IMA Journal of Numerical Analysis (2010) **30**, 1206–1234
doi:10.1093/imanum/drn083
Advance Access publication on July 10, 2009

A hybrid mixed discontinuous Galerkin finite-element method for convection–diffusion problems

HERBERT EGGER† AND JOACHIM SCHÖBERL
Centre for Computational Engineering Science, RWTH Aachen University, Aachen, Germany

[Received on 15 March 2008; revised on 8 November 2008]

We propose and analyse a new finite-element method for convection–diffusion problems based on the combination of a mixed method for the elliptic and a discontinuous Galerkin (DG) method for the hyperbolic part of the problem. The two methods are made compatible via hybridization and the combination of both is appropriate for the solution of intermediate convection–diffusion problems. By construction, the discrete solutions obtained for the limiting subproblems coincide with the ones obtained by the mixed method for the elliptic and the DG method for the limiting hyperbolic problem. We present a new type of analysis that explicitly takes into account the Lagrange multipliers introduced by hybridization. The use of adequate energy norms allows us to treat the purely diffusive, the convection-dominated and the hyperbolic regimes in a unified manner. In numerical tests we illustrate the efficiency of our approach and make a comparison with results obtained using other methods for convection–diffusion problems.

Keywords: convection–diffusion; upwind; finite-element method; discontinuous Galerkin methods; mixed methods; hybridization.

1. Introduction

In this paper we consider stationary convection–diffusion problems of the form

$$\begin{gathered} \operatorname{div}(-\epsilon \nabla u + \beta u) = f \quad \text{in } \Omega, \\ u = g_{\mathrm{D}} \quad \text{on } \partial\Omega_{\mathrm{D}}, \qquad -\epsilon \frac{\partial u}{\partial \nu} + \beta\nu u = g_{\mathrm{N}} \quad \text{on } \partial\Omega_{\mathrm{N}}, \end{gathered} \tag{1.1}$$

where Ω is a bounded open domain in $\mathbb{R}^d$, for $d = 2, 3$, with boundary $\partial\Omega = \partial\Omega_{\mathrm{D}} \cup \partial\Omega_{\mathrm{N}}$ consisting of a Dirichlet and a Neumann part, ϵ is a non-negative function and $\beta\colon \Omega \to \mathbb{R}^d$ is a d-dimensional vector field.

Similar problems arise in many applications, for example, in the modelling of contaminant transport, in electrohydrodynamics or macroscopic models for semiconductor devices. A feature that makes the numerical solution difficult is that convection often plays the dominant role. In the case of vanishing diffusion, solutions of (1.1) will, in general, not be smooth, i.e., discontinuities are propagated along the characteristic direction β. Nonlinear problems may even lead to discontinuities or blow up in a finite time when starting from smooth initial data. So appropriate numerical schemes for the convection-dominated regime have to be able to deal with almost discontinuous solutions in an accurate but stable manner. Another property that is also desirable to be reflected on the discrete level is the conservation structure inherent in the divergence form of (1.1).

†Corresponding author. Email: herbert.egger@rwth-aachen.de

Due to the variety of applications, there has been significant interest in the design and analysis of numerical schemes for convection-dominated problems. Much work has been devoted to devise accurate and stable finite-difference and finite-volume methods for the solution of hyperbolic systems by means of appropriate upwind techniques including flux or slope limiters in the nonlinear case.

A different approach to the stable solution of (almost) hyperbolic problems is offered by discontinuous Galerkin (DG) methods, introduced originally for a linear hyperbolic equation in neutron transport (Reed & Hill, 1973; LeSaint & Raviart, 1974; Johnson & Pitkäranta, 1986). Starting from the 1970s, DG methods have been investigated intensively and applied to the solution of various linear and nonlinear hyperbolic and convection-dominated elliptic problems with great success (cf. Bassi & Rebay (1997a,b), Aizinger *et al.* (2000) and Cockburn *et al.* (2000) for an overview and further references). Since in practical applications convection and diffusion phenomena may dominate in different parts of the computational domain, several attempts have been made to also generalize DG methods to elliptic problems (Richter, 1992; Oden *et al.*, 1998; Houston & Süli, 2001), yielding numerical schemes very similar to interior penalty methods studied much earlier (Nitsche, 1971; Babuska & Zlámal, 1973; Arnold, 1982). For further references on this topic and a unified analysis of several DG methods for elliptic problems we refer to Arnold *et al.* (2002). For DG methods applied to convection–diffusion problems we refer to Cockburn (1988), Baumann & Oden (1999), Castillo *et al.* (2002) and Buffa *et al.* (2006) for a multiscale version. Two disadvantages of DG methods applied to problems with diffusion are that, compared to a standard conforming discretization, the overall number of unknowns is increased substantially and that the resulting linear systems are much less sparse.

Another very successful approach for the solution of convection-dominated problems is the streamline diffusion method (Hughes & Brooks, 1979; Johnson & Saranen, 1986), where standard conforming finite-element discretizations are stabilized by adding in a conforming way an appropriate amount of artificial diffusion in the streamline direction. This method is easy to implement and yields stable discretizations in many situations, but may lead to unphysically large layers near discontinuities and boundaries. For a comparison of high-order DG and streamline diffusion methods we refer to Houston *et al.* (2000). For an appropriate treatment of boundary layers via Nitsche's method see Freund & Stenberg (1995). In contrast to DG methods, the streamline diffusion method does not yield conservative discretizations.

Here we follow a different approach, namely, the combination of upwind techniques used in DG methods for hyperbolic problems with conservative discretizations of mixed methods for elliptic problems. Other extensions of mixed finite-element methods to convection–diffusion problems were considered in Chen *et al.* (1995) and Dawson & Aizinger (1999).

In order to make the two different methods compatible we will utilize hybrid formulations for the mixed and the DG methods. It is well known (Arnold & Brezzi, 1985; Brezzi & Fortin, 1991; Cockburn *et al.*, 2009) that hybridization can be used for the efficient implementation of mixed finite elements for elliptic problems. Also introducing the Lagrange multipliers in the DG methods allows us to couple both methods naturally and yields a stable mixed hybrid DG method with the following properties.

- For $\beta \equiv 0$ the numerical solution coincides with that of a mixed method (cf. Arnold & Brezzi, 1985; Brezzi & Fortin, 1991), and postprocessing techniques can be used to increase the accuracy of the solution.
- For $\epsilon \equiv 0$ the solution coincides with that obtained by a DG method for hyperbolic problems (LeSaint & Raviart, 1974; Johnson & Pitkäranta, 1986).
- The intermediate convection–diffusion regime is treated automatically with no need to choose stabilization parameters.

For diffusion-dominated regions the stabilization can be omitted, yielding a scheme that was studied numerically in one dimension in Farhoul & Mounim (2005). Our analysis in Section 4.2 also includes this case.

A particular advantage of our method is that it is formulated and can be implemented element-wise, i.e., it allows for static condensation: local degrees of freedom (dofs) can be eliminated on the element level (see Brezzi & Fortin, 1991, Section V.1 or Section 5 for details), yielding global systems for the dofs on the mesh skeleton only. In this way, we can obtain global systems with less unknowns and sparser stencils than that of other DG methods, at the price of a somewhat more demanding assembling process. Further remarks and a comparison with the interior penalty are given in Section 5.2.

The relaxation of the coupling terms of DG methods has also been investigated recently by other authors. In Buffa *et al.* (2006) a method was proposed that, after the elimination of local dofs, yields a global system corresponding to that of a continuous Galerkin method (see also Brix *et al.* (2008) for similar ideas used for the construction of multilevel preconditioners). A further comparison with this method is given in Section 5.2. The hybridization of several DG methods has already been proposed in Cockburn *et al.* (2009), but without convergence analysis.

The outline of this article is as follows. In Section 2 we review the hybrid formulation of the mixed method for the Poisson equation and then introduce a hybrid version of the DG method for the hyperbolic subproblem. The scheme for the intermediate convection–diffusion regime then results from a combination of the two methods for the limiting subproblems, and we show consistency and conservation of all three methods under consideration. Section 3 presents the main stability and boundedness estimates for the corresponding bilinear forms and contains an *a priori* error analysis in the energy norm with emphasis on the convection-dominated regime. Details on superconvergence results and postprocessing for the diffusion-dominated case are presented in Section 4. Results of numerical tests, including a comparsion with the streamline diffusion method, are presented in Section 5.

2. Hybrid mixed DG methods for convection–diffusion problems

The aim of this section is to formulate the problem under consideration in detail and to fix the relevant notation and some basic assumptions. By introducing the *diffusive flux* $\sigma = -\epsilon \nabla u$ as a new variable, we rewrite (1.1) in mixed form as follows:

$$\sigma + \epsilon \nabla u = 0, \quad \operatorname{div}(\sigma + \beta u) = f \quad \text{in } \Omega$$
$$u = g_{\mathrm{D}} \quad \text{on } \partial\Omega_{\mathrm{D}}, \qquad -\epsilon \tfrac{\partial u}{\partial \nu} + \beta\nu u = g_{\mathrm{N}} \quad \text{on } \partial\Omega_{\mathrm{N}}, \tag{2.1}$$

which will be the starting point for our considerations. Here and below ν denotes the outward unit normal vector on the boundary of some domain. We refer to βu as the *convective flux* and call $\sigma + \beta u$ the *total flux*. The existence and uniqueness of a solution to (2.1) follow under standard assumptions on the coefficients. For ease of presentation, let us make some simplifying assumptions.

2.1 *Basic assumptions and notation*

We assume that Ω is a polyhedral domain and that $\partial\Omega_{\mathrm{D}} = \partial\Omega$, i.e., $\partial\Omega_{\mathrm{N}} = \emptyset$. Let $\mathcal{T}_h$ be a shape regular partition of Ω into simplices T and let $\mathcal{E}_h$ denote the set of facets E. By the term *facets* we denote interfaces between elements or to the boundary, i.e., faces or edges in three or two dimensions, respectively. We assume that each element T and facet E are generated by an affine map Φ_T or Φ_E

from a corresponding reference element $\hat{T}$ or $\hat{E}$, respectively. With $\partial\mathcal{T}_h$ we denote the set of all element boundaries ∂T (with outward normal ν). Finally, by χ_S we denote the characteristic function of a set $S \subset \Omega$.

Regarding the coefficients, we assume for simplicity that $g_D = 0$ and that $\epsilon \geqslant 0$ is constant on elements $T \in \mathcal{T}_h$. Furthermore, the vector field β is assumed to be piecewise constant with continuous normal components across element interfaces, which implies that $\mathrm{div}\beta = 0$. Moreover, such a vector field β induces a natural splitting of element boundaries into inflow and outflow parts, i.e., we define the outflow boundary $\partial T^{\mathrm{out}} := \{x \in \partial T : \beta\nu > 0\}$ and $\partial T^{\mathrm{in}} = T \setminus \partial T^{\mathrm{out}}$. The unions of the element inflow and outflow boundaries will be denoted by $\partial\mathcal{T}_h^{\mathrm{in}}$ and $\partial\mathcal{T}_h^{\mathrm{out}}$, respectively, and, similarly, the symbols $\partial\Omega^{\mathrm{in}}$ and $\partial\Omega^{\mathrm{out}}$ are used for the inflow and outflow regions, respectively of the boundary $\partial\Omega$.

For our analysis we will utilize the broken Sobolev spaces

$$H^s(\mathcal{T}_h) := \{u : u \in H^s(T),\ \forall\, T \in \mathcal{T}_h\}, \quad s \geqslant 0,$$

and for functions $u \in H^{s+1}(\mathcal{T}_h)$ we define $\nabla u \in [H^s(\mathcal{T}_h)]^d$ to be the piecewise gradient. In a natural manner, we define the inner products

$$(u, v)_T := \int_T uv \,\mathrm{d}x \quad \text{and} \quad (u, v)_{\mathcal{T}_h} := \sum_{T \in \mathcal{T}_h} (u, v)_T,$$

with the obvious modifications for vector-valued functions. The norm induced by the volume integrals $(\cdot,\cdot)_{\mathcal{T}_h}$ is denoted by $\|u\|_{\mathcal{T}_h} := \sqrt{(u, u)_{\mathcal{T}_h}}$, and for piecewise constant α we define $\alpha(u, v)_{\mathcal{T}_h} := \sum_T (\alpha u, v)_T$ and $\alpha\|u\|_{\mathcal{T}_h} := \sqrt{\alpha^2(u, u)_{\mathcal{T}_h}}$. Norms and seminorms on the broken Sobolev spaces $H^s(\mathcal{T}_h)$ will be denoted by $\|\cdot\|_{s,\mathcal{T}_h}$ and $|\cdot|_{s,\mathcal{T}_h}$.

For the element interfaces we consider the function spaces

$$L^2(\mathcal{E}_h) := \{\mu : \mu \in L^2(E),\ \forall\, E \in \mathcal{E}_h\}$$

and

$$L^2(\partial\mathcal{T}_h) := \{v : v \in L^2(\partial T),\ \forall\, T \in \mathcal{T}_h\}.$$

Note that functions in $L^2(\partial\mathcal{T}_h)$ are double valued on element interfaces and may be considered as traces of element-wise defined functions. Moreover, we can identify $\mu \in L^2(\mathcal{E}_h)$ with a function $v \in L^2(\partial\mathcal{T}_h)$ by duplicating the values at element interfaces, and so in this sense $L^2(\mathcal{E}_h) \subset L^2(\partial\mathcal{T}_h)$. For $u, v \in L_2(\partial\mathcal{T}_h)$ we denote integrals over element interfaces by

$$\langle\lambda, \mu\rangle_{\partial T} := \int_{\partial T} \lambda\mu \,\mathrm{d}s \quad \text{and} \quad \langle\lambda, \mu\rangle_{\partial\mathcal{T}_h} := \sum_T \langle\lambda, \mu\rangle_{\partial T},$$

and the corresponding norms are denoted by $|u|_{\partial\mathcal{T}_h} := \sqrt{\langle u, u\rangle_{\partial\mathcal{T}_h}}$. Again, we write $\alpha\langle u, v\rangle_{\partial\mathcal{T}_h}$ with the meaning $\sum_{\partial T}\langle\alpha u, v\rangle_{\partial T}$.

