcal{T}\},$$

respectively.

4. HYBRID ESTIMATOR FOR DIFFUSION EQUATIONS

A version of this chapter has been published in [12].

Adaptive mesh refinement is necessary in the discretization of partial differential equations (PDEs) in order to handle computational challenges [1]. A posteriori error estimates play a crucial role in adaptive mesh refinement, where one tries to estimate the error by computing quantities (called error estimators) based on numerical solution as well as data from the underlying PDE. It is well known that the explicit residual error estimators (see, e.g., [8, 16, 36–38]) are computationally inexpensive with applications to a large class of problems. Moreover, for computationally challenging problems such as interface problems, proper weighted residual estimators (see, e.g., [8, 9]) generate efficient meshes. However, it is also known that residual estimators usually overestimate the true error by a large margin compared to estimators of the Zienkiewicz-Zhu (ZZ) type (cf. [39]). In this chapter, we introduce the so-called "hybrid a posteriori error estimator" for the conforming finite element method. The name "hybrid" comes from the fact that the estimator is defined as a sum of two terms: the flux error term and the modified element residual term, where the flux term is commonly seen in recovery-based estimators and the modified residual term is inspired by the standard residual estimator. The explicit hybrid estimator is proved to be robust, reliable on all meshes (unlike the ZZ-type estimators), valid for higher order finite element approximations, and numerical results indicate that it is more accurate than the standard residual estimator.

The Zienkiewicz-Zhu (ZZ) estimator [3] is defined as the L^2 norm of the difference between the recovered and the numerical gradients, where the recovered gradient is a continuous piecewise linear vector field computed by averaging local numerical gradients. Due to its simplicity, universality, and asymptotic exactness for smooth problems, the ZZ estimator enjoys a high popularity in the engineering community

(cf. [4,38,40,41]). Extensions and variants of the ZZ estimator to higher order elements have been studied recently (see [5–7], etc.). Those estimators are often shown to display the superconvergence property when the exact solution is smooth enough. However, ZZ-type estimators are known to have several major drawbacks, especially for non-smooth problems. Firstly, adaptive mesh refinement (AMR) algorithms using the ZZ estimator are not efficient to reduce global error for non-smooth problems, e.g., interface problems (see, e.g., [42]). By exploring the mathematical structure of the underlying problem, [43] identified the reason for this failure and proposed a robust estimator for the lowest order finite element discretization. Secondly, estimators of the ZZ type are not reliable on coarse meshes relative to the underlying problem. A simple one-dimensional example in [16] shows that the ZZ estimator equals zero but the true error is arbitrarily large. A two-dimensional example is presented in Section 11.3. Moreover, unlike the original easy-to-compute ZZ estimator [3, 4], the extensions or variants for higher order elements are not explicit, as they usually require solving either a global problem or many local least square problems (cf. [5–7]). To ensure the reliability on coarse meshes and the applicability for higher order elements, [11] introduced a new estimator by simply adding an appropriately weighted element residual to the flux error, where the recovered flux is the L^2 projection of the numerical flux in an $H(\mathrm{div})$-conforming space. A major drawback, however, is that the computation of the recovered flux requires solving a global problem and hence is more expensive than the residual estimator.

The purpose of this chapter is to introduce an explicit flux recovery in an $H(\mathrm{div})$-conforming space so that the resulting hybrid error estimator is robust and is more accurate than the residual estimator with similar computational cost and applicability. The recovered flux is chosen to satisfy a compatible divergence equation on each element. In particular, we are able to derive an explicit formula for a recovered flux in an $H(\mathrm{div})$-conforming space and the formula is automatically valid for higher order finite element approximations. Unlike existing ZZ-type estimators, which are not reliable on coarse meshes and are not robust with respect to discontinuous diffusion

coefficients, the hybrid estimator is proved to be reliable on all meshes and robust with respect to diffusion coefficient jump.

The hybrid estimator displays a strong connection to the explicit residual estimator as we can prove that the proposed estimator is actually equivalent to the residual estimator [8] with constants independent of the diffusion coefficients (see Section 4.3). As a result, the robustness of the residual estimator with respect to the coefficient jump carries over to the hybrid estimator. Despite the theoretical equivalence, numerical results show that the hybrid estimator is more accurate than the residual estimator. Hence the hybrid estimator can be viewed as a substitute of the residual estimator with an improved accuracy. The innate link to the residual estimator lends comparable generality to the hybrid estimator and the technique can also be applied to convection-diffusion-reaction problems (see Chapter 5).

The chapter is organized as follows. In Section 4.1, we introduce the model problem with a conforming finite element discretization and some notation. In Section 4.2, we present the hybrid estimator. The equivalence between the hybrid estimator and the standard residual estimator is established in Section 4.3. Section 4.4 summarizes the robust a posteriori error estimates, which follow immediately from the equivalence result in Section 4.3. Numerical results are presented in Section 4.5 to demonstrate the performance of the hybrid estimator, including its advantage over the standard residual estimator.

4.1 Problem And Discretization

Let Ω be a bounded polygonal domain in $\mathbb{R}^d$ $(d = 2, 3)$ with Lipschitz boundary $\partial\Omega$, where $\partial\Omega = \bar{\Gamma}_D \cup \bar{\Gamma}_N$, $\Gamma_D \cap \Gamma_N = \emptyset$ and $\text{meas}(\Gamma_D) > 0$. Consider the diffusion equation

$$\begin{cases} -\text{div}(A\nabla u) = f, & \text{in } \Omega, \\ u = 0, & \text{on } \Gamma_D, \\ -A\nabla u \cdot \boldsymbol{n} = g_N, & \text{on } \Gamma_N. \end{cases} \tag{4.1.1}$$

Here, $f \in L^2(\Omega)$, $g_{_N} \in L^2(\Gamma_N)$, $\boldsymbol{n}$ denotes the unit outward vector normal to $\partial\Omega$; A is piecewise constant in Ω and for almost all $x \in \Omega$, $A(x)$ is a symmetric positive definite matrix. If A is a scalar multiple of the identity matrix in a region, we would simply say A is a scalar in that region. Let $\bigcup_{i=1}^{L} \Omega_i$ be a disjoint partition of the domain Ω such that A is constant on each subdomain Ω_i.

The weak formulation for the problem in (4.1.1) is to find $u \in H^1_D(\Omega)$ such that

$$a(u,v) := \int_\Omega A\nabla u \cdot \nabla v dx = \int_\Omega fv dx - \int_{\Gamma_N} g_{_N} v ds, \quad \forall\, v \in H^1_D(\Omega).$$

The well-posedness of the weak formulation follows from the Riesz Representation Theorem.

Let the mesh information (elements,vertices,edges/faces,etc.) be given in Section 3.1. We assume that for each $K \in \mathcal{T}$, $A|_K$ is a symmetric, positive definite constant matrix that is nearly scalar, i.e., there exists a moderate constant $\kappa > 0$ such that

$$\rho_{\max}(A|_K) \le \kappa \rho_{\min}(A|_K), \quad \forall\, K \in \mathcal{T}, \tag{4.1.2}$$

where $\rho_{\max}$ and $\rho_{\min}$ denote maximal and minimal eigenvalues of a matrix, respectively. Obviously, if A is piecewise scalar with respect to $\mathcal{T}$, then (4.1.2) holds with $\kappa \equiv 1$. Let $(\cdot,\cdot)_S$ and $\|\cdot\|_S$ denote the L^2 inner product and norm on set S, respectively, and the subscript S is omitted when $S = \Omega$. Set

$$\alpha_{\max} := \max_{K\in\mathcal{T}} \rho_{\max}(A|_K), \quad \alpha_{\min} := \min_{K\in\mathcal{T}} \rho_{\min}(A|_K),$$

and define the piecewise constant function $\alpha \in L^2(\Omega)$ such that

$$\alpha = \alpha_{_K} := \rho_{\max}(A|_K) \quad \text{in } K \in \mathcal{T}. \tag{4.1.3}$$

Furthermore, as in [8], for each $e \in \mathcal{E}$, define

$$\alpha_e := \max_{K \subseteq \omega_e} \alpha_{_K},$$

where ω_e denotes the union of elements adjacent to e.

The conforming finite element space of order k is defined by

$$V_{\mathcal{T}} = \{v \in H^1_D(\Omega) : v|_{_K} \in P_k(K), \quad \forall\, K \in \mathcal{T}\}.$$

The finite element solution $u_{\mathcal{T}} \in V_{\mathcal{T}}$ satisfies

$$a(u_{\mathcal{T}}, v) = \int_{\Omega} f v dx - \int_{\Gamma_N} g_N v ds, \quad \forall\, v \in V_{\mathcal{T}}.$$

The well-posedness follows from the Riesz Representation Theorem.

Let $\bar{f}$ be the L^2 projection of f onto the piecewise polynomial space of degree $k-1$ with respect to $\mathcal{T}$ and denote by $\bar{g}_N$ the L^2 projection of g_N onto the piecewise polynomial space of degree $k-1$ with respect to $\mathcal{E}_N$. That is,

$$\bar{f}|_K := \Pi_K^{k-1} f, \quad \forall\, K \in \mathcal{T} \quad \text{and} \quad \bar{g}_N|_e := \Pi_e^{k-1} g_N, \quad \forall\, e \in \mathcal{E}_N,$$

where Π_K^{k-1} and Π_e^{k-1} denote the L^2 projection from $L^2(K)$ to $P_{k-1}(K)$ and from $L^2(e)$ to $P_{k-1}(e)$, respectively. The so-called *data oscillations* in f and g_N are defined by

$$\operatorname{osc}(f, K) := \frac{h_K}{\alpha_K^{1/2}} \|f - \bar{f}\|_K, \quad \operatorname{osc}(f, \mathcal{T}) := \left(\sum_{K \in \mathcal{T}} \operatorname{osc}(f, K)^2 \right)^{1/2},$$

and

$$\operatorname{osc}(g_N, e) := \frac{h_e^{1/2}}{\alpha_e^{1/2}} \|g_N - \bar{g}_N\|_e, \quad \operatorname{osc}(g_N, \mathcal{E}_N) := \left(\sum_{e \in \mathcal{E}_N} \operatorname{osc}(g_N, e)^2 \right)^{1/2},$$

respectively.

4.2 Hybrid estimator

Denote the true and the numerical fluxes by

$$\boldsymbol{\sigma} = -A \nabla u \ \text{ and } \ \boldsymbol{\sigma}_{\mathcal{T}} = -A \nabla u_{\mathcal{T}},$$

respectively. Since $\boldsymbol{\sigma} \in H(\operatorname{div}; \Omega)$ and $\boldsymbol{\sigma}_{\mathcal{T}} \notin H(\operatorname{div}; \Omega)$ in general, it is reasonable to find a flux $\hat{\boldsymbol{\sigma}} \in H(\operatorname{div}; \Omega)$ as a better approximation to $\boldsymbol{\sigma}$ than $\boldsymbol{\sigma}_{\mathcal{T}}$. In [43], $\hat{\boldsymbol{\sigma}}$ is chosen as the projection of the numerical flux $\boldsymbol{\sigma}_{\mathcal{T}}$ into some $H(\operatorname{div}; \Omega)$-conforming space with respect to the weighted L^2 inner product $(A^{-1}\cdot, \cdot)$. Precisely, with

$$H_{N,g}(\operatorname{div}; \Omega) := \{ \boldsymbol{\tau} \in H(\operatorname{div}; \Omega) : \boldsymbol{\tau} \cdot \boldsymbol{n} = g_N \text{ on } \Gamma_N \},$$

the projection $\hat{\boldsymbol{\sigma}} \in H_{N,g}(\mathrm{div}\,;\Omega)$ satisfies

$$\|A^{-1/2}(\hat{\boldsymbol{\sigma}} - \boldsymbol{\sigma}_{\mathcal{T}})\| = \min_{\boldsymbol{\tau} \in H_{N,g}(\mathrm{div}\,;\Omega)} \|A^{-1/2}(\boldsymbol{\tau} - \boldsymbol{\sigma}_{\mathcal{T}})\|. \tag{4.2.1}$$

Since $H_{N,g}(\mathrm{div}\,;\Omega)$ is a nonempty closed convex subset of the Hilbert space $H(\mathrm{div};\Omega)$, there exists a unique solution $\hat{\boldsymbol{\sigma}} \in H_{N,g}(\mathrm{div}\,;\Omega)$ for (4.2.1).

In the discrete setting, corresponding to the finite element space $V_{\mathcal{T}}$ of order $k\,(k = 1, 2, \dots)$, let $\mathcal{W}(K) := \mathrm{RT}_{k-1}(K)$ or $\mathrm{BDM}_k(K)$ and define

$$\mathcal{W}_g := \{\boldsymbol{\tau} \in H(\mathrm{div};\Omega) : \boldsymbol{\tau}|_K \in \mathcal{W}(K),\, \forall\, K \in \mathcal{T} \text{ and } \boldsymbol{\tau} \cdot \boldsymbol{n}|_e = \bar{g}_N,\, \forall\, e \in \mathcal{E}_N\}.$$

The finite element approximation of problem (4.2.1) is to find $\hat{\boldsymbol{\sigma}} \in \mathcal{W}_g$ such that

$$\|A^{-1/2}(\hat{\boldsymbol{\sigma}} - \boldsymbol{\sigma}_{\mathcal{T}})\| = \min_{\boldsymbol{\tau} \in \mathcal{W}_g} \|A^{-1/2}(\boldsymbol{\tau} - \boldsymbol{\sigma}_{\mathcal{T}})\|. \tag{4.2.2}$$

(4.2.2) was employed in [11] to perform flux recovery for higher order finite element approximations. However, a global minimization is computationally expensive in engineering practice, and it is preferable to find a local procedure, especially an element-based one.

We present a local approach by looking for a flux $\hat{\boldsymbol{\sigma}} \in \mathcal{W}_g$ that satisfies in each element $K \in \mathcal{T}$:

$$\begin{cases} \mathrm{div}\ \hat{\boldsymbol{\sigma}} = \hat{f}_K & \text{in } K, \\ \hat{\boldsymbol{\sigma}} \cdot \boldsymbol{n}_e = \hat{g}_e & \text{on } e \in \mathcal{E}_K, \end{cases} \tag{4.2.3}$$

where $\hat{g}_e \in P_{k-1}(e)$ is a weighted average of the normal components of the numerical fluxes

$$\hat{g}_e := \begin{cases} (1-\lambda_e)\boldsymbol{\sigma}_{\mathcal{T}}|_{K_e^+} \cdot \boldsymbol{n}_e + \lambda_e \boldsymbol{\sigma}_{\mathcal{T}}|_{K_e^-} \cdot \boldsymbol{n}_e, & \text{if } e \in \mathcal{E}_I, \\ \bar{g}_N, & \text{if } e \in \mathcal{E}_N, \\ \boldsymbol{\sigma}_{\mathcal{T}} \cdot \boldsymbol{n}_e, & \text{if } e \in \mathcal{E}_D, \end{cases} \tag{4.2.4}$$

with weight

$$\lambda_e := \frac{\alpha_{K_e^-}^{-1} h_{K_e^-}}{\alpha_{K_e^+}^{-1} h_{K_e^+} + \alpha_{K_e^-}^{-1} h_{K_e^-}} \tag{4.2.5}$$

and $\hat{f}_{K} \in P_{k-1}(K)$ is given by

$$\hat{f}_{K} := \bar{f} + J_K, \quad \text{with } J_K := |K|^{-1} \left(\sum_{e \in \mathcal{E}_K} \int_e s_{K} \hat{g}_e ds - \int_K \bar{f} dx \right). \tag{4.2.6}$$

Note that the choice of $\hat{f}_{K}$ guarantees that (4.2.3) is compatible. Since the solution to (4.2.3) is not unique in general, we have to choose a particular one. Two choices are presented below.

A unique $\hat{\boldsymbol{\sigma}}_{\mathcal{T}}$ can be chosen as the minimizer of the following problem in each element

$$\|A^{-1/2}(\hat{\boldsymbol{\sigma}}_{\mathcal{T}} - \boldsymbol{\sigma}_{\mathcal{T}})\|_K = \min_{\boldsymbol{\tau} \in \mathcal{S}(K)} \|A^{-1/2}(\boldsymbol{\tau} - \boldsymbol{\sigma}_{\mathcal{T}})\|_K, \quad \forall\, K \in \mathcal{T}, \tag{4.2.7}$$

where

$$\mathcal{S}(K) := \{\boldsymbol{\tau} \in \mathcal{W}(K) : \text{div } \boldsymbol{\tau} = \hat{f}_{K} \text{ in } K \text{ and } \boldsymbol{\tau} \cdot \boldsymbol{n}_e = \hat{g}_e \text{ on } e \in \mathcal{E}_K\}.$$

An alternative, which is explicit, is to define $\widetilde{\boldsymbol{\sigma}}_{\mathcal{T}} \in \mathcal{W}_g$ as follows:

$$\begin{cases} \widetilde{\boldsymbol{\sigma}}_{\mathcal{T}} \cdot \boldsymbol{n}_e = \hat{g}_e, \quad \forall\, e \in \mathcal{E}_K, \\ \displaystyle\int_K \widetilde{\boldsymbol{\sigma}}_{\mathcal{T}} \cdot \nabla p dx = \sum_{e \in \mathcal{E}_K} (s_{K} \hat{g}_e, p)_e - (\hat{f}_{K}, p)_K, \quad \forall\, p \in P_{k-1}(K), \\ \displaystyle\int_K \widetilde{\boldsymbol{\sigma}}_{\mathcal{T}} \cdot \boldsymbol{q} dx = \int_K \boldsymbol{\sigma}_{\mathcal{T}} \cdot \boldsymbol{q} dx, \quad \forall\, \boldsymbol{q} \in \mathcal{Q}(K), \end{cases} \tag{4.2.8}$$

where

$$\mathcal{Q}(K) = \begin{cases} Q_{k-2}(K), & \text{if } \mathcal{W}(K) = \mathrm{RT}_{k-1}(K),\ k \geq 2, \\ H_k(K), & \text{if } \mathcal{W}(K) = \mathrm{BDM}_k(K). \end{cases} \tag{4.2.9}$$

Here

$$H_k(K) := \{\boldsymbol{q} \in P_k(K)^d : \text{div } \boldsymbol{q} = 0 \text{ in } K \text{ and } \boldsymbol{q} \cdot \boldsymbol{n}|_{\partial K} = 0\} \quad (k \geq 1) \tag{4.2.10}$$

and

$$Q_{k-2}(K) := \{\boldsymbol{q} \in P_{k-2}(K)^d : (\boldsymbol{q}, \nabla p)_K = 0,\ \forall\, p \in P_{k-1}(K)\} \quad (k \geq 2). \tag{4.2.11}$$

The space $H_k(K)$ is defined in [28] to fix the degrees of freedom in $\mathrm{BDM}_k(K)$, and the space $Q_{k-2}(K)$ gives the following orthogonal decomposition of $P_{k-2}(K)^d$ with respect to the L^2 inner product:

$$P_{k-2}(K)^d = \{\nabla p : p \in P_{k-1}(K)\} \oplus Q_{k-2}(K). \tag{4.2.12}$$

Corresponding to two choices $\hat{\boldsymbol{\sigma}}_{\mathcal{T}}$ and $\widetilde{\boldsymbol{\sigma}}_{\mathcal{T}}$, we define two local indicators by

$$\xi_K := \left(h_K^2 \alpha_K^{-1} \|\bar{f} - \hat{f}_K\|_K^2 + \|A^{-1/2}(\hat{\boldsymbol{\sigma}}_{\mathcal{T}} - \boldsymbol{\sigma}_{\mathcal{T}})\|_K^2\right)^{1/2} \tag{4.2.13}$$

and

$$\tilde{\xi}_K := \left(h_K^2 \alpha_K^{-1} \|\bar{f} - \hat{f}_K\|_K^2 + \|A^{-1/2}(\widetilde{\boldsymbol{\sigma}}_{\mathcal{T}} - \boldsymbol{\sigma}_{\mathcal{T}})\|_K^2\right)^{1/2}, \tag{4.2.14}$$

respectively. The corresponding global estimators are

$$\xi := \left(\sum_{K\in\mathcal{T}} \xi_K^2\right)^{1/2} \quad \text{and} \quad \tilde{\xi} := \left(\sum_{K\in\mathcal{T}} \tilde{\xi}_K^2\right)^{1/2}. \tag{4.2.15}$$

Those estimators are called "hybrid" since they incorporate not only flux errors as in recovery-based estimators [3, 4, 6, 7, 43, 44] but also errors related to f as in residual-based estimators [8, 9, 38].

Due to the minimization property of $\hat{\boldsymbol{\sigma}}_{\mathcal{T}}$ in (4.2.7), it immediately follows that

$$\xi_K \le \tilde{\xi}_K \quad \text{and} \quad \xi \le \tilde{\xi}. \tag{4.2.16}$$

4.3 Equivalence to the robust residual estimator

Let η_K denote the standard residual-based error indicator in K (cf. [8, 9, 38]), i.e.,

$$\eta_K^2 := \frac{h_K^2}{\alpha_K}\|\bar{f} - \operatorname{div} \boldsymbol{\sigma}_{\mathcal{T}}\|_K^2 + \frac{1}{2}\sum_{e\in\mathcal{E}_K\cap\mathcal{E}_I} \frac{h_K}{\alpha_e}\|j_e\|_e^2 + \sum_{e\in\mathcal{E}_K\cap\mathcal{E}_N} \frac{h_K}{\alpha_e}\|j_e\|_e^2, \tag{4.3.1}$$

where the flux jump j_e is given by

$$j_e := \begin{cases} (\boldsymbol{\sigma}_{\mathcal{T}}|_{K_e^+} - \boldsymbol{\sigma}_{\mathcal{T}}|_{K_e^-}) \cdot \boldsymbol{n}_e, & \text{if } e \in \mathcal{E}_I, \\ \boldsymbol{\sigma}_{\mathcal{T}} \cdot \boldsymbol{n}_e - \bar{g}_N, & \text{if } e \in \mathcal{E}_N, \\ 0, & \text{if } e \in \mathcal{E}_D. \end{cases} \tag{4.3.2}$$

The corresponding global error estimator is

$$\eta_R := \left(\sum_{K \in \mathcal{T}} \eta_K^2 \right)^{1/2}.$$

We prove in this section that the two hybrid estimators are equivalent to the residual estimator with constants independent of α.

Theorem 4.3.1. *Let η_K be the residual-based error indicator in* (4.3.1) *and ξ_K and $\tilde{\xi}_K$ be the hybrid error indicators in* (4.2.13) *and* (4.2.14)*, respectively. Then there exist positive constants C_1 and C_2 independent of α such that*

$$C_1 \xi_K \le \eta_K \le C_2 \sum_{K' \subset \omega_K} \xi_{K'} \quad \text{and} \quad C_1 \tilde{\xi}_K \le \eta_K \le C_2 \sum_{K' \subset \omega_K} \tilde{\xi}_{K'}. \tag{4.3.3}$$

In order to bound $\tilde{\xi}_K$ from above by η_K, the following lemma (cf. [12]) is needed.

Lemma 4.3.1. *For $k \ge 1$, let $f \in P_{k-1}(K)$ and $g|_e \in P_{k-1}(e)$ for all $e \in \mathcal{E}_K$ satisfy the following compatibility condition*

$$\int_K f dx = \int_{\partial K} g ds.$$

With $\mathcal{W}(K) = \mathrm{RT}_{k-1}(K)$ or $\mathrm{BDM}_k(K)$, if the vector field $\boldsymbol{\tau} \in \mathcal{W}(K)$ is defined by

$$\begin{cases} \boldsymbol{\tau} \cdot \boldsymbol{n} = g, \quad \text{on } \partial K, \\ \displaystyle\int_K \boldsymbol{\tau} \cdot \nabla p \, dx = (g, p)_{\partial K} - (f, p)_K, \quad \forall\, p \in P_{k-1}(K), \\ \displaystyle\int_K \boldsymbol{\tau} \cdot \boldsymbol{q} \, dx = 0, \quad \forall\, \boldsymbol{q} \in \mathcal{Q}(K), \end{cases} \tag{4.3.4}$$

where $\mathcal{Q}(K)$ is defined in (4.2.9)*, then $\boldsymbol{\tau}$ satisfies the following divergence equation*

$$\begin{cases} \mathrm{div}\ \boldsymbol{\tau} = f \quad \text{in } K, \\ \boldsymbol{\tau} \cdot \boldsymbol{n} = g \quad \text{on } \partial K \end{cases} \tag{4.3.5}$$

and the following stability estimate

$$\|\boldsymbol{\tau}\|_K \le c \left(h_K \|f\|_K + h_K^{1/2} \|g\|_{\partial K} \right).$$

In the proof below, we will use c (or C) to denote a generic positive constant, possibly different at different occurrences, that is independent of α, but may depend on the shape parameter of mesh $\mathcal{T}$, the polynomial degree k, and κ in (4.1.2).

Proof of Theorem 4.3.1. We first estimate J_K in terms of flux jump and element residual. By the divergence theorem, J_K in (4.2.6) can be expressed as

$$J_K = |K|^{-1} \left(\int_K (\operatorname{div} \boldsymbol{\sigma}_{\mathcal{T}} - \bar{f}) dx + \sum_{e \in \mathcal{E}_K} \int_e s_{K}(\hat{g}_e - \boldsymbol{\sigma}_{\mathcal{T}}|_K \cdot \boldsymbol{n}_e) ds \right). \tag{4.3.6}$$

It can be computed that (by checking the two cases $K = K_e^+$ and $K = K_e^-$ when $e \in \mathcal{E}_I$)

$$\|\hat{g}_e - \boldsymbol{\sigma}_{\mathcal{T}}|_K \cdot \boldsymbol{n}_e\|_e \leq c \left(\frac{\alpha_{K}}{\alpha_e} \right)^{1/2} \|j_e\|_e, \quad \forall\, e \in \mathcal{E}_K, \tag{4.3.7}$$

where c may depend on κ, but is independent of α. Then it follows from (4.3.6), the triangle and the Cauchy-Schwarz inequalities that

$$\|J_K\|_K \leq \|\operatorname{div} \boldsymbol{\sigma}_{\mathcal{T}} - \bar{f}\|_K + c h_K^{-1/2} \sum_{e \in \mathcal{E}_K} \left(\frac{\alpha_{K}}{\alpha_e} \right)^{1/2} \|j_e\|_e. \tag{4.3.8}$$

Applying Lemma 4.3.1 to $\widetilde{\boldsymbol{\sigma}}_{\mathcal{T}} - \boldsymbol{\sigma}_{\mathcal{T}}$, together with the triangle inequality, (4.3.8) and (4.3.7), yields that

$$\|A^{-1/2}(\widetilde{\boldsymbol{\sigma}}_{\mathcal{T}} - \boldsymbol{\sigma}_{\mathcal{T}})\|_K \leq c \left(h_K \alpha_K^{-1/2} \|\operatorname{div} \boldsymbol{\sigma}_{\mathcal{T}} - \bar{f}\|_K + \sum_{e \in \mathcal{E}_K} \frac{h_K^{1/2}}{\sqrt{\alpha_e}} \|j_e\|_e \right), \tag{4.3.9}$$

where c may depend on κ, but is independent of α.

Analogously, it follows from (4.2.6) and (4.3.8) that

$$h_K \alpha_K^{-1/2} \|\bar{f} - \hat{f}_{K}\|_K \leq c \left(h_K \alpha_K^{-1/2} \|\operatorname{div} \boldsymbol{\sigma}_{\mathcal{T}} - \bar{f}\|_K + \sum_{e \in \mathcal{E}_K} \frac{h_K^{1/2}}{\sqrt{\alpha_e}} \|j_e\|_e \right). \tag{4.3.10}$$

It is easily seen from the definition of $\tilde{\xi}_K$ in (4.2.14), the property in (4.2.16) and the estimates in (4.3.9) and (4.3.10) that

$$\xi_K^2 \leq \tilde{\xi}_K^2 \leq c\, \eta_K^2, \tag{4.3.11}$$

with c independent of α.

To prove the upper bounds in (4.3.3), we adopt the idea on the proof of local efficiency of η_K in [8]. To this end, let ϕ_e denote the edge/face bubble function associated with edge ($d = 2$)/face ($d = 3$) e and ϕ_K the element bubble function associated with element K. For each element K contained in ω_e, by extending j_e to be a polynomial in K with stability estimate (cf. [38, 45])

$$\|j_e\|_K \leq Ch_e^{1/2}\|j_e\|_e, \quad K \subseteq \omega_e,$$

we deduce from integration by parts, the Cauchy-Schwarz and the inverse inequalities that

$$\begin{aligned}\|j_e\|_e^2 &\leq C\int_e \phi_e j_e^2 ds = C\sum_{K\subseteq\omega_e}\left(\int_K (\boldsymbol{\sigma}_\tau - \hat{\boldsymbol{\sigma}}_\tau)\cdot\nabla(\phi_e j_e)dx + \int_K \phi_e j_e \mathrm{div}\,(\boldsymbol{\sigma}_\tau - \hat{\boldsymbol{\sigma}}_\tau)dx\right)\\ &\leq C\sum_{K\subseteq\omega_e}\left(\alpha_K^{1/2}h_e^{-1/2}\|A^{-1/2}(\boldsymbol{\sigma}_\tau - \hat{\boldsymbol{\sigma}}_\tau)\|_K + h_e^{1/2}\|\mathrm{div}\ \boldsymbol{\sigma}_\tau - \hat{f}_K\|_K\right)\|j_e\|_e\end{aligned}$$

and that

$$\begin{aligned}\|\hat{f}_K - \mathrm{div}\ \boldsymbol{\sigma}_\tau\|_K^2 &\leq C\int_K \phi_K(\mathrm{div}\,\hat{\boldsymbol{\sigma}}_\tau - \mathrm{div}\,\boldsymbol{\sigma}_\tau)^2 dx\\ &= -C\int_K \nabla(\phi_K(\mathrm{div}\,\hat{\boldsymbol{\sigma}}_\tau - \mathrm{div}\,\boldsymbol{\sigma}_\tau))\cdot(\hat{\boldsymbol{\sigma}}_\tau - \boldsymbol{\sigma}_\tau)dx\\ &\leq C\alpha_K^{1/2}h_K^{-1}\|\hat{f}_K - \mathrm{div}\ \boldsymbol{\sigma}_\tau\|_K\|A^{-1/2}(\hat{\boldsymbol{\sigma}}_\tau - \boldsymbol{\sigma}_\tau)\|_K,\end{aligned}$$

where C may depend on κ, but is independent of α. Hence we conclude that

$$h_K\alpha_K^{-1/2}\|\hat{f}_K - \mathrm{div}\ \boldsymbol{\sigma}_\tau\|_K \leq C\|A^{-1/2}(\hat{\boldsymbol{\sigma}}_\tau - \boldsymbol{\sigma}_\tau)\|_K,$$

and

$$h_e^{1/2}\alpha_e^{-1/2}\|j_e\|_e \leq C\|A^{-1/2}(\hat{\boldsymbol{\sigma}}_\tau - \boldsymbol{\sigma}_\tau)\|_{\omega_e}$$

with C independent α. Now the upper bounds in (4.3.3) are a direct consequence of the definition of η_K and ξ_K in the respective (4.3.1) and (4.2.13), the triangle inequality, and (4.2.16). This completes the proof of the theorem. □

4.4 A Posteriori Error Estimate

As a consequence of the equivalence in Theorem 4.3.1, the a posteriori error estimate of hybrid estimators follows from that of the residual estimator, where the

robustness of the upper bound estimate (i.e., reliability) is proved under an assumption on the distribution of α. Such an assumption can be the monotonicity assumption in [8, Hypothesis 2.7] or the quasi-monotonicity condition in [9], which is weaker than [8, Hypothesis 2.7]. The assumption in [8, Hypothesis 2.7] reads:

> for any two different subdomains $\overline{\Omega}_i$ and $\overline{\Omega}_j$ that share at least one point, there exists a connected path passing from $\overline{\Omega}_i$ to $\overline{\Omega}_j$ through adjacent subdomains such that α is monotone along this path.

The quasi-monotonicity condition in [9] is cited as Definition 7.4.1.

Theorem 4.4.1. *For the local indicators ξ_K and $\tilde{\xi}_K$ defined in* (4.2.13) *and* (4.2.14)*, respectively, there exists a constant c independent of $\alpha_{\max}/\alpha_{\min}$ such that*

$$\xi_K^2 \leq \tilde{\xi}_K^2 \leq c \left(\|A^{1/2}\nabla(u - u_{\mathcal{T}})\|_{\omega_K}^2 + \sum_{K' \subset \omega_K} \operatorname{osc}(f, K')^2 + \sum_{e \in \mathcal{E}_K \cap \mathcal{E}_N} \operatorname{osc}(g_N, e)^2 \right).$$

If α defined in (4.1.3) *is quasi-monotone, then the estimators ξ and $\tilde{\xi}$ defined in* (4.2.15) *satisfy the following global reliability estimate:*

$$\begin{aligned} \|A^{1/2}\nabla(u - u_{\mathcal{T}})\| &\leq C\left(\xi + \operatorname{osc}(f, \mathcal{T}) + \operatorname{osc}(g_N, \mathcal{E}_N)\right) \\ &\leq C\left(\tilde{\xi} + \operatorname{osc}(f, \mathcal{T}) + \operatorname{osc}(g_N, \mathcal{E}_N)\right) \end{aligned}$$

with C independent of $\alpha_{\max}/\alpha_{\min}$.

Theorem 4.4.1 is an immediate result of Theorem 4.3.1 and the robust a posteriori error estimates of the residual estimator (cf. [8, 9, 38]).

4.5 Numerical Experiments

We consider adaptive finite element for solving three examples. For all examples, $\Omega = (-1, 1)^2$ and the initial mesh consists of 4×4 congruent squares, each of which is partitioned into two triangles connecting bottom-left and top-right corners. Dörfler's

marking strategy [46] is used with $\theta_D = 0.5$ as in [46, 47]. Namely, in the refinement of $\mathcal{T}$, a minimal subset $\hat{\mathcal{T}}$ of $\mathcal{T}$ is constructed such that

$$\left(\sum_{K\in\hat{\mathcal{T}}} \zeta_K^2\right)^{1/2} \geq \theta_D \left(\sum_{K\in\mathcal{T}} \zeta_K^2\right)^{1/2}, \tag{4.5.1}$$

where ζ_K denotes an error indicator on element K. The newest-vertex bisection [44] is used in the refinement. The following notation will be used:

- exact error $e := u - u_{\mathcal{T}}$;
- effectivity index is denoted by: eff-ind;
- degrees of freedom: DOFs;
- stopping criterion: $\|A^{1/2}\nabla e\| \leq \epsilon_{\text{rel}}\|A^{1/2}\nabla u\|$ where ϵ_{rel} is a prescribed tolerance.

Examples in Section 4.5.1 and Section 4.5.2 are designed to illustrate the well-known fact that residual estimator usually overestimates the true error by a large margin [39] and also to show that the hybrid estimator is more accurate. In Section 4.5.3 we solve the numerical benchmark known as Kellogg's example [47, 48] to justify the robustness and generality of the proposed estimator.

4.5.1 Example 1

To illustrate the effectivity of residual estimator, we first consider the Poisson equation with Dirichlet boundary condition and the data is chosen such that the exact solution is a quadratic polynomial

$$u(x, y) = -x^2 - y^2.$$

The tolerance in the stopping criterion is chosen as $\epsilon_{\text{rel}} = 0.01$. With P_1 finite element approximation, the numerical results are shown in Table 4.1 and Figure 4.1 - 4.2. It can be seen from Table 4.1 that, for such a smooth problem, the residual estimator η is much less accurate compared to the proposed hybrid estimator ξ.

Table 4.1.
Example 1 - quadratic solution and P_1 discretization

estimator	DOFs	$\|A^{1/2}\nabla e\|/\|A^{1/2}\nabla u\|$	eff-ind
η	7926	9.6E-3	5.35
ξ	8520	9.4E-3	1.11

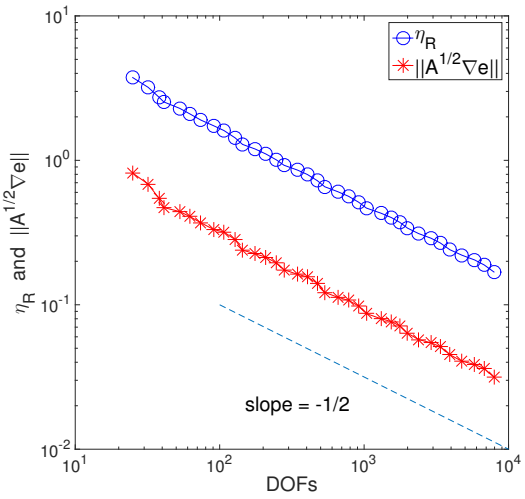

Figure 4.1. Example 1: P_1 - η

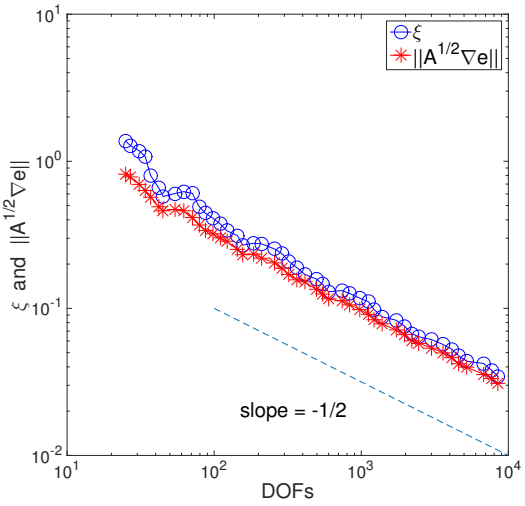

Figure 4.2. Example 1: P_1 - ξ

Table 4.2.
Example 2 - quartic solution and P_2 discretization

estimator	DOFs	$\|A^{1/2}\nabla e\|/\|A^{1/2}\nabla u\|$	eff-ind
η	3037	9.9E-4	9.87
ξ	3193	9.5E-4	2.36

4.5.2 Example 2

We consider again the Poisson equation as in Section 4.5.1, where the exact solution is now a quartic polynomial

$$u(x, y) = -(x^2 - 1)(y^2 - 1).$$

The tolerance is chosen as $\epsilon_{\text{rel}} = 0.001$ and P_2 discretization is used. The numerical results are presented in Table 4.2 and Figure 4.3 - 4.4, from which we see that the residual estimator η overestimates the true error by a factor of 9.87, while the proposed estimator ξ remains accurate.

4.5.3 Example 3 - Kellogg's example

We consider solving the Kellogg's example [48]. The parameters are same as in [47] and are collected here for completeness. The domain is $\Omega = (-1, 1)^2$ and the diffusion coefficient is chosen as $A = \alpha_1 I$ in the first and third quadrants, and $A = \alpha_2 I$ in the second and fourth quadrants, where

$$\alpha_1 \approx 161.4476387975881 \quad \text{and} \quad \alpha_2 = 1.$$

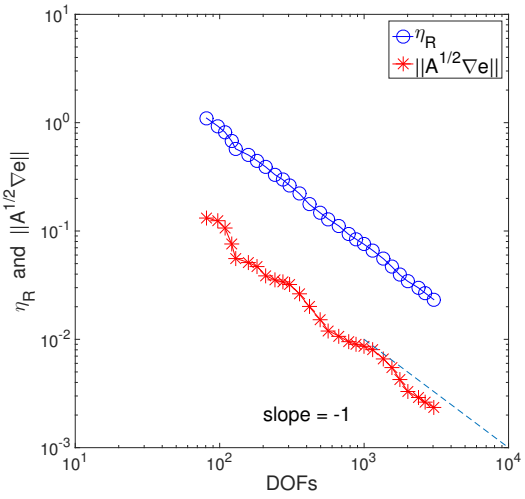

Figure 4.3. Example 2: P_2 - η

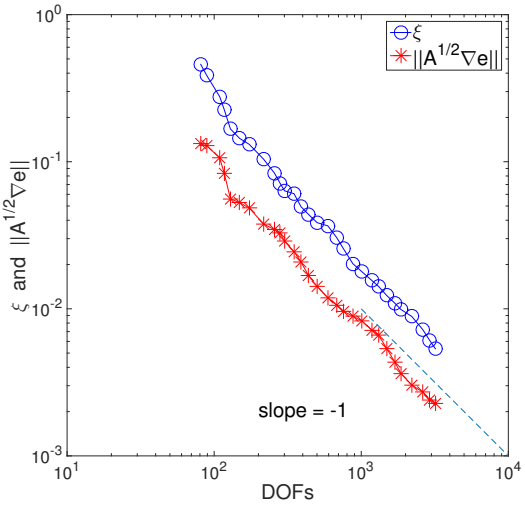

Figure 4.4. Example 2: P_2 - ξ

For $f = 0$, an exact solution in polar coordinate is given by $u(r, \theta) = r^{\beta}\psi(\theta)$ with

$$\psi(\theta) := \begin{cases} \cos((\pi/2 - \tau)\beta)\cos((\theta - \pi/4)\beta), & \text{if } 0 \le \theta \le \pi/2, \\ \cos(\pi\beta/4)\cos((\theta - \pi + \tau)\beta), & \text{if } \pi/2 \le \theta \le \pi, \\ \cos(\tau\beta)\cos((\theta - 5\pi/4)\beta), & \text{if } \pi \le \theta \le 3\pi/2, \\ \cos((\pi/2 - \tau)\beta)\cos((\theta - 3\pi/2 - \tau)\beta), & \text{if } 0 \le \theta \le \pi/2, \end{cases}$$

$$\beta = 0.1 \quad \text{and} \quad \tau \approx 14.92256510455152.$$

The regularity of u is quite low as $u \notin H^{1.1}(\Omega)$.

We perform numerical tests by using both P_1 and P_2 conforming elements. For $P_k (k = 1, 2)$ element and RT_{k-1} flux recovery, the solution to (4.2.3) is unique, so, $\hat{\boldsymbol{\sigma}}_{\mathcal{T}} = \widetilde{\boldsymbol{\sigma}}_{\mathcal{T}}$. Consequently, $\xi_K = \tilde{\xi}_K$. The relative error tolerance is chosen as $\epsilon_{\mathrm{rel}} = 0.05$.

The numerical results are collected in Table 4.3 and Fig. 4.5 – 4.12.

From the meshes in Fig. 4.5 – 4.6 and Fig. 4.9 – 4.10, we see that both estimators are robust with respect to the large jump of the diffusion coefficient. Moreover, optimal convergence rates are observed from Fig. 4.7 – 4.8 for P_1 element and from Fig. 4.11 – 4.12 for P_2 element. From Table 4.3, Fig. 4.7 – 4.8 and Fig. 4.11 – 4.12, we observe that, although the two estimators are proved to be equivalent in Theorem 4.3.1 ξ (or $\tilde{\xi}$) is more accurate than η.

Next we present numerical results to demonstrate that, for the estimator ξ(as well as $\tilde{\xi}$), the term ξ_f defined below, which measures the divergence error between recovered flux and exact flux (actually J_K in (4.2.6)), is not a higher order term.

$$\begin{aligned} \xi_f :&= \left(\sum_{K \in \mathcal{T}} h_K^2 \alpha_K^{-1} \| \bar{f} - \hat{f}_K \|_K^2 \right)^{1/2} \\ &= \left(\sum_{K \in \mathcal{T}} h_K^2 \alpha_K^{-1} \| J_K \|_K^2 \right)^{1/2} = \left(\sum_{K \in \mathcal{T}} h_K^2 \alpha_K^{-1} |K| |J_K|^2 \right)^{1/2}. \end{aligned}$$

It can be seen from Fig. 4.13 and Fig. 4.14 that ξ_f is of the same order as the true error.

