HAL
archives-ouvertes.fr

Flux-Upwind Stabilization of the Discontinuous Petrov–Galerkin Formulation with Lagrangian Multipliers for Advection-Diffusion Problems

Carlo L. Bottasso, Paola Causin, Riccardo Sacco

▶ **To cite this version:**

Carlo L. Bottasso, Paola Causin, Riccardo Sacco. Flux-Upwind Stabilization of the Discontinuous Petrov–Galerkin Formulation with Lagrangian Multipliers for Advection-Diffusion Problems. [Research Report] RR-5157, INRIA. 2004. inria-00077044

HAL Id: inria-00077044
https://hal.inria.fr/inria-00077044

Submitted on 29 May 2006

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

INRIA

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Flux-Upwind Stabilization of the Discontinuous Petrov–Galerkin Formulation with Lagrangian Multipliers for Advection-Diffusion Problems

Carlo L. Bottasso — Paola Causin — Riccardo Sacco

N° 5157

Mars 2004

THÈME 4

ISSN 0249-6399 ISRN INRIA/RR--5157--FR+ENG

INRIA
ROCQUENCOURT

Flux-Upwind Stabilization of the Discontinuous Petrov–Galerkin Formulation with Lagrangian Multipliers for Advection-Diffusion Problems

Carlo L. Bottasso *, Paola Causin† , Riccardo Sacco ‡

Thème 4 — Simulation et optimisation
de systèmes complexes
Projet BANG

Rapport de recherche n° 5157 — Mars 2004 — 32 pages

Abstract: In this work we consider the dual-primal Discontinuous Petrov-Galerkin (DPG) method for the advection-diffusion model problem. Since in the DPG method both mixed internal variables are discontinuous, a static condensation procedure can be carried out, leading to a single-field nonconforming discretization scheme. For this latter formulation, we propose a flux-upwind stabilization technique to deal with the advection-dominated case. The resulting scheme is conservative and satisfies a discrete maximum principle under standard geometrical assumptions on the computational grid. A convergence analysis is developed, proving first-order accuracy of the method in a discrete H^1-norm, and the numerical performance of the scheme is validated on benchmark problems with sharp internal and boundary layers.

Key-words: Finite element methods, mixed and hybrid methods, Discontinuous Galerkin and Petrov-Galerkin methods, nonconforming finite elements, stabilized finite elements, upwinding, advection-diffusion problems.

* D. Guggenheim School of Aerospace Engineering, Georgia Institute of Technology, USA
† INRIA Rocquencourt, France
‡ MOX - Modeling and Scientific Computing, Dipartimento di Matematica, Politecnico di Milano, Italy

Unité de recherche INRIA Rocquencourt
Domaine de Voluceau, Rocquencourt, BP 105, 78153 Le Chesnay Cedex (France)
Téléphone : +33 1 39 63 55 11 —Télécopie : +33 1 39 63 53 30

Stabilisation flux-upwind de la formulation Discontinuous Petrov–Galerkin avec multiplicateurs de Lagrange pour des problémes de convection-diffusion

Résumé : Dans cette étude nous considérons la méthode duale-primale Discontinuous Petrov-Galerkin (DPG) pour le problème de convection-diffusion en deux dimensions. Comme dans la méthode DPG les variables internes mixtes sont discontinues, nous pouvons effectuer un procédé de condensation statique et ansi parvenir à un schéma discret non-conforme en un seul champ. Pour cette dernière formulation, nous proposons une technique de stabilisation de type flux-upwind adaptée au cas de convection dominante. Le schéma résultant est conservatif et satisfait un principe de maximum discret sous hypothèses standard pour le maillage de calcul. Une analyse de convergence de la méthode est présentée, afin de démontrer une précision à l'ordre un dans une norme discrète H^1. La performance numérique du schéma est validée sur des cas test qui sont caractérisés à l'intérieur et sur la frontière du domaine par des couches limites présentants de très fortes variations.

Mots-clés : méthodes des éléments finis; méthodes mixtes-hybrides; méthodes non-conformes; méthodes Galerkin discontinues, méthodes Petrov-Galerkin; multiplicateurs de Lagrange; éléments finis stabilisées; flux-upwinding; problem de convection-diffusion

Introduction

It is well known that there exist several physical problems (for example, flow in porous media or semiconductor charge transport) where, at the same time, it is desirable to preserve interelement flux continuity and to account for the presence of strongly varying coefficients. The numerical approximation of these kinds of problems can significantly benefit from the use of mixed discretizations. These latter methodologies are well established for the approximation of elliptic problems, but they still lack a robust extension to deal with advective-diffusive problems.

The idea proposed in [18, 17, 16, 31] is to handle the diffusive term with a standard mixed approach and to introduce an upwind technique (or a Riemann solver) to deal with the advective term. Using the terminology introduced in [17], we will denote these approaches as *Upwind Mixed* (UM) methods. The UM methods proposed in the above references were proved to be stable and convergent. Nonetheless, mixed methods may suffer from the computational cost associated with the solution of the corresponding linear algebraic system. Lumping procedures of the stress mass matrix can be designed to eliminate the mixed variable from the system, but these are typically limited to finite element approximations of lowest order [5, 23]. In [31] an hybridized method is proposed that allows for a significant reduction of the computational cost. With this aim, a fractional step algorithm is introduced, that leads to the solution of a sequence of explicit problems. In this context, the hybrid variable is recovered as a purely post-processed quantity, its role merely being that of producing a discontinuous mixed field.

In this paper we propose an approximate formulation of the advective-diffusive problem that is indeed solved as a function of the sole hybrid variable, as in standard mixed methods after static condensation. The method here discussed is based on the Discontinuous Petrov-Galerkin (DPG) formulation, discussed and analyzed in references [6, 11, 12, 10, 7]. The DPG method is a dual-primal hybrid formulation that, after static condensation, reduces to a nonconforming single-field method. It is on this form that we apply an upwinding technique in order to stabilize the discrete scheme, as discussed in Sect. 5. This latter upwind formulation does not require to introduce any secondary partition of the computational domain, as is the case with the upwind-based nonconforming scheme proposed in [25]. Once the problem on the interface variable is solved, we can recover the mixed structure of the method by applying a simple element-by-element post-processing procedure, which provides an approximation of the advective-diffusive flux that is both self-equilibrated and conservative over the computational grid.

The work is organized as follows. In Sect. 1 we introduce the advective-diffusive model. In Sect. 2 we provide the DPG weak formulation of the advective-diffusive equation and its corresponding finite element discretization in the lowest-order case. The static condensation procedure which allows one to derive a nonconforming single-field Galerkin formulation is described in Sect. 3. The construction of the stiffness matrix of this latter method, as well as its stability analysis, are carried out in Sect. 4, where a standard limitation on the Peclèt number is shown to be a sufficient condition to obtain a numerical scheme enjoying a

discrete maximum principle. This latter condition, which can be quite severe in terms of the choice of the mesh size, is overcome in Sect. 5, where an upwind treatment of the convective fluxes across the element boundary is proposed. This, in turn, yields a conservative and monotone nonconforming finite element method. A convergence analysis of the stabilized DPG formulation is carried out in Sect. 6, where it is proved that the discretization error satisfies first-order accuracy measured in a discrete H^1-norm. We illustrate in Sect. 7 the post-processing procedure which allows for an element-by-element recovery of the interface fluxes. Finally, the numerical performance of the proposed method is demonstrated in the concluding Sect. 8, where the scheme is applied to representative benchmark test problems of advection-dominated flows. Some concluding remarks and future work are addressed in Sect. 9.

1 The Advection-diffusion problem

1.1 Mathematical setting of the problem

Let Ω be an open, bounded set of $\mathbb{R}^2$, and let $\Gamma = \partial\Omega$ be the piecewise smooth boundary of Ω, $\mathbf{n}$ being the unit outward normal vector to Γ. With reference to Fig. 1, we set

$$\begin{cases} \Gamma^- = \{x \in \Gamma \,|\, \boldsymbol{b}(x) \cdot \mathbf{n} < 0\}, & \Gamma^+ = \Gamma - \Gamma^-, \\ \Gamma_D^\pm = \Gamma_D \cap \Gamma^\pm, & \Gamma_N^\pm = \Gamma_N \cap \Gamma^\pm, \\ \Gamma_D = \Gamma_D^+ \cup \Gamma_D^-, & \Gamma_N = \Gamma_N^+ \cup \Gamma_N^-, \end{cases}$$

where the subscript D indicates the Dirichlet part of the boundary, while the subscript N indicates the Neumann part of the boundary.

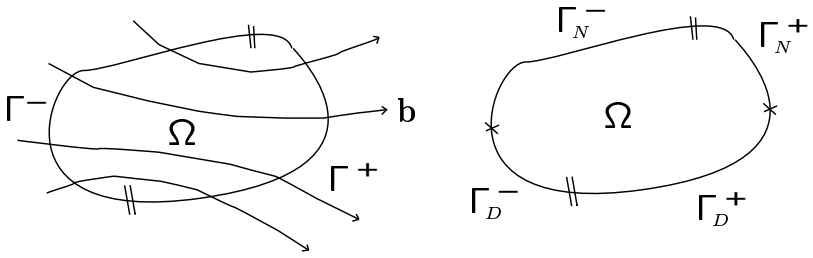

Figure 1: Computational domain and partition of its boundary.

We consider the following model problem

$$\begin{cases} \mathcal{L}_C(u) = f & \text{in } \Omega, \\ u = g_D & \text{on } \Gamma_D, \\ \varepsilon \nabla u \cdot \mathbf{n} - b_n^- u = g_N & \text{on } \Gamma_N, \end{cases} \tag{1}$$

where $\mathcal{L}_C(u) = -\text{div}\,(\varepsilon \nabla u) + \text{div}\,(\boldsymbol{b}u)$ is the linear advection-diffusion operator in *conservative form*, and $\boldsymbol{b}$ is a given advective field with

$$b_n = \boldsymbol{b} \cdot \mathbf{n}, \qquad b_n^+ = \frac{b_n + |b_n|}{2}, \qquad b_n^- = \frac{b_n - |b_n|}{2}.$$

The function f is a given source term, and the boundary data are

$$g_D = \begin{cases} g_D^- & \text{on } \Gamma_D^-, \\ g_D^+ & \text{on } \Gamma_D^+, \end{cases} \qquad g_N = \begin{cases} g_N^- & \text{on } \Gamma_N^-, \\ g_N^+ & \text{on } \Gamma_N^+. \end{cases}$$

On the inflow portion of the Neumann boundary the total advective-diffusive flux $(\varepsilon \nabla u - \boldsymbol{b}u) \cdot \mathbf{n}$ is prescribed, while on the outflow portion of the Neumann boundary only the diffusive flux $\varepsilon \nabla u \cdot \mathbf{n}$ is prescribed (see Fig. 1).

The conservative form of the linear advection-diffusion operator $\mathcal{L}_C(u)$ represents a simplified model for the compressible Navier-Stokes equations or the Drift-Diffusion transport model in semiconductor device simulation [22, 19].

Under the assumption that $\boldsymbol{b}$ is solenoidal, and that ε and $\boldsymbol{b}$ are sufficiently smooth functions, the conservative form is completely equivalent to the semi-conservative form of the advection-diffusion operator, *i.e.*

$$\mathcal{L}_C(u) = \mathcal{L}_{SC}(u) = -\text{div}\,(\varepsilon \nabla u) + \boldsymbol{b} \cdot \nabla u. \tag{2}$$

The semi-conservative form represents a simplified model for incompressible fluid-dynamics problems in the presence of a variable viscosity [21].

