



HAL
open science

Convergence of numerical schemes for a conservation equation with convection and degenerate diffusion

Robert Eymard, Cindy Guichard, Xavier Lhébrard

► **To cite this version:**

Robert Eymard, Cindy Guichard, Xavier Lhébrard. Convergence of numerical schemes for a conservation equation with convection and degenerate diffusion. *Journal of Computational Mathematics -International Edition-*, 2020, 39, pp.428-452. 10.4208/jcm.2002-m2018-0287 . hal-01939144v2

HAL Id: hal-01939144

<https://hal.science/hal-01939144v2>

Submitted on 27 Aug 2019

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.

Convergence of numerical schemes for a conservation equation with convection and degenerate diffusion *

R. Eymard † C. Guichard ‡ Xavier Lhébrard †

August 27, 2019

Abstract

The approximation of problems with linear convection and degenerate nonlinear diffusion, which arise in the framework of the transport of energy in porous media with thermodynamic transitions, is done using a θ -scheme based on the centred gradient discretisation method. The convergence of the numerical scheme is proved, although the test functions which can be chosen are restricted by the weak regularity hypotheses on the convection field, owing to the application of a discrete Gronwall lemma and a general result for the time translate in the gradient discretisation setting. Some numerical examples, using both the Control Volume Finite Element method and the Vertex Approximate Gradient scheme, show the role of θ for stabilising the scheme.

Keywords: linear convection, degenerate diffusion, gradient discretisation method, θ -scheme.

1 Introduction

The development of geothermal energy leads to increasing needs for simulating the displacement of the water in a porous medium, accounting for the liquid-vapour change of phase [5]. This is achieved by writing the system of the conservation equation of the mass of water and that of the conservation of energy, together with a system of equations and inequalities expressing the thermodynamic equilibrium between the two phases when they are simultaneously present [10]. Let us consider a simplification of this system, which may be considered as a reasonable approximation in some physical cases:

$$\partial_t(\rho_l(1-S) + \rho_v S) - \operatorname{div} \left(\frac{K}{\mu} (\rho_l(1-S) + \rho_v S) \nabla P \right) = w, \quad (1)$$

$$\partial_t(e_l(1-S) + e_v S) - \operatorname{div} \left(\frac{K}{\mu} (e_l(1-S) + e_v S) \nabla P + \Lambda \nabla T \right) = f, \quad (2)$$

$$(T < T_e \text{ and } S = 0) \text{ or } (T = T_e \text{ and } 0 \leq S \leq 1) \text{ or } (T > T_e \text{ and } S = 1). \quad (3)$$

In the preceding system, the indices l, v respectively stand for the liquid and vapour phases, S is the saturation of the vapour phase (hence $1-S$ is that of the liquid phase), P the pressure assumed to be common for the two fluids (we neglect the capillary pressure), and for $\alpha = l, v$, ρ_α and e_α are respectively the density and the internal energy per mass unit of the phase α , assumed to be given functions of T . In System (2)-(3), the mobilities of the phases l and v are assumed to be equal to $(1-S)/\mu$ and S/μ , assuming the same viscosity μ for the two phases, and K is the absolute permeability field. The thermal conductivity is denoted by Λ . The right hand sides w and f are respectively the source terms of water and energy. The thermodynamic equilibrium between

*This work was supported by the French Agence Nationale de la Recherche (CHARMS project, ANR-16-CE06-0009).

†Université Paris-Est, Laboratoire d'Analyse et de Mathématiques Appliquées (UMR 8050), UPEM, UPEC, CNRS, F-77454, Marne-la-Vallée, France (Robert.Eymard@u-pem.fr), (xavier.lhebrard@u-pem.fr)

‡Sorbonne Université, Université Paris-Diderot SPC, CNRS, Inria, Laboratoire Jacques-Louis Lions, équipe ANGE, F-75005 Paris (cindy.guichard@sorbonne-universite.fr).

the two fluid phases l and v is assumed to hold when the temperature is equal to the equilibrium temperature T_e , assumed to be a constant; otherwise, one of the two fluid phases is missing.

Now denoting by $\bar{u} = e_l(1 - S) + e_v S$, we notice that, from (3), it is possible to express $T - T_e$ as a function ν of \bar{u} . For example, if $e_l = C_l(T - T_e)$ and $e_v = L + C_v(T - T_e)$, where L is the latent heat and C_α the thermal capacity of phase α , then there holds,

$$\nu(\bar{u}) = \begin{cases} \frac{\bar{u}}{C_l} & \bar{u} < 0, \\ 0 & 0 \leq \bar{u} \leq L, \\ \frac{\bar{u} - L}{C_v} & \bar{u} > L. \end{cases}$$

Therefore, denoting by $\vec{v} = -\frac{K}{\mu}\nabla P$ and only focusing on the energy conservation (we assume that the water conservation equation (1) is in some way decoupled from this problem), we consider the following linear convection – degenerate diffusion problem, issued from (2)-(3):

$$\partial_t \bar{u}(\mathbf{x}, t) + \operatorname{div}(\bar{u}(\mathbf{x}, t)\vec{v}(\mathbf{x}, t) - \Lambda(\mathbf{x})\nabla\nu(\bar{u}(\mathbf{x}, t))) = f(\mathbf{x}, t), \text{ for a.e. } (\mathbf{x}, t) \in \Omega \times (0, T), \quad (4a)$$

with the initial condition:

$$\bar{u}(\mathbf{x}, 0) = u_{\text{ini}}(\mathbf{x}), \text{ for a.e. } \mathbf{x} \in \Omega, \quad (4b)$$

together with the homogeneous Dirichlet boundary condition:

$$\nu(\bar{u}(\mathbf{x}, t)) = 0 \text{ on } \partial\Omega \times (0, T). \quad (4c)$$

In (4), we consider the following hypotheses.

- Ω is an open bounded connected polyhedral subset of \mathbb{R}^d , $d \in \mathbb{N}^*$ and $T > 0$ is now the final time, (5a)

- $u_{\text{ini}} \in L^2(\Omega)$, (5b)

- $\vec{v} \in L^\infty(\Omega \times (0, T))$, (5c)

- Λ is a measurable function from Ω to the set of $d \times d$ symmetric matrices and there exist $\underline{\lambda}, \bar{\lambda} > 0$ such that, for a.e. $\mathbf{x} \in \Omega$, $\Lambda(\mathbf{x})$ has eigenvalues in $[\underline{\lambda}, \bar{\lambda}]$, (5d)

- $f \in L^2(\Omega \times (0, T))$, (5e)

- $\nu \in C^0(\mathbb{R})$ is non-decreasing, Lipschitz continuous with constant \bar{L}_ν and s.t. $\nu(0) = 0$, and (5f)

- $|\nu(s)| \geq \underline{L}_\nu |s| - C_\nu$ for all $s \in \mathbb{R}$ for some given values $\underline{L}_\nu, C_\nu \in (0, +\infty)$. (5g)

We remark that (5f) and (5g) imply $0 < \underline{L}_\nu \leq \bar{L}_\nu$. Note that, inspired by the properties which can be expected from (1), Hypothesis (5c) only prescribes poor regularity properties for the velocity, that we only assume to be bounded, without regularity hypotheses on its derivatives. This weak regularity at least implies that Problem (4) must be considered in a weak sense (see (6) below). Hypotheses (5f)-(5g) on ν are classically satisfied in the framework described in the beginning of this section. The assumptions on Λ are taking into account heterogeneous and anisotropic porous media.

A function \bar{u} is said to be a weak solution of Problem (4) if the following holds:

$$\begin{aligned} & \bar{u} \in L^2(\Omega \times (0, T)), \nu(\bar{u}) \in L^2(0, T; H_0^1(\Omega)), \forall \varphi \in C_c^\infty(\Omega \times [0, T]), \\ & - \int_0^T \int_\Omega \bar{u}(\mathbf{x}, t) \partial_t \varphi(\mathbf{x}, t) d\mathbf{x} dt - \int_\Omega u_{\text{ini}}(\mathbf{x}) \varphi(\mathbf{x}, 0) d\mathbf{x} \\ & + \int_0^T \int_\Omega (\Lambda(\mathbf{x})\nabla\nu(\bar{u})(\mathbf{x}, t) - \bar{u}(\mathbf{x}, t)\vec{v}(\mathbf{x}, t)) \cdot \nabla\varphi(\mathbf{x}, t) d\mathbf{x} dt = \int_0^T \int_\Omega f(\mathbf{x}, t) \varphi(\mathbf{x}, t) d\mathbf{x} dt, \end{aligned} \quad (6)$$

where we denote by $C_c^\infty(\Omega \times [0, T])$ the set of the restrictions of functions of $C_c^\infty(\Omega \times (-\infty, T))$ to $\Omega \times [0, T]$.

Let us first comment the question of the existence and uniqueness of a solution to (6), which enters into the class of the degenerate parabolic equations of the Stefan type, with a linear convection term, but with an irregular velocity, since we only assume $\vec{v} \in L^\infty(\Omega \times (0, T))$. For a regular velocity, and using a weak formulation based on Krushkov entropy pairs which implies (6), the existence and uniqueness of a solution is proved in [8], and, for the triply degenerate problem in [4].

Here, although the convection term is linear, there can exist infinitely many solutions to (6), due to the fact that the test functions are assumed to vanish at the boundary. Indeed, the case where ν is constant on an interval is included in our study, and therefore (6) includes the case of linear scalar hyperbolic equations whose solutions are valued in such an interval: assume that $\Omega = (0, 1)$, $\vec{v} = 1$, $\nu(s) = 0$ for all $s \in [0, 1]$, $u_{\text{ini}} = 0$, $f = 0$. Then, for any $\alpha \in [0, 1]$, the function $\bar{u}(x, t) = \alpha$ for $x < t$ and $\bar{u}(x, t) = 0$ for $x > t$ is solution to (6) since it verifies $\partial_t \bar{u} + \partial_x \bar{u} = 0$ in the distribution sense. On the contrary, if the velocity is regular and if $\nu^{-1}(\{0\}) = \{0\}$ or $\vec{v} \cdot \mathbf{n} \geq 0$ a.e. on $\partial\Omega \times (0, T)$ (homogeneous Neumann boundary conditions are often chosen in the framework of fluid flows in porous media), one can expect that the solution to (6) is unique. But, only assuming (5c), the lack of regularity of \vec{v} prevents from applying the existence and uniqueness results of the literature. In [11], existence is proved with very low regularity for \vec{v} , but uniqueness only holds if $\vec{v} \in W^{1,1}$ and $\text{div} \vec{v} \in L^\infty$, and in [2], the preceding results are extended to the case where $\vec{v} \in BV$, and $\text{div} \vec{v} \in L^\infty$.

Hence in this paper, we only provide an existence result (obtained here through the convergence of a numerical scheme). Recall that, in the case where ν is affine, (6) enters into the framework of linear and quasilinear second order differential equations, for which there are numerous numerical approximations (see for example [23, 9] for approximations using mixed finite element methods in the quasilinear case). But the case where ν is no longer affine leads to additional difficulties, some of which have been solved only after the seminal paper by Alt and Luckhaus [1] (as recalled below). Without a convection term, the approximation of this problem by conforming methods is studied in [3, 13, 14, 24]. In the case of a convection term with a regular velocity, see [26] for an approach based on monotone graphs, and refer to [7] for a numerical approach leading to an entropy diminishing scheme.

Let us now describe the specific difficulties for the convergence analysis of the numerical approximation, that are issued from the linear convection term with irregular velocity and a non-affine function ν .

We follow in this paper the same mathematical steps for this proof as the ones which would be considered in the study of the convergence of u_ε , solution in a weak sense to the problem

$$\partial_t u_\varepsilon + \text{div}(u_\varepsilon \vec{v}(\mathbf{x}, t) - \Lambda \nabla \nu(u_\varepsilon)) = f_\varepsilon, \text{ in } \Omega \times (0, T),$$

together with the same initial and boundary conditions as those included in Problem (4), assuming that f_ε converges to f in $L^2(\Omega \times (0, T))$ as $\varepsilon \rightarrow 0$. In the case where the velocity \vec{v} is sufficiently regular (for example, $\vec{v} \in H^1(\Omega)^d$, or even more regular as in [8]), one can multiply the equation by u_ε , in order to get an estimate on u_ε^2 .

Assuming that $u_\varepsilon = 0$ holds on the part $\partial\Omega^-$ of $\partial\Omega$ where $\vec{v} \cdot \mathbf{n} < 0$ holds, we can apply the calculation chain

$$\int_{\Omega} u_\varepsilon \text{div}(u_\varepsilon \vec{v}) d\mathbf{x} = \int_{\Omega} \frac{u_\varepsilon^2}{2} \text{div} \vec{v} d\mathbf{x} + \int_{\partial\Omega^-} \frac{u_\varepsilon^2}{2} \vec{v} \cdot \mathbf{n} ds(\mathbf{x}) + \int_{\partial\Omega \setminus \partial\Omega^-} \frac{u_\varepsilon^2}{2} \vec{v} \cdot \mathbf{n} ds(\mathbf{x}) \geq \int_{\Omega} \frac{u_\varepsilon^2}{2} \text{div} \vec{v} d\mathbf{x},$$

and an estimate on $\int_{\Omega} u_\varepsilon^2 d\mathbf{x}$ can be obtained.

In the present situation, due to the less regular hypothesis (5c), the above calculation chain no longer holds, and this choice for the test function does not lead to any estimate. Moreover, the only test function that we can take at the discrete level, even in the case of more regular \vec{v} , is $\nu(u_\varepsilon)$, using the equation which provides a control on $\nabla \nu(u_\varepsilon)$ in $L^2(\Omega \times (0, T))^d$. Indeed, our aim is to approximate Problem (6) in some framework which includes, in addition to conforming finite elements, non-conforming methods such as mixed finite element methods, or discontinuous Galerkin methods, and other more recent methods [12]. In such a general framework, the continuous gradient operator is replaced by a discrete one, denoted by $\nabla_{\mathcal{D}}$. Then Stampacchia's result [25], which allows to write in the continuous case $\Lambda \nabla \nu(u_\varepsilon) \cdot \nabla \eta(u_\varepsilon) = \Lambda \nabla \varphi(u_\varepsilon) \cdot \nabla \varphi(u_\varepsilon)$ with $(\varphi')^2 = \nu' \eta'$, does not hold in the discrete framework (note that Stampacchia's result has a discrete counterpart if the scheme is based on the two-point flux approximation, but in this case, the meshes are restricted, and Λ should be isotropic, see the discussion in [17]).