Table 4.3.
Example 3 - Kellogg's Example (P_1, P_2 elements)

u_τ	estimator	DOFs	$\|A^{1/2}\nabla e\|/\|A^{1/2}\nabla u\|$	eff-ind
P_1	η	35707	4.9E-2	1.96
	ξ	29072	4.8E-2	1.35
P_2	η	5133	4.9E-2	2.48
	ξ	4429	4.9E-2	1.50

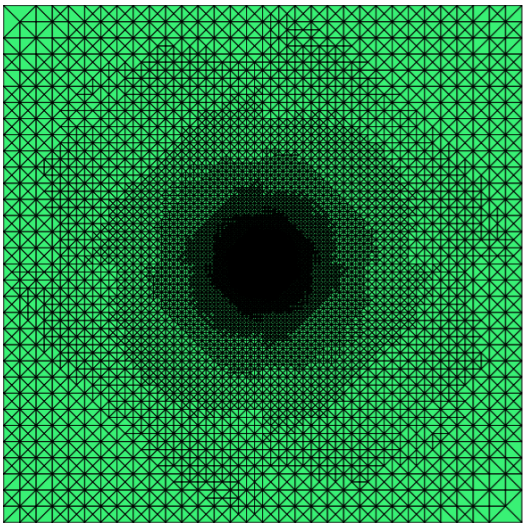

Figure 4.5. Example 3: P_1 - mesh - η

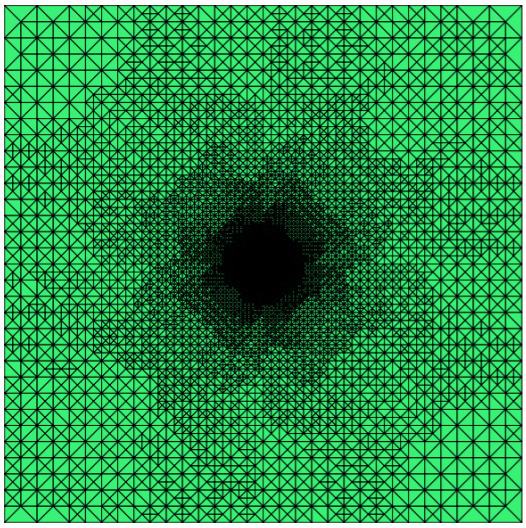

Figure 4.6. Example 3: P_1 - mesh - ξ

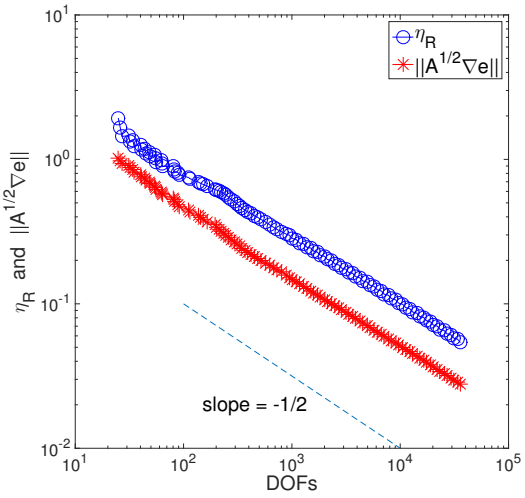

Figure 4.7. Example 3: P_1 - η

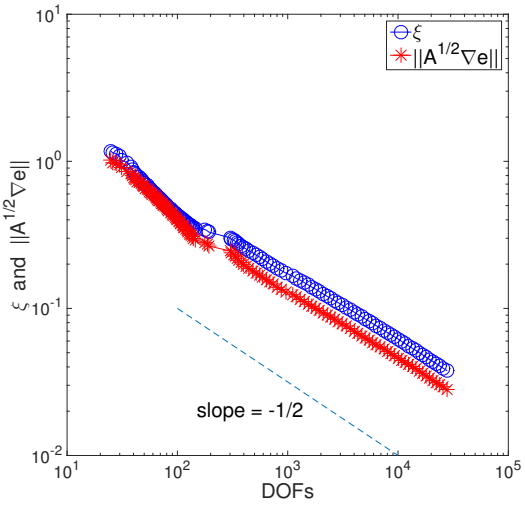

Figure 4.8. Example 3: P_1 - ξ

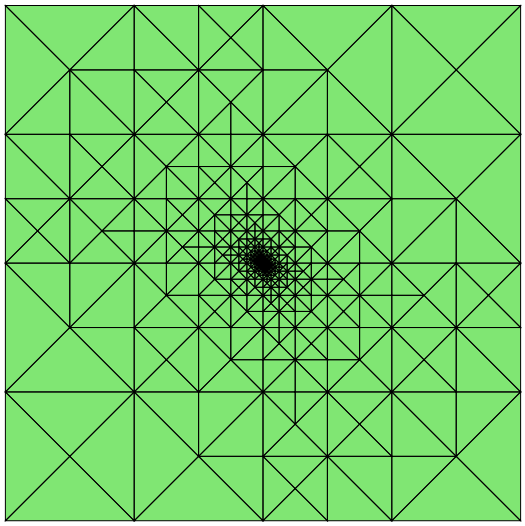

Figure 4.9. Example 3: P_2 - mesh - η

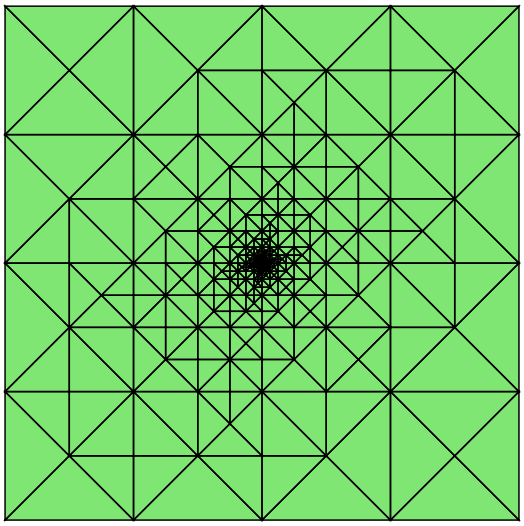

Figure 4.10. Example 3: P_2 - mesh - ξ

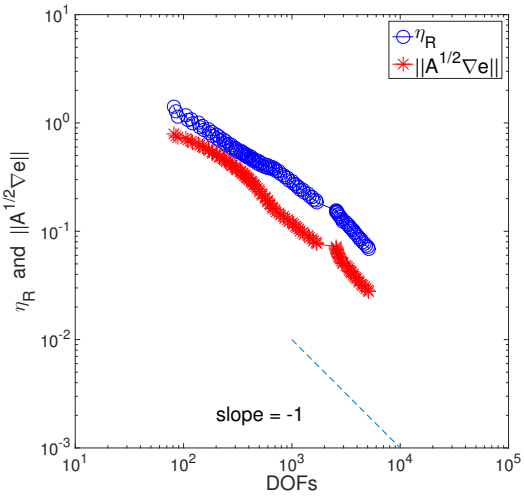

Figure 4.11. Example 3: P_2 - η

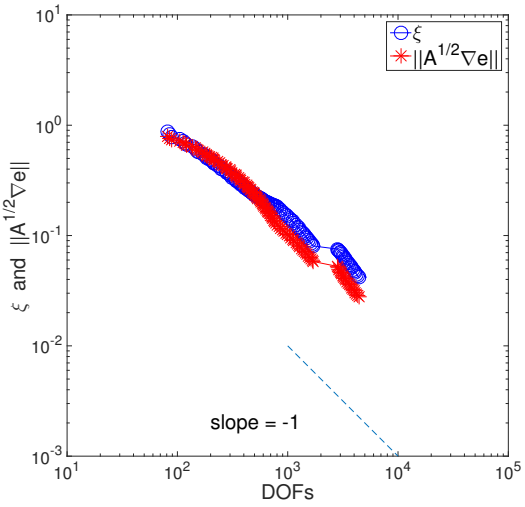

Figure 4.12. Example 3: P_2 - ξ

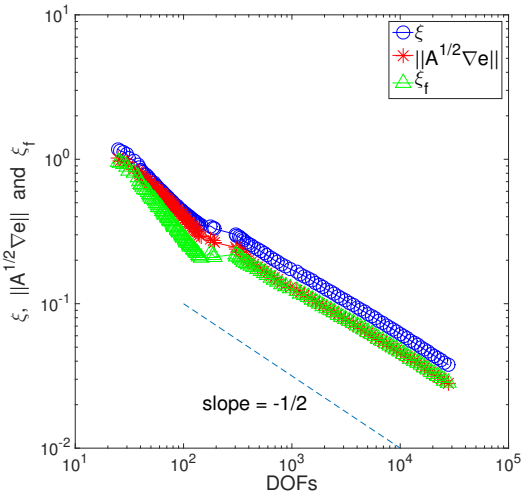

Figure 4.13. Example 3: P_1 - error, ξ and ξ_f

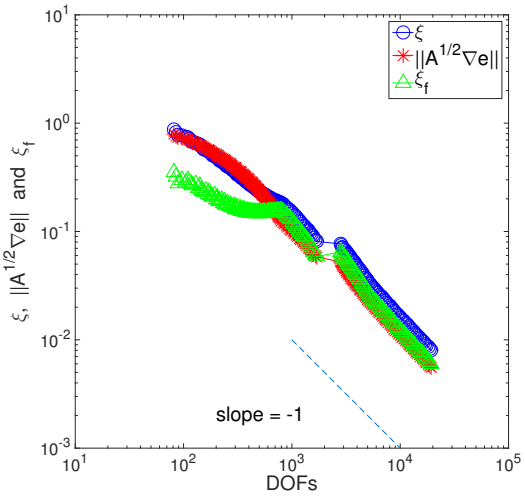

Figure 4.14. Example 3: P_2 - error, ξ and ξ_f

5. HYBRID ESTIMATOR FOR CONVECTION-DIFFUSION-REACTION EQUATIONS

A version of this chapter has been submitted for publication [14].

For elliptic partial differential equations consisting of terms characterizing diffusion, reaction and convection, the solution may display a strong interface singularity, interior or boundary layers, etc., due to discontinuous coefficients or terms in significantly different scales, etc. Consequently, it is a challenging task to design a *general* a posteriori error estimator that is accurate and robust enough to resolve local behaviors of the exact solution without spending much computational resource.

Residual estimators are directly related to the error and have been well-studied since 1970s (cf. [2,8,9,36,37,45,49,50]). They enjoy many appealing features including explicit formula, validity to a large class of problems, natural extension to higher order elements, etc. More importantly, *robust* residual estimators have been proposed for various problems. For diffusion problems with discontinuous coefficient, [8, 9] derived robustness with respect to coefficient jump and concepts of monotonicity and quasi-monotonicity of the diffusion coefficient were introduced. For singularly perturbed reaction-diffusion problems, Verfürth [49] pioneered a residual estimator that is robust with respect to the size of reaction. For convection/reaction-dominated convection-reaction-diffusion problems, it is still under debate on how to choose a suitable norm to measure the approximation error (cf. [2, 50–53]). Sangalli [52, 53] proposed a norm incorporating the standard energy norm and a seminorm of order 1/2 and developed the a posteriori error analysis for the one-dimensional setting. Verfürth [2] introduced a norm incorporating the standard energy norm and a dual norm of the convective derivative. With respect to this norm, the general residual estimator in [2] was proved to be robust. Those developments make residual estimators extremely

competitive when dealing with challenging problems. One drawback, however, is that residual estimators tend to overestimate the true error by a large margin (cf. [12,39]). This calls for the need of an estimator as general as the residual estimator but with improved accuracy.

Recovery-based estimators via the averaging techniques, e.g., the ZienkiewiczZhu (ZZ) estimator and its variations (cf. [3, 4, 6, 7, 44, 54], etc.), are quite popular in the engineering community. However, unlike residual estimators, the robustness of those estimators with respect to issues like coefficient jump, dominated convection or reaction, etc., has not been emphasized or studied in detail yet (cf. [42]). On coarse meshes, it is known that ZZ-type estimators are in general unreliable, and counterexamples can be easily constructed where the estimator vanishes but the true error is large (cf. Section 11.3). Moreover, the straightforward extension of the original ZZ estimator [3, 4] from lowest order elements to higher order elements is nontrivial.

The *hybrid estimator* introduced in [12] (see also Chapter 4) for diffusion problems incorporates the advantages of both residual and ZZ-type estimators. Numerical results in Section 4.5 indicate that the hybrid estimator is more accurate than the residual estimator. This opens a door of finding a substitute of the residual estimator with improved accuracy. Thus one may ask if it is possible to construct hybrid estimators for more general problems and if the hybrid estimator is still more accurate than the residual estimator.

In the following, we introduce the hybrid estimator as well as flux recoveries for convection-reaction-diffusion equations. In diffusion-dominated regime, the flux recovery is a natural extension of the one in (4.2.8). In convection/reaction-dominated regime, the flux recovery in each element depends on the size of diffusion. Roughly speaking, in elements with resolved diffusion, the recovered flux is same to the diffusion-dominated case; in elements where diffusion is not resolved, inspired in part by [55, Section 3.4], the recovered flux is defined piecewisely in each element (see Section 5.3.2). Once the recovered flux is obtained, the hybrid estimator is analogously

defined as in Section 4.2. We prove that the hybrid estimator is equivalent to the robust residual estimator in [2] and then the robustness follows immediately from that of the residual estimator. Similar to the residual estimator, the hybrid estimator is explicit and valid for higher order elements. Various numerical results show that, compared to the residual-based estimator, the hybrid estimator is more accurate and the corresponding effectivity index is significantly less sensitive to the magnitude of ϵ.

The chapter is organized as follows. The model problems and finite element discretizations are introduced in Section 5.1. Section 5.2 collects results on robust residual estimators. After the flux recovery presented in Section 5.3, the hybrid estimator is defined in Section 5.4. Section 5.5 gives the proof of the local equivalence between the residual estimator and the hybrid estimator. Numerical results are shown in Section 5.6.

5.1 Problem And Discretization

Let Ω be a polygonal domain in $\mathbb{R}^d$ $(d = 2, 3)$ with Lipschitz boundary $\partial\Omega$ consisting of two disjoint components Γ_D and Γ_N. By convention, assume that $\mathrm{diam}(\Omega) = O(1)$. Consider the stationary convection-reaction-diffusion equation:

$$\begin{cases} -\mathrm{div}(\alpha\nabla u) + \boldsymbol{a}\cdot\nabla u + bu = f, & \text{in } \Omega, \\ u = 0, & \text{on } \Gamma_D, \\ -\alpha\nabla u\cdot\boldsymbol{n} = g_N, & \text{on } \Gamma_N, \end{cases} \tag{5.1.1}$$

with $\alpha(x) \geq \delta$, for almost all $x \in \Omega$ and for some constant $\delta > 0$. Assume that:

(A1) $\boldsymbol{a} \in W^{1,\infty}(\Omega)^d$ and $b \in L^\infty(\Omega)$;

(A2) there are two constants $\beta \geq 0$ and $c_b \geq 0$, independent of α, such that

$$b - \frac{1}{2}\mathrm{div}\ \boldsymbol{a} \geq \beta \text{ in } \Omega \quad \text{and} \quad \|b\|_\infty \leq c_b\beta;$$

(A3) $\text{meas}(\Gamma_D) > 0$ and Γ_D contains the inflow boundary

$$\{x \in \partial\Omega : \boldsymbol{a}(x) \cdot \boldsymbol{n}(x) < 0\}.$$

Depending on the magnitude of the α (with respect to $\boldsymbol{a}$ and b), two regimes are studied in this chapter:

1. **diffusion-dominated regime**: there exists a constant $C_b \geq 0$ such that

$$|\boldsymbol{a}(x)/\alpha(x)| \leq C_b \quad \text{and} \quad |b(x)/\alpha(x)| \leq C_b \quad \text{for almost all } x \in \Omega;$$

2. **convection/reaction-dominated regime**: $\alpha(x) \equiv \epsilon \ll 1$ for a constant ϵ.

Define the bilinear form on $H^1_D(\Omega)$ by

$$B(u, v) := (\alpha \nabla u, \nabla v) + (\boldsymbol{a} \cdot \nabla u, v) + (bu, v), \quad \forall\, u, v \in H^1_D(\Omega),$$

where $(\cdot, \cdot)_S$ denote the L^2 inner product on set S and the subscript S is omitted when $S = \Omega$. The L^2 norm on S is denoted by $\|\cdot\|_S$.

The weak formulation of (5.1.1) is to find $u \in H^1_D(\Omega)$ such that

$$B(u, v) = (f, v) - (g_N, v)_{\Gamma_N}, \quad \forall\, v \in H^1_D(\Omega). \tag{5.1.2}$$

According to integration by parts and the assumptions in (A2) – (A3), it follows that, for any $v \in H^1_D(\Omega)$,

$$(\boldsymbol{a} \cdot \nabla v, v) + (bv, v) = \frac{1}{2}(v^2, \boldsymbol{a} \cdot \boldsymbol{n})_{\Gamma_N} + (v^2, b - \frac{1}{2}\text{div } \boldsymbol{a}) \geq \beta \|v\|^2, \tag{5.1.3}$$

where $\boldsymbol{n}$ denotes the unit outward vector normal to Γ_N. The energy norm induced by $B(\cdot, \cdot)$ is defined by

$$|||v||| = \left(\|\alpha^{1/2} \nabla v\|^2 + \beta \|v\|^2\right)^{1/2}, \quad \forall\, v \in H^1_D(\Omega),$$

where $|||\cdot|||_S$ denotes the energy norm over S and the subscript S is omitted when $S = \Omega$.

5.1.1 Finite element approximations

Let the mesh information (elements,vertices,edges/faces,etc.) be given in Section 3.1. Denote by R_K the inradius of the element $K \in \mathcal{T}$. The shape regularity of the triangulation requires the existence of a generic constant $C_0 > 1$ such that

$$h_K \leq C_0 R_K, \quad \forall\, K \in \mathcal{T} \tag{5.1.4}$$

holds true for each mesh $\mathcal{T}$ in the adaptive mesh refinement procedure.

The conforming finite element space of order k ($k \geq 1$) is given by

$$V_{\mathcal{T}} := \{v \in C(\Omega) : v|_K \in P_k(K),\ \forall K \in \mathcal{T},\ \text{and } v|_{\Gamma_D} = 0\}.$$

The projected data $\bar{f}$ and $\bar{g}_{_N}$ are defined by

$$\bar{f}|_K := \Pi_K^{k-1} f, \quad \forall\, K \in \mathcal{T} \quad \text{and} \quad \bar{g}_{_N}|_e := \Pi_K^{k-1} g_{_N}, \quad \forall\, e \in \mathcal{E}_N,$$

respectively.

The standard finite element approximation for problem (5.1.1) is to find $u_{_\mathcal{T}} \in V_{\mathcal{T}}$ such that

$$B(u_{_\mathcal{T}}, v) = (f, v) - (g_{_N}, v)_{\Gamma_N}, \quad \forall\, v \in V_{\mathcal{T}}. \tag{5.1.5}$$

In the case that the convection is dominant, one often adds a stabilization term along the convective direction. For example, the so-called SUPG method in [56] is to find $u_{_\mathcal{T}} \in V_{\mathcal{T}}$ such that

$$B_\delta(u_{_\mathcal{T}}, v) = l_\delta(v), \quad \forall\, v \in V_{\mathcal{T}}, \tag{5.1.6}$$

where the stabilized bilinear and linear forms are given by

$$B_\delta(u_{_\mathcal{T}}, v) = B(u_{_\mathcal{T}}, v) + \sum_{K \in \mathcal{T}} \delta_K(-\epsilon \Delta u_{_\mathcal{T}} + \boldsymbol{a} \cdot \nabla u_{_\mathcal{T}} + b u_{_\mathcal{T}}, \boldsymbol{a} \cdot \nabla v)_K,\ \forall\, u_{_\mathcal{T}}, v \in V_{\mathcal{T}}$$

and

$$l_\delta(v) = (f, v) - (g_{_N}, v)_{\Gamma_N} + \sum_{K \in \mathcal{T}} \delta_K(f, \boldsymbol{a} \cdot \nabla v)_K,$$

respectively. Here, the stabilization parameters δ_K are nonnegative and satisfy

$$\delta_K \|\boldsymbol{a}\|_{L^\infty(K)} \leq Ch_K, \quad \forall\, K \in \mathcal{T}.$$

To measure the convective derivative, the following dual norm was used in [2]:

$$|||\phi|||_* = \sum_{v \in H_D^1(\Omega)\setminus\{0\}} \frac{\langle \phi, v\rangle}{|||v|||},$$

where ϕ is in the dual space of $H_D^1(\Omega)$ and $\langle\cdot,\cdot\rangle$ denotes the duality pairing. Following [2], the dual norm will be combined with the energy norm to measure the approximation error in the convection/reaction-dominated regime.

To keep the exposition simple, we ignore data oscillation in coefficients by further assuming that for each $K \in \mathcal{T}$, $\alpha_{_K} := \alpha|_K$ is constant, $\boldsymbol{a}|_K$ is constant and $b|_K$ is constant. The algorithm and analysis remain valid without this assumption if we replace those quantities with their proper projections and add the corresponding oscillation error in the estimates.

Define $\alpha_e := \max_{K\subseteq\omega_e} \alpha_{_K}$, $\alpha_{\max} := \max_{K\in\mathcal{T}} \alpha_{_K}$ and $\alpha_{\min} := \min_{K\in\mathcal{T}} \alpha_{_K}$.

For diffusion-dominated case, we are interested in the case where α may be discontinuous and the hybrid estimator is robust with respect to the discontinuity. For convection/reaction-dominated case ($\alpha = \epsilon \ll 1$), we design a hybrid estimator that is robust with respect to ϵ in appropriate norms.

5.2 Residual Estimator

Let $\boldsymbol{\sigma}_{_\mathcal{T}} = -\alpha\nabla u_{_\mathcal{T}}$ be the numerical flux. For $S \in \mathcal{T}\cup\mathcal{E}$, define the weight $\gamma_{_S}$ as below

$$\gamma_{_S} = \begin{cases} 1, & \text{in diffusion-dominated regime,} \\ \min\{1, h_{_S}^{-1}\alpha_{_S}^{1/2}\beta^{-1/2}\}, & \text{in convection/reaction-dominated regime.} \end{cases} \tag{5.2.1}$$

The residual estimator is defined by

$$\eta = \left(\sum_{K\in\mathcal{T}} \eta_K^2\right)^{1/2} \text{ with } \eta_K^2 = \gamma_{_K}^2 h_K^2 \alpha_{_K}^{-1}\|r_K\|_K^2 + \frac{1}{2}\sum_{e\in\mathcal{E}_K} \gamma_e h_e \alpha_e^{-1}\|j_e\|_e^2, \tag{5.2.2}$$

where the element residual $r_K \in L^2(K)$ is given by

$$r_K := \bar{f} - \boldsymbol{a} \cdot \nabla u_\tau - b u_\tau - \operatorname{div} \boldsymbol{\sigma}_\tau \tag{5.2.3}$$

and the flux jump j_e is defined in (4.3.2). Note that for diffusion-dominated problems, the η is a simple extension of the one in [8] or [9] for pure diffusion problems; for convection/reaction-dominated problems, the η is introduced by Verfürth in [2]. In fact, $\gamma_S h_S \epsilon_S^{-1/2}$ is same to the weight defined in [2, Eq.(3.4)]. Here we use γ_S only for notational purpose, which does not change the definition of η_K in [2].

In the remainder of this section, we will discuss reliability and efficiency bounds of the estimator η.

5.2.1 Convection/reaction-dominated regime

The global reliability and efficiency bounds of the η were established by by Verfürth in [2, Theorem 4.1]. Those estimates are uniform with respect to the perturbation parameter ϵ. For reader's convenience, they are cited below.

Theorem 5.2.1. *Define the data oscillation on K by*

$$\Theta_K^2 = \gamma_K^2 h_K^2 \alpha_K^{-1} \|f - \bar{f}\|_K^2 + \sum_{e \in \mathcal{E}_K \cap \mathcal{E}_N} \gamma_e h_e \alpha_e^{-1} \|g_N - \bar{g}_N\|_e^2.$$

Let u and u_τ be the solutions of (5.1.2) *and* (5.1.6)*, respectively. Let η_K be defined in* (5.2.2)*. Then*

$$|||u - u_\tau|||^2 + |||\boldsymbol{a} \cdot \nabla(u - u_\tau)|||_*^2 \le C_1 \sum_{K \in \mathcal{T}} \left(\eta_K^2 + \Theta_K^2\right)$$

and

$$\sum_{K \in \mathcal{T}} \eta_K^2 \le C_2 \left(|||u - u_\tau|||^2 + |||\boldsymbol{a} \cdot \nabla(u - u_\tau)|||_*^2 + \sum_{K \in \mathcal{T}} \Theta_K^2 \right),$$

where the constants C_1 and C_2 are independent of ϵ, β, and any mesh-size.

The reaction-dominated diffusion problem, i.e., with $\boldsymbol{a} = 0$, $b = 1$ and $\Gamma_D = \partial\Omega$, corresponds to the singularly perturbed reaction-diffusion equation:

$$\begin{cases} -\epsilon \Delta u + u = f & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega. \end{cases} \tag{5.2.4}$$

The assumptions in Section 5.1 are fulfilled with $\beta = c_b = 1$. For the finite element approximation in (5.1.5) to the problem in (5.2.4), the residual estimator η in (5.2.2) is proved by Verfürth in [49] to be globally reliable (see Theorem 5.2.1); moreover, it is not only globally but also locally efficient.

Theorem 5.2.2. *Let u and u_τ be the respective exact and finite element solutions of the reaction-diffusion equation in* (5.2.4)*. The residual error indicator η_K in* (5.2.2) *satisfies*

$$\eta_K^2 \leq C \left(|||u - u_\tau|||_{\omega_K}^2 + \sum_{K' \subset \omega_K} \Theta_{K'}^2 \right),$$

where C is independent of ϵ and any mesh size.

5.2.2 Diffusion-dominated regime

For the diffusion-dominated case, following the pure diffusion case in [8] or [9], this section establishes global reliability and local efficiency of the estimator η. These estimates are proved to be robust in terms of the discontinuity of the diffusion coefficient α.

To this end, an alternative and often used expression of the element residual r_K in (5.2.3) is given in terms of the true error and the data oscillation:

$$r_K = \operatorname{div}(\alpha \nabla(u_\tau - u)) + \boldsymbol{a} \cdot \nabla(u - u_\tau) + b(u - u_\tau) + \bar{f} - f. \tag{5.2.5}$$

In the following, we will use C with or without subscripts to denote a generic non-negative constant, possibly different at different occurrences, that is independent of $\alpha_{\max}/\alpha_{\min}$ and any mesh-size, but may depend on shape parameter of the mesh $\mathcal{T}$, the polynomial degree k.

Theorem 5.2.3. *Let u be the exact solution of* (5.1.2) *and u_τ be the finite element solution in* (5.1.5)*. Under the monotonicity assumption* [8, Hypothesis 2.7] *of α, the residual estimator η satisfies*

$$|||u - u_\tau|||^2 \leq C \left(\eta^2 + \sum_{K \in \mathcal{T}} \Theta_K^2 \right)$$

with constant C independent of $\alpha_{\max}/\alpha_{\min}$ and any mesh-size.

Proof. Let $w := u - u_{\mathcal{T}}$. To prove the reliability bound, according to [8], it suffices to derive an estimate of the error in the form below:

$$\begin{aligned}|||w|||^2 \leq &\sum_{K\in\mathcal{T}}(r_K, w - w_{\mathcal{T}})_K + \sum_{e\in\mathcal{E}}(j_e, w - w_{\mathcal{T}})_e \\ &+ \sum_{K\in\mathcal{T}}(\bar{f} - f, w - w_{\mathcal{T}})_K + \sum_{e\in\mathcal{E}_N}(\bar{g}_N - g_N, w - w_{\mathcal{T}})_e, \quad \forall\, w_{\mathcal{T}} \in V_{\mathcal{T}}.\end{aligned} \tag{5.2.6}$$

To do so, it follows from integration by parts, (5.2.5), and the error equation $B(w, w_{\mathcal{T}}) = 0$ for all $w_{\mathcal{T}} \in V_{\mathcal{T}}$ that

$$\begin{aligned}(\alpha\nabla w, \nabla w) =&(\alpha\nabla w, \nabla(w - w_{\mathcal{T}})) + (\alpha\nabla w, \nabla w_{\mathcal{T}}) \\ &= \sum_{K\in\mathcal{T}}(r_K, w - w_{\mathcal{T}})_K + \sum_{e\in\mathcal{E}}(j_e, w - w_{\mathcal{T}})_e - (\boldsymbol{a}\cdot\nabla w, w) - (bw, w) \\ &\quad + \sum_{K\in\mathcal{T}}(\bar{f} - f, w - w_{\mathcal{T}})_K + \sum_{e\in\mathcal{E}_N}(\bar{g}_N - g_N, w - w_{\mathcal{T}})_e,\end{aligned}$$

which, together with (5.1.3), yields (5.2.6). This completes the proof of the theorem. □

Theorem 5.2.4. *Let u be the exact solution of* (5.1.2) *and $u_{\mathcal{T}}$ be the finite element solution in* (5.1.5)*. Then there exists a constant C independent of $\alpha_{\max}/\alpha_{\min}$ and any mesh-size such that*

$$\eta_K \leq C\left(|||u_{\mathcal{T}} - u|||_{\omega_K} + \sum_{K'\subseteq\omega_K}\Theta_{K'}\right), \quad \forall\, K \in \mathcal{T}.$$

Proof. The proof of local efficiency follows [8]. Let $w = u - u_{\mathcal{T}}$. Let ψ_K denote the element bubble supported on $K \in \mathcal{T}$ and define $v_K = \psi_K r_K$. The property of the bubble function, (5.2.5), integration by parts, and the Cauchy-Schwarz inequality imply

$$\|r_K\|_K^2 \leq C\|r_K\|_K\left(h_K^{-1}\|\alpha\nabla w\|_K + \|\boldsymbol{a}\cdot\nabla w\|_K + \|bu\|_K + \|\bar{f} - f\|_K\right).$$

The assumptions in Section 5.1 give that

$$h_K\alpha_K^{-1/2}\|\boldsymbol{a}\cdot\nabla w\|_K \leq h_K C_b\|\alpha^{1/2}\nabla w\|_K \leq C|||w|||_K \tag{5.2.7}$$

and that

$$h_K \alpha_K^{-1/2} \|bw\|_K \leq h_K C_b^{1/2} c_b^{1/2} \beta^{1/2} \|w\|_K \leq C |||w|||_K. \tag{5.2.8}$$

Combining the above three inequalities yield Hence it follows that

$$h_K \alpha_K^{-1/2} \|r_K\|_K \leq C \left(|||w|||_K + \Theta_K \right). \tag{5.2.9}$$

Let ψ_e denote the face bubble associated with $e \in \mathcal{E}$ and define $v_e = \psi_e j_e$. It follows from the property of the bubble function, integration by parts, (5.2.5), (5.2.7), (5.2.8), and (5.2.9) that

$$\begin{aligned}
\|j_e\|_e^2 &\leq C \int_e j_e v_e ds \\
&= C \sum_{K \subseteq \omega_e} \int_K \alpha \nabla w \cdot \nabla v_e dx + \int_K v_e \mathrm{div}\,(\alpha \nabla w) dx + \int_{e \cap \Gamma_N} (\bar{g}_N - g_N) v_e ds \\
&\leq C \|j_e\|_e \left(\alpha_e^{1/2} h_e^{-1/2} |||w|||_{\omega_e} + \sum_{K \subseteq \omega_e} h_e^{1/2} \|\bar{f} - f\|_K + \|\bar{g}_N - g_N\|_{e \cap \Gamma_N} \right),
\end{aligned}$$

which implies

$$h_e^{1/2} \alpha_e^{-1/2} \|j_e\|_e \leq C \left(|||w|||_{\omega_e} + \sum_{K \subseteq \omega_e} \Theta_K \right).$$

This, together with (5.2.9), shows the local efficiency bound. □

5.3 Flux Recovery

We show in this section how to recover a suitable flux in $H(\mathrm{div}; \Omega)$, denoted by $\hat{\boldsymbol{\sigma}}_{\mathcal{T}}$, such that the resulting hybrid estimator is robust.

For the diffusion-dominated case, $\hat{\boldsymbol{\sigma}}_{\mathcal{T}}$ is defined in (5.3.2), which is analogous to (4.2.8) for pure diffusion problems. For convection/reaction-dominated case, the definition for $\hat{\boldsymbol{\sigma}}_{\mathcal{T}}$ in an element K depends on the size the element K. Its construction remains the same as in (5.3.2) if R_K is relatively small. Otherwise, it is essentially the numerical flux with transitional regions to be in $H(\mathrm{div}; \Omega)$.

5.3.1 Diffusion-dominated regime

In the diffusion-dominated regime, the flux recovery is a natural extension of the one in (4.2.8) for the pure diffusion problems. Same to Section 4.2, the normal component $\hat{\boldsymbol{\sigma}}_{\mathcal{T}}|_K \cdot \boldsymbol{n} = \hat{g}_{_K} \in L^2(\partial K)$ is defined as

$$\hat{g}_{_K}|_e := \begin{cases} \lambda_{K,e}\boldsymbol{\sigma}_{\mathcal{T}}|_K \cdot \boldsymbol{n} + (1-\lambda_{K,e})\boldsymbol{\sigma}_{\mathcal{T}}|_{K'_e} \cdot \boldsymbol{n}, & \text{if } e \in \mathcal{E}_K \cap \mathcal{E}_I, \\ \bar{g}_{_N}, & \text{if } e \in \mathcal{E}_K \cap \mathcal{E}_N, \\ \boldsymbol{\sigma}_{\mathcal{T}} \cdot \boldsymbol{n}, & \text{if } e \in \mathcal{E}_K \cap \mathcal{E}_D, \end{cases} \tag{5.3.1}$$

where

$$\lambda_{K,e} = \frac{\alpha_{_K}^{-1} h_K}{\alpha_{_K}^{-1} h_K + \alpha_{_{K'_e}}^{-1} h_{K'_e}}.$$

We define $\hat{\boldsymbol{\sigma}}_{\mathcal{T}}|_K \in \mathrm{RT}_{k-1}(K)$ by assigning the degrees of freedom:

$$\begin{cases} \hat{\boldsymbol{\sigma}}_{\mathcal{T}} \cdot \boldsymbol{n} = \hat{g}_{_K} \quad \text{on } \partial K, \\ \operatorname{div} \hat{\boldsymbol{\sigma}}_{\mathcal{T}} = \Pi_K^{k-1}(\bar{f} - \boldsymbol{a} \cdot \nabla u_{\mathcal{T}} - b u_{\mathcal{T}}) + J_K \quad \text{in } K, \\ \displaystyle\int_K \hat{\boldsymbol{\sigma}}_{\mathcal{T}} \cdot \boldsymbol{q} dx = \int_K \boldsymbol{\sigma}_{\mathcal{T}} \cdot \boldsymbol{q} dx \quad \forall\, \boldsymbol{q} \in Q_{k-2}(K), \quad k \geq 2, \end{cases} \tag{5.3.2}$$

where

$$J_K := |K|^{-1} \left(\int_{\partial K} \hat{g}_{_K} ds + \int_K \boldsymbol{a} \cdot \nabla u_{\mathcal{T}} + b u_{\mathcal{T}} - f dx \right) \tag{5.3.3}$$

and $Q_{k-2}(K)$ is defined in (4.2.11). It is easily seen that $\hat{\boldsymbol{\sigma}}_{\mathcal{T}} \in H(\mathrm{div}; \Omega)$ and $\hat{\boldsymbol{\sigma}}_{\mathcal{T}}$ coincides with the one in (4.2.8) when $\boldsymbol{a}$ and b vanish.

5.3.2 Flux recovery in convection/reaction-dominated Regime

In the case that $\alpha = \epsilon$ is a constant, the weight is $\lambda_{K,e} = \frac{h_K}{h_K + h_{K'_e}}$. One may simplify it as $\lambda_{K,e} = \frac{1}{2}$ for all $K \in \mathcal{T}$ if the size of each element is close to sizes of its neighboring elements. The flux recovery on element $K \in \mathcal{T}$ depends on the size of K.

Flux recovery in K with $R_K \leq \kappa^{-1}$. If $R_K \leq \kappa^{-1}$, the recovered flux $\hat{\boldsymbol{\sigma}}_\tau|_K \in \mathrm{RT}_{k-1}(K)$ is defined as in (5.3.2).

Remark 5.3.1. *The case of absent reaction corresponds to $\beta = 0$, or equivalently, $\beta^{-1/2} = \infty$. Consequently, $R_K \leq \epsilon^{1/2}\beta^{-1/2}$ holds true for all $K \in \mathcal{T}$. Therefore, the recovered flux in each element is given by* (5.3.2).

Flux recovery in K with $R_K > \kappa^{-1}$. Here we assume for simplicity that $\Omega \subset \mathbb{R}^2$. The three dimensional case will be discussed in the future.

If K is an element with $R_K > \kappa^{-1}$, then as in [55], we partition K into a triangle K_Δ and three trapezoids illustrated in Figure 5.1. The edges of the triangle K_Δ are parallel to corresponding edges of K and the distance between each pair of parallel edges from K_Δ to K is equal to κ^{-1}. Figure 5.2 illustrates the partition of the trapezoid on edge $e \in \mathcal{E}_K$, where the trapezoid is partitioned into a rectangle, denoted by Q, and two triangles, denoted by T and S, respectively.

The flux recovery in Q shares the similar idea as in [55], while the recovery in T or S is different from [55].

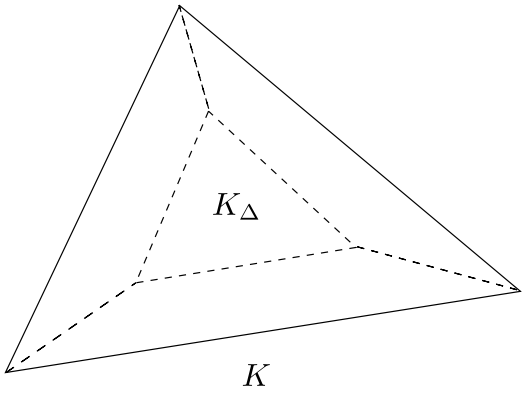

Figure 5.1. partition of element K

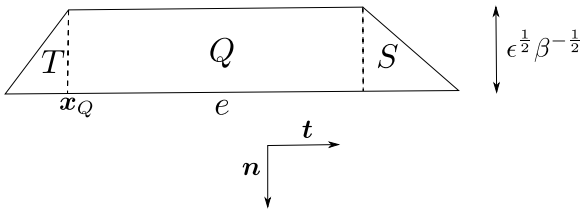

Figure 5.2. partition of a trapezoid in K

We define

$$\hat{\boldsymbol{\sigma}}_\tau = \boldsymbol{\sigma}_\tau \quad \text{in } K_\Delta, \tag{5.3.4}$$

and in each trapezoid as illustrated in Figure 5.2, with $\boldsymbol{x}_Q = \partial T \cap \partial Q \cap \partial K$ and $\boldsymbol{t}$ being a unit vector normal to $\boldsymbol{n}$, the recovered flux $\hat{\boldsymbol{\sigma}}_{\mathcal{T}}$ is given by:

- in Q,

$$\hat{\boldsymbol{\sigma}}_{\mathcal{T}} = \boldsymbol{\sigma}_{\mathcal{T}} + (\kappa y - 1)\left(\boldsymbol{\sigma}_{\mathcal{T}}(\boldsymbol{x}_Q + x\boldsymbol{t}) \cdot \boldsymbol{n} - \hat{g}_{K}(\boldsymbol{x}_Q + x\boldsymbol{t})\right)\boldsymbol{n}, \tag{5.3.5}$$

  where $x = (\boldsymbol{x} - \boldsymbol{x}_Q) \cdot \boldsymbol{t}$ and $y = (\boldsymbol{x}_Q - \boldsymbol{x}) \cdot \boldsymbol{n}$ are local coordinates of point $\boldsymbol{x} \in Q$;

- in $\omega = T$ or S, $\hat{\boldsymbol{\sigma}}_{\mathcal{T}}|_{\omega} \in \mathrm{RT}_{k-1}(\omega)$ has the following degrees of freedom:

$$\begin{cases} \hat{\boldsymbol{\sigma}}_{\mathcal{T}} \cdot \boldsymbol{n}_{\omega} = \boldsymbol{\sigma}_{\mathcal{T}} \cdot \boldsymbol{n}_{\omega} & \text{on } \partial\omega \backslash \partial K, \\ \hat{\boldsymbol{\sigma}}_{\mathcal{T}} \cdot \boldsymbol{n}_{\omega} = \hat{g}_{K} & \text{on } \partial\omega \cap \partial K, \\ \operatorname{div} \hat{\boldsymbol{\sigma}}_{\mathcal{T}} = \operatorname{div} \boldsymbol{\sigma}_{\mathcal{T}} + J_{\omega} & \text{in } \omega, \\ \displaystyle\int_{\omega} \hat{\boldsymbol{\sigma}}_{\mathcal{T}} \cdot \boldsymbol{q} dx = \int_{\omega} \boldsymbol{\sigma}_{\mathcal{T}} \cdot \boldsymbol{q} dx & \forall\, \boldsymbol{q} \in Q_{k-2}(\omega), \quad k \geq 2, \end{cases} \tag{5.3.6}$$

  where $\boldsymbol{n}_{\omega}$ denotes the unit outward normal for ω,

$$J_{\omega} := |\omega|^{-1} \left(\int_{\partial\omega \cap \partial K} \hat{g}_{K} - \boldsymbol{\sigma}_{\mathcal{T}}|_{K} \cdot \boldsymbol{n}_{\omega} ds \right) \tag{5.3.7}$$

  and the space $Q_{k-2}(\omega)$ is defined as in (4.2.11).

Inside K, it is easy to verify that the normal component of $\hat{\boldsymbol{\sigma}}_{\mathcal{T}}$ is continuous over $\partial\omega \backslash \partial K$ and $\partial Q \cap \partial K_{\Delta}$, whose value is actually equal to that of $\boldsymbol{\sigma}_{\mathcal{T}}$. Hence $\hat{\boldsymbol{\sigma}}_{\mathcal{T}} \in H(\mathrm{div}; K)$. On ∂K, the normal component of $\hat{\boldsymbol{\sigma}}_{\mathcal{T}}$ is always given by $\hat{g}_{K}$. Therefore, we see that $\hat{\boldsymbol{\sigma}}_{\mathcal{T}} \in H(\mathrm{div}; \Omega)$.

Remark 5.3.2. *It can be seen that the magnitude of the quantity* $\boldsymbol{\sigma}_{\mathcal{T}}(\boldsymbol{x}_Q + x\boldsymbol{t}) \cdot \boldsymbol{n} - \hat{g}_{K}(\boldsymbol{x}_Q + x\boldsymbol{t})$ *in* (5.3.5) *is bounded by the flux jump* $|j_e|$.

5.4 Hybrid Estimator

The hybrid estimator is defined as

$$\xi = \left(\sum_{K \in \mathcal{T}} \xi_K^2 \right)^{1/2} \quad \text{with} \quad \xi_K^2 = \|\alpha^{-1/2}(\hat{\boldsymbol{\sigma}}_{\mathcal{T}} - \boldsymbol{\sigma}_{\mathcal{T}})\|_K^2 + \gamma_K^2 h_K^2 \alpha_K^{-1} \|\hat{r}_K\|_K^2, \tag{5.4.1}$$

where $\gamma_{_K}$ is defined in (5.2.1) and the modified element residual $\hat{r}_K$ is given by

$$\hat{r}_K = \bar{f} - \operatorname{div} \hat{\boldsymbol{\sigma}}_{_\mathcal{T}} - \boldsymbol{a} \cdot \nabla u_{_\mathcal{T}} - b u_{_\mathcal{T}}, \quad \forall\, K \in \mathcal{T}. \tag{5.4.2}$$

To establish the reliability and the efficiency of the hybrid estimator ξ, we use the following local equivalence between ξ and η.

Lemma 5.4.1. *Let η_K be the residual estimator in* (5.2.2) *and ξ_K be the hybrid estimator in* (5.4.1)*, respectively. Then there exist positive constants C_1 and C_2 such that*

$$C_1 \xi_K \le \eta_K \le C_2 \sum_{K' \subset \omega_K} \xi_{K'}, \tag{5.4.3}$$

where C_1 and C_2 are independent of h_K and the parameters either α for dominant diffusion or ϵ and β for dominant convection/reaction.

See Section 5.5 for the proof of Lemma 5.4.1.

5.4.1 A posteriori error estimates

Thanks to the equivalence result in Lemma 5.4.1, the reliability and efficiency of the hybrid estimator ξ follow immediately from the corresponding results of the residual estimator η.

The estimates below of ξ in the diffusion-dominated case are parallel to Theorem 5.2.4 and Theorem 5.2.3 of η.

Theorem 5.4.1. *In the* diffusion-dominated *regime, let u be the exact solution of* (5.1.2) *and $u_{_\mathcal{T}}$ be the finite element solution in* (5.1.5)*. For the hybrid estimator ξ defined in* (5.4.1)*, there exists a constant C_1 independent of $\alpha_{\max}/\alpha_{\min}$ and any mesh-size such that*

$$\xi_K \le C_1 \left(|||u_{_\mathcal{T}} - u|||_{\omega_K} + \sum_{K' \subseteq \omega_K} \Theta_{K'} \right), \quad \forall\, K \in \mathcal{T}.$$

Furthermore, under the monotonicity assumption [8, Hypothesis 2.7] *of α, ξ satisfies*

$$|||u - u_{_\mathcal{T}}|||^2 \le C_2 \left(\xi^2 + \sum_{K \in \mathcal{T}} \Theta_K^2 \right)$$

with constant C_2 independent of $\alpha_{\max}/\alpha_{\min}$ and any mesh-size.

The estimates of ξ in the convection/reaction-dominated case follow from Theorem 5.2.1.

Theorem 5.4.2. *In the* convection/reaction-dominated *regime, let u and $u_{\mathcal{T}}$ be the solutions of* (5.1.2) *and* (5.1.6)*, respectively. The hybrid estimator ξ defined in* (5.4.1) *satisfies*

$$|||u - u_{\mathcal{T}}|||^2 + |||\boldsymbol{a} \cdot \nabla(u - u_{\mathcal{T}})|||_*^2 \leq C_1 \left(\xi^2 + \sum_{K \in \mathcal{T}} \Theta_K^2 \right)$$

and

$$\xi^2 \leq C_2 \left(|||u - u_{\mathcal{T}}|||^2 + |||\boldsymbol{a} \cdot \nabla(u - u_{\mathcal{T}})|||_*^2 + \sum_{K \in \mathcal{T}} \Theta_K^2 \right),$$

where the constants C_1 and C_2 are independent of ϵ, β and any mesh-size.

For the singularly perturbed reaction-diffusion equation (5.2.4), in addition to the *global* efficiency bound in Theorem 5.4.2, ξ_K satisfies a *local* efficiency bound as a counterpart of Theorem 5.2.2.

Theorem 5.4.3. *Let u and $u_{\mathcal{T}}$ be the respective exact and finite element solutions of the singularly perturbed reaction-diffusion equation in* (5.2.4)*. The hybrid error indicator ξ_K in* (5.4.1) *satisfies the following estimate*

$$\xi_K^2 \leq C \left(|||u - u_{\mathcal{T}}|||_{\omega_K}^2 + \sum_{K' \subset \omega_K} \Theta_{K'}^2 \right),$$

where C is independent of ϵ and any mesh size.

Remark 5.4.1. *For the singularly perturbed reaction-diffusion equation* (5.2.4)*, in addition to the robust residual estimator in [49], other types of estimators have been proposed over the years. A general recovery-based estimator was proposed in [11] based on projecting the numerical flux onto an $H(\mathrm{div}; \Omega)$-conforming space. The global projection, however, can be computationally expensive. A series of estimators based on flux equilibration were explored in [55, 57, 58], where the estimators yield guaranteed upper bounds of the true error. The flux equilibrations in [55, 58] require solving local*

least square problems associated with each vertex patch and only piecewise linear finite element approximations were discussed.

