NORGES TEKNISK-NATURVITENSKAPELIGE
UNIVERSITET

Explicit Error Estimators for Bilinear Finite Elements on Anisotropic Rectangular Meshes

by

Hallgeir Melbø

PREPRINT
NUMERICS NO. 7/2001

NORWEGIAN UNIVERSITY OF SCIENCE AND
TECHNOLOGY
TRONDHEIM, NORWAY

This report has URL
http://www.math.ntnu.no/preprint/numerics/2001/N7-2001.ps
Address: Department of Mathematical Sciences, Norwegian University of Science and Technology, N-7491 Trondheim, Norway.

EXPLICIT ERROR ESTIMATORS FOR BILINEAR FINITE ELEMENTS ON ANISOTROPIC RECTANGULAR MESHES

HALLGEIR MELBØ

Department of Mathematical Sciences, NTNU, N-7491 Trondheim, Norway.
email: hallgeir@math.ntnu.no

Abstract.

After several decades of research a posteriori error estimation for finite element methods on isotropic meshes is now generally well understood. However, the question of how to treat anisotropic meshes is still not very well answered. Previously, several authors have concluded that one can find estimators that work well, also for anisotropic meshes, provided the anisotropy of the grid matches the anisotropy of the solution. In practice this may not be the case, and it is also difficult to verify whether it is satisfied or not. This article presents estimators that in typical situations will give very accurate results, and that under certain conditions will be asymptotically exact, *independently* of the anisotropy of the grid.

Key words: Error estimation, Anisotropic meshes, Finite element method, Adaptivity.

1 Introduction

The solutions of several important classes of problems exhibit localised and highly directional behaviour that is difficult to resolve using a standard finite element approach. Anisotropic, or stretched, finite elements offer obvious advantages, but fall outside the standard theory of finite element analysis. Further details and extensions to the anisotropic case are found in [2].

Adaptive finite element algorithms based on a posteriori error estimators are now routinely used for the efficient resolution of localised features. The behaviour and properties of a posteriori error estimators is relatively well understood for isotropic finite elements [1]. However, in the case of anisotropic elements, a number of fundamental issues remain open despite considerable theoretical and numerical investigation [4, 5, 6, 9, 10, 11]. The most important question concerns robustness of the estimators with respect to the aspect ratio ρ of the elements, where for rectangular elements ρ is defined as the longest edge of an element divided by the shortest edge of the same element. Specifically, it is desirable to have an estimator which is asymptotically exact independently of the aspect ratio. Letting the estimator be denoted by η, it should therefore ideally satisfy

$$\left| \frac{\eta}{|||e|||} - 1 \right| \leq \psi,$$

where $|||e|||$ denotes the error measured in some suitable norm, and where $\psi \to 0$, independently of ρ, when the longest element edge in the mesh approaches zero. We have found no proof showing that any known explicit estimator generally satisfies this requirement.

This article presents two types of residual based estimators that, under certain conditions, provide very good approximations of the true error, even for extremely graded grids. Moreover, the estimators seem to be robust in the sense that it is unnecessary to assume that the mesh is well refined towards the particular solution at hand. This means that the estimators are suitable for driving an adaptive refinement procedure. Under certain conditions some of the estimators are shown to be asymptotically exact, even on an element by element basis, independently of the anisotropy of the grid. To give an overview of the approach presented here, a short summary of the article is given below.

In Section 2 our model problem, which is the Poisson equation, and some well known error estimators are presented. An important quantity defined in this section is the *matching function*, suggested in [6], which enters typical robustness estimates on anisotropic grids, i.e.

$$\frac{1}{C}\eta \leq |||e||| \leq C\, m(e, \mathcal{P})\eta,$$

where C is a constant and $m(e, \mathcal{P})$ is the matching function which depends on the error e and the finite element partition $\mathcal{P}$. As will be seen, $m(e, \mathcal{P})$ behaves like the aspect ratio ρ on highly anisotropic grids, causing the above error bounds to degenerate.

The new estimators are defined in Section 3. They are constructed to be exact, independently of the aspect ratio, when the true solution locally is a second order polynomial and when there is no local change in pollution error. Results concerning the behaviour of the estimators when this is not satisfied are presented in Section 5. All the new estimators rely on making a reconstruction of the error on each element based on the changes in the derivatives of the finite element solution (known as the boundary residuals). Once the reconstructed error is in hand, it will be used in two different ways: One possibility is to simply compute the norm of the reconstructed error to obtain an error estimator. This kind of estimators will be referred to as R-estimators (e.g. η_R). Another possibility is to rewrite the norm of the error by the help of a residual equation and use the reconstructed error in the final expression. Estimators based on this strategy will be referred to as r-estimators (e.g. η_r).

Section 4 contains numerical examples comparing the new estimators to well known estimators, including upper bounds based on the matching function.

In Section 5 we give theoretical results for some of the new estimators. The results show how the estimators can be expected to behave in the presence of changing pollution errors, and when the true solution is *not* a second order polynomial locally. Furthermore, it is shown how parts of the standard analysis can be applied to the new estimators and also what goes wrong if the complete standard analysis is applied. This clearly shows where the matching function comes into play and why it should be avoided. In the last part of this section we prove that some of the new estimators are asymptotically exact.

Section 6 addresses the question of optimal refinement directions. In cases where the new estimators work well, i.e. cases where the true solution locally is close to a second order polynomial and the pollution error is close to constant, the new estimators carry information about the optimal direction for refining the grid. Our analytical results show that in typical situations previously suggested strategies will not produce optimal grids. In fact, the aspect ratio suggested by those strategies will in such situations, asymptotically, be $\sqrt{\rho_{opt}}$, where ρ_{opt} is the true optimal element aspect ratio.

In Section 7 the new estimators are extended from the Poisson problem to a reaction-diffusion problem, and numerical tests are performed.

2 Model problem and standard estimators

Let $\Omega \subset \mathbb{R}^2$ be a bounded domain with a piecewise linear Lipschitz boundary, and consider the problem

$$\begin{aligned} -\Delta u &= f \quad \text{in } \Omega \\ u &= 0 \quad \text{on } \partial\Omega, \end{aligned} \tag{2.1}$$

where $f \in L^2(\Omega)$. The weak form of this problem is: Find $u \in V_0$ such that

$$B(u,v) = L(v) \qquad \forall v \in V_0 \ , \tag{2.2}$$

where

$$\begin{aligned} V_0 &= \left\{ v \in H^1(\Omega) \ : \ v = 0 \text{ on } \partial\Omega \right\} \\ B(u,v) &= \int_\Omega \nabla u \cdot \nabla v \, dx \\ L(v) &= \int_\Omega f v \ dx. \end{aligned}$$

The bilinear form defines an energy norm for a function $v \in V_0$ in the following way,

$$|||v|||^2 = B(v,v).$$

Let $X_0 \subset V_0$ denote a finite element space constructed using continuous piecewise bilinear functions on a mesh $\mathcal{P}$ of rectangular elements. The set of inter-element edges is denoted by $\partial\mathcal{P}$. The finite element approximation of the problem is then to find $u_X \in X_0$ such that

$$B(u_X, v_X) = L(v_X) \quad \forall v_X \in X_0. \tag{2.3}$$

The error, defined as

$$e = u - u_X, \tag{2.4}$$

belongs to the space V_0 and satisfies

$$\begin{aligned} B(e,v) &= L(v) - B(u_X, v) && \forall\, v \in V_0 \\ B(e, v_X) &= 0 && \forall\, v_X \in X_0. \end{aligned}$$

The last equation is the characteristic Galerkin orthogonality property. The bilinear form may be written as a sum over all elements,

$$B(e,v) = \sum_{K\in\mathcal{P}} \left\{ \int_K fv\ dx - \int_K \nabla u_X \nabla v\ dx \right\}. \tag{2.5}$$

For each element K in the partition $\mathcal{P}$ of Ω and each edge $\gamma \in \partial\mathcal{P} \setminus \partial\Omega$ we define the following residuals,

$$\begin{aligned} r &= f + \Delta u_X && \text{in } K \\ R &= -\left[\tfrac{\partial u_X}{\partial n}\right]\Big|_\gamma && \text{on } \gamma, \end{aligned}$$

where r is known as the interior residual, R is called the boundary residual, and on an edge γ separating elements K and K' we define the "jump" as

$$\left[\frac{\partial u_X}{\partial n}\right]\Bigg|_\gamma = \mathbf{n}_K \cdot (\nabla u_X)_K + \mathbf{n}_{K'} \cdot (\nabla u_X)_{K'}\ .$$

Applying Green's theorem to the last integral in equation (2.5) and introducing these two types of residuals, leads to the *residual equation*,

$$B(e,v) = \sum_{K\in\mathcal{P}} \int_K rv\ dx + \sum_{\gamma\in\partial\mathcal{P}} \int_\gamma Rv\ ds \quad \forall v \in V_0. \tag{2.6}$$

Consider the class of residual type estimators of the form

$$\eta_K^2 = \eta_K^2(h_1, h_2, \mathbf{h}_3) = h_1 h_2 ||r||_K^2 + \frac{1}{2} \sum_{\gamma\subset\partial K} h_{3,\gamma} ||R||_\gamma^2\ , \tag{2.7}$$

where $h_1, h_2, \mathbf{h}_3$ are weights to be chosen, $||\cdot||$ is the L^2-norm over the indicated domain, and where the corresponding global estimator is defined by

$$\eta^2 = \sum_{K\in\mathcal{P}} \eta_K^2\ . \tag{2.8}$$

Note that applying a scalar as the last argument in η_K should be interpreted as giving all boundary residuals the same weight. For a rectangular element K, let $h_{-,K}$ (or simply h_-) be the length of the shortest edge of K, let $h_{+,K}$ (or simply h_+) be the length of the longest edge of K, and let $\mathbf{h}_{\perp\gamma,K}$ (or simply $\mathbf{h}_{\perp\gamma}$) be the vector composed of the lengths of the edges perpendicular to each edge of K. The simplified notation will be used when there can be no confusion. The standard explicit estimator now takes the form

$$\eta^2_{max,K} = \eta^2_K(h_+, h_+, h_+),$$

and the corresponding global estimator η_{max} can be shown to give an upper bound for the energy norm of the error. However, this estimator will provide a lower bound for the error *only* on isotropic meshes. An alternative estimator, η_{min}, defined by

$$\eta^2_{min,K} = \eta^2_K(h_-, h_-, h_-),$$

provides a lower bound for the error, but on anisotropic grids it does not give an upper bound. It has been shown, see Siebert [11], that the global estimator corresponding to

$$\eta^2_{S,K} \quad = \quad \eta^2_K(h_-, h_-, \mathbf{h}_{\perp\gamma}) \tag{2.9}$$

satisfies

$$C_1\eta_S \leq |||e||| \leq C_2 m(e, \mathcal{P})\eta_S, \tag{2.10}$$

where higher order terms have been omitted for simplicity. Similar results have been shown for triangular and tetrahedral meshes in [6], where the function m appearing in (2.10), the so-called *matching function*, was defined:

$$[m(v, \mathcal{P})]^2 = \frac{\sum\limits_{K\in\mathcal{P}} \sum\limits_{i=1}^{2} \frac{h^2_{i,K}}{h^2_{-,K}} ||\nabla_i v||^2_K}{\sum\limits_{K\in\mathcal{P}} \sum\limits_{i=1}^{2} ||\nabla_i v||^2_K}\ , \tag{2.11}$$

where ∇_i denotes differentiation in direction i. This function measures how well the grid matches the function v, and it may take on any real value larger than or equal to 1. The problem with this approach is, however, twofold. First, the actual error, which is of course unknown, needs to be estimated to compute the matching function. This may be difficult, and if one actually has a good approximation of the error, it may be regarded as inefficient use of information only to use it for computing this function. More seriously, however, even if m is computed exactly, the ratio between the upper and lower bound on the error equals m, and it can therefore not be bounded. Note that if for any direction i and any element K

$$\frac{h_{i,K}}{h_{-,K}} \to \infty \quad \text{ without } \nabla_i e|_K \to 0,$$

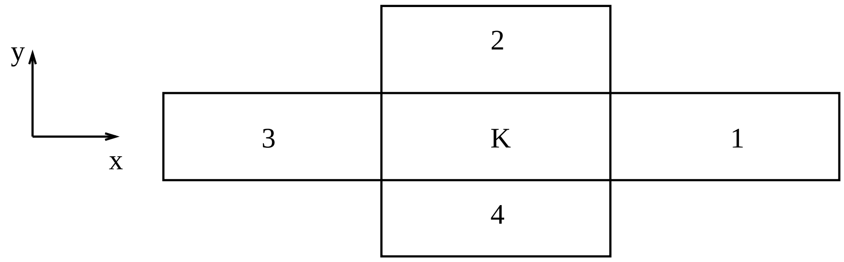

Figure 3.1: The patch $\tilde{K}$ of elements surrounding K.

then

$$m(e, \mathcal{P}) \to \infty.$$

The result presented in [11] does not actually contain the matching function. Instead it is assumed that $m(e, \mathcal{P})$ is bounded by a constant. The behaviour of the above kind of estimator and the importance of the matching function seems not to be very well documented in the literature.

3 New estimators

In this section new estimators are introduced which will be exact, independently of the anisotropy of the grid, if the true solution locally is a second order polynomial, and if there locally is no change in the pollution error. The estimators will be derived under these conditions. However, in later sections they will be analysed for the general case when the conditions are relaxed. A key issue when deriving the estimators will be to reconstruct a second order polynomial approximation, e_K, to the error on each element, which means that the estimators will work well when such an approximation captures the main features of the exact error.

For a patch $\tilde{K}$ of elements surrounding an element K (see Figure 3.1), let u^* be the second order polynomial

$$u^* = a_1 + a_2 x + a_3 y + a_4 xy + a_5 x^2 + a_6 y^2, \tag{3.1}$$

defined on $\tilde{K}$. For any patch P of elements, let $\mathcal{N}(P)$ denote the nodes in the patch. Assuming u is continuous, suppose that u^* is the best approximation to u in the sense that

$$\begin{array}{rcll} u|_n & = & u^*|_n & \text{for } n \in \mathcal{N}(K) \\ u & = & u^* + \tilde{\epsilon}(x, y) & \text{for } (x, y) \in K, \end{array} \tag{3.2}$$

where $\tilde{\epsilon}$ is small according to some measure which may be chosen later. Let the vertices of element K be numbered from 1 to 4, and for element K let the

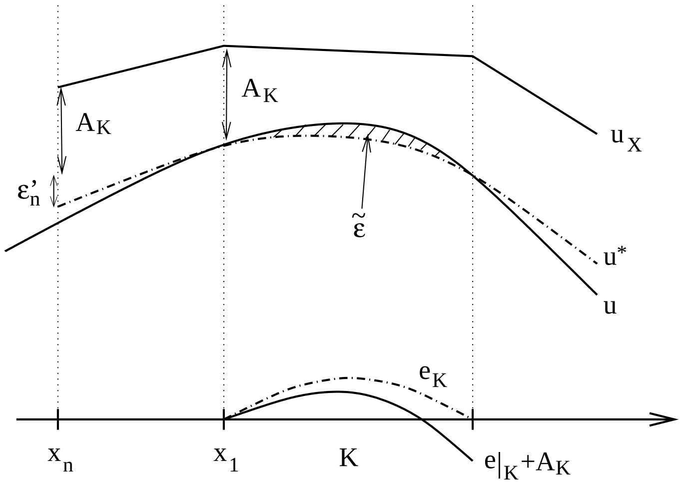

Figure 3.2: A schematic one dimensional drawing showing the relation between some of the variables involved. The axis overlaps with an element edge of element K and two of its neighbours. The position of the local node number 1 in element K, for which the pollution error is A_K, is denoted by x_1, whereas x_n is the position of a node $n \in \mathcal{N}(\tilde{K})$, where the distance between u_X and u^* is $A_K + \epsilon'_n$. Note that $u^* = u$ at the nodes in $\mathcal{N}(K)$, but they are not necessarily equal at all the nodes in $\mathcal{N}(\tilde{K})$. Furthermore, $e|_K$ is the true error on element K, and e_K is the reconstructed error on element K.

constant A_K denote the pollution error in (the local) node number 1, i.e.,

$$u_X|_{n_1} = u|_{n_1} + A_K.$$

The local numbering is chosen such that each node in the mesh locally is node number 1 in only one element. Then at least one element has its node number 1 situated on $\partial\Omega$, which means that A_K vanishes for this element. For element K, let ϵ'_n be defined through the relation

$$u_X|_n = u^*|_n + A_K + \epsilon'_n \qquad \text{for each } n \in \mathcal{N}(\tilde{K}), \tag{3.3}$$

which implies $\epsilon'_1 = 0$. With this definition the vector $\boldsymbol{\epsilon}'$, containing the nodal values $\{\epsilon'_n\}$ for this patch, keeps track of how well the numerical solution locally fits the shape of the second order polynomial which approximates the true solution. A schematic drawing of some of the variables involved can be seen in Figure 3.2.

