



**HAL**  
open science

# Residual flux-based a posteriori error estimates for finite volume and related locally conservative methods

Martin Vohralík

► **To cite this version:**

Martin Vohralík. Residual flux-based a posteriori error estimates for finite volume and related locally conservative methods. *Numerische Mathematik*, 2008, 111 (1), pp.121-158. 10.1007/s00211-008-0168-4. hal-00079221v2

**HAL Id: hal-00079221**

**<https://hal.science/hal-00079221v2>**

Submitted on 10 Jul 2008

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

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

Martin Vohralík

# Residual flux-based a posteriori error estimates for finite volume and related locally conservative methods

Received: date / Revised: date

**Abstract** We derive in this paper a posteriori error estimates for discretizations of convection–diffusion–reaction equations in two or three space dimensions. Our estimates are valid for any cell-centered finite volume scheme, and, in a larger sense, for any locally conservative method such as the mimetic finite difference, covolume, and other. We consider meshes consisting of simplices or rectangular parallelepipeds and also provide extensions to nonconvex cells and nonmatching interfaces. We allow for the cases of inhomogeneous and anisotropic diffusion–dispersion tensors and of convection dominance. The estimates are established in the energy (semi-)norm for a locally postprocessed approximate solution preserving the conservative fluxes and are of residual type. They are fully computable (all occurring constants are evaluated explicitly) and locally efficient (give a local lower bound on the error times an efficiency constant), so that they can serve both as indicators for adaptive refinement and for the actual control of the error. They are semi-robust in the sense that the local efficiency constant only depends on local variations in the coefficients and becomes optimal as the local Péclet number gets sufficiently small. Numerical experiments confirm their accuracy.

**Keywords** a posteriori error estimates · convection–diffusion–reaction equation · inhomogeneous and anisotropic diffusion · convection dominance · locally conservative methods · finite volume method · mimetic finite difference method · covolume method · nonmatching grids

**Mathematics Subject Classification (2000)** 65N15, 76M12, 76S05

---

This work was supported by the GdR MoMaS project “Numerical Simulations and Mathematical Modeling of Underground Nuclear Waste Disposal”, PACEN/CNRS, ANDRA, BRGM, CEA, EdF, IRSN, France.

The main part of this work was carried out during the author’s post-doc stay at Laboratoire de Mathématiques, Analyse Numérique et EDP, Université de Paris-Sud and CNRS, Orsay, France

---

M. Vohralík  
Laboratoire Jacques-Louis Lions, Université Pierre et Marie Curie (Paris 6), B.C. 187, 4 place Jussieu, 75252 Paris, France.  
Tel.: +33-1-44 27 72 01  
Fax: +33-1-44 27 72 00  
E-mail: vohralik@ann.jussieu.fr

## 1 Introduction

We consider in this paper the convection–diffusion–reaction problem

$$-\nabla \cdot (\mathbf{S}\nabla p) + \nabla \cdot (p\mathbf{w}) + rp = f \quad \text{in } \Omega, \quad (1.1a)$$

$$p = g \quad \text{on } \Gamma_D, \quad (1.1b)$$

$$-\mathbf{S}\nabla p \cdot \mathbf{n} = u \quad \text{on } \Gamma_N, \quad (1.1c)$$

where  $\mathbf{S}$  is a diffusion–dispersion tensor (which can be inhomogeneous and anisotropic),  $\mathbf{w}$  is a possibly dominating velocity field,  $r$  is a reaction function,  $f$  is a source term, and  $g$  and  $u$  prescribe the Dirichlet and Neumann boundary conditions, respectively. We suppose that  $\Omega \subset \mathbb{R}^d$ ,  $d = 2, 3$ , is a polygonal (meaning alternatively also polyhedral) domain (open, bounded, and connected set), that  $\Gamma_D \cap \Gamma_N = \emptyset$ ,  $\Gamma_D \cup \Gamma_N = \Gamma := \partial\Omega$ , that  $|\Gamma_D| \neq 0$ , where  $|\Gamma_D|$  is the measure of the set  $\Gamma_D$ , and that  $\Gamma_{\text{in}} := \{\mathbf{x} \in \partial\Omega; \mathbf{w} \cdot \mathbf{n} < 0\} \subset \Gamma_D$ . Finally,  $\mathbf{n}$  stands for the unit normal vector of  $\partial\Omega$ , outward to  $\Omega$ . Our purpose is to derive a posteriori error estimates for finite volume, and, in a larger sense, any locally conservative discretizations of the problem (1.1a)–(1.1c).

Residual a posteriori error estimates are nowadays well established for discretizations of the pure diffusion form of (1.1a)–(1.1c) (i.e.,  $\mathbf{w} = r = 0$ ) by the finite element method, cf. for example the survey by Verfürth [45]. In most cases, however, the analysis is only given for  $\mathbf{S}$  being an identity matrix; an in-depth analysis for the general inhomogeneous (and anisotropic) diffusion tensor was presented by Bernardi and Verfürth [14]. In recent years a posteriori error estimates have been extended to convection–diffusion problems as well. We cite in particular Verfürth [46], who derived estimates in the energy norm for the conforming Galerkin method and its stabilized SUPG version. His estimates are both reliable and locally efficient and, moreover, the efficiency constant becomes optimal as the local Péclet number gets sufficiently small. Similar results have been obtained in the framework of nonconforming finite element methods by Ainsworth [7] for the inhomogeneous pure diffusion case and by El Alaoui et al. in [26] for the convection–diffusion case. Recently, Verfürth [47] improved his results while giving estimates which are fully robust with respect to convection dominance in a norm incorporating a dual norm of the convective derivative. The new norm is, however, not computable, there is no local lower bound, and the estimators do not change with respect to [46] and hence the adaptive strategies will remain the same.

The theory of a posteriori error estimation is much less developed for finite volume methods. For vertex-centered schemes, the analogy with the finite element case is usually exploited in order to obtain a posteriori error estimates—this is, e.g., the case of the works of Afif et al. [4], Bergam et al. [13] or Lazarov and Tomov [35] (cf. also [18]). Still less work has been done for cell-centered schemes. Agouzal and Oudin [6] simply note that one can exploit the relations between the lowest-order Raviart–Thomas mixed finite element, the lowest-order nonconforming finite element, and the cell-centered finite volume methods on triangular meshes in order to obtain an error indicator under a saturation-like hypothesis. Rigorous a posteriori error estimates are obtained by Achdou et al. [2], however, only for two particular schemes. Equivalence of the discrete forms of the schemes with some finite element ones is used for this purpose. Nicaise [36, 37] gives a posteriori estimates for Morley-type interpolants of the original piecewise

constant finite volume approximation. Finally, yet a different approach, yielding an estimate in the  $L^1$ -norm, independent of the size of the diffusion tensor, is given by Ohlberger [38,39], whereas one of the first results on a posteriori error estimation in finite volume methods were obtained by Angermann [8] and Angermann et al. [9]. In the two years refereeing period of this paper, a few new references appeared. Kim [34] presents estimates applicable to any locally conservative method, as we do it here. Both approximation to  $-\mathbf{S}\nabla p$  (which we shall term here as “velocity”) and to  $p$  (which we shall term here as “potential”) have to be specified first. Then bounds up to an undetermined constant are given for this couple in a mesh-dependent norm, which contains a weighted jump term for the potential. Beirão da Veiga [44] gives a posteriori error estimates for the mimetic finite difference method, more precisely for the known velocity and a piecewise linear postprocessing of the originally piecewise constant potential. Again a mesh-dependent norm is used and the estimator features an unknown generic constant.

The purpose of this paper is to develop a sufficiently general and unified framework for a posteriori error estimation in the finite volume method, and, in a larger sense, in any locally conservative method. The derived estimates are first of all independent of particular schemes. Any cell-centered finite volume scheme, cf. Eymard et al. [27], the schemes proposed in [22,29,30,32], “multi-point flux-approximation” schemes [1,25], “discrete duality finite volume” schemes [23], and “mixed finite volume” schemes [24] can be considered. Similarly, the estimates are valid for the mimetic finite difference, cf. [16], covolume, cf. [21], and other locally conservative methods. The only requirement we have is that the scheme satisfies a conservation equation over each computational cell, and prescribes in particular the discrete fluxes. We focus on bounds for the error measured in the natural energy norm for the (postprocessed) potential only. Bounds for the error in the approximate velocity only are established in a general setting covering all the methods considered here in [54, Section 6]. Similarly, bounds for the error in the couple of velocity and potential approximations are stated therein. Next, our estimates hold and have the same form from pure diffusion cases to the full convection–diffusion–reaction ones. As an important point, we derive them for very general meshes containing nonconvex cells and nonmatching interfaces. For the sake of completeness of the analysis, we also consider general inhomogeneous Dirichlet or Neumann boundary conditions. The derived estimators are associated with the mesh cells, are locally and easily computable, and they are also fully computable in the sense that all occurring constants are evaluated explicitly. This in particular means that they can serve not only as indicators for adaptive refinement, which is the usual practice, but also for the actual control of the error. In the developed independent theory, no equivalence with a different (finite element or mixed finite element) case is necessary. Also, no convexity of the domain, no additional regularity of the weak solution, and no saturation assumption are needed.

The basis of our approach is to exploit the particular feature of the considered schemes, which is the conservativity of the discrete fluxes across the sides (edges if  $d = 2$ , faces if  $d = 3$ ) of the mesh. Inspired by the results of Eymard et al. [28] and of the author [51], we first build a postprocessed approximate potential  $\tilde{p}_h$  which preserves exactly the given discrete diffusive fluxes and whose mean or point value is in each cell fixed by the original constant approximation. The interest of such a postprocessing is twofold. First of all, we obtain an approximate potential suit-

able for energy error measuring (recall that the piecewise gradient of the original cellwise constant approximation is zero, so that it gives no sense to measure the energy error in it). Secondly, by such a construction,  $-\mathbf{S}\nabla\tilde{p}_h$  lies in the  $\mathbf{H}(\operatorname{div}, \Omega)$  space. In the finite volume case, moreover, using the fixation of  $\tilde{p}_h$  by the original cell values, we are able to prove the convergence of  $\tilde{p}_h$  and to give a priori error estimates, under the condition that the original scheme satisfies some necessary properties. The construction of the postprocessed potential is particularly easy for simplices (triangles if  $d = 2$  and tetrahedra if  $d = 3$ ) or, when  $\mathbf{S}$  is diagonal, for rectangular parallelepipeds (rectangles if  $d = 2$ , rectangular parallelepipeds if  $d = 3$ ); in this case  $\tilde{p}_h$  is a second-order polynomial whose gradient (times minus  $\mathbf{S}$ ) is constructed as in the lowest-order Raviart–Thomas mixed finite element method, cf. [15, 42]. The crucial advantage  $-\mathbf{S}\nabla\tilde{p}_h \in \mathbf{H}(\operatorname{div}, \Omega)$  is, however, compensated by the fact that  $\tilde{p}_h$  is nonconforming in the sense that  $\tilde{p}_h \notin H^1(\Omega)$ . We describe all the above results in detail in Section 3, after collecting some preliminary remarks, notation, assumptions, and details on the continuous problem in Section 2. In this section, we also recall the Oswald interpolation operator and describe its generalization to arbitrary grids and boundary conditions.

Section 4 is then devoted to our a posteriori error estimates. They consist of several independent estimators, the principal of which penalize the fact that  $\tilde{p}_h$  is nonconforming and that its residual is nonzero. For pure diffusion problems, only these estimators (plus possibly still a Neumann boundary one) are present. When there is some convection, additional convection and upwinding estimators appear, and for cases with reaction, a reaction quadrature estimator may be present as well. We next prove that the principal (nonconformity, convection, and residual) estimators represent local lower bounds for the error as well, where in particular the efficiency constants are of the form  $c_1 + c_2 \min\{\operatorname{Pe}, \rho\}$ , where  $\operatorname{Pe}$  (the local Péclet number) and  $\rho$  are given below by (4.12) and where  $c_1, c_2$  only depend on local variations in  $\mathbf{S}$  (i.e., on local inhomogeneities and anisotropies), on local variations in  $\mathbf{w}$  and  $r$ , on the space dimension, on the polynomial degree of  $\tilde{p}_h, f, \mathbf{w}, r$ , and on the shape-regularity parameter of the mesh. These estimators are thus in particular optimally efficient as the local Péclet number gets sufficiently small. We are not able to obtain similar results for the upwinding estimator but numerical experiments suggest that this estimator represents a higher-order term as soon as the local Péclet number gets sufficiently small. A more detailed discussion, as well as several other remarks, is given in Section 4.3.

We finally in Section 5 discuss a particular approach to meshes consisting of arbitrary polygons (we use this notation also for polyhedrons) which may be non-matching and consist of elements which are not necessarily convex. It consists in subdividing each mesh element into a simplicial submesh and then directly using the theory established previously. Finally, we illustrate the accuracy of the derived estimates in Section 6 on several numerical experiments and conclude by some technical lemmas in Section 7. This paper is a detailed description, completion, and extension of the results previously announced in [50] and [53]. Robust a posteriori error estimates for vertex-centered-like finite volume methods are then presented in [52] and [19]. For a complementary approach to a posteriori error estimation in locally conservative methods, evaluating the error in the velocity only (and alternatively in the velocity and the potential), we refer to [54].

## 2 Preliminaries

We introduce here the notation, define admissible partitions of the domain  $\Omega$ , recall some inequalities which will be important in the sequel, state the assumptions on the data, give details on the continuous problem (1.1a)–(1.1c), and finally recall and give some extensions on the Oswald interpolation operator. In what follows, we conceptually denote by  $C_A, c_A$  constants basically dependent on a quantity  $A$  but always independent of the discretization parameter  $h$ .

### 2.1 Function spaces

For a domain  $S \subset \mathbb{R}^d$ , we denote by  $L^l(S)$  and  $\mathbf{L}^l(S) = [L^l(S)]^d$  the Lebesgue spaces, by  $(\cdot, \cdot)_S$  the  $L^2(S)$  or  $\mathbf{L}^2(S)$  inner product, and by  $\|\cdot\|_S$  the associated norm. We mean by  $|S|$  the Lebesgue measure of  $S$ , by  $|\sigma|$  the  $(d-1)$ -dimensional Lebesgue measure of  $\sigma \subset \mathbb{R}^{d-1}$ , and in particular by  $|\mathbf{s}|$  the length of a segment  $\mathbf{s}$ . At the same time,  $|A|$  is the cardinality of a set  $A$ . Next,  $H^l(S)$  is the Sobolev space of functions with square-integrable weak derivatives up to order  $l$ ;  $H_{\mathbb{D}}^1(S)$  is a subspace of  $H^1(S)$  of functions with traces vanishing on  $\Gamma_{\mathbb{D}}$  and  $H^{1/2}(\Gamma_{\mathbb{D}})$  is the trace space on  $\Gamma_{\mathbb{D}}$ . Next,  $\mathbf{H}(\text{div}, S)$  is the space of functions with square-integrable weak divergences,  $\mathbf{H}(\text{div}, S) = \{\mathbf{v} \in \mathbf{L}^2(S); \nabla \cdot \mathbf{v} \in L^2(S)\}$ , and  $\langle \cdot, \cdot \rangle_{\partial S}$  stands for  $(d-1)$ -dimensional inner product on  $\partial S$ . Finally, for a partition  $\mathcal{T}_h$  of  $\Omega$  as defined below, we will use the ‘‘broken Sobolev space’’  $H^l(\mathcal{T}_h)$ ,  $H^l(\mathcal{T}_h) := \{\varphi \in L^2(\Omega); \varphi|_K \in H^l(K) \forall K \in \mathcal{T}_h\}$ . We refer to Adams [3] for more details.

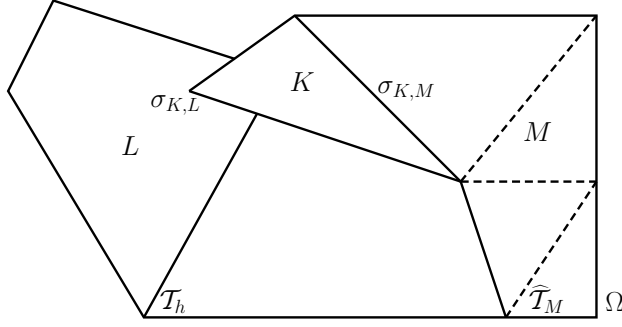
### 2.2 Partitions of the domain

We consider in this paper partitions  $\mathcal{T}_h$  of  $\Omega$  into closed polygons such that  $\overline{\Omega} = \bigcup_{K \in \mathcal{T}_h} K$  and such that the intersections of their interiors are pairwise empty. For  $K \in \mathcal{T}_h$ , we denote by  $\mathcal{E}_K$  the set of such subsets  $\sigma$  of  $\partial K$  that there exists  $L \in \mathcal{T}_h$ ,  $L \neq K$ , such that  $\sigma = \sigma_{K,L} := \partial K \cap \partial L$  has a positive  $(d-1)$ -dimensional Lebesgue measure. We call  $\sigma$  a side and  $K$  and  $L$  neighbors. We note that for matching meshes (containing no hanging nodes) consisting of convex elements,  $\sigma$  are simply the sides in the geometrical sense, which is not necessarily the case on general polygonal meshes, see Figure 2.1. The set  $\mathcal{E}_K$  is completed by the (geometrical) sides of  $K$  lying at the boundary; we suppose in this case that each  $\sigma$  lies entirely either in  $\overline{\Gamma_{\mathbb{D}}}$  or in  $\overline{\Gamma_{\mathbb{N}}}$ . We next denote by  $\mathcal{E}_h$  the set of all, by  $\mathcal{E}_h^{\text{int}}$  the set of interior, and by  $\mathcal{E}_h^{\text{ext}}$  the set of exterior sides of  $\mathcal{T}_h$ . Also, let  $\mathcal{E}_h^{\mathbb{N}}$  stand for the sides contained in  $\overline{\Gamma_{\mathbb{N}}}$ ,  $\mathcal{E}_h^{\mathbb{D}}$  for those contained in  $\overline{\Gamma_{\mathbb{D}}}$ ,  $\mathcal{E}_K$  for  $\sigma \in \mathcal{E}_h^{\text{int}}$  which share at least a vertex with  $K \in \mathcal{T}_h$ , and  $\mathfrak{T}_K$  for all elements sharing at least a vertex with  $K \in \mathcal{T}_h$ . A family of meshes  $\{\mathcal{T}_h\}_h$  is parameterized by  $h := \max_{K \in \mathcal{T}_h} h_K$ , where  $h_K$  stands for the diameter of  $K$ ; we also denote by  $h_\sigma$  the diameter of  $\sigma \in \mathcal{E}_h$ .

For matching meshes consisting of simplices or rectangular parallelepipeds, we will in some parts of the paper need the following assumption ( $\kappa_K := |K|/h_K^d$ ):

*Assumption (A) (Shape regularity of matching meshes consisting of simplices or rectangular parallelepipeds)*

*There exists a constant  $\kappa_{\mathcal{T}} > 0$  such that  $\min_{K \in \mathcal{T}_h} \kappa_K \geq \kappa_{\mathcal{T}}$  for all  $h > 0$ .*



**Fig. 2.1** Considered meshes and notation

Let  $\rho_K$  denote the diameter of the largest ball inscribed in  $K$ . Then Assumption (A) is equivalent to the usual requirement of the existence of a constant  $\theta_{\mathcal{T}} > 0$  such that  $\max_{K \in \mathcal{T}_h} h_K / \rho_K \leq \theta_{\mathcal{T}}$  for all  $h > 0$ .

