

A POSTERIORI ERROR ESTIMATION FOR INTERIOR PENALTY FINITE ELEMENT APPROXIMATIONS OF THE ADVECTION-REACTION EQUATION*

ERIK BURMAN†

Abstract. In this note we consider residual-based a posteriori error estimation for finite element approximations of the transport equation. For the discretization we use piecewise affine continuous or discontinuous finite elements and symmetric stabilization of interior penalty type. The lowest order discontinuous Galerkin method using piecewise constant approximation is included as a special case. The key elements in the analysis are a saturation assumption and an approximation result for interpolation between discrete spaces.

Key words. a posteriori error estimation, advection-reaction equations, finite element method, stabilization, discontinuous Galerkin, saturation assumption

AMS subject classifications. 65N30, 35A35, 35L02, 65N15

DOI. 10.1137/080733899

1. Introduction. A posteriori error estimation and adaptivity for finite element approximations of elliptic problems has been an active research field for a long time [2] and is now reaching maturity [26]. These results are fully satisfactory only in cases where the second order elliptic operator dominates. Indeed, the a posteriori error estimates that have been proposed for the elliptic problem degenerate for advection–diffusion-type problems when the advection becomes dominant. When the hyperbolic character of the problem becomes important, interior and outflow layers form, causing local large gradients in the solution. In this regime the standard Galerkin method is unstable, and various stabilized methods have been proposed.

An early attempt to derive a posteriori error estimates in the advection-dominated regime was proposed by Eriksson and Johnson in [16], using regularization and duality techniques. Improved energy norm techniques were then proposed by Verfürth in [31], where semirobust estimates were obtained. Extensions of these ideas to some different methods were proposed in [5, 23, 27, 15, 33]. However, when the advection dominates, estimates based on the ideas of [31] fail. More recently Verfürth [32] and Sangalli [29] have proposed different norms for the a posteriori error estimation that allow for robust estimators; however, these approaches are not yet completely satisfactory: in the one case the norm is too weak, and in the other the analysis is only valid in the one-dimensional case.

In this paper we will consider the extreme case of the advection-reaction problem for our a posteriori error estimates. This can be seen as the first step to robust a posteriori error estimation for advection-diffusion equations, handling the difficult advection dominated case separately, but is also an interesting problem in its own right.

The case of the advection-reaction equation is less studied than the advection-diffusion equation. This is mainly due to the fact that whereas the elliptic problem

*Received by the editors August 29, 2008; accepted for publication (in revised form) June 22, 2009; published electronically November 13, 2009.

<http://www.siam.org/journals/sinum/47-5/73389.html>

†Department of Mathematics, University of Sussex, Brighton, UK–BN1 9RF, United Kingdom (E.N.Burman@sussex.ac.uk).

has some smoothing properties, the transport equation does not. This lack of smoothing and the nonsymmetry of the problem is what makes standard techniques using coercivity suboptimal. If the error is measured in weaker norms, some results may be obtained, but for most cases these weak norms do not measure the quantities of interest and are difficult to use with adaptivity. Another possibility is to use duality techniques to derive a posteriori error estimates. However, then the solution of a continuous dual problem must be analyzed or approximated. For work using duality techniques or nonstandard norms, see Süli et al. [30, 22].

None of the above cited works, however, yields results comparable to those obtained by the a priori error analysis for the transport equation; in particular they do not use the stability properties of the method that are crucial to obtain optimal convergence order.

The aim of the present note is to exploit these stability properties of interior penalty finite element methods to obtain improved estimates. The idea is to derive a posteriori error estimates in the same weighted norms that are used for the a priori error analysis for stabilized methods, without making any assumptions on the regularity of the exact solution. When finite element spaces of globally continuous functions are used, this norm is given by

$$\|v\|_{h,0,\Xi}^2 := \|v_h\|_{L^2(\Omega)}^2 + \|\beta_n^{\frac{1}{2}} v_h\|_{L^2(\partial\Omega)}^2 + \Xi \|\tilde{h}^{\frac{1}{2}} \beta \cdot \nabla v_h\|_{L^2(\Omega)}^2,$$

where Ω denotes the computational domain with boundary $\partial\Omega$, h denotes the mesh-size, and β the transport velocity vector field. When $\Xi = 1$, this corresponds to the h -weighted graph norm and for $\Xi = 0$ it reduces to the L^2 -norm. The lack of smoothing that usually leads to a seemingly unsolvable interpolation problem is circumvented by the introduction of a saturation hypothesis that allows us to consider only the interpolation error between two discrete spaces. The a posteriori error estimates we propose are optimal both in the graph norm and the L^2 -norm, provided the saturation assumption holds. We prove global upper bounds and local lower bounds. The key ingredients of our analysis is the inf-sup stability of the interior penalty method, the saturation assumption, and an interpolation estimate between discrete spaces.

The saturation assumption has been extensively used in a posteriori error estimation; see [3, 7, 34, 4, 1]. It is typically needed where the standard techniques for a posteriori error estimation based on coercivity or smoothing fail, either due to an indefinite problem or nonconforming approximation. Of special interest in our case is the reference [7] where the saturation assumption is introduced, due to the use of the anisotropic $H(\text{div}, \Omega)$ -norm and lack of regularity of the exact solution. In a similar fashion as in [7] we introduce an anisotropic mesh-dependent norm and estimate the discrete error in this norm. In our case continuity of the bilinear form is guaranteed thanks only to the presence of the stabilizing terms that gives inf-sup control of the h -weighted streamline derivative. Therefore, the estimators that we propose herein are invalid for the standard Galerkin method, and numerical experiments show that, without stabilization, they fail to produce appropriate adaptive meshes *even in cases where the exact solution is smooth*. For elliptic problems, the saturation assumption can be proved to be equivalent with resolving the problem data on the computational mesh [13].

The analysis presented here is valid both in the case of continuous, piecewise affine approximations (using the continuous interior penalty method (CIP) [14, 11, 8]) and discontinuous, piecewise affine or constant approximations (the discontinuous Galerkin (DG) method DG0, DG1 [28, 25]). In all these cases stability is obtained

using interior penalty: in the CIP case the gradient jumps over element faces are penalized, and in the DG case, the jumps of the solution. A salient feature of the analysis in the case of discontinuous approximation is that it does not use comparison with a continuous approximation. We also use our estimate in the framework of goal-oriented a posteriori error estimation to derive a duality-based estimate that is rigorous and computable, up to the constant from the saturation assumption. The upshot here is that the residual of the finite element solution of the dual problem is used in the estimator.

2. The problem setting. Let Ω be a polygon in \mathbb{R}^2 with outer normal n . We consider the following advection-reaction equation:

Find $u : \Omega \rightarrow \mathbb{R}$ such that

$$(2.1) \quad \begin{cases} \beta \cdot \nabla u + \mu u = f, \\ u|_{\partial\Omega^-} = g, \end{cases}$$

where $f \in L^2(\Omega)$, $g \in L^2(\partial\Omega^-)$, the vector field β is chosen in $[\text{Lip}(\Omega)]^2$, and $\mu \in \mathbb{R}$, $\mu > 0$. Assume that $\mu - \frac{1}{2}\nabla \cdot \beta \geq \mu_0 > 0$ and that $\inf_{x \in \Omega} |\beta| \geq c > 0$. The inflow boundary is defined by $\partial\Omega^\pm = \{x \in \partial\Omega; \pm\beta(x) \cdot n(x) > 0\}$. Define

$$W = \{w \in L^2(\Omega); \beta \cdot \nabla w \in L^2(\Omega)\},$$

and observe that functions in W have traces in $L^2(\partial\Omega; \beta \cdot n)$. The problem (2.1) is well posed in W (see [17]).

2.1. Definitions. Let \mathcal{T}_h be a conforming subdivision of $\Omega \subset \mathbb{R}^2$ into non-overlapping triangles. For an element $T \in \mathcal{T}_h$, \tilde{h}_T denotes its diameter. Set $h = \max_{T \in \mathcal{T}_h} \tilde{h}_T$, and let \tilde{h} be the function such that $\tilde{h}|_T = \tilde{h}_T$. Assume that the family of meshes $\{\mathcal{T}_h\}_{h>0}$ is shape-regular, i.e., there exists $c_{\mathcal{T}}$ such that for all $h > 0$ and for all $T \in \mathcal{T}_h$

$$\frac{\tilde{h}_T}{\rho_T} \leq c_{\mathcal{T}},$$

where ρ_T denotes the largest inscribed ball in T . Let \mathcal{F}_i denote the set of interior faces of \mathcal{T}_h , i.e., the set of faces that are not included in the boundary $\partial\Omega$, let \mathcal{F}_e denote the faces that are included in $\partial\Omega$, and define $\mathcal{F}_h = \mathcal{F}_i \cup \mathcal{F}_e$. For a face $F \in \mathcal{F}_h$, h_F denotes its length, and let $h_{\mathcal{F}}$ be the function such that $h_{\mathcal{F}}|_F = h_F$. We will also use the set of faces in the neighborhood of a triangle $\mathcal{F}_T = \{F : \bar{F} \cap \bar{T} \neq \emptyset\}$.

The scalar product on a subset $X \subset \Omega$ will be denoted $(u, v)_X = \int_X uv \, dx$ on the elements $(u, v)_{\mathcal{T}_h} = \sum_{T \in \mathcal{T}_h} \int_T uv \, dx$ and on the faces of the mesh $(u, v)_{\mathcal{F}_h} = \sum_{F \in \mathcal{F}_h} \int_F uv \, ds$, with the corresponding norms $\|u\|_X = (u, u)_X^{\frac{1}{2}}$, $\|u\|_{\mathcal{T}_h} = (u, u)_{\mathcal{T}_h}^{\frac{1}{2}}$, and $\|u\|_{\mathcal{F}_h} = (u, u)_{\mathcal{F}_h}^{\frac{1}{2}}$. For $s \geq 0$, let $H^s(\mathcal{T}_h)$ be the space of elementwise Sobolev H^s -functions, and denote its norm by $\|\cdot\|_{H^s(\mathcal{T}_h)}$.

For $v \in H^2(\mathcal{T}_h)$ and an interior face $F = T_1 \cap T_2 \in \mathcal{F}_i$, where T_1 and T_2 are two distinct elements of \mathcal{T}_h with respective outer normals n_1 and n_2 , also define $\Delta_F := T_1 \cup T_2$ and $\Delta_T := \cup_{F \in \partial T} \Delta_F$. Define the jump of the normal component of the gradient by $[\nabla v]_F = \nabla v|_{T_1} \cdot n_1 + \nabla v|_{T_2} \cdot n_2$ and the jump of the tangential component of the gradient by $[\nabla v]_{\tau} = \nabla v|_{T_1} \times n_1 + \nabla v|_{T_2} \times n_2$. Similarly, for $v \in H^1(\mathcal{T}_h)$, define the jump $[v]_F = v|_{T_1} n_1 + v|_{T_2} n_2$ and the average $\{v\}_F = \frac{1}{2}(v|_{T_1} + v|_{T_2})$. On outer faces $F = \partial T \cap \partial\Omega$, with normal n_E pointing out from Ω , the vector valued jump and the scalar average are defined as $[v]_F = v|_T n_E$ and $\{v\}_F = v|_T$, respectively.

On each triangle we define $\beta_{\infty,T} := \|\beta\|_{L^\infty(T)}$, $\beta_{inf,T} := \min_{x \in T} |\beta|$, and on each face $\beta_{\infty,F} := \|\beta\|_{L^\infty(F)}$, $\beta_n = |\beta \cdot n_F|$, and $e_\beta = \beta|\beta|^{-1}$. By $\bar{\beta}_T$ we denote the average value of $\beta|_T$ over the element T . We will assume that the variation of β is bounded so that $\exists c_1, c_2 > 0$ such that $\beta_{\infty,T} \leq c_1 \beta_{inf,T}$ and $\beta_{\infty,T} \leq c_2 |\bar{\beta}_T|$. By the Lipschitz continuity of β , we deduce that $|\beta(x_1) - \beta(x_2)| \leq L_\beta h_T$ for all $x_1, x_2 \in T$.