Asymptotically u can be well approximated by a second order polynomial, which justifies $\tilde{\epsilon}$ being small. Away from singularities the change in pollution error is typically very small, sometimes even of higher order, so it also seems reasonable to assume that the entries of $\boldsymbol{\epsilon}'$ are small.

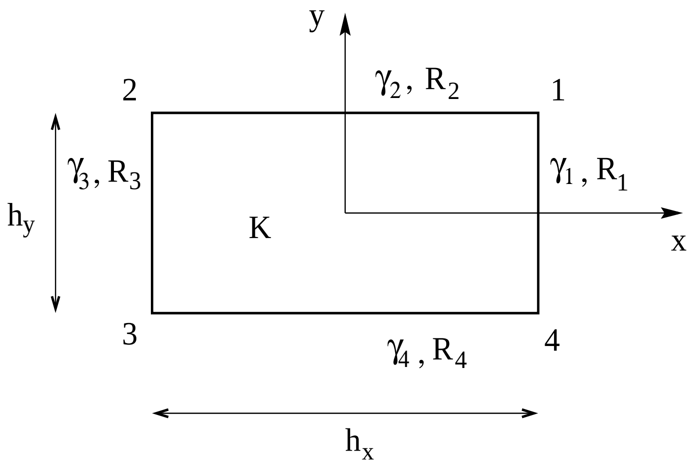

Figure 3.3: The element K with local numbering of nodes, edges, and boundary residuals.

Let $h_{x,K}$ and $h_{y,K}$ (or simply h_x, h_y) be as shown in Figure 3.3. Then the element size will only be allowed to change slowly according to the following assumption:

(A1) *There exists a global constant δ, $0 \leq \delta < 1$, such that for every pair of neighbouring elements K and K'*

$$|h_{x,K} - h_{x,K'}| < h_{x,K}\delta \qquad \text{and} \qquad |h_{y,K} - h_{y,K'}| < h_{y,K}\delta\ . \tag{3.4}$$

For each element K in the partition $\mathcal{P}$, let the global constant ϵ satisfy

$$\max_{(x,y)\in K} |\tilde{\epsilon}(x,y)| < \epsilon \qquad \text{and} \qquad \max_{n\in\mathcal{N}(\tilde{K})} |\epsilon'_n| < \epsilon.$$

In order to present the ideas in the simplest possible setting, we initially assume

$$\epsilon = 0,$$

as was mentioned at the beginning of this section. The choice of local numbering then implies $A_K = 0$ for all elements. This means that the solution is assumed to be a second order polynomial, i.e.

$$u|_K = u^*|_K,$$

and that u_X is exact at the nodes. Note that if the true solution *is* a second order polynomial (on some subdomain) and the boundary values (for that subdomain) are exact, then the finite element solution will indeed be exact at the nodes (in that subdomain). A simple calculation reveals that under the above assumption the boundary residuals will be constant on each edge, and the interior residual

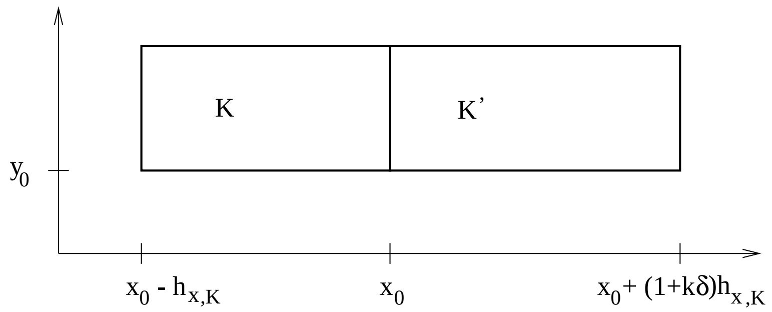

Figure 3.4: Two neighbouring elements K and K'.

will be constant on each element. Equation (2.6) may therefore be rewritten as

$$B(e,e) \quad = \quad \sum_{K\in\mathcal{P}} \left\{ \bar{r}_K \int_K e\, dx + \frac{1}{2} \sum_{\gamma\in\partial K} \bar{R}_\gamma \int_\gamma e\, ds \right\}, \tag{3.5}$$

where variables marked with a bar denote average values. The assumption $\epsilon = 0$ allows u to be expressed as

$$u|_K = u^*|_K = a_1 + a_2 x + a_3 y + a_4 xy + a_5 x^2 + a_6 y^2, \qquad \text{where } (x,y) \in K.$$

Consider the element shown in Figure 3.3. Since $\epsilon = 0$ we have

$$u\big(\pm\frac{h_x}{2}\ ,\ \pm\frac{h_y}{2}\big) = u_X\big(\pm\frac{h_x}{2}, \pm\frac{h_y}{2}\big)\ ,$$

and consequently, since the finite element functions are bilinear, the error is given by

$$e|_K = a_5 \left(x^2 - \frac{h_x^2}{4} \right) + a_6 \left(y^2 - \frac{h_y^2}{4} \right). \tag{3.6}$$

Let the two neighbouring elements K and K' be as shown in Figure 3.4. Since u_X interpolates u^* at the nodes, the boundary residual on the vertical edge between the two elements is given by

$$R = -\left(\frac{u^*(x_0,y_0) - u^*(x_0 - h_{x,K}, y_0)}{h_{x,K}} - \frac{u^*(x_0 + (1+k\delta)h_{x,K}, y_0) - u^*(x_0,y_0)}{(1+k\delta)h_{x,K}} \right).$$

A simple computation then reveals that

$$R = (2 + k\delta) h_{x,K} a_5.$$

More generally, using the element numbering in Figure 3.1, the factors k_i relating the length scales of neighbouring elements are defined through the following equations,

$$\begin{aligned} h_{x,1} &= (1+k_1\delta)h_{x,K} \\ h_{x,3} &= (1+k_3\delta)h_{x,K} \\ h_{y,2} &= (1+k_2\delta)h_{y,K} \\ h_{y,4} &= (1+k_4\delta)h_{y,K}, \end{aligned}$$

where it is clear from the assumptions that $|k_i| \leq 1$. For simplicity we omit the subscript K in what follows. Using the notation from Figure 3.3, the boundary residual across edge 1 is

$$R|_{\gamma_1} = \bar{R}_1 = (2+k_1\delta)a_5 h_x, \tag{3.7}$$

since the jump is constant on each edge. Here we allow for a slight abuse of notation as $\bar{R}$ can be indexed both by the local edge number and by the edge itself. Corresponding results are obtained for the other edges. We also define the following average boundary residuals,

$$\begin{aligned} \bar{R}_x &= (\bar{R}_1 + \bar{R}_3)/2 \\ \bar{R}_y &= (\bar{R}_2 + \bar{R}_4)/2\,. \end{aligned} \tag{3.8}$$

If the local edge number i lies on the global boundary $\partial\Omega$, then R_i and k_i are replaced by the values from the opposite edge in the element. Consequently, we arrive at the following expressions for the coefficients in the error (3.6),

$$\begin{aligned} a_5 &= \tfrac{1}{2h_x}\left(\tfrac{\bar{R}_1}{2+k_1\delta} + \tfrac{\bar{R}_3}{2+k_3\delta}\right) \\ a_6 &= \tfrac{1}{2h_y}\left(\tfrac{\bar{R}_2}{2+k_2\delta} + \tfrac{\bar{R}_4}{2+k_4\delta}\right)\,. \end{aligned} \tag{3.9}$$

Here we have used the fact that a_5 may be expressed both through $\bar{R}_1$ and $\bar{R}_3$, and that a_6 may be expressed through $\bar{R}_2$ and $\bar{R}_4$. We emphasise that these expressions hold only for the case $\epsilon = 0$. The reconstruction e_K of the elementwise error, which under these special circumstances is exact, is taken to be

$$\begin{aligned} e_K &= \frac{1}{2h_x}\left(\frac{\bar{R}_1}{2+k_1\delta} + \frac{\bar{R}_3}{2+k_3\delta}\right)\left(x^2 - \frac{h_x^2}{4}\right) \\ &\quad + \frac{1}{2h_y}\left(\frac{\bar{R}_2}{2+k_2\delta} + \frac{\bar{R}_4}{2+k_4\delta}\right)\left(y^2 - \frac{h_y^2}{4}\right). \end{aligned} \tag{3.10}$$

Note that the true error on element K is denoted by $e|_K$, whereas the reconstructed error on element K is denoted by e_K. However, under the current assumptions the two are identical. Based on (3.5) we define the local estimator $\alpha^*_{r,K}$ by

$$\alpha^*_{r,K} = \bar{r}_K \int_K e_K \, dx + \frac{1}{2}\sum_{\gamma\in\partial K} \bar{R}_\gamma \int_\gamma e_K \, ds\,. \tag{3.11}$$

Substitution and direct integration reveals that the expressions appearing in (3.11) can be evaluated as

$$\begin{aligned}\int_K e_K\, dxdy &= -\frac{h_x h_y}{12}\Big\{ h_x\Big(\frac{\bar{R}_1}{2+k_1\delta}+\frac{\bar{R}_3}{2+k_3\delta}\Big)+h_y\Big(\frac{\bar{R}_2}{2+k_2\delta}+\frac{\bar{R}_4}{2+k_4\delta}\Big)\Big\}\\ \int_{\gamma_1} e_K\, ds &= \int_{\gamma_3} e_K\, ds = -\frac{h_y^2}{12}\Big(\frac{\bar{R}_2}{2+k_2\delta}+\frac{\bar{R}_4}{2+k_4\delta}\Big)\\ \int_{\gamma_2} e_K\, ds &= \int_{\gamma_4} e_K\, ds = -\frac{h_x^2}{12}\Big(\frac{\bar{R}_1}{2+k_1\delta}+\frac{\bar{R}_3}{2+k_3\delta}\Big).\end{aligned}$$

This leads to the following estimator,

$$\alpha_r^* = \sum_{K\in\mathcal{P}} \alpha_{r,K}^*\ , \tag{3.12}$$

where

$$\begin{aligned}\alpha_{r,K}^* &= -\frac{1}{12}\Big\{\ h_x^2(\bar{r}_K h_y+\bar{R}_y)\Big(\frac{\bar{R}_1}{2+k_1\delta}+\frac{\bar{R}_3}{2+k_3\delta}\Big)\\ &\qquad + h_y^2(\bar{r}_K h_x+\bar{R}_x)\Big(\frac{\bar{R}_2}{2+k_2\delta}+\frac{\bar{R}_4}{2+k_4\delta}\Big)\Big\}\ .\end{aligned}\tag{3.13}$$

The derivation of the estimator shows that it is exact if the true solution is a quadratic, since in that case $\epsilon = 0$.

A simpler expression is obtained if the variation of the grid is ignored, i.e., we put $\delta = 0$ and define,

$$\alpha_{r,K} = -\frac{1}{12}\Big\{h_x^2(\bar{r}_K h_y+\bar{R}_y)\bar{R}_x + h_y^2(\bar{r}_K h_x+\bar{R}_x)\bar{R}_y\Big\}\ . \tag{3.14}$$

This estimator and the corresponding global estimator α_r are exact if $\epsilon = \delta = 0$. There is, however, no guarantee that $\alpha_r^*, \alpha_r, \alpha_{r,K}^*$, and $\alpha_{r,K}$ are positive quantities if δ and ϵ are nonzero. We will therefore also define the corresponding positive valued local and global estimators,

$$\begin{aligned}(\eta_{r,K}^*)^2 &= |\alpha_{r,K}^*| \qquad (\eta_r^*)^2 = \sum_{K\in\mathcal{P}}(\eta_{r,K}^*)^2\\ \eta_{r,K}^2 &= |\alpha_{r,K}| \qquad \eta_r^2 = \sum_{K\in\mathcal{P}}\eta_{r,K}^2.\end{aligned}\tag{3.15}$$

Note that α estimates $|||e|||^2$, whereas η estimates $|||e|||$.

Since equation (3.10) gives an expression for the reconstructed error on element K, which is exact when $\epsilon = 0$, we can easily derive a second estimator simply by computing the energy norm of that expression. This leads to

$$\begin{aligned}(\eta_{R,K}^*)^2 &= B(e_K, e_K)\\ &= \frac{h_x h_y}{12}\Bigg\{\Big(\frac{\bar{R}_1}{2+k_1\delta}+\frac{\bar{R}_3}{2+k_3\delta}\Big)^2+\Big(\frac{\bar{R}_2}{2+k_2\delta}+\frac{\bar{R}_4}{2+k_4\delta}\Big)^2\Bigg\}.\end{aligned}\tag{3.16}$$

The global estimator is chosen to be

$$(\eta_R^*)^2 = \sum_{K\in\mathcal{P}} (\eta_{R,K}^*)^2. \tag{3.17}$$

From the construction of the estimator it is clear that it will be exact in the case when $\epsilon = 0$.

Also here we can easily obtain a simpler estimator by ignoring the changes in the grid. This gives the compact estimator

$$\eta_{R,K}^2 = \frac{h_x h_y}{12}\big(\bar{R}_x^2 + \bar{R}_y^2\big), \tag{3.18}$$

with the global estimator η_R defined in the same manner as η_R^*.

Note that $\eta_R^*, \eta_R, \eta_{R,K}^*$, and $\eta_{R,K}$ are all non-negative quantities, so there is no need for extra definitions similar to those made for the α_r- and α_r^*-estimators. It is also worth noticing that η_R^* and η_R do not contain the interior residual. Nevertheless, they are exact when $\epsilon = 0$ and when $\epsilon = \delta = 0$, respectively. For these special cases it is therefore evident that all the information about the error lies in the boundary residuals.

It is also worth noticing that when the boundary residuals are constant on each edge and $\bar{R}_x = \bar{R}_1 = \bar{R}_3$ in addition to $\bar{R}_y = \bar{R}_2 = \bar{R}_4$, the estimator $\eta_{R,K}^2$ may be rewritten in the following way,

$$\eta_{R,K}^2 = \frac{1}{24} \sum_{\gamma\in\partial K} h_{\perp\gamma} ||R||_\gamma^2\,,$$

where $h_{\perp\gamma}$ denotes the length of the element edge perpendicular to γ, e.g. $h_{\perp\gamma_1} = h_x$. This is, within a multiplicative constant, the dominant part of the estimator η_S defined in (2.9), which was suggested by Siebert [11]. The term involving the interior residual in the estimator η_S is weighted by the square of the shortest length scale, so that term will typically show optimal convergence when the grid is refined, independently of the anisotropy of the grid. For η_S it has been shown that the error is bounded from below by this estimator and from above by this estimator multiplied by the matching function. However, our derivation suggests that in most cases the matching function may not be necessary.

It is obvious that one can construct situations where the error is nonzero but where $\bar{R}_x$ and $\bar{R}_y$ vanish. This would make $\eta_{R,K}$ vanish without the actual error being zero. However, this problem should not be over-emphasised as it can only occur in very special situations, provided the finite element solution is convergent: It may happen if the average boundary residuals on the element edges are pairwise equal but have opposite signs, if the boundary residuals are odd functions centred at the midpoint of the edge, and if the true solution oscillates on a sub-grid scale. The first two situations are easily detected, and asymptotically the third situation cannot occur.

4 Numerical examples

The first group of estimators that was derived will be referred to as r-estimators, while the second group will be referred to as R-estimators. The r-estimators are further divided into α- and η-estimators.

From an implementational point of view, the estimators not marked with an asterisk are preferred, since in those cases the local estimators do not require information about the length scales of the surrounding elements. For the $r-$estimators it is not quite clear if one should use the α- or the η-estimators. However, since we would like non-negative refinement indicators we suggest to use $\eta_{r,K}$ instead of $\alpha_{r,K}$ as the refinement indicator. As the different r-estimators have shown almost identical behaviour in our numerical tests, we will only consider the $\eta-$estimators in the examples that follow.

In addition to the two new estimators η_r and η_R, three estimators, η_{max}, η_{min}, and $\eta_{natural}$, will be considered in the numerical examples. With the help of the notation in Section 2, the three estimators can be described by equations (2.7) and (2.8) as follows,

$$\begin{aligned} \eta^2_{h_{max},K} &= \eta^2_K(h_+, h_+, h_+) \\ \eta^2_{natural,K} &= \eta^2_K(h_-, h_+, \mathbf{h}_{edge}) \\ \eta^2_{h_{min},K} &= \eta^2_K(h_-, h_-, h_-), \end{aligned}$$

where $\mathbf{h}_{edge}$ is the vector composed of the lengths of the element sides. The reason for considering $\eta_{natural}$ is that this choice lies between $\eta_{h_{max}}$ and $\eta_{h_{min}}$, and it could be seen as a natural guess for a weighting between the different residuals. Note that, except for the constant, the estimator η_S defined in Section 2 has approximately the same behaviour as η_R, and it is therefore left out.

