Universität des Saarlandes

Fachrichtung 6.1 – Mathematik

Preprint Nr. 156

A comparison of spurious oscillations at layers diminishing (SOLD) methods for convection–diffusion equations: Part I

Volker John and Petr Knobloch

Saarbrücken 2005

Fachrichtung 6.1 – Mathematik
Universität des Saarlandes

Preprint No. 156
submitted: November 2, 2005

A comparison of spurious oscillations at layers diminishing (SOLD) methods for convection–diffusion equations: Part I

Volker John

Saarland University
Department of Mathematics
Postfach 15 11 50
D–66041 Saarbrücken
Germany
john@math.uni-sb.de

Petr Knobloch

Charles University
Faculty of Mathematics and Physics
Department of Numerical Mathematics
Sokolovská 83
186 75 Praha 8
Czech Republic
knobloch@karlin.mff.cuni.cz

Edited by
FR 6.1 – Mathematik
Universität des Saarlandes
Postfach 15 11 50
66041 Saarbrücken
Germany

Fax:	+ 49 681 302 4443
e-Mail:	preprint@math.uni-sb.de
WWW:	http://www.math.uni-sb.de/

Abstract

An unwelcome feature of the popular streamline upwind/Petrov–Galerkin (SUPG) stabilization of convection–dominated convection–diffusion equations is the presence of spurious oscillations at layers. Since the mid of the 1980-ies, a number of methods have been proposed to remove or, at least, to diminish these oscillations without leading to excessive smearing of the layers. The paper gives a review and state of the art of these methods, discusses their derivation, proposes some alternative choices of parameters in the methods and categorizes them. Some numerical studies give a first insight into the advantages and drawbacks of the methods.

Keywords: Convection–diffusion equations, Streamline upwind/Petrov–Galerkin (SUPG) method, Spurious oscillations at layers diminishing (SOLD) methods

1 Introduction

This paper is devoted to the numerical solution of the scalar convection–diffusion equation

$$-\varepsilon\,\Delta u + \boldsymbol{b}\cdot\nabla u = f \quad \text{in } \Omega, \qquad\qquad u = u_b \quad \text{on } \partial\Omega, \tag{1}$$

where $\Omega \subset \mathbb{R}^d$, $d = 2, 3$, is a bounded domain with a polygonal (resp. polyhedral) boundary $\partial\Omega$, $\varepsilon > 0$ is the constant diffusivity, $\boldsymbol{b} \in W^{1,\infty}(\Omega)^2$ is a given convective field satisfying the incompressibility condition $\operatorname{div}\boldsymbol{b} = 0$, $f \in L^2(\Omega)$ is an outer force, and $u_b \in H^{1/2}(\partial\Omega)$ represents the Dirichlet boundary condition. In our numerical tests we shall also consider less regular functions u_b.

Problem (1) describes the stationary distribution of a physical quantity u (e.g., temperature or concentration) determined by two basic physical mechanisms, namely the convection and diffusion. The broad interest in solving problem (1) is caused not only by its physical meaning just explained but also (and perhaps mainly) by the fact that it is a simple model problem for convection–diffusion effects which appear in many more complicated problems arising in applications (e.g. in various fluid flow problems).

Despite the apparent simplicity of problem (1), its numerical solution is still a challenge when convection is strongly dominant (i.e., when $\varepsilon \ll |\boldsymbol{b}|$). The basic difficulty is that, in this case, the solution of (1) typically possesses interior and boundary layers, which are small subregions where the derivatives of the solution are very large. The widths of these layers are usually significantly smaller than the mesh size and hence the layers cannot be resolved properly. This leads to unwanted spurious (nonphysical) oscillations in the numerical solution, the attenuation of which has been the subject of extensive research for more than three decades.

In this paper, we concentrate on the solution of (1) using the finite element method which proved to be a very efficient tool for the numerical solution of various boundary value problems in science and engineering. Unfortunately, the classical Galerkin formulation of (1) is inappropriate since, in case of dominant convection, the discrete solution is usually globally polluted by spurious oscillations causing a severe loss of accuracy and stability. This is not surprising since, in simple settings, the standard Galerkin finite element method is equivalent to a central finite difference discretization and it is well known that central difference approximations of the convective term give rise to spurious oscillations in convection dominated regimes (cf. e.g. Roos *et al.* [50]).

To enhance the stability and accuracy of the Galerkin discretization of (1) in the convection dominated regime, various stabilization strategies have been developed. Initially, these techniques imitated the upwind finite difference techniques. An important contribution to this development was made by Christie *et al.* [15], who showed that, in the one–dimensional case, a stabilization can be achieved using asymmetric test functions in a weighted residual finite element formulation. Choosing these test functions in a suitable way, they recovered the usual one–sided differences used for the approximation of the convective term in the finite difference method. Two–dimensional upwind finite element discretizations were derived by Heinrich *et al.* in [28, 29] and by Tabata [54]. Many other finite element discretizations of upwind type have been proposed later.

Like in the finite difference method, the upwind finite element discretizations remove the unwanted oscillations but the accuracy attained is often poor since too much numerical diffusion is introduced. In addition, if the flow field $\boldsymbol{b}$ is directed skew to the mesh, an excessive artificial diffusion perpendicular to the flow (crosswind diffusion) can be observed. A further important drawback is that these methods are not consistent, i.e., the solution of (1) is no longer a solution to the variational problem as it is the case for a Galerkin formulation. Consequently, the accuracy is limited to first order. Moreover, non–consistent formulations are also known to produce inaccurate or wrong solutions when f (or the time derivative in case of transient problems) is significant. It can even happen that the discrete solution is then less accurate than that one produced by the Galerkin method (cf. e.g. Brooks and Hughes [8] for a discussion on shortcomings of upwind methods).

A significant improvement came with the streamline upwind/Petrov–Galerkin

(SUPG) method developed by Brooks and Hughes [8] which substantially eliminates almost all the difficulties mentioned above. In contrast with upwind methods proposed earlier, the SUPG method introduces numerical diffusion along streamlines only and hence it possesses no spurious crosswind diffusion. Moreover, the streamline diffusion is added in a consistent manner. Consequently, stability is obtained without compromising accuracy and convergence results may be derived for a wide class of finite elements. In view of its stability properties and higher–order accuracy, the SUPG method is regarded as one of the most efficient procedures for solving convection–dominated equations. A detailed description of the SUPG method will be provided in Section 3.

An alternative to the SUPG method is the Galerkin/least–squares method introduced by Hughes *et al.* [31] who observed that stabilization terms can be obtained by minimizing the square of the equation residual. A variant to this method was proposed by Franca *et al.* [23] using the idea of Douglas and Wang [20] to change the sign of the Laplacian in the test function. To simplify the presentation, we shall restrict ourselves to the SUPG method in the following.

The SUPG method produces to a great extent accurate and oscillation–free solutions but it does not preclude spurious oscillations (overshooting and undershooting) localized in narrow regions along sharp layers. It was observed by Almeida and Silva [3] that these oscillations can even be amplified if high–order finite elements are used in these regions. This indicates that using the streamlines as upwind direction is not always sufficient. Although the remaining nonphysical oscillations are usually small in magnitude, they are not permissible in many applications. An example are chemically reacting flows where it is essential to guarantee that the concentrations of all species are positive. Another example are free–convection computations where temperature oscillations create spurious sources and sinks of momentum that effect the computation of the flow field. The small spurious oscillations may also deteriorate the solution of nonlinear problems, e.g., in two–equations turbulence models or in numerical simulations of compressible flow problems, where the solution may develop discontinuities (shocks) whose poor resolution may effect the global stability of the numerical calculations.

The oscillations along sharp layers are caused by the fact that the SUPG method is neither monotone nor monotonicity preserving. Therefore, various terms introducing artificial crosswind diffusion in the neighborhood of layers have been proposed to be added to the SUPG formulation in order to obtain a method which is monotone, at least in some model cases, or which at least reduces the local oscillations. This procedure is referred to as discontinuity capturing or shock capturing. However, these names are not really appropriate in our opinion for several reasons. First, the solution of (1) does not posses shocks or discontinuities because of the presence of diffusion. Instead, steep but continuous layers are formed. Second, the position of these layers is in general already captured well by the SUPG formulation. And third, a confusion might arise with shock capturing methods which are used in the numerical simulation of compressible flows. For these reasons, we propose to call the methods *spurious oscillations at layers diminishing (SOLD) methods* and this name is used throughout the paper.

Since linear monotone methods can be at most first–order accurate, it is natural to look for SOLD terms which depend on the discrete solution in a nonlinear way. However, linear SOLD terms applicable to first–order finite elements have also been developed. A basic problem of all these methods is to find the proper amount of artificial diffusion which leads to sufficiently small nonphysical oscillations (requiring that artificial diffusion is not 'too small') and to a sufficiently high accuracy (requiring that artificial diffusion is not 'too large').

The literature on SOLD methods is rather extended and the various numerical tests published in the literature do not allow to draw a clear conclusion concerning their advantages and drawbacks. Therefore, the aim of the present paper is to provide a review of the most published SOLD methods and to compare these methods computationally at two test problems whose solutions possess characteristic features of solutions of (1). The numerical results will give a first insight into the behavior of the SOLD methods. Comprehensive numerical studies will be presented in the second part of the paper. The aspects which will be covered by these studies are summarized at the beginning of Section 6. In order to keep the paper in a reasonable length, we do not consider a reaction term in equation (1) since special techniques are necessary if this term is dominant.

Sometimes, it is claimed that the SUPG method applied on adaptively refined meshes should be preferred to SOLD methods considered in this paper. However, if convection strongly dominates diffusion, the spurious oscillations of the SUPG method disappear only if extremely fine meshes are used along inner and boundary layers. This leads to a high computational cost which further increases if systems of equations or transient problems are considered. The numerical comparison of the SUPG method on adaptively refined grids and several SOLD methods will be a topic of the second part of the paper. Let us also mention that a further reason for using SOLD methods is that they try to preserve the inverse monotonicity property of the continuous problem.

The plan of the paper is as follows. In the next section, we describe the usual Galerkin discretization of (1) and, in Section 3, we introduce the SUPG method. The accuracy of the SUPG method is greatly influenced by the choice of the stabilizing parameter, which is discussed in Section 4. Then, a detailed review of SOLD methods follows in Section 5. Results of our numerical tests with the SOLD methods at two typical examples are reported in Section 6. Finally, the paper is closed by Section 7 containing our conclusions.

Throughout the paper, we use the standard notations $L^p(\Omega)$, $W^{k,p}(\Omega)$, $H^k(\Omega) = W^{k,2}(\Omega)$, $C(\overline{\Omega})$, etc. for the usual function spaces, see e.g. Ciarlet [16]. The norm and seminorm in the Sobolev space $H^k(\Omega)$ will be denoted by $\|\cdot\|_{k,\Omega}$ and $|\cdot|_{k,\Omega}$, respectively. The inner product in the space $L^2(\Omega)$ or $L^2(\Omega)^d$ will be denoted by $(\cdot,\cdot)$. For a vector $\boldsymbol{a} \in \mathbb{R}^d$, we denote by $|\boldsymbol{a}|$ its Euclidean norm.

2 Galerkin's finite element discretization

The starting point of defining any finite element discretization is a weak (or variational) formulation of the respective problem. Denoting by $\widetilde{u}_b \in H^1(\Omega)$ an extension of u_b, a natural weak formulation of the convection–diffusion equation (1) reads:

Find $u \in H^1(\Omega)$ such that $u - \widetilde{u}_b \in H^1_0(\Omega)$ and

$$a(u,v) = (f,v) \qquad \forall\ v \in H^1_0(\Omega)\,, \tag{2}$$

where

$$a(u,v) = \varepsilon\,(\nabla u, \nabla v) + (\boldsymbol{b}\cdot\nabla u, v)\,.$$

Since $a(v,v) = \varepsilon\,|v|^2_{1,\Omega}$ for any $v \in H^1_0(\Omega)$, it easily follows from the Lax–Milgram theorem that this weak formulation has a unique solution (cf. e.g. Ciarlet [16]).

To define a finite element discretization of (1), we introduce a triangulation $\mathcal{T}_h$ of the domain Ω consisting of a finite number of open polygonal resp. polyhedral elements K. The discretization parameter h in the notation $\mathcal{T}_h$ is a positive real number satisfying $\mathrm{diam}(K) \le h$ for any $K \in \mathcal{T}_h$. We assume that $\overline{\Omega} = \bigcup_{K\in\mathcal{T}_h} \overline{K}$ and that the closures of any two different elements K, $\widetilde{K} \in \mathcal{T}_h$ are either disjoint or possess either a common vertex or a common edge or, if $d = 3$, a common face. In what follows, we shall confine ourselves to simplicial elements and to elements which are images of a d–dimensional cube under a d–linear mapping (these are general convex quadrilaterals for $d = 2$ and suitable convex hexahedra for $d = 3$). In order to prevent the elements from degenerating when h tends to zero, the elements have to satisfy certain shape–regularity assumptions.

