

ADAPTIVITY IN ANISOTROPIC FINITE ELEMENT CALCULATIONS

Der Fakultät für Mathematik der Technischen Universität Chemnitz vorgelegte

DISSERTATION

zur Erlangung des akademischen Grades
Doctor rerum naturalium
(Dr. rer. nat.)

überarbeitete Version

von

M. SC. SERGEY GROSMAN

geboren am 15.08.1978 in Rostow am Don, Russland

Institut für Mathematik und Bauinformatik
Fakultät für Bauingenieur- und Vermessungswesen
Universität der Bundeswehr München

Contents

1	Introduction	3
1.1	Anisotropic adaptation in finite element method	3
1.2	Error estimation and convergence of adaptive procedure	4
1.3	Aim of this work and outline	5
2	Preliminaries	8
2.1	The model problem and its discretization	8
2.2	Some notations	9
2.3	Notation on the triangle and general mesh requirements	10
2.4	Bubble functions and their inverse inequalities	10
2.5	Special edge bubble functions	11
2.6	Refinement functions and their inverse inequalities	13
3	The strengthened Cauchy-Schwarz Inequality	15
3.1	Theoretical background	15
3.2	Pure Laplace problem $\kappa = 0, \varepsilon = 1$	15
3.3	Squeezed case	17
3.4	Additional pair of spaces	22
4	Convergence results for Poisson problem on isotropic meshes	24
4.1	Isotropic residual error estimator	24
4.2	The convergent algorithm due to Dörfler	25
4.3	Convergent algorithm due to Morin, Nochetto and Siebert	26
4.4	Avoidance of the new inner nodes	30
4.5	Error reduction	32
4.6	Advantage of edge dominance in error estimation regarding adaptive refinement	35
5	Convergence results for Poisson problem on anisotropic meshes	37
5.1	Residual anisotropic error estimator	37
5.2	On the possible choice of weights for the edge error indicator	39
5.3	Marking strategies and convergent algorithm on anisotropic meshes	44
5.4	Error reduction theorem	46
6	Error estimation for the reaction-diffusion problem	49
6.1	Equilibrated residual method with computable error approximation	49
6.1.1	The equilibrated residual method	49
6.1.2	Construction of the equilibrated fluxes	51
6.1.3	Minimum energy extensions	53
6.1.4	Estimates for element and face residuals in the anisotropic case	56
6.1.5	Lower error bound of the original Ainsworth-Babuška estimator in the anisotropic singularly perturbed case	61
6.1.6	Modified equilibrated residual method	63
6.1.7	Computable approximation for the solution to the local problem	64
6.1.8	Numerical experiments	68
6.2	Residual error estimator	69
6.3	Error reduction property	71
6.4	Hierarchical error estimator	74
6.5	Numerical experiments	79

7	Towards a convergent algorithm for the reaction-diffusion problem	81
7.1	Alternative residual a posteriori error estimation	81
7.2	Marking strategy	82
7.3	Error reduction theorem	83
7.4	Numerical experiments	87
7.5	Discussion of the alignment measure m_1	88

1 Introduction

1.1 Anisotropic adaptation in finite element method

The most physical problems considered within the area of civil engineering may be described by partial differential equations. Examples include diffusion and heat conduction problems, the calculation of electrostatic potential distributions, and the calculation displacement fields in linear elasticity. The aforementioned problems, as well as many others, are in general too complex to be solved analytically. The finite element method is one of the numerical methods that can be used in such situations. Other methods include the finite difference method, the finite volume method and the boundary element method. In this work we will exclusively consider the finite element method (FEM). It is the most common numerical tool for analysis of diverse structural systems. It was first formulated for the problem of torsion of a cylinder in 1943 by Courant [19]. Since then, it has come a long way.

When investigating real-world problems one is primarily interested in an accurate solution computed with as little effort as possible. The presence of numerical error in calculations has been a principal source of concern since the beginning of computer simulations of physical phenomena. In the numerical solution of practical problems of physics and engineering such as, e.g., computational fluid dynamics, elasticity, or semiconductor device simulation one often encounters the difficulty that the overall accuracy of the numerical approximation is deteriorated by local singularities such as, e.g., singularities arising from re-entrant corners, interior or boundary layers, or sharp shock-like fronts.

The topic of our work is a special class of problems which can be solved very efficiently by a non-classical finite element method. Some boundary value problems yield a solution which exhibits little variation in one direction but much change in an another direction. Such solutions are called *anisotropic*. Examples include functions which are almost constant or linear on one direction, and which have a boundary or interior layer. An equivalent description is that an anisotropic function shows an almost one-dimensional behavior. By this we mean that the function varies significantly only perpendicularly to a certain manifold. Typical problems with anisotropic solutions include the diffusion-convection-reaction equation (see e.g. [8]), the Poisson equation in a three-dimensional domain with an edge of an interior angle larger than π (see e.g. [10]), and other problems arising from fluid dynamics or weather simulation, for example. Functions which are not anisotropic are called *isotropic* - clearly, this distinction is not a strict mathematical partitioning of the set of functions but rather a matter of degree.

One feature of the classical finite element method is that the ratio of the diameters of the circumscribed and inscribed sphere of a finite element is bounded. Such meshes are referred to as *isotropic meshes*. But when an anisotropic solution as mentioned above occurs it is sensible to violate this condition and to employ highly stretched elements instead. From a heuristic point of view, as well as from anisotropic interpolation analysis (see [6]), it is natural to use a small mesh size in the direction of the rapid variation of the solution, and a larger mesh size in the direction of little variation (i.e. along the manifold of anisotropy). We also say the mesh is anisotropically aligned with the solution, and we refer to it as *anisotropic mesh*. In this way one hopes to capture the important features of the solution with much less elements than when using an isotropic mesh. Numerical evidence confirms that problems with anisotropic solutions can indeed be solved more efficiently on anisotropic meshes (i.e. with less degrees of freedom, less computational effort, or less memory to achieve the same accuracy). If the anisotropy

of a solution occurs along a curved manifold then the anisotropic mesh (i.e. stretched elements) has to follow that curved manifold.

Some important problem classes which frequently yield anisotropic solutions include diffusion-convection-reaction equations and flow simulation (see e.g. [41]). In our work we consider the *singularly perturbed reaction-diffusion equation* $-\varepsilon^2 \Delta u + \kappa^2 u = f$ as a model problem. It displays certain typical features of the aforementioned problems, for example boundary layers which can be discretized advantageously with anisotropic meshes. In the case of singularly perturbed problems the special mesh adaptivity is desirable. Triangles should not only adapt in size but also in shape, to fit the function to be approximated better. The singularly perturbed reaction-diffusion problem typically requires triangles stretched along the boundary or in the direction of the interior layer [5, 6, 9].

1.2 Error estimation and convergence of adaptive procedure

Numerical error is intrinsic in such situations: The discretization process of transforming a continuum model of physical phenomena into one manageable by digital computers cannot capture all of the information of embodied models characterized by partial differential equations or integral equations. The questions of the approximation error and how the error can be controlled, measured, and effectively minimized have confronted computational mechanicians, practitioners, and theorists alike since the earliest applications of numerical methods to problems in engineering and science.

Concrete advances toward the resolution of such questions have been made in the form of theories and methods of *a posteriori* error estimation, whereby the computed solution itself is used to assess the accuracy. The remarkable success of some a posteriori error estimators has opened a new chapter in computational mathematics and mechanics that could revolutionize the subject. By effectively estimating the error, the possibility of controlling the entire computational process through new adaptive algorithms emerges.

Of course, the calculation of the a posteriori error estimate should be far less expensive than the computation of the numerical solution. Moreover, the error estimator should be local and should yield reliable upper and lower bounds for the true error in a user-specified norm. In this context one should note, that global upper bounds are sufficient to obtain the numerical solution with an accuracy below a prescribed tolerance. Local lower bounds, however, are necessary to ensure that the grid is correctly refined so that one obtains a numerical solution with a prescribed tolerance using a (nearly) minimal number of grid points.

When the quality of error estimation is to be assessed, one often encounters the terms *reliable*, *efficient* and *asymptotically exact*. To explain these terms, define the so-called *effectivity index* to be the ratio of the estimated error and the true error (in some norm). Primarily one is interested in estimators that reliably bound the error, i.e. the error is guaranteed to be smaller than some estimated value. Such estimators are called *reliable*. In other words, the error is bounded from above, and the effectivity index is bounded from below. Secondly, an estimator is said to be efficient if it bounds the error from below. This corresponds to the effectivity index, bounded from above. Efficient local estimators are desirable in order to reduce the global error by refining elements with large local error contributions. In our exposition we will refer to the lower and upper bound of the error respectively. Lastly, an estimator is said to be *asymptotically exact* if the effectivity index tends to one as the discretization becomes finer. There is a large variety of a posteriori error estimation techniques. We do not aim at giving an

overview of all related works here, instead we refer to [4, 47] and citations therein.

Disposing of an a posteriori error estimator, an adaptive mesh refinement process has the following general structure:

Step 1. Construct an initial mesh \mathcal{T}_0 representing sufficiently well the geometry of the problem. Put $k := 0$.

Step 2. Solve the discrete problem on \mathcal{T}_k .

Step 3. For each element K of \mathcal{T}_k compute the a posteriori error estimate.

Step 4. If the estimated global error is sufficiently small then **stop**.

Step 5. Based on the information obtained in *Step 3* reconstruct the old mesh (refinement and possibly coarsening) to get the next mesh \mathcal{T}_{k+1} . Replace k by $k + 1$ and return to *Step 2*.

Beside the robustness of the error estimation the success of the adaptive process rests itself on the appropriate mesh reconstruction procedure. In other words, the relation between *Step 3* and *Step 5* has to guarantee in some suitable way the quality of the solution/mesh after some steps of the adaptive process. Long time the restructuring *Step 5* was based only on the heuristic ideas, see George [25, 26], Liseikin [38, 39], Dolejší [22] and many others.

Only recently mathematicians could prove, starting with the pioneer work of Dörfler [23], that using some special refinement techniques the error on the new, adaptively generated mesh is significantly smaller than the error on the actual mesh. This idea was then further developed by Morin, Nochetto and Siebert [40], Binev, Dahmen and Devore [15], Stevenson [43] and the work for the reaction-diffusion problem by Stevenson [44].

In this context we should mention the work by Vassilevski, Dyadechko and Lipnikov [45] where not itself the convergence of the adaptive process is discussed, but the quasi-optimality of the resulting mesh. This paper is based on the aforementioned ideas stemming from George [25, 26] and Liseikin [38, 39].

1.3 Aim of this work and outline

The present work includes three main contributions to the theory of the adaptive finite element method on anisotropic meshes:

1. The hierarchical a posteriori error estimator is developed for the reaction-diffusion problem on anisotropic meshes.
2. The lack of the equilibrated residual method regarding insolvability of the auxiliary problems is considered and the solution possibility is provided.
3. Diverse adaptive algorithms based on the a posteriori error estimation for the Poisson and reaction-diffusion problems on anisotropic meshes are provided, starting with the simplest idea for the Poisson problem on isotropic meshes. The corresponding error reduction estimates are proved for each case.

After this introduction we start with the basics needed for the rest of the work in Section 2, including the model problem description, the mesh requirements, some basic inequalities and various notation.

In Chapter 3 the validity of the strengthened Cauchy-Schwarz inequality is confirmed. In fact, it is shown that there exists a positive constant $\gamma < 1$, such that

$$(x, y) \leq \gamma \|x\| \|y\|, \quad \forall x \in V_1, y \in \tilde{V}_2,$$

where V_1 is the original piecewise linear finite element space, \tilde{V}_2 is the enrichment space, needed in Section 6.3. We emphasize that the constant γ in the strengthened Cauchy-Schwarz inequality for the chosen pair of spaces is strictly smaller than 1 and independent of the aspect ratio and the perturbation parameters. This inequality is one of the main ingredients for the proof of the robustness of the hierarchical error estimator. Furthermore, we verify the Cauchy-Schwarz inequality for two additional pairs of spaces which are needed for the analysis of the hierarchical error estimator and the analysis of adaptive algorithm for the reaction-diffusion problem.

Chapter 4 starts with the repetition of the isotropic residual a posteriori error estimator for the Poisson problem (Section 4.1) and then deals with adaptive convergent algorithms on isotropic meshes. The convergence of the adaptive finite element method is understood in the sense that there exists a positive constant $\beta < 1$, such that in some norm $\|u - u_k\| \leq C\beta^k$ holds, where u is the exact solution, u_k is the finite element approximation in k steps of the adaptive algorithm. Such algorithms with the special marking and refinement strategies are known for the case of Poisson and reaction-diffusion problems on isotropic meshes. Two known adaptive algorithms are described in Sections 4.2 and 4.3. The marking strategy not only takes into account the usual information given by the edge and element error indicators, but also involves the additional control of the data oscillation terms in order to guarantee the convergence. The refinement strategy usually includes the creation of a new node inside each marked triangle. This, however, is not a must – the creation of a new node can be avoided for all triangles where the edge error indicator dominates the element error indicator. We describe this possibility in Section 4.4 and successively prove the convergence of the corresponding algorithm in Section 4.5. At the end of this chapter, in Section 4.6 it is noted that the edge error indicators usually dominate in the error estimator for the Poisson problem.

Chapter 5 starts with repeating the a posteriori residual error estimator for the anisotropic case in Section 5.1. We find out in Section 5.2 that the weights in the edge error indicator on anisotropic meshes is not uniquely defined. Furthermore we provide the range of possible values of this weight, for which all the robustness estimates hold true. Based on the idea from the isotropic case, the adaptive algorithm allowing anisotropic mesh refinement is developed and analyzed in this chapter – one can analyze the edge and element error indicators separately, and according to this information only the marked entities (edges/elements) should be appropriately refined. In Section 5.3 the adaptive algorithm is given. It satisfies the convergence property for the Poisson problem with a parameter β depending on the alignment measure m_1 , as shown in Section 5.4.

In Chapter 6 we deal with the a posteriori error estimators for the singularly perturbed reaction-diffusion problem. Section 6.1 states the equilibrated residual method. Among others, the equilibrated residual method involves the solution of an infinite dimensional local problem on each element. In practical computations an approximate solution to this local problem was successfully computed. Nevertheless, up to now no rigorous analysis has been done showing the appropriateness of any computable approximation. This demands special attention since an improper approximate solution to the local problem can be fatal for the robustness of the whole method. In the present work we provide one of the possible approximations in Subsection 6.1.7 and prove that the method is not affected by the approximate solution of the local problem. As for the rest, Section 6.1 consists of the repetition of the known results needed for the proof.

In Section 6.3 we give a proof of the error reduction property. The error reduction property signifies that using the quadratic finite element basis we achieve strictly

higher accuracy than with linear ones. That is, in some norm $\|\cdot\|$:

$$\|u - u_2\| \leq \alpha \|u - u_1\|, \quad \text{where } \alpha < 1,$$

where u_1 is the usual linear finite element solution, u_2 is the solution using the enriched finite element space. However, as it was shown in the paper by Dörfler and Nochetto [24], there are examples where the error reduction fails in this form (the equation $f = -\Delta u$ was set under consideration). The modification done there concerns on additional term – the so-called data oscillation appears in the right hand side. For more details on data oscillation see [24]. Their proof of the error reduction property was based on the residual a posteriori error estimator. More recently, Agouzal [1] has given a proof for the error reduction property for the reaction-diffusion equation. The proof in this case does not involve any theory of residual a posteriori error estimators. Since [1] is not easy to transfer to anisotropic case the proof of the present work mainly follows the lines of the work [24], but appears to be much more technical. The estimate obtained in Section 6.3 is not only uniform with respect to the mesh size, but also with respect to the aspect ratio and the perturbation parameters κ and ε . Furthermore, in Section 6.4 the error reduction property and the strengthened Cauchy-Schwarz inequality are utilized in order to show the reliability and the efficiency of the proposed hierarchical error estimator. The final estimates are in accordance with [35] and [27]. The numerical experiments presented in Section 6.5 confirm our formulas for the robustness of the error estimator and show the validity of the saturation assumption.

In Chapter 7 we deal with the adaptive algorithm that allows anisotropic triangulations works in addition for the reaction-diffusion problem. The error reduces in each adaptive step, but the convergence property does not seem to be possible to be proven for this algorithm because of additional data oscillation terms. Numerical experiments in Section 7.4 confirm the theory for the adaptive algorithm. The adaptive algorithm shows its potential by creating the anisotropic mesh for the problem with the boundary layer starting with a very coarse isotropic mesh.

2 Preliminaries

2.1 The model problem and its discretization

Let $\Omega \subset \mathbb{R}^2$ be an open domain with polyhedral boundary $\partial\Omega$. Consider the reaction-diffusion equation with homogeneous Dirichlet boundary conditions

$$-\varepsilon^2 \Delta u + \kappa^2 u = f \text{ in } \Omega, \quad u = 0 \text{ on } \partial\Omega, \quad (2.1)$$

where κ and ε is the nonnegative constants. If $\varepsilon \ll 1$, then we have a singularly perturbed problem. Many physical phenomena lead to singularly perturbed problems, for instance, boundary value problems formulated on thin domains [49], where ε is proportional to the domain thickness. They also arise in mathematical models of physical problems, where diffusion is small compared with reaction and convection.

Such problems yield solutions with local anisotropic behavior, e.g. boundary and/or interior layers. Indeed, if $\kappa = O(1)$ and $\varepsilon \ll 1$ then clearly $u \approx f/\kappa$ inside the domain, but due to the diffusion term it is still possible to satisfy the homogeneous boundary conditions. From the heuristical point of view the solution in such a situation could exhibit boundary layers, which is also shown theoretically. The literature on this topic is vast and we are not going to make here a complete overview, but refer the book of Apel [6] and references therein instead. The case $\varepsilon = O(1)$ and $\kappa \gg 1$ arises when discretizing in time [2]. To be general we incorporate both cases in one equation and do not differ them later in work.

Assume $f \in L_2(\Omega)$. The Sobolev space of functions from $H^1(\Omega)$ that vanish on $\partial\Omega$ is denoted by $H_0^1(\Omega)$ as usual. The corresponding variational formulation for (2.1) becomes:

$$\text{Find } u \in H_0^1(\Omega) : \quad B(u, v) = (f, v) \quad \forall v \in H_0^1(\Omega), \quad (2.2)$$

where

$$\begin{aligned} B(u, v) &:= \int_{\Omega} \left(\varepsilon^2 \nabla^\top u \nabla v + \kappa^2 uv \right) dx, \\ (f, v) &:= \int_{\Omega} f v dx. \end{aligned}$$

We utilize a family $\mathcal{F} = \{\mathcal{T}_h\}$ of triangulations \mathcal{T}_h of Ω . By \mathcal{E}_h we denote the collection of edges in the triangulation \mathcal{T}_h . Let $V_h \subset H_0^1(\Omega)$ be the space of continuous, piecewise linear functions over \mathcal{T}_h that vanish on $\partial\Omega$. Then the finite element solution $u_h \in V_h$ is uniquely defined by

$$B(u_h, v_h) = (f, v_h) \quad \forall v_h \in V_h. \quad (2.3)$$

Due to the Lax-Milgram Lemma both problems (2.2) and (2.3) admit unique solutions.

2.2 Some notations

We will use the following notations

L_2 -norm:	$\ v\ _{L_2(\Omega)}$	$:= \left(\int_{\Omega} v^2 dx \right)^{1/2},$
Energy norm:	$\ v\ $	$:= \left(\varepsilon^2 \ \nabla u\ _{L_2(\Omega)}^2 + \kappa^2 \ u\ _{L_2(\omega)}^2 \right)^{1/2},$
Local energy scalar product:	$B_K(u, v)$	$:= \int_K (\varepsilon^2 \nabla^\top u \nabla v + \kappa^2 uv) dx,$
Local L_2 -norm:	$\ v\ _{L_2(K)}$	$:= \left(\int_K v^2 dx \right)^{1/2},$
Local energy norm:	$\ v\ _K$	$:= \left(\varepsilon^2 \ \nabla u\ _{L_2(K)}^2 + \kappa^2 \ u\ _{L_2(K)}^2 \right)^{1/2},$
Length of an edge γ	$ \gamma $	$:= \text{meas}_1(\gamma),$
Volume of an element K	$ K $	$:= \text{meas}_2(K),$
Patch of an element K	ω_K	$:= \text{int} \{ \text{Uclosure}(K'), K' \in \mathcal{T}_h : \partial K \cap \partial K' \text{ is nonempty} \},$
Patch of an edge γ	ω_γ	$:= \text{int} \{ \text{Uclosure}(K), K \in \mathcal{T}_h : \gamma \subset \partial K \}.$

Let \mathcal{N} be the set of all the nodes in triangulation \mathcal{T}_h , then we denote by $\mathcal{N}(K)$ and $\mathcal{N}(\gamma)$ the set of all vertices of a triangle K and an edge γ respectively. Let $n \in \mathcal{N}$ be any node and let θ_n be the Lagrange function associated with the node. The set of elements \mathcal{T}_{hn} influenced by this function is defined by

$$\mathcal{T}_{hn} = \{K \in \mathcal{T}_h : n \in \mathcal{N}(K)\},$$

or, simply stated, \mathcal{T}_{hn} is the patch of elements with a vertex at x_n . Similarly, the set of edges influenced by the basis function is

$$\mathcal{E}_n = \{\gamma \in \partial \mathcal{T}_h : n \in \mathcal{N}(\gamma)\}, \quad (2.4)$$

and consists of those edges having a vertex at x_n .

Let $\{\theta_n : n \in \mathcal{N}\}$ be a Lagrange basis for the space V_h including the basis functions corresponding to the nodes on the Dirichlet boundary $\partial\Omega$, characterized by the conditions

$$\theta_m(x_n) = \delta_{mn} \quad m, n \in \mathcal{N}$$

so that

$$\sum_{n \in \mathcal{N}} \theta_n(x) = 1 \text{ in } \Omega.$$

The nodes on an element K are denoted by $\mathcal{N}(K)$, and it follows that the Lagrange basis functions on the element satisfy

$$\sum_{n \in \mathcal{N}(K)} \theta_n(x) = 1 \text{ in } K. \quad (2.5)$$

Similarly the nodes on an edge γ are denoted by $\mathcal{N}(\gamma)$ and the basis functions satisfy

$$\sum_{n \in \mathcal{N}(\gamma)} \theta_n(x) = 1 \text{ on } \gamma.$$

We will require an extension operator $F_{ext} : \mathbb{P}^0(\gamma) \mapsto \mathbb{P}^0(K)$ defined by

$$F_{ext}(\varphi)(x) := \varphi|_\gamma \equiv \text{const.}$$

We will also use the following notations

$$a \succeq b \Leftrightarrow a \geq Cb,$$

$$a \preceq b \Leftrightarrow a \leq Cb.$$

2.3 Notation on the triangle and general mesh requirements

Let a triangulation \mathcal{T} be given which satisfies the usual conformity condition (see [18], Chapter 2). Following the notation of Kunert [31], the three vertices of an arbitrary triangle $K \in \mathcal{T}$ are denoted by P_0, P_1, P_2 such that P_0P_1 is the longest edge of K . Additionally define two orthogonal vectors p_i with lengths $h_{i,K} := |p_i|$, see Figure 1. Observe that $h_{1,K} > h_{2,K}$ and set $h_{min,K} := h_{2,K}$, $h_{max,K} := h_{1,K}$.

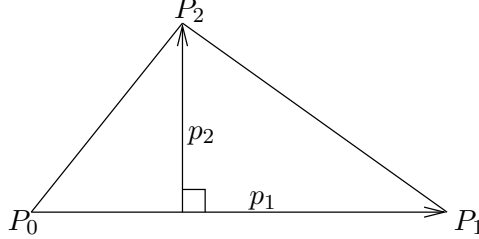


Figure 1: Notation of a triangle K .

In addition to the usual conformity conditions of the mesh we assume that the following three properties hold.

1. The number of triangles containing a node x_n is bounded uniformly.
2. The dimensions of adjacent triangles must not change rapidly, i.e.

$$h_{i,K'} \sim h_{i,K} \quad \forall K, K' \text{ with } \overline{K} \cap \overline{K'} \neq \emptyset, \quad i = 1, 2.$$

3. The orientation of adjacent triangles must not change rapidly.

Define the matrices A_K and $C_K \in \mathbb{R}^{2 \times 2}$ by

$$A_K := (\overrightarrow{P_0P_1}, \overrightarrow{P_0P_2}) \quad \text{and} \quad C_K := (p_1, p_2)$$

and introduce an affine linear mapping

$$F_A(\mathbf{x}) := A_K \cdot \mathbf{x} + \overrightarrow{P_0}, \quad \mathbf{x} \in \mathbb{R}^2.$$

We will use the notion of the *reference triangle* $\hat{K} = F_A^{-1}(K)$ which is uniquely defined by the mapping F_A .

Furthermore, for any interior face $\gamma = \partial K \cap \partial K'$ define the quantity $h_{min,\gamma}$ by

$$h_{min,\gamma} := \frac{h_{min,K} + h_{min,K'}}{2}.$$

The advantage of this notation is clear, we need a value that relates to the edge, in contrast with others related to triangles. Note that $h_{min,\gamma} \sim h_{min,K} \sim h_{min,K'}$ due to the mesh assumptions. In addition we introduce the following notation:

$$\begin{aligned} \alpha_K &:= \min(\varepsilon^{-1} h_{min,K}, \kappa^{-1}), \\ \alpha_\gamma &:= \min(\varepsilon^{-1} h_{min,\gamma}, \kappa^{-1}). \end{aligned}$$

2.4 Bubble functions and their inverse inequalities

Now we introduce so-called bubble functions which are defined as usual, cf. [42]. They play an important role in deriving lower error bounds. Denote by $\lambda_{K,1}, \lambda_{K,2}, \lambda_{K,3}$ the barycentric coordinates of an arbitrary triangle K . The *element bubble function* b_K is defined by

$$b_K := 27\lambda_{K,1} \cdot \lambda_{K,2} \cdot \lambda_{K,3} \text{ on } K$$

Let $\gamma = \text{int}(\overline{K_1} \cap \overline{K_2})$ be an inner face (edge) of \mathcal{T}_h . Enumerate the vertices of K_1 and K_2 such that the vertices of γ are numbered first. Define the *edge bubble function* b_γ and the *edge spline function* s_γ by

$$\begin{aligned} b_\gamma &:= 4\lambda_{K_i,1}\lambda_{K_i,2} && \text{on } K_i, i = 1, 2, \\ s_\gamma &:= \frac{18}{\sqrt{3}}\lambda_{K_i,1}\lambda_{K_i,2}(\lambda_{K_i,2} - \lambda_{K_i,1}) && \text{on } K_i, i = 1, 2, \end{aligned}$$

with the obvious modification for a boundary face $\gamma \subset \partial\Omega$. For simplicity we assume that b_K , b_γ and s_γ are extended by zero outside their original domain of definition. It holds that $0 \leq b_K(x), b_\gamma(x) \leq 1$, $-1 \leq s_\gamma(x) \leq 1$ and $\|b_K\|_{L_\infty(K)} = \|b_\gamma\|_{L_\infty(K)} = \|s_\gamma\|_{L_\infty(K)} = 1$.

The following anisotropic equivalences/inverse inequalities can be derived easily.

Lemma 2.1. *Assume that $\varphi_K \in \mathbb{P}^0(K)$ and $\varphi_\gamma \in \mathbb{P}^0(\gamma)$. Then*

$$\|b_K^{1/2}\varphi_K\|_{L_2(K)} \sim \|\varphi_K\|_{L_2(K)} \quad (2.6)$$

$$\|b_K\varphi_K\|_{L_2(K)} \sim \|\varphi_K\|_{L_2(K)} \quad (2.7)$$

$$\|\nabla(b_K\varphi_K)\|_{L_2(K)} \preceq h_{\min,K}^{-1}\|\varphi_K\|_{L_2(K)} \quad (2.8)$$

$$\|C_K^\top \nabla(b_K\varphi_K)\|_{L_2(K)} \preceq \|\varphi_K\|_{L_2(K)} \quad (2.9)$$

$$\| \|b_K\varphi_K \| \|_K \preceq \alpha_K^{-1} \|\varphi_K\|_{L_2(K)} \quad (2.10)$$

$${}_m \| \|b_K\varphi_K \| \|_K \preceq \alpha_K^{-1} \|\varphi_K\|_{L_2(K)} \quad (2.11)$$

Proof. Inequalities (2.6) and (2.8) are copied from [31] (page 27, Lemma 2.7) without changes. Inequality (2.7) can be obtained analogously to (2.6). Inequality (2.9) is a refined version of (2.8) obtained by avoiding additional estimation on page 26, lines 3–4 from above in [31]. Indeed, using the inverse inequality on the reference triangle and the standard scaling argument we get:

$$\|C_K^\top \nabla(b_K\varphi_K)\|_{L_2(K)}^2 \sim 6|K| \|\hat{\nabla}(\hat{b}_K\varphi_K)\|_{L_2(\hat{K})}^2 \preceq \|\varphi_K\|_{L_2(K)}^2.$$

Another refined inequality (2.11) may be obtained in the same manner. Finally, combining (2.7) and (2.8) we show (2.10):

$$\begin{aligned} \| \|b_K\varphi_K \| \|_K^2 &= \varepsilon^2 \|\nabla(b_K\varphi_K)\|_{L_2(K)}^2 + \kappa^2 \|b_K\varphi_K\|_{L_2(K)}^2 \\ &\preceq (\varepsilon^2 h_{\min,K}^{-2} + \kappa^2) \|\varphi_K\|_{L_2(K)}^2 \preceq \alpha_K^{-2} \|\varphi_K\|_{L_2(K)}^2. \end{aligned}$$

□

2.5 Special edge bubble functions

Following [35], we define the squeezed edge bubble functions, extend the definition also for the squeezed spline functions and state the corresponding inverse inequalities. The definitions are given first for the reference triangle \hat{K} and then for the actual triangle K .

Consider the reference triangle \hat{K} and an edge $\hat{\gamma}$ thereof. Without loss of generality, assume that it lies on the axis $O\hat{y}$. By γ we denote the corresponding edge on the boundary of actual triangle K . For a real number $\delta \in (0, 1]$ define a linear mapping $F_\delta : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ by

$$F_\delta(\mathbf{x}) := B_\delta \cdot \mathbf{x} \quad \text{with } B_\delta = \begin{pmatrix} \delta & 0 \\ \frac{1-\delta}{2} & 1 \end{pmatrix} \in \mathbb{R}^{2 \times 2}.$$

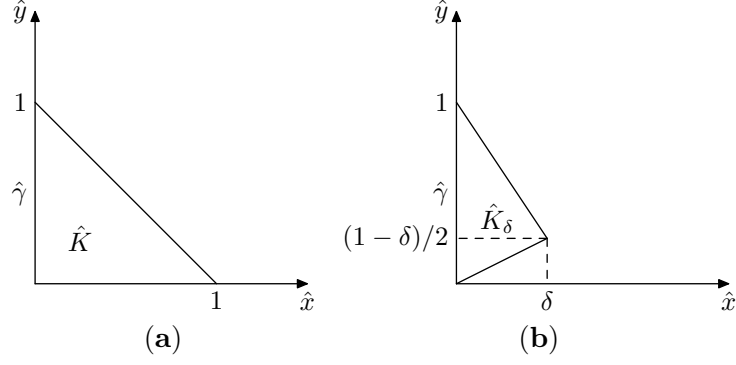


Figure 2: Definition of squeezed edge bubble functions: (a) – reference triangle \hat{K} , (b) – squeezed reference triangle \hat{K}_δ .

Set $\hat{K}_\delta := F_\delta(\hat{K})$, i.e it is the triangle with the edge $\hat{\gamma}$ and a vertex at $\left(\delta, \frac{1-\delta}{2}\right)$, see Figure 2.

Let $b_{\hat{\gamma}}$ be the usual edge bubble function of $\hat{\gamma}$ on \hat{K} . Define the squeezed bubble function $b_{\hat{\gamma},\delta}$ by $b_{\hat{\gamma},\delta} := b_{\hat{\gamma}} \circ F_\delta^{-1}$, i.e. $b_{\hat{\gamma},\delta}$ is the usual face bubble function of $\hat{\gamma}$ on the triangle \hat{K}_δ . For clarity we recall that $b_{\hat{\gamma},\delta} = 0$ on $\hat{K} \setminus \hat{K}_\delta$.

Consider now an actual triangle K . The squeezed edge bubble function $b_{\gamma,\delta} \in H^1(K)$ of an edge γ of K is defined by $b_{\gamma,\delta} := b_{\hat{\gamma},\delta} \circ F_A^{-1}$. Analogously we can define the squeezed edge spline function $s_{\gamma,\delta} := s_{\hat{\gamma},\delta} \circ F_\delta^{-1} \circ F_A^{-1}$. The actual value of parameter δ will be specified later. The usual and squeezed edge bubble/spline functions are drawn in Figure 3.

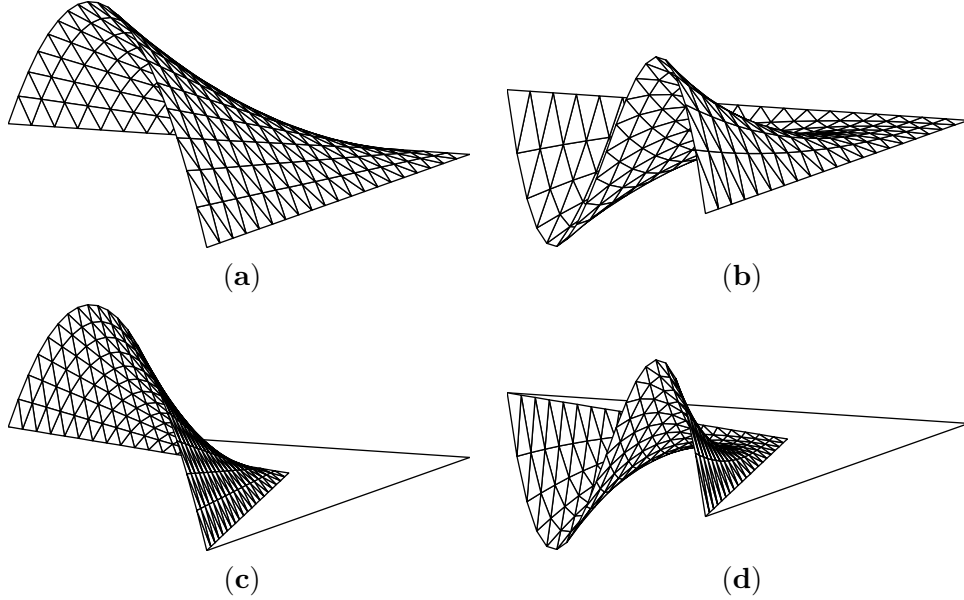


Figure 3: Bubble and spline functions: (a) – edge bubble function b_γ , (b) – edge spline function s_γ , (c) – squeezed edge bubble function $b_{\gamma,\delta}$, (d) – squeezed edge spline function $s_{\gamma,\delta}$.

The value for the parameter δ will differ for various applications and will be specified

when needed.

2.6 Refinement functions and their inverse inequalities

For the proof of the diverse convergence results we will utilize hat functions of the next refinement level and inverse inequalities/equivalences thereof.

From now on, while speaking about the refinement, we will conventionally use the notation \mathcal{T}_H for the course mesh and \mathcal{T}_h for the corresponding refined mesh. Consider a triangle $K \in \mathcal{T}_H$. The different ways of possible refinements are shown in Figure 4. We do not differ the hat functions corresponding to distinct resulting meshes \mathcal{T}_h since they have the same behavior regarding to the inverse inequalities. We can mention that there are cases when the refinement based on marking is not identical. The lack of uniqueness can be avoided in practical computations, e.g. by connection the center of the longest (marked) edge or of the edge with dominating error indicator with the opposite vertex. According to the author's experience the way of resolution of this lack of uniqueness does not essentially influence the numerical results.

We define by ψ_K the hat function on the triangulation \mathcal{T}_h associated with the node lying inside the triangle K , if such a function exists in \mathcal{T}_h . For the edge $\gamma \subset \partial K$ we analogously define a function ψ_γ as a hat function on the triangulation \mathcal{T}_h associated with the node lying on the edge γ .

The following anisotropic equivalences/inverse inequalities can be derived analogously to the corresponding edge and interior bubble functions in Section 2.4.

Lemma 2.2. (Inverse relations for refinement functions). *Assume that $\varphi_K \in \mathbb{P}^0(K)$ and $\varphi_\gamma \in \mathbb{P}^0(\gamma)$. Then*

$$\|\psi_K \varphi_K\|_{L_2(K)} = \|\varphi_K\|_{L_2(K)} / \sqrt{6} \quad (2.12)$$

$$\|\nabla(\psi_K \varphi_K)\|_{L_2(K)} \preceq h_{min,K}^{-1} \|\varphi_K\|_{L_2(K)} \quad (2.13)$$

$$\|F_{ext}(\varphi_\gamma) \psi_\gamma\|_{L_2(K)} \leq \frac{\sqrt{6}}{6} \left(\frac{|K|}{|\gamma|} \right)^{1/2} \|\varphi_\gamma\|_{L_2(\gamma)} \quad \text{for } \gamma \subset \partial K \quad (2.14)$$

$$\|\nabla(F_{ext}(\varphi_\gamma) \psi_\gamma)\|_{L_2(K)} \preceq \left(\frac{|K|}{|\gamma|} \right)^{1/2} h_{min,K}^{-1} \|\varphi_\gamma\|_{L_2(\gamma)} \quad \text{for } \gamma \subset \partial K \quad (2.15)$$

$$\|\psi_\gamma^{1/2} \varphi_\gamma\|_{L_2(\gamma)} \sim \|\varphi_\gamma\|_{L_2(\gamma)}. \quad (2.16)$$

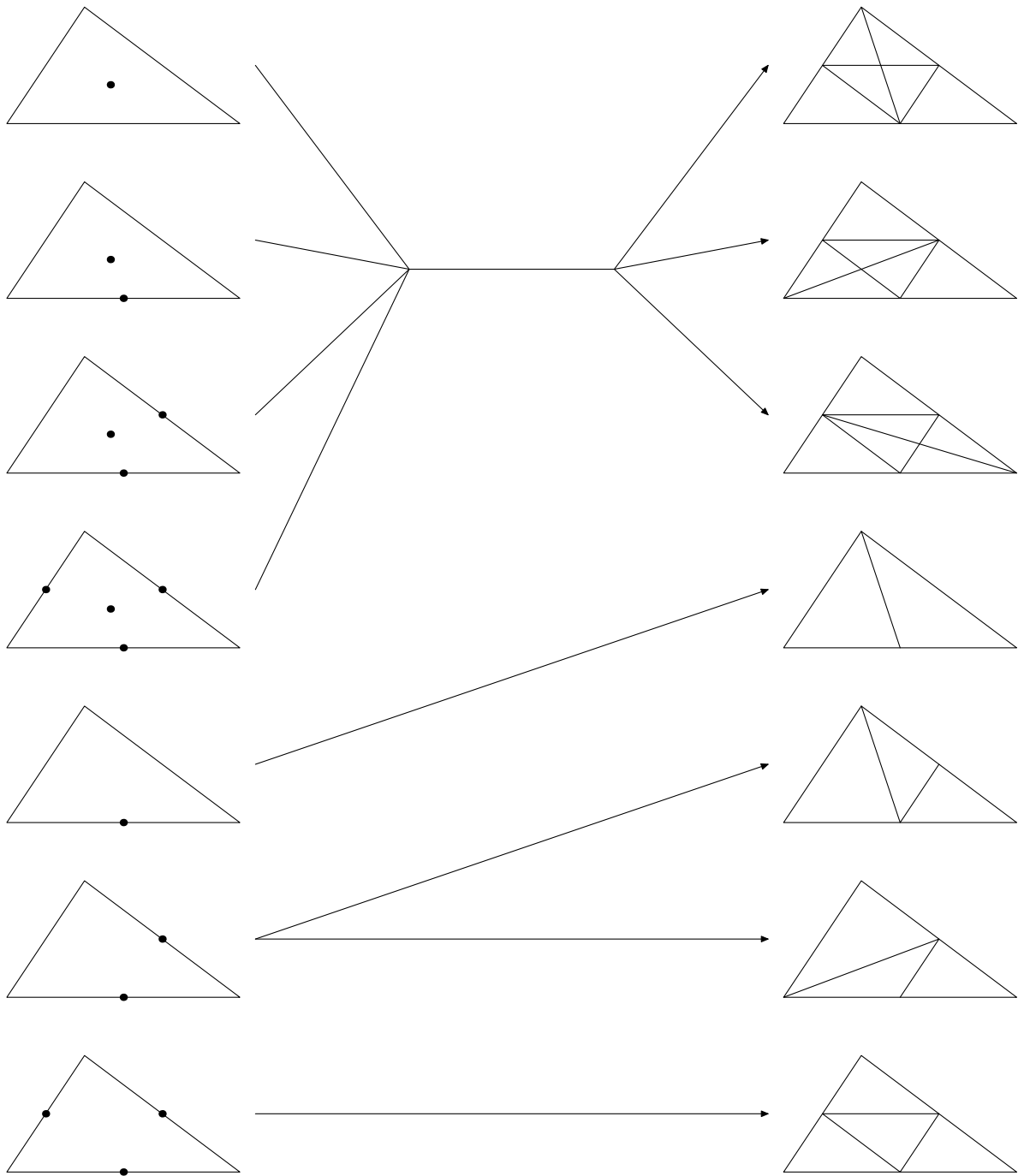


Figure 4: The possible triangle refinement possibilities according to the marked entities.

