Adaptive defect correction methods for convection dominated, convection diffusion problems

Document status and date:
Published: 01/01/1998

Document Version:
Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

Download date: 16. Nov. 2023

EINDHOVEN UNIVERSITY OF TECHNOLOGY
Department of Mathematics and Computing Science

RANA 98-29
December1998

Adaptive Defect Correction Methods for Convection Dominated, Convection Diffusion Problems

by

V. Ervin, W. Layton, and J. Maubach

ADAPTIVE DEFECT CORRECTION METHODS

FOR CONVECTION DOMINATED,

CONVECTION DIFFUSION PROBLEMS

V. Ervin[1], W. Layton[2] and J. Maubach[3]

ABSTRACT. We present a posteriori error estimators for a defect correction method for approximating solutions of convection diffusion problems. The algorithms and estimators include the possibility of using in the discretization a nonlinear selection mechanism, which we find improves solution quality in and near layers. Energy norm and L^2 a posteriori error estimates are proven for the full algorithm. Two examples of fully adaptive finite element-defect correction calculations are presented. These examples illustrate the scheme as well as the reliability of the derived estimators.

1. Introduction.

This report considers the problem of computing efficiently and to within a preassigned error tolerance an approximate solution to the singularly perturbed, that is, convection dominated, convection diffusion equations:

$$\mathcal{L}_\epsilon u := -\epsilon \Delta u + \mathbf{v} \cdot \nabla u + gu = f, \text{ in } \Omega, \tag{1.1}$$

$$u = 0, \text{ on } \Gamma. \tag{1.2}$$

In (1.1), (1.2) $\Omega \subset \mathbb{R}^2$ is a polygonal domain with boundary Γ, $\mathbf{v}$ is a given vector field on $\bar{\Omega}$ and f and g are known functions on $\bar{\Omega}$. We specifically focus on the case when (1.1) is convection dominated, i.e. $\epsilon << O(h)$, where h is a realizable global (or outer) meshwidth.

[1] Dept. of Mathematical Sciences, Clemson University, Clemson, S.C. 29634; email: ervin@math.clemson.edu

[2] Partially supported by NSF Grant DMS-9400057; Dept. of Mathematics, University of Pittsburgh, Pittsburgh, PA 15260; email:wjl+@pitt.edu

[3] Dept. of Mathematics and Computer Science, Eindhoven University of Technology, NL-5600MB Eindhoven, the Netherlands; email: maubach@win.tue.nl

It is well known that discretization of (1.1), (1.2) for small ϵ is fraught with difficulties: low order "upwind" or "donor cell" type discretizations are quite inefficient and produce grossly smeared solutions of low quality while higher order ("centered") approximations typically exhibit nonphysical oscillations. Further, even with a "good" discretization method such as the streamline diffusion method or a defect correction method, the overall accuracy is deteriorated by the presence of sharp boundary and interior layers in the true solution of (1.1), (1.2). The clear solution to this problem is to place more mesh points in the small regions where the solution is less regular. To find the regions in which a finer mesh is needed, and the degree of refinement needed there, requires an error estimator which can be computed from the approximate solution and the problem data.

Reliability of the a posteriori error estimator, meaning that the estimated error is a true upper bound to the true error, is essential. Further, for efficiency, the computation of the estimator should be far less expensive than simply calculating another approximate solution on a further refined mesh.

It is necessary that error estimation and mesh redistribution takes place in the context of a "good" discretization method. Minimally, the method should have high accuracy in smooth regions, well-supported by local error analysis to elucidate the essential requirements, and be globally stable. Further, it is also highly desirable that some sort of nonlinearity is introduced in the scheme to control over- and undershoots near the layers, [29, 30]. (Otherwise a reliable mesh refinement process will refine around these nonphysical oscillations until they are reduced by brute force - clearly not optimally efficient.) We therefore introduce into our approximation a mechanism to nearly eliminate these over and undershoots and thereby control excessive refinement *near* layers.

The streamline diffusion finite element method [9,10,15,22,23,24] possibly coupled with a nonlinear shock capturing mechanism is a powerful technique for the approximate solution of (1.1), (1.2) - especially coupled with the a posteriori error estimators developed by Eriksson and Johnson [10]. In this report we consider instead a discretization strategy based upon a defect correction, finite element method for (1.1), (1.2). This method was developed by Hemker [12] and used extensively, see, for example, [12,13,14,16], to solve high Reynolds number compressible flow problems. For local and global a priori error estimates

for defect correction methods see [2,3,11,18]. Because of the simple structure of the basic defect correction procedure, Algorithm 1.1, we are able to introduce a nonlinear selection mechanism into the scheme, Algorithm 1.2, without increasing its overall complexity.

To present the basic defect correction algorithm and the modification we study, let $\Pi_j^h(\Omega), j \geq 1$, denote a series of edge-to-edge finite element triangulation of Ω, with $X_j^h \subset \overset{\circ}{H}^1(\Omega)$ denoting a conforming finite element space based upon that mesh. (In the computational experiments we present X_j^h will be either conforming linears, quadratics or cubics.) Define the usual and artificial viscosity bilinear forms:

$$a_\epsilon(u,v) := \int_\Omega \epsilon \nabla u \cdot \nabla v + (\mathbf{v} \cdot \nabla u + gu)v \, dx$$

$$a_{\epsilon_0}(u,v) := \sum_{T \in \Pi^h(\Omega)} \int_T \epsilon_0(T, \epsilon) \nabla u \cdot \nabla v + (\mathbf{v} \cdot \nabla u + gu)v \, dx,$$

where, for example, the artificial viscosity parameter $\epsilon_0(T, \epsilon)$ can be chosen as

$$\epsilon_0(T, \epsilon) = \max\{|\mathbf{v}|_{L^\infty(T)} \operatorname{diam}(T), \epsilon\}.$$

The basic defect correction algorithm [2,3,11,12-14,16] then proceeds as follows. First the global solution envelop is captured via an artificial viscosity approximation. This is then "anti-diffused" J (= polynomial degree $(X^h) + 1$) times. Note that at each step only the matrix arising from the artificial viscosity discretization need be inverted.

3

> **Algorithm 1.1: Basic Defect Correction Method (D.C.M.) for (1.1).**
>
> 1. Calculate $U^1 \in X_1^h$ satisfying
>
> $$a_{\epsilon_0}(U^1, v) = (f, v), \text{ for all } v \in X_1^h.$$
>
> 2. For $j = 2, \cdots, J$ calculate $U^j \in X_j^h$ satisfying
>
> $$a_{\epsilon_0}(U^j - U^{j-1}, v) = (f, v) - a_\epsilon(U^{j-1}, v), \text{ for all } v \in X_j^h.$$

It has been proven (see [2,3,11] for details) this algorithm produces an approximate solution u^J which converges, uniformly in ϵ in smooth regions $\Omega' \subset \Omega$, to u at rate $O(h^k)$ in $H^1(\Omega')$ and $O(h^{k+1/2})$ in $L^2(\Omega')$ where k = polynomial degree (X_j^h). It has also been observed [11,12] that this basic algorithm tends to antidiffuse too much near layers (resulting in oscillations near layers) and needs to be modified to incorporate some sort of nonlinear selection mechanism.

The nonlinear selection mechanism we shall employ involves the use of a (nonlinear) p-Laplacian, incorporated into the residual calculation, to limit the antidiffusion in regions where $\text{grad}(u^j) = O(h^{-1})$. This use of the p-Laplacian is natural for finite element methods. Analogous uses of p-Laplacians have occurred quite early in the global circulation models of Smagorinski [26] and other subgridscale models [19]. Define the nonlinear functional $AV_p(\cdot, \cdot)$ by:

$$AV_p(u, v) := \sum_{T \in \Pi^h(\Omega)} \int_T \mu_0 \, \text{diam}\,(T) |\text{diam}\,(T) \nabla u|^p \nabla u \cdot \nabla v \, dx.$$

In all our experiments we take $p = k$, where k is the polynomial degree of X^h.

1. Calculate $U^1 \in X_1^h$ satisfying

$$a_{\epsilon_0}(U^1, v) = (f, v), \text{ for all } v \in X_1^h.$$

2. For $j = 2, \cdots, J$ calculate $U^j \in X_j^h$ satisfying

$$a_{\epsilon_0}(U^j - U^{j-1}, v) = (f, v) - a_\epsilon(U^{j-1}, v) - AV_p(U^{j-1}, v), \text{ for all } v \in X_j^h.$$

The main result of this report is a reliable a posteriori error estimator for Algorithm 1.2 (and also for Algorithm 1.1 by taking $\mu_0 = 0$). To present the estimator, let r^j denote the strong (local) residual of U^j:

$$r^j := f - (-\epsilon \Delta U^j + \mathbf{v} \cdot \nabla U^j + g U^j),$$

and let $[w]_e$ denote the jump of a function w a cross an edge e in the finite element triangulation $\Pi_j^h(\Omega)$. Let E_j^h denote the set of edges in the finite element mesh $\Pi_j^h(\Omega)$. The estimators we derive in Corollaries 3.1 and 3.2 take the following form: for computable constants C_R, C_J and C_U

$$\|u - U^j\|_{L^2(\Omega)}^2 \leq C_R \sum_{T \in \Pi_j^h(\Omega)} \epsilon^{-1} h_T^2 \|r^j\|_{L^2(T)}^2 +$$

$$+ C_J \sum_{e \in E_j^h} \min\{\epsilon^{-1} h_e^3, \epsilon h_e\} \|[\nabla U^j \cdot \hat{n}_e]\|_{L^2(e)}^2$$

$$+ C_U \sum_{T \in \Pi_j^h(\Omega)} \min\{\epsilon^{-1}(\epsilon_0(T) - \epsilon)^2, h_T^{-2}(\epsilon_0(T) - \epsilon)^2\} * \qquad (1.3)$$

$$* \|\nabla(U^j - U^{j-1}) + \mu_0 h |h \nabla U^{j-1}|^p \nabla U^{j-1}\|_{L^2(T)}^2 \equiv Est^2.$$

The constants $C_{R,J,U}$ are computable in terms of tabulated upper bound constants in various inequalities and ground states of certain simple eigenvalue problems. Since the worse case in each step of the derivation of (1.3) seldom occur simultaneously, it can be usual in practical settings to estimate $C_{R,J,U}$ by solving several small problems with known solutions followed by data fitting.

