



# Numerical analysis of a robust entropy-diminishing Finite Volume scheme for parabolic equations with gradient structure

Clément Cancès, Cindy Guichard

## ► To cite this version:

Clément Cancès, Cindy Guichard. Numerical analysis of a robust entropy-diminishing Finite Volume scheme for parabolic equations with gradient structure. 2015. hal-01119735v2

**HAL Id: hal-01119735**

**<https://hal.science/hal-01119735v2>**

Preprint submitted on 18 Mar 2015 (v2), last revised 7 Jul 2016 (v4)

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

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

# NUMERICAL ANALYSIS OF A ROBUST ENTROPY-DIMINISHING FINITE VOLUME SCHEME FOR PARABOLIC EQUATIONS WITH GRADIENT STRUCTURE

CLÉMENT CANCÈS AND CINDY GUICHARD

**ABSTRACT.** We present a numerical method for approximating the solutions of degenerate parabolic equations with a formal gradient flow structure. The numerical method we propose preserves at the discrete level the formal gradient flow structure, allowing the use of some nonlinear test function in the analysis. The existence of a solution to and the convergence of the scheme are proved under very general assumptions on the continuous problem (nonlinearities, anisotropy, heterogeneity) and on the mesh. Moreover, we provide numerical evidences of the efficiency and of the robustness of our approach.

## 1. INTRODUCTION

Many problem coming from physical (like e.g. for porous media flows modeling [8, 7, 17]) or biological modeling (like e.g. for chemotaxis modeling [32]) lead to degenerate parabolic equations or systems. Many of these models can be interpreted as gradient flows in appropriate geometries. The gradient flow structure can be encountered in many situation of great interest, like for instance porous media flows [37, 16], in chemotaxis processes in biology [10], in supraconductivity [3, 2], or in semiconductor devices modeling [35] (this list is far from being complete).

Designing accurate numerical schemes for approximating their solutions is therefore a major issue. In the case of porous media flow models — used e.g. in oil-engineering, water resources management or nuclear waste repository management — the problems may moreover be highly anisotropic and heterogeneous. As an additional difficulty, the meshes are often prescribed by geological data, yielding non-conformal grids made of elements of various shapes. This situation can also be encountered in mesh adaptation procedures. Hence, the robustness of the method w.r.t. anisotropy and to the grid is an important quality criterion for a numerical method in view of practical applications.

In this contribution, which is mainly academical, we focus on the numerical approximation of a single nonlinear Fokker-Planck equation. Since it contains crucial difficulties arising in the applications, namely degeneracy and possibly strong anisotropy, the discretization of this nonlinear Fokker-Planck equation appears to be a keystone for the approximation of more complex problems.

**1.1. Presentation of the continuous problem.** Let  $\Omega$  be a polyhedral connected open bounded subset of  $\mathbb{R}^d$  ( $d = 2$  or  $3$ ), and let  $t_f > 0$  be a finite time

---

This work was supported by the French National Research Agency ANR (project GeoPor, grant ANR-13-JS01-0007-01).

horizon. In this contribution, we focus on the discretization of the model problem

$$(1) \quad \begin{cases} \partial_t u - \nabla \cdot (\eta(u) \mathbf{\Lambda} \nabla (p(u) + V)) = 0 & \text{in } Q_{t_f} := \Omega \times (0, t_f), \\ (\eta(u) \mathbf{\Lambda} \nabla (p(u) + V)) \cdot \mathbf{n} = 0 & \text{on } \partial\Omega \times (0, t_f), \\ u|_{t=0} = u_0 & \text{in } \Omega, \end{cases}$$

which appears to be a keystone before discretizing more complex problems. We do the following assumptions on the data of the continuous problem (1).

- (A1) The function  $\eta : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  is a continuous function such that  $\eta(0) = 0$ ,  $\eta(u) > 0$  if  $u \neq 0$  and  $\eta$  is non-decreasing on  $\mathcal{I}$ . The function  $\eta$  is extended on whole  $\mathbb{R}$  into an even function.
  - (A2) The function  $p \in L^1_{\text{loc}}(\mathbb{R}_+)$  is absolutely continuous and increasing on  $(0, +\infty)$  (i.e.,  $0 < p' \in L^1_{\text{loc}}((0, +\infty))$ ), and satisfies  $\lim_{u \rightarrow +\infty} p(u) = +\infty$ . In the case where  $p(0) = \lim_{u \searrow 0} p(u)$  is finite, the function  $p$  is extended into an increasing absolutely continuous function  $p : \mathbb{R} \rightarrow \mathbb{R}$  defined by
- $$(2) \quad p(u) = 2p(0) - p(-u), \quad \forall u \leq 0.$$

From now on, we denote by

$$\mathcal{I}_p = \begin{cases} \mathbb{R}_+^* & \text{if } p(0) = -\infty, \\ \mathbb{R} & \text{if } p(0) > -\infty. \end{cases}$$

and by  $\overline{\mathcal{I}_p}$  its closure in  $\mathbb{R}$ . We additionally require that the function  $\sqrt{\eta}p'$  belongs to  $L^1_{\text{loc}}(\mathbb{R}_+)$  (and is in particular integrable near 0) and that  $\lim_{u \searrow 0} \sqrt{\eta(u)}p(u) = 0$ .

- (A3) The tensor field  $\mathbf{\Lambda} : \Omega \rightarrow (L^\infty(\mathbb{R}))^{d \times d}$  is such that  $\mathbf{\Lambda}(\mathbf{x})$  is symmetric for almost every  $\mathbf{x} \in \Omega$ . Moreover, we assume that there exist  $\lambda_* > 0$  and  $\lambda^* \in [\lambda_*, +\infty)$  such that
- $$(3) \quad \lambda_* |\mathbf{v}|^2 \leq \mathbf{\Lambda}(\mathbf{x}) \mathbf{v} \cdot \mathbf{v} \leq \lambda^* |\mathbf{v}|^2, \quad \forall \mathbf{v} \in \mathbb{R}^d, \text{ for a.e. } \mathbf{x} \in \Omega.$$

- (A4) The initial data  $u_0$  is supposed to belong to  $L^1(\Omega)$ . Moreover, defining the convex function  $\Gamma : \overline{\mathcal{I}_p} \rightarrow \mathbb{R}_+$  by

$$(4) \quad \Gamma(u) = \int_1^u (p(a) - p(1)) da, \quad \forall u \in \mathcal{I}_p,$$

we assume that the following positivity and finite entropy conditions are fulfilled:

$$(5) \quad u_0 \geq 0 \text{ a.e. in } \Omega, \quad \int_\Omega u_0 d\mathbf{x} > 0, \quad \int_\Omega \Gamma(u_0) d\mathbf{x} < +\infty.$$

- (A5) The exterior potential  $V : \Omega \rightarrow \mathbb{R}$  is Lipschitz continuous.

Throughout this paper, we adopt the convention

$$(6) \quad \Gamma(u) = +\infty \text{ if } p(0) = -\infty \text{ and } u < 0.$$

In order to give a proper mathematical sense to the solution of (1), we need to introduce the function  $\xi : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  defined by

$$(7) \quad \xi(u) = \int_0^u \sqrt{\eta(a)} p'(a) da, \quad \forall u \geq 0.$$

Note that  $\xi$  is well defined since we assumed that  $\sqrt{\eta}p'$  belongs to  $L^1_{\text{loc}}(\mathbb{R}_+)$ . Moreover, in the case where  $p(0)$  is finite, then the formula (7) can be extended to the whole  $\mathbb{R}$ , leading to an odd function. We additionally suppose that the following relations between  $\xi$ ,  $\eta$  and  $\Gamma$  hold:

(A6) There exists  $C > 0$  such that

$$(8) \quad 0 \leq \xi(u) \leq C(1 + \Gamma(u)), \quad \forall u \in [0, +\infty).$$

Moreover, we assume that

$$(9) \quad \frac{\Gamma(u)}{\eta(u)} \rightarrow +\infty \quad \text{as } u \rightarrow +\infty,$$

and that

$$(10) \quad \sqrt{\eta \circ \xi^{-1}} \text{ is uniformly continuous on the range of } \xi.$$

**Definition 1** (weak solution). *A measurable function  $u$  is said to be a weak solution to problem (1) if*

- i. *the functions  $u$  and  $\eta(u)$  belong to  $L^\infty((0, T); L^1(\Omega))$ ;*
- i. *the function  $\xi(u)$  belongs to  $L^2((0, t_f); H^1(\Omega))$ ;*
- i. *for all function  $\psi \in C_c^\infty(\overline{\Omega} \times [0, t_f]; \mathbb{R})$ , one has*

$$(11) \quad \iint_{Q_{t_f}} u \partial_t \psi \, d\mathbf{x} dt + \int_{\Omega} u_0 \psi(\cdot, 0) \, d\mathbf{x} \\ - \iint_{Q_{t_f}} \eta(u) \mathbf{\Lambda} \nabla V \cdot \nabla \psi \, d\mathbf{x} dt - \iint_{Q_{t_f}} \mathbf{\Lambda} \nabla \xi(u) \cdot \sqrt{\eta(u)} \nabla \psi \, d\mathbf{x} dt = 0.$$

Following the seminal work of [1], there exists at least one weak solution  $u$  to the problem (1). Denoting by

$$\phi(u) = \int_0^u \eta(a) p'(a) \, da, \quad \forall u \in \mathcal{I}_p,$$

the uniqueness of the solution (and even a  $L^1$ -contraction principle) is ensured as soon as  $\eta \circ \phi^{-1} \in C^{0,1/2}$  (see [36], and [5] for a slightly weaker condition in the case of a smooth domain  $\Omega$ ). Moreover,  $u$  belongs to  $C([0, T]; L^1_+(\Omega))$  (cf. [13]).

**1.2. Formal gradient flow structure of the continuous problem.** Let us first highlight the gradient flow structure of the system (1). Following the path proposed in [37, §1.3] (see also [35, 38]), the calculations carried out in this section are formal. They can be made rigorous under the non-degeneracy assumption  $\eta(u) \geq \alpha > 0$  for all  $u \geq 0$ .

Define the affine space

$$\mathfrak{M} = \left\{ u : \Omega \rightarrow \mathbb{R} \mid \int_{\Omega} u(\mathbf{x}) \, d\mathbf{x} = \int_{\Omega} u_0(\mathbf{x}) \, d\mathbf{x} \right\}$$

of the admissible states, called *state space*.

In order to define a Riemannian geometry on  $\mathfrak{M}$ , we need to introduce the tangent space  $T_u \mathfrak{M}$ , given by

$$T_u \mathfrak{M} = \left\{ w : \Omega \rightarrow \mathbb{R} \mid \int_{\Omega} w(\mathbf{x}) \, d\mathbf{x} = 0 \right\}.$$

We also need to define the metric tensor  $\mathfrak{g}_u : T_u\mathfrak{M} \times T_u\mathfrak{M} \rightarrow \mathbb{R}$ , which consists in a scalar product on  $T_u\mathfrak{M}$  (depending on the state  $u$ )

$$(12) \quad \mathfrak{g}_u(w_1, w_2) = \int_{\Omega} \phi_1 w_2 \, d\mathbf{x} = \int_{\Omega} w_1 \phi_2 \, d\mathbf{x} = \int_{\Omega} \eta(u) \nabla \phi_1 \cdot \Lambda \nabla \phi_2 \, d\mathbf{x},$$

for all  $w_1, w_2 \in T_u\mathfrak{M}$ , where  $\phi_i$  are defined *via* the elliptic problem

$$(13) \quad \begin{cases} -\nabla \cdot (\eta(u) \Lambda \nabla \phi_i) = w_i & \text{in } \Omega, \\ \eta(u) \Lambda \nabla \phi_i \cdot \mathbf{n} = 0 & \text{on } \partial\Omega, \\ \int_{\Omega} \phi_i \, d\mathbf{x} = 0. \end{cases}$$

Note that  $T_u\mathfrak{M}$  does not depend on  $u$ , but that the metric tensor  $\mathfrak{g}_u(\cdot, \cdot)$  does, so that we are not in a hilbertian framework.

Define the *entropy* functional

$$(14) \quad \mathfrak{E} : \begin{cases} \mathfrak{M} \rightarrow \mathbb{R} \cup \{+\infty\} \\ u \mapsto \mathfrak{E}(u) = \int_{\Omega} (\Gamma(u(\mathbf{x})) + u(\mathbf{x})V(\mathbf{x})) \, d\mathbf{x}, \end{cases}$$

and the *free-enthalpy* function

$$\mathfrak{h} : \begin{cases} \mathcal{I}_p \times \Omega \rightarrow \mathbb{R} \\ (u, \mathbf{x}) \mapsto \mathfrak{h}(u, \mathbf{x}) = p(u) + V(\mathbf{x}) = D_u \mathfrak{E}(u), \end{cases}$$

then given  $w \in T_u\mathfrak{M}$ , one has

$$(15) \quad D_u \mathfrak{E}(u) \cdot w = \int_{\Omega} \mathfrak{h}(u(\mathbf{x}), \mathbf{x}) w(\mathbf{x}) \, d\mathbf{x} = \int_{\Omega} \eta(u) \nabla \mathfrak{h}(u, \cdot) \cdot \Lambda \nabla \phi \, d\mathbf{x},$$

where  $\phi$  is deduced from  $w$  thanks to the elliptic problem (13). Moreover, thanks to (12), one has

$$(16) \quad \mathfrak{g}_u(\partial_t u, w) = \int_{\Omega} \partial_t u \phi \, d\mathbf{x}, \quad \forall w \in T_u\mathfrak{M}.$$

In view of (15)–(16), the problem (1) turns to

$$(17) \quad \mathfrak{g}_u(\partial_t u, w) = -D_u \mathfrak{E}(u) \cdot w = -\mathfrak{g}_u(\nabla_u \mathfrak{E}, w), \quad \forall w \in T_u\mathfrak{M},$$

where the cotangent vector  $D_u \mathfrak{E}(u) \in (T_u\mathfrak{M})^*$  has been identified to the tangent vector  $\nabla_u \mathfrak{E}(u) \in T_u\mathfrak{M}$  thanks to Riesz theorem applied with the scalar product  $g_u$ .

Choosing  $w = \partial_t u$  in (17) yields the classical entropy/dissipation relation:  $\forall t \in [0, t_f]$ ,

$$(18) \quad \begin{aligned} & \mathfrak{E}(u(\cdot, t)) - \mathfrak{E}(u_0) \\ & + \int_0^t \int_{\Omega} \eta(u(\mathbf{x}, \tau)) \Lambda(\mathbf{x}) \nabla \mathfrak{h}(u(\mathbf{x}, \tau), \mathbf{x}) \cdot \nabla \mathfrak{h}(u(\mathbf{x}, \tau), \mathbf{x}) \, d\mathbf{x} \, d\tau = 0. \end{aligned}$$

**1.3. Goal and positioning of the paper.** The goal of this paper is to provide a numerical scheme that mimics at the discrete level the gradient flow structure highlighted in §1.2. Since the point of view adopted in our presentation concerning the gradient flow structure is formal, the rigorous numerical analysis of the scheme will rather rely on the well established theory of weak solutions in the sense of Definition 1.

Our scheme is able to handle highly anisotropic and heterogeneous diffusion tensor  $\Lambda$  and very general grids (non-conformal grids, cells of various shapes), relying on the recently developed VAG method [25, 27, 26, 12]. This was already the case of a nonlinear scheme proposed in our previous contributions [15, 14], but it appeared in the numerical simulations that the convergence of the scheme proposed in [15, 14] was too slow, in particular in presence of strong anisotropy. The main goal of this paper was to propose a scheme that preserved some important features of the scheme introduced in [15, 14] (allowance of some nonlinear test function, decay of physically motivated entropy, convergence proof for discretization parameters tending to 0), without jeopardizing the accuracy of the scheme compared to the more classical approach consisting in approximation the weak formulation (11). Convincing numerical results are provided in §5 as an evidence of the efficiency of our approach, while two theorems are stated in §2.4 (and proved in §3 and §4) in order to ensure the following properties.

- (1) Theorem 2.3. At a fixed mesh, the scheme, that consists in a nonlinear system, admits (at least) one solution. This allows in particular to speak about the discrete solution provided by the scheme. Moreover, we take advantage of the gradient structure of the scheme for deriving some nonlinear stability estimates.
- (2) Theorem 2.4. Letting the discretization parameters tend to 0 (while controlling some regularity factors related to the discretization), the discrete solution converges in some appropriate sense towards the unique weak solution to the problem (1) in the sense of Definition 1.

## 2. DEFINITION OF THE SCHEME AND MAIN RESULTS

As already mentioned, the scheme we propose is based on the so-called VAG scheme [25]. In §2.1, we state our assumptions on the spatial mesh and the time discretization of  $(0, t_f)$ . Then in §2.2, we define the nonlinear scheme we will study in this paper. The gradient flow structure of the discretized problem is highlighted in §2.3, where a variational interpretation is given to the scheme. Finally, in §2.4, we state the existence of discrete solutions to the scheme and its convergence towards the unique weak solution as the discretization parameters tend to 0.

### 2.1. Discretization of $Q_{t_f}$ and discrete functional spaces.

**2.1.1. Spatial discretization and discrete reconstruction operators.** Following [26, 12], we consider generalized polyhedral discretizations of  $\Omega$ . Let  $\mathcal{M}$  be the set of the cell, that are disjoint polyhedral open subsets of  $\Omega$  such that  $\bigcup_{\kappa \in \mathcal{M}} \bar{\kappa} = \bar{\Omega}$ . Each cell  $\kappa \in \mathcal{M}$  is supposed to be star-shaped with respect to its “center”, denoted by  $\mathbf{x}_\kappa$ . We denote by  $\mathcal{F}$  the set of the faces of the mesh, which are not assumed to be planar if  $d = 3$  (whence the term “generalized polyhedral”). We denote by  $\mathcal{V}$  the set of the vertices of the mesh. We denote by  $\mathbf{x}_s \in \bar{\Omega}$  the location of the vertex  $s \in \mathcal{V}$ . The sets  $\mathcal{V}_\kappa$ ,  $\mathcal{F}_\kappa$  and  $\mathcal{V}_\sigma$  denote respectively the vertices and faces of a cell  $\kappa$ ,

and the vertices of a face  $\sigma$ . For any face  $\sigma \in \mathcal{F}_\kappa$ , one has  $\mathcal{V}_\sigma \subset \mathcal{V}_\kappa$ . Let  $\mathcal{M}_s$  denote the set of the cells sharing the vertex  $s$ . The set of edges of the mesh (defined only if  $d = 3$ ) is denoted by  $\mathcal{E}$  and  $\mathcal{E}_\sigma$  denotes the set of edges of the face  $\sigma \in \mathcal{F}$ , while  $\mathcal{E}_\kappa$  denotes the set of the edges of the cell  $\kappa$ . The set  $\mathcal{V}_e$  denotes the pair of vertices at the extremities of the edge  $e \in \mathcal{E}$ . In the 3-dimensional case, it is assumed that for each face  $\sigma \in \mathcal{F}$ , there exists a so-called “center” of the face  $\mathbf{x}_\sigma$  such that

$$(19) \quad \mathbf{x}_\sigma = \sum_{s \in \mathcal{V}_\sigma} \beta_{\sigma,s} \mathbf{x}_s, \quad \text{with} \quad \sum_{s \in \mathcal{V}_\sigma} \beta_{\sigma,s} = 1,$$

and  $\beta_{\sigma,s} \geq 0$  for all  $s \in \mathcal{V}_\sigma$ . The face  $\sigma$  is then assumed to match with the union of the triangles  $T_{\sigma,e}$  defined as by the face center  $\mathbf{x}_\sigma$  and each of its edge  $e \in \mathcal{E}_\sigma$ .

The previous discretization is denoted by  $\mathcal{D}$ , and we define the discrete space

$$W_{\mathcal{D}} = \left\{ \mathbf{v} = (v_\kappa, v_s)_{\kappa \in \mathcal{M}, s \in \mathcal{V}} \in \mathbb{R}^{\#\mathcal{M} + \#\mathcal{V}} \right\}.$$

In the 3-dimensional case, we introduce for all  $\sigma \in \mathcal{F}$  the operator  $I_\sigma : W_{\mathcal{D}} \rightarrow \mathbb{R}$  defined by

$$I_\sigma(\mathbf{v}) = \sum_{s \in \mathcal{V}_\sigma} \beta_{\sigma,s} v_s, \quad \forall \mathbf{v} \in W_{\mathcal{D}},$$

yielding a second order interpolation at  $\mathbf{x}_\sigma$  thanks to the definition (19) of  $\mathbf{x}_\sigma$ .

We introduce the simplicial submesh  $\mathcal{T}$  defined by

- $\mathcal{T} = \{T_{\kappa,\sigma}, \kappa \in \mathcal{M}, \sigma \in \mathcal{F}_\kappa\}$  in the two-dimensional case, where  $T_{\kappa,\sigma}$  denotes the triangle whose vertices are  $\mathbf{x}_\kappa$  and  $\mathbf{x}_s$  for  $s \in \mathcal{V}_\sigma$ ;
- $\mathcal{T} = \{T_{\kappa,\sigma,e}, \kappa \in \mathcal{M}, \sigma \in \mathcal{F}_\kappa, e \in \mathcal{E}_\sigma\}$  in the three-dimensional case, where  $T_{\kappa,\sigma,e}$  denotes the tetrahedron whose vertices are  $\mathbf{x}_\kappa$ ,  $\mathbf{x}_\sigma$  and  $\mathbf{x}_s$  for  $s \in \mathcal{V}_e$ .

We define the regularity  $\theta_{\mathcal{T}}$  of the simplicial mesh  $\mathcal{T}$  by

$$(20) \quad \theta_{\mathcal{T}} = \max_{T \in \mathcal{T}} \frac{h_T}{\rho_T},$$

where  $h_T$  and  $\rho_T$  respectively denote the diameter of  $T$  and the insphere diameter of  $T$ . We denote by

$$(21) \quad h_{\mathcal{T}} = \max_{T \in \mathcal{T}} h_T$$

the maximum diameter of the simplicial mesh. We also define the quantities  $\ell_\kappa$  and  $\ell_s$  quantifying the number of vertices of the cell  $\kappa$  and the number of neighboring cells for the vertex  $s$  respectively:

$$(22) \quad \ell_\kappa = \#\mathcal{V}_\kappa, \quad \ell_s = \#\mathcal{M}_s, \quad \forall \kappa \in \mathcal{M}, \forall s \in \mathcal{V}.$$

This allows to introduce the quantity

$$(23) \quad \ell_{\mathcal{D}} = \max \left( \max_{\beta \in \mathcal{M} \cup \mathcal{V}} \ell_\beta, \max_{\kappa \in \mathcal{M}} \{\#\mathcal{F}_\kappa\} \right),$$

controlling the regularity of the general discretization  $\mathcal{D}$  of  $\Omega$ .