5.5 Proof Of Lemma 5.4.1

In diffusion-dominated regime, the proof of (5.4.3) is the same as that of Theorem 4.3.1. In convection/reaction-dominated regime, the proof of (5.4.3) is more complicated. We first prove the upper bound of (5.4.3) in Section 5.5.1. The lower bound of (5.4.3) will be justified in Section 5.5.2.

In the proof below, C (with or without subscript) denotes a constant (possibly different at different occurrences) that depends only on the shape parameter or the polynomial degree k, but is independent of ϵ, β and any mesh-size.

5.5.1 Upper bound

The proof of the upper bound is analogous to that in Theorem 4.3.1, which is essentially based on the proof of local efficiency of residual estimator using properly chosen bubble functions.

Let ψ_K and ψ_e be the element and face bubble functions defined in [2], associated with element K and face e, respectively, It is known from [2, Lemma 3.6] that

$$\begin{cases} \|\psi_K r_K\|_K \le \|r_K\|_K \le C(\psi_K r_K, r_K)_K, \\ \|\nabla(\psi_K r_K)\|_K \le C\gamma_K^{-1} h_K^{-1} \|r_K\|_K, \\ \|j_e\|_e^2 \le C(j_e, \psi_e j_e)_e, \\ \|\nabla(\psi_e j_e)\|_K \le C\gamma_e^{-1/2} h_e^{-1/2} \|j_e\|_e, \quad K \subseteq \omega_e, \\ \text{and} \quad \|\psi_e j_e\|_K \le C\gamma_e^{1/2} h_e^{1/2} \|j_e\|_e, \quad K \subseteq \omega_e. \end{cases} \tag{5.5.1}$$

Proof of the upper bound in (5.4.3). With the help of (5.5.1), integration by parts, and the Cauchy-Schwarz inequality, we deduce that

$$\begin{aligned}\|r_K\|_K^2 &\le C\,(r_K, \psi_K r_K)_K = C\,((\mathrm{div}\,(\hat{\boldsymbol{\sigma}}_\tau - \boldsymbol{\sigma}_\tau), \psi_K r_K)_K + (\hat{r}_K, \psi_K r_K)_K)\\ &= C\,((\boldsymbol{\sigma}_\tau - \hat{\boldsymbol{\sigma}}_\tau, \nabla(\psi_K r_K))_K + (\hat{r}_K, \psi_K r_K)_K)\\ &\le C\|r_K\|_K \left(\gamma_K^{-1} h_K^{-1} \|\hat{\boldsymbol{\sigma}}_\tau - \boldsymbol{\sigma}_\tau\|_K + \|\hat{r}_K\|_K\right),\end{aligned}$$

which implies

$$\gamma_K h_K \epsilon^{-1/2} \|r_K\|_K \le C\left(\epsilon^{-1/2}\|\hat{\boldsymbol{\sigma}}_\tau - \boldsymbol{\sigma}_\tau\|_K + \gamma_K h_K \epsilon^{-1/2}\|\hat{r}_K\|_K\right) \le C\xi_K. \tag{5.5.2}$$

To estimate $\|j_e\|_e$, it follows from (5.5.1), integration by parts, the Cauchy-Schwarz and the triangle inequalities, and (5.5.2) that

$$\begin{aligned}\|j_e\|^2 &\le C(j_e, \psi_e j_e)_e\\ &= C\sum_{K\subseteq\omega_e} ((\boldsymbol{\sigma}_\tau - \hat{\boldsymbol{\sigma}}_\tau, \nabla(\psi_e j_e))_K + (\mathrm{div}\,(\boldsymbol{\sigma}_\tau - \hat{\boldsymbol{\sigma}}_\tau), \psi_e j_e)_K)\\ &\le C\gamma_e^{-1/2} h_e^{-1/2} \epsilon^{1/2} \|j_e\|_e \sum_{K\subseteq\omega_e} \left(\epsilon^{-1/2}\|\hat{\boldsymbol{\sigma}}_\tau - \boldsymbol{\sigma}_\tau\|_K + \gamma_e h_e \epsilon^{-1/2}(\|\hat{r}_K\|_K + \|r_K\|_K)\right)\\ &\le C\gamma_e^{-1/2} h_e^{-1/2} \epsilon^{1/2} \|j_e\|_e \sum_{K\subseteq\omega_e} \xi_K.\end{aligned} \tag{5.5.3}$$

Now, the upper bound in (5.4.3) is a direct consequence of (5.5.2), (5.5.3), and the definition of η_K in (5.2.2). □

5.5.2 Lower bound

In this section, we prove the lower bound in (5.4.3) in the convection/reaction-dominated regime. That is, for any $K \in \mathcal{T}$, there exists a positive constant C independent of h_K, ϵ, and β such that

$$\xi_K \le C\eta_K. \tag{5.5.4}$$

This is proceeded in two cases: (i) $R_k \le \kappa^{-1}$ and (ii) $R_k > \kappa^{-1}$.

Proof in Case (i). In this case, the weight $\gamma_{_K} = O(1)$ independent of h_k, ϵ, and β. In fact, it follows from (5.1.4) that $\kappa^{-1} \geq R_K \geq \frac{1}{C_0} h_K$ and

$$1 \geq \gamma_{_K} = h_K^{-1} \min\{h_K, \epsilon^{1/2}\beta^{-1/2}\} \geq h_K^{-1} \min\{h_K, \frac{h_K}{C_0}\} = \frac{1}{C_0}. \tag{5.5.5}$$

Hence, to show the validity of (5.5.4), it suffices to show that

$$\|\hat{\boldsymbol{\sigma}}_{\tau} - \boldsymbol{\sigma}_{\tau}\|_K^2 + h_K^2 \|\hat{r}_K\|_K^2 \leq C \left(h_K^2 \|r_K\|_K^2 + \sum_{e \in \mathcal{E}_K} h_e \|j_e\|_e^2 \right). \tag{5.5.6}$$

To this end, note first that $\operatorname{div} \boldsymbol{\sigma}_{\tau}|_K \in P_{k-1}(K)$, the recovered flux defined in (5.3.2) satisfies

$$\operatorname{div} (\hat{\boldsymbol{\sigma}}_{\tau} - \boldsymbol{\sigma}_{\tau}) = \Pi_K^{k-1} r_K + J_K, \tag{5.5.7}$$

which, together with (4.3.1) and the triangle inequality, implies

$$\begin{aligned} \|\hat{\boldsymbol{\sigma}}_{\tau} - \boldsymbol{\sigma}_{\tau}\|_K &\leq C \left(h_K \|\operatorname{div} (\hat{\boldsymbol{\sigma}}_{\tau} - \boldsymbol{\sigma}_{\tau})\|_K + \sum_{e \in \mathcal{E}_K} h_e^{1/2} \|j_e\|_e \right) \\ &\leq C \left(h_K \|r_K\|_K + h_K \|J_K\| + \sum_{e \in \mathcal{E}_K} h_e^{1/2} \|j_e\|_e \right). \end{aligned} \tag{5.5.8}$$

To bound the modified element residual $\hat{r}_K$ in (5.4.2), by (5.5.7) we have

$$\hat{r}_K = r_K - \operatorname{div} (\hat{\boldsymbol{\sigma}}_{\tau} - \boldsymbol{\sigma}_{\tau}) = \left(I - \Pi_K^{k-1} \right) r_K - J_K, \tag{5.5.9}$$

which yields

$$\|\hat{r}_K\|_K \leq \|r_K\|_K + \|J_K\|_K. \tag{5.5.10}$$

Now, (5.5.6) is a direct consequence of (5.5.8), (5.5.10), and the following bound

$$\|J_K\| \leq C \left(\|r_K\|_K + \sum_{e \in \mathcal{E}_K} h_e^{-1/2} \|j_e\|_e \right),$$

which follows from the divergence theorem the triangle and the Cauchy-Schwarz inequalities that

$$\begin{aligned} |J_K| &= |K|^{-1} \left| \int_{\partial K} (\hat{g}_{_K} - \boldsymbol{\sigma}_{\tau}|_K \cdot \boldsymbol{n}) ds - \int_K r_K dx \right| \\ &\leq |K|^{-1} \left(\sum_{e \in \mathcal{E}_K} h_e^{1/2} \|j_e\|_e + |K|^{-1/2} \|r_K\|_K \right). \end{aligned}$$

This completes the proof of (5.5.6) and, hence, (5.5.4). □

Proof in Case (ii). When $R_K > \kappa^{-1}$, the fact that $h_S > R_S$ implies

$$\gamma_s = \epsilon^{1/2} h_S^{-1} \beta^{-1/2}, \quad S = K \text{ or } S \in \mathcal{E}_K. \tag{5.5.11}$$

To prove (5.5.4), it suffices to show that

$$\epsilon^{-1} \|\hat{\boldsymbol{\sigma}}_\tau - \boldsymbol{\sigma}_\tau\|_K^2 + \beta^{-1} \|\hat{r}_K\|_K^2 \leq C \left(\beta^{-1} \|r_K\|_K^2 + \sum_{e \in \mathcal{E}_K} \epsilon^{-1/2} \beta^{-1/2} \|j_e\|_e^2 \right). \tag{5.5.12}$$

First we estimate $\|\hat{\boldsymbol{\sigma}}_\tau - \boldsymbol{\sigma}_\tau\|_K$. In K_Δ, by definition, $\hat{\boldsymbol{\sigma}}_\tau - \boldsymbol{\sigma}_\tau = 0$. In each trapezoid illustrated in Figure 5.2, we estimate the error in Q and $\omega = T, S$ separately. From (5.3.5) and (5.5.11), it can be computed that (cf. [55, Lemma 3])

$$\epsilon^{-1/2} \|\hat{\boldsymbol{\sigma}}_\tau - \boldsymbol{\sigma}_\tau\|_Q \leq \frac{1}{\sqrt{12}} \epsilon^{-1/4} \beta^{-1/4} \|j_e\|_e. \tag{5.5.13}$$

The estimate of $\|\hat{\boldsymbol{\sigma}}_\tau - \boldsymbol{\sigma}_\tau\|_\omega$ is analogous to Case (i). Note that $\operatorname{diam}(\omega) = O(\kappa^{-1})$. It follows from (5.3.7) that

$$\|\operatorname{div}(\hat{\boldsymbol{\sigma}}_\tau - \boldsymbol{\sigma}_\tau)\|_\omega = \|J_\omega\|_\omega \leq C \epsilon^{-1/4} \beta^{1/4} \|j_e\|_e, \tag{5.5.14}$$

which, together with Lemma 4.3.1 and (5.5.11), implies

$$\epsilon^{-1/2} \|\hat{\boldsymbol{\sigma}}_\tau - \boldsymbol{\sigma}_\tau\|_\omega \leq C \left(\beta^{-1/2} \|J_\omega\|_\omega + \epsilon^{-1/4} \beta^{-1/4} \|j_e\|_e \right) \leq C \epsilon^{-1/4} \beta^{-1/4} \|j_e\|_e.$$

Hence

$$\epsilon^{-1} \|\hat{\boldsymbol{\sigma}}_\tau - \boldsymbol{\sigma}_\tau\|_K^2 \leq C \sum_{e \in \mathcal{E}_K} \epsilon^{-1/2} \beta^{-1/2} \|j_e\|_e^2. \tag{5.5.15}$$

It remains to estimate $\beta^{-1} \|\hat{r}_K\|_K^2$. In K_Δ, $\hat{\boldsymbol{\sigma}}_\tau = \boldsymbol{\sigma}_\tau$, so $\hat{r}_K = r_K$. Consider then each trapezoid illustrated in Figure 5.2. In ω, analogous to (5.5.9), we deduce from the triangle inequality, (5.5.11), and (5.5.14) that

$$\beta^{-1/2} \|\hat{r}_K\|_\omega \leq \beta^{-1/2} \left(\|J_\omega\|_\omega + \|r_K\|_\omega \right) \leq C \epsilon^{-1/4} \beta^{-1/4} \|j_e\|_e + \beta^{-1/2} \|r_K\|_K. \tag{5.5.16}$$

In Q, similar to [55], it can be computed that

$$\operatorname{div}(\hat{\boldsymbol{\sigma}}_\tau - \boldsymbol{\sigma}_\tau) = -\kappa \left(\boldsymbol{\sigma}_\tau(\boldsymbol{x}_Q + x\boldsymbol{t}) \cdot \boldsymbol{n} - \hat{g}_{_K}(\boldsymbol{x}_Q + x\boldsymbol{t}) \right) \quad \text{in } Q$$

and

$$\|\mathrm{div}\,(\hat{\boldsymbol{\sigma}}_{\mathcal{T}} - \boldsymbol{\sigma}_{\mathcal{T}})\|_Q \leq \frac{1}{2}\epsilon^{-1/4}\beta^{1/4}\|j_e\|_e.$$

The triangle inequality then yields

$$\begin{aligned}\beta^{-1/2}\|\hat{r}_K\|_Q &\leq \beta^{-1/2}\left(\|\mathrm{div}\,(\hat{\boldsymbol{\sigma}}_{\mathcal{T}} - \boldsymbol{\sigma}_{\mathcal{T}})\|_Q + \|r_K\|_Q\right)\\ &\leq C\epsilon^{-1/4}\beta^{-1/4}\|j_e\|_e + \beta^{-1/2}\|r_K\|_K.\end{aligned} \tag{5.5.17}$$

Therefore,

$$\beta^{-1}\|\hat{r}_K\|_K^2 \leq C\left(\sum_{e\in\mathcal{E}_K}\epsilon^{-1/2}\beta^{-1/2}\|j_e\|_e^2 + \beta^{-1}\|r_K\|_K^2\right),$$

which, together with (5.5.15), completes the proof of (5.5.12). □

5.6 Numerical Experiments

In the numerical experiments, we consider the singularly perturbed reaction-diffusion equation in (5.2.4). Here the domain is chosen as $\Omega = [-1,1]^2$. The initial mesh in the adaptive mesh refinement consists of 4×4 congruent squares, each of which is partitioned into two triangles connecting bottom-left and top-right corners. As in [46, 47], we use Dörfler's marking strategy [46] with $\theta_D = 0.5$ (cf. (4.5.1)). The newest-vertex bisection [44] is used in the refinement. For the finite element discretization, P_1 conforming element is used in all examples. The exact error is denoted by $e = u - u_{\mathcal{T}}$. "DOFs" denotes the degrees of freedom.

5.6.1 Example 1

We first consider an example as in [55, Example 1], where the solution is smooth but, as pointed out in [55], non-robust estimators do not perform well . The exact solution is chosen as

$$u(x,y) = \cos(\pi x/2)\cos(\pi y/2)/(1+\epsilon\pi^2/2)$$

Table 5.1.
Example 1 - Effectivity indices for different ϵ on a uniform mesh

ϵ	1E-5	1E-4	5E-4	1E-3	5E-3	1E-2	5E-2	1E-1	1	10	100
η	0.66	0.66	0.93	1.21	2.22	2.81	4.83	5.58	5.57	5.56	5.56
ξ	0.80	0.84	1.09	1.35	1.20	1.28	1.38	1.38	1.36	1.36	1.36

and the data is $f(x, y) = \cos(\pi x/2)\cos(\pi y/2)$. The aim of this example is to show that the hybrid estimator ξ is more accurate than the residual estimator η and to demonstrate that ξ is less sensitive with respect to the variation of ϵ.

Effectivity with respect to ϵ . On a fixed uniform mesh composed of 200 isosceles right triangles, we vary ϵ and investigate the behavior of effectivity index for each estimator. The numerical results are collected in Table 5.1. It is easily seen that the hybrid estimator ξ is more accurate and than the residual estimator η. When ϵ changes gradually from 10^{-5} to 100, the effectivity indices for ξ remain close to 1, while the effectivity indices for η increase significantly from 0.66 to 5.56, by a factor of 8.4. This implies that, compared to the explicit residual estimator η, the hybrid estimator ξ is less sensitive to the size of reaction.

Effectivity during adaptive mesh refinement . In this subsection, for each fixed ϵ, we investigate the effectivity of each estimator during adaptive mesh refinement. The stopping criterion for AMR: DOFs > 50000.

The numerical results for residual estimator η are shown in Figure 5.3 – 5.5. It is obviously seen that the effectivity of η strongly depends on the relation between ϵ and mesh-size. Hence even though theoretically the constants in the a posteriori error estimates for η are independent of ϵ, the practical performance does display a difference for different choices of ϵ. This is because the theoretical results can only

provide lower and upper bounds of $\eta/|||e|||$, i.e., an interval that $\eta/|||e|||$ lies in, and the exact value may vary in such a (possibly large) interval. The size of this interval could be quite large even though it is independent of ϵ.

The numerical results for the hybrid estimator ξ are shown in Figure 5.6 – 5.8. Unlike η, numerical results indicate that ϵ does not have much influence on the effectivity of ξ during the adaptive mesh refinement. Therefore, we see that the hybrid estimator is indeed less sensitive than the residual estimator with respect to ϵ.

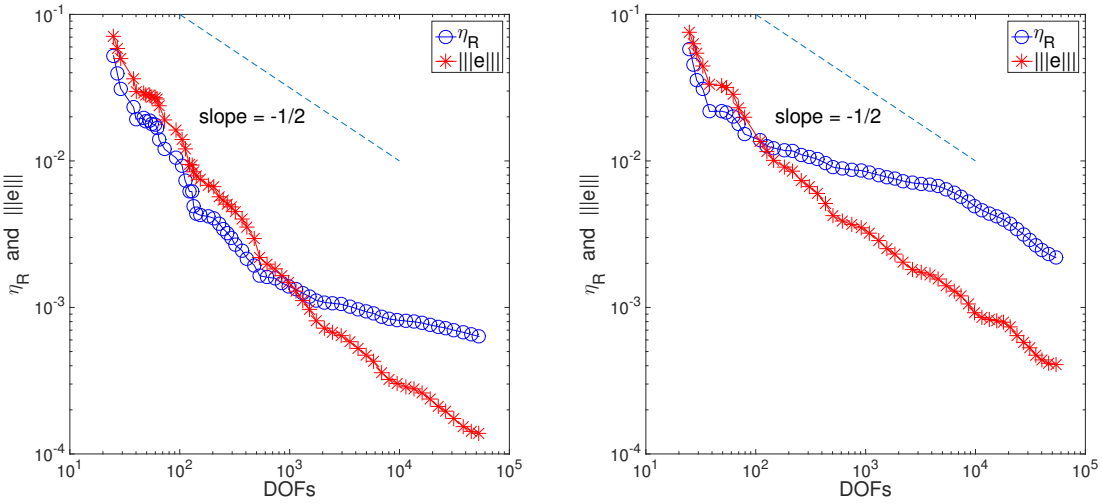

Figure 5.3. Example 1 - η and $|||e|||$ with $\epsilon = 10^{-4}$(left), $\epsilon = 10^{-3}$(right)

5.6.2 Example 2

Choose the exact solution as

$$u(x, y) = e^{-(x+1)/\sqrt{\epsilon}} + e^{-(y+1)/\sqrt{\epsilon}} \quad \text{with} \quad \epsilon = 10^{-4}.$$

Then $f = 0$ and u displays boundary layers along $x = -1$ and $y = -1$.

We perform adaptive mesh refinement with stopping criterion $|||e||| \leq 0.1|||u|||$ and compare the results obtained using η and ξ.

The meshes generated by η and ξ in Figure 5.9 are quite similar, where major refinements are along the boundary layers. From Figure 5.10 and Table 5.2, it is

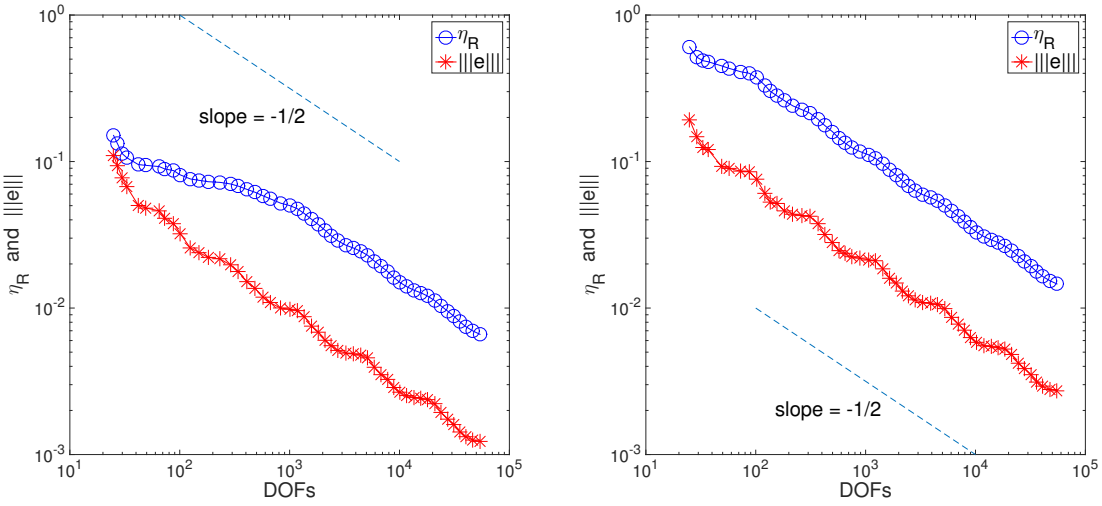

Figure 5.4. Example 1 - η and $|||e|||$ with $\epsilon = 10^{-2}$(left), $\epsilon = 10^{-1}$(right)

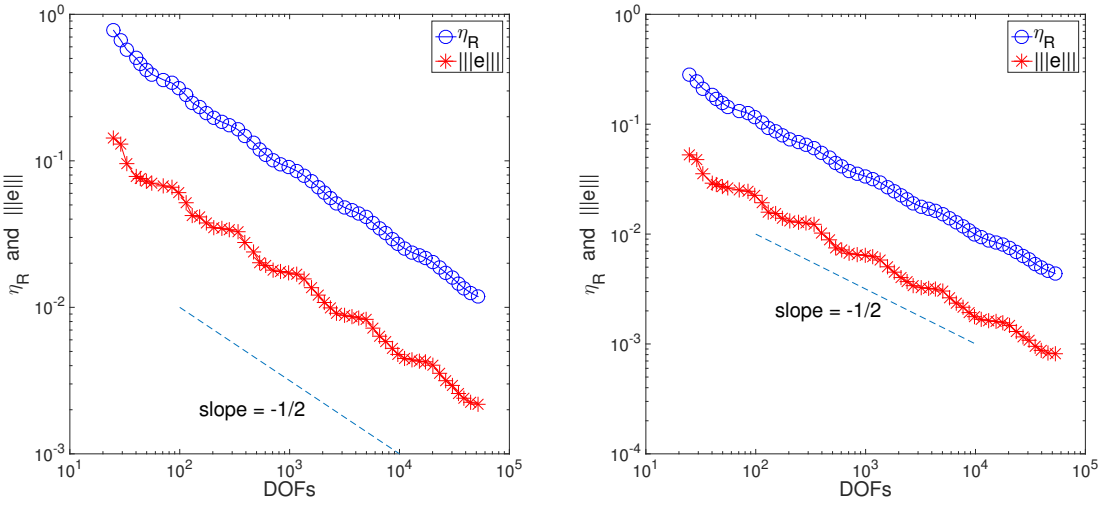

Figure 5.5. Example 1 - η and $|||e|||$ with $\epsilon = 1$(left), $\epsilon = 10$(right)

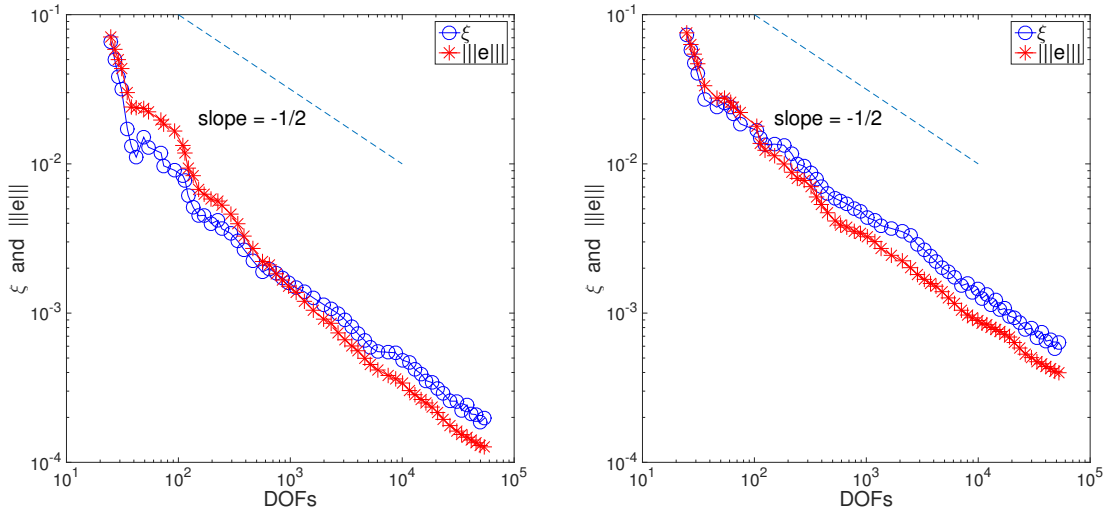

Figure 5.6. Example 1 - ξ and $|||e|||$ with $\epsilon = 10^{-4}$(left), $\epsilon = 10^{-3}$(right)

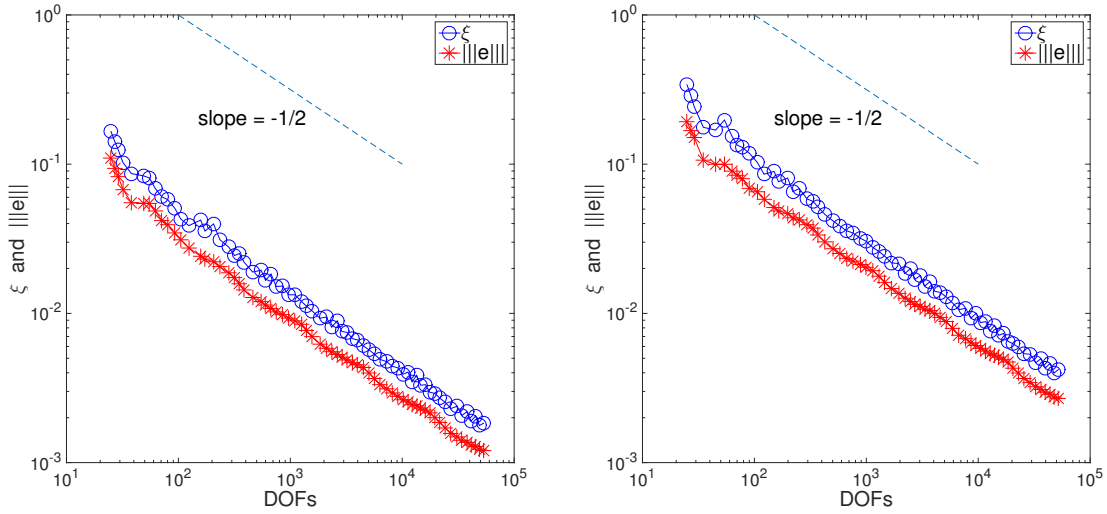

Figure 5.7. Example 1 - ξ and $|||e|||$ with $\epsilon = 10^{-2}$(left), $\epsilon = 10^{-1}$(right)

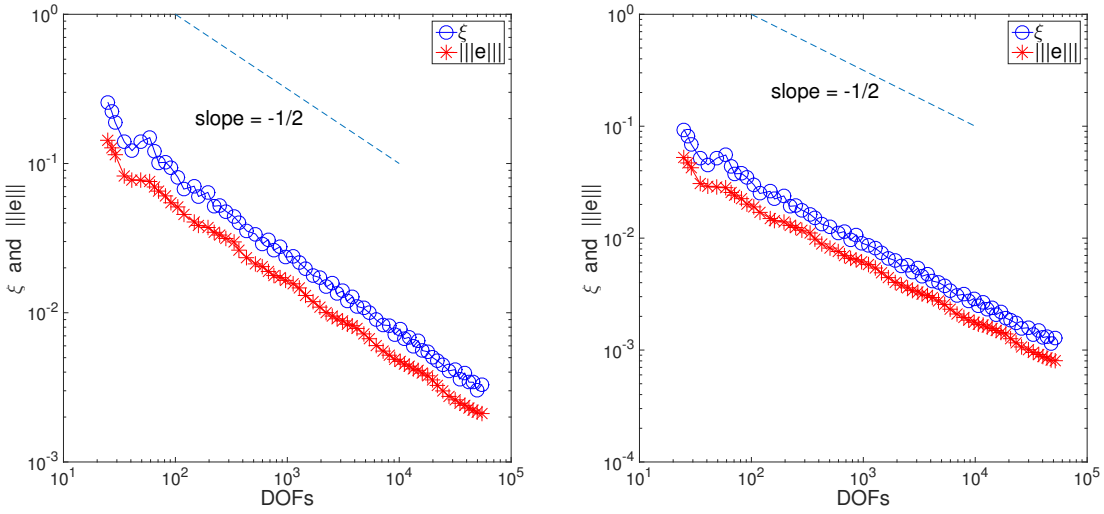

Figure 5.8. Example 1 - ξ and $|||e|||$ with $\epsilon = 1$(left), $\epsilon = 10$(right)

easy to see that the residual estimator is less accurate than the hybrid estimator. Moreover, Figure 5.10 again shows that the effectivity index of η varies more widely than that of ξ during the mesh refinement procedure.

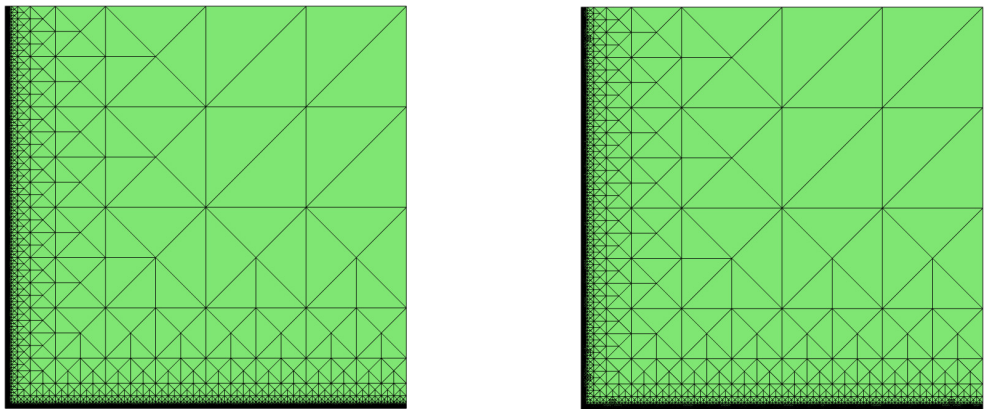

Figure 5.9. Example 2 - meshes generated using η(left) and ξ(right)

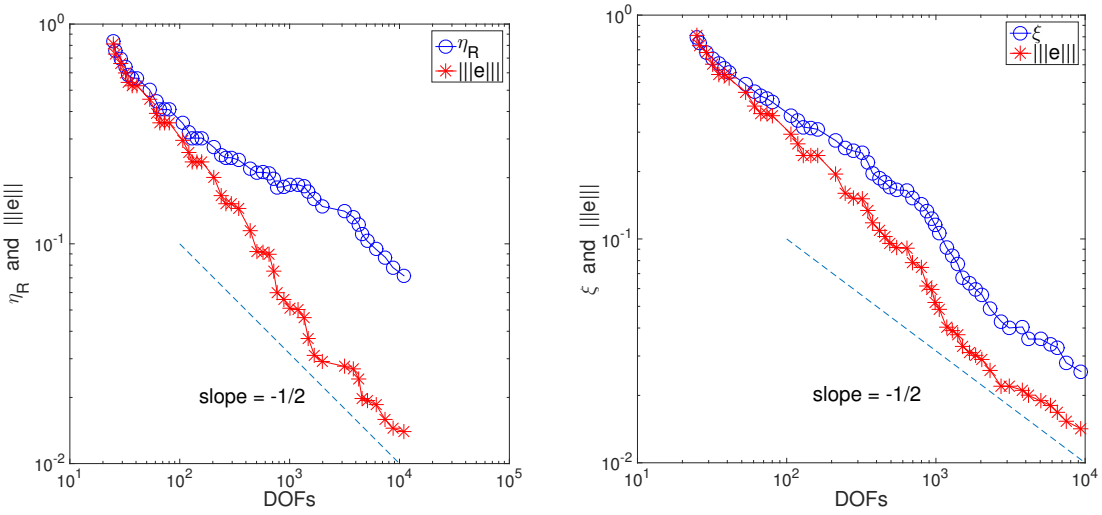

Figure 5.10. Example 2 - error plot for η(left) and ξ(right)

Table 5.2.
Example 2 - results for η and ξ

estimator	DOFs	$\|\|\|e\|\|\|/\|\|\|u\|\|\|$	eff-ind
η	10987	9.8E-2	5.11
ξ	9383	9.9E-2	1.80

5.6.3 Example 3

The exact solution is chosen as

$$u(x,y) = \tanh\left(\epsilon^{-1/2}(x^2+y^2-\frac{1}{4})\right) \quad \text{with} \quad \epsilon = 10^{-4},$$

which displays an interior layer on a circle with radius $\frac{1}{2}$.

The stopping criterion for the adaptive mesh refinement is chosen as $|||e||| < 0.01|||u|||$. The meshes generated by η and ξ in Figure 5.11 are quite similar, where major refinements are along the interior layer. In terms of accuracy or effectivity, from Table 5.3 and Figure 5.12, same conclusion can be drawn as in Example 2: the residual estimator is less accurate than the hybrid estimator and its effectivity index varies more widely during the adaptive mesh refinement.

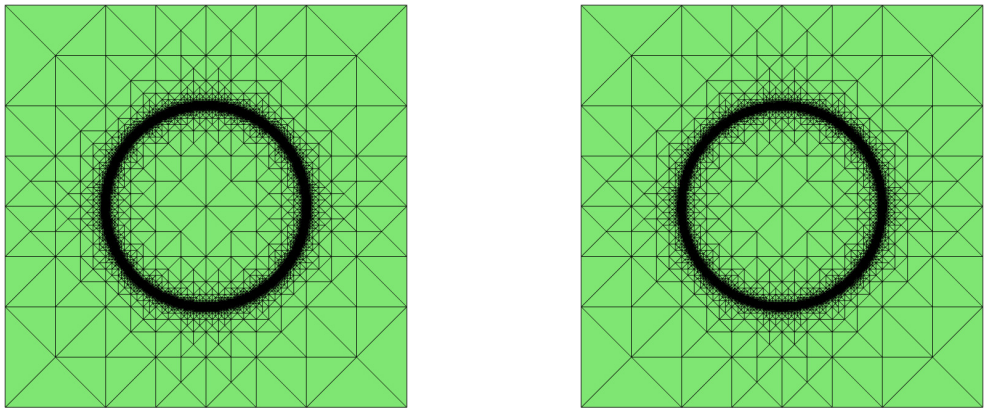

Figure 5.11. Example 3 - meshes generated using η(left) and ξ(right)

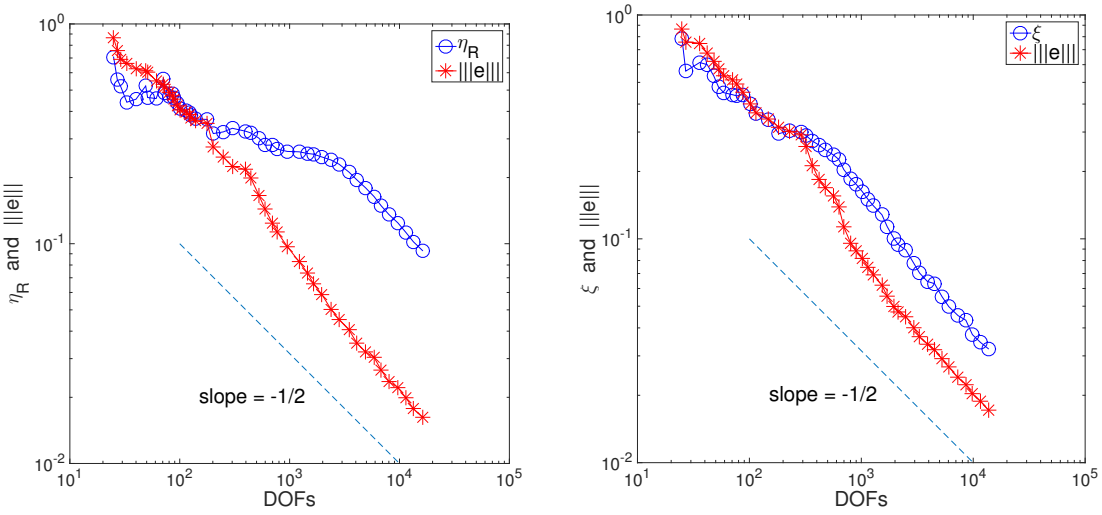

Figure 5.12. Example 3 - error and estimator plot for η(left) and ξ(right)

Table 5.3.
Example 3 - results for η and ξ

estimator	DOFs	$\|\|\|e\|\|\|/\|\|\|u\|\|\|$	eff-ind
η	16217	9.2E-3	5.74
ξ	13664	9.8E-3	1.88

6. EQUILIBRATED ESTIMATOR FOR CONFORMING FINITE ELEMENT APPROXIMATIONS

A posteriori error estimators based on equilibrated flux recovery have been popular recently, since they yield guaranteed upper bounds of the true error (i.e., the reliability constant exactly equals 1) as a result of the Prager-Synge Theorem (cf. [29] or Corollary 2.3.2). Estimators of this type are perfect for discretization error control on both fine and coarse meshes, and error control on pre-asymptotic meshes is important in practice but very difficult. The intial work in this direction by Ladeveze and Leguillon [15], used a partition of unity to reduce the construction of an equilibrated flux to vertex patch based local computations. Hence this approach is computationally more efficient compared to global flux recovery procedures.

Driven by the advantage of such a local procedure, a vast number of approaches on equilibrated flux recovery for diffusion problems have been proposed (cf. [16,17], etc.). In particular, for the conforming linear finite element approximation to the Poisson equation in two dimensions, an equilibrated flux in the lowest order Raviart-Thomas space was explicitly constructed by Braess and Schöberl [18]. Their procedure starts with a decomposition of the error flux into local error fluxes via a partition of unity. The local error fluxes are then approximated by vertex patch problems, and finally each vertex patch problem is solved explicitly by computing the normal components of the recovered local error flux on each edge through circling elements around the vertex at the center.

Extensions of this simple procedure to three dimensions and to higher order elements are non-trivial. Nevertheless, there are many efforts in this regards recently. An attempt was made in [10] on extension to higher order elements, but the resulting admissible flux (constructed in [10, p.157]) is actually not equilibrated in general.

The newest result on extension to three dimensions for the linear elements is reported in an unpublished manuscript by Ern and Vohralik [59] using techniques from the polytopes [60]. Their method is based on a specific enumeration of all faces in a vertex patch. However, it is not easy to obtain such an enumeration in practice even though it exists theoretically. One byproduct of this chapter is a simple approach for the efficient computation of an admissible equilibrated flux for arbitrary order finite elements in both two and three dimensions.

For singularly-perturbed diffusion-reaction and elliptic interface problems, the equilibrated error estimator in [18] is not robust with respect to parameters of the underlying problem (see [10,61]). To guarantee robustness for elliptic interface problems, Cai and Zhang [10] followed the ideas in [18,61] and introduced an additional minimization on each vertex patch to select the recovered equilibrated flux. The approach is applicable for arbitrary order finite element approximations and three dimensional problems. The resulting equilibrated estimator in [10] was shown to be robust under the quasi-monotonicity condition [9] of the diffusion coefficient.

In this chapter, we provide a general algorithm for computing the equilibrated flux via solving patch-based minimizations and prove that the resulting estimator is robust with respect to coefficient jump under the quasi-monotonicity condition of the diffusion coefficient. The new algorithm corrects a mistake in [10], where the recovered flux computed from the algorithm in [10, p.156] is equilibrated for lowest order elements but is *not* equilibrated for higher order elements. Compared to the procedures in [17, 18], which only work for the lowest order discretizations in two dimensions, the new algorithm is valid for arbitrary order conforming finite elements in both two and three dimensions. Theoretically, the proof of robustness is much simpler than the one in [10] thanks to a simpler mixed formulation of the constrained minimization problem on each vertex patch.

The chapter is organized as follows. In Section 6.1, the model problem is introduced. In Section 6.2, the equilibrated flux recovery is reviewed in detail. Section 6.3 formulates local flux equilibration with a minimization over each vertex patch, and

presents a complete procedure on solving the constrained minimization efficiently. The resulting error estimator is defined in Section 6.4. The robust local efficiency bound is proved in Section 6.5. Numerical experiments are presented in Section 6.6 to illustrate the performance of the proposed estimator for both P_1 and P_2 elements.

Remark 6.0.1. *Since the recovered flux in this chapter is required to satisfy the equilibrium condition and to yield an estimator with provable robustness, the flux reconstruction for conforming elements is more complicated than those without the equilibrium constraint, e.g., ZZ-type estimators, hybrid estimators, etc. For interface problems, the development of* robust *equilibrated estimators is not mature. To the best of our knowledge, the one in [10] is the only equilibrated estimator with proved robustness. Moreover, in existing literatures, there does not exist a robust equilibrated estimator that is* explicit*. For the two dimensional case, Chapter 7 presents for the first time an* explicit *equilibrated estimator for conforming finite element approximations with proved robustness with respect to the coefficient jump. For nonconforming elements and discontinuous Galerkin method, however, the flux equilibration is much easier. In Chapter 8 and Chapter 9, we present simple* explicit *equilibrated estimators valid for arbitrary order discretizations.*

6.1 Problem And Conforming Finite Element Approximation

Let Ω be a bounded polygonal domain in $\mathbb{R}^d$ $(d = 2, 3)$ with Lipschitz boundary $\partial\Omega$, where $\partial\Omega = \bar{\Gamma}_D \cup \bar{\Gamma}_N$ and $\Gamma_D \cap \Gamma_N = \emptyset$. For simplicity, assume that $\text{meas}(\Gamma_D) > 0$. Consider the following diffusion problem:

$$\begin{cases} -\text{div}(A\nabla u) = f, & \text{in } \Omega, \\ u = 0, & \text{on } \Gamma_D, \\ (-A\nabla u)\cdot \boldsymbol{n} = g_N, & \text{on } \Gamma_N, \end{cases} \tag{6.1.1}$$

where for almost all $x \in \Omega$, $A(x)$ is a symmetric, positive definite matrix whose smallest eigenvalue is no less than a positive constant independent of x, $f \in L^2(\Omega)$, $g_N \in L^2(\Gamma_N)$, and $\boldsymbol{n}$ is the unit outward vector normal to Γ_N.

The weak formulation for the problem in (7) is to find $u \in H^1_D(\Omega)$ such that

$$a(u,v) := \int_\Omega A\nabla u \cdot \nabla v dx = \int_\Omega f v dx - \int_{\Gamma_N} g_{_N} v ds, \quad \forall\, v \in H^1_D(\Omega).$$

Let the mesh information (elements,vertices,edges/faces,etc.) be given in Section 3.1. Assume that for each $K \in \mathcal{T}$, $A|_K$ is a symmetric, positive definite, constant matrix. For $S \in \mathcal{T} \cup \mathcal{E}$, let Π^k_S denote the L^2 projection from $L^2(S)$ to $P_k(S)$. Assume that $f \in L^2(\Omega)$ and $g_{_N} \in L^2(\Gamma_N)$ are piecewise polynomials such that

$$f|_K \in P_{k-1}(K), \quad \forall\, K \in \mathcal{T} \quad \text{and} \quad g_{_N}|_e \in P_{k-1}(e), \quad \forall\, e \in \mathcal{E}_N.$$

We define the continuous finite element space of order $k\,(k \geq 1)$ by

$$V^k_{\mathcal{T}} = \{v \in H^1_D(\Omega) : v|_K \in P_k(K), \quad \forall\, K \in \mathcal{T}\}.$$

The finite element solution $u_{_\mathcal{T}} \in V^k_{\mathcal{T}}$ satisfies

$$a(u_{_\mathcal{T}}, v) = \int_\Omega f v dx - \int_{\Gamma_N} g_{_N} v ds, \quad \forall\, v \in V^k_{\mathcal{T}}. \tag{6.1.2}$$

6.2 Equilibrated Flux Recovery

The idea of an equilibrated flux recovery is to construct a flux with certain properties that are satisfied by the true flux, $\hat{\boldsymbol{\sigma}} = -A\nabla u$. Specifically, we notice that the true flux enjoys the following properties:

$$\hat{\boldsymbol{\sigma}} \in H(\mathrm{div};\Omega), \quad \mathrm{div}\,\hat{\boldsymbol{\sigma}} = f \text{ in } \Omega, \quad \text{and } \ \hat{\boldsymbol{\sigma}} \cdot \boldsymbol{n} = g_{_N} \text{ on } \Gamma_N, \tag{6.2.1}$$

where $H(\mathrm{div};\Omega)$ is the space of all square-integrable vector fields whose divergence is also square-integrable over Ω. Moreover, the Prager-Synge identity (cf. [17, 29]) implies that

$$\|A^{1/2}\nabla(u_{_\mathcal{T}} - u)\| \leq \|A^{1/2}\nabla u_{_\mathcal{T}} + A^{-1/2}\boldsymbol{\tau}\|, \tag{6.2.2}$$

as long as $\boldsymbol{\tau}$ satisfies the conditions in (6.2.1). Thus we aim to find a flux that fulfills (6.2.1) and that minimizes the right-hand side of (6.2.2). However, it would be expensive to solve a global constraint minimization problem, so a local procedure is usually preferred (see, e.g., [10, 18]).

