# Guaranteed and robust discontinuous Galerkin a posteriori error estimates for convection-diffusion-reaction problems

Alexandre Ern, Annette Stephansen, Martin Vohralík

Mathematical Models and Methods in Applied Sciences


# GUARANTEED AND ROBUST DISCONTINUOUS GALERKIN A POSTERIORI ERROR ESTIMATES FOR CONVECTION–DIFFUSION–REACTION PROBLEMS*

ALEXANDRE ERN

*Université Paris-Est, CERMICS, Ecole des Ponts, 77455 Marne la Vallée cedex 2, France*
*ern@cermics.enpc.fr*

ANNETTE F. STEPHANSEN

*ANDRA, Parc de la Croix-Blanche, 1–7 rue Jean Monnet, 92298 Châtenay-Malabry, France*
*stephansen@cermics.enpc.fr*

MARTIN VOHRALÍK

*Laboratoire Jacques-Louis Lions, Université Pierre et Marie Curie (Paris 6), B.C. 187, 4 place Jussieu, 75252 Paris cedex 5, France*
*vohralik@ann.jussieu.fr*

Received (Day Month Year)
Revised (Day Month Year)
Communicated by (xxxxxxxxxx)

We propose and study a posteriori error estimates for convection–diffusion–reaction problems with inhomogeneous and anisotropic diffusion approximated by weighted interior-penalty discontinuous Galerkin methods. Our twofold objective is to derive estimates without undetermined constants and to analyze carefully the robustness of the estimates in singularly perturbed regimes due to dominant convection or reaction. We first derive locally computable estimates for the error measured in the energy (semi)norm. These estimates are evaluated using $\mathbf{H}(\mathrm{div}, \Omega)$-conforming diffusive and convective flux reconstructions, thereby extending previous work on pure diffusion problems. The resulting estimates are semi-robust in the sense that local lower error bounds can be derived using suitable cutoff functions of the local Péclet and Damköhler numbers. Fully robust estimates are obtained for the error measured in an augmented norm consisting of the energy (semi)norm, a dual norm of the skew-symmetric part of the differential operator, and a suitable contribution of the interelement jumps of the discrete solution. Numerical experiments are presented to illustrate the theoretical results.

*Keywords*: convection–diffusion–reaction, discontinuous Galerkin methods, a posteriori error estimates, dominant convection, dominant reaction, robustness

AMS Subject Classification: 65N15, 65N30, 76S05

*This work was partly supported by the Groupement MoMaS (PACEN/CNRS, ANDRA, BRGM, CEA, EdF, IRSN).

## 1. Introduction

We consider the convection–diffusion–reaction problem

$$-\nabla\cdot(\mathbf{K}\nabla u) + \boldsymbol{\beta}\cdot\nabla u + \mu u = f \quad \text{in } \Omega, \tag{1.1a}$$

$$u = 0 \quad \text{on } \partial\Omega, \tag{1.1b}$$

where $\Omega \subset \mathbb{R}^d$, $d \geq 2$, is a polyhedral domain, $\mathbf{K}$ the diffusion tensor, $\boldsymbol{\beta}$ the velocity field, $\mu$ the reaction coefficient, and $f$ the source term. We only consider homogeneous Dirichlet boundary conditions for the sake of simplicity; extensions to inhomogeneous Dirichlet and Neumann boundary conditions are possible. Our intention is to study a posteriori error estimates for the approximation of (1.1a)–(1.1b) by weighted interior-penalty discontinuous Galerkin (DG) methods with the twofold objective of deriving estimates without undetermined constants and analyzing carefully the robustness of the estimates in singularly perturbed regimes due to dominant convection or reaction. We have chosen to address the convection–diffusion–reaction problem in a general setting for the parameters $\mathbf{K}$, $\boldsymbol{\beta}$, and $\mu$ so that our results can be readily used in practical simulations. The reader interested in simplified situations can for instance take $\mathbf{K}$ equal to $\epsilon$ times the identity matrix ($\epsilon \ll 1$), $\boldsymbol{\beta}$ a divergence-free veclocity field of order unity, and $\mu$ of order unity.

For the pure diffusion problem ((1.1a)–(1.1b) with $\boldsymbol{\beta} = \mu = 0$), residual-based a posteriori energy (semi)norm error estimates for DG methods can be traced back to Ref. 6, 21; see also Ref. 11 for a unified analysis. Although the estimates derived therein are both reliable (that is, they yield an upper bound on the difference between the exact and approximate solution) and locally efficient (that is, they give local lower bounds for the error as well), they feature various undetermined constants. This shortcoming has been remedied recently in Ref. 2 upon introducing estimators based on equilibrated fluxes (for the first-order symmetric interior-penalty DG scheme in the case $d = 2$). Such estimates can be reformulated upon introducing a reconstructed $\mathbf{H}(\mathrm{div}, \Omega)$-conforming diffusive flux, say $\mathbf{t}_h$, associated with the approximate DG diffusive flux $-\mathbf{K}\nabla_h u_h$.[22,13,32,19,14] We also mention Ref. 25 where numerical experiments for similar estimators are presented. Error estimates for continuous finite element methods using reconstructed $\mathbf{H}(\mathrm{div}, \Omega)$-conforming fluxes can be traced back to the seminal work of Prager and Synge,[27] while more recent developments include Ref. 23, 24, 15.

A posteriori error estimates based on flux reconstruction for DG approximations to *convection–diffusion–reaction* problems appear to be a novel topic. Our first intermediate, yet practically important, result delivers a locally computable, global upper bound for the error measured in the energy (semi)norm $|||\cdot|||$ defined by (2.4). Letting $u$ be the exact solution of (1.1a)–(1.1b) and letting $u_h$ be its DG approximation, Theorem 3.1 states that

$$|||u - u_h||| \leq \eta,$$

where $\eta$ collects various locally computable contributions with only known constants, the leading terms for low enough local Péclet numbers having constant equal to one. These contributions are evaluated using a $H_0^1(\Omega)$-conforming reconstruction of the primal solution $u_h$ and $\mathbf{H}(\mathrm{div}, \Omega)$-conforming reconstructions of its diffusive flux $-\mathbf{K}\nabla_h u_h$ and convective flux $\boldsymbol{\beta} u_h$, thereby extending previous work on pure diffusion problems. Theorem 3.2 then states that the elementwise contributions in $\eta$ can be bounded by the local error in the energy (semi)norm augmented by the natural DG jump seminorm $|||\cdot|||_{*,\mathcal{F}_h}$ defined by (3.12) times suitable cutoff functions of the local Péclet and Damköhler numbers. In particular, this yields

$$\eta \leq C\chi(|||u - u_h||| + |||u_h|||_{*,\mathcal{F}_h}),$$

where the constant $C$ is independent of any mesh size and mildly depends on the data $\mathbf{K}$, $\boldsymbol{\beta}$, and $\mu$ as specified below, whereas $\chi$ collects the above-mentioned cutoff functions. This result is in its form similar to that derived by Verfürth for stabilized conforming finite elements in Ref. 33 and to the results in Ref. 18, 37, 38 for DG, mixed finite element, and finite volume methods, respectively. The difference with Ref. 33 is that the present $\eta$ features no undetermined constant. Moreover, $\eta$ represents a lower bound for the DG residual-based a posteriori estimate derived in Ref. 18.

To achieve full robustness in singularly perturbed regimes resulting from dominant advection or reaction, we follow the approach proposed again by Verfürth for stabilized conforming finite elements in Ref. 34 and which consists in measuring the error in an augmented norm including a suitable dual norm of the skew-symmetric part of the differential operator. Another approach to robust a posteriori error estimation has been proposed by Sangalli[28,29,30]; it consists in evaluating the convective derivative using a fractional order norm. For DG methods, the augmented norm $|||\cdot|||_{\oplus}$ defined by (3.13) differs from that considered in the conforming case and features an additional contribution which depends on the interelement jumps of the discrete solution. By proceeding this way, see Theorem 3.3, an upper bound is derived in the form

$$|||u - u_h|||_{\oplus} \leq \widetilde{\eta},$$

where $\widetilde{\eta}$ again collects various locally computable contributions (with only known constants as for $\eta$) which are evaluated using the above-mentioned reconstructions. Theorem 3.4 then states that $\widetilde{\eta}$ can be globally bounded by the error measured in the augmented norm supplemented by a suitable jump seminorm $|||\cdot|||_{\#,\mathcal{F}_h}$ defined by (3.18), that is,

$$\widetilde{\eta} \leq \tilde{C}(|||u - u_h||| + |||u_h|||_{\#,\mathcal{F}_h}),$$

where the constant $\tilde{C}$ has dependencies similar to those of $C$. By adding this jump seminorm to the error measure as well, we arrive at the final result of this paper, see Theorem 3.5, namely a fully robust equivalence result between the error and the a posteriori estimate, namely

$$|||u - u_h|||_{\oplus} + |||u - u_h|||_{\#,\mathcal{F}_h} \leq \widetilde{\eta} + |||u_h|||_{\#,\mathcal{F}_h} \leq \tilde{C}(|||u - u_h|||_{\oplus} + |||u - u_h|||_{\#,\mathcal{F}_h}).$$

This result is in its form similar to the one derived recently in Ref. 31 for DG methods using residual-based techniques instead of flux reconstruction. However, there are two important differences between the present results and those in Ref. 31. First, the latter contain undetermined constants; furthermore, the present jump seminorm features an additional cutoff function to lower its contribution in the singularly perturbed regimes.

This paper is organized as follows. We introduce the setting in Section 2, including the main notation and assumptions, the formulation of the continuous problem and its DG approximation, the reconstructed $\mathbf{H}(\mathrm{div},\Omega)$-conforming diffusive and convective fluxes for the DG solution, and the cutoff functions needed to formulate our results. We then present our main results in Section 3 while the proofs are collected in Section 4. Some numerical experiments illustrating the theoretical analysis are presented in Section 5. Finally, Appendix A briefly describes the modifications needed to handle nonmatching meshes.

## 2. The setting

### 2.1. *Main notation and assumptions*

Let $\{\mathcal{T}_h\}_{h>0}$ be a family of simplicial meshes of the domain $\Omega$. A generic element in $\mathcal{T}_h$ is denoted by $T$, $h_T$ stands for its diameter, $|T|$ for its measure, and $\mathbf{n}_T$ for its unit outward normal. The family $\{\mathcal{T}_h\}_{h>0}$ is assumed to be shape-regular in the sense that there exists a constant $\kappa_{\mathcal{T}} > 0$ such that $\min_{T\in\mathcal{T}_h} |T|/h_T^d > \kappa_{\mathcal{T}}$ for all $h > 0$. The shape-regularity is actually only necessary to prove the lower error bounds. We also suppose that the meshes cover $\overline{\Omega}$ exactly. For the sake of simplicity, we assume until Appendix A that meshes do not possess "hanging nodes". All the mesh faces are collected in the set $\mathcal{F}_h$. It is convenient to define the following sets: For all $T \in \mathcal{T}_h$,

$$\begin{aligned} \mathcal{F}_T &= \{F \in \mathcal{F}_h;\ F \subset \partial T\}, & \mathfrak{F}_T &= \{F \in \mathcal{F}_h;\ F \cap \partial T \neq \emptyset\}, \\ \mathcal{T}_T &= \{T' \in \mathcal{T}_h;\ \mathcal{F}_T \cap \mathcal{F}_{T'} \neq \emptyset\}, & \mathfrak{T}_T &= \{T' \in \mathcal{T}_h;\ T \cap T' \neq \emptyset\}, \end{aligned}$$

and for all $F \in \mathcal{F}_h$,

$$\mathcal{T}_F = \{T \in \mathcal{T}_h;\ F \in \mathcal{F}_T\}, \qquad \mathfrak{T}_F = \{T \in \mathcal{T}_h;\ F \cap \partial T \neq \emptyset\}.$$

Thus, $\mathcal{F}_T$ collects the faces of $T$, $\mathfrak{F}_T$ the faces having a non-empty intersection with $T$, $\mathcal{T}_T$ the elements sharing a face with $T$, $\mathfrak{T}_T$ the elements having a non-empty intersection with $T$, $\mathcal{T}_F$ the elements of which $F$ is a face, and $\mathfrak{T}_F$ the elements having a non-empty intersection with $F$.

We will be using the so-called broken Sobolev space

$$H^s(\mathcal{T}_h) := \{v \in L^2(\Omega); v|_T \in H^s(T) \quad \forall T \in \mathcal{T}_h\}, \tag{2.1}$$

along with its DG approximation space

$$V^k(\mathcal{T}_h) := \{v_h \in L^2(\Omega); v_h|_T \in \mathbb{P}_k(T) \quad \forall T \in \mathcal{T}_h\}, \tag{2.2}$$

where $\mathbb{P}_k(T)$, $k \geq 0$, is the set of polynomials of degree less than or equal to $k$ on an element $T$. The $L^2$-orthogonal projection onto $V^k(\mathcal{T}_h)$ is denoted by $\Pi_k$. The $L^2$-scalar product and its associated norm on a region $R \subset \Omega$ are indicated by the subscript $0, R$; shall $R$ coincide with $\Omega$, this subscript will be dropped. For $s \geq 1$, a norm (seminorm) with the subscript $s, R$ stands for the usual norm (seminorm) in $H^s(R)$. Finally, $\nabla_h$ denotes the broken gradient operator, that is, for $v \in H^1(\mathcal{T}_h)$, $\nabla_h v \in [L^2(\Omega)]^d$ and for all $T \in \mathcal{T}_h$, $(\nabla_h v)|_T = \nabla(v|_T)$.

We assume that $\mathbf{K} \in [L^\infty(\Omega)]^{d\times d}$ is a symmetric, uniformly positive definite, and piecewise constant tensor and for all $T \in \mathcal{T}_h$, we denote by $c_{\mathbf{K},T}$ and $C_{\mathbf{K},T}$, respectively, its minimum and maximum eigenvalue on $T$. We also assume that $\boldsymbol{\beta} \in [L^\infty(\Omega)]^d$ with $\nabla{\cdot}\boldsymbol{\beta} \in L^\infty(\Omega)$, $\mu \in L^\infty(\Omega)$ and $\mu - \frac{1}{2}\nabla{\cdot}\boldsymbol{\beta} \geq 0$ a.e. in $\Omega$. For all $T \in \mathcal{T}_h$, $c_{\boldsymbol{\beta},\mu,T}$ indicates the (essential) minimum value of $\mu - \frac{1}{2}\nabla{\cdot}\boldsymbol{\beta}$ on $T$; we suppose that if $c_{\boldsymbol{\beta},\mu,T} = 0$, then $\|\mu\|_{\infty,T} = \|\frac{1}{2}\nabla{\cdot}\boldsymbol{\beta}\|_{\infty,T} = 0$. We also assume $f \in L^2(\Omega)$. For all $T \in \mathcal{T}_h$, the local Péclet and Damköhler numbers can be defined as $h_T\|\boldsymbol{\beta}\|_{\infty,T}c_{\mathbf{K},T}^{-1}$ and $h_T^2 c_{\boldsymbol{\beta},\mu,T}c_{\mathbf{K},T}^{-1}$, respectively. The simplified setting discussed in the Introduction leads to $C_{\mathbf{K},T} = c_{\mathbf{K},T} = \epsilon$, $\|\boldsymbol{\beta}\|_{\infty,T} \simeq 1$, $c_{\boldsymbol{\beta},\mu,T} \simeq 1$, so that the local Péclet and Damköhler numbers reduce to $h_T\epsilon^{-1}$ and $h_T^2\epsilon^{-1}$, respectively.