In the general case, we make the following shape regularity assumption:

*Assumption (B) (Shape regularity of arbitrary polygonal meshes)*

- (B1) There exists a constant  $\kappa_{\mathcal{T}} > 0$  such that  $\min_{K \in \mathcal{T}_h} \kappa_K \geq \kappa_{\mathcal{T}}$  for all  $h > 0$ ;
- (B2) there exists a constant  $\xi_{\mathcal{T}} > 0$  such that  $\max_{K \in \mathcal{T}_h} |\mathcal{E}_K| \leq \xi_{\mathcal{T}}$  for all  $h > 0$ ;
- (B3) there exist matching refinements  $\widehat{\mathcal{T}}_h$  of  $\mathcal{T}_h$  which consist of simplices or rectangular parallelepipeds and satisfy Assumption (A).

By Assumption (B), the elements of  $\mathcal{T}_h$  are not “flat”, the number of their sides is limited, and each  $\mathcal{T}_h$  admits a matching refinement  $\widehat{\mathcal{T}}_h$ . In the sequel, as for the mesh  $\mathcal{T}_h$ , we will use the notation  $\widehat{\mathcal{E}}_h^{\text{int}}$ ,  $\widehat{\mathcal{E}}_h^{\text{ext}}$ ,  $\widehat{\mathcal{E}}_K$ ,  $\widehat{\mathcal{E}}_K$ , and  $\widehat{\mathcal{T}}_K$  also for  $\widehat{\mathcal{T}}_h$ . We also denote by  $\widehat{\mathcal{T}}_K$  the partition of  $K \in \mathcal{T}_h$  by the elements of  $\widehat{\mathcal{T}}_h$ .

### 2.3 Poincaré, Friedrichs, and trace inequalities

Let  $K$  be a convex polygon,  $\sigma$  its (geometrical) side, and  $\varphi \in H^1(K)$ . Three inequalities will play an essential role in the derivation of our a posteriori error estimates.

First, the Poincaré inequality states that

$$\|\varphi - \varphi_K\|_K^2 \leq C_{P,K} h_K^2 \|\nabla \varphi\|_K^2, \quad (2.1)$$

where  $\varphi_K$  is the mean of  $\varphi$  over  $K$  given by  $\varphi_K := (\varphi, 1)_K / |K|$  and where the constant  $C_{P,K}$  can for each convex  $K$  be evaluated as  $1/\pi^2$ , cf. [40, 12].

Second, the generalized Friedrichs inequality states that

$$\|\varphi - \varphi_\sigma\|_K^2 \leq C_{F,K,\sigma} h_K^2 \|\nabla \varphi\|_K^2, \quad (2.2)$$

where  $\varphi_\sigma$  is the mean of  $\varphi$  over  $\sigma \in \mathcal{E}_K$  given by  $\varphi_\sigma := \langle \varphi, 1 \rangle_\sigma / |\sigma|$ . It follows from [48, Lemma 4.1] that  $C_{F,K,\sigma} = 3d$  for a simplex  $K$  and its side  $\sigma$ . For a

general convex  $K$ ,  $C_{F,K,\sigma}$  can be evaluated from the proof of [27, Lemma 9.4] and [28, Lemma 3] in the form  $Ch_K^{d-1}/|\sigma|$  with  $C$  only depending on  $d$  and  $\kappa_K$ .

Finally, the trace inequality states that

$$\|\varphi - \varphi_\sigma\|_\sigma^2 \leq C_{t,K,\sigma} h_K \|\nabla \varphi\|_K^2, \quad (2.3a)$$

$$\|\varphi - \varphi_K\|_\sigma^2 \leq \tilde{C}_{t,K,\sigma} h_K \|\nabla \varphi\|_K^2, \quad (2.3b)$$

cf. the proof of [27, Lemma 9.4] and [28, Lemma 1]. It has in particular been shown in [36, Lemma 3.5] that  $C_{t,K,\sigma} = C_{t,d} h_K / h_\sigma$  for a simplex, where  $C_{t,d} \approx 1.55416$  if  $d = 2$  and  $C_{t,d} \approx 11.53557$  if  $d = 3$ . Similarly, it follows from the proof of [48, Lemma 4.1] that  $\tilde{C}_{t,K,\sigma} = 3dh_K |\sigma| / |K|$  for a simplex. Finally,  $\tilde{C}_{t,K,\sigma}$  only depends on  $d$  and  $\kappa_K$  for a general convex  $K$  and its side  $\sigma$ .

The above inequalities are also valid for nonconvex polygons  $K$  or when  $\sigma$  is only a general part of  $\partial K$  such that  $|\sigma| > 0$ , cf. [27, Lemma 10.4]. The dependencies on  $K$  and  $\sigma$  are in this case, however, more complicated in general.

## 2.4 Assumptions on the data

We suppose that there exists a basic partition  $\tilde{\mathcal{T}}_h$  of  $\Omega$  such that the data of the problem (1.1a)–(1.1c) are related to  $\tilde{\mathcal{T}}_h$  in the following way:

*Assumption (C) (Data)*

- (C1)  $\mathbf{S}_K := \mathbf{S}|_K$  is a constant, symmetric, bounded, and uniformly positive definite tensor such that  $c_{\mathbf{S},K} \mathbf{v} \cdot \mathbf{v} \leq \mathbf{S}_K \mathbf{v} \cdot \mathbf{v} \leq C_{\mathbf{S},K} \mathbf{v} \cdot \mathbf{v}$ ,  $c_{\mathbf{S},K} > 0$ ,  $C_{\mathbf{S},K} > 0$ , for all  $\mathbf{v} \in \mathbb{R}^d$  and all  $K \in \tilde{\mathcal{T}}_h$ ;
- (C2)  $\mathbf{w} \in \mathbf{H}(\text{div}, \Omega)$  and each component of  $\mathbf{w}$  is a polynomial of degree at most  $k$  on each  $K \in \tilde{\mathcal{T}}_h$  such that  $|\mathbf{w}|_K| \leq C_{\mathbf{w},K}$ ,  $C_{\mathbf{w},K} \geq 0$ , for all  $K \in \tilde{\mathcal{T}}_h$ ;
- (C3)  $r$  is a polynomial of degree at most  $k$  on each  $K \in \tilde{\mathcal{T}}_h$ ;
- (C4)  $\frac{1}{2} \nabla \cdot \mathbf{w}|_K + r|_K \geq c_{\mathbf{w},r,K}$  and  $|\nabla \cdot \mathbf{w}|_K + r|_K| \leq C_{\mathbf{w},r,K}$ ,  $c_{\mathbf{w},r,K} \geq 0$ ,  $C_{\mathbf{w},r,K} \geq 0$ , for all  $K \in \tilde{\mathcal{T}}_h$ ;
- (C5)  $f$  is a polynomial of degree at most  $k$  on each  $K \in \tilde{\mathcal{T}}_h$ ;
- (C6)  $g \in H^{1/2}(\Gamma_D)$ ;
- (C7)  $u \in L^2(\Gamma_N)$ ;
- (C8) if  $c_{\mathbf{w},r,K} = 0$ , then  $C_{\mathbf{w},r,K} = 0$ .

The assumptions that  $\mathbf{S}$  is piecewise constant and that  $\mathbf{w}$ ,  $r$ , and  $f$  are piecewise polynomial are made only for the sake of simplicity and are usually satisfied in practice. They are in fact only needed for the local efficiency proofs of our a posteriori error estimates. If the functions at hand do not fulfill these requirements, interpolation can be used. Also, note that Assumption (C8) allows  $c_{\mathbf{w},r,K} = 0$  but  $\mathbf{w}|_K \neq 0$ .



## 2.5 Continuous problem

Let  $\mathcal{T}_h$  be, as throughout the whole paper, a refinement of  $\widetilde{\mathcal{T}}_h$ . We define a bilinear form  $\mathcal{B}$  by

$$\mathcal{B}(p, \varphi) := \sum_{K \in \mathcal{T}_h} \{ (\mathbf{S}\nabla p, \nabla \varphi)_K + (\nabla \cdot (p\mathbf{w}), \varphi)_K + (rp, \varphi)_K \} \quad p, \varphi \in H^1(\mathcal{T}_h) \quad (2.4)$$

and the corresponding energy (semi-)norm by

$$\begin{aligned} \|\varphi\|_{\Omega}^2 &:= \sum_{K \in \mathcal{T}_h} \|\varphi\|_K^2, \\ \|\varphi\|_K^2 &:= \|\mathbf{S}^{1/2}\nabla\varphi\|_K^2 + \|(\frac{1}{2}\nabla \cdot \mathbf{w} + r)^{1/2}\varphi\|_K^2 \quad \varphi \in H^1(\mathcal{T}_h). \end{aligned} \quad (2.5)$$

In this way  $\mathcal{B}(\cdot, \cdot)$  and  $\|\cdot\|_{\Omega}$  are well-defined for  $p, \varphi \in H^1(\Omega)$  as well as for  $p, \varphi$  that are only piecewise regular. We remark that  $\|\cdot\|_{\Omega}$  is always a norm on  $H_D^1(\Omega)$ , whereas it is a norm on  $H^1(\mathcal{T}_h)$  only when  $c_{\mathbf{w},r,K} > 0$  for all  $K \in \mathcal{T}_h$ .

The weak formulation of the problem (1.1a)–(1.1c) is then to find  $p \in H^1(\Omega)$  with  $p|_{\Gamma_D} = g$  in the sense of traces such that

$$\mathcal{B}(p, \varphi) = (f, \varphi)_{\Omega} - \langle u, \varphi \rangle_{\Gamma_N} \quad \forall \varphi \in H_D^1(\Omega). \quad (2.6)$$

Assumptions (C1)–(C7), the Green theorem, and the Cauchy–Schwarz inequality imply that

$$\mathcal{B}(\varphi, \varphi) \geq \|\varphi\|_{\Omega}^2 \quad \forall \varphi \in H_D^1(\Omega), \quad (2.7)$$

$$\mathcal{B}(\varphi, \varphi) = \|\varphi\|_{\Omega}^2 + \frac{1}{2} \sum_{K \in \mathcal{T}_h} \langle \varphi^2, \mathbf{w} \cdot \mathbf{n} \rangle_{\partial K} \quad \forall \varphi \in H^1(\mathcal{T}_h), \quad (2.8)$$

$$\begin{aligned} \mathcal{B}(p, \varphi) &\leq \max \left\{ 1, \max_{K \in \mathcal{T}_h} \left\{ \frac{C_{\mathbf{w},r,K}}{c_{\mathbf{w},r,K}} \right\} \right\} \|p\|_{\Omega} \|\varphi\|_{\Omega} \\ &\quad + \max_{K \in \mathcal{T}_h} \left\{ \frac{C_{\mathbf{w},K}}{\sqrt{c_{\mathbf{S},K}}} \right\} \|p\|_{\Omega} \|\varphi\|_{\Omega} \quad \forall p, \varphi \in H^1(\mathcal{T}_h) \end{aligned} \quad (2.9)$$

and problem (2.6) under Assumption (C) in particular admits a unique solution.

*Remark 2.1 (Notation)* In estimate (2.9), if  $c_{\mathbf{w},r,K} = 0$ , then the term  $C_{\mathbf{w},r,K}/c_{\mathbf{w},r,K}$  should be evaluated as zero; since Assumption (C8) for this case gives  $C_{\mathbf{w},r,K} = 0$ , the term with  $C_{\mathbf{w},r,K}$  in fact does not even enter the estimate. To simplify notation, we will systematically use the convention  $0/0 = 0$  throughout the text.

*Remark 2.2 (Different partitions)* We clarify here the relations between the different partitions of  $\Omega$  considered. First of all, the partition  $\widetilde{\mathcal{T}}_h$  is only used so as to describe the different parts of  $\Omega$  with different data (materials). Next,  $\mathcal{T}_h$  is the “computational” partition and we suppose that  $\mathcal{T}_h$  is a refinement of  $\widetilde{\mathcal{T}}_h$ . Finally, for general (nonmatching) polygonal meshes, we shall sometimes need a matching refinement  $\widehat{\mathcal{T}}_h$  of  $\mathcal{T}_h$  consisting of simplices or rectangular parallelepipeds.

## 2.6 Oswald interpolation operator

The Oswald operator, previously considered, e.g., in [2, 33] and studied in detail in [33, 17], associates to a discontinuous piecewise polynomial function  $\tilde{p}_h$  its conforming (continuous, contained in  $H^1(\Omega)$ ) interpolant. We briefly recall it here and present its extension to arbitrary grids and boundary conditions.

Let  $\mathcal{T}_h$  be matching, consisting of simplices or rectangular parallelepipeds, and satisfying Assumption (A). If  $\mathcal{T}_h$  consists of simplices, let  $\mathbb{R}_l(\mathcal{T}_h) := \mathbb{P}_l(\mathcal{T}_h)$  denote the space of polynomials of total degree at most  $l$  on each simplex (without any continuity requirement at the sides). Similarly, if  $\mathcal{T}_h$  consists of rectangular parallelepipeds, let  $\mathbb{R}_l(\mathcal{T}_h) := \mathbb{Q}_l(\mathcal{T}_h)$  denote the space of piecewise polynomials of degree at most  $l$  in each variable. The Oswald interpolation operator  $\mathcal{I}_{\text{Os}} : \mathbb{R}_l(\mathcal{T}_h) \rightarrow \mathbb{R}_l(\mathcal{T}_h) \cap H^1(\Omega)$  is defined as follows: given a function  $\varphi_h \in \mathbb{R}_l(\mathcal{T}_h)$ , the value of  $\mathcal{I}_{\text{Os}}(\varphi_h)$  is prescribed at the Gauss–Lobatto nodes on rectangular parallelepipeds and suitable (e.g., Lagrangian on triangles when  $l = 2$ ) nodes on simplices, see [17, Sections 3 and 5] of  $\mathbb{R}_l(\mathcal{T}_h) \cap H^1(\Omega)$  by the average of the values of  $\varphi_h$  at this node,  $\mathcal{I}_{\text{Os}}(\varphi_h)(V) = \sum_{K \in \mathcal{T}_V} \varphi_h|_K(V) / |\mathcal{T}_V|$ , where  $\mathcal{T}_V$  is the set of such  $K \in \mathcal{T}_h$  that contain the node  $V$ . Note that the interpolant is in particular equal to  $\varphi_h|_K(V)$  at a node  $V$  lying in the interior of some  $K \in \mathcal{T}_h$ . Let  $[[\varphi_h]]$  be the jump of a function  $\varphi_h$  across a side  $\sigma$ : if  $\sigma = \partial K \cap \partial L$ , then  $[[\varphi_h]]$  is the difference of the value of  $\varphi_h$  in  $K$  and  $L$  (the order of  $K$  and  $L$  has no influence on what follows) and  $[[\varphi_h]] := 0$  otherwise. The following result has been proved in [17], see Lemmas 3.2 and 5.3 and Remark 3.2 in this reference:

**Lemma 2.1 (Oswald interpolation operator)** *Let  $\mathcal{T}_h$  be matching, consisting of simplices or rectangular parallelepipeds, and satisfying Assumption (A). Let next  $\varphi_h \in \mathbb{R}_l(\mathcal{T}_h)$  and let  $\mathcal{I}_{\text{Os}}(\varphi_h)$  be constructed as described above. Then*

$$\begin{aligned} \|\varphi_h - \mathcal{I}_{\text{Os}}(\varphi_h)\|_K^2 &\leq C_1 \sum_{\sigma \in \mathcal{E}_K} h_\sigma \|[[\varphi_h]]\|_\sigma^2, \\ \|\nabla(\varphi_h - \mathcal{I}_{\text{Os}}(\varphi_h))\|_K^2 &\leq C_2 \sum_{\sigma \in \mathcal{E}_K} h_\sigma^{-1} \|[[\varphi_h]]\|_\sigma^2 \end{aligned}$$

for all  $K \in \mathcal{T}_h$ , where the constants  $C_1, C_2$  only depend on the space dimension  $d$ , on the maximal polynomial degree  $l$ , and on the shape regularity parameter  $\kappa_{\mathcal{T}}$ .

If  $\mathcal{T}_h$  is nonmatching or consists of arbitrary polygons, let  $\widehat{\mathcal{T}}_h$  be its conforming submesh of Assumption (B3). Then each  $\tilde{p}_h \in \mathbb{R}_l(\mathcal{T}_h)$  also belongs to  $\mathbb{R}_l(\widehat{\mathcal{T}}_h)$ . Hence we can define  $\mathcal{I}_{\text{Os}}(\tilde{p}_h)$  on  $\widehat{\mathcal{T}}_h$  as in the previous paragraph. We finally define  $\mathcal{I}_{\text{Os}}^\Gamma(\tilde{p}_h) \in H^1(\Omega)$ , differing from  $\mathcal{I}_{\text{Os}}(\tilde{p}_h)$  only on such  $K \in \mathcal{T}_h$  that contain a boundary side, by

$$\mathcal{I}_{\text{Os}}^\Gamma(\tilde{p}_h)|_{\Gamma_D} = g \quad \text{in the sense of traces,} \quad (2.10a)$$

$$\langle \mathcal{I}_{\text{Os}}^\Gamma(\tilde{p}_h) \mathbf{w} \cdot \mathbf{n}, 1 \rangle_\sigma = W_{K,\sigma} \quad \forall \sigma \in \mathcal{E}_h^N. \quad (2.10b)$$

Here  $W_{K,\sigma}$  are the given scheme convective fluxes, cf. Sections 3 and 4 below. In particular, for the finite volume scheme (3.1)–(3.12) below, if  $\mathbf{w} \cdot \mathbf{n}$  is constant but nonzero on a given  $\sigma \in \mathcal{E}_h^N$ , then (2.10b) prescribes the mean value of  $\mathcal{I}_{\text{Os}}^\Gamma(\tilde{p}_h)$  on this side by  $W_{K,\sigma} / \langle \mathbf{w} \cdot \mathbf{n}, 1 \rangle_\sigma$ , i.e., by  $p_\sigma$ .

### 3 Cell-centered finite volume schemes and postprocessing

This section is devoted to the example of the cell-centered finite volume method for the problem (1.1a)–(1.1c). We also demonstrate here the construction of the locally postprocessed approximation  $\tilde{p}_h$  which will serve later as the basis for our a posteriori error estimates.