3 The strengthened Cauchy-Schwarz Inequality

The strengthened Cauchy-Schwarz Inequality plays an important role in the analysis of the hierarchical error estimator as well as in some other parts of this work. In the case of the hierarchical error estimator we enrich the space V_1 by the squeezed bubble functions for all edges and the interior bubbles. Namely, we introduce the space

$$V_2 := \{v_h \in H_0^1(\Omega) : \forall K \in \mathcal{T}, v_h|_K \in P_1(K) + \text{span}\{b_K, b_{\gamma, \delta_\gamma} : \gamma \in \partial K \setminus \partial\Omega\}\} = V_1 \oplus \tilde{V}_2,$$

where \tilde{V}_2 is the space spanned by the interior bubble functions for each triangle and the squeezed edge bubble functions for each edge.

3.1 Theoretical background

Definition 3.1. *Let X, Y be two subspaces of a Hilbert space equipped with a scalar product (\cdot, \cdot) and induced norm $\|\cdot\|$. A strengthened Cauchy-Schwarz Inequality is said to hold for this couple of spaces if there exist a non-negative constant $\gamma < 1$ such that:*

$$(x, y) \leq \gamma \|x\| \|y\|, \quad \forall x \in X, y \in Y. \quad (3.1)$$

Let X, Y be finite dimensional spaces. Consider a stiffness matrix \mathbf{B} corresponding to the space $X^* = X \oplus Y$,

$$\mathbf{B} = \begin{bmatrix} \mathbf{B}_{XX} & \mathbf{B}_{XY} \\ \mathbf{B}_{YX} & \mathbf{B}_{YY} \end{bmatrix}.$$

We state here without a proof the following theorem from [4].

Theorem 3.2. *The constant γ in the Cauchy-Schwarz inequality 3.1 may be expressed in the following way:*

$$\gamma^2 = \max_{\mathbf{x} \in \mathbb{R}^{\dim(X)}} \frac{\mathbf{x}^T \mathbf{B}_{XY} \mathbf{B}_{YY}^{-1} \mathbf{B}_{YX} \mathbf{x}}{\mathbf{x}^T \mathbf{B}_{XX} \mathbf{x}}$$

Now come back to our problem. Suppose that for each element K of triangulation the constant γ_K is known:

$$B_K(u, v) \leq \gamma_K \|u\|_K \|v\|_K, \quad \forall u \in X_K, v \in Y_K,$$

where X_K and Y_K are restrictions of corresponding spaces to the element K . Now, prescribing $\gamma = \max_K \gamma_K$, we obtain the constant γ for the whole mesh:

$$B(u, v) = \sum_K B_K(u, v) \leq \gamma \sum_K \|u\|_K \|v\|_K \leq \gamma \|u\| \cdot \|v\|, \quad (3.2)$$

where we utilized the discrete Cauchy-Schwarz inequality.

3.2 Pure Laplace problem $\kappa = 0, \varepsilon = 1$

We state this result here because it could be used in other applications. In the case of a pure Laplace problem the Cauchy-Schwarz constant has a nice structure and may be expressed explicitly (see below). We assume here that $\delta_\gamma = 1$, for all edges γ in triangulation. In other words \tilde{V}_2 is the space spanned by three usual edge bubble functions together with interior bubble functions for each triangle. Recall that $V_2 = V_1 \oplus \tilde{V}_2$.

Let a triangle have angles ϕ_1, ϕ_2, ϕ_3 . After straight forward maple calculations we get the matrices needed to obtain the strengthened Cauchy-Schwarz constant for H_1 semi-norm:

$$B_{V_1 V_1} = \frac{1}{2} \begin{pmatrix} \frac{\sin \phi_1}{\sin \phi_2 \sin \phi_3} & -\cot \phi_3 & -\cot \phi_2 \\ -\cot \phi_3 & \frac{\sin \phi_2}{\sin \phi_3 \sin \phi_1} & -\cot \phi_1 \\ -\cot \phi_2 & -\cot \phi_1 & \frac{\sin \phi_3}{\sin \phi_1 \sin \phi_2} \end{pmatrix},$$

$$B_{\tilde{V}_2 \tilde{V}_2} = \frac{1}{2} \begin{pmatrix} D & -\cot \phi_3 & -\cot \phi_2 & \frac{\cot \phi_1}{5} \\ -\cot \phi_3 & D & -\cot \phi_1 & \frac{\cot \phi_2}{5} \\ -\cot \phi_2 & -\cot \phi_1 & D & \frac{\cot \phi_3}{5} \\ \frac{\cot \phi_1}{5} & \frac{\cot \phi_2}{5} & \frac{\cot \phi_3}{5} & \frac{D}{15} \end{pmatrix},$$

$$\text{where } D = \frac{\cos \phi_1 \cos \phi_2 \cos \phi_3 + 1}{\sin \phi_1 \sin \phi_2 \sin \phi_3}.$$

$$B_{\tilde{V}_2 V_1} = -\frac{1}{6} \begin{pmatrix} \frac{\sin \phi_1}{\sin \phi_2 \sin \phi_3} & -\cot \phi_3 & -\cot \phi_2 \\ -\cot \phi_3 & \frac{\sin \phi_2}{\sin \phi_3 \sin \phi_1} & -\cot \phi_1 \\ -\cot \phi_2 & -\cot \phi_1 & \frac{\sin \phi_3}{\sin \phi_1 \sin \phi_2} \\ 0 & 0 & 0 \end{pmatrix},$$

$$B_{V_1 \tilde{V}_2} = B_{\tilde{V}_2 V_1}^T.$$

Compute the matrix $L = B_{V_1 \tilde{V}_2} B_{\tilde{V}_2 \tilde{V}_2}^{-1} B_{\tilde{V}_2 V_1} \in \mathbb{R}^{3 \times 3}$:

$$\begin{pmatrix} A_{1,2,3} & B_{3,1,2} & B_{2,3,1} \\ B_{3,1,2} & A_{2,3,1} & B_{1,2,3} \\ B_{2,3,1} & B_{1,2,3} & A_{3,1,2} \end{pmatrix},$$

where we used the notation:

$$A_{i,j,k} := \frac{(2 - \cos^2 \phi_i - 2 \cos \phi_i \cos \phi_j \cos \phi_k) \sin \phi_i}{\sin \phi_j \sin \phi_k},$$

$$B_{i,j,k} := \frac{\cos \phi_j \cos \phi_k - 2 \cos \phi_i}{\sin \phi_i}.$$

What remains is to find

$$\gamma^2 = \max_{\mathbf{x} \in \mathbb{R}^3} \frac{\mathbf{x}^T L \mathbf{x}}{\mathbf{x}^T B_{V_1 V_1} \mathbf{x}}.$$

If $B_{V_1 V_1}$ was a non-singular matrix then we would solve an eigenvalue problem $L \mathbf{x} = \lambda B_{V_1 V_1} \mathbf{x}$ and maximal eigenvalue λ_{max} would be our wanted constant γ^2 . However, it is not possible in this case since $B_{V_1 V_1}$ has a vector $e = [1, 1, 1]^T$ in its kernel. It turns out that e lies in the kernel of the matrix L as well. There exist two vectors e_1 and e_2

such that $\{e, e_1, e_2\}$ is the basis in \mathbb{R}^3 . We have

$$\begin{aligned}\gamma^2 &= \max_{\alpha, \beta_1, \beta_2 \in \mathbb{R}} \frac{(\alpha e + \beta_1 e_1 + \beta_2 e_2)^T L (\alpha e + \beta_1 e_1 + \beta_2 e_2)}{(\alpha e + \beta_1 e_1 + \beta_2 e_2)^T B_{V_1 V_1} (\alpha e + \beta_1 e_1 + \beta_2 e_2)} \\ &= \max_{\alpha, \beta_1, \beta_2 \in \mathbb{R}} \frac{(\beta_1 e_1 + \beta_2 e_2)^T L (\beta_1 e_1 + \beta_2 e_2)}{(\beta_1 e_1 + \beta_2 e_2)^T B_{V_1 V_1} (\beta_1 e_1 + \beta_2 e_2)} \\ &= \max_{\beta \in \mathbb{R}^2} \frac{\beta^T E^T L E \beta}{\beta^T E^T B_{V_1 V_1} E \beta},\end{aligned}$$

where the matrix $E = (e_1, e_2) \in \mathbb{R}^{3 \times 2}$ can be chosen in such a way that $\{e, e_1, e_2\}$ is the basis in \mathbb{R}^3 . Choosing for simplicity

$$E := \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix},$$

and solving eigenvalue problem $E^T L E \beta = \lambda E^T B_{V_1 V_1} E \beta$ we get two eigenvalues

$$\beta_{1,2} = \frac{1}{2} \pm \frac{1}{6} \sqrt{1 - 8 \cos \phi_1 \cos \phi_2 \cos \phi_3}.$$

Choosing the maximal of these two numbers we formulate the following lemma.

Theorem 3.3. *The constant γ in the strengthened Cauchy-Schwarz inequality for the spaces V_1 and \tilde{V}_2 is expressed by*

$$\gamma^2 = \frac{1}{2} + \frac{1}{6} \sqrt{1 - 8 \cos \phi_1 \cos \phi_2 \cos \phi_3}.$$

Proof. For the proof see above arguments. □

In the inequality of Theorem 3.3 maximum angle condition naturally appears in the sense that, if one angle of a triangle goes to π then the strengthened Cauchy-Schwarz constant goes to 1. It leads to the fact that induced hierarchical error estimator will fail in general on meshes where the maximum angle condition is not satisfied.

3.3 Squeezed case

Divide a triangle K into three parts $K_{\frac{1}{3}}$, where the central point is the center of mass. Mention that support of each special edge bubble function lies inside exactly one part $K_{\frac{1}{3}}$. Evaluate (estimate from above) the Cauchy-Schwarz constant for each part independently assuming that the squeezing parameter δ of the special edge bubble function can be any number from 0 to $1/3$. The constant for the whole triangle and subsequently for the whole mesh may be chosen as the maximum value of three corresponding constants as we did in (3.2).

Lemma 3.4. *Let $K \in \mathcal{T}_h$ and $K_{\frac{1}{3}}$ be a part thereof described above (see Figure 5). Then the following estimate for the H^1 -semi-norm holds*

$$(u, v)_{H^1(K_{\frac{1}{3}})} \leq \frac{2\sqrt{2}}{3} |u|_{H^1(K_{\frac{1}{3}})} |v|_{H^1(K_{\frac{1}{3}})}, \quad \forall u \in V_1, v \in \tilde{V}_2.$$

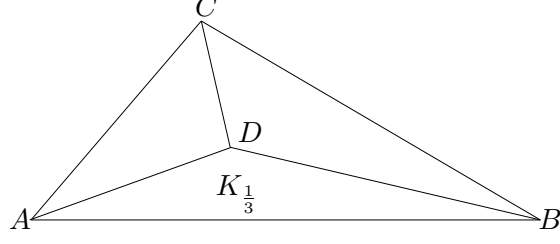


Figure 5: Notation of a triangle, the vertices have the following coordinates: $A = (0, 0)$, $B = (a, 0)$, $C = ((a + q)/2, b)$, $D = (A + B + C)/3$.

Proof. Write down matrices in notation from Figure 2.

$$B_{V_1 V_1} = \begin{pmatrix} \frac{4b^2 + (1-q)^2}{24b} & -\frac{4b^2 - 1 + q^2}{24b} & -\frac{1-q}{12b} \\ -\frac{4b^2 - 1 + q^2}{24b} & \frac{4b^2 + (1-q)^2}{24b} & -\frac{1+q}{12b} \\ -\frac{1-q}{12b} & -\frac{1+q}{12b} & \frac{1}{6b} \end{pmatrix},$$

$$B_{\tilde{V}_2 \tilde{V}_2} = \begin{pmatrix} \frac{\delta^2 q^2 + 4\delta^2 b^2 + 3}{48b\delta} & \frac{\delta^2 q^2 - 3\delta^2 + 4\delta^2 b^2 + 10\delta - 8}{240b} \\ \frac{\delta^2 q^2 - 3\delta^2 + 4\delta^2 b^2 + 10\delta - 8}{240b} & \frac{q^2 + 4b^2 + 51}{19440b} \end{pmatrix},$$

$$B_{\tilde{V}_2 V_1} = \begin{pmatrix} \frac{1-q}{12b} & \frac{1+q}{12b} & -\frac{1}{6b} \\ -\frac{1-q}{108b} & -\frac{1+q}{108b} & \frac{1}{54b} \end{pmatrix}.$$

Compute the matrix $L = B_{V_1 \tilde{V}_2} B_{\tilde{V}_2 \tilde{V}_2}^{-1} B_{\tilde{V}_2 V_1} \in \mathbb{R}^{3 \times 3}$ and using the same trick as in the proof of Theorem 3.3 we solve eigenvalue problem $E^T L E \beta = \lambda E^T B_{V_1 V_1} E \beta$. One of two eigenvalues of this problem equals zero, the second one is the constant of interest:

$$\begin{aligned} \gamma^2 &= \frac{8}{9} - \frac{8}{9}[-90 - 4860\delta^4 + 9558\delta^3 + 729\delta^5 - 4860\delta^2 + 999\delta \\ &+ (-5\delta^2 + 81\delta^5)q^4 \\ &+ (-1944\delta^3 + 6480\delta^4 + 180\delta - 60 - 1944\delta^5 - 120\delta^2)b^2 \\ &+ (-30\delta^2 + 1620\delta^4 - 15 + 45\delta - 486\delta^5 - 486\delta^3)q^2 \\ &+ (1296\delta^5 - 80\delta^2)b^4 + (648\delta^5 - 40\delta^2)q^2 b^2] \\ &/ [5184\delta - 12960\delta^2 - 4860\delta^4 - 765 + 729\delta^5 + 11988\delta^3 \\ &+ (-5\delta^2 + 81\delta^5)q^4 + (6480\delta^4 - 60 - 1020\delta^2 - 5184\delta^3 - 1944\delta^5)b^2 \\ &+ (-1296\delta^3 - 255\delta^2 + 1620\delta^4 - 486\delta^5 - 15)q^2 + (1296\delta^5 - 80\delta^2)b^4 + (648\delta^5 - 40\delta^2)q^2 b^2] \end{aligned}$$

Observing that for $\delta \in [0, \frac{1}{3}]$

$$\begin{aligned}
& -90 - 4860\delta^4 + 9558\delta^3 + 729\delta^5 - 4860\delta^2 + 999\delta \\
& = 9(3\delta - 1) \left(243 \left(\delta - \frac{1}{6} \right)^2 + \frac{13}{4} + 9\delta^2(3\delta - 1)(\delta - 6) \right) \leq 0 \\
& \quad -5\delta^2 + 81\delta^5 = \delta^2(81\delta^3 - 5) \leq -2\delta^2 \leq 0 \\
& \quad -1944\delta^3 + 6480\delta^4 + 180\delta - 60 - 1944\delta^5 - 120\delta^2 \\
& = -\frac{40}{3} - \frac{4}{3}(3\delta - 1) \left(5(15\delta + 7)(3\delta - 1) + 9\delta^2 \left(54 \left(\frac{3}{2} - \delta \right)^2 - \frac{293}{2} \right) \right) \\
& \quad \leq -\frac{40}{3} - \frac{4}{3}(3\delta - 1) (5(15\delta + 7)(3\delta - 1) - 25 \cdot 9\delta^2) \leq 0 \\
& \quad \quad -30\delta^2 + 1620\delta^4 - 15 + 45\delta - 486\delta^5 - 486\delta^3 \\
& = -\frac{10}{3} - \frac{1}{3}(3\delta - 1) \left(5(15\delta + 7)(3\delta - 1) + 9\delta^2 \left(54 \left(\frac{3}{2} - \delta \right)^2 - \frac{293}{2} \right) \right) \leq 0 \\
& \quad \quad 1296\delta^5 - 80\delta^2 = 16\delta^2(81\delta^3 - 5) \leq 0 \\
& \quad \quad 648\delta^5 - 40\delta^2 = 8\delta^2(81\delta^3 - 5) \leq 0 \\
& \quad 5184\delta - 12960\delta^2 - 4860\delta^4 - 765 + 729\delta^5 + 11988\delta^3 \\
& = 9(3\delta - 1) \left(477 \left(\delta - \frac{107}{318} \right)^2 + \frac{6571}{212} \right) + 81\delta^3 \left(9 \left(\frac{10}{3} - \delta \right)^2 - 111 \right) \leq 0 \\
& \quad \quad -5\delta^2 + 81\delta^5 \leq 0 \\
& \quad \quad 6480\delta^4 - 60 - 1020\delta^2 - 5184\delta^3 - 1944\delta^5 \\
& = -60 - 12\delta^2 \left(85 + \delta \left(162 \left(\frac{5}{3} - \delta \right)^2 - 18 \right) \right) \leq 0 \\
& \quad \quad -1296\delta^3 - 255\delta^2 + 1620\delta^4 - 486\delta^5 - 15 \\
& = -15 - 3\delta^2 \left(85 + \delta \left(162 \left(\frac{5}{3} - \delta \right)^2 - 18 \right) \right) \leq 0 \\
& \quad \quad 1296\delta^5 - 80\delta^2 \leq 0 \\
& \quad \quad 648\delta^5 - 40\delta^2 \leq 0
\end{aligned}$$

we finish the proof. □

Lemma 3.5. *Let $K \in \mathcal{T}_h$ and $K_{\frac{1}{3}}$ be a part thereof described above (see Figure 5). Then the following estimate for the L_2 -norm holds*

$$(u, v)_{L_2(K_{\frac{1}{3}})} \leq \sqrt{\frac{31927}{35680} + \frac{7\sqrt{193953}}{35680}} \|u\|_{L_2(K_{\frac{1}{3}})} \|v\|_{L_2(K_{\frac{1}{3}})}, \quad \forall u \in V_1, v \in \tilde{V}_2.$$

The approximate value of the Cauchy-Schwarz constant is given by

$$\sqrt{\frac{31927}{35680} + \frac{7\sqrt{193953}}{35680}} \approx 0.9905637561 < 1.$$

Proof. The matrices in the L_2 case has the following form:

$$B_{V_1 V_1} = a^2 b \begin{pmatrix} \frac{13}{324} & \frac{17}{648} & \frac{5}{648} \\ \frac{17}{648} & \frac{13}{324} & \frac{5}{648} \\ \frac{5}{648} & \frac{5}{648} & \frac{1}{324} \end{pmatrix},$$

$$B_{\tilde{V}_2 \tilde{V}_2} = a^2 b \begin{pmatrix} \frac{\delta}{180} & \frac{\delta^2(3\delta^2 - 14\delta + 19)}{10080} \\ \frac{\delta^2(3\delta^2 - 14\delta + 19)}{10080} & \frac{1}{15120} \end{pmatrix},$$

$$B_{\tilde{V}_2 V_1} = a^2 b \begin{pmatrix} \frac{\delta(5-\delta)}{240} & \frac{\delta(5-\delta)}{240} & \frac{\delta^2}{120} \\ \frac{17}{14580} & \frac{17}{14580} & \frac{13}{29160} \end{pmatrix}.$$

Computing the matrix L and then solving the eigenvalue problem $Lx = \lambda B_{V_1 V_1} x$ we get three eigenvalues, from which one is zero and the maximal is

$$\begin{aligned} \gamma^2 &= \frac{1}{2}(-1876770\delta^4 + 408240\delta^5 + 3241350\delta^2 - 302400\delta^3 - 221480 - 1275750\delta \\ &- 70(612301545\delta^2 - 114647400\delta + 120105666\delta^{10} - 1254436956\delta^9 \\ &+ 4978952631\delta^8 - 9594124056\delta^7 + 10853696088\delta^6 - 8279204886\delta^5 \\ &+ 4695055083\delta^4 - 2017608642\delta^3 + 10010896)^{1/2}) \\ &/ (54675\delta^7 - 510300\delta^6 + 1883250\delta^5 - 3231900\delta^4 + 2193075\delta^3 - 226800). \end{aligned}$$

In Figure 6 we can see dependence of the strengthened Cauchy-Schwarz constant γ on

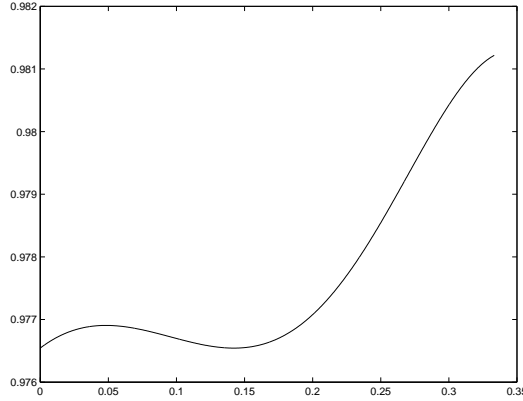


Figure 6: dependence of the strengthened Cauchy-Schwarz constant on the parameter δ .

the parameter δ . The maximum is reached for $\delta = 1/3$ and is equal to the aforementioned constant. \square

Theorem 3.6. Let $K \in \mathcal{T}_h$ and $K_{\frac{1}{3}}$ be a part thereof described above (see Figure 5). Then the following estimate for the energy norm holds

$$B_{K_{\frac{1}{3}}}(u, v) \leq \sqrt{\frac{31927}{35680} + \frac{7\sqrt{193953}}{35680}} \|u\|_{K_{\frac{1}{3}}} \|v\|_{K_{\frac{1}{3}}}, \quad \forall u \in V_1, v \in \tilde{V}_2.$$

Proof. This is an immediate consequence of the previous two lemmas. \square

Corollary 3.7. *The global strengthened Cauchy-Schwarz Inequality holds in the form*

$$B(u, v) \leq \sqrt{\frac{31927}{35680} + \frac{7\sqrt{193953}}{35680}} \|u\| \cdot \|v\|, \quad \forall u \in V_1, v \in \tilde{V}_2.$$

Proof. For the proof see (3.2). \square

We will also need the Cauchy-Schwarz constant between the space of edge bubble functions and the space of interior bubble functions. Introduce the following notation:

$$\begin{aligned} V_{eb} &= \bigoplus_{\gamma \in \partial T} \text{span}\{b_{\gamma, \delta_\gamma}\}, \\ V_{ib} &= \bigoplus_{K \in \mathcal{T}} \text{span}\{b_K\}. \end{aligned}$$

We express the resulting strengthened Cauchy-Schwarz inequality in the following theorem.

Theorem 3.8. *Let the spaces V_{eb} and V_{ib} be defined as above. Then*

$$B_K(u, v) \leq \frac{2\sqrt{2}}{3} \|u\|_K \|v\|_K, \quad \forall u \in V_{eb}, v \in V_{ib}.$$

Proof. Prove this inequality for $K_{\frac{1}{3}}$. The statement of the lemma will be a direct consequence.

1. Verify first the inequality for H_1 semi-norm and corresponding scalar product. The matrices for computing the strengthened Cauchy-Schwarz constant are now reduced to scalars. Therefore we simply state the constant and estimate it from above.

$$\begin{aligned} \gamma^2 &= \frac{8}{9} - \frac{1}{45} [(40\delta^2 - 729\delta^5)(q^2 + 4b^2)^2 \\ &+ (4374\delta^5 - 14580\delta^4 + 120 + 11664\delta^3 + 2040\delta^2)(q^2 + 4b^2) \\ &- 6561\delta^5 + 6120 - 107892\delta^3 + 43740\delta^4 + 116640\delta^2 - 46656\delta] \\ &/ (3 + \delta^2(q^2 + 4b^2))(q^2 + 4b^2 + 51) \end{aligned}$$

Observing that for $\delta \in [0, \frac{1}{3}]$

$$\begin{aligned} 40\delta^2 - 729\delta^5 &= \delta^2(40 - 729\delta^3) \geq 13\delta^2 \geq 0 \\ 4374\delta^5 - 14580\delta^4 + 120 + 11664\delta^3 + 2040\delta^2 & \\ &= (120 + 7290\delta^3 + 2040\delta^2) + 1458\delta^3(3\delta - 1)(\delta - 3) \geq 0 \\ -6561\delta^5 + 6120 - 107892\delta^3 + 43740\delta^4 + 116640\delta^2 - 46656\delta & \\ \left\{ 5 \left[213 + 18\sqrt{105} + (1 - 3\delta)(60 + 2\sqrt{105} - 15\delta) \right] (1 - 3\delta) + 1755 + 204\sqrt{105} \right\} & \\ \times \frac{27}{625} (15\delta - 15 + \sqrt{105})^2 + \frac{15048}{5} - \frac{36288}{125} \sqrt{105} &\geq \frac{15048}{5} - \frac{36288}{125} \sqrt{105} \approx 34.8 \geq 0 \end{aligned}$$

we conclude

$$(u, v)_{H^1(K_{\frac{1}{3}})} \leq \frac{2\sqrt{2}}{3} |u|_{H^1(K_{\frac{1}{3}})} |v|_{H^1(K_{\frac{1}{3}})}, \quad \forall u \in V_{eb}, v \in V_{ib}. \quad (3.3)$$

2. It remains to verify the inequality for L_2 scalar product. In this case the constant of interest has the following structure:

$$\gamma^2 = \frac{3}{112} \delta^3 (3\delta^2 - 14\delta + 19)^2.$$

It is easy to verify that the maximum is reached for $\delta = 1/3$ and is equal to $121/567$. Thus,

$$(u, v)_{L_2(K_{\frac{1}{3}})} \leq \sqrt{\frac{121}{567}} \|u\|_{L_2(K_{\frac{1}{3}})} \|v\|_{L_2(K_{\frac{1}{3}})}, \quad \forall u \in V_{eb}, v \in V_{ib}. \quad (3.4)$$

Combining (3.3) and (3.4) and choosing the maximal constant among two we get the result claimed. \square

3.4 Additional pair of spaces

In this subsection we investigate the strengthened Cauchy-Schwarz constant that will be used to prove the error reduction theorem for the adaptive algorithm for the case of singularly perturbed problem in Section 7.3. Consider an inner edge e together with two triangles K_1, K_2 sharing this edge. Define the space W in the following way

$$W = \{w \in C(K_1 \cup K_2) : w = 0 \text{ on } K_2, w \in P_1(K_1)\}. \quad (3.5)$$

The second space V is expressed via

$$V = P_1(K_1 \cup K_2) + \prod_{i=1}^2 P_0(K_i). \quad (3.6)$$

Lemma 3.9. *There is a constant $\tilde{C} < 1$, such that for any inner edge e the Strengthened Cauchy-Schwarz inequality*

$$(v, w)_{L_2(K_1 \cup K_2)} \leq \gamma \|v\|_{L_2(K_1 \cup K_2)} \|w\|_{L_2(K_1 \cup K_2)}, \quad \forall v \in V, w \in W$$

holds, where $\gamma \leq \tilde{C} < 1$.

Proof. Place the coordinate system so, that the edge e lies on the axis Oy and the origin is exactly at the center of this edge (see Figure 7). Moreover, by homogeneity argument, it is sufficient to make the proof for $|e| = 1$.

Proceeding as before, according to the theory from Section 3.1, we get the expression for the Cauchy-Schwarz constant in dependance on parameters from Figure 7.

$$\gamma^2 = 1 - \frac{\alpha^3(\alpha + 1)}{3\alpha^4 + 3\alpha^3 + 3\alpha + 3 + 4\alpha(\alpha b_1 + b_2)^2},$$

where $\alpha = h_2/h_1$. Due to the requirements to the mesh described in Section 2.3 $\alpha = O(1)$ and $|\alpha b_1 + b_2|$ is bounded from above. Thus, if the mesh requirements are satisfied the Cauchy-Schwarz constant is bounded from above by some constant $\tilde{C} < 1$. \square

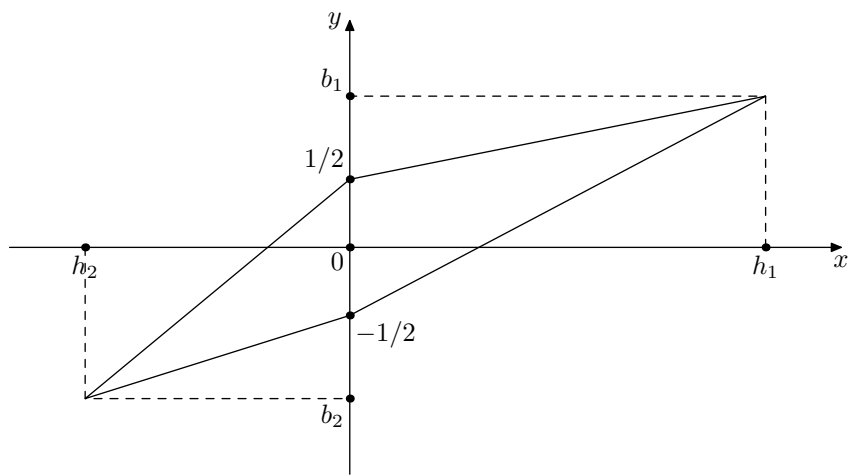


Figure 7: Notation of two neighboring triangles.

4 Convergence results for Poisson problem on isotropic meshes

In this chapter we consider the pure Poisson problem ($\varepsilon = 1$, $\kappa = 0$) and introduce some convergent adaptive algorithms on two-dimensional domains Ω . We restrict ourselves to isotropic meshes, which in terms of Section 2.3 means

$$\sup_{T_h \in \mathcal{F}} \max_{K \in \mathcal{T}_h} \frac{h_{max,K}}{h_{min,K}} \leq C < \infty,$$

with a moderate constant C . The anisotropic case will be considered in the next chapter.

In regard to the converging algorithms, we should mention the work by Babuška and Rheinboldt [13], where the one dimensional elliptic equation was considered and convergence for an adaptive procedure was shown. The authors used there the exact solution representation for controlling the error. However, the computable approximation of the error arising from the a posteriori error estimation also provides sufficient information for our convergent algorithm as we show in the current chapter.

The technique of this chapter was developed by Dörfler in his pioneering work [23]. We recall the residual a posteriori error estimator in Section 4.1 and proceed with his results in Section 4.2. In [23] a kind of preadaptation of the mesh before applying the algorithm is needed. This preadaptation ensures the sufficient resolution of the input data on the initial mesh. A further modification of his approach, without additional preadaptation, is done by Morin, Nochetto and Siebert in [40]. We describe the main differences in Section 4.3.

Further in work, in Section 4.4 with proofs given in Section 4.5, we discuss a possibility of avoiding a new inner node among some of the marked elements. This possibility stems from a more precise analysis of the information provided by local error indicators. Moreover, for the Poisson problem we can fully avoid inner nodes as we describe in Section 4.6. The resulting algorithm converges as soon as we make some additional assumptions on the mesh.

4.1 Isotropic residual error estimator

We recall shortly the model problem (2.1) with $\kappa = 0$ and $\varepsilon = 1$. For any $f \in L_2(\Omega)$ there exists a unique weak solution u of the following problem:

$$u \in H_0^1(\Omega) : \quad (\nabla u, \nabla v)_{L_2(\Omega)} = (f, v)_{L_2(\Omega)}, \quad \forall v \in H_0^1(\Omega). \quad (4.1)$$

Let V_h be a space of continuous piecewise linear functions over \mathcal{T}_h that vanish at the boundary. Let u_h denote the solution of the discrete problem

$$u_h \in V_h : \quad (\nabla u_h, \nabla v_h)_{L_2(\Omega)} = (f, v_h)_{L_2(\Omega)}, \quad \forall v_h \in V_h. \quad (4.2)$$

Denote by f_K the mean value of the function f on element $K \in \mathcal{T}_h$:

$$f_K = \frac{1}{|K|} \int_K f.$$

The jump discontinuity in the approximation of the normal flux at an interelement boundary $\gamma = \partial K \cap \partial K'$ is defined by

$$\left[\frac{\partial u_h}{\partial n} \right]_{\gamma} = n_K \cdot (\nabla u_h)_K + n_{K'} \cdot (\nabla u_h)_{K'},$$

and the usual boundary residual R is given by

$$R_\gamma = \begin{cases} \left[\frac{\partial u_h}{\partial n} \right]_\gamma & \text{for } \gamma \notin \partial\Omega \\ 0 & \text{for } \gamma \subset \partial\Omega \end{cases}$$

which are defined as in e.g. book by Ainsworth and Oden [4].

We recall now a residual a posteriori error estimator for (4.1) and (4.2) according to the work by Babuška and Miller [11] and the book by Verfürth [47]. Define the error indicator η_γ associated with an edge γ by

$$\eta_\gamma := |\gamma|^{1/2} \|R_\gamma\|_{L_2(\gamma)},$$

and the error indicator η_K associated with an element K by

$$\eta_K := |K|^{1/2} \|f_K\|_{L_2(K)}.$$

Then the local residual error estimator $\eta_{R,K}$ and the local data oscillation (local approximation term) ζ_K for the element K are defined via

$$\begin{aligned} \eta_{R,K}^2 &:= \eta_K^2 + \frac{1}{2} \sum_{\gamma \subset \partial K} \eta_\gamma^2, \\ \zeta_K^2 &:= h_K^2 \|f - f_K\|_{L_2(K)}^2. \end{aligned}$$

Let η_R and ζ be their global counterparts, given by

$$\begin{aligned} \eta_R^2 &:= \sum_{K \in \mathcal{T}_h} \eta_{R,K}^2, \\ \zeta^2 &:= \sum_{K \in \mathcal{T}_h} \zeta_K^2. \end{aligned}$$

The following robustness result is well known [11, 47].

Theorem 4.1 (Global upper error bound). *There exists a constant C depending only on the minimum angle of \mathcal{T}_h , such that*

$$\|\nabla(u - u_h)\|_{L_2(\Omega)}^2 \preceq \eta_R^2 + \zeta^2.$$

Theorem 4.2 (Local lower error bound). *There exists a constant C depending only on the minimum angle of \mathcal{T}_h , such that*

$$\eta_{R,K}^2 \preceq \|\nabla(u - u_h)\|_{L_2(\omega_K)}^2 + \sum_{K \subset \omega_K} \zeta_K^2.$$

4.2 The convergent algorithm due to Dörfler

In this section we repeat the main results of [23]. We do not give the proof for Theorem 4.3 since we prove the analogous Theorem 4.8 for the new modification of the approach in Section 4.5. As it was already described in the introductory Section 1.2 the adaptive procedure involves usually *a marking strategy* and *a refinement strategy*. For the method in the current section an initial triangulation also plays an important role. The mesh \mathcal{T}_h is said to have *fineness* μ with respect to χ if

$$\left(\sum_{K \in \mathcal{T}_H} h_K^2 \|f\|_{L_2(K)}^2 \right)^{1/2} \leq \mu \chi. \quad (4.3)$$

Mesh fineness is the condition on the data approximation that we require to be satisfied for the initial mesh. We assume first that the initial mesh satisfying (4.3) is given. In the next section, in Remark 4.6, we will describe a rigorous way of creating an appropriate, “fine” mesh from any given macro triangulation.

We present now the marking strategy from [23] and then the corresponding error reduction theorem. We do not discuss in detail the refinement strategy but refer to [46] for the appropriate algorithm using so-called red, green and blue refinement.

Marking Strategy B

Given a parameter $0 < \theta < 1$ and a mesh \mathcal{T}_H , mark a set of elements $\hat{\mathcal{T}}_H \subset \mathcal{T}_H$ for which

$$\left(\sum_{K \in \hat{\mathcal{T}}_H} \frac{1}{2} \sum_{\gamma \subset \partial K} \eta_\gamma^2 \right)^{1/2} \geq \theta \left(\sum_{\gamma \in \mathcal{E}_H} \eta_\gamma^2 \right)^{1/2} .$$

The Marking Strategy B ensures that we mark sufficiently many elements $\hat{\mathcal{T}}_h$ with total contribution from edge associated indicators proportional to $\left(\sum_{\gamma \in \mathcal{E}_H} \eta_\gamma^2 \right)^{1/2}$. In this marking strategy and in all further strategies of the present work the collection of elements/edges is not uniquely defined. Although the concrete choice does not influence theoretical considerations, in practical computations it makes sense to mark elements/edges with the largest contribution to the estimated error. For this strategy we have the following error reduction result shown in [23].

Theorem 4.3. *Let \mathcal{T}_H be a triangulation of Ω , and let \mathcal{T}_h be the triangulation obtained from \mathcal{T}_H by refining every element marked according to Strategy B in such a way that all three edges of any marked triangle are divided. There exist constants $\mu > 0$ and $0 < \alpha < 1$, depending only on the minimum angle and θ , such that for $\chi > 0$ satisfying*

$$\left(\sum_{K \in \mathcal{T}_H} h_K^2 \|f\|_{L_2(K)}^2 \right)^{1/2} \leq \mu \chi$$

one of two statements holds: either $\|\nabla(u - u_H)\|_{L_2(\Omega)} \leq \chi$, or the solution u_h on the mesh \mathcal{T}_h satisfies

$$\|\nabla(u - u_h)\|_{L_2(\Omega)} \leq \alpha \|\nabla(u - u_H)\|_{L_2(\Omega)} .$$

One could mention a lack of asymptotic convergence in the following sense. As soon as the prescribed tolerance χ is reached the further error reduction cannot be guaranteed by the theory. In order to achieve further error reduction one needs a new initial mesh having fineness μ with respect to $\chi_{new} < \chi$. This leads to the 2-stage algorithm which may be not flexible enough for some practical simulation. In the next section we describe an approach that does not have this difficulty and leads to an asymptotically convergent algorithm.