The Galerkin finite element discretization of (1) is now obtained by replacing the space $H_0^1(\Omega)$ in (2) by a finite element subspace V_h (cf. e.g. Ciarlet [16]). In addition, we approximate the function $\widetilde{u}_b$ by a finite element interpolate $\widetilde{u}_{bh}$. Thus, we may say that $u_h \in H^1(\Omega)$ is a discrete solution of (1) if $u_h - \widetilde{u}_{bh} \in \mathrm{V}_h$ and

$$a(u_h, v_h) = (f, v_h) \qquad \forall\ v_h \in \mathrm{V}_h\,.$$

Again, the discrete problem is uniquely solvable.

3 The SUPG method

Since the Galerkin method lacks stability if convection dominates diffusion, we enrich it by a stabilization term proposed by Brooks and Hughes [8] yielding the SUPG method. For doing this, we change the assumptions on the space V_h. First, to introduce the SUPG method, the functions from V_h have to be at least of class H^2 inside each element $K \in \mathcal{T}_h$. To simplify further considerations, we shall assume that they are infinitely smooth inside each element, which can be justified by the fact that typical finite element functions are piecewise polynomial. Second, we shall not require that the functions from V_h are continuous across element edges (resp. faces), in order to include nonconforming finite element spaces into the formulation below. Thus, from now on, we assume that V_h is a finite–dimensional space satisfying

$$\mathrm{V}_h \subset \{v \in L^2(\Omega)\,;\ v|_K \in C^\infty(\overline{K}) \ \ \forall\ K \in \mathcal{T}_h\}\,.$$

Defining the discrete operators ∇_h and Δ_h by

$$(\nabla_h\, v)|_K = \nabla(v|_K)\,, \qquad (\Delta_h\, v)|_K = \Delta(v|_K) \qquad \forall\ K \in \mathcal{T}_h\,,$$

the bilinear form

$$a_h(u,v) = \varepsilon\,(\nabla_h\, u, \nabla_h\, v) + (\boldsymbol{b}\cdot\nabla_h\, u, v)$$

and the residual

$$R_h(u) = -\varepsilon\,\Delta_h\, u + \boldsymbol{b}\cdot\nabla_h\, u - f$$

are well defined for $u, v \in \mathrm{V}_h$.

Then, the streamline upwind/Petrov–Galerkin (SUPG) method of Brooks and Hughes [8] reads:

Find $u_h \in L^2(\Omega)$ such that $u_h - \widetilde{u}_{bh} \in \mathrm{V}_h$ and

$$a_h(u_h, v_h) + (R_h(u_h), \tau\,\boldsymbol{b}\cdot\nabla_h\, v_h) = (f, v_h) \qquad \forall\ v_h \in \mathrm{V}_h\,, \tag{3}$$

where $\tau \in L^\infty(\Omega)$ is a nonnegative stabilization parameter.
For the SUPG method, many theoretical results have been derived. Since the SUPG method itself is not the subject of this paper, we shall not present any details and only refer to the monograph by Roos *et al.* [50].

4 Choice of the SUPG stabilization parameter

An important drawback of many stabilized methods (including the SUPG method) is that they contain stabilization parameters for which a general 'optimal' choice is not known. Since the SUPG method attracted a considerable attention over the last two decades, much research has also been devoted to the choice of the parameter τ. Theoretical investigations of the SUPG method provide certain bounds for τ for which the SUPG method is stable and leads to (quasi–)optimal convergence of the discrete solution u_h. However, it has been reported many times that the choice of τ inside these bounds may dramatically influence the accuracy of the discrete solution.
It follows from the results of Christie *et al.* [15] that, for the one–dimensional case of (1) with constant data, the SUPG solution with continuous piecewise linear finite elements on a uniform division of Ω is nodally exact if

$$\tau = \frac{h}{2\,|b|}\,\xi_0(\mathrm{Pe}) \qquad \text{with} \qquad \xi_0(\alpha) = \coth\alpha - \frac{1}{\alpha}\,, \quad \mathrm{Pe} = \frac{|b|\,h}{2\,\varepsilon}\,. \tag{4}$$

Here, h is the element length, ξ_0 is the so–called upwind function and Pe is the local Péclet number which determines whether the problem is locally (i.e., within a particular element) convection dominated or diffusion dominated. The parameter τ is often called 'intrinsic time scale' since $h/(2\,|b|)$ is the time for a particle to travel the distance $h/2$ at a speed equal to $|b|$. Since $\xi_0(\alpha) \to 1$ for $\alpha \to \infty$ and $\xi_0(\alpha)/\alpha \to 1/3$ for $\alpha \to 0+$ (and the SUPG stabilization is not necessary for $\alpha \to 0+$), the function ξ_0 is often approximated by

$$\xi_1(\alpha) = \max\left\{0, 1 - \frac{1}{\alpha}\right\}$$

or

$$\xi_2(\alpha) = \min\left\{1, \frac{\alpha}{3}\right\}\,.$$

Brooks and Hughes [8] call these functions 'critical' and 'doubly asymptotic' approximations of ξ_0, respectively. If the right–hand side of (1) is not constant, the choice (4) generally does not lead to a nodally exact discrete solution. Nevertheless, our numerical tests (not reported in this paper) indicate, that, in the most cases, the function ξ_0 leads to better results than

ξ_1 and ξ_2. However, it should be stressed that, for large values of Pe, the results for these three upwind functions are very close. This is particularly true for ξ_0 and ξ_1, for which $|\xi_0(\alpha) - \xi_1(\alpha)|/\xi_0(\alpha) < 10^{-3}$ for $\alpha > 4$ and $|\xi_0(\alpha) - \xi_1(\alpha)|/\xi_0(\alpha) < 10^{-10}$ for $\alpha > 12$ so that the corresponding discrete solutions are virtually indistinguishable for Pe > 10.

Many researchers have tried to find a suitable generalization of (4) to the multidimensional case and to more general finite element spaces V_h. Usually, for any element $K \in \mathcal{T}_h$, this generalization takes the form

$$\tau|_K \equiv \tau_K = \mu_K \, \frac{h_K}{2\,\|\boldsymbol{b}\|_K}\, \xi(\mathrm{Pe}_K) \qquad \text{with} \qquad \mathrm{Pe}_K = \nu_K \, \frac{\|\boldsymbol{b}\|_K \, h_K}{2\,\varepsilon}\,, \tag{5}$$

where μ_K, ν_K are constants depending on the definition of V_h, h_K is a characteristic dimension of K (also called 'local length scale' or 'element length'), $\|\boldsymbol{b}\|_K$ is a suitable norm of $\boldsymbol{b}$, ξ is an upwind function (such that $\xi(\alpha)/\alpha$ is bounded for $\alpha \to 0+$) and Pe_K is the local Péclet number. This generalization seems to be reasonable since, for linear or d–linear finite elements on certain uniform meshes aligned with a constant velocity $\boldsymbol{b}$, the discrete problem corresponds to the one–dimensional case and hence the formula for τ should reduce to (4).

The mentioned correspondence between the one–dimensional and d–dimensional cases particularly implies that one should take $\mu_K = \nu_K = 1$ for linear and d–linear finite elements. Moreover, if K is a rectangle and $\boldsymbol{b}$ is constant on K and aligned with one of its edges, we deduce that one should choose $\|\boldsymbol{b}\|_K = |(\boldsymbol{b}|_K)|$ and h_K equal to the length of the edge $\boldsymbol{b}$ is aligned with. The same holds if K is a right triangle and the vector $\boldsymbol{b}$ is aligned with one of its legs.

Another hint for choosing $\|\boldsymbol{b}\|_K$ and h_K follows from the necessary conditions for uniform convergence of $\|u - u_h\|_{0,\Omega}$ of order greater than 1/2 introduced by Stynes and Tobiska [53]. Let $d = 2$, $\boldsymbol{b} = (b, b)$ with some constant $b \in \mathbb{R}$ and let $\mathcal{T}_h$ be a uniform triangulation of $\Omega = (0,1)^2$ consisting of equal squares or of equal right triangles with hypotenuses in the direction $(1,1)$. Then, for (bi)linear finite elements, the necessary conditions are satisfied if and only if

$$\tau_K = \frac{\mathrm{diam}(K)}{2\,|\boldsymbol{b}|}\, \xi_0(\mathrm{Pe}_K/2) \qquad \text{with} \qquad \mathrm{Pe}_K = \frac{|\boldsymbol{b}|\,\mathrm{diam}(K)}{2\,\varepsilon}\,,$$

where $\mathrm{diam}(K) \equiv \sup\{|x - y|\,;\; x, y \in K\}$ is the diameter of K (see Stynes and Tobiska [53] and Shih and Elman [52] for details). The necessary conditions of Stynes and Tobiska were designed for the convection dominated case where $\xi_0(\mathrm{Pe}_K/2) \approx \xi_0(\mathrm{Pe}_K)$. This suggests to set $\|\boldsymbol{b}\|_K = |\boldsymbol{b}|$ and $h_K = \mathrm{diam}(K)$.

In view of the above considerations, it seems to be reasonable to define h_K as the diameter of K in the direction of the convection $\boldsymbol{b}$. Generally, given a vector $\boldsymbol{s} \in \mathbb{R}^d$, $\boldsymbol{s} \neq \boldsymbol{0}$, the diameter of K in the direction of $\boldsymbol{s}$ is defined by

$$\operatorname{diam}(K, \boldsymbol{s}) = \sup\{|x - y|\,;\ x, y \in K,\, x - y = \alpha\, \boldsymbol{s},\ \alpha \in \mathbb{R}\}\,.$$

This value may be sometimes difficult to compute and therefore we consider a slightly different definition which was used by Tezduyar and Park [56]. Let N_K be the number of vertices of K and let $\varphi_1, \dots, \varphi_{N_K}$ be the usual basis functions of $P_1(K)$ (if K is a simplex) or of $Q_1([0,1]^d)$ mapped onto K (if K is a quadrilateral or a hexahedron). We set

$$\operatorname{diam}^*(K, \boldsymbol{s}) = \frac{2\,|\boldsymbol{s}|}{\sum_{i=1}^{N_K} |\boldsymbol{s} \cdot \nabla\varphi_i(C_K)|}\,,$$

where C_K is the barycentre of K. Then $\operatorname{diam}^*(K, \boldsymbol{s}) = \operatorname{diam}(K, \boldsymbol{s})$ if K is a simplex or a parallelogram. If K is a hexahedron, then generally $\operatorname{diam}^*(K, \boldsymbol{s}) \neq \operatorname{diam}(K, \boldsymbol{s})$ (even not for a cube), but the value of $\operatorname{diam}^*(K, \boldsymbol{s})$ is still reasonable. If $\boldsymbol{s} = \boldsymbol{0}$, we set $\operatorname{diam}^*(K, \boldsymbol{s}) = \operatorname{diam}(K)$. Using this notation, we define

$$h_K = \operatorname{diam}^*(K, \boldsymbol{b})\,. \tag{6}$$

The norm $\|\boldsymbol{b}\|_K$ will be defined as the Euclidean norm of $\boldsymbol{b}$, i.e.,

$$\|\boldsymbol{b}\|_K = |\boldsymbol{b}|\,. \tag{7}$$

Note that, in view of (5)–(7), the parameters h_K, $\|\boldsymbol{b}\|_K$ and, consequently, Pe_K and τ_K are generally functions of the points $x \in K$.

Usually, the criterion for choosing τ is the accuracy of the discrete solution measured in some suitable norm. Nevertheless, it is also possible to look for τ such that the stiffness matrix corresponding to the discrete problem is well conditioned and enables an efficient application of iterative solvers. This idea was followed by Fischer *et al.* [24] and Ramage [47, 48]. In these papers, Q_1–discretizations of model problems in both two and three dimensions were investigated and it was observed that there is a close relationship between 'best' solution approximation and fast convergence of iterative methods. Particularly, for constant $\boldsymbol{b}$ aligned with a uniform mesh consisting of squares with side length h, an analysis of the structure of eigenvalues of the stiffness matrix reveals that one should choose $\tau = h/(2\,|\boldsymbol{b}|)$ for $h/\varepsilon \to \infty$ and provides the formula $\tau = h/(2\,|\boldsymbol{b}|)\,\xi_1(\mathrm{Pe})$ with $\mathrm{Pe} = |\boldsymbol{b}|\,h/(2\,\varepsilon)$ as a significant value with respect to the changes in the eigenvalue structure. In the general case, the choice of h_K as element size in the direction of $\boldsymbol{b}$ is advocated.

In [21], Elman and Ramage examined how the choice of τ influences the oscillations in a bilinear discrete solution and demonstrated that, generally, τ cannot be chosen in such a way that the discrete solution is simultaneously oscillation–free and accurate. The analysis gives a theoretical justification to the formula for τ given by (5)–(7) with $\mu_K = \nu_K = 1$ and $\xi = \xi_1$.

In Harari *et al.* [27], a formula for τ was found by requiring that the bilinear discrete solution on a uniform mesh is nodally exact for the equation (1) with $\boldsymbol{b} = const.$, $f = 0$ and $\Omega = \mathbb{R}^d$. It is interesting that, for $\boldsymbol{b}$ aligned with the element diagonals and $h/\varepsilon \to \infty$, the formula of Harari *et al.* gives only 2/5 of the value obtained from (5)–(7). Nevertheless, due to the absence of boundary conditions, the investigations of Harari *et al.* do not seem to be relevant for problems with boundary layers, which is the type of problems the SUPG method was designed for.

