Available online at www.sciencedirect.com

ScienceDirect

J. Differential Equations 260 (2016) 7821–7860

Journal of Differential Equations

www.elsevier.com/locate/jde

Convergence in $C([0, T]; L^2(\Omega))$ of weak solutions to perturbed doubly degenerate parabolic equations

Jérôme Droniou [a], Robert Eymard [b], Kyle S. Talbot [a,*]

[a] School of Mathematical Sciences, Monash University, Clayton, Victoria 3800, Australia
[b] 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

Received 16 June 2015; revised 13 January 2016

Available online 23 February 2016

Abstract

We study the behaviour of solutions to a class of nonlinear degenerate parabolic problems when the data are perturbed. The class includes the Richards equation, Stefan problem and the parabolic p-Laplace equation. We show that, up to a subsequence, weak solutions of the perturbed problem converge uniformly-in-time to weak solutions of the original problem as the perturbed data approach the original data. We do not assume uniqueness or regularity. When uniqueness is known, our result demonstrates that the weak solution is uniformly temporally stable to perturbations of the data. Beginning with a proof of temporally-uniform, spatially-weak convergence, we strengthen the latter by relating the unknown to an underlying convex structure that emerges naturally from energy estimates. The double degeneracy — shown to be equivalent to a maximal monotone operator framework — is handled with techniques inspired by a classical monotonicity argument and a simple variant of the compensated compactness phenomenon.

MSC: 35D30; 35K20; 35K55; 35K65; 35K92

Keywords: Uniform temporal convergence; Degenerate parabolic equation; Leray–Lions operator; Maximal monotone operator; Richards equation; Stefan problem

---

[*] Corresponding author.
E-mail addresses: jerome.droniou@monash.edu (J. Droniou), robert.eymard@u-pem.fr (R. Eymard), kyle.talbot@monash.edu (K.S. Talbot).

http://dx.doi.org/10.1016/j.jde.2016.02.004
0022-0396/

1. Introduction

Consider the class of doubly nonlinear parabolic problems

$$\begin{cases} \partial_t \beta(u) - \operatorname{div}[a(x, \nu(u), \nabla\zeta(u))] = f & \text{in } \Omega \times (0,T), \\ \beta(u)(x,0) = \beta(u^{\mathrm{ini}})(x) & \text{in } \Omega, \\ \zeta(u) = 0 & \text{on } \partial\Omega \times (0,T) \end{cases} \tag{P}$$

on a bounded open subset Ω of $\mathbb{R}^d$. The functions β and ζ are nondecreasing and the function ν satisfies $\nu' = \beta'\zeta'$. The operator a is of Leray–Lions type and $u^{\mathrm{ini}} \in L^2(\Omega)$. In applications one may have only approximate knowledge of the data $(\beta, \zeta, \nu, a, f, u^{\mathrm{ini}})$, and one is interested in the value of the solution at a particular instant in time. The main result of this article concerns the continuity of $\nu(u)$ with respect to perturbations of the data. For $n \in \mathbb{N}$, consider perturbed problems with corresponding solutions u_n:

$$\begin{cases} \partial_t \beta_n(u_n) - \operatorname{div}[a_n(x, \nu_n(u_n), \nabla\zeta_n(u_n))] = f_n & \text{in } \Omega \times (0,T), \\ \beta_n(u_n)(x,0) = \beta_n(u_n^{\mathrm{ini}})(x) & \text{in } \Omega, \\ \zeta_n(u_n) = 0 & \text{on } \partial\Omega \times (0,T). \end{cases} \tag{P$_n$}$$

If the data $(\beta_n, \zeta_n, \nu_n, a_n, f_n, u_n^{\mathrm{ini}})$ converge to $(\beta, \zeta, \nu, a, f, u^{\mathrm{ini}})$ in suitable manners, we show that, up to a subsequence, $\nu_n(u_n)$ converges to $\nu(u)$ in $C([0,T]; L^2(\Omega))$.

