Available online at www.sciencedirect.com
ScienceDirect
Comput. Methods Appl. Mech. Engrg. 197 (2008) 1997–2014
Computer methods in applied mechanics and engineering
www.elsevier.com/locate/cma

On spurious oscillations at layers diminishing (SOLD) methods for convection–diffusion equations: Part II – Analysis for P_1 and Q_1 finite elements

Volker John [a], Petr Knobloch [b,*]

[a] Universität des Saarlandes, Fachbereich 6.1 – Mathematik, Postfach 15 11 50, 66041 Saarbrücken, Germany
[b] Charles University, Faculty of Mathematics and Physics, Department of Numerical Mathematics, Sokolovská 83, 186 75 Praha 8, Czech Republic

Received 27 September 2007; received in revised form 20 December 2007; accepted 27 December 2007
Available online 15 January 2008

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. A review and a comparison of the most methods which have been proposed to remove or, at least, to diminish these oscillations without leading to excessive smearing of the layers are given in Part I, [V. John, P. Knobloch, On spurious oscillations at layers diminishing (SOLD) methods for convection–diffusion equations: Part I – A review, Comput. Methods Appl. Mech. Engrg. 196 (2007) 2197–2215]. In the present paper, the most promising of these SOLD methods are investigated in more detail for P_1 and Q_1 finite elements. In particular, the dependence of the results on the mesh, the data of the problems and parameters of the methods are studied analytically and numerically. Furthermore, the numerical solution of the nonlinear discrete problems is discussed and the capability of adaptively refined grids for reducing spurious oscillations is examined. Our conclusion is that, also for simple problems, any of the SOLD methods generally provides solutions with non-negligible spurious oscillations.

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

1. Introduction

This paper is a continuation of [26], in the following cited as Part I, which was devoted to a review and a comparison of finite element techniques developed to diminish spurious oscillations in discrete solutions of convection-dominated problems. Like in Part I, we consider the steady scalar convection–diffusion equation

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

We assume that Ω is a bounded domain in $\mathbb{R}^2$ with a polygonal boundary $\partial\Omega$, $\varepsilon > 0$ is the constant diffusivity, $\boldsymbol{b} \in W^{1,\infty}(\Omega)^2$ is a given convective field, $f \in L^2(\Omega)$ is an outer source of u, 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.

A popular finite element discretization technique for (1) is the streamline upwind/Petrov–Galerkin (SUPG) method which is frequently used because of its stability properties and higher-order accuracy. Since, in the convection-dominated regime, the SUPG solutions typically contain oscillations in layer regions, various stabilizing terms have been proposed to be added to the SUPG discretization in order to obtain discrete solutions in which the local oscillations are suppressed. In Part I, we called such techniques *spurious oscillations at layers diminishing (SOLD)* methods.

Part I presented a review of most SOLD methods published in the literature, discussed their derivation, proposed some alternative choices of parameters in the methods and categorized them. Some numerical studies gave a

* Corresponding author. Tel.: +420 737314937; fax: +420 224811036.
E-mail addresses: john@math.uni-sb.de (V. John), knobloch@karlin.mff.cuni.cz (P. Knobloch).

0045-7825/$ - see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.cma.2007.12.019

first impression of the behavior of the SOLD methods. These numerical tests were performed in a two-dimensional domain using the conforming P_1 finite element and it was observed that there are large differences between the SOLD methods. In some cases, the SOLD methods were able to significantly improve the SUPG solution and to provide a discrete solution with negligible spurious oscillations and without an excessive smearing of layers. However, it was not possible to identify a method which could be preferred in all the test cases. There are some methods which never produced good results since they either do not suppress the oscillations sufficiently or they are very diffusive and smear the layers considerably.

The aim of the present paper is to perform deeper investigations of those SOLD methods which gave acceptable results in Part I. We shall formulate the SOLD methods in the two-dimensional case and for conforming linear and bilinear finite elements. Formulations valid also in the three-dimensional case and for more general finite element spaces can be found in Part I. We do not consider the Mizukami–Hughes method [35,33] investigated in Part I since its applicability is rather limited. We shall investigate how strongly the methods depend on the computational mesh and the data of the problem. For methods containing parameters, we shall seek their optimal values and study the dependence of the results on the parameters. Since most of the SOLD methods are nonlinear, we shall also address algorithms for computing the discrete solution. Finally, the question will be studied whether adaptively refined grids help to suppress the spurious oscillations in SUPG solutions.

Our investigations will be performed on academic test examples whose solutions possess characteristic features of solutions of convection–diffusion equations. These academic problems allow to study the SOLD methods analytically, at least in the limit $\varepsilon \to 0+$. The analysis enables us to identify clearly those methods which can be expected to suppress the spurious oscillations and to study the dependence of the results on parameters in some of the methods.

The analysis presented in this paper will include the consideration of moderately anisotropic grids. Using such grids might not be reasonable for the considered examples since these grids are not adapted to the layers of the solution. Our motivation for looking at moderately anisotropic grids comes from applications. First, the meshing of complicated domains leads easily to anisotropic elements with moderate aspect ratio. Second, convection–diffusion equations are often just a part of a coupled system of equations, like in the $k-\varepsilon$ turbulence model [36] or in the simulation of precipitation processes [29]. For such problems, an adaptation of the grid is performed rather with respect to other equations in the system, for instance with respect to the Navier–Stokes equations in the mentioned examples. Thus, one has to face the situation that the grids might be not particularly well adapted with respect to the convection–diffusion equation but the SOLD methods still should provide satisfactory results.

The paper is organized in the following way. In the next section, we formulate the usual Galerkin discretization of (1) and introduce the SUPG method. In Section 3, the SOLD methods investigated in this paper are briefly reviewed. Then, in Section 4, we shall investigate the properties of the SOLD methods for three model problems. Section 5 is devoted to the computation of the discrete solution and, in Section 6, the usefulness of adaptively refined grids for the suppression of spurious oscillations is studied. Finally, Section 7 presents our conclusions.

Throughout the paper, we use the standard notations P_1, Q_1, $L^2(\Omega)$, $H^1(\Omega) = W^{1,2}(\Omega)$, etc. for the usual function spaces, see, e.g., Ciarlet [9]. The inner product in the space $L^2(\Omega)$ or $L^2(\Omega)^2$ will be denoted by $(\cdot,\cdot)$. For a vector $\boldsymbol{a} \in \mathbb{R}^2$, the symbol $|\boldsymbol{a}|$ stands for its Euclidean norm.

2. The Galerkin method and the SUPG method

To define a finite element discretization of (1), we introduce a triangulation $\mathscr{T}_h$ of the domain Ω consisting of a finite number of open elements K. We shall assume that all elements of $\mathscr{T}_h$ are either triangles or convex quadrilaterals. The discretization parameter h in the notation $\mathscr{T}_h$ is a positive real number satisfying $\operatorname{diam}(K) \leqslant h$ for any $K \in \mathscr{T}_h$. We assume that $\overline{\Omega} = \bigcup_{K\in\mathscr{T}_h} \overline{K}$ and that the closures of any two different elements of $\mathscr{T}_h$ are either disjoint or possess either a common vertex or a common edge.

We introduce the finite element space

$$V_h = \{v \in H_0^1(\Omega); v|_K \in R(K) \quad \forall K \in \mathscr{T}_h\},$$

where $R(K) = P_1(K)$ if K is a triangle and $R(K) = Q_1(K)$ if K is a rectangle. If K is a general convex quadrilateral, then $R(K)$ is defined by transforming the space $Q_1((0,1)^2)$ onto K by means of a bilinear one-to-one mapping, see, e.g., Ciarlet [9]. Finally, let $u_{bh} \in H^1(\Omega)$ be a function whose trace approximates the boundary condition u_b. Then the usual Galerkin finite element discretization of the convection–diffusion equation (1) reads:

Find $u_h \in H^1(\Omega)$ such that $u_h - u_{bh} \in V_h$ and

$$a(u_h, v_h) = (f, v_h) \quad \forall v_h \in V_h,$$

where

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

It is well known that this discretization is inappropriate if convection dominates diffusion since then the discrete solution is usually globally polluted by spurious oscillations. An improvement can be achieved by adding a stabilization term to the Galerkin discretization. One of the most efficient procedures of this type is the streamline upwind/Petrov–Galerkin (SUPG) method developed by Brooks and Hughes [3]. To formulate this method, we define the residual

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

where Δ_h is the Laplace operator defined elementwise, i.e., $(\Delta_h v)|_K = \Delta(v|_K)$ for any $K \in \mathscr{T}_h$ and any piecewise smooth function v. Then the SUPG method reads:

Find $u_h \in H^1(\Omega)$ such that $u_h - u_{bh} \in V_h$ and

$$a(u_h, v_h) + (R_h(u_h), \tau \boldsymbol{b} \cdot \nabla v_h) = (f, v_h) \quad \forall v_h \in V_h, \tag{2}$$

where $\tau \in L^\infty(\Omega)$ is a nonnegative stabilization parameter. The choice of τ may dramatically influence the accuracy of the discrete solution and therefore it has been a subject of an extensive research over the last three decades, see, e.g., the review in Part I. Unfortunately, a general optimal definition of τ is still not known. In our computations, we define τ, on any element $K \in \mathscr{T}_h$, by the formula

$$\tau|_K = \frac{h_K}{2|\boldsymbol{b}|}\left(\coth Pe_K - \frac{1}{Pe_K}\right) \quad \text{with} \quad Pe_K = \frac{|\boldsymbol{b}|h_K}{2\varepsilon}, \tag{3}$$

where h_K is the element diameter in the direction of the convection vector $\boldsymbol{b}$. We refer to Part I for various justifications of this formula and for a precise definition of h_K. If convection strongly dominates diffusion in Ω and hence the local Péclet numbers Pe_K are very large, the parameter τ is basically given by

$$\tau|_K = \frac{h_K}{2|\boldsymbol{b}|} \quad \forall K \in \mathscr{T}_h. \tag{4}$$

Note that, generally, the parameters h_K, Pe_K and $\tau|_K$ are functions of the points $\mathbf{x} \in K$.

An alternative to the SUPG method is the Galerkin/least-squares method introduced by Hughes et al. [21] or its modification proposed by Franca et al. [16]. A similar stabilization can also be obtained using the subgrid scale method of Hughes [20]. In addition, for transient problems, stabilization terms of the discussed type also result by applying the characteristic Galerkin method of Douglas and Russell [15] or the Taylor–Galerkin method of Donéa [14]. See also Codina [11] for a comparison of these methods. However, all these methods are identical to the SUPG method (up to the choice of the stabilization parameter) if problem (1) has constant coefficients and is discretized using linear triangular or bilinear rectangular finite elements. Since this will be the case in all the model problems discussed in this paper, we confine ourselves to the SUPG method in the following.

3. Spurious oscillations at layers diminishing methods

Because the SUPG method is not monotone, a discrete solution satisfying (2) usually still contains spurious oscillations. Although these oscillations are localized in narrow regions along sharp layers, they are often not negligible and they are not permissible in many applications. A possible remedy is to add a suitable artificial diffusion term to the SUPG method. In Part I, methods of this type are called *spurious oscillations at layers diminishing* (*SOLD*) methods. Here, we describe these methods only very briefly and refer to the review in Part I for details. To make similarities and differences between the methods better visible, we shall formulate the methods in a slightly different way than in Part I.

There are three basic classes of SOLD methods: methods adding isotropic artificial diffusion, methods adding crosswind artificial diffusion, and methods where the additional artificial diffusion stems from an edge stabilization. The amount of the artificial diffusion in these methods typically depends on the unknown discrete solution u_h. Thus, the resulting methods are nonlinear (although the original problem (1) is linear).