4.3 Convergent algorithm due to Morin, Nchetto and Siebert

In Section 4.2 we dealt with the marking strategy taking into account only the edge associated error indicators (see [40]). The resulting algorithm needs an additional mesh

preadaption and does not have asymptotic convergence. We can avoid any preadaption of the mesh and get an asymptotically convergent algorithm. Two ingredients are to be involved in marking and remeshing strategies. Namely,

1. data oscillation ζ should be controlled on each iteration step,
2. constructing a new mesh, one should introduce new nodes not only on all the edges of marked triangles, but also at least one node in the interior of each marked triangle.

To this end we are in need of a marking strategy that takes into account together with the edge error indicators also the element error indicators.

Marking Strategy E

Given a parameter $0 < \theta < 1$ and a mesh \mathcal{T}_H , mark the set of elements $\hat{\mathcal{T}}_H \subset \mathcal{T}_H$ for which

$$\left(\sum_{K \in \hat{\mathcal{T}}_H} \eta_{R,K}^2 \right)^{1/2} \geq \theta \eta_R.$$

Marking Strategy E differs from the original in [40] where the authors suggest to mark the edges first and then based thereof mark the triangles. We consider only triangle based marking since it is widely used. The differences are only technical.

Theorem 4.4. *Let \mathcal{T}_H be a triangulation of Ω , and let \mathcal{T}_h be the triangulation obtained from \mathcal{T}_H by refining every element marked according to Strategy E in such a way that all three edges of any marked triangle are divided and a new node inside any marked triangle is introduced. Then, there exist constants $\mu > 0$ and $0 < \alpha < 1$, depending only on the minimum angle and θ , such that, for $\chi > 0$ satisfying*

$$\zeta \leq \mu \chi$$

one of two statements holds: either $\|\nabla(u - u_H)\|_{L_2(\Omega)} \leq \chi$, or the solution u_h on the mesh \mathcal{T}_h satisfies

$$\|\nabla(u - u_h)\|_{L_2(\Omega)} \leq \alpha \|\nabla(u - u_H)\|_{L_2(\Omega)}.$$

Proof. See [40] for the proof. □

The following lemma ensures the resolution of data oscillation ζ according to a decreasing tolerance [40]. Since the proof is simple it is given.

Lemma 4.5. *Let $0 < \gamma < 1$ be the reduction factor of element size associated with one refinement step. Given $0 < \hat{\theta} < 1$, let $\hat{\alpha} := (1 - (1 - \gamma^2)\hat{\theta}^2)^{1/2}$. Let $\hat{\mathcal{T}}_H \subset \mathcal{T}_H$ be a collection of elements, such that*

$$\left(\sum_{K \in \hat{\mathcal{T}}_H} \zeta_K^2 \right)^{1/2} \geq \hat{\theta} \zeta. \tag{4.4}$$

If \mathcal{T}_h results from \mathcal{T}_H after refining at least all the elements in $\hat{\mathcal{T}}_H$, then the following data oscillation reduction holds:

$$\zeta_{new} \leq \hat{\alpha} \zeta,$$

where $\zeta_{new} = \left(\sum_{K \in \mathcal{T}_h} \zeta_K^2 \right)^{1/2}$ is the value of data oscillation on the refined mesh \mathcal{T}_h .

Proof. Let $K \in \hat{\mathcal{T}}_H$ be an element to refine. Since $f_K = |K|^{-1} \int_K f$ is an L_2 -projection of f on piecewise constant functions, then the following inequality holds:

$$\sum_{K' \in \mathcal{T}_h, K' \subset K} \|f - f_{K'}\|_{L_2(K')}^2 \leq \|f - f_K\|_{L_2(K)}^2$$

Using this inequality, the assumption of Theorem 4.4 and $h_{new} \leq \gamma h_{old}$ we have:

$$\begin{aligned} \zeta_{new}^2 &= \sum_{K \in \mathcal{T}_h} h_K^2 \|f - f_K\|_{L_2(K)}^2 \\ &\leq \gamma^2 \sum_{K \in \hat{\mathcal{T}}_H} h_K^2 \|f - f_K\|_{L_2(K)}^2 + \sum_{K \in \mathcal{T}_H \setminus \hat{\mathcal{T}}_H} h_K^2 \|f - f_K\|_{L_2(K)}^2 \\ &= (\gamma^2 - 1) \sum_{K \in \hat{\mathcal{T}}_H} \zeta_K^2 + \zeta^2 \leq (\gamma^2 - 1) \hat{\theta}^2 \zeta^2 + \zeta^2 \leq \hat{\alpha}^2 \zeta^2. \end{aligned}$$

□

Based on the Lemma 4.5 we can formulate marking strategy D leading to a reduction of data oscillation on each iteration step.

Marking Strategy D

Given a parameter $0 < \hat{\theta} < 1$ and the subset $\hat{\mathcal{T}}_H \subset \mathcal{T}_H$ produced by marking strategy E:
 Enlarge $\hat{\mathcal{T}}_H$ such that

$$\left(\sum_{K \in \hat{\mathcal{T}}_H} \zeta_K^2 \right) \geq \hat{\theta} \zeta.$$

Remark 4.6. We are now in a position to explain the idea of making an initial mesh from any given starting triangulation \mathcal{T}_h . To this end we can utilize Lemma 4.5 together with the marking strategy D, where we always replace ζ_K and ζ by $h_k \|f\|_{L_2(K)}$ and $\left(\sum_{K \in \mathcal{T}_H} h_k^2 \|f\|_{L_2(K)}^2 \right)^{1/2}$ respectively. Successively refining all elements marked by strategy D we can obviously come to the initial mesh with any prescribed fineness μ , which is guaranteed by Lemma 4.5.

Combining the a posteriori error control and data oscillation control we end up with the following algorithm.

Convergent Algorithm C

Choose parameters $0 < \theta, \hat{\theta} < 1$.

1. Take some initial mesh \mathcal{T}_0 .
2. Solve the discrete problem on \mathcal{T}_0 , denote by u_0 its solution by u_0 .
3. Let $k = 0$.
4. Compute the local indicators η_K .
5. If the global estimated error is small then STOP.
6. Construct $\hat{\mathcal{T}}_k \in \mathcal{T}_k$ by marking strategy E and parameter θ .
7. Enlarge $\hat{\mathcal{T}}_k$ by marking strategy D and parameter $\hat{\theta}$.
8. Let \mathcal{T}_{k+1} be a refinement of \mathcal{T}_k such that each element of $\hat{\mathcal{T}}_k$ as well as each of its sides contain a node of \mathcal{T}_{k+1} in its interior.
9. Solve the discrete problem on \mathcal{T}_{k+1} , denote its solution by u_{k+1} .
10. Let $k := k + 1$ and go to Step 4.

Theorem 4.7 (Convergence result). For $0 < \theta, \hat{\theta} < 1$, let $0 < \alpha < 1$, $\mu > 0$ be given by Theorem 4.4 and $0 < \hat{\alpha} < 1$ by Lemma 4.5. Let

$$\beta := \max(\alpha, \hat{\alpha}), \quad C_0 := \max\left(\|\nabla(u - u_0)\|_{L_2(\Omega)}, \frac{\zeta_0}{\alpha\mu}\right).$$

Algorithm C produces a convergent sequence $\{u_k\}_{k \in \mathbb{N}_0}$ of discrete solutions satisfying for all $k \geq 0$

$$\|\nabla(u - u_k)\|_{L_2(\Omega)} \leq C_0 \beta^k. \quad (4.5)$$

Proof. We perform this proof by means of mathematical induction. The desired inequality (4.5) holds evidently for $k = 0$. Assume (4.5) holds for k and show it for $k + 1$. Then we have either

$$(i) \|\nabla(u - u_k)\|_{L_2(\Omega)} > C_0 \beta^{k+1} \quad \text{or} \quad (ii) \|\nabla(u - u_k)\|_{L_2(\Omega)} \leq C_0 \beta^{k+1}.$$

In case (i), we see from Step 6 of Algorithm C that $\zeta_k \leq \hat{\alpha}^k \zeta_0 \leq \beta^k \zeta_0$, and, consequently, that for $\chi := C_0 \beta^{k+1}$

$$\begin{aligned} \zeta_k &\leq \beta^k \zeta_0 \leq \beta^k \alpha \mu \max\left(\|\nabla(u - u_0)\|_{L_2(\omega)}, \frac{\zeta_0}{\alpha\mu}\right) \\ &= \beta^k \alpha \mu C_0 \leq \beta^{k+1} \mu C_0 = \mu \chi. \end{aligned}$$

Since $\|\nabla(u - u_k)\|_{L_2(\Omega)} > \chi$, we may then combine Theorem 4.4 with (4.5) to arrive at

$$\|\nabla(u - u_{k+1})\|_{L_2(\Omega)} \leq \beta \|\nabla(u - u_k)\|_{L_2(\Omega)} \leq C_0 \beta^{k+1}.$$

On the other hand, exploiting that \mathcal{T}_{k+1} is a refinement of \mathcal{T}_k , and thus the error must not increase, we can handle case (ii) as follows:

$$\|\nabla(u - u_{k+1})\|_{L_2(\Omega)} \leq \|\nabla(u - u_k)\|_{L_2(\Omega)} \leq C_0 \beta^{k+1}.$$

Which completes the proof. \square

We would like to emphasize that the creation of a new inner node per marked triangle is required in the Algorithm C.

4.4 Avoidance of the new inner nodes

We can avoid the creation of a new node inside some marked triangles. Two possibilities are at hand, first we can carefully analyze data gained from the a posteriori error estimator. The second possibility is to use the fact that the error estimation for the Poisson equation is edge dominant, we describe this approach in Section 4.6. In this section we focus on the careful analysis of data stemming from the error indicators. Namely, in Algorithm C' we prescribe the creation of a new inner node only for the elements K with dominating element error indicator η_K . This approach still leads to a convergent algorithm with the same convergence properties as the original algorithm described in Section 4.3.

Marking Strategy E'

Given a parameter $0 < \theta < 1$ and a mesh \mathcal{T}_H , select the sets of elements $\hat{\mathcal{T}}_H \subset \mathcal{T}_H$ and $\tilde{\mathcal{T}}_H \subset \mathcal{T}_H$ ($\hat{\mathcal{T}}_H \cap \tilde{\mathcal{T}}_H = \emptyset$) for which

$$\left(\frac{1}{2} \sum_{K \in \hat{\mathcal{T}}_H} \sum_{\gamma \subset \partial K} \eta_\gamma^2 + \sum_{K \in \tilde{\mathcal{T}}_H} \eta_{R,K}^2 \right)^{1/2} \geq \theta \eta_R.$$

The marking strategy E' ensures that we mark sufficiently many elements $\hat{\mathcal{T}}_H \cup \tilde{\mathcal{T}}_H$ with a total contribution from the edge associated indicators of $\hat{\mathcal{T}}_H$ and the edge and element associated indicators of $\tilde{\mathcal{T}}_H$ proportional to η_R . We have the following error reduction theorem.

Theorem 4.8. *Let \mathcal{T}_H be a triangulation of Ω , and let \mathcal{T}_h be the triangulation obtained from \mathcal{T}_H by refining every element marked according to Strategy E' in such a way that all three edges of any triangle $K \in \hat{\mathcal{T}}_H \cup \tilde{\mathcal{T}}_H$ are divided, and additionally a new node is created inside each triangle $K \in \tilde{\mathcal{T}}_H$. Then, there exist constants $\mu > 0$ and $0 < \alpha < 1$, depending only on the minimum angle and θ , such that for $\chi > 0$ satisfying*

$$\zeta \leq \mu \chi \tag{4.6}$$

one of two statements holds: either $\|\nabla(u - u_H)\|_{L_2(\Omega)} \leq \chi$, or the solution u_h on the mesh \mathcal{T}_h satisfies

$$\|\nabla(u - u_h)\|_{L_2(\Omega)} \leq \alpha \|\nabla(u - u_H)\|_{L_2(\Omega)}.$$

Proof. This theorem will be proved in Section 4.5. \square

We keep the same idea for controlling the data oscillation leading to the marking strategy similar to marking strategy D.

Marking Strategy D'

Given a parameter $0 < \hat{\theta} < 1$ and the subsets $\hat{\mathcal{T}}_H, \tilde{\mathcal{T}}_H \subset \mathcal{T}_H$ produced by Marking Strategy E'.
Enlarge $\hat{\mathcal{T}}_H$ and $\tilde{\mathcal{T}}_H$ such that

$$\left(\sum_{K \in \hat{\mathcal{T}}_H} \zeta_K^2 + \sum_{K \in \tilde{\mathcal{T}}_H} \zeta_K^2 \right) \geq \hat{\theta} \zeta.$$

Although it is not important for theory which collection $\tilde{\mathcal{T}}_H, \hat{\mathcal{T}}_H$ is enlarged, it is recommended to enlarge in strategy D' the collection $\hat{\mathcal{T}}_H$ in order to avoid unnecessary nodes.

Convergent Algorithm C'

Choose parameters $0 < \theta, \hat{\theta} < 1$.

1. Take some initial mesh \mathcal{T}_0 .
2. Solve the discrete problem on \mathcal{T}_0 , denote its solution by u_0 .
3. Let $k = 0$.
4. Compute the local indicators η_K .
5. If the global estimated error is small then STOP.
6. Construct $\hat{\mathcal{T}}_k, \tilde{\mathcal{T}}_k \in \mathcal{T}_k$ by Marking Strategy E' and parameter θ .
7. Enlarge $\hat{\mathcal{T}}_k, \tilde{\mathcal{T}}_k$ by Marking Strategy D' and parameter $\hat{\theta}$.
8. Let \mathcal{T}_{k+1} be a refinement of \mathcal{T}_k such that
 - each element of $\hat{\mathcal{T}}_k$ contains a node of \mathcal{T}_{k+1} in the interior of every edge and
 - each element of $\tilde{\mathcal{T}}_k$, as well as each of its sides, contains a node of \mathcal{T}_{k+1} in its interior.
9. Solve the discrete problem on \mathcal{T}_{k+1} , denote its solution by u_{k+1} .
10. Let $k := k + 1$ and go to Step 4.

The following convergence result holds for Algorithm C'. We present it without a proof due to its similarity with Theorem 4.7.

Theorem 4.9 (Convergence result). For $0 < \theta, \hat{\theta} < 1$, let $0 < \alpha < 1$, $\mu > 0$ be given by Theorem 4.4 and $0 < \hat{\alpha} < 1$ by Lemma 4.5. Let

$$\beta := \max(\alpha, \hat{\alpha}), \quad C_0 := \max \left(\|\nabla(u - u_0)\|_{L_2(\Omega)}, \frac{\zeta_0}{\alpha\mu} \right).$$

Algorithm C' produces a convergent sequence $\{u_k\}_{k \in N_0}$ of discrete solutions satisfying for all $k \geq 0$

$$\|\nabla(u - u_k)\|_{L_2(\Omega)} \leq C_0 \beta^k. \quad (4.7)$$

4.5 Error reduction

This section is aimed to prove Theorem 4.8. We follow the lines of [23] and [40] in the proof.

Lemma 4.10. *In the notation of Theorem 4.8 the following error reduction holds:*

$$\|\nabla(u - u_h)\|_{L_2(\Omega)}^2 = \|\nabla(u - u_H)\|_{L_2(\Omega)}^2 - \|\nabla(u_h - u_H)\|_{L_2(\Omega)}^2$$

Proof. The proof follows directly from the Galerkin orthogonality. \square

This lemma gives the idea for the proof of Theorem 4.8. We have to show that $\|\nabla(u_h - u_H)\|_{L_2(\Omega)}$ is a certain part of the error $\|\nabla(u - u_H)\|_{L_2(\Omega)}$.

Lemma 4.11. *Let $K \in \mathcal{T}_H$ be any element and $\gamma \in \mathcal{E}_H$ be any edge of triangulation \mathcal{T}_H . In the notation of Theorem 4.8 the following estimates hold:*

$$\eta_K^2 \preceq \|\nabla(u_h - u_H)\|_{L_2(K)}^2 + \zeta_K^2, \quad (4.8)$$

$$\eta_\gamma^2 \leq \sum_{K \subset \omega_\gamma} \left(C_1 \|\nabla(u_h - u_H)\|_{L_2(K)}^2 + \frac{1}{6} \eta_K^2 + C_1 \zeta_K^2 \right) \quad (4.9)$$

provided there exist the basis functions ψ_K and ψ_γ on the new mesh \mathcal{T}_h . We denote by ψ_K and ψ_γ the hat functions corresponding to the new nodes introduced in the interior of the triangle K and on the edge γ respectively.

Proof. We show (4.8) first. Using the fact that $\int_K \nabla u_H \cdot \nabla \psi_K dx = - \int_K \Delta u_H \cdot \psi_K dx = 0$, we have

$$\begin{aligned} f_K \frac{|\text{supp}(\psi_K)|}{3} &= \int_K f_K \psi_K dx = \int_K f \psi_K dx - \int_K (f - f_K) \psi_K dx \\ &= \int_K \nabla u_h \cdot \nabla \psi_K dx - \int_K (f - f_K) \psi_K dx \\ &= \int_K \nabla(u_h - u_H) \cdot \nabla \psi_K dx - \int_K (f - f_K) \psi_K dx \end{aligned}$$

Squaring and using the Cauchy-Schwarz inequality we get:

$$|K| \|f_K\|_{L_2(K)}^2 \preceq \|\nabla(u_h - u_H)\|_{L_2(K)}^2 \|\nabla \psi_K\|_{L_2(K)}^2 + \|f - f_K\|_{L_2(K)}^2 \|\psi_K\|_{L_2(K)}^2$$

The desired inequality (4.8) follows now from the estimates ensured by the isotropy of the mesh:

$$\begin{aligned} \|\nabla \psi_K\|_{L_2(K)} &= O(1), \\ \|\psi_K\|_{L_2(K)} &\approx h_K. \end{aligned}$$

We show now the second inequality (4.9). Taking into account that R_γ is a constant over any edge γ , we have:

$$\begin{aligned} \frac{1}{2}|\gamma|R_\gamma &= \int_\gamma R_\gamma \psi_\gamma ds = - \int_{\omega_\gamma} \nabla u_H \cdot \nabla \psi_\gamma dx \\ &= \int_{\omega_\gamma} \nabla(u_h - u_H) \cdot \nabla \psi_\gamma dx - \int_{\omega_\gamma} f \psi_\gamma dx \\ &= \int_{\omega_\gamma} \nabla(u_h - u_H) \cdot \nabla \psi_\gamma dx - \sum_{K \subset \omega_\gamma} \int_K f_K \psi_\gamma - \sum_{K \subset \omega_\gamma} \int_K (f - f_K) \psi_\gamma dx \end{aligned}$$

Squaring and using the Cauchy-Schwarz inequality we come to the following:

$$\begin{aligned} \frac{1}{2}|\gamma|\|R_\gamma\|_{L_2(\gamma)}^2 &\leq \sum_{K \subset \omega_\gamma} \left(\|\nabla(u_h - u_H)\|_{L_2(K)}^2 \|\nabla \psi_\gamma\|_{L_2(K)}^2 + \|f_K\|_{L_2(K)}^2 \|\psi_\gamma\|_{L_2(K)}^2 \right. \\ &\quad \left. + \|f - f_K\|_{L_2(K)}^2 \|\psi_\gamma\|_{L_2(K)}^2 \right) \end{aligned}$$

We finish the proof observing that for $K \subset \omega_\gamma$ the following relations hold:

$$\begin{aligned} \|\nabla \psi_\gamma\|_{L_2(K)}^2 &= O(1), \\ \|\psi_\gamma\|_{L_2(K)}^2 &\leq \frac{|K|}{6}. \end{aligned}$$

□

Lemma 4.12. *Let \mathcal{T}_h be a triangulation resulting from \mathcal{T}_H by applying one cycle of algorithm C' . Then:*

1. *We have the following global lower bound for the error reduction*

$$\|\nabla(u_h - u_H)\|_{L_2(\Omega)}^2 \geq \frac{\theta^2 - \frac{1}{4}}{C_1} \|\nabla(u - u_H)\|_{L_2(\Omega)}^2 - C_2 \zeta^2.$$

If an additional condition on the marking holds:

$$\forall K \in \hat{\mathcal{T}}_h : \quad \eta_K^2 \leq \frac{1}{2} \sum_{\gamma \in \partial K} \eta_\gamma^2, \quad (4.10)$$

then

$$\|\nabla(u_h - u_H)\|_{L_2(\Omega)}^2 \geq \frac{\theta^2}{C_1} \|\nabla(u - u_H)\|_{L_2(\Omega)}^2 - C_2 \zeta^2.$$

Proof.

1. By Lemma 4.11 and the first step of the marking strategy E' we have

$$\begin{aligned} \theta^2 \eta_R^2 &\leq \frac{1}{2} \sum_{K \in \tilde{\mathcal{T}}_H} \sum_{\gamma \subset \partial K} \eta_\gamma^2 + \sum_{K \in \tilde{\mathcal{T}}_H} \eta_{R,K}^2 \\ &= \frac{1}{2} \sum_{K \in \tilde{\mathcal{T}}_H} \sum_{\gamma \subset \partial K} \eta_\gamma^2 + \frac{1}{2} \sum_{K \in \tilde{\mathcal{T}}_H} \sum_{\gamma \subset \partial K} \eta_\gamma^2 + \sum_{K \in \tilde{\mathcal{T}}_H} \eta_K^2 \\ &\leq C_1 \left(\sum_{K \in \tilde{\mathcal{T}}_H \cup \tilde{\mathcal{T}}_H} \sum_{\gamma \subset \partial K} \|u_h - u_H\|_{L_2(\omega_\gamma)}^2 + \sum_{K \in \tilde{\mathcal{T}}_H \cup \tilde{\mathcal{T}}_H} \sum_{\gamma \subset \partial K} \sum_{K' \subset \omega_\gamma} \zeta_{K'}^2 \right) \\ &\quad + \frac{1}{12} \sum_{K \in \tilde{\mathcal{T}}_H \cup \tilde{\mathcal{T}}_H} \sum_{\gamma \subset \partial K} \sum_{K' \subset \omega_\gamma} \eta_{K'}^2. \end{aligned}$$

Each triangle K' is counted at most three times in the last sum. Thus,

$$\|\nabla(u_h - u_H)\|_{L_2(\Omega)}^2 \geq \frac{\theta^2 - \frac{1}{4}}{C_1} \eta_R^2 - C_2 \zeta^2.$$

2. Proceeding analogously to the first part and utilizing the condition 4.10 we obtain

$$\begin{aligned} \theta^2 \eta_R^2 &\leq \frac{1}{2} \sum_{K \in \tilde{\mathcal{T}}_H} \sum_{\gamma \subset \partial K} \eta_\gamma^2 + \sum_{K \in \tilde{\mathcal{T}}_H} \eta_{R,K}^2 \\ &= \frac{1}{2} \sum_{K \in \tilde{\mathcal{T}}_H} \sum_{\gamma \subset \partial K} \eta_\gamma^2 + \frac{1}{11} \sum_{K \in \tilde{\mathcal{T}}_H} \eta_K^2 - \frac{1}{11} \sum_{K \in \tilde{\mathcal{T}}_H} \eta_K^2 + \frac{1}{2} \sum_{K \in \tilde{\mathcal{T}}_H} \sum_{\gamma \subset \partial K} \eta_\gamma^2 + \sum_{K \in \tilde{\mathcal{T}}_H} \eta_K^2 \\ &\leq \frac{6}{11} \sum_{K \in \tilde{\mathcal{T}}_H} \sum_{\gamma \subset \partial K} \eta_\gamma^2 - \frac{1}{11} \sum_{K \in \tilde{\mathcal{T}}_H} \eta_K^2 + \frac{1}{2} \sum_{K \in \tilde{\mathcal{T}}_H} \sum_{\gamma \subset \partial K} \eta_\gamma^2 + \sum_{K \in \tilde{\mathcal{T}}_H} \eta_K^2 \\ &\leq C_1 \left(\sum_{K \in \tilde{\mathcal{T}}_H \cup \tilde{\mathcal{T}}_H} \sum_{\gamma \subset \partial K} \|u_h - u_H\|_{L_2(\omega_\gamma)}^2 + \sum_{K \in \tilde{\mathcal{T}}_H \cup \tilde{\mathcal{T}}_H} \sum_{\gamma \subset \partial K} \sum_{K' \subset \omega_\gamma} \zeta_{K'}^2 \right). \end{aligned}$$

□

We are now in position to prove Theorem 4.8.

Proof of Theorem 4.8. Using Lemma 4.10 and Lemma 4.12, we have

$$\begin{aligned} \|\nabla(u - u_h)\|_{L_2(\Omega)}^2 &= \|\nabla(u - u_H)\|_{L_2(\Omega)}^2 - \|\nabla(u_h - u_H)\|_{L_2(\Omega)}^2 \\ &\leq \|\nabla(u - u_H)\|_{L_2(\Omega)}^2 \left(1 - \frac{\theta^2 - \frac{1}{4}}{C_1} \right) + C_2 \zeta^2 \end{aligned}$$

Assume now that $\|\nabla(u - u_H)\|_{L_2(\Omega)} > \chi$. Since $\zeta \leq \mu \chi$, then

$$\|\nabla(u - u_h)\|_{L_2(\Omega)}^2 \leq \|\nabla(u - u_H)\|_{L_2(\Omega)}^2 \left(1 - \frac{\theta^2 - \frac{1}{4}}{C_1} + C_2 \mu^2 \right).$$

Therefore,

$$\|\nabla(u - u_h)\|_{L_2(\Omega)} \leq \alpha \|\nabla(u - u_H)\|_{L_2(\Omega)},$$

where $\alpha := \left(1 - \frac{\theta^2 - \frac{1}{4}}{C_1} + C_2 \mu^2 \right)^{1/2} < 1$ provided $\mu > 0$ is sufficiently small. Moreover, if we assume the condition 4.10 to hold the error reduction constant reduces to

$$\alpha := \left(1 - \frac{\theta^2}{C_1} + C_2 \mu^2 \right)^{1/2}.$$

□

Remark 4.13. The smaller we choose μ the smaller is α but the larger C_0 . But most important is that μ is independent of \mathcal{T}_h .

4.6 Advantage of edge dominance in error estimation regarding adaptive refinement

It was shown in [17] that the edge residuals are dominating in a posteriori error estimation for the pure Poisson problem (see [36] for anisotropic case). However, to ensure this, the mesh properties formulated in [17] are required. To define these properties we need some additional notation.

Given an arbitrary triangle $K_0 \in \mathcal{T}_h$, define $R_0(K_0) := K_0$ and then, recursively, for $j \geq 1$, $R_j(K_0)$ as the union of all $K \in \mathcal{T}_h$ which are not contained in $\bigcup_{i < j} R_i(K_0)$ but which have at least one vertex in $R_{j-1}(K_0)$. Thus $R_j(K_0)$ is the union of all triangles K which may be connected by a path P_1, P_2, \dots, P_j with a vertex P_1 of K_0 , where P_j is a vertex of K , and $P_i P_{i+1}$ is an edge in \mathcal{E}_h for $1 \leq i < j$, and not by any shorter such a path. For $K \in R_j(K_0)$ we set $l(K_0, K) := j$. It follows, in particular, that $l(K_0, K)$ is symmetric in K_0 and K . Denote by $n_j(K_0)$ the number of triangles in $R_j(K_0)$. We assume that \mathcal{T}_h satisfies the following condition:

(A) There are constants $c_1, c_2, \alpha, \beta, r$ with $\alpha \geq 1, \beta \geq 1$ and

$$\alpha^{1/2} \beta < \sqrt{3} + \sqrt{2} \approx 3.146$$

such that, uniformly for small h ,

$$\begin{aligned} |K|/|K_0| &\leq c_1 \alpha^{l(K_0, K)} \quad \forall K, K_0 \in \mathcal{T}_h, \\ n_j(K) &\leq c_2 j^r \beta^j \quad \forall K \in \mathcal{T}_h. \end{aligned}$$

Remark 4.14. For the discussion of the condition (A), the examples and the counter examples of appropriate meshes we refer to [17] and [20].

The condition (A) leads to the following form of the error bounds.

Theorem 4.15 (Global upper error bound). *Assume that the condition (A) holds true. Then there exists a constant C depending only on the minimum angle of \mathcal{T}_h , such that*

$$\|\nabla(u - u_h)\|_{L_2(\Omega)}^2 \leq \sum_{\gamma \in \mathcal{E}} \eta_\gamma^2 + \zeta^2.$$

Theorem 4.16 (Local lower error bound). *Assume that the condition (A) holds true. There exists a constant C depending only on the minimum angle of \mathcal{T}_h , such that*

$$\eta_\gamma^2 \leq \|\nabla(u - u_h)\|_{L_2(\omega_\gamma)}^2 + \sum_{K \subset \omega_\gamma} \zeta_K^2.$$

We are now in position to formulate the algorithm C'' where we do not any more pay attention to the element error indicators.

Convergent Algorithm C''

Choose parameters $0 < \theta, \hat{\theta} < 1$.

1. Take some initial mesh \mathcal{T}_0 .
2. Solve the discrete problem on \mathcal{T}_0 , denote its solution by u_0 .
3. Let $k = 0$.
4. Compute the local indicators η_K .
5. If the global estimated error is small then STOP.
6. Construct $\hat{\mathcal{T}}_k \in \mathcal{T}_k$ by marking strategy B and parameter θ .
7. Enlarge $\hat{\mathcal{T}}_k$ by marking strategy D and parameter $\hat{\theta}$.
8. Let \mathcal{T}_{k+1} be a refinement of \mathcal{T}_k such that all the sides of each element in $\hat{\mathcal{T}}_k$, are divided.
9. Solve the discrete problem on \mathcal{T}_{k+1} , denote its solution by u_{k+1} .
10. Let $k := k + 1$ and go to Step 4.

For the algorithm C'' we can formulate analog of the convergence Theorem 4.7. The proof can be derived analogously.

Theorem 4.17 (Convergence result). *For $0 < \theta, \hat{\theta} < 1$, let $0 < \alpha < 1$, $\mu > 0$ be given by Theorem 4.3 and $0 < \hat{\alpha} < 1$ by Lemma 4.5. Let*

$$\beta := \max(\alpha, \hat{\alpha}), \quad C_0 := \max \left(\|\nabla(u - u_0)\|_{L_2(\Omega)}, \frac{\zeta_0}{\alpha\mu} \right).$$

Algorithm C' produces a convergent sequence $\{u_k\}_{k \in N_0}$ of discrete solutions satisfying for all $k \geq 0$

$$\|\nabla(u - u_k)\|_{L_2(\Omega)} \leq C_0 \beta^k. \quad (4.11)$$

Proof. Proof follows the lines of the proof of Theorem 4.8 by utilizing the bounds for the error and the error indicators formulated above. □

5 Convergence results for Poisson problem on anisotropic meshes

We restrict ourself in this chapter to the case of the pure Poisson problem $\varepsilon = 1$, $\kappa = 0$. We start with repeating the a posteriori residual error estimator for the anisotropic case. Then we investigate the possibilities to choose the weight in the edge error indicator on anisotropic meshes. Furthermore we provide the range of possible values of this weight, for which all the robustness estimates hold true. Based on the idea from the isotropic case, the adaptive algorithm allowing anisotropic mesh refinement is developed and analyzed in this chapter – one can analyze the edge and element error indicators separately, and according to this information only the marked entities (edges/elements) should be appropriately refined. In Section 5.3 the adaptive algorithm is given. It satisfies the convergence property for the Poisson problem with a parameter β depending on the alignment measure m_1 , as shown in Section 5.4.

5.1 Residual anisotropic error estimator

For any $f \in L_2(\Omega)$ there exist a unique weak solution u to the following problem:

$$u \in H_0^1(\Omega) : \quad (\nabla u, \nabla v)_{L_2(\Omega)} = (f, v)_{L_2(\Omega)}, \quad \forall v \in H_0^1(\Omega). \quad (5.1)$$

From a heuristic point of view one should stretch the triangle in that direction where the (directional) derivative of the function shows little change. The better the anisotropic mesh \mathcal{T} is aligned with the anisotropic function v , the more accurate one would expect the error estimates to be. In order to measure the alignment of \mathcal{T} with v , Kunert [31, 32] has introduced the *alignment measure* $m_1(v, \mathcal{T})$ which is defined as follows.

Definition 5.1 (Alignment measure m_1). *Let $v \in H^1(\Omega)$ be an arbitrary non-constant function, and \mathcal{F} be a family of triangulations of Ω . Define the alignment measure $m_1(\cdot, \cdot) : H^1(\Omega) \times \mathcal{F} \rightarrow \mathbb{R}$ by*

$$m_1(v, \mathcal{T}) := \frac{\left(\sum_{K \in \mathcal{T}} h_{min,K}^{-2} \cdot \|C_K^\top \nabla v\|_{L_2(K)}^2 \right)^{1/2}}{\|\nabla v\|}. \quad (5.2)$$

Furthermore the local matching function $m_1(\cdot, \cdot) : H^1(\Omega) \times \mathcal{T} \rightarrow \mathbb{R}$ is obviously defined by

$$m_1(v, K) := h_{min,K}^{-1} \frac{\|C_K^\top \nabla v\|_{L_2(K)}}{\|\nabla v\|_{L_2(K)}}$$

The alignment measure satisfies the following property:

$$1 \leq m_1(v, \mathcal{T}) \leq C \max_{K \in \mathcal{T}} \frac{h_{max,K}}{h_{min,K}}$$

The definition implies that a mesh \mathcal{T} which is well aligned with an anisotropic function v , results in a small alignment measure $m_1(v, \mathcal{T})$.

Let V_h be the space of continuous piecewise linear functions over a triangulation \mathcal{T}_h that vanish at the boundary. Let u_h denote the solution of the discrete problem

$$u_h \in V_h : \quad (\nabla u_h, \nabla v_h)_{L_2(\Omega)} = (f, v_h)_{L_2(\Omega)}, \quad \forall v_h \in V_h. \quad (5.3)$$

Denote by f_K the mean value of the function f on element $K \in \mathcal{T}_h$: $f_K = \frac{1}{|K|} \int_K f$. The jump discontinuity in the approximation of the normal flux at an interelement boundary $\gamma = \partial K \cap \partial K'$ is defined by

$$\left[\frac{\partial u_h}{\partial n} \right]_\gamma = n_K \cdot (\nabla u_h)_K + n_{K'} \cdot (\nabla u_h)_{K'},$$

and the edge residual R is given by

$$R_\gamma = \begin{cases} \left[\frac{\partial u_h}{\partial n} \right]_\gamma & \text{for } \gamma \not\subset \partial\Omega \\ 0 & \text{for } \gamma \subset \partial\Omega \end{cases}$$

which is defined as usual (see [4]).

We recall now residual the a posteriori error estimator for (5.1) and (5.3) from [32]. Define the error indicator η_γ associated with an edge γ by

$$\eta_\gamma := h_{\min, \gamma} \left(\frac{|\gamma|}{|\omega_\gamma|} \right)^{1/2} \|R_\gamma\|_{L_2(\gamma)},$$

and the error indicator η_K associated with an element K by

$$\eta_K := h_{\min, K} \|f_K\|_{L_2(K)}.$$

Then the local residual error estimator $\eta_{R,K}$ and the local data oscillation (local approximation term) ζ_K for the element K are defined via

$$\begin{aligned} \eta_{R,K}^2 &:= \eta_K^2 + \frac{1}{2} \sum_{\gamma \subset \partial K} \eta_\gamma^2, \\ \zeta_K^2 &:= h_{\min, K}^2 \|f - f_K\|_{L_2(K)}^2. \end{aligned}$$

Let η_R and ζ be their global counter parts, given by

$$\begin{aligned} \eta_R^2 &:= \sum_{K \in \mathcal{T}_h} \eta_{R,K}^2, \\ \zeta^2 &:= \sum_{K \in \mathcal{T}_h} \zeta_K^2. \end{aligned}$$

The following robustness results was shown in [32].

Theorem 5.2 (Global upper error bound). *There exists a constant C that does not depend on the aspect ratio of the mesh, such that*

$$\|\nabla(u - u_1)\|_{L_2(\Omega)}^2 \preceq m_1^2(u - u_h, \mathcal{T}_h) (\eta_R^2 + \zeta^2).$$

Theorem 5.3 (Local lower error bound). *There exists a constant C that does not depend on the aspect ratio of the mesh, such that*

$$\eta_{R,K}^2 \preceq \|\nabla(u - u_1)\|_{L_2(\omega_K)}^2 + \sum_{K \subset \omega_K} \zeta_K^2.$$

5.2 On the possible choice of weights for the edge error indicator

As one could see in the previous section the a posteriori residual error estimator can underestimate the exact error by a factor of $m_1(u - u_h, \mathcal{T}_h)$. In discussions on this concern the idea of choosing larger weights in edge indicators was proposed. We investigate this idea in this section and provide the answer how large the weights in edge indicators can be still remaining smaller than the exact error. Write the edge error indicator in the form

$$\eta_\gamma = \rho_\gamma \|R_\gamma\|_{L_2(\gamma)},$$

where ρ_γ is a weight. For the isotropic meshes $\rho_\gamma = h^{1/2}$ is typically used. In this section we are concerned with the possible values for ρ_γ , such that Theorems 5.2 and 5.3 hold true. One of the possible values for anisotropic meshes, $\rho_\gamma = h_{\min, \gamma} \left(\frac{|\gamma|}{|\omega_\gamma|} \right)^{1/2}$, was given in Section 5.1. In the current section we derive the upper bound $\rho_{\gamma, \max}$ for ρ_γ which ensures Theorems 5.2 and 5.3 to hold. It turns out that $\rho_{\gamma, \max} = h_{\min, \gamma}^{1/2}$, which is quantitatively larger than ρ_γ given in Section 5.1. We assume in this section the uniform maximum angle condition, i.e.

$$\sup_{\mathcal{T}_h \in \mathcal{F}} \max_{K \in \mathcal{T}_h} \max_{\alpha \in \text{angles}(K)} \alpha \leq C < \pi.$$

From the standard derivation of the a posteriori error estimator we know that the weight of the edge error indicator is directly defined by an edge bubble function $\phi \in H_0^1(\omega_\gamma)$ in the following way

$$\rho_\gamma^{-1}(\phi) = \frac{\|\nabla \phi\|_{L_2(K \cup K')}}{\|\phi\|_{L_2(\gamma)}}. \quad (5.4)$$

We would like to find a largest value of ρ_γ , or in another words we wish to solve the minimization problem

$$\rho_{\gamma, \max}^{-1} = \min_{\phi \in H_0^1(K \cup K')} \frac{\|\nabla \phi\|_{L_2(K \cup K')}}{\|\phi\|_{L_2(\gamma)}}. \quad (5.5)$$

Since we can not solve this minimization problem exactly, we do the following. First we make some simplifications, then we estimate the largest possible value of ρ_γ and at the end we give a function that brings the asymptotic optimum.