There are at least two important issues connected with the estimator (1.3). The first issue concerns the reliability and efficiency of the estimator. Is $\|u - U^j\|_{L^2(\Omega)} \leq Est$ in all cases and does Est accurately reflect the true error? The second issue is how the information in Est is used to generate an adaptive algorithm based upon equidistribution of Est. To this end, we have used a simple and robust n-dimensional mesh refinement and de-refinement procedure developed in [20,21]. In the test reported herein the refinement criteria used in conjunction with the mesh generator of [20,21] is as follows. Define, for T a triangle in $\Pi_j^h(\Omega)$,

$$Est(T)^2 := \tilde{C}_R \epsilon^{-1} h_T \|r^j\|_{L^2(T)}^2 + \tag{1.3:b}$$
$$+ \frac{1}{2} \tilde{C}_J \sum_{\substack{\text{all edges} \\ e \text{ of } T}} \min\{\epsilon^{-1} h_e^3, \epsilon h_e\} \int_e [\nabla U^j \cdot \hat{n}_e]_e^2 de +$$
$$+ \tilde{C}_U \min\{\epsilon^{-1}, h_T^{-2}\} * \|(\epsilon_0(T) - \epsilon)\nabla(U^j - U^{j-1}) + \mu_0 h |h\nabla U^{j-1}|^p \nabla U^{j-1}\|_{L^2(T)}^2 .$$

The decision procedure used is then quite simply given Tol (global tolerance) and $LocTol_j^2 = Tol^2 / (\text{no. of triangles in } \Pi_j^h(\Omega))$

```
Algorithm 1.3: j = 1

Compute U^j, Est(T)

  If Est < Tol, then STOP
  Else
    For T ∈ Π_j^h(Ω)
    If Est^2(T) > LocTol_j^2, mark T for refinement.
    If Est^2(T) < ½ LocTol_j^2, mark T for de-refinement.
    (*) Refine and de-refine Π_j^h(Ω) to obtain Π_{j+1}^h(Ω),
       j = j + 1
       Continue.
```

The mesh generator used in step (*) in the previous algorithm is critical; we have used a conservative remeshing strategy biased to refinement in our tests. Different final meshes are generated depending upon the remeshing algorithm used at (*). (Sections 4 and 5 present some of our conclusions (and speculations) on possible improvements to the procedure we used.)

Although changes in how the information contained in the estimator (1.3;b) is *used* in (*) is possible as noted above, (1.3;b) did prove to be a reliable estimator and quite reasonable grids were generated by this simple procedure, as Section 4 illustrates.

2.1 Notation and Preliminaries.

This preliminary section records some basic notation and introduces a posteriori error estimation through energy norm error analysis for the usual Galerkin formulation. This preliminary error analysis shows that, although the estimator derived provides both upper and lower estimates of the true errors, for problems with large skew symmetric parts ($\epsilon \ll O(1)$) the respective constants differ by $O(\epsilon^{-1/2})$. This effect is shared by our estimators for the defect correction procedure. (It seems to arise essential from the mathematical framework used for a posteriori error analysis and Lemma 2.1.) We thus focus in the remainder of this report on computable upper estimates.

The space X throughout will denote $X := \overset{\circ}{H}{}^1(\Omega)$. The finite element spaces $X_j^h, j = 1, \cdots, J$, are assumed to be conforming $X_j^h \subset X$. They are associated with meshes $\Pi_j^h(\Omega)$ which are constructed based upon refinement and de-refinement of an initial mesh. The set of internal edges of triangles in $\Pi_j^j(\Omega)$ is denoted E_j^h. For D a subdomain (typically one triangle), $\| \cdot \|_D$ will denote the $L^2(D)$ norm and $(\cdot, \cdot), \| \cdot \|$ the usual $L^2(\Omega)$ inner product and norm. The operator $I^h : X \rightarrow X_j^h$, (henceforth, suppressing the "j"), is defined by interpolating suitable local averages of v in X_j^h. This interpolant was introduced by Clément [7] and satisfies (see [7]) for T in Π_j^h, e in $E_j^h, h_T = \text{diam}\,(T)$, and $h_e = \text{diam}\,(e)$

$$\|\nabla^i v^h\|_T \leq C_1 \|\nabla^i v\|_{N(T)}, i = 0, 1, \tag{2.1}$$

$$\|v - v^h\|_T \leq C_2 h_T^i \|\nabla^i v\|_{N(T)}, i = 1, 2, \tag{2.2}$$

$$\|\nabla(v - v^h)\|_T \leq C_3 h_T^{i-1} \|\nabla^i v\|_{N(T)}, i = 1, 2, \tag{2.3}$$

$$\|\nabla v^h\|_T \leq C_4 h_T^{-1} \|v^h\|_{N(T)}, \tag{2.4}$$

$$\|v - v^h\|_e \leq C_5 h_e^{i-1/2} \|\nabla^i v\|_{N(e)}, i = 1, 2, \tag{2.5}$$

where $\nabla^0 v = v$, $\nabla_v^1 := \nabla v$ and $|\nabla^2 v| := (v_{xx}^2 + 2v_{xy} + v_{yy}^2)^{1/2}$ and $N(\nu)$ denotes the neighborhood of the element or edge ν consisting of those elements K in Π_j^h sharing at least one node with ν. We also introduce the constant $C_7 := C_7(\theta_{\min}(\Pi_j^h))$ which satisfies the conditions

$$\left(\sum_{T \in \Pi_j^h} \|w\|_{N(T)}^2 \right)^{1/2} \leq C_7 \|w\|, \quad \left(\sum_{e \in E_j^h} \|w\|_{N(e)}^2 \right)^{1/2} \leq C_7 \|w\|. \tag{2.6}$$

The usual Galerkin finite element approximation to (1.1), (1.2) is the function $U^h \in X^h$ satisfying

$$a_\epsilon(U^h, v) = (f, v), \text{ for all } v \in X^h.$$

It is well known that U^h is typically an oscillatory, low quality approximation to $u(x, y)$. Nevertheless, the estimators for U^h share some common features with those for both the streamline diffusion method, [10], and Algorithms 1.1 and 1.2. Thus, we shall introduce some notation and techniques by considering estimators for the standard Galerkin method.

8

One universal effect that the analysis of the "centered" method illustrates is that there is an $O(\epsilon^{-1/2})$ gap between computable upper and lower error bounds for ϵ small, see Theorem 2.1 below. The cause of this gap is that one bound comes naturally weighted with $\|\mathcal{L}_\epsilon\|_{\mathcal{L}(H^1, H^{-1})}$ (which is $O(1)$), and the other bound with $\|\mathcal{L}_\epsilon\|_{\mathcal{L}(H^{-1}, H^1)}$, which is $O(\epsilon^{-1/2})$). This gap does not occur for symmetric problems.

Lemma 2.1. *Define the norm on X, for $a > 0$ $b > 0$*

$$\||w\||_{a,b}^2 := \int_\Omega a\|\nabla w\|^2 + bw^2 \, dx.$$

Assume $0 < g_0 \leq g - \frac{1}{2}\nabla \cdot \mathbf{v}$, and $|g - \nabla \cdot \mathbf{v}| \leq g_1$. Then

$$\||w\||_{\epsilon,g_0} \leq \sup_{0 \neq v \in X} \frac{a_\epsilon(w,v)}{\||v\||_{\epsilon,g_0}}. \tag{2.7}$$

Further,

$$\sup_{0 \neq v \in X} \frac{a_\epsilon(w,v)}{\|v\|_1} \leq \sqrt{2}\max\{\epsilon^{1/2}, g_0^{-1/2}(g_1 + \|\mathbf{v}\|_\infty)\}\||w\||_{\epsilon,g_0}, \tag{2.8}$$

$$\sup_{0 \neq v \in X} \frac{a_\epsilon(w,v)}{\||v\||_{\epsilon,g_0}} \leq \sqrt{2}\max\{\epsilon^{1/2}, g_0^{-1/2}(\|g\|_\infty + \|\mathbf{v}\|_\infty)\}\|w\|_1. \tag{2.9}$$

Proof. The inequality (2.7) follows trivially from coercivity of $a_\epsilon(\cdot,\cdot)$ in the $\||\cdot\||_{\epsilon,g_0}$ norm. For the upper bound (2.8), note that

$$a_\epsilon(w,v) = \int_\Omega \epsilon\nabla w \cdot \nabla v + (\mathbf{v}\cdot\nabla w + gw)v dx$$

$$= \int_\Omega \epsilon\nabla w \cdot \nabla v + (g - \nabla\cdot\mathbf{v})wv dx - \int_\Omega \mathbf{v}\cdot\nabla v w dx$$

$$\leq \epsilon\|\nabla w\|\,\|\nabla v\| + \|g - \nabla\cdot\mathbf{v}\|_\infty\|w\|\|v\| + \|\mathbf{v}\|_\infty\|\nabla v\|\|w\|,$$

$$\leq [\epsilon\|\nabla w\| + (g_1 + \|\mathbf{v}\|_\infty)\|w\|]\|v\|_1,$$

$$\leq \sqrt{2}[\epsilon^2\|\nabla w\|^2 + (g_1 + \|\mathbf{v}\|_\infty)^2\|w\|^2]^{1/2}\|v\|_1,$$

from which (2.8) then follows.