1.2 Primal weak formulation of the advection-diffusion problem

Let $S \subset \mathbb{R}^2$ be an open bounded set with Lipschitz continuous boundary ∂S. For a non-negative integer m, let $H^m(S)$ be the usual m-th order Sobolev space defined over S and equipped with the norm and seminorm

$$||v||_{m,S} = \left(\sum_{|\alpha| \le m} ||D^\alpha v||_{0,S}^2 \right)^{1/2}, \quad |v|_{m,\Omega} = \left(\sum_{|\alpha| = m} ||D^\alpha v||_{0,S}^2 \right)^{1/2},$$

where $D^\alpha v$ is the distributional derivative of order α of a function v and $||\cdot||_{0,S}$ is the norm in $L^2(S)$. We refer to [1, 20] for definitions and properties of Sobolev spaces. We set

$$V = \left\{ v \in H^1(\Omega) \mid v = 0 \text{ on } \Gamma_D \right\}$$

and we define the bilinear form on $V \times V$ as

$$B(u,v) = \int_\Omega (\varepsilon \nabla u - \mathbf{b}u) \cdot \nabla v \; dx \qquad u, v \in V,$$

where we assume that $\varepsilon \in L^\infty(\Omega)$ and $\mathbf{b} \in (W^{1,\infty}(\Omega))^2$. The *weak primal problem* associated with (1) reads: find $u_0 \in V$ such that

$$B(u_0,v) + \int_{\Gamma_N^+} u_0 \, v \, \mathbf{b}\cdot\mathbf{n} \, ds = (f,v)_{0,\Omega} - B(u_D,v) - \int_{\Gamma_N^+} u_D \, v \, \mathbf{b}\cdot\mathbf{n} \, ds + \int_{\Gamma_N^-} v \, g_N^- \, ds + \int_{\Gamma_N^+} v \, g_N^+ \, ds \qquad \forall v \in V, \tag{3}$$

where $u_D \in H^1(\Omega)$ is a function such that $u_D = g_D$ on Γ_D in the sense of traces, $f \in L^2(\Omega)$, g_D and g_N belong to appropriate trace spaces on Γ_D and Γ_N and $(\cdot,\cdot)_{0,\Omega}$ denotes the L^2 inner product. The primal problem (3) has a unique solution under the condition that there exists a positive constant α such that

$$\frac{\varepsilon_0}{C_\Omega^2} - \|\mathrm{div}\, \mathbf{b}\|_{\infty,\Omega} > \alpha > 0,$$

where C_Ω is the Poincaré constant and $\varepsilon_0 = \inf_{x \in \Omega} \varepsilon(x) > 0$.

2 DPG Formulation of the Advection-Diffusion Problem

Before introducing the DPG formulation, we need some additional notation. Let $\mathcal{T}_h$ be a given triangulation of Ω into triangles K, with area $|K|$, boundary ∂K and outward unit normal vector $\mathbf{n}_{\partial K}$ on ∂K. We denote by h_K the diameter of K and by ρ_K the diameter of the largest ball inscribed in K. We assume henceforth that $\mathcal{T}_h$ is regular [13], i.e. that there exists a positive constant κ independent of h such that

$$\frac{h_K}{\rho_K} \leq \kappa \qquad \forall K \in \mathcal{T}_h. \tag{4}$$

It is immediate to check that the previous inequality implies also that

$$\frac{h_K^2}{|K|} \leq \frac{4\,\kappa^2}{\pi} \qquad \forall K \in \mathcal{T}_h. \tag{5}$$

Let $\mathcal{E}_h$ denote the set of edges in $\mathcal{T}_h$ and for each edge $\mathbf{e} \in \mathcal{E}_h$, let $|\mathbf{e}|$ represent the edge length. Moreover, let $\partial K_{int} = \partial K \cap \Omega$, $\partial K_D = \partial K \cap \Gamma_D$, $\partial K_{N^+} = \partial K \cap \Gamma_N^+$ and $\partial K_{N^-} = \partial K \cap \Gamma_N^-$, so that for each $K \in \mathcal{T}_h$, $\partial K = \partial K_{int} \cup \partial K_D \cup \partial K_{N^+} \cup \partial K_{N^-}$.

Proceeding as in standard Discontinuous Galerkin formulations [3], we introduce the mixed variable $\boldsymbol{\sigma} = \varepsilon \nabla u$ associated with the diffusive flux and we formally integrate by parts both equations in (1), yielding the following one-element Discontinuous Petrov-Galerkin (DPG) weak formulation of the model advection-diffusion problem:

find $(u, \boldsymbol{\sigma}, (\lambda, \mu))$ such that for all $K \in \mathcal{T}_h$ and for all $(\boldsymbol{\tau}, v)$, we have

$$\left\{ \begin{array}{l} \displaystyle\int_K \varepsilon^{-1} \boldsymbol{\sigma} \cdot \boldsymbol{\tau} \; dx + \int_K u \, \mathrm{div}\, \boldsymbol{\tau} \; dx - \int_{\partial K \setminus \partial K_D} \lambda \, \boldsymbol{\tau} \cdot \mathbf{n}_{\partial K} \; ds = \int_{\partial K_D} \mathcal{P} u_D \, \boldsymbol{\tau} \cdot \mathbf{n}_{\partial K} \; ds \qquad \forall \boldsymbol{\tau}, \\ \displaystyle\int_K (\boldsymbol{\sigma} - \boldsymbol{b}\, u) \cdot \nabla v \; dx - \int_{\partial K_{int}} (\mu - \boldsymbol{b}\, \lambda \cdot \mathbf{n}_{\partial K}) \, v \; ds - \int_{\partial K_D} \mu \, v \; ds + \int_{\partial K_{N+}} \boldsymbol{b}\, \lambda \cdot \mathbf{n}_{\partial K} \, v \; ds \\ \qquad \displaystyle = \int_K f \, v \; dx - \int_{\partial K_D} \boldsymbol{b}\, \mathcal{P} u_D \cdot \mathbf{n}_{\partial K} \, v \; ds + \int_{\partial K_{N-}} \mathcal{P} g_N^- \, v \; ds + \int_{\partial K_{N+}} \mathcal{P} g_N^+ \, v \; ds \qquad \forall v, \end{array} \right. \tag{6}$$

where $\mathcal{P}$ is the L^2-projection over the constant functions, $\boldsymbol{\tau}$ and v are smooth functions inside each element $K \in \mathcal{T}_h$ and

$$\boldsymbol{\sigma}|_K = (\varepsilon \nabla u)|_K, \qquad \lambda = u|_{\partial K}, \qquad \mu = \boldsymbol{\sigma} \cdot \mathbf{n}|_{\partial K} \qquad \forall K \in \mathcal{T}_h.$$

Two different kinds of variables are present in the DPG one-element formulation (6): the internal – discontinuous – variables u and $\boldsymbol{\sigma}$ that express (in a weak sense) the constitutive and the equilibrium relations at the interior of each element, and the boundary variables λ and μ that represent (in weak sense) the trace of the internal variables on the element boundary. Formulation (6) is of Petrov-Galerkin type since the functional spaces for the shape and test functions are different. We refer to [6] and [12] for a presentation of the DPG method and its convergence and stability analysis in the case of diffusion problems.

In view of the finite element approximation of (6), we introduce some notation for the polynomial spaces and the projection operators. For a given nonnegative integer k, we denote by $\mathbb{P}_k(K)$ the space of all polynomials of degree $\leq k$ on K, and by $R_k(\partial K)$ the space of all functions defined over the boundary ∂K of K whose restrictions to any side $\mathbf{e} \in \partial K$ are polynomials of degree $\leq k$. Functions in $R_k(\partial K)$ can be discontinuous at the vertices of ∂K. Moreover, denoting by $\mathbf{x}$ the position vector in $\mathbb{R}^2$, we let

$$\mathbb{RT}_k(K) = (\mathbb{P}_k(K))^2 \oplus \mathbf{x}\mathbb{P}_k(K) \qquad \forall K \in \mathcal{T}_h, \tag{7}$$

be the Raviart-Thomas (RT) finite element space of degree k [27]. In the case $k = 0$, we define $\mathbb{RT}_0(\mathcal{T}_h) \subset H(\mathrm{div}\,, \Omega)$ the space of RT polynomials of lowest degree having continuous normal component across each internal edge of $\mathcal{E}_h$. Finally, we denote by $\mathbb{P}_1^{CR}(K)$ the Crouzeix-Raviart space of linear polynomials over the element K [15] and by $\mathbb{P}_1^{CR}(\mathcal{T}_h)$ the space of affine functions that are continuous at the midpoints of each edge of $\mathcal{E}_h$ and whose restriction on each element K belongs to $\mathbb{P}_1^{CR}(K)$.

We define the projection operator P_K from $L^2(K)$ onto $\mathbb{P}_0(K)$ such that, for all $v \in L^2(K)$, we have

$$\int_K (P_K v - v)\, p_0 \; dx = 0 \qquad \forall p_0 \in \mathbb{P}_0(K), \qquad \forall K \in \mathcal{T}_h. \tag{8}$$

The operator P_K associates a scalar function with its mean integral value over K. From the operator P_K, for all $v \in L^2(\Omega)$, we construct the global operator P_h as

$$P_h v|_K = P_K v \qquad \forall K \in \mathcal{T}_h.$$

Then, we define the projection operator $\Pi_K^{RT} : H(\mathrm{div};\mathrm{K}) \to \mathbb{RT}_0(K)$ satisfying the orthogonality relation

$$\int_{\partial K} \left(\Pi_K^{RT}\boldsymbol{\tau} - \boldsymbol{\tau}\right) \cdot \mathbf{n}_{\partial K}\, r_0 \; ds = 0 \qquad \forall r_0 \in R_0(\partial K), \qquad \forall K \in \mathcal{T}_h. \tag{9}$$

The operator Π_K^{RT} associates a vector function with its fluxes across the boundary ∂K. From the operator Π_K^{RT}, for all $\boldsymbol{\tau} \in H(\mathrm{div}\,;\Omega)$, we construct the global operator $\Pi_h^{RT} : H(\mathrm{div}\,;\Omega) \to \mathbb{RT}(\mathcal{T}_h)$ as

$$\Pi_h^{RT}\boldsymbol{\tau}|_K = \Pi_K^{RT}\boldsymbol{\tau} \qquad \forall K \in \mathcal{T}_h.$$

Finally, given a function $\boldsymbol{\tau} \in H(\mathrm{div}\,;\Omega)$ such that $\mathrm{div}\;\boldsymbol{\tau} = 0$, we define $\widetilde{\boldsymbol{\tau}} = \Pi_h^{RT}\boldsymbol{\tau}$ with $\mathrm{div}\;\Pi_K^{RT}\boldsymbol{\tau} = 0$ for each $K \in \mathcal{T}_h$. Function $\widetilde{\boldsymbol{\tau}}$ is a piecewise constant vector over $\mathcal{T}_h$, with a continuous normal component across each internal edge of $\mathcal{E}_h$.

The finite element discretization of (6) using the lowest-order DPG method reads:

find $(u_h, \boldsymbol{\sigma}_h, (\lambda_h, \mu_h)) \in (U_h \times \Sigma_h \times (\Lambda_h \times M_h))$ such that for all $K \in \mathcal{T}_h$ we have

$$\begin{cases} \displaystyle\int_K \varepsilon^{-1}\boldsymbol{\sigma}_h \cdot \boldsymbol{\tau}_h \; dx + \int_K u_h \,\mathrm{div}\,\boldsymbol{\tau}_h \; dx - \int_{\partial K \setminus \partial K_D} \lambda_h\, \boldsymbol{\tau}_h \cdot \mathbf{n}_{\partial K} \; ds = \int_{\partial K_D} \mathcal{P}u_D\, \boldsymbol{\tau}_h \cdot \mathbf{n}_{\partial K} \; ds \qquad \forall \boldsymbol{\tau}_h \in Q_h(K), \\ \displaystyle\int_K (\boldsymbol{\sigma}_h - \widetilde{\boldsymbol{b}}\, u_h) \cdot \nabla v_h \; dx - \int_{\partial K_{int}} (\mu_h - \widetilde{\boldsymbol{b}}\, \lambda_h \cdot \mathbf{n}_{\partial K})\, v_h \; ds - \int_{\partial K_D} \mu_h\, v_h \; ds + \int_{\partial K_{N^+}} \widetilde{\boldsymbol{b}}\, \lambda_h \cdot \mathbf{n}_{\partial K}\, v_h \; ds \\ \displaystyle = \int_K f\, v_h \; dx - \int_{\partial K_D} \widetilde{\boldsymbol{b}}\, \mathcal{P}u_D \cdot \mathbf{n}_{\partial K}\, v_h \; ds + \int_{\partial K_{N^-}} \mathcal{P}g_N^-\, v_h \; ds + \int_{\partial K_{N^+}} \mathcal{P}g_N^+\, v_h \; ds \qquad \forall v_h \in W_h(K). \end{cases} \tag{10}$$

The discrete local trial spaces are defined as

$$\begin{aligned} &U_h(K) = \mathbb{P}_0(K), && \Sigma_h = (\mathbb{P}_0(K))^2 && \forall K \in \mathcal{T}_h, \\ &\Lambda_h(\partial K) = \gamma_{0,K}(\mathbb{P}_1^{CR}(K)), && M_h(\partial K) = R_0(\partial K) && \forall K \in \mathcal{T}_h, \end{aligned} \tag{11}$$

where $\gamma_{0,K} : H^1(K) \to H^{1/2}(\partial K)$ is the linear continuous operator which associates with a function defined on K its trace on ∂K.

