



**HAL**  
open science

# A posteriori error estimates for combined finite volume–finite element discretizations of reactive transport equations on nonmatching grids

Danielle Hilhorst, Martin Vohralík

► **To cite this version:**

Danielle Hilhorst, Martin Vohralík. A posteriori error estimates for combined finite volume–finite element discretizations of reactive transport equations on nonmatching grids. *Computer Methods in Applied Mechanics and Engineering*, 2011, 200, pp.5-8 | 597-613. 10.1016/j.cma.2010.08.017. hal-00461154

**HAL Id: hal-00461154**

**<https://hal.science/hal-00461154>**

Submitted on 4 Mar 2010

**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.

# A posteriori error estimates for combined finite volume–finite element discretizations of reactive transport equations on nonmatching grids\*

Danielle Hilhorst<sup>1</sup>, and Martin Vohralík<sup>2</sup>

<sup>1</sup> Laboratoire de Mathématiques, Analyse Numérique et EDP, Université de Paris-Sud et CNRS, Bât. 425, 91 405 Orsay, France. [danielle.hilhorst@math.u-psud.fr](mailto:danielle.hilhorst@math.u-psud.fr)

<sup>2</sup> UPMC Univ. Paris 06, UMR 7598, Laboratoire Jacques-Louis Lions, 75005, Paris, France & CNRS, UMR 7598, Laboratoire Jacques-Louis Lions, 75005, Paris, France. [vohralik@ann.jussieu.fr](mailto:vohralik@ann.jussieu.fr)

## Abstract

We derive fully computable a posteriori error estimates for vertex-centered finite volume-type discretizations of transient convection–diffusion–reaction equations. Our estimates enable actual control of the error measured either in the energy norm or in the energy norm augmented by a dual norm of the skew-symmetric part of the differential operator. Lower bounds, global-in-space but local-in-time, are also derived. These lower bounds are fully robust with respect to convection or reaction dominance and the final simulation time in the augmented norm setting. On the basis of the derived estimates, we propose an adaptive algorithm which enables to automatically achieve a user-given relative precision. Moreover, this algorithm leads to optimal efficiency as it balances the time and space error contributions. As an example, we apply our estimates to the combined finite volume–finite element scheme, including such features as use of mass lumping for the time evolution or reaction terms, of upwind weighting for the convection term, and discretization on nonmatching meshes possibly containing nonconvex and non-star-shaped elements. Numerical experiments illustrate the theoretical developments.

**Key words:** transient convection–diffusion–reaction equation, finite volume method, finite element method, convection dominance, inhomogeneous and anisotropic diffusion, nonmatching grids, a posteriori error estimates, guaranteed upper bound, adaptive mesh refinement

**AMS subject classifications:** 65M15, 65M60, 76M12, 76S05

---

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

# 1 Introduction

We consider in this paper the time-dependent convection–diffusion–reaction problem

$$u_t - \nabla \cdot (\mathbf{S}\nabla u) + \nabla \cdot (u\mathbf{v}) + ru = f \quad \text{in } \Omega \times (0, T), \quad (1.1a)$$

$$u(\cdot, 0) = u_0 \quad \text{in } \Omega, \quad (1.1b)$$

$$u = 0 \quad \text{on } \partial\Omega \times (0, T), \quad (1.1c)$$

where  $\mathbf{S}$  is in general an inhomogeneous and anisotropic (nonconstant full-matrix) diffusion–dispersion tensor,  $\mathbf{v}$  is a (dominating) velocity field,  $r$  is a reaction function,  $f$  is a source term,  $u_0$  prescribes the initial condition,  $\Omega \subset \mathbb{R}^d$ ,  $d = 2, 3$ , is a polygonal (polyhedral for  $d = 3$ ) domain (open, bounded, and connected set), and  $(0, T)$  is a time interval. Our purpose is to derive a posteriori error estimates for conforming locally conservative discretizations of problem (1.1a)–(1.1c). A particular attention will be paid to the combined finite volume–finite element scheme [13] on nonmatching grids.

One of the first works on a posteriori error estimates for finite element discretizations of steady convection–diffusion–reaction problems are those of Angermann [2] and of Eriksson and Johnson [9]. In these works, the overestimation factor depends unfavorably on the ratio between convection and diffusion. Estimates with semi-robust lower bounds in the energy norm and estimates with robust lower bounds in the energy norm augmented by the dual norm of the material derivative were then derived by Verfürth respectively in [23] and [26]. The robustness result has been extended to the unsteady case in [25]. Recently, attention has also been paid to vertex-centered finite volume methods. Let us mention, in the steady convection–diffusion–reaction case and energy norm setting, Lazarov and Tomov [17], Carstensen et al. [7], Nicaise [18], [30], and Ju et al. [16]. Fewer results are known in the unsteady case.  $L^1$ -norm estimates for nonlinear problems are derived by Ohlberger [19], whereas the energy norm setting has been pursued in, e.g., Felcman and Kubera [15] or Amaziane et al. [1]. Typically, the estimate only gives the error upper bound up to an undetermined constant, so that the actual error control is not possible.

The present paper aims at enriching the known results for vertex-centered finite volume-type methods in several directions. Firstly, we derive fully and easily computable estimates, not featuring any undetermined constants. We achieve this by introducing  $\mathbf{H}(\text{div}, \Omega)$ -conforming locally conservative diffusive and convective flux reconstructions, following [29, 31, 10, 11]. Consequently, the estimates allow for actual error control in unsteady convection–diffusion–reaction problems. Secondly, we propose a space-time adaptive algorithm which equilibrates properly the time and space contributions in the sense that they are in actual balance, not weighted by any undetermined quantities. Moreover, this algorithm is proposed to guarantee a user-given relative precision in the simulation. Thirdly, following [25], we also prove a lower error bound. Fourthly, our results are valid for a larger family of conforming locally conservative discretizations, in the framework of the so-called combined finite volume–finite element method, cf. Feistauer et al. [14] and the references therein. Consequently, the analysis includes such features as use of mass lumping for the time evolution or reaction terms. Lastly, we also treat the case of very general nonmatching meshes possibly containing nonconvex and non-star-shaped elements; convergence analysis of the combined finite volume–finite element method on such meshes (for degenerate parabolic convection–diffusion–reaction problems) was carried out in [13].

The paper is organized as follows: in Section 2, the continuous problem is described. Section 3 collects the notation of the discrete setting and introduces the combined finite volume–finite element scheme. In Sections 4 and 5, the a posteriori error estimate and its efficiency is respectively stated and proved. Finally, an adaptive algorithm is presented in Section 6 and numerical experiments in Section 7.

## 2 The continuous setting

We state in this section our assumptions on the data and define a weak solution of problem (1.1a)–(1.1c).

### 2.1 Assumptions on the data

Let the time interval  $(0, T]$  be split such that  $0 = t_0 < \dots < t_n < \dots < t_N = T$  and define  $\tau_n := t_n - t_{n-1}$ ,  $n \in \{1, \dots, N\}$ . On each time interval  $(t_{n-1}, t_n]$ ,  $n \in \{1, \dots, N\}$ , consider a partition  $\mathcal{D}_h^n$  of  $\Omega$  into closed polygons such that  $\bar{\Omega} = \bigcup_{D \in \mathcal{D}_h^n} D$  and such that the intersection of the interiors of two different polygons is empty. Denote the set of its sides (edges for  $d = 2$  and faces for  $d = 3$ ) by  $\mathcal{F}_h^n$ . Also set  $Q_T := \Omega \times (0, T)$ . In the sequel, we will use the notation  $|R|$  for the  $d'$ -dimensional Lebesgue measure of  $R \subset \mathbb{R}^{d'}$ ; at the same time  $|R|$  stands for the cardinality of a set  $R$ .

We suppose that the data of problem (1.1a)–(1.1c) satisfy:

*Assumption (A) (Data)*

(A1)  $\mathbf{S}_{ij} \in L^\infty(Q_T)$ ,  $1 \leq i, j \leq d$ , is a symmetric, bounded, and uniformly positive definite tensor such that for all  $n \in \{1, \dots, N\}$  and for all  $D \in \mathcal{D}_h^n$ ,

$$C_{\mathbf{S},D}^n \mathbf{u} \cdot \mathbf{u} \geq \mathbf{S}(\mathbf{x}, t) \mathbf{u} \cdot \mathbf{u} \geq c_{\mathbf{S},D}^n \mathbf{u} \cdot \mathbf{u}, \quad C_{\mathbf{S},D}^n > 0, c_{\mathbf{S},D}^n > 0 \quad \forall \mathbf{u} \in \mathbb{R}^d$$

for a.e.  $\mathbf{x} \in D$  and a.e.  $t \in (t_{n-1}, t_n]$ ;

(A2)  $\mathbf{v} \in \mathbf{L}^\infty(Q_T)$  such that  $\nabla \cdot \mathbf{v} \in L^2(0, T; L^\infty(\Omega))$  such that for all  $n \in \{1, \dots, N\}$  and for all  $D \in \mathcal{D}_h^n$ ,

$$|\mathbf{v}(\mathbf{x}, t)| \leq C_{\mathbf{v},D}^n, \quad C_{\mathbf{v},D}^n \geq 0 \quad \text{for a.e. } \mathbf{x} \in D \text{ and a.e. } t \in (t_{n-1}, t_n];$$

(A3)  $r \in L^\infty(Q_T)$ ;

(A4) For all  $n \in \{1, \dots, N\}$  and for all  $D \in \mathcal{D}_h^n$ ,

$$\begin{aligned} \frac{1}{2} \nabla \cdot \mathbf{v}(\mathbf{x}, t) + r(\mathbf{x}, t) &\geq c_{\mathbf{v},r,D}^n, \quad \nabla \cdot \mathbf{v}(\mathbf{x}, t) + r(\mathbf{x}, t) \leq C_{\mathbf{v},r,D}^n, \\ r(\mathbf{x}, t) &\leq C_{r,D}^n, \quad c_{\mathbf{v},r,D}^n \geq 0, \quad C_{\mathbf{v},r,D}^n \geq 0, \quad C_{r,D}^n \geq 0 \\ &\text{for a.e. } \mathbf{x} \in D \text{ and a.e. } t \in (t_{n-1}, t_n]; \end{aligned}$$

(A5)  $f \in L^2(Q_T)$ ;

(A6)  $u_0 \in L^\infty(\Omega)$ ;

(A7) if  $c_{\mathbf{v},r,D}^n = 0$ , then  $C_{\mathbf{v},r,D}^n = 0$ .

### 2.2 Continuous problem

Let  $X := L^2(0, T; H_0^1(\Omega))$ ,  $X' = L^2(0, T; H^{-1}(\Omega))$ , and  $Z := \{v \in X; v_t \in X'\}$ . Let  $v \in Z$ . We define  $\mathcal{B}(v) \in X'$  by

$$\langle \mathcal{B}(v), \varphi \rangle := \langle v_t, \varphi \rangle + (\mathbf{S} \nabla v, \nabla \varphi) + (\nabla \cdot (v \mathbf{v}), \varphi) + (rv, \varphi) \quad (2.1)$$

for all  $\varphi \in H_0^1(\Omega)$  and a.e.  $t \in (0, T)$ . Here  $\langle \cdot, \cdot \rangle$  stands for the duality pairing between  $H^{-1}(\Omega)$  and  $H_0^1(\Omega)$  and  $(\cdot, \cdot)$  for the  $L^2(\Omega)$  scalar product. We will also use  $(\cdot, \cdot)_R$  for the  $L^2(R)$  scalar

product on  $R \subset \Omega$  and  $\langle \cdot, \cdot \rangle_R$  for the  $L^2(R)$  scalar product on  $R \subset \mathbb{R}^{d-1}$ . We say that a function  $u$  is a weak solution of problem (1.1a)–(1.1c) if  $u \in Z$ ,  $u(\cdot, 0) = u_0$  in  $H^{-1}(\Omega)$ , and  $u$  satisfies the equality

$$\langle \mathcal{B}(u), \varphi \rangle = (f, \varphi) \quad (2.2)$$

for a.e.  $t \in (0, T)$  and for all  $\varphi \in H_0^1(\Omega)$ .

For  $v \in X$ , we introduce the space-time energy norm

$$\begin{aligned} \|v\|_X^2 &:= \int_0^T \|v(\cdot, t)\|^2 dt, \\ \|v\|^2 &:= \|\mathbf{S}^{\frac{1}{2}} \nabla v\|^2 + \|(\frac{1}{2} \nabla \cdot \mathbf{v} + r)^{\frac{1}{2}} v\|^2, \end{aligned} \quad (2.3)$$

where  $\|\cdot\|$  denotes the  $L^2(\Omega)$  norm.

Assumption (A) then implies

$$\frac{1}{2} \|v(\cdot, T)\|^2 + \|v\|_X^2 \leq \int_0^T \langle \mathcal{B}(v), v \rangle dt + \frac{1}{2} \|v(\cdot, 0)\|^2 \quad \forall v \in Z. \quad (2.4)$$

Under Assumptions (A), there in particular exists a unique solution of (2.2).

Define, for  $v \in Z$ ,

$$b_A(v) := \nabla \cdot (v\mathbf{v}) - \frac{1}{2}(\nabla \cdot \mathbf{v})v, \quad (2.5)$$

and, for all  $\varphi \in H_0^1(\Omega)$  and a.e.  $t \in (0, T)$ ,  $\mathcal{B}_S(v), \mathcal{B}_A(v) \in X'$  by

$$\begin{aligned} \langle \mathcal{B}_S(v), \varphi \rangle &:= (\mathbf{S} \nabla v, \nabla \varphi) + ((\frac{1}{2} \nabla \cdot \mathbf{v} + r)v, \varphi), \\ \langle \mathcal{B}_A(v), \varphi \rangle &:= (b_A(v), \varphi). \end{aligned} \quad (2.6)$$

We will in the sequel also need a space  $Y$ , inspired by that of [25]:  $Y := \{v \in X; \partial_t v + b_A(v) \in X'\}$  that we equip with its graph norm

$$\|v\|_Y^2 := \|v\|_X^2 + \|\partial_t v + b_A(v)\|_{X'}^2. \quad (2.7)$$

Recall that

$$\|\partial_t v + b_A(v)\|_{X'} = \sup_{\varphi \in X; \|\varphi\|_X=1} \int_0^T \langle \partial_t v + b_A(v), \varphi \rangle(t) dt.$$

We will also localize the  $\|\cdot\|_X$  and  $\|\cdot\|_Y$  norms on  $(t_{n-1}, t_n]$ ,  $n \in \{1, \dots, N\}$ , as follows:

$$\begin{aligned} \|v\|_{X(t_{n-1}, t_n)}^2 &:= \int_{t_{n-1}}^{t_n} \|v(\cdot, t)\|^2 dt, \\ \|v\|_{Y(t_{n-1}, t_n)}^2 &:= \|v\|_{X(t_{n-1}, t_n)}^2 + \|\partial_t v + b_A(v)\|_{X'(t_{n-1}, t_n)}^2. \end{aligned}$$

### 3 The combined finite volume–finite element method

We define in this section admissible grids and the combined finite volume–finite element scheme. This section can be skipped if the reader is only interested in a posteriori error estimates for conforming locally conservative methods and not in the details of the present scheme.