Theorem 5.4. *In notation of Figure 8 the inequality*

$$\min_{\phi \in H_0^1(K \cup K')} \frac{\|\nabla \phi\|_{L_2(K \cup K')}}{\|\phi\|_{L_2(\gamma)}} \geq \min_{\phi \in H_0^1(\tilde{K} \cup \tilde{K}')} \frac{\|\nabla \phi\|_{L_2(\tilde{K} \cup \tilde{K}')}}{\|\phi\|_{L_2(\gamma)}} \quad (5.6)$$

holds.

Proof. Without loss of generality we assume that $\alpha = \max(\alpha, \beta)$. For the proof we use the transform $F : \tilde{K} \rightarrow K$ (the transformation $F : \tilde{K}' \rightarrow K'$ can be considered analogously) defined via

$$F(x) = \begin{pmatrix} x \\ y + \frac{b_2}{a_2}x \end{pmatrix}.$$

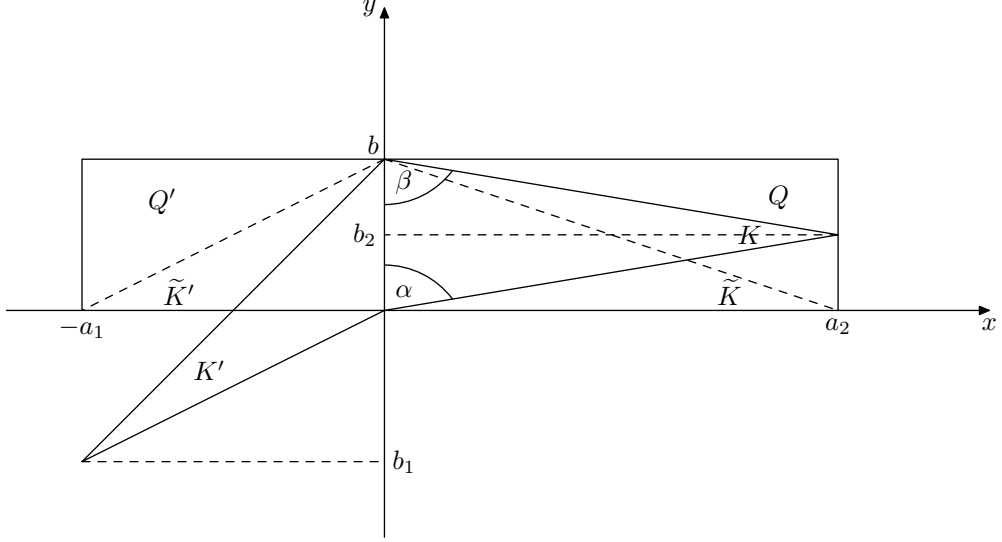


Figure 8: Edge and two neighboring triangles. K, K' – actual triangles; \tilde{K}, \tilde{K}' – triangles after applying the transformation F^{-1} ; Q, Q' – the smallest rectangles containing triangles \tilde{K} and \tilde{K}' respectively.

This transform has the following jacobian matrix

$$J = \begin{pmatrix} 1 & 0 \\ \frac{b_2}{a_2} & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \cot(\alpha) & 1 \end{pmatrix},$$

with $|J| = 1$. We also will use the fact that $\|J\|$ is bounded. Indeed,

$$\|J\|^2 = \lambda_{\max}(J^T J) = 1 + \frac{\cot^2(\alpha)}{2} + \sqrt{\cot^2(\alpha) + \frac{\cot^4(\alpha)}{4}} \leq C < \infty \quad (5.7)$$

due to the uniform maximal angle condition. Now we proceed estimating the enumerator of the target functional

$$\begin{aligned} \|\nabla\phi\|_{L_2(K)}^2 &= \int_K \|\nabla\phi\|^2 dx dy = \int_{\tilde{K}} \|J^{-T}\nabla\phi\|^2 |J| dx dy \\ &\geq \|J\|^{-2} \int_{\tilde{K}} \|\nabla\phi\|^2 dx dy \geq \int_{\tilde{K}} \|\nabla\phi\|^2 dx dy. \end{aligned}$$

Performing the same computations for $F : \tilde{K} \rightarrow K$ we complete the proof. \square

Theorem 5.5. *In notations of Figure 8*

$$\min_{\phi \in H_0^1(\tilde{K} \cup \tilde{K}')} \frac{\|\nabla\phi\|_{L_2(\tilde{K} \cup \tilde{K}')}}{\|\phi\|_{L_2(\gamma)}} \geq \min_{\phi \in H_0^1(Q \cup Q')} \frac{\|\nabla\phi\|_{L_2(Q \cup Q')}}{\|\phi\|_{L_2(\gamma)}}. \quad (5.8)$$

Proof. The proof is evident since $H_0^1(\tilde{K} \cup \tilde{K}') \subset H_0^1(Q \cup Q')$. \square

The last minimum in Theorem 5.5 is nothing else, but the Rayleigh quotient of the eigenvalue problem

$$\text{Find } (\lambda, \phi) \in \mathbb{R} \times H_0^1(Q \cup Q') : (\nabla\phi, \nabla\psi)_{L_2(Q \cup Q')} = \lambda^2 (\phi, \psi)_{L_2(\gamma)} \quad \forall \psi \in H_0^1(Q \cup Q'),$$

or in the classical formulation

$$-\Delta\phi = \lambda^2\delta(x)\phi,$$

where $\delta(x)$ is the Dirac delta function of the variable x . Search for the eigenfunction in the form $\phi = \alpha(y)\beta(x)$

$$-\frac{\alpha''(y)}{\alpha(y)} - \frac{\beta''(x)}{\beta(x)} = \lambda^2\delta(x)$$

The problem is now split into two ordinary differential equations. The first of them involves constant parameter ν to be defined later.

$$-\frac{\alpha''(y)}{\alpha(y)} = \nu^2, \quad \alpha(0) = 0, \quad \alpha(b) = 0.$$

The solution thereto is obtained directly

$$\alpha(y) = \sin \nu y, \quad \nu = \pi k/b, \quad k \in \mathbb{N} \quad (5.9)$$

The second equation reads:

$$\nu^2 - \frac{\beta''(x)}{\beta(x)} = \lambda^2\delta(x),$$

or

$$-\beta''(x) + \nu^2\beta(x) = \lambda^2\delta(x)\beta(x)$$

We search for the solutions in the form

$$\beta(x) = \begin{cases} \beta_1(x), & x < 0, \\ \beta_2(x), & x > 0, \end{cases}$$

which satisfies the system of equations

$$\begin{cases} -\beta_i''(x) + \nu^2\beta_i(x) = 0, & i = 1, 2, \\ \beta_1(-0) = \beta_2(+0), \\ \beta_2'(0) - \beta_1'(0) = \lambda^2\beta_1(0), \\ \beta_1(-a_1) = 0, \\ \beta_2(a_2) = 0. \end{cases}$$

Solving this system we get:

$$\beta(x) = \begin{cases} \beta_1(x) = \frac{\sinh(\nu(a_1 + x))}{\sinh(\nu a_1)}, \\ \beta_2(x) = \frac{\sinh(\nu(a_2 - x))}{\sinh(\nu a_2)}, \end{cases}$$

and

$$\lambda^2 = \frac{\nu \sinh(\nu(a_1 + a_2))}{\sinh(\nu a_1) \sinh(\nu a_2)} = \nu \frac{\cosh(\nu a_1)}{\sinh(\nu a_1)} + \nu \frac{\cosh(\nu a_2)}{\sinh(\nu a_2)}. \quad (5.10)$$

Proposition 5.6. *If a_1, a_2 are two positive parameters then the function*

$$f(\xi) = \frac{\xi \sinh(\xi(a_1 + a_2))}{\sinh(\xi a_1) \sinh(\xi a_2)} = \xi \frac{\cosh(\xi a_1)}{\sinh(\xi a_1)} + \xi \frac{\cosh(\xi a_2)}{\sinh(\xi a_2)}$$

monotonically increases for $\xi > 0$.

Proof. We can rewrite the function in the following way

$$f(\xi) = \xi \frac{\cosh(\xi a_1)}{\sinh(\xi a_1)} + \xi \frac{\cosh(\xi a_2)}{\sinh(\xi a_2)} = g(\xi, a_1) + g(\xi, a_2),$$

where

$$g(\xi, a) = \xi \frac{\cosh(\xi a)}{\sinh(\xi a)}.$$

It is enough to show that the function $g(\xi, a)$ is a monotonically increasing function with respect to ξ for any positive parameter a . Therefore we differentiate $g(\xi, a)$:

$$\frac{\partial g(\xi, a)}{\partial \xi} = \frac{\frac{1}{2} \sinh(2\xi a) - \xi a}{\sinh^2(\xi a)}.$$

It is easy to see that $\frac{\partial g(\xi, a)}{\partial \xi} > 0$ for all $\xi a > 0$ since $x < \sinh x$ for all $x > 0$. It means that $g(\xi, a)$ and consequently $f(\xi)$ are monotonically increasing functions which was claimed. \square

We are interested in the smallest possible value for λ^2 . Due to Proposition 5.6 the minimum is reached for $k = 1$ in (5.9) and the corresponding value

$$\nu_1 = \frac{\pi}{b}.$$

The function that minimizes the last functional in Theorem 5.5 reads as follows:

$$\phi(x, y) = \begin{cases} \frac{\sinh\left(\frac{\pi}{b}(a_1 + x)\right)}{\sinh\left(\frac{\pi}{b}a_1\right)} \sin\left(\frac{\pi}{b}y\right), & (x, y) \in \tilde{Q}', \\ \frac{\sinh\left(\frac{\pi}{b}(a_2 - x)\right)}{\sinh\left(\frac{\pi}{b}a_2\right)} \sin\left(\frac{\pi}{b}y\right), & (x, y) \in \tilde{Q}. \end{cases}$$

This function together with the traditional bubble function are drawn in Figure 9.

We arrive at the following result:

Theorem 5.7. *For any function $\phi \in H_0^1(K \cup K')$ the following estimate hold:*

$$\frac{\|\nabla \phi\|_{L_2(K \cup K')}^2}{\|\phi\|_{L_2(\gamma)}^2} \geq \frac{\pi}{b} \coth\left(\frac{\pi}{b}a_1\right) + \frac{\pi}{b} \coth\left(\frac{\pi}{b}a_2\right).$$

Proof. To show this we only need to substitute the optimal value $\nu_1 = \frac{\pi}{b}$ into (5.10). \square

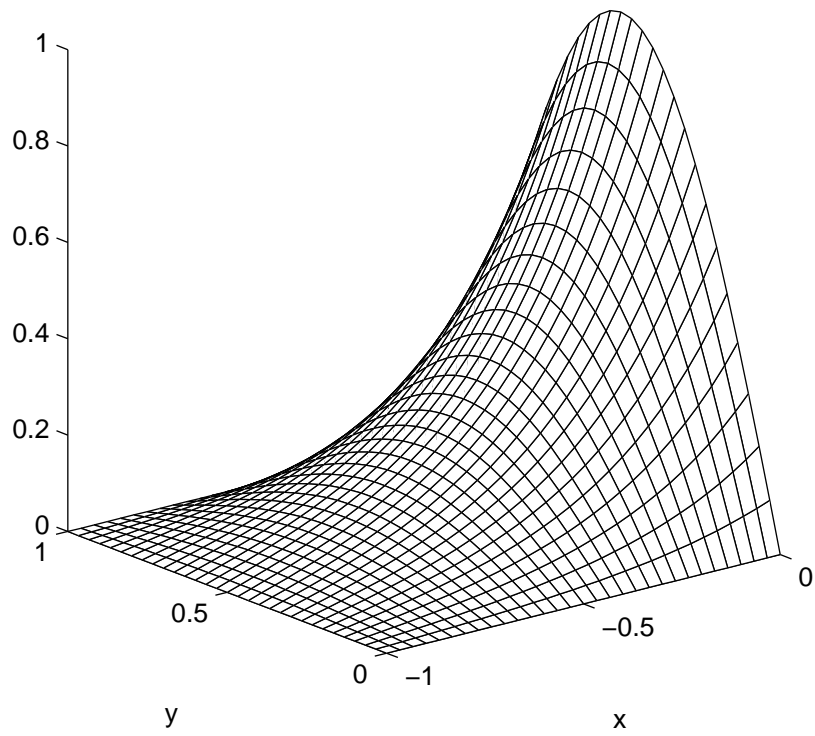
It remains only to give an example of a function, for which the equality or in our case equivalence relation takes place. As an example of such a function we may consider the function $b_{\gamma, \delta}$ with

$$\delta = \frac{|\gamma|}{|\omega_\gamma|} h_{min, \gamma}.$$

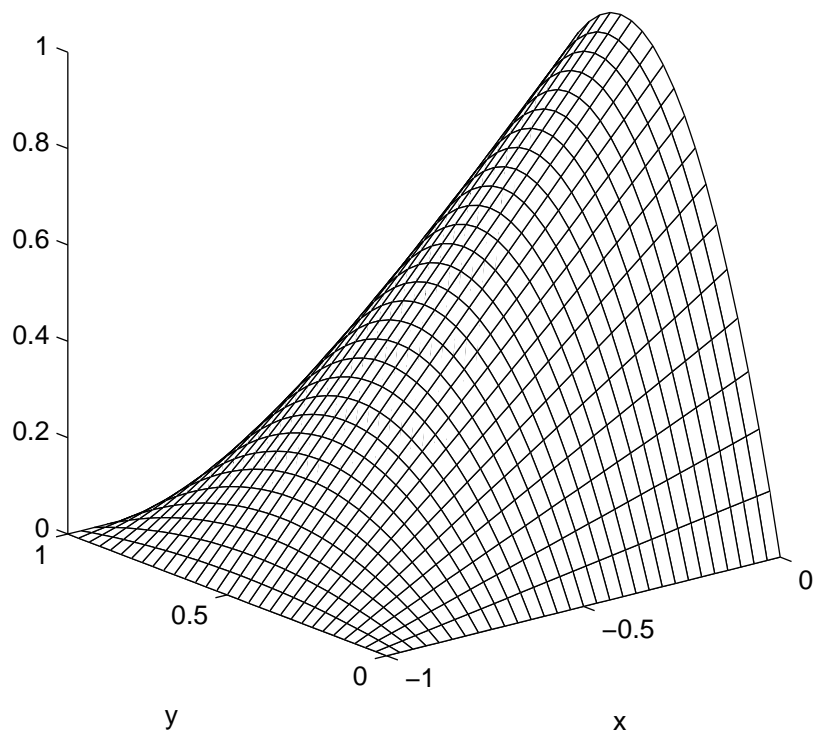
Theorem 5.8. *The squeezed edge bubble function $b_{\gamma, \delta}$ with the squeezing parameter*

$$\delta = \frac{|\gamma|}{|\omega_\gamma|} h_{min, \gamma} \text{ satisfies}$$

$$\frac{\|\nabla b_{\gamma, \delta}\|_{L_2(K)}^2}{\|b_{\gamma, \delta}\|_{L_2(\gamma)}^2} \sim \frac{\pi}{b} \coth\left(\frac{\pi}{b}a_2\right)$$



Bubble function obtained in this section



Traditional edge bubble function

Figure 9: Two bubble functions on quadrilateral.

Proof. For the bubble function $b_{\gamma,\delta}$ we have

$$\frac{\|\nabla b_{\gamma,\delta}\|_{L_2(K)}^2}{\|b_{\gamma,\delta}\|_{L_2(\gamma)}^2} \sim h_{min,\gamma}^{-1}.$$

We show now that the right hand side of the desired equivalence relation is also of the same order as $h_{min,\gamma}^{-1}$. To this end we consider two cases:

1. $a_2 \sim h_{min,\gamma}$. Then

$$\frac{\pi}{b} \coth\left(\frac{\pi}{b}a_2\right) \sim \frac{\pi}{b} \left(\frac{\pi}{b}a_2\right)^{-1} = a_2^{-1} \sim h_{min,\gamma}^{-1}.$$

2. $b \sim h_{min,\gamma}$. We have

$$\frac{\pi}{b} \coth\left(\frac{\pi}{b}a_2\right) \sim \frac{\pi}{b} \sim b^{-1} \sim h_{min,\gamma}^{-1},$$

where we used the knowledge on the behavior of the function \coth :

$$\coth(x) \sim \begin{cases} x^{-1}, & \text{for small } x, \\ 1, & \text{for large } x. \end{cases}$$

□

This theorem shows that the optimal value of the weight $\rho_{\gamma,\max} = h_{min,\gamma}^{1/2}$ is asymptotically reached by the squeezed edge bubble function. There is a natural mapping between the possible bubble function and weights in the edge error indicators. Although the optimal weight $h_{min,\gamma}^{1/2}$ is good for the hierarchical error estimators it cannot be used in the convergence proofs of the adaptive algorithms. The functions appearing in the refined mesh cannot be any kind of squeezed functions, and thus we use the weight $h_{min,\gamma} \left(\frac{|\gamma|}{|K|}\right)^{1/2}$ in the adaptive algorithm.

5.3 Marking strategies and convergent algorithm on anisotropic meshes

In this section we present marking and refinement strategies which produce suitable anisotropic meshes.

Marking Strategy $\tilde{\mathbf{E}}$

Given a parameter $0 < \theta < 1$ and a mesh \mathcal{T}_H , select the set of elements $\hat{\mathcal{T}}_H \subset \mathcal{T}_H$ and the set of edges $\hat{\mathcal{E}}_H \subset \mathcal{E}_H$ for which

$$\left(\sum_{K \in \hat{\mathcal{T}}_H} \eta_K^2 + \sum_{\gamma \in \hat{\mathcal{E}}_H} \eta_\gamma^2 \right)^{1/2} \geq \theta \eta_R.$$

Theorem 5.9. *Let \mathcal{T}_H be a triangulation of Ω , and let \mathcal{T}_h be the triangulation obtained from \mathcal{T}_H by refining every element marked according to Strategy $\tilde{\mathbf{E}}$ in such a way that new node is created in the interior of each marked triangle $K \in \hat{\mathcal{T}}_H$ and on the each marked edge $\gamma \in \hat{\mathcal{E}}_H$.*

Then, there exist constants $\mu > 0$ and $0 < \alpha < 1$, depending on θ and the alignment measure, such that for $\chi > 0$ satisfying

$$\zeta \leq \mu\chi \tag{5.11}$$

one of two statements holds: either $\|\nabla(u - u_H)\|_{L_2(\Omega)} \leq \chi$, or the solution u_h on the mesh \mathcal{T}_h satisfies

$$\|\nabla(u - u_h)\|_{L_2(\Omega)} \leq \alpha\|\nabla(u - u_H)\|_{L_2(\Omega)}.$$

The dependence on the alignment measure has the following form:

$$\alpha = \left(1 - \frac{1}{Cm_1(u - u_H, \mathcal{T}_H)}\right)^{1/2}$$

Proof. Proof is postponed to Section 5.4. □

Remark 5.10. It is clear that if the alignment measure $m_1(u - u_H, \mathcal{T}_H)$ is of moderate size then the error reduction factor α will be strictly smaller than 1. Thus it makes sense to control the alignment measure $m_1(u - u_H, \mathcal{T}_H)$ in order to have guaranteed error reduction.

Marking Strategy $\tilde{\mathbf{D}}$

Given a parameter $0 < \hat{\theta} < 1$ and the subsets $\hat{\mathcal{T}}_H \subset \mathcal{T}_H$ and $\hat{\mathcal{E}}_H \subset \mathcal{E}_H$ produced by Marking Strategy $\tilde{\mathbf{E}}$:

Enlarge $\hat{\mathcal{T}}_H$ so that

$$\sum_{K \in \hat{\mathcal{T}}_H} \zeta_K^2 \geq \hat{\theta}\zeta.$$

Combining the marking strategy we formulate the adaptive algorithm.

Convergent Algorithm \tilde{C}

Choose parameters $0 < \theta, \hat{\theta} < 1$.

1. Take some initial mesh \mathcal{T}_0 .
2. Solve the discrete problem on \mathcal{T}_0 , denote by u_0 its solution.
3. Let $k = 0$.
4. Compute the local indicators η_K .
5. If the global estimated error is small then STOP.
6. Construct $\hat{\mathcal{T}}_k \in \mathcal{T}_k$ and $\hat{\mathcal{E}}_k \in \partial\mathcal{T}_k$ by Marking Strategy \tilde{E} and parameter θ .
7. Enlarge $\hat{\mathcal{T}}_k$ by Marking Strategy \tilde{D} and parameter $\hat{\theta}$.
8. Let \mathcal{T}_{k+1} be a refinement of \mathcal{T}_k such that
 - each element of $\hat{\mathcal{T}}_k$ contains a node of \mathcal{T}_{k+1} in its interior and
 - each edge of $\hat{\mathcal{E}}_k$, contains a node of \mathcal{T}_{k+1} in its interior.
9. Solve the discrete problem on \mathcal{T}_{k+1} , denote its solution by u_{k+1} .
10. Let $k := k + 1$ and go to Step 4.

In the anisotropic case the analog of convergence Theorem 4.7 reads similarly.

Theorem 5.11 (Convergence result). *For $0 < \theta, \hat{\theta} < 1$, let $0 < \alpha < 1$, $\mu > 0$ be given by Theorem 4.4 and $0 < \hat{\alpha} < 1$ by Lemma 4.5. Let*

$$\beta := \max(\alpha, \hat{\alpha}), \quad C_0 := \max\left(\|\nabla(u - u_0)\|_{L_2(\Omega)}, \frac{\zeta_0}{\alpha\mu}\right).$$

Algorithm \tilde{C} produces a convergent sequence $\{u_k\}_{k \in N_0}$ of discrete solutions satisfying for all $k \geq 0$

$$\|\nabla(u - u_k)\|_{L_2(\Omega)} \leq C_0 \beta^k, \tag{5.12}$$

provided that there exists a constant M such that the alignment measure satisfies

$$m_1(u - u_k, \mathcal{T}_k) < M, \text{ for all } k \geq 0.$$

5.4 Error reduction theorem

This section is aimed to prove Theorem 5.9. Mention that Lemma 4.10 holds true for the anisotropic case as well as for the isotropic one.

Lemma 5.12. *Let $K \in \mathcal{T}_H$ be any element and $\gamma \in \mathcal{E}_H$ be any edge of triangulation. In the notation of Theorem 4.8 the following estimates hold:*

$$\eta_K^2 \leq \|\nabla(u_h - u_H)\|_{L_2(K)}^2 + \zeta_K^2, \quad (5.13)$$

$$\eta_\gamma^2 \leq \sum_{K \subset \omega_\gamma} \left(C_1 \|\nabla(u_h - u_H)\|_{L_2(K)}^2 + \frac{1}{6} \eta_K^2 + C_1 \zeta_K^2 \right), \quad (5.14)$$

provided there exist the basis functions ψ_K and ψ_γ on the new mesh \mathcal{T}_h . We denote by ψ_K and ψ_γ the hat functions corresponding to the new nodes introduced in the interior of the triangle K and on the edge γ respectively.

Proof. We show (5.13) first. Using the fact that $\int_K \nabla u_H \cdot \nabla \psi_K dx = \int_K \nabla u_h \cdot \nabla \psi_K dx = 0$, we have

$$\begin{aligned} f_K \frac{|K|}{3} &= \int_K f_K \psi_K dx = \int_K f \psi_K dx - \int_K (f - f_K) \psi_K dx \\ &= \int_K \nabla u_h \cdot \nabla \psi_K dx - \int_K (f - f_K) \psi_K dx \\ &= \int_K \nabla(u_h - u_H) \cdot \nabla \psi_K dx - \int_K (f - f_K) \psi_K dx \end{aligned}$$

Squaring and using Cauchy-Schwarz inequality we get:

$$|K| \|f_K\|_{L_2(K)}^2 \leq \|\nabla(u_h - u_H)\|_{L_2(K)}^2 \|\nabla \psi_K\|_{L_2(K)}^2 + \|f - f_K\|_{L_2(K)}^2 \|\psi_K\|_{L_2(K)}^2$$

The desired inequality (5.13) follows now from the estimates (2.12) and (2.13).

Show now the second inequality (5.14). Taking into account that R_γ is a constant over any edge γ , we have:

$$\begin{aligned} \frac{1}{2} |\gamma| R_\gamma &= \int_\gamma R_\gamma \psi_\gamma ds = - \int_{\omega_\gamma} \nabla u_H \cdot \nabla \psi_\gamma dx \\ &= \int_{\omega_\gamma} \nabla(u_h - u_H) \cdot \nabla \psi_\gamma dx - \int_{\omega_\gamma} f \psi_\gamma dx \\ &= \int_{\omega_\gamma} \nabla(u_h - u_H) \cdot \nabla \psi_\gamma dx - \sum_{K \subset \omega_\gamma} \int_K f_K \psi_\gamma - \sum_{K \subset \omega_\gamma} \int_K (f - f_K) \psi_\gamma dx \end{aligned}$$

We finish the proof observing that for $K \subset \omega_\gamma$ the relations (2.14), (2.15) hold. \square

Lemma 5.13. *Let \mathcal{T}_h be a triangulation resulting from \mathcal{T}_H by applying one cycle of algorithm \tilde{C} . Then:*

1. *we have the following global lower bound for the error reduction*

$$\|\nabla(u_h - u_H)\|_{L_2(\Omega)}^2 \geq \frac{\theta^2 - \frac{1}{4}}{C_1 m_1^2(u - u_H, \mathcal{T}_H)} \|\nabla(u - u_H)\|_{L_2(\Omega)}^2 - C_2 \zeta^2.$$

If an additional condition on the marking holds:

$$\forall \gamma \in \hat{\mathcal{E}}_h, \forall K : \gamma \in \partial K : \quad \eta_K^2 \leq \eta_\gamma^2, \text{ or } K \in \hat{\mathcal{T}}_h \quad (5.15)$$

then

$$\|\nabla(u_h - u_H)\|_{L_2(\Omega)}^2 \geq \frac{\theta^2}{C_1 m_1^2(u - u_H, \mathcal{T}_H)} \|\nabla(u - u_H)\|_{L_2(\Omega)}^2 - C_2 \zeta^2.$$

Proof. **1.** By Lemma 5.12 and the first step of the marking strategy $\tilde{\mathbf{E}}$ we have

$$\begin{aligned} \theta^2 \eta_R^2 &\leq \sum_{\gamma \in \hat{\mathcal{E}}_H} \eta_\gamma^2 + \sum_{K \in \hat{\mathcal{T}}_H} \eta_K^2 \\ &\leq C_1 \left(\sum_{\gamma \in \hat{\mathcal{E}}_H} \|u_h - u_H\|_{L_2(\omega_\gamma)}^2 + \sum_{K \in \hat{\mathcal{T}}_H} \|u_h - u_H\|_{L_2(K)}^2 + \zeta^2 \right) + \frac{1}{12} \sum_{\gamma \in \hat{\mathcal{E}}_H} \sum_{K \subset \omega_\gamma} \eta_K^2. \end{aligned}$$

Each triangle K is counted at most three times in the last sum. Thus,

$$\|\nabla(u_h - u_H)\|_{L_2(\Omega)}^2 \geq \frac{\theta^2 - \frac{1}{4}}{C_1 m_1^2(u - u_H, \mathcal{T}_H)} \eta_R^2 - C_2 \zeta^2.$$

2. Proceeding analogously to the first part and utilizing the condition 5.15 we obtain

$$\begin{aligned} \theta^2 \eta_R^2 &\leq \sum_{\gamma \in \hat{\mathcal{E}}_H} \eta_\gamma^2 + \sum_{K \in \hat{\mathcal{T}}_H} \eta_K^2 \\ &= \sum_{\gamma \in \hat{\mathcal{E}}_H} \eta_\gamma^2 + \frac{1}{5} \sum_{\gamma \in \hat{\mathcal{E}}_H} \sum_{K \subset \omega_\gamma, K \notin \hat{\mathcal{T}}_H} \eta_K^2 - \frac{1}{5} \sum_{\gamma \in \hat{\mathcal{E}}_H} \sum_{K \subset \omega_\gamma, K \notin \hat{\mathcal{T}}_H} \eta_K^2 + \sum_{K \in \hat{\mathcal{T}}_H} \eta_K^2 \\ &\leq \frac{6}{5} \sum_{K \in \hat{\mathcal{T}}_H} \sum_{\gamma \subset \partial K} \eta_\gamma^2 - \frac{1}{5} \sum_{\gamma \in \hat{\mathcal{E}}_H} \sum_{K \subset \omega_\gamma, K \notin \hat{\mathcal{T}}_H} \eta_K^2 + \sum_{K \in \hat{\mathcal{T}}_H} \eta_K^2 \\ &\leq C_1 \left(\sum_{\gamma \in \hat{\mathcal{E}}_H} \|u_h - u_H\|_{L_2(\omega_\gamma)}^2 + \sum_{K \in \hat{\mathcal{T}}_H} \|u_h - u_H\|_{L_2(K)}^2 + \zeta^2 \right). \end{aligned}$$

□

We are now in a position to prove Theorem 5.9.

Proof of Theorem 5.9. Using Lemma 4.10 and Lemma 5.13, we have

$$\begin{aligned} \|\nabla(u - u_h)\|_{L_2(\Omega)}^2 &= \|\nabla(u - u_H)\|_{L_2(\Omega)}^2 - \|\nabla(u_h - u_H)\|_{L_2(\Omega)}^2 \\ &\leq \|\nabla(u - u_H)\|_{L_2(\Omega)}^2 \left(1 - \frac{\theta^2 - \frac{1}{4}}{C_1 m_1^2(u - u_H, \mathcal{T}_H)} \right) + C_2 \zeta^2 \end{aligned}$$

Assume now that $\|\nabla(u - u_H)\|_{L_2(\Omega)} > \chi$. Since $\zeta \leq \mu \chi$, then

$$\|\nabla(u - u_h)\|_{L_2(\Omega)}^2 \leq \|\nabla(u - u_H)\|_{L_2(\Omega)}^2 \left(1 - \frac{\theta^2 - \frac{1}{4}}{C_1 m_1^2(u - u_H, \mathcal{T}_H)} + C_2 \mu^2 \right).$$

Therefore,

$$\|\nabla(u - u_h)\|_{L_2(\Omega)} \leq \alpha \|\nabla(u - u_H)\|_{L_2(\Omega)},$$

where $\alpha := \left(1 - \frac{\theta^2 - \frac{1}{4}}{C_1 m_1^2(u - u_H, \mathcal{T}_H)} + C_2 \mu^2 \right)^{1/2} < 1$ provided $\mu > 0$ and $m_1^2(u - u_H, \mathcal{T}_H)$ are sufficiently small. Moreover, if we assume the condition (5.15) to hold the error reduction constant reduces to

$$\alpha := \left(1 - \frac{\theta^2}{C_1 m_1^2(u - u_H, \mathcal{T}_H)} + C_2 \mu^2 \right)^{1/2}.$$

□

6 Error estimation for the reaction-diffusion problem

Some types of a posteriori error estimation methods have already been generalized for anisotropic meshes. They include the works on residual error estimators [32, 34, 36], hierarchical error estimator [21] and Dirichlet local problem error estimators [33, 35]. In this chapter we give

6.1 Equilibrated residual method with computable error approximation

The modified equilibrated residual method [2] has been shown to provide one of the most reliable error estimates for the singularly perturbed reaction-diffusion problem. We will refer to this method as Ainsworth-Babuška estimator. The equilibrated residual method dates back to [16] and was further developed in [37, 14, 2, 3]. The purpose of the current section is to consider the estimator from [2] on anisotropic meshes and to give upper and lower error bounds on anisotropic meshes obtained in [27]. The equilibrated residual method is shown to fail on anisotropic meshes due to a (potentially unbounded) factor appearing in the lower bound. This factor is $O(1)$ on isotropic meshes, but it can be of size of the maximum aspect ratio on anisotropic meshes. We recall the modification for anisotropic elements from [27] leading to a robust error estimator. The upper error bound of the modification contains an alignment measure which is in accordance with the results by Kunert [35].

Among others, the robustness of the equilibrated residual method relies on the solution of a local residual problem on each element. This problem is infinite dimensional and does not allow an exact solution in general. In practical computations it is therefore solved approximately by means of the finite element method. According to the existing theory, however, the equilibrated residual method is guaranteed to be robust if the exact solution to the local problem can be computed. The potentially fatal role of the approximation of this solution has been poorly investigated up to now, which gives rise to heavy criticism of the whole method. In this section we provide an appropriate basis for the solution of the local problem and show that the whole method is not affected by this approximation.

Regarding to the history of the equilibrated residual method, the reaction diffusion equation was always put under consideration depending only on one parameter κ , in other words $\varepsilon = 1$. In this section we also restrict ourselves to the case $\varepsilon = 1$ although all the considerations may be generalized to arbitrary ε .

6.1.1 The equilibrated residual method

In this section a brief overview over the equilibrated residual method is given since we strongly require parts of this method for our subsequent analysis. The equilibrated residual method may be found in [4] and its modification for the singularly perturbed case in [2].

Consider the model problem (2.1). Then the error $e := u - u_X$ belongs to the space $H_0^1(\Omega)$ and satisfies the variational formulation

$$B(e, v) = B(u, v) - B(u_X, v) = (f, v) - B(u_X, v) \quad \forall v \in H_0^1(\Omega). \quad (6.1)$$

For an element K with boundary ∂K , let n_K be the outer normal vector. Next we introduce a set of boundary fluxes $\{g_K : K \in \mathcal{T}\}$ which approximate the actual fluxes of the exact solution on the element boundaries, $g_K \approx n_K \nabla u|_K$. Taking into account that

the trace of the true solution is continuous on the edges, we construct the approximate fluxes g_K so that the condition

$$g_K + g_{K'} = 0 \text{ on } \partial K \cap \partial K', \quad K, K' \in \mathcal{T} \quad (6.2)$$

holds true. With this definition the residual on the right hand side of (6.1) can be decomposed into contributions from the individual elements

$$(f, v) - B(u_X, v) = \sum_{K \in \mathcal{T}} \left\{ (f, v)_K - B_K(u_X, v) + \int_{\partial K} g_K v ds \right\} \quad \forall v \in H_0^1(\Omega). \quad (6.3)$$

The term in parentheses defines a linear functional on the space of the locally admissible functions

$$V_K = \{v : v = w|_\Omega \text{ for some } w \in H_0^1(\Omega)\}.$$

If the parameter κ in (2.1) is not zero, there is a unique solution $\phi_K \in V_K$ to the local residual problem

$$B_K(\phi_K, v) = (f, v)_K - B_K(u_X, v) + \int_{\partial K} g_K v ds \quad \forall v \in V_K. \quad (6.4)$$

If κ vanishes then the problem will have a solution if and only if the collection of fluxes $\{g_K : K \in \mathcal{T}\}$ satisfies the so-called *equilibration condition*

$$0 = (f, 1)_K - B_K(u_X, 1) + \int_{\partial K} g_K ds. \quad (6.5)$$

This condition means that the boundary flux g_K is in equilibrium with the interior load. Note that the local problem (6.4) is infinite dimensional. The solution ϕ_K is treated as an approximation of the true error on the element K . It yields the a posteriori error estimation $\|e\|^2 \sim \sum_{K \in \mathcal{T}} \|\phi_K\|_K^2$, which will be shown later.

The substitution of (6.4) into (6.3) implies

$$B(e, v) = (f, v) - B(u_X, v) = \sum_{K \in \mathcal{T}} B_K(\phi_K, v), \text{ for all } v \in V.$$

An immediate consequence of this result is the upper bound on the true error. We obtain from the Cauchy-Schwarz inequality that

$$|B(e, v)| \leq \sum_{K \in \mathcal{T}} \|\phi_K\|_K \|v\|_K \leq \left\{ \sum_{K \in \mathcal{T}} \|\phi_K\|_K^2 \right\}^{1/2} \|v\|,$$

and conclude that

$$\|e\| = \sup_{v \in H_0^1(\Omega) : \|v\|=1} B(e, v) \leq \left\{ \sum_{K \in \mathcal{T}} \|\phi_K\|_K^2 \right\}^{1/2}.$$

These developments lead to the following theorem:

Theorem 6.1 (Upper error bound). *Let $\{g_K : K \in \mathcal{T}\}$ be any set of boundary fluxes satisfying condition (6.2). Additionally, if κ vanishes, then (6.5) is assumed to hold on all elements that do not abut the boundary $\partial\Omega$. Then, the global error in the finite element approximation is bounded by*

$$\|e\|^2 \leq \sum_{K \in \mathcal{T}} \|\phi_K\|_K^2.$$

Proof. For the proof see the argumentation above. □

6.1.2 Construction of the equilibrated fluxes

For the construction of the equilibrated fluxes we adopt the theory of [4] which we will briefly repeat in this section. It will be assumed that the finite element subspace X is constructed using linear elements on a partition \mathcal{T} of the domain Ω into triangular elements. The key issue of the lower bound of the error is the construction of appropriate approximate fluxes. The procedure that will be developed produces a set of fluxes $\{g_K\}$ that satisfy the *first-order equilibration conditions*:

$$\left. \begin{aligned} (f, \theta_n)_K - B_K(u_X, \theta_n) + \int_{\partial K} g_K \theta_n ds &= 0 \quad \forall n \in \mathcal{N}(K) \\ g_K + g_{K'} &= 0 \text{ on } \partial K \cap \partial K'. \end{aligned} \right\} \quad (6.6)$$

It is convenient to look for $\overline{g_K}|_\gamma$ belonging to $\text{span}\{\theta_n : n \in \mathcal{N}(\gamma)\}$ on all edges. Ainsworth and Oden [4] suggest to choose the degrees of freedom for the fluxes to be the moments $\mu_{K,n}^\gamma = \int_\gamma g_K \theta_n ds$ with respect to the FEM basis functions θ_n associated with $x_n \in \mathcal{N}(\gamma)$. Thereby we avoid a global problem by reducing the construction of fluxes to computation of the moments over local patches of elements.

Let $\mathcal{N}(\gamma) = \{x_l, x_r\}$, then it can be shown that the actual flux may be reconstructed from its moments:

$$g_K|_\gamma = \frac{2}{|\gamma|} \left\{ (2\mu_{K,l}^\gamma - \mu_{K,r}^\gamma)\theta_l + (-\mu_{K,l}^\gamma + 2\mu_{K,r}^\gamma)\theta_r \right\}. \quad (6.7)$$

Note that (6.7) could be rewritten in the form

$$g_K|_\gamma = \mu_{K,l}^\gamma \psi_l + \mu_{K,r}^\gamma \psi_r,$$

where ψ_l and ψ_r are the dual basis functions corresponding to θ_l and θ_r , i.e. $(\psi_i, \theta_j)_{L_2(\gamma)} = \delta_{ij}$ for $i, j \in \{l, r\}$:

$$\psi_l = \frac{2}{|\gamma|}(2\theta_l - \theta_r), \quad \psi_r = \frac{2}{|\gamma|}(-\theta_l + 2\theta_r).$$

In order to determine the boundary fluxes, it is sufficient to determine the moments of the flux with respect to the basis functions. The first-order equilibration conditions (6.6) for the flux g_K may be rewritten in terms of the flux moments in the form

$$\left. \begin{aligned} \sum_{\gamma \subset \partial K} \mu_{K,n}^\gamma &= \Delta_K(\theta_n) \quad \forall n \in \mathcal{N}(K) \\ \mu_{K,n}^\gamma + \mu_{K',n}^\gamma &= 0 \quad \forall n \in \mathcal{N}(\gamma), \quad \gamma = \partial K \cap \partial K' \end{aligned} \right\} \quad (6.8)$$

where

$$\Delta_K(\theta_n) = B_K(u_X, \theta_n) - (f, \theta_n)_K. \quad (6.9)$$

In (6.8) we used the convention that $\mu_{K,n}^\gamma = 0$ if $n \notin \mathcal{N}(\gamma)$.