3. The interior penalty finite element formulation. Let

$$V_h^k = \{v : v|_T \in P_k(T), \forall T \in \mathcal{T}_h\}, k \in \{0, 1\} \text{ and } V_h^C = V_h^1 \cap C^0(\bar{\Omega}).$$

The analysis is the same for continuous and discontinuous approximations, and the two cases will not be distinguished unless necessary. Indeed we will use the generic space W_h which may be chosen as either one of V_h^0 , V_h^1 , or V_h^C . The interior penalty method then takes the following form: find $u_h \in W_h$ such that

$$(3.1) \quad a_h(u_h, v_h) = \int_{\mathcal{T}_h} f v_h dx + \int_{\partial\Omega^-} \beta_n g v_h ds \quad \forall v_h \in W_h.$$

Where the form $a_h(\cdot, \cdot)$ is given by $a_h(v, w) := a(v, w) + \sum_{i=0}^1 \gamma_i j_h^i(v, w)$, with the bilinear forms $a(\cdot, \cdot)$ and $j(\cdot, \cdot)$ defined by

$$a(v, w) := (\mu v, w)_{\mathcal{T}_h} + (\beta \cdot \nabla v, w)_{\mathcal{T}_h} + ([v], \{\beta v\})_{\mathcal{F}_i \cup \partial\Omega^-},$$

$$(3.2) \quad j_h^0(v, w) := (\beta_n [v], [w])_{\mathcal{F}_i}, \quad j_h^1(v, w) := (h_{\mathcal{F}}^2 \beta_n |e_\beta \cdot n| [\nabla v], [\nabla w])_{\mathcal{F}_i}.$$

Here γ_0 and γ_1 denote positive parameters. When discontinuous approximation is being used, one may take $\gamma_1 = 0$, and we will assume this is the case below. We will assume that $\gamma_0 := \frac{1}{2}$, corresponding to the standard upwind flux.

3.0.1. Technical results. For the analysis, we need the following results of finite element analysis. For a proof, see a standard textbook such as [17] or [12].

LEMMA 3.1 (trace inequality). *Let $v|_T \in H^1(T)$; then*

$$(3.3) \quad \|v\|_{\partial T} \leq C_T (h_T^{-\frac{1}{2}} \|v\|_T + h_T^{\frac{1}{2}} \|\nabla v\|_T).$$

LEMMA 3.2 (inverse inequality). *Let $v|_T \in P_k(T)$, $k \geq 0$; then*

$$(3.4) \quad \|\nabla v\|_T \leq C_I h_T^{-1} \|v\|_T.$$

In this paper $c > 0$ denotes a generic constant that can change at each occurrence. c is independent of the meshsize h , but not necessarily of γ_1 , C_T , C_I , or the mesh regularity.

3.1. A priori analysis. Optimal a priori error estimates have been derived for the formulation (3.1). Typically, such estimates are obtained in the triple norm,

$$(3.5) \quad \|v_h\|_{h,\gamma,\Xi}^2 := \mu_0 \|v_h\|_{\mathcal{T}_h}^2 + \frac{1}{2} \|\beta_n^{\frac{1}{2}} [v_h]\|_{\mathcal{F}_h}^2 + \Xi \|\tilde{h}^{\frac{1}{2}} |\beta|^{\frac{1}{2}} e_\beta \cdot \nabla v_h\|_{\mathcal{T}_h}^2 + \gamma_1 j_h^1(v_h, v_h),$$

$$\Xi \in \{0, 1\}.$$

Following [11, 18, 9] we may show the following inf-sup condition that gives control of the h -weighted graph norm (under suitable conditions on the variation of β).

THEOREM 3.3 (stability). *For all $v_h \in W_h$ there holds*

$$(3.6) \quad \|v_h\|_{h,\gamma,\Xi} \leq c_\gamma \sup_{w_h \in W_h} \frac{a_h(v_h, w_h)}{\|w_h\|_{h,\gamma,\Xi}}.$$

Using the stability given by the inf-sup condition of (3.6) in combination with the consistency of the method (3.1) and the approximation properties of the space W_h , the following a priori error estimate is obtained.

LEMMA 3.4. *Let u be the solution of problem (2.1) and u_h the solution of problem (3.1); then there holds*

$$(3.7) \quad \|u - u_h\|_{h,0,\Xi} + \gamma_1^{\frac{1}{2}} j_h^1(u_h, u_h)^{\frac{1}{2}} \leq ch^{s-\frac{1}{2}} \|u\|_{H^s(\Omega)},$$

where $s = \min\{k + 1, r\}$ for $u \in H^r(\Omega)$ and $r \geq 1$.

Remark 1. The case $\mu = 0$ and $\nabla \cdot \beta = 0$ may also be treated. Indeed for these stabilized methods control of the L^2 -norm of the solution may be recovered from the streamline derivative using a suitably chosen weight function. This has been proven for the DG method in [20] and may be proven for the CIP method by adapting the analysis of [10].

4. A posteriori error analysis. Denote by \mathcal{T}_h a uniform refinement of \mathcal{T}_h with local meshsize $\tilde{h} = \tilde{h}/2$ and corresponding finite element space $W_{\tilde{h}}$ such that $W_h \subset W_{\tilde{h}}$. For a $T \in \mathcal{T}_h$, we will denote by \mathcal{F}_T° the set of faces of triangles in $\mathcal{T}_{\tilde{h}}$ such that $\tilde{F} \subset \overset{\circ}{T}$, i.e., the subgrid faces, the faces in $\mathcal{F}_{\tilde{h}}$ that are not included in a face in \mathcal{F}_h . We assume that there exists $c_\rho > 0$ such that for all faces $F \in \mathcal{F}_T^{\circ}$ and for all $\{\mathcal{T}_h\}_h$ there holds $|\tilde{\beta}_T \cdot n_F| \geq c_\rho |\beta_T|$. Note that this condition depends on the subdivision. It is easy to see that in two space dimensions there is always one subdivision of each T such that the condition on $\tilde{\beta}$ holds true. For instance, compare the two subdivisions in Figure 1. If $|\beta_T \cdot n| = 0$ for some face in one of the subdivisions, it will be larger than zero on all faces in the other, since no two interior faces are parallel between the two types of refinements. It follows that the constant c_ρ depends only on the shape regularity constant $c_{\mathcal{T}}$ of the mesh family. A consequence of this assumption is that $\exists c > 0$, depending only on $c_{\mathcal{T}}$ such that $|\tilde{\beta} \times n| \leq c |\beta \cdot n|$ on the subgrid faces. Finally, let $u_{\tilde{h}} \in W_{\tilde{h}}$ satisfy

$$(4.1) \quad a_{\tilde{h}}(u_{\tilde{h}}, v_{\tilde{h}}) = \int_{\mathcal{T}_{\tilde{h}}} f v_{\tilde{h}} \, dx + \int_{\partial\Omega^-} \beta_n g v_{\tilde{h}} \, ds \quad \forall v_{\tilde{h}} \in W_{\tilde{h}}.$$

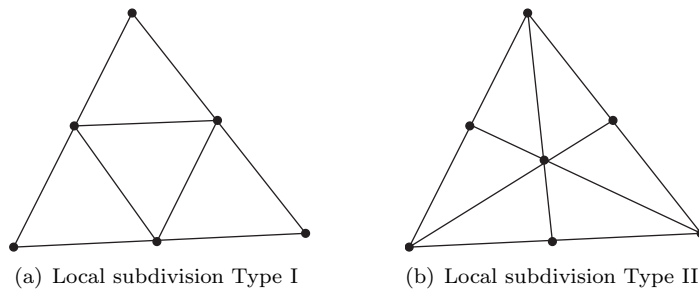


FIG. 1. Two different types of subgrids. $|\tilde{\beta} \cdot n|$ can not vanish on one (interior) face in both the Type I and the Type II subdivisions simultaneously.

We assume that the following saturation assumption holds uniformly on the family of meshes $\{\mathcal{T}_h\}_h$:

$$[\mathbf{SA}] \text{ There exists } \delta < 1 \text{ such that } \|u - u_h\|_{h,0,\Xi} \leq \delta \|u - u_h\|_{h,0,\Xi}.$$

Note that the subscript h is the same on both sides of the equality, meaning that the saturation is assumed to hold for the error in the streamline derivative without weight. This means that for $\Xi = 1$ there must hold

$$\|\beta \cdot \nabla(u - u_h)\|_{\tau_h} \leq \delta \|\beta \cdot \nabla(u - u_h)\|_{\tau_h}.$$

Remark 2. Clearly, by (3.7) $[\mathbf{SA}]$ is expected to hold for smooth u , i.e., $r \geq 1$. For irregular solutions with interior layers, the saturation assumption may fail for the streamline derivative, but is expected to hold for the L^2 -norm of the error on the interior and the boundary of Ω . Then the analysis proposed below gives an a posteriori error estimate with $\Xi = 0$. Note that in this case, for the DG method the norm also includes the jumps of the discrete solution over element faces. These can not trivially be bounded by the L^2 -norm. To show that it is not unreasonable to believe that saturation in the L^2 -norm implies saturation also for the jumps, we here sketch a simple argument showing that convergence of the error in L^2 -norm must imply convergence of the stabilization, regardless of the regularity of the underlying solution. For simplicity, we assume that $g = 0$. It follows by taking $v_h = u_h$ in (3.1) that for both the method using discontinuous approximation and the one using continuous approximation there holds

$$\mu_0 \|u_h\|_{\Omega}^2 + \frac{1}{2} \|\beta_n^{\frac{1}{2}} u_h\|_{\partial\Omega}^2 + \sum_{i=0}^1 \gamma_i j^i(u_h, u_h) \leq (f, u_h),$$

and hence

$$\mu_0 \|u_h\|_{\Omega} + \frac{1}{2} \|\beta_n^{\frac{1}{2}} u_h\|_{\partial\Omega} \leq \|f\|_{\Omega}.$$

Similarly, there holds for the continuous problem (2.1)

$$\mu_0 \|u\|_{\Omega}^2 + \frac{1}{2} \|\beta_n^{\frac{1}{2}} u\|_{\partial\Omega}^2 \leq (f, u), \text{ and hence } \mu_0 \|u\|_{\Omega} + \frac{1}{2} \|\beta_n^{\frac{1}{2}} u\|_{\partial\Omega} \leq \|f\|_{\Omega}.$$

Taking the difference of the two inequalities and using the inverse triangle inequality yields

$$\begin{aligned} \sum_{i=0}^1 j^i(u_h, u_h) &\leq (f, u_h - u) + \mu_0 (\|u\|_{\Omega}^2 - \|u_h\|_{\Omega}^2) + \frac{1}{2} (\|\beta_n^{\frac{1}{2}} u\|_{\partial\Omega}^2 - \|\beta_n^{\frac{1}{2}} u_h\|_{\partial\Omega}^2) \\ (4.2) \quad &= (f, u_h - u) + \mu_0 (\|u\|_{\Omega} + \|u_h\|_{\Omega}) (\|u\|_{\Omega} - \|u_h\|_{\Omega}) \\ &\quad + \frac{1}{2} (\|\beta_n^{\frac{1}{2}} u\|_{\partial\Omega} + \|\beta_n^{\frac{1}{2}} u_h\|_{\partial\Omega}) (\|\beta_n^{\frac{1}{2}} u\|_{\partial\Omega} - \|\beta_n^{\frac{1}{2}} u_h\|_{\partial\Omega}) \\ &\leq c \|f\|_{\Omega} (\|u_h - u\|_{\Omega} + \|\beta_n^{\frac{1}{2}}(u_h - u)\|_{\partial\Omega}). \end{aligned}$$

It follows that if we have convergence of the error in the L^2 -norm (on the volume and on the boundary), the stabilization terms must converge at least with a rate that is the square root of that of the L^2 -norm, even in the presence of sharp layers.