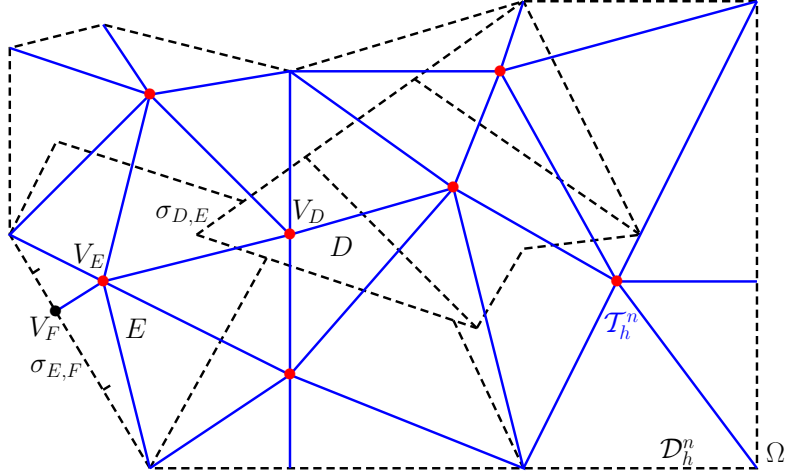


Figure 1: Primal nonmatching grid  $\mathcal{D}_h^n$  (dashed) and dual triangular grid  $\mathcal{T}_h^n$  (solid) with  $D, E \in \mathcal{D}_h^n$ ,  $V_D, V_E \in \mathcal{V}_h^{n,\text{int}}$ ,  $V_F \in \mathcal{V}_h^{n,\text{ext}}$ ,  $\sigma_{D,E} = \partial D \cap \partial E \in \mathcal{F}_h^{n,\text{int}}$ , and  $\sigma_{E,F} \in \mathcal{F}_h^{n,\text{ext}}$

### 3.1 Grids

We call the grid  $\mathcal{D}_h^n$  introduced in Section 2.1 a *primal grid* of  $\Omega$ . In particular we admit non-matching grids, i.e., the case where there exist two different polygons  $D, E \in \mathcal{D}_h^n$  such that their intersection is not an empty set but it is not a common vertex, edge, or face of  $D$  and  $E$ . We also allow for nonconvex elements and elements which are not star-shaped. An example of an admissible primal grid is given in Figure 1 by the dashed line. We suppose that there exists a family of points  $\mathcal{V}_h^{n,\text{int}}$  such that there is one point  $V_D$  in the interior of  $D$  associated with each  $D \in \mathcal{D}_h^n$ . For  $D \in \mathcal{D}_h^n$ , we denote by  $\mathcal{F}_D$  the set of subsets  $\sigma$  of  $\partial D$  such that there exists  $E \in \mathcal{D}_h^n$  such that  $\sigma = \sigma_{D,E} := \partial D \cap \partial E$  has a positive  $(d-1)$ -dimensional Lebesgue measure. If there is a part of  $\partial D$  with a positive  $(d-1)$ -dimensional Lebesgue measure lying on the boundary, then  $\mathcal{F}_D$  contains in addition a union of  $\sigma = \sigma_{D,E} \subset \partial D \cap \partial \Omega$  covering  $\partial D \cap \partial \Omega$  and such that each  $\sigma_{D,E}$  has a positive  $(d-1)$ -dimensional Lebesgue measure and contains exactly one point  $V_E \in \mathcal{V}_h^{n,\text{ext}}$  defined below. We remark that  $\partial D = \sum_{\sigma_{D,E} \in \mathcal{F}_D} \sigma_{D,E}$ , that  $\sigma_{D,E}$  is not necessarily a (whole) side of  $D$ , and that  $\sigma_{D,E}$  not necessarily lies in a hyperplane of  $\mathbb{R}^d$ , see Figure 1. We denote by  $\mathcal{F}_h^{n,\text{int}}$  the union of all  $\sigma_{D,E} = \partial D \cap \partial E$  for some  $D, E \in \mathcal{D}_h^n$ , and by  $\mathcal{F}_h^{n,\text{ext}}$  the union of all  $\sigma \subset \partial \Omega$ .

A *dual grid* of  $\Omega$  is a partition  $\mathcal{T}_h^n$  of  $\Omega$  into closed simplices which satisfies the following properties: (i) The set of points  $\mathcal{V}_h^{n,\text{int}}$  is contained in the set of vertices of  $\mathcal{T}_h^n$ , denoted by  $\mathcal{V}_h^n$ ; (ii) The vertices from  $\mathcal{V}_h^{n,\text{ext}} := \mathcal{V}_h^n \setminus \mathcal{V}_h^{n,\text{int}}$  lie on the boundary of  $\Omega$ ; (iii)  $\mathcal{T}_h^n$  is conforming, i.e. the intersection of two different simplices is either an empty set or their common vertex, edge, or face; (iv)  $\bar{\Omega} = \bigcup_{K \in \mathcal{T}_h^n} K$ . This definition is not unique: we have a choice in connecting the different points  $V_D \in \mathcal{V}_h^{n,\text{int}}$  and also a choice in the definition of the vertices on the boundary. The general intention is to find a triangulation such that the transmissibilities  $\mathbb{S}_{D,E}^n$  defined below by (3.4) were non-negative, since this implies the discrete maximum principle, see Remark 3.2 below. We do not impose any particular simultaneous requirement on  $\mathcal{D}_h^n$  and  $\mathcal{T}_h^n$  except of the assumption that the intersection of the sides of  $\mathcal{D}_h^n$  and  $\mathcal{T}_h^n$  has a zero  $(d-2)$ -dimensional Lebesgue measure (in practice, the sides of  $\mathcal{D}_h^n$  and  $\mathcal{T}_h^n$  can intersect but not coincide). An example of a dual grid to a primal nonmatching grid is given in Figure 1 by the solid line. For  $K \in \mathcal{T}_h^n$ , we denote by  $h_K$  the diameter of  $K$ . For a vertex  $V_D \in \mathcal{V}_h^{n,\text{int}}$ , we denote by  $\mathcal{M}(V_D)$  the set of all vertices  $V_E \in \mathcal{V}_h^n$  such that there exists an edge of the dual grid  $\mathcal{T}_h^n$  between  $V_D$  and  $V_E$ .

### 3.2 The combined scheme

The combined finite volume–finite element scheme is obtained by the discretization of the diffusion term of (1.1a) by means of the piecewise linear conforming finite element method on  $\mathcal{T}_h^n$ , the discretization of the other terms of (1.1a) by means of the cell-centered finite volume method on  $\mathcal{D}_h^n$ , and using an implicit finite difference time stepping.

For the ease of notation, let us denote by  $V_E$  the boundary vertices, even if there are no corresponding  $E \in \mathcal{D}_h^n$ , and let us set  $u_E^n = 0$  for all  $n \in \{1, \dots, N\}$  and all  $V_E \in \mathcal{V}_h^{n, \text{ext}}$  (this corresponds to the discretization of the homogeneous Dirichlet boundary condition). The scheme then reads: find the values  $u_D^n$ ,  $n \in \{1, \dots, N\}$ ,  $D \in \mathcal{D}_h^n$ , such that

$$\begin{aligned} \frac{u_D^n - \bar{u}_D^{n-1}}{\tau_n} |D| - \sum_{V_E \in \mathcal{M}(V_D)} \mathbb{S}_{D,E}^n (u_E^n - u_D^n) + \sum_{\sigma_{D,E} \in \mathcal{F}_D} \langle \mathbf{v}^n \cdot \mathbf{n}, 1 \rangle_{\sigma_{D,E}} \bar{u}_{D,E}^n + r_D^n u_D^n |D| = f_D^n |D| \\ \forall n \in \{1, \dots, N\}, \forall D \in \mathcal{D}_h^n. \end{aligned} \quad (3.1)$$

The value  $\bar{u}_D^0$  for  $D \in \mathcal{D}_h^1$  is given by  $u_D^0 := (u_0, 1)_D / |D|$  and it corresponds to the discretization of the initial condition. When  $\mathcal{D}_h^{n-1}$  and  $\mathcal{D}_h^n$  (and  $\mathcal{T}_h^{n-1}$  and  $\mathcal{T}_h^n$ ) coincide,  $\bar{u}_D^{n-1} = u_D^{n-1}$ , the approximate value from the previous time step. In the opposite case, we define  $\bar{u}_D^{n-1}$  by the value of  $u_h^{n-1}$  (cf. its definition (3.8) below) at the point  $V_D$  corresponding to  $D \in \mathcal{D}_h^n$ . Let

$$\begin{aligned} f^n(\mathbf{x}) &:= \frac{1}{\tau_n} \int_{t_{n-1}}^{t_n} f(\mathbf{x}, t) dt & \mathbf{x} \in \Omega, n \in \{1, \dots, N\}, \\ r^n(\mathbf{x}) &:= \frac{1}{\tau_n} \int_{t_{n-1}}^{t_n} r(\mathbf{x}, t) dt & \mathbf{x} \in \Omega, n \in \{1, \dots, N\}, \\ \mathbf{v}^n(\mathbf{x}) &:= \frac{1}{\tau_n} \int_{t_{n-1}}^{t_n} \mathbf{v}(\mathbf{x}, t) dt & \mathbf{x} \in \Omega, n \in \{1, \dots, N\}, \end{aligned}$$

where  $\mathbf{n}$  is the unit normal vector of  $\partial D$ , outward to  $D$ . In the above definition, we have also set

$$f_D^n := \frac{(f^n, 1)}{|D|} \quad n \in \{1, \dots, N\}, D \in \mathcal{D}_h^n, \quad (3.2)$$

$$r_D^n := \frac{(r^n, 1)}{|D|} \quad n \in \{1, \dots, N\}, D \in \mathcal{D}_h^n. \quad (3.3)$$

We next define

$$\mathbf{S}^n(\mathbf{x}) := \frac{1}{\tau_n} \int_{t_{n-1}}^{t_n} \mathbf{S}(\mathbf{x}, t) dt \quad \mathbf{x} \in \Omega, n \in \{1, \dots, N\}.$$

The elements  $\mathbb{S}_{D,E}^n$ ,  $V_D \in \mathcal{V}_h^{n, \text{int}}$ ,  $V_E \in \mathcal{V}_h^n$ , of the finite element diffusion matrix are then given by

$$\mathbb{S}_{D,E}^n := - \sum_{K \in \mathcal{T}_h^n} (\mathbf{S}_h^n \nabla \psi_{V_E}, \nabla \psi_{V_D})_K \quad n \in \{1, \dots, N\}. \quad (3.4)$$

Here  $\psi_{V_D}$  is the finite element basis function associated with  $V_D \in \mathcal{V}_h^n$ : the function linear on each  $K \in \mathcal{T}_h^n$ , equal to 1 at  $V_D$ , and equal to zero at all other  $V_E \in \mathcal{V}_h^n$ . There are two basic choices for the space-time discrete diffusion–dispersion tensor  $\mathbf{S}_h^n$ : Firstly, we may set

$$\mathbf{S}_h^n|_K := \frac{1}{|K|} (\mathbf{S}^n, 1)_K \quad n \in \{1, \dots, N\}, K \in \mathcal{T}_h^n. \quad (3.5)$$

Note that since  $\nabla\psi_{V_D}, \nabla\psi_{V_E}$  are constant on all  $K \in \mathcal{T}_h^n$ , the above definition is equivalent to directly putting  $\mathbf{S}_h^n = \mathbf{S}^n$  in (3.4). An alternative choice is to define  $\mathbf{S}_h^n$  by

$$\mathbf{S}_h^n|_K := \left( \frac{1}{|K|}([\mathbf{S}^n]^{-1}, 1)_K \right)^{-1} \quad n \in \{1, \dots, N\}, K \in \mathcal{T}_h^n. \quad (3.6)$$

In fact, the members  $\mathbb{S}_{D,E}^n$  for  $V_E \in \mathcal{V}_h^{n,\text{ext}}$  do not occur in the definition of the scheme (3.1), but it shows convenient to define them. Finally, we define the value  $\overline{u_{D,E}^n}$ , used for the evaluation of the convective flux across  $\sigma_{D,E} \in \mathcal{F}_h^n$ ,  $n \in \{1, \dots, N\}$ ,  $D \in \mathcal{D}_h^n$ , by

$$\overline{u_{D,E}^n} := \begin{cases} u_D^n + \alpha_{D,E}^n(u_E^n - u_D^n) & \text{if } \langle \mathbf{v}^n \cdot \mathbf{n}, 1 \rangle_{\sigma_{D,E}} \geq 0 \\ u_E^n + \alpha_{D,E}^n(u_D^n - u_E^n) & \text{if } \langle \mathbf{v}^n \cdot \mathbf{n}, 1 \rangle_{\sigma_{D,E}} < 0 \end{cases}. \quad (3.7)$$

Here  $\alpha_{D,E}^n$  is the coefficient of the amount of upstream weighting which is defined by

$$\alpha_{D,E}^n := \begin{cases} \frac{\max\{\min\{\mathbb{S}_{D,E}^n, \frac{1}{2}|\langle \mathbf{v}^n \cdot \mathbf{n}, 1 \rangle_{\sigma_{D,E}}|\}, 0\}}{|\langle \mathbf{v}^n \cdot \mathbf{n}, 1 \rangle_{\sigma_{D,E}}|} & \text{if } \langle \mathbf{v}^n \cdot \mathbf{n}, 1 \rangle_{\sigma_{D,E}} \neq 0 \text{ and } \sigma_{D,E} \in \mathcal{F}_h^{n,\text{int}} \\ & \text{or if } \sigma_{D,E} \in \mathcal{F}_h^{n,\text{ext}} \text{ and } \langle \mathbf{v}^n \cdot \mathbf{n}, 1 \rangle_{\sigma_{D,E}} > 0 \\ 0 & \text{if } \langle \mathbf{v}^n \cdot \mathbf{n}, 1 \rangle_{\sigma_{D,E}} = 0 \text{ or if } \sigma_{D,E} \in \mathcal{F}_h^{n,\text{ext}} \\ & \text{and } \langle \mathbf{v}^n \cdot \mathbf{n}, 1 \rangle_{\sigma_{D,E}} < 0 \end{cases}.$$

The values  $u_D^n$  define, for each  $n \in \{1, \dots, N\}$ , the usual piecewise linear (on  $\mathcal{T}_h^n$ ) finite element approximation  $u_h^n$  by

$$u_h^n := \sum_{D \in \mathcal{D}_h^n} u_D^n \psi_{V_D}. \quad (3.8)$$

By the approximate solution of problem (1.1a)–(1.1c) by means of the combined finite volume–finite element scheme (3.1), we understand a function  $u_{h,\tau}$  such that

$$\begin{aligned} u_{h,\tau}(\cdot, t_n) &:= u_h^n \quad n \in \{0, \dots, N\}, \\ u_{h,\tau}(\mathbf{x}, \cdot) &\text{ is linear on } [t_{n-1}, t_n], n \in \{1, \dots, N\}, \text{ for all } \mathbf{x} \in \Omega. \end{aligned} \quad (3.9)$$

**Remark 3.1** (Convergence). *The convergence of  $u_{h,\tau}$ , when the meshes  $\mathcal{D}_h^n, \mathcal{T}_h^n$  do not change with time, towards the weak solution of problem (1.1a)–(1.1c) given by (2.2) has been proved (for a more general nonlinear degenerate case) in [13].*

**Remark 3.2** (Discrete maximum principle). *If  $\mathbb{S}_{D,E}^n \geq 0$  for all  $n \in \{1, \dots, N\}$ ,  $D \in \mathcal{D}_h^n$ , and  $V_E \in \mathcal{M}(V_D)$  and under suitable conditions on the data, the discrete maximum principle for the combined scheme holds, see [13, Theorem 4.11].*

**Remark 3.3** (Arithmetic versus harmonic averaging). *We remark that the choice (3.5) for  $\mathbf{S}_h^n$  corresponds to the arithmetic average of the diffusion–dispersion tensor  $\mathbf{S}$  in space, whereas the choice (3.6) corresponds to the harmonic one.*

**Remark 3.4** (Relation to the vertex-centered finite volume method). *Let the meshes  $\mathcal{T}_h^n$  be given first and let the meshes  $\mathcal{D}_h^n$  be constructed using the face, edge, and element barycentres. Then the combined finite volume–finite element scheme (3.1) is equivalent to the classical vertex-centered finite volume method (cf., e.g., [19] and the references therein), where mass lumping has been used in the time evolution and reaction terms. This follows easily using [3, Lemma 3] for the diffusion term. Hence, all the results of the present paper apply to the vertex-centered finite volume method as a particular case.*



## 4 A posteriori error estimate and its efficiency

We summarize in this section our a posteriori estimate on the error between the weak solution  $u$  and the approximate solution  $u_{h,\tau}$  and its efficiency.