Now let us turn our attention to the choice of τ for higher order finite elements. If the polynomial degree of shape functions used on a given element $K \in \mathcal{T}_h$ is increased, the resulting effect is, in a certain sense, similar as if the element K is refined. Therefore, both Pe_K and τ_K should decrease for increasing polynomial degree on K. However, the kind of dependence of Pe_K and τ_K on the polynomial degree is not very clear. This is also caused by the fact that the most research on the choice of τ has been performed for linear and bilinear finite elements.

For the one–dimensional case with continuous piecewise quadratic finite elements, Codina *et al.* [18] showed that a nodally exact discrete solution can be obtained only if a different upwind function is used for each of the two types of shape functions. This is, of course, not very convenient, particularly because an extension to the multidimensional case is rather complicated. Using another definition of the optimality of the discrete solution, Codina *et al.* derived a formula for a unique upwind function equal to $\xi_0/2$. This was also proposed before on the basis of numerical experiments.

In case of general polynomial approximations, Almeida and Silva [3] report that, for $\mathrm{Pe}_K \to \infty$, numerical experiments suggest to divide τ_K corresponding to the linear case by the respective polynomial degree k. Galeão *et al.* [25] proposed to divide by k both Pe_K and τ_K corresponding to the linear case and a similar definition is also used here. We set

$$\mu_K = \nu_K = \frac{1}{k}\,, \tag{8}$$

where k is the order of approximation of $\mathrm{V}_h|_K$ with respect to the norm $\|\cdot\|_{1,K}$, i.e., it is the largest integer k such that

$$\inf_{v_h \in \mathrm{V}_h} \|v - v_h\|_{1,K} \le C\,[\mathrm{diam}(K)]^k\,\|v\|_{k+1,K} \qquad \forall\ v \in H^{k+1}(\Omega) \cap H^1_0(\Omega)$$

with some constant C independent of h and v. Thus, particularly, we consider (8) if V_h consists of piecewise polynomial functions of degree k.

Another definition of Pe_K and τ_K for general finite element spaces, which is based on error analysis considerations, was introduced by Franca *et al.* [23] who rescaled Pe_K using the constant from an inverse inequality. This corresponds to a rescaling with k^4 instead of k.

The relations (5)–(8) with $\xi = \xi_0$ represent a complete definition of the stabilization parameter τ used in this paper. Let us stress that this definition mostly relies on heuristic arguments and the 'best' way of choosing τ for general convection–diffusion problems is not known, particularly in case of higher order finite elements. Also, many other ways of computing τ have been proposed in the literature. Let us briefly mention a few of them.

Tezduyar and Osawa [55] proposed to compute stabilization parameters using element–level matrices and vectors. In this way, the local length scales, convection field and Péclet number are automatically taken into account. A similar idea was also used by Mizukami [43] for linear finite elements. A comparison of various definitions of local length scales and stabilization parameters can be found in Akin *et al.* [2]. Let us also mention the work of Akin and Tezduyar [1] where a comparative investigation of various ways of calculating the advective limit of τ is performed.

Roos *et al.* [50] propose to set

$$\tau|_K = \begin{cases} \tau_0\, h_K & \text{if } \mathrm{Pe}_K > 1, \\ \tau_1\, h_K^2/\varepsilon & \text{if } \mathrm{Pe}_K \le 1, \end{cases}$$

where τ_0 and τ_1 are appropriate positive constants. This definition of τ leads to the best possible convergence rate of the discrete solution with respect to the streamline diffusion norm. However, an 'optimal' choice of the constants τ_0 and τ_1 is unsolved.

Another possibility of defining the parameter τ is based on the observation that adding bubbles to the finite element space and eliminating them from the Galerkin discretization by static condensation is equivalent to the addition of a stabilizing term of streamline diffusion type. In this way, the question how to define τ is transformed into the question how to define suitable bubbles (cf. e.g. Brezzi and Russo [7]). This question was partially answered by introducing the concept of residual–free bubbles, see e.g. Brezzi *et al.* [5, 6, 7]. Using a similar approach in the framework of multiscale methods, an analytical formula for τ in terms of element Green's function was derived by Hughes [30]. Another method for stabilizing convection–dominated problems was proposed by Oñate [46], who introduced higher order terms into the continuous problem using the concept of flow balance over a finite domain. Applying the Galerkin method, the SUPG method can be recovered, which also provides a formula for computing the stabilization parameter τ.

5 A review of SOLD methods

In this section, we review most of the SOLD methods introduced during the last two decades to diminish the oscillations arising in the solution of the SUPG discretization (3). Let us recall that these oscillations appear along sharp layers of the solution to the continuous problem (1) due to the fact that the SUPG method is neither monotone nor monotonicity preserving.

One of the first successful monotone methods for solving (1) was introduced by Mizukami and Hughes [44] for conforming linear triangular finite elements. This method is based on the observation that the convection vector $\boldsymbol{b}$ in (1) can be changed in a perpendicular direction to ∇u without affecting the solution u of (1). This suggests that the streamline may not always be the appropriate upwind direction, an idea which have also been used to derive other SOLD methods later. Mizukami and Hughes used this idea to introduce a Petrov–Galerkin method which, due to the arbitrariness in $\boldsymbol{b}$, can be viewed as a method satisfying the discrete maximum principle. In contrast with other upwinding methods for conforming linear triangular finite elements satisfying the discrete maximum principle published earlier (cf. Tabata [54], Kanayama [39], Baba and Tabata [4], Ikeda [33]), the Mizukami–Hughes method adds much less numerical diffusion and provides rather accurate discrete solutions in the most cases. Recently some improvements of the Mizukami–Hughes method were introduced by Knobloch [40]. Unfortunately, it is not clear how to generalize the Mizukami–Hughes method to other types of finite elements.

Let us mention that the discrete maximum principle (which was also considered by many other authors to design SOLD methods) is an important property of a numerical scheme since it ensures monotonicity and that no spurious oscillations will appear, not even in the vicinity of sharp layers. Moreover, it enables to prove uniform convergence and pointwise stability estimates.

At the time as the Mizukami–Hughes scheme was published, Rice and Schnipke [49] proposed another monotone method which is based on a direct streamline upwind approximation to the convective term, rather than modifying the weighting function. This method was developed for bilinear finite elements and again a generalization does not seem to be easy.

Hughes *et al.* [32] came with the idea to change the upwind direction in the SUPG term of (3) by adding a multiple of the function

$$\boldsymbol{b}_h^{\|} = \begin{cases} \dfrac{(\boldsymbol{b}\cdot\nabla u_h)\nabla u_h}{|\nabla u_h|^2} & \text{if } \nabla u_h \neq \mathbf{0}, \\ \mathbf{0} & \text{if } \nabla u_h = \mathbf{0}, \end{cases}$$

which corresponds to the direction in which oscillations in SUPG solutions are observed. This leads to the additional term

$$(R_h(u_h), \sigma\, \boldsymbol{b}_h^{\|} \cdot \nabla_h v_h) \tag{9}$$

on the left–hand side of (3), where σ is a nonnegative stabilization parameter. This additional term controls the derivatives in the direction of the solution gradient, thus increasing the robustness of the SUPG method in the presence of sharp layers. Since $\boldsymbol{b}_h^{\|}$ depends on the unknown discrete solution u_h, the resulting method is nonlinear.

Of course, the key point here and in many other SOLD methods is how to choose the parameter σ. Unfortunately, due to the large number of various SOLD methods and the comparatively small amount of theoretical research on them, the correct choice of the respective stabilization parameters is even less clear than for the SUPG method. Often, the definition of these parameters is related to the choice of the parameter τ in the SUPG stabilization. Therefore, it is convenient to introduce the notation $\tau(\boldsymbol{b}^\star)$ representing τ determined by (5)–(8) with $\boldsymbol{b}$ replaced by some function $\boldsymbol{b}^\star$. Note that $\boldsymbol{b}^\star$ influences the value of $\tau_K(\boldsymbol{b}^\star)$ not only through the norm $\|\boldsymbol{b}^\star\|_K$ but also through the definition of h_K.

Now let us return to the choice of σ from (9). One could think of using the value $\tau(\boldsymbol{b}_h^{\|})$ but this would lead to a doubling of the SUPG stabilization if $\boldsymbol{b}_h^{\|} = \boldsymbol{b}$. Therefore, Hughes *et al.* [32] proposed to set

$$\sigma = \max\{0, \tau(\boldsymbol{b}_h^{\|}) - \tau(\boldsymbol{b})\}\,. \tag{10}$$

Although, for linear triangular finite elements, the method does not attain the precision of the Mizukami–Hughes scheme mentioned above (see Hughes *et al.* [32]), it has the important property that it is applicable to general finite elements.

Tezduyar and Park [56] proposed to redefine $\tau(\boldsymbol{b}_h^{\|})$, which leads to

$$\sigma = \mu_K \, \frac{h_K^{\|}}{2\,|\boldsymbol{b}_h^{\|}|} \, \eta\left(\frac{|\boldsymbol{b}_h^{\|}|}{|\boldsymbol{b}|}\right) \tag{11}$$

with

$$h_K^{\|} = \mathrm{diam}^*(K, \boldsymbol{b}_h^{\|})\,, \qquad \eta(\alpha) = 2\,\alpha\,(1-\alpha)\,. \tag{12}$$

This definition assures that the SUPG effect is not doubled if $\boldsymbol{b}_h^{\|} = \boldsymbol{b}$ and hence an ad hoc correction like (10) is not needed. Tezduyar and Park also observed that the SOLD term (9) with the above definitions of σ depends only on the direction of ∇u_h but not on its magnitude. Since the SOLD term is required only along steep gradients of the solution, they suggested to use

$$\sigma = \mu_K \, \frac{h_K^{\|}}{2\,|\boldsymbol{b}_h^{\|}|} \, \eta \left(\frac{|\boldsymbol{b}_h^{\|}|}{|\boldsymbol{b}|} \right) h_K^{\|} \, \frac{|\nabla u_h|}{u_0} \,, \tag{13}$$

where u_0 is a global scaling value for u_h.
An approach related to the above–described method of Hughes *et al.* [32] was used by de Sampaio and Coutinho [51], who introduced the concept of the effective transport velocity $\boldsymbol{b}^{\|}$ defined on the continuum level analogously as $\boldsymbol{b}_h^{\|}$ (i.e, with u instead of u_h). Before performing a discretization, the convective field $\boldsymbol{b}$ in (1) is replaced by $\tilde{\boldsymbol{b}} = \gamma\,\boldsymbol{b} + (1-\gamma)\,\boldsymbol{b}^{\|}$ with $\gamma \in [0,1]$. Then, an application of a standard discretization technique like the Galerkin/least–squares or, in our case, SUPG method yields a Petrov–Galerkin method containing a SOLD term. The method uses only one stabilization parameter (defined using the discrete counterpart of $\tilde{\boldsymbol{b}}$) and hence an alignment of $\boldsymbol{b}$ and ∇u does not create the undesirable doubling effect discussed above. However, it is not clear how to choose the parameter γ and, therefore, the value $\gamma = 0.5$ is recommended except for regions where $\nabla u = \mathbf{0}$.
An alternative approach to the above nonlinear SOLD methods is to modify the SUPG discretization (3) by adding artificial diffusion in the crosswind direction as considered by Johnson *et al.* [37] for the two–dimensional case with $\boldsymbol{b} = (1,0)$ and $u_b = 0$. A straightforward generalization of this approach leads to the additional term

$$(\widetilde{\varepsilon}\, D\, \nabla_h\, u_h, \nabla_h\, v_h) \tag{14}$$

on the left–hand side of (3), where

$$\widetilde{\varepsilon}|_K = \max\{0, |\boldsymbol{b}|\, h_K^{3/2} - \varepsilon\} \qquad \forall\ K \in \mathcal{T}_h \tag{15}$$

and D is the projection onto the line or plane orthogonal to $\boldsymbol{b}$ defined by

$$D = \begin{cases} I - \dfrac{\boldsymbol{b} \otimes \boldsymbol{b}}{|\boldsymbol{b}|^2} & \text{if } \boldsymbol{b} \neq \mathbf{0}, \\ 0 & \text{if } \boldsymbol{b} = \mathbf{0}, \end{cases}$$

I being the identity tensor. The value $h_K^{3/2}$ was motivated by a careful analysis of the numerical crosswind spread in the discrete problem, i.e., of the maximal distance in which the right–hand side f significantly influences the discrete solution. The resulting method is linear but non–consistent and hence it is restricted to finite elements of first order of accuracy. For the two–dimensional case with $\boldsymbol{b} = (1,0)$, $u_b = 0$ and a reaction term in (1), Johnson *et al.* [37] proved pointwise error estimates of order $O(h^{5/4})$ in regions of smoothness and a global L^1–estimate of order $O(h^{1/2})$. Later, these results were improved by Niijima [45], Zhou and Rannacher [58] and Zhou [57]. Note that, in the two–dimensional case, the SOLD term can be written in the form

$$(\widetilde{\varepsilon}\,\boldsymbol{b}^\perp\cdot\nabla_h\,u_h,\boldsymbol{b}^\perp\cdot\nabla_h\,v_h)\qquad\text{with}\qquad \boldsymbol{b}^\perp=\frac{(-b_2,b_1)}{|\boldsymbol{b}|}\,. \tag{16}$$