The methods of the first class add the isotropic artificial diffusion term

$$(\tilde{\varepsilon}\nabla u_h, \nabla v_h) \tag{5}$$

to the left-hand side of the SUPG discretization (2). The parameter $\tilde{\varepsilon}$ is nonnegative and usually depends on u_h. For the first time, a SOLD term which can be written in the form (5) was introduced by Hughes et al. [22]. Further approaches were proposed by Tezduyar and Park [38] and Galeão and do Carmo [17]. According to the criteria and tests in Part I (and according to further numerical experiments we have performed in [24,25]), one of the best choices of $\tilde{\varepsilon}$ in (5) is to set

$$\tilde{\varepsilon} = \max\left\{0, \frac{\tau|\boldsymbol{b}||R_h(u_h)|}{|\nabla u_h|} - \tau\frac{|R_h(u_h)|^2}{|\nabla u_h|^2}\right\}, \tag{6}$$

as proposed by do Carmo and Galeão [8], abbreviated with dCG91 in Part I. Here and in the following, we always assume that $\tilde{\varepsilon} = 0$ if the denominator of a formula defining $\tilde{\varepsilon}$ vanishes. Almeida and Silva [1] suggested to multiply the negative term in (6) by

$$\zeta_h = \max\left\{1, \frac{\boldsymbol{b} \cdot \nabla u_h}{R_h(u_h)}\right\},$$

which is method AS97 in Part I. However, in our tests, we often observed no significant differences to the results obtained with (6). Another $\tilde{\varepsilon}$, motivated by assumptions needed for theoretical investigations, can be found in Knopp et al. [34]. Further modifications of the above approaches were proposed by do Carmo and Galeão [8] and do Carmo and Alvarez [7], who introduced rather complicated definitions of $\tilde{\varepsilon}$ which should suppress the addition of the artificial diffusion in regions where the solution of (1) is smooth. The SOLD term (5) was also used by Johnson [30], who proposed to set

$$\tilde{\varepsilon}|_K = \max\{0, C[\mathrm{diam}(K)]^2|R_h(u_h)| - \varepsilon\} \quad \forall K \in \mathscr{T}_h, \tag{7}$$

where C is a nonnegative parameter (method J90 in Part I).

Johnson et al. [32] modified the SUPG discretization (2) by adding artificial diffusion in the crosswind direction only. This corresponds to the additional term

$$(\tilde{\varepsilon}\boldsymbol{b}^\perp \cdot \nabla u_h, \boldsymbol{b}^\perp \cdot \nabla v_h) \quad \text{with} \quad \boldsymbol{b}^\perp = \frac{(-b_2, b_1)}{|\boldsymbol{b}|} \tag{8}$$

on the left-hand side of (2). In [32], the parameter $\tilde{\varepsilon}$ was defined by

$$\tilde{\varepsilon}|_K = \max\{0, |\boldsymbol{b}|h_K^{3/2} - \varepsilon\} \quad \forall K \in \mathscr{T}_h \tag{9}$$

so that the resulting method (JSW87 in Part I) is linear but non-consistent and hence it is restricted to finite elements of first order of accuracy. Moreover, the numerical tests from Part I show that this method is very diffusive.

Codina [10] proposed to define $\tilde{\varepsilon}$ in (8), for any $K \in \mathscr{T}_h$, by

$$\tilde{\varepsilon}|_K = \max\left\{0, C\frac{\operatorname{diam}(K)|R_h(u_h)|}{2|\nabla u_h|} - \varepsilon\frac{|R_h(u_h)|}{|\boldsymbol{b}\cdot\nabla u_h|}\right\}, \tag{10}$$

where C is a suitable constant, and he recommended to set $C \approx 0.7$ for (bi)linear finite elements. This is method C93 in Part I. For $f \neq 0$, we observed that, in some cases, this choice of $\tilde{\varepsilon}$ does not lead to a reduction of the oscillations (see the discussion to Example 1 in the next section). Therefore, in Part I, we replaced (10) by

$$\tilde{\varepsilon}|_K = \max\left\{0, C\frac{\operatorname{diam}(K)|R_h(u_h)|}{2|\nabla u_h|} - \varepsilon\right\}, \tag{11}$$

called method KLR02_3 in Part I. Here, we shall also call this method *modified method of Codina*. If $f = 0$ and $\Delta_h u_h = 0$, it is equivalent to the original method (10). A modification of (10), leading to properties convenient for theoretical investigations, was proposed by Knopp et al. [34].

For triangulations consisting of weakly acute triangles, Burman and Ern [4] proposed to use (8) with $\tilde{\varepsilon}$ defined, on any $K \in \mathscr{T}_h$, by

$$\tilde{\varepsilon}|_K = \frac{\tau|\boldsymbol{b}||R_h(u_h)|}{|\nabla u_h|}\frac{|\boldsymbol{b}||\nabla u_h|}{|\boldsymbol{b}||\nabla u_h| + |R_h(u_h)|} \times \frac{|\boldsymbol{b}||\nabla u_h| + |R_h(u_h)| + \tan\alpha_K|\boldsymbol{b}||\boldsymbol{b}^\perp\cdot\nabla u_h|}{|R_h(u_h)| + \tan\alpha_K|\boldsymbol{b}||\boldsymbol{b}^\perp\cdot\nabla u_h|}. \tag{12}$$

The parameter α_K is equal to $\pi/2 - \beta_K$ where β_K is the largest angle of K. If $\beta_K = \pi/2$, it is recommended in [4] to set $\alpha_K = \pi/6$. To improve the convergence of the nonlinear iterations, we replaced in Part I $|R_h(u_h)|$ by $|R_h(u_h)|_{\text{reg}}$ with $|x|_{\text{reg}} \equiv x\tanh(x/2)$ as proposed already in [4]. The resulting method was called BE02_1.

In Part I, we also introduced a simplification of (12), called BE02_2, defined by

$$\tilde{\varepsilon} = \frac{\tau|\boldsymbol{b}||R_h(u_h)|}{|\nabla u_h|}\frac{|\boldsymbol{b}||\nabla u_h|}{|\boldsymbol{b}||\nabla u_h| + |R_h(u_h)|}, \tag{13}$$

which adds less artificial diffusion than (12). In (13), we do not apply any regularization of the absolute values. We call this method *modified method of Burman and Ern*. Based on the evaluation of the numerical studies in Part I and [24,25], in our opinion, this method and the modified method of Codina are the best methods among the methods adding crosswind artificial diffusion.

It is also possible to add both isotropic and crosswind artificial diffusion terms to the left-hand side of (2). Denoting the parameters in (5) and (8) by $\tilde{\varepsilon}^{\text{iso}}$ and $\tilde{\varepsilon}^{\text{cross}}$, respectively, Codina and Soto [12] proposed to set

$$\tilde{\varepsilon}^{\text{iso}} = \max\{0, \tilde{\varepsilon}^{\text{dc}} - \tau|\boldsymbol{b}|^2\}, \quad \tilde{\varepsilon}^{\text{cross}} = \tilde{\varepsilon}^{\text{dc}} - \tilde{\varepsilon}^{\text{iso}},$$

where $\tilde{\varepsilon}^{\text{dc}}$ is defined by a formula similar to (11). However, in the numerical tests we have performed up to now, we have not observed an advantage in using this approach instead of (8) with $\tilde{\varepsilon}$ given by (11).

There are some similarities between the definitions of $\tilde{\varepsilon}$ in (6), (7) and (10)–(13). Particularly, the presence of a term of the type $h|R_h(u_h)|/|\nabla u_h|$ seems to be important. Indeed, if convection is strongly dominant (and hence (4) approximately holds), we have in (6), (12) and (13)

$$\frac{\tau|\boldsymbol{b}||R_h(u_h)|}{|\nabla u_h|} \approx \frac{h_K|R_h(u_h)|}{2|\nabla u_h|}. \tag{14}$$

Remark 1. The recently published $YZ\beta$ scheme for scalar convection–diffusion equations [2], originally proposed by Tezduyar [37] for compressible flows, gives for $\beta = 1$ exactly the parameter (14) if, in contrast to [2], in the definition of the local element length the convection is used instead of the gradient of the solution. Using the latter replaces h_K by the element size orthogonal to the convection, see the discussion of this choice in Section 4.

The third class of SOLD methods is based on so-called edge stabilizations, which add the term

$$\sum_{K\in\mathscr{T}_h}\int_{\partial K}\Psi_K(u_h)\operatorname{sign}\left(\frac{\partial u_h}{\partial \boldsymbol{t}_{\partial K}}\right)\frac{\partial v_h}{\partial \boldsymbol{t}_{\partial K}}\,\mathrm{d}\sigma \tag{15}$$

to the left-hand side of (2), $\boldsymbol{t}_{\partial K}$ being a tangent vector to the boundary ∂K of K. Various choices of the nonnegative function Ψ_K were proposed by Burman and Hansbo [6] and Burman and Ern [5]. To make the convergence of the nonlinear iterative process possible, the sign operator is regularized by replacing it by the hyperbolic tangent as recommended in [6]. Our numerical tests in Part I and in [27] indicate that some SOLD methods based on edge stabilizations work comparatively well on unstructured grids with acute triangles, but still away from being perfect. In general, these methods lead to a more pronounced smearing of layers in comparison with the best methods of the previous two classes. The best edge stabilization method in the numerical studies of Part I is defined by $\Psi_K(u_h) = \gamma|(R_h(u_h)|_K)|$, where γ is a nonnegative parameter. This method was called BE05_2 in Part I. We shall see in the next section that the parameter γ should be proportional to the area $|K|$ of the respective element K, i.e., $\gamma|_K = C|K|$ with some $C \geqslant 0$. Then (15) can be written in the form

$$\sum_{K\in\mathscr{T}_h}|K|\int_{\partial K} C\frac{|R_h(u_h)|_K|}{\left|\frac{\partial u_h}{\partial \boldsymbol{t}_{\partial K}}\right|}\frac{\partial u_h}{\partial \boldsymbol{t}_{\partial K}}\frac{\partial v_h}{\partial \boldsymbol{t}_{\partial K}}\,\mathrm{d}\sigma, \tag{16}$$

which has a similar structure like many of the SOLD terms discussed above.

4. Properties of SOLD methods for model problems

In this section, we shall investigate the properties of the SOLD methods described in the previous section by applying them to three model problems whose solutions possess characteristic features of solutions of (1), in particular, parabolic and exponential boundary layers and interior layers. The goal of these investigations consists in understanding why the methods work well or not. All numerical results have been double-checked by computing them with two different codes, one of them was *MooNMD*, [28].

In all model problems, we shall consider (1) with

$$\Omega = (0,1)^2 \quad \text{and} \quad \varepsilon = 10^{-8}. \tag{17}$$

Moreover, we shall confine ourselves to the two types of triangulations depicted in Fig. 1. To characterize these triangulations, we shall use the notion '$N_1 \times N_2$ mesh' where N_1 and N_2 are the numbers of vertices in the horizontal and vertical directions, respectively. The corresponding mesh widths will be denoted by h_1 and h_2, i.e., $h_1 = 1/(N_1 - 1)$ and $h_2 = 1/(N_2 - 1)$.

Example 1 (*Solution with parabolic and exponential boundary layers*). We consider the convection–diffusion equation (1) with (17) and

$$\boldsymbol{b} = (1,0)^{\mathrm{T}}, \quad f = 1, \quad u_b = 0.$$

The solution $u(x,y)$ of this problem, see Fig. 2a, possesses an exponential boundary layer at $x = 1$ and parabolic (characteristic) boundary layers at $y = 0$ and $y = 1$. Outside the layers, the solution $u(x,y)$ is very close to x.

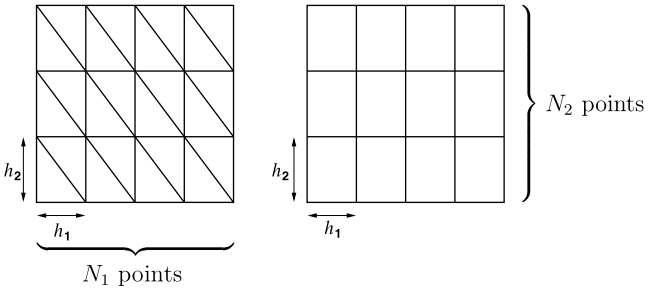

Fig. 1. Triangulations used in Section 4.

This test problem was used, e.g., by Mizukami and Hughes [35].

