Uniform-in-time convergence result of numerical methods for non-linear parabolic equations

Jérôme Droniou, Robert Eymard

Uniform-in-time convergence result of numerical methods for non-linear parabolic equations

J. Droniou* and R. Eymard†

October 28, 2014

Abstract

We prove that all Gradient Schemes – which include Finite Element, some Mixed Finite Element and Finite Volume methods – converge uniformly in time when applied to a family of nonlinear parabolic equations which contains the Richards, Stefan and Leray-Lions models.

AMS Subject Classification: 46N40

1 Introduction

1.1 Motivation

The following generic nonlinear parabolic model

$$\begin{aligned}
&\partial_t\beta(\overline{u}) - \operatorname{div}\left(\boldsymbol{a}(\boldsymbol{x},\nu(\overline{u}),\nabla\zeta(\overline{u}))\right) = f \text{ in } \Omega\times(0,T),\\
&\beta(\overline{u})(\boldsymbol{x},0) = \beta(u_{\rm ini})(\boldsymbol{x}) \text{ in } \Omega,\\
&\zeta(\overline{u}) = 0 \text{ on } \partial\Omega\times(0,T)
\end{aligned} \tag{1}$$

where β,ζ are non-decreasing, ν is such that $\nu'=\beta'\zeta'$ and $\boldsymbol{a}$ is a Leray-Lions operator, arises in various frameworks (see next section for precise hypotheses on the data). This model includes

1. Richards' model, setting $\zeta(s)=s$, $\nu=\beta$ and $\boldsymbol{a}(\boldsymbol{x},\nu(\overline{u}),\nabla\zeta(\overline{u}))=\Lambda(\boldsymbol{x})K(\boldsymbol{x},\beta(\overline{u}))\nabla\overline{u}$, which describes the flow of water in a heterogeneous anisotropic underground medium,
2. Stefan's model [4], setting $\beta(s)=s$, $\nu=\zeta$, $\boldsymbol{a}(\boldsymbol{x},\nu(\overline{u}),\nabla\zeta(\overline{u}))=K(\zeta(\overline{u}))\nabla\zeta(\overline{u})$, which arises in the study of a simplified heat diffusion in a melting medium,
3. $p-$Laplace problem (and $p-$Laplace-like problems), setting $\beta(s)=\zeta(s)=\nu(s)=s$ and $\boldsymbol{a}(\boldsymbol{x},\nu(\overline{u}),\nabla\zeta(\overline{u}))=|\nabla\overline{u}|^{p-2}\nabla\overline{u}$, which is involved in the motion of glaciers [29] or flows of incompressible turbulent fluids through porous media [11].

The numerical approximation of these models has been extensively studied in the literature (see the fundamental work on the Stefan problem [34], and [35] for a review of some numerical approximations, see [32] for the Richards problem and see [12, 16] and references therein for some studies of convergence of numerical methods for the Leray-Lions problem). However, the convergence analysis of the considered schemes received a much reduced coverage and consists mostly in

*School of Mathematical Sciences, Monash University, Victoria 3800, Australia. jerome.droniou@monash.edu.

†Université Paris-Est, Laboratoire d'Analyse et de Mathématiques Appliquées, UMR 8050, 5 boulevard Descartes, Champs-sur-Marne 77454 Marne-la-Vallée Cedex 2, France. Robert.Eymard@univ-mlv.fr.

establishing space-time averaged results (e.g. in $L^2(\Omega \times (0,T))$), see for example [23, 26]). Yet, the quantity of interest is often not $\overline{u}$ on $\Omega \times (0,T)$ but $\overline{u}$ at a given time, for example $t = T$. Existing numerical analysis results therefore do not ensure that this quantity of interest is indeed properly approximated by numerical methods.
The usual way to obtain pointwise-in-time approximation results for numerical schemes is to prove estimates in $L^\infty(0,T;L^2(\Omega))$ on $u - \overline{u}$, where u is the approximated solution. Establishing such error estimates is however only feasible when uniqueness of the solution $\overline{u}$ to (1) can be proved, which is the case for Richards' and Stefan's problems but not for more complex non-linear parabolic problems as (1) or even p-Laplace problems. It moreover requires some regularity assumptions on $\overline{u}$, which clearly fail for (1) (and simpler p-Laplace problems) for which, because of the possible plateaux of β and ζ, the solution can develop jumps in its gradient.
The purpose of this article is to prove that, using Discrete Functional Analysis techniques (i.e. the translation to numerical analysis of nonlinear analysis techniques), an $L^\infty(0,T;L^2(\Omega))$ convergence result can be established for numerical approximations of (1), without having to assume non-physical regularity assumptions on the data. Note that, although Richards' and Stefan's models are formally equivalent when β and ζ are strictly increasing (consider $\beta = \zeta^{-1}$ to pass from one model to the other), they change nature when these functions are allowed to have plateaux. Stefan's model can degenerate to an ODE (if ζ is constant on the range of the solution) and Richards' model can become a non-transient elliptic equation (if β is constant on this range). The innovative technique we develop in this paper is nonetheless generic enough to work directly on (1) and with a vast number of numerical methods.
That being said, a particular numerical framework must be selected in order to write precise equations and estimates. The framework we choose is that of Gradient Schemes, which has the double benefit of covering a vast number of numerical methods and of having already been studied for many models – elliptic, parabolic, linear or non-linear, possibly degenerate, etc. – with various boundary conditions. The schemes or family of schemes included in the Gradient Schemes framework, and to which our results therefore directly apply, currently are:

- Galerkin methods, including conforming Finite Element schemes,
- Finite Element with mass lumping [7],
- The Crouzeix-Raviart non-conforming Finite Element, with or without mass lumping [9, 20],
- The Raviart-Thomas Mixed Finite Elements [5],
- The Vertex Approximate Gradient scheme [24],
- The Hybrid Mimetic Mixed family [15], which includes Mimetic Finite Differences [6], Mixed Finite Volume [13] and the SUSHI scheme [22],
- The Discrete Duality Finite Volume scheme in dimension 2 [30, 2], and the CeVeFE-Discrete Duality Finite Volume scheme in dimension 3 [8],
- The Multi-Point Flux Approximation O-method [1, 18].

We refer the reader to [14, 16, 21, 27, 25] for more details. Let us finally emphasize that the unified convergence study of numerical schemes for Problem (1), which combines a general Leray-Lions operator and nonlinear functions β or ζ, seems to be new even without the uniform-in-time convergence result.

The paper is organised as follows. In Section 1.2, we present the assumptions and the notion of weak solution for (1) and, in Section 1.3, we give an overview of the ideas involved in the proof

of uniform-in-time convergence. This overview is given not in a numerical analysis context but in in a context of pure stability analysis of (1) with very little regularity on the data, for which the uniform-in-time convergence result also seems to be new. Section 2 presents the Gradient Schemes for our generic model (1). We give in Section 3 some preliminaries to the convergence study. Section 4 contains the complete convergence proof of Gradient Schemes for (1), including the uniform-in-time convergence result. An appendix, Section 5, concludes the article with important technical results, and in particular a generalisation of Ascoli-Arzela compactness result to discontinuous functions and a characterisation of uniform convergence of sequences of functions which is critical to establishing our uniform-in-time convergence result.
Note that these results and their proofs have been sketched and illustrated by some numerical examples in [17], for $\boldsymbol{a}(\boldsymbol{x},\nu(\overline{u}),\nabla\zeta(\overline{u}))=\nabla\zeta(\overline{u})$.

1.2 Hypotheses and weak sense for the continuous problem

We consider the evolution problem (1) under the following hypotheses.

$$\Omega \text{ is an open bounded connected polyhedral subset of } \mathbb{R}^d \ (d\in\mathbb{N}^\star) \text{ and } T>0, \tag{2a}$$

$$\begin{array}{l}\zeta\in C^0(\mathbb{R}) \text{ is non--decreasing, Lipschitz continuous with Lipschitz constant } L_\zeta>0\\ \text{such that } \zeta(0)=0 \text{ and, for some } M_0,M_1>0,\ |\zeta(s)|\ge M_0|s|-M_1 \text{ for all } s\in\mathbb{R}.\end{array} \tag{2b}$$

$$\begin{array}{l}\beta \text{ is a non-decreasing Lipschitz continuous function with Lipschitz constant } L_\beta>0\\ \text{and } \beta(0)=0.\end{array} \tag{2c}$$

$$\forall s\in\mathbb{R}\,,\quad \nu(s)=\int_0^s \zeta'(q)\beta'(q)\mathrm{d}q. \tag{2d}$$

$$\boldsymbol{a}\ :\ \Omega\times\mathbb{R}\times\mathbb{R}^d\to\mathbb{R}^d,\ \text{with } p\in(1,+\infty),\ \text{is a Caratheodory function,} \tag{2e}$$

(i.e. a function such that, for a.e. $\boldsymbol{x}\in\Omega$, $(s,\boldsymbol{\xi})\mapsto\boldsymbol{a}(\boldsymbol{x},s,\boldsymbol{\xi})$ is continuous and, for any $(s,\boldsymbol{\xi})\in\mathbb{R}\times\mathbb{R}^d$, $\boldsymbol{x}\mapsto\boldsymbol{a}(\boldsymbol{x},s,\boldsymbol{\xi})$ is measurable)

$$\exists\underline{a}\in(0,+\infty)\ :\ \boldsymbol{a}(\boldsymbol{x},s,\boldsymbol{\xi})\cdot\boldsymbol{\xi}\ge\underline{a}|\boldsymbol{\xi}|^p,\ \text{for a.e. } \boldsymbol{x}\in\Omega,\ \forall s\in\mathbb{R},\ \forall\boldsymbol{\xi}\in\mathbb{R}^d, \tag{2f}$$

$$(\boldsymbol{a}(\boldsymbol{x},s,\boldsymbol{\xi})-\boldsymbol{a}(\boldsymbol{x},s,\boldsymbol{\chi}))\cdot(\boldsymbol{\xi}-\boldsymbol{\chi})\ge 0,\ \text{for a.e. } \boldsymbol{x}\in\Omega,\ \forall s\in\mathbb{R},\ \forall\boldsymbol{\xi},\boldsymbol{\chi}\in\mathbb{R}^d, \tag{2g}$$

$$\begin{array}{l}\exists\overline{a}\in L^{p'}(\Omega)\,,\ \exists\mu\in(0,+\infty)\ :\\ \quad|\boldsymbol{a}(\boldsymbol{x},s,\boldsymbol{\xi})|\le\overline{a}(\boldsymbol{x})+\mu|\boldsymbol{\xi}|^{p-1},\ \text{for a.e. } \boldsymbol{x}\in\Omega,\ \forall s\in\mathbb{R},\ \forall\boldsymbol{\xi}\in\mathbb{R}^d,\end{array} \tag{2h}$$

and

$$u_{\rm ini}\in L^2(\Omega),\quad f\in L^{p'}(\Omega\times(0,T)). \tag{2i}$$

We denote by R_β the range of β and define the pseudo-inverse function $\beta_r:R_\beta\to\mathbb{R}$ of β by

$$\begin{aligned}\forall s\in R_\beta\,,\ \beta_r(s)\ &=\ \begin{cases}\inf\{t\in\mathbb{R}\,|\,\beta(t)=s\} & \text{if } s\ge 0,\\ \sup\{t\in\mathbb{R}\,|\,\beta(t)=s\} & \text{if } s<0,\end{cases}\\ &=\ \text{closest } t \text{ to } 0 \text{ such that } \beta(t)=s.\end{aligned} \tag{3}$$

Since $\beta(0)=0$, we notice that $\beta_r\ge 0$ on $R_\beta\cap\mathbb{R}^+$ and $\beta_r\le 0$ on $R_\beta\cap\mathbb{R}^-$. We then define $B:R_\beta\to[0,\infty]$ by

$$B(z)=\int_0^z\zeta(\beta_r(s))\,ds.$$

Since β_r is non-decreasing, this expression is always well-defined in $[0,\infty)$. The signs of β_r also ensure that that B is non-decreasing on $R_\beta\cap\mathbb{R}^+$ and non-increasing on $R_\beta\cap R^-$. We can therefore extend B to $\overline{R_\beta}$ by these limits (possibly $+\infty$) at the potential endpoints of R_β.

The precise notion of solution to (1) that we consider is then the following:

$$
\left\{\begin{array}{l}
\overline{u} \in L^p(0,T;L^p(\Omega))\,,\ \zeta(\overline{u}) \in L^p(0,T;W_0^{1,p}(\Omega))\,,\\
B(\beta(\overline{u})) \in L^\infty(0,T;L^1(\Omega)),\ \beta(\overline{u}) \in C([0,T];L^2(\Omega)\text{-w}), \partial_t\beta(\overline{u}) \in L^{p'}(0,T;W^{-1,p'}(\Omega)),\\
\beta(\overline{u})(\cdot,0) = \beta(u_{\rm ini}) \text{ in } L^2(\Omega),\\
\displaystyle\int_0^T \langle \partial_t\beta(\overline{u})(\cdot,t), \overline{v}(\cdot,t)\rangle_{W^{-1,p'},W_0^{1,p}} {\rm d}t\\
\quad\displaystyle + \int_0^T\int_\Omega \boldsymbol{a}(\boldsymbol{x},\nu(\overline{u}(\boldsymbol{x},t)),\nabla\zeta(\overline{u})(\boldsymbol{x},t))\cdot\nabla\overline{v}(\boldsymbol{x},t){\rm d}\boldsymbol{x}{\rm d}t = \int_0^T\int_\Omega f(\boldsymbol{x},t)\overline{v}(\boldsymbol{x},t){\rm d}\boldsymbol{x}{\rm d}t\,,\\
\quad\forall \overline{v} \in L^p(0;T;W_0^{1,p}(\Omega)).
\end{array}\right. \tag{4}
$$

where $C([0,T];L^2(\Omega)\text{-w})$ denotes the space of continuous functions $[0,T] \mapsto L^2(\Omega)$ for the weak-$*$ topology of $L^2(\Omega)$. Here and in the following, we denote by p' the dual exponent $\frac{p}{p-1}$ to p and we remove the mention of Ω in the duality bracket $\langle\cdot,\cdot\rangle_{W^{-1,p'},W_0^{1,p}} = \langle\cdot,\cdot\rangle_{W^{-1,p'}(\Omega),W_0^{1,p}(\Omega)}$.

Remark 1.1 *The derivative $\partial_t\beta(\overline{u})$ is to be understood in the usual sense of distributions on $\Omega\times(0,T)$. Since the set $\mathcal{T} = \{\sum_{i=1}^q \varphi_i(t)\gamma_i(\boldsymbol{x}) \,:\, q\in\mathbb{N}, \varphi_i \in C_c^\infty(0,T), \gamma_i \in C_c^\infty(\Omega)\}$ of tensorial functions in $C^\infty(\Omega\times(0,T))$ is dense in $L^p(0,T;W_0^{1,p}(\Omega))$, one can ensure that this distribution derivative $\partial_t\beta(\overline{u})$ belongs to $L^{p'}(0,T;W^{-1,p'}(\Omega)) = (L^p(0,T;W_0^{1,p}(\Omega))'$ by checking that the linear form*

$$\varphi\in\mathcal{T} \mapsto \langle\partial_t\beta(\overline{u}),\varphi\rangle_{\mathcal{D}',\mathcal{D}} = -\int_0^T\int_\Omega \beta(\overline{u})(\boldsymbol{x},t)\partial_t\varphi(\boldsymbol{x},t){\rm d}\boldsymbol{x}{\rm d}t$$

is continuous for the norm of $L^p(0,T;W_0^{1,p}(\Omega))$.

Note that the continuity property of $\beta(\overline{u})$ natural. Indeed, the PDE in the sense of distributions shows that $T_\varphi : t \mapsto \langle\beta(\overline{u})(t),\varphi\rangle_{L^2}$ belongs to $W^{1,1}(0,T)$, and is therefore continuous, for any $\varphi\in C_c^\infty(\Omega)$. The density in $L^2(\Omega)$ of such φ, combined with the fact that $\beta(\overline{u})\in L^\infty(0,T;L^2(\Omega))$ (coming from $B(\beta(\overline{u}))\in L^\infty(0,T;L^1(\Omega))$ and (26)), proves the continuity of T_φ for any $\varphi\in L^2(\Omega)$, that is to say the continuity of $\beta(\overline{u}) : [0,T]\to L^2(\Omega)$-w.
This notion of $\beta(\overline{u})$ as a function continuous in time is nevertheless a subtle one. It is to be understood in the sense that the function $(\boldsymbol{x},t)\mapsto\beta(\overline{u}(\boldsymbol{x},t))$ has an a.e. representative which is continuous $[0,T]\mapsto L^2(\Omega)$-w. In other words, there is a function $Z\in C([0,T];L^2(\Omega)\text{-w})$ such that $Z(t)(\boldsymbol{x}) = \beta(\overline{u}(\boldsymbol{x},t))$ for a.e. $(\boldsymbol{x},t)\in\Omega\times(0,T)$. We must however make sure, when dealing with pointwise values in time to separate Z from $\beta(\overline{u}(\cdot,\cdot))$ as $\beta(\overline{u}(\cdot,t_1))$ may not make sense for a particular $t_1\in[0,T]$.
That being said, in order to adopt a simple notation, we will denote by $\beta(\overline{u})(\cdot,\cdot)$ the function Z, and by $\beta(\overline{u}(\cdot,\cdot))$ the a.e.-defined composition of β and $\overline{u}$. Hence, it will make sense to talk about $\beta(\overline{u})(\cdot,t)$ for a particular $t_1\in[0,T]$, and we will only write $\beta(\overline{u})(\boldsymbol{x},t) = \beta(\overline{u}(\boldsymbol{x},t))$ for a.e. $(\boldsymbol{x},t)\in\Omega\times(0,T)$. Note that from this a.e. equality we can ensure that $\beta(\overline{u})(\cdot,\cdot)$ takes its values in the closure $\overline{R_\beta}$ of the range of β.

1.3 General principle for the uniform-in-time convergence result

As explained in the introduction, the main innovative result of this article is the uniform-in-time convergence result (Theorem 2.14 below). Although it's stated and proved in the context of numerical approximations of (1), we emphasize that its principle is also applicable to theoretical analysis of PDEs. Let us informally present this principle on the following continuous approximation of (1):

$$
\begin{array}{l}
\partial_t\beta(\overline{u}_\varepsilon) - {\rm div}\left(\boldsymbol{a}_\varepsilon(\boldsymbol{x},\nu(\overline{u}_\varepsilon),\nabla\zeta(\overline{u}_\varepsilon))\right) = f \text{ in } \Omega\times(0,T),\\
\beta(\overline{u}_\varepsilon)(\boldsymbol{x},0) = \beta(u_{\rm ini})(\boldsymbol{x}) \text{ in } \Omega,\\
\zeta(\overline{u}_\varepsilon) = 0 \text{ on } \partial\Omega\times(0,T)
\end{array} \tag{5}
$$

where $\boldsymbol{a}_\varepsilon$ satisfies Assumptions (2e)–(2h) with constants not depending on ε and $\boldsymbol{a}_\varepsilon \to \boldsymbol{a}$ pointwise as $\varepsilon \to 0$.
We want to show here how to deduce from averaged convergences a strong uniform-in-time convergence result. We therefore assume the following convergences (up to a subsequence as $\varepsilon \to 0$), which are compatible with basic compactness results that can be obtained on $(\overline{u}_\varepsilon)_\varepsilon$ and also correspond to the initial convergences (18) that can be obtained on numerical approximations of (1):

$$\begin{aligned}&\beta(\overline{u}_\varepsilon) \to \beta(\overline{u}) \text{ in } C([0,T];L^2(\Omega)\text{-w})\,,\ \nu(\overline{u}_\varepsilon) \to \nu(\overline{u}) \text{ strongly in } L^1(\Omega\times(0,T)),\\ &\zeta(\overline{u}_\varepsilon) \to \zeta(\overline{u}) \text{ weakly in } L^p(0,T;W_0^{1,p}(\Omega))\,,\\ &\boldsymbol{a}_\varepsilon(\cdot,\nu(\overline{u}_\varepsilon),\nabla\zeta(\overline{u}_\varepsilon)) \to \boldsymbol{a}(\cdot,\nu(\overline{u}),\nabla\zeta(\overline{u})) \text{ weakly in } L^p(\Omega\times(0,T))^d.\end{aligned} \tag{6}$$

We will prove from these convergences that, along the same subsequence, $\nu(\overline{u}_\varepsilon) \to \nu(\overline{u})$ strongly in $C([0,T];L^2(\Omega))$, which is our uniform-in-time convergence result.
We start by noticing that the weak-in-space uniform-in-time convergence of $\beta(\overline{u}_\varepsilon)$ gives, for any $T_0 \in [0,T]$ and any family $(T_\varepsilon)_{\varepsilon>0}$ converging to T_0 as $\varepsilon \to 0$, $\beta(\overline{u}_\varepsilon)(T_\varepsilon,\cdot) \to \beta(\overline{u})(T_0,\cdot)$ weakly in $L^2(\Omega)$. Classical strong-weak semi-continuity properties of convex functions (see Lemma 3.4) and the convexity of B (see Lemma 3.3) then ensure that

$$\int_\Omega B(\beta(\overline{u})(\boldsymbol{x},T_0))\mathrm{d}\boldsymbol{x} \le \liminf_{\varepsilon\to 0}\int_\Omega B(\beta(\overline{u}_\varepsilon)(\boldsymbol{x},T_\varepsilon))\mathrm{d}\boldsymbol{x}. \tag{7}$$

The second step is to notice that, by (2g) for $\boldsymbol{a}_\varepsilon$,

$$\int_0^{T_\varepsilon}\int_\Omega [\boldsymbol{a}_\varepsilon(\cdot,\nu(\overline{u}_\varepsilon),\nabla\zeta(\overline{u}_\varepsilon) - \boldsymbol{a}_\varepsilon(\cdot,\nu(\overline{u}_\varepsilon),\nabla\zeta(\overline{u}))]\cdot[\nabla\zeta(\overline{u}_\varepsilon) - \nabla\zeta(\overline{u})]\,\mathrm{d}\boldsymbol{x}\mathrm{d}t \ge 0.$$

Developing this expression and using the convergences (6), we find that

$$\liminf_{\varepsilon\to 0}\int_0^{T_\varepsilon}\int_\Omega \boldsymbol{a}_\varepsilon(\cdot,\nu(\overline{u}_\varepsilon),\nabla\zeta(\overline{u}_\varepsilon)\cdot\nabla\zeta(\overline{u}_\varepsilon)(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x}\mathrm{d}t \ge \int_0^{T_0}\int_\Omega \boldsymbol{a}(\cdot,\nu(\overline{u}),\nabla\zeta(\overline{u}))\cdot\nabla\zeta(\overline{u})\mathrm{d}\boldsymbol{x}\mathrm{d}t. \tag{8}$$

We then establish the following formula:

$$\begin{aligned}\int_\Omega B(\beta(\overline{u}_\varepsilon(\boldsymbol{x},T_\varepsilon)))\mathrm{d}\boldsymbol{x} + \int_0^{T_\varepsilon}\int_\Omega \boldsymbol{a}_\varepsilon(\boldsymbol{x},\nu(\overline{u}_\varepsilon(\boldsymbol{x},t)),\nabla\zeta(\overline{u}_\varepsilon)(\boldsymbol{x},t))\cdot\nabla\zeta(\overline{u}_\varepsilon)(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x}\mathrm{d}t\\ = \int_\Omega B(\beta(u_{\rm ini}(\boldsymbol{x})))\mathrm{d}\boldsymbol{x} + \int_0^{T_\varepsilon}\int_\Omega f(\boldsymbol{x},t)\zeta(\overline{u}_\varepsilon)(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x}\mathrm{d}t.\end{aligned} \tag{9}$$

This energy estimate is formally obtained by multiplying (5) by $\zeta(\overline{u}_\varepsilon)$ and integrating by parts, using the fact that $(B\circ\beta)' = \zeta\beta'$ (see Lemma 3.3). The rigorous justification of (9) is however quite technical, see Lemma 3.6 and Corollary 3.8. Thanks to (8), we can pass to the lim sup in (9) and we find, using the same energy estimate with $(\overline{u},\boldsymbol{a},T_0)$ instead of $(\overline{u}_\varepsilon,\boldsymbol{a}_\varepsilon,T_\varepsilon)$,

$$\limsup_{\varepsilon\to 0}\int_\Omega B(\beta(\overline{u}_\varepsilon(\boldsymbol{x},T_\varepsilon)))\mathrm{d}\boldsymbol{x} \le \int_\Omega B(\beta(\overline{u}(\boldsymbol{x},T_0)))\mathrm{d}\boldsymbol{x}. \tag{10}$$

Combined with (7), this shows that $\int_\Omega B(\beta(\overline{u}_\varepsilon(\boldsymbol{x},T_\varepsilon)))\mathrm{d}\boldsymbol{x} \to \int_\Omega B(\beta(\overline{u}(\boldsymbol{x},T_0)))\mathrm{d}\boldsymbol{x}$. The uniform convexity of B (see (28)) then allows us to deduce that $\nu(\overline{u}_\varepsilon(\cdot,T_\varepsilon)) \to \nu(\overline{u}(\cdot,T_0))$ strongly in $L^2(\Omega)$ and thus that $\nu(\overline{u}_\varepsilon) \to \nu(\overline{u})$ strongly in $C([0,T];L^2(\Omega))$ (see Lemma 6.3).