Let us now turn to the formulation of appropriate finite-element spaces. We start from piecewise polynomials on the reference elements and define the finite-element spaces via appropriate mappings (cf. Brenner & Scott, 2002). By $\mathcal{P}_k(\hat{T})$ and $\mathcal{P}_k(\hat{E})$ we denote the sets of all polynomials of order at most k on the reference elements, and by $RT_k(\hat{T}) := \mathcal{P}_k(\hat{T}) \oplus \vec{x} \cdot \mathcal{P}_k(\hat{T})$ we denote the Raviart–Thomas (–Nedelec) element (cf. Raviart & Thomas, 1977; Nedelec, 1980; Brezzi & Fortin, 1991). Here the symbol $\oplus$ is used to denote the union of two vector spaces. For our finite-element methods we will utilize the following function spaces:

$$\Sigma_h := \left\{ \tau_h \in [L_2(\Omega)]^d \colon \tau_h|_T = \frac{1}{\det \Phi_T'} \Phi_T' \hat{\tau} \circ \Phi_T^{-1}, \hat{\tau} \in RT_k(\hat{T}) \right\},$$

$$\mathcal{V}_h := \{ v_h \in L_2(\Omega) \colon v_h|_T = \hat{v} \circ \Phi_T^{-1}, \hat{v} \in \mathcal{P}_k(\hat{T}) \},$$

$$\mathcal{M}_h := \{ \mu_h \in L_2(\mathcal{E}_h) \colon \mu|_E = \hat{\mu} \circ \Phi_E^{-1}, \mu = 0 \text{ on } \partial\Omega, \hat{\mu} \in \mathcal{P}_k(\hat{E}) \}.$$

For convenience, we will sometimes use the notation $\mathcal{W}_h := \Sigma_h \times \mathcal{V}_h \times \mathcal{M}_h$. Since we assumed that our elements T are generated by affine maps Φ_T, the finite-element spaces could be defined equivalently as the appropriate polynomial spaces on the mapped triangles (cf. Brezzi & Fortin, 1991). This would, however, complicate a generalization to nonaffine elements.

Let us now turn to the formulation of the finite-element methods. We will start by recalling the hybrid mixed formulation for the elliptic subproblem ($\beta \equiv 0$) and then introduce a hybrid version for the DG method for the hyperbolic subproblem ($\epsilon \equiv 0$). The scheme for the intermediate convection–diffusion problem then results by simply adding up the bilinear and linear forms of the limiting subproblems.

2.2 *Diffusion*

For $\beta \equiv 0$ equation (2.1) reduces to the mixed form of the Dirichlet problem

$$\sigma = -\epsilon \nabla u, \quad \operatorname{div}\sigma = f \quad \text{in } \Omega, \quad u = 0 \quad \text{on } \partial\Omega, \tag{2.2}$$

and the corresponding (dual) mixed variational problem reads

$$\begin{aligned} \frac{1}{\epsilon}(\sigma, \tau)_{\mathcal{T}_h} - (u, \operatorname{div}\tau)_{\mathcal{T}_h} &= 0 && \forall\, \tau \in H(\operatorname{div}, \Omega), \\ (\operatorname{div}\sigma, v)_{\mathcal{T}_h} &= (f, v)_{\mathcal{T}_h} && \forall\, v \in L^2(\Omega). \end{aligned}$$

While a conforming discretization of (2.2) allows us to also easily obtain conservation on the discrete level, it also has some disadvantages: the resulting linear system is a saddle-point problem and involves considerably more dofs than a standard (primal) H^1-conforming discretization of (2.2). Both difficulties can be overcome by hybridization (cf. Arnold & Brezzi, 1985; Brezzi & Fortin, 1991; Cockburn *et al.*, 2009). Let us briefly sketch the main ideas: instead of requiring the discrete fluxes to be in $H(\operatorname{div}, \Omega)$, one can use completely discontinuous piecewise polynomial ansatz functions and ensure the continuity of the normal fluxes over element interfaces by adding appropriate constraints. The corresponding discretized variational problem reads

$$\begin{aligned} \frac{1}{\epsilon}(\sigma_h, \tau_h)_{\mathcal{T}_h} - (u_h, \operatorname{div}\tau_h)_{\mathcal{T}_h} + \langle \lambda_h, \tau_h \nu \rangle_{\partial\mathcal{T}_h} &= 0 && \forall\, \tau_h \in \Sigma_h, \\ (\operatorname{div}\sigma_h, v_h)_{\mathcal{T}_h} &= (f, v)_{\mathcal{T}_h} && \forall\, v_h \in \mathcal{V}_h, \\ \langle \sigma_h \nu, \mu_h \rangle_{\partial\mathcal{T}_h} &= 0 && \forall\, \mu_h \in \mathcal{M}_h. \end{aligned}$$

Note that the choice of finite-element spaces allows us to eliminate the dual and primal variables on the element level, yielding a global (positive definite) system for the Lagrange multipliers only. The global system has an optimal sparsity pattern and information on the Lagrange multipliers can be used further to obtain better reconstructions by local postprocessing. We refer to Arnold & Brezzi (1985), Brezzi & Fortin (1991) and Stenberg (1991) for further discussion of these issues and come back to postprocessing later in Section 4.

After integration by parts, we arrive at the following hybrid mixed finite-element method.

METHOD 2.1 (Diffusion) Find $(\sigma_h, u_h, \lambda_h) \in \Sigma_h \times \mathcal{V}_h \times \mathcal{M}_h$ such that

$$\mathcal{B}_{\mathrm{D}}(\sigma_h, u_h, \lambda_h; \tau_h, v_h, \mu_h) = \mathcal{F}_{\mathrm{D}}(\tau_h, v_h, \mu_h) \tag{2.3}$$

for all $\tau_h \in \Sigma_h$, $v_h \in \mathcal{V}_h$ and $\mu_h \in \mathcal{M}_h$, where $\mathcal{B}_{\mathrm{D}}$ and $\mathcal{F}_{\mathrm{D}}$ are defined by

$$\begin{aligned}&\mathcal{B}_{\mathrm{D}}(\sigma_h, u_h, \lambda_h; \tau_h, v_h, \mu_h)\\ &\quad := \frac{1}{\epsilon}(\sigma_h, \tau_h)_{\mathcal{T}_h} + (\nabla u_h, \tau_h)_{\mathcal{T}_h} + \langle \lambda_h - u_h, \tau_h \nu\rangle_{\partial\mathcal{T}_h} + (\sigma_h, \nabla v_h)_{\mathcal{T}_h} + \langle \sigma_h \nu, \mu_h - v_h\rangle_{\partial\mathcal{T}_h}\end{aligned} \tag{2.4}$$

and

$$\mathcal{F}_{\mathrm{D}}(\tau_h, v_h, \mu_h) := -(f, v_h)_{\mathcal{T}_h}. \tag{2.5}$$

We only mention that the case $\epsilon = 0$ on some elements T can be allowed in principle. For these elements the term $\frac{1}{\epsilon}(\sigma_h, \tau_h)_T$ just has to be interpreted as $\sigma_h|_T \equiv 0$.

REMARK 2.2 Let $\Sigma := [H^1(\mathcal{T}_h)]^d$, $\mathcal{V} := H^1(\mathcal{T}_h)$ and $\mathcal{M} := \{\mu \in L^2(\mathcal{E}_h)\colon \mu = 0 \text{ on } \partial\Omega\}$, and let $\mathcal{W} := \Sigma \times \mathcal{V} \times \mathcal{M}$ denote the continuous analogue to $\mathcal{W}_h$. The above bilinear form is then defined for all $(\sigma, u, \lambda; \tau_h, v_h, \mu_h) \in \mathcal{W} \oplus \mathcal{W}_h \times \mathcal{W}_h$. This property will be used below to show consistency of the method and to obtain Galerkin orthogonality. Using appropriate lifting operators $L\colon L^2(\mathcal{T}_h) \to \Sigma_h$, the terms involving integrals over the boundary can be replaced by volume integrals, for example, $\langle \lambda_h - u_h, \tau_h \nu\rangle_{\partial\mathcal{T}_h} = (L(\lambda_h - u_h), \tau_h)_{\mathcal{T}_h}$, and in this way Method 2.1 can be well defined on $(\mathcal{W}_h \oplus \mathcal{W}) \times (\mathcal{W}_h \oplus \mathcal{W})$. Such *extensions* are used, for example, in Perugia & Schötzau (2002) and Houston *et al.* (2007) for the *hp*-error analysis of DG methods under minimal regularity assumptions.

Method 2.1 is algebraically equivalent to the conforming $RT_k \times \mathcal{P}_k$ discretization of the dual mixed formulation of (2.2) and can be seen as a pure implementation trick. Below we will analyse Method 2.1 in a somewhat nonstandard way, including the gradient of the primal variable and the Lagrange multipliers explicitly in the energy norm. This kind of analysis is quite close to that of DG methods for elliptic problems and allows us to investigate the mixed method together with the DG method for the hyperbolic subproblem in a uniform framework.

2.3 *Convection*

By setting $\epsilon \equiv 0$ in (2.1), we arrive at the limiting hyperbolic problem

$$\operatorname{div}(\beta u) = f \quad \text{in } \Omega, \qquad u = 0 \quad \text{on } \partial\Omega^{\text{in}}. \tag{2.6}$$

Multiplying (2.6) by a test function $v \in H^1(\mathcal{T}_h)$, and adding upwind stabilization, we obtain the DG method for hyperbolic problems (Reed & Hill, 1973; LeSaint & Raviart, 1974; Johnson & Pitkäranta, 1986)

$$(\operatorname{div}(\beta u), v)_{\mathcal{T}_h} + \langle \beta\nu(u^+ - u), v\rangle_{\partial\mathcal{T}_h^{\text{in}}} = (f, v)_{\mathcal{T}_h},$$

where $u^+ := u|_{\partial T^+}$ denotes the upwind value and T^+ is the upwind element, i.e., the element attached to E where $\beta\nu = \beta \cdot \nu_T \geqslant 0$. To incorporate the boundary condition we define $u^+ = 0$ on $\partial\Omega^{\text{in}}$. After integration by parts and noting that $u = u^+$ on ∂T^{out}, we obtain that

$$(u, \beta\nabla v)_{\mathcal{T}_h} - \langle \beta\nu u^+, v\rangle_{\partial\mathcal{T}_h^{\text{in}}} - \langle \beta\nu u, v\rangle_{\partial\mathcal{T}_h^{\text{out}}} = -(f, v)_{\mathcal{T}_h}.$$

In order to make the DG method compatible with the hybrid mixed method formulated in the Section 2.2 let us introduce the upwind value as a new variable $\lambda := u^+$, and let us define the symbol

$$\{\lambda/u\} := \begin{cases} \lambda, & E \subset \partial T^{\text{in}}, \\ u, & E \subset \partial T^{\text{out}}, \end{cases}$$

for all $T \in \mathcal{T}_h$. Note that $\lambda = \{\lambda/u\} = u^+$ on both sides of E, and so $\{\lambda/u\}$ is just a new characterization of the upwind value. After discretization, we now arrive at the following hybrid version of the DG method.

METHOD 2.3 (Convection) Find $(u_h, \lambda_h) \in \mathcal{V}_h \times \mathcal{M}_h$ such that

$$\mathcal{B}_{\text{C}}(u_h, \lambda_h; v_h, \mu_h) = \mathcal{F}_{\text{C}}(v_h, \mu_h) \tag{2.7}$$

for all $(v_h, \mu_h) \in \mathcal{V}_h \times \mathcal{M}_h$ with

$$\mathcal{B}_{\text{C}}(u_h, \lambda_h; v_h, \mu_h) := (u_h, \beta\nabla v_h)_{\mathcal{T}_h} + \langle \beta\nu\{\lambda_h/u_h\}, \mu_h - v_h\rangle_{\partial\mathcal{T}_h} \tag{2.8}$$

and

$$\mathcal{F}_{\text{C}}(v_h, \mu_h) := -(f, v)_{\mathcal{T}_h}. \tag{2.9}$$

By construction, Method 2.3 is algebraically equivalent to the classical DG method. This can easily be seen by testing with $\mu_h = \chi_E$, which yields that $\lambda_h = u_h{}^+$ on the element interfaces. All terms of the bilinear form are again defined element-wise, which allows us to use static condensation on the element level. Moreover, as in the case of pure diffusion, the bilinear form $\mathcal{B}_{\text{C}}$ can be extended onto $\mathcal{W} \oplus \mathcal{W}_h \times \mathcal{W}_h$, which then allows us to derive consistency and use Galerkin orthogonality arguments. On facets E where $\beta\nu = 0$, the Lagrange multiplier is not uniquely defined, and we set $\lambda = 0$ there.

2.4 *Convection–diffusion regime*

Let us now return to the original convection–diffusion problem and consider the system

$$\sigma + \epsilon\nabla u = 0, \quad \operatorname{div}(\sigma + \beta u) = f \quad \text{in } \Omega, \quad u = 0 \quad \text{on } \partial\Omega. \tag{2.10}$$

Since we used the same spaces for the discretization of the elliptic and hyperbolic subproblems, the two hybrid methods can be coupled in a very natural way by simply adding up their bilinear and linear forms. This yields the following hybrid mixed DG method for the intermediate convection–diffusion regime.

METHOD 2.4 (Convection–diffusion) Find $(\sigma_h, u_h, \lambda_h) \in (\Sigma_h, \mathcal{V}_h, \mathcal{M}_h)$ such that

$$\mathcal{B}(\sigma_h, u_h, \lambda_h; \tau_h, v_h, \mu_h) = \mathcal{F}(\sigma_h, u_h, \lambda_h) \tag{2.11}$$

for all $\tau_h \in \Sigma_h$, $v_h \in \mathcal{V}_h$ and $\mu_h \in \mathcal{M}_h$, where $\mathcal{B}$ and $\mathcal{F}$ are defined by

$$\begin{aligned}\mathcal{B}(\sigma_h, u_h, \lambda_h; \tau_h, v_h, \mu_h) :=& \frac{1}{\epsilon}(\sigma_h, \tau_h)_{\mathcal{T}_h} + (\nabla u_h, \tau_h)_{\mathcal{T}_h} + \langle \lambda_h - u_h, \tau_h \nu \rangle_{\partial \mathcal{T}_h} \\ &+ (\sigma_h + \beta u_h, \nabla v_h)_{\mathcal{T}_h} + \langle \sigma_h \nu + \beta \nu \{\lambda_h / u_h\}, \mu_h - v_h \rangle_{\partial \mathcal{T}_h} \end{aligned} \tag{2.12}$$

and

$$\mathcal{F}(\tau_h, v_h) := -(f, v_h)_{\mathcal{T}_h}. \tag{2.13}$$

By testing with $\mu_h = \chi_E$ for $E \in \mathcal{E}_h$, we obtain that $\sigma_h \nu_E + \beta \nu_E \{\lambda_h / u_h\}$ is continuous across element interfaces. Here ν_E denotes the unit normal vector on E with fixed orientation. Thus λ_h and $\sigma_h \nu_E + \beta \nu_E \{\lambda_h / u_h\}$ have unique values on the element interfaces and can be considered as discrete traces for u and the total flux $\sigma + \beta u$.