Let us sketch the steps of the convergence proof, only using the test function $\nu(u_\varepsilon)$. We introduce the function $\xi(s) = \int_0^s \nu(x) dx$ (as in (15)), and we obtain, after integration on $\Omega \times (0, T)$, and an integration by parts,

$$\int_{\Omega} \xi(u_\varepsilon(\mathbf{x}, T)) d\mathbf{x} - \int_{\Omega} \xi(u_{\text{ini}}) d\mathbf{x} + \int_0^T \int_{\Omega} (\Lambda \nabla \nu(u_\varepsilon) - u_\varepsilon \vec{\nu}(\mathbf{x}, t)) \cdot \nabla \nu(u_\varepsilon) d\mathbf{x} dt = \int_0^T \int_{\Omega} f_\varepsilon \nu(u_\varepsilon) d\mathbf{x} dt. \quad (7)$$

Owing to Hypothesis (5g), we can show that $\int_{\Omega} \xi(u_\varepsilon(\mathbf{x}, T)) d\mathbf{x} \geq C(\|u_\varepsilon(\cdot, T)\|_{L^2(\Omega)}^2 - 1)$ for some constant C . We then write, using the Young inequality,

$$\left| \int_0^T \int_{\Omega} u_\varepsilon \vec{\nu}(\mathbf{x}, t) \cdot \nabla \nu(u_\varepsilon) d\mathbf{x} dt \right| \leq \alpha \int_0^T \|u_\varepsilon(\cdot, t)\|_{L^2(\Omega)}^2 dt + \beta \|\nabla \nu(u_\varepsilon)\|_{L^2(\Omega \times (0, T))}^2,$$

with β small enough for controlling the second term by $\int_0^T \int_{\Omega} \Lambda \nabla \nu(u_\varepsilon) \cdot \nabla \nu(u_\varepsilon) d\mathbf{x} dt$. We can then use the Cauchy-Schwarz and the Poincaré inequalities for handling the right hand side of (7). This leads to an equation under the form

$$\|u_\varepsilon(\cdot, T)\|_{L^2(\Omega)}^2 \leq A \int_0^T \|u_\varepsilon(\cdot, t)\|_{L^2(\Omega)}^2 dt + B,$$

allowing for the application of Gronwall's lemma. Gathering these results, we obtain the following estimates:

(E1) an estimate on $\nu(u_\varepsilon)$ in $L^2(0, T; H_0^1(\Omega))$,

(E2) an estimate on u_ε in $L^\infty(0, T; L^2(\Omega))$,

(E3) an estimate on $\partial_t u_\varepsilon$ in $L^2(0, T; H^{-1}(\Omega))$.

Once these estimates are proved, it only remains to prove that, under the extraction of a subsequence from a sequence of approximate solutions, u_ε converges to a weak solution of the problem. One remarkable idea in [1] is the use of the following result:

$$\|\nu(u_\varepsilon(\cdot, \cdot + \tau)) - \nu(u_\varepsilon)\|_{L^2(\Omega \times (0, T - \tau))}^2 \leq \tau 2\bar{L}_\nu \|\partial_t u_\varepsilon\|_{L^2(0, T; H^{-1}(\Omega))} \|\nu(u_\varepsilon)\|_{L^2(0, T; H_0^1(\Omega))}$$

(we give a discrete equivalent of this result in Lemma A.1). This result, in addition to (E1) and (E3), allows to apply Kolmogorov's theorem, and therefore to extract a sequence $(\nu(u_{\varepsilon_m}))_{m \in \mathbb{N}}$ which converges to some function χ in $L^2(\Omega \times (0, T))$. Estimate (E2) allows to extract a subsequence from the preceding one such that there exists $\bar{u} \in L^\infty(0, T; L^2(\Omega))$ with $(u_{\varepsilon_m})_{m \in \mathbb{N}}$ weakly converges to \bar{u} . Then, owing to the monotonicity of ν , Minty's trick provides that $\chi = \nu(\bar{u})$.

In the convergence part of this paper, we therefore derive discrete equivalents of (E1)-(E2)-(E3), from similar computations only resulting from the multiplication of the discrete scheme by $\nu(u_\varepsilon)$.

We emphasise that the scheme which is considered below includes a parameter θ , such that, if $\theta = 0$, the convection term is explicit, if $\theta = 1/2$, the convection term is centred in time, and if $\theta = 1$, the convection term is implicit. We consider in this paper the case $\theta \in \mathbb{R}$ since we show in the numerical examples that values $\theta > 1$ lead to a kind of stabilisation in the case where oscillations occur for $\theta \in [0, 1]$. In all cases, the degenerate diffusion term is taken implicit. We recall that, for a centred finite volume scheme for the convection, the θ -scheme is L^2 -stable only if $\theta \geq 1/2$. We prove below that this limitation does not apply in the framework of this paper: for any $\theta \in \mathbb{R}$, the degenerate diffusion is sufficient for leading to weakly convergent schemes and in some particular cases to a strong convergence property.

This paper is organised as follows. We first apply the gradient discretisation tools to the continuous Problem in Section 2, and derive some estimates, which are used in Section 3 for the convergence analysis. Finally in Section 4, numerical examples show the behaviour of the Control Volume Finite Element scheme (CVFE) and the Vertex Approximate Gradient (VAG) scheme [18] which present some interesting characteristics for coupled flows in porous media.

2 Approximation by space-time gradient discretisations

In the same manner as in [16], the discretisation of Problem (6) is obtained by simply replacing the continuous operators by their discrete equivalent, following the Gradient Discretisation Method [12]. Let $\mathcal{D}_T := (X_{\mathcal{D},0}, \Pi_{\mathcal{D}}, \nabla_{\mathcal{D}}, \mathcal{I}_{\mathcal{D}}, (t^{(n)})_{n=0,\dots,N})$ be a space-time discretisation in the sense of [12, Definition 4.1], such that $\mathcal{D} := (X_{\mathcal{D},0}, \Pi_{\mathcal{D}}, \nabla_{\mathcal{D}})$ is a space gradient discretisation for Dirichlet boundary conditions in the sense of [12, Definition 2.1] (we then denote by $\delta^{(n+\frac{1}{2})} := t^{(n+1)} - t^{(n)}$ and $\delta_{\mathcal{D}} = \max_{n=0,\dots,N-1} \delta^{(n+\frac{1}{2})}$). Then, [12, Definition 2.1] specifies that $\|\cdot\|_{\mathcal{D}} := \|\nabla_{\mathcal{D}} \cdot\|_{L^2(\Omega)^d}$ is a norm on $X_{\mathcal{D},0}$, and the following quantities: $C_{\mathcal{D}} \in [0, +\infty)$, $S_{\mathcal{D}} : H_0^1(\Omega) \rightarrow [0, +\infty)$ and $W_{\mathcal{D}} : H_{\text{div}}(\Omega) \rightarrow [0, +\infty)$, are defined by

$$C_{\mathcal{D}} = \max_{v \in X_{\mathcal{D},0} \setminus \{0\}} \frac{\|\Pi_{\mathcal{D}} v\|_{L^2(\Omega)}}{\|v\|_{\mathcal{D}}}, \quad (8)$$

$$\forall \varphi \in H_0^1(\Omega), \quad S_{\mathcal{D}}(\varphi) = \min_{v \in X_{\mathcal{D},0}} (\|\Pi_{\mathcal{D}} v - \varphi\|_{L^2(\Omega)} + \|\nabla_{\mathcal{D}} v - \nabla \varphi\|_{L^2(\Omega)^d}), \quad (9)$$

$$\forall \varphi \in H_{\text{div}}(\Omega), \quad W_{\mathcal{D}}(\varphi) = \max_{u \in X_{\mathcal{D},0} \setminus \{0\}} \frac{1}{\|u\|_{\mathcal{D}}} \left| \int_{\Omega} (\nabla_{\mathcal{D}} u(\mathbf{x}) \cdot \varphi(\mathbf{x}) + \Pi_{\mathcal{D}} u(\mathbf{x}) \text{div} \varphi(\mathbf{x})) \, d\mathbf{x} \right|. \quad (10)$$

These quantities are then used in the following definitions.

Definition 2.1 (coercivity, GD-consistency, limit-conformity and compactness). *Under Hypotheses (5), let $((\mathcal{D}_T)_m)_{m \in \mathbb{N}}$ be a sequence of space-time gradient discretisations. The sequence $((\mathcal{D}_T)_m)_{m \in \mathbb{N}}$ is said to be:*

- coercive if there exists $C_P > 0$ such that $C_{\mathcal{D}_m} \leq C_P$ for all $m \in \mathbb{N}$ (see (8)).
- GD-consistent if, for all $\varphi \in H_0^1(\Omega)$, $\lim_{m \in \mathbb{N}} S_{\mathcal{D}_m}(\varphi) = 0$ (see (9)), and in addition: for all $u \in L^2(\Omega)$, $\lim_{m \rightarrow \infty} \|u - \Pi_{\mathcal{D}_m} \mathcal{I}_{\mathcal{D}_m} u\|_{L^2(\Omega)} = 0$ and $\lim_{m \rightarrow \infty} \delta_{\mathcal{D}_m} = 0$.
- limit-conforming if, for all $\varphi \in H_{\text{div}}(\Omega)$, $\lim_{m \in \mathbb{N}} W_{\mathcal{D}_m}(\varphi) = 0$ (see (10)).
- compact if, for any $(v_m)_{m \in \mathbb{N}}$ such that $v_m \in X_{\mathcal{D}_m,0}$ for all $m \in \mathbb{N}$ and $(\nabla_{\mathcal{D}_m} v_m)_{m \in \mathbb{N}}$ is bounded in $L^2(\Omega)^d$, the set $\{\Pi_{\mathcal{D}_m} v_m : m \in \mathbb{N}\}$ is relatively compact in $L^2(\Omega)$.

We assume that $\Pi_{\mathcal{D}}$ is a piecewise constant function reconstruction in the sense of the following definition:

Definition 2.2 (Piecewise constant reconstruction and nonlinear functions of the elements of $X_{\mathcal{D},0}$). *We say that the reconstruction $\Pi_{\mathcal{D}}$ is piecewise constant if there exists a finite set I and a family of disjoint subsets $(\Omega_i)_{i \in I}$ of Ω (some of the Ω_i can be empty) such that $X_{\mathcal{D},0} = \mathbb{R}^I$ and*

$$\forall v = (v_i)_{i \in I} \in X_{\mathcal{D},0}, \quad \Pi_{\mathcal{D}} v = \sum_{i \in I} v_i \mathbf{1}_{\Omega_i}, \quad (11)$$

where $\mathbf{1}_{\Omega_i}$ is the characteristic function of Ω_i . In other words, $(\Pi_{\mathcal{D}} v)|_{\Omega_i} = v_i$ for all $i \in I$.

For any $g \in C^0(\mathbb{R})$ and $v = (v_i)_{i \in I} \in X_{\mathcal{D},0}$, denoting $g(v) \in X_{\mathcal{D},0}$ the element $g(v) := (g(v_i))_{i \in I} \in X_{\mathcal{D},0}$, we then have

$$\text{For a.e. } \mathbf{x} \in \Omega, \quad g(\Pi_{\mathcal{D}} v(\mathbf{x})) = \Pi_{\mathcal{D}} g(v)(\mathbf{x}). \quad (12)$$

We introduce the following notations for the definition of discrete space-time dependent functions:

$$\begin{aligned} & \text{for a.e. } \mathbf{x} \in \Omega, \forall \theta \in \mathbb{R}, \forall v := (v^{(n)})_{n=0,\dots,N} \in X_{\mathcal{D},0}^{N+1} : \\ & \Pi_{\mathcal{D}}^{(\theta)} v(\mathbf{x}, 0) = \Pi_{\mathcal{D}} v^{(0)}(\mathbf{x}), \\ & \forall t \in (t^{(n)}, t^{(n+1)}], \forall n = 0, \dots, N-1 : \\ & \delta_{\mathcal{D}}^{(n+\frac{1}{2})} v(\mathbf{x}) = \Pi_{\mathcal{D}} \frac{v^{(n+1)} - v^{(n)}}{\delta^{(n+\frac{1}{2})}}(\mathbf{x}), \quad \delta_{\mathcal{D}} v(\mathbf{x}, t) = \delta_{\mathcal{D}}^{(n+\frac{1}{2})} v(\mathbf{x}), \\ & v^{(n+\theta)} = \theta v^{(n+1)} + (1-\theta) v^{(n)}, \quad \Pi_{\mathcal{D}}^{(\theta)} v(\mathbf{x}, t) = \Pi_{\mathcal{D}} v^{(n+\theta)}(\mathbf{x}), \quad \nabla_{\mathcal{D}}^{(\theta)} v(\mathbf{x}, t) = \nabla_{\mathcal{D}} v^{(n+\theta)}(\mathbf{x}). \end{aligned} \quad (13)$$

Note that letting $\theta = 1$ in (13) provides implicit expressions.

The approximate solution $u := (u^{(n)})_{n=0,\dots,N}$ is computed through the following θ -scheme, for a given $\theta \in \mathbb{R}$.

$$\left\{ \begin{array}{l} u^{(0)} = \mathcal{I}_{\mathcal{D}} u_{\text{ini}} \in X_{\mathcal{D},0} \text{ and for all } n = 0, \dots, N-1 : \\ u^{(n+1)} \in X_{\mathcal{D},0}, \\ \int_{\Omega} \left(\delta_{\mathcal{D}}^{(n+\frac{1}{2})} u(\mathbf{x}) \Pi_{\mathcal{D}} v(\mathbf{x}) - \Pi_{\mathcal{D}} u^{(n+\theta)}(\mathbf{x}) \bar{\mathbf{v}}(\mathbf{x}, t) \cdot \nabla_{\mathcal{D}} v(\mathbf{x}) \right. \\ \quad \left. + \Lambda(\mathbf{x}) \nabla_{\mathcal{D}} \nu(u^{(n+1)})(\mathbf{x}) \cdot \nabla_{\mathcal{D}} v(\mathbf{x}) \right) d\mathbf{x} \\ = \frac{1}{\delta_{\mathcal{D}}^{(n+\frac{1}{2})}} \int_{t^{(n)}}^{t^{(n+1)}} \int_{\Omega} f(\mathbf{x}, t) \Pi_{\mathcal{D}} v(\mathbf{x}) d\mathbf{x} dt, \quad \forall v \in X_{\mathcal{D},0}, \forall n = 0, \dots, N-1. \end{array} \right. \quad (14)$$

Let us observe that the time discretisation of the degenerate diffusion term is implicit, whereas a θ -scheme is used for the convection term. We finally introduce the function

$$\xi(s) = \int_0^s \nu(x) dx, \quad \forall s \in \mathbb{R}. \quad (15)$$

We then have

$$\xi(s) = \int_0^s \nu(x) dx = \int_0^s (\nu(x) - \nu(0)) dx \leq \bar{L}_{\nu} \int_0^s x dx = \bar{L}_{\nu} \frac{s^2}{2}, \quad \forall s \in \mathbb{R}, \quad (16)$$

and, from Hypotheses (5f) and (5g), and using Young's inequality,

$$\xi(s) \geq \int_0^{|s|} (\underline{L}_{\nu} x - C_{\nu}) dx = \frac{1}{2} \underline{L}_{\nu} |s|^2 - C_{\nu} |s| \geq \frac{1}{4} \underline{L}_{\nu} s^2 - \frac{C_{\nu}^2}{\underline{L}_{\nu}}, \quad \forall s \in \mathbb{R}. \quad (17)$$