### 2.2. *The continuous problem*

For all $u, v \in H^1(\mathcal{T}_h)$, we define the bilinear form

$$\mathcal{B}(u,v) := (\mathbf{K}\nabla_h u, \nabla_h v) + (\boldsymbol{\beta}{\cdot}\nabla_h u, v) + (\mu u, v), \tag{2.3}$$

and the corresponding energy (semi)norm

$$|||v|||^2 := \sum_{T\in\mathcal{T}_h} |||v|||_T^2, \qquad |||v|||_T^2 := \|\mathbf{K}^{\frac{1}{2}}\nabla v\|_{0,T}^2 + \big\|\big(\mu - \tfrac{1}{2}\nabla{\cdot}\boldsymbol{\beta}\big)^{\frac{1}{2}} v\big\|_{0,T}^2. \tag{2.4}$$

We remark that $|||\cdot|||$ is always a norm on $H_0^1(\Omega)$, whereas it is a norm on $H^1(\mathcal{T}_h)$ only if $c_{\boldsymbol{\beta},\mu,T} > 0$ for all $T \in \mathcal{T}_h$. For all $u, v \in H^1(\mathcal{T}_h)$, we also define

$$\mathcal{B}_{\mathrm{S}}(u,v) := (\mathbf{K}\nabla_h u, \nabla_h v) + \big(\big(\mu - \tfrac{1}{2}\nabla{\cdot}\boldsymbol{\beta}\big)u, v\big), \tag{2.5}$$

$$\mathcal{B}_{\mathrm{A}}(u,v) := \big(\boldsymbol{\beta}{\cdot}\nabla_h u + \tfrac{1}{2}(\nabla{\cdot}\boldsymbol{\beta})u, v\big). \tag{2.6}$$

Observe that $\mathcal{B}_{\mathrm{A}}$ is skew-symmetric on $H_0^1(\Omega)$ (but not on $H^1(\mathcal{T}_h)$), that $\mathcal{B}_{\mathrm{S}}(v,v) = |||v|||^2$ for all $v \in H^1(\mathcal{T}_h)$, and that

$$\mathcal{B} = \mathcal{B}_{\mathrm{S}} + \mathcal{B}_{\mathrm{A}}. \tag{2.7}$$

The weak formulation of (1.1a)–(1.1b) consists in finding $u \in H_0^1(\Omega)$ such that

$$\mathcal{B}(u,v) = (f,v) \qquad \forall v \in H_0^1(\Omega). \tag{2.8}$$

The above assumptions, the Green theorem, and the Cauchy–Schwarz inequality imply that $\mathcal{B}(v,v) = |||v|||^2$ for all $v \in H_0^1(\Omega)$ and that for all $u, v \in H^1(\mathcal{T}_h)$,

$$\mathcal{B}(u,v) \leq \max\left\{1, \max_{T\in\mathcal{T}_h}\left\{\frac{\|\mu\|_{\infty,T}}{c_{\boldsymbol{\beta},\mu,T}}\right\}\right\} |||u||| |||v||| + \max_{T\in\mathcal{T}_h}\left\{\frac{\|\boldsymbol{\beta}\|_{\infty,T}}{c_{\mathbf{K},T}^{1/2}}\right\} |||u||| \|v\|. \tag{2.9}$$

Hence, the problem (2.8) admits a unique solution.

**Remark 2.1 (Notation).** If $c_{\boldsymbol{\beta},\mu,T} = 0$, the term $\|\mu\|_{\infty,T}/c_{\boldsymbol{\beta},\mu,T}$ in estimate (2.9) should be evaluated as zero, since in this case we assume $\|\mu\|_{\infty,T} = 0$. To simplify the notation, we will systematically use the convention $0/0 = 0$.

### 2.3. *The discontinuous Galerkin method*

To formulate the DG method, we need to introduce jumps and (weighted) averages on mesh faces. We say that $F$ is an interior face of a given mesh if it has positive $(d-1)$-dimensional measure and if there are distinct $T^-(F)$ and $T^+(F)$ in $\mathcal{T}_h$ such that $F = \partial T^-(F) \cap \partial T^+(F)$ and we define $\mathbf{n}_F$ as the unit normal vector to $F$ pointing from $T^-(F)$ towards $T^+(F)$. Similarly, we say that $F$ is a boundary face of the mesh if it has positive $(d-1)$-dimensional measure and there is $T(F) \in \mathcal{T}_h$ such that $F = \partial T(F) \cap \partial\Omega$ and we define $\mathbf{n}_F$ as the unit outward normal to $\partial\Omega$ (the arbitrariness in the orientation of $\mathbf{n}_F$ is irrelevant in the sequel). All the interior (resp., boundary) faces of the mesh are collected into the set $\mathcal{F}_h^{\text{int}}$ (resp., $\mathcal{F}_h^{\text{ext}}$) and we define $\mathcal{F}_h := \mathcal{F}_h^{\text{int}} \cup \mathcal{F}_h^{\text{ext}}$. For a function $v$ that is double-valued on a face $F \in \mathcal{F}_h^{\text{int}}$, its jump and arithmetic average on $F$ are defined as

$$[\![v]\!]_F := v|_{T^-(F)} - v|_{T^+(F)}, \qquad \{\!\{v\}\!\}_F := \tfrac{1}{2}(v|_{T^-(F)} + v|_{T^+(F)}). \tag{2.10}$$

We set $[\![v]\!]_F := v|_F$ and $\{\!\{v\}\!\}_F := \frac{1}{2}v|_F$ on boundary faces. The subscript $F$ in the above jumps and averages is omitted if there is no ambiguity. To achieve robustness with respect to diffusion inhomogeneities, diffusivity-dependent weighted averages are considered.[20,16] For all $F \in \mathcal{F}_h^{\text{int}}$, let

$$\omega_{T^-(F),F} := \frac{\delta_{\mathbf{K},F+}}{\delta_{\mathbf{K},F+} + \delta_{\mathbf{K},F-}}, \qquad \omega_{T^+(F),F} := \frac{\delta_{\mathbf{K},F-}}{\delta_{\mathbf{K},F+} + \delta_{\mathbf{K},F-}}, \tag{2.11}$$

where $\delta_{\mathbf{K},F\mp} := \mathbf{n}_F{\cdot}\mathbf{K}|_{T^\mp(F)}\mathbf{n}_F$, and define

$$\{\!\{v\}\!\}_\omega := \omega_{T^-(F),F} v|_{T^-(F)} + \omega_{T^+(F),F} v|_{T^+(F)}. \tag{2.12}$$

On boundary faces, we set $\{\!\{v\}\!\}_\omega := v|_F$ and $\omega_{T(F),F} := 1$.

The interior-penalty DG methods considered herein are associated with the bilinear form

$$\begin{aligned} \mathcal{B}_h(u,v) := (\mathbf{K}\nabla_h u, \nabla_h v) + ((\mu - \nabla{\cdot}\boldsymbol{\beta})u, v) - (u, \boldsymbol{\beta}{\cdot}\nabla_h v) \\ - \sum_{F\in\mathcal{F}_h} \{(\mathbf{n}_F{\cdot}\{\!\{\mathbf{K}\nabla_h u\}\!\}_\omega, [\![v]\!])_{0,F} + \theta(\mathbf{n}_F{\cdot}\{\!\{\mathbf{K}\nabla_h v\}\!\}_\omega, [\![u]\!])_{0,F}\} \\ + \sum_{F\in\mathcal{F}_h} \{(\gamma_F[\![u]\!], [\![v]\!])_{0,F} + (\boldsymbol{\beta}{\cdot}\mathbf{n}_F\{\!\{u\}\!\}, [\![v]\!])_{0,F}\}. \end{aligned} \tag{2.13}$$

The discrete problem consists in finding $u_h \in V^k(\mathcal{T}_h)$ with $k \geq 1$ such that

$$\mathcal{B}_h(u_h, v_h) = (f, v_h) \qquad \forall v_h \in V^k(\mathcal{T}_h). \tag{2.14}$$

Taking in (2.13) the weights on interior faces equal to 1/2 and letting $\theta \in \{-1, 0, 1\}$ leads to the well-known Nonsymmetric, Incomplete, or Symmetric Interior-Penalty DG methods. The penalty parameter $\gamma_F$ takes the form

$$\gamma_F := \alpha_F \gamma_{\mathbf{K},F} h_F^{-1} + \gamma_{\boldsymbol{\beta},F} \qquad \forall F \in \mathcal{F}_h, \tag{2.15}$$

where $\alpha_F$ is a (user-dependent) positive parameter,

$$\gamma_{\mathbf{K},F} := \frac{\delta_{\mathbf{K},F+}\delta_{\mathbf{K},F-}}{\delta_{\mathbf{K},F+} + \delta_{\mathbf{K},F-}}, \tag{2.16}$$

$h_F$ the diameter of $F$, and $\gamma_{\boldsymbol{\beta},F}$ a nonnegative scalar-valued function depending on $\boldsymbol{\beta}$ and vanishing if $\boldsymbol{\beta} = 0$; we suppose here that $\gamma_{\boldsymbol{\beta},F} = \frac{1}{2}|\boldsymbol{\beta}{\cdot}\mathbf{n}_F|$, which amounts to so-called upwinding. As usual with interior-penalty methods, the parameters $\alpha_F$ must be taken large enough to ensure the coercivity of the discrete bilinear form $\mathcal{B}_h$ on $V^k(\mathcal{T}_h)$ whenever $\theta \neq -1$.

### 2.4. *Diffusive and convective flux reconstruction*

The approximate DG diffusive flux $-\mathbf{K}\nabla_h u_h$ and convective flux $\boldsymbol{\beta} u_h$ are nonconforming since they do not belong to the space $\mathbf{H}(\mathrm{div}, \Omega)$ as their exact counterparts do. For pure diffusion problems, $\mathbf{H}(\mathrm{div}, \Omega)$-conforming reconstructions of the approximate DG diffusive flux have been investigated in Ref. 4, 17, 22. We generalize here the approach of Ref. 17, 22 to convection–diffusion–reaction problems.

The reconstructed diffusive and convective fluxes will belong to the Raviart–Thomas–Nédélec spaces of vector functions on the mesh $\mathcal{T}_h$,

$$\mathbf{RTN}^l(\mathcal{T}_h) = \left\{\mathbf{v}_h \in \mathbf{H}(\mathrm{div}, \Omega)\,;\mathbf{v}_h|_T \in \mathbf{RTN}^l_T \quad \forall T \in \mathcal{T}_h\right\},$$

where $l \in \{k-1, k\}$ (recall that $k$ is the polynomial degree used for the DG approximation) and $\mathbf{RTN}^l_T = \mathbb{P}^d_l(T) + \mathbf{x}\mathbb{P}_l(T)$. In particular, $\mathbf{v}_h \in \mathbf{RTN}^l(\mathcal{T}_h)$ is such that $\nabla{\cdot}\mathbf{v}_h \in \mathbb{P}_l(T)$ for all $T \in \mathcal{T}_h$, $\mathbf{v}_h{\cdot}\mathbf{n}_F \in \mathbb{P}_l(F)$ for all $F \in \mathcal{F}_T$ and all $T \in \mathcal{T}_h$, and such that its normal trace is continuous, *cf.* Ref. 8. Using the specification of the degrees of freedom of functions in $\mathbf{RTN}^l_T$, our $\mathbf{H}(\mathrm{div}, \Omega)$-conforming flux reconstructions $\mathbf{t}_h \in \mathbf{RTN}^l(\mathcal{T}_h)$ and $\mathbf{q}_h \in \mathbf{RTN}^l(\mathcal{T}_h)$ are prescribed locally on all $T \in \mathcal{T}_h$ as follows: For all $F \in \mathcal{F}_T$ and all $q_h \in \mathbb{P}_l(F)$,

$$(\mathbf{t}_h{\cdot}\mathbf{n}_F, q_h)_{0,F} = \left(-\mathbf{n}_F{\cdot}\{\!\{\mathbf{K}\nabla_h u_h\}\!\}_\omega + \alpha_F \gamma_{\mathbf{K},F} h_F^{-1} [\![u_h]\!], q_h\right)_{0,F}, \tag{2.17}$$

$$(\mathbf{q}_h{\cdot}\mathbf{n}_F, q_h)_{0,F} = (\boldsymbol{\beta}{\cdot}\mathbf{n}_F \{\!\{u_h\}\!\} + \gamma_{\boldsymbol{\beta},F}[\![u_h]\!], q_h)_{0,F}, \tag{2.18}$$

and for all $\mathbf{r}_h \in \mathbb{P}^d_{l-1}(T)$,

$$(\mathbf{t}_h, \mathbf{r}_h)_{0,T} = -(\mathbf{K}\nabla u_h, \mathbf{r}_h)_{0,T} + \theta \sum_{F \in \mathcal{F}_T} \omega_{T,F} (\mathbf{n}_F{\cdot}\mathbf{K}\mathbf{r}_h, [\![u_h]\!])_{0,F}, \tag{2.19}$$

$$(\mathbf{q}_h, \mathbf{r}_h)_{0,T} = (u_h, \boldsymbol{\beta}{\cdot}\mathbf{r}_h)_{0,T}. \tag{2.20}$$

Observe that the quantities prescribing the moments of $\mathbf{t}_h{\cdot}\mathbf{n}_F$ and $\mathbf{q}_h{\cdot}\mathbf{n}_F$ are univocally defined for each face $F \in \mathcal{F}_h$, whence the continuity of the normal traces of $\mathbf{t}_h$ and $\mathbf{q}_h$. The above construction is motivated by the following important result:

**Lemma 2.1 (Local conservativity).** *There holds*

$$(\nabla\cdot\mathbf{t}_h + \nabla\cdot\mathbf{q}_h + \Pi_l((\mu - \nabla\cdot\boldsymbol{\beta})u_h))|_T = \Pi_l f|_T \qquad \forall T \in \mathcal{T}_h. \tag{2.21}$$

**Proof.** Let $T \in \mathcal{T}_h$ and let $\xi_h \in \mathbb{P}_l(T)$. Owing to the Green theorem,

$$(\nabla\cdot\mathbf{t}_h + \nabla\cdot\mathbf{q}_h, \xi_h)_{0,T} = -(\mathbf{t}_h + \mathbf{q}_h, \nabla\xi_h)_{0,T} + \sum_{F\in\mathcal{F}_T} ((\mathbf{t}_h + \mathbf{q}_h)\cdot\mathbf{n}_T, \xi_h)_{0,F}.$$

Using (2.17)–(2.20) along with the definition (2.13) of the bilinear form $B_h$ leads to

$$(\nabla\cdot\mathbf{t}_h + \nabla\cdot\mathbf{q}_h, \xi_h)_{0,T} = B_h(u_h, \xi_h 1_T) - ((\mu - \nabla\cdot\boldsymbol{\beta})u_h, \xi_h)_{0,T}. \tag{2.22}$$