The inequality (2.9) is established in an analogous fashion. $\square$

Theorem 2.1. *Assume* $0 < g_0 \leq g(x) - 1/2 \nabla \cdot \mathbf{v}(x)$. *Then, there is a constant C, independent of h and ϵ, such that*

$$\||u - U^h\||_{\epsilon,g_0} \leq \sup_{0 \neq v \in X} \inf_{v^h \in X^h} \frac{(f, v - v^h) - a_\epsilon(U^h, v - v^h)}{\||v\||_{\epsilon,g_0}} \leq C\epsilon^{-1/2} \||u - U^h\||_{\epsilon,g_0}. \quad (2.10)$$

Proof. Using the identity $a_\epsilon(u - U^h, v) = (f, v) - a_\epsilon(U^h, v)$ for all $v \in X$, and Galerkin orthogonality, $a_\epsilon(u - U^h, v^h) = 0$, for all $v^h \in X^h$, it follows that

$$\begin{aligned}
\||u - U^h\||_{\epsilon,g_0} &\leq \sup_{0 \neq v \in X} \frac{a_\epsilon(u - U^h, v)}{\||v\||_{\epsilon,g_0}} \\
&= \sup_{0 \neq v \in X} \frac{a_\epsilon(u - U^h, v - v^h)}{\||v\||_{\epsilon,g_0}} \\
&= \sup_{0 \neq v \in X} \frac{(f, v - v^h) - a_\epsilon(U^h, v - v^h)}{\||v\||_{\epsilon,g_0}}.
\end{aligned}$$

The left inequality in (2.10) thus follows, as $v^h \in X^h$ is arbitrary. The right inequality in (2.10), follows the same approach, beginning with: for any $v \in X, v^h \in X^h$ $u = U^h$ satisfies

$$a_\epsilon(u - U^h, v - v^h) = (f, v - v^h) - a_\epsilon(U^h, v - v^h).$$

Applying (2.9) and the arbitrariness of v^h yields the right hand side of (2.10). $\square$

Remark: The difference between the norms of the error on the left and right hand side of (2.10) arise from the relatively large skew symmetric part of the bilinear form $a_\epsilon(\cdot, \cdot)$ through the different scalings of ϵ in the continuity and coercivity constants of $a_\epsilon(\cdot, \cdot)$.

For u satisfying (1.1), (1.2), we have in general that as $\epsilon \to 0$ that $\|u\|_1 \to \infty$. As we are interested in constructing error estimators valid for ϵ small it is thus more appropriate to bound the error in the weighted energy norm $\|| \cdot \||_{\epsilon,g_0}$. We therefore proceed by taking the first inequality in (2.10) and developing its right hand side into a sum of computable quantities.

Consider now the residual term in (2.10). For $v \in X, v^h \in X^h$, applying the divergence theorem on each triangle T gives

$$
\begin{aligned}
(f, v - v^h) &- a_\epsilon(U^h, v - v^h) \\
&= \sum_{T \in \Pi^h} \int_T [f - (-\epsilon \Delta U^h + \mathbf{v} \cdot \nabla U^h + g U^h)](v - v^h) dT \\
&\quad - \epsilon \sum_{e \in E^h} \int_e [\nabla U^h \cdot \mathbf{n}_e](v - v^h) de.
\end{aligned}
$$

Choosing $v^h = I^h(v) \in X^h$, applying (2.1) - (2.5), and using the Cauchy–Schwarz inequality gives

$$
\begin{aligned}
(f, v - v^h) - a_\epsilon(U^h, v - v^h) &\leq \sum_{T \in \Pi^h} C_2 \|r^h\|_T h_T \|\nabla^1 v\|_{N(T)} \\
&\quad + \epsilon \sum_{e \in E^h} C_5 h_T^{1/2} \|[\nabla U^h \cdot \mathbf{n}_e]\|_e \|\nabla^1 v\|_{N(e)} \\
&\leq \sqrt{2} \max\{C_2, C_5\} \Bigg[\sum_{T \in \Pi^h} h_T^2 \|r^h\|_T^2 \\
&\quad + \sum_{e \in E^h} h_T \epsilon^2 \|[\nabla U^h \cdot \mathbf{n}_e]\|_e^2 \Bigg]^{1/2} C_7 \|\nabla^1 v\|.
\end{aligned}
$$

Note that $\|\nabla v\|^2 \leq \epsilon^{-1} \|\|v\|\|_{\epsilon, g_0}^2$. Thus,

$$
\begin{aligned}
\frac{(f, v - v^h) - a_\epsilon(U^h, v - v^h)}{\|\|v\|\|_{\epsilon, g_0}} &\leq \sqrt{2} \max\{C_2, C_5\} C_7 \epsilon^{-1/2} * \\
&\quad \Bigg[\sum_{T \in \Pi^h} h_T^2 \|r^h\|_T^2 + \sum_{e \in E^h} h_T \epsilon^2 \|[\nabla U^h \cdot \mathbf{n}_e]\|_e^2 \Bigg]^{1/2}
\end{aligned}
$$

which gives the following estimate:

Theorem 2.2. *Assume $0 < g_0 \leq g(x) - 1/2 \nabla \cdot \mathbf{v}(x)$. Then*

$$
\|\|u - U^h\|\|_{\epsilon, g_0} \leq \sqrt{2} \epsilon^{-1/2} \max\{C_2, C_5\} C_7 \Bigg[\sum_{T \in \Pi^h} h_T^2 \|r^h\|_T^2 + \sum_{e \in E^h} h_T \epsilon^2 \|[\nabla U^h \cdot \mathbf{n}_e]\|_e^2 \Bigg]^{1/2}.
$$

$$\tag{2.11}$$

11

An adaptive procedure can be based upon the usual Galerkin finite element method and equidistribution of the bracketed terms on the R.H.S. of (2.11). Such a procedure is quite poor compared to one based upon a "better" discretization, because it leads to nearly uniform overrefinement as the indicators try to eliminate the nonphysical oscillations by refinement.

3. A Posteriori Error Analysis of Algorithms 1.1 and 1.2.

This section considers the defect correction discretization with and without the additional p-Laplacian term. We first give an a posteriori error analysis for Algorithms 1.1 and 1.2 in the energy norm, then (the main result of our work) in the L^2-norm. The $L^2(\Omega)$ analysis follows the pioneering path set forth in the work of Eriksson and Johnson (see, e.g., [9,10] and the references therein). For the convection dominated case, it seems to give essentially a better estimator (even accounting for the differences in the norm) than energy norm estimators.

Recall that ϵ_0 is a piecewise constant function (which is $O(h_T)$ for $\epsilon << h_T$) on each triangle given by

$$\epsilon_0(T) = \max\{|v|_{L^\infty(T)} \cdot \text{diam } (T), \epsilon\}. \qquad (3.1)$$

The dependence of $\epsilon_0(T)$ upon T and ϵ will be supressed in the manipulations that follow. For later reference we note that the update step in Algorithms 1.1 and 1.2 can be rewritten as, for all $v^h \in X_j^h$:

$$a_{\epsilon_0}(U^j, v^h) = (f, v^h) + ((\epsilon_0 - \epsilon)\nabla U^{j-1} - \mu_0 h|h\nabla U^{j-1}|^p \nabla U^{j-1}, \nabla v^h), \qquad (3.2)$$

and

$$a_\epsilon(U^j, v^h) = (f, v^h) - ((\epsilon_0 - \epsilon)\nabla(U^j - U^{j-1}) + \mu_0 h|h\nabla U^{j-1}|^p \nabla U^{j-1}, \nabla v^h). \qquad (3.3)$$

Energy Norm Estimates

There are two natural energy norms, $||| \cdot |||_{\epsilon,g_0}$ for the continuous problem and $||| \cdot |||_{\epsilon_0,g_0}$ for the discrete problem. Using $||| \cdot |||_{\epsilon,g_0}$ results in a fully computable estimator analogous

to the estimator for the centered Galerkin method (Section 2) which, for ϵ small, appears worse than the $L^2(\Omega)$ estimators we consider next. The $\||\cdot\||_{\epsilon,g_0}$ norm is the "natural" norm arising in the approach pioneered by Babuska, Rheinboldt and Banks (see [4,5,6, also 27] and the references therein). The norm bounded by the estimator is "improvable" by using the $\||\cdot\||_{\epsilon_0,g_0}$ norm. However, the result is not fully computable without a heuristic step of replacing $\|\nabla u\|_{L^2(T)}$ by $\|\nabla U^j\|_{L^2(T)}$ on the RHS (compare (3.13) and (3.14)). Full reliability is therefore sacrificed. The derivation of the estimates for Algorithm 1.1 (i.e. $\mu_0 = 0$) begins with the identity

$$a_\epsilon(u - U^j, v) = (f, v) - a_\epsilon(U^j, v) \quad \text{for all} \quad v \in X. \tag{3.4}$$

From (3.3), $(f, v^h) - a_\epsilon(U^j, v^h) = ((\epsilon_0 - \epsilon)\nabla(U^j - U^{j-1}), \nabla v^h)$ for all $v^h \in X_j^h$. Thus, (3.4) can be written as

$$a_\epsilon(u - U^j, v) = (f, v - v^h) - a_\epsilon(U^j, v - v^h) + ((\epsilon_0 - \epsilon)\nabla(U^j - U^{j-1}), \nabla v^h),$$

for all $v \in X, v^h \in X_j^h$.

