Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Comput. Methods Appl. Mech. Engrg. 193 (2004) 1437–1453
Computer methods in applied mechanics and engineering
www.elsevier.com/locate/cma

Edge stabilization for Galerkin approximations of convection–diffusion–reaction problems

Erik Burman [a], Peter Hansbo [b,*]

[a] *Department of Mathematics, Ecole Polytechnique Fédérale de Lausanne, Switzerland*
[b] *Department of Applied Mechanics, Chalmers University of Technology, SE–41296 Göteborg, Sweden*

Received 15 December 2002; received in revised form 6 September 2003; accepted 9 December 2003

Abstract

In this paper we recall a stabilization technique for finite element methods for convection–diffusion–reaction equations, originally proposed by Douglas and Dupont [Computing Methods in Applied Sciences, Springer-Verlag, Berlin, 1976]. The method uses least square stabilization of the gradient jumps a across element boundaries. We prove that the method is stable in the hyperbolic limit and prove optimal a priori error estimates. We address the question of monotonicity of discrete solutions and present some numerical examples illustrating the theoretical results.

Keywords: Stabilized methods; Finite element; Interior penalty

1. Introduction

The standard Galerkin for convection–diffusion–reaction problems is not stable if implemented without stabilization. Over the years many different stabilization methods have been proposed and it is by now a well-established discipline with different well-explored methods like the SUPG/SD-method [8], the residual free bubbles [2] and more recent contributions like subviscosity models for convection diffusion problems [7]. The relation between the different approaches is also well understood in most cases. However for complex flow problems, like the ones arising in combustion problems, most of these methods have drawbacks. The SUPG stabilization becomes non-symmetric and the formulation does not permit lumped mass; the residual free bubbles add additional degrees of freedom; the projection methods introduce the need of hierarchical meshes for the projection or the sub viscosity model. In this paper we recall a method due to Douglas and Dupont [6] which stabilizes convection–diffusion–reaction problems by adding a least-squares term based on the jump in the gradient over element boundaries. Unlike [6], we also consider the crucial case of a vanishing diffusion parameter.

* Corresponding author. Fax: +31-772-3872.
E-mail address: hansbo@solid.chalmers.se (P. Hansbo).

0045-7825/$ - see front matter
doi:10.1016/j.cma.2003.12.032

The method can be seen as a continuous, higher order interior penalty method. We also add a non-linear term adding diffusion on the element edges in the tangential direction, in order to guarantee monotonicity. We prove that the shock-capturing parameter can be chosen in such a way that a discrete maximum principle holds. The method has many of the advantages of the above methods, but no additional degrees of freedom are added, no hierarchical meshes are needed, the formulation remains symmetric, and the mass can be lumped for efficient time marching and treatment of stiff source terms. Furthermore the method allows for the introduction of crosswind diffusion which is consistent for solutions in $H^2(\Omega)$. The price to pay is an increased number of non-zero elements in the stiffness matrix due to the fact that the gradient jump term couple neighboring elements. However for systems of PDEs (like the ones in combustion problems) where a large number of unknowns are associated with each node, these additional blocks are diagonal, making the increased memory cost reasonable.

From a practical point of view, the difference in implementation compared to a standard finite element code is the need for a data structure containing the elements neighboring to a given element; such a data structure (needed for computing the gradient jumps) is typically not required in a standard code.

2. Convection–diffusion–reaction

As a first model problem, we consider, in $\Omega \subset \mathbb{R}^d$, $d = 2, 3$, the problem of solving

$$\sigma u + \beta \cdot \nabla u - \nabla \cdot (\varepsilon \nabla u) = f \quad \text{in } \Omega \tag{1}$$

with, for simplicity, $u = 0$ on $\partial\Omega$. Here, f is a given source term, β is a given smooth velocity field, satisfying $\nabla \cdot \beta = 0$, and σ and ε are bounded positive functions.

The weak form of this problem is to find $u \in H_0^1(\Omega)$ such that

$$A(u, v) = (f, v) \quad \forall v \in H_0^1(\Omega), \tag{2}$$

where

$$A(u, v) := \int_\Omega (\sigma uv + \varepsilon \nabla u \cdot \nabla v + \beta \cdot \nabla uv)\, \mathrm{d}x \quad \text{and} \quad (f, v) := \int_\Omega fv\, \mathrm{d}x.$$

We denote the L_2-scalar product by $(\cdot,\cdot)$ and the corresponding norm by $\|\cdot\|$. The finite element method consists of seeking a piecewise polynomial approximation $U \in V_h \subset H_0^1(\Omega)$. It is well known that the standard Galerkin approximation, in the convection dominated case, results in a wildly oscillating solution in the presence of sharp layers. To stabilize the method we propose, following [6], to add a term penalizing the gradient jumps across element boundaries of the type

$$J(U, v) = \sum_K \frac{1}{2} \int_{\partial K} \gamma h_{\partial K}^2 [\nabla U] \cdot [\nabla v]\, \mathrm{d}s = \sum_K \frac{1}{2} \int_{\partial K} \gamma h_{\partial K}^2 [n \cdot \nabla U][n \cdot \nabla v]\, \mathrm{d}s. \tag{3}$$

Here, $h_{\partial K}$ is the size of ∂K, $[q]$ denotes the jump of q across ∂K for $\partial K \cap \partial\Omega = \emptyset$, $[q] = 0$ on $\partial K \cap \partial\Omega \neq \emptyset$, n is the outward pointing unit normal to K, and γ is a constant. We also introduce the local mesh size

$$h_K := \max_K h_{\partial K},$$

and we will assume that $h_K / h_{\partial K} < C$ where C is a fixed constant. Our finite element method then reads, find $U \in V_h$ such that

$$A(U, v) + J(U, v) = (f, v) \quad \forall v \in V_h. \tag{4}$$

To simplify the analysis we will assume that the exact solution belongs to $H^2(\Omega)$; it then follows that the formulation (4) is consistent, as put forth in the following lemma:

Lemma 1. *For $u \in H^2(\Omega)$ there holds*

$$A(u - U, v) + J(u - U, v) = 0$$

for all $v \in V^h$.

Proof. This is an immediate consequence of the regularity hypothesis: if $u \in H^2(\Omega)$ then the trace of ∇u is well defined and hence $J(u, v) = 0$. $\square$

Remark 2. Another possible choice of $J(U, v)$ is

$$J(U,v) = \sum_K \frac{1}{2} \int_{\partial K} \gamma_\beta h_{\partial K}^2 [\beta \cdot \nabla U][\beta \cdot \nabla v] \, \mathrm{d}s + \sum_K \frac{1}{2} \int_{\partial K} \gamma_{\beta^\perp} h_{\partial K}^2 [\beta^\perp \cdot \nabla U][\beta^\perp \cdot \nabla v] \, \mathrm{d}s. \tag{5}$$

This way the streamline and the crosswind stabilizations may be tuned independently. Note that (3) corresponds to the case $\gamma_\beta = \gamma_{\beta^\perp}$.