Remark 1.2 *A close examination of this proof indicates that equality in the energy estimate* (9) *is not required for* $\overline{u}_\varepsilon$. *An inequality* $\le$ *would be sufficient. This is particularly important in the*

context of numerical methods which may introduce additional numerical diffusion (for example due to an implicit-in-time discretisation) and therefore only provide an upper bound in this energy estimate, see for example Estimate (39)*.*
It is however crucial that the limit solution $\overline{u}$ satisfies the equivalent of (9) *with an equal sign (or $\geq$).*

2 Gradient discretisations and gradient schemes

2.1 Definitions

We give here a minimal presentation of gradient discretisations and gradient schemes, limiting ourselves to what is necessary to study the discretisation of (1). We refer the reader to [14, 24, 16] for more details.
A gradient scheme can be viewed as a general formulation of several discretisations of (1) which are based on a nonconforming approximation of the weak formulation of the problem. The approximation of the weak formulation of (1) is based on some discrete spaces and mappings, the set of which we call a gradient discretisation. Throughout this paper, Ω is an open bounded subset of $\mathbb{R}^d$, $d \in \mathbb{N}^\star$, and $p \in (1, +\infty)$.

Definition 2.1 (Space-Time gradient discretisation for homogeneous Dirichlet boundary conditions)
We say that $\mathcal{D} = (X_{\mathcal{D},0}, \Pi_{\mathcal{D}}, \nabla_{\mathcal{D}}, \mathcal{I}_{\mathcal{D}}, (t^{(n)})_{n=0,\ldots,N})$ is a space-time gradient discretisation for homogeneous Dirichlet boundary conditions if

1. *the set of discrete unknowns $X_{\mathcal{D},0}$ is a finite dimensional real vector space,*
2. *the linear mapping $\Pi_{\mathcal{D}} : X_{\mathcal{D},0} \to L^\infty(\Omega)$ is a piecewise constant reconstruction operator in the sense that there exists a set I of degrees of freedom such that $X_{\mathcal{D},0} = \mathbb{R}^I$ and there exists a family $(\Omega_i)_{i\in I}$ of disjoint subsets of Ω such that $\overline{\Omega} = \bigcup_{i\in I} \overline{\Omega_i}$ and, for all $u = (u_i)_{i\in I} \in X_{\mathcal{D},0}$ and all $i \in I$,$\Pi_{\mathcal{D}} u = u_i$ on Ω_i,*
3. *the linear mapping $\nabla_{\mathcal{D}} : X_{\mathcal{D},0} \to L^p(\Omega)^d$ gives a reconstructed discrete gradient. It must be chosen such that $\|\nabla_{\mathcal{D}} \cdot \|_{L^p(\Omega)^d}$ is a norm on $X_{\mathcal{D},0}$,*
4. *$\mathcal{I}_{\mathcal{D}} : L^2(\Omega) \to X_{\mathcal{D},0}$ is a linear interpolation operator,*
5. *$t^{(0)} = 0 < t^{(1)} < t^{(2)} < \ldots < t^{(N)} = T$.*

We then set $\delta t^{(n+\frac{1}{2})} = t^{(n+1)} - t^{(n)}$, for $n = 0, \ldots, N-1$, and $\delta t_{\mathcal{D}} = \max_{n=0,\ldots,N-1} \delta t^{(n+\frac{1}{2})}$, and we define the dual semi-norm $|w|_{\star,\mathcal{D}}$ of $w \in X_{\mathcal{D},0}$ by

$$|w|_{\star,\mathcal{D}} = \sup\left\{ \int_\Omega \Pi_{\mathcal{D}} w(\boldsymbol{x}) \Pi_{\mathcal{D}} z(\boldsymbol{x}) \mathrm{d}\boldsymbol{x} \,:\, z \in X_{\mathcal{D},0}\,,\ ||\nabla_{\mathcal{D}} z||_{L^p(\Omega)^d} = 1 \right\}. \tag{11}$$

Remark 2.2 (Boundary conditions) *Other boundary conditions can be seamlessly handled by Gradient Schemes, see [14].*

Remark 2.3 (Nonlinear function of the elements of $X_{\mathcal{D},0}$) *Let $\mathcal{D}$ be a gradient discretisation in the sense of Definition 2.1. For any $\chi : \mathbb{R} \mapsto \mathbb{R}$ and any $u = (u_i)_{i\in I} \in X_{\mathcal{D},0}$, we define $\chi_I(u) \in X_{\mathcal{D},0}$ by $\chi_I(u) = (\chi_I(u)_i)_{i\in I}$ with $\chi_I(u)_i = \chi(u_i)$. As indicated by the subscript I, this definition depends on the choice of the degrees of freedom in $X_{\mathcal{D},0}$. That said, these degrees of* *freedom are usually canonical and we therefore drop the index I. An important consequence of the fact that* $\Pi_{\mathcal{D}}$ *is a* piecewise constant *reconstruction is the following:*

$$\forall \chi : \mathbb{R} \mapsto \mathbb{R}\,,\ \forall u \in X_{\mathcal{D},0}\,,\quad \Pi_{\mathcal{D}}\chi(u) = \chi(\Pi_{\mathcal{D}}u). \tag{12}$$

It is customary to also use the notations $\Pi_{\mathcal{D}}$ and $\nabla_{\mathcal{D}}$ for space-time dependent functions. We will also need a notation for the jump-in-time of piecewise constant functions in time. Hence, if $(v^{(n)})_{n=0,\ldots,N} \subset X_{\mathcal{D},0}$, we set

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

Thanks to Remark 2.3, the related gradient scheme is merely the discretisation of the weak formulation of (1) obtained by using the discrete space and mappings of the gradient discretisation. If $\mathcal{D} = (X_{\mathcal{D},0},\Pi_{\mathcal{D}},\nabla_{\mathcal{D}},\mathcal{I}_{\mathcal{D}},(t^{(n)})_{n=0,\ldots,N})$ is a space-time gradient discretisation in the sense of Definition 2.1, we define the following gradient scheme for Problem (1): we consider a sequence $(u^{(n)})_{n=0,\ldots,N} \subset X_{\mathcal{D},0}$ such that

$$\left\{\begin{aligned}
&u^{(0)} = \mathcal{I}_{\mathcal{D}}u_{\rm ini} \text{ and, for all } v = (v^{(n)})_{n=1,\ldots,N} \subset X_{\mathcal{D},0},\\
&\int_0^T\!\!\int_\Omega [\Pi_{\mathcal{D}}\delta_{\mathcal{D}}\beta(u)(\boldsymbol{x},t)\Pi_{\mathcal{D}}v(\boldsymbol{x},t) + \boldsymbol{a}\left(\boldsymbol{x},\Pi_{\mathcal{D}}\nu(u)(\boldsymbol{x},t),\nabla_{\mathcal{D}}\zeta(u)(\boldsymbol{x},t)\right)\cdot\nabla_{\mathcal{D}}v(\boldsymbol{x},t)]\,\mathrm{d}\boldsymbol{x}\mathrm{d}t\\
&\qquad = \int_0^T\!\!\int_\Omega f(\boldsymbol{x},t)\Pi_{\mathcal{D}}v(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x}\mathrm{d}t.
\end{aligned}\right. \tag{14}$$

Remark 2.4 *We could as well consider, instead of a fully implicit method, a Crank-Nicolson scheme or any scheme in between those two. Such schemes are defined by taking* $\alpha \in [\frac{1}{2},1]$ *and replacing the terms* $u^{(n+1)}$ *appearing in* $\boldsymbol{a}(\boldsymbol{x},\cdot,\cdot)$ *in* (14) *for* $t \in (t^{(n)},t^{(n+1)}]$ *with* $u^{(n+\alpha)} = \alpha u^{(n+1)} + (1-\alpha)u^{(n)}$. *All the results we establish for* (14) *would hold for such a scheme (see the treatment in [16]).*

2.2 Properties of gradient discretisations

In order to establish the convergence of the associated gradient schemes, sequences of space-time gradient discretisations are required to satisfy four properties: *coercivity*, *consistency*, *limit-conformity* and *compactness*.

Definition 2.5 (Coercivity) *A sequence* $(\mathcal{D}_m)_{m\in\mathbb{N}}$ *of space-time gradient discretisations in the sense of Definition 2.1 is said to be coercive if there exists* $C_{\mathcal{D}}$ *such that, for any* $m \in \mathbb{N}$ *and any* $v \in X_{\mathcal{D}_m,0}$, $||\Pi_{\mathcal{D}_m}v||_{L^p(\Omega)} \le C_{\mathcal{D}}||\nabla_{\mathcal{D}_m}v||_{L^p(\Omega)^d}$.

Definition 2.6 (Consistency) *If* $\mathcal{D}$ *is a space-time gradient discretisation in the sense of Definition 2.1, we define*

$$\forall\varphi \in L^2(\Omega)\cap W_0^{1,p}(\Omega),\ \widehat{S}_{\mathcal{D}}(\varphi) = \min_{w\in X_{\mathcal{D},0}}\left(||\Pi_{\mathcal{D}}w - \varphi||_{L^{\max(p,2)}(\Omega)} + ||\nabla_{\mathcal{D}}w - \nabla\varphi||_{L^p(\Omega)}\right). \tag{15}$$

A sequence $(\mathcal{D}_m)_{m\in\mathbb{N}}$ *of space-time gradient discretisations in the sense of Definition 2.1 is said to be consistent if*

- *for all* $\varphi \in L^2(\Omega)\cap W_0^{1,p}(\Omega)$, $\widehat{S}_{\mathcal{D}_m}(\varphi) \to 0$ *as* $m \to \infty$,

- *for all* $\varphi \in L^2(\Omega)$, $\Pi_{\mathcal{D}_m}\mathcal{I}_{\mathcal{D}_m}\varphi \to \varphi$ *in* $L^2(\Omega)$ *as* $m \to \infty$, *and*
- $\eth_{\mathcal{D}_m} \to 0$ *as* $m \to \infty$.

Definition 2.7 (Compactness) *If* $\mathcal{D}$ *is a space-time gradient discretisation in the sense of Definition 2.1, we define*

$$\forall \boldsymbol{\xi} \in \mathbb{R}^d\,,\ T_{\mathcal{D}}(\boldsymbol{\xi}) = \max_{v\in X_{\mathcal{D},0}\setminus\{0\}} \frac{||\Pi_{\mathcal{D}}v(\cdot+\boldsymbol{\xi}) - \Pi_{\mathcal{D}}v||_{L^p(\mathbb{R}^d)}}{||\nabla_{\mathcal{D}}v||_{L^p(\Omega)}},$$

where $\Pi_{\mathcal{D}}v$ *has been extended by* 0 *outside* Ω.
A sequence $(\mathcal{D}_m)_{m\in\mathbb{N}}$ *of space-time gradient discretisations is said to be compact if*

$$\lim_{\boldsymbol{\xi}\to 0} \sup_{m\in\mathbb{N}} T_{\mathcal{D}}(\boldsymbol{\xi}) = 0.$$

Definition 2.8 (Limit-conformity) *If* $\mathcal{D}$ *is a space-time gradient discretisation in the sense of Definition 2.1, we define*

$$\forall \boldsymbol{\varphi} \in W^{\mathrm{div},p'}(\Omega)\,,\ W_{\mathcal{D}}(\boldsymbol{\varphi}) = \max_{u\in X_{\mathcal{D},0}\setminus\{0\}} \frac{\left|\displaystyle\int_\Omega (\nabla_{\mathcal{D}}u(\boldsymbol{x})\cdot\boldsymbol{\varphi}(\boldsymbol{x}) + \Pi_{\mathcal{D}}u(\boldsymbol{x})\mathrm{div}\boldsymbol{\varphi}(\boldsymbol{x}))\,\mathrm{d}\boldsymbol{x}\right|}{\|\nabla_{\mathcal{D}}u\|_{L^p(\Omega)^d}}. \tag{16}$$

A sequence $(\mathcal{D}_m)_{m\in\mathbb{N}}$ *of space-time gradient discretisations in the sense of Definition 2.1 is said to be limit-conforming if, for all* $\boldsymbol{\varphi} \in W^{\mathrm{div},p'}(\Omega)$, $W_{\mathcal{D}_m}(\boldsymbol{\varphi})$ *tends to 0 as* $m \to \infty$.

We refer the reader to [16, 14] for a proof of the following lemma.

Lemma 2.9 (Consequence of limit-conformity) *Let* $(\mathcal{D}_m)_{m\in\mathbb{N}}$ *be a sequence of space-time gradient discretisations in the sense of Definition 2.1, which is* limit-conforming *in the sense of Definition 2.8. Let, for any* $m \in \mathbb{N}$, $v_m = (v_m^{(n)})_{n=0,\ldots,N_m} \subset X_{\mathcal{D}_m,0}$ *be such that, with the notations in* (13), $(\nabla_{\mathcal{D}_m}v_m)_{m\in\mathbb{N}}$ *is bounded in* $L^p(\Omega\times(0,T))$.
Then there exists $v \in L^p(0,T;W_0^{1,p}(\Omega))$ *such that, up to a subsequence as* $m \to \infty$, $\Pi_{\mathcal{D}_m}v_m \to v$ *weakly in* $L^p(\Omega\times(0,T))$ *and* $\nabla_{\mathcal{D}_m}v_m \to \nabla v$ *weakly in* $L^p(\Omega\times(0,T))^d$.

2.3 Main results

Our first theorem states weak or space-time averaged convergence properties of Gradient Schemes for (1). These results are quite classical and have already been established for Richards' and Stefan's models, see [21, 25]. The convergence proof we provide afterwards however covers more non-linear model, as $\boldsymbol{a}$ do not need to be linear with respect to $\boldsymbol{\xi}$, and is more compact than the ones available in the literature.

Theorem 2.10 (Convergence of the Gradient Scheme) *Under Assumptions* (2), *let* $(\mathcal{D}_m)_{mp\in\mathbb{N}}$ *be a sequence of space-time gradient discretisations in the sense of Definition 2.1, which is coercive, consistent, compact and limit-conforming (see Section 2.2).*
Then for any $m \in \mathbb{N}$ *there exists a solution* u_m *to* (14) *with* $\mathcal{D} = \mathcal{D}_m$.
Moreover, if we assume that

$$(\forall s \in \mathbb{R}\,,\ \beta(s) = s) \quad \textit{or} \quad (\forall s \in \mathbb{R}\,,\ \zeta(s) = s), \tag{17}$$

then there exists a solution $\overline{u}$ *to* (4) *such that, up to a subsequence, the following convergences hold as* $m \to \infty$*:*

$$\begin{array}{l}\Pi_{\mathcal{D}_m}\beta(u_m) \to \beta(\overline{u}) \text{ weakly in } L^2(\Omega) \text{ uniformly on } [0,T] \text{ (see Definition 6.4)},\\ \Pi_{\mathcal{D}_m}\nu(u_m) \to \nu(\overline{u}) \text{ strongly in } L^1(0,T;L^1(\Omega)),\\ \Pi_{\mathcal{D}_m}\zeta(u_m) \to \zeta(\overline{u}) \text{ weakly in } L^p(\Omega\times(0,T)).\\ \nabla_{\mathcal{D}_m}\zeta(u_m) \to \nabla\zeta(\overline{u}) \text{ weakly in } L^p(\Omega\times(0,T))^d.\end{array} \tag{18}$$

Remark 2.11 *Since* $|\nu| \le L_\zeta|\beta|$ *and* $|\nu| \le L_\beta|\zeta|$*, the* $L^\infty(0,T;L^2(\Omega))$ *bound on* $\beta(u_m)$ *and the* $L^p(\Omega\times(0,T))$ *bound on* $\zeta(u_m)$ *(see Lemma 4.1 and Definition 2.5) allows us to see that the strong convergence of* $\Pi_{\mathcal{D}_m}\nu(u_m)$ *is also valid in* $L^q(0,T;L^r(\Omega))$ *for any* (q,r) *such that* $q<\infty$ *and* $r<2$ *or* $q,r<p$ *(and, of course, any space interpolated between the two cases).*

Remark 2.12 *Note that we do not assume the existence of a solution* $\overline{u}$ *to the continuous problem, our convergence analysis will establish this existence.*

Remark 2.13 *Assumption* (17) *covers Richards' and Stefan's models, as well as many other non-linear parabolic equations. This assumption is actually not mandatory if* $p \ge 2$*, see Section 5. We decide however to first state and prove Theorem 2.10 with this assumption to simplify the presentation. See also Remark 2.17.*

The main innovation of this paper is the following theorem, which states the *uniform-in-time* convergence of numerical methods for fully non-linear parabolic equations with no regularity assumptions on the data.

Theorem 2.14 (Uniform-in-time convergence) *Under Assumptions* (2)*, let* $(\mathcal{D}_m)_{mp\in\mathbb{N}}$ *be a sequence of space-time gradient discretisations in the sense of Definition* (2.1)*, which is coercive, consistent, compact and limit-conforming (see Section 2.2).*
We assume that, for any $m \in \mathbb{N}$*,* u_m *is a solution to* (14) *with* $\mathcal{D} = \mathcal{D}_m$*, which converges to a solution* $\overline{u}$ *of* (4) *in the sense* (18)*.*
Then, as $m \to \infty$*,* $\Pi_{\mathcal{D}_m}\nu(u_m) \to \nu(\overline{u})$ *strongly in* $L^\infty(0,T;L^2(\Omega))$*.*

Remark 2.15 *Note that since* $(\Pi_{\mathcal{D}_m}\nu(u_m))_{m\in\mathbb{N}}$ *are piecewise constant in time, their convergence in* $L^\infty(0,T;L^2(\Omega))$ *is actually a uniforme-in-time convergence (not "uniform a.e. in time").*

The last theorem completes our convergence result by stating the strong space-time averaged convergence of the discrete gradients. Its proof is inspired by the study of Gradient Schemes for Leray-Lions operators made in [16].

Theorem 2.16 (Strong convergence of the gradients) *Under Assumptions* (2)*, let* $(\mathcal{D}_m)_{mp\in\mathbb{N}}$ *be a sequence of space-time gradient discretisations in the sense of Definition* (2.1)*, which is coercive, consistent, compact and limit-conforming (see Section 2.2).*
We assume that, for any $m \in \mathbb{N}$*,* u_m *is a solution to* (14) *with* $\mathcal{D} = \mathcal{D}_m$*, which converges to a solution* $\overline{u}$ *of* (4) *in the sense* (18)*. We also assume that* $\boldsymbol{a}$ *is strictly monotone in the sense:*

$$(\boldsymbol{a}(\boldsymbol{x},s,\boldsymbol{\xi}) - \boldsymbol{a}(\boldsymbol{x},s,\boldsymbol{\chi}))\cdot(\boldsymbol{\xi}-\boldsymbol{\chi}) > 0, \text{ for a.e. } \boldsymbol{x}\in\Omega,\ \forall s\in\mathbb{R},\ \forall\boldsymbol{\xi}\neq\boldsymbol{\chi}\in\mathbb{R}^d. \tag{19}$$

Then, as $m\to\infty$*,* $\Pi_{\mathcal{D}_m}\zeta(u_m) \to \zeta(\overline{u})$ *strongly in* $L^p(\Omega\times(0,T))$ *and* $\nabla_{\mathcal{D}_m}\zeta(u_m) \to \nabla\zeta(\overline{u})$ *strongly in* $L^p(\Omega\times(0,T))^d$*.*

Remark 2.17 *Theorems 2.14 and 2.16 do not require the structural assumption* (17)*; these theorems only require that the convergences* (18) *hold.*

3 Preliminaries

We establish here a few results which will be used in the analysis of the Gradient Scheme (14).

3.1 Uniform-in-time compactness for space-time Gradient Discretisations

The first result is a consequence, in the framework of gradient discretisations, of the results in Section 6.

Theorem 3.1 (Uniform-in-time $L^2(\Omega)$-weak compactness result) *Let $T > 0$, and $(\mathcal{D}_m)_{m\in\mathbb{N}} = (X_{\mathcal{D}_m,0}, \Pi_{\mathcal{D}_m}, \nabla_{\mathcal{D}_m}, \mathcal{I}_{\mathcal{D}_m}, (t_m^{(n)})_{n=0,\ldots,N_m})_{m\in\mathbb{N}}$ be a sequence of space-time discretisation in the sense of Definition 2.1 which is consistent in the sense of Definition 2.6.*
For any $m \in \mathbb{N}$, let $v_m = (v_m^{(n)})_{n=0,\ldots,N_m} \subset X_{\mathcal{D}_m,0}$. If there exists $q > 1$ and $C > 0$ such that, for any $m \in \mathbb{N}$,

$$||\Pi_{\mathcal{D}_m} v_m||_{L^\infty(0,T;L^2(\Omega))} \le C \quad \text{and} \quad \int_0^T |\delta_m v_m(t)|^q_{\star,\mathcal{D}_m} \mathrm{d}t \le C, \tag{20}$$

then the sequence $(\Pi_{\mathcal{D}_m} v_m)_{m\in\mathbb{N}}$ is relatively compact uniformly-in-time and weakly in $L^2(\Omega)$, i.e. it has a subsequence which converges according to Definition 6.4.
Moreover, any limit v of such a subsequence is continuous $[0,T] \to L^2(\Omega)$ for the weak topology.