Lemma 2.3 (Discrete versions of (E1) and (E2) and existence of a discrete solution). *Under Hypotheses (5), let $\mathcal{D}_T = (X_{\mathcal{D},0}, \Pi_{\mathcal{D}}, \nabla_{\mathcal{D}}, \mathcal{I}_{\mathcal{D}}, (t^{(n)})_{n=0,\dots,N})$ be a space-time gradient discretisation such that $\Pi_{\mathcal{D}}$ is a piecewise constant function reconstruction (see Definition 2.2). Let $\gamma \in (0, 1)$ be given and let $\theta \in \mathbb{R}$ be such that*

$$4 \|\bar{\mathbf{v}}\|_{L^{\infty}(\Omega \times (0, T))}^2 \theta^2 \delta_{\mathcal{D}} \leq \gamma \underline{L}_{\nu} \lambda. \quad (18)$$

Let $C_{\text{ini}} > 0$ be such that

$$C_{\text{ini}} > \|u_{\text{ini}} - \Pi_{\mathcal{D}} u^{(0)}\|_{L^2(\Omega)}, \quad (19)$$

and let $C_P > 0$ be such that

$$C_P > C_{\mathcal{D}}. \quad (20)$$

Then there exists at least one solution to Scheme (14), and there exists $C_1 > 0$, only depending on \bar{L}_{ν} , \underline{L}_{ν} , C_{ν} , C_P , C_{ini} , f , $\bar{\mathbf{v}}$, $\underline{\lambda}$, θ and γ such that, for any solution u to this scheme,

$$\|\Pi_{\mathcal{D}}^{(1)} \nu(u)\|_{L^{\infty}(0, T; L^2(\Omega))} \leq C_1, \quad \|\Pi_{\mathcal{D}}^{(1)} u\|_{L^{\infty}(0, T; L^2(\Omega))} \leq C_1, \quad \text{and} \quad \|\Pi_{\mathcal{D}}^{(\theta)} u\|_{L^{\infty}(0, T; L^2(\Omega))} \leq C_1, \quad (21)$$

and

$$\|\nabla_{\mathcal{D}}^{(1)} \nu(u)\|_{L^2(\Omega \times (0, T))} \leq C_1. \quad (22)$$

Remark 2.4 (On condition (18)). *For any consistent sequence $((\mathcal{D}_T)_m)_{m \in \mathbb{N}}$ of space-time gradient discretisations, condition (18) is necessarily satisfied for m large enough for any $\theta \in \mathbb{R}$, since the consistency property implies that $\delta_{\mathcal{D}_m}$ tends to 0 as $m \rightarrow \infty$. For a given space-time gradient discretisation, it is always possible to choose $\theta \in \mathbb{R}$ such that condition (18) holds (see Section 4 for an example of numerical maximum value for $|\theta|$ such that this condition holds).*

Proof. Before showing the existence of at least one discrete solution to Scheme (14), let us first prove (21) and (22). From properties (16) and (17), and using $\int_a^b \nu(s) ds = \xi(b) - \xi(a) = \nu(b)(b-a) - \int_a^b \nu'(s)(s-a) ds$, we

get, since (5f) implies $\nu'(s) \geq 0$ and using the piecewise constant reconstruction hypothesis which implies that (12) holds, that

$$(\Pi_{\mathcal{D}}u^{(n+1)}(\mathbf{x}) - \Pi_{\mathcal{D}}u^{(n)}(\mathbf{x})) \Pi_{\mathcal{D}}\nu(u^{(n+1)})(\mathbf{x}) \geq \Pi_{\mathcal{D}}\xi(u^{(n+1)})(\mathbf{x}) - \Pi_{\mathcal{D}}\xi(u^{(n)})(\mathbf{x}) \text{ for a.e. } \mathbf{x} \in \Omega.$$

We then let $v = \delta^{\alpha(n+\frac{1}{2})}\nu(u^{(n+1)})$ in (14), and we sum the obtained equation on $n = 0, \dots, m-1$ for a given $m = 1, \dots, N$. Accounting for the above inequality yields that there holds

$$\begin{aligned} & \int_{\Omega} (\xi(\Pi_{\mathcal{D}}u^{(m)}(\mathbf{x})) - \xi(\Pi_{\mathcal{D}}u^{(0)}(\mathbf{x})))d\mathbf{x} + \lambda \int_0^{t^{(m)}} \int_{\Omega} |\nabla_{\mathcal{D}}^{(1)}\nu(u)(\mathbf{x}, t)|^2 d\mathbf{x}dt \\ & \leq \int_0^{t^{(m)}} \int_{\Omega} \left(\Pi_{\mathcal{D}}^{(\theta)}u(\mathbf{x}, t)\vec{\mathbf{v}}(\mathbf{x}, t) \cdot \nabla_{\mathcal{D}}^{(1)}\nu(u)(\mathbf{x}, t) + f(\mathbf{x}, t)\Pi_{\mathcal{D}}^{(1)}\nu(u)(\mathbf{x}, t) \right) d\mathbf{x}dt, \quad \forall m = 0, \dots, N, \end{aligned} \quad (23)$$

since for $m = 0$ the above inequality reduces to $0 \leq 0$. Thanks to Young's inequality and applying the definition of the coercivity constant (8), we get that, for any $n = 0, \dots, m-1$, for every $\eta_1 > 0$ and a.e. $(\mathbf{x}, t) \in \Omega \times (t^{(n)}, t^{(n+1)})$,

$$\begin{aligned} f(\mathbf{x}, t) \Pi_{\mathcal{D}}\nu(u^{(n+1)})(\mathbf{x}) & \leq \frac{1}{2\eta_1} |f(\mathbf{x}, t)|^2 + \frac{\eta_1}{2} |\Pi_{\mathcal{D}}\nu(u^{(n+1)})(\mathbf{x})|^2 \\ & \leq \frac{1}{2\eta_1} |f(\mathbf{x}, t)|^2 + \frac{C_{\mathcal{D}}^2 \eta_1}{2} |\nabla_{\mathcal{D}}\nu(u^{(n+1)})(\mathbf{x})|^2. \end{aligned} \quad (24)$$

Summing two Young's inequalities applied to both terms $\theta \Pi_{\mathcal{D}}u^{(n+1)}(\mathbf{x})\vec{\mathbf{v}}(\mathbf{x}, t) \cdot \nabla_{\mathcal{D}}\nu(u^{(n+1)})(\mathbf{x})$ and $(1-\theta) \Pi_{\mathcal{D}}u^{(n)}(\mathbf{x})\vec{\mathbf{v}}(\mathbf{x}, t) \cdot \nabla_{\mathcal{D}}\nu(u^{(n+1)})(\mathbf{x})$, we get that for every $\eta_2 > 0$ and a.e. $(\mathbf{x}, t) \in \Omega \times (t^{(n)}, t^{(n+1)})$,

$$\begin{aligned} & |\Pi_{\mathcal{D}}u^{(n+\theta)}(\mathbf{x})\vec{\mathbf{v}}(\mathbf{x}, t) \cdot \nabla_{\mathcal{D}}\nu(u^{(n+1)})(\mathbf{x})| \\ & \leq \|\vec{\mathbf{v}}\|_{L^\infty(\Omega \times (0, T))} \left(\frac{1}{2\eta_2} \left(\theta^2 (\Pi_{\mathcal{D}}u^{(n+1)}(\mathbf{x}))^2 + (1-\theta)^2 (\Pi_{\mathcal{D}}u^{(n)}(\mathbf{x}))^2 \right) + \eta_2 |\nabla_{\mathcal{D}}\nu(u^{(n+1)})(\mathbf{x})|^2 \right). \end{aligned} \quad (25)$$

Thus from (23), using (24) with $\eta_1 = \lambda/(2C_{\mathcal{D}}^2)$ and (25) with $\eta_2 = \lambda/(2\|\vec{\mathbf{v}}\|_{L^\infty(\Omega \times (0, T))})$, we get that, for all $m = 0, \dots, N$ (with the convention that an empty sum is equal to 0),

$$\begin{aligned} & \|\Pi_{\mathcal{D}}\xi(u^{(m)})\|_{L^1(\Omega)} + \frac{1}{4}\lambda \|\nabla_{\mathcal{D}}^{(1)}\nu(u)\|_{L^2(\Omega \times (0, t^{(m)}))}^2 \\ & \leq \frac{\|\vec{\mathbf{v}}\|_{L^\infty(\Omega \times (0, T))}^2}{\lambda} \sum_{n=0}^{m-1} \delta^{\alpha(n+\frac{1}{2})} \left(\theta^2 \|\Pi_{\mathcal{D}}u^{(n+1)}\|_{L^2(\Omega)}^2 + (1-\theta)^2 \|\Pi_{\mathcal{D}}u^{(n)}\|_{L^2(\Omega)}^2 \right) \\ & \quad + \frac{C_{\mathcal{D}}^2}{\lambda} \|f\|_{L^2(\Omega \times (0, t^{(m)}))}^2 + \|\Pi_{\mathcal{D}}\xi(u^{(0)})\|_{L^1(\Omega)}. \end{aligned} \quad (26)$$

This in turn yields, thanks to (16) and (17) and using $\frac{1}{4}\lambda \|\nabla_{\mathcal{D}}^{(1)}\nu(u)\|_{L^2(\Omega \times (0, t^{(m)}))}^2 \geq 0$,

$$\begin{aligned} & \left(\frac{\bar{L}_\nu}{4} - \frac{\|\vec{\mathbf{v}}\|_{L^\infty(\Omega \times (0, T))}^2}{\lambda} \delta^{\alpha(m-\frac{1}{2})} \theta^2 \right) \|\Pi_{\mathcal{D}}u^{(m)}\|_{L^2(\Omega)}^2 \\ & \leq \frac{\|\vec{\mathbf{v}}\|_{L^\infty(\Omega \times (0, T))}^2}{\lambda} \sum_{n=1}^{m-1} \left(\theta^2 \delta^{\alpha(n-\frac{1}{2})} + (1-\theta)^2 \delta^{\alpha(n+\frac{1}{2})} \right) \|\Pi_{\mathcal{D}}u^{(n)}\|_{L^2(\Omega)}^2 \\ & \quad + \frac{C_{\mathcal{D}}^2}{\lambda} \|f\|_{L^2(\Omega \times (0, T))}^2 + \left(\frac{\|\vec{\mathbf{v}}\|_{L^\infty(\Omega \times (0, T))}^2}{\lambda} (1-\theta)^2 \delta^{\alpha(\frac{1}{2})} + \frac{\bar{L}_\nu}{2} \right) \|\Pi_{\mathcal{D}}u^{(0)}\|_{L^2(\Omega)}^2 + \frac{C_\nu^2}{\bar{L}_\nu}, \quad \forall m = 1, \dots, N. \end{aligned} \quad (27)$$

Let us notice that (18) has been designed in order that the coefficient of $\|\Pi_{\mathcal{D}}u^{(m)}\|_{L^2(\Omega)}^2$ at the left hand side remains strictly positive. Indeed, (18) is equivalent to

$$\frac{\underline{L}_\nu}{4} - \theta^2 \frac{\|\vec{v}\|_{L^\infty(\Omega \times (0, T))}^2}{\lambda} \delta_{\mathcal{D}} > \frac{\underline{L}_\nu(1 - \gamma)}{4}.$$

Since, owing to (19), we have

$$\|\Pi_{\mathcal{D}}u^{(0)}\|_{L^2(\Omega)} \leq \|u_{\text{ini}}\|_{L^2(\Omega)} + C_{\text{ini}}, \quad (28)$$

we obtain, denoting $a_m = \|\Pi_{\mathcal{D}}u^{(m)}\|_{L^2(\Omega)}^2$ for all $m = 1, \dots, N$, that

$$\forall m \in \{1, \dots, N\}, \quad a_m \leq \sum_{n=1}^{m-1} b_n a_n + B, \quad (29)$$

with the convention that an empty sum is equal to zero, and denoting by

$$B = \frac{4}{\underline{L}_\nu(1 - \gamma)} \left(\frac{C_{\mathcal{D}}^2}{\lambda} \|f\|_{L^2(\Omega \times (0, T))}^2 + \left(\frac{\bar{L}_\nu}{2} + (1 - \theta)^2 T \frac{\|\vec{v}\|_{L^\infty(\Omega \times (0, T))}^2}{\lambda} \right) (\|u_{\text{ini}}\|_{L^2(\Omega)} + C_{\text{ini}})^2 + \frac{C_\nu^2}{\underline{L}_\nu} \right),$$

and

$$b_n = A(\theta^2 \delta^{(n-\frac{1}{2})} + (1 - \theta)^2 \delta^{(n+\frac{1}{2})}) \quad \text{with} \quad A = \frac{4\|\vec{v}\|_{L^\infty(\Omega \times (0, T))}^2}{(1 - \gamma)\underline{L}_\nu\lambda}.$$

Using that $\sum_{n=1}^m \delta^{(n-\frac{1}{2})} = t^{(m)}$, we get that $\sum_{n=1}^{m-1} b_n \leq AT(\theta^2 + (1 - \theta)^2)$, which, introduced into the discrete Gronwall inequality

$$\forall m \in \{1, \dots, N\}, \quad a_m \leq B \exp\left(\sum_{n=1}^{m-1} b_n\right), \quad (30)$$

provides

$$\|\Pi_{\mathcal{D}}u^{(m)}\|_{L^2(\Omega)}^2 \leq B \exp(AT(\theta^2 + (1 - \theta)^2)), \quad \forall m = 1, \dots, N.$$

This concludes the proof of (21), and, reporting this estimate in (26), allows to deduce (22).

For the sake of completeness, let us give the short proof of (30). We define $v_m = \sum_{n=1}^{m-1} b_n a_n$, and we prove by induction that $v_m \leq B \exp\left(\sum_{n=1}^{m-1} b_n\right) - B$. Since we have

$$v_{m+1} - v_m = b_m a_m \leq b_m (v_m + B) \leq b_m B \exp\left(\sum_{n=1}^{m-1} b_n\right),$$

by applying (29) and the induction hypothesis (second inequality), we get, using again the induction hypothesis,

$$v_{m+1} \leq B(1 + b_m) \exp\left(\sum_{n=1}^{m-1} b_n\right) - B \leq B \exp(b_m) \exp\left(\sum_{n=1}^{m-1} b_n\right) - B,$$

hence proving the induction hypothesis for $m + 1$. This concludes the proof of (30).