### 4.1 Some useful inequalities

Let  $D \subset \mathbb{R}^d$  be a polygon. The Poincaré inequality states that

$$\|\varphi - \varphi_D\|_D^2 \leq C_{P,D} h_D^2 \|\nabla \varphi\|_D^2 \quad \forall \varphi \in H^1(D), \quad (4.1)$$

where  $\varphi_D$  is the mean of  $\varphi$  over  $D$  given by  $\varphi_D := (\varphi, 1)_D / |D|$  and where  $h_D$  is the diameter of  $D$ . The constant  $C_{P,D}$  can be evaluated as  $1/\pi^2$  if  $D$  is convex, cf. [20, 4], and only depends on the geometry of  $D$  if  $D$  is nonconvex, cf. [12, Lemma 10.4].

Let  $D \subset \mathbb{R}^d$  be a polygon such that  $|\partial\Omega \cap \partial D| \neq 0$ . Then the Friedrichs inequality states that

$$\|\varphi\|_D^2 \leq C_{F,D,\partial\Omega} h_D^2 \|\nabla \varphi\|_D^2 \quad \forall \varphi \in H^1(D) \text{ such that } \varphi = 0 \text{ on } \partial\Omega \cap \partial D. \quad (4.2)$$

As long as  $\partial\Omega$  is such that there exists a vector  $\mathbf{b} \in \mathbb{R}^d$  such that for almost all  $\mathbf{x} \in D$ , the first intersection of  $\mathcal{B}_{\mathbf{x}}$  and  $\partial D$  lies in  $\partial\Omega$ , where  $\mathcal{B}_{\mathbf{x}}$  is the straight semi-line defined by the origin  $\mathbf{x}$  and the vector  $\mathbf{b}$ ,  $C_{F,D,\partial\Omega} = 1$ , cf. [28, Remark 5.8]. To evaluate  $C_{F,D,\partial\Omega}$  in the general case is more complicated but it still can be done, cf. [28, Remark 5.9] or [6, Section 3]. As a particular case, when  $D = \Omega$ , the Friedrichs inequality reads

$$\|\varphi\|^2 \leq C_F h_\Omega^2 \|\nabla \varphi\|^2 \quad \forall \varphi \in H_0^1(\Omega), \quad (4.3)$$

and  $C_F = 1$ .

Finally, for a simplex  $K \subset \mathbb{R}^d$ , the trace inequality states that

$$\|\varphi\|_\sigma^2 \leq C_{t,K,\sigma} (h_K^{-1} \|\varphi\|_K^2 + \|\varphi\|_K \|\nabla \varphi\|_K) \quad \forall \varphi \in H^1(K). \quad (4.4)$$

It follows from [22, Lemma 3.12] that the constant  $C_{t,K,\sigma}$  can be evaluated as  $|\sigma| h_K / |K|$ .

### 4.2 Diffusive and convective flux reconstructions

Let  $\mathcal{D}_h^n$  and  $\mathcal{T}_h^n$  be the primal and dual meshes as introduced in Section 3.1. For  $n \in \{2, \dots, N\}$ , let  $\mathcal{S}_h^{n-1,n}$  be an additional tertial simplicial mesh, a conforming refinement of all  $\mathcal{D}_h^{n-1}$ ,  $\mathcal{D}_h^n$ ,  $\mathcal{T}_h^{n-1}$ , and  $\mathcal{T}_h^n$ , and for  $n = 1$ , let  $\mathcal{S}_h^{0,1}$  be a simplicial conforming refinement of both  $\mathcal{D}_h^1$  and  $\mathcal{T}_h^1$ . An example of the three types of meshes  $\mathcal{D}_h^n$ ,  $\mathcal{T}_h^n$ , and  $\mathcal{S}_h^{n-1,n}$  is given in Figure 2 below. We denote by  $\mathcal{S}_D$  the partition of  $D \in \mathcal{D}_h^n$  by the elements of  $\mathcal{S}_h^{n-1,n}$ , by  $K \in \mathcal{S}_h^{n-1,n}$  a generic element of  $\mathcal{S}_h^{n-1,n}$ , by  $\mathcal{G}_K$  its sides, and by  $\mathcal{G}_K^{\text{int}}$  its sides not contained in  $\partial\Omega$ . By  $\mathcal{G}_D$ , we denote all the sides of the partition  $\mathcal{S}_D$  for a given  $D$ , except those included in the boundary  $\partial\Omega$ . We will also use the obvious notation  $\mathcal{S}_h^{n-1,n,\text{int}}$  and  $\mathcal{S}_h^{n-1,n,\text{ext}}$ .

Let  $X_h$  denote the space of scalar functions continuous in space and piecewise linear on the meshes  $\mathcal{S}_h^{n-1,n}$  and piecewise linear and continuous in time. A function  $v_h \in X_h$  is, as in (3.9), uniquely prescribed by the functions  $v_h^n$  defined on the meshes  $\mathcal{S}_h^{n-1,n}$ , the values of  $v_h$  at the times  $t_n$ ,  $n \in \{1, \dots, N\}$ , and by the function  $v_h^0$  defined on  $\mathcal{S}_h^{0,1}$ , the values of  $v_h$  at the time 0. Typically, the values  $u_D^n$ ,  $D \in \mathcal{D}_h^n$ ,  $n \in \{1, \dots, N\}$ , define through (3.8)–(3.9) the function  $u_{h,\tau} \in X_h$ . Let  $\mathbf{V}_h$  denote the space of vector functions belonging to the Raviart–Thomas–Nédélec spaces of lowest order (cf. [5]) on the meshes  $\mathcal{S}_h^{n-1,n}$  in space and piecewise constant in time. A

function  $\mathbf{u}_h \in \mathbf{V}_h$  is uniquely prescribed by the functions  $\mathbf{u}_h^n$  defined on the meshes  $\mathcal{S}_h^{n-1,n}$ , the values of  $\mathbf{u}_h$  on the time intervals  $(t_{n-1}, t_n]$ ,  $n \in \{1, \dots, N\}$ . Recall that the distinctive feature is that the normal component of the functions  $\mathbf{u}_h$  is continuous ([5]); the values  $\mathbf{u}_h^n \cdot \mathbf{n}|_\sigma = \langle \mathbf{u}_h^n \cdot \mathbf{n}, 1 \rangle_\sigma$  on the sides  $\sigma$  of the meshes  $\mathcal{S}_h^{n-1,n}$ ,  $n \in \{1, \dots, N\}$ , represent the degrees of freedom in the space  $\mathbf{V}_h$ . Following [29, 31, 11], our a posteriori error estimates rely on the concept of the *diffusive flux reconstruction*  $\boldsymbol{\theta}_h \in \mathbf{V}_h$ ; following [10], we also introduce a *convective flux reconstruction*  $\mathbf{w}_h \in \mathbf{V}_h$ .

In order to obtain the upper bounds of Section 4.3 below, we merely suppose that there exist values  $u_D^n$  and  $\bar{u}_D^{n-1}$  and functions  $\boldsymbol{\theta}_h, \mathbf{w}_h \in \mathbf{V}_h$  such that

$$\frac{u_D^n + \bar{u}_D^{n-1}}{\tau_n} |D| + \langle \boldsymbol{\theta}_h^n \cdot \mathbf{n}, 1 \rangle_{\partial D} + \langle \mathbf{w}_h^n \cdot \mathbf{n}, 1 \rangle_{\partial D} + r_D^n u_D^n |D| + q_D^n |D| = f_D^n |D| \quad (4.5)$$

$$\forall n \in \{1, \dots, N\}, \forall D \in \mathcal{D}_h^n.$$

Here  $r_D^n$  and  $f_D^n$  are given by (3.2)–(3.3). Relation (4.5) is a local conservation property: the term  $(u_D^n + \bar{u}_D^{n-1})|D|/\tau_n$  represents the time accumulation in the volume  $D$ ,  $\langle \boldsymbol{\theta}_h^n \cdot \mathbf{n}, 1 \rangle_{\partial D}$  the diffusive flux over  $\partial D$ ,  $\langle \mathbf{w}_h^n \cdot \mathbf{n}, 1 \rangle_{\partial D}$  the convective flux over  $\partial D$ , and  $r_D^n u_D^n |D|$  and  $f_D^n |D|$ , respectively, the reaction and sources accumulation in the volume  $D$ . Finally, the term  $q_D^n |D|$  represents a quadrature error when the local conservation is not satisfied exactly for the given scheme. A vast class of locally conservative methods enters the framework of (4.5). In particular, the combined finite volume–finite element scheme (3.1) can be written in the form (4.5) with

$$\langle \boldsymbol{\theta}_h^n \cdot \mathbf{n}, 1 \rangle_\sigma := -\langle \mathbf{S}_h^n \nabla u_h^n \cdot \mathbf{n}, 1 \rangle_\sigma \quad \sigma \subset \partial D, D \in \mathcal{D}_h^n \quad (4.6)$$

and

$$\langle \mathbf{w}_h^n \cdot \mathbf{n}, 1 \rangle_{\sigma_{D,E}} := \langle \mathbf{v}^n \cdot \mathbf{n}, 1 \rangle_{\sigma_{D,E}} \overline{u_{D,E}^n} \quad \sigma_{D,E} \in \mathcal{F}_D, D \in \mathcal{D}_h^n; \quad (4.7)$$

here, by the notation  $\sigma$ , we mean a side of  $\mathcal{G}_K$  for some  $K \in \mathcal{S}_D$  which is such that  $\sigma \subset \partial D$ . As the diffusive fluxes  $\langle \boldsymbol{\theta}_h \cdot \mathbf{n}, 1 \rangle_\sigma$  given by (4.6) do not necessarily coincide completely with those stemming from the combined scheme (3.1), we use the additional factor  $q_D^n |D|$  to fix the balance. In numerical experiments in Section 7 below, the values  $q_D^n |D|$  were negligible. Equations (4.6) and (4.7) do not prescribe  $\boldsymbol{\theta}_h$  and  $\mathbf{w}_h$  uniquely; they are sufficient, altogether with (4.12) below, for the present theoretical analysis. Practically, there exist several possibilities for fixing the remaining degrees of freedom of  $\boldsymbol{\theta}_h$  and  $\mathbf{w}_h$ , following [31].

### 4.3 A posteriori error estimate

In this section, we suppose that  $u_{h,\tau} \in X_h$  and  $\boldsymbol{\theta}_h, \mathbf{w}_h \in \mathbf{V}_h$  are arbitrary but such that (4.5) holds. A particular example is the combined finite volume–finite element scheme (3.1) with the construction (4.6)–(4.7). Let  $n \in \{1, \dots, N\}$  and  $D \in \mathcal{D}_h^n$ . We set

$$m_D^n := \min \left\{ C_{P,D}^{\frac{1}{2}} h_D (c_{\mathbf{S},D}^n)^{-\frac{1}{2}}, (c_{\mathbf{v},r,D}^n)^{-\frac{1}{2}} \right\}$$

and define the *residual estimator*  $\eta_{R,D}^n$  by

$$\eta_{R,D}^n := m_D^n \left\| f_D^n - \frac{u_D^n - \bar{u}_D^{n-1}}{\tau_n} - \nabla \cdot \boldsymbol{\theta}_h^n - \nabla \cdot \mathbf{w}_h^n - r_D^n u_D^n - q_D^n \right\|_D. \quad (4.8)$$

The *diffusive and convective flux estimator*  $\eta_{DCF,D}^n$  is given by

$$\eta_{DCF,D}^n(t) := \left\| \mathbf{S}^{\frac{1}{2}} \nabla u_{h,\tau} + \mathbf{S}^{-\frac{1}{2}} \boldsymbol{\theta}_h^n - \mathbf{S}^{-\frac{1}{2}} u_{h,\tau} \mathbf{v} + \mathbf{S}^{-\frac{1}{2}} \mathbf{w}_h^n \right\|_D(t). \quad (4.9)$$

Let  $K \in \mathcal{S}_h^{n-1,n}$ . We then define the cutoff factors

$$\begin{aligned} m_K^n &:= \min \left\{ C_{\text{P},\text{F},K}^{\frac{1}{2}} h_K (c_{\mathbf{S},K}^n)^{-\frac{1}{2}}, (c_{\mathbf{v},r,K}^n)^{-\frac{1}{2}} \right\}, \\ \tilde{m}_K^n &:= \min \left\{ (C_{\text{P},\text{F},K} + C_{\text{P},\text{F},K}^{\frac{1}{2}}) h_K (c_{\mathbf{S},K}^n)^{-1}, h_K^{-1} (c_{\mathbf{v},r,K}^n)^{-1} + \frac{1}{2} (c_{\mathbf{v},r,K}^n)^{-\frac{1}{2}} (c_{\mathbf{S},K}^n)^{-\frac{1}{2}} \right\}, \end{aligned}$$

where  $C_{\text{P},\text{F},K} := C_{\text{P},K}$  for  $K \in \mathcal{S}_h^{n-1,n,\text{int}}$  and  $C_{\text{P},\text{F},K} := C_{\text{F},K,\partial\Omega}$  for  $K \in \mathcal{S}_h^{n-1,n,\text{ext}}$ . We will need below also the *diffusive flux estimator*  $\eta_{\text{DF},D}^n$

$$\begin{aligned} \eta_{\text{DF},D}^n(t) &:= \left\{ \sum_{K \in \mathcal{S}_D} \left( m_K^n \|\nabla \cdot (\mathbf{S}\nabla u_{h,\tau} + \boldsymbol{\theta}_h^n)\|_K \right. \right. \\ &\quad \left. \left. + (\tilde{m}_K^n)^{\frac{1}{2}} \sum_{\sigma \in \mathcal{G}_K^{\text{int}}} (C_{t,K,\sigma})^{\frac{1}{2}} \|(\mathbf{S}\nabla u_{h,\tau} + \boldsymbol{\theta}_h^n) \cdot \mathbf{n}\|_{\sigma} \right)^2 \right\}^{\frac{1}{2}}(t). \end{aligned}$$

Lastly, set

$$\bar{m}^n := \min \left\{ C_{\text{F}}^{\frac{1}{2}} h_{\Omega} (c_{\mathbf{S},\Omega}^n)^{-\frac{1}{2}}, (c_{\mathbf{v},r,\Omega}^n)^{-\frac{1}{2}} \right\}.$$