Adding $((\epsilon_0 - \epsilon)\nabla(u - U^j), \nabla v)$ to both sides of the above gives

$$a_{\epsilon_0}(u - U^j, v) = (f, v - v^h) - a_\epsilon(U^j, v - v^h) + ((\epsilon_0 - \epsilon)\nabla(U^j - U^{j-1}), \nabla v)$$
$$- ((\epsilon_0 - \epsilon)\nabla(U^j - U^{j-1}), \nabla(v - v^h)) + ((\epsilon_0 - \epsilon)\nabla(u - U^j), \nabla v)$$

or, for all $v \in X, v^h \in X_j^h$,

$$a_{\epsilon_0}(u - U^j, v) = (f, v - v^h) - a_\epsilon(U^j, v - v^h) + ((\epsilon_0 - \epsilon)\nabla(u - U^{j-1}), \nabla v)$$
$$- ((\epsilon_0 - \epsilon)\nabla(U^j - U^{j-1}), \nabla(v - v^h)). \tag{3.5}$$

Consider the first two terms on the right hand side of (3.5). Applying the divergence theorem upon each triangle T gives

$$(f, v - v^h) - a_\epsilon(U^j, v - v^h) = \sum_{T \in \Pi_j^h} \int_T (f - (-\epsilon\Delta U^j + \mathbf{v} \cdot \nabla U^j + gU^j))(v - v^h)dx$$
$$- \epsilon \sum_{\epsilon \in E_j^h} \int_\epsilon [\nabla U^j \cdot \mathbf{n}_\epsilon]\nabla(v - v^h)ds. \tag{3.6}$$

13

Choosing $v^h = I^h(v)$, and letting $r^j := f - (-\epsilon \Delta U^j + v \cdot \nabla U^j + g U^j)$ denote the local residue for U^j, we obtain

$$\sum_T \int_T r^j(v - v^h) dT \leq \sum_T \|r^j\|_T \|v - v^h\|_T$$

$$\leq C_2 \sum_T \|r^j\|_T h_T \|\nabla v\|_{N(T)}$$

$$\leq C_2 \sum_T \epsilon_0^{-1/2} \|r^j\|_T h_T \epsilon_0^{1/2} \|\nabla v\|_{N(T)}$$

$$\leq C_2 \left(\sum_T \epsilon_0^{-1} h_T^2 \|r^j\|_T^2 \right)^{1/2} \left(\sum_T \epsilon_0 \|\nabla v\|_{N(T)}^2 \right)^{1/2}$$

$$\leq C_2 \left(\sum_T \epsilon_0^{-1} h_T^2 \|r^j\|_T^2 \right)^{1/2} \left(\sum_T \epsilon_0 \|\nabla v\|_{N(T)}^2 + g_0 \|v\|_{N(T)}^2 \right)^{1/2}$$

$$\leq C_2 C_7 \left(\sum_T \epsilon_0^{-1} h_T^2 \|r^j\|_T^2 \right)^{1/2} \|\|v\|\|_{\epsilon_0, g_0}. \tag{3.7}$$

The jump integral terms in (3.6) may be bounded similarly as

$$\left| \epsilon \sum_{e \in E_j^h} \int_e [\nabla U^j \cdot \mathbf{n}_e](v - v^h) ds \right|$$

$$\leq C_5 C_7 \left(\sum_{e \in E_j^h} \epsilon^2 \epsilon_0^{-1} h_T \|[\nabla U^j \cdot \mathbf{n}_e]\|_e^2 \right)^{1/2} \cdot \|\|v\|\|_{\epsilon_0, g_0}. \tag{3.8}$$

Consider now the fourth term on the right hand side of (3.5).

$$((\epsilon_0 - \epsilon)\nabla(U^j - U^{j-1}), \nabla(v - v^h))$$

$$= \sum_T ((\epsilon_0 - \epsilon)\nabla(U^j - U^{j-1}), \nabla(v - v^h))_T$$

$$\leq \sum_T (\epsilon_0 - \epsilon)\epsilon_0^{-1/2} \|\nabla(U^j - U^{j-1})\|_T \epsilon_0^{1/2} \|\nabla(v - v^h)\|_T$$

$$\leq \sum_T (\epsilon_0 - \epsilon)\epsilon_0^{-1/2} \|\nabla(U^j - U^{j-1})\|_T \epsilon_0^{1/2} C_3 \|\nabla v\|_{N(T)}$$

$$\leq C_3 C_7 \left(\sum_T (\epsilon_0 - \epsilon)^2 \epsilon_0^{-1} \|\nabla(U^j - U^{j-1})\|_T^2 \right)^{1/2} \cdot \|\|v\|\|_{\epsilon_0, g_0}. \tag{3.9}$$

14

Analogously, the third term on the right hand side of (3.5) satisfies

$$
((\epsilon_0 - \epsilon)\nabla(u - U^{j-1}), \nabla v) \leq \sum_T (\epsilon_0 - \epsilon)\epsilon_0^{-1/2}\|\nabla(u - U^{j-1})\|_T \epsilon_0^{1/2}\|\nabla v\|_T
$$

$$
\leq \left(\sum_T (\epsilon_0 - \epsilon)^2 \epsilon_0^{-1}\|\nabla(u - U^{j-1})\|_T^2\right)^{1/2} \cdot \|\|v\|\|_{\epsilon_0, g_0}.
$$

$$(3.10)$$

Using (3.6) - (3.10) in (3.5) gives for any $v \in X, v \neq 0$,

$$
\frac{a_{\epsilon_0}(u - U^j, v)}{\|\|v\|\|_{\epsilon_0, g_0}} \leq C_2 C_7 \left(\sum_T \epsilon_0^{-1} h_T^2 \|r^j\|_T^2\right)^{1/2}
$$

$$
+ C_5 C_7 \left(\sum_{e \in E_j^h} \epsilon^2 \epsilon_0^{-1} h_T \|[\nabla U^j \cdot \mathbf{n}_e]\|_e^2\right)^{1/2}
$$

$$
+ C_3 C_7 \left(\sum_T (\epsilon_0 - \epsilon)^2 \epsilon_0^{-1}\|\nabla(U^j - U^{j-1})\|_T^2\right)^{1/2}
$$

$$
+ \left(\sum_T (\epsilon_0 - \epsilon)^2 \epsilon_0^{-1}\|\nabla(u - U^{j-1})\|_T^2\right)^{1/2}.
$$

Define $\bar{\epsilon}_0 = \max_T \epsilon_0(T)$. Using the previous inequality and (2.7) in Lemma 2.1 (with ϵ replaced by ϵ_0 on both sides of (2.7)), and $(\epsilon_0 - \epsilon)^2 \epsilon_0^{-1} \leq [(\bar{\epsilon}_0 - \epsilon)/\bar{\epsilon}_0]^2 \epsilon_0$ gives

$$
\|\|u - U^j\|\|_{\epsilon_0, g_0}^2 \leq c_j \sum_T \epsilon_0^{-1} h_T^2 \|r^j\|_T^2 +
$$

$$
+ c_{jj} \sum_{e \in E_j^h} \epsilon^2 \epsilon_0^{-1} h_T \|[\nabla U^j \cdot \mathbf{n}_e]\|_e^2 +
$$

$$
+ c_{jjj} \sum_T (\epsilon_0 - \epsilon)^2 \epsilon_0^{-1}\|\nabla(U^j - U^{j-1})\|_T^2 +
$$

$$
+ \left(\frac{\bar{\epsilon}_0 - \epsilon}{\bar{\epsilon}_0}\right)^2 \|\|u - U^{j-1}\|\|_{\epsilon_0, g_0}^2.
$$

Thus,

$$\||u - U^j\||^2_{\epsilon_0,g_0} \leq \sum_{\ell=1}^{j}\left(\frac{\epsilon_0 - \epsilon}{\epsilon_0}\right)^{2(j-\ell)}\left[\sum_{T}(c_l\epsilon_0^{-1}h_T^2\|r^\ell\|_T^2 + \right.$$

$$+ c_{\ell\ell\ell}(\epsilon_0 - \epsilon)^2\epsilon_0^{-1}\|\nabla(U^\ell - U^{\ell-1})\|_T \qquad (3.11)$$

$$\left. + c_{\ell\ell}\sum_{e\in E_j^h}\epsilon^2\epsilon_0^{-1}h_T\|[\nabla U^\ell \cdot \mathbf{n}_e]\|_e^2] + \left(\frac{\bar{\epsilon}_0 - \epsilon}{\bar{\epsilon}_0}\right)^{2(j-1)}\||u - U^1\||^2_{\epsilon_0,g_0}.$$

This inequality (3.11), together with an estimate for $\||u - U^1\||_{\epsilon_0,g_0}$, will give a complete error estimate for DCM.

Consider the error in the first step in DCM. U^1 satisfies

$$a_{\epsilon_0}(U^1, v^h) = (f, v^h), \text{ for all } v^h \in X^h.$$

This can be rewritten as : for all $v^h \in X^h$

$$a_\epsilon(U^1, v^h) = (f, v^h) - ((\epsilon_0 - \epsilon)\nabla U^1, \nabla v^h).$$

Consider, now for all $v \in X$ and $v^h \in X^h$

$$a_\epsilon(u - U^1, v) = (f, v) - a_\epsilon(U^1, v)$$

$$= (f, v - v^h) - a_\epsilon(U^1, v - v^h) + ((\epsilon_0 - \epsilon)\nabla U^1, \nabla v^h). \qquad (3.12)$$

Repeating the operations of (3.6) - (3.8) gives :

$$a_\epsilon(u - U^1, v) \leq C_2 C_7 \left(\sum_{T}\epsilon^{-1}h_T^2\|r^1\|_T^2\right)^{1/2} + C_5 C_7 \left(\sum_{e\in E}\epsilon h_T\|[\nabla U^1 \cdot \mathbf{n}_e]\|_e^2\right)^{1/2}$$

$$+ C_1\left(\sum_{T}(\epsilon_0 - \epsilon)^2\epsilon^{-1}\|\nabla U^1\|_T^2\right)^{1/2} \cdot \|v\|_{\epsilon,g_0}.$$

The gives immediately the following error estimator for $u - U^1$:

$$\|u - U^1\|_{\epsilon,g_0} \leq C_2 C_7 \left(\sum_{T}\epsilon^{-1}h_T^2\|r^1\|_T^2\right)^{1/2} + C_5 C_7 \left(\sum_{e\in E}\epsilon h_T\|[\nabla U^1 \cdot \mathbf{n}_e]\|_e^2\right)^{1/2}$$

$$+ C_1\left(\sum_{T}(\epsilon_0 - \epsilon)^2\epsilon^{-1}\|\nabla U^1\|_T^2\right)^{1/2}. \qquad (3.13)$$

We note for future reference, that by taking a slightly different path the following estimate can also be derived :

$$\|u - U^1\|_{\epsilon_0,g_0} \leq C_2 C_7 \left(\sum_T \epsilon_0^{-1} h_T^2 \|r^1\|_T^2 \right)^{1/2} + C_5 C_7 \left(\sum_{e \in E} \epsilon^2 \epsilon_0^{-1} h_T \|[\nabla U^1 \cdot \mathbf{n}_e]\|_e^2 \right)^{1/2}$$
$$+ C_1 \left(\sum_T (\epsilon_0 - \epsilon)^2 \epsilon_0^{-1} \|\nabla u\|_T^2 \right)^{1/2}. \quad (3.14)$$

These two estimates for $(u - U^1)$, (3.13), (3.14), differ in that (3.14) uses a "better" norm and has "better" constants but has one term (the $\|\nabla u\|_T^2$ term) which is not computable. Using (3.14) would require a heuristic substitution of, e.g., $\|\nabla U^j\|_T^2$ for $\|\nabla u\|_T^2$ in the R.H.S. of (3.14). If the estimate (3.13) or (3.14) is inserted into (3.11) we obtain an estimator for $\||u - U^j\||_{\epsilon_0,g_0}$. This estimator obtained is unattractive for practical computations due to the sum of residuals, jumps and updates calculated for each update. We could use instead $\||\cdot\||_{\epsilon,g_0}$ in the analysis. However, since ϵ is typically very small $\||\cdot\||_{\epsilon,g_0}$ should be thought of as very close to the $L^2(\Omega)$ norm. We therefore proceed directly to estimating $\|u - U^j\|$.