6.2.1 Localization

Let ϕ_z be the Lagrange nodal basis function associated with vertex z. That is, ϕ_z is the continuous piecewise linear function over mesh $\mathcal{T}$ with $\phi_z(z) = 1$ and $\phi_z(z') = 0$, $\forall\, z' \in \mathcal{N}\backslash\{z\}$. Consider a partition of unity via ϕ_z:

$$\sum_{z\in\mathcal{N}} \phi_z(x) = 1, \quad \forall\, x \in \overline{\Omega}. \tag{6.2.3}$$

The vertex patch associated to a vertex z is given by

$$\omega_z := \text{supp } \phi_z.$$

The partition of unity in (6.2.3) produces a decomposition of $\hat{\boldsymbol{\sigma}}$:

$$\hat{\boldsymbol{\sigma}} = \sum_{z\in\mathcal{N}} \phi_z \hat{\boldsymbol{\sigma}} = \sum_{z\in\mathcal{N}} \boldsymbol{\sigma}_z,$$

where $\boldsymbol{\sigma}_z$ vanishes outside ω_z. Thus we try to construct an approximation of $\boldsymbol{\sigma}_z$ such that its sum over all vertices satisfies (6.2.1).

6.2.2 Conditions at continuous level

To derive conditions on approximations to $\boldsymbol{\sigma}_z$, let us look at necessary conditions for the true flux. Note that $\boldsymbol{\sigma}_z = -\phi_z A\nabla u$ and that

$$\text{div}\, \boldsymbol{\sigma}_z = \nabla\phi_z \cdot \hat{\boldsymbol{\sigma}} + \phi_z f.$$

However, the true $\hat{\boldsymbol{\sigma}}$ is unknown, so we replace it by its approximation, i.e., the numerical flux: $\boldsymbol{\sigma}_{\mathcal{T}} = -A\nabla u_{\mathcal{T}}$, to obtain the first condition

$$\text{div}\, \hat{\boldsymbol{\sigma}}_z = \nabla\phi_z \cdot \boldsymbol{\sigma}_{\mathcal{T}} + \phi_z f. \tag{6.2.4}$$

The boundary conditions for $\hat{\boldsymbol{\sigma}}_z \in H(\text{div}; \omega_z)$ can be immediately seen as below:

$$\begin{cases} \hat{\boldsymbol{\sigma}}_z|_e \cdot \boldsymbol{n}_e = \phi_z g_{_N}, & \text{if } e \in \mathcal{E}_N, \\ \hat{\boldsymbol{\sigma}}_z|_e \cdot \boldsymbol{n}_e = 0, & \text{if } e \in \mathcal{E}_z. \end{cases} \tag{6.2.5}$$

Here $\mathcal{E}_z$ is a subset of $\mathcal{E}$ associated with ω_z where ϕ_z vanishes:

$$\mathcal{E}_z := \{e \in \mathcal{E} : e \subset \partial\omega_z \text{ if } z \notin \partial\Omega \text{ or } e \subset \partial\omega_z \backslash \partial\Omega \text{ if } z \in \partial\Omega\}. \tag{6.2.6}$$

Moreover, $\hat{\boldsymbol{\sigma}}_z$ vanishes outside ω_z.

Define

$$\hat{\boldsymbol{\sigma}} = \sum_{z\in\mathcal{N}} \hat{\boldsymbol{\sigma}}_z. \tag{6.2.7}$$

Then we we have the following proposition.

Proposition 6.2.1. *For each vertex z, assume that $\hat{\boldsymbol{\sigma}}_z$ satisfies the conditions in* (6.2.4) *and* (6.2.5). *Then the flux $\hat{\boldsymbol{\sigma}}$ defined in* (6.2.7) *satisfies* (6.2.1).

Proof. According to (6.2.4) and (6.2.3), we have

$$\operatorname{div} \hat{\boldsymbol{\sigma}} = \sum_{z\in\mathcal{N}} \operatorname{div} \hat{\boldsymbol{\sigma}}_z = \sum_{z\in\mathcal{N}} (\nabla\phi_z \cdot \boldsymbol{\sigma}_{\mathcal{T}} + \phi_z f) = f.$$

From (6.2.5), it follows immediately that $\hat{\boldsymbol{\sigma}} \in H(\operatorname{div};\Omega)$ and that $\hat{\boldsymbol{\sigma}} \cdot \boldsymbol{n} = g_{_N}$ on Γ_N. Therefore, we conclude that $\hat{\boldsymbol{\sigma}}$ satisfies (6.2.1). □

We next show that the conditions on $\hat{\boldsymbol{\sigma}}_z$ are well-posed. In other words, we prove the existence of $\hat{\boldsymbol{\sigma}}_z$ with conditions (6.2.4) and (6.2.5) at the continuous level.

Theorem 6.2.1. *Suppose that $u_{\mathcal{T}}$ is the finite element solution defined in* (6.1.2). *For each vertex $z \in \mathcal{N}$, there exists a flux $\hat{\boldsymbol{\sigma}}_z \in H(\operatorname{div};\Omega)$ satisfying* (6.2.4) *and* (6.2.5).

Proof. Consider the following Neumann problem:

$$\begin{cases} \operatorname{div}(\nabla v) = \nabla\phi_z \cdot \boldsymbol{\sigma}_{\mathcal{T}} + \phi_z f, & \text{in } \omega_z, \\ \nabla v \cdot \boldsymbol{n} = 0, & \text{on } \partial\omega_z \backslash \partial\Omega, \\ \nabla v \cdot \boldsymbol{n} = \phi_z g_{_N}, & \text{on } \partial\omega_z \cap \Gamma_N, \\ \nabla v \cdot \boldsymbol{n} = C, & \text{on } \partial\omega_z \cap \Gamma_D, \end{cases} \tag{6.2.8}$$

where C is a constant defined by

$$C := \operatorname{meas}(\partial\omega_z \cap \Gamma_D)^{-1} \left(\int_{\omega_z} \nabla\phi_z \cdot \boldsymbol{\sigma}_{\mathcal{T}} + \phi_z f dx - \int_{\partial\omega_z\cap\Gamma_N} \phi_z g_{_N} ds \right)$$

if $z \in \partial\Omega$ and $\text{meas}(\partial\omega_z \cap \Gamma_D) > 0$; otherwise, $C = 0$.

We show that the compatibility condition below for the Neumann problem in (6.2.8) always holds true:

$$\int_{\omega_z} \nabla\phi_z \cdot \boldsymbol{\sigma}_{\mathcal{T}} + \phi_z f dx = \int_{\partial\omega_z \cap \Gamma_D} C ds + \int_{\partial\omega_z \cap \Gamma_N} \phi_z g_{_N} ds. \tag{6.2.9}$$

According to the definition of C, it is obvious that (6.2.9) is true when $z \in \partial\Omega$ and $\text{meas}(\partial\omega_z \cap \Gamma_D) > 0$. Otherwise, we have $\phi_z \in V_{\mathcal{T}}^1$, and (6.2.9) follows immediately from the weak formulation for $u_{\mathcal{T}}$ in (6.1.2) by choosing $v = \phi_z \in V_{\mathcal{T}}^1$. Therefore, we conclude that the Neumann problem in (6.2.8) is solvable.

Let $v \in H^1(\omega_z)$ be a solution to (6.2.8). By setting $\nabla v = 0$ outside ω_z, we see that $\hat{\boldsymbol{\sigma}}_z := \nabla v \in H(\text{div}; \Omega)$ satisfies (6.2.4) and (6.2.5). □

6.3 Local Equilibrated Flux Recovery Via Minimization

In this section, we consider conditions (6.2.4) and (6.2.5) at the discrete level. To approximate the flux, we consider the Raviart-Thomas space associated with the triangulation $\mathcal{T}$. Let $\text{RT}_{k-1}(K)$ be the Raviart-Thomas space of index $k-1$ on element K defined in (3.3.1). The corresponding $H(\text{div}; \Omega)$-conforming and the broken Raviart-Thomas spaces are RT_{k-1} and RT_{k-1}^{-1} defined in (3.3.2) and (3.3.3), respectively. Corresponding to the conforming finite element space $V_{\mathcal{T}}^k$, approximation $\hat{\boldsymbol{\sigma}}_{_{z,\mathcal{T}}}$ to $\hat{\boldsymbol{\sigma}}_z$ is required to satisfy

$$\text{div}\, \hat{\boldsymbol{\sigma}}_{_{z,\mathcal{T}}}|_K \in P_{k-1}(K),\ \forall\, K \in \mathcal{T} \text{ and } \hat{\boldsymbol{\sigma}}_{_{z,\mathcal{T}}} \cdot \boldsymbol{n}_e \in P_{k-1}(e),\ \forall\, e \in \mathcal{E}.$$

Therefore, at the discrete level, the conditions are the discrete equilibrium equation:

$$\text{div}\, \hat{\boldsymbol{\sigma}}_{_{z,\mathcal{T}}}|_K = \bar{f}_z|_K := \Pi_K^k \left(\nabla\phi_z \cdot \boldsymbol{\sigma}_{\mathcal{T}} + \phi_z f\right), \quad \forall\, K \in \mathcal{T} \tag{6.3.1}$$

and the boundary conditions

$$\begin{cases} \hat{\boldsymbol{\sigma}}_{_{z,\mathcal{T}}} \cdot \boldsymbol{n}_e = \Pi_e^k(\phi_z g_{_N}), & \text{if } e \in \mathcal{E}_N, \\ \hat{\boldsymbol{\sigma}}_{_{z,\mathcal{T}}} \cdot \boldsymbol{n}_e = 0, & \text{if } e \in \mathcal{E}_z. \end{cases} \tag{6.3.2}$$

By incorporating the constraints in (6.3.2), we define the following subset of RT_{k-1} associated with vertex z:

$$\mathrm{RT}_{z,g} := \{\boldsymbol{\tau} \in \mathrm{RT}_{k-1} : \operatorname{supp} \boldsymbol{\tau} \subseteq \omega_z,\ \boldsymbol{\tau} \text{ satisfies (6.3.2)}\}.$$

Moreover, we define below the set Σ_z as the collection of admissible fluxes satisfying (6.3.1) and (6.3.2):

$$\Sigma_z := \{\boldsymbol{\tau} \in \mathrm{RT}_{z,g} : \operatorname{div} \boldsymbol{\tau} = \bar{f}_z\}.$$

For each vertex $z \in \mathcal{N}$, we look for a flux $\hat{\boldsymbol{\sigma}}_{z,\mathcal{T}} \in \Sigma_z$ such that $\hat{\boldsymbol{\sigma}}_{z,\mathcal{T}}$ solves the following constrained minimization problem on ω_z:

$$\|A^{-1/2}(\hat{\boldsymbol{\sigma}}_{z,\mathcal{T}} - \boldsymbol{\sigma}_{z,\mathcal{T}})\|_{\omega_z} = \min_{\boldsymbol{\tau}\in\Sigma_z} \|A^{-1/2}(\boldsymbol{\tau} - \boldsymbol{\sigma}_{z,\mathcal{T}})\|_{\omega_z}, \tag{6.3.3}$$

where

$$\boldsymbol{\sigma}_{z,\mathcal{T}}|_K = \Pi_{\mathrm{RT}_{k-1}(K)} \left(\phi_z \boldsymbol{\sigma}_{\mathcal{T}}\right), \quad \forall\, K \in \mathcal{T}, \tag{6.3.4}$$

and $\Pi_{\mathrm{RT}_{k-1}(K)}$ denotes the interpolation from $H(\operatorname{div}; K)$ to $\mathrm{RT}_{k-1}(K)$.

Since Σ_z is a closed convex subset of RT_{k-1}, the minimization problem in (6.3.3) is uniquely solvable whenever Σ_z is non-empty.

We prove the existence of an equilibrated local flux $\hat{\boldsymbol{\sigma}}_{z,\mathcal{T}} \in \Sigma_z$ first in Theorem 6.3.1, then we construct one $\hat{\boldsymbol{\sigma}}_{z,\mathcal{T}} \in \Sigma_z$ in Section 6.3.1.

As a discrete version of Theorem 6.2.1, the existence of an equilibrated local flux $\hat{\boldsymbol{\sigma}}_{z,\mathcal{T}} \in \Sigma_z$ can be easily proved.

Theorem 6.3.1. *Suppose that $u_{\mathcal{T}} \in V_{\mathcal{T}}^k\, (k \geq 1)$ is the finite element solution defined in* (6.1.2)*. For each vertex $z \in \mathcal{N}$, there exists a $\hat{\boldsymbol{\sigma}}_{z,\mathcal{T}} \in \mathrm{RT}_{k-1}$ satisfying* (6.3.1) *and* (6.3.2)*. Hence Σ_z is non-empty.*

Proof. According to Theorem 6.2.1, there exists a $\boldsymbol{\tau}_z \in H(\operatorname{div}; \Omega)$ satisfying the conditions in (6.2.4) and (6.2.5) at the continuous level. Define $\hat{\boldsymbol{\sigma}}_{z,\mathcal{T}}$ by setting $\hat{\boldsymbol{\sigma}}_{z,\mathcal{T}}|_K := \Pi_{\mathrm{RT}_{k-1}(K)}\boldsymbol{\tau}_z$ for all $K \in \mathcal{T}$. Then it is easy to see that $\hat{\boldsymbol{\sigma}}_{z,\mathcal{T}} \in \mathrm{RT}_{k-1}$ and $\hat{\boldsymbol{\sigma}}_{z,\mathcal{T}}$ satisfies (6.3.1) and (6.3.2). □

Remark 6.3.1. *It can be easily verified that there is a relation between* $\hat{\boldsymbol{\sigma}}_{z,\mathcal{T}} \in \Sigma_z$ *in* Section 6.3 *and the recovered error flux* $\boldsymbol{\sigma}_z^{\Delta} \in \mathrm{RT}_{k-1}^{-1}$ *in* [10] (*see also* [17,18] *without imposing the minimization in* (6.3.3)) *given by*

$$\boldsymbol{\sigma}_z^{\Delta} = \hat{\boldsymbol{\sigma}}_{z,\mathcal{T}} - \boldsymbol{\sigma}_{z,\mathcal{T}},$$

where $\boldsymbol{\sigma}_{z,\mathcal{T}}$ *is defined in* (6.3.4).

6.3.1 Solution of the constrained minimization problem (6.3.3)

To compute the recovered local flux on each vertex patch, [17] used an explicit procedure in two dimensions to compute the recovered flux, which requires an enumeration of elements in a specific direction around the center vertex. This simple approach is only valid for the RT_0 recovery in the lowest order finite element discretization and the extension to three dimensions is not straightforward due to the enumeration issues. Moreover, without imposing any minimization in the recovery, the computed estimator in [17,18] is not robust with respect to coefficient jump. To ensure the robustness of the estimator, a constrained minimization as in (6.3.3), valid for higher order discretizations, was introduced in [10]. Since the nonhomogeneous constraint in (6.3.3) poses more computational challenges than the homogeneous constraint, following the procedure in [17], an attempt to construct an admissible equilibrated flux was made in [10], but again the computed flux was equilibrated only for the lowest order discretization in two dimensions. Recently, for Poisson equations, Ern and Vohralík [59] extended the idea in [17] to three dimensions based on a specific enumeration of all faces in a vertex patch, but it is not straightforward to obtain such an enumeration in practice even though it exists theoretically. Hence, for interface problems, no simple procedure was presented so far regarding the robust equilibrated flux recovery for higher order discretizations in both two and three dimensions.

In this section, we first present a simple algorithm, which only requires to solve a local linear system, to construct an admissible equilibrated flux $\hat{\boldsymbol{\sigma}}_{z,\mathcal{T}}^{f} \in \Sigma_z$, valid for discretizations of arbitrary order in d $(d = 2, 3)$ dimensions. Then it suffices to solve

the minimization below over the divergence free subspace of $\mathrm{RT}_{z,0}$, where the basis functions are known explicitly (see (6.3.8)):

$$\hat{\boldsymbol{\sigma}}^0_{z,\mathcal{T}} = \underset{\substack{\boldsymbol{\tau}\in \mathrm{RT}_{z,0}\\ \operatorname{div}\boldsymbol{\tau}=0}}{\arg\min} \|A^{-1/2}(\boldsymbol{\tau} + \hat{\boldsymbol{\sigma}}^f_{z,\mathcal{T}} - \boldsymbol{\sigma}_{z,\mathcal{T}})\|_{\omega_z}. \tag{6.3.5}$$

Setting

$$\hat{\boldsymbol{\sigma}}_{z,\mathcal{T}} = \hat{\boldsymbol{\sigma}}^f_{z,\mathcal{T}} + \hat{\boldsymbol{\sigma}}^0_{z,\mathcal{T}}$$

gives the minimizer in (6.3.3).

We define

$$\tau_e = 0 \quad \text{for } e \in \mathcal{E}_z \quad \text{and} \quad \tau_e = \int_e \phi_z g_{_N} ds \quad \text{for } e \in \mathcal{E}_N,$$

and τ_e ($e \in \mathcal{E}_{\omega_z}\backslash(\mathcal{E}_z \cup \mathcal{E}_N)$) to be a solution of the following linear system

$$\sum_{e\in\mathcal{E}_K\backslash\mathcal{E}_z} s_{_K}\tau_e = \int_K \bar{f}_z dx, \quad \forall\, K \in \mathcal{T}_z, \tag{6.3.6}$$

which is indeed solvable according to Proposition 6.3.1. In practice, it can be solved via singular value decomposition (SVD).

Once all τ_e ($e \in \mathcal{E}_{\omega_z}$) are known, an equilibrated local flux $\hat{\boldsymbol{\sigma}}^f_{z,\mathcal{T}} \in \Sigma_z$ can be constructed below by assigning degrees of freedom in $\mathrm{RT}_{k-1}(K)$ on each $K \in \mathcal{T}_z$:

$$\begin{cases} \hat{\boldsymbol{\sigma}}^f_{z,\mathcal{T}}\cdot\boldsymbol{n}_e = \tau_e/|e|, & \forall\, e \in \mathcal{E}_K\backslash\mathcal{E}_N,\\ \hat{\boldsymbol{\sigma}}^f_{z,\mathcal{T}}\cdot\boldsymbol{n}_e = \Pi^k_e(\phi_z g_{_N}), & \forall\, e \in \mathcal{E}_K\cap\mathcal{E}_N,\\ \displaystyle\int_K \hat{\boldsymbol{\sigma}}^f_{z,\mathcal{T}}\cdot\nabla p\, dx = \sum_{e\in\mathcal{E}_K}(s_{_K}\hat{\boldsymbol{\sigma}}^f_{z,\mathcal{T}}\cdot\boldsymbol{n}_e, p)_e - (\bar{f}_z, p)_K, & \forall\, p \in P_{k-1}(K),\\ \displaystyle\int_K \hat{\boldsymbol{\sigma}}^f_{z,\mathcal{T}}\cdot\boldsymbol{q}\, dx = 0, & \forall\, \boldsymbol{q} \in Q_{k-2}(K), \end{cases} \tag{6.3.7}$$

where

$$Q_{k-2}(K) := \left\{\boldsymbol{q} \in P_{k-2}(K)^d : (\boldsymbol{q}, \nabla p) = 0,\ \forall\, p \in P_{k-1}(K)\right\}.$$

Now in view of the minimization in (6.3.5) with the homogeneous constraint, the solution of $\hat{\boldsymbol{\sigma}}^0_{z,\mathcal{T}}$ in (6.3.5) has been considered in [10] and we briefly review it here. According to [10], we know that

$$N_z := \{\boldsymbol{\tau} \in \mathrm{RT}_{z,0} : \operatorname{div} \boldsymbol{\tau} = 0\} = \begin{cases} \nabla^\perp S^k_{z,0}, & \text{if } d = 2, \\ \nabla \times N d^k_{z,0}, & \text{if } d = 3, \end{cases} \tag{6.3.8}$$

where

$$\nabla^\perp v = (-\frac{\partial v}{\partial y}, \frac{\partial v}{\partial x}),$$

$$S^k_{z,0} := \left\{ v \in H^1(\omega_z) : v|_K \in P_k(K)\ \forall K \in \mathcal{T}_z \text{ and } v|_e = 0 \text{ on } e \in \mathcal{E}_z \cup \mathcal{E}_N \right\},$$

$$N d^k_{z,0} := \left\{ \boldsymbol{\tau} \in H(\mathrm{curl}; \omega_z) : \boldsymbol{\tau}|_K \in N d^k(K)\ \forall K \in \mathcal{T}_z \text{ and } \boldsymbol{\tau} \times \boldsymbol{n}_e = 0 \text{ on } e \in \mathcal{E}_z \cup \mathcal{E}_N \right\},$$

$$N d^k(K) := \left\{ \boldsymbol{\tau} \in L^2(K)^d : \boldsymbol{\tau} = \boldsymbol{a} + \boldsymbol{b}, \boldsymbol{a} \in P_k(K)^d, \boldsymbol{b} \in P^h_{k+1}(K)^d \text{ and } \boldsymbol{b} \cdot \boldsymbol{x} = 0 \right\}$$

with $P^h_k(K)^d$ the space of homogeneous polynomials of order k on element K. The minimizer $\hat{\boldsymbol{\sigma}}^0_{z,\mathcal{T}} \in N_z$ of (6.3.5) is then characterized by

$$\left(A^{-1} \hat{\boldsymbol{\sigma}}^0_{z,\mathcal{T}}, \boldsymbol{\tau} \right)_{\omega_z} = \left(A^{-1} (\boldsymbol{\sigma}_{z,\mathcal{T}} - \hat{\boldsymbol{\sigma}}^f_{z,\mathcal{T}}), \boldsymbol{\tau} \right)_{\omega_z}, \quad \forall\, \boldsymbol{\tau} \in N_z. \tag{6.3.9}$$

The algorithm for computing the desired flux $\hat{\boldsymbol{\sigma}}_{z,\mathcal{T}}$ is summarized below.

1. Solve the linear system of τ_e in (6.3.6) for $e \in \mathcal{E}_{\omega_z} \backslash (\mathcal{E}_z \cup \mathcal{E}_N)$ using SVD.
2. Define $\hat{\boldsymbol{\sigma}}^f_{z,\mathcal{T}}$ as in (6.3.7).
3. Solve the linear system associated with (6.3.9) and then obtain $\hat{\boldsymbol{\sigma}}^0_{z,\mathcal{T}}$.
4. Set $\hat{\boldsymbol{\sigma}}_{z,\mathcal{T}} = \hat{\boldsymbol{\sigma}}^f_{z,\mathcal{T}} + \hat{\boldsymbol{\sigma}}^0_{z,\mathcal{T}}$.

Proposition 6.3.1. *The linear system of τ_e for $e \in \mathcal{E}_{\omega_z} \backslash (\mathcal{E}_z \cup \mathcal{E}_N)$ in (6.3.6) is solvable.*

Proof. According to Theorem 6.3.1, there exists a $\hat{\boldsymbol{\sigma}}^*_{z,\mathcal{T}} \in \Sigma_z$. By applying the divergence theorem to $\operatorname{div} \hat{\boldsymbol{\sigma}}^*_{z,\mathcal{T}} = \bar{f}_z$ in each element $K \in \mathcal{T}_z$, we obtain the linear system in (6.3.6) with

$$\tau_e = \int_e \hat{\boldsymbol{\sigma}}^*_{z,\mathcal{T}} \cdot \boldsymbol{n}_e ds.$$

This implies that (6.3.6) is solvable. □

Remark 6.3.2. *As long as the first three equations in* (6.3.7) *are satisfied,* $\hat{\boldsymbol{\sigma}}^f_{z,\mathcal{T}} \in \Sigma_z$. *Hence the last equation in* (6.3.7) *can actually be arbitrary.*

6.4 Equilibrated Error Estimator

Consider the global error estimator

$$\xi := \|A^{-1/2}(\hat{\boldsymbol{\sigma}}_{\mathcal{T}} - \boldsymbol{\sigma}_{\mathcal{T}})\|, \tag{6.4.1}$$

where $\hat{\boldsymbol{\sigma}}_{\mathcal{T}} = \sum_{z\in\mathcal{N}} \hat{\boldsymbol{\sigma}}_{z,\mathcal{T}}$ is the recovered global flux. According to Proposition 6.2.1 and inequality (6.2.2), the estimator above has guaranteed reliability, i.e.,

$$\|A^{1/2}\nabla(u_{\mathcal{T}} - u)\| \le \|A^{-1/2}(\hat{\boldsymbol{\sigma}}_{\mathcal{T}} - \boldsymbol{\sigma}_{\mathcal{T}})\|.$$

The local error indicator is given by

$$\xi_K := \|A^{-1/2}(\hat{\boldsymbol{\sigma}}_{\mathcal{T}} - \boldsymbol{\sigma}_{\mathcal{T}})\|_K.$$

We prove the local efficiency of ξ_K by proving it first for the following indicator

$$\xi_z := \|A^{-1/2}(\hat{\boldsymbol{\sigma}}_{z,\mathcal{T}} - \boldsymbol{\sigma}_{z,\mathcal{T}})\|,$$

and then using the triangle inequality. Note that ξ_z measures the error inside the vertex patch ω_z and is needed in the proof only, while ξ_K is used in the adaptive mesh refinement.

Remark 6.4.1. *It is tempting for one to choose the local error indicator as*

$$\|A^{-1/2}(\hat{\boldsymbol{\sigma}}_{z,\mathcal{T}} - \phi_z\boldsymbol{\sigma}_{\mathcal{T}})\|_{\omega_z},$$

where $u_{\mathcal{T}}$ *is the finite element solution in* $V^k_{\mathcal{T}}\,(k \ge 1)$ *and* $\hat{\boldsymbol{\sigma}}_{z,\mathcal{T}}$ *is the recovered flux in* RT_{k-1}. *However, this is not correct as it can not guarantee the local efficiency in general. For example, suppose that A is the identity matrix and the linear conforming finite element solution coincides with the exact solution such that* $\nabla u_{\mathcal{T}} = \nabla u \neq 0$ *in* $K \subset \omega_z$, *then the true error is* 0, *but* $\hat{\boldsymbol{\sigma}}_{z,\mathcal{T}} + \phi_z\nabla u_{\mathcal{T}} \neq 0$ *because* $\hat{\boldsymbol{\sigma}}_{z,\mathcal{T}}|_K \in \mathrm{RT}_0(K)$ *while* $\phi_z\nabla u_{\mathcal{T}}|_K \notin \mathrm{RT}_0(K)$.

Remark 6.4.2. *In existing literature, corresponding to the P_k element, both* RT_{k-1} *and* RT_k *spaces are considered in the recovery of* $\hat{\boldsymbol{\sigma}}_{z,\mathcal{T}}$. *One of the earlier work in this direction, i.e.,* [18] (see also [17]) *performed flux recovery in* RT_0 *space for* P_1 *element, which is a natural choice as used in most literature on recovery-based estimators* (cf. [11,16,34,43,62]). *A generalization in* [10] *handles* P_k *element with* RT_{k-1} *flux recovery. In those cases, as pointed out in* Remark 6.4.1, *the mapping of* $\phi_z \boldsymbol{\sigma}_{\mathcal{T}}|_K$ *to* $\mathrm{RT}_{k-1}(K)$ *(see* (6.3.4)*) is necessary in order to guarantee the local efficiency of the resulting error indicator. On the other hand, the work in* [19] *chose to use* RT_k *flux recovery for the* P_k *element. In that case, the mapping in* (6.3.4) *is not necessary, but due to the increased degrees of freedom, it would be computationally more expensive.*

6.5 Robust Local Efficiency

We prove robust local efficiency (stated in Theorem 6.5.1) for ξ_z as well as ξ_K. For simplicity, assume that $A = \alpha I$, where $\alpha > 0$ is piecewise constant with respect to $\mathcal{T}$, i.e., $\alpha|_K = \alpha_K > 0$, $\forall\, K \in \mathcal{T}$. Define

$$\alpha_{\max} := \max_{K\in\mathcal{T}} \alpha_K \quad \text{and} \quad \alpha_{\min} := \min_{K\in\mathcal{T}} \alpha_K.$$

Furthermore, for each $e \in \mathcal{E}$, define

$$\alpha_{e,M} := \max_{K\subseteq\omega_e} \alpha_K \quad \text{and} \quad \alpha_{e,m} := \min_{K\subseteq\omega_e} \alpha_K, \tag{6.5.1}$$

where ω_e denotes the union of elements adjacent to e.

For each vertex z, let $\mathcal{T}_z$ be a subset of $\mathcal{T}$ such that $\mathcal{T}_z$ is the collection of all elements contained in the vertex patch ω_z. Let K_z be an element in $\mathcal{T}_z$ such that

$$\alpha_{K_z} = \max_{K\in\mathcal{T}_z} \alpha_K. \tag{6.5.2}$$

We define

$$P_z := \left\{ v \in L^2(\omega_z) : v|_K \in P_{k-1}(K),\ \forall\, K \in \mathcal{T}_z \right\}$$

and

$$\overline{P}_z := \begin{cases} \{v \in P_z : \int_{K_z} v\, dx = 0\}, & \text{if } z \in \mathcal{N}\backslash\Gamma_D, \\ \{v \in P_z : v|_{\Gamma_D} = 0\}, & \text{if } z \in \mathcal{N} \cap \Gamma_D. \end{cases} \tag{6.5.3}$$

We justify the local efficiency of ξ_z via proving the stability of mixed formulations corresponding to the constrained minimization problem in (6.3.3). The mixed formulation will be based on the Hilbert spaces: $\mathrm{RT}_{z,0}$ and $\overline{P}_z$, where $\mathrm{RT}_{z,0}$ is equipped with the $H(\mathrm{div};\Omega)$ inner product and $\overline{P}_z$ is equipped with the usual L^2 inner product.

6.5.1 Mixed formulation for the constrained minimization

By choosing a $\hat{\boldsymbol{\sigma}}_{z,g} \in \mathrm{RT}_{z,g}$, any $\boldsymbol{\tau} \in \mathrm{RT}_{z,g}$ can be written as

$$\boldsymbol{\tau} = \hat{\boldsymbol{\sigma}}_{z,g} + \boldsymbol{\tau}_0, \quad \boldsymbol{\tau}_0 \in \mathrm{RT}_{z,0}.$$

Hence the minimization problem in (6.3.3) is equivalent to: find $\hat{\boldsymbol{\sigma}}_{z,0} \in \mathrm{RT}_{z,0}$ such that

$$\|A^{-1/2}(\hat{\boldsymbol{\sigma}}_{z,0} + \hat{\boldsymbol{\sigma}}_{z,g} - \boldsymbol{\sigma}_{z,\mathcal{T}})\| = \min_{\substack{\boldsymbol{\tau}_0 \in \mathrm{RT}_{z,0} \\ \mathrm{div}\,\boldsymbol{\tau}_0 = \bar{f}_z - \mathrm{div}\,\hat{\boldsymbol{\sigma}}_{z,g}}} \|A^{-1/2}(\boldsymbol{\tau}_0 + \hat{\boldsymbol{\sigma}}_{z,g} - \boldsymbol{\sigma}_{z,\mathcal{T}})\|. \tag{6.5.4}$$

The corresponding mixed formulation is to find $(\hat{\boldsymbol{\sigma}}_{z,0}, u_z) \in \mathrm{RT}_{z,0} \times \overline{P}_z$ such that

$$\begin{cases} \left(A^{-1}\hat{\boldsymbol{\sigma}}_{z,0}, \boldsymbol{\tau}\right) + (\mathrm{div}\,\boldsymbol{\tau}, u_z) = \left(A^{-1}(\boldsymbol{\sigma}_{z,\mathcal{T}} - \hat{\boldsymbol{\sigma}}_{z,g}), \boldsymbol{\tau}\right), & \forall\, \boldsymbol{\tau} \in \mathrm{RT}_{z,0}, \\ (\mathrm{div}\,\hat{\boldsymbol{\sigma}}_{z,0}, v) = (\bar{f}_z - \mathrm{div}\,\hat{\boldsymbol{\sigma}}_{z,g}, v), & \forall\, v \in \overline{P}_z. \end{cases} \tag{6.5.5}$$

$\hat{\boldsymbol{\sigma}}_{z,\mathcal{T}} = \hat{\boldsymbol{\sigma}}_{z,0} + \hat{\boldsymbol{\sigma}}_{z,g} \in \mathrm{RT}_{z,g}$ is then the desired flux that solves (6.3.3).

6.5.2 *Inf-sup* conditions

Given a vertex $z \in \mathcal{N}$, we define a mesh-dependent norm on $\overline{P}_z$:

$$\|v\|_{\mathcal{T}_z}^2 := \sum_{K \in \mathcal{T}_z} \|\alpha^{1/2}\nabla v\|_K^2 + \sum_{e \in \mathcal{E}_{\omega_z}\backslash\mathcal{E}_{\partial\omega_z}} \alpha_{e,m} h_e^{-1} \|[\![v]\!]\|_e^2, \quad \forall\, v \in \overline{P}_z, \tag{6.5.6}$$

where $[\![\phi]\!]_e := \phi_+ - \phi_-$ with ϕ_+ and ϕ_- the limits of ϕ on e from K_e^+ and K_e^-, respectively. It is easy to verify that $\|\cdot\|_{\mathcal{T}_z}$ is a well-defined norm on $\overline{P}_z$.

Here and thereafter, we will use C with or without subscripts to denote a generic positive constant, possibly different at different occurrences, that is independent of $\alpha_{\max}/\alpha_{\min}$, but may depend on the shape parameter of the mesh $\mathcal{T}$ and the polynomial degree k. The result below is proved in a similar fashion as [63, Lemma 2.3].

Lemma 6.5.1. *There exists a positive constant C, depending only on polynomial degree k and the shape parameter of $\mathcal{T}$, such that the following inf-sup condition holds:*

$$\|v\|_{\mathcal{T}_z} \leq C \sup_{\boldsymbol{\tau}\in \mathrm{RT}_{z,0}} \frac{(\operatorname{div}\boldsymbol{\tau}, v)}{\|\alpha^{-1/2}\boldsymbol{\tau}\|}, \quad \forall\, v \in \overline{P}_z. \tag{6.5.7}$$

Proof. For any given $v \in \overline{P}_z$, to establish the *inf-sup* condition in (6.5.7), it suffices to construct a $\boldsymbol{\tau} \in \mathrm{RT}_{z,0}$ such that

$$\|\alpha^{-1/2}\boldsymbol{\tau}\| \leq C\|v\|_{\mathcal{T}_z} \quad \text{and} \quad (\operatorname{div}\boldsymbol{\tau}, v) = \|v\|^2_{\mathcal{T}_z}. \tag{6.5.8}$$

To this end, according to the degrees of freedom for $\mathrm{RT}_{k-1}(K)$, there is a unique vector field $\boldsymbol{\tau} \in \mathrm{RT}_{z,0}$ such that

$$\begin{cases} \boldsymbol{\tau}\cdot\boldsymbol{n}_e = \alpha_{e,m} h_e^{-1}[\![v]\!]_e, & \text{on } e \in \mathcal{E}_{\omega_z}\backslash\mathcal{E}_{\partial\omega_z}, \\ \boldsymbol{\tau}\cdot\boldsymbol{n}_e = 0, & \text{on } e \in \mathcal{E}_{\partial\omega_z}, \\ (\boldsymbol{\tau}, \boldsymbol{q})_K = (-\alpha\nabla v, \boldsymbol{q})_K, & \forall\, \boldsymbol{q} \in P_{k-2}(K)^d, \quad \forall\, K \in \mathcal{T}_z. \end{cases} \tag{6.5.9}$$

A standard scaling argument implies that (cf. [63])

$$\|\boldsymbol{\tau}\|^2_K \leq C\left(\|\alpha\nabla v\|^2_K + h_K\|\boldsymbol{\tau}\cdot\boldsymbol{n}\|^2_{\partial K}\right), \tag{6.5.10}$$

which, together with (6.5.1), yields that

$$\|\alpha^{-1/2}\boldsymbol{\tau}\|^2_K \leq C\left(\|\alpha^{1/2}\nabla v\|^2_K + \sum_{e\in\mathcal{E}_K\backslash\mathcal{E}_{\partial\omega_z}} \alpha_{e,m} h_e^{-1}\|[\![v]\!]\|^2_e\right).$$

Summing over $K \in \mathcal{T}_z$ implies the inequality in (6.5.8). The equality in (6.5.8) is a direct consequence of integration by parts element-wisely. This proves that the $\boldsymbol{\tau}$ defined in (6.5.9) satisfies (6.5.8) and, hence, the lemma.

□

6.5.3 Proof of local efficiency via mixed formulation

We prove the local efficiency of ξ_z via bounding ξ_z from above by the residual-based local indicator defined below, which is known to have local efficiency (cf. [8, 9, 38]):

$$\eta_K := \left(\frac{h_K^2}{\alpha_{_K}} \|f - f_K\|_K^2 + \frac{1}{2} \sum_{e \in \mathcal{E}_K \cap \mathcal{E}_I} \frac{h_K}{\alpha_{e,M}} \|j_e\|_e^2 + \sum_{e \in \mathcal{E}_K \cap \mathcal{E}_N} \frac{h_K}{\alpha_{e,M}} \|j_e\|_e^2 \right)^{1/2},$$

where

$$f_K := \operatorname{div} \boldsymbol{\sigma}_{\mathcal{T}}|_K \quad \text{and} \quad j_e := \begin{cases} [\![\boldsymbol{\sigma}_{\mathcal{T}} \cdot \boldsymbol{n}_e]\!]_e, & \text{if } e \in \mathcal{E}_I, \\ \boldsymbol{\sigma}_{\mathcal{T}}|_e \cdot \boldsymbol{n}_e - g_{_N}, & \text{if } e \in \mathcal{E}_N, \\ 0, & \text{if } e \in \mathcal{E}_D. \end{cases}$$

It is well-known [8, 38] that the local residual indicator has the following robust efficiency bound:

$$\eta_K \leq C \|\alpha^{1/2} \nabla (u - u_{_{\mathcal{T}}})\|_{\omega_K}, \tag{6.5.11}$$

where ω_K denotes the union of all elements that share an edge $(d = 2)$/face $(d = 3)$ with K.

To bound ξ_z from above by η_K, we need the quasi-monotonicity condition on α [9], which is cited as Definition 7.4.1.

Under the quasi-monotonicity condition, the following result was proved in [10, Corollary 5.10].

Lemma 6.5.2. *Assume that α is quasi-monotone. For each $v \in \overline{P}_z$, there exists a constant $C > 0$ such that*

$$\sum_{K \in \mathcal{T}_z} h_K^{-2} \|\alpha^{1/2} v\|_K^2 \leq C \, \|v\|_{\mathcal{T}_z}^2.$$

To show the local efficiency of ξ_z, the following result is needed.

Proposition 6.5.1. *For each vertex $z \in \mathcal{N}$, there exists a vector field $\boldsymbol{\tau}_z \in \mathrm{RT}_{z,g}$ such that for each element $K \in \mathcal{T}_z$,*

$$\operatorname{div} \boldsymbol{\tau}_z = \operatorname{div} \boldsymbol{\sigma}_{_{z,\mathcal{T}}} + J_z, \quad \textit{in } K \quad \textit{and} \quad \|\alpha^{-1/2} (\boldsymbol{\tau}_z - \boldsymbol{\sigma}_{_{z,\mathcal{T}}})\|_K \leq C \eta_K, \tag{6.5.12}$$

where

$$J_z|_K = \sum_{e\in\mathcal{E}_K} |K|^{-1} \int_e -\nu_{z,K,e}\Pi_e^k(\phi_z j_e)ds \tag{6.5.13}$$

and

$$\nu_{z,K,e} := \begin{cases} \frac{\sqrt{\alpha_K}}{\sqrt{\alpha_{K_e^+}}+\sqrt{\alpha_{K_e^-}}}, & \text{if } e \in \mathcal{E}_{\omega_z}\backslash\mathcal{E}_{\partial\omega_z}, \\ 1, & \text{otherwise.} \end{cases} \tag{6.5.14}$$

The construction of a vector field $\boldsymbol{\tau}_z \in \mathrm{RT}_{z,g}$ satisfying the estimates (6.5.18) and (6.5.12) in Proposition 6.5.1 is similar to the flux recovery in Section 4.2. Namely, we pose a boundary value problem for $\boldsymbol{\tau}_z$ in each element in vertex patch ω_z and choose a solution that fulfills the stability estimate.

Let $|K|$ and $|e|$ denote the area ($d = 2$) / volume ($d = 3$) of an element K and length of an edge e ($d = 2$) / area of a face e ($d = 3$), respectively.

Proof of Proposition 6.5.1. For each vertex $z \in \mathcal{N}$, consider the following boundary value problem for $\boldsymbol{\tau}_z \in \mathrm{RT}_{z,g}$ on each element $K \in \mathcal{T}_z$:

$$\begin{cases} \quad \mathrm{div}\,(\boldsymbol{\tau}_z - \boldsymbol{\sigma}_{z,\mathcal{T}}) = J_z, & \text{in } K, \\ (\boldsymbol{\tau}_z - \boldsymbol{\sigma}_{z,\mathcal{T}}|_K)\cdot \boldsymbol{n}_e = -s_K|_e \nu_{z,K,e}\Pi_e^k(\phi_z j_e), & \text{on } e \in \mathcal{E}_K. \end{cases} \tag{6.5.15}$$

It can be verified that the choice of J_z in (6.5.13) guarantees the solvability of each local problem in (6.5.15) and $\boldsymbol{\tau}_z \in \mathrm{RT}_{z,g}$.

According to [61, Lemma 3.1], there exists a $\boldsymbol{\tau}_z|_K \in \mathrm{RT}_{k-1}(K)$ such that the following stability estimate holds

$$\|\boldsymbol{\tau}_z - \boldsymbol{\sigma}_{z,\mathcal{T}}\|_K \le C\left(h_K\|J_z\|_K + h_K^{1/2}\sum_{e\in\mathcal{E}_K}\nu_{z,K,e}\|\Pi_e^k(\phi_z j_e)\|_e\right). \tag{6.5.16}$$

Due to the facts that

$$\|\Pi_e^k(\phi_z j_e)\|_e \le \|j_e\|_e \quad \text{and} \quad \nu_{z,K,e} \le \frac{\sqrt{\alpha_K}}{\sqrt{\alpha_{e,M}}}, \tag{6.5.17}$$

by the Cauchy-Schwarz inequality, we have

$$h_K\alpha^{-1/2}\|J_z\|_K \le \sum_{e\in\mathcal{E}_K}\frac{h_K^{1/2}}{\sqrt{\alpha_{e,M}}}\|j_e\|_e \le C\,\eta_K. \tag{6.5.18}$$

Now, the inequality in (6.5.12) is a direct consequence of (6.5.16) and (6.5.17). This completes the proof of the proposition. □

Now we are in a position to state the local efficiency of ξ_z as well as ξ_K.

Theorem 6.5.1. *Assume that α is quasi-monotone. The following estimates hold true:*

$$\xi_z \le C_1 \left(\sum_{K \in \mathcal{T}_z} \eta_K^2 \right)^{1/2} \le C_2 \, \|\alpha^{1/2} \nabla (u - u_\tau)\|_{\hat{\omega}_z}, \tag{6.5.19}$$

$$\xi_K \le \sum_{z \in \mathcal{N} \cap \partial K} \xi_z \le \sum_{z \in \mathcal{N} \cap \partial K} C \|\alpha^{1/2} \nabla (u - u_\tau)\|_{\hat{\omega}_z}. \tag{6.5.20}$$

where $\hat{\omega}_z$ denotes the union of elements that share at least one edge ($d = 2$) or one face ($d = 3$) with an element in ω_z.

Proof. Note that (6.5.20) is an immediate result of the triangle inequality and (6.5.19), so it suffices to show (6.5.19).