Then the *data oscillation–quadrature estimator*  $\eta_{\text{DOQ},D}^n$  is given by

$$\eta_{\text{DOQ},D}^n(t) := \bar{m}^n \left\| f - f_D^n - (u_{h,\tau})_t + \frac{u_D^n - \bar{u}_D^{n-1}}{\tau_n} - r u_{h,\tau} + r_D^n u_D^n + q_D^n \right\|_D(t). \quad (4.10)$$

We finally denote

$$\begin{aligned} (\eta^{(1),n})^2 &:= \int_{t_{n-1}}^{t_n} \left( \left\{ \sum_{D \in \mathcal{D}_h^n} (\eta_{\text{R},D}^n + \eta_{\text{DCF},D}^n(t))^2 \right\}^{\frac{1}{2}} + \left\{ \sum_{D \in \mathcal{D}_h^n} (\eta_{\text{DOQ},D}^n(t))^2 \right\}^{\frac{1}{2}} \right)^2 dt, \\ (\eta^{(2),n})^2 &:= \int_{t_{n-1}}^{t_n} \left( \left\{ \sum_{D \in \mathcal{D}_h^n} (\eta_{\text{R},D}^n)^2 \right\}^{\frac{1}{2}} + \left\{ \sum_{D \in \mathcal{D}_h^n} (\eta_{\text{DF},D}^n(t))^2 \right\}^{\frac{1}{2}} \right. \\ &\quad \left. + \sup_{\varphi \in H_0^1(\Omega); \|\varphi\|=1} (\nabla \cdot (u_{h,\tau} \mathbf{v} - \mathbf{w}_h^n), \varphi) + \left\{ \sum_{D \in \mathcal{D}_h^n} (\eta_{\text{DOQ},D}^n(t))^2 \right\}^{\frac{1}{2}} \right)^2 dt \end{aligned}$$

and

$$\eta^n := \min \{ \eta^{(1),n}, \eta^{(2),n} \}, \quad (4.11a)$$

$$\eta := \left\{ \sum_{n=1}^N (\eta^n)^2 \right\}^{\frac{1}{2}}. \quad (4.11b)$$

**Remark 4.1** (Data oscillation–quadrature estimators). *In a posteriori estimates for finite element, vertex-centered finite volume, or mixed finite element methods, cf. [25, 11], data oscillation estimators of the form  $\bar{m}^n \|f - f_D^n\|$  appear. The other terms of the present estimators  $\eta_{\text{DOQ},D}^n$  are related to the non-variational general nature of (4.5) (or of (3.1)). They clearly represent a numerical quadrature error, which is illustrated below on numerical experiments. If the vertex-centered finite volume method, cf., e.g. [31], was used instead of (4.5), then  $\eta_{\text{DOQ},D}^n$  would reduce to the usual data oscillation estimator  $\bar{m}^n \|f - f_D^n\|$ .*

**Remark 4.2** (Estimators  $\eta^{(1),n}$  and  $\eta^{(2),n}$ ). *The estimator  $\eta^{(1),n}$  is fully computable and is based on the constant-free estimator (4.9). The estimator  $\eta^{(1),n}$  is, however, not robust with respect to the convection dominance, as noted in [8, 27]. Using the estimator  $\eta^{(2),n}$ , on the contrary, robust lower bound will be proven below following [25]. The estimator  $\eta^{(2),n}$  is, however, not computable. Its computable upper bound (still ensuring robustness) can be established following [25, Section 8], for the price of a solution of an auxiliary stationary reaction–diffusion problem on each time step. Such an approach is not pursued here in view of its complexity.*

Using the above definitions, we will prove in Section 5.1 below the following a posteriori error estimate:

**Theorem 4.3** (Energy norm a posteriori error estimate). *Let  $u$  be the weak solution of problem (1.1a)–(1.1c) given by (2.2) and let  $u_{h,\tau} \in X_h$  and  $\boldsymbol{\theta}_h, \mathbf{w}_h \in \mathbf{V}_h$  be arbitrary but such that (4.5) holds. Then*

$$\|(u - u_{h,\tau})(\cdot, T)\|^2 + \|u - u_{h,\tau}\|_X^2 \leq \eta^2 + \|u_0 - u_{h,\tau}(\cdot, 0)\|^2.$$

**Theorem 4.4** (Augmented norm a posteriori error estimate). *Let the assumptions of Theorem 4.3 hold. Then*

$$3\|(u - u_{h,\tau})(\cdot, T)\|^2 + \|u - u_{h,\tau}\|_Y^2 \leq 5\eta^2 + 3\|u_0 - u_{h,\tau}(\cdot, 0)\|^2.$$

#### 4.4 Distinguishing the space and time error contributions

The estimate  $\eta^{(1),n}$  of (4.11a) is fully computable and allows to control the overall error. There is, however, no distinction between the space and time errors. Using the triangle inequality, we have, as in [11], for the example of  $\eta_{\text{DCF},D}^n(t)$ ,

$$\begin{aligned} \eta_{\text{DCF},D}^n(t) &\leq \left\| \mathbf{S}^{\frac{1}{2}} \nabla u_h^n + \mathbf{S}^{-\frac{1}{2}} \boldsymbol{\theta}_h^n - \mathbf{S}^{-\frac{1}{2}} u_h^n \mathbf{v}^n + \mathbf{S}^{-\frac{1}{2}} \mathbf{w}_h^n \right\|_D(t) \\ &\quad + \left\| \mathbf{S}^{\frac{1}{2}} \nabla u_{h,\tau} - \mathbf{S}^{\frac{1}{2}} \nabla u_h^n - \mathbf{S}^{-\frac{1}{2}} u_{h,\tau} \mathbf{v} + \mathbf{S}^{-\frac{1}{2}} u_h^n \mathbf{v}^n \right\|_D(t). \end{aligned}$$

Note that whenever  $\mathbf{S}$  is constant in time on  $(t_{n-1}, t_n]$ , the first of the above terms gets independent of time; it clearly corresponds to the spatial part of the error. The second of the above terms then corresponds to the temporal part of the error, so that we have the usual space-time contributions division as in [21, 25]. Note as well that when both  $\mathbf{S}$  and  $\mathbf{v}$  are constant in time on  $(t_{n-1}, t_n]$ , we get, using the definition of  $u_{h,\tau}$  by (3.9) as in [24, Equation (6.5)], the following easily computable upper bound on  $\left\{ \int_{t_{n-1}}^{t_n} (\eta_{\text{DCF},D}^n(t))^2 dt \right\}^{\frac{1}{2}}$ :

$$\begin{aligned} \left\{ \int_{t_{n-1}}^{t_n} (\eta_{\text{DCF},D}^n(t))^2 dt \right\}^{\frac{1}{2}} &\leq \tau_n^{\frac{1}{2}} \left\| (\mathbf{S}^n)^{\frac{1}{2}} \nabla u_h^n + (\mathbf{S}^n)^{-\frac{1}{2}} \boldsymbol{\theta}_h^n - (\mathbf{S}^n)^{-\frac{1}{2}} u_h^n \mathbf{v}^n + (\mathbf{S}^n)^{-\frac{1}{2}} \mathbf{w}_h^n \right\|_D \\ &\quad + \left( \frac{\tau_n}{3} \right)^{\frac{1}{2}} \left\| (\mathbf{S}^n)^{\frac{1}{2}} \nabla (u_h^{n-1} - u_h^n) - (\mathbf{S}^n)^{-\frac{1}{2}} \mathbf{v}^n (u_h^{n-1} - u_h^n) \right\|_D. \end{aligned}$$

We now proceed similarly for the other estimators, for general  $\mathbf{S}$  and  $\mathbf{v}$ . We obtain

$$\eta^n \leq \eta_{\text{sp}}^n + \eta_{\text{tm}}^n$$

with

$$\begin{aligned} \eta_{\text{sp}}^n &:= \left\{ \tau_n \sum_{D \in \mathcal{D}_h^n} (\eta_{\text{R},D}^n)^2 \right\}^{\frac{1}{2}} + \left\{ \int_{t_{n-1}}^{t_n} \sum_{D \in \mathcal{D}_h^n} \left\| \mathbf{S}^{\frac{1}{2}} \nabla u_h^n + \mathbf{S}^{-\frac{1}{2}} \boldsymbol{\theta}_h^n - \mathbf{S}^{-\frac{1}{2}} u_h^n \mathbf{v}^n + \mathbf{S}^{-\frac{1}{2}} \mathbf{w}_h^n \right\|_D^2(t) dt \right\}^{\frac{1}{2}} \\ &+ \left\{ \int_{t_{n-1}}^{t_n} \sum_{D \in \mathcal{D}_h^n} (\bar{m}^n)^2 \left\| f - f_h - (u_{h,\tau})_t + \frac{u_D^n - \bar{u}_D^{n-1}}{\tau_n} - r^n u_h^n + r_D^n u_D^n + q_D^n \right\|_D^2(t) dt \right\}^{\frac{1}{2}} \end{aligned}$$

and

$$\begin{aligned} \eta_{\text{tm}}^n &:= \left\{ \int_{t_{n-1}}^{t_n} \sum_{D \in \mathcal{D}_h^n} \left\| \mathbf{S}^{\frac{1}{2}} \nabla u_{h,\tau} - \mathbf{S}^{\frac{1}{2}} \nabla u_h^n - \mathbf{S}^{-\frac{1}{2}} u_{h,\tau} \mathbf{v} + \mathbf{S}^{-\frac{1}{2}} u_h^n \mathbf{v}^n \right\|_D^2(t) dt \right\}^{\frac{1}{2}} \\ &+ \left\{ \int_{t_{n-1}}^{t_n} \sum_{D \in \mathcal{D}_h^n} (\bar{m}^n)^2 \left\| f_h - f_D^n - r u_{h,\tau} + r^n u_h^n \right\|_D^2(t) dt \right\}^{\frac{1}{2}}. \end{aligned}$$

Consequently, we have the following corollary:

**Corollary 4.5** (Energy norm a posteriori error estimate distinguishing the space and time errors). *Let the assumptions of Theorem 4.3 hold. Then*

$$\|(u - u_{h,\tau})(\cdot, T)\|^2 + \|u - u_{h,\tau}\|_X^2 \leq \sum_{n=1}^N (\eta_{\text{sp}}^n + \eta_{\text{tm}}^n)^2 + \|(u - u_{h,\tau})(\cdot, 0)\|^2.$$

## 4.5 Efficiency of the estimate

Henceforth, let  $a \lesssim b$  denote  $a \leq Cb$  for a generic positive constant  $C$ , not necessarily the same at each occurrence, dependent on the shape regularity of the meshes  $\mathcal{S}_h^{n-1,n}$  and  $d$  but independent of the space domain  $\Omega$ , final time  $T$ , and the parameters  $\mathbf{S}$ ,  $\mathbf{v}$ ,  $r$ ,  $f$ , and  $u_0$ ;  $C$  can only depend on the local oscillation of these parameters. Similarly, the notation  $a \approx b$  will be employed when simultaneously  $a \lesssim b$  and  $b \lesssim a$ . From now on, we also omit denoting explicitly by  $(t)$  the dependence of the estimators on the time. We will suppose below that the following *approximation property* of the diffusive flux reconstruction  $\boldsymbol{\theta}_h$  is satisfied:

$$\|\mathbf{S}_h^n \nabla u_h^n + \boldsymbol{\theta}_h^n\|_D \lesssim \left\{ \sum_{\sigma \in \mathcal{G}_D} h_\sigma \|\llbracket \mathbf{S}_h^n \nabla u_h^n \cdot \mathbf{n} \rrbracket_\sigma\|_2^2 \right\}^{\frac{1}{2}} \quad \forall n \in \{1, \dots, N\}, \forall D \in \mathcal{D}_h^n. \quad (4.12)$$

Here  $\llbracket \cdot \rrbracket$  denotes the jump across a side. This property can be shown, in dependence on the construction of  $\boldsymbol{\theta}_h^n$  from  $\nabla u_h^n$ , as in [31]. We also introduce some additional notation. Let  $f_h, r_h \in X_h$  and  $\mathbf{v}_h \in \mathbf{V}_h$  be space-time discrete approximations of the source term, of the reaction function, and of the velocity field  $\mathbf{v}$ , respectively. Let  $n \in \{1, \dots, N\}$  and  $D \in \mathcal{D}_h^n$ . We define the *data oscillation estimator*  $\eta_{\text{DO},D}^n$  by

$$\begin{aligned} \eta_{\text{DO},D}^n &:= m_D^n \|f_D^n - f_h^n\|_D + m_D^n \|u_h^n (r_h^n - r_D^n)\|_D \\ &+ \left\{ \sum_{K \in \mathcal{S}_D} \left( (m_K^n)^2 \|\nabla \cdot (\mathbf{S} - \mathbf{S}_h^n) \nabla u_{h,\tau}\|_K^2 + \tilde{m}_K^n \sum_{\sigma \in \mathcal{G}_K^{\text{int}}} \|(\mathbf{S} - \mathbf{S}_h^n) \nabla u_{h,\tau} \cdot \mathbf{n}\|_\sigma^2 \right) \right\}^{\frac{1}{2}} \\ &+ \left\{ \sum_{K \in \mathcal{S}_D} \left( (m_K^n)^2 \|\nabla \cdot ((\mathbf{v} - \mathbf{v}_h^n) u_{h,\tau})\|_K^2 + \tilde{m}_K^n \sum_{\sigma \in \mathcal{G}_K^{\text{int}}} \|(\mathbf{v} - \mathbf{v}_h^n) u_{h,\tau} \cdot \mathbf{n}\|_\sigma^2 \right) \right\}^{\frac{1}{2}} \end{aligned}$$

and the quadrature estimator  $\eta_{Q,D}^n$  by

$$\eta_{Q,D}^n := m_D^n \left\| \frac{u_h^n - u_h^{n-1}}{\tau_n} - \frac{u_D^n - \bar{u}_D^{n-1}}{\tau_n} + r_D^n (u_h^n - u_D^n) - q_D^n \right\|_D + \left\{ \sum_{K \in \mathcal{S}_D} \left( (m_K^n)^2 \|\nabla \cdot (\mathbf{v}_h^n u_h^n - \mathbf{w}_h^n)\|_K^2 + \tilde{m}_K^n \sum_{\sigma \in \mathcal{G}_K^{\text{int}}} \|(\mathbf{v}_h^n u_h^n - \mathbf{w}_h^n) \cdot \mathbf{n}\|_\sigma^2 \right) \right\}^{\frac{1}{2}}.$$

Finally, for the sake of simplicity, we make in this section an additional assumption that  $\nabla \cdot \mathbf{v} = 0$ . In this case, the augmented norm (2.7) is closely related to that of [25], so that we can use the results of [25] for the lower bound proof. We then have the following theorem:

**Theorem 4.6** (Efficiency of the a posteriori error estimators in the augmented norm). *Let  $\nabla \cdot \mathbf{v} = 0$ , let  $u$  be the weak solution of problem (1.1a)–(1.1c) given by (2.2) and let  $u_{h,\tau} \in X_h$  and  $\boldsymbol{\theta}_h, \mathbf{w}_h \in \mathbf{V}_h$  be arbitrary but such that (4.12) holds. Let  $n \in \{1, \dots, N\}$ , let  $\mathcal{S}_h^{n-1,n}$  be shape-regular, and let, for all  $D \in \mathcal{D}_h^n$ ,  $h_D \approx h_K$  for all  $K \in \mathcal{S}_D$ . Then*

$$(\eta^n)^2 \lesssim \|u - u_{h,\tau}\|_{Y(t_{n-1}, t_n)}^2 + \int_{t_{n-1}}^{t_n} \sum_{D \in \mathcal{D}_h^n} \{(\eta_{\text{DOQ},D}^n)^2 + (\eta_{\text{DO},D}^n)^2 + (\eta_{Q,D}^n)^2\} dt + \|f - f^n - (r - r^n)u_{h,\tau}\|_{X'(t_{n-1}, t_n)}^2.$$

## 5 Proof of the a posteriori error estimate and of its efficiency

We shall prove in this section the a posteriori error estimates stated in Theorems 4.3 and 4.4, as well as their efficiency discussed in Theorem 4.6.