L_2 Norm Estimates.

In this section we present the main theoretical result of our report: L_2 norm estimates for $u - U^j$ for Algorithms 1.1 and 1.2. We estimate $\|u - U^j\|$ by duality techniques following the approach of Eriksson and Johnson [10].

Theorem 3.1. *For u satisfying (1.1), (1.2), with $0 < g_0 \leq (g(x) - 1/2\nabla \cdot \mathbf{v}(x))$ and Γ either a convex polygon or smooth, and U^j given by Algorithm 1.1, we have*

$$\|u - U^j\| \leq C_8 \epsilon^{-1/2} \left(\sum_{T \in \Pi_j^h} h_T^2 \|r^j\|_T^2 \right)^{1/2}$$
$$+ C_{12} \epsilon^{-1/2} \left(\sum_{e \in E_j^h} \int_e h_e^3 [\nabla U^j \cdot \mathbf{n}_e]^2 de \right)^{1/2}$$
$$+ C_{14} \epsilon^{-1/2} \left(\sum_{T \in \Pi_j^h} (\epsilon_0 - \epsilon)^2 \|\nabla(U^j - U^{j-1})\|_T^2 \right)^{1/2}, \quad (3.15)$$

where C_8, C_{12}, and C_{14}, are computable constants dependent upon the given data and the partition of the domain.

The proof of this theorem uses the following regularity result of Navert [23], see also [2].

Lemma 3.1. *Assume that $(g - 1/2\nabla \cdot v) \geq g_0 > 0$. If Γ is either a convex polygon or smooth, then the solution u of (1.1), (1.2), satisfies*

$$\epsilon^{3/2}\|u\|_2 + \epsilon^{1/2}\|u\|_1 + \|u\| \leq C_6\|f\|, \tag{3.16}$$

where C_6 is independent of u and ϵ. $\square$

Proof of Theorem 3.1: We begin the a posteriori error analysis by introducing the associated adjoint problem of (1.1), (1.2). Since $\mu_0 = 0$

$$a_{\epsilon_0}(U^j, v^h) = (f, v^h) - ((\epsilon_0 - \epsilon)\nabla(U^j - U^{j-1}), \nabla v^h). \tag{3.17}$$

For $\theta := u - U^j$, define z as the solution to the associated adjoint problem of (1.1), (1.2) with right hand side θ. Specifically,

$$\mathcal{L}_\epsilon^* z := -\epsilon\Delta z - \nabla \cdot (zv) + gz = \theta \quad \text{in } \Omega, z = 0, \text{ on } \Gamma. \tag{3.18}$$

Then,

$$\|\theta\|^2 = (\theta, \theta) = (\theta, \mathcal{L}_\epsilon^* z) = a_\epsilon(\theta, z). \tag{3.19}$$

With $v^h = \tilde{z} = I^h(z)$, and $\mu_0 = 0$, using (3.3) in (3.19) we obtain:

$$\begin{aligned}
\|\theta\|^2 &= a_\epsilon(\theta, z) - a_\epsilon(u, \tilde{z}) + a_{\epsilon_0}(U^j, \tilde{z}) + ((\epsilon_0 - \epsilon)\nabla(U^j - U^{j-1}), \nabla\tilde{z}) \\
&= a_\epsilon(\theta, z - \tilde{z}) + ((\epsilon_0 - \epsilon)\nabla(U^j - U^{j-1}), \nabla\tilde{z}) \\
&= (f, z - \tilde{z}) - (gU^j + v \cdot \nabla U^j, z - \tilde{z}) - (\epsilon\nabla U^j, \nabla(z - \tilde{z})) \\
&\quad + ((\epsilon_0 - \epsilon)\nabla(U^j - U^{j-1}), \nabla\tilde{z}) \\
&= \sum_{T \in \Pi_j^h} (r^j, z - \tilde{z})_T - \sum_{e \in E_j^h} \int_e \epsilon[\nabla U^j \cdot n_e](z - \tilde{z})ds \\
&\quad + ((\epsilon_0 - \epsilon)\nabla(U^j - U^{j-1}), \nabla\tilde{z}) \\
&= I + II + III.
\end{aligned} \tag{3.20}$$

Using the approximaton properties of the Clément interpolant and (3.16) we can bound expression I by :

$$|I| = |\sum_{T \in \Pi_j^h} (r^j, z - \tilde{z})_T|$$

$$\leq \sum_{T \in \Pi_j^h} h_T \|r^j\|_T \|h_T^{-1}(z - \tilde{z})\|_T$$

$$\leq \sum_{T \in \Pi_j^h} C_2 \epsilon^{-1/2} h_T \|r^j\|_T \|\epsilon^{1/2} \nabla z\|_{N(T)}$$

$$\leq C_2 \epsilon^{-1/2} \left(\sum_{T \in \Pi_j^h} h_T^2 \|r^j\|_T^2 \right)^{1/2} \left(\sum_{T \in \Pi_j^h} \|\epsilon^{1/2} \nabla z\|_{N(T)}^2 \right)^{1/2}$$

$$\leq C_2 \epsilon^{-1/2} C_7 \left(\sum_{T \in \Pi_j^h} h_T^2 \|r^j\|_T^2 \right)^{1/2} \|\epsilon^{1/2} \nabla z\|$$

$$\leq C_2 \epsilon^{-1/2} C_7 C_6 \left(\sum_{T \in \Pi_j^h} h_T^2 \|r^j\|_T^2 \right)^{1/2} \|\theta\|.$$

Thus,

$$|I| \leq C_8 \epsilon^{-1/2} \left(\sum_{T \in \Pi_j^h} h_T^2 \|r^j\|_T^2 \right)^{1/2}, \tag{3.21}$$

where $C_8 = C_2 C_6 C_7$. Consider now II.

$$|II| = |\sum_{e \in E_j^h} \int_e \epsilon^{-1/2} h_e^{3/2} [\nabla U^j \cdot \mathbf{n}_e] \epsilon^{3/2} h_e^{-3/2} (z - \tilde{z}) ds|$$

$$\leq \epsilon^{-1/2} \sum_{e \in E_j^h} \left(\int_e h_e^3 [\nabla U^j \cdot \mathbf{n}_e]^2 ds \right)^{1/2} \left(\int_e \epsilon^3 h_e^{-3} (z - \tilde{z})^2 ds \right)^{1/2}. \tag{3.22}$$

To bound the second term in (3.22) we use (2.5). Hence,

$$|II| \leq \epsilon^{-1/2} \left(\sum_{e \in E_j^h} \int_e h_e^3 [\nabla U^j \cdot \mathbf{n}_e]^2 ds \right)^{1/2} \left(\sum_{e \in E_j^h} C_5^2 \|\epsilon^{3/2} \nabla^2 z\|_{N(e)}^2 \right)^{1/2}$$

$$\leq \epsilon^{-1/2} C_5 C_7 \left(\sum_{e \in E_j^h} \int_e h_e^3 [\nabla U^j \cdot \mathbf{n}_e]^2 ds \right)^{1/2} \|\epsilon^{3/2} \nabla^2 z\|$$

$$\leq C_{12} \epsilon^{-1/2} \left(\sum_{e \in E_j^h} \int_e h_e^3 [\nabla U^j \cdot \mathbf{n}_e]^2 ds \right)^{1/2} \|\theta\|, \tag{3.23}$$

where $C_{12} = C_7 C_5$.

The remaining term to be bounded in (3.20) is III.

$$|III| = |((\epsilon_0 - \epsilon)\nabla(U^j - U^{j-1}), \nabla \tilde{z})|$$

$$\leq \sum_{T \in \Pi_j^h} \|(\epsilon_0 - \epsilon)\epsilon^{-1/2}\nabla(U^j - U^{j-1})\|_T \|\epsilon^{1/2}\nabla \tilde{z}\|_T$$

$$\leq \sum_{T \in \Pi_j^h} \|(\epsilon_0 - \epsilon)\epsilon^{-1/2}\nabla(U^j - U^{j-1})\|_T C_{13} \|\epsilon^{1/2}\nabla z\|_{N(T)},$$

where C_{13} satisfies $\|\nabla \tilde{z}\|_T \leq C_{13} \|\nabla z\|_{N(T)}$,

$$\leq C_{13} \epsilon^{-1/2} \left(\sum_{T \in \Pi_j^h} (\epsilon_0 - \epsilon)^2 \|\nabla(U^j - U^{j-1})\|_T^2 \right)^{1/2} \left(\sum_{T \in \Pi_j^h} \|\epsilon^{1/2}\nabla z\|_{N(T)}^2 \right)^{1/2}$$

$$\leq C_{13} \epsilon^{-1/2} C_7 \left(\sum_{T \in \Pi_j^h} (\epsilon_0 - \epsilon)^2 \|\nabla(U^j - U^{j-1})\|_T^2 \right)^{1/2} \|\epsilon^{1/2}\nabla z\|$$

$$\leq C_{14} \epsilon^{-1/2} \left(\sum_{T \in \Pi_j^h} (\epsilon_0 - \epsilon)^2 \|\nabla(U^j - U^{j-1})\|_T^2 \right)^{1/2} \|\theta\|, \tag{3.24}$$

where $C_{14} = C_6 C_7 C_{13}$.