2.1. Stability

The main point of any stabilized method is of course that it enhances stability. The stability estimate obtained using edge stabilization is less immediate than that obtained in the case of streamline-diffusion or discontinuous Galerkin. However we will show that we, thanks to the term $J(U, v)$, get the control of $\|h_K^{1/2} \beta \cdot \nabla U\|^2$ crucial for the analysis. To prove stability in a discontinuous Galerkin method one exploits the fact that $h_K \beta \cdot \nabla U$ is in the finite element test space and hence can be chosen as test function. In the case of edge stabilization we proceed in a similar way. Indeed, even if $h_K \beta \cdot \nabla U$ is not in the finite element space something which is close is, and the difference is controlled by the edge stabilization term. We denote by π_h the Clément quasi-interpolant [5], $\pi_h : L_2(\Omega) \to V_h$.

We shall frequently use the following inequalities, which we collect in a lemma.

Lemma 3. *For the Clément operator there holds*

$$\|\pi_h u\|_{H^s(\Omega)} \leqslant C \|u\|_{H^s(\Omega)} \quad \forall u \in H^s(\Omega), \tag{6}$$

for $s = 0, 1$. Further,

$$\|\pi_h h_K \beta \cdot \nabla U\| \leqslant C \|U\| \quad \forall U \in V_h. \tag{7}$$

Finally, we have the trace inequality

$$\|v\|_{L_2(\partial K)}^2 \leqslant C(h_K^{-1} \|v\|_{L_2(K)}^2 + h_K \|v\|_{H^1(K)}^2) \quad \forall v \in H^1(K). \tag{8}$$

Here, C is a generic constant independent of h_K.

Proof. Inequality (6) follows from the interpolation estimate

$$\|u - \pi_h u\|_{H^s(\Omega)} \leqslant C \|u\|_{H^s(\Omega)}, \quad s = 0, 1,$$

cf. [5], and (7) follows from (6) and the well-known inverse inequality

$$\|v\|_{H^1(K)} \leqslant C h_K^{-1} \|v\|_{L_2(K)} \quad \forall v \in V_h. \tag{9}$$

Finally, a proof of (8) is given in [9]. $\square$

As a model example we choose $\epsilon = 0$ and we assume that $h_K < \sigma^{-1/2}$ corresponding to a convection–reaction problem. Furthermore let us assume that h_K is constant throughout the domain. The problem takes the form: find $U \in V_h$ such that

$$(\beta \cdot \nabla U, v) + (\sigma U, v) + J(U, v) = (f, v) \quad \forall v \in V_h. \tag{10}$$

Taking $v = U$ we obtain the basic stability estimate

$$J(U, U) + \|\sigma^{1/2} U\|^2 = (f, U). \tag{11}$$

Clearly we may use the fact that $\pi_h h_K \beta \cdot \nabla U \in V_h$ to write

$$\|h_K^{1/2} \beta \cdot \nabla U\|^2 + (\beta \cdot \nabla U, \pi_h h_K \beta \cdot \nabla U - h_K \beta \cdot \nabla U) = -J(U, \pi_h h_K \beta \cdot \nabla U) + (-\sigma U + f, \pi_h h_K \beta \cdot \nabla U). \tag{12}$$

We use Cauchy–Schwartz inequality followed by the arithmetic–geometric inequality for the left-hand side to obtain

$$\tfrac{3}{4}\|h_K^{1/2} \beta \cdot \nabla U\|^2 - \|h_K^{1/2}(\pi_h \beta \cdot \nabla U - \beta \cdot \nabla U)\|^2 \leqslant |J(U, \pi_h h_K \beta \cdot \nabla U)| + C\|\sigma^{1/2} U\|^2 + C\|f\|^2. \tag{13}$$

Comparing the two expressions (11) and (13) we find that we need the following two results:

(1) Proof that there exists some $\zeta \geqslant \zeta_0 > 0$ such that

$$\|h_K^{1/2}(\pi_h \beta \cdot \nabla U - \beta \cdot \nabla U)\|^2 \leqslant \zeta J(U, U).$$

(2) The inverse estimate

$$J(\pi_h h_K \beta \cdot \nabla U, \pi_h h_K \beta \cdot \nabla U) \leqslant C_i \|h_K^{1/2} \beta \cdot \nabla U\|^2. \tag{14}$$

The inverse estimate is immediately proven by noting that

$$J(\pi_h h_K \beta \cdot \nabla U, \pi_h h_K \beta \cdot \nabla U) = \sum_K \int_{\partial K} h_K^4 [\nabla \pi_h \beta \cdot \nabla U]^2 \,\mathrm{d}s \leqslant \tilde{C} \|h_K^{3/2} \nabla \pi_h \beta \cdot \nabla U\|^2 \leqslant C \|h_K^{1/2} \beta \cdot \nabla U\|^2$$

by virtue of (9) and (6).

2.1.1. Bounding the projection error by the stabilization term

The stability of the method is obtained by the fact that the edge operator controls the projection error of $h_K \beta \cdot \nabla U$ in the case of convection–diffusion. By $\{\varphi_i\}$ we denote the set of finite element basis functions spanning the space V_h. Let $\mathcal{N}_i$ be the set of all triangles K^i containing node i and assume that the cardinality of $\mathcal{N}_i$ is bounded uniformly in i. Let $\mathcal{F}_K$ be the set of all test functions φ_i such that $K \subset \operatorname{supp} \varphi_i$ and $\Omega_i = \bigcup_{\mathcal{N}_i} K^i$. We will consider a function $p \in [P_0(K)]^2$, and its representation in the finite element basis $\tilde{p}$ defined by

$$\tilde{p}|_K = p|_K \sum_{i \in \mathcal{F}_K} \varphi_i. \tag{15}$$

It follows that $\tilde{p} = p$ everywhere except on elements adjacent to Dirichlet boundaries where the boundary nodes are not included in the finite element space. We note that, with $p := h_K^{1/2} \beta \cdot \nabla U$, we have on the left-hand side of (12) the expression $\|p\|^2 + (p, \pi_h p - p)$, and we wish to bound the second term using the first term and the jumps. This cannot be done exactly since $\pi_h p$ must obey the boundary conditions, unlike p. However, the left-hand side of (12) can equally well be written $(p, \tilde{p}) + (p, \pi_h p - \tilde{p})$, and if we can show that $c\|p\|^2 \leqslant (p, \tilde{p})$ we have

$$c\|p\|^2 + (p, \pi_h p - \tilde{p}) \leqslant (p, \tilde{p}) + (p, \pi_h p - \tilde{p}),$$

and we can proceed to bound the second term on the left-hand side in terms of the first together with the jumps. Thus, we need:

Lemma 4. *Suppose that K is an element with at least one node on a Dirichlet boundary then*

$$\|p\|_K^2 = \frac{d+1}{n_i}(p,\tilde{p}), \tag{16}$$

where n_i denotes the number of interior nodes of the element.

Proof. The proof is immediate noting that

$$(p,\tilde{p}) = |p_K|^2 \int_K \sum_{i\in\mathcal{F}_K} \varphi_i \,\mathrm{d}x = \frac{n_i}{d+1}|p_K|^2 m(K).$$

We will now proceed to prove that

$$\|h^{s/2}(\tilde{p} - \pi_h p)\|^2 \leqslant C\tilde{J}_s(p,p)$$

with

$$\tilde{J}_s(p,p) = \sum_K \int_{\partial K} h^{s+1}[p]^2 \,\mathrm{d}s.$$

The operator $\pi_h : [P_0(K)]^2 \to [V_h]^2$, which denotes the lowest order Clément operator is constructed as follows:

$$\pi_h p = \sum_i p_i \varphi_i \tag{17}$$

with

$$p_i = \frac{1}{m(\Omega_i)} \sum_{\mathcal{N}_i} p|_{K^i} m(K^i). \tag{18}$$

In the following we will also write $p|_{K^i} - p|_K = \sum_{K_i}^K [p]$, with $[p]$ denoting the jump across element boundaries and the sum is taken over the shortest "path" from element K^i to element K.