### 5.1 Proof of the a posteriori error estimate

**Lemma 5.1** (Bound by the residual). *There holds*

$$\|(u - u_{h,\tau})(\cdot, T)\|^2 + \|u - u_{h,\tau}\|_X^2 \leq \left( \sup_{\varphi \in X, \|\varphi\|_X=1} \int_0^T \langle \mathcal{B}(u - u_{h,\tau}), \varphi \rangle dt \right)^2 + \|(u - u_{h,\tau})(\cdot, 0)\|^2.$$

*Proof.* Using the obvious relation

$$\int_0^T \langle \mathcal{B}(v), v \rangle dt \leq \sup_{\varphi \in X, \|\varphi\|_X=1} \int_0^T \langle \mathcal{B}(v), \varphi \rangle dt \|v\|_X \leq \frac{1}{2} \left( \sup_{\varphi \in X, \|\varphi\|_X=1} \int_0^T \langle \mathcal{B}(v), \varphi \rangle dt \right)^2 + \frac{1}{2} \|v\|_X^2,$$

valid for any  $v \in Z$ , the assertion follows from (2.4) and the fact that  $(u - u_{h,\tau}) \in Z$ .  $\square$

Let from now on,  $\varphi \in X$  with  $\|\varphi\|_X = 1$  be fixed. Then, on the time interval  $(t_{n-1}, t_n]$ ,  $n \in \{1, \dots, N\}$ , we have the following results.

**Lemma 5.2** (Equivalent form of the residual). *There holds*

$$\langle \mathcal{B}(u - u_{h,\tau}), \varphi \rangle = T_R(\varphi) + T_{\text{DCF}}(\varphi) + T_Q(\varphi)$$

with

$$\begin{aligned}
T_R(\varphi) &:= \sum_{D \in \mathcal{D}_h^n} \left( f_D^n - \frac{u_D^n - \bar{u}_D^{n-1}}{\tau_n} - \nabla \cdot \boldsymbol{\theta}_h^n - \nabla \cdot \mathbf{w}_h^n - r_D^n u_D^n - q_D^n, \varphi - \varphi_D \right)_D, \\
T_{\text{DCF}}(\varphi) &:= -(\mathbf{S} \nabla u_{h,\tau} + \boldsymbol{\theta}_h^n - u_{h,\tau} \mathbf{v} + \mathbf{w}_h^n, \nabla \varphi), \\
T_Q(\varphi) &:= \sum_{D \in \mathcal{D}_h^n} \left( f - f_D^n - (u_{h,\tau})_t + \frac{u_D^n - \bar{u}_D^{n-1}}{\tau_n} - r u_{h,\tau} + r_D^n u_D^n + q_D^n, \varphi \right)_D,
\end{aligned}$$

where  $\varphi_D := (\varphi, 1)_D / |D|$ ,  $D \in \mathcal{D}_h^n$ .

*Proof.* We first use the definition of the weak solution (2.2) and that of  $\mathcal{B}$  (2.1) to infer that

$$\langle \mathcal{B}(u - u_{h,\tau}), \varphi \rangle = (f, \varphi) - ((u_{h,\tau})_t, \varphi) - (\mathbf{S} \nabla u_{h,\tau}, \nabla \varphi) - (\nabla \cdot (u_{h,\tau} \mathbf{v}), \varphi) - (r u_{h,\tau}, \varphi).$$

We then add and subtract  $(\nabla \cdot \boldsymbol{\theta}_h^n, \varphi)$  and  $(\nabla \cdot \mathbf{w}_h^n, \varphi)$  and use the Green theorem to see that

$$\begin{aligned}
\langle \mathcal{B}(u - u_{h,\tau}), \varphi \rangle &= (f, \varphi) - ((u_{h,\tau})_t, \varphi) - (\nabla \cdot \boldsymbol{\theta}_h^n, \varphi) - (\nabla \cdot \mathbf{w}_h^n, \varphi) \\
&\quad - (\mathbf{S} \nabla u_{h,\tau} + \boldsymbol{\theta}_h^n - u_{h,\tau} \mathbf{v} + \mathbf{w}_h^n, \nabla \varphi) - (r u_{h,\tau}, \varphi) \\
&= T_R^*(\varphi) + T_{\text{DCF}}(\varphi) + T_Q(\varphi),
\end{aligned}$$

where  $T_R^*(\varphi)$  is as  $T_R(\varphi)$  with, however, the second argument replaced by  $\varphi$ . Finally, it suffices to multiply (4.5) by  $\varphi_D$ , to use therein the Green theorem for the terms involving  $\boldsymbol{\theta}_h^n$  and  $\mathbf{w}_h^n$ , and add it to  $T_R^*(\varphi)$  for each  $D \in \mathcal{D}_h^n$  to arrive at the assertion of the lemma.  $\square$

**Lemma 5.3** (Residual estimate). *There holds*

$$T_R(\varphi) \leq \sum_{D \in \mathcal{D}_h^n} \eta_{R,D}^n \|\varphi\|_D.$$

*Proof.* The assertion follows immediately from an application of the Cauchy–Schwarz inequality, the Poincaré inequality (4.1), and the definition (2.3) of the energy norm  $\|\cdot\|$ , cf. [30, Lemmas 4.3 and 7.1].  $\square$

**Lemma 5.4** (Diffusive and convective flux estimates). *There holds*

$$T_{\text{DCF}}(\varphi) \leq \sum_{D \in \mathcal{D}_h^n} \eta_{\text{DCF},D}^n \|\varphi\|_D, \quad (5.1a)$$

$$T_{\text{DCF}}(\varphi) \leq \left\{ \sum_{D \in \mathcal{D}_h^n} (\eta_{\text{DF},D}^n)^2 \right\}^{\frac{1}{2}} \|\varphi\| + \sup_{\psi \in H_0^1(\Omega); \|\psi\|=1} (\nabla \cdot (u_{h,\tau} \mathbf{v} - \mathbf{w}_h^n), \psi) \|\varphi\|. \quad (5.1b)$$

*Proof.* Estimate (5.1a) follows easily by the Cauchy–Schwarz inequality and (2.3). The first part of estimate (5.1b), bounding the term  $-(\mathbf{S} \nabla u_{h,\tau} + \boldsymbol{\theta}_h^n, \nabla \varphi)$ , is established as in [8, proof of Theorem 4.4]. Two following simple modifications are necessary: firstly, the Friedrichs inequality (4.2) has to be used for  $K \in \mathcal{S}_h^{n-1, n, \text{ext}}$  instead of the Poincaré inequality (4.1) in [8, proof of Theorem 4.4]; consequently, only a sum over  $\sigma \in \mathcal{G}_K^{\text{int}}$ , and not a sum over  $\sigma \in \mathcal{G}_K$ , appears in the definition of  $\eta_{\text{DF},D}^n$ . Secondly, the scaling by  $\mathbf{S}$  has to be added, leading to the factors  $c_{\mathbf{S},K}^n$  in the definition of  $m_K^n$  and  $\tilde{m}_K^n$ . Finally, by the Green theorem, we easily have

$$(u_{h,\tau} \mathbf{v} - \mathbf{w}_h^n, \nabla \varphi) = -(\nabla \cdot (u_{h,\tau} \mathbf{v} - \mathbf{w}_h^n), \varphi) \leq \sup_{\psi \in H_0^1(\Omega); \|\psi\|=1} (\nabla \cdot (u_{h,\tau} \mathbf{v} - \mathbf{w}_h^n), \psi) \|\varphi\|.$$

$\square$

**Lemma 5.5** (Data oscillation–quadrature estimate). *There holds*

$$T_Q(\varphi) \leq \left\{ \sum_{D \in \mathcal{D}_h^n} (\eta_{\text{DOQ},D}^n)^2 \right\}^{\frac{1}{2}} \|\varphi\|.$$

*Proof.* Denote by  $v_D$  the argument in the left part of  $(\cdot, \cdot)_D$  in  $T_Q(\varphi)$ . Then, on the one hand,

$$T_Q(\varphi) \leq \sum_{D \in \mathcal{D}_h^n} \|v_D\|_D \|\varphi\|_D \leq \sum_{D \in \mathcal{D}_h^n} (c_{\mathbf{v},r,D}^n)^{-\frac{1}{2}} \|v_D\|_D \|\varphi\|_D \leq (c_{\mathbf{v},r,\Omega}^n)^{-\frac{1}{2}} \left\{ \sum_{D \in \mathcal{D}_h^n} \|v_D\|_D^2 \right\}^{\frac{1}{2}} \|\varphi\|$$

by the Cauchy–Schwarz inequality and definition (2.3) of the energy norm. On the other hand,

$$T_Q(\varphi) \leq \left\{ \sum_{D \in \mathcal{D}_h^n} \|v_D\|_D^2 \right\}^{\frac{1}{2}} \|\varphi\| \leq C_F^{\frac{1}{2}} h_\Omega (c_{\mathbf{S},\Omega}^n)^{-\frac{1}{2}} \left\{ \sum_{D \in \mathcal{D}_h^n} \|v_D\|_D^2 \right\}^{\frac{1}{2}} \|\varphi\|$$

by the Cauchy–Schwarz inequality, the Friedrichs inequality (4.3), and definition (2.3) of the energy norm.  $\square$

*Proof of Theorem 4.3.* Let  $\varphi \in X$ ,  $\|\varphi\|_X = 1$ , be fixed. Lemmas 5.2–5.5 and the Cauchy–Schwarz inequality imply

$$\langle \mathcal{B}(u - u_{h,\tau}), \varphi \rangle dt \leq \left( \left\{ \sum_{D \in \mathcal{D}_h^n} (\eta_{\mathbf{R},D}^n + \eta_{\text{DCF},D}^n)^2 \right\}^{\frac{1}{2}} + \left\{ \sum_{D \in \mathcal{D}_h^n} (\eta_{\text{DOQ},D}^n)^2 \right\}^{\frac{1}{2}} \right) \|\varphi\|,$$

using the bound (5.1a) on  $T_{\text{DCF}}(\varphi)$ . Similarly, using the bound (5.1b) on  $T_{\text{DCF}}(\varphi)$ , we arrive at

$$\begin{aligned} \langle \mathcal{B}(u - u_{h,\tau}), \varphi \rangle dt &\leq \left( \left\{ \sum_{D \in \mathcal{D}_h^n} (\eta_{\mathbf{R},D}^n)^2 \right\}^{\frac{1}{2}} + \left\{ \sum_{D \in \mathcal{D}_h^n} (\eta_{\text{DF},D}^n)^2 \right\}^{\frac{1}{2}} \right. \\ &\quad \left. + \sup_{\psi \in H_0^1(\Omega); \|\psi\|=1} (\nabla \cdot (u_{h,\tau} \mathbf{v} - \mathbf{w}_h^n), \psi) + \left\{ \sum_{D \in \mathcal{D}_h^n} (\eta_{\text{DOQ},D}^n)^2 \right\}^{\frac{1}{2}} \right) \|\varphi\|. \end{aligned}$$

Another use of the Cauchy–Schwarz inequality gives

$$\int_0^T \langle \mathcal{B}(u - u_{h,\tau}), \varphi \rangle dt \leq \sum_{n=1}^N \left( \min\{\eta^{(1),n}, \eta^{(2),n}\} \left\{ \int_{t_{n-1}}^{t_n} \|\varphi\|^2 dt \right\}^{\frac{1}{2}} \right) \leq \eta \|\varphi\|_X = \eta \quad (5.2)$$

and using Lemma 5.1 concludes the proof.  $\square$

*Proof of Theorem 4.4.* Let us first bound the term  $\|\partial_t(u - u_{h,\tau}) + b_A(u - u_{h,\tau})\|_{X'}$ . Following [10, 11], using definitions (2.1) and (2.6), estimate (5.2), and the Cauchy–Schwarz inequality, we readily have

$$\begin{aligned} &\sup_{\varphi \in X; \|\varphi\|_X=1} \int_0^T \langle \partial_t(u - u_{h,\tau}) + b_A(u - u_{h,\tau}), \varphi \rangle(t) dt \\ &= \sup_{\varphi \in X; \|\varphi\|_X=1} \int_0^T \{ \langle \mathcal{B}(u - u_{h,\tau}), \varphi \rangle - \langle \mathcal{B}_S(u - u_{h,\tau}), \varphi \rangle \}(t) dt \\ &\leq \eta + \|u - u_{h,\tau}\|_X. \end{aligned}$$



Consequently, using Theorem 4.3,

$$\|\partial_t(u - u_{h,\tau}) + b_A(u - u_{h,\tau})\|_{X'}^2 \leq 4\eta^2 + 2\|(u - u_{h,\tau})(\cdot, 0)\|^2 - 2\|(u - u_{h,\tau})(\cdot, T)\|^2.$$

The assertion of the theorem follows from definition (2.7) of the augmented norm.  $\square$

## 5.2 Proof of the efficiency of the estimate

Let  $n \in \{1, \dots, N\}$  and  $K \in \mathcal{S}_h^{n-1, n}$  be fixed. Denote by

$$\begin{aligned} \eta_{\mathbb{R}, K}^n &:= m_K^n \left\| f_h^n - \frac{u_h^n - u_h^{n-1}}{\tau_n} + \nabla \cdot (\mathbf{S}_h^n \nabla u_h^n) - \nabla \cdot (\mathbf{v}_h^n u_h^n) - r_h^n u_h^n \right\|_K, \\ \eta_{\mathbb{J}, K}^n &:= (m_K^n)^{\frac{1}{2}} (c_{\mathbf{S}, K}^n)^{-\frac{1}{4}} \sum_{\sigma \in \mathcal{G}_K^{\text{int}}} \|[\mathbf{S}_h^n \nabla u_h^n \cdot \mathbf{n}]\|_{\sigma}, \end{aligned}$$

respectively, the usual element and face residual estimators, cf. [25]. In order to prove Theorem 4.6, we show in the following lemmas that our a posteriori error estimate represents a lower bound for the residual a posteriori error estimate of [25, Lemma 7.1]. Consequently, we will be able to establish the lower bound by estimate (7.2) of this reference.