For this special example, the stabilization parameter τ given in (3) is optimal along lines $y = \text{const.}$ outside the parabolic layers. Therefore, for both the P_1 and Q_1 finite elements, the SUPG method gives a nodally exact solution outside the parabolic layers. However, there are strong oscillations at the parabolic layers, see Fig. 2b, which shows a SUPG solution for the Q_1 finite element. For the P_1 finite element, the solution is similar. To measure the quality of a discrete solution u_h at the parabolic layers, we define the values

$$osc := \max_{y\in[0,1]} \{u_h(0.5,y) - u_h(0.5,0.5)\}, \tag{18}$$

$$smear := \max_{y\in[h_2,1-h_2]} \{u_h(0.5,0.5) - u_h(0.5,y)\}, \tag{19}$$

see also Part I. The first value measures the oscillations at the parabolic layers. In the case that the oscillations are suppressed to the most part, the second value measures the smearing of these layers.

To investigate the optimality of the definitions of $\tilde{\varepsilon}$ presented in the previous section, we introduce a parameter η such that, for any $K \in \mathscr{T}_h$,

$$\tilde{\varepsilon}|_K = \eta \frac{\operatorname{diam}(K)|R_h(u_h)|}{2|\nabla u_h|} \quad \text{if } \nabla u_h \neq \mathbf{0}. \tag{20}$$

This ansatz is based on the similarities between the SOLD methods discussed at the end of Section 3. The relation (20) can be satisfied provided that $\tilde{\varepsilon} = 0$ if $R_h(u_h) = 0$, which is true in all the cases except for (9). Of course, η generally depends on u_h, $\mathscr{T}_h$ and the data of (1). Nevertheless, we can also consider $\tilde{\varepsilon}$ defined by (20) with a constant value of η, which resembles the first term of (10) and (11). Fig. 3 shows how the value of η influences the oscillations and smearing along the line $x = 0.5$ in a discrete solution of Example 1 defined using the crosswind artificial diffusion term (8). We observe that there is a clear optimal value of η which, however, depends on the used triangulation. We also see that the optimal values of η are nearly the same for both the P_1 and Q_1 finite elements. Using (20) together with the isotropic artificial diffusion term (5), the curves and the optimal values of η are very similar to those in Fig. 3.

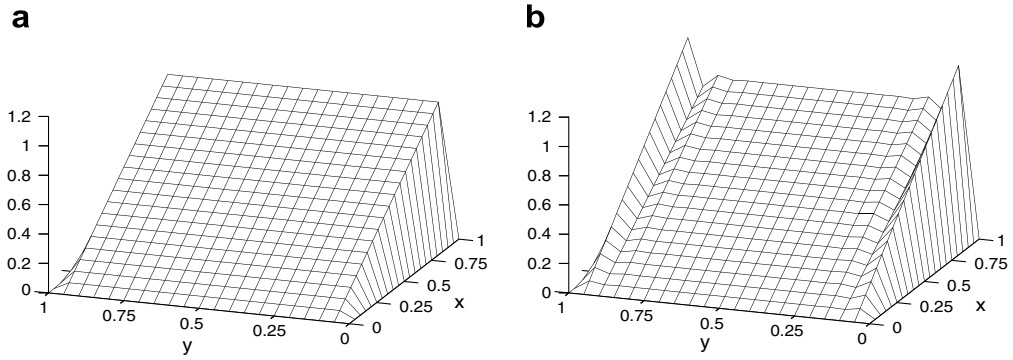

Fig. 2. Example 1: (a) solution u and (b) discrete solution u_h obtained using the SUPG method with the Q_1 finite element on a 21×21 mesh.

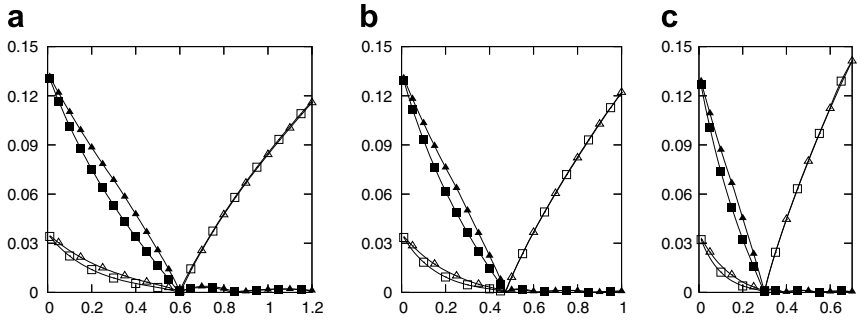

Fig. 3. Example 1, discretization with a crosswind SOLD term given by (8) and (20), dependence of the measures for oscillations (■ P_1, ▲ Q_1) and smearing (□ P_1, △ Q_1) defined by (18) and (19), respectively, on the parameter η: (a) 65×33 mesh; (b) 65×65 mesh; (c) 33×65 mesh.

The optimal values of η from Fig. 3 correspond to discrete solutions which are nodally exact along the line $x = 0.5$. We would like to derive now an analytic expression for the optimal value of η by requiring that the discrete solution be nodally exact outside the exponential boundary layer. For simplicity, we shall consider the case $\varepsilon \to 0+$ so that the nodally exact discrete solution satisfies $u_h(x,y) = x$ for $(x,y) \in [0, 1-h_1] \times [h_2, 1-h_2]$, where h_1 and h_2 are defined in Fig. 1. By the definition of the SOLD methods, we have, for any $v_h \in V_h$,

$$(R_h(u_h), v_h + \tau \boldsymbol{b} \cdot \nabla v_h) + (\tilde{\varepsilon} \nabla u_h, \nabla v_h) = 0 \tag{21}$$

or

$$(R_h(u_h), v_h + \tau \boldsymbol{b} \cdot \nabla v_h) + (\tilde{\varepsilon} \boldsymbol{b}^\perp \cdot \nabla u_h, \boldsymbol{b}^\perp \cdot \nabla v_h) = 0. \tag{22}$$

In what follows, we shall assume that $\operatorname{supp} v_h \subset [0, 1-h_1] \times [0,1]$. Then it is easy to verify that, for both the P_1 and Q_1 finite elements, the nodally exact discrete solution satisfies $(R_h(u_h), \tau \boldsymbol{b} \cdot \nabla v_h) = 0$ provided that τ is independent of x (for the P_1 finite element, this is true even for any $\tau \in L^\infty(\Omega)$ and it follows from the fact that, for any $K \in \mathscr{T}_h$, either $R_h(u_h)|_K = 0$ or $\boldsymbol{b} \cdot \nabla v_h|_K = 0$ – see below). Therefore, the optimal value of η is independent of the choice of τ. It also shows that the SUPG method alone is not able to provide an oscillation-free solution.

Let us consider the P_1 finite element. Then for elements K lying in $[0, 1-h_1] \times [h_2, 1-h_2]$ or having exactly one vertex at the boundary $y = 0$ or $y = 1$, we have $\boldsymbol{b} \cdot \nabla u_h|_K = 1$ and hence $R_h(u_h)|_K = 0$. Thus, the only elements K in $[0, 1-h_1] \times [0,1]$ which may lead to non-vanishing parameters $\tilde{\varepsilon}|_K$ are elements with two vertices at $y = 0$ or $y = 1$. If K is such an element, we may assume that the vertex of K not lying on $y = 0$ or $y = 1$ has the coordinates (ih_1, h_2) or $(ih_1, 1-h_2)$ with $i \in \{1, \ldots, N_1 - 3\}$ since the two elements which have all three vertices on the boundary of $[0, 1-h_1] \times [0,1]$ do not have to be considered. Then $\nabla u_h|_K = (0, \pm i h_1/h_2)$ and, consequently, for any η, we get $(\tilde{\varepsilon} \nabla u_h, \nabla v_h) = (\tilde{\varepsilon} \boldsymbol{b}^\perp \cdot \nabla u_h, \boldsymbol{b}^\perp \cdot \nabla v_h)$ so that we do not have to distinguish between (21) and (22). If v_h equals 1 at the interior vertex of K and vanishes at all other vertices of the triangulation, the conditions (21) and (22) reduce to

$$(R_h(u_h), v_h)_K + (\tilde{\varepsilon} \nabla u_h, \nabla v_h)_K = 0,$$

where $(\cdot,\cdot)_K$ denotes the inner product in $L^2(K)$ or $L^2(K)^2$. Since $(\nabla u_h \cdot \nabla v_h)|_K = i h_1/h_2^2$ and $R_h(u_h)|_K = -f = -1$, we deduce that the optimal value of $\tilde{\varepsilon}$ is

$$\tilde{\varepsilon}_{\mathrm{opt}}|_K = \frac{h_2^2}{3 i h_1}$$

and that the optimal value of η is

$$\eta_{\mathrm{opt}} = \frac{2}{3\sqrt{1 + \left(\dfrac{h_1}{h_2}\right)^2}}. \tag{23}$$

This formula is in a very good agreement with the optimal values of η observed in Fig. 3. Note also that η_{opt} does not depend on K and it depends on the used triangulation only through the aspect ratio of the elements of the triangulation defined by

$$v := \frac{h_1}{h_2}. \tag{24}$$

The graphs in Fig. 3 indicate that a SOLD term of the form (5) or (8) can be expected to lead to an oscillation-free solution only if, on any element $K \subset [0, 1-h_1] \times [0,1]$ with two vertices at $y = 0$ or $y = 1$, the value of $\tilde{\varepsilon}$ corresponding to the nodally exact discrete solution u_h is at least $\tilde{\varepsilon}_{\mathrm{opt}}$. Inserting u_h into the formulas (6), (7) and (10)–(13) from Section 3, we obtain the following relations between $\tilde{\varepsilon}$ and $\tilde{\varepsilon}_{\mathrm{opt}}$ (we drop the notation for restriction to K):

$$(6): \tilde{\varepsilon} = \frac{3}{2}\left(v - \frac{1}{i}\right)\tilde{\varepsilon}_{\mathrm{opt}},$$

$$(7): \tilde{\varepsilon} = 3iv\left(Ch_2(1+v^2) - \frac{\varepsilon}{h_2}\right)\tilde{\varepsilon}_{\mathrm{opt}},$$

$$(9): \tilde{\varepsilon} = 3iv^2\left(\sqrt{h_1} - \frac{\varepsilon}{h_1}\right)\tilde{\varepsilon}_{\mathrm{opt}},$$

$$(10): \tilde{\varepsilon} = 0 \quad \text{since } \boldsymbol{b} \cdot \nabla u_h = 0,$$

$$(11): \tilde{\varepsilon} = \left(C\frac{3}{2}\sqrt{1+v^2} - \frac{3iv\varepsilon}{h_2}\right)\tilde{\varepsilon}_{\mathrm{opt}},$$

$$(12): \tilde{\varepsilon} = \frac{3iv^2}{2(1+iv)}\,\frac{\sqrt{3} + iv(1+\sqrt{3})}{\sqrt{3}+iv}\,\tilde{\varepsilon}_{\mathrm{opt}},$$

$$(13): \tilde{\varepsilon} = \frac{3iv^2}{2(1+iv)}\tilde{\varepsilon}_{\text{opt}}.$$

These relations have to be understood in the way that a right-hand side is replaced by zero if it is negative. As we see, $\tilde{\varepsilon}$ of the original method by Codina defined by (10) cannot be expected to lead to an oscillation-free discrete solution since, for the nodally exact discrete solution, we have $\tilde{\varepsilon} = 0$ on any element in $[0, 1-h_1] \times [0,1]$. On the other hand, using $C = \eta_{\text{opt}}$ in the modified method of Codina with $\tilde{\varepsilon}$ given by (11), we have $\tilde{\varepsilon} \approx \tilde{\varepsilon}_{\text{opt}}$ (provided that the ε-dependent term can be neglected) and hence we obtain nearly the nodally exact solution. The methods with $\tilde{\varepsilon}$ defined by (7) and (9) do not seem to be practical since the ratio $\tilde{\varepsilon}/\tilde{\varepsilon}_{\text{opt}}$ decreases when refining the mesh while keeping the aspect ratio fixed. The remaining three definitions of $\tilde{\varepsilon}$, i.e., (6), (12) and (13), enable to satisfy the condition $\tilde{\varepsilon} \geqslant \tilde{\varepsilon}_{\text{opt}}$ for sufficiently large aspect ratios, in particular, for $v \geqslant 5/3$, $v \geqslant 0.9$ and $v \geqslant (1+\sqrt{7})/3$, respectively.