Combining (3.21), (3.23), (3.24), with (3.20), we obtain the claimed a posteriori error estimate for $u - U^j$.

Remarks

1. As noted in [10], in the presence of boundary layers, the *unweighted* L_2 norm of the residual will, in general, not converge to zero under local refinement of the partition.

Hence, it cannot be used as an effective error estimator for problems having boundary layers. However, when weighted with the local mesh parameter the product does then converge to zero under local refinement of the mesh.

2. With an analogous argument, replacing $h_e^{3/2}$ by $h_e^{1/2}$, one can derive for II the bound

$$|II| \le \epsilon^{1/2} C_{12} \left(\sum_{e \in E_j^h} \int_e h_e [\nabla U^j \cdot \mathbf{n}_e]^2 ds \right)^{1/2} \|\theta\|. \tag{3.25}$$

3. Introducing the local grid parameter, h_T, instead of ϵ, and using an inverse estimate for $\nabla \tilde{z}$, $\|\nabla \tilde{z}\|_T \le C_{15} h_T^{-1} \|\tilde{z}\|_T$, an alternate estimate for III can be derived

$$|III| \le C_{16} \left(\sum_{T \in \Pi_j^h} (\epsilon_0 - \epsilon)^2 h_T^{-2} \|\nabla(U^j - U^{j-1})\|_T^2 \right)^{1/2} \|\theta\|, \tag{3.26}$$

where $C_{16} = C_1 C_6 C_7 C_{15}$.

Using remarks 2 and 3 appropriately in the last proof we obtain the following.

Corollary 3.1. *Under the conditions of Theorem 3.1, the following error bound holds:*

$$\|u - U^j\| \le C_8 \left(\sum_{T \in \Pi_j^h} \epsilon^{-1} h_T^2 \|r^j\|_T^2 \right)^{1/2}$$

$$+ 2C_{12} \left(\sum_{e \in E_j^h} \min\{h_e^2/\epsilon, \epsilon\} \int_e h_e [\nabla U^j \cdot \hat{n}_e]_e^2 de \right)^{1/2}$$

$$+ (C_{14} + C_{16}) \left(\sum_{T \in \Pi_j^h} \min\{(\epsilon_0 - \epsilon)^2/\epsilon, (\epsilon_0 - \epsilon)^2/h_T^2\} \|\nabla(U^j - U^{j-1})\|_T^2 \right)^{1/2}. \tag{3.27}$$

3.3 Incorporation of the p-Laplacian.

If a p-Laplacian is used in the algorithm to limit the amount of anti-diffusion performed in transition regions the analysis leading to (3.27) must be modified. The required modification is simply to replace, at each step, the update term $\|(\epsilon_0 - \epsilon)\nabla(U^j - U^{j-1})\|_T^2$ by a

modified update term reading

$$\|(\epsilon_0 - \epsilon)\nabla(U^j - U^{j-1}) + \mu_0 h|h\nabla U^{j-1}|^p\nabla U^{j-1}\|_T^2 =$$

$$\|(\epsilon_0 - [\epsilon + \mu_0 h_T|h_T\nabla U^{j+1}|^p])\nabla U^{j-1}\|_T^2.$$

Repeating the steps leading to Corollary 3.1 with this modification gives the following estimator.

Corollary 3.2. Suppose the hypotheses of Theorem 3.1 hold. Then, for u satisfying (1.1), (1.2) and U^j given by Algorithm 1.2 the error $\|u - U^j\|$ satisfies:

$$\|u - U^j\| \le C_8 \left(\sum_{T \in \Pi_j^h} h_T^2 \epsilon^{-1} \|r^j\|_T^2 \right)^{1/2}$$

$$+ 2C_{12} \left(\sum_{e \in E_j^h} \min\{h_e^2/\epsilon, \epsilon\} \int_e h_e[\nabla U^j \cdot \hat{n}_e]_e^2 de \right)^{1/2}$$

$$+ (C_{14} + C_{16}) \left(\sum_{T \in \Pi_j^h} \min\{\epsilon^{-1}, h^{-2}(T)\}\|(\epsilon_{0(T)} - \epsilon)\nabla(U^j - U^{j-1})\right.$$

$$\left. + \mu_0 h_T|h_T\nabla U^{j-1}|^p\nabla U^{j-1}\|_T^2 \right)^{1/2}.$$

4. Numerical Illustrations.

Two numerical illustrations of Algorithms 1.1 and 1.2 are presented. The first illustrates the reliability of the error estimators. Next, we present an additional example for which an exact solution is not known. It is traditional to exhibit the grids generated by adaptive procedures (and we shall do so shortly). Nevertheless it has been noted in many places that the final adaptively generated grid can change dramatically with small changes in the input parameters. Thus grids should be judged with regard to their general "reasonableness" rather than their exact configuration.

The two examples we consider here are, (1) a "skew step" type problem with a known exact solution and, (2) the "rotating pulse" problem of [19] (modelling some features of

internal flow problems) with a circular plateau-like solution and $O(\sqrt{\epsilon})$ internal transition region. In each case we use continuous, piecewise linear elements.

In these illustrations, we take the coefficients in the estimator (1.3,b) to be equal:

$$\tilde{C}_J = \tilde{C}_R = \tilde{C}_U = 1/15.$$

A better estimate of these coefficients can certainly be obtained. Since we are illustrating the estimators utility rather than optimizing the overall algorithm we do not pursue this point.

In all our tests herein, we take $\epsilon_0 = 2h(x)$, which is almost certainly overly diffuse and $\mu_0 = 0$ or $\mu_0 = 4$. Only one antidiffusion/corrector step was performed. No p-Laplacian was needed in Example 4.2 since only one corrector steps were performed.

In all the tests presented herein, a simple ILU_ϕ preconditioned CGS (conjugate gradient squared) solver was used. The number of CGS iterations was observed for each approximate solution U^j to be always quite small due to the good numerical stability of the linear system arising from the artificial viscosity operator $a_{\epsilon_0}(\cdot, \cdot)$.

Example 4.1: Let $\Omega = (0,1)^2, \epsilon = 10^{-4}$ $\mathbf{v} = (1, 1/4)^{tr}, g = 2$ and $h = \frac{1}{16}, \frac{1}{32}, \frac{1}{64}, \frac{1}{128}$ and $\frac{1}{250}$ (uniform refinement). We take

$$u_{\text{TRUE}}(x, y) = \arctan[\alpha(x/4 - y + 1/2]$$

with $\alpha = 100$, substitute u_{TRUE} into the differential equation and calculate a RHS $f(x, y, \epsilon)$ and take nonhomogeneous boundary conditions $u = u_{\text{TRUE}}$ on $\partial\Omega$. In the table below we present the ratio of the estimated and true errors for $\mu_0 = 0$ (Algorithm 1.1) and $\mu_0 = 4$ (Algorithm 1.2).

h	Algorithm 1.1 Estimated Error TRUE ERROR	Algorithm 1.2 Estimated Error TRUE ERROR
$\frac{1}{16}$	2.83	3.47
$\frac{1}{32}$	2.11	2.96
$\frac{1}{64}$	1.53	2.45
$\frac{1}{128}$	1.37	2.19
$\frac{1}{256}$	1.05	1.25

Table 4.1

In Table 4.1 we see that the estimators become closer to the true error (Est/Error → 1) as the mesh is refined. This is entirely as expected. The ratio approaches 1 more slowly in Algorithm 1.2 since the combined effects of a too diffuse stable operator (see the conclusions section) with the p-Laplacian slows resolution of the internal layer.

Example 4.2: The "Rotating Pulse" In (1.1) we take $\epsilon = 10^{-4}, \Omega = (-1,1)^2$ and for $r = \sqrt{x^2 + y^2}$ and $c = 2, f(r) \equiv 1$ if $r \le \frac{1}{2}$ with $f \equiv 0$ for $r > \frac{1}{2}$. The velocity field $\mathbf{v}$ is given by

$$\mathbf{v} = [-y(1-r^2), x(1-r^2)]^{tr} \text{ for } 0 \le r \le 1 \text{ and } \mathbf{v} = 0 \text{ otherwise.}$$

At the boundary, $u(x,y) \equiv 0$.

In Example 4.2's experiments we take $\mu_0 = 0$, Tol $= 0.02$ and, on each successive grid, perform one artificial viscosity step followed by one defect correction update.

Table 4.2 presents the number of triangles and estimated error on the seven meshes. Note the de-refinement occurring between Steps 6 and 7.

Mesh Number	No. of Triangles	Estimated Error
1	512	0.19
2	3136	0.19
3	9568	0.087
4	21504	0.045
5	39552	0.035
6	65996	0.023
7	53240	0.027

Table 4.2: Estimated errors for the rotating pulse problem (Example 4.2).