The conditions in (6.8) have one of two distinct structures depending on the location of the node x_n in the interior or on the boundary of Ω . Here we omit the case of a boundary vertex, see [4] for details. Assume x_n to be an interior vertex. The elements and edges are labeled as shown in Figure 10. The moment equilibration conditions (6.8) for the elements $K \in \tilde{x}_n$ associated with the node x_n may be rewritten in the form

$$\left. \begin{aligned} \mu_{1,n}^{\gamma_1} + \mu_{1,n}^{\gamma_2} &= \Delta_1(\theta_n) \\ \mu_{2,n}^{\gamma_2} + \mu_{2,n}^{\gamma_3} &= \Delta_2(\theta_n) \\ &\vdots \\ \mu_{N,n}^{\gamma_N} + \mu_{N,n}^{\gamma_1} &= \Delta_N(\theta_n) \end{aligned} \right\}$$

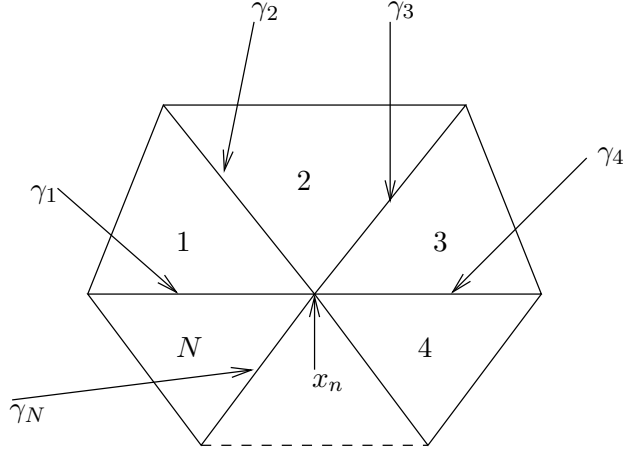


Figure 10: The patch of elements influenced by the basis function θ_n .

with constraints on the edges

$$\left. \begin{aligned} \mu_{1,n}^{\gamma_1} + \mu_{N,n}^{\gamma_1} &= 0 \\ \mu_{2,n}^{\gamma_2} + \mu_{1,n}^{\gamma_2} &= 0 \\ &\vdots \\ \mu_{N,n}^{\gamma_N} + \mu_{N-1,n}^{\gamma_N} &= 0. \end{aligned} \right\}$$

It is shown in [4] that this system of $2N$ equations for $2N$ unknowns has a one-parametric family of solutions. We recall from [4] which solution should be selected.

The ideal situation would be to choose the flux moments $\{g_K\}$ satisfying $\mu_{K,n}^{\gamma} \approx \int_{\gamma} \theta_n n_K \nabla u \, ds$. Since the true fluxes are unknown, the flux moments are selected so that

$$\mu_{K,n}^{\gamma} \approx \tilde{\mu}_{K,n}^{\gamma} := \int_{\gamma} \theta_n n_K \nabla u_X|_K \, ds. \quad (6.10)$$

We seek flux moments that minimize the objective

$$\frac{1}{2} \sum_{K \in \tilde{x}_n} \sum_{\gamma \subset \partial K} \left(\mu_{K,n}^{\gamma} - \tilde{\mu}_{K,n}^{\gamma} \right)^2. \quad (6.11)$$

Introducing Lagrange multipliers we come to the optimality condition. The Lagrangian is given by

$$\begin{aligned} \mathcal{L} \left(\{ \tilde{\mu}_{K,n}^{\gamma} \}, \{ \lambda_{\gamma} \}, \{ \sigma_K \} \right) &= \frac{1}{2} \sum_{K \in \tilde{x}_n} \sum_{\gamma \subset \partial K} \left(\mu_{K,n}^{\gamma} - \tilde{\mu}_{K,n}^{\gamma} \right)^2 \\ &+ \sum_{K \in \tilde{x}_n} \sigma_{K,n} \left(\Delta_K(\theta_n) - \sum_{\gamma \subset \partial K} \mu_{K,n}^{\gamma} \right) + \sum_{\gamma = \partial K \cap \partial K'} \lambda_{\gamma,n} \left(\mu_{K,n}^{\gamma} + \mu_{K',n}^{\gamma} \right). \end{aligned}$$

Here we used the convention that $\lambda_{\gamma,n} = 0$ on $\gamma \subset \partial\Omega$. We conclude that the conditions for a stationary point consist of two parts; the first part is (6.8), the second part is

$$\mu_{K,n}^{\gamma} - \tilde{\mu}_{K,n}^{\gamma} - \sigma_{K,n} + \lambda_{\gamma,n} = 0. \quad (6.12)$$

Using the second part of (6.8) we obtain

$$\lambda_{\gamma,n} = \begin{cases} \frac{1}{2} \left(\sigma_{K,n} + \sigma_{K',n} + \tilde{\mu}_{K,n}^{\gamma} + \tilde{\mu}_{K',n}^{\gamma} \right) & \gamma = \partial K \cap \partial K', \\ 0 & \gamma = \partial K \cap \partial\Omega. \end{cases}$$

Using the last formula together with (6.12) the flux moments are expressed as

$$\mu_{K,n}^\gamma = \begin{cases} \frac{1}{2} (\sigma_{K,n} - \sigma_{K',n} + \tilde{\mu}_{K,n}^\gamma - \tilde{\mu}_{K',n}^\gamma) & \gamma = \partial K \cap \partial K', \\ \sigma_{K,n} + \tilde{\mu}_{K,n}^\gamma & \gamma = \partial K \cap \partial \Omega. \end{cases} \quad (6.13)$$

Substituting this into the first equation of (6.8) we obtain the following conditions for $\{\sigma_{K,n} : K \in \tilde{x}_n\}$:

$$\frac{1}{2} \sum_{\gamma=\partial K \cap \partial K'} (\sigma_{K,n} - \sigma_{K',n}) + \sum_{\gamma \subset \partial K \cap \partial \Omega} \sigma_{K,n} = \tilde{\Delta}_K(\theta_n) \quad \forall K \in \tilde{x}_n, \quad (6.14)$$

where

$$\tilde{\Delta}_K(\theta_n) := B_K(u_X, \theta_n) - (f, \theta_n)_K - \int_{\partial K} \left\langle \frac{\partial u_X}{\partial n_K} \right\rangle \theta_n ds, \quad (6.15)$$

$$\left\langle \frac{\partial u_X}{\partial n_K} \right\rangle := \begin{cases} \frac{1}{2} n_K \{ (\nabla u_X)_K + (\nabla u_X)_{K'} \} & \text{on } \partial K \cap \partial K' \\ n_K (\nabla u_X)_K & \text{on } \partial K \cap \partial \Omega. \end{cases} \quad (6.16)$$

The conditions (6.14) form a linear system of equations over the element patches \tilde{x}_n with unknowns $\{\sigma_{K,n} : K \in \tilde{x}_n\}$ corresponding to the elements in the patch. The specific form for an interior vertex is

$$\frac{1}{2} \begin{bmatrix} 2 & -1 & \dots & -1 \\ -1 & 2 & -1 & \dots & 0 \\ \vdots & & & & \vdots \\ 0 & \dots & -1 & 2 & -1 \\ -1 & \dots & & -1 & 2 \end{bmatrix} \begin{bmatrix} \sigma_1 \\ \sigma_2 \\ \vdots \\ \sigma_{N-1} \\ \sigma_N \end{bmatrix} = \begin{bmatrix} \tilde{\Delta}_1(\theta_n) \\ \tilde{\Delta}_2(\theta_n) \\ \vdots \\ \tilde{\Delta}_{N-1}(\theta_n) \\ \tilde{\Delta}_N(\theta_n) \end{bmatrix}.$$

The kernel of this matrix is the vector $\mathbf{1} = [1, 1, \dots, 1]^\top$, which implies that a solution exists if and only if the sum of the components of the right-hand data vanishes. This may be easily verified thanks to the Galerkin property (see [4]).

Since the system (6.14) is singular the least square solution is selected. As a consequence, there exists a constant C , depending only on the number of elements in the patch \tilde{x}_n surrounding a vertex x_n , such that (for proof see for ex. [28])

$$\sum_{K \in \tilde{x}_n} \sigma_{K,n}^2 \preceq \sum_{K \in \tilde{x}_n} \tilde{\Delta}_K(\theta_n)^2. \quad (6.17)$$

6.1.3 Minimum energy extensions

Minimum energy extensions were first introduced in the work of Ainsworth and Babuška [2]. These extensions play a key role in the construction of an estimator which is stable with respect to the perturbation parameter κ . The original equilibrated residual method is described in the work of Ainsworth and Oden [3]. However, as it is shown in [2], it is not stable with respect to κ . Ainsworth and Babuška [2] propose the following modification of the previous method for the singularly perturbed case. The functions θ_n in (6.14) are replaced by an approximate minimum energy extension θ_n^* to $\theta_n|_{\partial K}$. The system (6.14) then is solved in a least-square sense, since it has no solution in general. As in (6.17), one gets the solution that depends continuously on the data:

$$\sum_{K \in \tilde{x}_n} \sigma_{K,n}^2 \preceq \sum_{K \in \tilde{x}_n} \tilde{\Delta}_K(\theta_n^*)^2. \quad (6.18)$$

The error estimator we propose is derived from the estimator of the work [2] but differs in two details. Firstly, we pay more attention to the minimization of the appropriate function energy norm and even obtain the minimum. We will develop this in this section. The second modification is described in §6.1.6.

Let K be any element and let $v \in H^{1/2}(\partial K)$. The *minimum energy extension* $\mathcal{E}v$ of v to the interior of the element is characterized by the conditions

$$\mathcal{E}v \in H^1(K) : \mathcal{E}v = v \text{ on } \partial K, \quad B_K(\mathcal{E}v, \omega) = 0 \quad \forall \omega \in H_0^1(K).$$

The definition of the minimum energy extension has an advantageous property. Let $v \in H^{1/2}(\partial K)$. The *minimum energy extension* $\mathcal{E}v$ of v to the interior of the element has the *minimal energy norm among all functions coinciding with v on the boundary ∂K* . Indeed, consider the energy norm of the function $\mathcal{E}v + \omega$:

$$\|\|\|\mathcal{E}v + \omega\|\|^2 = \|\|\|\mathcal{E}v\|\|^2 + \|\|\|\omega\|\|^2 + 2B_K(\mathcal{E}v, \omega) = \|\|\|\mathcal{E}v\|\|^2 + \|\|\|\omega\|\|^2 \geq \|\|\|\mathcal{E}v\|\|^2. \quad (6.19)$$

The proof easily follows from (6.19) observing that $\mathcal{E}v + \omega$ coincides with $\mathcal{E}v$ on the boundary ∂K .

For the one-dimensional case it is possible to find a minimum energy extension explicitly (see [2]).

Consider now the two-dimensional case. We look for an approximation for the minimum energy extension of the first-order basis function. Let the element $K = \triangle ABC$ be a triangle. Consider the basis function θ corresponding to the vertex A . We seek an approximation to the minimum energy extension $\mathcal{E}\theta$ in the following class Λ of functions. Set

$$\Lambda : = \{v \in C^0(K) : v = \theta \text{ on } \partial K, v = 0 \text{ in } \triangle CDB, \\ v \text{ is linear in each triangle } \triangle CAD \text{ and } \triangle BAD, D \in \triangle ABC\}.$$

We obtain now an approximation for the minimum energy extension of this basis function. To this end we put an arbitrary point D in the triangle (see Figure 11).

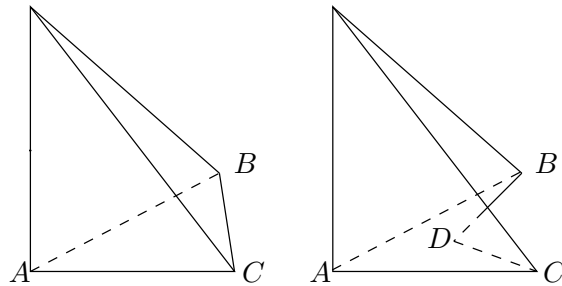


Figure 11: Original hat function and the family of functions used to approximate the minimum energy extension $\mathcal{E}\theta$.

Next we choose that function from the set Λ that minimizes energy norm. Our developments here differs from the original paper [2]. There a point $(1/\kappa, 1/\kappa)$ is introduced in the *reference* triangle and D is the image of this point after the corresponding affine transformation. The corresponding function does not necessarily minimize the energy over Λ but it is shown to be sufficiently accurate. For us, however *this is not sufficient* and we instead consider the point D to be in the *actual* triangle in order to obtain the optimal position of this point.

Introduce a local coordinate system such that the vertex A coincides with the origin and the edge AC lies on the axis Ox . Let $D = (a, b)$, $C = (h_1, 0)$ and $B = (h_3, h_2)$ (see Figure 12).

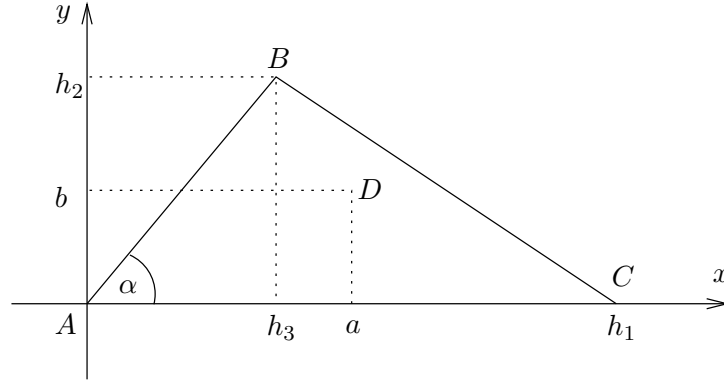


Figure 12: Notations for the parameters of an arbitrary triangle.

Let $\Theta(a, b, x, y) \in \Lambda$ be an admissible function, see Figure 11. The squared energy norm of this function is

$$\begin{aligned} \Phi(a, b) &= B_K(\Theta(a, b, x, y), \Theta(a, b, x, y)) \\ &= \frac{\kappa^2}{12}(b(h_1 - h_3) + ah_2) + \frac{2h_1h_2h_3 - ah_1h_2 - bh_1h_3 + bh_3^2}{2h_1h_3^2} \\ &\quad + \frac{(h_2^2 + h_3^2)(a - h_3)^2}{2h_3^2(ah_2 - bh_3)} + \frac{(a - h_1)^2}{2h_1b} \end{aligned}$$

For given parameters h_1, h_2, h_3 we want to minimize $\Phi(a, b)$ with respect to a and b . A number of calculations leads to a stationary point of this function

$$\begin{cases} \bar{a}^* = \frac{\sqrt{6h_1}(h_3 + \sqrt{h_2^2 + h_3^2})}{\sqrt{12h_3 + 12\sqrt{h_2^2 + h_3^2} + \kappa^2h_2^2h_1}} \\ \bar{b}^* = \left(-\frac{h_3}{h_2} + \sqrt{\left(\frac{h_3}{h_2}\right)^2 + 1}\right) \bar{a}^*. \end{cases}$$

To prove that this solution is a minimum point it is sufficient to show convexity of the function. Therefore consider the Hessian matrix $D^2\Phi$ of the second-order derivatives. By direct computations one obtains

$$\begin{aligned} \frac{\partial^2 \Phi}{\partial a^2} &= \frac{1}{bh_1} + \frac{(h_2^2 + h_3^2)(b - h_2)^2}{(h_2a - h_3b)^3} \geq 0, \\ \det D^2\Phi &= \frac{(h_2^2 + h_3^2)(h_1b - h_3b + h_2a - h_1h_2)^2}{(h_2a - h_3b)^3 b^3 h_1} \geq 0. \end{aligned}$$

Hence $\Phi(a, b)$ is convex and (\bar{a}^*, \bar{b}^*) is the unique minimum.

We have found the function $\Theta^*(\bar{a}^*, \bar{b}^*, x, y)$ which minimizes the energy norm over the set Λ . However, in practice it is sufficient to take not the exact values of \bar{a}^* and \bar{b}^* , but some values a^*, b^* that are equivalent for $\kappa \rightarrow \infty$, namely

$$\begin{cases} a^* = \frac{\sqrt{6}}{\kappa} \frac{(h_3 + \sqrt{h_2^2 + h_3^2})}{h_2} \\ b^* = \frac{\sqrt{6}}{\kappa}. \end{cases}$$

Note that the corresponding point $D = (a^*, b^*)$ lies on the bisector of the angle $\angle BAC =: \alpha$ and $|AD| = \frac{\sqrt{6}}{\kappa \sin \alpha/2}$. The analysis given neglects the fact that (a, b) should be contained in K . Therefore, we construct the function θ^* as follows

$$\theta^* := \begin{cases} \Theta(a^*, b^*, x, y), & \text{if } (a^*, b^*) \in K, \\ \theta, & \text{otherwise.} \end{cases}$$

Lemma 6.2. *Under the above notations and assumptions the following holds*

$$\|\theta^*\|_{L_2(K)}^2 \preceq |K| \min(1, h_{min,K}^{-1} \kappa^{-1}) \sim \text{meas}(\partial K) \min(h_{min,K}, \kappa^{-1}),$$

where $h_{min,K}$ is the height corresponding to the largest edge of the triangle K .

Proof. Consider K for which $(a^*, b^*) \in K$. A short calculation yields $\kappa^{-1} \preceq h_{min,K}$ and $\min(h_{min,K}, \kappa^{-1}) \sim \kappa^{-1}$. Furthermore one obtains

$$\begin{aligned} \|\theta^*\|_{L_2(K)}^2 &= \frac{\sqrt{6} \left(h_1 + \sqrt{h_2^2 + h_3^2} \right)}{12\kappa} \\ &\preceq \text{meas}(\partial K) \kappa^{-1} \sim \text{meas}(\partial K) \min(h_{min,K}, \kappa^{-1}). \end{aligned}$$

It remains to consider the case $\kappa^{-1} \gg h_{min,K}$. In this case we have $\min(1, h_{min,K}^{-1} \kappa^{-1}) = 1$ and θ^* coincides with θ . The estimate

$$\|\theta^*\|_{L_2(K)}^2 = \|\theta\|_{L_2(K)}^2 \sim |K| = |K| \min(1, h_{min,K}^{-1} \kappa^{-1})$$

completes the proof. □

6.1.4 Estimates for element and face residuals in the anisotropic case

Further we will need the anisotropic trace inequality and some more facts concerning approximation properties on an anisotropic triangle.

Lemma 6.3 (Anisotropic trace inequality). *Let K be an arbitrary triangle and γ be an edge of it. For $v \in H^1(K)$ the following trace inequality holds:*

$$\|v\|_{L_2(\gamma)}^2 \preceq \frac{|\gamma|}{|K|} \|v\|_{L_2(K)} \left(\|v\|_{L_2(K)} + \|C_K^\top \nabla v\|_{L_2(K)} \right).$$

Lemma 6.4 (Anisotropic approximation properties). *Let K be any triangle, $\gamma \subset \partial K$ be any edge thereof and $v \in H^1(K)$. Denote by $\bar{v} = \frac{1}{|K|} \int_K v$ the mean value of v over an element K . Then*

$$\|v - \bar{v}\|_{L_2(K)} \leq \|v\|_{L_2(K)}, \tag{6.20}$$

$$\|v - \bar{v}\|_{L_2(K)} \preceq \|C_K^\top \nabla v\|_{L_2(K)}, \tag{6.21}$$

$$\|v - \bar{v}\|_{L_2(K)} \preceq \alpha_K m \|v\|_K \preceq \alpha_K m_1(v, K) \|v\|_K, \tag{6.22}$$

$$\|v - \bar{v}\|_{L_2(\gamma)} \preceq \left(\frac{|\gamma|}{|K|} \right)^{1/2} \|C_K^\top \nabla v\|_{L_2(K)}, \tag{6.23}$$

$$\begin{aligned} \|v - \bar{v}\|_{L_2(\gamma)} &\preceq \left(\frac{|\gamma|}{|K|} \right)^{1/2} h_{min,K}^{1/2} \alpha_K^{1/2} m \|v\|_K \\ &\preceq \left(\frac{|\gamma|}{|K|} \right)^{1/2} h_{min,K}^{1/2} \alpha_K^{1/2} m_1(v, K) \|v\|_K. \end{aligned} \tag{6.24}$$

Proof. Estimate (6.20) is obvious. For estimate (6.21) see for instance [32]. Estimate (6.22) evidently follows from (6.20) and (6.21). Combining Lemma 6.3 and estimate (6.20) of the current lemma and observing that $\|v\|_{L_2(K)} \|C_K^\top \nabla v\|_{L_2(K)} \leq \|v\|_{L_2(K)}^2 + \|C_K^\top \nabla v\|_{L_2(K)}^2$ we verify (6.23). In order to show (6.24) we use (6.23) to obtain

$$\|v - \bar{v}\|_{L_2(\gamma)}^2 \preceq \frac{|\gamma|}{|K|} \|C_K^\top \nabla v\|_{L_2(K)}^2 \leq \frac{|\gamma|}{|K|} h_{min,K}^2 \|v\|_K^2.$$

Furthermore, with the aid of Lemma 6.3 and estimates (6.20), (6.21) we get the following:

$$\|v - \bar{v}\|_{L_2(\gamma)}^2 \preceq \frac{|\gamma|}{|K|} \kappa^{-1} \sqrt{\kappa^2 \|v\|_{L_2(K)}^2} h_{min,K} \sqrt{h_{min,K}^{-2} \|C_K^\top \nabla v\|_{L_2(K)}^2} \leq \frac{|\gamma|}{|K|} h_{min,K} \kappa^{-1} \|v\|_K^2.$$

Combining the two previous estimates we get the result claimed. \square

We are interested in a particular value of parameter δ depending on the edge γ for which the squeezed functions are defined. From now on we let

$$\delta_\gamma := \min(1, \kappa^{-1} h_{min,\gamma}^{-1}) = h_{min,\gamma}^{-1} \min(h_{min,\gamma}, \kappa^{-1}).$$

For the modification of the equilibrated residual method that will be done in Section 6.1.6 we will need the modified energy scalar product and the modified energy norms defined as follows.

Definition 6.5 (Mesh dependent energy scalar product). *Let $K \in \mathcal{T}$ be any triangle, $u \in H_0^1(\Omega)$ and $v \in H^1(K)$, then we define the mesh-dependent energy scalar product and norms by*

$$\begin{aligned} mB_K(u, v) &:= h_{min,K}^{-2} \left(C_K^\top \nabla u, C_K^\top \nabla v \right)_K + \kappa^2 (u, v)_K, \\ m\|u\|_K &:= mB_K(u, u)^{1/2}, \\ m\|u\| &:= \left(\sum_{K \in \mathcal{T}} m\|u\|_K^2 \right)^{1/2}. \end{aligned}$$

The local mesh-dependent energy norm satisfies the following property

$$\|u\|_K \leq m\|u\|_K \leq \frac{h_{max,K}}{h_{min,K}} \|u\|_K.$$

Note that the standard and mesh dependent energy norms are equivalent in the case of isotropic elements.

Lemma 6.6 (Inverse inequalities for squeezed edge bubble/spline functions). *Let γ be an arbitrary face of K . Assume that $\varphi_\gamma \in \mathbb{P}^0(\gamma)$, $\mu_{\gamma,\delta_\gamma} \in \{b_{\gamma,\delta_\gamma}, s_{\gamma,\delta_\gamma}\}$ and*

$\nu_{\gamma, \delta_\gamma} \in \text{span}\{b_{\gamma, \delta_\gamma}, s_{\gamma, \delta_\gamma}\}$. Then the following inverse inequalities hold:

$$\|b_{\gamma, \delta_\gamma}^{1/2} \varphi_\gamma\|_{L_2(\gamma)} \sim \|\varphi_\gamma\|_{L_2(\gamma)} \quad (6.25)$$

$$\|\mu_{\gamma, \delta_\gamma} \varphi_\gamma\|_{L_2(\gamma)} \sim \|\varphi_\gamma\|_{L_2(\gamma)} \quad (6.26)$$

$$\|\mu_{\gamma, \delta_\gamma} F_{ext}(\varphi_\gamma)\|_{L_2(K)} \preceq \left(\frac{|K|}{|\gamma|}\right)^{1/2} h_{min, K}^{-1/2} \alpha_K^{1/2} \|\varphi_\gamma\|_{L_2(\gamma)} \quad (6.27)$$

$$\|\nabla(\mu_{\gamma, \delta_\gamma} F_{ext}(\varphi_\gamma))\|_{L_2(K)} \preceq \left(\frac{|K|}{|\gamma|}\right)^{1/2} h_{min, K}^{-1/2} \alpha_K^{-1/2} \|\varphi_\gamma\|_{L_2(\gamma)} \quad (6.28)$$

$$h_{min, K}^{-1} \|C_K^\top \nabla(\mu_{\gamma, \delta_\gamma} F_{ext}(\varphi_\gamma))\|_{L_2(K)} \preceq \left(\frac{|K|}{|\gamma|}\right)^{1/2} h_{min, K}^{-1/2} \alpha_K^{-1/2} \|\varphi_\gamma\|_{L_2(\gamma)} \quad (6.29)$$

$$\|\mu_{\gamma, \delta_\gamma} F_{ext}(\varphi_\gamma)\|_K \preceq \left(\frac{|K|}{|\gamma|}\right)^{1/2} h_{min, K}^{-1/2} \alpha_K^{-1/2} \|\varphi_\gamma\|_{L_2(\gamma)} \quad (6.30)$$

$$m \|\mu_{\gamma, \delta_\gamma} F_{ext}(\varphi_\gamma)\|_K \preceq \left(\frac{|K|}{|\gamma|}\right)^{1/2} h_{min, K}^{-1/2} \alpha_K^{-1/2} \|\varphi_\gamma\|_{L_2(\gamma)} \quad (6.31)$$

$$\|\nu_{\gamma, \delta_\gamma}\|_{L_2(K)} \preceq \left(\frac{|K|}{|\gamma|}\right)^{1/2} h_{min, K}^{-1/2} \alpha_K^{1/2} \|\nu_{\gamma, \delta_\gamma}\|_{L_2(\gamma)} \quad (6.32)$$

$$m \|\nu_{\gamma, \delta_\gamma}\|_K \preceq \left(\frac{|K|}{|\gamma|}\right)^{1/2} h_{min, K}^{-1/2} \alpha_K^{-1/2} \|\nu_{\gamma, \delta_\gamma}\|_{L_2(\gamma)}. \quad (6.33)$$

Proof. We observe first that (6.25) is identical to (2.26) in [31] (page 27, Lemma 2.7) and (6.26) can be obtained analogously. Inequalities (6.27) and (6.28) follow directly from the corresponding inequalities in [35] (page 247, Lemma 3.7) extended also to the spline functions $s_{\gamma, \delta_\gamma}$. Inequality (6.30) follows from (6.27) and (6.28) in the similar way as (2.10) in Lemma 2.1. The proof of refined estimates (6.29) and (6.31) follows the lines of the corresponding proof of (2.9). In order to show (6.32) we express $\nu_{\gamma, \delta_\gamma}$ in the form $\nu_{\gamma, \delta_\gamma} = C_b b_{\gamma, \delta_\gamma} + C_s s_{\gamma, \delta_\gamma}$, where C_b and C_s are two constants, and show (6.32) utilizing the triangle inequality, (6.27) and (6.26) subsequently:

$$\begin{aligned} \|\nu_{\gamma, \delta_\gamma}\|_{L_2(K)}^2 &= \|C_b b_{\gamma, \delta_\gamma} + C_s s_{\gamma, \delta_\gamma}\|_{L_2(K)}^2 \leq \|C_b b_{\gamma, \delta_\gamma}\|_{L_2(K)}^2 + \|C_s s_{\gamma, \delta_\gamma}\|_{L_2(K)}^2 \\ &\preceq \frac{|K|}{|\gamma|} h_{min, K}^{-1} \alpha_K \left(\|C_b\|_{L_2(\gamma)}^2 + \|C_s\|_{L_2(\gamma)}^2 \right) \\ &\sim \frac{|K|}{|\gamma|} h_{min, K}^{-1} \alpha_K \left(\|C_b b_{\gamma, \delta_\gamma}\|_{L_2(\gamma)}^2 + \|C_s s_{\gamma, \delta_\gamma}\|_{L_2(\gamma)}^2 \right) \\ &= \frac{|K|}{|\gamma|} h_{min, K}^{-1} \alpha_K \|C_b b_{\gamma, \delta_\gamma} + C_s s_{\gamma, \delta_\gamma}\|_{L_2(\gamma)}^2, \end{aligned}$$

where we used the orthogonality of functions $b_{\gamma, \delta_\gamma}$ and $s_{\gamma, \delta_\gamma}$ over edge γ . Estimate (6.33) may be obtained analogously. \square

In this section we prove two lemmas which we will need later. Namely, we derive the upper bounds for interior and face residuals. The jump discontinuity in the approximation of the normal flux at an interelement boundary is defined by

$$\left[\frac{\partial u_X}{\partial n} \right] := n_K (\nabla u_X)_K + n_{K'} (\nabla u_X)_{K'},$$

and the usual interior and boundary residuals r and R are given by

$$r := f + \Delta u_X - \kappa^2 u_X$$

and

$$R := \begin{cases} - \left[\frac{\partial u_X}{\partial n} \right] & \text{on } \partial K \cap \partial K' \\ 0 & \text{on } \partial K \cap \partial \Omega \end{cases}$$

Lemma 6.7 (Residuals estimates). *Let $K \in \mathcal{T}$ and γ be any interior edge. Then*

$$\|r\|_{L_2(K)} \preceq \alpha_K^{-1} \|e\|_K + \|r - \bar{r}\|_{L_2(K)}, \quad (6.34)$$

$$\|R\|_{L_2(\gamma)} \preceq \sum_{K' \in \omega_\gamma} \left(\frac{|K'|}{|\gamma|} \right)^{1/2} h_{\min, K'}^{-1/2} \alpha_{K'}^{-1/2} (\|e\|_{K'} + \alpha_{K'} \|r - \bar{r}\|_{L_2(K')}). \quad (6.35)$$

Proof. Let $v \in H_0^1(\Omega)$. Integrating by parts on each element yields

$$B(e, v) = \sum_{K \in \mathcal{T}} \int_K r v \, dx - \sum_{\gamma \in \partial \mathcal{T}} \int_\gamma R v \, ds, \quad (6.36)$$

where $\partial \mathcal{T}$ denotes the collection of interelement edges. Hence for any $v \in H_0^1(\Omega)$

$$B(e, v) = \sum_{K \in \mathcal{T}} \int_K \bar{r} v \, dx - \sum_{\gamma \in \partial \mathcal{T}} \int_\gamma R v \, ds + \sum_{K \in \mathcal{T}} \int_K (r - \bar{r}) v \, dx.$$

Now, choosing $v := b_K \bar{r}$ in the previous equality gives

$$\int_K b_K \bar{r}^2 \, dx = B_K(e, b_K \bar{r}) - \int_K (r - \bar{r}) b_K \bar{r} \, dx.$$

Using (2.6), with the aid of Cauchy-Schwarz inequality we obtain

$$\|\bar{r}\|_{L_2(K)}^2 \preceq \int_K b_K \bar{r}^2 \, dx \leq \|e\|_K \|b_K \bar{r}\|_K + \|r - \bar{r}\|_{L_2(K)} \|b_K \bar{r}\|_{L_2(K)}.$$

Estimates (2.7), (2.10) together with the triangle inequality imply (6.34).

We show now (6.35). Let $\gamma \in \partial \mathcal{T}$. Suppose that $\gamma = \bar{K}_1 \cap \bar{K}_2$. Then $\omega_\gamma = \text{int}(\bar{K}_1 \cup \bar{K}_2)$. Choosing $v := F_{\text{ext}}(R) b_{\gamma, \delta_\gamma} \in H_0^1(\Omega)$ in (6.36) implies

$$\int_\gamma b_{\gamma, \delta_\gamma} R^2 \, ds = \sum_{K \subset \omega_\gamma} \int_K r F_{\text{ext}}(R) b_{\gamma, \delta_\gamma} \, dx - B_{\omega_\gamma}(e, F_{\text{ext}}(R) b_{\gamma, \delta_\gamma}).$$

Furthermore, applying the Cauchy-Schwarz inequality and (6.25), one obtains

$$\|R\|_{L_2(\gamma)}^2 \preceq \int_\gamma b_{\gamma, \delta_\gamma} R^2 \, ds \leq \sum_{K \subset \omega_\gamma} \|r\|_{L_2(K)} \|F_{\text{ext}}(R) b_{\gamma, \delta_\gamma}\|_{L_2(K)} + \|e\|_{\omega_\gamma} \|F_{\text{ext}}(R) b_{\gamma, \delta_\gamma}\|_{\omega_\gamma}.$$

The desired inequality (6.35) follows now from (6.27), (6.30) and the first result of the current lemma (6.34). \square

Recall that we use the procedure for finding approximate fluxes described in Section 6.1.2 with the functions θ_n replaced by θ_n^* in the system (6.14). In the singularly perturbed case and using anisotropic elements we have the following theorem.

Theorem 6.8 (Stability of the approximate fluxes). *Suppose that the finite element subspace X is constructed using first-order (linear) elements on a partition \mathcal{T} of the domain Ω into triangular elements. Let $\{g_K\}$ be the set of approximate fluxes, produced*

by the algorithm described in Section 6.1.2 with the functions θ_n replaced by θ_n^* , $n \in \mathcal{N}$. Then, for each edge γ of any element K ,

$$\left\| g_K - \left\langle \frac{\partial u_X}{\partial n_K} \right\rangle \right\|_{L_2(\gamma)} \preceq \sum_{K' \subset \omega_K} \left(\frac{|K'|}{|\gamma|} \right)^{1/2} h_{\min, K'}^{-1/2} \alpha_{K'}^{-1/2} \times (\|e\|_{K'} + \alpha_{K'} \|r - \bar{r}\|_{L_2(K')}).$$

Proof. Let $K \in \mathcal{T}$ be a fixed element and $\gamma \subset K$ be an edge thereof. Then

$$\left(g_K - \left\langle \frac{\partial u_X}{\partial n_K} \right\rangle \right) \Big|_{\gamma} \in \mathbb{P}_1(\gamma).$$

Following Section 6.1.2 the moments of this quantity are

$$\mu_{K,n}^{*\gamma} = \int_{\gamma} \left(g_K - \left\langle \frac{\partial u_X}{\partial n_K} \right\rangle \right) \theta_n ds.$$

By analogy with (6.7),

$$\left(g_K - \left\langle \frac{\partial u_X}{\partial n_K} \right\rangle \right) \Big|_{\gamma} = \mu_{K,l}^{*\gamma} \psi_l + \mu_{K,r}^{*\gamma} \psi_r.$$

Therefore,

$$\left\| g_K - \left\langle \frac{\partial u_X}{\partial n_K} \right\rangle \right\|_{L_2(\gamma)} \leq |\mu_{K,l}^{*\gamma}| \|\psi_l\|_{L_2(\gamma)} + |\mu_{K,r}^{*\gamma}| \|\psi_r\|_{L_2(\gamma)}$$

and since

$$\|\psi_l\|_{L_2(\gamma)}^2 = \|\psi_r\|_{L_2(\gamma)}^2 = C|\gamma|^{-1},$$

it follows that

$$\left\| g_K - \left\langle \frac{\partial u_X}{\partial n_K} \right\rangle \right\|_{L_2(\gamma)}^2 \leq |\gamma|^{-1} \sum_{n \in \mathcal{N}(\gamma)} |\mu_{K,n}^{*\gamma}|^2. \quad (6.37)$$

With the aid of (6.16), we conclude that

$$\int_{\gamma} \left\langle \frac{\partial u_X}{\partial n_K} \right\rangle \theta_n ds = \begin{cases} \frac{1}{2} (\tilde{\mu}_{K,n}^{\gamma} - \tilde{\mu}_{K',n}^{\gamma}) & \text{on } \gamma = \partial K \cap \partial K' \\ \tilde{\mu}_{K,n}^{\gamma} & \text{on } \gamma = \partial K \cap \partial \Omega \end{cases}$$

with $\tilde{\mu}_{K,n}^{\gamma}$ defined in (6.10). Hence, thanks to (6.13),

$$\mu_{K,n}^{*\gamma} = \begin{cases} \frac{1}{2} (\sigma_{K,n} - \sigma_{K',n}) & \text{on } \gamma = \partial K \cap \partial K' \\ \sigma_{K,n} & \text{on } \gamma = \partial K \cap \partial \Omega \end{cases}$$

where the unknowns $\{\sigma_{K,n}\}$ are determined from conditions (6.14) and satisfy (6.17). It follows that

$$|\mu_{K,n}^{*\gamma}|^2 \preceq \sum_{K' \in \bar{x}_n} \sigma_{K',n}^2 \preceq \sum_{K' \in \bar{x}_n} \tilde{\Delta}_{K'}(\theta_n^*)^2. \quad (6.38)$$

The terms appearing on the right-hand side may be bounded by first recalling (6.15),

$$\tilde{\Delta}_{K'}(\theta_n^*) = B_{K'}(u_X, \theta_n^*) - (f, \theta_n^*)_{K'} - \int_{\partial K'} \left\langle \frac{\partial u_X}{\partial n_{K'}} \right\rangle \theta_n^* ds;$$

then, integrating by parts reveals that

$$\tilde{\Delta}_{K'}(\theta_n^*) = -(r, \theta_n^*)_{K'} - \int_{\partial K'} R \theta_n^* ds.$$

We proceed applying the Cauchy-Schwarz inequality and using Lemma 6.2 and Lemma 6.7:

$$\begin{aligned} \left| \tilde{\Delta}_{K'}(\theta_n^*) \right| &\leq \|r\|_{L_2(K')} \|\theta_n^*\|_{L_2(K')} + \sum_{\gamma' \subset \partial K' \cap \mathcal{E}_n} \|R\|_{L_2(\gamma')} \|\theta_n^*\|_{L_2(\gamma')} \\ &\preceq (\alpha_{K'}^{-1} \|e\|_{K'} + \|r - \bar{r}\|_{L_2(K')}) |K'|^{1/2} h_{min, K'}^{-1/2} \alpha_{K'}^{1/2} \\ &+ \sum_{\gamma' \subset \partial K' \cap \mathcal{E}_n} \sum_{K'' \subset \omega_{\gamma'}} |K''|^{1/2} h_{min, K''}^{-1/2} \alpha_{K'}^{-1/2} (\|e\|_{K''} + \alpha_{K'} \|r - \bar{r}\|_{L_2(K'')}) \\ &\preceq \sum_{K'' \in \tilde{\mathcal{X}}_n} |K''|^{1/2} h_{min, K''}^{-1/2} \left(\alpha_{K'}^{-1/2} \|e\|_{K''} + \alpha_{K'}^{1/2} \|r - \bar{r}\|_{L_2(K'')} \right). \end{aligned}$$

Hence,

$$\sum_{K' \in \tilde{\mathcal{X}}_n} \left| \tilde{\Delta}_{K'}(\theta_n^*) \right|^2 \preceq \sum_{K' \in \tilde{\mathcal{X}}_n} |K'| h_{min, K'}^{-1} \left(\alpha_{K'}^{-1} \|e\|_{K'}^2 + \alpha_{K'} \|r - \bar{r}\|_{L_2(K'')}^2 \right). \quad (6.39)$$

Combining (6.37), (6.38) and (6.39) leads to the result claimed. \square

6.1.5 Lower error bound of the original Ainsworth-Babuška estimator in the anisotropic singularly perturbed case

Describing in §6.1.1 the equilibrated residual method, we derived the upper error bound. The original analysis of the lower error bound for *isotropic* triangles dates back to the work by Ainsworth and Babuška [2]. Here we analyze the *anisotropic* case. It turns out that the original error estimator described in [2] has degenerating lower error bound.