Remark 3.2 *The bound on $|\delta_m v_m|_{\star,\mathcal{D}_m}$ is often a consequence on a numerical scheme satisfied by v_m and on bounds on $||\nabla_{\mathcal{D}_m} v_m||_{L^p(\Omega\times(0,T))}$, see the proof of Lemma 4.3 for example.*

Proof. This result is a consequence of the generalised Ascoli-Arzela theorem (Theorem 6.2) with $K = [0,T]$ and E the ball of radius C in $L^2(\Omega)$, endowed with the weak topology. We let $(\varphi_l)_{l\in\mathbb{N}} \subset C_c^\infty(\Omega)$ be a dense sequence in $L^2(\Omega)$ and endow E with the metric (78) from these φ_l, which indeed defines the weak $L^2(\Omega)$ topology (see Proposition 6.5).
The set E is metric compact and therefore complete, and all $\Pi_{\mathcal{D}_m} v_m$ take their values in E. It remains to estimate $d_E(v_m(s), v_m(s'))$. We drop the index m in the spaces for legibility of notations.
Let us define the interpolant $P_{\mathcal{D}}\varphi_l \in X_{\mathcal{D},0}$ by

$$P_{\mathcal{D}}\varphi_l = \underset{w\in X_{\mathcal{D},0}}{\operatorname{argmin}} \left(||\Pi_{\mathcal{D}} w - \varphi_l||_{L^{\max(p,2)}(\Omega)} + ||\nabla_{\mathcal{D}} w - \nabla\varphi_l||_{L^p(\Omega)} \right). \tag{21}$$

For any $0 \le s \le s' \le T$, writing $\Pi_{\mathcal{D}} v_m(s') - \Pi_{\mathcal{D}} v_m(s)$ as the sum of its jumps $\delta t^{(n+\frac{1}{2})} \Pi_{\mathcal{D}} \delta_{\mathcal{D}}^{(n+\frac{1}{2})} v_m$ at the points $(t^{(n)})_{n=n_1,\ldots,n_2}$ lying between s and s', the definition of $|\cdot|_{\star,\mathcal{D}}$ and Estimate (20) give

$$\begin{aligned}
&\left| \int_\Omega \left(\Pi_{\mathcal{D}} v_m(\boldsymbol{x}, s') - \Pi_{\mathcal{D}} v_m(\boldsymbol{x}, s) \right) \Pi_{\mathcal{D}} P_{\mathcal{D}}\varphi_l(\boldsymbol{x}) \mathrm{d}\boldsymbol{x} \right| \\
&\quad = \left| \int_{t^{(n_1)}}^{t^{(n_2+1)}} \int_\Omega \Pi_{\mathcal{D}} \delta_{\mathcal{D}} v(t)(\boldsymbol{x}) \Pi_{\mathcal{D}} P_{\mathcal{D}}\varphi_l(\boldsymbol{x}) \mathrm{d}\boldsymbol{x} \mathrm{d}t \right| \le C^{1/q} (t^{(n_2+1)} - t^{(n_1)})^{1/q'} ||\nabla_{\mathcal{D}} P_{\mathcal{D}}\varphi_l||_{L^p(\Omega)^d}.
\end{aligned}$$

By definition of $P_{\mathcal{D}}$, we have

$$||\Pi_{\mathcal{D}} P_{\mathcal{D}}\varphi_l - \varphi_l||_{L^2(\Omega)} \le \widehat{S}_{\mathcal{D}}(\varphi_l)$$

and

$$||\nabla_{\mathcal{D}} P_{\mathcal{D}}\varphi_l||_{L^p(\Omega)^d} \le \widehat{S}_{\mathcal{D}}(\varphi_l) + ||\varphi_l||_{L^p(\Omega)} + ||\nabla\varphi_l||_{L^p(\Omega)^d} \le C_{\varphi_l}$$

with C_{φ_l} not depending on $\mathcal{D}$ (and therefore on m). Since $t^{(n_2+1)} - t^{(n_1)} \le |s'-s| + \delta t$ and $(\Pi_{\mathcal{D}} v_m)_{m\in\mathbb{N}}$ is bounded in $L^\infty(0,T;L^2(\Omega))$, we deduce that

$$\left|\int_\Omega \left(\Pi_{\mathcal{D}} v_m(\boldsymbol{x},s') - \Pi_{\mathcal{D}} v_m(\boldsymbol{x},s)\right)\varphi_l(\boldsymbol{x})\mathrm{d}\boldsymbol{x}\right| \\ \le 2C\widehat{S}_{\mathcal{D}}(\varphi_l) + C^{1/q}C_{\varphi_l}|s'-s|^{1/q'} + C^{1/q}C_{\varphi_l}\delta t^{1/q'}. \quad (22)$$

Plugged into the definition (78) of the distance in E, this shows that

$$\begin{aligned} & d_E\Big(\Pi_{\mathcal{D}} v_m(s'), \Pi_{\mathcal{D}} v_m(s)\Big) \\ \le\; & \sum_{l\in\mathbb{N}} \frac{\min(1, C^{1/q'}C_{\varphi_l}|s'-s|^{1/q'})}{2^l} + \sum_{l\in\mathbb{N}} \frac{\min(1, 2C\widehat{S}_{\mathcal{D}_m}(\varphi_l) + C^{1/q'}C_{\varphi_l}\delta t_m^{1/q'})}{2^l} \\ =:\; & \omega(s,s') + \delta_m. \end{aligned}$$

Using the dominated convergence theorem for series, we see that $\omega(s,s') \to 0$ as $s - s' \to 0$ and that $\delta_m \to 0$ as $m \to \infty$ (we invoke the space consistency to establish that $\lim_{m\to\infty} \widehat{S}_{\mathcal{D}_m}(\varphi_l) \to 0$ for any l). Hence, the assumptions of Theorem 6.2 are satisfied and the proof is complete. ■

3.2 Technical results

We state here a family of technical lemmas, starting with a few useful properties on ν and B.

Lemma 3.3 *Under Assumptions* (2)*, there holds*

$$|\nu(a) - \nu(b)| \le L_\beta |\zeta(a) - \zeta(b)|, \quad (23)$$

$$(\nu(a) - \nu(b))^2 \le L_\beta L_\zeta (\zeta(a) - \zeta(b))(\beta(a) - \beta(b)). \quad (24)$$

The function B is convex on $\overline{R_\beta}$, the function $B\circ\beta : \mathbb{R} \to [0,\infty)$ is continuous,

$$\forall s \in \mathbb{R}, \quad B(\beta(s)) = \int_0^s \zeta(q)\beta'(q)\mathrm{d}q, \quad (25)$$

$$\exists K_0, K_1, K_2 > 0 \text{ such that } \forall s \in \mathbb{R}, \quad K_0\beta(s)^2 - K_1 \le B(\beta(s)) \le K_2 s^2, \quad (26)$$

$$\forall a \in \mathbb{R}, \forall r \in \overline{R_\beta}, \quad B(r) - B(\beta(a)) \ge \zeta(a)(r - \beta(a)), \quad (27)$$

and

$$\forall s,s' \in \mathbb{R},\ (\nu(s) - \nu(s'))^2 \le 4L_\beta L_\zeta \left[B(\beta(s)) + B(\beta(s')) - 2B\left(\frac{\beta(s)+\beta(s')}{2}\right)\right]. \quad (28)$$

Proof.
Inequality (23) is a straightforward consequence of the estimate $\nu' = \zeta'\beta' \le L_\beta\zeta'$. Note that the same inequality also holds with β and ζ swapped. Inequality (24) is then a direct application of (23) and the similar inequality with β and ζ swapped.
Let us first notice that, since $\beta \ge 0$ on $\mathbb{R}^+$ and $\beta \le 0$ on $\mathbb{R}^-$, $\beta_r(s)$ is always a real number when $s \in R_\beta$. Moreover, since β is non-decreasing, β_r is also non-decreasing on R_β and therefore locally bounded on R_β. Hence, B is well defined and locally Lipschitz-continuous, with an a.e. derivative $B' = \zeta(\beta_r)$. B' is therefore non-decreasing and B is convex.

To prove (25), we denote by $P \subset R_\beta$ the countable set of plateaux values of β, i.e. the $y \in \mathbb{R}$ such that $\beta^{-1}(\{y\})$ is not reduced to a singleton. If $s \notin \beta^{-1}(P)$ then $\beta^{-1}(\{\beta(s)\})$ is the singleton $\{s\}$ and therefore $\beta_r(\beta(s)) = s$. Moreover, β_r is continuous at $\beta(s)$ and thus B is differentiable at $\beta(s)$. Since β is differentiable a.e., we therefore deduce that, for a.e. $s \notin \beta^{-1}(P)$, $(B(\beta))'(s) = B'(\beta(s))\beta'(s) = \zeta(\beta_r(\beta(s)))\beta'(s) = \zeta(s)\beta'(s)$. The set $\beta^{-1}(P)$ is a union of intervals on which β and thus $B(\beta)$ are locally constant; hence, for a.e. s in this set, $(B(\beta))'(s) = 0$ and $\zeta(s)\beta'(s) = 0$. Hence, the locally Lipschitz-continuous functions $B(\beta)$ and $s \to \int_0^s \zeta(q)\beta'(q)\mathrm{d}q$ have identical derivatives a.e. on $\mathbb{R}$ and take the same value at $s = 0$. They are thus equal on $\mathbb{R}$ and the proof of (25) is complete.

The continuity of $B \circ \beta$ is an obvious consequence of (25). The second inequality in (26) can also be easily deduced from (25) by noticing that $|\zeta(s)\beta'(s)| \le L_\zeta L_\beta |s|$ (we can take $K_2 = \frac{L_\beta L_\zeta}{2}$). For the first inequality in (26), we first infer from (2b) the existence of $S > 0$ such that $|\zeta(q)| \ge \frac{M_0}{2}|q| \ge \frac{M_0}{2L_\beta}|\beta(q)|$ whenever $|q| \ge S$. We then write, for $s \ge S$,

$$B(\beta(s)) = \int_0^S \zeta(q)\beta'(q)\mathrm{d}q + \int_S^s \zeta(q)\beta'(q)\mathrm{d}q \ge \frac{M_0}{2L_\beta}\int_S^s \beta(q)\beta'(q)\mathrm{d}q = \frac{M_0}{4L_\beta}\left(\beta(s)^2 - \beta(S)^2\right).$$

A similar inequality holds for $s \le -S$ (with $\beta(-S)$ instead of $\beta(S)$) and the first inequality in (26) therefore holds with $K_0 = \frac{M_0}{4L_\beta}$ and $K_1 = \frac{M_0}{4L_\beta}\max_{[-S,S]}\beta^2$.

We now prove (27), which states that a belongs to the convex sub-differential of B at $\beta(a)$. We first start with the case $r \in R_\beta$, that is $r = \beta(b)$ for some $b \in \mathbb{R}$. If β_r is continuous at $\beta(a)$ then this inequality is an obvious consequence of the convexity of B since B is then differentiable at $\beta(a)$ with $B'(\beta(a)) = \zeta(\beta_r(\beta(a))) = \zeta(a)$. Otherwise, a plain reasoning also does the job as

$$B(r) - B(\beta(a)) = B(\beta(b) - B(\beta(a))$$
$$\int_a^b \zeta(q)\beta'(q)\mathrm{d}q = \int_a^b (\zeta(q) - \zeta(a))\beta'(q)\mathrm{d}q + \zeta(a)(\beta(b) - \beta(a)) \ge \zeta(a)(r - \beta(a)),$$

the inequality coming from the fact that $\beta' \ge 0$ and $\zeta(q) - \zeta(a)$ has the same sign as $b - a$ when q is between a and b. The general case $r \in \overline{R_\beta}$ is obtained by passing to the limit on b_n such that $\beta(b_n) \to r$ and using the fact that B has limits (possibly $+\infty$) at the endpoints of R_β.

Let us now take $s, s' \in \mathbb{R}$. Let $\bar{s} \in \mathbb{R}$ be such that $\beta(\bar{s}) = \frac{\beta(s)+\beta(s')}{2}$. We notice that

$$B(\beta(s)) + B(\beta(s')) - 2B(\beta(\bar{s})) = \int_{\bar{s}}^s (\zeta(q) - \zeta(\bar{s}))\beta'(q)\mathrm{d}q + \int_{\bar{s}}^{s'} (\zeta(q) - \zeta(\bar{s}))\beta'(q)\mathrm{d}q.$$

We then use $|\zeta(q) - \zeta(\bar{s})| \ge \frac{1}{L_\beta}|\nu(q) - \nu(\bar{s})|$ and $\beta'(q) \ge \beta'(q)\frac{\zeta'(q)}{L_\zeta} = \frac{\nu'(q)}{L_\zeta}$ to write

$$\int_{\bar{s}}^s (\zeta(q) - \zeta(\bar{s}))\beta'(q)\mathrm{d}q \ge \frac{1}{L_\beta L_\zeta}\int_{\bar{s}}^s \nu'(q)(\nu(q) - \nu(\bar{s}))\mathrm{d}q = \frac{1}{2L_\beta L_\zeta}(\nu(s) - \nu(\bar{s}))^2.$$

Thanks to

$$(\nu(s) - \nu(s'))^2 \le 2((\nu(s) - \nu(\bar{s}))^2 + (\nu(s') - \nu(\bar{s}))^2),$$

we deduce that (28) follows. ■

The next lemma is an easy consequence of Fatou's lemma and the fact that strongly lower semi-continuous convex functions are also weakly lower semi-continuous. We all the same provide its short proof.

Lemma 3.4 *Let I be an interval of $\mathbb{R}$ and $H : I \to \mathbb{R}$ be a convex function. We denote by $L^2(\Omega; I)$ the convex set of functions in $L^2(\Omega)$ with values in I. Let $v \in L^2(\Omega; I)$ and $(v_m)_{m\in\mathbb{N}}$ a sequence of functions in $L^2(\Omega; I)$ which converges weakly to v in $L^2(\Omega)$. Then*

$$\int_\Omega H(v(\boldsymbol{x}))\mathrm{d}\boldsymbol{x} \le \liminf_{m\to\infty} \int_\Omega H(v_m(\boldsymbol{x}))\mathrm{d}\boldsymbol{x}.$$

Proof.
Let $\Phi : L^2(\Omega; I) \to (-\infty, \infty]$ be defined by $\Phi(w) = \int_\Omega H(w(\boldsymbol{x}))\mathrm{d}\boldsymbol{x}$. Since H is convex, it is greater than a linear functional and $\Phi(w)$ is thus well defined in $(\infty, \infty]$. Moreover, if $w_k \to w$ strongly in $L^2(\Omega; I)$ then, up to a subsequence, $w_k \to w$ a.e. on Ω and therefore $H(w_k) \to H(w)$ a.e. on Ω. Combined with the linear lower bound of H, we can apply Fatou's lemma to see that $\Phi(w) \le \liminf_{k\to\infty} \Phi(w_k)$.
Hence, Φ is lower semi-continuous for the strong topology of $L^2(\Omega; I)$. Since Φ (as H) is convex, we deduce that this lower semi-continuity property is also valid for the weak topology of $L^2(\Omega; I)$, see [19]. The result of the lemma is just the translation of this weak lower semi-continuity of Φ. ∎

The last technical result is a consequence of the Minty trick. It has been proved and used in the L^2 case in [21, 14], but we need here an extension to the non-Hilbertian case.

Lemma 3.5 (Minty's trick) *Let $H \in C^0(\mathbb{R})$ be a nondecreasing function. Let (X, μ) be a measurable set with finite measure and let $(u_n)_{n\in\mathbb{N}} \subset L^p(X)$ with $p > 1$ such that*

1. *there exists $u \in L^p(X)$ such that $(u_n)_{n\in\mathbb{N}}$ weakly converges to u in $L^p(X)$;*
2. *$(H(u_n))_{n\in\mathbb{N}} \subset L^1(X)$ and there exists $w \in L^1(X)$ such that $(H(u_n))_{n\in\mathbb{N}}$ strongly converges to w in $L^1(X)$;*

Then $w = H(u)$ a.e. on X.

Proof.
Let $k > 0$ and $T_k(s) = \max(-k, \min(s, k))$ be the usual truncation at level k. Since H is non-decreasing, there exists $h_k \to \infty$ as $k \to \infty$ such that $H(T_k(s)) = T_{h_k}(H(s))$. Thus, $H(T_k(u_n)) \to T_{h_k}(w)$ in $L^1(X)$ as $n \to \infty$. Given that $(H(T_k(u_n)))_{n\in\mathbb{N}}$ remains bounded in $L^\infty(X)$, its convergence to $T_{h_k}(w)$ also holds in $L^{p'}(X)$.
Using fact that $H \circ T_k$ is non-decreasing, we write, for any $g \in L^p(X)$,

$$\int_X (H(T_k(u_n)) - H(T_k(g)))(u_n - g)\mathrm{d}\mu \ge 0.$$

By strong convergence of $H(T_k(u_n))$ in $L^{p'}(X)$ and weak convergence of u_n in $L^p(X)$, as well as the fact that $H \circ T_k$ is bounded, we can take the limit of this expression as $n \to \infty$ and we find

$$\int_X (T_{h_k}(w) - H(T_k(g)))(u - g)\mathrm{d}\mu \ge 0. \tag{29}$$

We then use Minty's trick, that is pick a generic $\varphi \in L^p(X)$, apply (29) to $g = u - t\varphi$, divide by t and let $t \to \pm 0$ (using the dominated convergence theorem and the fact that $H \circ T_k$ is continuous and bounded) to find

$$\int_X (T_{h_k}(w) - H(T_k(u)))\varphi\mathrm{d}\mu = 0.$$

Selecting $\varphi = \mathrm{sign}(T_{h_k}(w) - H(T_k(u)))$, we deduce that $T_{h_k}(w) = H(T_k(u))$ a.e. on X. Letting $k \to \infty$, we conclude that $w = H(u)$ a.e. on X. ∎

3.3 Integration-by-parts for the continuous solution

The last series of preliminary results are properties on the solution to (4), all based on the following integration-by-part property. This property, used in the proof of Theorem 2.10 and 2.14, allows to compute the value of the linear form $\partial_t\beta(v) \in L^{p'}(0,T;W^{-1,p'}(\Omega))$ on the function $\zeta(\overline{u})$. Because of the lack of regularity on $\overline{u}$ and the many non-linearities in (1), justifying this integration-by-parts is however not straightforward at all...

Lemma 3.6 *Let us assume* (2b) *and* (2c). *Let $v : \Omega \times (0,T) \mapsto \mathbb{R}$ be measurable such that $\zeta(v) \in L^p(0,T;W_0^{1,p}(\Omega))$, $B(\beta(v)) \in L^\infty(0,T;L^1(\Omega))$, $\beta(v) \in C([0,T];L^2(\Omega)\text{-}w)$ and $\partial_t\beta(v) \in L^{p'}(0,T;W^{-1,p'}(\Omega))$. Then $t \in [0,T] \to \int_\Omega B(\beta(v)(\boldsymbol{x},t))\mathrm{d}\boldsymbol{x} \in [0,\infty)$ is continuous and, for all $t_1,t_2 \in [0,T]$,*

$$\int_{t_1}^{t_2} \langle \partial_t\beta(v)(t), \zeta(v(t))\rangle_{W^{-1,p'},W_0^{1,p}}\mathrm{d}t = \int_\Omega B(\beta(v)(\boldsymbol{x},t_2))\mathrm{d}\boldsymbol{x} - \int_\Omega B(\beta(v)(\boldsymbol{x},t_1))\mathrm{d}\boldsymbol{x}. \tag{30}$$

Remark 3.7 *Like at the end of Section 1.2, it is important to in mind the separation of $\beta(v(\cdot,\cdot))$ from its continuous representative $\beta(v)(\cdot,\cdot)$.*

Proof.
Note that we obviously only need to make the proof when $0 \le t_1 < t_2 \le T$.
Step 1: truncation, extensions and approximation of $\beta(v)$.
We define $\overline{\beta(v)} : \mathbb{R} \to L^2(\Omega)$ by setting

$$\overline{\beta(v)}(t) = \begin{cases} \beta(v)(t) & \text{if } t \in [t_1,t_2], \\ \beta(v)(t_1) & \text{if } t \le t_1, \\ \beta(v)(t_2) & \text{if } t \ge t_2. \end{cases}$$

By the continuity property of $\beta(v)$, this definition and gives $\overline{\beta(v)} \in C(\mathbb{R};L^2(\Omega)\text{-w})$ such that $\partial_t\overline{\beta(v)} = \mathbf{1}_{(t_1,t_2)}\partial_t\beta(v) \in L^{p'}(\mathbb{R};W^{-1,p'}(\Omega))$ (no Dirac masses have been introduced at $t = t_1$ or $t = t_2$). This regularity of $\partial_t\overline{\beta(v)}$ ensures that the function

$$t \in \mathbb{R} \mapsto D_h\overline{\beta(v)} := \frac{1}{h}\int_t^{t+h} \partial_t\overline{\beta(v)}(s)\mathrm{d}s = \frac{\overline{\beta(v)}(t+h) - \overline{\beta(v)}(t)}{h} \text{ in } W^{-1,p'}(\Omega) \tag{31}$$

tend to $\partial_t\overline{\beta(v)}$ in $L^{p'}(\mathbb{R};W^{-1,p'}(\Omega))$ as $h \to 0$.

Step 2: we prove that $||B(\overline{\beta(v)}(t)||_{L^1(\Omega)} \le ||B(\beta(v))||_{L^\infty(0,T;L^1(\Omega))}$ for *all* $t \in \mathbb{R}$ (not only for a.e. t).
Let $t \in [t_1,t_2]$. Since $\beta(v)(\cdot,\cdot) = \beta(v(\cdot,\cdot))$ a.e. on $\Omega \times (t_1,t_2)$, there exists a sequence $t_n \to t$ such that $\beta(v)(\cdot,t_n) = \beta(v(\cdot,t_n))$ in L^2 for all n and $||B(\beta(v)(\cdot,t_n))||_{L^1(\Omega)} \le ||B(\beta(v))||_{L^\infty(0,T;L^1(\Omega))}$. As $\beta(v) \in C([0,T];L^2(\Omega)\text{-w})$, we have $\beta(v)(\cdot,t_n) \to \beta(v)(\cdot,t)$ weakly in $L^2(\Omega)$. We then use the convexity of B and Lemma 3.4 to write, thanks to our choice of t_n,

$$\int_\Omega B(\beta(v)(\boldsymbol{x},t))\mathrm{d}\boldsymbol{x} \le \liminf_{n\to\infty}\int_\Omega B(\beta(v)(\boldsymbol{x},t_n))\mathrm{d}\boldsymbol{x} \le ||B(\beta(v))||_{L^\infty(0,T;L^1(\Omega))}$$

and the proof is complete for $t \in [t_1,t_2]$. The result for $t \le t_1$ or $t \ge t_2$ is obvious since $\overline{\beta(v)}(t)$ is then either $\beta(v)(t_1)$ or $\beta(v)(t_2)$.