Denoting by  $H_{\mathcal{T}} \subset W^{1,\infty}(\Omega)$  the usual  $\mathbb{P}_1$ -finite element space on the simplicial mesh  $\mathcal{T}$ , we define the reconstruction operator  $\pi_{\mathcal{T}} : W_{\mathcal{D}} \rightarrow H_{\mathcal{T}}$  by setting, for all  $\mathbf{v} \in W_{\mathcal{D}}$  and all  $(s, \kappa, \sigma) \in \mathcal{V} \times \mathcal{M} \times \mathcal{F}$ ,

$$(24) \quad \pi_{\mathcal{T}} \mathbf{v}(\mathbf{x}_s) = v_s, \quad \pi_{\mathcal{T}} \mathbf{v}(\mathbf{x}_\kappa) = v_\kappa, \quad \text{and} \quad \pi_{\mathcal{T}} \mathbf{v}(\mathbf{x}_\sigma) = I_\sigma(\mathbf{v}).$$

This allows to define the operator  $\nabla_{\mathcal{T}} : W_{\mathcal{D}} \rightarrow (L^\infty(\Omega))^d$  by

$$(25) \quad \nabla_{\mathcal{T}} \mathbf{v} = \nabla \pi_{\mathcal{T}} \mathbf{v}, \quad \forall \mathbf{v} \in W_{\mathcal{D}}.$$

We aim now to reconstruct piecewise constant functions. To this end, we introduce additional parameters that appear to play an important role in practical applications [27]. Let  $\kappa \in \mathcal{M}$ , then we introduce some weights  $(\alpha_{\kappa,s})_{s \in \mathcal{V}_\kappa}$  such that

$$(26) \quad \alpha_{\kappa,s} \geq 0, \quad \text{and} \quad \sum_{s \in \mathcal{V}_\kappa} \alpha_{\kappa,s} \leq 1, \quad \forall \kappa \in \mathcal{M}.$$

Denoting by  $\text{meas}(\kappa) = \int_\kappa d\mathbf{x}$  the volume of  $\kappa$ , then we define the quantities

$$(27) \quad \begin{cases} m_{\kappa,s} = \alpha_{\kappa,s} \text{meas}(\kappa), & \forall \kappa \in \mathcal{M}, \forall s \in \mathcal{V}_\kappa, \\ m_s = \sum_{\kappa \in \mathcal{M}_s} m_{\kappa,s}, & \forall s \in \mathcal{V}, \\ m_\kappa = \text{meas}(\kappa) - \sum_{s \in \mathcal{V}_\kappa} m_{\kappa,s}, & \forall \kappa \in \mathcal{M}, \end{cases}$$

so that one has

$$\sum_{\beta \in \mathcal{M} \cup \mathcal{V}} m_\beta = \text{meas}(\Omega).$$

For all  $\kappa \in \mathcal{M}$ , we denote by  $\omega_\kappa$  and  $\omega_{\kappa,s}$  some disjointed open subsets of  $\kappa$ , such that

$$\overline{\omega}_\kappa \cup \left( \bigcup_{s \in \mathcal{V}_\kappa} \overline{\omega}_{\kappa,s} \right) = \overline{\kappa},$$

and such that

$$\text{meas}(\omega_\kappa) = m_\kappa \quad \text{and} \quad \text{meas}(\omega_{\kappa,s}) = m_{\kappa,s}, \quad \forall \kappa \in \mathcal{M}, \forall s \in \mathcal{V}_\kappa.$$

Note that such a decomposition always exists thanks to (26)–(27). Then we denote by

$$\omega_s = \bigcup_{\kappa \in \mathcal{M}_s} \omega_{\kappa,s}, \quad \forall s \in \mathcal{V}.$$

In what follows, we denote by

$$(28) \quad \zeta_{\mathcal{D}} = \min_{\beta \in \mathcal{M} \cup \mathcal{V}} \frac{m_\beta}{\int_\Omega \pi_{\mathcal{T}} \mathbf{e}_\beta d\mathbf{x}},$$

where  $\mathbf{e}_\beta$ ,  $\beta \in \mathcal{M} \cup \mathcal{V}$  is the unique element of  $W_{\mathcal{D}}$  such that

$$\pi_{\mathcal{T}} \mathbf{e}_\beta(\mathbf{x}_\beta) = 1, \quad \pi_{\mathcal{T}} \mathbf{e}_\beta(\mathbf{x}_{\beta'}) = 0, \quad \forall \beta' \in \mathcal{M} \cup \mathcal{V}, \beta' \neq \beta.$$

We can now define the piecewise constant reconstruction operators  $\pi_{\mathcal{D}} : W_{\mathcal{D}} \rightarrow L^\infty \cap \text{BV}(\Omega)$  and  $\pi_{\mathcal{M}} : W_{\mathcal{D}} \rightarrow L^\infty \cap \text{BV}(\Omega)$  by

$$(29) \quad \pi_{\mathcal{D}}(\mathbf{v})(\mathbf{x}) = \sum_{\kappa \in \mathcal{M}} v_\kappa \mathbf{1}_{\omega_\kappa}(\mathbf{x}) + \sum_{s \in \mathcal{V}} v_s \mathbf{1}_{\omega_s}(\mathbf{x}), \quad \forall \mathbf{x} \in \Omega, \quad \forall \mathbf{v} \in W_{\mathcal{D}}.$$

and

$$(30) \quad \pi_{\mathcal{M}}(\mathbf{v})(\mathbf{x}) = \sum_{\kappa \in \mathcal{M}} v_\kappa \mathbf{1}_\kappa(\mathbf{x}) \quad \forall \mathbf{x} \in \Omega, \quad \forall \mathbf{v} \in W_{\mathcal{D}}.$$

Let  $f : \mathbb{R} \rightarrow \mathbb{R}$  be a possibly nonlinear function, then we denote by

$$f(\mathbf{v}) = (f(v_\kappa), f(v_s))_{\kappa \in \mathcal{M}, s \in \mathcal{V}}, \quad \forall \mathbf{v} = (v_\kappa, v_s)_{\kappa \in \mathcal{M}, s \in \mathcal{V}} \in W_{\mathcal{D}}.$$

Notice that in general,

$$\pi_{\mathcal{T}}(f(\mathbf{v})) \neq f(\pi_{\mathcal{T}}(\mathbf{v})) \quad \text{and} \quad \nabla_{\mathcal{T}}(f(\mathbf{v})) \neq \nabla f(\pi_{\mathcal{T}}(\mathbf{v})),$$



but that

$$(31) \quad \pi_{\mathcal{D}}(f(\mathbf{v})) = f(\pi_{\mathcal{D}}(\mathbf{v})) \quad \text{and} \quad \pi_{\mathcal{M}}(f(\mathbf{v})) = f(\pi_{\mathcal{M}}(\mathbf{v})), \quad \forall \mathbf{v} \in W_{\mathcal{D}}.$$

**2.1.2. Time-and-space discretizations and discrete functions.** Let  $N \geq 1$  and let  $0 = t_0 < t_1 < \dots < t_{N-1} < t_N = t_f$  be some subdivision of  $[0, t_f]$ . We denote by  $\Delta t_n = t_n - t_{n-1}$  for all  $n \in \{1, \dots, N\}$ , by  $\Delta t = (\Delta t_1, \dots, \Delta t_N)^T \in \mathbb{R}^N$ , and by

$$(32) \quad \overline{\Delta t} = \max_{1 \leq n \leq N} \Delta t_n.$$

The time and space discrete space is then defined by

$$W_{\mathcal{D}, \Delta t} = \left\{ \mathbf{v} = (v_{\kappa}^n, v_s^n)_{\kappa \in \mathcal{M}, s \in \mathcal{V}, 1 \leq n \leq N} \in \mathbb{R}^{N(\#\mathcal{M} + \#\mathcal{V})} \right\}.$$

For  $\mathbf{v} \in W_{\mathcal{D}, \Delta t}$  and  $n \in \{1, \dots, N\}$ , we denote by

$$\mathbf{v}^n = (v_{\kappa}^n, v_s^n)_{\kappa \in \mathcal{M}, s \in \mathcal{V}} \in W_{\mathcal{D}}.$$

The time and space reconstructions operators  $\pi_{\mathcal{D}, \Delta t}, \pi_{\mathcal{T}, \Delta t} : W_{\mathcal{D}, \Delta t} \rightarrow L^\infty(Q_{t_f})$  yield to be constant w.r.t. time functions defined by

$$\pi_{\mathcal{D}, \Delta t} \mathbf{v}(\cdot, t) = \pi_{\mathcal{D}}(\mathbf{v}^n) \quad \text{and} \quad \pi_{\mathcal{T}, \Delta t} \mathbf{v}(\cdot, t) = \pi_{\mathcal{T}}(\mathbf{v}^n) \quad \text{if } t \in (t_{n-1}, t_n],$$

the gradient reconstruction operator  $\nabla_{\mathcal{T}, \Delta t} : W_{\mathcal{D}, \Delta t} \rightarrow (L^\infty(Q_{t_f}))^d$  is then defined by

$$\nabla_{\mathcal{T}, \Delta t} \mathbf{v} = \nabla \pi_{\mathcal{T}, \Delta t} \mathbf{v}, \quad \forall \mathbf{v} \in W_{\mathcal{D}, \Delta t}.$$

**2.2. The nonlinear scheme for degenerate parabolic equations.** For  $\kappa \in \mathcal{M}$ , we denote by  $\mathbf{A}_{\kappa} = (a_{s,s'}^{\kappa})_{s,s' \in \mathcal{V}_{\kappa}} \in \mathbb{R}^{\ell_{\kappa} \times \ell_{\kappa}}$  the symmetric definite positive matrix whose coefficients are defined by

$$(33) \quad a_{s,s'}^{\kappa} = \int_{\kappa} \Lambda(\mathbf{x}) \nabla_{\mathcal{T}} \mathbf{e}_s(\mathbf{x}) \cdot \nabla_{\mathcal{T}} \mathbf{e}_{s'}(\mathbf{x}) d\mathbf{x} = a_{s',s}^{\kappa}.$$

It results from the relation

$$\pi_{\mathcal{T}} \mathbf{e}_{\kappa}(\mathbf{x}) + \sum_{s \in \mathcal{V}_{\kappa}} \pi_{\mathcal{T}} \mathbf{e}_s(\mathbf{x}) = 1, \quad \forall \mathbf{x} \in \kappa, \quad \forall \kappa \in \mathcal{M}.$$

that, for all  $\mathbf{u}, \mathbf{v} \in W_{\mathcal{D}}$  and all  $\kappa \in \mathcal{M}$ , one has

$$(34) \quad \int_{\kappa} \Lambda(\mathbf{x}) \nabla_{\mathcal{T}} \mathbf{u}(\mathbf{x}) \cdot \nabla_{\mathcal{T}} \mathbf{v}(\mathbf{x}) d\mathbf{x} = \sum_{s \in \mathcal{V}_{\kappa}} \sum_{s' \in \mathcal{V}_{\kappa}} a_{s,s'}^{\kappa} (u_{\kappa} - u_s)(v_{\kappa} - v_{s'}).$$

For  $\kappa \in \mathcal{M}$ , we denote by  $\delta_{\kappa} : W_{\mathcal{D}} \rightarrow \mathbb{R}^{\ell_{\kappa}}$  the linear operator defined by

$$(\delta_{\kappa} \mathbf{v})_s = v_{\kappa} - v_s, \quad \forall s \in \mathcal{V}_{\kappa}, \quad \forall \mathbf{v} \in W_{\mathcal{D}}.$$

With this notation, we obtain that (34) rewrites

$$\int_{\kappa} \Lambda(\mathbf{x}) \nabla_{\mathcal{T}} \mathbf{u}(\mathbf{x}) \cdot \nabla_{\mathcal{T}} \mathbf{v}(\mathbf{x}) d\mathbf{x} = \delta_{\kappa} \mathbf{v} \cdot \mathbf{A}_{\kappa} \delta_{\kappa} \mathbf{u}, \quad \forall \mathbf{u}, \mathbf{v} \in W_{\mathcal{D}}, \quad \forall \kappa \in \mathcal{M}.$$

In order to deal with the nonlinearities of the problem, we introduce the sets  $W_{\mathcal{D}}^{\text{ad}} \subset W_{\mathcal{D}}$  and  $W_{\mathcal{D}, \Delta t}^{\text{ad}} \subset W_{\mathcal{D}, \Delta t}$  of the admissible states defined by

$$\mathbf{v} \in W_{\mathcal{D}}^{\text{ad}} \quad \text{iff} \quad v_{\nu} \in \mathcal{I}_p, \quad \forall \nu \in \mathcal{M} \cup \mathcal{V},$$

and

$$\mathbf{v} \in W_{\mathcal{D}, \Delta t}^{\text{ad}} \quad \text{iff} \quad \mathbf{v}^n \in W_{\mathcal{D}}^{\text{ad}}, \quad \forall n \in \{1, \dots, N\},$$

while we denote by  $W_{\mathcal{D}}^{\text{en}} \subset W_{\mathcal{D}}$  the set of finite entropy vectors:

$$(35) \quad \mathbf{v} \in W_{\mathcal{D}}^{\text{en}} \quad \text{iff} \quad \mathfrak{E}_{\mathcal{D}}(\mathbf{u}) := \int_{\Omega} \Gamma(\pi_{\mathcal{D}} \mathbf{u}) + \pi_{\mathcal{D}}(\mathbf{u}) \pi_{\mathcal{D}}(\mathbf{V}) d\mathbf{x} < \infty,$$

where  $\mathbf{V} = (V_{\kappa}, V_{\mathbf{s}})_{\kappa, \mathbf{s}} \in W_{\mathcal{D}}$  is defined by

$$(36) \quad V_{\kappa} = V(\mathbf{x}_{\kappa}) \quad V_{\mathbf{s}} = V(\mathbf{x}_{\mathbf{s}}) \quad \forall \kappa \in \mathcal{M}, \forall \mathbf{s} \in \mathcal{V}.$$

It is easy to check that, thanks to assumptions (A2) and to the definition (4) of the convex function  $\Gamma$ , one has  $W_{\mathcal{D}}^{\text{ad}} \subset W_{\mathcal{D}}^{\text{en}}$ .

Given  $\mathbf{u} \in W_{\mathcal{D}}^{\text{ad}}$  and  $\mathbf{V} \in W_{\mathcal{D}}$ , we define the discrete free-enthalpy  $\mathbf{h}(\mathbf{u}) = (\mathfrak{h}_{\kappa}(u_{\kappa}), \mathfrak{h}_{\mathbf{s}}(u_{\mathbf{s}}))_{\kappa, \mathbf{s}} \in W_{\mathcal{D}}$  by

$$(37) \quad \mathfrak{h}_{\kappa}(u_{\kappa}) = p(u_{\kappa}) + V_{\kappa}, \quad \mathfrak{h}_{\mathbf{s}}(u_{\mathbf{s}}) = p(u_{\mathbf{s}}) + V_{\mathbf{s}}, \quad \forall \kappa \in \mathcal{M}, \forall \mathbf{s} \in \mathcal{V}.$$

The initial data  $u_0$  is discretized into an element  $\mathbf{u}^0 \in W_{\mathcal{D}}^{\text{en}}$  by

$$(38) \quad u_{\beta}^0 = \frac{1}{m_{\beta}} \int_{\omega_{\beta}} u_0(\mathbf{x}) d\mathbf{x}, \quad \forall \beta \in \mathcal{M} \cup \mathcal{V},$$

so that, thanks to (5),

$$(39) \quad \int_{\Omega} \pi_{\mathcal{D}}(\mathbf{u}^0) d\mathbf{x} = \int_{\Omega} u_0 d\mathbf{x} > 0.$$

Let us state a first lemma that ensure that the discretized initial data has a finite discrete entropy.

**Lemma 2.1.** *Let  $u_0 \in L^1(\Omega)$  be such that (A4) holds, let  $V$  be such that (AA5) holds. Let  $\mathbf{u}^0$  be defined by (38) and  $\mathbf{V}$  defined by (36), then there exists  $C$  depending only on  $\|u_0\|_{L^1(\Omega)}$  and  $\|\nabla V\|_{L^{\infty}(\Omega)^d}$  such that*

$$(40) \quad \mathfrak{E}_{\mathcal{D}}(\mathbf{u}^0) \leq \mathfrak{E}(u_0) + Ch_{\mathcal{T}} \leq \mathfrak{E}(u_0) + C \text{diam}(\Omega),$$

where the entropy functional  $\mathfrak{E}$  is defined by (14) and its discrete counterpart  $\mathfrak{E}_{\mathcal{D}}$  is defined by (35). In particular,  $\mathbf{u}^0$  belongs to  $W_{\mathcal{D}}^{\text{en}}$ .

*Proof.* We deduce from Jensen inequality that

$$\Gamma(u_{\beta}^0) \leq \frac{1}{m_{\beta}} \int_{\omega_{\beta}} \Gamma(u_0) d\mathbf{x},$$

whence, thanks to (AA4) and to the definition (35) of the discrete entropy functional  $\mathfrak{E}_{\mathcal{D}}$ , one has

$$\mathfrak{E}_{\mathcal{D}}(\mathbf{u}^0) \leq \mathfrak{E}(u_0) + \int_{\Omega} u_0(\pi_{\mathcal{D}} \mathbf{V} - V) d\mathbf{x} + \int_{\Omega} (\pi_{\mathcal{D}} \mathbf{u}^0 - u) \pi_{\mathcal{D}} \mathbf{V} d\mathbf{x}.$$

The last term in the above inequality is equal to zero thanks to the definition (38) of  $\mathbf{u}^0 = (u_{\kappa}^0, u_{\mathbf{s}}^0)_{\kappa, \mathbf{s}}$ . The Lipschitz regularity of  $V$  yields

$$\|\pi_{\mathcal{D}} \mathbf{V} - V\|_{L^{\infty}(\Omega)} \leq \|\nabla V\|_{L^{\infty}(\Omega)^d} h_{\mathcal{T}},$$

so that

$$\mathfrak{E}_{\mathcal{D}}(\mathbf{u}^0) \leq \mathfrak{E}(u_0) + \|u_0\|_{L^1(\Omega)} \|\nabla V\|_{L^{\infty}(\Omega)^d} h_{\mathcal{T}}.$$

□

With all this material, we can present the scheme we will analyze in this contribution. For  $\mathbf{u} \in W_{\mathcal{D}, \Delta t}^{\text{ad}}$ , we introduce the notation

$$\eta_{\kappa, s}^n = \frac{\eta(u_{\kappa}^n) + \eta(u_s^n)}{2}, \quad \forall \kappa \in \mathcal{M}, \forall s \in \mathcal{V}_{\kappa}, \forall n \in \{1, \dots, N\}.$$

Given  $\mathbf{u}^{n-1} \in W_{\mathcal{D}}^{\text{en}}$ , the vector  $\mathbf{u}^n \in W_{\mathcal{D}}^{\text{ad}}$  is obtained by solving the following nonlinear system:

$$(41a) \quad m_{\kappa} \frac{u_{\kappa}^n - u_{\kappa}^{n-1}}{\Delta t_n} + \sum_{s \in \mathcal{V}_{\kappa}} F_{\kappa, s}(\mathbf{u}^n) = 0, \quad \forall \kappa \in \mathcal{M},$$

$$(41b) \quad m_s \frac{u_s^n - u_s^{n-1}}{\Delta t_n} + \sum_{\kappa \in \mathcal{M}_s} F_{s, \kappa}(\mathbf{u}^n) = 0, \quad \forall s \in \mathcal{V},$$

$$(41c) \quad F_{\kappa, s}(\mathbf{u}^n) = \sqrt{\eta_{\kappa, s}^n} \sum_{s' \in \mathcal{V}_{\kappa}} a_{s, s'}^{\kappa} \sqrt{\eta_{\kappa, s'}^n} (p(u_{\kappa}^n) + V_{\kappa} - p(u_{s'}^n) - V_{s'}),$$

$$\forall \kappa \in \mathcal{M}, \forall s \in \mathcal{V}_{\kappa},$$

$$(41d) \quad F_{\kappa, s}(\mathbf{u}^n) + F_{s, \kappa}(\mathbf{u}^n) = 0, \quad \forall \kappa \in \mathcal{M}, \forall s \in \mathcal{V}_{\kappa}.$$

The scheme (41) is clearly a finite volume scheme. As a direct consequence of the conservativity of the scheme, one has

$$(42) \quad \int_{\Omega} \pi_{\mathcal{D}}(\mathbf{u}^n) d\mathbf{x} = \int_{\Omega} \pi_{\mathcal{D}}(\mathbf{u}^{n-1}) d\mathbf{x}.$$

Defining, for all  $\kappa \in \mathcal{M}$  and  $\mathbf{u} = (u_{\kappa}, u_s)_{\kappa, s} \in W_{\mathcal{D}}$ , the diagonal matrix  $\mathbf{M}_{\kappa}(\mathbf{u}) \in \mathbb{R}^{\ell_{\kappa} \times \ell_{\kappa}}$  by

$$(\mathbf{M}_{\kappa}(\mathbf{u}))_{s, s'} = \begin{cases} \sqrt{\frac{\eta(u_{\kappa}) + \eta(u_s)}{2}} & \text{if } s = s', \\ 0 & \text{otherwise,} \end{cases}$$

the systems (41) is equivalent to the following compact formulation:  $\forall \mathbf{v} \in W_{\mathcal{D}}$ ,

$$(43) \quad \int_{\Omega} \pi_{\mathcal{D}} \mathbf{u}^n \pi_{\mathcal{D}} \mathbf{v} d\mathbf{x} + \Delta t_n \sum_{\kappa \in \mathcal{M}} \delta_{\kappa} \mathbf{h}(\mathbf{u}^n) \cdot \mathbf{B}_{\kappa}(\mathbf{u}^n) \delta_{\kappa} \mathbf{v} = \int_{\Omega} \pi_{\mathcal{D}} \mathbf{u}^{n-1} \pi_{\mathcal{D}} \mathbf{v} d\mathbf{x},$$

where

$$(44) \quad \mathbf{B}_{\kappa}(\mathbf{u}) := \mathbf{M}_{\kappa}(\mathbf{u}) \mathbf{A}_{\kappa} \mathbf{M}_{\kappa}(\mathbf{u}), \quad \forall \kappa \in \mathcal{M}, \forall \mathbf{u} \in W_{\mathcal{D}}$$

is a symmetric semi-positive matrix since  $\mathbf{A}_{\kappa}$  and  $\mathbf{M}_{\kappa}(\mathbf{u})$  are.

**2.3. Gradient flow interpretation for the scheme.** The goal of this section is to transpose the formal variational structure pointed out in §1.2 to the discrete setting. A natural discretization of the manifold  $\mathfrak{M}$  consists in

$$(45) \quad \mathfrak{M}_{\mathcal{D}} = \left\{ \mathbf{u} \in W_{\mathcal{D}} \mid \int_{\Omega} \pi_{\mathcal{D}} \mathbf{u} d\mathbf{x} = \int_{\Omega} u_0 d\mathbf{x} \right\},$$

leading to

$$(46) \quad T_{\mathbf{u}} \mathfrak{M}_{\mathcal{D}} = \left\{ \mathbf{v} \in W_{\mathcal{D}} \mid \int_{\Omega} \pi_{\mathcal{D}} \mathbf{v} d\mathbf{x} = 0 \right\}.$$

In order to define the discrete counterpart  $\mathbf{g}_{\mathcal{D}, \mathbf{u}}$  of the metric tensor  $\mathbf{g}_{\mathbf{u}}$  defined by (12)–(13), one needs a discrete counterpart of

- the classical  $L^2(\Omega)$  scalar product: for reasons appearing in the analysis, one will use

$$(\mathbf{w}_1, \mathbf{w}_2) \mapsto \int_{\Omega} \pi_{\mathcal{D}} \mathbf{w}_1 \pi_{\mathcal{D}} \mathbf{w}_2 \, d\mathbf{x}, \quad \forall \mathbf{w}_1, \mathbf{w}_2 \in W_{\mathcal{D}};$$

- the weighted  $H^1(\Omega)$  “scalar product” with weight  $\eta(u)$  : this can be done by choosing

$$(\mathbf{w}_1, \mathbf{w}_2) \mapsto \sum_{\kappa \in \mathcal{M}} \delta_{\kappa} \mathbf{w}_1 \cdot \mathbf{B}_{\kappa}(u) \delta_{\kappa} \mathbf{w}_2, \quad \forall \mathbf{w}_1, \mathbf{w}_2 \in W_{\mathcal{D}}.$$

This allows to define the discrete metric tensor  $\mathbf{g}_{\mathcal{D},u}$  by:  $\forall u \in W_{\mathcal{D}}, \forall \mathbf{w}_1, \mathbf{w}_2 \in T_u \mathfrak{M}_{\mathcal{D}}$ ,  
(47)

$$\mathbf{g}_{\mathcal{D},u}(\mathbf{w}_1, \mathbf{w}_2) = \int_{\Omega} \pi_{\mathcal{D}} \mathbf{w}_1 \pi_{\mathcal{D}} \phi_2 \, d\mathbf{x} = \int_{\Omega} \pi_{\mathcal{D}} \phi_1 \pi_{\mathcal{D}} \mathbf{v}_2 \, d\mathbf{x} = \sum_{\kappa \in \mathcal{M}} \delta_{\kappa} \phi_1 \cdot \mathbf{B}_{\kappa}(u) \delta_{\kappa} \phi_2,$$

where  $\phi_i \in W_{\mathcal{D}}$  solves the discrete counterpart of (13), that is

$$(48) \quad \sum_{\kappa \in \mathfrak{M}} \delta_{\kappa} \phi_i \cdot \mathbf{B}_{\kappa}(u) \delta_{\kappa} \psi = \int_{\Omega} \pi_{\mathcal{D}} \mathbf{w}_i \pi_{\mathcal{D}} \psi \, d\mathbf{x}, \quad \forall \psi \in W_{\mathcal{D}}.$$

In this setting, we can define the semi-discrete in space gradient flow by

$$(49) \quad \begin{aligned} \mathbf{g}_{\mathcal{D},u}(\partial_t \mathbf{u}, \mathbf{w}) &= \int_{\Omega} \pi_{\mathcal{D}}(\partial_t \mathbf{u}) \pi_{\mathcal{D}} \mathbf{v} \, d\mathbf{x} \\ &= \mathbf{g}_{\mathcal{D},u}(-\nabla_u \mathfrak{E}(u), \mathbf{w}) = \int_{\Omega} \pi_{\mathcal{D}} \mathbf{h}(u) \pi_{\mathcal{D}} \mathbf{w} \, d\mathbf{x} = \sum_{\kappa} \delta_{\kappa} \mathbf{h}(u) \cdot \mathbf{B}_{\kappa}(u) \delta_{\kappa} \mathbf{v}, \end{aligned}$$

where  $\mathbf{v}$  solve the discrete elliptic problem

$$\sum_{\kappa \in \mathfrak{M}} \delta_{\kappa} \mathbf{v} \cdot \mathbf{B}_{\kappa}(u) \delta_{\kappa} \psi = \int_{\Omega} \pi_{\mathcal{D}} \mathbf{w} \pi_{\mathcal{D}} \psi \, d\mathbf{x}, \quad \forall \psi \in W_{\mathcal{D}}.$$

In order to recover (43) from (49), one applies the implicit Euler scheme.