Instances of (P) arise in various contexts. We focus our attention upon three models in particular: the Richards equation, the Stefan problem, and the parabolic p-Laplace equation. By taking ζ to be the identity, $\nu = \beta$ and $a(x, \nu(u), \nabla\zeta(u)) = K(x, \beta(u))\nabla u$, we recover the first of these, which describes the flow of water in an unsaturated porous medium [25,29]. The quantity of interest is the pressure-dependent saturation $\beta(u)$, with $K(x, \beta(u))$ the mobility. A model of the Stefan problem [6] of heat diffusion in a medium undergoing phase transition is realised by taking β to be the identity, $\nu = \zeta$ and $a(x, \nu(u), \nabla\zeta(u)) = K(x, \zeta(u))\nabla\zeta(u)$. Here we are interested in the enthalpy-dependent temperature $\zeta(u)$, with $K(x, \zeta(u))$ representing the thermal conductivity. To recover the parabolic p-Laplace equation, take each of β, ζ and ν to be the identity and $a(x, \nu(u), \nabla\zeta(u)) = |\nabla u|^{p-2}\nabla u$. The parabolic p-Laplace equation features in, for example, the theory of non-Newtonian filtration; see E. DiBenedetto's monograph [12] and the references therein.

In each of these examples the quantity of practical interest is $\nu(u)$. More specifically, it is the value of $\nu(u)$ at a particular instant in time, say $t = T$. Pragmatically speaking, it is therefore critical that $\nu(u)(T)$ be stable to perturbations of the data. Our main result shows this to be the case in each of the above examples, where uniqueness of the solution is known (at least if K depends only upon x; see Appendix C). For general problems (P), uniqueness appears to be open, so we can only assert that a subsequence of $\nu_n(u_n)$ converges to $\nu(u)$, where u is a solution of the limit problem.

The existence and uniqueness of weak solutions to (P) with $\zeta = \mathrm{Id}$ is studied in the seminal article of H.W. Alt and S. Luckhaus [1]. F. Otto [28] subsequently improved their uniqueness result by removing a linearity assumption on the diffusion operator $a(\nu(u), \nabla u)$, and by assuming independence with respect to x, strict monotony with respect to ∇u and Hölder continuity with respect to $\nu(u)$. However, to our knowledge there are no existence and uniform temporal-strong spatial stability results for parabolic equations with as many nonlinearities and degeneracies as (P).

Stability results do exist for simplified models. Using techniques from nonlinear semigroup theory, P. Bénilan and M.G. Crandall [5] show that solutions to the Cauchy problem for $\partial_t u - \Delta\varphi(u) = 0$ on the whole space are stable in $C([0,T]; L^1(\mathbb{R}^d))$ with respect to pointwise perturbations of φ and $L^1(\mathbb{R}^d)$-perturbations of the initial datum. D. Blanchard and A. Porretta [8] demonstrate the $L^\infty(0,T; L^1(\Omega))$-stability of renormalised solutions to the initial-boundary value problem for $\partial_t b(u) - \mathrm{div}(a(x,u,\nabla u)) + \mathrm{div}(\Phi(u)) = f$, under L^1-perturbations of the source and initial datum. The authors assume that b is a maximal monotone graph on $\mathbb{R}$, $b^{-1} \in C(\mathbb{R})$ and a is a Leray–Lions operator. We refer the reader to Section 3 for further comparisons of our work to this reference.

Stability for other notions of solution to degenerate parabolic problems has also been considered. In the framework of entropy solutions, B. Andreianov et al. [2] demonstrate the stability in $L^1(\Omega \times (0,T))$ of solutions to (P) with additional convection and reaction terms, but with specific assumptions on the monotonicity of ζ and a. I.C. Kim and N. Požár [20] show that viscosity solutions to the Richards equation are stable. One can also consider stability of solutions to the parabolic p-Laplace equation with respect to perturbations of p. To this end, we refer the reader to the work of J. Kinnunen and M. Parviainen [21] and subsequently T. Lukkari and Parviainen [24].