In the quadrilateral case, it is not possible to derive simple formulas for $\tilde{\varepsilon}_{\text{opt}}$ and η_{opt}, but the results in Fig. 3 suggest that the optimal values of η do not differ much from (23). Therefore, conditions for obtaining an oscillation-free solution can be derived by requiring that the parameters $\tilde{\varepsilon}$ in (5) and (8) satisfy

$$\tilde{\varepsilon}|_K \geqslant \eta_{\text{opt}} \frac{\text{diam}(K)|R_h(u_h)|}{2|\nabla u_h|} = \frac{h_2}{3}\frac{|R_h(u_h)|}{|\nabla u_h|} \quad \forall K \in \mathcal{T}_h \tag{25}$$

for any function u_h. The resulting relations also apply to the P_1 finite element but are less sharp than above. It is obvious that, for the method of do Carmo and Galeão and for the modified method of Burman and Ern, i.e., for $\tilde{\varepsilon}$ given by (6) or (13), respectively, the inequality (25) may hold only if

$$\tau|\boldsymbol{b}| > \frac{h_2}{3}, \tag{26}$$

which is equivalent to $v > 2/3$. If $v \leqslant 2/3$, we have to expect spurious oscillations in the discrete solution as it is demonstrated in Fig. 4. The inequality (26) suggests to define τ in (6) and (13) using the element diameter $h_K^{\perp}$ in the direction orthogonal to the convection vector $\boldsymbol{b}$ instead of using h_K. For instance, in the convection-dominated case, we can use the formula

$$\tau|_K = \frac{h_K^{\perp}}{2|\boldsymbol{b}|} \quad \forall K \in \mathcal{T}_h, \tag{27}$$

which in fact removes the spurious oscillations visible in Fig. 4. For $\tilde{\varepsilon}$ given by (12), the necessary condition obtained from (25) is weaker than (26) but, for a 41×21 mesh, we get a similar discrete solution as in Fig. 4 (slightly better for the P_1 finite element and slightly worse for the Q_1 finite element). On the other hand, if we use $\tilde{\varepsilon}$ given by (11), spurious oscillations should not appear for $C > 2/3 > \eta_{\text{opt}}$, which is particularly satisfied by the value $C \approx 0.7$ recommended in [10]. However, for certain triangulations, the layers can be smeared as Fig. 3 indicates.

As we already showed, $\tilde{\varepsilon}$ defined by (10) is not appropriate in case of the P_1 finite element. The situation is different for the Q_1 finite element for which similar results can be obtained as with (11) provided that the term (8) is evaluated using a quadrature formula with nodes which are not 'too near' to the boundary of Ω.

Finally, let us mention a further drawback of $\tilde{\varepsilon}$ defined by (7). If the functions f and u_b in (1) are multiplied by a constant α, then the solution u changes to αu. For the SOLD methods defined using the terms (5) and (8), this property is valid if and only if the value of $\tilde{\varepsilon}$ does not change after replacing u_h, f by αu_h, αf, respectively. This is true for most of the definitions of $\tilde{\varepsilon}$ mentioned in Section 3, however not for the formula (7). Let us assume that, for a given mesh, the parameter C in (7) is defined in such a way that the corresponding discrete solution is a good approximation to the solution of Example 1. Now, replacing $f = 1$ by $f = \alpha$, we typically obtain with (7) either an oscillatory solution (if $|\alpha| < 1$) or a solution excessively smearing the layers (if $|\alpha| > 1$). This shows that the formula (7) cannot be expected to lead to a qualitatively correct discrete solution unless C depends on u_h or the data of problem (1). This was probably also recognized by Johnson [31] who proposed to set $C = \beta/\max_{\Omega}|u_h|$ in (7) where β is a constant. However, a constant value of β allows to remove spurious oscillations only at the price of a significant smearing of the layers and hence the method does

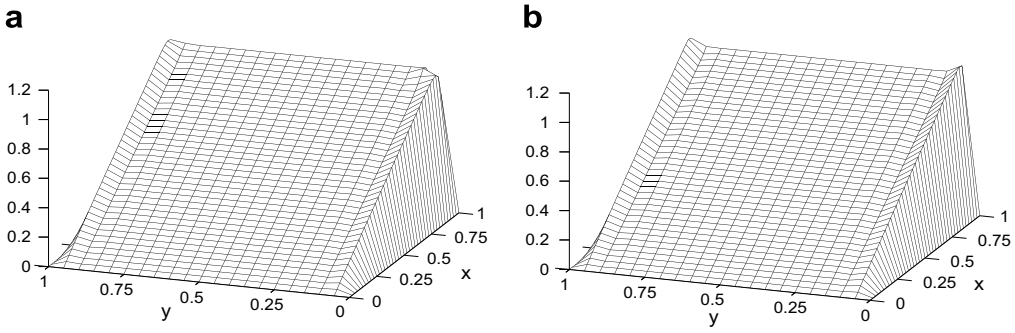

Fig. 4. Example 1, discrete solutions on 41×21 meshes: (a) P_1 finite element, isotropic artificial diffusion given by (6) and (b) Q_1 finite element, crosswind artificial diffusion given by (13).

not attain the quality of the best SOLD methods (see also Part I).

For the edge stabilization term (16) and both the P_1 and Q_1 finite elements, it is easy to derive that the function $u_h(x,y) = x$ satisfies the respective discrete problem for $\varepsilon \to 0+$ and test functions $v_h \in V_h$ with $\operatorname{supp} v_h \subset [0, 1-h_1] \times [0,1]$ if $C = 1/6$. However, in practice, the discrete solution is slightly worse at the parabolic boundary layers due to the regularization of the sign operator. Moreover, in contrast to the modified method of Codina, the discrete solution is significantly smeared along the exponential boundary layer. A sharp approximation of this layer requires to set $C = 0$ in this region.

To summarize the discussion to Example 1, among the SOLD methods adding the isotropic diffusion term (5) or the crosswind diffusion term (8), the only SOLD method which gives satisfactory results seems to be the modified method of Codina defined by (8) and (11), but only with an appropriately chosen constant C. The edge stabilization (16) enables to compute a satisfactory solution if the parameter C is layer-adapted.

Example 2 (*Solution with interior layer and exponential boundary layers*). We consider the convection–diffusion equation (1) with (17) and

$$\boldsymbol{b} = (\cos(-\pi/3), \sin(-\pi/3))^{\mathrm{T}}, \quad f = 0,$$
$$u_b(x,y) = \begin{cases} 0 & \text{for } x = 1 \text{ or } y \leqslant 0.7, \\ 1 & \text{else.} \end{cases}$$

The solution, see Fig. 5a, possesses an interior (characteristic) 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 was used, e.g., by Hughes et al. [22].

The position of spurious oscillations in the solutions obtained with the SUPG method depends on h_1 and h_2. If the mesh is constructed such that

$$h_1 b_2 + h_2 b_1 < 0, \tag{28}$$

then, for both the P_1 and Q_1 finite elements, the SUPG solution contains oscillations along the interior layer and along the boundary layer at $x = 1$. However, there are no oscillations along the boundary layer at $y = 0$ and this layer is not smeared. This is illustrated in Fig. 5b which shows a SUPG solution for the P_1 finite element. For the Q_1 finite element the discrete solution is very similar. If $h_1 b_2 + h_2 b_1 > 0$, then the SUPG solution contains oscillations along the interior layer and along the boundary layer at $y = 0$ but no oscillations and no smearing occur along the boundary layer at $x = 1$. For shortness of presentation, we shall consider only the case (28) in the following.

For a nodally exact solution, the SUPG term will not vanish in Example 2 (in contrast to Example 1). Thus, for obtaining a nodally exact solution with a SOLD method, the choice of the SUPG parameter τ will be of importance, too. The chosen parameter has to ensure that there is no smearing of layers since smeared layers cannot be corrected with SOLD methods. With the approach presented in Section 2, the SUPG parameter in Example 2 will be the same on each element. We found that the choice (3) is optimal in the class of globally constant parameters in the sense that any larger value leads to a smearing of the layer at $y = 0$ and any smaller value results in spurious oscillations at this layer and increases the oscillations at $x = 1$.

Let us first investigate the quality of the approximation of the interior layer. For simplicity, we shall confine ourselves to the P_1 finite element unless stated otherwise. To measure the oscillations of a discrete solution u_h at the interior layer, we define the value

$$osc_{\mathrm{int}} := \max \left\{ \max_{(x,y)\in G} u_h(x,y) - 1, \left| \min_{(x,y)\in G} u_h(x,y) \right| \right\}, \tag{29}$$

where (x,y) are the nodes in $G := [0, 0.5] \times [0.25, 1]$. Let us again consider SOLD methods defined using the term (5) or (8) with $\tilde{\varepsilon}$ given by (20). Numerical tests show that the value of osc_{int} is a non-increasing function of η on a given mesh. Given an integer m, we define

$$\eta_m := \min\{\eta \in \mathbb{R}_0^+; osc_{\mathrm{int}}(\eta) \leqslant 10^{-m}\}.$$

This value depends on the aspect ratio ν defined in (24). In view of (28), we have $\nu > \sqrt{3}/3$. Fig. 6 presents the dependence of η_2, η_3 and η_4 on the aspect ratio for both the isotropic and the crosswind artificial diffusion and for $h_1 = 1/64$. Of course, h_2 and consequently the number of degrees of freedom is different for different aspect ratios.

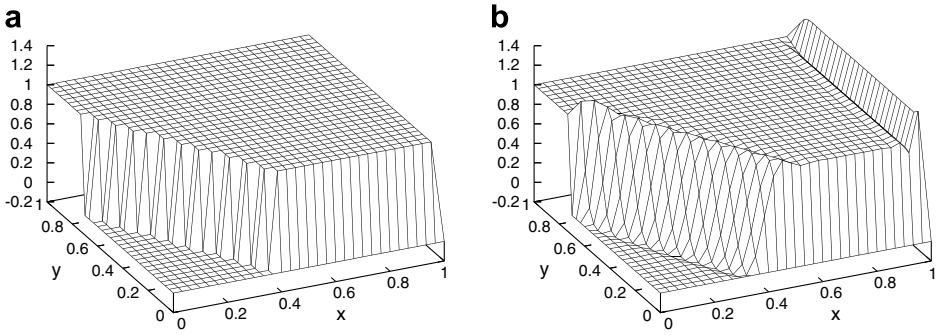

Fig. 5. Example 2: (a) solution u and (b) discrete solution u_h obtained using the SUPG method with the P_1 finite element on a 31×31 mesh.

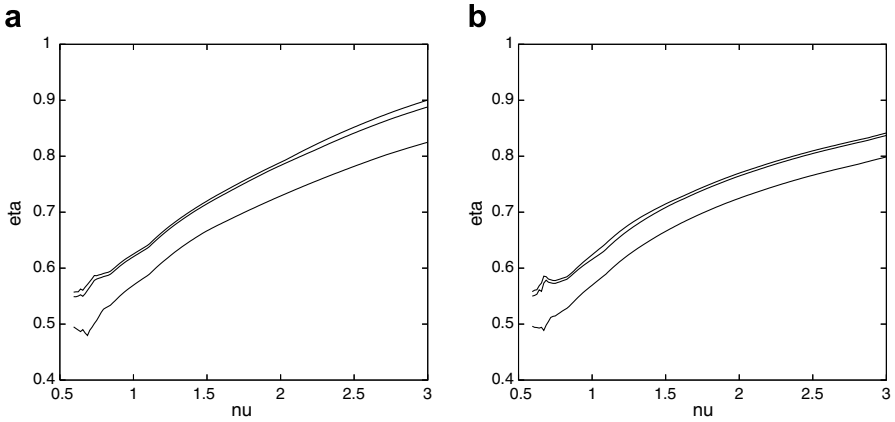

Fig. 6. Example 2, dependence of η_2, η_3 and η_4 on v (from bottom to top) for the P_1 finite element and meshes with $h_1 = 1/64$: (a) isotropic artificial diffusion (5) and (b) crosswind artificial diffusion (8).

We checked with several values for h_1 that the results presented in Fig. 6 depend only on v. Thus, one would get the same results for a fixed number of degrees of freedom with varying h_1 and h_2. Fig. 6 shows that the smallest value of η assuring that oscillations will not exceed a given tolerance increases with increasing aspect ratio. Qualitatively, the results for the Q_1 finite element are the same as for the P_1 finite element: increasing aspect ratios require increasing parameters η to suppress the oscillations below given thresholds.