It is now straightforward to show that the projection error is controlled by the operator $\tilde{J}_s(p,p)$,

$$\begin{aligned}
\|h_K^{s/2}(\pi_h p - \tilde{p})\|^2 &= \sum_K \int_K h_K^s \left(\sum_{i\in\mathcal{F}_K} \left(\frac{1}{m(\Omega_i)} \sum_{K^i\in\mathcal{N}_i} p|_{K^i} m(K^i) \right) \varphi_i - \tilde{p} \right)^2 \mathrm{d}x \\
&= \sum_K \int_K h_K^s \left(\sum_{i\in\mathcal{F}_K} \frac{1}{m(\Omega_i)} \sum_{K^i\in\mathcal{N}_i} (p|_{K^i} - p|_K) m(K^i) \varphi_i \right)^2 \mathrm{d}x \\
&\leqslant C \sum_K \int_K h_K^s \sum_{i\in\mathcal{F}_K} \frac{1}{m(\Omega_i)^2} \sum_{K^i\in\mathcal{N}_i} \left\{ \sum_{K_i}^{K} [p] \right\}^2 m(K^i)^2 \,\mathrm{d}x \\
&\leqslant C \sum_K \int_{\partial K} h_K^{s+1} [p]^2 \,\mathrm{d}s \leqslant C\tilde{J}_s(p,p). \qquad \square
\end{aligned}$$

Here we used the upper bound on the number of triangles neighboring to a node and a scaling argument. We have proved the following:

Lemma 5. *If p is some piecewise constant function, $\tilde{p}$ is defined by* (15) *and π_h is the Clément interpolant on V_h, then the edge stabilization term satisfies*

$$\|h_K^{s/2}(\pi_h p - \tilde{p})\|^2 \leqslant \gamma \tilde{J}_s(p,p) \tag{19}$$

for some $\gamma \geqslant \gamma_0 > 0$ independent of h_K.

From this the stability of our method now follows noting that by Lemma 4 we have $c\|p\| \leqslant (p,\tilde{p})$.

Remark 6. Note that by the construction of $\tilde{p}$ we get less stabilization in elements adjacent to Dirichlet boundaries than in the interior of the domain, hence we expect to get poorer stabilizing properties close to sharp out flow layers (when diffusion is present), something which is confirmed by the numerical experiments.

2.1.2. The inf–sup condition

We may now combine the above results to prove a discrete inf–sup condition for our method. We consider the following mesh dependent norm:

$$|||u|||^2 = \|c^{1/2}h_K^{1/2}\beta\cdot\nabla u\|^2 + \|\varepsilon^{1/2}\nabla u\|^2 + \|\sigma^{1/2}u\|^2 + J(u,u). \tag{20}$$

Theorem 7. *Assume that $\sigma < h^{-1/2}$ and $\varepsilon < h$. With the triple norm defined above we have, for some α independent of h,*

$$\alpha|||U||| \leqslant \sup_{w_h\in V_h} \frac{A(U,w_h)+J(U,w_h)}{|||w_h|||} \quad \forall U \in V_h. \tag{21}$$

Proof. The proof is straightforward using the inverse inequalities and Lemma 5 of the previous section. We will start by proving that

$$\tilde{\alpha}|||U|||^2 \leqslant A(U, U + c\pi_h h_K\beta\cdot\nabla U) + J(U, U + c\pi_h h_K\beta\cdot\nabla U). \tag{22}$$

First note that using $(\beta\cdot\nabla U, U) = 0$ we have the coercivity

$$A(U,U) = \|\varepsilon^{1/2}\nabla U\|^2 + \|\sigma^{1/2}U\|^2 + J(U,U). \tag{23}$$

Then by choosing $v = \pi_h h_K\beta\cdot\nabla U$ in (4) we have by adding and subtracting $\|h_K^{1/2}\beta\cdot\nabla U\|^2$ and applying Cauchy–Schwartz inequality, the stability of the projection and an inverse inequality:

$$\begin{aligned} A(U,\pi_h h_K\beta\cdot\nabla U) + J(U,\pi_h h_K\beta\cdot\nabla U) \geqslant{}& \|h_K^{1/2}\beta\cdot\nabla U\|^2 + (\beta\cdot\nabla U, \pi_h h_K\beta\cdot\nabla U - h_K\beta\cdot\nabla U) \\ &- C\|h_K\beta\cdot\nabla U\|A(U,U)^{1/2}, \end{aligned}$$

where the constant C essentially depends on the stability of the projection (6) and the constant in the inverse inequalities (14) and (9). Using Lemma 5 we may control the second term on the right-hand side in the following fashion:

$$(\beta\cdot\nabla U, \pi_h h_K\beta\cdot\nabla U - h_K\beta\cdot\nabla U) \geqslant -\|h_K^{1/2}\beta\cdot\nabla U\|J(U,U)^{1/2}.$$

It follows that

$$A(U,\pi_h h_K\beta\cdot\nabla U) + J(U,\pi_h h_K\beta\cdot\nabla U) \geqslant \frac{1}{2}\|h_K^{1/2}\beta\cdot\nabla U\|^2 - \left(\frac{1}{2}+\frac{1}{2C}\right)A(U,U).$$

We conclude that (22) holds true with $\tilde{\alpha} = \frac{1}{2}$ and $c = \frac{C}{1+C}$. To conclude we need to show that $\exists c$ such that $|||U + C\pi_h h_K \beta \cdot \nabla U||| \leqslant \tilde{c}|||U|||$, but this is immediate by the inverse inequalities (7) and (14) and the claim is proved with $\alpha = \frac{1}{2\tilde{c}}$. $\square$

2.2. A priori error estimates

We now proceed to prove a priori error estimates for the discrete solution using the triple norm and the inf–sup condition defined above. For the a priori analysis we need the following approximation result:

Lemma 8. *The following interpolation estimate holds*:

$$|||u - \pi_h u||| \leqslant C(\varepsilon^{1/2} h + h^{3/2} + \sigma^{1/2} h^2)\|u\|_{H^2(\Omega)}. \tag{24}$$

Proof. The estimates

$$\|\varepsilon^{1/2}\nabla(u - \pi_h u)\|_{L_2(\Omega)} \leqslant Ch\varepsilon^{1/2}\|u\|_{H^2(\Omega)}$$

and

$$\|\sigma^{1/2}(u - \pi_h u)\|_{L_2(\Omega)} \leqslant Ch^2\sigma^{1/2}\|u\|_{H^2(\Omega)}$$

follow from standard interpolation theory. Further, we have, using (8),

$$\|\nabla(u - \pi_h u)\|^2_{L_2(\partial K)} \leqslant C(h_K^{-1}\|\nabla(u - \pi_h u)\|^2_{L_2(K)} + h_K\|u\|^2_{H^2(K)}) \leqslant Ch_K\|u\|^2_{H^2(K)},$$

and it follows by summation that $J(u - \pi_h u, u - \pi_h u)^{1/2} \leqslant Ch^{3/2}\|u\|_{H^2(\Omega)}$. $\square$

Using this interpolation estimate and the consistency we prove the following a priori estimate in the convection dominated case when $\varepsilon < h$.