Shih and Elman [52] considered the SUPG discretization (3) with the additional term (16) for $\Omega=(0,1)^2$ and a constant vector $\boldsymbol{b}$. They used bilinear finite elements on a uniform triangulation of Ω and proposed two choices of the parameters τ and $\widetilde{\varepsilon}$ based on the requirement that the necessary conditions for uniform convergence of $\|u-u_h\|_{0,\Omega}$ of order greater than 1/2 introduced by Stynes and Tobiska [53] hold. However, both methods of Shih and Elman reduce to the SUPG discretization (3) whenever the flow vector $\boldsymbol{b}$ is aligned with the mesh, which indicates that the methods generally cannot work properly. Therefore, we do not consider them in our numerical tests.
Now, let us return to the SOLD term (9) which can be written in the form

$$(\widetilde{\varepsilon}\,\nabla_h\,u_h,\nabla_h\,v_h) \tag{17}$$

with

$$\widetilde{\varepsilon}=\begin{cases}\sigma\,\dfrac{R_h(u_h)\,\boldsymbol{b}\cdot\nabla u_h}{|\nabla u_h|^2} & \text{if } \nabla u_h\neq\mathbf{0},\\ 0 & \text{if } \nabla u_h=\mathbf{0}\,.\end{cases} \tag{18}$$

Galeão and do Carmo [26] observed that, when $f\neq 0$ in (1), this SOLD term does not prevent localized oscillations in the discrete solution. The reason is that this term introduces a negative artificial diffusion $\widetilde{\varepsilon}$ if $R_h(u_h)\,\boldsymbol{b}\cdot\nabla u_h<0$. As a remedy, Galeão and do Carmo proposed to replace the flow velocity $\boldsymbol{b}$ in the SUPG stabilization term by an approximate upwind direction

$$\boldsymbol{b}_h^{up}=\alpha_1\,\boldsymbol{b}+\alpha_2\,\boldsymbol{b}_h\,,$$

where $\boldsymbol{b}_h$ is an approximate streamline direction such that, for any $K\in\mathcal{T}_h$, the discrete solution u_h satisfies

$$-\varepsilon\,\Delta u_h+\boldsymbol{b}_h\cdot\nabla u_h=f\quad\text{in } K. \tag{19}$$

Of course, such $\boldsymbol{b}_h$ generally does not exist at those points of K at which $\nabla u_h=\mathbf{0}$. Therefore, we replace (19) by

$$(-\varepsilon\,\Delta u_h+\boldsymbol{b}_h\cdot\nabla u_h-f)\,|\nabla u_h|=0\quad\text{in } K. \tag{20}$$

A reasonable choice of $\boldsymbol{b}_h$ is $\boldsymbol{b}_h = \boldsymbol{b} - \boldsymbol{z}_h$ with

$$\boldsymbol{z}_h = \begin{cases} \dfrac{R_h(u_h)\,\nabla u_h}{|\nabla u_h|^2} & \text{if } \nabla u_h \neq \boldsymbol{0}, \\ \boldsymbol{0} & \text{if } \nabla u_h = \boldsymbol{0}, \end{cases}$$

since it minimizes $|\boldsymbol{b}_h - \boldsymbol{b}|$ in any $K \in \mathcal{T}_h$ among all functions $\boldsymbol{b}_h$ satisfying (20). Defining the SUPG stabilization using the approximate upwind direction $\boldsymbol{b}_h^{up}$, we obtain the discretization (3) with the additional term

$$(R_h(u_h), \sigma\, \boldsymbol{z}_h \cdot \nabla_h\, v_h) \tag{21}$$

on the left–hand side. The parameter $\tau \equiv \alpha_1 + \alpha_2$ is defined as before and the choice of $\sigma \equiv -\alpha_2$ will be discussed in the following. Note that, if $f = 0$ and $\Delta_h\, u_h = 0$ (which holds for (bi,tri)linear finite elements), we have $\boldsymbol{z}_h = \boldsymbol{b}_h^{\|}$. Hence, the terms (9) and (21) are the same provided that the parameters σ are defined appropriately. Galeão and do Carmo [26] use (21) with

$$\sigma = \max\{0, \tau(\boldsymbol{z}_h) - \tau(\boldsymbol{b})\}\,, \tag{22}$$

which is identical with (10) if $\boldsymbol{z}_h = \boldsymbol{b}_h^{\|}$. Do Carmo and Galeão [14] proposed to simplify (22) to

$$\sigma = \tau(\boldsymbol{b})\, \max\left\{0, \frac{|\boldsymbol{b}|}{|\boldsymbol{z}_h|} - 1\right\}\,, \tag{23}$$

which assures that the term (21) is added only if $|\boldsymbol{b}| > |\boldsymbol{z}_h|$, i.e., only if the above–introduced vector $\boldsymbol{b}_h$ satisfies the natural requirement $\boldsymbol{b} \cdot \boldsymbol{b}_h > 0$.

For problems with regular solutions, it was observed that the SOLD term (21) adds an undesirable crosswind diffusion and that the discrete solution is less accurate than for the SUPG method. Therefore, do Carmo and Galeão [14] introduced a feedback function which should minimize the influence of the SOLD term (21) in regions where the solution of (1) is smooth. Since the definition of the feedback function is rather involved, we only refer to [14].

The intricacy of the feedback approach of do Carmo and Galeão [14] motivated do Carmo and Alvarez [12] to introduce a simpler expression for the parameter σ. For this, the following parameters are used on any element $K \in \mathcal{T}_h$:

$$\begin{aligned} \alpha_K &= \frac{|\boldsymbol{z}_h|}{|\boldsymbol{b}|}\,, \qquad \beta_K = \min\{1, h_K\}^{1-\alpha_K^2}\,, \\ \gamma_K &= \min\{\beta_K, \tfrac{1}{2}\,(\alpha_K + \beta_K)\}\,, \qquad \lambda_K = \frac{\max\{\alpha_K, |R_h(u_h)|\}^{3+\alpha_K/2+\alpha_K^2}}{\gamma_K^{\max\{1/2, 1/4+\alpha_K\}}}\,, \\ \kappa_K &= |2 - \lambda_K|^{\frac{1-\lambda_K}{1+\lambda_K}} - 1\,, \qquad \omega_K = \frac{\alpha_K^2\, \gamma_K^{2-\alpha_K^2}}{\tau_K(\boldsymbol{b})}\,. \end{aligned}$$

Now, denoting by $\bar{\sigma}$ the value of σ defined by (22), do Carmo and Alvarez [12] consider (21) with

$$\sigma = \varrho\,\bar{\sigma}\,, \tag{24}$$

where

$$\varrho|_K = \begin{cases} 1 & \text{if } \alpha_K \geq 1 \text{ or } \lambda_K \geq 1, \\ [\omega_K\,\bar{\sigma}]^{\kappa_K} & \text{if } \alpha_K < 1 \text{ and } \lambda_K < 1 \end{cases} \qquad \forall\ K \in \mathcal{T}_h. \tag{25}$$

Like the above–mentioned feedback function, the parameter ϱ should suppress the addition of the artificial diffusion in regions where the solution of (1) is smooth.

In [13], do Carmo and Alvarez introduced a finer tuning of the parameters τ and σ by multiplying them by a factor τ_0 on those elements $K \in \mathcal{T}_h$ whose boundary intersects the outflow part of the boundary of Ω. The value of τ_0 on an element K depends on the geometry of K and the polynomial degree of shape functions on K. Based on numerical experiments, do Carmo and Alvarez set $\tau_0 = 1$ for bilinear shape functions on quadrilaterals, $\tau_0 = 0.5$ for biquadratic shape functions on quadrilaterals or linear shape functions on triangles and $\tau_0 = 0.25$ for quadratic shape functions on triangles.

A remedy for the above–mentioned loss of accuracy which appears when (21) with (22) or (23) is used was also proposed by Almeida and Silva [3], who conjectured that this loss of accuracy was mainly caused by the incapability of the formulas (22) and (23) to avoid the doubling effect. They observed that, setting $v_h = u_h$, the SUPG term in (3) becomes

$$(R_h(u_h), \tau\,\boldsymbol{b}\cdot\nabla_h\,u_h) = (R_h(u_h), \tau\,\vartheta_h\,\boldsymbol{z}_h\cdot\nabla_h\,u_h) \quad \text{with} \quad \vartheta_h = \frac{\boldsymbol{b}\cdot\nabla_h\,u_h}{R_h(u_h)}\,.$$

Therefore, they proposed to replace (23) by

$$\sigma = \tau(\boldsymbol{b})\,\max\left\{0, \frac{|\boldsymbol{b}|}{|\boldsymbol{z}_h|} - \zeta_h\right\} \quad \text{with} \quad \zeta_h = \max\left\{1, \frac{\boldsymbol{b}\cdot\nabla_h\,u_h}{R_h(u_h)}\right\}\,, \tag{26}$$

which provides a reduction of the amount of artificial diffusion along the $\boldsymbol{z}_h$ direction, which is the direction of the approximate solution gradient.

The SOLD term (21) can be written in the form (17) with

$$\widetilde{\varepsilon} = \begin{cases} \sigma\,\dfrac{|R_h(u_h)|^2}{|\nabla u_h|^2} & \text{if } \nabla u_h \neq \mathbf{0}, \\ 0 & \text{if } \nabla u_h = \mathbf{0}, \end{cases} \tag{27}$$

and hence it introduces an isotropic artificial diffusion. Since the streamline diffusion introduced by the SUPG method seems to be enough along the streamlines, Codina [17] proposed to add the artificial diffusion $\widetilde{\varepsilon}$ only in the crosswind direction like in (14) and, for any $K \in \mathcal{T}_h$, to set its amount to

$$\widetilde{\varepsilon}|_K = \frac{1}{2}\max\left\{0, C - \frac{2\,\varepsilon}{|\boldsymbol{b}_h^{\|}|\,\mathrm{diam}(K)}\right\}\mathrm{diam}(K)\,\frac{|R_h(u_h)|}{|\nabla u_h|} \quad \text{if } \nabla u_h \neq \mathbf{0}, \quad (28)$$

where C is a suitable constant. Codina [17] reports that two–dimensional numerical experiments suggest to set $C \approx 0.7$ for (bi)linear finite elements and $C \approx 0.35$ for (bi)quadratic finite elements. The design of (28) is based on investigations of the validity of the discrete maximum principle for several simple model problems and on the requirements that $\widetilde{\varepsilon}$ should be small in regions where $|\boldsymbol{b}\cdot\nabla u_h|$ is small (to avoid excessive overdamping) and proportional to the element residual (to guarantee consistency).

In order to be able to prove some theoretical results on SOLD methods of the above type, Knopp *et al.* [41] proposed to use (14) with $\widetilde{\varepsilon}$ defined, for any $K \in \mathcal{T}_h$, by

$$\widetilde{\varepsilon}|_K = \frac{1}{2}\max\left\{0, C - \frac{2\,\varepsilon}{Q_K(u_h)\,\mathrm{diam}(K)}\right\}\mathrm{diam}(K)\,Q_K(u_h)\,. \quad (29)$$

Here,

$$Q_K(u_h) = \frac{\|R_h(u_h)\|_{0,K}}{S_K + \|u_h\|_{1,K}} \quad (30)$$

and S_K are appropriate constants (equal to 1 in numerical experiments of [41]). This definition of $\widetilde{\varepsilon}$ was also motivated by a posteriori error estimates which show that the action of the SOLD stabilization should be restricted to regions where the local residual is not small.

Combining the above two definitions of $\widetilde{\varepsilon}$, we further propose to use (14) with $\widetilde{\varepsilon}$ defined by (29) where

$$Q_K(u_h) = \frac{|R_h(u_h)|}{|\nabla u_h|} \qquad \text{if } \nabla u_h \neq \mathbf{0}. \quad (31)$$

This is equivalent to (28) if $f = 0$ and $\Delta_h\, u_h = 0$. Another possibility is to set

$$Q_K(u_h) = \frac{\|R_h(u_h)\|_{0,K}}{|u_h|_{1,K}}\,. \quad (32)$$

Knopp *et al.* [41] also suggested to replace the isotropic artificial diffusion in (17) by

$$\widetilde{\varepsilon}|_K = \sigma_K(u_h)\,|Q_K(u_h)|^2 \qquad \forall\ K \in \mathcal{T}_h \quad (33)$$

with $Q_K(u_h)$ defined by (30) and some appropriate constants $\sigma_K(u_h) \geq 0$ (e.g., defined by (22) or (23)). Like in case of (29) with (30), this definition of $\widetilde{\varepsilon}$ was introduced to satisfy assumptions enabling Knopp *et al.* [41] to perform a priori and a posteriori error analyses of a rather general class of nonlinear discretizations of (1) which include SOLD discretizations with stabilizing terms defined by (14), (29), (30) or (17), (33), (30).