Let us now turn to the proof of existence of a solution to Scheme (14). Let us introduce, for any $\mu \in [0, 1]$, $\nu_\mu(s) = (1 - \mu)\underline{L}_\nu s + \mu\nu(s)$. Then all the hypotheses of the lemma are satisfied with the same values \bar{L}_ν , \underline{L}_ν and C_ν , which implies that the bounds in (21)-(22) hold with the same constant for all $\mu \in [0, 1]$, if we replace u_{ini} by μu_{ini} and f by μf . Identifying $X_{\mathcal{D},0}$ with \mathbb{R}^M , let $(v_i)_{i=1, \dots, M}$ be a basis of $X_{\mathcal{D},0}$. Let $F : \mathbb{R}^{MN} \times [0, 1] \rightarrow \mathbb{R}^{MN}$ be defined, for any $u := (u_i^{(n)})_{\substack{i=1, \dots, M \\ n=1, \dots, N}}$ and $\mu \in [0, 1]$, by

$$\begin{aligned} (F(u, \mu))_i^{(n)} &= \int_{\Omega} \left(\delta_{\mathcal{D}}^{(n+\frac{1}{2})} u(\mathbf{x}) \Pi_{\mathcal{D}} v_i(\mathbf{x}) - \Pi_{\mathcal{D}} u^{(n+\theta)}(\mathbf{x}) \vec{v}(\mathbf{x}, t) \cdot \nabla_{\mathcal{D}} v_i(\mathbf{x}) + \Lambda(\mathbf{x}) \nabla_{\mathcal{D}} \nu_\mu(u^{(n+1)})(\mathbf{x}) \cdot \nabla_{\mathcal{D}} v_i(\mathbf{x}) \right) d\mathbf{x} \\ &\quad - \mu \frac{1}{\delta_{\mathcal{D}}^{(n+\frac{1}{2})}} \int_{t^{(n)}}^{t^{(n+1)}} \int_{\Omega} f(\mathbf{x}, t) \Pi_{\mathcal{D}} v_i(\mathbf{x}) d\mathbf{x} dt, \end{aligned}$$

with $u^{(0)} = \mu \mathcal{I}_{\mathcal{D}} u_{\text{ini}}$. Then F is continuous.

Noticing that the set $S = \{(v, \lambda) : F(v, \lambda) = 0\}$ is not empty since it contains the pair $(0, 0)$, let $(u, \mu) \in S$.

Owing to (22), by equivalence of norms in finite dimensional spaces, there exists $C > 0$, independent on (u, μ) , such that $|\nu(u_i^{(n)})| \leq C$ for all $i = 1, \dots, M$ and $n = 1, \dots, N$. Applying (5g), we get that

$$\forall i = 1, \dots, M, \forall n = 1, \dots, N, |u_i^{(n)}| \leq (C + C_\nu) / \underline{L}_\nu. \quad (31)$$

For $\mu = 0$, the system is linear with a right-hand-side equal to 0. The bound (31) shows that it is invertible. Therefore, we apply the constancy of the topological degree by an homotopy which allows to conclude that $F(u, 1) = 0$ has at least one solution $u \in \mathbb{R}^{MN}$ (see [15] for the topological degree theory). This proves the existence of at least one solution to Scheme (14). \square

In view of the study of the time translates, in order to fulfil the hypotheses of Kolmogorov's compactness theorem, let us prove an estimate on the dual norm of the discrete time derivative (this dual norm is defined by (43) and (44)).

Lemma 2.5 (Discrete version of (E3)).

Under Hypotheses (5), let $\mathcal{D}_T = (X_{\mathcal{D},0}, \Pi_{\mathcal{D}}, \nabla_{\mathcal{D}}, \mathcal{I}_{\mathcal{D}}, (t^{(n)})_{n=0,\dots,N})$ be a space-time gradient discretisation such that $\Pi_{\mathcal{D}}$ is a piecewise constant function reconstruction (see Definition 2.2). Let $\gamma \in (0, 1)$ be given and let $\theta \in \mathbb{R}$ be such that condition (18) holds.

Let $C_{\text{ini}} > 0$ and $C_P > 0$ be such that (19) and (20) hold.

Then there exists $C_2 > 0$, only depending on $\bar{L}_\nu, \underline{L}_\nu, C_\nu, C_P, C_{\text{ini}}, f, \vec{v}, \underline{\lambda}, \bar{\lambda}, \theta$ and γ such that, for any solution $u := (u^{(n)})_{n=0,\dots,N}$ to Scheme (14),

$$|\delta_{\mathcal{D}} u|_{L^2(0,T; \star, \mathcal{D})} \leq C_2. \quad (32)$$

Proof. From (14), we can write, for any $v \in X_{\mathcal{D},0}$ and any $n = 0, \dots, N-1$,

$$\begin{aligned} \int_{\Omega} \delta_{\mathcal{D}}^{(n+\frac{1}{2})} u(\mathbf{x}) \Pi_{\mathcal{D}} v(\mathbf{x}) d\mathbf{x} &= \int_{\Omega} \Pi_{\mathcal{D}} u^{(n+\theta)}(\mathbf{x}) \vec{v}(\mathbf{x}, t) \cdot \nabla_{\mathcal{D}} v(\mathbf{x}) d\mathbf{x} \\ &+ \frac{1}{\delta^{(n+\frac{1}{2})}} \int_{t^{(n)}}^{t^{(n+1)}} \int_{\Omega} f(\mathbf{x}, t) \Pi_{\mathcal{D}} v(\mathbf{x}) d\mathbf{x} dt - \int_{\Omega} \Lambda(\mathbf{x}) \nabla_{\mathcal{D}} \nu(u^{(n+1)})(\mathbf{x}) \nabla_{\mathcal{D}} v(\mathbf{x}) d\mathbf{x}. \end{aligned}$$

Applying the Cauchy-Schwarz inequality and applying the definition of the coercivity constant (8), we obtain

$$\begin{aligned} |\delta_{\mathcal{D}}^{(n+\frac{1}{2})} u|_{\star, \mathcal{D}} &\leq \|\vec{v}\|_{L^\infty(\Omega \times (0, T))} \|\Pi_{\mathcal{D}} u^{(n+\theta)}\|_{L^2(\Omega)} \\ &+ C_{\mathcal{D}} \frac{1}{\delta^{(n+\frac{1}{2})}} \int_{t^{(n)}}^{t^{(n+1)}} \|f(\cdot, t)\|_{L^2(\Omega)} dt + \bar{\lambda} \|\nabla_{\mathcal{D}} \nu(u^{(n+1)})\|_{L^2(\Omega)}. \end{aligned}$$

Then on one hand we use that for all $x, y, z \in \mathbb{R}$, $(x + y + z)^2 \leq 3(x^2 + y^2 + z^2)$. On the other hand, the definition $|\delta_{\mathcal{D}} u|_{L^2(0,T; \star, \mathcal{D})}^2 = \sum_{n=0}^{N-1} \delta^{(n+\frac{1}{2})} |\delta_{\mathcal{D}}^{(n+\frac{1}{2})} u|_{\star, \mathcal{D}}^2$ leads to

$$|\delta_{\mathcal{D}} u|_{L^2(0,T; \star, \mathcal{D})}^2 \leq 3 \left(\|\vec{v}\|_{L^\infty(\Omega \times (0, T))}^2 \|\Pi_{\mathcal{D}}^{(\theta)} u\|_{L^2(\Omega \times (0, T))}^2 + C_{\mathcal{D}}^2 \|f\|_{L^2(\Omega \times (0, T))}^2 + \bar{\lambda}^2 \|\nabla_{\mathcal{D}}^{(1)} \nu(u)\|_{L^2(\Omega \times (0, T))}^2 \right).$$

Thus (21) and (22) of Lemma 2.3 imply (32), where the dependence of C_2 with respect to the data of the problem are resulting from the one arising in Lemma 2.3. \square

The next lemma concerns the study of the time translates of $\Pi_{\mathcal{D}}^{(1)} \nu(u)$. Note that the estimate (22) provides an estimate on the space translate of the same function, and the combination of these two estimates allows the application of Kolmogorov's theorem for deriving a strong convergence property.

Lemma 2.6 (Estimate on the time translates).

Under Hypotheses (5), let $\mathcal{D}_T = (X_{\mathcal{D},0}, \Pi_{\mathcal{D}}, \nabla_{\mathcal{D}}, \mathcal{I}_{\mathcal{D}}, (t^{(n)})_{n=0,\dots,N})$ be a space-time gradient discretisation such that $\Pi_{\mathcal{D}}$ is a piecewise constant function reconstruction (see Definition 2.2). Let $\gamma \in (0,1)$ be given and let $\theta \in \mathbb{R}$ be such that condition (18) holds.

Let $C_{\text{ini}} > 0$ and $C_P > 0$ be such that (19) and (20) hold.

Then there exists $C_3 > 0$, only depending on $\bar{L}_\nu, \underline{L}_\nu, C_\nu, C_P, C_{\text{ini}}, f, \vec{\nu}, \underline{\lambda}, \bar{\lambda}, \theta$ and γ such that, for any solution $u := (u^{(n)})_{n=0,\dots,N}$ to Scheme (14),

$$\|\Pi_{\mathcal{D}}^{(1)}\nu(u)(\cdot, \cdot + \tau) - \Pi_{\mathcal{D}}^{(1)}\nu(u)(\cdot, \cdot)\|_{L^2(\Omega \times (0, T-\tau))}^2 \leq C_3 \sqrt{\tau(\tau + \delta_{\mathcal{D}})}, \forall \tau \in (0, T). \quad (33)$$

Proof. Let $\tau \in (0, T)$. Similarly using that \bar{L}_ν is a Lipschitz constant of ν and ν is non-decreasing, and using the fact that $\Pi_{\mathcal{D}}$ is piecewise constant, the following inequality holds:

$$\int_{\Omega \times (0, T-\tau)} \left(\Pi_{\mathcal{D}}^{(1)}\nu(u)(\mathbf{x}, t + \tau) - \Pi_{\mathcal{D}}^{(1)}\nu(u)(\mathbf{x}, t) \right)^2 d\mathbf{x}dt \leq \bar{L}_\nu \int_0^{T-\tau} A(t)dt, \quad (34)$$

where, for almost every $t \in (0, T - \tau)$,

$$A(t) = \int_{\Omega} \left(\Pi_{\mathcal{D}}^{(1)}\nu(u)(\mathbf{x}, t + \tau) - \Pi_{\mathcal{D}}^{(1)}\nu(u)(\mathbf{x}, t) \right) \left(\Pi_{\mathcal{D}}^{(1)}u(\mathbf{x}, t + \tau) - \Pi_{\mathcal{D}}^{(1)}u(\mathbf{x}, t) \right) d\mathbf{x}.$$

We apply lemma A.1 and we get

$$\int_0^{T-\tau} A(t)dt \leq 2\sqrt{\tau(\tau + \delta)} |\delta_{\mathcal{D}}u|_{L^2(0, T; \star, \mathcal{D})} \|\nabla_{\mathcal{D}}^{(1)}\nu(u)\|_{L^2(0, T; L^2(\Omega)^d)}. \quad (35)$$

Using (32), (22) in (35), we get the result. \square

3 Convergence analysis

Let us begin with the weak convergence of $\Pi_{\mathcal{D}}^{(\theta)}u(t)$ and $\Pi_{\mathcal{D}}^{(1)}u(t)$, for all $t \in [0, T]$ to an element of $C_w([0, T]; L^2(\Omega))$, denoting the set of functions from $[0, T]$ to $L^2(\Omega)$, continuous for the weak topology of $L^2(\Omega)$.

Lemma 3.1 (Time pointwise weak convergence of $\Pi_{\mathcal{D}}^{(\theta)}u(t)$ and $\Pi_{\mathcal{D}}^{(1)}u(t)$).

Let Hypotheses (5) be fulfilled. Let $((\mathcal{D}_T)_m)_{m \in \mathbb{N}}$ be a consistent sequence of space-time gradient discretisations which is limit-conforming, and such that, for all $m \in \mathbb{N}$, $\Pi_{\mathcal{D}_m}$ is a piecewise constant function reconstruction (see Definitions 2.1 and 2.2). Let $\gamma \in (0,1)$ be given and let $\theta \in \mathbb{R}$ be such that condition (18) holds for all $m \in \mathbb{N}$. For any $m \in \mathbb{N}$, let u_m be a solution to Scheme (14).

Then there exists $\bar{u} \in L^\infty(0, T; L^2(\Omega)) \cap C_w([0, T]; L^2(\Omega))$ such that, up to a subsequence, for all $t \in [0, T]$, $\Pi_{\mathcal{D}_m}^{(\theta)}u_m(t)$ and $\Pi_{\mathcal{D}_m}^{(1)}u_m(t)$ weakly converges in $L^2(\Omega)$ to $\bar{u}(t)$ as $m \rightarrow \infty$.

Proof. The existence of $C_{\text{ini}} > 0$ such that (19) holds, letting $\mathcal{D} = \mathcal{D}_m$ for all $m \in \mathbb{N}$, is a consequence of the consistency hypothesis, which implies that $\|u_{\text{ini}} - \Pi_{\mathcal{D}_m}\mathcal{I}_{\mathcal{D}_m}u_{\text{ini}}\|_{L^2(\Omega)}$ tend to 0 as $m \rightarrow \infty$, and the existence of $C_P > 0$ such that (20) holds, letting $\mathcal{D} = \mathcal{D}_m$ for all $m \in \mathbb{N}$, is resulting from the limit-conformity hypothesis, which implies the coercivity property (see [12, Lemma 2.6]).

Applying Lemma 2.3, we get that there exists $\bar{u}^{(\theta)} \in L^\infty(0, T; L^2(\Omega))$ (resp. $\bar{u}^{(1)} \in L^\infty(0, T; L^2(\Omega))$) such that $\Pi_{\mathcal{D}_m}^{(\theta)}u_m$ (resp. $\Pi_{\mathcal{D}_m}^{(1)}u_m$) weakly converges, up again to the extraction of a subsequence, to $\bar{u}^{(\theta)}$ (resp. $\bar{u}^{(1)}$) in $L^2(\Omega \times (0, T))$.

