Adaptivity and Variational Stabilization for Convection-Diffusion Equations *

Albert Cohen, Wolfgang Dahmen, Gerrit Welper

January 31, 2011

Abstract

In this paper we propose and analyze stable variational formulations for convection diffusion problems starting from concepts introduced by Sangalli. We derive efficient and reliable a posteriori error estimators that are based on these formulations. The analysis of resulting adaptive solution concepts reveals partly unexpected phenomena related to the specific nature of the norms induced by the variational formulation. Several remedies are explored and illustrated by numerical experiments.

Key Words: Variational problems, adaptivity, a-posteriori error estimators, stabilization.

AMS Subject Classification: 65N12, 35J50, 65N30

1 Introduction

A notorious obstruction to an accurate and efficient numerical simulation of transport dominated processes is the interplay of transport in competition with diffusion. Its perhaps simplest manifestation is the classical linear convection-diffusion-reaction equation which arises in numerous contexts, in particular, in Oseen-type implicit discretizations of non-stationary incompressible Navier Stokes equations. It also serves as a guiding model when developing subgrid scale concepts such as the "variational multiscale method".

To be specific, we shall be concerned in what follows with the boundary value problem

$$-\epsilon\Delta u + b\cdot\nabla u + cu = f \quad \text{in } \Omega, \quad u = 0 \text{ on } \partial\Omega. \tag{1.1}$$

or rather the corresponding weak formulation: find $u \in H_0^1(\Omega)$ s.t.

$$a(u,v) := \epsilon\langle\nabla u, \nabla v\rangle + \langle b\cdot\nabla u, v\rangle + \langle cu, v\rangle = \langle f, v\rangle, \qquad v \in H_0^1(\Omega), \tag{1.2}$$

which, under well known conditions, admits a unique solution in $u \in H_0^1(\Omega)$. The central theme of this paper is the interplay between stabilization and adaptivity for problems of this type.

1.1 A Conceptual Preview

The current understanding of adaptive solution concepts with rigorous convergence and complexity estimates is confined to what we call (X,Y)-*stable* variational problems. By this we mean the following: let $a(\cdot,\cdot) : X \times Y \to \mathbb{R}$ be a bilinear form on a pair of Hilbert spaces X, Y with norms $\|\cdot\|_X, \|\cdot\|_Y$. Given a continuous linear functional $f \in Y'$, the normed dual of Y, find $u \in X$ such that

$$a(u,v) = f(v) \quad \forall\, v \in Y. \tag{1.3}$$

Here "(X,Y)-stable" means that the operator $A : X \to Y'$, induced by $\langle Au, v\rangle = a(u,v)$, $u \in X, v \in Y$, is a norm-isomorphism, i.e.

$$\|A\|_{X\to Y'} \le C_A, \quad \|A^{-1}\|_{Y'\to X} \le 1/\alpha \tag{1.4}$$

*This work has been supported in part by the Priority Program SPP 1324 funded by the German Research Foundation

holds for some constants $0 < \alpha, C_A < \infty$, independent of possibly varying problem parameters. Thus, the "condition" $\kappa_{X,Y}(A) := \|A\|_{X\to Y'}\|A^{-1}\|_{Y'\to X}$ of the operator is bounded by C_A/α.

An essential consequence of (X, Y)-stability is that errors in the "energy norm" $\|\cdot\|_X$ are bounded from below and above by the *dual norm of the residual*

$$\|A\|_{X\to Y'}^{-1}\|Au_h - f\|_{Y'} \leq \|u - u_h\|_X \leq \|A^{-1}\|_{Y'\to X}\|Au_h - f\|_{Y'}, \quad u_h \in X_h \subset X. \tag{1.5}$$

Thus, as long as $\kappa_{X,Y}(A)$ is of moderate size the residual bounds the error in a good way. All presently known adaptive methods rely in one way or the other on bounding the residuals $\|Au_h - f\|_{Y'}$ from above and below by sums of local terms whose size suggests further local refinements.

This has been realized so far primarily for symmetric problems that are known to be (X, X)-stable where X is a classical Sobolev space or a product of such. In principle, the convection-diffusion equation is in the above sense (H_0^1, H_0^1)-stable, but *not* robustly so, since for dominating convection, i.e. when $|b|/\epsilon \gg 1$, the condition behaves essentially like $\kappa_{H_0^1,H_0^1}(A) \sim |b|/\epsilon$, so that the relations (1.5) become useless.

This ill-conditioning already on the infinite dimensional level causes (at least) two major problem areas, namely (a) the development of *robust solvers* that are able to solve the finite dimensional systems of equations arising from a given discretization with an efficiency that is, for instance, independent of the parameters ϵ, b, c in (1.2), and (b) the choice of the discretization itself. Although these issues are not independent of each other we focus here exclusively on (b). In this regard, it is well-known that, unless a possibly unrealistically small mesh size (depending on ϵ) is chosen, a standard Galerkin discretization of (1.2) will be unstable. Substantial effort has been spent therefore on the development of *stabilization techniques*. One way is to add artificial diffusion in a consistent way and preferably only along stream lines, in order not to smear the possibly near singular (often anisotropic) features exhibited by the true solution. The perhaps most prominent example is the concept of streamline diffusion e.g. in the form of SUPG which in special cases can be understood also via the so called bubble function approach [1, 2]. In spite of the great success of such concepts it is fair to say that from several perspectives the current state of the art is not completely satisfactory. In particular, the choice of stabilization parameters is still a subtle issue that is not fully understood. This is reflected either by remaining unphysical oscillations in the numerical solution or by smearing solution features too much. In brief, stabilization can ameliorate somewhat but not avoid ill-conditioning.

Some of the stabilization techniques can be interpreted as mimicking, on a discrete level, a *Petrov-Galerkin* formulation, suggesting to choose also on the infinite dimensional level X different from Y which, after all seems natural since the underlying equation is not symmetric. In fact, such norms, giving rise in our terminology to an (X, Y)-stable formulation, have already been used in [17] and in [19, 20] for the purpose of error estimators. However, these error estimators have been derived for the SUPG, not for a discretization based on the (X, Y)-setting.

The central objective of this paper to explore (X, Y)-stable weak formulations for the convection diffusion equation (1.1) along with discretizations directly based on them. Our two main principal reasons for thereby proposing an alternative to mesh-dependent (a-priori) stabilization techniques may be summarized as follows.

First, it is in our opinion desirable to avoid *mixing of numerical and physical stabilization concepts*. In fact, convection-diffusion equations often arise in connection with the *Variational Multiscale Method* ([13, 14]) that aims at capturing the effect of unresolved scales on a given macro scale that might entirely depend on the given numerical budget. This still requires at the end of the day solving a possibly convection dominated problem in the range of resolved scales endowed with a correction term that is to capture the effect of the unresolved scales on the resolved macro-scale. It is then unsatisfactory that several stabilization terms are ultimately superimposed.

The second aspect, which is the primary focus of the present research, is to explore the combination of (X, Y)-stability with adaptive solution concepts based on the corresponding (X, Y)-discretizations and the interplay of their stability with adaptivity. It was experimentally demonstrated in [16] that an adaptive full multigrid scheme for plain Galerkin discretizations of convection dominated problems lead to perfectly stable solutions as long as layers are resolved. In the present context the stabilizing effect of adaptivity enters in a somewhat different way. (X, Y)-stability comes at a prize, namely the resulting variational formulation involves an inner product $\langle\cdot,\cdot\rangle_Y$ that is not readily numerically accessible. We propose to introduce an auxiliary variable (on the infinite dimensional level), somewhat in the spirit of mixed formulations, that require solving an auxiliary elliptic problem. The exact solution of this problem corresponds to the perfectly stable formulation and the access resolution of its discretization compared with the trial space for the "primal" variable can be interpreted as the "amount of stabilization". The point is that efficient and reliable a-posteriori bounds for errors in the X-norm allow us to determine the right level of stabilization *adaptively*.

Nevertheless, resulting adaptive schemes turn out *not* to give rise to a fixed error reduction. A closer look reveals that this is to a great extent due to some summands contained in most norms for the analysis of convection-diffusion problems including the ones used in this paper as well as the norms used for SUPG-schemes. This issue will be discussed along with possible remedies.

We emphasize that the central objective of this paper is *not* to propose a specific definitive scheme that should deal with these obstructions. In fact, we are content here with simple low order finite elements on isotropic refinements, although it will be clear from the experiments that higher order trial functions in layer regions and anisotropic refinements would ameliorate somewhat the observed adverse effects. However, we prefer here to focus here on some principal mechanisms that are in our opinion relevant in this context and may help to shed some more light on the central issues. One such central issue is to allow arbitrarily small diffusion without going to the reduced problem though since viscosity is essential for important small scale effects.

1.2 Layout of the Paper

The paper is organized as follows: in Section 2 we recall the infinite dimensional setting of [20, 17]. A general strategy is introduced by which one gets numerical schemes for such well conditioned settings. Furthermore relations to subgrid modeling and the SUPG scheme are discussed. Section 3 focuses on adaptive strategies for the convection-diffusion problem. One the one hand it is used to get adaptive approximations of the solution and on the other hand it is used to choose a proper amount of stabilization. Unfortunately it turns out that these schemes generate some artifacts. In Section 4 it is argued that these artefacts are induced by the specific norms and some remedies are proposed and illustrated numerically.

After completion of this paper we became aware of recent related work reported in [9, 10]. It centers on the notion of "optimal test spaces" for Petrov-Galerkin discretizations which is closely related to the functional analytic framework developed below in Section 2. However, it is pursued in a quite different direction than in the present study.

2 A stable elliptic formulation and its discretization

2.1 Construction of norms

In this section we describe a functional analytic setting essentially following [17] and [20]. To this end assume $\Omega \subset \mathbb{R}^n$ is a domain, $b \in W^{1,\infty}(\Omega)$, $c \in L_\infty(\Omega)$ with

$$c - \frac{1}{2}\mathrm{div} b \geq \bar{c} \qquad\qquad \|b\|_\infty \leq c^* \bar{c} \tag{2.1}$$

where $\bar{c}, c^* \geq 0$ are two constants and c and b are the coefficients of the convection-diffusion problem (1.2). Under these assumptions it is well known that the bilinear form a from (1.2) induces an isomorphism $A : H_0^1(\Omega) \to H^{-1}(\Omega)$ through $\langle Av, w\rangle = a(v, w)$ for all $v, w \in H_0^1(\Omega)$. Here and below $\langle \cdot, \cdot \rangle$ always denote the dual pairing induced by the standard L_2-inner product.

We wish to to define now new equivalent norms $\|\cdot\|_X$, $\|\cdot\|_Y$ on $H_0^1(\Omega)$, considered as a set, in such a way that the operator A has condition number $\kappa_{X,Y}(A)$ equal to one. Of course, such norms and their equivalence constants to the standard norm $H_0^1(\Omega)$ must depend on the problem parameters ϵ, c and b.

We denote by $\langle \cdot, \cdot \rangle_X = \langle \cdot, R_X \cdot \rangle$ and $\langle \cdot, \cdot \rangle_Y = \langle \cdot, R_Y \cdot \rangle$ the respective scalar products and R_X where R_Y are the corresponding Riesz maps.

The general procedure can be described as follow: Fix any norm $\|\cdot\|_Y$ on $H_0^1(\Omega)$ (as a set). Then for any other norm $\|\cdot\|_X$ which is also equivalent to $\|\cdot\|_{H^1(\Omega)}$, the operator A, defined above, is still an isomorphism from X onto Y'. Therefore, the problem: for $f \in Y'$ find $u \in X$ such that

$$a(u, v) = \langle f, v \rangle, \quad \forall\ v \in Y, \tag{2.2}$$

is a weak formulation of (1.1) that is equivalent to

$$\langle Au, Av\rangle_{Y'} = \langle f, Av\rangle_{Y'} \quad \forall\, v \in X. \tag{2.3}$$

Choosing now the X-scalar product as

$$\langle v, w\rangle_X := \langle Av, Aw\rangle_{Y'}, \quad v, w \in X, \tag{2.4}$$

we have

$$\|v\|_X^2 = \|Av\|_{Y'}^2, \tag{2.5}$$

which means that (1.5) holds with $\|A\|_{X\to Y'} = \|A^{-1}\|_{Y'\to X} = 1$ and (2.3) is perfectly conditioned.