Since $u_h$ solves (2.14), this yields (2.21). □

## 2.5. *Cutoff functions*

The following local approximation results for $L^2$-projections hold: For all $\varphi \in H_0^1(\Omega)$,

$$\|\varphi - \Pi_0\varphi\|_{0,T} \le m_T |||\varphi|||_T \qquad \forall T \in \mathcal{T}_h, \tag{2.23}$$

$$\|\varphi - \Pi_0\varphi|_T\|_{0,F} \le C_{\mathrm{t},T,F}^{1/2} \widetilde{m}_T^{1/2} |||\varphi|||_T \qquad \forall T \in \mathcal{T}_h,\ \forall F \in \mathcal{F}_T, \tag{2.24}$$

$$\|[\![\Pi_0\varphi]\!]\|_{0,F} \le m_F \sum_{T\in\mathcal{T}_F} |||\varphi|||_T \qquad \forall F \in \mathcal{F}_h, \tag{2.25}$$

with the cutoff functions

$$m_T^2 := \min\{C_{\mathrm{P}} h_T^2 c_{\mathbf{K},T}^{-1}, c_{\boldsymbol{\beta},\mu,T}^{-1}\}, \tag{2.26}$$

$$\widetilde{m}_T := \min\{(C_{\mathrm{P}} + C_{\mathrm{P}}^{1/2}) h_T c_{\mathbf{K},T}^{-1}, h_T^{-1} c_{\boldsymbol{\beta},\mu,T}^{-1} + c_{\boldsymbol{\beta},\mu,T}^{-1/2} c_{\mathbf{K},T}^{-1/2}/2\}, \tag{2.27}$$

$$m_F^2 := \min\left\{\max_{T\in\mathcal{T}_F}\left\{C_{\mathrm{F},T,F}\frac{|F|h_T^2}{|T|c_{\mathbf{K},T}}\right\}, \max_{T\in\mathcal{T}_F}\left\{\frac{|F|}{|T|c_{\boldsymbol{\beta},\mu,T}}\right\}\right\}, \tag{2.28}$$

where $|F|$ denotes the measure of $F$. Here, $C_{\mathrm{P}}$ is the constant from the Poincaré inequality

$$\|\varphi - \Pi_0\varphi\|_{0,T}^2 \le C_{\mathrm{P}} h_T^2 \|\nabla\varphi\|_{0,T}^2 \qquad \forall\varphi \in H^1(T), \tag{2.29}$$

which can be evaluated as $C_{\mathrm{P}} = 1/\pi^2$ owing to the convexity of simplices.[26,5] In addition, $C_{\mathrm{t},T,F}$ and $C_{\mathrm{F},T,F}$ are respectively the constants from the following trace and generalized Friedrichs inequalities:

$$\|\varphi\|_{0,F}^2 \le C_{\mathrm{t},T,F}(h_T^{-1}\|\varphi\|_T^2 + \|\varphi\|_T\|\nabla\varphi\|_T), \tag{2.30}$$

$$\|\varphi - \Pi_{0,F}\varphi\|_{0,T}^2 \le C_{\mathrm{F},T,F} h_T^2 \|\nabla\varphi\|_{0,T}^2, \tag{2.31}$$

valid for all $T \in \mathcal{T}_h$, $\varphi \in H^1(T)$, and $F \in \mathcal{F}_T$; here for $l \ge 0$, $\Pi_{l,F}$ denotes the $L^2$-orthogonal projection onto $\mathbb{P}_l(F)$. It follows from Lemma 3.12 in Ref. 32 that $C_{\mathrm{t},T,F} = |F|h_T/|T|$ for a simplex $T$ and its face $F$; see also Ref. 10. Furthermore, it follows from Lemma 4.1 in Ref. 36 that $C_{\mathrm{F},T,F} = 3d$ for a simplex $T$ and its face $F$. The estimate (2.23) is readily inferred from the Poincaré inequality (2.29) and the fact that $\|\varphi - \Pi_0\varphi\|_{0,T} \le \|\varphi\|_{0,T}$. The estimate (2.24) is established in Ref. 12. Finally, the estimate (2.25) is proved in Lemma 4.5 of Ref. 38.

## 3. Main results

This section exposes the main results of this work; their proofs are collected in the next section. For the sake of clarity, this section is split into three subparts. The first one contains intermediate, yet practically important, results, namely global upper bounds and local, semi-robust, lower bounds for the error estimated in the energy norm. The second one contains global upper bounds and global, fully robust, lower bounds for the error estimated in an augmented norm. All the upper bounds below are valid for *arbitrary* $H^1_0(\Omega)$-conforming reconstructions of the primal unknown. The lower bounds instead are proven for a specific choice of this reconstruction. The third subpart contains the final, fully robust result.

### 3.1. *Energy norm estimates*

This section is devoted to energy norm error estimates.

3.1.1. *Locally computable estimate*

Let $s_h \in H^1_0(\Omega)$ and let $\mathbf{t}_h, \mathbf{q}_h \in \mathbf{H}(\mathrm{div}, \Omega)$ be defined by (2.17)–(2.20). Let $T \in \mathcal{T}_h$. The *nonconformity estimator* $\eta_{\mathrm{NC},T}$, the *residual estimator* $\eta_{\mathrm{R},T}$, and the *diffusive flux estimator* $\eta_{\mathrm{DF},T}$ are defined as

$$\eta_{\mathrm{NC},T} := |||u_h - s_h|||_T, \tag{3.1}$$

$$\eta_{\mathrm{R},T} := m_T \|f - \nabla\cdot\mathbf{t}_h - \nabla\cdot\mathbf{q}_h - (\mu - \nabla\cdot\boldsymbol{\beta})u_h\|_{0,T}, \tag{3.2}$$

$$\eta_{\mathrm{DF},T} := \min\left\{\eta^{(1)}_{\mathrm{DF},T}, \eta^{(2)}_{\mathrm{DF},T}\right\}, \tag{3.3}$$

where

$$\eta^{(1)}_{\mathrm{DF},T} := \|\mathbf{K}^{\frac{1}{2}}\nabla u_h + \mathbf{K}^{-\frac{1}{2}}\mathbf{t}_h\|_{0,T}, \tag{3.4}$$

$$\begin{aligned}\eta^{(2)}_{\mathrm{DF},T} := {} & m_T \|(Id - \Pi_0)(\nabla\cdot(\mathbf{K}\nabla u_h + \mathbf{t}_h))\|_{0,T} \\ & + \widetilde{m}_T^{1/2} \sum_{F\in\mathcal{F}_T} C^{1/2}_{\mathrm{t},T,F} \|(\mathbf{K}\nabla u_h + \mathbf{t}_h)\cdot\mathbf{n}_F\|_{0,F}.\end{aligned} \tag{3.5}$$

Furthermore, we define the two *convection estimators* $\eta_{\mathrm{C},1,T}$ and $\eta_{\mathrm{C},2,T}$ and the *upwinding estimator* $\eta_{\mathrm{U},T}$ as

$$\eta_{\mathrm{C},1,T} := m_T \|(Id - \Pi_0)(\nabla\cdot(\mathbf{q}_h - \boldsymbol{\beta}s_h))\|_{0,T}, \tag{3.6}$$

$$\eta_{\mathrm{C},2,T} := c^{-1/2}_{\boldsymbol{\beta},\mu,T} \left\|\tfrac{1}{2}(\nabla\cdot\boldsymbol{\beta})(u_h - s_h)\right\|_{0,T}, \tag{3.7}$$

$$\eta_{\mathrm{U},T} := \sum_{F\in\mathcal{F}_T} m_F \|\Pi_{0,F}((\mathbf{q}_h - \boldsymbol{\beta}s_h)\cdot\mathbf{n}_F)\|_{0,F}. \tag{3.8}$$

Recall that the constants $m_T$, $\widetilde{m}_T$, and $m_F$ are defined by (2.26)–(2.28). We can now state the main result of this section.

**Theorem 3.1 (Energy norm estimate).** *Let $u$ be the solution of* (2.8) *and let $u_h$ be its DG approximation solving* (2.14)*. Then,*

$$|||u - u_h||| \le \eta,$$

*where*

$$\eta := \left\{ \sum_{T\in\mathcal{T}_h} \eta_{\mathrm{NC},T}^2 \right\}^{1/2} + \left\{ \sum_{T\in\mathcal{T}_h} (\eta_{\mathrm{R},T} + \eta_{\mathrm{DF},T} + \eta_{\mathrm{C},1,T} + \eta_{\mathrm{C},2,T} + \eta_{\mathrm{U},T})^2 \right\}^{1/2}.$$

**Remark 3.1 (Properties of the estimate of Theorem 3.1).** The estimate of Theorem 3.1 yields a guaranteed upper bound, the estimate is valid uniformly with respect to the polynomial degree $k$, no polynomial data form is needed for $f$, and, finally, the estimate is valid more generally for any $\mathbf{t}_h, \mathbf{q}_h \in \mathbf{H}(\mathrm{div}, \Omega)$ such that $(\nabla\cdot\mathbf{t}_h + \nabla\cdot\mathbf{q}_h + (\mu - \nabla\cdot\boldsymbol{\beta})u_h, 1)_{0,T} = (f, 1)_{0,T}$ for all $T \in \mathcal{T}_h$; this is a local (conservation) property, in contrast to the global Galerkin orthogonality used traditionally for conforming finite element methods.

**Remark 3.2 (Form of $\eta_{\mathrm{DF},T}$).** The idea of defining the diffusive flux estimator $\eta_{\mathrm{DF},T}$ as a minimum between two quantities has been proposed in Ref. 12. The purpose is to obtain in singularly perturbed regimes resulting from dominant convection or reaction appropriate cutoff functions in the expression for $\eta^{(2)}_{\mathrm{DF},T}$. This way of proceeding is coherent with the recent observation made by Verfürth that the diffusive flux estimator $\eta^{(1)}_{\mathrm{DF},T}$ alone cannot be shown to be robust.[35]

**Remark 3.3 (Superconvergence of $\eta_{\mathrm{R},T}$).** For pure diffusion problems, Lemma 2.1 implies $\eta_{\mathrm{R},T} = m_T\|f - \Pi_l f\|_{0,T}$ and hence, $\eta_{\mathrm{R},T}$ takes the form of a data oscillation term that superconverges by one ($l = k - 1$) or two ($l = k$) orders in mesh size if $f$ is piecewise smooth. In the general case, taking $l = k$ and $\mu$ and $\nabla\cdot\boldsymbol{\beta}$ piecewise constant, Lemma 2.1 still implies the superconvergent form $\eta_{\mathrm{R},T} = m_T\|f - \Pi_k f\|_{0,T}$. In practice, $\eta_{\mathrm{R},T}$ should not be neglected since it can be significant on coarse grids or for singularly perturbed regimes.

### 3.1.2. *Local efficiency*

To state the local efficiency of the estimate derived in Theorem 3.1, we choose a specific reconstruction $s_h \in H_0^1(\Omega)$ of $u_h$ and introduce some additional notation. Firstly, we consider the so-called Oswald interpolation operator $\mathcal{I}_{\mathrm{Os}} : V^k(\mathcal{T}_h) \to V^k(\mathcal{T}_h)\cap H_0^1(\Omega)$ defined as follows: For a function $v_h \in V^k(\mathcal{T}_h)$, $\mathcal{I}_{\mathrm{Os}}(v_h)$ is prescribed through its values at suitable (Lagrange) nodes of the simplices of $\mathcal{T}_h$. At the nodes located inside $\Omega$, the average of the values of $v_h$ at this node is used,

$$\mathcal{I}_{\mathrm{Os}}(v_h)(V) = \frac{1}{\#(\mathcal{T}_V)} \sum_{T\in\mathcal{T}_V} v_h|_T(V),$$

where $\mathcal{T}_V$ is the set of those $T \in \mathcal{T}_h$ to which the node $V$ belongs and where for any set $S$, $\#(S)$ denotes its cardinality. Note that $\mathcal{I}_{\mathrm{Os}}(v_h)(V) = v_h(V)$ at those nodes $V$ lying in the interior of some $T \in \mathcal{T}_h$. At boundary nodes, the value of $\mathcal{I}_{\mathrm{Os}}(v_h)$ is set to zero. Furthermore, we consider the following residual-based a posteriori error estimators[18]: For all $T \in \mathcal{T}_h$,

$$\rho_{1,T} := m_T \|f + \nabla\cdot(\mathbf{K}\nabla_h u_h) - \boldsymbol{\beta}\cdot\nabla_h u_h - \mu u_h\|_{0,T}, \tag{3.9}$$

$$\rho_{2,T} := m_T^{1/2} c_{\mathbf{K},T}^{-1/4} \sum_{F\in\mathcal{F}_T} \bar{\omega}_{T,F} \|\mathbf{n}_F\cdot[\![\mathbf{K}\nabla u_h]\!]\|_{0,F}, \tag{3.10}$$

where $\bar{\omega}_{T,F} = (1-\omega_{T,F})$. We can now state the main result of this section.