#### 3.1 The cell-centered finite volume method

A general cell-centered finite volume scheme for the problem (1.1a)–(1.1c) can be written in the following form: find  $p_K, K \in \mathcal{T}_h$ , the approximations to  $p$  such that

$$\sum_{\sigma \in \mathcal{E}_K} S_{K,\sigma} + \sum_{\sigma \in \mathcal{E}_K} W_{K,\sigma} + r_K p_K |K| = f_K |K| \quad \forall K \in \mathcal{T}_h, \quad (3.1)$$

where  $f_K := (f, 1)/|K|$ ,  $r_K := (r, 1)/|K|$ , and where  $S_{K,\sigma}$  and  $W_{K,\sigma}$  are, respectively, the diffusive and convective fluxes through the sides  $\sigma$  of an element  $K$ , functions of  $p_K, K \in \mathcal{T}_h$ , of the mesh, and of the data. For the a posteriori error estimates presented in this paper, we do not need the specific form of the diffusive and convective fluxes; our analysis, however, relies on the finite volume concept of their continuity, imposing that  $S_{K,\sigma_{K,L}} = -S_{L,\sigma_{K,L}}$  and  $W_{K,\sigma_{K,L}} = -W_{L,\sigma_{K,L}}$  for all  $\sigma_{K,L} \in \mathcal{E}_h^{\text{int}}$ . Note also that for nonmatching grids,  $\sigma$  is not necessarily a side of  $K$  in the geometrical sense, see Section 2.2. To fix ideas, we give an example.

When  $\mathcal{T}_h$  is admissible in the sense of [27, Definition 9.1], which is satisfied, e.g., when  $\mathcal{T}_h$  is matching, consists of convex elements, and when there exist points  $\mathbf{x}_K$  in the interior of each  $K \in \mathcal{T}_h$  such that the straight lines connecting  $\mathbf{x}_K$  and  $\mathbf{x}_L$  for two neighboring elements  $K$  and  $L$  are orthogonal to  $\sigma_{K,L} = \partial K \cap \partial L$  and finally when an analogous orthogonality condition holds on the Dirichlet part of the boundary, and under the additional assumption that  $\mathbf{S}_K = s_K Id$  for all  $K \in \mathcal{T}_h$ , the following choices for  $S_{K,\sigma}$  and  $W_{K,\sigma}$  are possible:

$$S_{K,\sigma} = -s_{K,L} \frac{|\sigma_{K,L}|}{d_{K,L}} (p_L - p_K) \quad \sigma = \sigma_{K,L} \in \mathcal{E}_h^{\text{int}}, \quad (3.2)$$

$$S_{K,\sigma} = -s_K \frac{|\sigma|}{d_{K,\sigma}} (g_\sigma - p_K) \quad \sigma \in \mathcal{E}_K \cap \mathcal{E}_h^{\text{D}}, \quad (3.3)$$

$$S_{K,\sigma} = u_\sigma |\sigma| \quad \sigma \in \mathcal{E}_K \cap \mathcal{E}_h^{\text{N}}, \quad (3.4)$$

$$W_{K,\sigma} = p_\sigma w_{K,\sigma} \quad \sigma \in \mathcal{E}_K. \quad (3.5)$$

Here  $d_{K,L} = |\mathbf{x}_K - \mathbf{x}_L|$ , the Euclidean distance of  $\mathbf{x}_K$  and  $\mathbf{x}_L$ ,  $d_{K,\sigma}$  is the Euclidean distance of  $\mathbf{x}_K$  and  $\sigma \in \mathcal{E}_K \cap \mathcal{E}_h^{\text{D}}$ , and

$$g_\sigma := \frac{\langle g, 1 \rangle_\sigma}{|\sigma|}, \quad u_\sigma := \frac{\langle u, 1 \rangle_\sigma}{|\sigma|}, \quad w_{K,\sigma} := \langle \mathbf{w} \cdot \mathbf{n}, 1 \rangle_\sigma \quad (3.6)$$

are, respectively, the mean value of the Dirichlet boundary condition  $g$  on a side  $\sigma \in \mathcal{E}_h^{\text{D}}$ , the mean value of the Neumann boundary condition  $u$  on a side  $\sigma \in \mathcal{E}_h^{\text{N}}$ , and the flux of the velocity field  $\mathbf{w}$  through a side  $\sigma$  of  $K$ . For the value  $s_{K,L}$ ,

we have two choices, corresponding respectively to the arithmetic and harmonic averaging in the diffusion–dispersion tensor,

$$s_{K,L} := \frac{s_K + s_L}{2}, \quad (3.7a)$$

$$s_{K,L} := \frac{2s_K s_L}{s_K + s_L}. \quad (3.7b)$$

Finally,  $p_\sigma$  is the weighted upwind value, given for interior or Dirichlet boundary sides by

$$p_\sigma := v_\sigma \hat{p}_\sigma + (1 - v_\sigma) \bar{p}_\sigma, \quad (3.8)$$

where  $\hat{p}_\sigma$  is the upwind value and  $\bar{p}_\sigma$  the centered value, given respectively by

$$\hat{p}_\sigma := \begin{cases} p_K & \text{if } w_{K,\sigma} \geq 0 \\ p_L & \text{if } w_{K,\sigma} < 0 \end{cases}, \quad \bar{p}_\sigma := \frac{1}{2}(p_K + p_L) \quad (3.9)$$

if  $\sigma$  is an interior side between elements  $K$  and  $L$  and

$$\hat{p}_\sigma := \begin{cases} p_K & \text{if } w_{K,\sigma} \geq 0 \\ g_\sigma & \text{if } w_{K,\sigma} < 0 \end{cases} \quad \bar{p}_\sigma := \frac{1}{2}(p_K + g_\sigma) \quad (3.10)$$

if  $\sigma \in \mathcal{E}_K$  is a Dirichlet boundary side. Finally,

$$p_\sigma := p_K \quad (3.11)$$

if  $\sigma \in \mathcal{E}_K$  is a Neumann boundary side. In the above definitions,  $v_\sigma \in [0, 1]$  is a coefficient of the amount of upstream weighting. The full-upwind scheme (with  $v_\sigma = 1$  for all  $\sigma \in \mathcal{E}_h$ ) is classical, see [27, Section 9]. The introduction of the parameter  $v_\sigma$  is done in order to guarantee the stability of the scheme while simultaneously reducing the excessive numerical diffusion added by the full upstream weighting. An optimal choice for  $v_\sigma$  (cf. [20, 31]) is (for  $w_{K,\sigma} \neq 0$ , otherwise  $v_\sigma := 1$ )

$$v_\sigma := \begin{cases} 1 - \min \left\{ \frac{2s_{K,L} |\sigma_{K,L}|}{d_{K,L} |w_{K,\sigma}|}, 1 \right\} & \text{if } \sigma_{K,L} \in \mathcal{E}_h^{\text{int}}, \\ 1 - \min \left\{ \frac{2s_K |\sigma|}{d_{K,\sigma} |w_{K,\sigma}|}, 1 \right\} & \text{if } \sigma \in \mathcal{E}_h^{\text{D}} \text{ and } w_{K,\sigma} > 0, \\ 1 & \text{if } \sigma \in \mathcal{E}_h^{\text{D}} \text{ and } w_{K,\sigma} < 0. \end{cases} \quad (3.12)$$

*Remark 3.1 (Local Péclet upstream weighting)* The local Péclet upstream weighting (3.8)–(3.12) for the scheme (3.1)–(3.7b) on polygonal meshes  $\mathcal{T}_h$  admissible in the sense of [27, Definition 9.1] guarantees, when  $\mathbf{S}_K = s_K \text{Id}$  for all  $K \in \mathcal{T}_h$ ,  $\Gamma_N = \emptyset$ , and under appropriate assumptions on  $\mathbf{w}$ ,  $r$ , and  $f$ , the discrete maximum principle, while reducing the excessive numerical diffusion added by the full upstream weighting, cf. [31, Theorem 4.5] and [27, Proposition 9.2]. Moreover, the inequality

$$\sum_{K \in \mathcal{T}_h} \sum_{\sigma \in \mathcal{E}_K} p_\sigma w_{K,\sigma} p_K \geq \frac{1}{2} \sum_{K \in \mathcal{T}_h} p_K^2 (\nabla \cdot \mathbf{w}, 1)_K \quad (3.13)$$

holds, see [51, Lemma 6.5]. Another possibility for still more precise discretization of the convection term would be to define the centered value  $\bar{p}_\sigma$  by the value

that takes a linear function given by  $p_K$  in  $\mathbf{x}_K$  and by  $p_L$  in  $\mathbf{x}_L$  on  $\sigma_{K,L} \in \mathcal{E}_h^{\text{int}}$  common to  $K, L \in \mathcal{T}_h$ , i.e., to take into account the distances of  $\mathbf{x}_K$  and  $\mathbf{x}_L$  from the side  $\sigma_{K,L}$  instead of putting  $\bar{p}_\sigma := (p_K + p_L)/2$ , and similarly at the boundary. While under appropriate modification of (3.12) and under the above-mentioned assumptions, the discrete maximum principle would still hold, this would not in general be the case for the inequality (3.13). This inequality is, however, very useful already while proving the existence and uniqueness of the discrete solution when there is no discrete maximum principle (recall that this principle can in general only be satisfied if  $\mathbf{S}$  is not anisotropic and under some conditions on the mesh), cf. [31, Theorems 4.3 and 4.4] and [27, Lemma 9.2].

*Remark 3.2 (Inhomogeneous diffusion–dispersion tensors)* Similarly as for the convection term, instead of directly using (3.7a) or (3.7b), we may take into account the distances of  $\mathbf{x}_K$  and  $\mathbf{x}_L$  from the side  $\sigma_{K,L} \in \mathcal{E}_h^{\text{int}}$  common to  $K, L \in \mathcal{T}_h$  for the discretization of inhomogeneous diffusion–dispersion tensors.

### 3.2 Postprocessing for cell-centered finite volumes

The approximate finite volume solution is only piecewise constant, given by the value  $p_K$  in each  $K \in \mathcal{T}_h$ . In order to derive a posteriori error estimates, we construct in this paper a postprocessed approximation which has more regularity. We use for this purpose the additional knowledge that we have from a finite volume scheme: the fluxes.

Developing the ideas of Eymard et al. [28] or of the author [51], the most general postprocessed approximation  $\tilde{p}_h \in H^1(\mathcal{T}_h)$  that we consider in this paper is defined as the weak solution of the following local Neumann problems:

$$-\nabla \cdot (\mathbf{S}\nabla \tilde{p}_h) = \frac{1}{|K|} \sum_{\sigma \in \mathcal{E}_K} S_{K,\sigma} \quad \forall K \in \mathcal{T}_h, \quad (3.14a)$$

$$(1 - \mu_K) \frac{(\tilde{p}_h, 1)_K}{|K|} + \mu_K \tilde{p}_h(\mathbf{x}_K) = p_K \quad \forall K \in \mathcal{T}_h, \quad (3.14b)$$

$$-\mathbf{S}\nabla \tilde{p}_h|_K \cdot \mathbf{n} = \frac{S_{K,\sigma}}{|\sigma|} \quad \forall \sigma \in \mathcal{E}_K, \quad \forall K \in \mathcal{T}_h. \quad (3.14c)$$

Here,  $\mu_K = 0$  or  $1$ , depending on whether the particular scheme (3.1) represents by  $p_K$  the approximate mean value on  $K \in \mathcal{T}_h$  or the approximate point value in a point  $\mathbf{x}_K$  (for simplicity assumed inside  $K$ ) and also on the expected regularity of  $\tilde{p}_h$ . It is immediate to see that such  $\tilde{p}_h$  verifies  $-\mathbf{S}\nabla \tilde{p}_h \in \mathbf{H}(\text{div}, \Omega)$ . For general elements, it is not clear how to practically find such  $\tilde{p}_h$ ; it is, however, easy in two particular cases treated below. We also show below that  $\tilde{p}_h$  is a good approximation of the weak solution  $p$  under some conditions on the original scheme.

#### 3.2.1 Postprocessing on simplices for general diffusion–dispersion tensors

For matching triangular or tetrahedral meshes, it is immediate to show that if  $\mathbf{S} = Id$ ,  $\tilde{p}_h$  given by (3.14a)–(3.14c) is a piecewise second-order polynomial of the form  $\tilde{p}_h|_K = a_K(x^2 + y^2) + b_Kx + c_Ky + d_K$  if  $d = 2$  and  $\tilde{p}_h|_K = a_K(x^2 + y^2 + z^2) +$

$b_K x + c_K y + d_K z + e_K$  if  $d = 3$ . If  $\mathbf{S} \neq Id$ , then  $\tilde{p}_h$  verifying (3.14a)–(3.14c) is a full second-order polynomial on each  $K \in \mathcal{T}_h$  (the symmetry of  $\mathbf{S}$  is important in this case). Simply, (3.14c) imposes the normal derivative of  $\tilde{p}_h$  on the sides of  $K$ , (3.14b) imposes the mean of  $\tilde{p}_h$  over  $K$  or a point value in  $\mathbf{x}_K$ , and (3.14a) follows by the Green theorem and by the fact that  $-\nabla \cdot (\mathbf{S}_K \nabla \tilde{p}_h)$  is constant. Note also that the proposed postprocessing is local on each element and that one only has to build a second-order polynomial on each  $K \in \mathcal{T}_h$  from the prescribed degrees of freedom, so that its cost is negligible. The postprocessed approximation  $\tilde{p}_h$  is, however, in general nonconforming, not included in  $H^1(\Omega)$  but only in  $H^1(\mathcal{T}_h)$ .

### 3.2.2 Postprocessing on rectangular parallelepipeds for diagonal diffusion–dispersion tensors

For  $\mathbf{S}$  diagonal and matching meshes consisting of rectangles or rectangular parallelepipeds,  $\tilde{p}_h$  verifying (3.14a)–(3.14c) is a piecewise second-order polynomial of the form  $\tilde{p}_h|_K = a_K x^2 + b_K y^2 + c_K x + d_K y + e_K$  if  $d = 2$  and  $\tilde{p}_h|_K = a_K x^2 + b_K y^2 + c_K z^2 + d_K x + e_K y + f_K z + g_K$  if  $d = 3$ . Again, the postprocessing is local and the new variable  $\tilde{p}_h$  is nonconforming in the sense that it is in general only in  $H^1(\mathcal{T}_h)$  (but  $-\mathbf{S} \nabla \tilde{p}_h$  belongs to  $\mathbf{H}(\text{div}, \Omega)$ ).

*Remark 3.3 (Comparison with mixed finite elements)* Similar postprocessing was used in [51] for the lowest-order Raviart–Thomas mixed finite element method on simplicial meshes. If  $\Gamma_D = \partial\Omega$  and  $g = 0$ , it is proved in the above reference that in this case, the means of  $\tilde{p}_h$  on interior sides are continuous and its means on exterior sides are equal to zero, so that we have at least some continuity of  $\tilde{p}_h$ . For pure diffusion problems ( $\mathbf{w} = r = 0$ ), it is possible to show the same result also for the cell-centered finite volume scheme (3.1)–(3.7b), but only under the condition  $f = 0$ . It turns out that in general,  $\tilde{p}_h \in H^1(\mathcal{T}_h)$  only. Note also that for pure diffusion problems, cell-centered finite volumes and lowest-order Raviart–Thomas mixed finite elements only differ by numerical integration, cf., e.g., Agouzal et al. [5] or Arbogast et al. [10]. In the general case and on simplicial meshes, lowest-order mixed finite elements are directly, without any numerical integration, equivalent to a particular finite volume scheme, see [49] and also Younès et al. [55].

### 3.2.3 A priori error estimate for the postprocessed approximation

We give in this section an a priori error estimate for the finite volume postprocessed approximation  $\tilde{p}_h$  given by (3.14a)–(3.14c).

**Theorem 3.1 (A priori error estimate for  $\tilde{p}_h$ )** *Let a basic partition  $\tilde{\mathcal{T}}_h$  be given and let its refinement  $\mathcal{T}_h$  be matching, consisting of convex elements, and satisfying Assumptions (B1)–(B2). Let  $p$ , the weak solution of the problem (1.1a)–(1.1c) given by (2.6), be such that  $p \in H^2(\tilde{\mathcal{T}}_h)$  and  $-\mathbf{S} \nabla p \in \mathbf{H}(\text{div}, \Omega)$ . Next, let the finite volume scheme (3.1) be such that*

$$\sum_{K \in \mathcal{T}_h} h_K \left( \sum_{\sigma \in \mathcal{E}_K} \|(S_{K,\sigma} + \langle \mathbf{S} \nabla p \cdot \mathbf{n}, 1 \rangle_\sigma) |\sigma|^{-1} \|\sigma\| \right)^2 \leq Ch^2,$$

$$\sum_{K \in \mathcal{T}_h} \|p - p_K\|_K^2 \leq Ch^2,$$

where the constant  $C > 0$  only depends on  $p$ ,  $d$ ,  $\kappa_{\mathcal{T}}$ ,  $\xi_{\mathcal{T}}$ , the data, and  $\Omega$ . Finally, let  $\tilde{p}_h$  be the postprocessed solution given by (3.14a)–(3.14c). Then

$$\begin{aligned} \sum_{K \in \mathcal{T}_h} \|\nabla(p - \tilde{p}_h)\|_K^2 &\leq Ch^2, \\ \|p - \tilde{p}_h\|_{\Omega}^2 &\leq Ch^2, \end{aligned}$$

where the constant  $C > 0$  only depends on  $p$ ,  $d$ ,  $\kappa_{\mathcal{T}}$ ,  $\xi_{\mathcal{T}}$ , the data, and  $\Omega$ .