The discrete local test spaces are defined as

$$Q_h(K) = \mathbb{RT}_0(K), \qquad W_h(K) = \mathbb{P}_1(K) \qquad \forall K \in \mathcal{T}_h. \tag{12}$$

The global finite element spaces of the DPG method of lowest degree are constructed as

$$U_h = \{u_h|_K \in U_h(K)\,\forall K \in \mathcal{T}_h\}, \qquad \Sigma_h = \{\boldsymbol{\sigma}_h|_K \in \Sigma_h(K)\,\forall K \in \mathcal{T}_h\},$$

$$\Lambda_h = \{\lambda_h \in \gamma_{0,\mathcal{T}_h}(\mathbb{P}_1^{CR}(\mathcal{T}_h)) \,|\, \lambda_h = \mathcal{P}g_D \text{ at the midpoints of } \Gamma_D\}, \qquad M_h = \{\mu_h|_K \in M_h(\partial K)\,\forall K \in \mathcal{T}_h\},$$

$$Q_h = \{\boldsymbol{\tau}_h|_K \in Q_h(K)\,\forall K \in \mathcal{T}_h\}, \qquad W_h = \{v_h|_K \in W_h(K)\,\forall K \in \mathcal{T}_h\},$$

where $\gamma_{0,\mathcal{T}_h} : \prod_{K\in\mathcal{T}_h} H^1(K) \to \prod_{K\in\mathcal{T}_h} H^{1/2}(\partial K)$ is the linear continuous operator which associates with a piecewise smooth function defined on $\mathcal{T}_h$ its trace on $\mathcal{E}_h$ in such a way that this trace is continuous at the midpoint of each internal edge.

3 The single-field problem associated with the DPG formulation

In this section, we describe the static condensation procedure carried out on an element-by-element basis, which allows for the elimination of the internal variables u_h^K, $\boldsymbol{\sigma}_h^K$ and also of the boundary variable $\mu_h^{\partial K}$ from the DPG formulation (10) in favor of the boundary variable λ_h. As a matter of fact, from the definition of the space Λ_h, one can notice that the Lagrange multiplier λ_h represents the trace on the edges of the triangulation of a nonconforming finite element basis. Exploiting this feature, we will end up with a *nonconforming single-field scheme* in the sole unknown λ_h, which makes the formulation computationally convenient (see [2] and [14] for a discussion on the procedures to perform static condensation on hybridized mixed methods for elliptic problems).

For ease of presentation, we assume that $\partial K \cap \Gamma = \emptyset$, i.e. that the element is in the interior of the domain Ω, with a straightforward extension of the elimination procedure to the case where also the boundary conditions in (1) are accounted for. Moreover, we indicate from now on, with a slight abuse of notation, with the symbol λ_h the element itself of $\mathbb{P}_1^{CR}(K)$ (and not only its trace on ∂K). Integrating by parts in $(1)_1$ the boundary term, then gives

$$\int_K (\varepsilon^{-1}\boldsymbol{\sigma}_h - \nabla\lambda_h)\cdot\boldsymbol{\tau}_h\; dx + \int_K (u_h - \lambda_h)\,\mathrm{div}\,\boldsymbol{\tau}_h\; dx = 0 \qquad \forall \boldsymbol{\tau}_h \in Q_h(K). \tag{13}$$

Taking first $\boldsymbol{\tau}_h \in (\mathbb{P}_0(K))^2$ in (13), yields

$$\boldsymbol{\sigma}_h^K = \widetilde{\varepsilon}^K\,\nabla\lambda_h^K \qquad \forall K \in \mathcal{T}_h, \tag{14}$$

where

$$\widetilde{\varepsilon}^K = \left(\int_K \varepsilon^{-1}(x)\; dx\,/|K|\right)^{-1} \tag{15}$$

is the *harmonic average* of the diffusion coefficient ε over K.

Taking then $\boldsymbol{\tau}_h = (x,y)^T$ in (13), and replacing the function $(\varepsilon^K)_K^{-1}$ with its average $(\widehat{\varepsilon}^K)^{-1}$, yields

$$u_h^K = \int_K \lambda_h \, dx \, / |K| = P_K \lambda_h = \frac{1}{3} \sum_{i=1}^{3} \lambda_i \qquad \forall K \in \mathcal{T}_h, \tag{16}$$

where λ_i are the nodal values of λ_h at the midpoints of each edge of K. Notice that (14) and (16) imply

$$\int_K u_h \, dx = \int_K \lambda_h \, dx \qquad \forall K \in \mathcal{T}_h. \tag{17}$$

Let us now consider equation $(10)_2$ and take $v_h \in \mathbb{P}_1^{CR}(\mathcal{T}_h)$. Summing $(10)_2$ over the elements of the triangulation, and using the fact that μ_h is constant on each edge of $\mathcal{E}_h$, automatically eliminates this latter variable when each element boundary contribution $\int_{\partial K} \mu_h \, v_h \, ds$ is assembled together over all the internal edges. Then, substituting (14) and (16) into $(10)_2$, incorporating the boundary conditions and using (17), yields the discrete problem:

find $\lambda_h \in V_{h,g_D}$ such that

$$\begin{aligned} &\sum_{K \in \mathcal{T}_h} \left\{ \int_K (\widehat{\varepsilon} \nabla \lambda_h - \widetilde{\boldsymbol{b}} \lambda_h) \cdot \nabla v_h \, dx - \int_{\partial K_{int} \cup \partial K_{N+}} \widetilde{\boldsymbol{b}} \, \lambda_h \cdot \mathbf{n}_{\partial K} \, v_h \, ds \right\} \\ &= \sum_{K \in \mathcal{T}_h} \left\{ \int_K f \, v_h \, dx + \int_{\partial K_{N-}} \mathcal{P} g_N^- \, v_h \, ds + \int_{\partial K_{N+}} \mathcal{P} g_N^+ \, v_h \, ds \right\} \qquad \forall v_h \in V_{h,0}, \end{aligned}$$

where, for a given function $\xi \in L^2(\Gamma_D)$, we have defined

$$V_{h,\xi} = \{ v_h \in \mathbb{P}_1^{CR}(\mathcal{T}_h) \, | \, v_h = \mathcal{P}\xi \, \text{at the midpoints of the edges of} \, \Gamma_D \}.$$

Finally, using the fact that $\boldsymbol{b}$ is divergence-free, an integration by parts of the convective term $-\int_K \widetilde{\boldsymbol{b}} \lambda_h \cdot \nabla v_h \, dx$ in the previous equation gives the following single–field form of the DPG method (10):

find $\lambda_h \in V_{h,g_D}$ such that

$$\begin{aligned} &\sum_{K \in \mathcal{T}_h} \left\{ \int_K (\widehat{\varepsilon} \nabla \lambda_h \cdot \nabla v_h + \widetilde{\boldsymbol{b}} \cdot \nabla \lambda_h \, v_h) \, dx - \int\limits_{\partial K_{N-}} \widetilde{\boldsymbol{b}} \, \lambda_h \cdot \mathbf{n}_{\partial K} \, v_h \, ds \right\} \\ &= \sum_{K \in \mathcal{T}_h} \left\{ \int_K f \, v_h \, dx + \int\limits_{\partial K_{N-}} \mathcal{P} g_N^- \, v_h \, ds + \int\limits_{\partial K_{N+}} \mathcal{P} g_N^+ \, v_h \, ds \right\} \qquad \forall v_h \in V_{h,0}. \end{aligned} \tag{18}$$

The Galerkin problem (18) *can be interpreted as the nonconforming finite element approximation of the advection-diffusion boundary-value problem* (1) in semi-conservative form

and with harmonic averaging of the diffusion coefficient ε. Notice that, the solution λ_h of (18) differs from the solution λ_h^{NC} of the standard nonconforming approximation of problem (1) in semi-conservative form, which would in fact read:

find $\lambda_h^{NC} \in V_{h,g_D}$ such that

$$\begin{aligned}\sum_{K\in\mathcal{T}_h}\left\{\int_K(\overline{\varepsilon}\nabla\lambda_h^{NC}\cdot\nabla v_h+\widetilde{\boldsymbol{b}}\cdot\nabla\lambda_h^{NC}\,v_h)\;dx-\int\limits_{\partial K_{N^-}}\widetilde{\boldsymbol{b}}\,\lambda_h^{NC}\cdot\mathbf{n}_{\partial K}\,v_h\;ds\right\}\\ =\sum_{K\in\mathcal{T}_h}\left\{\int_K f\,v_h\;dx+\int\limits_{\partial K_{N^-}}\mathcal{P}g_N^-\,v_h\;ds+\int\limits_{\partial K_{N^+}}\mathcal{P}g_N^+\,v_h\;ds\right\}\qquad\forall v_h\in V_{h,0},\end{aligned}\tag{19}$$

where

$$\overline{\varepsilon}^K=\int_K\varepsilon(x)\;dx\,/|K|$$

is the usual (note, not the harmonic) average of ε on K. It is well known that in the presence of rough (or strongly varying) coefficients, the use of harmonic averaging provides superior accuracy and stability than standard averaging (see [4] for the 1D case, and [9, 23, 12] for applications in 2D).

4 The plain DPG discrete formulation

In this section we explicitly construct the finite element equations that arise from the nonconforming DPG formulation (18) and analyze the properties of the stiffness matrix $\boldsymbol{K}$ of the associated linear algebraic system

$$\boldsymbol{K}\,\boldsymbol{\lambda}=\boldsymbol{f},\tag{20}$$

where $\boldsymbol{\lambda}$ and $\boldsymbol{f}$ are the vectors of nodal unknowns and the right-hand side, respectively. For ease of presentation, we consider the special case $\Gamma=\Gamma_D$ (i.e., nonhomogeneous Dirichlet boundary conditions in (1)).

Let us denote by `Nel` the number of triangles in $\mathcal{T}_h$ and by `Ned` the number of edges in $\mathcal{E}_h$, with `Ni` internal edges and `Nb` boundary edges. Correspondingly, we denote by $\widetilde{\varphi}_i$, $i=1,\ldots,$`Ni`, the global basis function of the space V_h. The function $\widetilde{\varphi}_i$ has its support on the two triangles K_1, K_2 that share the common edge $\mathbf{e}_i$ (see Fig. 2 for the notation), and satisfies the following property

$$\int_{e_p}\widetilde{\varphi}_i\;ds=\delta_{ip}|\mathbf{e}_i|,\qquad p=i,j,k,l,m.\tag{21}$$

We let henceforth $S_i=K_1\cup K_2$ denote the support of $\widetilde{\varphi}_i$, and write

$$\lambda_h=\sum_{r=1}^{\texttt{Ned}}\lambda_r\,\widetilde{\varphi}_r,$$

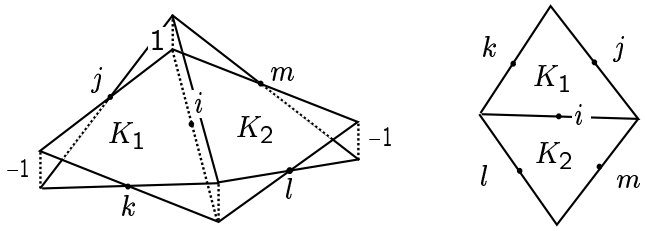

Figure 2: Basis function $\widetilde{\varphi}_i$ for V_h (left) and notation (right).

where λ_r is the nodal value of λ_h at the midpoint of edge $\mathbf{e}_r$, with $\lambda_r = \mathcal{P}u_D|_{\mathbf{e}_r}$ for any $\mathbf{e}_r \in \Gamma_D$. For each element $K \in S_i$, we assume a counterclockwise orientation over ∂K and denote by $\boldsymbol{b}_r$ the value of $\boldsymbol{b}$ at the midpoint of edge $\mathbf{e}_r \in \partial K$. Then, we set $\Phi_r^{\partial K} = \widetilde{\boldsymbol{b}}_r^K \cdot \mathbf{n}_{r,\partial K}|\mathbf{e}_r| = \boldsymbol{b}_r \cdot \mathbf{n}_{r,\partial K}|\mathbf{e}_r|$ to be the convective flux across edge $\mathbf{e}_r \in \partial K$ with outward unit normal vector $\mathbf{n}_{r,\partial K}$, such that

$$\sum_{\mathbf{e}_r \in \partial K} \Phi_r^{\partial K} = 0. \tag{22}$$

Taking now $v_h = \widetilde{\varphi}_i$ in (18) and using the two-dimensional midpoint rule to compute the right-hand side contribution $\int_{S_i} f\, v_h\, dx$, we obtain the following finite element equation associated with each edge $\mathbf{e}_i \in \mathcal{E}_i$

$$\sum_{p=i,j,k,l,m} K_{ip}\lambda_p = f_i, \qquad i = 1, \ldots, \mathtt{Ni}. \tag{23}$$

