

# A combined finite volume–finite element scheme for the discretization of strongly nonlinear convection–diffusion–reaction problems on nonmatching grids\*

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

<sup>1</sup> Département de Mathématiques, Université de Marne-la-Vallée, 5 boulevard Descartes, Champs-sur-Marne, 77 454 Marne-la-Vallée, France. [eynard@math.univ-mlv.fr](mailto:eynard@math.univ-mlv.fr)

<sup>2</sup> Laboratoire de Mathématiques, Analyse Numérique et EDP, Université de Paris-Sud et CNRS, Bât. 425, 91 405 Orsay, France. [[danielle.hilhorst](mailto:danielle.hilhorst), [martin.vohralik](mailto:martin.vohralik)][@math.u-psud.fr](mailto:math.u-psud.fr)

## Abstract

We propose and analyze a numerical scheme for nonlinear degenerate parabolic convection–diffusion–reaction equations in two or three space dimensions. We discretize the time evolution, convection, reaction, and sources terms on a given grid, which can be nonmatching and can contain nonconvex elements, by means of the cell-centered finite volume method. To discretize the diffusion term, we construct a conforming simplicial mesh with the vertices given by the original grid and use the finite element method. In this way, the scheme is fully consistent and the discrete solution is naturally continuous across the interfaces between the subdomains with nonmatching grids, without introducing any supplementary equations and unknowns or using any interpolation of the discrete solutions at the interfaces. We allow for general inhomogeneous and anisotropic diffusion–dispersion tensors and use the local Péclet upstream weighting in order to only add the minimal numerical diffusion necessary to avoid spurious oscillations in the convection-dominated case. The scheme is robust, efficient, locally conservative, and satisfies the discrete maximum principle under some conditions on the mesh and the diffusion tensor. We prove its convergence using a priori estimates and the Kolmogorov relative compactness theorem and illustrate its behavior on a numerical experiment.

**Key words:** degenerate parabolic problem, convection–diffusion–reaction equation, inhomogeneous and anisotropic diffusion, convection dominance, nonmatching grids, finite volume method, finite element method, convergence of approximate solutions

**AMS subject classifications:** 65M12, 76M10, 76M12, 35K65, 76S05

## 1 Introduction

Nonlinear convection–diffusion–reaction equations arise in many contexts, such as flow in porous media, heat conduction, or free boundary problems. This paper is motivated by contaminant transport in porous media with equilibrium adsorption reaction, described by the equation

$$\frac{\partial \beta(c)}{\partial t} - \nabla \cdot (\mathbf{S} \nabla c) + \nabla \cdot (c \mathbf{v}) + F(c) = q, \quad (1.1)$$

---

\*This work was supported by the GdR MoMaS, CNRS-2439, ANDRA, BRGM, CEA, EdF, France.

cf. [4, 6], where  $c$  is the unknown concentration of the contaminant, the function  $\beta(\cdot)$  represents time evolution and equilibrium adsorption reaction,  $\mathbf{S}$  is the diffusion–dispersion tensor,  $\mathbf{v}$  is the velocity field,  $F(\cdot)$  is a reaction function, and finally,  $q$  stands for sources. In (1.1), the tensor  $\mathbf{S}$  is usually inhomogeneous and anisotropic (nonconstant full-matrix) and the velocity field  $\mathbf{v}$  is dominating. Moreover, if  $\beta'$  is unbounded, the equation (1.1) is degenerate parabolic.

A large variety of methods has been proposed for the discretization of the problem (1.1). The conforming piecewise linear finite element method has been studied e.g. by Barrett and Knabner [4], Ebmeyer [12], and Rulla and Walkington [30], the cell-centered finite volume method by Baughman and Walkington [5] and Eymard *et al* [15], the vertex-centered finite volume method by Ohlberger [26], the combined finite volume–nonconforming finite element method by the authors [18], the finite difference method by Karlsen *et al* [23], and the mixed finite element method by Arbogast *et al* [3]. In all these references, only matching (conforming) grids have been considered.

In recent years, there was a growing interest in discretization methods for nonmatching grids, usually proposed in the context of domain decomposition methods (cf. Quarteroni and Valli [29]). The mortar method, developed for elliptic problems discretized by the finite element or spectral methods by Bernardi *et al* [7], enforces a weak continuity of the discrete solution across the interfaces between the subdomains with nonmatching grids with the help of auxiliary spaces defined at the interfaces. It is recognized for its stability and accuracy and it has been later extended to other methods, such as to the mixed finite element method by Arbogast *et al* [2] and to the finite volume element method by Ewing *et al* [13]. Analogous approaches for cell-centered finite volume methods have been proposed by Achdou *et al* [1] or Faille *et al* [19]. Alternatively, other techniques use interpolation of the discrete solutions at the interfaces. A different approach, avoiding the need to introduce supplementary equations and unknowns or to interpolate at the interfaces and leading to a simple and stable cell-centered finite volume scheme for nonoverlapping nonmatching grids has been proposed by Cautrès *et al* [9]. Its disadvantage however is that the diffusive flux approximation across the nonmatching interfaces is nonconsistent, which leads to the introduction of an additional numerical error. In all the above references, only linear elliptic problems are considered.

The intention of this paper is twofold. Firstly, we want to propose a simple, stable, consistent, and efficient method permitting the use of very general and in particular nonmatching grids. Secondly, we want this method to enable the discretization of degenerate parabolic convection–diffusion–reaction problems of type (1.1). We employ for this purpose the ideas of combined finite volume–finite element schemes (cf. Sonier and Eymard [31], Feistauer *et al* [20], or Debiez *et al* [11]) and in particular those of the scheme proposed and studied by the authors [18]. In these schemes, one supposes that there is a given simplicial (consisting of triangles/tetrahedra) mesh and discretizes the diffusion term of (1.1) by means of a finite element method on this mesh, whereas the other terms are discretized by means of the cell-centered finite volume method on a dual partition constructed on the basis of the simplicial one. In this paper, we rather start from the finite volume partition, being motivated by the following consideration: for a pure cell-centered finite volume discretization of the equation (1.1) without the diffusion term, the mesh can in fact be nonmatching and can even contain nonconvex control volumes, cf. [14, Chapter VI]. The mesh is required to match along hyperplanes and to consist of convex control volumes only when the diffusion term is present, cf. [14, Chapter III]. At the same time, however, we notice that given a set of points, we can always construct a simplicial mesh with vertices given by this set of points. Hence an intuitive idea is as follows: given a nonmatching mesh (with possibly nonconvex elements) and with a set of points associated with these elements, construct a simplicial mesh having this set of points as vertices. Then consider a finite element discretization of the diffusion term of (1.1) on the simplicial mesh and a finite volume discretization of the other terms of (1.1) on the original mesh.

We believe that our approach to nonmatching grids is in some sense the simplest, yet very efficient. In particular, we do not introduce any supplementary equations and unknowns, nor do we use any interpolation of the discrete solutions at the interfaces. There is no need for this for the finite volume part of the scheme and the finite element part uses a conforming mesh. The proposed scheme is in this respect similar to that from [9]. The essential difference is that we replace the finite volume diffusion fluxes by the finite element ones. This is very important in the present case, since the diffusion fluxes through the interfaces between the subdomains with nonmatching grids of the scheme proposed in [9] are not consistent, whereas our discretization of the diffusion term (over the interfaces) is fully consistent. We however have to construct a dual simplicial mesh, which is not necessary in [9].

We can mention the following advantages of our scheme with respect to the convection–diffusion–reaction equation (1.1). Our scheme is stable but not excessively diffusive in the convection-dominated case since we avoid the spurious oscillations by checking the local Péclet number and by adding exactly the necessary amount of upstream weighting. It inherits the diffusion properties of the finite element method, enabling in particular an easy discretization of inhomogeneous and anisotropic diffusion tensors. Moreover, in order to avoid the smearing of the fronts for highly inhomogeneous problems, we give, in addition to the classical finite element discretization leading to the arithmetic average of the diffusion–dispersion tensor, a second variant corresponding to the harmonic average. Our scheme next possesses the discrete maximum principle in the case where all transmissibilities are non-negative. This happens for instance when the diffusion tensor reduces to a constant scalar function and when the simplicial mesh is Delaunay in two space dimensions, or when the diffusion tensor is scalar and when the angles between the outward normal vectors of sides of each simplex in the triangulation are greater or equal to  $\pi/2$  in two or three space dimensions. Our scheme only has the number of unknowns equal to the number of grid cells and the arising matrices are positive definite (but generally nonsymmetric because of the upstream weighting). It is also locally conservative. It next permits to efficiently discretize degenerate parabolic problems: when we search for the discrete unknowns corresponding to  $\beta(c)$ , the resulting system of nonlinear algebraic equations can be solved by the Newton method without any parabolic regularization (cf. [4]) or perturbation of initial and boundary conditions (cf. [27]), which make the equation uniformly parabolic. Moreover, the resulting matrices are diagonal for the part of the unknowns corresponding to the region where the approximate solution is zero. Finally, differing only in the discretization of the diffusion term from standard cell-centered finite volume schemes, our scheme can easily be implemented in any finite volume code, in order to permit a (nonmatching) local refinement of the mesh and an easy discretization of inhomogeneous and anisotropic tensors, a highly desirable feature e.g. in contaminant transport modeling. Finally, especially on rectangular grids, the ease of (nonmatching) local refinement using our scheme makes it appealing for adaptive simulations.

Our numerical scheme permits to construct approximate solutions that are piecewise constant on the given mesh or piecewise linear and continuous on the simplicial mesh. We prove the convergence of both these approximations to a weak solution of the continuous problem in this paper. The methods of proof follow the lines previously given by the authors [18] and are based upon the Kolmogorov relative compactness theorem and the finite volume tools from [14]. We only suppose that  $\beta$  is continuous with the growth bounded from below in the case where the discrete maximum principle is satisfied. In the general case we require in addition  $\beta$  to be bounded on some interval and Lipschitz-continuous outside this interval. There is no restriction on the maximal time step in the case where  $F$  is nondecreasing. If  $F$  does not possess this property, we impose an appropriate maximal time step condition. For the sake of simplicity, we only consider the case of a homogeneous Dirichlet boundary condition. Extensions to other types of boundary conditions

and to the case where the equation (1.1) involves a nonlinear convection term are possible, using the techniques from [14] and [15].

The rest of the paper is organized as follows. In Section 2 we state the assumptions on the data and present a weak formulation of the continuous problem. In Section 3 we define admissible grids, introduce the combined finite volume–finite element scheme, and give several remarks on the scheme and its possible extensions. In Section 4 we present some properties of our scheme and prove that it possesses a unique solution, which satisfies the discrete maximum principle under the hypotheses stated above. In Section 5 we derive a priori estimates and estimates on differences of time and space translates for the approximate solutions. Finally, in Section 6, using the Kolmogorov relative compactness theorem, we prove the convergence of a subsequence of the sequence of approximate solutions to a weak solution of the continuous problem. The results of a numerical experiment are presented in Section 7.

## 2 The nonlinear degenerate parabolic problem

Let  $\Omega \subset \mathbb{R}^d$ ,  $d = 2, 3$ , be a polygonal (we use this term for  $d = 3$  as well instead of polyhedral) domain (open, bounded, and connected set) with boundary  $\partial\Omega$ , let  $(0, T)$ ,  $0 < T < \infty$ , be a time interval, and let us define  $Q_T := \Omega \times (0, T)$ . We consider the equation (1.1) in  $Q_T$  together with the homogeneous Dirichlet boundary condition

$$c = 0 \quad \text{on } \partial\Omega \times (0, T) \quad (2.1)$$

and the initial condition

$$c(\cdot, 0) = c_0 \quad \text{in } \Omega. \quad (2.2)$$

Suppose that  $S$  is a domain of  $\mathbb{R}^d$ . We use the standard notation  $L^p(S)$  and  $\mathbf{L}^p(S) = [L^p(S)]^d$  for the Lebesgue spaces on  $S$ ,  $(\cdot, \cdot)_S$  stands for the  $L^2(S)$  or  $\mathbf{L}^2(S)$  inner product, and  $\|\cdot\|_S$  for the associated norm. We use  $dx$  as the integration symbol for the Lebesgue measure on  $S$ ,  $d\gamma(\mathbf{x})$  for the Lebesgue measure on a hyperplane of  $S$ , and  $dt$  for the Lebesgue measure on  $(0, T)$ . We denote by  $|S|$  the  $d$ -dimensional 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,  $H^1(S)$  and  $H_0^1(S)$  are the Sobolev spaces of functions with square-integrable weak derivatives and  $\mathbf{H}(\text{div}, S)$  is the space of vector functions with square-integrable weak divergences,  $\mathbf{H}(\text{div}, S) = \{\mathbf{v} \in \mathbf{L}^2(S); \nabla \cdot \mathbf{v} \in L^2(S)\}$ . In the subsequent text we will denote by  $C_A$ ,  $c_A$  a constant basically dependent on a quantity  $A$  but always independent of the discretization parameters  $h$  and  $\Delta t$  whose definition we shall give later. We make the following assumption on the data:

*Assumption (A) (Data)*

(A1)  $\beta \in C(\mathbb{R})$ ,  $\beta(0) = 0$ , is a strictly increasing function such that, for all  $a, b \in \mathbb{R}$ ,

$$|\beta(a) - \beta(b)| \geq c_\beta |a - b|, \quad c_\beta > 0$$

or

(A2) in addition to (A1), there exists  $P \in \mathbb{R}$ ,  $P > 0$ , such that  $|\beta(s)| \leq C_\beta$  in  $[-P, P]$ ,  $C_\beta > 0$ , and  $\beta$  is Lipschitz-continuous with a constant  $L_\beta$  on  $(-\infty, -P]$  and  $[P, +\infty)$ ;

(A3)  $\mathbf{S}_{ij} \in L^\infty(Q_T)$ ,  $|\mathbf{S}_{ij}| \leq C_{\mathbf{S}}/d$  a.e. in  $Q_T$ ,  $1 \leq i, j \leq d$ ,  $C_{\mathbf{S}} > 0$ ,  $\mathbf{S}$  is a symmetric and uniformly positive definite tensor for almost all  $t \in (0, T)$  with a constant  $c_{\mathbf{S}} > 0$ , i.e.

$$\mathbf{S}(\mathbf{x}, t)\boldsymbol{\eta} \cdot \boldsymbol{\eta} \geq c_{\mathbf{S}} \boldsymbol{\eta} \cdot \boldsymbol{\eta} \quad \forall \boldsymbol{\eta} \in \mathbb{R}^d, \text{ for a.e. } (\mathbf{x}, t) \in Q_T;$$

(A4)  $\mathbf{v} \in L^2(0, T; \mathbf{H}(\text{div}, \Omega)) \cap \mathbf{L}^\infty(Q_T)$  satisfies  $\nabla \cdot \mathbf{v} = r \geq 0$  and  $|\mathbf{v}| \leq C_{\mathbf{v}}$ ,  $C_{\mathbf{v}} > 0$ , a.e. in  $Q_T$ ;

(A5)  $F(0) = 0$ ,  $F$  is a nondecreasing, Lipschitz-continuous function with a constant  $L_F$

or

(A6)  $F(0) = 0$ ,  $F$  is a Lipschitz-continuous function with a constant  $L_F$  and there holds  $sF(s) \geq 0$  for  $s < 0$  and  $s > M$ ,  $M > 0$ ;

(A7)  $q \in L^2(Q_T)$ , where  $q = r\bar{c}$  with  $\bar{c} \in L^\infty(Q_T)$ ,  $0 \leq \bar{c} \leq M$  a.e. in  $Q_T$ ;

(A8)  $c_0 \in L^\infty(\Omega)$ ,  $0 \leq c_0 \leq M$  a.e. in  $\Omega$ .

**Remark 2.1** (Hypotheses on  $\beta$ ). In contaminant transport problems one typically has  $\beta(c) = c + c^\alpha$ ,  $\alpha \in (0, 1)$ . Assumption (A1) generalizes this type of functions; we in particular do not limit the number of points where  $\beta'$  explodes. As we shall see, we will be able to prove the convergence of the combined scheme with this assumption only for the case where the discrete maximum principle (cf. Theorem 4.11 below) holds. In the general case we add Assumption (A2), which is however still satisfied by all realistic functions  $\beta$ . Also, it is necessary that the function  $\beta$  was defined for negative values since our scheme can take them in this latter case.

We now give the definition of a weak solution of (1.1)–(2.2), following essentially Knabner and Otto [24].

**Definition 2.2** (Weak solution). We say that a function  $c$  is a weak solution of the problem (1.1)–(2.2) if

- (i)  $c \in L^2(0, T; H_0^1(\Omega))$ ,
- (ii)  $\beta(c) \in L^\infty(0, T; L^2(\Omega))$ ,
- (iii)  $c$  satisfies the integral equality
 
$$\begin{aligned}
 & - \int_0^T \int_\Omega \beta(c) \varphi_t \, d\mathbf{x} \, dt - \int_\Omega \beta(c_0) \varphi(\cdot, 0) \, d\mathbf{x} + \int_0^T \int_\Omega \mathbf{S} \nabla c \cdot \nabla \varphi \, d\mathbf{x} \, dt - \\
 & - \int_0^T \int_\Omega c \mathbf{v} \cdot \nabla \varphi \, d\mathbf{x} \, dt + \int_0^T \int_\Omega F(c) \varphi \, d\mathbf{x} \, dt = \int_0^T \int_\Omega q \varphi \, d\mathbf{x} \, dt
 \end{aligned}$$
 for all  $\varphi \in L^2(0, T; H_0^1(\Omega))$  with  $\varphi_t \in L^\infty(Q_T)$ ,  $\varphi(\cdot, T) = 0$ .

**Remark 2.3** (Existence of a weak solution). The existence of at least one weak solution is proved e.g. in Theorem 6.2 below (cf. also [18, Theorem 6.2]).

**Remark 2.4** (Uniqueness of a weak solution). For a slightly more restrictive hypothesis on the data than that given in Assumption (A), the uniqueness of a weak solution given by Definition 2.2 is guaranteed in [24]. Namely, no time-dependency of the diffusion–dispersion tensor  $\mathbf{S}$  is still required in [24].

### 3 Combined finite volume–finite element scheme

We will describe the space and time discretizations, define the approximation spaces, introduce the combined finite volume–finite element scheme, and give several remarks in this section.