*Proof* We first study the estimate on the gradient, following the proof of [28, Theorem 2]. We have, for each  $K \in \mathcal{T}_h$ ,

$$\begin{aligned} c_{S,K} \|\nabla(p - \tilde{p}_h)\|_K^2 &\leq (\mathbf{S}\nabla(p - \tilde{p}_h), \nabla(p - \tilde{p}_h))_K \\ &= -(\nabla \cdot (\mathbf{S}\nabla p), (p - \tilde{p}_h))_K + \langle \mathbf{S}\nabla(p - \tilde{p}_h) \cdot \mathbf{n}, (p - \tilde{p}_h) \rangle_{\partial K} \\ &\leq \|\nabla \cdot (\mathbf{S}\nabla p)\|_K C_{P,K}^{1/2} h_K \|\nabla(p - \tilde{p}_h)\|_K \\ &\quad + \sum_{\sigma \in \mathcal{E}_K} \tilde{C}_{t,K,\sigma}^{1/2} h_K^{1/2} \|\mathbf{S}\nabla(p - \tilde{p}_h) \cdot \mathbf{n}\|_{\sigma} \|\nabla(p - \tilde{p}_h)\|_K, \end{aligned}$$

using Assumption (C1), the Green theorem, the fact that  $\|\nabla(p - \tilde{p}_h)\|_K$  does not depend on shifting  $\tilde{p}_h$  by a constant and fixing temporarily  $(\tilde{p}_h, 1)_K$  by  $(p, 1)_K$  instead of (3.14b), which makes the term  $(\nabla \cdot (\mathbf{S}\nabla \tilde{p}_h), (p - \tilde{p}_h))_K$  disappear by virtue of (3.14a), and finally using the Schwarz inequality, the Poincaré inequality (2.1), and the trace inequality (2.3b). Thus

$$c_{S,K} \|\nabla(p - \tilde{p}_h)\|_K \leq C_{P,K}^{1/2} h_K \|\nabla \cdot (\mathbf{S}\nabla p)\|_K + h_K^{1/2} \sum_{\sigma \in \mathcal{E}_K} \tilde{C}_{t,K,\sigma}^{1/2} \|\mathbf{S}\nabla(p - \tilde{p}_h) \cdot \mathbf{n}\|_{\sigma}. \quad (3.15)$$

Next, the triangle inequality and (3.14c) imply that

$$\begin{aligned} \|\mathbf{S}\nabla(p - \tilde{p}_h) \cdot \mathbf{n}\|_{\sigma} &\leq \|\mathbf{S}\nabla p \cdot \mathbf{n} - \langle \mathbf{S}\nabla p \cdot \mathbf{n}, 1 \rangle_{\sigma} |\sigma|^{-1}\|_{\sigma} \\ &\quad + \|(\langle \mathbf{S}\nabla p \cdot \mathbf{n}, 1 \rangle_{\sigma} + S_{K,\sigma}) |\sigma|^{-1}\|_{\sigma} \end{aligned}$$

and we have

$$\begin{aligned} \|\mathbf{S}\nabla p \cdot \mathbf{n} - \langle \mathbf{S}\nabla p \cdot \mathbf{n}, 1 \rangle_{\sigma} |\sigma|^{-1}\|_{\sigma} &\leq C_{S,K} \|\nabla p \cdot \mathbf{n} - \langle \nabla p \cdot \mathbf{n}, 1 \rangle_{\sigma} |\sigma|^{-1}\|_{\sigma} \\ &\leq C_{S,K} C_K^{1/2} h_K^{1/2} \|p\|_{H^2(K)} \end{aligned}$$

using Assumption (C1) and [28, Lemma 2]; here the constant  $C_K$  only depends on  $d$  and  $\kappa_K$ . Thus, squaring (3.15), summing over all  $K \in \mathcal{T}_h$ , and using the triangle, Cauchy–Schwarz, and the above inequalities gives

$$\begin{aligned} \sum_{K \in \mathcal{T}_h} c_{S,K} \|\nabla(p - \tilde{p}_h)\|_K^2 &\leq \sum_{K \in \mathcal{T}_h} \frac{3}{c_{S,K}} \left\{ C_{P,K} h_K^2 \|\nabla \cdot (\mathbf{S}\nabla p)\|_K^2 \right. \\ &\quad \left. + C_{S,K}^2 C_K h_K^2 \|p\|_{H^2(K)}^2 \left( \sum_{\sigma \in \mathcal{E}_K} \tilde{C}_{t,K,\sigma}^{1/2} \right)^2 \right. \\ &\quad \left. + h_K \left( \sum_{\sigma \in \mathcal{E}_K} \tilde{C}_{t,K,\sigma}^{1/2} \| (S_{K,\sigma} + \langle \mathbf{S}\nabla p \cdot \mathbf{n}, 1 \rangle_{\sigma}) |\sigma|^{-1} \|_{\sigma} \right)^2 \right\}. \end{aligned}$$

Using the assumptions of the theorem, this concludes the proof of the bound for the error in the gradient.

We now turn to the  $L^2(\Omega)$ -estimate on  $p - \tilde{p}_h$ . We have

$$\begin{aligned} \|p - \tilde{p}_h\|_{\Omega}^2 &= \sum_{K \in \mathcal{T}_h} \|p - \tilde{p}_h\|_K^2 \leq 2 \sum_{K \in \mathcal{T}_h} \{ \|p - p_K\|_K^2 + \|p_K - \tilde{p}_h\|_K^2 \} \\ &\leq 2 \sum_{K \in \mathcal{T}_h} \{ \|p - p_K\|_K^2 + C_{P,K} h_K^2 \|\nabla \tilde{p}_h\|_K^2 \} \\ &\leq 2 \sum_{K \in \mathcal{T}_h} \{ \|p - p_K\|_K^2 + 2C_{P,K} h_K^2 \|\nabla(\tilde{p}_h - p)\|_K^2 + 2C_{P,K} h_K^2 \|\nabla p\|_K^2 \}, \end{aligned} \quad (3.16)$$

using the triangle and Cauchy–Schwarz inequalities, (3.14b), and the Poincaré inequality (2.1) when  $\mu_K = 0$  or its continuous version when  $\mu_K = 1$ . The assumptions of the theorem and the previously obtained bound on the gradient of  $p - \tilde{p}_h$  conclude the proof of this second estimate and thus of the whole theorem.  $\square$

*Remark 3.4 (Meaning of Theorem 3.1 and an example)* Theorem 3.1 simply guarantees that if  $S_{K,\sigma}$  from the finite volume scheme (3.1) are  $O(h)$  approximations of the diffusive fluxes through the sides  $\sigma$  of all elements  $K \in \mathcal{T}_h$  and if the original piecewise constant approximation given by  $p_K$  in each  $K \in \mathcal{T}_h$  converges to  $p$  as  $O(h)$  in the  $L^2(\Omega)$ -norm, then  $\tilde{p}_h$  converges to  $p$  as  $O(h)$  in the broken  $H^1(\Omega)$ -norm. Let  $\mathbf{S} = Id$ ,  $\Gamma_N = \emptyset$ , and  $p \in H^2(\Omega)$ . Then this is for example the case, under some additional regularity assumptions on  $\mathbf{w}$ ,  $r$ , and  $g$  and under the assumption that  $d_{K,L} \approx h_K \approx h_L$  for all  $\sigma_{K,L} \in \mathcal{E}_h^{\text{int}}$  and  $d_{K,\sigma} \approx h_K$  for all  $\sigma \in \mathcal{E}_h^{\text{D}}$  and appropriate  $K$ , for the scheme (3.1)–(3.10) (with  $v_\sigma = 1$  for all  $\sigma \in \mathcal{E}_h$ ) on polygonal meshes  $\mathcal{T}_h$  admissible in the sense of [27, Definition 9.1], see [27, Theorem 9.4 and Remark 9.12].

### 3.2.4 Convergence for the postprocessed approximation

For the sake of completeness, we state here the following result, discussing the convergence of  $\tilde{p}_h$  when  $p \in H^1(\Omega)$  only.

**Theorem 3.2 (Convergence of  $\tilde{p}_h$ )** *Let  $\mathbf{w} = r = 0$ ,  $\mathbf{S} = Id$ ,  $\Gamma_N = \emptyset$ , and  $g = 0$ . Let  $\mathcal{T}_h$  be admissible in the sense of [27, Definition 9.1] and satisfying Assumptions (B1)–(B2) and let the finite volume scheme be given by (3.1)–(3.7b). Let finally  $p$  be the weak solution of the problem (1.1a)–(1.1c) given by (2.6) and let  $\tilde{p}_h$  be the postprocessed solution given by (3.14a)–(3.14c). Then*

$$\begin{aligned} \sum_{K \in \mathcal{T}_h} \|\nabla(p - \tilde{p}_h)\|_K^2 &\rightarrow 0 \quad \text{as } h \rightarrow 0, \\ \|p - \tilde{p}_h\|_{\Omega}^2 &\rightarrow 0 \quad \text{as } h \rightarrow 0. \end{aligned}$$

The convergence of the gradient of  $\tilde{p}_h$  to the gradient of  $p$  was proved in [28]. The convergence of  $\tilde{p}_h$  to  $p$  then follows using the estimate (3.16).



#### 4 A posteriori error estimates for locally conservative methods

In this section, we present our a posteriori error estimates for the discretization of the problem (1.1a)–(1.1c) by locally conservative methods. The developments here are completely independent of Section 3 on finite volumes; the only assumption that we make is that we are given a function  $\tilde{p}_h$ , cell values  $p_K$ ,  $K \in \mathcal{T}_h$ , and convective fluxes  $W_{K,\sigma}$ ,  $K \in \mathcal{T}_h$ ,  $\sigma \in \mathcal{E}_K$  (locally conservative, i.e.,  $W_{K,\sigma_{K,L}} = -W_{L,\sigma_{K,L}}$  for all  $\sigma_{K,L} \in \mathcal{E}_h^{\text{int}}$ ) such that

$$\tilde{p}_h \in \mathbb{R}_l(\mathcal{T}_h), \quad (4.1)$$

$$-\mathbf{S}\nabla\tilde{p}_h \in \mathbf{H}(\text{div}, \Omega), \quad (4.2)$$

$$(f + \nabla \cdot (\mathbf{S}\nabla\tilde{p}_h) - rp_K, 1)_K - \sum_{\sigma \in \mathcal{E}_K} W_{K,\sigma} = 0 \quad \forall K \in \mathcal{T}_h, \quad (4.3)$$

$$-\mathbf{S}\nabla\tilde{p}_h \cdot \mathbf{n} = u_\sigma \quad \forall \sigma \in \mathcal{E}_h^{\text{N}}, \quad (4.4)$$

where  $u_\sigma$  are the mean values of the Neumann boundary condition given by (3.6). In the finite volume case, the way to obtain  $\tilde{p}_h$  satisfying (4.1), (4.2), and (4.4) was shown in Section 3.2. The relation (4.3) then easily follows from (3.1) by the Green theorem and by (3.14c), which gives  $(\nabla \cdot (\mathbf{S}\nabla\tilde{p}_h), 1)_K = \langle \mathbf{S}\nabla\tilde{p}_h \cdot \mathbf{n}, 1 \rangle_{\partial K} = -\sum_{\sigma \in \mathcal{E}_K} S_{K,\sigma}$ . We consider general meshes  $\mathcal{T}_h$  but we are aware of the fact that finding  $\tilde{p}_h$  satisfying (4.1)–(4.4) may not be practically feasible for nonmatching meshes or elements other than simplices or rectangular parallelepipeds. Thus an alternative approach for general meshes is developed in Section 5 below. We first state here the a posteriori error estimate and its local efficiency, then give several remarks, and finally present the proofs of the estimate and of its local efficiency.

##### 4.1 A posteriori error estimate

We state here our a posteriori error estimate. Let us first put

$$m_K^2 := \min \left\{ C_{P,K} \frac{h_K^2}{c_{S,K}}, \frac{1}{c_{\mathbf{w},r,K}} \right\}$$

for all  $K \in \mathcal{T}_h$ , where  $C_{P,K}$  is the constant from the Poincaré inequality (2.1). We define the *residual estimator*  $\eta_{R,K}$  by

$$\eta_{R,K} := m_K \|f + \nabla \cdot (\mathbf{S}\nabla\tilde{p}_h) - \nabla \cdot (\tilde{p}_h \mathbf{w}) - r\tilde{p}_h\|_K. \quad (4.5)$$

Let  $\mathcal{I}_{\text{Os}}^{\Gamma}(\tilde{p}_h)$  be the modification of the Oswald interpolate  $\mathcal{I}_{\text{Os}}(\tilde{p}_h)$  at the boundary given by (2.10a)–(2.10b), see Section 2.6. Denote  $v := \tilde{p}_h - \mathcal{I}_{\text{Os}}^{\Gamma}(\tilde{p}_h)$ . Then the *nonconformity estimator*  $\eta_{\text{NC},K}$  is given by

$$\eta_{\text{NC},K} := \|v\|_K \quad (4.6)$$

and the *convection estimator*  $\eta_{C,K}$  by

$$\eta_{C,K} := \min \left\{ \frac{\|\nabla \cdot (v\mathbf{w}) - \frac{1}{2}v\nabla \cdot \mathbf{w}\|_K + \|\nabla \cdot (v\mathbf{w})\|_K}{\sqrt{c_{\mathbf{w},r,K}}}, \left( \frac{C_{P,K}h_K^2\|\nabla v \cdot \mathbf{w}\|_K^2}{c_{S,K}} + \frac{9\|v\nabla \cdot \mathbf{w}\|_K^2}{4c_{\mathbf{w},r,K}} \right)^{1/2} \right\}. \quad (4.7)$$

Next, let

$$m_\sigma^2 := \min \left\{ \max_{K; \sigma \in \mathcal{E}_K} \left\{ C_{F,K,\sigma} \frac{|\sigma| h_K^2}{|K| c_{S,K}} \right\}, \max_{K; \sigma \in \mathcal{E}_K} \left\{ \frac{|\sigma|}{|K| c_{w,r,K}} \right\} \right\} \quad (4.8)$$

for interior or Dirichlet boundary sides. Here  $C_{F,K,\sigma}$  is the constant from the generalized Friedrichs inequality (2.2). The *upwinding estimator*  $\eta_{U,K}$  is given by

$$\eta_{U,K} := \sum_{\sigma \in \mathcal{E}_K \setminus \mathcal{E}_h^N} m_\sigma \| (W_{K,\sigma} - \langle \mathcal{J}_{O_s}^\Gamma(\tilde{p}_h) \mathbf{w} \cdot \mathbf{n}, 1 \rangle_\sigma) |\sigma|^{-1} \|_\sigma. \quad (4.9)$$

Finally, the *reaction quadrature estimator*  $\eta_{RQ,K}$  is given by

$$\eta_{RQ,K} := \frac{1}{\sqrt{c_{w,r,K}}} \| r_K p_K - (r \tilde{p}_h, 1)_K |K|^{-1} \|_K \quad (4.10)$$

and the *Neumann boundary estimator*  $\eta_{I_N,K}$  by

$$\eta_{I_N,K} := 0 + \frac{\sqrt{h_K}}{\sqrt{c_{S,K}}} \sum_{\sigma \in \mathcal{E}_K \cap \mathcal{E}_h^N} \sqrt{C_{t,K,\sigma}} \| u_\sigma - u \|_\sigma, \quad (4.11)$$

where  $C_{t,K,\sigma}$  is the constant from the trace inequality (2.3a). We then have the following a posteriori error estimate:

**Theorem 4.1 (A posteriori error estimate on the mesh  $\mathcal{T}_h$ )** *Let  $p$  be the weak solution of the problem (1.1a)–(1.1c) given by (2.6) and let  $\tilde{p}_h$  satisfy (4.1)–(4.4). Then*

$$\| \| p - \tilde{p}_h \| \|_\Omega \leq \left\{ \sum_{K \in \mathcal{T}_h} \eta_{NC,K}^2 \right\}^{1/2} + \left\{ \sum_{K \in \mathcal{T}_h} (\eta_{R,K} + \eta_{C,K} + \eta_{U,K} + \eta_{RQ,K} + \eta_{I_N,K})^2 \right\}^{1/2}.$$

## 4.2 Local efficiency of the estimate

Let the local Péclet number  $\text{Pe}_K$  and  $\rho_K$  be given by

$$\text{Pe}_K := h_K \frac{C_{w,K}}{c_{S,K}}, \quad \rho_K := \frac{C_{w,K}}{\sqrt{c_{w,r,K}} \sqrt{c_{S,K}}}. \quad (4.12)$$

Next, let

$$\alpha_{*,K} := c_{S,K} \left( \frac{C_{S,K}}{c_{S,K}} + 8\rho_K^2 \right), \quad \beta_{*,K} := \left\| \frac{1}{2} \nabla \cdot \mathbf{w} + r \right\|_{\infty,K} + \frac{5 \|\nabla \cdot \mathbf{w}\|_{\infty,K}^2}{c_{w,r,K}},$$

$$\alpha_{\#,K} := c_{S,K} \left( \frac{C_{S,K}}{c_{S,K}} + C_{P,K} \text{Pe}_K^2 \right), \quad \beta_{\#,K} := \left\| \frac{1}{2} \nabla \cdot \mathbf{w} + r \right\|_{\infty,K} + \frac{9 \|\nabla \cdot \mathbf{w}\|_{\infty,K}^2}{4c_{w,r,K}},$$

and, for  $\varphi \in H^1(K)$ ,

$$\|\varphi\|_{*,K}^2 := \alpha_{*,K} \|\nabla \varphi\|_K^2 + \beta_{*,K} \|\varphi\|_K^2, \quad \|\varphi\|_{\#,K}^2 := \alpha_{\#,K} \|\nabla \varphi\|_K^2 + \beta_{\#,K} \|\varphi\|_K^2.$$

Finally, let

$$c_{\mathbf{S},\mathfrak{T}_K} := \min_{L \in \mathfrak{T}_K} c_{\mathbf{S},L}, \quad c_{\mathbf{w},r,\mathfrak{T}_K} := \min_{L \in \mathfrak{T}_K} c_{\mathbf{w},r,L},$$

and, for  $\varphi \in H^1(\mathcal{T}_h)$ ,

$$\begin{aligned} \|\varphi\|_{b,\mathfrak{E}_K}^2 &:= c_{\mathbf{S},\mathfrak{T}_K} \sum_{\sigma \in \mathfrak{E}_K} h_\sigma^{-1} \|\langle [\![\varphi]\!] , 1 \rangle_\sigma |\sigma|^{-1}\|_\sigma^2 \\ &\quad + c_{\mathbf{w},r,\mathfrak{T}_K} \sum_{\sigma \in \mathfrak{E}_K} h_\sigma \|\langle [\![\varphi]\!] , 1 \rangle_\sigma |\sigma|^{-1}\|_\sigma^2. \end{aligned}$$

The theorem below discusses the local efficiency of the a posteriori error estimators of Theorem 4.1.

**Theorem 4.2 (Local efficiency of the a posteriori error estimators on Theorem 4.1)** *Suppose that the assumptions of Theorem 4.1 are verified. Let moreover  $\mathcal{T}_h$  satisfy Assumption (B). Then, for the residual estimator  $\eta_{\mathbf{R},K}$  (4.5) on each  $K \in \mathcal{T}_h$ , there holds*

$$\eta_{\mathbf{R},K} \leq C_3 \|\|p - \tilde{p}_h\|\|_K \left\{ \sqrt{\frac{C_{\mathbf{S},K}}{c_{\mathbf{S},K}}} \max \left\{ 1, \frac{C_{\mathbf{w},r,K}}{c_{\mathbf{w},r,K}} \right\} + \min \left\{ \text{Pe}_K, \sqrt{\frac{C_{\mathbf{S},K}}{c_{\mathbf{S},K}}} \rho_K \right\} \right\}, \quad (4.13)$$

where the constant  $C_3$  only depends on the space dimension  $d$ , on  $\max_{K \in \mathcal{T}_h} C_{\mathbf{P},K}$ , on the shape of the elements  $K \in \mathcal{T}_h$  and in particular on the regularity parameter  $\kappa_{\mathcal{T}}$ , on the polynomial degree  $l$  of  $\tilde{p}_h$ , and on the polynomial degree  $k$  of  $f$ ,  $\mathbf{w}$ , and  $r$  (see Lemma 4.8 below). Let next for  $K \in \widehat{\mathcal{T}}_h$ , the nonconformity and convection estimators  $\eta_{\text{NC},K}$  and  $\eta_{\text{C},K}$  be respectively given by (4.6) and (4.7). Then

$$\begin{aligned} \eta_{\text{NC},K}^2 + \eta_{\text{C},K}^2 &\leq C_4 \min \left\{ \frac{\alpha_{*,K}}{c_{\mathbf{S},\widehat{\mathfrak{T}}_K}} + \min \left\{ \frac{\beta_{*,K}}{c_{\mathbf{w},r,\widehat{\mathfrak{T}}_K}}, \frac{\beta_{*,K} h_K^2}{c_{\mathbf{S},\widehat{\mathfrak{T}}_K}} \right\}, \frac{\alpha_{\#,K}}{c_{\mathbf{S},\widehat{\mathfrak{T}}_K}} \right. \\ &\quad \left. + \min \left\{ \frac{\beta_{\#,K}}{c_{\mathbf{w},r,\widehat{\mathfrak{T}}_K}}, \frac{\beta_{\#,K} h_K^2}{c_{\mathbf{S},\widehat{\mathfrak{T}}_K}} \right\} \right\} (\|\|p - \tilde{p}_h\|\|_{\widehat{\mathfrak{T}}_K}^2 + \|\|p - \tilde{p}_h\|\|_{b,\widehat{\mathfrak{E}}_K}^2) \\ &\quad + C_4 \beta_{*,K} \inf_{s_h \in \mathbb{R}_l(\widehat{\mathcal{T}}_h) \cap H^1(\Omega)} \|p - s_h\|_{\widehat{\mathfrak{T}}_K}^2 \\ &\quad + 2 \max \left\{ \|\|\mathcal{I}_{\text{Os}}(\tilde{p}_h) - \mathcal{I}_{\text{Os}}^\Gamma(\tilde{p}_h)\|\|_{*,K}^2, \|\|\mathcal{I}_{\text{Os}}(\tilde{p}_h) - \mathcal{I}_{\text{Os}}^\Gamma(\tilde{p}_h)\|\|_{\#,K}^2 \right\} \end{aligned} \quad (4.14)$$

with the constant  $C_4$  only depending on  $d$ ,  $\kappa_{\widehat{\mathcal{T}}}$ , and the polynomial degree  $l$  of  $\tilde{p}_h$  (see Lemma 4.9 below).