For small ε, formula (20) for $\tilde{\varepsilon}$ is the main part of the method of Codina given by (8) and (10). Particularly, the results in Fig. 6 show that, in contrast to Example 1, the recommended value $C \approx 0.7$ does not generally lead to sufficiently small spurious oscillations.

Now let us turn our attention to the method of do Carmo and Galeão given by (5) and (6) and the modified method of Burman and Ern given by (8) and (13). Comparing the formulas (6) and (13) with (20), one finds, using (4), that for obtaining comparable results as for $\tilde{\varepsilon}$ defined by (20) with a given value of η, the condition

$$\eta \leqslant \frac{h_K}{\mathrm{diam}(K)} = \frac{2}{\sqrt{3}\sqrt{1+v^2}} \tag{30}$$

should be satisfied. The investigations of Example 1 suggested to define τ in (6) and (13) by (27). Since an interior layer is a characteristic layer, it is natural to ask whether this modification is reasonable also in the present example. Then, instead of (30), we obtain the condition

$$\eta \leqslant \frac{h_K^{\perp}}{\mathrm{diam}(K)} = \frac{2v}{(\sqrt{3}+v)\sqrt{1+v^2}}. \tag{31}$$

Fig. 7 compares the curves $\eta_2 = \eta_2(v)$ for both the isotropic and the crosswind artificial diffusion with the functions on the right-hand sides of (30) and (31). Values of the right-hand sides of (30) and (31) below the curves of $\eta_2(v)$ indicate that the values of (6) and (13) are too small to suppress the oscillations at the interior layer below the value 10^{-2}. Thus, Fig. 7 shows that the method

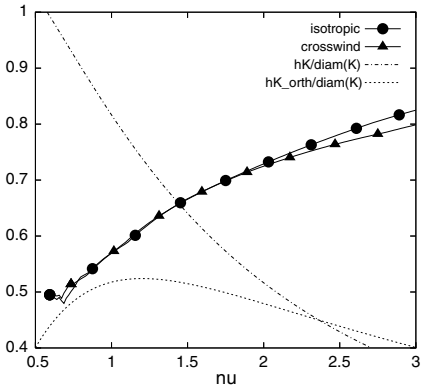

Fig. 7. Example 2, dependence on v of η_2 for the isotropic and the crosswind artificial diffusion and of the functions from the right-hand sides of (30) and (31).

of do Carmo and Galeão and the modified method of Burman and Ern will generally lead to non-negligible spurious oscillations at the interior layer of Example 2. Replacing h_K by $h_K^{\perp}$ in the definition of τ used in (6) and (13), oscillations of size at least 10^{-2} should appear for any aspect ratio and they should be mostly even larger than for τ defined using h_K. Thus, in contrast to Example 1, τ in (6) and (13) should be defined rather using h_K for small aspect ratios ($v \lesssim 1.5$) and using even a measure larger than h_K, for instance $\mathrm{diam}(K)$, for larger aspect ratios.

Next, the usefulness of the curves presented in Fig. 7 will be demonstrated. Considering, e.g., $v = 2$, one expectation is that the method of Codina given by (8) and (10) with $C = 0.7$, whose parameter $\tilde{\varepsilon}$ corresponds to the solid lines, leads to a solution with small spurious oscillations at the interior layer (less than 10^{-2}). In contrast, the methods of do Carmo and Galeão, (5) and (6), and of Burman and Ern, (8) and (13), whose parameters correspond to the dash-dot line, should produce solutions with larger oscillations at the interior layer. Fig. 8 shows numerical examples which confirm both expectations. For the methods (5), (6) and (8), (13), the results obtained

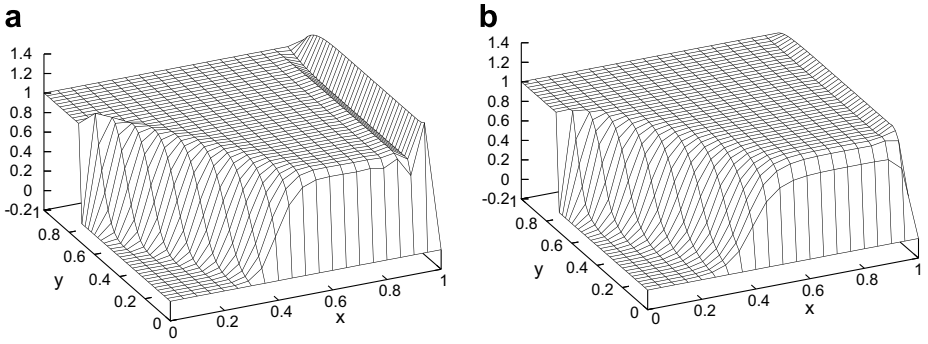

Fig. 8. Example 2, discrete solution u_h obtained on 21×41 meshes using (a) the method of do Carmo and Galeão and the Q_1 finite element and (b) the method of Codina with $C = 0.7$ and the P_1 finite element.

with both the P_1 and the Q_1 finite elements are similar. In particular, these solutions possess non-negligible spurious oscillations at the beginning of the interior layer. Considering the method of Codina and the Q_1 finite element, the violation of the discrete maximum principle at the beginning of the interior layer is larger and mainly in form of undershoots. For the method of Burman and Ern given by (8) and (12), the results are similar as for the method of Codina but slightly worse with respect to the spurious oscillations.

As pointed out above, the results of a SOLD method depend not only on the definition of $\tilde{\varepsilon}$ but also on the definition of τ in the SUPG term. In addition, we explained that the formula (3) is optimal with respect to the boundary layer at $y = 0$. Neglecting for the moment the quality of the solution at this boundary layer, one can ask whether increasing τ can help to reduce the spurious oscillations at the characteristic layer. However, the expectations are rather low because, in case of a characteristic layer, the influence of the choice of τ is usually weak since the SUPG method stabilizes in the streamline direction which is nearly perpendicular to the direction in which oscillations appear. Fig. 9 shows a comparison of η_4 for both the isotropic and the crosswind artificial diffusion and for two choices of τ. One choice of τ is the same as before and the other one is given by the formula (3) where h_K is replaced by $\mathrm{diam}(K)$. The use of the element diameter in the definition of τ is quite common in practice. It can be seen that increasing the amount of the streamline diffusion provided by the SUPG method requires to introduce more crosswind diffusion by the SOLD term if larger aspect ratios are used to reduce the oscillations at the characteristic layer below 10^{-4}. In summary, generally, the spurious oscillations at the interior layer present in the solution of a SOLD method cannot be expected to become smaller if higher values of the SUPG parameter τ are used.

Let us now consider the boundary layers. One can observe in Fig. 8 that the boundary layer at $y = 0$ is slightly smeared and that oscillations appear along the boundary layer at $x = 1$. The smearing is not surprising since the SUPG solution approximates the boundary layer

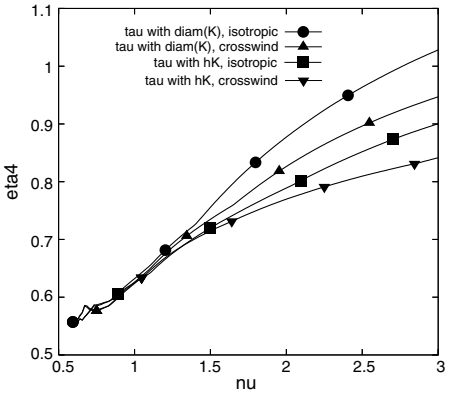

Fig. 9. Example 2, dependence of η_4 on ν for the isotropic and the crosswind artificial diffusion and for the SUPG parameter τ defined either by (3) or by (3) with h_K replaced by $\mathrm{diam}(K)$.

at $y = 0$ nodally exactly for $\varepsilon \to 0+$. Thus, along the boundary layer at $y = 0$, the optimal choice of $\tilde{\varepsilon}$ in a SOLD term is $\tilde{\varepsilon} = 0$, i.e., $\eta_{\mathrm{opt}} = 0$ in (20). To investigate the optimality of $\tilde{\varepsilon}$ for the boundary layer at $x = 1$ with $y \in [h_2, 1]$, let us again consider $\varepsilon \to 0+$ and $\tilde{\varepsilon}$ given by (20). The optimal solution has the values $u_h = 1$ at the nodes with $x = 1 - h_1$. A straightforward computation reveals that the value of η for obtaining this optimal solution is

$$\eta_{\mathrm{opt}} = \frac{h_1 b_2 + h_2 b_1}{\mathrm{diam}(K) b_2}$$

for the isotropic artificial diffusion (5) and

$$\eta_{\mathrm{opt}} = \frac{(h_1 b_2 + h_2 b_1)|\boldsymbol{b}|^2}{\mathrm{diam}(K) b_2^3}$$

for the crosswind artificial diffusion (8). These formulas hold for both the P_1 and the Q_1 finite elements. One can see that the optimal choice of η depends not only on the aspect ratio of the elements of the triangulation but also on the direction of the convection vector $\boldsymbol{b}$. The most important conclusion is that different values of η should be used in different regions of the computational domain.

To find a universal formula for the optimal value of η is very difficult or even impossible. This will be demonstrated by studying a limit case of Example 2 where the limit is approached in two different ways. First, consider the limit case $b_1 \to 0+$ and $b_2 \to -1$, $|\boldsymbol{b}| = 1$. Then, for both SOLD terms (5) and (8), we get

$$\eta_{\text{opt}} = \frac{h_1}{\text{diam}(K)} \quad \text{along the boundary } x = 1 \text{ if } \boldsymbol{b} = (0,-1).$$

On the other hand, consider $\boldsymbol{b} = (0,-1)$, the boundary conditions of Example 2 and a constant right-hand side $f > 0$ of (1). The optimal solution on the mesh line at $x = 1 - h_1$ has the form $u(x,y) = f(1-y) + 1$ (away from the lower boundary). Now, using the considerations leading to (23) gives

$$\eta_{\text{opt}} = \frac{2h_1}{3\text{diam}(K)} \quad \text{along the boundary } x = 1 \text{ if } \boldsymbol{b} = (0,-1) \tag{32}$$

independently of the choice of f. In particular, (32) holds for $f \to 0+$ and hence we obtained two different limit values of η_{opt}.

For the edge stabilization term (16) and both the P_1 and Q_1 finite elements, one can show similarly as above that the optimal value of the parameter C at $x = 1$ is

$$C_{\text{opt}} = \frac{h_1 b_2 + h_2 b_1}{4 h_1 b_2}.$$

For $\boldsymbol{b} = (0,-1)$, the limit values of the optimal C at $x = 1$ are 1/6 for Example 1 and 1/4 for Example 2 and hence they also differ by the factor 2/3. Choosing $C = C_{\text{opt}}$ in Example 2 still leads to oscillations at the interior layer. These can be suppressed by increasing the value of C in this region. This shows once again that different values of the parameter should be used in different regions of the computational domain to obtain a globally satisfactory solution.

The above discussion supports our conclusion to Example 1 that the best SOLD methods are the modified method of Codina and the edge stabilization (16), however, only if the parameter C is chosen appropriately, i.e., layer-adapted. Nevertheless, one generally cannot expect that the discrete solutions will be without any spurious oscillations.

Example 3 (*Solution with two interior layers*). We consider the convection–diffusion equation (1) with (17) and

$$\boldsymbol{b} = (1,0)^{\mathrm{T}}, \quad u_b = 0,$$

$$f(x,y) = \begin{cases} 16(1-2x) & \text{for } (x,y) \in [0.25, 0.75]^2, \\ 0 & \text{else.} \end{cases}$$

The solution, see Fig. 10a, possesses two interior (characteristic) layers at $(0.25, 0.75) \times \{0.25\}$ and $(0.25, 0.75) \times \{0.75\}$. In $(0.25, 0.75)^2$, the solution $u(x,y)$ is very close to the quadratic function $(4x-1)(3-4x)$. This example was first considered by John and Knobloch [25].

This is an example of a problem for which all the SOLD methods mentioned in Section 3 fail. Note that, in contrast to Example 2, the data of Example 3 satisfy the requirements for defining the standard weak formulation of (1). Moreover, the solution of Example 3 belongs to $H^2(\Omega)$, cf. Grisvard [18].