The convergence of $\nu_n(u_n)$ to $\nu(u)$ in $C([0,T]; L^2(\Omega))$ cannot be deduced by mere interpolation from the uniform-in-time $L^1(\Omega)$ stability results in the previous references, since the best uniform-in-time bound that we can obtain for $\nu(u)$ is in $L^2(\Omega)$. From the viewpoint of uniform-in-time estimates, establishing a convergence result in this “limit” space $L^2(\Omega)$ therefore requires new ideas. The first step is the uniform-$[0,T]$, weak-$L^2(\Omega)$ convergence of $\beta_n(u_n)$ to $\beta(u)$. A key ingredient of the proof of this fact, and indeed much of our paper, is the function B (and its perturbed analogue B_n) defined below in (2.4). The importance of B was previously observed in [1] when $\zeta = \mathrm{Id}$. It enables energy estimates on the solution via an integration-by-parts formula for the action of $\partial_t\beta(u)$ on $\zeta(u)$. These estimates are sufficient for us to deduce the aforementioned convergence of $\beta_n(u_n)$ thanks to Proposition 4.9, a uniform-in-time, weak-in-space analogue of the Aubin–Simon compactness theorem. The spatial compactness is weak here since (P) does not provide any information on the gradient of $\beta(u)$. The convexity of B yields lower semi-continuity of certain integral functionals, that when combined with the energy identity satisfied by the limit solution, enables us to prove the uniform convergence of $\int_\Omega B_n(\beta_n(u_n)(x,\cdot))\,\mathrm{d}x$ on $[0,T]$. A uniform convexity property of B connects the convergence of these integrals to that of $\nu_n(u_n)$ in $L^2(\Omega)$, thus enhancing the convergence of $\nu_n(u_n)$ to prove the main result, Theorem 2.3.

We anticipate that these ideas for obtaining uniform-temporal, strong-L^2 spatial dependence of solutions upon the data may generalise to systems of equations as in [1], and to convection–diffusion–reaction equations of the form studied in [2], but in the variational setting.

We obtain the existence of solutions to (P) as a straightforward corollary to Theorem 2.3. When $a(x,\nu(u),\nabla\zeta(u)) = \Lambda(x)\nabla\zeta(u)$, we give a short uniqueness proof in Appendix C. We do not, however, address uniqueness or regularity for general a. With the nonlinearities in (P) and the irregularities in the data seen in the applications described above, one cannot expect to obtain such properties in these instances. Indeed, examples of non-uniqueness of weak solutions exist, see [15, Remark 3.4] for stationary Leray–Lions equations (corresponding to $\beta = 0$ and $\zeta = \mathrm{Id}$).

Since β and ζ may share common plateaux, one of the challenges in studying compactness properties of solutions to (P) is identifying weak limits. Our method handles this difficulty principally using a monotonicity argument. However, the double degeneracy necessitates the use of a compensated compactness lemma (see Remark 5.1), which in our setting is actually a direct consequence of the Aubin–Simon theorem.

These tools enable us to generalise some aspects of [8], at least when the regularity index p is not too small; see the concluding remarks to Section 3 for additional discussion on this point. The first two authors of the current article use similar techniques [14] for the convergence analysis of numerical approximations of (P). Discrete compensated compactness was recently employed by B. Andreianov, C. Cancès and A. Moussa [3] to identify the limits of numerical schemes in the framework of maximal monotone operators.

The article is organised as follows. In Section 2 we list the hypotheses on the model (P) and state the main result, Theorem 2.3. In Section 3 we recast the problem in the framework of maximal monotone operators and give the analogue of Theorem 2.3 in this setting. In Section 4.1 we note some technical properties of the function B. To focus attention on the convergence problem, some of these results are only stated. For proofs, the reader should consult [14]. Section 4.2 establishes our estimates. Section 4.3 presents two lemmas that play an important role in the proof of Theorem 2.3, and which may be of independent interest. Our temporally-uniform, spatially-weak analogue of the Aubin–Simon compactness theorem occupies Section 4.4. Section 5 is the proof of the convergence results, including the $C([0,T];L^2(\Omega))$ convergence. Appendix A lists several minor lemmas that we employ throughout the article. Aubin–Simon compactness appears again in Appendix B, where we use it to prove a compensated compactness lemma adapted for our current setting. Appendix C is a self-contained uniqueness proof when the Leray–Lions operator a is linear.