Theorem 9. *Let $u \in H^2(\Omega)$ be the solution of* (2) *and $U \in V_h$ the finite element solution of* (4). *Then*

$$|||u - U||| \leqslant C(\varepsilon^{1/2} h + h^{3/2} + \sigma^{1/2} h^2)\|u\|_{H^2(\Omega)}. \tag{25}$$

Proof. We decompose the error into

$$|||u - U||| \leqslant |||u - \pi_h u||| + |||U - \pi_h u|||$$

the first part is bounded by Lemma 8 and for the second part we use the inf–sup condition of Theorem 7 and the consistency to obtain, using the notation $\tilde{e} = U - \pi_h u$

$$\alpha|||\tilde{e}||| \leqslant \sup_{w_h \in V_h} \frac{A(\tilde{e}, w_h) + J(\tilde{e}, w_h)}{|||w_h|||} = \sup_{w_h \in V_h} \frac{A(u - \pi_h u, w_h) + J(u - \pi_h u, w_h)}{|||w_h|||},$$

where the nominator may be written

$$\begin{aligned} A(u - \pi_h u, w_h) + J(u - \pi_h u, w_h) &= (\varepsilon\nabla(u - \pi_h u), \nabla w_h) + (\sigma(u - \pi_h u), w_h) + (\beta \cdot \nabla(u - \pi_h u), w_h) \\ &\quad + J(u - \pi_h u, w_h) = \text{(i)} + \text{(ii)} + \text{(iii)} + \text{(iv)}. \end{aligned} \tag{26}$$

We now bound the four contributions. The first and second terms are handled by applying Cauchy–Schwartz inequality followed by the inverse inequality (7)

$$\text{(i)} \leqslant \tilde{c}\|\varepsilon^{1/2}\nabla(u-\pi_h u)\|\|\varepsilon^{1/2}\nabla w_h\| \leqslant \tilde{c}\|\varepsilon^{1/2}\nabla(u-\pi_h u)\|\,|||w_h|||,$$
$$\text{(ii)} \leqslant \|\sigma^{1/2}(u-\pi_h u)\|\|\sigma^{1/2}w_h\| \leqslant \tilde{c}\|\sigma^{1/2}(u-\pi_h u)\|\,|||w_h|||.$$

In the third term we integrate by parts in the $\tilde{e}$ part and use (6) in the second part to obtain

$$\text{(iii)} \leqslant (u-\pi_h u, \beta\cdot\nabla\tilde{w}_h) \leqslant \tilde{c}\|h_K^{-1/2}(u-\pi_h u)\|\|h_K^{1/2}\beta\cdot\nabla w_h\| \leqslant \tilde{c}\|h_K^{-1/2}(u-\pi_h u)\|\,|||w_h|||.$$

For the edge penalty term finally we simply apply Cauchy–Schwartz inequality

$$\text{(iv)} \leqslant \tilde{c}J(u-\pi_h u, u-\pi_h u)^{1/2}J(w_h,w_h)^{1/2} \leqslant cJ(u-\pi_h u, u-\pi_h u)^{1/2}|||w_h|||.$$

The claim now follows by applying Lemma 8. □

It was pointed out in the introduction that the method gives increased control of crosswind derivatives. We may use this to prove an a priori error estimate on the gradient which is superior to that of most stabilized methods where the best estimate on the error of the gradient in L^2-norm is $\|\nabla(u-U)\| \leqslant Ch^{1/2}$ obtained by an inverse inequality. Using the added crosswind diffusion we gain a factor $h^{1/4}$ as proved in the following lemma:

Lemma 10. *With $\sigma > C > 0$, let U be the solution of* (4) *and u a solution to* (1) *fulfilling $u \in H^2(\Omega)$. Then there holds*

$$\|\nabla(u-U)\| \leqslant Ch^{3/4}. \tag{27}$$

Proof. First we note that $\|\nabla(u-U)\| \leqslant \|\nabla(\pi_h u - U)\| + \|\nabla(u-\pi_h u)\|$. Let us now consider the first term, setting $\eta = \pi_h u - U$. Using integration by parts and the fact that $\nabla\eta$ is constant on each element we have

$$\|\nabla\eta\|^2 = \int_\Omega (\nabla\eta)^2\,\mathrm{d}x = \frac{1}{2}\left(\sum_K \int_{\partial K} h_K[\nabla\eta]h_K^{-1}\eta\,\mathrm{d}s\right)^{1/2}.$$

Using Cauchy–Schwartz inequality we obtain

$$\|\nabla\eta\|^2 \leqslant \frac{1}{2}J(\eta,\eta)^{1/2}\sum_K \int_{\partial K}(h_K^{-1}\eta)^2\,\mathrm{d}s,$$

and we may apply (8) and (9) to arrive at

$$\|\nabla\eta\|^2 = \frac{1}{2}J(\eta,\eta)^{1/2}\left(\sum_K h_K^{-3}\|\eta\|_K^2\right)^{1/2}.$$

By the estimate of Theorem 9 we know that $\|\eta\|^2 \leqslant Ch^3$ and that $J(\eta,\eta) \leqslant |||\eta|||^2 \leqslant Ch^3$, and the claim follows. □

We proceed to prove an a priori L_2-estimate for the diffusion dominated case using a duality argument. Consider the following adjoint problem: find $\varphi \in H_0^1(\Omega)$ such that

$$A(v,\varphi) = (\psi, v) \quad \forall v \in H_0^1(\Omega). \tag{28}$$

This problem is well-posed and if $\psi = u - U$ then $\|\varepsilon\varphi\|_{H^2(\Omega)} \leqslant C\|u-U\|$. Using the Aubin-Nitsche duality argument we obtain

Theorem 11. *Assume that ε is bounded away from zero and let $u \in H^2(\Omega)$ be the solution of* (2) *and $U \in V_h$ the finite element solution of* (4). *Then*

$$\|u - U\| \leqslant Ch^2\|u\|_{H^2(\Omega)}. \tag{29}$$

Proof. Choosing $\psi = v = u - U$ in Eq. (28) we obtain, since $\varphi \in H^2(\Omega)$,

$$\begin{aligned}\|u - U\|^2 &= A(u - U, \varphi) = A(u - U, \varphi - \pi_h\varphi) + J(u - U, \pi_h\varphi)\\ &= A(u - U, \varphi - \pi_h\varphi) + J(u - U, \varphi - \pi_h\varphi).\end{aligned}$$

Writing out the different contributions and applying Cauchy–Schwartz in the last term we have

$$\begin{aligned}\|u - U\|^2 &\leqslant (\beta \cdot \nabla(u - U), \varphi - \pi_h\varphi) + (\sigma(u - U), \varphi - \pi_h\varphi) + (\varepsilon\nabla(u - U), \nabla(\varphi - \pi_h\varphi))\\ &\quad + J(u - U, u - U)^{1/2}J(\varphi - \pi_h\varphi, \varphi - \pi_h\varphi)^{1/2}\\ &\leqslant C|||u - U|||(\|h_K^{-1/2}(\varphi - \pi_h\varphi)\| + \|\varepsilon^{1/2}(\varphi - \pi_h\varphi)\|_{H^1(\Omega)} + J(\varphi - \pi_h\varphi, \varphi - \pi_h\varphi)^{1/2}).\end{aligned}$$