2.5 *Consistency and conservation*

Before we turn to a detailed analysis of the finite-element Methods 2.1, 2.3 and 2.4, let us summarize two important properties that follow almost directly from the corresponding properties of the mixed and the DG methods for limiting subproblems. For the sake of completeness, we sketch the proofs in the present framework.

PROPOSITION 2.5 (Consistency) Methods 2.1, 2.3 and 2.4 are consistent. That is, let u denote the solution of the problems (2.2), (2.6) and (2.10), respectively, and define $\sigma = -\epsilon \nabla u$ and $\lambda = u$. Then the corresponding variational equations (2.3), (2.7) and (2.11) hold if σ_h, u_h and λ_h are replaced by σ, u and λ.

Proof. We first consider Method 2.1. Let u denote the solution of (2.2) and make the substitutions as mentioned in the proposition. Then we obtain by testing the bilinear form $\mathcal{B}_{\mathrm{D}}$ with $(\tau_h, 0, 0)$ that

$$\begin{aligned}&\mathcal{B}_{\mathrm{D}}(-\epsilon \nabla u, u, u; \tau_h, 0, 0) \\ &\quad = -(\nabla u, \tau_h)_{\mathcal{T}_h} + (\nabla u, \tau_h)_{\mathcal{T}_h} - \langle u - u, \tau_h \nu \rangle_{\partial \mathcal{T}_h \setminus \partial \Omega} - \langle u, \tau_h \nu \rangle_{\partial \Omega} \\ &\quad = -\langle u, \tau_h \nu \rangle_{\partial \Omega} = 0.\end{aligned}$$

Next we test with $(0, v_h, 0)$ and integrate by parts to recover

$$\mathcal{B}_{\mathrm{D}}(-\epsilon \nabla u, u, u; 0, v_h, 0) = -(\mathrm{div}(-\epsilon \nabla u), v_h)_{\mathcal{T}_h} = -(f, v_h)_{\mathcal{T}_h},$$

which follows since u is the solution of (2.2). Finally, testing with $(0, 0, \mu_h)$ we obtain that

$$\mathcal{B}_{\mathrm{D}}(-\epsilon \nabla u, u, u; 0, 0, \mu_h) = \left\langle -\epsilon \frac{\partial u}{\partial n}, \mu_h \right\rangle_{\partial \mathcal{T}_h} = 0,$$

which holds since $\mathrm{div}(\epsilon \nabla u) = f \in L^2$ implies that $\epsilon \nabla u \in H(\mathrm{div}; \Omega)$ and thus the normal flux $-\epsilon \frac{\partial u}{\partial n}$ is continuous across element interfaces. Note that, at this point, we formally require some extra regularity, for example, $u \in H^1(\Omega) \cap H^{3/2+\varepsilon}(\mathcal{T}_h)$ or $\sigma = -\epsilon \nabla u \in L^s(\Omega)$ for some $s > 2$, in order to ensure that the moments $\left\langle \epsilon \frac{\partial u}{\partial n}, \mu_h \right\rangle$ are well defined for $\mu_h \in \mathcal{M}_h$ (cf. Brezzi & Fortin, 1991). As already mentioned in Remark 2.2, this extra regularity assumption can be dropped by appropriately extending B_{D}. In summary, we have shown that Method 2.1 is consistent.

Next consider Method 2.3 and let u denote the solution of (2.6). Substituting u for u_h and λ_h in (2.7)–(2.9) and testing with $(v_h, 0)$, we obtain after integration by parts that

$$\mathcal{B}_{\mathrm{C}}(u, u; v_h, 0) = (\mathrm{div}(\beta u), v_h)_{\mathcal{T}_h} - \langle \beta \nu u, v_h \rangle_{\partial \Omega^{\mathrm{in}}} = -(f, v_h)_{\mathcal{T}_h}.$$

Now test with $(0, \mu_h)$. Then we have

$$\mathcal{B}_{\mathrm{C}}(u, u; 0, \mu_h) = \langle \beta \nu u, \mu_h \rangle_{\partial \mathcal{T}_h} = 0$$

since u and μ_h are single valued and $\beta \nu$ appears two times with different signs for each element interface. Thus we have proven consistency of Method 2.3.

Finally, Method 2.4 is consistent as it is the sum of two consistent methods. □

While consistency is a key ingredient for the derivation of *a priori* error estimates, conservation is a property of the discrete methods that is desired for physical reasons since it inhibits unphysical increase of mass or total charge. This is particularly important for time-dependent problems. If a finite-element scheme allows us to test with piecewise-constant functions, then conservation can be shown to hold locally (for each element) as well as globally as long as the discrete fluxes are single valued on element interfaces.

PROPOSITION 2.6 (Conservation) Methods 2.1, 2.3 and 2.4 are locally and globally conservative.

Proof. Let us first show the local conservation of Method 2.1 by testing (2.3) with $(0, \chi_T, 0)$. This yields

$$-(f, 1)_T = \mathcal{B}_{\mathrm{D}}(u_h, \lambda_h, \sigma_h; 0, \chi_T, 0) = -\langle \sigma_h \nu, 1 \rangle_{\partial T},$$

that is, the total flux over an element boundary equals the sum of internal sources, and hence the method is locally conservative. By testing with $(0, 0, \chi_E)$ for some $E \in \mathcal{E}_h$, we obtain continuity of the normal fluxes $\sigma_h \nu$ across element interfaces, and so the scheme is also globally conservative. Now consider Method 2.3. Testing with $(\chi_T, 0)$, we get

$$(f, 1)_T = \mathcal{B}_{\mathrm{C}}(u_h, \lambda_h; \chi_T, 0) = \langle \beta \nu \lambda_h, 1 \rangle_{\partial T^{\mathrm{in}}} + \langle \beta \nu u_h, 1 \rangle_{\partial T^{\mathrm{out}}},$$

and so the total flux over the element boundaries equals the sum of internal sources and fluxes over the boundary of the domain. Note that $\beta \nu \{\lambda_h / u_h\}$ defines a unique flux on element interfaces. Now let $E \in \mathcal{E}_h$ such that $E = \partial T_1^{\mathrm{out}} \cap \partial T_2^{\mathrm{in}}$. By testing with $(0, \chi_E)$, we obtain that

$$\begin{aligned} 0 = \mathcal{B}_{\mathrm{C}}(u_h, \lambda_h; 0, \chi_E) &= \langle \beta \nu \{\lambda_h / u_h\}, 1 \rangle_{\partial T_1^{\mathrm{out}}} + \langle \beta \nu \{\lambda_h / u_h\}, 1 \rangle_{\partial T_2^{\mathrm{in}}} \\ &= \langle \beta \nu u_h, 1 \rangle_{\partial T_1^{\mathrm{out}}} + \langle \beta \nu \lambda_h, 1 \rangle_{\partial T_2^{\mathrm{in}}}, \end{aligned}$$

and so the total outflow over a facet on one element balances the inflow over the same facet on the neighbouring element.

Finally, Method 2.4 is conservative as it is the sum of two conservative methods. □

3. *A priori* error analysis

As already mentioned previously, our analysis of the hybrid methods under consideration is inspired by that of DG methods (Johnson & Pitkäranta, 1986; Arnold *et al.*, 2002). In particular, we will utilize similar mesh-dependent energy norms for proving the stability and boundedness of the bilinear and linear forms. We will show the stability of Method 2.1 in the norm

$$|||(\tau, v, \mu)|||_{\mathrm{D}} := \left(\frac{1}{\epsilon}\|\tau\|^2_{\mathcal{T}_h} + \epsilon\|\nabla v\|^2_{\mathcal{T}_h} + \frac{\epsilon}{h}|\lambda - u|^2_{\partial\mathcal{T}_h}\right)^{1/2}, \tag{3.1}$$

and the stability of Method 2.3 will be analysed with respect to the norm

$$|||(u, \lambda)|||_{\mathrm{C}} := \left(\frac{h}{|\beta|}\|\beta\nabla u\|^2_{\mathcal{T}_h} + |\beta\nu||\lambda - u|^2_{\partial\mathcal{T}_h}\right)^{1/2}. \tag{3.2}$$

Here by $|\beta|$ and $|\beta\nu|$ we understand appropriate bounds for β and $\beta\nu$, respectively, on single elements or facets. Note that, for $\epsilon \sim h\beta$ (the crossover from the diffusion-dominated to the convection-dominated regime), all terms in (3.1) and (3.2) scale uniformly with respect to ϵ, β and h. For proving the boundedness of the bilinear forms we require the following slightly different norms:

$$|||(\tau, v, \mu)|||_{\mathrm{D},*} := \left(|||(\tau, v, \mu)|||^2_{\mathrm{D}} + \frac{h}{\epsilon}|\tau\nu|^2_{\partial\mathcal{T}_h}\right)^{1/2} \tag{3.3}$$

and

$$|||(u, \lambda)|||_{\mathrm{C},*} := \left(\frac{|\beta|}{h}\|u\|^2_{\mathcal{T}_h} + |\beta\nu||\{\lambda/u\}|^2_{\partial\mathcal{T}_h}\right)^{1/2}. \tag{3.4}$$

These norms scale again in the same manner with respect to h, ϵ and β as their counterparts (3.1) and (3.2), and therefore it can be shown easily that the additional terms do not disturb the approximation.

3.1 *Pure diffusion—Method 2.1*

Below we will require the following preparatory result.

LEMMA 3.1 Let $v_h \in \mathcal{V}_h$ and $\mu_h \in \mathcal{M}_h$ be given. Then there exists a unique solution $\tilde{\tau} \in \Sigma_h$ defined element-wise by the variational problems

$$\begin{aligned} (\tilde{\tau}, p)_T &= (\nabla v_h, p)_T \quad \forall\, p \in [\mathcal{P}_{k-1}(T)]^d, \\ \langle \tilde{\tau}\nu, q\rangle_{\partial T} &= \langle \mu_h, q\rangle_{\partial T} \quad \forall\, q \in \mathcal{P}_k(\partial T). \end{aligned}$$

Moreover, there exists a constant c_I only depending on the shape of the elements such that

$$\|\tilde{\tau}\|_{\mathcal{T}_h} \leqslant c_I \left(\|\nabla v_h\|^2_{\mathcal{T}_h} + h|\mu_h|^2_{\partial\mathcal{T}_h}\right)^{1/2} \tag{3.5}$$

holds.

Proof. The existence of a unique solution $\tilde{\tau}$ follows with standard arguments, and the norm estimate then follows by the usual scaling argument and the equivalence of norms on finite-dimensional spaces (cf. Brezzi & Fortin (1991) for details). □

Since the estimate (3.5) uses an inverse inequality, the constant c_I depends on the shapes of the elements. Lemma 3.1 now allows us to construct a suitable test function for establishing the following stability estimate.

PROPOSITION 3.2 (Stability) There exists a positive constant c_{D} that is independent of the mesh size h such that the estimate

$$\sup_{(\tau_h,v_h,\mu_h)} \frac{\mathcal{B}_{\mathrm{D}}(\sigma_h,u_h,\lambda_h;\tau_h,v_h,\mu_h)}{|||(\tau_h,v_h,\mu_h)|||_{\mathrm{D}}} \geqslant c_{\mathrm{D}} |||(\sigma_h,u_h,\lambda_h)|||_{\mathrm{D}} \tag{3.6}$$

holds for all $(\sigma_h,u_h,\lambda_h) \in \Sigma_h \times \mathcal{V}_h \times \mathcal{M}_h$.

Proof. Let us start with testing the bilinear form (2.4) with $(\sigma_h,-u_h,-\lambda_h)$, which yields

$$\mathcal{B}_{\mathrm{D}}(\sigma_h,u_h,\lambda_h;\sigma_h,-v_h,-\mu_h) = \frac{1}{\epsilon}\|\sigma_h\|^2_{\mathcal{T}_h}.$$

Now let $\tilde{\tau}$ be defined as in Lemma 3.1 with μ_h replaced by $\frac{\epsilon}{h}(\lambda_h-u_h)$ and ∇v_h replaced by $\epsilon\nabla u_h$, so that

$$\|\tilde{\tau}\|_{\mathcal{T}_h} \leqslant c_I\left(\frac{\epsilon^2}{h}|\lambda_h-u_h|^2_{\partial\mathcal{T}_h}+\epsilon^2\|\nabla u_h\|^2_{\mathcal{T}_h}\right)^{1/2} \tag{3.7}$$

holds with a constant c_I that is independent of the mesh size h. For $\gamma > 0$ we then obtain

$$\begin{aligned}
&\mathcal{B}_{\mathrm{D}}(\sigma_h,u_h,\lambda_h;\gamma\,\tilde{\tau},0,0)\\
&\quad= \gamma\,\frac{1}{\epsilon}(\sigma_h,\tilde{\tau})_{\mathcal{T}_h}+\gamma\,(\nabla u_h,\tilde{\tau})_{\mathcal{T}_h}+\gamma\,\langle\lambda_h-u_h,\tilde{\tau}\rangle_{\partial\mathcal{T}_h}\\
&\quad\geqslant -\frac{1}{2\epsilon}\|\sigma_h\|^2_{\mathcal{T}_h}-\frac{\gamma^2}{2\epsilon}\|\tilde{\tau}\|^2_{\mathcal{T}_h}+\gamma\,\left(\epsilon\|\nabla u_h\|^2_{\mathcal{T}_h}+\frac{\epsilon}{h}|\lambda_h-u_h|^2_{\partial\mathcal{T}_h}\right)\\
&\quad\geqslant -\frac{1}{2\epsilon}\|\sigma_h\|^2_{\mathcal{T}_h}+\left(\gamma-\frac{c_I\gamma^2}{2}\right)\left(\epsilon\|\nabla u_h\|^2_{\mathcal{T}_h}+\frac{\epsilon}{h}|\lambda_h-u_h|^2_{\partial\mathcal{T}_h}\right),
\end{aligned}$$

where we have used (3.7) for the last estimate. The assertion of the proposition now follows by choosing $\gamma = 1/c_I$ and combining the estimates for the two choices of test functions. □

REMARK 3.3 The constant c_{D} in (3.6) depends on the constant c_I of (3.5) and thus on an inverse inequality. To make the dependence on the polynomial degree k explicit let us slightly change the definition of $\tilde{\tau}$ by requiring that $\tilde{\tau}\nu = h^{-1}k^2(\lambda_h-u_h)$ and define the energy norm by $|||\sigma_h,u_h,\lambda_h|||^2_{\mathrm{D}} := \|\sigma_h\|^2_{\mathcal{T}_h}+\|\nabla u_h\|^2_{\mathcal{T}_h}+h^{-1}k^2|\lambda_h-u_h|^2_{\partial\mathcal{T}_h}$. Then one can show that the ellipticity estimate holds with $c_{\mathrm{D}} = \tilde{c}_{\mathrm{D}}k^{-s}$ for $s > 1/2$ and $\tilde{c}_{\mathrm{D}}$ is independent of k. Therefore we will observe suboptimality of the error estimates with respect to the polynomial degree k. Note that the scaling of the jump terms $|\lambda_h-u_h|_{\partial\mathcal{T}_h}$ is the same as the one used in the hp-error analysis of DG methods (cf. Perugia & Schötzau, 2002; Houston *et al.*, 2007).