Remark 2.1. *Obviously, (2.3) is the normal equation for the infinite dimensional least squares problem*

$$u = \mathrm{argmin}_{v\in X}\|Av - f\|_{Y'}. \tag{2.6}$$

Moreover, for any subspace $X_h \subseteq X$ one has

$$u_h = \mathrm{argmin}_{v_h\in X_h}\|u - u_h\|_X \iff u_h = \mathrm{argmin}_{v_h\in X_h}\|f - Au_h\|_{Y'}, \tag{2.7}$$

and replacing X by X_h in (2.3), this variational problem is the normal equation for the least squares problem (2.7).

Following [18, 20], a canonical, but by no means mandatory choice for $\|\cdot\|_Y$ is obtained by splitting the original form $a(\cdot,\cdot)$ into its symmetric and skew-symmetric parts

$$a_s(u, v) := \frac{1}{2}\big(a(u, v) + a(v, u)\big), \qquad a_{sk}(u, v) := \frac{1}{2}\big(a(u, v) - a(v, u)\big),$$

which gives

$$a(v, w) = a_s(v, w) + a_{sk}(v, w) =: \langle A_s v, w\rangle + \langle A_{sk} v, w\rangle, \quad (v, w) \in X \times Y, \tag{2.8}$$

where, in particular,

$$A_{sk}^* = -A_{sk} \quad \text{so that} \quad a_{sk}(v, v) = 0 \quad \text{and } A = A_s + A_{sk}. \tag{2.9}$$

Taking

$$\|y\|_Y^2 := a_s(y, y), \tag{2.10}$$

this norm is indeed equivalent to $\|\cdot\|_{H^1(\Omega)}$ and the corresponding Riesz map is $R_Y = A_s$ and specifically, for problem (1.2) one obtains

$$\langle v, w\rangle_Y := \langle A_s v, w\rangle = \epsilon\langle\nabla v, \nabla w\rangle + \langle(c - \frac{1}{2}\mathrm{div}(b)v, w\rangle. \tag{2.11}$$

Note that in this case (2.5), by cancellation of the skew-symmetric term, takes the form

$$\|v\|_X^2 = \|v\|_Y^2 + \|A_{sk} v\|_{Y'}^2. \tag{2.12}$$

The main obstacle for an implementation of the scheme (2.3) is that in general the Y' norm is a dual norm so that the scalar product cannot easily be evaluated. This issue is addressed in the next section.

2.2 Resolving the Y'-scalar product

Of course, a direct Galerkin formulation for (2.3) is not possible since the inner product $\langle\cdot,\cdot\rangle_{Y'}$ is not easily evaluated. To deal with this fact first note that

$$\|R_Y v\|_{Y'} = \|v\|_Y \qquad\qquad \|w\|_{Y'} = \|R_Y^{-1} w\|_Y, \tag{2.13}$$

and recall that by the definition of the Riesz map, the inner product of Y' is given by

$$\langle v, w\rangle_{Y'} = \langle v, R_Y^{-1} w\rangle = \langle R_Y^{-1} v, w\rangle. \tag{2.14}$$

This explicit formula for the Y'-scalar product can be used to restate the variational problem (2.3) as

$$\langle Au - f, R_Y^{-1}Av\rangle = 0, \quad \forall\, v \in X, \tag{2.15}$$

which is now based on regular L_2-inner products but involves the inverse of R_Y. Before we proceed note that this variational formulation is essentially an (infinite dimensional) Petrov-Galerkin formulation with the ideal test space $R_Y^{-1}AX$. However, introducing an auxiliary variable

$$y = R_Y^{-1}(Au - f) \quad \Leftrightarrow \quad R_Y y - Au = -f,$$

the problem (2.15) is equivalent to finding $(u, y) \in X \times Y$ such that

$$\begin{aligned} \langle y, Av\rangle &= 0 \qquad && \forall\, v \in X, \\ \langle Au, z\rangle - \langle R_Y y, z\rangle &= \langle f, z\rangle && \forall\, z \in Y. \end{aligned} \tag{2.16}$$

Thus at the expense of an additional variable y we arrive at a variational problem that could be treated, for instance, by standard finite element discretizations. First let us confirm though that the system (2.16) is well posed. To this end, let

$$\bar{a}([u, y], [v, z]) := \langle y, Av\rangle + \langle Au, z\rangle - \langle R_Y y, z\rangle, \tag{2.17}$$

which is the bilinear form $\bar{a} : (X \times Y) \times (X \times Y) \to \mathbb{R}$ corresponding to (2.16).

Proposition 2.2. *The bilinear form (2.17) defines an isomorphism $\bar{A} : (X \times Y') \to (X' \times Y')$.*

Proof: We denote by $\bar{A} : X \times Y \to X' \times Y'$ the operator corresponding to the bilinear form (2.17). We have to show that this operator is an isomorphism. To this end we consider the equation $\bar{A}[u, y] = [g, f]$. According to (2.16) this is equivalent to the following system

$$\begin{aligned} A^* y &= g \\ Au - R_Y y &= f. \end{aligned} \tag{2.18}$$

To see that this system possesses a unique solution, recall that $A^* : Y \to X'$ is an isomorphism. Therefore $y := A^{-*}g$ satisfies the first row of the system. Furthermore, since $A : X \to Y$ is an isomorphism we infer form the second row $u = A^{-1}(f - R_Y y)$, showing that (2.18) has a solution.

To show its uniqueness we prove that the operator $\bar{A}$ has a trivial kernel. To this end, assume $\bar{A}[u, y] = 0$. Again by the first row of (2.18) we get $y = 0$. Plugging this into the second row yields $Au = 0$ which implies $u = 0$. Thus $\bar{A}$ is injective.

Now one can explicitly write down the inverse of $\bar{A}$ which is given by

$$\bar{A}^{-1}[g, f] = [A^{-1}(f - R_Y A^{-*}g), A^{-*}g] \tag{2.19}$$

Finally we have to show that the norms of $\bar{A}$ and its inverse $\bar{A}^{-1}$ are bounded. First we get

$$\|\bar{A}[u, y]\|^2_{X' \times Y'} = \|A^* y\|^2_{X'} + \|Au - R_Y y\|^2_{Y'} \leq 2\|u\|^2_X + 3\|y\|^2_Y, \tag{2.20}$$

where we have used the definition $\|u\|_X = \|Au\|_{Y'}$ of the X-norm and its consequence $\|A^* y\|_{X'} = \|y\|_Y$. Finally, by the explicit formula (2.19) for the inverse, we obtain, again using $\|A^{-*}g\|_Y = \|g\|_{X'}$,

$$\begin{aligned} \|\bar{A}^{-1}[g, f]\|^2_{X \times Y} &= \|A^{-1}(f - R_Y A^{-*}g)\|^2_X + \|A^{-*}g\|^2_Y \\ &\leq \|f - R_Y A^{-*}g\|^2_{Y'} + \|g\|^2_{X'} \leq 2\|f\|_{Y'} + 3\|g\|^2_{Y'}. \end{aligned}$$

□

We conclude this section with the simple observation that the above strategy of first prescribing Y and then choosing X through setting $\|v\|_X := \|Av\|_{Y'}$ can be reversed which is the point of view taken in [8] for a different problem class. In fact, first prescribing X and setting $\|y\|_Y = \|A^* y\|_{X'}$ yields the dual space $\|f\|_{Y'} = \|A^{-1}f\|_X$ so that one gets again the desired mapping property $\|Au\|_{Y'} = \|u\|_X$ and (X, Y)-stability with perfect condition $\kappa_{X,Y}(A) = 1$.

2.3 Discretization and Stabilization

As in the case of the infinite dimensional convection-diffusion problem (1.1) we always assume that for any $X_h \subseteq X$

$$a(u_h, v_h) = 0,\ v_h \in X_h \quad \Rightarrow \quad u_h = 0. \tag{2.21}$$

The problem (2.16), or equivalently

$$\bar{a}([u,y],[v,z]) = \langle f, z\rangle \quad \forall\ [v,z] \in X \times Y, \tag{2.22}$$

can be treated in the usual way. Specifically, suppose that $X_h \subset X$ and $Y_h \subset Y$ are finite dimensional trial spaces for the two solution components. To understand the roles of X_h and Y_h with regard to the stability of the discretization, note first that the resulting finite dimensional problem

$$\bar{a}([u_h,y_h],[v_h,z_h]) = \langle f, z_h\rangle \quad \forall\ [v_h,z_h] \in X_h \times Y_h, \tag{2.23}$$

or equivalently written as a block system,

$$\begin{array}{rcll} \langle y_h, Av_h\rangle & = & 0 & \forall\ v_h \in X_h \\ \langle Au_h, z_h\rangle - \langle R_Y y_h, z_h\rangle & = & \langle f, z_h\rangle & \forall\ z_h \in Y_h. \end{array} \tag{2.24}$$

could, in principle, be very ill-conditioned. In fact, consider first the case $X_h = Y_h$. Testing with $[v_h, 0]$, $v_h \in Y_h = X_h$, reveals that, by (2.21), $y_h = 0$, which, by testing with $[0, z_h]$ gives $a(u_h, z_h) = \langle f, z_h\rangle$, $z_h \in Y_h = X_h$. This is simply the original Galerkin discretization which is unstable. On the other hand, choosing for any finite dimensional $X_h \subset X$ the second component as the whole infinite dimensional space Y, and calling the resulting solution $[u_h, \hat{y}_h]$, we can redo the steps (2.15) to (2.16) for the derivation of the block system. Namely setting v_h to zero gives $\hat{y}_h = R_Y^{-1}(Au_h - f)$. Plugging this again into the discrete system (2.23) and setting $z_h = 0$ gives the original least squares problem

$$\langle Au_h - f, Av_h\rangle = 0 \qquad \text{for all } v_h \in X_h$$

which by Remark 2.1 gives the optimal discrete approximation in the X-norm.

In summary, the same discretization for both components u and y is unstable while an infinite resolution for the auxiliary variable y gives rise to a stable formulation with condition equal to one. This suggests that the excess resolution of y relative to u acts as a *stabilization*.

That raises two questions: (i) can we find a-priori criteria to determine how large should Y_h be for a given X_h, or more generally, how Y_h should be related to X_h to warrant uniform stability of the discrete problems? (ii) are there practical a-posteriori error indicators that tell us at any given stage how much stabilization is needed to warrant a desired target accuracy of an approximate solution. This would give rise to another manifestation of adaptivity as a stabilizing concept. We defer (i) and address (ii) in Section 3.3 below. It is important to note that in this context a discrete inf-sup condition is *not* needed.

2.4 Subgrid Modeling and SUPG

Next we briefly pause to put the above numerical scheme into the context of subgrid modeling and compare it to the SUPG scheme.

Accepting the fact that turbulent flows involve a range of relevant scales that can usually not be resolved by a numerical scheme, one may try to still capture the effect of unresolved scales on the macro scale by modeling them. Examples are low parameter models, Large Eddy Simulation or modern variants such as Variational Multiscale Method and Subgrid Modeling. As pointed out in [11] all these approaches may be interpreted as *regularizing* the Navier Stokes equations. It is therefore perhaps interesting to see how this fits into the present setting. To this end, let $X_h \subset X$ be again a finite dimensional trial space while in general Y_h should be a larger space which, for simplicity of exposition is for the moment taken to be all of Y. This is the second scenario considered in the previous subsection leading to a conceptually stable formulation for the semi-discrete problem on $X_h \times Y$.