4.1. Upper bounds. First, we need a result on interpolation between discrete spaces in the following auxilliary norm:

$$(4.3) \quad \|w_h\|_{h,*}^2 := \sum_{T \in \mathcal{T}_h} \max\{\mu_0, \beta_{\infty,T} h_T^{-1}\} \|w_h\|_T^2 + \|\beta_{\infty,F}^{1/2}\{w_h\}\|_{\mathcal{F}_h}^2 + \sum_{i=0}^1 \gamma_i j_h^i(w_h, w_h).$$

LEMMA 4.1. *Let $v_h \in W_h$, and let $v_h = i_h^k v_h \in W_h$, where i_h^k denotes the standard (elementwise) Lagrange interpolant when $W_h = V_h^C$ and the L^2 -projection onto V_h^k when $W_h = V_h^k$. There holds*

$$(4.4) \quad \|v_h - v_h\|_{h,*} \leq c \|v_h\|_{h,0,k}.$$

Proof. First note that $\mu_0 \|v_h - v_h\|_T^2 \leq c \mu_0 \|v_h\|_T^2$, so we can assume that the max is taken by $\beta_{\infty,T} h_T^{-1}$. Using the trace inequality (3.3) and the inverse inequality (3.4), we obtain

$$(4.5) \quad \sum_{F \in \mathcal{F}_h} \|\beta_{\infty,F}^{1/2}\{v_h - v_h\}\|_F^2 + \sum_{i=0}^1 \gamma_i j_h^i(v_h - v_h, v_h - v_h) \leq c \sum_{T \in \mathcal{T}_h} \beta_{\infty,T} h_T^{-1} \|v_h - v_h\|_T^2.$$

The following local interpolation estimate on $T \in \mathcal{T}_h$ may be proved using norm equivalence on the reference element and scaling (see [7] for details):

$$\beta_{\infty,T} h_T^{-1} \|v_h - v_h\|_T^2 \leq c \beta_{\infty,T} h_T^{-1} \sum_{\bar{F} \in \mathcal{F}_T} \left(\|h_{\bar{F}}^{3/2} [\nabla v_h]\|_{\bar{F}}^2 + \|h_{\bar{F}}^{1/2} [v_h]\|_{\bar{F}}^2 \right).$$

By the assumptions on β and the subgrid mesh, we have $\beta_{\infty,T} \leq c |\bar{\beta}_T \cdot n|$ on each face in \mathcal{F}_T . It then follows, since $|\bar{\beta}_T - \beta|_T \leq L_\beta h$, that

$$\beta_{\infty,T} h_T^{-1} \|h_{\bar{F}}^{1/2} [v_h]\|_{\bar{F}}^2 \leq c \|\bar{\beta}_T \cdot n\|^{1/2} \|v_h\|_{\bar{F}}^2 \leq c (\|\beta_n^{1/2} [v_h]\|_{\bar{F}}^2 + L_\beta \|v_h\|_T^2).$$

For the term with the gradient jumps we note that, using the equality

$$\bar{\beta}_T \cdot \nabla u_h = \bar{\beta}_T \cdot n \nabla u_h \cdot n + \bar{\beta}_T \times n \cdot \nabla u_h \times n$$

and since $\bar{\beta}_T$ is constant on the triangle T , we may write

$$\begin{aligned} \beta_{\infty,T} h_T^{-1/2} \|h_{\bar{F}}^{3/2} [\nabla v_h]\|_{\bar{F}}^2 &\leq c \|h_{\bar{F}} |\bar{\beta} \cdot n|^{1/2} [\nabla v_h]\|_{\bar{F}}^2 \\ &\leq \frac{c}{|\bar{\beta}|} \|h_T [\bar{\beta} \cdot \nabla v_h]\|_{\bar{F}}^2 + c \|h_T |\bar{\beta} \times n| [\nabla v_h]_\tau\|_{\bar{F}}^2 \\ &\leq c (\|h_T^{1/2} |\beta|^{1/2} e_\beta \cdot \nabla v_h\|_T^2 + \|\beta_n^{1/2} [v_h]\|_{\bar{F}}^2 + L_\beta \|v_h\|_T^2). \end{aligned}$$

In the last inequality we have also used the inverse inequality $h_F \|[\nabla v_h]_\tau\|_F \leq c \|v_h\|_F$ and the bound $|\bar{\beta} \times n| \leq c |\bar{\beta} \cdot n|$. We conclude by summing over $T \in \mathcal{T}_h$. \square

LEMMA 4.2 (Galerkin orthogonality). *Let u_h be the solution of (4.1) in W_h and u_h be the solution of (3.1) in W_h ; then, in case $k = 0$,*

$$a_h(u_h - u_h, v_h) = 0 \quad \forall v_h \in W_h,$$

and, in case $k = 1$,

$$a_h(u_h - u_h, v_h) = -\gamma_1 \frac{3}{4} j_h^1(u_h, v_h) \quad \forall v_h \in W_h.$$

Proof. The argument is standard in the case $k = 0$, since in that case $a_{\mathfrak{h}}(\cdot, \cdot)$ is independent of \mathfrak{h} . For $k = 1$ observe that by definition there holds

$$(4.6) \quad j_{\mathfrak{h}}^1(u_{\mathfrak{h}}, v_h) = \frac{1}{4}j_h^1(u_{\mathfrak{h}}, v_h) \quad \forall u_{\mathfrak{h}} \in W_{\mathfrak{h}} \text{ and } \forall v_h \in W_h.$$

It follows that

$$a_{\mathfrak{h}}(u_{\mathfrak{h}} - u_h, v_h) = a_{\mathfrak{h}}(u_{\mathfrak{h}}, v_h) - a_h(u_h, v_h) + \gamma_1 j_h^1(u_h, v_h) - \gamma_1 j_{\mathfrak{h}}^1(u_h, v_h).$$

The claim now follows by (3.1), (4.1), and (4.6). \square

THEOREM 4.3 (a posteriori error estimate). *If assumption [SA] holds for some $\Xi \in \{0, 1\}$ (if $k = 0$, then $\Xi = 0$), then there exists a constant c_{δ} independent of h , but not δ or the local mesh geometry such that*

$$\|u - u_h\|_{h,0,\Xi}^2 \leq c_{\delta} \sum_{T \in \mathcal{T}_h} (\eta_{1,T}^2 + \eta_{2,T}^2 + \eta_{3,T}^2 + \eta_{4,T}^2) =: \eta(u_h)^2,$$

where $\eta_{1,T} = \alpha_{1,T} \|f - \beta \cdot \nabla u_h - \mu u_h\|_T$, $\eta_{2,T} = \alpha_2 \|\beta_n^{\frac{1}{2}} [u_h]\|_{\partial T \setminus \partial \Omega}$, $\eta_{3,T} = (1 - \alpha_2) \gamma_1^{\frac{1}{2}} \|\beta_n^{\frac{1}{2}} |e_{\beta} \cdot n_T|^{\frac{1}{2}} h_T [\nabla u_h]\|_{\partial T \setminus \partial \Omega}$, and $\eta_{4,T} = \|\beta_n^{\frac{1}{2}} (u_h - g)\|_{\partial \Omega \cap \partial T}$, where $\alpha_{1,T} = \min \{\mu_0^{-\frac{1}{2}}, h_T^{\frac{1}{2}} \beta_{\infty,T}^{-\frac{1}{2}}\}$, $\alpha_2 = 1$ for $W_h = V_h^k$ and $\alpha_2 = 0$ for $W_h = V_h^C$.

Proof. Using [SA] we get the bound

$$\|u - u_h\|_{h,0,\Xi} \leq \|u - u_{\mathfrak{h}}\|_{h,0,\Xi} + \|u_{\mathfrak{h}} - u_h\|_{h,0,\Xi} \leq \delta \|u - u_h\|_{h,0,\Xi} + \|u_{\mathfrak{h}} - u_h\|_{h,0,\Xi}.$$

Consequently, $\|u - u_h\|_{h,0,\Xi} \leq 2^{\frac{1}{2}}(1 - \delta)^{-1} \|u_{\mathfrak{h}} - u_h\|_{h,0,\Xi} \leq c \|u_{\mathfrak{h}} - u_h\|_{\mathfrak{h},\gamma,\Xi}$. By Theorem 3.3, a uniform inf-sup condition holds in $W_{\mathfrak{h}}$ for $\|u_{\mathfrak{h}} - u_h\|_{\mathfrak{h},\gamma,\Xi}$:

$$(4.7) \quad \|u_{\mathfrak{h}} - u_h\|_{\mathfrak{h},\gamma,\Xi} \leq c_{\gamma} \sup_{v_{\mathfrak{h}} \in W_{\mathfrak{h}}} \frac{a_{\mathfrak{h}}(u_{\mathfrak{h}} - u_h, v_{\mathfrak{h}})}{\|v_{\mathfrak{h}}\|_{\mathfrak{h},\gamma,\Xi}}.$$

By Galerkin orthogonality and Lemma 4.2,

$$a_{\mathfrak{h}}(u_{\mathfrak{h}} - u_h, v_{\mathfrak{h}}) = a_{\mathfrak{h}}(u_{\mathfrak{h}} - u_h, v_{\mathfrak{h}} - i_h^k v_{\mathfrak{h}}) + k \frac{3}{4} \gamma_1 j_h^1(u_h, i_h^k v_{\mathfrak{h}}) = I + II,$$

where $i_h^k v_{\mathfrak{h}} \in W_h$ is defined in Lemma 4.1. Now (3.1) is used to eliminate the $u_{\mathfrak{h}}$ in the left slot of $a_{\mathfrak{h}}(\cdot, \cdot)$. The error is bounded by the residual of u_h weighted with the discrete error $v_{\mathfrak{h}} - i_h^k v_{\mathfrak{h}}$:

$$(4.8) \quad \begin{aligned} I &= a_{\mathfrak{h}}(u_{\mathfrak{h}} - u_h, v_{\mathfrak{h}} - i_h^k v_{\mathfrak{h}}) = (f, v_{\mathfrak{h}} - i_h^k v_{\mathfrak{h}})_{\mathcal{T}_{\mathfrak{h}}} + (\beta_n g, v_{\mathfrak{h}} - i_h^k v_{\mathfrak{h}})_{\partial \Omega} - \\ &\quad - a_{\mathfrak{h}}(u_h, v_{\mathfrak{h}} - i_h^k v_{\mathfrak{h}}) \\ &= (f - \beta \cdot \nabla u_h - \mu u_h, v_{\mathfrak{h}} - i_h^k v_{\mathfrak{h}})_{\mathcal{T}_{\mathfrak{h}}} + (\beta_n (g - u_h), v_{\mathfrak{h}} - i_h^k v_{\mathfrak{h}})_{\partial \Omega} - \\ &\quad - ([u_h], \{\beta(v_{\mathfrak{h}} - i_h^k v_{\mathfrak{h}})\})_{\mathcal{F}_i} - \gamma_0 j_{\mathfrak{h}}^0(u_h, v_{\mathfrak{h}} - i_h^k v_{\mathfrak{h}}) - \gamma_1 j_{\mathfrak{h}}^1(u_h, v_{\mathfrak{h}} - i_h^k v_{\mathfrak{h}}). \end{aligned}$$

For the weakly consistent jump term, note that (4.6) allows us exchange the subscripts \mathfrak{h} and h :

$$(4.9) \quad \begin{aligned} II &= \frac{3}{4} \gamma_1 j_h^1(u_h, i_h^k v_{\mathfrak{h}}) = \frac{3}{4} \gamma_1 (j_h^1(u_h, i_h^k v_{\mathfrak{h}} - v_{\mathfrak{h}}) + j_h^1(u_h, v_{\mathfrak{h}})) \\ &\leq \frac{3}{2} \gamma_1 j_h^1(u_h, u_h)^{\frac{1}{2}} (j_{\mathfrak{h}}^1(i_h^k v_{\mathfrak{h}} - v_{\mathfrak{h}}, i_h^k v_{\mathfrak{h}} - v_{\mathfrak{h}})^{\frac{1}{2}} + j_{\mathfrak{h}}^1(v_{\mathfrak{h}}, v_{\mathfrak{h}})^{\frac{1}{2}}). \end{aligned}$$

Now observe that by (4.8) and the Cauchy–Schwarz inequality

$$(4.10) \quad |(f, v_h - i_h^k v_h)_{\mathcal{T}_h} + (\beta_n g, v_h - i_h^k v_h)_{\partial\Omega^-} - a_h(u_h, v_h - i_h^k v_h)| \leq \left(\sum_{T \in \mathcal{T}_h} \sum_{i=1}^4 \eta_{i,T}^2 \right)^{\frac{1}{2}} \|v_h - i_h^k v_h\|_{h,*}.$$

Apply Lemma 4.1 in the right-hand side of (4.10):

$$\|v_h - i_h^k v_h\|_{h,*} \leq c \|v_h\|_{h,\gamma,\Xi},$$

and in the remaining weakly consistent contribution in the right-hand side of (4.9):

$$\gamma_1 J_h^1(i_h^k v_h - v_h, i_h^k v_h - v_h) \leq \|v_h - i_h^k v_h\|_{h,*} \leq c \|v_h\|_{h,\gamma,\Xi}.$$

The result follows by dividing by $\|v_h\|_{h,\gamma,\Xi}$. \square

This estimate is optimal for smooth solutions. For $u \in H^r(\Omega)$, $r \geq 1$, and with $s = \min\{k + 1, r\}$, $\eta(u_h) \leq ch^{s-\frac{1}{2}} \|u\|_{H^s(\Omega)}$ is easily shown using (3.7).