The SOLD term (17) was also used by Johnson [35], who proposed to set

$$\widetilde{\varepsilon}|_K = \max\{0, \alpha\,[\mathrm{diam}(K)]^\nu\,|R_h(u_h)| - \varepsilon\} \qquad \forall\ K \in \mathcal{T}_h \tag{34}$$

with some constants α and $\nu \in (3/2, 2)$. He suggested to take $\nu \sim 2$. Johnson [36] replaced α by $\beta/\max_\Omega |u_h|$ and proposed to set $\beta = 0.1$. A similar approach was also used by Johnson *et al.* [38]. A priori and a posteriori error estimates for this type of SOLD discretizations can be found in the papers by Johnson [35] and Eriksson and Johnson [22]. The mentioned papers [36] and [38] contain convergence results for space–time elements.

Burman and Ern [9] derived formulas for $\widetilde{\varepsilon}$ in (14) and (17) that guarantee a discrete maximum principle for strictly acute meshes and linear simplicial finite elements. However, they observed that, from a numerical viewpoint, the stronger one wishes to enforce a discrete maximum principle, the more ill behaved the nonlinear discrete equations become. Therefore, they slightly changed the formulas implied by the theoretical investigations and recommended to use (14) with $\widetilde{\varepsilon}$ defined, on any $K \in \mathcal{T}_h$, by

$$\widetilde{\varepsilon}|_K = \frac{\tau(\boldsymbol{b})\,|\boldsymbol{b}|^2\,|R_h(u_h)|}{|\boldsymbol{b}|\,|\nabla_h u_h| + |R_h(u_h)|}\;\frac{|\boldsymbol{b}|\,|\nabla_h u_h| + |R_h(u_h)| + \tan\alpha_K\,|\boldsymbol{b}|\,|D\,\nabla_h u_h|}{|R_h(u_h)| + \tan\alpha_K\,|\boldsymbol{b}|\,|D\,\nabla_h u_h|} \tag{35}$$

($\widetilde{\varepsilon} = 0$ if one of the denominators vanishes). The parameter α_K is equal to $\pi/2 - \beta_K$ where β_K is the largest angle of K if K is a triangle and β_K is the largest angle among the six pairs of faces of K if K is a tetrahedron. If $\beta_K = \pi/2$ (and hence the strictly acute condition is violated), it is recommended to set $\alpha_K = \pi/6$. Further, to improve the convergence of the nonlinear iterations, it is recommended to replace the absolute value $|x|$ of a real number x by the regularized expression $|x|_{\mathrm{reg}} \equiv x\,\tanh(x/\varepsilon_{\mathrm{reg}})$. We apply this regularization only to $|R_h(u_h)|$ and set $\varepsilon_{\mathrm{reg}} = 2$.

Our numerical experiments in Section 6 indicate that the above artificial diffusion $\widetilde{\varepsilon}$ is too large and therefore we also consider (14) with $\widetilde{\varepsilon}$ defined, on any $K \in \mathcal{T}_h$, by

$$\widetilde{\varepsilon}|_K = \frac{\tau(\boldsymbol{b})\,|\boldsymbol{b}|^2\,|R_h(u_h)|}{|\boldsymbol{b}|\,|\nabla_h u_h| + |R_h(u_h)|}\,. \tag{36}$$

In this case, we do not apply any regularization of the absolute values.

Another SOLD strategy for linear simplicial finite elements was introduced by Burman and Hansbo [11]. The SOLD term to be added to the left–hand side of (3) is defined by

$$\sum_{K\in\mathcal{T}_h}\int_{\partial K}\Psi_K(u_h)\,\mathrm{sign}(\boldsymbol{t}_{\partial K}\cdot\nabla(u_h|_K))\,\boldsymbol{t}_{\partial K}\cdot\nabla(v_h|_K)\,\mathrm{d}\sigma, \tag{37}$$

where $\boldsymbol{t}_{\partial K}$ is a tangent vector to the boundary ∂K of K,

$$\Psi_K(u_h)=\mathrm{diam}(K)\,(C_1\,\varepsilon+C_2\,\mathrm{diam}(K))\,\max_{E\subset\partial K}|\,[|\boldsymbol{n}_E\cdot\nabla u_h|]_E\,|\,, \tag{38}$$

$\boldsymbol{n}_E$ are normal vectors to edges E of K, $[|v|]_E$ denotes the jump of a function v across the edge E and C_1, C_2 are appropriate constants (note that C_2 has to be proportional to $|\boldsymbol{b}|$). Burman and Hansbo proved that, using an edge stabilization instead of the SUPG term, the discrete maximum principle is satisfied provided that $C_1\geq 1/2$ and C_2 is sufficiently large. In their numerical tests with $|\boldsymbol{b}|=1$, they used $C_2=10$. To improve the convergence of the nonlinear iterative process, they further regularize the sign operator in (37) by replacing it by tanh.

Burman and Ern [10] proposed to use the SOLD term (37) with $\Psi_K(u_h)$ defined by

$$\Psi_K(u_h)|_E=C\,|\boldsymbol{b}|\,[\mathrm{diam}(K)]^2\,|\,[|\nabla u_h|]_E\,|\qquad\forall\,E\subset\partial K\,, \tag{39}$$

where C is a suitable constant. For linear simplicial finite elements on weakly acute triangulations satisfying a local quasi–uniformity property, they proved the validity of a discrete maximum principle. Another definition of $\Psi_K(u_h)$ proposed in [10] is

$$\Psi_K(u_h)=C\,|R_h(u_h)|\,. \tag{40}$$

Let us mention that establishing a discrete maximum principle for higher order stabilized Galerkin methods still remains an open problem.

A further possibility of suppressing oscillations arising in the SUPG solution along boundary layers was considered by Lube [42], who presented an asymptotically fitted variant of the SUPG method. This method consists in replacing the Dirichlet boundary conditions on the downstream (if $\varepsilon<C\,h$) and characteristic (if $\varepsilon<h^{3/2}$) parts of the boundary by homogeneous Neumann's conditions. Existence, stability and convergence results are proved for (1) containing a suitable reaction term.

6 Numerical studies

The SOLD methods presented in Section 5 can be divided into three classes. These are SOLD methods which add isotropic additional diffusion (17),

Table 1: Summary of SOLD methods

name	citation	add. diffusion	method param.	user param.
MH85	[40]	upwind	-	-
HMM86	[32]	iso. (17)	(18), (10)	-
TP86_1	[56]	iso. (17)	(18), (11), (12)	-
TP86_2	[56]	iso. (17)	(18), (12), (13)	u_0
JSW87	[37]	orth. (14)	(15)	-
GdC88	[26]	iso. (17)	(27), (22)	-
dCG91	[14]	iso. (17)	(27), (23)	-
dCA03	[12]	iso. (17)	(27), (24), (25)	-
AS97	[3]	iso. (17)	(27), (26)	-
C93	[17]	orth. (14)	(28)	C
KLR02_1	[41]	orth. (14)	(29), (30)	C, S_K
KLR02_2	[41], here	orth. (14)	(29), (31)	C
KLR02_3	[41], here	orth. (14)	(29), (32)	C
KLR02_4	[41]	iso. (17)	(33), (22), (30)	S_K
J90	[35]	iso. (17)	(34)	α, ν
BE02_1	[9]	orth. (14)	(35)	α_K
BE02_2	[9], here	orth. (14)	(36)	-
BH04	[11]	edge (37)	(38)	C_1, C_2
BE05_1	[10]	edge (37)	(39)	C
BE05_2	[10]	edge (37)	(40)	C

SOLD methods which add the additional diffusion only orthogonally to the streamlines (14) and SOLD methods which rely upon edge stabilization. A summary of the most SOLD methods considered in Section 5, introducing also their abbreviations which will be used in the evaluation of the numerical examples, is presented in Table 1.

This paper presents results of two numerical examples which are defined in a two–dimensional domain and which are discretized by conforming piecewise linear finite elements. The only criterion for the evaluation of the SOLD methods will be the quality of the computed solution. This evaluation is twofold: the suppressing of spurious oscillations and the smearing of layers will be rated. Since spurious oscillations are far more undesirable than moderately smeared layers, the results concerning spurious oscillations will be weighted higher. We would like to note that the evaluation of the many computational results is rather complicated. The difficulty is that not errors to a known solution are of interest but the size of oscillations and the extent of smearing of layers. Measuring the size of oscillations is only easy if the

solution should be constant on both sides of the layer. Often, pictures of the computed solutions give a good impression of their quality. However, due to the considerable potential length of the paper, it is not possible to support each computation with one or even more pictures. Several measures for evaluating the results were tested in our numerical studies. We found out that the measures used below are appropriate ones.
The numerical results presented in this paper give only a first impression of the capabilities of the SOLD methods. Comprehensive numerical studies will be postponed to the second part of this paper. Open questions not treated in the present paper include:

- other two–dimensional and also three–dimensional examples,
- finite elements on simplices and quadrilaterals/hexahedra,
- higher order finite elements,
- nonconforming finite elements,
- parameter studies for SOLD methods with user–defined parameters,
- the efficiency of iterative schemes for solving the nonlinearities in the SOLD methods,
- the effect of replacing the parameter h_K from (6) by the diameter of K,
- the robustness of SOLD methods with respect to the Péclet number,
- the grid independency of the quality of the solution obtained with the SOLD methods,
- nonconstant data ε, $\boldsymbol{b}$ and f in (1),
- a comparison with results obtained on adaptively refined grids.

Now, let us come to a discussion of our computational results obtained using the methods from Table 1. The underlying SUPG method (3) was applied with τ defined by (5)–(8) using the upwind function ξ_0 from (4). We shall not mention any results for the method KLR02_3 since it is identical with KLR02_2 for the conforming P_1 finite element and constant data $\boldsymbol{b}$ and f in (1). The nonlinear problems were solved accurately, up to a norm of the residual lower than 10^{-10}. Methods which worked best in our opinion are printed boldly in the tables. Italic is used for methods which also produced acceptable results but which were clearly worse than the best methods. Most

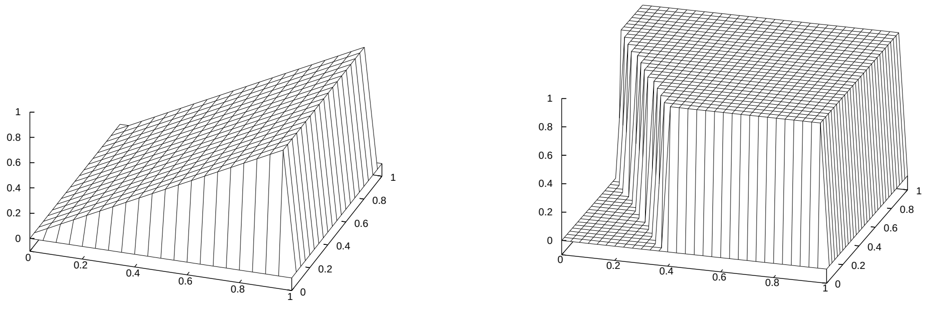

Figure 1: Solution of Example 1 (left) and of Example 2 (right).

of the numerical results have been double–checked by computing them with two different codes, one of them was *MooNMD*, [34].

Example 1. Solution with parabolic and exponential boundary layers. We consider the convection–diffusion equation (1) in $\Omega = (0,1)^2$ with $\varepsilon = 10^{-8}$, $\boldsymbol{b} = (1,0)^T$, $f = 1$ and $u_b = 0$. The solution $u(x,y)$ of this problem, see Figure 1, possesses an exponential boundary layer at $x = 1$ and parabolic boundary layers at $y = 0$ and $y = 1$. In the interior grid points, the solution $u(x,y)$ is very close to x.

The numerical tests were performed on a regular and on an unstructured triangular grid, see Figure 2 for the initial regular grid (Grid 1) and the final unstructured grid (Grid 3). The latter was obtained using the anisotropic mesh adaptation technique of [19].

First, we present computations on Grid 1 where the length of the legs of the triangles was 1/64. Thus, from (6) follows $h_K = 1/64$ and the Péclet number is $\mathrm{Pe}_K = 10^8/128 = 781250$. The number of degrees of freedom is

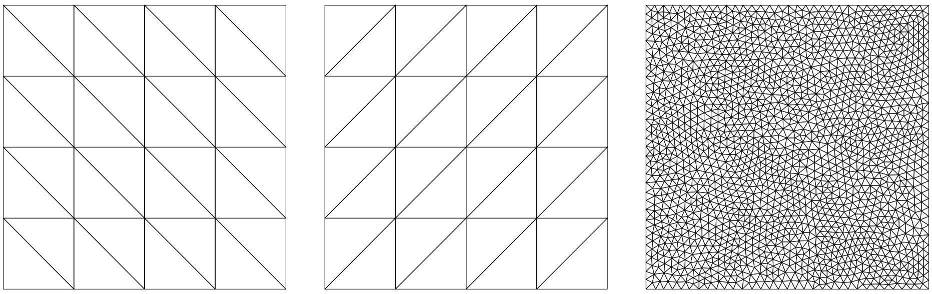

Figure 2: The grids used in the computations: Grid 1, Grid 2 and Grid 3 (left to right). The structured grids are refined till the length of the legs of the triangles is 1/64.