2. Hypotheses and main result

We assume that $T>0$, Ω is a bounded open subset of $\mathbb{R}^d$ ($d\in\mathbb{N}$), and

$$\beta:\mathbb{R}\to\mathbb{R} \text{ is nondecreasing, Lipschitz continuous with Lipschitz constant } L_\beta>0, \text{ and satisfies } \beta(0)=0. \tag{2.1a}$$

$$\zeta:\mathbb{R}\to\mathbb{R} \text{ is nondecreasing, Lipschitz continuous with Lipschitz constant } L_\zeta>0, \text{ and satisfies } \zeta(0)=0. \text{ Furthermore, there are positive constants } M_1, M_2 \text{ such that for every } s\in\mathbb{R},\ |\zeta(s)|\geq M_1|s|-M_2. \tag{2.1b}$$

$$\text{For all } s\in\mathbb{R},\qquad \nu(s)=\int_0^s \zeta'(q)\beta'(q)\,\mathrm{d}q. \tag{2.1c}$$

Fix $p\in(1,\infty)$ and denote by $p'=\frac{p}{p-1}$ its Hölder conjugate. We assume that $a:\Omega\times\mathbb{R}\times\mathbb{R}^d\to\mathbb{R}^d$ is Carathéodory, and that there are constants $\underline{a}$, $\mu>0$ and a function $\overline{a}\in L^{p'}(\Omega)$ such that for almost every $x\in\Omega$, every $s\in\mathbb{R}$ and for all $\xi,\chi\in\mathbb{R}^d$,

$$a(x,s,\xi)\cdot\xi\geq\underline{a}|\xi|^p, \tag{2.1d}$$

$$|a(x,s,\xi)|\leq\overline{a}(x)+\mu|\xi|^{p-1}, \tag{2.1e}$$

$$(a(x,s,\xi)-a(x,s,\chi))\cdot(\xi-\chi)\geq 0. \tag{2.1f}$$

The source term and initial trace satisfy

$$f \in L^{p'}(0,T;W^{-1,p'}(\Omega)), \quad u^{\mathrm{ini}} \in L^2(\Omega). \tag{2.1g}$$

Due to the double degeneracy (from β and ζ), identifying weak limits obtained by compactness results is challenging and requires monotonicity and compensated compactness techniques. To prove that weak limits of sequences of solutions to (P) are also solutions to (P), we consider three separate cases for p:

$$\left\{\begin{array}{ll} \text{(I)} & p \geq 2, \\ \text{or} & \\ \text{(II)} & \dfrac{2d}{d+2} < p < 2 \text{ and there are positive constants } M_3, M_4 \text{ such that for all } s \in \mathbb{R}, \\ & |\beta(s)| \geq M_3|s| - M_4, \\ \text{or} & \\ \text{(III)} & 1 < p \leq \dfrac{2d}{d+2}, \text{ there are positive constants } M_3, M_4 \text{ such that for all } s \in \mathbb{R}, \\ & |\beta(s)| \geq M_3|s| - M_4, \text{ and } \beta \text{ is (strictly) increasing.} \end{array}\right. \tag{2.2}$$

Remark 2.1.

(i) The assumption $\beta(0) = \zeta(0) = 0$ is not restrictive, since replacing β and ζ with $\beta - \beta(0)$ and $\zeta - \zeta(0)$ (respectively) does not change the problem.

(ii) Hypotheses (2.1d) and (2.1e) can be relaxed to

$$a(x,s,\xi)\cdot\xi \geq \underline{a}|\xi|^p - \Theta(x) \qquad \text{with } \Theta \in L^1(\Omega),$$

and

$$|a(x,s,\xi)| \leq \overline{a}(x) + \mu|s|^q + \mu|\xi|^{p-1} \qquad \text{with } q < \max(2/p', p-1).$$

(iii) The condition $p > \frac{2d}{d+2}$ in (2.2) is equivalent to $p^* > 2$, where p^* is the Sobolev exponent of p; i.e., $p^* = \frac{dp}{d-p}$ if $p < d$ and $p^* = +\infty$ if $p \geq d$.