After using Galerkin orthogonality in the analysis below, we will need the boundedness of $\mathcal{B}_{\mathrm{D}}$ on the larger space $\mathcal{W}\oplus\mathcal{W}_h\times\mathcal{W}_h$.

PROPOSITION 3.4 (Boundedness) There exists a constant C_{D} that is independent of h such that the estimate

$$|\mathcal{B}_{\mathrm{D}}(\sigma,u,\lambda;\tau_h,v_h,\mu_h)| \leqslant C_{\mathrm{D}}|||(\sigma,u,\lambda)|||_{\mathrm{D},*}|||(\tau_h,v_h,\mu_h)|||_{\mathrm{D}} \tag{3.8}$$

holds for all $(\sigma,u,\lambda)\in\mathcal{W}\oplus\mathcal{W}_h$ and $(\tau_h,v_h,\mu_h)\in\mathcal{W}_h$.

Proof. We only consider the term $\langle \lambda - u, \tau_h \nu\rangle_{\partial\mathcal{T}_h}$ in detail. Using the Cauchy–Schwarz and a discrete trace inequality $|\tau_h \nu|_{\partial T} \leqslant \frac{c}{\sqrt{h}}\|\tau_h\|_T$, we obtain $|\langle \lambda - u, \tau_h\nu\rangle_{\partial T}| \leqslant \frac{c}{\sqrt{h}}|\lambda - u|_{\partial\mathcal{T}_h}\|\tau_h\|_T$. The result then follows by standard estimates for the remaining terms and summing up over all elements. □

The above discrete trace inequality cannot be used for the term involving $\sigma\nu$ since $\sigma \in \mathcal{W} \otimes \mathcal{W}_h$. Therefore an additional term appears in the norm $|||\cdot|||_{\mathrm{D},*}$.

3.2 *Pure convection—Method 2.3*

Since Method 2.3 is equivalent to the DG method for hyperbolic problems, our analysis is carried out in a similar manner to that presented in Johnson & Pitkäranta (1986).

PROPOSITION 3.5 (Stability) There exists a constant c_{C} that is independent of the mesh size h such that the estimate

$$\sup_{(v_h,\mu_h)} \frac{\mathcal{B}_{\mathrm{C}}(u_h,\lambda_h; v_h,\mu_h)}{|||(v_h,\mu_h)|||_{\mathrm{C}}} \geqslant c_{\mathrm{C}} |||(u_h,\lambda_h)|||_{\mathrm{C}} \tag{3.9}$$

holds for all $(u_h,\lambda_h) \in \mathcal{V}_h \times \mathcal{M}_h$.

Proof. We start by choosing test functions $v_h = -u_h$ and $\mu_h = -\lambda_h$. Since $\mathrm{div}\beta = 0$, we have $(u_h, \beta\nabla u_h)_T = \frac{1}{2}\langle \beta\nu u_h, u_h\rangle_{\partial T}$ on each element, and thus

$$\begin{aligned}
&\mathcal{B}_{\mathrm{C}}(u_h,\lambda_h; -u_h, -\lambda_h)\\
&\quad = -\frac{1}{2}\langle \beta\nu u_h, u_h\rangle_{\partial\mathcal{T}_h} + \langle \beta\nu\{\lambda_h/u_h\}, u_h\rangle_{\partial\mathcal{T}_h} - \langle \beta\nu\{\lambda_h/u_h\}, \lambda_h\rangle_{\partial\mathcal{T}_h}\\
&\quad = (1) + (2) + (3) = (*).
\end{aligned}$$

Recall that λ_h equals 0 on $\partial\Omega$, and let us rearrange the terms (1)–(3) in the following way:

$$\begin{aligned}
(1) &= -\frac{1}{2}\langle \beta\nu u_h, u_h\rangle_{\partial\mathcal{T}_h} = \frac{1}{2}|\beta\nu||u_h|^2_{\partial\mathcal{T}_h^{\mathrm{in}}} - \frac{1}{2}|\beta\nu||u_h|^2_{\partial\mathcal{T}_h^{\mathrm{out}}},\\
(2) &= \langle \beta\nu\{\lambda_h/u_h\}, u_h\rangle_{\partial\mathcal{T}_h} = |\beta\nu||u_h|^2_{\partial\mathcal{T}_h^{\mathrm{out}}} - |\beta\nu|\langle \lambda_h, u_h\rangle_{\partial\mathcal{T}_h^{\mathrm{in}}},\\
(3) &= -\langle \beta\nu\{\lambda_h/u_h\}, \lambda_h\rangle_{\partial\mathcal{T}_h} = |\beta\nu||\lambda_h|^2_{\partial\mathcal{T}_h^{\mathrm{in}}} - |\beta\nu|\langle \lambda_h, u_h\rangle_{\partial\mathcal{T}_h^{\mathrm{out}}}.
\end{aligned}$$

Now let T_1 and T_2 denote two elements sharing the facet $E = \partial T_1^{\mathrm{out}} \cap \partial T_2^{\mathrm{in}}$. Since λ_h is single valued on E by definition, we have $\lambda_h|_{\partial T_1^{\mathrm{out}}} = \lambda_h|_{\partial T_2^{\mathrm{in}}}$, which means that we can shift the terms only involving the Lagrange multiplier between neighbouring elements. Summing up, we obtain that

$$(*) = \frac{1}{2}|\beta\nu||\lambda_h - u_h|^2_{\partial\mathcal{T}_h}.$$

Let us now include a second term in the stability estimate by testing the bilinear form with $v_h = -\gamma\frac{h}{|\beta|}\beta\nabla u_h$ for some $\gamma \geqslant 0$, which yields

$$
\begin{aligned}
\mathcal{B}_{\mathrm{C}}(u_h, \lambda_h; v_h, 0) &= -\frac{\gamma h}{|\beta|}(u_h, \beta\nabla(\beta\nabla u_h))_{\mathcal{T}_h} + \frac{\gamma h}{|\beta|}\langle \beta\nu\{\lambda_h/u_h\}, \beta\nabla u_h\rangle_{\partial\mathcal{T}_h} \\
&= \frac{\gamma h}{|\beta|}\|\beta\nabla u_h\|^2_{\mathcal{T}_h} + \frac{\gamma h}{|\beta|}\langle \beta\nu(\lambda_h - u_h), \beta\nabla u_h\rangle_{\partial\mathcal{T}_h^{\mathrm{in}}} \\
&\geqslant c\gamma\left(\frac{h}{|\beta|}\|\beta\nabla u_h\|^2_{\mathcal{T}_h} - |\beta\nu||\lambda_h - u_h|^2_{\partial\mathcal{T}_h}\right).
\end{aligned}
$$

For the last estimate we used Young's inequality and a discrete trace inequality. The result now follows by choosing $\gamma = \frac{1}{4c}$ and combining the estimates for the two different test functions. Note that, by inverse inequalities and due to our scaling of v_h with $h/|\beta|$, it follows that $|||(v_h, 0)|||_{\mathrm{C}} \leqslant C|||(u_h, 0)|||_{\mathrm{C}}$ with a constant C that is independent of the mesh size. □

PROPOSITION 3.6 (Boundedness) There exists a constant C_{C} that is independent of h such that the estimate

$$|\mathcal{B}_{\mathrm{C}}(u, \lambda; v_h, \mu_h)| \leqslant C_{\mathrm{C}}|||(u, \lambda)|||_{\mathrm{C},*}|||(v_h, \mu_h)|||_{\mathrm{C}} \tag{3.10}$$

holds for all $u \in \mathcal{V} \oplus \mathcal{V}_h$, $\lambda \in \mathcal{M} \oplus \mathcal{M}_h$ and $(v_h, \mu_h) \in \mathcal{V}_h \times \mathcal{M}_h$.

Proof. The assertion follows directly from the definition of the norms and the Cauchy–Schwarz inequality. □

3.3 *Convection–diffusion—Method 2.4*

Due to the structure of Method 2.4 as the combination of Methods 2.1 and 2.3, the stability and boundedness of the bilinear form (2.12) follow almost directly from the corresponding properties of the bilinear forms for the limiting subproblems. The appropriate norms for the analysis of Method 2.4 are given by

$$|||(\sigma_h, u_h, \lambda_h)||| = (|||(\sigma_h, u_h, \lambda_h)|||^2_{\mathrm{D}} + |||(u_h, \lambda_h)|||^2_{\mathrm{C}})^{1/2} \tag{3.11}$$

and

$$|||(\sigma, u, \lambda)|||_* = (|||(\sigma, u, \lambda)|||^2_{\mathrm{D},*} + |||(u, \lambda)|||^2_{\mathrm{C},*})^{1/2}, \tag{3.12}$$

i.e., they are just assembled from the norms used for the analysis of the elliptic and hyperbolic subproblems. Note that all terms in the norm scale appropriately. For example, in the diffusion-dominated case ($|\beta|h \leqslant \epsilon$) the terms coming from the convective part can be absorbed by the terms stemming from the stability of the diffusion part. Let us now state the properties of $\mathcal{B}$ in detail.

PROPOSITION 3.7 (Stability) There exists a positive constant c_{B} not depending on the mesh size h such that

$$\sup_{(\tau_h, v_h, \mu_h)} \frac{\mathcal{B}(\sigma_h, u_h, \lambda_h; \tau_h, v_h, \mu_h)}{|||(\tau_h, v_h, \mu_h)|||} \geqslant c_{\mathrm{B}}|||(\sigma_h, u_h, \lambda_h)||| \tag{3.13}$$

holds for all $(\sigma_h, u_h, \lambda_h) \in \Sigma_h \times \mathcal{V}_h \times \mathcal{M}_h$.

Proof. We will show the inf–sup stability by testing with the functions used in the previous stability estimates, i.e., $\tau_h = \sigma_h + \alpha\tilde{\tau}$, $v_h = -u_h + \gamma\frac{h}{|\beta|}\beta\nabla u_h$ and $\mu_h = -\lambda_h$. In view of Propositions 3.2 and

3.5, it only remains to estimate the additional term coming from the test function $\gamma \frac{h}{|\beta|}\beta\nabla u_h$ inserted in the diffusion bilinear form, namely,

$$\begin{aligned}\mathcal{B}_{\mathrm{D}0}\left(\sigma_h, u_h, \lambda_h; 0, \gamma \frac{h}{|\beta|}\beta\nabla u_h, 0\right) &= -\gamma \frac{h}{|\beta|}(\sigma_h, \nabla(\beta\nabla u_h)) + \gamma \frac{h}{|\beta|}\langle \sigma_h \nu, \beta\nabla u_h\rangle \\ &= \gamma \frac{h}{|\beta|}(\mathrm{div}\sigma_h, \beta\nabla u_h) \geqslant -\gamma \frac{h}{|\beta|}\|\mathrm{div}\sigma_h\|\|\beta\nabla u_h\| \\ &\geqslant -c\gamma\left(\frac{1}{\epsilon}\|\sigma_h\|^2 + \epsilon\|\nabla u_h\|^2\right) \geqslant -c\gamma\, |||(\sigma_h, u_h, \lambda_h)|||^2_{\mathrm{D}}.\end{aligned}$$

This term can be absorbed by the stability estimate for the diffusion problem as long as γ is chosen to be sufficiently small. Note that γ does not depend on h, ϵ or β, i.e., the stability constant c_{B} does not depend on these parameters. □

The boundedness of the bilinear form follows directly by combining the two results for the limiting subproblems.

COROLLARY 3.8 (Boundedness) There exists a constant C_{B} that is independent of the mesh size h such that

$$|\mathcal{B}(\sigma, u, \lambda; \tau_h, v_h, \mu_h)| \leqslant C_{\mathrm{B}} |||(\sigma, u, \lambda)|||_* |||(\tau_h, v_h, \mu_h)||| \tag{3.14}$$

holds for all $(\sigma, u, \lambda) \in \mathcal{W} \oplus \mathcal{W}_h$ and $(\tau_h, v_h, \mu_h) \in \mathcal{W}_h$.

As a last ingredient for deriving the *a priori* error estimates, we have to establish some approximation properties of our finite-dimensional spaces with respect to the norms under consideration.

3.4 *Interpolation operators and approximation properties*

Let us start by introducing appropriate interpolation operators and then recall some basic interpolation error estimates. For $T \in \mathcal{T}_h$, $E \in \mathcal{E}_h$ and functions $u \in L^2(T)$ and $\lambda \in L^2(E)$ we define the local L^2-projections $\Pi_k^T u$ and $\Pi_k^E \lambda$ by

$$(u - \Pi_k^T u, v_h)_T = 0 \quad \forall v_h \in \mathcal{P}_k(T)$$

and

$$(\lambda - \Pi_k^E \lambda, \mu_h)_E = 0 \quad \forall \mu_h \in \mathcal{P}_k(E),$$

respectively. These interpolation operators satisfy the following error estimates (cf. Brenner & Scott, 2002).