**Remark 2.2.** *In their seminal paper [31], Jordan, Kinderlehrer and Otto proposed to approximate the solution of gradient flows thanks to the minimizing movement scheme*

$$(50) \quad \mathbf{u}^n \in \operatorname{argmin}_{\mathbf{u} \in \mathfrak{M}_{\mathcal{D}}} \left\{ \frac{\mathfrak{d}(\mathbf{u}, \mathbf{u}^{n-1})}{2\Delta t} + \mathfrak{E}_{\mathcal{D}}(\mathbf{u}) \right\},$$

where  $\mathfrak{d}$  denotes the distance on  $\mathfrak{M}_{\mathcal{D}}$  induced by the metric tensor field  $\mathbf{g}_{\mathcal{D}}$ . Several practical and theoretical difficulties arise when one aims at using (50). First of all, the Riemannian structure is formal, even in the continuous case. It is unclear if one can define rigorously a distance  $\mathfrak{d}$  if  $\eta(0) = 0$ . But even if  $\mathfrak{d}$  is a distance, yielding a metric structure for  $\mathfrak{M}_{\mathcal{D}}$ , computing this distance is a complex problem we avoid by using an implicit Euler scheme rather than (50).

**2.4. Main results.** The first result we want to point out concerns the scheme for a fixed mesh. The following Proposition states that the scheme (41) admits at least one solution, and justifies the *entropy-diminishing* denomination for the scheme.

**Theorem 2.3.** *Let  $\mathbf{u}^{n-1} \in W_{\mathcal{D}}^{\text{en}}$ , then there exists (at least) one vector  $\mathbf{u}^n \in W_{\mathcal{D}}^{\text{ad}}$  solution to the system (41), and the following dissipation property holds:*

$$(51) \quad \mathfrak{E}_{\mathcal{D}}(\mathbf{u}^n) + \Delta t_n \sum_{\kappa \in \mathcal{M}} \delta_{\kappa} \mathbf{h}(\mathbf{u}^n) \cdot \mathbf{B}_{\kappa}(\mathbf{u}^n) \delta_{\kappa} \mathbf{h}(\mathbf{u}^n) \leq \mathfrak{E}_{\mathcal{D}}(\mathbf{u}^{n-1}),$$

where  $\mathfrak{E}_{\mathcal{D}}$  is defined by (35) and  $\mathbf{h}(\mathbf{u}^n) = (\mathfrak{h}_{\kappa}(u_{\kappa}^n), \mathfrak{h}_s(u_s^n))_{\kappa, s}$  is defined by (37).

Since  $\mathbf{u}^0 \in W_{\mathcal{D}}^{\text{en}}$  and since  $W_{\mathcal{D}}^{\text{ad}} \subset W_{\mathcal{D}}^{\text{en}}$ , Theorem 2.3 allows to define the *iterated solution*  $\mathbf{u} = (\mathbf{u}^n)_{1 \leq n \leq N} \in W_{\mathcal{D}, \Delta t}^{\text{ad}}$  to the scheme (41).

The proof of Theorem 2.3 is contained in §3, together with some supplementary material that allows to carry out the convergence analysis when the discretization steps tend to 0. More precisely, we consider a sequence  $(\mathcal{D}_m)_{m \geq 1} = (\mathcal{M}_m, \mathcal{T}_m)_{m \geq 1}$  of discretizations of  $\Omega$  as introduced in §2.1.1, such that

$$(52a) \quad h_{\mathcal{T}_m} = \max_{T \in \mathcal{T}_m} h_T \xrightarrow{m \rightarrow +\infty} 0,$$

and such that there exists  $\theta^* > 0$  and  $\ell^* > 0$  satisfying

$$(52b) \quad \sup_{m \geq 1} \theta_{\mathcal{T}_m} \leq \theta^*, \quad \sup_{m \geq 1} \ell_{\mathcal{D}_m} \leq \ell^*,$$

where  $\theta_{\mathcal{T}_m}$  and  $\ell_{\mathcal{D}_m}$  are defined by (20) and (23) respectively.

Despite it can be avoided in some specific situations, we do the following assumption, allowing to circumvent some technical difficulties:

$$(52c) \quad \inf_{m \geq 1} \zeta_{\mathcal{D}_m} = \zeta^* > 0.$$

This means that there is a minimum ratio of volume allocated to the cell centers and to the nodes in the mass lumping procedure.

Concerning the time-discretizations, we consider a sequence  $(\Delta t_m)_{m \geq 1}$  of discretizations of  $(0, t_f)$  as prescribed in §2.1.2 :

$$\Delta t_m = (\Delta t_{1,m}, \dots, \Delta t_{1,N_m}), \quad \forall m \geq 1.$$

We assume that the time discretization step tends to 0, i.e.,

$$(52d) \quad \overline{\Delta t}_m = \max_{1 \leq n \leq N_m} \Delta t_{n,m} \xrightarrow{m \rightarrow +\infty} 0,$$

where  $\Delta t_m^*$  is defined by (32).

**Theorem 2.4.** *Let  $(\mathcal{D}_m, \Delta t_m)_{m \geq 1}$  be a sequence of discretizations of  $Q_{t_f}$  satisfying Assumptions (52), and let  $(\mathbf{u}_m)_{m \geq 1}$ , with  $\mathbf{u}_m \in W_{\mathcal{D}_m, \Delta t_m}^{\text{ad}}$ , be a corresponding sequence of iterated discrete solutions, then*

$$\pi_{\mathcal{D}_m, \Delta t_m} \mathbf{u}_m \xrightarrow{m \rightarrow +\infty} u \quad \text{strongly in } L^1(Q_{t_f}),$$

where  $u$  is the unique weak solution to (1) in the sense of Definition 1.

Proving the Theorem 2.4 is the purpose of §4. The practical implementation of the scheme (41) is discussed in §5, where we also give evidences of the efficiency of the scheme.

### 3. PROOF OF THEOREM 2.3 AND ADDITIONAL ESTIMATES

In order to ease the reading of the paper, several technical lemmas have been postponed to Appendix.

### 3.1. One-step *A priori* estimates.

**Lemma 3.1.** *Let  $\mathbf{u}^{n-1} \in W_{\mathcal{D}}^{\text{en}}$ , and let  $\mathbf{u}^n \in W_{\mathcal{D}}^{\text{ad}}$  be a solution to the scheme (41), then (51) holds.*

*Proof.* Substituting  $\mathbf{v}$  by  $\mathbf{h}(\mathbf{u}^n) = (\mathbf{h}_{\kappa}(u_{\kappa}^n), \mathbf{h}_{\mathbf{s}}(u_{\mathbf{s}}^n))_{\kappa, \mathbf{s}}$ , defined by (37), in (43) yields

$$(53) \quad \int_{\Omega} (\pi_{\mathcal{D}} \mathbf{u}^n - \pi_{\mathcal{D}} \mathbf{u}^{n-1}) \pi_{\mathcal{D}} \mathbf{h}(\mathbf{u}^n) d\mathbf{x} + \Delta t_n \sum_{\kappa \in \mathcal{M}} \delta_{\kappa} \mathbf{h}(\mathbf{u}^n) \cdot \mathbf{B}_{\kappa}(\mathbf{u}^n) \delta_{\kappa} \mathbf{h}(\mathbf{u}^n) = 0.$$

It follows from the convexity of  $\Gamma$  that

$$\Gamma(a) - \Gamma(b) \leq (a - b)p(a), \quad \forall a, b \in \mathbb{R} \text{ s.t. } \Gamma(a), \Gamma(b) < +\infty,$$

whence

$$\begin{aligned} & \int_{\Omega} (\pi_{\mathcal{D}} \mathbf{u}^n - \pi_{\mathcal{D}} \mathbf{u}^{n-1}) \pi_{\mathcal{D}} \mathbf{h}(\mathbf{u}^n) d\mathbf{x} \\ &= \int_{\Omega} (\pi_{\mathcal{D}} \mathbf{u}^n - \pi_{\mathcal{D}} \mathbf{u}^{n-1}) (p(\pi_{\mathcal{D}} \mathbf{u}^n) + \pi_{\mathcal{D}}(\mathbf{V})) d\mathbf{x} \\ &\leq \int_{\Omega} (\Gamma(\pi_{\mathcal{D}} \mathbf{u}^n) - \Gamma(\pi_{\mathcal{D}} \mathbf{u}^{n-1}) + \pi_{\mathcal{D}}(\mathbf{u}^n - \mathbf{u}^{n-1}) \pi_{\mathcal{D}} \mathbf{V}) d\mathbf{x} \\ &= \mathfrak{E}_{\mathcal{D}}(\mathbf{u}^n) - \mathfrak{E}_{\mathcal{D}}(\mathbf{u}^{n-1}). \end{aligned}$$

Using this inequality in (53) provides (51).  $\square$

**Lemma 3.2.** *For all  $\epsilon > 0$ , there exists  $C_{\epsilon} \in \mathbb{R}$  depending on  $\epsilon$  and  $p$  such that*

$$|u| \leq \epsilon \Gamma(u) + C_{\epsilon}, \quad \forall u \in W_{\mathcal{D}}^{\text{en}}.$$

*Proof.* Fix  $\epsilon > 0$ , then in view of Assumption (A2), the intermediate value theorem ensures the existence of  $u_{\epsilon} \geq 1$  such that  $p(u_{\epsilon}) = p(1) + 1/\epsilon$ . Then for all  $u \in \overline{\mathcal{I}}_p$ , one has

$$\Gamma(u) = \int_1^u (p(a) - p(1)) da = \Gamma(u_{\epsilon}) + \int_{u_{\epsilon}}^u (p(a) - p(1)) da.$$

The function  $p$  being increasing, we deduce that

$$\Gamma(u) \geq \Gamma(u_{\epsilon}) + (p(u_{\epsilon}) - p(1))|u - u_{\epsilon}| \geq \Gamma(u_{\epsilon}) + \frac{1}{\epsilon} (|u| - |u_{\epsilon}|), \quad \forall u \in \overline{\mathcal{I}}_p,$$

Lemma 3.2 follows with  $C_{\epsilon} = |u_{\epsilon}| - \epsilon \Gamma(u_{\epsilon})$ .  $\square$

**Lemma 3.3.** *For all  $\epsilon > 0$ , there exists  $C_{\epsilon} \in \mathbb{R}$  depending on  $\epsilon$  and  $\eta$  and  $\Gamma$  such that*

$$\eta(u) \leq \epsilon \Gamma(u) + C_{\epsilon}, \quad \forall u \in \overline{\mathcal{I}}_p.$$

*Proof.* The function  $u \mapsto \frac{\eta(u)}{\Gamma(u)}$  tends to 0 as  $|u| \rightarrow \infty$  thanks to Assumption (A6). Let  $\epsilon > 0$ , then there exists  $r_{\epsilon} > 0$  such that

$$|u| > r_{\epsilon} \implies 0 \leq \eta(u) \leq \epsilon \Gamma(u).$$

The function  $\eta$  being continuous and nonnegative following Assumption (A1), we know that

$$0 \leq C_{\epsilon} := \max_{u \in [-r_{\epsilon}, r_{\epsilon}]} \eta(u) < +\infty.$$

The result of Lemma 3.3 follows.  $\square$

**Lemma 3.4.** *There exist  $C_1$  and  $C_2$  depending only on  $p$ ,  $V$  and  $\Omega$  such that*

$$(54) \quad \frac{1}{2} \mathfrak{E}_{\mathcal{D}}(\mathbf{u}) + C_1 \leq \int_{\Omega} \Gamma(\pi_{\mathcal{D}} \mathbf{u}) d\mathbf{x} \leq 2 \mathfrak{E}_{\mathcal{D}}(\mathbf{u}) + C_2, \quad \forall \mathbf{u} \in W_{\mathcal{D}}^{\text{en}}.$$

*In particular, the discrete entropy functional  $\mathfrak{E}_{\mathcal{D}}$  is bounded from below uniformly w.r.t. the discretization  $\mathcal{D}$ .*

*Proof.* Recall that the discrete entropy functional  $\mathfrak{E}_{\mathcal{D}}$  is defined by

$$\mathfrak{E}_{\mathcal{D}}(\mathbf{u}) = \int_{\Omega} (\Gamma(\pi_{\mathcal{D}} \mathbf{u}) + \pi_{\mathcal{D}} \mathbf{V} \pi_{\mathcal{D}} \mathbf{u}) d\mathbf{x}, \quad \forall \mathbf{u} \in W_{\mathcal{D}}^{\text{en}}.$$

whence

$$(55) \quad \int_{\Omega} \Gamma(\pi_{\mathcal{D}} \mathbf{u}) d\mathbf{x} \leq \mathfrak{E}_{\mathcal{D}}(\mathbf{u}) + \|\pi_{\mathcal{D}} \mathbf{V}\|_{L^{\infty}(\Omega)} \|\pi_{\mathcal{D}} \mathbf{u}\|_{L^1(\Omega)}, \quad \forall \mathbf{u} \in W_{\mathcal{D}}^{\text{en}}.$$

Let  $\epsilon > 0$  a parameter to be fixed later on. Thanks to Lemma 3.2, there exists a quantity  $C_{\epsilon}$  depending only on  $p$  and  $\epsilon$  such that

$$|u| \leq \epsilon \Gamma(u) + C_{\epsilon}, \quad \forall u \in \bar{\mathcal{I}}_p,$$

ensuring that

$$(56) \quad \|\pi_{\mathcal{D}} \mathbf{u}\|_{L^1(\Omega)} \leq \epsilon \Gamma(\pi_{\mathcal{D}} \mathbf{u}) + C_{\epsilon} \text{meas}(\Omega), \quad \forall \mathbf{u} \in W_{\mathcal{D}}^{\text{en}}.$$

On the other hand, Assumption (A4) together with the definition (36) of  $\mathbf{V} = (V_{\kappa}, V_s)_{\kappa, s}$  provide that

$$\|\pi_{\mathcal{D}} \mathbf{V}\|_{L^{\infty}(\Omega)} \leq \|\mathbf{V}\|_{\infty}.$$

Setting  $\epsilon = \frac{1}{2\|\mathbf{V}\|_{\infty}}$  in (56) and injecting the resulting estimate in (55) ends the proof of the second inequality of (54). The proof of the first inequality of (54) is similar, hence left to the reader.  $\square$

**Lemma 3.5.** *There exists  $C$  depending only on  $\mathbf{\Lambda}$ ,  $\Omega$ ,  $\theta_{\mathcal{T}}$ ,  $\zeta_{\mathcal{D}}$ ,  $\ell_{\mathcal{D}}$ ,  $\eta$ ,  $p$  and  $V$  such that, for all  $\mathbf{v} = (v_{\kappa}, v_s)_{\kappa, s} \in W_{\mathcal{D}}^{\text{ad}}$ , one has*

$$(57) \quad \sum_{\kappa \in \mathcal{M}} \sum_{s \in \mathcal{V}_{\kappa}} \left( \sum_{s' \in \mathcal{V}_{\kappa}} |a_{s, s'}^{\kappa}| \right) \eta_{\kappa, s}(\mathbf{v}) (p(v_{\kappa}) - p(v_s))^2 \leq C \left( 1 + \mathfrak{E}_{\mathcal{D}}(\mathbf{v}) + \sum_{\kappa \in \mathcal{M}} \delta_{\kappa} \mathbf{h}(\mathbf{v}) \cdot \mathbf{B}_{\kappa}(\mathbf{v}) \delta_{\kappa} \mathbf{h}(\mathbf{v}) \right),$$

where we have set  $\eta_{\kappa, s}(\mathbf{v}) = \frac{\eta(v_{\kappa}) + \eta(v_s)}{2}$  for all  $\kappa \in \mathcal{M}$  and all  $s \in \mathcal{V}_{\kappa}$ .

*Proof.* Let  $\mathbf{v} \in W_{\mathcal{D}}^{\text{ad}} \subset W_{\mathcal{D}}^{\text{en}}$ , then it follows from the definition (37) of  $\mathbf{h}(\mathbf{v}) = (\mathbf{h}_{\kappa}(v_{\kappa}), \mathbf{h}_s(v_s))_{\kappa, s} \in W_{\mathcal{D}}$  that

$$\delta_{\kappa} p(\mathbf{v}) \cdot \mathbf{B}_{\kappa}(\mathbf{v}) \delta_{\kappa} p(\mathbf{v}) \leq 2 \delta_{\kappa} \mathbf{h}(\mathbf{v}) \cdot \mathbf{B}_{\kappa}(\mathbf{v}) \delta_{\kappa} \mathbf{h}(\mathbf{v}) + 2 \delta_{\kappa} \mathbf{V} \cdot \mathbf{B}_{\kappa}(\mathbf{v}) \delta_{\kappa} \mathbf{V}, \quad \forall \kappa \in \mathcal{M}.$$

It follows from Lemma A.2 stated in appendix that there exists  $C$  depending only on  $\mathbf{\Lambda}$ ,  $\theta_{\mathcal{T}}$  and  $\ell_{\mathcal{D}}$  such that

$$\left( \sum_{s' \in \mathcal{V}_{\kappa}} |a_{s, s'}^{\kappa}| \right) \eta_{\kappa, s}(\mathbf{v}) (p(v_{\kappa}) - p(v_s))^2 \leq C \delta_{\kappa} p(\mathbf{v}) \cdot \mathbf{B}_{\kappa}(\mathbf{v}) \delta_{\kappa} p(\mathbf{v}), \quad \forall \kappa \in \mathcal{M}.$$

Therefore, it only remains to prove that

$$(58) \quad \sum_{\kappa \in \mathcal{M}} \delta_{\kappa} \mathbf{V} \cdot \mathbf{B}_{\kappa}(\mathbf{v}) \delta_{\kappa} \mathbf{V} \leq \mathfrak{E}_{\mathcal{D}}(\mathbf{v}) + C$$

for some  $C$  depending only on the prescribed data. Using Lemma A.3, we get that for all

$$(59) \quad \sum_{\kappa \in \mathcal{M}} \delta_\kappa \mathbf{V} \cdot \mathbf{B}_\kappa(\mathbf{v}) \delta_\kappa \mathbf{V} \leq \sum_{\kappa \in \mathcal{M}} \max_{s \in \mathcal{V}_\kappa} \eta_{\kappa,s}(\mathbf{v}) \sum_{s \in \mathcal{V}_\kappa} \left( \sum_{s' \in \mathcal{V}_\kappa} |a_{s,s'}^\kappa| \right) (V_\kappa - V_s)^2.$$

It results from Lemma A.2 that for all  $\kappa \in \mathcal{M}$ ,

$$(60) \quad \sum_{s \in \mathcal{V}_\kappa} \left( \sum_{s' \in \mathcal{V}_\kappa} |a_{s,s'}^\kappa| \right) (V_\kappa - V_s)^2 \leq C \int_\kappa \nabla_{\mathcal{T}} \mathbf{V} \cdot \mathbf{\Lambda} \nabla_{\mathcal{T}} \mathbf{V} d\mathbf{x} \leq C \lambda^* \|\nabla V\|_\infty^2,$$

whence, denoting by  $\bar{\eta}(\mathbf{v}) = (\bar{\eta}_\kappa(\mathbf{v}), \bar{\eta}_s(\mathbf{v}))_{\kappa,s} \in W_{\mathcal{D}}$  the vector defined by

$$\bar{\eta}_\kappa(\mathbf{v}) = \max \left( \eta(v_\kappa); \max_{s' \in \mathcal{V}_\kappa} \eta(v_{s'}) \right), \quad \bar{\eta}_s(\mathbf{v}) = 0, \quad \forall \kappa \in \mathcal{M}, \forall s \in \mathcal{V},$$

and remarking that

$$\max_{s \in \mathcal{V}_\kappa} \eta_{\kappa,s}(\mathbf{v}) \leq \bar{\eta}_\kappa(\mathbf{v}), \quad \forall \kappa \in \mathcal{M},$$

we deduce from using (60) in (59) that

$$\sum_{\kappa \in \mathcal{M}} \delta_\kappa \mathbf{V} \cdot \mathbf{B}_\kappa(\mathbf{v}) \delta_\kappa \mathbf{V} \leq C \int_\Omega \pi_{\mathcal{M}} \bar{\eta}(\mathbf{v}) d\mathbf{x},$$

for some  $C$  depending only on  $\theta_{\mathcal{T}}$ ,  $\mathbf{\Lambda}$ ,  $\ell_{\mathcal{D}}$  and  $\|\nabla V\|_\infty$ , the operator  $\pi_{\mathcal{M}}$  being defined by (30). Let us now use Lemma A.8 to obtain that

$$(61) \quad \sum_{\kappa \in \mathcal{M}} \delta_\kappa \mathbf{V} \cdot \mathbf{B}_\kappa(\mathbf{v}) \delta_\kappa \mathbf{V} \leq \tilde{C} \int_\Omega \pi_{\mathcal{D}} \eta(\mathbf{v}) d\mathbf{x}$$

for some  $\tilde{C}$  depending only on the prescribed data, namely  $\theta_{\mathcal{T}}$ ,  $\mathbf{\Lambda}$ ,  $\ell_{\mathcal{D}}$ ,  $\zeta_{\mathcal{D}}$  and  $\|\nabla V\|_\infty$ . Using Lemma 3.3, we know that for all  $\epsilon > 0$ , there exists  $C_\epsilon$  depending only on  $\epsilon$ ,  $\eta$ ,  $\Gamma$  and  $\text{meas}(\Omega)$  such that

$$\int_\Omega \pi_{\mathcal{D}} \eta(\mathbf{v}) d\mathbf{x} \leq \epsilon \int_\Omega \Gamma(\pi_{\mathcal{D}} \mathbf{v}) d\mathbf{x} + C_\epsilon.$$

Combining this result with Lemma 3.4 and (61), we deduce that for all  $\epsilon > 0$ , there exists  $C_\epsilon$  depending only on  $\epsilon$ ,  $\mathbf{\Lambda}$ ,  $\Omega$ ,  $\theta_{\mathcal{T}}$ ,  $\zeta_{\mathcal{D}}$ ,  $\ell_{\mathcal{D}}$ ,  $\eta$ ,  $p$  and  $V$  such that

$$\sum_{\kappa \in \mathcal{M}} \delta_\kappa \mathbf{V} \cdot \mathbf{B}_\kappa(\mathbf{v}) \delta_\kappa \mathbf{V} \leq \epsilon \tilde{C} \mathfrak{E}_{\mathcal{D}}(\mathbf{v}) + C_\epsilon, \quad \forall \mathbf{v} \in W_{\mathcal{D}}^{\text{en}}.$$

We obtain (58) by choosing  $\epsilon = \frac{1}{\tilde{C}}$ . This ends the proof of Lemma 3.5.  $\square$

**Lemma 3.6.** *Let  $\mathbf{u}^{n-1} \in W_{\mathcal{D}}^{\text{en}}$ , and let  $\mathbf{u}^n \in W_{\mathcal{D}}^{\text{ad}}$  be a solution to the scheme (41), then there exist  $C_1$  and  $C_2$  depending only on  $\mathbf{\Lambda}$ ,  $\Omega$ ,  $\theta_{\mathcal{T}}$ ,  $\zeta_{\mathcal{D}}$ ,  $\ell_{\mathcal{D}}$ ,  $\eta$ ,  $p$  and  $V$  such that*

$$\begin{aligned} \sum_{\kappa \in \mathcal{M}} \sum_{s \in \mathcal{V}_\kappa} \left( \sum_{s' \in \mathcal{V}_\kappa} |a_{s,s'}^\kappa| \right) \eta_{\kappa,s}^n (p(u_\kappa^n) - p(u_s^n))^2 \\ \leq C_1 (1 + \mathfrak{E}(\mathbf{u}^{n-1})) \leq C_2 \left( 1 + \int_\Omega \Gamma(\pi_{\mathcal{D}} \mathbf{u}^{n-1}) d\mathbf{x} \right). \end{aligned}$$



*Proof.* Since  $\mathbf{u}^n$  is a solution of the scheme (41), the nonlinear discrete stability estimate (51) holds. Therefore, taking (51) into account in (57) yields

$$\sum_{\kappa \in \mathcal{M}} \sum_{s \in \mathcal{V}_\kappa} \left( \sum_{s' \in \mathcal{V}_\kappa} |a_{s,s'}^\kappa| \right) \eta_{\kappa,s}^n (p(u_\kappa^n) - p(u_s^n))^2 \leq C_1 (1 + \mathfrak{E}(\mathbf{u}^{n-1}))$$

for some  $C_1$  depending on the prescribed data. Then it only remains to use Lemma 3.4 to conclude the proof of Lemma 3.6.  $\square$

**3.2. Existence of a discrete solution.** The scheme (41) can be rewritten in the form of a nonlinear system

$$\mathcal{F}(\mathbf{u}^n) = \mathbf{0}_{\mathbb{R}^{\#\mathcal{M} + \#\mathcal{V}}}.$$

In the case where  $p(0) = -\infty$ , the function  $\mathcal{F}$  is continuous on  $W_{\mathcal{D}}^{\text{ad}}$ , but not uniformly continuous. The existence proof for a discrete solution we propose relies on a topological degree argument (see e.g. [33, 21]), whence we need to restrict the set of the possible  $\mathbf{u}^n$  for recovering the uniform continuity by avoiding the singularity near 0. This is the purpose of the following lemma, which is an adaptation of [14, Lemma 3.10].