(iv) Since the basic energy estimates on (P) provide strong compactness for $\nu(u)$ (see (5.5)), we can just as easily handle source terms of the form $f(x,t,\nu(u))$ as in [1].

Denote by R_β the range of β and for $s \in R_\beta$ define the right inverse $\beta^r : R_\beta \to \mathbb{R}$ of β by

$$\beta^r(s) = \begin{cases} \inf\{t \in \mathbb{R} \mid \beta(t) = s\} & \text{if } s > 0, \\ 0 & \text{if } s = 0, \\ \sup\{t \in \mathbb{R} \mid \beta(t) = s\} & \text{if } s < 0. \end{cases} \tag{2.3}$$

That is, $\beta^r(s)$ is the closest t to 0 such that $\beta(t) = s$. Since $\beta(0) = 0$, note that β^r is nondecreasing, nonnegative on $R_\beta \cap \mathbb{R}^+$ and nonpositive on $R_\beta \cap \mathbb{R}^-$. We can therefore extend β^r as a function $\overline{R_\beta} \to [-\infty, \infty]$. We then define $B : \overline{R_\beta} \to [0, \infty]$ by

$$B(z) = \int_0^z \zeta(\beta^r(s))\,\mathrm{d}s. \tag{2.4}$$

The signs of ζ and β^r ensure that B is nonnegative on $\overline{R_\beta}$, nondecreasing on $\overline{R_\beta} \cap \mathbb{R}^+$ and nonincreasing on $\overline{R_\beta} \cap \mathbb{R}^-$. Moreover, since ζ and β^r are non-decreasing, B is convex on $\overline{R_\beta}$. This calls for extending B as a function $\mathbb{R} \to [0, +\infty]$ by setting $B = +\infty$ outside $\overline{R_\beta}$. This function is still nondecreasing on $\mathbb{R}^+$ and nonincreasing on $\mathbb{R}^-$.

Our notion of solution to (P) is as follows.

Definition 2.2. Under Hypotheses (2.1), a *solution* to (P) is a function u satisfying

$$\left\{\begin{array}{l} u : \Omega \times (0,T) \to \mathbb{R} \text{ is measurable}, \quad \zeta(u) \in L^p(0,T; W_0^{1,p}(\Omega)), \\ B(\beta(u)) \in L^\infty(0,T; L^1(\Omega)), \quad \beta(u) \in C([0,T]; L^2(\Omega)\text{-w}), \\ \partial_t \beta(u) \in L^{p'}(0,T; W^{-1,p'}(\Omega)), \quad \beta(u)(\cdot,0) = \beta(u^{\mathrm{ini}}) \text{ in } L^2(\Omega), \\ \displaystyle\int_0^T \langle \partial_t \beta(u)(\cdot,t), v(\cdot,t)\rangle_{W^{-1,p'},W_0^{1,p}}\,\mathrm{d}t \\ \qquad \displaystyle + \int_0^T \int_\Omega a(x, \nu(u(x,t)), \nabla\zeta(u)(x,t)) \cdot \nabla v(x,t)\,\mathrm{d}x\,\mathrm{d}t \\ \quad \displaystyle = \int_0^T \langle f(\cdot,t), v(\cdot,t)\rangle_{W^{-1,p'},W_0^{1,p}}\,\mathrm{d}t \quad \forall v \in L^p(0,T; W_0^{1,p}(\Omega)). \end{array}\right. \tag{2.5}$$