From Theorem 9 we know that $|||u - U||| \leqslant C(\varepsilon^{1/2}h + h^{3/2})$. Applying standard interpolation estimates we bound the norms of the dual solution by

$$\|h_K^{-1/2}(\varphi - \pi_h\varphi)\| + \|\varepsilon^{1/2}(\varphi - \pi_h\varphi)\|_{H^1(\Omega)} + J(\varphi - \pi_h\varphi, \varphi - \pi_h\varphi)^{1/2} \leqslant C\left(\frac{h^{3/2}}{\varepsilon} + \frac{h}{\varepsilon^{1/2}}\right)\|\varepsilon\varphi\|_{H^2(\Omega)}.$$

and consequently $\|u - U\| \leqslant Ch^2(\frac{h}{\varepsilon} + \frac{h^{1/2}}{\varepsilon^{1/2}} + 1)\|u\|_{H^2(\Omega)}$. □

2.3. A posteriori error estimates

We consider estimates of general linear functionals of the error, following Becker and Rannacher [1].

Theorem 12. *Let u be a solution to* (4), *let ψ and φ be the data and solution to* (28), *and define $I_\psi(u - U) = (u - U, \psi)$. Then*

$$|I_\psi(u - U)| \leqslant \sum_K (\rho_K\omega_K + \tilde{\rho}_K\tilde{\omega}_K), \tag{30}$$

where

$$\rho_K = \|\sigma U + \beta \cdot \nabla U - f\|_K + h_K^{-1/2}\|[\varepsilon n \cdot \nabla U]\|_{\partial K}, \tilde{\rho}_K = \gamma h_K^{1/2}\|[n \cdot \nabla U]\|_{\partial K}$$

and

$$\omega_K = \max\{\|\varphi - \pi_h\varphi\|_K, h_K^{1/2}\|\varphi - \pi_h\varphi\|_{\partial K}\}, \quad \tilde{\omega}_K = h_K^{3/2}\|[n \cdot \nabla\pi_h\varphi]\|_{\partial K}.$$

Proof. We have, using Lemma 1, that

$$\begin{aligned}I_\psi(u - U) &= A(u - U, \varphi) = A(u - U, \varphi - \pi_h\varphi) - J(U, \pi_h\varphi) = A(u - U, \varphi - \pi_h\varphi) - J(U, \pi_h\varphi)\\ &= (f, \varphi - \pi_h\varphi) - (\beta \cdot \nabla U + \sigma U, \varphi - \pi_h\varphi) - (\varepsilon\nabla U, \nabla(\varphi - \pi_h\varphi)) - J(U, \pi_h\varphi).\end{aligned}$$

We note that the stabilizing term is bounded by

$$J(U, \pi_h\varphi) \leqslant \sum_K \frac{1}{2}\gamma h_K^2\|[n \cdot \nabla U]\|_{\partial K}\|[n \cdot \nabla\pi_h\varphi]\|_{\partial K}.$$

The desired estimate is then obtained by an integration by parts in the third term together with an element wise application of the Cauchy–Schwartz inequality. □

Remark 13. The sum over $\tilde{\rho}_K\tilde{\omega}_K$ may be replaced by $|J(U,\pi_h\varphi)|$.

We now prove that for sufficiently regular solution and adjoint solution the stabilizing term contribution is of the right order.

Corollary 14. *If* $\varphi \in H^2(\Omega)$ *and* $u \in H^2(\Omega)$, *then*

$$|J(U,\pi_h\varphi)| \leqslant Ch^2. \tag{31}$$

Proof. The result is an immediate consequence of the consistency and the interpolation:

$$\begin{aligned}|J(U,\pi_h\varphi)| &= |J(U-u,\pi_h\varphi-\varphi)|\\ &\leqslant |J(U-u,U-u)|^{1/2}|J(\pi_h\varphi-\varphi,\pi_h\varphi-\varphi)|^{1/2}\\ &\leqslant C(\varepsilon^{1/2}h+h^{3/2}+\sigma^{1/2}h^2)^2,\end{aligned}$$

where the last inequality follows from Lemma 8 and Theorem 9. □

Remark 15. We note that the stabilization term $J(U,U)$ will not make convergence deteriorate when $\varepsilon > h$; hence there is no need to tune the stabilization parameter in such a way that it tends to zero when the fine scales of the flow are resolved to preserve order. This is another advantage of our method compared with the SUPG method.

2.4. Discrete maximum principle

In a recent paper [3] the authors constructed shock-capturing terms, for which they rigorously proved a discrete maximum principle (DMP). This was in the case of the streamline diffusion method and only for strictly acute meshes. These monotonicity results were then developed further in [4]. In particular a DMP satisfying shock-capturing edge stabilization method was proposed and its convergence was proved. Here we will recall these results and outline a proof showing that the shock-capturing edge stabilized scheme satisfies a discrete maximum principle. This edge stabilization shock-capturing term is defined on the edges of the elements and uses the jump of the gradient between adjacent elements instead of the residual internal to the elements. Moreover the diffusion is given in the edge tangent direction to avoid acute conditions on the mesh. This has the advantage of making the shock-capturing term independent of data. We recall the notation of Section 2.1, consider some node S_i, let $\mathcal{N}_i$ be the set of all triangles K containing node i, $\Omega_i = \bigcup_{K\in\mathcal{N}_i} K$, $\mathcal{S}_i$ the set of all edges connected to S_i and $\tilde{\mathcal{S}}_i$ the set of all edges in $\overline{\Omega}_i$. Furthermore we denote by v_i the function in V_h, such that $v_i = \delta_{ij}$ in node S_j and by $[x]_e$ we denote the jump of the quantity x across the edge e. The argument of [4] is based on the notion of the *DMP property* defined by

Definition 2.1. We say that the semi-linear form $\tilde{a}(U;v)$ satisfies the *strong DMP property* if the following holds true: $\forall U \in V_h^g$ and for all interior vertex S_i, if U is locally minimal (resp. maximal) on vertex S_i over macro-element Ω_i, then there exist positive quantities $(\alpha_e)_{e\in\mathcal{E}_i}$ such that

$$\tilde{a}(U;w_i) \leqslant -\sum_{e\in\mathcal{E}_i}\alpha_e|[\nabla U]_e| \tag{32}$$

(resp. $\tilde{a}(U;w_i) \geqslant \sum_{e\in\mathcal{E}_i}\alpha_e|[\nabla U]_e|$).

Using these definitions we now recall a lemma from [4].

Proposition 16. *Assume that the semi-linear form $\tilde{a}(U;v)$ satisfies the strong DMP property, that $f \geqslant 0$, and let $U \in V_h^g$ be the solution of the abstract problem*

$$\tilde{a}(U;v) = (f,v) \quad \forall v \in V_h^g.$$

Then U reaches its minimum on the boundary $\partial\Omega$.