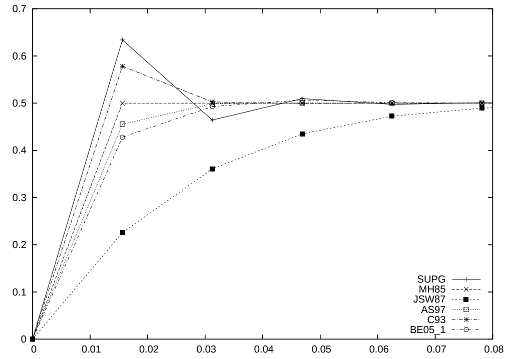

Figure 3: Example 1, Grid 1, the parabolic boundary layer computed with different schemes.

4225 (including Dirichlet nodes).
For this special example, the stabilization parameter τ used in this paper is optimal along lines y = const outside the parabolic layers. Applying the SUPG method on Grid 1, one finds that there are no oscillations at the exponential layer. However, there are still strong oscillations at the parabolic layers and for this reason we will concentrate on these layers in the evaluation of the SOLD methods on Grid 1. Particularly, we consider the cut line $x = 0.5$ and the values

$$\begin{aligned} osc &:= \max_{y\in\left\{\frac{1}{64},\frac{2}{64},\ldots,\frac{63}{64}\right\}} \left\{u_h(0.5,y) - u_h(0.5,0.5)\right\}, && (41) \\ smear &:= \min_{y\in\left\{\frac{1}{64},\frac{2}{64},\ldots,\frac{63}{64}\right\}} \left\{u_h(0.5,y) - u_h(0.5,0.5)\right\}. && (42) \end{aligned}$$

The first value measures the oscillations in the parabolic layers. In the case that the oscillations are suppressed to the most part, the second value measures the smearing of these layers. The computational results are given in Table 2 and Figure 3. To simplify their evaluation and the ranking of the methods, we scored each result. The scores are as follows:

$osc \in$	score	$smear \in$	score
$[0, 1\text{e-}3)$	4	$(-1\text{e-}5, 0]$	2
$[1\text{e-}3, 1\text{e-}2)$	2	$(-1\text{e-}3, -1\text{e-}5]$	1
$[1\text{e-}2, 1\text{e-}1)$	0	$(-1\text{e-}1, -1\text{e-}3])$	0
$[1\text{e-}1, 1)$	-4	$(-1, -1\text{e-}1]$	-2

Values which are close to the interval with the next higher score will get an intermediate score.

Table 2: Example 1, Grid 1, *osc* and *smear* defined in (41) and (42).

name	*osc*	score	*smear*	score	total
SUPG	1.340e-1	-4	-	-	-4
MH85	0	4	-5.280e-6	2	6
HMM86	8.737e-2	0	-1.141e-2	0	0
TP86_1	1.150e-1	-4	-	-	-4
TP86_2; $u_0 = 1$	1.312e-1	-4	-	-	-4
JSW87	1.479e-6	4	-2.743e-1	-2	2
GdC88	2.179e-3	2	-4.860e-2	0	2
dCG91	5.992e-4	4	-4.515e-2	0	4
dCA03	1.316e-2	1	-4.387e-2	0	1
AS97	4.742e-4	4	-4.494e-2	0	4
C93; $C = 0.6$	7.816e-2	0	-8.076e-4	1	1
KLR02_1; $C = 0.6, S_K = 1$	9.654e-2	0	-2.383e-2	0	0
KLR02_2; $C = 0.6$	2.469e-4	4	-3.680e-2	0	4
KLR02_4; $S_K = 1$	1.241e-1	-4	-	-	-4
J90; $\alpha = 0.3, \nu = 2$	5.465e-2	0	-1.299e-2	0	0
BE02_1; $\alpha_K = \pi/6$	1.528e-2	1	-9.184e-2	0	1
BE02_2	6.942e-4	4	-4.729e-2	0	4
BH04; $C_1 = 0.5, C_2 = 0.01$	2.477e-3	2	-2.168e-1	-2	0
BE05_1; $C = 0.05$	6.765e-3	2	-7.212e-2	0	2
BE05_2; $C = 5\text{e-}5$	2.826e-3	2	-1.489e-1	-2	0

Clearly the best method is MH85. Good results were computed also with dCG91, AS97, KLR02_2 and BE02_2. All other methods, save JSW87, still exhibit non–negligible spurious oscillations at the parabolic layers. These layers are smeared considerably in the solution computed with JSW87. In addition, we want to note that the solutions obtained with J90, BH04 and BE05_2 show, in contrast to all other methods, a smearing of the exponential boundary layer.

Table 3, Figure 4 and Figure 5 present results obtained on the unstructured Grid 3 from Figure 2. This grid possesses 3312 triangles and 1721 vertices (degrees of freedom). Introducing the sets

$$\begin{aligned} \Omega_1 &= \Omega_2 \cup \Omega_3\,, \qquad \Omega_2 = (0,0.9) \times (0,0.1]\,, \qquad \Omega_3 = (0,0.9) \times [0.9,1)\,, \\ \Omega_4 &= [0.9,1) \times (0.1,0.9), \end{aligned}$$

the spurious oscillations are measured by

$$\begin{aligned} osc_{\text{para}(1)} &:= \max_{(x,y)\in\Omega_1} (u_h(x,y) - x)\,, & (43)\\ osc_{\text{para}(2)} &:= \max\left\{ \max_{(x_s,y_s)\in\Omega_2} \left(-\frac{\partial u_h(x_s,y_s)}{\partial y}\right), \max_{(x_s,y_s)\in\Omega_3} \frac{\partial u_h(x_s,y_s)}{\partial y}\right\}, & (44)\\ osc_{\text{exp}} &:= \max_{(x_s,y_s)\in\Omega_4} \frac{\partial u_h(x_s,y_s)}{\partial x}\,, & (45) \end{aligned}$$

where (x, y) are the nodes in Ω_1 and (x_s, y_s) are the coordinates of the barycentres of the triangles. The optimal value of $osc_{\text{para}(2)}$ is zero and of osc_{exp} is one. The larger these values are, the stronger are the oscillations in the parabolic and exponential layer, respectively. For evaluating the extent of the global smearing, the value

$$smear := \left(\sum_{\text{interior nodes } (x,y)} \Big(\min\left\{0, u_h(x,y) - x\right\} \Big)^2 \right)^{1/2} \qquad (46)$$

is computed. The rating of the results is as follows:

$osc_{\text{para}(1)} \in$	sc.	$osc_{\text{para}(2)} \in$	sc.	$osc_{\text{exp}} \in$	sc.	$smear \in$	sc.
[0, 1e-3)	2	[0, 1e-1)	2	$[1, 1.25)$	4	$[0, 1.25)$	2
[1e-3, 1e-2)	1	[1e-1, 3e-1)	1	$[1.25, 2)$	2	$[1.25, 2)$	1
[1e-2, 1e-1)	0	[3e-1, 1)	0	$[2, 3)$	0	$[2, 3)$	0
[1e-1, 1)	-2	$[1, 10)$	-2	$[3, 5)$	-4	$[3, 5)$	-2

Values which are close to the interval with the next higher score will get an intermediate score. Since there are two criteria for the oscillations in the parabolic layers, the score of each is half of the score of osc_{exp}.
For MH85 and HMM86, we were not able to solve the nonlinear problems. It is remarkable that only the edge stabilization schemes BH04, BE05_1 and BE05_2 and the method J90 were able to compute solutions almost without spurious oscillations at the exponential layer, see Table 3 and Figure 4. The results at the exponential layer obtained with the most other methods are similar to the result of KLR02_2 in the middle of Figure 4. However, the edge stabilization schemes lead to a larger smearing of layers, see Figure 5 for the parabolic layer at $y = 0$. The method J90 produces rather large spurious oscillations in the parabolic layers. Altogether, BH04, BE05_1 and BE05_2 worked best on the unstructured Grid 3 since these methods suppressed the spurious oscillations at the exponential layer well and they

Table 3: Example 1, Grid 3, the measures for evaluating the oscillations and the smearing are defined in (43)–(46), the parameters in the SOLD methods are the same as in Table 2.

name	$osc_{\text{para}(1)}$	sc.	$osc_{\text{para}(2)}$	sc.	osc_{exp}	sc.	$smear$	sc.	total
SUPG	1.545e-1	-2	7.883e+0	-2	4.972	-4	8.550e-1	2	-6
MH85	no conv.		–		–		–		
HMM86	no conv.		–		–		–		
TP86_1	9.225e-2	0	3.612e+0	-2	2.771	0	9.164e-1	2	0
TP86_2	1.291e-1	-2	6.369e+0	-2	2.968	0	9.125e-1	2	-2
JSW87	6.167e-4	2	2.002e-2	2	2.250	0	4.247e+0	-2	2
GdC88	7.103e-3	1	2.679e-1	1	2.702	0	1.711e+0	1	3
dCG91	7.048e-3	1	2.746e-1	1	2.675	0	1.846e+0	1	3
dCA03	1.191e-2	0.5	5.550e-1	0	2.695	0	1.720e+0	1	1.5
AS97	8.961e-3	1	4.336e-1	0	2.876	0	1.849e+0	1	2
C93	2.416e-2	0	8.591e-1	0	2.823	0	1.131e+0	2	2
KLR02_1	9.862e-2	0	4.741e+0	-2	2.420	0	1.047e+0	2	0
KLR02_2	2.829e-3	1.5	1.112e-1	1.5	2.823	0	1.549e+0	1	4
KLR02_4	1.313e-1	-2	6.786e+0	-2	4.563	-4	9.508e-1	2	-6
J90	5.759e-2	0	2.128e+0	-2	1.183	4	2.253e+0	0	2
BE02_1	5.336e-3	1	2.189e-1	1	3.224	-2	2.177e+0	0	0
BE02_2	2.604e-3	1.5	1.030e-1	1.5	2.320	0	1.826e+0	1	4
BH04	8.941e-3	1	3.549e-1	0.5	1.086	4	2.309e+0	0	5.5
BE05_1	5.431e-3	1	1.998e-1	1	1.075	4	2.211e+0	0	6
BE05_2	8.367e-3	1	3.417e-1	0.5	1.080	4	2.013e+0	0.5	6

worked also relatively well in the parabolic layers. A second group of methods, GdC88, dCG91, KLR02_2 and BE02_2, computed good results outside the exponential layer.

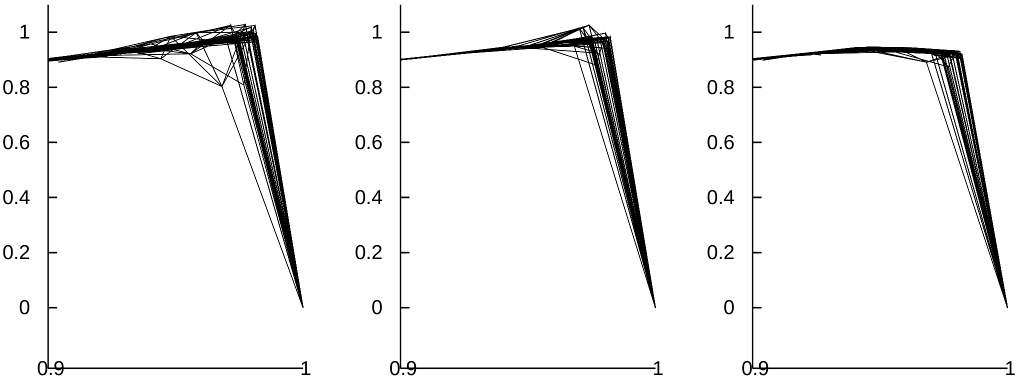

Figure 4: Example 1, the exponential boundary layer computed with SUPG, KLR02_2 and BH04 (left to right) on Grid 3, $(x, y) \in [0.9, 1] \times [0.1, 0.9]$.

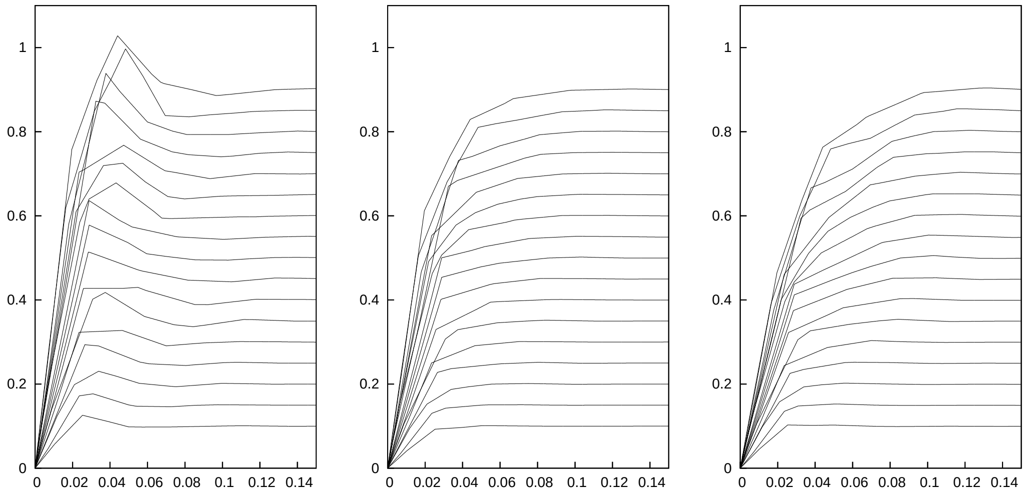

Figure 5: Example 1, the parabolic boundary layer at $y = 0$ computed with SUPG, KLR02_2 and BH04 (left to right) on Grid 3, cuts of the solution at $x \in \{0.1, 0.15, 0.2, \ldots, 0.9\}$.