### 4.3 Various remarks

We give several remarks in this section. For further comments on this type of estimators, we refer to [51, Section 5].

### 4.3.1 Nature of the estimate and comparisons with different estimators

The basis for our a posteriori error estimates is the construction of the postprocessed approximation  $\tilde{p}_h$  satisfying (4.1)–(4.4). This approximation is more regular than the original piecewise constant one and gives sense to the energy error. In comparison with a posteriori error estimates used in the Galerkin finite element method,  $\tilde{p}_h$  has the crucial advantage that the normal traces of  $-\mathbf{S}\nabla\tilde{p}_h$  are by (4.2) continuous across interior sides. Hence the side error estimators penalizing the mass balance common in Galerkin finite element methods (cf. [46]) do not appear here at all. This advantage is, however, compensated by the fact that  $\tilde{p}_h$  is nonconforming in the sense that it is not included in  $H^1(\Omega)$ , so that the estimator (4.6), penalizing in fact the jumps in  $\tilde{p}_h$  (cf. (4.17) in the proof of Lemma 4.9 below) comes in place. Next, whereas in the lowest-order Galerkin finite element method,  $\nabla \cdot (\mathbf{S}_K \nabla p_h)|_K$  is always equal to zero on all  $K \in \mathcal{T}_h$ , the element residuals (4.5) are relevant even when the original solution is elementwise constant.

Nicaise [36,37] also proposed a posteriori error estimators for the finite volume method. His basic idea is also to first postprocess the original piecewise constant finite volume approximation. He uses for this purpose Morley-type interpolants. However, only the means of the fluxes of this interpolant through the mesh sides are continuous, so that, in the general case, one has to penalize both the improper mass balance of  $-\mathbf{S}\nabla\tilde{p}_h$  and the nonconformity of  $\tilde{p}_h$ . We note, however, that in certain cases, the Morley interpolant is conforming (contained in  $H^1(\Omega)$ ), so that the nonconformity penalization disappears. Another remark in this comparison may be that the postprocessed approximation presented in [36,37] has to be constructed differently in dependence on whether convection and reaction are present. This on the one hand permits to prove the local efficiency of the estimates (see the next section for the discussion of the efficiency of our estimates), but it on the other one complicates the implementation. Finally, the question of the a priori error estimates (convergence) of the postprocessed approximation is not investigated in [36,37].

### 4.3.2 The estimate and its efficiency with respect to $\mathbf{S}$ and $\mathbf{w}$

Concerning the estimators of Theorem 4.1, we first remark that the only constants appearing are those of the inequalities (2.1), (2.2), and (2.3a). They are known for matching meshes consisting of simplices or rectangular parallelepipeds. Note, however, that their evaluation may be more complicated in the general case and that some precaution should be made here. In particular, on nonmatching meshes when convection is present, the assumption that  $|\sigma| \approx h_K^{d-1}$  for all  $\sigma \in \mathcal{E}_K$  (now not necessarily the sides of  $K$  in the geometrical sense, see Section 2.2) it is practically necessary in order that the constant  $C_{F,K,\sigma}$  of (4.8) does not explode. Let us also note that, as the numerical experiments of Section 6 below and in particular Figure 6.5 show, all the estimators featuring these constants are superconvergent and only important on rough grids or in the singular limits.

Concerning the efficiency of Theorem 4.2, we remark that the residual, nonconformity, and convection estimators are semi-robust with respect to the inhomogeneities in the sense that the local efficiency does not depend on the maximal ratio of the inhomogeneities over the whole domain, but for a given element  $K$ , only on

the maximal ratio over  $K$  and elements sharing a vertex with  $K$ ; this is the meaning of the term  $\sqrt{\alpha_{*,K}/c_{\mathbf{S},\widehat{\mathcal{T}}_K}}$  from Theorem 4.2. The local efficiency with respect to anisotropy is then expressed by  $\sqrt{C_{\mathbf{S},K}/c_{\mathbf{S},K}}$  and is always local in a given  $K$ . Next, one can easily see that the local efficiency gets to optimal values as soon as only the local Péclet number (4.12), not the global one, gets sufficiently small. We finally remark that for the nonconformity and convection estimators, the above statements are only valid up to higher-order terms (the part  $\inf_{s_h \in \mathbb{R}_l(\widehat{\mathcal{T}}_h) \cap H^1(\Omega)}$ ), which is, however, only present when  $\frac{1}{2}\nabla \cdot \mathbf{w} + r > 0$ . Moreover, it follows from Lemma 4.9 that these terms disappear when only the terms  $\beta_{*,K}h_K^2/c_{\mathbf{S},\widehat{\mathcal{T}}_K}$  and not the minimum of these terms and of  $\beta_{*,K}/c_{\mathbf{w},r,\widehat{\mathcal{T}}_K}$  (and similarly for #) in (4.14) are present. The efficiency on the original grid  $\mathcal{T}_h$  (if different from  $\widehat{\mathcal{T}}_h$ ) could be easily obtained using the fact that  $\|\|\|\tilde{p}_h - \mathcal{I}_{\text{Os}}(\tilde{p}_h)\|\|\|_{*,K}^2 = \sum_{L \in \widehat{\mathcal{T}}_K} \|\|\|\tilde{p}_h - \mathcal{I}_{\text{Os}}(\tilde{p}_h)\|\|\|_{*,L}^2$  for any  $K \in \mathcal{T}_h$ . Note also that the error in the approximation of the boundary conditions is left aside from these considerations by the last term of the estimate (4.14) on  $\eta_{\text{NC},K}^2 + \eta_{\text{C},K}^2$ . We finally point on the conceptual difference with the mixed finite element methods, see [51, Theorem 4.4]. There  $\tilde{p}_h$  is such that  $\langle \llbracket \tilde{p}_h \rrbracket, 1 \rangle_\sigma = 0$  for all  $\sigma \in \mathcal{E}_h^{\text{int}}$  and one has the local efficiency in the original energy norm. Here, in contrast, the additional term  $\|\|p - \tilde{p}_h\|\|_{b,\widehat{\mathcal{E}}_K}^2$  appears, see also Remark 3.3 in this respect.

Next, note that the reaction quadrature estimator is zero whenever  $r$  is piecewise constant and  $\mu_K$  from (3.14b) is set to zero. If this is not the case, its influence is as that of a quadrature formula—see Section 6.2 below for an example of its influence in a numerical experiment. One could eliminate this estimator by considering directly  $(r\tilde{p}_h, 1)_K$  in (4.3). In what concerns the Neumann boundary estimator, it only penalizes the fact that the Neumann boundary condition is not exactly satisfied by the approximate solution  $\tilde{p}_h$  when it is not given by a piecewise constant function. It has a form similar to Neumann boundary estimators from the finite element method, see [46].

Finally, the fact that the upwinding estimator cannot in general give a lower bound for the error is quite obvious and explained in [51, Section 5.2]. The numerical experiments for finite volumes presented below in Section 6.2, however, show that this estimator represents a higher-order term as soon as the local Péclet number gets small and when the upstream weighting (3.8)–(3.12), modified in the sense of Remark 3.1, is used. Similarly to the reaction quadrature estimator, this estimator would completely disappear while employing  $\mathcal{I}_{\text{Os}}^\Gamma(\tilde{p}_h)$  in the discretization of the convective term. We, however, notice that such a convective flux does not seem too much appealing for practical computations: whereas the scheme (3.1)–(3.12) on admissible triangular meshes is very simple and leads to a 4-point stencil, employing  $\mathcal{I}_{\text{Os}}^\Gamma(\tilde{p}_h)$  in the discretization of the convection term would complicate the definition of the scheme and imply instead of 4 at around 20 nonzero values on each matrix row.

### 4.3.3 The estimate for pure diffusion problems

When  $\mathbf{w} = r = 0$ , on simplices or on a rectangular parallelepipeds when  $\mathbf{S}$  is diagonal and with  $\tilde{p}_h$  constructed as in Section 3.2.1 or 3.2.2, (4.3) implies  $-\nabla \cdot$

$(\mathbf{S}_K \nabla \tilde{p}_h) = f_K$  for all  $K \in \mathcal{T}_h$ . Thus using the fact that  $s \in H^1(\Omega)$  such that  $s|_{\Gamma_D} = g$  in the sense of traces in Lemma 4.2 below can be chosen arbitrarily, our a posteriori error estimate becomes

$$\| \|p - \tilde{p}_h\| \|_{\Omega} \leq \inf_{s \in H^1(\Omega), s|_{\Gamma_D} = g} \| \tilde{p}_h - s \| \|_{\Omega} + \left\{ \sum_{K \in \mathcal{T}_h} (\eta_{R,K} + \eta_{I_N,K})^2 \right\}^{1/2}. \quad (4.15)$$

Note that the residual estimator becomes  $\eta_{R,K} = h_K \sqrt{C_{P,K}} / \sqrt{c_{S,K}} \|f - f_K\|_K$ , a superconvergent term usually considered separately and called ‘‘data oscillation’’. Apparently, it is for this reason that our estimators are almost asymptotically exact and robust in the numerical experiments presented below in Section 6. For further remarks on this point, we refer to [51, Section 5.3] and to [54].

#### 4.3.4 Further generalizations of Theorem 4.1

We remark that for the a posteriori estimate of Theorem 4.1 itself,  $\tilde{p}_h$  does not necessarily have to be a polynomial on each  $K \in \mathcal{T}_h$ ,  $\tilde{p}_h \in H^1(\mathcal{T}_h)$  would be sufficient. The assumption that  $\tilde{p}_h \in \mathbb{R}_l(\mathcal{T}_h)$  is only necessary in the given proof of the local efficiency. The postprocessed approximation  $\tilde{p}_h$  could thus be, in the finite volume setting, obtained from (3.14a)–(3.14c) on general elements. As this, however, not at all practical, a different approach is developed in Section 5 below.

#### 4.4 Proof of the a posteriori error estimate

We give in this section a proof of Theorem 4.1. To begin with, we recall the following lemma ([51, Lemma 7.1]):

**Lemma 4.1 (Abstract framework)** *Let  $p, s \in H^1(\Omega)$  be such that  $p - s \in H_D^1(\Omega)$  and let  $\tilde{p} \in H^1(\mathcal{T}_h)$  be arbitrary. Then*

$$\| \|p - \tilde{p}\| \|_{\Omega} \leq \| \tilde{p} - s \| \|_{\Omega} + \left| \mathcal{B} \left( p - \tilde{p}, \frac{p - s}{\| \|p - s\| \|_{\Omega}} \right) + \sum_{K \in \mathcal{T}_h} \left( \nabla \cdot ((\tilde{p} - s) \mathbf{w}) - \frac{1}{2} (\tilde{p} - s) \nabla \cdot \mathbf{w}, \frac{p - s}{\| \|p - s\| \|_{\Omega}} \right)_K \right|.$$

Consequently, the following bound for the error  $\| \|p - \tilde{p}_h\| \|_{\Omega}$  holds:

**Lemma 4.2 (Abstract error estimate)** *Let the assumptions of Theorem 4.1 be verified and let  $s \in H^1(\Omega)$  such that  $s|_{\Gamma_D} = g$  in the sense of traces be arbitrary. Then*

$$\| \|p - \tilde{p}_h\| \|_{\Omega} \leq \| \tilde{p}_h - s \| \|_{\Omega} + \sup_{\varphi \in H_D^1(\Omega), \| \varphi \|_{\Omega} = 1} \{ T_R(\varphi) + T_C(\varphi) + T_U(\varphi) + T_{RQ}(\varphi) + T_{I_N}(\varphi) \},$$

where

$$\begin{aligned}
T_R(\varphi) &:= \sum_{K \in \mathcal{T}_h} (f + \nabla \cdot (\mathbf{S}\nabla \tilde{p}_h) - \nabla \cdot (\tilde{p}_h \mathbf{w}) - r\tilde{p}_h, \varphi - \varphi_K)_K, \\
T_C(\varphi) &:= \sum_{K \in \mathcal{T}_h} \left\{ (\nabla \cdot ((\tilde{p}_h - s)\mathbf{w}) - \frac{1}{2}(\tilde{p}_h - s)\nabla \cdot \mathbf{w}, \varphi)_K \right. \\
&\quad \left. + \langle (s - \tilde{p}_h)\mathbf{w} \cdot \mathbf{n}, \varphi_K \rangle_{\partial K} \right\}, \\
T_U(\varphi) &:= \sum_{K \in \mathcal{T}_h} \sum_{\sigma \in \mathcal{E}_K} \langle (W_{K,\sigma} - \langle s\mathbf{w} \cdot \mathbf{n}, 1 \rangle_\sigma) |\sigma|^{-1}, \varphi_K \rangle_\sigma, \\
T_{RQ}(\varphi) &:= \sum_{K \in \mathcal{T}_h} (r_K p_K - (r\tilde{p}_h, 1)_K |K|^{-1}, \varphi_K)_K, \\
T_{I_N}(\varphi) &:= \sum_{\sigma \in \mathcal{E}_h^N} \langle u_\sigma - u, \varphi - \varphi_\sigma \rangle_\sigma,
\end{aligned}$$

and where  $\varphi_K$  is the mean of  $\varphi$  over  $K \in \mathcal{T}_h$ ,  $\varphi_K := (\varphi, 1)_K / |K|$ , and  $\varphi_\sigma$  is the mean of  $\varphi$  over  $\sigma \in \mathcal{E}_h$ ,  $\varphi_\sigma := \langle \varphi, 1 \rangle_\sigma / |\sigma|$ .

*Proof* Let us consider an arbitrary  $\varphi \in H_D^1(\Omega)$ . We have, using the bilinearity of  $\mathcal{B}(\cdot, \cdot)$ , the definition (2.6) of the weak solution  $p$ , and the Green theorem in each  $K \in \mathcal{T}_h$ ,

$$\begin{aligned}
&\mathcal{B}(p - \tilde{p}_h, \varphi) \tag{4.16} \\
&= (f, \varphi)_\Omega - \langle u, \varphi \rangle_{\Gamma_N} - \sum_{K \in \mathcal{T}_h} \{ (\mathbf{S}\nabla \tilde{p}_h, \nabla \varphi)_K + (\nabla \cdot (\tilde{p}_h \mathbf{w}), \varphi)_K + (r\tilde{p}_h, \varphi)_K \} \\
&= \sum_{K \in \mathcal{T}_h} \left\{ (f + \nabla \cdot (\mathbf{S}\nabla \tilde{p}_h) - \nabla \cdot (\tilde{p}_h \mathbf{w}) - r\tilde{p}_h, \varphi)_K + \sum_{\sigma \in \mathcal{E}_K \cap \Gamma_N} \langle u_\sigma - u, \varphi \rangle_\sigma \right\}.
\end{aligned}$$

Note that we have in particular used the continuity of the normal trace of  $\mathbf{S}\nabla \tilde{p}_h$  across interior sides following from (4.2), the fact that  $\langle \mathbf{S}\nabla \tilde{p}_h \cdot \mathbf{n}, \varphi \rangle_\sigma = 0$  for  $\sigma \in \mathcal{E}_h^D$  following by  $\varphi \in H_D^1(\Omega)$ , and finally (4.4) for Neumann boundary sides. If  $\sigma \in \mathcal{E}_h^N$ , notice, moreover, that  $\langle u_\sigma - u, \varphi \rangle_\sigma = \langle u_\sigma - u, \varphi - \varphi_\sigma \rangle_\sigma$ , since  $\langle u_\sigma - u, \varphi_\sigma \rangle_\sigma = 0$  by (3.6) (recall that  $\varphi_\sigma$  is a constant).

Now by the local conservativity assumption (4.3), it follows that

$$\begin{aligned}
&(f + \nabla \cdot (\mathbf{S}\nabla \tilde{p}_h) - \nabla \cdot (\tilde{p}_h \mathbf{w}) - r\tilde{p}_h, \varphi)_K + \langle \tilde{p}_h \mathbf{w} \cdot \mathbf{n}, \varphi_K \rangle_{\partial K} - \varphi_K \sum_{\sigma \in \mathcal{E}_K} W_{K,\sigma} \\
&+ (r\tilde{p}_h, \varphi)_K - \varphi_K r_K p_K |K| \pm \langle s\mathbf{w} \cdot \mathbf{n}, \varphi_K \rangle_{\partial K} = 0 \quad \forall K \in \mathcal{T}_h,
\end{aligned}$$

using that  $\varphi_K$  is the constant mean of  $\varphi$  over  $K$  and that  $(\nabla \cdot (\tilde{p}_h \mathbf{w}), 1)_K = \langle \tilde{p}_h \mathbf{w} \cdot \mathbf{n}, 1 \rangle_{\partial K}$  by the Green theorem. Hence we can subtract this term from each summand in (4.16). To conclude the proof, it now suffices to use Lemma 4.1.  $\square$

*Remark 4.1 (Upwinding and reaction quadrature estimators)* Using the fact that  $\varphi_K$  is constant, we have subtracted in the above proof the term  $\langle s\mathbf{w} \cdot \mathbf{n}, 1 \rangle_\sigma |\sigma|^{-1}$  instead of  $s\mathbf{w} \cdot \mathbf{n}$  in  $T_U(\varphi)$ . As for a function  $\psi \in L^2(\sigma)$ ,  $\|\langle \psi, 1 \rangle_\sigma |\sigma|^{-1}\|_\sigma \leq \|\psi\|_\sigma$ , this can considerably reduce the size of the upwinding estimator and it was indeed necessary in the numerical experiments of Section 6.2 below. A similar remark applies to  $T_{RQ}(\varphi)$  as well.

We now estimate the terms  $T_R$ ,  $T_C$ ,  $T_U$ ,  $T_{RQ}$ , and  $T_{I_N}$  separately, putting  $s = \mathcal{J}_{O_S}^\Gamma(\tilde{p}_h)$  in Lemma 4.2.