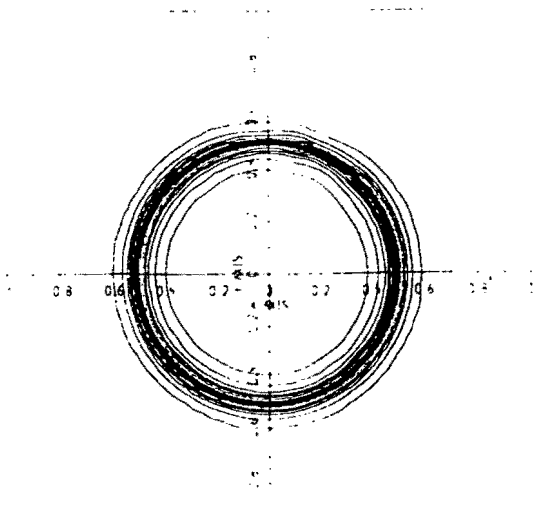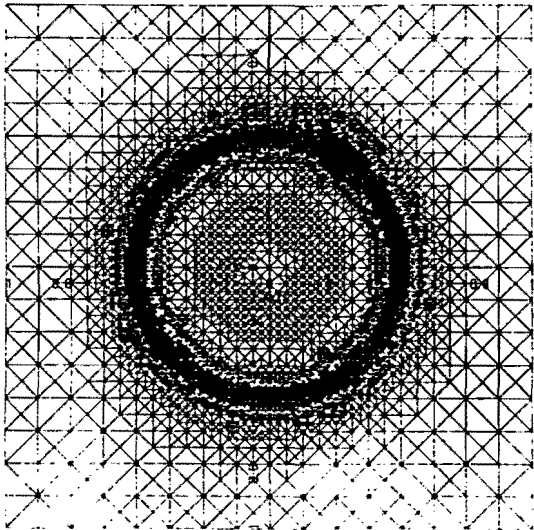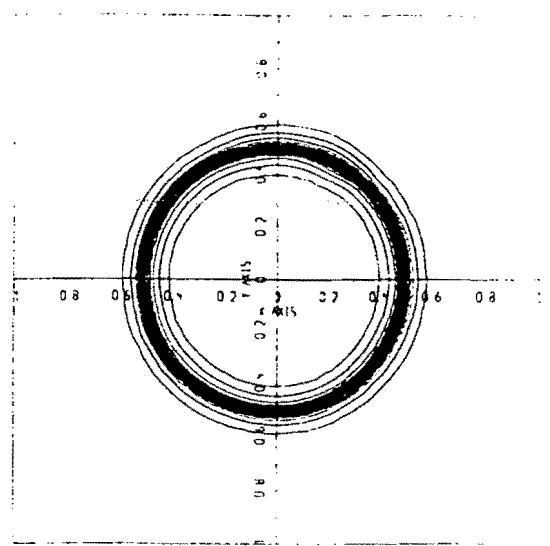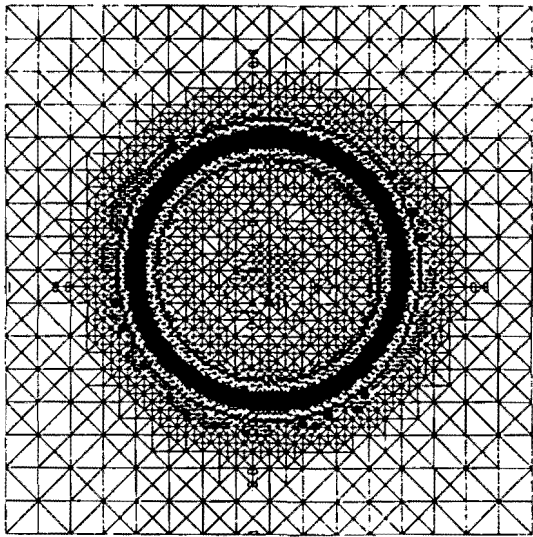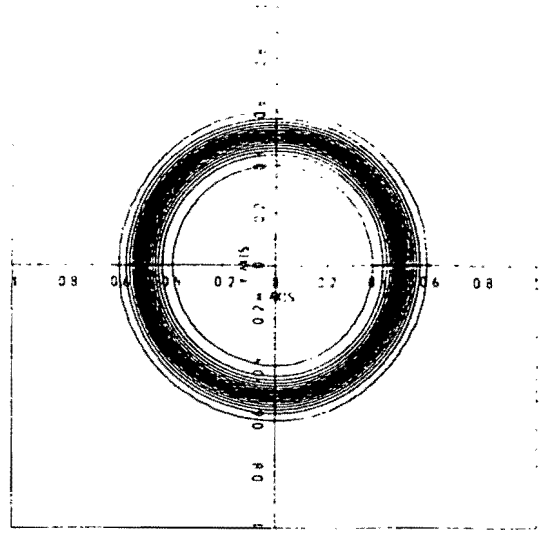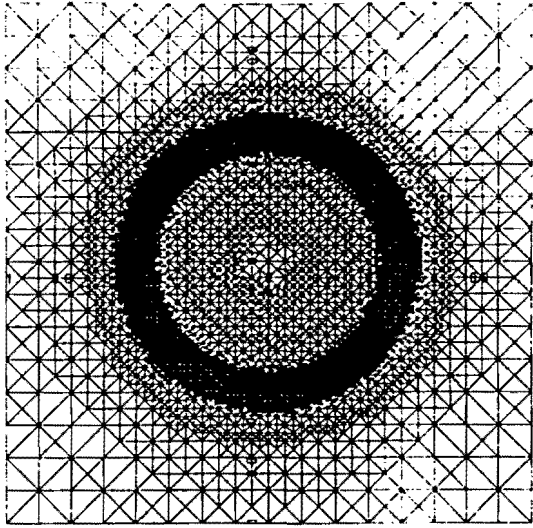In Figure 4.1 we present the seven meshes and approximate solutions.

26a
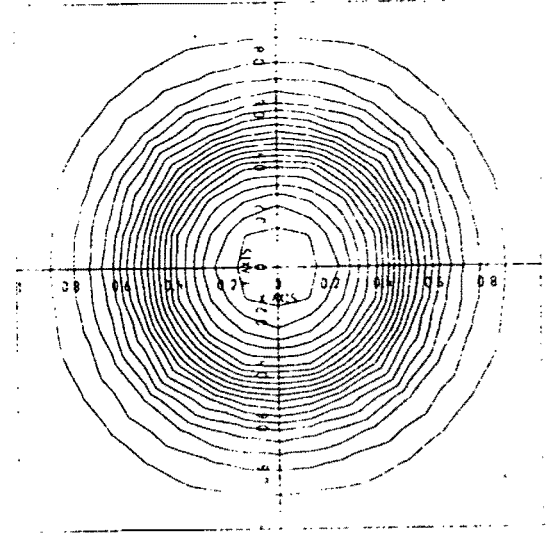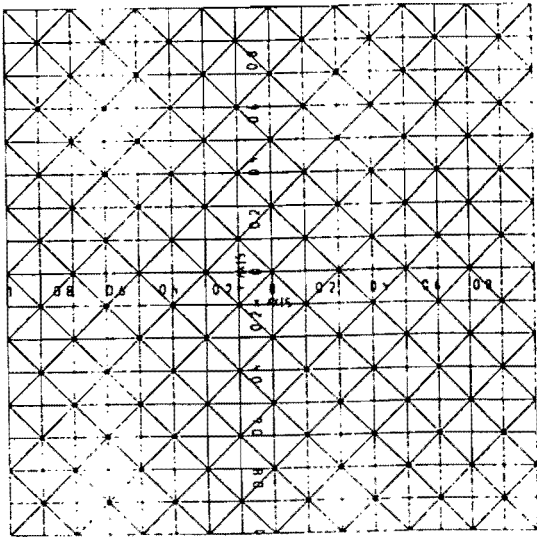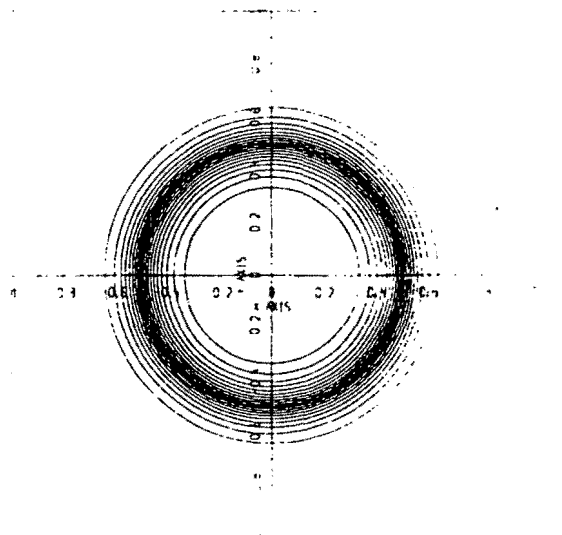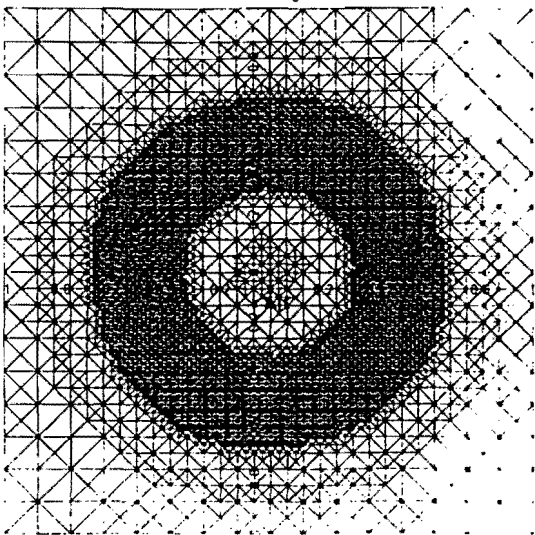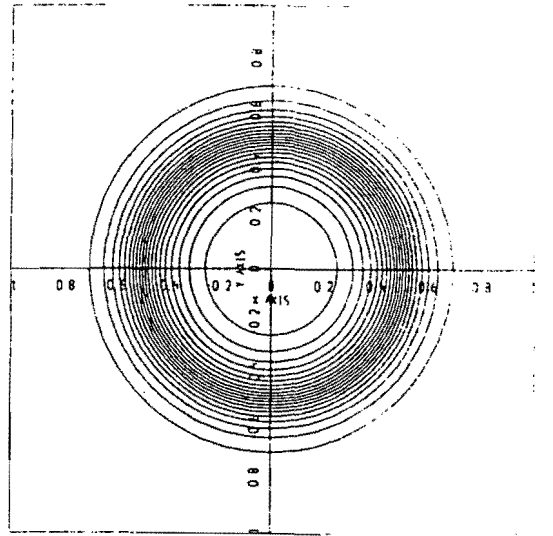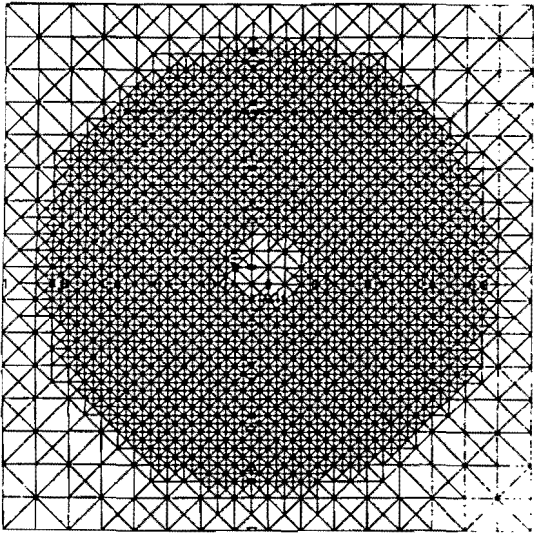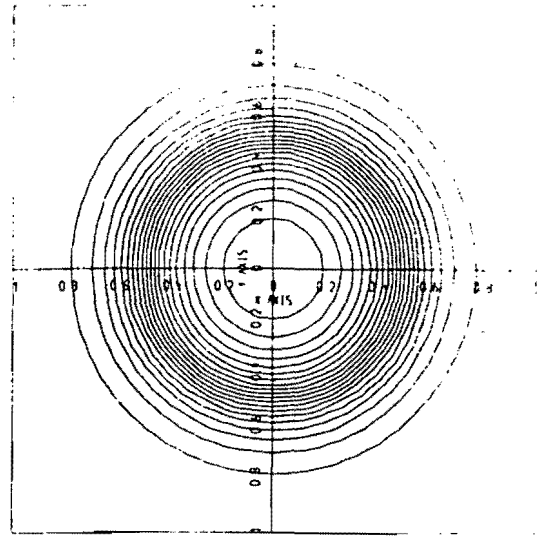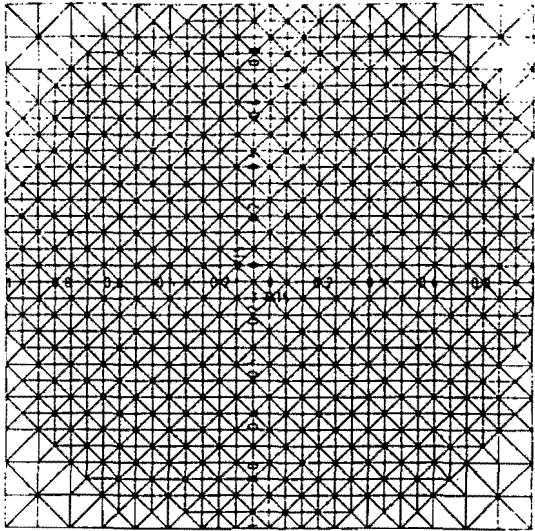
FIGURE 4.1: ADAPTIVE APPROXIMATIONS TO EXAMPLE 4.2.