The second inequality in (6.5.19) is a direct consequence of the local efficiency bound of η_K in (6.5.11). To prove the first inequality in (6.5.19), let $\boldsymbol{\tau}_z \in \mathrm{RT}_{z,g}$ be the vector field in Proposition 6.5.1. The Cauchy-Schwarz inequality and (6.5.12) imply that

$$\begin{aligned}
\xi_z^2 &= \big(\alpha^{-1}(\hat{\boldsymbol{\sigma}}_{z,\tau} - \boldsymbol{\sigma}_{z,\tau}), \hat{\boldsymbol{\sigma}}_{z,\tau} - \boldsymbol{\sigma}_{z,\tau}\big) \\
&= \big(\alpha^{-1}(\hat{\boldsymbol{\sigma}}_{z,\tau} - \boldsymbol{\sigma}_{z,\tau}), \hat{\boldsymbol{\sigma}}_{z,\tau} - \boldsymbol{\tau}_z\big) + \big(\alpha^{-1}(\hat{\boldsymbol{\sigma}}_{z,\tau} - \boldsymbol{\sigma}_{z,\tau}), \boldsymbol{\tau}_z - \boldsymbol{\sigma}_{z,\tau}\big) \\
&\le \big(\alpha^{-1}(\hat{\boldsymbol{\sigma}}_{z,\tau} - \boldsymbol{\sigma}_{z,\tau}), \hat{\boldsymbol{\sigma}}_{z,\tau} - \boldsymbol{\tau}_z\big) + C \|\alpha^{-1/2}(\hat{\boldsymbol{\sigma}}_{z,\tau} - \boldsymbol{\sigma}_{z,\tau})\| \left(\sum_{K \in \mathcal{T}_z} \eta_K^2 \right)^{1/2}.
\end{aligned}$$

Now it suffices to show that

$$b \equiv \big(\alpha^{-1}(\hat{\boldsymbol{\sigma}}_{z,\tau} - \boldsymbol{\sigma}_{z,\tau}), \hat{\boldsymbol{\sigma}}_{z,\tau} - \boldsymbol{\tau}_z\big) \le C \|\alpha^{-1/2}(\hat{\boldsymbol{\sigma}}_{z,\tau} - \boldsymbol{\sigma}_{z,\tau})\| \left(\sum_{K \in \mathcal{T}_z} \eta_K^2 \right)^{1/2}. \tag{6.5.21}$$

To this end, it follows from the first equation in (6.5.5) with $\boldsymbol{\tau} = \hat{\boldsymbol{\sigma}}_{z,\mathcal{T}} - \boldsymbol{\tau}_z \in \mathrm{RT}_{z,0}$, (6.5.12), the Cauchy-Schwarz and the triangle inequalities, (6.5.18), and Lemma 6.5.2 that

$$
\begin{aligned}
b &= \big(-\mathrm{div}\,(\hat{\boldsymbol{\sigma}}_{z,\mathcal{T}} - \boldsymbol{\tau}_z), u_z\big) = \sum_{K\in\mathcal{T}_z} \big(J_z - \mathrm{div}\,(\hat{\boldsymbol{\sigma}}_{z,\mathcal{T}} - \boldsymbol{\sigma}_{z,\mathcal{T}}), u_z\big)_K \\
&= \sum_{K\in\mathcal{T}_z} (J_z - \phi_z(f - f_K), u_z)_K \le \sum_{K\in\mathcal{T}_z} \alpha_K^{-1/2} \left(\|J_z\|_K + \|f - f_K\|_K\right) \|\alpha^{1/2} u_z\|_K \\
&\le C \left(\sum_{K\in\mathcal{T}_z} \eta_K^2\right)^{1/2} \left(\sum_{K\in\mathcal{T}_z} h_K^{-2} \|\alpha^{1/2} u_z\|_K^2\right)^{1/2} \le C \|u_z\|_{\mathcal{T}_z} \left(\sum_{K\in\mathcal{T}_z} \eta_K^2\right)^{1/2}.
\end{aligned}
$$

By Lemma 6.5.1, the first equation in (6.5.5), and the Cauchy-Schwarz inequality, we have

$$
\begin{aligned}
\|u_z\|_{\mathcal{T}_z} &\le C \sup_{\boldsymbol{\tau}\in\mathrm{RT}_{z,0}} \frac{(\mathrm{div}\,\boldsymbol{\tau}, u_z)}{\|\alpha^{-1/2}\boldsymbol{\tau}\|} = C \sup_{\boldsymbol{\tau}\in\mathrm{RT}_{z,0}} \frac{(-\alpha^{-1}(\hat{\boldsymbol{\sigma}}_{z,\mathcal{T}} - \boldsymbol{\sigma}_{z,\mathcal{T}}), \boldsymbol{\tau})}{\|\alpha^{-1/2}\boldsymbol{\tau}\|} \\
&\le C \|\alpha^{-1/2}(\hat{\boldsymbol{\sigma}}_{z,\mathcal{T}} - \boldsymbol{\sigma}_{z,\mathcal{T}})\|.
\end{aligned}
$$

Combining the above two inequalities gives (6.5.21). This proves the first inequality in (6.5.19) and, hence, the theorem. □

6.6 Numerical Experiments

We consider solving the Kellogg's problem. The problem setup and parameters in adaptive mesh refinement are same as in Section 4.5. The initial mesh consists of 4×4 congruent squares, each of which is partitioned into two triangles connecting bottom-left and top-right corners. We use Dörfler's marking strategy [46] with $\theta_D = 0.5$ as in Section 4.5. The newest-vertex bisection [44] is used in the refinement.

Numerical tests were performed with both P_1 and P_2 elements. RT_{k-1} flux recovery is used for P_k element ($k = 1, 2$). The following notation will be used:

- exact error $e := u - u_{\mathcal{T}}$;
- effectivity index: eff-ind;
- degrees of freedom: DOFs;

Table 6.1.
Kellogg's example - P_1 and P_2 elements

u_τ	ϵ_{rel}	DOFs	$\|A^{1/2}\nabla e\|/\|A^{1/2}\nabla u\|$	eff-ind
P_1	0.05	12410	4.9E-2	1.47
P_2	0.01	10237	9.9E-3	1.62

- stopping criterion: $\|A^{1/2}\nabla e\| \leq \epsilon_{\text{rel}}\|A^{1/2}\nabla u\|$ with

$$\epsilon_{\text{rel}} = \begin{cases} 0.05, & \text{for } P_1 \text{ element,} \\ 0.01, & \text{for } P_2 \text{ element.} \end{cases}$$

Numerical results are shown in Table 6.1 and in Figure 6.1 – Figure 6.4. First we notice from the plots in Figure 6.2 and Figure 6.4 that the estimator is always larger than the true error, confirming the guaranteed error control. It can be seen from Figure 6.1 and Figure 6.3 that the mesh refinement is homogeneous with respect to the singularity regardless of different scales of the diffusion coefficient in different quadrants, which implies the robustness of the estimator with respect to the coefficient jump. Optimal convergence rates are observed for both P_1 element in Figure 6.2 and P_2 element in Figure 6.4. Table 6.1 shows that the effectivity index is close to 1, so the estimator is considered accurate.

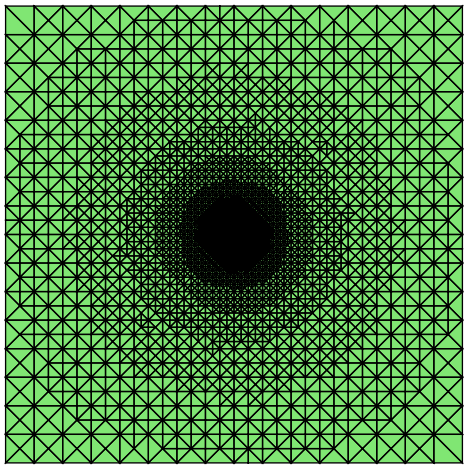

Figure 6.1. P_1 element - mesh

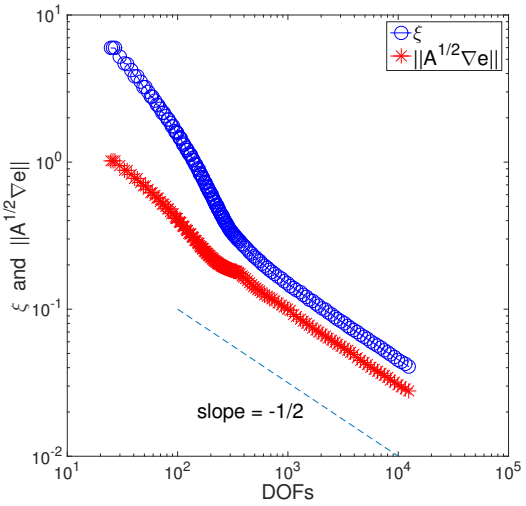

Figure 6.2. P_1 element - error

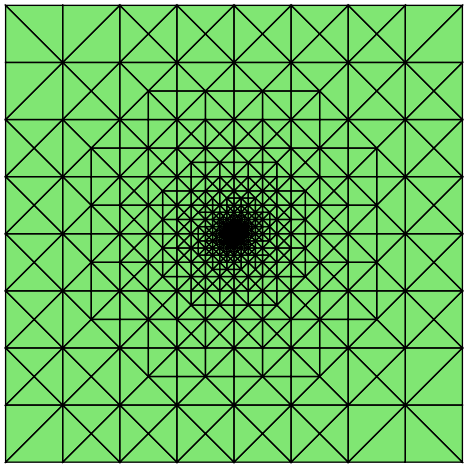

Figure 6.3. P_2 element - mesh

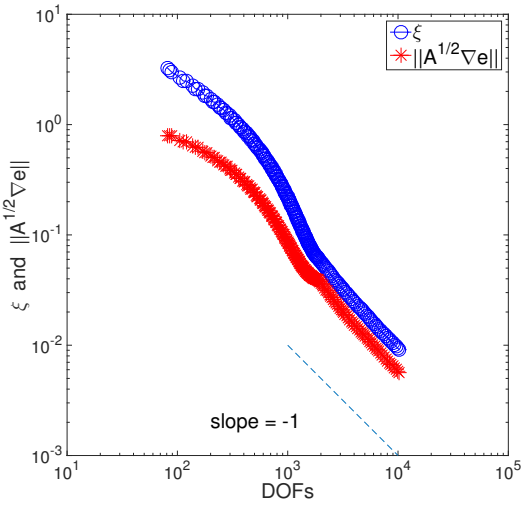

Figure 6.4. P_2 element - error

7. EXPLICIT EQUILIBRATED ESTIMATOR FOR CONFORMING FINITE ELEMENT APPROXIMATIONS IN TWO DIMENSIONS

Consider the same problem and conforming finite element discretization as in Chapter 6. That is,

$$\begin{cases} -\mathrm{div}(\alpha\nabla u) = f, & \text{in } \Omega, \\ u = 0, & \text{on } \Gamma_D, \\ (-\alpha\nabla u)\cdot \boldsymbol{n} = g_{_N}, & \text{on } \Gamma_N, \end{cases}$$

where $\alpha > 0$ is assumed to be piecewise constant with respect to the mesh of discretization and we are interested in the a posteriori error estimation that is independent of the discontinuities of α. The idea of flux equilibration is to find a $H(\mathrm{div};\Omega)$-conforming flux $\hat{\boldsymbol{\sigma}}_{_\mathcal{T}}$ that satisfies $\mathrm{div}\,\hat{\boldsymbol{\sigma}}_{_\mathcal{T}} = f$ in Ω (as well as certain boundary condition). For computational efficiency, the global problem is decomposed into local ones (via a partition of unity using nodal basis functions), where for each vertex z, one constructs a flux $\hat{\boldsymbol{\sigma}}_{_{z,\mathcal{T}}}$ supported on ω_z (vertex patch at z) that satisfies certain local equilibration $\mathrm{div}\,\hat{\boldsymbol{\sigma}}_{_{z,\mathcal{T}}} = \bar{f}_z$ for some appropriate $\bar{f}_z$ with $\sum_z \bar{f}_z = f$. Then the desired equilibrated flux is defined as $\hat{\boldsymbol{\sigma}}_{_\mathcal{T}} = \sum_z \hat{\boldsymbol{\sigma}}_{_{z,\mathcal{T}}}$.

Once $\hat{\boldsymbol{\sigma}}_{_\mathcal{T}}$ is found, the resulting estimator is given by $\|\hat{\boldsymbol{\sigma}}_{_\mathcal{T}} + \alpha\nabla u_{_\mathcal{T}}\|$, where $u_{_\mathcal{T}}$ is the finite element solution. For the conforming finite element approximation to the model problem, the main difficulty (both algorithmically and theoretically) is to construct an equilibrated flux $\hat{\boldsymbol{\sigma}}_{_\mathcal{T}}$ such that the resulting local error indicator is provably robust with respect to the discontinuity of α (under certain assumptions if necessary). To the best of our knowledge, however, there is only one equilibration-based estimator, proposed in [10], that yields a robust local efficiency bound (i.e., lower bound of the true error). The flux equilibration in [10] imposes a constrained

minimization over each vertex patch and the recovered local flux $\hat{\boldsymbol{\sigma}}_{z,\mathcal{T}}$ is defined as the solution to the minimization problem. The proof of robust local efficiency relies on an *inf-sup* inequality for the mixed formulation of the minimization problem, which is rather complicated. An improvement was given in [20] (see also Chapter 6) to simplify both the algorithm and the proof, but it still requires solving constrained minimizations on vertex patches, which can be costly for higher order discretizations (see Section 7.5).

So far no algorithm can be found to explicitly compute an equilibrated flux that yields a robust estimator and it is unknown if a straightforward proof of the robust local efficiency exists. Moreover, the proofs of robust local efficiency in [10,20] require the quasi-monotonicity assumption of α (cf. [9] or Definition 7.4.1) but it is not clear why it is needed since the assumption is not embodied in the flux equilibration.

In this chapter, we present an entirely different approach to address the questions above. The main contributions include:

1. an explicit construction of the local flux $\hat{\boldsymbol{\sigma}}_{z,\mathcal{T}}$ supported on the vertex patch ω_z without solving any minimization problem;

2. a straightforward proof of the robust local efficiency of the error indicator;

3. an interpretation of the quasi-monotonicity assumption of α in the computation of $\hat{\boldsymbol{\sigma}}_{z,\mathcal{T}}$.

This chapter is organized as follows. Section 7.1 reviews the idea of flux equilibration and existing work on robust equilibrated estimators. The proposed explicit equilibrated estimator is defined in Section 7.2 – 7.3. Section 7.4 is dedicated to the proof of robust local efficiency. The numerical experiments are presented in Section 7.5, where the proposed estimator is compared to the minimization-based estimator in [10, 20].

7.1 Review Of Equilibrated Estimators

7.1.1 Idea of flux equilibration

Let the mesh information (elements,vertices,edges/faces,etc.) be given in Section 3.1. Recall that ω_z denotes the support of the P_1-conforming nodal basis function ϕ_z associated with vertex z and $\mathcal{T}_z$ is the set of elements contained in ω_z. Let $\mathcal{E}_z$ be the set of edges as defined (6.2.6) (see Figure 7.1 for an illustration).

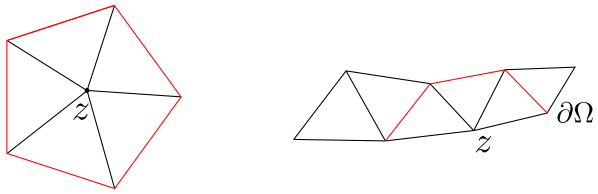

Figure 7.1. Illustrations of $\mathcal{E}_z$(red): $z \in \Omega$(left) and $z \in \partial\Omega$(right)

Below we briefly review the idea of constructing an equilibrated flux in RT_{k-1}. The equilibrated flux is written in the form $\hat{\boldsymbol{\sigma}}_{\mathcal{T}} = \sum_{z\in\mathcal{N}} \hat{\boldsymbol{\sigma}}_{z,\mathcal{T}}$, where for each vertex z, $\hat{\boldsymbol{\sigma}}_{z,\mathcal{T}}$ vanishes outside ω_z. In order for $\hat{\boldsymbol{\sigma}}_{\mathcal{T}}$ to satisfy the equilibrium equation and the boundary condition, the following requirement is imposed for $\hat{\boldsymbol{\sigma}}_{z,\mathcal{T}}$:

$$\begin{cases} \hat{\boldsymbol{\sigma}}_{z,\mathcal{T}} \in H(\mathrm{div};\Omega), & \\ \hat{\boldsymbol{\sigma}}_{z,\mathcal{T}} \cdot \boldsymbol{n}_e = \Pi_e^k(\phi_z g_N), & \text{if } e \in \mathcal{E}_N, \\ \hat{\boldsymbol{\sigma}}_{z,\mathcal{T}} \cdot \boldsymbol{n}_e = 0, & \text{if } e \in \mathcal{E}_z, \\ \mathrm{div}\, \hat{\boldsymbol{\sigma}}_{z,\mathcal{T}} = \bar{f}_z, & \text{in } \Omega, \end{cases} \tag{7.1.1}$$

where

$$\bar{f}_z|_K := \Pi_K^k \left(\nabla\phi_z \cdot (-\alpha \nabla u_{\mathcal{T}}) + \phi_z f \right), \quad \forall\, K \in \mathcal{T}. \tag{7.1.2}$$

Definition 7.1.1. *For a vertex z, we call a vector field $\hat{\boldsymbol{\sigma}}_{z,\mathcal{T}}$ (supported on ω_z)* admissible *if it satisfies* (7.1.1). *The vector field $\hat{\boldsymbol{\sigma}}_{\mathcal{T}} = \sum_{z\in\mathcal{N}} \hat{\boldsymbol{\sigma}}_{z,\mathcal{T}}$ is called* admissible *if each $\hat{\boldsymbol{\sigma}}_{z,\mathcal{T}}$ is admissible.*

For each vertex z, the set of admissible local fluxes in RT_{k-1} is defined as

$$\Sigma_z := \{\boldsymbol{\tau} \in \mathrm{RT}_{k-1} : \operatorname{supp} \boldsymbol{\tau} \subseteq \omega_z,\ \boldsymbol{\tau} \text{ satisfies (7.1.1)}\}. \tag{7.1.3}$$

It is known that Σ_z is not empty (see [20, Theorem 3.2]). Note that if we choose $\hat{\boldsymbol{\sigma}}_{z,\mathcal{T}} \in \Sigma_z$ for each vertex z, then $\hat{\boldsymbol{\sigma}}_{\mathcal{T}} = \sum_{z\in\mathcal{N}} \hat{\boldsymbol{\sigma}}_{z,\mathcal{T}}$ is equilibrated.

It is not hard to construct an equilibrated flux. For example, in two dimensions, to find an admissible $\hat{\boldsymbol{\sigma}}_{z,\mathcal{T}}$, a simple approach was proposed in [17] by circling around the vertex z to fulfill the compatibility condition in each element in ω_z. However, this approach is limited to the lowest order discretization, and more importantly, the resulting a posteriori error estimator is not robust with respect to the discontinuity of α. Hence the challenge lies in the construction of an equilibrated flux that yields a robust error estimator.

7.1.2 Existing work on robust equilibrated estimators

Let $\boldsymbol{\sigma}_{\mathcal{T}} := -\alpha\nabla u_{\mathcal{T}}$ be the numerical flux and $\boldsymbol{\sigma}_{z,\mathcal{T}}$ be the local flux associated with vertex patch ω_z, as defined in (6.3.4).

In order to obtain a robust estimator based on flux equilibration, [10, 20] seek to find a $\hat{\boldsymbol{\sigma}}_{z,\mathcal{T}} \in \Sigma_z$ such that $\hat{\boldsymbol{\sigma}}_{z,\mathcal{T}}$ is the minimizer of the following constrained minimization problem:

$$\|\alpha^{-1/2}\left(\hat{\boldsymbol{\sigma}}_{z,\mathcal{T}} - \boldsymbol{\sigma}_{z,\mathcal{T}}\right)\|_{\omega_z} = \min_{\boldsymbol{\tau}\in\Sigma_z} \|\alpha^{-1/2}\left(\boldsymbol{\tau} - \boldsymbol{\sigma}_{z,\mathcal{T}}\right)\|_{\omega_z}. \tag{7.1.4}$$

The resulting a posteriori error estimator based on $\hat{\boldsymbol{\sigma}}_{\mathcal{T}} = \sum_{z\in\mathcal{N}} \hat{\boldsymbol{\sigma}}_{z,\mathcal{T}}$ was shown to be robust with respect to α (under the quasi-monotonicity assumption).

Since the flux equilibration and the proofs of robustness in [10, 20] rely on the minimization property in (7.1.4), it is natural to ask:

1. Theoretically, is it necessary for $\hat{\boldsymbol{\sigma}}_{z,\mathcal{T}}$ to satisfy (7.1.4) in order to yield a robust error estimator ? If not, can we prove the robustness directly ?

2. Algorithmically, is it possible to construct an admissible $\hat{\boldsymbol{\sigma}}_{z,\mathcal{T}} \in \Sigma_z$ explicitly (without solving any minimization problem) such that the resulting error estimator is robust ?

7.2 The Explicit Flux Equilibration And Error Estimator

In this section, we present an *explicit* approach to compute an equilibrated flux such that the resulting a posteriori estimator is *robust*. The explicit formula for the local flux $\hat{\boldsymbol{\sigma}}_{z,\mathcal{T}}$ is given in (7.2.3).

Since we require $\operatorname{div} \hat{\boldsymbol{\sigma}}_{z,\mathcal{T}} = \bar{f}_z$, to guarantee both compatibility and robustness, a judicious choice of the normal component $\hat{g}_{z,e} := \hat{\boldsymbol{\sigma}}_{z,\mathcal{T}} \cdot \boldsymbol{n}_e$ on each $e \in \mathcal{E}_{\omega_z}$ is made.

As in [12], we first define a weighted average of normal component of $\boldsymbol{\sigma}_{z,\mathcal{T}}$ as below:

$$\tilde{g}_{z,e} := \begin{cases} (1-\lambda_e)\boldsymbol{\sigma}_{z,\mathcal{T}}|_{K_e^+} \cdot \boldsymbol{n}_e + \lambda_e \boldsymbol{\sigma}_{z,\mathcal{T}}|_{K_e^-} \cdot \boldsymbol{n}_e, & \text{if } e \in \mathcal{E}_I, \\ \Pi_e^k(\phi_z g_{_N}), & \text{if } e \in \mathcal{E}_N, \\ \boldsymbol{\sigma}_{z,\mathcal{T}} \cdot \boldsymbol{n}_e, & \text{if } e \in \mathcal{E}_D, \end{cases} \tag{7.2.1}$$

where the weight λ_e is defined in [12, (3.5)]. Note that since $\phi_z = 0$ on $e \in \mathcal{E}_z$, $\tilde{g}_{z,e} = 0$ for $e \in \mathcal{E}_z$.

For the local flux $\hat{\boldsymbol{\sigma}}_{z,\mathcal{T}}$, the normal component $\hat{\boldsymbol{\sigma}}_{z,\mathcal{T}} \cdot \boldsymbol{n}_e$ is chosen to be

$$\hat{g}_{z,e} := \begin{cases} \tilde{g}_{z,e} + J_{z,e}, & \text{if } e \in \mathcal{E}_{\omega_z} \setminus (\mathcal{E}_z \cup \mathcal{E}_N), \\ \tilde{g}_{z,e}, & \text{if } e \in \mathcal{E}_z \cup \mathcal{E}_N, \end{cases} \tag{7.2.2}$$

where $J_{z,e}$ is a constant that will be defined later in Section 7.3.

We define $\hat{\boldsymbol{\sigma}}_{z,\mathcal{T}} \in \Sigma_z$ by imposing all degrees of freedom in $\mathrm{RT}_{k-1}(K)$. For each $K \in \mathcal{T}_z$, $\hat{\boldsymbol{\sigma}}_{z,\mathcal{T}}$ is defined by

$$\begin{cases} \hat{\boldsymbol{\sigma}}_{z,\mathcal{T}} \cdot \boldsymbol{n}_e = \hat{g}_{z,e}, \quad \forall\, e \in \mathcal{E}_K, \\ \displaystyle\int_K \hat{\boldsymbol{\sigma}}_{z,\mathcal{T}} \cdot \nabla p dx = \sum_{e \in \mathcal{E}_K} (s_{_K} \hat{g}_{z,e}, p)_e - (\bar{f}_z, p)_K, \quad \forall\, p \in P_{k-1}(K), \\ \displaystyle\int_K \hat{\boldsymbol{\sigma}}_{z,\mathcal{T}} \cdot \boldsymbol{q} dx = \int_K \boldsymbol{\sigma}_{z,\mathcal{T}} \cdot \boldsymbol{q} dx, \quad \forall\, \boldsymbol{q} \in Q_{k-2}(K), \quad \text{if } k \geq 2, \end{cases} \tag{7.2.3}$$

where $Q_{k-2}(K)$ is defined in (4.2.11). The global flux is defined by

$$\hat{\boldsymbol{\sigma}}_{\mathcal{T}} = \sum_{z\in\mathcal{N}} \hat{\boldsymbol{\sigma}}_{z,\mathcal{T}}. \tag{7.2.4}$$

It is easy to check that $\hat{\boldsymbol{\sigma}}_{z,\mathcal{T}}$ is admissible, i.e., $\hat{\boldsymbol{\sigma}}_{z,\mathcal{T}} \in \Sigma_z$. In general, $\hat{\boldsymbol{\sigma}}_{z,\mathcal{T}}$ may not necessarily satisfy the minimization property in (7.1.4).

The compatibility condition for the first two requirements in (7.2.3), i.e.,

$$\int_K \bar{f}_z dx = \sum_{e\in\mathcal{E}_K} \int_e s_K \hat{g}_{z,e} ds, \tag{7.2.5}$$

is fulfilled due to the choice of $\hat{g}_{z,e}$ (or, equivalently, $J_{z,e}$) in Section 7.3.

The patch-based and element-based error indicators are then defined as

$$\xi_z := \|\alpha^{-1/2}\left(\hat{\boldsymbol{\sigma}}_{z,\mathcal{T}} - \boldsymbol{\sigma}_{z,\mathcal{T}}\right)\|_{\omega_z} \quad \text{and} \quad \xi_K := \|\alpha^{-1/2}\left(\hat{\boldsymbol{\sigma}}_{\mathcal{T}} - \boldsymbol{\sigma}_{\mathcal{T}}\right)\|_K, \tag{7.2.6}$$

respectively. The global error estimator is given by

$$\xi := \|\alpha^{-1/2}\left(\hat{\boldsymbol{\sigma}}_{\mathcal{T}} - \boldsymbol{\sigma}_{\mathcal{T}}\right)\|.$$

Since $\hat{\boldsymbol{\sigma}}_{\mathcal{T}}$ is equilibrated, the guaranteed reliability of ξ follows immediately from the Prager-Synge identity (cf. [29]). The robust local efficiency holds under the quasi-monotonicity assumption of α and the proof is given in Section 7.4.

Theorem 7.2.1. *Let u be the exact weak solution to the model problem in* (7) *and $u_{\mathcal{T}} \in V_{\mathcal{T}}$ be the finite element solution. The recovered flux $\hat{\boldsymbol{\sigma}}_{\mathcal{T}}$ defined in* (7.2.4) *satisfies*

$$\operatorname{div}\hat{\boldsymbol{\sigma}}_{\mathcal{T}} = f \ \text{ in } \Omega \quad \text{and} \quad \hat{\boldsymbol{\sigma}}_{\mathcal{T}}\cdot\boldsymbol{n} = g_{_N} \ \text{ on } \Gamma_N.$$

Consequently,

$$\|\alpha^{-1/2}\nabla(u_{\mathcal{T}} - u)\| \le \|\alpha^{-1/2}\left(\hat{\boldsymbol{\sigma}}_{\mathcal{T}} - \boldsymbol{\sigma}_{\mathcal{T}}\right)\| = \xi.$$

Additionally, if α is quasi-monotone with respect to vertex z, then

$$\xi_z \le C\|\alpha^{1/2}\nabla(u - u_{\mathcal{T}})\|_{\hat{\omega}_z} \quad \text{and} \quad \xi_K \le C\sum_{z\in\mathcal{N}\cap\partial K} \|\alpha^{1/2}\nabla(u - u_{\mathcal{T}})\|_{\hat{\omega}_z}, \tag{7.2.7}$$

where $\hat{\omega}_z$ denotes the the union of elements that share at least one edge with an element in ω_z.

7.3 Choice Of $J_{z,e}$

In this section, we show how to choose the constant $J_{z,e}$ in (7.2.2). The idea is to choose $J_{z,e}$ such that the following two requirements are fulfilled:

1. the compatibility condition in (7.2.5);
2. the robust local efficiency in (7.2.7).

The proof of robustness will be presented in Section 7.4.

7.3.1 Notation in a vertex patch

We first introduce some notations that will facilitate the reasoning later in Section 7.4. Definie $\alpha_{_K} := \alpha|_K$ and

$$\alpha_{\max} := \max_{K\in\mathcal{T}} \alpha_{_K} \quad \text{and} \quad \alpha_{\min} := \min_{K\in\mathcal{T}} \alpha_{_K}.$$

Furthermore, for each $e \in \mathcal{E}$, define

$$\alpha_e := \max_{K\subseteq\omega_e} \alpha_{_K}, \tag{7.3.1}$$

where ω_e denotes the union of elements adjacent to e.

Suppose ω_z is composed of r elements $K_1, \ldots, K_r$ ($r \geq 1$), ordered counterclockwisely with respect to vertex z. The following notation regarding the indexing of elements in ω_z will be used throughout this chapter:

- let K_m be an element in ω_z such that

$$\alpha_{K_m} = \max_{K\subseteq\omega_z} \alpha_{_K}; \tag{7.3.2}$$

- if $z \in \Omega$, K_1 is assumed to be an element in ω_z such that

$$\alpha_{K_1} = \min_{K\subseteq\omega_z} \alpha_{_K}; \tag{7.3.3}$$

- if $z \in \overline{\Gamma}_D \cap \overline{\Gamma}_N$, K_1 is assumed to be on Γ_D, i.e., $\text{meas}(\partial K_1 \cap \Gamma_D) > 0$ as illustrated in Figure 7.5;

- if $z \in \Gamma_D \backslash \bar{\Gamma}_N$ as illustrated in Figure 7.3, let K_s be an element in ω_z such that

$$\alpha_{K_s} = \min_{K \subseteq \omega_z} \alpha_K. \tag{7.3.4}$$

We specify the edges adjacent to z as follows. If $z \in \Omega$, there are r such edges $e_1, \ldots, e_r$, defined by

$$e_i = \partial K_i \cap \partial K_{i+1} (\text{ with } K_{r+1} := K_1), \quad i = 1, \ldots, r.$$

If $z \in \partial\Omega$, there are $r+1$ such edges $e_0, e_1, \ldots, e_r$, given by

$$e_i = \partial K_i \cap \partial K_{i+1} (i = 1, \ldots, r-1), \quad e_0 \cup e_r \subseteq \partial\omega_z \cap \partial\Omega.$$

We associate each edge e_i with a unit vector $\boldsymbol{n}_i$ normal to e_i such that

- $\boldsymbol{n}_i$ points from K_i to K_{i+1} if $e_i \in \mathcal{E}_I$;
- $\boldsymbol{n}_i$ points to the exterior of Ω if $e_i \in \mathcal{E}_N$;
- $\boldsymbol{n}_0$ points to the interior of Ω if $e_0 \in \mathcal{E}_D$;
- $\boldsymbol{n}_r$ points to the exterior of Ω if $e_r \in \mathcal{E}_D$.

Without loss of generality, we assume that if $z \in \bar{\Gamma}_D \cap \bar{\Gamma}_N$, the path in ω_z from Γ_D to Γ_N is counterclockwise as in Figure 7.5. Then there are in total four types of patches depending on the location of vertex z:

$$z \in \Omega, \quad z \in \Gamma_D \backslash \bar{\Gamma}_N, \quad z \in \Gamma_N \backslash \bar{\Gamma}_D, \quad z \in \bar{\Gamma}_D \cap \bar{\Gamma}_N.$$

Illustrations for those four cases are shown in Figure 7.2 ($r = 5$), Figure 7.3 ($r = 3$), Figure 7.4 ($r = 3$) and Figure 7.5 ($r = 3$), respectively.

7.3.2 Choice of $J_{z,e}$ in an interior vertex patch

Assume z is an interior vertex. To fulfill the first requirement at the beginning of Section 7.3, we require

$$\int_{e_i} \hat{g}_{z,e_i} ds - \int_{e_{i-1}} \hat{g}_{z,e_{i-1}} ds = \int_{K_i} \bar{f}_z dx, \quad i = 1, \ldots, r \ (e_0 := e_r),$$

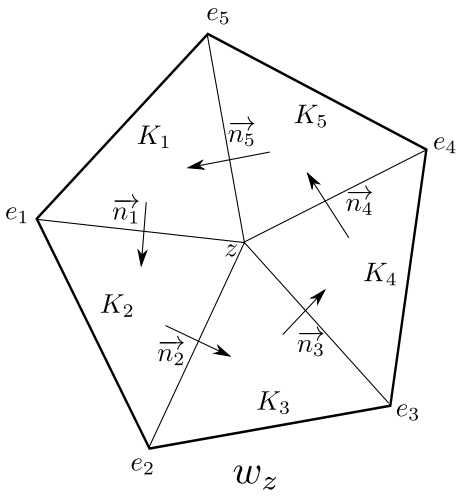

Figure 7.2. An interior vertex patch ω_z

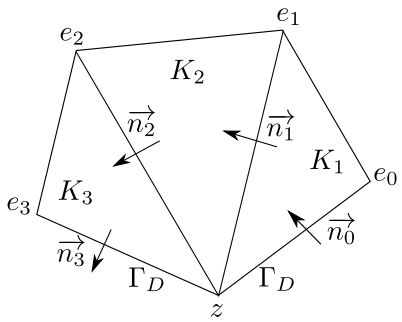

Figure 7.3. A Dirichlet boundary vertex patch ω_z

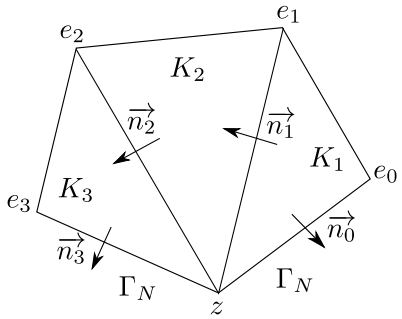

Figure 7.4. A Neumann boundary vertex patch ω_z

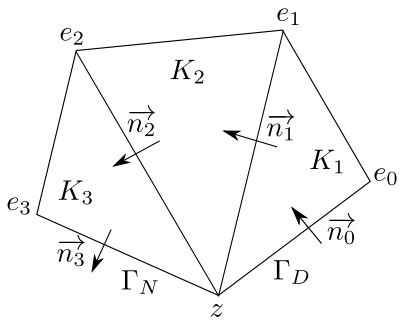

Figure 7.5. A mixed boundary vertex patch ω_z

which, according to the definition of $\hat{g}_{z,e}$ in (7.2.2), is equivalent to

$$|e_i|J_{z,e_i} - |e_{i-1}|J_{z,e_{i-1}} = B_i, \quad i = 1, \ldots, r \ (e_0 := e_r), \tag{7.3.5}$$

with

$$B_i := \int_{K_i} \bar{f}_z dx - \sum_{e \in \mathcal{E}_{K_i}} \int_e s_{K_i}|_e \tilde{g}_{z,e} ds. \tag{7.3.6}$$

It can be seen that (cf. [16, 17, 20]) the r-by-r linear system of $\{J_{z,e_i}\}_{i=1}^r$ in (7.3.5) is solvable and a particular solution is uniquely determined by assigning a value to J_{z,e_1}. We choose

$$J_{z,e_1} = 0 \tag{7.3.7}$$

and $J_{z,e_2}, \ldots, J_{z,e_r}$ can be determined immediately from (7.3.5). Following the indexing in Section 7.3.1, it will be shown in Section 7.4.2 that, under the quasi-monotonicity condition of α, this particular solution $\{J_{z,e_i}\}_{i=1}^r$ satisfies (7.4.3), which implies a robust local efficiency bound of the estimator.

7.3.3 Choice of $J_{z,e}$ in a Dirichlet boundary vertex patch

For $z \in \Gamma_D \backslash \overline{\Gamma}_N$ as illustrated in Figure 7.3, the equilibration of $\hat{\boldsymbol{\sigma}}_{\mathcal{T}}$ requires

$$\int_{e_i} \hat{g}_{z,e_i} ds - \int_{e_{i-1}} \hat{g}_{z,e_{i-1}} ds = \int_{K_i} \bar{f}_z dx, \quad i = 1, \ldots, r,$$

which, according to (7.2.2), is equivalent to

$$|e_i|J_{z,e_i} - |e_{i-1}|J_{z,e_{i-1}} = B_i, \quad i = 1, \ldots, r, \tag{7.3.8}$$

where B_i is defined in (7.3.6). It can be seen that (cf. [16, 20]) the r-by-$r+1$ linear system of $\{J_{z,e_i}\}_{i=0}^r$ is solvable. We choose a particular solution by assigning a value to J_{z,e_s}, where s is the index in (7.3.4). We define

$$J_{z,e_s} = 0 \tag{7.3.9}$$

and $J_{z,e_i} (i = 1, \ldots, s-1, s+1, \ldots, r)$ can be determined by (7.3.8) along the paths illustrated in Figure 7.7. It will be shown in Section 7.4.3 that, under the quasi-monotonicity condition of α, this particular solution $\{J_{z,e_i}\}_{i=0}^r$ satisfies (7.4.3), which contributes to the robust local efficiency of the resulting estimator.

7.3.4 Choice of $J_{z,e}$ in a Neumann boundary vertex patch

For $z \in \Gamma_N \backslash \bar{\Gamma}_D$ as illustrated in Figure 7.4, according to [20], the compatibility condition in (7.2.5) requires

$$\begin{aligned} \int_{e_1} \hat{g}_{z,e_1} ds &= \int_{K_1} \bar{f}_z dx - \int_{e_0} \phi_z g_N ds, \\ \int_{e_i} \hat{g}_{z,e_i} ds - \int_{e_{i-1}} \hat{g}_{z,e_{i-1}} ds &= \int_{K_i} \bar{f}_z dx, \quad i = 2, \ldots, r-1, \\ -\int_{e_{r-1}} \hat{g}_{z,e_{r-1}} ds &= \int_{K_r} \bar{f}_z dx - \int_{e_r} \phi_z g_N ds, \end{aligned}$$

which, according to (7.2.2), is equivalent to

$$\begin{aligned} |e_1| J_{z,e_1} &= B_1, \\ |e_i| J_{z,e_i} - |e_{i-1}| J_{z,e_{i-1}} &= B_i, \quad i = 2, \ldots, r-1, \\ -|e_{r-1}| J_{z,e_{r-1}} &= B_r, \end{aligned} \tag{7.3.10}$$

where B_i is defined in (7.3.6). It is known that (cf. [16, 20]) the r-by-$(r-1)$ linear system of $\{J_{z,e_i}\}_{i=1}^{r-1}$ is uniquely solvable. Therefore, according to (7.3.10), we can obtain inductively $J_{z,e_1}, J_{z,e_2}, \ldots, J_{z,e_{m-1}}$ and $J_{z,e_{r-1}}, J_{z,e_{r-2}}, \ldots, J_{z,e_m}$, via the paths illustrated in Figure 7.8. It will be shown in Section 7.4.4 that, under the quasi-monotonicity condition of α, $\{J_{z,e_i}\}_{i=1}^{r-1}$ satisfies (7.4.3), which implies the robust local efficiency of the resulting error estimator.

7.3.5 Choice of $J_{z,e}$ in a mixed boundary vertex patch

For $z \in \bar{\Gamma}_D \cap \bar{\Gamma}_N$ as illustrated in Figure 7.5, according to [20], the compatibility condition in (7.2.5) requires

$$\begin{aligned} \int_{e_i} \hat{g}_{z,e_i} ds - \int_{e_{i-1}} \hat{g}_{z,e_{i-1}} ds &= \int_{K_i} \bar{f}_z dx, \quad i = 1, \ldots, r-1, \\ -\int_{e_{r-1}} \hat{g}_{z,e_{r-1}} ds &= \int_{K_r} \bar{f}_z dx - \int_{e_r} \phi_z g_N ds, \end{aligned}$$

which, according to (7.2.2), is equivalent to

$$\begin{aligned} |e_i| J_{z,e_i} - |e_{i-1}| J_{z,e_{i-1}} &= B_i, \quad i = 1, \ldots, r-1, \\ -|e_{r-1}| J_{z,e_{r-1}} &= B_r, \end{aligned} \tag{7.3.11}$$

where B_i is defined in (7.3.6). It is known that (cf. [16, 20]) the r-by-r linear system of $\{J_{z,e_i}\}_{i=0}^{r-1}$ is uniquely solvable. Hence from (7.3.11), we can evaluate inductively $J_{z,e_{r-1}}, J_{z,e_{r-2}}, \ldots, J_{z,e_0}$, along the path illustrated in Figure 7.9. It will be shown in Section 7.4.5 that, under the quasi-monotonicity condition of α, $\{J_{z,e_i}\}_{i=0}^{r-1}$ satisfies (7.4.3), which implies the robust local efficiency of the resulting error estimator.

7.4 Proof Of Robust Local Efficiency

We prove the robust local efficiency of ξ_z as well as ξ_K in Theorem 7.2.1 under the quasi-monotonicity assumption of α. It suffices to show that the error indicator in each element can be bounded from above by the robust residual estimator [8, 9], which is stated as Lemma 7.4.1. Then the robust local efficiency follows from that of the residual estimator.

We will use C with or without subscripts to denote a generic positive constant, possibly different at different occurrences, that only depends on shape parameter of the mesh $\mathcal{T}$, the polynomial degree k, and is independent of $\alpha_{\max}/\alpha_{\min}$, ω_z, mesh size, etc.

Lemma 7.4.1. *Let η_K be the residual-based error indicator over K defined in* (4.3.1)*. For each vertex $z \in \mathcal{N}$, let $\boldsymbol{\sigma}_{z,\mathcal{T}}$ be the local numerical flux in* (6.3.4) *and $\hat{\boldsymbol{\sigma}}_{z,\mathcal{T}}$ be the recovered local flux in* (7.2.3)*. If α is quasi-monotone with respect to z, then*

$$\|\alpha^{-1/2}\left(\hat{\boldsymbol{\sigma}}_{z,\mathcal{T}} - \boldsymbol{\sigma}_{z,\mathcal{T}}\right)\|_K \leq C \sum_{T\in\mathcal{T}_z} \eta_T, \quad \forall\, K \in \mathcal{T}_z. \tag{7.4.1}$$

Proof. According to [12, Lemma 4.1], $\hat{\boldsymbol{\sigma}}_{z,\mathcal{T}}$ defined in (7.2.3) satisfies the following estimate

$$\|\hat{\boldsymbol{\sigma}}_{z,\mathcal{T}} - \boldsymbol{\sigma}_{z,\mathcal{T}}\|_K \leq C\left(h_K\|\bar{f}_z - \operatorname{div}\boldsymbol{\sigma}_{z,\mathcal{T}}\|_K + h_K^{1/2}\sum_{e\in\mathcal{E}_K}\|\hat{g}_{z,e} - \boldsymbol{\sigma}_{z,\mathcal{T}}|_K\cdot\boldsymbol{n}_e\|_e\right),$$

where constant C only depends on the polynomial degree k and the shape regularity of $\mathcal{T}$. Together with (7.2.2) and the triangle inequality, it follows that

$$\|\alpha^{-1/2}\left(\hat{\boldsymbol{\sigma}}_{z,\mathcal{T}}-\boldsymbol{\sigma}_{z,\mathcal{T}}\right)\|_K \leq Ch_K\alpha_K^{-1/2}\|\bar{f}_z-\operatorname{div}\boldsymbol{\sigma}_{z,\mathcal{T}}\|_K + Ch_K^{1/2}\alpha_K^{-1/2}\sum_{e\in\mathcal{E}_K}\|\tilde{g}_{z,e}-\boldsymbol{\sigma}_{z,\mathcal{T}}|_K\cdot\boldsymbol{n}_e\|_e$$
$$+ Ch_K^{1/2}\alpha_K^{-1/2}\sum_{e\in\mathcal{E}_K\setminus(\mathcal{E}_z\cup\mathcal{E}_N)}\|J_{z,e}\|_e.$$

The first term in the upper bound above can be estimated as below

$$\|\bar{f}_z-\operatorname{div}\boldsymbol{\sigma}_{z,\mathcal{T}}\|_K = \|\Pi_K^k(\phi_z f-\phi_z\operatorname{div}\boldsymbol{\sigma}_{\mathcal{T}})\|_K \leq h_K^{-1}\alpha_K^{1/2}\|f-\operatorname{div}\boldsymbol{\sigma}_{\mathcal{T}}\|_K \leq \eta_K.$$

The second term is estimated by using (4.3.7):

$$\|\tilde{g}_{z,e}-\boldsymbol{\sigma}_{z,\mathcal{T}}|_K\cdot\boldsymbol{n}_e\|_e \leq C\left(\frac{\alpha_K}{\alpha_e}\right)^{1/2}\|j_e\|_e,\quad\forall\, e\in\mathcal{E}_K, \tag{7.4.2}$$

which implies

$$h_K^{1/2}\alpha_K^{-1/2}\sum_{e\in\mathcal{E}_K}\|\tilde{g}_{z,e}-\boldsymbol{\sigma}_{z,\mathcal{T}}|_K\cdot\boldsymbol{n}_e\|_e \leq C\sum_{e\in\mathcal{E}_K}h_K^{1/2}\alpha_e^{-1/2}\|j_e\|_e \leq C\eta_K.$$

Therefore, it remains to bound the third term from above by the residual estimator. This is true thanks to Lemma 7.4.2. Hence (7.4.1) is justified and the proof is complete. □

Lemma 7.4.2. *Let ω_z be the patch associated with vertex $z\in\mathcal{N}$ and assume that the elements in ω_z are indexed according to Section 7.3.1. Let $J_{z,e}$ be given in Section 7.3. If α is quasi-monotone with respect to z, then*

$$h_K^{1/2}\alpha_K^{-1/2}\|J_{z,e}\|_e \leq C\sum_{T\in\mathcal{T}_z}\eta_T,\quad\forall\, K\subseteq\omega_z,\, e\in\mathcal{E}_K\setminus(\mathcal{E}_z\cup\mathcal{E}_N). \tag{7.4.3}$$

The rest of this section is dedicated to the proof of Lemma 7.4.2.

Remark 7.4.1. *The existing proofs of robust local efficiency* (cf., [10, 20]) *of the indicator in* (7.2.6) *is based on the mixed formulation of* (7.1.4)*. Therefore, the proofs in* [10, 20] *are only valid for recovered flux $\hat{\boldsymbol{\sigma}}_{z,\mathcal{T}}$ with the minimization property in* (7.1.4) *and are not applicable to the flux $\hat{\boldsymbol{\sigma}}_{z,\mathcal{T}}$ constructed in Section 7.2.*

7.4.1 The quasi-monotonicity condition

The quasi-monotonicity condition of α [9] is required for the robust local efficiency of the indicator defined in (7.2.6). This condition is weaker than the Hypothesis 2.7 in [8] and we cite the quasi-monotonicity condition [9, Definition 4.1] below for completeness.

Definition 7.4.1. *For $z \in \mathcal{N}$, denote by K_z an element in $\mathcal{T}_z$ such that*

$$\alpha_{K_z} = \max_{K \in \mathcal{T}_z} \alpha_K.$$

α is called quasi-monotone with respect to z *if the following conditions are satisfied.*

For each $K \in \mathcal{T}_z$, there exists a Lipschitz set $\tilde{\omega}_{K,z}$ containing only elements from $\mathcal{T}_z$ such that

- *if $z \notin \bar{\Gamma}_D$, then $K \cup K_z \subseteq \tilde{\omega}_{K,z}$ and $\alpha_K \leq \alpha_{K'}$, $\forall\, K' \subseteq \tilde{\omega}_{K,z}$;*
- *if $z \in \bar{\Gamma}_D$, then $K \subseteq \tilde{\omega}_{K,z}$, $meas(\partial\tilde{\omega}_{K,z} \cap \Gamma_D) > 0$ and $\alpha_K \leq \alpha_{K'}$, $\forall\, K' \subseteq \tilde{\omega}_{K,z}$.*

Following [9], we say that α is quasi-monotone if it is quasi-monotone with respect to all $z \in \mathcal{N}$.