**Lemma 3.7.** *Let  $\mathbf{u}^{n-1} \in W_{\mathcal{D}}^{\text{en}}$  be such that  $\int_{\Omega} \pi_{\mathcal{D}} \mathbf{u}^{n-1} d\mathbf{x} > 0$  and let  $\mathbf{u}^n$  be a solution of the scheme (41). Assume that  $p(0) = -\infty$ , then there exists  $\epsilon_{\mathcal{D}, \Delta t_n} > 0$  depending on  $\Delta t_n$ ,  $\mathcal{D}$ ,  $\Lambda$ ,  $\Omega$ ,  $\eta$ ,  $p$ ,  $V$ , and  $\mathfrak{E}_{\mathcal{D}}(\mathbf{u}^{n-1})$  such that*

$$u_\nu^n \geq \epsilon_{\mathcal{D}, \Delta t_n}, \quad \forall \nu \in \mathcal{M} \cup \mathcal{V}.$$

*Proof.* First of all, remark that proving Lemma 3.7 is equivalent to proving that there exists  $C_{\mathcal{D}, \Delta t_n} > 0$  such that

$$(62) \quad p(u_\nu^n) \geq -C_{\mathcal{D}, \Delta t_n}, \quad \forall \nu \in \mathcal{M} \cup \mathcal{V}.$$

Because of the conservation of mass (42), we have

$$\int_{\Omega} \pi_{\mathcal{D}} \mathbf{u}^n d\mathbf{x} = \int_{\Omega} \pi_{\mathcal{D}} \mathbf{u}^{n-1} d\mathbf{x} > 0.$$

Therefore, we can claim that there exists  $\nu_i \in \mathcal{M} \cup \mathcal{V}$  such that

$$(63) \quad u_{\nu_i}^n \geq \frac{1}{\text{meas}(\Omega)} \int_{\Omega} \pi_{\mathcal{D}} \mathbf{u}^{n-1} d\mathbf{x} > 0.$$

Let  $(\nu_f) \in \mathcal{M} \cup \mathcal{V}$  be arbitrary, and let  $(\nu_q)_{q=0, \dots, \ell}$  be a path from  $\nu_i$  to  $\nu_f$ , i.e.

- $\nu_0 = \nu_i$ ,  $\nu_\ell = \nu_f$ , and  $\nu_p \neq \nu_q$  if  $p \neq q$ ;
- for all  $q \in \{0, \dots, \ell-1\}$ , one has:

$$\nu_q \in \mathcal{M} \implies \nu_{q+1} \in \mathcal{V}_{\nu_{q+1}}, \quad \text{and} \quad \nu_q \in \mathcal{V} \implies \nu_{q+1} \in \mathcal{M}_{\nu_{q+1}}.$$

Let  $q \in \{0, \dots, \ell-1\}$

It follows from Lemma 3.6 that there exists  $C_{\mathcal{D}, \Delta t_n}$  depending on  $\mathcal{D}$ ,  $\Delta t_n$ ,  $\Lambda$ ,  $\Omega$ ,  $\eta$ ,  $p$ ,  $V$ , and  $\mathfrak{E}_{\mathcal{D}}(\mathbf{u}^{n-1})$  such that

$$\sum_{\kappa \in \mathcal{M}} \sum_{s \in \mathcal{V}_\kappa} \eta_{\kappa,s}^n (p(u_\kappa^n) - p(u_s^n))^2 \leq C_{\mathcal{D}, \Delta t_n}.$$

This ensures in particular that

$$(64) \quad \sum_{q=0}^{\ell-1} \eta_{\nu_q, \nu_{q+1}}^n (p(u_{\nu_q}^n) - p(u_{\nu_{q+1}}^n))^2 \leq C_{\mathcal{D}, \Delta t_n},$$

where we have set  $\eta_{\nu_q, \nu_{q+1}}^n = \eta_{\kappa, s}^n$  if  $\{\nu_q, \nu_{q+1}\} = \{\kappa, s\}$ .

We can now prove (62) thanks to an induction along the path. Assume that  $u_{\nu_q}^n > \epsilon_{\mathcal{D}, \Delta t_n}$  for some  $\epsilon_{\mathcal{D}, \Delta t_n} > 0$ , whence  $\eta_{\nu_q, \nu_{q+1}}^n \geq \frac{\eta(u_{\nu_q}^n)}{2} \geq \epsilon'_{\mathcal{D}, \Delta t_n} > 0$ . Then it follows from (64) that

$$p(u_{\nu_{q+1}}^n) \geq p(u_{\nu_q}^n) - \sqrt{\frac{C_{\mathcal{D}, \Delta t_n}}{\epsilon'_{\mathcal{D}, \Delta t_n}}} \geq -C'_{\mathcal{D}, \Delta t_n} \implies u_{\nu_{q+1}}^n \geq \epsilon''_{\mathcal{D}, \Delta t_n} > 0.$$

We conclude as in [14, Lemma 3.10] thanks to the finite number of possible paths.  $\square$

Thanks to Lemma 3.7, one can apply the same strategy as in [14] for proving the existence of a solution to the scheme (41).

**Proposition 3.8.** *Let  $\mathbf{u}^{n-1} \in W_{\mathcal{D}}^{\text{en}}$ , then there exists (at least) one vector  $\mathbf{u}^n \in W_{\mathcal{D}}^{\text{ad}}$  solution to the system (41).*

The proof is similar to the one proposed in [14, Proposition 3.11], it is therefore omitted. Putting together the results of Lemma 3.1, Lemma 3.10, we get Theorem 2.3.

**3.3. Multistep *a priori* estimates.** As a by-product of the existence of a discrete solution  $\mathbf{u}^n$  for all  $n \in \{1, \dots, N\}$ , we can now derive *a priori* estimates on functions reconstructed thanks that the discrete solution  $\mathbf{u} \in W_{\mathcal{D}, \Delta t}$ .

The first estimate we get is obtained by summing Ineq. (51) w.r.t.  $n$ , and by using the positivity of the dissipation. This provides

$$(65) \quad \max_{n \in \{1, \dots, N\}} \mathfrak{E}_{\mathcal{D}}(\mathbf{u}^n) \leq \mathfrak{E}_{\mathcal{D}}(\mathbf{u}^0) \leq \mathfrak{E}(u_0) + C \leq C,$$

and where  $C$  only depending on  $V$ ,  $u_0$ ,  $p$  and  $\Omega$  thanks to Lemma 2.1. Since the discrete entropy functional  $\mathfrak{E}_{\mathcal{D}}$  is bounded from below by a quantity depending only on  $p$ ,  $V$  and  $\Omega$  (cf. Lemma 3.4), we deduce also from the summation of (51) w.r.t.  $n$  that there exists  $C$  depending only on  $p$ ,  $V$ ,  $\Omega$ , and  $u_0$  (but not on  $\mathcal{D}$ ) such that

$$(66) \quad \sum_{n=1}^N \Delta t_n \sum_{\kappa \in \mathcal{M}} \delta_{\kappa} \mathbf{h}(\mathbf{u}^n) \cdot \mathbf{B}_{\kappa}(\mathbf{u}^n) \delta_{\kappa} \mathbf{h}(\mathbf{u}^n) \leq C.$$

Mimicking the proof of Lemma 3.5, this yields

$$(67) \quad \sum_{n=1}^N \Delta t_n \sum_{\kappa \in \mathcal{M}} \delta_{\kappa} p(\mathbf{u}^n) \cdot \mathbf{B}_{\kappa}(\mathbf{u}^n) \delta_{\kappa} p(\mathbf{u}^n) \leq C$$

for some quantity  $C$  depending on  $\mathbf{\Lambda}$ ,  $\Omega$ ,  $\theta_{\mathcal{T}}$ ,  $\zeta_{\mathcal{D}}$ ,  $\ell_{\mathcal{D}}$ ,  $\eta$ ,  $p$ ,  $V$  and  $t_f$ .

The following lemma is a direct consequence of Estimate (65) and Lemma 3.4. Its detailed proof is left to the reader.

**Lemma 3.9.** *There exists  $C$  depending only on  $\Omega$ ,  $p$ ,  $V$ ,  $u_0$  and  $\Omega$  (but not on the discretization) such that*

$$\|\Gamma(\pi_{\mathcal{D}, \Delta t} \mathbf{u})\|_{L^{\infty}((0, T); L^1(\Omega))} \leq C.$$

**Lemma 3.10.** *There exists  $C > 0$  depending only on  $\mathbf{\Lambda}$ ,  $\eta$ ,  $V$ ,  $\theta_{\mathcal{T}}$  and  $\ell_{\mathcal{D}}$ ,  $\Omega$ ,  $t_f$ ,  $u_0$  such that*

$$\iint_{Q_{t_f}} \mathbf{\Lambda} \nabla_{\mathcal{T}, \Delta t} \xi(\mathbf{u}) \cdot \nabla_{\mathcal{T}, \Delta t} \xi(\mathbf{u}) d\mathbf{x} \leq C,$$

where the function  $\xi$  is defined by (7).

*Proof.* Since the function  $\eta$  was supposed to be nondecreasing, cf. Assumption (A1), we know that for all interval  $[a, b] \subset \mathcal{I}_p$ ,  $\max_{c \in [a, b]} \eta(c) = \max\{\eta(a), \eta(b)\}$ , hence, denoting by  $\mathcal{I}_{\kappa, s}^n$  the interval with extremities  $u_{\kappa}^n$  and  $u_s^n$ , we obtain that

$$(68) \quad \eta_{\kappa, s}^n \geq \frac{1}{2} \max_{c \in \mathcal{I}_{\kappa, s}^n} \eta(c), \quad \forall \kappa \in \mathcal{M}, \forall s \in \mathcal{V}_{\kappa}.$$

The definition (7) of the function  $\xi$  implies that

$$(\xi(u_s^n) - \xi(u_{\kappa}^n))^2 \leq \left( \max_{c \in \mathcal{I}_{\kappa, s}^n} \eta(c) \right) (p(u_s^n) - p(u_{\kappa}^n))^2,$$

whence we obtain that for all  $\kappa \in \mathcal{M}$ ,

$$\begin{aligned} |\mathbf{M}_{\kappa}(\mathbf{u}) \delta_{\kappa} p(\mathbf{u}^n)|^2 &= \sum_{s \in \mathcal{V}_{\kappa}} \eta_{\kappa, s}^n (p(u_s^n) - p(u_{\kappa}^n))^2 \\ &\geq \frac{1}{2} \sum_{s \in \mathcal{V}_{\kappa}} (\xi(u_s^n) - \xi(u_{\kappa}^n))^2 = \frac{1}{2} |\delta_{\kappa} \xi(\mathbf{u}^n)|^2. \end{aligned}$$

Using that

$$\mathbf{v} \cdot \mathbf{A}_{\kappa} \mathbf{v} \geq \mathbf{w} \cdot \mathbf{A}_{\kappa} \mathbf{w}, \quad \forall \mathbf{v}, \mathbf{w} \in \mathbb{R}^{\ell_{\kappa}} \text{ s.t. } |\mathbf{v}|^2 \geq \text{Cond}_2(\mathbf{A}_{\kappa}) |\mathbf{w}|^2,$$

and that  $\mathbf{B}_{\kappa}(\mathbf{u}) = \mathbf{M}_{\kappa}(\mathbf{u}) \mathbf{A}_{\kappa} \mathbf{M}_{\kappa}(\mathbf{u})$ , we get that for all  $\kappa \in \mathcal{M}$ ,

$$\begin{aligned} \delta_{\kappa} p(\mathbf{u}^n) \cdot \mathbf{B}_{\kappa}(\mathbf{u}^n) \delta_{\kappa} p(\mathbf{u}^n) &\geq \frac{1}{2 \text{Cond}_2(\mathbf{A}_{\kappa})} \delta_{\kappa} \xi(\mathbf{u}^n) \cdot \mathbf{A}_{\kappa} \delta_{\kappa} \xi(\mathbf{u}^n) \\ &= \frac{1}{2 \text{Cond}_2(\mathbf{A}_{\kappa})} \int_{\kappa} \mathbf{\Lambda} \nabla_{\mathcal{T}} \xi(\mathbf{u}^n) \cdot \nabla_{\mathcal{T}} \xi(\mathbf{u}^n) d\mathbf{x}. \end{aligned}$$

Thanks to Lemma A.1 stated in appendix, we know that  $C > 0$  depending only on  $\mathbf{\Lambda}$ ,  $\theta_{\mathcal{T}}$  and  $\ell_{\mathcal{D}}$  such that  $\text{Cond}_2(\mathbf{A}_{\kappa}) \leq C$ , for all  $\kappa \in \mathcal{M}$ , so that:  $\forall \kappa \in \mathcal{M}$ ,  $\forall n \in \{1, \dots, N\}$ ,

$$(69) \quad \int_{\kappa} \nabla_{\mathcal{T}} \xi(\mathbf{u}^n) \cdot \mathbf{\Lambda} \nabla_{\mathcal{T}} \xi(\mathbf{u}^n) d\mathbf{x} \leq C \delta_{\kappa} p(\mathbf{u}^n) \cdot \mathbf{B}_{\kappa}(\mathbf{u}^n) \delta_{\kappa} p(\mathbf{u}^n).$$

In order to conclude the proof, it only remains to multiply (69) by  $\Delta t_n$  and to sum over  $\kappa \in \mathcal{M}$  and  $n \in \{1, \dots, N\}$  and to use (67).  $\square$

Combining Estimate (67) and Lemma A.2 yields the following lemma, whose complete proof is left to the reader.

**Lemma 3.11.** *There exists  $C$  depending only on  $\mathbf{\Lambda}$ ,  $\Omega$ ,  $\theta_{\mathcal{T}}$ ,  $\zeta_{\mathcal{D}}$ ,  $\ell_{\mathcal{D}}$ ,  $\eta$ ,  $p$ ,  $V$  and  $t_f$  such that*

$$\sum_{n=1}^N \Delta t_n \sum_{\kappa \in \mathcal{M}} \sum_{s \in \mathcal{V}_{\kappa}} \left( \sum_{s' \in \mathcal{V}_{\kappa}} |a_{s, s'}^{\kappa}| \right) \eta_{\kappa, s}^n (p(u_{\kappa}^n) - p(u_s^n))^2 \leq C.$$

## 4. PROOF OF THEOREM 2.4

In what follows, we consider a sequence  $(\mathcal{D}_m, \Delta t_m)_{m \geq 1}$  of discretizations of  $Q_{t_f}$  such that (52) holds. In order to prove the convergence of the reconstructed discrete solution  $\pi_{\mathcal{D}_m, \Delta t_m} \mathbf{u}_m$  towards the weak solution of (1) as  $m$  tends to  $\infty$ , we adopt the classical strategy that consists in showing firstly that the family  $(\pi_{\mathcal{D}_m, \Delta t_m} \mathbf{u}_m)_{m \geq 1}$  is precompact in  $L^1(Q_{t_f})$  (this is the purpose of §4.1), then to identify in §4.2 the limit as a weak solution of (1) in the sense of Definition 1.

As a direct consequence of Theorem 2.3, one knows that the scheme admits a solution  $\mathbf{u}_m = (u_{\kappa, m}^n, u_{s, m}^n)$  that, thanks to the regularity assumptions (52b) – (52c) on the discretization and thanks to Lemmas 3.9, 3.10 and 3.11, satisfies the following uniform estimates w.r.t.  $m$ :

$$(70) \quad \|\pi_{\mathcal{D}_m, \Delta t_m} \Gamma(\mathbf{u}_m)\|_{L^\infty((0, t_f); L^1(\Omega))} \leq C,$$

$$(71) \quad \iint_{Q_{t_f}} \nabla_{\mathcal{T}_m, \Delta t_m} \xi(\mathbf{u}_m) \cdot \Lambda \nabla_{\mathcal{T}_m, \Delta t_m} \xi(\mathbf{u}_m) dx dt \leq C,$$

$$(72) \quad \sum_{n=1}^{N_m} \Delta t_{n, m} \sum_{\kappa \in \mathcal{M}_m} \sum_{s \in \mathcal{V}_\kappa} \left( \sum_{s' \in \mathcal{V}_\kappa} |a_{s, s'}^\kappa| \right) \eta_{\kappa, s}^n (p(u_{\kappa, m}^n) - p(u_{s, m}^n))^2 \leq C,$$

where  $C$  may depend on the data of the continuous problem, and on the discretization regularity factors  $\theta^*$ ,  $\ell^*$  and  $\zeta^*$  but not on  $m$ .

## 4.1. Compactness properties of the discrete solutions.

**Lemma 4.1.** *Let  $(\mathcal{D}_m, \Delta t_m)$  be a sequence of discretizations of  $Q_{t_f}$  satisfying Assumptions (52), there exists  $C$  depending only on  $\Lambda$ ,  $\theta^*$ ,  $\ell^*$ ,  $\Omega$ ,  $t_f$ ,  $p$  and  $u_0$  such that, for all  $m \geq 1$ , one has*

$$\|\pi_{\mathcal{T}_m, \Delta t_m} \xi(\mathbf{u}_m)\|_{L^2((0, T); H^1(\Omega))} \leq C \quad \text{and} \quad \|\pi_{\mathcal{D}_m, \Delta t_m} \xi(\mathbf{u}_m)\|_{L^2(Q_{t_f})} \leq C.$$

*Proof.* It follows from the Estimate (71) and from Lemma A.5 stated in appendix that for all  $m \geq 1$ ,

$$(73) \quad \begin{aligned} \|\pi_{\mathcal{D}_m} \xi(\mathbf{u}_m) - \pi_{\mathcal{T}_m} \xi(\mathbf{u}_m)\|_{L^2((0, T); L^1(\Omega))} \\ \leq \text{meas}(\Omega)^{1/2} \|\pi_{\mathcal{D}_m} \xi(\mathbf{u}_m) - \pi_{\mathcal{T}_m} \xi(\mathbf{u}_m)\|_{L^2(Q_{t_f})} \leq C \end{aligned}$$

for some  $C$  depending only on  $\Lambda$ ,  $\Omega$ ,  $t_f$ ,  $\ell^*$  and  $\theta^*$ ,  $p$  and  $u_0$  (but not on  $m$ ).

Moreover, it follows from Assumption (8) that

$$\|\pi_{\mathcal{D}_m, \Delta t_m} \xi(\mathbf{u}_m)\|_{L^\infty((0, T); L^1(\Omega))} \leq C \left( 1 + \|\pi_{\mathcal{D}_m, \Delta t_m} \Gamma(\mathbf{u}_m)\|_{L^\infty((0, T); L^1(\Omega))} \right) \leq C$$

thanks to the estimate (70). Combining this inequality with (73) provides that

$$\|\pi_{\mathcal{T}_m, \Delta t_m} \xi(\mathbf{u}_m)\|_{L^1(Q_{t_f})} \leq C,$$

whence the sequence  $(\pi_{\mathcal{T}_m, \Delta t_m} \xi(\mathbf{u}_m))_{m \geq 1}$  is bounded in  $L^2((0, T); W^{1,1}(\Omega))$  thanks to (71). A classical bootstrap argument using Sobolev inequalities allows to claim that it is bounded in  $L^2((0, T); H^1(\Omega))$ , thus in particular in  $L^2(Q_{t_f})$ . One concludes that  $(\pi_{\mathcal{T}_m, \Delta t_m} \xi(\mathbf{u}_m))_{m \geq 1}$  is also bounded in  $L^2(Q_{t_f})$  thanks to (73).  $\square$

**Remark 4.2.** *An alternative way to prove the key-point of Lemma 4.1, namely*

$$\|\pi_{\mathcal{T}_m, \Delta t_m} \xi(\mathbf{u}_m)\|_{L^2((0,T); H^1(\Omega))} \leq C,$$

*would consist in using [30, Lemma A.1], that claims that*

$$v \in L^1(\Omega) \text{ and } \nabla \xi(v) \in L^2(\Omega)^d \implies \xi(v) \in H^1(\Omega).$$

As a consequence of Lemma 4.1, we know that the sequence  $(\pi_{\mathcal{T}_m, \Delta t_m} \xi(\mathbf{u}_m))_{m \geq 1}$  is relatively compact for the  $L^2((0, t_f); H^1(\Omega))$ -weak topology. Moreover, the space  $H^1(\Omega)$  being locally compact in  $L^2(\Omega)$ , a uniform information on the time translates of  $\pi_{\mathcal{T}_m, \Delta t_m} \xi(\mathbf{u}_m)$  will provide the relative compactness of  $(\pi_{\mathcal{T}_m, \Delta t_m} \xi(\mathbf{u}_m))_{m \geq 1}$  in the  $L^2(Q_{t_f})$ -strong topology (see e.g. [39]). Such a uniform time-translate estimate can be obtained by using directly the numerical scheme (see e.g. [24, 14]). One can also make use of black-boxes like e.g. [4, 6]. Note that the result of [28] does not apply here because of the degeneracy of the problem. We do not provide the proof of next proposition here, since a suitable black-box will be contained in the forthcoming contribution [6]. A more calculative possibility would consist in mimicking the proof of [14, Lemmas 4.3 and 4.5].