Proof. Suppose that U reaches its minimum at an interior vertex S_i. Since $\tilde{a}(U;v)$ satisfies the strong DMP property and $f \geqslant 0$, we readily deduce from (32) that ∇U is constant over macro-element Ω_i. Therefore, the minimum is reached on a further vertex and we eventually deduce that the minimum is reached on the boundary. $\square$

Furthermore we recall the following geometric lemma for the proof of which we refer to [4]:

Lemma 17. *If $U \in V_h$ and U has a local minimum in the node S_i, then*

$$\|U - U(S_i)\|_{L_1(\Omega_i)} \leqslant C_0 \|h_K \nabla U\|_{L_1(\Omega_i)} \leqslant C_1 \|h_{\partial K}^2 [\nabla U]\|_{L_1(\mathscr{S}_i)}. \tag{33}$$

Lemma 17 and Proposition 16 are the fundamental ingredients in the following theorem which shows that the shock-capturing edge stabilization method satisfies the DMP. We consider the case $\sigma = 0$; the case $\sigma > 0$ requires additional assumptions, cf. Remark 19.

Theorem 18. *If $U \in V_h^0$ is a function such that*

$$A(U,v) + J(U,v) + J_{sc}(U,v) = (f,v) \quad \forall v \in V_h^0, \tag{34}$$

with $\sigma = 0$ in $A(U,v)$, $f \geqslant 0$, and

$$J_{sc}(U,v) = \sum_K \int_{\partial K} \Psi(U) \operatorname{sign}(\tau \cdot \nabla U) \tau \cdot \nabla v \, \mathrm{d}s, \tag{35}$$

where

$$\Psi(U)|_K = h_K (C_\varepsilon \varepsilon + C_{\beta,\gamma} h_K) \max_{e \in K} |[n \cdot \nabla U]_e|, \tag{36}$$

then $U \geqslant 0$.

Proof. First some remarks are in order. We note that the shock-capturing term is divided into two parts: one of order $h_K \varepsilon$ and the other of order h_K^2. The first contribution is needed to control violations of the DMP due to the Laplace operator discretized on non-strictly acute meshes, the other term controls violations of the DMP provoked by the convective term and the stabilization. To prove the theorem is sufficient to show that the semi-linear form has the DMP-property. We assume that there is a local minimum in the node S_i and test (34) with the corresponding test function v_i. First, we integrate by parts to obtain

$$A(U, v_i) = \int_{\mathscr{S}_i} [\varepsilon n \cdot \nabla U] v_i \, \mathrm{d}s + \int_{S_i} (U(S_i) - U) \beta \cdot \nabla v_i \, \mathrm{d}x.$$

Next, we use that $\mathrm{sign}(\tau \cdot \nabla U)\tau \cdot \nabla v_i < 0$ on every edge $e \in \mathscr{S}_i$ to bound the first term

$$\int_{\mathscr{S}_i} [\varepsilon n \cdot \nabla U] v_i \,\mathrm{d}s + C_\varepsilon \varepsilon \int_{\tilde{\mathscr{S}}_i} h_{\partial K} \max_{e \in K} |[n \cdot \nabla U]_e| \mathrm{sign}(\tau \cdot \nabla U)\tau \cdot \nabla v_i \,\mathrm{d}s$$
$$\leqslant \varepsilon \frac{1}{2} \|[n \cdot \nabla U]\|_{L_1(\mathscr{S}_i)} - C_\varepsilon \sum_{K \in \Omega_i} \| \max_{e \in K} |[n \cdot \nabla U]_e| \|_{L_1(\partial K)} \leqslant 0$$

with $C_\varepsilon \geqslant \frac{1}{2}$. In the same manner we write for the second term

$$\int_{\Omega_i} (U(S_i) - U)\beta \cdot \nabla v_i \,\mathrm{d}x + C_{\gamma,\beta} \sum_{K \in \Omega_i} \int_{\partial K} h^2_{\partial K} \max_{e \in K} |[n \cdot \nabla U]_e| \mathrm{sign}(\tau \cdot \nabla U)\tau \cdot \nabla v_i \,\mathrm{d}s$$
$$\leqslant \|U(S_i) - U\|_{L_1(\Omega_i)} \|\beta \cdot \nabla v_i\|_{L_\infty(\Omega_i)} - C_{\gamma,\beta} h_K^2 \sum_{K \in \Omega_i} \| \max_{e \in K} |[n \cdot \nabla U]_e| \|_{L_1(\partial K)} \min_{\mathscr{S}_i} |\tau \cdot \nabla v_i| \leqslant 0,$$

where the last inequality is a consequence of Lemma 17 and a scaling argument. Consider finally the least-squares stabilization term $J(U, v_i)$:

$$\sum_K \frac{1}{2} \int_{\partial K} \gamma h^2_{\partial K} [n \cdot \nabla U][n \cdot \nabla v_i] \,\mathrm{d}s + C_{\gamma,\beta} \int_{\mathscr{S}_i} h^2_{\partial K} \max_{e \in K} |[n \cdot \Delta U]_e| \mathrm{sign}(\tau \cdot \nabla U)\tau \cdot \nabla v_i \,\mathrm{d}s$$
$$\leqslant h_K^2 \frac{\gamma}{2} \|[n \cdot \nabla U]\|_{L_1(\tilde{\mathscr{S}}_i)} \|\nabla v_i\|_{L_\infty(S_i)} - C_{\gamma,\beta} h_K^2 \| \max_{e \in K} |[n \cdot \Delta U]_e| \|_{L_1(\tilde{\mathscr{S}}_i)} \min_{\mathscr{S}_i} |\tau \cdot \nabla v_i| \leqslant 0.$$

It follows that the semi-linear form has the DMP-property for $C_{\gamma,\beta}$ big enough and hence $U \geqslant 0$ by Proposition 16. □

Remark 19. We note that this holds true also for elliptic problems, allowing for the discrete maximum principle to hold in this case on meshes that are not strictly acute. The $\sigma > 0$ case may be included in the above framework, either by using nodal quadrature (lumped mass) for the source term, or by adding a shock-capturing term tailored to control the source term. For further detail on these issues we refer to [4].

Remark 20. The above form of $\Psi(U)$ has been chosen in order to enhance clarity of the argument, however it is not the minimal coefficient assuring a DMP. Indeed a more detailed study allows for a minimal shock-capturing term where each of the terms is accounted for separately.

3. Numerical examples

In this section we will illustrate the theoretical results obtained above with some computational experiments.

3.1. Convection–diffusion–reaction

The model problem (1) is considered, choosing $\sigma = 1$, $\beta = (1, 0)$ and $\varepsilon = 10^{-5}$, corresponding to the convection dominated case. We let $\Omega = [0, 1] \times [0, 1]$ and use two different source terms f in order to get the following exact solutions, see Fig. 1:

- Test case 1: $u = \exp\left(-\frac{(x-0.5)^2}{a_w} - \frac{3(y-0.5)^2}{a_w}\right)$,
- Test case 2: $u = \frac{1}{2}\left(1 - \tanh\left(\frac{x-0.5}{a_w}\right)\right)$.