### 3.1 Space discretization

As a *primal mesh* of  $\Omega$ , we understand a partition  $\mathcal{D}_h$  of  $\Omega$  into closed polygons such that  $\bar{\Omega} = \bigcup_{D \in \mathcal{D}_h} D$  and such that the intersection of interiors of two different polygons is empty. We in particular admit nonmatching grids, i.e. the case where there exist two different polygons  $D, E \in \mathcal{D}_h$  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  $D$ . 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  $\mathcal{V}_h^{\text{int}}$  such that there is one point  $V_D$  in the interior of  $D$  associated with each  $D \in \mathcal{D}_h$ . For  $D \in \mathcal{D}_h$ , we denote by  $\mathcal{F}_D$  the set of such subsets  $\sigma$  of  $\partial D$  that there exists  $E \in \mathcal{D}_h$  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  $\partial\Omega$ , 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  $V_E \in \mathcal{V}_h^{\text{ext}}$  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 (edge if  $d=2$ , face if  $d=3$ ) of  $D$ , and that  $\sigma_{D,E}$  not necessarily lies in a hyperplane of  $\mathbb{R}^d$ , see Figure 1. We next denote by  $\mathcal{F}_h$  the union of all  $\sigma$ , by  $\mathcal{F}_h^{\text{int}}$  the union of all  $\sigma_{D,E} = \partial D \cap \partial E$  for some  $D, E \in \mathcal{D}_h$ , and by  $\mathcal{F}_h^{\text{ext}}$  the union of all  $\sigma \subset \partial\Omega$ . We also denote, for  $\sigma_{D,E} \in \mathcal{F}_h$ , by  $\mathbf{b}_{D,E}$  the line segment connecting the vertices  $V_D$  and  $V_E$  and put  $d_{D,E} := |V_D - V_E|$ . Finally, denoting by  $h_D$  the diameter of  $D$ , the primal mesh discretization parameter  $h_{\mathcal{D}}$  is given by  $h_{\mathcal{D}} := \max_{D \in \mathcal{D}_h} h_D$ . We make the following admissibility assumption on the family of primal meshes  $\{\mathcal{D}_h\}_h$ :

*Assumption (B) (Admissibility of the primal meshes)*

(B1) *There exists a constant  $C_{1,\mathcal{D}} > 0$  such that*

$$|\mathcal{F}_D| \leq C_{1,\mathcal{D}} \quad \forall D \in \mathcal{D}_h, \forall h > 0;$$

(B2) *There exists a constant  $C_{2,\mathcal{D}} > 0$  such that*

$$|\sigma_{D,E}| d_{D,E} \leq C_{2,\mathcal{D}} |D| \quad \forall D \in \mathcal{D}_h, \forall \sigma_{D,E} \in \mathcal{F}_D, \forall h > 0.$$

Assumption (B1) simply states that each  $D \in \mathcal{D}_h$  has at most  $C_{1,\mathcal{D}}$  “neighbors”. Assumption (B2) then expresses that the measure of each “side”  $\sigma_{D,E} \in \mathcal{F}_D$  is bounded with respect to the measure of  $D$  and the distance of  $V_D$  and  $V_E$ .

A *dual mesh* of  $\Omega$  is a partition  $\mathcal{T}_h$  of  $\Omega$  into closed simplices which satisfies the following properties: (i) The set of points  $\mathcal{V}_h^{\text{int}}$  is contained in the set of vertices of  $\mathcal{T}_h$ , denoted by  $\mathcal{V}_h$ ; (ii) The vertices from  $\mathcal{V}_h^{\text{ext}} := \mathcal{V}_h \setminus \mathcal{V}_h^{\text{int}}$  lie on the boundary of  $\Omega$ ; (iii)  $\mathcal{T}_h$  is conforming, i.e. the intersection of two different simplices is either an empty set or their common vertex, edge, or face; (iv)  $\bar{\Omega} = \bigcup_{K \in \mathcal{T}_h} K$ . This definition is not unique: we have a choice in connecting the different points  $V_D \in \mathcal{V}_h^{\text{int}}$  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 (3.3) were non-negative, since this will imply the discrete maximum principle, see Remark 3.7 and Theorem 4.11 below. An example of a dual grid to a primal nonmatching grid is given in Figure 1 by the solid line. For  $K \in \mathcal{T}_h$ , we denote by  $h_K$  the diameter of  $K$ ; the simplicial mesh discretization parameter  $h_{\mathcal{T}}$  is given by  $h_{\mathcal{T}} := \max_{K \in \mathcal{T}_h} h_K$ . For a vertex  $V_D \in \mathcal{V}_h$ , we denote by  $\mathcal{M}(V_D)$  the set of all vertices  $V_E \in \mathcal{V}_h$  such that there exists an edge of the dual grid  $\mathcal{T}_h$  between  $V_D$  and  $V_E$ . We finally make the following shape regularity assumption on the family of dual meshes  $\{\mathcal{T}_h\}_h$ , denoting  $\kappa_K := |K|/h_K^d$ :

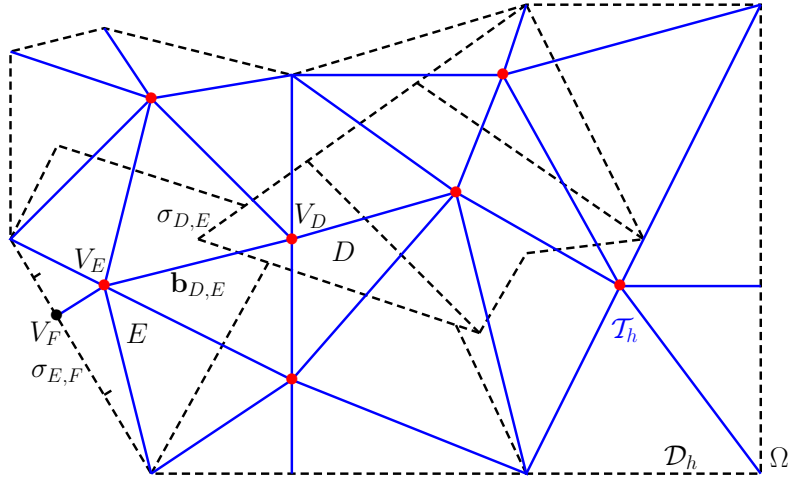


Figure 1: Primal nonmatching grid  $\mathcal{D}_h$  (dashed) and dual triangular grid  $\mathcal{T}_h$  (solid) with  $D, E \in \mathcal{D}_h$ ,  $V_D, V_E \in \mathcal{V}_h^{\text{int}}$ ,  $V_F \in \mathcal{V}_h^{\text{ext}}$ ,  $\sigma_{D,E} = \partial D \cap \partial E \in \mathcal{F}_h^{\text{int}}$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

*Assumption (C) (Shape regularity of the dual meshes)*

There exists a constant  $\kappa_{\mathcal{T}} > 0$  such that

$$\min_{K \in \mathcal{T}_h} \kappa_K \geq \kappa_{\mathcal{T}} \quad \forall h > 0.$$

Let  $\rho_K$  denote the diameter of the largest ball inscribed in  $K \in \mathcal{T}_h$ . Then in view of the inequalities  $|K| \geq \text{diam}(K)^{d-1} \rho_K / (d-1)d$ ,  $|K| \leq (d+1) \text{diam}(K)^{d-1} \rho_K / (d-1)d$  following from geometrical properties of a triangle (tetrahedron)  $K$ , Assumption (C) is equivalent to the more common requirement of the existence of a constant  $\theta_{\mathcal{T}} > 0$  such that

$$\max_{K \in \mathcal{T}_h} \frac{\text{diam}(K)}{\rho_K} \leq \theta_{\mathcal{T}} \quad \forall h > 0. \quad (3.1)$$

We finally denote  $h := \max\{h_{\mathcal{D}}, h_{\mathcal{T}}\}$  and make the following simultaneous assumption on the meshes  $\mathcal{D}_h$  and  $\mathcal{T}_h$ :

*Assumption (D) (Primal and dual meshes)*

(D1) There exists a constant  $C_{1,\mathcal{D},\mathcal{T}} > 0$  such that for all  $\sigma_{D,E} \in \mathcal{F}_h$ ,

$$|\sigma_{D,E}| |\mathbf{b}_{D,E} \cap K| \leq C_{1,\mathcal{D},\mathcal{T}} |K| \quad \text{for all } K \in \mathcal{T}_h \text{ such that } \mathbf{b}_{D,E} \cap K \neq \emptyset;$$

(D2) There exists a constant  $C_{2,\mathcal{D},\mathcal{T}} \geq 0$  such that each  $K \in \mathcal{T}_h$  is intersected by at most  $C_{2,\mathcal{D},\mathcal{T}}$  segments  $\mathbf{b}_{D,E}$  for some  $\sigma_{D,E} \in \mathcal{F}_h$ .

Assumption (D1) expresses that the measure of each “side”  $\sigma_{D,E} \in \mathcal{F}_h$  is bounded with respect to the measure of all  $K \in \mathcal{T}_h$  “between” the vertices  $V_D$  and  $V_E$  and the length of the intersection of the line segment connecting  $V_D$  and  $V_E$  and  $K$ . Together with Assumption (D2), it is satisfied for any “reasonable” dual mesh  $\mathcal{T}_h$  to  $\mathcal{D}_h$ .

**Remark 3.1** (Terminology). We remark that in correspondence with the order of the construction of the meshes, the terminology is reversed here with respect to standard finite volume–finite element schemes [18, 20], where the simplicial mesh is termed primal, whereas the mesh for the finite volume part of the scheme is termed dual.

We define the following finite-dimensional spaces:

$$\begin{aligned} X_h &:= \{ \varphi_h \in C(\Omega); \varphi_h|_K \text{ is linear } \forall K \in \mathcal{T}_h \}, \\ X_h^0 &:= \{ \varphi_h \in X_h; \varphi_h(V_D) = 0 \quad \forall V_D \in \mathcal{V}_h^{\text{ext}} \}. \end{aligned}$$

The basis of  $X_h$  is spanned by the standard finite element shape functions  $\varphi_{V_D}$ ,  $V_D \in \mathcal{V}_h$ , such that  $\varphi_{V_D}(V_E) = \delta_{DE}$ ,  $V_E \in \mathcal{V}_h$ ,  $\delta$  being the Kronecker delta. We have the following lemma:

**Lemma 3.2.** *Let  $c_D = 0$  for all  $V_D \in \mathcal{V}_h^{\text{ext}}$  and let  $c_h = \sum_{V_D \in \mathcal{V}_h} c_D \varphi_{V_D} \in X_h^0$ . Then*

$$\sum_{\sigma_{D,E} \in \mathcal{F}_h} \frac{|\sigma_{D,E}|}{d_{D,E}} (c_E - c_D)^2 \leq C_{1,D,T} C_{2,D,T} \|\nabla c_h\|_{\Omega}^2.$$

PROOF:

Let  $\sigma_{D,E} \in \mathcal{F}_h$  and recall that  $\mathbf{b}_{D,E}$  is the line segment connecting the vertices  $V_D$  and  $V_E$ . There holds

$$|c_E - c_D| \leq \sum_{K \in \mathcal{T}_h} \chi_{K,D,E} |c_{2,K,D,E} - c_{1,K,D,E}|,$$

where, for  $K \in \mathcal{T}_h$ ,  $\chi_{K,D,E} = 1$  if  $\mathbf{b}_{D,E} \cap K \neq \emptyset$  and 0 otherwise and where  $c_{1,K,D,E}$  is the value of  $c_h$  at the point where  $\mathbf{b}_{D,E}$  for the first time intersects (or “enters”)  $K$  and  $c_{2,K,D,E}$  is the value of  $c_h$  at the point where  $\mathbf{b}_{D,E}$  for the second time intersects (or “leaves”)  $K$ . In the simplest case where  $\mathbf{b}_{D,E}$  is an edge of the simplicial mesh  $\mathcal{T}_h$ ,  $c_E - c_D = c_{2,K,D,E} - c_{1,K,D,E}$  for any of the simplices  $K$  sharing this edge; we in such cases only put  $\chi_{K,D,E} = 1$  for one  $K \in \mathcal{T}_h$ . Note that the above inequality holds true in the general case as well and in particular also if, for a nonconvex  $\Omega$ ,  $\mathbf{b}_{D,E}$  “leaves” and “reenters”  $\Omega$  (in this case, the fact that  $c_h$  is zero on  $\partial\Omega$  is important). Hence, by virtue of the Cauchy–Schwarz inequality and of the fact that  $|c_{2,K,D,E} - c_{1,K,D,E}| \leq |\nabla c_h|_K| |\mathbf{b}_{D,E} \cap K|$ ,

$$|c_E - c_D|^2 \leq \sum_{K \in \mathcal{T}_h} \chi_{K,D,E} |\mathbf{b}_{D,E} \cap K| \sum_{K \in \mathcal{T}_h} \chi_{K,D,E} |\mathbf{b}_{D,E} \cap K| |\nabla c_h|_K^2.$$

Next,

$$\sum_{K \in \mathcal{T}_h} \chi_{K,D,E} |\mathbf{b}_{D,E} \cap K| \leq d_{D,E}$$

(the inequality can appear at the boundary for nonconvex  $\Omega$ ), so that

$$\sum_{\sigma_{D,E} \in \mathcal{F}_h} \frac{|\sigma_{D,E}|}{d_{D,E}} (c_E - c_D)^2 \leq \sum_{\sigma_{D,E} \in \mathcal{F}_h} \sum_{K \in \mathcal{T}_h} \chi_{K,D,E} |\sigma_{D,E}| |\mathbf{b}_{D,E} \cap K| |\nabla c_h|_K^2 \leq C_{1,D,T} C_{2,D,T} \|\nabla c_h\|_{\Omega}^2,$$

using Assumption (D).

### 3.2 Time discretization

We suppose the partition of the time interval  $(0, T)$  such that  $0 = t_0 < \dots < t_n < \dots < t_N = T$  and define  $\Delta t_n := t_n - t_{n-1}$  and  $\Delta t := \max_{1 \leq n \leq N} \Delta t_n$ . In the case where Assumption (A5) is satisfied we do not impose any restriction on the time step. When only Assumption (A6) is satisfied, we suppose in addition:

*Assumption (E) (Maximum time step for decreasing  $F$ )*

*The following maximum time step condition is satisfied:*

$$\Delta t < \frac{c_\beta}{L_F}.$$

### 3.3 The combined scheme

We are now ready to define the combined scheme.

**Definition 3.3** (Combined scheme). *The fully implicit combined finite volume–finite element scheme for the problem (1.1)–(2.2) on the grids  $\mathcal{D}_h$  and  $\mathcal{T}_h$  reads: find the values  $c_D^n$ ,  $V_D \in \mathcal{V}_h$ ,  $n \in \{0, 1, \dots, N\}$ , such that*

$$c_D^0 = \frac{1}{|D|} \int_D c_0(\mathbf{x}) \, d\mathbf{x} \quad D \in \mathcal{D}_h, \quad (3.2a)$$

$$c_D^n = 0 \quad V_D \in \mathcal{V}_h^{\text{ext}}, \quad n \in \{1, \dots, N\}, \quad (3.2b)$$

$$\frac{\beta(c_D^n) - \beta(c_D^{n-1})}{\Delta t_n} |D| - \sum_{E \in \mathcal{D}_h} \mathbb{S}_{D,E}^n c_E^n + \sum_{\sigma_{D,E} \in \mathcal{F}_D} v_{D,E}^n \overline{c_{D,E}^n} + F(c_D^n) |D| = q_D^n |D| \quad D \in \mathcal{D}_h, \quad n \in \{1, \dots, N\}. \quad (3.2c)$$

In (3.2a)–(3.2c) we have denoted

$$v_{D,E}^n := \frac{1}{\Delta t_n} \int_{t_{n-1}}^{t_n} \int_{\sigma_{D,E}} \mathbf{v}(\mathbf{x}, t) \cdot \mathbf{n}_{D,E} \, d\gamma(\mathbf{x}) \, dt$$

for  $D \in \mathcal{D}_h$ ,  $\sigma_{D,E} \in \mathcal{F}_D$ , and  $n \in \{1, \dots, N\}$ , with  $\mathbf{n}_{D,E}$  the unit normal vector of  $\partial D$ , outward to  $D$ , and

$$q_D^n := \frac{1}{\Delta t_n |D|} \int_{t_{n-1}}^{t_n} \int_D q(\mathbf{x}, t) \, d\mathbf{x} \, dt \quad D \in \mathcal{D}_h, \quad n \in \{1, \dots, N\}.$$

Next, we first set

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

The elements  $\mathbb{S}_{D,E}^n$  of the finite element diffusion matrix are then given by

$$\mathbb{S}_{D,E}^n := - \sum_{K \in \mathcal{T}_h} (\mathbf{S}^n \nabla \varphi_{V_E}, \nabla \varphi_{V_D})_K \quad V_D, V_E \in \mathcal{V}_h, \quad n \in \{1, \dots, N\}, \quad (3.3)$$

where

$$\mathbf{S}^n(\mathbf{x}) = \tilde{\mathbf{S}}^n(\mathbf{x}) \quad n \in \{1, \dots, N\}, \quad \mathbf{x} \in \Omega. \quad (3.4)$$

An alternative choice is to define

$$\mathbf{S}^n(\mathbf{y}) = \left( \frac{1}{|K|} \int_K [\tilde{\mathbf{S}}^n(\mathbf{x})]^{-1} \, d\mathbf{x} \right)^{-1} \quad \mathbf{y} \in K, \quad K \in \mathcal{T}_h, \quad n \in \{1, \dots, N\}. \quad (3.5)$$

In fact, the terms  $\mathbb{S}_{D,E}^n$  for  $V_D \in \mathcal{V}_h^{\text{ext}}$  or  $V_E \in \mathcal{V}_h^{\text{ext}}$  do not occur in the scheme (3.2a)–(3.2c). It will however show convenient to define these values. Finally, we define  $\overline{c_{D,E}^n}$  for  $D \in \mathcal{D}_h$ ,  $\sigma_{D,E} \in \mathcal{F}_D$ , and  $n \in \{1, \dots, N\}$  as follows:

$$\overline{c_{D,E}^n} := \begin{cases} c_D^n + \alpha_{D,E}^n (c_E^n - c_D^n) & \text{if } v_{D,E}^n \geq 0 \\ c_E^n + \alpha_{D,E}^n (c_D^n - c_E^n) & \text{if } v_{D,E}^n < 0 \end{cases}. \quad (3.6)$$

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

$$\alpha_{D,E}^n := \begin{cases} \frac{\max\left\{\min\left\{\mathbb{S}_{D,E}^n, \frac{1}{2}|v_{D,E}^n|\right\}, 0\right\}}{|v_{D,E}^n|} & \text{if } v_{D,E}^n \neq 0 \text{ and } \sigma_{D,E} \in \mathcal{F}_h^{\text{int}} \\ & \text{or if } \sigma_{D,E} \in \mathcal{F}_h^{\text{ext}} \text{ and } v_{D,E}^n > 0 \\ 0 & \text{if } v_{D,E}^n = 0 \text{ or if } \sigma_{D,E} \in \mathcal{F}_h^{\text{ext}} \text{ and } v_{D,E}^n < 0 \end{cases} \quad (3.7)$$

We remark that  $\overline{c_{D,E}^n} = \widehat{c_{D,E}^n} + \text{sign}(v_{D,E}^n)\alpha_{D,E}^n(c_E^n - c_D^n)$ , where  $\widehat{c_{D,E}^n}$  stands for full upstream weighting.