Let $\varphi \in C_c^\infty([0, T])$ and $w \in C_c^\infty(\Omega)$, and let $w_m \in X_{\mathcal{D}_m,0}$ be such that

$$w_m = \operatorname{argmin}_{z \in X_{\mathcal{D}_m,0}} S_{\mathcal{D}_m}(z). \quad (36)$$

Since the definitions of $|\cdot|_{\star, \mathcal{D}}$ and of $\delta_{\mathcal{D}}^{(n+\frac{1}{2})}u$ imply

$$\left| \int_{\Omega} (\Pi_{\mathcal{D}_m} u_m^{(n+1)} - \Pi_{\mathcal{D}_m} u_m^{(n+\theta)}) \Pi_{\mathcal{D}_m} w_m d\mathbf{x} \right| \leq \delta_{\mathcal{D}_m} |1 - \theta| |\delta_{\mathcal{D}_m}^{(n+\frac{1}{2})}u|_{\star, \mathcal{D}_m} \|w_m\|_{\mathcal{D}_m},$$

we obtain, multiplying the above inequality by $\delta_{\mathcal{D}_m}^{(n+\frac{1}{2})}\varphi(t^{(n)})w_m$, summing the resulting equation on $n = 0, \dots, N_m - 1$ and using the Cauchy-Schwarz inequality, that

$$\begin{aligned} \left| \sum_{n=0}^{N_m-1} \int_{\Omega} (\Pi_{\mathcal{D}_m} u_m^{(n+1)} - \Pi_{\mathcal{D}_m} u_m^{(n+\theta)}) \delta_{\mathcal{D}_m}^{(n+\frac{1}{2})}\varphi(t^{(n)}) \Pi_{\mathcal{D}_m} w_m d\mathbf{x} \right| \\ \leq \delta_{\mathcal{D}_m} |1 - \theta| |\delta_{\mathcal{D}_m} u|_{L^2(0, T; \star, \mathcal{D}_m)} \|w_m\|_{\mathcal{D}_m} \sqrt{T} \|\varphi\|_{L^\infty([0, T])}. \end{aligned}$$

Denoting by $\psi_m(\mathbf{x}, t) = \varphi(t^{(n)}) \Pi_{\mathcal{D}_m} w_m(\mathbf{x})$ for a.e. $t \in (t^{(n)}, t^{(n+1)})$ and $\mathbf{x} \in \Omega$, we get

$$\left| \int_0^T \int_{\Omega} (\Pi_{\mathcal{D}_m}^{(1)} u_m - \Pi_{\mathcal{D}_m}^{(\theta)} u_m) \psi_m d\mathbf{x} dt \right| \leq \delta_{\mathcal{D}_m} |1 - \theta| |\delta_{\mathcal{D}_m} u|_{L^2(0, T; \star, \mathcal{D}_m)} \|w_m\|_{\mathcal{D}_m} \sqrt{T} \|\varphi\|_{L^\infty([0, T])}.$$

Using Lemma 2.5, we get that the right hand side of the above inequality tends to 0 as $m \rightarrow \infty$. Passing to the limit in the above inequality, and using weak/strong convergence in the left hand side, we obtain that

$$\int_0^T \int_{\Omega} (\bar{u}^{(1)}(\mathbf{x}, t) - \bar{u}^{(\theta)}(\mathbf{x}, t)) \varphi(t) w(\mathbf{x}) d\mathbf{x} dt = 0.$$

Since the set $\mathcal{T} = \{\sum_{i=1}^q \varphi_i(t) w_i(\mathbf{x}) : q \in \mathbb{N}, \varphi_i \in C_c^\infty[0, T], w_i \in C_c^\infty(\Omega)\}$ is dense in $C_c^\infty(\Omega \times [0, T])$, we conclude that $\bar{u}^{(1)} = \bar{u}^{(\theta)}$. We now denote by $\bar{u} \in L^\infty(0, T; L^2(\Omega))$ the common limit of $\Pi_{\mathcal{D}_m}^{(1)} u_m$ and $\Pi_{\mathcal{D}_m}^{(\theta)} u_m$.

The fact that $\bar{u} \in C_w([0, T]; L^2(\Omega))$ and that, up to a subsequence, for all $t \in [0, T]$, $\Pi_{\mathcal{D}_m}^{(\theta)} u_m(t)$ and $\Pi_{\mathcal{D}_m}^{(1)} u_m(t)$ weakly converges in $L^2(\Omega)$ to $\bar{u}(t)$ as $m \rightarrow \infty$ is proved by [12, theorem 4.19], since its hypotheses hold thanks to Lemmas 2.3 and 2.5. \square

We can now state the convergence result in the general case.

Theorem 3.2 (Convergence of Scheme (14)).

Let Hypotheses (5) be fulfilled. Let $((\mathcal{D}_T)_m)_{m \in \mathbb{N}}$ be a consistent sequence of space-time gradient discretisations, such that the associated sequence of approximate gradient approximations is limit-conforming and compact (see Definition 2.1), and such that, for all $m \in \mathbb{N}$, $\Pi_{\mathcal{D}_m}$ is a piecewise constant function reconstruction (see Definition 2.2). Let $\gamma \in (0, 1)$ be given and let $\theta \in \mathbb{R}$ be such that condition (18) holds for all $m \in \mathbb{N}$. For any $m \in \mathbb{N}$, let u_m be a solution to Scheme (14).

Then there exists $\bar{u} \in L^\infty(0, T; L^2(\Omega)) \cap C_w([0, T]; L^2(\Omega))$ such that \bar{u} is a solution of Problem (6) and, up to a subsequence,

1. for all $t \in [0, T]$, $\Pi_{\mathcal{D}_m}^{(\theta)} u_m(t)$ and $\Pi_{\mathcal{D}_m}^{(1)} u_m(t)$ weakly converges in $L^2(\Omega)$ to $\bar{u}(t)$ as $m \rightarrow \infty$,
2. $\Pi_{\mathcal{D}_m}^{(1)} \nu(u_m)$ converges in $L^2(\Omega \times (0, T))$ to $\nu(\bar{u})$ as $m \rightarrow \infty$,
3. $\nabla_{\mathcal{D}_m}^{(1)} \nu(u_m)$ weakly converges in $L^2(\Omega \times (0, T))^d$ to $\nabla \nu(\bar{u})$ as $m \rightarrow \infty$.

Remark 3.3. Note that this theorem shows, in the case where ν is strictly increasing, that the convergence of $\Pi_{\mathcal{D}_m}^{(1)} u_m$ to \bar{u} holds in $L^2(\Omega \times (0, T))$.

Proof. We notice that the hypotheses of the theorem enable the application of Lemma 3.1, and we therefore consider the corresponding extracted subsequence. The compactness hypothesis of $(\mathcal{D}_m)_{m \in \mathbb{N}}$ allows to enter into the framework of Kolmogorov's theorem. Indeed, prolonging $\Pi_{\mathcal{D}}^{(1)} \nu(u)$ by 0 outside $\Omega \times (0, T)$, from [12,

Lemma 2.21], we get that the space translates of $\Pi_{\mathcal{D}}^{(1)}\nu(u)$ uniformly tend to 0. For the time translates, in addition to Lemma A.1, we show that the terms $\int_{\Omega \times (-\tau, 0)} \left(\Pi_{\mathcal{D}}^{(1)}\nu(u)(\mathbf{x}, t + \tau) - \Pi_{\mathcal{D}}^{(1)}\nu(u)(\mathbf{x}, t) \right)^2 d\mathbf{x}dt$ and $\int_{\Omega \times (T-\tau, T)} \left(\Pi_{\mathcal{D}}^{(1)}\nu(u)(\mathbf{x}, t + \tau) - \Pi_{\mathcal{D}}^{(1)}\nu(u)(\mathbf{x}, t) \right)^2 d\mathbf{x}dt$ are of order less than τ thanks to Estimate (21). Therefore, there exists $\chi \in L^2(\Omega \times (0, T))$ such that $\Pi_{\mathcal{D}_m}^{(1)}\nu(u_m)$ converges, up to the extraction of a subsequence, to χ in $L^2(\Omega \times (0, T))$. Thanks to the limit-conformity of the sequence $(\mathcal{D}_m)_{m \in \mathbb{N}}$, we get that $\chi \in L^2(0, T; H_0^1(\Omega))$. This allows to apply to Minty's trick [12, Lemma D.10], for concluding that $\chi(\mathbf{x}, t) = \nu(\bar{u}(\mathbf{x}, t))$ for a.e. $(\mathbf{x}, t) \in \Omega \times (0, T)$. The proof that \bar{u} is the weak solution of Problem (6) follows similar steps to that of [16, Theorem 3.1]: indeed, passing to the limit on the convection term is a consequence of the weak convergence of $\Pi_{\mathcal{D}_m}^{(\theta)}u_m(t)$. \square

Let us now provide a stronger convergence result. For a solution \bar{u} of Problem (6) and for all $t \in [0, T]$, the ‘‘mushy’’ zone $\mathcal{M}(t)$ is defined as the set of a.e. $\mathbf{x} \in \Omega$ such that there exist $a < b \in \mathbb{R}$ such that ν is constant on $[a, b]$ and $\bar{u}(\mathbf{x}, t) \in]a, b[$. This can also be expressed by

$$\text{for a.e. } \mathbf{x} \in \mathcal{M}(t), \exists a, b \in \nu^{-1}(\{\nu(\bar{u}(\mathbf{x}, t))\}), (a - \bar{u}(\mathbf{x}, t))(b - \bar{u}(\mathbf{x}, t)) < 0.$$

This last criterion is used in Corollary 3.4 and in Lemma 3.5.

Recall that, in the case of regular divergence free velocity and null right hand side f , the solution \bar{u} of (4a) is such that the measure of the mushy zone $|\mathcal{M}(t)|$ is decreasing with t , and therefore, if $\mathcal{M}(0) = 0$, we have $\mathcal{M}(t) = 0$ for all $t \in [0, T]$ (see for example [6, 22]). Denoting by $\widehat{\mathcal{M}} = \{(\mathbf{x}, t) : t \in [0, T], \mathbf{x} \in \mathcal{M}(t)\}$, we show that the convergence theorem 3.2 can be improved in $Q := \Omega \times (0, T) \setminus \widehat{\mathcal{M}}$. This convergence result implies that, for strictly increasing functions ν , or in the case of continuous solutions with negligible mushy zone (see the numerical results of Section 4), a strong convergence result holds on the whole domain.

Corollary 3.4 (Strong convergence outside the mushy region). *Under the hypotheses of Theorem 3.2, let \bar{u} be given by Theorem 3.2. Let us define Q as the largest subset of $\Omega \times (0, T)$ (up to a subset with null measure) such that:*

$$\text{for a.e. } (\mathbf{x}, t) \in Q, \forall a, b \in \nu^{-1}(\{\nu(\bar{u}(\mathbf{x}, t))\}), (a - \bar{u}(\mathbf{x}, t))(b - \bar{u}(\mathbf{x}, t)) \geq 0.$$

Then, $\int_Q (\Pi_{\mathcal{D}_m}^{(\theta)}u_m - \bar{u})^2 d\mathbf{x}dt$ and $\int_Q (\Pi_{\mathcal{D}_m}^{(1)}u_m - \bar{u})^2 d\mathbf{x}dt$ tend to 0 as $m \rightarrow \infty$.

Proof. If the measure of Q is strictly positive, we apply Lemma 3.5 letting $w_m = \Pi_{\mathcal{D}_m}^{(\theta)}u_m$ or $w_m = \Pi_{\mathcal{D}_m}^{(1)}u_m$ and $w = \bar{u}$. \square

Lemma 3.5. *Let $N \in \mathbb{N}^*$ be given, let ν be such that Hypotheses (5f) and (5g) hold, and let Q be a bounded measurable subset of \mathbb{R}^N with strictly positive measure. Let $w \in L^2(Q)$ be such that*

$$\text{for a.e. } \mathbf{x} \in Q, \forall a, b \in \nu^{-1}(\{\nu(w(\mathbf{x}))\}), (a - w(\mathbf{x}))(b - w(\mathbf{x})) \geq 0. \quad (37)$$

Let $(w_n)_{n \in \mathbb{N}}$ be a sequence of functions of $L^2(Q)$ such that, as $m \rightarrow \infty$:

1. $(w_n)_{n \in \mathbb{N}}$ weakly converges to w in $L^2(Q)$,
2. $(\nu(w_n))_{n \in \mathbb{N}}$ converges to $\nu(w)$ in $L^2(Q)$.

Then $(w_n)_{n \in \mathbb{N}}$ converges to w in $L^2(Q)$.

Proof. For all $n \in \mathbb{N}$, let us denote by \widehat{w}_n a representative of w_n , and let us denote by \widehat{w} a representative of w . Since $(\nu(w_n))_{n \in \mathbb{N}}$ converges to $\nu(w)$ in $L^2(Q)$, we extract a subsequence, that we again denote $(w_n)_{n \in \mathbb{N}}$, such that $(\nu(\widehat{w}_n(\mathbf{x})) - \nu(\widehat{w}(\mathbf{x})))^2$ converges to 0 almost everywhere with domination in $L^1(Q)$ by [20, Theorem 4.49]. We then denote by \widehat{Q} the subset of Q such that this convergence holds (we then have that the measure of $Q \setminus \widehat{Q}$ is null). Applying (5g), we get that

$$\forall n \in \mathbb{N}, \forall \mathbf{x} \in \widehat{Q}, |\widehat{w}_n(\mathbf{x})| \leq \frac{1}{\underline{L}_\nu} (|\nu(\widehat{w}_n(\mathbf{x}))| + C_\nu), \quad (38)$$

showing that

$$\forall \mathbf{x} \in \widehat{Q}, \quad |\limsup_{n \rightarrow \infty} \widehat{w}_n(\mathbf{x})| \leq \frac{1}{L_\nu} (|\nu(\widehat{w}(\mathbf{x}))| + C_\nu) \text{ and } |\liminf_{n \rightarrow \infty} \widehat{w}_n(\mathbf{x})| \leq \frac{1}{L_\nu} (|\nu(\widehat{w}(\mathbf{x}))| + C_\nu). \quad (39)$$

Denoting by $\widehat{w}_{\inf} := \liminf_{n \rightarrow \infty} \widehat{w}_n$ and $\widehat{w}_{\sup} := \limsup_{n \rightarrow \infty} \widehat{w}_n$, and owing to the monotonicity property of ν , we have that

$$\forall \mathbf{x} \in \widehat{Q}, \quad \nu(\widehat{w}(\mathbf{x})) \leq \nu(\widehat{w}_{\inf}(\mathbf{x})) \leq \nu(\widehat{w}_{\sup}(\mathbf{x})) \leq \nu(\widehat{w}(\mathbf{x})),$$

which implies

$$\forall \mathbf{x} \in \widehat{Q}, \quad \nu(\widehat{w}(\mathbf{x})) = \nu(\widehat{w}_{\inf}(\mathbf{x})) = \nu(\widehat{w}_{\sup}(\mathbf{x})) = \nu(\widehat{w}(\mathbf{x})).$$

Hence, owing to Hypothesis (37) and to $\widehat{w}_{\inf}(\mathbf{x}), \widehat{w}_{\sup}(\mathbf{x}) \in \nu^{-1}(\{\nu(\widehat{w}(\mathbf{x}))\})$, we get that

$$\text{for a.e. } \mathbf{x} \in \widehat{Q}, \quad (\widehat{w}_{\inf}(\mathbf{x}) - \widehat{w}(\mathbf{x}))(\widehat{w}_{\sup}(\mathbf{x}) - \widehat{w}(\mathbf{x})) \geq 0.$$

We now denote by \widetilde{Q} the subset of \widehat{Q} such that the above inequality holds for any $\mathbf{x} \in \widetilde{Q}$, and such that, owing to the weak convergence of w_n to w in Q , we have

$$\forall \mathbf{x} \in \widetilde{Q}, \quad \widehat{w}_{\inf}(\mathbf{x}) \leq \widehat{w}(\mathbf{x}) \leq \widehat{w}_{\sup}(\mathbf{x}).$$

We again have that the measure of $Q \setminus \widetilde{Q}$ is null. Gathering the two previous relations, we get that

$$\forall \mathbf{x} \in \widetilde{Q}, \quad \widehat{w}_{\inf}(\mathbf{x}) = \widehat{w}(\mathbf{x}) \text{ or } \widehat{w}(\mathbf{x}) = \widehat{w}_{\sup}(\mathbf{x}).$$