Next we assume that $X_h \subset Y$ which is of course true for the above choice $X = Y = H_0^1(\Omega)$ endowed with suitably modified norms. In order to interpret the scheme (2.24) b in terms of subgrid modeling consider its first row. Defining the Galerkin projection $P_h : Y \to X_h$ by

$$a(v_h, P_h z - z) = 0 \quad \forall\, v_h \in X_h, \tag{2.25}$$

it states that $P_h \hat{y}_h = 0$, i.e. $\hat{y}_h$, defined in the previous subsection, is a *fluctuation* in nature. In order to use this observation in the second row of (2.24) we first decompose the space Y into

$$Y = P_h \oplus (I - P_h)Y = X_h \oplus X_h^{\perp}.$$

Now we can split the test functions of (2.24) into test functions from X_h and from $X_h^{\perp}$ which gives

$$\begin{aligned} \langle Au_h, z_h\rangle - \langle R_Y y, z_h\rangle &= \langle f, z_h\rangle && \forall\, z_h \in X_h \\ -\langle R_Y y, z_h\rangle &= \langle f, z_h\rangle && \forall\, z_h \in X_h^{\perp}, \end{aligned}$$

where we have used that $\langle Au_h, z_h\rangle = 0$ for all $z_h \in X_h^{\perp}$. In this equation we can interpret the first row as a plain Galerkin scheme with a correction $-\langle R_y \hat{y}_h, z_h\rangle$. Since $P_h \hat{y}_h = 0$ this represents essentially the influence of the unresolved scales on the resolved scales.

In order to compare the scheme (2.23) with the SUPG scheme we first rewrite the least squares problem (2.3) in a slightly different form. To this end for the choice $R_Y = A_s$ by (2.14) and the decomposition $A = A_s + A_{sk}$ for the least squares problem (2.3) one gets

$$\langle Au, v\rangle + \langle Au, A_{sk}\rangle_{Y'} = \langle f, v\rangle + \langle f, A_s k\rangle_{Y'}.$$

Note that the stabilization terms added in SUPG could be viewed as mimicking the terms $\langle Au, A_{sk} v\rangle_{Y'}$ and $\langle f, A_{sk} v\rangle_{Y'}$ by suitably weighted element-wise defined inner products.

3 Adaptive strategies

In this section we investigate adaptive strategies for the system (2.24). The reason for adaptivity for this problem is twofold. On the one hand solutions of convection-diffusion problems typically have layers which we want to resolve adaptively. On the other hand we have already seen in Section 2.3 that the resolution of the auxiliary variable y determines the amount of stabilization.

To keep this article self contained in the next section we briefly recall the error estimators from [20]. Base on those estimators in the Section 3.2 we derive a-posteriori error estimators for the block system (2.24). Then in Section 3.3 we prove suitable a-posteriori conditions for stability and in Section 3.4 we give a short numerical experiment.

3.1 The error estimators of Verfürth

In this section we recall the error estimators from [20] applied to the model problem: find $y_h \in Y_h$ s.t.

$$\langle R_Y y_h, z_h\rangle = \langle f - Au_h, z_h\rangle \qquad \text{for all } z_h \in Y_h \tag{3.1}$$

where we use the choice $R_Y = A_s$ in this section. This problem is, of course, the second block row of (2.16) or rather the discrete case (2.23). In this section we treat u_h as a fixed function. In the next section we use this estimator for the full system where also u_h is an unknown variable. This setting is slightly more general than the results in [20] because there only right hand sides which are piecewise constant on a finite element gird are allowed. However the statements of this section with this more general right hand side can be proved in identical ways.

To state the main result of [20] we first need some notation. So far we have not considered the type of finite dimensional spaces for X_h and Y_h. In this section we will fix it to finite element spaces. To this end let $\mathcal{T}_h^X$, $\mathcal{T}_h^Y$ be admissible and shape regular triangulation and $\mathcal{E}_h^X$, $\mathcal{E}_h^Y$ be the corresponding set of edges. Then X_h and Y_h are the corresponding finite element space

$$\begin{aligned} X_h &= \{u \in H_0^1(\Omega) : \, u|_T \in \mathcal{P}^k, \, T \in \mathcal{T}_h^X\} \\ Y_h &= \{y \in H_0^1(\Omega) : \, y|_T \in \mathcal{P}^l, \, T \in \mathcal{T}_h^Y\} \end{aligned}$$

of continuous piecewise polynomial function of degree k and l respectively. In order to be able to apply the theory from [20] we have to assume that the spaces X_h and Y_h are nested, i.e. $X_h \subset Y_h$. The reason is the appearance of the variable u_h in the right hand side. Furthermore n_E is the outward normal to the edge E and $[\cdot]_E$ is the jump across E. Because for the error estimators we are mainly concerned with the finite element space Y_h we will usually drop the superscript ^{Y}. Next define

$$\alpha_S := \min\{\epsilon^{-1/2} h_S, \bar{c}^{-1/2}\}, \quad S \in \{T, E\},\ h_S := \operatorname{diam} S, \tag{3.2}$$

where $\bar{c}$ is the constant form (2.1). Here the diameters h_S and cells T and E might correspond to either triangulation $\mathcal{T}_h^X$ or $\mathcal{T}_h^Y$ which will be clear from the context.

Along the lines of [20] we now define the element and edge residuals

$$\begin{aligned} \rho_T(u_h, y_h) &:= f - Au_h + R_Y y_h|_T \\ &= \left(f + \epsilon\Delta(u_h - y_h) - b\nabla u_h + c(y_h - u_h) - \frac{1}{2}\operatorname{div}(b) y_h\right)\Big|_T \end{aligned} \tag{3.3}$$

$$\rho_E(u_h, y_h) := \begin{cases} \epsilon n_E [\nabla(y_h - u_h)]_E & \text{if } E \not\subset \Gamma, \\ 0 & \text{if } E \subset \Gamma. \end{cases} \tag{3.4}$$

Next we define an error indicator for one element T by

$$\eta_T(u_h, y_h, T)^2 := \alpha_T^2 \|\rho_T(u_h, y_h)\|_{0;T}^2 + \epsilon^{-1/2}\alpha_E \|\rho_{\partial T}(u_h, y_h)\|_{0;\partial T}^2 \tag{3.5}$$

as well as for a set $\mathcal{T}' \subset \mathcal{T}_h$ of elements by

$$\eta_{\mathcal{T}_h}(u_h, y_h, \mathcal{T}')^2 = \sum_{T \in \mathcal{T}'} \eta_{\mathcal{T}_h}(u_h, y_h, T)^2.$$

Finally we define the data error

$$osc(f, \mathcal{T}_h^Y)^2 = \inf_{z_h \in Y_h} \|f - z_h\|_{Y'}^2, \tag{3.6}$$

where errors in b and c are excluded for simplicity. Then one gets form [20] the following theorem

Theorem 3.1. *Assume that the error indicators $\eta(u_h, y_h, T)$ and the data errors $osc(f, \mathcal{T}_h^Y)$ are defined as in (3.5) and (3.6) respectively. Furthermore assume that $X_h \subset Y_h$. Then we have the upper bound*

$$\|\hat{y} - y_h\|_Y^2 \le \eta(u_h, y_h, \mathcal{T}_h^Y)^2 + osc(f, \mathcal{T}_h^Y)^2$$

and the lower bound

$$\eta(u_h, y_h, \mathcal{T}_h^Y)^2 \le \|\hat{y} - y_h\|_Y^2 + osc(f, \mathcal{T}_h^Y)^2$$

where $\hat{y}$ is the solution of (3.1) for fixed $u_h \in X_h$.

3.2 Error estimators for the block system

In this section we analyze a-posteriori error estimators for the full discrete system

$$\bar{a}([u_h, y_h], [v_h, z_h]) = \langle f, z_h \rangle \quad \forall\ [v_h, z_h] \in X_h \times Y_h. \tag{3.7}$$

with the choice $R_Y = A_s$. Especially we want to derive efficient and reliable bounds for the error

$$\|u - u_h\|_X^2 + \|y - y_h\|_Y^2 = \|u - u_h\|_X^2 + \|y_h\|_Y^2, \tag{3.8}$$

where we have used that $y = R_Y^{-1}(Au - f) = 0$ holds for the solution u of (2.2). Moreover, since by (2.5) we have $\|u - u_h\|_X = \|A(u - u_h)\|_{Y'} = \|f - Au_h\|_{Y'}$, the term we need to estimate is just

$$\|Au_h - f\|_{Y'}^2 + \|y_h\|_Y^2. \tag{3.9}$$

The second summand is already a computable quantity. However, it is always dominated by the first summand in (3.9). In fact, choosing $v_h = 0$ in (3.7) implies

$$\langle R_Y y_h, z_h\rangle = a(u_h, z_h) - \langle f, z_h\rangle = \langle Au_h - f, z_h\rangle,$$

so that by standard reasoning

$$\begin{aligned}
\|y_h\|_Y = \frac{\langle R_Y y_h, y_h\rangle}{\|y_h\|_Y} &= \sup_{z_h\in Y_h} \frac{\langle R_Y y_h, z_h\rangle}{\|z_h\|_Y} \\
&= \sup_{z_h\in Y_h} \frac{\langle Au_h - f, z_h\rangle}{\|z_h\|_Y} \le \sup_{z\in Y} \frac{\langle Au_h - f, z\rangle}{\|z\|_Y} \\
&= \|Au_h - f\|_{Y'},
\end{aligned} \tag{3.10}$$

which therefore gives for any $Y_h \subseteq Y$

$$\|Au_h - f\|_{Y'}^2 \le \|u - u_h\|_X^2 + \|y - y_h\|_Y^2 \le 2\|Au_h - f\|_{Y'}^2. \tag{3.11}$$

Again it is instructive to look at the following two choices of Y_h. If $Y_h = X_h$ we have $y_h = 0$ which corresponds to an unstable discretization. Nonetheless, the a-posteriori indicators from [19, 20] would yield sharp lower and upper bounds for $\|Au_h - f\|_{Y'}$. This may, however, not give useful information for refinements since these indicators reflect also the unphysical oscillations in regions where no refinement would be needed. The other extreme case is $Y_h = Y$. Then (3.10) actually gives

$$\|y_h\|_Y = \|Au_h - f\|_{Y'} \quad \text{whenever} \quad Y_h = Y. \tag{3.12}$$

Hence in the "stable" situation both terms in the error bound (3.9) are the same. Thus a reasonable refinement strategy should aim at balancing both contributions. We take up this point of view again later in connection with a concrete refinement strategy.