The right hand side of the local problem (6.4) is originally defined as a linear functional only for the functions $v \in V_K$. We will need, however, to apply this functional also to the functions outside of V_K , namely to those not preserving the Dirichlet boundary conditions. To this end we introduce a new notation for the residual functional on the right hand side of (6.4):

$$\mathcal{B}_K(v) := (f, v)_K - B_K(u_X, v) + \int_{\partial K} g_K v ds \quad \forall v \in H^1(K). \quad (6.40)$$

From this notation it is clear that $\mathcal{B}_K(v) = B_K(\phi_K, v)$ for all $v \in V_K$, but the domain of definition of the functional \mathcal{B}_K is larger for the elements K touching the boundary. The next lemma states some stability properties of the residual functional \mathcal{B}_K which we will require in the lower error bound estimates.

Lemma 6.9. *Let ϕ_K denote the solution of the local residual problem (6.4) for the error estimator on element K . Then, for any $v \in H^1(K)$,*

$$|\mathcal{B}_K(v - \bar{v})| \preceq m_1(v, K) (\|e\|_{\omega_K} + \alpha_K \|r - \bar{r}\|_{L_2(\omega_K)}) \|v\|_K. \quad (6.41)$$

Furthermore, if $\kappa \succeq h_{min, K}^{-1}$, then

$$\begin{aligned} |\bar{\phi}_K \mathcal{B}_K(1)| &\preceq (\|e\|_{\omega_K} + \kappa^{-1} \|r - \bar{r}\|_{L_2(\omega_K)}) m \|\phi_K\|_K \\ &\leq m_1(\phi_K, K) (\|e\|_{\omega_K} + \kappa^{-1} \|r - \bar{r}\|_{L_2(\omega_K)}) \|\phi_K\|_K. \end{aligned} \quad (6.42)$$

Proof. 1. Integrating (6.40) by parts yields

$$\mathcal{B}_K(v - \bar{v}) = \int_K r(v - \bar{v}) dx + \frac{1}{2} \int_{\partial K} R(v - \bar{v}) ds + \int_{\partial K} \left(g_K - \left\langle \frac{\partial u_X}{\partial n_K} \right\rangle \right) (v - \bar{v}) ds.$$

and it therefore follows that

$$\begin{aligned} |\mathcal{B}_K(v - \bar{v})| &\leq \sum_{\gamma \subset \partial K} \left\| g_K - \left\langle \frac{\partial u_X}{\partial n_K} \right\rangle \right\|_{L_2(\gamma)} \|v - \bar{v}\|_{L_2(\gamma)} \\ &+ \|r\|_{L_2(K)} \|v - \bar{v}\|_{L_2(K)} + \frac{1}{2} \sum_{\gamma \subset \partial K} \|R\|_{L_2(\gamma)} \|v - \bar{v}\|_{L_2(\gamma)}. \end{aligned} \quad (6.43)$$

Combining results from Lemma 6.7, Theorem 6.8 and Lemma 6.4 we get (6.41).

2. Suppose that $\kappa \geq h_{min,K}^{-1}$. Then

$$\mathcal{B}_K(1) = (f, 1)_K - B_K(u_X, 1) + \int_{\partial K} g_K ds.$$

Integrating by parts, applying the Cauchy-Schwarz inequality, and estimating each term using Lemma 6.7 and Theorem 6.8 yield

$$\begin{aligned} |\mathcal{B}_K(1)| &\leq |K|^{1/2} \|r\|_{L_2(K)} + \frac{1}{2} \sum_{\gamma \in \partial K} |\gamma|^{1/2} \|R\|_{L_2(\gamma)} + \sum_{\gamma \in \partial K} |\gamma|^{1/2} \left\| g_K - \left\langle \frac{\partial u_X}{\partial n_K} \right\rangle \right\|_{L_2(\gamma)} \\ &\leq \kappa |K|^{1/2} (\|e\|_{\omega_K} + \kappa^{-1} \|r - \bar{r}\|_{L_2(\omega_K)}), \end{aligned}$$

where the inequality $\alpha_K \leq \kappa^{-1}$ has been used. Inequality (6.42) now can be easily obtained

$$\begin{aligned} |\bar{\phi}_K \mathcal{B}_K(1)| &\preceq \kappa |K|^{1/2} |\bar{\phi}_K| (\|e\|_{\omega_K} + \kappa^{-1} \|r - \bar{r}\|_{L_2(\omega_K)}) \\ &\preceq \kappa \|\phi_K\|_{L_2(K)} (\|e\|_{\omega_K} + \kappa^{-1} \|r - \bar{r}\|_{L_2(\omega_K)}) \leq (\|e\|_{\omega_K} + \kappa^{-1} \|r - \bar{r}\|_{L_2(\omega_K)}) m \|\phi_K\|_K. \end{aligned}$$

□

For the lower bound we have the following result.

Theorem 6.10 (Lower error bound). *Let g_K be the set of fluxes produced by the algorithm described in Section 6.1.2 with the functions θ_n replaced by θ_n^* , and let $\phi_K \in V_K$ denote the solution of the local residual problem (6.4). Then,*

$$\|\phi_K\|_K \preceq m_1(\phi_K, K) (\|e\|_{\omega_K} + \alpha_K \|r - \bar{r}\|_{L_2(\omega_K)}).$$

If κ vanishes, then α_K is replaced by $h_{min,K}$.

Proof. Observe that for any $v \in V_K$,

$$B_K(\phi_K, v) = \mathcal{B}_K(v) = \mathcal{B}_K(v - \bar{v}) + \bar{v} \mathcal{B}_K(1). \quad (6.44)$$

First, suppose $\kappa h_{min,K} \geq 1$ so that, in particular, κ is positive and $\alpha_K^{-1} \sim \kappa$. Therefore, with the aid of Lemma 6.9,

$$|\mathcal{B}_K(\phi_K - \bar{\phi}_K)| \leq m_1(\phi_K, K) \{ \|e\|_{\omega_K} + \alpha_K \|r - \bar{r}\|_{L_2(\omega_K)} \} \|\phi_K\|_K.$$

Choosing v to be equal to ϕ_K in (6.44), together with the above estimate, proves that the result holds for all elements K satisfying $\kappa h_{min,K} \geq 1$.

The remaining elements satisfy $\kappa h_{\min,K} \ll 1$. Thanks to the assumptions on the partition, the condition $\kappa h_{K'} \ll 1$ is satisfied by all elements K' contained in the patch ω_K . Therefore, Lemma 6.2 reveals that the modified basis functions reduce to the standard basis functions on the patch. Consequently, the approximate fluxes will actually satisfy the equilibration conditions (6.6) exactly. Moreover, since

$$B_K(1) = (f, 1)_K - B_K(u_X, 1) + \int_{\partial K} g_K ds = 0, \quad (6.45)$$

the second term in (6.44) vanishes. The first estimate in Lemma 6.9 then completes the proof. \square

Theorem 6.10 gives the lower error bound of the true error. The main danger for reliability of the estimator is the function $m_1(\phi_K, K)$ presented on the right hand side. One cannot guarantee that the approximation for the error ϕ_K is aligned as well as the true error e . Unfortunately, it may happen so that the alignment of the approximation ϕ_K on the element K is much worse than e : $m_1(\phi_K, K) \gg m_1(e, K)$. To avoid this problem a modification is proposed in the next paragraph.

6.1.6 Modified equilibrated residual method

For finding the equilibrated fluxes we use again the equilibrated residual method described in Section 6.1.2. In this paragraph, we propose an alternative method by changing the local problem, namely, instead of (6.4) we use

$${}_m B_K(\phi_K, v) = (f, v)_K - B_K(u_X, v) + \int_{\partial K} g_K v ds \quad \forall v \in V_K, \quad (6.46)$$

where on the left hand side stays the modified energy scalar product (see Definition 6.5). The modified local problem (6.46) differs from the original local problem (6.4) only in the scalar product ${}_m B_K(u, v)$ on the left hand side. The quantity ϕ_K is then not equivalent to the error e , but we will show that the ${}_m \|\phi_K\|_K$ is related to $\|e\|_K$. The following two theorems give upper and lower bounds for the error.

Theorem 6.11 (Reliability). *Let $\{g_K : K \in \mathcal{T}\}$ be any set of boundary fluxes satisfying condition (6.2). In addition, if the absolute term κ vanishes, then it is assumed that the fluxes satisfy the equilibration condition (6.5) on all elements that do not abut the boundary $\partial\Omega$. Then, the global error residual may be decomposed into local contributions*

$$B(e, v) = L(v) - B(u_X, v) = \sum_{K \in \mathcal{T}} {}_m B_K(\phi_K, v) \quad v \in H^1(K)$$

where $\phi_K \in V_K$ is the solution of the local problem (6.46). The global error in the finite element approximation may be bounded by

$$\|e\|^2 \leq m_1(e, \mathcal{T})^2 \sum_{K \in \mathcal{T}} {}_m \|\phi_K\|_K^2,$$

where $m_1(e, \mathcal{T})$ is the matching function introduced by (5.2), page 37.

Proof. Using the representation of $B(e, v)$ in the local terms and subsequently applying the Cauchy-Schwarz inequality and the definition of the matching function, we have:

$$\begin{aligned}
|B(e, v)| &= \left| \sum_{K \in \mathcal{T}} \{(f, v)_K - B_K(u_X, v)\} \right| \\
&= \left| \sum_{K \in \mathcal{T}} \{(f, v)_K - B_K(u_X, v) + \int_{\partial K} g_K v ds\} \right| \\
&= \left| \sum_{K \in \mathcal{T}} {}_m B_K(\phi_K, v) \right| \\
&\leq \sum_{K \in \mathcal{T}} m \|\phi_K\|_K m \|v\|_K \\
&\leq \sqrt{\sum_{K \in \mathcal{T}} m \|\phi_K\|_K^2} \sqrt{\sum_{K \in \mathcal{T}} (h_{\min, K}^{-2} \|C_K^\top \nabla v\|_{L_2(K)}^2 + \kappa^2 \|v\|_{L_2(K)}^2)} \\
&\leq \sqrt{\sum_{K \in \mathcal{T}} m \|\phi_K\|_K^2} \sqrt{m_1(v, \mathcal{T})^2 \|\nabla v\|_{L_2(\Omega)}^2 + \kappa^2 \|v\|_{L_2(\Omega)}^2} \\
&\leq m_1(v, \mathcal{T}) \|v\| \sqrt{\sum_{K \in \mathcal{T}} m \|\phi_K\|_K^2}.
\end{aligned}$$

The substitution $v := e$ completes the proof. \square

Theorem 6.11 gives the usual result for anisotropic error estimators. See for instance [35, 32].

Theorem 6.12 (Efficiency). *Let g_K be the set of approximate fluxes produced by the algorithm described in Section 6.1.2 with the functions θ replaced by θ^* , and let $\phi_K \in V_K$ denote the solution of the local residual problem (6.46). Then,*

$$m \|\phi_K\|_K \preceq \|e\|_{\tilde{K}} + \alpha_K \|r - \bar{r}\|_{L_2(\tilde{K})}.$$

Proof. The proof follows the same lines as the proof of the Theorem 6.10. \square

These two theorems are part of the main result of this work and guarantee the reliability and efficiency of the estimator, assuming an exact solution of the local problems.

6.1.7 Computable approximation for the solution to the local problem

Up to this time we considered the infinite dimensional local problems (6.4) and (6.46). The author has not found any result in the literature saying that some computable approximation $\tilde{\phi}_K$ is equivalent to ϕ_K in the energy norm,

$$\|\tilde{\phi}_K\| \sim \|\phi_K\|, \quad (6.47)$$

even not for isotropic elements. In the current section we construct an approximation $\tilde{\phi}_K$ so that (6.47) holds. To this end we restrict the space V_K to the space of bubbles/splines $V_b(K)$ defined in the following way:

$$V_b(K) := \text{span}\{b_K, b_{\gamma, \delta_\gamma}, s_{\gamma, \delta_\gamma} : \gamma \in \partial K \setminus \partial\Omega\} \subset V_K.$$

By means of the space $V_b(K)$ we can define the function $\tilde{\phi}_K \in V_b(K)$ as the solution of the local finite dimensional problem

$${}_m B_K(\tilde{\phi}_K, v) = (f, v)_K - B_K(u_X, v) + \int_{\partial K} g_K v ds \quad \forall v \in V_b(K) \quad (6.48)$$

We prove the lower bound for ϕ_K first.

Theorem 6.13. Let ϕ_K and $\tilde{\phi}_K$ be the solutions to the problems (6.46) and (6.48), respectively. Then

$$m \|\tilde{\phi}_K\|_K \leq m \|\phi_K\|_K.$$

Proof. Substituting $v = \tilde{\phi}_K \in V_b(K) \subset V_K$ in (6.48) and subsequently utilizing the Cauchy-Schwarz inequality we estimate $mB_K(\tilde{\phi}_K, \tilde{\phi}_K)$ as follows:

$$\begin{aligned} mB_K(\tilde{\phi}_K, \tilde{\phi}_K) &= (f, \tilde{\phi}_K)_K - B_K(u_X, \tilde{\phi}_K) + \int_{\partial K} g_K \tilde{\phi}_K \, ds \\ &= mB_K(\phi_K, \tilde{\phi}_K) \leq m \|\phi_K\|_K m \|\tilde{\phi}_K\|_K. \end{aligned}$$

Dividing both sides by $m \|\tilde{\phi}_K\|_K$ we get the result claimed. \square

For further investigations we need the following preliminary result.

Lemma 6.14. Let ϕ_K denote the solution to the local residual problem (6.46) for the error estimator on the element K . If $\kappa \geq h_{\min, K}^{-1}$, then

$$|\bar{\phi}_K \mathcal{B}_K(1)| \preceq \left(\kappa^{-1} \|r\|_{L_2(K)} + \sum_{\gamma \subset \partial K} \kappa^{-1} \left(\frac{|\gamma|}{|K|} \right)^{1/2} \left\| g_K - \frac{\partial u_x}{\partial n_K} \right\|_{L_2(\gamma)} \right) m \|\phi_K\|_K, \quad (6.49)$$

where \mathcal{B}_K is the residual functional defined in (6.40).

Proof. The proof is done analogously to the proof of (6.42) in Lemma 6.9. \square

For the upper bound we employ the technique that is usually used in obtaining the lower error bound in the residual a posteriori error estimation (see e.g. [12]).

Lemma 6.15. Let ϕ_K be the solution of (6.46) and let $\bar{\phi}_K$ be its mean value over the triangle K . Then the following estimate holds:

$$\begin{aligned} \mathcal{B}_K(\phi_K - \bar{\phi}_K) &\preceq m \|\phi_K\|_K \left(\alpha_K \|r\|_{L_2(K)} \right. \\ &\quad \left. + \sum_{\gamma \subset \partial K} h_{\min, K}^{1/2} \left(\frac{|\gamma|}{|K|} \right)^{1/2} \alpha_K^{1/2} \left\| g_K - \frac{\partial u_X}{\partial n_K} \right\|_{L_2(\gamma)} \right). \end{aligned}$$

Proof. We use the definition of the residual functional (6.40) and apply subsequently the Cauchy-Schwarz inequality and the anisotropic approximation properties (6.22) and (6.24) obtained in Lemma 6.4:

$$\begin{aligned} \mathcal{B}_K(\phi_K - \bar{\phi}_K) &= (f, \phi_K - \bar{\phi}_K)_K - B_K(u_X, \phi_K - \bar{\phi}_K) + \int_{\partial K} g_K (\phi_K - \bar{\phi}_K) \, ds \\ &= (r, \phi_K - \bar{\phi}_K)_K + \int_{\partial K} \left(g_K - \frac{\partial u_X}{\partial n_K} \right) (\phi_K - \bar{\phi}_K) \, ds \\ &\leq \|r\|_{L_2(K)} \|\phi_K - \bar{\phi}_K\|_{L_2(K)} + \sum_{\gamma \subset \partial K} \left\| g_K - \frac{\partial u_X}{\partial n_K} \right\|_{L_2(\gamma)} \|\phi_K - \bar{\phi}_K\|_{L_2(\gamma)} \\ &\preceq m \|\phi_K\|_K \alpha_K \|r\|_{L_2(K)} \\ &\quad + m \|\phi_K\|_K \sum_{\gamma \subset \partial K} h_{\min, K}^{1/2} \left(\frac{|\gamma|}{|K|} \right)^{1/2} \alpha_K^{1/2} \left\| g_K - \frac{\partial u_X}{\partial n_K} \right\|_{L_2(\gamma)}, \end{aligned}$$

which completes the proof. \square

Lemma 6.16. *Let ϕ_K be the solution to (6.46). Then the following estimate holds:*

$$m \|\phi_K\|_K \leq \alpha_K \|r\|_{L_2(K)} + \sum_{\gamma \subset \partial K} h_{\min,K}^{1/2} \left(\frac{|\gamma|}{|K|} \right)^{1/2} \alpha_K^{1/2} \left\| g_K - \frac{\partial u_X}{\partial n_K} \right\|_{L_2(\gamma)}$$

Proof. We represent $m \|\phi_K\|_K^2$ as a sum of two terms:

$$m \|\phi_K\|_K^2 = {}_m B_K(\phi_K, \phi_K) = \mathcal{B}_K(\phi_K) = \mathcal{B}_K(\phi_K - \bar{\phi}_K) + \mathcal{B}_K(\bar{\phi}_K). \quad (6.50)$$

In Lemma 6.15 we constructed already the estimate from above for the first term. We proceed estimating the second term analogously to the proof of Theorem 6.10.

Consider the case $\kappa h_{\min,K} \geq 1$. Thus, the second assertion of Lemma 6.9 holds, i.e.

$$\begin{aligned} \mathcal{B}_K(\bar{\phi}_K) &\leq m \|\phi_K\|_K \left(\kappa^{-1} \|r\|_{L_2(K)} + \sum_{\gamma \subset \partial K} \left(\frac{|\gamma|}{|K|} \right)^{1/2} \kappa^{-1} \left\| g_K - \frac{\partial u_X}{\partial n_K} \right\|_{L_2(\gamma)} \right) \\ &\leq m \|\phi_K\|_K \left(\alpha_K \|r\|_{L_2(K)} \right. \\ &\quad \left. + \sum_{\gamma \subset \partial K} h_{\min,K}^{1/2} \left(\frac{|\gamma|}{|K|} \right)^{1/2} \alpha_K^{1/2} \left\| g_K - \frac{\partial u_X}{\partial n_K} \right\|_{L_2(\gamma)} \right), \end{aligned}$$

where we used $\kappa \geq h_{\min,K}^{-1}$.

The remaining elements satisfy $\kappa h_{\min,K} \ll 1$. With the same arguments as in the proof of Theorem 6.10 we verify that for these elements (6.45) holds, and thus, the second term in (6.50) vanishes.

Summing up the contributions from the two terms on the right hand side of representation (6.50) and dividing the concluding inequality by $m \|\phi_K\|_K$ we get the result claimed. \square

In order to prove the main theorem of this section we need some additional elementary facts.

Lemma 6.17. *Let $\phi \in P_1([-1, 1])$ be a linear function. For the L_2 -projection operator $I : P_1([-1, 1]) \rightarrow \text{span}\{1 - x^2, x(1 - x^2)\}$ the following inequality holds:*

$$\sqrt{\frac{6}{5}} \|I\phi\|_{L_2([-1, 1])} \leq \|\phi\|_{L_2([-1, 1])} \leq \sqrt{\frac{10}{7}} \|I\phi\|_{L_2([-1, 1])}.$$

Proof. Denote $\psi_1 := 1 - x$ and $\psi_2 := 1 + x$. The desired constants are the square roots of the maximal and minimal eigenvalues of the eigenvalue problem $Ax = \lambda Bx$, where $A = \left\{ \int_{-1}^1 \psi_i \psi_j dx \right\} \in \mathbb{R}^{2 \times 2}$, $B = \left\{ \int_{-1}^1 I\psi_i I\psi_j dx \right\} \in \mathbb{R}^{2 \times 2}$ and $x \in \mathbb{R}^2$. \square

Lemma 6.18. *Let γ be an edge of a triangle K . We define the operator $I_\gamma : P_1(\gamma) \rightarrow \text{span}\{b_{\gamma, \delta_\gamma}, s_{\gamma, \delta_\gamma}\}$ so that for any function $\phi \in P_1(\gamma)$ the restriction of the resulting function $I_\gamma \phi|_\gamma$ to the edge γ is the L_2 projection of ϕ to the space $\text{span}\{b_{\gamma, \delta_\gamma}|_\gamma, s_{\gamma, \delta_\gamma}|_\gamma\}$. In other words, we project a function ϕ onto the space of two functions on the edge γ and then take the corresponding constant to produce the function $I_\gamma \phi$ inside the triangle K . For the operator I_γ and any function $\phi \in P_1(\gamma)$ the following estimate holds:*

$$\sqrt{\frac{6}{5}} \|I_\gamma \phi\|_{L_2(\gamma)} \leq \|\phi\|_{L_2(\gamma)} \leq \sqrt{\frac{10}{7}} \|I_\gamma \phi\|_{L_2(\gamma)}.$$

Proof. The functions $1 - x^2$ and $x(1 - x^2)$ coincide with the functions $b_{\gamma, \delta_\gamma}|_\gamma$ and $s_{\gamma, \delta_\gamma}|_\gamma$ whenever $\gamma = [-1, 1]$. Thus, the assertion follows from Lemma 6.17 and standard transformation techniques. \square

We proceed with the main result of this section, that guarantees the estimate from above for the solution of the local problem (6.46).

Theorem 6.19. *Let ϕ_K and $\tilde{\phi}_K$ be the solutions to the problems (6.46) and (6.48) respectively. Then:*

$$m \|\phi_K\|_K \leq m \|\tilde{\phi}_K\|_K + \alpha_K \|r - \bar{r}\|_{L_2(K)}.$$

Proof. We will essentially use the estimate of Lemma 6.16 and bound the terms on the right hand side by $m \|\tilde{\phi}_K\|_K$.

Performing the partial integration we can rewrite the finite dimensional local problem (6.48) as

$$\begin{aligned} m B_K(\tilde{\phi}_K, v) &= (r, v)_K + \int_{\partial K} \left(g_K - \frac{\partial u_X}{\partial n_K} \right) v \, ds \\ &= (r_K, v)_K + \int_{\partial K} \left(g_K - \frac{\partial u_X}{\partial n_K} \right) v \, ds + (r - r_K, v)_K. \end{aligned} \quad (6.51)$$

Since (6.51) holds for all $v \in V_b(K)$, we substitute $v := b_K \bar{r}$,

$$\int_K b_K \bar{r}^2 \, dx = m B_K(\phi_K, b_K \bar{r}) - (r - r_K, b_K \bar{r})_K.$$

Using (2.6), we obtain with the aid of Cauchy-Schwarz inequality

$$\|\bar{r}\|_{L_2(K)}^2 \leq m \|\tilde{\phi}_K\|_K m \|b_K \bar{r}\|_K + \|r - \bar{r}\|_{L_2(K)} \|b_K \bar{r}\|_{L_2(K)}.$$

Applying estimates (2.7) and (2.11) of Lemma 2.1 and dividing both parts by $\alpha_K^{-1} \|\bar{r}\|_{L_2(K)}$, we get

$$\alpha_K \|\bar{r}\|_{L_2(K)} \leq m \|\tilde{\phi}_K\|_K + \alpha_K \|r - \bar{r}\|_{L_2(K)},$$

or, with the aid of the triangle inequality we derive the upper bound for the term involving the element residual:

$$\alpha_K \|r\|_{L_2(K)} \leq m \|\tilde{\phi}_K\|_K + \alpha_K \|r - \bar{r}\|_{L_2(K)}.$$

Suppose $\gamma \subset \partial K \setminus \partial\Omega$ is one of the edges of K , which is not a Dirichlet edge. We choose now $v := I_\gamma \left(g_K - \frac{\partial u_X}{\partial n_K} \right) \in V_b(K)$ in (6.51), where the operator I_γ was defined in Lemma 6.18.

$$m B_K \left(\tilde{\phi}_K, I_\gamma \left(g_K - \frac{\partial u_X}{\partial n_K} \right) \right) = \left(r, I_\gamma \left(g_K - \frac{\partial u_X}{\partial n_K} \right) \right)_K + \int_{\partial K} \left(g_K - \frac{\partial u_X}{\partial n_K} \right) I_\gamma \left(g_K - \frac{\partial u_X}{\partial n_K} \right) \, ds. \quad (6.52)$$

Since I_γ is the L_2 -projection operator over γ , we rewrite the last equality in the form:

$$\left\| I_\gamma \left(g_K - \frac{\partial u_X}{\partial n_K} \right) \right\|_{L_2(\gamma)}^2 = m B_K \left(\tilde{\phi}_K, I_\gamma \left(g_K - \frac{\partial u_X}{\partial n_K} \right) \right) - \left(r, I_\gamma \left(g_K - \frac{\partial u_X}{\partial n_K} \right) \right)_K, \quad (6.53)$$

and utilizing the Cauchy-Schwarz inequality, we obtain:

$$\left\| I_\gamma \left(g_K - \frac{\partial u_X}{\partial n_K} \right) \right\|_{L_2(\gamma)}^2 \leq m \|\tilde{\phi}_K\|_K m \left\| I_\gamma \left(g_K - \frac{\partial u_X}{\partial n_K} \right) \right\|_K + \|r\|_{L_2(K)} \left\| I_\gamma \left(g_K - \frac{\partial u_X}{\partial n_K} \right) \right\|_{L_2(K)}. \quad (6.54)$$

Evaluating (6.32) and (6.33) of Lemma 6.6 together with Lemma 6.18, we complete the proof. \square

To complete the discussion of the equilibrated residual method it only remains to give a definition for the error estimator that can be used in practical computations and to give the resulting bounds provided by the preceding theory.

Definition 6.20. We define the local estimator of the equilibrated residual method and the local higher order term corresponding to the element K by

$$\begin{aligned}\eta_{ER,K} &:= m \|\tilde{\phi}_K\|_K, \\ \zeta_K &:= \alpha_K \|r - \bar{r}\|_{L_2(K)},\end{aligned}$$

with the global counterparts

$$\begin{aligned}\eta_{ER} &:= \left(\sum_{K \in \mathcal{T}} \eta_{ER,K}^2 \right)^{1/2} \quad \text{and} \\ \zeta &:= \left(\sum_{K \in \mathcal{T}} \zeta_K^2 \right)^{1/2},\end{aligned}$$

respectively.

Combining the bounds from Theorems 6.11, 6.12, 6.13 and 6.19 we arrive at the robustness result for the error estimator defined above.

Theorem 6.21. In notation of Definition 6.20 the upper and lower error bounds hold

$$\begin{aligned}\|e\|^2 &\preceq m_1^2(e, \mathcal{T}) (\eta_{ER}^2 + \zeta^2), \\ \eta_{ER,K}^2 &\preceq \|e\|_{\omega_K}^2 + \sum_{K \subset \omega_K} \zeta_K^2.\end{aligned}$$

This theorem is the final result of this work and guarantees the reliability and efficiency of the estimator. The bounds are in accordance with those provided by Kunert [35] for the Dirichlet local problem error estimator for the singularly perturbed reaction-diffusion equation.

6.1.8 Numerical experiments

In Section 6.1.7 we gave an example for the bases for the local problem (6.46) consisting of the very specially squeezed functions. We have good experience, however, in solving this problem with a finite element method, where we choose the nodal basis corresponding to a division of the triangles into n^2 parts, see Figure 13.

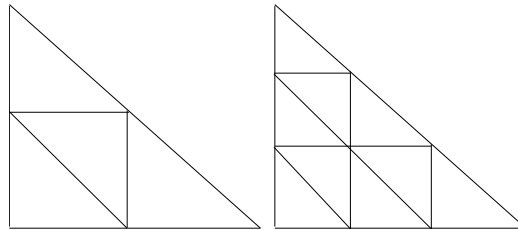


Figure 13: Triangle subdivisions. $n = 2$ and $n = 3$, respectively.

Let us consider the 2D model problem

$$-\Delta u + \kappa^2 u = 0 \text{ in } \Omega := [0, 1]^2, \quad u = u_0 \text{ on } \partial\Omega.$$

Prescribe the exact solution

$$u = e^{-\kappa x} + e^{-\kappa y}$$

which displays typical boundary layers along the sides $x = 0$ and $y = 0$. The Dirichlet boundary data u_0 are chosen accordingly.

We use a sequence of finite element meshes generated by the algorithm described in [7]. The idea of adaptive procedure is that the choice of a refinement direction is done according to the components of energy norm of an error $\left\| \frac{\partial e}{\partial x} \right\|_{L_2(K)}$, $\left\| \frac{\partial e}{\partial y} \right\|_{L_2(K)}$, and $\kappa^2 \|e\|_{L_2(K)}$. One of the resulting meshes of this program is displayed in Figure 14.

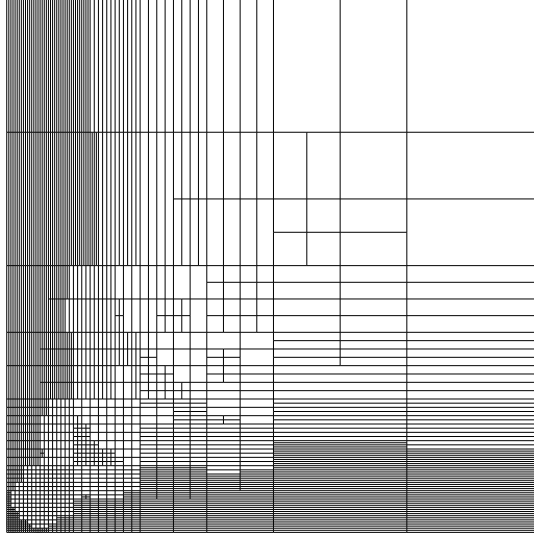


Figure 14: Mesh refinement.

Tables 1 and 2 show the behavior of the estimators in the singularly perturbed case on anisotropic meshes. We observe that the new error estimator is robust while the original one overestimates the true error when the aspect ratio is large enough.

6.2 Residual error estimator

After some additional notation we formulate an upper error bound for the error measured in the energy norm. The jump discontinuity in the approximation of the normal flux at an interelement boundary is defined by

$$\left[\frac{\partial u_1}{\partial n} \right] = n_K (\nabla u_1)_K + n_{K'} (\nabla u_1)_{K'},$$

and the usual interior and boundary residuals r and R are given by

$$r = f + \varepsilon^2 \Delta u_1 - \kappa^2 u_1$$

and

$$R = \begin{cases} \varepsilon^2 \left[\frac{\partial u_1}{\partial n} \right] & \text{on } \partial K \cap \partial K' \\ 0 & \text{on } \partial K \cap \partial \Omega \end{cases}$$

which are defined as usually (see [4]). Define by $r_K := \frac{1}{|K|} \int_K r \, dx$ the mean value of r over an element K .

Iteration N	Unknowns N	Maximal aspect ratio	$\frac{AB \text{ error}}{\text{exact err}}$	$\frac{AB \text{ er(mod)}}{\text{exact err}}$
1	25	71	1.392	1.102
2	51	71	1.301	0.896
3	97	71	1.565	0.962
4	116	142	1.903	1.032
5	157	285	2.457	1.094
6	217	571	3.151	1.153
7	382	1142	4.264	1.169
8	714	2284	5.955	1.167
9	1481	4568	8.903	1.157
10	3274	9137	14.475	1.130
11	6847	18273	23.967	1.112
12	15187	36547	44.111	1.090
13	35536	73095	81.237	1.062
14	106819	146191	138.711	1.005

Table 1: Results for $n = 4$, $\kappa = 1000$. The fourth column represents the ratio between the Ainsworth and Babuška estimator and the energy norm of the true solution, while the fifth column represents the similar ratio for the estimator defined in the current work.

Iteration N	Unknowns N	Maximal aspect ratio	$\frac{AB \text{ error}}{\text{exact err}}$	$\frac{AB \text{ er(mod)}}{\text{exact err}}$
1	25	541	2.665	2.150
2	51	541	2.177	1.587
3	120	541	1.802	1.198
4	143	541	1.995	1.187
5	192	1083	2.410	1.243
6	217	2167	3.062	1.302
7	283	4335	4.273	1.392
8	446	8669	6.207	1.436
9	814	17339	9.748	1.405
10	1553	34679	16.646	1.411
11	3053	69359	30.149	1.433
12	5809	138718	54.363	1.429
13	11357	277436	101.47	1.420
14	23376	554873	211.02	1.407
15	104916	1109745	423.10	1.383

Table 2: Results for $n = 4$, $\kappa = 10000$.

In addition we recall the following notation:

$$\begin{aligned}\alpha_K &:= \min(\varepsilon^{-1}h_{\min,K}, \kappa^{-1}), \\ \alpha_\gamma &:= \min(\varepsilon^{-1}h_{\min,\gamma}, \kappa^{-1}).\end{aligned}$$

Theorem 6.22. *There exists a constant C that depends neither on the geometry of the mesh \mathcal{T} nor on the perturbation parameters, such that*

$$\begin{aligned}\|u - u_1\| &\preceq m_1(u - u_1, \mathcal{T}) \left\{ \sum_{K \in \mathcal{T}} \alpha_K^2 \|r_K\|_{L_2(K)}^2 \right. \\ &\quad \left. + \sum_{\gamma \in \partial\mathcal{T}} \varepsilon^{-1} \alpha_\gamma \|R\|_{L_2(\gamma)}^2 + \sum_{K \in \mathcal{T}} \alpha_K^2 \|r - r_K\|_{L_2(K)}^2 \right\}^{1/2},\end{aligned}$$

where $\partial\mathcal{T}$ denote the collection of all edges in the triangulation \mathcal{T} .

Proof. See the proof for the anisotropic case in Kunert [34]. For the isotropic case it appeared first in Verfürth [48]. \square

6.3 Error reduction property

In order to obtain an error reduction property in this chapter we will require the inverse inequalities with for the squeezed bubble function with the squeezing parameter which differs to the one defined in Section 6.1.

Lemma 6.23. (Inverse inequalities for bubble functions and special edge bubble functions). *Let γ be an arbitrary edge of K . Assume that $\varphi_K \in \mathbb{P}^0(K)$ and $\varphi_\gamma \in \mathbb{P}^0(\gamma)$. Then the following inverse inequalities hold:*

$$\|\nabla(b_K \varphi_K)\|_{L_2(K)} \sim h_{\min,K}^{-1} \|\varphi_K\|_{L_2(K)} \quad (6.55)$$

$$\|F_{ext}(\varphi_\gamma) b_{\gamma,\delta}\|_{L_2(K)} \sim \left(\frac{|K|}{|\gamma|}\right)^{1/2} \delta^{1/2} \|\varphi_\gamma\|_{L_2(\gamma)} \quad (6.56)$$

$$\|\nabla(F_{ext}(\varphi_\gamma) b_{\gamma,\delta})\|_{L_2(K)} \sim \left(\frac{|K|}{|\gamma|}\right)^{1/2} \delta^{1/2} \min\left\{\delta \frac{|K|}{|\gamma|}, h_{\min,K}\right\}^{-1} \|\varphi_\gamma\|_{L_2(\gamma)}. \quad (6.57)$$

Proof. See [34]. \square

We are in a position to specify our parameter $\delta = \delta(\gamma)$. From now on in this section we use

$$\delta_\gamma := \frac{1}{3} \frac{|\gamma|}{|\omega_\gamma|} \varepsilon \min(\varepsilon^{-1}h_{\min,\gamma}, \kappa^{-1}). \quad (6.58)$$

Note that if $\gamma = \partial K \cap \partial K'$, then

$$\delta_\gamma \sim \frac{|\gamma|}{|K|} \varepsilon \min(\varepsilon^{-1}h_{\min,K}, \kappa^{-1}) \sim \frac{|\gamma|}{|K'|} \varepsilon \min(\varepsilon^{-1}h_{\min,K'}, \kappa^{-1}).$$

We should mention that the definition (6.58) differs from the original definition in Kunert [34] by a factor of $\frac{1}{3}$, which however does not disturb the estimates. This modification is done in order to avoid overlapping supports of special edge bubble functions.

In the case of a singularly perturbed problem the choice of space enrichment is crucial. First, recall the definition of the space V_1 :

$$V_1 := \{v_h \in H_0^1(\Omega) : \forall K \in \mathcal{T}, v_h|_K \in P_1(K)\}.$$

We enrich the space V_1 by the squeezed bubble functions for all edges and the interior bubbles. Namely,

$$V_2 := \{v_h \in H_0^1(\Omega) : \forall K \in \mathcal{T}, v_h|_K \in P_1(K) + \text{span}\{b_K, b_{\gamma, \delta_\gamma} : \gamma \in \partial K \setminus \partial\Omega\}\}.$$

Then the finite element solution $u_2 \in V_2$ is uniquely defined by

$$B(u_2, v) = (f, v) \quad \forall v \in V_2. \quad (6.59)$$

It is not clear at the moment whether we get the estimate similar to the estimate of Theorem 6.25 using the usual bubbles as it was done for example in [24] for the Laplace problem.

The proof of the error reduction property is based on the following lemma.

Lemma 6.24. *Let the functions u_1 and u_2 be the solutions to (2.3) and (6.59) respectively. Then the upper error bound*

$$\|u - u_1\|^2 \leq Cm_1(u - u_1, \mathcal{T})^2 \left(\|u_1 - u_2\|^2 + \sum_{K \in \mathcal{T}} \alpha_K^2 \|r - r_K\|_{L_2(K)}^2 \right)$$

holds.

Proof. Using the Theorem 6.55 we estimate the terms involving boundary and interior residual subsequently.