In the sequel we shall consider apart the following special case:

*Assumption (F) (Diffusion matrix)*

*All transmissibilities are non-negative, i.e.*

$$\mathbb{S}_{D,E}^n \geq 0 \quad \forall D \in \mathcal{D}_h, V_E \in \mathcal{M}(V_D) \quad \forall n \in \{1, \dots, N\}.$$

### 3.4 Remarks

We end this section by several remarks on the scheme and its possible extensions.

**Remark 3.4** (Comparison with standard cell-centered finite volume schemes). *We recall that the scheme (3.2a)–(3.2c) only differs from standard cell-centered finite volume schemes (cf. [14]) in the discretization of the diffusion term (and in the treatment of the boundary condition). This difference is however essential: not only that it allows the discretization on grids not necessarily satisfying the “orthogonality condition” (see [14, Definition 9.1.(iv)]) and of full diffusion–dispersion tensors, but the family of admissible grids is virtually unlimited and in particular comprises the grids with nonmatching interfaces.*

**Remark 3.5** (Numerical flux). *We can easily see from (3.7) that  $0 \leq \alpha_{D,E}^n \leq 1/2$ , i.e. the numerical flux defined by (3.6) ranges from the full upstream weighting to the centered scheme. The amount of upstream weighting is set with respect to the local proportion of convection and diffusion.*

**Remark 3.6** (Arithmetic versus harmonic averaging). *We remark that the choice (3.4) for  $\mathbf{S}^n$  corresponds to the arithmetic average of the diffusion–dispersion tensor  $\mathbf{S}$  in space, whereas the choice (3.5) corresponds to the harmonic one.*

**Remark 3.7** (Discrete maximum principle). *We shall see in Theorem 4.11 below that the discrete maximum principle for the combined scheme holds under Assumption (F). This is e.g. the case, in two space dimensions, when  $\mathbf{S}$  reduces to a constant scalar function and when  $\mathcal{T}_h$  is Delaunay, that is the circumcircle of each triangle does not contain any vertex in its interior, and under the additional condition that no circumcenters of boundary triangles lie outside the domain, cf. [21, 28]. Remark that given a set of points, we can always construct a Delaunay triangulation. In three space dimensions, however, a Delaunay tetrahedral mesh in general does not guarantee the non-negativity of the finite element transmissibilities, cf. [25, 28]. We refer to Remark 3.10 for the modification of the proposed scheme, which guarantees the discrete maximum principle in both two and three space dimensions under the condition that  $\mathbf{S}$  is a constant scalar function.*

**Remark 3.8** (Dual Delaunay triangulation for a locally refined square grid). *Let us consider a locally refined square grid, where a square is refined into 9 subsquares and where the difference of levels of refinement of two neighboring squares is at most one, such as that given in Figure 2 in Section 7 below by the dashed line. Then an example of a dual Delaunay triangulation is given in Figure 2 by the solid line.*

**Remark 3.9** (Necessity to construct the dual mesh). *In the otherwise straightforward implementation of the combined scheme, the only complex point may be the construction of the dual mesh  $\mathcal{T}_h$  and the evaluation of the elements  $\mathbb{S}_{D,E}^n$  of the finite element diffusion matrix, especially when  $\mathbf{S}$  is nonconstant. This may however not always be necessary: recall for example that the finite volume stiffness matrix of a Laplacian on a Voronoï mesh and its finite element counterpart on a dual Delaunay triangulation in two space dimensions coincide, see [14, Section III.12.2]. So in particular on square meshes of the type of Figure 2 below, we may use the cell-centered finite volume discretization in the unrefined parts and only construct the dual mesh at the nonmatching refined interfaces.*

**Remark 3.10** (A two-grid finite volume scheme verifying the discrete maximum principle for  $\mathbf{S}$  scalar). *When  $\mathbf{S}$  is a scalar function, we can in spirit of the previous remark replace the discretization of the diffusion term by the finite element method on a dual simplicial grid by a finite volume discretization on a Voronoï grid given by the points from  $\mathcal{V}_h$ . Whereas in two space dimensions for  $\mathcal{T}_h$  a Delaunay triangulation and under the additional assumption that  $\mathbf{S}$  is constant, this would lead to the same scheme, in three space dimensions, such a scheme would verify the discrete maximum principle, since the finite volume discretization on a Voronoï grid still leads to positive transmissibilities; compare this with Remark 3.7.*

## 4 Existence, uniqueness, and discrete properties

In this section we first present some technical lemmas. We then show the conservativity of the scheme, the coercivity and boundedness of the bilinear diffusion form corresponding to the diffusion term, and an a priori estimate for an extended scheme, which is needed later in the proof of the existence of the solution of the discrete problem. Finally, we show the uniqueness of this solution and prove the discrete maximum principle when Assumption (F) is satisfied.

### 4.1 Discrete properties of the scheme

**Lemma 4.1** (Finite element diffusion matrix). *For all  $V_D \in \mathcal{V}_h$  and  $n \in \{1, \dots, N\}$ ,  $\mathbb{S}_{D,D}^n = -\sum_{V_E \in \mathcal{M}(V_D)} \mathbb{S}_{D,E}^n$ .*

PROOF:

We will show the assertion for  $d = 2$ ; the case  $d = 3$  is similar. Let us consider a fixed vertex  $V_D \in \mathcal{V}_h$  and a  $K \in \mathcal{T}_h$  having  $V_D$  for vertex. We denote the two other vertices of  $K$  by  $V_E$  and  $V_F$ . Now

$$-(\mathbf{S}^n \nabla \varphi_{V_D}, \nabla \varphi_{V_D})_K = (\mathbf{S}^n \nabla \varphi_{V_E}, \nabla \varphi_{V_D})_K + (\mathbf{S}^n \nabla \varphi_{V_F}, \nabla \varphi_{V_D})_K,$$

which follows immediately by

$$-\varphi_{V_D}|_K = (\varphi_{V_E} + \varphi_{V_F})|_K - 1.$$

Since  $K \in \mathcal{T}_h$  creating a support for  $\varphi_{V_D}$  was arbitrary and noticing the definition (3.3) of  $\mathbb{S}_{D,E}^n$ , this concludes the proof.  $\square$

**Corollary 4.2** (Equivalent form of the diffusion term). *Let  $V_D \in \mathcal{V}_h$ . Using the fact that  $\mathbb{S}_{D,E}^n \neq 0$  only if  $V_E \in \mathcal{M}(V_D)$  or if  $E = D$  and Lemma 4.1, one has*

$$\sum_{V_E \in \mathcal{V}_h} \mathbb{S}_{D,E}^n c_E^n = \sum_{V_E \in \mathcal{M}(V_D)} \mathbb{S}_{D,E}^n c_E^n + \mathbb{S}_{D,D}^n c_D^n = \sum_{V_E \in \mathcal{M}(V_D)} \mathbb{S}_{D,E}^n (c_E^n - c_D^n).$$

**Theorem 4.3** (Conservativity of the scheme). *The scheme (3.2a)–(3.2c) is conservative with respect to the mesh  $\mathcal{D}_h$ .*

PROOF:

The equation (3.2c) defining the scheme and Corollary 4.2 imply that the combined finite volume–finite element scheme is conservative on each time level and on each cell of the mesh  $\mathcal{D}_h$  as the pure finite volume is, cf. [14].  $\square$

**Lemma 4.4** (Equivalent form of the upwind part of the convection term). *For all  $D \in \mathcal{D}_h$  and  $n \in \{1, \dots, N\}$ ,*

$$\sum_{\sigma_{D,E} \in \mathcal{F}_D} v_{D,E}^n \widehat{c_{D,E}^n} = \sum_{\sigma_{D,E} \in \mathcal{F}_D} (v_{D,E}^n)^- (c_E^n - c_D^n) + r_D^n c_D^n |D|,$$

where  $(v_{D,E}^n)^- := \min\{v_{D,E}^n, 0\}$  and

$$r_D^n := \frac{1}{\Delta t_n |D|} \int_{t_{n-1}}^{t_n} \int_D r(\mathbf{x}, t) \, d\mathbf{x} \, dt \quad \forall D \in \mathcal{D}_h, \forall n \in \{1, \dots, N\}.$$

The assertion of this lemma is a simple consequence of Assumption (A4). The proof can be found in [32].

**Lemma 4.5** (Coercivity of the diffusion form). *For all  $c_h = \sum_{V_D \in \mathcal{V}_h} c_D \varphi_{V_D} \in X_h$  and  $n \in \{1, \dots, N\}$ ,*

$$- \sum_{V_D \in \mathcal{V}_h} c_D \sum_{V_E \in \mathcal{V}_h} \mathbb{S}_{D,E}^n c_E \geq c_S \|\nabla c_h\|_{\Omega}^2.$$

PROOF:

We have

$$- \sum_{V_D \in \mathcal{V}_h} c_D \sum_{V_E \in \mathcal{V}_h} \mathbb{S}_{D,E}^n c_E = \sum_{K \in \mathcal{T}_h} (\mathbf{S}^n \nabla c_h, \nabla c_h)_K \geq c_S \|\nabla c_h\|_{\Omega}^2,$$

using (3.3) and Assumption (A3) and the subsequent uniform positive definiteness of the diffusion tensors (3.4) and (3.5).  $\square$

**Lemma 4.6** (Boundedness of the diffusion form). *For all  $c_h = \sum_{V_D \in \mathcal{V}_h} c_D \varphi_{V_D} \in X_h$ ,  $u_h = \sum_{V_D \in \mathcal{V}_h} u_D \varphi_{V_D} \in X_h$ , and  $n \in \{1, \dots, N\}$ ,*

$$\left| - \sum_{V_D \in \mathcal{V}_h} c_D \sum_{V_E \in \mathcal{V}_h} \mathbb{S}_{D,E}^n u_E \right| \leq C_S \|\nabla c_h\|_{\Omega} \|\nabla u_h\|_{\Omega}.$$

PROOF:

We have

$$\left| - \sum_{V_D \in \mathcal{V}_h} c_D \sum_{V_E \in \mathcal{V}_h} \mathbb{S}_{D,E}^n u_E \right| = \left| \sum_{K \in \mathcal{T}_h} (\mathbf{S}^n \nabla c_h, \nabla u_h)_K \right| \leq C_S \|\nabla c_h\|_{\Omega} \|\nabla u_h\|_{\Omega},$$

using (3.3) and Assumption (A3).  $\square$

**Lemma 4.7** (Estimate on the convection term). *For all values  $c_D, V_D \in \mathcal{V}_h$ , such that  $c_D = 0$  for all  $V_D \in \mathcal{V}_h^{\text{ext}}$  and  $n \in \{1, \dots, N\}$ ,*

$$\sum_{D \in \mathcal{D}_h} c_D \sum_{\sigma_{D,E} \in \mathcal{F}_D} v_{D,E}^n \overline{c_{D,E}} \geq 0.$$

PROOF:

We can write

$$\begin{aligned} & \sum_{D \in \mathcal{D}_h} c_D \sum_{\sigma_{D,E} \in \mathcal{F}_D} v_{D,E}^n \overline{c_{D,E}} \\ = & \sum_{\sigma_{D,E} \in \mathcal{F}_h^{\text{int}}, v_{D,E}^n \geq 0} v_{D,E}^n (c_D(c_D - c_E) - \alpha_{D,E}^n (c_E - c_D)^2) + \sum_{\sigma_{D,E} \in \mathcal{F}_h^{\text{ext}}} c_D v_{D,E}^n \overline{c_{D,E}} \\ = & \frac{1}{2} \sum_{\sigma_{D,E} \in \mathcal{F}_h^{\text{int}}, v_{D,E}^n \geq 0} v_{D,E}^n (c_D^2 - c_E^2) + \sum_{\sigma_{D,E} \in \mathcal{F}_h^{\text{int}}} |v_{D,E}^n| (c_E - c_D)^2 \left( \frac{1}{2} - \alpha_{D,E}^n \right) \\ & + \sum_{\sigma_{D,E} \in \mathcal{F}_h^{\text{ext}}} \left\{ \frac{1}{2} v_{D,E}^n c_D^2 + |v_{D,E}^n| c_D^2 \left( \frac{1}{2} - \alpha_{D,E}^n \right) \right\} \geq \frac{1}{2} \sum_{D \in \mathcal{D}_h} c_D^2 \sum_{\sigma_{D,E} \in \mathcal{F}_D} v_{D,E}^n \\ = & \frac{1}{2} \sum_{D \in \mathcal{D}_h} c_D^2 r_D^n |D| \geq 0, \end{aligned}$$

where we have used the relation  $2(a-b)a = (a-b)^2 + a^2 - b^2$ , rewritten the summation over interior  $\sigma_{D,E}$  with fixed denotation of the control volumes sharing given set  $\sigma_{D,E}$  such that  $v_{D,E}^n \geq 0$  and over exterior  $\sigma_{D,E}$ , and the fact that  $c_D = 0$  for all  $V_D \in \mathcal{V}_h^{\text{ext}}$ . In the last two estimates we have used, respectively, the fact that  $0 \leq \alpha_{D,E}^n \leq 1/2$ , which follows from (3.7), and Assumption (A4).  $\square$

**Theorem 4.8** (A priori estimate for an extended scheme). *Let us define an extended scheme by*

$$c_D^0 = \frac{1}{|D|} \int_D c_0(\mathbf{x}) \, d\mathbf{x} \quad D \in \mathcal{D}_h, \quad (4.1a)$$

$$c_D^n = 0 \quad V_D \in \mathcal{V}_h^{\text{ext}}, \quad n \in \{1, \dots, N\}, \quad (4.1b)$$

$$\begin{aligned} & u \frac{\beta(c_D^n) - \beta(c_D^{n-1})}{\Delta t_n} |D| - \sum_{E \in \mathcal{D}_h} \mathbb{S}_{D,E}^n c_E^n + u \sum_{\sigma_{D,E} \in \mathcal{F}_D} v_{D,E}^n \overline{c_{D,E}^n} + u F(c_D^n) |D| \\ & = u q_D^n |D| \quad D \in \mathcal{D}_h, \quad n \in \{1, \dots, N\} \end{aligned} \quad (4.1c)$$

with  $u \in [0, 1]$ . Then

$$\sum_{D \in \mathcal{D}_h} (c_D^n)^2 |D| \leq C_{\text{es}} \quad \forall n \in \{1, \dots, N\}$$

with

$$C_{\text{es}} := \frac{8}{c_\beta} M \beta(M) |\Omega| + \frac{16T}{c_\beta^2} \|q\|_{Q_T}^2 + \frac{8}{c_\beta} L_F M^2 T |\Omega|.$$

PROOF:

The proof closely follows the proof of [18, Theorem 4.2]. We however detail it for the sake of completeness.