4.1 Prescribed refinement

First we run a test where the assumptions from the beginning of Section 3 are known to be satisfied, i.e., $\epsilon = \delta = 0$. Since any smooth solution asymptotically can be well approximated by a second order polynomial this should constitute a generic test case. Moreover, it is also a very simple problem, so it is unlikely that an estimator will perform well in more complex practical cases if it fails to give satisfactory results here. The exact solution is chosen to be

$$u = x^2 + ay^2, \qquad a = 1/10\ ,$$

such that most of the action is in the x-direction. This can be seen in Figure 4.1. We will first keep a fixed number of elements in the y-direction and refine uniformly in the x-direction. Thereafter, the same procedure will be repeated keeping the number of elements in the x-direction fixed and refining in the y-direction. In Figure 4.2 the error and the different estimators are plotted, and it is also shown what happens if the matching function is applied. What is seen

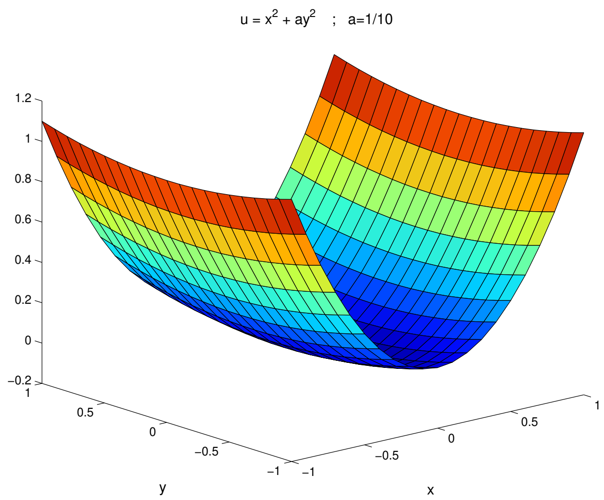

Figure 4.1: Plot of solution, first test case.

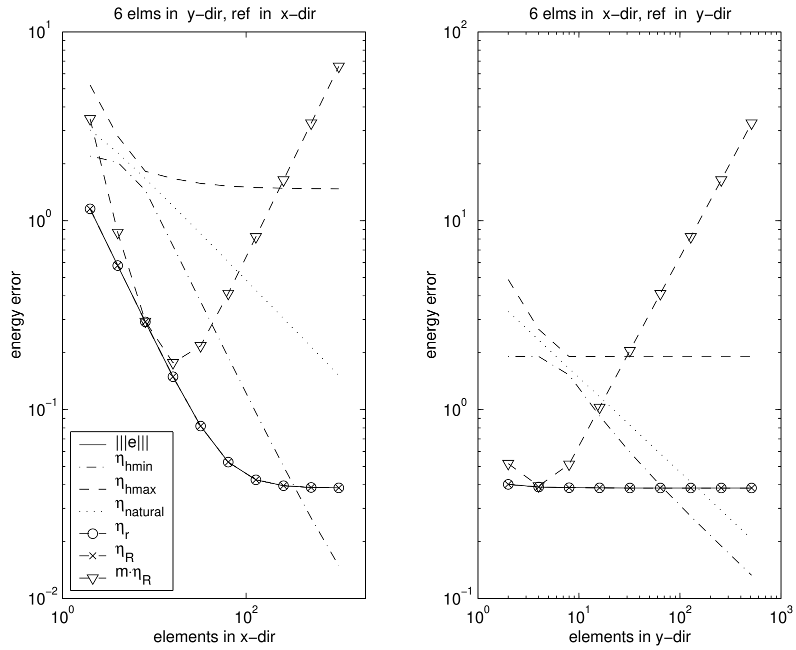

Figure 4.2: The various estimators applied to the test case $u = x^2 + ay^2$ with $a = 1/10$. The estimators η_r and η_R are exact in this case.

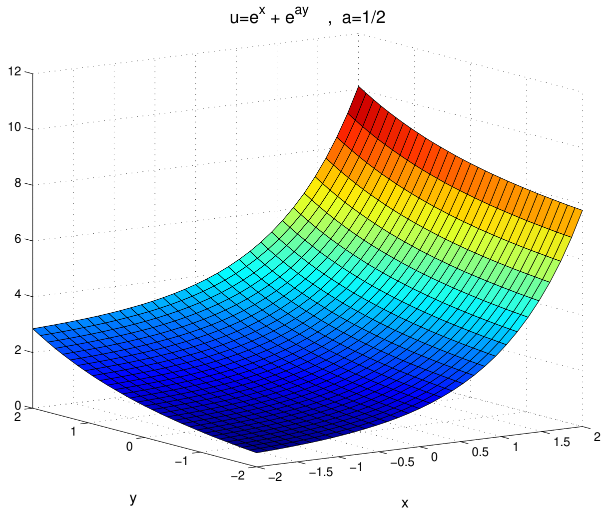

Figure 4.3: Plot of solution, second test case.

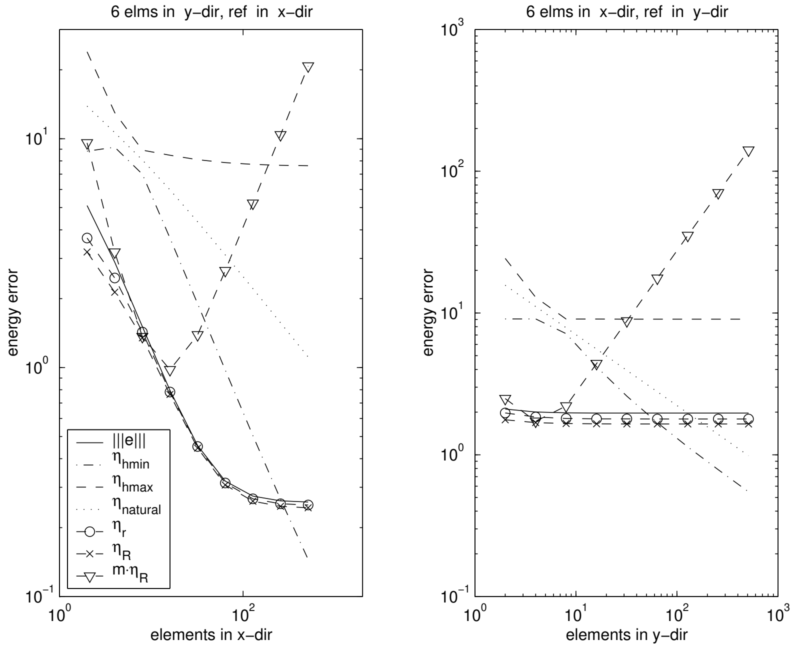

Figure 4.4: The results for the second test case, $u = e^x + e^{ay}$ with $a = 1/2$.

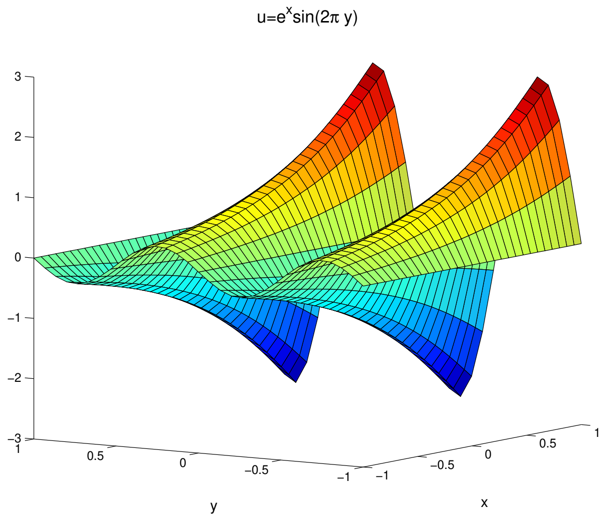

Figure 4.5: Plot of solution, third test case.

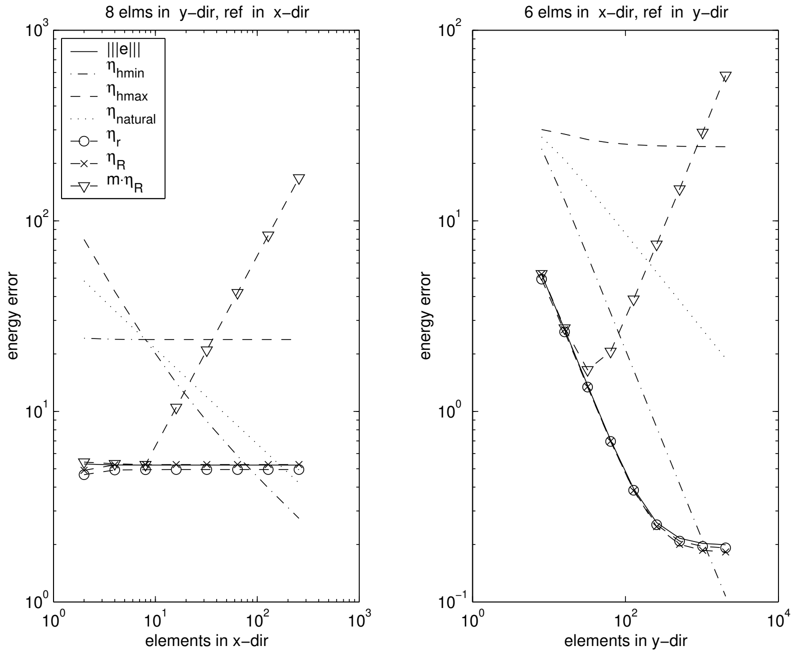

Figure 4.6: The results for the third test case, $u = e^x \sin(2\pi y)$.

from the plots is that none of the traditional error estimators give satisfactory results when the grid is highly graded, while the new estimators, as expected, give exact estimates for this case. The application of the matching function gives an upper bound which increases with increasing aspect ratio. This shows that any assumption of the type

$$m(e, \mathcal{P}) \leq C,$$

cannot be expected to hold on anisotropic grids.

In the second test example the exact solution has the following form,

$$u = e^x + e^{ay} \qquad a = 1/2\,.$$

A plot of the solution is shown in Figure 4.3. Here ϵ clearly cannot be zero, but as the solution is not too far from a second order polynomial, we can still expect the estimators to give quite good results. The refinement strategy is the same as in the first example, and the results are shown in Figure 4.4. It is clear from the plot that also here the new estimators are behaving close to perfectly, while the traditional estimators show the same behaviour as in the first case. Again we observe that the matching function seems to be redundant.

The third example has a more complex solution, namely

$$u = e^x \sin(2\pi y)\,,$$

which can be seen in Figure 4.5. Now the solution is no longer close to a second order polynomial globally, but if it can be locally well approximated by a second order polynomial, the new estimators should still give reasonable results. The refinement strategy is the same as in the first two test cases: The number of elements is kept fixed in one direction while the elements are refined uniformly in the other direction. The results are shown in Figure 4.6. Again the new estimators are close to exact, while the standard estimators do not give much information about the behaviour of the actual error. Also in this case the matching function seems to be unnecessary.

4.2 Adaptive refinement

So far the refinements have been prescribed in all the examples, and the new estimators have given promising results. We will now run two examples where the estimators are used for adaptive refinement, which is what local error estimators are usually aimed at. In the examples the refinement is done uniformly in the sense that if an element is marked for refinement it will be split into 4 new equally shaped elements, and the anisotropy of the elements is provided by the initial grid. The examples will therefore highlight how the various estimators may differ in their choice of elements marked for refinement when the aspect ratio of (some of) the elements is high. Clearly, the estimator η_S would indicate approximately the same relative distribution of the error as η_R, even if it was multiplied by the

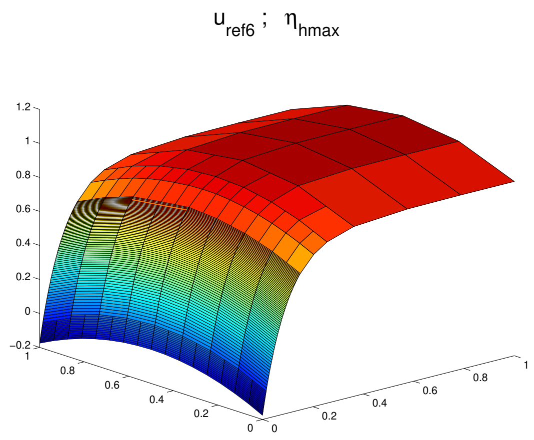

Figure 4.7: The solution and the grid produced after 6 refinements using $\eta_{h_{max}}$ as refinement indicator.

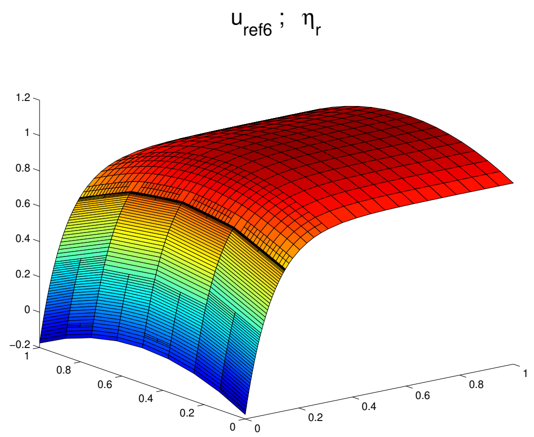

Figure 4.8: The solution and the grid produced after 6 refinements using η_r as refinement indicator.

matching function. However, it seems that so far no such comparison between η_S and standard estimators has been made.

First we consider a case where the new estimators are compared to $\eta_{h_{max}}$. The exact solution is chosen to be

$$u = -e^{-x/a} + b\Big(y - \frac{1}{2}\Big)^2 + c\,,$$
$$a = 1/10, \quad b = -7/10, \quad c = 1\,.$$

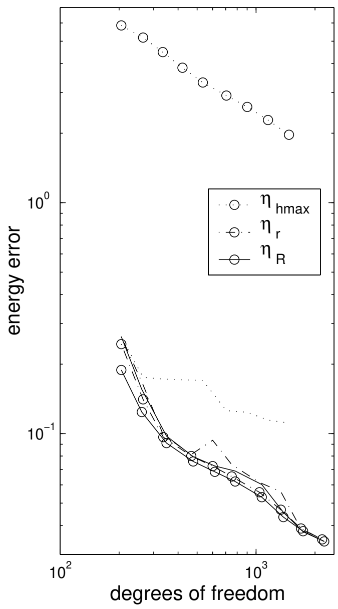

Figure 4.9: The results for the first adaptive test case. The estimators η_r, η_R, and $\eta_{h_{max}}$ are used as refinement indicators for the adaptive procedure. The lines with circles are the estimated values, and the corresponding lines without circles are the actual errors produced when using each of the estimators for refinement.

The domain is $[0,1]^2$, and the initial grid has some anisotropic elements in the rather weak boundary layer. Figures 4.7 and 4.8 show the grids and the solutions produced after 6 refinements using $\eta_{h_{max},K}$ and $\eta_{r,K}$ as refinement indicators. Note that $\eta_{h_{max}}$ fails to see the large errors outside the boundary layer since it only pays attention to the long edges where the grid is anisotropic. Figure 4.9 shows the resulting convergence plots. There we see that $\eta_{h_{max}}$ over estimates the actual error associated with it by a large factor, and it does not manage to bring down the error efficiently, whereas the new estimators give a rapid decrease in the actual error, and the estimators are almost identical to the true errors associated with the grids they produce.

The second adaptive test case compares the new estimators to $\eta_{h_{min}}$. The example is very similar to the previous one, and this time the solution is chosen to be

$$u = -e^{-x/a} + b\Big(y - \frac{1}{2}\Big)^2 + c\,,$$
$$a = 1/10, \quad b = -4, \quad c = 1\,.$$

The domain is $[0,1]^2$, and the initial grid has anisotropic elements between

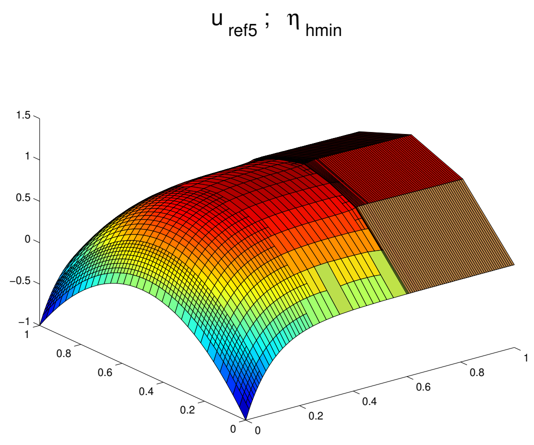

Figure 4.10: The solution and the grid produced after 5 refinements using $\eta_{h_{min}}$ as refinement indicator.

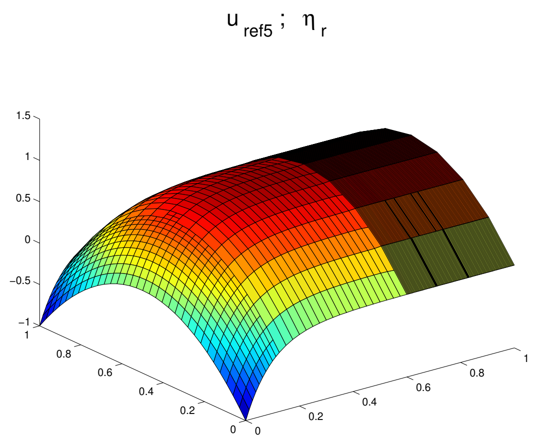

Figure 4.11: The solution and the grid produced after 5 refinements using η_r as refinement indicator.