**Proposition 4.3.** *Let  $(\mathcal{D}_m, \Delta t_m)$  be a sequence of discretizations of  $Q_{t_f}$  satisfying Assumptions (52), and let  $(\mathbf{u}_m)_{m \geq 1}$  be the corresponding sequence of solutions to the scheme (41), then there exists a measurable function  $u : Q_{t_f} \rightarrow \mathbb{R}$  with  $\xi(u) \in L^2((0, T); H^1(\Omega))$  such that, up to a subsequence, one has*

$$(74) \quad \pi_{\mathcal{D}_m, \Delta t_m} \mathbf{u}_m \xrightarrow{m \rightarrow \infty} u \quad \text{a.e. in } Q_{t_f},$$

**Corollary 4.4.** *Keeping the assumption and notations of Proposition 4.3, one has*

$$\pi_{\mathcal{D}_m, \Delta t_m} \mathbf{u}_m \xrightarrow{m \rightarrow \infty} u \quad \text{strongly in } L^1(Q_{t_f}).$$

*Proof.* As a result of Proposition 4.3, the almost everywhere convergence property (74) holds. On the other hand, it follows from Assumption (A2), more precisely from the fact that  $\lim_{u \rightarrow \infty} p(u) = +\infty$  that the function  $\Gamma$  defined by (4) is superlinear, i.e.,

$$\lim_{u \rightarrow +\infty} \frac{\Gamma(u)}{u} = +\infty.$$

Therefore, Estimate (70) implies that  $(\pi_{\mathcal{D}_m, \Delta t_m} \mathbf{u}_m)_{m \geq 1}$  is uniformly equi-integrable. Hence we can apply Vitali's convergence theorem to conclude the proof of Corollary 4.4.  $\square$

**Lemma 4.5.** *Let  $u$  be the limit value of  $\pi_{\mathcal{D}_m, \Delta t_m} \mathbf{u}_m$  exhibited in Proposition 4.3, then*

$$(75) \quad \pi_{\mathcal{D}_m, \Delta t_m} \eta(\mathbf{u}_m) \xrightarrow{m \rightarrow \infty} \eta(u) \quad \text{strongly in } L^1(Q_{t_f}).$$

and

$$(76) \quad \pi_{\mathcal{M}_m, \Delta t_m} \eta(\mathbf{u}_m) \xrightarrow{m \rightarrow \infty} \eta(u) \quad \text{strongly in } L^1(Q_{t_f}).$$

*Proof.* Let us first establish (75). Thanks to the entropy estimate (70), we know that the sequence  $(\pi_{\mathcal{D}_m, \Delta t_m} \Gamma(\mathbf{u}_m))_{m \geq 1}$  is uniformly bounded in  $L^\infty((0, t_f); L^1(\Omega))$ ,

thus in  $L^1(Q_{t_f})$ . Then Assumption (9) allows to use the de la Vallée-Poussin theorem to claim that  $(\pi_{\mathcal{D}_m, \Delta t_m} \eta(\mathbf{u}_m))_{m \geq 1}$  is uniformly equi-integrable on  $Q_{t_f}$ . Moreover, the continuity of  $\eta$  and Proposition 4.3 provide that

$$(\pi_{\mathcal{D}_m, \Delta t_m} \eta(\mathbf{u}_m))_{m \geq 1} \xrightarrow{m \rightarrow \infty} \eta(u) \quad \text{a.e. in } Q_{t_f}.$$

Therefore, we obtain (75) by applying Vitali's theorem.  $\square$

**Lemma 4.6.** *Under the assumptions of Proposition 4.3, one has*

$$\pi_{\mathcal{T}_m, \Delta t_m} \xi(\mathbf{u}_m) \xrightarrow{m \rightarrow \infty} \xi(u) \quad \text{weakly in } L^2((0, T); H^1(\Omega)),$$

where  $u$  is the solution exhibited in Proposition 4.3

*Proof.* Thanks to Lemma 4.1, the sequence  $(\nabla_{\mathcal{T}_m, \Delta t_m} \xi(\mathbf{u}_m))_{m \geq 1}$  is uniformly bounded in  $L^2(Q_{t_f})^d$ . Therefore, there exists  $\Xi \in L^2((0, T); H^1(\Omega))$  such that

$$\pi_{\mathcal{T}_m, \Delta t_m} \xi(\mathbf{u}_m) \xrightarrow{m \rightarrow \infty} \Xi \quad \text{weakly in } L^2((0, T); H^1(\Omega)).$$

But in view of Proposition 4.3 and of the continuity of  $\xi$ , we know that

$$\xi(\pi_{\mathcal{D}_m, \Delta t_m} \mathbf{u}_m) = \pi_{\mathcal{D}_m, \Delta t_m} \xi(\mathbf{u}_m) \xrightarrow{m \rightarrow \infty} \xi(u) \quad \text{a.e. in } Q_{t_f}.$$

Since  $\pi_{\mathcal{T}_m, \Delta t_m} \xi(\mathbf{u}_m)$  and  $\pi_{\mathcal{D}_m, \Delta t_m} \xi(\mathbf{u}_m)$  have the same limit (cf. Lemma A.5), we get that  $\Xi = \xi(u)$ .  $\square$

#### 4.2. Identification of the limit as a weak solution.

**Proposition 4.7.** *Let  $u$  be a limit value of the sequence  $(\pi_{\mathcal{D}_m, \Delta t_m} \mathbf{u}_m)_{m \geq 1}$  exhibited in Proposition 4.3, then  $u$  is the unique weak solution to the problem (1) in the sense of Definition 1.*

*Proof.* In order to check that  $u$  is a weak solution, it only remains to check that the weak formulation (11) holds. Let  $\psi \in C_c^\infty(\bar{\Omega} \times [0, T])$ , then, for all  $m \geq 1$ , for all  $\beta \in \mathcal{M}_m \cup \mathcal{V}_m$  and all  $n \in \{0, \dots, N_m\}$ , we denote by  $\Delta t_m = (\Delta t_{1,m}, \dots, \Delta t_{N_m,m})$ , by  $t_{n,m} = \sum_{i=1}^n \Delta t_{i,m}$ , by  $\psi_\beta^n = \psi(\mathbf{x}_\beta, t_{n,m})$ , by  $\psi_m^n = \left( \psi_\beta^n \right)_{\beta \in \mathcal{M}_m \cup \mathcal{V}_m} \in W_{\mathcal{D}_m}$ , and by  $\psi_m = (\psi_m^n)_{0 \leq n \leq N_m} \in W_{\mathcal{D}_m, \Delta t_m}$ . Note that since  $\psi(\cdot, t_f) = 0$ , one has  $\psi_m^{N_m} = \mathbf{0}$  for all  $m \geq 1$ .

Setting  $\mathbf{v} = \psi_m^{n-1}$  in (43) and summing over  $n \in \{1, \dots, N_m\}$  leads, after a classical reorganization of the sums [24], to

$$(77) \quad A_m + B_m + C_m + D_m = 0,$$

where we have set

$$\begin{aligned} A_m &= \sum_{n=1}^{N_m} \Delta t_{n,m} \int_{\Omega} \pi_{\mathcal{D}_m} \mathbf{u}_m^n(\mathbf{x}) \pi_{\mathcal{D}_m} \left( \frac{\psi_m^{n-1} - \psi_m^n}{\Delta t_{n,m}} \right) (\mathbf{x}) d\mathbf{x}, \\ B_m &= - \int_{\Omega} \pi_{\mathcal{D}_m} \mathbf{u}_m^0(\mathbf{x}) \pi_{\mathcal{D}_m} \psi_m^0(\mathbf{x}) d\mathbf{x}, \\ C_m &= \sum_{n=1}^{N_m} \Delta t_{n,m} \sum_{\kappa \in \mathcal{M}_m} \delta_\kappa p(\mathbf{u}_m^n) \cdot \mathbf{B}_\kappa(\mathbf{u}_m^n) \delta_\kappa \psi_m^{n-1}, \\ D_m &= \sum_{n=1}^{N_m} \Delta t_{n,m} \sum_{\kappa \in \mathcal{M}_m} \delta_\kappa \mathbf{V}_m \cdot \mathbf{B}_\kappa(\mathbf{u}_m^n) \delta_\kappa \psi_m^{n-1}, \end{aligned}$$

where  $\mathbf{V}_m = (V(\mathbf{x}_\kappa), V(\mathbf{x}_s))_{\kappa \in \mathcal{M}_m, s \in \mathcal{V}_m}$ .

Since for all  $m \geq 1$  and all  $n \in \{0, \dots, N_m\}$ , one has

$$\int_{\Omega} \pi_{\mathcal{D}_m} \Gamma(\mathbf{u}_m^n) d\mathbf{x} = \int_{\Omega} \Gamma(\pi_{\mathcal{D}_m} \mathbf{u}_m^n) d\mathbf{x} \leq \int_{\Omega} \Gamma(u_0(\mathbf{x})) d\mathbf{x} < +\infty,$$

the sequence  $(\pi_{\mathcal{D}_m, \Delta t_m} \mathbf{u}_m)_{m \geq 1}$  is uniformly equi-integrable on  $Q_{t_f}$ . Therefore, Proposition 4.3 together with Vitali's theorem yield that

$$\pi_{\mathcal{D}_m, \Delta t_m} \mathbf{u}_m \xrightarrow{m \rightarrow \infty} u \quad \text{strongly in } L^1(Q_{t_f}).$$

The regularity of  $\psi$  yields

$$\sum_{n=1}^{N_m} \pi_{\mathcal{D}_m} \left( \frac{\psi^{n-1} - \psi^n}{\Delta t_{n,m}} \right) \mathbf{1}_{[t_{n-1,m}, t_{n,m})} \xrightarrow{m \rightarrow \infty} -\partial_t \psi \quad \text{uniformly on } Q_{t_f}$$

where  $t_{n,m} = \sum_{i=1}^n \Delta t_{i,m}$ , so that

$$(78) \quad \lim_{m \rightarrow \infty} A_m = - \iint_{Q_T} u \partial_t \psi \, d\mathbf{x} dt.$$

The function  $\pi_{\mathcal{D}_m} \mathbf{u}_m^0(\mathbf{x})$  tends strongly in  $L^1(\Omega)$  towards  $u_0$  and  $\pi_{\mathcal{D}_m} \psi^0$  converges uniformly towards  $\psi(\cdot, 0)$  as  $m$  tends to  $+\infty$ , leading to

$$(79) \quad \lim_{m \rightarrow \infty} B_m = - \iint_{\Omega} u_0(\mathbf{x}) \psi(\mathbf{x}, 0) \, d\mathbf{x}.$$

We split the term  $C_m$  into three parts

$$(80) \quad C_m = C_{1,m} + C_{2,m} + C_{3,m}, \quad m \geq 1,$$

where, setting  $\hat{\psi}_m^0 = \psi_m^0$ ,  $\hat{\psi}_m^n = \psi_m^{n-1} \in W_{\mathcal{D}_m}$  for all  $n \in \{1, \dots, N_m\}$ , and  $\hat{\psi}_m = (\hat{\psi}_m^n)_{0 \leq n \leq N_m} \in W_{\mathcal{D}_m, \Delta t_m}$ , one has

$$\begin{aligned} C_{1,m} &= \iint_{Q_{t_f}} \pi_{\mathcal{M}_m, \Delta t_m} \sqrt{\eta(\mathbf{u}_m)} \nabla_{\mathcal{T}_m, \Delta t_m} \xi(\mathbf{u}_m) \cdot \mathbf{\Lambda} \nabla_{\mathcal{T}_m, \Delta t_m} \hat{\psi}_m \, d\mathbf{x} dt, \\ C_{2,m} &= \sum_{n=1}^{N_m} \Delta t_{n,m} \sum_{\kappa \in \mathcal{M}_m} \sum_{s \in \mathcal{V}_\kappa} \sum_{s' \in \mathcal{V}_\kappa} (\xi(u_\kappa^n) - \xi(u_s^n)) a_{s,s'}^\kappa \\ &\quad \times \left( \sqrt{\eta_{\kappa,s'}^n} - \sqrt{\eta(u_\kappa^n)} \right) (\psi_\kappa^{n-1} - \psi_{s'}^{n-1}), \\ C_{3,m} &= \sum_{n=1}^{N_m} \Delta t_{n,m} \sum_{\kappa \in \mathcal{M}_m} \sqrt{\eta(u_\kappa^n)} \sum_{s \in \mathcal{V}_\kappa} \left( \sqrt{\eta_{\kappa,s}^n} (p(u_\kappa^n) - p(u_s^n)) - (\xi(u_\kappa^n) - \xi(u_s^n)) \right) \\ &\quad \times \sum_{s' \in \mathcal{V}_\kappa} a_{s,s'}^\kappa (\psi_\kappa^{n-1} - \psi_{s'}^{n-1}). \end{aligned}$$

Thanks to Lemma 4.5, we know that

$$\pi_{\mathcal{M}_m, \Delta t_m} \sqrt{\eta(\mathbf{u}_m)} \xrightarrow{m \rightarrow \infty} \sqrt{\eta(u)} \quad \text{strongly in } L^2(Q_{t_f}).$$

Hence, it follows from the weak convergence in  $L^2(Q_{t_f})$  of  $\nabla_{\mathcal{T}_m, \Delta t_m} \xi(\mathbf{u}_m)$  towards  $\nabla \xi(u)$  (cf. Lemma 4.6) and from the uniform convergence of  $\nabla_{\mathcal{T}_m, \Delta t_m} \hat{\psi}_m$  towards

$\nabla\psi$  as  $m$  tends to  $+\infty$  (see for instance [20, Theorem 16.1]) that

$$(81) \quad \lim_{m \rightarrow \infty} C_{1,m} = \iint_{Q_{t_f}} \sqrt{\eta(u)} \nabla \xi(u) \cdot \Lambda \nabla \psi \, d\mathbf{x} dt.$$

Let us focus now of  $C_{2,m}$ . Using the inequality  $ab \leq \epsilon a^2 + \frac{1}{4\epsilon} b^2$ , one gets that

$$(82) \quad C_{2,m} \leq \epsilon C'_{2,m} + \frac{C''_{2,m}}{4\epsilon}, \quad \forall \epsilon > 0,$$

where

$$\begin{aligned} C'_{2,m} &= \sum_{n=1}^{N_m} \Delta t_{n,m} \sum_{\kappa \in \mathcal{M}_m} \sum_{s \in \mathcal{V}_\kappa} \left( \sum_{s' \in \mathcal{V}_\kappa} |a_{s,s'}^\kappa| \right) (\xi(u_\kappa^n) - \xi(u_s^n))^2, \\ C''_{2,m} &= \sum_{n=1}^{N_m} \Delta t_{n,m} \sum_{\kappa \in \mathcal{M}_m} \sum_{s \in \mathcal{V}_\kappa} \left( \sum_{s' \in \mathcal{V}_\kappa} |a_{s,s'}^\kappa| \right) \left( \sqrt{\eta_{\kappa,s}^n} - \sqrt{\eta(u_\kappa^n)} \right)^2 (\psi_\kappa^{n-1} - \psi_s^{n-1})^2. \end{aligned}$$

Using Lemma A.2 stated in appendix, one gets that there exists  $C$  depending only on  $\Lambda$ ,  $\theta^*$  and  $\ell^*$  such that

$$C'_{2,m} \leq C \iint_{Q_{t_f}} \nabla_{\mathcal{T}_m, \Delta t_m} \xi(\mathbf{u}_m) \cdot \Lambda \nabla_{\mathcal{T}_m, \Delta t_m} \xi(\mathbf{u}_m) d\mathbf{x} dt, \quad \forall m \geq 1.$$

Hence, in view of Estimate (71), one gets that

$$(83) \quad C'_{2,m} \leq C, \quad \forall m \geq 1,$$

for some  $C$  depending only on  $u_0$ ,  $p$ ,  $\Omega$ ,  $\Lambda$ ,  $\theta^*$  and  $\ell^*$ .

Define  $\boldsymbol{\mu}_m = (\mu_\kappa^n, \mu_s^n)_{\kappa, s, n} \in W_{\mathcal{D}_m, \Delta t_m}$  by

$$(84) \quad \begin{cases} \mu_s^n = 0, & \forall s \in \mathcal{V}_m, \\ \mu_\kappa^n = \max_{s \in \mathcal{V}_\kappa} \left( \sqrt{\eta_{\kappa,s}^n} - \sqrt{\eta(u_\kappa^n)} \right)^2, & \forall \kappa \in \mathcal{M}_m, \end{cases} \quad \forall n \in \{0, \dots, N_m\},$$

then thanks to Lemma A.2 again, there exists  $C$  depending only on  $\Lambda$ ,  $\theta^*$  and  $\ell^*$  such that

$$C''_{2,m} \leq C \iint_{Q_{t_f}} \pi_{\mathcal{M}_m, \Delta t_m} \boldsymbol{\mu}_m \nabla_{\mathcal{T}_m, \Delta t_m} \hat{\psi}_m \cdot \Lambda \nabla_{\mathcal{T}_m, \Delta t_m} \hat{\psi}_m d\mathbf{x} dt, \quad \forall m \geq 1.$$

Since  $\pi_{\mathcal{M}_m, \Delta t_m} \boldsymbol{\mu}_m$  converges to 0 strongly in  $L^1(Q_{t_f})$  as  $m$  tends to  $\infty$  (this is the purpose of Lemma 4.8 hereafter), and since  $\nabla_{\mathcal{T}_m, \Delta t_m} \hat{\psi}_m$  remains bounded in  $L^\infty(Q_{t_f})$  uniformly w.r.t.  $m \geq 1$ , one gets that

$$(85) \quad \lim_{m \rightarrow \infty} C''_{2,m} = 0.$$

Therefore, it follows from (82)–(85) that

$$\limsup_{m \rightarrow \infty} C_{2,m} \leq C\epsilon, \quad \forall \epsilon > 0,$$

whence

$$(86) \quad \lim_{m \rightarrow \infty} C_{2,m} = 0.$$



As a preliminary before considering  $C_{3,m}$ , let us set, for all  $\kappa \in \mathcal{M}_m$ , all  $s \in \mathcal{V}_\kappa$ , and all  $n \in \{1, \dots, N_m\}$ ,

$$\tilde{\eta}_{\kappa,s}^n = \begin{cases} \left( \frac{\xi(u_\kappa^n) - \xi(u_s^n)}{p(u_\kappa^n) - p(u_s^n)} \right)^2 & \text{if } u_\kappa^n \neq u_s^n, \\ \eta(u_\kappa^n) & \text{if } u_\kappa^n = u_s^n. \end{cases}$$

Thanks to the mean value theorem, we can claim that, for all  $\kappa \in \mathcal{M}_m$ , all  $s \in \mathcal{V}_\kappa$  and all  $n \in \{1, \dots, N_m\}$ , there exists  $\tilde{u}_{\kappa,s}^n \in \mathcal{I}_{\kappa,s}^n = [\min(u_\kappa^n, u_s^n), \max(u_\kappa^n, u_s^n)]$  such that  $\tilde{\eta}_{\kappa,s}^n = \eta(\tilde{u}_{\kappa,s}^n)$ . Then it follows from (68) that

$$(87) \quad 0 \leq \tilde{\eta}_{\kappa,s}^n \leq 2\eta_{\kappa,s}^n, \quad \forall \kappa \in \mathcal{M}_m, \forall s \in \mathcal{V}_\kappa, \forall n \in \{1, \dots, N_m\}.$$

Introducing  $\nu_m = (\nu_\kappa^n, \nu_s^n)_{\kappa,s,n} \in W_{\mathcal{D}_m, \Delta t_m}$ , defined by

$$(88) \quad \nu_s^n = 0 \text{ and } \nu_\kappa^n = \max_{s' \in \mathcal{V}_\kappa} \left( \frac{\sqrt{\eta_{\kappa,s'}^n} - \sqrt{\tilde{\eta}_{\kappa,s'}^n}}{\sqrt{\eta_{\kappa,s'}^n}} \right)^2, \quad \forall \kappa \in \mathcal{M}_m, \forall s \in \mathcal{V}_\kappa, \forall n \geq 1,$$

we deduce from (87) that

$$(89) \quad 0 \leq \nu_\beta^n \leq 1, \quad \forall \beta \in \mathcal{M}_m \cup \mathcal{V}_m, \forall n \in \{1, \dots, N_m\}.$$

We can now overestimate  $C_{3,m}$  by

$$\begin{aligned} C_{3,m} &\leq \sum_{n=1}^{N_m} \Delta t_{n,m} \sum_{\kappa \in \mathcal{M}_m} \sqrt{\nu_\kappa^n \eta(u_\kappa^n)} \\ &\quad \times \sum_{s \in \mathcal{V}_\kappa} \sqrt{\eta_{\kappa,s}^n} |p(u_\kappa^n) - p(u_s^n)| \sum_{s' \in \mathcal{V}_\kappa} |a_{s,s'}^\kappa| |\psi_\kappa^{n-1} - \psi_{s'}^{n-1}|, \quad \forall m \geq 1. \end{aligned}$$

Using again that  $ab \leq \epsilon a^2 + \frac{1}{4\epsilon} b^2$ , one gets that

$$(90) \quad C_{3,m} \leq \epsilon C'_{3,m} + \frac{1}{4\epsilon} C''_{3,m}, \quad \forall m \geq 1, \forall \epsilon > 0,$$

where we have set

$$\begin{aligned} C'_{3,m} &= \sum_{n=1}^{N_m} \Delta t_{n,m} \sum_{\kappa \in \mathcal{M}_m} \sum_{s \in \mathcal{V}_\kappa} \left( \sum_{s' \in \mathcal{V}_\kappa} |a_{s,s'}^\kappa| \right) \eta_{\kappa,s}^n (p(u_\kappa^n) - p(u_s^n))^2, \\ C''_{3,m} &= \sum_{n=1}^{N_m} \Delta t_{n,m} \sum_{\kappa \in \mathcal{M}_m} \nu_\kappa^n \eta(u_\kappa^n) \sum_{s \in \mathcal{V}_\kappa} \left( \sum_{s' \in \mathcal{V}_\kappa} |a_{s,s'}^\kappa| \right) (\psi_\kappa^{n-1} - \psi_s^{n-1})^2. \end{aligned}$$

Estimate (72) yields the existence of  $C$  depending only on the data of the continuous problem and on the regularity factors  $\theta^*$  and  $\ell^*$  such that

$$(91) \quad C'_{3,m} \leq C, \quad \forall m \geq 1.$$

It follows from Lemma A.2 stated in appendix that there exists  $C$  depending only on  $\mathbf{\Lambda}$ ,  $\theta^*$ ,  $\ell^*$ , and  $\|\nabla \psi\|_{L^\infty(Q_{t_f})}$  such that

$$\sum_{s \in \mathcal{V}_\kappa} \left( \sum_{s' \in \mathcal{V}_\kappa} |a_{s,s'}^\kappa| \right) (\psi_\kappa^{n-1} - \psi_s^{n-1})^2 \leq C \text{meas}(\kappa),$$

so that

$$C_{3,m}'' \leq C \iint_{Q_{t_f}} \pi_{\mathcal{M}_m, \Delta t_m} (\nu_m \eta(\mathbf{u}_m)) \, d\mathbf{x}.$$

Then thanks to Lemma 4.9 stated hereafter, we can claim that

$$\lim_{m \rightarrow \infty} C_{3,m}'' = 0,$$

and thus, thanks to (90) and (91), that

$$\limsup_{m \rightarrow \infty} C_{3,m} \leq C\epsilon, \quad \forall \epsilon > 0.$$

This ensures that

$$(92) \quad \lim_{m \rightarrow \infty} C_{3,m} = 0.$$

Putting (80) together with (81), (86), and (92), one gets that

$$(93) \quad \lim_{m \rightarrow \infty} C_m = \iint_{Q_{t_f}} \sqrt{\eta(u)} \nabla \xi(u) \cdot \Lambda \nabla \psi \, d\mathbf{x} dt.$$