Let us define $\widetilde{Q}_{\inf} = \{\mathbf{x} \in \widetilde{Q}, \widehat{w}_{\inf}(\mathbf{x}) = \widehat{w}(\mathbf{x})\}$ (we similarly define \widetilde{Q}_{\sup}). We get that $\widetilde{Q} = \widetilde{Q}_{\inf} \cup \widetilde{Q}_{\sup}$. We observe that the weak convergence of w_n to w in Q implies

$$\lim_{n \rightarrow \infty} \int_{\widetilde{Q}_{\inf}} \widehat{w}_n(\mathbf{x}) d\mathbf{x} = \int_{\widetilde{Q}_{\inf}} \widehat{w}(\mathbf{x}) d\mathbf{x} = \int_{\widetilde{Q}_{\inf}} \widehat{w}_{\inf}(\mathbf{x}) d\mathbf{x},$$

(the same property holds for \widetilde{Q}_{\sup}) and therefore

$$\lim_{n \rightarrow \infty} \int_{\widetilde{Q}_{\inf}} (\widehat{w}_n(\mathbf{x}) - \widehat{w}_{\inf}(\mathbf{x})) d\mathbf{x} = \lim_{n \rightarrow \infty} \int_{\widetilde{Q}_{\sup}} (\widehat{w}_{\sup}(\mathbf{x}) - \widehat{w}_n(\mathbf{x})) d\mathbf{x} = 0.$$

Let us denote by $z_n(\mathbf{x}) = \widehat{w}_n(\mathbf{x}) - \widehat{w}_{\inf}(\mathbf{x})$ for all $\mathbf{x} \in \widetilde{Q}_{\inf}$. Since $\liminf_{n \rightarrow \infty} z_n(\mathbf{x}) = 0$ for all $\mathbf{x} \in \widetilde{Q}_{\inf}$, we get that $\lim_{n \rightarrow \infty} \min(z_n(\mathbf{x}), 0) = 0$ for all $\mathbf{x} \in \widetilde{Q}_{\inf}$, and the domination property of $|z_n(\mathbf{x})|$, resulting from (38) and (39), implies that $\lim_{n \rightarrow \infty} \int_{\widetilde{Q}_{\inf}} \min(z_n(\mathbf{x}), 0) d\mathbf{x} = 0$. Therefore, since $\int_{\widetilde{Q}_{\inf}} |z_n(\mathbf{x})| d\mathbf{x} = \int_{\widetilde{Q}_{\inf}} z_n(\mathbf{x}) d\mathbf{x} - 2 \int_{\widetilde{Q}_{\inf}} \min(z_n(\mathbf{x}), 0) d\mathbf{x}$, we get that

$$\lim_{n \rightarrow \infty} \int_{\widetilde{Q}_{\inf}} |z_n(\mathbf{x})| d\mathbf{x} = \lim_{n \rightarrow \infty} \int_{\widetilde{Q}_{\inf}} |\widehat{w}_n(\mathbf{x}) - \widehat{w}_{\inf}(\mathbf{x})| d\mathbf{x} = \lim_{n \rightarrow \infty} \int_{\widetilde{Q}_{\inf}} |\widehat{w}_n(\mathbf{x}) - \widehat{w}(\mathbf{x})| d\mathbf{x} = 0.$$

A similar reasoning shows that

$$\lim_{n \rightarrow \infty} \int_{\widetilde{Q}_{\sup}} |\widehat{w}_n(\mathbf{x}) - \widehat{w}_{\sup}(\mathbf{x})| d\mathbf{x} = \lim_{n \rightarrow \infty} \int_{\widetilde{Q}_{\sup}} |\widehat{w}_n(\mathbf{x}) - \widehat{w}(\mathbf{x})| d\mathbf{x} = 0.$$

This proves, since $\widetilde{Q} = \widetilde{Q}_{\inf} \cup \widetilde{Q}_{\sup}$, that

$$\lim_{n \rightarrow \infty} \int_{\widetilde{Q}} |\widehat{w}_n(\mathbf{x}) - \widehat{w}(\mathbf{x})| d\mathbf{x} = 0.$$

Therefore, again up to the extraction of a subsequence, we get that $\widehat{w}_n(\mathbf{x})$ converges to $\widehat{w}(\mathbf{x})$ for almost every $\mathbf{x} \in \widetilde{Q}$. Applying the dominated convergence theorem (using the domination of $(\nu(\widehat{w}_n(\mathbf{x})) - \nu(\widehat{w}(\mathbf{x})))^2$ and (38)), we then get that

$$\lim_{n \rightarrow \infty} \int_{\widetilde{Q}} (\widehat{w}_n(\mathbf{x}) - \widehat{w}(\mathbf{x}))^2 d\mathbf{x} = 0.$$

This proves the convergence of an extracted subsequence of the initial sequence $(w_n)_{n \in \mathbb{N}}$ to w in $L^2(Q)$. We then get that the whole sequence converges by a standard argument of uniqueness of the limit. \square

4 Numerical examples

We compare in this section two schemes which satisfy the assumptions of this paper. Let us briefly recall each of them.

The Control Volume Finite Element (CVFE) scheme

This scheme, also called the mass-lumped \mathbb{P}^1 finite element method, is detailed in [12, Section 8.4]. We consider a conforming triangular mesh of Ω , and we define a dual mesh by joining the centre of gravity of the triangles with the middle of the edges. Denote by \mathcal{V} the set of the vertices of the mesh, and for $\mathbf{v} \in \mathcal{V}$ define $K_{\mathbf{v}}$ as the dual cell around the vertex \mathbf{v} .

1. We then define $X_{\mathcal{D},0}$ as the set of all families $u = (u_{\mathbf{v}})_{\mathbf{v} \in \mathcal{V}}$ such that $u_{\mathbf{v}} = 0$ for all vertices \mathbf{v} located on the boundary of the domain.
2. For every $u \in X_{\mathcal{D},0}$, $\mathbf{v} \in \mathcal{V}$ and for almost-every $x \in K_{\mathbf{v}}$, $\Pi_{\mathcal{D}}u(x) = u_{\mathbf{v}}$ (piecewise constant reconstruction in all $K_{\mathbf{v}}$);
3. For every $u \in X_{\mathcal{D},0}$, define $\nabla_{\mathcal{D}}u$ as the gradient of the conforming piecewise affine function reconstructed in the triangles from the values at the vertices of the triangles.

The Vertex Approximate Gradient (VAG) scheme

The properties of the VAG scheme, introduced in [18], are detailed in [12, Section 8.5]. In the 2D case, a polygonal mesh \mathcal{M} is given, such that each element $K \in \mathcal{M}$ is strictly star-shaped with respect to some point \mathbf{x}_K . We denote by \mathcal{V} the set of all the vertices of the mesh, and by \mathcal{E} the set of all the edges of the elements of the mesh, assumed to be linear segments. For any $\sigma \in \mathcal{E}$, we denote by \mathbf{x}_{σ} its middle point. We consider two meshes of Ω . The first one is a triangular mesh, where the vertices of the triangles are the points $\mathbf{x}_K, \mathbf{x}_{\sigma}, \mathbf{v}$, for all $K \in \mathcal{M}$, for all σ of K and \mathbf{v} common vertex of K and σ . The second one is a dual mesh, associated to all points $(\mathbf{x}_K)_{K \in \mathcal{M}}$ and $(\mathbf{v})_{\mathbf{v} \in \mathcal{V}}$.

1. We then define $X_{\mathcal{D},0}$ as the set of all families $u = ((u_K)_{K \in \mathcal{M}}, (u_{\mathbf{v}})_{\mathbf{v} \in \mathcal{V}})$ such that $u_{\mathbf{v}} = 0$ for all $\mathbf{v} \in \mathcal{V} \cap \partial\Omega$.
2. The mapping $\Pi_{\mathcal{D}}$ is defined by piecewise constant functions having the values u_K in the dual control volume associated to \mathbf{x}_K for all $K \in \mathcal{M}$ and $u_{\mathbf{v}}$ in the dual control volume associated to $\mathbf{v} \in \mathcal{V}$.
3. Considering the value $u_{\sigma} = \frac{1}{2}(u_{\mathbf{v}} + u_{\mathbf{v}'})$ at the middle \mathbf{x}_{σ} of an edge $\sigma = [\mathbf{v}, \mathbf{v}']$, the mapping $\nabla_{\mathcal{D}}$ is defined as the gradient of the \mathbb{P}^1 affine reconstruction with the values $u_K, u_{\sigma}, u_{\mathbf{v}}$ at the vertices $\mathbf{x}_K, \mathbf{x}_{\sigma}, \mathbf{v}$ of any triangle of the triangular mesh.

This scheme has two advantages. Firstly, adjusting the dual mesh with respect to the heterogeneous properties of the domain, it allows accurate computations of coupled conservation equations in porous media [18]. Secondly, it leads to cheap computations with the elimination of all values $(u_K)_{K \in \mathcal{M}}$ with respect to the values $(u_{\mathbf{v}})_{\mathbf{v} \in \mathcal{V}}$.

Data for the numerical tests

In the following numerical examples, we consider the function ν presented in the introduction of this paper, letting $C_v = C_l = 1$, $T_f = 0$ and $L = 1$, which leads to

$$\forall s \in \mathbb{R}, \nu(s) = \begin{cases} s & \text{if } s < 0, \\ 0 & \text{if } 0 \leq s \leq 1, \\ s - 1 & \text{if } 1 < s. \end{cases} \quad (40)$$

In this case, the data involved in (5f) and (5g) can be chosen as $\bar{L}_{\nu} = 1$, $\underline{L}_{\nu} = 1$ and $C_{\nu} = 1$.

We let $d = 2$, $\Omega = (0, 1) \times (0, 1)$, $\vec{\nu} = (1, 1)$ (hence $\|\vec{\nu}\|_{\infty} = \sqrt{2}$), $\Lambda = \text{Id}$ (we then choose $\underline{\lambda} = \bar{\lambda} = 1$ in (5d)). Letting γ tend to 1, and considering the case $\delta_{\mathcal{D}} = 0.001$ (this value is selected in most of the numerical tests

below) the maximum value of θ^2 such that Condition (18) holds is therefore equal to $125\sqrt{2}$, which enables $|\theta| \leq 10$.

We finally let $f = 0$ and $u_{\text{ini}}(\mathbf{x}) = \alpha_{\text{ini}}$ if $\mathbf{x} = (x_1, x_2) \in (0.1, 0.4) \times (0.1, 0.4)$ and $u_{\text{ini}}(\mathbf{x}) = 0$ elsewhere, with the three cases $\alpha_{\text{ini}} = 1.5$, $\alpha_{\text{ini}} = 1$ and $\alpha_{\text{ini}} = .5$. These three choices are aimed to provide the behaviour of the numerical schemes in three cases: in the first and second (which is a limit case) cases a strong convergence property holds for $\Pi_{\mathcal{D}}u$ (see Lemma 3.5), and in the third one is a case where oscillations are expected. The computations are done with a final time $T = 0.5$, with a constant time step equal to 0.001 and a family of triangular meshes extracted from the benchmark [21]. The coarsest grid, respectively the finest one, has a space step equal to 0.25, respectively 0.016.

Since $\vec{\mathbf{v}} \cdot \mathbf{n} < 0$ on a part of $\partial\Omega$, any value $\alpha \in [0, 1]$ can be considered as an entering boundary condition for the linear hyperbolic problem, and there are infinitely many solutions to (6). Nevertheless, all of them are identical on the domain $(t, 1) \times (t, 1)$ at any time $t \in (0, T)$ since $\vec{\mathbf{v}} = (1, 1)$. The following comparisons with an analytical solution are therefore done in the domain $Q_T := \bigcup_{t \in (0, T)} (t, 1) \times (t, 1) \times \{t\} \subset \Omega \times (0, T)$.

4.1 Case $\alpha_{\text{ini}} = 1.5$

The mushy zone is such that $\mathcal{M}(0) = 0$. So any solution of (6) is given at any time $t \in (0, T]$ by $\bar{u}(\cdot, t) > 1$ a.e. inside a moving domain $\tilde{\Omega}(t)$, with $\tilde{\Omega}(0) = (0.1, 0.4) \times (0.1, 0.4)$, and $\bar{u}(\cdot, t) = 0$, a.e. inside $[t, 1] \times [t, 1] \setminus \tilde{\Omega}(t)$. Due to the choice of the velocity $\vec{\mathbf{v}}$, this domain $\tilde{\Omega}(t)$ moves along the diagonal $[(0, 0), (1, 1)]$.

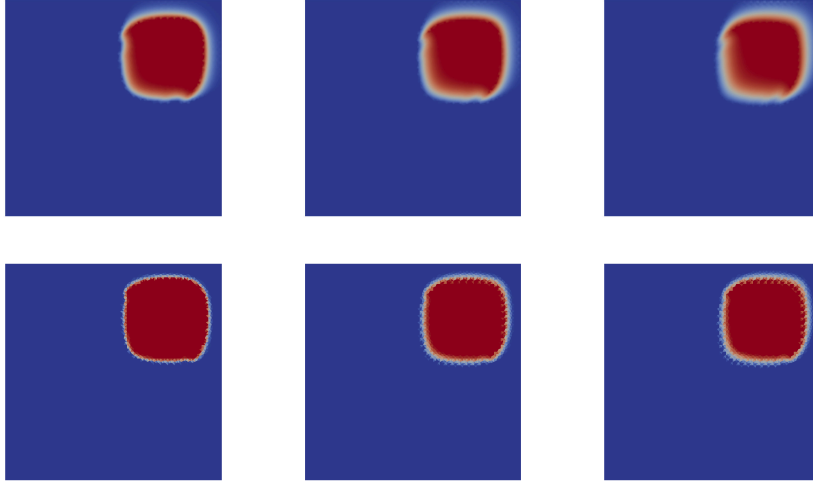


Figure 1: Case $\alpha_{\text{ini}} = 1.5$, solutions at $T = 0.5$. Top: CVFE scheme (from left to right: $\theta = 0$, $\theta = 0.5$, $\theta = 1$). Bottom: same for the VAG scheme (cell center values). The colour blue is associated to the value 0, whereas the colour red corresponds to 1. Results obtained on the finest grid.

We then observe in Figures 1 and 2 that the numerical solutions obtained with the CVFE and VAG schemes respect the expected physical features of the problem and seem to numerically converge to this solution. There is a good agreement between the two schemes, for the different values of θ on both figures. Recall that, in this case, Theorem 3.2 and Corollary 3.4 prove the convergence in $L^2(Q_T)$ of the numerical solution to the continuous one, since $\bar{u}(\mathbf{x}, t) \notin (0, 1)$ for a.e. $(\mathbf{x}, t) \in Q_T$.