$x = 0.6$ and $x = 1$. The solutions and the grids produced after 5 refinements using $\eta_{h_{min},K}$ and $\eta_{r,K}$ as refinement indicators are shown in Figures 4.10 and 4.11. We can clearly see that $\eta_{h_{min}}$ fails to detect the large errors that are present where the elements are anisotropic, while the new estimator does refine in this area. It is evident that the two estimators suggest very different refinement strategies. The comparisons of estimators and the actual errors produced are found in Figure 4.12. It shows that the estimator $\eta_{h_{min}}$ indicates there is a good rate of convergence, whereas the actual error associated with it levels out and hardly

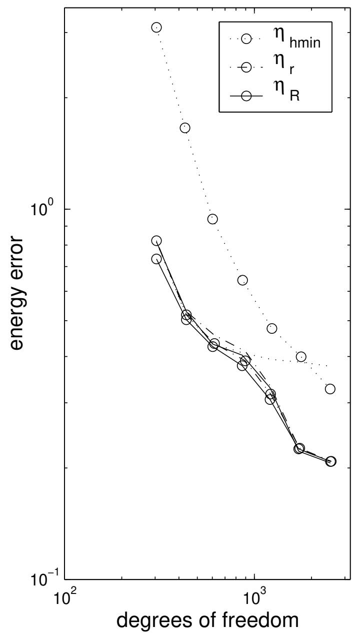

Figure 4.12: The results for the second adaptive test case. The estimators η_r, η_R, and $\eta_{h_{min}}$ are used as refinement indicators for the adaptive procedure. The lines with circles are the estimated values, and the corresponding lines without circles are the actual errors produced when using each of the estimators for refinement.

decreases at all. The new estimators manage to bring down the error a lot faster than $\eta_{h_{min}}$, and they are both very close to the actual error all the way through the adaptive procedure.

If we keep on refining, both $\eta_{h_{min}}$ and $\eta_{h_{max}}$ will eventually suggest refinements all over the domain, which will bring the error down. However, it is clear that they may suggest far from optimal refinement strategies, and they may totally fail to describe the behaviour of the actual error.

5 Theoretical results for some of the estimators

In this section we present some theoretical results for the new estimators. First, upper bounds for the difference between the error and some of the estimators are derived. Thereafter, a standard framework for finding upper bounds will be applied to η_R, and in this part also the applicability of the actual framework will be investigated. Finally, it will be shown that under certain conditions some of the new estimators are asymptotically exact.

For an element K we define the following maximum finite differences,

$$\begin{aligned}
\hat{D}_x(\mathbf{v}) &= \max_{n\in\mathcal{N}(\tilde{K})} |D_x(v_n)| \\
\hat{D}^2_x(\mathbf{v}) &= \max_{n\in\mathcal{N}(\tilde{K})} |D^2_x(v_n)| \\
\hat{D}_{xy}(\mathbf{v}) &= |D_{xy}(\mathbf{v})|,
\end{aligned}$$

where $D_x(v_n)$ and $D^2_x(v_n)$ are the first order (forward or backward) and the second order (central) finite differences in the x-direction, centred at node n, and involving only nodes in $\tilde{K}$. Further, $D_{xy}(\mathbf{v})$ is the cross-term finite difference centred in the middle of element K. The first two maximum finite differences above may also be equipped with the subscripts y, γ, or $\perp\gamma$ instead of x. The subscript γ should be interpreted as the difference taken in the direction parallel to γ, whereas $\perp\gamma$ means the direction perpendicular to γ. For a more detailed description, see Appendix A.

For a simpler notation we introduce the function $\phi_{r,K}(\boldsymbol{\epsilon}')$, defined by

$$\phi_{r,K}(\boldsymbol{\epsilon}') = h_x\hat{D}_x(\boldsymbol{\epsilon}') + h_y\hat{D}_y(\boldsymbol{\epsilon}') + h_x^2\hat{D}^2_x(\boldsymbol{\epsilon}') + h_y^2\hat{D}^2_y(\boldsymbol{\epsilon}') + h_xh_y\hat{D}_{xy}(\boldsymbol{\epsilon}'), \qquad (5.1)$$

where $\boldsymbol{\epsilon}'$ was defined in (3.3). With this definition $\phi_{r,K}(\boldsymbol{\epsilon}')$ will be small if the entries of $\boldsymbol{\epsilon}'$ are small. Note that the value of ϵ'_n for $n \in \mathcal{N}(K)$ expresses the change in pollution error from (the local) node 1 to (the local) node n. Therefore, in parts of the domain where the pollution error is changing slowly $\phi_{r,K}$ will usually be small. This typically happens where ∇u changes slowly. Further, let the quantities Q_0, Q_1, and Q_2 be defined by

$$\left.\begin{aligned}
Q_0^2 &= \sum_{K\in\mathcal{P}} h_+^2\Big\{ ||r-\bar{r}_K||^2_K \\
&\qquad + \sum_{\gamma\subset\partial K} h_\gamma^3\Big(h_{\perp\gamma}\big|D^2_{\perp\gamma}(D_\gamma(\boldsymbol{\epsilon}'))\big|^2 + \tfrac{\delta^2}{h_{\perp\gamma}}\big|D_{xy}(\boldsymbol{\epsilon}')\big|^2\Big)\Big\} \\
Q_1 &= \sum_{K\in\mathcal{P}} \phi_{r,K}(\boldsymbol{\epsilon}')\Big\{|\bar{r}_K|h_xh_y + \sum_{\gamma\subset\partial K} h_\gamma|\bar{R}_\gamma|\Big\} \\
Q_2 &= \sum_{K\in\mathcal{P}} ||\bar{r}_K||_K\,||\tilde{\epsilon}_K||_K + \sum_{\gamma\subset\partial K} ||\bar{R}_\gamma||_\gamma\,||\tilde{\epsilon}_K||_\gamma\,.
\end{aligned}\right\} \qquad (5.2)$$

We will only briefly comment on the typical behaviour of these quantities provided the solution is smooth enough. In Q_0, the contribution from the first term will, during anisotropic refinement, first decrease and then become a constant if h_+ is kept fixed. The two other terms will also typically decrease as long as anisotropic refinement reduces the change in the pollution error sufficiently in the direction of refinement.

The quantity Q_1 will not vanish in the anisotropic limit ($\rho\to\infty$, $h_-\to 0$) unless $\phi_{r,K}$ vanishes. However, down to a certain point $\phi_{r,K}$ typically *does* decrease when the grid is refined asymptotically, until the differences in the under-refined direction become dominant.

As long as the second order approximation of the true solution improves as the grid is refined anisotropically, the quantity Q_2 will decrease during the refinement. However, at a certain point the approximation error associated with the under-refined direction typically becomes dominant, which prevents $|\tilde{\epsilon}|$ from decreasing any further, and the term becomes a constant.

Summarising, since we will prove that the absolute difference between the global error and the estimator can be bounded mainly by $|||e|||Q_0$, Q_1, and Q_2, we can expect the estimator to have very much the same behaviour as the true error. If the grid is refined in only one direction, one will then typically end up with a constant difference between the error and the estimator. This happens when the local change in pollution error in the under-refined direction becomes dominant, and when the second order approximation u^* does not improve during further anisotropic refinement. When this happens, either the approximation is perfect, in which case the constant vanishes, or neither the true error nor the estimator vanishes during further anisotropic refinement. This also shows that the choice of u^* should be a compromise between paying attention to $\tilde{\epsilon}$ and $\boldsymbol{\epsilon}'$, although the two will be very much correlated.

Before stating the results, we need to introduce the following notation: Let $[A_\gamma]$ be the difference in pollution error between elements K and K' sharing the edge γ, i.e.

$$[A_\gamma] = A_K - A_{K'} \qquad \gamma = \bar{K} \cap \bar{K}',$$

and define $< v >_\gamma$ to be the average value of the (possibly discontinuous) function v on γ, in the sense that

$$< v >_\gamma = \frac{1}{2}\big(v|_K + v|_{K'}\big)\big|_\gamma \qquad \gamma = \bar{K} \cap \bar{K}'.$$

For a global boundary edge $\gamma \subset \partial K$ we define $[A_\gamma]$ to be A_K and $< v >_\gamma$ to be $v|_K$ evaluated on γ. The results can now be stated for the general case when $\epsilon \geq 0$ and $\delta \geq 0$.

Theorem 5.1 *Let e be as defined in (2.4), let the finite element mesh consist of rectangular bilinear elements, and suppose assumption (A1) is satisfied. Let the error estimators α_r^* and $\alpha_{r,K}^*$ be defined by (3.12) and (3.13), and let Q_0, Q_1, and Q_2 be as defined in (5.2). Then there exists a constant C such that*

$$\Big| |||e|||^2 - \alpha_r^* \Big| \leq C \left\{ |||e|||\, Q_0 + Q_1 + Q_2 + \sum_{\gamma \in \partial\mathcal{P}} h_\gamma^{1/2} \left\| \frac{\partial u}{\partial n} - \left\langle \frac{\partial u_X}{\partial n} \right\rangle_\gamma \right\|_\gamma \big|[A_\gamma]\big| \right\}.$$

PROOF. The theorem is proved in Appendix A. □

We define $\phi_{R,K}(\boldsymbol{\epsilon}')$ by

$$\phi_{R,K}(\boldsymbol{\epsilon}') = \hat{D}_x(\boldsymbol{\epsilon}') + \hat{D}_y(\boldsymbol{\epsilon}') + h_x \hat{D}_x^2(\boldsymbol{\epsilon}') + h_y \hat{D}_y^2(\boldsymbol{\epsilon}') + (h_x + h_y) D_{xy}(\boldsymbol{\epsilon}'). \quad (5.3)$$

Note that $\phi_{R,K}$ has discrete derivatives, whereas $\phi_{r,K}$ only has differences. The requirements on $\boldsymbol{\epsilon}'$ that will make $\phi_{R,K}$ decrease during refinement can therefore be seen as stricter than what is needed for $\phi_{r,K}$.

Theorem 5.2 *Let e be as defined in (2.4), let $\tilde{\epsilon}$ be defined by (3.2), let the finite element mesh consist of rectangular bilinear elements, and suppose assumption (A1) is satisfied. Define the global error estimators η_R, η_R^*, and their corresponding local estimators as in equations (3.16) to (3.18), and let $\phi_{R,K}(\boldsymbol{\epsilon}')$ be defined by (5.3). Then there exists a constant C such that*

$$\begin{aligned} \Big| |||e|||_K - \eta^*_{R,K} \Big| &\leq C\left\{ |K|^{1/2}\phi_{R,K}(\boldsymbol{\epsilon}') + |||\tilde{\epsilon}_K|||_K \right\} \\ \Big| |||e|||_K - \eta_{R,K} \Big| &\leq C\left\{ |K|^{1/2}\phi_{R,K}(\boldsymbol{\epsilon}') + |||\tilde{\epsilon}_K|||_K + \eta_{R,K}\,\mathcal{O}\big(\delta^{1/2}\big) \right\}, \end{aligned}$$

where $|K|$ is the area of element K.
PROOF. The theorem is proved in Appendix A. $\square$

Note that the above result is *local* in the sense that it indications how good each local error estimator is instead of only stating what happens globally. Therefore this result is stronger than what was obtained for α_r^*. The result can be discussed along the same lines as was done prior to Theorem 5.1. Since the result is local we conclude that $\eta^*_{R,K}$ will give very good results in the parts of Ω where u is smooth and the pollution error is changing slowly, provided the main features of the solution are captured also in the under-refined direction. It is clear that in these situations also the estimator $\eta_{R,K}$ will work well if δ is small.

5.1 The estimator η_R viewed in the standard framework

The derivation of η_R was based on quite strong assumptions, but still the estimator seems to work well in all our test cases. It has been shown that generally this estimator (or rather η_S, which is larger than η_R) gives an upper bound for the error if it is multiplied by the matching function. However, our experiments indicate that in practice this seems unnecessary, and in this section we show why this is usually what is observed. The estimator will be analysed using the standard framework, only considering typical cases.

Without loss of generality, assume that for any fixed element K, h_y is the length of the short edges, such that

$$h_x \gg h_y.$$

Then, only the following cases need to be analysed:

- Under-refined/Well refined: All elements satisfy

$$\frac{h_x}{h_y}\left|\left|\frac{\partial e}{\partial x}\right|\right|_K \leq C\left|\left|\frac{\partial e}{\partial y}\right|\right|_K.$$

- Over-refined: At least one element satisfies

$$\frac{h_x}{h_y}\left\|\frac{\partial e}{\partial x}\right\|_K \gg \left\|\frac{\partial e}{\partial y}\right\|_K.$$

Again we point out that some authors have chosen to assume that the matching function is bounded by a constant, whereas we have chosen to explicitly show the dependency on this function. In the first case (under-refined/well refined) the matching function (2.11) *is* bounded by a constant, so according to (2.10),

$$\frac{1}{C}\eta_S \leq |||e||| \leq C\eta_S.$$

The main difference between η_S and η_R is the term involving the interior residual, which in η_S is given the weight $h_-^2 \equiv h_y^2$. With this weight it converges with optimal rate regardless of the anisotropy of the grid, so this term will not play a very important role. It therefore seems reasonable that η_R should give satisfactory results for the first case, so we will concentrate on the over-refined case. In the following analysis it will be assumed that the whole domain has the same anisotropy, but a similar analysis could equally well be done for a subdomain.

Define the Clément interpolation operator $\mathcal{R}$ by

$$\mathcal{R}v = \sum_{n\in\mathcal{N}} \theta_n \bar{v}_{\omega_n}, \tag{5.4}$$

where $\mathcal{N}$ is the set of all interior nodes, θ_n is the piecewise bilinear basis function associated with node n, and ω_n is the nodal patch comprised of the elements having node n in common. With this notation $\bar{v}_{\omega_n}$ should be understood as the average of v taken over ω_n.

In the over-refined case it seems reasonable to assume that there exist small positive (global) parameters ϵ_u and ϵ_r, possibly depending on the longest element length scale, such that for all elements $K \in \mathcal{P}$,

$$\begin{aligned}
\max_{(x,y)\in K} \left|\overline{e|}_K - \mathcal{R}e\right| &\leq \epsilon_u \\
\left|\overline{e|}_K - \overline{e|}_{\gamma_2}\right| &\leq \epsilon_r \\
\left|\overline{e|}_K - \overline{e|}_{\gamma_4}\right| &\leq \epsilon_r,
\end{aligned}$$

where the edge numbering in Figure 3.3 has been applied, and where $\overline{e|}_K$ and $\overline{e|}_{\gamma_i}$ are averages of the true error taken over the domain indicated in the subscripts. The first inequality will be referred to as the *patchwise uniformity assumption*, whereas the last two inequalities will be referred to as the *over-refinement assumption*. In the case where the actual solution is a second order polynomial over the patch and the finite element solution is exact at the nodes, up to an additional constant for each patch, the first inequality holds with $\epsilon_u = 0$. Since the grid is assumed to be over-refined, the average of the error taken over one of the long edges is typically close to the average value for the element, which justifies ϵ_r being small. In the anisotropic limit ($h_y \to 0$, $h_x/h_y \to \infty$) the last two inequalities hold with $\epsilon_r = 0$.

Before going through the analysis, we quote two results from [11].