Example 2. Solution with interior layer and exponential boundary layer. The convection–diffusion equation (1) is considered in $\Omega = (0,1)^2$ with the data $\varepsilon = 10^{-8}$, $\boldsymbol{b} = (\cos(-\pi/3), \sin(-\pi/3))^T$, $f = 0$ and

$$u_b(x,y) = \begin{cases} 0 & \text{for } x = 1 \text{ or } y \le 0.7, \\ 1 & \text{else.} \end{cases}$$

The solution, see Figure 1, possesses an interior layer in the direction of the convection starting at $(0, 0.7)$. On the boundary $x = 1$ and on the right part of the boundary $y = 0$, exponential layers are developed. This example has been used, e.g., in [32].

The computations were performed on the regular triangular Grid 1 and Grid 2, see Figure 2 for the initial grids. For these grids, the convection is skew to the grid lines. The grid size in the computations was chosen to be 1/64 (length of the legs of the triangles) such that the Péclet number is $\text{Pe}_K = 781250$ and the number of degrees of freedom 4225. Since the right–hand side of (1) vanishes, the following methods are the identical ones: HMM86 and GdC88; dCG91 and AS97; C93 and KLR02_2.

Denoting

$$\Omega_1 = \{(x,y) \in \Omega\,;\; x \le 0.5,\, y \ge 0.1\}\,, \qquad \Omega_2 = \{(x,y) \in \Omega\,;\; x \ge 0.7\}\,,$$

the following quantities are considered for assessing the computational results:

$$osc_{\text{int}} := \left(\sum_{(x,y)\in\Omega_1} \left(\min\{0, u_h(x,y)\}\right)^2 + \left(\max\{0, u_h(x,y)-1\}\right)^2 \right)^{1/2} \tag{47}$$

$$osc_{\text{exp}} := \left(\sum_{(x,y)\in\Omega_2} \left(\max\{0, u_h(x,y)-1\}\right)^2 \right)^{1/2}, \tag{48}$$

$$smear_{\text{int}} := x_2 - x_1\,, \tag{49}$$

$$smear_{\text{exp}} := \left(\sum_{(x,y)\in\Omega_2} \left(\min\{0, u_h(x,y)-1\}\right)^2 \right)^{1/2}, \tag{50}$$

where x_1 is the x–coordinate of the first point on the cut line $(x, 0.25)$ with $u_h(x_1, 0.25) \geq 0.1$ and x_2 is the x–coordinate of the first point with $u_h(x_2, 0.25)$
≥ 0.9. Thus, (49) gives a measure for the thickness of the interior layer. The evaluation of x_1 and x_2 used a grid with mesh width 10^{-5} on the cut line. The summations are performed over the nodes (x, y) of the meshes. Results of the computations on Grid 1 are presented in Table 4. The scoring of the results is as follows:

$osc_{\text{int}} \in$	sc.	$osc_{\text{exp}} \in$	sc.	$smear_{\text{int}} \in$	sc.	$smear_{\text{exp}} \in$	sc.
[0, 1e-4)	4	[0, 1e-5)	4	[0, 4e-2)	2	[0, 1e-4)	2
[1e-4, 1e-2)	2	[1e-5, 1e-3)	2	[4e-2, 6e-2)	1	[1e-4, 1e-2)	1
[1e-2, 1e-1)	0	[1e-3, 1e-1)	0	[6e-2, 8e-2)	0	[1e-2, 5e-1)	0
[1e-1, 1)	-4	[1e-1, 10)	-4	[8e-2, 1)	-2	[5e-1, 10)	-2

Again, intermediate scores are used.
The method MH85 gives an almost perfect result. Only the interior layer is smeared somewhat. Quite good results are obtained also with dCG91, AS97 and BE02_2. We observed for all SOLD methods that there are no spurious oscillations in the exponential layer at $y = 0$ on Grid 1, see also Figure 6. Comparing the results on Grid 1 and Grid 2, one finds that the results on Grid 2 are considerably worse, see Tables 4 and 5. Because of this, the conditions for rating the methods are relaxed somewhat:

$osc_{\text{int}} \in$	sc.	$osc_{\text{exp}} \in$	sc.	$smear_{\text{int}} \in$	sc.	$smear_{\text{exp}} \in$	sc.
[0, 1e-3)	4	[0, 1e-3)	4	[0, 5e-2)	2	[0, 1e-4)	2
[1e-3, 1e-2)	2	[1e-3, 2.5e-1)	2	[5e-2, 8e-2)	1	[1e-4, 1e-2)	1
[1e-2, 1e-1)	0	[2.5e-1, 1)	0	[8e-2, 1.1e-1)	0	[1e-2, 5e-1)	0
[1e-1, 1)	-4	[1, 10)	-4	[1.1e-1, 1)	-2	[5e-1, 10)	-2

Table 4: Example 2, Grid 1 from Figure 2, the measures for evaluating the oscillations and the smearing are defined in (47)–(50), the parameters in the SOLD methods are the same as in Table 2.

name	$osc_{\rm int}$	sc.	$osc_{\rm exp}$	sc.	$smear_{\rm int}$	sc.	$smear_{\rm exp}$	sc.	total
SUPG	5.891e-1	-4	2.124e+0	-4	3.747e-2	2	5.666e-1	-1	-7
MH85	6.081e-13	4	0	4	5.792e-2	1	1.083e-5	2	11
HMM86, GdC88	1.185e-1	-2	3.010e-2	0	5.927e-2	1	2.921e-3	1	0
TP86_1	2.038e-1	-4	2.581e-6	4	4.020e-2	1.5	5.445e-1	-1	0.5
TP86_2	4.700e-1	-4	5.972e-2	0	3.852e-2	2	4.768e-1	0	-2
JSW87	5.440e-11	4	1.007e-4	2	1.473e-1	-2	2.656e-1	0	4
dCG91, AS97	1.248e-5	4	1.482e-10	4	7.090e-2	0	6.479e-1	-1	7
dCA03	1.299e-1	-2	3.019e-2	0	6.074e-2	0.5	3.220e-3	1	-0.5
C93, KLR02_2	4.278e-3	2	1.959e-5	3	6.677e-2	0	9.042e-1	-2	3
KLR02_1	2.990e-1	-4	6.240e-1	-4	4.247e-2	1	2.292e-1	0	-7
KLR02_4	5.256e-1	-4	1.589e+0	-4	3.852e-2	2	4.118e-1	0	-6
J90	1.276e-1	-2	1.325e-1	-2	5.106e-2	1	2.108e+0	-2	-5
BE02_1	1.083e-2	1	9.488e-4	2	7.527e-2	0	2.274e+0	-2	1
BE02_2	2.470e-8	4	2.546e-5	3	7.132e-2	0	6.723e-1	-1	6
BH04	1.754e-2	1	5.063e-1	-4	7.106e-2	0	3.793e-1	0	-3
BE05_1	4.906e-3	2	1.904e+0	-4	9.685e-2	-2	4.520e-1	0	-4
BE05_2	4.580e-3	2	1.648e-4	2	7.930e-2	0	3.867e+0	-2	2

To obtain a better classification of the methods, intermediate values are used as in the other tests.

The only method which worked still very good was MH85. Only the smearing of the interior layer became somewhat larger in comparison to Grid 1. None of the other SOLD schemes produced a satisfactory solution with respect to all criteria of evaluation. It is remarkable that methods which worked well on Grid 1 completely failed on Grid 2, see Figure 6 for dCG91 and AS97. Two other results are presented in Figure 7. It can be seen that the solution computed with HMM86, GdC88 has a big oscillation at the starting point of the interior layer and another one in a vicinity of the corner $(1, 0)$ of Ω. The smearing of the layers which led to bad scores for BE05_2 is clearly visible in the right picture of Figure 7.

A reason for the bad results obtained with the SOLD methods on Grid 2 can be found, in our opinion, already in the underlying SUPG stabilization. Since the SUPG method gives on Grid 2 considerably worse results than on Grid 1, there is not sufficient diffusion introduced in the streamline direction. However, the SOLD methods introduce additional diffusion above all orthogonally to the streamlines and rely upon the assumption that the SUPG method has done a good job in the streamline direction. If this is not the case, the SOLD methods give rather poor results as this example shows.

Table 5: Example 2, Grid 2 from Figure 2, the measures for evaluating the oscillations and the smearing are defined in (47)–(50), the parameters in the SOLD methods are the same as in Table 2.

name	osc_{int}	sc.	osc_{exp}	sc.	$smear_{\mathrm{int}}$	sc.	$smear_{\mathrm{exp}}$	sc.	total
SUPG	6.925e-1	-4	3.847e+0	-4	6.206e-2	1	1.698e+0	-2	-9
MH85	0	4	0	4	1.024e-1	0	1.161e-5	2	10
HMM86, GdC88	2.176e-1	-3	1.279e-1	2	1.037e-1	0	2.480e-3	1	0
TP86_1	2.719e-1	-3	6.713e-1	0	7.424e-2	1	4.586e-2	0	-2
TP86_2	5.509e-1	-4	5.489e-1	0	6.498e-2	1	1.952e-1	0	-3
JSW87	2.444e-1	-3	2.133e+0	-4	1.117e-1	-1	5.005e-1	0	-8
dCG91, AS97	2.971e-1	-3	1.406e+0	-4	8.544e-2	0	2.114e-1	0	-7
dCA03	2.204e-1	-3	1.279e-1	2	1.060e-1	0	2.527e-3	1	0
C93, KLR02_2	1.386e-1	-2	3.606e-1	0	9.750e-2	0	3.126e-2	0	-2
KLR02_1	5.125e-1	-4	1.773e+0	-4	6.671e-2	1	5.941e-1	-1	-8
KLR02_4	6.629e-1	-4	2.681e+0	-4	6.309e-2	1	1.080e+0	-2	-9
J90	3.911e-1	-4	3.053e-1	1	7.241e-2	1	1.601e+0	-2	-4
BE02_1	1.496e-1	-2	4.306e-1	0	1.034e-1	0	3.651e-1	0	-2
BE02_2	2.214e-1	-3	1.396e+0	-4	8.634e-2	0	2.102e-1	0	-7
BH04	9.224e-2	0	1.548e+0	-4	9.966e-2	0	1.408e-1	0	-4
BE05_1	6.153e-3	2	3.514e+0	-4	1.528e-1	-2	1.402e+0	-2	-6
BE05_2	6.470e-3	2	2.163e-3	3	1.435e-1	-2	3.411e+0	-2	1

Summary of the numerical studies. The numerical tests were performed in a two–dimensional domain using the conforming P_1 finite element. Under these conditions, the upwind method MH85 was always the best method on regular grids. Among the other SOLD methods, no one could be preferred in all cases. The methods dCG91 and BE02_2 were often among the best ones. There are also some methods which never produced good results, e.g., TP86_1 and TP86_2 introduce in general not enough artificial diffusion to damp the oscillations sufficiently or JSW87 and J90 are very diffusive and smear the layers considerably. Altogether, there are still many open questions to be answered (see the beginning of this section), which will be started in the second part of this paper.

7 Conclusions and outlook

A characteristic feature of numerical solutions of scalar convection–dominated convection–diffusion equations computed with the popular SUPG stabilization is the presence of quite large spurious oscillations at layers. The main goal of SOLD methods consists in suppressing these oscillations without an excessive smearing of the layers. The present paper gave a review of the state of the art of SOLD methods. These methods, save MH85, can be classified

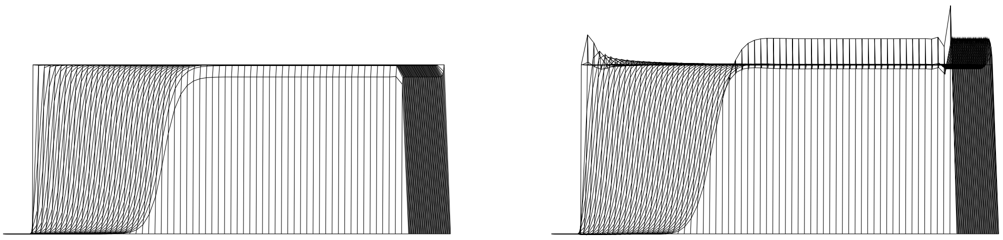

Figure 6: Example 2, solutions obtained with dCG91 (AS97); left: on Grid 1, right: on Grid 2.