**Lemma 4.3 (Residual estimate)** *Let  $\varphi \in H_D^1(\Omega)$  be arbitrary. Then*

$$T_R(\varphi) \leq \sum_{K \in \mathcal{T}_h} \eta_{R,K} \|\varphi\|_K,$$

where  $\eta_{R,K}$  is given by (4.5).

*Proof* The Schwarz inequality and Lemma 7.1 from Section 7 below imply

$$\begin{aligned} T_R(\varphi) &\leq \sum_{K \in \mathcal{T}_h} \|f + \nabla \cdot (\mathbf{S}\nabla \tilde{p}_h) - \nabla \cdot (\tilde{p}_h \mathbf{w}) - r\tilde{p}_h\|_K \|\varphi - \varphi_K\|_K \\ &\leq \sum_{K \in \mathcal{T}_h} \eta_{R,K} \|\varphi\|_K. \quad \square \end{aligned}$$

**Lemma 4.4 (Convection estimate)** *Let  $\varphi \in H_D^1(\Omega)$  be arbitrary. Then*

$$T_C(\varphi) \leq \sum_{K \in \mathcal{T}_h} \eta_{C,K} \|\varphi\|_K,$$

where  $\eta_{C,K}$  is given by (4.7).

*Proof* Denote  $v := \tilde{p}_h - \mathcal{J}_{O_S}^\Gamma(\tilde{p}_h)$ . Then, for each  $K \in \mathcal{T}_h$ ,

$$\begin{aligned} &(\nabla \cdot (v\mathbf{w}) - \frac{1}{2}v\nabla \cdot \mathbf{w}, \varphi)_K - \langle v\mathbf{w} \cdot \mathbf{n}, \varphi_K \rangle_{\partial K} \\ &= (\nabla \cdot (v\mathbf{w}) - \frac{1}{2}v\nabla \cdot \mathbf{w}, \varphi)_K - (\nabla \cdot (v\mathbf{w}), \varphi_K)_K \\ &\leq \frac{\|\nabla \cdot (v\mathbf{w}) - \frac{1}{2}v\nabla \cdot \mathbf{w}\|_K + \|\nabla \cdot (v\mathbf{w})\|_K}{\sqrt{c_{\mathbf{w},r,K}}} \|\varphi\|_K, \end{aligned}$$

using the Green theorem. Alternatively,

$$\begin{aligned} &(\nabla v \cdot \mathbf{w}, \varphi - \varphi_K)_K + (\frac{1}{2}v\nabla \cdot \mathbf{w}, \varphi)_K - (v\nabla \cdot \mathbf{w}, \varphi_K)_K \\ &\leq \frac{\sqrt{C_{P,K}}h_K \|\nabla v \cdot \mathbf{w}\|_K}{\sqrt{c_{S,K}}} \sqrt{c_{S,K}} \|\nabla \varphi\|_K + \frac{3\|v\nabla \cdot \mathbf{w}\|_K}{2\sqrt{c_{\mathbf{w},r,K}}} \sqrt{c_{\mathbf{w},r,K}} \|\varphi\|_K \\ &\leq \left( \frac{C_{P,K}h_K^2 \|\nabla v \cdot \mathbf{w}\|_K^2}{c_{S,K}} + \frac{9\|v\nabla \cdot \mathbf{w}\|_K^2}{4c_{\mathbf{w},r,K}} \right)^{1/2} \|\varphi\|_K, \end{aligned}$$

employing the Cauchy–Schwarz inequality and the Poincaré inequality (2.1).  $\square$

**Lemma 4.5 (Upwinding estimate)** *Let  $\varphi \in H_D^1(\Omega)$  be arbitrary. Then*

$$T_U(\varphi) \leq \sum_{K \in \mathcal{T}_h} \eta_{U,K} \|\varphi\|_K,$$

where  $\eta_{U,K}$  is given by (4.9).



*Proof* We have

$$\begin{aligned} T_U(\varphi) &= \sum_{\sigma_{K,L} \in \mathcal{E}_h^{\text{int}}} \langle (W_{K,\sigma_{K,L}} - \langle \mathcal{J}_{\text{Os}}^\Gamma(\tilde{p}_h) \mathbf{w} \cdot \mathbf{n}_K, 1 \rangle_{\sigma_{K,L}}) |\sigma_{K,L}|^{-1}, \varphi_K - \varphi_L \rangle_{\sigma_{K,L}} \\ &\quad + \sum_{\sigma_K \in \mathcal{E}_h^{\text{D}}} \langle (W_{K,\sigma_K} - \langle \mathcal{J}_{\text{Os}}^\Gamma(\tilde{p}_h) \mathbf{w} \cdot \mathbf{n}_K, 1 \rangle_{\sigma_K}) |\sigma_K|^{-1}, \varphi_K \rangle_{\sigma_K}, \end{aligned}$$

using that  $\mathbf{w} \cdot \mathbf{n}$  as well as  $\mathcal{J}_{\text{Os}}^\Gamma(\tilde{p}_h)$  are continuous across interelement sides, a similar continuity of the finite volume convective fluxes  $W_{K,\sigma}$ , and (2.10b). Let  $\sigma_{K,L} \in \mathcal{E}_h^{\text{int}}$  and let us put  $\varphi_{\sigma_{K,L}} := \langle \varphi, 1 \rangle_{\sigma_{K,L}} / |\sigma_{K,L}|$ . Then

$$\begin{aligned} \|\varphi_K - \varphi_L\|_{\sigma_{K,L}} &\leq \|\varphi_K - \varphi_{\sigma_{K,L}}\|_{\sigma_{K,L}} + \|\varphi_L - \varphi_{\sigma_{K,L}}\|_{\sigma_{K,L}} \\ &\leq \max_{M=\{K,L\}} \left\{ C_{F,M,\sigma_{K,L}} \frac{|\sigma_{K,L}| h_M^2}{|M| c_{S,M}} \right\}^{1/2} (\|\varphi\|_K + \|\varphi\|_L) \end{aligned}$$

by the triangle inequality and the first estimate of Lemma 7.2 from Section 7 below. At the same time,

$$\begin{aligned} \|\varphi_K - \varphi_L\|_{\sigma_{K,L}} &\leq \|\varphi_K\|_{\sigma_{K,L}} + \|\varphi_L\|_{\sigma_{K,L}} \\ &\leq \max_{M=\{K,L\}} \left\{ \frac{|\sigma_{K,L}|}{|M| c_{w,r,M}} \right\}^{1/2} (\|\varphi\|_K + \|\varphi\|_L), \end{aligned}$$

using the triangle inequality and the second estimate of Lemma 7.2. Similar estimates on  $\|\varphi_K\|_{\sigma_K}$  for  $\sigma_K \in \mathcal{E}_h^{\text{D}}$  follow directly from Lemma 7.2 using that  $\varphi_{\sigma_K} = 0$  on Dirichlet boundary sides by  $\varphi \in H_D^1(\Omega)$ . Hence, with  $m_\sigma$  given by (4.8),

$$\begin{aligned} T_U(\varphi) &\leq \sum_{\sigma \in \mathcal{E}_h \setminus \mathcal{E}_h^{\text{N}}} \left\{ m_\sigma \| (W_{K,\sigma} - \langle \mathcal{J}_{\text{Os}}^\Gamma(\tilde{p}_h) \mathbf{w} \cdot \mathbf{n}, 1 \rangle_\sigma) |\sigma|^{-1} \|_\sigma \sum_{K;\sigma \in \mathcal{E}_K} \|\varphi\|_K \right\} \\ &= \sum_{K \in \mathcal{T}_h} \left\{ \sum_{\sigma \in \mathcal{E}_K \setminus \mathcal{E}_h^{\text{N}}} m_\sigma \| (W_{K,\sigma} - \langle \mathcal{J}_{\text{Os}}^\Gamma(\tilde{p}_h) \mathbf{w} \cdot \mathbf{n}, 1 \rangle_\sigma) |\sigma|^{-1} \|_\sigma \right\} \|\varphi\|_K. \end{aligned}$$

Noticing the definition of  $\eta_{U,K}$  by (4.9) concludes the proof.  $\square$

**Lemma 4.6 (Reaction quadrature estimate)** *Let  $\varphi \in H_D^1(\Omega)$  be arbitrary. Then*

$$T_{\text{RQ}}(\varphi) \leq \sum_{K \in \mathcal{T}_h} \eta_{\text{RQ},K} \|\varphi\|_K,$$

where  $\eta_{\text{RQ},K}$  is given by (4.10).

*Proof* The Schwarz inequality and the definition of  $\|\cdot\|_K$  by (2.5) imply

$$T_{\text{R}}(\varphi) \leq \sum_{K \in \mathcal{T}_h} \|r_K p_K - (r \tilde{p}_h, 1)_K |K|^{-1}\|_K \|\varphi\|_K \leq \sum_{K \in \mathcal{T}_h} \eta_{\text{RQ},K} \|\varphi\|_K. \quad \square$$

**Lemma 4.7 (Neumann boundary estimate)** *Let  $\varphi \in H_D^1(\Omega)$  be arbitrary. Then*

$$T_{\Gamma_N}(\varphi) \leq \sum_{K \in \mathcal{T}_h} \eta_{\Gamma_N, K} \|\varphi\|_K,$$

where  $\eta_{\Gamma_N, K}$  is given by (4.11).

*Proof* We have, using the trace inequality (2.3a) and the definition of  $\|\cdot\|_K$  by (2.5),

$$\begin{aligned} T_{\Gamma_N}(\varphi) &= \sum_{K \in \mathcal{T}_h} \sum_{\sigma \in \mathcal{E}_K \cap \mathcal{E}_h^N} \langle u_\sigma - u, \varphi - \varphi_\sigma \rangle_\sigma \\ &\leq \sum_{K \in \mathcal{T}_h} \|\nabla \varphi\|_K h_K^{1/2} \sum_{\sigma \in \mathcal{E}_K \cap \mathcal{E}_h^N} C_{t, K, \sigma}^{1/2} \|u_\sigma - u\|_\sigma \leq \sum_{K \in \mathcal{T}_h} \eta_{\Gamma_N, K} \|\varphi\|_K. \quad \square \end{aligned}$$

Lemmas 4.1–4.7 and the Cauchy–Schwarz inequality prove Theorem 4.1.

#### 4.5 Proof of the local efficiency of the estimate

We give in this section a proof of Theorem 4.2.

**Lemma 4.8 (Local efficiency of the residual estimator)** *Let  $K \in \mathcal{T}_h$ , and let  $\eta_{R, K}$  be the residual estimator given by (4.5). Then (4.13) holds true.*

The proof of this lemma follows the one given in [46], is based on the equivalence of norms on finite-dimensional spaces, the inverse inequality, and the definition of  $\|\cdot\|_K$  by (2.5), see [51, Lemma 7.6] for the simplicial case.

**Lemma 4.9 (Local efficiency of the nonconformity and velocity estimators)** *Let  $K \in \widehat{\mathcal{T}}_h$  and let  $\eta_{NC, K}$  and  $\eta_{C, K}$  be the nonconformity and velocity estimators given respectively by (4.6) and (4.7). Then (4.14) holds true.*

*Proof* Let us first recall that  $K \in \widehat{\mathcal{T}}_h$ , where  $\widehat{\mathcal{T}}_h$  is the matching submesh of  $\mathcal{T}_h$  consisting of simplices or rectangular parallelepipeds of Assumption (B). Recall also that by construction of Section 2.6,  $\mathcal{I}_{Os}(\tilde{p}_h) \in \mathbb{R}_I(\widehat{\mathcal{T}}_h) \cap H^1(\Omega)$  is the Oswald interpolate of  $\tilde{p}_h$  taken as an element of  $\mathbb{R}_I(\widehat{\mathcal{T}}_h)$ . We will use this fact in the proof.

To begin, it is easy to show that

$$\eta_{NC, K}^2 + \eta_{C, K}^2 \leq \min \{ \|\tilde{p}_h - \mathcal{I}_{Os}^\Gamma(\tilde{p}_h)\|_{*, K}^2, \|\tilde{p}_h - \mathcal{I}_{Os}^\Gamma(\tilde{p}_h)\|_{\#, K}^2 \}$$

with  $\|\cdot\|_{*, K}$  and  $\|\cdot\|_{\#, K}$  defined in Section 4.2. Thus

$$\begin{aligned} \eta_{NC, K}^2 + \eta_{C, K}^2 &\leq 2 \min \{ \|\tilde{p}_h - \mathcal{I}_{Os}(\tilde{p}_h)\|_{*, K}^2, \|\tilde{p}_h - \mathcal{I}_{Os}(\tilde{p}_h)\|_{\#, K}^2 \} \\ &\quad + 2 \max \{ \|\mathcal{I}_{Os}(\tilde{p}_h) - \mathcal{I}_{Os}^\Gamma(\tilde{p}_h)\|_{*, K}^2, \|\mathcal{I}_{Os}(\tilde{p}_h) - \mathcal{I}_{Os}^\Gamma(\tilde{p}_h)\|_{\#, K}^2 \} \end{aligned}$$

follows immediately. The rest of the proof, devoted to showing a bound on  $\|\tilde{p}_h - \mathcal{I}_{Os}(\tilde{p}_h)\|_{*, K}$ , follows the path of that given in [51, Lemma 7.7]. The proof for  $\|\tilde{p}_h - \mathcal{I}_{Os}(\tilde{p}_h)\|_{\#, K}$  can be established likewise.

Let henceforth  $C$  denote a constant only depending on  $d$ ,  $\kappa_{\widehat{\mathcal{T}}}$ , and  $l$ , not necessarily the same at each occurrence. It follows directly from Lemma 2.1 and the definition of  $\|\cdot\|_{*,K}$  that

$$\|\tilde{p}_h - \mathcal{I}_{\text{Os}}(\tilde{p}_h)\|_{*,K}^2 \leq C \left( \alpha_{*,K} \sum_{\sigma \in \widehat{\mathcal{E}}_K} h_\sigma^{-1} \|[[\tilde{p}_h]]\|_\sigma^2 + \beta_{*,K} \sum_{\sigma \in \widehat{\mathcal{E}}_K} h_\sigma \|[[\tilde{p}_h]]\|_\sigma^2 \right). \quad (4.17)$$

We will next use the inequality

$$h_\sigma^{-1/2} \|[[\psi_h]]\|_\sigma \leq C \sum_{L: \sigma \in \widehat{\mathcal{E}}_L} \|\nabla(\psi_h - \varphi)\|_L, \quad (4.18)$$

where  $\sigma \in \widehat{\mathcal{E}}_h^{\text{int}}$ ,  $\psi_h \in H^1(\widehat{\mathcal{T}}_h)$  such that  $\langle [[\psi_h]], 1 \rangle_\sigma = 0$ , and  $\varphi \in H^1(\Omega)$  are arbitrary. This inequality was established in [2, Theorem 10] for simplicial meshes and it generalizes easily to rectangular parallelepipeds. The triangle inequality, the fact that  $\langle [[p]], 1 \rangle_\sigma = 0$ , and a few manipulations then imply

$$h_\sigma^{-1/2} \|[[\tilde{p}_h]]\|_\sigma \leq C \sum_{L: \sigma \in \widehat{\mathcal{E}}_L} \|\nabla(\tilde{p}_h - \varphi)\|_L + h_\sigma^{-1/2} \|\langle [[p - \tilde{p}_h]], 1 \rangle_\sigma | \sigma |^{-1} \|_\sigma.$$

Thus, finally,

$$h_\sigma^\gamma \|[[\tilde{p}_h]]\|_\sigma^2 \leq \frac{Ch_\sigma^{\gamma+1}}{\min_{L: \sigma \in \widehat{\mathcal{E}}_L} c_{S,L}} \sum_{L: \sigma \in \widehat{\mathcal{E}}_L} \|\mathbf{S}\nabla(\tilde{p}_h - p)\|_L^2 + 2h_\sigma^\gamma \|\langle [[p - \tilde{p}_h]], 1 \rangle_\sigma | \sigma |^{-1} \|_\sigma^2,$$

where we put  $\gamma = -1, 1$ . Next, for an arbitrary  $\psi_h \in \mathbb{R}_l(\widehat{\mathcal{T}}_h)$  such that  $\langle [[\psi_h]], 1 \rangle_\sigma = 0$  and an arbitrary  $s_h \in \mathbb{R}_l(\widehat{\mathcal{T}}_h) \cap H^1(\Omega)$ ,

$$\begin{aligned} h_\sigma^{1/2} \|[[\psi_h]]\|_\sigma &\leq h_\sigma C \sum_{L: \sigma \in \widehat{\mathcal{E}}_L} \|\nabla(\psi_h - s_h)\|_L \leq C \sum_{L: \sigma \in \widehat{\mathcal{E}}_L} h_L \|\nabla(\psi_h - s_h)\|_L \\ &\leq C \sum_{L: \sigma \in \widehat{\mathcal{E}}_L} \|\psi_h - s_h\|_L \leq C \sum_{L: \sigma \in \widehat{\mathcal{E}}_L} \|\psi_h - p\|_L + C \sum_{L: \sigma \in \widehat{\mathcal{E}}_L} \|p - s_h\|_L, \end{aligned}$$

using (4.18), the inverse inequality, and the triangle inequality. Hence

$$\begin{aligned} h_\sigma \|[[\tilde{p}_h]]\|_\sigma^2 &\leq C \frac{1}{\min_{L: \sigma \in \widehat{\mathcal{E}}_L} c_{\mathbf{w},r,L}} \sum_{L: \sigma \in \widehat{\mathcal{E}}_L} \|(\tfrac{1}{2}\nabla \cdot \mathbf{w} + r)^{1/2}(\tilde{p}_h - p)\|_L^2 \\ &\quad + C \sum_{L: \sigma \in \widehat{\mathcal{E}}_L} \|p - s_h\|_L^2 + 2h_\sigma \|\langle [[p - \tilde{p}_h]], 1 \rangle_\sigma | \sigma |^{-1} \|_\sigma^2 \end{aligned} \quad (4.19)$$

holds as well, which gives a sense when all  $c_{\mathbf{w},r,L}$  for  $L$  such that  $\sigma \in \widehat{\mathcal{E}}_L$  are nonzero. Combining estimates (4.17)–(4.19) while estimating  $\min_{L: \sigma \in \widehat{\mathcal{E}}_L} c_L$  for a side  $\sigma$  such that  $\sigma \cap K \neq \emptyset$  from below by  $\min_{L \in \widehat{\mathcal{T}}_K} c_L$  concludes the proof.  $\square$

Lemmas 4.8–4.9 together prove Theorem 4.2.

## 5 Estimates on arbitrary polygonal grids by the solution of local Neumann problems

We propose here an approach alternative to finding  $\tilde{p}_h$  satisfying (4.1)–(4.4), which may be difficult (impossible) to carry out practically on arbitrary polygonal grids. This approach is motivated by the simple postprocessing (3.14a)–(3.14c) of the finite volume case on simplices. We only consider in this section pure diffusion problems ( $\mathbf{w} = r = 0$ ).