As expected, the SUPG solution of Example 3 possesses spurious oscillations along the interior layers, see Fig. 10b. To visualize both undershoots and overshoots, we present the SUPG solution at an angle for which the plane $z = 0$ reduces to a line. Applying the modified method of Codina with $C = 0.7$, the spurious oscillations present in the SUPG solution are significantly suppressed, however, the solution is wrong in the region $(0.75, 1) \times (0, 1)$, see Fig. 11. Very similar results are obtained for *any* of the SOLD methods mentioned in Section 3 and for both the P_1 and Q_1 finite elements.

Note that, in view of the discontinuous right-hand side f, the SOLD methods should be implemented using quadrature formulas whose nodes do not lie on the edges of the triangulations. However, such nodes cannot be avoided when evaluating the edge stabilization term (16), which complicates the implementation of this method.

To measure the spurious oscillations of a discrete solution u_h to Example 3, we define the values

$$min := -\min_{0.4 \leqslant x \leqslant 0.6} u_h(x,y), \quad diff := \max_{x \geqslant 0.8} u_h(x,y) - \min_{x \geqslant 0.8} u_h(x,y), \tag{33}$$

where $y \in [0,1]$ and $\min u_h$ and $\max u_h$ are computed using values of u_h at the vertices of $\mathcal{T}_h$. Tables 1 and 2 show the values of *min* and *diff*, respectively, for the P_1 finite element, most of the SOLD methods discussed above and several meshes. The abbreviations denoting the methods can be

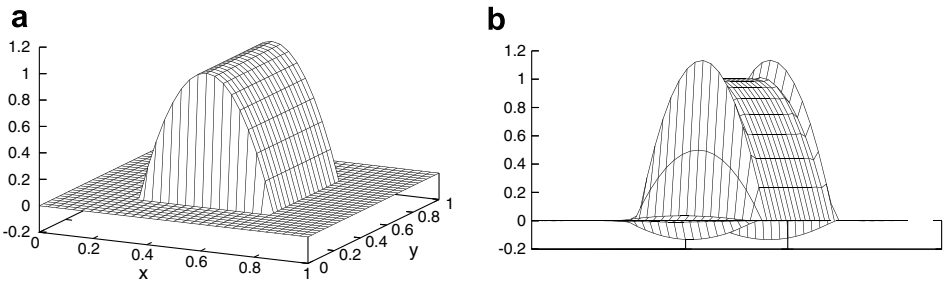

Fig. 10. Example 3: (a) solution u and (b) discrete solution u_h obtained using the SUPG method with the P_1 finite element on a 33×33 mesh.

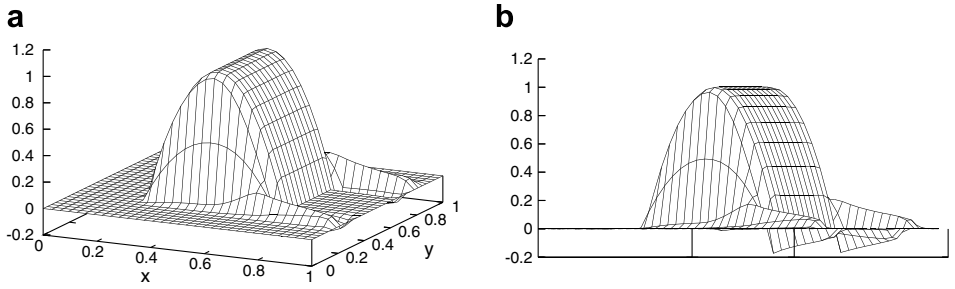

Fig. 11. Example 3, discrete solution u_h obtained on a 33×33 mesh using the modified method of Codina with $C = 0.7$ and the P_1 finite element: (a) view as in Fig. 10a and (b) view as in Fig. 10b.

Table 1
Example 3, values of *min* defined in (33) obtained for the P_1 finite element using the methods from Sections 2 and 3

Method	Mesh			
	17×17	33×33	65×65	129×129
SUPG	1.31e−1	1.33e−1	1.34e−1	1.34e−1
dCG91	2.37e−2	1.27e−2	2.42e−3	n.c.
KLR02_3, $C = 0.4714$	1.93e−2	1.88e−2	1.22e−2	6.85e−3
KLR02_3, $C = 0.7$	8.52e−3	1.38e−3	2.65e−4	n.c.
BE02_1	1.37e−2	9.33e−3	n.c.	n.c.
BE02_2	1.85e−2	7.74e−3	1.20e−3	n.c.
BE05_2, $C = 1/6$	1.06e−2	6.77e−3	3.98e−3	2.04e−3
BE05_2, $C = 0.4$	2.79e−3	1.59e−3	8.24e−4	n.c.

Table 2
Example 3, values of *diff* defined in (33) obtained for the P_1 finite element using the methods from Sections 2 and 3

Method	Mesh			
	17×17	33×33	65×65	129×129
SUPG	3.30e−3	9.52e−5	3.83e−5	1.53e−4
dCG91	2.62e−1	2.95e−1	2.81e−1	n.c.
KLR02_3, $C = 0.4714$	2.88e−1	3.24e−1	3.37e−1	3.37e−1
KLR02_3, $C = 0.7$	2.82e−1	2.74e−1	2.42e−1	n.c.
BE02_1	3.77e−1	4.36e−1	n.c.	n.c.
BE02_2	2.78e−1	2.94e−1	2.76e−1	n.c.
BE05_2, $C = 1/6$	2.76e−1	3.05e−1	3.25e−1	3.36e−1
BE05_2, $C = 0.4$	2.53e−1	2.56e−1	2.43e−1	n.c.

found in Section 3 and are the same as in Part I. The abbreviation nc means that the nonlinear iterative process did not converge, see the next section. This happens mainly for the finest mesh. Generally, the convergence of the nonlinear iterations deteriorates if the mesh becomes finer or the parameter C in (11) or (16) increases. We consider two values of C for each method. First, since interior layers are characteristic layers, we use the optimal values of C found in the investigations of Example 1. For (11), we further use the value $C = 0.7$ recommended in [10]. For (16), the value $C = 0.4$ corresponds to the choice of C in Part I. Table 1 shows that all the SOLD methods significantly reduce the undershoots along the interior layers present in the SUPG solution (the same holds for overshoots). For the considered meshes, the maximal undershoots of the SUPG method are not influenced by the size of the mesh width. In contrast to this, for all the SOLD methods, the undershoots become smaller if the mesh is refined. The undershoots also decrease if the parameter C in (11) or (16) increases. However, for larger values of C, the smearing of the discrete solution is more pronounced and, as we mentioned, the convergence of the nonlinear iterative process deteriorates.

Table 2 shows that the wrong part of the discrete solution in $(0.8, 1) \times (0, 1)$ is of comparable magnitude for all the SOLD methods and does not improve significantly if the mesh is refined or C is increased (both in the range where the nonlinear iterative schemes converge). Therefore, we conclude that, using the SOLD methods described in Section 3, it is not feasible to obtain a qualitatively correct approximation of the solution to Example 3. An open question is whether appropriately defined non-constant parameters in the modified method of Codina (11) or the edge stabilization (16) might lead to satisfactory solutions.

5. The solution of the nonlinear discrete problems

The discrete SOLD problems can be written in the form

$$a_h(u_h; u_h, v_h) = \langle f, v_h \rangle \quad \forall v_h \in V_h,$$

where $a_h(u_h; \cdot, \cdot)$ is a bilinear form and the first argument of a_h enters the definition of a_h through the parameter $\tilde{\varepsilon}$ or the respective term in (16). Thus, it is straightforward to compute the discrete solution by means of the following iterative scheme. Given an approximation u_h^k of the solution of the SOLD system, compute $\tilde{u}_h^{k+1}$ by solving

$$\tilde{u}_h^{k+1}: \quad a_h(u_h^k; \tilde{u}_h^{k+1}, v_h) = \langle f, v_h \rangle \quad \forall v_h \in V_h. \tag{34}$$

The next iterate is defined as

$$u_h^{k+1} := u_h^k + \omega_{k+1}(\tilde{u}_h^{k+1} - u_h^k)$$

with the damping factor $\omega_{k+1} > 0$.

As initial iterate u_h^0, we use the solution obtained with the SUPG method. Thus, apart from the spurious oscillations, the initial iterate coincides already rather well with the solution wished to be obtained with the SOLD methods.

Our experiences are that an appropriate choice of the damping factors $\{\omega_k\}$ is often essential for the convergence

of the iterative process and the number of iterations. Appropriate damping factors depend on the SOLD scheme, the problem and its data, the grid and the choice of parameters in parameter-dependent SOLD schemes and these damping factors might be very different. Since it is not practicable in applications that the user should find every time an appropriate damping factor, it is necessary to use a strategy for an automatic and dynamic choice of this factor.

The dynamic choice of the damping factor which we used in our computations is illustrated with the pseudo code in Fig. 12. Our approach contains a number of parameters, whose values for the results presented in this section are given on lines 1–2. These values seemed reasonable choices in our opinion and we did not try to optimize them for the examples considered in this paper. Our strategy for the dynamic choice of the damping factor is based on the following principles:

- There is an upper bound $\omega_{\max}$ for the damping factor. The upper bound is adjusted dynamically in the course of the iterative process. Initially, we set $\omega_{\max} = 1$, i.e., no damping.
- There is a lower bound $\omega_{\min}$ for the damping factor. This bound is fixed. We used in the computations presented in this paper $\omega_{\min} = 0.01$. Note that very small damping factors lead in general to a very large number of iterations and thus to inefficient schemes.
- The iterate u_h^{k+1} is accepted if the norm $|R_h(u_h^{k+1})|$ of its residual

$$\langle R_h(u_h^{k+1}), v_h\rangle := a_h(u_h^{k+1}; u_h^{k+1}, v_h) - \langle f, v_h\rangle, \quad v_h \in V_h,$$

is smaller than $|R_h(u_h^k)|$ or if ω is not allowed to decrease any more, see the pseudo code presented in Fig. 12, lines 10–14. If $|R_h(u_h^{k+1})| < |R_h(u_h^k)|$ and if there was no rejection of an iterate u_h^{k+1} for a larger value of ω before, the maximal damping factor will be increased, see line l2, and then the damping factor will be increased, too, see line l3.
- If the proposal for the iterate u_h^{k+1} is not accepted, ω will be decreased, see line 16. In addition, if in the step $k+1$ an iterate is rejected the first time, $\omega_{\max}$ will be decreased too, see lines 17–20. Now, a new proposal for u_h^{k+1} is computed with the new value of the damping factor. The acceptance or rejection of this new proposal is checked the same way as for the former damping factor.

```
 1. ω_min := 0.01;  ω_max := 1
 2. c_1 := 1.001;  c_2 := 1.1;  c_3 := 1.001;  c_4 := 0.9
 3. compute SUPG solution u_h^0 and residual r^0
 4. ω := ω_max;  k := 0
 5. while r^k > tolerance do
 6.   compute ũ_h^{k+1} satisfying (34)
 7.   first_damp := 1
 8.   u_h^{k+1} := u_h^k + ω(ũ_h^{k+1} − u_h^k)
 9.   compute residual r^{k+1}
10.   if r^{k+1} < r^k or ω ≤ c_1 ω_min then
11.     if r^{k+1} < r_k and first_damp = 1 then
12.       ω_max := min{1, c_3 ω_max}
13.       ω := min{ω_max, c_2 ω}
14.     endif
15.   else
16.     ω := max{ω_min, ω/2}
17.     if first_damp = 1 then
18.       ω_max := max{ω_min, c_4 ω_max}
19.       first_damp := 0
20.     endif
21.     goto line 8
22.   endif
23.   k := k + 1
24. endwhile
```

Fig. 12. Dynamic choice of the damping factor.

The main features of this approach are as follows:

- The damping factor decreases in general if the residual increases.
- The decrease of the damping factor stops at the threshold $\omega_{\min}$ so that also a non-monotone sequence with respect to the norm of the residual can be computed.
- The damping factor as well as the maximal damping parameter increase if the residual decreases to improve the efficiency of the nonlinear iteration scheme. Thus, a strong damping, which might be necessary only at the beginning of the iterative process, influences the damping factor at the end of the process only slightly.

In the simulations presented in this paper, the linear systems were solved by a sparse direct solver (UMFPACK, [13]). Since the costs for solving the linear systems are always the same, this leads to a fair comparison of the costs of the iterative process for all SOLD schemes by simply giving the number of nonlinear iterations.