Lemma 5.3 *For $v \in H^1(K)$*

$$||v||_{\gamma_i} \le C\Big\{h_{\perp\gamma_i}^{-1/2}||v||_K + h_{\perp\gamma_i}^{1/2}||\nabla_{\perp\gamma_i} v||_K\Big\},$$

where $\gamma_i \in \partial K$ and $\perp\gamma_i$ is the direction orthogonal to the edge γ_i. $\square$

Lemma 5.4 *Let assumption (A1) in Section 3 be satisfied. Then, for all $v \in H_0^1(\Omega)$*

$$||v - \mathcal{R}v||_K \le C\sum_i h_{i,K}||\nabla_i v||_{\tilde{\tilde{K}}}$$

$$||\nabla_k(v - \mathcal{R}v)||_K \le C\sum_i \frac{h_{i,K}}{h_{k,K}}||\nabla_i v||_{\tilde{\tilde{K}}},$$

where $\mathcal{R} : H_0^1(\Omega) \to X_0$ is defined in (5.4), and $\tilde{\tilde{K}}$ is the patch of elements having at least one node in common with K. $\square$

Together, Galerkin orthogonality and the residual equation state that

$$B(e, e) = \sum_{K\in\mathcal{P}} \int_K r(e - \mathcal{R}e)\, dx + \frac{1}{2}\sum_{\gamma\in\partial K}\int_\gamma R(e - \mathcal{R}e)\, ds.$$

The term involving the interior residual may be rewritten as

$$\begin{aligned}\int_K r(e - \mathcal{R}e)dx &= \bar{r}_K\int_K(e - \overline{e}|_K)dx + \bar{r}_K\int_K(\overline{e}|_K - \mathcal{R}e)dx + \int_K(r - \bar{r}_K)(e - \mathcal{R}e)dx \\ &= \bar{r}_K\int_K(\overline{e}|_K - \mathcal{R}e)dx + \int_K(r - \bar{r}_K)(e - \mathcal{R}e)dx\ .\end{aligned}$$

By the aid of the patchwise uniformity assumption, the Cauchy-Schwarz inequality, and Lemma 5.4, we obtain

$$\Big|\int_K r(e - \mathcal{R}e)dx\Big| \le \epsilon_u|K|^{1/2}||\bar{r}_K||_K + h_x||r - \bar{r}_K||_K\, |||e|||_{\tilde{\tilde{K}}}\,. \tag{5.5}$$

Next, we consider one of the short edges. From Lemma 5.3 and Lemma 5.4 we obtain

$$\begin{aligned}||e - \mathcal{R}e||_{\gamma_1} &\le C\Big\{h_x^{-1/2}||e - \mathcal{R}e||_K + h_x^{1/2}||\nabla_x(e - \mathcal{R}e)||_K\Big\} \\ &\le Ch_x^{-1/2}\sum_i h_i||\nabla_i e||_{\tilde{\tilde{K}}}\ .\end{aligned}$$

Since $h_x > h_y$ this leads to

$$||e - \mathcal{R}e||_{\gamma_1} \le C h_x^{1/2} |||e|||_{\tilde{\tilde{K}}} \ ,$$

and hence,

$$\begin{aligned} \left| \int_{\gamma_1} R(e - \mathcal{R}e) ds \right| &\le \Big(||\bar{R}_1||_{\gamma_1} + ||R - \bar{R}_1||_{\gamma_1} \Big) ||e - \mathcal{R}e||_{\gamma_1} \\ &\le C h_x^{1/2} \Big(||\bar{R}_1||_{\gamma_1} + ||R - \bar{R}_1||_{\gamma_1} \Big) |||e|||_{\tilde{\tilde{K}}}. \qquad (5.6) \end{aligned}$$

The contribution from the long edge γ_2 can be expressed as follows,

$$\int_{\gamma_2} R(e - \mathcal{R}e)\, ds = \int_{\gamma_2} R(e - \overline{e}|_{\gamma_2}) ds + \int_{\gamma_2} R(\overline{e}|_{\gamma_2} - \overline{e}|_K) ds + \int_{\gamma_2} R(\overline{e}|_K - \mathcal{R}e) ds.$$

The first of these integrals can again be written as

$$\int_{\gamma_2} R(e - \overline{e}|_{\gamma_2}) ds = \bar{R}_2 \int_{\gamma_2} (e - \overline{e}|_{\gamma_2}) ds + \int_{\gamma_2} (R - \bar{R}_2)(e - \overline{e}|_{\gamma_2}) ds,$$

in which the first integral clearly vanishes. In the second integral we note that $(R - \bar{R}_2)$ is linear, and it is also an odd function centred at the midpoint of γ_2. In addition, it typically takes on small values since the over-refinement of the grid makes the boundary residuals in this direction small. The other part of the integrand, $(e - \overline{e}|_{\gamma_2})$, is *not* necessarily small valued. However, away from singularities the change in pollution error is typically small, or even of higher order, across each element. In that case this function is close to being an even function on γ_2, unless the double derivative of u along γ_2 changes sign. If it *is* an even function the integral will vanish. Again, this will happen if the true solution is a second order polynomial over the patch and the finite element solution is exact at the nodes, within a patchwise additional constant. Based on these observations we define

$$\phi_{\gamma_i} = \Big| \int_{\gamma_i} (R - \bar{R}_i)(e - \overline{e}|_{\gamma_i})\, ds \Big|,$$

which according to the above arguments can be assumed to be small for $\gamma_i = \gamma_2, \gamma_4$. In the anisotropic limit ϕ_{γ_i} will vanish for these two edges. With this definition and with the help of the over-refinement assumption and the patchwise uniformity assumption, we obtain

$$\begin{aligned} \Big| \int_{\gamma_2} R(e - \mathcal{R}e)\, ds \Big| &\le \phi_{\gamma_2} + ||R||_{\gamma_2} \big(||\overline{e}|_{\gamma_2} - \overline{e}|_K||_{\gamma_2} + ||\overline{e}|_K - \mathcal{R}e||_{\gamma_2} \big) \\ &\le \phi_{\gamma_2} + h_x^{1/2} (\epsilon_u + \epsilon_r) ||R||_{\gamma_2}. \qquad (5.7) \end{aligned}$$

Summarising the above results yields

$$\begin{aligned} B(e,e) \le \ & \sum_{K \in \mathcal{P}} \Big\{ h_x ||r - \bar{r}_K||_K + h_x^{1/2} \sum_{i=1,3} \Big(||\bar{R}_i||_{\gamma_i} + ||R - \bar{R}_i||_{\gamma_i} \Big) \Big\} |||e|||_{\tilde{\tilde{K}}} \\ & + \sum_{K \in \mathcal{P}} \Big\{ \epsilon_u |K|^{1/2} ||\bar{r}_K||_K + \sum_{i=2,4} \Big((\epsilon_u + \epsilon_r) h_x^{1/2} ||R||_{\gamma_i} + \phi_{\gamma_i} \Big) \Big\}. \end{aligned}$$

After applying the Cauchy-Schwarz inequality and for each element in the first sum adding the term

$$h_y\Big(||\bar{R}_y||^2_{\gamma_2} + ||\bar{R}_y||^2_{\gamma_4}\Big), \tag{5.8}$$

the following is obtained,

$$|||e||| \leq C\left\{\sum_{K\in\mathcal{P}}\left\{\eta^2_{R,K} + h_x^2||r-\bar{r}_K||^2_K + h_x\sum_{i=1,3}\Big(||\bar{R}_1-\bar{R}_3||^2_{\gamma_i} + ||R-\bar{R}_i||^2_{\gamma_i}\Big)\right\}\right\}^{1/2}$$
$$+\frac{1}{|||e|||}\sum_{K\in\mathcal{P}}\left\{\epsilon_u|K|^{1/2}||\bar{r}_K||_K + \sum_{i=2,4}\Big((\epsilon_u+\epsilon_r)h_x^{1/2}||R||_{\gamma_i} + \phi_{\gamma_i}\Big)\right\}. \tag{5.9}$$

Here the inequality

$$||\bar{R}_1||^2_{\gamma_1} + ||\bar{R}_3||^2_{\gamma_3} \leq C\sum_{i=1,3}||\bar{R}_x||^2_{\gamma_i} + ||\bar{R}_1 - \bar{R}_3||^2_{\gamma_i}$$

has been used. We once again point out that $h_x \gg h_y$ and that the side numbering from Figure 3.3 is applied. The extra term (5.8) was added only to restore the estimator. For an over-refined mesh this is typically only a higher order perturbation of the estimator, which is dominated by the term involving the boundary residuals across the short edges.

Except for the error estimator and the term $||\bar{R}_1 - \bar{R}_3||_\gamma$, all terms in the first line of the inequality (5.9) are approximation errors which can be assumed to be small. Provided δ is small, the term $||\bar{R}_1 - \bar{R}_3||_\gamma$ will be negligible if the finite element solution locally is close to being the nodal interpolant of *any* second order polynomial in the x-direction. As discussed previously, the terms ϕ_{γ_2} and ϕ_{γ_4} will vanish if the error is an even function along the two long edges, centred at the midpoint of each edge. The contributions from $||R||_{\gamma_2}$ and $||R||_{\gamma_4}$ are typically small since these boundary residuals vanish in the anisotropic limit (provided u is smooth and provided the finite element solution does not oscillate where u does not oscillate), in which case ϵ_r also vanishes. The term involving the average of the interior residual can be neglected only if ϵ_u is small. This clearly depends on the solution at hand, but as already mentioned, if u is well approximated by a second order polynomial over each patch, this term will be tiny, and if u is exactly a second order polynomial, it will vanish.

If the exact solution is a second order polynomial and $\delta = 0$, then (5.9) simplifies into

$$|||e||| \leq C\eta_R + \frac{1}{|||e|||}\sum_{K\in\mathcal{P}}\sum_{i=2,4}\epsilon_r h_x^{1/2}||R||_{\gamma_i},$$

and in the anisotropic limit the second term will vanish. This does not prove that the matching function is redundant, but it gives a strong indication of why the estimator will give good results, without the matching function, in typical situations.

To clarify what goes wrong with the standard analysis, which leads to the matching function, we will now repeat the analysis and plot the results of each of the most important estimates. This will clearly show what is lost in each step. The estimates will be performed on an example where the solution is a paraboloid, which, as already mentioned, should be a generic test case. Our results are therefore not general, but definitely typical. The true solution is chosen to be

$$u = x^2 + ay^2, \qquad a = 10^4,$$

leading to the following local error,

$$e|_K = \left(x^2 - \frac{h_x^2}{4}\right) + a\left(y^2 - \frac{h_y^2}{4}\right).$$

The problem is solved on the domain $\Omega = [0,1]^2$ with $\delta = 0$, and we investigate the behaviour of the different estimated terms when h_x is kept fixed and h_y is decreased. The parameter a is chosen to be large to make the plots as clear as possible. Based on Lemma 5.3 and Lemma 5.4, the following results are easily derived,

$$\begin{aligned} ||v - \bar{v}_K||_{\gamma_1} &\leq C\Big\{h_x^{1/2}||\nabla_x v||_{\tilde{\tilde{K}}} + \tfrac{h_y}{h_x^{1/2}}||\nabla_y v||_{\tilde{\tilde{K}}}\Big\} \\ ||v - \bar{v}_K||_{\gamma_2} &\leq C\Big\{\tfrac{h_x}{h_y^{1/2}}||\nabla_x v||_{\tilde{\tilde{K}}} + h_y^{1/2}||\nabla_y v||_{\tilde{\tilde{K}}}\Big\}. \end{aligned} \tag{5.10}$$

For the chosen example we start by rewriting the bilinear form in the following way,

$$\begin{aligned} B(e,e) &= \sum_{K\in\mathcal{P}} \int_K r(e - \mathcal{R}e)dx + \frac{1}{2}\sum_{\gamma\in\partial K}\int_\gamma R(e-\mathcal{R}e)ds \\ &= \sum_{K\in\mathcal{P}} \bar{R}_x \int_{\gamma_1}(e - \overline{e|}_K)ds + \bar{R}_y\int_{\gamma_2}(e - \overline{e|}_K)ds \\ &=: \psi_1, \end{aligned}$$

where we have used the fact that when the solution is a paraboloid, then $r|_K = \bar{r}_K$, $(\mathcal{R}e)|_K = \overline{e|}_K$, and the boundary residual contributions are pairwise equal. So far no estimates have been made, and the exact value for ψ_1, as h_y is decreased, is plotted in Figure 5.1. With the aid of the estimates (5.10), only considering the case when $h_x \geq h_y$, we proceed with the next steps in the

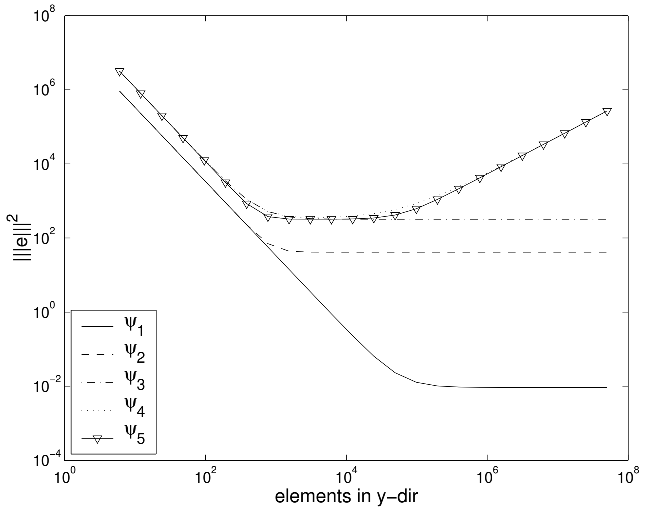

Figure 5.1: The exact expressions from each of the steps in the standard estimation procedure are plotted as h_x is kept fixed and h_y is decreased. Initially $h_x = h_y = 1/6$. The exact solution is $u = x^2 + ay^2$ with $a = 10^4$.

standard argument used by several authors,

$$\begin{aligned}
B(e,e) &\leq \sum_{K\in\mathcal{P}} ||\bar{R}_x||_{\gamma_1} ||e - \overline{e|}_K||_{\gamma_1} + ||\bar{R}_y||_{\gamma_2} ||e - \overline{e|}_K||_{\gamma_2} && =: \psi_2 \\
&\leq C \sum_{K\in\mathcal{P}} h_x^{1/2} ||\bar{R}_x||_{\gamma_1} \left(||\nabla_x e||_{\tilde{\tilde{K}}} + \tfrac{h_y}{h_x} ||\nabla_y e||_{\tilde{\tilde{K}}} \right) \\
&\qquad + h_y^{1/2} ||\bar{R}_y||_{\gamma_2} \left(\tfrac{h_x}{h_y} ||\nabla_x e||_{\tilde{\tilde{K}}} + ||\nabla_y e||_{\tilde{\tilde{K}}} \right) && =: C\psi_3 \\
&\leq C \sum_{K\in\mathcal{P}} \left(h_x^{1/2} ||\bar{R}_x||_{\gamma_1} + h_y^{1/2} ||\bar{R}_y||_{\gamma_2} \right) \\
&\qquad \cdot \left(\tfrac{h_x}{h_y} ||\nabla_x e||_{\tilde{\tilde{K}}} + ||\nabla_y e||_{\tilde{\tilde{K}}} \right) && =: C\psi_4 \\
&\leq C \Big\{ \sum_{K\in\mathcal{P}} h_x ||\bar{R}_x||_{\gamma_1}^2 + h_y ||\bar{R}_y||_{\gamma_2}^2 \Big\}^{1/2} \\
&\qquad \cdot \Big\{ \sum_{K\in\mathcal{P}} \tfrac{h_x^2}{h_y^2} ||\nabla_x e||_K^2 + ||\nabla_y e||_K^2 \Big\}^{1/2} \\
&= C\eta_R \, m(e,\mathcal{P}) |||e||| && =: C\psi_5
\end{aligned}$$

These steps are independent of the chosen exact solution. The exact values for $\psi_1, \ldots, \psi_5$ are plotted in Figure 5.1, and there it is clearly seen what is lost in each of the estimation steps. Already when norms are introduced ($\psi_1 \to \psi_2$) the information about the optimal aspect ratio is lost. The difference between where ψ_1 and ψ_2 are levelling out can be made arbitrarily large by changing the value of a. This is due to the fact that the estimate no longer takes into account that as h_y decreases, $\overline{e|}_K$ approaches $\overline{e|}_{\gamma_2}$, which causes the second integral in ψ_1 to approach zero, whereas $||e - \overline{e|}_K||_{\gamma_2}$ does *not* approach zero. The other dramatic step is when the residuals are separated from the properly weighted derivatives of the error ($\psi_3 \to \psi_4$). This causes the asymptotically dominant residual, $\bar{R}_x$, to be multiplied by the aspect ratio, h_x/h_y, which makes the term explode. This is the step which actually introduces the matching function, although it is defined as a global function after the last application of the Cauchy-Schwarz inequality. However, this last step ($\psi_4 \to \psi_5$) does not affect the asymptotic behaviour of the estimates. The step from ψ_2 to ψ_3 makes no difference except for a change in the constant, which means the estimates (5.10) are sharp.

This shows that the traditional way of finding an upper bound for the error cannot be expected to give reasonable results on anisotropic grids, and the explanation for this is *not* that (5.10) needs to be improved. Moreover, the reason why η_S and the new estimators give good results on anisotropic grids is *not* that the matching function remains bounded, but it has to do with the actual behaviour of the different terms that are present before going from ψ_1 to ψ_2, as was seen in the previous analysis. This also shows that the step $\psi_3 \to \psi_4$, which introduces the matching function or a matching condition, throws away all the remaining information on anisotropically over-refined grids.

