RICE UNIVERSITY

Adaptive finite element methods for linear-quadratic convection dominated elliptic optimal control problems

by

Eelco Nederkoorn

A Thesis Submitted
in Partial Fulfillment of the
Requirements for the Degree

Master of Arts

Approved, Thesis Committee:

Dr. Matthias Heinkenschloss, Chairman
Professor of Computational and Applied Mathematics

Dr. Tim Warburton
Associate Professor of Computational and Applied Mathematics

Dr. Béatrice Rivière
Associate Professor of Computational and Applied Mathematics

Houston, Texas

August, 2009

Abstract

Adaptive finite element methods for linear-quadratic convection dominated elliptic optimal control problems

by

Eelco Nederkoorn

The numerical solution of linear-quadratic elliptic optimal control problems requires the solution of a coupled system of elliptic partial differential equations (PDEs), consisting of the so-called state PDE, the adjoint PDE and an algebraic equation. Adaptive finite element methods (AFEMs) attempt to locally refine a base mesh in such a way that the solution error is minimized for a given discretization size. This is particularly important for the solution of convection dominated problems where inner and boundary layers in the solutions to the PDEs need to be sufficiently resolved to ensure that the solution of the discretized optimal control problem is a good approximation of the true solution.

This thesis reviews several AFEMs based on energy norm based error estimates for single convection dominated PDEs and extends them to the solution of the coupled

system of convection dominated PDEs arising from the optimality conditions for optimal control problems.

Keywords Adaptive finite element methods, optimal control problems, convection-diffusion equations, local refinement, error estimation.

Acknowledgements

I'd like to thank my thesis committee, Dr Heinkenschloss, Dr Rivière and Dr Warburton.

This research was supported in part by NSF Grant DMS-0511624.

Contents

List of Figures

Chapter 1

Introduction

This thesis analyzes and compares different adaptive finite element methods (AFEMs) for linear-quadratic convection dominated elliptic optimal control problems of the form

$$\min \ \frac{1}{2}\|y-\widehat{y}\|^2_{L^2(\Omega)}+\frac{\omega}{2}\|u\|^2_{L^2(\Omega)} \tag{1.1a}$$

subject to

$$-\epsilon\Delta y+b\cdot\nabla y+cy=f+u \qquad \text{in } \Omega, \tag{1.1b}$$

$$y=g_D \qquad \text{on } \Gamma_D, \tag{1.1c}$$

$$\epsilon\frac{\partial y}{\partial n}=g_N \qquad \text{on } \Gamma_N, \tag{1.1d}$$

where the boundary of $\Omega\subset\mathbb{R}^2$ is divided such that $\partial\Omega=\Gamma_D\cup\Gamma_N$ and $\Gamma_D\cap\Gamma_N=\emptyset$ (note that Γ_D and Γ_N are respectively the Dirichlet and Neumann boundary). Assume that the constants $\epsilon>0$, $\omega>0$ and the functions $b:\mathbb{R}^2\to\mathbb{R}^2$, $c:\mathbb{R}^2\to\mathbb{R}$, $f:\mathbb{R}^2\to\mathbb{R}$, $g_D:\mathbb{R}^2\to\mathbb{R}$, $g_N:\mathbb{R}^2\to\mathbb{R}$ and $\widehat{y}:\mathbb{R}^2\to\mathbb{R}$ are sufficiently smooth.

This thesis first reviews AFEMs for the uncontrolled problem, i.e. the convection-diffusion equation (1.1b) - (1.1d) with $u \equiv 0$. In the second part of his thesis the theory of AFEMs for the single convection-diffusion equation (1.1b) - (1.1d) is extended to the optimal control problem (1.1).

For the single equation case, this thesis focuses on the convection dominated problems, i.e. $\epsilon \ll \|b\|_\infty$. Solutions to such convection dominated diffusion equations typically have localized features. These so-called layers, local regions with steep gradients, tend to get steeper when ϵ gets smaller. In general these layers occur at the 'outflow' boundary $\{x \in \partial\Omega \ : \ b(x) \cdot n(x) > 0\}$ (so-called boundary layers) or in the interior of Ω (inner layers) ([32], [33]). Boundary layers arise because the interior solution, driven by strong convection, suddenly has to match the Dirichlet boundary conditions at the outflow boundary. Inner layers typically stem from a discontinuity at the inflow boundary data. Despite this discontinuity at the boundary, the solution is continuous in the interior because of the, albeit small, diffusive nature of the problem. However, due to the dominating convection, such discontinuities at the boundary cause a sharp gradient in a small band throughout the interior of Ω.

Standard Galerkin finite element (FEM) approximations produce oscillatory solutions unless the mesh size is small relative to $\epsilon/\|b\|_\infty$. Since $\epsilon/\|b\|_\infty$ is very small in the problems of interest, the standard Galerkin FEM has to be modified to compute 'good' approximate solutions at moderate mesh sizes. This thesis considers the streamline upwind / Petrov-Galerkin (SUPG) method of Brooks and Hughes ([18],

[19]) . This thesis numerically shows that even though this stabilization method significantly improves the numerical solution, this solution may still contain oscillations in cross wind direction in a small region around the layers. The errors caused by these oscillations may be propagated downwind by the convection flow into the interior of Ω. Hence, even the stabilized variational problem needs to resolve the layers in order to achieve satisfying results.

Uniform meshes with sufficiently small elements tend to get impractically large. This gives rise to the idea of using locally refined meshes around the layers. Adaptive finite element methods (AFEMs) generate such locally refined meshes. A typical AFEM computes a numerical solution on a triangulation, estimates the local error on each single element, marks a selection of the elements and refines the selection. This iteration is repeated until a desired accuracy or a maximum number of nodes is reached.

Key in the adaptive process is estimating the error of a computed numerical solution. This thesis reviews and numerically compares three existing classes of error indicators: the Zienkiewicz-Zhu estimator ([7], [39]), the norm-residual based estimator and the local Neumann estimator ([23], [34], [35], [36], [37]). The results clearly indicate that choosing an error estimator is highly problem dependent, which coincides with existing numerical studies ([20], [23], [30]).

The solution to the linear-quadratic convection dominated elliptic optimal control problem (1.1) is characterized by the first order optimality conditions. This coupled

system ([26], [27], [13]) consists of the the state PDE

$$-\epsilon\Delta y + b\cdot\nabla y + cy = f + u \qquad \text{in } \Omega, \tag{1.2a}$$

$$y = g_D \qquad \text{on } \Gamma_D, \tag{1.2b}$$

$$\epsilon\frac{\partial y}{\partial n} = g_N \qquad \text{on } \Gamma_N, \tag{1.2c}$$

the adjoint PDE

$$-\epsilon\Delta p - b\cdot\nabla p + (c - \nabla\cdot b)p = -(y - y_0) \qquad \text{in } \Omega, \tag{1.3a}$$

$$p = 0 \qquad \text{on } \Gamma_D, \tag{1.3b}$$

$$\epsilon\frac{\partial p}{\partial n} + (b\cdot n)p = 0 \qquad \text{on } \Gamma_N, \tag{1.3c}$$

and the gradient equation

$$p = \omega u \qquad \text{in } \Omega. \tag{1.4}$$

The solution of this system has three components, the state y, the adjoint p and the control u. Since the two PDEs in this system are convection dominated (with convection in opposite direction), some of the components y, u, p may exhibit layers. Again, when these layers are not resolved, the standard FEM solution contains spurious oscillations. Heinkenschloss and Collis [13] introduce the SUPG stabilization to this coupled system, which significantly reduces the oscillations. However, as this thesis shows by computation, oscillations may still occur in cross-wind direction in a small band around the layers. Because of the coupling of the system and the opposite convection in state and adjoint PDE, errors caused by spurious oscillations are

propagated up- and downwind from the layers. These propagated errors might even show in components of the solution which do not have layers [15]. Therefore, in the optimal control setting it is even more important to resolve the local features of the solution.

This thesis applies adaptive finite element methods, combined with SUPG stabilization, to linear-quadratic convection dominated elliptic optimal control problems. The goal is to obtain numerical solutions defined on meshes which are able to resolve the layers in all components of the solution. Like in the single equation case, key components in an adaptive strategy are a posteriori error indicators. Though error estimation for SUPG solutions to single convection dominated diffusion equations is an established technique, only recently a few of such indicators were proposed for optimal control problems ([8], [17], [38]). The estimators in [17], [38] are essentially a generalization of the norm-residual based error estimators. This thesis uses this existing work and in addition extends the ZZ and local Neumann estimator to the optimal control setting.

Recall that for the single convection diffusion equation there was no preferred estimator which excels in all situations. Because of the presence of multiple layers in different components of the solution and the increased capacity of error propagation, there is even less of a clear choice in error estimators in the optimal control setting. The numerical results shown in this thesis clearly illustrate in which situations the classes of estimators excel or under perform.

The restriction to three classes of estimators leaves out many proposed alternatives. One of the methods out of scope of this thesis is goal-oriented estimation ([4], [8]). The central idea in these methods is to reduce the error of a given functional of the error e_h using a duality based approach. However, the functional can be chosen such that the goal-oriented estimator is in essence the same as the norm-residual based approach mentioned above. In optimal control setting, the functional can be chosen the same as the objective function [8]. Resolving layers using such techniques is well worth exploring.

The adaptive strategy used in this work follows the iteration solve, estimate, mark and refine ([9], [28]). Though the title of this thesis suggests all these aspects of adaptive finite element methods are discussed, this work mainly focuses on solving the problem (i.e. SUPG/Galerkin approximation) and error estimation. The mark and refine steps, which are also essential to the adaptive process, are beyond the scope of this thesis.

A popular strategy for selecting elements for refinement based on an error estimator, is bulk marking ([9], [10], [11], [12], [28]). In this work a fixed percentage of elements is marked each refinement step. In this approach there is a guaranteed increase in elements, which has its computational benefits.

This thesis uses continuous finite element methods, which requires the underlying mesh to be conform. Local refinement of triangular meshes is done through bisection. This approach has been extensively studied in order to ensure that refined meshes

are conform ([1], [5], [6], [10], [11], [12]). Another option is to use quadrilateral meshes and interpolate the solution on hanging nodes ([2], [3]). One could also use discontinuous Galerkin methods ([16], [31]), which do not require conformity of the mesh.

Convergence of a similar adaptive strategy as outlined above has been proven ([28], [29]). However, these articles do not specifically target convection-diffusion equations and SUPG stabilization. Moreover, they only use the norm-residual based error estimator combined with the bulk marking strategy. Therefore, a solid mathematical foundation for the convergence of all AFEMs used in this thesis still has to be developed.

This work uses continuous finite element methods in conjuncture with SUPG stabilization because it is so commonly used in literature. Hence, extending this known theory is a logic step in applying adaptive FEM to optimal control problems. Though methods have been developed ([21], [22]) which reduce the cross wind spurious oscillations around the layers, these require additional artificial diffusion, and therefore also reduce the sharpness of the layers.

1.1 Organization of the Thesis

This thesis first analyzes and compares different AFEMs for the single convection-diffusion equation, and extends this theory to the optimal control setting.

The first section, Chapter 2, describes the finite element discretization and SUPG

stabilization of a single convection-diffusion equation. Recall that key in an adaptive regime is estimating the error of an obtained numerical solution. This thesis reviews the existing theory on the ZZ, norm-residual based and local Neumann estimator (Chapter 3). Local refinements driven by these different estimators yield different results. Chapter 4 explains several numerical experiments, and compares the effectivity of the three estimators when applied to these examples.

The results in Chapters 2 – 3 form a foundation for the analysis of AFEMs for linear-quadratic convection dominated elliptic optimal control problems. Chapter 5 first reviews the finite element discretization in conjuncture with SUPG stabilization, and secondly extends the three estimators to the optimal control setting. Similar experiments as in the single equation case are performed with these optimal control problems. The second numerics chapter (Chapter 6) explains these examples, states the results and compares the different estimators.

Chapter 2

Finite element discretization

This section discusses the solution of a single convection diffusion equation

$$-\epsilon \Delta y + b \cdot \nabla y + cy = f \qquad \text{in } \Omega, \tag{2.1a}$$

$$y = g_D \qquad \text{on } \Gamma_D, \tag{2.1b}$$

$$\epsilon \frac{\partial y}{\partial n} = g_N \qquad \text{on } \Gamma_N, \tag{2.1c}$$

using the finite element method with SUPG stabilization.

2.1 Weak formulation

Given a subset $E \subseteq \bar{\Omega}$ define

$$\|v\|_{L^2(E)} = \left[\int_E |v|^2\right]^{1/2}, \tag{2.2a}$$

$$\|v\|_{H^1(E)} = \left[\int_E |v|^2 + |\nabla v|^2\right]^{1/2}, \tag{2.2b}$$

$$\|v\|_{\epsilon} = \left[\int_E |v|^2 + \epsilon|\nabla v|^2\right]^{1/2}. \tag{2.2c}$$

Furthermore, let

$$H^1_D(\Omega) = \left\{v \in H^1(\Omega) \ : \ v = 0 \text{ on } \Gamma_D\right\},$$

$$H^1_g(\Omega) = \left\{v \in H^1(\Omega) \ : \ v = g_D \text{ on } \Gamma_D\right\},$$

and for w in the dual $[H^1_D(\Omega)]^*$ of $H^1_D(\Omega)$ the norm is defined by

$$\|w\|_{\epsilon,*} = \sup_{v \in H^1_D \setminus \{0\}} \frac{(w, v)}{\|v\|_{\epsilon}}. \tag{2.3}$$

The weak form corresponding to (2.1) is obtained by multiplying (2.1a) with a test function $v \in H^1_D(\Omega)$, applying integration by parts and using (2.1c). This leads to the following variational problem:

Find $y \in H^1_g(\Omega)$ such that

$$a(y, v) = \ell(v) \qquad \forall v \in H^1_D(\Omega), \tag{2.4}$$

where

$$a(y, v) = \int_\Omega \epsilon \nabla y \cdot \nabla v + \int_\Omega b \cdot \nabla y \ v + \int_\Omega cyv, \tag{2.5}$$

$$\ell(v) = \int_\Omega fv + \int_{\Gamma_N} g_N v. \tag{2.6}$$

Consider the following assumptions

(A1) $\Omega \subset \mathbb{R}^d$, is a polygonal domain with boundary $\partial\Omega$ decomposed into Γ_N and $\Gamma_D = \partial\Omega \setminus \Gamma_N$.

(A2) $b \in [W^{1,\infty}(\Omega)]^d$, $c \in L^\infty(\Omega)$, $f \in L^2(\Omega)$, $g_N \in L^2(\Gamma_N)$, there exists $y_D \in H^1(\Omega)$ such that $y_D = g_D$ on Γ_D.

(A3) $\Gamma_N \subseteq \{x \in \partial\Omega \ : \ b(x) \cdot n_{\partial\Omega} \geq 0\}$

(A4) $c - \frac{1}{2}\nabla \cdot b \geq \gamma > 0$ a.e. in Ω.

If Γ_D has a nonzero $d-1$ dimensional measure, then (A4) can be replaced by

(A4)' $c - \frac{1}{2}\nabla \cdot b \geq \gamma \geq 0$ a.e. in Ω.

Under the assumptions (A1)-(A4) existence and uniqueness of the solution $y \in H^1(\Omega)$ of (2.4) can be shown using the Lax-Milgram lemma. This work shows coercivity and continuity of the bilinear form with respect to the $\epsilon-$weighted $H^1(\Omega)$ norm defined by

$$\|v\|_\epsilon = \left[\int_\Omega |v|^2 + \epsilon|\nabla v|^2\right]^{1/2}.$$

Since $0 < \epsilon \ll 1$, the following relation holds

$$\|v\|^2_{H^1(\Omega)} = \frac{1}{\epsilon}\left[\epsilon\|v\|^2_{L^2(\Omega)} + \epsilon\|\nabla v\|^2_{L^2(\Omega)}\right] \leq \frac{1}{\epsilon}\|v\|^2_\epsilon. \tag{2.7}$$

Lemma 2.1.1 *If the assumptions (A1)-(A4) hold, then there exist $\alpha, M > 0$ which are independent of ϵ such that*

$$a(v,v) \geq \alpha\|v\|^2_\epsilon,$$

and

$$|a(v,w)| \leq M\|v\|_\epsilon \|w\|_{H^1(\Omega)},$$
$$|a(v,w)| \leq \frac{M}{\sqrt{\epsilon}}\|v\|_\epsilon \|w\|_\epsilon.$$

for all $v, w \in H^1_D(\Omega)$.

Proof: For ellipticity, note that for $v \in H^1(\Omega)$

$$\begin{aligned} a(v,v) &= \epsilon \int_\Omega \nabla v \cdot \nabla v + \int_\Omega (b \cdot \nabla v)v + \int_\Omega cv^2 \\ &= \epsilon \|\nabla v\|^2_{L^2(\Omega)} + \frac{1}{2}\int_\Omega (b \cdot \nabla v)v + \frac{1}{2}\int_\Omega (b \cdot \nabla v)v + \int_\Omega cv^2 \\ &= \epsilon \|\nabla v\|^2_{L^2(\Omega)} + \frac{1}{2}\int_\Omega (b \cdot \nabla v)v + \frac{1}{2}\int_\Omega (vb) \cdot \nabla v + \int_\Omega cv^2 \end{aligned}$$

Using integration by parts yields

$$\begin{aligned} a(v,v) &= \epsilon \|\nabla v\|^2_{L^2(\Omega)} + \frac{1}{2}\int_\Omega (b \cdot \nabla v)v - \frac{1}{2}\int_\Omega \nabla \cdot (vb)\ v + \int_{\partial\Omega} v^2 b \cdot n_{\partial\Omega} + \int_\Omega cv^2 \\ &= \epsilon \|\nabla v\|^2_{L^2(\Omega)} + \frac{1}{2}\int_\Omega (b \cdot \nabla v)v - \frac{1}{2}\int_\Omega (v \cdot \nabla b + b \cdot \nabla v)\ v + \int_{\Gamma_N} v^2 b \cdot n_{\partial\Omega} + \int_\Omega cv^2 \\ &= \epsilon \|\nabla v\|^2_{L^2(\Omega)} + \int_\Omega (c - \frac{1}{2}\nabla \cdot b)v^2 + \int_{\Gamma_N} v^2 b \cdot n_{\partial\Omega}. \end{aligned}$$

Assumption (A4) implies

$$a(v,v) \geq \epsilon \|\nabla v\|^2_{L^2(\Omega)} + \gamma \|v\|^2_{L^2(\Omega)} \geq \alpha \|v\|^2_\epsilon,$$

where $\alpha = \min(1, \gamma)$. This inequality implies that the bilinear form $a(\cdot,\cdot)$ is elliptic, independent of ϵ, in the $\|\cdot\|_\epsilon$ norm. Continuity follows from

$$\begin{aligned} |a(v,w)| &= \left| \epsilon \int_\Omega \nabla v \cdot \nabla w + \int_\Omega (b \cdot \nabla v)w + \int_\Omega cvw \right|, \\ &= \left| \epsilon \int_\Omega \nabla v \cdot \nabla w + \int_\Omega (\nabla \cdot (bv) - (\nabla \cdot b)v)w + \int_\Omega cvw \right|. \end{aligned}$$

Integration by parts and the Cauchy-inequality give

$$\begin{aligned}
|a(v,w)| &\le \epsilon \int_\Omega |\nabla v \cdot \nabla w| + \int_\Omega |v(b\cdot\nabla w)| + \int_\Omega |(\nabla\cdot b)vw| + \int_\Omega |cvw| \\
&\le \epsilon \|\nabla v\|_{L^2(\Omega)} \|\nabla w\|_{L^2(\Omega)} + \|b\|_\infty \|v\|_{L^2(\Omega)} \|\nabla w\|_{L^2(\Omega)} \\
&\qquad + \ \|\nabla b\|_\infty \|v\|_{L^2(\Omega)} \|w\|_{L^2(\Omega)} + \|c\|_\infty \|v\|_{L^2(\Omega)} \|w\|_{L^2(\Omega)} \\
&\le \left(\sqrt{\epsilon}\|\nabla v\|_{L^2(\Omega)} + \|v\|_{L^2(\Omega)}\right) \left(M_1 \|\nabla w\|_{L^2(\Omega)} + M_2 \|w\|_{L^2(\Omega)}\right) \\
&\le M \|v\|_\epsilon \|w\|_{H^1(\Omega)},
\end{aligned}$$

with $M_1 = \sqrt{\epsilon} + \|b\|_\infty$, $M_2 = \|\nabla b\|_\infty + \|c\|_\infty$ and $M = \max(M_1, M_2)$. Using relation (2.7) we obtain $|a(v,w)| \le \frac{M}{\sqrt{\epsilon}} \|v\|_\epsilon \|w\|_\epsilon$. □

The continuity and ellipticity guarantee the existence of a unique solution by the Lax-Milgram theorem (see, e.g., [14, 24]).

Theorem 2.1.2 *If the assumptions (A1)-(A4) hold, then* (2.4) *has a unique solution* $y \in H^1(\Omega)$.

Under additional regularity conditions additional smoothness results can be proved. See, e.g., [32, 33]. The following result is proved in [33, L. 7.2].

Theorem 2.1.3 *Assume that* $\Omega \subset \mathbb{R}^2$ *is bounded and convex or has smooth boundary and that* $\Gamma_D = \partial\Omega$. *Furthermore, let* b, c, f *be Hölder continuous on* $\overline{\Omega}$, $c \ge 0$, *and let* $g \in H^{3/2}(\partial\Omega)$. *There exists a constant* C *independent of* ϵ *such that*

$$\epsilon^{3/2}|y|_2 + \epsilon^{1/2}|y|_1 + \|y\|_{L^2(\Omega)} \le C.$$

In particular, Theorem 2.1.3 indicates that the norm of the second derivative of the solution y behaves like $\epsilon^{-3/2}$,

$$\|y\|_{H^2(\Omega)} = O(\epsilon^{-3/2}), \tag{2.8}$$

which is true in many examples where the solution exhibits layers. See [32, 33].

2.2 Galerkin approximation

Let $\mathcal{T}_h$ be a conforming triangulation and let the following function spaces on $\mathcal{T}_h$ be defined by:

$$Y_h = \left\{ y_h \in H^1(\Omega) \ : \ y_h|_\tau \in P_p(\tau), \ \forall \tau \in \mathcal{T}_h \right\},$$

$$Y_h^0 = \left\{ y_h \in Y_h \ : \ y_h = 0 \text{ on } \Gamma_D \right\},$$

$$Y_h^g = \left\{ y_h \in Y_h \ : \ y_h = g_D \text{ on } \Gamma_D \right\}.$$

Assume that g_D is a piecewise polynomial on Γ_D. The finite element method reduces the weak form (2.4) to:

Find $y_h \in Y_h^g$ such that

$$a(y_h, v_h) = \ell(v_h) \qquad \forall v_h \in Y_h^0. \tag{2.9}$$

This thesis uses piecewise linear functions ($p = 1$). All theory presented in this work can be generalized to higher order piecewise continuous polynomials.

Let y be the solution to the variational problem and let y_h be the Galerkin ap-

proximation. By Lemma 2.1.1 and orthogonality of the error

$$\alpha\|y-y_h\|_\epsilon^2 \leq a(y-y_h, y-y_h) = a(y-y_h, y-v_h) \leq M\|y-y_h\|_\epsilon \|y-v_h\|_{H^1(\Omega)}.$$

for any arbitrary $v_h \in Y_h^0$. Hence,

$$\|y-y_h\|_\epsilon \leq \frac{M}{\alpha} \inf_{v\in Y_h} \|y-v_h\|_{H^1(\Omega)}.$$

This equation implies that the error of the Galerkin approximation y_h is dependent on ϵ. This becomes even more evident if in addition $y \in H^2(\Omega)$. A standard a priori error bound can be applied to this result ([14, p.135], [24, p.382]):

$$\|y-y_h\|_\epsilon \leq Ch\|y\|_{H^2(\Omega)}, \tag{2.10}$$

for constant $C > 0$, and h an indicator of the largest element size in $\mathcal{T}_h$. Equation (2.8) shows that the constant term in this error bound is large in convection dominated situations (i.e. $0 < \epsilon \ll 1$). Therefore, when $\epsilon < h$ the constant term in this bound dominates the mesh size h. Hence, h needs to be sufficiently small, especially around the layers, in order to significantly reduce the error.

Besides this disadvantageous property of convection dominated diffusion equations, there is another, Galerkin specific, reason why it is hard to compute a numerical solution with the standard finite element method. Recall the definition of $\|\cdot\|_\epsilon$, which expands the bound (2.10) to:

$$\|y-y_h\|_{L^2(\Omega)}^2 + \epsilon\|\nabla(y-y_h)\|_{L^2(\Omega)}^2 \leq C^2h^2\|y\|_{H^2(\Omega)}^2. \tag{2.11}$$

This bound implies that large errors in the gradient are possible since they are weighted by $\epsilon \ll 1$. Therefore, when $\epsilon < h$, the Galerkin approximation may not

be able to control the error in its gradient. Errors that arise from not resolving boundary layers are propagated throughout the entire numerical solution. In practice this results in Galerkin approximations which may contain spurious oscillations when the mesh-size is not fine enough.