A possibly sharper estimate can be obtained if one makes use of the orthogonality of the fine to coarse projection in the case of the DG method or uses the global L^2 -projection and its orthogonality (instead of the Lagrange interpolant) for the CIP method. We collect these results in two corollaries.

COROLLARY 4.4. *Under the same assumptions as in Theorem 4.3, when $k = 0, 1$, and discontinuous approximation is used, the volume estimator $\eta_{1,T}$ may be replaced by*

$$\eta_{1,T}^* = \alpha_{1,T} (\|f - i_h^k f\|_T + c \|(\beta - i_h^k \beta) \cdot \nabla u_h\|_T).$$

Proof. In the error representation, clearly by the orthogonality of the projection i_h^k we have

$$(f - \beta \cdot \nabla u_h - \mu u_h, v_h - i_h^k v_h)_{\mathcal{T}_h} = (f - i_h^k f - \beta \cdot \nabla u_h + i_h^k \beta \cdot \nabla u_h, v_h - i_h^k v_h)_{\mathcal{T}_h},$$

where μu_h vanishes, since here $\mu \in \mathbb{R}$. The result follows as in the proof of Theorem 4.3, noting that, for $k = 1$,

$$\begin{aligned} \|\beta \cdot \nabla u_h - i_h^1(\beta \cdot \nabla u_h)\|_T &\leq \|(\beta - i_h^1 \beta) \cdot \nabla u_h\|_T + \|i_h^1 \beta \cdot \nabla u_h - i_h^1(\beta \cdot \nabla u_h)\|_T \\ &= \|(\beta - i_h^1 \beta) \cdot \nabla u_h\|_T + \|i_h^1(i_h^1 \beta \cdot \nabla u_h - (\beta \cdot \nabla u_h))\|_T \\ &\leq c \|(\beta - i_h^1 \beta) \cdot \nabla u_h\|_T. \quad \square \end{aligned}$$

A similar result holds for the CIP method. However, in this case, due to the continuous approximation space used, the proof needs stronger restrictions on the variations of the data and the mesh, since we need stability in weighted norm of the global L^2 -projection $\pi_h : L^2(\Omega) \rightarrow V_h^C$. In particular we assume that the estimate

$$(4.11) \quad \|\pi_h(v_h - i_h^k v_h)\|_{h,*} \leq c \|v_h - i_h^k v_h\|_{h,*}$$

holds. This can be proven using the techniques of [6], provided the local variation of the mesh parameter and the velocity is sufficiently small. We will also use a quasi interpolant based on nodal local averages $i_h^O : V_h^1 \rightarrow V_h^C$, defined in each node x_i as

$$(i_h^O v)(x_i) = \frac{1}{n_i} \sum_{T: T \cap x_i \neq \emptyset} v|_T(x_i),$$

where n_i denotes the cardinality of the set $\{T : T \cap x_i \neq \emptyset\}$.

COROLLARY 4.5. *Under the same assumptions as in Theorem 4.3 and assuming that (4.11) holds, when continuous approximation is used, the volume estimator $\eta_{1,T}$ may be replaced by*

$$\eta_{1,T}^* = \alpha_{1,T} \left(\|f - f_h\|_T + c \|(\beta - i_h^1 \beta) \cdot \nabla u_h\|_T + c \sum_{F \in \mathcal{F}_T} \|h_{\bar{T}} [i_h^1 \beta \cdot \nabla u_h]\|_{F \setminus \partial \Omega} \right),$$

where f_h denotes some suitable approximation of f in W_h .

Proof. This time we proceed as in the proof of Theorem 4.3, but we use $\pi_h v_h$ instead of the nodal interpolant. As before, we may then use orthogonality of the L^2 -projection to write

$$(4.12) \quad \begin{aligned} & (f - \beta \cdot \nabla u_h - \mu u_h, v_h - \pi_h v_h)_{\mathcal{T}_h} \\ &= (f - f_h - (\beta - i_h^1 \beta) \cdot \nabla u_h + (i_h^1 \beta) \cdot \nabla u_h - i_h^O((i_h^1 \beta) \cdot \nabla u_h), v_h - \pi_h v_h)_{\mathcal{T}_h}. \end{aligned}$$

We now use the following discrete interpolation estimate (see [8]):

$$\|h_T^{\frac{1}{2}}((i_h^1 \beta) \cdot \nabla u_h - i_h^O((i_h^1 \beta) \cdot \nabla u_h))\|_T \leq c \sum_{F \in \mathcal{F}_h(\Delta_T)} \|h_{\bar{T}}[(i_h^1 \beta) \cdot \nabla u_h]\|_{F \setminus \partial \Omega}$$

and conclude in the same fashion as in Theorem 4.3, noting that by (4.11)

$$\|v_h - \pi_h v_h\|_{h,*} \leq \|v_h - i_h^1 v_h\|_{h,*} + \|\pi_h(v_h - i_h^1 v_h)\|_{h,*} \leq (1 + c) \|v_h - i_h^1 v_h\|_{h,*}. \quad \square$$

4.2. Lower bounds. We now prove that the estimator also is a lower bound of the local error. Indeed, when $\Xi = 1$, the estimators are upper bounded by the error and some terms measuring oscillation of data. In the case of the DG method, the constants in the lower bounds depend only on the size of $\nabla \cdot \beta$. For the CIP method, the stabilization introduces an additional coupling between the convection and reaction terms, and robustness is only obtained for small enough h_T .

LEMMA 4.6 (lower bounds). *The following lower bounds hold for the estimators of Theorem 4.3:*

$$(4.13) \quad \eta_{1,T} \leq c (\|h^{\frac{1}{2}} |\beta|^{\frac{1}{2}} e_\beta \cdot \nabla(u - u_h)\|_T + c_\mu \|\mu_0^{\frac{1}{2}}(u - u_h)\|_T),$$

$$(4.14) \quad \eta_{2,T} + \eta_{4,T} \leq c \|\beta_n^{\frac{1}{2}} [u - u_h]\|_{\partial T},$$

and for $h_T < \beta_{\infty,T}$

$$(4.15) \quad \begin{aligned} \eta_{3,T} &\leq c \left(\|h^{\frac{1}{2}} |\beta|^{\frac{1}{2}} e_\beta \cdot \nabla(u - u_h)\|_{\Delta_T} + c_\mu \|\mu_0^{\frac{1}{2}}(u - u_h)\|_{\Delta_T} \right) \\ &+ c \beta_{\infty,T}^{-\frac{1}{2}} \left(\|h^{\frac{1}{2}}(f - f_h)\|_{\Delta_T} \right. \\ &\left. + \|\tilde{h}^{\frac{1}{2}}(\beta - i_h^1 \beta) \cdot \nabla u_h\|_{\Delta_T} + \|\tilde{h}^{\frac{3}{2}} |\nabla(\beta - i_h^1 \beta)| \nabla u_h\|_{\Delta_T} \right), \end{aligned}$$

where f_h denotes some suitable discrete representation of f and $c_\mu = \frac{\mu}{\mu_0^{\frac{1}{2}}}$.

Proof. The proof of the bound (4.13) follows by using the equation in the estimator

$$\eta_{1,T} = \alpha_{1,T} \|f - \beta \cdot \nabla u_h - \mu u_h\|_T \leq \alpha_{1,T} (\|\beta \cdot \nabla(u - u_h)\|_T + \|\mu(u - u_h)\|_T).$$

For the first term, we may assume that the min in $\alpha_{1,T}$ is taken for $\beta_{\infty,T}^{-\frac{1}{2}} h_T^{\frac{1}{2}}$, and it follows that

$$\alpha_{1,T} \|\beta \cdot \nabla(u - u_h)\|_T \leq \|h_T^{\frac{1}{2}} |\beta|^{\frac{1}{2}} e_\beta \cdot \nabla(u - u_h)\|_T.$$

For the second term, assume that the min is taken for $\mu_0^{-\frac{1}{2}}$, note that

$$\|\mu(u - u_h)\|_T \leq \left(\frac{\mu}{\mu_0^{\frac{1}{2}}}\right) \|\mu_0^{\frac{1}{2}}(u - u_h)\|_T,$$

and choose $c_\mu \geq \frac{\mu}{\mu_0^{\frac{1}{2}}}$. The second bound (4.14) follows by definition. To prove (4.15), first note that, since $\beta_n |e_\beta \cdot n| = \beta_n^2 |\beta|^{-1}$ and $W_h = V_h^C$, we have

$$\eta_{3,T}^2 = \gamma_1^{\frac{1}{2}} \|\beta_n |\beta|^{-\frac{1}{2}} h_T [\nabla u_h]\|_{\partial T \setminus \partial \Omega}^2 \leq \gamma_1 \beta_{inf,T}^{-1} \int_{\partial T \setminus \partial \Omega} \beta_n^2 h_T^2 [\nabla u_h]^2 ds.$$

Now we add and subtract the Lagrange interpolant of the velocity field $i_h^1 \beta$:

$$\eta_{3,T}^2 \leq c \beta_{inf,T}^{-1} \int_{\partial T \setminus \partial \Omega} (h_T^2 |i_h^1 \beta \cdot n|^2 [\nabla u_h]^2 + h_T^2 |(\beta - i_h^1 \beta) \cdot n|^2 [\nabla u_h]^2) ds = I + II.$$

Consider the terms I and II . For I , note that since $i_h^1 \beta, u_h, f_h \in C^0(\bar{\Omega})$, there holds $|i_h^1 \beta \cdot n|^2 [\nabla u_h]^2 = [\mu u_h + i_h^1 \beta \cdot \nabla u_h - f_h]^2$, and as a consequence, again using (3.3) and (3.4),

$$\begin{aligned} I &= c \beta_{inf,T}^{-1} \gamma_1 \int_{\partial T \setminus \partial \Omega} h_T^2 [\mu u_h + i_h^1 \beta \cdot \nabla u_h - f_h]^2 ds \\ &\leq c \beta_{inf,T}^{-1} \|h_T^{\frac{1}{2}} (\mu u_h + (i_h^1 \beta) \cdot \nabla u_h - f_h)\|_{\Delta_T}^2 \\ &\leq c \beta_{\infty,T}^{-1} \|h_T^{\frac{1}{2}} (\mu u_h + \beta \cdot \nabla u_h - f)\|_{\Delta_T}^2 \\ &\quad + c \beta_{\infty,T}^{-1} \|h_T^{\frac{1}{2}} (\beta - i_h^1 \beta) \cdot \nabla u_h\|_{\Delta_T}^2 + 2c \beta_{\infty,T}^{-1} \|h_T^{\frac{1}{2}} (f - f_h)\|_{\Delta_T}^2. \end{aligned}$$

The upper bound of I now follows after applying the bound for $\eta_{1,T}$ to the first term in the right-hand side and the assumed upper bound on h_T . For the term II , we use a trace inequality

$$II \leq C_T^2 (\|\tilde{h}^{\frac{1}{2}} (\beta - i_h^1 \beta) \cdot \nabla u_h\|_{L^2(\Delta_T)}^2 + \|\tilde{h}^{\frac{3}{2}} |\nabla(\beta - i_h^1 \beta)| \nabla u_h\|_{L^2(\Delta_T)}^2). \quad \square$$

Remark 3. Using the stability estimate (3.6) and the inverse inequality (3.4), one may prove that the oscillation term coming from the transport term is at worst $O(h^{\frac{1}{2}})$ and hence is no worse than the oscillation term for f .