Let us now turn to identifying corresponding computable indicators for the error (3.8). Using Galerkin orthogonality one has

$$\bar{a}([u - u_h, y - y_h], [v_h, z_h]) = 0 \quad \forall\ [v_h, z_h] \in X_h \times Y_h, \tag{3.13}$$

for the discrete solution $[u_h, y_h] \in X_h \times Y_h$ of (3.7) so that we obtain by (2.2) in the usual manner

$$\Big(\|u - u_h\|_X^2 + \|y_h\|_Y^2\Big)^{1/2} \le 2 \sup_{[v,z]\in X\times Y} \frac{\bar{a}([u - u_h, -y_h], [v - v_h, z - z_h])}{\Big(\|v\|_X^2 + \|z\|_Y^2\Big)^{1/2}}. \tag{3.14}$$

Abbreviating $e_r := r - r_h$, $r \in \{u, y, v, z\}$, straightforward calculations yield

$$\begin{aligned}
\bar{a}([e_u, -y_h], [e_v, e_z]) &= -\langle y_h, Ae_v\rangle + \langle Ae_u + R_Y y_h, e_z\rangle \\
&= -\langle y_h, Ae_v\rangle + \langle f - Au_h + R_Y y_h, e_z\rangle \\
&= -\langle y_h, Av\rangle + \langle f - Au_h + R_Y y_h, e_z\rangle,
\end{aligned} \tag{3.15}$$

where we have used that by (3.7)

$$\langle y_h, Av_h\rangle = 0 \qquad \forall\ v_h \in X_h. \tag{3.16}$$

Now with Theorem 3.1 we can estimate

$$|\langle f - Au_h + R_Y y_h, e_z\rangle| \le C\eta_{\mathcal{T}}(u_h, y_h, \mathcal{T}')\|z\|_Y + osc(f, \mathcal{T}_h^Y)\|z\|_Y. \tag{3.17}$$

Moreover,by (2.4), we have

$$|\langle y_h, Av\rangle| \le \|y_h\|_Y \|Av\|_{Y'} = \|y_h\|_Y \|v\|_X. \tag{3.18}$$

Thus combining (3.17), (3.18) and (3.15), yields

$$\begin{aligned}
|\bar{a}([e_u, e_y], [e_v, e_z])| &\le C\|y_h\|_Y\|v\|_X + \eta_{\mathcal{T}}(u_h, y_h, \mathcal{T})\|z\|_Y + osc(f, \mathcal{T}_h^Y)\|z\|_Y \\
&\le C'\Big(\|y_h\|_Y^2 + \eta_{\mathcal{T}}(u_h, y_h, \mathcal{T})^2 + osc(f, \mathcal{T}_h^Y)^2\Big)^{1/2} \\
&\qquad \times \Big(\|v\|_X^2 + \|z\|_Y^2\Big)^{1/2}.
\end{aligned} \tag{3.19}$$

Thus, we can infer from (3.14) that

$$\|u-u_h\|_X^2+\|y-y_h\|_Y^2 \leq C\left(\|y_h\|_Y^2+\eta_{\mathcal{T}}(u_h,y_h,\mathcal{T})^2+osc(f,\mathcal{T}_h^Y)^2\right). \tag{3.20}$$

Note that $\|y_h\|_Y^2+\eta_{\mathcal{T}}(u_h,y_h,\mathcal{T}')^2$ are computable and in fact localizable quantities that can, in principle, be used for steering adaptive refinements for both X_h and Y_h.

In view of (3.10) and (3.11), to derive lower bounds, it suffices to show that $\eta_{\mathcal{T}}(u_h,y_h,\mathcal{T}')^2$ is bounded by a constant multiple of $\|u-u_h\|_X^2+\|y_h\|^2$ plus data oscillations. That this is indeed the case follows from Theorem 3.1 which yields

$$\eta(u_h,y_h,\mathcal{T}_h^Y)^2 \leq \|\hat{y}-y_h\|_Y^2+osc(f,\mathcal{T}_h^Y)^2$$

Now we can use that $\|\hat{y}\|=\|R_Y^{-1}(Au-Au_h)\|_Y=\|u-u_h\|_X$ which gives

$$\eta(u_h,y_h,\mathcal{T}_h^Y)^2 \lesssim \|u-u_h\|_X^2+2\|y_h\|_Y^2+osc(f,\mathcal{T}_h^Y)^2$$

as desired.

In summary we therefore obtain the following result.

Proposition 3.2. *In the above terms one has for a uniform constant C*

$$\|u-u_h\|_X^2+\|y_h\|_Y^2 \leq C(\eta(u_h,y_h,\mathcal{T}_h^Y)^2+osc(f,\mathcal{T}_h^Y)^2),$$

as well as

$$\eta(u_h,y_h,\mathcal{T}_h^Y)^2 \leq C(\|u-u_h\|_X^2+\|y_h\|_Y^2+osc(f,\mathcal{T}_h^Y)^2).$$

3.3 Adaptive Stabilization

When trying to use the above error indicators for steering an adaptive refinement process, it is not quite clear how exactly should one treat the terms involving the auxiliary variable y relative to the terms involving u. In fact, as we have already seen in Section 2.3, the resolutions of the space Y_h in relation to X_h determines the level of stabilization. The following lemma gives a sufficient condition for Y_h to ensure a stable discretization.

Lemma 3.1. *Assume $u_h \in X_h$ and $y_h \in Y_h$ are solutions of the scheme* (2.24) *and that for some fixed $\delta \in (0,2)$ we have*

$$\|R_Y^{-1}(Au_h-f)-y_h\|_Y \leq \delta\|y_h\|_Y. \tag{3.21}$$

Then we get

$$\|u-u_h\|_X+\|y-y_h\|_Y \leq 4\left(1-\frac{\delta}{2}\right)^{-2}\inf_{\phi\in X_h}\|u-\phi\|_X \tag{3.22}$$

Before proving this lemma, some comments on assumption (3.21) are in order. Recall that for a fixed u_h the corresponding auxiliary variable y_h is given by the variational problem

$$\langle Au_h,z_h\rangle-\langle R_Y y_h,z_h\rangle=\langle f,z_h\rangle \tag{3.23}$$

for all $z_h \in Y_h$. Thus y_h is the Y-orthogonal projection of the exact infinite dimensional solution $\hat{y}_h = R_Y^{-1}(Au_h-f)$, i.e. it is a Galerkin approximation for the elliptic variational problem $a_Y(y_h,z_h) = \langle Au_h,z_h\rangle-\langle f,z_h\rangle$. Therefore, we can apply the a-posteriori error estimators from [19, 20] to control the left hand side of (3.21). Furthermore, the right hand side of (3.21) can be computed explicitly. Thus, one can check a-posteriori whether the resolution of Y_h is sufficiently high and, if necessary, refine the gird of Y_h.

Also note that despite the fact that we have two variables u and y the overall error is governed by the approximation error of u alone. Intuitively this is reasonable because the approximation error $\inf_{\varphi\in Y_h}\|y-\varphi\|_Y=0$ because $y=0$.

Proof of Lemma 3.1: We employ the following abbreviations

$$\begin{aligned} e_u &= u-u_h & e_y &= y-y_h \\ d_u &= u-P_{X_h}u & d_y &= y-P_{Y_h}y, \end{aligned}$$

and as before

$$\hat{y}_h = R_Y^{-1}(Au_h - f) = -R_Y^{-1} A e_u,$$

where P_{X_h} and P_{Y_h} are the X and Y orthogonal projectors onto X_h and Y_h, respectively, i.e. d_u and d_y are the residuals of the best approximation, respectively.

With the definition (2.4) of the X-norm we may proceed as in the proof of the Céa lemma.

$$\|e_u\|_X^2 = \langle Ae_u, Ae_u\rangle_{Y'} = -\langle Ae_u, \hat{y}_h\rangle = -\langle Ae_u, y_h\rangle + \langle Ae_u, y_h - \hat{y}_h\rangle$$

and using the second block row of (2.24) we have

$$\|e_y\|_Y^2 = \langle R_Y y_h, y_h\rangle = \langle Au_h - f, y_h\rangle = -\langle Ae_u, y_h\rangle.$$

This yields

$$\|e_u\|_X^2 + \|e_y\|_Y^2 = -2\langle Ae_u, y_h\rangle + \langle Ae_u, y_h - \hat{y}_h\rangle. \tag{3.24}$$

Using Galerkin orthogonality in the first block row of (2.24), we obtain $\langle Ae_u, y_h\rangle = \langle Ad_u, y_h\rangle$. Furthermore, by the assumption (3.21) we have

$$\|y_h - \hat{y}_h\|_Y \le \delta\|y_h\|_Y = \delta\|e_y\|_Y,$$

where we have used that $y = 0$. Thus, with (3.24) and (2.5) we conclude that

$$\|e_u\|^2 + \|e_y\|^2 \le 2\|d_u\|_X\|e_y\|_Y + \delta\|e_u\|_X\|e_y\|_Y.$$

Now Young's inequality $ab \le \frac{1}{2c}a^2 + \frac{c}{2}b^2$ with $c = \frac{1}{2}\left(1 - \frac{\delta}{2}\right)$ for the first summand and $c = 1$ for the second summand of the right hand side yields the desired error estimate (3.22) for $\delta < 2$. $\square$

The above observations suggest now the following organization of typical adaptive refinement cycle

$$solve \to estimate \to refine$$

for the numerical solution of the system (2.16) regarding proper refinements of the two individual spaces X_h and Y_h. As we have already seen in (3.11) the error $\|y - y_h\|_Y$ of the auxiliary variable is always bounded by the error $\|u - u_h\|_X$ of the solution. Thus, in the $solve \to estimate \to refine$ cycle we use the estimator $\eta(u_h, y_h, \mathcal{T})$ only to refine the grid of X_h. As we have argued above we only expect a stable scheme if Y_h is somewhat more refined than X_h. The simplest strategy is to generate the new grid for Y_h by refining all cells of the grid $\mathcal{T}$ of X_h a fixed number of times. In our experiments two refinement levels have always been sufficient. The following more sophisticated variant is suggested by Lemma 3.1. After each refinement of X_h one sets first $Y_h = X_h$. Surely this will not give rise to a stable scheme. Now one can use an inner $solve \to estimate \to refine$ cycle to refine the grid Y_h until the condition (3.21) is satisfied. The complete algorithm 1 is outlined in Algorithm 1.

Algorithm 1 Adaptive stabilization

```
Choose initial spaces X_h, Y_h.
Choose the stability parameter δ and the error bound ϵ.
while η(u_h, y_h, T_h^Y) + osc(f, T_h^Y) ≥ ϵ^2 do
    Compute u_h and y_h by (2.16)
    Compute the estimators η(u_h, y_h, T_h^Y).
    Refine T_h^X by bulk chasing.
    Set T_h^X = T_h^Y.
    while Condition (3.21) not ture do
        Compute u_h and y_h by (2.16)
        Estimate ‖R_Y^{-1}(Au_h − f) − y_h‖_Y.
        Refine T_h^Y by bulk chasing.
    end while
end while
```

3.4 Some Experiments

In this section a few numerical experiments for Algorithm 1 are presented. As a simple test problem we consider

$$-\epsilon\Delta u + \begin{pmatrix}1\\1\end{pmatrix} u = 1 \quad \text{in } \Omega = (0,1)^2,\ u = 0 \text{ on } \partial\Omega, \tag{3.25}$$

for $\epsilon = 10^{-5}$. We prescribe Y as in the setting (2.8) and (2.10) which leads to the according infinite dimensional block system (2.16) or rather (2.22). We apply Algorithm 1 of the last section where we use quadrilateral grids with bilinear finite elements for X_h and Y_h.