Computing a solution on uniform meshes of these sufficiently small $h < \epsilon$ elements is often unfeasible. Therefore stabilization terms need to be added to the standard Galerkin method. The stabilized FEM aims to generate 'good' solutions for moderate mesh sizes. Additionally locally refined meshes, preferably around the layers, could solve this issue, which motivates the adaptive strategy used in this thesis.

2.3 SUPG stabilization

A popular stabilization method is the streamline upwind / Petrov Galerkin (SUPG) method, sometimes referred to as streamline diffusion FEM (SDFEM) which was introduced by Hughes and Brooks ([18], [19]). This method adds stabilization terms to the weak form in (2.9). The stabilized problem is:

Find $y_h \in Y_h^g$ *such that*

$$a_h(y_h, v_h) = \ell_h(v_h) \qquad \forall v_h \in Y_h^0, \tag{2.12a}$$

where

$$a_h(y_h, v_h) = a(y_h, v_h) + \sum_{\tau \in T_h} \delta_\tau (-\epsilon \Delta y_h + b \cdot \nabla y_h + c y_h, b \cdot \nabla v_h)_\tau \tag{2.12b}$$

$$\ell_h(v_h) = \ell(v_h) + \sum_{\tau \in T_h} \delta_\tau (f, b \cdot \nabla v_h)_\tau, \tag{2.12c}$$

and δ_τ, $\tau \in \mathcal{T}_h$, are stabilization parameters. The inner product $(\cdot, \cdot)_\tau$ in the above definition is the $L^2(\tau)$ inner product. The stabilization parameters highly influence the efficiency of the method and need to be chosen carefully. Let the element Peclet number of an element τ be defined as

$$\mathcal{P}_\tau = \frac{h_\tau}{2\epsilon} \|b\|_{\tau,\infty},$$

where h_τ is an indicator of the size of τ. A common choice for h_τ is the diameter of τ along the convection flow b. John and Knobloch [21] introduced an approximation which is exact for the triangular elements used in this thesis. Let C_τ be the barycenter of τ, then

$$h_\tau = \frac{2|b(C_\tau)|}{\sum_{i=1}^{3} |b(C_\tau) \cdot \nabla \phi_i(C_\tau)|},$$

where $\{\phi_i\}_{i=1}^3$ are the three nonzero linear basis functions on τ. The mesh-Peclet number indicates in what degree convection is dominating over diffusion on a element τ. Large Peclet numbers, $\mathcal{P}_\tau \gg 1$, correspond to a convection dominated regime, while low Peclet numbers, $0 < \mathcal{P}_\tau \ll 1$, are associated with diffusion dominated situations. For such low Peclet numbers, the Galerkin approximation will yield satisfactory results, and one should refrain from introducing any stabilizing terms. The choice of the stabilization parameters are an ongoing discussion in the scientific community.

There does not seem to be an optimal choice for every problem. Elman et al. [14, p.132] published a simple parameter that yields satisfactory numerical results:

$$\delta_\tau = \begin{cases} \frac{h_\tau}{2\|b\|}\left(1 - \frac{1}{\mathcal{P}_\tau}\right) & \mathcal{P}_\tau \geq 1, \\ 0, & \mathcal{P}_\tau < 1. \end{cases} \tag{2.13}$$

Note that for low Peclet numbers the SUPG approximation is equal to the Galerkin approximation. For other parameter choices, see [21], [22], [25].

There exist a unique solution of (2.12), since $a_h(\cdot,\cdot)$ is coercive and continuous in the SUPG norm ([14], [24])

$$\|v\|_{\text{supg}} = \left[\|v\|^2_{L^2(\Omega)} + \epsilon\|\nabla v\|^2_{L^2(\Omega)} + \sum_{\tau\in\mathcal{T}_h} \delta_\tau \|b\cdot\nabla v\|^2_{L^2(\Omega)}\right]^{1/2}.$$

Since $\|v\|_\epsilon \leq \|v\|_{\text{supg}}$ for $v \in H^1_D(\Omega)$ the SUPG norm is stronger than the weighted norm. If $P_\tau \gg 1$ for all τ and if in addition $y \in H^2(\Omega)$, then ([14, p.136], [24, p.381]):

$$\|y - y_h\|_{\text{supg}} \leq Ch^{3/2}\|y\|_{H^2(\Omega)}. \tag{2.14}$$

Note that the $\|y\|_{H^2(\Omega)}$ is still inversely dependent on ϵ by equation (2.8). The benefit of using SUPG over Galerkin becomes evident when expanding the SUPG norm:

$$\|y-y_h\|^2_{L^2(\Omega)} + \epsilon\|\nabla(y-y_h)\|^2_{L^2(\Omega)} + \sum_{\tau\in\mathcal{T}_h}\delta_\tau\|b\cdot\nabla(y-y_h)\|^2_{L^2(\Omega)} \leq C^2h^3\|y\|^2_{H^2(\Omega)}. \tag{2.15}$$

Again, the error in the gradient is weighted by an $\epsilon \ll 1$ term. However, the SUPG stabilization yields a term which controls the error the gradient along the convection flow b. Since δ_τ is not dependent on ϵ, a large error in the gradient along the convection lines is not possible. This summarizes the power of SUPG stabilization. In practice

the SUPG stabilization significantly reduces the oscillations in the direction of the convection. This implies that the gradient error in the cross wind direction is not controlled, and the solution could still have spurious oscillations perpendicular to the convection. However, the SUPG stabilization leads to a smaller error in the gradient perpendicular to the convection b than the standard Galerkin approximation. Note that the the bound (2.15) contains a h^3 term instead of the h^2 term of the Galerkin approximation in equation (2.10). Hence, when h approaches ϵ, the gradient error in all directions becomes more influential, and is reduced faster. This double play is exceptionally effective and makes this method the preferred choice in many applications.

SUPG stabilization does not make adaptive refinement superfluous. Recall that the solutions to convection dominated diffusion equations typically exhibit layers. The artificial diffusion along the convection lines that SUPG introduces, also reduces the sharpness of the layers. A remedy is local refinement to resolve layers sufficiently (i.e. such that $\mathcal{P}_\tau < 1$ along the layers), and than switch to Galerkin approximation in these areas. This justifies our choice of stabilization parameter, which enables this switch by setting $\delta_\tau = 0$.

Chapter 3

A posteriori error estimation for convection-diffusion equations

Adaptive finite element methods repeatedly improve a numerical solution by locally refining the underlying mesh. The selection of elements for refinement is based on the local error of the numerical solution. The exact error is never known, since it requires the true solution, which FEM intends to approximate in the first place. Therefore, approximation techniques are necessary. This chapter focuses on such a posteriori error estimation.

There is much written about error estimation over the last few decades. This thesis considers three estimators. These specific classes have been studied the most extensively the specific case of convection dominated diffusion equations. The three methods are the Zienkiewicz-Zhu error, the norm-residual based and the local Neu-

mann problem estimator.

Let y be the solution to the variational problem, let y_h be the Galerkin or SUPG approximation and define the error by $e_y = y - y_h$. Typically an estimator η_τ tries to approximate $\|e_y\|_\tau$ in some norm. The overall global error η is commonly defined by

$$\eta = \left[\sum_{\tau \in \mathcal{T}_h} \eta_\tau^2 \right]^{1/2}.$$

One of the difficulties in error estimation is to find an appropriate measure of error. That is, to find a norm such that $\|e_y\|$ can be approximated properly. In common practice ([9], [20], [34], [35]), an error estimator efficient and reliable if

$$\|y - y_h\| \leq C_1 \eta, \tag{3.1a}$$

$$\eta_\tau \leq c_\tau \|y - y_h\|_{N(\tau)}, \ \forall \tau \in \mathcal{T}_h, \tag{3.1b}$$

where $C_1 > 0$ is independent of the triangulation T_h and $N(\tau)$ is some neighborhood of τ (definition will follow). One usually assumes that c_τ is nearly constant ($c_\tau \approx C_2, \forall \tau \in T_h$). In many cases the last inequality is hard to prove and a similar bound is used:

$$\eta \leq C_2 \|y - y_h\|. \tag{3.1c}$$

This work reviews the three estimators. The first class, the Zienkiewicz-Zhu estimator is probably the most simple, and therefore a good start of this chapter.

3.1 Zienkiewicz-Zhu estimator

One of the first error estimators was introduced by Zienkiewicz and Zhu [39]. This so-called averaging technique obtains a high order recovery Gy of the gradient of y such that $Gy \approx \nabla y$, where y is the solution to the variational problem (2.4). The aim is to estimate $\|\nabla y - \nabla y_h\|_{L^2(\Omega)}$ by $\|Gy - \nabla y_h\|_{L^2(\Omega)}$, where y_h is either a Galerkin or SUPG approximation. Zienkiewicz and Zhu proposed to use the following piecewise linear function, defined at the nodes $\{x_j\}$ of the triangulation $\mathcal{T}_h$:

$$Gy(x_j) = \sum_{\tau \in N(x_j)} \frac{|\tau|}{|N(x_j)|} \nabla y_h|_\tau, \tag{3.2}$$

where $N(x_j)$ are the adjacent elements to x_j. The element-wise ZZ estimator is defined by

$$\eta_\tau^{\mathrm{ZZ}} = \|Gy - \nabla y_h\|_{L^2(\tau)}.$$

Showing upper and lower bounds in the sense of (3.1) can be done when assuming Gy is a higher order approximation of ∇y than ∇y_h ([20]). Carstensen and Bartels published an extensive study [7] on averaging techniques, where the ZZ estimator is a subclass of. In their work they prove these bounds in a wider context.

This simple estimator only uses y_h and does not need the problem data. Hence, it seems likely that with more information, especially when the convection term dominates, more accurate estimators are feasible. A class of estimators which incorporates the problem data is based on the residual of the strong partial differential equation.

3.2 Norm-residual based error estimation

There are different approaches to build an error estimator based on the residual of the strong form of the PDE. The indicators introduced in this section incorporate the norm of the residuals, and will therefore be referred to as norm-residual based estimators.

Definition 3.2.1 *Let y_h be a Galerkin or SUPG approximation, then its interior residual is defined by*

$$R^{\text{int}}(y_h) = f + \epsilon \Delta y_h - b \cdot \nabla y_h - c y_h$$

and its edge residual by

$$R^{\text{edge}}(y_h) = \begin{cases} 0 & \text{if } e \in \mathcal{E}_h^D, \\ g - \epsilon \nabla y_h \cdot n & \text{if } e \in \mathcal{E}_h^N, \\ \epsilon \left[\nabla y_h \cdot n \right] & \text{if } e \in \mathcal{E}_h^\Omega, \end{cases}$$

or with slight modification:

$$\widetilde{R}^{\text{edge}}(y_h) = \begin{cases} 0 & \text{if } e \in \mathcal{E}_h^D, \\ g - \epsilon \nabla y_h \cdot n & \text{if } e \in \mathcal{E}_h^N, \\ \frac{\epsilon}{2} \left[\nabla y_h \cdot n \right] & \text{if } e \in \mathcal{E}_h^\Omega. \end{cases}$$

Here $\mathcal{E}_h^\Omega$ are interior edges, $\mathcal{E}_h^D$ are Dirichlet edges and $\mathcal{E}_h^N$ are Neumann edges of $\mathcal{T}_h$.

Let y be the solution to the variational problem, let y_h be the Galerkin or SUPG approximation and define the error by $e_y = y - y_h$. Consider $a(e_h, v)$ for $v \in H_0^1(\Omega)$:

$$\begin{aligned}
a(e_h, v) &= a(y, v) - a(y_h, v) = \ell(v) - a(y_h, v), \\
&= \int_\Omega fv + \int_{\Gamma_N} gv - \left[\int_\Omega \epsilon \nabla y_h \cdot \nabla v + b \cdot \nabla y_h v + c y_h v \right], \\
&= \int_\Omega fv + \int_{\Gamma_N} gv - \sum_{\tau \in \mathcal{T}_h} \left[\int_\tau \epsilon \nabla y_h \cdot \nabla v + b \cdot \nabla y_h v + c y_h v \right], \\
&= \int_\Omega fv + \int_{\Gamma_N} gv - \sum_{\tau \in \mathcal{T}_h} \left[\int_\tau -\epsilon \Delta y_h v + b \cdot \nabla y_h v + c y_h v + \int_{\partial\tau} \epsilon \nabla y_h \cdot n \; v \right].
\end{aligned}$$

Inserting the definition of the residual yields

$$\begin{aligned}
a(e_h, v) &= \left[\sum_{\tau \in \mathcal{T}_h} \int_\tau R_\tau^{\text{int}}(y_h) v \right] + \left[\int_{\Gamma_N} gv - \sum_{\tau \in \mathcal{T}_h} \int_{\partial\tau} \epsilon \nabla y_h \cdot n \; v \right], \\
&= \left[\sum_{\tau \in \mathcal{T}_h} \int_\tau R^{\text{int}}(y_h) v \right] + \left[\sum_{e \in \mathcal{E}_h^N} \int_e gv - \sum_{\tau \in \mathcal{T}_h} \int_{\partial\tau} \epsilon \nabla y_h \cdot n \; v \right], \\
&= \left[\sum_{\tau \in \mathcal{T}_h} \int_\tau R^{\text{int}}(y_h) v \right] + \left[\sum_{e \in \mathcal{E}_h^N} \int_e (g - \epsilon \nabla y_h \cdot n) v \right] - \left[\sum_{e \in \mathcal{E}_h^\Omega} \int_e [\epsilon \nabla y_h \cdot n] \; v \right].
\end{aligned}$$

This leads to the following formula:

$$a(e_h, v) = \sum_{\tau \in \mathcal{T}_h} \int_\tau R^{\text{int}}(y_h) v + \sum_{e \in \mathcal{E}_h} \int_e R^{\text{edge}}(y_h) v \qquad \forall v \in H_0^1(\Omega). \tag{3.3}$$

This equation is fundamental to most residual based error estimators. Before formally defining the norm-residual based error indicators, consider a technical result used to analyze the residual based error estimator. We use the following notation. For a triangle $\tau \in \mathcal{T}_h$ and an edge e in the conforming triangulation $\mathcal{T}_h$, h_τ is the length of

the longest edge of τ, and h_e is the length of e and

$$N(\tau) = \cup \{\mu \in \mathcal{T}_h \ : \ \bar{\mu} \cap \bar{\tau} \neq \emptyset\},$$

$$N(e) = \cup \{\mu \in \mathcal{T}_h \ : \ \bar{\mu} \cap e \neq \emptyset\}.$$

are the neighborhoods of an element τ and edge e (see Figure 3.1).

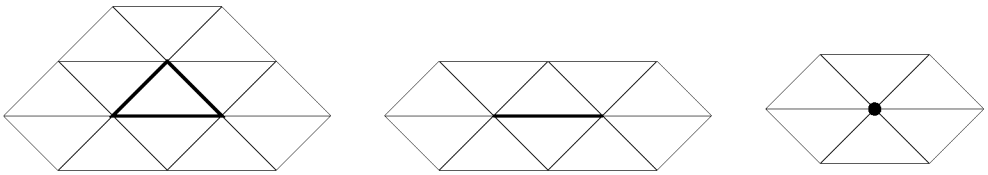

Figure 3.1: The neighborhood of a highlighted triangular element (left), an edge (middle) and a vertex (right), is the union of the displayed elements.

Lemma 3.2.2 *(Verfürth, [34, p.11], [35], [37]) Let $\mathcal{T}_h$ be a conforming triangulation of a domain $\Omega \subset \mathbb{R}^2$. For any arbitrary function $v \in H^1(\Omega)$ there exists a function $I_h v \in Y_h$ such that for all elements $\tau \in \mathcal{T}_h$ and edges $e \in \mathcal{E}_h$ the following is true:*

$$\|v - I_h v\|_{L^2(\tau)} \leq C_3 h_\tau \ \|v\|_{H^1(N(\tau))},$$

$$\|v - I_h v\|_{L^2(e)} \leq C_4 \sqrt{h_e} \ \|v\|_{H^1(N(e))},$$

$$\|I_h v\|_{L^2(\tau)} \leq C_5 \|v\|_{H^1(N(\tau))},$$

where C_3, C_4, C_5 are independent of $\mathcal{T}_h$.

This lemma provides the tools for norm-residual based error estimators, formulated in Definition 3.2.3.

Definition 3.2.3 *The norm-residual based element-wise error estimator in the H^1 sense is defined by*

$$\eta_\tau^{\text{Res}-\text{H1}} \equiv \left[h_\tau^2 \|R^{\text{int}}(y_h)\|^2_{L^2(\tau)} + \sum_{e\in\mathcal{E}(\tau)} h_e \|\widetilde{R}^{\text{edge}}(y_h)\|^2_{L^2(e)} \right]^{1/2}.$$

The element-wise estimator in the L^2 norm is defined by

$$\eta_\tau^{\text{Res}-\text{L2}} \equiv \left[h_\tau^4 \|R^{\text{int}}(y_h)\|^2_{L^2(\tau)} + \sum_{e\in\mathcal{E}(\tau)} h_e^3 \|\widetilde{R}^{\text{edge}}(y_h)\|^2_{L^2(e)} \right]^{1/2}.$$

Moreover, the global error estimates are defined by:

$$\eta^{\text{Res}-\text{H1}} = \left[\sum_{\tau\in\mathcal{T}_h} \left(\eta_\tau^{\text{Res}-\text{H1}}\right)^2 \right], \text{ and } \qquad \eta^{\text{Res}-\text{L2}} = \left[\sum_{\tau\in\mathcal{T}_h} \left(\eta_\tau^{\text{Res}-\text{L2}}\right)^2 \right].$$

The next result shows that the norm-residual based error estimators satisfy (3.1a) - (3.1c).

Theorem 3.2.4 *(Verfürth, [37, p.1772-1773]) Let $\eta^{\text{Res}-\text{H1}}$ be as in Definition 3.2.3, let y be the solution to the variational problem (2.4) and let y_h be the SUPG approximation from (2.12), then there exists constants $C_9, C_{10} > 0$ such that:*

$$\|y - y_h\|_\epsilon + \|b\cdot\nabla(y-y)h\|_* \leq \frac{C_9}{\sqrt{\epsilon}}\, \eta_\tau^{\text{Res}-\text{H1}},$$

$$\eta_\tau^{\text{Res}-\text{H1}} \leq C_{10} \left[\, \|y - y_h\|_\epsilon + \|b\cdot\nabla(y-y)h\|_* \, \right].$$

Proof: Let y be the exact solution to the variational problem, let y_h be the SUPG approximation and define $e_h = y - y_h$. For a function $v \in H^1_D(\Omega)$ let its interpolant

$I_h v \in Y_h^0$ be defined by Lemma 3.2.2. Then,

$$
\begin{aligned}
a(e_h, v) &= a(e_h, v - I_h v) + a(e_h, I_h v) \\
&= a(e_h, v - I_h v) + a(y, I_h v) - a(y_h, I_h v) \\
&= a(e_h, v - I_h v) + \ell(I_h v) - \ell(I_h v) - \sum_{\tau \in \mathcal{T}_h} \delta_\tau (f + \epsilon \Delta y_h - b \cdot \nabla y_h - c y_h, b \cdot I_h v)_\tau \\
&= a(e_h, v - I_h v) - \sum_{\tau \in \mathcal{T}_h} \delta_\tau (R^{\text{int}}(y_h), b \cdot I_h v)_\tau
\end{aligned}
$$

Replacing $a(e_h, v - I_h v)$ by the identity of equation (3.3) yields:

$$
\begin{aligned}
a(e_h, v) = &\left[\sum_{\tau \in \mathcal{T}_h} \int_\tau R^{\text{int}}(y_h)\ (v - I_h v) \right] + \left[\sum_{e \in \mathcal{E}_h} \int_e R^{\text{edge}}(y_h)\ (v - I_h v) \right] \\
&- \left[\sum_{\tau \in \mathcal{T}_h} \delta_\tau \int_\tau R^{\text{int}}(y_h)\ (b \cdot I_h v) \right], \\
\leq &\left[\sum_{\tau \in \mathcal{T}_h} \| R^{\text{int}}(y_h) \|_{L^2(\tau)} \| v - I_h v \|_{L^2(\tau)} \right] + \left[\sum_{e \in \mathcal{E}_h} \| R^{\text{edge}}(y_h) \|_{L^2(e)} \| v - I_h v \|_{L^2(e)} \right] \\
&+ \left[\sum_{\tau \in \mathcal{T}_h} \delta_\tau \| R^{\text{int}} \|_{L^2(\tau)} \| b \cdot I_h v \|_{L^2(\tau)} \right]
\end{aligned}
$$

Because $\| b \cdot I_h v \|_{L^2(\tau)} \leq \| b \|_{\infty, \tau} \| I_h v \|_{L^2(\tau)}$ and Lemma 3.2, this inequality becomes:

$$
\begin{aligned}
a(e_h, v) \leq &\sum_{\tau \in \mathcal{T}_h} C_3 h_\tau \| R^{\text{int}}(y_h) \|_{L^2(\tau)} \| v \|_{H^1(N(\tau))} + \sum_{e \in \mathcal{E}_h} C_4 \sqrt{h_e} \| R^{\text{edge}}(y_h) \|_{L^2(e)} \| v \|_{H^1(N(e))} \\
&+ \sum_{\tau \in \mathcal{T}_h} C_5 \delta_\tau \| R^{\text{int}}(y_h) \|_{L^2(\tau)} \| v \|_{H^1(N(\tau))} \| b \|_{\infty, \tau}\ ,
\end{aligned}
$$

with $C_3, C_4 > 0$ from Lemma 3.2.2 and $C_5 > 0$. Recall that $\delta_\tau \leq \frac{h_\tau}{2\|b\|_{\infty,\tau}}$ by equation (2.13). Note that the terms $\|b\|_{\infty,\tau}$ cancel out, which leads to

$$\begin{aligned} a(e_h, v) &\leq \left[(C_3 + \frac{1}{2}C_5) \sum_{\tau \in \mathcal{T}_h} h_\tau^2 \|R^{\text{int}}(y_h)\|^2_{L^2(\tau)} + C_4 \sum_{e \in \mathcal{E}_h} h_e \|R^{\text{edge}}(y_h)\|^2_{L^2(e)} \right]^{1/2} \\ &\qquad \times \left[\sum_{\tau \in \mathcal{T}_h} \|v\|^2_{H^1(N(\tau))} \right]^{1/2}, \\ &\leq C_6 \|v\|_{H^1(\Omega)} \left[\sum_{\tau \in \mathcal{T}_h} h_\tau^2 \|R^{\text{int}}(y_h)\|^2_{L^2(\tau)} + \sum_{e \in \mathcal{E}_h} h_e \|R^{\text{edge}}(y_h)\|^2_{L^2(e)} \right]^{1/2}, \end{aligned}$$

where $C_7 = C_6 \ \max\{C_3 + \frac{1}{2}C_5, C_4\}^{1/2}$ and C_6 is defined by inequality

$$\|v\|_{H^1(\Omega)} \leq C_6 \sum_{\tau \in \mathcal{T}_h} \|v\|_{H^1(\tau)},$$

Verfürth [34, p.13] proves that this constant only depends on the smallest angle in the mesh. This leads to

$$\frac{a(e_h, v)}{\|v\|_\epsilon} \leq \frac{C_7}{\sqrt{\epsilon}} \, \eta^{\text{Res}-\text{H1}}, \qquad \forall v \in H^1_D(\Omega).$$

Select $\widetilde{v} \in H^1_D \backslash \{0\}$ such that the supremum

$$\sup_{v \in H^1_D \backslash \{0\}} \frac{a(e_h, v)}{\|v\|_\epsilon}$$

is attained. Verfürth proves [37, l.3.1] there exist a $C_8 > 0$ such that $\forall w \in H^1_D(\Omega)$

$$\frac{a(w, \widetilde{v})}{\|\widetilde{v}\|_\epsilon} \geq C_8 \ [\|w\|_\epsilon + \|b \cdot \nabla w\|_*].$$

Setting $w = y - y_h$ results to the upper bound of the theorem (with $C_9 = C_7/C_8$):

$$\|y - y_h\|_\epsilon + \|b \cdot \nabla (y - y_h)\|_* \leq \frac{C_9}{\sqrt{\epsilon}} \, \eta^{\text{Res}-\text{H1}}.$$

The proof of the lower bound in the theorem is beyond the scope of this thesis, but is published by Verfürth [37, p.1774-1776]. □

Remark 3.2.5 *Similar bounds (as those shown in Theorem 3.2.4) hold for the Res-L2 estimator of Definition 3.2.3. Their derivation is based on a duality technique ([4], [20]), and is beyond the scope of this thesis.*

How well the Res-H1 and Res-L2 estimators approximate the true error depends heavily on the problem dependent constants C_3, C_4, C_5, C_6. These four constants are combined in C_7, which implicitly assumes that both interior and edge residual have equal weights. Therefore combining interior and edge residual has to be done explicitly in this approach.

Another, maybe even more compromising result of Theorem 3.2.4, is that the upper bound of the error estimator is inversely dependent on ϵ. Hence, for the convection-dominated equations studied in this work, the error might be grossly over-estimated.

A different, maybe more natural approach is the use of a local Neumann estimator. This indicator combines the residuals implicitly, and could potentially overcome the issues mentioned above.