**Theorem 3.2 (Local efficiency of the energy norm estimate).** *Let $u$ be the solution of (2.8) and let $u_h$ be its DG approximation solving (2.14). Assume for simplicity that $\nabla\cdot(\mathbf{q}_h - \boldsymbol{\beta} s_h) \in \mathbb{P}_l(T)$ and $\nabla\cdot(\mathbf{q}_h - \boldsymbol{\beta} u_h) \in \mathbb{P}_l(T)$ for all $T \in \mathcal{T}_h$ and that $\gamma_{\boldsymbol{\beta},F}$ is facewise constant. For all $T \in \mathcal{T}_h$, let $c_{\mathbf{K},\mathfrak{T}_T} := \min_{T'\in\mathfrak{T}_T} c_{\mathbf{K},T'}$, $c_{\mathbf{K},\mathcal{T}_T} := \min_{T'\in\mathcal{T}_T} c_{\mathbf{K},T'}$, $c_{\boldsymbol{\beta},\mathfrak{F}_T} := \min_{F\in\mathfrak{F}_T}\gamma_{\boldsymbol{\beta},F}$ and $c_{\boldsymbol{\beta},\mathcal{F}_T} := \min_{F\in\mathcal{F}_T}\gamma_{\boldsymbol{\beta},F}$, and introduce the cutoff functions*

$$\chi_{\mathfrak{T}_T} := \min(h_T c_{\mathbf{K},\mathfrak{T}_T}^{-1/2}, h_T^{1/2} c_{\boldsymbol{\beta},\mathfrak{F}_T}^{-1/2}), \qquad \chi_{\mathcal{T}_T} := \min(h_T c_{\mathbf{K},\mathcal{T}_T}^{-1/2}, h_T^{1/2} c_{\boldsymbol{\beta},\mathcal{F}_T}^{-1/2}), \tag{3.11}$$

*as well as $m_{\mathcal{T}_T} := \min(h_T c_{\mathbf{K},\mathcal{T}_T}^{-1/2}, c_{\boldsymbol{\beta},\mu,\mathcal{T}_T}^{-1/2})$ where $c_{\boldsymbol{\beta},\mu,\mathcal{T}_T} := \min_{T'\in\mathcal{T}_T} c_{\boldsymbol{\beta},\mu,T'}$. For any subset $\mathcal{F}$ of $\mathcal{F}_h$, define the jump seminorm*

$$|||v|||_{*,\mathcal{F}}^2 := \sum_{F\in\mathcal{F}} \|\gamma_F^{1/2}[\![v]\!]\|_{0,F}^2 \qquad v \in H^1(\mathcal{T}_h). \tag{3.12}$$

*Let $\eta_{\mathrm{NC},T}$, $\eta_{\mathrm{R},T}$, $\eta_{\mathrm{DF},T}$, $\eta_{\mathrm{C},1,T}$, $\eta_{\mathrm{C},2,T}$, and $\eta_{\mathrm{U},T}$ be defined by (3.1)–(3.8) with $s_h = \mathcal{I}_{\mathrm{Os}}(u_h)$ and let $\mathbf{t}_h, \mathbf{q}_h \in \mathbf{H}(\mathrm{div},\Omega)$ be defined by (2.17)–(2.20). Then,*

$$\begin{aligned}
\eta_{\mathrm{NC},T} &\le C\left(\frac{C_{\mathbf{K},T}^{1/2}}{c_{\mathbf{K},\mathfrak{T}_T}^{1/2}} + \left\|\mu - \tfrac{1}{2}\nabla\cdot\boldsymbol{\beta}\right\|_{\infty,T}^{1/2}\chi_{\mathfrak{T}_T}\right)|||u-u_h|||_{*,\mathfrak{F}_T},\\
\eta_{\mathrm{C},2,T} &\le C\left\|\tfrac{1}{2}\nabla\cdot\boldsymbol{\beta}\right\|_{\infty,T} c_{\boldsymbol{\beta},\mu,T}^{-1/2}\chi_{\mathfrak{T}_T}|||u-u_h|||_{*,\mathfrak{F}_T},\\
\eta_{\mathrm{U},T} &\le C m_{\mathcal{T}_T} h_T^{-1}\|\boldsymbol{\beta}\|_{\infty,T}\chi_{\mathfrak{T}_T}|||u-u_h|||_{*,\mathfrak{F}_T},\\
\eta_{\mathrm{C},1,T} &\le C m_T h_T^{-1}\|\boldsymbol{\beta}\|_{\infty,T}\chi_{\mathfrak{T}_T}|||u-u_h|||_{*,\mathfrak{F}_T},\\
\eta_{\mathrm{R},T} &\le \rho_{1,T} + C\varsigma_T\rho_{2,T} + C\left(\varsigma_T^2\frac{C_{\mathbf{K},T}^{1/2}}{c_{\mathbf{K},T}^{1/2}} + m_T h_T^{-1}\|\boldsymbol{\beta}\|_{\infty,T}\chi_{\mathcal{T}_T}\right)|||u-u_h|||_{*,\mathcal{F}_T},\\
\eta_{\mathrm{DF},T} &\le C\rho_{2,T} + C\varsigma_T\frac{C_{\mathbf{K},T}^{1/2}}{c_{\mathbf{K},T}^{1/2}}|||u-u_h|||_{*,\mathcal{F}_T},
\end{aligned}$$

*where $\varsigma_T := m_T^{1/2} h_T^{-1/2} c_{\mathbf{K},T}^{1/4} \le C_{\mathrm{P}}^{1/4}$ by construction. The constant $C$ only depends on the space dimension $d$, the polynomial degree $k$ of $u_h$, the shape-regularity parameter $\kappa_{\mathcal{T}}$, and the DG parameters $\alpha_F$ and $\theta$.*

**Remark 3.4 (Estimates on $\rho_{1,T}$ and $\rho_{2,T}$).** The following semi-robust bounds are proved in Propositions 3.3 and 3.4 of Ref. 18 under the assumption that $f$, $\boldsymbol{\beta}$, and $\mu$ are piecewise polynomials of degree $m$:

$$\rho_{1,T} \le C m_T (C_{\mathbf{K},T}^{1/2} h_T^{-1} + \min(\alpha_{1,T}, \alpha_{2,T})) |||u - u_h|||_T,$$

$$\rho_{2,T} \le C \frac{C_{\mathbf{K},T}^{1/2}}{c_{\mathbf{K},T}^{1/2}} m_T^{1/2} c_{\mathbf{K},T}^{1/4} \sum_{T' \in \mathcal{T}_T} m_{T'}^{-1/2} c_{\mathbf{K},T'}^{-1/4} \left( \frac{C_{\mathbf{K},T'}^{1/2}}{c_{\mathbf{K},T'}^{1/2}} + m_{T'} \alpha_{1,T'} \right) |||u - u_h|||_{T'},$$

with $\alpha_{1,T} := \|\mu\|_{\infty,T} c_{\boldsymbol{\beta},\mu,T}^{-1/2} + \|\boldsymbol{\beta}\|_{\infty,T} c_{\mathbf{K},T}^{-1/2}$ and $\alpha_{2,T} := c_{\boldsymbol{\beta},\mu,T}^{-1/2}(\|\mu - \nabla\cdot\boldsymbol{\beta}\|_{\infty,T} + \|\boldsymbol{\beta}\|_{\infty,T} h_T^{-1})$. The constant $C$ only depends on $d$, $k$, $m$, and $\kappa_{\mathcal{T}}$.

**Remark 3.5 (Comments on the results of Theorem 3.2).** In the DG energy norm, the a posteriori error estimate of Theorem 3.1 is semi-robust in the sense that the bounds on the estimators involve cutoff functions of the local Péclet and Damköhler numbers in various forms. This result is of the same quality as those achieved in Ref. 33, 37, 38, 18. Moreover, as $h \to 0$, the estimators $\eta_{\mathrm{C},1,T}$, $\eta_{\mathrm{C},2,T}$, and $\eta_{\mathrm{U},T}$ will loose influence, whereas $\eta_{\mathrm{NC},T}$ and $\eta_{\mathrm{DF},T}$ will become optimally efficient. Numerical experiments suggest that $\eta_{\mathrm{DF},T}^{(1)}$ is often well-behaved.

**Remark 3.6 (Pure diffusion).** Theorems 3.1 and 3.2 obviously apply to the pure diffusion case and deliver similar results to Ref. 2, 22, 13, 32, 19, 14. One salient feature of the present estimate is that owing to the bounds in Remark 3.4, the diffusion estimator $\eta_{\mathrm{DF},T}$ is fully robust with respect to diffusion inhomogeneities.

### 3.2. *Augmented norm estimates*

The so-called augmented norm that we will be using for error control is defined as

$$|||v|||_{\oplus} := |||v||| + \sup_{\varphi \in H_0^1(\Omega),\, |||\varphi|||=1} \{\mathcal{B}_{\mathrm{A}}(v,\varphi) + \mathcal{B}_{\mathrm{D}}(v,\varphi)\} \qquad v \in H^1(\mathcal{T}_h), \tag{3.13}$$

with $\mathcal{B}_{\mathrm{A}}$ defined by (2.6) and where for all $u, v \in H^1(\mathcal{T}_h)$,

$$\mathcal{B}_{\mathrm{D}}(u,v) := -\sum_{F \in \mathcal{F}_h} (\boldsymbol{\beta}\cdot\mathbf{n}_F [\![u]\!], \{\!\{\Pi_0 v\}\!\})_{0,F}. \tag{3.14}$$

Whenever $\|\nabla\cdot\boldsymbol{\beta}\|_{\infty,T}$ is controlled by $c_{\boldsymbol{\beta},\mu,T}$ for all $T \in \mathcal{T}_h$, the zero-order contribution in $\mathcal{B}_{\mathrm{A}}$ can be discarded in the definition of the augmented norm, recovering the dual norm introduced by Verfürth for conforming finite elements.[34] The additional contribution from $\mathcal{B}_{\mathrm{D}}$ in the augmented norm is specific to the DG setting and has been introduced in the present work to sharpen the global efficiency result; see Remark 4.2 below.

#### 3.2.1. *Locally computable estimate*

Let $s_h \in H_0^1(\Omega)$ and let $\mathbf{t}_h, \mathbf{q}_h \in \mathbf{H}(\mathrm{div},\Omega)$ be defined by (2.17)–(2.20). Let $T \in \mathcal{T}_h$. Let $\eta$, $\eta_{\mathrm{R},T}$, and $\eta_{\mathrm{DF},T}$ be as in Section 3.1. We define the *modified convection* *estimator* $\widetilde{\eta}_{\mathrm{C},1,T}$ and the *modified upwinding estimator* $\widetilde{\eta}_{\mathrm{U},T}$ as

$$\widetilde{\eta}_{\mathrm{C},1,T} := m_T \|(Id - \Pi_0)(\nabla\cdot(\mathbf{q}_h - \boldsymbol{\beta} u_h))\|_{0,T}, \tag{3.15}$$

$$\widetilde{\eta}_{\mathrm{U},T} := \sum_{F\in\mathcal{F}_T} m_F \|\Pi_{0,F}(\gamma_{\boldsymbol{\beta},F}[\![u_h]\!])\|_{0,F}. \tag{3.16}$$

**Theorem 3.3 (Augmented norm estimate).** *Let $u$ be the solution of* (2.8) *and let $u_h$ be its DG approximation solving* (2.14)*. Then,*

$$|\!|\!| u - u_h |\!|\!|_{\oplus} \le \widetilde{\eta} := 2\eta + \left\{ \sum_{T\in\mathcal{T}_h} (\eta_{\mathrm{R},T} + \eta_{\mathrm{DF},T} + \widetilde{\eta}_{\mathrm{C},1,T} + \widetilde{\eta}_{\mathrm{U},T})^2 \right\}^{1/2}. \tag{3.17}$$

**Remark 3.7 (Comparison of $\eta$ and $\widetilde{\eta}$).** We observe that the estimator $\widetilde{\eta}$ is fully computable and that it has the same structure as the estimator $\eta$ derived in Theorem 3.1, so that for practical purposes, $\eta$ can often be sufficient.

3.2.2. *Global efficiency*

We show here that the $|\!|\!|\cdot|\!|\!|_{\oplus}$-norm a posteriori error estimate of Theorem 3.3 is globally efficient and fully robust.

**Theorem 3.4 (Global efficiency of the augmented norm estimate).** *Along with the assumptions of Theorem 3.2, assume that $f$, $\boldsymbol{\beta}$, and $\mu$ are piecewise polynomials of degree $m$. For all $v \in H^1(\mathcal{T}_h)$, define*

$$\begin{aligned} |\!|\!| v |\!|\!|^2_{\#,\mathcal{F}_h} := \sum_{T\in\mathcal{T}_h} \sum_{F\in\mathfrak{F}_T} \frac{1}{\#(\mathfrak{T}_F)} \Bigg\{ & \frac{C_{\mathbf{K},T}}{c_{\mathbf{K},\mathfrak{T}_T}} \alpha_F \gamma_{\mathbf{K},F} h_F^{-1} \|[\![v]\!]\|^2_{0,F} + c_{\boldsymbol{\beta},\mu,T} h_F \|[\![v]\!]\|^2_{0,F} \\ & + m^2_{\mathcal{T}_T} \|\boldsymbol{\beta}\|^2_{\infty,\mathcal{T}_T} h_F^{-1} \|[\![v]\!]\|^2_{0,\mathcal{F}_F\cap\mathfrak{F}_T} \Bigg\}, \end{aligned} \tag{3.18}$$

*where $m_{\mathcal{T}_T}$ is defined in Theorem 3.2 and $\mathcal{F}_F$ collects the faces of the one or two elements in $\mathcal{T}_F$. Then,*

$$\widetilde{\eta} \le \tilde{C}(|\!|\!| u - u_h |\!|\!|_{\oplus} + |\!|\!| u - u_h |\!|\!|_{\#,\mathcal{F}_h}), \tag{3.19}$$

*where the constant $\tilde{C}$ depends on the same parameters as the constant $C$ in Theorem 3.2 and in addition on the polynomial degree $m$ of $f$, $\boldsymbol{\beta}$, and $\mu$, the ratios $C_{\mathbf{K},T}/c_{\mathbf{K},T}$ and $(\|\mu\|_{\infty,T} + \|\frac{1}{2}\nabla\cdot\boldsymbol{\beta}\|_{\infty,T})/c_{\boldsymbol{\beta},\mu,T}$ for all $T \in \mathcal{T}_h$, and the ratios $c_{\boldsymbol{\beta},\mu,T}/c_{\boldsymbol{\beta},\mu,T'}$ for all $T, T' \in \mathcal{T}_h$ sharing a face.*

### 3.3. *Fully robust equivalence result*

This section contains the final result of this paper, namely a fully robust equivalence result between the error measured in the $(|\!|\!|\cdot|\!|\!|_{\oplus} + |\!|\!|\cdot|\!|\!|_{\#,\mathcal{F}_h})$-norm and a suitable a posteriori estimate. This result is an immediate consequence of Theorems 3.3 and 3.4.

**Theorem 3.5 (Fully robust equivalence between error and a posteriori estimate).** *Let $u$ be the solution of* (2.8) *and let $u_h$ be its DG approximation solving* (2.14). *Then,*

$$|||u - u_h|||_{\oplus} + |||u - u_h|||_{\#,\mathcal{F}_h} \leq \widetilde{\eta} + |||u_h|||_{\#,\mathcal{F}_h} \leq \tilde{C}(|||u - u_h|||_{\oplus} + |||u - u_h|||_{\#,\mathcal{F}_h}), \tag{3.20}$$

*where $\tilde{C}$ is the constant in* (3.19).

**Remark 3.8 (Comparison with the results of Ref. 31).** The result of Theorem 3.5 is in its form comparable with that reported in Ref. 31. One essential difference is, however, that our discrete jump seminorm $|||\cdot|||_{\#,\mathcal{F}_h}$ contains the cutoff factors $m_{\mathcal{T}_T}$ in front of $\|\boldsymbol{\beta}\|_{\infty,\mathcal{T}_T} h_F^{-1/2} \|[\![v]\!]\|_{0,\mathcal{F}_F \cap \mathfrak{F}_T}$, which can considerably reduce the size of this term. Moreover, we stress that the a posteriori estimate $\widetilde{\eta} + |||u_h|||_{\#,\mathcal{F}_h}$ is fully computable with no undetermined constants.

**Remark 3.9 ($|||\cdot|||_{\#,\mathcal{F}_h}$-seminorm).** It can be argued that the discrete seminorm $|||\cdot|||_{\#,\mathcal{F}_h}$ is not fully satisfactory since it does not appear in the natural DG stability norm. In particular, a priori error estimates including this new seminorm have not been established. Moreover, the $|||\cdot|||_{\#,\mathcal{F}_h}$-seminorm is not easily localizable with respect to data.

**Remark 3.10 (Pure diffusion).** In the pure diffusion case, the augmented norm $|||\cdot|||_{\oplus}$ coincides with the energy norm $|||\cdot|||$ and the jump seminorm $|||\cdot|||_{\#,\mathcal{F}_h}$ reduces to the first term in the right-hand side of (3.18). The result of Theorem 3.5 then provides a mean to circumvent any assumption on the distribution of diffusion inhomogeneities (such as those in Ref. 1, 7) to infer a robust equivalence result with respect to diffusion inhomogeneities.