Let  $n \in \{1, \dots, N\}$  and  $D \in \mathcal{D}_h^n$  be fixed. We then have the following results:

**Lemma 5.6** (Upper bound on  $\eta_{\mathbb{R}, D}^n$ ). *There holds*

$$\eta_{\mathbb{R}, D}^n \lesssim \left\{ \sum_{K \in \mathcal{S}_D} (\eta_{\mathbb{R}, K}^n)^2 \right\}^{\frac{1}{2}} + \left\{ \sum_{K \in \mathcal{S}_D} (\eta_{\mathbb{J}, K}^n)^2 \right\}^{\frac{1}{2}} + \eta_{\text{DO}, D}^n + \eta_{\text{Q}, D}^n.$$

*Proof.* The triangle inequality yields

$$\begin{aligned} \eta_{\mathbb{R}, D}^n &\leq m_D^n \left\| f_h^n - \frac{u_h^n - u_h^{n-1}}{\tau_n} + \nabla \cdot (\mathbf{S}_h^n \nabla u_h^n) - \nabla \cdot (\mathbf{v}_h^n u_h^n) - r_h^n u_h^n \right\|_D + m_D^n \|f_D^n - f_h^n\|_D \\ &\quad + m_D^n \left\| \frac{u_h^n - u_h^{n-1}}{\tau_n} - \frac{u_D^n - \bar{u}_D^{n-1}}{\tau_n} + r_D^n (u_h^n - u_D^n) - q_D^n \right\|_D \\ &\quad + m_D^n \|u_h^n (r_h^n - r_D^n)\|_D + m_D^n \|\nabla \cdot (\mathbf{S}_h^n \nabla u_h^n + \boldsymbol{\theta}_h^n)\|_D + m_D^n \|\nabla \cdot (\mathbf{v}_h^n u_h^n - \mathbf{w}_h^n)\|_D. \end{aligned}$$

The fact that  $\mathbf{S}_h^n \nabla u_h^n + \boldsymbol{\theta}_h^n$  is a piecewise polynomial in  $D$ , the inverse inequality, the fact that  $h_D \approx h_K$  for all  $K \in \mathcal{S}_D$ , and the approximation property (4.12) lead to

$$\|\nabla \cdot (\mathbf{S}_h^n \nabla u_h^n + \boldsymbol{\theta}_h^n)\|_D \lesssim h_D^{-1} \|\mathbf{S}_h^n \nabla u_h^n + \boldsymbol{\theta}_h^n\|_D \lesssim h_D^{-1} \left\{ \sum_{\sigma \in \mathcal{G}_D} h_{\sigma} \|[\mathbf{S}_h^n \nabla u_h^n \cdot \mathbf{n}]\|_{\sigma}^2 \right\}^{\frac{1}{2}}.$$

From here, we easily get

$$m_D \|\nabla \cdot (\mathbf{S}_h^n \nabla u_h^n + \boldsymbol{\theta}_h^n)\|_D \lesssim \left\{ \sum_{K \in \mathcal{S}_D} (\eta_{\mathbb{J}, K}^n)^2 \right\}^{\frac{1}{2}},$$

using the previous arguments and the bound  $m_K^n h_K^{-\frac{1}{2}} \lesssim (m_K^n)^{\frac{1}{2}} (c_{\mathbf{S}, K}^n)^{-\frac{1}{4}}$  (equivalent to the obvious  $(m_K^n)^{\frac{1}{2}} h_K^{-\frac{1}{2}} (c_{\mathbf{S}, K}^n)^{\frac{1}{4}} \lesssim 1$ ) in each  $K \in \mathcal{S}_D$ . Combining the above results, the assertion follows.  $\square$

**Lemma 5.7** (Upper bound on  $\eta_{\text{DF},D}^n$ ). *There holds*

$$\eta_{\text{DF},D}^n \lesssim \left\{ \sum_{K \in \mathcal{S}_D} (\eta_{\text{J},K}^n)^2 \right\}^{\frac{1}{2}} + \left\{ \sum_{K \in \mathcal{S}_D} m_K^n (c_{\mathbf{S},K}^n)^{-\frac{1}{2}} h_K^{-1} \|\mathbf{S}_h^n \nabla(u_{h,\tau} - u_h^n)\|_K^2 \right\}^{\frac{1}{2}} + \eta_{\text{DO},D}^n.$$

*Proof.* Let  $K \in \mathcal{S}_D$ . We have, separating the data oscillation, time evolution, and diffusive flux approximation parts,

$$\begin{aligned} & m_K^n \|\nabla \cdot (\mathbf{S} \nabla u_{h,\tau} + \boldsymbol{\theta}_h^n)\|_K + (\tilde{m}_K^n)^{\frac{1}{2}} \sum_{\sigma \in \mathcal{G}_K^{\text{int}}} (C_{t,K,\sigma})^{\frac{1}{2}} \|(\mathbf{S} \nabla u_{h,\tau} + \boldsymbol{\theta}_h^n) \cdot \mathbf{n}\|_{\sigma} \\ & \lesssim m_K^n \|\nabla \cdot (\mathbf{S} \nabla u_{h,\tau} + \boldsymbol{\theta}_h^n)\|_K + (\tilde{m}_K^n)^{\frac{1}{2}} \sum_{\sigma \in \mathcal{G}_K^{\text{int}}} \|(\mathbf{S} \nabla u_{h,\tau} + \boldsymbol{\theta}_h^n) \cdot \mathbf{n}\|_{\sigma} \\ & \leq m_K^n (\|\nabla \cdot (\mathbf{S} - \mathbf{S}_h^n) \nabla u_{h,\tau}\|_K + \|\nabla \cdot \mathbf{S}_h^n \nabla(u_{h,\tau} - u_h^n)\|_K + \|\nabla \cdot (\mathbf{S}_h^n \nabla u_h^n + \boldsymbol{\theta}_h^n)\|_K) \\ & \quad + (\tilde{m}_K^n)^{\frac{1}{2}} \sum_{\sigma \in \mathcal{G}_K^{\text{int}}} (\|(\mathbf{S} - \mathbf{S}_h^n) \nabla u_{h,\tau} \cdot \mathbf{n}\|_{\sigma} + \|\mathbf{S}_h^n \nabla(u_{h,\tau} - u_h^n) \cdot \mathbf{n}\|_{\sigma} + \|(\mathbf{S}_h^n \nabla u_h^n + \boldsymbol{\theta}_h^n) \cdot \mathbf{n}\|_{\sigma}) \\ & \lesssim m_K^n \|\nabla \cdot (\mathbf{S} - \mathbf{S}_h^n) \nabla u_{h,\tau}\|_K + (\tilde{m}_K^n)^{\frac{1}{2}} \sum_{\sigma \in \mathcal{G}_K^{\text{int}}} \|(\mathbf{S} - \mathbf{S}_h^n) \nabla u_{h,\tau} \cdot \mathbf{n}\|_{\sigma} \\ & \quad + (m_K^n)^{\frac{1}{2}} (c_{\mathbf{S},K}^n)^{-\frac{1}{4}} h_K^{-\frac{1}{2}} (\|\mathbf{S}_h^n \nabla(u_{h,\tau} - u_h^n)\|_K + \|\mathbf{S}_h^n \nabla u_h^n + \boldsymbol{\theta}_h^n\|_K), \end{aligned}$$

using the inverse inequalities

$$\begin{aligned} \|\nabla \cdot \mathbf{q}_h\|_K & \lesssim h_K^{-1} \|\mathbf{q}_h\|_K, \\ \|\mathbf{q}_h \cdot \mathbf{n}\|_{\sigma} & \lesssim h_K^{-\frac{1}{2}} \|\mathbf{q}_h\|_K \end{aligned}$$

for  $\mathbf{q}_h = \mathbf{S}_h^n \nabla(u_{h,\tau} - u_h^n)$  and  $\mathbf{q}_h = \mathbf{S}_h^n \nabla u_h^n + \boldsymbol{\theta}_h^n$  and the estimate

$$m_K^n h_K^{-\frac{1}{2}} + (\tilde{m}_K^n)^{\frac{1}{2}} \lesssim (m_K^n)^{\frac{1}{2}} (c_{\mathbf{S},K}^n)^{-\frac{1}{4}}$$

following from [8, Proof of Theorem 5.1]. To conclude, it now suffices to use the Cauchy–Schwarz inequality, approximation property (4.12), and the definitions of  $\eta_{\text{J},K}^n$  and  $\eta_{\text{DO},D}^n$ .  $\square$

Let  $n \in \{1, \dots, N\}$  and  $K \in \mathcal{S}_h^{n-1,n}$  be fixed. We then have:

**Lemma 5.8** (Equivalent form of  $\int_{t_{n-1}}^{t_n} \|\mathbf{S}_h^n \nabla(u_{h,\tau} - u_h^n)\|_K^2 dt$ ). *There holds*

$$\int_{t_{n-1}}^{t_n} \|\mathbf{S}_h^n \nabla(u_{h,\tau} - u_h^n)\|_K^2 dt = \frac{\tau_n}{3} \|\mathbf{S}_h^n \nabla(u_h^n - u_h^{n-1})\|_K^2.$$

*Proof.* Follows from the definition of  $u_{h,\tau}$  by (3.9), cf. [24, Equation (6.5)].  $\square$

Let  $n \in \{1, \dots, N\}$  be fixed. We then have:

**Lemma 5.9** (Upper bound on the convective time evolution term). *There holds*

$$\begin{aligned} \int_{t_{n-1}}^{t_n} \left( \sup_{\varphi \in H_0^1(\Omega); \|\varphi\|=1} (\nabla \cdot (u_{h,\tau} \mathbf{v} - \mathbf{w}_h^n), \varphi) \right)^2 dt & \lesssim \tau_n \sup_{\varphi \in H_0^1(\Omega); \|\varphi\|=1} (\nabla \cdot ((u_h^n - u_h^{n-1}) \mathbf{v}_h^n), \varphi)^2 \\ & \quad + \int_{t_{n-1}}^{t_n} \sum_{D \in \mathcal{D}_h^n} \{(\eta_{\text{DO},D}^n)^2 + (\eta_{\text{Q},D}^n)^2\} dt. \end{aligned}$$

*Proof.* Let  $t \in (t_{n-1}, t_n]$  and  $\varphi \in H_0^1(\Omega)$ . Then, separating the data oscillation, time evolution, and quadrature parts,

$$\begin{aligned} (\nabla \cdot (u_{h,\tau} \mathbf{v} - \mathbf{w}_h^n), \varphi) &= (\nabla \cdot (u_{h,\tau}(\mathbf{v} - \mathbf{v}_h^n)), \varphi) + (\nabla \cdot ((u_{h,\tau} - u_h^n) \mathbf{v}_h^n), \varphi) + (\nabla \cdot (u_h^n \mathbf{v}_h^n - \mathbf{w}_h^n), \varphi) \\ &= -(u_{h,\tau}(\mathbf{v} - \mathbf{v}_h^n), \nabla \varphi) + (\nabla \cdot ((u_{h,\tau} - u_h^n) \mathbf{v}_h^n), \varphi) - (u_h^n \mathbf{v}_h^n - \mathbf{w}_h^n, \nabla \varphi) \end{aligned}$$

Bounding the first and last terms of the above expression as in (5.1b) in Lemma 5.4 and integrating the middle term in time as in Lemma 5.8, the assertion of the lemma follows.  $\square$

**Lemma 5.10** (Upper bound on  $(\eta^n)^2$  using the usual residual estimators). *There holds*

$$\begin{aligned} (\eta^n)^2 &\leq (\eta^{(2),n})^2 \lesssim \tau_n \sum_{D \in \mathcal{D}_h^n} \sum_{K \in \mathcal{S}_D} \{(\eta_{R,K}^n)^2 + (\eta_{J,K}^n)^2 + (c_{\mathbf{S},K}^n)^{-1} \|\mathbf{S}_h^n \nabla (u_h^n - u_h^{n-1})\|_K^2\} \\ &\quad + \int_{t_{n-1}}^{t_n} \sum_{D \in \mathcal{D}_h^n} \{(\eta_{\text{DOQ},D}^n)^2 + (\eta_{\text{DO},D}^n)^2 + (\eta_{\text{Q},D}^n)^2\} dt \\ &\quad + \tau_n \sup_{\varphi \in H_0^1(\Omega); \|\varphi\|=1} (\nabla \cdot ((u_h^n - u_h^{n-1}) \mathbf{v}_h^n), \varphi)^2. \end{aligned}$$

*Proof.* Follows by combining the results of Lemmas 5.6–5.9.  $\square$

*Proof of Theorem 4.6.* Follows by combining Lemma 5.10 and [25, Lemma 7.1].  $\square$

## 6 Adaptive algorithm

We present here an adaptive algorithm based on our a posteriori error estimates which is designed to ensure that the relative energy error between the exact and approximate solutions will be below a prescribed tolerance  $\varepsilon$ , i.e., such that

$$\frac{\sum_{n=1}^N (\eta_{\text{sp}}^n + \eta_{\text{tm}}^n)^2}{\sum_{n=1}^N \|u_{h,\tau}\|_{X(t_{n-1}, t_n)}^2} \leq \varepsilon^2. \quad (6.1)$$

On a given time level  $t_{n-1}$ , we in particular choose the space mesh  $\mathcal{D}_h^n$  and time step  $\tau_n$  such that

$$\eta_{\text{sp}}^n \leq \varepsilon \frac{\|u_{h,\tau}\|_{X(t_{n-1}, t_n)}}{2}, \quad \eta_{\text{tm}}^n \leq \varepsilon \frac{\|u_{h,\tau}\|_{X(t_{n-1}, t_n)}}{2}.$$

At the same time, using the fact that there are no unknown constants hidden in both  $\eta_{\text{sp}}^n$  and  $\eta_{\text{tm}}^n$ , we intend to equilibrate the space and time errors. For practical implementation purposes, we introduce the maximal refinement level parameters  $N_{\text{sp}}$  and  $N_{\text{tm}}$ . Altogether with some other parameters of the algorithm, they are listed in Table 1. We also denote by  $\text{SpTmUnkn}$  the total number of space-time unknowns.