Step 3: We prove that for all $\tau \in \mathbb{R}$ and a.e. $t \in (t_1,t_2)$,

$$\langle\overline{\beta(v)}(\tau) - \beta(v)(t), \zeta(v(\cdot,t))\rangle_{W^{-1,p'},W_0^{1,p}} \le \int_\Omega B(\overline{\beta(v)}(\boldsymbol{x},\tau)) - B(\beta(v)(\boldsymbol{x},t))\mathrm{d}\boldsymbol{x}. \tag{32}$$

Note that if we could just replace the duality product $W^{-1,p'}$–$W_0^{1,p}$ with an L^2 inner product, then this formula would be a straightforward consequence of (27). The problem is that nothing ensures that $\zeta(v)(t) \in L^2(\Omega)$ for a.e. t.
We first notice that $\overline{\beta(v)}(\tau) - \beta(v)(t) = \int_t^\tau \partial_t \overline{\beta(v)}(s)\mathrm{d}s$ indeed belongs to $W^{-1,p'}(\Omega)$ so the right-hand side of (32) makes sense provided that t is chosen such that $\zeta(v(\cdot,t)) \in W_0^{1,p}(\Omega)$ (which we do from here on). To deal with the fact that $\zeta(v(\cdot,t))$ does not necessarily belong to $L^2(\Omega)$, we replace it with a truncation. Denoting by $T_k(s) = \max(-k, \min(s,k))$ the classical truncation at level k, by the growth assumption (2b) on ζ we see that there exists $r_k \to +\infty$ as $k \to +\infty$ such that $\zeta(T_k(v(\cdot,t))) = T_{r_k}(\zeta(v(\cdot,t)))$. Hence, $\zeta(T_k(v(\cdot,t)) \in W_0^{1,p}(\Omega)$ and converges, as $k \to \infty$, to $\zeta(v(\cdot,t))$ in $W_0^{1,p}(\Omega)$.
We can therefore write

$$\begin{aligned}\langle \overline{\beta(v)}(\tau) - \beta(v)(t), \zeta(v)(t)\rangle_{W^{-1,p'},W_0^{1,p}} &= \lim_{k\to\infty} \langle \overline{\beta(v)}(\tau) - \beta(v)(t), \zeta(T_k(v(\cdot,t))\rangle_{W^{-1,p'},W_0^{1,p}} \\ &= \lim_{k\to\infty} \int_\Omega \left(\overline{\beta(v)}(\boldsymbol{x},\tau) - \beta(v(\boldsymbol{x},t))\right) \zeta(T_k(v(\boldsymbol{x},t))\mathrm{d}\boldsymbol{x}, \quad (33)\end{aligned}$$

the replacement of the duality product by an $L^2(\Omega)$ inner product being justified since $\overline{\beta(v)}(\tau) - \beta(v)(t)$ and $\zeta(v(\cdot,t))$ both belong to $L^2(\Omega)$. We also used the fact that, for a.e. $t \in (t_1,t_2)$, $\beta(v)(\cdot,t) = \beta(v(\cdot,t))$ a.e. on Ω, so (33) is valid for a.e. $t \in (t_1,t_2)$.
We then write $\beta(v(\boldsymbol{x},t)) = \beta(T_k(v(\boldsymbol{x},t))) + (\beta(v(\boldsymbol{x},t)) - \beta(T_k(v(\boldsymbol{x},t))))$ and apply (27) with $r = \overline{\beta(v)}(\boldsymbol{x},\tau)$ and $a = T_k(\beta(v(\boldsymbol{x},t)))$ to find

$$\begin{aligned}&\int_\Omega \left(\overline{\beta(v)}(\boldsymbol{x},\tau) - \beta(v(\boldsymbol{x},t))\right) \zeta(T_k(v(\boldsymbol{x},t))\mathrm{d}\boldsymbol{x} \\ &\qquad = \int_\Omega \left(\overline{\beta(v)}(\boldsymbol{x},\tau) - \beta(T_k(v(\boldsymbol{x},t)))\right) \zeta(T_k(v(\boldsymbol{x},t))\mathrm{d}\boldsymbol{x} \\ &\qquad\quad - \int_\Omega (\beta(v(\boldsymbol{x},t)) - \beta(T_k(v(\boldsymbol{x},t))))\, \zeta(T_k(v(\boldsymbol{x},t))\mathrm{d}\boldsymbol{x} \\ &\le \int_\Omega B(\overline{\beta(v)}(\boldsymbol{x},\tau)) - B(\beta(T_k(v(\boldsymbol{x},t))))\mathrm{d}\boldsymbol{x} - \int_\Omega (\beta(v(\boldsymbol{x},t)) - \beta(T_k(v(\boldsymbol{x},t))))\, \zeta(T_k(v(\boldsymbol{x},t))\mathrm{d}\boldsymbol{x}.\end{aligned}$$

Studying the cases $v(\boldsymbol{x},t) \ge k$ or $v(\boldsymbol{x},t) \le -k$, we notice that the last integrand is always non-negative, so we actually can write

$$\int_\Omega \left(\overline{\beta(v)}(\boldsymbol{x},\tau) - \beta(v(\boldsymbol{x},t))\right) \zeta(T_k(v(\boldsymbol{x},t))\mathrm{d}\boldsymbol{x} \le \int_\Omega B(\overline{\beta(v)}(\boldsymbol{x},\tau)) - B(\beta(T_k(v(\boldsymbol{x},t))))\mathrm{d}\boldsymbol{x}.$$

We then use the continuity of $B \circ \beta$ and Fatou's lemma to deduce

$$\begin{aligned}&\limsup_{k\to\infty} \int_\Omega \left(\overline{\beta(v)}(\boldsymbol{x},\tau) - \beta(v(\boldsymbol{x},t))\right) \zeta(T_k(v(\boldsymbol{x},t))\mathrm{d}\boldsymbol{x} \\ &\qquad \le \int_\Omega B(\overline{\beta(v)}(\boldsymbol{x},\tau))\mathrm{d}\boldsymbol{x} - \liminf_{k\to\infty} \int_\Omega B(\beta(T_k(v(\boldsymbol{x},t))))\mathrm{d}\boldsymbol{x} \\ &\qquad\qquad \le \int_\Omega B(\overline{\beta(v)}(\boldsymbol{x},\tau))\mathrm{d}\boldsymbol{x} - \int_\Omega B(\beta(v(\boldsymbol{x},t)))\mathrm{d}\boldsymbol{x}\end{aligned}$$

which, combined with (33), concludes the proof of (32) (recall that t has been chosen such that $\beta(v(\cdot,t)) = \beta(v)(\cdot,t)$ a.e. on Ω).

Step 4: proof of the formula

By convergence of $D_h\overline{\beta(v)}$ to $\partial_t\overline{\beta(v)}$ in $L^{p'}(0,T;W^{-1,p'}(\Omega))$ and since $\mathbf{1}_{(t_1,t_2)}\zeta(v) \in L^p(\mathbb{R};W_0^{1,p}(\Omega))$, we have

$$\begin{aligned}\int_{t_1}^{t_2} \langle \partial_t\beta(v)(t), \zeta(v(t))\rangle_{W^{-1,p'},W_0^{1,p}}\mathrm{d}t &= \int_{\mathbb{R}} \langle \partial_t\overline{\beta(v)}(t), \mathbf{1}_{(t_1,t_2)}(t)\zeta(v(\cdot,t))\rangle_{W^{-1,p'},W_0^{1,p}}\mathrm{d}t \\ &= \lim_{h\to 0}\int_{\mathbb{R}} \langle D_h\overline{\beta(v)}(t), \mathbf{1}_{(t_1,t_2)}(t)\zeta(v(\cdot,t))\rangle_{W^{-1,p'},W_0^{1,p}}\mathrm{d}t \\ &= \lim_{h\to 0}\frac{1}{h}\int_{t_1}^{t_2} \langle\overline{\beta(v)}(s+h) - \overline{\beta(v)}(t), \zeta(v(\cdot,t)\rangle_{W^{-1,p'},W_0^{1,p}}\mathrm{d}t. \quad (34)\end{aligned}$$

We then use (32) for a.e. $t \in (t_1,t_2)$ to obtain, for h small enough such that $t_1 + h < t_2$,

$$\begin{aligned}&\frac{1}{h}\int_{t_1}^{t_2} \langle\overline{\beta(v)}(t+h) - \overline{\beta(v)}(t), \zeta(v(\cdot,t))\rangle_{W^{-1,p'},W_0^{1,p}}\mathrm{d}t \\ &\quad\leq \frac{1}{h}\int_{t_1}^{t_2}\int_\Omega B(\overline{\beta(v)}(\boldsymbol{x},t+h)) - B(\overline{\beta(v)}(\boldsymbol{x},t))\mathrm{d}\boldsymbol{x}\mathrm{d}t \\ &\quad= \frac{1}{h}\int_{t_2}^{t_2+h}\int_\Omega B(\overline{\beta(v)}(\boldsymbol{x},t))\mathrm{d}\boldsymbol{x}\mathrm{d}t - \frac{1}{h}\int_{t_1}^{t_1+h}\int_\Omega B(\overline{\beta(v)}(\boldsymbol{x},t))\mathrm{d}\boldsymbol{x}\mathrm{d}t \\ &\quad= \int_\Omega B(\beta(v)(\boldsymbol{x},t_2))\mathrm{d}\boldsymbol{x} - \frac{1}{h}\int_{t_1}^{t_1+h}\int_\Omega B(\beta(v)(\boldsymbol{x},t))\mathrm{d}\boldsymbol{x}\mathrm{d}t\end{aligned}$$

We now take the lim sup of this inequality, using the fact that $B(\beta(v)(\cdot,t_2))$ is integrable (Step 2) to take its integral out of the lim sup. Coming back to (34) we obtain

$$\begin{aligned}&\int_{t_1}^{t_2} \langle \partial_t\beta(v)(t), \zeta(v(t))\rangle_{W^{-1,p'},W_0^{1,p}}\mathrm{d}t \\ &\qquad\leq \int_\Omega B(\beta(v)(\boldsymbol{x},t_2))\mathrm{d}\boldsymbol{x} - \liminf_{h\to 0}\frac{1}{h}\int_{t_1}^{t_1+h}\int_\Omega B(\beta(v)(\boldsymbol{x},t))\mathrm{d}\boldsymbol{x}\mathrm{d}t. \quad (35)\end{aligned}$$

But since $\beta(v) \in C([0,T];L^2(\Omega)\text{-w})$, as $h \to 0$ we have $\frac{1}{h}\int_{t_1}^{t_1+h}\beta(v)(t)\mathrm{d}t \to \beta(v)(t_1)$ weakly in $L^2(\Omega)$. Hence, the convexity of B, Lemma 3.4 and Jensen's inequality give

$$\begin{aligned}\int_\Omega B(\beta(v)(\boldsymbol{x},t_1))\mathrm{d}\boldsymbol{x} &\leq \liminf_{h\to 0}\int_\Omega B\left(\frac{1}{h}\int_{t_1}^{t_1+h}\beta(v)(\boldsymbol{x},t)\mathrm{d}t\right)\mathrm{d}\boldsymbol{x} \\ &\leq \liminf_{h\to 0}\int_\Omega\frac{1}{h}\int_{t_1}^{t_1+h} B(\beta(v)(\boldsymbol{x},t))\mathrm{d}t\mathrm{d}\boldsymbol{x}.\end{aligned}$$

Plugged into (35), this inequality shows that (30) holds with $\leq$ instead of $=$. The reverse inequality is obtained by reversing the time. We consider $\widetilde{v}(t) = v(t_1+t_2-t)$. Then $\zeta(\widetilde{v})$, $B(\beta(\widetilde{v}))$ and $\beta(\widetilde{v})$ have the same properties as $\zeta(v)$, $B(\beta(v))$ and $\beta(v)$, and $\beta(\widetilde{v})$ takes values $\beta(v)(t_1)$ at $t = t_2$ and $\beta(v)(t_2)$ at $t = t_1$. Applying (30) with "$\leq$" instead of "$=$" to $\widetilde{v}$ and using the fact that $\partial_t\beta(\widetilde{v})(t) = -\partial_t\beta(v)(t_1+t_2-t)$, we obtain (30) with "$\geq$" instead of "$=$" and the proof of (30) is complete.
The continuity of $t \in [0,T] \mapsto \int_\Omega B(\beta(v)(\boldsymbol{x},t))\mathrm{d}\boldsymbol{x}$ is straightforward from (30) as the left-hand side of this relation is continuous with respect to t_1 and t_2. ■

The following corollary states continuity properties and an essential formula on the solution to (4).

Corollary 3.8 *Under Assumptions* (2a)–(2i)*, if* $\overline{u}$ *is a solution of* (4) *then:*

1. *the function* $t \in [0,T] \mapsto \int_\Omega B(\beta(\overline{u})(\boldsymbol{x},t))\mathrm{d}\boldsymbol{x} \in [0,\infty)$ *is continuous and bounded,*
2. *for any* $T_0 \in [0,T]$*,*

$$\int_\Omega B(\beta(\overline{u})(\boldsymbol{x},T_0))\mathrm{d}\boldsymbol{x} + \int_0^{T_0}\int_\Omega \boldsymbol{a}(\boldsymbol{x},\nu(\overline{u}(\boldsymbol{x},t)),\nabla\zeta(\overline{u})(\boldsymbol{x},t))\cdot\nabla\zeta(\overline{u})(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x}\mathrm{d}t$$
$$= \int_\Omega B(\beta(u_{\mathrm{ini}}(\boldsymbol{x})))\mathrm{d}\boldsymbol{x} + \int_0^{T_0}\int_\Omega f(\boldsymbol{x},t)\zeta(\overline{u})(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x}\mathrm{d}t, \quad (36)$$

3. $\nu(\overline{u})$ *is continuous* $[0,T] \to L^2(\Omega)$*.*

Remark 3.9 *The continuity of* $\nu(\overline{u})$ *has to be understood in the same sense as the continuity of* $\beta(\overline{u})$*, that is* $\nu(\overline{u})$ *is a.e. on* $\Omega\times(0,T)$ *equal to a continuous function* $[0,T]\to L^2(\Omega)$*. We use in particular the same notation* $\nu(\overline{u}(\cdot,\cdot))$ *for the continuous representative* $\nu(\overline{u})(\cdot,\cdot)$ *as we did for the continuous representative of* $\beta(\overline{u})$*.*

Proof.
The continuity of $t \in [0,T] \mapsto \int_\Omega B(\beta(\overline{u})(\boldsymbol{x},t))\mathrm{d}\boldsymbol{x} \in [0,\infty)$ and Formula (36) are straightforward consequences of Lemma 3.6 with $v = \overline{u}$ and using (4) with $\boldsymbol{v} = \zeta(\overline{u})$. Note that the bound on $\int_\Omega B(\beta(\overline{u})(\boldsymbol{x},t))\mathrm{d}\boldsymbol{x}$ can be seen as a consequence of (36), or from Step 2 in the proof of Lemma 3.6.
Let us prove the strong continuity of $\nu(\overline{u}) : [0,T] \mapsto L^2(\Omega)$. Let $\mathcal{T}$ be the set of $\tau\in[0,T]$ such that $\beta(\overline{u}(\cdot,\tau)) = \beta(\overline{u})(\cdot,\tau)$ a.e. on Ω, and let $(s_l)_{l\in\mathbb{N}}$ and $(t_k)_{k\in\mathbb{N}}$ be two sequences in $\mathcal{T}$ which converge to the same value s. Invoking (28) we can write

$$\int_\Omega(\nu(\overline{u}(\boldsymbol{x},s_l)) - \nu(\overline{u}(\boldsymbol{x},t_k)))^2\mathrm{d}\boldsymbol{x} \le 4L_\beta L_\zeta\left(\int_\Omega B(\beta(\overline{u})(\boldsymbol{x},s_l))\mathrm{d}\boldsymbol{x} + \int_\Omega B(\beta(\overline{u})(\boldsymbol{x},t_k))\mathrm{d}\boldsymbol{x}\right)$$
$$- 8L_\beta L_\zeta\int_\Omega B\left(\frac{\beta(\overline{u})(\boldsymbol{x},s_l)+\beta(\overline{u})(\boldsymbol{x},t_k)}{2}\right)\mathrm{d}\boldsymbol{x}. \quad (37)$$

Since $\frac{\beta(\overline{u})(\cdot,s_l)+\beta(\overline{u})(\cdot,t_k)}{2} \to \beta(\overline{u})(\cdot,s)$ weakly in $L^2(\Omega)$ as $l,k\to\infty$, Lemma 3.4 gives

$$\int_\Omega B\left(\beta(\overline{u})(\boldsymbol{x},s)\right)\mathrm{d}\boldsymbol{x} \le \liminf_{l,k\to\infty}\int_\Omega B\left(\frac{\beta(\overline{u})(\boldsymbol{x},s_l)+\beta(\overline{u})(\boldsymbol{x},t_k)}{2}\right)\mathrm{d}\boldsymbol{x}.$$

Taking the lim sup as $l,k\to\infty$ of (37) and using the continuity of $t\mapsto\int_\Omega B(\beta(\overline{u})(\boldsymbol{x},t))\mathrm{d}\boldsymbol{x}$ thus shows that

$$||\nu(\overline{u}(\cdot,s_l)) - \nu(\overline{u}(\cdot,t_k))||_{L^2(\Omega)} \to 0 \quad \text{as } l,k\to\infty. \quad (38)$$

The existence of an a.e. representative of $\nu(\overline{u}(\cdot,\cdot))$ which is continuous $[0,T]\mapsto L^2(\Omega)$ is a direct consequence of this convergence. Let $s\in[0,T]$ and $(s_l)_{l\in\mathbb{N}}\subset\mathcal{T}$ which converges to s. Applied with $t_k = s_k$, (38) shows that $(\nu(\overline{u}(\cdot,s_l)))_{l\in\mathbb{N}}$ is a Cauchy sequence in $L^2(\Omega)$ and therefore that $\lim_{l\to\infty}\nu(\overline{u}(\cdot,s_l))$ exists in $L^2(\Omega)$. Relation (38) moreover also shows that this limit, that we can call $\nu(\overline{u})(\cdot,s)$, does not depend on the Cauchy sequence in $\mathcal{T}$ which converges to s. With $t_k = s$, we also see that whenever $s\in\mathcal{T}$ we have $\nu(\overline{u}(\cdot,s)) = \nu(\overline{u})(\cdot,s)$ a.e. on Ω, and $\nu(\overline{u})(\cdot,\cdot)$ is therefore equal to $\nu(\overline{u}(\cdot,\cdot))$ a.e. on $\Omega\times(0,T)$.
It remains to establish that $\nu(\overline{u})$ thus defined is continuous $[0,T]\mapsto L^2(\Omega)$. For any $(\tau_r)_{r\in\mathbb{N}}\subset[0,T]$ which converges to $\tau\in[0,T]$, we can pick $s_r\in\mathcal{T}\cap(\tau_r-\frac{1}{r},\tau_r+\frac{1}{r})$ and $t_r\in\mathcal{T}\cap(\tau-\frac{1}{r},\tau+\frac{1}{r})$ such that

$$||\nu(\overline{u})(\cdot,\tau_r) - \nu(\overline{u}(\cdot,s_r))||_{L^2(\Omega)} \le \frac{1}{r}, \quad ||\nu(\overline{u})(\cdot,\tau) - \nu(\overline{u}(\cdot,t_r))||_{L^2(\Omega)} \le \frac{1}{r}.$$

We therefore have

$$||\nu(\overline{u})(\cdot,\tau_r) - \nu(\overline{u})(\cdot,\tau)||_{L^2(\Omega)} \le \frac{2}{r} + ||\nu(\overline{u}(\cdot,s_r)) - \nu(\overline{u}(\cdot,t_r))||_{L^2(\Omega)},$$

which proves by (38) with $l = k = r$ that $\nu(\overline{u})(\cdot,\tau_r) \to \nu(\overline{u})(\cdot,\tau)$ in $L^2(\Omega)$ as $r \to \infty$ and completes the proof. ■

4 Proof of the convergence theorems

4.1 Estimates on the approximate solution

As it is usual in the study of numerical methods for PDE with strong non-linearities or without regularity assumptions on the data, everything starts with *a priori* estimates.

Lemma 4.1 ($L^\infty(0,T;L^2(\Omega))$ estimate and discrete $L^p(0,T;W_0^{1,p}(\Omega))$ estimate) *Under Assumptions* (2)*, let $\mathcal{D}$ be a space-time gradient discretisation in the sense of Definition 2.1. Let u be a solution to Scheme* (14)*.*
Then, for any $T_0 \in (0,T]$, denoting by $k = 1,\ldots,N$ the index such that $T_0 \in (t^{(k-1)}, t^{(k)}]$ we have

$$\int_\Omega B(\Pi_\mathcal{D}\beta(u)(\boldsymbol{x},T_0))\mathrm{d}\boldsymbol{x} + \int_0^{T_0}\int_\Omega \boldsymbol{a}(\boldsymbol{x},\Pi_\mathcal{D}\nu(u)(\boldsymbol{x},t),\nabla_\mathcal{D}\zeta(u)(\boldsymbol{x},t))\cdot\nabla_\mathcal{D}\zeta(u)(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x}\mathrm{d}t$$
$$\le \int_\Omega B(\Pi_\mathcal{D}\beta(\mathcal{I}_\mathcal{D}u_{\rm ini})(\boldsymbol{x}))\mathrm{d}\boldsymbol{x} + \int_0^{t^{(k)}}\int_\Omega f(\boldsymbol{x},t)\Pi_\mathcal{D}\zeta(u)(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x}\mathrm{d}t. \quad (39)$$

Consequently, there exists $C_1 > 0$, only depending on p, L_β, $C_P \ge C_\mathcal{D}$ (see Definition 2.5), $C_{\rm ini} \ge \|\Pi_\mathcal{D}\mathcal{I}_\mathcal{D}u_{\rm ini}\|_{L^2(\Omega)}$, f and $\underline{a}$ such that

$$\|\Pi_\mathcal{D}B(\beta(u))\|_{L^\infty(0,T;L^1(\Omega))} \le C_1\,,\ \|\nabla_\mathcal{D}\zeta(u)\|_{L^p(\Omega\times(0,T))^d} \le C_1$$
$$\text{and } \|\Pi_\mathcal{D}\beta(u)\|_{L^\infty(0,T;L^2(\Omega))} \le C_1. \quad (40)$$

Proof. Using (12) and (27), we notice that for any $n = 0,\ldots,N-1$, any $t \in (t^{(n)}, t^{(n+1)}]$,

$$\begin{aligned}\Pi_\mathcal{D}\delta_\mathcal{D}\beta(u)(t)\Pi_\mathcal{D}\zeta(u^{(n+1)}) &= \frac{1}{\delta t^{(n+\frac12)}}\Big(\beta(\Pi_\mathcal{D}u^{(n+1)}) - \beta(u^{(n)})\Big)\zeta(\Pi_\mathcal{D}u^{(n+1)})\\ &\ge \frac{1}{\delta t^{(n+\frac12)}}\Big(B(\Pi_\mathcal{D}\beta(u^{(n+1)})) - B(\Pi_\mathcal{D}\beta(u^{(n)}))\Big).\end{aligned}$$

Hence, taking $v = (\zeta(u^{(0)}),\zeta(u^{(1)}),\ldots,\zeta(u^{(k)}),0,\ldots,0) \subset X_{\mathcal{D},0}$ in (14), we find

$$\int_\Omega B(\Pi_\mathcal{D}\beta(u)(\boldsymbol{x},t^{(k)}))\mathrm{d}\boldsymbol{x} + \int_0^{t^{(k)}}\int_\Omega \boldsymbol{a}(\boldsymbol{x},\Pi_\mathcal{D}\nu(u)(\boldsymbol{x},t),\nabla_\mathcal{D}\zeta(u)(\boldsymbol{x},t))\cdot\nabla_\mathcal{D}\zeta(u)(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x}\mathrm{d}t$$
$$\le \int_\Omega B(\Pi_\mathcal{D}\beta(u^{(0)})(\boldsymbol{x}))\mathrm{d}\boldsymbol{x} + \int_0^{t^{(k)}} f(\boldsymbol{x},t)\Pi_\mathcal{D}\zeta(u)(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x}\mathrm{d}t \quad (41)$$

Equation (39) is a straightforward consequence of this estimate, of the relation $\beta(u)(\cdot,T_0) = \beta(u)(\cdot,t^{(k)})$ (see (13)) and of the fact that the integrand involving $\boldsymbol{a}$ is nonnegative on $[T_0,t^{(k)}]$. Using the Young inequality, we can write

$$\int_0^{t^{(k)}}\int_\Omega f(\boldsymbol{x},t)\Pi_{\mathcal D}\zeta(u)(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x}\mathrm{d}t$$
$$\leq \frac{2^{1/(p-1)}C_{\mathcal D}^{p'}}{(p\underline{a})^{1/(p-1)}\,p'}\|f\|^{p'}_{L^{p'}(\Omega\times(0,t^{(k)}))}+\frac{\underline{a}}{2C_{\mathcal D}^p}\|\Pi_{\mathcal D}\zeta(u)\|^p_{L^p(\Omega\times(0,t^{(k)}))}$$

and the first two estimates in (40) therefore follow from (41), the coercivity assumption (2f) on $\boldsymbol{a}$ and (26). The estimate on $\Pi_{\mathcal D}\beta(u)=\beta(\Pi_{\mathcal D}u)$ in $L^\infty(0,T;L^2(\Omega))$ is a consequence of the estimate on $B(\beta(\Pi_{\mathcal D}u))$ in $L^\infty(0,T;L^1(\Omega))$ and of (26). ■

Corollary 4.2 (Existence of a solution to the Gradient Scheme) *Under Assumptions* (2), *if $\mathcal D$ is a gradient discretisation in the sense of Definition 2.1 then there exists at least a solution to the Gradient Scheme* (14).