3.3 Estimators based on solving local Neumann problems

Though this estimator is also based on the residual, it takes quite a different approach than the previous section. Recall equation (3.3) in a slightly different form:

$$\sum_{\tau\in\mathcal{T}_h} a(e_h, v)_\tau = \sum_{\tau\in\mathcal{T}_h} \left[\int_\tau R^{\text{int}}(y_h)v + \sum_{e\in\mathcal{E}(\tau)} \int_e \widetilde{R}^{\text{edge}}(y_h)v \right] \qquad \forall v \in H_0^1(\Omega),$$

where $a(\cdot,\cdot)_\tau$ is the standard bilinear form of (2.5) restricted to element τ. Removing the sums from both sides, results in a local Neumann problem, with the error as solution.

Definition 3.3.1 *Let $Q_\tau \subset H^1(\tau)$ be finite dimensional. If $\widehat{e}_\tau \in Q$ is the solution of*

$$a(\widehat{e}_\tau, v)_\tau = \int_\tau R^{\text{int}}(y_h)v + \sum_{e\in\mathcal{E}(\tau)} \int_e \widetilde{R}^{\text{edge}}(y_h)v \qquad \forall v \in Q_\tau^0 \tag{3.4}$$

then the element-wise local Neumann estimators in H^1 and L^2 are defined by

$$\eta_\tau^{\text{Neu-H1}} \equiv \|\widehat{e}_\tau\|_{H^1(\tau)}, \quad \eta_\tau^{\text{Neu-L2}} \equiv \|\widehat{e}_\tau\|_{L^2(\tau)}.$$

and the global estimator are defined by

$$\eta^{\text{Neu-H1}} = \left(\sum_{\tau\in\mathcal{T}_h} \left(\eta_\tau^{\text{Neu-H1}}\right)^2 \right)^{1/2}, \quad \eta^{\text{Neu-L2}} = \left(\sum_{\tau\in\mathcal{T}_h} \left(\eta_\tau^{\text{Neu-L2}}\right)^2 \right)^{1/2}.$$

Solving the local Neumann problem in Definition 3.3.1 requires a finite dimensional subspace Q_τ of $H^1(\tau)$. Verfürth [34] proposes the following four bubble functions on a reference triangle (see Figure 3.2):

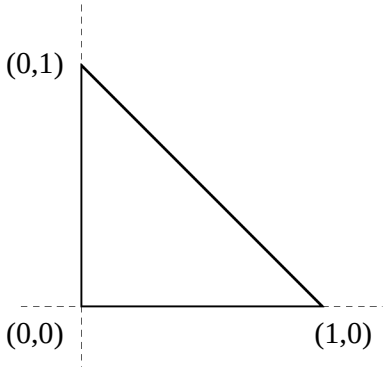

Figure 3.2: A reference triangle.

$$\psi_1(x_1, x_2) = 4x_1x_2, \tag{3.5a}$$

$$\psi_2(x_1, x_2) = 4x_1(1 - x_1 - x_2), \tag{3.5b}$$

$$\psi_3(x_1, x_2) = 4x_2(1 - x_1 - x_2), \tag{3.5c}$$

$$\psi_4(x_1, x_2) = 27x_1x_2(1 - x_1 - x_2). \tag{3.5d}$$

The solution to (3.4) can be approximated from the span of these basis functions $\{\psi_i\}$ mapped to element τ. This leads to a 4×4 linear system that needs to be solved for each element.

The details of the error bounds (3.1) for these error estimators are beyond the scope of this thesis, but have extensively been studied by Verfürth ([34], [35]). He recently extended this work [37] in the context of the bounds described in Theorem 3.2.4.

Elman et al. [14] reason that the $b \cdot \nabla uv$ and cuv terms can be dropped from the restricted bilinear form $a(\cdot, \cdot)_\tau$ without significant loss of accuracy. Kay and Silvester [23] also adopt such strategy as extension of the work of Verfürth [35].

However, dropping these terms does not yield a significant computational advantage, and therefore they are included in this work.

Note that combining interior and edge residual and the diffusion constant ϵ is done implicitly. Hence, there are no constants that need to be estimated. This implicit character is a considerable benefit of this estimator in theory.

Chapter 4

Numerics I: Single convection diffusion equations

This chapter contains numerical results for single convection diffusion equations. Its aim is to illustrate the theory introduced in the previous chapters. Papastravrou and Verfürth [30], as well as John [20], have published numerical studies on AFEM combined with SUPG for convection dominated diffusion equations. The shown results in this chapter are compared to their findings.

4.1 Notes on numerical experiments

The theory introduced so far has been focused on the SUPG approximation and a posteriori error estimation. To implement AFEMs, elements need to be selected for refinement and the mesh must be refined based on this selection and in such a way

that the refined mesh is conforming. Strategies to excecute these tasks are discussed on this section.

Let y be the solution to the variational problem and let y_h be the Galerkin or the SUPG approximation. The steps of an adaptive finite element method are usually repeated until some stopping criterion is reached. There are two options, given some tolerance TOL and number of nodes $N_{\max}$, either

1. $\|y - y_h\| < TOL$, or

2. $\mathcal{T}_h$ has more than $N_{\max}$ nodes.

Since y is commonly not know, in many applications one sets $N_{\max}$ in advance and uses the second rule. Of course, if η is an indicator which estimates the error in some norm $\|\cdot\|$, also $\eta < TOL$ could be a stopping criterion. However, this assumes that η is a good approximation of $\|y - y_h\|$. If the exact solution is known, one can compute the effectivity index

$$I_{\text{eff}}(\eta) \equiv \frac{\eta}{\|y - y_h\|}$$

to see how well the estimator approximates the true error.

The work of Bänsch ([5], [6]) is the basis of many local refinement algorithms for triangular meshes (for refinement of tetrahedral meshes see [1]). The mesh refinement used in this thesis is based on Bänsch's ideas. The implementation used in this work is derived from AFEM@MATLAB from Chen [12]. This software package is the predecessor of iFEM [10], which contains an advanced refinement tool.

Let the current triangulation $\mathcal{T}_h$ contain N_t elements. After a local error estimator η_τ has been obtained for each element τ, some elements need to be selected for refinement. Many options have been proposed ([5], [10], [12], [34]), however this work only uses one. First order the triangles τ_k such that $\eta_{\tau_1}, \eta_{\tau_2}, \ldots, \eta_{\tau_{N_t}}$ is descending. Let $\theta \in (0,1)$ (typically $\theta \approx .2$), refine elements τ_k, $k = 1, \ldots, \lfloor \theta N_t \rfloor$. The main benefit of this approach is that each refinement will have a guaranteed percentage of increase in elements. For a numerical comparison of different marking strategies see [30].

4.2 Numerical examples

The examples studied in this chapter are more commonly used in the literature. Example 1A shows how a boundary layer is reseloved by local refinement. In the second example, 1B, an inner layer needs to resolved, which in this case is slightly more complex than the first example. Finally, in Example 1C, an inner and boundary layer need to resolved.

4.2.1 Example 1A: Boundary layer

This example originates from work by Collis, Heinkenschloss and Leykekhman ([13], [15]). Consider the function

$$\eta(x) = x - \frac{\exp((x-1)/\epsilon) - \exp(-1/\epsilon)}{1 - \exp(-1/\epsilon)}. \tag{4.1}$$

$\epsilon = 10^{-1}$ $\epsilon = 10^{-2}$ $\epsilon = 10^{-3}$

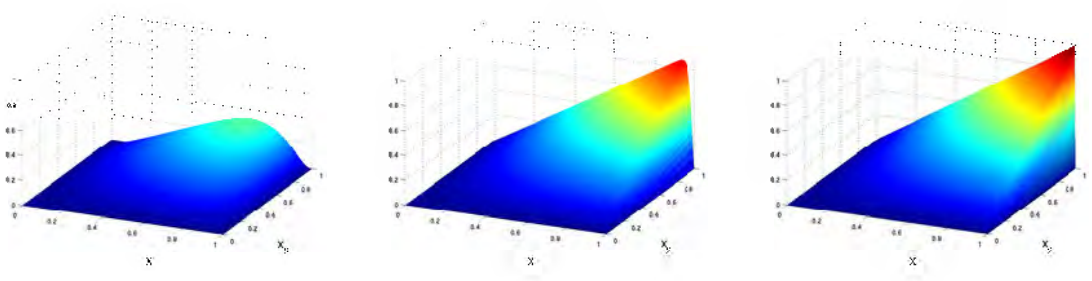

Figure 4.1: Example 1A: The effect of increasing convection on a solution ($\theta = 45^o$) containing a boundary layer ($\epsilon = 10^{-1}, 10^{-2}, 10^{-3}$), surface plots.

Let $\Omega = (0,1)^2$, $\theta = 45^o$, $b = (\cos\theta, \sin\theta)$ and $c = 0$ and let the true solution be:

$$y(x_1, x_2) = \eta(x_1)\ \eta(x_2).$$

The right hand side can now be determined by inserting these identities in (2.1), which leads to:

$$\begin{aligned} f(x_1, x_2) &= -\epsilon \Delta y(x_1, x_2) + b \cdot \Delta y(x_1, x_2), \\ &= -\epsilon\,[\ \eta''(x_1)\eta(x_2) + \eta(x_1)\eta''(x_2)\ ] + [\ \cos\theta\ \eta'(x_1)\eta(x_2) + \sin\theta\ \eta(x_1)\eta'(x_2)\ ]\,. \end{aligned}$$

Figure 4.1 shows the exact solution for various diffusion parameters ϵ. As ϵ becomes smaller, a boundary layer forms on the boundary $\{x \in \partial\Omega\ :\ x_1 = 1 \text{ or } x_2 = 1\}$. The region in which the gradient of y is large becomes the smaller the smaller ϵ. Ideally, an adaptive refinement procedure would pick up these layers, and refine around the upper and right boundary.

Applying the Neu-L2 estimator (others yield very similar results), yields numerical

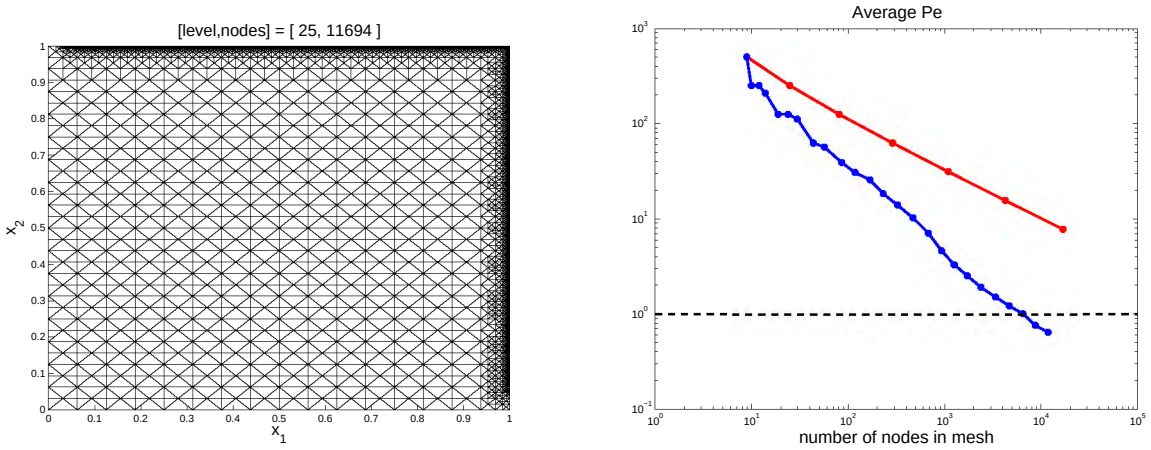

Figure 4.2: Example 1A with $\epsilon = 10^{-3}$: Mesh generated by the Neu-L2 estimator (left) with roughly 10^5 nodes. The figure on the right displays the average mesh Peclet numbers around the boundary layer.

solutions on a sequence of refined meshes. The refined mesh, after $N_{\max} = 10^5$ has been reached, is shown in Figure 4.2. Note how the refinement picks out the boundary layer and places all its nodes in this region. This figure also shows the average mesh Peclet number around the boundary layer throughout the refinement. Uniform refinement, indicated by the red line, yields a straight line in this figure. The reason is that ϵ and $\|b\|_\infty$ are constant in this example, and hence $\mathcal{P}_\tau$ is constant in the mesh and is linearly dependent on h_τ.

Figure 4.2 shows that local refinement reduces the mesh Peclet numbers around the boundary layer much faster than the uniform refinement, which is exaclty to be expected, since the locally refined meshes have many more elements around the layer.

A few of the computed solutions for $\epsilon = 10^{-3}$ are shown in Figure 4.3. The

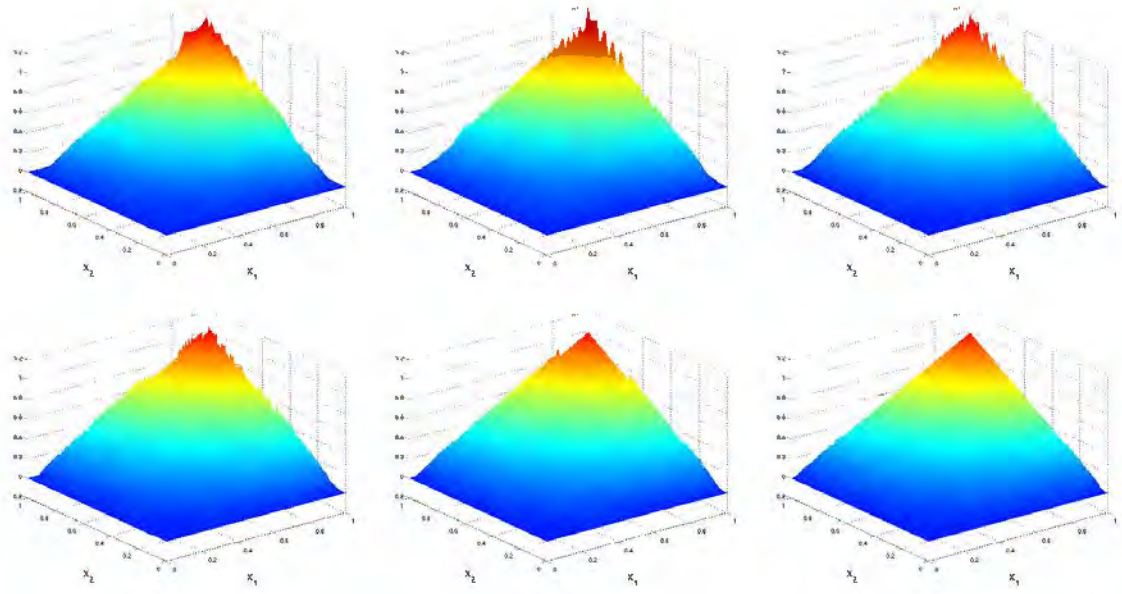

Figure 4.3: Example 1A with $\epsilon = 10^{-3}$: Surface plots of the SUPG solution after $N_{\max} \approx 200, 400, 800, 1600, 3200, 6400$ (one value of $N_{\max}$ from left to right).

SUPG solution approximates the true solution well away from the boundary layer on all meshes. On coarser meshes, the solution in the boundary layer cannot be resolved. The computed SUPG solution exhibits small oscillations perpendicular to the direction of convection. After more refinements, the mesh size around the layer is sufficiently small so that the boundary layer can be resolved and the spurious oscillations near the boundary layer are reduced..

Figure 4.4 shows the H^1 and L^2 errors during the refinements. Note that initially the global H^1 errors go up, which is do to the oscillations at the boundary. These oscillations dominate the global H^1 error. It seems likely that the H^1 global error restricted to the region away from the boundary layer is much smaller. Interesting to

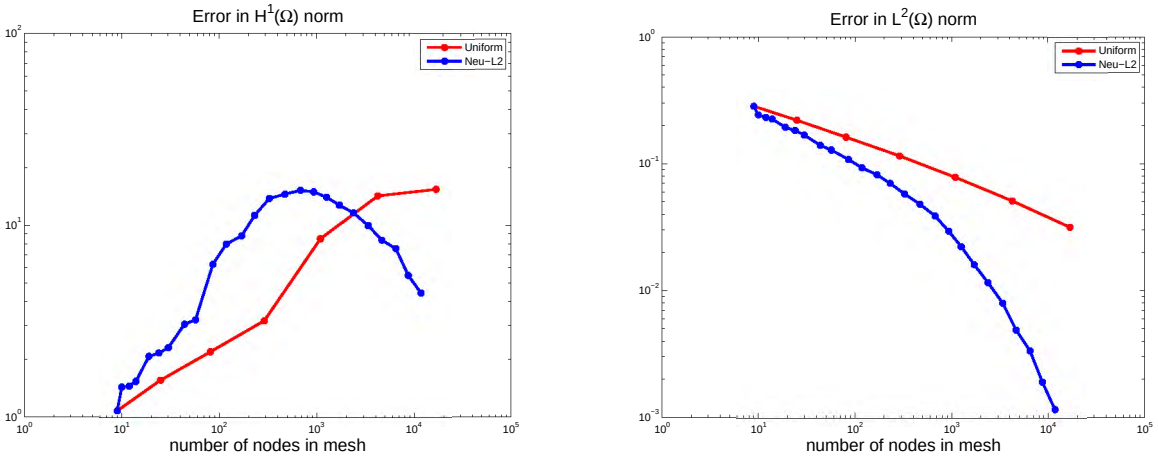

Figure 4.4: Example 1A with $\epsilon = 10^{-3}$: Errors of uniform refinement (red) and local (blue) refinement using Neu-L2 estimator, measured in $H^1(\Omega)$ (left) and $L^2(\Omega)$ (right) norm.

see is the convergence of the L^2 norm of the error. The better the boundary layer is resolved, i.e. the lower the mesh Peclet numbers around the boundary, the faster the convergence.

The localized behavior of the oscillations becomes even more evident when looking at a cross section of the numerical solution. Figures 4.5 and 4.6 show such a cross section along the line $x_1 = x_2$. This cross section is along the convection flow. The SUPG solution approximates the true solution very well away from the boundary layer on all meshes. Unless the mesh is sufficiently refined, the SUPG solution cannot resolve the boundary layer. Eventually, local refinement enables an excellent approximation of the solution in the boundary layer.

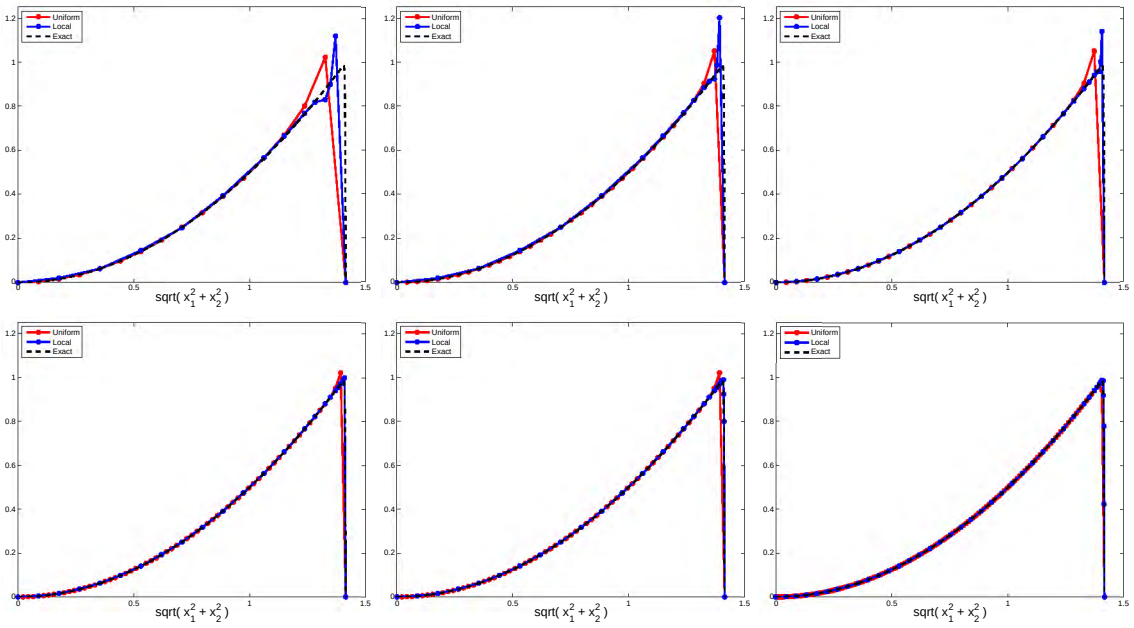

Figure 4.5: Example 1A with $\epsilon = 10^{-3}$: Cross section of the solution along the constant convection flow after $N_{\max} \approx 200, 400, 800, 1600, 3200, 6400$ (one value of $N_{\max}$ from left to right). Displayed lines are the true solution (black), the solution with uniform refinement (red) and the local refinement solution (blue).

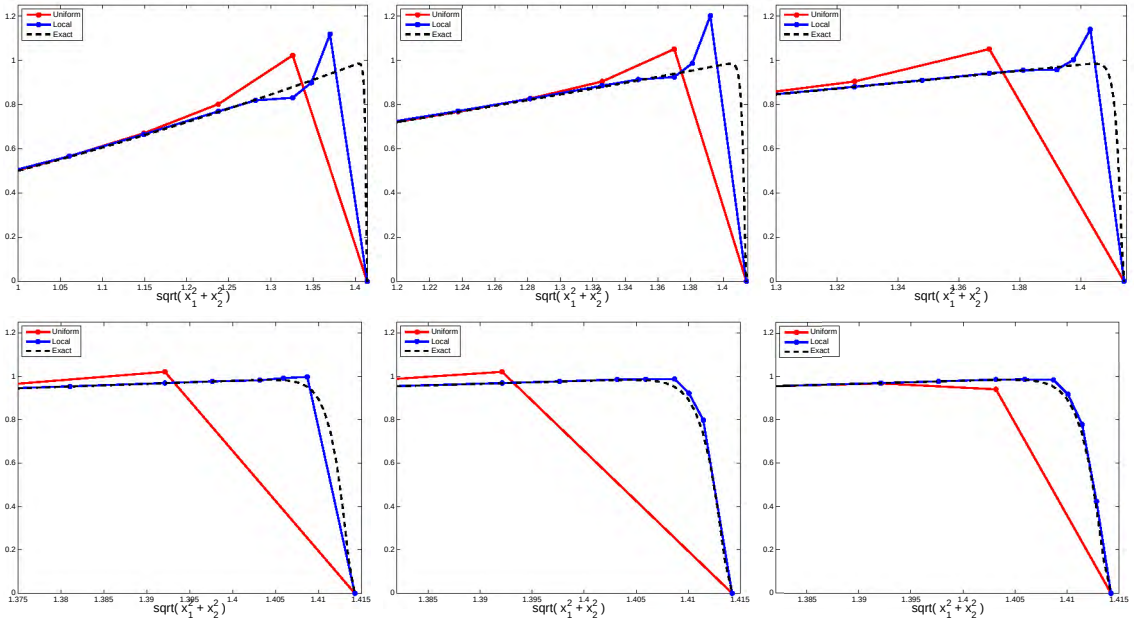

Figure 4.6: Example 1A with $\epsilon = 10^{-3}$: Zoomed cross section (around the boundary layer) of the solution along the constant convection flow after $N_{\max} \approx 200, 400, 800, 1600, 3200, 6400$ (one value of $N_{\max}$ from left to right). Displayed lines are the true solution (black), the solution with uniform refinement (red) and the local refinement solution (blue).

4.2.2 Example 1B: Circular inner layer

The convection-diffusion equation of this example has been studied by John [20] and in optimal control setting by Hinze, Yan and Zhou [17]. Again, in this example the domain is defined by $\Omega = (0,1)^2$. Let $b = -(2,3)$, let $c = 1$, and let the true solution be

$$\begin{aligned} y(x_1,x_2) =& 16x_1(1-x_1)x_2(1-x_2) \\ & \times \left(\frac{1}{2} + \frac{1}{\pi}\arctan\left[\frac{2}{\sqrt{\epsilon}}\left(\frac{1}{16} - \left(x_1 - \frac{1}{2}\right)^2 - \left(x_2 - \frac{1}{2}\right)^2\right)\right]\right). \end{aligned}$$

In this problem there is only a Dirichlet boundary, hence, $\Gamma_D = \partial\Omega$, $\Gamma_N = \emptyset$ (note $y(x_1,x_2) = 0$ on $\partial\Omega$). The right hand side is now determined by

$$f(x_1,x_2) = -\epsilon\Delta y(x_1,x_2) - 2\frac{\partial}{\partial x_1}y(x_1,x_2) - 3\frac{\partial}{\partial x_2}y(x_1,x_2) + y(x_1,x_2).$$

Figure 4.7 shows the exact solution for various diffusion parameters ϵ. As ϵ becomes smaller, an inner layer forms.