We multiply (4.1c) by  $\Delta t_n c_D^n$ , sum over all  $D \in \mathcal{D}_h$  and  $n \in \{1, \dots, k\}$ , and use the fact that  $u \geq 0$  and Lemmas 4.5 and 4.7. Further, for  $c_D^n < 0$  or  $c_D^n > M$ ,  $F(c_D^n)c_D^n \geq 0$  follows from Assumption (A5) or (A6). When  $0 \leq c_D^n \leq M$ ,  $-F(c_D^n)c_D^n \leq |F(c_D^n)||c_D^n| \leq L_F M^2$ , which altogether yields

$$\begin{aligned} & u \sum_{n=1}^k \sum_{D \in \mathcal{D}_h} [\beta(c_D^n) - \beta(c_D^{n-1})] c_D^n |D| + c_S \sum_{n=1}^k \Delta t_n \|\nabla c_h^n\|_{\Omega}^2 \\ & \leq u \sum_{n=1}^k \Delta t_n \sum_{D \in \mathcal{D}_h} c_D^n q_D^n |D| + u L_F M^2 \sum_{n=1}^k \sum_{D \in \mathcal{D}_h} \Delta t_n |D| \end{aligned} \quad (4.2)$$

with  $c_h^n = \sum_{D \in \mathcal{D}_h} c_D^n \varphi_{V_D}$ . Let us now introduce a function  $B$ ,

$$B(s) := \beta(s)s - \int_0^s \beta(\tau) \, d\tau \quad s \in \mathbb{R}.$$

One then can derive

$$B(c_D^n) - B(c_D^{n-1}) = [\beta(c_D^n) - \beta(c_D^{n-1})] c_D^n - \int_{c_D^{n-1}}^{c_D^n} [\beta(\tau) - \beta(c_D^{n-1})] \, d\tau.$$

Using that  $\beta$  is nondecreasing, one can easily show that

$$\int_{c_D^{n-1}}^{c_D^n} [\beta(\tau) - \beta(c_D^{n-1})] \, d\tau \geq 0.$$

In view of the two last expressions, one has

$$\sum_{n=1}^k \sum_{D \in \mathcal{D}_h} [B(c_D^n) - B(c_D^{n-1})] |D| \leq \sum_{n=1}^k \sum_{D \in \mathcal{D}_h} [\beta(c_D^n) - \beta(c_D^{n-1})] c_D^n |D|,$$

which yields

$$\sum_{D \in \mathcal{D}_h} B(c_D^k) |D| - \sum_{D \in \mathcal{D}_h} B(c_D^0) |D| \leq \sum_{n=1}^k \sum_{D \in \mathcal{D}_h} [\beta(c_D^n) - \beta(c_D^{n-1})] c_D^n |D|.$$

Using the growth condition on  $\beta$  from Assumption (A1), one can derive  $B(s) \geq s^2 c_\beta / 2$  for all  $s \in \mathbb{R}$ , see [18, Lemma 8.2]. Thus, using in addition Assumption (A8)

$$\frac{c_\beta}{2} \sum_{D \in \mathcal{D}_h} (c_D^k)^2 |D| - M \beta(M) |\Omega| \leq \sum_{n=1}^k \sum_{D \in \mathcal{D}_h} [\beta(c_D^n) - \beta(c_D^{n-1})] c_D^n |D|.$$

We notice that

$$\sum_{n=1}^N \sum_{D \in \mathcal{D}_h} \Delta t_n |D| (q_D^n)^2 \leq \|q\|_{Q_T}^2 \quad (4.3)$$

by the Cauchy–Schwarz inequality. Hence extending the summation over all  $n \in \{1, \dots, N\}$  in the first term of the right-hand side of (4.2) and using the Cauchy–Schwarz and Young inequality, we

have

$$\begin{aligned} \sum_{n=1}^k \Delta t_n \sum_{D \in \mathcal{D}_h} c_D^n q_D^n |D| &\leq \left( \sum_{n=1}^N \Delta t_n \sum_{D \in \mathcal{D}_h} (c_D^n)^2 |D| \right)^{\frac{1}{2}} \|q\|_{Q_T} \\ &\leq \frac{\varepsilon}{2} \sum_{n=1}^N \Delta t_n \sum_{D \in \mathcal{D}_h} (c_D^n)^2 |D| + \frac{1}{2\varepsilon} \|q\|_{Q_T}^2. \end{aligned}$$

Hence, substituting these estimates into (4.2), we obtain

$$\begin{aligned} u \frac{c_\beta}{2} \max_{n \in \{1, \dots, N\}} \sum_{D \in \mathcal{D}_h} (c_D^n)^2 |D| + c_s \sum_{n=1}^N \Delta t_n \|\nabla c_h^n\|_\Omega^2 &\leq 2uM\beta(M)|\Omega| \\ + u\varepsilon T \max_{n \in \{1, \dots, N\}} \sum_{D \in \mathcal{D}_h} (c_D^n)^2 |D| + u \frac{1}{\varepsilon} \|q\|_{Q_T}^2 + 2uL_F M^2 T |\Omega|, \end{aligned} \quad (4.4)$$

considering the fact that  $k$  was arbitrarily chosen. We now put  $\varepsilon = c_\beta/(4T)$ . When  $u \neq 0$ , this already yields the assertion of the lemma. When  $u = 0$ , it follows from (4.4) that  $c_D^n = 0$  for all  $D \in \mathcal{D}_h$  and all  $n \in \{1, \dots, N\}$ , since in view of (4.1b),  $\|\nabla(\cdot)\|_\Omega$  is a norm on  $X_h$ . Thus the assertion of the lemma is trivially satisfied in this case.  $\square$

## 4.2 Existence, uniqueness, and the discrete maximum principle

The proof of the following theorem follows the ideas of the proof given in [17]. It makes use of the a priori estimate for the extended scheme given by Theorem 4.8 and of the (Brouwer) topological degree argument:

**Theorem 4.9** (Existence of the solution of the discrete problem). *The problem (3.2a)–(3.2c) has at least one solution.*

We state the uniqueness without a proof as well (cf. [18, Theorem 4.4]):

**Theorem 4.10** (Uniqueness of the solution of the discrete problem). *The solution of the problem (3.2a)–(3.2c) is unique.*

In standard cell-centered finite volume schemes (cf. [14]), as well as in the combined finite volume–nonconforming finite element scheme [18], there can be nonzero convective and diffusive fluxes between  $D$  and  $E$  only if the control volumes  $D$  and  $E$  neighbors. This is however not the case with the scheme (3.2a)–(3.2c): there can be a nonzero convective flux between  $D$  and  $E$  only if  $D$  and  $E$  neighbors, but there can be a nonzero diffusive flux between  $D$  and  $E$  even if  $D$  and  $E$  are not neighbors (because the transmissibility between  $D$  and  $E$  is given by the grid  $\mathcal{T}_h$ ). The following theorem shows that the local Péclet upstream weighting (3.6) still guarantees, adding minimal numerical diffusion, the stability of the scheme:

**Theorem 4.11** (Discrete maximum principle). *Under Assumption (F), the solution of the problem (3.2a)–(3.2c) satisfies*

$$0 \leq c_D^n \leq M$$

for all  $D \in \mathcal{D}_h$ ,  $n \in \{1, \dots, N\}$ .

PROOF:

Setting  $\mathbb{T}_{D,E}^n := \mathbb{S}_{D,E}^n - |v_{D,E}^n| \alpha_{D,E}^n$ ,  $D \in \mathcal{D}_h$ ,  $E \in \mathcal{L}(D)$ , where  $E \in \mathcal{L}(D)$  if  $\sigma_{D,E} \in \mathcal{F}_D$  or if  $V_E \in \mathcal{M}(V_D)$ , and using Corollary 4.2 and Lemma 4.4, we can rewrite (3.2c) as

$$\begin{aligned} & \frac{\beta(c_D^n) - \beta(c_D^{n-1})}{\Delta t_n} |D| - \sum_{E \in \mathcal{L}(D)} \mathbb{T}_{D,E}^n (c_E^n - c_D^n) + \sum_{\sigma_{D,E} \in \mathcal{F}_D} (v_{D,E}^n)^- (c_E^n - c_D^n) \\ & + r_D^n c_D^n |D| + F(c_D^n) |D| = q_D^n |D| \quad D \in \mathcal{D}_h, n \in \{1, \dots, N\}. \end{aligned} \quad (4.5)$$

In view of Assumption (F) and of (3.7), one has  $\mathbb{T}_{D,E}^n \geq 0$  for all  $D \in \mathcal{D}_h$ ,  $E \in \mathcal{L}(D)$ , and  $n \in \{1, \dots, N\}$ . We now make use of an induction argument. We remark that  $0 \leq c_D^n \leq M$  is satisfied for  $n = 0$  by Assumption (A8) and (3.2a). Let us suppose that  $0 \leq c_D^{n-1} \leq M$  for all  $D \in \mathcal{D}_h$  and for a fixed  $(n-1) \in \{0, 1, \dots, N-1\}$ . Since  $|\mathcal{D}_h|$  is finite, there exist  $D_0, D_1 \in \mathcal{D}_h$  such that  $c_{D_0}^n \leq c_D^n \leq c_{D_1}^n$  for all  $D \in \mathcal{D}_h$ . Using a contradiction argument we prove below that  $c_{D_0}^n \geq 0$  and  $c_{D_1}^n \leq M$ . Suppose that  $c_{D_0}^n < 0$ . Then, since  $\mathbb{T}_{D_0,E}^n \geq 0$  and  $-(v_{D_0,E}^n)^- \geq 0$ , we have

$$\sum_{E \in \mathcal{L}(D_0)} \mathbb{T}_{D_0,E}^n (c_E^n - c_{D_0}^n) + \sum_{\sigma_{D_0,E} \in \mathcal{F}_{D_0}} -(v_{D_0,E}^n)^- (c_E^n - c_{D_0}^n) \geq 0.$$

This yields, using (4.5),

$$\frac{\beta(c_{D_0}^n) - \beta(c_{D_0}^{n-1})}{\Delta t_n} |D_0| + r_{D_0}^n c_{D_0}^n |D_0| + F(c_{D_0}^n) |D_0| - q_{D_0}^n |D_0| \geq 0.$$

Now  $c_{D_0}^n < 0$  implies  $r_{D_0}^n c_{D_0}^n \leq 0$  and  $F(c_{D_0}^n) \leq 0$  using, respectively, Assumption (A4) and (A5) or (A6). Also  $-q_{D_0}^n \leq 0$ , using Assumption (A7). Hence  $\beta(c_{D_0}^n) \geq \beta(c_{D_0}^{n-1})$ , which is a contradiction, since  $\beta$  is strictly increasing from Assumption (A1).

Let us now suppose  $c_{D_1}^n > M$ . Similarly as in the previous case, one comes to

$$\frac{\beta(c_{D_1}^n) - \beta(c_{D_1}^{n-1})}{\Delta t_n} |D_1| + r_{D_1}^n c_{D_1}^n |D_1| + F(c_{D_1}^n) |D_1| - q_{D_1}^n |D_1| \leq 0.$$

We can estimate

$$-q_{D_1}^n |D_1| \geq -M r_{D_1}^n |D_1| \geq -c_{D_1}^n r_{D_1}^n |D_1|$$

using, respectively, Assumption (A7) and (A4). It follows from (A5) or (A6) that  $F(c_{D_1}^n) \geq 0$ . This implies  $\beta(c_{D_1}^n) \leq \beta(c_{D_1}^{n-1})$ , which is again a contradiction, using Assumption (A1).  $\square$

## 5 A priori estimates

In this section we give a priori estimates and estimates on differences of time and space translates of the approximate solutions that we shall define.

### 5.1 Discrete energy-type estimates

We now give energy-type a priori estimates for the approximate solution values  $c_D^n$ ,  $D \in \mathcal{D}_h$ ,  $n \in \{1, \dots, N\}$ .

**Theorem 5.1** (A priori estimates). *The solution of the combined scheme (3.2a)–(3.2c) satisfies*

$$c_\beta \max_{n \in \{1, \dots, N\}} \sum_{D \in \mathcal{D}_h} (c_D^n)^2 |D| \leq C_{\text{ae}}, \quad (5.1)$$

$$\max_{n \in \{1, \dots, N\}} \sum_{D \in \mathcal{D}_h} [\beta(c_D^n)]^2 |D| \leq C_{\text{ae}\beta}, \quad (5.2)$$

$$c_S \sum_{n=1}^N \Delta t_n \|\nabla c_h^n\|_\Omega^2 \leq C_{\text{ae}} \quad (5.3)$$

with  $c_h^n = \sum_{D \in \mathcal{D}_h} c_D^n \varphi_{V_D}$ ,

$$C_{\text{ae}} := 8M\beta(M)|\Omega| + \frac{16T}{c_\beta} \|q\|_{Q_T}^2 + 8L_F M^2 T |\Omega|,$$

$$C_{\text{ae}\beta} := [\beta(M)]^2 |\Omega|$$

when Assumption (F) is satisfied and only Assumption (A1) holds and

$$C_{\text{ae}\beta} := (2C_\beta^2 + 4L_\beta^2 P^2) |\Omega| + \frac{4L_\beta^2}{c_\beta} C_{\text{ae}}$$

when Assumption (F) is not satisfied but Assumption (A2) holds.

PROOF:

Estimates (5.1) and (5.3) follow immediately from (4.4) for  $\varepsilon = c_\beta/(4T)$ , since for  $u = 1$  the extended scheme (4.1a)–(4.1c) completely coincides with the scheme (3.2a)–(3.2c). To see the boundedness of the term on the left-hand side of (5.2) under Assumption (F) is immediate, using the discrete maximum principle stated by Theorem 4.11. In this case Assumption (A1) suffices. In the general case one has to use Assumption (A2) to show  $[\beta(s)]^2 \leq 2C_\beta^2 + 4L_\beta^2 P^2 + 4L_\beta^2 s^2$ , see [18, Lemma 8.3]. Hence, for all  $n \in \{1, \dots, N\}$ ,

$$\sum_{D \in \mathcal{D}_h} [\beta(c_D^n)]^2 |D| \leq (2C_\beta^2 + 4L_\beta^2 P^2) |\Omega| + 4L_\beta^2 \sum_{D \in \mathcal{D}_h} (c_D^n)^2 |D|. \quad \square$$

Using the values  $c_D^n$ ,  $D \in \mathcal{D}_h$ ,  $n \in \{0, 1, \dots, N\}$ , we now define two approximate solutions.

**Definition 5.2** (Approximate solutions). *Let the values  $c_D^n$  with  $V_D \in \mathcal{V}_h$ ,  $n \in \{0, 1, \dots, N\}$ , be the solutions to (3.2a)–(3.2c). As the approximate solutions of the problem (1.1)–(2.2) by means of the combined finite volume–finite element scheme, we understand:*

(i) *The function  $c_{h,\Delta t}$  defined by*

$$\begin{aligned} c_{h,\Delta t}(\mathbf{x}, 0) &= c_h^0(\mathbf{x}) \text{ for } \mathbf{x} \in \Omega, \\ c_{h,\Delta t}(\mathbf{x}, t) &= c_h^n(\mathbf{x}) \text{ for } \mathbf{x} \in \Omega, t \in (t_{n-1}, t_n] \quad n \in \{1, \dots, N\}, \end{aligned} \quad (5.4)$$

where  $c_h^n = \sum_{D \in \mathcal{D}_h} c_D^n \varphi_{V_D}$ ;

(ii) *The function  $\tilde{c}_{h,\Delta t}$  defined by*

$$\begin{aligned} \tilde{c}_{h,\Delta t}(\mathbf{x}, 0) &= c_D^0 \text{ for } \mathbf{x} \in D^\circ, D \in \mathcal{D}_h, \\ \tilde{c}_{h,\Delta t}(\mathbf{x}, t) &= c_D^n \text{ for } \mathbf{x} \in D^\circ, D \in \mathcal{D}_h, t \in (t_{n-1}, t_n] \quad n \in \{1, \dots, N\}. \end{aligned} \quad (5.5)$$

The function  $c_{h,\Delta t}$  is continuous and piecewise linear in space and piecewise constant in time. The function  $\tilde{c}_{h,\Delta t}$  is given by the values of  $c_{h,\Delta t}$  in the vertices from  $\mathcal{V}_h^{\text{int}}$  and is piecewise constant on the control volumes in space and piecewise constant in time. The following important relation between  $c_{h,\Delta t}$  and  $\tilde{c}_{h,\Delta t}$  is a simple consequence of the a priori estimate (5.3):

**Lemma 5.3** (Relation between  $c_{h,\Delta t}$  and  $\tilde{c}_{h,\Delta t}$ ). *There holds*

$$\|c_{h,\Delta t} - \tilde{c}_{h,\Delta t}\|_{Q_T} \longrightarrow 0 \text{ as } h_T \rightarrow 0.$$

**Remark 5.4** (Interpretation of the values  $c_D^n$ ). *We remark that the approximate solutions  $c_{h,\Delta t}$ ,  $\tilde{c}_{h,\Delta t}$  are only interpretations of the values  $c_D^n$ ,  $V_D \in \mathcal{V}_h$ ,  $n \in \{0, 1, \dots, N\}$ . In particular, we may work with  $\tilde{c}_{h,\Delta t}$  as in the finite volume method and then use Lemma 5.3 to extend the convergence results also to  $c_{h,\Delta t}$ .*

## 5.2 Estimates on differences of time and space translates

Estimates on differences of time and space translates have been used in [16, 17, 18] to prove the relative compactness property of the sequence of approximate solutions. We give below the time translate estimate for  $\tilde{c}_{h,\Delta t}$  given by (5.5). The main difficulty is in the simultaneous treatment of the general grid  $\mathcal{D}_h$  and the dual simplicial partition  $\mathcal{T}_h$  and of the approximate solution  $\tilde{c}_{h,\Delta t}$  on them; Lemma 3.2 is important in this respect.

**Lemma 5.5** (Time translate estimate). *There exists a constant  $C_{\text{tt}} > 0$  such that*

$$\int_0^{T-\tau} \int_{\Omega} (\tilde{c}_{h,\Delta t}(\mathbf{x}, t + \tau) - \tilde{c}_{h,\Delta t}(\mathbf{x}, t))^2 \, d\mathbf{x} \, dt \leq C_{\text{tt}}(\tau + \Delta t)$$

for all  $\tau \in (0, T)$ .

PROOF:

We set

$$T_T := \int_0^{T-\tau} \int_{\Omega} (\tilde{c}_{h,\Delta t}(\mathbf{x}, t + \tau) - \tilde{c}_{h,\Delta t}(\mathbf{x}, t))^2 \, d\mathbf{x} \, dt.$$

Using the definition of  $\tilde{c}_{h,\Delta t}$  given by (5.5), we can rewrite  $T_T$  as

$$T_T = \int_0^{T-\tau} \sum_{D \in \mathcal{D}_h} |D| \left( c_D^{n_1(t)} - c_D^{n_2(t)} \right)^2 \, dt,$$

where

$$\begin{aligned} n_1(t) &\in \{1, \dots, N\} \text{ is such that } t_{n_1-1} < t + \tau \leq t_{n_1}, \\ n_2(t) &\in \{1, \dots, N\} \text{ is such that } t_{n_2-1} < t \leq t_{n_2}. \end{aligned}$$

We now use the growth condition imposed on  $\beta$  in Assumption (A1) and estimate

$$\begin{aligned} T_T &\leq \frac{1}{c_\beta} \int_0^{T-\tau} \sum_{D \in \mathcal{D}_h} |D| \left( c_D^{n_1(t)} - c_D^{n_2(t)} \right) \left( \beta(c_D^{n_1(t)}) - \beta(c_D^{n_2(t)}) \right) \, dt \\ &= \frac{1}{c_\beta} \int_0^{T-\tau} \sum_{D \in \mathcal{D}_h} |D| \left( c_D^{n_1(t)} - c_D^{n_2(t)} \right) \sum_{n=1}^N \chi(n, t) \left( \beta(c_D^n) - \beta(c_D^{n-1}) \right) \, dt, \end{aligned}$$

where the function  $\chi(n, t)$  is defined as

$$\chi(n, t) := \begin{cases} 1 & \text{if } t \leq t_{n-1} < t + \tau \\ 0 & \text{otherwise} \end{cases}.$$

In view of the definition (3.2a)–(3.2c) of the combined scheme, we have

$$\begin{aligned} T_T &\leq \frac{1}{c_\beta} \sum_{n=1}^N \Delta t_n \int_0^{T-\tau} \chi(n, t) \sum_{D \in \mathcal{D}_h} \left( c_D^{n_1(t)} - c_D^{n_2(t)} \right) \left( \sum_{E \in \mathcal{D}_h} \mathbb{S}_{D,E}^n c_E^n \right. \\ &\quad \left. - \sum_{\sigma_{D,E} \in \mathcal{F}_D} v_{D,E}^n \overline{c_{D,E}^n} - F(c_D^n) |D| + q_D^n |D| \right) dt. \end{aligned} \quad (5.6)$$

We now estimate each term separately.