The stiffness matrix coefficients K_{ip}, $p = i, j, k, l, m$, are given by

$$K_{ip} = K_{ip}^{\mathrm{diff}} + K_{ip}^{\mathrm{adv}}, \qquad p = i, j, k, l, m, \tag{24}$$

where

$$K_{ip}^{\mathrm{diff}} = \begin{cases} \widetilde{\varepsilon}^K \dfrac{\mathbf{e}_i \cdot \mathbf{e}_p}{|K|} & \text{if } p \neq i, \\ \left(\dfrac{\widetilde{\varepsilon}^{K_1}}{|K_1|} + \dfrac{\widetilde{\varepsilon}^{K_2}}{|K_2|} \right) \mathbf{e}_i \cdot \mathbf{e}_i & \text{if } p = i, \end{cases} \tag{25}$$

$$K_{ip}^{\mathrm{adv}} = \begin{cases} \dfrac{1}{3}\Phi_p^{\partial K} & \text{if } p \neq i, \\ 0 & \text{if } p = i, \end{cases}$$

are the separate contributions due to the diffusive and convective fluxes across each edge $\mathbf{e}_p$, $p = i, j, k, l, m$. Notice that the i-th row of $\boldsymbol{K}$ has (a priori) five nonzero matrix entries. The i-th component of the right-hand side $\boldsymbol{f}$ is given by

$$f_i = \frac{1}{3}\left(f^{K_1}|K_1| + f^{K_2}|K_2|\right), \qquad i = 1, \ldots, \mathtt{Ni}, \tag{26}$$

where f^{K_1}, f^{K_2} are the values of the source function f at the centroids of K_1, K_2. From (24)-(25), and using (22), it is immediate to check that:

(P1) $\widetilde{\boldsymbol{b}} \in \mathbb{RT}_0(\mathcal{T}_h)$ implies that $\Phi_i^{\partial K_1} + \Phi_i^{\partial K_2} = 0$, which explains why only a diffusive contribution is present in the diagonal matrix entry K_{ii}, unlike the case of the off-diagonal matrix terms K_{ip}, $p \neq i$;

(P2) when $\partial S_i \cap \Gamma = \emptyset$

$$\sum_{p=i,j,k,l,m} K_{ip}^{\mathrm{diff}} = \sum_{p=i,j,k,l,m} K_{ip}^{\mathrm{adv}} = 0, \tag{27}$$

that is, both the *net* diffusive and advective fluxes across ∂S_i are zero when λ_h is constant and $\mathbf{b}$ is divergence-free;

(P3) when $\mathbf{e}_p \in \partial S_i$ is such that $\mathbf{e}_p \in \Gamma$, $p \neq i$, then the associated nodal unknown λ_p is eliminated from the system by setting $\lambda_p = \mathcal{P} u_D|_{\mathbf{e}_p}$.

Having characterized the basic properties of the stiffness matrix of the DPG nonconforming formulation (18), we are in a position to address the stability analysis of the discretization scheme. In particular, we aim at obtaining sufficient conditions for $\boldsymbol{K}$ to be an M-matrix. This, in turn, allows one to ensure that the solution of system (20) satisfies a *Discrete Maximum Principle* (DMP). In other words, when $f = 0$, the discrete solution $\boldsymbol{\lambda}$ is nonnegative over $\overline{\Omega}$ and attains its maximum value on the boundary Γ.

To start with, let us recall the definition of an M-matrix [29].

Definition 4.1 *A matrix $\mathbf{A} \in \mathbb{R}^{n\times n}$, with $n \geq 1$, is an M-matrix if it is invertible, its entries A_{ij} satisfy $A_{ij} \leq 0$, $i \neq j$, and $(\mathbf{A}^{-1})_{ij} \geq 0$, $i, j = 1, \ldots, n$.*

The diagonal entries of an M-matrix are positive and an M-matrix is inverse-monotone, i.e., $\mathbf{Ax} \leq \mathbf{Ay}$ implies that $\mathbf{x} \leq \mathbf{y}$, $\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. The next result provides a useful sufficient condition which allows to check whether a given matrix is an M-matrix ([29], Thm. 3.1, p.202).

Theorem 4.1 *Let $\mathbf{A} \in \mathbb{R}^{n\times n}$ be an irreducible matrix such that*

$$\begin{aligned}
&\text{(a)} \quad A_{ii} > 0, \quad A_{ij} \leq 0, \qquad i, j = 1, \ldots, n,\\
&\text{(b)} \quad |A_{ii}| \geq \sum_{j\neq i} |A_{ij}|, \qquad i = 1, \ldots, n,\\
&\text{(c)} \quad |A_{kk}| > \sum_{j\neq k} |A_{kj}|, \qquad \textit{for at least one row index } k \in [1, \ldots, n].
\end{aligned} \tag{28}$$

Then, $\mathbf{A}$ *is an M-matrix with* $(\mathbf{A}^{-1})_{ij} > 0$, $i, j = 1, \ldots, n$.

Notice that a matrix fulfilling the assumptions in Theorem 4.1 satisfies a DMP.

From now on, we assume that the triangulation $\mathcal{T}_h$ is of weakly acute type (i.e., all the angles of the triangles are less or equal to $\pi/2$). Therefore, $\mathbf{e}_i \cdot \mathbf{e}_p \leq 0$, $p \neq i$, so that the diffusion contribution to the off-diagonal matrix entries is nonpositive. As a consequence, since $K_{ii} > 0$ for every $i = 1, \ldots, n$, it is easy to check that if $\boldsymbol{b} = (0,0)^T$ (i.e. when (1) is a purely diffusive boundary-value problem) then stiffness matrix $\boldsymbol{K}$ satisfies Theorem 4.1, and, therefore, a DMP ([29], p. 203). When $\boldsymbol{b} \neq (0,0)^T$, then the request that the off-diagonal entries of $\boldsymbol{K}$ are nonpositive is satisfied if

$$\frac{\widetilde{\boldsymbol{b}}_p^K \cdot \mathbf{n}_p}{3\widetilde{\varepsilon}^K} \frac{|K|}{|\mathbf{e}_i| \cos(\theta_{ip})} \leq 1, \qquad p \in \{j,k,l,m\}, \qquad K \in \{K_1, K_2\},$$

where $p = \{j,k\}$ if $K = K_1$ or $p = \{l,m\}$ if $K = K_2$, and θ_{ip} is the angle between the edges $\mathbf{e}_i$ and $\mathbf{e}_p$. Denoting by $h_r > 0$ the height relative to edge $\mathbf{e}_r$, noting that $|K| = h_p|\mathbf{e}_p|/2$, and letting $\gamma_{ip} = |\mathbf{e}_i| \cos(\theta_{ip})/|\mathbf{e}_p|$, with $\gamma_{ip} \in [0,1)$, we immediately obtain that the off-diagonal entries of $\boldsymbol{K}$ are nonpositive if

$$\frac{|\widetilde{\boldsymbol{b}}_p^K \cdot \mathbf{n}_p| \, h_p}{6\widetilde{\varepsilon}^K} \leq 1, \qquad p \in \{j,k,l,m\}, \qquad K \in \{K_1, K_2\}. \tag{29}$$

Observing that h_p can be interpreted as the length of K in the direction of the convective flow across edge $\mathbf{e}_p$, we can define the *Péclet number associated with edge* $\mathbf{e}_p$ as

$$\alpha_p = \frac{|\widetilde{\boldsymbol{b}}_p^K \cdot \mathbf{n}_p| \, h_p}{6\widetilde{\varepsilon}^K}, \qquad p \in \{j,k,l,m\}, \qquad K \in \{K_1, K_2\},$$

and conclude that, as usual in the finite element approximation of advection-diffusion problems, the DPG nonconforming formulation (18) is stable (i.e. the associated stiffness matrix is an M-matrix) if the local Péclet number is less than 1. Clearly, condition (29) can be too restrictive on the mesh size when the flow is advection-dominated. For this reason, we introduce in the next section a suitable stabilization of the plain DPG method (18), which allows to compute a reasonably accurate solution even on a coarse mesh $\mathcal{T}_h$.

5 The Stabilized DPG Formulation

In this section we introduce a stabilization technique for the nonconforming DPG formulation (18). The technique is based on a suitable treatment of the convective term in the finite element equation (23) associated with each internal edge $\mathbf{e}_i \in \mathcal{E}_h$.

From relation (22) and (P1), it follows that the DPG nonconforming scheme is *conservative* with respect to both the single element K (K_1 or K_2) and to the control volume $K_1 \cup K_2$. At the same time, it is clearly seen that (22) prevents *all* the convective fluxes

$\Phi_p^{\partial K}$, $p \neq i$, to have the same (negative) sign, except in the (trivial) case $\boldsymbol{b} = (0,0)^T$. This eventually prevents the stiffness matrix from being an M-matrix for any value of the Péclet number α_p.

A possible remedy is suggested by the *flux-balance interpretation* of the nonconforming DPG formulation discussed in Sect.4, and proceeds as follows. For every $K \in \mathcal{T}_h$, we let

$$\partial K^{\rm in} = \bigcup_r \{\mathbf{e}_r \in \partial K \,|\, \boldsymbol{b}_r \cdot \mathbf{n}_{r,\partial K} \leq 0\}\,, \qquad \partial K^{\rm out} = \bigcup_r \{\mathbf{e}_r \in \partial K \,|\, \boldsymbol{b}_r \cdot \mathbf{n}_{r,\partial K} > 0\}$$

denote the *inflow* and *outflow* boundaries of K, respectively. Moreover, we associate with every edge internal $\mathbf{e}_i$ an absolute unit normal vector $\mathbf{n}_i$ by setting, for instance, $\mathbf{n}_i = \mathbf{n}_{\partial K_1}|_{\mathbf{e}_i}$ (cf. Fig. 3). Accordingly, we define the *upstream* triangle $K_i^{\rm upstrm}$ associated with edge $\mathbf{e}_i$ as

$$K_i^{\rm upstrm} = \begin{cases} K_1 & \text{if } \boldsymbol{b}_i \cdot \mathbf{n}_i > 0, \\ K_2 & \text{if } \boldsymbol{b}_i \cdot \mathbf{n}_i < 0. \end{cases} \tag{30}$$

The definition of $K_i^{\rm upstrm}$ in the special case $\boldsymbol{b}_i \cdot \mathbf{n}_i = 0$ will be addressed at the end of this section. Then, we introduce the following *min-max* treatment of the edge convective fluxes

$$K_{ip}^{\rm adv,upw} = \begin{cases} \min\left(0, \dfrac{1}{3}\Phi_p^{\partial K}\right) & \text{if } p \neq i, \\ \displaystyle\sum_{p \neq i} \max\left(0, \frac{1}{3}\Phi_p^{\partial K}\right) & \text{if } p = i, \end{cases} \tag{31}$$

where the sum is taken over all the edges of ∂S_i. The flux-upwind stabilized DPG finite element equation associated with $\mathbf{e}_i \in \mathcal{E}_h \setminus \Gamma$ reads

$$\sum_{p=i,j,k,l,m} K_{ip}^{\rm stab} \lambda_p = f_i^{\rm stab}, \qquad i = 1, \ldots, \mathtt{Ni}, \tag{32}$$

where the stabilized stiffness matrix coefficients $K_{ip}^{\rm stab}$, $p = i, j, k, l, m$, are given by

$$K_{ip}^{\rm stab} = K_{ip}^{\rm diff} + K_{ip}^{\rm adv,upw}, \qquad p = i, j, k, l, m. \tag{33}$$

Definition (31) obeys the classical *upwind* philosophy. Precisely, relation $(31)_1$ amounts to setting to zero the convective flux associated with an edge, whenever this latter edge belongs to $\partial K^{\rm out}$. This procedure is equivalent to *subtracting* some edge contributions to the whole net convective flux balance across ∂S_i. Accordingly, relation $(31)_2$ allows to redistribute the missing convective fluxes, in order to satisfy *at the same time* the net flux conservation property (27) *and* the request of positive diagonal matrix entries, as stated by Thm. 4.1.

The redistribution procedure of the outflow convective fluxes is schematically illustrated in Fig. 3 (left) and results into a *flux upwind-modified* DPG nonconforming scheme, which

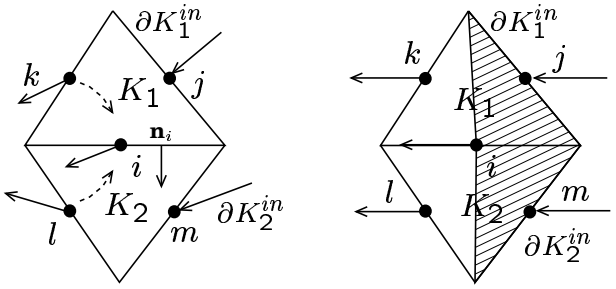

Figure 3: Left: redistribution procedure of the outflow convective fluxes. Right: definition of K_i^{upstrm} when $\boldsymbol{b}_i \cdot \mathbf{n}_i = 0$ (the upstream triangle is the shaded area in the figure).

is conservative with respect to both the single element K (K_1 or K_2) and to the control volume $K_1 \cup K_2$, as was the case with the plain formulation (18). Moreover, it is immediate to check that the stiffness matrix $\boldsymbol{K}$ of the flux upwind-modified DPG method satisfies the conditions in Thm. 4.1, which allows one to conclude that the scheme satisfies a DMP irrespectively of the Pèclet number α_p.