Now, we will focus on the term  $D_m$ , that can be decomposed into

$$(94) \quad D_m = D_{1,m} + D_{2,m} + D_{3,m}, \quad \forall m \geq 1,$$

where we have set

$$\begin{aligned} D_{1,m} &= \iint_{Q_{t_f}} \pi_{\mathcal{M}_m, \Delta t_m} \eta(\mathbf{u}_m) \nabla_{\mathcal{T}_m} \mathbf{V}_m \cdot \Lambda \nabla_{\mathcal{T}_m, \Delta t_m} \hat{\psi}_m \, d\mathbf{x} dt, \\ D_{2,m} &= \frac{1}{2} \sum_{n=1}^{N_m} \Delta t_{n,m} \sum_{\kappa \in \mathcal{M}_m} \sum_{s \in \mathcal{V}_\kappa} \sum_{s' \in \mathcal{V}_\kappa} a_{s,s'}^\kappa \left( \sqrt{\eta_{\kappa,s}^n} - \sqrt{\eta(u_\kappa^n)} \right) (V_\kappa - V_s) \\ &\quad \times \left( \sqrt{\eta_{\kappa,s'}^n} + \sqrt{\eta(u_\kappa^n)} \right) (\psi_\kappa^{n-1} - \psi_{s'}^{n-1}) \\ D_{3,m} &= \frac{1}{2} \sum_{n=1}^{N_m} \Delta t_{n,m} \sum_{\kappa \in \mathcal{M}_m} \sum_{s \in \mathcal{V}_\kappa} \sum_{s' \in \mathcal{V}_\kappa} a_{s,s'}^\kappa \left( \sqrt{\eta_{\kappa,s}^n} + \sqrt{\eta(u_\kappa^n)} \right) (V_\kappa - V_s) \\ &\quad \times \left( \sqrt{\eta_{\kappa,s'}^n} - \sqrt{\eta(u_\kappa^n)} \right) (\psi_\kappa^{n-1} - \psi_{s'}^{n-1}) \end{aligned}$$

It follows from Lemma 4.5, from the uniform convergence of  $\nabla_{\mathcal{T}_m} \mathbf{V}_m$  towards  $\nabla V$  and of  $\nabla_{\mathcal{T}_m, \Delta t_m} \hat{\psi}_m$  towards  $\nabla \psi$  as  $m$  tends to  $+\infty$  that

$$(95) \quad \lim_{m \rightarrow \infty} D_{1,m} = \int_{Q_{t_f}} \eta(u) \nabla V \cdot \Lambda \nabla \psi \, d\mathbf{x} dt.$$

Let  $\epsilon > 0$ , then it follows from the inequality  $|ab| \leq \epsilon a^2 + \frac{b^2}{4\epsilon}$  that

$$(96) \quad |D_{2,m}| \leq \epsilon D'_{2,m} + \frac{1}{16\epsilon} D''_{2,m}, \quad \forall m \geq 1,$$

where we have set

$$D'_{2,m} = \sum_{n=1}^{N_m} \Delta t_{n,m} \sum_{\kappa \in \mathcal{M}_m} \sum_{s \in \mathcal{V}_\kappa} \left( \sum_{s' \in \mathcal{V}_\kappa} |a_{s,s'}^\kappa| \right) \left( \sqrt{\eta_{\kappa,s}^n} + \sqrt{\eta(u_\kappa^n)} \right)^2 (\psi_\kappa^{n-1} - \psi_s^{n-1})^2,$$

$$D''_{2,m} = \sum_{n=1}^{N_m} \Delta t_{n,m} \sum_{\kappa \in \mathcal{M}_m} \sum_{s \in \mathcal{V}_\kappa} \left( \sum_{s' \in \mathcal{V}_\kappa} |a_{s,s'}^\kappa| \right) \left( \sqrt{\eta_{\kappa,s}^n} - \sqrt{\eta(u_\kappa^n)} \right)^2 (V_\kappa - V_s)^2,$$

Define by  $\bar{\eta}_m = (\bar{\eta}_\kappa^n, \bar{\eta}_s^n)_{\kappa,s} \in W_{\mathcal{D}_m, \Delta t_m}$  by

$$\bar{\eta}_\kappa^n = \max \left( \eta(u_\kappa^n), \max_{s \in \mathcal{V}_\kappa} u_s^n \right), \quad \bar{\eta}_s^n = 0, \quad \forall \kappa \in \mathcal{M}_m, \forall s \in \mathcal{V}_m,$$

then one has

$$\left( \sqrt{\eta_{\kappa,s}^n} + \sqrt{\eta(u_\kappa^n)} \right)^2 \leq \bar{\eta}_\kappa^n, \quad \forall \kappa \in \mathcal{M}_m, \forall s \in \mathcal{V}_\kappa,$$

whence

$$D'_{2,m} \leq C \iint_{Q_{t_f}} \pi_{\mathcal{M}_m, \Delta t_m} \bar{\eta}_m \Lambda \nabla_{\mathcal{T}_m, \Delta t_m} \psi_m \cdot \nabla_{\mathcal{T}_m, \Delta t_m} \psi_m \, dx dt, \quad \forall m \geq 1$$

thanks to Lemma A.2. Using Lemma A.8, we know that there exists  $C$  depending on the data of the continuous problem and of the regularity factors  $\theta^*$ ,  $\ell^*$  and  $\zeta^*$  such that

$$\|\pi_{\mathcal{M}_m, \Delta t_m} \bar{\eta}_m\|_{L^1(Q_{t_f})} \leq C, \quad \forall m \geq 1,$$

while the regularity of  $\psi$  ensures that

$$\|\nabla_{\mathcal{T}_m, \Delta t_m} \psi_m\|_{L^\infty(Q_{t_f})} \leq C, \quad \forall m \geq 1.$$

Therefore, there exists  $C$  depending only on the data of the continuous problem and the regularity factors  $\theta^*$ ,  $\ell^*$  and  $\zeta^*$  such that

$$(97) \quad D'_{2,m} \leq C, \quad \forall m \geq 1.$$

The term  $D''_{2,m}$  can be studied as  $C''_{2,m}$  was, leading to

$$(98) \quad \lim_{m \rightarrow \infty} D''_{2,m} = 0,$$

whence, taking (97)–(98) into account in (96), one gets that

$$(99) \quad \lim_{m \rightarrow \infty} D_{2,m} = 0.$$

Reproducing the calculations carried out for dealing with  $D_{2,m}$  allows to show that

$$(100) \quad \lim_{m \rightarrow \infty} D_{3,m} = 0.$$

Combining (94)–(95) and (99)–(100), we obtain that

$$(101) \quad \lim_{m \rightarrow \infty} D_m = \iint_{Q_{t_f}} \eta(u) \Lambda \nabla V \cdot \nabla \psi \, dx dt.$$

Finally, it follows from (77), (78)–(79), (93) and (101) that the limit  $u$  of the discrete reconstructions  $(\pi_{\mathcal{D}_m, \Delta t_m} \mathbf{u}_m)_{m \geq 1}$  is a weak solution to the problem (1) in the sense of Definition 1.  $\square$

**Lemma 4.8.** *Let  $\mu_m = (\mu_\kappa^n, \mu_s^n)_{\kappa,s,n} \in W_{\mathcal{D}_m, \Delta t_m}$  be defined by (84), then*

$$\pi_{\mathcal{M}_m, \Delta t_m} \mu_m \xrightarrow{m \rightarrow \infty} 0 \quad \text{strongly in } L^1(Q_{t_f}),$$

*Proof.* Notice that for all  $\kappa \in \mathcal{M}_m$ , all  $s \in \mathcal{V}_\kappa$  and all  $n \in \{1, \dots, N_m\}$ , one has

$$\left(\sqrt{\eta_{\kappa,s}^n} - \sqrt{\eta(u_\kappa^n)}\right)^2 \leq \left(\sqrt{\eta(u_s^n)} - \sqrt{\eta(u_\kappa^n)}\right)^2 \leq 2(\eta(u_\kappa^n) + \eta(u_s^n)).$$

Therefore, one gets that

$$(102) \quad \mu_\kappa^n = \max_{s \in \mathcal{V}_\kappa} \left(\sqrt{\eta_{\kappa,s}^n} - \sqrt{\eta(u_\kappa^n)}\right)^2 \leq \max_{s \in \mathcal{V}_\kappa} \left(\sqrt{\eta(u_s^n)} - \sqrt{\eta(u_\kappa^n)}\right)^2 \leq 2 \left(\eta(u_\kappa^n) + \sum_{s \in \mathcal{V}_\kappa} \eta(u_s^n)\right).$$

Then, one obtains that

$$0 \leq \pi_{\mathcal{M}_m, \Delta t_m} \boldsymbol{\mu}_m \leq 2 \sum_{n=1}^{N_m} \Delta t_{n,m} \sum_{\kappa \in \mathcal{M}} \text{meas}(\kappa) \left(\eta(u_\kappa^n) + \sum_{s \in \mathcal{V}_\kappa} \eta(u_s^n)\right).$$

Using (52b)–(52c), which ensure that

$$m_\kappa \geq \frac{\zeta^\star}{d} \text{meas}(\kappa) \quad \text{and} \quad m_s \geq \frac{\zeta^\star}{d\ell^\star} \text{meas}(\kappa),$$

we deduce that there exists  $C$  depending on  $d$ ,  $\ell^\star$  and  $\zeta^\star$  such that

$$0 \leq \pi_{\mathcal{M}_m, \Delta t_m} \boldsymbol{\mu}_m \leq C \pi_{\mathcal{D}_m, \Delta t_m} \eta(\mathbf{u}_m), \quad \forall m \geq 1.$$

As a particular consequence of Lemma 4.5, we know that  $(\pi_{\mathcal{D}_m, \Delta t_m} \eta(\mathbf{u}_m))_{m \geq 1}$  is uniformly equi-integrable, whence

$$(103) \quad (\pi_{\mathcal{M}_m, \Delta t_m} \boldsymbol{\mu}_m)_{m \geq 1} \text{ is uniformly equi-integrable.}$$

Let us introduce  $\mathbf{w}_m = (w_\kappa^n, w_s^n)_{\kappa,s,n} \in W_{\mathcal{D}_m, \Delta t_m}$  defined for all  $\kappa \in \mathcal{M}_m$ , all  $s \in \mathcal{V}_m$  and all  $n \in \{1, \dots, N_m\}$  by

$$(104) \quad w_s^n = 0, \quad w_\kappa^n = \max_{s \in \mathcal{V}_\kappa} |\xi(u_\kappa^n) - \xi(u_s^n)|.$$

It follows from a straightforward generalization of Lemma A.9 and from estimate (71) that  $\pi_{\mathcal{M}_m, \Delta t_m} \mathbf{w}_m$  converges strongly in  $L^2(Q_{t_f})$  towards 0. Therefore, up to an unlabeled subsequence, it converges almost everywhere. As a consequence,

$$(105) \quad \pi_{\mathcal{M}_m, \Delta t_m} \phi(\mathbf{w}_m) \xrightarrow{m \rightarrow \infty} 0 \quad \text{a.e. in } Q_{t_f}$$

for all continuous function  $\phi : \mathbb{R}_+ \rightarrow \mathbb{R}$  such that  $\phi(0) = 0$ .

Since the function  $\sqrt{\eta \circ \xi^{-1}}$  is supposed to be uniformly continuous (cf. (10)), it admits a non-decreasing modulus of continuity  $\varpi \in C(\mathbb{R}_+; \mathbb{R}_+)$  with  $\varpi(0) = 0$  such that, for all  $v, \hat{v}$  in the range of  $\xi$ ,

$$(106) \quad \left| \sqrt{\eta \circ \xi^{-1}}(v) - \sqrt{\eta \circ \xi^{-1}}(\hat{v}) \right| \leq C \varpi(|v - \hat{v}|).$$

Using (102), we obtain that

$$\mu_\kappa^n \leq (\varpi(w_\kappa^n))^2, \quad \forall \kappa \in \mathcal{M}_m, \quad \forall n \in \{1, \dots, N_m\},$$

so that

$$0 \leq \pi_{\mathcal{M}_m, \Delta t_m} \boldsymbol{\mu}_m \leq \pi_{\mathcal{M}_m, \Delta t_m} (\varpi(\mathbf{w}_m))^2.$$

Thanks to (105), we obtain that

$$(107) \quad \pi_{\mathcal{M}_m, \Delta t_m} \boldsymbol{\mu}_m \xrightarrow{m \rightarrow \infty} 0 \quad \text{a.e. in } Q_{t_f}.$$

In order to conclude, it only remains to remark that (103) and (107) allow us to use Vitali's convergence theorem.  $\square$

**Lemma 4.9.** *Let  $\mathbf{u}_m \in W_{\mathcal{D}_m, \Delta t_m}$  be the solution to the numerical scheme, and let  $\boldsymbol{\nu}_m \in W_{\mathcal{D}_m, \Delta t_m}$  be defined by (88), then, up to a subsequence,*

$$\pi_{\mathcal{M}_m, \Delta t_m}(\boldsymbol{\nu}_m \eta(\mathbf{u}_m)) \xrightarrow{m \rightarrow \infty} 0 \quad \text{strongly in } L^1(Q_{t_f}).$$

*Proof.* On the one hand, remark that due to the mass lumping, one has

$$\pi_{\mathcal{M}_m, \Delta t_m}(\boldsymbol{\nu}_m \eta(\mathbf{u}_m)) = (\pi_{\mathcal{M}_m, \Delta t_m} \boldsymbol{\nu}_m) \times (\pi_{\mathcal{M}_m, \Delta t_m} \eta(\mathbf{u}_m)), \quad \forall m \geq 1.$$

Thanks to Lemma 4.5, we know that  $(\pi_{\mathcal{M}_m, \Delta t_m} \eta(\mathbf{u}_m))_{m \geq 1}$  is relatively compact in  $L^1(Q_{t_f})$ , whence it is uniformly equi-integrable. Now it follows from (89) that  $(\pi_{\mathcal{M}_m, \Delta t_m} \boldsymbol{\nu}_m)_{m \geq 1}$  is uniformly bounded in  $L^\infty(Q_{t_f})$ . Therefore, the sequence

$$(108) \quad (\pi_{\mathcal{M}_m, \Delta t_m}(\boldsymbol{\nu}_m \eta(\mathbf{u}_m)))_{m \geq 1} \text{ is uniformly equi-integrable.}$$

On the other hand, since  $\eta(u_\kappa^n) \leq 2\eta_{\kappa, s}^n$ , we know that

$$(109) \quad \nu_\kappa^n \eta(u_\kappa^n) \leq 2 \max_{s \in \mathcal{V}_\kappa} \left( \sqrt{\eta_{\kappa, s}^n} - \sqrt{\tilde{\eta}_{\kappa, s}^n} \right)^2, \quad \forall \kappa \in \mathcal{M}_m, \forall n \in \{1, \dots, N_m\}.$$

The continuity of  $\sqrt{\eta \circ \xi^{-1}}$  provides the existence, for all  $\kappa \in \mathcal{M}_m$ , all  $s \in \mathcal{V}_\kappa$  and all  $n \geq 1$ , of  $\xi_{\kappa, s}^n$  and  $\tilde{\xi}_{\kappa, s}^n$  in  $[\min(\xi(u_\kappa^n), \xi(u_s^n)), \max(\xi(u_\kappa^n), \xi(u_s^n))]$  such that

$$\eta \circ \xi^{-1}(\xi_{\kappa, s}^n) = \eta_{\kappa, s}^n \quad \text{and} \quad \eta \circ \xi^{-1}(\tilde{\xi}_{\kappa, s}^n) = \tilde{\eta}_{\kappa, s}^n.$$

In particular, we know that

$$\left| \xi_{\kappa, s}^n - \tilde{\xi}_{\kappa, s}^n \right| \leq w_\kappa^n, \quad \forall \kappa \in \mathcal{M}_m, \forall s \in \mathcal{V}_\kappa, \forall n \in \{1, \dots, N_m\},$$

where  $w_\kappa^n$  is defined by (104). Since  $\sqrt{\eta \circ \xi^{-1}}$  is uniformly continuous, we deduce from (109) that

$$\nu_\kappa^n \eta(u_\kappa^n) \leq 2(\varpi(w_\kappa^n))^2, \quad \forall \kappa \in \mathcal{M}_m, \forall n \in \{1, \dots, N_m\},$$

where  $\varpi$  denotes the modulus of continuity of  $\sqrt{\eta \circ \xi^{-1}}$  introduced in (106). We conclude as in the proof of Lemma 4.8 that, up to an unlabeled subsequence,

$$(110) \quad (\pi_{\mathcal{M}_m, \Delta t_m}(\boldsymbol{\nu}_m \eta(\mathbf{u}_m)))_{m \geq 1} \xrightarrow{m \rightarrow \infty} 0 \quad \text{a.e. in } Q_{t_f}.$$

Combining (108) with (110) to apply Vitali's convergence theorem allows to conclude the proof.  $\square$

## 5. NUMERICAL IMPLEMENTATION AND RESULTS

This section is devoted to the numerical resolution of the nonlinear system (41). First, we discuss in §5.1 the strategy that we used for solving the nonlinear system (41). Then we present in §5.2 two 2-dimensional cases with analytical solutions in order to illustrate the numerical convergence of the method.

**5.1. Newton method, Schur complement and time-step adaptation.** The nonlinear system (41) obtained at each time step is solved by a Newton-Raphson algorithm. Given  $\mathbf{u}^{n-1} \in W_{\mathcal{D}}$ , this leads to the computation of a sequence  $(\mathbf{u}^{n,i})_{i \geq 0} \subset W_{\mathcal{D}}$  such that  $\mathbf{u}^n = \lim_{i \rightarrow \infty} \mathbf{u}^{n,i}$  is a solution to (41). The variation of the discrete unknowns between two Newton-Raphson algorithm iterations is denoted as follows,

$$d\mathbf{u}^{n,i} = (du_s^{n,i}, du_{\kappa}^{n,i})_{s \in \mathcal{V}, \kappa \in \mathcal{M}} = \mathbf{u}^{n,i+1} - \mathbf{u}^{n,i}, \quad \forall i \geq 0.$$

Let us briefly detail the practical implementation of the iterative procedure allowing to deduce  $\mathbf{u}^n$  from  $\mathbf{u}^{n-1}$ .

- (1) In the case where  $p(0)$  is finite, the initial guess for the Newton algorithm is, as usual, taken as  $\mathbf{u}^{n,0} = (u_s^{n-1}, u_{\kappa}^{n-1})_{\kappa, s}$  for all  $s \in \mathcal{V}, \kappa \in \mathcal{M}$ . In the singular case  $p(0) = -\infty$ , it was proved in Lemma 3.7 that the solution  $\mathbf{u}^n = (u_{\kappa}^n, u_s^n)_{\kappa, s}$  of (41) is such that  $\min_{\beta \in \mathcal{M} \cup \mathcal{V}} u_{\beta}^n > 0$ . Therefore, we can initialize the Newton algorithm by

$$\mathbf{u}^{n,0} = (\max(\epsilon, u_s^{n-1}), \max(\epsilon, u_{\kappa}^{n-1}))_{\kappa, s}.$$

In the computations, we fixed  $\epsilon = 10^{-10}$ .

- (2) The Newton-Raphson algorithm iterations are done until a convergence criterion on the  $L^{\infty}(\Omega)$  norm of the variation of the discrete unknowns is reached or until the maximum number of iterations is reached. At each iteration, the Jacobian matrix resulting of (41) is computed and has the following block structure

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} d\mathbf{u}^{n,i} = \begin{pmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \end{pmatrix},$$

where the sub-matrices have the following sizes:  $\mathbf{A} \in \mathbb{R}^{\#\mathcal{V}} \otimes \mathbb{R}^{\#\mathcal{V}}$ ,  $\mathbf{B} \in \mathbb{R}^{\#\mathcal{V}} \otimes \mathbb{R}^{\#\mathcal{M}}$ ,  $\mathbf{C} \in \mathbb{R}^{\#\mathcal{M}} \otimes \mathbb{R}^{\#\mathcal{V}}$ , and  $\mathbf{D} \in \mathbb{R}^{\#\mathcal{M}} \otimes \mathbb{R}^{\#\mathcal{M}}$ . The sub-vectors at the right hand side have thus the following sizes:  $\mathbf{b}_1 \in \mathbb{R}^{\#\mathcal{V}}$  and  $\mathbf{b}_2 \in \mathbb{R}^{\#\mathcal{M}}$ . The dependence of the sub-matrices and the sub-vectors w.r.t.  $n$  and  $i$  was not highlighted here for the ease of notations. A main characteristic of this block structure is that the block  $\mathbf{D}$  is a non singular diagonal matrix, thus the Schur complement can be easily computed without fill-in to eliminate the variation of the cell unknowns. This allow to reduce the linear system to the variation of the vertices unknowns as usual when using the VAG scheme. The resulting linear system that we have to solve in order to obtain the variation of the vertices unknowns is given by,

$$(111) \quad (\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C})(du_s^{n,i})_{s \in \mathcal{V}} = \mathbf{b}_1 - \mathbf{B}\mathbf{D}^{-1}\mathbf{b}_2,$$

and then the variation of the cell unknowns can be easily deduced by the matrix-vector product below,

$$(112) \quad (du_{\kappa}^{n,i})_{\kappa \in \mathcal{M}} = \mathbf{D}^{-1}(\mathbf{b}_2 - \mathbf{C}(du_s^{n,i})_{s \in \mathcal{V}}).$$

As for the initial step, we have to take into account the singular case at each Newton-Raphson iteration by,

$$\mathbf{u}^{n,i+1} = \max(\mathbf{u}^{n,i} + d\mathbf{u}^{n,i}, \epsilon).$$

- (3) If the Newton-Raphson algorithm iterations stop before the maximum number of iterations is reached, the next time iteration is proceeded by increasing the time step. Otherwise, the current time iteration is recomputed by

reducing the time step.

**Remark.** The increase of the time step is bounded by a maximum value denoted  $\Delta t_{\max}$  and a maximum number of convergence failures of the nonlinear methods imposed in order to abort the simulation in case of a non-convergence.

**5.2. Definition of the test-cases and numerical results.** We present here three 2-dimensional numerical cases where  $\Omega$  is the unit square. The space domain is discretized by using meshes issued from a benchmark on anisotropic diffusion problem [29]. In the following numerical experiments, the tensor is defined by

$$\mathbf{\Lambda} = \begin{pmatrix} l_x & 0 \\ 0 & l_y \end{pmatrix}$$

where  $l_x$  and  $l_y$  are chosen constant in  $\Omega$ , and the exterior potential is defined by  $V(\mathbf{x}) = -\mathbf{g} \cdot \mathbf{x}$  for all  $\mathbf{x} \in \Omega$  where  $\mathbf{g} = (g, 0)^t$  with  $g \in \mathbb{R}_+$ . For these cases, the weights of the VAG scheme defined in (26) are computed here by  $\alpha_{\kappa, s} = \frac{0.1}{\#\mathcal{V}_\kappa}$  for all  $\kappa \in \mathcal{M}$ ,  $s \in \mathcal{V}_\kappa$ . Such an equi-distribution of the measure of  $\Omega$  is suitable in our experiments since we deal with a constant in space tensor  $\mathbf{\Lambda}$ . The question of the mass distribution in the case of heterogeneous problems is discussed in [27, 11]. The linear solver applied to resolve (111) is a home-made direct solver using a gaussian elimination with an optimal reordering. Thanks to the construction of analytical solutions, errors can be computed in the classical discrete  $L^2(Q_{t_f})$ ,  $L^1(Q_{t_f})$  and  $L^\infty(Q_{t_f})$  norms. All the results are presented in the Tables below. Each table provides the mesh size  $h$ , the initial and maximum time steps, the discrete errors, their associated convergence rate and the minimum value of the discrete solution.

**5.2.1. Test 1: Linear Fokker-Planck equation with no-flux boundary condition.** This first test case matches with the problem defined by (1) with the functions  $\eta(u) = u$  on  $\mathbb{R}_+$  and  $p(u) = \log(u)$  and the constant  $g$  is fixed to 1. In this case, note that (1) leads to the linear equation

$$(113) \quad \partial_t u - \nabla \cdot (\mathbf{\Lambda} (\nabla u - u \mathbf{g})) = 0 \quad \text{in } Q_{t_f}$$

with the same boundary and initial conditions as in (1). Therefore we compare the results obtained by (41) with those obtained using, instead of (41c), this particular definition of the fluxes

$$(114) \quad \tilde{F}_{\kappa, s}(\mathbf{u}^n) = \sum_{s' \in \mathcal{V}_\kappa} a_{s, s'}^\kappa (u_\kappa^n - u_{s'}^n) + \frac{u_\kappa^n + u_s^n}{2} \sum_{s' \in \mathcal{V}_\kappa} a_{s, s'}^\kappa (V_\kappa - V_{s'}),$$

$\forall \kappa \in \mathcal{M}, \forall s \in \mathcal{V}_\kappa.$

The resulting scheme is called the “linear scheme”. The numerical convergence of both scheme has been compared on the following analytical solution (built on a 1-dimensional case),

$$(115) \quad \tilde{u}(x, y, t) = \exp\left(-\alpha t + \frac{g}{2}x\right) \left(\pi \cos(\pi x) + \frac{g}{2} \sin(\pi x)\right) + \pi \exp\left(g\left(x - \frac{1}{2}\right)\right),$$

$\forall ((x, y), t) \in \Omega \times (0, t_f),$

where  $\alpha = l_x (\pi^2 + \frac{9}{4})$ . This function satisfies the homogeneous Neumann boundary condition and the property  $\tilde{u}(x, y, t) > 0$  for all  $(x, y, t) \in Q_{t_f}$ .