### Diffusion term

We set

$$T_D := \sum_{n=1}^N \Delta t_n \int_0^{T-\tau} \chi(n, t) \sum_{D \in \mathcal{D}_h} \left( c_D^{n_1(t)} - c_D^{n_2(t)} \right) \sum_{E \in \mathcal{D}_h} \mathbb{S}_{D,E}^n c_E^n dt.$$

Using Lemma 4.6 and the inequality  $ab \leq a^2/2 + b^2/2$ , we can write

$$T_D \leq C_{\mathbf{S}} \left( T_{X_1}^* + \frac{1}{2} T_{X_2}^* + \frac{1}{2} T_{X_3}^* \right), \quad (5.7)$$

where

$$T_{X_1}^* := \sum_{n=1}^N \Delta t_n \int_0^{T-\tau} \chi(n, t) \|\nabla c_h^n\|_\Omega^2 dt$$

and

$$T_{X_i}^* := \sum_{n=1}^N \Delta t_n \int_0^{T-\tau} \chi(n, t) \|\nabla (c_h^{n_{i-1}(t)})\|_\Omega^2 dt \quad i \in \{2, 3\}.$$

We will need below the additional terms

$$T_{L_1}^* := \sum_{n=1}^N \Delta t_n \int_0^{T-\tau} \chi(n, t) \sum_{D \in \mathcal{D}_h} (c_D^n)^2 |D| dt$$

and

$$T_{L_i}^* := \sum_{n=1}^N \Delta t_n \int_0^{T-\tau} \chi(n, t) \sum_{D \in \mathcal{D}_h} \left( c_D^{n_{i-1}(t)} \right)^2 |D| dt \quad i \in \{2, 3\}.$$

It follows from [18, Lemma 5.2] that

$$T_{X_1}^* \leq \tau \frac{C_{\text{ae}}}{c_{\mathbf{S}}}, \quad (5.8)$$

$$T_{L_1}^* \leq \tau T \frac{C_{\text{ae}}}{c_\beta} \quad (5.9)$$

and that

$$T_{X_i}^* \leq \tau \frac{C_{\text{ae}}}{c_{\mathbf{S}}}, \quad T_{L_i}^* \leq \tau T \frac{C_{\text{ae}}}{c_\beta} \quad i \in \{2, 3\} \quad (5.10)$$

for a constant time step and

$$T_{X_i}^* \leq (\tau + \Delta t) \frac{C_{ae}}{c_S}, \quad T_{L_i}^* \leq (\tau + \Delta t) T \frac{C_{ae}}{c_\beta} \quad i \in \{2, 3\} \quad (5.11)$$

for a generally nonconstant time step.

### Convection term

We set

$$T_C := - \sum_{n=1}^N \Delta t_n \int_0^{T-\tau} \chi(n, t) \sum_{D \in \mathcal{D}_h} \left( c_D^{n_1(t)} - c_D^{n_2(t)} \right) \sum_{\sigma_{D,E} \in \mathcal{F}_D} v_{D,E}^n \overline{c_{D,E}^n} dt.$$

Now using that  $c_D^n = c_D^{n_1(t)} = c_D^{n_2(t)} = 0$  for  $V_D \in \mathcal{V}_h^{\text{ext}}$ , we can rewrite  $T_C$  as a summation over  $\sigma_{D,E} \in \mathcal{F}_h$ ,

$$T_C = - \sum_{n=1}^N \Delta t_n \int_0^{T-\tau} \chi(n, t) \sum_{\sigma_{D,E} \in \mathcal{F}_h} v_{D,E}^n \overline{c_{D,E}^n} \left( c_D^{n_1(t)} - c_E^{n_1(t)} + c_E^{n_2(t)} - c_D^{n_2(t)} \right) dt.$$

Using the inequality  $\pm ab \leq \varepsilon a^2/2 + b^2/(2\varepsilon)$ ,  $\varepsilon > 0$ , where we put  $\varepsilon = d_{D,E}$ , we come to

$$T_C \leq T_{C_1} + T_{C_2} + T_{C_3}$$

with

$$\begin{aligned} T_{C_1} &:= \sum_{n=1}^N \Delta t_n \int_0^{T-\tau} \chi(n, t) \sum_{\sigma_{D,E} \in \mathcal{F}_h} |v_{D,E}^n| d_{D,E} (\overline{c_{D,E}^n})^2 dt, \\ T_{C_2} &:= \frac{1}{2} \sum_{n=1}^N \Delta t_n \int_0^{T-\tau} \chi(n, t) \sum_{\sigma_{D,E} \in \mathcal{F}_h} \frac{|v_{D,E}^n|}{d_{D,E}} \left( c_E^{n_1(t)} - c_D^{n_1(t)} \right)^2 dt, \\ T_{C_3} &:= \frac{1}{2} \sum_{n=1}^N \Delta t_n \int_0^{T-\tau} \chi(n, t) \sum_{\sigma_{D,E} \in \mathcal{F}_h} \frac{|v_{D,E}^n|}{d_{D,E}} \left( c_E^{n_2(t)} - c_D^{n_2(t)} \right)^2 dt. \end{aligned}$$

Using the definition of  $\overline{c_{D,E}^n}$  by (3.6), the fact that  $0 \leq \alpha_{D,E}^n \leq 1/2$  following from (3.7), and the inequality  $(a+b)^2 \leq 2a^2 + 2b^2$ ,

$$(\overline{c_{D,E}^n})^2 \leq 3(c_D^n)^2 + 3(c_E^n)^2.$$

Hence

$$\begin{aligned} \sum_{\sigma_{D,E} \in \mathcal{F}_h} |v_{D,E}^n| d_{D,E} (\overline{c_{D,E}^n})^2 &\leq 3C_v \sum_{\sigma_{D,E} \in \mathcal{F}_h} |\sigma_{D,E}| d_{D,E} \{ (c_D^n)^2 + (c_E^n)^2 \} \\ &\leq 3C_v C_{1,\mathcal{D}} C_{2,\mathcal{D}} \sum_{D \in \mathcal{D}_h} (c_D^n)^2 |D|, \end{aligned}$$

where we have used Assumption (A4), which implies  $|v_{D,E}^n| \leq C_v |\sigma_{D,E}|$ , and Assumption (B). Thus, we have

$$T_{C_1} \leq 3C_v C_{1,\mathcal{D}} C_{2,\mathcal{D}} T_{L_1}^*.$$

Finally, using  $|v_{D,E}^n| \leq C_{\mathbf{v}}|\sigma_{D,E}|$  and Lemma 3.2, we have

$$T_{C_i} \leq \frac{C_{\mathbf{v}}}{2} C_{1,\mathcal{D},\mathcal{T}} C_{2,\mathcal{D},\mathcal{T}} T_{X_i}^* \quad i \in \{2, 3\},$$

which altogether leads to

$$T_C \leq C_{\mathbf{v}} \left( 3C_{1,\mathcal{D}} C_{2,\mathcal{D}} T_{L_1}^* + \frac{1}{2} C_{1,\mathcal{D},\mathcal{T}} C_{2,\mathcal{D},\mathcal{T}} (T_{X_2}^* + T_{X_3}^*) \right). \quad (5.12)$$

### Reaction term

We denote

$$T_R := - \sum_{n=1}^N \Delta t_n \int_0^{T-\tau} \chi(n, t) \sum_{D \in \mathcal{D}_h} \left( c_D^{n_1(t)} - c_D^{n_2(t)} \right) F(c_D^n) |D| dt.$$

We estimate

$$-F(c_D^n)(c_D^{n_1} - c_D^{n_2}) \leq \frac{(c_D^{n_1} - c_D^{n_2})^2}{2} + \frac{(F(c_D^n))^2}{2} \leq (c_D^{n_1})^2 + (c_D^{n_2})^2 + \frac{L_F^2 (c_D^n)^2}{2},$$

using the inequalities  $ab \leq a^2/2 + b^2/2$ ,  $(a-b)^2/2 \leq a^2 + b^2$ , the Lipschitz continuity of  $F$  with the constant  $L_F$ , and the fact that  $F(0) = 0$ , following either from Assumption (A5) or (A6). This implies

$$T_R \leq \left( \frac{L_F^2}{2} T_{L_1}^* + T_{L_2}^* + T_{L_3}^* \right). \quad (5.13)$$

### Sources term

We denote

$$T_S := \sum_{n=1}^N \Delta t_n \int_0^{T-\tau} \chi(n, t) \sum_{D \in \mathcal{D}_h} \left( c_D^{n_1(t)} - c_D^{n_2(t)} \right) q_D^n |D| dt.$$

Using the same estimate as for the reaction term and (4.3), we come to

$$T_S \leq \frac{1}{2} \tau \|q\|_{Q_T}^2 + T_{L_2}^* + T_{L_3}^*. \quad (5.14)$$

The proof of the lemma is concluded by introducing (5.7), (5.12), (5.13), and (5.14) into (5.6), while using the estimates (5.8), (5.9), and (5.11).  $\square$

**Remark 5.6** (Time translate estimate for a constant time step). *For a constant time step, we have indeed an  $O(\tau)$  estimate, using (5.10) instead of (5.11).*

We give below a space translate estimate for  $\tilde{c}_{h,\Delta t}$  given by (5.5). It is a variant of the estimates from [16, 17, 18] for nonmatching grids  $\mathcal{D}_h$  possibly containing nonconvex elements, where the gradient is given by the associated piecewise linear continuous approximation on  $T_h$ .

**Lemma 5.7** (Space translate estimate). *Let us define  $\tilde{c}_{h,\Delta t}(\mathbf{x}, t)$  by zero outside of  $\Omega$ . Then there exists a constant  $C_{\text{st}} > 0$  such that*

$$\int_0^T \int_{\Omega} \left( \tilde{c}_{h,\Delta t}(\mathbf{x} + \boldsymbol{\xi}, t) - \tilde{c}_{h,\Delta t}(\mathbf{x}, t) \right)^2 d\mathbf{x} dt \leq C_{\text{st}} (|\boldsymbol{\xi}|(|\boldsymbol{\xi}| + h_T) + 2h_{\mathcal{D}}(h_{\mathcal{D}} + h_T))$$

for all  $\boldsymbol{\xi} \in \mathbb{R}^d$ .

PROOF:

We first estimate

$$\int_0^T \int_{\Omega} (\tilde{c}_{h,\Delta t}(\mathbf{x} + \boldsymbol{\xi}, t) - \tilde{c}_{h,\Delta t}(\mathbf{x}, t))^2 d\mathbf{x} dt \leq 3 \sum_{n=1}^N \Delta t_n \int_{\Omega} \{ (c_h^n(\mathbf{x} + \boldsymbol{\xi}) - c_h^n(V_D(\mathbf{x} + \boldsymbol{\xi})))^2 + (c_h^n(\mathbf{x} + \boldsymbol{\xi}) - c_h^n(\mathbf{x}))^2 + (c_h^n(\mathbf{x}) - c_h^n(V_D(\mathbf{x})))^2 \} d\mathbf{x},$$

using the definition (5.5) of  $\tilde{c}_{h,\Delta t}$ , the fact that  $c_h^n(V_D(\mathbf{x})) = c_D^n$ , where  $V_D(\mathbf{x})$  is such  $V_D \in \mathcal{V}_h$  that  $\mathbf{x} \in D$ , and the Cauchy–Schwarz inequality. We next define a function  $\chi_K(\mathbf{x})$  for each  $K \in \mathcal{T}_h$  by

$$\chi_K(\mathbf{x}) := \begin{cases} 1 & \text{if } K \cap [\mathbf{x}, \mathbf{x} + \boldsymbol{\xi}] \neq \emptyset \\ 0 & \text{if } K \cap [\mathbf{x}, \mathbf{x} + \boldsymbol{\xi}] = \emptyset \end{cases}.$$

The fact that  $c_h^n$  is continuous leads to,

$$c_h^n(\mathbf{x} + \boldsymbol{\xi}) - c_h^n(\mathbf{x}) = \sum_{K \in \mathcal{T}_h} (c_{2,K,\mathbf{x}}^n - c_{1,K,\mathbf{x}}^n) \chi_K(\mathbf{x})$$

for a.e.  $\mathbf{x} \in \Omega$ , where  $c_{1,K,\mathbf{x}}^n$  is the value of  $c_h$  at the point where  $[\mathbf{x}, \mathbf{x} + \boldsymbol{\xi}]$  for the first time intersects (or “enters”)  $K$  and  $c_{2,K,\mathbf{x}}^n$  is the value of  $c_h$  at the point where  $[\mathbf{x}, \mathbf{x} + \boldsymbol{\xi}]$  for the second time intersects (or “leaves”)  $K$ , with the obvious modification for the points  $\mathbf{x}$  and  $\mathbf{x} + \boldsymbol{\xi}$  themselves (cf. also the proof of Lemma 3.2). Hence, using that  $|c_{2,K,\mathbf{x}}^n - c_{1,K,\mathbf{x}}^n| \leq |\nabla c_h^n|_K |[\mathbf{x}, \mathbf{x} + \boldsymbol{\xi}] \cap K|$  and the Cauchy–Schwarz inequality,

$$(c_h^n(\mathbf{x} + \boldsymbol{\xi}) - c_h^n(\mathbf{x}))^2 \leq \sum_{K \in \mathcal{T}_h} \chi_K(\mathbf{x}) |[\mathbf{x}, \mathbf{x} + \boldsymbol{\xi}] \cap K| \sum_{K \in \mathcal{T}_h} \chi_K(\mathbf{x}) |[\mathbf{x}, \mathbf{x} + \boldsymbol{\xi}] \cap K| |\nabla c_h^n|_K^2.$$

Next,

$$\sum_{K \in \mathcal{T}_h} \chi_K(\mathbf{x}) |[\mathbf{x}, \mathbf{x} + \boldsymbol{\xi}] \cap K| \leq |\boldsymbol{\xi}|$$

for a.e.  $\mathbf{x} \in \Omega$  and

$$\int_{\Omega} \chi_K(\mathbf{x}) |[\mathbf{x}, \mathbf{x} + \boldsymbol{\xi}] \cap K| d\mathbf{x} \leq (|\boldsymbol{\xi}| + h_{\mathcal{T}}) |K|,$$

so that

$$\sum_{n=1}^N \Delta t_n \int_{\Omega} (c_h^n(\mathbf{x} + \boldsymbol{\xi}) - c_h^n(\mathbf{x}))^2 d\mathbf{x} \leq |\boldsymbol{\xi}| (|\boldsymbol{\xi}| + h_{\mathcal{T}}) \sum_{n=1}^N \Delta t_n \sum_{K \in \mathcal{T}_h} |\nabla c_h^n|_K^2 |K| \leq |\boldsymbol{\xi}| (|\boldsymbol{\xi}| + h_{\mathcal{T}}) \frac{C_{ae}}{c_{\mathbf{S}}},$$

using in addition the a priori estimate (5.3). The remaining estimates follow similarly, using that  $|\mathbf{x} - V_D(\mathbf{x})| \leq h_{\mathcal{D}}$ .  $\square$

**Remark 5.8** (Space translate estimate). *In space translate estimates in [16, 17, 18], the upper bound is given by  $C_{st} |\boldsymbol{\xi}| (|\boldsymbol{\xi}| + h)$  for some constant  $C_{st}$ . We pay for the grid  $\mathcal{D}_h$  by an additional  $O(h)$  factor. This is however still sufficient for the convergence proof presented in the next section.*

## 6 Convergence

Using the a priori estimates of the previous section and the Kolmogorov relative compactness theorem, we show in this section that the approximate solutions converge strongly in  $L^2(Q_T)$  to a function  $c$  and we prove that  $c$  is a weak solution of the continuous problem.