An important issue in upwind finite element procedures is related to the appropriate treatment of the integral $\int_{S_i} f\,\widetilde{\varphi}_i\,dx$. It is in fact well-known that upwind methods can produce a physically uncorrect solution in the presence of a non-zero source term f [24, 8]. For this reason, we have devised the following "upwind" rule for the evaluation of the above integral if $|\boldsymbol{b}_i| \neq 0$:

$$\int_{S_i} f\,\widetilde{\varphi}_i\,dx \simeq f_i^{\mathrm{stab}} = \begin{cases} \dfrac{f^{K_i^{\mathrm{upstrm}}}|K_i^{\mathrm{upstrm}}|}{3} & \text{if } \boldsymbol{b}_i \cdot \mathbf{n}_i \neq 0, \\ \dfrac{1}{2}\left(f^{K_1}\dfrac{|K_1|}{3} + f^{K_2}\dfrac{|K_2|}{3}\right) & \text{if } \boldsymbol{b}_i \cdot \mathbf{n}_i = 0, \end{cases} \tag{34}$$

while we obviously set $f_i^{\mathrm{stab}} = f_i$ if $|\boldsymbol{b}_i| = 0$. Fig. 3 (left) helps in providing a simple and immediate interpretation for this upwind rule, by observing that in the case $\Phi_i \neq 0$ the integral in (34) is computed *only* over K_i^{upstrm} according to a full-upwind treatment (in the present case, we have $K_i^{\mathrm{upstrm}} = K_1$), while in the case $\Phi_i = 0$ the upstream triangle is defined as the union of the two triangles having as sides the half of edge $\mathbf{e}_i$, the two edges lying on the inflow boundary of K_1 and K_2 and the two segments joining node i and the two vertices of K_1 and K_2 opposite to i (see Fig. 3, right). The beneficial effect of the use of (34), in contrast to the application of the two-dimensional midpoint rule as done in (23) to compute the approximate right-hand side, will be examined in Sect. 8.3. A suitably different, although equivalent, interpretation will be provided in Sect. 6 for the upwind relations (31) and (34). This will cast the DPG stabilized formulation into a more conventional upwind framework, allowing a simpler error analysis of the method.

6 Convergence Analysis

In this section we provide a convergence analysis of the stabilized DPG formulation introduced in Sect. 5. Our approach follows the guideline of [25].

6.1 Bilinear and linear forms

Throughout this section, for ease of presentation, we assume that homogeneous Dirichlet boundary conditions are enforced in (1), i.e., $\Gamma_D = \Gamma$, $\Gamma_N = \emptyset$ and $g_D = 0$, and for brevity we shall write V_h instead of $V_{h,0}$. Since the discrete DPG upwind-stabilized formulation is of nonconforming type, we have that the finite element space $V_h \not\subset H_0^1(\Omega)$. However, functions in V_h satisfy the following *compatibility conditions*:

(C.1) For any $K_1, K_2 \in \mathcal{T}_h$ with $\mathbf{e} = \partial K_1 \cap \partial K_2$, we have

$$\int_{\mathbf{e}} (v_h^{K_1} - v_h^{K_2})\, ds = 0 \qquad \forall v_h \in V_h.$$

(C.2) For any $K \in \mathcal{T}_h$, we have

$$\int_{\partial K \cap \Gamma} v_h^K\, ds = 0 \qquad \forall v_h \in V_h.$$

Let us introduce the following quantities

$$||v_h||_{1,h} = \left(\sum_{K \in \mathcal{T}_h} ||v_h||_{1,K}^2 \right)^{1/2}, \quad |v_h|_{1,h} = \left(\sum_{K \in \mathcal{T}_h} |v_h|_{1,K}^2 \right)^{1/2} \qquad \forall v_h \in V_h.$$

It can be shown that functions in V_h satisfy the following discrete Poincarè inequality ([32], Proposition 4.13)

$$\|v_h\|_{0,\Omega} \leq C_P |v_h|_{1,h} \qquad \forall v_h \in V_h, \tag{35}$$

where $C_P = C_P(\Omega)$ is a positive constant. As a consequence, conditions (C.1)-(C.2) and (35) imply that $|\cdot|_{1,h}$ is a norm over the space V_h (equivalent to $\|\cdot\|_{1,h}$).

We associate with each edge mid-point M_i, $i = 1, \ldots, \texttt{Ned}$, the following index sets

$$\mathcal{I}_i = \left\{ \text{the pair of neighboring triangles } K^j \in \mathcal{T}_h,\ j \in [1, \texttt{Nel}],\ \text{that share the edge } \mathbf{e}_i \right\},$$

$$\mathcal{J}_i = \left\{ j \neq i;\ M_j \text{ is the mid-point of the side of a triangle having } M_i \text{ as another one} \right\},$$

and denote any edge that is adjacent to $\mathbf{e}_i$ by

$$\Gamma_{is} = \left\{ s \neq i;\ \mathbf{e}_s \in \mathcal{E}_h \mid \mathbf{e}_s \text{ shares a vertex with } \mathbf{e}_i \right\}.$$

Then, we define the following bilinear forms on $V_h \times V_h$ for all $w_h, v_h \in V_h$

$$a(w_h, v_h) = \sum_{K \in \mathcal{T}_h} \int_K \widetilde{\varepsilon} \nabla w_h \cdot \nabla v_h \, dx \,, \qquad b(w_h, v_h) = \sum_{K \in \mathcal{T}_h} \int_K (\widetilde{\mathbf{b}} \cdot \nabla w_h) \, v_h \, dx \,, \tag{36}$$

and the linear form on V_h for all $v_h \in V_h$

$$F(v_h) = \frac{1}{3} \sum_{i=1}^{\mathtt{Ni}} \sum_{j \in \mathcal{I}_i} \left(\int_{K^j} f_h \, dx \right) v_h(M_i), \tag{37}$$

where $f_h|_K = P_K f$ for all $K \in \mathcal{T}_h$. Finally, we define for any $w_h, v_h \in V_h$ the bilinear form

$$\mathcal{B}(w_h, v_h) = a(w_h, v_h) + b(w_h, v_h),$$

in such a way that the plain DPG formulation (18) applied to the advection-diffusion problem reads:

find $\lambda_h \in V_h$ such that

$$\mathcal{B}(\lambda_h, v_h) = F(v_h) \qquad \forall v_h \in V_h. \tag{38}$$

Recalling (P1) of Sect. 4, it can be easily verified that

$$b(w_h, v_h) = \frac{1}{3} \sum_{i=1}^{\mathtt{Ni}} \sum_{j \in \mathcal{J}_i} \left(\int_{\Gamma_{ij}} \mathbf{b}_{ij} \cdot \mathbf{n}_{ij} \, w_h \, ds \right) v_h(M_i) \qquad \forall v_h \in V_h, \tag{39}$$

where $\mathbf{b}_{ij}$ is the value of $\mathbf{b}$ at the midpoint of edge Γ_{ij} and $\mathbf{n}_{ij}$ is the unit outward normal vector along $\Gamma_{ij} \cap \partial K^j$, $j \in \mathcal{J}_i$. We define the following modification of $b(\cdot, \cdot)$

$$b_h(w_h, v_h) = \frac{1}{3} \sum_{i=1}^{\mathtt{Ni}} \sum_{j \in \mathcal{J}_i} \left(\int_{\Gamma_{ij}} \mathbf{b}_{ij} \cdot \mathbf{n}_{ij} \, w_h^{ij} \, ds \right) v_h(M_i) \qquad \forall w_h, v_h \in V_h, \tag{40}$$

where

$$\begin{aligned} w_h^{ij} &= \alpha^{ij} w_h(M_i) + (1 - \alpha^{ij}) w_h(M_j), \\ \alpha^{ij} &= \begin{cases} 1 & \text{if } \displaystyle\int_{\Gamma^{ij}} \mathbf{b}_{ij} \cdot \mathbf{n}_j \, ds \geq 0, \\ 0 & \text{otherwise.} \end{cases} \end{aligned} \tag{41}$$

Moreover, we define the modified form of $F(v_h)$ as follows

$$F_h(v_h) = \frac{1}{3} \sum_{i=1}^{\mathtt{Ni}} \sum_{j \in \mathcal{I}_i} \left(\int_{K^j} f_h^j \, dx \right) v_h(M_i) \qquad \forall v_h \in V_h, \tag{42}$$

where we set

$$f_h^j = \beta^j P_{K_1} f + (1-\beta^j) P_{K_2} f \qquad \text{if } |\mathbf{b}_i| \neq 0,$$
$$\beta^j = \begin{cases} 1 & \text{if } \mathbf{b}_i \cdot \mathbf{n}_i > 0, \\ 0 & \text{if } \mathbf{b}_i \cdot \mathbf{n}_i < 0, \\ \dfrac{1}{2} & \text{if } \mathbf{b}_i \cdot \mathbf{n}_i = 0, \end{cases} \tag{43}$$

and we set $f_h^j = f_h$ if $|\mathbf{b}_i| = 0$. Finally, we define our modified form of $\mathcal{B}(\cdot,\cdot)$ as

$$\mathcal{B}_h(w_h, v_h) = a(w_h, v_h) + b_h(w_h, v_h) \qquad \forall v_h, w_h \in V_h,$$

in such a way that the DPG upwind-stabilized formulation introduced in Sect. 5 reads:

find $\lambda_h^* \in V_h$ such that

$$\mathcal{B}_h(\lambda_h^*, v_h) = F_h(v_h) \qquad \forall v_h \in V_h. \tag{44}$$

6.2 Consistency analysis for $b_h(\cdot,\cdot)$

The following result shows that the modified bilinear form $b_h(\cdot,\cdot)$ is *consistent* with $b(\cdot,\cdot)$.

Theorem 6.1 *Under the regularity assumption* (4)*, there exists a positive constant C_b independent of h such that*

$$|b(w_h, v_h) - b_h(w_h, v_h)| \leq C_b\, h\, |w_h|_{1,h} |v_h|_{1,h} \qquad \forall\, w_h, v_h \in V_h. \tag{45}$$

Proof.
From (39) and (40), we find that

$$|b(w_h, v_h) - b_h(w_h, v_h)| = \frac{1}{3} \sum_{i=1}^{\mathtt{Ni}} \sum_{j \in \mathcal{J}_i} \int_{\Gamma_{ij}} \mathbf{b}_{ij} \cdot \mathbf{n}_{ij}\, (w_h(s) - w_h^{ij}(s))\, v_h(M_i)\, ds$$
$$= \frac{1}{6} \sum_{i=1}^{\mathtt{Ni}} \sum_{j \in \mathcal{J}_i} \int_{\Gamma_{ij}} \mathbf{b}_{ij} \cdot \mathbf{n}_{ij}\, [\alpha^{ij}(w_h(s) - w_h(M_i)) + (1-\alpha^{ij})(w_h(s) - w_h(M_j))](v_h(M_i) - v_h(M_j))\, ds$$
$$= \frac{1}{6} \sum_{i=1}^{\mathtt{Ni}} \sum_{j \in \mathcal{J}_i} \int_{\Gamma_{ij}} \mathbf{b}_{ij} \cdot \mathbf{n}_{ij}\, \alpha^{ij} (w_h(s) - w_h(M_i))(v_h(M_i) - v_h(M_j))\, ds\,.$$

The last equality follows by noticing that $\int_{\Gamma_{ij}} (w_h(s) - w_h(M_j))\, ds = 0$. Using the mean value theorem and noting that ∇w_h is constant over K, we have

$$|w_h(\mathbf{x}) - w_h(M_i)| = |\nabla w_h \cdot (\mathbf{x} - \mathbf{x}_i)| \leq |\nabla w_h| h_K = \frac{\|\nabla w_h\|_{0,K} h_K}{|K|^{1/2}},$$

from which, using (5) and the fact that $\alpha^{ij} \leq 1$, we obtain

$$\left| \int_{\Gamma_{ij}} \mathbf{b}_{ij} \cdot \mathbf{n}_{ij}\, \alpha^{ij} (w_h(s) - w_h(M_i))(v_h(M_i) - v_h(M_j))\, ds \right| \leq \frac{4\,\kappa^2}{\pi} \, |\mathbf{b}_{ij} \cdot \mathbf{n}_{ij}| \, \|\nabla w_h\|_{0,K} \|\nabla v_h\|_{0,K} \, h_K.$$

The consistency estimate (45) then immediately follows summing over $\mathcal{E}_h$ and using the discrete Cauchy-Schwarz inequality, with C_b depending on κ and $\|\mathbf{b} \cdot \mathbf{n}\|_{\infty,\mathcal{E}_h}$ but not on h.