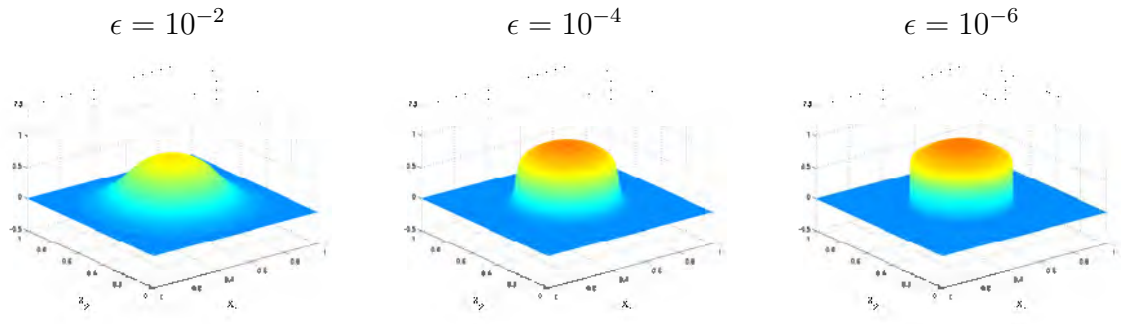

Figure 4.7: Example 1B: Surface plots of the analytic solution for diffusion parameters $\epsilon = 10^{-2}, 10^{-4}, 10^{-6}$. As ϵ becomes smaller, a circular inner layer forms.

Solving this problem on a uniform mesh gets increasingly difficult as shown in

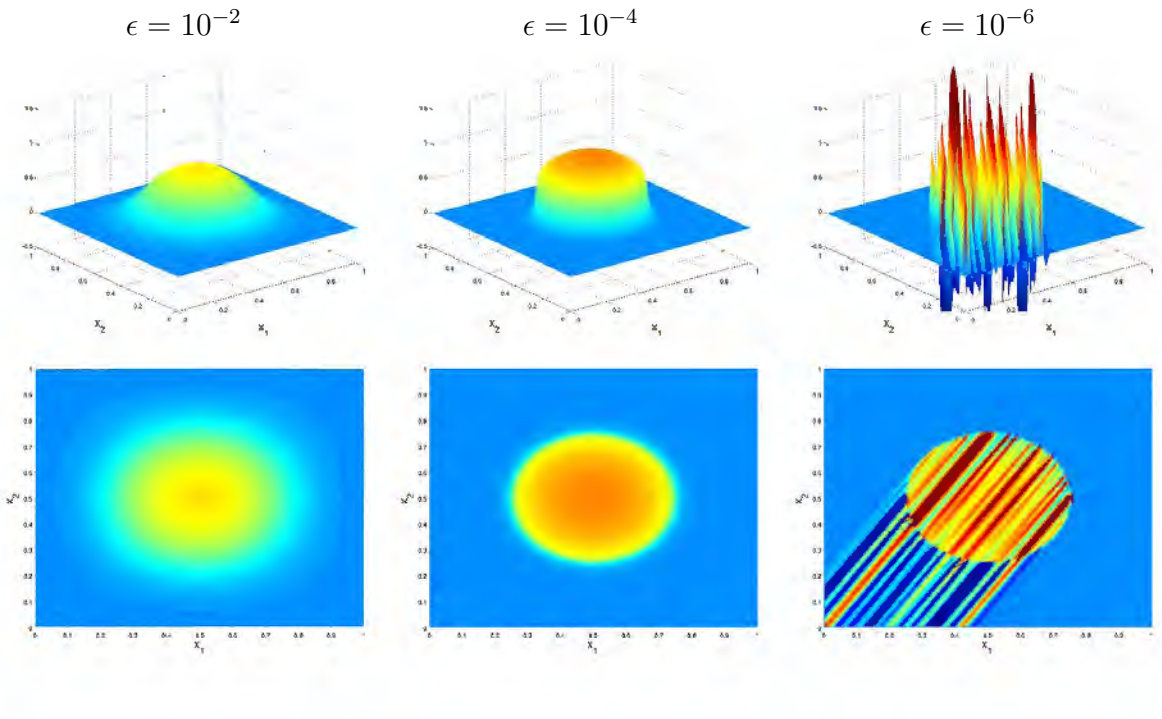

Figure 4.8: Example 1B: Solution on a uniform mesh with (approximately 16641 nodes) of a problem with circular interior layer for $\epsilon = 10^{-2}, 10^{-4}, 10^{-6}$ (fixed ϵ for each column). Top row are surface plots, bottom row are top-down views.

Figure 4.8. When $\epsilon = 10^{-6}$ the SUPG solution with a mesh of about 16641 nodes (129×129) cannot resolve the inner layer. The solution exhibits spurious oscillations perpendicular to the direction of convection near the inner layer. These oscillations are also convected through part of the domain. This transport of the error due to the strong convection is a main difference between this example and Example 1A. Any error in the layer is propagated downwind from the layer. Since Example 1A only had a boundary layer errors in the layer could not pollute the solution into the domain. In the present example, the solution downwind from the layer is polluted. This behavior will cause significant differences in the performance of the error estimators.

The propagation of errors from the circular inner layer into the interior of the domain has drastic implications for adaptive finite element methods. To compute a good approximation of the solution, the mesh near the inner layer needs to be refined. Only if the circular inner layer can be resolved there will be no propagation of errors and the resulting errors downwind are reduced. However, the errors downwind from the inner layer cause the numerical solution to have a large gradient and residual downstream. This will cause most error estimators to refine the mesh away from the inner layer, thus trying to cure the symptoms but not the cause.

Figure 4.9 shows the meshes of refinements driven by the ZZ, Res-H1, Res-L2, Neu-H1 and Neu-L2 estimators. For $\epsilon = 10^{-2}, 10^{-4}$ all estimators appear to generate satisfactory meshes. In the third case, when $\epsilon = 10^{-6}$, the effect mentioned above seems to influence the result. Especially the ZZ estimator places a higher emphasis on the propagated errors than on the oscillations at the layer. Since the ZZ estimator only uses regularity information this seems logical. The propagated oscillations downwind have steep gradients and will therefore yield high values in $\|\nabla y - \nabla y_h\|_{L^2(\Omega)}$.

The residual based estimators also seem to be influenced by this effect, although to a lesser degree. The local Neumann estimators do generate the expected meshes. However, in his numerical study, John [20] also considers $\epsilon = 10^{-8}$ and in this case none of the error estimators, not even the norm-residual and the local Neumann estimators, deliver a satisfactory result.

For more insight into the errors see Figure 4.10. The top two rows contain the H^1

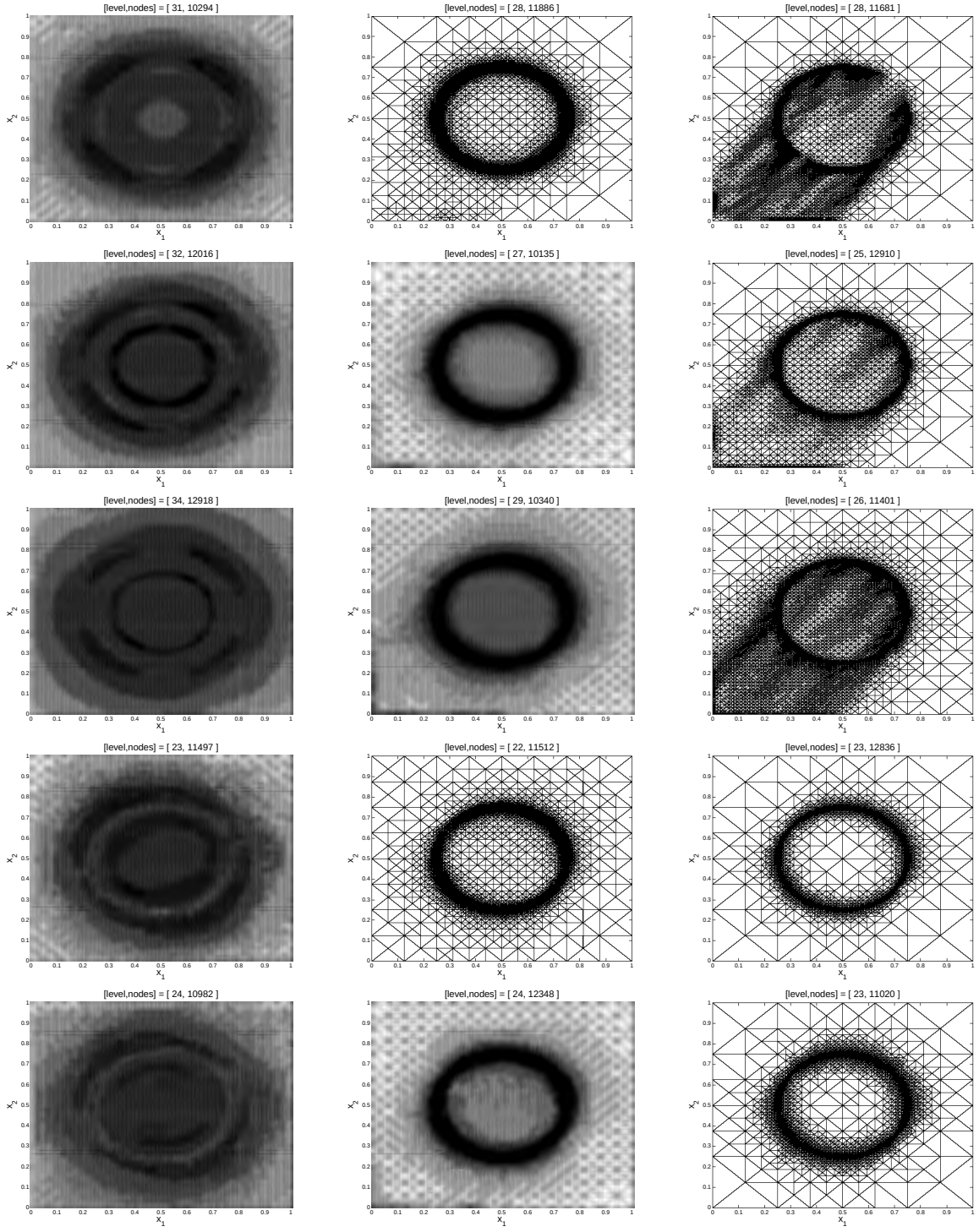

Figure 4.9: Example 1B: Generated locally refined meshes ($N_{\max} = 12000$) for $\epsilon = 10^{-2}, 10^{-4}, 10^{-6}$ (fixed ϵ for each column). Row 1: ZZ-meshes, row 2: Res-H1 meshes, row 3: Res-L2 meshes, row 4: Neu-H1 meshes, row 5: Neu-L2 meshes.

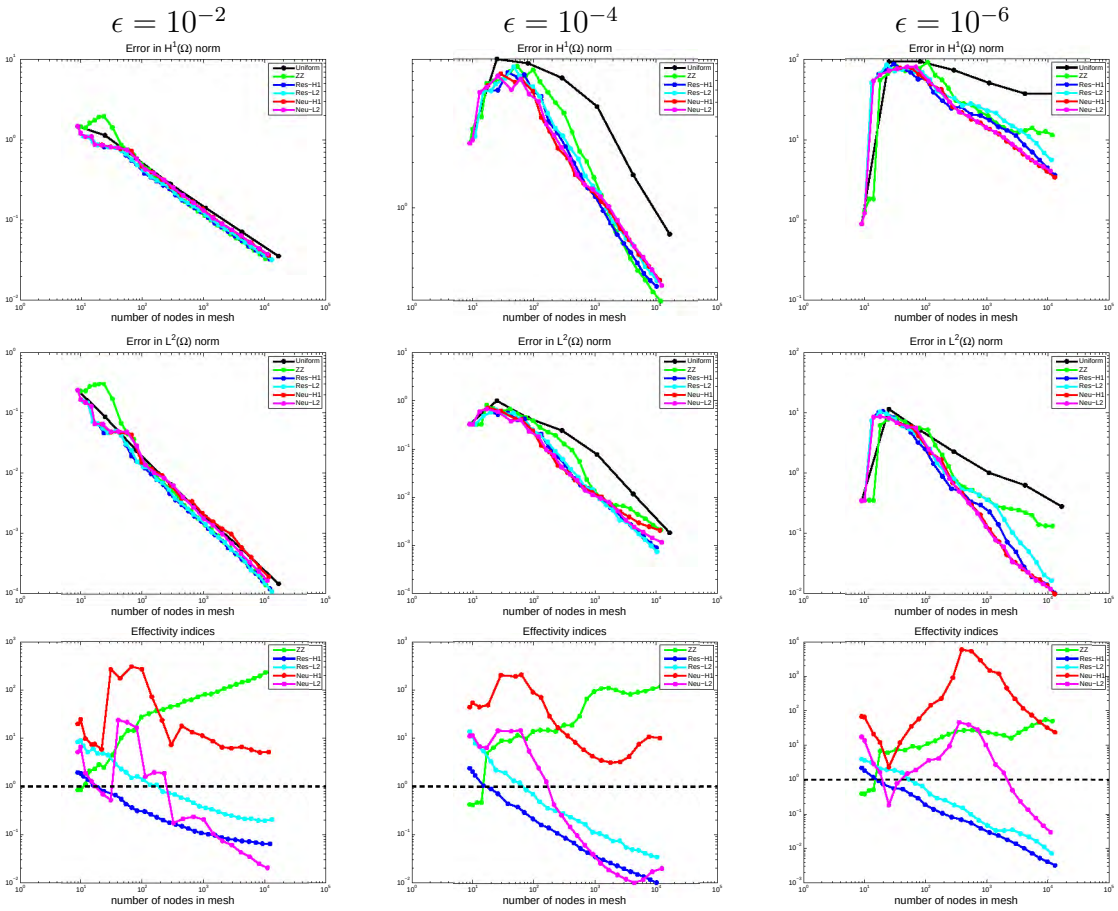

Figure 4.10: Example 1B: Errors of uniform and local refinement using ZZ, Res-H1, Res-L2, Neu-H1 and Neu-L2 estimator for $\epsilon = 10^{-2}, 10^{-4}, 10^{-6}$ (fixed ϵ for each column), measured in $H^1(\Omega)$ (top row) and $L^2(\Omega)$ (middle row) norm. The final row are the effectivity indices of the estimators.

and L^2 error of the computed solutions during refinement. The black line indicates uniform refinement. Note that for $\epsilon = 10^{-2}$ the solution computed with uniform refinement does not appear to be much worse than the solutions computed with locally refinement. Observe the computed meshes for $\epsilon = 10^{-2}$ shown in the left column in Figure 4.9. The elements in these meshes are not localized, and therefore close to uniform refinement. When ϵ decreases the differences become more apparent.

The third row in Figure 4.10 show the effectivity indices of the global error estimators. Ideally, one would like the effectivity indices to converge to one. At first sight there does not seem to be any convergence in the indices, especially not to the desired case of $I_{\text{eff}} = 1$ (the dotted line). John [20] does observe convergence in his experiments on this example with $\epsilon = 10^{-6}$ and $\epsilon = 10^{-8}$ when the number of nodes is approximately 10^4 to 10^5. The number of nodes in the meshes used in this thesis is lower. Papastravrou and Verfürth do use small mesh size (on the order of 10^3) for experiments with $\epsilon = 10^{-2}$ and also observe fluctuations in effectivity indices comparable to the ones shown in the lower left plot of Figure 4.10. Elman et al. [14, p.140-148] show that the effectivity index for the local Neumann estimator is poluted by the errors the layers. When they use Shishkin meshes, which can resolve the boundary layer, the effectivity index of the local Neumann estimator on these special meshes is close to one.

4.2.3 Example 1C: Interior and boundary layer

For the next example is constructed in such a way that the solution exhibits an interior layer and a boundary layer. The analytical solution is not known. This example is adapted from Heinkenschloss and Leykekhman [15] .

Consider the problem data $\Omega = (0,1)^2$, $\theta = 47.3^o$, $b = (\cos\theta, \sin\theta)$, $c = 0$, $f = 0$. Dirichlet conditions re imposed on the entire boundary (i.e. $\Gamma_D = \partial\Omega$, $\Gamma_N = \emptyset$) and the boundary data are

$$g_D(x_1, x_2) = \begin{cases} 1 & \text{if } x_1 = 0 \text{ and } x_2 \leq .25 \\ 1 & \text{if } x_2 = 0 \\ 0 & \text{else.} \end{cases}$$

Figure 4.11 shows a sketch of the problem data.

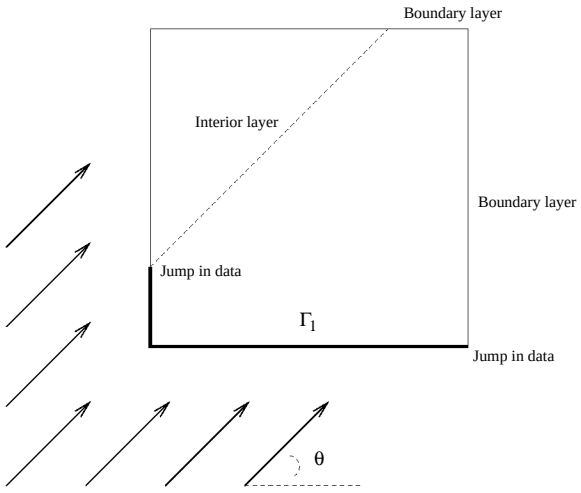

Figure 4.11: Example 1C: Sketch of the problem data.

Depending on the strength of the convection (i.e. the value of ϵ) the solution will have a boundary layer at the outflow region and a straight inner layer. The

SUPG solution is a good approximation of the true solution away from the inner and the boundary layer. If the inner layer or the boundary layer are not resolved, then the errors due the unresolved layers are propagated essentially along the direction of convection. In this case, unlike Example 1B, the errors in the interior layer will be propagated along the interior layer. Thus, large errors in the SUPG solution only occur in a band around the interior layer and around the boundary layer.

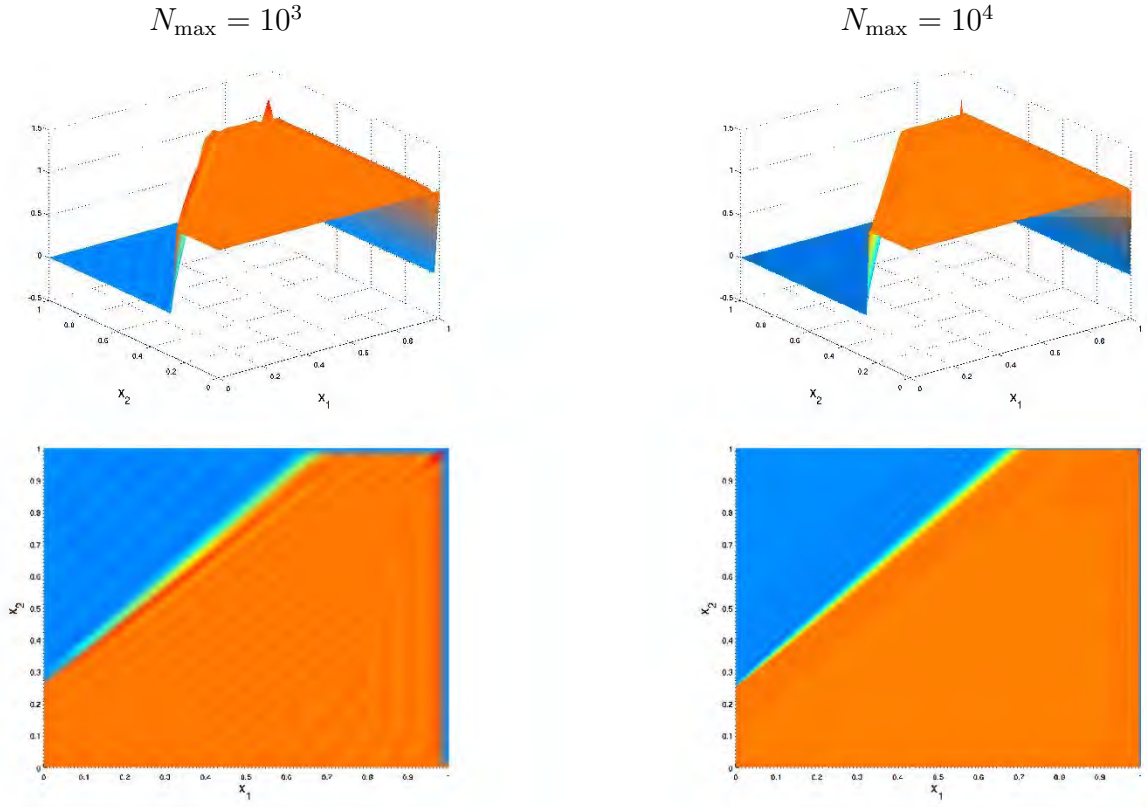

Figure 4.12: Example 1C with $\epsilon = 10^{-4}$: SUPG solution on uniform meshes with 1089 nodes (left) and 16641 nodes (right). Surfaces (top row) and top down views (bottom row).

The SUPG solutions for Example 1C with $\epsilon = 10^{-4}$ computed on uniform meshes

are shown in Figure 4.12. Note that the boundary layer is much steeper than the inner layer.

Figure 4.13 shows the locally refined meshes with roughly 10^3 nodes (left column) and 10^4 nodes (middle column). Note that all estimators initially put the emphasis on the boundary layer; initially only a few elements are refined along the inner layer. Once the boundary layer is sufficiently resolved, the elements along the inner layer are refined.

Even with a locally refined mesh of approximately 10^4 nodes, the numerical solutions still display oscillations in the boundary layer. The H^1 error is dominated by the error in the boundary layer. Therefore the Neu-H1 (and in lesser degree the Res-H1) estimator only refines in this area, completely neglecting the inner layer. This observation was also made by Kay and Silvester [23]. The Neu-L2 and Res-l2 estimator do pick out the interior layer, because the L^2 error is not dominated by the oscillations on the boundary layer. In this case with multiple layers, it is not clear how to combine H^1 and L^2 error information in order to resolve all layers equally at all refinement levels.

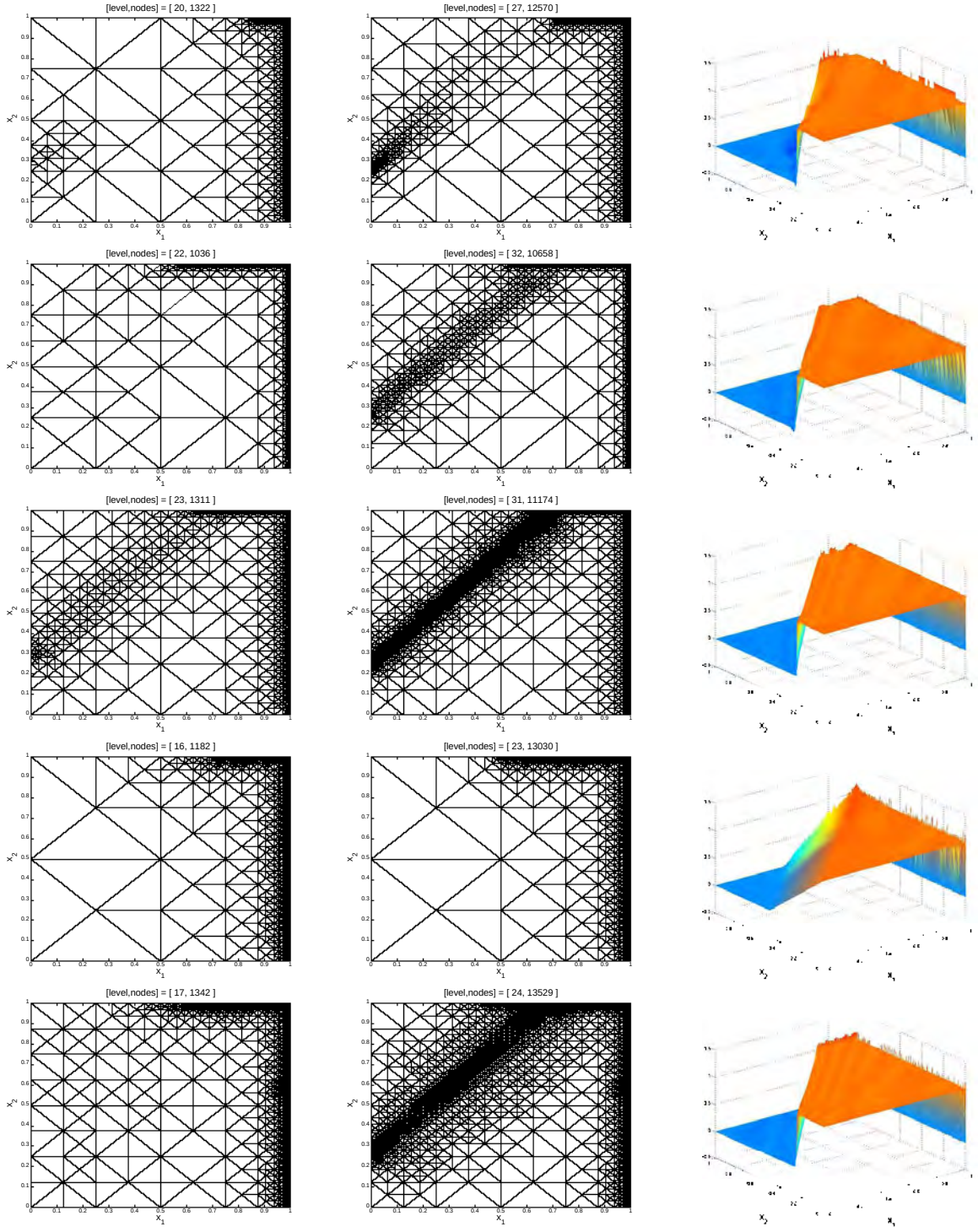

Figure 4.13: Example 1C with $\epsilon = 10^{-4}$: Generated locally refined meshes with $N_{\max} = 10^3$ (left column), $N_{\max} = 10^3$ (middle column), and fine grid solution (right column). Row 1: ZZ, row 2: Res-H1, row 3: Res-L2, row 4: Neu-H1, row 5: Neu-L2.