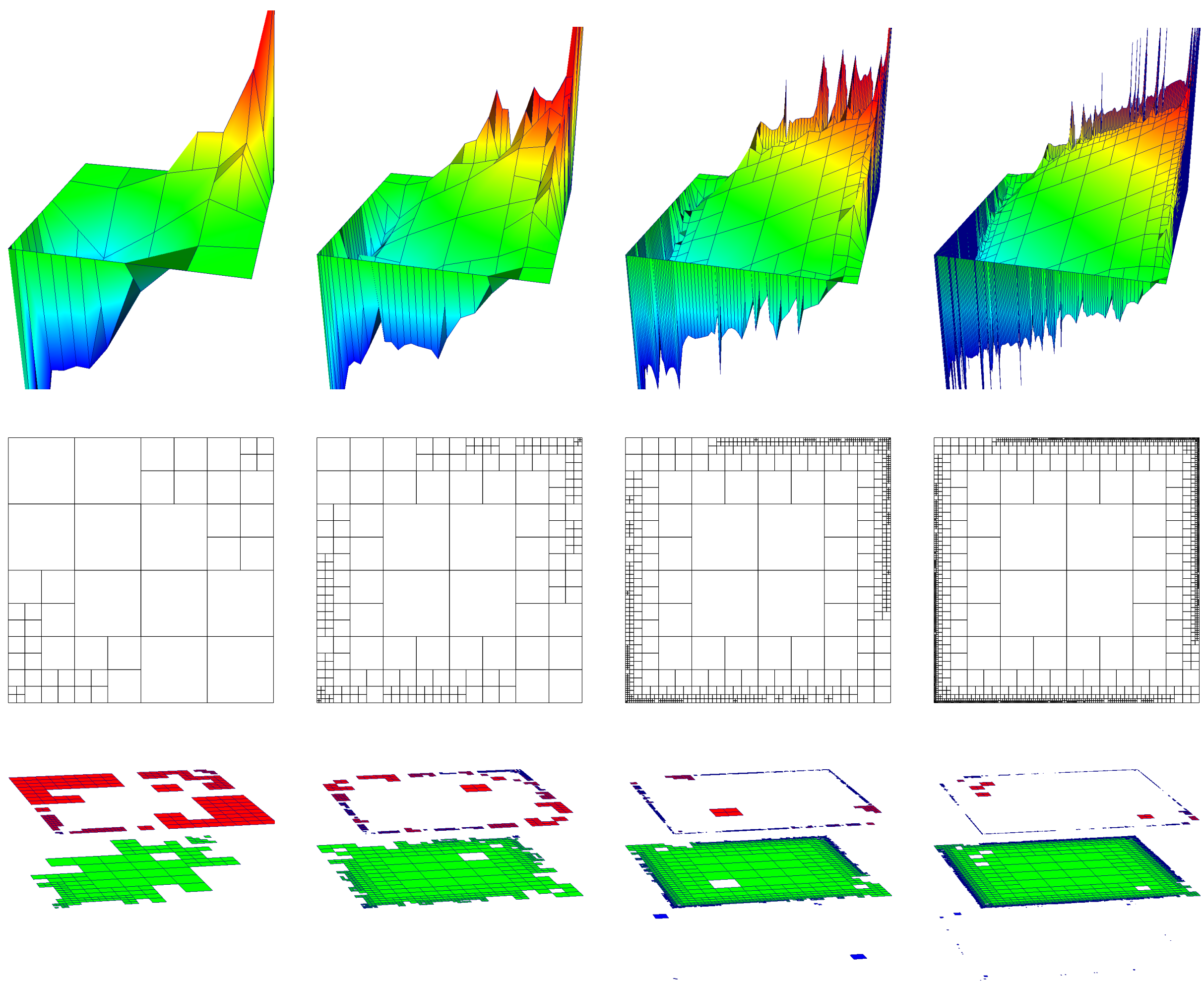


Figure 3.1: These plots show every third adaptive cycle of Algorithm 1 for $\epsilon = 10^{-5}$ starting from cycle 5 for test problem (3.25). The first row depicts the finite element solution u_h. The second row shows the corresponding grids of u_h. The third row shows the relative refinements of the spaces X_h and Y_h. Here the individual planes correspond to one additional refinement level of Y_h starting from zero additional refinements.

The approximate solutions shown in Figure 3.1 approximately fulfill the partial differential equation in the interior of Ω but miss the boundary conditions completely by introducing a second boundary layer at the inflow part of the boundary. This effect is analyzed in more detail in the next section.

4 A Closer Look at the Norms

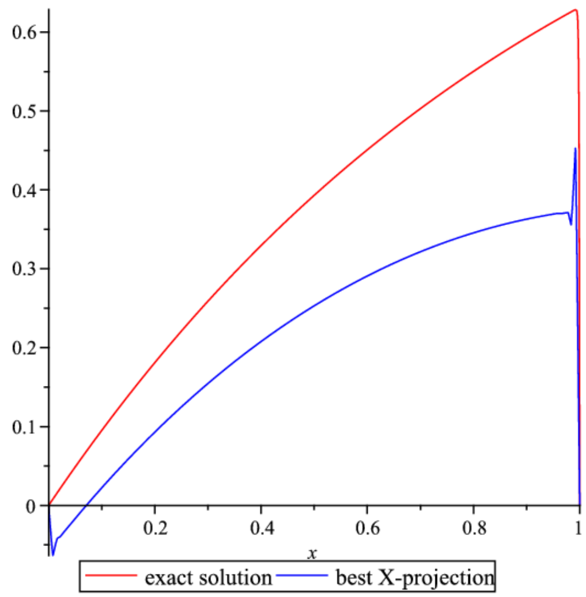
In this section we first analyze in more detail why we see the shift artifacts in the numerical examples of Section 3.4. Then in Section 4.2 we give a motivation for some possible remedies which are discussed in Sections 4.3 and 4.4. Finally some examples for the estimation of drag and lift coefficients are given in Section 4.5.

4.1 Analysis of a 1d model problem

In spite of tight lower and upper a posteriori bounds the standard refinement strategies do apparently not necessarily imply a constant error reduction per step. To understand this better we consider a simple $1D$-example shown in Figure 4.1.

$$-10^{-3}u'' + u' + u = 1, \quad \text{on } (0,1), \quad u(0) = u(1) = 0. \tag{4.1}$$

As it will be seen, the issue we are facing is actually not directly related to specific adaptive refinements.



Moreover, since R_Y^{-1} can be applied directly in this case we choose $Y_h = Y$. In view of the remarks from Section 2.3, this ensures stability of a corresponding discretization for any choice of X_h. For simplicity we choose X_h as the space of piecewise linear finite elements on an equidistant mesh of mesh size $h = 2^{-7}$. Figure 4.1 displays an approximate solution, which is the numerical evaluation of the exact X-orthogonal projection of the exact solution to X_h. Perhaps two phenomena look somewhat surprising, namely the occurrence of an oscillation at the inflow boundary and a distinct downward shift of the approximate solution. The implications on the approximation error become even more pronounced for the analogous example with smaller diffusion

$$-10^{-5}u'' + u' + u = 1, \text{on } (0,1), \quad u(0) = u(1) = 0.$$

Again an exact computation of the Y'-scalar product gives rise to errors in the X-norm displayed in Table 1 showing that there is essentially no error reduction for the first ten (uniform) refinement steps in the X-norm. Of course, ten steps of *local* refinements cannot do better either.

The reasons for this behavior seem to be the following. On one hand, the error is very much concentrated in the layer region which is hardly affected as long as the layer is not resolved. This explains the poor error reduction for the grids under consideration. To understand the strange downward shift it is instructive to look at the following example

$$-\epsilon u'' + bu' = f \quad \text{on } (0,1), \quad u(0) = u(1) = 0, \tag{4.2}$$

#cells	$\|u-u_h\|_X$	$\frac{\|u-u_h\|_X}{\|u-u_H\|_X}$	$\|u-P_Y u\|_Y$
4	0.960		0.478
8	0.958	0.998	0.463
16	0.957	0.999	0.455
32	0.957	1.000	0.451
64	0.956	0.999	0.449
128	0.954	0.998	0.447
256	0.949	0.995	0.446
512	0.939	0.989	0.445

Table 1: Error in the X-norm. Here u_h and u_H are the best X-projections where the grid of u_h has one additional refinement level to the one of u_H. Additionally P_Y is the best Y-approximation to the discrete space of u_h.

where we assume for simplicity that ϵ and b are constant. Straightforward calculations show that in this case, for $\|v\|_Y^2 = \epsilon\|v'\|_{L_2(0,1)}^2$, one has

$$\|v\|_X^2 = \epsilon\|v'\|_{L_2(0,1)}^2 + \frac{b^2}{\epsilon}\inf_{c\in\mathbb{R}}\|v-c\|_{L_2(0,1)}^2. \tag{4.3}$$

Obviously, the X-norm hardly sees any constant shift of the solution because constants in the heavily (for small ϵ) emphasized L_2-term are factored out. So apparently the projection chooses a shift that reduces the layer errors by trading a single high layer against two lower layers.

Thus, it is the particular nature of the norm $\|\cdot\|_X$ causes the shift effect observed above together with the difficulty of quickly resolving the layer which seems to be necessary for error reduction with respect to this norm. Also it is clear that the shift effect is stronger when the equation does not involve any zero order term. Furthermore it seems that also the error of the best Y-projection is not reduced by a reasonable fixed constant so that this error does not become small on an affordable grid which does not resolve yet the layers. Note that this Y-norm is also contained in most other norms commonly used in connection with convection-diffusion problems as for example the norms of the SUPG-scheme. Thus in the given example such schemes might admit an error reduction but by table 1 one cannot expect that the error becomes smaller that 0.44 for the given resolutions.

4.2 Modifying the boundary conditions

One possible (semi-)remedy is motivated by searching for pairs X, Y that allow one to stably pass the viscosity ϵ to zero. In the limit case $\epsilon = 0$ one has to split the boundary $\Gamma = \partial\Omega$ into the pieces:

$$\begin{aligned} \Gamma^+ &= \{x \in \Gamma : b\cdot n \geq 0\} \\ \Gamma^- &= \{x \in \Gamma : b\cdot n < 0\} \end{aligned} \tag{4.4}$$

where Γ^- is the inflow boundary and Γ^+ stands for the outflow boundary (containing possibly characteristic boundary portions). Since the limiting PDE is a first order equation one can only prescribe boundary conditions on the inflow boundary Γ^- to obtain a well posed problem. Depending on the variational formulations there seems to be two natural ways to impose these boundary conditions which are in some sense dual to each other. Here we only give a short motivation. The details are given in [8]. To get a weak formulation we can multiply the reduced equation $b\cdot\nabla u + cu = f$ with a test function to get

$$\langle b\cdot\nabla u, v\rangle + \langle cu, v\rangle = \langle f, v\rangle \tag{4.5}$$

In this form it is natural to build for example zero boundary conditions into the spaces for u and v. For convection-diffusion problems this gives rise to modified outflow boundary conditions which are discussed in Section 4.4.

An alternative variational setting for the reduced problem is obtained if we apply integration by parts to the problem (4.5) which gives

$$\langle u, b\nabla v\rangle + \langle u, c - \operatorname{div}(bu)\rangle = \langle f, v\rangle + \int_{\partial\Omega} nbuv$$

To treat the boundary integral we can plug in the inflow boundary condition $u = u_0$ on Γ_-. Then we are left with a boundary integral on Γ_+. This one can be treated by imposing zero boundary conditions on the outflow for the test function v. That this strategy indeed gives a well-posed problem and additional details can be found in [8]. Some modified boundary conditions for the convection-diffusion problem based on this variational formulation are treated in Section 4.3

4.3 Modifying the inflow boundary

Because the boundary layers are induced by the different boundary conditions for the convection-diffusion problem and the reduced problem and the scheme (2.24) does not find the correct layers in this section we want to treat the boundary conditions in a way that is more related to the reduced problem. Namely, as motivated in the last section we change the test space Y to $Y^+ = \{y \in H^1 : u_{\Gamma_+} = 0\}$ where the boundary condition is of course in the trace sense. Because u should still have zero boundary conditions we now have to many test functions for a well-posed system. To this end we split the discrete space $Y_h^+ \subset Y^+$ into $Y_h^+ \oplus Y_h^c$. To get a suitable variational formulation first recall (2.16) which is

$$\begin{aligned} \langle y, Av\rangle &= 0 && \forall\ v \in X \\ \langle Au, z\rangle - \langle R_Y y, z\rangle &= \langle f, z\rangle && \forall\ z \in Y. \end{aligned}$$

What is missing are the tests for test functions in Y_h^c in the second row. To find some consistent conditions note that for the true solution u we want $\langle Au, z\rangle = \langle f, z\rangle$ for all $z \in Y^+$. However there are two difficulties: firstly $\langle f, z\rangle$ might not be well defined for all $z \in Y^c$ and secondly this variational problem is overdetermined. Now for this ideal situation the second row of the block system reads $\langle R_Y y, z\rangle = 0$ for $y \in Y^c$. Thus we augment the block system with this condition to get

$$\begin{aligned} \langle y_h^+, Av_h^+\rangle &= 0 && \forall v_h^+ \in X_h \\ \langle Au_h^+, z_h\rangle - \langle R_Y y_h^+, z_h\rangle &= \langle f, z_h\rangle && \forall z_h \in Y_h \\ -\langle R_Y y_h^+, z_h^+\rangle &= 0 && \forall z_h^+ \in Y_h^c. \end{aligned} \tag{4.6}$$

which we have written immediately in discrete form. Note that for the true solution u of the original variational problem $\langle Au, z\rangle = \langle f, z\rangle$ for all $z \in Y$ the second and third block row give $\langle R_h y_h^+, z_h^+\rangle = 0$ which yields $y_h^+ = 0$ so that this system is consistent with the original problem.