In order to make a numerical convergence study, we have used a family of triangular meshes. These triangle meshes show no symmetry which could artificially increase the convergence rate. This family of meshes is built through the same pattern, which is reproduced at different scales: the first (coarsest) mesh and the third mesh are shown by Figure 1. For the tests on triangular grid, the final time

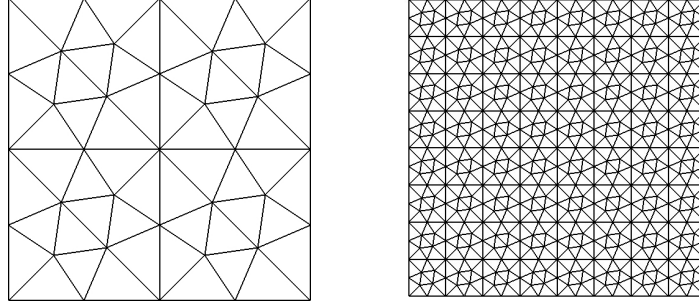


FIGURE 1. First and third mesh used in the numerical examples.

$t_f$  has been chosen to 0.25 and an anisotropic tensor has been consider:  $l_x = 1$  and  $l_y = 10$ .

$h$	$\#\mathcal{V}$	$\Delta t_{\text{init}}$	$\Delta t_{\text{max}}$	$\text{err}_{L^2}$	rate	$\text{err}_{L^1}$	rate	$\text{err}_{L^\infty}$	rate	$u_{\min}$
0.250	37	0.001	0.01024	0.196E-01	-	0.754E-02	-	0.216E+00	-	0.022
0.125	129	0.00025	0.00256	0.512E-02	1.935	0.178E-02	2.084	0.600E-01	1.848	0.004
0.063	481	0.00006	0.00064	0.129E-02	1.986	0.430E-03	2.050	0.157E-01	1.931	0.001
0.031	1857	0.00002	0.00016	0.324E-03	1.997	0.107E-03	2.007	0.473E-02	1.734	0.000

TABLE 1. Triangles. Nonlinear scheme (41).

$h$	$\#\mathcal{V}$	$\Delta t_{\text{init}}$	$\Delta t_{\text{max}}$	$\text{err}_{L^2}$	rate	$\text{err}_{L^1}$	rate	$\text{err}_{L^\infty}$	rate	$u_{\min}$
0.250	37	0.001	0.01024	0.187E-01	-	0.708E-02	-	0.225E+00	-	-0.155
0.125	129	0.00025	0.00256	0.469E-02	1.993	0.165E-02	2.100	0.786E-01	1.515	-0.046
0.063	481	0.00006	0.00064	0.117E-02	1.999	0.406E-03	2.023	0.228E-01	1.784	-0.012
0.031	1857	0.00002	0.00016	0.293E-03	1.999	0.102E-03	1.999	0.611E-02	1.901	-0.003

TABLE 2. Triangles. Linear scheme, fluxes defined by (114).

Let us first observe that, in this case where the continuous solution is regular, the numerical order of convergence is close to 2, using the scheme developed in this paper or the scheme (114). Moreover the nonlinear scheme has better approximation of the lower bound of the solution, since its minimum value is always nonnegative, unlike the linear scheme. This numerical behavior is a verification of the theoretical result mentioned in the Lemma 3.7.

Now, in order to exhibit the ability of the VAG scheme to deal with all type of meshes, the same test case has been applied on a so-called Kershaw grid illustrated on the Figure 2. A numerical convergence study is not relevant on such family



of grids due to the difficulty of defined a factor of refinement. On this grid, the final time  $t_f$  has been chosen to 250 and an anisotropic tensor has been consider:  $l_x = 0.001$  and  $l_y = 1$ . The results are listed on the Table 3 and we can check again the property of the nonlinear scheme which satisfies the lower bound.

	$\#\mathcal{V}$	$\Delta t_{\text{init}}$	$\Delta t_{\text{max}}$	$\text{err}_{L^2}$	$\text{err}_{L^1}$	$\text{err}_{L^\infty}$	$u_{\min}$
nonlinear scheme	324	2.E-04	1	3.99E-02	0.404	1.42E-02	8.92E-04
linear scheme	324	2.E-04	1	3.47E-02	0.377	2.01E-02	-1.49E-02

TABLE 3. Kershaw grid. Nonlinear and linear scheme, with an anisotropic tensor.

Denoting by  $w = \pi \exp(g(x - \frac{1}{2}))$  the long-time asymptotic of  $\tilde{u}$  defined by (115), then the relative entropy of a function  $u : \Omega \rightarrow \mathbb{R}_+$  w.r.t.  $w$  is defined by

$$(116) \quad E^w(u) = \int_{\Omega} \left( u \log \left( \frac{u}{w} \right) - u + w \right) d\mathbf{x}.$$

Note that  $E^w$  is undefined (or is set to  $+\infty$ ) if  $u < 0$  on a positive measure set. Following [34], the anisotropic Fokker-Planck equation has a rigorous gradient flow structure for some Wasserstein metric, for which the entropy  $u \mapsto \int_{\Omega} (u \log(u) - u) d\mathbf{x}$  is geodesically  $\delta$ -convex for some  $\delta > 0$  (this relies on the fact that  $\mathbf{\Lambda}$  does not depend on  $\mathbf{x}$ , see [34, Remark 1.6]). Moreover, the relative entropy  $E^w(\tilde{u}(\cdot, t))$  converges exponentially fast towards 0 as  $t$  tends to  $+\infty$ . A closely related exponential convergence result was transposed to the discrete setting in [19, 18] in the case of a monotone discretization of the diffusion equation (see also [9]). In order to check this asymptotic behavior at the discrete level, we introduce the discrete relative entropy  $E_{\mathcal{D}}^w(\mathbf{u})$  defined for all nonnegative  $\mathbf{u} = (u_{\kappa}, u_s)_{\kappa, s} \in W_{\mathcal{D}}$  (i.e., such that  $u_{\beta} \geq 0$  for all  $\beta \in \mathcal{M} \cup \mathcal{V}$ ) by

$$(117) \quad E_{\mathcal{D}}^w(\mathbf{u}) = \sum_{\beta \in \mathcal{M} \cup \mathcal{V}} m_{\beta} \left( u_{\beta} \log \left( \frac{u_{\beta}}{w(\mathbf{x}_{\beta})} \right) - u_{\beta} + w(\mathbf{x}_{\beta}) \right).$$

The exponential convergence towards equilibrium is recovered as it appears clearly on Figure 2.

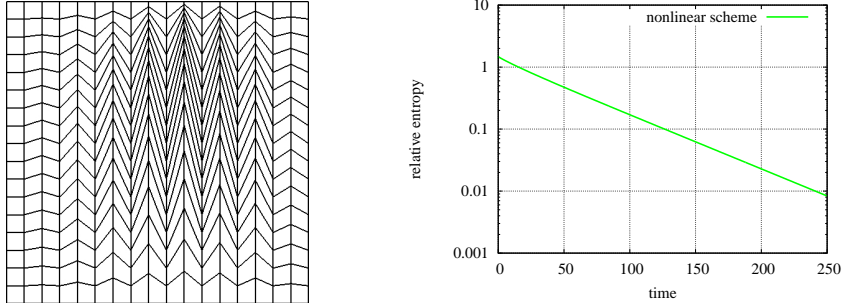


FIGURE 2. Left: Kershaw mesh. Right: Evolution of the relative entropy  $t \mapsto E_{\mathcal{D}}^w(\mathbf{u}(\cdot, t))$  on a logarithmic scale in function of time.

5.2.2. *Test 2: Porous medium equation with Dirichlet boundary condition.* In this second test case, the functions are  $\eta(u) = 2u^2$  on  $\mathbb{R}_+$  and  $p(u) = \log(u)$ , and the potential  $V$  is set to 0, so that the problem turns to

$$(118) \quad \partial_t u - \nabla \cdot (\Lambda \nabla(u^2)) = 0 \quad \text{in } Q_{t_f}$$

Note that since the function  $\eta$  is strictly convex, the rigorous gradient flow structure of the problem corresponding to this choice of mobility function  $\eta$  is unclear [22].

The problem is closed here with Dirichlet boundary conditions. We compare the results obtained by (41) with those obtained using, instead of (41c), this particular definition of the fluxes

$$(119) \quad \hat{F}_{\kappa,s}(\mathbf{u}^n) = \sum_{s' \in \mathcal{V}_\kappa} a_{s,s'}^\kappa ((u_\kappa^n)^2 - (u_{s'}^n)^2), \quad \forall \kappa \in \mathcal{M}, \forall s \in \mathcal{V}_\kappa.$$

The resulting scheme is called the “quasilinear scheme”. The numerical convergence of both schemes has been compared thanks to the following analytical solution (again built in 1-dimension),

$$(120) \quad \hat{u}(x, y, t) = \max(2l_x x - t, 0), \quad \forall ((x, y), t) \in Q_{t_f}.$$

Note that (120) is the unique weak solution corresponding to the initial condition  $u_0(x, y) = \hat{u}(x, y, 0)$  and to the Dirichlet boundary condition  $u_D(x, y, t) = \hat{u}(x, y, t)$  on  $\partial\Omega \times (0, t_f)$ . Our numerical convergence study makes use of the family of triangular meshes presented on the previous test case. Once again, the final time  $t_f$  is fixed to 0.25 and an anisotropic tensor is given by  $l_x = 1$  and  $l_y = 10$ .

$h$	$\#\mathcal{V}$	$\Delta t_{\text{init}}$	$\Delta t_{\text{max}}$	$\text{err}_{L^2}$	rate	$\text{err}_{L^1}$	rate	$\text{err}_{L^\infty}$	rate	$u_{\min}$
0.306	37	0.001	0.01024	0.523E-02	-	0.997E-03	-	0.105E+00	-	0.000
0.153	129	0.00025	0.00256	0.205E-02	1.352	0.344E-03	1.535	0.522E-01	1.013	0.000
0.077	481	0.00006	0.00064	0.898E-03	1.190	0.123E-03	1.490	0.259E-01	1.012	0.000
0.038	1857	0.00002	0.00016	0.380E-03	1.240	0.417E-04	1.554	0.128E-01	1.012	0.000

TABLE 4. Test 2. Nonlinear scheme (41).

$h$	$\#\mathcal{V}$	$\Delta t_{\text{init}}$	$\Delta t_{\text{max}}$	$\text{err}_{L^2}$	rate	$\text{err}_{L^1}$	rate	$\text{err}_{L^\infty}$	rate	$u_{\min}$
0.306	37	0.001	0.01024	0.116E-01	-	0.371E-02	-	0.764E-01	-	-0.065
0.153	129	0.00025	0.00256	0.423E-02	1.461	0.116E-02	1.672	0.388E-01	0.977	-0.039
0.077	481	0.00006	0.00064	0.149E-02	1.501	0.337E-03	1.788	0.233E-01	0.737	-0.021
0.038	1857	0.00002	0.00016	0.524E-03	1.513	0.932E-04	1.856	0.129E-01	0.856	-0.010

TABLE 5. Test 2. Quasilinear scheme, with fluxes defined by (119).

We first observe that second order convergence is destroyed because of the lack of regularity of the exact solution. Here again, the discrete solution of the nonlinear scheme studied in this paper remains positive as expected by the Lemma 3.7, while the discrete solution to the quasilinear scheme suffers from undershoots. Despite the convergence for the nonlinear scheme seems to be slightly slower than the convergence for the quasilinear scheme, the accuracy of the nonlinear scheme is better, cf. Tables 4 and 5.

The Figures 3 illustrate the iso-values of the piecewise affine function defined on the triangular mesh  $\mathcal{M}$  reconstructed thanks to its nodal values  $(u_s^n)_{s \in \mathcal{V}}$  for the coarsest triangle grid at the final time  $t_f$ . In the case of the nonlinear scheme, the iso-values are chosen from 0 to 0.025 by step of 0.05 and then from 0.1 to 0.5 by

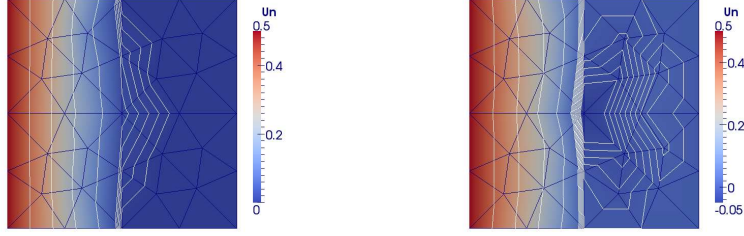


FIGURE 3. Test 2. Coarsest grid. Discrete unknown  $(u_s)_{s \in \mathcal{V}_\kappa}$  and its iso-values. Left: Nonlinear scheme. Right: Quasilinear scheme.

step of 0.1. For the quasilinear scheme, the iso-values are taken from  $-0.025$  to  $0.025$  by step of  $0.05$  and also from  $0.1$  to  $0.5$  by step of  $0.1$ .

5.2.3. *Test 3: Porous medium equation with drift.* In this third test case, we have set  $\eta(u) = u$  on  $\mathbb{R}_+$  and  $p(u) = 2u$  and  $g = 1$ , leading to the degenerate problem

$$(121) \quad \partial_t u - \nabla \cdot (\mathbf{\Lambda} (\nabla(u^2) - u\mathbf{g})) = 0 \quad \text{in } Q_{t_f}.$$

The problem is endowed with Dirichlet boundary conditions. The tensor  $\mathbf{\Lambda}$  is chosen to be diagonal with  $l_x = 1$  and  $l_y = 100$ . We compare the results obtained by (41) with those obtained using, instead of (41c), this particular definition of the fluxes

$$(122) \quad \hat{F}_{\kappa,s}(\mathbf{u}^n) = \sum_{s' \in \mathcal{V}_\kappa} a_{s,s'}^\kappa ((u_\kappa^n)^2 - (u_{s'}^n)^2) + \frac{u_\kappa^n + u_s^n}{2} \sum_{s' \in \mathcal{V}_\kappa} a_{s,s'}^\kappa (V_\kappa - V_{s'}),$$

$$\forall \kappa \in \mathcal{M}, \forall s \in \mathcal{V}_\kappa.$$

The resulting scheme is again called the “quasilinear scheme”. The numerical convergence of both schemes has been compared on the sequence of triangular meshes already used in the previous tests, thanks to the following analytical solution (again built in 1-dimension),

$$(123) \quad \hat{u}(x, y, t) = \max(\beta x - t, 0), \quad \forall ((x, y), t) \in \Omega \times (0, t_f),$$

with  $\beta = l_x(2 + g)$ . The profile (123) is the unique weak solution corresponding to the initial condition  $u_0(x, y) = \hat{u}(x, y, 0)$  in  $\Omega$  and the Dirichlet boundary condition  $u_D(x, y, t) = \hat{u}(x, y, t)$  on  $\partial\Omega \times (0, t_f)$ .

$h$	$\#\mathcal{V}$	$\Delta t_{\text{init}}$	$\Delta t_{\text{max}}$	$\text{err}_{L^2}$	rate	$\text{err}_{L^1}$	rate	$\text{err}_{L^\infty}$	rate	$u_{\min}$
0.306	37	0.001	0.01024	0.130E-01	-	0.423E-02	-	0.890E-01	-	-0.046
0.153	129	0.00025	0.00256	0.495E-02	1.398	0.133E-02	1.675	0.496E-01	0.843	-0.032
0.077	481	0.00006	0.00064	0.184E-02	1.428	0.397E-03	1.741	0.283E-01	0.808	-0.017
0.038	1857	0.00002	0.00016	0.660E-03	1.479	0.116E-03	1.771	0.145E-01	0.970	-0.009

TABLE 6. Test 3. Nonlinear scheme (41).

Here again, the convergence orders of both scheme are similar, but strictly lower than 2 because of the lack of regularity of the exact solution. We also remark that both schemes violate the positivity of the solution in this case, but that the amplitude of the undershoots is slightly smaller for the nonlinear scheme. There is no contradiction here with Lemma 3.7 since  $p$  is not singular at  $u = 0$ .

$h$	$\#\mathcal{V}$	$\Delta t_{\text{init}}$	$\Delta t_{\text{max}}$	$\text{err}_{L^2}$	rate	$\text{err}_{L^1}$	rate	$\text{err}_{L^\infty}$	rate	$u_{\min}$
0.306	37	0.001	0.01024	0.154E-01	-	0.568E-02	-	0.939E-01	-	-0.068
0.153	129	0.00025	0.00256	0.671E-02	1.201	0.213E-02	1.416	0.613E-01	0.615	-0.048
0.077	481	0.00006	0.00064	0.271E-02	1.309	0.702E-03	1.600	0.326E-01	0.910	-0.027
0.038	1857	0.00002	0.00016	0.104E-02	1.384	0.212E-03	1.725	0.170E-01	0.938	-0.015

TABLE 7. Test 3. Quasilinear scheme, fluxes defined by (122).

## APPENDIX A. SOME LEMMAS RELATED TO THE VAG DISCRETIZATION

In this section,  $\mathcal{D} = (\mathcal{M}, \mathcal{T})$  denote a discretization of  $\Omega$  as prescribed in §2.1.1.

**Lemma A.1.** *For  $\kappa \in \mathcal{M}$ , let  $\mathbf{A}_\kappa = (a_{s,s'}^\kappa)_{s,s' \in \mathcal{V}_\kappa}$  be the matrix defined by (33), then there exists  $C$  depending only on  $\mathbf{A}$ ,  $\theta_{\mathcal{T}}$  and  $\ell_{\mathcal{D}}$  (but not on  $\kappa$ ) such that  $\text{Cond}_2(\mathbf{A}_\kappa) \leq C$ .*

*Proof.* Following [12, Lemma 3.2], there exist  $C_1, C_2 > 0$  depending only on  $\theta_{\mathcal{T}}$  and  $\ell_{\mathcal{D}}$  such that, for all  $\mathbf{u} \in W_{\mathcal{D}}$  and all  $\kappa \in \mathcal{M}$ , one has

$$C_1 \frac{\text{meas}(\kappa)}{(h_\kappa)^2} \sum_{s \in \mathcal{V}_\kappa} (u_s - u_\kappa)^2 \leq \|\nabla_{\mathcal{T}} \mathbf{u}\|_{L^2(\kappa)}^2 \leq C_2 \frac{\text{meas}(\kappa)}{(h_\kappa)^2} \sum_{s \in \mathcal{V}_\kappa} (u_s - u_\kappa)^2,$$

where  $h_\kappa$  denotes the diameter of the cell  $\kappa \in \mathcal{M}$ . As a consequence, one has

$$\lambda_* C_1 \frac{\text{meas}(\kappa)}{(h_\kappa)^2} |\delta_\kappa \mathbf{u}|^2 \leq \delta_\kappa \mathbf{u} \cdot \mathbf{A}_\kappa \delta_\kappa \mathbf{u} = \int_\kappa \mathbf{A} \nabla_{\mathcal{T}} \mathbf{u} \cdot \nabla_{\mathcal{T}} \mathbf{u} \, dx \leq \lambda^* C_2 \frac{\text{meas}(\kappa)}{(h_\kappa)^2} |\delta_\kappa \mathbf{u}|^2.$$

Since the application  $\delta_\kappa : W_{\mathcal{D}} \rightarrow \mathbb{R}^{\ell_\kappa}$  is onto, we deduce that

$$\lambda_* C_1 \frac{\text{meas}(\kappa)}{(h_\kappa)^2} |\mathbf{v}|^2 \leq \mathbf{v} \cdot \mathbf{A}_\kappa \mathbf{v} \leq \lambda^* C_2 \frac{\text{meas}(\kappa)}{(h_\kappa)^2} |\mathbf{v}|^2, \quad \forall \mathbf{v} \in \mathbb{R}^{\ell_\kappa},$$

and thus that  $\text{Cond}_2(\mathbf{A}_\kappa) \leq \frac{\lambda^* C_2}{\lambda_* C_1}$ .  $\square$

**Lemma A.2.** *There exists  $C > 0$  depending only on  $\mathbf{A}$ ,  $\theta_{\mathcal{T}}$  and  $\ell_{\mathcal{D}}$  such that, for all  $\kappa \in \mathcal{M}$  and all  $\mathbf{v} = (v_s)_{s \in \mathcal{V}_\kappa} \in \mathbb{R}^{\ell_\kappa}$ , one has*

$$\sum_{s \in \mathcal{V}_\kappa} \left( \sum_{s' \in \mathcal{V}_\kappa} |a_{s,s'}^\kappa| \right) (v_s)^2 \leq C \mathbf{v} \cdot \mathbf{A}_\kappa \mathbf{v}.$$

*Proof.* Denoting by  $\|\cdot\|_q$  the usual matrix  $q$ -norm, one has

$$\sum_{s \in \mathcal{V}_\kappa} \left( \sum_{s' \in \mathcal{V}_\kappa} |a_{s,s'}^\kappa| \right) (v_s)^2 \leq \|\mathbf{A}_\kappa\|_1 |\mathbf{v}|^2.$$

Since the dimension of the space  $\mathbb{R}^{\ell_\kappa}$  is bounded by  $\ell_{\mathcal{D}}$ , there exists  $C_1$  depending only on  $\ell_{\mathcal{D}}$  such that  $\|\mathbf{A}_\kappa\|_1 \leq C_1 \|\mathbf{A}_\kappa\|_2$ , so that

$$(124) \quad \sum_{s \in \mathcal{V}_\kappa} \left( \sum_{s' \in \mathcal{V}_\kappa} |a_{s,s'}^\kappa| \right) (v_s)^2 \leq C_1 \|\mathbf{A}_\kappa\|_2 |\mathbf{v}|^2.$$

On the other hand, since  $\mathbf{A}_\kappa$  is symmetric definite and positive, one has

$$\mathbf{v} \cdot \mathbf{A}_\kappa \mathbf{v} \geq \frac{\|\mathbf{A}_\kappa\|_2}{\text{Cond}_2(\mathbf{A}_\kappa)} |\mathbf{v}|^2.$$

Using Lemma A.1, we obtain that there exists  $C_2 > 0$  depending only on  $\mathbf{A}$ ,  $\theta_{\mathcal{T}}$  and  $\ell_{\mathcal{D}}$  such that

$$(125) \quad \mathbf{v} \cdot \mathbf{A}_{\kappa} \mathbf{v} \geq C_2 \|\mathbf{A}_{\kappa}\|_2 |\mathbf{v}|^2.$$

Putting (124) and (125) together, we conclude the proof of Lemma A.2 by choosing  $C = \frac{C_1}{C_2}$ .  $\square$

**Lemma A.3.** *Let  $\kappa \in \mathcal{M}$  and  $\mathbf{A}_{\kappa} = (a_{s,s'}^{\kappa})_{s,s' \in \mathcal{V}_{\kappa}} \in \mathbb{R}^{\ell_{\kappa} \times \ell_{\kappa}}$  be the matrix defined by (33). Let  $\mu_{\kappa} = (\mu_{\kappa,s})_{s \in \mathcal{V}_{\kappa}} \in \mathbb{R}^{\ell_{\kappa}}$  and  $\mathbf{v} \in W_{\mathcal{D}}$ , then*

$$\sum_{s \in \mathcal{V}_{\kappa}} \sum_{s' \in \mathcal{V}_{\kappa}} (v_s - v_{\kappa}) \mu_{\kappa,s} a_{s,s'}^{\kappa} \mu_{\kappa,s'} (v_{s'} - v_{\kappa}) \leq \max_{s \in \mathcal{V}_{\kappa}} (\mu_{\kappa,s})^2 \sum_{s \in \mathcal{V}_{\kappa}} \left( \sum_{s' \in \mathcal{V}_{\kappa}} |a_{s,s'}^{\kappa}| \right) (v_s - v_{\kappa})^2.$$

*Proof.* Using  $ab \leq \frac{a^2}{2} + \frac{b^2}{2}$ , we obtain that

$$\begin{aligned} & \sum_{s \in \mathcal{V}_{\kappa}} \sum_{s' \in \mathcal{V}_{\kappa}} (v_s - v_{\kappa}) \mu_{\kappa,s} a_{s,s'}^{\kappa} \mu_{\kappa,s'} (v_{s'} - v_{\kappa}) \\ & \leq \max_{s \in \mathcal{V}_{\kappa}} (\mu_{\kappa,s})^2 \sum_{s \in \mathcal{V}_{\kappa}} \sum_{s' \in \mathcal{V}_{\kappa}} |v_s - v_{\kappa}| |a_{s,s'}^{\kappa}| |v_{s'} - v_{\kappa}| \\ & \leq \frac{\max_{s \in \mathcal{V}_{\kappa}} (\mu_{\kappa,s})^2}{2} \sum_{s \in \mathcal{V}_{\kappa}} \left( \sum_{s' \in \mathcal{V}_{\kappa}} |a_{s,s'}^{\kappa}| \right) (v_s - v_{\kappa})^2 \\ & \quad + \frac{\max_{s \in \mathcal{V}_{\kappa}} (\mu_{\kappa,s})^2}{2} \sum_{s' \in \mathcal{V}_{\kappa}} \left( \sum_{s \in \mathcal{V}_{\kappa}} |a_{s,s'}^{\kappa}| \right) (v_{s'} - v_{\kappa})^2. \end{aligned}$$

One concludes the proof of Lemma A.3 by noticing that, since  $\mathbf{A}_{\kappa}$  is symmetric, the two terms in the right-hand side of the above inequality are equal.  $\square$

**Lemma A.4.** *There exists  $C$  depending only on  $\theta_{\mathcal{T}}$  and  $\ell_{\mathcal{D}}$  such that*

$$\text{meas}(T) \leq \text{meas}(\kappa) \leq C \text{meas}(T), \quad \forall \kappa \in \mathcal{M}, \forall T \in \mathcal{T} \text{ with } T \subset \kappa.$$

*Proof.* Let  $\kappa \in \mathcal{M}$ , then there exist  $T_1, \dots, T_r$  simplexes, with  $r = \ell_{\kappa}$  if  $d = 2$  and  $r = 2\#\mathcal{E}_{\kappa}$  if  $d = 3$ , such that

$$\bigcup_{i=1}^{r_{\kappa}} \overline{T}_i = \overline{\kappa}, \quad T_i \cap T_j = \emptyset \text{ if } i \neq j.$$

The Euler-Descartes theorem ensures that  $r \leq 4(\ell_{\mathcal{D}} - 1)$  if  $d = 3$ .