6.3 Coercivity analysis for $B_h(\cdot,\cdot)$

We have the following result.

Theorem 6.2 *Let* $0 < \varepsilon_0 = \inf_{\mathbf{x}\in\Omega} \varepsilon(\mathbf{x})$. *Then, we have*

$$\mathcal{B}_h(w_h, w_h) \geq \varepsilon_0 \, |w_h|^2_{1,h} \qquad \forall w_h \in V_h. \tag{46}$$

As a consequence, assuming α^{ij} *defined as in* (41), *problem* (44) *has a unique solution* $\lambda_h^* \in V_h$.

Proof.
In order to prove the coercivity of $\mathcal{B}_h(\cdot,\cdot)$, we show that the modified bilinear form $b_h(\cdot,\cdot)$ is *coercive* with respect to the norm $|\cdot|_{1,h}$. For all $w_h \in V_h$ we have

$$b_h(w_h, w_h) = \frac{1}{3} \sum_{i=1}^{\mathtt{Ni}} \sum_{j\in\mathcal{J}_i} \left(\int_{\Gamma_{ij}} \mathbf{b}_{ij} \cdot \mathbf{n}_{ij}\, w_h^{ij}\, ds \right) w_h(M_i)$$

$$= (\text{since } \mathbf{n}_{ij} = -\mathbf{n}_{ji}) = \frac{1}{6} \sum_{i=1}^{\mathtt{Ni}} \sum_{j\in\mathcal{J}_i} \int_{\Gamma_{ij}} \mathbf{b}_{ij} \cdot \mathbf{n}_{ij}\, (w_h^{ij}\, v_h(M_i) - w_h^{ij}\, v_h(M_j))\, ds$$

$$= \frac{1}{6} \sum_{i=1}^{\mathtt{Ni}} \sum_{j\in\mathcal{J}_i} \int_{\Gamma_{ij}} \mathbf{b}_{ij} \cdot \mathbf{n}_{ij}\, \left(\alpha^{ij} w_h^2(M_i) - \alpha^{ji} w_h^2(M_j) + w_h(M_i) w_h(M_j)(1 - \alpha^{ij} - 1 + \alpha^{ji}) \right) ds$$

$$= (\text{since } \alpha^{ji} = -\alpha^{ij}) = \frac{1}{6} \sum_{i=1}^{\mathtt{Ni}} \sum_{j\in\mathcal{J}_i} \int_{\Gamma_{ij}} \mathbf{b}_{ij} \cdot \mathbf{n}_{ij}\, \left(\alpha^{ij} w_h^2(M_i) + \alpha^{ij} w_h^2(M_j) - 2\alpha^{ij} w_h(M_i) w_h(M_j) \right) ds$$

$$= \frac{1}{6} \sum_{i=1}^{\mathtt{Ni}} \sum_{j\in\mathcal{J}_i} \int_{\Gamma_{ij}} \mathbf{b}_{ij} \cdot \mathbf{n}_{ij}\, \alpha^{ij}\, (w_h(M_i) - w_h(M_j))^2\, ds \geq 0,$$

from which (46) immediately follows. Using this latter result and (45), we can also show that the plain DPG formulation is stable *only* for small values of the Péclet number, while the stabilized DPG method is stable *irrespectively* of the size of the Péclet number. We have

$$\mathcal{B}(w_h, w_h) = \mathcal{B}_h(w_h, w_h) + b(w_h, w_h) - b_h(w_h, w_h) \geq (\varepsilon_0 - C_b h) |w_h|^2_{1,h}$$

from which it follows that, taking a sufficiently small value $h^* > 0$ of the mesh size, there exists a suitable constant $\varepsilon^* > 0$ such that, for all $h \leq h^*$, we have

$$\mathcal{B}(w_h, w_h) \geq \varepsilon^* \, |w_h|^2_{1,h} \qquad \forall w_h \in V_h.$$

Using the Lax-Milgram Lemma and (35), we immediately get the a priori estimate

$$|\lambda_h|_{1,h} \leq \frac{C_P \, \|f\|_{0,\Omega}}{\varepsilon^*}. \tag{47}$$

6.4 Consistency analysis for $F_h(\cdot,\cdot)$

The following result shows that the modified linear form $F_h(\cdot)$ (as well as $F(\cdot)$) are *consistent* with the *exact linear form* $(f,\cdot)_{0,\Omega}$.

Theorem 6.3 *Assuming that $f \in W^{1,\infty}(\Omega)$, there exists a positive constant C_f independent of h such that*

$$|F(v_h) - F_h(v_h)| \leq C_f \, h \, |v_h|_{1,h} \qquad \forall \, v_h \in V_h. \tag{48}$$

Proof.
For all $v_h \in V_h$ we have

$$F(v_h) - F_h(v_h) = \frac{1}{3} \sum_{i=1}^{\mathtt{Ni}} \sum_{j \in \mathcal{J}_i} \left(\int_{K^j} (f_h - f_h^j) \, dx \right) v_h(M_i)$$

$$= \frac{1}{3} \sum_{i=1}^{\mathtt{Ni}} \sum_{j \in \mathcal{J}_i} \left(\int_{K^j} (f_h - f(\mathbf{x})) \, dx \right) v_h(M_i) + \frac{1}{3} \sum_{i=1}^{\mathtt{Ni}} \sum_{j \in \mathcal{J}_i} \left(\int_{K^j} (f(\mathbf{x}) - f_h^j) \, dx \right) v_h(M_i).$$

Using standard interpolation estimates ([26], Sect. 6.2.3), noticing that $\beta^j \leq 1$ in the definition of f_h^j and using the discrete Cauchy-Schwarz inequality, we obtain

$$|F(v_h) - F_h(v_h)| \leq C \, h \, |\Omega| \, |f|_{1,\infty,\Omega} \, \|v_h\|_{0,\Omega},$$

where $C > 0$ is a constant independent of h. The estimate (48) then immediately follows using (35).

6.5 Error estimates

In this section we prove optimal error estimates for the nonconforming flux-upwind stabilized DPG method. We start proving that the solutions of the two DPG problems (plain and stabilized) are close in the discrete H^1-norm $|\cdot|_{1,h}$.

Theorem 6.4 *Under the assumptions of Theorems 6.1, 6.2 and 6.3, we have*

$$|\lambda_h - \lambda_h^*|_{1,h} \leq \frac{1}{\varepsilon_0} \left(C_b \frac{\|f\|_{0,\Omega}}{\varepsilon^*} + C_f \right) h, \tag{49}$$

where λ_h and λ_h^* denote respectively the solutions of problems (38) *and* (44).

Proof.
We have

$$\mathcal{B}_h(\lambda_h - \lambda_h^*, v_h) = b_h(\lambda_h, v_h) - b(\lambda_h, v_h) + F(v_h) - F_h(v_h) \qquad \forall v_h \in V_h,$$

from which we immediately obtain

$$|\mathcal{B}_h(\lambda_h - \lambda_h^*, v_h)| \leq C_b h |\lambda_h|_{1,h} |v_h|_{1,h} + C_f h |v_h|_{1,h} \qquad \forall v_h \in V_h.$$

Taking $v_h = \lambda_h - \lambda_h^*$ we get

$$\varepsilon_0 |\lambda_h - \lambda_h^*|_{1,h}^2 \leq C_b h |\lambda_h|_{1,h} |\lambda_h - \lambda_h^*|_{1,h} + C_f h |\lambda_h - \lambda_h^*|_{1,h},$$

from which, using (47), we immediately get inequality (49).

The error analysis of the stabilized DPG formulation can now be easily concluded using the convergence result of [27] for the primal-hybrid nonconforming finite element approximation λ_h^{NC} solution of (19), estimating the difference $|\lambda_h^{NC} - \lambda_h|_{1,h}$ (by means of Strang Lemma, [23]) and then using the triangle inequality.

Theorem 6.5 *Under the assumptions of Theorems 6.1, 6.2 and 6.3, and assuming also that* $u \in H^2(\Omega) \cap H_0^1(\Omega)$*, we have*

$$|u - \lambda_h^*|_{1,h} \leq (C_1 \, |u|_{2,\Omega} + C_2 + C_3) \, h, \tag{50}$$

where C_1*,* C_2 *and* C_3 *are positive constants independent of* h *and depending only on* $|\Omega|$*,* ε*,* $\mathbf{b}$ *and* f*.*

7 Recovery of interface fluxes

Once the nonconforming single field problem (in its plain or stabilized form) has been solved, the convective flux is immediately available, while the diffusive flux μ_h can computed using the element-by-element recovery procedure illustrated in the following.

Going back to equation $(10)_2$, the diffusive flux is obtained by solving on each $K \in \mathcal{T}_h$ such that $\partial K \cap \Gamma = \emptyset$ the following local subproblem of dimension 3

$$\int_{\partial K} \mu_h \, v_h \, ds = \int_K (\widetilde{\varepsilon} \nabla \lambda_h - \widetilde{\mathbf{b}} \lambda_h) \cdot \nabla v_h \, dx + \int_{\partial K} \widetilde{\mathbf{b}} \cdot \mathbf{n}_{\partial K} \lambda_h \, v_h \, ds - \int_K f \, v_h \, dx \qquad \forall v_h \in W_h(K). \tag{51}$$

On the Neumann boundaries we have

$$\mu_h = \begin{cases} \mathcal{P} g_N^+ & \text{on every } \mathbf{e} \in \Gamma_N^+ \\ \mathcal{P} g_N^- + \widetilde{\mathbf{b}}^K \cdot \mathbf{n}_{\partial K} \mathcal{P} \lambda_h & \text{on every } \mathbf{e} \in \Gamma_N^-. \end{cases}$$

Notice that standard linear Lagrangian nodal-based test functions are used in (51), this being the same procedure adopted in primal-hybrid formulations implemented as nonconforming finite elements (see [28], p. 691).

Problem (51) can be written in matrix form as

$$\boldsymbol{M}_K \, \boldsymbol{\mu}^K = \mathbf{r}^K$$

where

$$\boldsymbol{M}_K = \frac{1}{2} \begin{pmatrix} 0 & |\mathbf{e}_2| & |\mathbf{e}_3| \\ |\mathbf{e}_1| & 0 & |\mathbf{e}_3| \\ |\mathbf{e}_1| & |\mathbf{e}_2| & 0 \end{pmatrix}, \quad \boldsymbol{\mu}^K = (\mu_1, \mu_2, \mu_3)^T, \quad \mathbf{r}^K = (\mathcal{A}_K^{diff} + \mathcal{A}_K^{conv})\boldsymbol{\lambda}^K - \mathbf{f}^K,$$

having introduced a local counterclockwise numbering of the edges of ∂K, denoted by $\mathbf{e}_i$, $i = 1, 2, 3$, and where

$$\boldsymbol{A}_K^{\text{diff}} = -\frac{\widetilde{\varepsilon}^K}{2\,|K|} \begin{pmatrix} \mathbf{e}_1 \cdot \mathbf{e}_1 & \mathbf{e}_1 \cdot \mathbf{e}_2 & \mathbf{e}_1 \cdot \mathbf{e}_3 \\ \mathbf{e}_1 \cdot \mathbf{e}_2 & \mathbf{e}_2 \cdot \mathbf{e}_2 & \mathbf{e}_2 \cdot \mathbf{e}_3 \\ \mathbf{e}_1 \cdot \mathbf{e}_3 & \mathbf{e}_2 \cdot \mathbf{e}_3 & \mathbf{e}_3 \cdot \mathbf{e}_3 \end{pmatrix}, \quad \boldsymbol{A}_K^{\text{conv}} = \frac{1}{3} \begin{pmatrix} \Phi_1^{\partial K} & \Phi_2^{\partial K} & \Phi_3^{\partial K} \\ \Phi_1^{\partial K} & \Phi_2^{\partial K} & \Phi_3^{\partial K} \\ \Phi_1^{\partial K} & \Phi_2^{\partial K} & \Phi_3^{\partial K} \end{pmatrix},$$

$$\mathbf{f}^K = \frac{|K|}{3} P_K f \, (1, 1, 1)^T, \qquad \boldsymbol{\lambda}^K = (\lambda_1, \lambda_2, \lambda_3)^T.$$

Once μ_h is available on each edge $\mathbf{e}_p \in \mathcal{E}_h$ (denoted by μ_p, $p = 1, \ldots, \texttt{Ned}$), it is possible to compute the approximate advective-diffusive edge flux $J_p \equiv (\mu_p - \boldsymbol{b}_p \cdot \mathbf{n}_p \lambda_p)\,|\mathbf{e}_p|$ and then the corresponding approximate advective-diffusive vector field $\mathbf{J}_h = \sum_{\mathbf{e}_p \in \mathcal{E}_h} J_p \, \boldsymbol{\tau}_p(\mathbf{x})$ over $\mathcal{T}_h$ using the $\mathbb{RT}_0(\mathcal{T}_h)$ finite element space, where $\boldsymbol{\tau}_p$ is the RT basis function associated with edge $\mathbf{e}_p$. The discrete advective-diffusive field $\mathbf{J}_h$ computed by the DPG formulation (18) enjoys the conservation property at each element $K \in \mathcal{T}_h$. An example of the flux recovery procedure will be given in Sect. 8.1, while we refer to [12] for a comparison with other kinds of flux-recovery procedures proposed in a standard primal-based Galerkin framework.