Let thus  $\mathcal{T}_h$  be a (nonmatching) grid consisting of arbitrary polygons and let  $\widehat{\mathcal{T}}_h$  be its matching refinement consisting of simplices and satisfying Assumption (A) (cf. Assumption (B3)). Let a locally conservative scheme be given on  $\mathcal{T}_h$ . This scheme then should define discrete diffusive fluxes  $S_{K,\sigma}$ ,  $K \in \mathcal{T}_h$  and  $\sigma \in \mathcal{E}_K$ , and relation like (3.1) should hold. Now instead of directly searching  $\tilde{p}_h$  verifying (4.1)–(4.4), let us consider approximations by, e.g., the given locally conservative method to the local Neumann problems (3.14a)–(3.14c) on the simplicial grids  $\widehat{\mathcal{T}}_K$  of  $K$ , for all  $K \in \mathcal{T}_h$ . Suppose moreover that for  $K \in \mathcal{T}_h$ , the given scheme preserve exactly the given Neumann boundary fluxes  $S_{K,\sigma}$ ,  $\sigma \in \mathcal{E}_K$  (this is the case of the majority of locally conservative methods) and note that by (3.14b), the approximate values  $p_L$ ,  $L \in \widehat{\mathcal{T}}_K$ , will be closely related to the original  $p_K$ .

Now since the mesh  $\widehat{\mathcal{T}}_h$  is simplicial, we can immediately use the postprocessing (3.14a)–(3.14c) on  $\widehat{\mathcal{T}}_h$  in order to construct a piecewise second-order polynomial  $\tilde{p}_h$  satisfying (4.1)–(4.4) on the mesh  $\widehat{\mathcal{T}}_h$  and have all the results of Section 4 for this  $\tilde{p}_h$ . Note in particular that if  $\widehat{\mathcal{T}}_K$  for some  $K \in \mathcal{T}_h$  only contains  $K$ , i.e., when  $K$  is itself a simplex, then the present approach coincides with constructing  $\tilde{p}_h|_K$  directly by (3.14a)–(3.14c).

We summarize the above results in the following theorem:

**Theorem 5.1 (A posteriori error estimate for pure diffusion problems on arbitrary polygonal grids by the solution of local Neumann problems)** *Let  $p$  be the weak solution of the problem (1.1a)–(1.1c) given by (2.6), with  $\mathbf{w} = r = 0$ . Let  $\mathcal{T}_h$  satisfy Assumption (B) and let  $\widehat{\mathcal{T}}_h$  be its matching refinement, consisting of simplices. Next, let a locally conservative scheme of the form (3.1) on  $\mathcal{T}_h$  be given and let  $\tilde{p}_h$  be given by the approximations of (3.14a)–(3.14c) on the simplicial grids  $\widehat{\mathcal{T}}_K$  of  $K$  for all  $K \in \mathcal{T}_h$  and the subsequent postprocessing by (3.14a)–(3.14c) on  $\widehat{\mathcal{T}}_h$  in the sense described above. Then  $\tilde{p}_h$  satisfies (4.1)–(4.4) and all the results of Section 4 hold true for this  $\tilde{p}_h$  on the mesh  $\widehat{\mathcal{T}}_h$ .*

*Remark 5.1 (Relation to a posteriori error estimates based on the solution of local Neumann problems in the finite element method)* Bank and Weiser [11] derive a posteriori error estimates in the finite element method on the basis of approximate solution of local Neumann problems. In their case, the solutions of the local problems serve to define an a posteriori error estimator for the original finite element approximation. Contrarily, we have in this section used the solutions of the local problems in order to define a new approximate solution, for which we can easily give an a posteriori estimate.

## 6 Numerical experiments

We test our a posteriori error estimates on two model problems in this section. The first problem contains a strongly inhomogeneous diffusion–dispersion tensor and the second one is convection-dominated. In both cases, the analytical solution is known.

We use the finite volume scheme (3.1)–(3.12), which we extend from triangular grids admissible in the sense of [27, Definition 9.1] to strictly Delaunay triangular meshes (the circumcircle of each triangle does not contain any vertex in its interior or on its boundary and no circumcenters of boundary triangles lie outside the domain or on its boundary), cf. [27, Example 9.1]. The harmonic averaging (3.7b) for the diffusion–dispersion tensor is employed while modifying it in the sense of Remark 3.2, i.e., by taking into account the distances of the circumcenters  $\mathbf{x}_K$ ,  $K \in \mathcal{T}_h$ , from the sides of  $K$ . When  $\sigma_{K,L} \in \mathcal{E}_h^{\text{int}}$  and  $\mathbf{x}_K$  lies in  $L$  (which may happen for strictly Delaunay triangular meshes),  $s_{K,L} = s_L$ . When convection is present, we use the local Péclet upstream weighting (3.8)–(3.12), modified in the sense of Remark 3.1. The postprocessed potential  $\tilde{p}_h$  is given by (3.14a)–(3.14c) with  $\mu_K = 1$  for all  $K \in \mathcal{T}_h$  and  $\mathbf{x}_K$  being the circumcenter of  $K$ , see Section 3.2.1. All these choices seem to be important for the presented numerical results of our a posteriori error estimates.

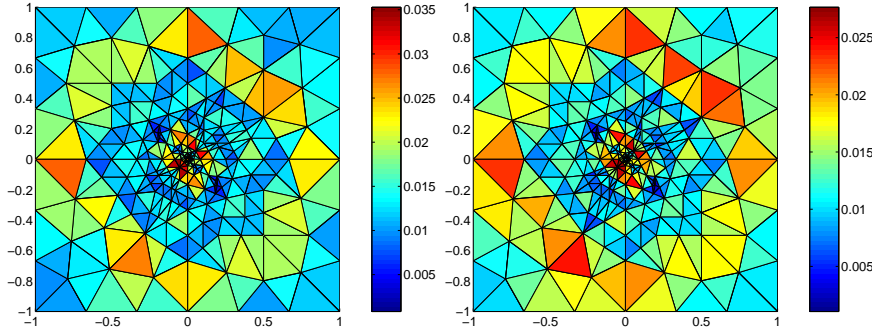
### 6.1 Model problem with strongly inhomogeneous diffusion–dispersion tensor

This model problem is taken from [41, 25] and is motivated by the fact that in real-life applications, the diffusion–dispersion tensor  $\mathbf{S}$  may be discontinuous and strongly inhomogeneous. We consider in particular  $\Omega = (-1, 1) \times (-1, 1)$  and the equation (1.1a) with  $\mathbf{w} = 0$ ,  $r = 0$ , and  $f = 0$ . We suppose that  $\Omega$  is divided into four subdomains  $\Omega_i$  corresponding to the axis quadrants (in the counterclockwise direction) and that  $\mathbf{S}$  is constant and equal to  $s_i Id$  in  $\Omega_i$ . Under such conditions, analytical solution writing

$$p(r, \theta)|_{\Omega_i} = r^\alpha (a_i \sin(\alpha\theta) + b_i \cos(\alpha\theta))$$

in each  $\Omega_i$  can be found. Here  $(r, \theta)$  are the polar coordinates in  $\Omega$ ,  $a_i$  and  $b_i$  are constants depending on  $\Omega_i$ , and  $\alpha$  is a parameter. This solution is continuous across the interfaces but only the normal component of its flux  $\mathbf{u} = -\mathbf{S}\nabla p$  is continuous; it finally exhibits a singularity at the origin. We assume Dirichlet boundary conditions given by this solution and consider two sets of the coefficients, with  $s_1 = s_3 = 5$ ,  $s_2 = s_4 = 1$  in the first case and  $s_1 = s_3 = 100$ ,  $s_2 = s_4 = 1$  in the second one:

$$\begin{aligned} \alpha &= 0.53544095 \\ a_1 &= 0.44721360 & b_1 &= 1 \\ a_2 &= -0.74535599 & b_2 &= 2.33333333 \\ a_3 &= -0.94411759 & b_3 &= 0.55555556 \\ a_4 &= -2.40170264 & b_4 &= -0.48148148, \end{aligned}$$



**Fig. 6.1** Estimated (left) and actual (right) error distribution,  $\alpha = 0.53544095$

$$\alpha = 0.12690207$$

$a_1 = 0.1$	$b_1 = 1$
$a_2 = -9.60396040$	$b_2 = 2.96039604$
$a_3 = -0.48035487$	$b_3 = -0.88275659$
$a_4 = 7.70156488$	$b_4 = -6.45646175$

The original grid consisted of 112 triangles and we have refined it either uniformly (up to 4 refinements) or adaptively on the basis of our estimator. The latter case is a little complicated, since we need the refined mesh to still be strictly Delaunay, and hence the usual “longest edge” refinement or its variants cannot be used. We thus use the following concept: to the given set of vertices, we first add those which correspond to edges midpoints of such triangles where the estimated  $\|\cdot\|_{\Omega}$ -error is greater than the half of the maximum of the estimators. Then the *Triangle* mesh generator, see Shewchuk [43], is used to produce a mesh which comprises the given set of vertices, respects the four subdomains, and guarantees a minimal angle (20 degrees in the given case). This mesh is then checked for the (uniform) strict Delaunay condition: the sum of the two opposite angles for each interior edge has to be less than or equal to  $\pi - \alpha$  where  $\alpha$  is a positive constant and similarly at the boundary. If it does not satisfy this condition, further vertices are added and the process is repeated until the mesh is (uniformly) strictly Delaunay. Since  $f = 0$ , the residual estimators  $\eta_{R,K}$  (4.5) are zero for each  $K \in \mathcal{T}_h$  (recall that this would be the case for general piecewise constant  $f$ , cf. Section 4.3.3), and hence the a posteriori error estimate is entirely given by the nonconformity estimator  $\eta_{NC,K}$  (4.6).

We give in Figure 6.1 an example of our a posteriori estimate on the error and its distribution and the actual error and its distribution on an adaptively refined mesh for the first test case. We can see that the predicted distribution is excellent and that in particular even in this case where the solution is smoother, the singularity is well recognized. Next, Figure 6.2 gives an example of the approximate solution on an adaptively refined mesh and this mesh in the second test case. Here, the singularity is much more important and consequently the grid is highly refined around the origin (for an adaptively refined grid of 2000 triangles, the diameter of the smallest triangles near the origin is  $10^{-15}$  and 80% of the triangles are contained in the circle of radius 0.1). Figure 6.3 then reports the estimated and

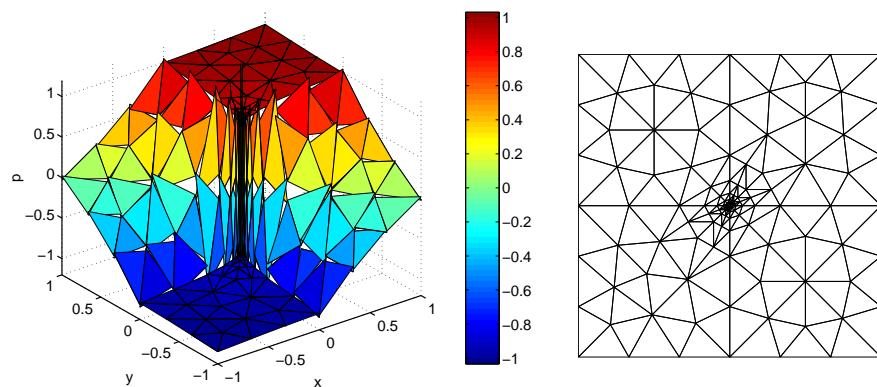


Fig. 6.2 Approximate solution and the corresponding adaptively refined mesh,  $\alpha = 0.12690207$

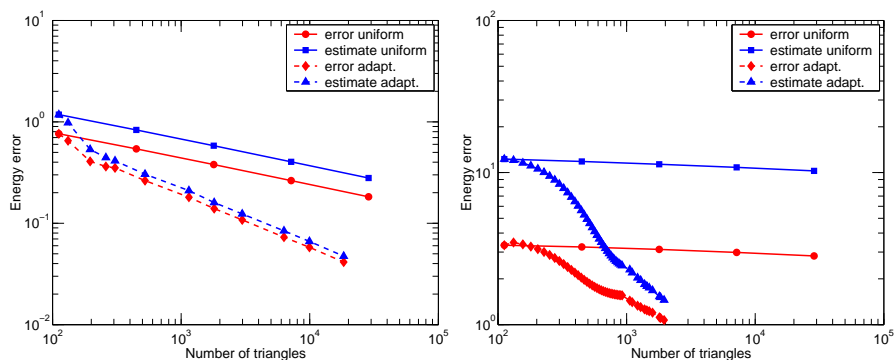


Fig. 6.3 Estimated and actual error against the number of elements in uniformly/adaptively refined meshes for  $\alpha = 0.53544095$  (left) and  $\alpha = 0.12690207$  (right)

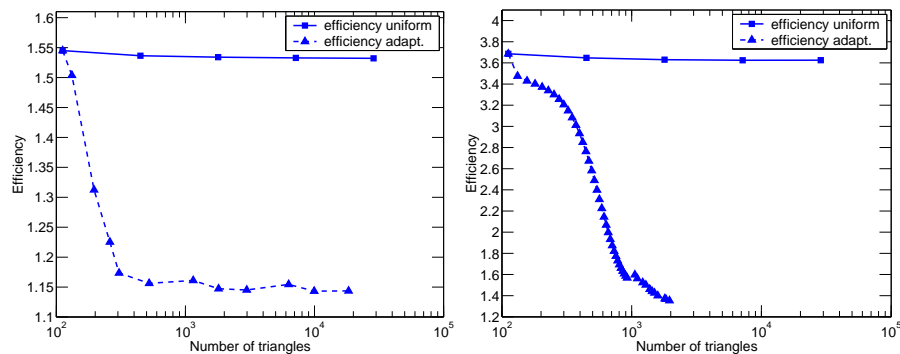
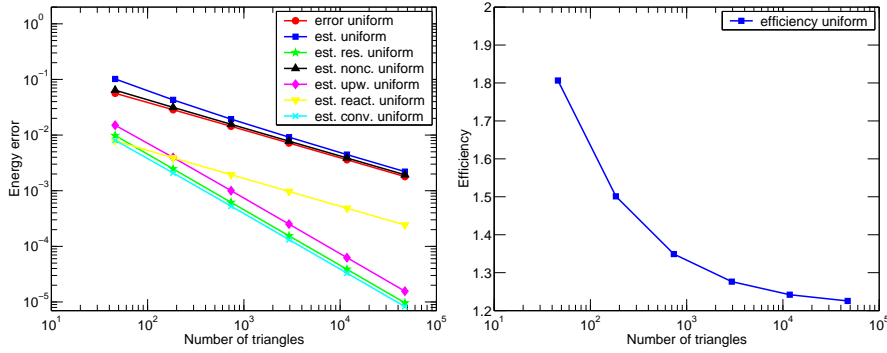


Fig. 6.4 Overall efficiency of the a posteriori error estimates against the number of elements in uniformly/adaptively refined meshes for  $\alpha = 0.53544095$  (left) and  $\alpha = 0.12690207$  (right)



**Fig. 6.5** Estimated and actual error and the different estimators (left) and overall efficiency (right) against the number of elements,  $\varepsilon = 1$ ,  $a = 0.5$

actual errors on uniformly/adaptively refined grids in the two test cases. The energy seminorm (2.5) was approximated with a 7-point quadrature formula in each triangle. One can clearly substantially reduce the number of unknowns necessary to attain the prescribed precision using the derived a posteriori error estimates and adaptively refined grids. Finally, we can see in Figure 6.4 the efficiency plots for the two cases, giving the ratio of the estimated  $\|\cdot\|_{\Omega}$ -error to the actual  $\|\cdot\|_{\Omega}$ -error. This quantity simply expresses how many times we have overestimated the actual error—recall that there are no undetermined multiplicative constants in our estimates. These plots confirm the theoretical results of Section 4.3.3. Even while only using  $\mathcal{J}_{\text{Os}}(\tilde{p}_h)$  instead of evaluating the infimum in (4.15), (approximate) asymptotic exactness and robustness with respect to inhomogeneities is confirmed.

## 6.2 Convection-dominated model problem

This problem is a modification of a problem considered in [26]. We put  $\Omega = (0, 1) \times (0, 1)$ ,  $\mathbf{w} = (0, 1)$ , and  $r = 1$  in (1.1a) and consider three cases with  $\mathbf{S} = \varepsilon Id$  and  $\varepsilon$  equal to, respectively, 1,  $10^{-2}$ , and  $10^{-4}$ . The right-hand side term  $f$ , Neumann boundary conditions on the upper side, and Dirichlet boundary conditions elsewhere are chosen so that the solution was

$$p(x, y) = 0.5 \left( 1 - \tanh \left( \frac{0.5 - x}{a} \right) \right).$$

This solution is in fact one-dimensional and possesses an internal layer of width  $a$  which we set, respectively, equal to 0.5, 0.05, and 0.02. We start the computations from an unstructured grid of  $\Omega$  consisting of 46 triangles and refine it either uniformly (up to 5 refinements) or adaptively.

For  $\varepsilon = 1$  and  $a = 0.5$  (diffusion-dominated regime), estimated and actual errors in the energy norm (2.5), the different estimators, and the efficiency on uniformly refined grids are reported in Figure 6.5. Note in particular that in this regime, the residual, convection, and upwinding estimators represent higher-order terms and that the influence of the reaction quadrature estimator is limited. Finally, our estimator reproduced very precisely the distribution of the error in this case.