In the following we want to analyze this modified scheme in the general framework of Section 2. Because the test space has to many test functions the operator $A : X \to (Y^+)'$ is no longer an isomorphism. However one easily gets this property if we also treat an enlarged space $X^+ = \{u \in H^1(\Omega) : u_{\Gamma_+} = 0\}$ for the solution u. In this situation we have essentially changed the boundary conditions for u to Neumann boundary conditions on the inflow and thus we find by standard theory that $A : X^+ \to Y^+$ is an isomorphism.

Before we proceed in analyzing this setting note that for problem (4.2) the X-norm as defined in (2.4) with the modified boundary conditions is

$$\|v\|_X^2 = \epsilon\|v'\|_{L_2(0,1)}^2 + \frac{b^2}{\epsilon}\|v\|_{L_2(0,1)}^2.$$

In contrast to 4.3 shifting the solution by a constant is now strongly penalized by $\|\cdot\|_X$.

For the new operator $A : X^+ \to Y^+$ we now have to face two obstructions. Firstly u might completely miss the zero boundary conditions on the inflow. Secondly originally we only have data in $H^{-1}(\Omega)$. So we have to assign a meaning to $\langle f, y^c\rangle$ for some $y^c \in Y^c$.

Both problems can be overcome by defining a suitable injection $E : Y' \to (Y^+)'$ which maps the original data to some new data for the modified system

$$Au^+ = f^+ \tag{4.7}$$

with $u^+ \in X^+$ and $f^+ = Ef$.

One general construction for such an extension is as follows: assume $a_e : H^1(\Omega) \times H^1(\Omega) \to \mathbb{R}$ is a continuous and $H_0^1(\Omega)$-elliptic bilinear form. This induces an isomorphism $A_e : H_0^1(\Omega) \to H^{-1}$. In this way we can easily define the extension as $\langle Ef, v\rangle = a_e(A_e^{-1} f, v)$. Here we make use of the fact that although for the inversion of A_e we made use of the $H_0^1(\Omega)$-ellipticity, i.e. we used the zero boundary conditions, the bilinear form a_e is still well defined for all functions in $H^1(\Omega)$.

Now we have to find a bilinear form a_e. One simple choice is $a_e = a$. This implies that by construction the solution of (4.7) has zero boundary conditions everywhere even though these conditions are not implied by the operator $A : X^+ \to Y^+$ itself. Surely this infinite dimensional setting is only a tautology but as we will see it leads to the modified discrete system (4.6). To this end, using finite dimensional subspaces $X_h^+ \subset X^+$ and $Y_h^+ \subset Y^+$ as trial and test spaces in the discrete block system (2.23), the corresponding version for the modified boundary conditions reads: find $u_h^+ \in X_h^+$ and $y_h^+ \in Y_h^+$ s.t.

$$\begin{aligned} \langle y_h^+, A v_h^+\rangle &= 0 && \forall v_h^+ \in X_h^+ \\ \langle A u_h^+, z_h^+\rangle - a_Y(y_h^+, z_h^+) &= \langle f^+, z_h^+\rangle && \forall z_h^+ \in Y_h^+. \end{aligned}$$

Since we know that the correct solution u^+ for our injection $f^+ = Ef$ has trace zero on the whole boundary, we choose $X_h^+ = X_h \subset X$ that enforces zero boundary conditions for the numerical solution as well. As above we decompose $Y_h^+ = Y_h \oplus Y_h^c$ into a direct sum of two spaces. For $z_h \in Y_h$ by construction we have

$$\langle f^+, z_h\rangle = \langle f, z_h\rangle.$$

For the test functions in Y_h^c we use the approximation

$$\langle A u_h^+, z_h^+\rangle - \langle f^+, z_h^+\rangle \approx \langle A u^+, z_h^+\rangle - \langle f^+, z_h^+\rangle = 0$$

Finally we get the modified block system (4.6): find $u_h^+ \in X_h$ and $y_h^+ \in Y_h^+$ s.t.

$$\begin{aligned} \langle y_h^+, A v_h^+\rangle &= 0 && \forall v_h^+ \in X_h \\ \langle A u_h^+, z_h\rangle - a_Y(y_h^+, z_h) &= \langle f, z_h\rangle && \forall z_h \in Y_h \\ -a_Y(y_h^+, z_h^+) &= 0 && \forall z_h^+ \in Y_h^c. \end{aligned}$$

Figure 4.2 shows the results of these modifications applied to the problem and numerical scheme described in Section 3.4 and Section 3.3, respectively. We see that the grid is well adapted to the solution and this time there is no unphysical additional layer. Nevertheless, at an early refinement stage the oscillations at the layer still look unnecessarily strong. However, one should note that they disappear as soon as the refinement level permits a layer resolution. In fact, it seems that the intermediate oscillations do not mislead the refinement process. In this sense the scheme is stable. Moreover, the particular form (2.12) suggests that oscillations in transversal direction to the streamlines are primarily penalized in dual norm $\|\cdot\|_{Y'}$ that permits them to some extent. In order to address the behavior prior to the layer resolution we shall explore next further alternatives.

4.4 Modifying the outflow boundary

The poor resolution of very narrow boundary layers seems to remain as a serious obstruction to error decay even when the space Y_h is chosen in such a way that a discrete inf-sup condition holds as described in Section 2.3. However, if we do not impose any Dirichlet boundary conditions on the outflow boundary the corresponding solution u^n of this new problem will not have a boundary layer and thus one expects a much smaller approximation error in that case. One example of such a boundary condition would be a Neumann boundary condition on the outflow. To see how this could be exploited, denoting by u^d the solution with zero boundary conditions on the whole boundary, we can decompose it as

$$u^d = u^n + k, \tag{4.8}$$

where u^n is the solution for Neumann conditions on the outflow boundary. From the weak formulation of both boundary value problems we obtain

$$a(k, v) = a(u^d - u^n, v) = 0 \qquad \text{for all } v \in Y. \tag{4.9}$$

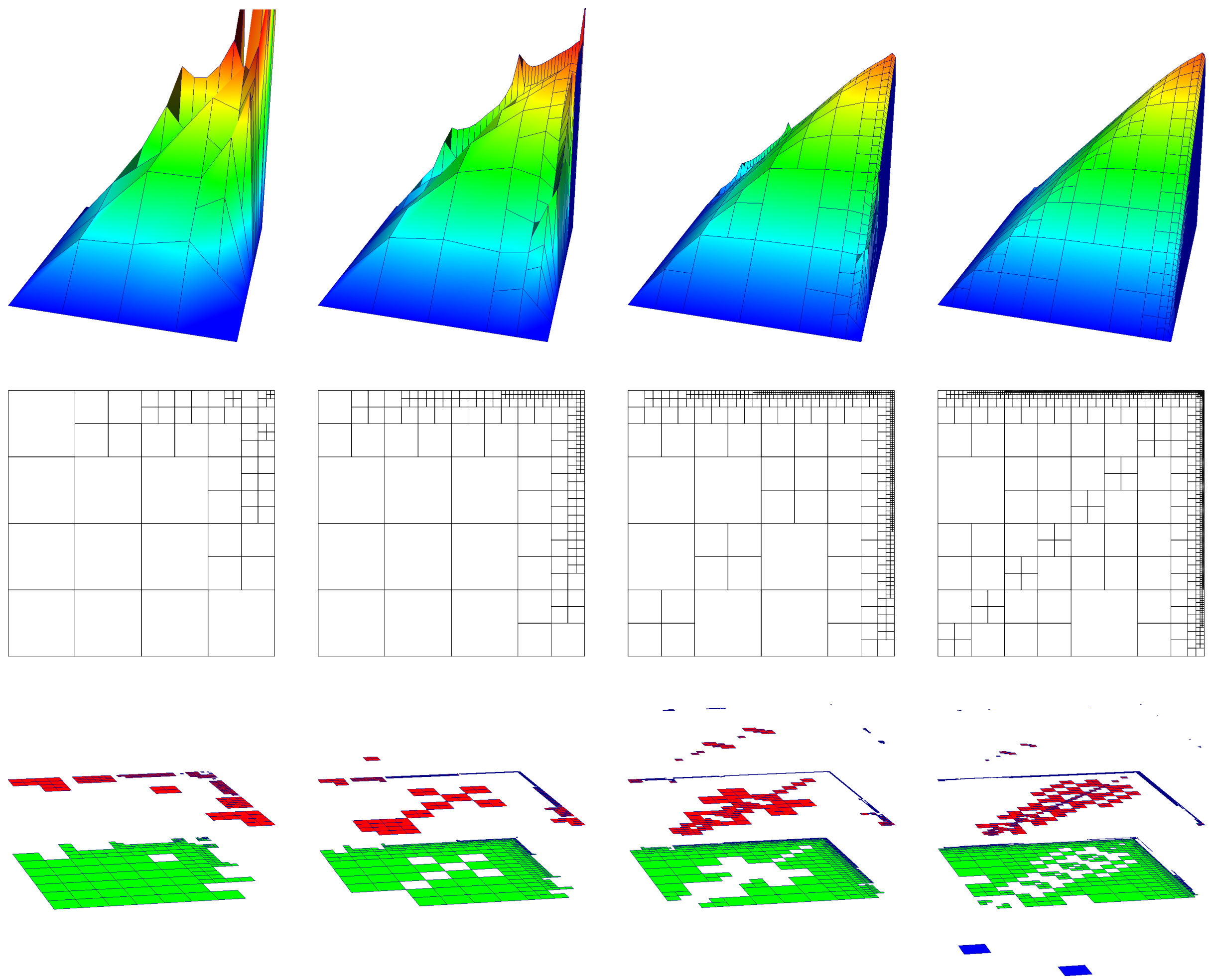


Figure 4.2: These pictures show every third adaptive cycle of Algorithm 1 with the above modification of the inflow boundary condition and $\epsilon = 5 \cdot 10^{-2}$, starting from cycle 5 for the test problem (3.25). The first row depicts the finite element solution u_h. The second row displays the corresponding grids of u_h. The third row shows the relative refinements of the spaces X_h and Y_h. Here the individual planes correspond to one additional refinement level of Y_h starting from zero additional refinements.

In order to specify the meaning of the component k recall that Γ^- and Γ^+ denote the inflow and outflow boundary, respectively, as defined in (4.4). Let X^- denote the closure of smooth functions on Ω with respect to $\|\cdot\|_X$ which vanish on Γ^-. Now define the operator $\bar{A} : X^- \to Y'$ by

$$\langle \bar{A}u, v \rangle = a(u, v), \quad \forall\, v \in Y$$

where $u \in X^-$. With this definition we see form (4.9) that k is in the kernel of $\bar{A}$. Recall that generally we expect to get good approximations of u^n but only lousy approximations of u^d. It follows from (4.8) that it is generally hard to approximate k. We try next to exploit this for a numerical scheme: we try to find a good approximation to u^n and disregard k which we cannot approximate anyway. This suggests considering the problem

$$\|\bar{A}u - f\|_{Y'} \to \min \tag{4.10}$$

for $u \in X^-$. Obviously u^d and u^n are solutions of this problem. But what happens for the discrete optimization problem? We have already seen that the expression we minimize is equivalent to the error of the solution in the X-norm. Assume $u_h^d \in X_h$ is a solution of the optimization problem with zero boundary conditions and u^n is one solution with nonzero boundary conditions on the outflow. As seen in table 1 we expect $\|\bar{A}u_h^d - f\|_{Y'}$ to be large while, as we argued above, there are solutions u^n for which

$\|\bar{A}u_h^n - f\|_{Y'}$ is small. Thus a discretization of this kind automatically approximates one element u^n that can be approximated best. Furthermore we know that $u^n = u^d - k$ for some k in the kernel of $\bar{A}$. To better understand the difference of u^n and u^d we characterize the kernel of $\bar{A}$.