1. *Boundary residual.* Due to the fact that R is constant over each edge γ applying partial integration we get:

$$\begin{aligned} \frac{2}{3}|\gamma|R &= \int_{\gamma} R b_{\gamma, \delta_\gamma} ds = - \int_{\omega_\gamma} \nabla u_1 \nabla b_{\gamma, \delta_\gamma} dx \\ &= \varepsilon^2 \int_{\omega_\gamma} \nabla(u_2 - u_1) \nabla b_{\gamma, \delta_\gamma} dx + \kappa^2 \int_{\omega_\gamma} (u_2 - u_1) b_{\gamma, \delta_\gamma} dx - \int_{\omega_\gamma} f b_{\gamma, \delta_\gamma} dx + \kappa^2 \int_{\omega_\gamma} u_1 b_{\gamma, \delta_\gamma} dx \\ &= B_{\omega_\gamma}(u_2 - u_1, b_{\gamma, \delta_\gamma}) - \int_{\omega_\gamma} r b_{\gamma, \delta_\gamma} dx, \end{aligned}$$

where ω_γ is the union of two triangles sharing the edge γ (see Section 2.2). Squaring and integrating over γ we get

$$\begin{aligned} |\gamma| \|R\|_{L_2(\gamma)}^2 &\leq B_{\omega_\gamma}(u_2 - u_1, b_{\gamma, \delta_\gamma})^2 + \left(\int_{\omega_\gamma} r b_{\gamma, \delta_\gamma} dx \right)^2 \\ &\leq \|u_2 - u_1\|_{\omega_\gamma}^2 \|b_{\gamma, \delta_\gamma}\|_{\omega_\gamma}^2 + \|r\|_{L_2(\omega_\gamma)}^2 \|b_{\gamma, \delta_\gamma}\|_{L_2(\omega_\gamma)}^2 \end{aligned}$$

Estimate the first term using the inequalities for the special bubble functions (6.56), (6.57) and the definition of δ_γ (6.58) as follows:

$$\begin{aligned} \|b_{\gamma, \delta_\gamma}\|_{\omega_\gamma}^2 &= \varepsilon^2 \|\nabla b_{\gamma, \delta_\gamma}\|_{L_2(\omega_\gamma)}^2 + \kappa^2 \|b_{\gamma, \delta_\gamma}\|_{L_2(\omega_\gamma)}^2 \\ &\leq \sum_{K \subset \omega_\gamma} \left(\varepsilon^2 \frac{|K|}{|\gamma|} \delta_\gamma \min \left\{ \delta_\gamma \frac{|K|}{|\gamma|}, h_{\min, K} \right\}^{-2} |\gamma| + \kappa^2 \frac{|K|}{|\gamma|} \delta_\gamma |\gamma| \right) \\ &\sim \sum_{K \subset \omega_\gamma} \left(\varepsilon^3 \alpha_K \min \{ \varepsilon \alpha_K, h_{\min, K} \}^{-2} |\gamma| + \kappa^2 \varepsilon \alpha_K |\gamma| \right) \\ &\leq \varepsilon |\gamma| \alpha_\gamma^{-1} \end{aligned}$$

Estimate the second term using (6.56):

$$\|b_{\gamma, \delta_\gamma}\|_{L_2(\omega_\gamma)}^2 \preceq \frac{|K|}{|\gamma|} \delta_\gamma |\gamma| \sim |\gamma| \varepsilon \alpha_\gamma$$

Combining three previous estimates we come to the following:

$$\begin{aligned} \varepsilon^{-1} \alpha_\gamma \|R\|_{L_2(\gamma)}^2 &\preceq \|u_2 - u_1\|_{\omega_\gamma}^2 + \alpha_\gamma^2 \|r\|_{L_2(\omega_\gamma)}^2 \\ &\preceq \|u_2 - u_1\|_{\omega_\gamma}^2 + \sum_{K \subset \omega_\gamma} \alpha_K^2 \|r_K\|_{L_2(K)}^2 + \sum_{K \subset \omega_\gamma} \alpha_K^2 \|r - r_K\|_{L_2(K)}^2. \end{aligned}$$

2. *Interior residual.* It remains to estimate the term $\alpha_K^2 \|r_K\|_{L_2(\omega_\gamma)}^2$. We have

$$\begin{aligned} r_K \frac{|K|}{5!} &= \int_K r_K b_K dx \\ &= \int_K f b_K dx - \kappa^2 \int_K u_1 b_K dx - \int_K r b_K dx + \int_K r_K b_K dx \\ &= B_K(u_2, b_K) - B_K(u_1, b_K) - \int_K (r - r_K) b_K dx \\ &= B_K(u_2 - u_1, b_K) - \int_K (r - r_K) b_K dx, \end{aligned}$$

because $\int_K \nabla u_1 \nabla b_K dx = - \int_K \Delta u_1 b_K dx = 0$. Squaring and integrating over an element K we get:

$$\begin{aligned} |K| \|r_K\|_{L_2(K)}^2 &\preceq B_K(u_2 - u_1, b_K)^2 + \left(\int_K (r - r_K) b_K dx \right)^2 \\ &\leq \|u_2 - u_1\|_K^2 \|b_K\|_K^2 + \|r - r_K\|_{L_2(K)}^2 \|b_K\|_{L_2(K)}^2 \end{aligned}$$

Now we use (6.55) for $\|b_K\|_K$ as follows

$$\begin{aligned} \|b_K\|_K^2 &= \varepsilon^2 \|\nabla(b_K)\|_{L_2(K)}^2 + \kappa^2 \|b_K\|_{L_2(K)}^2 \\ &\preceq \left(\varepsilon^2 h_{\min, K}^{-2} + \kappa^2 \right) \|b_K\|_{L_2(K)}^2 \\ &\preceq \left(\varepsilon^2 h_{\min, K}^{-2} + \kappa^2 \right) |K| \end{aligned}$$

or,

$$\|b_K\|_K^2 \preceq \alpha_K^{-2} |K|.$$

Thus, it follows that

$$\alpha_K^2 \|r_K\|_{L_2(K)}^2 \preceq \|u_2 - u_1\|_K^2 + \alpha_K^2 \|r - r_K\|_{L_2(K)}^2 \quad (6.60)$$

Now, applying the inequalities (6.60) and (6.60) to the estimate of the Theorem 6.55 we get the result claimed. \square

Theorem 6.25 (Error reduction property on anisotropic meshes). *The following inequality takes place:*

$$\|u - u_2\| \leq \sqrt{1 - \frac{1}{C m_1(u - u_1, \mathcal{T})^2}} \|u - u_1\| + \left(\sum_{K \in \mathcal{T}} \alpha_K^2 \|r - r_K\|_{L_2(K)}^2 \right)^{1/2}. \quad (6.61)$$

Proof. Using the identity

$$\|u - u_1\|^2 = \|u - u_2\|^2 + \|u_1 - u_2\|^2$$

we get

$$\|u - u_2\|^2 \leq \left(1 - \frac{1}{Cm_1(u - u_1, \mathcal{T})^2}\right) \|u - u_1\|^2 + \sum_{K \in \mathcal{T}} \alpha_K^2 \|r - r_K\|_{L_2(K)}^2.$$

Taking the square root we finish the proof. \square

The estimate (6.61) we call the error reduction property on anisotropic meshes. As it could be mentioned the constant in (6.61) depends strongly on the value of the matching function m_1 , and only bounding m_1 one can claim that the error reduces its value significantly while using refined finite element space.

6.4 Hierarchical error estimator

Let $V_2 = V_1 \oplus \tilde{V}_2$. The true error e satisfies the following variational formulation:

$$B(e, v) = (f, v) - B(u_1, v) \quad \forall v \in V.$$

Now let us try to reduce the space V to the space V_2 , namely consider e_2 satisfying

$$B(e_2, v) = (f, v) - B(u_1, v) \quad \forall v \in V_2.$$

It is clear that $e_2 = u_2 - u_1$.

Now by means of the error reduction property we prove that the approximation of the error e_2 is equivalent in the energy norm to the true error e .

Theorem 6.26. *There exists a constant C that depends neither on the geometry of the mesh \mathcal{T} nor on the perturbation parameters, such that*

$$\|e_2\| \leq \|e\| \preceq m_1(e, \mathcal{T}) \left(\|e_2\|^2 + \sum_{K \in \mathcal{T}} \alpha_K^2 \|r - r_K\|_{L_2(K)}^2 \right)^{1/2}.$$

Proof. Let us verify the first inequality:

$$e = u - u_1 = u - u_2 + e_2,$$

and due to orthogonal property we get

$$\|e\|^2 = \|u - u_2\|^2 + \|e_2\|^2,$$

which leads to the first inequality $\|e\| \geq \|e_2\|$.

The second inequality is nothing else, but Lemma 6.24. \square

Represent error in the form $e_2 = e_{21} + e_{22}$, where $e_{21} \in V_1$, $e_{22} \in \tilde{V}_2$, where e_{21} , e_{22} satisfy

$$\begin{cases} B(e_{21}, v_1) + B(e_{22}, v_1) = (f, v_1) - B(u_1, v_1) & \forall v_1 \in V_1 \\ B(e_{21}, v_2) + B(e_{22}, v_2) = (f, v_2) - B(u_1, v_2) & \forall v_2 \in \tilde{V}_2, \end{cases}$$

or,

$$\begin{cases} B(e_{21}, v_1) + B(e_{22}, v_1) = 0 & \forall v_1 \in V_1 \\ B(e_{21}, v_2) + B(e_{22}, v_2) = (f, v_2) - B(u_1, v_2) & \forall v_2 \in \tilde{V}_2. \end{cases}$$

Ignoring the coupling terms we get

$$\begin{cases} B(\bar{e}_{21}, v_1) = 0 & \forall v_1 \in V_1 \\ B(\bar{e}_{22}, v_2) = (f, v_2) - B(u_1, v_2) & \forall v_2 \in \tilde{V}_2. \end{cases}$$

From the first equation we immediately get $\bar{e}_{21} = 0$. Denote $\bar{e} := \bar{e}_{22}$. For \bar{e} we have the following equation:

$$B(\bar{e}, v_2) = (f, v_2) - B(u_1, v_2) \quad \forall v_2 \in \tilde{V}_2.$$

It is useful to know that

$$\begin{cases} B(e_{21}, v_1) + B(e_{22}, v_1) = 0 & \forall v_1 \in V_1 \\ B(e_{21}, v_2) + B(e_{22}, v_2) = B(\bar{e}, v_2) & \forall v_2 \in \tilde{V}_2, \end{cases}$$

or,

$$\begin{cases} B(e_2, v_1) = 0 & \forall v_1 \in V_1 \\ B(e_2, v_2) = B(\bar{e}, v_2) & \forall v_2 \in \tilde{V}_2. \end{cases} \quad (6.62)$$

Now by means of strengthened Cauchy-Schwarz inequality we prove that the approximation of the error \bar{e} is equivalent in the energy norm to the approximation e_2 .

Theorem 6.27. *The approximation of the error \bar{e} stemming from (6.62) satisfies*

$$\|\bar{e}\| \leq \|e_2\| \leq \frac{1}{\sqrt{1-\gamma^2}} \|\bar{e}\|,$$

where $\gamma = \sqrt{\frac{31927}{35680} + \frac{7\sqrt{193953}}{35680}}$ is the constant from Corollary 3.7.

Proof. We have

$$\begin{aligned} \|e_2\|^2 &= B(e_{21} + e_{22}, e_{21} + e_{22}) = \|e_{21}\|^2 + 2B(e_{21}, e_{22}) + \|e_{22}\|^2 \\ &\geq \|e_{21}\|^2 - 2\gamma \|e_{21}\| \|e_{22}\| + \|e_{22}\|^2. \end{aligned}$$

Utilizing the inequality

$$2\gamma \|e_{21}\| \|e_{22}\| \leq \|e_{21}\|^2 + \gamma \|e_{22}\|^2,$$

we get

$$\|e_2\|^2 \geq (1 - \gamma^2) \|e_{22}\|^2.$$

Applying the first inequality of (6.62) we get

$$\|e_2\|^2 = B(e_2, e_{21}) + B(e_2, e_{22}) = B(e_2, e_{22}).$$

Applying the second inequality of (6.62) we get

$$\|e_2\|^2 = (\bar{e}, e_{22}) \leq \|\bar{e}\| \|e_{22}\| \leq \frac{1}{\sqrt{1-\gamma^2}} \|\bar{e}\| \|e_2\|.$$

So we get

$$\|e_2\| \leq \frac{1}{\sqrt{1-\gamma^2}} \|\bar{e}\|.$$

Second inequality of the theorem is shown as follows:

$$\|\bar{e}\|^2 = B(\bar{e}, \bar{e}) = B(e_2, \bar{e}) \leq \|\bar{e}\| \|e_2\|,$$

and thus,

$$\|\bar{e}\| \leq \|e_2\|,$$

□

Definition 6.28 (Error estimator). For all triangles K and edges γ define the following terms

$$a_\gamma := \frac{B(u_1, b_{\gamma, \delta_\gamma}) - \int_\Omega f b_{\gamma, \delta_\gamma} dx}{\|b_{\gamma, \delta_\gamma}\|^2} = -\frac{\int_\gamma R b_{\gamma, \delta_\gamma} ds + \int_{\omega_\gamma} r b_{\gamma, \delta_\gamma} dx}{\|b_{\gamma, \delta_\gamma}\|^2},$$

$$c_K := \frac{B(u_1, b_K) - \int_\Omega f b_K dx}{\|b_K\|^2} = -\frac{\int_K r b_K dx}{\|b_K\|^2}.$$

By means of these terms we define approximation function to the error:

$$\tilde{e} := \sum_{\gamma \in \partial \mathcal{T}} a_\gamma b_{\gamma, \delta_\gamma} + \sum_{K \in \mathcal{T}} c_K b_K,$$

$\|\tilde{e}\|$ is then the hierarchical a posteriori error estimator.

Let $v, w \in \tilde{V}_2$.

$$v = \sum_{\gamma \in \partial \mathcal{T}} v_\gamma b_{\gamma, \delta_\gamma} + \sum_{K \in \mathcal{T}} v_K b_K$$

$$w = \sum_{\gamma \in \partial \mathcal{T}} w_\gamma b_{\gamma, \delta_\gamma} + \sum_{K \in \mathcal{T}} w_K b_K.$$

Define a bilinear form $d(\cdot, \cdot) : \tilde{V}_2^2 \rightarrow \mathbb{R}$ as follows:

$$d(v, w) := \sum_{\gamma \in \partial \mathcal{T}} v_\gamma w_\gamma \|b_{\gamma, \delta_\gamma}\|^2 + \sum_{K \in \mathcal{T}} v_K w_K \|b_K\|^2$$

We need also a local analogue of this bilinear form for any triangle K :

$$d_K(v, w) := \sum_{\gamma \subset \partial K} v_\gamma w_\gamma \|b_{\gamma, \delta_\gamma}\|_K^2 + v_K w_K \|b_K\|_K^2$$

This bilinear form has the following properties:

$$d(v, w) = \sum_{K \in \mathcal{T}} d_K(v, w) \tag{6.63}$$

$$d(\tilde{e}, v) = -B(\tilde{e}, v) \quad \forall v \in \tilde{V}_2. \tag{6.64}$$

The first relation is clear, let us prove the second one. Indeed,

$$\begin{aligned} d(\tilde{e}, v) &= \sum_{\gamma \in \partial \mathcal{T}} \left[B(u_1, b_{\gamma, \delta_\gamma}) - \int_{\omega_\gamma} f b_{\gamma, \delta_\gamma} dx \right] v_\gamma + \sum_{K \in \mathcal{T}} \left[B(u_1, b_K) - \int_K f b_K dx \right] v_K \\ &= B(u_1, v) - \int_\Omega f v dx = -B(\tilde{e}, v). \end{aligned}$$

In subsequent analysis we need a kind of stability property for the bilinear form $d(\cdot, \cdot)$ which we formulate in the following lemma.

Lemma 6.29. There exists a constant C that depends neither on the geometry of the mesh \mathcal{T} nor on the perturbation parameters, such that

$$d(v, v) \leq \|v\|^2 \quad \forall v \in \tilde{V}_2. \tag{6.65}$$

Proof. First we prove the claimed result locally, namely that for any triangle K the following inequality holds:

$$d_K(v, v) \preceq \|v\|_K^2 \quad \forall v \in \tilde{V}_2,$$

from which the inequality (6.65) evidently follows. Let $v \in \tilde{V}_2$, then $v|_K$ can be represented in the following way:

$$v|_K = \sum_{\gamma \subset \partial K} v_\gamma b_{\gamma, \delta_\gamma} + v_K b_K.$$

We have

$$\begin{aligned} \|v\|_K^2 &= B_K \left(\sum_{\gamma \subset \partial K} v_\gamma b_{\gamma, \delta_\gamma} + v_K b_K, \sum_{\gamma \subset \partial K} v_\gamma b_{\gamma, \delta_\gamma} + v_K b_K \right) \\ &= \sum_{\gamma \subset \partial K} v_\gamma^2 \|b_{\gamma, \delta_\gamma}\|_K^2 + v_K^2 \|b_K\|_K^2 + 2B_K \left(v_K b_K, \sum_{\gamma \subset \partial K} v_\gamma b_{\gamma, \delta_\gamma} \right) \\ &\geq \sum_{\gamma \subset \partial K} v_\gamma^2 \|b_{\gamma, \delta_\gamma}\|_K^2 + v_K^2 \|b_K\|_K^2 - 2\gamma \|v_K b_K\|_K \left\| \sum_{\gamma \subset \partial K} v_\gamma b_{\gamma, \delta_\gamma} \right\|_K \\ &\geq (1 - \gamma) \left(\sum_{\gamma \subset \partial K} v_\gamma^2 \|b_{\gamma, \delta_\gamma}\|_K^2 + v_K^2 \|b_K\|_K^2 \right) = (1 - \gamma) d_K(v, v), \end{aligned}$$

where γ is the strengthened Cauchy-Schwarz constant from Theorem 3.8. Dividing both sides by $1 - \gamma$ we get the result claimed. \square

We need also the estimates from above for interior and edge residuals. The following lemma is taken from [27].

Lemma 6.30 (Interior residual). *Let $K \in \mathcal{T}$. Then*

$$\|r\|_{L_2(K)} \preceq \varepsilon h_{\min, K}^{-1} \|e\|_K + \|r - \bar{r}\|_{L_2(K)}$$

Proof. For the proof see Grosman [27]. \square

The following lemma is an improved version of the one from [27].

Lemma 6.31 (Face residual). *Let γ be any interior interface. Then,*

$$\|R\|_{L_2(\gamma)} \preceq \sum_{K' \in \omega_\gamma} \left\{ \varepsilon^{1/2} \alpha_{K'}^{-1/2} \|e\|_{K'} + \left(\frac{|K'|}{|\gamma|} \right)^{1/2} \delta_\gamma^{1/2} \|r - \bar{r}\|_{L_2(K')} \right\}.$$

Proof. Let $v \in H_0^1(\Omega)$. Integrating by parts on each element yields

$$B(e, v) = \sum_{K \in \mathcal{T}} \int_K r v \, dx - \sum_{\gamma \in \partial \mathcal{T}} \int_\gamma R v \, ds, \quad (6.66)$$

where $\partial \mathcal{T}$ denotes the collection of interelement faces.

Let $\gamma \in \partial \mathcal{T}$. Suppose that $\gamma = \bar{K}_1 \cap \bar{K}_2$. Then $\omega_\gamma = \text{int}(\bar{K}_1 \cup \bar{K}_2)$. Choosing $v := F_{\text{ext}}(R) b_{\gamma, \delta_\gamma} \in H_0^1(\Omega)$ in (6.66) implies

$$\int_\gamma b_{\gamma, \delta_\gamma} R^2 \, ds = \sum_{K \subset \omega_\gamma} \int_K r F_{\text{ext}}(R) b_{\gamma, \delta_\gamma} \, dx - B_{\omega_\gamma}(e, F_{\text{ext}}(R) b_{\gamma, \delta_\gamma}).$$

Furthermore, applying the Cauchy-Schwarz inequality, one obtains

$$|B_K(e, F_{ext}(R)b_{\gamma, \delta_\gamma})| \leq \|e\|_K \|F_{ext}(R)b_{\gamma, \delta_\gamma}\|_K.$$

Using (6.56) and (6.57) one estimates the second factor as follows:

$$\begin{aligned} \|F_{ext}(R)b_{\gamma, \delta_\gamma}\|_K^2 &= \varepsilon^2 \|\nabla(F_{ext}(R)b_{\gamma, \delta_\gamma})\|_{L_2(K)}^2 + \kappa^2 \|F_{ext}(R)b_{\gamma, \delta_\gamma}\|_{L_2(K)}^2 \\ &\leq \left(\varepsilon^2 \min \left\{ \delta_\gamma \frac{|K|}{|\gamma|}, h_{min, K} \right\}^{-2} \delta_\gamma \frac{|K|}{|\gamma|} + \kappa^2 \delta_\gamma \frac{|K|}{|\gamma|} \right) \|R\|_{L_2(\gamma)}^2. \end{aligned}$$

Thus, we have

$$\begin{aligned} |B_K(e, F_{ext}(R)b_{\gamma, \delta_\gamma})| &\leq \left(\frac{|K|}{|\gamma|} \right)^{1/2} \|e\|_K \|R\|_{L_2(\gamma)} \\ &\quad * \left(\varepsilon \min \left\{ \delta_\gamma \frac{|K|}{|\gamma|}, h_{min, K} \right\}^{-1} \delta_\gamma^{1/2} + \kappa \delta_\gamma^{1/2} \right). \end{aligned}$$

Applying the Cauchy-Schwarz inequality, Lemma 6.30 and (6.56) to the second term we have

$$\begin{aligned} \left| \int_K r F_{ext}(R)b_{\gamma, \delta_\gamma} dx \right| &\leq \|r\|_{L_2(K)} \|F_{ext}(R)b_{\gamma, \delta_\gamma}\|_{L_2(K)} \\ &\leq \left[\left(\varepsilon h_{min, K}^{-1} + \kappa \right) \|e\|_K + \|r - \bar{r}\|_{L_2(K)} \right] \delta_\gamma^{1/2} \left(\frac{|K|}{|\gamma|} \right)^{1/2} \|R\|_{L_2(\gamma)}. \end{aligned}$$

Combining two previous estimates we get

$$\begin{aligned} \|R\|_{L_2(\gamma)} &\leq \sum_{K' \in \omega_\gamma} \left\{ \left(\frac{|K'|}{|\gamma|} \right)^{1/2} \delta_\gamma^{1/2} \|r - \bar{r}\|_{L_2(K')} + \left(\frac{|K'|}{|\gamma|} \right)^{1/2} \|e\|_{K'} \right. \\ &\quad \left. * \left(\varepsilon \min \left\{ \delta_\gamma \frac{|K|}{|\gamma|}, h_{min, K} \right\}^{-1} \delta_\gamma^{1/2} + h_{min, K}^{-1} \delta_\gamma^{1/2} + \kappa \delta_\gamma^{1/2} \right) \right\}. \end{aligned}$$

By simple manipulations we get (with δ_γ from (6.58)),

$$\varepsilon \min \left\{ \delta_\gamma \frac{|K|}{|\gamma|}, h_{min, K} \right\}^{-1} \delta_\gamma^{1/2} + h_{min, K}^{-1} \delta_\gamma^{1/2} + \kappa \delta_\gamma^{1/2} \leq 4\delta_\gamma^{1/2} \alpha_K^{-1},$$

which finishes the proof. \square

We are now in a position to formulate a main result of the present paper, namely the robustness of the error estimator.

Theorem 6.32. *In foregoing notation and assumptions the following inequalities hold*

$$\begin{aligned} \|e\| &\leq m_1(e, \mathcal{T}) \left(\|\tilde{e}\|^2 + \sum_{K \in \mathcal{T}} \alpha_K^2 \|r - r_K\|_{L_2(K)}^2 \right)^{1/2}, \\ \|\tilde{e}\|_K &\leq \left(\|e\|_K^2 + \sum_{K \in \tilde{K}} \alpha_K^2 \|r - r_K\|_{L_2(K)}^2 \right)^{1/2}, \end{aligned}$$

where \tilde{K} is a unit of four triangles including K itself and three triangles sharing with K its edges.

Proof. 1. First inequality.

$$\begin{aligned} \|e\| &\preceq m_1(e, \mathcal{T}) \left(\|e_2\|^2 + \sum_{K \in \mathcal{T}} \alpha_K^2 \|r - r_K\|_{L_2(K)}^2 \right)^{1/2} \\ &\preceq m_1(e, \mathcal{T}) \left(\frac{1}{\sqrt{1-\gamma^2}} \|\bar{e}\|^2 + \sum_{K \in \mathcal{T}} \alpha_K^2 \|r - r_K\|_{L_2(K)}^2 \right)^{1/2} \end{aligned}$$

It remains to show that $\|\bar{e}\| \leq \|\tilde{e}\|$. Indeed,

$$\|\bar{e}\|^2 = d(\tilde{e}, \bar{e}) \leq \sqrt{d(\tilde{e}, \bar{e})} \sqrt{d(\bar{e}, \bar{e})} \preceq \|\tilde{e}\| \|\bar{e}\|$$

2. Second inequality is obtained more or less straight forward as it is shown below.

$$\|\tilde{e}\|_K \leq \sum_{\gamma \subset \partial K} |a_\gamma| \|b_{\gamma, \delta_\gamma}\|_K + |c_K| \|b_K\|_K$$

Estimate first and second term subsequently. Utilize (6.56), (6.57), Lemma 6.30 and Lemma 6.31 subsequently.

$$\begin{aligned} |a_\gamma| \|b_{\gamma, \delta_\gamma}\|_K &\leq \frac{\|R\|_{L_2(\gamma)} \|b_{\gamma, \delta_\gamma}\|_{L_2(\gamma)} + \|r\|_{L_2(\omega_\gamma)} \|b_{\gamma, \delta_\gamma}\|_{L_2(\omega_\gamma)}}{\|b_{\gamma, \delta_\gamma}\|_K} \\ &\preceq \frac{\|R\|_{L_2(\gamma)}}{\kappa \alpha_\gamma^{1/2} + \varepsilon \alpha_\gamma^{-1/2}} + \alpha_\gamma \|r\|_{L_2(\omega_\gamma)} \\ &\preceq \varepsilon^{-1/2} \alpha_\gamma^{1/2} \|R\|_{L_2(\gamma)} + \alpha_\gamma \|r\|_{L_2(\omega_\gamma)} \\ &\preceq \|e\|_{\omega_\gamma} + \sum_{K \in \omega_\gamma} \alpha_K \|r - r_K\|_{L_2(K)}. \end{aligned}$$

Similarly estimating the second term we get the result claimed.

$$\begin{aligned} |c_K| \|b_K\|_K &\leq \frac{\|r\|_{L_2(K)} \|b_K\|_{L_2(K)}}{\|b_K\|_K} \\ &\preceq \alpha_K \|r\|_{L_2(K)} \\ &\preceq \|e\|_K + \alpha_K \|r - r_K\|_{L_2(K)}. \end{aligned}$$

□

Combining two previous estimates we finish the proof.

6.5 Numerical experiments

Let us consider the 2D model problem

$$-\Delta u + \kappa^2 u = 0 \text{ in } \Omega := [0, 1]^2, \quad u = u_0 \text{ on } \partial\Omega.$$

Prescribe the exact solution

$$u = e^{-\kappa x} + e^{-\kappa y}$$

which displays typical boundary layers along the sides $x = 0$ and $y = 0$. The Dirichlet boundary data u_0 are chosen accordingly.

Mesh	Elements	$\frac{\ e\ }{\ u-u_1\ }$	$\frac{\ u-u_2\ }{\ u-u_1\ }$	$m_1(u-u_1, \mathcal{T})$	$\frac{\ u-u_1\ }{\ u\ }$
1	8	0.80	0.60	1.42	1.92
2	32	0.73	0.64	1.41	1.14
3	128	0.64	0.69	1.42	0.65
4	512	0.56	0.65	1.42	0.34
5	2048	0.50	0.55	1.42	0.17
6	8192	0.50	0.55	1.42	0.09
7	32768	0.51	0.54	1.42	0.04
8	131072	0.51	0.54	1.43	0.02

Table 3: Results for $\kappa = 1000$ with transition parameter $\tau = 2 \ln(\kappa)/\kappa$.

γ	Mesh	Elements	$\frac{\ e\ }{\ u-u_1\ }$	$\frac{\ u-u_2\ }{\ u-u_1\ }$	$m_1(u-u_1, \mathcal{T})$	$\frac{\ u-u_1\ }{\ u\ }$
1	4	512	0.56	0.65	1.42	0.34
0.1	4	512	0.41	0.91	16.6	0.45
0.01	4	512	0.40	0.97	24.8	2.56
10	4	512	0.72	0.65	1.44	2.05

Table 4: Results for $\kappa = 1000$ with various transition parameters $\tau = 2\gamma \ln(\kappa)/\kappa$.

In the first table we use a sequence of Shishkin meshes with transition parameter $\tau = 2 \ln(\kappa)/\kappa$.

The second table has the results for the various transition parameters $\tau = 2\gamma \ln(\kappa)/\kappa$, perturbed from the original value by additional factor γ . $\gamma \sim 1$ corresponds to the appropriate transition parameter τ . $\gamma \gg 1$ means that the mesh is unnecessarily course, while $\gamma \ll 1$ produces an overrefinement leading to the large values of the matching function m_1 . We can observe the influence of the matching function to the error estimator and the constant in the error reduction property. It could be argued that moderate values of the matching function yield the constant in the error reduction property actually smaller than 1, while the large values destroy the error reduction for enlarged finite element space. Thus it demonstrates that the estimate (6.61) is sharp.

7 Towards a convergent algorithm for the reaction-diffusion problem

In this chapter we provide the adaptive algorithm that allows anisotropic triangulations works in addition for the reaction-diffusion problem. The error reduces in each adaptive step, but the convergence property does not seem to be possible to be proven for this algorithm because of additional data oscillation terms. Numerical experiments in Section 7.4 confirm the theory for the adaptive algorithm. The adaptive algorithm shows its potential by creating the anisotropic mesh for the problem with the boundary layer starting with a very coarse isotropic mesh.

7.1 Alternative residual a posteriori error estimation

As it was already mentioned in Section 5.2 a weight of the edge error indicator is not uniquely defined in general. In this section we will formulate the residual a posteriori error estimator obtained in [31].

Consider the model problem (2.1). Recall the notation for the interior and boundary residuals r and R

$$r = f + \varepsilon^2 \Delta u_h - \kappa^2 u_h$$

and

$$R_\gamma = \begin{cases} \varepsilon^2 \left[\frac{\partial u_h}{\partial n} \right] & \text{on } \partial K \cap \partial K' \\ 0 & \text{on } \partial K \cap \partial \Omega. \end{cases}$$

The mean value is defined via $r_K := \frac{1}{|K|} \int_K r \, dx$ as before. In addition we recall the following notation:

$$\begin{aligned} \alpha_K &:= \min(\varepsilon^{-1} h_{\min, K}, \kappa^{-1}), \\ \alpha_\gamma &:= \min(\varepsilon^{-1} h_{\min, \gamma}, \kappa^{-1}). \end{aligned}$$

Define the error indicator η_γ associated with an edge γ by

$$\eta_\gamma := \varepsilon^{-1/2} \alpha_\gamma^{1/2} h_{\min, \gamma}^{1/2} \left(\frac{|\gamma|}{|\omega_\gamma|} \right)^{1/2} \|R_\gamma\|_{L_2(\gamma)},$$

and the error indicator η_K associated with an element K by

$$\eta_K := \alpha_K \|r_K\|_{L_2(K)}.$$

Then the local residual error estimator $\eta_{R,K}$ and the local data oscillation (local approximation term) ζ_K for the element K are defined via

$$\begin{aligned} \eta_{R,K}^2 &:= \eta_K^2 + \frac{1}{2} \sum_{\gamma \subset \partial K} \eta_\gamma^2, \\ \zeta_K^2 &:= \alpha_K^2 \|r - r_K\|_{L_2(K)}^2. \end{aligned}$$

Let η_R and ζ be their global counter parts, given by

$$\begin{aligned} \eta_R^2 &:= \sum_{K \in \mathcal{T}_h} \eta_{R,K}^2, \\ \zeta^2 &:= \sum_{K \in \mathcal{T}_h} \zeta_K^2. \end{aligned}$$

The following robustness results was shown in [31] for $\kappa = 1$. The proof is identical for general $\kappa > 0$.

Theorem 7.1 (Global upper error bound). *There exists a constant C that does not depend on the aspect ratio of the mesh, such that*

$$\| \|u - u_h\| \|_{\Omega}^2 \leq m_1^2(u - u_h, \mathcal{T}_h) (\eta_R^2 + \zeta^2).$$

Theorem 7.2 (Local lower error bound). *There exists a constant C that does not depend on the aspect ratio of the mesh, such that*

$$\eta_{R,K}^2 \leq \| \|u - u_h\| \|_{\omega_K}^2 + \sum_{K \subset \omega_K} \zeta_K^2.$$

7.2 Marking strategy

The marking and refinement strategies are the simplified variants of the corresponding strategies for the case of the pure Poisson problem.

Marking Strategy $\tilde{\tilde{E}}$

Given a parameter $0 < \theta < 1$ and a mesh \mathcal{T}_H , select the set of elements $\hat{\mathcal{T}}_H \subset \mathcal{T}_H$ and the set of edges $\hat{\mathcal{E}}_H \subset \mathcal{E}_H$ for which

$$\left(\sum_{K \in \hat{\mathcal{T}}_H} \eta_K^2 + \sum_{\gamma \in \hat{\mathcal{E}}_H} \eta_{\gamma}^2 \right)^{1/2} \geq \theta \eta_R.$$

In order to formulate an error reduction theorem we need an additional condition on the marking procedure :

$$\forall \gamma \in \hat{\mathcal{E}}_h \text{ with } h_{\min, \gamma} \geq \varepsilon \kappa^{-1} \text{ holds: } \forall K \subset \omega_{\gamma} : K \in \hat{\mathcal{T}}_H. \quad (7.1)$$

This condition makes sense since the anisotropic refinement is usually appropriate only inside the boundary/interior layer where $h_{\min, \gamma} \ll \varepsilon \kappa^{-1}$.

Theorem 7.3. *Let \mathcal{T}_H be a triangulation of Ω , and let \mathcal{T}_h be the triangulation obtained from \mathcal{T}_H by refining every element marked according to Strategy $\tilde{\tilde{E}}$ in such a way that new node is created in the interior of each marked triangle $K \in \hat{\mathcal{T}}_H$ and on the each marked edge $\gamma \in \hat{\mathcal{E}}_H$. If condition (7.1) is satisfied, there exist a constant $0 < \alpha < 1$, depending on θ and the alignment measure, such that*

$$\| \|u - u_h\| \|_{\Omega}^2 \leq \alpha^2 \| \|u - u_H\| \|_{\Omega}^2 + C\zeta^2 + C \sum_{K \in \mathcal{T}_H} \alpha_K^2 \|f - f_K\|_{L_2(K)}^2.$$

The dependence on the alignment measure has the following form:

$$\alpha = \left(1 - \frac{1}{Cm_1^2(u - u_H, \mathcal{T}_H)} \right)^{1/2}$$

The proof will be given in Section 7.3. Before, we will state the algorithm $\tilde{\tilde{C}}$.

Remark 7.4. It is not possible to show the asymptotic convergence of the adaptive algorithm for the reaction-diffusion equation due to the additional higher order terms. It would be possible if the algorithm guaranteed that three new nodes are introduced per each marked triangle each iteration step, cf. [44]. However, an algorithm with this property that additionally allows anisotropic refinement is unknown for the author. That is why in this chapter we do not talk about the convergence but only restrict ourselves to the error reduction.

Algorithm $\tilde{\mathbb{C}}$

Choose parameters $0 < \theta, \hat{\theta} < 1$.

1. Take some initial mesh \mathcal{T}_0 .
2. Solve the discrete problem on \mathcal{T}_0 , denote by u_0 its solution.
3. Let $k = 0$.
4. Compute the local indicators η_K .
5. If the global estimated error is small then STOP.
6. Construct $\hat{\mathcal{T}}_k \in \mathcal{T}_k$ by Marking Strategy $\tilde{\mathbb{E}}$ and parameter θ .
7. Let \mathcal{T}_{k+1} be a refinement of \mathcal{T}_k such that
 - each element of $\hat{\mathcal{T}}_k$ contains a node of \mathcal{T}_{k+1} in its interior and
 - each edge of $\hat{\mathcal{E}}_k$, contains a node of \mathcal{T}_{k+1} in its interior and
 - Condition (7.1) holds.
8. Solve the discrete problem on \mathcal{T}_{k+1} , denote by u_{k+1} its solution.
9. Let $k := k + 1$ and go to Step 4.

7.3 Error reduction theorem

This section is aimed to prove Theorem 7.3. To this end we follow Stevenson [44] and formulate a lemma that will allow us to bound an edge residual for element K on which $h_{\min, K} \succeq \varepsilon \kappa^{-1}$.

Lemma 7.5. *For any two triangles K_1, K_2 that share an edge γ , for any continuous piecewise linear function u_h over $K_1 \cup K_2$, and for any piecewise constant function f over $K_1 \cup K_2$ it holds*

$$\left\| \left[\frac{\partial u_h}{\partial n} \right]_{\gamma} \right\|_{L_2(\gamma)} \preceq \left(\frac{|\omega_{\gamma}|}{|\gamma|} \right)^{-3/2} \|f - u_h\|_{L_2(K_1 \cup K_2)}.$$

Proof. Define the continuous piecewise linear function w by $w(A_1) = \frac{|2K_1|}{|\gamma|}$ for the vertex

A_1 of K_1 not on γ , and $w = 0$ on K_2 and note that $\left[\frac{\partial w}{\partial n} \right] = 1$. Since any continuous, piecewise linear function u can be rewritten as a linear combination of w and a linear over $K_1 \cup K_2$ function p thus having a zero jump in the normal derivative over γ , it suffices to prove that

$$\left\| \left[\frac{\partial w}{\partial n} \right]_{\gamma} \right\|_{L_2(\gamma)} \preceq \left(\frac{|\omega_{\gamma}|}{|\gamma|} \right)^{-3/2} \inf_{p \in P_1(K_1 \cup K_2), f \in \prod_{i=1}^2 P_0(K_i)} \|f - (w + p)\|_{L_2(K_1 \cup K_2)}.$$

Since

$$\left\| \left[\frac{\partial w}{\partial n} \right]_{\gamma} \right\|_{L_2(\gamma)} = |\gamma|^{1/2} \sim \left(\frac{|\omega_{\gamma}|}{|\gamma|} \right)^{-3/2} \|w\|_{L_2(K_1 \cup K_2)},$$

where $V = P_1(K_1 \cup K_2) + \prod_{i=1}^2 P_0(K_i)$, it suffices to show that

$$\|w\|_{L_2(K_1 \cup K_2)} \preceq \inf_{v \in V} \|w - v\|_{L_2(K_1 \cup K_2)}.$$

We can note that the last infimum is reached for the L_2 -projection of w onto the space V , in other words v should satisfy the relation:

$$(v - w, \phi)_{L_2(K_1 \cup K_2)} = 0, \quad \forall \phi \in V.$$

Due to Lemma 3.9 there exists a constant $\tilde{C} < 1$ such that:

$$\begin{aligned} \|w - v\|_{L_2(K_1 \cup K_2)}^2 &= \|w\|_{L_2(K_1 \cup K_2)}^2 + \|v\|_{L_2(K_1 \cup K_2)}^2 - 2(v, w)_{L_2(K_1 \cup K_2)} \\ &\geq \|w\|_{L_2(K_1 \cup K_2)}^2 + \|v\|_{L_2(K_1 \cup K_2)}^2 - 2\tilde{C}\|w\|_{L_2(K_1 \cup K_2)}\|v\|_{L_2(K_1 \cup K_2)} \\ &\geq (1 - \tilde{C})(\|w\|_{L_2(K_1 \cup K_2)}^2 + \|v\|_{L_2(K_1 \cup K_2)}^2) \geq (1 - \tilde{C})\|w\|_{L_2(K_1 \cup K_2)}^2. \end{aligned}$$

which finishes the proof. \square

We conclude results for the edge error indicator separately for large ($h_{\min, \gamma} \succeq \varepsilon \kappa^{-1}$) and small ($h_{\min, \gamma} \ll \varepsilon \kappa^{-1}$) elements.