In practice, it suffices to solve the linear systems only approximately by a few steps of an iterative method without affecting the convergence of the nonlinear iterative method much. This approach might be faster, depending on the iterative linear system solver. However, different numbers of iterations for solving the linear systems are in general necessary for different SOLD schemes, which makes it harder to perform a fair comparison.

Below, our experiences with respect to the solution of the nonlinear discrete problems corresponding to the examples of Section 4 are reported. Tables with characteristic results are presented, where besides the dynamic approach for computing the damping factor also numbers of iterations with fixed factors are given. The computations were carried out for the P_1 and the Q_1 finite elements on 65×65, 33×65 and 65×33 meshes. The iterative processes were stopped if the l^2-norm of the residual vector was smaller than 10^{-8} or after 100,000 iterations (n.c. = not convergent in the tables). Again, the abbreviations of the SOLD methods given in Section 3 are used.

The numbers of iterations generally depend on the quadrature formula used and this dependence is stronger

for the Q_1 finite element than for the P_1 finite element. All results in this paper were computed using Gaussian quadrature formulas of order 5 (with 7 nodes in case of triangles and 9 nodes in case of rectangles). Of course, for Examples 1 and 2 discretized using the P_1 finite element, the results are independent of the used quadrature formula since all integrands are constant or linear.

We would like to emphasize that analytical results concerning the existence and uniqueness of solutions to the nonlinear discrete problems are not available. Thus, it cannot be excluded that a failure of all used damping strategies has its reason in the non-existence of the solution of the nonlinear discrete problem.

Example 1. The nonlinear discrete problems on the 65×65 and the 65×33 meshes could be usually solved without damping, see Table 3. Apart from dCG91 and BE05_2 with $C = 0.4$, the iterative schemes converged in only few iterations. Solving the problems on the 33×65 mesh required for some SOLD methods considerable damping, see Table 4 for the Q_1 finite element. For the P_1 finite element, the convergence was mostly even worse than in Table 4 and dCG91 did not converge at all. Except the latter case, the dynamic choice of the damping factor was always successful, but often more iterations were needed than with the best fixed damping factor, cf. also the last row in Table 3. In these computations, the dynamic approach proposes many damping factors close to $\omega_{\min}$ because the norm of the residual is slightly oscillating, before finally convergence is achieved. Note that the numbers of iterations for the optimal constant in KLR02_3 are very small on both meshes.

Table 3
Example 1, number of iterations for solving the nonlinear SOLD problems, 65×65 mesh, P_1 finite element

Method	$\omega = 0.25$	$\omega = 0.5$	$\omega = 0.75$	$\omega = 1$	Dynamic
dCG91	472	236	161	169	169
KLR02_3, $C = 0.4714$	71	32	18	9	9
KLR02_3, $C = 0.7$	108	50	32	22	22
BE02_1	76	36	24	28	28
BE02_2	92	44	27	19	19
BE05_2, $C = 1/6$	164	78	50	29	29
BE05_2, $C = 0.4$	1010	506	345	n.c.	10943

Table 4
Example 1, number of iterations for solving the nonlinear SOLD problems, 33×65 mesh, Q_1 finite element

Method	$\omega = 0.25$	$\omega = 0.5$	$\omega = 0.75$	$\omega = 1$	Dynamic
dCG91	394	n.c.	n.c.	n.c.	935
KLR02_3, $C = 0.2981$	73	33	20	13	13
KLR02_3, $C = 0.7$	119	64	63	157	66
BE02_1	235	173	218	n.c.	339
BE02_2	213	380	n.c.	n.c.	353
BE05_2, $C = 1/6$	78	36	23	72	72
BE05_2, $C = 0.4$	n.c.	n.c.	n.c.	n.c.	n.c.

Example 2. The nonlinear discrete SOLD problems in this example were harder to solve than for Example 1, in particular for the P_1 finite element. Even on the equidistant mesh, strong damping was necessary, see Table 5. The dynamic choice of the damping factor always led to the convergence of the iterative process on this mesh. Using the P_1 finite element on the 65×33 mesh, the nonlinear problems could be solved only for KLR02_3 and BE05_2 with sufficiently small parameters. The solution of the discrete problems with the Q_1 finite element was much easier on all grids, see Table 6 for representative results.

Example 3. Using the equidistant 65×65 mesh with the P_1 and Q_1 finite element, the discrete equations could be solved without damping for most of the SOLD methods, see Table 7. Only for BE02_1 and BE05_2 with $C = 0.4$, it was not possible to solve them at all, see also Tables 1 and 2. These tables show also that the solution of the nonlinear problems for the P_1 finite element on the next finer equidistant grid became more difficult. We could obtain convergence only for the parameter-dependent SOLD schemes with sufficiently small parameters. For the P_1 finite element on the 33×65 mesh, the iterative processes was

Table 5
Example 2, number of iterations for solving the nonlinear SOLD problems, 65×65 mesh, P_1 finite element

Method	$\omega = 0.25$	$\omega = 0.5$	$\omega = 0.75$	$\omega = 1$	Dynamic
dCG91	160	n.c.	n.c.	n.c.	340
KLR02_3, $C = 0.7$	194	n.c.	n.c.	n.c.	408
BE02_1	n.c.	n.c.	n.c.	n.c.	389
BE02_2	210	n.c.	n.c.	n.c.	412
BE05_2, $C = 0.4$	362	n.c.	n.c.	n.c.	536

Table 6
Example 2, number of iterations for solving the nonlinear SOLD problems, 65×65 mesh, Q_1 finite element

Method	$\omega = 0.25$	$\omega = 0.5$	$\omega = 0.75$	$\omega = 1$	Dynamic
dCG91	67	36	29	33	33
KLR02_3, $C = 0.7$	102	58	51	60	60
BE02_1	213	275	n.c.	n.c.	203
BE02_2	84	47	39	45	45
BE05_2, $C = 0.4$	689	n.c.	n.c.	n.c.	7520

Table 7
Example 3, number of iterations for solving the nonlinear SOLD problems, 65×65 mesh, P_1 finite element

Method	$\omega = 0.25$	$\omega = 0.5$	$\omega = 0.75$	$\omega = 1$	Dynamic
dCG91	158	86	59	49	49
KLR02_3, $C = 0.4714$	157	74	46	33	33
KLR02_3, $C = 0.7$	199	115	89	115	110
BE02_1	n.c.	n.c.	n.c.	n.c.	n.c.
BE02_2	178	93	65	62	62
BE05_2, $C = 1/6$	173	83	53	37	37
BE05_2, $C = 0.4$	n.c.	n.c.	n.c.	n.c.	n.c.

Table 8
Example 3, number of iterations for solving the nonlinear SOLD problems, 33×65 mesh, Q_1 finite element

Method	$\omega = 0.25$	$\omega = 0.5$	$\omega = 0.75$	$\omega = 1$	Dynamic
dCG91	475	n.c.	n.c.	n.c.	599
KLR02_3, $C = 0.2981$	123	58	36	25	25
KLR02_3, $C = 0.7$	247	168	n.c.	n.c.	345
BE02_1	332	974	n.c.	n.c.	461
BE02_2	317	432	n.c.	n.c.	381
BE05_2, $C = 1/6$	150	72	46	33	33
BE05_2, $C = 0.4$	565	277	184	146	1640

not convergent for dCG91, BE02_2 and KLR02_3 with $C = 0.7$. The results for the Q_1 finite element and the 33×65 mesh are presented in Table 8. Again, the need of damping can be observed as well as the successfulness of the dynamic approach (however, on the expense of somewhat more iterations than for the best fixed damping factors). On 65×33 meshes, the only method which did not converge at all was BE05_2 with $C = 0.4$.

Remark 2. The numerical studies show that even for the academic test problems considered in this paper, it was sometimes difficult to solve the nonlinear SOLD problems. Considering more challenging problems, like the one defined by Hemker [19], the difficulties in the solution of the nonlinear problems became even greater. For instance, convergence for KLR02_3 on reasonably structured grids could be achieved only for rather small constants C.

Remark 3. Another possibility for solving the nonlinear discrete problems is to apply Newton's method. However, it is rather difficult to implement since one deals with non-smooth operators. Therefore, usually it is convenient to use some simplified version of Newton's method. In any case, a good initial approximation is typically needed. Hence a general strategy is first to apply the iterative scheme given above and then to switch to Newton's method, possibly by performing several special iterations assuring a smooth transition between the two iterative processes. An appropriate switching point or transition strategy depend on the solved problem, the SOLD scheme, the grid, etc. If the norm of the residual increases after switching to Newton's method, it is advisable to return to the original iterative process without employing the results of the Newton iterations and to try to switch to Newton's method at a later stage of the iterative process. Applying this alternative strategy instead of the iterative scheme studied in this paper, the numbers of iterations change of course but the ranking of the methods basically remains the same. This can be explained by our observation that the methods with a large number of iterations in Tables 3–8 usually show a slow rate of convergence from the beginning of the iterative process. Thus, these methods also require a large number of iterations for obtaining a good initial approximation for Newton's method.

Our experiences concerning the solution of the nonlinear SOLD problems can be summarized as follows:

- Generally, it was easier to solve the problems for the Q_1 finite element than for the P_1 finite element.
- The larger the constant in the SOLD methods KLR02_3 and BE05_2, the more iterations were needed. If the constant became too large (size depended on the problem, the grid, etc.), the iterative process did not solve the nonlinear problem any more.
- It was often easier to solve the problems arising from the SOLD method BE02_2 than those coming from BE02_1.
- Solving the problems obtained with the edge stabilization BE05_2 required in general somewhat more iterations than solving the problems coming from KLR02_3, if in both SOLD methods reasonable constants with respect to the reduction of the spurious oscillations have been chosen. Moreover, the convergence of BE05_2 was much more sensitive to the choice of the parameter C than it was for the method KLR02_3.
- If the nonlinear discrete problems could be solved at all, the dynamic choice of the damping factor was generally among the successful approaches. If damping was necessary, the dynamic approach needed often more iterations than an appropriately chosen fixed damping factor.

6. Numerical results obtained with adaptive methods

In several discussions with our colleagues about Part I, the question arose whether the application of adaptive methods is useful for the reduction of spurious oscillations. In this section, we shall study this question for the SUPG method and adaptive grids obtained with two residual-based error estimators, which are typically used in applications.

There are different ways of defining criteria for a fair comparison of the results obtained with adaptive methods and with SOLD schemes. One possible criterion is to require that the number of degrees of freedom is roughly the same. A different one might be that the computing times are similar. Since the solution of the nonlinear discrete problems of the SOLD methods often is rather time-consuming (because of the large number of iterations), it is possible to solve the linear problems on adaptive meshes with much more degrees of freedom in the same time. Both criteria might be of interest and thus, we will present results on adaptive meshes starting with a few thousand degrees of freedom up to more than 100,000 degrees of freedom.

Computational studies for Example 2 will be presented. As starting grid for the adaptive refinement, we used the triangular grid from Fig. 1 with $h_1 = h_2 = 1/16$ (289 degrees of freedom). The control of the adaptive refinement process was performed analogously to the way described in Section 4 of [23]. The oscillations at the interior layer were

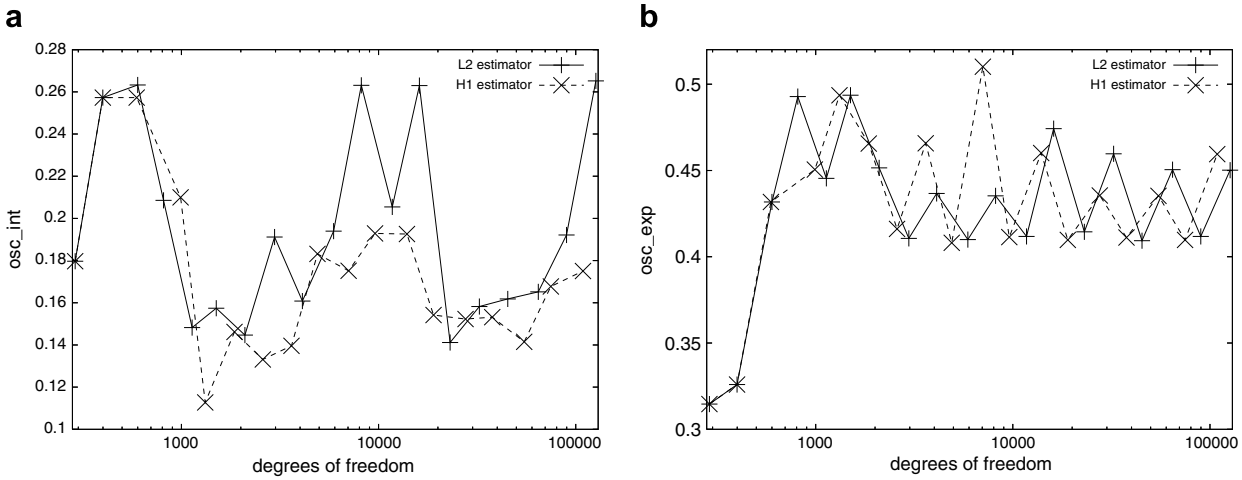

Fig. 13. Example 2, oscillations on adaptively refined grids: (a) interior layer; (b) exponential layer.