The actual algorithm is as follows:

- let an initial mesh  $\mathcal{D}_h^0$  and an initial time step  $\tau_1$  be given
- set up the initial conditions on  $\mathcal{D}_h^0$
- set  $t_0 = t_1 = 0$ ,  $\mathcal{D}_h^1 = \mathcal{D}_h^0$ , and  $n = 1$
- set  $\text{EstSpPrev} = 1$ ,  $\text{EstTmPrev} = 0$
- set  $\text{LevTmRef} = 0$ ,  $\text{SpTmUnkn} = 0$
- set  $\eta = 0$

Parameter	Meaning
$N_{\text{sp}}$	maximal level of space refinement
$N_{\text{tm}}$	maximal level of time refinement
Ref	fraction of cells for the space mesh refinement
Deref	fraction of cells for the space mesh derefinement
Comp	parameter for comparison of $\eta_{\text{sp}}$ and $\eta_{\text{tm}}$
StepsSpDeref	number of steps after which the space mesh is derefined
StepsTmDeref	number of steps after which the time mesh is derefined

Table 1: Different parameters of the adaptive algorithm and their meaning

**while**  $t_n < T$

- set Count = 0
- set  $t_n = t_{n-1} + \tau_n$
- set up the boundary conditions on  $\mathcal{D}_h^n$
- set  $\eta_{\text{sp}}^n = \text{Crit} = 1$ ,  $\text{ItSpRef} = 1$
- **while**  $\eta_{\text{sp}}^n \geq \text{Crit}$ ,  $\text{ItSpRef} \leq N_{\text{sp}} + 1$ , and  $\text{EstSpPrev} > \text{Comp} \cdot \text{EstTmPrev}$  when  $\text{ItSpRef} \neq 1$ 
  - **if**  $\text{ItSpRef} > 1$ 
    - refine such cells  $D \in \mathcal{D}_h^n$  where  $\eta_{D,\text{sp}}^n \geq \text{Ref} \cdot \max_{E \in \mathcal{D}_h^n} \eta_{E,\text{sp}}^n$  and such that their level of refinement is less than  $N_{\text{sp}}$
    - create a new mesh  $\mathcal{D}_h^n$  and interpolate the data onto this new mesh
  - solve (3.1) on  $\mathcal{D}_h^n$  with the time step  $\tau_n$  to get new  $u_{h,\tau}|_{[t_{n-1}, t_n]}$
  - compute the space a posteriori error estimate  $\eta_{\text{sp}}^n$
  - set  $\text{EstSpPrev} = \eta_{\text{sp}}^n / \sqrt{\tau_n}$
  - compute the norm of the approximate solution  $\|u_{h,\tau}\|_{X(t_{n-1}, t_n)}$  and set  $\text{Crit} = \varepsilon \cdot \|u_{h,\tau}\|_{X(t_{n-1}, t_n)} / 2$
  - set  $\text{ItSpRef} = \text{ItSpRef} + 1$
- compute the time a posteriori error estimate  $\eta_{\text{tm}}^n$
- set  $\text{EstTmPrev} = \eta_{\text{tm}}^n / \sqrt{\tau_n}$
- **if**  $\eta_{\text{tm}}^n \geq \text{Crit}$ ,  $\text{LevTmRef} < N_{\text{tm}}$ , and  $\text{EstTmPrev} > \text{Comp} \cdot \text{EstSpPrev}$ 
  - set  $t_n = t_n - \tau_n$ ,  $\tau_n = \tau_n / 3$ , and  $\text{LevTmRef} = \text{LevTmRef} + 1$
- **else**
  - $\eta^2 = \eta^2 + (\eta_{\text{tm}}^n + \eta_{\text{sp}}^n)^2$
  - $\text{SpTmUnkn} = \text{SpTmUnkn} + |\mathcal{D}_h^n|$
  - Count = Count + 1
  - **if** Count is a multiple of StepsSpDeref
    - derefine such cells  $D \in \mathcal{D}_h^n$  where  $\eta_{D,\text{sp}}^n \leq \text{Deref} \cdot \max_{E \in \mathcal{D}_h^n} \eta_{E,\text{sp}}^n$
    - create a new mesh  $\mathcal{D}_h^n$  and interpolate the data onto this mesh
  - **if** Count is a multiple of StepsTmDeref, set  $\tau_n = 3\tau_n$  and  $\text{LevTmRef} = \text{LevTmRef} - 1$
  - set  $\mathcal{D}_h^{n+1} = \mathcal{D}_h^n$ ,  $\tau_{n+1} = \tau_n$ , and  $n = n + 1$

## 7 Numerical experiments

As a model problem for numerical experiments, we consider (1.1a)–(1.1c) with

$$\begin{aligned}\Omega &= (0, 3) \times (0, 3), \\ \mathbf{S} &= \nu \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \\ \mathbf{v} &= (v_1, v_2), \\ r &= 0, \\ f &= 0,\end{aligned}$$

where  $v_1 = 0.8$  and  $v_2 = 0.4$  are two constant convective field components and  $\nu > 0$  determines the amount of diffusion. We will consider below the case  $\nu = 0.001$  and  $\nu = 10$ . The initial condition  $u_0$ , as well as the Dirichlet boundary condition, are given by the exact solution

$$u(x, y, t) = \frac{1}{200\nu(t + t_0) + 1} e^{-50 \frac{(x-x_0-v_1(t+t_0))^2 + (y-y_0-v_2(t+t_0))^2}{200\nu(t+t_0)+1}}.$$

Here,  $x_0 = 0.5$  and  $y_0 = 1.35$ . Moreover,  $t_0 \geq 0$ , set to  $t_0 = 0$  unless otherwise specified, is an additional parameter. Finally, we will test two cases with  $T = 0.6$  and  $T = 1.5$ , respectively.

We consider the scheme (3.1) on square meshes with possibly nonmatching refinements as indicated in Figure 2. Diffusive and convective flux reconstructions  $\boldsymbol{\theta}_h$  and  $\mathbf{w}_h$  are defined respectively by (4.6) and (4.7); the additional degrees of freedom are fixed following [31, Section 3.3.3] so that the residual estimators (4.8) are equal to zero. We use the estimator  $\eta$  given by (4.11b) (we only evaluate  $\eta^{(1),n}$ ) and present the results in the energy norm setting of Theorem 4.3. We do not present the results of Theorem 4.4 for the augmented norm since this norm is not easily calculable. We neglect the additional error from the inhomogeneous Dirichlet boundary condition.

The starting mesh is given by a uniform  $10 \times 10$  space grid and a uniform division of the time interval into 2 time steps when  $T = 0.6$  and 5 time steps when  $T = 1.5$ . The meshes are refined either uniformly or adaptively. In the first case, each square is always divided into nine subsquares, and the time step is cut by three. In the second case, the adaptive algorithm proposed in Section 6 is employed. We set  $\text{Ref} = 0.5$ ,  $\text{Deref} = 0.15$ ,  $\text{Comp} = 0.7$ ,  $\text{StepsSpDeref} = 6$ ,  $\text{StepsTmDeref} = 6$ . We will be choosing different values of  $N_{\text{sp}}$  and  $N_{\text{tm}}$  (recall their definition in Table 1). As we limit the maximal level of space and time refinements, we usually do not achieve the prescribed tolerance  $\varepsilon$  in (6.1). We define the experimental order of convergence  $\xi$  by

$$\xi := \frac{\log(e_N) - \log(e_{N-1})}{\frac{1}{3} \log |\mathcal{V}_{N-1}| - \frac{1}{3} \log |\mathcal{V}_N|},$$

where  $e_N$  is the error on the last space-time mesh,  $e_{N-1}$  is the error on the last but one space-time mesh, and  $|\mathcal{V}_N|$  and  $|\mathcal{V}_{N-1}|$  denote the corresponding numbers of total space-time unknowns given by  $\sum_{n=1}^N |\mathcal{D}_h^n|$ .

We first consider  $\nu = 0.001$  and  $T = 0.6$ . In this case, the problem is strongly convection-dominated and the exact solution takes a form of a steep Gaussian peak moving through the domain and only very slowly diffusing; the initial maximal value is 1, whereas the maximal value at the end of the simulation is roughly 0.9. Figure 6 below shows some approximate solutions. In Figure 3, we compare the actual error distribution and the one predicted by our a posteriori error estimate. The result is presented on the final time, on an adaptively refined mesh with  $N_{\text{sp}} = N_{\text{tm}} = 4$ . We can see that a correct form of the error distribution is predicted (circular

with smaller error in the middle of the circle), whereas this is not the case for the localization—the predicted error distribution is much more spread. We anticipate that this is caused by the fact that the solution itself is rather diffused; increasing  $N_{\text{sp}}$  and  $N_{\text{tm}}$  shall improve this considerably.

In the left part of Figure 4, we compare the actual error  $(\|(u - u_{h,\tau})(\cdot, T)\|^2 + \|u - u_{h,\tau}\|_X^2)^{\frac{1}{2}}$  with the estimate  $(\eta^2 + \|u_0 - u_{h,\tau}(\cdot, 0)\|^2)^{\frac{1}{2}}$  on uniformly/adaptively refined meshes. In the right part of Figure 4, we present the corresponding effectivity indices, given as the ratios of the estimate over the error. In the adaptive refinement strategy, we obtain the same precision for much fewer (roughly 12 times less for the last meshes) space-time unknowns than in the uniform one. Concerning the experimental order of convergence, we have found  $\xi = 0.64$  and  $\xi = 0.71$  in the uniform and adaptive cases, respectively. The effectivity indices depend on the local Péclet number and improve as the mesh is refined (and the local Péclet number decreased), as expected for the energy norm setting which is not robust with respect to convection dominance. We then in Figure 5 plot the same results for a 2.5 times longer final time  $T = 1.5$ . We find very similar results, with in particular very similar effectivity indices. These thus seem to be rather independent of the simulation duration. In the uniform and adaptive case, respectively, we find  $\xi = 0.47$  and  $\xi = 0.54$ .

We next compare the uniform and adaptive refinement strategies visually, for  $\nu = 0.001$  and  $T = 0.6$ . Figure 6 shows the approximate solution at the final time obtained in the adaptive case with  $N_{\text{sp}} = N_{\text{tm}} = 2$  (left) and  $N_{\text{sp}} = N_{\text{tm}} = 4$  (right). We can see that whereas in the first case, the numerical diffusion is extremely strong (notice that it only applies in the streamline direction by the definition of the local Péclet upstream weighting (3.7)), in the second one the approximate solution starts to capture the exact one rather well. Figure 7 then compares the uniform refinement strategy with the adaptive one. In its left part, the uniform refinement approximate solution at the final time with  $N_{\text{sp}} = N_{\text{tm}} = 3$  is shown, whereas in its right part, we present its adaptive refinement counterpart for  $N_{\text{sp}} = N_{\text{tm}} = 4$ . The adaptive case gives clearly much better results, and this for roughly the same total number of space-time unknowns.

We finally focus on the diffusion-dominant case  $\nu = 10$  (we consider  $T = 1.5$ ). In this case, the adaptive refinement strategy does not lead to important improvements and hence we only present results for the uniform refinement. In Figure 8, the effectivity index is excellent as very close to the optimal value of one. This is probably caused by the fact that the exact solution almost instantaneously diffuses from the maximal value of 1 to the final maximal value of order  $10^{-4}$ . Coincidentally,  $\xi = 0.53$  only. The solution presents much smoother behavior for the parameter  $t_0 = 0.5$ , in which case  $\xi$  takes the almost optimal value 0.97. We illustrate this case in Figure 9.

## References

- [1] AMAZIANE, B., BERGAM, A., EL OSSMANI, M., AND MGHAZLI, Z. A posteriori estimators for vertex centred finite volume discretization of a convection-diffusion-reaction equation arising in flow in porous media. *Internat. J. Numer. Methods Fluids* 59, 3 (2009), 259–284.
- [2] ANGERMANN, L. An a posteriori estimation for the solution of elliptic boundary value problems by means of upwind FEM. *IMA J. Numer. Anal.* 12, 2 (1992), 201–215.
- [3] BANK, R. E., AND ROSE, D. J. Some error estimates for the box method. *SIAM J. Numer. Anal.* 24, 4 (1987), 777–787.
- [4] BEBENDORF, M. A note on the Poincaré inequality for convex domains. *Z. Anal. Anwendungen* 22, 4 (2003), 751–756.

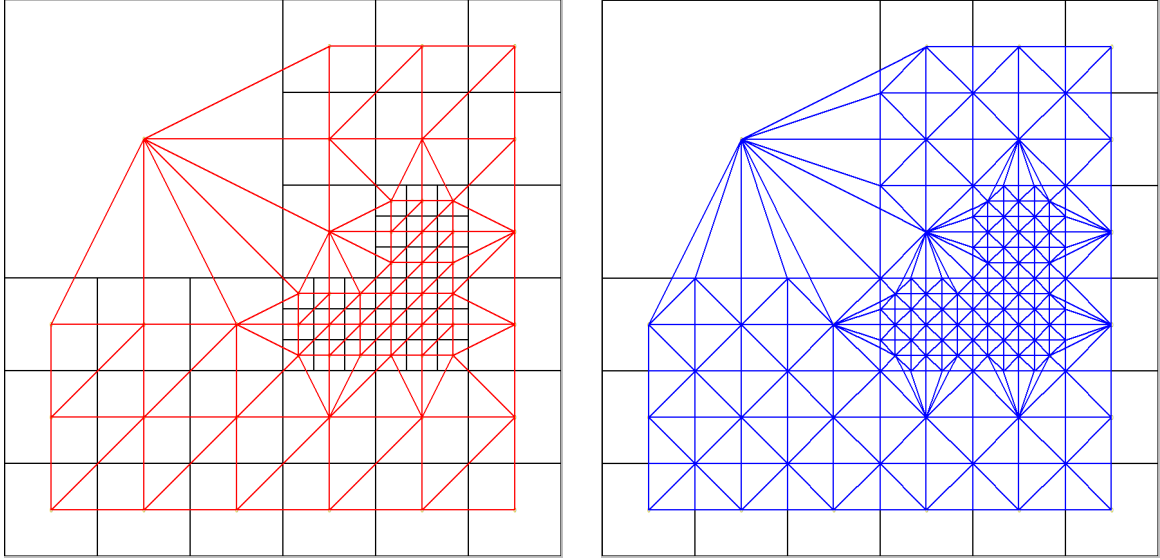


Figure 2: Nonmatching primal square mesh  $\mathcal{D}_h^n$ , the corresponding dual triangular mesh  $\mathcal{T}_h^n$  (left), and the corresponding tertial triangular mesh  $\mathcal{S}_h^{n-1,n}$

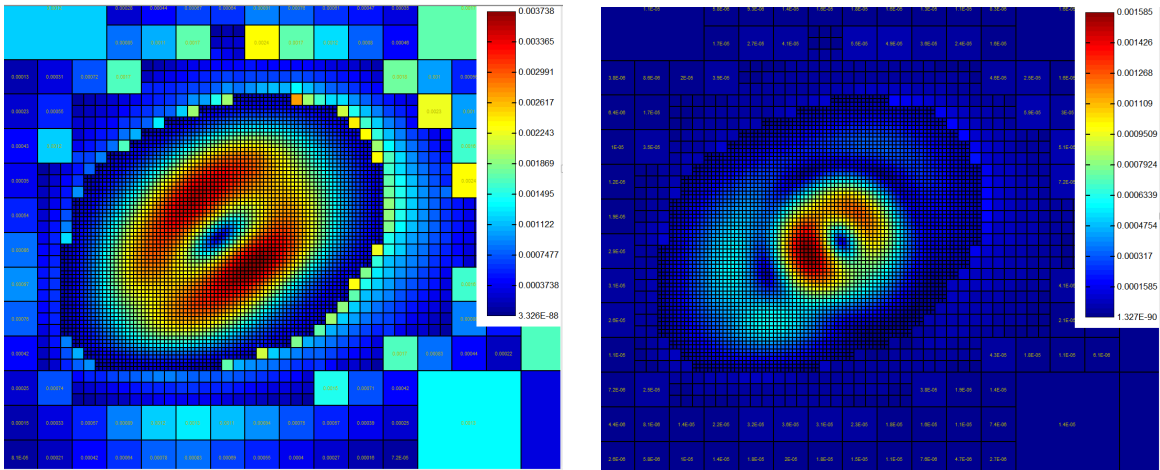


Figure 3: Estimated (left) and actual (right) error distribution, adaptive refinement,  $\nu = 0.001$ ,  $T = 0.6$ , adaptively refined mesh with  $N_{sp} = N_{tm} = 4$