Chapter 5

Optimal control problems governed by convection diffusion equations

The previous chapters focus on adaptive finite element methods for single convection dominated diffusion equations. The goal of this thesis is to extend such methods to optimal control problems. This chapter extends the error estimation of linear-quadratic elliptic optimal control problem of the form

$$\min \ J(y,u) = \frac{1}{2}\|y-\widehat{y}\|^2_{L^2(\Omega)} + \frac{\omega}{2}\|u\|^2_{L^2(\Omega)} \tag{5.1a}$$

subject to

$$-\epsilon\Delta y + b\cdot\nabla y + cy = f + u \qquad \text{in } \Omega, \tag{5.1b}$$

$$y = g_D \qquad \text{on } \Gamma_D, \tag{5.1c}$$

$$\epsilon\frac{\partial y}{\partial n} = g_N \qquad \text{on } \Gamma_N, \tag{5.1d}$$

where the boundary of Ω is divided such that $\partial\Omega = \Gamma_D \cup \Gamma_N$ and $\Gamma_D \cap \Gamma_N = \emptyset$. In addition to assumptions (A1) - (A4), let

(A5) $\omega > 0$, $\widehat{y} \in L^2(\Omega)$.

The solution (y, u) of this problem has two components, the state y and the control function u. This setup first introduces the optimality conditions, then addresses the Galerkin and SUPG approximation to these equations. Finally, the three error estimators from chapter 3 are generalized to this optimal control setting.

5.1 Optimality conditions

From now on the inner product $(\cdot, \cdot)$ is the standard L^2 inner product. The weak form of the PDE in the constraint of (5.1) for $y \in H^1_g(\Omega)$ and $u \in L^2(\Omega)$ is

$$a(y, v) - (u, v) = \ell(v) \qquad \forall v \in H^1_D(\Omega), \tag{5.2}$$

with $a(y, v)$ and $\ell(v)$ as defined in (2.4). Define $Y = H^1(\Omega)$, $Y^g = H^1_g(\Omega)$, $Y^0 = H^1_D(\Omega)$ and $U = L^2(\Omega)$. We seek a solution $(y, u) \in Y \times U$.

$$\min\ J(y, u) = \frac{1}{2}\|y - \widehat{y}\|^2_{L^2(\Omega)} + \frac{\omega}{2}\|u\|^2_{L^2(\Omega)} \tag{5.3a}$$

$$\text{s.t. } a(y, v) - (u, v) = \ell(v) \qquad \forall v \in H^1_D(\Omega). \tag{5.3b}$$

Under the assumptions (A1)-(A4) stated in Section 2.1 the bilinear form a is continuous on $H^1 \times H^1$ and H^1_D-elliptic. See Lemma 2.1.1. Hence the theory in [26, Sec. II.1] guarantees the existence of a unique solution $(y, u) \in Y^g \times U$ of (5.3).

Theorem 5.1.1 *If Assumptions (A1)-(A5) are satisfied, the optimal control problem* (5.3) *has a unique solution* $(y, u) \in Y^g \times U$.

The theory in [26, Sec. II.1] also provides necessary and sufficient optimality conditions, which can be best described using the Lagrangian

$$L(y, u, p) = J(y, u) + a(y, p) - (u, p) - \ell(p),$$

where p is the Lagrangian multiplier. Setting the partial Fréchet-derivatives of this Lagrangian to zero results in the following KKT-system:

$$a(y, v) - (u, v) = \ell(v) \qquad \forall v \in Y^0, \tag{5.4a}$$

$$a(v, p) + (y, v) = k(v) \qquad \forall v \in Y^0, \tag{5.4b}$$

$$-(w, p) + \omega(u, w) = 0 \qquad \forall w \in U, \tag{5.4c}$$

where $k(v) = (\widehat{y}, v)$. These identities are commonly referred to as the state, adjoint and gradient equation. For the specific problem (5.1) these conditions are necessary and sufficient.

Theorem 5.1.2 *If assumptions (A1) - (A5) hold, then* $(y, u) \in Y^g \times U$ *solves* (5.3) *if and only if there exists a* $p \in Y^0$ *such that the optimality conditions (5.4) are satisfied. Furthermore, the optimality conditions (5.4) have a unique solution* $(y, u, p) \in Y^g \times U \times Y^0$.

Equation (5.4b) can be interpreted as the weak form of a PDE in its own right:

$$-\epsilon\Delta p - b\cdot\nabla p + (c - \nabla\cdot b)p = -(y - \widehat{y}) \qquad \text{in } \Omega, \tag{5.5a}$$

$$p = 0 \qquad \text{on } \Gamma_D, \tag{5.5b}$$

$$\epsilon\frac{\partial p}{\partial n} + (b\cdot n)p = 0 \qquad \text{on } \Gamma_N, \tag{5.5c}$$

In similar fashion, equation (5.4c) corresponds to

$$p(x) = \omega\ u(x). \tag{5.6}$$

Hence, solving the optimality system requires the solution to the weak form of two convection-diffusion equations and an algebraic equation.

5.2 Galerkin approximation

Let T_h be a conforming triangulation of the domain Ω. Consider the following spaces of continuous piece-wise linear functions on T_h:

$$Y_h = \left\{ y_h \in H^1(\Omega)\ :\ y_h|_\tau \in \mathcal{P}^1(\tau),\ \forall\tau \in T_h \right\},$$

$$Y_h^0 = \left\{ y_h \in Y_h\ :\ y_h = 0 \text{ on } \Gamma_D \right\},$$

$$Y_h^g = \left\{ y_h \in Y_h\ :\ y_h = g_D \text{ on } \Gamma_D \right\},$$

$$U_h = \left\{ u_h \in L^2(\Omega)\ :\ u_h|_\tau \in \mathcal{P}^1(\tau),\ \forall\tau \in T_h \right\}.$$

The Galerkin approach approximates (y, u, p) by solving:

Find $(y_h, u_h, p_h) \in Y_h^g \times U_h \times Y_h^0$ *such that*

$$a(y_h, v_h) - (u_h, v_h) = \ell(v_h) \qquad \forall v_h \in Y_h^0, \tag{5.7a}$$

$$a(w_h, p_h) + (y_h, w_h) = k(w_h) \qquad \forall w_h \in Y_h^0, \tag{5.7b}$$

$$-(q_h, p_h) + \omega(u_h, q_h) = 0 \qquad \forall q_h \in U_h. \tag{5.7c}$$

Recall that such Galerkin approximations yield highly oscillatory numerical solutions for convection dominated diffusion equations (Section 2.2). Since this system is coupled, errors are no longer localized to state, adjoint or control. Hence, there is even more reason to modify the problem in order to control the errors (especially the gradient errors). Again, this work applies SUPG stabilization.

5.3 SUPG stabilization

This generalization of SUPG is not trivial. Collis and Heinkenschloss [13] propose two different tactics. This work uses the optimize-then-discretize approach, which reduces the problem to:

Find $(y_h, u_h, p_h) \in Y_h^g \times U_h \times Y_h^0$ *such that*

$$a_h^s(y_h, v_h) - (u_h, v_h)_h^s = \ell_h^s(v_h) \qquad \forall v_h \in Y_h^0, \tag{5.8a}$$

$$a_h^a(p_h, w_h) + (y_h, w_h)_h^a = k_h^a(w_h) \qquad \forall w_h \in Y_h^0, \tag{5.8b}$$

$$-(q_h, p_h) + \omega(u_h, q_h) = 0 \qquad \forall q_h \in U_h. \tag{5.8c}$$

where

$$
\begin{aligned}
(u_h, v_h)_h^s &= (u_h, v_h) + \sum_{\tau \in T_h} \delta_\tau (u_h, b \cdot \nabla v_h)_\tau, \\
a_h^s(y_h, v_h) &= \int_\Omega [\epsilon \nabla u_h \cdot \nabla v_h + (b \cdot \nabla u_h) v_h + c u_h v_h] \\
&\qquad + \sum_{\tau \in T_h} \delta_\tau (-\epsilon \Delta p_h + b \cdot \nabla y_h + c y_h, b \cdot \nabla v_h)_\tau, \\
\ell_h^s(v) &= \ell(v_h) + \sum_{\tau \in T_h} \delta_\tau (f, b \cdot \nabla v_h)_\tau, \\
(y_h, v_h)_h^a &= (y_h, v_h) + \sum_{\tau \in T_h} \delta_\tau (y_h, -b \cdot \nabla v_h)_\tau, \\
a_h^a(p_h, v_h) &= \int_\Omega [\epsilon \nabla p_h \cdot \nabla v_h - (b \cdot \nabla p_h) v_h + (c - \nabla \cdot b) p_h v_h] \\
&\qquad + \sum_{\tau \in T_h} \delta_\tau (-\epsilon \Delta p_h - b \cdot \nabla p_h + (c - \nabla b) p_h, -b \cdot \nabla v_h)_\tau, \\
k_h^a(v_h) &= k(v_h) + \sum_{\tau \in T_h} \delta_\tau (\hat{y}, -b \cdot \nabla v_h)_\tau.
\end{aligned}
$$

The parameters δ_τ, $\tau \in \mathcal{T}_h$, are defined by (2.13). The state equation (5.8a) is stabilized using the SUPG method described for single convection-diffusion equation (with an added term for the control). Recall that the adjoint is the weak form of (5.5). In the optimize-then-discretize approach [13] the adjoint is discretized and stabilized by the problem functions from this equation. This is clearly reflected by the definition of $a_h^a(\cdot, \cdot)$, $k_h^a(\cdot)$ and $(\cdot, \cdot)_h^a$.

Recall that in an adaptive regime the error needs to be estimated after a numerical solution of the SUPG problem is computed. Instead of a single equation, in this SUPG problem error estimation is based on the coupled system of equations (5.4) and (5.8).

5.4 A posteriori error estimation

The coupling of convection diffusion equations complicates error estimation. This section explains the extension of the Zienkiewicz-Zhu, the norm-residual based and the local Neumann problem estimator of Chapter 3. In literature, so far only one error estimator for the SUPG approximation has been proposed [38] (which is also extended to edge stabilization [17]), and is equivalent to the norm-residual based H^1 estimator of this thesis.

5.4.1 Zienkiewicz-Zhu estimation

The ZZ estimator can be trivially generalized to the coupled system, since it only uses local regularity information. Define the operator G by (3.2), and let the state and adjoint estimators be

$$\eta_\tau^y \equiv \|Gy - \nabla y_h\|_{L^2(\tau)}, \quad \eta_\tau^p \equiv \|Gp - \nabla p_h\|_{L^2(\Omega)}.$$

Since u is not necessarily a function in $H^1(\Omega)$, it makes no sense to apply the same approach. In this thesis $\eta_\tau^u \equiv 0$. However, other approaches can certainly be used to locally recover L^2 error in u [7].

The remaining issue now is how these estimators should be combined. One option is to add the three estimators:

$$\eta^{\mathrm{ZZ}} = \sqrt{\sum_{\tau \in \mathcal{T}_h} \left(\eta_\tau^y\right)^2 + \left(\eta_\tau^u\right) + \left(\eta_\tau^p\right)^2}.$$

This estimator only uses regularity information of the approximation (y_h, u_h, p_h). None of the problem data is used, nor any information of the coupling of the system. Therefore, just as in the previous chapters, it seems that a more advanced tool can be developed which incorporates more properties of the optimal control problem. One of such methods is the norm-residual based estimator.

5.4.2 Norm-residual based estimation

Obviously, one of the key concepts in this sort of error estimation is the residual of a numerical solution (y_h, u_h, p_h).

Definition 5.4.1 *If (y_h, u_h, p_h) is a Galerkin or SUPG approximation, then the gradient equations interior residual is defined by*

$$R_u^{\text{int}}(u_h, p_h) = -p_h + \omega u_h,$$

the state equations interior residual is defined by

$$R_y^{\text{int}}(y_h, u_h) = f + u_h - \epsilon \Delta y_h + b \cdot \nabla y_h + c y_h,$$

its edge residual is defined by

$$R_y^{\text{edge}}(y_h) = \begin{cases} 0 & \text{if } e \in \mathcal{E}_h^D, \\ g - \epsilon \nabla y_h \cdot n & \text{if } e \in \mathcal{E}_h^N, \\ \epsilon \left[\nabla y_h \cdot n\right] & \text{if } e \in \mathcal{E}_h^\Omega, \end{cases}$$

the adjoint equations interior residual is defined by

$$R_p^{\text{int}}(p_h, y_h) = -(y_h - y_0) - \epsilon \Delta p_h + b \cdot \nabla p_h + (c - \nabla \cdot b) p_h,$$

and its edge residual by

$$R_p^{\text{edge}}(p_h) = \begin{cases} 0 & \text{if } e \in \mathcal{E}_h^D, \\ \epsilon \nabla p_h \cdot n + (b \cdot n) p_h & \text{if } e \in \mathcal{E}_h^N, \\ \epsilon \left[\nabla p_h \cdot n \right] & \text{if } e \in \mathcal{E}_h^\Omega. \end{cases}$$

Here $\mathcal{E}_h^\Omega$ are interior edges, $\mathcal{E}_h^D$ are Dirichlet edges and $\mathcal{E}_h^N$ are Neumann edges of $\mathcal{T}_h$.

Based on these residuals a norm-residual based error estimator can be formulated.

Definition 5.4.2 *If (y, u, p) is the solution to (5.4), (y_h, u_h, p_h) is the solution to (5.8) and $R_*^{\text{int}}(\cdot,\cdot)$, $R_*^{\text{edge}}(\cdot)$ be as Definition 5.4.1, let*

$$\eta_\tau^y \equiv \left[h_\tau^2 \| R_y^{\text{int}}(y_h, u_h) \|_{L^2(\Omega)}^2 + \sum_{e \in \mathcal{E}(\tau)} h_e \| \widetilde{R}_y^{\text{edge}}(y_h) \|_{L^2(\Omega)}^2 \right]^{1/2},$$

$$\eta_\tau^u \equiv \| R_u^{\text{int}}(u_h, p_h) \|_{L^2(\tau)},$$

$$\eta_\tau^p \equiv \left[h_\tau^2 \| R_p^{\text{int}}(p_h, y_h) \|_{L^2(\Omega)}^2 + \sum_{e \in \mathcal{E}(\tau)} h_e \| \widetilde{R}_p^{\text{edge}}(p_h) \|_{L^2(\Omega)}^2 \right]^{1/2},$$

then the norm-residual based estimator in the H^1 sense is defined by [38]

$$\eta_\tau^{\text{Neu-H1}} = \left[\frac{1}{\epsilon} (\eta_\tau^y)^2 + (\eta_\tau^u)^2 + \frac{1}{\epsilon} (\eta_\tau^p)^2 \right]^{1/2}.$$

Let,

$$\widetilde{\eta}_\tau^y \equiv \left[h_\tau^4 \| R_y^{\text{int}}(y_h, u_h) \|_{L^2(\Omega)}^2 + \sum_{e \in \mathcal{E}(\tau)} h_e^3 \| \widetilde{R}_y^{\text{edge}}(y_h) \|_{L^2(\Omega)}^2 \right]^{1/2},$$

$$\widetilde{\eta}_\tau^p \equiv \left[h_\tau^4 \| R_p^{\text{int}}(p_h, y_h) \|_{L^2(\Omega)}^2 + \sum_{e \in \mathcal{E}(\tau)} h_e^3 \| \widetilde{R}_p^{\text{edge}}(p_h) \|_{L^2(\Omega)}^2 \right]^{1/2},$$

then the norm-residual based estimator in the L^2 sense is defined by

$$\eta_\tau^{\text{Neu}-\text{L2}} = \left[\frac{1}{\epsilon} \left(\widetilde{\eta}_\tau^y\right)^2 + \left(\eta_\tau^u\right)^2 + \frac{1}{\epsilon} \left(\widetilde{\eta}_\tau^p\right)^2 \right]^{1/2}.$$

The global error estimators are defined in the usual way:

$$\eta^{\text{Neu}-\text{H1}} = \left(\sum_{\tau \in \mathcal{T}_h} \left(\eta_\tau^{\text{Neu}-\text{H1}}\right)^2 \right)^{1/2}, \quad \eta^{\text{Neu}-\text{L2}} = \left(\sum_{\tau \in \mathcal{T}_h} \left(\eta_\tau^{\text{Neu}-\text{L2}}\right)^2 \right)^{1/2}.$$

Ideally, it would be desirable obtain bounds in the sense of (3.1) for this estimator. Such an upper bound exists for the H^1 estimator and is shown in Theorem 5.4.3, a lower bound has not been proved to date.

Theorem 5.4.3 *[38, Th.5.5] If (y, u, p) is the solution to (5.4), (y_h, u_h, p_h) is the solution to (5.8) and $\eta^{\text{Neu}-\text{H1}}$ is defined by Definition 5.4.2, then there exists $C_{12} > 0$ such that*

$$\|y - y_h\|_\epsilon^2 + \|u - u_h\|_{L^2(\Omega)}^2 + \|p - p_h\|_\epsilon^2 \leq C_{12} \ \left(\eta^{\text{Neu}-\text{H1}}\right)^2.$$

Proof: Given $v \in U$, let $y[v]$ be the unique solution to

$$a(y[v], w) - (v, w) = \ell(w) \qquad \forall w \in Y^0. \tag{5.9}$$

The optimal control problem can now be viewed as an unconstrained optimization problem:

$$\min_{v \in U} \widehat{J}(v) \equiv J(y[v], v).$$

This implicit formulation opens the door to further analysis. For given v, and $y[v]$ defined by (5.9) and let $p[v]$ be the solution to

$$a(w, p[v]) + (y[v], w) = k(w) \qquad \forall w \in Y^0. \tag{5.10}$$

Let u be the optimal control, then $(y, u, p) = (y[u], u, p[u])$ is the solution to (5.4). The gradient of the objective functions $\widehat{J}(v)$ satisfies

$$(\nabla \widehat{J}(v), w) = -(w, p[v]) + \alpha(v, w) \qquad \forall w \in U. \tag{5.11}$$

Consider the following for any $w \in U$

$$\begin{aligned}(\nabla \widehat{J}(u) - \nabla \widehat{J}(u_h), w) &= (\nabla \widehat{J}(u), w) - (\nabla \widehat{J}(u_h), w), \\ &= -(w, p) + \omega(u, w) + (w, p[u_h]) - \omega(u_h, w), \\ &= (w, p[u_h] - p) + \omega(u - u_h, w).\end{aligned}$$

Setting $w = u - u_h$ leads to

$$(\nabla \widehat{J}(u) - \nabla \widehat{J}(u_h), u - u_h) = (u - u_h, p[u_h] - p) + \omega(u - u_h, u - u_h), \tag{5.12}$$

$$= (u - u_h, p[u_h] - p) + \omega \|u - u_h\|^2_{L^2(\Omega)}. \tag{5.13}$$

$$\begin{aligned}(u - u_h, p[u_h] - p) &= (u, p[u_h] - p) - (u_h, p[u_h] - p), \\ &= a(y, p[u_h] - p) - \ell(p[u_h] - p) \\ &\quad - a(y[u_h], p[u_h] - p) + \ell(p[u_h] - p), \\ &= a(y[u_h] - y, p[u_h] - p),\end{aligned}$$

by equation (5.9). Furthermore equation (5.10) yields

$$\begin{aligned}(u-u_h, p[u_h]-p) &= a(y-y[u_h], p[u_h]) - a(y-y[u_h], p),\\ &= k(y-y[u_h]) + (y-y[u_h], y)\\ &\qquad - k(y-y[u_h]) - (y-y[u_h], y[u_h]),\\ &= (y-y[u_h], y-y[u_h]) = \|y-y[u_h]\|^2 \geq 0.\end{aligned}$$

The optimality condition of the unconstrained optimization problem, $\nabla\hat{J}(u) = 0$, reduces equation (5.13) to

$$\begin{aligned}\omega\|u-u_h\|^2_{L^2(\Omega)} &\leq (\nabla\widehat{J}(u) - \nabla\widehat{J}(u_h), u-u_h),\\ &= -(\nabla\widehat{J}(u_h), u-u_h),\\ &= (u-u_h, p[u_h]) - \omega(u_h, u-u_h),\\ &= (p[u_h], u-u_h) - \omega(u_h, u-u_h) + (p_h, u-u_h) - (p_h, u-u_h),\\ &= (p[u_h]-p_h, u-u_h) + (p_h - \omega u_h, u-u_h).\end{aligned}$$

Invoking Young's inequality with parameter $\gamma > 0$ yields:

$$\begin{aligned}\omega\|u-u_h\|^2_{L^2(\Omega)} \leq \frac{1}{2\gamma}\|p_h - p[u_h]\|^2_{L^2(\Omega)} + \frac{\gamma}{2}\|u-u_h\|^2_{L^2(\Omega)}\\ + \frac{1}{2\gamma}\|p_h - \omega u_h\|^2_{L^2(\Omega)} + \frac{\gamma}{2}\|u-u_h\|^2_{L^2(\Omega)}.\end{aligned}$$

Since γ is an arbitrary positive constant, set $\gamma = \omega/2$:

$$\frac{\omega^2}{2}\|u-u_h\|^2_{L^2(\Omega)} \leq \|p_h - p[u_h]\|^2_{L^2(\Omega)} + (\eta^u)^2. \tag{5.14}$$

Consider the following for $v \in Y$

$$\begin{aligned}
a(v, p[u_h] - p_h) &= -(y[u_h] - y_0, v) - a(v, p_h), \\
&= \int_\Omega (-y[u_h] + y_0 + y_h - y_h)v - \left[\int_\Omega \epsilon \nabla v \cdot \nabla p_h + (b \cdot \nabla v) p_h + c v p_h \right], \\
&= \int_\Omega (y_h - y[u_h])v - \left[\int_\Omega (y_h - y_0) + \epsilon \nabla v \cdot \nabla p_h + (b p_h) \cdot \nabla v + c v p_h \right].
\end{aligned}$$

Applying integration by parts:

$$\begin{aligned}
a(v, p[u_h] - p_h) &= \int_\Omega (y_h - y[u_h])v - \sum_{\tau \in \mathcal{T}_h} \left[\int_\tau (y_h - y_0) - \epsilon \Delta p_h v - \nabla (b p_h) v + c v p_h \right. \\
&\qquad \left. + \int_{\partial \tau} \epsilon \nabla p_h \cdot n \; v + (b p_h) \cdot n \; v \right] \\
&= \int_\Omega (y_h - y[u_h])v - \sum_{\tau \in \mathcal{T}_h} \left[\int_\tau (y_h - y_0) - \epsilon \Delta p_h v - b \cdot \nabla p_h v + (c - \nabla \cdot b) p_h v \right. \\
&\qquad \left. + \int_{\partial \tau} \epsilon \nabla p_h \cdot n \; v + (b \cdot n) p_h v \right], \\
&= \int_\Omega (y_h - y[u_h])v + \sum_{\tau \in \mathcal{T}_h} \int_\tau R_p^{\text{int}}(p_h, y_h)v + \sum_{e \in \mathcal{E}_h} \int_e R_p^{\text{edge}}(p_h, y_h)v.
\end{aligned}$$

Section 3.2 has shown how the right hand side of this equation can be bounded. Setting $v = p[u_h] - p_h$ yields

$$\begin{aligned}
a(p[u_h] - p_h, p[u_h] - p) &\leq (y[u_h] - y, p[u_h] - p_h) + \|p[u_h] - p_h\|_\epsilon \; \eta^p. \\
&\leq \|y[u_h] - y\|_\epsilon \|p[u_h] - p_h\|_\epsilon + \frac{C_{10}}{\sqrt{\epsilon}} \|p[u_h] - p_h\|_\epsilon \; \eta^p.
\end{aligned}$$

The coercivity of $a(\cdot, \cdot)$ and Young's inequality with $\gamma > 0$ reduce this equation to:

$$\begin{aligned}
\alpha \|p[u_h] - p_h\|_\epsilon^2 \leq \frac{\gamma}{2} \|p[u_h] - p_h\|_\epsilon^2 + \frac{1}{2\gamma} \|y[u_h] - y_h\|_\epsilon^2 \\
+ \frac{\gamma}{2} \|p[u_h] - p_h\|_\epsilon^2 \; + \frac{C_{10}^2}{2\gamma\epsilon} \left(\eta^p\right)^2.
\end{aligned}$$

Setting $\gamma = \alpha/2$ yields

$$\|p[u_h] - p_h\|_\epsilon^2 \leq \frac{2}{\alpha^2}\|y[u_h] - y_h\|_\epsilon^2 + \frac{2C_{10}^2}{(\alpha^2\epsilon)}\ (\eta^p)^2. \tag{5.15}$$

A similar argument holds for $y[u_h] - y_h$:

$$a(y[u_h] - y_h, v) = l(v) + (u_h, v) - a(y_h, v).$$

Section 3.2 shows how this equation can be transformed to:

$$\begin{aligned} a(y[u_h] - y_h, v) &= \sum_{\tau\in\mathcal{T}_h}\int_\tau R_y^{\text{int}}(y_h, u_h)v + \sum_{e\in\mathcal{E}_h}\int_e R_y^{\text{edge}}(y_h, u_h)v, \\ &\leq \frac{C_{12}}{\sqrt{\epsilon}}\|y[u_h] - y_h\|_\epsilon \eta^y. \end{aligned}$$

Hence, by coercivity of $a(\cdot,\cdot)$

$$\|y[u_h] - y\|_\epsilon^2 \leq \frac{C_{12}^2}{\alpha^2\epsilon}\ \left(\eta^y\right)^2. \tag{5.16}$$

Combining equations (5.14), (5.15) and (5.16)

$$\|u - u_h\|_{L^2(\Omega)}^2 \leq C_{12}\left[\frac{1}{\epsilon}(\eta^y)^2 + (\eta^u)^2 + \frac{1}{\epsilon}(\eta^p)^2\right], \tag{5.17}$$

with

$$C_{12} = \frac{1}{\alpha^2\omega^2}\max\left\{C_{10}^2, \frac{2C_{11}^2}{\alpha^2}\right\}.$$

Note that:

$$a(y - y[u_h], v) = (u - u_h, v) \qquad \forall v \in Y_h^0, \tag{5.18}$$

$$a(w, p - p[u_h]) = -(y - y[u_h], w) \qquad \forall w \in Y_h^0. \tag{5.19}$$

If $v = y - y[u_h]$ and $w = p - p[u_h]$, then there exists a constant C_{12} such that:

$$\|p - p[u_h]\|_\epsilon^2 \leq \|y - y[u_h]\|_\epsilon^2 \leq C_{12}\|u - u_h\|_{L^2(\Omega)}^2.$$

Therefore the following holds:

$$\|y - y_h\|_\epsilon^2 \leq \|y - y[u_h\|_\epsilon^2 + \|y[u_h] - y_h\|_\epsilon^2 \leq C_{13} \left(\eta^{\text{Neu}-\text{H1}}\right)^2, \tag{5.20a}$$

$$\|p - p_h\|_\epsilon^2 \leq \|p - p[u_h\|_\epsilon^2 + \|p[u_h] - p_h\|_\epsilon^2 \leq C_{14} \left(\eta^{\text{Neu}-\text{H1}}\right)^2. \tag{5.20b}$$

Equations (5.17) and (5.20) result to the final bound:

$$\|y - y_h\|_\epsilon^2 + \|u - u_h\|_{L^2(\Omega)}^2 + \|p - p_h\|_\epsilon^2 \leq C_{12} \left(\eta^{\text{Neu}-\text{H1}}\right)^2.$$

□

Remark 5.4.4 *A similar bound as Theorem 5.4.3 for the Neu-L2 estimator has not been proven to date.*

In this bound constant C_{12} relies heavily on $\omega, \alpha, C_{10}, C_{11}$, where these last constant on their turn depend on the problem. Combining all these constants explicitly in C_{12} neglects many, sometimes vital, properties of the error. In Chapter 3 this problem was fixed by setting up a different estimator based on local Neumann problems. This technique can be generalized to the optimal control context.