Proof. For $\rho\in[0,1]$ we let $\beta_\rho(u)=\rho u+(1-\rho)\beta(u)$, $\zeta_\rho(u)=\rho u+(1-\rho)\zeta(u)$ and $\boldsymbol{a}_\rho(\boldsymbol{x},s,\boldsymbol{\xi})=\rho\boldsymbol{\xi}+(1-\rho)\boldsymbol{a}(\boldsymbol{x},s,\boldsymbol{\xi})$. It is clear that β_ρ, ζ_ρ and $\boldsymbol{a}_\rho$ satisfy the same assumptions as β, ζ and $\boldsymbol{a}$ for some L_β, M_0, M_1 and $\underline{a}$ not depending on ρ. We can therefore apply Lemma 4.1 to see that there exists C_2 not depending on s such that any solution u_ρ to (14) with $\beta=\beta_\rho$, $\zeta=\zeta_\rho$ and $\boldsymbol{a}=\boldsymbol{a}_\rho$ satisfies
$$||\nabla_{\mathcal D}\zeta_\rho(u_\rho)||_{L^p((0,T)\times\Omega)^d}\leq C_2.$$
Since $||\nabla_{\mathcal D}\cdot||_{L^p(\Omega\times(0,T))^d}$ is a norm on $X_{\mathcal D,0}$, this shows that $(\zeta_\rho(u_\rho))_{\rho\in[0,1]}$ remains bounded in this finite dimensional space, which implies in particular that for all $i\in I$, $(\zeta_\rho(u_\rho)_i)_{\rho\in[0,1]}$. Using Assumption 2b for ζ_ρ with constants not depending on ρ, we deduce that $((u_\rho)_i)_{\rho\in[0,1]}$ remains bounded for any $i\in I$, and thus that $(u_\rho)_{\rho\in[0,1]}$ is bounded in $X_{\mathcal D,0}$.
But if $\rho=0$ then (14) is a linear scheme. Any solution to this scheme being bounded in $X_{\mathcal D,0}$, this shows that the underlying linear system is invertible. A topological degree argument [10] then shows, combined with the uniform bound on $(u_\rho)_{\rho\in[0,1]}$, that the scheme corresponding to $\rho=1$, that is (14), possesses at least one solution. ■

Lemma 4.3 (Estimate on the dual semi-norm of the discrete time derivative) *Under Assumptions* (2), *let $\mathcal D$ be a space-time gradient discretisation in the sense of Definition 2.1. Let u be a solution to Scheme* (14). *Then there exists C_3, only depending on p, L_β, $C_P\geq C_{\mathcal D}$, $C_{\rm ini}\geq\|\Pi_{\mathcal D}I_{\mathcal D}u_{\rm ini}\|_{L^2(\Omega)}$, f, $\underline{a}$, μ, $\overline{a}$ and T, such that*

$$\int_0^T|\delta_{\mathcal D}\beta(u)(t)|^{p'}_{\star,\mathcal D}\mathrm{d}t\leq C_3. \tag{42}$$

Proof. Let us take a generic $v=(v^{(n)})_{n=1,\ldots,N}\subset\in X_{\mathcal D,0}$ as test function in Scheme (14). We have, thanks to Assumption (2h) on $\boldsymbol{a}$,

$$\int_0^T\int_\Omega\Pi_{\mathcal D}\delta_{\mathcal D}\beta(u)(\boldsymbol{x},t)\Pi_{\mathcal D}v(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x}\mathrm{d}t\leq\int_0^T\int_\Omega(\overline{a}(\boldsymbol{x})+\mu|\nabla_{\mathcal D}\zeta(u)(\boldsymbol{x},t)|^{p-1})|\nabla_{\mathcal D}v(\boldsymbol{x},t)|\mathrm{d}\boldsymbol{x}\mathrm{d}t$$
$$+\int_0^T\int_\Omega f(\boldsymbol{x},t)\Pi_{\mathcal D}v(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x}\mathrm{d}t.$$

Using Hölder's inequality, Definition 2.5 and Estimates (40), this leads to the existence of $C_4>0$ only depending on p, L_β, C_P, $C_{\rm ini}$, f, $\underline{a}$, $\overline{a}$ and μ such that

$$\int_0^T\int_\Omega\Pi_{\mathcal D}\delta_{\mathcal D}\beta(u)(\boldsymbol{x},t)\Pi_{\mathcal D}v(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x}\mathrm{d}t\leq C_4\|\nabla_{\mathcal D}v\|_{L^p(0,T;L^p(\Omega))^d}.$$

The proof of (42) is completed by selecting $v = (|\delta_{\mathcal{D}}^{(n+\frac{1}{2})}\beta(u)|_{\star,\mathcal{D}}^{p'-1} z^{(n)})_{n=0,\ldots,N}$ with $(z^{(n)})_{n=0,\ldots,N} \subset X_{\mathcal{D},0}$ such that, for any $n = 0, \ldots, N-1$, $z^{(n+1)}$ realises the supremum in (11) with $w = \delta_{\mathcal{D}}^{(n+\frac{1}{2})}\beta(u)$. ■

Lemma 4.4 (Estimate on the time translates of $\nu(u)$)
Under Assumptions (2), *let $\mathcal{D}$ be a space-time gradient discretisation in the sense of Definition 2.1. Let u be a solution to Scheme* (14). *Then there exists C_5, only depending on p, L_β, L_ζ, $C_P \geq C_{\mathcal{D}}$, $C_{\rm ini} \geq \|\Pi_{\mathcal{D}} I_{\mathcal{D}} u_{\rm ini}\|_{L^2(\Omega)}$, f, $\underline{a}$, μ, $\overline{a}$ and T, such that*

$$\|\Pi_{\mathcal{D}}\nu(u)(\cdot, \cdot+\tau) - \Pi_{\mathcal{D}}\nu(u)(\cdot,\cdot)\|_{L^2(\Omega\times(0,T-\tau))}^2 \leq C_5(\tau + \delta\!t), \quad \forall \tau \in (0,T). \tag{43}$$

Proof. Let $\tau \in (0,T)$. Thanks to (24), we get that

$$\int_{\Omega\times(0,T-\tau)} \Big(\Pi_{\mathcal{D}}\nu(u)(\boldsymbol{x}, t+\tau) - \Pi_{\mathcal{D}}\nu(u)(\boldsymbol{x},t)\Big)^2 \mathrm{d}\boldsymbol{x}\mathrm{d}t \leq L_\beta L_\zeta \int_0^{T-\tau} A(t)\mathrm{d}t, \tag{44}$$

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

$$A(t) = \int_\Omega \Big(\Pi_{\mathcal{D}}\zeta(u)(\boldsymbol{x}, t+\tau) - \Pi_{\mathcal{D}}\zeta(u)(\boldsymbol{x},t)\Big)\Big(\Pi_{\mathcal{D}}\beta(u)(\boldsymbol{x}, t+\tau) - \Pi_{\mathcal{D}}\beta(u)(\boldsymbol{x},t)\Big)\mathrm{d}\boldsymbol{x}.$$

Let $t \in (0, T-\tau)$. Letting $n_0(t)$, $n_1(t) = 0, \ldots, N-1$ be 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

$$A(t) = \int_\Omega \Big(\Pi_{\mathcal{D}}\zeta(u^{(n_1(t)+1)})(\boldsymbol{x}) - \Pi_{\mathcal{D}}\zeta(u^{(n_0(t)+1)})(\boldsymbol{x})\Big)\Big(\sum_{n=n_0(t)+1}^{n_1(t)} \delta\!t^{(n+\frac{1}{2})}\Pi_{\mathcal{D}}\delta_{\mathcal{D}}^{(n+\frac{1}{2})}\beta(u)(\boldsymbol{x})\Big)\mathrm{d}\boldsymbol{x},$$

which also reads

$$A(t) = \int_\Omega \Big(\Pi_{\mathcal{D}}\zeta(u^{(n_1(t)+1)})(\boldsymbol{x}) - \Pi_{\mathcal{D}}\zeta(u^{(n_0(t)+1)})(\boldsymbol{x})\Big) \\ \times \Big(\sum_{n=1}^{N-1} \chi_n(t, t+\tau)\delta\!t^{(n+\frac{1}{2})}\Pi_{\mathcal{D}}\delta_{\mathcal{D}}^{(n+\frac{1}{2})}\beta(u)(\boldsymbol{x})\Big)\mathrm{d}\boldsymbol{x}, \tag{45}$$

with $\chi_n(t,t+\tau) = 1$ if $t^{(n)} \in (t, t+\tau]$ and $\chi_n(t,t+\tau) = 0$ if $t^{(n)} \notin (t,t+\tau]$. We then let $v = (\chi_{n-1}(t,t+\tau)(\zeta(u^{(n_1(t)+1)}) - \zeta(u^{(n_0(t)+1)})))_{n=0,\ldots,N}$ in Scheme (14). Using (45), we get

$$A(t) = A_3(t) - \sum_{n=1}^{N-1} \chi_n(t,t+\tau)\delta\!t^{(n+\frac{1}{2})} \int_\Omega \boldsymbol{a}\Big(\boldsymbol{x}, \Pi_{\mathcal{D}}\nu(u^{(n+1)}), \nabla_{\mathcal{D}}\zeta(u^{(n+1)})(\boldsymbol{x})\Big) \\ \cdot \Big(\nabla_{\mathcal{D}}\zeta(u^{(n_1(t)+1)})(\boldsymbol{x}) - \nabla_{\mathcal{D}}\zeta(u^{(n_0(t)+1)})(\boldsymbol{x})\Big)\mathrm{d}\boldsymbol{x}$$

with

$$A_3(t) = \sum_{n=1}^{N-1} \chi_n(t,t+\tau) \int_\Omega \int_{t^{(n)}}^{t^{(n+1)}} f(\boldsymbol{x},t)\mathrm{d}t\Big(\Pi_{\mathcal{D}}\zeta(u^{(n_1(t)+1)})(\boldsymbol{x}) - \Pi_{\mathcal{D}}\zeta(u^{(n_0(t)+1)})(\boldsymbol{x})\Big)\mathrm{d}\boldsymbol{x}.$$

Using the Young inequality and the inequality $a_1a_2 \le \frac{(p-1)^{1/p'}}{p}(a_1^p + a_2^{p'})$ for $a_1 \ge 0$ and $a_2 \ge 0$, this yields:

$$A(t) \le \frac{(p-1)^{1/p'}}{p}(A_0(t) + A_1(t) + 2A_2(t)) + A_3(t), \tag{46}$$

with

$$A_0(t) = \sum_{n=1}^{N-1} \chi_n(t,t+\tau)\delta t^{(n+\frac{1}{2})} \int_\Omega \left|\nabla_{\mathcal{D}}\zeta(u^{(n_0(t)+1)})(\boldsymbol{x})\right|^p \mathrm{d}\boldsymbol{x},$$

$$A_1(t) = \sum_{n=1}^{N-1} \chi_n(t,t+\tau)\delta t^{(n+\frac{1}{2})} \int_\Omega \left|\nabla_{\mathcal{D}}\zeta(u^{(n_1(t)+1)})(\boldsymbol{x})\right|^p \mathrm{d}\boldsymbol{x},$$

and

$$A_2(t) = \sum_{n=1}^{N-1} \chi_n(t,t+\tau)\delta t^{(n+\frac{1}{2})} \int_\Omega \left|\boldsymbol{a}\left(\boldsymbol{x}, \Pi_{\mathcal{D}}\nu(u^{(n+1)}), \nabla_{\mathcal{D}}\zeta(u^{(n+1)})(\boldsymbol{x})\right)\right|^{p'} \mathrm{d}\boldsymbol{x}.$$

Applying (80) of Lemma 6.6 yields

$$\int_0^{T-\tau} A_0(t) + A_1(t)\mathrm{d}t \le 2(\tau + \delta t)\|\nabla_{\mathcal{D}}\zeta(u)\|^p_{L^p(\Omega\times(0,T))}. \tag{47}$$

Applying (79) of Lemma 6.6 yields

$$\int_0^{T-\tau} A_2(t)\mathrm{d}t \le \tau\|\boldsymbol{a}\left(\cdot, \Pi_{\mathcal{D}}\nu(u), \nabla_{\mathcal{D}}\zeta(u)\right)\|^{p'}_{L^{p'}(\Omega\times(0,T))}, \tag{48}$$

and, similarly, we obtain using Definition (2.5) and (40), the existence of C_6 such that

$$\textstyle\int_0^{T-\tau} A_3(t)\mathrm{d}t \le C_6((\tau + \delta t) + \tau\|f\|^{p'}_{L^{p'}(\Omega\times(0,T))}). \tag{49}$$

Using inequalities (44), (46) and (47)-(49), (43) is proved. ∎

4.2 Proof of Theorem 2.10

Step 1 Application of compactness results.
Thanks to Theorem 3.1 and Estimates (40) and (42), we first extract a subsequence such that $(\Pi_{\mathcal{D}_m}\beta(u_m))_{m\in\mathbb{N}}$ converges weakly in $L^2(\Omega)$ uniformly in $[0,T]$ (in the sense of Definition 6.4) to some function $\overline{\beta} \in C([0,T]; L^2(\Omega)\text{-w})$ which satisfies $\overline{\beta}(\cdot,0) = \beta(u_{\rm ini})$ in $L^2(\Omega)$. Using again Estimates (40) and applying Lemma 2.9, we again extract a subsequence such that, for some $\overline{\zeta} \in L^p(0,T; W_0^{1,p}(\Omega))$, $\Pi_{\mathcal{D}_m}\zeta(u_m) \to \overline{\zeta}$ weakly in $L^p(\Omega\times(0,T))$ and $\nabla_{\mathcal{D}_m}\zeta(u_m) \to \nabla\overline{\zeta}$ weakly in $L^p(\Omega\times(0,T))^d$. From Estimates (40), Definition 2.5 and the growth assumption (2b) on ζ, we also see that $(\Pi_{\mathcal{D}_m}u_m)_{m\in\mathbb{N}}$ is bounded in $L^p(\Omega\times(0,T))$ and we can therefore assume, up to a subsequence, that it converges weakly to some $\overline{u}$ in this space.
We then remark that $|\nu(a) - \nu(b)| \le L_\beta|\zeta(a) - \zeta(b)|$, which implies that, using Definition 2.7 with $v = \zeta(u_m)$ and (40),

$$||\Pi_{\mathcal{D}_m}\nu(u_m)(\cdot + \boldsymbol{\xi}) - \Pi_{\mathcal{D}_m}\nu(u)||_{L^p(\mathbb{R}^d\times(0,T))} \le L_\beta C_1 T_{\mathcal{D}_m}(\boldsymbol{\xi}) \tag{50}$$

where $\Pi_{\mathcal{D}}\nu(u)$ has been extended by 0 outside Ω and $\lim_{\boldsymbol{\xi}\to 0}\sup_{m\in\mathbb{N}} T_{\mathcal{D}_m}(\boldsymbol{\xi}) = 0$. Invoking Lemma 4.4, we also see that the time translates of $\Pi_{\mathcal{D}_m}\nu(u_m)$ tend uniformly to 0 in $L^1(\Omega\times(0,T))$. We

use the fact that $\Pi_{\mathcal{D}_m}\beta(u_m)$, and therefore also $\Pi_{\mathcal{D}_m}\nu(u_m)$, remains bounded in $L^\infty(0,T;L^2(\Omega))$ to control the time translates at both ends of $[0,T]$. Hence, applying Kolmogorov's theorem, we deduce that, up to the extraction of another subsequence, $\Pi_{\mathcal{D}_m}\nu(u_m) \to \overline{\nu}$ in $L^1(\Omega\times(0,T))$.
Under the first case in the structural hypothesis (17), we have $\beta = \mathrm{Id}$ and therefore $\overline{\beta} = \overline{u} = \beta(\overline{u})$, and $\nu = \zeta$. The strong convergence of $\Pi_{\mathcal{D}_m}\nu(u_m) = \Pi_{\mathcal{D}_m}\zeta(u_m)$ to $\overline{\nu} = \overline{\zeta}$ thus allows us to apply Lemma 3.5 to see that $\overline{\zeta} = \zeta(\overline{u})$. Exchanging the roles of β and ζ, we see that $\overline{\beta} = \beta(\overline{u})$ and $\overline{\zeta} = \zeta(\overline{u})$ still hold in the second case of (17). We notice that this is the only place where we need this structural assumption (17) on β,ζ.
Upon extraction of another subsequence, we can also assume that $\boldsymbol{a}(\cdot,\Pi_{\mathcal{D}}\nu(u),\nabla_{\mathcal{D}}\zeta(u))$ has a weak limit in $L^{p'}(\Omega\times(0,T))^d$, which we denote by $\boldsymbol{A}$.
Finally, for any $T_0\in[0,T]$, since $\Pi_{\mathcal{D}_m}\beta(u_m(\cdot,T_0)) \to \beta(\overline{u})(\cdot,T_0)$ weakly in $L^2(\Omega)$, Lemma 3.4 gives

$$\int_\Omega B(\beta(\overline{u})(\boldsymbol{x},T_0))\mathrm{d}\boldsymbol{x} \le \liminf_{m\to\infty}\int_\Omega B(\beta(\Pi_{\mathcal{D}_m}u_m)(\boldsymbol{x},T_0))\mathrm{d}\boldsymbol{x}. \tag{51}$$

This shows in particular that $B(\beta(\overline{u}))\in L^\infty(0,T;L^1(\Omega))$.

Step 2 Passing to the limit in the scheme. We drop the indices m for legibility reasons.
Let $\varphi\in C_c^1(-\infty,T)$ and $w\in W_0^{1,p}(\Omega)\cap L^2(\Omega)$. We introduce $v = (\varphi(t^{(n-1)})P_{\mathcal{D}}w)_{n=0,\ldots,N}$ (where $t^{(-1)}=0$ for example, this value being irrelevant) as test function in (14), with $P_{\mathcal{D}}$ defined by (21). We get $T_1^{(m)}+T_2^{(m)}=T_3^{(m)}$ with

$$T_1^{(m)} = \sum_{n=0}^{N-1}\varphi(t^{(n)})\delta\!t^{(n+\frac12)}\int_\Omega \Pi_{\mathcal{D}}\delta_{\mathcal{D}}^{(n+\frac12)}\beta(u)(\boldsymbol{x})\Pi_{\mathcal{D}}P_{\mathcal{D}}w(\boldsymbol{x})\mathrm{d}\boldsymbol{x},$$

$$T_2^{(m)} = \sum_{n=0}^{N-1}\varphi(t^{(n)})\delta\!t^{(n+\frac12)}\int_\Omega \boldsymbol{a}\left(\boldsymbol{x},\Pi_{\mathcal{D}}\nu(u^{(n+1)}),\nabla_{\mathcal{D}}\zeta(u^{(n+1)})(\boldsymbol{x})\right)\cdot\nabla_{\mathcal{D}}P_{\mathcal{D}}w(\boldsymbol{x})\mathrm{d}\boldsymbol{x},$$

and

$$T_3^{(m)} = \sum_{n=0}^{N-1}\varphi(t^{(n)})\int_{t^{(n)}}^{t^{(n+1)}}\int_\Omega f(\boldsymbol{x},t)\Pi_{\mathcal{D}}P_{\mathcal{D}}w(\boldsymbol{x})\mathrm{d}\boldsymbol{x}\mathrm{d}t.$$

Using discrete integrate-by-parts to transform the terms $\varphi(t^{(n)})(\Pi_{\mathcal{D}}\beta(u^{(n+1)})-\Pi_{\mathcal{D}}\beta(u^{(n)}))$ appearing in $T_1^{(m)}$ into $(\varphi(t^{(n)})-\varphi(t^{(n+1)}))\Pi_{\mathcal{D}}\beta(u^{(n+1)})$, we have

$$T_1^{(m)} = -\int_0^T\varphi'(t)\int_\Omega \Pi_{\mathcal{D}}\beta(u)(\boldsymbol{x},t)\Pi_{\mathcal{D}}P_{\mathcal{D}}w(\boldsymbol{x})\mathrm{d}\boldsymbol{x}\mathrm{d}t - \varphi(0)\int_\Omega \Pi_{\mathcal{D}}\beta(u^{(0)})(\boldsymbol{x})\Pi_{\mathcal{D}}P_{\mathcal{D}}w(\boldsymbol{x})\mathrm{d}\boldsymbol{x}.$$

Setting $\varphi_{\mathcal{D}}(t)=\varphi(t^{(n)})$ for $t\in(t^{(n)},t^{(n+1)})$, we have

$$\begin{aligned} T_2^{(m)} &= \int_0^T\varphi_{\mathcal{D}}(t)\int_\Omega \boldsymbol{a}\left(\boldsymbol{x},\Pi_{\mathcal{D}}\nu(u)(\boldsymbol{x},t),\nabla_{\mathcal{D}}\zeta(u)(\boldsymbol{x},t)\right)\cdot\nabla_{\mathcal{D}}P_{\mathcal{D}}w(\boldsymbol{x})\mathrm{d}\boldsymbol{x}\mathrm{d}t\\ T_3^{(m)} &= \int_0^T\varphi_{\mathcal{D}}(t)\int_\Omega f(\boldsymbol{x},t)\Pi_{\mathcal{D}}P_{\mathcal{D}}w(\boldsymbol{x})\mathrm{d}\boldsymbol{x}\mathrm{d}t.\end{aligned}$$

We may thus let $m\to\infty$ in $T_1^{(m)}+T_2^{(m)}=T_3^{(m)}$ to see that $\overline{u}$ satisfies

$$\left\{\begin{array}{l} \overline{u}\in L^p(\Omega\times(0,T))\,,\ \zeta(\overline{u})\in L^p(0,T;W_0^{1,p}(\Omega))\,,\ B(\beta(\overline{u}))\in L^\infty(0,T;L^1(\Omega)),\\ \displaystyle -\int_0^T\varphi'(t)\int_\Omega\beta(\overline{u}(\boldsymbol{x},t))w(\boldsymbol{x})\mathrm{d}\boldsymbol{x}\mathrm{d}t-\varphi(0)\int_\Omega u_{\mathrm{ini}}(\boldsymbol{x})w(\boldsymbol{x})\mathrm{d}\boldsymbol{x}\\ \displaystyle +\int_0^T\varphi(t)\int_\Omega\boldsymbol{A}(\boldsymbol{x},t)\cdot\nabla w(\boldsymbol{x})\mathrm{d}\boldsymbol{x}\mathrm{d}t=\int_0^T\varphi(t)\int_\Omega f(\boldsymbol{x},t)w(\boldsymbol{x})\mathrm{d}\boldsymbol{x}\mathrm{d}t,\\ \qquad\forall w\in W_0^{1,p}(\Omega)\cap L^2(\Omega),\ \forall\varphi\in C_c^\infty(-\infty,T).\end{array}\right. \tag{52}$$

Linear combinations of this relation show that it also holds with $\varphi(t)w(\boldsymbol{x})$ replaced by a tensorial functions in $C_c^\infty(\Omega\times(0,T))$. This proves that $\partial_t\beta(\overline{u})\in L^{p'}(0,T;W^{-1,p'}(\Omega))$. Standard arguments then show that $\beta(\overline{u})$ can be identified with an element of $C^0([0,T];L^2(\Omega)\text{-w})$ with the property $\beta(\overline{u}(\cdot,0))=\beta(u_{\rm ini})$ (cf. discussion at the end of Section 1.2). Using the density of tensorial functions in $L^p(0,T;W_0^{1,p}(\Omega))$, we then see that $\overline{u}$ satisfies

$$\begin{aligned}&\int_0^T\langle\partial_t\beta(\overline{u})(\cdot,t),\overline{v}(\cdot,t)\rangle_{(W^{-1,p'}(\Omega))'W_0^{1,p}(\Omega)}\mathrm{d}t\\&+\int_0^T\int_\Omega\boldsymbol{A}(\boldsymbol{x},t)\cdot\nabla\overline{v}(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x}\mathrm{d}t=\int_0^T\int_\Omega f(\boldsymbol{x},t)\overline{v}(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x}\mathrm{d}t\,,\quad\forall\overline{v}\in L^p(0,T;W_0^{1,p}(\Omega)).\end{aligned}\tag{53}$$

Step 3 Proof that $\overline{u}$ is solution to (4).