4.2.1. Lower bounds when $\Xi = 0$. If the exact solution contains sharp layers, the saturation assumption may fail for the streamline derivative. This is illustrated in the numerical section. A possibility then is to consider the residual estimator as an upper bound for $\|u - u_h\|_{h,0,0}$, implying that we control only the L^2 -norm of the error on the domain and its boundary for CIP. For DG, we also control the internal jumps. In this case, however, the above derived lower bounds fail, since we have eliminated the error in the streamline derivative in the left-hand side, but it is still present in the right-hand side. Indeed at first sight it does not seem possible to upper bound $\eta_{1,T}$ of Theorem 4.3 by $\|u - u_h\|_{h,0,0}$.

In this section we will therefore sketch how we still may prove a global lower bound for the case of the estimator of Corollaries 4.4 and 4.5. Indeed in this case the volume estimator η_1 instead takes the form of data oscillation and penalty terms. Since the

jumps of the solution for the DG method are an error quantity in $\|u - u_h\|_{h,0,0}$, we have the desired lower bound immediately in this case. On the other hand, the internal jumps are also part of the a posteriori error bound, and hence this lower bound may seem artificial and of little use. However, for both continuous and discontinuous approximations, we may now use the bound (4.2) to upper bound the penalty term by the L^2 -norms of the error only. This last step, however, introduces a gap that prevents full optimality. Indeed the L^2 -norm may at worst converge at double the rate of the estimator, without violating the lower bound.

LEMMA 4.7. *Consider the DG method with $k = 0, 1$. Then the following lower bound holds:*

$$\sum_{T \in \mathcal{T}_h} ((\eta_{1,T}^*)^2 + \eta_{2,T}^2) \leq c(\text{Osc}(f, \beta) + \|\beta_n^{\frac{1}{2}}[u_h]\|_{\mathcal{F}_i}^2),$$

where

$$\text{Osc}(f, \beta) = \sum_{T \in \mathcal{T}_h} \alpha_{1,T}^2 \left(\|f - f_h\|_T^2 + \|(\beta - i_h^1 \beta) \nabla u_h\|_T^2 \right).$$

Moreover,

$$\sum_{T \in \mathcal{T}_h} ((\eta_{1,T}^*)^2 + \eta_{2,T}^2) \leq \text{Osc}(f, \beta) + c\|f\|_{\Omega} (\|u - u_h\|_{\Omega} + \|\beta_n^{\frac{1}{2}}(u - u_h)\|_{\partial\Omega}).$$

Proof. The first upper bound holds by definition. The second follows from the energy inequality (4.2). \square

LEMMA 4.8. *Consider the CIP method. Then the following lower bound holds:*

$$\sum_{T \in \mathcal{T}_h} ((\eta_{1,T}^*)^2 + \eta_{3,T}^2) \leq \text{Osc}(f, \beta) + c\|f\|_{\Omega} (\|u - u_h\|_{\Omega} + \|\beta_n^{\frac{1}{2}}(u - u_h)\|_{\partial\Omega}),$$

where

$$\text{Osc}(f, \beta) = \sum_{T \in \mathcal{T}_h} \alpha_{1,T}^2 \left(\|f - f_h\|_T^2 + \|(\beta - i_h^1 \beta) \nabla u_h\|_T^2 + h_T \|\nabla(\beta - i_h^1 \beta) \nabla u_h\|_T^2 \right).$$

Proof. It is straightforward to show that

$$\begin{aligned} \eta_{1,T}^* \leq & \alpha_{1,T} \left(\|f - f_h\|_T + \|(\beta - i_h^1 \beta) \nabla u_h\|_T + h_T \|\nabla(\beta - i_h^1 \beta) \nabla u_h\|_T \right) \\ & + c \sum_{F \in \mathcal{F}_T} \|h_F \beta_n^{\frac{1}{2}} |e_{\beta} \cdot n|^{\frac{1}{2}} [\nabla u_h]\|_{F \setminus \partial\Omega}. \end{aligned}$$

We conclude by taking the square of $\eta_{1,T}^*$, summing over the triangles in \mathcal{T}_h , and applying the estimate (4.2), for the continuous approximation, to bound the penalty term. \square

4.3. Summary of upper and lower a posteriori bounds. Let us here briefly summarize what results can be expected for different types of solutions.

1. Smooth solutions (possibly with smooth but steep internal layers):

When the adaptive algorithm has resolved the internal layers, the saturation assumption [SA] is expected to hold with $\Xi = 1$. Then applying the analysis above we have the upper and lower bounds

$$\|u - u_h\|_{h,0,1} \leq \eta(u_h) \leq c\|u - u_h\|_{h,0,1},$$

where the lower bound may be written on local form.

2. Nonsmooth solutions:

In the case of nonsmooth solutions, i.e., solutions that have only the minimal regularity, it is only reasonable to assume saturation for the L^2 -norm. Thanks to (4.2) we may assume saturation also for the jumps. We then have the following global upper and lower bounds:

$$\|u - u_h\|_{h,0,0} \leq \eta(u_h) \leq c(\|u - u_h\|_{\Omega}^{\frac{1}{2}} + \|\beta_n(u - u_h)\|_{\partial\Omega}^{\frac{1}{2}}).$$

Note that this time there is a gap between the upper and lower bounds. Indeed the estimator can be proven to converge only at half the order of the L^2 -norm. This is because the estimator also includes the jumps of the solution or of the gradient. These can only be proven to converge at a lower rate than the L^2 -norm of the error. If this estimate is sharp, the gap will lead to a nonuniform efficiency index. This is not observed in our numerical experiments.

Remark 4 (the role of crosswind diffusion). It is known that the addition of some weakly consistent crosswind diffusion introduces additional smoothing and can improve local estimates (see [10]). In the above analysis the addition of crosswind diffusion changes the assumptions needed. Indeed if we choose $\beta_n := \beta_{\infty,F}$ in the penalty terms, then the assumption $c_{\rho}|\bar{\beta}_T| \leq |\bar{\beta}_T \cdot n|$ on all faces in the subgrid can be lifted when using continuous approximation or the low order discontinuous approximation. In this case the right-hand side of (4.5) is immediately bounded by the penalty term. On the other hand, the crosswind part of the error estimator does not allow for an optimal lower bound, since it cannot be upper bounded by $\|u - u_h\|_{h,0,\Xi}$. It can only be upper bounded using (4.2).

4.4. Goal-oriented a posteriori error estimation. In many cases a posteriori error estimates of functionals of the error are of interest. In this section we will focus on linear error functionals on the form

$$\Lambda(u - u_h) = \int_{\mathcal{T}_h} (u - u_h)\Psi dx + \int_{\partial\Omega^+} \beta \cdot n(u - u_h)\psi ds,$$

where Ψ and ψ are some user-specified functions defining the quantity of interest. We will here outline how a duality argument may be used to obtain rigorous a posteriori error bounds under the saturation assumption [SA] with $\Xi = 0$. The constant from the saturation assumption enters this estimate explicitly. Recall that in the dual weighted residual method it is assumed that the local norms of certain derivatives of the dual problem may be accurately approximated, which make a saturation assumption enter implicitly (see, i.e., [22]). In the analysis we propose below, [SA] is applied to a dual solution φ and its finite element solution φ_h . If the goal function is some average with smooth weight, the dual solution will be smooth, implying that the saturation assumption is satisfied. Note that the convergence rates of the estimates proposed below are completely independent of the regularity of u . Indeed the rate of convergence will depend only on the smoothness of φ and hence on Ψ , the error quantity. First, we will present the fundamental duality property that holds, thanks to the dual consistency of the methods considered. Then we will use this dual consistency and show an a priori error estimate in the dual norm of H^1 that does not need regularity of u . Finally, we will derive the goal-oriented a posteriori error estimate.

To this aim we introduce the following dual problem:

$$(4.16) \quad \begin{aligned} (\mu - \nabla \cdot \beta)\varphi - \beta \cdot \nabla \varphi &= \Psi \text{ in } \Omega, \\ \varphi|_{\partial\Omega^+} &= \psi. \end{aligned}$$

The variational formulation of this problem is as follows: Find $\varphi \in W$ such that $a(v, \varphi) = \Lambda(v)$ for all $v \in W$. The finite element approximation is given by the following: Find $\varphi_h \in W_h$ such that

$$(4.17) \quad a_h(v_h, \varphi_h) = \Lambda(v_h) \quad \forall v_h \in W_h.$$

For the interior penalty method (3.1), the following duality relation holds.

LEMMA 4.9 (duality). *Let u, u_h, φ , and φ_h be the solutions of (2.1), (3.1), (4.16), and (4.17), respectively. Then there holds*

$$(4.18) \quad \Lambda(u - u_h) = \int_{\mathcal{T}_h} (\varphi - \varphi_h) f \, dx + \int_{\partial\Omega^-} \beta_n (\varphi - \varphi_h) g \, ds.$$

Proof. Using the linearity of the functionals and the definitions of the problems (2.1), (3.1), (4.16), and (4.17) we have

$$\Lambda(u - u_h) = a(u, \varphi) - a_h(u_h, \varphi_h) = \int_{\mathcal{T}_h} (\varphi - \varphi_h) f \, dx + \int_{\partial\Omega^-} \beta_n (\varphi - \varphi_h) g \, ds. \quad \square$$

LEMMA 4.10 (a priori error estimate in a dual norm). *Assume that $f \in L^2(\Omega)$ and $\beta_n^{\frac{1}{2}} g \in L^2(\partial\Omega^-)$; then there holds*

$$\|u - u_h\|_{(H^1(\Omega))'} := \sup_{\substack{\Psi \in H^1(\Omega) \\ \|\Psi\|_{H^1(\Omega)} \neq 0}} \frac{|(u - u_h, \Psi)_\Omega|}{\|\Psi\|_{H^1(\Omega)}} \leq c \left(\|f\|_{\mathcal{T}_h}^2 + \|\beta_n^{\frac{1}{2}} g\|_{\partial\Omega^-}^2 \right)^{\frac{1}{2}} h^{\frac{1}{2}},$$

with c independent of u and h , but not of the mesh regularity.

Proof. By (4.18) with $\psi = 0$ we may write

$$|(u - u_h, \Psi)_\Omega| \leq \left(\|f\|_{\mathcal{T}_h}^2 + \|\beta_n^{\frac{1}{2}} g\|_{\partial\Omega^-}^2 \right)^{\frac{1}{2}} \left(\|\varphi - \varphi_h\|_{\mathcal{T}_h}^2 + \|\beta_n^{\frac{1}{2}} (\varphi - \varphi_h)\|_{\partial\Omega^-}^2 \right)^{\frac{1}{2}}.$$

It follows from the a priori error estimate (3.7) that

$$\left(\|\varphi - \varphi_h\|_{\mathcal{T}_h}^2 + \|\beta_n^{\frac{1}{2}} (\varphi - \varphi_h)\|_{\partial\Omega^-}^2 \right)^{\frac{1}{2}} \leq ch^{\frac{1}{2}} \|\varphi\|_{H^1(\Omega)}.$$

We conclude by applying the following regularity estimate $\|\varphi\|_{H^1(\Omega)} \leq c\|\Psi\|_{H^1(\Omega)}$ (see [19, 24]), dividing with $\|\Psi\|_{H^1(\Omega)}$, and taking the supremum over all nonzero Ψ in $H^1(\Omega)$. \square

THEOREM 4.11 (a posteriori estimate for linear functionals). *Let $\Psi \in L^2(\Omega)$ and $\beta_n^{\frac{1}{2}} \psi \in L^2(\partial\Omega^+)$. Assume that the saturation assumption [SA], with $\Xi = 0$, is satisfied for the dual solution φ and its approximation φ_h . Then there holds*

$$|\Lambda(u - u_h)| \leq c_{f,g,\delta} \left\{ \sum_{T \in \mathcal{T}_h} (\eta_{1,T}^2 + \eta_{2,T}^2 + \eta_{3,T}^2 + \eta_{4,T}^2) \right\}^{\frac{1}{2}},$$

where $\eta_{1,T} = \alpha_{1,T} \|\Psi + \beta \cdot \nabla \varphi_h - (\mu - \nabla \cdot \beta) \varphi_h\|_T$, $\eta_{2,T} = \alpha_2 \|\beta_n^{\frac{1}{2}} [\varphi_h]\|_{\partial T \setminus \partial\Omega}$, $\eta_{3,T} = (1 - \alpha_2) \gamma_1^{\frac{1}{2}} \|\beta_n^{\frac{1}{2}} |e_\beta \cdot n_T|^{\frac{1}{2}} h_T [\nabla \varphi_h]\|_{\partial T \setminus \partial\Omega}$, and $\eta_{4,T} = \|\beta_n^{\frac{1}{2}} (\varphi_h - \psi)\|_{\partial\Omega^+ \cap \partial T}$, with the weights $\alpha_{1,T} = \min \{ \mu^{-\frac{1}{2}}, h_T^{\frac{1}{2}} \beta_{\infty,T}^{-\frac{1}{2}} \}$, $\alpha_2 = 1$ for $W_h = V_h^k$ and $\alpha_2 = 0$ for $W_h = V_h^C$.