5.2 Asymptotic exactness

In this section it will be proved that under certain conditions some of the estimators derived in Section 3 will be asymptotically exact independently of the aspect ratio of the elements. It will be assumed that the data are smooth such that

$$h_x, h_y \to 0 \quad \Rightarrow \quad |\bar{f}_K| \to C,$$

and

$$||f - \bar{f}_K||_K^2 \leq C h_+ h_x h_y.$$

Assuming the true solution is smooth, the second order polynomial u^* can be chosen such that in each element the difference between u^* and the true solution is, asymptotically, of third order, i.e.,

$$\max_K |\tilde{\epsilon}| \leq \mathcal{O}(h_x^3 + h_y^3). \tag{5.11}$$

The finite element solution will be referred to as *superconvergent* with respect to the second order approximation u^* if there exists a parameter $\tau > 0$ such that, without (5.11) being violated, for each element K

$$\begin{aligned} |u_X - I_X u^* - A_K|_{\tilde{K}} &\leq \left(h_x^{2+\tau} + h_y^{2+\tau}\right) C(u^*) \\ \hat{D}_x(\boldsymbol{\epsilon}') &\leq h_x^{1+\tau} C(u^*) \\ \hat{D}_y(\boldsymbol{\epsilon}') &\leq h_y^{1+\tau} C(u^*), \end{aligned}$$

where $I_X u^*$ is the nodal interpolant of u^* from the finite element space. If the finite element solution is superconvergent with respect to u^*, we directly obtain

$$\begin{aligned} |\epsilon_n'| &\leq \mathcal{O}\big(h_x^{2+\tau} + h_y^{2+\tau}\big) \\ \hat{D}_\gamma^2(\boldsymbol{\epsilon}') &\leq \mathcal{O}\big(h_\gamma^\tau\big) \\ \big|D_{\perp\gamma}^2\big(D_\gamma(\boldsymbol{\epsilon}')\big)\big| &\leq \mathcal{O}\big(h_{\perp\gamma}^\tau h_\gamma^{1+\tau}\big) \\ |D_{xy}(\boldsymbol{\epsilon}')| &\leq \mathcal{O}\big(h_x^{1+\tau} h_y^{1+\tau}\big) \\ |[A_\gamma]| &\leq \mathcal{O}\big(h_x^{2+\tau} + h_y^{2+\tau}\big) \\ |\bar{R}_\gamma| &\leq \mathcal{O}(h_{\perp\gamma}) \\ \max_\gamma \left| \frac{\partial u}{\partial n} - \left\langle \frac{\partial u_X}{\partial n} \right\rangle_\gamma \right| &\leq \mathcal{O}(h_{\perp\gamma}). \end{aligned}$$

Let

$$h_{max} = \max_{K\in\mathcal{P}} h_+,$$

where we recall that h_+ is a local quantity even if the subscript K is omitted, and note that for nontrivial cases we then have

$$C_1 h_{max}^2 \leq |||e|||^2 \leq C_2 h_{max}^2, \tag{5.12}$$

and

$$C_1 h_x h_y h_+^2 \leq |||e|||_K^2 \leq C_2 h_x h_y h_+^2. \tag{5.13}$$

Theorem 5.5 *Let the finite element mesh consist of rectangular bilinear elements and let assumption (A1) in Section 3 be satisfied. Suppose the true solution is smooth, and assume that the finite element solution is superconvergent with respect to u^*. Then the global estimators α_r^* and η_R^*, and the local estimator $\eta_{R,K}^*$ are asymptotically exact independently of the aspect ratio of the elements.*

If $\delta = 0$ this result holds also for the global estimators α_r and η_R, and for the local estimator $\eta_{R,K}$. (The estimators are defined in Section 3.)

PROOF. First, the statement regarding α_r^* is proved. With the aid of the previous results of this section and the superconvergence assumption, we estimate the quantities Q_0, Q_1, and Q_2 defined in (5.2),

$$\begin{aligned} Q_0^2 &\leq \sum_K C h_+^2 \Big\{ h_+ h_x h_y + \sum_{\gamma\subset\partial K} \mathcal{O}\big(h_\gamma^{5+2\tau} h_{\perp\gamma}^{1+2\tau}\big) \Big\} \\ &\leq \mathcal{O}(h_{max}^3) \\ Q_1 &\leq \sum_K \mathcal{O}\big(h_x^{2+\tau} + h_y^{2+\tau} + h_x^{2+\tau} h_y^{2+\tau}\big)\big(h_x h_y + \sum_{\gamma\subset\partial K} h_\gamma h_{\perp\gamma}\big) \\ &\leq \mathcal{O}\big(h_{max}^{2+\tau}\big), \end{aligned}$$

and

$$\begin{aligned} Q_2 &\leq C \sum_K \big(h_x h_y \max_K |\tilde{\epsilon}| + \sum_{\gamma \subset \partial K} h_\gamma h_{\perp\gamma} \max_K |\tilde{\epsilon}| \big) \\ &\leq \mathcal{O}\big(h_{max}^3\big). \end{aligned}$$

Moreover,

$$\begin{aligned} \sum_K \sum_{\gamma \subset \partial K} h_\gamma^{1/2} \Big\| \frac{\partial u}{\partial n} - \Big\langle \frac{\partial u_X}{\partial n} \Big\rangle_\gamma \Big\|_\gamma \, \big|[A_\gamma]\big| &\leq C \sum_K \sum_{\gamma \subset \partial K} h_\gamma h_{\perp\gamma} \mathcal{O}\big(h_x^{2+\tau} + h_y^{2+\tau}\big) \\ &\leq \mathcal{O}\big(h_{max}^{2+\tau}\big). \end{aligned}$$

From Theorem 5.1 and equation (5.12) we obtain

$$\big|\, |||e|||^2 - \alpha_r^* \,\big| \leq \mathcal{O}(h_{max}^{2+\tau} + h_{max}^{5/2}),$$

so after once again applying (5.12),

$$\left| \frac{\alpha_r^*}{|||e|||^2} - 1 \right| \leq \mathcal{O}(h_{max}^\tau + h_{max}^{1/2}),$$

which shows that α_r^* is asymptotically exact independently of the aspect ratio of the elements. If $\delta = 0$ then $\alpha_r^* = \alpha_r$, in which case α_r is also asymptotically exact.

Turning now to η_R^*, the foregoing results and Theorem 5.2 show that

$$\begin{aligned} \big|\, |||e|||_K - \eta_{R,K}^* \,\big| &\leq |K|^{1/2} \mathcal{O}\big(h_x^{1+\tau} + h_y^{1+\tau}\big) + |K|^{1/2} \mathcal{O}\big(h_x^3 + h_y^3\big) \\ &\leq |K|^{1/2} \mathcal{O}\big(h_+^{1+\tau} + h_+^3\big). \end{aligned}$$

Therefore, by the aid of (5.13),

$$\left| \frac{\eta_{R,K}^*}{|||e|||_K} - 1 \right| \leq \mathcal{O}(h_+^\tau + h_+^2),$$

which proves that $\eta_{R,K}^*$ is asymptotically exact. This immediately implies that η_R^* is asymptotically exact, and if $\delta = 0$ these estimators will be identical to $\eta_{R,K}$ and η_R, respectively. The theorem is hereby proved. □

Note that the results regarding $\eta_{R,K}$ and $\eta_{R,K}^*$ prove local asymptotic exactness on element level.

6 Finding optimal refinement directions

On the basis of the new estimators presented in Section 3 we will now derive optimal directions for local refinement. The result will be two indicators, x_{ref} and y_{ref}, that will vary between 0 and 1 and have the property

$$x_{ref} + y_{ref} = 1. \tag{6.1}$$

Assuming that the simplest versions of the error estimators are working properly, which means we assume both ϵ and δ to be negligible, the final indicators will carry the following information:

x_{ref} close to 1: Optimal to refine in the x-direction.

x_{ref} close to 0: Optimal to refine in the y-direction.

Else: There is no optimal direction. Refine uniformly.

The first step will be to derive expressions for how much the local estimators will decrease if the element is split into p parts in either the x- or the y-direction. Typically a refinement means that $p = 2$, but as will be seen, p will not be present in the final indicators. Before deriving the indicators it is worth noting that, according to (3.7), the boundary residuals will be reduced by the same factor as the corresponding mesh sizes.

Another point worth noticing, is that standard differentiation will not give a useful result. This is due to the fact that splitting an element in the x-direction would not give the same change in h_x as splitting it in the y-direction would do to h_y, if the two length scales were not initially the same.

Define d_x as the difference between the estimator on the current element and the estimator we would obtain on an x-refined element with new length scales h_x/p. The latter element estimator must then be multiplied by p to account for the smaller area of the new elements. Let d_y be defined similarly for y-refinement. Then

$$\begin{aligned} d_x(\alpha_{r,K}) &\stackrel{def}{=} \alpha_{r,K}(h_x, h_y, \bar{R}_x, \bar{R}_y) - p\,\alpha_{r,K}(\frac{h_x}{p}, h_y, \frac{\bar{R}_x}{p}, \bar{R}_y) \\ &= -\frac{p^2-1}{12p^2} h_x^2 \bar{R}_x (r h_y + \bar{R}_y), \end{aligned} \tag{6.2}$$

and

$$\begin{aligned} d_y(\alpha_{r,K}) &= -\frac{p^2-1}{12p^2} h_y^2 \bar{R}_y (\bar{r}_K h_x + \bar{R}_x) \\ d_x(\eta^2_{R,K}) &= \frac{p^2-1}{12p^2}\, h_x h_y \bar{R}_x^2 \\ d_y(\eta^2_{R,K}) &= \frac{p^2-1}{12p^2}\, h_x h_y \bar{R}_y^2. \end{aligned}$$

If these expressions are compared to the estimators, we find that by letting p tend to infinity we completely remove the contributions associated with the direction of refinement. Now define the direction indicators to be

$$\begin{aligned} x_{ref} &= \frac{|d_x|}{|d_x| + |d_y|} \\ y_{ref} &= \frac{|d_y|}{|d_x| + |d_y|} = 1 - x_{ref}. \end{aligned} \tag{6.3}$$

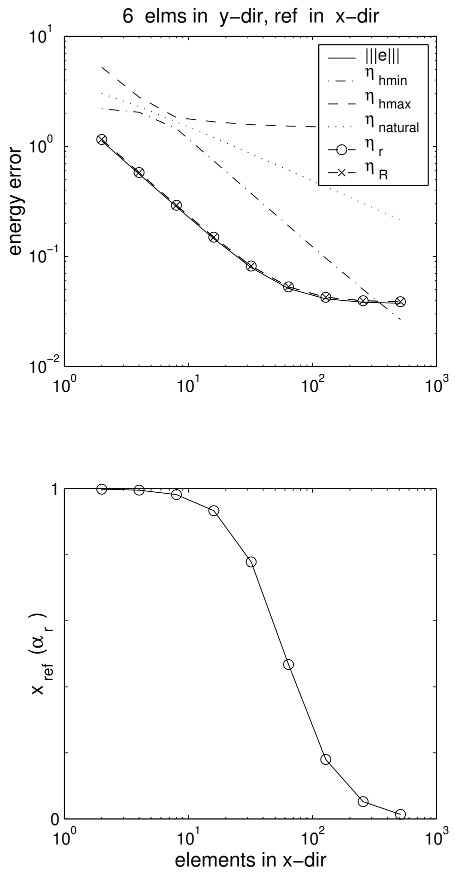

Figure 6.1: The direction indicator for the test case $u = x^2 + ay^2$; $a = 1/10$. The upper plot shows the actual error and the different error estimators. The lower plot shows the corresponding direction indicator. In this case $\alpha_r \equiv \eta_r^2$, $x_{ref}(\alpha_{r,K}) \equiv x_{ref}(\eta_{R,K}^2)$, and the indicators are identical for all elements.

From these definitions we obtain the following direction indicators for the two estimators,

$$\begin{aligned} x_{ref}(\alpha_{r,K}) &= \frac{1}{1+\beta_r} \\ x_{ref}(\eta_{R,K}^2) &= \frac{1}{1+\beta_R}, \end{aligned}$$

where

$$\beta_r = \frac{|h_y^2 \bar{R}_y(\bar{r}_K h_x + \bar{R}_x)|}{|h_x^2 \bar{R}_x(\bar{r}_K h_y + \bar{R}_y)|}$$

and

$$\beta_R = \frac{\bar{R}_y^2}{\bar{R}_x^2}.$$

The y-indicators are found from the relation (6.3). Figure 6.1 shows a plot of the x-indicator resulting from the test where the solution is a paraboloid. Then the direction indicators are equal for all elements, and the plot shows that as the error is levelling out the indicator quickly drops from indicating the x-direction as the optimal direction for refinement to indicating that the error will not decrease if further x-refinement is performed.

For the η_R indicator this strategy suggests that the element aspect ratio is optimal when

$$\bar{R}_x^2 = \bar{R}_y^2. \tag{6.4}$$

When the solution locally is a second order polynomial and the pollution error does not increase across the element the estimator is exact, so in that case this is the optimal strategy. However, this is not the same strategy as previously has been suggested, where the aim is to produce a grid which minimises a discrete counterpart of the matching function. In the case where η_R is exact, that strategy would strive to satisfy

$$h_x^2 \bar{R}_x^2 = h_y^2 \bar{R}_y^2 \tag{6.5}$$

on each element. This condition suggests a grid which has less anisotropy than a grid satisfying (6.4). According to equations (3.7) and (3.8) the criterion (6.5) suggests an aspect ratio of $\sqrt{\rho_{opt}}$, where ρ_{opt} is the true optimal aspect ratio satisfying (6.4). The criterion (6.5) suggests the grid is optimal approximately when $\eta_R m(e, \mathcal{P})$ hits its minimum in Figure 4.2, and when ψ_5 changes from having a negative slope to being constant (number of elements in y-direction ≈ 600) in Figure 5.1. The criterion (6.4) is satisfied at the point where the true error is levelling out in Figure 4.2, and where ψ_1 is levelling out in Figure 5.1. Note that until the element aspect ratio satisfying (6.4) is reached, any refinement in the non-optimal direction would only give a negligible decrease in the error, i.e., it would be inefficient use of degrees of freedom. However, once a matching condition is introduced, the natural choice for an optimal grid is to have (6.5) satisfied, see [11].

7 Reaction-diffusion equation

The new estimators developed in the previous sections can easily be modified to estimate the error in a reaction-diffusion problem of the type

$$\begin{aligned} -\Delta u + \kappa^2 u &= f \quad \text{in } \Omega \\ u &= 0 \quad \text{on } \partial\Omega, \end{aligned} \tag{7.1}$$

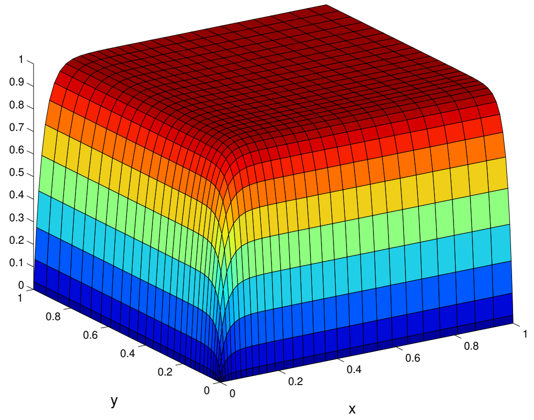

Figure 7.1: The solution to the reaction-diffusion test example. The problem considered is: $-\Delta u + \kappa^2 u = 1, \quad \kappa^2 = 10^3$.

where κ is a constant. The problem has the following weak formulation: Find $u \in V_0$ such that

$$B^*(u, v) = L(v) \qquad \forall v \in V_0 \ , \tag{7.2}$$

where

$$B^*(u, v) \quad = \quad \int_\Omega \nabla u \cdot \nabla v + \kappa^2 uv \, dx \ .$$

As previously, the bilinear form defines an energy norm for a function $v \in V_0$ in the following way,

$$|||v|||^2 = B^*(v, v).$$

Note that this norm includes not only the derivatives of the function, but also the actual values. Hence, this choice of norm should not be very well suited for the new estimators since the reconstructed error vanishes at the nodes. For the $R-$estimators we would therefore expect better results if instead the H^1-semi-norm was chosen. However, only the above energy norm will be used since it is the natural norm for this problem. The $r-$estimators depend on the residual equation, so for those estimators one can not easily replace this natural norm. The interior residual is now defined as

$$r = f + \Delta u_X - \kappa^2 u_X \qquad \text{in } K.$$

With this definition of r and thereby also a new definition of $\bar{r}_K$, the estimator $\eta_{r,K}$ remains as in equation (3.15). Also the standard estimators $\eta_{h_{max}}, \eta_{h_{min}}$, and $\eta_{natural}$ remain unchanged, modulus the new definition of r.

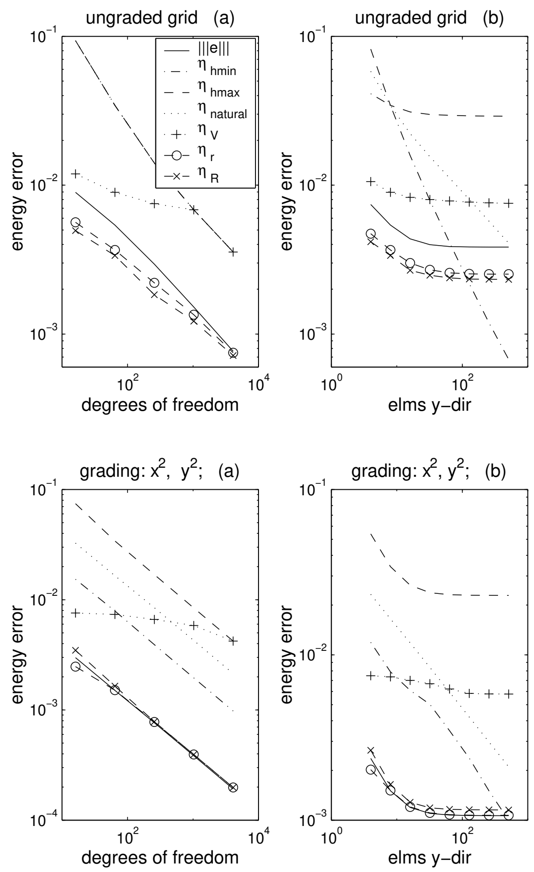

Figure 7.2: The results for the reaction-diffusion test case. Uniform refinement is marked with (a) and refinement in only one direction is marked with (b). Two of the cases start with an ungraded grid, whereas the two others start with an initially well graded grid.