The proof will be completed by showing that

$$\boldsymbol{A}(\boldsymbol{x},t)=\boldsymbol{a}(\boldsymbol{x},\nu(\overline{u})(\boldsymbol{x},t),\nabla\zeta(\overline{u})(\boldsymbol{x},t)),\ \text{for a.e. }(\boldsymbol{x},t)\in\Omega\times(0,T).\tag{54}$$

We take $T_0\in[0,T]$, write (39) with $\mathcal{D}=\mathcal{D}_m$ and take the supremum limit as $m\to\infty$. We notice that the $t^{(k)}=T_m$ from Lemma 4.1 converges to T_0 as $m\to\infty$. Hence, using the quadratic growth of B we obtain

$$\begin{aligned}\limsup_{m\to\infty}\int_0^{T_0}\int_\Omega&\boldsymbol{a}(\boldsymbol{x},\Pi_{\mathcal{D}_m}\nu(u_m)(\boldsymbol{x},t),\nabla_{\mathcal{D}_m}\zeta(u_m)(\boldsymbol{x},t))\cdot\nabla_{\mathcal{D}_m}\zeta(u_m)(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x}\mathrm{d}t\\&\le\int_\Omega B(\beta(u_{\rm ini})(\boldsymbol{x}))\mathrm{d}\boldsymbol{x}+\int_0^{T_0}\int_\Omega f(\boldsymbol{x},t)\zeta(\overline{u})(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x}\mathrm{d}t\\&\qquad-\liminf_{m\to\infty}\int_\Omega B(\beta(\Pi_{\mathcal{D}_m}u_m)(\boldsymbol{x},T_0))\mathrm{d}\boldsymbol{x}.\end{aligned}\tag{55}$$

We then apply Lemma 3.6 and take $\overline{v}=\zeta(\overline{u})$ in (53) to get

$$\begin{aligned}&\int_\Omega B(\beta(\overline{u})(\boldsymbol{x},T_0))\mathrm{d}\boldsymbol{x}-\int_\Omega B(\beta(\overline{u})(\boldsymbol{x},0))\mathrm{d}\boldsymbol{x}\\&+\int_0^{T_0}\int_\Omega\boldsymbol{A}(\boldsymbol{x},t)\cdot\nabla\zeta(\overline{u})(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x}\mathrm{d}t=\int_0^{T_0}\int_\Omega f(\boldsymbol{x},t)\zeta(\overline{u})(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x}\mathrm{d}t.\end{aligned}$$

This relation, combined with (55) and using (51), shows that

$$\begin{aligned}\limsup_{m\to\infty}\int_0^{T_0}\int_\Omega&\boldsymbol{a}(\boldsymbol{x},\Pi_{\mathcal{D}_m}\nu(u_m)(\boldsymbol{x},t),\nabla_{\mathcal{D}_m}\zeta(u_m)(\boldsymbol{x},t))\cdot\nabla_{\mathcal{D}_m}\zeta(u_m)(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x}\mathrm{d}t\\&\le\int_0^{T_0}\int_\Omega\boldsymbol{A}(\boldsymbol{x},t)\cdot\nabla\zeta(\overline{u})\boldsymbol{x},t)\mathrm{d}\boldsymbol{x}\mathrm{d}t.\end{aligned}\tag{56}$$

It is now possible to apply Minty's trick. Consider, for $\boldsymbol{G}\in L^p(0,T;L^p(\Omega))^d$,

$$\int_0^{T_0}\int_\Omega\quad[\boldsymbol{a}(\cdot,\Pi_{\mathcal{D}}\nu(u),\nabla_{\mathcal{D}}\zeta(u))-\boldsymbol{a}(\cdot,\Pi_{\mathcal{D}}\nu(u),\boldsymbol{G})]\cdot[\nabla_{\mathcal{D}}\zeta(u)-\boldsymbol{G}]\,\mathrm{d}\boldsymbol{x}\mathrm{d}t\ge0.\tag{57}$$

By strong convergence of $\Pi_{\mathcal{D}_m}\nu(u_m)$ to $\nu(\overline{u})$ in $L^1(\Omega\times(0,T))$ and Assumptions (2e), (2h) on $\boldsymbol{a}$, we see that $\boldsymbol{a}(\cdot,\Pi_{\mathcal{D}_m}\nu(u_m),\boldsymbol{G})\to\boldsymbol{a}(\cdot,\nu(\overline{u}),\boldsymbol{G})$ strongly in $L^{p'}(\Omega\times(0,T))$. We can therefore develop

(57) (with $T_0 = T$) and use (56) to pass to the supremum limit on the only "weak-weak" term, and we find, for any $\boldsymbol{G} \in L^p(0,T;L^p(\Omega))^d$,

$$\int_0^T \int_\Omega [\boldsymbol{A}(\boldsymbol{x},t) - \boldsymbol{a}(\boldsymbol{x},\nu(\overline{u})(\boldsymbol{x},t),\boldsymbol{G}(\boldsymbol{x},t))] \cdot [\nabla\zeta(\overline{u})(\boldsymbol{x},t) - \boldsymbol{G}(\boldsymbol{x},t)]\,\mathrm{d}\boldsymbol{x}\mathrm{d}t \geq 0.$$

Application of Minty's method [33] (i.e. taking $\boldsymbol{G} = \nabla\zeta(\overline{u}) + r\boldsymbol{\varphi}$ for $\boldsymbol{\varphi} \in L^p(0,T;L^p(\Omega))^d$ and letting $r \to 0$) then shows that (54) holds and concludes the proof that $\overline{u}$ is a weak solution to (4).

4.3 Proof of Theorem 2.14

Let $T_0 \in [0,T]$ and $(T_m)_{m\geq 1}$ be a sequence in $[0,T]$ which converges to T_0. By setting $T_0 = T_m$ and $\boldsymbol{G} = \nabla\zeta(\overline{u})$ in the developed form of (57), by taking the infimum limit (thanks to the strong convergence of $\boldsymbol{a}(\cdot,\Pi_{\mathcal{D}_m}\nu(u_m),\nabla\zeta(\overline{u}))$) and by using (54), we find

$$\liminf_{m\to\infty} \int_0^{T_m} \int_\Omega \boldsymbol{a}(\boldsymbol{x},\Pi_{\mathcal{D}_m}\nu(u_m)(\boldsymbol{x},t),\nabla_{\mathcal{D}_m}\zeta(u_m)(\boldsymbol{x},t)) \cdot \nabla_{\mathcal{D}_m}\zeta(u_m)(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x}\mathrm{d}t$$
$$\geq \int_0^{T_0} \int_\Omega \boldsymbol{a}(\boldsymbol{x},\nu(\overline{u})(\boldsymbol{x},t),\nabla\zeta(\overline{u})(\boldsymbol{x},t)) \cdot \nabla\zeta(\overline{u})(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x}\mathrm{d}t. \quad (58)$$

We then write (39) with T_m instead of T_0. We notice that the $t^{(k)} = t^{(k(m))}$ such that $T_m \in (t^{(k-1)},t^{(k)}]$ converges to T_0 as $m \to \infty$. Using (58) and Corollary 3.8, we therefore obtain

$$\limsup_{m\to\infty} \int_\Omega B(\beta(\Pi_{\mathcal{D}_m}u_m(\boldsymbol{x},T_m)))\mathrm{d}\boldsymbol{x} \leq \int_\Omega B(\beta(\overline{u})(\boldsymbol{x},T_0))\mathrm{d}\boldsymbol{x}. \quad (59)$$

By Lemma 6.3, the uniform-in-time weak convergence of $\beta(\Pi_{\mathcal{D}_m}u_m)$ to $\beta(\bar{u})$ and the weak continuity of $\beta(\bar{u}) : [0,T] \to L^2(\Omega)$, we have $\beta(\Pi_{\mathcal{D}_m}u_m)(T_m) \to \beta(\bar{u})(T_0)$ weakly in $L^2(\Omega)$ as $m \to \infty$. Therefore, for any $(s_m)_{m\in\mathbb{N}}$ converging to T_0, $\frac{1}{2}(\beta(\Pi_{\mathcal{D}_m}u_m(T_m)) + \beta(\overline{u})(s_m)) \to \beta(\overline{u})(T_0)$ weakly in $L^2(\Omega)$ as $m \to \infty$ and Lemma 3.4 gives, by convexity of B,

$$\int_\Omega B(\beta(\overline{u})(\boldsymbol{x},T_0))\mathrm{d}\boldsymbol{x} \leq \liminf_{m\to\infty} \int_\Omega B\left(\frac{\beta(\Pi_{\mathcal{D}_m}u_m(\boldsymbol{x},T_m)) + \beta(\overline{u})(\boldsymbol{x},s_m)}{2}\right)\mathrm{d}\boldsymbol{x}. \quad (60)$$

Property (28) of B and the two inequalities (59) and (60) allow us to conclude the proof. Let $(s_m)_{m\in\mathbb{N}}$ be a sequence in $\mathcal{T}$ (see the proof of Corollary 3.8) which converges to T_0. Then $\nu(\overline{u}(\cdot,s_m)) \to \nu(\overline{u})(\cdot,T_0)$ in $L^2(\Omega)$ as $m \to \infty$. Using (28), we get

$$\begin{aligned}
&\|\nu(\Pi_{\mathcal{D}_m}u_m(\cdot,T_m)) - \nu(\overline{u})(\cdot,T_0)\|^2_{L^2(\Omega)} \\
&\quad\leq 2\|\nu(\Pi_{\mathcal{D}_m}u_m(\cdot,T_m)) - \nu(\overline{u}(\cdot,s_m))\|^2_{L^2(\Omega)} + 2\|\nu(\overline{u}(\cdot,s_m)) - \nu(\overline{u})(\cdot,T_0))\|^2_{L^2(\Omega)} \\
&\quad\leq 8L_\beta L_\zeta \int_\Omega [B(\beta(\Pi_{\mathcal{D}_m}u_m(\boldsymbol{x},T_m))) + B(\beta(\overline{u}(\boldsymbol{x},s_m)))]\,\mathrm{d}\boldsymbol{x} \\
&\qquad - 16L_\beta L_\zeta \int_\Omega B\left(\frac{\beta(\Pi_{\mathcal{D}_m}u_m(\boldsymbol{x},T_m)) + \beta(\overline{u}(\boldsymbol{x},s_m))}{2}\right)\mathrm{d}\boldsymbol{x} \\
&\qquad + 2\|\nu(\overline{u}(\cdot,s_m)) - \nu(\overline{u})(\cdot,T_0))\|^2_{L^2(\Omega)}.
\end{aligned}$$

We then take the lim sup as $m \to \infty$ of this expression. Thanks to (59) and the continuity of $t \in [0,T] \mapsto \int_\Omega B(\beta(\overline{u})(\boldsymbol{x},t))\mathrm{d}\boldsymbol{x} \in [0,\infty)$ (see Corollary 3.8), the first term in the right-hand side has a finite lim sup, bounded above by $16L_\beta L_\zeta \int_\Omega B(\beta(\overline{u})(\boldsymbol{x},T_0))\mathrm{d}\boldsymbol{x}$. We can therefore split the lim sup of this right-hand side without risking writing $\infty - \infty$ and we get, thanks to (60),

$$\limsup_{m\to\infty} \|\nu(\Pi_{\mathcal{D}_m} u_m(\cdot,T_m)) - \nu(\overline{u})(\cdot,T_0)\|^2_{L^2(\Omega)} \le 0.$$

Thus, $\nu(\Pi_{\mathcal{D}_m} u_m(\cdot,T_m)) \to \nu(\overline{u})(T_0)$ strongly in $L^2(\Omega)$. By Lemma 6.3 and the continuity of $\nu(\overline{u})$ stated in Corollary 3.8, this concludes the proof of the convergence of $\nu(\Pi_{\mathcal{D}_m} u_m)$ to $\nu(\overline{u})$ in $L^\infty(0,T;L^2(\Omega))$.

Remark 4.5 *Since $\beta(\Pi_{\mathcal{D}_m} u_m)(T_m) \to \beta(\bar{u})(T_0)$ weakly in $L^2(\Omega)$ as $m\to\infty$, Lemma 3.4 also shows that $\int_\Omega B(\beta(\overline{u})(\boldsymbol{x},T_0))\mathrm{d}\boldsymbol{x} \le \liminf_{m\to\infty}\int_\Omega B(\beta(\Pi_{\mathcal{D}_m} u_m)(\boldsymbol{x},T_m))\mathrm{d}\boldsymbol{x}$, and therefore, combined with* (59)*, that*

$$\lim_{m\to\infty}\int_\Omega B(\beta(\Pi_{\mathcal{D}_m} u_m(\boldsymbol{x},T_m)))\mathrm{d}\boldsymbol{x} = \int_\Omega B(\beta(\overline{u})(\boldsymbol{x},T_0))\mathrm{d}\boldsymbol{x}. \tag{61}$$

4.4 Proof of Theorem 2.16

Writing (39) for u_m with $T_0 = T$, taking the lim sup as $m\to\infty$, using (61) (with $T_m = T$) and the continuous integration-by-part formula (36) we find that

$$\limsup_{m\to\infty}\int_0^T\int_\Omega \boldsymbol{a}(\boldsymbol{x},\Pi_{\mathcal{D}_m}\nu(u_m)(\boldsymbol{x},t),\nabla_{\mathcal{D}_m}\zeta(u_m)(\boldsymbol{x},t))\cdot\nabla_{\mathcal{D}_m}\zeta(u_m)(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x}\mathrm{d}t$$
$$\le \int_0^{T_0}\int_\Omega \boldsymbol{a}(\boldsymbol{x},\nu(\overline{u})(\boldsymbol{x},t),\nabla\zeta(\overline{u})(\boldsymbol{x},t))\cdot\nabla\zeta(\overline{u})(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x}\mathrm{d}t.$$

Combined with (58), this shows that

$$\lim_{m\to\infty}\int_0^T\int_\Omega \boldsymbol{a}(\boldsymbol{x},\Pi_{\mathcal{D}_m}\nu(u_m)(\boldsymbol{x},t),\nabla_{\mathcal{D}_m}\zeta(u_m)(\boldsymbol{x},t))\cdot\nabla_{\mathcal{D}_m}\zeta(u_m)(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x}\mathrm{d}t$$
$$= \int_0^{T_0}\int_\Omega \boldsymbol{a}(\boldsymbol{x},\nu(\overline{u})(\boldsymbol{x},t),\nabla\zeta(\overline{u})(\boldsymbol{x},t))\cdot\nabla\zeta(\overline{u})(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x}\mathrm{d}t. \tag{62}$$

Let us define

$$f_m = [\boldsymbol{a}(\boldsymbol{x},\Pi_{\mathcal{D}_m}\nu(u_m),\nabla_{\mathcal{D}_m}\zeta(u_m)) - \boldsymbol{a}(\boldsymbol{x},\Pi_{\mathcal{D}_m}\nu(u_m)(\cdot,t),\nabla\zeta(\overline{u}))]\cdot[\nabla_{\mathcal{D}_m}\zeta(u_m) - \nabla\zeta(\overline{u})] \ge 0.$$

Developing this expression and using (62) and (18), we see that $\int_0^T\int_\Omega f_m(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x}\mathrm{d}t \to 0$ as $m\to\infty$. This shows that $f_m \to 0$ in $L^1(\Omega\times(0,T))$ and therefore a.e. up to a subsequence. We can then reason as in [16], using the strict monotony (19) of $\boldsymbol{a}$, the coercivity assumption (2f) and Vitali's theorem, to deduce that $\nabla_{\mathcal{D}_m}\zeta(u_m) \to \nabla\zeta(\overline{u})$ strongly in $L^p(\Omega\times(0,T))^d$ as $m\to\infty$.

5 About Assumption (17)

We discuss show here that, if $p\ge 2$, Theorem 2.10 can be proved without the structural assumption (17) – i.e. without assuming that $\beta = \mathrm{Id}$ or $\zeta = \mathrm{Id}$.
We first notice that we can always assume that

$$\beta + \zeta \text{ is strictly increasing.} \tag{63}$$

Indeed, if this is not the case then we have $s_1 < s_2$ such that $(\beta+\zeta)(s_1) = (\beta+\zeta)(s_2)$ and, since β and ζ are non-decreasing, that $[s_1, s_2]$ is a common plateau of β and ζ. Denoting by $\widetilde{\beta}$, $\widetilde{\zeta}$ and $\widetilde{\nu}$ the functions obtained from β, ζ and ν by removing this common plateau (by a contraction of the s-ordinate), we see that u is a solution to (1) if and only if u is a solution of the same problem with β, ζ and ν replaced with $\widetilde{\beta}$, $\widetilde{\zeta}$ and $\widetilde{\nu}$.

Remark 5.1 *Note that a rigorous and global way to remove all common plateaux of β and ζ at once is to actually consider, for $f = \beta$, ζ or ν, $\widetilde{f} = f \circ (\beta+\zeta) \circ (\beta+\zeta)_r$, where $(\beta+\zeta)_r$ is the pseudo-inverse of $\alpha+\beta$ constructed as in* (3)*.*

Once this reduction has been applied, we can state the following theorem.

Theorem 5.2 *Under Assumptions* (2)*, let $(\mathcal{D}_m)_{mp\in\mathbb{N}}$ be a sequence of space-time gradient discretisations in the sense of Definition 2.1, which is coercive, consistent, compact and limit-conforming (see Section 2.2). Let, for any $m \in \mathbb{N}$, u_m be the solution to* (14) *with $\mathcal{D} = \mathcal{D}_m$, provided by Theorem 2.10.*
If we assume that Assumption (63) *hold and that $p \geq 2$, then there exists a solution $\overline{u}$ to* (4) *such that, up to a subsequence:*

- *the convergences stated in* (18) *hold,*
- $\Pi_{\mathcal{D}_m}\nu(u_m) \to \nu(\overline{u})$ *strongly in* $L^\infty(0,T;L^2(\Omega))$ *as* $m \to \infty$,
- *under the strict monotony of* $\boldsymbol{a}$ *(i.e. Assumption* (19)*), as $m \to \infty$, $\Pi_{\mathcal{D}_m}\zeta(u_m) \to \zeta(\overline{u})$ strongly in $L^p(\Omega\times(0,T))$ and $\nabla_{\mathcal{D}_m}\zeta(u_m) \to \nabla\zeta(\overline{u})$ strongly in $L^p(\Omega\times(0,T))^d$.*

Proof.
We only need to prove the first conclusion of the theorem, i.e. that the convergences (18) hold. Indeed, Theorems 2.14 and 2.16 then provide the last two conclusions.
The only difference in the assumptions of Theorem 5.2 and 2.10 is the absence, here, of the structural assumption (17) and the assumption that $\beta+\zeta$ is strictly increasing. The only place in the proof of Theorem (2.10) where we used Assumption (17) is in Step 1, to identify the limits $\overline{\beta}$, $\overline{\zeta}$ and $\overline{\nu}$ of $\Pi_{\mathcal{D}_m}\beta(u_m)$, $\Pi_{\mathcal{D}_m}\zeta(u_m)$ and $\Pi_{\mathcal{D}_m}\nu(u_m)$. We will show that without assuming (17) but by assuming (63), we can still identify those limits.
We first work with the first two limits. Lemmas 4.1 and 4.3 show that $\beta_m = \beta(u_m)$ and $\zeta_m = \zeta(u_m)$ satisfy the assumptions of the discrete compensated compactness theorem 5.4 below (we also use that $p \geq 2$ here). Hence, $\Pi_{\mathcal{D}_m}\beta(u_m)\Pi_{\mathcal{D}_m}\zeta(u_m) \to \overline{\beta}\,\overline{\zeta}$ in the sense of measures on $\Omega\times(0,T)$. We can then apply Lemma 5.6 with $\varphi \equiv 1$ and $w_m = \Pi_{\mathcal{D}_m}u_m$. This gives a measurable $\overline{u}$ such that $\overline{\beta} = \beta(\overline{u})$ and $\overline{\zeta} = \zeta(\overline{u})$ a.e. on $\Omega\times(0,T)$.
We now turn to $\overline{\nu}$. Estimates (50) and (43) and Kolmogorov's compactness theorem show that the convergence of $\Pi_{\mathcal{D}_m}\nu(u_m)$ towards $\overline{\nu}$ is actually strong on $L^2(\Omega\times(0,T))$ (we use $p \geq 2$ here). Setting $\mu = \beta+\zeta$, we saw that $\mu(\Pi_{\mathcal{D}_m}u_m) \to \overline{\mu} := \beta(\overline{u}) + \zeta(\overline{u})$ weakly in $L^2(\Omega\times(0,T))$. We can therefore apply Lemma 5.6 with $\varphi \equiv 1$ and μ, ν instead of β, ζ, and we find a measurable w such that $\overline{\nu} = \nu(w)$ and $\overline{\mu} = \mu(w)$. The second relation translates into $(\beta+\zeta)(\overline{u}) = (\beta+\zeta)(w)$, that is $w = \overline{u}$ since $\beta+\zeta$ is strictly increasing. Hence, $\overline{\nu} = \nu(\overline{u})$.
To summarise, the limits of $\Pi_{\mathcal{D}_m}\beta(u_m)$, $\Pi_{\mathcal{D}_m}\zeta(u_m)$ and $\Pi_{\mathcal{D}_m}\nu(u_m)$ have been identified as $\beta(\overline{u})$, $\zeta(\overline{u})$ and $\nu(\overline{u})$ for some $\overline{u}$. This allows to take over the proof of Theorem 2.10 from after the usage of (17) and conclude that $\overline{u}$ is a solution to (4) and that the convergences (18) hold. The last two conclusions of the theorem follow from Theorems 2.14 and 2.16. ■

Remark 5.3 *Note that it is not proved, in this context, that $\overline{u}$ is a weak limit of $\Pi_{\mathcal{D}_m} u_m$, but such a limit is not stated in* (18) *and can actually be considered as irrelevant for the model* (1)*, in which the quantities of interest (physically relevant when this PDE models a natural phenomenon) are $\beta(\overline{u})$, $\zeta(\overline{u})$ and $\nu(\overline{u})$.*

We now state the two key results that allowed us to replace Assumption (17) by Assumption (17). The first is a discrete version of a compensated compactness result in [31]. The second is a Minty-like result, useful to identify weak non-linear limits.
We note that Theorem 5.4 is a more general version than the one needed in the proof of Theorem 5.2 (which only requires $\varphi \equiv 1$), but we state this more general version nevertheless as it is the genuine discrete equivalent of the compensated compactness result in [31].