8 Numerical Results

To test the numerical performance of the upwind–stabilized DPG method discussed in Sect. 5, we solve several benchmark test problems for advective-diffusive flows, both on structured and unstructured meshes, characterized by the presence of steep interior and boundary layers.

8.1 Test case nr. 1: the Smith and Hutton test problem

We consider the classical Smith and Hutton benchmark model problem, with $f = 0$. In this test case, a fluid enters the lower left edge of the rectangle $\Omega = [-1, 1] \times [0, 1]$ and exits at

the lower right edge of the domain, where a homogeneous boundary condition is enforced on the diffusive flux. On the remaining sides of the rectangle, Dirichlet boundary conditions are prescribed so that the total advective-diffusive normal flux is zero (see Fig. 4). Precisely, we set

$$\boldsymbol{b} = (2y(1-x^2), -2x(1-y^2))^T,$$

and

$$\begin{cases} u(x,y) = \begin{cases} 1+\tanh(10(2x+1)) & \text{on } \Gamma_D^- = \{(x,y)\in\Gamma \mid x\in[-1,0],\, y=0\}, \\ 0 & \text{on } \Gamma\setminus(\Gamma_D^-\cup\Gamma_N^+), \end{cases} \\ \dfrac{\partial u(x,y)}{\partial n} = 0 \qquad \text{on } \Gamma_N^+. \end{cases}$$

Numerical computations have been performed on a structured uniform triangulation with

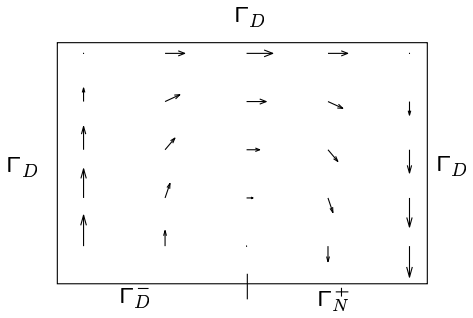

Figure 4: Computational domain and prescribed convective field **b** for the Smith and Hutton test case.

40 subdivisions in both x and y directions, corresponding to $h_x = 1/20$ and $h_y = 1/40$, respectively.

In Figs. 5 and 6 we show the numerical results in the case $\varepsilon = 10^{-6}$, corresponding to a nondimensional Peclèt number $Pe = (h_x||\mathbf{b}||_{\infty,\Omega})/(2\,\varepsilon)$ equal to $5\cdot 10^4$. Fig. 5 displays the surface plot of λ_h^*. A nodally continuous interpolation of the nonconforming finite element solution is employed for graphical purposes. The stabilizing effects of the flux-upwind procedure are clearly visible (right), in contrast with the severe oscillations arising in the non-stabilized case (left). Fig. 6 (left) shows the contour lines of the computed solution, with no appreciable numerical dissipation in the crosswind direction, as expected in this quasi-hyperbolic problem. In Fig. 6 (right) the profile of the solution along the inlet/outlet boundary of the domain is illustrated. No oscillations arise in the computed profile, which is in good agreement with other results in the literature.

Finally, Fig. 7 shows the vector plot of the approximate advective-diffusive vector field $\mathbf{J}_h$, reconstructed using the recovery procedure discussed in Sect. 7. Two values of the diffusion coefficient have been used in the numerical experiments, namely $\varepsilon = 10^{-1}$ (left) and $\varepsilon = 10^{-6}$ (right), in order to better emphasize the (different) role played by the diffusive flux in the computation of $\mathbf{J}_h$. In both cases, an accurate and smooth representation of the advection-diffusion field is achieved, with continuous interelement fluxes over $\mathcal{T}_h$.

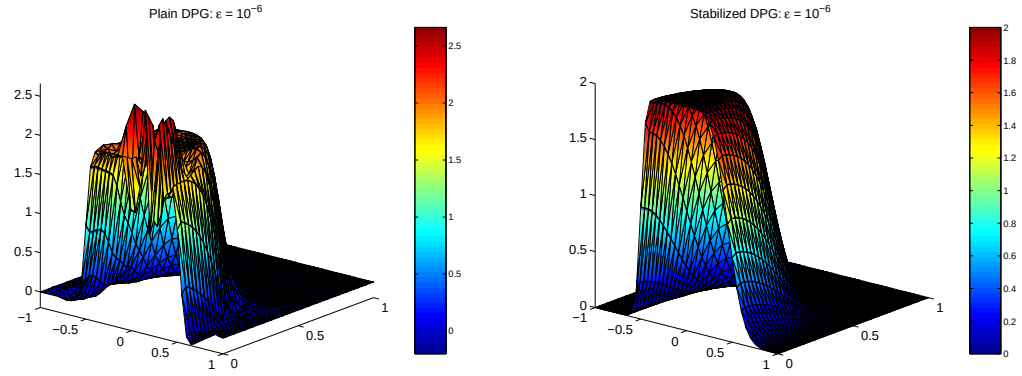

Figure 5: Surface plot of λ_h^*. Left: plain DPG formulation, right: stabilized DPG formulation.

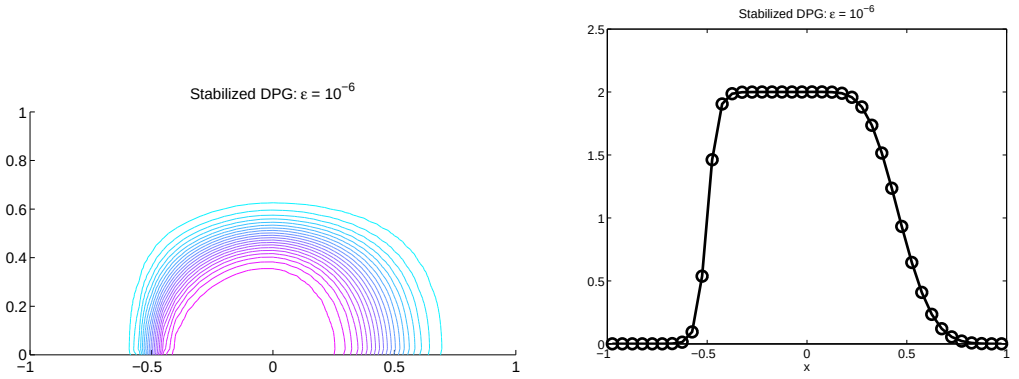

Figure 6: Contour lines (left) and profile of λ_h^* along the inflow-outflow boundary (right).

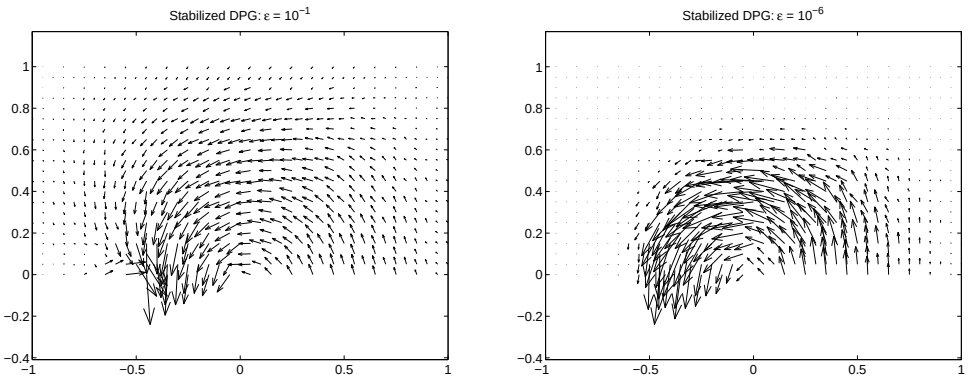

Figure 7: Vector plots of $\mathbf{J}_h$. Left: $\varepsilon = 10^{-1}$, right: $\varepsilon = 10^{-6}$.

8.2 Test case nr.2: advective transport of a discontinuity in the boundary data

The domain Ω is the unit square, where we set $f = 0, \boldsymbol{b} = (\cos(\theta), \sin(\theta))^T$ with $\theta = \tan^{-1}(3)$, and prescribe the Dirichlet boundary conditions

$$\begin{cases} u(x,y) = 1, & \text{for } x = 0, \quad y < 1 \text{ and } x < 1/3, \quad y = 0, \\ u(x,y) = 0, & \text{elsewhere.} \end{cases}$$

The presence of a discontinuity in the boundary data, together with a small value of the viscosity ε, gives rise to an almost-hyperbolic transport problem along the characteristic direction of $\mathbf{b}$. The corresponding solution in this latter case is very close to a discontinuous function, jumping from the value 0 to the value 1 along the line $y = 3x - 1$, with a steep outflow boundary layer along $x = 1$, due to the abrupt change in the boundary data from the (transported) value 1 to the (imposed) value 0. The results are shown in Fig. 8 for $\varepsilon = 10^{-9}$ and with mesh discretization parameter $h = 0.05$. The solution computed by the stabilized DPG method is again unaffected by spurious oscillation, and the internal layer is well approximated, without introducing an excessive smearing in the crosswind direction (i.e., the direction orthogonal to $\mathbf{b}$).

8.3 Test case nr.3: flow with a non-zero source term

The computational domain Ω is again the unit square, where we prescribe $\boldsymbol{b} = (1, 0)^T$, $f = 1$ and homogeneous Dirichlet boundary conditions, in such a way that the solution is a bubble function with an outflow ("hyperbolic") boundary layer along $x = 1$, the width of the layer becoming stronger as the viscosity gets smaller, and two "parabolic" boundary layers along $y = 0$ and $y = 1$ [29].

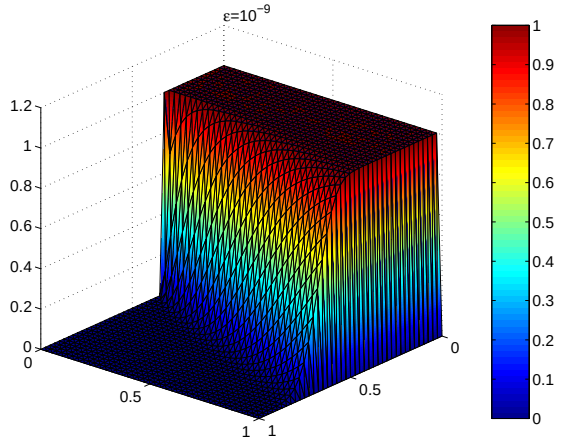

Figure 8: Solution for the discontinuity transport test case, using a structured grid.

In Fig. 9 and Fig. 10 we show the computed solution λ_h (after a suitable re-interpolation over the space of continuous piecewise linear functions for graphical purposes), for different values of ε, using structured ($h = 0.08$) and unstructured meshes ($h = 0.05$), respectively. The results show the ability of the scheme in capturing the steep outflow layer in the solution without introducing neither any spurious oscillation, nor any appreciable amount of extra-viscosity along the direction of the flow. Morever, notice how, in the case of a structured grid, the method can handle without difficulties the case of a convective field aligned with the mesh itself (corresponding to the special case $\mathbf{b} \cdot \mathbf{n}_i = 0$ discussed in Sect. 5).