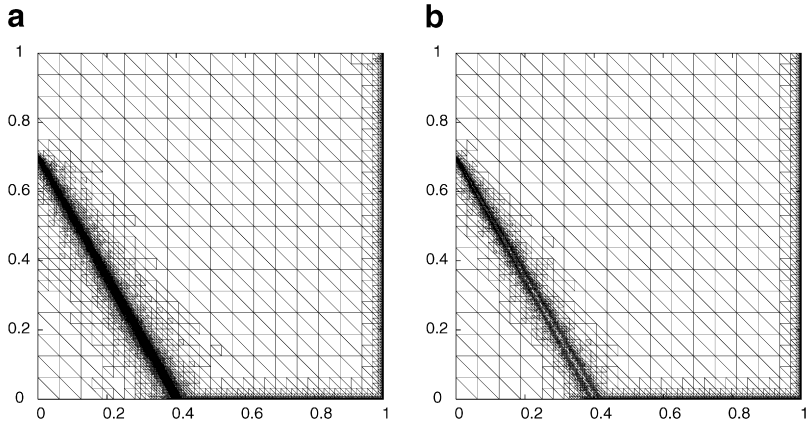

Fig. 14. Example 2, adaptive grids with more than 100,000 degrees of freedom: (a) L^2-error estimator; (b) H^1-seminorm error estimator.

measured with osc_{int} defined in (29) and the oscillations at the exponential boundary layer with

$$osc_{\text{exp}} := \max_{x \geqslant 0.7}(\max\{0, u_h(x,y) - 1\}).$$

We will present results for residual-based error estimators in the H^1-semi norm and the L^2-norm, see [39]. For a detailed description of these estimators and their implementation, we refer to [23]. The gradient indicator and a residual-based error estimator in the energy norm considered in [23] failed to refine the region of the interior layer. This coincides with their behavior observed in Examples 6.4 and 6.5 of [23].

The computational results for osc_{int} and osc_{exp} are presented in Fig. 13 and the final grids for both error estimators in Fig. 14. The meshes match the expectations on the error estimators since the regions of all layers are refined and a deeper refinement occurs at the exponential boundary layers. The graphs in Fig. 13 show that the adaptive refinement of the layer regions neither reduces the spurious oscillations at the interior layer nor at the boundary layers. The adaptively refined meshes are still too coarse in these regions to resolve the layers and to suppress the oscillations.

This section showed exemplarily that a suppression of spurious oscillations cannot be achieved with adaptively refined grids whose elements do not resolve the layers.

7. Conclusions

This paper studied in detail SOLD methods which were identified in Part I as the best ones. In particular, the limits of the available methods were demonstrated. Analytical and numerical studies showed that SOLD methods without user-chosen parameters are in general not able to remove the spurious oscillations of the solution obtained with the SUPG discretization. For the two studied methods involving a parameter, the modified method of Codina (8), (11) and the edge stabilization (16), values of the parameter could be derived in two examples such that the spurious oscillations were almost removed. It turned out that a spatially constant choice of the parameters was not sufficient in general and that the optimal parameters depended on the data of the problem and on the

grid. In addition, an example was presented for which none of the investigated methods provided a qualitatively correct discrete solution.

The iterative solution of the nonlinear discrete problems was also studied. The number of iterations or the convergence of the iterative process depended again on the problem, the grid and the parameters of the SOLD methods. In particular, the convergence of the nonlinear iterative process for the edge stabilization (16) proved to be rather sensitive to this parameter. It could be observed that the convergence is often strongly influenced by the choice of an appropriate damping factor and a strategy was proposed for an automatic and dynamic computation of this factor.

Finally, it was demonstrated that adaptive grid refinement generally does not lead to a suppression of the spurious oscillations of the solutions computed with the SUPG discretization.

Considering the reduction of the spurious oscillations, the sharpness of the layers and the computational overhead for solving the nonlinear discrete problem, the SOLD methods involving parameters, i.e., the modified method of Codina (8), (11) and the edge stabilization method (16), seem to be the only reasonably promising approaches among the studied SOLD methods. However, the appropriate definition of the generally non-constant parameters in these methods will represent a great difficulty in more complicated problems and in applications. Future research should develop an a posteriori algorithm for an automatic choice of these parameters.

The current situation can be summarized as follows: it is in general completely open how to obtain oscillation-free solutions using the considered classes of methods.

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] R.C. Almeida, R.S. Silva, A stable Petrov–Galerkin method for convection-dominated problems, Comput. Methods Appl. Mech. Engrg. 140 (1997) 291–304.
[2] Y. Bazilevs, V.M. Calo, T.E. Tezduyar, T.J.R. Hughes, $YZ\beta$ discontinuity capturing for advection-dominated processes with application to arterial drug delivery, Int. J. Numer. Methods Fluids 54 (2007) 593–608.
[3] 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. Engrg. 32 (1982) 199–259.
[4] E. Burman, A. Ern, Nonlinear diffusion and discrete maximum principle for stabilized Galerkin approximations of the convection–diffusion–reaction equation, Comput. Methods Appl. Mech. Engrg. 191 (2002) 3833–3855.
[5] E. Burman, A. Ern, Stabilized Galerkin approximation of convection–diffusion–reaction equations: discrete maximum principle and convergence, Math. Comput. 74 (2005) 1637–1652.
[6] E. Burman, P. Hansbo, Edge stabilization for Galerkin approximations of convection–diffusion–reaction problems, Comput. Methods Appl. Mech. Engrg. 193 (2004) 1437–1453.
[7] 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. Engrg. 192 (2003) 3379–3396.
[8] E.G.D. do Carmo, A.C. Galeão, Feedback Petrov–Galerkin methods for convection-dominated problems, Comput. Methods Appl. Mech. Engrg. 88 (1991) 1–16.
[9] P.G. Ciarlet, Basic error estimates for elliptic problems, in: P.G. Ciarlet, J.L. Lions (Eds.), Handbook of Numerical Analysis, Finite Element Methods, vol. 2 (pt. 1), North-Holland, Amsterdam, 1991, pp. 17–351.
[10] R. Codina, A discontinuity-capturing crosswind-dissipation for the finite element solution of the convection–diffusion equation, Comput. Methods Appl. Mech. Engrg. 110 (1993) 325–342.
[11] R. Codina, Comparison of some finite element methods for solving the diffusion–convection–reaction equation, Comput. Methods Appl. Mech. Engrg. 156 (1998) 185–210.
[12] R. Codina, O. Soto, Finite element implementation of two-equation and algebraic stress turbulence models for steady incompressible flows, Int. J. Numer. Meth. Fluids 30 (1999) 309–333.
[13] T.A. Davis, Algorithm 832: UMFPACK V4.3 – an unsymmetric-pattern multifrontal method, ACM Trans. Math. Software 30 (2004) 196–199.
[14] J. Donéa, A Taylor–Galerkin method for convection transport problems, Int. J. Numer. Methods Engrg. 20 (1984) 101–119.
[15] J. Douglas, T.F. Russell, Numerical methods for convection dominated diffusion problems based on combining the method of characteristics with finite element or finite difference procedures, SIAM J. Numer. Anal. 19 (1982) 871–885.
[16] 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. Engrg. 95 (1992) 253–276.
[17] A.C. Galeão, E.G.D. do Carmo, A consistent approximate upwind Petrov–Galerkin method for convection-dominated problems, Comput. Methods Appl. Mech. Engrg. 68 (1988) 83–95.
[18] P. Grisvard, Elliptic Problems in Nonsmooth Domains, Pitman, 1985.
[19] P.W. Hemker, A singularly perturbed model problem for numerical computation, J. Comput. Appl. Math. 76 (1996) 277–285.
[20] 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. Engrg. 127 (1995) 387–401.
[21] 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. Engrg. 73 (1989) 173–189.
[22] T.J.R. Hughes, M. Mallet, A. Mizukami, A new finite element formulation for computational fluid dynamics: II. Beyond SUPG, Comput. Methods Appl. Mech. Engrg. 54 (1986) 341–355.
[23] V. John, A numerical study of a posteriori error estimators for convection–diffusion equations, Comput. Methods Appl. Mech. Engrg. 190 (2000) 757–781.
[24] V. John, P. Knobloch, On discontinuity-capturing methods for convection–diffusion equations, in: A. Bermúdez de Castro, D. Gómez, P. Quintela, P. Salgado (Eds.), Numerical Mathematics and Advanced Applications, Proceedings of ENUMATH 2005, Springer-Verlag, Berlin, 2006, pp. 336–344.
[25] V. John, P. Knobloch, A computational comparison of methods diminishing spurious oscillations in finite element solutions of convection–diffusion equations, in: J. Chleboun, K. Segeth, T. Vejchodský (Eds.), Proceedings of the International Conference

Programs and Algorithms of Numerical Mathematics, vol. 13, Academy of Science of the Czech Republic, pp. 122–136.
[26] V. John, P. Knobloch, On spurious oscillations at layers diminishing (SOLD) methods for convection–diffusion equations: Part I – A review, Comput. Methods Appl. Mech. Engrg. 196 (2007) 2197–2215.
[27] V. John, P. Knobloch, On the performance of SOLD methods for convection–diffusion problems with interior layers, Int. J. Comput. Sci. Math. 1 (2007) 245–258.
[28] V. John, G. Matthies, MooNMD – A program package based on mapped finite element methods, Comput. Visual. Sci. 6 (2004) 163–170.
[29] V. John, M. Roland, T. Mitkova, K. Sundmacher, L. Tobiska, A. Voigt, Simulations of population balance systems with one internal coordinate using finite element methods, Technical Report, Universität des Saarlandes, FR. 6.1 – Mathematik, 2007.
[30] C. Johnson, Adaptive finite element methods for diffusion and convection problems, Comput. Methods Appl. Mech. Engrg. 82 (1990) 301–322.
[31] C. Johnson, A new approach to algorithms for convection problems which are based on exact transport + projection, Comput. Methods Appl. Mech. Engrg. 100 (1992) 45–62.
[32] 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.
[33] P. Knobloch, Improvements of the Mizukami–Hughes method for convection–diffusion equations, Comput. Methods Appl. Mech. Engrg. 196 (2006) 579–594.
[34] T. Knopp, G. Lube, G. Rapin, Stabilized finite element methods with shock capturing for advection–diffusion problems, Comput. Methods Appl. Mech. Engrg. 191 (2002) 2997–3013.
[35] 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. Engrg. 50 (1985) 181–193.
[36] B. Mohammadi, O. Pironneau, Analysis of the K–Epsilon Turbulence Model, John Wiley & Sons, 1994.
[37] T.E. Tezduyar, Finite element methods for fluid dynamics with moving boundaries and interfaces, in: E. Stein, R. De Borst, T.J.R. Hughes (Eds.), Encyclopedia of Computational Mechanics, Fluids, vol. 3, Wiley, New York, 2004 (Chapter 17).
[38] T.E. Tezduyar, Y.J. Park, Discontinuity-capturing finite element formulations for nonlinear convection–diffusion–reaction equations, Comput. Methods Appl. Mech. Engrg. 59 (1986) 307–325.
[39] R. Verfürth, A Review of a Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques, Wiley and Teubner, 1996.