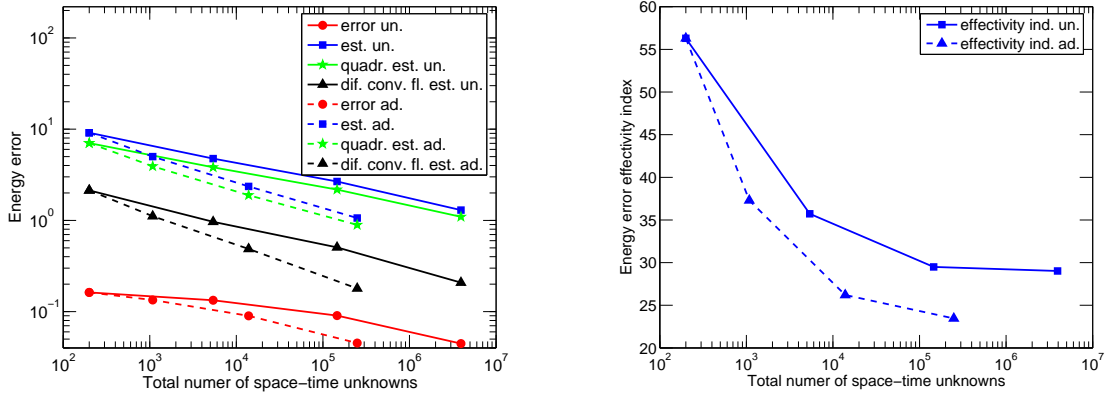


Figure 4: Estimated and actual errors (left) and corresponding effectivity indices (right),  $\nu = 0.001$ ,  $T = 0.6$

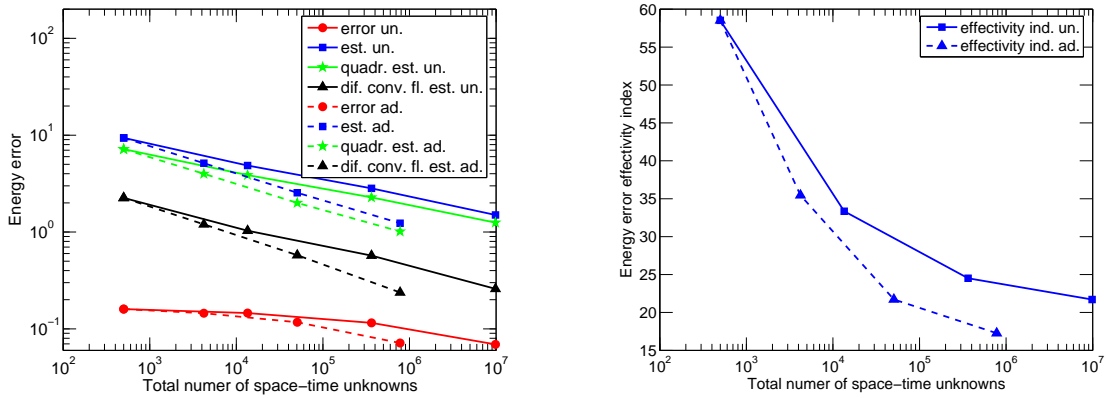


Figure 5: Estimated and actual errors (left) and corresponding effectivity indices (right),  $\nu = 0.001$ ,  $T = 1.5$

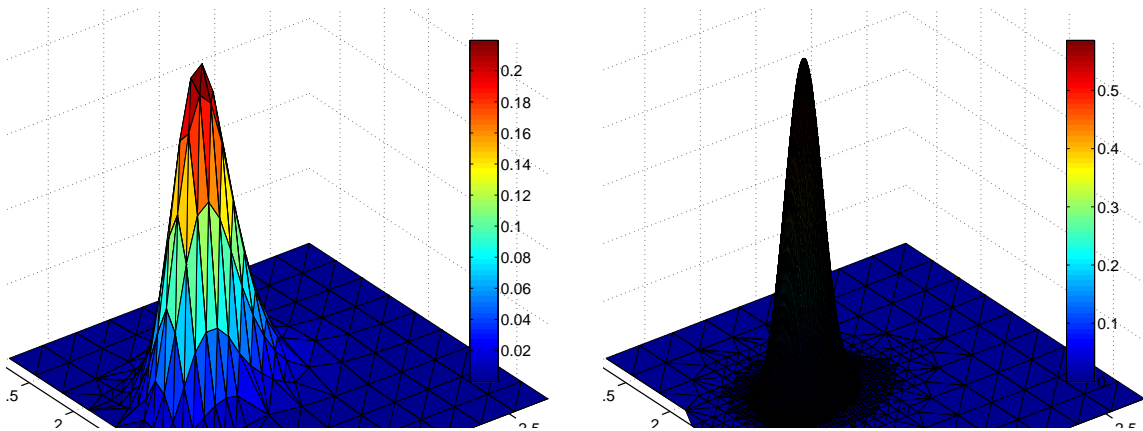


Figure 6: Adaptive refinement approximate solutions for  $\nu = 0.001$ ,  $T = 0.6$ , and  $N_{sp} = N_{tm} = 2$  (left) and  $N_{sp} = N_{tm} = 4$  (right)



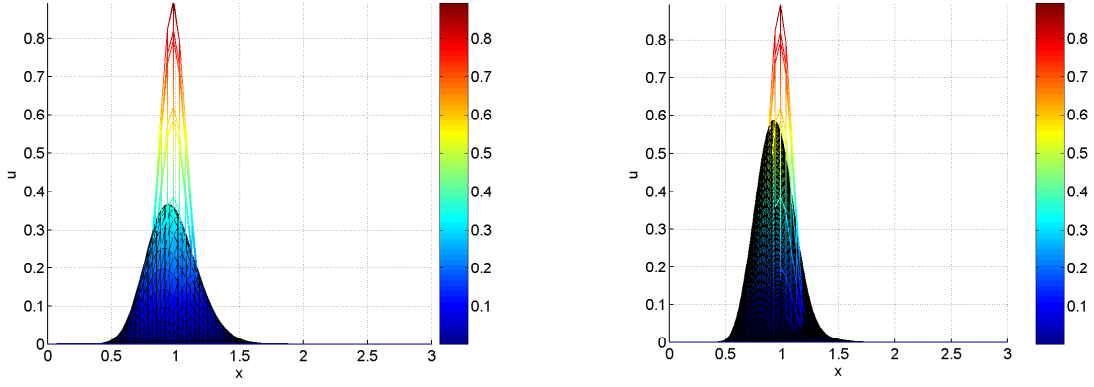


Figure 7: Uniform refinement approximate solution for  $N_{\text{sp}} = N_{\text{tm}} = 3$  (left) and adaptive refinement approximate solution for  $N_{\text{sp}} = N_{\text{tm}} = 4$  (right),  $\nu = 0.001$  and  $T = 0.6$

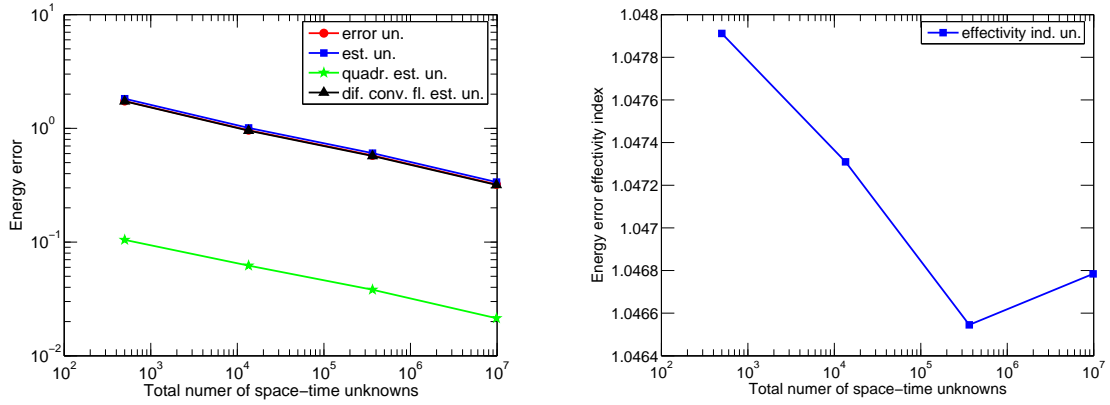


Figure 8: Estimated and actual errors (left) and corresponding effectivity indices (right),  $\nu = 10$ ,  $T = 1.5$

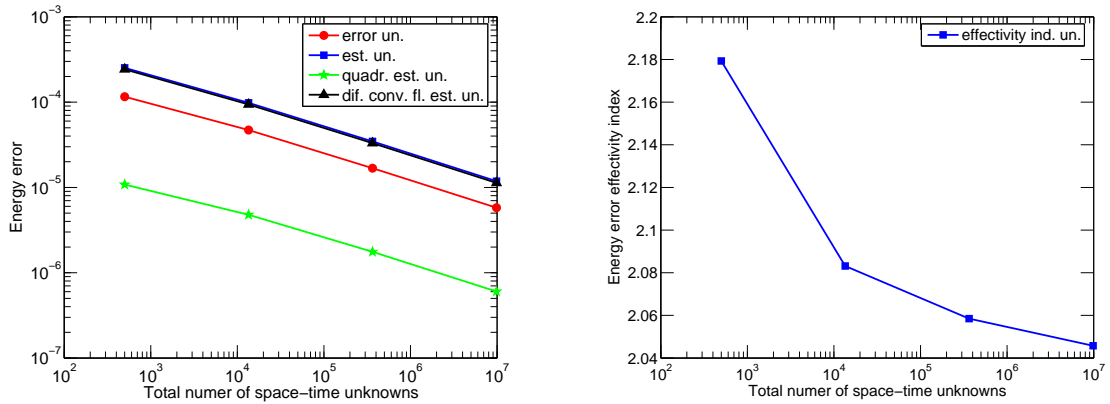


Figure 9: Estimated and actual errors (left) and corresponding effectivity indices (right),  $\nu = 10$ ,  $T = 1.5$ , and  $t_0 = 0.5$

- [5] BREZZI, F., AND FORTIN, M. *Mixed and hybrid finite element methods*, vol. 15 of *Springer Series in Computational Mathematics*. Springer-Verlag, New York, 1991.
- [6] CARSTENSEN, C., AND FUNKEN, S. A. Constants in Clément-interpolation error and residual based a posteriori error estimates in finite element methods. *East-West J. Numer. Math.* 8, 3 (2000), 153–175.
- [7] CARSTENSEN, C., LAZAROV, R., AND TOMOV, S. Explicit and averaging a posteriori error estimates for adaptive finite volume methods. *SIAM J. Numer. Anal.* 42, 6 (2005), 2496–2521.
- [8] CHEDDADI, I., FUČÍK, R., PRIETO, M. I., AND VOHRALÍK, M. Guaranteed and robust a posteriori error estimates for singularly perturbed reaction–diffusion problems. *M2AN Math. Model. Numer. Anal.* 43, 5 (2009), 867–888.
- [9] ERIKSSON, K., AND JOHNSON, C. Adaptive streamline diffusion finite element methods for stationary convection–diffusion problems. *Math. Comp.* 60, 201 (1993), 167–188, S1–S2.
- [10] ERN, A., STEPHANSEN, A. F., AND VOHRALÍK, M. Guaranteed and robust discontinuous Galerkin a posteriori error estimates for convection–diffusion–reaction problems. *J. Comput. Appl. Math.* (2009). Accepted for publication, DOI 10.1016/j.cam.2009.12.009.
- [11] ERN, A., AND VOHRALÍK, M. A posteriori error estimation based on potential and flux reconstruction for the heat equation. *SIAM J. Numer. Anal.* (2010). Accepted for publication.
- [12] EYMARD, R., GALLOUËT, T., AND HERBIN, R. Finite volume methods. In *Handbook of Numerical Analysis, Vol. VII*. North-Holland, Amsterdam, 2000, pp. 713–1020.
- [13] EYMARD, R., HILHORST, D., AND VOHRALÍK, M. A combined finite volume–finite element scheme for the discretization of strongly nonlinear convection–diffusion–reaction problems on nonmatching grids. *Numer. Methods Partial Differential Equations* (2009). DOI 10.1002/num.20449.
- [14] FEISTAUER, M., FELCMAN, J., LUKÁČOVÁ-MEDVID’OVÁ, M., AND WARNECKE, G. Error estimates for a combined finite volume–finite element method for nonlinear convection–diffusion problems. *SIAM J. Numer. Anal.* 36, 5 (1999), 1528–1548.
- [15] FELCMAN, J., AND KUBERA, P. An adaptive finite volume method for non-stationary problems. In *Finite volumes for complex applications V*. ISTE, London, 2008, pp. 407–414.
- [16] JU, L., WU, W., AND ZHAO, W. Adaptive finite volume methods for steady convection–diffusion equations with mesh optimization. *Discrete Contin. Dyn. Syst. Ser. B* 11, 3 (2009), 669–690.
- [17] LAZAROV, R., AND TOMOV, S. A posteriori error estimates for finite volume element approximations of convection–diffusion–reaction equations. *Comput. Geosci.* 6, 3-4 (2002), 483–503. Locally conservative numerical methods for flow in porous media.
- [18] NICAISE, S. A posteriori error estimations of some cell centered finite volume methods for diffusion–convection–reaction problems. *SIAM J. Numer. Anal.* 44, 3 (2006), 949–978.
- [19] OHLBERGER, M. A posteriori error estimates for vertex centered finite volume approximations of convection–diffusion–reaction equations. *M2AN Math. Model. Numer. Anal.* 35, 2 (2001), 355–387.

- [20] PAYNE, L. E., AND WEINBERGER, H. F. An optimal Poincaré inequality for convex domains. *Arch. Rational Mech. Anal.* 5 (1960), 286–292.
- [21] PICASSO, M. Adaptive finite elements for a linear parabolic problem. *Comput. Methods Appl. Mech. Engrg.* 167, 3-4 (1998), 223–237.
- [22] STEPHANSEN, A. F. *Méthodes de Galerkin discontinues et analyse d’erreur a posteriori pour les problèmes de diffusion hétérogène*. Ph.D. thesis, Ecole Nationale des Ponts et Chaussées, 2007.
- [23] VERFÜRTH, R. A posteriori error estimators for convection–diffusion equations. *Numer. Math.* 80, 4 (1998), 641–663.
- [24] VERFÜRTH, R. A posteriori error estimates for finite element discretizations of the heat equation. *Calcolo* 40, 3 (2003), 195–212.
- [25] VERFÜRTH, R. Robust a posteriori error estimates for nonstationary convection-diffusion equations. *SIAM J. Numer. Anal.* 43, 4 (2005), 1783–1802.
- [26] VERFÜRTH, R. Robust a posteriori error estimates for stationary convection-diffusion equations. *SIAM J. Numer. Anal.* 43, 4 (2005), 1766–1782.
- [27] VERFÜRTH, R. A note on constant-free a posteriori error estimates. *SIAM J. Numer. Anal.* 47, 4 (2009), 3180–3194.
- [28] VOHRALÍK, M. On the discrete Poincaré–Friedrichs inequalities for nonconforming approximations of the Sobolev space  $H^1$ . *Numer. Funct. Anal. Optim.* 26, 7–8 (2005), 925–952.
- [29] VOHRALÍK, M. A posteriori error estimation in the conforming finite element method based on its local conservativity and using local minimization. *C. R. Math. Acad. Sci. Paris* 346, 11–12 (2008), 687–690.
- [30] VOHRALÍK, M. Residual flux-based a posteriori error estimates for finite volume and related locally conservative methods. *Numer. Math.* 111, 1 (2008), 121–158.
- [31] VOHRALÍK, M. Guaranteed and fully robust a posteriori error estimates for conforming discretizations of diffusion problems with discontinuous coefficients. HAL Preprint 00235810, version 2, submitted for publication, 2009.