The corresponding f's are obtained by inserting these exact solutions into the partial differential equation. For the Gaussian the parameter controlling the slope was chosen to $a_w = 0.2$ and for the

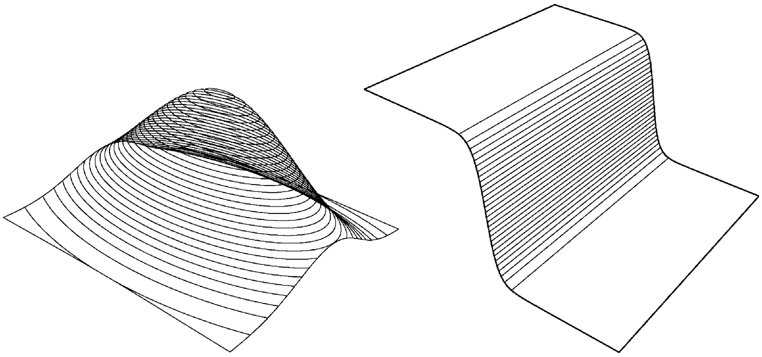

Fig. 1. The two exact solutions—left: the Gaussian, right: the hyperbolic tangent.

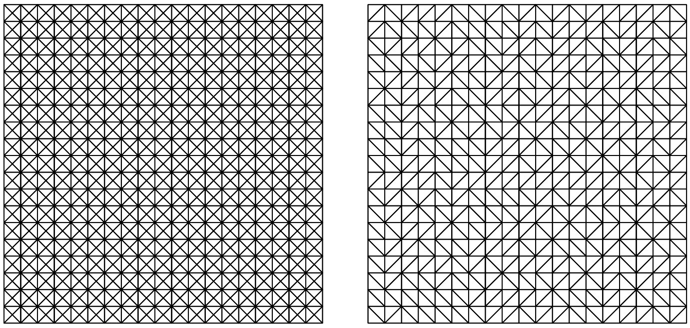

Fig. 2. The two different meshes used—left: mesh 1, right: mesh 2.

hyperbolic tangent the parameter was chosen to $a_w = 0.05$. Two different types of meshes have been used, illustrated in Fig. 2, both are based on square elements, in the first case (denoted *mesh 1*) a they are cut into four triangles and in the other (denoted *mesh 2*) the square elements are cut into two triangles, with the diagonal chosen randomly.

We have computed the solution using the streamline diffusion method, edge stabilization, with the term given by (3) (abbreviated EC) and the one given by (5) with $\gamma_{\beta^\perp} = 0$ (abbreviated ES). The stabilization parameter for the edge stabilization was chosen to $\gamma = 0.025$ and no shock capturing was used. The solutions were computed on four consecutive meshes having $N = 20$, 40, 80 and 160 elements on each side respectively. We present the errors in the L_2 norm, the H^1 norm and the L_∞ norm for the three methods applied to the two test cases in Tables 1–4.

For the first test case we note the following approximate convergence orders:

- $\|u - u_h\|_{0,\Omega} \approx \mathrm{O}(h^2)$
- $\|\nabla(u - u_h)\|_{0,\Omega} \approx \mathrm{O}(h)$
- $\|u - u_h\|_{0,\infty} \approx \mathrm{O}(h^{1.7})$

on both meshes. We note that the edge stabilization method ES, using the jumps only in the streamline derivative gives results very similar to that of the streamline-diffusion method, whereas the method EC, where the jump of the whole gradient is used for stabilization gives slightly larger errors on the coarsest mesh. On finer meshes the errors of all three methods are comparable. The method EC has slightly smaller error in the H^1 norm, possibly a consequence of the stronger control of the gradient reflected in Lemma 10.

Table 1
Convergence results for test case 1 on mesh 1

N	SD			ES			EC		
	L_2	H^1	L_∞	L_2	H^1	L_∞	L_2	H^1	L_∞
20	1.3e−3	1.6e−1	5.9e−3	1.3e−3	1.6e−1	6.1e−3	1.9e−3	1.4e−1	4.4e−3
40	3.2e−4	7.9e−2	1.5e−3	3.2e−4	7.9e−2	1.5e−3	3.6e−4	6.8e−2	1.1e−3
80	8.0e−5	3.9e−2	3.7e−4	7.9e−5	3.9e−2	3.8e−4	8.2e−5	3.4e−2	3.3e−4
160	2.0e−5	2.0e−2	9.3e−5	2.0e−5	2.0e−2	9.6e−5	2.0e−5	1.7e−2	9.7e−5

Table 2
Convergence results for test case 1 on mesh 2

N	SD			ES			EC		
	L_2	H^1	L_∞	L_2	H^1	L_∞	L_2	H^1	L_∞
20	2.1e−3	1.9e−1	9.7e−3	2.3e−3	1.9e−1	1.1e−2	4.0e−3	1.9e−1	1.3e−2
40	5.0e−4	9.6e−1	2.7e−3	5.1e−4	9.7e−2	2.8e−4	7.3e−4	1.0e−1	3.3e−3
80	1.2e−4	4.8e−2	8.0e−4	1.4e−4	4.8e−2	8.0e−4	1.6e−4	4.8e−2	1.1e−3
160	3.1e−5	2.4e−2	2.0e−4	3.2e−5	2.4e−2	2.1e−4	3.9e−5	2.4e−2	3.1e−4

Table 3
Convergence results for test case 2 on mesh 1

N	SD			ES			EC		
	L_2	H^1	L_∞	L_2	H^1	L_∞	L_2	H^1	L_∞
20	5.2e−3	6.8e−1	2.1e−2	5.8e−3	7.5e−1	3.2e−2	8.2e−3	6.4e−1	3.8e−2
40	1.2e−3	3.3e−1	6.2e−3	1.2e−3	3.6e−1	8.7e−3	1.3e−3	2.8e−1	8.4e−3
80	2.8e−4	1.7e−1	1.7e−3	3.0e−4	1.8e−1	2.0e−3	2.8e−4	1.4e−1	2.5e−3
160	7.0e−5	8.3e−2	4.2e−4	7.4e−5	9.0e−2	5.0e−4	6.8e−5	6.8e−2	8.1e−4

Table 4
Convergence results for test case 2 on mesh 2

N	SD			ES			EC		
	L_2	H^1	L_∞	L_2	H^1	L_∞	L_2	H^1	L_∞
20	1.1e−2	7.0e−1	6.9e−2	1.2e−2	7.8e−1	9.6e−2	1.2e−2	7.5e−1	7.5e−2
40	2.5e−3	3.8e−1	2.6e−2	2.5e−3	3.7e−1	2.5e−2	2.1e−3	3.4e−1	1.8e−2
80	6.2e−4	1.9e−1	7.2e−3	6.0e−4	1.9e−1	6.9e−3	4.6e−4	1.6e−1	4.5e−3
160	1.5e−4	9.5e−2	2.0e−3	1.4e−4	9.3e−2	2.0e−3	1.1e−4	8.2e−2	1.2e−3

3.2. Outflow layers and discrete maximum principle

In this section we will show qualitatively the loss of stability in outflow layers discussed in Remark 6 and how this instability can be countered using the shock-capturing term proposed in Section 2.4. We propose a classical testcase with a convection–diffusion problem ($\sigma = 0, |\beta| = 1, \eta = 10^{-5}$). The geometry, the boundary conditions and the orientation of β are resumed in Fig. 3.