For a vertex patch ω_z in two dimensions, the following consequence is straightforward to derive if α is quasi-monotone.

Proposition 7.4.1. *Assume α is quasi-monotone with respect to z in two dimensions. With notation introduced in Section 7.3.1, the following relations hold true:*

- *if $z \in \Omega$ as illustrated in Figure 7.6, then*

$$\alpha_{K_1} \leq \alpha_{K_2} \leq \cdots \leq \alpha_{K_m} \quad \textit{and} \quad \alpha_{K_1} \leq \alpha_{K_r} \leq \alpha_{K_{r-1}} \cdots \leq \alpha_{K_m}; \tag{7.4.4}$$

- *if $z \in \Gamma_D \backslash \bar{\Gamma}_N$ as illustrated in Figure 7.7, then*

$$\alpha_{K_s} \leq \alpha_{K_{s+1}} \leq \cdots \leq \alpha_{K_r} \quad \textit{and} \quad \alpha_{K_s} \leq \alpha_{K_{s-1}} \leq \cdots \leq \alpha_{K_1}; \tag{7.4.5}$$

- *if* $z \in \Gamma_N \backslash \bar{\Gamma}_D$ *as illustrated in Figure 7.8, then*

$$\alpha_{K_1} \leq \alpha_{K_2} \leq \cdots \leq \alpha_{K_m} \quad \textit{and} \quad \alpha_{K_r} \leq \alpha_{K_{r-1}} \leq \cdots \leq \alpha_{K_m}; \tag{7.4.6}$$

- *if* $z \in \bar{\Gamma}_D \cap \bar{\Gamma}_N$ *as illustrated in Figure 7.9, then*

$$\alpha_{K_r} \leq \alpha_{K_{r-1}} \leq \cdots \leq \alpha_{K_1}. \tag{7.4.7}$$

Proof. For $z \in \Omega$, we prove (7.4.4) by contradiction. Assume there exists an element K_i with $1 < i < m$ such that $\alpha_{K_i} < \alpha_{K_{i-1}}$. Since α_{K_1} is the minimum according to (7.3.3), it follows that $\alpha_{K_1} \leq \alpha_{K_i} < \alpha_{K_{i-1}}$, which violates the quasi-monotonicity of α if we choose $K_z = K_m$ and $K = K_{i-1}$ in Definition 7.4.1. Hence the first inequality in (7.4.4) is justified. The second inequality in (7.4.4) can be proved analogously.

For $z \in \Gamma_D \backslash \bar{\Gamma}_N$, (7.4.5) can be proved in a similar fashion.

For $z \in \Gamma_N \backslash \bar{\Gamma}_D$, consider an arbitrary element $K = K_i$ with $1 \leq i < m$ and choose $K_z = K_m$ in Definition 7.4.1. Then $K_i \cup K_m \subseteq \tilde{\omega}_{K_i,z}$ and consequently $K_{i+1} \subseteq \tilde{\omega}_{K_i,z}$, as $\tilde{\omega}_{K_i,z}$ is connected. Therefore, Definition 7.4.1 implies that $\alpha_{K_i} \leq \alpha_{K_{i+1}}$, which justifies the first inequality in (7.4.6). The second inequality in (7.4.6) can be proved analogously.

For $z \in \bar{\Gamma}_D \cap \bar{\Gamma}_N$, since K_1 is the only element in ω_z such that $\text{meas}(\partial K_1 \cap \Gamma_D) > 0$, it readily follows from Definition 7.4.1 that $K_1 \subseteq \tilde{\omega}_{K,z}, \quad \forall\, K \subseteq \omega_z$. Consider an arbitrary element K_i with $1 < i \leq r$. Then $K_1 \cup K_i \subseteq \tilde{\omega}_{K_i,z}$ and consequently $K_{i-1} \subseteq \tilde{\omega}_{K_i,z}$, as $\tilde{\omega}_{K_i,z}$ is connected. Therefore, Definition 7.4.1 implies that $\alpha_{K_i} \leq \alpha_{K_{i-1}}$, which concludes the proof of (7.4.7). □

7.4.2 Proof of Lemma 7.4.2 - interior vertex patch

Suppose $z \in \Omega$. We first estimate $\alpha_{K_i}^{-1/2}|B_i|$. It follows from the definitions of $B_i, \bar{f}_z$ in (7.3.6), (7.1.2), respectively, and divergence theorem that

$$B_i = \sum_{e \in \mathcal{E}_{K_i}} \int_e s_{K_i}|_e (\boldsymbol{\sigma}_{z,\mathcal{T}}|_{K_i} \cdot \boldsymbol{n}_e - \tilde{g}_{z,e}) ds + \int_{K_i} \phi_z (f - \operatorname{div} \boldsymbol{\sigma}_{\mathcal{T}}) dx.$$

Applying the triangle and the Cauchy-Schwarz inequalities yields

$$\alpha_{K_i}^{-1/2}|B_i| \leq C\alpha_{K_i}^{-1/2}\left(\sum_{e\in\mathcal{E}_{K_i}}|e|^{1/2}\|\boldsymbol{\sigma}_{z,\mathcal{T}}|_{K_i}\cdot\boldsymbol{n}_e-\tilde{g}_{z,e}\|_e+|K_i|^{1/2}\|f-\operatorname{div}\boldsymbol{\sigma}_{\mathcal{T}}\|_{K_i}\right),$$

which, together with (7.4.2), implies that

$$h_{K_i}^{1/2}\alpha_{K_i}^{-1/2}|e_i|^{-1}\|B_i\|_{e_i}\leq C\,\eta_{K_i}, \tag{7.4.8}$$

where C is independent of $\alpha_{\max}/\alpha_{\min}$.

We next prove (7.4.3) by using relation (7.3.5) and Proposition 7.4.1 inductively.

Consider the paths from K_1 to K_m as illustrated in Figure 7.6 which constitute a traversal of all elements contained in ω_z. We first estimate $h_K^{1/2}\alpha_K^{-1/2}\|J_{z,e}\|_e, e\in\mathcal{E}_K\backslash\mathcal{E}_z$ for each K along the path

$$K_1\to K_2\to\cdots\to K_m. \tag{7.4.9}$$

The quasi-monotonicity of α in (7.4.4) and the choice $J_{z,e_1}=0$ in (7.3.7) imply

$$h_{K_2}^{1/2}\alpha_{K_2}^{-1/2}\|J_{z,e_1}\|_{e_1}\leq Ch_{K_1}^{1/2}\alpha_{K_1}^{-1/2}\|J_{z,e_1}\|_{e_1}=0\leq C\,\eta_{K_1}. \tag{7.4.10}$$

The relation in (7.3.5), the estimates in (7.4.8) and (7.4.10), the triangle inequality and the shape regularity of $\mathcal{T}$ yield

$$\begin{aligned}h_{K_2}^{1/2}\alpha_{K_2}^{-1/2}\|J_{z,e_2}\|_{e_2}&\leq C\left(h_{K_2}^{1/2}\alpha_{K_2}^{-1/2}\|J_{z,e_1}\|_{e_1}+h_{K_2}^{1/2}\alpha_{K_2}^{-1/2}|e_2|^{-1}\|B_2\|_{e_2}\right)\\&\leq C\,(\eta_{K_1}+\eta_{K_2}).\end{aligned}$$

By repeating this argument (i.e., using (7.4.4), (7.3.5), (7.4.8)) along the path in (7.4.9), we deduce for $i=2,\ldots,m$,

$$h_{K_{i+1}}^{1/2}\alpha_{K_{i+1}}^{-1/2}\|J_{z,e_i}\|_{e_i}\leq C\,h_{K_i}^{1/2}\alpha_{K_i}^{-1/2}\|J_{z,e_i}\|_{e_i}\leq C\,(\eta_{K_1}+\cdots+\eta_{K_i}) \tag{7.4.11}$$

and

$$\begin{aligned}h_{K_{i+1}}^{1/2}\alpha_{K_{i+1}}^{-1/2}\|J_{z,e_{i+1}}\|_{e_{i+1}}&\leq C\left(h_{K_{i+1}}^{1/2}\alpha_{K_{i+1}}^{-1/2}\|J_{z,e_i}\|_{e_i}+h_{K_{i+1}}^{1/2}\alpha_{K_{i+1}}^{-1/2}|e_{i+1}|^{-1}\|B_{i+1}\|_{e_{i+1}}\right)\\&\leq C\,(\eta_{K_1}+\cdots+\eta_{K_{i+1}}).\end{aligned} \tag{7.4.12}$$

For another path

$$K_1 \to K_r \to K_{r-1} \cdots \to K_m,$$

same argument yields that, for $i = 1, r, r-1, \ldots, m+1$ ($e_0 := e_r$),

$$h_{K_i}^{1/2} \alpha_{K_i}^{-1/2} \|J_{z,e_{i-1}}\|_{e_{i-1}} \leq C \left(\eta_{K_1} + \eta_{K_r} + \cdots + \eta_{K_i}\right), \tag{7.4.13}$$

and for $i = r, r-1, \ldots, m$,

$$h_{K_i}^{1/2} \alpha_{K_i}^{-1/2} \|J_{z,e_i}\|_{e_i} \leq C \left(\eta_{K_1} + \eta_{K_r} + \cdots + \eta_{K_i}\right). \tag{7.4.14}$$

The estimates (7.4.11)–(7.4.14) conclude the proof of (7.4.3).

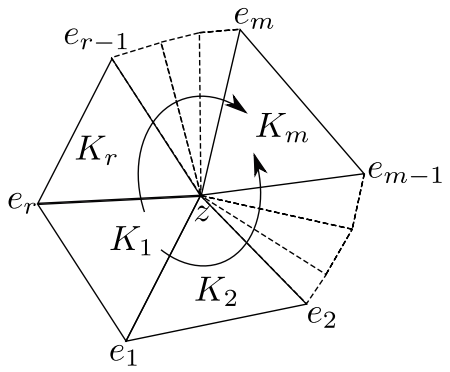

Figure 7.6. Paths from K_1 to K_m in an interior vertex patch ω_z

7.4.3 Proof of Lemma 7.4.2 - Dirichlet boundary vertex patch

The proof of (7.4.3) is in a same fashion as the one in Section 7.4.2 by considering instead the paths

$$K_s \to K_{s+1} \to \cdots \to K_r \quad \text{and} \quad K_s \to K_{s-1} \to \cdots \to K_1$$

illustrated in Figure 7.7 and using (7.3.8), (7.3.9), (7.4.5) and (7.4.8).

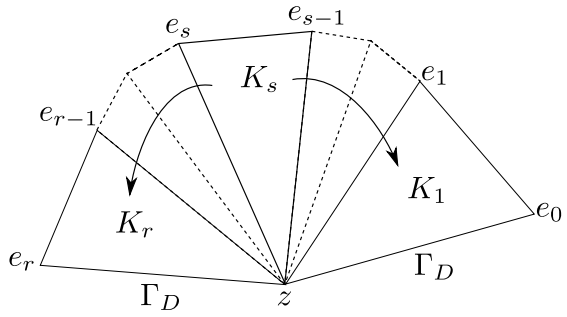

Figure 7.7. Paths from K_s to boundary in a vertex patch ω_z on the Dirichlet boundary

7.4.4 Proof of Lemma 7.4.2 - Neumann boundary vertex patch

The proof of (7.4.3) is same to Section 7.4.2 by considering instead the paths

$$K_1 \to K_2 \to \cdots \to K_m \quad \text{and} \quad K_r \to K_{r-1} \to \cdots \to K_m$$

illustrated in Figure 7.8 and using (7.3.10), (7.4.6) and (7.4.8).

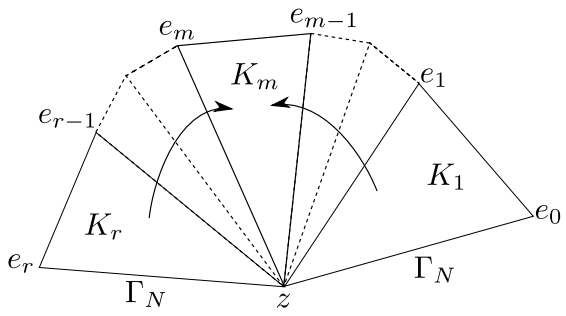

Figure 7.8. Paths from K_1 to K_m and from K_r to K_m in a vertex patch ω_z on the Neumann boundary

7.4.5 Proof of Lemma 7.4.2 - mixed boundary vertex patch

The proof of (7.4.3) is same to Section 7.4.2 by considering instead the path

$$K_r \to K_{r-1} \to \cdots \to K_1$$

illustrated in Figure 7.9 and using (7.3.11), (7.4.7) and (7.4.8).

7.5 Numerical Experiments

Let $\eta_{\min}$ denote the minimization-based estimator in [10, 20]. We perform numerical experiments to first illustrate the robustness of ξ for the Kellogg numerical benchmark, and then demonstrate the advantage of ξ over $\eta_{\min}$.

Consider the Kellogg problem as in Section 4.5 with same setup and parameters. The solution u is barely differentiable ($u \notin H^{1.1}(\Omega)$), so the example is perfect for

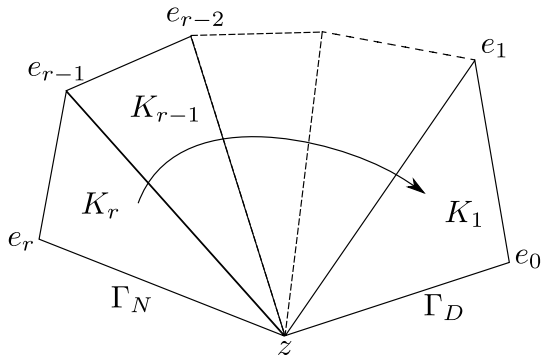

Figure 7.9. Path from K_1 to K_r in a vertex patch ω_z on both Γ_D and Γ_N

testing adaptive mesh refinement. We use RT_{k-1} flux recovery for conforming P_k element. The initial mesh consists of 4×4 congruent squares, each of which is partitioned into two triangles connecting bottom-left and top-right corners. We use Dörfler's marking strategy [46] as in [12, 46, 47] and the newest-vertex bisection [44].

The following notation will be used:

- exact error $e := u - u_{\mathcal{T}}$;
- effectivity index: eff-ind;
- degrees of freedom: DOFs;
- stopping criterion: $\|A^{1/2}\nabla e\| \leq \epsilon_{\mathrm{rel}}\|A^{1/2}\nabla u\|$ with $\epsilon_{\mathrm{rel}} = 0.05$ for P_1 element and $\epsilon_{\mathrm{rel}} = 0.01$ for P_2 element.

For P_1 discretization, the mesh and error are shown in Figure 7.10. Figure 7.11 shows the results for P_2 discretization. We see that the meshes generated using ξ are refined towards the center and there is no over-refinement along the interfaces where α is discontinuous. This justifies the robustness of ξ with respect to the jump of α. Optimal convergence rates are observed from the error plots in Figure 7.10 – 7.11 and it is clear that ξ(blue) provides a guaranteed upper bound of the true error (red).

We compare the performance of two estimators in adaptive mesh refinement for the Kellogg problem considered before. The results are collected in Table 7.1, where the last column "time" shows the total time to compute each estimator in the mesh refinement procedure. According to Table 7.1, it can be seen that: (1) in terms of accuracy(or effectivity index), $\eta_{\min}$ is slightly more accurate than ξ; (2) in terms of computational cost, ξ is much better than $\eta_{\min}$. For P_1 element, it is more costly to compute $\eta_{\min}$ than ξ but the margin is not large. For P_2 element, however, $\eta_{\min}$ is significantly more expensive to compute than ξ. The cost for computing $\eta_{\min}$ scales poorly with respect to the polynomial degree k.

Now we compare the two estimators in terms of computational costs. We use a sequence of uniform meshes and record the time to compute the estimator for each

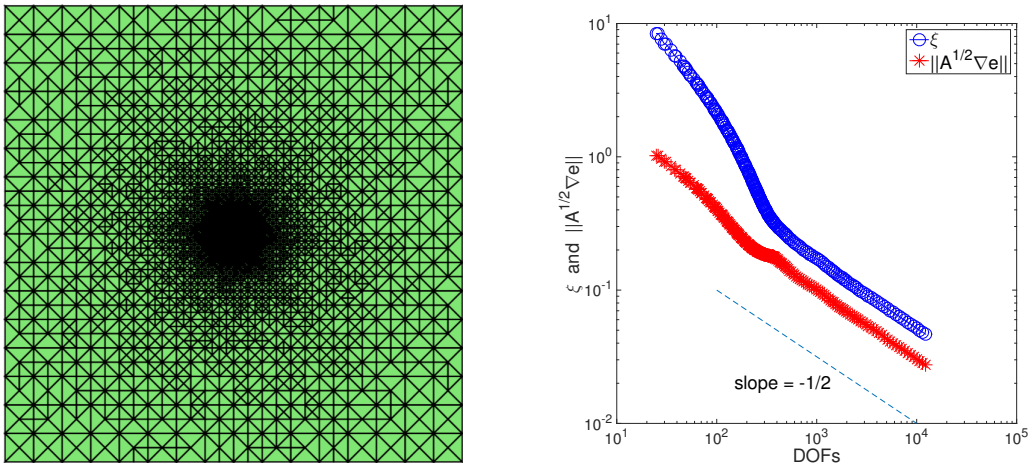

Figure 7.10. Left: mesh; Right: error and ξ (P_1 discretization).

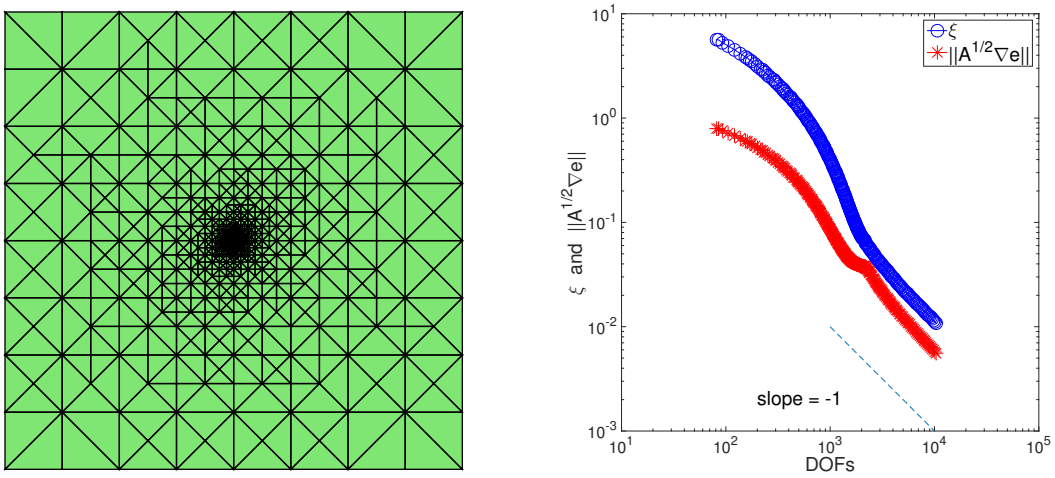

Figure 7.11. Left: mesh; Right: error and ξ (P_2 discretization).

Table 7.1.
Comparison of $\eta_{\min}$ [10, 20] and ξ in adaptive mesh refinement

$u_{\mathcal{T}}$	estimator	DOFs	$\|A^{1/2}\nabla e\|/\|A^{1/2}\nabla u\|$	eff-ind	time
P_1	$\eta_{\min}$	12410	4.9E-2	1.47	571s
	ξ	12303	4.9E-2	1.69	310s
P_2	$\eta_{\min}$	10237	1E-2	1.62	1611s
	ξ	10401	1E-2	1.92	251s

Table 7.2.
Timing for $\eta_{\min}$ [10, 20] and ξ (P_1 element)

DOFs	time($\eta_{\min}$)	time(ξ)
441	1.5s	0.6s
841	2.5s	1.2s
1681	4.9s	2.3s
3249	9.5s	4.5s
6561	18.9s	8.9s

mesh. Table 7.2 shows that, for P_1 element, the cost of computing $\eta_{\min}$ is about twice that of ξ. Table 7.3 shows that, for P_2 element, the cost of computing $\eta_{\min}$ is about eight times that of ξ.

We conclude that: (1) $\eta_{\min}$ and ξ are both robust with respect to the discontinuity of the diffusion coefficient; (2) they have similar effectivity indices; (3) ξ is much cheaper to compute than $\eta_{\min}$: the computational cost for $\eta_{\min}$ grows rapidly as the polynomial degree k increases, while the cost for ξ is not sensitive to k.

Table 7.3.
Timing for $\eta_{\min}$ $[10, 20]$ and ξ (P_2 element)

DOFs	time($\eta_{\min}$)	time(ξ)
1681	6.6s	0.9s
3249	12.5s	1.6s
6561	25.4s	3.3s
12769	50.5s	6.0s
25921	97.7s	12.7s

8. EQUILIBRATED ESTIMATOR FOR NONCONFORMING FINITE ELEMENT APPROXIMATIONS

The a posteriori error estimation based on equilibrated flux recovery has received increasing popularity recently since one is able to derive a *guaranteed* upper bound of the true error. As a result, "equilibrated estimators" are perfect for the error control in numerical simulations. For conforming finite element discretizations, numerous methods have been proposed since the initial work in [15] and we refer to the discussion in Chapter 6. For nonconforming finite element discretizations, it was only recently that the problem has attracted attention. See, for example, [22, 23, 31, 32]. This chapter will focus on nonconforming elements.

In existing literatures, most of the equilibrated estimators for nonconforming elements are only designed for lower order elements, e.g., the Crouzeix–Raviart element [64] (P_1 element) or the Fortin–Soulie element [65] (P_2 element). To the best of our knowledge, the only paper that provides a robust equilibrated error estimator for *arbitrary* order nonconforming triangular elements is [22]. The estimator in [22] was shown to provide a guaranteed upper bound and a robust local lower bound for diffusion problems with discontinuous coefficients, assuming the diffusion coefficient is quasi-monotone. Different from the a posteriori error estimation for conforming elements, existing error estimators for nonconforming elements strongly depend on the odd/even order of the approximation space. For example, the definitions of the estimator in [22] (as well as the analysis) for odd and even order nonconforming elements are considered separately, as pointed out in [22] that

> "The fundamental difference in the structure of the odd and even order nonconforming elements also manifests itself in the structure of the estimators."

and that

> "The essential difference in the structures of the degrees of freedom in the case of odd versus even order elements is well known for nonconforming finite elements. This phenomenon will manifest itself throughout the subsequent analysis."

Contrary to the common belief that the structure of nonconforming elements will impact the a posteriori error estimation, we ask the following questions:

> Can we design an explicit equilibrated estimator *independent* of the odd/even order of the nonconforming element? If, so, can we prove the robustness with respect to the diffusion coefficient jump without distinguishing the odd and even orders?

We show in this the chapter that the answers are YES. We present an explicit equilibrated estimator that is *universal* for *arbitrary* order nonconforming elements and is robust with respect to the discontinuities of the diffusion coefficient. That is to say, both the explicit estimator and the proof of robust estimates are *independent* of the odd/even order of the nonconforming element. Hence there is no need to distinguish the odd and even orders and jump into the immense technicalities, for example, as in [22]. To the best of our knowledge, the estimator introduced here is the *first* estimator for nonconforming elements that satisfies the all properties below:

- explicit;
- valid for arbitrary order nonconforming elements;
- independent of the odd/even order of nonconforming element;
- provides a guaranteed upper bound of the true error;

- yields a robust local lower bound if the diffusion coefficient is quasi-monotone.

The chapter is organized as follows. Section 8.1 introduces the nonconforming finite element approximation to the model problem. Section 8.2 presents the general formula for the equilibrated flux, valid for arbitrary order nonconforming elements. The robust a posteriori error estimates are derived in Section 8.3. Section 8.4 is auxiliary, offering a choice of a function used in Section 8.2.

8.1 Problem And Nonconforming Finite Element Approximation

Let Ω be a bounded open polygonal domain on the plane with boundary $\partial\Omega = \bar{\Gamma}_D \cup \bar{\Gamma}_N$ and $\Gamma_D \cap \Gamma_N = \emptyset$. Assume that $\mathrm{meas}(\Gamma_D) > 0$. Consider the same equation as in (7). That is,

$$\begin{cases} -\mathrm{div}(A\nabla u) = f & \text{in } \Omega, \\ u = 0 & \text{on } \Gamma_D, \\ -(A\nabla u)\cdot \boldsymbol{n} = g_{_N} & \text{on } \Gamma_N, \end{cases} \tag{8.1.1}$$

where $\boldsymbol{n}$ denotes the unit outward normal on $\partial\Omega$. Denote by $\alpha(x)$ the smallest eigenvalue of $A(x)$.

Let the mesh information (elements,vertices,edges/faces,etc.) be given in Section 3.1. To ignore data oscillation in the finite element approximation, for simplicity, it will be assumed that for each $K \in \mathcal{T}$, $\alpha_{_K} := \alpha|_K$ is constant and

$$f|_K \in P_{k-1}(K),\, \forall\, K \in \mathcal{T}, \quad g_{_N}|_e \in P_{k-1}(e),\, \forall\, e \in \mathcal{E}_N.$$

Of particular interest is the case where the value of α undergoes large jumps across interelement edges.

The $P_k (k \geq 1)$ nonconforming finite element space $V_{\mathcal{T}}$ is defined by

$$V_{\mathcal{T}} := \{v \in L^2(\Omega) : v|_K \in P_k(K)\, \forall\, K \in \mathcal{T}, \int_e [\![v]\!]_e p ds = 0,\, \forall\, p \in P_{k-1}(e), e \in \mathcal{E}_I\},$$

where $[\![v]\!]_e$ denotes the jump of v across the edge e if $e \in \mathcal{E}_I$. Corresponding to the Dirichlet boundary condition in (8.1.1), the nonconforming finite element space is

$$V_{\mathcal{T},D} := \left\{ v \in V_{\mathcal{T}} : \int_e [\![v]\!]_e p\, ds = 0,\, \forall\, p \in P_{k-1}(e), e \in \mathcal{E}_D \right\}. \tag{8.1.2}$$

Since the Gauss quadrature of order k is exact for polynomials of degree up to $2k-1$, from the definition in (8.1.2), it can be seen that if a piecewise P_k function v is continuous at k Gauss points on $e \in \mathcal{E}_I$ and also vanishes at k Gauss points on $e \in \mathcal{E}_D$, then $v \in V_{\mathcal{T},D}$.

The finite element approximation of (8.1.1) is to find $u_{\mathcal{T}} \in V_{\mathcal{T},D}$ such that

$$\int_\Omega A\nabla_h u_{\mathcal{T}} \cdot \nabla_h v dx = \int_\Omega f v dx - \int_{\Gamma_N} v g_{_N} ds, \quad \forall\, v \in V_{\mathcal{T},D}, \tag{8.1.3}$$

where $\nabla_h v$ is defined piecewisely with respect to $\mathcal{T}$, i.e., $(\nabla_h v)|_K := \nabla(v|_K)$ for each $K \in \mathcal{T}$.

Here we recall the notation introduced in Section 3.1. For each edge $e \in \mathcal{E}$, let ω_e be the union of elements that share e in common. We associate each e with a unit vector $\boldsymbol{n}_e$ normal to e and $\boldsymbol{n}_e$ is chosen to be the unit outward normal if $e \subset \partial\Omega$. In the rest of this section, $\boldsymbol{n}$ denotes the unit outward normal with respect to ∂K. For $e \in \mathcal{E}_I$, denote by K_e^+ and K_e^- the two elements adjacent to e such that $\boldsymbol{n}_e$ is outward for K_e^+. For $e \subset \partial\Omega$, we set $K_e^+ = K_e^- = \omega_e = K$. Recall the sign function $s_{_K}$ defined in Section 3.1:

$$s_{_K} : \partial K \to \{-1, 1\}, \quad s_{_K}|_e = \begin{cases} 1, & \text{if } \boldsymbol{n}_e \text{ is an outward normal on } \partial K, \\ -1, & \text{if } \boldsymbol{n}_e \text{ is an inward normal on } \partial K. \end{cases}$$

8.2 Equilibrated Flux Recovery For Nonconforming Elements Of Arbitrary Order

In this section, we give explicit formula for the recovered flux $\hat{\boldsymbol{\sigma}}_{\mathcal{T}}$ in terms of degrees of freedom in the $H(\mathrm{div};\Omega)$-conforming Raviart-Thomas spaces. The formula is valid for arbitrary polynomial degree $k \geq 1$. $\hat{\boldsymbol{\sigma}}_{\mathcal{T}}$ is shown to be equilibrated in Theorem 8.2.1.

Denote the numerical flux by

$$\boldsymbol{\sigma}_{\mathcal{T}} := -A\nabla_h u_{\mathcal{T}}.$$

First we define some functions and quantities associated with each edge. For each $e \in \mathcal{E}$, we choose a function $v_e \in V_{\mathcal{T}}$ supported in ω_e such that $|v_e| \leq 1$ in Ω, $v_e = 1$ at Gauss points on e and $v_e = 0$ at Gauss points on $e' \neq e$. Note that if $e \in \mathcal{E}_I \cup \mathcal{E}_N$, then $v_e \in V_{\mathcal{T},D}$. One choice of such a function v_e is given in Section 8.4. For each $e \in \mathcal{E}$, define the function $\hat{f}_e$ and the weight λ_e by

$$\hat{f}_e := f v_e - A\nabla_h u_{\mathcal{T}} \cdot \nabla_h v_e \quad \text{and} \quad \lambda_e := \frac{\alpha_{K_e^-}^{-1} h_{K_e^-}}{\alpha_{K_e^+}^{-1} h_{K_e^+} + \alpha_{K_e^-}^{-1} h_{K_e^-}}, \tag{8.2.1}$$

respectively. For $e \in \mathcal{E} \backslash \mathcal{E}_N$, define the constant J_e by

$$J_e := |e|^{-1} \left(\int_{K_e^+} \hat{f}_e dx - \int_e (1-\lambda_e) \boldsymbol{\sigma}_{\mathcal{T}}|_{K_e^+} \cdot \boldsymbol{n}_e + \lambda_e \boldsymbol{\sigma}_{\mathcal{T}}|_{K_e^-} \cdot \boldsymbol{n}_e ds \right). \tag{8.2.2}$$

Additionally, we set $J_e = 0$ for $e \in \mathcal{E}_N$.

Next we define the recovered flux $\hat{\boldsymbol{\sigma}}_{\mathcal{T}}$. Let $\mathrm{RT}_{k-1}(K)$ denote the Raviart-Thomas space of index $k-1$ on $K \in \mathcal{T}$, i.e.,

$$\mathrm{RT}_{k-1}(K) := P_{k-1}(K)^d + \boldsymbol{x} P_{k-1}(K) \quad (k \geq 1).$$

The degrees of freedom of $\hat{\boldsymbol{\sigma}}_{\mathcal{T}}|_K \in \mathrm{RT}_{k-1}(K)$ are given by

$$\begin{cases} \hat{\boldsymbol{\sigma}}_{\mathcal{T}} \cdot \boldsymbol{n}_e = \hat{g}_e, \quad \forall\, e \in \mathcal{E}_K, \\ \displaystyle\int_K \hat{\boldsymbol{\sigma}}_{\mathcal{T}} \cdot \nabla p dx = \sum_{e \in \mathcal{E}_K} (s_K \hat{g}_e, p)_e - (f,p)_K, \quad \forall\, p \in P_{k-1}(K), \\ \displaystyle\int_K \hat{\boldsymbol{\sigma}}_{\mathcal{T}} \cdot \boldsymbol{q} dx = \int_K \boldsymbol{\sigma}_{\mathcal{T}} \cdot \boldsymbol{q} dx, \quad \forall\, \boldsymbol{q} \in Q_{k-2}(K), \ \text{if} \ k \geq 2, \end{cases} \tag{8.2.3}$$

where $Q_{k-2}(K)$ is defined in (4.2.11) and

$$\hat{g}_e := \begin{cases} g_N, & \text{if } e \in \mathcal{E}_N, \\ (1-\lambda_e) \boldsymbol{\sigma}_{\mathcal{T}}|_{K_e^+} \cdot \boldsymbol{n}_e + \lambda_e \boldsymbol{\sigma}_{\mathcal{T}}|_{K_e^-} \cdot \boldsymbol{n}_e + J_e, & \text{if } e \notin \mathcal{E}_N. \end{cases} \tag{8.2.4}$$

Since the normal component of $\hat{\boldsymbol{\sigma}}_{\mathcal{T}}$ is continuous across each edge, we see that $\hat{\boldsymbol{\sigma}}_{\mathcal{T}} \in H(\mathrm{div}; \Omega)$. Note that the second identity in (8.2.3) does hold true when $k = 1$ (thus the left-hand side is 0) and $\hat{\boldsymbol{\sigma}}_{\mathcal{T}}$ is indeed equilibrated, as stated in the theorem below.

Theorem 8.2.1. *Let $\hat{g}_e$ be defined in* (8.2.4) *and $\hat{\boldsymbol{\sigma}}_{\mathcal{T}}$ be defined in* (8.2.3)*. Then*

$$\sum_{e\in\mathcal{E}_K}\int_e s_{_K}\hat{g}_e ds = \int_K f dx \tag{8.2.5}$$

and

$$\begin{cases} \operatorname{div}\hat{\boldsymbol{\sigma}}_{\mathcal{T}} = f & \text{in } \Omega, \\ \hat{\boldsymbol{\sigma}}_{\mathcal{T}}\cdot\boldsymbol{n} = g_{_N} & \text{on } \Gamma_N. \end{cases} \tag{8.2.6}$$

To prove Theorem 8.2.1, we need the following two propositions.

Proposition 8.2.1. *Let $\hat{f}_e, J_e, \hat{g}_e$ be given in* (8.2.1)*,* (8.2.2)*,* (8.2.4)*, respectively. Then for each $e\in\mathcal{E}$, the following identity holds:*

$$s_{_K}|_e\int_e \hat{g}_e ds = \int_K \hat{f}_e dx, \quad K = K_e^+ \text{ or } K_e^-. \tag{8.2.7}$$

Proof. As $v_e\in V_{\mathcal{T},D}$ for $e\in\mathcal{E}_I\cup\mathcal{E}_N$, inserting v_e into (8.1.3) yields

$$\int_{\omega_e}\hat{f}_e dx = \int_{\Gamma_N} v_e g_{_N} ds = \begin{cases} 0, & \text{if } e\in\mathcal{E}_I, \\ \int_e g_{_N} ds, & \text{if } e\in\mathcal{E}_N. \end{cases} \tag{8.2.8}$$

If $e\in\mathcal{E}_N$, then $\hat{g}_e = g_{_N}$ and (8.2.7) follows from (8.2.8). If $e\in\mathcal{E}_D$, then $K_e^+ = K_e^- = K$ and (8.2.7) automatically holds true in view of the definition of J_e. If $e\in\mathcal{E}_I$, for the case $K = K_e^+$, (8.2.7) again holds true thanks to the definition of J_e. For the case $K = K_e^-$, note that $s_{_{K_e^-}}|_e = -s_{_{K_e^+}}|_e$ and that from (8.2.8),

$$\int_{K_e^-}\hat{f}_e dx = -\int_{K_e^+}\hat{f}_e dx. \tag{8.2.9}$$

Thus the identity in (8.2.7) for $K = K_e^-$ follows from that of $K = K_e^+$, which completes the proof. □

Proposition 8.2.2. *Let $\hat{f}_e$ be defined in* (8.2.1)*. Then*

$$\sum_{e\in\mathcal{E}_K}\int_K \hat{f}_e dx = \int_K f dx, \quad \forall\, K\in\mathcal{T}.$$

Proof. For each $K \in \mathcal{T}$, let v_K be the function defined by

$$v_K := \begin{cases} 1 - \sum_{e \in \mathcal{E}_K} v_e, & \text{in } K, \\ 0, & \text{in } \Omega \backslash K. \end{cases}$$

It can be verified that $v_K \in V_{\mathcal{T},D}$. Inserting v_K into (8.1.3) yields the desired identity and the proof is complete. □

Now we are ready to prove Theorem 8.2.1.

Proof of Theorem 8.2.1. The compatibility condition in (8.2.5) is an immediate consequence of Proposition 8.2.1 and Proposition 8.2.2. Next we prove (8.2.6). Integration by parts and the definition of $\hat{\boldsymbol{\sigma}}_{\mathcal{T}}$ in (8.2.3) imply

$$\int_K p \operatorname{div} \hat{\boldsymbol{\sigma}}_{\mathcal{T}} dx = \sum_{e \in \mathcal{E}_K} \int_e p s_{_K} \hat{g}_e ds - \int_K \hat{\boldsymbol{\sigma}}_{\mathcal{T}} \cdot \nabla p dx = \int_K p f dx, \quad \forall\, p \in P_{k-1}(K).$$

Hence $\operatorname{div} \hat{\boldsymbol{\sigma}}_{\mathcal{T}} = f$. The fact that $\hat{\boldsymbol{\sigma}}_{\mathcal{T}} \cdot \boldsymbol{n} = g_{_N}$ on Γ_N follows directly from the definition. The proof is now complete. □

8.3 A Posteriori Error Estimates

Let $\mathcal{S}$ be the interpolation operator defined in [22] so that $\mathcal{S}(u_{\mathcal{T}}) \in H^1_D(\Omega)$. The error indicator on $K \in \mathcal{T}$ is defined by

$$\xi_K := \left(\|A^{-1/2}(\hat{\boldsymbol{\sigma}}_{\mathcal{T}} - \boldsymbol{\sigma}_{\mathcal{T}})\|_K^2 + \|A^{1/2} \nabla_h (u_{\mathcal{T}} - \mathcal{S}(u_{\mathcal{T}}))\|_K^2 \right)^{1/2}. \tag{8.3.1}$$

The global error estimator is then given by

$$\xi := (\sum_{K \in \mathcal{T}} \xi_K^2)^{1/2}.$$

The quantity

$$\xi_{nc,K} := \|A^{1/2} \nabla_h (u_{\mathcal{T}} - \mathcal{S}(u_{\mathcal{T}}))\|_K \tag{8.3.2}$$

estimates the so-called nonconforming error on K.

8.3.1 Local efficiency

Assume A is scalar, i.e., $A = \alpha I$. We prove the robust local efficiency of ξ_K under the monotonicity assumption [8, Hypothesis 2.7] of α. The monotonicity assumption is only needed in estimating the nonconforming error indicator $\xi_{nc,K}$ as in [22,31]. We will use C with or without subscripts in this section to denote a generic positive constant, possibly different at different occurrences, that is independent of $\max_\Omega \alpha / \min_\Omega \alpha$, but may depend on the shape parameter of $\mathcal{T}$ and the polynomial degree k.

Theorem 8.3.1. *Let u be the exact solution of* (8.1.1) *and $u_\mathcal{T}$ be the nonconforming finite element solution in* (8.1.3). *Then the flux $\hat{\boldsymbol{\sigma}}_\mathcal{T}$ defined in* (8.2.3) *satisfies*

$$\|A^{-1/2}(\hat{\boldsymbol{\sigma}}_\mathcal{T} - \boldsymbol{\sigma}_\mathcal{T})\|_K \leq C\|A^{1/2}\nabla_h(u - u_\mathcal{T})\|_{\hat{\omega}_K}, \tag{8.3.3}$$

where $\hat{\omega}_K$ denotes the union of elements sharing one edge with K. Let ξ_K and $\xi_{nc,K}$ be defined in (8.3.1) *and* (8.3.2), *respectively. Under the monotonicity assumption of α, the following inequalities hold:*

$$\xi_{nc,K} \leq C\|A^{1/2}\nabla_h(u_\mathcal{T} - u)\|_{\omega_K} \tag{8.3.4}$$

and

$$\xi_K \leq C\|A^{1/2}\nabla_h(u_\mathcal{T} - u)\|_{\omega_K}, \tag{8.3.5}$$

where ω_K denotes the union of elements sharing a common vertex with K.

Proof. Note that (8.3.5) is an immediate consequence of (8.3.3) and (8.3.4) according to the definition of ξ_K and meanwhile (8.3.4) is a known result in [22, Lemma 4.2]. So it suffices to prove (8.3.3).

We prove (8.3.3) by bounding $\|A^{-1/2}(\hat{\boldsymbol{\sigma}}_\mathcal{T} - \boldsymbol{\sigma}_\mathcal{T})\|_K$ from above by the standard residual estimator (cf. [8, 9]). Hence the proof is in fact similar to that of Theorem 4.3.1. To this end, for each $e \in \mathcal{E}$, we define α_e by

$$\alpha_e := \max_{K \subseteq \omega_e} \alpha_K$$

and the flux jump by

$$j_e := \begin{cases} (\boldsymbol{\sigma}_\tau|_{K_e^+} - \boldsymbol{\sigma}_\tau|_{K_e^-}) \cdot \boldsymbol{n}_e, & \text{if } e \in \mathcal{E}_I, \\ \boldsymbol{\sigma}_\tau \cdot \boldsymbol{n}_e - g_N, & \text{if } e \in \mathcal{E}_N, \\ 0, & \text{if } e \in \mathcal{E}_D. \end{cases}$$

We first estimate J_e by writing it in terms of element residual $f - \operatorname{div} \boldsymbol{\sigma}_\tau$ and flux jump j_e.

The definition of $\hat{f}_e$ in (8.2.1) and integration by parts imply that

$$\int_{K_e^+} \hat{f}_e dx = \int_{K_e^+} (f - \operatorname{div} \boldsymbol{\sigma}_\tau) v_e dx + \int_e \boldsymbol{\sigma}_\tau|_{K_e^+} \cdot \boldsymbol{n}_e ds, \tag{8.3.6}$$

where we have used the fact that

$$\int_{\partial K_e^+} v_e \boldsymbol{\sigma}_\tau|_{K_e^+} \cdot \boldsymbol{n} ds = \int_e v_e \boldsymbol{\sigma}_\tau|_{K_e^+} \cdot \boldsymbol{n}_e ds = \int_e \boldsymbol{\sigma}_\tau|_{K_e^+} \cdot \boldsymbol{n}_e ds.$$

Inserting (8.3.6) into (8.2.2) yields that for $e \notin \mathcal{E}_N$,

$$J_e = |e|^{-1} \left(\int_{K_e^+} (f - \operatorname{div} \boldsymbol{\sigma}_\tau) v_e dx + \lambda_e \int_e j_e ds \right). \tag{8.3.7}$$

Analogously, thanks to (8.2.9), it can be deduced that

$$J_e = |e|^{-1} \left(\int_{K_e^-} (f - \operatorname{div} \boldsymbol{\sigma}_\tau) v_e dx + (1 - \lambda_e) \int_e j_e ds \right). \tag{8.3.8}$$

If follows from (8.3.7), (8.3.8), (8.2.1), the Cauchy-Schwarz inequality and $|v_e| \leq 1$ that

$$\|J_e\|_e \leq C h_K^{1/2} \|f - \operatorname{div} \boldsymbol{\sigma}_\tau\|_K + C \left(\frac{\alpha_K}{\alpha_e} \right)^{1/2} \|j_e\|_e, \quad K \subseteq \omega_e. \tag{8.3.9}$$

The stability estimate in Lemma 4.3.1 implies that

$$\|\hat{\boldsymbol{\sigma}}_\tau - \boldsymbol{\sigma}_\tau\|_K \leq C \left(h_K \|f - \operatorname{div} \boldsymbol{\sigma}_\tau\|_K + h_K^{1/2} \sum_{e \in \mathcal{E}_K} \|\hat{g}_e - \boldsymbol{\sigma}_\tau|_K \cdot \boldsymbol{n}_e\|_e \right).$$

According to the triangle inequality and (8.2.4), it can be computed that (see, for exmaple, (4.3.7)):

$$\|\hat{g}_e - \boldsymbol{\sigma}_\tau|_K \cdot \boldsymbol{n}_e\|_e \leq \|J_e\|_e + C \left(\frac{\alpha_K}{\alpha_e} \right)^{1/2} \|j_e\|_e. \tag{8.3.10}$$

Combining (8.3.9) – (8.3.10), we arrive at

$$\|A^{-1/2}(\hat{\boldsymbol{\sigma}}_{\mathcal{T}} - \boldsymbol{\sigma}_{\mathcal{T}})\|_K \leq C \left(h_K \alpha_K^{-1/2} \|f - \operatorname{div} \boldsymbol{\sigma}_{\mathcal{T}}\|_K + h_K^{1/2} \sum_{e \in \mathcal{E}_K} \alpha_e^{-1/2} \|j_e\|_e \right).$$

The local efficiency for the residual estimator implies that (cf. [8, 9])

$$h_K \alpha_K^{-1/2} \|f - \operatorname{div} \boldsymbol{\sigma}_{\mathcal{T}}\|_K \leq C \|A^{1/2} \nabla (u - u_{\mathcal{T}})\|_K,$$
$$h_K^{1/2} \alpha_e^{-1/2} \|j_e\|_e \leq C \|A^{1/2} \nabla_h (u - u_{\mathcal{T}})\|_{\omega_e}, \quad e \in \mathcal{E}_K.$$

Hence we conclude that

$$\|A^{-1/2}(\hat{\boldsymbol{\sigma}}_{\mathcal{T}} - \boldsymbol{\sigma}_{\mathcal{T}})\|_K \leq C \|A^{1/2} \nabla_h (u - u_{\mathcal{T}})\|_{\hat{\omega}_K} \leq C \|A^{1/2} \nabla_h (u - u_{\mathcal{T}})\|_{\omega_K},$$

which proves (8.3.3). The local efficiency of ξ_K in (8.3.5) follows immediately from (8.3.3) and (8.3.4).