LEMMA 3.9 Let Π_k^T and Π_k^E be defined as above. Then the estimates

$$\begin{aligned}\|u - \Pi_k^T u\|_T &\leqslant Ch^s|u|_{s,T}, && 0 \leqslant s \leqslant k+1,\\ \|\nabla(u - \Pi_k^T u)\|_T &\leqslant Ch^s|u|_{s+1,T}, && 0 \leqslant s \leqslant k,\\ \|u - \Pi_k^T u\|_{\partial T} + \|u - \Pi_k^E u\|_{\partial T} &\leqslant Ch^{s+1/2}|u|_{s+1,T}, && 0 \leqslant s \leqslant k,\end{aligned}$$

hold with a constant C that is independent of h.

The corresponding interpolation operators for functions on $\mathcal{T}_h$ and $\mathcal{E}_h$ are defined element-wise and are denoted by the same symbols.

For the flux function σ we utilize the Raviart–Thomas interpolant defined by

$$(\sigma - \Pi_k^{\mathrm{RT}}\sigma, p_h)_T = 0 \qquad \forall\, p_h \in [\mathcal{P}_{k-1}(T)]^d,$$

$$((\sigma - \Pi_k^{\mathrm{RT}}\sigma)\nu, \mu_h)_E = 0 \qquad \forall\, \mu_h \in \mathcal{P}_k(E), \quad E \subset \partial T.$$

In order to make moments of $\sigma\nu$ to be well defined on single facets E, one has to require some extra regularity, for example, $\sigma \in H(\mathrm{div}, T) \cap L^s(T)$ for some $s > 2$ or $\sigma \in H^{1/2+\varepsilon}(T)$ (cf. Brezzi & Fortin, 1991). Under such an assumption, the following interpolation error estimates hold (Brezzi & Fortin, 1991; Toselli & Widlund, 2005).

LEMMA 3.10 Let Π_k^{RT} be defined as above. Then the estimates

$$\|\sigma - \Pi_k^{\mathrm{RT}}\sigma\|_T + h^{1/2}\|(\sigma - \Pi_k^{\mathrm{RT}}\sigma)\nu\|_{\partial T} \leqslant Ch^s|\sigma|_{s,T}, \qquad 1/2 < s \leqslant k+1,$$

$$\|\mathrm{div}(\sigma - \Pi_k^{\mathrm{RT}}\sigma)\|_T \leqslant Ch^s|\mathrm{div}\sigma|_{s,T}, \quad 1 \leqslant s \leqslant k+1,$$

hold with a constant C that is independent of h.

Applying these results element-wise, we immediately obtain the following interpolation error estimates for the mesh-dependent norms used above.

PROPOSITION 3.11 Let $u \in H^1(\Omega) \cap H^{3/2+\varepsilon}(\mathcal{T}_h)$ and set $\sigma := -\epsilon\nabla u$. Then

$$|||(\sigma - \Pi_k^{\mathrm{RT}}\sigma, u - \Pi_k^T u, \lambda - \Pi_k^{\mathrm{RT}}u)|||_{\mathrm{D},*} \leqslant Ch^s\sqrt{\epsilon}|u|_{s+1,\mathcal{T}_h}, \quad 1/2 < s \leqslant k, \tag{3.15}$$

and for $u \in H^1(\Omega)$ we have

$$|||(u - \Pi_k^T u, \lambda - \Pi_k^{\mathrm{RT}}u)|||_{\mathrm{C},*} \leqslant Ch^{s+1/2}\sqrt{|\beta|}|u|_{s+1,\mathcal{T}_h}, \quad 0 \leqslant s \leqslant k, \tag{3.16}$$

with constants C not depending on u or h. The same estimates hold if the $*$-norms are replaced by their counterparts without $*$.

REMARK 3.12 The estimates of Proposition 3.11 hold with obvious modifications if the smoothness s or the polynomial degree k varies locally. We assume uniform polynomial degree and smoothness only for ease of notation here.

The interpolation error estimate (3.15) is suboptimal regarding the approximation capabilities of the flux interpolant. In fact, by Lemma 3.10, one can obtain

$$\frac{1}{\sqrt{\epsilon}}\|\sigma - \Pi_k^{\mathrm{RT}}\sigma\| \leqslant Ch^s\sqrt{\epsilon}|u|_{s+1,\mathcal{T}_h} \quad \text{for } 1/2 < s \leqslant k+1,$$

and so the best possible rate is h^{k+1} instead of h^k as for $|||\cdot|||_{\mathrm{D}}$ in (3.15). We will use this fact in Section 4 to derive superconvergence results for the primal variable u_h.

3.5 *A priori error estimates*

The error of the finite-element approximation can be decomposed into an approximation error and a discrete error. Let $(\sigma_h, u_h, \lambda_h)$ denote the discrete solution of (2.11), and let u be the solution of (2.10) and define $\sigma := -\epsilon \nabla u$. Then we have

$$
\begin{aligned}
&|||(\sigma - \sigma_h, u - u_h, u - \lambda_h)||| \\
&\quad \leqslant |||(\sigma - \Pi_k^{\mathrm{RT}}\sigma, u - \Pi_k^T u, u - \Pi_k^E u)||| + |||(\Pi_k^{\mathrm{RT}}\sigma - \sigma_h, \Pi_k^T u - u_h, \Pi_k^E u - \lambda_h)|||. \qquad (3.17)
\end{aligned}
$$

Using the stability and boundedness of the bilinear form and applying Galerkin orthogonality, the second term can now also be estimated by the interpolation error.

PROPOSITION 3.13 Let $(\sigma_h, u_h, \lambda_h) \in \mathcal{W}_h$ denote the solution of (2.11), and let $u \in H^1(\Omega) \cap H^{3/2+\varepsilon}(\mathcal{T}_h)$ be the solution of the convection–diffusion problem (2.10). Then there exists a constant C that is independent of the mesh size h such that the estimate

$$
|||(\Pi_k^{\mathrm{RT}}(-\epsilon\nabla u) - \sigma_h, \Pi_k^T u - u_h, \Pi_k^E u - \lambda_h)||| \leqslant Ch^s(\sqrt{\epsilon} + h^{1/2}\sqrt{|\beta|})|u|_{s+1,\mathcal{T}_h}
$$

holds for $1/2 < s \leqslant k$.

Proof. Let us define $\sigma = -\epsilon\nabla u$ and $\lambda = u$. By an application of the stability estimate (3.6), Galerkin orthogonality and the boundedness (3.8) of the bilinear form, we obtain that

$$
\begin{aligned}
&c_{\mathrm{B}}|||(\Pi_k^{\mathrm{RT}}\sigma - \sigma_h, \Pi_k^T u - u_h, \Pi_k^E u - \lambda_h)||| \\
&\quad \leqslant \sup_{(\tau_h, v_h, \mu_h)} \mathcal{B}(\Pi_k^{\mathrm{RT}}\sigma - \sigma_h, \Pi_k^T u - u_h, \Pi_k^E u - \lambda_h; \tau_h, v_h, \mu_h)/|||(\tau_h, v_h, \mu_h)||| \\
&\quad = \sup_{(\tau_h, v_h, \mu_h)} \mathcal{B}(\Pi_k^{\mathrm{RT}}\sigma - \sigma, \Pi_k^T u - u, \Pi_k^E u - u; \tau_h, v_h, \mu_h)/|||(\tau_h, v_h, \mu_h)||| \\
&\quad \leqslant C_{\mathrm{B}}|||(\Pi_k^{\mathrm{RT}}\sigma - \sigma, \Pi_k^T u - u, \Pi_k^E u - u)|||_*.
\end{aligned}
$$

The assertion follows directly from (3.15). □

The complete error estimate can now be derived by combining (3.17) and Proposition 3.11.

THEOREM 3.14 (Energy norm estimate) Let $(\sigma_h, u_h, \lambda_h)$ be the finite-element solution of Method 2.4, and let $u \in H^1(\Omega) \cap H^{3/2+\varepsilon}(\mathcal{T}_h)$ denote the solution of (2.10) and $\sigma := -\epsilon\nabla u$. Then

$$
|||(\sigma - \sigma_h, u - u_h, u - \lambda_h)||| \leqslant Ch^s(\sqrt{\epsilon} + h^{1/2}\sqrt{|\beta|})|u|_{s+1,\mathcal{T}_h}
$$

holds for $1/2 < s \leqslant k$ with a constant C that is independent of the mesh size h.

In the convection-dominated case the error estimate coincides with the well-known error estimates for the DG and the streamline diffusion method for hyperbolic problems (cf. Johnson & Pitkäranta, 1986; Johnson & Saranen, 1986).

COROLLARY 3.15 Let $\epsilon \leqslant |\beta|h$ on each element, and let the conditions of Theorem 3.14 hold. Then the estimate

$$
|||(\sigma - \sigma_h, u - u_h, u - \lambda_h)||| \leqslant Ch^{s+1/2}\sqrt{|\beta|}|u|_{s+1,\mathcal{T}_h}
$$

holds for $1/2 < s \leqslant k$ with a constant C that is independent of the parameters ϵ, β and h.

This estimate holds, in particular, for the limiting hyperbolic problem ($\epsilon \equiv 0$), in which case $\sigma = \sigma_h \equiv 0$ and $|||(\tau, v, \mu)||| = |||(v, \mu)|||_{\mathrm{C}}$, and so Method 2.4 collapses to Method 2.3, i.e., the DG method for hyperbolic problems.

By analogy with standard error estimates for mixed methods for the Poisson problem, we obtain the following convergence result in the diffusion-dominated regime.

COROLLARY 3.16 Let $\epsilon \geqslant |\beta| h$ and let the conditions of Theorem 3.14 hold. Then the estimate

$$|||(\sigma - \sigma_h, u - u_h, \lambda - \lambda_h)||| \leqslant C h^s \sqrt{\epsilon} |u|_{s+1,\mathcal{T}_h}$$

holds for $1/2 < s \leqslant k$ with a constant C that is independent of ϵ, β and h. Moreover, we have $|||\cdot|||_{\mathrm{D}} \sim |||\cdot|||$.

Clearly, this estimate also holds for Method 2.1 in the case of pure diffusion. Let us remark once again that all terms in the *a priori* error estimates are defined locally, and so the smoothness index s and the polynomial degree k can vary locally, allowing for hp-adaptivity.

4. Superconvergence and postprocessing for diffusion-dominated problems

The best possible rate for $\frac{1}{\sqrt{\epsilon}}\|\sigma - \sigma_h\|$ guaranteed by Theorem 3.14 and Corollary 3.16 is h^k, which is one order suboptimal regarding the interpolation error estimate of Lemma 3.10. It is well known, however, that in the purely elliptic case the optimal rate h^{k+1} can be obtained by a refined analysis, and we will derive corresponding results below. Since we consider the case of dominating diffusion in this section, we assume for ease of notation that $\epsilon \equiv 1$ in what follow.

4.1 *Refined analysis for pure diffusion*

Although the estimate (3.15) is optimal concerning the approximation error with respect to the norm $|||\cdot|||_{\mathrm{D}}$, we can obtain better error estimates for $\sigma = -\nabla u$, i.e., we will show that $\|\sigma - \sigma_h\|$ depends only on the interpolation error $\|\sigma - \Pi_k^{\mathrm{RT}}\sigma\|$, and thus optimal convergence for σ_h can be expected. We refer to Arnold & Brezzi (1985), Brezzi & Fortin (1991) and Stenberg (1991) for corresponding results in the mixed framework.

PROPOSITION 4.1 Let $(\sigma_h, u_h, \lambda_h)$ denote the solution of (2.3) and let u and $\sigma := -\nabla u$ be the solution of problem (2.2). Then

$$|||(\sigma_h - \sigma, u_h - \Pi_k^T u, \lambda_h - \Pi_k^E u)|||_{\mathrm{D}} \leqslant C h^s |u|_{s+1,\mathcal{T}_h} \tag{4.1}$$

holds for $1/2 < s \leqslant k+1$ with a constant C that is independent of h.

Proof. Let us first consider the following term:

$$\begin{aligned}
&B_{\mathrm{D}}(\Pi_k^{\mathrm{RT}}\sigma - \sigma, \Pi_k^T u - u, \Pi_k^E u - u; \tau_h, v_h, \lambda_h) \\
&= (\Pi_k^{\mathrm{RT}}\sigma - \sigma, \tau_h)_{\mathcal{T}_h} - (\Pi_k^T u - u, \mathrm{div}\tau_h)_{\partial\mathcal{T}_h} + \langle \Pi_k^E u - u, \tau_h \nu\rangle_{\partial\mathcal{T}_h} \\
&\quad + (\mathrm{div}(\Pi_k^{\mathrm{RT}}\sigma - \sigma), v_h)_{\mathcal{T}_h} + \langle (\Pi_k^{\mathrm{RT}}\sigma - \sigma)\nu, \mu_h\rangle_{\partial\mathcal{T}_h} \\
&= (\Pi_k^{\mathrm{RT}}\sigma - \sigma, \tau_h)_{\mathcal{T}_h},
\end{aligned}$$

where the last equality follows from the definition of the interpolants. Then, in the same manner as in the proof of Proposition 3.13, we obtain that

$$c_{\mathrm{D}} |||(\Pi_k^{\mathrm{RT}}\sigma - \sigma_h, \Pi_k^T u - u_h, \Pi_k^E u - \lambda_h)|||_{\mathrm{D}} \leqslant \|\Pi_k^{\mathrm{RT}}\sigma - \sigma\|_{\mathcal{T}_h},$$

and the statement follows by an application of the triangle inequality and the interpolation error estimate (3.15). □

Note that, for the modified error (4.1), the best possible rate now is h^{k+1}, which is optimal in view of the interpolation error estimates. As we show next, the estimates for $(\Pi_k^T u - u_h)$ and $(\Pi_k^E u - \lambda_h)$ can even be improved if we assume that the domain Ω is convex (cf. Stenberg (1991) for similar results in the mixed framework).

PROPOSITION 4.2 Let Ω be convex and $u \in H^1(\Omega) \cap H^{3/2+\varepsilon}(\mathcal{T}_h)$ be the solution of (2.2). Moreover, let u_h denote the discrete solution obtained by Method 2.1. Then the estimate

$$\|\Pi_k^T u - u_h\|_0 \leqslant Ch^{s+1} \begin{cases} |u|_{s+2,\mathcal{T}_h}, & k = 0, \\ |u|_{s+1,\mathcal{T}_h}, & k > 0, \end{cases} \tag{4.2}$$

holds for $1/2 < s \leqslant k+1$ when $k > 0$ and $0 \leqslant s \leqslant 1$ when $k = 0$. If, in addition, f is piecewise constant then

$$\|\Pi_0^T u - u_h\|_0 \leqslant Ch^{s+1}|u|_{s+1,\mathcal{T}_h} \tag{4.3}$$

also holds for $k = 0$.