Lemma 7.6. *For any edge $\gamma \in \partial \mathcal{T}_h$ with $h_{\min, \gamma} \succeq \varepsilon \kappa^{-1}$, we have*

$$\eta_{\gamma}^2 \preceq \eta_{K_1}^2 + \eta_{K_2}^2 + \zeta_{K_1}^2 + \zeta_{K_2}^2 + \alpha_{K_1}^2 \|f - f_{K_1}\|_{L_2(K_1)}^2 + \alpha_{K_2}^2 \|f - f_{K_2}\|_{L_2(K_2)}^2.$$

Proof. By $h_{\min, \gamma} \succeq \varepsilon \kappa^{-1}$ the utilization of Lemma 7.5 shows that

$$\begin{aligned} \eta_{\gamma}^2 &= \varepsilon^{-1} \alpha_{\gamma} h_{\min, \gamma} \frac{|\gamma|}{|\omega_{\gamma}|} \|R_{\gamma}\|_{L_2(\gamma)}^2 = \varepsilon^3 \alpha_{\gamma} h_{\min, \gamma} \frac{|\gamma|}{|\omega_{\gamma}|} \left\| \left[\frac{\partial u_h}{\partial n} \right]_{\gamma} \right\|_{L_2(\gamma)}^2 \\ &\preceq \varepsilon^3 \kappa^{-1} \left\| \left[\frac{\partial u_h}{\partial n} \right]_{\gamma} \right\|_{L_2(\gamma)}^2 \\ &\preceq \sum_{i=1}^2 \varepsilon^3 \kappa^{-1} \left(\frac{|\omega_{\gamma}|}{|\gamma|} \right)^{-3} \left\| \frac{1}{\kappa^2} f_{K_i} - u_h \right\|_{L_2(K_i)}^2 \\ &\preceq \sum_{i=1}^2 (\varepsilon h_{\min, \gamma}^{-1})^3 \kappa^{-1} \left\| \frac{1}{\kappa^2} f_{K_i} - u_h \right\|_{L_2(K_i)}^2 \\ &\preceq \sum_{i=1}^2 \kappa^2 \left\| \frac{1}{\kappa^2} f_{K_i} - u_h \right\|_{L_2(K_i)}^2 \\ &\preceq \eta_{K_1}^2 + \eta_{K_2}^2 + \zeta_{K_1}^2 + \zeta_{K_2}^2 + \alpha_{K_1}^2 \|f - f_{K_1}\|_{L_2(K_1)}^2 + \alpha_{K_2}^2 \|f - f_{K_2}\|_{L_2(K_2)}^2. \end{aligned}$$

\square

Lemma 7.7. *Let $K \in \mathcal{T}_H$ be any element and $\gamma \in \mathcal{E}_H$ be any edge of the triangulation. In the notation of Theorem 7.3 the following estimate holds:*

$$\eta_K^2 \preceq \|u_h - u_H\|_K^2 + \zeta_K^2, \quad (7.2)$$

if in addition $h_{\min,\gamma} \ll \varepsilon \kappa^{-1}$, then

$$\eta_\gamma^2 \leq \sum_{K \subset \omega_\gamma} \left(C_1 \|u_h - u_H\|_K^2 + \frac{1}{6} \eta_K^2 + C_1 \zeta_K^2 \right), \quad (7.3)$$

provided there exist the basis functions ψ_K and ψ_γ on the new mesh \mathcal{T}_h . We denote by ψ_K and ψ_γ the hat functions corresponding to the new nodes introduced in the interior of the triangle K and on the edge γ , respectively.

Proof. We show (7.2) first. Using the fact that $\int_K \nabla u_H \nabla \psi_K dx = -\int_K \Delta u_H \psi_K dx = 0$, we have

$$\begin{aligned} r_K \frac{|K|}{3} &= \int_K r_K \psi_K dx \\ &= \int_K f \psi_K dx - \kappa^2 \int_K u_H \psi_K dx - \int_K r \psi_K dx + \int_K r_K \psi_K dx \\ &= B_K(u_h, \psi_K) - B_K(u_H, \psi_K) - \int_K (r - r_K) \psi_K dx \\ &= B_K(u_h - u_H, \psi_K) - \int_K (r - r_K) \psi_K dx. \end{aligned}$$

Squaring and using the Cauchy-Schwarz inequality we get:

$$|K| \|r_K\|_{L_2(K)}^2 \preceq \|u_h - u_H\|_K^2 \|\psi_K\|_K^2 + \|r - r_K\|_{L_2(K)}^2 \|\psi_K\|_{L_2(K)}^2$$

The desired inequality (7.2) follows now from the estimates (2.12) and (2.13).

We show now the second inequality (7.3). Taking into account that R_γ is a constant over any edge γ , we have:

$$\begin{aligned} \frac{1}{2} |\gamma| R_\gamma &= \int_\gamma R_\gamma \psi_\gamma ds = -\varepsilon^2 \int_{\omega_\gamma} \nabla u_H \nabla \psi_\gamma dx \\ &= \varepsilon^2 \int_{\omega_\gamma} \nabla (u_h - u_H) \nabla \psi_\gamma dx + \kappa^2 \int_{\omega_\gamma} (u_h - u_H) \psi_\gamma dx - \int_{\omega_\gamma} f \psi_\gamma dx + \kappa^2 \int_{\omega_\gamma} u_H \psi_\gamma dx \\ &= B_{\omega_\gamma}(u_h - u_H, \psi_\gamma) - \int_{\omega_\gamma} r \psi_\gamma dx. \end{aligned}$$

Observing that $\frac{1}{2} |\gamma| R_\gamma = \frac{1}{2} |\gamma|^{1/2} \|R_\gamma\|_{L_2(\gamma)}$ we estimate the edge error indicator:

$$\begin{aligned} \eta_\gamma^2 &= \varepsilon^{-1} \alpha_\gamma h_{\min,\gamma} \frac{|\gamma|}{|\omega_\gamma|} \|R_\gamma\|_{L_2(\gamma)}^2 = \varepsilon^{-2} h_{\min,\gamma}^2 \frac{|\gamma|}{|\omega_\gamma|} \|R_\gamma\|_{L_2(\gamma)}^2 \\ &\leq \frac{\varepsilon^{-2} h_{\min,\gamma}^2}{|\omega_\gamma|} \sum_{K \subset \omega_\gamma} \left(\|u_h - u_H\|_K^2 \|\psi_\gamma\|_K^2 + \|r\|_{L_2(K)}^2 \|\psi_\gamma\|_{L_2(K)}^2 \right) \\ &\leq \frac{\varepsilon^{-2} h_{\min,\gamma}^2}{|\omega_\gamma|} \sum_{K \subset \omega_\gamma} \left(C_1 \|u_h - u_H\|_K^2 (\varepsilon^2 |K| h_{\min,K}^2 + \kappa^2 |K|) + \|r\|_{L_2(K)}^2 \|\psi_\gamma\|_{L_2(K)}^2 \right) \\ &\leq \sum_{K \subset \omega_\gamma} \left(C_1 \|u_h - u_H\|_K^2 + \frac{1}{6} \eta_K^2 + C_1 \zeta_K^2 \right), \end{aligned}$$

where we used (2.14), (2.15) and the assumption of the lemma. \square

Lemmata 7.5 and 7.6 are now used to prove estimates for $u_h - u_H$.

Lemma 7.8. *Let \mathcal{T}_h be a triangulation resulting from \mathcal{T}_H by applying one cycle of algorithm $\tilde{\mathcal{C}}$. Then:*

1. *we have the following global lower bound for the error reduction*

$$\|u_h - u_H\|_{\Omega}^2 \geq \frac{\theta^2 - \frac{1}{4}}{C_1 m_1^2(u - u_H, \mathcal{T}_H)} \|u - u_H\|_{\Omega}^2 - C_2 \zeta^2 - C_2 \sum_{K \in \mathcal{T}_H} \alpha_K^2 \|f - f_K\|_{L_2(K)}^2.$$

2. *If an additional condition on the marking holds:*

$$\forall \gamma \in \hat{\mathcal{E}}_h, \forall K : \gamma \subset \partial K : \quad \eta_K^2 \leq \eta_{\gamma}^2, \text{ or } K \in \hat{\mathcal{T}}_h, \quad (7.4)$$

then

$$\|u_h - u_H\|_{\Omega}^2 \geq \frac{\theta^2}{C_1 m_1^2(u - u_H, \mathcal{T}_H)} \|u - u_H\|_{\Omega}^2 - C_2 \zeta^2 - C_2 \sum_{K \in \mathcal{T}_H} \alpha_K^2 \|f - f_K\|_{L_2(K)}^2.$$

In the second case the factor in the first term of the right hand side is not so small.

Proof. **1.** By Lemmata 7.5, 7.7 and the marking strategy $\tilde{\mathbf{E}}$ we have

$$\begin{aligned} \theta^2 \eta_R^2 &\leq \sum_{\gamma \in \hat{\mathcal{E}}_H} \eta_{\gamma}^2 + \sum_{K \in \hat{\mathcal{T}}_H} \eta_K^2 = \sum_{\gamma \in \hat{\mathcal{E}}_H, h_{\min, \gamma} \ll \varepsilon \kappa^{-1}} \eta_{\gamma}^2 + \sum_{\gamma \in \hat{\mathcal{E}}_H, h_{\min, \gamma} \geq \varepsilon \kappa^{-1}} \eta_{\gamma}^2 + \sum_{K \in \hat{\mathcal{T}}_H} \eta_K^2 \\ &\leq \sum_{\gamma \in \hat{\mathcal{E}}_H, h_{\min, \gamma} \ll \varepsilon \kappa^{-1}} \eta_{\gamma}^2 + C \sum_{K \in \hat{\mathcal{T}}_H} \eta_K^2 + C \sum_{\gamma \in \hat{\mathcal{E}}_H} \sum_{K \subset \omega_{\gamma}, h_{\min, \gamma} \geq \varepsilon \kappa^{-1}} \alpha_K^2 \|f - f_K\|_{L_2(K)}^2 + \zeta^2 \\ &\leq C_1 \left(\sum_{\gamma \in \hat{\mathcal{E}}_H} \|u_h - u_H\|_{\omega_{\gamma}}^2 + \sum_{K \in \hat{\mathcal{T}}_H} \|u_h - u_H\|_K^2 + \zeta^2 + \sum_{K \in \mathcal{T}_H} \alpha_K^2 \|f - f_K\|_{L_2(K)}^2 \right) \\ &\quad + \frac{1}{12} \sum_{\gamma \in \hat{\mathcal{E}}_H} \sum_{K \subset \omega_{\gamma}, h_{\min, \gamma} \ll \varepsilon \kappa^{-1}} \eta_K^2. \end{aligned}$$

Each triangle K is counted at most three times in the last sum. Thus,

$$\begin{aligned} \|u_h - u_H\|_{\Omega}^2 &\geq \frac{\theta^2 - \frac{1}{4}}{C_1} \eta_R^2 - C_2 \zeta^2 - C_2 \sum_{K \in \mathcal{T}_H} \alpha_K^2 \|f - f_K\|_{L_2(K)}^2 \\ &\geq \frac{\theta^2 - \frac{1}{4}}{C_1 m_1^2(u - u_H, \mathcal{T}_H)} \|u_h - u_H\|_{\Omega}^2 - C_2 \zeta^2 - C_2 \sum_{K \in \mathcal{T}_H} \alpha_K^2 \|f - f_K\|_{L_2(K)}^2. \end{aligned}$$

2. Proceeding analogously to the first part and utilizing the condition (7.4) we obtain

$$\begin{aligned}
\theta^2 \eta_R^2 &\leq \sum_{\gamma \in \hat{\mathcal{E}}_H} \eta_\gamma^2 + \sum_{K \in \hat{\mathcal{T}}_H} \eta_K^2 = \sum_{\gamma \in \hat{\mathcal{E}}_H, h_{min, \gamma} \ll \varepsilon \kappa^{-1}} \eta_\gamma^2 + \sum_{\gamma \in \hat{\mathcal{E}}_H, h_{min, \gamma} \geq \varepsilon \kappa^{-1}} \eta_\gamma^2 + \sum_{K \in \hat{\mathcal{T}}_H} \eta_K^2 \\
&= \sum_{\gamma \in \hat{\mathcal{E}}_H, h_{min, \gamma} \ll \varepsilon \kappa^{-1}} \eta_\gamma^2 + \frac{1}{5} \sum_{\gamma \in \hat{\mathcal{E}}_H, h_{min, \gamma} \ll \varepsilon \kappa^{-1}} \sum_{K \subset \omega_\gamma, K \notin \hat{\mathcal{T}}_H} \eta_K^2 \\
&\quad - \frac{1}{5} \sum_{\gamma \in \hat{\mathcal{E}}_H, h_{min, \gamma} \ll \varepsilon \kappa^{-1}} \sum_{K \subset \omega_\gamma, K \notin \hat{\mathcal{T}}_H} \eta_K^2 + \sum_{\gamma \in \hat{\mathcal{E}}_H, h_{min, \gamma} \geq \varepsilon \kappa^{-1}} \eta_\gamma^2 + \sum_{K \in \hat{\mathcal{T}}_H} \eta_K^2 \\
&\leq \frac{6}{5} \sum_{K \in \hat{\mathcal{T}}_H} \sum_{\gamma \subset \partial K, h_{min, \gamma} \ll \varepsilon \kappa^{-1}} \eta_\gamma^2 - \frac{1}{5} \sum_{\gamma \in \hat{\mathcal{E}}_H, h_{min, \gamma} \ll \varepsilon \kappa^{-1}} \sum_{K \subset \omega_\gamma, K \notin \hat{\mathcal{T}}_H} \eta_K^2 \\
&\quad + C \sum_{K \in \hat{\mathcal{T}}_H} \eta_K^2 + C \sum_{\gamma \in \hat{\mathcal{E}}_H} \sum_{K \subset \omega_\gamma, h_{min, \gamma} \geq \varepsilon \kappa^{-1}} \alpha_K^2 \|f - f_K\|_{L_2(K)} + \zeta^2 \\
&\leq C_1 \left(\sum_{\gamma \in \hat{\mathcal{E}}_H} \|u_h - u_H\|_{\omega_\gamma}^2 + \sum_{K \in \hat{\mathcal{T}}_H} \|u_h - u_H\|_K^2 + \zeta^2 + \sum_{K \in \hat{\mathcal{T}}_H} \alpha_K^2 \|f - f_K\|_{L_2(K)}^2 \right).
\end{aligned}$$

□

We are now in a position to prove Theorem 7.3.

Proof of Theorem 7.3. Using the Galerkin orthogonality property and Lemma 7.8, we have

$$\begin{aligned}
\|u - u_h\|_\Omega^2 &= \|u - u_H\|_\Omega^2 - \|u_h - u_H\|_\Omega^2 \\
&\leq \|u - u_H\|_\Omega^2 \left(1 - \frac{\theta^2 - \frac{1}{4}}{C_1 m_1^2(u - u_H, \mathcal{T}_H)} \right) + C_2 \zeta^2 + C_2 \sum_{K \in \mathcal{T}_H} \alpha_K^2 \|f - f_K\|_{L_2(K)}^2.
\end{aligned}$$

If the condition (7.4) is satisfied then we have a similar estimate:

$$\begin{aligned}
\|u - u_h\|_\Omega^2 &= \|u - u_H\|_\Omega^2 - \|u_h - u_H\|_\Omega^2 \\
&\leq \|u - u_H\|_\Omega^2 \left(1 - \frac{\theta^2}{C_1 m_1^2(u - u_H, \mathcal{T}_H)} \right) + C_2 \zeta^2 + C_2 \sum_{K \in \mathcal{T}_H} \alpha_K^2 \|f - f_K\|_{L_2(K)}^2.
\end{aligned}$$

□

7.4 Numerical experiments

We are not able to prove the error reduction for the Algorithm $\tilde{\mathcal{C}}$ without the condition (7.1), which is used to bound the terms involving the element residuals. The essential point of this condition is that the new inner nodes are created per each "large" ($h_{min, K} \leq \varepsilon \kappa^{-1}$) triangle with at least one marked edge. We can heuristically observe that, inside each triangle will be sooner or later created a new node throughout the refinement process. Thus we leave this condition when implementing the adaptive algorithm in practice and implement the simplified algorithm.

As a test example for the adaptive algorithm we consider the reaction-diffusion equation

$$-\Delta u + \kappa^2 u = 0; \quad u = u_0 \text{ on } \partial\Omega$$

with known exact solution:

$$u = e^{-\kappa x} + e^{-\kappa y}.$$

For large κ the solution has the typical boundary layers along the Ox and Oy axes.

A sequence of meshes produced by the adaptive algorithm described in this chapter is shown in Figure 15. As we can mention the algorithm shows its ability to produce anisotropic meshes starting with a very coarse isotropic mesh, which is the crucial and very important property of this algorithm. Up to now there were only a few algorithms (including [7]) that can produce the anisotropic meshes starting from isotropic ones.

The convergence graph that shows the dependence of the error measured in the energy on the number of triangles is plotted in Figure 7.4. The convergence line can be conventionally split into three stages. At the beginning, the error looks proportional to the one obtained with the Shishkin meshes (with a factor of order 10). But later at the second stage, the algorithm starts to refine anisotropically and convergences definitely better than the a priori designed Shishkin meshes. After the optimal aspect ratio is reached the isotropic refinement does the job.

7.5 Discussion of the alignment measure m_1

In this section we follow Kunert [32] and state it here to comment the matching which appears on the lower error bound of the error estimator and in the error reduction constant.

The alignment measure $m_1(v, \mathcal{T}_h)$ (see [32]) has been introduced in (5.2) to measure how good an anisotropic mesh is aligned with an anisotropic function v . The worse this alignment is, the larger m_1 will be (recall that $m_1 \geq 1$). Simultaneously the quality of the error estimation and the error reduction constants deteriorates. Hence it is vital to determine $m_1(u - u_h, \mathcal{T}_h)$ to obtain a reliable estimates. Yet this is impossible since $u - u_h$ is not known. Two (partly overlapping) remedies seem possible.

1. Construct the mesh such that it is aligned with the anisotropic solution [42, 41, 30, 29, 50]. One can hope that using a refinement strategy given in this chapter with isotropic initial mesh results in a small $m_1(u - u_h, \mathcal{T}_h)$ and, in turn, in a useful upper error bound and an error reduction constant. Numerical experience strengthens this point.

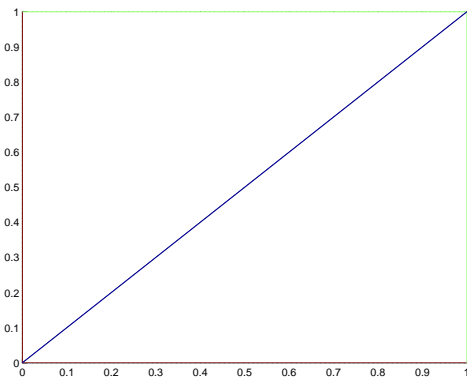
2. One can try to approximate $m_1(u - u_h, \mathcal{T}_h)$. For example, the gradient ∇u of the solution could be replaced by a recovered gradient $\nabla^R u_h$ (cf. e.g. [47]) giving

$$m_1^R(u - u_h, \mathcal{T}_h) := \frac{\left(\sum_{K \in \mathcal{T}_h} h_{min,K}^{-2} \cdot \|C_K^\top (\nabla^R u_h - \nabla u_h)\|_{L_2(K)}^2 \right)^{\frac{1}{2}}}{\|\nabla^R u_h - \nabla u_h\|_{L_2(\Omega)}}$$

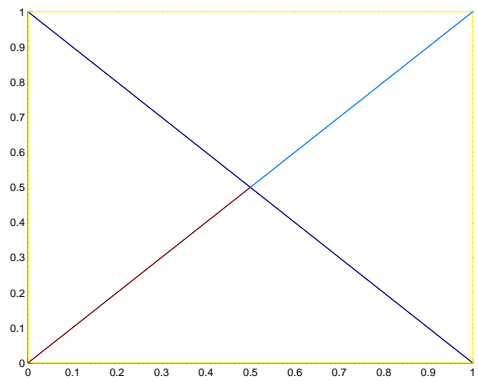
Indeed, numerous numerical examples (see [32]) yield that $m_1^R(u - u_h, \mathcal{T}_h) \approx m_1(u - u_h, \mathcal{T}_h)$ with only a small deviation (independent of the actual alignment of the mesh).

Acknowledgements

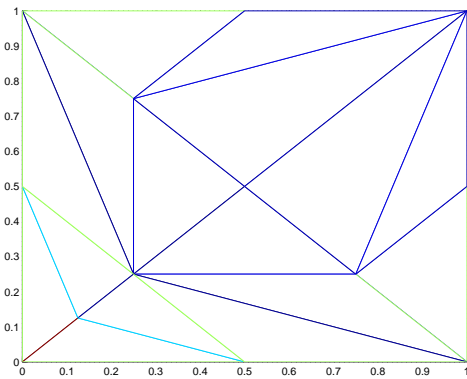
I would like to thank the Deutsche Forschungsgemeinschaft, SFB393 and the project AP 72/3–1 for their support. I would also like to express my sincere thanks to Thomas Apel, Cornelia Pester and Gerd Kunert for the fruitful discussions.



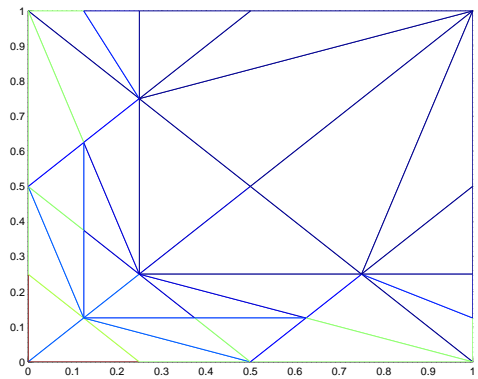
0



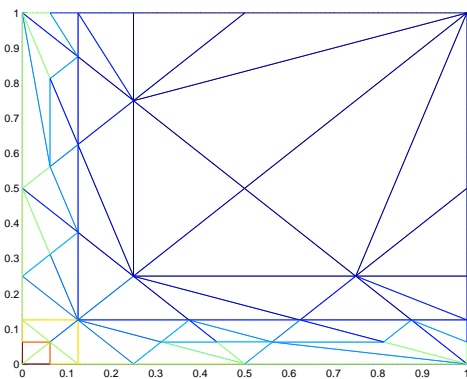
1



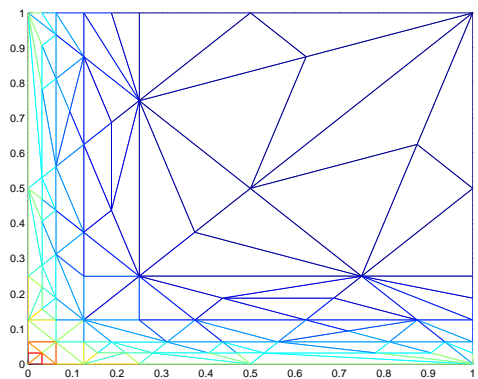
3



5



7



9

Figure 15: Example of mesh sequence starting from the isotropic mesh of two triangles. A number under each Figure stands for the number of iteration.

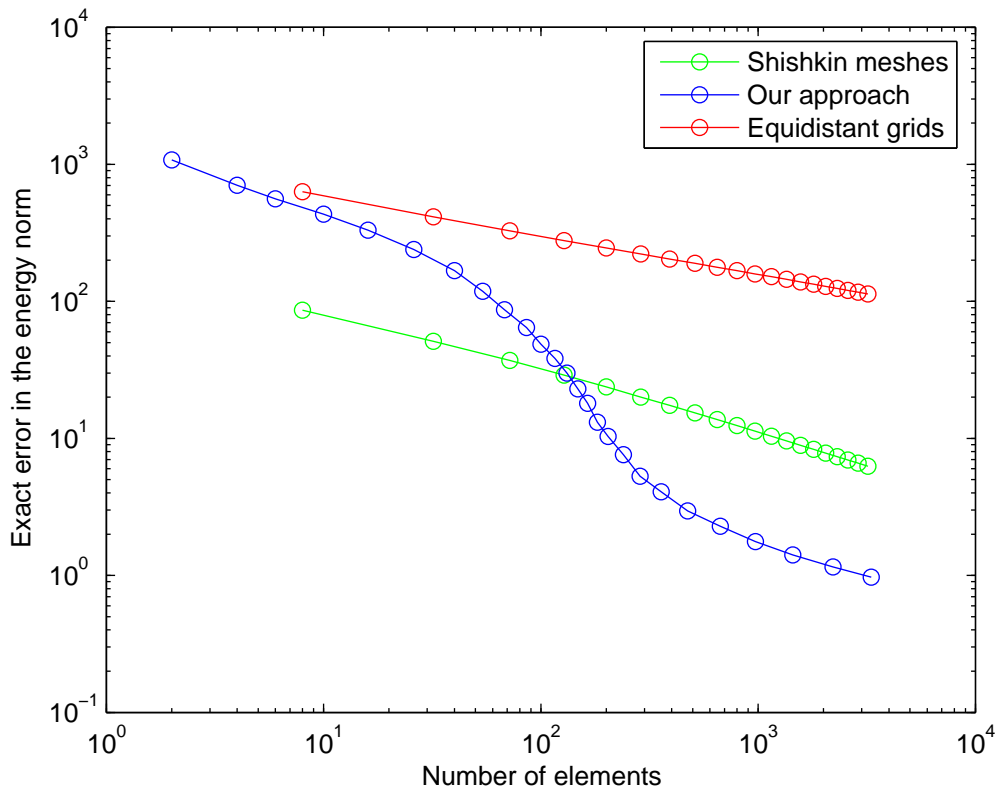


Figure 16: Convergence of the adaptive algorithm for $\kappa = 1000$.

References

- [1] A. Agouzal. On the saturation assumption and hierarchical a posteriori error estimator. *Comput. Methods Appl. Math.*, 2(2):125–131, 2002.
- [2] Mark Ainsworth and Ivo Babuška. Reliable and robust a posteriori error estimation for singularly perturbed reaction-diffusion problems. *SIAM J. Numer. Anal.*, 36(2):331–353, 1999. See also Corrigendum at <http://www.maths.strath.ac.uk/~aas98107/papers.html>.
- [3] Mark Ainsworth and J. Tinsley Oden. A unified approach to a posteriori error estimation using element residual methods. *Numer. Math.*, 65(1):23–50, 1993.
- [4] Mark Ainsworth and J. Tinsley Oden. *A posteriori error estimation in finite element analysis*. Pure and Applied Mathematics. A Wiley-Interscience Series of Texts, Monographs, and Tracts. Wiley, Chichester, 2000.
- [5] Thomas Apel. Treatment of boundary layers with anisotropic finite elements. *Z. Angew. Math. Mech.*, 1998.
- [6] Thomas Apel. *Anisotropic finite elements: local estimates and applications*. B. G. Teubner, Stuttgart, 1999.

- [7] Thomas Apel, Sergei Grosman, Peter K. Jimack, and Arnd Meyer. A new methodology for anisotropic mesh refinement based upon error gradients. *Appl. Numer. Math.*, 50(3–4):329–341, 2004.
- [8] Thomas Apel and Gert Lube. Anisotropic mesh refinement in stabilized Galerkin methods. *Numer. Math.*, 74(3):261–282, 1996.
- [9] Thomas Apel and Gert Lube. Anisotropic mesh refinement for a singularly perturbed reaction diffusion model problem. *Appl. Numer. Math.*, 26(4):415–433, 1998.
- [10] Thomas Apel and Serge Nicaise. Elliptic problems in domains with edges: anisotropic regularity and anisotropic finite element meshes. In *Partial differential equations and functional analysis*, pages 18–34. Birkhäuser, Boston, MA, 1996.
- [11] Ivo Babuška and A. Miller. A feedback finite element method with a posteriori error estimation. I: The finite element method and some basic properties of the a posteriori error estimator. *Comput. Methods Appl. Mech. Eng.*, 61:1–40, 1987.
- [12] Ivo Babuška and W.C. Rheinboldt. A posteriori error estimates for the finite element method. *Int. J. Numer. Methods Eng.*, 12:1597–1615, 1978.
- [13] Ivo Babuška and M. Vogelius. Feedback and adaptive finite element solution of one-dimensional boundary value problems. *Numer. Math.*, 44:75–102, 1984.
- [14] R.E. Bank and A. Weiser. Some a posteriori error estimators for elliptic partial differential equations. *Math. Comput.*, 44:283–301, 1985.
- [15] Peter Binev, Wolfgang Dahmen, and Ron DeVore. Adaptive finite element methods with convergence rates. *Numer. Math.*, 97(2):219–268, 2004.
- [16] H. Bufler and E. Stein. Zur Plattenberechnung mittels finiter Elemente. *Ingenieur Archiv*, (39):248–260, 1970.
- [17] Carsten Carstensen and Rüdiger Verfürth. Edge residuals dominate a posteriori error estimates for low order finite element methods. *SIAM J. Numer. Anal.*, 36(5):1571–1587, 1999.
- [18] Philippe G. Ciarlet. *The finite element method for elliptic problems*. North-Holland Publishing Co., Amsterdam, 1978. Studies in Mathematics and its Applications, Vol. 4.
- [19] Richard Courant. Variational methods for the solution of problems of equilibrium and vibrations. *Bulletin of the American Mathematical Society*, 49:1–23, 1943.
- [20] M. Crouzeix and V. Thomée. The stability in L_p and W_p^1 of the L_2 -projection onto finite element function spaces. *Math. Comput.*, 48:521–532, 1987.
- [21] Manfred Dobrowolski, Steffen Gräf, and Christoph Pflaum. On a posteriori error estimators in the infinite element method on anisotropic meshes. *ETNA, Electron. Trans. Numer. Anal.*, 8:36–45, 1999.
- [22] Vít Dolejší. Anisotropic mesh adaptation for finite volume and finite element methods on triangular meshes. *Comput. Vis. Sci.*, 1(3):165–178, 1998.
- [23] Willy Dörfler. A convergent adaptive algorithm for Poisson’s equation. *SIAM J. Numer. Anal.*, 33(3):1106–1124, 1996.

- [24] Willy Dörfler and Ricardo H. Nochetto. Small data oscillation implies the saturation assumption. *Numer. Math.*, 91(1):1–12, 2002.
- [25] Pascal Jean Frey and Paul-Louis George. *Mesh generation. Application to finite elements. Transl. from the 1999 French original by the authors.* Hermes Science Publishing, Oxford, 2000.
- [26] Paul-Louis George and Houman Borouchaki. *Delaunay triangulation and meshing. Application to finite elements. Transl. from the French.* Edition Hermes, Paris, 1998.
- [27] Sergey Grosman. An equilibrated residual method with a computable error approximation for a singularly perturbed reaction-diffusion problem on anisotropic finite element meshes. *Mathematical Modelling and Numerical Analysis (accepted for publication)*.
- [28] Roland Hagen, Steffen Roch, and Bernd Silbermann. *C^* -algebras and numerical analysis.* Marcel Dekker Inc., New York, 2001.
- [29] Ralf Kornhuber and Rainer Roitzsch. On adaptive grid refinement in the presence of internal and boundary layers. *IMPACT of computing in Sci. and Engrg.*, 2:40–72, 1990.
- [30] Dietmar Kröner and Rüdiger Beinert. In adaptive methods - algorithms, theory and applications. *Notes on Numerical Fluid Mechanics*, 46:38–53, 1994. Braunschweig, Vieweg.
- [31] Gerd Kunert. *A posteriori error estimation for anisotropic tetrahedral and triangular finite element meshes.* Logos Verlag, Berlin, 1999. Also PhD thesis, TU Chemnitz, <http://archiv.tu-chemnitz.de/pub/1999/0012/index.html>.
- [32] Gerd Kunert. An a posteriori residual error estimator for the finite element method on anisotropic tetrahedral meshes. *Numer. Math.*, 86(3):471–490, 2000.
- [33] Gerd Kunert. A local problem error estimator for anisotropic tetrahedral finite element meshes. *SIAM J. Numer. Anal.*, 39(2):668–689, 2001.
- [34] Gerd Kunert. Robust *a posteriori* error estimation for a singularly perturbed reaction-diffusion equation on anisotropic tetrahedral meshes. *Adv. Comp. Math.*, 15(1-4):237–259, 2001.
- [35] Gerd Kunert. Robust local problem error estimation for a singularly perturbed problem on anisotropic finite element meshes. *M2AN Math. Model. Numer. Anal.*, 35(6):1079–1109, 2001.
- [36] Gerd Kunert and Rüdiger Verfürth. Edge residuals dominate a posteriori error estimates for linear finite element methods on anisotropic triangular and tetrahedral meshes. *Numer. Math.*, 86(2):283–303, 2000.
- [37] P. Ladevèze and D. Leguillon. Error estimate procedure in the finite element method and applications. *SIAM J. Numer. Anal.*, 20:485–509, 1983.
- [38] Vladimir D. Liseikin. *Grid generation methods.* Scientific Computation. Springer, Berlin, 1999.
- [39] Vladimir D. Liseikin. *A computational differential geometry approach to grid generation.* Springer, Berlin, 2004.

- [40] Pedro Morin, Ricardo H. Nochetto, and Kunibert G. Siebert. Data oscillation and convergence of adaptive FEM. *SIAM J. Numer. Anal.*, 38(2):466–488, 2000.
- [41] J. Peraire, M. Vahdati, K. Morgan, and O. C. Zienkiewicz. Adaptive remeshing for compressible flow computation. *Journal of Computational Physics*, (72):449–466, 1987.
- [42] Kunibert G. Siebert. An a posteriori error estimator for anisotropic refinement. *Numer. Math.*, 73(3):373–398, 1996.
- [43] Rob Stevenson. An optimal adaptive finite element method. *SIAM J. Numer. Anal.*, 42(5):2188–2217, 2005.
- [44] Rob Stevenson. The uniform saturation property for a singularly perturbed reaction-diffusion equation. *Numer. Math.*, 101(2):355–379, 2005.
- [45] Yu.V. Vassilevski, V.G. Dyadechko, and K.N. Lipnikov. Hessian-based anisotropic mesh adaption in domains with discrete boundaries. *Russ. J. Numer. Anal. Math. Model.*, 20(4):391–402, 2005.
- [46] Rüdiger Verfürth. A posteriori error estimation and adaptive mesh-refinement techniques. *J. Comput. Appl. Math.*, 50(1-3):67–83, 1994.
- [47] Rüdiger Verfürth. *A review of a posteriori error estimation and adaptive mesh-refinement techniques*. Wiley-Teubner, Chichester; Stuttgart, 1996.
- [48] Rüdiger Verfürth. Robust a posteriori error estimators for a singularly perturbed reaction-diffusion equation. *Numer. Math.*, 78:479–493, 1998.
- [49] Michael Vogelius and Ivo Babuška. On a dimensional reduction method. I. The optimal selection of basis functions. *Math. Comp.*, 37(155):31–46, 1981.
- [50] Olgierd C. Zienkiewicz and J. Wu. Automatic directional refinement in adaptive analysis of compressible flows. *Internat. J. Numer. Methods Engrg.*, 37(13):2189–2210, 1994.

ERKLÄRUNG

Ich erkläre an Eides Statt, daß ich die vorliegende Arbeit (entsprechend der genannten Verantwortlichkeit) selbständig und nur unter Verwendung der angegebenen Literatur und Hilfsmittel angefertigt habe.

Chemnitz, May 3, 2006

THESES

1. When the finite element method is used to solve boundary value problems, the corresponding finite element mesh is appropriate if it reflects the behavior of the true solution. A posteriori error estimators are suited to construct adequate meshes. They are useful to measure the quality of an approximate solution and to design adaptive solution algorithms.
2. There exist different possibilities to measure the a posteriori error in the energy norm for the singularly perturbed reaction-diffusion equation $-\varepsilon^2 \Delta u + \kappa^2 u = f$. They include the residual a posteriori error estimator, the local Neumann problem error estimator (the equilibrated residual method), the local Dirichlet problem error estimator and the hierarchical error estimator.
3. Singularly perturbed problems yield in general solutions with anisotropic features, e.g. strong boundary or interior layers. For such problems it is useful to use anisotropic meshes in order to reach maximal order of convergence. Triangles should both be adapted in size and in shape in order to fit the function to be approximated. Moreover, the quality of the numerical solution rests on the robustness of the a posteriori error estimation with respect to both the anisotropy of the mesh and the perturbation parameters.
4. The equilibrated residual method is known to provide lower and upper error bounds as long as one solves auxiliary local Neumann problems exactly on each element. It is possible to provide a basis for an approximate solution of the aforementioned auxiliary problem and to show that this approximation does not affect the quality of the error estimation. A simple example of such a basis includes seven functions of third order over a triangle.
5. The simplest local error estimator from the implementation point of view is the so-called hierarchical error estimator. The reliability proof is usually based on two prerequisites: the error reduction property and the strengthened Cauchy-Schwarz inequality.
6. The strengthened Cauchy-Schwarz inequality holds for the following pair of spaces: the space of refinement functions used in hierarchical a posteriori error estimators and the space of piecewise linear finite element functions. The corresponding constant depends neither on the aspect ratio of the elements nor on the perturbation parameters. Furthermore, the strengthened Cauchy-Schwarz inequality holds also for two additional pairs of spaces which are needed for the hierarchical error estimation and the analysis of the adaptive algorithm for the reaction diffusion problem.
7. The error reduction property (also known as the saturation assumption as long as it is not proved) is shown to hold for the chosen enrichment space. This space includes the specially squeezed, based on the geometry of the mesh and the perturbation parameters, bubble functions of second order on each edge of the mesh and the usual bubble function of third order. The error reduction property together with the strengthened Cauchy-Schwarz inequality leads to the robustness proof of the hierarchical error estimator.
8. The convergence of the adaptive finite element method is understood in the sense that there exists a positive constant $\beta < 1$, such that in some norm $\|u - u_k\| \leq C\beta^k$

holds, where u is the exact solution and u_k is the finite element approximation in k steps of the adaptive algorithm. Such algorithms with the special marking and refinement strategies are known for the case of Poisson and reaction-diffusion problems on isotropic meshes. The marking strategy not only takes into account the usual information given by the edge and element error indicators, but also involves the additional control of the data oscillation terms in order to guarantee the convergence. The refinement strategy usually includes the creation of a new node inside each marked triangle. This, however, is not a must – the creation of a new node can be avoided for all triangles where the edge error indicator dominates the element error indicator.

9. Based on the idea from the isotropic case, the adaptive algorithm allowing anisotropic mesh refinement is based on analyzing the edge and element error indicators separately. According to this information only the marked entities (edges/elements) should be appropriately refined. The resulting algorithm satisfies the convergence property for the Poisson problem with a parameter β depending on some alignment measure m_1 .
10. The adaptive algorithm allowing anisotropic triangulations works in addition for the reaction-diffusion problem. The error reduces in each adaptive step, but the convergence property does not seem to be possible to be proven for this algorithm because of additional data oscillation terms.
11. Numerical experiments for the equilibrated residual method, for the hierarchical error estimator and for the adaptive algorithm confirm the theory. The adaptive algorithm shows its potential by creating the anisotropic mesh for the problem with the boundary layer starting with a very coarse isotropic mesh.