5.4.3 Local Neumann estimation

In order to derive a local Neumann estimator, consider a slightly different approach from the previous section. If (y, u, p) is the the solution to the weak form, (y_h, u_h, p_h)

is the SUPG approximation and the error is defined by $(e^y, e^u, e^p) = (y - y_h, u - u_h, p - p_h)$, then for $v \in H^1_D(\Omega)$

$$\begin{aligned} a(e^y, v) &= a(y, v) - a(y_h, v), \\ &= \ell(v) + (u, v) - a(y_h, v) + (u_h - u_h, v), \\ &= (e^u, v) + \ell(v) + (u_h, v) - a(y_h, v), \\ &= (e^u, v) + \sum_{\tau \in \mathcal{T}_h} \int_\tau R_y^{\text{int}}(y_h, u_h)v + \sum_{e \in \mathcal{E}_h} \int_e R_y^{\text{edge}}(y_h)v. \end{aligned}$$

The same principle can be applied to the error of the adjoint and $w \in H^1_D(\Omega)$:

$$\begin{aligned} a(w, e^p) &= a(w, p) - a(w, p_h), \\ &= k(v) - (y, w) - a(w, p_h) + (y_h - y_h, w), \\ &= -(e^y, w) + k(w) - (y_h, w) - a(w, p_h), \\ &= -(e^y, w) + \sum_{\tau \in \mathcal{T}_h} \int_\tau R_p^{\text{int}}(p_h, y_h)w + \sum_{e \in \mathcal{E}_h} \int_e R_p^{\text{edge}}(p_h)w. \end{aligned}$$

Finally, a similar bound holds for the control error, let $q \in L^2(\Omega)$:

$$\begin{aligned} \omega(e^u, q) &= \omega(u, q) - \omega(u_h, q), \\ &= (p, q) - \omega(u_h, q) + (p_h - p_h, q), \\ &= (p - p_h, q) + (p_h - \omega u_h, q), \\ &= (e^p, q) + \sum_{\tau \in \mathcal{T}_h} \int_\tau R_u^{\text{int}}(u_h, p_h)q. \end{aligned}$$

These three expressions form the heart of the local Neumann estimator.

Definition 5.4.5 *Let $Q_\tau \subset H^1(\tau)$, $S_\tau \subset L^2(\tau)$ be finite dimensional, let $(\widehat{e}_\tau^y, \widehat{e}_\tau^u, \widehat{e}_\tau^p) \in Q_\tau \times S_\tau \times Q_\tau$ be the solution of*

$$a(e_\tau^y, v)_\tau - (e_\tau^u, v)_\tau = \int_\tau R_y^{\text{int}}(y_h, u_h)v + \sum_{e\in\mathcal{E}(\tau)} \int_e \widetilde{R}_y^{\text{edge}}(y_h)v, \qquad \forall v \in Q_\tau^0,$$

$$(e_\tau^y, w)_\tau + a(w, e_\tau^p)_\tau = \int_\tau R_p^{\text{int}}(p_h, y_h)w + \sum_{e\in\mathcal{E}(\tau)} \int_e \widetilde{R}_p^{\text{edge}}(p_h)w, \qquad \forall w \in S_\tau,$$

$$\omega(e_\tau^u, q)_\tau - (e_\tau^p, q)_\tau = \int_\tau R_u^{\text{int}}(u_h, p_h)q, \qquad \forall q \in Q_\tau^0.$$

then

$$\eta_\tau^y \equiv \|\widehat{e}_\tau^y\|_{H^1(\Omega)}, \quad \eta_\tau^u \equiv \|\widehat{e}_\tau^u\|_{L^2(\Omega)} \quad \text{and} \quad \eta_\tau^p \equiv \|\widehat{e}_\tau^p\|_{H^1(\Omega)},$$

and the element-wise local Neumann estimator in H^1 is defined by

$$\eta^{\text{Neu}-\text{H1}} = \left[(\eta^y)^2 + (\eta^u)^2 + (\eta^p)^2\right]^{1/2}$$

If

$$\eta_\tau^y \equiv \|\widehat{e}_\tau^y\|_{L^2(\Omega)}, \quad \eta_\tau^u \equiv \|\widehat{e}_\tau^u\|_{L^2(\Omega)} \quad \text{and} \quad \eta_\tau^p \equiv \|\widehat{e}_\tau^p\|_{L^2(\Omega)}.$$

then the element-wise local Neumann estimator in L^2 is defined by

$$\eta^{\text{Neu}-\text{L2}} = \left[(\eta^y)^2 + (\eta^u)^2 + (\eta^p)^2\right]^{1/2}$$

The global estimators are defined by

$$\eta^{\text{Neu}-\text{H1}} = \left(\sum_{\tau\in\mathcal{T}_h} \left(\eta_\tau^{\text{Neu}-\text{H1}}\right)^2\right)^{1/2}, \quad \eta^{\text{Neu}-\text{L2}} = \left(\sum_{\tau\in\mathcal{T}_h} \left(\eta_\tau^{\text{Neu}-\text{L2}}\right)^2\right)^{1/2}.$$

Solving the local Neumann problem in Definition 5.4.5 relies on finding subspaces Q_τ and S_τ. This work uses the bubble functions $\{\psi_i\}$ defined in (3.5) as finite basis for

both spaces. This leads to a 12×12 system, which needs to be solved for each single element.

Chapter 6

Numerical results II: Linear-quadratic elliptic optimal control problems

This section contains numerical results on AFEMs for linear-quadratic elliptic optimal control problems governed by convection dominated diffusion equations.

6.1 Notes on numerical experiments

The SUPG discretized optimality conditions (5.8) are equivalent to

Find $(y_h, p_h) \in Y_h^g \times Y_h^0$ *such that*

$$a_h^s(y_h, v_h) - \frac{1}{\omega}\,(p_h, v_h)_h^s = \ell_h^s(v_h) \qquad \forall v_h \in Y_h^0, \tag{6.1a}$$

$$a_h^a(p_h, w_h) + (y_h, w_h)_h^a = k_h^a(w_h) \qquad \forall w_h \in Y_h^0, \tag{6.1b}$$

and setting

$$u_h = \frac{1}{\omega}\,p_h. \tag{6.1c}$$

Hence, the interior residual of the gradient equation is $R_u^{\text{int}}(u_h, p_h) = 0$. In the norm-residual based estimators, defined in Section 5.4.2, this implies $\eta_\tau^u = 0$ for all $\tau \in \mathcal{T}_h$. In the Zienkiewicz-Zhu estimator defined in Section 5.4.1 we set $\eta_\tau^u = 0$, for all $\tau \in \mathcal{T}_h$. Hence the Zienkiewicz-Zhu estimator and the norm-residual based estimators only use η_τ^y and η_τ^p to estimate the error.

The local Neumann estimator is also simplified by solving the system in this way. Since $R_u^{\text{int}}(u_h, p_h) = 0$ and the since the true control and adjoints and the SUG computed control and adjoint satisfy $\omega u = p$ and $\omega u_h = p_h$, respectively, the third equation in the Neumann problem in Definition 5.4.5 is true and the local Neumann estimator also yields $\eta_\tau^u = 0$, for all $\tau \in \mathcal{T}_h$.

Note that $\eta_\tau^u = 0$ in general would not hold if control constraints are present or if the discretized-then-optimize approach is used [13].

6.2 Numerical examples

6.2.1 Example 2A: Boundary layer

This example is based on Example 1A. Recall function $\eta(x)$ defined in (4.1). Now consider $\mu(x) = \eta(1-x)$:

$$\mu(x) = 1 - x - \frac{\exp(-x/\epsilon) - \exp(-1/\epsilon)}{1 - \exp(-1/\epsilon)}.$$

Consider the same problem as in 1A with $\Omega = [0,1]^2$, $\theta = 45^o$, $b = (\cos\theta, \sin\theta)^T$ and $c = 0$. Let the state and adjoint be defined by

$$y(x_1, x_2) = \eta(x_1)\ \eta(x_2),$$

$$p(x_1, x_2) = \mu(x_1)\ \mu(x_2).$$

Hence, the state is the same as the PDE solution in Example 1A. The control is defined by (5.6):

$$u(x_1, x_2) = \frac{1}{\omega}\ p(x_1, x_2) = \frac{1}{\omega}\ \mu(x_1)\ \mu(x_2).$$

Hence, by the constraint in (5.1):

$$f(x_1, x_2) = -u(x_1, x_2) - \epsilon\Delta y(x_1, x_2) + b \cdot \Delta y(x_1, x_2).$$

Since $c = \nabla \cdot b = 0$, equation (5.5) yields

$$\widehat{y}(x_1, x_2) = y(x_1, x_2) - \epsilon\Delta p(x_1, x_2) - b \cdot \Delta p(x_1, x_2).$$

The true state is shown in Figure 4.1 for $\epsilon = 10^{-1}, 10^{-2}, 10^{-3}$. Note that the control and adjoint are merely mirrored (and scaled by ω) versions of the state. Therefore the plots are omitted in this section.

The difference between Example 1A and this example is that the state exhibits a boundary layer at the boundary where $x_1 = 1$ or $x_2 = 1$, whereas the adjoint and the control exhibit a boundary layer at the boundary where $x_1 = 0$ or $x_2 = 0$. Both the boundary layer for the state and the adjoint/control need to be resolved. For example, if the boundary layer in the state is not resolved, then the coupling in the optimality conditions, causes this error to be propagated though the adjoint equation into the domain. This is very different from the situation when the state equation is solved alone for fixed u as in Example 1A.

Let $\epsilon = 10^{-3}$ and $\omega = 1$, Figure 6.1 shows the generated mesh when the refinement is driven by the Res-H1 estimator. All refinement is done around the boundary layers of the state and the control. Hence, the error indicator is able to pick up both boundary layers, and combine them to an effective estimate.

Since boundary layer occur near the entire boundary, for given number of nodes the the mesh-sizes around the layers are bigger than in Example 1A. In the single convection equation case, an estimator could put all its focus on the layer for the state equation. In the optimal control case, it has to distribute the refinement over the the layer for the state y as well as over the the layer for the adjoint p. Hence, each layer will (ideally) only get half the refinements.

Figures 6.2 and 6.3 show cross sections of the state and the control along the line $x_1 = x_2$ for $\epsilon = 10^{-3}$ and $\omega = 1$ and various refinement levels. The upper left plots in both figures correspond to the coarsest mesh and show that an unresolved boundary

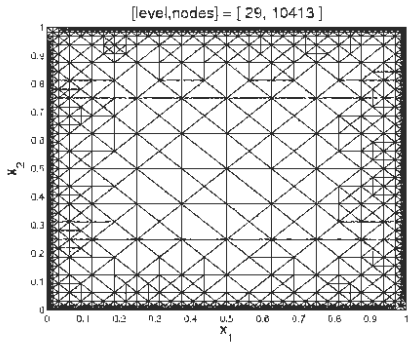

Figure 6.1: Example 2A with $\omega = 1$ and $\epsilon = 10^{-3}$: Locally refined mesh ($N_{\max} = 10000$) driven by Res-H1 estimator.

layer in the control/adjoint will pollute the state. If we compare the upper left plot in Figure 6.2 with the upper left plot in Figure 4.5 we see that the error in the state for $x_1 = x_2 \ll 1$ is much smaller in the case of a fixed u (Figure 4.5) as it is in the optimal control case (Figure 6.2).

Recall the cross sections of the solution along the convection flow (on $x_1 = x_2$) in Figure 4.5. Evident from this figure is that spurious oscillations occur in a small region around the boundary layer when the layer is not resolved. Figures 6.2 and 6.3 are the same cross sections, only now of respectively the state and control. Also in this coupled system the spurious oscillations occur around the boundary layer. For the coarsest meshes, there is also a visible error in the interior. The errors in the boundary layers are propagated into the interior due to the coupling with an equation with convection in opposite direction. Once the boundary layers are sufficiently resolved, the oscillations are reduced, and no oscillations are propagated, which also results in a higher accuracy in the interior.

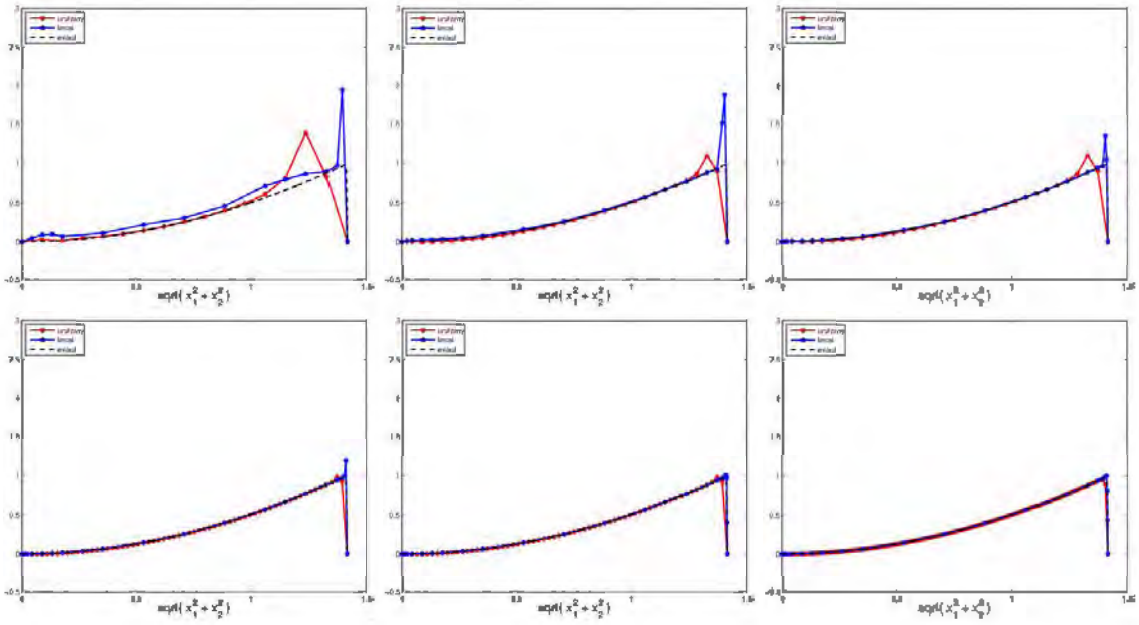

Figure 6.2: Example 2A with $\epsilon = 10^{-3}$ and $\omega = 1$: Cross section of the state along the constant convection flow after $N_{\max} = 200, 400, 800, 1600, 3200, 6400$ (one value of $N_{\max}$ from left to right). Displayed lines are the true solution (black), the solution with uniform refinement (red), the local refinement solution (blue) and the exact solution (black).

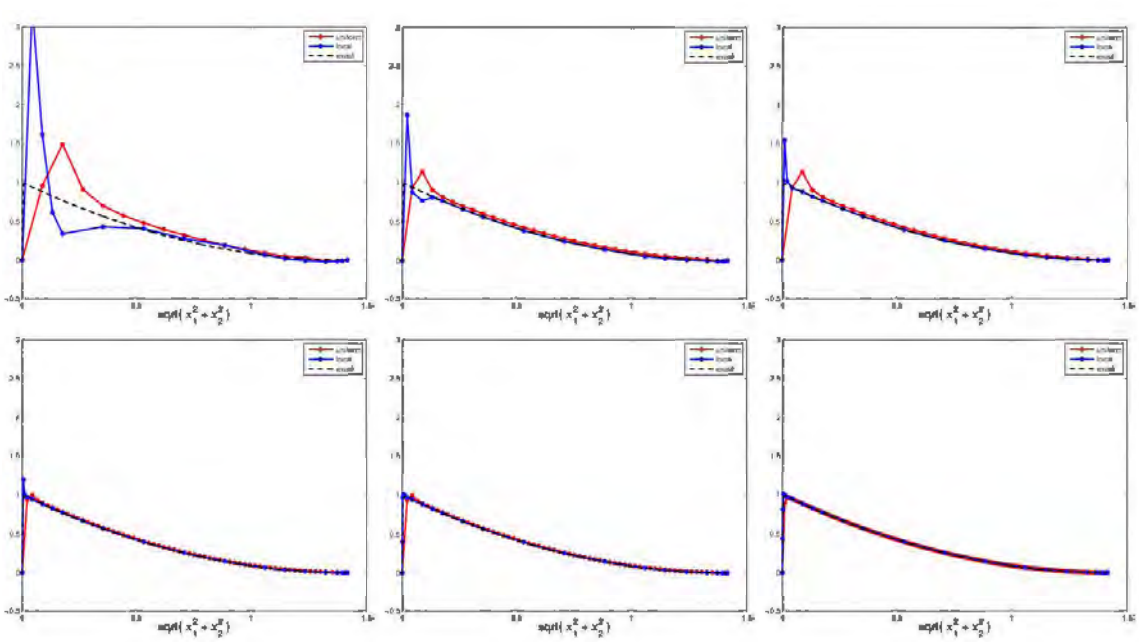

Figure 6.3: Example 2A with $\epsilon = 10^{-3}$ and $\omega = 1$: Cross section of the control along the constant convection flow after $N_{\max} = 200, 400, 800, 1600, 3200, 6400$ (one value of $N_{\max}$ from left to right). Displayed lines are the true solution (black), the solution with uniform refinement (red), the local refinement solution (blue) and the exact solution (black).

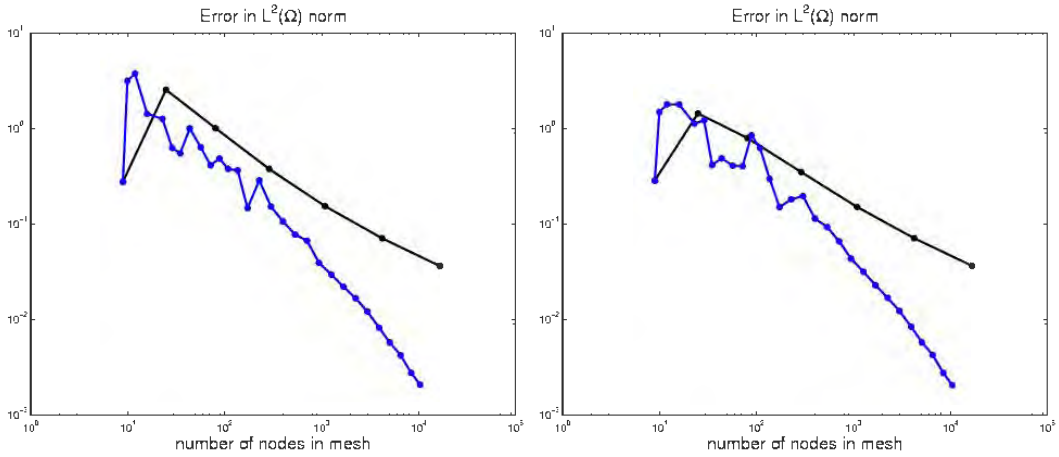

Figure 6.4: Example 2A with $\omega = 1$ and $\epsilon = 10^{-3}$: Global errors in L^2 sense of state (left) and control (right) of solutions uniform (black) and local (blue) refinement driven by Res-H1.

Figure 6.4 shows how the global L^2 error is influenced by the propagated errors. The global L^2 initially does not decrease monotonically. Only when the layers are sufficiently resolved do the errors decrease monotonically. Since in the optimal control case two layers need to be resolved, this seems to agree with the observations for the solution of the single equation in Example 1A.

6.2.2 Example 2B: Circular and straight inner layer

Hinze et al. use this example in their analysis [17] of the norm-residual based estimator in combination with an edge stabilization technique. The problem data are given by $\Omega = (0,1)^2$,

$$b = (2,3), \quad c = 1, \quad \Gamma_D = \partial\Omega, \quad \Gamma_N = \emptyset.$$

The optimal state is

$$y(x_1, x_2) = \frac{2}{\pi} \arctan\left(\frac{1}{\sqrt{\epsilon}}\left[-\frac{1}{2}x_1 + x_2 - \frac{1}{4}\right]\right),$$

which is a function with a straight inner layer which needs to be resolved. The corresponding adjoint is given by

$$p(x_1, x_2) = 16x_1(1-x_1)x_2(1-x_2) \times \left(\frac{1}{2} + \frac{1}{\pi}\arctan\left[\frac{2}{\sqrt{\epsilon}}\left(\frac{1}{16} - \left(x_1 - \frac{1}{2}\right)^2 - \left(x_2 - \frac{1}{2}\right)^2\right)\right]\right).$$

(and corresponds to the function in Example 1B). The optimal control is given by

$$u(x_1, x_2) = \frac{1}{\omega}\, p(x_1, x_2).$$

The Dirichlet data y_D, the right hand side f and the observation $\hat{y}$ are computed from (5.1d) and (5.5) respectively. In particular

$$f(x_1, x_2) = -u(x_1, x_2) - \epsilon\Delta y(x_1, x_2) + 2\frac{\partial}{\partial x_1}y(x_1, x_2) + 3\frac{\partial}{\partial x_2}y(x_1, x_2) + y(x_1, x_2),$$
$$\widehat{y}(x_1, x_2) = y(x_1, x_2) - \epsilon\Delta p(x_1, x_2) - 2\frac{\partial}{\partial x_1}p(x_1, x_2) - 3\frac{\partial}{\partial x_2}p(x_1, x_2) + p(x_1, x_2).$$

Figure 6.5 shows the optimal state and adjoints for $\epsilon = 10^{-2}, 10^{-4}, 10^{-6}$.