Proof. Let $\phi \in H_0^1(\Omega)$ denote the solution of the Poisson equation $\Delta\phi = \Pi_k^T u - u_h$ with homogeneous Dirichlet conditions and let $z := \nabla\phi$. Due to the convexity of Ω, we have

$$\|\phi\|_{2,\Omega} \leqslant c\|\Pi_k^T u - u_h\|_0 \quad \text{and} \quad \|\phi - \Pi_k^T \phi\| \leqslant ch^{\min(k+1,2)}\|\Pi_k^T u - u_h\|_0.$$

Using the definition of ϕ and z, we obtain that

$$\begin{aligned} \|\Pi_k^T u - u_h\|_0^2 &= (\Pi_k^T u - u_h, \mathrm{div} z) = (\Pi_k^T u - u_h, \mathrm{div}(\Pi_k^{\mathrm{RT}} z)) = (\sigma - \sigma_h, \Pi_k^{\mathrm{RT}} z) \\ &= (\sigma - \sigma_h, \Pi_k^{\mathrm{RT}} z - \nabla\phi) - (\mathrm{div}(\sigma - \sigma_h), \phi - \Pi_k^T \phi) \\ &\leqslant \|\sigma - \sigma_h\|_0 \|\Pi_k^{\mathrm{RT}} z - \nabla\phi\|_0 + \|\mathrm{div}(\sigma - \sigma_h)\|_0 \|\phi - \Pi_k^T \phi\|_0. \end{aligned}$$

The first estimate now follows by Lemma 3.10. If f is piecewise polynomial of order k then $\mathrm{div}(\sigma - \sigma_h) \equiv 0$, and so the last term in the above estimate vanishes and we conclude the second assertion. □

4.2 *The diffusion-dominated case*

Let us now show that similar results still hold in the presence of convection as long as diffusion is sufficiently dominating. In this case we can discretize the convective term without upwind stabilization, and we therefore consider the following bilinear form instead of (2.8):

$$\mathcal{B}_{\mathrm{C}}^{\mathrm{NU}}(u_h, \lambda_h; v_h, \mu_h) := (u_h, \beta\nabla v_h)_{\mathcal{T}_h} + \langle \beta\nu\lambda_h, \mu_h - v_h\rangle_{\partial\mathcal{T}_h}. \tag{4.4}$$

Such a discretization for the convective part was previously investigated numerically but not analysed in Farhoul & Mounim (2005) for a one-dimensional problem. There the authors conjectured that this discretization already introduces some stabilization, which is not the case, as is clear from our analysis.

Consistency and conservation. Substituting the continuous solution u for u_h and λ_h in (4.4), we obtain after integration by parts that

$$-(\mathrm{div}(\beta u), v_h)_{\mathcal{T}_h} + \langle \beta\nu u, \mu_h\rangle_{\partial\mathcal{T}_h} = -(\mathrm{div}(\beta u), v_h)_{\mathcal{T}_h} = (-f, v_h)_{\mathcal{T}_h},$$

and so the bilinear form $\mathcal{B}_{\mathrm{C}}^{\mathrm{NU}}$ is consistent. The scheme is also conservative since the flux $\beta\nu\lambda_h$ in (4.4) is single valued on element interfaces. Moreover, we have

$$\mathcal{B}_{\mathrm{C}}(u_h, \lambda_h; v_h, \mu_h) = \mathcal{B}_{\mathrm{C}}^{\mathrm{NU}}(u_h, \lambda_h; v_h, \mu_h) + |\beta\nu|\langle \lambda_h - u_h, \mu_h - v_h\rangle_{\partial\mathcal{T}_h^{\mathrm{out}}}, \tag{4.5}$$

which clarifies what kind of upwind was used for the DG stabilization in (2.8).

Stability. Testing the bilinear form $\mathcal{B}_{\mathrm{C}}^{\mathrm{NU}}$ with $v_h = -u_h$ and $\mu_h = -\lambda_h$, we obtain that

$$\begin{aligned}\mathcal{B}_{\mathrm{C}}^{\mathrm{NU}}(u_h, \lambda_h; -u_h, -\lambda_h) &= -(u_h, \beta\nabla u_h)_{\mathcal{T}_h} - \langle \beta\nu\lambda_h, \lambda_h - u_h\rangle_{\partial\mathcal{T}_h}\\ &= -\frac{1}{2}\langle \beta\nu u_h, u_h\rangle_{\partial\mathcal{T}_h} - \langle \beta\nu\lambda_h, \lambda_h - u_h\rangle_{\partial\mathcal{T}_h}\\ &= \frac{1}{2}|\beta\nu||u_h - \lambda_h|^2_{\partial\mathcal{T}_h^{\mathrm{in}}} - \frac{1}{2}|\beta\nu||u_h - \lambda_h|^2_{\partial\mathcal{T}_h^{\mathrm{out}}}.\end{aligned}$$

Note that, by adding the stabilization term $|\beta\nu||u_h - \lambda_h|^2_{\partial\mathcal{T}_h^{\mathrm{out}}}$, the last term becomes strictly positive, i.e.,

$$\begin{aligned}\mathcal{B}_{\mathrm{C}}(u_h, \lambda_h; -u_h, -\lambda_h) &= \mathcal{B}_{\mathrm{C}}^{\mathrm{NU}}(u_h, \lambda_h; -u_h, -\lambda_h) + |\beta\nu||\lambda_h - u_h|^2_{\partial\mathcal{T}_h^{\mathrm{out}}}\\ &= \frac{1}{2}|\beta\nu||\lambda_h - u_h|^2_{\partial\mathcal{T}_h},\end{aligned}$$

and we recover the first part of the stability estimate of Proposition 3.5.

Following the approach for the convection-dominated case, we now consider the following method for the diffusion-dominated regime (cf. also Farhoul & Mounim, 2005).

METHOD 4.3 (No upwind) Find $(\sigma_h, u_h, \lambda_h) \in \mathcal{W}_h$ such that

$$\mathcal{B}^{\mathrm{NU}}(\sigma_h, u_h, \lambda_h; \tau_h, v_h, \mu_h) = \mathcal{F}(v_h, \mu_h) \tag{4.6}$$

holds for all $(\tau_h, v_h, \mu_h) \in \mathcal{W}_h$, where $\mathcal{B}^{\mathrm{NU}} := \mathcal{B}_{\mathrm{D}} + \mathcal{B}_{\mathrm{C}}^{\mathrm{NU}}$.

For the proof of stability of the bilinear form B^{NU} we require that the convection is sufficiently small. A sufficient condition is given by

$$|\beta\nu||\lambda_h - u_h|^2_{\partial\mathcal{T}_h} \leqslant c_{\mathrm{D}}|||(\sigma_h, u_h, \lambda_h)|||^2_{\mathrm{D}} \quad \forall\,(\sigma_h, u_h, \lambda_h) \in \mathcal{W}_h. \tag{4.7}$$

REMARK 4.4 Recall that the stability constant c_{D} and thus the validity of condition (4.7) depend only on the constant of an inverse inequality and thus on the shape of the elements. Moreover, since both norms are defined element-wise, it is possible to decide for each element separately if stabilization should be added or not. Clearly, (4.7) can be shown to hold if $|\beta|h \leqslant c_T\epsilon$ is valid on each element, with the constant c_T only depending on the shape of the individual elements.

Using (4.7) as the characterization of dominating diffusion, we can now prove the following stability result.

PROPOSITION 4.5 Let (4.7) be valid. Then the estimate

$$\sup_{(\tau_h, v_h, \mu_h)} \frac{\mathcal{B}^{\mathrm{NU}}(\sigma_h, u_h, \lambda_h; \tau_h, v_h, \mu_h)}{|||(\tau_h, v_h, \mu_h)|||_{\mathrm{D}}} \geqslant \frac{c_{\mathrm{D}}}{2} |||(\sigma_h, u_h, \lambda_h)|||_{\mathrm{D}}^2 \tag{4.8}$$

holds for all $(\sigma_h, u_h, \lambda_h) \in \mathcal{W}_h$ with c_{D} denoting the stability constant of Proposition 3.2.

Since the convective terms can be absorbed by the diffusion terms, the boundedness result of Corollary 3.8 applies with $|||\cdot|||_{(*)}$ replaced by $|||\cdot|||_{\mathrm{D},(*)}$. Using the stability estimate (4.8), the following *a priori* error estimate is obtained in a similar manner as Proposition 4.1 for the purely elliptic case.

PROPOSITION 4.6 Let condition (4.7) be valid and $(\sigma_h, u_h, \lambda_h)$ denote the solution of Method 4.3. Moreover, let $u \in H^1(\Omega) \cap H^{3/2+\varepsilon}(\mathcal{T}_h)$ denote the solution of problem (2.10) and set $\sigma := -\nabla u$. Then

$$|||(\sigma_h - \sigma, u_h - \Pi_k^T u, \lambda_h - \Pi_k^E u)|||_{\mathrm{D}} \leqslant C h^s |u|_{s+1,\Omega}$$

holds for all $1/2 < s \leqslant k+1$ with a constant C that is independent of h.

Proof. In view of Proposition 4.1, we only have to ensure that the convective term does not disturb the estimate. Following the proof of Proposition 4.1, i.e., testing with the same test functions as there, we obtain the additional term

$$\mathcal{B}_{\mathrm{C}}^{\mathrm{NU}}(\Pi_k^T u - u, \Pi_k^E u - u; v_h, \mu_h) = (\Pi_k^T u - u, \beta \nabla v_h)_{\mathcal{T}_h} + \langle \beta\nu(\Pi_k^E u - u), \mu_h - v_h \rangle_{\partial\mathcal{T}_h} = 0$$

since $\beta\nabla v_h \in \mathcal{P}_k(T)$ on each element and $\beta\nu(\mu_h - v_h) \in \mathcal{P}_k(E)$ for each facet. The result now follows along the lines of the proof of Proposition 4.1. □

Proposition 4.6 allows us to derive a superconvergence estimate for $\|\Pi_k^T u - u_h\|_{\mathcal{T}_h}$ as in the purely elliptic case.

PROPOSITION 4.7 Let Ω be convex and u be the solution of (2.10) with β satisfying (4.7). Moreover, let u_h denote the discrete solution of Method 4.3. Then

$$\|\Pi_k^T u - u_h\|_{\mathcal{T}_h} \leqslant C h^{s+1} \begin{cases} |u|_{s+2,\mathcal{T}_h}, & k = 0, \\ |u|_{s+1,\mathcal{T}_h}, & k > 0, \end{cases}$$

holds for $1/2 < s \leqslant k+1$ when $k > 0$ and $0 \leqslant s \leqslant 1$ when $k = 0$.

Proof. By means of Proposition 4.6, the result follows in the same way as for Proposition 4.2. □

Due to the lack of a condition $\operatorname{div}(\sigma - \sigma_h) \equiv 0$, which is valid in the purely elliptic case, we cannot obtain (4.3) here. So, in the lowest order case, superconvergence holds only under some additional smoothness of the solution u.

4.3 *Postprocessing*

The superconvergence results of the Section 4.2 can now be utilized to construct better approximations $\tilde{u}_h \in \mathcal{P}_{k+1}(\mathcal{T}_h)$ by local postprocessing. Here we follow an approach proposed by Stenberg (1991) for the mixed discretization of the Poisson equation (2.2) and construct our postprocessed solution from the approximations of the primal and the dual variables. Alternative approaches based on the Lagrange multipliers can be found in Arnold & Brezzi (1985) and Brezzi & Fortin (1991).

Let us define $\tilde{u}_h \in \mathcal{P}_{k+1}(\mathcal{T}_h)$ element-wise by the variational problems

$$(\nabla u_h^*, \nabla v)_T = -(\sigma_h, \nabla v)_T \quad \forall\, v \in \mathcal{P}_{k+1}(T)\colon (v, 1)_T = 0,$$
$$(u_h^*, 1)_E = (u_h, 1)_T.$$

Then the following order optimal error estimate holds.

PROPOSITION 4.8 Let Ω be convex and u denote the solution of (2.10) with (4.7) being valid. Moreover, let $(\sigma_h, u_h, \lambda_h)$ be the solution of Method 4.3 and u_h^* be defined as above. Then

$$\|\nabla(u_h^* - u)\|_{\mathcal{T}_h} \leqslant Ch^s |u|_{s+1,\mathcal{T}_h}$$

and

$$\|u_h^* - u\|_{\mathcal{T}_h} \leqslant Ch^{s+1} \begin{cases} |u|_{s+2,\mathcal{T}_h}, & k = 0, \\ |u|_{s+1,\mathcal{T}_h}, & k > 0, \end{cases}$$

for all $1/2 < s \leqslant k+1$ with a constant C that is independent of the mesh size h. For $k = 0$ the second estimate holds for $0 \leqslant s \leqslant 1$.