If  $T_i$  and  $T_j$  share a common edge, one gets that

$$\text{meas}(T_i) \leq \theta^d \text{meas}(T_j).$$

Let  $i_0, i_1 \in \{1, \dots, r_{\kappa}\}$  be arbitrary but different, we deduce from the previous inequality the following non-optimal estimate:

$$\text{meas}(T_{i_0}) \leq \theta^{4(\ell_{\mathcal{D}}-1)d} \text{meas}(T_{i_1}).$$

Let  $i_{\max}$  be such that  $\text{meas}(T_{i_{\max}}) = \max_{1 \leq i \leq r} \text{meas}(T_i)$ , then

$$\text{meas}(\kappa) \leq r \text{meas}(T_{i_{\max}}) \leq 4(\ell_{\mathcal{D}} - 1) \theta^{4(\ell_{\mathcal{D}}-1)d} \text{meas}(T_i), \quad \forall i \in \{1, \dots, r\}.$$

$\square$

We state now a slight generalization of [12, Lemma 3.4], where the same result is proven in the particular case  $q = 2$ . The straightforward adaptation of the proof given in [12] to the case  $q \neq 2$  is left to the reader.

**Lemma A.5.** *There exists  $C$  depending only on  $\ell_{\mathcal{D}}$  and  $\theta_{\mathcal{T}}$  defined in (23) and (20) respectively such that*

$$\|\pi_{\mathcal{D}}\mathbf{v} - \pi_{\mathcal{T}}\mathbf{v}\|_{L^q(\Omega)} \leq Ch_{\mathcal{T}} \|\nabla_{\mathcal{T}}\mathbf{v}\|_{L^q(\Omega)}, \quad \forall \mathbf{v} \in W_{\mathcal{D}}, \quad \forall q \in [1, \infty].$$

**Lemma A.6.** *Let  $\mathcal{D}$  be a discretization of  $\Omega$  as introduced in §2.1.1 such that  $\zeta_{\mathcal{D}} > 0$ , then there exist  $C_1 > 0$  depending only on  $q$ ,  $\theta_{\mathcal{T}}$  and  $\ell_{\mathcal{D}}$  and  $C_2$  depending moreover on  $\zeta_{\mathcal{D}}$  such that*

$$(126) \quad C_1 \|\pi_{\mathcal{D}}\mathbf{v}\|_{L^q(\Omega)} \leq \|\pi_{\mathcal{T}}\mathbf{v}\|_{L^q(\Omega)} \leq C_2 \|\pi_{\mathcal{D}}\mathbf{v}\|_{L^q(\Omega)}, \quad \forall \mathbf{v} \in W_{\mathcal{D}}.$$

*Proof.* Let  $\hat{T}$  be a reference tetrahedron, and let  $\hat{v} : \hat{T} \rightarrow \mathbb{R}$  be an affine function with nodal values  $v_i$ ,  $i \in \{1, \dots, 4\}$ , then for all  $q > 0$ , there exists  $C$  depending on  $q$  such that

$$\frac{1}{C} \sum_{i=1}^4 |v_i|^q \leq \|\hat{v}\|_{L^q(\hat{T})}^q \leq C \sum_{i=1}^4 |v_i|^q.$$

Therefore, using classical properties of the affine change of variable between simplexes, one gets the existence of  $C$  depending only on  $q$ ,  $\theta_{\mathcal{T}}$ , and  $\ell_{\mathcal{D}}$  such that, for all  $\mathbf{v} \in W_{\mathcal{D}}$ ,

$$(127) \quad \frac{1}{C} \sum_{\kappa \in \mathcal{M}} \text{meas}(\kappa) \left( |v_{\kappa}|^q + \sum_{s \in \mathcal{V}_{\kappa}} |v_s|^q \right) \leq \|\pi_{\mathcal{T}}\mathbf{v}\|_{L^q(\Omega)}^q \leq C \sum_{\kappa \in \mathcal{M}} \text{meas}(\kappa) \left( |v_{\kappa}|^q + \sum_{s \in \mathcal{V}_{\kappa}} |v_s|^q \right).$$

On the other hand, one has

$$\|\pi_{\mathcal{D}}\mathbf{v}\|_{L^q(\Omega)}^q = \sum_{\kappa \in \mathcal{M}} m_{\kappa} |v_{\kappa}|^q + \sum_{s \in \mathcal{V}} m_s |v_s|^q.$$

A classical geometrical property and (28) yield

$$(128) \quad m_{\kappa} \leq \text{meas}(\kappa) = d \int_{\Omega} \pi_{\mathcal{T}} \mathbf{e}_{\kappa}(\mathbf{x}) d\mathbf{x} \leq \frac{d}{\zeta_{\mathcal{D}}} m_{\kappa} \quad \forall \kappa \in \mathcal{M},$$

and similarly

$$m_s \leq d \int_{\Omega} \pi_{\mathcal{T}} \mathbf{e}_s(\mathbf{x}) d\mathbf{x} \leq \frac{d}{\zeta_{\mathcal{D}}} m_s, \quad \forall s \in \mathcal{V}.$$

Notice now that the following geometrical identity holds:

$$d \int_{\Omega} \pi_{\mathcal{T}} \mathbf{e}_s(\mathbf{x}) d\mathbf{x} = \sum_{\substack{T \in \mathcal{T} \\ \mathbf{x}_s \in \partial T}} \text{meas}(T), \quad \forall s \in \mathcal{V}.$$

Lemma A.4 yields the existence of  $C > 0$  depending on  $\theta_{\mathcal{T}}$  and  $\ell_{\mathcal{D}}$  such that

$$\frac{1}{C} \sum_{\kappa \in \mathcal{M}_s} \text{meas}(\kappa) \leq d \int_{\Omega} \pi_{\mathcal{T}} \mathbf{e}_s(\mathbf{x}) d\mathbf{x} \leq \sum_{\kappa \in \mathcal{M}_s} \text{meas}(\kappa), \quad \forall s \in \mathcal{V},$$

and the result of Lemma A.6 follows.  $\square$

**Lemma A.7.** *Let  $\mathcal{D}$  be a discretization of  $\Omega$  as introduced in §2.1.1 such that  $\zeta_{\mathcal{D}} > 0$ , then, for all  $q \in [1, \infty]$ , one has*

$$\|\pi_{\mathcal{M}} \mathbf{v}\|_{L^q(\Omega)} \leq \left(\frac{d}{\zeta_{\mathcal{D}}}\right)^{1/q} \|\pi_{\mathcal{D}} \mathbf{v}\|_{L^q(\Omega)}, \quad \forall \mathbf{v} \in W_{\mathcal{D}}.$$

*Proof.* Let  $\mathbf{v} = (v_{\kappa}, v_s)_{\kappa \in \mathcal{M}, s \in \mathcal{V}} \in W_{\mathcal{D}}$ , then it follows from (128) that

$$\begin{aligned} \|\pi_{\mathcal{M}} \mathbf{v}\|_{L^q(\Omega)}^q &= \sum_{\kappa \in \mathcal{M}} \text{meas}(\kappa) |v_{\kappa}|^q \\ &\leq \left(\frac{d}{\zeta_{\mathcal{D}}}\right) \sum_{\kappa \in \mathcal{M}} m_{\kappa} |v_{\kappa}|^q \leq \left(\frac{d}{\zeta_{\mathcal{D}}}\right) \|\pi_{\mathcal{D}} \mathbf{v}\|_{L^q(\Omega)}^q. \end{aligned}$$

□

**Lemma A.8.** *Let  $\mathbf{v} = (v_{\kappa}, v_s)_{\kappa, s} \in W_{\mathcal{D}}$  be such that  $v_{\beta} \geq 0$  for all  $\beta \in \mathcal{M} \cup \mathcal{V}$ , and define  $\bar{\mathbf{v}} = (\bar{v}_{\kappa}, \bar{v}_s)_{\kappa, s} \in W_{\mathcal{D}}$  by*

$$\bar{v}_s = v_s, \quad \bar{v}_{\kappa} = \max \left( v_{\kappa}, \max_{s' \in \mathcal{V}_{\kappa}} v_{s'} \right), \quad \forall s \in \mathcal{V}, \forall \kappa \in \mathcal{M}.$$

*Then there exists  $C$  depending only on  $\theta_{\mathcal{T}}$ ,  $\ell_{\mathcal{D}}$  and  $\zeta_{\mathcal{D}}$  such that*

$$\|\pi_{\mathcal{M}} \bar{\mathbf{v}}\|_{L^1(\Omega)} \leq C \|\pi_{\mathcal{D}} \mathbf{v}\|_{L^1(\Omega)}.$$

*Proof.* Let  $\mathbf{v} \in W_{\mathcal{D}}$  be a vector with positives coordinates, and let  $\bar{\mathbf{v}}$  be constructed as above. It follows from the construction of  $\bar{\mathbf{v}}$  that

$$\bar{v}_{\kappa} \leq v_{\kappa} + \sum_{s \in \mathcal{V}_{\kappa}} v_s, \quad \forall \kappa \in \mathcal{M},$$

whence, applying (127) with  $q = 1$ , one gets

$$\|\pi_{\mathcal{M}} \bar{\mathbf{v}}\|_{L^1(\Omega)} \leq C \|\pi_{\mathcal{T}} \mathbf{v}\|_{L^1(\Omega)}.$$

The result now directly follows from Lemma A.6. □

**Lemma A.9.** *Let  $\mathbf{u} = (u_{\kappa}, u_s)_{\kappa \in \mathcal{M}, s \in \mathcal{V}} \in W_{\mathcal{D}}$ , then for all  $\kappa \in \mathcal{M}$ , we define  $\bar{\delta} \mathbf{u} = (\bar{\delta}_{\kappa} \mathbf{u}, \bar{\delta}_s \mathbf{u})_{\kappa \in \mathcal{M}, s \in \mathcal{V}} \in W_{\mathcal{D}}$  by*

$$\bar{\delta}_s \mathbf{u} = 0 \quad \text{and} \quad \bar{\delta}_{\kappa} \mathbf{u} = \max_{s' \in \mathcal{V}_{\kappa}} |u_{\kappa} - u_{s'}|, \quad \forall \kappa \in \mathcal{M}, \forall s \in \mathcal{V},$$

*then, for all  $q \in [1, \infty]$ , there exists  $C$  depending only on  $q$ ,  $\theta_{\mathcal{T}}$ , and  $\ell_{\mathcal{D}}$  such that*

$$(129) \quad \|\pi_{\mathcal{M}} \bar{\delta} \mathbf{u}\|_{L^q(\Omega)} \leq C h_{\mathcal{T}} \|\nabla_{\mathcal{T}} \mathbf{u}\|_{L^q(\Omega)}.$$

*Proof.* Let  $\kappa \in \mathcal{M}$  and  $s \in \mathcal{V}_{\kappa}$ , then there exists a simplicial sub-element  $T \in \mathcal{T}$  of  $\kappa \in \mathcal{M}$  such that  $\mathbf{x}_{\kappa}$  and  $\mathbf{x}_s$  are vertices of  $T$ . Then it follows from classical finite element arguments (see e.g. [20, 23]) that

$$\text{meas}(T)^{1/q} |u_{\kappa} - u_s| \leq c \frac{(h_T)^2}{\rho_T} \|\nabla_{\mathcal{T}} \mathbf{u}\|_{L^q(T)} \leq C h_{\mathcal{T}} \|\nabla_{\mathcal{T}} \mathbf{u}\|_{L^q(\kappa)},$$

where  $c$  depends only on the dimension  $d$  and on  $q$ , while  $C$  depends additionally on  $\theta_{\mathcal{T}}$ . Thanks to Lemma A.4, we get the existence of  $C$  depending on  $d$ ,  $q$ ,  $\theta_{\mathcal{T}}$  and  $\ell_{\mathcal{D}}$  such that,

$$\text{meas}(\kappa)^{1/q} |u_{\kappa} - u_s| \leq C h_{\mathcal{T}} \|\nabla_{\mathcal{T}} \mathbf{u}\|_{L^q(\kappa)}, \quad \forall \kappa \in \mathcal{M}, \forall s \in \mathcal{V}_{\kappa}.$$

Summing over  $\kappa \in \mathcal{M}$  provides that (129) holds. □

## REFERENCES

- [1] H. W. Alt and S. Luckhaus. Quasilinear elliptic-parabolic differential equations. *Math. Z.*, 183(3):311–341, 1983.
- [2] L. Ambrosio, E. Mainini, and S. Serfaty. Gradient flow of the Chapman-Rubinstein-Schatzman model for signed vortices. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 28(2):217–246, 2011.
- [3] L. Ambrosio and S. Serfaty. A gradient flow approach to an evolution problem arising in superconductivity. *Comm. Pure Appl. Math.*, 61(11):1495–1539, 2008.
- [4] B. Andreianov. Time compactness tools for discretized evolution equations and applications to degenerate parabolic PDEs. In *Finite volumes for complex applications. VI. Problems & perspectives. Volume 1, 2*, volume 4 of *Springer Proc. Math.*, pages 21–29. Springer, Heidelberg, 2011.
- [5] B. Andreianov and F. Bouhssiss. Uniqueness for an elliptic-parabolic problem with Neumann boundary condition. *J. Evol. Equ.*, 4(2):273–295, 2004.
- [6] B. Andreianov, C. Cancès, and A. Moussa. A nonlinear time compactness result and applications to discretization of degenerate parabolic-elliptic PDEs. In preparation.
- [7] S. N. Antontsev, A. V. Kazhikhov, and V. N. Monakhov. *Boundary value problems in mechanics of nonhomogeneous fluids*, volume 22 of *Studies in Mathematics and its Applications*. North-Holland Publishing Co., Amsterdam, 1990. Translated from the Russian.
- [8] J. Bear. *Dynamic of Fluids in Porous Media*. American Elsevier, New York, 1972.
- [9] Marianne Bessemoulin-Chatard. *Développement et analyse de schémas volumes finis motivés par la présentation de comportements asymptotiques. Application à des modèles issus de la physique et de la biologie*. PhD thesis, Université Blaise Pascal - Clermont-Ferrand II, 2012.
- [10] A. Blanchet. A gradient flow approach to the Keller-Segel systems. RIMS Kokyuroku’s lecture notes, 2014.
- [11] K. Brenner, Groza M., C. Guichard, and R. Masson. Vertex approximate gradient scheme for hybrid dimensional two-phase Darcy flows in fractured porous media. *ESAIM Math. Model. Numer. Anal.*, 49(2):303–330, 2015.
- [12] K. Brenner and R. Masson. Convergence of a vertex centered discretization of two-phase darcy flows on general meshes. *Int. J. Finite Vol.*, 10:1–37, 2013.
- [13] C. Cancès and T. Gallouët. On the time continuity of entropy solutions. *J. Evol. Equ.*, 11(1):43–55, 2011.
- [14] C. Cancès and C. Guichard. Convergence of a nonlinear entropy diminishing control volume finite element scheme for solving anisotropic degenerate parabolic equations. HAL : hal-00955091, to appear in *Math. Comp.*, 2014.
- [15] C. Cancès and C. Guichard. Entropy-diminishing CVFE scheme for solving anisotropic degenerate diffusion equations. In *Finite volumes for complex applications. VII. Methods and theoretical aspects*, volume 77 of *Springer Proc. Math. Stat.*, pages 187–196. Springer, Cham, 2014.
- [16] C. Cancès, L. Monsaigeon, and T. O. Gallouët. The gradient flow structure of immiscible incompressible two-phase flows in porous media. In preparation.
- [17] C. Cancès and M. Pierre. An existence result for multidimensional immiscible two-phase flows with discontinuous capillary pressure field. *SIAM J. Math. Anal.*, 44(2):966–992, 2012.
- [18] C. Chainais-Hillairet. Entropy method and asymptotic behaviours of finite volume schemes. In *Finite volumes for complex applications. VII. Methods and theoretical aspects*, volume 77 of *Springer Proc. Math. Stat.*, pages 17–35. Springer, Cham, 2014.
- [19] C. Chainais-Hillairet, A. Jüngel, and Schuchnigg S. Entropy-dissipative discretization of nonlinear diffusion equations and discrete Beckner inequalities. HAL : hal-00924282, 2014.
- [20] P. G. Ciarlet. Basic error estimates for elliptic problems. Ciarlet, P. G. & Lions, J.-L. (ed.), in *Handbook of numerical analysis*. North-Holland, Amsterdam, pp. 17–351, 1991.
- [21] K. Deimling. *Nonlinear functional analysis*. Springer-Verlag, Berlin, 1985.
- [22] J. Dolbeault, B. Nazaret, and G. Savaré. A new class of transport distances between measures. *Calc. Var. Partial Differential Equations*, 34(2):193–231, 2009.
- [23] A. Ern and J.L. Guermond. *Theory and Practice of Finite Elements*, volume 159 of *Applied Mathematical Series*. Springer, New York, 2004.
- [24] R. Eymard, T. Gallouët, and R. Herbin. Finite volume methods. Ciarlet, P. G. (ed.) et al., in *Handbook of numerical analysis*. North-Holland, Amsterdam, pp. 713–1020, 2000.



- [25] R. Eymard, C. Guichard, and R. Herbin. Benchmark 3D: the VAG scheme. In J. Fořt, J. Fürst, J. Halama, R. Herbin, and F. Hubert, editors, *Finite Volumes for Complex Applications VI Problems & Perspectives*, volume 4 of *Springer Proceedings in Mathematics*, pages 1013–1022. Springer Berlin Heidelberg, 2011.
- [26] R. Eymard, C. Guichard, and R. Herbin. Small-stencil 3D schemes for diffusive flows in porous media. *ESAIM Math. Model. Numer. Anal.*, 46(2):265–290, 2012.
- [27] R. Eymard, C. Guichard, R. Herbin, and R. Masson. Vertex-centred discretization of multi-phase compositional Darcy flows on general meshes. *Comput. Geosci.*, 16(4):987–1005, 2012.
- [28] T. Gallouët and J.-C. Latché. Compactness of discrete approximate solutions to parabolic PDEs—application to a turbulence model. *Commun. Pure Appl. Anal.*, 11(6):2371–2391, 2012.
- [29] R. Herbin and F. Hubert. Benchmark on discretization schemes for anisotropic diffusion problems on general grids. In R. Eymard and J.-M. Herard, editors, *Finite Volumes for Complex Applications V*, pages 659–692. Wiley, 2008.
- [30] N. Igbida. Hele-Shaw type problems with dynamical boundary conditions. *J. Math. Anal. Appl.*, 335(2):1061–1078, 2007.
- [31] R. Jordan, D. Kinderlehrer, and F. Otto. The variational formulation of the Fokker-Planck equation. *SIAM J. Math. Anal.*, 29(1):1–17, 1998.
- [32] E. F. Keller and L. A. Segel. Model for chemotaxis. *Journal of Theoretical Biology*, 30(2):225–234, 1971.
- [33] J. Leray and J. Schauder. Topologie et équations fonctionnelles. *Ann. Sci. École Norm. Sup. (3)*, 51:45–78, 1934.
- [34] S. Lisini. Nonlinear diffusion equations with variable coefficients as gradient flows in Wasserstein spaces. *ESAIM Control Optim. Calc. Var.*, 15(3):712–740, 2009.
- [35] A. Mielke. A gradient structure for reaction-diffusion systems and for energy-drift-diffusion systems. *Nonlinearity*, 24(4):1329–1346, 2011.
- [36] F. Otto.  $L^1$ -contraction and uniqueness for quasilinear elliptic-parabolic equations. *J. Differential Equations*, 131:20–38, 1996.
- [37] F. Otto. The geometry of dissipative evolution equations: the porous medium equation. *Comm. Partial Differential Equations*, 26(1-2):101–174, 2001.
- [38] M. A. Peletier. Variational modelling: Energies, gradient flows, and large deviations. Lecture Notes, Würzburg. Available at <http://www.win.tue.nl/~mpeletie>, Feb. 2014.
- [39] J. Simon. Compact sets in the space  $L^p(0, T; B)$ . *Ann. Mat. Pura Appl. (4)*, 146:65–96, 1987.

CLÉMENT CANCES ([cances@ljl11.math.upmc.fr](mailto:cances@ljl11.math.upmc.fr))

- (1) SORBONNE UNIVERSITÉS, UPMC UNIV PARIS 06, UMR 7598, LABORATOIRE JACQUES-LOUIS LIONS, F-75005, PARIS, FRANCE
- (2) CNRS, UMR 7598, LABORATOIRE JACQUES-LOUIS LIONS, F-75005, PARIS, FRANCE

CINDY GUICHARD ([guichard@ljl11.math.upmc.fr](mailto:guichard@ljl11.math.upmc.fr))

- (1) SORBONNE UNIVERSITÉS, UPMC UNIV PARIS 06, UMR 7598, LABORATOIRE JACQUES-LOUIS LIONS, F-75005, PARIS, FRANCE
- (2) CNRS, UMR 7598, LABORATOIRE JACQUES-LOUIS LIONS, F-75005, PARIS, FRANCE
- (3) INRIA, ANGE PROJECT-TEAM, ROCQUENCOURT - B.P. 105, F78153 LE CHESNAY CEDEX, FRANCE
- (4) CEREMA, ANGE PROJECT-TEAM, 134 RUE DE BEAUVAIS, F-60280 MARGNY-LÈS-COMPIÈGNE, FRANCE