## 4. Proofs

This section collects the proofs of the results presented in Section 3.

### 4.1. *Energy norm estimates*

**Lemma 4.1 (Abstract energy norm estimate).** *Let $u$ be the solution of* (2.8) *and let $u_h \in H^1(\mathcal{T}_h)$ be arbitrary. Then,*

$$\begin{aligned} |||u - u_h||| \leq &\inf_{s \in H_0^1(\Omega)} \Big\{ |||u_h - s||| \\ &+ \inf_{\mathbf{t},\mathbf{q} \in \mathbf{H}(\mathrm{div},\Omega)} \sup_{\varphi \in H_0^1(\Omega),\, |||\varphi|||=1} \big\{ (f - \nabla\cdot\mathbf{t} - \nabla\cdot\mathbf{q} - (\mu - \nabla\cdot\boldsymbol{\beta}) u_h, \varphi) \\ &- (\mathbf{K}\nabla_h u_h + \mathbf{t}, \nabla\varphi) + (\nabla\cdot\mathbf{q} - \nabla\cdot(\boldsymbol{\beta} s), \varphi) - \big(\tfrac{1}{2}(\nabla\cdot\boldsymbol{\beta})(u_h - s), \varphi\big) \big\} \Big\} \\ \leq &\, 2|||u - u_h|||. \end{aligned} \tag{4.1}$$

**Proof.** It has been proved in Lemma 7.1 of Ref. 37 and Lemma 3.1 of Ref. 18 that

$$|||u-u_h||| \le \inf_{s\in H_0^1(\Omega)}\left\{|||u_h-s||| + \sup_{\varphi\in H_0^1(\Omega),\,|||\varphi|||=1}\big\{\mathcal{B}(u-u_h,\varphi)+\mathcal{B}_{\rm A}(u_h-s,\varphi)\big\}\right\}.$$

It suffices to use (2.8) therein, to introduce arbitrary fields $\mathbf{t},\mathbf{q}\in\mathbf{H}(\mathrm{div},\Omega)$, add and subtract $(\mathbf{t},\nabla\varphi)$ and $(\mathbf{q},\nabla\varphi)$, and to employ the Green theorem to infer the upper error bound in (4.1). For the lower error bound, put $s=u$, $\mathbf{t}=-\mathbf{K}\nabla u$, and $\mathbf{q}=\boldsymbol{\beta}u$ and use the Cauchy–Schwarz inequality and the fact that $|||\varphi|||=1$. □

**Proof.** [Proof of Theorem 3.1] We start by putting $s=s_h$, $\mathbf{t}=\mathbf{t}_h$, and $\mathbf{q}=\mathbf{q}_h$ in the upper error bound (4.1). We next write

$$\begin{aligned}
&(f-\nabla\cdot\mathbf{t}_h-\nabla\cdot\mathbf{q}_h-(\mu-\nabla\cdot\boldsymbol{\beta})u_h,\varphi)-(\mathbf{K}\nabla_h u_h+\mathbf{t}_h,\nabla\varphi)+(\nabla\cdot\mathbf{q}_h-\nabla\cdot(\boldsymbol{\beta}s_h),\varphi)\\
&-\big(\tfrac12(\nabla\cdot\boldsymbol{\beta})(u_h-s_h),\varphi\big)=\sum_{T\in\mathcal{T}_h}\Big\{(f-\nabla\cdot\mathbf{t}_h-\nabla\cdot\mathbf{q}_h-(\mu-\nabla\cdot\boldsymbol{\beta})u_h,\varphi-\Pi_0\varphi)_{0,T}\\
&-(\mathbf{K}\nabla u_h+\mathbf{t}_h,\nabla\varphi)_{0,T}-\big(\tfrac12(\nabla\cdot\boldsymbol{\beta})(u_h-s_h),\varphi\big)_{0,T}+(\nabla\cdot(\mathbf{q}_h-\boldsymbol{\beta}s_h),\varphi-\Pi_0\varphi)_{0,T}\\
&+\sum_{F\in\mathcal{F}_T}((\mathbf{q}_h-\boldsymbol{\beta}s_h)\cdot\mathbf{n}_T,\Pi_0\varphi)_{0,F}\Big\},
\end{aligned}\tag{4.2}$$

using Lemma 2.1 in the first term and subtracting $(\nabla\cdot(\mathbf{q}_h-\boldsymbol{\beta}s_h),\Pi_0\varphi)_{0,T}$ and adding the same quantity rewritten using the Green theorem in the last two terms. Next, in these last two terms, we replace $\nabla\cdot(\mathbf{q}_h-\boldsymbol{\beta}s_h)$ by $(Id-\Pi_0)(\nabla\cdot(\mathbf{q}_h-\boldsymbol{\beta}s_h))$ and $(\mathbf{q}_h-\boldsymbol{\beta}s_h)\cdot\mathbf{n}_T$ by $\Pi_{0,F}((\mathbf{q}_h-\boldsymbol{\beta}s_h)\cdot\mathbf{n}_T)$. Furthermore, following Ref. 12, there are two ways to bound the term $-(\mathbf{K}\nabla u_h+\mathbf{t}_h,\nabla\varphi)_{0,T}$. Either one simply uses

$$-(\mathbf{K}\nabla u_h+\mathbf{t}_h,\nabla\varphi)_{0,T}\le\eta^{(1)}_{\mathrm{DF},T}|||\varphi|||_T,$$

or one notices using (2.23) and (2.24) that

$$\begin{aligned}
-(\mathbf{K}\nabla u_h+\mathbf{t}_h,\nabla\varphi)_{0,T}&=-(\mathbf{K}\nabla u_h+\mathbf{t}_h,\nabla(\varphi-\Pi_0\varphi))_{0,T}\\
&=(\nabla\cdot(\mathbf{K}\nabla u_h+\mathbf{t}_h),\varphi-\Pi_0\varphi)_{0,T}\\
&\quad-\sum_{F\in\mathcal{F}_T}((\mathbf{K}\nabla u_h+\mathbf{t}_h)\cdot\mathbf{n}_T,\varphi-\Pi_0\varphi)_{0,F}\le\eta^{(2)}_{\mathrm{DF},T}|||\varphi|||_T.
\end{aligned}$$

Finally, using (2.25) and the continuity of the normal component of $(\mathbf{q}_h-\boldsymbol{\beta}s_h)$ for the last term in (4.2), it is inferred that

$$\sum_{T\in\mathcal{T}_h}\sum_{F\in\mathcal{F}_T}(\Pi_{0,F}((\mathbf{q}_h-\boldsymbol{\beta}s_h)\cdot\mathbf{n}_T),\Pi_0\varphi)_{0,F}\le\sum_{T\in\mathcal{T}_h}\eta_{\mathrm{U},T}|||\varphi|||_T.$$

Collecting the above bounds leads to

$$\begin{aligned}
&(f-\nabla\cdot\mathbf{t}_h-\nabla\cdot\mathbf{q}_h-(\mu-\nabla\cdot\boldsymbol{\beta})u_h,\varphi)-(\mathbf{K}\nabla_h u_h+\mathbf{t}_h,\nabla\varphi)+(\nabla\cdot\mathbf{q}_h-\nabla\cdot(\boldsymbol{\beta}s_h),\varphi)\\
&-\big(\tfrac12(\nabla\cdot\boldsymbol{\beta})(u_h-s_h),\varphi\big)\le\sum_{T\in\mathcal{T}_h}\Big(\eta_{\mathrm{R},T}+\eta_{\mathrm{DF},T}+\eta_{\mathrm{C},1,T}+\eta_{\mathrm{C},2,T}+\eta_{\mathrm{U},T}\Big)|||\varphi|||_T,
\end{aligned}$$

whence the conclusion is straightforward. □

**Proof.** [Proof of Theorem 3.2] Let $C$ denote a generic constant depending on the parameters as in the statement of the theorem. Let $T \in \mathcal{T}_h$. The proof is decomposed in two parts.
(1) Bounds on the estimators involving $s_h = \mathcal{I}_{\mathrm{Os}}(u_h)$. First, consider $\eta_{\mathrm{NC},T}$ and recall the estimate

$$\|\nabla(u_h - s_h)\|_{0,T} \leq C \sum_{F \in \mathfrak{F}_T} h_F^{-1/2} \|[\![u_h]\!]\|_{0,F},$$

proved in Theorem 2.2 of Ref. 21. Using this bound, the fact that $[\![u - u_h]\!] = -[\![u_h]\!]$ and owing to (2.16), it is easy to see that

$$\|\mathbf{K}^{\frac{1}{2}} \nabla(u_h - s_h)\|_{0,T} \leq C \frac{C_{\mathbf{K},T}^{1/2}}{c_{\mathbf{K},\mathfrak{T}_T}^{1/2}} |||u - u_h|||_{*,\mathfrak{F}_T}.$$

Furthermore, it is well-known (see, *e.g.*, Lemma 3.2 in Ref. 9) that

$$\|u_h - s_h\|_{0,T} \leq C \sum_{F \in \mathfrak{F}_T} h_F^{1/2} \|[\![u_h]\!]\|_{0,F},$$

and it follows from (2.15) that

$$\sum_{F \in \mathfrak{F}_T} \|[\![u_h]\!]\|_{0,F} \leq C h_T^{-1/2} \chi_{\mathfrak{T}_T} |||u - u_h|||_{*,\mathfrak{F}_T}, \tag{4.3}$$

with $\chi_{\mathfrak{T}_T}$ defined by (3.11). Hence,

$$\|u_h - s_h\|_{0,T} \leq C \chi_{\mathfrak{T}_T} |||u - u_h|||_{*,\mathfrak{F}_T}. \tag{4.4}$$

The bound on $\eta_{\mathrm{NC},T}$ is now straightforward. Moreover, the bound on $\eta_{\mathrm{C},2,T}$ is readily inferred from (4.4). Considering next $\eta_{\mathrm{U},T}$, we observe that owing to (2.18) and the fact that $\|\Pi_{0,F} g\|_{0,F} \leq \|g\|_{0,F}$ for all $F \in \mathcal{F}_T$ and $g \in L^2(F)$,

$$\begin{aligned}
\|\Pi_{0,F}((\mathbf{q}_h - \boldsymbol{\beta} s_h)\cdot\mathbf{n}_F)\|_{0,F} &= \|\Pi_{0,F}(\boldsymbol{\beta}\cdot\mathbf{n}_F \{\!\{u_h\}\!\} + \gamma_{\boldsymbol{\beta},F}[\![u_h]\!] - \boldsymbol{\beta}\cdot\mathbf{n}_F s_h)\|_{0,F} \\
&\leq \|\boldsymbol{\beta}\cdot\mathbf{n}_F \{\!\{u_h\}\!\} + \gamma_{\boldsymbol{\beta},F}[\![u_h]\!] - \boldsymbol{\beta}\cdot\mathbf{n}_F s_h\|_{0,F} \\
&\leq C \|\boldsymbol{\beta}\|_{\infty,T} \sum_{F' \in \mathfrak{F}_T} \|[\![u_h]\!]\|_{0,F'},
\end{aligned}$$

since $\|u_h - s_h\|_{0,F} \leq C \sum_{F' \in \mathfrak{F}_T} \|[\![u_h]\!]\|_{0,F'}$ for the Oswald interpolate. Hence, using (4.3) and the fact that $m_F \leq C h_T^{-1/2} m_{\mathcal{T}_T}$, the bound on $\eta_{\mathrm{U},T}$ is inferred. Finally, to prove the bound on $\eta_{\mathrm{C},1,T}$, we observe that

$$\|(Id - \Pi_0)(\nabla\cdot(\mathbf{q}_h - \boldsymbol{\beta} s_h))\|_{0,T} \leq \|\nabla\cdot(\mathbf{q}_h - \boldsymbol{\beta} s_h)\|_{0,T} = \sup_{\xi \in \mathbb{P}_l(T)} \frac{(\nabla\cdot(\mathbf{q}_h - \boldsymbol{\beta} s_h), \xi)_{0,T}}{\|\xi\|_{0,T}},$$

using the assumption that $\nabla\cdot(\mathbf{q}_h - \boldsymbol{\beta} s_h) \in \mathbb{P}_l(T)$. Using the Green theorem and (2.20) yields

$$(\nabla\cdot(\mathbf{q}_h - \boldsymbol{\beta} s_h), \xi)_{0,T} = -(u_h - s_h, \boldsymbol{\beta}\cdot\nabla\xi)_{0,T} + \sum_{F \in \mathcal{F}_T} ((\mathbf{q}_h - \boldsymbol{\beta} s_h)\cdot\mathbf{n}_T, \xi)_{0,F}.$$

Using the Cauchy–Schwarz inequality, the bound (4.4), the fact that $(\mathbf{q}_h - \boldsymbol{\beta} s_h)\cdot\mathbf{n}_F = \boldsymbol{\beta}\cdot\mathbf{n}_F\{\!\{u_h\}\!\} + \gamma_{\boldsymbol{\beta},F}[\![u_h]\!] - \boldsymbol{\beta}\cdot\mathbf{n}_F s_h$ has been bounded above, and inverse inequalities to estimate $\|\nabla\xi\|_{0,T}$ and $\|\xi\|_{0,F}$, the bound on $\eta_{\mathrm{C},1,T}$ is inferred.
(2) Bounds on $\eta_{\mathrm{R},T}$ and $\eta_{\mathrm{DF},T}$. Using the triangle inequality yields

$$\eta_{\mathrm{R},T} \le \rho_{1,T} + m_T\|\nabla\cdot(\mathbf{K}\nabla u_h + \mathbf{t}_h)\|_{0,T} + m_T\|\nabla\cdot(\mathbf{q}_h - \boldsymbol{\beta} u_h)\|_{0,T},$$

with $\rho_{1,T}$ defined by (3.9). To bound the last two terms in the right-hand side, we proceed as we did above for $\nabla\cdot(\mathbf{q}_h - \boldsymbol{\beta} s_h)$. Since $\nabla\cdot(\mathbf{q}_h - \boldsymbol{\beta} u_h) \in \mathbb{P}_l(T)$, it is easy to see that

$$m_T\|\nabla\cdot(\mathbf{q}_h - \boldsymbol{\beta} u_h)\|_{0,T} \le C m_T h_T^{-1}\|\boldsymbol{\beta}\|_{\infty,T}\chi_{\mathcal{T}_T}|\!|\!|u - u_h|\!|\!|_{*,\mathcal{F}_T}.$$

Similarly,

$$\sup_{\xi\in\mathbb{P}_l(T)}\frac{(\mathbf{K}\nabla u_h + \mathbf{t}_h, \nabla\xi)_{0,T}}{\|\xi\|_{0,T}} \le C\sum_{F\in\mathcal{F}_T}\gamma_{\mathbf{K},F}h_F^{-3/2}\|[\![u_h]\!]\|_{0,F},$$

and for all $F\in\mathcal{F}_T$, (2.17) yields

$$\|(\mathbf{K}\nabla u_h + \mathbf{t}_h)\cdot\mathbf{n}_F\|_{0,F} \le C(\bar{\omega}_{T,F}\|\mathbf{n}_F\cdot[\![\mathbf{K}\nabla u_h]\!]\|_{0,F} + \gamma_{\mathbf{K},F}h_F^{-1}\|[\![u_h]\!]\|_{0,F}). \quad (4.5)$$

Hence,

$$\|\nabla\cdot(\mathbf{K}\nabla u_h + \mathbf{t}_h)\|_{0,T} \le C\sum_{F\in\mathcal{F}_T}(\gamma_{\mathbf{K},F}h_F^{-3/2}\|[\![u_h]\!]\|_{0,F} + h_F^{-1/2}\bar{\omega}_{T,F}\|\mathbf{n}_F\cdot[\![\mathbf{K}\nabla u_h]\!]\|_{0,F}).$$

As a result,

$$m_T\|\nabla\cdot(\mathbf{K}\nabla u_h + \mathbf{t}_h)\|_{0,T} \le C\left(\varsigma_T^2\frac{C_{\mathbf{K},T}^{1/2}}{c_{\mathbf{K},T}^{1/2}}|\!|\!|u - u_h|\!|\!|_{*,\mathcal{F}_T} + \varsigma_T\rho_{2,T}\right),$$

with $\rho_{2,T}$ defined by (3.10), whence the bound on $\eta_{\mathrm{R},T}$ is inferred. Finally, since $\eta_{\mathrm{DF},T} \le \eta_{\mathrm{DF},T}^{(2)}$ owing to (3.3), it suffices to bound $\eta_{\mathrm{DF},T}^{(2)}$. The volume term in (3.5) can be bounded as above since $\|(Id - \Pi_0)g\|_{0,T} \le \|g\|_{0,T}$ for all $g \in L^2(T)$. For the face term, we use (4.5) and the estimate $\widetilde{m}_T \le C m_T c_{\mathbf{K},T}^{-1/2}$ proven in Ref. 12. □

**Remark 4.1 (Estimators $\eta_{\mathrm{C},1,T}$ and $\eta_{\mathrm{U},T}$).** As observed in Remark 4.1 of Ref. 38, subtracting or using mean values in the estimators $\eta_{\mathrm{C},1,T}$ and $\eta_{\mathrm{U},T}$ can only lower these quantities, with noteworthy improvements in some situations. These improvements were however not taken into account in the proof of Theorem 3.2. Hence, the actual efficiency of these estimators may still be better.