□

8.3.2 Guaranteed reliability

The guaranteed reliability of ξ follows immediately from the results in Theorem 2.3.10 on the decomposition of vector fields.

Theorem 8.3.2. *Let u be the exact solution of* (8.1.1) *and $u_{\mathcal{T}}$ be the nonconforming finite element solution in* (8.1.3)*. The estimator ξ defined in Section 8.3 satisfies*

$$\|A^{1/2} \nabla_h (u_{\mathcal{T}} - u)\| \leq \xi.$$

Proof. Since $\hat{\boldsymbol{\sigma}}_{\mathcal{T}} \in \Sigma_f$ and $\mathcal{S}(u_{\mathcal{T}}) \in H_D^1(\Omega)$, the inequality follows immediately from (2.3.17). □

8.4 Choice Of v_e

Let $L_k(x) \in P_k([-1, 1])$ denote the Legendre polynomial of order k. For each triangle K, let μ_1, μ_2, μ_3 denote the barycentric coordinates such that $\mu_i = 0$ corresponds to edge $e_i \in \mathcal{E}_K$. Define the polynomial $v_i(x) \in P_k(K)$ by

$$v_i(x) := L_k(1 - 2\mu_i) \quad i = 1, 2, 3. \tag{8.4.1}$$

It is easy to verify that $v_i|_{e_i} = L_k(1) = 1$ and v_i vanishes at k mapped Gauss points on $e_j\,(j \neq i)$. For each $K \subseteq \omega_e$, we define $v_e|_K$ as in (8.4.1) associated with edge $e \in \mathcal{E}_K$ and we set $v_e = 0$ outside ω_e. Hence, $v_e \in V_{\mathcal{T}}$, $|v_e| \leq 1$ in Ω, $v_e = 1$ at Gauss points on e and $v_e = 0$ at Gauss points on $e' \neq e$.

9. EQUILIBRATED ESTIMATOR FOR DISCONTINUOUS GALERKIN APPROXIMATIONS

In this chapter, we consider the same problem as in Chapter 6 and discuss explicit flux recoveries for discontinuous Galerkin(DG) approximations. The flux equilibration for DG elements is significantly easier than conforming elements or nonconforming elements discussed earlier. This is due to the fact that DG finite element spaces impose minimal restriction on test functions. Unlike conforming elements or nonconforming elements in (8.1.2), no requirement is imposed over interelement boundaries.

In [34], an explicit flux equilibration for DG elements was proposed and robust a posteriori error estimates were derived. We provide in the following a slightly different a posteriori error estimator and show that the estimator is more accurate than the one in [34].

9.1 Discontinuous Galerkin Approximation

Let Ω be a bounded polygonal domain in $\mathbb{R}^d$ $(d = 2, 3)$ with Lipschitz boundary $\partial\Omega$, where $\partial\Omega = \bar{\Gamma}_D \cup \bar{\Gamma}_N$ and $\Gamma_D \cap \Gamma_N = \emptyset$. For simplicity, assume that $\text{meas}(\Gamma_D) > 0$. Consider the following diffusion problem:

$$\begin{cases} -\text{div}(A\nabla u) = f, & \text{in } \Omega, \\ u = 0, & \text{on } \Gamma_D, \\ (-A\nabla u)\cdot \boldsymbol{n} = g_N, & \text{on } \Gamma_N, \end{cases} \tag{9.1.1}$$

where for almost all $x \in \Omega$, $A(x)$ is a symmetric, positive definite matrix whose smallest eigenvalue is no less than a positive constant independent of x, $f \in L^2(\Omega)$, $g_N \in L^2(\Gamma_N)$, and $\boldsymbol{n}$ is the unit outward vector normal to Γ_N.

Let the mesh information (elements,vertices,edges/faces,etc.) be given in Section 3.1. The DG finite element space of order k is

$$V_{\mathcal{T}}^{\mathrm{DG}} := \{v \in L^2(\Omega) : v|_K \in P_k(K), \forall K \in \mathcal{T}\}.$$

The DG approximation for (9.1.1) reads: find $u_{\mathcal{T}} \in V_{\mathcal{T}}$ such that

$$\begin{aligned}
&\sum_{K\in\mathcal{T}}\int_K \alpha\nabla u_{\mathcal{T}}\cdot\nabla v dx - \sum_{e\notin\mathcal{E}_N}\int_e \{\{\alpha\nabla_h v\}\}\cdot[\![u_{\mathcal{T}}]\!] + \{\{\alpha\nabla_h u_{\mathcal{T}}\}\}\cdot[\![v]\!]ds + \gamma\sum_{e\notin\mathcal{E}_N} h_e^{-1}\int_e [\![u_{\mathcal{T}}]\!][\![v]\!]ds \\
&= \int_\Omega fv dx + \sum_{e\in\mathcal{E}_D}\int_e g_D(\gamma h_e^{-1}v - \alpha\nabla v\cdot\boldsymbol{n})ds + \int_{\Gamma_N} g_{_N} v ds, \quad \forall\, v \in V_{\mathcal{T}},
\end{aligned} \tag{9.1.2}$$

where γ is a positive constant,

$$[\![q]\!] = q^+\boldsymbol{n}_+ + q^-\boldsymbol{n}_-$$

denotes the vector-valued jump of a scalar q, and

$$\{\{w\}\} = \frac{w^+ + w^-}{2}$$

is the average of a scalar w or a vector w on e. If e is on the boundary, then $\{\{w\}\} = w$ and $[\![q]\!] = q\boldsymbol{n}$.

9.2 Equilibrated Flux Recovery And A Posteriori Error Estimates

In this section, we give explicit formula for the recovered flux $\hat{\boldsymbol{\sigma}}_{\mathcal{T}}$ in terms of degrees of freedom in the $H(\mathrm{div};\Omega)$-conforming Raviart-Thomas spaces. Firstly, for each $K \in \mathcal{T}$, define $\hat{g}_K \in L^2(\partial K)$ by

$$\begin{aligned}
\hat{g}_K|_e &= (-\{\{\alpha\nabla_h u_{\mathcal{T}}\}\} + \gamma h_e^{-1}[\![u_{\mathcal{T}}]\!])\cdot\boldsymbol{n} \quad \text{if } e \in \mathcal{E}_I, \\
\hat{g}_K|_e &= (-\{\{\alpha\nabla_h u_{\mathcal{T}}\}\} + \gamma h_e^{-1}[\![u_{\mathcal{T}}]\!] - \gamma h_e^{-1} g_D)\cdot\boldsymbol{n} \quad \text{if } e \in \mathcal{E}_D, \\
\hat{g}_K|_e &= g_{_N}|_e \quad \text{if } e \in \mathcal{E}_N,
\end{aligned}$$

where $\boldsymbol{n}$ denotes the outward vector normal to ∂K. The above choice of $\hat{g}_K$ coincides with the one in [34]. It can be verified that $\hat{g}_K$ fulfills the equilibrium condition:

$$\int_{\partial K}\hat{g}_K ds = \int_K f dx.$$

Let $\boldsymbol{\sigma}_{\mathcal{T}} := -A\nabla u_{\mathcal{T}}$ be the numerical flux. The degrees of freedom of the recovered flux $\hat{\boldsymbol{\sigma}}_{\mathcal{T}}|_K \in \mathrm{RT}_{k-1}(K)$ is defined by

$$\begin{cases} \hat{\boldsymbol{\sigma}}_{\mathcal{T}} \cdot \boldsymbol{n} = \hat{g}_K, \quad \text{on } \partial K, \\ \displaystyle\int_K \hat{\boldsymbol{\sigma}}_{\mathcal{T}} \cdot \nabla p dx = (\hat{g}_K, p)_{\partial K} - (f, p)_K, \quad \forall\, p \in P_{k-1}(K), \\ \displaystyle\int_K \hat{\boldsymbol{\sigma}}_{\mathcal{T}} \cdot \boldsymbol{q} dx = \int_K \boldsymbol{\sigma}_{\mathcal{T}} \cdot \boldsymbol{q} dx, \quad \forall\, \boldsymbol{q} \in Q_{k-2}(K), \text{ if } k \geq 2, \end{cases} \tag{9.2.1}$$

where $Q_{k-2}(K)$ is defined in (4.2.11)

The flux given by (9.2.1) and the one defined in [34] share the same first two conditions while differ in the third condition in (9.2.1). In [34], the third condition is given by:

$$\int_K \hat{\boldsymbol{\sigma}} \cdot \boldsymbol{q} dx = -\int_K \alpha \nabla u_{\mathcal{T}} \cdot \boldsymbol{q} dx + \frac{1}{2}\alpha_K \sum_{e \in \mathcal{E}_K} \int_e [\![u_{\mathcal{T}}]\!] \cdot \boldsymbol{q} ds, \quad \forall\, \boldsymbol{q} \in Q_{k-2}(K).$$

The error estimator can be defined as in [34] with recovered flux given in (9.2.1). That is, following the notation in [34],

$$\eta_{CF,T} := \|A^{-1/2}(\hat{\boldsymbol{\sigma}}_{\mathcal{T}} - \boldsymbol{\sigma}_{\mathcal{T}})\|_T, \quad \forall T \in \mathcal{T}.$$

It can be easily verified that the proofs in [34] all carry over and we refer to [34] for detailed statements. Particularly, in view of the proof of Theorem 3.6(local efficiency of $\eta_{CF,T}$) in [34], it can be seen that, compared to [34], the choice of flux in (9.2.1) gives a smaller upper bound of $\eta_{CF,T}$ in general. (The contribution from the integral with $\boldsymbol{q} \in Q_{k-2}(K)$ is 0. See the last two lines in the first estimate of $\eta_{CF,T}$ in [34, p.1248].) Thus we expect that the estimator here can be smaller than the one in [34], which implies a better accuracy. This is because both estimators provide guaranteed global upper bound of the true error (due to flux equilibration), i.e.,

$$\|A^{1/2}\nabla(u - u_{\mathcal{T}})\|^2 \leq \sum_{T \in \mathcal{T}} \eta_{CF,T}^2 + \eta_{NC,T}^2,$$

where $\eta_{NC,_T}$ denotes the nonconforming error indicator on T (see [34, Theorem 3.4]). Hence the smaller the estimator is, the more accurate it is.

10. THE QUASI-MONOTONICITY CONDITION

A version of this chapter has been submitted for publication [66].

Consider the diffusion problem with discontinuous coefficients and the conforming finite element discretization. Estimators that are robust with respect to the coefficient jump have drawn more and more attention over the years. The earliest result about the robustness is derived for the residual estimator in [8, 9]. Recently, robust estimates have been investigated for the equilibrated estimator (an estimator based on flux equilibration) [10, 20], the hybrid estimator [12], etc. In the proof of robustness, one key assumption about the distribution of the diffusion coefficient is imposed: the monotonicity assumption [8, Hypothesis 2.7] or the quasi-monotonicity condition [9, Definition 4.1] (see also Definition 7.4.1), where the latter contains a larger class of coefficients. Specifically, the assumption is imposed in the robust global reliability bound of the residual estimator [8, 9], the robust local efficiency bound of the equilibrated estimator [10, 20], the robust global reliability bound of the hybrid estimator [12], etc. Though the assumption is used ubiquitously in obtaining robust estimates, one natural question has not been addressed (let alone answered) so far:

> Is the quasi-monotonicity condition *necessary* for the robustness with respect to coefficient jump?

In this chapter, we present a counter example to show that, for the equilibrated estimator in [10, 20], the quasi-monotonicity condition is indeed *necessary* to ensure robustness.

This chapter is organized as follows. The model problem and the finite element discretization are presented in Section 10.1. Section 10.2 reviews the equilibrated flux recovery and the robust a posteriori error estimates. The main result is stated as Proposition 10.2.1 in Section 10.2. In Section 10.3, a counter example is constructed to

illustrate the loss of robustness when the quasi-monotonicity condition is not satisfied. Section 10.4 is the conclusion.

10.1 Model Problem

For a bounded polygonal domain Ω in $\mathbb{R}^2$, where $\partial\Omega = \bar{\Gamma}_D \cup \bar{\Gamma}_N$ and $\Gamma_D \cap \Gamma_N = \emptyset$, consider the following problem:

$$\begin{aligned} -\mathrm{div}(\alpha\nabla u) &= f \quad \text{in } \Omega, \\ u|_{\Gamma_D} &= 0, \quad \mathrm{measure}(\Gamma_D) > 0, \\ -\alpha\nabla u \cdot \boldsymbol{n}|_{\Gamma_N} &= g_N, \end{aligned} \tag{10.1.1}$$

where $\inf_x \alpha(x) > 0$, $f \in L^2(\Omega), g_N \in L^2(\Gamma_N)$ and $\boldsymbol{n}$ is the unit outward vector normal to Γ_N.

The weak formulation for (10.1.1) is to find $u \in H^1_D(\Omega)$ such that

$$a(u,v) := \int_\Omega \alpha\nabla u \cdot \nabla v dx = \int_\Omega f v dx - \int_{\Gamma_N} g_N v ds, \quad \forall v \in H^1_D(\Omega).$$

Let the mesh information (elements,vertices,edges/faces,etc.) be given in Section 3.1. We define the conforming piecewise linear finite element space

$$V_{\mathcal{T}} = \{v \in H^1_D(\Omega) : v|_K \text{ is linear}, \forall\, K \in \mathcal{T}\}.$$

The finite element solution $u_{\mathcal{T}} \in V_{\mathcal{T}}$ satisfies

$$a(u_{\mathcal{T}}, v) = \int_\Omega f v dx - \int_{\Gamma_N} g_N v ds, \quad \forall\, v \in V_{\mathcal{T}}.$$

Assume $\alpha|_K = \alpha_K$ is constant for $K \in \mathcal{T}$. Furthermore, we ignore data oscillation by assuming that $f|_K$, $g_N|_e$ are constant for $K \in \mathcal{T}$, $e \in \mathcal{E}_N$.

10.2 Review Of Equilibrated Flux Recovery And Robust Estimates

In this section, we review the equilibrated flux recovery via local minimization. A detailed discussion of the approach is presented in Chapter 6. Let ϕ_z be the Lagrange

nodal basis function associated with vertex z. The vertex patch associated with z is $\omega_z := \text{supp}\ \phi_z$. Let $\mathcal{E}_z$ be the subset of boundary edges on ω_z where ϕ_z vanishes as defined in (6.2.6). The patch-based equilibrated flux recovery (cf. [10, 15–17, 19, 20]) computes $\hat{\boldsymbol{\sigma}}_{\mathcal{T}} = \sum_{z\in\mathcal{N}} \hat{\boldsymbol{\sigma}}_{z,\mathcal{T}}$. The local flux $\hat{\boldsymbol{\sigma}}_{z,\mathcal{T}}$ over ω_z must satisfy

$$\begin{cases} \operatorname{div} \hat{\boldsymbol{\sigma}}_{z,\mathcal{T}} = \bar{f}_z, & \\ \hat{\boldsymbol{\sigma}}_{z,\mathcal{T}} \cdot \boldsymbol{n}_e = \frac{1}{|e|}\int_e \phi_z g_{_N} ds, & \text{if } e \in \mathcal{E}_N, \\ \hat{\boldsymbol{\sigma}}_{z,\mathcal{T}} \cdot \boldsymbol{n}_e = 0, & \text{if } e \in \mathcal{E}_z, \end{cases} \tag{10.2.1}$$

where

$$\bar{f}_z := \frac{1}{|K|}\int_K \nabla\phi_z \cdot (-\alpha\nabla u_{_{\mathcal{T}}}) + \phi_z f dx$$

and $\boldsymbol{n}_e$ is the unit outward normal if $e \subset \partial\Omega$. In [10, 17–20], the recovered flux $\hat{\boldsymbol{\sigma}}_{\mathcal{T}}$ is sought in $H(\text{div};\Omega)$-conforming Raviart-Thomas spaces. Define the set of admissible fluxes as

$$\Sigma_z := \{\boldsymbol{\tau} \in \text{RT}_0 : \text{supp}\boldsymbol{\tau} \subseteq \omega_z, \boldsymbol{\tau} \text{ satisfy (10.2.1)}\}.$$

In [16, 17, 19], the recovered flux was chosen as an arbitrary equilibrated flux. The resulting error estimator, however, is not robust with respect to the discontinuity of the diffusion coefficient (cf. [10]). In [10, 20] (see also Chapter 6), $\hat{\boldsymbol{\sigma}}_{z,\mathcal{T}}$ is defined as the solution to the minimization problem:

$$\hat{\boldsymbol{\sigma}}_{z,\mathcal{T}} := \arg\min_{\boldsymbol{\tau}\in\Sigma_z} \|\alpha^{-1/2}(\boldsymbol{\tau} - \boldsymbol{\sigma}_{z,\mathcal{T}})\|_{\omega_z}, \tag{10.2.2}$$

where $\boldsymbol{\sigma}_{z,\mathcal{T}}|_K = \Pi_{\text{RT}_0(K)}(-\phi_z\alpha\nabla u_{_{\mathcal{T}}}), \quad \forall\, K \in \mathcal{T}$, and $\Pi_{\text{RT}_0(K)}$ denotes the interpolation from $H(\text{div}; K)$ to $\text{RT}_0(K)$. It is proved in [10, 20] (see also Chapter 6) that the resulting error estimator is robust under the quasi-monotonicity condition of the diffusion coefficient.

The recovered global flux

$$\hat{\boldsymbol{\sigma}}_{\mathcal{T}} = \sum_{z\in\mathcal{N}} \hat{\boldsymbol{\sigma}}_{z,\mathcal{T}}$$

satisfies the equilibrium equation $\operatorname{div} \hat{\boldsymbol{\sigma}}_{\mathcal{T}} = f$ in Ω and the boundary condition $\hat{\boldsymbol{\sigma}}_{\mathcal{T}} \cdot \boldsymbol{n} = g_{_N}$ on Γ_N. Hence $\hat{\boldsymbol{\sigma}}_{\mathcal{T}}$ (or this type of flux recovery) is called equilibrated.

Define the global error estimator by

$$\xi := \|\alpha^{-1/2}\hat{\boldsymbol{\sigma}}_{\mathcal{T}} + \alpha^{1/2}\nabla u_{\mathcal{T}}\|.$$

Since $\hat{\boldsymbol{\sigma}}_{\mathcal{T}}$ is equilibrated, according to the Prager-Synge theorem (cf. [17, 29]), the following inequality holds

$$\|\alpha^{1/2}\nabla(u - u_{\mathcal{T}})\| \leq \xi,$$

which gives the guaranteed reliability of ξ.

The patch-based and element-based local error indicators are defined as

$$\xi_z = \|\alpha^{-\frac{1}{2}}\left(\hat{\boldsymbol{\sigma}}_{z,\mathcal{T}} - \boldsymbol{\sigma}_{z,\mathcal{T}}\right)\| \quad \text{and} \quad \xi_K = \|\alpha^{-\frac{1}{2}}\left(\hat{\boldsymbol{\sigma}}_{\mathcal{T}} + \alpha\nabla u_{\mathcal{T}}\right)\|_K,$$

respectively.

For problems with discontinuous diffusion coefficient, a posteriori error estimates that are robust with respect to the coefficient jump are of particular importance. For robust residual estimators, the robust reliability bound is proved under either the monotonicity assumption [8, Hypothesis 2.7] or the quasi-monotonicity condition [9, Definition 4.1] about the distribution of α. The quasi-monotonicity condition admits a larger class of coefficients. It is not known, however, if the quasi-monotonicity condition is necessary for the robustness of residual estimators.

For estimators based on equilibrated flux recovery, the reliability bound is automatically robust with constant being exactly 1. In terms of the local efficiency bound, to the best of our knowledge, the estimator in [10, 20] via patch-based minimizations is the only one with proved robustness, and the proofs require the quasi-monotonicity condition. The result on robust local efficiency is stated in Theorem 6.5.1 and is cited below.

Theorem 10.2.1. *Let* $\alpha_{\max} := \max_{K\in\mathcal{T}} \alpha_{K}$ *and* $\alpha_{\min} := \min_{K\in\mathcal{T}} \alpha_{K}$. *Assume that* α *is quasi-monotone. The following estimates hold with constants* C_1 *and* C_2 *that only depend on the shape parameter of* $\mathcal{T}$ *and are independent of* $\alpha_{\max}/\alpha_{\min}$.

$$\xi_z \leq C_1\|\alpha^{1/2}\nabla(u - u_{\mathcal{T}})\|_{\hat{\omega}_z},$$
$$\xi_K \leq C_2 \sum_{z\in\mathcal{N}\cap\partial K} \|\alpha^{1/2}\nabla(u - u_{\mathcal{T}})\|_{\hat{\omega}_z},$$

where $\hat{\omega}_z$ denotes the union of elements that share at least one edge $(d = 2)$ or one face $(d = 3)$ with an element in ω_z.

We show in the next section that the quasi-monotonicity condition is necessary for the robust estimates in Theorem 10.2.1. That is, we will justify the following claim.

Proposition 10.2.1. *If α is* not *quasi-monotone, then there does* not *exist a constant C independent of $\alpha_{\max}/\alpha_{\min}$ such that*

$$\xi_z \leq C\|\alpha^{1/2}\nabla(u - u_{\mathcal{T}})\| \quad or \quad \xi_K \leq C\|\alpha^{1/2}\nabla(u - u_{\mathcal{T}})\| \tag{10.2.3}$$

for all $z \in \mathcal{N}$ and $K \in \mathcal{T}$. Consequently, the quasi-monotonicity condition is necessary for the robust local efficiency stated in Theorem 10.2.1.

Remark 10.2.1. *In [15–17, 19], the recovered flux was chosen as an arbitrary equilibrated flux. The resulting error estimator, however, is not robust with respect to the discontinuity of the diffusion coefficient (cf. [10]). [10] introduced the minimization in* (10.2.2) *to derive a robust estimator.*

10.3 A Counter Example

We first derive in Section 10.3.1 an explicit formula for the local error flux $\boldsymbol{\sigma}_z^{\Delta}$ associated with a Neumann boundary patch in a specific example. Then a counter example is constructed in Section 10.3.2 to demonstrate that the local efficiency bound is not robust without the quasi-monotonicity condition on diffusion coefficients.

10.3.1 Representation of error flux

We consider an example with $f = 0$ in Ω and focus on a Neumann boundary vertex patch with three elements. The goal of this subsection is to derive an explicit expression for the local error below

$$\boldsymbol{\sigma}_z^{\Delta} = \hat{\boldsymbol{\sigma}}_{z,\mathcal{T}} - \boldsymbol{\sigma}_{z,\mathcal{T}}.$$

We use $\boldsymbol{\psi}_{K,e}$ to denote the basis function for $\mathrm{RT}_0(K)$ associated with edge e. That is, $\boldsymbol{\psi}_{K,e} \cdot \boldsymbol{n}_{e'} = \delta_{ee'}$, $\forall\, e, e' \in \mathcal{E}_K$, where $\boldsymbol{n}_{e'}$ denotes the unit outward vector for K normal to edge e'.

Consider a vertex patch ω_z composed of three elements K_1, K_2, K_3, where $z \in \Gamma_N$. The notation is illustrated in Figure 10.1.

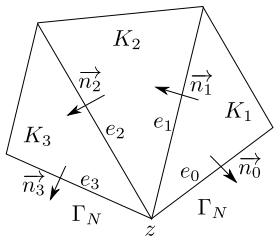

Figure 10.1. A general Neumann boundary vertex patch ω_z with 3 elements

In the following, we first determine the expression for $\hat{\boldsymbol{\sigma}}_{z,\mathcal{T}} \in \mathrm{RT}_0(K)$. It suffices to know the normal component on each edge, and in this case, only two normal components (on $e1$ and $e2$, respectively) are unknown: τ_1 and τ_2, where $\tau_i = \hat{\boldsymbol{\sigma}}_{z,\mathcal{T}} \cdot \boldsymbol{n}_i$. According to the condition in (10.2.1) and divergence theorem, τ_1 and τ_2 must satisfy

$$\begin{bmatrix} |e_1| & & \\ -|e_1| & |e_2| & \\ & -|e_2| & \end{bmatrix} \begin{bmatrix} \tau_1 \\ \tau_2 \end{bmatrix} = \begin{bmatrix} b_1 - |e_0|\tau_0 \\ b_2 \\ b_3 - |e_3|\tau_3 \end{bmatrix},$$

where

$$\begin{aligned} b_i &:= \int_{K_i} \nabla \phi_z \cdot (-\alpha \nabla u_\tau)\, dx = \int_{\partial K_i} \phi_z (-\alpha \nabla u_\tau|_{K_i}) \cdot \boldsymbol{n} ds \quad (i = 1, 2, 3), \\ \tau_i &:= \frac{1}{|e_i|} \int_{e_i} \phi_z g_N ds = \frac{1}{2} g_N|_{e_i} \quad (i = 0, 3). \end{aligned} \tag{10.3.1}$$

The solution is uniquely given by

$$\begin{aligned} \tau_1 &= |e_1|^{-1} b_1 - \frac{|e_0|}{2|e_1|} g_N|_{e_0} = 1/2(-\alpha \nabla u_\tau|_{K_1}) \cdot \boldsymbol{n}_1 + \frac{|e_0|}{2|e_1|} j_{e_0}, \\ \tau_2 &= -|e_2|^{-1} b_3 + \frac{|e_3|}{2|e_2|} g_N|_{e_3} = 1/2(-\alpha \nabla u_\tau|_{K_3}) \cdot \boldsymbol{n}_2 + \frac{|e_3|}{2|e_2|} j_{e_3}, \end{aligned}$$

where j_e denotes the flux jump of $-\alpha\nabla u_{\mathcal{T}}$ across edge e:

$$\begin{aligned} j_{e_0} &:= -\alpha\nabla u_{\mathcal{T}}|_{K_1}\cdot\boldsymbol{n}_0 - g_{_N}|_{e_0},\\ j_{e_i} &:= -\alpha\nabla u_{\mathcal{T}}|_{K_{i+1}}\cdot\boldsymbol{n}_i + \alpha\nabla u_{\mathcal{T}}|_{K_i}\cdot\boldsymbol{n}_i \quad (i=1,2),\\ j_{e_3} &:= g_{_N}|_{e_3} + \alpha\nabla u_{\mathcal{T}}|_{K_3}\cdot\boldsymbol{n}_3. \end{aligned}$$

Therefore, Σ_z only contains one element and consequently $\hat{\boldsymbol{\sigma}}_{_{z,\mathcal{T}}}$ is given by

$$\begin{aligned} \hat{\boldsymbol{\sigma}}_{_{z,\mathcal{T}}}|_{K_1} &= \tau_1\boldsymbol{\psi}_{K_1,e_1} + 1/2g_{_N}|_{e_0}\boldsymbol{\psi}_{K_1,e_0},\\ \hat{\boldsymbol{\sigma}}_{_{z,\mathcal{T}}}|_{K_2} &= -\tau_1\boldsymbol{\psi}_{K_2,e_1} + \tau_2\boldsymbol{\psi}_{K_2,e_2},\\ \hat{\boldsymbol{\sigma}}_{_{z,\mathcal{T}}}|_{K_3} &= -\tau_2\boldsymbol{\psi}_{K_3,e_2} + 1/2g_{_N}|_{e_3}\boldsymbol{\psi}_{K_3,e_3}, \end{aligned}$$

with τ_1,τ_2 in (10.3.1). Then one can derive $\boldsymbol{\sigma}_z^{\Delta} = \hat{\boldsymbol{\sigma}}_{_{z,\mathcal{T}}} - \boldsymbol{\sigma}_{_{z,\mathcal{T}}}$ as below:

$$\begin{aligned} \boldsymbol{\sigma}_z^{\Delta}|_{K_1} &= -\frac{1}{2}j_{e_0}\boldsymbol{\psi}_{K_1,e_0} + \frac{|e_0|}{2|e_1|}j_{e_0}\boldsymbol{\psi}_{K_1,e_1},\\ \boldsymbol{\sigma}_z^{\Delta}|_{K_2} &= \left(\frac{1}{2}j_{e_1} - \frac{|e_0|}{2|e_1|}j_{e_0}\right)\boldsymbol{\psi}_{K_2,e_1} + \left(\frac{1}{2}j_{e_2} + \frac{|e_3|}{2|e_2|}j_{e_3}\right)\boldsymbol{\psi}_{K_2,e_2},\\ \boldsymbol{\sigma}_z^{\Delta}|_{K_3} &= -\frac{|e_3|}{2|e_2|}j_{e_3}\boldsymbol{\psi}_{K_3,e_2} + \frac{1}{2}j_{e_3}\boldsymbol{\psi}_{K_3,e_3}. \end{aligned} \tag{10.3.2}$$

10.3.2 The counter example

The geometry of domain Ω is shown in Figure 10.2, where Ω contains three equilateral right-angled triangles K_1, K_2, K_3 with $|e_0| = |e_2| = 1, |e_1| = |e_3| = \sqrt{2}$. The notation here follows the previous section and $\Gamma_{_N} = e_0 \cup e_3$. Consider the model problem in (10.1.1) with

$$f = 0, \quad g_{_N}|_{e_0} = 0, \quad g_{_N}|_{e_3} = 1$$

and

$$\alpha|_{K_1} = \alpha_1, \quad \alpha|_{K_2} = \alpha_2, \quad \alpha|_{K_3} = \alpha_3.$$

Note that for ω_z illustrated in Figure 10.2, the quasi-monotonicity condition is *not* satisfied if $\alpha_1 = \alpha_3 > \alpha_2$.

The coordinate system is chosen as follows: the origin is at $e_0 \cap e_2$, e_0 is on the positive x-axis and e_2 is on the positive y-axis. Assume the triangulation of Ω is given

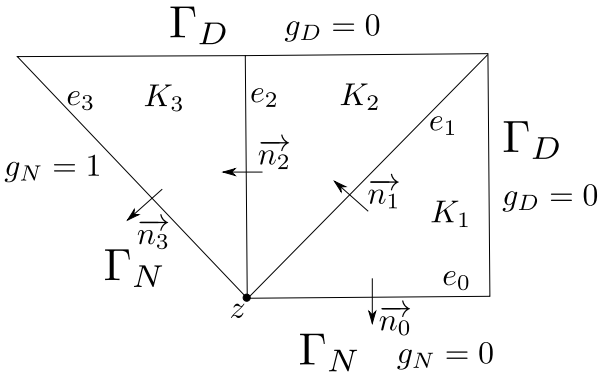

Figure 10.2. Domain Ω and patch ω_z with $z = e_0 \cap e_2$

by $\mathcal{T} = \{K_1, K_2, K_3\}$. We next characterize the finite element space $V_{\mathcal{T}}$. It is not hard to verify that $V_{\mathcal{T}}$ is a one-dimensional space spanned by

$$v(x,y) = \begin{cases} (x-1), & \text{if } (x,y) \in K_1, \\ (y-1), & \text{if } (x,y) \in K_2 \cup K_3. \end{cases}$$

Now we compute the finite element solution $u_{\mathcal{T}} = cv$ for some constant c. The weak formulation for $u_{\mathcal{T}}$ implies

$$\int_{K_1} \alpha_1 c dx + \int_{K_2} \alpha_2 c dx + \int_{K_3} \alpha_3 c dx = -\int_{e_0} g_{_N}(x-1)ds - \int_{e_3} g_{_N}(y-1)ds.$$

Then

$$c = \sqrt{2}(\alpha_1 + \alpha_2 + \alpha_3)^{-1},$$

and

$$u_{\mathcal{T}}(x,y) = \begin{cases} \sqrt{2}(\alpha_1 + \alpha_2 + \alpha_3)^{-1}(x-1), & \text{if } (x,y) \in K_1, \\ \sqrt{2}(\alpha_1 + \alpha_2 + \alpha_3)^{-1}(y-1), & \text{if } (x,y) \in K_2 \cup K_3. \end{cases} \tag{10.3.3}$$

It follows that

$$j_{e_0} = 0, \quad j_{e_1} = -\frac{\alpha_1 + \alpha_2}{\alpha_1 + \alpha_2 + \alpha_3}, \quad j_{e_2} = 0, \quad j_{e_3} = \frac{\alpha_1 + \alpha_2}{\alpha_1 + \alpha_2 + \alpha_3}. \tag{10.3.4}$$

Therefore,

$$|j_e| \leq 1, \quad \forall\, e \in \mathcal{E}, \tag{10.3.5}$$

where $j_e := 0$ for $e \subseteq \Gamma_D$.

Patch-based local error indicator. We show next that for the vertex patch illustrated in Figure 10.2, the local efficiency bound of $\xi_z = \|\alpha^{-\frac{1}{2}} \boldsymbol{\sigma}_z^{\Delta}\|$ is not robust with respect to the discontinuity of α. It suffices to show that the ratio of ξ_z over the global true error $\|\alpha^{1/2}\nabla(u_{\mathcal{T}} - u)\|$ tends to infinity as coefficient jump tends to infinity.

First we estimate ξ_z. According to (10.3.2),

$$\begin{aligned}
&\|\alpha_1^{-\frac{1}{2}}\boldsymbol{\sigma}_z^{\Delta}\|_{K_1} = 0,\\
&\|\alpha_2^{-\frac{1}{2}}\boldsymbol{\sigma}_z^{\Delta}\|_{K_2} = \alpha_2^{-\frac{1}{2}}\|\frac{1}{2}j_{e_1}\boldsymbol{\psi}_{K_2,e_1} + \frac{\sqrt{2}}{2}j_{e_3}\boldsymbol{\psi}_{K_2,e_2}\|_{K_2},\\
&\|\alpha_3^{-\frac{1}{2}}\boldsymbol{\sigma}_z^{\Delta}\|_{K_3} = \alpha_3^{-\frac{1}{2}}\|-\frac{\sqrt{2}}{2}j_{e_3}\boldsymbol{\psi}_{K_3,e_2} + \frac{1}{2}j_{e_3}\boldsymbol{\psi}_{K_3,e_3}\|_{K_3}.
\end{aligned}$$

It follows from (10.3.4) that

$$j_{e_1} \to -\frac{1}{2}, \quad j_{e_3} \to \frac{1}{2}, \quad \text{if } \alpha_1 = \alpha_3 \to +\infty \text{ and } \alpha_2 = 1,$$

and consequently

$$\|\alpha_1^{-\frac{1}{2}}\boldsymbol{\sigma}_z^{\Delta}\|_{K_1} = 0, \quad \|\alpha_2^{-\frac{1}{2}}\boldsymbol{\sigma}_z^{\Delta}\|_{K_2} \to C_*\alpha_2^{-\frac{1}{2}}, \quad \|\alpha_3^{-\frac{1}{2}}\boldsymbol{\sigma}_z^{\Delta}\|_{K_3} \to= 0, \tag{10.3.6}$$

where

$$C_* = \|-\frac{1}{4}\boldsymbol{\psi}_{K_2,e_1} + \frac{\sqrt{2}}{4}\boldsymbol{\psi}_{K_2,e_2}\|_{K_2} > 0. \tag{10.3.7}$$

Therefore,

$$\xi_z \to C_* \quad \text{if} \quad \alpha_1 = \alpha_3 \to \infty \text{ and } \alpha_2 = 1. \tag{10.3.8}$$

We show next that the true error tends to zero. The representation of the residual reads (cf. [8, 9, 38]):

$$\int_\Omega \alpha\nabla(u_\tau - u)\cdot\nabla v dx = \sum_{e\in\mathcal{E}}\int_e j_e v ds - \sum_{K\in\mathcal{T}}\int_K f v dx, \quad \forall\, v \in H_D^1(\Omega).$$

As $f = 0$, it follows that

$$E^2 = \|\alpha^{1/2}\nabla(u_\tau - u)\|^2 = \sum_{e\in\mathcal{E}}\int_e j_e(u_\tau - u)ds.$$

Since measure$(\partial K_i \cap \Gamma_D) > 0$ $(i = 1, 2, 3)$, Poincaré's inequality can be applied in each element K_i $(i = 1, 2, 3)$. Trace theorem (cf. [8]), Cauchy-Schwarz inequality, and Poincaré's inequality imply

$$\begin{aligned} E^2 &\leq \sum_{e\in\mathcal{E}} \|u - u_\tau\|_e \|j_e\|_e \\ &\leq \sum_{e\in\mathcal{E}} \left(h_e^{-\frac{1}{2}} \|u - u_\tau\|_{K_e} + h_e^{\frac{1}{2}} \|\nabla(u - u_\tau)\|_{K_e} \right) \|j_e\|_e \\ &\leq \sum_{e\in\mathcal{E}} C_1 \|\nabla(u - u_\tau)\|_{K_e} \|j_e\|_e \\ &\leq C_2 \left(\sum_{K\in\mathcal{T}} \alpha_K \|\nabla(u_\tau - u)\|_K^2 \right)^{\frac{1}{2}} \left(\sum_{e\in\mathcal{E}} \alpha_{K_e}^{-1} \|j_e\|_e^2 \right)^{\frac{1}{2}} = C_2 E \left(\sum_{e\in\mathcal{E}} \alpha_{K_e}^{-1} \|j_e\|_e^2 \right)^{\frac{1}{2}}, \end{aligned}$$

where C_1, C_2 are constants independent of α, u, u_τ, and K_e can be chosen arbitrarily as long as $e \subset \partial K_e$. We choose K_e as below

$$K_{e_0} = K_{e_1} = K_1, \quad K_{e_2} = K_{e_3} = K_3.$$

Then (recall that $j_e = 0$ for $e \subset \Gamma_D$)

$$E \leq C \left(\alpha_1^{-1} \|j_e\|_{e_0}^2 + \alpha_1^{-1} \|j_e\|_{e_1}^2 + \alpha_3^{-1} \|j_e\|_{e_2}^2 + \alpha_3^{-1} \|j_e\|_{e_3}^2 \right)^{\frac{1}{2}},$$

which is independent of α_2. Since $|j_e| \leq 1$ (cf. (10.3.5)), it follows that

$$E \to 0, \quad \text{if } \alpha_1 = \alpha_3 \to +\infty, \tag{10.3.9}$$

independent of α_2. Together with (10.3.8), we obtain the following

$$\xi_z / E \to \infty \quad \text{if } \alpha_1 = \alpha_3 \to \infty \text{ and } \alpha_2 = 1$$

Hence there does not exist a constant C independent α such that $\xi_z \leq CE$. We conclude that ξ_z in general can not have robust local efficiency if α is not quasi-monotone.

Element-based error indicator. We show in the following that the error indicator ξ_{K_2} on element K_2 is not robust with respect to α. That is, the second inequality in (10.2.3) fails to hold for $K = K_2$.

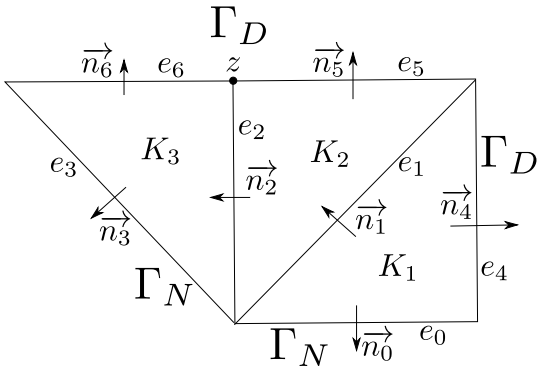

Figure 10.3. $z = \Gamma_D \cap \partial K_2 \cap \partial K_3$

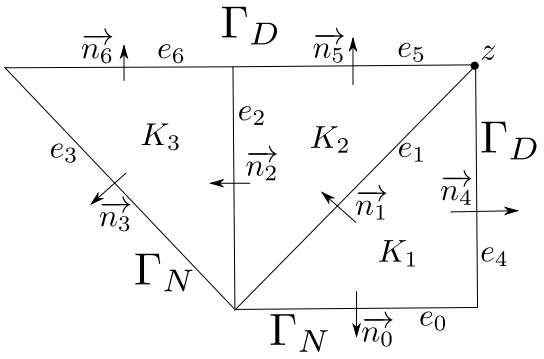

Figure 10.4. $z = \Gamma_D \cap \partial K_1 \cap \partial K_2$

If z is the vertex on Γ_N as illustrated in Figure 10.2, then this case has been considered in the section before.

If z is the vertex at the right angle in K_3 as illustrated in Figure 10.3, then ω_z contains only two elements: K_2 and K_3.

In this case, from the expression of $u_{\mathcal{T}}$ in (10.3.3), it can be computed that

$$\boldsymbol{\sigma}_{z,\mathcal{T}}|_{K_2} = -1/2\alpha\nabla u_{\mathcal{T}}|_{K_2}\cdot\boldsymbol{n}_5\boldsymbol{\psi}_{K_2,e_5} - 1/2\alpha\nabla u_{\mathcal{T}}|_{K_2}\cdot\boldsymbol{n}_2\boldsymbol{\psi}_{K_2,e_2} = \frac{-\sqrt{2}\alpha_2}{2(\alpha_1+\alpha_2+\alpha_3)}\boldsymbol{\psi}_{K_2,e_5},$$
$$\boldsymbol{\sigma}_{z,\mathcal{T}}|_{K_3} = -1/2\alpha\nabla u_{\mathcal{T}}|_{K_3}\cdot\boldsymbol{n}_6\boldsymbol{\psi}_{K_3,e_6} + 1/2\alpha\nabla u_{\mathcal{T}}|_{K_3}\cdot\boldsymbol{n}_2\boldsymbol{\psi}_{K_3,e_2} = \frac{-\sqrt{2}\alpha_3}{2(\alpha_1+\alpha_2+\alpha_3)}\boldsymbol{\psi}_{K_3,e_6}.$$

Thus $\boldsymbol{\sigma}_{z,\mathcal{T}}\in \mathrm{RT}_0$ as the normal component is continuous across e_2 (in fact, equal to 0). Note that $\boldsymbol{\sigma}_{z,\mathcal{T}}$ satisfies the conditions in (10.2.1) Hence $\boldsymbol{\sigma}_{z,\mathcal{T}}\in\Sigma_z$ and it follows that

$$\boldsymbol{\sigma}_z^{\Delta} = 0. \tag{10.3.10}$$

If z is the vertex at the top-right corner in Ω as illustrated in Figure 10.4, then ω_z contains only two elements: K_1 and K_2.

In this case, according to $u_{\mathcal{T}}$ in (10.3.3), we compute that

$$\begin{aligned}
\boldsymbol{\sigma}_{z,\mathcal{T}}|_{K_1} &= -1/2\alpha\nabla u_{\mathcal{T}}|_{K_1}\cdot\boldsymbol{n}_4\boldsymbol{\psi}_{K_1,e_4} - 1/2\alpha\nabla u_{\mathcal{T}}|_{K_1}\cdot\boldsymbol{n}_1\boldsymbol{\psi}_{K_1,e_1}\\
&= \frac{-\alpha_1}{\alpha_1+\alpha_2+\alpha_3}\left(\frac{\sqrt{2}}{2}\boldsymbol{\psi}_{K_1,e_4} - 1/2\boldsymbol{\psi}_{K_1,e_1}\right),\\
\boldsymbol{\sigma}_{z,\mathcal{T}}|_{K_2} &= -1/2\alpha\nabla u_{\mathcal{T}}|_{K_2}\cdot\boldsymbol{n}_5\boldsymbol{\psi}_{K_2,e_5} + 1/2\alpha\nabla u_{\mathcal{T}}|_{K_2}\cdot\boldsymbol{n}_1\boldsymbol{\psi}_{K_2,e_1}\\
&= \frac{-\alpha_2}{\alpha_1+\alpha_2+\alpha_3}\left(\frac{\sqrt{2}}{2}\boldsymbol{\psi}_{K_2,e_5} - 1/2\boldsymbol{\psi}_{K_2,e_1}\right),
\end{aligned}$$

and

$$\bar{f}_z = \nabla\phi_z\cdot(-\alpha\nabla u_{\mathcal{T}}) = 0 \quad \text{in } \Omega.$$

Hence $\boldsymbol{\tau} = 0\in\Sigma_z$. Next we compute the minimizer $\hat{\boldsymbol{\sigma}}_{z,\mathcal{T}}$ using the formula below (cf. [10])

$$\hat{\boldsymbol{\sigma}}_{z,\mathcal{T}} = \boldsymbol{\tau} - \frac{\left(\alpha^{-1}(\boldsymbol{\tau}-\boldsymbol{\sigma}_{z,\mathcal{T}}),\nabla^{\perp}\phi_z\right)_{\omega_z}}{(\alpha^{-1}\nabla^{\perp}\phi_z,\nabla^{\perp}\phi_z)_{\omega_z}}\nabla^{\perp}\phi_z,$$

where $\nabla^{\perp}v := (-\frac{\partial v}{\partial y},\frac{\partial v}{\partial x})$.