Proposition 4.1. *Let*

$$K := \{u \in X^- : \langle \bar{A}u, v\rangle = 0 \text{ for all } v \in Y\}.$$

Then K is nontrivial and is isomorphic to $H^{1/2}(\Gamma^+)$.

Proof: For any $0 \neq g \in H^{1/2}(\Gamma^+)$ let

$$\|g\|_{1/2,\Gamma^+} := \inf_{w\in X:\gamma w = g} \|w\|_X,$$

where γ is the trace operator on Γ^+. Let γ^* denote the adjoint of γ that produces the minimizer in the above definition of the trace norm. Now define $r \in Y'$ by

$$\langle r, v\rangle = \langle \bar{A}(\gamma^* g), v\rangle, \quad \forall\, v \in Y. \tag{4.11}$$

Then there exists a unique $w \in X$ such that

$$a(w, v) = -\langle r, v\rangle, \quad \forall\, v \in Y. \tag{4.12}$$

Clearly

$$u_0 := w + \gamma^* g \in X^- \tag{4.13}$$

and

$$\langle \bar{A}u_0, v\rangle = a(w, v) + a(\gamma^* g, v) = -\langle r, v\rangle + \langle r, v\rangle = 0,$$

so that $u_0 \in K$. Since by (4.12) that

$$w = -A^{-1}r = -A^{-1}\bar{A}\gamma^* g, \tag{4.14}$$

we have

$$u_0 = Lg := (I - A^{-1}\bar{A})\gamma^* g. \tag{4.15}$$

By definition $A^{-1}\bar{A}\gamma^* g \in X$ so that $\gamma u_0 = g$. Thus for $g \neq 0$ one has $Lg \neq 0$ so that L is injective. L is also surjective, since for $u_0 \in K \setminus \{0\}$ one must have $\gamma u_0 =: g \neq 0$ since otherwise $u_0 \in X$ and $0 = \bar{A}u_0 = Au_0$ implies $u_0 = 0$. Moreover,

$$\begin{aligned} \|Lg\|_X &\leq (1 + \|A^{-1}\bar{A}\|_{X^-\to X})\|\gamma^* g\|_{X^-} \\ &\lesssim (1 + \|A^{-1}\|_{Y'\to X}\|\bar{A}\|_{X^-\to Y'})\|g\|_{1/2,\Gamma^+}, \end{aligned} \tag{4.16}$$

i.e. L is bounded. By the Open-Mapping-Theorem its inverse is also bounded. □

Because $\bar{A}$ has a nontrivial kernel the minimization problem (4.10) is generally not uniquely solvable. Furthermore we have changed our initial convection-diffusion problem by changing the outflow boundary condition. We can solve both problems by adding another penalty term for the outflow boundary: find $u \in X^-$ s.t.

$$\|\bar{A}u - f\|_{Y'}^2 + \mu\|\gamma u\|_{1/2,\Gamma^+}^2 \to \min. \tag{4.17}$$

Because γu is isomorphic to the kernel of $\bar{A}$ which we generally cannot approximate very well we give μ a small weight like $\mu \sim \epsilon$. We get the following relation to the original convection-diffusion problem with zero boundary conditions:

Proposition 4.2. *One has $a(u, v) = \langle f, v\rangle$ for all $v \in Y$ if and only if one has for any positive μ*

$$u = \operatorname{argmin}_{v\in X^-}\big\{\|\bar{A}v - f\|_{Y'}^2 + \mu\|\gamma v\|_{1/2,\Gamma^+}^2\big\}. \tag{4.18}$$

Proof: Given $f \in Y'$ let $\tilde{u} \in X$ denote the solution of $a(\tilde{u}, v) = \langle f, v\rangle$, $v \in Y$ so that

$$\bar{A}\tilde{u} = A\tilde{u} = f \in Y', \quad \gamma\tilde{u} = 0.$$

Hence

$$0 = \|\bar{A}\tilde{u} - f\|_{Y'}^2 + \mu\|\gamma\tilde{u}\|_{1/2,\Gamma^+}^2$$

and $\tilde{u}$ is a minimizer of $\left\{\|\bar{A}v - f\|_{Y'}^2 + \mu\|\gamma v\|_{1/2,\Gamma^+}^2\right\}$. Conversely, since the minimal value of the functional is zero, a minimizer u must have a zero trace on Γ^+ and must satisfy $\bar{A}u = Au = f$, which completes the proof. $\square$

As for the normal equations the optimization problem (4.17) is equivalent to: find $u \in X^-$ s.t.

$$\langle \bar{A}u - f, \bar{A}v\rangle_{Y'} + \langle u, v\rangle_{1/2,\Gamma^+} = 0 \qquad \text{for all } v \in Y.$$

In analogy to the derivation of (2.16) this is equivalent to the block system: find $u \in X^-$ and $y \in Y^-$ s.t.

$$\begin{aligned} \mu\langle u, v\rangle_{1/2,\Gamma^+} + \langle y, Av\rangle &= 0 & \forall v \in X^- \\ \langle Au, z\rangle - a_Y(y, z) &= \langle f, z\rangle & \forall z \in Y^-. \end{aligned}$$

where Y^- is define analogously to X^- with the X-norm replaced by the Y-norm. For a first numerical test we replace the $\|\cdot\|_{1/2,\Gamma^+}$-norm by a weighted $\|\cdot\|_{L_2(\Gamma^+)}$-norm. Figure 4.3 shows the results for the problem and scheme described in Section 3.4 with the modifications of the present section. For the viscosity $\epsilon = 10^{-5}$ Figure 4.3 displays refinements that cannot resolve the layer. Yet, this time neither oscillations nor refinements near the layer region occur while away from the layer region excellent accuracy is observed. It will be shown in the next section that one can still derive accurate information about gradients in the layer.

Figure 4.4 shows for $\epsilon = 5 * 10^{-3}$ what happens when the local refinements permit to resolve the layer. Already for a the simple choice $\mu = \epsilon$ of the penalty parameter the resolution of the layer is triggered automatically and finally resolves it.

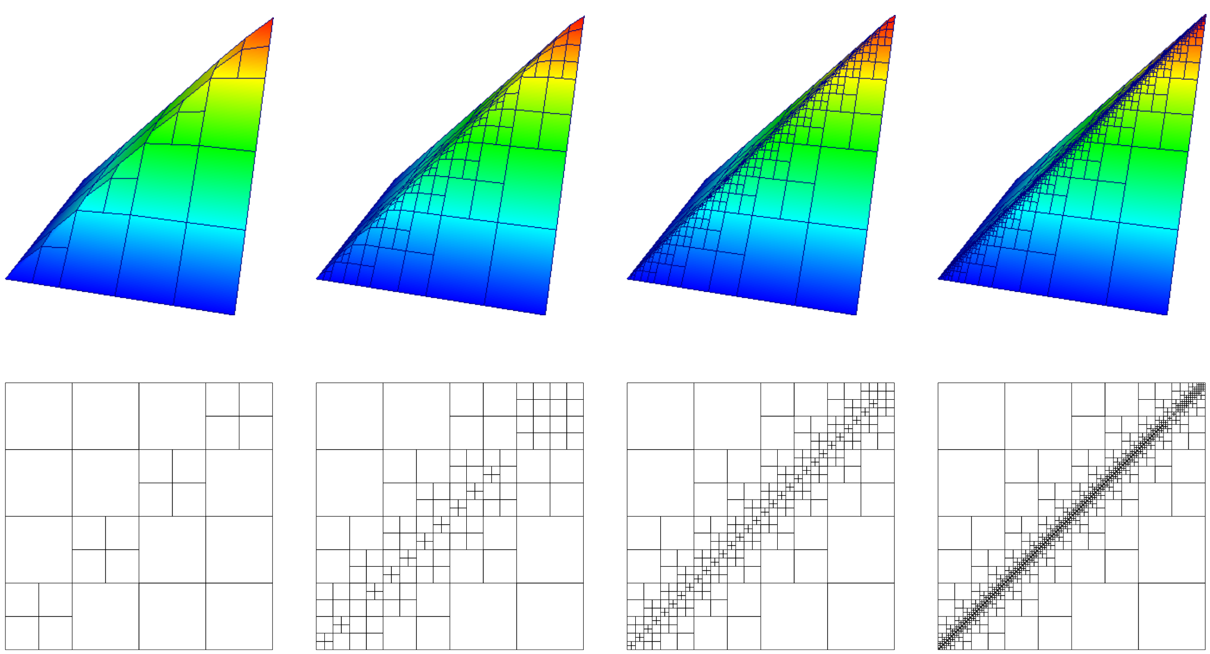


Figure 4.3: These pictures show every fourth adaptive cycle of Algorithm 1 with modification at the outflow boundary starting from cycle 3 for the test problem (3.25) with $\epsilon = 10^{-5}$. The first row depicts the finite element solution u_h. The second row displays the corresponding grids of u_h.

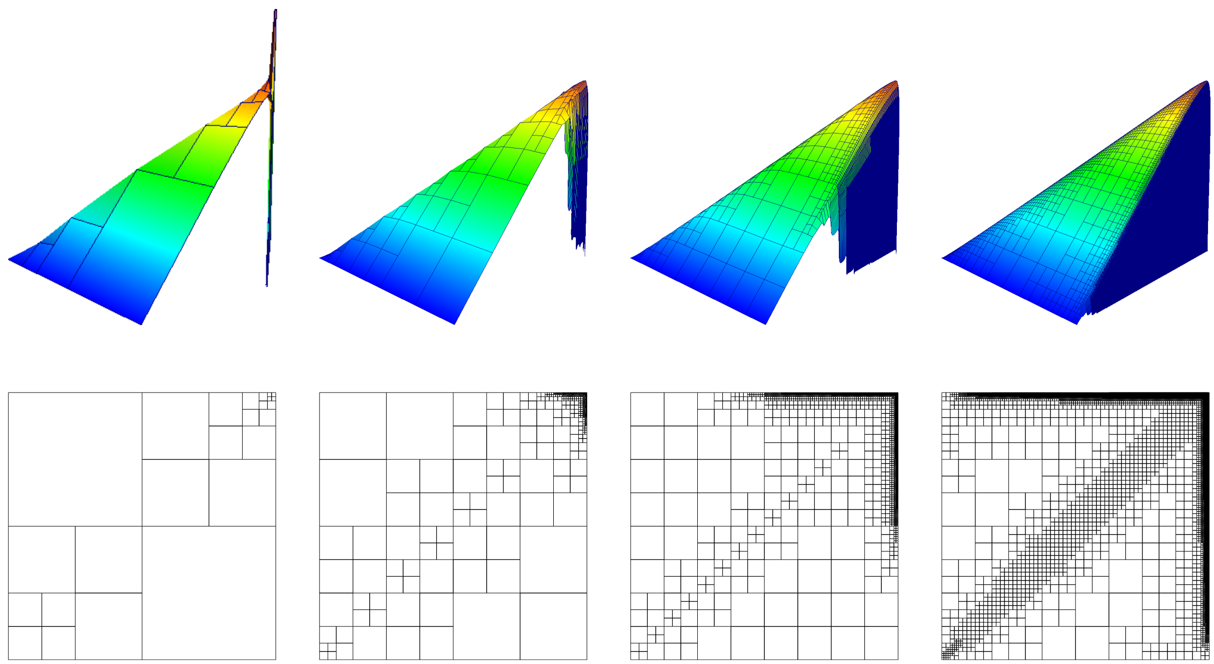


Figure 4.4: These pictures show every fourth adaptive cycle of Algorithm 1 with modification at the outflow boundary starting from cycle 3 for the test problem (3.25) with $\epsilon = 5 \cdot 10^{-3}$. The first row depicts the finite element solution u_h. The second row displays the corresponding grids of u_h.