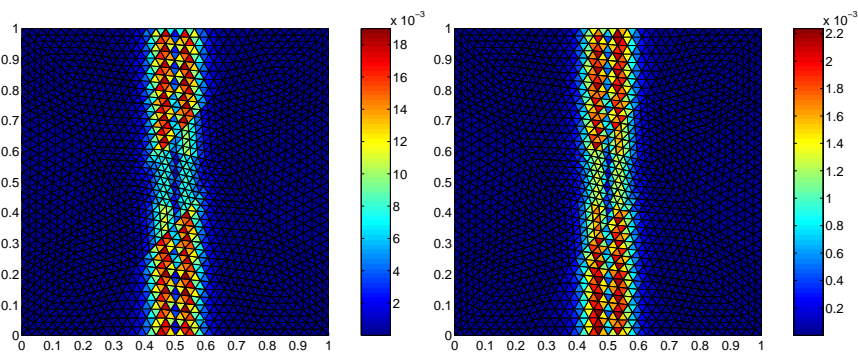


Fig. 6.6 Estimated (left) and actual (right) error distribution,  $\varepsilon = 10^{-2}$ ,  $a = 0.05$

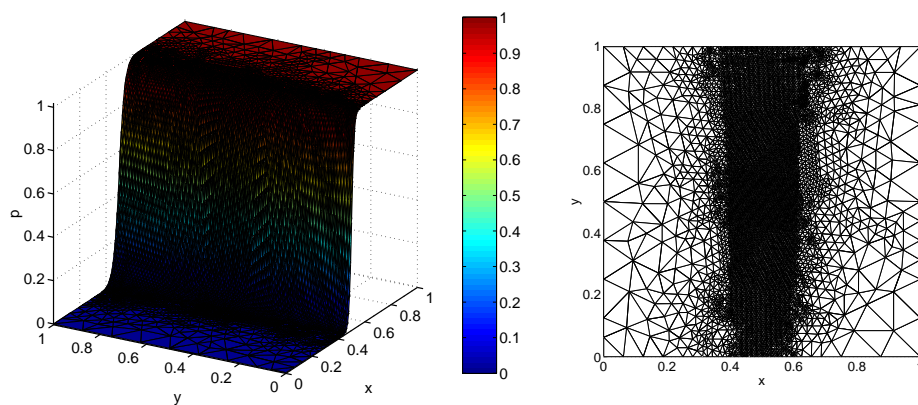


Fig. 6.7 Approximate solution and the corresponding adaptively refined mesh,  $\varepsilon = 10^{-4}$ ,  $a = 0.02$

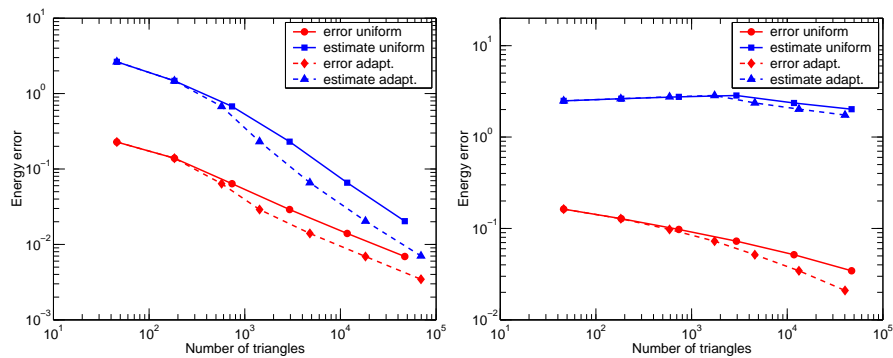
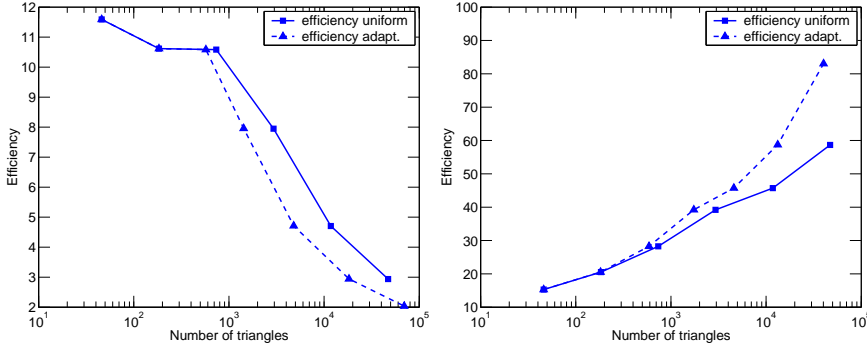


Fig. 6.8 Estimated and actual error against the number of elements in uniformly/adaptively refined meshes for  $\varepsilon = 10^{-2}$ ,  $a = 0.05$  (left) and  $\varepsilon = 10^{-4}$ ,  $a = 0.02$  (right)



**Fig. 6.9** Overall efficiency of the a posteriori error estimates against the number of elements in uniformly/adaptively refined meshes for  $\varepsilon = 10^{-2}$ ,  $a = 0.05$  (left) and  $\varepsilon = 10^{-4}$ ,  $a = 0.02$  (right)

For  $\varepsilon = 10^{-2}$  and  $a = 0.05$  (convection-dominated regime on coarse meshes and diffusion-dominated regime with progressive refinement), still the distribution of the error is predicted very well, cf. Figure 6.6. Note in particular the correct localization of the error away from the center of the shock, as well as the sensitivity of our estimator to the shape of the elements. Next, an example of an adaptively refined mesh and of the corresponding solution for  $\varepsilon = 10^{-4}$  and  $a = 0.02$  is given in Figure 6.7. The problem with keeping the refined mesh uniformly strictly Delaunay reveals as very severe in this case. For this reason, much more vertices are added at a time: the refinement criterion is set to 0.2- and 0.02-times the maximum of the estimators for  $\varepsilon = 10^{-2}$ ,  $a = 0.05$ , and  $\varepsilon = 10^{-4}$ ,  $a = 0.02$ , respectively. The estimated and actual errors for these cases are plotted against the number of elements in uniformly/adaptively refined meshes in Figure 6.8. Again, one can see that we can substantially reduce the number of unknowns necessary to attain the prescribed precision using the derived a posteriori error estimates and adaptively refined grids. Finally, the efficiency plots are given in Figure 6.9. In the first case, the efficiency is almost optimal for finest grids, whereas in the second one, not even the elements in the refined shock region start to leave the convection-dominated regime, a point where the efficiency would start to decrease.

## 7 Auxiliary results

We give in this section two auxiliary results that were needed in the paper.

**Lemma 7.1** *Let  $K \in \mathcal{T}_h$ , let  $\varphi \in H^1(K)$ , and let  $\varphi_K$  be the mean of  $\varphi$  over  $K$  given by  $\varphi_K := (\varphi, 1)_K / |K|$ . Then*

$$\|\varphi - \varphi_K\|_K^2 \leq \min \left\{ C_{P,K} \frac{h_K^2}{c_{S,K}}, \frac{1}{c_{w,r,K}} \right\} \|\varphi\|_K^2.$$

*Proof* The Poincaré inequality (2.1) and the definition of  $\|\cdot\|_K$  by (2.5) imply

$$\|\varphi - \varphi_K\|_K^2 \leq C_{P,K} h_K^2 \|\nabla \varphi\|_K^2 \leq C_{P,K} \frac{h_K^2}{c_{S,K}} \|\varphi\|_K^2.$$

Next, the estimate

$$\|\varphi - \varphi_K\|_K^2 \leq \|\varphi\|_K^2 \leq \frac{1}{c_{w,r,K}} \|\varphi\|_K^2$$

follows from the fact that  $\varphi_K$  is an orthogonal projection of  $\varphi$  onto a constant and from the definition of  $\|\cdot\|_K$  by (2.5).  $\square$

**Lemma 7.2** *Let  $K \in \mathcal{T}_h$ , let  $\varphi \in H^1(K)$ , and let  $\varphi_K$  be the mean of  $\varphi$  over  $K$  given by  $\varphi_K := (\varphi, 1)_K/|K|$  and  $\varphi_\sigma$  the mean of  $\varphi$  over  $\sigma \in \mathcal{E}_K$  given by  $\varphi_\sigma := \langle \varphi, 1 \rangle_\sigma/|\sigma|$ , respectively. Then*

$$\|\varphi_K - \varphi_\sigma\|_\sigma^2 \leq C_{F,K,\sigma} \frac{|\sigma|h_K^2}{|K|c_{S,K}} \|\varphi\|_K^2$$

and

$$\|\varphi_K\|_\sigma^2 \leq \frac{|\sigma|}{|K|c_{w,r,K}} \|\varphi\|_K^2.$$

*Proof* Let us put  $\tilde{\varphi} := \varphi - \varphi_\sigma$  and  $\tilde{\varphi}_K := (\tilde{\varphi}, 1)_K/|K|$ . We now note that  $\tilde{\varphi}_\sigma := \langle \tilde{\varphi}, 1 \rangle_\sigma/|\sigma| = 0$  and that  $\nabla \tilde{\varphi} = \nabla \varphi$ , which allows us to estimate

$$\|\varphi_K - \varphi_\sigma\|_\sigma^2 = \tilde{\varphi}_K^2 |\sigma| \leq \frac{|\sigma|}{|K|} \|\tilde{\varphi}\|_K^2 \leq C_{F,K,\sigma} \frac{|\sigma|h_K^2}{|K|} \|\nabla \varphi\|_K^2 \leq C_{F,K,\sigma} \frac{|\sigma|h_K^2}{|K|c_{S,K}} \|\varphi\|_K^2,$$

using the generalized Friedrichs inequality (2.2) and the definition (2.5) of  $\|\cdot\|_K$ . For the second estimate, we have

$$\|\varphi_K\|_\sigma^2 = |\sigma| \varphi_K^2 \leq \frac{|\sigma|}{|K|} \|\varphi\|_K^2 \leq \frac{|\sigma|}{|K|c_{w,r,K}} \|\varphi\|_K^2,$$

using the definition of  $\|\cdot\|_K$  by (2.5).  $\square$

## References

1. Aavatsmark, I., Barkve, T., Bøe, Ø., Mannseth, T.: Discretization on unstructured grids for inhomogeneous, anisotropic media. I. Derivation of the methods. *SIAM J. Sci. Comput.* **19**(5), 1700–1716 (1998)
2. Achdou, Y., Bernardi, C., Coquel, F.: A priori and a posteriori analysis of finite volume discretizations of Darcy's equations. *Numer. Math.* **96**(1), 17–42 (2003)
3. Adams, R.A.: Sobolev spaces. Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], New York-London (1975). Pure and Applied Mathematics, Vol. 65
4. Afif, M., Bergam, A., Mghazli, Z., Verfürth, R.: A posteriori estimators for the finite volume discretization of an elliptic problem. *Numer. Algorithms* **34**(2-4), 127–136 (2003). International Conference on Numerical Algorithms, Vol. II (Marrakesh, 2001)
5. Agouzal, A., Baranger, J., Maître, J.F., Oudin, F.: Connection between finite volume and mixed finite element methods for a diffusion problem with nonconstant coefficients. Application to a convection diffusion problem. *East-West J. Numer. Math.* **3**(4), 237–254 (1995)
6. Agouzal, A., Oudin, F.: A posteriori error estimator for finite volume methods. *Appl. Math. Comput.* **110**(2-3), 239–250 (2000)
7. Ainsworth, M.: Robust a posteriori error estimation for nonconforming finite element approximation. *SIAM J. Numer. Anal.* **42**(6), 2320–2341 (2005)

8. Angermann, L.: Balanced a posteriori error estimates for finite-volume type discretizations of convection-dominated elliptic problems. *Computing* **55**(4), 305–323 (1995)
9. Angermann, L., Knabner, P., Thiele, K.: An error estimator for a finite volume discretization of density driven flow in porous media. In: *Proceedings of the International Centre for Mathematical Sciences Conference on Grid Adaptation in Computational PDEs: Theory and Applications* (Edinburgh, 1996), vol. 26, pp. 179–191 (1998)
10. Arbogast, T., Dawson, C.N., Keenan, P.T., Wheeler, M.F., Yotov, I.: Enhanced cell-centered finite differences for elliptic equations on general geometry. *SIAM J. Sci. Comput.* **19**(2), 404–425 (1998)
11. Bank, R.E., Weiser, A.: Some a posteriori error estimators for elliptic partial differential equations. *Math. Comp.* **44**(170), 283–301 (1985)
12. Bebendorf, M.: A note on the Poincaré inequality for convex domains. *Z. Anal. Anwendungen* **22**(4), 751–756 (2003)
13. Bergam, A., Mghazli, Z., Verfürth, R.: Estimations a posteriori d'un schéma de volumes finis pour un problème non linéaire. *Numer. Math.* **95**(4), 599–624 (2003)
14. Bernardi, C., Verfürth, R.: Adaptive finite element methods for elliptic equations with non-smooth coefficients. *Numer. Math.* **85**(4), 579–608 (2000)
15. Brezzi, F., Fortin, M.: *Mixed and hybrid finite element methods*, *Springer Series in Computational Mathematics*, vol. 15. Springer-Verlag, New York (1991)
16. Brezzi, F., Lipnikov, K., Shashkov, M.: Convergence of the mimetic finite difference method for diffusion problems on polyhedral meshes. *SIAM J. Numer. Anal.* **43**(5), 1872–1896 (2005)
17. Burman, E., Ern, A.: Continuous interior penalty *hp*-finite element methods for advection and advection-diffusion equations. *Math. Comp.* **76**(259), 1119–1140 (2007)
18. Carstensen, C., Lazarov, R., Tomov, S.: Explicit and averaging a posteriori error estimates for adaptive finite volume methods. *SIAM J. Numer. Anal.* **42**(6), 2496–2521 (2005)
19. Cheddadi, I., Fučík, R., Prieto, M.I., Vohralík, M.: Guaranteed and robust a posteriori error estimates for singularly perturbed reaction–diffusion problems (2008). Preprint R08018, Laboratoire Jacques-Louis Lions, submitted for publication
20. Chénier, E., Eymard, R., Nicolas, X.: A finite volume scheme for the transport of radionuclides in porous media. *Comput. Geosci.* **8**(2), 163–172 (2004). Simulation of Transport Around a Nuclear Waste Disposal Site: The COUPLEX Test Cases
21. Chou, S.H., Kwak, D.Y., Kim, K.Y.: A general framework for constructing and analyzing mixed finite volume methods on quadrilateral grids: the overlapping covolume case. *SIAM J. Numer. Anal.* **39**(4), 1170–1196 (2001)
22. Coudière, Y., Vila, J.P., Villedieu, P.: Convergence rate of a finite volume scheme for a two-dimensional convection–diffusion problem. *M2AN Math. Model. Numer. Anal.* **33**(3), 493–516 (1999)
23. Domelevo, K., Omnes, P.: A finite volume method for the Laplace equation on almost arbitrary two-dimensional grids. *M2AN Math. Model. Numer. Anal.* **39**(6), 1203–1249 (2005)
24. Droniou, J., Eymard, R.: A mixed finite volume scheme for anisotropic diffusion problems on any grid. *Numer. Math.* **105**(1), 35–71 (2006)
25. Eigestad, G.T., Klausen, R.A.: On the convergence of the multi-point flux approximation O-method: numerical experiments for discontinuous permeability. *Numer. Methods Partial Differential Equations* **21**(6), 1079–1098 (2005)
26. El Alaoui, L., Ern, A., Burman, E.: A priori and a posteriori analysis of non-conforming finite elements with face penalty for advection-diffusion equations. *IMA J. Numer. Anal.* **27**(1), 151–171 (2007)
27. Eymard, R., Gallouët, T., Herbin, R.: *Finite volume methods*. In: *Handbook of Numerical Analysis*, Vol. VII, pp. 713–1020. North-Holland, Amsterdam (2000)
28. Eymard, R., Gallouët, T., Herbin, R.: Finite volume approximation of elliptic problems and convergence of an approximate gradient. *Appl. Numer. Math.* **37**(1-2), 31–53 (2001)
29. Eymard, R., Gallouët, T., Herbin, R.: A cell-centred finite-volume approximation for anisotropic diffusion operators on unstructured meshes in any space dimension. *IMA J. Numer. Anal.* **26**(2), 326–353 (2006)
30. Eymard, R., Gallouët, T., Herbin, R.: A new finite volume scheme for anisotropic diffusion problems on general grids: convergence analysis. *C. R. Math. Acad. Sci. Paris* **344**(6), 403–406 (2007)
31. Eymard, R., Hilhorst, D., Vohralík, M.: A combined finite volume–nonconforming/mixed-hybrid finite element scheme for degenerate parabolic problems. *Numer. Math.* **105**(1), 73–131 (2006)

32. Faille, I.: A control volume method to solve an elliptic equation on a two-dimensional irregular mesh. *Comput. Methods Appl. Mech. Engrg.* **100**(2), 275–290 (1992)
33. Karakashian, O.A., Pascal, F.: A posteriori error estimates for a discontinuous Galerkin approximation of second-order elliptic problems. *SIAM J. Numer. Anal.* **41**(6), 2374–2399 (2003)
34. Kim, K.Y.: A posteriori error analysis for locally conservative mixed methods. *Math. Comp.* **76**(257), 43–66 (2007)
35. Lazarov, R., Tomov, S.: A posteriori error estimates for finite volume element approximations of convection–diffusion–reaction equations. *Comput. Geosci.* **6**(3-4), 483–503 (2002). Locally conservative numerical methods for flow in porous media
36. Nicaise, S.: A posteriori error estimations of some cell-centered finite volume methods. *SIAM J. Numer. Anal.* **43**(4), 1481–1503 (2005)
37. Nicaise, S.: A posteriori error estimations of some cell centered finite volume methods for diffusion-convection-reaction problems. *SIAM J. Numer. Anal.* **44**(3), 949–978 (2006)
38. Ohlberger, M.: A posteriori error estimate for finite volume approximations to singularly perturbed nonlinear convection–diffusion equations. *Numer. Math.* **87**(4), 737–761 (2001)
39. Ohlberger, M.: A posteriori error estimates for vertex centered finite volume approximations of convection–diffusion–reaction equations. *M2AN Math. Model. Numer. Anal.* **35**(2), 355–387 (2001)
40. Payne, L.E., Weinberger, H.F.: An optimal Poincaré inequality for convex domains. *Arch. Rational Mech. Anal.* **5**, 286–292 (1960)
41. Rivière, B., Wheeler, M.F., Banas, K.: Part II. Discontinuous Galerkin method applied to single phase flow in porous media. *Comput. Geosci.* **4**(4), 337–349 (2000)
42. Roberts, J.E., Thomas, J.M.: Mixed and hybrid methods. In: *Handbook of Numerical Analysis*, Vol. II, pp. 523–639. North-Holland, Amsterdam (1991)
43. Shewchuk, J.R.: Triangle: engineering a 2d quality mesh generator and delaunay triangulator. In: *Applied Computational Geometry: Towards Geometric Engineering*, pp. 203–222. Lecture Notes in Computer Science, Vol. 1148. Springer-Verlag, Berlin (1996)
44. Beirão da Veiga, L.: A residual based error estimator for the mimetic finite difference method. *Numer. Math.* **108**(3), 387–406 (2008)
45. Verfürth, R.: A review of a posteriori error estimation and adaptive mesh-refinement techniques. Teubner-Wiley, Stuttgart (1996)
46. Verfürth, R.: A posteriori error estimators for convection–diffusion equations. *Numer. Math.* **80**(4), 641–663 (1998)
47. Verfürth, R.: Robust a posteriori error estimates for stationary convection-diffusion equations. *SIAM J. Numer. Anal.* **43**(4), 1766–1782 (2005)
48. Vohralík, M.: On the discrete Poincaré–Friedrichs inequalities for nonconforming approximations of the Sobolev space  $H^1$ . *Numer. Funct. Anal. Optim.* **26**(7–8), 925–952 (2005)
49. Vohralík, M.: Equivalence between lowest-order mixed finite element and multi-point finite volume methods on simplicial meshes. *M2AN Math. Model. Numer. Anal.* **40**(2), 367–391 (2006)
50. Vohralík, M.: A posteriori error estimates for finite volume and mixed finite element discretizations of convection–diffusion–reaction equations. *ESAIM Proc.* **18**, 57–69 (2007)
51. Vohralík, M.: A posteriori error estimates for lowest-order mixed finite element discretizations of convection-diffusion-reaction equations. *SIAM J. Numer. Anal.* **45**(4), 1570–1599 (2007)
52. Vohralík, M.: Guaranteed and fully robust a posteriori error estimates for conforming discretizations of diffusion problems with discontinuous coefficients (2008). Preprint R08009, Laboratoire Jacques-Louis Lions, submitted for publication
53. Vohralík, M.: Two types of guaranteed (and robust) a posteriori estimates for finite volume methods. In: *Finite Volumes for Complex Applications V*, pp. 649–656. ISTE and John Wiley & Sons, London, UK and Hoboken, USA (2008)
54. Vohralík, M.: Unified primal formulation-based a priori and a posteriori error analysis of mixed finite element methods (2008). Submitted for publication
55. Younès, A., Ackerer, P., Chavent, G.: From mixed finite elements to finite volumes for elliptic PDEs in two and three dimensions. *Internat. J. Numer. Methods Engrg.* **59**(3), 365–388 (2004)