The state and the adjoint exhibit inner layers. If these are not resolved, then errors in these layer are propagated along the directions of convection into the domain. Unlike in the case of single equation, the state equation propagates errors along the direction b, whereas the adjoint equation propagates errors along the direction $-b$. Since both equations are coupled, the state, the control, and the adjoint are

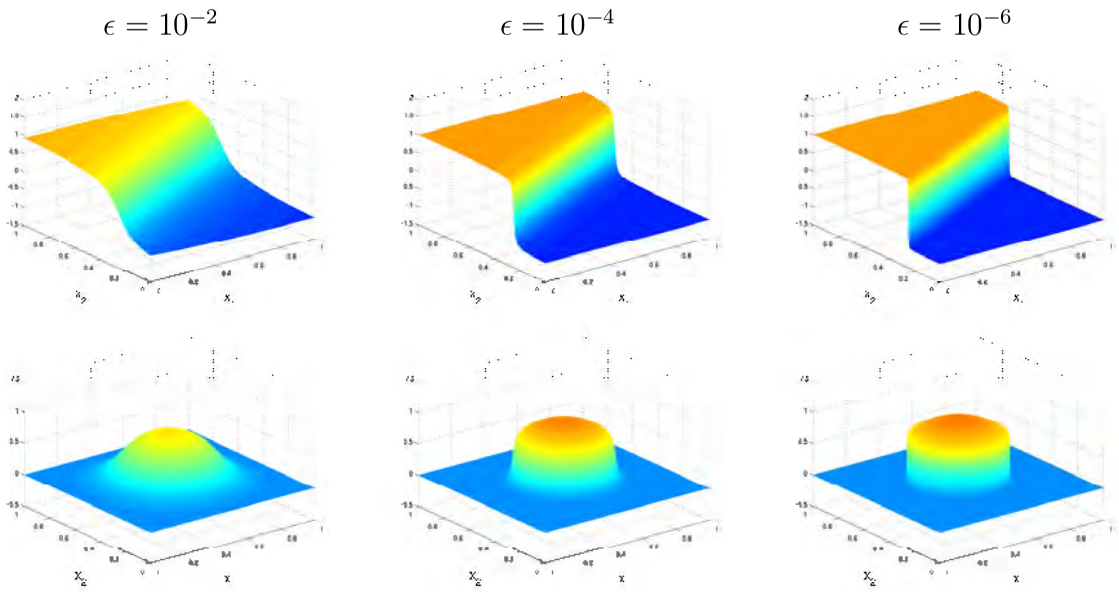

Figure 6.5: Example 2B: Surfaces of the exact state (top row) and adjoint (bottom row) for $\epsilon = 10^{-2}, 10^{-4}, 10^{-6}$ (fixed ϵ for each column).

all polluted by errors propagates along $\pm b$. This can be clearly seen in Figure 6.6, which shows the numerical solutions on uniform meshes (129×129 nodes) for different values of ϵ and fixed $\omega = 0.1$. In the approximations with $\epsilon = 10^{-6}$ the errors at the inner layers are clearly propagated in the convection b and opposite direction $-b$. This should be contrasted with the solution of a single equation shown in Figure 4.8, where the errors around the circular inner layer were propagated into the interior only in the direction of the convection b.

As we have already seen in Example 1B, these propagated errors make it difficult for error estimators to localize layers. Figure 6.7 shows the locally refined meshes generated by the five generalized error estimators. For high values of ϵ all error estimators seem to pick out the layers reasonably well. However, for small ϵ they

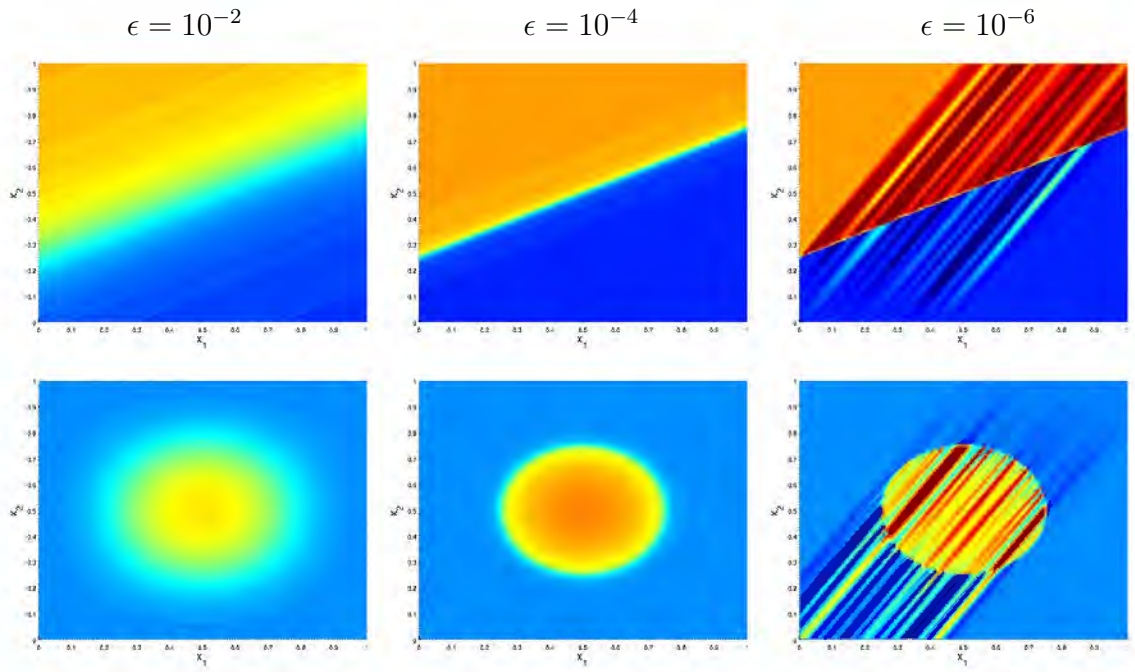

Figure 6.6: Example 2B with $\omega = 0.1$ and $\epsilon = 10^{-2}, 10^{-4}, 10^{-6}$: Computed state (top row) and control (bottom row) on a uniform mesh of 129×129 nodes (fixed ϵ per column).

refine the mesh also in the regions that are polluted by the transport of the error and hence tackle the symptoms rather than the cause of the errors.

Hinze, Yan and Zhou [17] obtain visually better meshes for $\epsilon = 10^{-4}$. They incorporate edge stabilization instead of SUPG, which could be the source of this difference. Furthermore, their total number of nodes is smaller than the ones used in the mehses shown in Figure 6.7. Yan and Zhou [38] use SUPG stabilization, but they use two separate meshes, one for the state, the other for the adjoint and control. The mesh refinement for the state is driven by η^y and the mesh refinement for the adjoint/control is driven by η^p and η^u. Although this seems to generate proper meshes

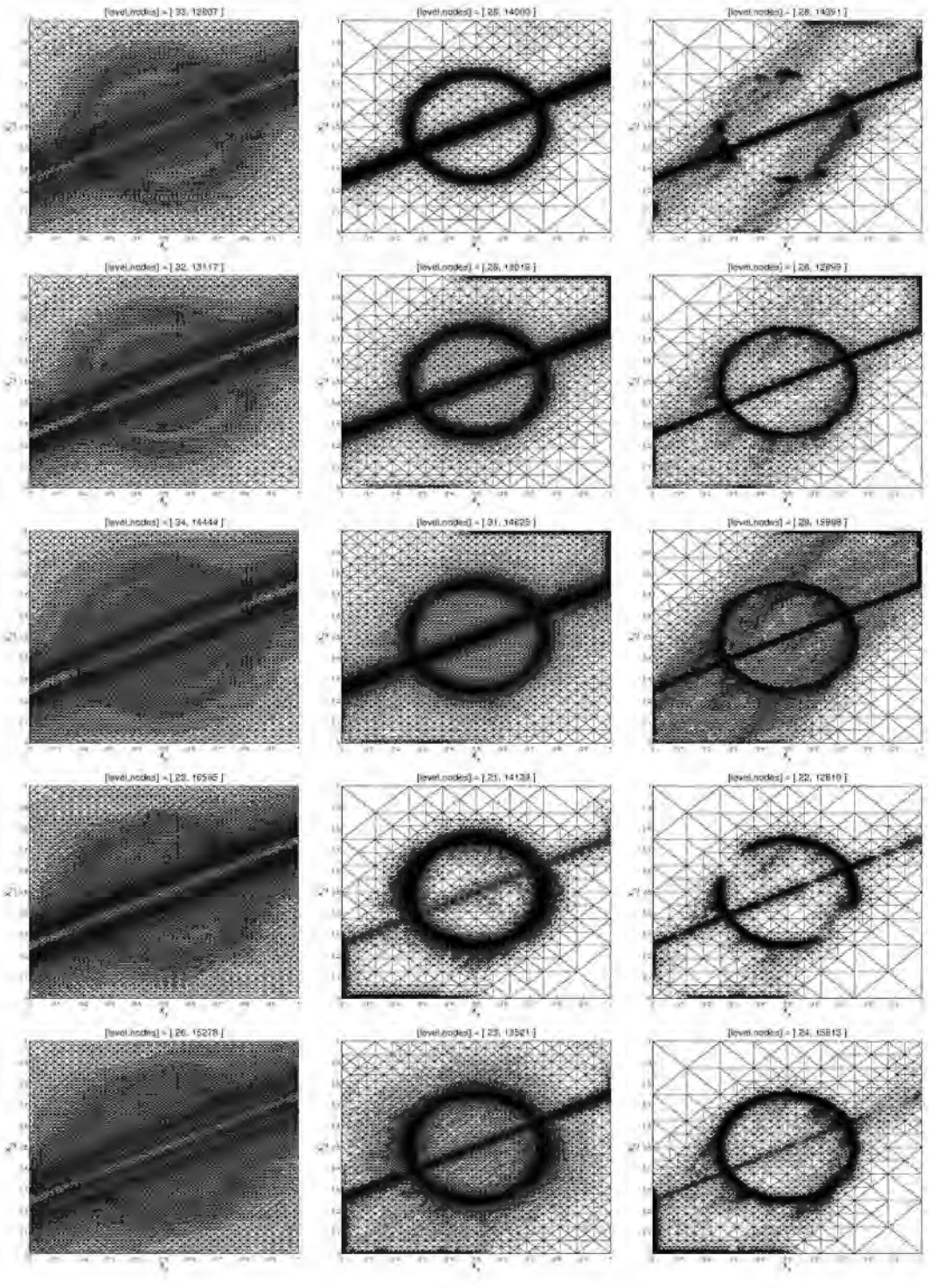

Figure 6.7: Example 2B with $\omega = 0.1$ and $\epsilon = 10^{-2}, 10^{-4}, 10^{-6}$: Generated locally refined meshes with $N_{\max} = 12000$ (fixed ϵ for each column). Row 1: ZZ-meshes, row 2: Res-H1 meshes, row 3: Res-L2 meshes, row 4: Neu-H1 meshes, row 5: Neu-L2 meshes.

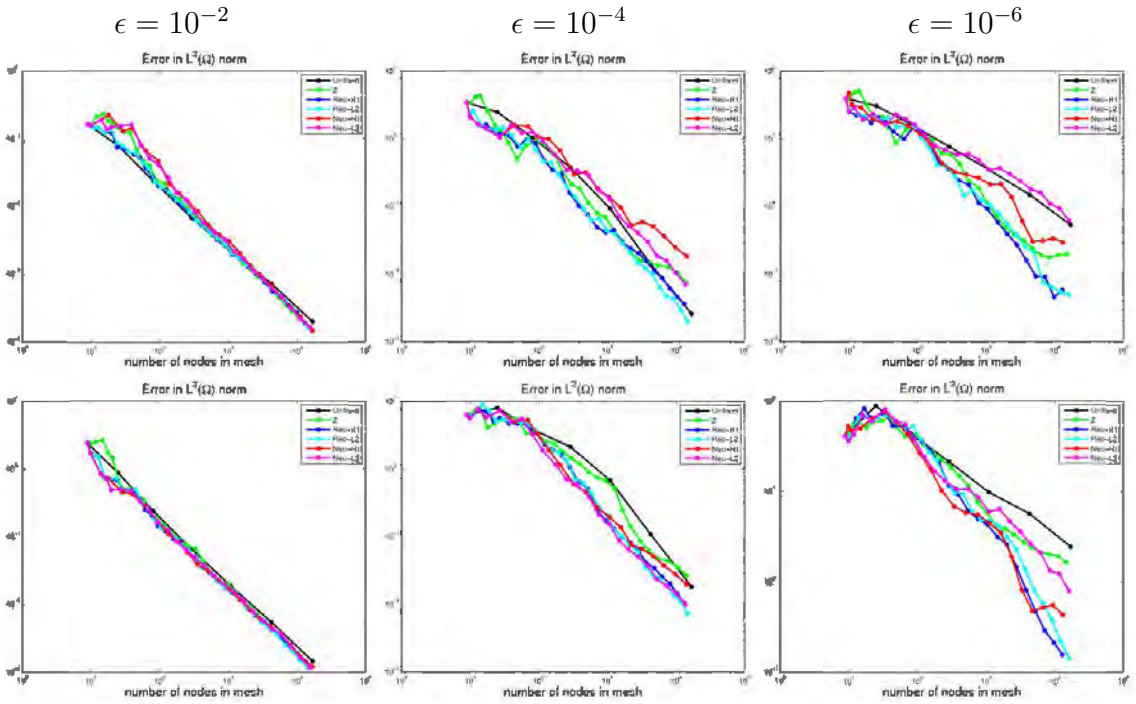

Figure 6.8: Example 2B with $\omega = 0.1$ and $\epsilon = 10^{-2}, 10^{-4}, 10^{-6}$: Errors of state (top row) and control (bottom row) uniform and local refinement using ZZ, Res-H1, Res-L2, Neu-H1 and Neu-L2 estimator (fixed ϵ for each column), measured in $L^2(\Omega)$ norm.

for the examples considered in [38], it implicitly assumes a more decoupled system.

6.2.3 Example 2C: Single straight inner layer

The next example is taken from Heinkenschloss and Leykekhman [15]. Let

$$\Omega = (0,1)^2, \quad b = (1,0), \quad c = 0,$$

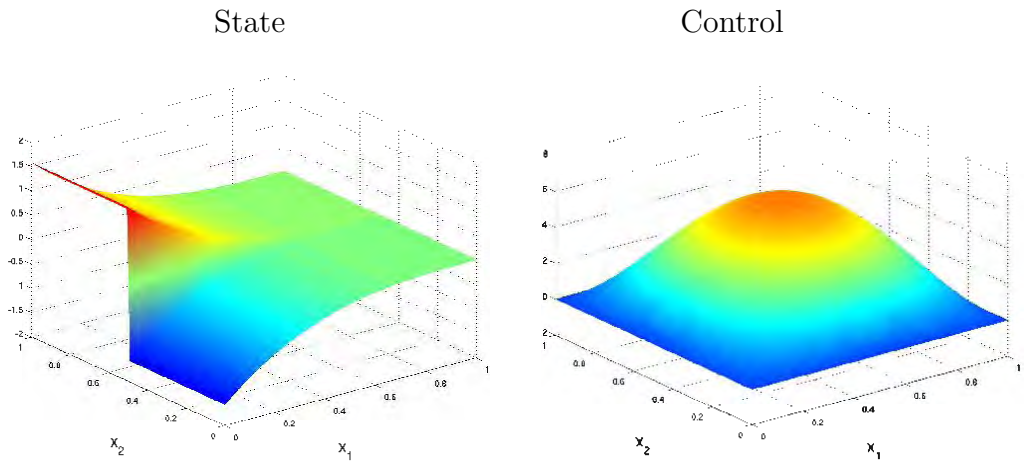

Figure 6.9: Example 2C with $\omega = 10^{-2}$ and $\epsilon = 10^{-7}$: Exact state (left) and exact control (right).

and let the true state, adjoint and control be given by

$$y(x_1, x_2) = (1 - x_1)^3 \arctan\left(\frac{x_2 - .5}{\epsilon}\right),$$

$$p(x_1, x_2) = x_1(1 - x_1)x_2(1 - x_2),$$

$$u(x_1, x_2) = \frac{1}{\omega}\, p(x_1, x_2).$$

We set $\Gamma_D = \partial\Omega$ and compute g_D, f, $\widehat{y}$ such that (5.1d) and (5.5) are satisfied. Figure 6.9 shows the exact state and control for $\epsilon = 10^{-7}$ and $\omega = 10^{-2}$. The state has a sharp inner layer, but the control is very smooth.

Figure 6.10 shows the meshes generated by the five generalized error estimators. The L^2 global errors of the computed state and control are given in Figure 6.11. In this case the Neu-L2 estimator reduces the global error in the state the fastest.

Note that all global control errors for adaptively refined meshes are higher than for uniform refinement. This polynomial solution does not have localized behavior

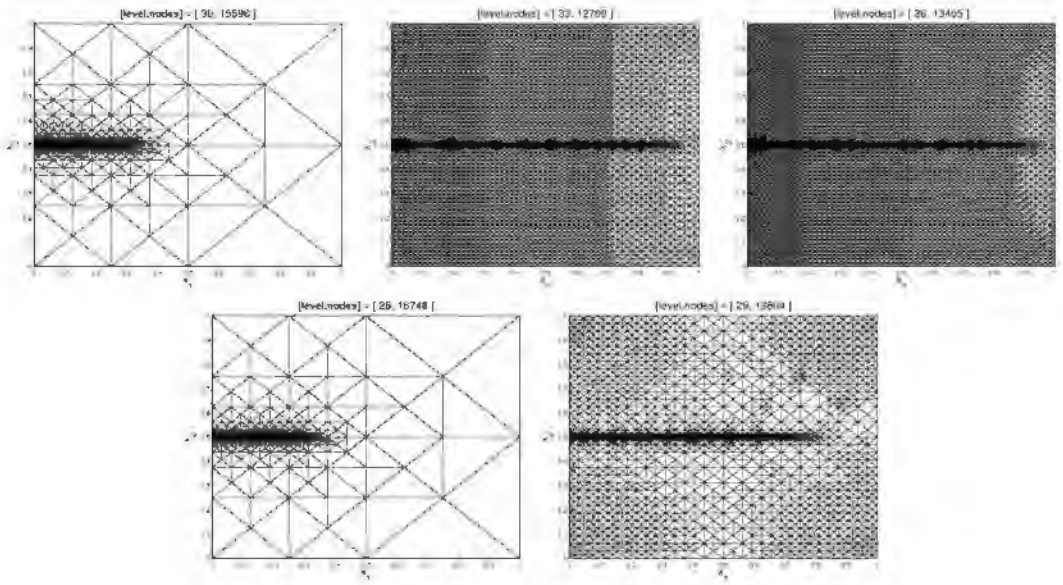

Figure 6.10: Example 2C with $\omega = 10^{-2}$ and $\epsilon = 10^{-7}$: Generated locally refined meshes, ($N_{\max} = 12000$). In order: ZZ, Res-H1, Res-L2, Neu-H1 and Neu-L2 meshes.

and therefore the L^2 error in the control is reduced the most when (close to) uniform refinement is applied.

6.2.4 Example 2D: Inner and boundary layer

This example is an extension of Example 1C and is taken from Heinkenschloss and Leykekhman [15]. The problem data are $\Omega = (0,1)^2$,

$$\epsilon = 10^{-4}, \quad \theta = 47.3^o, \quad b = (\cos\theta, \sin\theta), \quad c = 0, \quad f = 0.$$

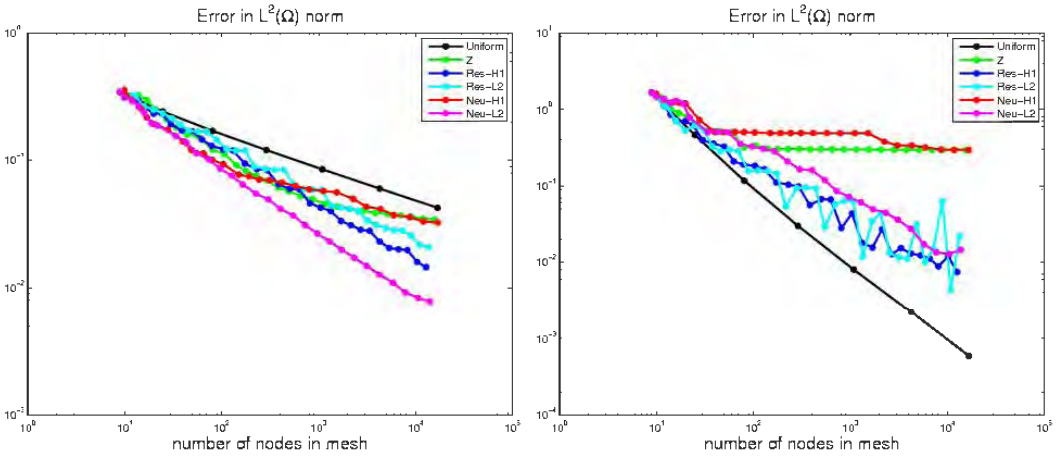

Figure 6.11: Example 2C with $\omega = 10^{-2}$ and $\epsilon = 10^{-7}$: Global errors of uniform and local refinement in the state (left) and control (right) in L^2 sense.

We apply Dirichlet conditions on the entire boundary (i.e. $\Gamma_D = \partial\Omega$, $\Gamma_N = \emptyset$) and the Dirichlet boundary data are

$$g_D(x_1, x_2) = \begin{cases} 1 & \text{if } x_1 = 0 \text{ and } x_2 \leq .25 \\ 1 & \text{if } x_2 = 0 \\ 0 & \text{else.} \end{cases}$$

Recall the physical setup of the uncontrolled problem in Figure 4.11. In the optimal control case we define the desired state to be

$$\widehat{y}(x_1, x_2) = 1.$$

Hence the optimization problem seeks (y, u) such that $\frac{1}{2}\|y - 1\|^2_{L^2(\Omega)} + \frac{\omega}{2}\|u\|^2_{L^2(\Omega)}$ is minimal. Since $f = 0$ the control can be seen as the sole forcing term to this system. The penalty parameter ω controls how 'big' u is allowed to be. The smaller ω the larger $\|u\|^2_{L^2(\Omega)}$.

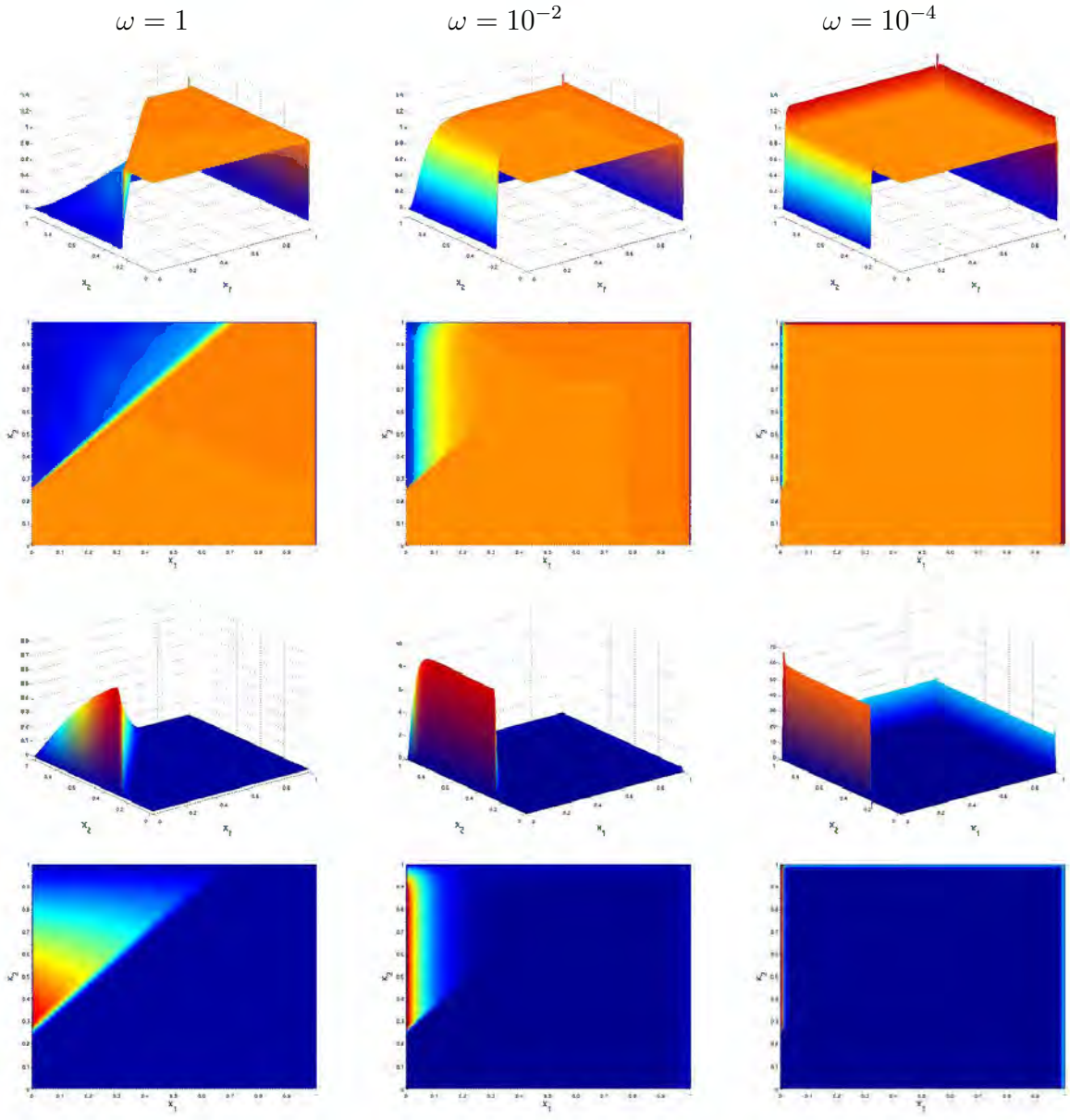

Figure 6.12: Example 2D with $\omega = 1, 10^{-2}, 10^{-4}$ and $\epsilon = 10^{-4}$: Computed state (top 2 rows) and control (bottom 2 rows) on uniform meshes of 129×129 nodes (fixed ω per column).