Proof. Let $\tilde{u}_h \in H^1(\Omega) \cap \mathcal{P}_{k+1}(\mathcal{T}_h)$ denote the finite-element solution of the standard H^1-conforming finite-element method applied to the solution of (2.2). Then $\|\nabla(u - \tilde{u}_h)\|_{\mathcal{T}_h} \leqslant Ch^s |u|_{s+1,\mathcal{T}_h}$ for $0 \leqslant s \leqslant k+1$. Moreover, $\|u - \tilde{u}_h\| \leqslant Ch^{s+1}|u|_{s+1,\mathcal{T}_h}$ for $0 \leqslant s \leqslant k+1$ since we assumed convexity of Ω and $f \in L_2$. Now define $\tilde{v}_h := (I - \Pi_0^T)(\tilde{u}_h - u_h^*)$. Then

$$\begin{aligned}\|\nabla \tilde{v}_h\|_T^2 &= (\nabla(I - \Pi_0^T)(\tilde{u}_h - u_h^*), \nabla \tilde{v}_h)_T = (\nabla(\tilde{u}_h - u_h^*), \nabla \tilde{v}_h)_T \\ &= (\nabla(\tilde{u}_h - u), \nabla \tilde{v}_h)_T + (\nabla u + \sigma_h, \nabla \tilde{v}_h)_T \\ &\leqslant \|\nabla \tilde{v}_h\|_T (\|\nabla(u - \tilde{u}_h)\|_T + \|\sigma_h + \nabla u\|)_T.\end{aligned}$$

Summing up over all elements and using the estimates for $(u - \tilde{u}_h)$ and Proposition 4.6 yields

$$\begin{aligned}\|\nabla(u - u_h^*)\|_{\mathcal{T}_h} &\leqslant \|\nabla(u - \tilde{u}_h)\|_{\mathcal{T}_h} + \|\nabla(\tilde{u}_h - u_h^*)\|_{\mathcal{T}_h} \\ &= \|\nabla(u - \tilde{u}_h)\|_{\mathcal{T}_h} + \|\nabla \tilde{v}_h\|_{\mathcal{T}_h} \\ &\leqslant Ch^s |u|_{s+1,\mathcal{T}_h},\end{aligned}$$

which is already the first part of the result. In order to establish the L^2-estimate we note that, by $\Pi_0^T \tilde{v}_h = 0$, we obtain $\|\tilde{v}_h\|_T \leqslant Ch\|\nabla \tilde{v}_h\|_T$ via an inverse inequality. Hence

$$\begin{aligned}\|u - u_h^*\|_T &\leqslant \|u - \tilde{u}_h\|_T + \|\tilde{u}_h - u_h^*\|_T \\ &\leqslant \|u - \tilde{u}_h\|_T + \|\tilde{v}_h\|_T + \|\Pi_0^T(\tilde{u}_h - u_h^*)\|_T \\ &= \|u - \tilde{u}_h\|_T + \|\tilde{v}_h\|_T + \|\Pi_0^T(\tilde{u}_h - u)\|_T + \|\Pi_0^T(u - u_h)\|_T.\end{aligned}$$

Summing up over all elements, and using that

$$\|\Pi_0^T(\tilde{u}_h - u)\|_{\mathcal{T}_h} \leqslant \|\tilde{u}_h - u\|_{\mathcal{T}_h} \leqslant Ch^{s+1}|u|_{s+1,\mathcal{T}_h}$$

and Proposition 4.7, we conclude the L^2-estimate. □

REMARK 4.9 In the purely elliptic case ($\beta \equiv 0$), with f piecewise constant, we can also obtain the optimal estimate $\|u - u_h^*\| \leqslant h^{s+1}|u|_{s+1,\mathcal{T}_h}$ for the case $k = 0$ by using the estimate (4.3) instead of Proposition 4.7.

5. Implementation and numerical tests

In Section 5 we want to illustrate the theoretical results derived in the previous section by some numerical tests. As a model problem, let us consider

$$\begin{aligned} -\epsilon\Delta u + \beta\nabla u &= f \quad \text{in } \Omega := (0,1)^2, \\ u &= g \quad \text{on } \partial\Omega, \end{aligned} \tag{5.1}$$

where ϵ and β are constant on Ω. Since for the limiting hyperbolic problem our method is equivalent to the DG method, we will compare our results mainly to those obtained by the streamline diffusion method (Hughes & Brooks, 1979; Johnson & Saranen, 1986; Johnson, 1987). We refer to Houston *et al.* (2000) for a detailed comparison of hp-versions of the streamline diffusion method with DG methods for first-order hyperbolic problems.

The variational form of the streamline diffusion method is formally derived by using $v + \alpha\beta\nabla v$ as a test function in the variational formulation of (5.1). Assuming that $g = 0$ for simplicity, this yields the following.

METHOD 5.1 (Streamline diffusion) Find $u \in H_0^1(\Omega) \cap H^2(\mathcal{T}_h)$ such that

$$\begin{aligned} &\epsilon(\nabla u, \nabla v)_{\mathcal{T}_h} + (\beta\nabla u, v)_{\mathcal{T}_h} + \alpha[-\epsilon(\Delta u, \beta\nabla v)_{\mathcal{T}_h} + (\beta\nabla u, \beta\nabla v)_{\mathcal{T}_h}] \\ &\quad = (f, v)_{\mathcal{T}_h} + \alpha(f, \beta\nabla v). \end{aligned}$$

In order to obtain stability of the method, the stabilization parameter has to be chosen appropriately, depending on the shape of the elements in the mesh. Typically, the stabilization parameter is of the order of $h/|\beta|$, where h is the local mesh size. For higher-order methods the polynomial degree also influences the choice of α (cf. Houston *et al.*, 2000). For our numerical tests below we use $\alpha = 1$ for problems with dominating convection and we set $\alpha = 0$ if diffusion dominates.

5.1 *Numerical tests*

With the following examples we want to illustrate the performance of the hybrid mixed DG method under the presence of boundary layers (Example 5.2), for discontinuous solutions and internal layers (Example 5.3), and for diffusion-dominated problems (Example 5.4). Throughout we will compare our method using polynomials of order k with the streamline diffusion method using polynomials of degree $k + 1$. Thus, formally, the approximation properties of our finite-element spaces are one order less. However, as our numerical results indicate, this affects the results only in the diffusion-dominated case, where, according to our theory, we can increase the approximations by local postprocessing.

EXAMPLE 5.2 (Boundary layers) In the first example we set $g = 0$ and

$$f = \beta_1[y + (\mathrm{e}^{\beta_2 y/\epsilon} - 1)/(1 - \mathrm{e}^{1/\epsilon})] + \beta_2[1 + (\mathrm{e}^{\beta_1 x/\epsilon} - 1)/(1 - \mathrm{e}^{\beta_1/\epsilon})].$$

For $\epsilon > 0$ the exact solution to (5.1) is then given by

$$u(x, y) = [x + (\mathrm{e}^{\beta_1 x/\epsilon} - 1)/(1 - \mathrm{e}^{\beta_1/\epsilon})] \cdot [y + (\mathrm{e}^{\beta_2 y/\epsilon} - 1)/(1 - \mathrm{e}^{\beta_2/\epsilon})],$$

i.e., the solution has boundary layers at the top and right outflow boundaries.

We set $\epsilon = 0.01$ and $\beta = (2, 1)$ and then solve the problem numerically for various mesh sizes h and polynomial degrees k. Table 1 displays the errors of the numerical solutions obtained with Method 2.4 and the streamline diffusion method.

Since the exact solution is essentially bilinear away from the boundary layers, one cannot expect to gain much from further increasing the polynomial degree. As the problem gets more and more diffusion dominated with decreasing mesh size h, the error of the hybrid mixed method decays with the rate h^{k+1}, which is the order of the best approximation error. While we showed that optimal rates hold if stabilization is omitted in the diffusion-dominant case, the optimal L^2-error estimate for the stabilized Method 2.4 is not yet covered by our theory.

Since in our example the location of boundary layers is determined *a priori*, one could, of course, also use locally refined meshes (see Fig. 1).

EXAMPLE 5.3 (Discontinuities and internal layers) For the second test case we set $f = 0$ and $\beta = (2, 1)$ as before, and $\epsilon = 10^{-6}$. So we are dealing with an (almost) hyperbolic problem. Additionally, we introduce a discontinuity in the boundary conditions, i.e., we set $u(0, y) = H(y - 0.5)$ on the left inflow boundary ($H(\cdot)$ denotes the Heavyside function) and we set $u = 0$ on the remaining part of the boundary. The exact solution for $\epsilon = 0$ (the boundary conditions at the outflow boundaries have to be

TABLE 1 *L^2-errors obtained for Example 5.2 with $\epsilon = 0.01$ and $\beta = (2, 1)$ on uniformly refined meshes with mesh size h using polynomials of order k*

	Streamline diffusion						Mixed hybrid DG					
h	$k=1$	Rate	$k=2$	Rate	$k=3$	Rate	$k=0$	Rate	$k=1$	Rate	$k=2$	Rate
1.0000	0.227		0.223		0.215		0.162		0.082		0.07188	
0.5000	0.199	0.19	0.177	0.33	0.160	0.42	0.089	0.87	0.064	0.35	0.02859	1.33
0.2500	0.142	0.48	0.114	0.64	0.097	0.72	0.070	0.33	0.029	1.14	0.00874	1.71
0.1250	0.089	0.68	0.059	0.94	0.048	1.01	0.044	0.66	0.011	1.41	0.00209	2.06
0.0625	0.050	0.81	0.025	1.22	0.017	1.48	0.025	0.81	0.003	1.71	0.00034	2.63
0.0313	0.027	0.89	0.009	1.47	0.004	2.06	0.013	0.92	0.001	1.91	0.00004	2.92

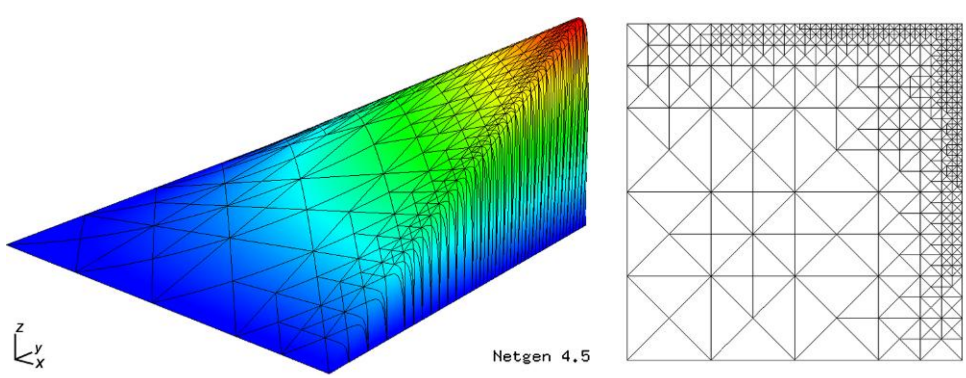

FIG. 1. Example 5.2: exact solution and locally adapted mesh with 878 elements.

TABLE 2 *L^2-errors of streamline diffusion(k) and hybrid mixed DG(k) method obtained for Example 5.3 with $\epsilon = 10^{-6}$ and $b = (2, 1)$ on uniformly refined meshes with mesh size h and polynomial degree k*

	Streamline diffusion						Mixed hybrid DG					
h	$k = 1$	Rate	$k = 2$	Rate	$k = 3$	Rate	$k = 0$	Rate	$k = 1$	Rate	$k = 2$	Rate
0.5000	0.408		0.300		0.275		0.299		0.182		0.133	
0.2500	0.328	0.31	0.243	0.30	0.227	0.28	0.222	0.43	0.139	0.39	0.098	0.44
0.1250	0.245	0.42	0.186	0.39	0.174	0.38	0.181	0.29	0.109	0.34	0.080	0.28
0.0625	0.179	0.45	0.138	0.43	0.129	0.43	0.150	0.27	0.087	0.33	0.064	0.32
0.0313	0.131	0.45	0.101	0.45	0.094	0.45	0.112	0.42	0.069	0.34	0.050	0.35

omitted in this case) is given by

$$u(x, y) = \begin{cases} 1, & y > 0.5(1 + x), \\ 0, & \text{otherwise.} \end{cases}$$

We use the solution of the purely hyperbolic problem for the calculation of the numerical errors of the finite-element solutions in Table 2. Again, we solve on uniform meshes (not aligned to the discontinuity) and compare the solutions obtained with Method 2.4 and the streamline upwind method for different polynomial degrees.

Since the exact solution has a line discontinuity at $y = 0.5(x + 1)$, one cannot expect to get better convergence rates than $h^{1/2}$. Moreover, since the solution is piecewise constant, the quality of the reconstructions can only be improved slightly by increasing the polynomial degree. Although the streamline diffusion method seems to provide better convergence rates, the actual reconstruction errors are smaller for the hybrid mixed method. In Fig. 2 we display the solutions obtained with the streamline diffusion and the hybrid mixed method. In both cases the crosswind diffusion is kept to a minimum, and so the jump of the exact solution is captured within one element layer, although the mesh is not aligned with the streamline velocity β.

Let us now turn to a diffusion-dominated problem and illustrate the increase in accuracy obtained by local postprocessing discussed in Section 4.3.

EXAMPLE 5.4 (Diffusion dominated) Consider problem (5.1) with $\beta = (2, 1)$, $\epsilon = 1$ and $f = 1$. Moreover, set $u = 0$ at the boundary. We solve problem (5.1) with Method 4.3 and compare the numerical results with those obtained by the streamline diffusion method. Since for the problem under consideration we do not have an analytical solution, we use the conforming finite-element solution with polynomial degree 8 as an approximation for the exact solution. The results of the numerical tests are summarized in Table 3.

Since in the diffusion-dominated case we omit stabilization, the streamline diffusion method coincides with the standard Galerkin method, and so we obtain optimal L^2-error estimates. The results obtained with the hybrid mixed method are also optimal with respect to the approximation properties of the finite-element space. For improving the approximation for the hybrid mixed method in that case, we can apply local postprocessing as discussed in Section 4. In Table 4 we list the results obtained after postprocessing. For comparison, we also list the L^2 best approximation errors for the corresponding finite-element spaces.

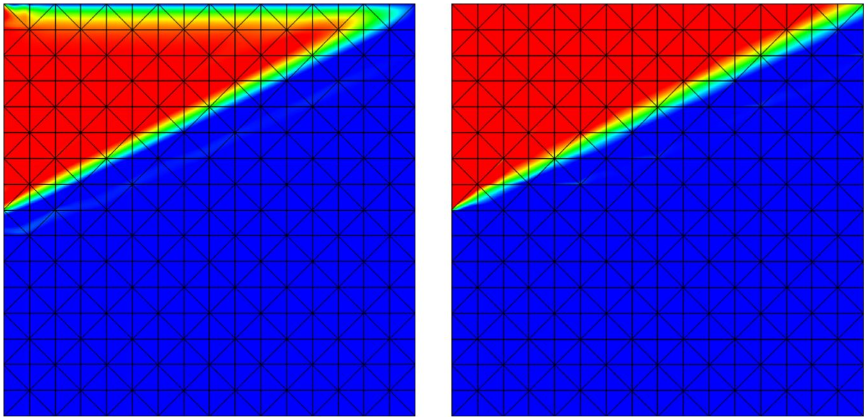

FIG. 2. Streamline diffusion(3) and hybrid mixed DG(2) solutions obtained on uniformly refined meshes with 512 elements. The streamline diffusion method develops boundary layers at the outflow boundaries. Both methods capture the discontinuity within one element layer.