## 6.1 Strong convergence in $L^2(Q_T)$

**Theorem 6.1** (Strong convergence in  $L^2(Q_T)$ ). *There exist subsequences of  $\tilde{c}_{h,\Delta t}$  and  $c_{h,\Delta t}$  which converge strongly in  $L^2(Q_T)$  to a function  $c \in L^2(0, T; H_0^1(\Omega))$ .*

PROOF:

Let us consider the sequence  $\tilde{c}_{h,\Delta t}$ . The a priori estimate (5.1) and Lemmas 5.5 and 5.7 imply that  $\tilde{c}_{h,\Delta t}$  satisfies the assumptions of [18, Lemma 8.4]. Thus  $\tilde{c}_{h,\Delta t}$  verifies the assumptions of the Kolmogorov theorem ([8, Theorem IV.25], [14, Theorem 14.1]) and consequently is relatively compact in  $L^2(Q_T)$ . This implies the existence of a subsequence of  $\tilde{c}_{h,\Delta t}$  which converges strongly to some function  $c$  in  $L^2(Q_T)$ . Moreover, due to the space translate estimate of Lemma 5.7, [14, Theorem 14.2] gives that  $c \in L^2(0, T; H_0^1(\Omega))$ . Finally, considering Lemma 5.3,  $c_{h,\Delta t}$  converges to the same  $c$ .  $\square$

## 6.2 Convergence to a weak solution

We have shown in Theorem 6.1 that subsequences of  $\tilde{c}_{h,\Delta t}$  and  $c_{h,\Delta t}$ , which we still denote by  $\tilde{c}_{h,\Delta t}$  and  $c_{h,\Delta t}$ , converge strongly in  $L^2(Q_T)$  to some function  $c \in L^2(0, T; H_0^1(\Omega))$ . We now show that  $c$  is a weak solution of the continuous problem. For this purpose, we introduce

$$\Psi := \{ \psi \in C^{2,1}(\bar{\Omega} \times [0, T]), \psi = 0 \text{ on } \partial\Omega \times [0, T], \psi(\cdot, T) = 0 \}. \quad (6.1)$$

We then take an arbitrary  $\psi \in \Psi$ , multiply (3.2c) by  $\Delta t_n \psi(V_D, t_{n-1})$ , and sum the result over  $D \in \mathcal{D}_h$  and  $n = 1, \dots, N$ . This gives

$$T_T + T_D + T_C + T_R = T_S \quad (6.2)$$

with

$$\begin{aligned} T_T &:= \sum_{n=1}^N \sum_{D \in \mathcal{D}_h} (\beta(c_D^n) - \beta(c_D^{n-1})) \psi(V_D, t_{n-1}) |D|, \\ T_D &:= \sum_{n=1}^N \Delta t_n \sum_{D \in \mathcal{D}_h} \sum_{E \in \mathcal{D}_h} c_E^n \sum_{K \in \mathcal{T}_h} (\mathbf{S}^n \nabla \varphi_{V_E}, \nabla \varphi_{V_D})_K \psi(V_D, t_{n-1}), \\ T_C &:= \sum_{n=1}^N \Delta t_n \sum_{D \in \mathcal{D}_h} \sum_{\sigma_{D,E} \in \mathcal{F}_D} v_{D,E}^n \overline{c_{D,E}^n} \psi(V_D, t_{n-1}), \\ T_R &:= \sum_{n=1}^N \Delta t_n \sum_{D \in \mathcal{D}_h} F(c_D^n) \psi(V_D, t_{n-1}) |D|, \\ T_S &:= \sum_{n=1}^N \Delta t_n \sum_{D \in \mathcal{D}_h} q_D^n \psi(V_D, t_{n-1}) |D|. \end{aligned}$$

The proofs of

$$T_T \longrightarrow - \int_0^T \int_{\Omega} \beta(c(\mathbf{x}, t)) \psi_t(\mathbf{x}, t) \, d\mathbf{x} \, dt - \int_{\Omega} \beta(c_0(\mathbf{x})) \psi(\mathbf{x}, 0) \, d\mathbf{x}, \quad (6.3)$$

$$T_R \longrightarrow \int_0^T \int_{\Omega} F(c(\mathbf{x}, t)) \psi(\mathbf{x}, t) \, d\mathbf{x} \, dt, \quad (6.4)$$

$$T_S \longrightarrow \int_0^T \int_{\Omega} q(\mathbf{x}, t) \psi(\mathbf{x}, t) \, d\mathbf{x} \, dt, \quad (6.5)$$

as  $h, \Delta t \rightarrow 0$  are completely analogous to those given in [18, Section 6.2]. We thus only show the convergence of  $T_D$  and  $T_C$  to their continuous versions.

### Diffusion term

We rewrite  $T_D$  as

$$T_D = \sum_{n=1}^N \Delta t_n \sum_{K \in \mathcal{T}_h} \int_K \mathbf{S}^n \nabla c_h^n(\mathbf{x}) \cdot \nabla \left( \sum_{D \in \mathcal{D}_h} \psi(V_D, t_{n-1}) \varphi_{V_D}(\mathbf{x}) \right) d\mathbf{x},$$

using the definition of  $c_h^n \in X_h$ , and define

$$\mathbf{S}_{\Delta t}(\mathbf{x}, t) := \mathbf{S}^n(\mathbf{x}) \text{ for } \mathbf{x} \in \Omega, t \in (t_{n-1}, t_n] \quad n \in \{1, \dots, N\}, \quad (6.6)$$

where  $\mathbf{S}^n$  is given by (3.4) or by (3.5). We will show the validity of two passages to the limit. We begin by proving that

$$\begin{aligned} & \sum_{n=1}^N \Delta t_n \sum_{K \in \mathcal{T}_h} \int_K \mathbf{S}^n \nabla c_h^n(\mathbf{x}) \cdot \nabla \left( \sum_{D \in \mathcal{D}_h} \psi(V_D, t_{n-1}) \varphi_{V_D}(\mathbf{x}) \right) d\mathbf{x} \\ & - \sum_{n=1}^N \Delta t_n \sum_{K \in \mathcal{T}_h} \int_K \mathbf{S}^n \nabla c_h^n(\mathbf{x}) \cdot \nabla \psi(\mathbf{x}, t_{n-1}) d\mathbf{x} \longrightarrow 0 \text{ as } h \rightarrow 0. \end{aligned} \quad (6.7)$$

We set

$$I_\psi(\cdot, t_{n-1}) := \sum_{D \in \mathcal{D}_h} \psi(V_D, t_{n-1}) \varphi_{V_D}$$

and

$$T_{D_1} := \sum_{n=1}^N \Delta t_n \sum_{K \in \mathcal{T}_h} \int_K \mathbf{S}^n \nabla c_h^n(\mathbf{x}) \cdot \nabla (I_\psi(\mathbf{x}, t_{n-1}) - \psi(\mathbf{x}, t_{n-1})) d\mathbf{x}.$$

We then estimate

$$|T_{D_1}| \leq C_S \sum_{n=1}^N \Delta t_n \|\nabla c_h^n\|_\Omega \|\nabla (I_\psi(\cdot, t_{n-1}) - \psi(\cdot, t_{n-1}))\|_\Omega,$$

using the Cauchy–Schwarz inequality. Next we use the interpolation estimate

$$\begin{aligned} \|\nabla (I_\psi(\cdot, t_{n-1}) - \psi(\cdot, t_{n-1}))\|_\Omega &= \left( \sum_{K \in \mathcal{T}_h} \int_K |\nabla (I_\psi(\cdot, t_{n-1}) - \psi(\cdot, t_{n-1}))|^2 d\mathbf{x} \right)^{\frac{1}{2}} \\ &\leq C_I \theta_{\mathcal{T}} h \left( \sum_{K \in \mathcal{T}_h} |\psi(\cdot, t_{n-1})|_{2,K}^2 \right)^{\frac{1}{2}} \leq C_I \theta_{\mathcal{T}} C_{1,\psi} h, \end{aligned}$$

where  $\theta_{\mathcal{T}}$  is given by the consequence (3.1) of Assumption (C),  $C_I$  does not depend on  $h$  (nor on  $\Delta t$ ), and  $|\cdot|_{2,K}$  denotes the  $H^2$  seminorm, see for instance [10, Theorem 3.1.5]. Finally, the Cauchy–Schwarz inequality yields

$$|T_{D_1}| \leq C_S C_I \theta_{\mathcal{T}} C_{1,\psi} h \left( \sum_{n=1}^N \Delta t_n \|\nabla c_h^n\|_\Omega^2 \right)^{\frac{1}{2}} \left( \sum_{n=1}^N \Delta t_n \right)^{\frac{1}{2}} = C_S C_I \theta_{\mathcal{T}} C_{1,\psi} T^{\frac{1}{2}} \left( \frac{C_{ae}}{C_S} \right)^{\frac{1}{2}} h,$$

using the a priori estimate (5.3). Hence (6.7) is fulfilled.

We next show that

$$\sum_{n=1}^N \Delta t_n \sum_{K \in \mathcal{T}_h} \int_K \mathbf{S}^n \nabla c_h^n(\mathbf{x}) \cdot \nabla \psi(\mathbf{x}, t_{n-1}) \, d\mathbf{x} \longrightarrow \int_0^T \int_{\Omega} \mathbf{S} \nabla c(\mathbf{x}, t) \cdot \nabla \psi(\mathbf{x}, t) \, d\mathbf{x} \, dt \quad (6.8)$$

as  $h, \Delta t \rightarrow 0$ . We see that both  $c_h^n(\mathbf{x})$  and  $\psi(\mathbf{x}, t_{n-1})$  are constant in time, so that we can easily introduce an integral with respect to time into the first term of (6.8). We further add and subtract  $\sum_{n=1}^N \int_{t_{n-1}}^{t_n} \int_{\Omega} \mathbf{S}^n \nabla c_h^n(\mathbf{x}) \nabla \psi(\mathbf{x}, t) \, d\mathbf{x} \, dt$  and introduce

$$\begin{aligned} T_{D_2} &:= \sum_{n=1}^N \int_{t_{n-1}}^{t_n} \sum_{K \in \mathcal{T}_h} \int_K \mathbf{S}^n \nabla c_h^n(\mathbf{x}) \cdot (\nabla \psi(\mathbf{x}, t_{n-1}) - \nabla \psi(\mathbf{x}, t)) \, d\mathbf{x} \, dt, \\ T_{D_3} &:= \int_0^T \sum_{K \in \mathcal{T}_h} \int_K \mathbf{S}_{\Delta t} \nabla c_{h, \Delta t}(\mathbf{x}, t) \cdot \nabla \psi(\mathbf{x}, t) \, d\mathbf{x} \, dt \\ &\quad - \int_0^T \int_{\Omega} \mathbf{S} \nabla c(\mathbf{x}, t) \cdot \nabla \psi(\mathbf{x}, t) \, d\mathbf{x} \, dt, \end{aligned}$$

where  $c_{h, \Delta t}$  is given by (5.4). Clearly, (6.8) is valid when  $T_{D_2}$  and  $T_{D_3}$  tend to zero as  $h, \Delta t \rightarrow 0$ . We first estimate  $T_{D_2}$ . We have, for  $t \in (t_{n-1}, t_n]$ ,

$$|\nabla \psi(\mathbf{x}, t_{n-1}) - \nabla \psi(\mathbf{x}, t)| \leq g(\Delta t),$$

where  $g$  satisfies  $g(\Delta t) > 0$  and  $g(\Delta t) \rightarrow 0$  as  $\Delta t \rightarrow 0$ . Thus

$$|T_{D_2}| \leq C_{\mathbf{S}} g(\Delta t) \sum_{n=1}^N \Delta t_n \sum_{K \in \mathcal{T}_h} |\nabla c_h^n|_K |K| \leq C_{\mathbf{S}} g(\Delta t) \left( \frac{C_{\text{ae}}}{c_{\mathbf{S}}} \right)^{\frac{1}{2}} T^{\frac{1}{2}} |\Omega|^{\frac{1}{2}},$$

using the Cauchy–Schwarz inequality and the a priori estimate (5.3).

To prove that  $T_{D_3} \rightarrow 0$  as  $h, \Delta t \rightarrow 0$ , we first show that

$$T'_{D_3} := \int_0^T \int_{\Omega} (\nabla c_{h, \Delta t}(\mathbf{x}, t) - \nabla c(\mathbf{x}, t)) \cdot \mathbf{w}(\mathbf{x}, t) \, d\mathbf{x} \, dt \longrightarrow 0 \quad (6.9)$$

as  $h, \Delta t \rightarrow 0$  for all  $\mathbf{w} \in [C^1(\overline{Q_T})]^d$ . This is however immediate, since

$$T'_{D_3} = \int_0^T \int_{\Omega} (c(\mathbf{x}, t) - c_{h, \Delta t}(\mathbf{x}, t)) \nabla \cdot \mathbf{w}(\mathbf{x}, t) \, d\mathbf{x} \, dt,$$

using the Green theorem for  $c_{h, \Delta t}$  and  $c$  (recall that  $c \in L^2(0, T; H_0^1(\Omega))$  by Theorem 6.1) and  $\mathbf{w}$ , which tends to zero by the strong  $L^2(Q_T)$  convergence of  $c_{h, \Delta t}$  to  $c$ .

We next show that the density of the set  $[C^1(\overline{Q_T})]^d$  in  $[L^2(Q_T)]^d$  and (6.9) implies a weak convergence of  $\nabla c_{h, \Delta t}$  (piecewise constant function in space and time) to  $\nabla c$ . Indeed, let  $\mathbf{w} \in [L^2(Q_T)]^d$  be given and let  $\mathbf{w}_n$  be a sequence of  $[C^1(\overline{Q_T})]^d$  functions converging in  $[L^2(Q_T)]^d$  to  $\mathbf{w}$ . Then

$$\begin{aligned} \int_0^T \int_{\Omega} (\nabla c_{h, \Delta t} - \nabla c) \cdot \mathbf{w} \, d\mathbf{x} \, dt &= \int_0^T \int_{\Omega} (\nabla c_{h, \Delta t} - \nabla c) \cdot \mathbf{w}_n \, d\mathbf{x} \, dt \\ &\quad + \int_0^T \int_{\Omega} (\nabla c_{h, \Delta t} - \nabla c) \cdot (\mathbf{w} - \mathbf{w}_n) \, d\mathbf{x} \, dt. \end{aligned}$$

Since both  $c_{h,\Delta t}$  and  $c$  belong to  $L^2(0, T; H_0^1(\Omega))$ , the second term of the above expression tends to zero as  $n \rightarrow \infty$  by the Cauchy–Schwarz inequality. Hence the whole expression tends to zero as  $h, \Delta t \rightarrow 0$  for each  $\mathbf{w} \in [L^2(Q_T)]^d$ , using (6.9) for the first term.

We now finally conclude that  $T_{D_3} \rightarrow 0$  as  $h, \Delta t \rightarrow 0$ . We can write

$$T_{D_3} = \int_0^T \int_{\Omega} (\mathbf{S}_{\Delta t} - \mathbf{S}) \nabla c_{h,\Delta t} \cdot \nabla \psi \, d\mathbf{x} \, dt - \int_0^T \int_{\Omega} \mathbf{S} (\nabla c - \nabla c_{h,\Delta t}) \cdot \nabla \psi \, d\mathbf{x} \, dt.$$

Since  $(\mathbf{S}_{\Delta t})_{i,j}$ ,  $1 \leq i, j \leq d$ , converge strongly in  $L^1(Q_T)$  to  $\mathbf{S}_{i,j}$  by its definition (6.6), the boundedness of  $\mathbf{S}_{\Delta t}$  and  $\mathbf{S}$  given by Assumption (A3) implies a strong  $L^2(Q_T)$  convergence as well. Hence the first term of the above expression tends to zero as  $h, \Delta t \rightarrow 0$ , using the boundedness of  $|\nabla \psi|$ , the a priori estimate (5.3), and the Cauchy–Schwarz inequality. The second term converges to zero by the  $L^\infty$  boundedness of  $\mathbf{S}$  and the weak convergence of  $\nabla c_{h,\Delta t}$  to  $\nabla c$  shown in the previous paragraph. Altogether, combining (6.7) and (6.8) gives

$$T_D \longrightarrow \int_0^T \int_{\Omega} \mathbf{S} \nabla c(\mathbf{x}, t) \cdot \nabla \psi(\mathbf{x}, t) \, d\mathbf{x} \, dt \text{ as } h, \Delta t \rightarrow 0. \quad (6.10)$$

### Convection term

We recall that

$$T_C = \sum_{n=1}^N \Delta t_n \sum_{D \in \mathcal{D}_h} \sum_{\sigma_{D,E} \in \mathcal{F}_D} v_{D,E}^n \overline{c_{D,E}^n} \psi(V_D, t_{n-1})$$

and denote

$$\mathbf{v}^n(\mathbf{x}) := \frac{1}{\Delta t_n} \int_{t_{n-1}}^{t_n} \mathbf{v}(\mathbf{x}, t) \, dt \quad n \in \{1, \dots, N\}, \mathbf{x} \in \Omega. \quad (6.11)$$

We first intend to show that

$$\begin{aligned} T_C + \sum_{n=1}^N \Delta t_n \sum_{D \in \mathcal{D}_h} c_D^n \sum_{\sigma_{D,E} \in \mathcal{F}_D} \int_{\sigma_{D,E}} \mathbf{v}^n(\mathbf{x}) \cdot \mathbf{n}_{D,E} \psi(\mathbf{x}, t_{n-1}) \, d\gamma(\mathbf{x}) \\ - \sum_{n=1}^N \Delta t_n \sum_{D \in \mathcal{D}_h} c_D^n \int_D \nabla \cdot \mathbf{v}^n(\mathbf{x}) \psi(\mathbf{x}, t_{n-1}) \, d\mathbf{x} \longrightarrow 0 \text{ as } h \rightarrow 0. \end{aligned} \quad (6.12)$$

We add and subtract  $c_D^n \psi(V_D, t_{n-1}) v_{D,E}^n$  and  $\overline{c_{D,E}^n} \int_{\sigma_{D,E}} \mathbf{v}^n(\mathbf{x}) \cdot \mathbf{n}_{D,E} \psi(\mathbf{x}, t_{n-1}) \, d\gamma(\mathbf{x})$  to the summations in the first two terms of (6.12). We denote

$$\begin{aligned} T_{C_1} &:= \sum_{n=1}^N \Delta t_n \sum_{D \in \mathcal{D}_h} \sum_{\sigma_{D,E} \in \mathcal{F}_D} (\overline{c_{D,E}^n} - c_D^n) \left( \psi(V_D, t_{n-1}) v_{D,E}^n \right. \\ &\quad \left. - \int_{\sigma_{D,E}} \mathbf{v}^n(\mathbf{x}) \cdot \mathbf{n}_{D,E} \psi(\mathbf{x}, t_{n-1}) \, d\gamma(\mathbf{x}) \right), \\ T_{C_2} &:= \sum_{n=1}^N \Delta t_n \sum_{D \in \mathcal{D}_h} \sum_{\sigma_{D,E} \in \mathcal{F}_D} \overline{c_{D,E}^n} \int_{\sigma_{D,E}} \mathbf{v}^n(\mathbf{x}) \cdot \mathbf{n}_{D,E} \psi(\mathbf{x}, t_{n-1}) \, d\gamma(\mathbf{x}), \\ T_{C_3} &:= \sum_{n=1}^N \Delta t_n \sum_{D \in \mathcal{D}_h} c_D^n \psi(V_D, t_{n-1}) \sum_{\sigma_{D,E} \in \mathcal{F}_D} v_{D,E}^n, \\ T_{C_4} &:= \sum_{n=1}^N \Delta t_n \sum_{D \in \mathcal{D}_h} c_D^n \int_D \nabla \cdot \mathbf{v}^n(\mathbf{x}) \psi(\mathbf{x}, t_{n-1}) \, d\mathbf{x}. \end{aligned}$$

One can easily verify that (6.12) is satisfied when  $T_{C_1} \rightarrow 0$ ,  $T_{C_2} \rightarrow 0$ , and  $(T_{C_3} - T_{C_4}) \rightarrow 0$  as  $h \rightarrow 0$ .