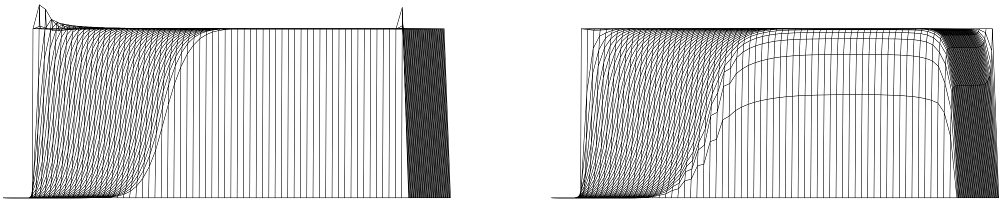

Figure 7: Example 2, solutions obtained on Grid 2 with HMM86, GdC88 (left) and BE05_2 (right).

into methods adding isotropic diffusion, methods adding diffusion orthogonally to the streamlines and into edge stabilization methods. Some numerical studies gave a first impression of the behavior of the SOLD methods.
The evaluation of comprehensive numerical tests, studying the aspects which are mentioned at the beginning of Section 6, will be subject of the second part of this paper.

Acknowledgments

The research of Petr Knobloch is a part of the project MSM 0021620839 financed by MSMT and it was partly supported by the Grant Agency of the Academy of Sciences of the Czech Republic under the grant No. IAA100190505.

References

[1] J.E. Akin, T.E. Tezduyar, Calculation of the advective limit of the SUPG stabilization parameter for linear and higher–order elements, Comput. Methods Appl. Mech. Eng. 193 (2004) 1909–1922.

[2] J.E. Akin, T.E. Tezduyar, M. Ungor, S. Mittal, Stabilization parameters and Smagorinski turbulence model, J. Appl. Mech. 70 (2003) 2–9.

[3] R.C. Almeida, R.S. Silva, A stable Petrov–Galerkin method for convection–dominated problems, Comput. Methods Appl. Mech. Eng. 140 (1997) 291–304.

[4] K. Baba, M. Tabata, On a conservative upwind finite element scheme for convective diffusion equations, RAIRO, Anal. Numér. 15 (1981) 3–25.

[5] F. Brezzi, L.P. Franca, A. Russo, Further considerations on residual–free bubbles for advective–diffusive equations, Comput. Methods Appl. Mech. Eng. 166 (1998) 25–33.

[6] F. Brezzi, D. Marini, E. Süli, Residual–free bubbles for advection–diffusion problems: The general error analysis, Numer. Math. 85 (2000) 31–47.

[7] F. Brezzi, A. Russo, Choosing bubbles for advection–diffusion problems, Math. Models Methods Appl. Sci. 4 (1994) 571–587.

[8] A.N. Brooks, T.J.R. Hughes, Streamline upwind/Petrov–Galerkin formulations for convection dominated flows with particular emphasis on the incompressible Navier–Stokes equations, Comput. Methods Appl. Mech. Eng. 32 (1982) 199–259.

[9] E. Burman, A. Ern, Nonlinear diffusion and discrete maximum principle for stabilized Galerkin approximations of the convection–diffusion–reaction equation, Comput. Methods Appl. Mech. Eng. 191 (2002) 3833–3855.

[10] E. Burman, A. Ern, Stabilized Galerkin approximation of convection–diffusion–reaction equations: Discrete maximum principle and convergence, Math. Comput. 74 (2005) 1637–1652.

[11] E. Burman, P. Hansbo, Edge stabilization for Galerkin approximations of convection–diffusion–reaction problems, Comput. Methods Appl. Mech. Eng. 193 (2004) 1437–1453.

[12] E.G.D. do Carmo, G.B. Alvarez, A new stabilized finite element formulation for scalar convection–diffusion problems: The streamline and approximate upwind/Petrov–Galerkin method, Comput. Methods Appl. Mech. Eng. 192 (2003) 3379–3396.

[13] E.G.D. do Carmo, G.B. Alvarez, A new upwind function in stabilized finite element formulations, using linear and quadratic elements for scalar convection–diffusion problems, Comput. Methods Appl. Mech. Eng. 193 (2004) 2383–2402.

[14] E.G.D. do Carmo, A.C. Galeão, Feedback Petrov–Galerkin methods for convection–dominated problems, Comput. Methods Appl. Mech. Eng. 88 (1991) 1–16.

[15] I. Christie, D.F. Griffiths, A.R. Mitchell, O.C. Zienkiewicz, Finite element methods for second order differential equations with significant first derivatives, Int. J. Numer. Methods Eng. 10 (1976) 1389–1396.

[16] P.G. Ciarlet, Basic error estimates for elliptic problems, in: P.G. Ciarlet, J.L. Lions (Eds.), Handbook of Numerical Analysis, v. 2 – Finite Element Methods (pt. 1), North–Holland, Amsterdam, 1991, pp. 17–351.

[17] R. Codina, A discontinuity–capturing crosswind–dissipation for the finite element solution of the convection–diffusion equation, Comput. Methods Appl. Mech. Eng. 110 (1993) 325–342.

[18] R. Codina, E. Oñate, M. Cervera, The intrinsic time for the streamline upwind/Petrov–Galerkin formulation using quadratic elements, Comput. Methods Appl. Mech. Eng. 94 (1992) 239–262.

[19] V. Dolejší, Anisotropic mesh adaptation for finite volume and finite element methods on triangular meshes, Comput. Vis. Sci. 1 (1998) 165–178.

[20] J. Douglas jun., J. Wang, An absolutely stabilized finite element method for the Stokes problem, Math. Comput. 52 (1989) 495–508.

[21] H.C. Elman, A. Ramage, An analysis of smoothing effects of upwinding strategies for the convection–diffusion equation, SIAM J. Numer. Anal. 40 (2002) 254–281.

[22] K. Eriksson, C. Johnson, Adaptive streamline diffusion finite element methods for stationary convection–diffusion problems, Math. Comput. 60 (1993) 167–188.

[23] L.P. Franca, S.L. Frey, T.J.R. Hughes, Stabilized finite element methods. I.: Application to the advective–diffusive model, Comput. Methods Appl. Mech. Eng. 95 (1992) 253–276.

[24] B. Fischer, A. Ramage, D.J. Silvester, A.J. Wathen, On parameter choice and iterative convergence for stabilised discretisations of advection–diffusion problems, Comput. Methods Appl. Mech. Eng. 179 (1999) 179–195.

[25] A.C. Galeão, R.C. Almeida, S.M.C. Malta, A.F.D. Loula, Finite element analysis of convection dominated reaction–diffusion problems, Appl. Numer. Math. 48 (2004) 205–222.

[26] A.C. Galeão, E.G.D. do Carmo, A consistent approximate upwind Petrov–Galerkin method for convection–dominated problems, Comput. Methods Appl. Mech. Eng. 68 (1988) 83–95.

[27] I. Harari, L.P. Franca, S.P. Oliveira, Streamline design of stability parameters for advection–diffusion problems, J. Comput. Phys. 171 (2001) 115–131.

[28] J.C. Heinrich, P.S. Huyakorn, O.C. Zienkiewicz, A.R. Mitchell, An 'upwind' finite element scheme for two–dimensional convective transport equation, Int. J. Numer. Methods Eng. 11 (1977) 131–143.

[29] J.C. Heinrich, O.C. Zienkiewicz, Quadratic finite element schemes for two–dimensional convective–transport problems, Int. J. Numer. Methods Eng. 11 (1977) 1831–1844.

[30] T.J.R. Hughes, Multiscale phenomena: Green's functions, the Dirichlet–to–Neumann formulation, subgrid scale models, bubbles and the origins of stabilized methods, Comput. Methods Appl. Mech. Eng. 127 (1995) 387–401.

[31] T.J.R. Hughes, L.P. Franca, G.M. Hulbert, A new finite element formulation for computational fluid dynamics. VIII. The Galerkin/least–squares method for advective–diffusive equations, Comput. Methods Appl. Mech. Eng. 73 (1989) 173–189.

[32] T.J.R. Hughes, M. Mallet, A. Mizukami, A new finite element formulation for computational fluid dynamics: II. Beyond SUPG, Comput. Methods Appl. Mech. Eng. 54 (1986) 341–355.

[33] T. Ikeda, Maximum Principle in Finite Element Models for Convection–Diffusion Phenomena. Lecture Notes in Numerical and Applied Analysis, Vol. 4. North–Holland, Amsterdam, 1983.

[34] V. John, G. Matthies, MooNMD – a program package based on mapped finite element methods, Comput. Visual. Sci. 6 (2004) 163–170.

[35] C. Johnson, Adaptive finite element methods for diffusion and convection problems, Comput. Methods Appl. Mech. Eng. 82 (1990) 301–322.

[36] C. Johnson, A new approach to algorithms for convection problems which are based on exact transport + projection, Comput. Methods Appl. Mech. Eng. 100 (1992) 45–62.

[37] C. Johnson, A.H. Schatz, L.B. Wahlbin, Crosswind smear and pointwise errors in streamline diffusion finite element methods, Math. Comput. 49 (1987) 25–38.

[38] C. Johnson, A. Szepessy, P. Hansbo, On the convergence of shock–capturing streamline diffusion finite element methods for hyperbolic conservation laws, Math. Comput. 54 (1990) 107–129.

[39] H. Kanayama, Discrete models for salinity distribution in a bay: conservation law and maximum principle, Theoretical Appl. Mech. 28 (1978) 559–579.

[40] P. Knobloch, Improvements of the Mizukami–Hughes method for convection–diffusion equations, Preprint No. MATH–knm–2005/6, Faculty of Mathematics and Physics, Charles University, Prague, 2005.

[41] T. Knopp, G. Lube, G. Rapin, Stabilized finite element methods with shock capturing for advection–diffusion problems, Comput. Methods Appl. Mech. Eng. 191 (2002) 2997–3013.

[42] G. Lube, An asymptotically fitted finite element method for convection dominated convection–diffusion–reaction problems, Z. Angew. Math. Mech. 72 (1992) 189–200.

[43] A. Mizukami, An implementation of the streamline–upwind/Petrov–Galerkin method for linear triangular elements, Comput. Methods Appl. Mech. Eng. 49 (1985) 357–364.

[44] A. Mizukami, T.J.R. Hughes, A Petrov–Galerkin finite element method for convection–dominated flows: An accurate upwinding technique for satisfying the maximum principle, Comput. Methods Appl. Mech. Eng. 50 (1985) 181–193.

[45] K. Niijima, Pointwise error estimates for a streamline diffusion finite element scheme, Numer. Math. 56 (1990) 707–719.

[46] E. Oñate, Derivation of stabilized equations for numerical solution of advective–diffusive transport and fluid flow problems, Comput. Methods Appl. Mech. Eng. 151 (1998) 233–265.

[47] A. Ramage, A note on parameter choice and iterative convergence for stabilised discretisations of advection–diffusion problems in three dimensions, Mathematics Research Report 32/98, University of Strathclyde, July 1998.

[48] A. Ramage, A multigrid preconditioner for stabilised discretisations of advection–diffusion problems, J. Comput. Appl. Math. 110 (1999) 187–203.

[49] J.G. Rice, R.J. Schnipke, A monotone streamline upwind finite element method for convection–dominated flows, Comput. Methods Appl. Mech. Eng. 48 (1985) 313–327.

[50] H.–G. Roos, M. Stynes, L. Tobiska, Numerical Methods for Singularly Perturbed Differential Equations. Convection–Diffusion and Flow problems, Springer–Verlag, Berlin, 1996.

[51] P.A.B. de Sampaio, A.L.G.A. Coutinho, A natural derivation of discontinuity capturing operator for convection–diffusion problems, Comput. Methods Appl. Mech. Eng. 190 (2001) 6291–6308.

[52] Y.–T. Shih, H.C. Elman, Modified streamline diffusion schemes for convection–diffusion problems, Comput. Methods Appl. Mech. Eng. 174 (1999) 137–151.

[53] M. Stynes, L. Tobiska, Necessary L^2–uniform convergence conditions for difference schemes for two–dimensional convection–diffusion problems, Comput. Math. Appl. 29 (1995) 45–53.

[54] M. Tabata, A finite element approximation corresponding to the upwind finite differencing, Mem. Numer. Math. 4 (1977) 47–63.

[55] T.E. Tezduyar, Y. Osawa, Finite element stabilization parameters computed from element matrices and vectors, Comput. Methods Appl. Mech. Eng. 190 (2000) 411–430.

[56] T.E. Tezduyar, Y.J. Park, Discontinuity–capturing finite element formulations for nonlinear convection–diffusion–reaction equations, Comput. Methods Appl. Mech. Eng. 59 (1986) 307–325.

[57] G. Zhou, How accurate is the streamline diffusion finite element method?, Math. Comput. 66 (1997) 31–44.

[58] G. Zhou, R. Rannacher, Pointwise superconvergence of the streamline diffusion finite–element method, Numer. Methods Partial Differ. Equations 12 (1996) 123–145.