The estimator $\eta_{R,K}$ must, however, be modified since it is defined as

$$\eta_{R,K}^2 = B^*(e_K, e_K),$$

where e_K is the reconstructed error obtained when neglecting the changes in the grid. A simple calculation reveals that this estimator now takes the form

$$\eta_{R,K}^2 = \frac{h_x h_y}{12}\left\{\bar{R}_x^2 + \bar{R}_y^2 + \frac{\kappa^2}{10}\left((h_x\bar{R}_x + h_y\bar{R}_y)^2 - \frac{1}{3}h_x h_y \bar{R}_x \bar{R}_y\right)\right\}.$$

In addition to the standard estimators we also compare the new estimators to an estimator, suggested by Verfürth [13], which is especially tailored towards a reaction-diffusion equation. This estimator can be expressed using the general estimator defined in equation (2.7). Let s_K be defined by

$$s_K = \min(\kappa^{-1}\ ,\ h_+)\ ,$$

where, as previously, h_+ is the length of the longest edge of the element. Then the local estimator can be expressed as

$$\eta^2_{V,K} = \eta^2_K(s_K, s_K, s_K), \tag{7.3}$$

with the global estimator defined by (2.8). It is obvious that this estimator differs from $\eta_{h_{max}}$ only for h_+ satisfying

$$\kappa h_+ > 1\ ,$$

so asymptotically it is identical to $\eta_{h_{max}}$. The original estimator suggested in [13] actually has $min(\kappa^{-1}, \mathbf{h}_{edge})$ as the last parameter in (7.3), but on anisotropic grids the above estimator appears to be the natural choice. In the test problem the following parameters are chosen,

$$\kappa^2 = 10^3 \qquad f = 1.$$

The problem is solved on a square with homogeneous Dirichlet boundary conditions, but because of symmetry we choose to consider only one quarter of the domain. This problem has no known analytical solution, but the finite element solution is shown in Figure 7.1. Note that there are rather sharp boundary layers present, and that the assumption $\epsilon = \delta = 0$, which would make the estimators exact, will be far from satisfied especially on the initial grid.

Four different tests are performed. First, starting with a uniform grid, we refine the grid both uniformly and only in one direction. Thereafter, the initial grid is taken to be a graded grid, where the grading function is x^2 for the x-direction and y^2 for the y-direction. This gives a relatively well adapted initial grid for the current problem. Just as was done for the uniform grid, we refine it both uniformly and only in one direction. The resulting plots are shown in Figure 7.2, where the uniformly refined cases are marked with (a) and refinement in only one direction is marked with (b). The plots show that the new estimators are very close to estimating the actual error correctly in all the cases, while most of the standard estimators give an erroneous picture of the true error. The special estimator, η_V, has a smaller initial constant than $\eta_{h_{max}}$, but in most of the cases it displays an incorrect convergence rate, until it finally becomes identical to $\eta_{h_{max}}$ when the mesh size is small enough.

8 Conclusion

Through analysis and numerical examples it has been shown that explicit error estimators can be constructed which in all test cases show promising results, even on extremely graded meshes, and which under certain conditions are asymptotically exact independently of the aspect ratio of the elements. Whenever the solution can be locally well approximated by a second order polynomial and the pollution error is under control the estimators will give good results. The estimators carry information about the optimal refinement directions, and our analytical findings contradict with anisotropic refinement algorithms previously suggested. It has also been demonstrated how error bounds derived within the standard framework degenerate in the anisotropic limit. Finally, it is shown that the new estimators are not limited to the test problems considered here, in fact some of the estimators, the R-estimators, have no connection to the actual differential equation. They only depend on the reconstructed error and the chosen error norm.

Acknowledgements

The author would like to thank professor Mark Ainsworth at the University of Strathclyde, for many constructive suggestions and comments during the work with this article.

References

[1] M.Ainsworth and J.T.Oden. *A posteriori error estimation in finite element analysis.* John Wiley & Sons Inc., New York, 2000.

[2] T.Apel. *Anisotropic finite elements: Local estimates and applications.* B.G.Teubner (Stuttgart, Leipzig), 1999.

[3] C.Carstensen and R.Verfürth. *Edge residuals dominate a posteriori error estimates for low order finite element methods.* SIAM J. Numer. Anal. **36**, (5):1571–1587, 1999.

[4] M.Dobrowolski, S.Graf, and C.Pflaum. *On a posteriori estimators in the finite element method on anisotropic meshes.* Elect. Trans. Num. Anal., **8**, 36–45, 1999.

[5] C.Johnson and P.Hansbo. *Adaptive finite element methods in computational mechanics.* Comput. Methods Appl. Mech. Engrg., **101**, (1-3):143–181, 1992.

[6] G.Kunert. *A posteriori error estimation for anisotropic tetrahedral and triangular finite element meshes.* PhD thesis, TU Chemnitz, 1999.

[7] G.Kunert. *An a posteriori residual error estimator for the finite element method on anisotropic tetrahedral meshes.* Numer. Math., **86**, (3):471–490, 2000.

[8] G.Kunert and R.Verfürth. *Edge residuals dominate a posteriori error estimates for linear finite element methods on anisoropic triangular and tetrahedral meshes.* Numer. Math., **86**, (2):283–303, 2000.

[9] J.Peraire, M.Vahdati, K.Morgan, and O.C.Zienkiewicz. *Adaptive remeshing for compressible flows.* J. Comp. Phys. **72**, 449–466, 1987.

[10] W.Rachowicz. *An anisotropic h-type mesh-refinement strategy.* Comp. Meth. in App. Mech. Eng., **109**, (1–2):169–181, 1993.

[11] K.G.Siebert. *An a posteriori estimator for anisotropic refinement.* Numer. Math., **73**, (3):373–398, 1996.

[12] R.Verfürth. *A review of a posteriori error estimation and adaptive mesh-refinement techniques.* Wiley and Teubner, Germany, 1996.

[13] R.Verfürth. *Robust a posteriori error estimators for a singularly perturbed reaction-diffusion equation.* Numer. Math., **78**, (3):479–493, 1998.

A Proofs for Theorem 5.1 and Theorem 5.2

A.1 Finite difference notation

First, we introduce some standard finite difference notation. Consider two neighbouring elements K and K' with a numbering of nodes and edges as shown in Figure A.1, and let $\mathbf{v}$ be a vector with one entry associated with each of the nodes in the patch $\tilde{K}$ surrounding element K. Define the first order difference along an edge γ with endpoints p_1 and p_2 as

$$D_\gamma(\mathbf{v}) = \frac{\mathbf{v}(p_2) - \mathbf{v}(p_1)}{h_\gamma} \qquad \text{e.g.} \qquad D_1(\mathbf{v}) = \frac{\mathbf{v}(e) - \mathbf{v}(b)}{h_y} \, .$$

It will be irrelevant whether forward or backward differences are used as long as it is done consistently within each patch. Define the average first order difference in the x-direction for element K as

$$D_x(\mathbf{v}) \quad = \quad \frac{D_2(\mathbf{v}) + D_4(\mathbf{v})}{2} \, , \tag{A.1}$$

and similarly for the y-direction.

Define the second order difference, centred at p, across an edge γ, using the additional points p_+ and p_-, as

$$D^2_{\perp\gamma,p}(\mathbf{v}) = \frac{\mathbf{v}(p_-) - 2\mathbf{v}(p) + \mathbf{v}(p_+)}{h^2_{\perp\gamma}} \qquad \text{e.g.} \qquad D^2_{\perp 1,b}(\mathbf{v}) = \frac{\mathbf{v}(a) - 2\mathbf{v}(b) + \mathbf{v}(c)}{h^2_x} \, .$$

The average second order difference across an edge γ with endpoints p_1 and p_2 is now naturally defined as

$$D^2_{\perp\gamma}(\mathbf{v}) = \frac{D^2_{\perp\gamma,p_1}(\mathbf{v}) + D^2_{\perp\gamma,p_2}(\mathbf{v})}{2} \qquad \text{e.g.} \qquad D^2_{\perp 1}(\mathbf{v}) = \frac{D^2_{\perp 1,b}(\mathbf{v}) + D^2_{\perp 1,e}(\mathbf{v})}{2}.$$

The second order difference across an edge may also take a first order difference along the same edge as its argument, and it is then to be understood in the sense

$$D^2_{\perp\gamma}\big(D_\gamma(\mathbf{v})\big) = \frac{D_{\gamma_-}(\mathbf{v}) - 2D_\gamma(\mathbf{v}) + D_{\gamma_+}(\mathbf{v})}{h^2_{\perp\gamma}},$$

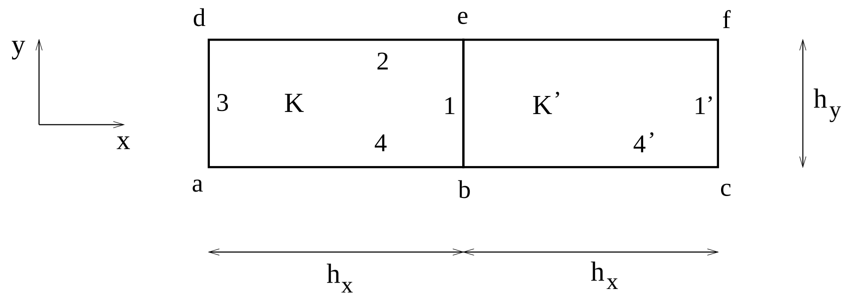

Figure A.1: Neighbouring elements K and K'.

e.g.

$$D^2_{\perp 1}\big(D_1(\mathbf{v})\big) = \frac{D_3(\mathbf{v}) - 2D_1(\mathbf{v}) + D_{1'}(\mathbf{v})}{h_x^2}.$$

The second order cross-term difference for element K in Figure A.1 is defined as

$$D_{xy}(\mathbf{v}) = \frac{\mathbf{v}(e) - \mathbf{v}(d) + \mathbf{v}(a) - \mathbf{v}(b)}{h_x h_y}.$$

We will later be needing the following results, here stated only for certain edges or pair of edges in Figure A.1, but with a natural generalisation to other edges:

$$\begin{aligned} D^2_{\perp 1,e}(\mathbf{v}) - D^2_{\perp 1,b}(\mathbf{v}) &= \frac{\mathbf{v}(d) - 2\mathbf{v}(e) + \mathbf{v}(f) - \mathbf{v}(a) + 2\mathbf{v}(b) - \mathbf{v}(c)}{h_x^2} \\ &= h_y \frac{D_3(\mathbf{v}) - 2D_1(\mathbf{v}) + D_{1'}(\mathbf{v})}{h_x^2} \\ &= h_y D^2_{\perp 1}\big(D_1(\mathbf{v})\big)\,, \end{aligned} \tag{A.2}$$

and

$$\begin{aligned} D_2(\mathbf{v}) - D_4(\mathbf{v}) &= \frac{\mathbf{v}(e) - \mathbf{v}(d) - \mathbf{v}(b) + \mathbf{v}(a)}{h_x} \\ &= h_y D_{xy}(\mathbf{v}). \end{aligned} \tag{A.3}$$

Suppose now that the edge length in the x-direction changes from element K to element K' in Figure A.1. As previously, let the edge length for K be h_x and let the edge length for K' be $(1 + k_1\delta)h_x$. With this notation we obtain

$$\begin{aligned} \frac{\mathbf{v}(b) - \mathbf{v}(a)}{h_x} - \frac{\mathbf{v}(c) - \mathbf{v}(b)}{(1 + k_1\delta)h_x} &\equiv D_4(\mathbf{v}) - D_{4'}(\mathbf{v}) \\ &= \frac{\mathbf{v}(b) - \mathbf{v}(a) - \mathbf{v}(c) + \mathbf{v}(b) + k_1\delta\big(\mathbf{v}(b) - \mathbf{v}(a)\big)}{(1 + k_1\delta)h_x} \\ &= \frac{-h_x D^2_{\perp 1,b}(\mathbf{v}) + k_1\delta D_4(\mathbf{v})}{1 + k_1\delta}\,, \end{aligned} \tag{A.4}$$

with similar results for the other sides of element K.

A.2 Preliminary results

Let $u, u^*, u_X, \tilde{\epsilon}$, and $\boldsymbol{\epsilon}'$ be as defined in Section 3. Then the true error on element K can be expressed as

$$\begin{aligned} e|_K &= u - u_X \\ &= u^* + \tilde{\epsilon} - u_X \\ &= c_1 + c_2 x + c_3 y + c_4 xy + a_5 x^2 + a_6 y^2 + \tilde{\epsilon}, \end{aligned}$$

where a_5x^2 and a_6y^2 are terms occurring in u^*, and where c_i are unknown coefficients. We will now determine the coefficients in this expression. Using the numbering in Figure 3.3, the error in node 1 can be expressed as

$$e(\frac{h_x}{2}, \frac{h_y}{2}) = u|_{n_1} - u_X|_{n_1} = u^*|_{n_1} - (u^*|_{n_1} + A_K + \epsilon'_1) \ ,$$

so

$$c_1 + c_2\,\frac{h_x}{2} + c_3\,\frac{h_y}{2} + c_4\,\frac{h_xh_y}{4} + a_5\,\frac{h_x^2}{4} + a_6\,\frac{h_y^2}{4} = -A_K - \epsilon'_1. \qquad \text{(A.5)}$$

Similar expressions for the other three corners give a total of four linear equations for the coefficients. These equations can be solved for $c_2, \ldots, c_4$ separately, and using the numbering from Figure 3.3, the solution is

$$c_2 = -\frac{\epsilon'_1 - \epsilon'_2 + \epsilon'_4 - \epsilon'_3}{2h_x} \qquad c_3 = -\frac{\epsilon'_1 - \epsilon'_4 + \epsilon'_2 - \epsilon'_3}{2h_y} \qquad c_4 = -\frac{\epsilon'_1 - \epsilon'_2 + \epsilon'_3 - \epsilon'_4}{h_xh_y}.$$

Applying the notation from the previous section, this may be expressed as

$$c_2 = -D_x(\boldsymbol{\epsilon}') \qquad c_3 = -D_y(\boldsymbol{\epsilon}') \qquad c_4 = -D_{xy}(\boldsymbol{\epsilon}'). \qquad \text{(A.6)}$$

Solving the linear system for c_1, simply by moving all other terms from equation (A.5) to the other side and recalling that $\epsilon'_1 = 0$, gives

$$c_1 = -\frac{1}{4}\big(a_5h_x^2 + a_6h_y^2\big) - \frac{1}{2}\Big(h_xD_x(\boldsymbol{\epsilon}') + h_yD_y(\boldsymbol{\epsilon}') + \frac{h_xh_y}{2}D_{xy}(\boldsymbol{\epsilon}')\Big) - A_K.$$

Consider the elements K and K' in Figure A.1, but let the element length scale in x-direction vary from K to K' such that

$$h_{x,K'} = (1 + k_1\delta)h_{x,K} \equiv (1 + k_1\delta)h_x.$$

Let node b have coordinates (x_0, y_0). Applying the notation from Section A.1, the boundary residual across edge 1 in node b is

$$\begin{aligned} R_{1,b} &= -\left(\frac{u_X(x_0,y_0) - u_X(x_0 - h_x, y_0)}{h_x} - \frac{u_X\big(x_0 + (1+k_1\delta)h_x, y_0\big) - u_X(x_0,y_0)}{(1+k_1\delta)h_x}\right) \\ &= -\Bigg(\frac{u^*(x_0,y_0) - u^*(x_0 - h_x, y_0)}{h_x} - \frac{u^*\big(x_0 + (1+k_1\delta)h_x, y_0\big) - u^*(x_0,y_0)}{(1+k_1\delta)h_x} \\ &\qquad + D_4(\boldsymbol{\epsilon}') - D_{4'}(\boldsymbol{\epsilon}')\Bigg) \\ &= (2 + k_1\delta)a_5h_x + \frac{h_xD^2_{\perp 1,b}(\boldsymbol{\epsilon}') - k_1\delta D_4(\boldsymbol{\epsilon}')}{1 + k_1\delta} \ , \end{aligned}$$

where we have used equation (A.4) and the fact that u^* has the form (3.1). Similarly,

$$R_{1,e} \ = \ (2 + k_1\delta)a_5h_x + \frac{h_xD^2_{\perp 1,e}(\boldsymbol{\epsilon}') - k_1\delta D_2(\boldsymbol{\epsilon}')}{1 + k_1\delta} \ .$$