We begin with  $T_{C_2}$ . We denote

$$\mathbf{v}_{\psi;D,E}^n := \int_{\sigma_{D,E}} \mathbf{v}^n(\mathbf{x}) \cdot \mathbf{n}_{D,E} \psi(\mathbf{x}, t_{n-1}) d\gamma(\mathbf{x}).$$

Since the summation in  $T_{C_2}$  is over all  $D \in \mathcal{D}_h$  and  $\sigma_{D,E} \in \mathcal{F}_D$ , each interior  $\sigma_{D,E}$  is in the summation just twice. We consider one fixed interior set  $\sigma_{D,E}$ , where we have denoted  $D$  and  $E$  such that  $v_{D,E}^n \geq 0$ , and have

$$(c_D^n + \alpha_{D,E}^n(c_E^n - c_D^n))\mathbf{v}_{\psi;D,E}^n + (c_D^n + \alpha_{D,E}^n(c_E^n - c_D^n))\mathbf{v}_{\psi;E,D}^n = 0,$$

considering the definition of the local Péclet upstream weighting (3.6) and the fact that  $\mathbf{v}_{\psi;D,E}^n = -\mathbf{v}_{\psi;E,D}^n$ . Thus

$$T_{C_2} = \sum_{n=1}^N \Delta t_n \sum_{\sigma_{D,E} \in \mathcal{F}_h^{\text{ext}}} \overline{c_{D,E}^n} \int_{\sigma_{D,E}} \mathbf{v}^n(\mathbf{x}) \cdot \mathbf{n}_{D,E} \psi(\mathbf{x}, t_{n-1}) d\gamma(\mathbf{x}).$$

Now  $c_E^n = 0$  for all  $V_E \in \mathcal{V}_h^{\text{ext}}$  and all  $n \in \{1, \dots, N\}$  implies  $\overline{c_{D,E}^n} \leq |c_D^n - c_E^n|$  for  $\sigma_{D,E} \in \mathcal{F}_h^{\text{ext}}$  by (3.6) and (3.7). Hence, using in addition  $|\psi| \leq C_{2,\psi}$ ,  $|v_{D,E}^n| \leq C_{\mathbf{v}}|\sigma_{D,E}|$  following from Assumption (A4), the Cauchy–Schwarz inequality, Lemma 3.2, once more the Cauchy–Schwarz inequality, and finally the a priori estimate (5.3),

$$\begin{aligned} |T_{C_2}| &\leq C_{\mathbf{v}}C_{2,\psi} \sum_{n=1}^N \Delta t_n \sum_{\sigma_{D,E} \in \mathcal{F}_h^{\text{ext}}} |c_E^n - c_D^n| |\sigma_{D,E}| \\ &\leq C_{\mathbf{v}}C_{2,\psi} \sum_{n=1}^N \Delta t_n \left( \sum_{\sigma_{D,E} \in \mathcal{F}_h^{\text{ext}}} \frac{|\sigma_{D,E}|}{d_{D,E}} (c_E^n - c_D^n)^2 \right)^{\frac{1}{2}} \left( \sum_{\sigma_{D,E} \in \mathcal{F}_h^{\text{ext}}} |\sigma_{D,E}| d_{D,E} \right)^{\frac{1}{2}} \\ &\leq C_{\mathbf{v}}C_{2,\psi} (C_{1,\mathcal{D},\mathcal{T}}C_{2,\mathcal{D},\mathcal{T}}|\partial\Omega|h)^{\frac{1}{2}} \sum_{n=1}^N \Delta t_n \|\nabla c_h^n\|_{\Omega} \leq C_{\mathbf{v}}C_{2,\psi} (C_{1,\mathcal{D},\mathcal{T}}C_{2,\mathcal{D},\mathcal{T}}|\partial\Omega|Th)^{\frac{1}{2}} \frac{C_{\text{ae}}}{c_{\mathbf{S}}}, \end{aligned}$$

so that  $T_{C_2} \rightarrow 0$  as  $h \rightarrow 0$ .

Next we consider  $T_{C_3}$  and  $T_{C_4}$ . We immediately have that

$$\sum_{\sigma_{D,E} \in \mathcal{F}_D} v_{D,E}^n = \int_D \nabla \cdot \mathbf{v}^n(\mathbf{x}) d\mathbf{x} \quad \forall D \in \mathcal{D}_h,$$

using the definition of  $v_{D,E}^n$ . We further estimate

$$\begin{aligned} |T_{C_3} - T_{C_4}| &= \left| \sum_{n=1}^N \Delta t_n \sum_{D \in \mathcal{D}_h} c_D^n \int_D \nabla \cdot \mathbf{v}^n(\mathbf{x}) (\psi(V_D, t_{n-1}) - \psi(\mathbf{x}, t_{n-1})) d\mathbf{x} \right| \\ &\leq C_{3,\psi} h \sum_{n=1}^N \sum_{D \in \mathcal{D}_h} |c_D^n| \int_{t_{n-1}}^{t_n} \int_D r(\mathbf{x}, t) d\mathbf{x} dt \tag{6.13} \\ &\leq C_{3,\psi} h \left( \sum_{n=1}^N \sum_{D \in \mathcal{D}_h} \Delta t_n |D| (c_D^n)^2 \right)^{\frac{1}{2}} \left( \sum_{n=1}^N \sum_{D \in \mathcal{D}_h} \frac{\left( \int_{t_{n-1}}^{t_n} \int_D r(\mathbf{x}, t) d\mathbf{x} dt \right)^2}{\Delta t_n |D|} \right)^{\frac{1}{2}} \\ &\leq C_{3,\psi} h \left( \frac{C_{\text{ae}}}{c_{\beta}} T \right)^{\frac{1}{2}} \|r\|_{Q_T}, \end{aligned}$$

considering

$$|\psi(V_D, t_{n-1}) - \psi(\mathbf{x}, t_{n-1})| \leq C_{3,\psi} h \quad (6.14)$$

for all  $\mathbf{x} \in D$ ,

$$\int_D |\nabla \cdot \mathbf{v}^n(\mathbf{x})| \, d\mathbf{x} = \frac{1}{\Delta t_n} \int_D \int_{t_{n-1}}^{t_n} \nabla \cdot \mathbf{v}(\mathbf{x}, t) \, dt \, d\mathbf{x} = \frac{1}{\Delta t_n} \int_D \int_{t_{n-1}}^{t_n} r(\mathbf{x}, t) \, dt \, d\mathbf{x},$$

which follows using Assumption (A4), the Cauchy–Schwarz inequality, and the a priori estimate (5.1). Thus  $(T_{C_3} - T_{C_4}) \rightarrow 0$  as  $h \rightarrow 0$ .

We finally turn to  $T_{C_1}$ . We first define

$$T_{C_5} := \sum_{n=1}^N \Delta t_n \sum_{D \in \mathcal{D}_h} \sum_{\sigma_{D,E} \in \mathcal{F}_D} \frac{|\sigma_{D,E}|}{d_{D,E}} (\overline{c_{D,E}^n} - c_D^n)^2.$$

We have

$$(\overline{c_{D,E}^n} - c_D^n)^2 = (\alpha_{D,E}^n (c_E^n - c_D^n))^2 \leq \frac{1}{4} (c_E^n - c_D^n)^2$$

when  $v_{D,E}^n \geq 0$ , considering the definition of the local Péclet upstream weighting (3.6) and Remark 3.5, which gives  $0 \leq \alpha_{D,E}^n \leq 1/2$ . Similarly, when  $v_{D,E}^n < 0$ , we come to

$$(\overline{c_{D,E}^n} - c_D^n)^2 = ((c_E^n - c_D^n)(1 - \alpha_{D,E}^n))^2 \leq (c_E^n - c_D^n)^2.$$

Thus

$$\begin{aligned} T_{C_5} &\leq 2 \sum_{n=1}^N \Delta t_n \sum_{\sigma_{D,E} \in \mathcal{F}_h} \frac{|\sigma_{D,E}|}{d_{D,E}} (c_E^n - c_D^n)^2 \\ &\leq 2C_{1,\mathcal{D},\mathcal{T}} C_{2,\mathcal{D},\mathcal{T}} \sum_{n=1}^N \Delta t_n \|\nabla c_h^n\|_{\Omega}^2 \leq 2C_{1,\mathcal{D},\mathcal{T}} C_{2,\mathcal{D},\mathcal{T}} \frac{C_{\text{ae}}}{C_{\text{S}}}, \end{aligned}$$

noticing that each interior  $\sigma_{D,E}$  is in the original summation twice and using Lemma 3.2 and the a priori estimate (5.3). We next define

$$\begin{aligned} T_{C_6} &:= \sum_{n=1}^N \Delta t_n \sum_{D \in \mathcal{D}_h} \sum_{\sigma_{D,E} \in \mathcal{F}_D} \frac{d_{D,E}}{|\sigma_{D,E}|} \\ &\quad \left( \int_{\sigma_{D,E}} \mathbf{v}^n(\mathbf{x}) \cdot \mathbf{n}_{D,E} (\psi(V_D, t_{n-1}) - \psi(\mathbf{x}, t_{n-1})) \, d\gamma(\mathbf{x}) \right)^2 \end{aligned}$$

and estimate

$$\begin{aligned} T_{C_6} &\leq C_{3,\psi}^2 h^2 C_{\mathbf{v}}^2 \sum_{n=1}^N \Delta t_n \sum_{D \in \mathcal{D}_h} \sum_{\sigma_{D,E} \in \mathcal{F}_D} d_{D,E} |\sigma_{D,E}| \\ &\leq C_{3,\psi}^2 h^2 C_{\mathbf{v}}^2 C_{1,\mathcal{D}} C_{2,\mathcal{D}} \sum_{n=1}^N \Delta t_n \sum_{D \in \mathcal{D}_h} |D| \leq C_{3,\psi}^2 h^2 C_{\mathbf{v}}^2 C_{1,\mathcal{D}} C_{2,\mathcal{D}} |\Omega| T, \end{aligned}$$

using (6.14),  $|v_{D,E}^n| \leq C_{\mathbf{v}} |\sigma_{D,E}|$  following from Assumption (A4), and finally Assumption (B). We now notice that

$$T_{C_1}^2 \leq T_{C_5} T_{C_6},$$

using the Cauchy–Schwarz inequality, and hence  $T_{C_1} \rightarrow 0$  as  $h \rightarrow 0$ . Thus (6.12) is satisfied.

Using the Green theorem, we easily come to

$$\begin{aligned} \sum_{n=1}^N \Delta t_n \sum_{D \in \mathcal{D}_h} c_D^n \sum_{\sigma_{D,E} \in \mathcal{F}_D} \int_{\sigma_{D,E}} \mathbf{v}^n(\mathbf{x}) \cdot \mathbf{n}_{D,E} \psi(\mathbf{x}, t_{n-1}) \, d\gamma(\mathbf{x}) &= \sum_{n=1}^N \Delta t_n \quad (6.15) \\ \sum_{D \in \mathcal{D}_h} c_D^n \int_D \mathbf{v}^n(\mathbf{x}) \nabla \psi(\mathbf{x}, t_{n-1}) \, d\mathbf{x} &+ \sum_{n=1}^N \Delta t_n \sum_{D \in \mathcal{D}_h} c_D^n \int_D \nabla \cdot \mathbf{v}^n(\mathbf{x}) \psi(\mathbf{x}, t_{n-1}) \, d\mathbf{x}. \end{aligned}$$

Therefore it follows from (6.12) that if we can prove that

$$\begin{aligned} &\sum_{n=1}^N \Delta t_n \sum_{D \in \mathcal{D}_h} c_D^n \int_D \mathbf{v}^n(\mathbf{x}) \cdot \nabla \psi(\mathbf{x}, t_{n-1}) \, d\mathbf{x} \quad (6.16) \\ &\longrightarrow \int_0^T \int_{\Omega} c(\mathbf{x}, t) \mathbf{v}(\mathbf{x}, t) \cdot \nabla \psi(\mathbf{x}, t) \, d\mathbf{x} \, dt \quad \text{as } h, \Delta t \rightarrow 0, \end{aligned}$$

then we will have that

$$T_C \longrightarrow - \int_0^T \int_{\Omega} c(\mathbf{x}, t) \mathbf{v}(\mathbf{x}, t) \cdot \nabla \psi(\mathbf{x}, t) \, d\mathbf{x} \, dt \quad \text{as } h, \Delta t \rightarrow 0. \quad (6.17)$$

To prove (6.16), we introduce

$$\begin{aligned} T_{C_7} &:= \sum_{n=1}^N \int_{t_{n-1}}^{t_n} \int_{\Omega} \tilde{c}_{h,\Delta t}(\mathbf{x}, t) \mathbf{v}^n(\mathbf{x}) \cdot (\nabla \psi(\mathbf{x}, t_{n-1}) - \nabla \psi(\mathbf{x}, t)) \, d\mathbf{x} \, dt, \\ T_{C_8} &:= \sum_{n=1}^N \int_{t_{n-1}}^{t_n} \int_{\Omega} (\tilde{c}_{h,\Delta t}(\mathbf{x}, t) - c(\mathbf{x}, t)) \mathbf{v}^n(\mathbf{x}) \cdot \nabla \psi(\mathbf{x}, t) \, d\mathbf{x} \, dt, \\ T_{C_9} &:= \sum_{n=1}^N \int_{t_{n-1}}^{t_n} \int_{\Omega} c(\mathbf{x}, t) (\mathbf{v}^n(\mathbf{x}) - \mathbf{v}(\mathbf{x}, t)) \cdot \nabla \psi(\mathbf{x}, t) \, d\mathbf{x} \, dt. \end{aligned}$$

We have

$$|\nabla \psi(\mathbf{x}, t_{n-1}) - \nabla \psi(\mathbf{x}, t)| \leq g(\Delta t)$$

for  $t \in (t_{n-1}, t_n]$  and thus

$$|T_{C_7}| \leq g(\Delta t) \sum_{n=1}^N \sum_{D \in \mathcal{D}_h} |c_D^n| \int_D \int_{t_{n-1}}^{t_n} |\mathbf{v}(\mathbf{x}, t)| \, d\mathbf{x} \, dt \leq g(\Delta t) \left( \frac{C_{ae}}{c_\beta} T \right)^{\frac{1}{2}} \|\mathbf{v}\|_{Q_T},$$

using the same estimate as in (6.13). Thus  $T_{C_7} \rightarrow 0$  as  $\Delta t \rightarrow 0$ . It is immediate that  $T_{C_8} \rightarrow 0$  as  $h, \Delta t \rightarrow 0$ , using the strong (and consequently weak) convergence of  $\tilde{c}_{h,\Delta t}$  to  $c$ . By Assumption (A4) and (6.11)  $\mathbf{v}$  and  $\mathbf{v}^n$  are bounded, and hence the piecewise constant in time approximation given by  $\mathbf{v}^n$  converges strongly in  $\mathbf{L}^2(Q_T)$  to  $\mathbf{v}$  as  $\Delta t \rightarrow 0$ . Since  $|\nabla \psi| \leq C_{3,\psi}$  and  $c \in L^2(Q_T)$ , it suffices to use the Cauchy–Schwarz inequality to conclude that  $T_{C_9} \rightarrow 0$  as  $\Delta t \rightarrow 0$ . Thus (6.16) and consequently (6.17) is fulfilled.

We are now ready to give the final theorem of this paper:

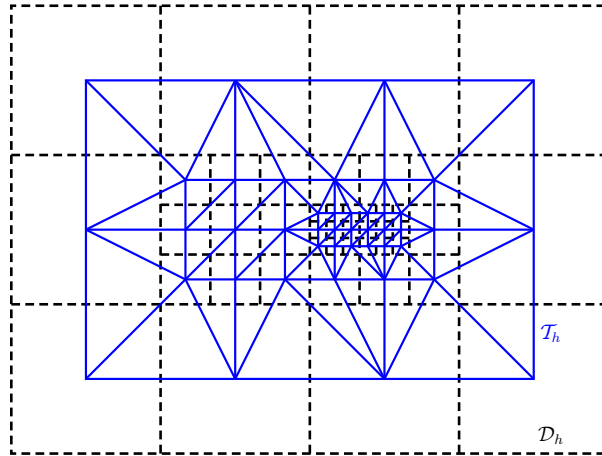


Figure 2: Primal locally refined square grid  $\mathcal{D}_h$  (dashed) and dual triangular grid  $\mathcal{T}_h$  (solid)

**Theorem 6.2** (Convergence to a weak solution). *There exist subsequences of  $\tilde{c}_{h,\Delta t}$  and  $c_{h,\Delta t}$ , the approximate solutions of the problem (1.1)–(2.2) by means of the combined finite volume–finite element scheme given by Definition 5.2, which converge strongly in  $L^2(Q_T)$  to a weak solution of the problem (1.1)–(2.2) given by Definition 2.2. If the weak solution is unique, then the whole sequences  $\tilde{c}_{h,\Delta t}$ ,  $c_{h,\Delta t}$  converge to the weak solution.*

PROOF:

We have from Theorem 6.1 that subsequences of  $\tilde{c}_{h,\Delta t}$  and  $c_{h,\Delta t}$  converge strongly in  $L^2(Q_T)$  to some function  $c \in L^2(0, T; H_0^1(\Omega))$ . The function  $c$  satisfies

$$\begin{aligned} & - \int_0^T \int_{\Omega} \beta(c(\mathbf{x}, t)) \psi_t(\mathbf{x}, t) \, d\mathbf{x} \, dt - \int_{\Omega} \beta(c_0(\mathbf{x})) \psi(\mathbf{x}, 0) \, d\mathbf{x} \\ & + \int_0^T \int_{\Omega} \mathbf{S}(\mathbf{x}, t) \nabla c(\mathbf{x}, t) \cdot \nabla \psi(\mathbf{x}, t) \, d\mathbf{x} \, dt - \int_0^T \int_{\Omega} c(\mathbf{x}, t) \mathbf{v}(\mathbf{x}, t) \cdot \nabla \psi(\mathbf{x}, t) \, d\mathbf{x} \, dt \\ & + \int_0^T \int_{\Omega} F(c(\mathbf{x}, t)) \psi(\mathbf{x}, t) \, d\mathbf{x} \, dt = \int_0^T \int_{\Omega} q(\mathbf{x}, t) \psi(\mathbf{x}, t) \, d\mathbf{x} \, dt \end{aligned}$$

for all test functions  $\psi \in \Psi$ , given by (6.1). This follows from (6.3), (6.10), (6.17), (6.4), (6.5), and (6.2). In addition,  $\beta(c) \in L^\infty(0, T; L^2(\Omega))$ , which follows from (5.2). Thus  $c$  is a weak solution of the problem (1.1)–(2.2), since  $\Psi$  is dense in the set  $\{\varphi; \varphi \in L^2(0, T; H_0^1(\Omega)), \varphi_t \in L^\infty(Q_T), \varphi(\cdot, T) = 0\}$ .  $\square$

## 7 Numerical experiment

We present the results of a numerical experiment in this section. The computations were done with the code TALISMAN [33], where our scheme is implemented.