4.5 Drag and lift

Of course, modifying the outflow boundary conditions, changes the problem. In this section we argue though that we can still retain some interesting information on the layers. In particular, we are interested in functionals of the form

$$l(u) = \epsilon \int_{\partial\Omega} n \nabla u w, \tag{4.19}$$

where $w \in H^{1/2}(\partial\Omega)$ and n is the outward normal vector. These functionals are a simple model for physical quantities like drag and lift coefficients in the case of fluid dynamics. It seems that, due to changing outflow boundary conditions in the last section, one has any useful information on such functionals in the layer regions. However, in this section we offer a heuristic argument for still estimating such functionals from the solutions of the last section. For convenience we assume that the convection-diffusion problem is given in the alternative form

$$Au = -\epsilon \Delta u + \operatorname{div}(\tilde{b} u) + \tilde{c} u.$$

and that all quantities are sufficiently smooth. Recall that in Section 4.4 in (4.8) we have split the solution as $u^d = u^n + k$ where u^d is the original solution with zero boundary conditions everywhere, u^n is a solution with free outflow boundary conditions and k is in the kernel of A. Thus, k is specified by the equations

$$Ak = 0 \qquad\qquad k|_{\partial\Omega} = -u^n|_{\partial\Omega}.$$

To evaluate the functional l form (4.19) the same splitting provides

$$l(u^d) = l(u^n) + l(k).$$

Here u^n is the solution one has computed with the free outflow boundary conditions so $l(u^n)$ can be easily computed. Since generally we cannot resolve the boundary layer we cannot expect to compute $l(k)$ directly. However, the following heuristics should offer a good estimator for this quantity. To this end, consider the harmonic extension w_e of w given by

$$\begin{aligned} -\Delta w_e &= 0 && \text{in } \Omega \\ w_e &= w && \text{on } \partial\Omega. \end{aligned}$$

In the following we do not distinguish between w and its extension w_e and always write w. Then we get

$$l(k) = \epsilon \int\limits_{\partial\Omega} n\nabla k w = -\epsilon \int\limits_{\Omega} \operatorname{div}(\nabla k w) = -\epsilon \int\limits_{\Omega} \Delta k w - \epsilon \int\limits_{\Omega} \nabla k \nabla w = -S_1 - S_2.$$

Since, by definition, $\Delta w = 0$ we obtain

$$S_2 = \epsilon \int\limits_{\Omega} \nabla k \nabla w = -\epsilon \int\limits_{\Omega} k\Delta w + \epsilon \int\limits_{\partial\Omega} nk\nabla w = \epsilon \int\limits_{\partial\Omega} nk\nabla w.$$

Next we treat S_1. Because k is in the kernel of A we have $\epsilon\Delta k = \operatorname{div}(\tilde{b}k) + \tilde{c}k$ which yields

$$\begin{aligned} S_1 &= \int\limits_{\Omega} (\operatorname{div}(\tilde{b}k) + \tilde{c}k)w \\ &= \int\limits_{\Omega} \operatorname{div}(\tilde{b}kw) - \int\limits_{\Omega} \tilde{b}k\nabla w + \int\limits_{\Omega} \tilde{c}kw \\ &= -\int\limits_{\partial\Omega} n\tilde{b}kw - \int\limits_{\Omega} \tilde{b}k\nabla w + \int\limits_{\Omega} \tilde{c}kw. \end{aligned}$$

Putting the pieces together, we obtain

$$\begin{aligned} l(u^d) &= \epsilon \int\limits_{\partial\Omega} n\nabla u^d w \\ &= \epsilon \int\limits_{\partial\Omega} n\nabla u^n w + \int\limits_{\partial\Omega} n\tilde{b}kw + \int\limits_{\Omega} \tilde{b}k\nabla w - \int\limits_{\Omega} \tilde{c}kw - \epsilon \int\limits_{\partial\Omega} nk\nabla w. \end{aligned}$$

Since k agrees with $-u^n$ on $\partial\Omega$ we can now compute the boundary integrals without knowing k explicitly. For the remaining two summands one could argue as follows: since $Ak = 0$ we expect k to be close to zero except for some layer regions of width ϵ. Thus, neglecting these terms, we obtain the following approximation to l, given by

$$l_h([u,\sigma]) = \epsilon \int\limits_{\partial\Omega} n\nabla u^n w - \int\limits_{\partial\Omega} n\tilde{b}u^n w + \epsilon \int\limits_{\partial\Omega} nu^n \nabla w.$$

This estimate might also be interesting form the perspective of turbulence modeling. There one of the leading questions is to what extend one can extract information on the unresolved scales form the resolved ones. In a much simpler model this is exactly what this estimator is doing: it can estimate the gradient in the layer region without resolving it.

As a simple example we use the model problem (3.25) with rigth hand side

$$u(x,y) = xy(1 - \exp\left(-\frac{1-x}{\epsilon}\right)\left(1 - \exp\left(-\frac{1-y}{\epsilon}\right)\right).$$

Table 4.5 shows the exact error of the estimation of the boundary integral $\int\limits_{\partial\Omega} n\nabla u$.

5 Conclusion

We have proposed and analyzed adaptive numerical schemes for convection-diffusion problems based on a uniformly well-posed variational formulation that gives rise to a mapping property. The main example are the norms used by Sangalli and Verfürth in [17, 20]. To construct a numerical scheme which finds near best approximations in these norms we introduce an auxiliary variable whose resolution compared to the resolution of the solution itself determines the amount of stabilization. We have developed a-posterior conditions that ensure a sufficient resolution of this auxiliary variable adaptively. Together

adaptive cycle	$\epsilon = 5 \cdot 10^{-2}$	$\epsilon = \cdot 10^{-5}$	adaptive cycle	$\epsilon = 5 \cdot 10^{-2}$	$\epsilon = \cdot 10^{-5}$
1	4.99e-01	9.51e-06	8	1.49e-01	9.95e-06
2	2.79e-01	9.99e-06	9	1.14e-01	9.91e-06
3	2.77e-01	9.99e-06	10	8.80e-02	9.89e-06
4	2.28e-01	9.98e-06	11	7.70e-02	9.88e-06
5	2.40e-01	9.98e-06	12	5.21e-02	9.87e-06
6	2.23e-01	9.97e-06	13	4.33e-02	9.78e-06
7	1.71e-01	9.96e-06	14	3.39e-02	9.69e-06

with the a-posteriori error estimators developed in this paper this yields a new adaptive scheme for convection-diffusion problems where, in contrast to earlier work, the norms used in the derivation of the error indicators agree with the norms in which accuracy is measured.

However, first numerical tests reveal the fact that the chosen norms tend to create - perhaps unexpected - artifacts in the numerical solution. It is perhaps somewhat surprising, at least at the first glance, that a uniformly stable variational formulation together with reliable and efficient a-posteriori bounds do not automatically guarantee to error histories exhibiting a quantitatively "ideal' behavior. More involved discretizations like hp-r or anisotropic efinements would certainly diminish the observed adverse effects. We have not pursued this line here. Insetaed, to understand the principal phenomena concerning the interplay between adaptivity and stability has been a core objective of the present work. Here, two remedies have been presented. Guided by the reduced problem with vanishing viscosity, one can modify the boundary conditions either at the inflow or outflow boundary to obtain norms which do no longer give rise to those artifacts. As long as the layer is not resolved oscillations remain in the first case which, however, remain within the layer region and do not mislead further mesh refinements. In the latter case the layer is no longer present and one obtains very accurate oscillation free solutions away from the layer region before the layer is resolved. An interesting point to be explored further is that one can still gain, somewhat in the spirit of subcell resolution, very accurate information on gradients in this region as is needed, for instance, for the computation of relevant functionals of the solution such as drag and lift.

References

[1] F. Brezzi, M. Fortin, Mixed and Hybrid Finite Element Methods, Springer Series in Computational Mathematics 15, Springer-Verlag, Berlin, 1991.

[2] F. Brezzi, T.J.R. Hughes, L.D. Marini, A. Russo, E. Süli, A priori analysis of residual-free bubbles for advection-diffusion problems, SIAM J. Numer. Anal., 36(1999), 1933–1948.

[3] A. Cohen, W. Dahmen, R. DeVore, Adaptive wavelet methods II - Beyond the elliptic case, Foundations of Computational Mathematics, 2 (2002), 203–245.

[4] A. Cohen, W. Dahmen, R. DeVore, Adaptive Wavelet Schemes for Nonlinear Variational Problems, SIAM J. Numer. Anal., (5)41(2003), 1785–1823.

[5] S. Dahlke, W. Dahmen, K. Urban, Adaptive wavelet methods for saddle point problems – Convergence rates, SIAM J. Numer. Anal., 40 (No. 4) (2002), 1230–1262.

[6] A. Cohen, W. Dahmen, R. DeVore, Compressed Sensing and best k-term approximation, J. Amer. Math. Soc. 22 (2009), 211-231.

[7] A. Cohen, W. Dahmen, R. DeVore, Instance Optimal Decoding by Thresholding in Compressed Sensing, Contemporary Mathematics, 505 (2010), 1–28.

[8] W. Dahmen, C. Huang, C. Schwab, G. Welper, Adaptive Petrov-Galerkin methods for first order transport equations, in preparation.

[9] L. Demkowicz, J. Gopalakrishnan, A class of discontinuous Petrov-Galerkin methods. Part II: Optimal test functions, Preprint.

[10] L. Demkowicz, J. Gopalakrishnan, A class of discontinuous Petrov-Galerkin methods. Part III: Adaptivity. Preprint

[11] J.L. Guermond, J.T. Oden, S. Prudhomme, An interpretation of the Navier-Stokes-alpha model as a frame-indifferent Leray regularization, Physica D, 177 (2003) 23-30.

[12] J.-L. Guermond, J.T. Oden, S. Prudhomme, Mathematical perspectives on Large Eddy Simulation models for turbulent flows , J. Math. Fluid Mech., 6 (2004), 194-248.

[13] T. Hughes, G. Sangalli, Variational Multiscale Analysis: the Fine-scale Green's Function, Projection, Optimization, Localization, and Stabilized Methods, SIAM Journal of Numerical Analysis, Vol. 45, (No. 2) (2007), 539–557.

[14] V John, S Kaya, W. Layton, A two-level variational multiscale method for convection-diffusion equations, Comp. Meth. Appl. Mech. Engrg., 195 (2006), 4594-4603.

[15] S. Jokar, V. Mehrmann, M. Pfetsch, H. Yserentant, Sparse Approximate Solution of Partial Differential Equations, Preprint 498 MATHEON, DFG Research Center Mathematics for key technologies in Berlin, Feb. 2009.

[16] W. Dahmen, S. Müller, T. Schlinkmann, On an adaptive multigrid solver for convection-dominated problems, SIAM J. Scient. Comp., 23 (No 3)(2001), 781–804.

[17] G. Sangalli, Robust a-posteriori estimators for advection-diffusion-reaction problems, Math. Comput., 77 (2008), 41–70, 2008.

[18] G. Sangalli, A uniform analysis of non-symmetric and coercive linear operators, SIAM J. Math. Anal., 36 (no 6) (2005), 2033-2048.

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

[20] R. Verfürth, Robust a posteriori error estimates for stationary convection-diffusion equations. SIAM J. Numer. Anal., 43, No. 4 (2005), 1766–1782.