Theorem 5.4 (Discrete compensated compactness) *We take $T > 0$, $p \geq 2$ and a sequence $(\mathcal{D}_m)_{m\in\mathbb{N}} = (X_{\mathcal{D}_m,0}, \Pi_{\mathcal{D}_m}, \nabla_{\mathcal{D}_m}, \mathcal{I}_{\mathcal{D}_m}, (t_m^{(n)})_{n=0,\ldots,N_m})_{m\in\mathbb{N}}$ of space-time discretisation in the sense of Definition 2.1 which is consistent, limit-conforming and compact in the sense of Definitions 2.6, 2.8 and 2.7.*
For any $m \in \mathbb{N}$, let $\beta_m = (\beta_m^{(n)})_{n=0,\ldots,N_m} \subset X_{\mathcal{D}_m,0}$ and $\zeta_m = (\zeta_m^{(n)})_{n=0,\ldots,N_m} \subset X_{\mathcal{D}_m,0}$ such that

- *The sequences $(\int_0^T |\delta_m \beta_m(t)|_{\star,\mathcal{D}_m})_{m\in\mathbb{N}}$ and $(||\nabla_{\mathcal{D}_m}\zeta_m||_{L^2(0,T;L^p(\Omega)^d)})_{m\in\mathbb{N}}$ are bounded,*
- *As $m \to \infty$, $\Pi_{\mathcal{D}_m}\beta_m \to \overline{\beta}$ and $\Pi_{\mathcal{D}_m}\zeta_m \to \overline{\zeta}$ weakly in $L^2(\Omega \times (0,T))$.*

Then $\Pi_{\mathcal{D}_m}\beta_m \Pi_{\mathcal{D}_m}\zeta_m \to \overline{\beta}\,\overline{\zeta}$ in the sense of measures on $\Omega\times(0,T)$, that is, for all $\varphi \in C(\overline{\Omega}\times[0,T])$,

$$\lim_{m\to\infty} \int_0^T \int_\Omega \Pi_{\mathcal{D}_m}\beta_m(\boldsymbol{x},t)\Pi_{\mathcal{D}_m}\zeta_m(\boldsymbol{x},t)\varphi(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x}\mathrm{d}t = \int_0^T \int_\Omega \overline{\beta}(\boldsymbol{x},t)\,\overline{\zeta}(\boldsymbol{x},t)\varphi(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x}\mathrm{d}t. \quad (64)$$

Proof.
The idea of the proof is to reduce to the case where $\Pi_{\mathcal{D}_m}\zeta_m$ are tensorial functions, in order to separate the space and time variables and make use of the separate compactness of $\Pi_{\mathcal{D}_m}\zeta_m$ and $\Pi_{\mathcal{D}_m}\beta_m$ with respect to each of these variables. Note that this proof also provides an apparently new proof for the continuous equivalent of this compensated compactness result.

Step 1: reduction of $\Pi_{\mathcal{D}_m}\zeta_m$ to tensorial functions.
Let us take $\delta > 0$ and consider a covering $(A_k^\delta)_{k=1,\ldots,K}$ of Ω in disjoint cubes of length δ. Let us denote $R_\delta : L^2(\Omega) \to L^2(\Omega)$ the operator defined by:

$$\forall k = 1,\ldots,K\,,\ \forall \boldsymbol{x} \in A_k^\delta\ :\ R_\delta f(\boldsymbol{x}) = \frac{1}{\mathrm{meas}(A_k^\delta)} \int_{A_\delta^k} f(\boldsymbol{y})\mathrm{d}\boldsymbol{y},$$

where f has been extended by 0 outside Ω. Let $\boldsymbol{x} \in A_k^\delta$. Using Jensen's inequality, fact that $\mathrm{meas}(A_k^\delta) = \delta^d$ and the change of variable $\boldsymbol{y} \in A_k^\delta \mapsto \boldsymbol{\xi} = \boldsymbol{y} - \boldsymbol{x} \in (-\delta,\delta)^d$, we can write

$$|R_\delta f(\boldsymbol{x}) - f(\boldsymbol{x})|^2 \leq \delta^{-d} \int_{A_k^\delta} |f(\boldsymbol{y}) - f(\boldsymbol{x})|^2 \mathrm{d}\boldsymbol{y} \leq \delta^{-d} \int_{(-\delta,\delta)^d} |f(\boldsymbol{x}+\boldsymbol{\xi}) - f(\boldsymbol{x})|^2 \mathrm{d}\boldsymbol{\xi}.$$

Integrating over $\boldsymbol{x} \in A_k^\delta$ and summing over $k = 1,\ldots,K$ then gives

$$\begin{aligned} ||R_\delta f - f||^2_{L^2(\Omega)} &\leq \delta^{-d} \int_{(-\delta,\delta)^d} \int_\Omega |f(\boldsymbol{x}+\boldsymbol{\xi}) - f(\boldsymbol{x})|^2 \mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{\xi} \\ &\leq 2^d \sup_{\boldsymbol{\xi}\in(-\delta,\delta)^d} \int_\Omega |f(\boldsymbol{x}+\boldsymbol{\xi}) - f(\boldsymbol{x})|^2 \mathrm{d}\boldsymbol{x}. \end{aligned} \quad (65)$$

On the other side, the compactness of $(\mathcal{D}_m)_{m\in\mathbb{N}}$ (Definition 2.7) and the fact that $p \geq 2$ give $\epsilon(\boldsymbol{\xi})$ such that $\epsilon(\boldsymbol{\xi}) \to 0$ as $\boldsymbol{\xi} \to 0$ and, for all $m \in \mathbb{N}$ and all $v \in X_{\mathcal{D}_m,0}$,

$$||\Pi_{\mathcal{D}_m} v(\cdot + \boldsymbol{\xi}) - \Pi_{\mathcal{D}_m} v||^2_{L^2(\Omega)} \leq \epsilon(\boldsymbol{\xi})||\nabla_{\mathcal{D}_m} v||^2_{L^p(\Omega)^d}.$$

Combining this with (65) and using the bound on $||\nabla_{\mathcal{D}_m}\zeta_m||_{L^2(0,T;L^p(\Omega)^d)}$ shows that

$$||R_\delta \Pi_{\mathcal{D}_m}\zeta_m - \Pi_{\mathcal{D}_m}\zeta_m||_{L^2(\Omega\times(0,T))} \leq C \sup_{|\boldsymbol{\xi}|\leq\delta} \sqrt{\epsilon(\boldsymbol{\xi})} =: \omega(\delta) \quad (66)$$

where C does not depend on m and $\omega(\delta) \to 0$ as $\delta \to 0$. Note that a similar estimate holds with $\Pi_{\mathcal{D}_m}\zeta_m$ replaced with $\overline{\zeta}$ since $\overline{\zeta} \in L^2(\Omega \times (0,T))$ and, therefore, its translates tend to 0 in $L^2(\Omega \times (0,T))$.
If we respectively denote by $A_m(\Pi_{\mathcal{D}_m}\zeta_m)$ and $A(\overline{\zeta})$ the left-hand side and right-hand side (64), then since $(\Pi_{\mathcal{D}_m}\beta_m)_{m\in\mathbb{N}}$ is bounded in $L^2(\Omega \times (0,T))$ we have, by (66),

$$|A_m(\Pi_{\mathcal{D}_m}\zeta_m) - A(\overline{\zeta})| \leq C\omega(\delta) + |A_m(R_\delta\Pi_{\mathcal{D}_m}\zeta_m) - A(R_\delta\overline{\zeta})|.$$

Let us assume that we can prove that, for a fixed δ,

$$A_m(R_\delta\Pi_{\mathcal{D}_m}\zeta_m) \to A(R_\delta\overline{\zeta}) \text{ as } m \to \infty. \quad (67)$$

Then the previous inequality gives $\limsup_{m\to\infty} |A_m(\Pi_{\mathcal{D}_m}\zeta_m) - A(\overline{\zeta})| \leq C\omega(\delta)$. Letting $\delta \to 0$ in this inequality gives $A_m(\Pi_{\mathcal{D}_m}\zeta_m) \to A(\overline{\zeta})$ as wanted.
Hence, we only need to prove 67. The definition of R_δ shows that $R_\delta f = \sum_{k=1}^K \mathbf{1}_{A_k^\delta}\text{meas}(A_k^\delta)^{-1}[f]_{A_k^\delta}$ where $\mathbf{1}_{A_k^\delta}$ is the characteristic function of A_k^δ and $[f]_A = \int_A f(\boldsymbol{x})\mathrm{d}\boldsymbol{x}$. Hence, (67) will follow if we can prove that, for any measurable set A of non-zero measure,

$$\lim_{m\to\infty}\int_0^T\int_\Omega \Pi_{\mathcal{D}_m}\beta_m(\boldsymbol{x},t)[\Pi_{\mathcal{D}_m}\zeta_m]_A(t)\varphi(t,\boldsymbol{x})\mathbf{1}_A(\boldsymbol{x})\mathrm{d}\boldsymbol{x}\mathrm{d}t$$
$$= \int_0^T\int_\Omega \overline{\beta}(\boldsymbol{x},t)[\,\overline{\zeta}\,]_A(t)\varphi(t,\boldsymbol{x})\mathbf{1}_A(\boldsymbol{x})\mathrm{d}\boldsymbol{x}\mathrm{d}t \quad (68)$$

where, for $g \in L^2(\Omega \times (0,T))$, we set $[g]_A(t) = \int_A g(t,\boldsymbol{y})\mathrm{d}\boldsymbol{y}$.

Step 2: further reductions.
We now reduce φ to a tensorial function and $\mathbf{1}_A$ to a smooth function. It is well-known that there exists tensorial functions $\varphi_r = \sum_{l=1}^{L_r}\theta_{l,r}(t)\gamma_{l,t}(\boldsymbol{x})$, with $\theta_{l,r} \in C^\infty([0,T])$ and $\gamma_{l,r} \in C^\infty(\overline{\Omega})$, such that $\varphi_r \to \varphi$ uniformly on $\Omega \times (0,T)$ as $r \to \infty$. Moreover, there exists $\rho_r \in C_c^\infty(\Omega)$ such that $\rho_r \to \mathbf{1}_A$ in $L^2(\Omega)$ as $r \to \infty$.
Hence, as $r \to \infty$ the function $(t,\boldsymbol{x}) \mapsto \varphi_r(t,\boldsymbol{x})\rho_r(\boldsymbol{x})$ converges in $L^\infty(0,T;L^2(\Omega))$ to the function $(t,\boldsymbol{x}) \mapsto \varphi(t,\boldsymbol{x})\mathbf{1}_A(\boldsymbol{x})$. Since the sequence of functions $(t,\boldsymbol{x}) \mapsto \Pi_{\mathcal{D}_m}\beta(t,\boldsymbol{x})[\Pi_{\mathcal{D}_m}\zeta_m]_A(t)$ is bounded in $L^1(0,T;L^2(\Omega))$ (notice that $([\Pi_{\mathcal{D}_m}\zeta_m]_A)_{m\in\mathbb{N}}$ is bounded in $L^2(0,T)$ since $(\Pi_{\mathcal{D}_m}\zeta_m)_{m\in\mathbb{N}}$ is bounded in $L^2(\Omega \times (0,T))$), a reasoning similar as the one used in Step 1 shows that we only need to prove (68) with $\varphi(t,\boldsymbol{x})\mathbf{1}_A(\boldsymbol{x})$ replaced with $\varphi_r(t,\boldsymbol{x})\rho_r(\boldsymbol{x})$ for a fixed r.
We have $\varphi_r(t,\boldsymbol{x})\rho_r(\boldsymbol{x}) = \sum_{l=1}^{L_r}\theta_{l,r}(t)(\gamma_{l,t}\rho)(\boldsymbol{x})$ and $\gamma_{l,r}\rho_r \in C_c^\infty(\Omega)$. Hence, (68) with $\varphi(t,\boldsymbol{x})\mathbf{1}_A(\boldsymbol{x})$ replaced with $\varphi_r(t,\boldsymbol{x})\rho_r(\boldsymbol{x})$ will follow if we can establish that: for any $\theta \in C^\infty([0,T])$, any $\psi \in C_c^\infty(\Omega)$ and any measurable set $A \subset \Omega$,

$$\lim_{m\to\infty}\int_0^T\int_\Omega \theta(t)\Pi_{\mathcal{D}_m}\beta_m(\boldsymbol{x},t)[\Pi_{\mathcal{D}_m}\zeta_m]_A(t)\psi(\boldsymbol{x})\mathrm{d}\boldsymbol{x}\mathrm{d}t = \int_0^T\int_\Omega \theta(t)\overline{\beta}(\boldsymbol{x},t)[\,\overline{\zeta}\,]_A(t)\psi(\boldsymbol{x})\mathrm{d}\boldsymbol{x}\mathrm{d}t. \quad (69)$$

Step 3: proof of (69).
We now use the estimates on $\delta_m\beta_m$ to conclude. We first write

$$\int_0^T\int_\Omega \theta(t)\Pi_{\mathcal{D}_m}\beta_m(\boldsymbol{x},t)[\Pi_{\mathcal{D}_m}\zeta_m]_A(t)\psi(\boldsymbol{x})\mathrm{d}\boldsymbol{x}\mathrm{d}t = \int_0^T \theta(t)[\Pi_{\mathcal{D}_m}\zeta_m]_A(t)F_m(t). \quad (70)$$

with $F_m(t) = \int_\Omega \Pi_{\mathcal{D}_m}\beta_m(\boldsymbol{x},t)\psi(\boldsymbol{x})\mathrm{d}\boldsymbol{x}$. It is clear from the weak convergence of $\Pi_{\mathcal{D}_m}\zeta_m$ that $[\Pi_{\mathcal{D}_m}\zeta_m]_A \to [\overline{\zeta}]_A$ weakly in $L^2(0,T)$. Hence, if we can prove that $F_m \to F := \int_\Omega \overline{\beta}(\boldsymbol{x},\cdot)\psi(\boldsymbol{x})\mathrm{d}\boldsymbol{x}$ strongly in $L^2(0,T)$, we can pass to the limit in (70) and obtain (69). Since F_m weakly converges to F in $L^2(0,T)$ (thanks to the weak convergence in $L^2(\Omega\times(0,T))$ of $\Pi_{\mathcal{D}_m}\beta_m$), we only have to prove that $(F_m)_{m\in\mathbb{N}}$ is relatively compact in $L^2(0,T)$.
To prove the strong convergence of F_m, we introduce the interpolant $P_{\mathcal{D}_m}$ defined by (21) and we define G_m as F_m with ψ replaced with $\Pi_{\mathcal{D}_m}P_{\mathcal{D}_m}\psi$. We then have

$$|F_m(t) - G_m(t)| \le ||\Pi_{\mathcal{D}_m}\beta_m(\cdot,t)||_{L^2(\Omega)}S_{\mathcal{D}_m}(\psi).$$

The consistency of $(\mathcal{D}_m)$ then shows that

$$F_m - G_m \to 0 \text{ strongly in } L^2(\Omega) \text{ as } m\to\infty. \quad (71)$$

We now study the strong convergence of G_m. This function is, as $\Pi_{\mathcal{D}_m}\beta_m$, piecewise constant on $(0,T)$ and, by definition of $|\cdot|_{\star,\mathcal{D}_m}$, its discrete derivative satisfies

$$|\delta_m G_m(t)| \le |\delta_m\beta_m(t)|_{\star,\mathcal{D}_m}||P_{\mathcal{D}_m}\psi||_{X_{\mathcal{D}_m,0}}.$$

Since $||P_{\mathcal{D}_m}\psi||_{X_{\mathcal{D}_m,0}} \le S_{\mathcal{D}_m}(\psi) + ||\nabla\psi||_{L^p(\Omega)^d}$ is bounded uniformly with respect to m, the assumption on $\delta_m\beta_m$ proves that $(||\delta_m G_m||_{L^1(0,T)})_{m\in\mathbb{N}}$ is bounded. Hence, G_m is bounded in $BV(0,T)\cap L^1(0,T)$ and therefore relatively compact in $L^2(0,T)$ (see [3, Theorem 10.1.4]). Combined with (71), this proves that $(F_m)_{m\in\mathbb{N}}$ is relatively compact in $L^2(0,T)$ and concludes the proof. ■

Remark 5.5 *If we assume that $(\Pi_{\mathcal{D}_m}\beta_m)_{m\in\mathbb{N}}$ is bounded in $L^\infty(0,T;L^2(\Omega))$ and that, for some $q>1$, $(\int_0^T |\delta_m\beta_m(t)|^q_{\star,\mathcal{D}_m})_{m\in\mathbb{N}}$ is bounded, then Step 2 becomes a trivial consequence of Theorem 3.1. Indeed, this theorem then shows that $(\Pi_{\mathcal{D}_m}\beta_m)_{m\in\mathbb{N}}$ is relatively compact uniformly-in-time and weakly in $L^2(\Omega)$, which translates into the relative compactness of $(F_m)_{m\in\mathbb{N}}$ in $L^\infty(0,T)$.*

Lemma 5.6 *Let $N\in\mathbb{N}^\star$ and V be a non-empty measurable subset of $\mathbb{R}^N$. Let $\beta,\zeta\in C^0(\mathbb{R})$ be two nondecreasing functions, such that $\beta(0)=\zeta(0)=0$ and $\beta+\zeta$ is strictly increasing. We assume that there exists a sequence $(w_m)_{m\in\mathbb{N}}$ of measurable functions on V and two functions $\overline{\beta},\overline{\zeta}\in L^2(V)$ such that:*

- *$\beta(w_m)\to\overline{\beta}$ and $\zeta(w_m)\to\overline{\zeta}$ weakly in $L^2(V)$,*
- *there exists $\varphi\in L^\infty(V)$ such that $\varphi>0$ a.e. on V and*

$$\lim_{m\to\infty}\int_V \varphi(\boldsymbol{z})\beta(w_m(\boldsymbol{z}))\zeta(w_m(\boldsymbol{z}))\mathrm{d}\boldsymbol{z} = \int_V \varphi(\boldsymbol{z})\overline{\beta}(\boldsymbol{z})\,\overline{\zeta}(\boldsymbol{z})\mathrm{d}\boldsymbol{z}. \quad (72)$$

Then

$$\overline{\beta}=\beta(w) \text{ and } \overline{\zeta}=\zeta(w) \text{ a.e. in } V, \text{ where } w = \left(\frac{\beta+\zeta}{2}\right)^{-1}\left(\frac{\overline{\beta}+\overline{\zeta}}{2}\right). \quad (73)$$

Proof. We first notice that $\beta(w), \zeta(w) \in L^2(V)$, given that $\frac{\overline{\beta}+\overline{\zeta}}{2} \in L^2(V)$ and $\beta \circ (\frac{\beta+\zeta}{2})^{-1}$ and $\zeta \circ (\frac{\beta+\zeta}{2})^{-1}$ are sub-linear (the sum of both is 2Id and each one has the same sign as s). Since β and ζ are non-decreasing, we can therefore write

$$\int_V \varphi(\boldsymbol{z}) \left[\beta(w_m(\boldsymbol{z})) - \beta(w(\boldsymbol{z}))\right] \left[\zeta(w_m(\boldsymbol{z})) - \zeta(w(\boldsymbol{z}))\right] \mathrm{d}\boldsymbol{z} \geq 0.$$

Letting $m \to \infty$ in the above inequality, and using the convergences of $\beta(w_m)$, $\zeta(w_m)$ and (72), we obtain

$$\int_V \varphi(\boldsymbol{z}) \left[\overline{\beta}(\boldsymbol{z}) - \beta(w(\boldsymbol{z}))\right] \left[\overline{\zeta}(\boldsymbol{z}) - \zeta(w(\boldsymbol{z}))\right] \mathrm{d}\boldsymbol{z} \geq 0. \tag{74}$$

We then remark that

$$\frac{\overline{\beta}+\overline{\zeta}}{2} = \frac{\beta(w)+\zeta(w)}{2}, \tag{75}$$

which gives $\beta(w) = \frac{\overline{\beta}+\overline{\zeta}}{2} + \left(\frac{\beta-\zeta}{2}\right)(w)$ and $\zeta(w) = \frac{\overline{\beta}+\overline{\zeta}}{2} - \left(\frac{\beta-\zeta}{2}\right)(w)$. Hence, (74) leads to

$$-\int_V \varphi(\boldsymbol{z}) \left(\frac{\overline{\beta}-\overline{\zeta}}{2}(\boldsymbol{z}) - \left(\frac{\beta-\zeta}{2}\right)(w(\boldsymbol{z}))\right)^2 \mathrm{d}\boldsymbol{z} \geq 0.$$

Since φ is almost everywhere strictly positive on V, we deduce that $\frac{\overline{\beta}-\overline{\zeta}}{2} = \frac{\beta(w)-\zeta(w)}{2}$ a.e. in V, and (73) follows from (75). ∎

6 Appendix: uniform-in-time compactness results for time-dependent problems

We establish in this appendix some generic results, unrelated to the framework of Gradient Schemes, that form the starting point for our uniform-in-time convergence results.

Solutions of numerical schemes for parabolic equations are usually piecewise constant, and therefore not continous, in time. As their jump nevertheless tend to become small with the time step, it is possible to establish some uniform-in-time convergence results using a generalisation to non-continuous functions of the classical Ascoli-Arzela theorem.