The exact optimal state, control and adjoint for this problem are not known. The computed solutions on a uniform grid (129×129) for $\omega = 1, 10^{-2}, 10^{-4}$ are shown in Figure 6.12. The computed state exhibits a sharp boundary layer, as seen in Example 1C. For small ω the control (and therefore the adjoint) is concentrated in a small region near the boundary. For small ω the state is equal to one, except for a small region around the boundary where $y = 0$. The inner layer in the state disappears, but the boundary layer become more pronounced.

Figures 6.13, 6.14 and 6.15 show the refined meshes for $\omega = 1, 10^{-2}, 10^{-4}$ respectively. For large ω, the mesh refinement has to resolve the boundary layer in the state as well as the interior layer. For small ω the mesh refinement has to resolve the boundary layers.

Except for the ZZ estimators, all estimators seem to locate the boundary layers fairly well. However, the interior layer is picked up better by the L^2 estimators. This is a result of all the spurious oscillations around the boundary layers of the state and the control. The oscillations dominate the H^1 norm, and therefore also their estimators.

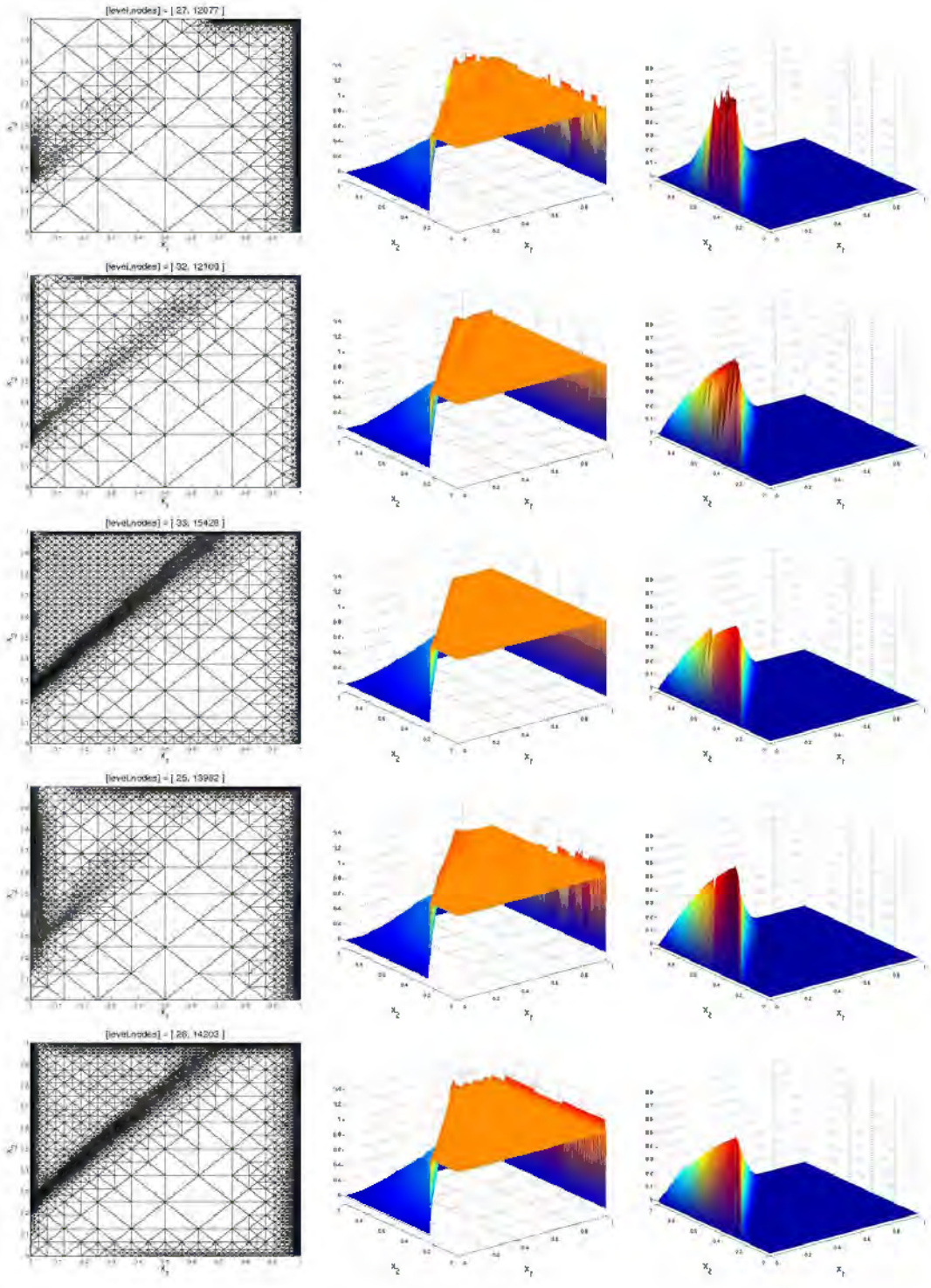

Figure 6.13: Example 2D with $\omega = 1$ and $\epsilon = 10^{-4}$: Computed locally refined mesh (left column), state (middle column) and control (right column) using $N_{\max} = 12000$. From top to bottom the ZZ, Res-H1, Res-L2, Neu-H1 and Neu-L2 estimator.

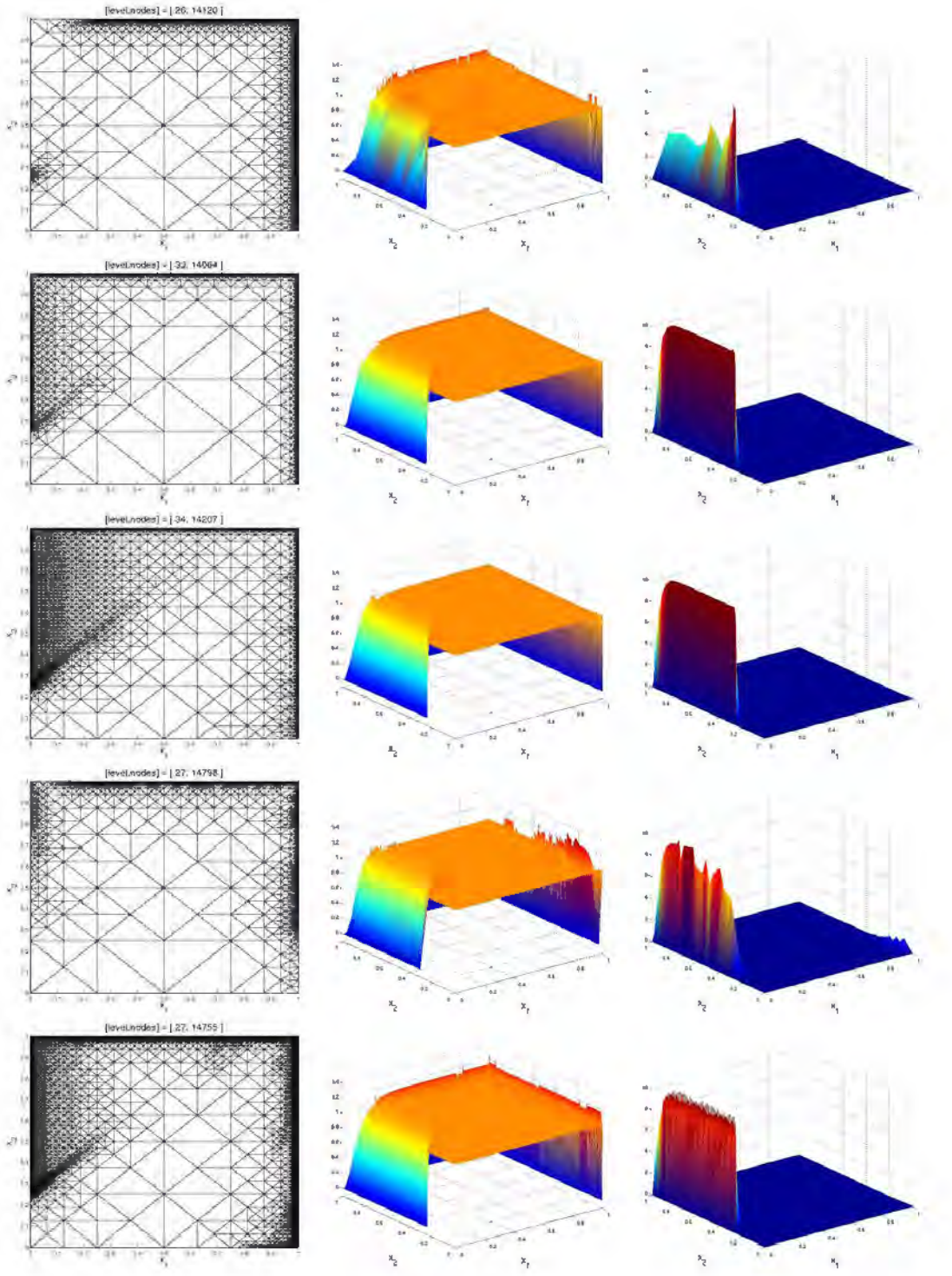

Figure 6.14: Example 2D with $\omega = 10^{-2}$ and $\epsilon = 10^{-4}$: Computed locally refined mesh (left column), state (middle column) and control (right column) using $N_{\max} = 12000$. From top to bottom the ZZ, Res-H1, Res-L2, Neu-H1 and Neu-L2 estimator.

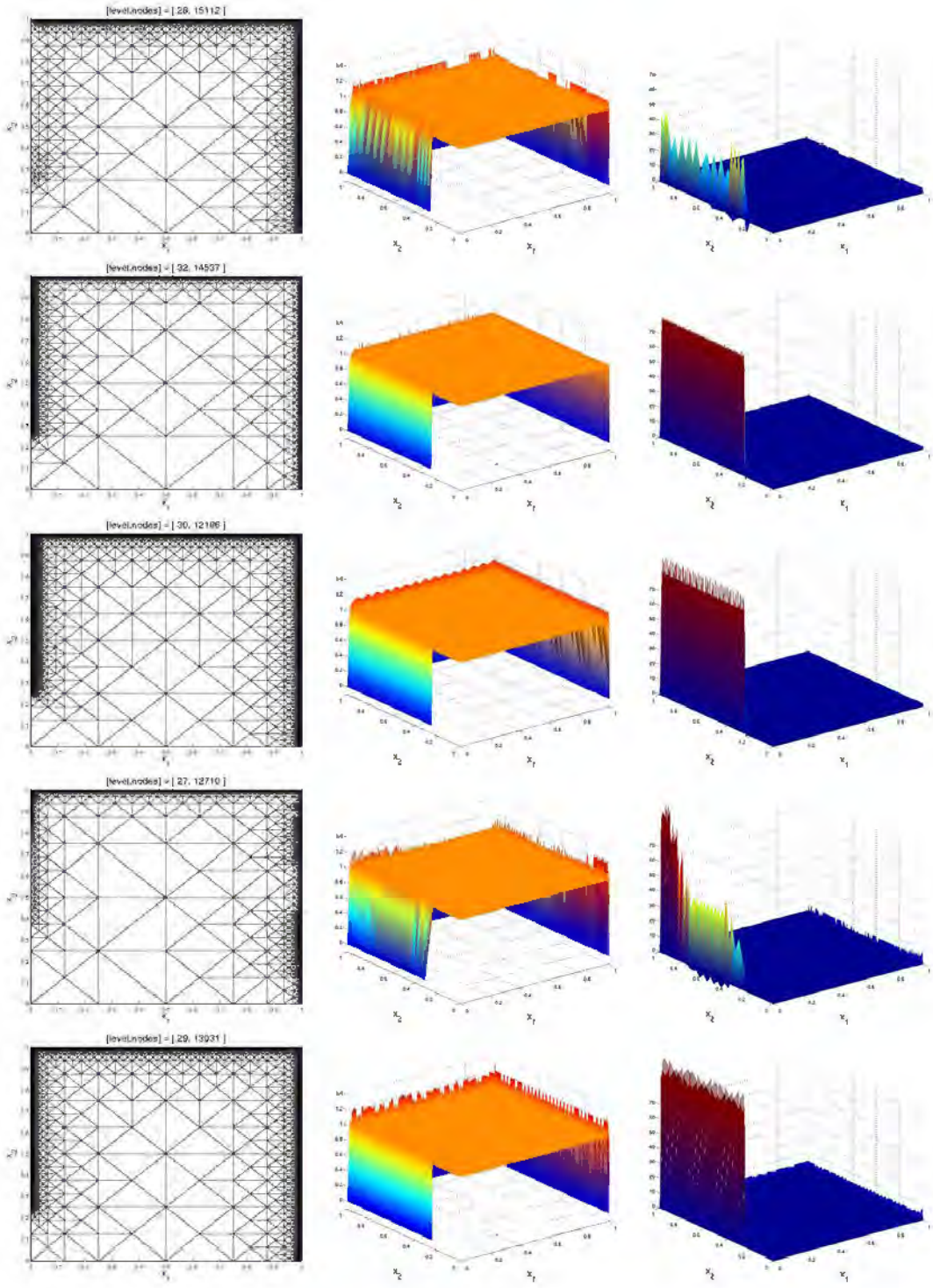

Figure 6.15: Example 2D with $\omega = 10^{-4}$ and $\epsilon = 10^{-4}$: Computed locally refined mesh (left column), state (middle column) and control (right column) using $N_{\max} = 12000$. From top to bottom the ZZ, Res-H1, Res-L2, Neu-H1 and Neu-L2 estimator.

In conclusion, solutions to linear-quadratic elliptic optimal control problems governed by convection diffusion equations typically exhibit layers in state and control. Locating these regions, and combining information on multiple layers is not a straightforward task. Though several estimators produce satisfying meshes in some degree, it appears that gaining more control over the spurious oscillations could result more accurate numerical solutions.

Chapter 7

Conclusions

Solutions to convection-diffusion equations typically exhibit small regions with steep gradients, so-called inner or boundary layers. The numerical solution of these problems using finite elements requires the introduction of stabilization terms, such as the SUPG stabilization. Although SUPG stabilized FEM produces significantly better solutions at moderate mesh sizes than standard Galerkin FEM, the numerical solution may still exhibit oscillations in a small neighborhoods around the layers when those are not resolved. These errors can be propagated downwind from the layer by the convection and can cause significant loss of accuracy of the numerical solution. Therefore this thesis uses SUPG stabilization in conjuncture with local mesh refinement. The goal of the local refinement is to resolve layers, so no oscillations occur and no errors can be propagated.

This work compared the adaptive finite element method driven by the ZZ, the

norm-residual based and the local Neumann error estimator. Because of the error propagation, all estimators place elements away from the layers. Though local Neumann estimators appeared to work better in resolving a single inner or boundary layer, it lacked the capacity to sufficiently resolve multiple layers at the same time. Hence, this thesis has shown that among these estimators there is no clear favorite for the single equation case.

Solving linear-quadratic convection dominated elliptic optimal control problems is more complicated since it requires the solution of a coupled system of two convection diffusion equations and an algebraic equation. The convection terms in the convection diffusion equations have opposite signs. The solution to this coupled system, with state, adjoint and control components, typically exhibits multiple layers. Again, when applying SUPG stabilization the solution contains spurious oscillations in a small band around the layers, unless these layers are resolved. Since the two PDEs have convection in opposite direction, the spurious oscillations are propagated down- and upwind. This work illustrated this effect by computation. Hence, in this coupled setting it is even more important to resolve the layers.

This work compared adaptive refinements driven by the same error estimators as in the single equation case. In the literature only the norm-residual based estimator has been extended to the optimal control setting. This thesis reviewed this theory, and proposed a generalization for the ZZ and local Neumann estimator. Though a solid mathematical justification is still needed, this work showed that the latter ap-

proach can compete with the norm-residual based estimator. Similar to the single equation case, local Neumann estimators are able to resolve layers better, however, they struggle when multiple layers are present. As expected there is not a single favorite estimator. Hence, choosing an error indicator is also highly problem dependent in the optimal control setting.

Results shown in this work indicate that there is still much work to be done in this field. As mentioned, the local Neumann estimator introduced in this work still needs to be justified by proving error bounds (3.1). Moreover, this estimator can be improved by incorporating more problem features in the local Neumann problem. Another option is to solve the local problem using SUPG stabilization rather than Galerkin alone. Goal-oriented error estimation [4] is another technique for error indication which is worth exploring.

The solution components, state, control, and adjoint, often have different scales and exhibit layers in different parts of the domain. Therefore, error estimators that estimate the error for all three solution components jointly need to scale the individual component error estimators carefully, which is difficult to do. Alternatively, one can thing of using different meshes for the state, the control, and the adjoint and refine them based on error estimates for the individual components. While some work in this area exists, the advantages and disadvantages needs to be explored more.

This thesis uses continuous finite element methods in conjuncture with SUPG stabilization because it is widely used in applications. However, there are other methods

which could yield good results. Different stabilization methods exists. Moreover, discontinuous Galerkin (DG) methods ([16], [31]) could be used as well.

Bibliography

[1] D. N. Arnold, A. Mukherjee, and L. Pouly. Locally adapted tetrahedral meshes using bisection. *SIAM J. Sci. Comput.*, 22(2):431–448 (electronic), 2000.

[2] W. Bangerth, R. Hartmann, and G. Kanschat. deal.II — a general-purpose object-oriented finite element library. *ACM Trans. Math. Softw.*, 33(4).

[3] W. Bangerth, R. Hartmann, and G. Kanschat. `deal.II` *Differential Equations Analysis Library, Technical Reference*. Interdisziplinäres Zentrum für Wissenschaftliches Rechnen (IWR). `http://www.dealii.org`.

[4] W. Bangerth and R. Rannacher. *Adaptive finite element methods for differential equations*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 2003.

[5] E. Bänsch. An adaptive finite-element strategy for the three-dimensional time-dependent Navier-Stokes equations. *J. Comput. Appl. Math.*, 36(1):3–28, 1991.

[6] E. Bänsch. Local mesh refinement in 2 and 3 dimensions. *Impact Comput. Sci. Engrg.*, 3(3):181–191, 1991.

[7] S. Bartels and C. Carstensen. Each averaging technique yields reliable a posteriori error control in FEM on unstructured grids. I. Low order conforming, nonconforming, and mixed FEM. *Math. Comp.*, 71(239):945–969 (electronic), 2002.

[8] R. Becker, M. Braack, D. Meidner, R. Rannacher, and B. Vexler. Adaptive finite element methods for PDE-constrained optimal control problems. In W. Jäger, R. Rannacher, and J. Warnatz, editors, *Reactive flows, diffusion and transport*, pages 177–205. Springer, Berlin, 2007.

[9] S. C. Brenner and C. Carstensen. Finite element methods. In E. Stein, R. de Borst, and T. J. R Hughes, editors, *Encyclopedia of Computational Mechanics*. John Wiley and Sons, New-York, Chicester, Brisbane, Toronto, 2004.

[10] L. Chen. *i*FEM: an innovative finite element methods package in MATLAB. *In Preparation*, 2008. Retrieved January 31, 2009, http://math.uci.edu/∼chenlong/iFEM.html.

[11] L. Chen. Short implementation of bisection in MATLAB. In Palle Jorgensen, Xiaoping Shen, Chi-Wang Shu, and Ningning Yan, editors, *Recent Advances in Computational Sciences – Selected Papers from the International Workship on Computational Sciences and Its Education*, pages 318 – 332, 2008.

[12] L. Chen and C.-S. Zhang. AFEM@matlab: a Matlab package of adaptive finite element methods. *Technique Report, Department of Mathemat-*

ics, University of Maryland at College Park, 2006. Retrieved June 1, 2008, http://math.uci.edu/∼chenlong/.

[13] S. S. Collis and M. Heinkenschloss. Analysis of the streamline upwind/petrov galerkin method applied to the solution of optimal control problems. Technical Report TR02–01, Department of Computational and Applied Mathematics, Rice University, Houston, TX 77005–1892, 2002. Retrieved Februari 4, 2009, http://www.caam.rice.edu/∼heinken.

[14] H. C. Elman, D. J. Silvester, and A. J. Wathen. *Finite Elements and Fast Iterative Solvers with applications in incompressible fluid dynamics*. Oxford University Press, Oxford, 2005.

[15] M. Heinkenschloss and D. Leykekhman. Local error estimates for supg solutions of advection-dominated elliptic linear-quadratic optimal control problems. Technical Report TR08–30, Department of Computational and Applied Mathematics, Rice University, 2008.

[16] J. S. Hesthaven and T. Warburton. *Nodal Discontinuous Galerkin Methods: Analysis, Algorithms, and Applications*. Springer-Verlag, Berlin, 2008.

[17] M. Hinze, N. Yan, and Z. Zhou. Variational discretization for optimal control governed by convection dominated diffusion equations. *J. Comp. Math.*, 27(2-3):237–253, 2009.

[18] T. J. R. Hughes and A. N. Brooks. Streamline upwind / Petrov Galerkin formulations for the convection dominated flows with particular emphasis on the incompressible navier-stokes equations. *Comput. Methods. Appl. Mech. Engrg.*, 54:199–259, 1982.

[19] T. J. R. Hughes and A. N. Brooks. A theoretical framework for Petrov–Galerkin methods with discontinuous weighting functions: Application to the streamline–upwind procedure. In R. H. Gallagher, D. H. Norrie, J. T. Oden, and O. C. Zienkiewicz, editors, *Finite Elements in Fluids - Volume 4*, pages 47–65. John Wiley & Sons, New York, 1982.

[20] V. John. A numerical study of a posteriori error estimators for convection-diffusion equations. *Comput. Methods Appl. Mech. Engrg.*, 190(5-7):757–781, 2000.

[21] V. John and P. Knobloch. On spurious oscillations at layers diminishing (SOLD) methods for convection-diffusion equations: Part I - a review. *Comput. Methods Appl. Mech. Engrg.*, 196:2197–2215, 2007.

[22] V. John and P. Knobloch. On spurious oscillations at layers diminishing (SOLD) methods for convection-diffusion equations: Part II - analysis for P_1 and Q_1 finite elements. *Comput. Methods Appl. Mech. Engrg.*, 197:1997–2014, 2008.

[23] D. Kay and D. Silvester. The reliability of local error estimators for convection diffusion equations. *IMA J. of Numer. Anal.*, 21:107–122, 2001.

[24] P. Knabner and L. Angermann. *Numerical Methods for Partial Differential Equations*. Texts in Applied Mathematics, Vol. 44. Springer–Verlag, Berlin, Heidelberg, New York, 2003.

[25] P. Knobloch. On the choice of the supg parameter at outflow boundary layers. Technical Report MATH-knm-2007/3, Charles University, Faculty of Mathematics and Physics, Prague, 2007.

[26] J.-L. Lions. *Optimal Control of Systems Governed by Partial Differential Equations*. Springer Verlag, Berlin, Heidelberg, New York, 1971.

[27] J.-L. Lions. *Some Aspects of the Optimal Control of Distributed Parameter Systems*. SIAM, Philadelphia, PA, 1972.

[28] K. Mekchay and R.H. Nochetto. Convergence of adaptive finite element methods for general second order linear elliptic PDEs. *SIAM J. Numer. Anal.*, 43(5):1803–1827, 2005.

[29] P. Morin, R.H. Nochetto, and K.G. Siebert. Data oscillation and convergence of adaptive FEM. *SIAM J. Numer. Anal.*, 38(2):466–488, 2001.

[30] A. Papastavrou and R. Verfürth. A posteriori error estimators for stationary convection-diffusion problems: a computational comparison. *Comput. Methods Appl. Mech. Engrg.*, 189(2):449–462, 2000.

[31] B. Rivière. *Discontinuous Galerkin methods for solving elliptic and parabolic equations. Theory and implementation*, volume 35 of *Frontiers in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2008.

[32] H. G. Roos, M. Stynes, and L. Tobiska. *Robust Numerical Methods for Singularly Perturbed Differential Equations*. Computational Mathematics, Vol. 24. Springer–Verlag, Berlin, second edition, 2008.

[33] M. Stynes. Steady-state convection-diffusion problems. In A. Iserles, editor, *Acta Numerica 2005*, pages 445–508. Cambridge University Press, Cambridge, London, New York, 2005.

[34] R. Verfürth. *A Review of A Posteriori Error Estimation and Adaptive Mesh–Refinement Techniques*. Wiley-Teubner Series: Advances in Numerical Mathematics. Wiley Teubner, Chicester, New York, Stuttgart, 1996.

[35] R. Verfürth. A posteriori error estimates for convection-diffusion equations. *Numer. Math.*, 80:641–663, 1998.

[36] R. Verfürth. Robust a posteriori error estimates for nonstationary convection-diffusion equations. *SIAM J. Numer. Anal.*, 43(4):1783–1802 (electronic), 2005.

[37] R. Verfürth. Robust a posteriori error estimates for stationary convection-diffusion equations. *SIAM J. Numer. Anal.*, 43(4):1766–1782 (electronic), 2005.

[38] N. Yan and Z. Zhou. A priori and a posteriori error estimates of streamline diffusion finite element method for optimal control problem governed by convection dominated diffusion equation. *Numer. Math. Theor. Meth. Appl.*, 1(3):297–320, 2008.

[39] O.C. Zienkiewicz and J.Z. Zhu. A simple error estimator and adaptive procedure for practical engineering analysis. *Internat. J. Numer. Meth. Engrg.*, 24(2):337–357, 1987.