26b

Mesh Refinement and De-refinement.

After the simplices are marked for refinement, de-refinement or no action via Algorithm 1.3, the particular mesh generation algorithm chosen is then applied. We have used a newest vertex bisection algorithm which produces simplical grids in any dimension with a small finite number of equivalence classes of elements. This method is developed and validated in detail in Maubach [20,21] (building upon the early work of Sewell [28]). Since this study focuses primarily upon Algorithm 1.2 and its error estimation we have at each step opted for simplicity, universality and reliability over economy. For example, a triangle being refined is simply cut in half twice (reducing its diameter by $\frac{1}{2}$) rather than using differing levels of refinement at each update according to the relative sizes of the individual indicators. De-refinement is also repeated twice at each step, when possible.

Due to the recursive and more global nature of de-refinement there are several options that must be selected in the de-refinement algorithm. Since our study has emphasized reliability over economy we have also, at each step, given refinement precedent over de-refinement. It is possible to include numerous "heuristics" into the mesh generation procedure with the aim of generating a better mesh at an earlier grid number. We have not done so in our tests.

The rotating pulse problem illustrates the fact that the adaptive defect correction algorithm can provide an accurate and high quality approximate solution. It further shows that the attractive features of the approach are independent of requiring a streamwise node ordering, or requiring a convection field without closed loops.

5. Conclusions.

To improve the performance of the defect correction method in adaptive computations our experience suggests the following refinements.

1) The amount of artificial viscosity used in the stable operator should be as small as possible consistent with stability and the error localization results of [2, 11]. (This observation has been borne out in a study of the authors of adaptive defect correction methods for the Navier-Stokes equation).

2) With less artifical viscosity in the stable operator or more correction steps, use of a p-Laplacian in the residual can control overshoots and undershoots, [29, 30]. The present paper shows that its use presents no problem for the mathematical basis of the error estimators.

3) For $\epsilon << O(h), L^2$ error estimators seem to be preferable to energy norm error estimators. Estimators using heuristics are possibly more *efficient* than the estimators we have used but they are certainly less reliable.

4) The mesh generator used, combined with a strategy restricted to mesh halving or doubling, is not optimal. It seems to initially over refine in smooth regions and then is slow to de-refine there because of the constraint of preserving mesh conformity. This question of *how* a posteriori information is best used in an adaptive procedure should be studied further since it influences the practical success of the overall computation.

6. References.

1. M. Ainsworth and J.T. Oden, *A unified approach to a posteriori error estimation using element residual methods*, Num. Math. 65 (1993), 23-50.

2. O. Axelsson and W. Layton, *Defect correction methods for convection dominated, convection diffusion equations*, RAIRO J. Numer. Anal. 24 (1990), 423-455.

3. O. Axelsson and W. Layton, *Iterative methods as discretization procedures in: Preconditioned Conjugate Gradient Methods*, Springer Berline 1457 (1991) (O. Axelsson and Yu. Kolotilina, eds.), Springer L.N.M.

4. I. Babuska, *Feedback, adaptivity and a posteriori estimates in finite elements*, Accuracy Estimates and Adaptive Refinement in Finite Element Computations (I. Babuska, O.C. Zienkiewicz, J. Gago and E.R. de A. Oliveira, eds.), Wiley, New York, pp. 3-23,(1986).

5. I. Babuska and W.C. Rheinboldt, *Error estimates for adaptive finite element computations*, SIAM J.N.A. 15 (1978), 736-754.

6. R.E. Bank and Weiser, *Some a posteriori error estimators for elliptic partial differential equations*, Math. Comp. 44 (1985), 283-301.

7. Ph. Clément, *Approximation by finite element functions using local regularization*, RAIRO Anal. Numer. 2 (1975), 77-84.

8. Q. Du and M. Gunzburger. *Analysis of a Ladzhenskaya model for incompressible viscous flow*. J.M.A.A. 155 (1991), 21-45.

9. K. Eriksson, D. Estep, P. Hansbo and C. Johnson, *Introduction to Adaptive Methods for Differential Equations*, Acta Numerica (1995), 1-54.

10. K. Eriksson and C. Johnson, *Adaptive streamline diffusion finite element methods for stationary convection-diffusion problems*, Math. Comp. 60 (1993), 167-188.

11. V. Ervin and W. Layton, *A study of defect correction finite difference methods for convection dominated, convection diffusion equations*, SIAM JNA 26 (1989), 169-179.

12. P. Hemker, *An accurate method without directional bias for the numerical solution of a 2-d elliptic singular perturbation problem*, Thy. and Appls. of Sing. Perts. (W. Eckhaus and E.M. de Jaeger, eds.), vol. 942, Springer LNM, Springer, Berlin (1982).

13. P.W. Hemker and B.Koren, *Defect correction and nonlinear multigrid for the steady Euler equations*, Advances in computational fluid dynamics (W.G. Habashi and M.M. Hafez, eds.), Cambridge University Press, Cambridge (1992).

14. P.W. Hemker and B. Koren, *Multigrid defect correction and upwind schemes for the steady Euler equations*, Num. Meth. for Fluid Dynamics III (K.W. Morton and M.J. Bainer, eds.), Clarendon, Oxford, pp. 153-170 (1988).

15. C. Johnson, *Adaptive finite element methods for diffusion and convection problems*, Comput. Methods Appl. Mech. Engrg. 82 (1990), 301-322.

16. B. Koren, *Multigrid and Defect Correction for the Steady Navier-Stokes Equations*, C.W.I. tract 74 Centrum vour Wiskunde en Informatica, Amsterdam (1991).

17. O.A. Ladyzhenkaya, *Modification of the Navier-Stokes Equations for large velocity gradients*, BVP's of Math, Phys. and Related Aspects of Funct. Thy., Consultants Bureau, New York (1970).

18. W. Layton, *Solution algorithms for incompressible viscous flows at high Reynolds number*, Comput. Math. and Cybernetics, series 15 (1996), 25-35.

19. W. Layton, *A nonlinear subgridscale model for incompressible, viscous flow problems*, SIAM J.Scientific Comput. (1996), 347-357.

20. J. Maubach, *Local bisection refinement for simplical grids generated by reflection*, SIAM J. Scientific Comput. 16 (1995), 210-227.

21. J. Maubach, *The amount of simplicial classes created by local n-simplicial bisection refinement*, preprint (1996).

22. J. Maubach and P. Rabier, *report in preparation*.

23. U. Navert, *A finite element method for convection diffusion problems*, Ph.D. Thesis, Chalmers Inst. of Tech. (1982).

24. J.T. Oden, L. Demkowicz. W. Rachiwicz and T.A. Westermann, *Toward a universal h − p adaptive finite element strategy, Part 2: a posteriori error estimation*, Comp. Meth. Appl. Mech. Engrg. 77 (1989), 113-180.

25. P.J. Rabier, *A posteriori error estimation "New" approach*, I.C.M.A. Report 94-188, Univ. of Pittsburgh (1994).

26. J.S. Smagorinsky, *General circulation model of the atmosphere*, Monthly Weather Review 91 (1963), 99-164.

27. E.G. Sewell, *Automatic generation of triangulations for piecewise polynomial approximations*, Ph.D. thesis, Purdue Univ., West Lafayette, Indiana (1972).

28. R. Verfürth, *A review of a posteriori error estimation and adaptive mesh refinement techniques*, Wiley-Teubner. Chichester (1996).