Definition 6.1 *If (K, d_K) and (E, d_E) are metric spaces, we denote by $\mathcal{F}(K, E)$ the space of functions $K \to E$, endowed with the uniform metric $d_{\mathcal{F}}(v, w) = \sup_{s\in K} d_E(v(s), w(s))$ (note that this metric may take infinite values).*

Theorem 6.2 (Generalised Ascoli-Arzela's theorem) *Let (K, d_K) be a compact metric space, (E, d_E) be a complete metric space and $(\mathcal{F}(K, E), d_{\mathcal{F}})$ as in Definition 6.1.*
Let $(v_m)_{m\in\mathbb{N}}$ be a sequence in $\mathcal{F}(K, E)$ such that there exists a function $\omega : K \times K \to [0, \infty]$ and a sequence $(\delta_m)_{m\in\mathbb{N}} \subset [0, \infty)$ satisfying

$$\lim_{d_K(s,s')\to 0} \omega(s, s') = 0\,, \quad \lim_{m\to\infty} \delta_m = 0\,,$$
$$\forall (s, s') \in K^2\,,\ \forall m \in \mathbb{N}\,,\ d_E(v_m(s), v_m(s')) \leq \omega(s, s') + \delta_m. \tag{76}$$

We also assume that, for all $s \in K$, $\{v_m(s) : m \in \mathbb{N}\}$ is relatively compact in (E, d_E).
Then $(v_m)_{m\in\mathbb{N}}$ is relatively compact in $(\mathcal{F}(K, E), d_{\mathcal{F}})$ and any adherence value of $(v_m)_{m\in\mathbb{N}}$ in this space is continuous $K \to E$.

Proof. Let us first notice that the last conclusion of the theorem, i.e. that any adherence value v of $(v_m)_{m\in\mathbb{N}}$ in $\mathcal{F}(K,E)$ is continuous, is trivially obtained by passing to the limit along this subsequence in (76), showing that the modulus of continuity of v is bounded above by ω.
The proof of the compactness result an easy generalisation of the proof of the classical Ascoli-Arzela compactness result. We start by taking a countable dense subset $\{s_l : l \in \mathbb{N}\}$ in K (the existence of this set is ensured since K is compact metric). Since each set $\{v_m(s_l) : m \in \mathbb{N}\}$ is relatively compact in E, by diagonal extraction we can select a subsequence of $(v_m)_{m\in\mathbb{N}}$, denoted the same way, such that for any $l \in \mathbb{N}$, $(v_m(s_l))_{m\in\mathbb{N}}$ converges in E. We then proceed in showing that $(v_m)_{m\in\mathbb{N}}$ is a Cauchy sequence in $(\mathcal{F}(K,E), d_{\mathcal{F}})$. Since this space is complete, this will show that this sequence converges in this space and will therefore complete the proof.
Let $\varepsilon > 0$ and, using (76), take $\delta > 0$ and $M \in \mathbb{N}$ such that $\omega(s,s') \le \varepsilon$ whenever $d_K(s,s') \le \delta$ and $\delta_m \le \varepsilon$ whenever $m \ge M$. Select a finite set $\{s_{l_1},\ldots,s_{l_N}\}$ such that any $s \in K$ is within distance δ of a s_{l_i}. Then, for any $m,m' \ge M$, by (76),

$$\begin{aligned} d_E(v_m(s),v_{m'}(s)) &\le d_E(v_m(s),v_m(s_{l_i})) + d_E(v_m(s_{l_i}),v_{m'}(s_{l_i})) \\ &\quad + d_E(v_{m'}(s_{l_i}),v_{m'}(s)) \\ &\le \omega(s,s_{l_i}) + \delta_m + d_E(v_m(s_{l_i}),v_{m'}(s_{l_i})) + \omega(s,s_{l_i}) + \delta_{m'} \\ &\le 4\varepsilon + d_E(v_m(s_{l_i}),v_{m'}(s_{l_i})). \end{aligned}$$

Since $\{(v_m(s_{l_i}))_{m\in\mathbb{N}} : i = 1,\ldots,N\}$ form a finite number of converging sequences in E, we can find $M' \ge M$ such that, whenever $m,m' \ge M'$ and $i = 1,\ldots,N$, $d_E(v_m(s_{l_i}),v_{m'}(s_{l_i})) \le \varepsilon$. This shows that, for all $m,m' \ge M'$ and all $s \in K$, $d_E(v_m(s),v_{m'}(s)) \le 5\varepsilon$ and concludes the proof that $(v_m)_{m\in\mathbb{N}}$ is a Cauchy sequence in $(\mathcal{F}(K,E), d_{\mathcal{F}})$. ∎

The following lemma states an equivalent condition for the uniform convergence of functions, which proves extremely useful to establish uniform-in-time convergence of numerical schemes for parabolic equations when no smoothness is assumed on the data.

Lemma 6.3 *Let (K,d_K) be a compact metric space, (E,d_E) be a metric space and $(\mathcal{F}(K,E),d_{\mathcal{F}})$ as in Definition 6.1. Let $(v_m)_{m\in\mathbb{N}}$ be a sequence in $\mathcal{F}(K,E)$ and $v : K \mapsto E$ be continuous.*
Then $v_m \to v$ for $d_{\mathcal{F}}$ if and only if, for any $s \in K$ and any sequence $(s_m)_{m\in\mathbb{N}} \subset K$ converging to s for d_K, we have $v_m(s_m) \to v(s)$ for d_E.

Proof. If $v_m \to v$ for $d_{\mathcal{F}}$ then for any sequence $(s_m)_{m\in\mathbb{N}}$ converging to s,

$$d_E(v_m(s_m),v(s)) \le d_E(v_m(s_m),v(s_m)) + d_E(v(s_m),v(s)) \le d_{\mathcal{F}}(v_m,v) + d(v(s_m),v(s))$$

and the right-hand side tends to 0 by definition of $v_m \to v$ for $d_{\mathcal{F}}$ and by continuity of v.
Let us now prove the converse by contraposition. If $(v_m)_{m\in\mathbb{N}}$ does not converge to v for $d_{\mathcal{F}}$ then there exists $\varepsilon > 0$ and a subsequence $(v_{m_k})_{k\in\mathbb{N}}$, such that, for any $k \in \mathbb{N}$, $\sup_{s\in K} d_E(v_{m_k}(s),v(s)) \ge \varepsilon$. We can then find a sequence $(r_k)_{k\in\mathbb{N}} \subset K$ such that, for any $k \in \mathbb{N}$,

$$d_E(v_{m_k}(r_k),v(r_k)) \ge \varepsilon/2. \tag{77}$$

K being compact, up to another subsequence denoted the same way, we can assume that $r_k \to s$ in K as $k \to \infty$. It is then trivial to construct a sequence $(s_m)_{m\in\mathbb{N}}$ converging to s and such that $s_{m_k} = r_k$ (just take $s_m = s$ when m is not an m_k). We then have $v_m(s_m) \to v(s)$ in E and, by continuity of v, $v(s_m) \to v(s)$ in E. This shows that $d_E(v_m(s_m),v(s_m)) \to 0$, which contradicts (77) and concludes the proof. ∎

Uniform-in-time convergence of numerical solutions to schemes for parabolic equations starts with a weak convergence with respect to the time variable. This weak convergence is then used to prove a stronger convergence. The following definition and proposition recall standard notions related to the weak topology on $L^2(\Omega)$.

Definition 6.4 (Uniform-in-time $L^2(\Omega)$-weak convergence) *A sequence of functions $u_m : [0,T] \to L^2(\Omega)$ converges weakly in $L^2(\Omega)$ uniformly on $[0,T]$ to a function $u : [0,T] \to L^2(\Omega)$ if, for all $\varphi \in L^2(\Omega)$, as $m \to \infty$ the sequence of functions $t \in [0,T] \to \langle u_m(t), \varphi\rangle_{L^2(\Omega)}$ converges uniformly on $[0,T]$ to $t \in [0,T] \to \langle u(t), \varphi\rangle_{L^2(\Omega)}$, where $\langle\cdot,\cdot\rangle_{L^2(\Omega)}$ is the inner product in $L^2(\Omega)$.*

The next result is classical, but its short proof is recalled for the reader's convenience.

Proposition 6.5 *Let E be closed bounded ball in $L^2(\Omega)$ and $(\varphi_l)_{l\in\mathbb{N}}$ be a dense sequence in $L^2(\Omega)$. Then, on E, the weak topology of $L^2(\Omega)$ is the topology given by the metric*

$$d_E(v,w) = \sum_{l\in\mathbb{N}} \frac{\min(1, |\langle v-w, \varphi_l\rangle_{L^2(\Omega)}|)}{2^l}. \tag{78}$$

Moreover, a sequence of functions $u_m : [0,T] \to E$ converges uniformly to $u : [0,T] \to E$ for the weak topology of $L^2(\Omega)$ if and only if, as $m \to \infty$, $d_E(u_m, u) : [0,T] \to [0,\infty)$ converges uniformly to 0.

Proof. The sets $E_{\varphi,\varepsilon} = \{v \in E : |\langle v,\varphi\rangle_{L^2(\Omega)}| < \varepsilon\}$, for $\varphi \in L^2(\Omega)$ and $\varepsilon > 0$, define a basis of neighborhood of 0 for the weak $L^2(\Omega)$ topology on E, and a basis of neighborhood of any other points is obtained by translation of this particular basis. If R is the radius of the ball E then for any $\varphi \in L^2(\Omega)$, $l \in \mathbb{N}$ and $v \in E$ we have

$$|\langle v,\varphi\rangle_{L^2(\Omega)}| \le R||\varphi - \varphi_l||_{L^2(\Omega)} + |\langle v,\varphi_l\rangle_{L^2(\Omega)}|.$$

By density of $(\varphi_l)_{l\in\mathbb{N}}$ we can select $l \in \mathbb{N}$ such that $||\varphi - \varphi_l||_{L^2(\Omega)} < \varepsilon/(2R)$ and we then see that $E_{\varphi_l,\varepsilon/2} \subset E_{\varphi,\varepsilon}$. Hence, a basis of neighborhood of 0 in E for the weak $L^2(\Omega)$ is also given by $(E_{\varphi_l,\varepsilon})_{l\in\mathbb{N},\,\varepsilon>0}$.
From the definition of d_E we see that, for any $l \in \mathbb{N}$, $\min(1, |\langle v,\varphi_l\rangle_{L^2(\Omega)}|) \le 2^l d_E(0,v)$. If $d_E(0,v) < 2^{-l}$ this shows that $|\langle v,\varphi_l\rangle_{L^2(\Omega)}| \le 2^l d_E(0,v)$ and therefore that

$$B_{d_E}(0, \min(2^{-l+1}, \varepsilon 2^{-l})) \subset E_{\varphi_l,\varepsilon}.$$

Hence, any neighborhood of 0 in E for the $L^2(\Omega)$ weak topology is a neighborhood of 0 for d_E. Conversely, for any $\varepsilon > 0$, selecting $N \in \mathbb{N}$ such that $\sum_{l\ge N+1} 2^{-l} < \varepsilon/2$ gives, from the definition (78) of d_E,

$$\bigcap_{l=1}^{N} E_{\varphi_l,\varepsilon/4} \subset B_{d_E}(0,\varepsilon).$$

Hence, any ball for d_E centered at 0 is a neighborhood of 0 for the $L^2(\Omega)$ weak topology. Since d_E and the $L^2(\Omega)$ weak neighborhood are invariant by translation, this concludes the proof that this weak topology is identical to the topology generated by d_E.
The conclusion on weak uniform convergence of sequences of functions follows from the preceding result, and more precisely by noticing that all previous inclusions are, when applied to $u_m(t)-u(t)$, uniform with respect to $t \in [0,T]$. ■

The following lemma has been established in [28, Proposition 9.3] but its proof is recalled for the reader's convenience.

Lemma 6.6
Let $(t^{(n)})_{n\in\mathbb{Z}}$ be a stricly increasing sequence of real values such that $\delta t^{(n+\frac{1}{2})} := t^{(n+1)} - t^{(n)}$ is uniformly bounded by $\delta t > 0$, $\lim_{n\to-\infty} t^{(n)} = -\infty$ and $\lim_{n\to\infty} t^{(n)} = \infty$. For all $t \in \mathbb{R}$, we denote by

$n(t)$ the element $n \in \mathbb{Z}$ such that $t \in [t^{(n)}, t^{(n+1)})$. Let $(a^{(n)})_{n\in\mathbb{Z}}$ be a family of non negative real values with a finite number of non zero values. Then

$$\int_{\mathbb{R}} \sum_{n=n(t)+1}^{n(t+\tau)} (\delta t^{(n+\frac{1}{2})} a^{(n+1)}) \mathrm{d}t = \tau \sum_{n\in\mathbb{Z}} (\delta t^{(n+\frac{1}{2})} a^{(n+1)}), \ \forall \tau \in (0, +\infty), \tag{79}$$

and

$$\int_{\mathbb{R}} \left(\sum_{n=n(t)+1}^{n(t+\tau)} \delta t^{(n+\frac{1}{2})} \right) a^{n(t+\zeta)+1} \mathrm{d}t$$
$$\leq (\tau + \delta t) \sum_{n\in\mathbb{Z}} (\delta t^{(n+\frac{1}{2})} a^{(n+1)}), \ \forall \tau \in (0, +\infty), \ \forall \zeta \in \mathbb{R}. \tag{80}$$

Proof.
Let us define the function $\chi(t, n, \tau)$ by $\chi(t, n, \tau) = 1$ if $t < t^{(n)}$ and $t + \tau \geq t^{(n)}$, else $\chi(t, n, \tau) = 0$. We have

$$\begin{aligned} \int_{\mathbb{R}} \sum_{n=n(t)+1}^{n(t+\tau)} (\delta t^{(n+\frac{1}{2})} a^{(n+1)}) \mathrm{d}t &= \int_{\mathbb{R}} \sum_{n\in\mathbb{Z}} (\delta t^{(n+\frac{1}{2})} a^{(n+1)} \chi(t, n, \tau)) \mathrm{d}t \\ &= \sum_{n\in\mathbb{Z}} \left(\delta t^{(n+\frac{1}{2})} a^{(n+1)} \int_{\mathbb{R}} \chi(t, n, \tau) \mathrm{d}t \right). \end{aligned}$$

Since $\int_{\mathbb{R}} \chi(t, n, \tau)\mathrm{d}t = \int_{t^{(n)}-\tau}^{t^{(n)}} \mathrm{d}t = \tau$, thus (79) is proven.
We now turn to the proof of (80). We define the function $\tilde{\chi}(n, t)$ by $\tilde{\chi}(n, t) = 1$ if $n(t) = n$, else $\tilde{\chi}(n, t) = 0$. We have

$$\begin{aligned} &\int_{\mathbb{R}} \left(\sum_{n=n(t)+1}^{n(t+\tau)} \delta t^{(n+\frac{1}{2})} \right) a^{(n(t+\zeta)+1)} \mathrm{d}t \\ &= \int_{\mathbb{R}} \left(\sum_{n=n(t)+1}^{n(t+\tau)} \delta t^{(n+\frac{1}{2})} \right) \sum_{m\in\mathbb{Z}} a^{(m+1)} \tilde{\chi}(m, t+\zeta) \mathrm{d}t, \end{aligned}$$

which yields

$$\begin{aligned} &\int_{\mathbb{R}} \left(\sum_{n=n(t)+1}^{n(t+\tau)} \delta t^{(n+\frac{1}{2})} \right) a^{(n(t+\zeta)+1)} \mathrm{d}t \\ &= \sum_{m\in\mathbb{Z}} a^{(m+1)} \int_{t^m-\zeta}^{t^{m+1}-\zeta} \left(\sum_{n=n(t)+1}^{n(t+\tau)} \delta t^{(n+\frac{1}{2})} \right) \mathrm{d}t. \end{aligned} \tag{81}$$

Since we have

$$\sum_{n=n(t)+1}^{n(t+\tau)} \delta t^{(n+\frac{1}{2})} = \sum_{n\in\mathbb{Z},\ t<t^{(n)}\leq t+\tau} (t^{(n+1)} - t^{(n)}) \leq \tau + \delta t,$$

we can write from (81)

$$\begin{aligned} \int_{\mathbb{R}} \left(\sum_{n=n(t)+1}^{n(t+\tau)} \delta t^{(n+\frac{1}{2})} \right) a^{(n(t+\zeta)+1)} \mathrm{d}t &\leq (\tau + \delta t) \sum_{m\in\mathbb{Z}} a^{(m+1)} \int_{t^{(m)}-\zeta}^{t^{(m+1)}-\zeta} \mathrm{d}t \\ &= (\tau + \delta t) \sum_{m\in\mathbb{Z}} a^{(m+1)} \delta t^{(m+\frac{1}{2})}, \end{aligned}$$

which is exactly (80).

■

References

[1] I. Aavatsmark, T. Barkve, O. Boe, and T. Mannseth. Discretization on non-orthogonal, quadrilateral grids for inhomogeneous, anisotropic media. *J. Comput. Phys.*, 127(1):2–14, 1996.

[2] B. Andreianov, F. Boyer, and F. Hubert. Discrete duality finite volume schemes for Leray-Lions-type elliptic problems on general 2D meshes. *Numer. Methods Partial Differential Equations*, 23(1):145–195, 2007.

[3] H. Attouch, G. Buttazzo, and G. Michaille. *Variational analysis in Sobolev and BV spaces*, volume 6 of *MPS/SIAM Series on Optimization*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA; Mathematical Programming Society (MPS), Philadelphia, PA, 2006.

[4] M. Bertsch, P. De Mottoni, and L. Peletier. The Stefan problem with heating: appearance and disappearance of a mushy region. *Trans. Amer. Math. Soc*, 293:677–691, 1986.

[5] F. Brezzi and M. Fortin. *Mixed and hybrid finite element methods*, volume 15 of *Springer Series in Computational Mathematics*. Springer-Verlag, New York, 1991.

[6] F. Brezzi, K. Lipnikov, and V. Simoncini. A family of mimetic finite difference methods on polygonal and polyhedral meshes. *Math. Models Methods Appl. Sci.*, 15(10):1533–1551, 2005.

[7] P. Ciarlet. The finite element method. In P. G. Ciarlet and J.-L. Lions, editors, *Part I*, Handbook of Numerical Analysis, III. North-Holland, Amsterdam, 1991.

[8] Y. Coudière and F. Hubert. A 3d discrete duality finite volume method for nonlinear elliptic equations. *SIAM Journal on Scientific Computing*, 33(4):1739–1764, 2011.

[9] M. Crouzeix and P.-A. Raviart. Conforming and nonconforming finite element methods for solving the stationary Stokes equations. I. *Rev. Française Automat. Informat. Recherche Opérationnelle Sér. Rouge*, 7(R-3):33–75, 1973.

[10] K. Deimling. *Nonlinear functional analysis*. Springer-Verlag, Berlin, 1985.

[11] J. Diaz and F. de Thelin. On a nonlinear parabolic problem arising in some models related to turbulent flows. *SIAM J. Math. Anal.*, 25(4):1085–1111, 1994.

[12] J. Droniou. Finite volume schemes for fully non-linear elliptic equations in divergence form. *ESAIM: Mathematical Modelling and Numerical Analysis*, 40(6):1069, 2006.

[13] J. Droniou and R. Eymard. A mixed finite volume scheme for anisotropic diffusion problems on any grid. *Numer. Math.*, 105(1):35–71, 2006.

[14] J. Droniou, R. Eymard, T. Gallouët, C. Guichard, and R. Herbin. Gradient schemes for elliptic and parabolic problems. 2014. To appear.

[15] J. Droniou, R. Eymard, T. Gallouët, and R. Herbin. A unified approach to mimetic finite difference, hybrid finite volume and mixed finite volume methods. *Math. Models Methods Appl. Sci.*, 20(2):265–295, 2010.

[16] J. Droniou, R. Eymard, T. Gallouët, and R. Herbin. Gradient schemes: a generic framework for the discretisation of linear, nonlinear and nonlocal elliptic and parabolic equations. *Math. Models Methods Appl. Sci. (M3AS)*, 23(13):2395–2432, 2013.

[17] J. Droniou, R. Eymard, and C. Guichard. Uniform-in-time convergence of numerical schemes for Richards' and Stefan's models. In *Finite Volumes for Complex Applications VII*. Springer, 2014.

[18] M. G. Edwards and C. F. Rogers. Finite volume discretization with imposed flux continuity for the general tensor pressure equation. *Comput. Geosci.*, 2(4):259–290, 1998.

[19] I. Ekeland and R. Témam. *Convex analysis and variational problems*, volume 28 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, english edition, 1999. Translated from the French.

[20] A. Ern and J.-L. Guermond. *Theory and practice of finite elements*, volume 159. Springer, 2004.

[21] R. Eymard, P. Féron, T. Gallouet, R. Herbin, and C. Guichard. Gradient schemes for the Stefan problem.

[22] R. Eymard, T. Gallouët, and R. Herbin. Discretization of heterogeneous and anisotropic diffusion problems on general nonconforming meshes SUSHI: a scheme using stabilization and hybrid interfaces. *IMA J. Numer. Anal.*, 30(4):1009–1043, 2010.

[23] R. Eymard, T. Gallouët, D. Hilhorst, and Y. Naït Slimane. Finite volumes and nonlinear diffusion equations. *RAIRO Modél. Math. Anal. Numér.*, 32(6):747–761, 1998.

[24] R. Eymard, C. Guichard, and R. Herbin. Small-stencil 3d schemes for diffusive flows in porous media. *M2AN*, 46:265–290, 2012.

[25] R. Eymard, C. Guichard, R. Herbin, and R. Masson. Gradient schemes for two-phase flow in heterogeneous porous media and Richards equation. *ZAMM*, page accepted for publication, 2013.

[26] R. Eymard, M. Gutnic, and D. Hilhorst. The finite volume method for Richards equation. *Comput. Geosci.*, 3(3-4):259–294, 1999.

[27] R. Eymard and R. Herbin. Gradient scheme approximations for diffusion problems. *Finite Volumes for Complex Applications VI Problems & Perspectives*, pages 439–447, 2011.

[28] R. Eymard, R. Herbin, and A. Michel. Mathematical study of a petroleum-engineering scheme. *M2AN Math. Model. Numer. Anal.*, 37(6):937–972, 2003.

[29] R. Glowinski and J. Rappaz. Approximation of a nonlinear elliptic problem arising in a non-newtonian fluid flow model in glaciology. *M2AN Math. Model. Numer. Anal.*, 37(1):175–186, 2003.

[30] F. Hermeline. Approximation of diffusion operators with discontinuous tensor coefficients on distorted meshes. *Computer methods in applied mechanics and engineering*, 192(16):1939–1959, 2003.

[31] A. V. Kazhikhov. Recent developments in the global theory of two-dimensional compressible Navier-Stokes equations. 25:ii+52, 1998.

[32] E. Maitre. Numerical analysis of nonlinear elliptic-parabolic equations. *M2AN Math. Model. Numer. Anal.*, 36(1):143–153, 2002.

[33] G. Minty. On a monotonicity method for the solution of non- linear equations in Banach spaces. *Proceedings of the National Academy of Sciences of the United States of America*, 50(6):1038, 1963.

[34] R. H. Nochetto and C. Verdi. Approximation of degenerate parabolic problems using numerical integration. *SIAM J. Numer. Anal.*, 25(4):784–814, 1988.

[35] I. S. Pop. Numerical schemes for degenerate parabolic problems. In *Progress in industrial mathematics at ECMI 2004*, volume 8 of *Math. Ind.*, pages 513–517. Springer, Berlin, 2006.