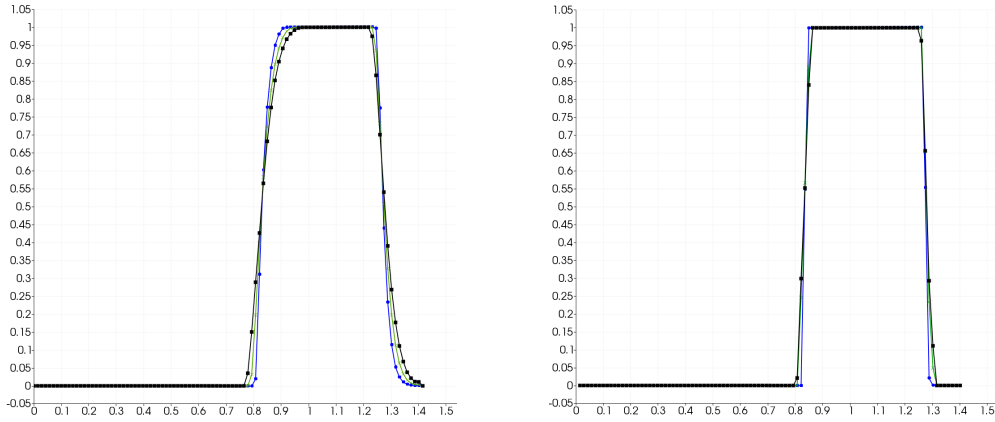


Figure 2: Case $\alpha_{\text{ini}} = 1.5$, solutions at $T = 0.5$. Comparison of profiles along the first diagonal. Circle blue for $\theta = 0$, cross green for $\theta = 0.5$ and square black for $\theta = 1$. Left: CVFE scheme. Right: same for the VAG scheme (cell center values). Results obtained on the finest grid.

4.2 Case $\alpha_{\text{ini}} = 1$

In this case, any solution \bar{u} to Problem (6) is also solution of the following pure convection problem:

$$\partial_t \bar{u} + \text{div}(\bar{u} \vec{v}) = 0, \text{ in } \Omega \times (0, T), \quad (41)$$

with the same initial condition, and is therefore equal, for any $t \in (0, T)$ to 1 in $[t+0.1, t+0.4] \times [t+0.1, t+0.4]$ and to 0 on $[t, 1] \times [t, 1] \setminus [t+0.1, t+0.4] \times [t+0.1, t+0.4]$. Therefore the measure of the mushy zone is null in Q_T , and Theorem 3.2 and Corollary 3.4 imply a strong convergence in Q_T .

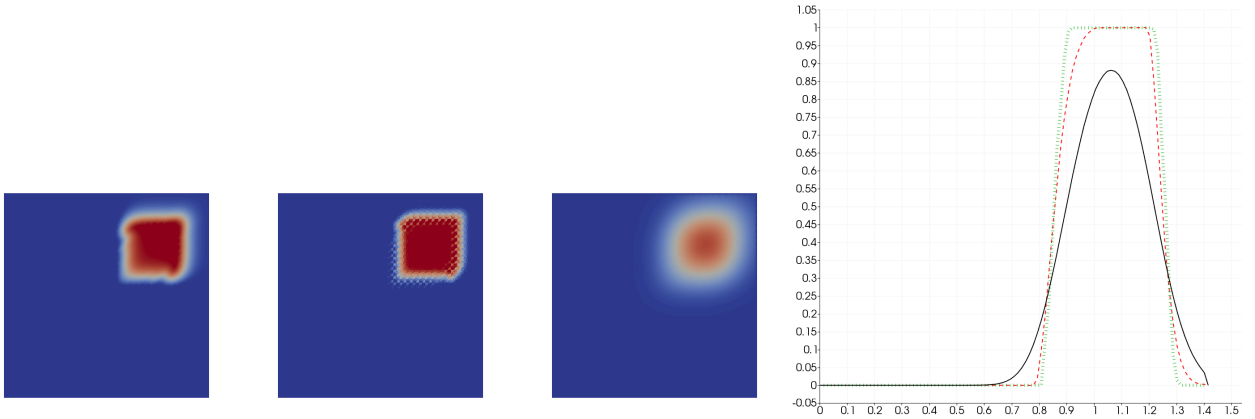


Figure 3: Case $\alpha_{\text{ini}} = 1$, solutions at $T = 0.5$. From left to right: CVFE, VAG (cell centre values), upstream schemes (the colour blue is associated to the value 0, whereas the colour red corresponds to 1) and comparison of profiles along the first diagonal (dash red for CVE, dot green for VAG, black line for upstream). Results obtained on the finest grid.

We compare in Figure 3 the numerical solution obtained at $T = 0.5$ with $\theta = 1$ by the CVFE and the VAG schemes, compared to an upstream implicit weighting scheme. The accurateness of centred schemes compared to upstream scheme is confirmed by the L^1 and L^2 errors respectively shown for these two schemes in Table 1. We also again observe that the bounds are not strictly respected by the centred scheme, but that the error committed on the bounds tends to 0 rapidly. As expected in the case of a discontinuous analytical solution, the numerical orders of convergence are nevertheless closer to $1/2$ than to 1.

h	δ	L^2 -error	ratio	L^1 -error	ratio	u_{\min}	u_{\max}
0.250	0.01024	0.202E+00	-	0.769E-01	-	-0.020	1.032
0.125	0.00512	0.136E+00	0.572	0.378E-01	1.023	-0.021	1.024
0.062	0.00256	0.992E-01	0.454	0.239E-01	0.662	-0.012	1.019
0.031	0.00128	0.748E-01	0.408	0.149E-01	0.685	-0.007	1.009
0.016	0.00064	0.574E-01	0.383	0.922E-02	0.689	-0.004	1.005
0.250	0.01024	0.182E+00	-	0.572E-01	-	-0.040	1.045
0.125	0.00512	0.126E+00	0.535	0.314E-01	0.864	-0.018	1.019
0.062	0.00256	0.942E-01	0.415	0.171E-01	0.876	-0.013	1.016
0.031	0.00124	0.708E-01	0.412	0.102E-01	0.751	-0.007	1.006
0.016	0.00064	0.550E-01	0.362	0.619E-02	0.717	-0.003	1.004
0.250	0.01024	0.170E+00	-	0.625E-01	-	0.000	0.995
0.125	0.00512	0.151E+00	0.170	0.488E-01	0.358	0.000	1.000
0.062	0.00256	0.124E+00	0.293	0.401E-01	0.281	0.000	1.000
0.031	0.00128	0.102E+00	0.275	0.311E-01	0.369	0.000	1.000
0.016	0.00064	0.854E-01	0.258	0.229E-01	0.442	0.000	1.000

Table 1: Case $\alpha_{\text{ini}} = 1$ with $\theta = 1$. $L^2(Q_T)$ and $L^1(Q_T)$ discrete errors. Results for CVFE method (top array), VAG scheme (middle array) and upstream scheme (bottom array).

In this situation where the degenerate diffusion numerically enables the strong convergence of the scheme, the scheme shows a very low dependence with respect to θ , as shown in Table 2. This is confirmed by the comparison of the profiles of the solution along the first diagonal (see Figure 4), which nevertheless shows that increasing the values of θ is similar to adding diffusion (as we also observe in Section 4.4).

h	δ	L^2 -error	ratio	L^1 -error	ratio	u_{\min}	u_{\max}
0.250	0.01024	0.277E+00	-	0.983E-01	-	-0.059	1.066
0.125	0.00512	0.169E+00	0.714	0.428E-01	1.199	-0.038	1.035
0.062	0.00256	0.106E+00	0.677	0.216E-01	0.990	-0.016	1.020
0.031	0.00128	0.745E-01	0.506	0.130E-01	0.727	-0.007	1.010
0.016	0.00064	0.561E-01	0.410	0.832E-02	0.646	-0.004	1.005
0.250	0.01024	0.192E+00	-	0.740E-01	-	-0.013	1.028
0.125	0.00512	0.136E+00	0.496	0.392E-01	0.916	-0.019	1.021
0.062	0.00256	0.100E+00	0.441	0.254E-01	0.630	-0.012	1.018
0.031	0.00128	0.756E-01	0.404	0.156E-01	0.703	-0.007	1.009
0.016	0.00064	0.579E-01	0.384	0.954E-02	0.708	-0.004	1.005
0.250	0.01024	0.189E+00	-	0.733E-01	-	-0.001	1.022
0.125	0.00512	0.140E+00	0.438	0.419E-01	0.807	-0.017	1.018
0.062	0.00256	0.102E+00	0.448	0.274E-01	0.613	-0.012	1.017
0.031	0.00128	0.769E-01	0.413	0.166E-01	0.726	-0.006	1.009
0.016	0.00064	0.588E-01	0.389	0.997E-02	0.733	-0.004	1.005

Table 2: Case $\alpha_{\text{ini}} = 1$. $L^2(Q_T)$ and $L^1(Q_T)$ discrete errors. Results for CVFE method. $\theta = -8$ (top array), $\theta = 4$ (middle array) and $\theta = 8$ (bottom array).

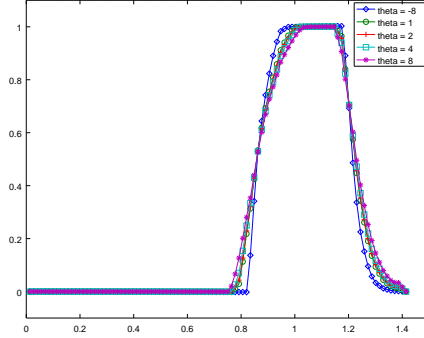


Figure 4: Case $\alpha_{\text{ini}} = 1$, solutions at $T = 0.5$ with $h = 0.031$. Comparison of profiles along the first diagonal for different values of θ .

4.3 Case $\alpha_{\text{ini}} = 0.5$

We now turn to the case where $\alpha_{\text{ini}} = 0.5$. In this case, any solution \bar{u} to Problem (6) is also solution of (41) with the same initial condition as the problem studied in this paper. It is therefore equal, for any $t \in (0, T)$, to 0.5 in $[t + 0.1, t + 0.4] \times [t + 0.1, t + 0.4]$ and to 0 on $[t, 1] \times [t, 1] \setminus [t + 0.1, t + 0.4] \times [t + 0.1, t + 0.4]$. Therefore the measure of the mushy zone is strictly positive and Theorem 3.2 and Corollary 3.4 imply a strong convergence only in $Q_T \setminus \bigcup_{t \in (0, T)} [t + 0.1, t + 0.4] \times [t + 0.1, t + 0.4] \times \{t\}$. Nevertheless, although the numerical results show oscillations, the weak convergence for $\Pi_{\mathcal{D}} u$ holds for any value $\theta \in \mathbb{R}$. We compare in Figure 5 the numerical solution obtained at $T = 0.5$ with the CVFE and VAG schemes on the finest mesh. We observe that the height of the oscillations is decreasing with respect to θ , and that high values for θ stabilise the oscillations, but introduce some numerical diffusion. The solution obtained with the VAG scheme and $\theta = 1.5$ appears to be quite good.

4.4 The θ -scheme: another way to introduce numerical diffusion

In numerical codes such as the one used in [5], the upstream weighting scheme is used for the energy equation, leading to the addition of numerical diffusion in order to stabilise the convection terms. Let us show how the parameter θ can play a similar role (as we noticed in the numerical Section 4.3), by using an analogy with continuous equations.

We assume here that \vec{v} is a constant vector and we consider formally the transport equation, in which $\bar{u}(t)$ is replaced by $\bar{u}(t + \theta \delta t)$ in the convection term:

$$\partial_t \bar{u}(\mathbf{x}, t) + \text{div}(\bar{u}(\mathbf{x}, t + \theta \delta t) \vec{v}) = \partial_t \bar{u}(\mathbf{x}, t) + \vec{v} \cdot \nabla \bar{u}(\mathbf{x}, t + \theta \delta t) = 0, \text{ in } \Omega \times (0, T). \quad (42)$$

We use the following approximation

$$\bar{u}(\mathbf{x}, t + \theta \delta t) \simeq \bar{u}(\mathbf{x}, t) + \theta \delta t \partial_t \bar{u}(\mathbf{x}, t) = \bar{u}(\mathbf{x}, t) - \theta \delta t \vec{v} \cdot \nabla \bar{u}(\mathbf{x}, t + \theta \delta t).$$

We then have, reporting this value in (42) and setting $\varepsilon = \theta \delta t |\vec{v}|$ (which is homogeneous with a length),

$$\partial_t \bar{u}(\mathbf{x}, t) + \vec{v} \cdot \nabla \bar{u}(\mathbf{x}, t) - \varepsilon \text{div}(D \nabla \bar{u}(\mathbf{x}, t + \theta \delta t)) = 0,$$

with

$$D = \frac{1}{|\vec{v}|} \vec{v} \otimes \vec{v}.$$

This equation shows a transport equation, with an anisotropic diffusion term which applies in the direction of the velocity, as does an upstream weighting scheme.

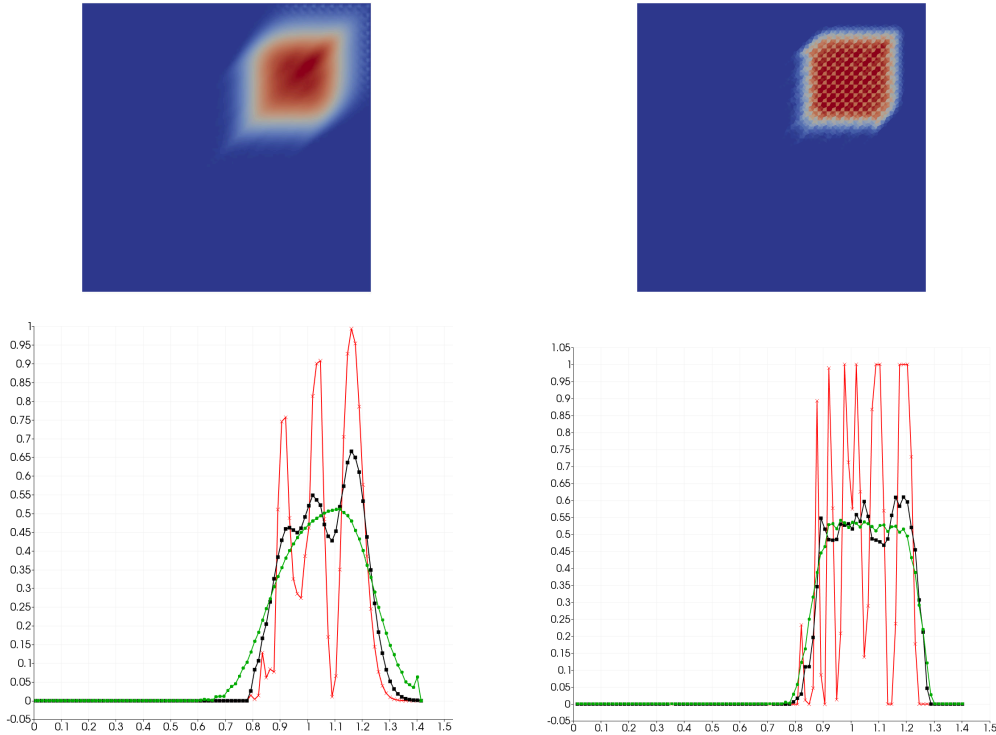


Figure 5: Case $\alpha_{ini} = 0.5$, solutions at $T = 0.5$. Results obtained on the finest grid. Left: CVFE scheme. Right: same for the VAG scheme (cell center values). Top: colour blue is associated to the value 0, whereas colour red corresponds to 0.5, CVFE for $\theta = 4$ and VAG for $\theta = 1.5$. Bottom: comparison of profiles along the diagonal, cross red for $\theta = 0$, square black for $\theta = 1$ and circle green for $\theta = 4$ (CVFE) or $\theta = 1.5$ (VAG).