TABLE 3 *L^2-errors of streamline diffusion(k) and hybrid mixed DG(k) method obtained for Example 5.3 with $\epsilon = 10^{-6}$ and $\beta = (2, 1)$ on uniformly refined meshes with mesh size h and polynomial degree k*

	Streamline diffusion				Mixed hybrid DG			
h	$k = 1$	Rate	$k = 2$	Rate	$k = 0$	Rate	$k = 1$	Rate
1.0000	0.040175		0.017043		0.022501		0.019833	
0.5000	0.009128	2.14	0.002682	2.67	0.022382	0.01	0.004392	2.18
0.2500	0.005720	0.67	0.000423	2.66	0.010841	1.05	0.001747	1.33
0.1250	0.001652	1.79	0.000061	2.81	0.005441	0.99	0.000487	1.84
0.0625	0.000428	1.95	0.000008	2.87	0.002722	1.00	0.000126	1.96

TABLE 4 *L^2-errors of postprocessed solution of the hybrid mixed DG$(k-1)$ method and the best piecewise polynomial approximation of order k on uniform meshes with mesh size h*

	Streamline diffusion				Mixed hybrid DG			
h	$k = 1$	Rate	$k = 2$	Rate	$k = 0$	Rate	$k = 1$	Rate
1.00000	0.022149		0.012169		0.018277		0.005064	
0.50000	0.012273	0.85	0.001657	2.88	0.004356	2.07	0.001108	2.19
0.25000	0.004598	1.42	0.000323	2.36	0.001741	1.32	0.000185	2.58
0.12500	0.001329	1.79	0.000048	2.74	0.000487	1.84	0.000027	2.81
0.06250	0.000347	1.94	0.000007	2.82	0.000126	1.95	0.000004	2.87

Throughout our numerical experiments the error of the postprocessed solution was always close to the best approximation error. Moreover, the hybrid mixed method with postprocessing always yielded slightly more accurate results than the standard conforming finite-element method with the corresponding polynomial degree.

5.2 *Comparison with other DG methods*

After the numerical experiments, we would like to compare the hybrid mixed method with other variants of DG methods, in particular, with the interior penalty method (Arnold, 1982) and the *multiscale DG* method presented in Buffa *et al.* (2006). The latter method is somewhat similar to the hybrid mixed method as it introduces new dofs at the skeleton and allows us to eliminate local dofs by the solution of local subproblems.

For the interior penalty Galerkin methods all dofs are present in the global system. The assembling of the element contributions requires only the dofs of one element, while the assembling of the coupling terms requires the dofs of neighbouring elements. Hence the dofs of one element are coupled to those of the neighbouring elements.

In the multiscale DG method the global dofs correspond to the trace (at the skeleton) of a continuous finite-element function. A vertex dof couples with all dofs belonging to the skeleton of all elements sharing that vertex, and dofs of one edge only couple to those belonging to the skeleton of the element sharing that edge. This carries over to three-dimensional problems, where vertex dofs couple with all dofs belonging to the skeleton of the vertex patch, and so on.

In the hybrid mixed method the global degrees belonging to one edge only couple with those of the skeleton of the neighbouring element. In three dimensions the global dofs correspond to single faces, and they couple only to those on the faces of the two neighbouring elements. The degrees of freedom for the three methods using linear polynomials for the primal variable are depicted in Figure. 3.

For a comparison of the computational effort required for the different methods we summarize the number of local and global dofs and the number of nonzero entries present in the global linear system in Table 5. For brevity, we only list the leading-order terms.

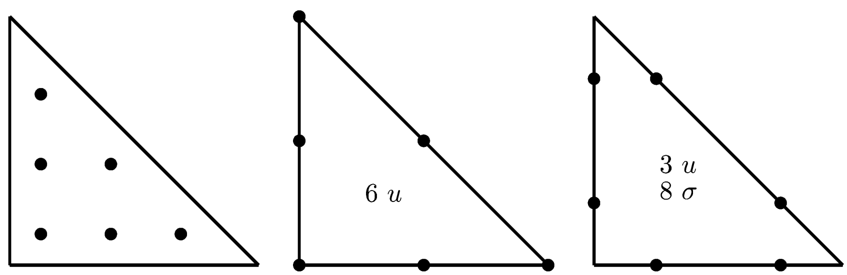

FIG. 3. Dofs for the interior penalty method and the multiscale DG method with order $k = 2$, and the hybrid mixed method with order $k = 1$. The global dofs are marked with $\bullet$, and local dofs for u and σ that can be eliminated by static condensation are depicted inside the elements.The solutions obtained by the hybrid mixed method can be improved by one order through local postprocessing (cf. Section 4).

TABLE 5 *Leading order of the number of dofs for the interior penalty method, the multiscale DG method and the hybrid mixed method of order k*

	Interior penalty	Multiscale	Hybrid mixed
Local element dofs	—	$\frac{1}{2}k^2$	$\frac{3}{2}k^2$
Global element dofs	$\frac{1}{2}k^2$	$3k$	$3k$
Global dofs	$\frac{1}{2}k^2 n_{\mathrm{el}}$	$\frac{9}{2}kn_{\mathrm{el}}$	$\frac{9}{2}kn_{\mathrm{el}}$
Nonzero entries	$k^4 n_{\mathrm{el}}$	$\frac{15}{2}k^2 n_{\mathrm{el}}$	$\frac{15}{2}k^2 n_{\mathrm{el}}$

The elimination of the internal dofs makes the assembling process of the multiscale DG and the hybrid mixed method more expensive than that of the interior penalty method. However, the coupling is decreased considerably, and therefore the local assembling can be done in parallel more easily. The global systems of the multiscale DG method and the hybrid mixed method involve less dofs and less coupling than the one for the interior penalty method.

5.3 *Concluding remarks*

In this paper we proposed a new finite-element method for convection–diffusion problems based on a mixed discretization for the elliptic part and a DG formulation for the convective part. The two methods are made compatible via hybridization, and the Lagrange multipliers play an essential role for the stabilization of the method and throughout the analysis.

Like other DG methods, but in contrast to the streamline diffusion method, the presented scheme is locally and globally conservative, which makes it a natural candidate for problems where conservation is important, for example, for time-dependent problems. Moreover, the treatment of boundary conditions is very natural and allows a seamless change from convection-dominated to purely hyperbolic regimes, where the outflow boundary conditions just disappear in the numerical scheme. In the hyperbolic limit our method corresponds to (a hybrid version of) the classical DG method and thus inherits the stabilizing features of DG methods for hyperbolic problems.

The hybrid mixed method allows a more natural treatment of elliptic operators than the DG methods. In particular, the discretization of diffusion terms does not increase the stencil of the scheme. In contrast to the streamline diffusion method and to several variants of DG methods, no tuning of a stabilization parameter is needed. In the diffusion-dominated regime the numerical solutions can be further improved by local postprocessing.

A particular advantage of our method from a computational point of view is that it is formulated and can be implemented purely element-wise. This allows static condensation of the primal and flux variables on the element level, and only the Lagrange multipliers appear in the global system. Thus the presented hybrid mixed DG method has smaller stencils as well as fewer dofs than standard DG methods, but still provides the same stability.

References

AIZINGER, V., DAWSON, C. N., COCKBURN, B. & CASTILLO, P. (2000) Local discontinuous Galerkin method for contaminant transport. *Adv. Water Resour.*, **24**, 73–87.

ARNOLD, D. N. (1982) An interior penalty finite element method with discontinuous elements. *SIAM J. Numer. Anal.*, **19**, 742–760.

ARNOLD, D. N. & BREZZI, F. (1985) Mixed and nonconforming finite element methods: implementation, postprocessing and error estimates. *Math. Model. Numer. Anal.*, **19**, 7–32.

ARNOLD, D. N., BREZZI, F., COCKBURN, B. & MARINI, D. (2002) Unified analysis of discontinuous Galerkin methods for elliptic problems. *SIAM J. Numer. Anal.*, **39**, 1749–1779.

BABUSKA, I. & ZLÁMAL, M. (1973) Nonconforming elements in the finite element method with penalty. *SIAM J. Numer. Anal.*, **10**, 863–875.

BASSI, F. & REBAY, S. (1997a) A high-order accurate discontinuous finite element method for the numerical solution of the compressible Navier–Stokes equations. *J. Comput. Phys.*, **131**, 267–279.

BASSI, F. & REBAY, S. (1997b) High-order accurate discontinuous finite element solution of the 2D Euler equations. *J. Comput. Phys.*, **138**, 251–285.

BAUMANN, C. E. & ODEN, J. T. (1999) A discontinuous *hp* finite element method for convection–diffusion problems. *Comput. Methods Appl. Mech. Eng.*, **175**, 311–341.

BRENNER, S. C. & SCOTT, L. R. (2002) *The Mathematical Theory of Finite Element Methods*. New York: Springer.

BREZZI, F. & FORTIN, M. (1991) *Mixed and Hybrid Finite Element Methods*. New York: Springer.

BRIX, K., PINTO, M. C. & DAHMEN, W. (2008) A multilevel preconditioner for the interior penalty discontinuous Galerkin method. *SIAM J. Numer. Anal.*, **46**, 2742–2768.

BUFFA, A., HUGHES, T. J. R. & SANGALLI, G. (2006) Analysis of a multiscale discontinuous Galerkin method for convection–diffusion problems. *SIAM J. Numer. Anal.*, **44**, 1420–1440.

CASTILLO, P., COCKBURN, B., SCHÖTZAU, D. & SCHWAB, C. (2002) An optimal a priori error estimate for the *hp*-version of the local discontinuous Galerkin method for convection–diffusion problems. *Math. Comput.*, **71**, 455–478.

CHEN, Z., COCKBURN, B., JEROME, J. W. & SHU, C.-W. (1995) Mixed-RKDG finite element methods for the 2-D hydrodynamic model for semiconductor device simulation. *VLSI Des.*, **3**, 145–158.

COCKBURN, B. (1988) An introduction to the discontinuous Galerkin method for convection-dominated problems. *Advanced Numerical Approximation of Nonlinear Hyperbolic Equations* (Lectures given at the 2nd Session of the Centro Internazionale Matematico Estivo (C.I.M.E.) held in Cetraro, Italy, June 23-28, 1997) (B. Cockburn, C. Johnson, C.-W. Shu, E. Tadmor eds). Berlin: Springer, pp. 151–268.

COCKBURN, B., GOPALAKRISHNAN, J. & LAZAROV, R. (2009) Unified hybridization of discontinuous Galerkin, mixed and conforming Galerkin methods for second order elliptic problems. *SIAM J. Numer. Anal.*, **47**, 1319–1365.

COCKBURN, B., KARNIADAKIS, G. E. & SHU, C.-W. (eds) (2000) *Discontinuous Galerkin Methods: Theory, Computation and Applications*. Berlin: Springer.

DAWSON, C. N. & AIZINGER, V. (1999) Upwind-mixed methods for transport equations. *Comput. Geosci.*, **3**, 93–110.

FARHOUL, M. & MOUNIM, A. S. (2005) A mixed-hybrid finite element method for convection–diffusion problems. *Appl. Math. Comput.*, **171**, 1037–1047.

FREUND, J. & STENBERG, R. (1995) On weakly imposed boundary conditions for second order problems. *Proceedings of the Ninth International Conference on Finite Elements in Fluids* (M. Morandi Cecchi, K. Morgan, J. Periaux, B. A. Screfler & O. C. Zienkiewicz eds). Venice, Italy, pp. 327–336. Available at http://math.tkk.fi/∼rstenber/Publications/Venice95.pdf

HOUSTON, P., SCHÖTZAU, D. & WIHLER, T. P. (2007) Energy norm a posteriori error estimation of *hp*-adaptive discontinuous Galerkin methods for elliptic problems. *Math. Model. Methods Appl. Sci.*, **17**, 33–62.

HOUSTON, P., SCHWAB, C. & SÜLI, E. (2000) Stabilized *hp*-finite element methods for first-order hyperbolic problems. *SIAM J. Numer. Anal.*, **37**, 1618–1643.

HOUSTON, P. & SÜLI, E. (2001) Stabilized *hp*-finite element approximation of partial differential equations with nonnegative characteristic form. *Computing*, **66**, 99–119.

HUGHES, T. J. R. & BROOKS, A. N. (1979) A multi-dimensional upwind scheme with no crosswind diffusion. *Finite Element Methods for Convection Dominated Flows* (T. Hughes ed.). Applied Mechanics Division, vol. 34. New York: American Society of Mechanical Engineers, pp. 19–35.

JOHNSON, C. (1987) *Numerical Solution of Partial Differential Equations by the Finite Element Method*. Cambridge: Cambridge University Press.

JOHNSON, C. & PITKÄRANTA, J. (1986) An analysis of the discontinuous Galerkin method for a scalar hyperbolic equation. *Math. Comput.*, **46**, 1–26.

JOHNSON, C. & SARANEN, J. (1986) Streamline diffusion methods for the incompressible Euler and Navier–Stokes equations. *Math. Comput.*, **47**, 1–18.

LESAINT, P. & RAVIART, P. A. (1974) On a finite element method for solving the neutron transport equation. *Mathematical Aspects of Finite Elements in Partial Differential Equations* (C. de Boor ed.). New York: Academic Press, pp. 89–123.

NEDELEC, J. C. (1980) Mixed finite elements in $\mathbb{R}^3$. *Numer. Math.*, **35**, 315–341.

NITSCHE, J. A. (1971) Über ein Variationsprinzip zur Lösung von Dirichlet-Problemen bei Verwendung von Telräumen, die keinen Randbedingungen unterworfen sind. *Abh. Math. Semin. Univ. Hamb.*, **36**, 9–15.

ODEN, J. T., BABUSKA, I. & BAUMANN, C. (1998) A discontinuous hp-FEM for diffusion problems. *J. Comput. Phys.*, **146**, 491–519.

PERUGIA, I. & SCHÖTZAU, D. (2002) An hp-analysis of the local discontinuous Galerkin method for diffusion problems. *J. Sci. Comput.*, **17**, 561–571.

RAVIART, P. A. & THOMAS, J. M. (1977) A mixed finite element method for second order elliptic problems. *Mathematical Aspects of the Finite Element Method* (I. Galligani & E. Magenes eds). Lecture Notes in Mathematics, vol. 606. Berlin: Springer, pp. 202–315.

REED, W. H. & HILL, T. R. (1973) Triangular mesh methods for the neutron transport equation. *Technical Report LA-UR-73-479*. Los Alamos, NM: Los Alamos Scientific Laboratory.

RICHTER, G. R. (1992) The discontinuous Galerkin method with diffusion. *Math. Comput.*, **58**, 631–643.

STENBERG, R. (1991) Postprocessing schemes for some mixed finite elements. *Math. Model. Numer. Anal.*, **25**, 151–168.

TOSELLI, A. & WIDLUND, O. (2005) *Domain Decomposition Methods—Algorithms and Theory*. Berlin: Springer.