### 4.2. *Augmented norm estimates*

**Lemma 4.2 (Abstract augmented norm estimate).** *Let $u$ be the solution of* (2.8) *and let $u_h \in H^1(\mathcal{T}_h)$ be arbitrary. Then,*

$$
\begin{aligned}
|||u-u_h|||_\oplus \le 2 \inf_{s\in H_0^1(\Omega)} \Bigg\{ & |||u_h - s||| \\
& + \inf_{\mathbf{t},\mathbf{q}\in\mathbf{H}(\mathrm{div},\Omega)} \sup_{\varphi\in H_0^1(\Omega),\,|||\varphi|||=1} \big\{(f-\nabla\cdot\mathbf{t}-\nabla\cdot\mathbf{q}-(\mu-\nabla\cdot\boldsymbol{\beta})u_h,\varphi) \\
& - (\mathbf{K}\nabla_h u_h+\mathbf{t},\nabla\varphi) + (\nabla\cdot\mathbf{q}-\nabla\cdot(\boldsymbol{\beta}s),\varphi) - \big(\tfrac{1}{2}(\nabla\cdot\boldsymbol{\beta})(u_h-s),\varphi\big)\big\}\Bigg\} \\
& + \inf_{\mathbf{t}\in\mathbf{H}(\mathrm{div},\Omega)} \sup_{\varphi\in H_0^1(\Omega),\,|||\varphi|||=1} \Bigg\{(f-\nabla\cdot\mathbf{t}-\boldsymbol{\beta}\cdot\nabla_h u_h-\mu u_h,\varphi) \\
& - (\mathbf{K}\nabla_h u_h+\mathbf{t},\nabla\varphi) - \mathcal{B}_\mathrm{D}(u_h,\varphi)\Bigg\} \\
\le {} & 5|||u-u_h|||_\oplus .
\end{aligned} \tag{4.6}
$$

**Proof.** Using the definition of the $|||\cdot|||$- and $|||\cdot|||_\oplus$-norms, (2.7), the Cauchy–Schwarz inequality, and the fact that $\mathcal{B}_\mathrm{D}(u-u_h,\cdot) = -\mathcal{B}_\mathrm{D}(u_h,\cdot)$, it is inferred that

$$
|||u-u_h|||_\oplus \le 2|||u-u_h||| + \sup_{\varphi\in H_0^1(\Omega),\,|||\varphi|||=1} \{\mathcal{B}(u-u_h,\varphi) - \mathcal{B}_\mathrm{D}(u_h,\varphi)\}.
$$

For the first term, we simply use Lemma 4.1. For the second term, we use (2.8), add and subtract $(\mathbf{t},\nabla\varphi)$ for an arbitrary $\mathbf{t}\in\mathbf{H}(\mathrm{div},\Omega)$, and employ the Green theorem. This yields the upper error bound. For the lower error bound, it suffices to use again Lemma 4.1 for the first term and the fact that

$$
\mathcal{B}(u-u_h,\varphi)-\mathcal{B}_\mathrm{D}(u_h,\varphi) = \mathcal{B}_\mathrm{S}(u-u_h,\varphi)+(\mathcal{B}_\mathrm{A}+\mathcal{B}_\mathrm{D})(u-u_h,\varphi) \le |||u-u_h|||_\oplus|||\varphi|||,
$$

for the second one. □

**Proof.** [Proof of Theorem 3.3] We start from the abstract estimate of Lemma 4.2. As the first term is bounded by $2\eta$ owing to Theorem 3.1, we only bound the second one where we put $\mathbf{t}=\mathbf{t}_h$. Proceeding as in the proof of Theorem 3.1 leads to

$$
\begin{aligned}
& (f-\nabla\cdot\mathbf{t}_h-\boldsymbol{\beta}\cdot\nabla_h u_h-\mu u_h,\varphi) - (\mathbf{K}\nabla_h u_h+\mathbf{t}_h,\nabla\varphi) - \mathcal{B}_\mathrm{D}(u_h,\varphi) \\
= & \sum_{T\in\mathcal{T}_h} \Bigg\{ (f-\nabla\cdot\mathbf{t}_h-\nabla\cdot\mathbf{q}_h-(\mu-\nabla\cdot\boldsymbol{\beta})u_h,\varphi-\Pi_0\varphi)_{0,T} - (\mathbf{K}\nabla u_h+\mathbf{t}_h,\nabla\varphi)_{0,T} \\
& + (\nabla\cdot(\mathbf{q}_h-\boldsymbol{\beta}u_h),\varphi-\Pi_0\varphi)_{0,T} + \sum_{F\in\mathcal{F}_T}((\mathbf{q}_h-\boldsymbol{\beta}u_h)\cdot\mathbf{n}_T,\Pi_0\varphi)_{0,F}\Bigg\} - \mathcal{B}_\mathrm{D}(u_h,\varphi) \\
\le & \sum_{T\in\mathcal{T}_h} (\eta_{\mathrm{R},T}+\eta_{\mathrm{DF},T}+\widetilde{\eta}_{\mathrm{C},1,T}+\widetilde{\eta}_{\mathrm{U},T})|||\varphi|||_T .
\end{aligned}
$$

For the last two terms, letting $\mathbf{y}_h = \mathbf{q}_h - \boldsymbol{\beta} u_h$, we have used the relation

$$\sum_{T\in\mathcal{T}_h}\sum_{F\in\mathcal{F}_T}(\mathbf{y}_h\cdot\mathbf{n}_T,\Pi_0\varphi)_{0,F} = \sum_{F\in\mathcal{F}_h}(\mathbf{n}_F\cdot[\![\mathbf{y}_h]\!],\{\!\{\Pi_0\varphi\}\!\})_{0,F} + (\mathbf{n}_F\cdot\{\!\{\mathbf{y}_h\}\!\},[\![\Pi_0\varphi]\!])_{0,F}$$
$$= \mathcal{B}_{\mathrm{D}}(u_h,\varphi) + \sum_{F\in\mathcal{F}_h}(\Pi_{0,F}(\gamma_{\boldsymbol{\beta},F}[\![u_h]\!]),[\![\Pi_0\varphi]\!])_{0,F},$$

and the right-hand side is estimated using (2.25), leading to the $\widetilde{\eta}_{\mathrm{U},T}$ estimator. □

**Remark 4.2 (Role of $\mathcal{B}_{\mathrm{D}}$ in the augmented norm).** Adding the bilinear form $\mathcal{B}_{\mathrm{D}}$ to the augmented norm plays an important role in that it eliminates the term $\mathcal{B}_{\mathrm{D}}(u_h,\varphi)$ from the above expression.

**Proof.** [Proof of Theorem 3.4] Let $\tilde{C}$ denote a generic constant depending on the parameters as in the statement of the theorem. Proceeding as in the proof of Theorem 3.2 and using similar bounds on the estimators $\widetilde{\eta}_{\mathrm{C},1,T}$ and $\widetilde{\eta}_{\mathrm{U},T}$ of Theorem 3.3, it is inferred that

$$\widetilde{\eta} \le \tilde{C}\left\{\sum_{T\in\mathcal{T}_h}(\rho_{1,T}^2+\rho_{2,T}^2)\right\}^{1/2} + \tilde{C}|||u_h|||_{\#,\mathcal{F}_h},$$

were $\rho_{1,T}$ and $\rho_{2,T}$ are defined by (3.9)–(3.10). Since $|||u_h|||_{\#,\mathcal{F}_h} = |||u-u_h|||_{\#,\mathcal{F}_h}$, it remains to bound the contributions from the residuals $\rho_{1,T}$ and $\rho_{2,T}$. For all $T\in\mathcal{T}_h$, let $\psi_T$ be the element bubble function introduced by Verfürth,[33] $R_T := (f+\nabla\cdot(\mathbf{K}\nabla u_h) - \boldsymbol{\beta}\cdot\nabla u_h - \mu u_h)|_T$ and $\Psi_T := \psi_T R_T$. Observe that

$$\sum_{T\in\mathcal{T}_h}\rho_{1,T}^2 \le \tilde{C}\sum_{T\in\mathcal{T}_h} m_T^2\big(\mathcal{B}_{\mathrm{S}}(u-u_h,\Psi_T) + (\mathcal{B}_{\mathrm{A}}+\mathcal{B}_{\mathrm{D}})(u-u_h,\Psi_T) - \mathcal{B}_{\mathrm{D}}(u-u_h,\Psi_T)\big).$$

Since $m_T|||\Psi_T|||_T \le \tilde{C}\|R_T\|_{0,T}$ with a constant $\tilde{C}$ depending on the local ratios $C_{\mathbf{K},T}/c_{\mathbf{K},T}$ and $(\|\mu\|_{\infty,T} + \|\frac{1}{2}\nabla\cdot\boldsymbol{\beta}\|_{\infty,T})/c_{\boldsymbol{\beta},\mu,T}$, it is easy to see that the first two terms in the above right-hand side are bounded by $\|u-u_h\|_{\oplus}\{\sum_{T\in\mathcal{T}_h}\rho_{1,T}^2\}^{1/2}$. Concerning the last term, we use an inverse inequality to infer

$$\sum_{T\in\mathcal{T}_h} m_T^2\mathcal{B}_{\mathrm{D}}(u-u_h,\Psi_T) \le \tilde{C}\sum_{T\in\mathcal{T}_h}\sum_{F\in\mathcal{F}_T} m_T\|R_T\|_{0,T} m_T\|\boldsymbol{\beta}\|_{\infty,T} h_F^{-1/2}\|[\![u_h]\!]\|_{0,F},$$

which can be bounded by $|||u-u_h|||_{\#,\mathcal{F}_h}\{\sum_{T\in\mathcal{T}_h}\rho_{1,T}^2\}^{1/2}$. Consider now $\rho_{2,T}$. For all $F\in\mathcal{F}_h$, let $\psi_F$ be the face bubble function introduced by Verfürth in Ref. 33 (see also Ref. 18), $R_F := \mathbf{n}_F\cdot[\![\mathbf{K}\nabla_h u_h]\!]$, and let $\Psi_F$ be the lifting of $\psi_F R_F$ to $\mathcal{T}_F$. Observe that

$$\sum_{T\in\mathcal{T}_h}\rho_{2,T}^2 \le \tilde{C}\sum_{T\in\mathcal{T}_h}\sum_{F\in\mathcal{F}_T} m_T c_{\mathbf{K},T}^{-1/2}\bar{\omega}_{T,F}^2\Bigg\{-\mathcal{B}_{\mathrm{S}}(u-u_h,\Psi_F) - (\mathcal{B}_{\mathrm{A}}+\mathcal{B}_{\mathrm{D}})(u-u_h,\Psi_F)$$
$$+\mathcal{B}_{\mathrm{D}}(u-u_h,\Psi_F) + \sum_{T'\in\mathcal{T}_F}(R_{T'},\Psi_F)_{0,T'}\Bigg\} := T_1+T_2+T_3+T_4.$$