It is easily seen that

$$\nabla^{\perp}\phi_z|_{K_1} = (-1,0) \quad \text{and} \quad \nabla^{\perp}\phi_z|_{K_2} = (0,1).$$

$$\begin{aligned}
&\left(-\alpha^{-1}\boldsymbol{\sigma}_{z,\mathcal{T}}, \nabla^{\perp}\phi_z\right)_{K_1} \\
&= -|K_1|\boldsymbol{\sigma}_{z,\mathcal{T}}|_{K_1}(x^*)\cdot(-1,0) \\
&= |K_1|(\alpha_1+\alpha_2+\alpha_3)^{-1}(\frac{\sqrt{2}}{2}\boldsymbol{\psi}_{K_1,e_4}(x^*) - 1/2\boldsymbol{\psi}_{K_1,e_1}(x^*))\cdot(-1,0) \\
&= (\alpha_1+\alpha_2+\alpha_3)^{-1}\left(\frac{\sqrt{2}}{2}\frac{|e_4|}{2}(\frac{1}{3}(1,1)+\frac{1}{3}(1,0)) - 1/2\frac{|e_1|}{2}(\frac{1}{3}(0,1)+\frac{1}{3}(-1,0))\right)\cdot(-1,0) \\
&= \frac{-\sqrt{2}}{4(\alpha_1+\alpha_2+\alpha_3)}, \\
&\left(-\alpha^{-1}\boldsymbol{\sigma}_{z,\mathcal{T}}, \nabla^{\perp}\phi_z\right)_{K_2} \\
&= -|K_2|\boldsymbol{\sigma}_{z,\mathcal{T}}|_{K_2}(x^*)\cdot(0,1) \\
&= (\alpha_1+\alpha_2+\alpha_3)^{-1}\left(\frac{\sqrt{2}}{2}\frac{|e_5|}{2}(\frac{1}{3}(0,1)+\frac{1}{3}(1,1)) - 1/2\frac{|e_1|}{2}(\frac{1}{3}(0,-1)+\frac{1}{3}(1,0))\right)\cdot(0,1) \\
&= \frac{\sqrt{2}}{4(\alpha_1+\alpha_2+\alpha_3)},
\end{aligned}$$

where x^* denotes the centroid of a triangle. Hence

$$\left(\alpha^{-1}(\boldsymbol{\tau}-\boldsymbol{\sigma}_{z,\mathcal{T}}), \nabla^{\perp}\phi_z\right)_{\omega_z} = 0.$$

Consequently, $\hat{\boldsymbol{\sigma}}_{z,\mathcal{T}} = 0$ and for $\alpha_1 = \alpha_3 \to \infty,\ \alpha_2 = 1$,

$$\alpha_2^{-1/2}\boldsymbol{\sigma}_z^{\Delta}|_{K_2} = \frac{\sqrt{\alpha_2}}{\alpha_1+\alpha_2+\alpha_3}\left(\frac{\sqrt{2}}{2}\boldsymbol{\psi}_{K_2,e_5} - 1/2\boldsymbol{\psi}_{K_2,e_1}\right) \to 0. \tag{10.3.11}$$

Collecting results in (10.3.6), (10.3.10), (10.3.11), we conclude that

$$\xi_{K_2} = \|\alpha^{-1/2}\sum_{z\in\partial K_2\cap\mathcal{N}}\boldsymbol{\sigma}_z^{\Delta}\|_{K_2} \to C_*, \quad \text{if } \alpha_1=\alpha_3\to\infty \text{ and } \alpha_2 = 1,$$

where C_* is the positive constant in (10.3.7). However, (10.3.9) states that

$$E \to 0 \quad \text{if } \alpha_1=\alpha_3\to\infty \text{ and } \alpha_2 = 1.$$

Hence there does not exist a constant C independent of α such that

$$\xi_{K_2} \leq CE.$$

We conclude that ξ in general does not have robust local efficiency if α is not quasi-monotone. Proposition 10.2.1 is now justified.

10.4 Conclusion

We considered diffusion problems with discontinuous coefficient and investigated the role of the quasi-monotonicity condition for robust estimates with respect to coefficient jump. Existing robust estimates are all proved under certain monotonicity assumption on the diffusion coefficient (cf. [8–10, 12, 20]) but it is not known if the assumption is necessary to ensure the robustness. In this chapter, we provide a counter example to demonstrate that, without the quasi-monotonicity condition, the local efficiency bound of the equilibrated estimator in [10,20] is not robust with respect to the coefficient jump. This then implies that the quasi-monotonicity condition is necessary for the robustness of the equilibrated estimator in [10, 20]. One future work is to investigate if this is also the case for the residual estimator [8, 9].

11. MISCONCEPTIONS AND COUNTER-EXAMPLES

Counterexamples or misconceptions

1. That j_e dominate (see [67]) does not mean total element residual is a higher order term or the edge estimator is always larger than the total element residual! This can be seen from Table 8.1 in [67] if we compute the total element residual $\sum_T \eta_T^2$ and compare all three quantities. Numerical examples are also given in the *fresolve* folder.

2. no reliability for recovery-based estimator if residual is not incorporated in estimator, e.g., ZZ, ZZ-variants, edgeEstimator, caizhang2009, etc. Therefore, convergence of AFEM is not guaranteed with any of those estimators..

3. Quasi-monotonicity is necessary for the robust local efficiency of estimator proposed in caizhang2012 and CaiCai2017

4. Discontinuous diffusion coefficient does not necessarily lead to a solution of low regularity and uniform refinement may be able to produce an optimal convergence, for example, when the exact solution is piecewise smooth and the mesh aligns with the interface

11.1 Discontinuous Coefficients Do Not Imply Low Regularity

It is known that solutions of elliptic PDEs may suffer from a low regularity due to discontinuities in PDE coefficients (cf. [68, 69]). An example is the Kellogg's example [47, 48] where the diffusion coefficient has large jumps across interfaces of quadrants and the solution is barely in $H^1(\Omega)$. However, it should be emphasized that there is no direct relation between coefficient jump and solution singularity. That is, coefficients

with large jumps may not imply a very low regularity of the PDE solution. In this section, we construct a simple example to show that the regularity of the PDE solution can be independent of the coefficient jump, where the underlying PDE is similar to the one in Kellogg's example.

Let $\Omega = [-1,1] \times [-1,1]$ and denote by Ω_i the intersection of Ω and the ith quadrant. Consider the function in $C(\Omega)$ defined below:

$$u(x,y) := \frac{1}{\alpha(x,y)} x \cos(\pi x/2) \sin(\pi y), \tag{11.1.1}$$

where

$$\alpha(x,y) := \begin{cases} 1, & \text{if } (x,y) \in \Omega_2 \cup \Omega_4, \\ \alpha_1, & \text{otherwise.} \end{cases} \tag{11.1.2}$$

Here $\alpha_1 \neq 1$ is a positive constant. We define U as an extension of u over $\mathbb{R}^2$ by setting $U(x,y) = 0$ for $(x,y) \notin \Omega$. It is easy to verify that $U \in C(\mathbb{R}^2)$.

11.1.1 Regularity in $H^s(\Omega)$

We first investigate the regularity of U, i.e., $U \in H^s(\mathbb{R}^2)$ for which s ?

Since U has a compact support, we know that its Fourier transform is smooth. In fact, its Fourier transform is entire in $\mathbb{C}^2$ according to Schwartz's version of Paley-Wiener Theorem. Therefore, if $\hat{U}$ denotes the Fourier transform of U, in order to show that the following quantity is bounded,

$$\|U\|^2_{H^s(\mathbb{R}^2)} = \int_{\mathbb{R}} \int_{\mathbb{R}} \left(1 + |\xi|^2 + |\mu|^2\right)^s \left|\hat{U}(\xi,\mu)\right|^2 d\xi d\mu,$$

it is equivalent to show the boundedness of the integral over $|\xi| > 1, |\mu| > 1$.

The Fourier transform of u is

$$\begin{aligned} \hat{U}(\xi,\mu) &= \int_{\mathbb{R}} \int_{\mathbb{R}} U(x,y) e^{-i\xi x} e^{-i\mu y} dx dy \\ &= \int_{-1}^{1} \int_{-1}^{1} u(x,y) e^{-i\xi x} e^{-i\mu y} dx dy. \end{aligned}$$

By inserting the expression of u, after some calculations ($\alpha_1 \neq 1$), we arrive at

$$\hat{U}(\xi,\mu) \sim \frac{1}{\xi^2 \mu^2}, \quad |\xi|, |\mu| \to \infty. \tag{11.1.3}$$

Note that for $s > 1$,

$$\left(\frac{\xi^2+\mu^2}{2}\right)^s \le \frac{\xi^{2s}+\mu^{2s}}{2} \le \frac{(\xi^2+\mu^2)^s + (\xi^2+\mu^2)^s}{2} = (\xi^2+\mu^2)^s.$$

Hence

$$(\xi^2+\mu^2)^s \sim \xi^{2s}+\mu^{2s}.$$

We deduce that

$$\begin{aligned}\int_{|\xi|>1}\int_{|\mu|>1}(|\xi|^2+|\mu|^2)^s\left(\frac{1}{\xi^2\mu^2}\right)^2 d\xi d\mu &\sim \int_{|\xi|>1}\int_{|\mu|>1}\frac{\xi^{2s}+\mu^{2s}}{\xi^4\mu^4}d\xi d\mu \\ &= \int_{|\xi|>1}\int_{|\mu|>1}\frac{1}{\xi^{4-2s}\mu^4}+\frac{1}{\xi^4\mu^{4-2s}}d\xi d\mu < \infty\end{aligned} \tag{11.1.4}$$

if and only if $4-2s>1$, i.e., $s<3/2$. Therefore, we conclude that $U \notin H^{3/2}(\mathbb{R}^2)$ and

$$U \in H^s(\mathbb{R}^2), \quad \forall\, s < 3/2. \tag{11.1.5}$$

Since

$$\|u\|_{H^s(\Omega)} := \inf_{\substack{v\in H^s(\mathbb{R}^2)\\ v|_\Omega = u}} \|v\|_{H^s(\mathbb{R}^2)},$$

we see that

$$u \in H^s(\Omega), \quad \forall\, s < 3/2.$$

Moreover, it can be shown that

$$u \notin H^{3/2+\epsilon}(\Omega), \quad \forall\, \epsilon > 0.$$

This is because if $u \in H^{3/2+\epsilon}(\Omega)$ for some $\epsilon > 0$, then it follows that $\frac{\partial u}{\partial x} \in H^{1/2+\epsilon}(\Omega)$ and $\frac{\partial u}{\partial x}|_\gamma \in H^\epsilon(\gamma)$ for any line segment γ in Ω, according to the trace theorem. (Note that for a Lipschitz domain, trace theorem holds for indices in $(\frac{1}{2}, \frac{3}{2})$.) This yields a contradiction because for $\alpha_1 \neq 1$, $\frac{\partial u}{\partial x}|_\gamma \notin L^2(\gamma)$ if we choose γ to be a vertical line segment centered at origin. In fact, the two-sided limits are not equal:

$$\frac{\partial u}{\partial x}|_{\gamma^+} \neq \frac{\partial u}{\partial x}|_{\gamma^-},$$

so $\frac{\partial u}{\partial x}|_\gamma$ does not exist.

Remark 11.1.1. *Can we prove or disprove that $u \notin H^{3/2}(\Omega)$?*

11.1.2 Regularity in periodic Sobolev space

Let H^s_{per} denote the periodic Sobolev space. The periodic Sobolev space in one dimension and biperiodic Sobolev space in two dimensions are defined in [70, p.141,p.317]. H^s_{per} consists of periodic(biperiodic in 2D) distributions, which can be constructed for any function defined in a bounded domain. For example, if u is a nonzero L^2-function defined in the unit square, by making copies of u over the copies of the unit square covering the entire $\mathbb{R}^2$, we obtain a biperiodic function, which is obviously not in $L^2(\mathbb{R}^2)$, but is in H^0_{per}. The norm in H^s_{per} is taken over a single period instead of the whole space. However, note that the regularity in H^s_{per} does not necessarily reflect the regularity in $H^s(\Omega)$ because regularity across the intersection of two adjacent periods is taken into account in H^s_{per} (which affects the index s). For example, the function $f(x) = x$ is in $C^\infty([0,1])$, and after making copies of f over $\mathbb{R}$, denoted by $\tilde{f}$, we have $\tilde{f} \in H^0_{\text{per}}$ but $\tilde{f} \notin H^\infty_{\text{per}}$. Therefore, it is thus beneficial to view H^s_{per} as a Sobolev space over two periods (in 1D) or two biperiods (in 2D), because this accounts for both the interior regularity and regularity across the boundary of a period.

Now we consider the function u in (11.1.1) with $\Omega = [0,1]^2$. We know that

$$\phi_{m,n}(x,y) := 1/2 e^{i\pi m x} e^{i\pi n y}, \quad m,n \in \mathbb{Z},$$

form an orthonormal basis of $L^2(\Omega)$, with $(\phi_{m,n}, \phi_{m',n'})_{L^2(\Omega)} = \delta_{mm'}\delta_{nn'}$. Since Ω is the unit square, $u \in L^2(\Omega)$ and $u|_{\partial\Omega} = 0$, we have

$$u = \sum_{m,n\in\mathbb{Z}} \hat{u}_{m,n}\phi_{m,n} \quad \text{in } L^2(\Omega) \quad \text{with } \hat{u}_{m,n} = (u, \phi_{m,n})_{L^2(\Omega)}.$$

Analougos to (11.1.3) in Section 11.1.1, we can compute that $\hat{u}_{m,n} = 0$ if $n \neq \pm 1$ is odd and otherwise,

$$|\hat{u}_{m,n}| \sim \frac{1}{m^2 n^2}, \quad m,n \to \infty. \tag{11.1.6}$$

Therefore,

$$\begin{aligned}\|u\|^2_{H^s_{\text{per}}} &\sim \sum_{m,n\in\mathbb{Z}^*} (m^2+n^2)^s \left(\frac{1}{m^2n^2}\right)^2 \\ &\sim \left(\sum_{n\neq 0}\frac{1}{n^4}\right)\left(\sum_{m\neq 0}\frac{1}{m^{4-2s}}\right) + \left(\sum_{m\neq 0}\frac{1}{m^4}\right)\left(\sum_{n\neq 0}\frac{1}{n^{4-2s}}\right) < \infty\end{aligned} \tag{11.1.7}$$

if and only if $s < 3/2$. Thus we conclude that

$$u \notin H^{3/2}_{\text{per}} \quad \text{and} \quad u \in H^s_{\text{per}}, \quad \forall\, s < 3/2. \tag{11.1.8}$$

Moreover, if $\alpha_1 = 1$, then $u \in C^\infty(\Omega)$. In this case, we can derive that

$$|\hat{u}_{m,n}| \sim \frac{1}{|m|^3 n^2}, \quad m, n \to \infty,$$

which implies the same regularity as in (11.1.8). This is because u_{xx} is not continuous across interface $x = \pm 1$ between two adjacent periods(∇u is continuous across periods).

$$\begin{aligned} u_x(x,y) &= \cos(\pi x/2)\sin(\pi y) - \frac{\pi}{2} x \sin(\pi x/2)\sin(\pi y), \\ u_y(x,y) &= \pi x \cos(\pi x/2)\cos(\pi y), \\ u_{xx}(x,y) &= -\left(\pi \sin(\pi x/2)\sin(\pi y) + \frac{\pi^2}{4} x \cos(\pi x/2)\sin(\pi y)\right), \\ u_{xx}|_{x=1} &\neq u_{xx}|_{x=-1}. \end{aligned}$$

11.1.3 Comparison to Kellogg's example

Consider again the domain $\Omega = [-1,1]^2$ in two dimensions. With Kellogg's formula [48], a solution in the form

$$u(r,\theta) = r^\lambda \mu(\theta)$$

with $\lambda = 0.1$ was constructed in [47] such that

$$u \notin H^{1.1}(\Omega) \quad \text{and} \quad u \in H^{1.1-\epsilon}(\Omega), \quad \forall\, \epsilon > 0,$$

and u satisfies the following equation

$$\begin{cases} -\mathrm{div}(\alpha\nabla u) = 0 & \text{in } \Omega, \\ u = u_D & \text{on } \partial\Omega, \end{cases} \tag{11.1.9}$$

where α is as in (11.1.2) with $\alpha_1 \approx 161.4476387975881$.

On the other hand, u defined in (11.1.1) satisfies the following equation

$$\begin{cases} -\mathrm{div}(\alpha\nabla u) = f & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega, \end{cases} \tag{11.1.10}$$

with $f \in C^\infty(\Omega)$.

Then we see that even though the diffusion coefficients in (11.1.9) and (11.1.10) are same, the regularities of the corresponding solutions differ a lot:

in (11.1.9), $u \notin H^{1.1}(\Omega)$, while in (11.1.10), $u \in H^{1.5-\epsilon}(\Omega)$, $\forall\, \epsilon > 0$.

Remark 11.1.2. *It is worth noting that the decomposition theorem in [69, Theorem 2.2] only gives a necessary, in other words, worst-case, regularity for the solution of* (11.1.10)*. In fact, this can be seen directly from the decomposition formula in [69, Eq.(2.4.4)]: since the coefficient $c_{l,j}$ of a singular function may equal to zero, only lowest regularity can be deduced from the decomposition formula.*

11.2 Edge Estimator

The residual estimator is composed of two parts: element residuals and edge residuals, i.e., jumps of numerical flux across interelement boundaries. In [67], it was shown that edge residuals dominate the error of linear finite elements up to higher order perturbation terms. One may tend to think that the element residuals serve as higher order terms and do not dominate the error, which, instead, is not true. The fact that edge residuals dominate the error does not exclude the possibility that element residuals also dominte the error. In fact, it can be observed in numerical practice that the total element residual is usually larger than the total edge residual.

We provide some numerical results in this section to show that the element residual term is in general not a higher order term and is usually larger than the edge residual term.

Consider the Poisson problem below

$$\begin{cases} -\Delta u = f & \text{in } \Omega, \\ u = u_D & \text{on } \partial\Omega, \end{cases}$$

where Ω is a square in $\mathbb{R}^2$ with vertices $(-1,-1),(1,-1),(1,1),(-1,1)$. The exact solution is chosen as

$$u(x,y) = -x^2 - y^2.$$

Note that $f = 4$, so there is no so-called data oscillation in f.

We use adaptive finite element with conforming linear elements to solve the problem. The total element residual, total edge residual, and the standard residual estimator are given by:

$$\begin{aligned} \eta_f &:= \left(\sum_{K \in \mathcal{T}} h_K^2 \| \bar{f} - \operatorname{div} \boldsymbol{\sigma}_{\mathcal{T}} \|_K^2 \right)^{1/2}, \\ \eta_{\mathcal{E}} &:= \left(\sum_{e \in \mathcal{E}_I} h_K \| j_e \|_e^2 \right)^{1/2}, \\ \eta &:= \left(\eta_f^2 + \eta_{\mathcal{E}}^2 \right)^{1/2}, \end{aligned}$$

respectively. Here $\bar{f}$ denotes the piecewise L^2 projection of f into P_{k-1} space with respect to mesh $\mathcal{T}$ corresponding to P_k elements. In this particular example, $\bar{f} = f$.

The setup of the experiment is given below:

- tolerance for relative error: 10^{-3} (stopping criterion);
- number of refinements: k;
- number of nodes: N;
- Dörfler's marking strategy: $\theta = 0.2$ in (4.5.1);
- eff-ind: effectivity index $\frac{\eta}{\|\nabla e\|}$, where $e = u_{\mathcal{T}} - u$.

Table 11.1.
Numerical results for element and edge residuals (P_1 elements)

k	N	$\|\nabla e\|$	η	eff-ind	η_f	$\eta_\mathcal{E}$
46	11471	0.0273	0.1447	5.29	0.1123	0.0911

Table 11.2.
Numerical results for element and edge residuals (P_2 elements)

k	N	$\|\nabla e\|$	η	eff-ind	η_f	$\eta_\mathcal{E}$
36	801	0.0023	0.0230	9.83	0.0205	0.0105

The numerical results are shown in Table 11.1 and Figure 11.1 (left). It is easy to see that the total element residual η_f is larger than edge estimator $\eta_\mathcal{E}$ for this simple problem with the lowest order finite element approximation. The phenomenon is more prominent for higher order elements. For example, consider the same problem with the exact solution

$$u(x, y) = -(x^2 - 1)(y^2 - 1)$$

and perform adaptive mesh refinement with P_2 conforming elements. The results for this case are shown in Figure 11.1 (right) and Table 11.2.

Remark 11.2.1. *From the above numerical results, we also see that the effectivity index of the residual estimator is quite large, even for smooth problems.*

Remark 11.2.2. *The results in [67] should be interpreted in asymptotic sense only. On coarse meshes, it is possible that the edge residuals vanish while the true error is large. This issue is discussed in the next section.*

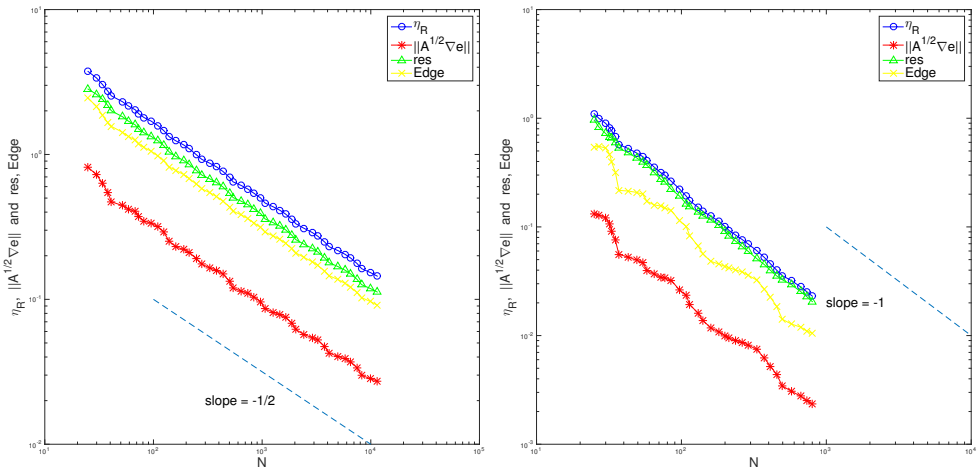

Figure 11.1. Comparison of $\|\nabla e\|$(red), η(blue), η_f(green) and $\eta_{\mathcal{E}}$(yellow). Left: P_1; Right: P_2.

11.3 Unreliability of Some Recovery-based Estimators

A version of this section has been published in [12].

This section aims to show that all ZZ-type estimators, without incorporating f, are in general not reliable on coarse meshes. Hybrid estimators are designed to resolve this issue. A counter example is constructed below, where all estimators solely computed from the finite element solution vanish everywhere while the true error is not zero.

Let Ω be the unit square. Consider a triangulation $\mathcal{T}$ of Ω as illustrated in Fig. 11.2, where the top-left element is denoted by K_1.

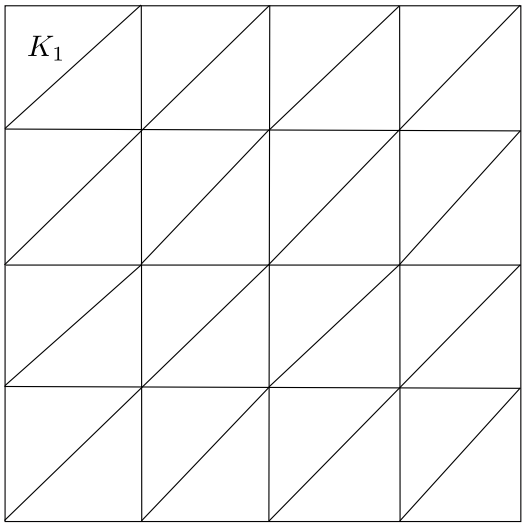

Figure 11.2. Initial mesh $\mathcal{T}$ and element K_1

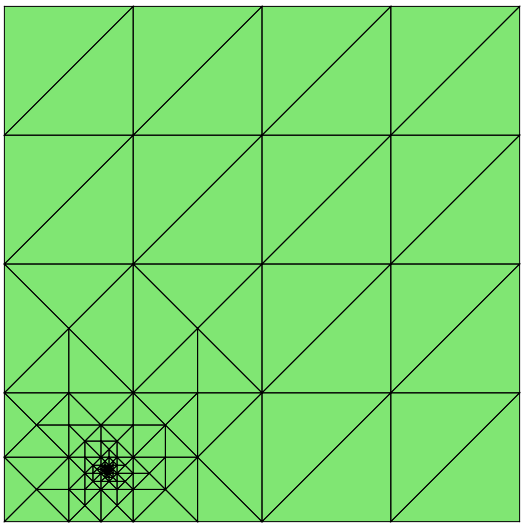

Figure 11.3. The mesh after 100 refinement steps with unreliable estimators

Consider the P_1 conforming finite element discretization associated with mesh $\mathcal{T}$ of the following homogeneous Dirichlet problem:

$$\begin{cases} \Delta u = f & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega, \end{cases}$$

where $f = 2018$ in K_1 and $f = 0$ in $\Omega \backslash K_1$. Note that for any $v \in V_{\mathcal{T}}$, $v|_{K_1} = 0$, which implies that

$$(f, v) = 0, \quad \forall\, v \in V_{\mathcal{T}}.$$

Therefore, the finite element solution is $u_{\mathcal{T}} = 0$ in Ω. All ZZ-type estimators computed solely from $u_{\mathcal{T}}$ are equal to zero. On the other hand, the true error in energy norm is $\|\nabla u\| > 0$. Hence those estimators are not reliable and consequently can not lead to a convergent adaptive finite element algorithm. In fact, based on Dörfler's marking strategy as before, since those estimators vanish everywhere, it suffices to choose an arbitrary element to form $\hat{\mathcal{T}}$ and (4.5.1) automatically holds true. With the initial mesh $\mathcal{T}$ in Fig. 11.2, if the element in $\hat{\mathcal{T}}$ is chosen as the bottom-left element on the horizontal boundary, then the mesh after 100 refinement steps based on unreliable estimators is shown in Fig. 11.3, where the numerical solution (as well as the resulting estimator) is always zero and convergence can never be achieved.

Remark 11.3.1. *Note that there is no oscillation error in f in the counter example above, so the MNS marking strategy proposed in [47] coincides with Dörfler's marking strategy. Also, it is easy to see that such a counter example can always be constructed by choosing f to be orthogonal to $V_{\mathcal{T}}$ no matter the mesh $\mathcal{T}$ is considered coarse or fine.*

11.4 Data Oscillations

The control of data oscillation is crucial in the convergence theory of adaptive finite element method and it is known that adaptive algorithms may fail to converge if the data oscillation is not resolved during the mesh refinement (cf. [47, 71, 72]). On the other hand, is it always necessary to resolve data oscillation ? The answer is obviously no as one can easily create an example where the true error is very small while the data oscillation is large. In this section, we present such an example to illustrate that the finite element solution is very accurate even though data oscillation is not resolved by the mesh. In the example, $f = \Delta u$ has a large L^2 norm while both the L^2

norm and the H^1 semi-norm of u are small. Hence $u_{\mathcal{T}} = 0$ is a good approximation to u.

Define the following function in $C^\infty(\mathbb{R})$ with a small parameter $c > 0$:

$$\psi_{1,c}(x) := \begin{cases} \exp(-\frac{c}{1-x^2}), & \text{if } |x| < 1, \\ 0, & \text{if } |x| \geq 1. \end{cases} \tag{11.4.1}$$

The first order and second order derivatives of $\psi_{1,c}$ are given by

$$\psi'_{1,c}(x) = \begin{cases} \frac{-2cx}{(1-x^2)^2}\psi_{1,c}(x), & \text{if } |x| < 1, \\ 0, & \text{if } |x| \geq 1, \end{cases} \tag{11.4.2}$$

$$\psi''_{1,c}(x) = \begin{cases} c\psi_{1,c}(x)[-\frac{2}{(1-x^2)^2} - \frac{8x^2}{(1-x^2)^3} + \frac{4cx^2}{(1-x^2)^4}], & \text{if } |x| < 1, \\ 0, & \text{if } |x| \geq 1. \end{cases} \tag{11.4.3}$$

For $\delta > 0$, we define

$$\psi_{\delta,c}(x) := \psi_{1,c}(x/\delta).$$

Then

$$\text{supp}\psi_{\delta,c} = [-\delta, \delta], \quad \psi'_{\delta,c}(x) = \frac{1}{\delta}\psi'_{1,c}(x/\delta) \quad \text{and} \quad \psi''_{\delta,c}(x) = \frac{1}{\delta^2}\psi''_{1,c}(x/\delta).$$

When c and δ are close to zero, it is easy to see that

$$\sup|\psi_{\delta,c}| \sim 1, \quad \sup\left|\psi'_{\delta,c}\right| \sim \frac{c}{\delta}, \quad \sup\left|\psi''_{\delta,c}\right| \sim \frac{c}{\delta^2},$$

where $\sim$ denotes an equivalence of two quantities up to constants independent of c and δ. Furthermore, we compute that

$$\begin{aligned} Q_1 &:= \int_{-\delta}^{\delta} |\psi_{\delta,c}|^2 dx \sim \delta, \\ Q_2 &:= \int_{-\delta}^{\delta} \left|\psi'_{\delta,c}\right|^2 dx \sim \frac{c^2}{\delta}, \\ Q_3 &:= \int_{-\delta}^{\delta} \left|\psi''_{\delta,c}\right|^2 dx \sim \frac{c^2}{\delta^3}. \end{aligned}$$

Now we define

$$u_0(x, y) := \psi_{\delta,c}(x)\psi_{\delta,c}(y) \in L^2(\mathbb{R}^2). \tag{11.4.4}$$

Hence $\mathrm{supp} u_0 = [-\delta, \delta] \times [-\delta, \delta]$ and

$$\frac{\partial u_0}{\partial x} = \psi'_{\delta,c}(x)\psi_{\delta,c}(y),$$
$$|\Delta u_0|^2 = \left|\psi''_{\delta,c}(x)\psi_{\delta,c}(y)\right|^2 + 2\psi''_{\delta,c}(x)\psi''_{\delta,c}(y)\psi_{\delta,c}(x)\psi_{\delta,c}(y) + \left|\psi_{\delta,c}(x)\psi''_{\delta,c}(y)\right|^2.$$

We compute that

$$\|u_0\|^2 = \int_{-\delta}^{\delta} |\psi_{\delta,c}(x)|^2 dx \int_{-\delta}^{\delta} |\psi_{\delta,c}(y)|^2 dy = Q_1^2 \sim \delta^2,$$
$$\|\nabla u_0\|^2 = 2\int_{-\delta}^{\delta}\int_{-\delta}^{\delta} \left|\frac{\partial u_0}{\partial x}\right|^2 dxdy = 2\int_{-\delta}^{\delta} \left|\psi'_{\delta,c}(x)\right|^2 dx \int_{-\delta}^{\delta} |\psi_{\delta,c}(y)|^2 dy = 2Q_2Q_1 \sim c^2,$$
$$\|\Delta u_0\|^2 = 2\int_{-\delta}^{\delta} \left|\psi''_{\delta,c}(x)\right|^2 dx \int_{-\delta}^{\delta} |\psi_{\delta,c}(y)|^2 dy + 2\left|\int_{-\delta}^{\delta} \psi''_{\delta,c}(x)\psi_{\delta,c}(x)dx\right|^2 \sim \frac{c^2}{\delta^2}.$$

If we choose $c = \delta^{1/2} \to 0$, then it follows that

$$\|u_0\|^2 \sim \delta^2 \to 0, \quad \|\nabla u_0\|^2 \sim \delta \to 0, \quad \|\Delta u_0\|^2 \sim \frac{1}{\delta} \to \infty. \tag{11.4.5}$$

If we choose $c = \delta \to 0$, then

$$\|u_0\|^2 \sim \delta^2 \to 0, \quad \|\nabla u_0\|^2 \sim \delta^2 \to 0, \quad \|\Delta u_0\|^2 \sim 1.$$

Next we construct an example where numerical solution $u_\tau = 0$. Consider the square domain $\Omega = [-2\delta, 1] \times [-1, 2\delta]$ with $0 < \delta < 10^{-2}$ and the following Dirichlet problem:

$$\begin{cases} \Delta u = f & \text{in} \quad \Omega, \\ u = 0 & \text{on} \quad \partial\Omega \end{cases}$$

with $u := u_0$ in (11.4.4). Define the piecewise linear conforming finite element space

$$V_{\mathcal{T}} = \left\{ v \in H_0^1(\Omega) : v|_K \in P_1(K), \quad \forall K \in \mathcal{T} \right\}$$

with mesh $\mathcal{T}$ shown in Figure 11.4. Since $\mathrm{supp} f = \mathrm{supp} u = [-\delta, \delta] \times [-\delta, \delta] \subset K_1$ and $v = 0$ on K_1 for any $v \in V_{\mathcal{T}}$, it follows that $(f, v) = 0$ for any test function $v \in V_{\mathcal{T}}$. Hence $u_\tau = 0$. Therefore, if $c = \delta^{1/2} \to 0$, then (11.4.5) implies that

$$\|\nabla(u - u_\tau)\| \to 0 \quad \text{and} \quad \|\Delta(u - u_\tau)\| \to \infty.$$

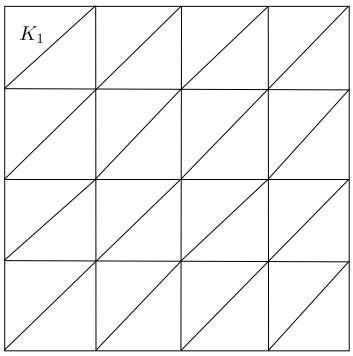

Figure 11.4. Mesh $\mathcal{T}$ over Ω and the element K_1 considered in Section 11.4

REFERENCES

REFERENCES

[1] U.S.N.C.T.A. Mechanics, C.E.T. Systems, B.M.E. Design, D.E.P. Sciences, and N.R. Council. *Research Directions in Computational Mechanics*. A Series. National Academies Press, 1991.

[2] R. Verfürth. Robust a posteriori error estimates for stationary convection-diffusion equations. *SIAM journal on numerical analysis*, 43(4):1766–1782, 2005.

[3] O.C. Zienkiewicz and J.Z. Zhu. A simple error estimator and adaptive procedure for practical engineerng analysis. *International Journal for Numerical Methods in Engineering*, 24(2):337–357, 1987.

[4] O.C. Zienkiewicz and J.Z. Zhu. The superconvergent patch recovery and a posteriori error estimates. part 1: The recovery technique. *International Journal for Numerical Methods in Engineering*, 33(7):1331–1364, 1992.

[5] S. Bartels and C. Carstensen. Each averaging technique yields reliable a posteriori error control in fem on unstructured grids. part ii: Higher order fem. *Mathematics of computation*, 71(239):971–994, 2002.

[6] A. Naga and Z. Zhang. The polynomial-preserving recovery for higher order finite element methods in 2D and 3D. *Discrete and Continuous Dynamical Systems Series B*, 5(3):769, 2005.

[7] R. Bank, J. Xu, and B. Zheng. Superconvergent derivative recovery for Lagrange triangular elements of degree p on unstructured grids. *SIAM Journal on Numerical Analysis*, 45(5):2032–2046, 2007.

[8] C. Bernardi and R. Verfürth. Adaptive finite element methods for elliptic equations with non-smooth coefficients. *Numerische Mathematik*, 85(4):579–608, 2000.

[9] M. Petzoldt. A posteriori error estimators for elliptic equations with discontinuous coefficients. *Advances in Computational Mathematics*, 16(1):47–75, 2002.

[10] Z. Cai and S. Zhang. Robust equilibrated residual error estimator for diffusion problems: Conforming elements. *SIAM Journal on Numerical Analysis*, 50(1):151–170, 2012.

[11] Z. Cai and S. Zhang. Flux recovery and a posteriori error estimators: conforming elements for scalar elliptic equations. *SIAM Journal on Numerical Analysis*, 48(2):578–602, 2010.

[12] D. Cai and Z. Cai. A hybrid a posteriori error estimator for conforming finite element approximations. *Computer Methods in Applied Mechanics and Engineering*, 339:320 – 340, 2018.

[13] M. Ainsworth, A. Allendes, G. R. Barrenechea, and R. Rankin. Fully computable a posteriori error bounds for stabilised FEM approximations of convection–reaction–diffusion problems in three dimensions. *International Journal for Numerical Methods in Fluids*, 73(9):765–790, 2013.

[14] D. Cai and Z. Cai. Hybrid a posteriori error estimators for conforming finite element approximations to stationary convection-diffusion-reaction equations. *submitted*.

[15] P. Ladeveze and D. Leguillon. Error estimate procedure in the finite element method and applications. *SIAM Journal on Numerical Analysis*, 20(3):485–509, 1983.

[16] M. Ainsworth and J.T. Oden. *A Posteriori Error Estimation in Finite Element Analysis*. A Wiley-Interscience publication. Wiley, 2000.

[17] D. Braess. *Finite Elements: Theory, Fast Solvers, and Applications in Solid Mechanics*. Cambridge University Press, 2007.

[18] D. Braess and J. Schöberl. Equilibrated residual error estimator for edge elements. *Mathematics of Computation*, 77(262):651–672, 2008.

[19] D. Braess, V. Pillwein, and J. Schöberl. Equilibrated residual error estimates are p-robust. *Computer Methods in Applied Mechanics and Engineering*, 198(13):1189–1197, 2009.

[20] D. Cai, Z. Cai, and S. Zhang. Robust equilibrated a posteriori error estimator for higher-order finite element approximations to diffusion problems. *submitted*.

[21] D. Cai, Z. Cai, and S. Zhang. Explicit and robust equilibrated a posteriori error estimator for conforming finite element approximations in two dimensions. *preprint*.

[22] M. Ainsworth and R. Rankin. Fully computable bounds for the error in nonconforming finite element approximations of arbitrary order on triangular elements. *SIAM Journal on Numerical Analysis*, 46(6):3207–3232, 2008.

[23] K.-Y. Kim. Flux reconstruction for the $P2$ nonconforming finite element method with application to a posteriori error estimation. *Applied Numerical Mathematics*, 62(12):1701 – 1717, 2012.

[24] D. Cai and Z. Cai. Explicit and robust equilibrated a posteriori error estimator for nonconforming finite element approximations of arbitrary order. *preprint*.

[25] V. Girault and P.-A. Raviart. *Finite element methods for Navier-Stokes equations: theory and algorithms*, volume 5. Springer Science & Business Media, 2012.

[26] L. Tartar. *An introduction to Sobolev spaces and interpolation spaces*, volume 3. Springer Science & Business Media, 2007.

[27] M. Costabel. Boundary integral operators on lipschitz domains: elementary results. *SIAM journal on Mathematical Analysis*, 19(3):613–626, 1988.

[28] D. Boffi, M. Fortin, and F. Brezzi. *Mixed finite element methods and applications*. Springer series in computational mathematics. Springer, Berlin, Heidelberg, 2013.

[29] W. Prager and J.L. Synge. Approximations in elasticity based on the concept of function space. *Q. Appl. Math.*, 5:241–269, 1947.

[30] E. Dari, R. Duran, C. Padra, and V. Vampa. A posteriori error estimators for nonconforming finite element methods. *RAIRO-Modélisation mathématique et analyse numérique*, 30(4):385–400, 1996.

[31] M. Ainsworth. Robust a posteriori error estimation for nonconforming finite element approximation. *SIAM Journal on Numerical Analysis*, 42(6):2320–2341, 2005.

[32] M. Ainsworth and R. Rankin. Robust a posteriori error estimation for the nonconforming Fortin–Soulie finite element approximation. *Mathematics of Computation*, 77(264):1917–1939, 2008.

[33] M. Schedensack. *A class of mixed finite element methods based on the Helmholtz decomposition in computational mechanics*. PhD thesis, Humboldt-Universität zu Berlin, Mathematisch-Naturwissenschaftliche Fakultät, 2015.

[34] S. Cochez-Dhondt and S. Nicaise. Equilibrated error estimators for discontinuous galerkin methods. *Numerical Methods for Partial Differential Equations*, 24(5):1236–1252, 2008.

[35] P. Ciarlet. *The Finite Element Method for Elliptic Problems*. Society for Industrial and Applied Mathematics, 2002.

[36] I. Babuška and W. C. Rheinboldt. Error estimates for adaptive finite element computations. *SIAM Journal on Numerical Analysis*, 15(4):736–754, 1978.

[37] I. Babuška and A. Miller. A feedback element method with a posteriori error estimation: Part i. the finite element method and some basic properties of the a posteriori error estimator. *Comput. Methods Appl. Mech. Eng.*, 61(1):1–40, March 1987.

[38] R. Verfürth. *A Posteriori Error Estimation Techniques for Finite Element Methods*. Numerical Mathematics and Scientific Computation. OUP Oxford, 2013.

[39] C. Carstensen and C. Merdon. Estimator competition for poisson problems. *Journal of Computational Mathematics*, pages 309–330, 2010.

[40] C. Carstensen and S. Bartels. Each averaging technique yields reliable a posteriori error control in FEM on unstructured grids. Part I: Low order conforming, nonconforming, and mixed FEM. *Mathematics of Computation*, 71(239):945–969, 2002.

[41] R. Bank and J. Xu. Asymptotically exact a posteriori error estimators, part i: Grids with superconvergence. *SIAM Journal on Numerical Analysis*, 41(6):2294–2312, 2003.

[42] J.S. Ovall. Two dangers to avoid when using gradient recovery methods for finite element error estimation and adaptivity. Technical report, Technical report 6, Max-Planck-Institute fur Mathematick in den Naturwissenschaften, Bonn, Germany, 2006.

[43] Z. Cai and S. Zhang. Recovery-based error estimator for interface problems: Conforming linear elements. *SIAM Journal on Numerical Analysis*, 47(3):2132–2156, 2009.

[44] E. G. Sewell. *Automatic generation of triangulations for piecewise polynomial approximation*. PhD thesis, Purdue University, West Lafayette, IN, 1972.

[45] R. Verfürth. A posteriori error estimation and adaptive mesh-refinement techniques. *Journal of Computational and Applied Mathematics*, 50(1):67 – 83, 1994.

[46] W. Dörfler. A convergent adaptive algorithm for poissons equation. *SIAM Journal on Numerical Analysis*, 33(3):1106–1124, 1996.

[47] P. Morin, R.H. Nochetto, and K.G. Siebert. Convergence of adaptive finite element methods. *SIAM review*, 44(4):631–658, 2002.

[48] R. B. Kellogg. On the poisson equation with intersecting interfaces. *Applicable Analysis*, 4(2):101–129, 1974.

[49] R. Verfürth. Robust a posteriori error estimators for a singularly perturbed reaction-diffusion equation. *Numerische Mathematik*, 78(3):479–493, 1998.

[50] R. Verfürth. A posteriori error estimators for convection-diffusion equations. *Numerische Mathematik*, 80(4):641–663, 1998.

[51] G. Sangalli. A robust a posteriori estimator for the residual-free bubbles method applied to advection-diffusion problems. *Numerische Mathematik*, 89(2):379–399, 2001.

[52] G. Sangalli. A uniform analysis of nonsymmetric and coercive linear operators. *SIAM journal on mathematical analysis*, 36(6):2033–2048, 2005.

[53] G. Sangalli. Robust a-posteriori estimator for advection-diffusion-reaction problems. *Mathematics of Computation*, 77(261):41–70, 2008.

[54] C. Carstensen. All first-order averaging techniques for a posteriori finite element error control on unstructured grids are efficient and reliable. *Mathematics of Computation*, 73(247):1153–1165, 2004.

[55] M. Ainsworth and T. Vejchodský. Fully computable robust a posteriori error bounds for singularly perturbed reaction–diffusion problems. *Numerische Mathematik*, 119(2):219–243, 2011.

[56] L.P. Franca, S.L. Frey, and T.J.R. Hughes. Stabilized finite element methods: I. application to the advective-diffusive model. *Computer Methods in Applied Mechanics and Engineering*, 95(2):253–276, 1992.

[57] M. Ainsworth and I. Babuška. Reliable and robust a posteriori error estimation for singularly perturbed reaction-diffusion problems. *SIAM journal on numerical analysis*, 36(2):331–353, 1999.

[58] M. Ainsworth and T. Vejchodský. Robust error bounds for finite element approximation of reaction–diffusion problems with non-constant reaction coefficient in arbitrary space dimension. *Computer Methods in Applied Mechanics and Engineering*, 281:184–199, 2014.

[59] A. Ern and M. Vohralík. Stable broken H^1 and $\mathbf{H}(\mathrm{div})$ polynomial extensions for polynomial-degree-robust potential and flux reconstruction in three space dimensions. *ArXiv e-prints*, January 2017.

[60] G. M. Ziegler. *Lectures on polytopes*, volume 152. Springer Science & Business Media, 2012.

[61] R. Verfürth. A note on constant-free a posteriori error estimates. *SIAM Journal on Numerical Analysis*, 47(4):3180–3194, 2009.

[62] R. Luce and B. Wohlmuth. A local a posteriori error estimator based on equilibrated fluxes. *SIAM Journal on Numerical Analysis*, 42(4):1394–1414, 2004.

[63] D. Braess and R. Verfürth. A posteriori error estimators for the raviart-thomas element. *SIAM Journal on Numerical Analysis*, 33(6):2431–2444, 1996.

[64] M. Crouzeix and P.-A. Raviart. Conforming and nonconforming finite element methods for solving the stationary stokes equations i. *ESAIM: Mathematical Modelling and Numerical Analysis-Modélisation Mathématique et Analyse Numérique*, 7(R3):33–75, 1973.

[65] M Fortin and M Soulie. A non-conforming piecewise quadratic finite element on triangles. *International Journal for Numerical Methods in Engineering*, 19(4):505–520, 1983.

[66] D. Cai and Z. Cai. A counter example concerning the robustness of equilibrated estimators and the quasi-monotonicity condition. *submitted*.

[67] C. Carstensen and R. Verfürth. Edge residuals dominate a posteriori error estimates for low order finite element methods. *SIAM Journal on Numerical Analysis*, 36(5):1571–1587, 1999.

[68] P. Grisvard. *Singularities in boundary value problems*, volume 22. Springer, 1992.

[69] M. Petzoldt. *Regularity and error estimators for elliptic problems with discontinuous coefficients*. PhD thesis, Freie Universität Berlin, 2001.

[70] J. Saranen and G. Vainikko. *Periodic integral and pseudodifferential equations with numerical approximation*. Springer monographs in mathematics. Springer, Berlin, New York, 2002.

[71] P. Morin, R. H. Nochetto, and K. G. Siebert. Data oscillation and convergence of adaptive FEM. *SIAM Journal on Numerical Analysis*, 38(2):466–488, 2000.

[72] J.M. Cascon, C. Kreuzer, R.H. Nochetto, and K.G. Siebert. Quasi-optimal convergence rate for an adaptive finite element method. *SIAM Journal on Numerical Analysis*, 46(5):2524–2550, 2008.