Concerning this latter aspect, it is interesting to investigate the role played by the "upwind" quadrature rule (34) in the performance of the stabilized DPG method. With this aim, assume to consider the case of a uniform grid of Friedrichs–Keller type with mesh size $h = 1/N$ ([29], p. 206). This is equivalent to constructing two sequences of one–dimensional parallel grids, the first (finer, identified by the label (A)) grid with mesh size equal to $h/2$, the second (coarser, identified by the label (B)) grid with mesh size equal to h. Then, it is easy to check that the use of exact integration of the right-hand side would produce the following discrete solution in the hyperbolic limit (which amounts to assuming $\varepsilon = 0$ in (1))

$$\begin{cases} \lambda_i^A = \lambda_{i-1}^A + h, & i = 1, \ldots, 2N-1, \\ \lambda_0^A = 0, & \\ \lambda_i^B = \dfrac{1}{2}(\lambda_{2i-1}^A + \lambda_{2(i-1)}^A) + \dfrac{h}{2}, & i = 1, \ldots, N, \\ \lambda_0^B = 0. & \end{cases}$$

The two sequences of discrete nodal values of λ_h are aligned on the straight line $y = 2x$, while the exact solution in the hyperbolic limit is $y = x$. In the same case, the stabilized DPG scheme with an upwind treatment of the source function f computes the following discrete solution

$$\begin{cases} \lambda_i^A = \lambda_{i-1}^A + \dfrac{h}{2}, & i = 1, \ldots, N, \\ \lambda_0^A = 0, & \\ \lambda_i^B = \dfrac{1}{2}(\lambda_{2i-1}^A + \lambda_{2(i-1)}^A) + \dfrac{h}{4}, & i = 1, \ldots, N, \\ \lambda_0^B = 0. & \end{cases}$$

In this case, the two sequences of discrete nodal values of λ_h are correctly aligned on the straight line $y = x$. This result can be interpreted as the exact fulfillment of the *patch-test* for consistency proof of the nonconforming formulation [30].

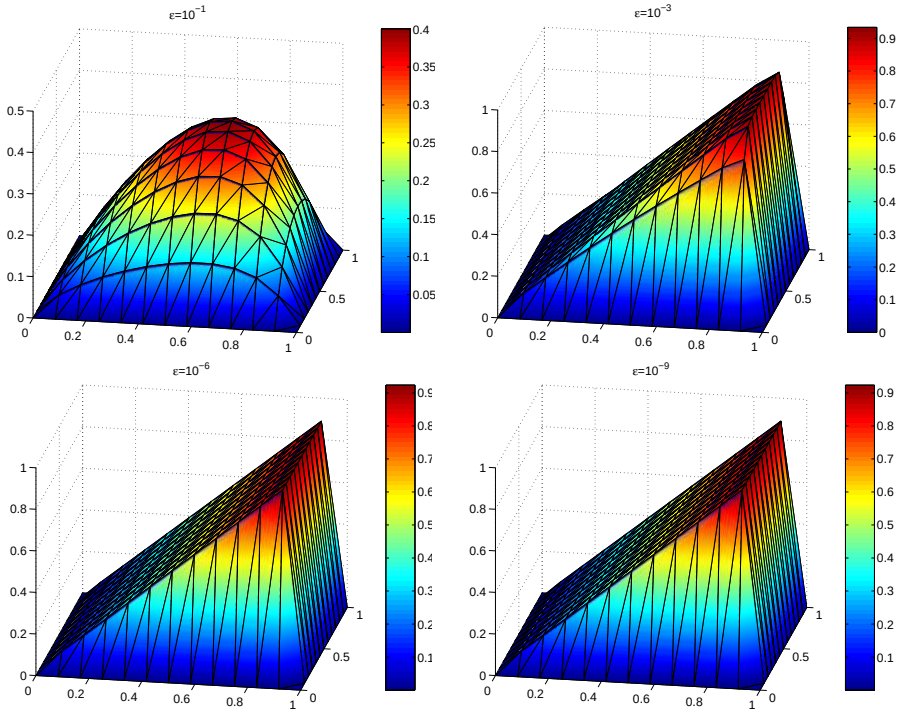

Figure 9: Solution for the bubble test case using structured meshes for different values of ε.

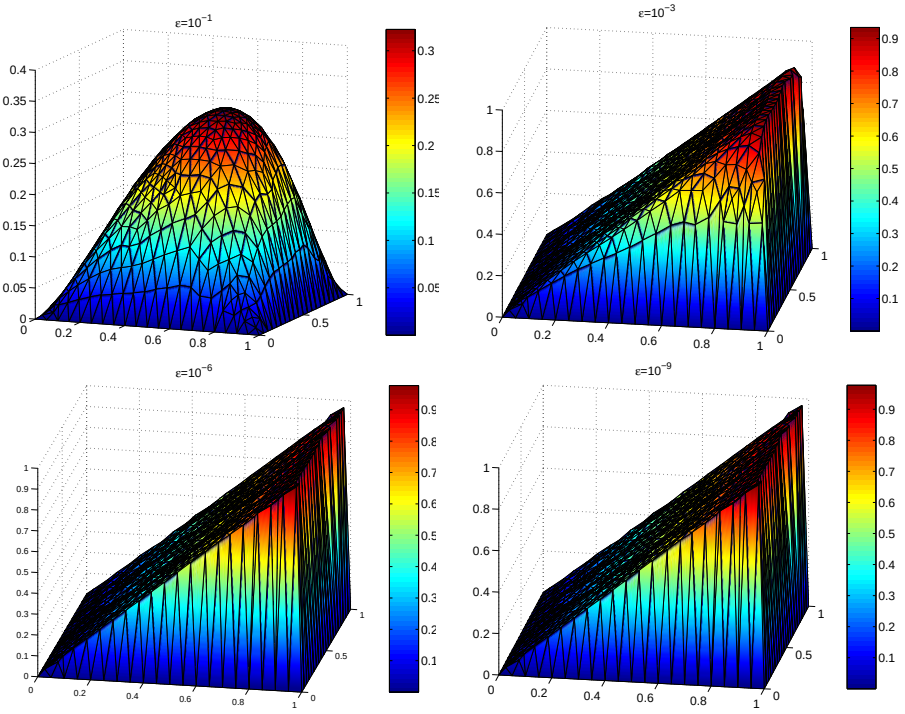

Figure 10: Solution for the bubble test case using unstructured meshes for $\varepsilon = [10^{-1}, 10^{-3}, 10^{-6}, 10^{-9}]$.

9 Conclusions

In this article we have extended the DPG finite element method to the numerical solution of the advection-diffusion equation. A static condensation procedure was used to eliminate the internal and the flux interface variables in favor of the remaining (hybrid) interface variable, leading to a nonconforming single-field formulation of strongly reduced size. In order to deal with the advection-dominated case, we have introduced a suitable flux-upwind stabilization technique, that has been proved to produce an optimally converging approximation measured in a discrete H^1-norm. The performance of the method and a simple (and conservative) flux-recovery post-processing have been successfully demonstrated in the numerical solution of several benchmark problems characterized by the presence of steep boundary and interior layers. The promising behaviour and economical (single-field) implementation of the proposed stabilized dual-primal DPG formulation suggest its use and extension to the solution of more complex problems in fluid mechanical applications.

References

[1] R.A. Adams. *Sobolev Spaces.* Academic Press, New York, 1975.

[2] D.N. Arnold and F. Brezzi. Mixed and nonconforming finite element methods: Implementation, postprocessing and error estimates. *Math. Modeling and Numer. Anal.*, 19-1:7–32, 1985.

[3] D.N. Arnold, F. Brezzi, B. Cockburn, and L.D. Marini. Discontinuous Galerkin methods. In *Lecture Notes in Computational Science and Engineering*, volume 11, pages 89–101. Springer-Verlag, 2000.

[4] I. Babuska and J. Osborn. Generalized finite element methods, their performance and their relation to mixed methods. *SIAM J. Numer. Anal.*, 20:510–536, 1983.

[5] J. Baranger, J.F. Maitre, and F. Oudin. Connection between finite volume and mixed finite element methods. *M^2AN*, 30:445–465, 1996.

[6] C. L. Bottasso, S. Micheletti, and R. Sacco. The Discontinuous Petrov-Galerkin method for elliptic problems. *Comput. Methods Appl. Mech. Engrg.*, 191:3391–3409, 2002.

[7] C. L. Bottasso, S. Micheletti, and R. Sacco. A multiscale formulation of the Discontinuous Petrov–Galerkin method for advective-diffusion problems. Submitted, MOX-Report No. 20, available at `http://mox.polimi.it/`, 2003.

[8] F. Brezzi, L.D. Marini, and P. Pietra. Numerical simulation of semiconductor devices. *Comput. Meths. Appl. Mech. and Engr.*, 75:493–514, 1989.

[9] F. Brezzi, L.D. Marini, and P. Pietra. Two-dimensional exponential fitting and applications to drift-diffusion models. *SIAM J. Numer. Anal.*, 26:1342–1355, 1989.

[10] P. Causin. *Mixed-hybrid Galerkin and Petrov-Galerkin finite element formulations in fluid mechanics.* PhD thesis, Università degli Studi di Milano, 2003.

[11] P. Causin and R. Sacco. Mixed-hybrid Galerkin and Petrov-Galerkin finite element formulations in continuum mechanics. In H.A. Mang, F.G. Rammerstorfer, and J. Eberhardsteiner, editors, *Proceedings of the Fifth World Congress on Computational Mechanics (WCCM V), Vienna, Austria.* Vienna University of Technology, Austria, ISBN 3-9501554-0-6, http://wccm.tuwien.ac.at, July 7-12 2002.

[12] P. Causin and R. Sacco. A Discontinuous Petrov–Galerkin method with Lagrangian multipliers for second order elliptic problems. MOX-Report No. 19, available at `http://mox.polimi.it/`, to appear in SIAM J. Numer. Anal., 2004.

[13] P.G. Ciarlet. *The Finite Element Method for Elliptic Problems.* North Holland, Amsterdam, 1978.

[14] B. Cockburn and J. Gopalakhrisnan. A characterization of hybridized mixed methods for second order elliptic problems. *SIAM Jour. Numer. Anal.*, 42 (1):283–301, 2003.

[15] M. Crouzeix and P.A. Raviart. Conforming and non-conforming finite element methods for solving the stationary Stokes equations. *R.A.I.R.O.*, R-3:33–76, 1973.

[16] C. Dawson. Godunov mixed methods for advection-diffusion equations in multidimensions. *SIAM J.Numer.Anal.*, 30:1315–1332, 1993.

[17] C. Dawson and V. Aizinger. Upwind-mixed methods for transport equations. *Comp. Geosc.*, 3:93–110, 1999.

[18] J. Jaffré. Decentrage et elements finis mixtes pour les equations de diffusion-convection. *Calcolo*, 2:171–197, 1984.

[19] J.W. Jerome. *Analysis of Charge Transport.* Springer-Verlag, Berlin Heidelberg, 1996.

[20] J.L. Lions and E. Magenes. *Problèmes aux limites non homogènes et applications.* Dunod, 1968.

[21] C. Liu and N.J. Walkington. Convergence of numerical approximations of the incompressible Navier-Stokes equations with variable density and viscosity. Submitted to Siam J. Numer. Anal., 2002.

[22] P.A. Markowich. *The Stationary Semiconductor Device Equations.* Springer-Verlag, Wien-New York, 1986.

[23] S. Micheletti, R. Sacco, and F.Saleri. On some mixed finite element methods with numerical integration. *SIAM J. Sci. Comput.*, 23-1:245–270, 2001.

[24] A. Mizukami and T.J.R. Hughes. A Petrov-Galerkin finite element method for convection–dominated flows: an accurate upwinding technique satisfying the discrete maximum principle. *Comput. Meth. Appl. Mech. Engrg.*, 50:181–193, 1985.

[25] K. Ohmori and T. Ushijima. A technique of upstream type applied to a linear nonconforming finite element approximation of convective diffusion equations. *R.A.I.R.O.*, 3:309–332, 1984.

[26] A. Quarteroni and A. Valli. *Numerical Approximation of Partial Differential Equations.* Springer-Verlag, New York, Berlin, 1994.

[27] P.A. Raviart and J.M. Thomas. Primal hybrid finite element methods for 2nd order elliptic equations. *Math. Comp.*, 31-138:391–413, 1977.

[28] J.E. Roberts and J.M. Thomas. Mixed and hybrid methods. In P.G. Ciarlet and J.L. Lions, editors, *Finite Element Methods, Part I.* North-Holland, Amsterdam, 1991. Vol.2.

[29] H. G. Roos, M. Stynes, and L. Tobiska. *Numerical methods for singularly perturbed differential equations*. Springer-Verlag, Berlin Heidelberg, 1996.

[30] R. Sacco, E. Gatti, and L. Gotusso. The patch test as a validation of a new finite element for the solution of convection-diffusion equations. *Comp. Meth. Appl. Mech. Engrg.*, 124:113–124, 1995.

[31] P. Siegel, R. Mosé, Ph. Ackerer, and J. Jaffré. Solution of the advection-diffusion equation using a combination of discontinuous and mixed finite elements. *Inter. J. Numer. Methods Fluids*, 24:593–613, 1997.

[32] R. Temam. *Navier-Stokes Equations*. North-Holland, Amsterdam, 1977.

Unité de recherche INRIA Rocquencourt
Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex (France)

Unité de recherche INRIA Lorraine : LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex (France)
Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)
Unité de recherche INRIA Rhône-Alpes : 655, avenue de l'Europe - 38330 Montbonnot-St-Martin (France)
Unité de recherche INRIA Sophia Antipolis : 2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399