We first consider $T_1$ and observe that (up to a multiplicative constant $\tilde{C}$)

$$
\begin{aligned}
|T_1| &\leq \sum_{T\in\mathcal{T}_h}\sum_{F\in\mathcal{F}_T} m_T c_{\mathbf{K},T}^{-1/2}\bar{\omega}_{T,F}^2 \sum_{T'\in\mathcal{T}_F} |||u-u_h|||_{T'}|||\Psi_F|||_{T'} \\
&\leq \sum_{T\in\mathcal{T}_h}\sum_{F\in\mathcal{F}_T} m_T^{1/2} c_{\mathbf{K},T}^{-1/4}\bar{\omega}_{T,F}\|R_F\|_{0,F}\sum_{T'\in\mathcal{T}_F}(m_T^{1/2}c_{\mathbf{K},T}^{-1/4}\bar{\omega}_{T,F}m_{T'}^{-1/2}c_{\mathbf{K},T'}^{1/4})|||u-u_h|||_{T'} \\
&\leq \sum_{T\in\mathcal{T}_h}\sum_{F\in\mathcal{F}_T} m_T^{1/2} c_{\mathbf{K},T}^{-1/4}\bar{\omega}_{T,F}\|R_F\|_{0,F}\sum_{T'\in\mathcal{T}_F}|||u-u_h|||_{T'},
\end{aligned}
$$

since $|||\Psi_F|||_{T'} \leq \tilde{C} m_{T'}^{-1/2} c_{\mathbf{K},T'}^{1/4}\|R_F\|_{0,F}$ and since, owing to (2.11),

$$
m_T^{1/2}c_{\mathbf{K},T}^{-1/4}\bar{\omega}_{T,F}m_{T'}^{-1/2}c_{\mathbf{K},T'}^{1/4} \leq m_T^{1/2}\bar{\omega}_{T,F}^{1/2}m_{T'}^{-1/2} \leq \tilde{C}, \tag{4.7}
$$

with $\tilde{C}$ depending on the ratios $c_{\boldsymbol{\beta},\mu,T}/c_{\boldsymbol{\beta},\mu,T'}$. The bound on $T_2$ is similar (details are skipped for brevity) leading to $|T_1|+|T_2| \leq \tilde{C}\|u-u_h\|_{\oplus}\{\sum_{T\in\mathcal{T}_h}\rho_{2,T}^2\}^{1/2}$. We next consider $T_3$ and observe that (up to a multiplicative constant $\tilde{C}$)

$$
\begin{aligned}
|T_3| &\leq \sum_{T\in\mathcal{T}_h}\sum_{F\in\mathcal{F}_T} m_T c_{\mathbf{K},T}^{-1/2}\bar{\omega}_{T,F}^2\|\boldsymbol{\beta}\|_{\infty,\mathcal{T}_T}\sum_{F'\in\mathcal{F}_F}\|[\![u_h]\!]\|_{0,F'}\|\{\!\{\Pi_0\Psi_F\}\!\}\|_{0,F'} \\
&\leq \sum_{T\in\mathcal{T}_h}\sum_{F\in\mathcal{F}_T} m_T^{1/2}c_{\mathbf{K},T}^{-1/4}\bar{\omega}_{T,F}\|R_F\|_{0,F}m_{\mathcal{T}_T}\|\boldsymbol{\beta}\|_{\infty,\mathcal{T}_T}\sum_{F'\in\mathcal{F}_F}h_{F'}^{-1/2}\|[\![u_h]\!]\|_{0,F'},
\end{aligned}
$$

where we have used the inverse inequality $\|\{\!\{\Pi_0\Psi_F\}\!\}\|_{0,F'} \leq \tilde{C}h_{F'}^{-1/2}\|\Psi_F\|_{0,\mathcal{T}_{F'}\cap\mathcal{T}_F}$, the fact that $\|\Psi_F\|_{0,T'} \leq \tilde{C}m_{T'}^{1/2}c_{\mathbf{K},T'}^{1/4}\|R_F\|_{0,F}$, and the bound (4.7). This yields $|T_3| \leq \tilde{C}|||u-u_h|||_{\#,\mathcal{F}_h}\{\sum_{T\in\mathcal{T}_h}\rho_{2,T}^2\}^{1/2}$. Finally, we proceed similarly to bound $T_4$ to obtain $|T_4| \leq \tilde{C}\{\sum_{T\in\mathcal{T}_h}\rho_{1,T}^2\}^{1/2}\{\sum_{T\in\mathcal{T}_h}\rho_{2,T}^2\}^{1/2}$. Using the previous estimate for $\{\sum_{T\in\mathcal{T}_h}\rho_{1,T}^2\}^{1/2}$ completes the proof. □

## 5. Numerical results

We consider the domain $\Omega = \{0 < x, y < 1\}$, the reaction coefficient $\mu = 1$, the velocity field $\boldsymbol{\beta} = (1,0)^t$, and an isotropic homogeneous diffusion tensor represented by a diffusion coefficient $\epsilon$. We run tests with $\epsilon = 10^{-2}$ and $\epsilon = 10^{-4}$. The source term $f$ is such that the exact solution with homogeneous Dirichlet boundary conditions is $u = \frac{1}{2}x(x-1)y(y-1)\,(1-\tanh(10-20x))$. For brevity, only results for uniformly refined structured meshes are presented. In the tables below, $N$ is the number of mesh elements. In the present setting, the jump seminorm $|||\cdot|||_{\#,\mathcal{F}_h}$ defined by (3.18) can be evaluated for $v \in H^1(\mathcal{T}_h)$ as

$$
|||v|||_{\#,\mathcal{F}_h}^2 = \sum_{F\in\mathcal{F}_h}(\tfrac{1}{2}\alpha_F\epsilon h_F^{-1} + h_F + m_F^2 h_F^{-1})\|[\![v]\!]\|_{0,F}^2, \tag{5.1}
$$

with $m_F = \min(h_F\epsilon^{-1/2}, 1)$ replacing $h_T$ by $h_F$ in the definition of $m_{\mathcal{T}_T}$. Moreover, observing that for all $\varphi \in H_0^1(\Omega)$,

$$
\mathcal{B}_{\mathrm{A}}(v,\varphi) + \mathcal{B}_{\mathrm{D}}(v,\varphi) = -(v,\boldsymbol{\beta}\cdot\nabla\varphi) + \sum_{F\in\mathcal{F}_h}(\boldsymbol{\beta}\cdot\mathbf{n}_F[\![v]\!],\{\!\{\varphi-\Pi_0\varphi\}\!\})_{0,F}, \tag{5.2}
$$

| | energy norm | | | augmented norm | | | |
|---|---|---|---|---|---|---|---|
| $N$ | err. | est. | eff. | err. | est. | eff. | $\|\|\|u_h\|\|\|_{\#,\mathcal{F}_h}$ |
| 128 | 7.74e-3 | 1.10e-1 | 14 | 1.40e-1 | 3.28e-1 | 2.3 | 3.40e-2 |
| 512 | 4.03e-3 | 4.35e-2 | 11 | 3.97e-2 | 1.29e-1 | 3.3 | 1.16e-2 |
| 2048 | 1.88e-3 | 1.43e-2 | 7.6 | 9.77e-3 | 4.14e-2 | 4.2 | 2.72e-3 |
| 8192 | 9.30e-4 | 3.58e-3 | 3.8 | 2.98e-3 | 1.02e-2 | 3.4 | 8.25e-4 |
| order | 1.0 | 2.0 | - | 1.7 | 2.0 | - | 1.7 |

Table 1. Errors ($|||u-u_h|||$ and $|||u-u_h|||_{\oplus'} + |||u-u_h|||_{\#,\mathcal{F}_h}$), estimates ($\eta$ and $\tilde{\eta} + |||u_h|||_{\#,\mathcal{F}_h}$), and effectivity indices as evaluated from (5.4) for the energy and augmented norms; $\epsilon = 10^{-2}$

| | | | $l=0$ | | $l=1$ | | |
|---|---|---|---|---|---|---|---|
| $N$ | $\eta_{\rm NC}$ | $\eta_{\rm U}$ | $\eta_{\rm R}$ | $\eta_{\rm DF}$ | $\eta_{\rm R}$ | $\eta_{\rm DF}$ | $\eta_{\rm C,1}$ |
| 128 | 4.29e-3 | 6.29e-2 | 3.81e-2 | 8.10e-3 | 1.03e-2 | 8.66e-3 | 3.24e-2 |
| 512 | 1.91e-3 | 2.87e-2 | 9.91e-3 | 3.79e-3 | 1.82e-3 | 4.71e-3 | 7.71e-3 |
| 2048 | 8.87e-4 | 9.77e-3 | 2.42e-3 | 1.42e-3 | 3.19e-4 | 2.16e-3 | 1.53e-3 |
| 8192 | 4.13e-4 | 2.11e-3 | 6.12e-4 | 4.97e-4 | 4.07e-5 | 8.40e-4 | 3.38e-4 |
| order | 1.1 | 2.2 | 2.0 | 1.5 | 3.0 | 1.4 | 2.2 |

Table 2. Estimators contributing to $\eta$ for $l=0$ and $l=1$; $\epsilon = 10^{-2}$

and using (2.24), the following upper bound on the augmented norm is inferred:

$$|||v|||_{\oplus} \le |||v|||_{\oplus'} := |||v||| + \epsilon^{-1/2}\|v\| + \left\{ \sum_{T\in\mathcal{T}_h} \sum_{F\in\mathcal{F}_T} C_{{\rm t},T,F} \widetilde{m}_T \|[\![v]\!]\|_{0,F}^2 \right\}^{1/2}. \quad (5.3)$$

We will use this computable bound on $|||v|||_{\oplus}$ and consider two effectivity indices,

$$\frac{\eta}{|||u-u_h|||} \quad \text{and} \quad \frac{\widetilde{\eta} + |||u_h|||_{\#,\mathcal{F}_h}}{|||u-u_h|||_{\oplus'} + |||u-u_h|||_{\#,\mathcal{F}_h}}, \quad (5.4)$$

illustrating the results of Theorems 3.1 and 3.5.

For $\epsilon = 10^{-2}$, convective effects dominate on the coarsest meshes, while the local Péclet number is of order unity on the finest mesh. Table 1 presents the errors, estimates, and effectivity indices as evaluated from (5.4) for the energy and augmented norms. The diffusive and convective fluxes are reconstructed using $l=0$; very similar results are obtained for $l=1$. For the energy norm, the effectivity index decreases from 14 to 3.8, reflecting the decrease in the local Péclet number. On the contrary, for the augmented norm, the effectivity index remains fairly stable and takes values around 3. We also observe that in the augmented norm, the energy norm contribution is very small and that the $|||\cdot|||_{\#,\mathcal{F}_h}$-seminorm contribution is not significant either. On the finest meshes, the energy norm and the $|||\cdot|||_{\#,\mathcal{F}_h}$-seminorm take similar values.

| | energy norm | | | augmented norm | | | |
|---|---|---|---|---|---|---|---|
| $N$ | err. | est. | eff. | err. | est. | eff. | $\|\|\|u_h\|\|\|_{\#,\mathcal{F}_h}$ |
| 128 | 1.70e-3 | 1.34e-1 | 79 | 3.67e-1 | 4.05e-1 | 1.10 | 4.02e-2 |
| 512 | 5.65e-4 | 7.01e-2 | 124 | 1.44e-1 | 2.11e-1 | 1.47 | 2.11e-2 |
| 2048 | 2.14e-4 | 3.09e-2 | 144 | 5.35e-2 | 9.36e-2 | 1.75 | 9.99e-3 |
| 8192 | 1.00e-4 | 1.25e-2 | 125 | 2.14e-2 | 3.89e-2 | 1.82 | 4.96e-3 |
| order | 1.1 | 1.3 | - | 1.3 | 1.3 | - | 1.0 |

Table 3. Errors ($|||u-u_h|||$ and $|||u-u_h|||_{\oplus'} + |||u-u_h|||_{\#,\mathcal{F}_h}$), estimates ($\eta$ and $\tilde{\eta} + |||u_h|||_{\#,\mathcal{F}_h}$), and effectivity indices as evaluated from (5.4) for the energy and augmented norms; $\epsilon = 10^{-4}$

| | | | $l=0$ | | $l=1$ | | |
|---|---|---|---|---|---|---|---|
| $N$ | $\eta_{\mathrm{NC}}$ | $\eta_{\mathrm{U}}$ | $\eta_{\mathrm{R}}$ | $\eta_{\mathrm{DF}}$ | $\eta_{\mathrm{R}}$ | $\eta_{\mathrm{DF}}$ | $\eta_{\mathrm{C},1}$ |
| 128 | 2.69e-3 | 6.91e-2 | 6.62e-2 | 3.42e-4 | 1.60e-2 | 6.25e-4 | 6.40e-2 |
| 512 | 6.76e-4 | 3.60e-2 | 3.43e-2 | 2.03e-4 | 4.55e-2 | 4.60e-4 | 3.39e-2 |
| 2048 | 1.66e-4 | 1.46e-2 | 1.63e-2 | 1.09e-4 | 2.01e-2 | 2.68e-4 | 1.60e-2 |
| 8192 | 6.78e-5 | 6.70e-3 | 5.81e-3 | 5.97e-5 | 3.66e-2 | 1.38e-4 | 5.68e-3 |
| order | 1.3 | 1.1 | 1.5 | 0.86 | 2.5 | 1.0 | 1.5 |

Table 4. Estimators contributing to $\eta$ for $l=0$ and $l=1$; $\epsilon = 10^{-4}$

A more detailed analysis of the estimators contributing to $\eta$ for $l=0$ and $l=1$ can be found in Table 2. The residual estimator $\eta_{\mathrm{R}}$ super-converges by one order for $l=0$ and by two orders for $l=1$. The diffusive flux estimator $\eta_{\mathrm{DF}}$ yields among the smallest contributions to the error estimate. The upwinding estimator $\eta_{\mathrm{U}}$ is dominant, along with the first convection estimator $\eta_{\mathrm{C},1}$ for $l=1$, while this latter estimator vanishes for $l=0$ since in this case, $\nabla\cdot(\mathbf{q}_h - \boldsymbol{\beta}\mathcal{I}_{\mathrm{Os}}(u_h))$ is by construction piecewise constant. Finally, the second convection estimator $\eta_{\mathrm{C},2}$ vanishes identically because $\boldsymbol{\beta}$ is divergence-free. All in all, there is little gain when going from $l=0$ to $l=1$.

Tables 3 and 4 report the results for $\epsilon = 10^{-4}$. In this case, the local Péclet number decreases from 1250 on the coarsest mesh to 150 on the finest mesh. For the energy norm, the effectivity index remains fairly constant, owing to the cutoff functions, but takes rather large values. On the contrary, for the augmented norm, the effectivity index is very close to the optimal value of 1 on all meshes. We also observe that the $|||\cdot|||_{\#,\mathcal{F}_h}$-seminorm contribution is larger than the energy norm, but smaller than the augmented norm. This important property is a consequence of the cutoff factors $m_{\mathcal{T}_T}$ in the $|||\cdot|||_{\#,\mathcal{F}_h}$-seminorm, see Remark 3.8. Finally, the results of Table 4 are similar to those of Table 2.

## Appendix A. Nonmatching meshes

This section briefly describes the modifications needed to extend the previous results to the case of nonmatching meshes.

### A.1. *The setting*

Let $\{\mathcal{T}_h\}_{h>0}$ be a family of simplicial, possibly nonmatching meshes of the domain $\Omega$. For each $\mathcal{T}_h$, there exists a matching simplicial submesh $\widehat{\mathcal{T}}_h$ of $\mathcal{T}_h$ such that $\widehat{\mathcal{T}}_h = \mathcal{T}_h$ if $\mathcal{T}_h$ is itself matching. For all $T \in \mathcal{T}_h$, we consider the refinement of $T$ by $\widehat{\mathcal{T}}_h$, namely

$$\mathfrak{R}_T = \{T' \in \widehat{\mathcal{T}}_h;\ T' \subset T\}.$$

Clearly, $\mathfrak{R}_T = \{T\}$ if $\mathcal{T}_h$ is matching. Furthermore, the set $\widehat{\mathcal{F}}_T$ collects the faces of $T \in \widehat{\mathcal{T}}_h$. We assume the following on the meshes:

(A1) $\{\widehat{\mathcal{T}}_h\}_{h>0}$ is shape-regular in the sense that there exists a constant $\kappa_{\widehat{\mathcal{T}}} > 0$ such that $\min_{T\in\widehat{\mathcal{T}}_h} |T|/h_T^d \geq \kappa_{\widehat{\mathcal{T}}}$ for all $h > 0$.

(A2) There exists a constant $\iota_{\mathcal{T}} > 0$ such that $\min_{T'\in\mathfrak{R}_T} h_{T'}/h_T \geq \iota_{\mathcal{T}}$ for all $T \in \mathcal{T}_h$ and all $h > 0$.

Observe that the above assumptions imply the shape-regularity of $\{\mathcal{T}_h\}_{h>0}$.