5 Conclusion

The mathematical study of the approximation of a linear convection – degenerate diffusion problem by a centred θ -scheme for the convection term shows the following features:

1. The strong convergence of the scheme is observed everywhere outside the mushy region.
2. A weak convergence always holds for any value of θ , due to the degenerate diffusion term, which is strongly different with the properties of a centred finite volume scheme without (degenerate) diffusion.
3. The centred scheme happens to be much more precise than upstream schemes in some situations.
4. The parameter θ can be numerically used for stabilising the centred scheme as an artificial viscosity.

These results show that a track which remains to be studied is the use of such a θ -centred scheme in practical applications (such as the ones handled in [5]), with the goal to tune the diffusion in such a way that the precision is sufficiently respected.

Acknowledgement

The authors thank Professor Thierry Gallouët for fruitful discussions.

A Gradient discretisations and time translates

Lemma A.1. *Let $(X_{\mathcal{D},0}, \Pi_{\mathcal{D}}, \nabla_{\mathcal{D}}, \mathcal{I}_{\mathcal{D}}, (t^{(n)})_{n=0,\dots,N})$ be a space-time gradient discretisation (we then denote by $\delta^{(n+\frac{1}{2})} := t^{(n+1)} - t^{(n)}$ and $\delta_{\mathcal{D}} = \max_{n=0,\dots,N-1} \delta^{(n+\frac{1}{2})}$). We define the following semi-norm on $L^2(\Omega)$:*

$$\forall w \in L^2(\Omega), |w|_{\star, \mathcal{D}} = \sup \left\{ \int_{\Omega} w(\mathbf{x}) \Pi_{\mathcal{D}} v(\mathbf{x}) d\mathbf{x}; v \in X_{\mathcal{D},0}, \|\nabla_{\mathcal{D}} v\|_{L^2(\Omega)^d} = 1 \right\}, \quad (43)$$

and we introduce the following semi-norm on $L^2(0, T; L^2(\Omega))$:

$$\forall w \in L^2(0, T; L^2(\Omega)), |w|_{L^2(0, T; \star, \mathcal{D})} = \left(\int_0^T |w(t)|_{\star, \mathcal{D}}^2 dt \right)^{1/2}. \quad (44)$$

For any sequence $(u^{(n)})_{n=0,\dots,N}$ of elements of $X_{\mathcal{D},0}$, we define

$$\begin{aligned} \text{for a.e. } \mathbf{x} \in \Omega, \forall t \in (t^{(n)}, t^{(n+1)}], \forall n = 0, \dots, N-1: \delta_{\mathcal{D}}^{(n+\frac{1}{2})} u(\mathbf{x}) &= \Pi_{\mathcal{D}} \frac{u^{(n+1)} - u^{(n)}}{\delta^{(n+\frac{1}{2})}}(\mathbf{x}), \\ \delta_{\mathcal{D}} u(\mathbf{x}, t) &= \delta_{\mathcal{D}}^{(n+\frac{1}{2})} u(\mathbf{x}), \quad \Pi_{\mathcal{D}}^{(1)} u(\mathbf{x}, t) = \Pi_{\mathcal{D}} u^{(n+1)}(\mathbf{x}), \quad \nabla_{\mathcal{D}}^{(1)} u(\mathbf{x}, t) = \nabla_{\mathcal{D}} u^{(n+1)}(\mathbf{x}), \end{aligned}$$

Then, for any sequences $(u^{(n)})_{n=0,\dots,N}$ and $(v^{(n)})_{n=0,\dots,N}$ of elements of $X_{\mathcal{D},0}$, there holds

$$\begin{aligned} \int_0^{T-\tau} \int_{\Omega} (\Pi_{\mathcal{D}}^{(1)} v(\mathbf{x}, t+\tau) - \Pi_{\mathcal{D}}^{(1)} v(\mathbf{x}, t)) (\Pi_{\mathcal{D}}^{(1)} u(\mathbf{x}, t+\tau) - \Pi_{\mathcal{D}}^{(1)} u(\mathbf{x}, t)) d\mathbf{x} dt \\ \leq 2\sqrt{\tau(\tau + \delta_{\mathcal{D}})} |\delta_{\mathcal{D}} v|_{L^2(0, T; \star, \mathcal{D})} \|\nabla_{\mathcal{D}}^{(1)} u\|_{L^2(0, T; L^2(\Omega)^d)}. \end{aligned} \quad (45)$$

Proof. Let us define

$$T_1(t) = \int_{\Omega} (\Pi_{\mathcal{D}}^{(1)} u(\mathbf{x}, t+\tau) - \Pi_{\mathcal{D}}^{(1)} u(\mathbf{x}, t)) (\Pi_{\mathcal{D}}^{(1)} v(\mathbf{x}, t+\tau) - \Pi_{\mathcal{D}}^{(1)} v(\mathbf{x}, t)) d\mathbf{x}.$$

Let $t \in (0, T-\tau)$. Denoting $n_0(t), n_1(t) = 0, \dots, N-1$ such that $t^{(n_0(t))} \leq t < t^{(n_0(t)+1)}$ and $t^{(n_1(t))} \leq t+\tau < t^{(n_1(t)+1)}$, we may write

$$T_1(t) = \int_{\Omega} \left(\Pi_{\mathcal{D}} u^{(n_1(t)+1)}(\mathbf{x}) - \Pi_{\mathcal{D}} u^{(n_0(t)+1)}(\mathbf{x}) \right) \left(\sum_{n=n_0(t)+1}^{n_1(t)} \delta^{(n+\frac{1}{2})} \delta_{\mathcal{D}}^{(n+\frac{1}{2})} v(\mathbf{x}) \right) d\mathbf{x},$$

which also reads

$$T_1(t) = \int_{\Omega} \left(\Pi_{\mathcal{D}} u^{(n_1(t)+1)}(\mathbf{x}) - \Pi_{\mathcal{D}} u^{(n_0(t)+1)}(\mathbf{x}) \right) \left(\sum_{n=1}^{N-1} \chi_n(t, t+\tau) \delta^{(n+\frac{1}{2})} \delta_{\mathcal{D}}^{(n+\frac{1}{2})} v(\mathbf{x}) \right) d\mathbf{x},$$

with $\chi_n(t, t+\tau) = 1$ if $t^{(n)} \in (t, t+\tau]$ and $\chi_n(t, t+\tau) = 0$ if $t^{(n)} \notin (t, t+\tau]$.

This leads to

$$T_1(t) \leq \left(\|\nabla_{\mathcal{D}} u^{(n_1(t)+1)}\|_{L^2(\Omega)^d} + \|\nabla_{\mathcal{D}} u^{(n_0(t)+1)}\|_{L^2(\Omega)^d} \right) \left(\sum_{n=1}^{N-1} \chi_n(t, t+\tau) \delta^{(n+\frac{1}{2})} |\delta_{\mathcal{D}}^{(n+\frac{1}{2})} v|_{\star, \mathcal{D}} \right).$$

Using the inequality $ab \leq \frac{1}{2}(\alpha a^2 + \frac{1}{\alpha} b^2)$ for some $\alpha > 0$ chosen later, this yields:

$$T_1(t) \leq \frac{\alpha}{2} (T_2(t) + T_3(t)) + \frac{1}{\alpha} T_4(t). \quad (46)$$

with

$$T_2(t) = \sum_{n=1}^{N-1} \chi_n(t, t + \tau) \delta^{(n+\frac{1}{2})} \|\nabla_{\mathcal{D}} u^{(n_0(t)+1)}\|_{L^2(\Omega)^d}^2,$$

$$T_3(t) = \sum_{n=1}^{N-1} \chi_n(t, t + \tau) \delta^{(n+\frac{1}{2})} \|\nabla_{\mathcal{D}} u^{(n_1(t)+1)}\|_{L^2(\Omega)^d}^2,$$

and

$$T_4(t) = \sum_{n=1}^{N-1} \chi_n(t, t + \tau) \delta^{(n+\frac{1}{2})} |\delta_{\mathcal{D}}^{(n+\frac{1}{2})} v|_{\star, \mathcal{D}}^2.$$

Applying [19, Proposition 9.3] yields

$$\int_0^{T-\tau} T_2(t) dt \leq (\tau + \delta_{\mathcal{D}}) \|\nabla_{\mathcal{D}}^{(1)} u\|_{L^2(0, T; L^2(\Omega)^d)}^2 \quad \text{and} \quad \int_0^{T-\tau} T_3(t) dt \leq (\tau + \delta_{\mathcal{D}}) \|\nabla_{\mathcal{D}}^{(1)} u\|_{L^2(0, T; L^2(\Omega)^d)}^2, \quad (47)$$

and again applying [19, Proposition 9.3] gives

$$\int_0^{T-\tau} T_4(t) dt \leq \tau \int_0^T |\delta_{\mathcal{D}} v(t)|_{\star, \mathcal{D}}^2 dt = \tau \|\delta_{\mathcal{D}} v\|_{L^2(0, T; \star, \mathcal{D})}^2, \quad (48)$$

hence leading to (45), letting $\alpha = \sqrt{\tau/(\tau + \delta_{\mathcal{D}})} \|\delta_{\mathcal{D}} v\|_{L^2(0, T; \star, \mathcal{D})} / \|\nabla_{\mathcal{D}}^{(1)} u\|_{L^2(0, T; L^2(\Omega)^d)}$. \square

References

- [1] H. W. Alt and S. Luckhaus. Quasilinear elliptic-parabolic differential equations. *Math. Z.*, 183(3):311–341, 1983.
- [2] L. Ambrosio. Transport equation and Cauchy problem for BV vector fields. *Invent. Math.*, 158(2):227–260, 2004.
- [3] G. Amiez and P.-A. Gremaud. On a numerical approach to Stefan-like problems. *Numer. Math.*, 59(1):71–89, 1991.
- [4] B. Andreianov, M. Bendahmane, K. H. Karlsen, and S. Ouaro. Well-posedness results for triply nonlinear degenerate parabolic equations. *J. Differential Equations*, 247(1):277–302, 2009.
- [5] L. Beaudé, K. Brenner, S. Lopez, R. Masson, and F. Smai. Non-isothermal compositional two-phase Darcy flow: formulation and outflow boundary condition. In *Finite volumes for complex applications VIII—hyperbolic, elliptic and parabolic problems*, volume 200 of *Springer Proc. Math. Stat.*, pages 317–325. Springer, Cham, 2017.
- [6] M. Bertsch, P. de Mottoni, and L. A. Peletier. The Stefan problem with heating: appearance and disappearance of a mushy region. *Trans. Amer. Math. Soc.*, 293(2):677–691, 1986.
- [7] C. Cancès and C. Guichard. Numerical analysis of a robust free energy diminishing finite volume scheme for parabolic equations with gradient structure. *Found. Comput. Math.*, 17(6):1525–1584, 2017.
- [8] J. Carrillo. Entropy solutions for nonlinear degenerate problems. *Arch. Ration. Mech. Anal.*, 147(4):269–361, 1999.
- [9] Z. Chen. Expanded mixed finite element methods for quasilinear second order elliptic problems. II. *RAIRO Modél. Math. Anal. Numér.*, 32(4):501–520, 1998.

- [10] D. De Vries. Thermal properties of soils. In W. R. van Wijk, editor, *Physics of Plant Environment (W. R. van Wijk ed.)*, pages 210–235. North-Holland Publishing Co., Amsterdam, 1964.
- [11] R. J. DiPerna and P.-L. Lions. On the Cauchy problem for Boltzmann equations: global existence and weak stability. *Ann. of Math. (2)*, 130(2):321–366, 1989.
- [12] J. Droniou, R. Eymard, T. Gallouët, C. Guichard, and R. Herbin. *The gradient discretisation method*, volume 82 of *Mathematics & Applications*. Springer International Publishing, 2018. <https://hal.archives-ouvertes.fr/hal-01382358>.
- [13] C. M. Elliott. On the finite element approximation of an elliptic variational inequality arising from an implicit time discretization of the Stefan problem. *IMA J. Numer. Anal.*, 1(1):115–125, 1981.
- [14] C. M. Elliott. Error analysis of the enthalpy method for the Stefan problem. *IMA J. Numer. Anal.*, 7(1):61–71, 1987.
- [15] L. C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition, 2010.
- [16] R. Eymard, P. Feron, T. Gallouët, R. Herbin, and C. Guichard. Gradient schemes for the Stefan problem. *International Journal on Finite Volumes*, page Volume 10 special, June 2013.
- [17] R. Eymard, T. Gallouët, C. Guichard, R. Herbin, and R. Masson. TP or not TP, that is the question. *Comput. Geosci.*, 18(3-4):285–296, 2014.
- [18] R. Eymard, C. Guichard, and R. Herbin. Small-stencil 3d schemes for diffusive flows in porous media. *M2AN Math. Model. Numer. Anal.*, 46:265–290, 2012.
- [19] R. Eymard, R. Herbin, and A. Michel. Mathematical study of a petroleum-engineering scheme. *M2AN Math. Model. Numer. Anal.*, 37(6):937–972, 2003.
- [20] T. Gallouët and R. Herbin. *Mesure, Intégration, Probabilités*. Références science. Ellipses, Paris, 2013.
- [21] R. Herbin and F. Hubert. Benchmark on discretization schemes for anisotropic diffusion problems on general grids. In *Finite volumes for complex applications V*, pages 659–692. ISTE, London, 2008.
- [22] I. C. Kim and N. Požár. Viscosity solutions for the two-phase Stefan problem. *Comm. Partial Differential Equations*, 36(1):42–66, 2011.
- [23] F. A. Milner. Mixed finite element methods for quasilinear second-order elliptic problems. *Math. Comp.*, 44(170):303–320, 1985.
- [24] R. Nochetto and C. Verdi. Approximation of degenerate parabolic problems using numerical integration. *SIAM Journal on Numerical Analysis*, 25(4):784–814, 1988.
- [25] G. Stampacchia. Le problème de Dirichlet pour les équations elliptiques du second ordre à coefficients discontinus. *Ann. Inst. Fourier (Grenoble)*, 15(fasc. 1):189–258, 1965.
- [26] F. Yi and T. M. Shih. Stefan problem with convection. *Appl. Math. Comput.*, 95(2-3):139–154, 1998.