As was noted in [4] the DMP satisfying shock-capturing methods result in very ill-conditioned non-linear equations due to the lack of continuity of the operator. We counter this by regularizing the sign operator, replacing it by sign_ϵ defined by $\mathrm{sign}_\epsilon(x) = \tanh(x/\epsilon)$, we choose $\epsilon = 1$ and $C_{\beta,\gamma} = 10$, a choice for which spurious oscillations are essentially eliminated. The results of the three methods applied with and without shock-capturing term is presented in Figs. 4–7 on the 40×40 mesh. In Fig. 6 we present the EC-method with and without shock capturing on the 20×20 mesh. We note the large oscillations on the outflow layer for both edge stabilization approaches. In the case of the streamline diffusion method the violation of the DMP is localized essentially at the inflow in this case. The maximal overshoot for the respective cases are reported in Table 5. Although the weaker outflow stability of the edge stabilization method results in huge overshoots

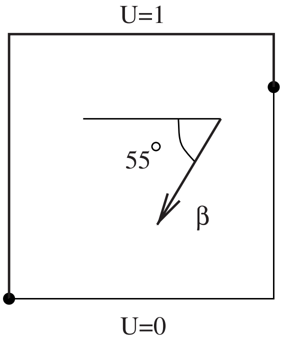

Fig. 3. Boundary conditions and flow orientation, for outflow layer test case: $U = 1$ along thick edge and $U = 0$ along thin edge.

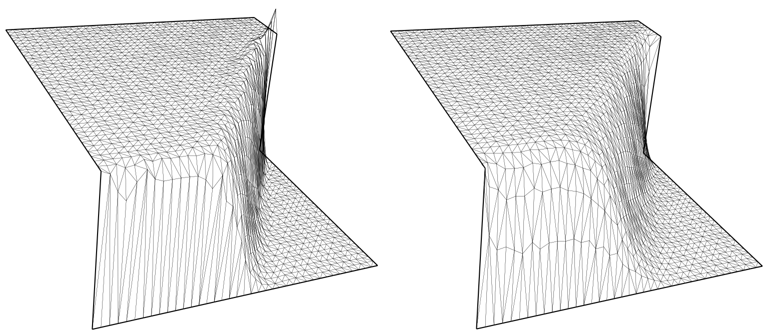

Fig. 4. Outflow boundary layer testcase using SD—left: without shock capturing, right: with shock capturing.

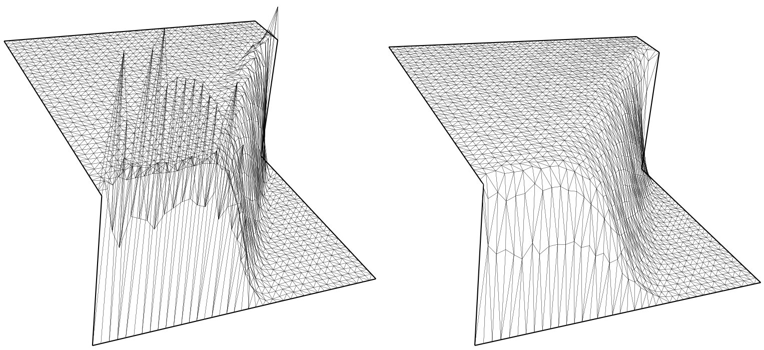

Fig. 5. Outflow boundary layer testcase using ES—left: without shock capturing, right: with shock capturing.

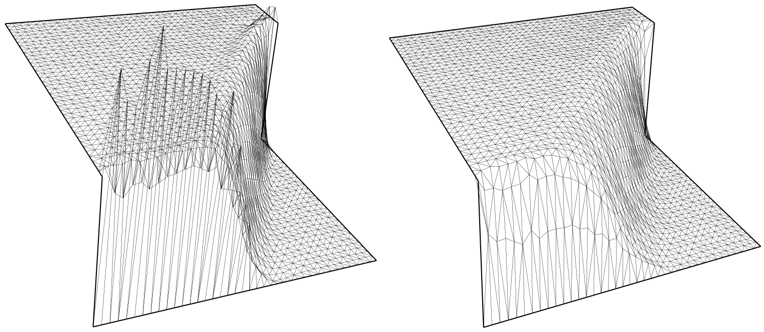

Fig. 6. Outflow boundary layer testcase using EC—left: without shock capturing, right: with shock capturing.

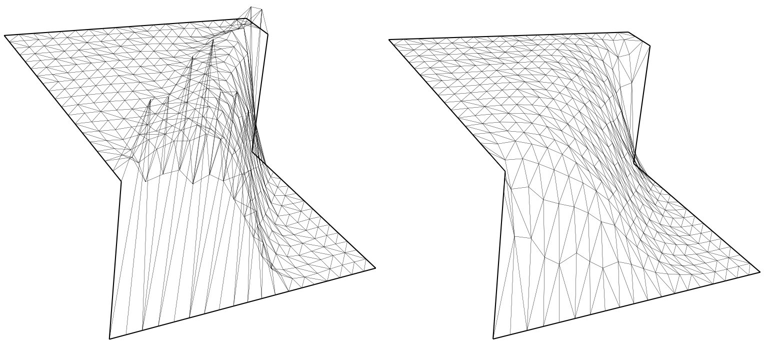

Fig. 7. Outflow boundary layer testcase using EC (20×20 mesh)—left: without shock capturing (max DMP violation: 71%), right: with shock capturing (max DMP violation: 0.6%).

Table 5
Maximum violation of the DMP in % for the different methods on the 40×40 mesh

Method	SD	ES	EC
SC	0.38	0.99	1.2
No SC	20	95	85

we see that the DMP satisfying shock-capturing term wipes them out almost entirely, the remaining violation of the DMP of about one percent is due to the regularization of the sign operator, see [3].

References

[1] R. Becker, R. Rannacher, A feed-back approach to error control in finite element methods: basic analysis and examples, East-West J. Numer. Math. 4 (1996) 237–264.
[2] F. Brezzi, T.J.R. Hughes, L.D. Marini, A. Russo, E.A. Süli, A priori error analysis of residual-free bubbles for advection–diffusion problems, SIAM J. Numer. Anal. 36 (1999) 1933–1948.
[3] E. Burman, A. Ern, Nonlinear diffusion and discrete maximum principle for stabilized Galerkin approximations of the advection–diffusion–reaction equation, Comput. Methods Appl. Mech. Engrg. 191 (2002) 3822–3855.

[4] E. Burman, A. Ern, Stabilized Galerkin approximation of convection–diffusion–reaction equations: discrete maximum principle and convergence, preprint: http://ima.epfl.ch/~burman/DMP.ps.gz.
[5] P. Clément, Approximation by finite element functions using local regularization, RAIRO Anal. Numer. 9 (R-2) (1975) 77–84.
[6] J. Douglas, T. Dupont, Interior penalty procedures for elliptic and parabolic Galerkin methods, in: R. Glowinski, J.L. Lions (Eds.), Computing Methods in Applied Sciences, Springer-Verlag, Berlin, 1976.
[7] J.L. Guermond, Stabilization of Galerkin approximations of transport equations by subgrid modeling, M2AN Math. Model. Numer. Anal. 33 (1999) 1293–1316.
[8] C. Johnson, U. Nävert, J. Pitkäranta, Finite element methods for linear hyperbolic equations, Comput. Methods Appl. Mech. Engrg. 45 (1984) 285–312.
[9] V. Thomée, Galerkin Finite Element Methods for Parabolic Problems, Springer-Verlag, Berlin, 1984.