Proof. First apply a Cauchy–Schwarz inequality to the right-hand side of (4.18) to obtain

$$(4.19) \quad \begin{aligned} |\Lambda(u - u_h)| &\leq \left(\|f\|_{\mathcal{T}_h}^2 + \|\beta_n^{\frac{1}{2}} g\|_{\partial\Omega^-}^2 \right)^{\frac{1}{2}} \\ &\quad \times \left(\|\varphi - \varphi_h\|_{\mathcal{T}_h}^2 + \|\beta_n^{\frac{1}{2}} (\varphi - \varphi_h)\|_{\partial\Omega^-}^2 \right)^{\frac{1}{2}}. \end{aligned}$$

Now use Theorem 4.3, with $\Xi = 0$, to derive an upper bound of the dual approximation error in the right-hand side of (4.19). \square

Remark 5. Upper bounds of the type given in Corollaries 4.4 and 4.5 may be derived for the goal-oriented case as well.

5. Numerical example. In this section we will present some numerical experiences using the a posteriori error estimate of Theorem 4.3. We will apply both the DG method and the CIP method to a problem with known solution. A simple adaptive algorithm is proposed, and computations are performed both on uniformly refined unstructured meshes and on adaptive meshes. For all the computations, we have used the finite element package FreeFem++ [21].

First, we introduce the model problem chosen for the study. Define the domain by $\Omega := \{(x, y) \in (0, 1) \times (0, 1)\}$. Consider the problem of seeking $u : \Omega \rightarrow \mathbb{R}$ such that

$$\begin{cases} \mu u + \beta \cdot \nabla u = 0 & \text{in } \Omega, \\ u = g & \text{on } \partial\Omega^-, \end{cases}$$

with

$$\beta(x, y) = \begin{pmatrix} y + 1 \\ -x \end{pmatrix} \frac{1}{\sqrt{x^2 + (y + 1)^2}}, \quad \mu = 0.1,$$

and $g(x, y)$ chosen so that the exact solution writes

$$u(x, y) = e^{\mu \sqrt{x^2 + (y+1)^2} \arcsin\left(\frac{y+1}{\sqrt{x^2 + (y+1)^2}}\right)} \arctan\left(\frac{\sqrt{x^2 + (y+1)^2} - 1.5}{\epsilon}\right).$$

The method (3.1) was used with

- DG0, $W_h = V_h^0$, $\gamma_0 = 0.5$, and $\gamma_1 = 0$;
- DG1, $W_h = V_h^1$, $\gamma_0 = 0.5$, and $\gamma_1 = 0$;
- CIP, $W_h = V_h^C$, $\gamma_0 = 0$, and $\gamma_1 = 0.01$.

To test how well the estimators work with adaptive refinement, we implemented a simple adaptive algorithm: all elements are refined for which $\eta_T^2 \geq 0.5 \max_{T \in \mathcal{T}_h} \eta_T^2$, where $\eta_T^2 := \sum_{i=1}^4 \eta_{i,T}^2$.

We present results on two problems with different stiffness of the layer. In the first case we choose $\epsilon = 0.01$. Here the layer is fully resolved under refinement, and the methods are optimally convergent. In the second case we take $\epsilon = 10^{-8}$, and the layer is never resolved on the scales considered here. Indeed in the second case the solution can be considered discontinuous on the meshsizes used.

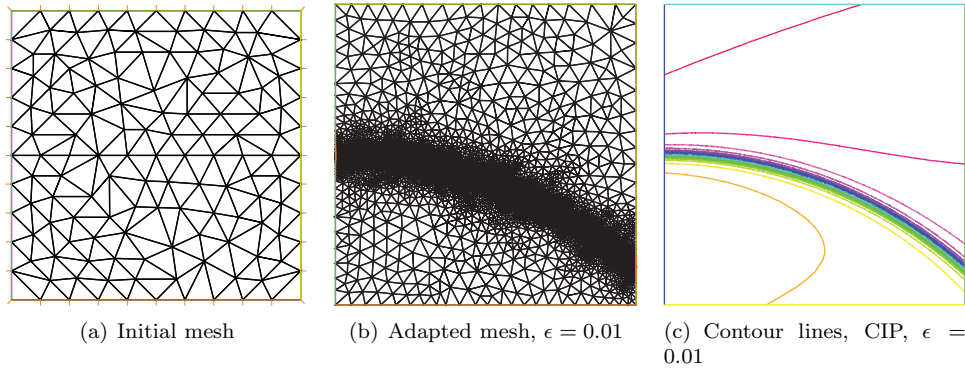


FIG. 2. Some figures illustrating the problem setup.

TABLE 5.1

Symbols used to distinguish the different methods in the graphics.

Method	Adaptive	Uniform
CIP	\diamond	filled \diamond
DG0	Δ	filled Δ
DG1	\square	filled \square
Galerkin	\circ	never used

TABLE 5.2

Guide to line types used to indicate different error quantities.

Error quantity	$\ i_h^k u - u_h\ _\Omega$	$\ \beta_n^{\frac{1}{2}} [i_h^k u - u_h]\ _{\mathcal{F}_h}$	$\ h^{\frac{1}{2}} \beta_n^{\frac{1}{2}} e_\beta \cdot \nabla(u - u_h)\ _\Omega$	$\eta(u_h)$
Line type	dotted	dash-dot	dashed	full

The initial mesh, an adapted mesh, and the contourlines of a solution are presented in Figure 2; the initial meshsize is $\tilde{h} \approx 0.1$.

In the a posteriori error estimate Theorem 4.3, we have set the constant $c_\delta = 1$ in all the cases considered. The efficiency index is defined by

$$\xi_{eff} = \frac{\eta(u_h)}{\|u - u_h\|_{h,0,\Xi}}.$$

In view of Remark 1 we have taken $\mu_0 = 1$ in $\|u - u_h\|_{h,0,\Xi}$. To study when the saturation assumption is satisfied, we have also performed computations on unstructured but uniformly refined meshes in both cases. In the smooth case the saturation assumption holds for the quantity $\|u - u_h\|_{h,0,1}$, but in the nonsmooth case it fails for the error in the streamline derivative. In that case we have used $\Xi = 0$ in the computation of the efficiency index. In all graphics the error quantities or efficiency indices have been plotted against the number of degrees of freedom N . Explications to the symbols in the graphics are given in Tables 5.1 and 5.2.

5.1. Smooth case, $\epsilon = 0.01$. In this case the problem exhibits a sharp layer, which can be completely resolved with a reasonable number of degrees of freedom. To-

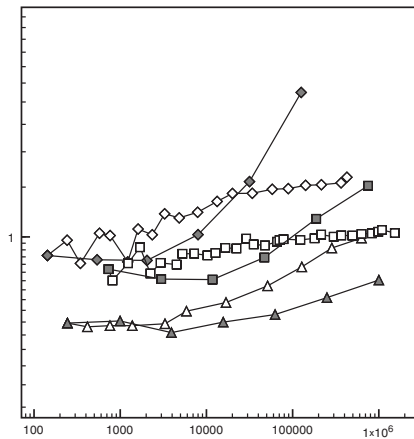
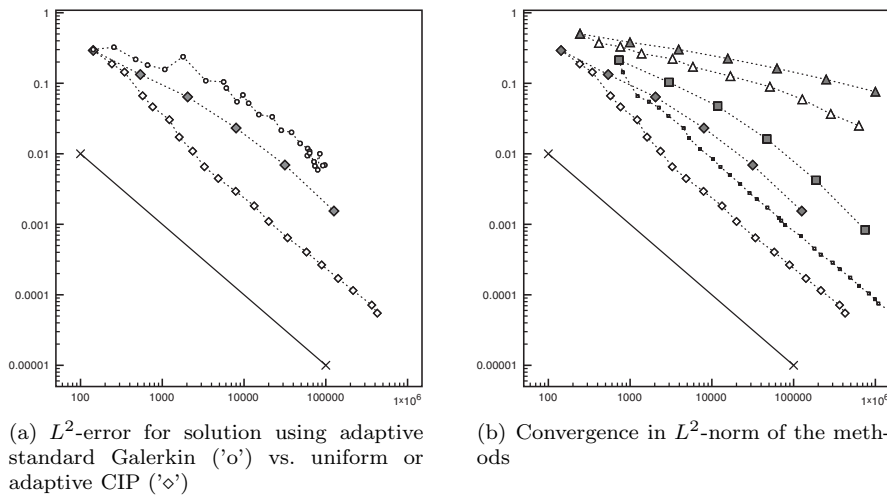
(c) Comparison of efficiency indices $\epsilon = 10^{-2}$

FIG. 3. Errors and efficiencies for smooth solutions, $\epsilon = 0.01$. See Tables 5.1 and 5.2 for a key to the graphics. All quantities are plotted against the number of degrees of freedom N . The full line without markers has slope N^{-1} , which is the optimal convergence in the L^2 -norm for smooth problems.

show that stabilization is necessary even in this case, we first run the adaptive algorithm using the standard Galerkin method (i.e., (3.1) with $W_h = V_h^C$ and $\gamma_1 = 0$) and compare with the CIP method using either uniform refinement or adaptive refinement. In Figure 3(a) we see that the standard Galerkin method with adaptivity does not manage to match the CIP method using uniform refinement. In Figure 3(b) we plot the error in the L^2 -norm for the three stabilized methods considered, both using adaptive refinement and using uniform refinement. The adaptive CIP method needs the least number of degrees of freedom to arrive at a certain precision. In Figures 4 and 5 we give the convergence of the error estimator and all the error quantities for uniform and adaptive refinement for all three methods. For reference, we give the slope of optimal convergence $N^{-1/2}$ for DG0 and N^{-1} for DG1 and CIP, where N denotes the number of degrees of freedom in the computation. These orders are

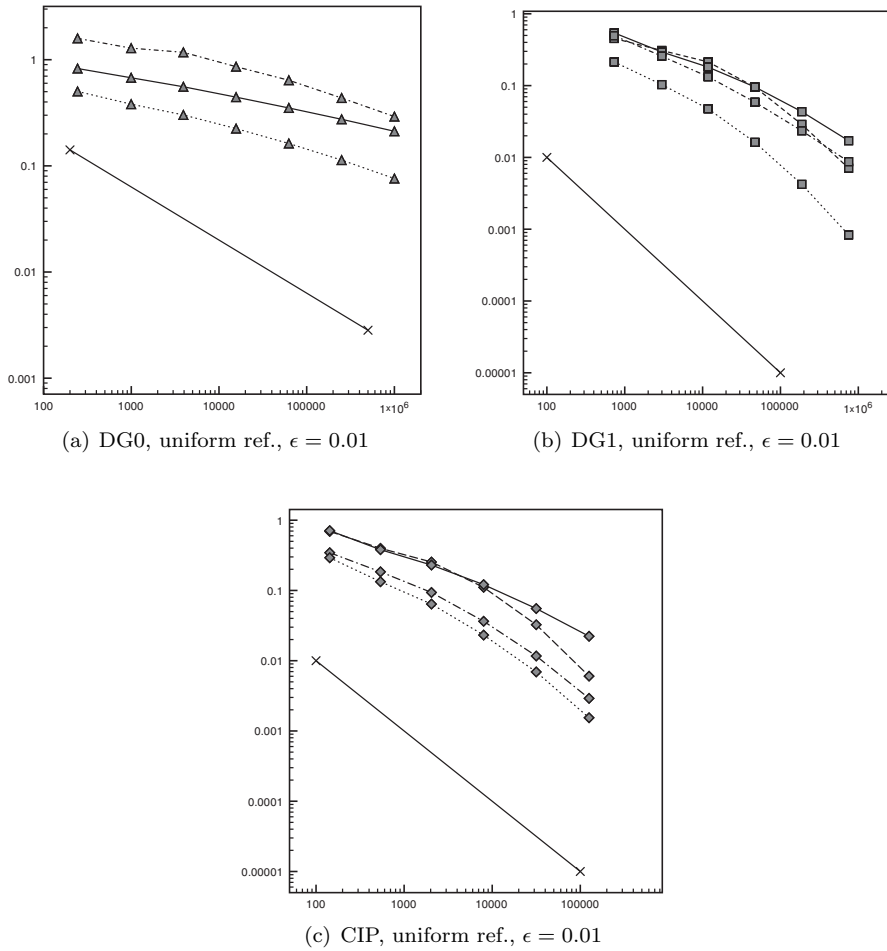


FIG. 4. Convergence under uniform refinement, $\epsilon = 0.01$. See Tables 5.1 and 5.2 for a key to the graphics. All quantities are plotted against the number of degrees of freedom N . The full line without markers represents the slope $N^{-1/2}$ for DG0 and N^{-1} for DG1 and CIP.

indicated in the graphics with full lines with \times -symbols at the endpoints, both in the cases of uniform and adaptive refinement. DG0 only reaches its optimal convergence at the end of the adaptive refinement.