Since the elements are bilinear, the boundary residuals are linear along each edge, so

$$\begin{aligned} \bar{R}_1 &= \frac{R_{1,b} + R_{1,e}}{2} \\ &= (2 + k_1\delta)a_5 h_x + \frac{h_x D^2_{\perp 1}(\boldsymbol{\epsilon}') - k_1\delta D_x(\boldsymbol{\epsilon}')}{1 + k_1\delta} \ . \end{aligned} \tag{A.7}$$

For the same reason we also obtain

$$\begin{aligned} \max_{\gamma_1} |R - \bar{R}_1| &\le |R_{1,e} - R_{1,b}| \\ &\le \frac{h_x \left|D^2_{\perp 1,e}(\boldsymbol{\epsilon}') - D^2_{\perp 1,b}(\boldsymbol{\epsilon}')\right| + \delta k_1 \left|D_2(\boldsymbol{\epsilon}') - D_4(\boldsymbol{\epsilon}')\right|}{1 + k_1\delta} \\ &\le C\Big(h_x h_y \left|D^2_{\perp 1}\big(D_1(\boldsymbol{\epsilon}')\big)\right| + \delta h_y \left|D_{xy}(\boldsymbol{\epsilon}')\right|\Big) \ , \end{aligned} \tag{A.8}$$

where (A.2) and (A.3) have been used. Solving (A.7) for a_5 yields,

$$a_5 = \frac{\bar{R}_1}{(2 + k_1\delta)h_x} - \frac{h_x D^2_{\perp 1}(\boldsymbol{\epsilon}') - k_1\delta D_x(\boldsymbol{\epsilon}')}{(1 + k_1\delta)(2 + k_1\delta)h_x} \ . \tag{A.9}$$

Similarly, we can find a_5 from the jumps across edge 3, so

$$a_5 = \frac{\bar{R}_3}{(2 + k_3\delta)h_x} - \frac{h_x D^2_{\perp 3}(\boldsymbol{\epsilon}') - k_3\delta D_x(\boldsymbol{\epsilon}')}{(1 + k_3\delta)(2 + k_3\delta)h_x} \ . \tag{A.10}$$

We define the approximate coefficients $\tilde{a}_5$, $\tilde{a}_6$, and $\tilde{c}_1$ as

$$\begin{aligned} \tilde{a}_5 &= \frac{1}{2h_x}\Big(\frac{\bar{R}_1}{2 + k_1\delta} + \frac{\bar{R}_3}{2 + k_3\delta}\Big) \\ \tilde{a}_6 &= \frac{1}{2h_y}\Big(\frac{\bar{R}_2}{2 + k_2\delta} + \frac{\bar{R}_4}{2 + k_4\delta}\Big) \\ \tilde{c}_1 &= -\frac{1}{4}(\tilde{a}_5 h_x^2 + \tilde{a}_6 h_y^2). \end{aligned}$$

Then, from (A.9) and (A.10) we obtain

$$|a_5 - \tilde{a}_5| \le C\Big(\hat{D}^2_x(\boldsymbol{\epsilon}') + \frac{\delta}{h_x}\,\hat{D}_x(\boldsymbol{\epsilon}')\Big), \tag{A.11}$$

and similarly

$$|a_6 - \tilde{a}_6| \le C\Big(\hat{D}^2_y(\boldsymbol{\epsilon}') + \frac{\delta}{h_y}\,\hat{D}_y(\boldsymbol{\epsilon}')\Big). \tag{A.12}$$

Furthermore,

$$\begin{aligned} |c_1 + A_K - \tilde{c}_1| \le \ & C\Big\{|a_5 - \tilde{a}_5|h_x^2 + |a_6 - \tilde{a}_6|h_y^2 \\ & + h_x\hat{D}_x(\boldsymbol{\epsilon}') + h_y\hat{D}_y(\boldsymbol{\epsilon}') + h_x h_y \hat{D}_{xy}(\boldsymbol{\epsilon}')\Big\}. \end{aligned}$$

We now define the reconstructed error, e_K, as

$$e_K = \tilde{c}_1 + \tilde{a}_5 x^2 + \tilde{a}_6 y^2,$$

which then gives the following upper bound,

$$\begin{aligned}\Big| e|_K + A_K - e_K \Big| &\leq C\Big\{ |c_1 + A_K - \tilde{c}_1| + |h_x c_2 + h_y c_3 + h_x h_y c_4| \\ &\qquad + |a_5 - \tilde{a}_5| h_x^2 + |a_6 - \tilde{a}_6| h_y^2 + |\tilde{\epsilon}| \Big\} \\ &\leq C\Big\{ \phi_{r,K}(\boldsymbol{\epsilon}') + |\tilde{\epsilon}| \Big\}, \qquad \text{(A.13)}\end{aligned}$$

where $\phi_{r,K}(\boldsymbol{\epsilon}')$ was defined in (5.1).

A.3 Proof for Theorem 5.1

In Section 3 the pollution error in node 1 of element K was denoted by A_K, and the fact that A_K is constant on K allows us to write

$$0 = \sum_{K \in \mathcal{P}} B_K(e, A_K) = \sum_{K \in \mathcal{P}} \int_K r A_K \, dx + \sum_{\gamma \in \partial K} \int_\gamma \frac{\partial e}{\partial n_K} A_K \, ds.$$

For real numbers a, b, c, d we have the identity

$$ab + cd = \frac{1}{2}\Big((a+c)(b+d) + (a-c)(b-d) \Big).$$

Hence, for the neighbouring elements K and K' sharing the edge γ,

$$\frac{\partial e}{\partial n_K}\bigg|_K A_K + \frac{\partial e}{\partial n_{K'}}\bigg|_{K'} A_{K'} = \frac{1}{2}\left\{ R(A_K + A_{K'}) + 2\left\langle \frac{\partial e}{\partial n_K} \right\rangle_\gamma [A_\gamma] \right\},$$

where the notation from Section 5 has been applied. Consequently,

$$\begin{aligned} 0 &= \sum_{K \in \mathcal{P}} B_K(e, A_K) \\ &= \sum_{K \in \mathcal{P}} \Big\{ \int_K r A_K \, dx + \frac{1}{2} \sum_{\gamma \in \partial K} \int_\gamma R \, A_K \, ds \Big\} + \sum_{\gamma \in \partial \mathcal{P}} \int_\gamma \left\langle \frac{\partial e}{\partial n_K} \right\rangle_\gamma [A_\gamma] \, ds. \end{aligned}$$

We now have the necessary results and notation to prove Theorem 5.1. Since ϵ is allowed to be nonzero, equation (3.5) is no longer valid because r may not be constant on each element, and R may not be constant on each edge. Instead of using (3.5) we therefore rewrite (2.6) as

$$
\begin{aligned}
B(e,e) = \sum_{K\in\mathcal{P}} \Big\{ & \bar{r}_K \int_K e_K\,dx + \frac{1}{2}\sum_{\gamma\in\partial K} \bar{R}_\gamma \int_\gamma e_K\,ds \\
& + \bar{r}_K \int_K (e + A_K - e_K)\,dx + \frac{1}{2}\sum_{\gamma\in\partial K} \bar{R}_\gamma \int_\gamma (e + A_K - e_K)\,ds \\
& + \int_K (r - \bar{r}_K)(e + A_K)\,dx + \frac{1}{2}\sum_{\gamma\in\partial K}\int_\gamma (R - \bar{R}_\gamma)(e + A_K)\,ds \Big\} \\
& + \sum_{\gamma\in\partial\mathcal{P}} \int_\gamma \left\langle \frac{\partial e}{\partial n_K}\right\rangle_\gamma [A_\gamma]\,ds, \qquad (A.14)
\end{aligned}
$$

where e_K is the reconstructed error defined in (A.13). The first two terms in (A.14) form the estimator $\alpha^*_{r,K}$ defined in (3.13).

The third last term in (A.14) is easily bounded by applying the Cauchy-Schwarz and the Poincaré inequalities,

$$
\begin{aligned}
\int_K (r - \bar{r}_K)(e + A_K)\,dx &= \int_K (r - \bar{r}_K)\big(e - \overline{e|}_K\big)\,dx \\
&\le ||r - \bar{r}_K||_K ||e - \overline{e|}_K||_K \\
&\le Ch_+ ||r - \bar{r}_K||_K\, |||e|||_K.
\end{aligned}
$$

Similarly, with the aid of Lemma 5.3,

$$
\begin{aligned}
\int_\gamma (R - \bar{R}_\gamma)(e + A_K)ds &= \int_\gamma (R - \bar{R}_\gamma)\big(e - \overline{e|}_K\big)ds \\
&\le ||R - \bar{R}_\gamma||_\gamma\, ||e - \overline{e|}_K||_\gamma \\
&\le C||R - \bar{R}_\gamma||_\gamma \Big\{ h_{\perp\gamma}^{-1/2} ||e - \overline{e|}_K||_K + h_{\perp\gamma}^{1/2} ||\nabla_{\perp\gamma} e||_K \Big\} \\
&\le C||R - \bar{R}_\gamma||_\gamma \Big\{ \frac{h_+}{h_{\perp\gamma}^{1/2}} + h_{\perp\gamma}^{1/2} \Big\} |||e|||_K \\
&\le C\frac{h_+}{h_{\perp\gamma}^{1/2}} ||R - \bar{R}_\gamma||_\gamma\, |||e|||_K.
\end{aligned}
$$

To bound the third and fourth term in (A.14), we need the following two results obtained simply by integrating (A.13) and applying the Cauchy-Schwarz inequality,

$$
\begin{aligned}
\bar{r}_K \int_K (e + A_K - e_K)\,dx &\le C\big(h_x h_y |\bar{r}_K| \phi_{r,K}(\boldsymbol{\epsilon}') + ||\bar{r}_K||_K\, ||\tilde{\epsilon}||_K\big) \\
\bar{R}_\gamma \int_\gamma (e + A_K - e_K)\,ds &\le C\big(h_\gamma |\bar{R}_\gamma| \phi_{r,K}(\boldsymbol{\epsilon}') + ||\bar{R}_\gamma||_\gamma\, ||\tilde{\epsilon}||_\gamma\big).
\end{aligned}
$$

Summarising and again applying the Cauchy-Schwarz inequality to the last term in (A.14) and to the terms involving $|||e|||_K$, we conclude that

$$\begin{aligned}|B(e,e)-\alpha_r^*| \quad \leq \quad & C\Big\{|||e|||\Big(\sum_{K\in\mathcal{P}} h_+^2||r-\bar{r}_K||_K^2+\sum_{\gamma\in\partial K}\frac{h_+^2}{h_{\perp\gamma}}||R-\bar{R}_\gamma||_\gamma^2\Big)^{1/2} \\ & +\sum_{K\in\mathcal{P}}\phi_{r,K}(\boldsymbol{\epsilon}')\Big(h_xh_y|\bar{r}_K|+\sum_{\gamma\in\partial K}h_\gamma|\bar{R}_\gamma|\Big) \\ & +\sum_{K\in\mathcal{P}}\Big(||\bar{r}_K||_K\,||\tilde{\epsilon}||_K+\sum_{\gamma\in\partial K}||\bar{R}_\gamma||_\gamma\,||\tilde{\epsilon}||_\gamma\Big) \\ & +\sum_{\gamma\in\partial\mathcal{P}}h_\gamma^{1/2}\left|\left|\frac{\partial u}{\partial n}-\left\langle\frac{\partial u_X}{\partial n}\right\rangle_\gamma\right|\right|_\gamma\Big|[A_\gamma]\Big|\Big\}.\end{aligned}$$

Furthermore, equation (A.8) shows that

$$||R-\bar{R}_\gamma||_\gamma^2\leq Ch_\gamma\Big\{h_\gamma^2h_{\perp\gamma}^2|D_{\perp\gamma}^2(D_\gamma(\boldsymbol{\epsilon}'))|^2+\delta^2h_\gamma^2|D_{xy}(\boldsymbol{\epsilon}')|^2\Big\},$$

so after substituting this and defining Q_0, Q_1 and Q_2 as in (5.2), Theorem 5.1 is proved.

A.4 Proof of Theorem 5.2

A.4.1 Proof of first statement

From Section A.2 we have the following expressions for the error and the reconstructed error on element K,

$$\begin{aligned}e|_K &= c_1+c_2x+c_3y+c_4xy+a_5x^2+a_6y^2+\tilde{\epsilon} \\ e_K &= \tilde{c}_1+\tilde{a}_5x^2+\tilde{a}_6y^2\,.\end{aligned}$$

Hence,

$$\begin{aligned}\nabla(e|_K-e_K)\cdot\ &\nabla(e|_K-e_K) \\ &\leq\ C\Big\{c_2^2+c_3^2+\big(c_4^2+|a_5-\tilde{a}_5|^2\big)x^2 \\ &\qquad +\big(c_4^2+|a_6-\tilde{a}_6|^2\big)y^2+\nabla\tilde{\epsilon}\cdot\nabla\tilde{\epsilon}\Big\},\end{aligned}$$

where the different terms can be bounded using (A.6), (A.11), and (A.12). Since in element K we have $x^2<h_x^2$ and $y^2<h_y^2$, the following upper bound is obtained,

$$\begin{aligned}\nabla(e|_K-e_K)\cdot\ &\nabla(e|_K-e_K) \\ &\leq\ C\Big\{\big(\hat{D}_x(\boldsymbol{\epsilon}')\big)^2+\big(\hat{D}_y(\boldsymbol{\epsilon}')\big)^2+(h_x^2+h_y^2)\big(\hat{D}_{xy}(\boldsymbol{\epsilon}')\big)^2 \\ &\qquad +h_x^2\big(\hat{D}_x^2(\boldsymbol{\epsilon}')\big)^2+h_y^2\big(\hat{D}_y^2(\boldsymbol{\epsilon}')\big)^2+\nabla\tilde{\epsilon}\cdot\nabla\tilde{\epsilon}\Big\}\,.\end{aligned}$$

Direct integration reveals that

$$|||e - e_K|||_K \leq C\Big\{|K|^{1/2}\phi_{R,K}(\boldsymbol{\epsilon}') + |||\tilde{\epsilon}|||_K\Big\},$$

where $\phi_{R,K}(\boldsymbol{\epsilon}')$ was defined in (5.3), and $|K|$ is the area of element K. Since $\eta^*_{R,K}$ is defined by

$$\eta^*_{R,K} = |||e_K|||,$$

the triangle inequality shows that

$$\Big| |||e|||_K - \eta^*_{R,K} \Big| \leq C\Big(|K|^{1/2}\phi_{R,K}(\boldsymbol{\epsilon}') + |||\tilde{\epsilon}|||_K\Big),$$

as claimed.

A.4.2 Proof of second statement

The difference between the estimator $\eta^*_{R,K}$, which takes into account the changes in the mesh size, and the estimator $\eta_{R,K}$, which neglects the changes in the mesh, is

$$\begin{aligned}(\eta^*_{R,K})^2 - \eta^2_{R,K} \quad = \quad & \frac{h_x h_y}{12}\Big\{\Big(\frac{\bar{R}_1}{2+k_1\delta} + \frac{\bar{R}_3}{2+k_3\delta}\Big)^2 - \Big(\frac{\bar{R}_1}{2} + \frac{\bar{R}_3}{2}\Big)^2 \\ & + \Big(\frac{\bar{R}_2}{2+k_4\delta} + \frac{\bar{R}_4}{2+k_2\delta}\Big)^2 - \Big(\frac{\bar{R}_2}{2} + \frac{\bar{R}_4}{2}\Big)^2\Big\}.\end{aligned}$$

From this we easily obtain

$$\begin{aligned}\Big|(\eta^*_{R,K})^2 - \eta^2_{R,K}\Big| \quad & \leq \quad \frac{h_x h_y}{12}\left\{\left(\frac{\bar{R}_1 + \bar{R}_3}{2}\right)^2 + \left(\frac{\bar{R}_2 + \bar{R}_4}{2}\right)^2\right\}\mathcal{O}(\delta) \\ & = \quad \eta^2_{R,K}\,\mathcal{O}(\delta).\end{aligned}$$

For non-negative real numbers a and b, the following inequality holds,

$$|a-b| \leq |a^2 - b^2|^{1/2},$$

so

$$\Big|\eta^*_{R,K} - \eta_{R,K}\Big| \quad \leq \quad \eta_{R,K}\,\mathcal{O}(\delta^{1/2})\ .$$

By the aid of the first statement in the theorem and the triangle inequality we obtain

$$\Big| |||e||| - \eta_{R,K}\Big| \leq C\Big(|K|^{1/2}\phi_{R,K}(\boldsymbol{\epsilon}') + |||\tilde{\epsilon}||| + \eta_{R,K}\,\mathcal{O}(\delta^{1/2})\Big)\ ,$$

as claimed.