### A.2. *Flux reconstruction on nonmatching meshes*

The $\mathbf{H}(\mathrm{div},\Omega)$-conforming diffusive and convective fluxes $\mathbf{t}_h$ and $\mathbf{q}_h$ belong to the space $\mathbf{RTN}^l(\widehat{\mathcal{T}}_h)$ and are prescribed locally on all $T \in \widehat{\mathcal{T}}_h$ (instead of $T \in \widehat{\mathcal{T}}_h$) as follows: For all $F \in \widehat{\mathcal{F}}_T$ (instead of $F \in \mathcal{F}_T$) and all $q_h \in \mathbb{P}_l(F)$, (2.17) and (2.18) hold and for all $\mathbf{r}_h \in \mathbb{P}^d_{l-1}(T)$, (2.19) and (2.20) hold. Observe that $\alpha_F$, $\gamma_{\mathbf{K},F}$, $\gamma_{\boldsymbol{\beta},F}$, and $\omega_{T,F}$ need only be evaluated on the faces of $\mathcal{T}_h$ (where they are actually defined) since $[\![u_h]\!] = 0$ and $\{\!\{\mathbf{K}\nabla_h u_h\}\!\}_\omega = \mathbf{K}\nabla u_h$ on the remaining faces of $\widehat{\mathcal{T}}_h$. The above construction leads to the following extension of Lemma 2.1.

**Lemma A.1 (Local conservativity on nonmatching meshes).** *There holds*

$$(\nabla\cdot\mathbf{t}_h + \nabla\cdot\mathbf{q}_h + (\mu - \nabla\cdot\boldsymbol{\beta})u_h, \xi_h)_{0,T} = (f, \xi_h)_{0,T} \qquad \forall T \in \mathcal{T}_h,\ \forall \xi_h \in \mathbb{P}_l(T).$$

**Proof.** Let $T \in \mathcal{T}_h$ and let $\xi_h \in \mathbb{P}_l(T)$. Owing to the Green theorem,

$$\begin{aligned}(\nabla\cdot\mathbf{t}_h + \nabla\cdot\mathbf{q}_h, \xi_h)_{0,T} &= \sum_{T'\in\mathfrak{R}_T} (\nabla\cdot\mathbf{t}_h + \nabla\cdot\mathbf{q}_h, \xi_h)_{0,T'} \\ &= \sum_{T'\in\mathfrak{R}_T} -(\mathbf{t}_h, \nabla\xi_h)_{0,T'} + \sum_{T'\in\mathfrak{R}_T}\sum_{F\in\widehat{\mathcal{F}}_{T'}} (\mathbf{t}_h\cdot\mathbf{n}_{T'}, \xi_h)_{0,F} \\ &\quad + \sum_{T'\in\mathfrak{R}_T} -(\mathbf{q}_h, \nabla\xi_h)_{0,T'} + \sum_{T'\in\mathfrak{R}_T}\sum_{F\in\widehat{\mathcal{F}}_{T'}} (\mathbf{q}_h\cdot\mathbf{n}_{T'}, \xi_h)_{0,F}.\end{aligned}$$

To handle the volumetric terms, we use (2.19) and (2.20), $\nabla\xi_h|_{T'} \in \mathbb{P}_{l-1}(T')^d$ for all $T' \in \mathfrak{R}_T$, and $[\![u_h]\!] = 0$ on those faces $F \in \widehat{\mathcal{F}}_{T'}$ that lie in the interior of $T$. To handle the face terms, we use (2.17) and (2.18), the continuity of $\xi_h$ and that of the normal component of $\mathbf{t}_h$ in the interior of $T$ and the fact that $\xi_h|_F \in \mathbb{P}_l(F)$ for all $F \in \widehat{\mathcal{F}}_{T'}$ and all $T' \in \mathfrak{R}_T$. This yields (2.22), which by (2.14) implies the statement of the lemma. □

Similar developments considering only flux equilibration on subfaces in nonconforming meshes can be found in Ref. 3.

### A.3. *Modification of the estimators*

The approximation results (2.23)–(2.25) need to be employed on $\widehat{\mathcal{T}}_h$ and the cutoff functions $m_T$, $\widetilde{m}_T$, and $m_F$ as well as the constants $C_{\mathrm{t},T,F}$ and $C_{\mathrm{F},T,F}$ are redefined accordingly for all $T \in \widehat{\mathcal{T}}_h$ and $F \in \widehat{\mathcal{F}}_T$. The $\mathbf{H}(\mathrm{div},\Omega)$-conforming diffusive and convective fluxes $\mathbf{t}_h$ and $\mathbf{q}_h$ are reconstructed as above, while the $H^1_0(\Omega)$-conforming primal reconstruction $s_h$ is evaluated using the Oswald interpolate on the matching submesh $\widehat{\mathcal{T}}_h$. Then, for all $T \in \mathcal{T}_h$, the definition of the estimators $\eta_{\mathrm{NC},T}$, $\eta_{\mathrm{R},T}$, $\eta^{(1)}_{\mathrm{DF},T}$, and $\eta_{\mathrm{C},2,T}$ is kept unchanged while we set

$$\eta_{\mathrm{C},1,T} := \left\{ \sum_{T' \in \mathfrak{R}_T} m^2_{T'} \|(Id - \widehat{\Pi}_0)(\nabla\cdot(\mathbf{q}_h - \boldsymbol{\beta} s_h))\|^2_{0,T'} \right\}^{1/2}, \tag{A.1}$$

$$\eta_{\mathrm{U},T} := \left\{ \sum_{T' \in \mathfrak{R}_T} \left( \sum_{F \in \widehat{\mathcal{F}}_{T'}, F \cap \partial T \neq \emptyset} m_F \|\widehat{\Pi}_{0,F}((\mathbf{q}_h - \boldsymbol{\beta} s_h)\cdot\mathbf{n}_F)\|_{0,F} \right)^2 \right\}^{1/2}, \tag{A.2}$$

where $\widehat{\Pi}_0$ denotes the $L^2$-orthogonal projection onto $V^0(\widehat{\mathcal{T}}_h)$ and $\Pi_{l,F}$ the $L^2$-orthogonal projection onto $\mathbb{P}_0(F)$, and we also set

$$\begin{aligned} \eta^{(2)}_{\mathrm{DF},T} := \Bigg\{ \sum_{T' \in \mathfrak{R}_T} \Bigg( & m_{T'} \|(Id - \widehat{\Pi}_0)(\nabla\cdot(\mathbf{K}\nabla u_h + \mathbf{t}_h))\|_{0,T'} \\ & + \widetilde{m}^{1/2}_{T'} \sum_{F \in \widehat{\mathcal{F}}_{T'}, F \subset \partial T} C^{1/2}_{\mathrm{t},T',F} \|(\mathbf{K}\nabla u_h + \mathbf{t}_h)\cdot\mathbf{n}_F\|_{0,F} \Bigg)^2 \Bigg\}^{1/2}. \end{aligned} \tag{A.3}$$

Then, it can be verified that the results of Theorems 3.1 and 3.2 still hold, with the constant $\kappa_{\mathcal{T}}$ replaced by $\kappa_{\widehat{\mathcal{T}}}$ and $\iota_{\mathcal{T}}$.

Finally, the bilinear form $\mathcal{B}_{\mathrm{D}}$ is modified as

$$\mathcal{B}_{\mathrm{D}}(u,v) := -\sum_{F \in \widehat{\mathcal{F}}'_h} (\boldsymbol{\beta}\cdot\mathbf{n}_F[\![u]\!], \{\!\{\widehat{\Pi}_0 v\}\!\})_{0,F},$$

where $\widehat{\mathcal{F}}'_h = \{F \in \widehat{\mathcal{F}}_h;\ \exists T \in T_h, F \subset \partial T\}$, while the estimators $\widetilde{\eta}_{\mathrm{C},1,T}$ and $\widetilde{\eta}_{\mathrm{U},T}$ are modified similarly to the estimators $\eta_{\mathrm{C},1,T}$ and $\eta_{\mathrm{U},T}$ above. Then, it can be verified that the results of Theorems 3.3 and 3.4 still hold, with again the constant $\kappa_{\mathcal{T}}$ replaced by $\kappa_{\widehat{\mathcal{T}}}$ and $\iota_{\mathcal{T}}$.

## References


1. Ainsworth, M. Robust a posteriori error estimation for nonconforming finite element approximation. *SIAM J. Numer. Anal. 42*, 6 (2005), 2320–2341.
2. Ainsworth, M. A posteriori error estimation for discontinuous Galerkin finite element approximation. *SIAM J. Numer. Anal. 45*, 4 (2007), 1777–1798.
3. Ainsworth, M., and Rankin, R. Fully computable error bounds for discontinuous Galerkin finite element approximations on meshes with an arbitrary number of levels of hanging nodes. Research Report 9, University of Strathclyde, 2008.
4. Bastian, P., and Rivière, B. Superconvergence and $H(\mathrm{div})$ projection for discontinuous Galerkin methods. *Internat. J. Numer. Methods Fluids 42*, 10 (2003), 1043–1057.
5. Bebendorf, M. A note on the Poincaré inequality for convex domains. *Z. Anal. Anwendungen 22*, 4 (2003), 751–756.
6. Becker, R., Hansbo, P., and Larson, M. G. Energy norm a posteriori error estimation for discontinuous Galerkin methods. *Comput. Methods Appl. Mech. Engrg. 192*, 5-6 (2003), 723–733.
7. Bernardi, C., and Verfürth, R. Adaptive finite element methods for elliptic equations with non-smooth coefficients. *Numer. Math. 85*, 4 (2000), 579–608.
8. Brezzi, F., and Fortin, M. *Mixed and hybrid finite element methods*, vol. 15 of *Springer Series in Computational Mathematics*. Springer-Verlag, New York, 1991.
9. Burman, E., and Ern, A. Continuous interior penalty *hp*-finite element methods for advection and advection-diffusion equations. *Math. Comp. 76*, 259 (2007), 1119–1140.
10. Carstensen, C., and Funken, S. A. Constants in Clément-interpolation error and residual based a posteriori error estimates in finite element methods. *East-West J. Numer. Math. 8*, 3 (2000), 153–175.
11. Carstensen, C., Gudi, T., and Jensen, M. A unifying theory of a posteriori error control for discontinuous Galerkin FEM. HIM preprint, Hausdorff Research Institute for Mathematics, 2008.
12. Cheddadi, I., Fučík, R., Prieto, M. I., and Vohralík, M. Guaranteed and robust a posteriori error estimates for singularly perturbed reaction–diffusion problems. Preprint R08018, Laboratoire Jacques-Louis Lions, submitted for publication, 2008.
13. Cochez-Dhondt, S. *Méthodes d'éléments finis et estimations d'erreur a posteriori*. Ph.D. thesis, Université de Valenciennes et du Hainaut-Cambrésis, 2007.
14. Cochez-Dhondt, S., and Nicaise, S. Equilibrated error estimators for discontinuous Galerkin methods. *Numer. Methods Partial Differential Equations 24*, 5 (2008), 1236–1252.
15. Destuynder, P., and Métivet, B. Explicit error bounds in a conforming finite element method. *Math. Comp. 68*, 228 (1999), 1379–1396.
16. Di Pietro, D. A., Ern, A., and Guermond, J.-L. Discontinuous Galerkin methods for anisotropic semidefinite diffusion with advection. *SIAM J. Numer. Anal. 46*, 2 (2008), 805–831.
17. Ern, A., Nicaise, S., and Vohralík, M. An accurate $\mathbf{H}(\mathrm{div})$ flux reconstruction for discontinuous Galerkin approximations of elliptic problems. *C. R. Math. Acad. Sci. Paris 345* (2007), 709–712.
18. Ern, A., and Stephansen, A. F. A posteriori energy-norm error estimates for advection-diffusion equations approximated by weighted interior penalty methods. *J. Comp. Math. 26*, 4 (2008), 488–510.
19. Ern, A., Stephansen, A. F., and Vohralík, M. Improved energy norm a posteriori error estimation based on flux reconstruction for discontinuous Galerkin methods. Research Report version 1 (04-12-2007), Université Paris 6 and Ecole des Ponts, 2007.
20. Ern, A., Stephansen, A. F., and Zunino, P. A discontinuous Galerkin method with

weighted averages for advection–diffusion equations with locally small and anisotropic diffusivity. IMA J. Numer. Anal., to appear, 2008.
21. Karakashian, O. A., and Pascal, F. A posteriori error estimates for a discontinuous Galerkin approximation of second-order elliptic problems. *SIAM J. Numer. Anal. 41*, 6 (2003), 2374–2399.
22. Kim, K. Y. A posteriori error estimators for locally conservative methods of nonlinear elliptic problems. *Appl. Numer. Math. 57* (2007), 1065–1080.
23. Ladevèze, P. *Comparaison de modèles de milieux continus*. Ph.D. thesis, Université Pierre et Marie Curie (Paris 6), 1975.
24. Ladevèze, P., and Leguillon, D. Error estimate procedure in the finite element method and applications. *SIAM J. Numer. Anal. 20*, 3 (1983), 485–509.
25. Lazarov, R., Repin, S., and Tomar, S. Functional a posteriori error estimates for discontinuous Galerkin approximations of elliptic problems. Report 2006-40, Ricam, Austria, 2006.
26. Payne, L. E., and Weinberger, H. F. An optimal Poincaré inequality for convex domains. *Arch. Rational Mech. Anal. 5* (1960), 286–292 (1960).
27. Prager, W., and Synge, J. L. Approximations in elasticity based on the concept of function space. *Quart. Appl. Math. 5* (1947), 241–269.
28. Sangalli, G. Analysis of the advection-diffusion operator using fractional order norms. *Numer. Math. 97*, 4 (2004), 779–796.
29. Sangalli, G. A uniform analysis of nonsymmetric and coercive linear operators. *SIAM J. Math. Anal. 36*, 6 (2005), 2033–2048 (electronic).
30. Sangalli, G. Robust a-posteriori estimator for advection-diffusion-reaction problems. *Math. Comp. 77*, 261 (2008), 41–70 (electronic).
31. Schötzau, D., and Zhu, L. A robust a-posteriori error estimator for discontinuous Galerkin methods for convection-diffusion equations. Appl. Numer. Math, to appear, 2008.
32. Stephansen, A. F. *Méthodes de Galerkine discontinues et analyse d'erreur a posteriori pour les problèmes de diffusion hétérogène*. Ph.D. thesis, Ecole Nationale des Ponts et Chaussées, 2007.
33. Verfürth, R. A posteriori error estimators for convection–diffusion equations. *Numer. Math. 80*, 4 (1998), 641–663.
34. Verfürth, R. Robust a posteriori error estimates for stationary convection-diffusion equations. *SIAM J. Numer. Anal. 43*, 4 (2005), 1766–1782.
35. Verfürth, R. A note on constant-free a posteriori error estimates. Tech. report, Ruhr-Universität Bochum, 2008.
36. Vohralík, M. On the discrete Poincaré–Friedrichs inequalities for nonconforming approximations of the Sobolev space $H^1$. *Numer. Funct. Anal. Optim. 26*, 7–8 (2005), 925–952.
37. Vohralík, M. A posteriori error estimates for lowest-order mixed finite element discretizations of convection-diffusion-reaction equations. *SIAM J. Numer. Anal. 45*, 4 (2007), 1570–1599.
38. Vohralík, M. Residual flux-based a posteriori error estimates for finite volume and related locally conservative methods. *Numer. Math., DOI 10.1007/s00211-008-0168-4 (electronic)* (2008).