We consider a model degenerate parabolic convection–diffusion problem with a known traveling wave solution (cf. [22]). In particular, we take the equation (1.1) for  $\Omega = (0, 1) \times (0, 1)$  and  $T = 0.5$

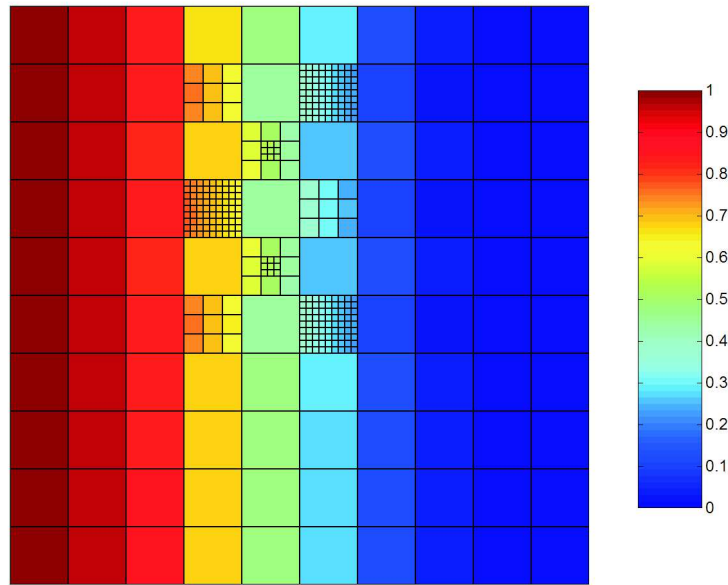


Figure 3: Initial space mesh  $\mathcal{D}_h$  and the approximate solution at  $t = 0.5$ ,  $\delta = 0.025$

with

$$\begin{aligned}\beta(c) &= c^{\frac{1}{2}} \text{ for } c \geq 0, \\ \mathbf{S} &= \delta \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \\ \mathbf{v} &= (v, 0), \\ F(c) &= 0, q = 0.\end{aligned}$$

Here,  $\delta > 0$  and  $v > 0$  are parameters. We fix  $v$  to 0.8 and consider two values of  $\delta$ : for  $\delta = 0.025$ , the problem is not too convection-dominated, which is the case for  $\delta = 0.0001$ . The initial and boundary conditions (Dirichlet on  $x = 0$  and  $x = 1$  and Neumann on  $y = 0$  and  $y = 1$ ) are given by the exact solution

$$c(x, y, t) = \left(1 - e^{\frac{v}{2\delta}(x-vt-p)}\right)^2 \text{ for } x \leq vt + p, \quad c(x, y, t) = 0 \text{ for } x \geq vt + p.$$

The shift  $p$  defines the position of the front of the wave at  $t = 0$  and is set to 0.2. Note that the problem is degenerate parabolic since  $\beta'(0) = +\infty$  and the solution takes the value of 0.

We perform the simulations on locally refined nonmatching square meshes with the initial one ( $r = 0$ ) given in Figure 3. The initial time step is  $T/2$ . We refine the space mesh (twice,  $r = 2$ , and four-times,  $r = 4$ ) by dividing each square into four subsquares and each time the space mesh is refined, the time step is divided by two. The simulated problem is in fact only one-dimensional and we use this fact to test the performance of the numerical scheme that we propose for this type of grids. Since the difference of levels of refinement of two neighboring squares in the given example is more than one (cf. Remark 3.8), the underlying triangulation is not Delaunay and hence the discrete maximum principle (see Theorem 4.11) does not necessarily hold. Hence we need to define the function  $\beta(c)$  for  $c < 0$ . To fulfill Assumptions (A1) and (A2), we set  $\beta(c) := -\beta(-c)$  for  $c < 0$ .

We have depicted the approximate solution for  $\delta = 0.025$  and  $r = 0$  at  $t = 0.5$  in Figure 4 and give the profiles of approximate solutions in  $y = 0.5$  for the different values of  $\delta$  and  $r$  in

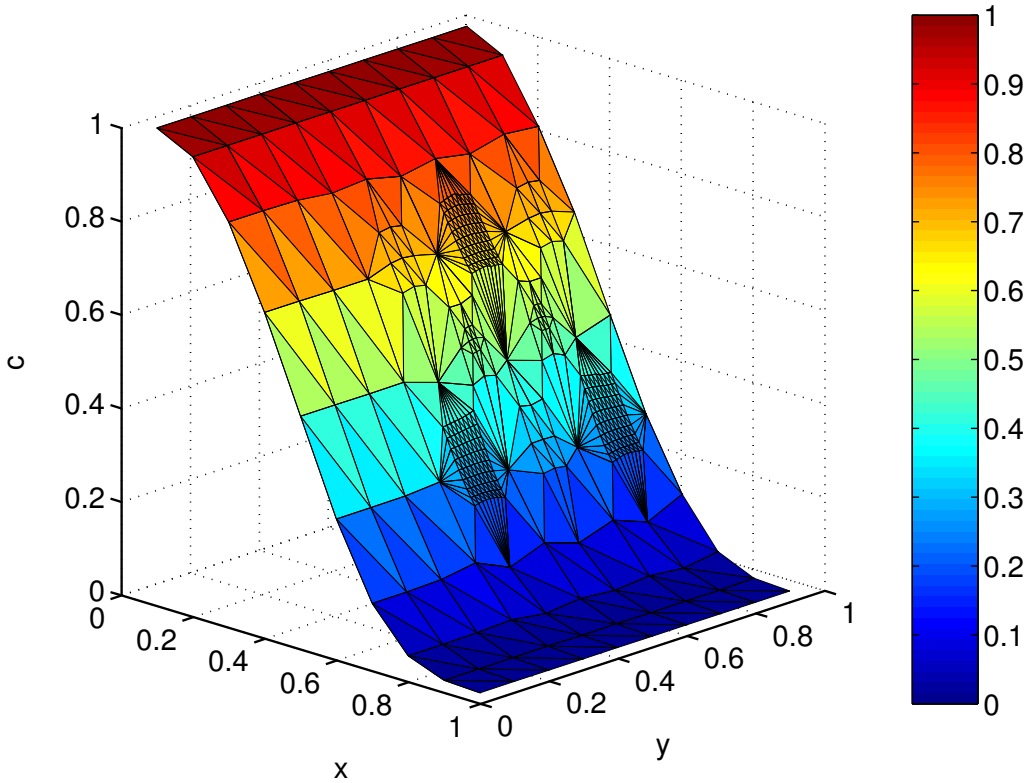


Figure 4: Approximate solution on the dual mesh  $\mathcal{T}_h$  to  $\mathcal{D}_h$  from Figure 3 at  $t = 0.5$ ,  $\delta = 0.025$

Figure 5. The profile is defined by all the calculated values  $c_D$  in squares whose centers  $Q_D$  satisfy  $|Q_D - l_{0.5}| \leq 0.05 \cdot 2^{-r}$ , where  $l_{0.5}$  is the line  $y = 0.5$ . We have purposely chosen the grid such that in its lower part ( $y \leq 0.4$ ), the combined scheme coincides with a pure finite volume one. We thus can clearly observe the influence of the nonmatching refinement. It can be seen from Figure 4 that not only that we do not introduce any error at the interfaces between the subdomains with nonmatching grids, but since the mesh size is smaller in the refined parts, the solution is sharper here. Similarly, the points with bigger error (more distant from the exact solution curve) in Figure 5 represent the original unrefined cells.

At each discrete time level, we had to solve the nonlinear system of algebraic equations given by (3.2a)–(3.2c). Since  $\beta'(0) = +\infty$  and since the solution takes the value 0, we could not directly apply the Newton method for this purpose. We have thus introduced new unknowns  $u_D^n = \beta(c_D^n)$  and rewritten this system for these new unknowns, cf. [18]. This allowed us to apply the Newton method directly and thus to avoid any parabolic regularization (cf. [4]) or perturbation of initial and boundary conditions (cf. [27]). Preconditioned bi-conjugate gradients stabilized method (Bi-CGStab) was used for the solution of the associated linear systems.

As the results demonstrate, our scheme works easily for the given locally refined nonmatching square grid, which would not be the case for the standard finite volume method, cf. [14]. Next, the local Péclet upstream weighting reduces the numerical diffusion of full upstream weighting but still guarantees the stability of the scheme and finally, the scheme is computationally very inexpensive, having just one unknown per cell and leading to positive definite matrices. We currently study the a posteriori error estimates for this scheme, with the perspective of using the adaptive grid refinement technique.

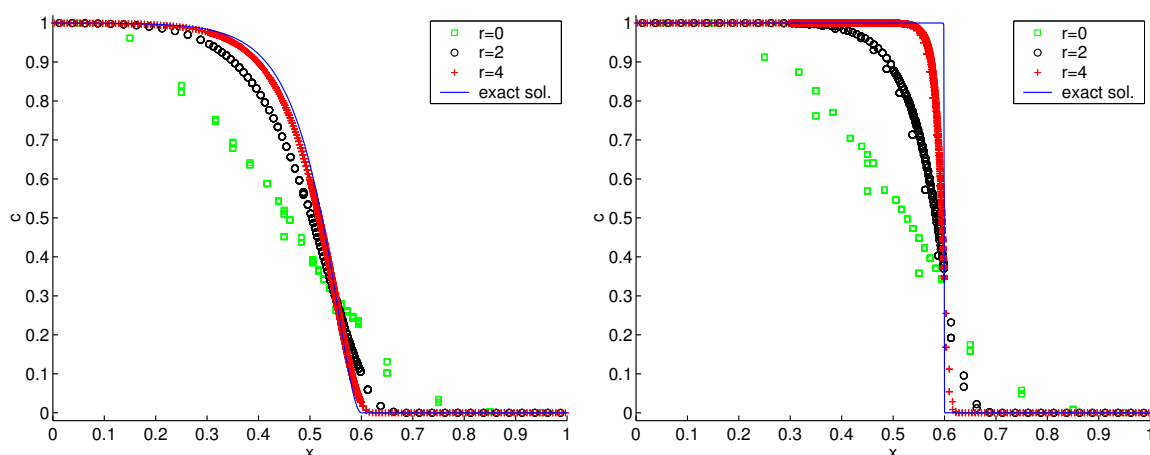


Figure 5: Solution profiles for  $y = 0.5$  at  $t = 0.5$ ,  $\delta = 0.025$  (left) and  $\delta = 0.0001$  (right)

## References

- [1] ACHDOU, Y., JAPHET, C., MADAY, Y., AND NATAF, F. A new cement to glue non-conforming grids with Robin interface conditions: the finite volume case. *Numer. Math.* *92*, 4 (2002), 593–620.
- [2] ARBOGAST, T., COWSAR, L. C., WHEELER, M. F., AND YOTOV, I. Mixed finite element methods on nonmatching multiblock grids. *SIAM J. Numer. Anal.* *37*, 4 (2000), 1295–1315.
- [3] ARBOGAST, T., WHEELER, M. F., AND ZHANG, N.-Y. A nonlinear mixed finite element method for a degenerate parabolic equation arising in flow in porous media. *SIAM J. Numer. Anal.* *33*, 4 (1996), 1669–1687.
- [4] BARRETT, J. W., AND KNABNER, P. Finite element approximation of the transport of reactive solutes in porous media. II. Error estimates for equilibrium adsorption processes. *SIAM J. Numer. Anal.* *34*, 2 (1997), 455–479.
- [5] BAUGHMAN, L. A., AND WALKINGTON, N. J. Co-volume methods for degenerate parabolic problems. *Numer. Math.* *64*, 1 (1993), 45–67.
- [6] BEAR, J., AND VERRUIJT, A. *Modelling Groundwater Flow and Pollution*, vol. 2 of *Theory and Applications of Transport in Porous Media*. Kluwer Academic Publisher, Dordrecht, Holland, 1987.
- [7] BERNARDI, C., MADAY, Y., AND PATERA, A. T. A new nonconforming approach to domain decomposition: the mortar element method. In *Nonlinear partial differential equations and their applications. Collège de France Seminar, Vol. XI (Paris, 1989–1991)*, vol. 299 of *Pitman Res. Notes Math. Ser.* Longman Sci. Tech., Harlow, 1994, pp. 13–51.
- [8] BREZIS, H. *Analyse fonctionnelle*. Collection Mathématiques Appliquées pour la Maîtrise. [Collection of Applied Mathematics for the Master’s Degree]. Masson, Paris, 1983. Théorie et applications. [Theory and applications].
- [9] CAUTRÈS, R., HERBIN, R., AND HUBERT, F. The Lions domain decomposition algorithm on non-matching cell-centred finite volume meshes. *IMA J. Numer. Anal.* *24*, 3 (2004), 465–490.

- [10] CIARLET, P. G. *The Finite Element Method for Elliptic Problems*, vol. 4 of *Studies in Mathematics and its Applications*. North-Holland, Amsterdam, 1978.
- [11] DEBIEZ, C., DERVIEUX, A., MER, K., AND NKONGA, B. Computation of unsteady flows with mixed finite volume/finite element upwind methods. *Internat. J. Numer. Methods Fluids* 27, 1-4, Special Issue (1998), 193–206. Finite elements in fluids.
- [12] EBMEYER, C. Error estimates for a class of degenerate parabolic equations. *SIAM J. Numer. Anal.* 35, 3 (1998), 1095–1112.
- [13] EWING, R., LAZAROV, R., LIN, T., AND LIN, Y. Mortar finite volume element approximations of second order elliptic problems. *East-West J. Numer. Math.* 8, 2 (2000), 93–110.
- [14] EYMARD, R., GALLOUËT, T., AND HERBIN, R. Finite volume methods. In *Handbook of Numerical Analysis, Vol. VII*. North-Holland, Amsterdam, 2000, pp. 713–1020.
- [15] EYMARD, R., GALLOUËT, T., HERBIN, R., AND MICHEL, A. Convergence of a finite volume scheme for nonlinear degenerate parabolic equations. *Numer. Math.* 92, 1 (2002), 41–82.
- [16] EYMARD, R., GALLOUËT, T., HILHORST, D., AND NAÏT SLIMANE, Y. Finite volumes and nonlinear diffusion equations. *RAIRO Modél. Math. Anal. Numér.* 32, 6 (1998), 747–761.
- [17] EYMARD, R., GUTNIC, M., AND HILHORST, D. The finite volume method for Richards equation. *Comput. Geosci.* 3, 3-4 (2000), 259–294.
- [18] EYMARD, R., HILHORST, D., AND VOHRALÍK, M. A combined finite volume–nonconforming/mixed-hybrid finite element scheme for degenerate parabolic problems. *To appear in Numer. Math.* (2006).
- [19] FAILLE, I., NATAF, F., SAAS, L., AND WILLIEN, F. Finite volume methods on non-matching grids with arbitrary interface conditions and highly heterogeneous media. In *Domain Decomposition Methods in Science and Engineering*, Lect. Notes Comput. Sci. Eng. 40. Springer-Verlag, Berlin, 2000, pp. 50–68.
- [20] FEISTAUER, M., FELCMAN, J., AND LUKÁČOVÁ-MEDVIĐOVÁ, M. On the convergence of a combined finite volume–finite element method for nonlinear convection–diffusion problems. *Numer. Methods Partial Differential Equations* 13, 2 (1997), 163–190.
- [21] JOE, B. Delaunay triangular meshes in convex polygons. *SIAM J. Sci. Statist. Comput.* 7, 2 (1986), 514–539.
- [22] KAČUR, J. Solution of degenerate convection–diffusion problems by the method of characteristics. *SIAM J. Numer. Anal.* 39, 3 (2001), 858–879.
- [23] KARLSEN, K. H., RISEBRO, N. H., AND TOWERS, J. D. Upwind difference approximations for degenerate parabolic convection–diffusion equations with a discontinuous coefficient. *IMA J. Numer. Anal.* 22, 4 (2002), 623–664.
- [24] KNABNER, P., AND OTTO, F. Solute transport in porous media with equilibrium and nonequilibrium multiple-site adsorption: uniqueness of weak solutions. *Nonlinear Anal.* 42, 3, Ser. A: Theory Methods (2000), 381–403.

- [25] LETNIEWSKI, F. W. Three-dimensional Delaunay triangulations for finite element approximations to a second-order diffusion operator. *SIAM J. Sci. Statist. Comput.* *13*, 3 (1992), 765–770.
- [26] 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.
- [27] POP, I. S., AND YONG, W.-A. A numerical approach to degenerate parabolic equations. *Numer. Math.* *92*, 2 (2002), 357–381.
- [28] PUTTI, M., AND CORDES, C. Finite element approximation of the diffusion operator on tetrahedra. *SIAM J. Sci. Comput.* *19*, 4 (1998), 1154–1168.
- [29] QUARTERONI, A., AND VALLI, A. *Domain decomposition methods for partial differential equations*. Numerical Mathematics and Scientific Computation. The Clarendon Press Oxford University Press, New York, 1999. Oxford Science Publications.
- [30] RULLA, J., AND WALKINGTON, N. J. Optimal rates of convergence for degenerate parabolic problems in two dimensions. *SIAM J. Numer. Anal.* *33*, 1 (1996), 56–67.
- [31] SONIER, F., AND EYMARD, R. Mathematical and numerical properties of control-volume finite element scheme for reservoir simulation. *SPE Reservoir Engineering November 1994* (1994), 283–289.
- [32] VOHRALÍK, M. *Numerical Methods for Nonlinear Elliptic and Parabolic Equations. Application to Flow Problems in Porous and Fractured media*. Ph.D. dissertation, Université de Paris-Sud & Czech Technical University in Prague, December 2004.
- [33] VOHRALÍK, M., AND RAMAROSY, N. TALISMAN, user guide. *HydroExpert company, 53 rue Charles Frérot, 94 250 Gentilly, France, www.hydroexpert.com* (2005).