When $\epsilon = 0.01$ and the problem is solved using uniform refinement, there is a plateau during which convergence is rather poor and the optimal convergence rate is obtained only as the layer is resolved. This can be seen in Figure 4(b) for DG1 and 4(c) for CIP. DG0 never reaches the asymptotic range when uniform refinement is used. Some discrete errors show superconvergence on uniformly refined meshes, probably because the finite element approximation is compared to the interpolant of the exact solution. Note in particular the superconvergence of the error in the streamline derivative for the CIP method when the meshes enter the asymptotic range. In the adaptive case, Figure 5, the plateau due to underresolution is no longer visible and both DG1, and CIP has optimal convergence in the L^2 -norm already on the coarsest meshes.

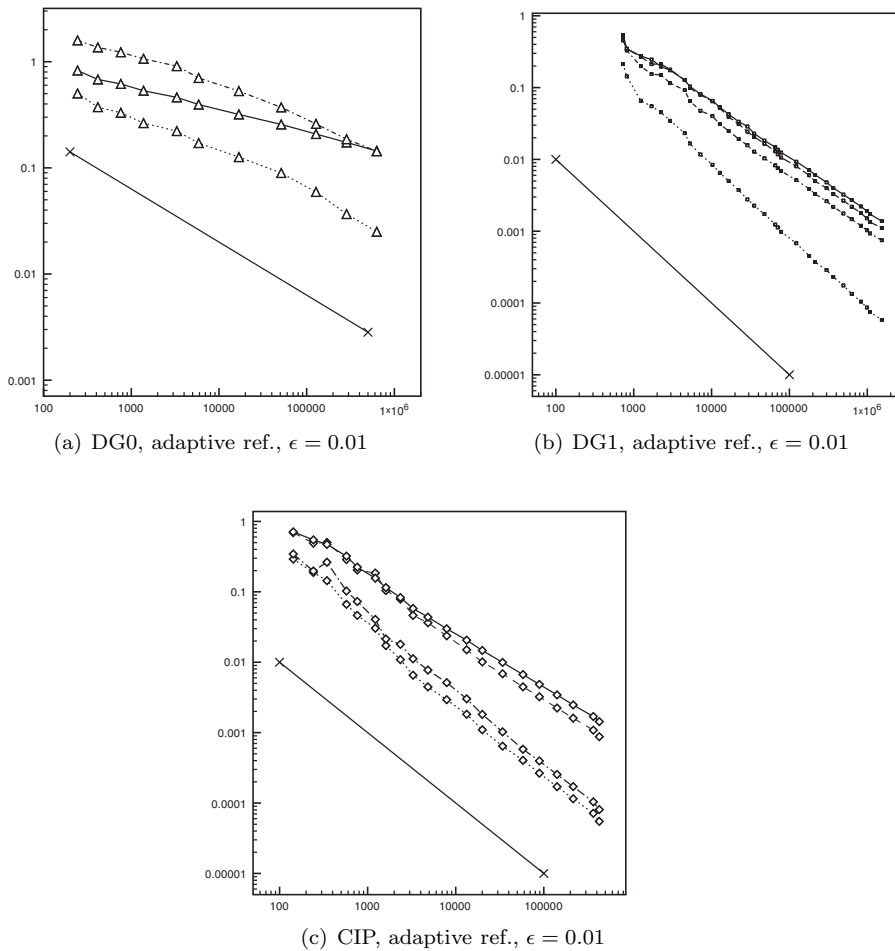


FIG. 5. Convergence under adaptive refinement, $\epsilon = 0.01$. See Tables 5.1 and 5.2 for a key to the graphics. All quantities are plotted against the number of degrees of freedom N . The full line without markers represents the slope $N^{-1/2}$ for DG0 and N^{-1} for DG1 and CIP.

The efficiency indices (with $\Xi = 1$) are given in Figure 3(c). The DG1 method using adaptive refinement has an efficiency index very close to 1 and DG0 with adaptive refinement also seems to converge to 1. The CIP method with adaptive refinement has an efficiency index of around 1.5. The strong growth of the efficiency index when using uniform refinement is due to the above-mentioned superconvergence.

5.2. Discontinuous case, $\epsilon = 10^{-8}$. Here the layer is so stiff that it is never resolved and can be considered as a discontinuity. Moreover, using uniform refinement, we show in Figure 6(a) that the saturation assumption [SA] is not satisfied for the error in the streamline derivative, neither for the CIP method nor for the DG1 method. Therefore, we consider only the a posteriori error estimate of Theorem 4.3 in the case $\Xi = 0$, and we compute the efficiency index using $\Xi = 0$. Although the estimator fails to capture the error in the streamline derivative, it does a good job of predicting the convergence of the error in the L^2 -norm in the interior and on the boundary, as can be seen in the efficiency indices of Figure 6(c). In Figure 6(b) we compare

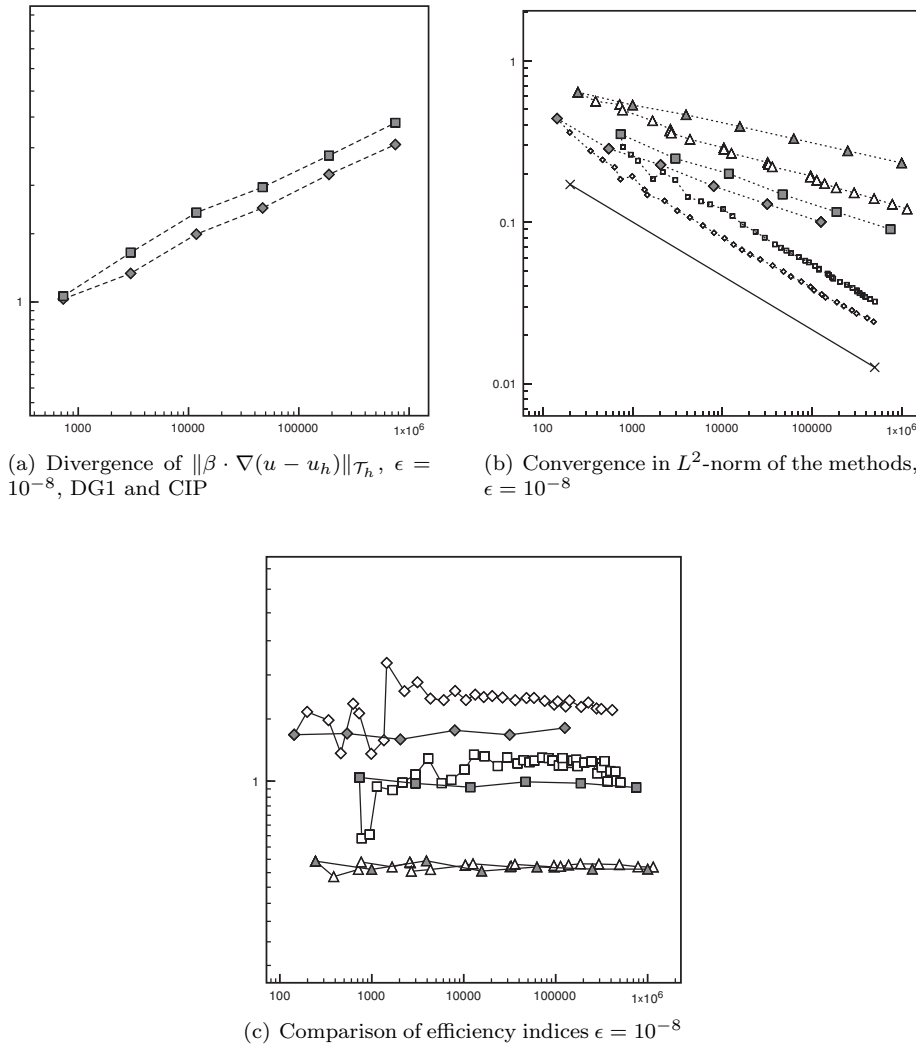


FIG. 6. Errors and efficiencies, $\epsilon = 10^{-8}$. See Tables 5.1 and 5.2 for a key to the graphics.

the error in the L^2 -norm for the different methods. In Figures 7 and 8 we give the values of the estimator $\eta(u_h)$ and the error quantities for all three methods ((a) DG0; (b) DG1; (c) CIP), both using uniform and adaptive refinement. The adaptive algorithm gives the order $O(N^{-\frac{1}{3}})$ when using the DG1 method and the CIP method, but only $O(N^{-\frac{1}{5}})$ for DG0. These orders are indicated in the graphics with full lines with \times -symbols at the endpoints, both in the cases of uniform and adaptive refinement.

6. Conclusion. In this paper we have derived a posteriori error estimates under a saturation assumption using an h -weighted graph norm. In the case where the saturation assumption holds for the graph norm, optimal upper and lower bounds are proven. For discontinuous solutions, numerical investigations show that the saturation assumption may fail for the error in the streamline derivative. In this case we proved upper and lower bounds for the L^2 -error only.

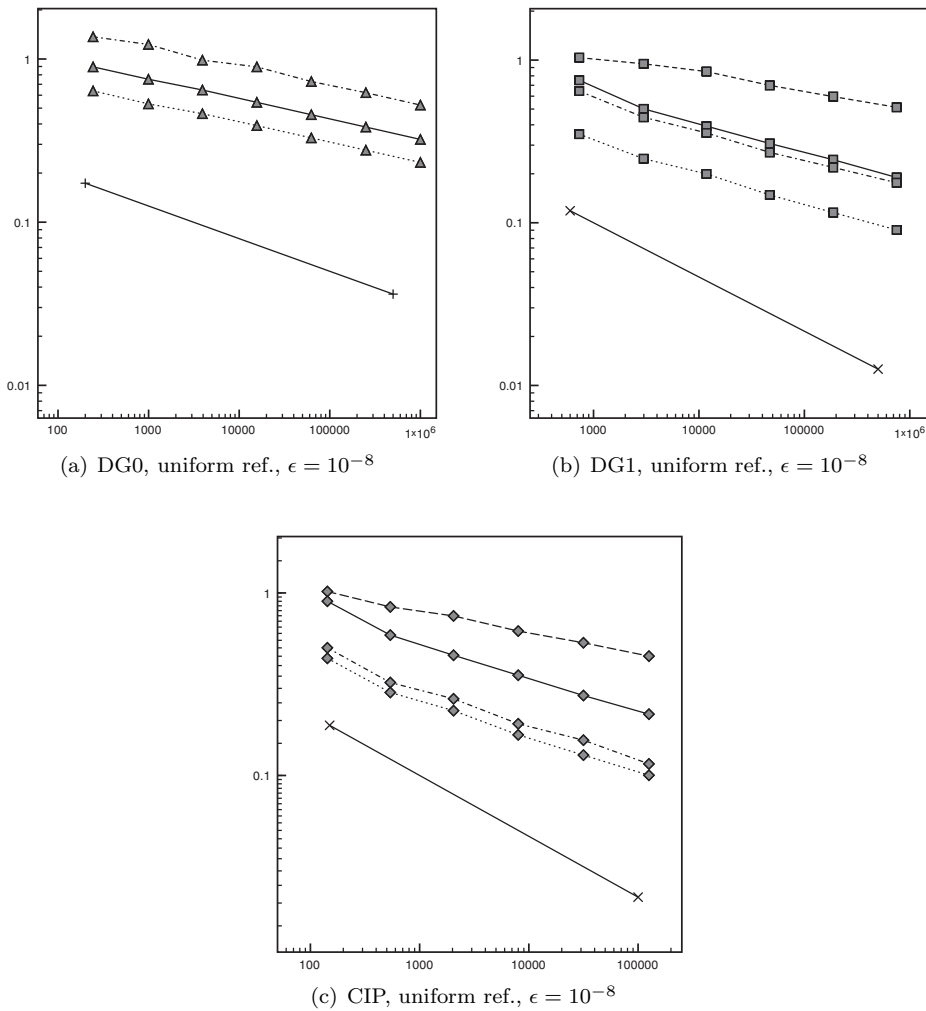


FIG. 7. Convergence under uniform refinement, $\epsilon = 10^{-8}$. See Tables 5.1 and 5.2 for a key to the graphics. All quantities are plotted against the number of degrees of freedom N . The full line without markers represents the slope $N^{-1/5}$ for DG0 and $N^{-1/3}$ for DG1 and CIP.

We then proved a goal-oriented a posteriori error estimate that is easily computable. In this case the saturation assumption must hold for the dual problem and introduces an unknown constant. Note that in this case the convergence of the estimator depends only on the regularity of the dual problem. If the data of the dual problem are smooth, indicating that the goal quantity is some average, the finite element solution of the dual problem is expected to have optimal convergence. This gives some more evidence that average quantities are easier to compute than norms of the error. Indeed averages are no more difficult to compute for very irregular u than for regular ones.

We have studied the performance of the above estimator numerically on a model problem and have found excellent agreement between the computed and estimated error.

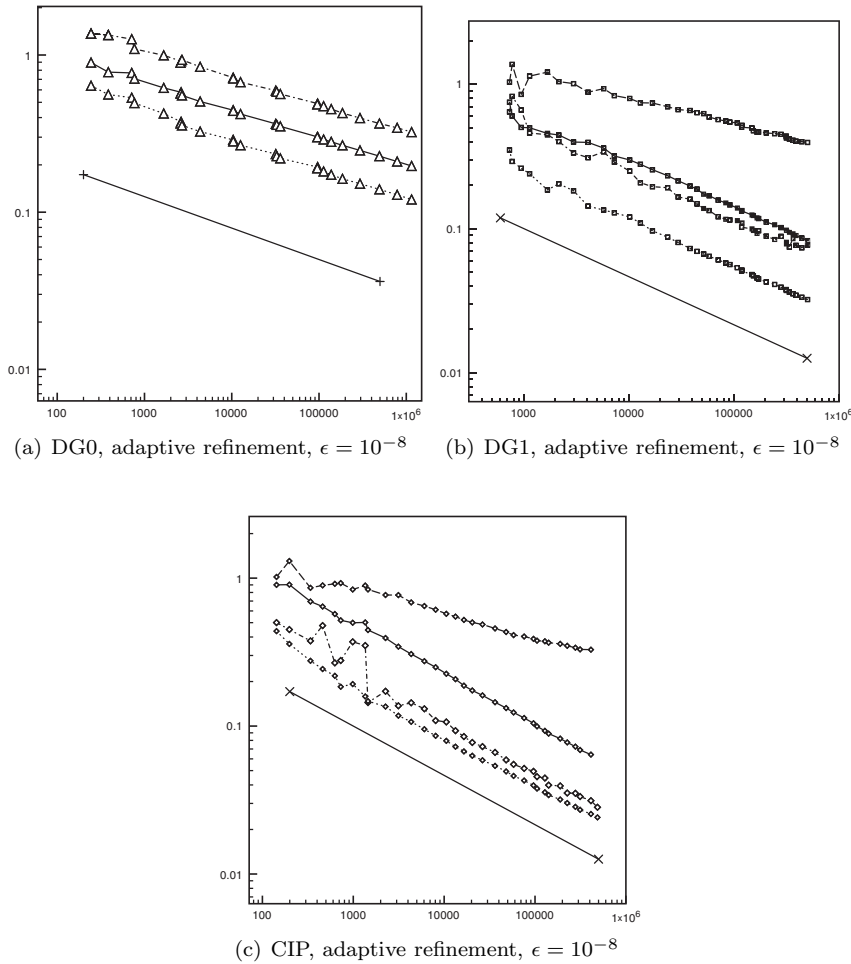


FIG. 8. Convergence under adaptive refinement, $\epsilon = 10^{-8}$. See Table 5.1 and 5.2 for a key to the graphics. All quantities are plotted against the number of degrees of freedom N . The full line without dots represents the slope $N^{-1/5}$ for DG0 and $N^{-1/3}$ for DG1 and CIP.

All the results use the control of $\|\tilde{h}^{\frac{1}{2}}\beta \cdot \nabla u_h\|_{\mathcal{T}_h}$ in a crucial way. Numerically, we give evidence that if this control is lacking, i.e., when using a standard Galerkin method, the above derived estimator fails to produce a reliable adaptive algorithm.

REFERENCES

- [1] B. ACHCHAB, S. ACHCHAB, AND A. AGOUZAL, *Some remarks about the hierarchical a posteriori error estimate*, Numer. Methods Partial Differential Equations, 20 (2004), pp. 919–932.
- [2] I. BABUŠKA, A. MILLER, AND M. VOGELIUS, *Adaptive methods and error estimation for elliptic problems of structural mechanics*, in Adaptive Computational Methods for Partial Differential Equations, SIAM, Philadelphia, 1983, pp. 57–73.
- [3] R. E. BANK AND A. WEISER, *Some a posteriori error estimators for elliptic partial differential equations*, Math. Comp., 44 (1985), pp. 283–301.
- [4] R. BECKER, P. HANSBO, AND R. STENBERG, *A finite element method for domain decomposition with non-matching grids*, M2AN Math. Model. Numer. Anal., 37 (2003), pp. 209–225.

- [5] S. BERRONE, *Robustness in a posteriori error analysis for FEM flow models*, Numer. Math., 91 (2002), pp. 389–422.
- [6] M. BOMAN, *Estimates for the L_2 -projection onto continuous finite element spaces in a weighted L_p -norm*, BIT, 46 (2006), pp. 249–260.
- [7] D. BRAESS AND R. VERFÜRTH, *A posteriori error estimators for the Raviart–Thomas element*, SIAM J. Numer. Anal., 33 (1996), pp. 2431–2444.
- [8] E. BURMAN, *A unified analysis for conforming and nonconforming stabilized finite element methods using interior penalty*, SIAM J. Numer. Anal., 43 (2005), pp. 2012–2033.
- [9] E. BURMAN AND A. ERN, *A continuous finite element method with face penalty to approximate Friedrichs’ systems*, M2AN Math. Model. Numer. Anal., 41 (2007), pp. 55–76.
- [10] E. BURMAN, J. GUZMÁN, AND D. LEYKHEMAN, *Weighted error estimates of the continuous interior penalty method for singularly perturbed problems*, IMA J Numer. Anal., 29 (2008), pp 284–314, doi:10.1093/imanum/drn001.
- [11] E. BURMAN AND P. HANSBO, *Edge stabilization for Galerkin approximations of convection-diffusion-reaction problems*, Comput. Methods Appl. Mech. Engrg., 193 (2004), pp. 1437–1453.
- [12] P. G. CIARLET, *The Finite Element Method for Elliptic Problems*, Classics Appl. Math. 40, SIAM, Philadelphia, 2002, Reprint of the 1978 original (North-Holland, Amsterdam; MR0520174 (58 #25001)).
- [13] W. DÖRFLER AND R. H. NOCHETTO, *Small data oscillation implies the saturation assumption*, Numer. Math., 91 (2002), pp. 1–12.
- [14] J. DOUGLAS AND T. DUPONT, *Interior penalty procedures for elliptic and parabolic Galerkin methods*, in Computing Methods in Applied Sciences (Second International Symposium, Versailles, 1975), Lecture Notes in Phys. 58, Springer, Berlin, 1976, pp. 207–216.
- [15] L. EL ALAOU, A. ERN, AND E. BURMAN, *A priori and a posteriori analysis of non-conforming finite elements with face penalty for advection-diffusion equations*, IMA J. Numer. Anal., 27 (2007), pp. 151–171.
- [16] K. ERIKSSON AND C. JOHNSON, *Adaptive streamline diffusion finite element methods for stationary convection-diffusion problems*, Math. Comp., 60 (1993), pp. 167–188.
- [17] A. ERN AND J.-L. GUERMOND, *Theory and Practice of Finite Elements*, Appl. Math. Sci. 159, Springer, New York, 2004.
- [18] A. ERN AND J.-L. GUERMOND, *Discontinuous Galerkin methods for Friedrichs’ systems. I. General theory*, SIAM J. Numer. Anal., 44 (2006), pp. 753–778.
- [19] K. O. FRIEDRICHS, *Symmetric positive linear differential equations*, Comm. Pure Appl. Math., 11 (1958), pp. 333–418.
- [20] J. GUZMÁN, *Local analysis of discontinuous Galerkin methods applied to singularly perturbed problems*, J. Numer. Math., 14 (2006), pp. 41–56.
- [21] F. HECHT AND O. PIRONNEAU, *FreeFEM++ Manual*, Laboratoire Jacques Louis Lions, Paris, 2005.
- [22] P. HOUSTON, R. RANNACHER, AND E. SÜLI, *A posteriori error analysis for stabilised finite element approximations of transport problems*, Comput. Methods Appl. Mech. Engrg., 190 (2000), pp. 1483–1508.
- [23] G. KUNERT, *A posteriori error estimation for convection dominated problems on anisotropic meshes*, Math. Methods Appl. Sci., 26 (2003), pp. 589–617.
- [24] P. D. LAX AND R. S. PHILLIPS, *Local boundary conditions for dissipative symmetric linear differential operators*, Comm. Pure Appl. Math., 13 (1960), pp. 427–455.
- [25] P. LESAINTE AND P.-A. RAVIART, *On a finite element method for solving the neutron transport equation*, in Mathematical Aspects of Finite Elements in Partial Differential Equations (Proc. Symposium, Math. Res. Center, Univ. Wisconsin, Madison, Wis., 1974) Publication 33, Math. Res. Center, Univ. of Wisconsin-Madison, Academic Press, New York, 1974, pp. 89–123.
- [26] K. MEKCHAY AND R. H. NOCHETTO, *Convergence of adaptive finite element methods for general second order linear elliptic PDEs*, SIAM J. Numer. Anal., 43 (2005), pp. 1803–1827.
- [27] G. RAPIN AND G. LUBE, *A stabilized scheme for the Lagrange multiplier method for advection-diffusion equations*, Math. Models Methods Appl. Sci., 14 (2004), pp. 1035–1060.
- [28] W.H. REED AND T.R. HILL, *Triangular mesh methods for the neutron transport equation*, Technical report LA-UR-73-479, Los Alamos Scientific Laboratory, Los Alamos, NM, 1973.
- [29] G. SANGALLI, *Robust a posteriori estimator for advection-diffusion-reaction problems*, Math. Comp., 77 (2008), pp. 41–70.
- [30] E. SÜLI AND P. HOUSTON, *Finite element methods for hyperbolic problems: A posteriori error analysis and adaptivity*, in The State of the Art in Numerical Analysis, Inst. Math. Appl. Conf. Ser. New Ser. 63, Oxford University Press, New York, 1997, pp. 441–471.

- [31] R. VERFÜRTH, *A posteriori error estimators for convection-diffusion equations*, Numer. Math., 80 (1998), pp. 641–663.
- [32] R. VERFÜRTH, *Robust a posteriori error estimates for stationary convection-diffusion equations*, SIAM J. Numer. Anal., 43 (2005), pp. 1766–1782.
- [33] M. VOHRALIK, *A posteriori error estimates for lowest order mixed finite element discretizations of convection-diffusion-reaction equations*, SIAM J. Numer. Anal., 45 (2007), pp. 1570–1599.
- [34] B. I. WOHLMUTH, *Hierarchical a posteriori error estimators for mortar finite element methods with Lagrange multipliers*, SIAM J. Numer. Anal., 36 (1999), pp. 1636–1658.