Here $C([0,T]; L^2(\Omega)\text{-w})$ denotes the space of continuous functions from $[0,T]$ into $L^2(\Omega)$, where the latter is equipped with the weak topology. This notion of continuity for $\beta(u)$ can be understood as a natural consequence of the integrability of $\beta(u)$ and the PDE itself. Fix $\varphi \in C_c^\infty(\Omega)$ and consider the map $\mathcal{L}_\varphi : [0,T] \to \mathbb{R}$, $t \mapsto \langle \beta(u)(t), \varphi\rangle_{L^2(\Omega)}$. One can show, using the PDE in the sense of distributions and the fact that $\beta(u) \in L^\infty(0,T; L^2(\Omega))$ (see Estimate (4.5) below), that $\mathcal{L}_\varphi \in W^{1,1}(0,T) \subset C([0,T])$. From the density of $C_c^\infty(\Omega)$ in $L^2(\Omega)$ one deduces that for every $\varphi \in L^2(\Omega)$, $\mathcal{L}_\varphi \in C([0,T])$. That is, $\beta(u) : [0,T] \to L^2(\Omega)$-w is continuous.

The main result of this paper is the following convergence theorem.

Theorem 2.3. *Let $(\beta_n, \zeta_n, \nu_n, a_n, f_n, u_n^{\mathrm{ini}})_{n\in\mathbb{N}}$ be a sequence converging to $(\beta, \zeta, \nu, a, f, u^{\mathrm{ini}})$ in the following sense:*

$$\left\{\begin{array}{l} \beta_n, \zeta_n \textit{ and } \nu_n \textit{ converge locally uniformly on } \mathbb{R} \textit{ to } \beta, \zeta \textit{ and } \nu, \textit{ respectively}; \\ \textit{for almost every } x \in \Omega,\ a_n(x,\cdot,\cdot) \to a(x,\cdot,\cdot) \textit{ locally uniformly on } \mathbb{R} \times \mathbb{R}^d; \\ f_n \to f \textit{ in } L^{p'}(0,T; W^{-1,p'}(\Omega)) \textit{ and } u_n^{\mathrm{ini}} \to u^{\mathrm{ini}} \textit{ in } L^2(\Omega). \end{array}\right. \tag{2.6}$$

Assume that $(\beta, \zeta, \nu, a, f, u^{\mathrm{ini}})$ and $(\beta_n, \zeta_n, \nu_n, a_n, f_n, u_n^{\mathrm{ini}})$ (for every $n \in \mathbb{N}$) satisfy (2.1) and (2.2), and that the constants L_β, L_ζ, M_1, M_2, M_3, M_4, $\underline{a}$, μ and the function $\overline{a}$ are in-

dependent of n. *Let* u_n *be a solution to* (P$_n$). *Then there is a solution* u *of* (P) *such that, up to a subsequence,*

$$\begin{cases} \beta_n(u_n) \to \beta(u) & \text{in } C([0,T]; L^2(\Omega)\text{-w}), \\ \nu_n(u_n) \to \nu(u) & \text{in } C([0,T]; L^2(\Omega)), \text{ and} \\ \zeta_n(u_n) \rightharpoonup \zeta(u) & \text{weakly in } L^p(0,T; W_0^{1,p}(\Omega)). \end{cases} \tag{2.7}$$

If in addition we assume that a *is strictly monotone, that is, the inequality in* (2.1f) *with* $\chi \neq \xi$ *is strict, then*

$$\zeta_n(u_n) \to \zeta(u) \quad \text{strongly in } L^p(0,T; W_0^{1,p}(\Omega)). \tag{2.8}$$

Remark 2.4. This theorem provides a stability result for any subclass of problem (P) for which uniqueness of the weak solution is known. This is indeed the case for simplified versions of the Richards equation and Stefan problem; see Appendix C. In these settings, Theorem 2.3 shows that the *whole sequence* $\beta_n(u_n)$ (respectively $\zeta_n(u_n)$) converges uniformly-in-time to $\beta(u)$ (respectively $\zeta(u)$).

Remark 2.5. Since β_n, ζ_n and ν_n are nondecreasing and β, ζ and ν are continuous, Dini's theorem shows that we only need to assume that β_n, ζ_n and ν_n converge pointwise. The Arzelà–Ascoli theorem can be used to arrive at the same conclusion, since β_n, ζ_n and ν_n are uniformly Lipschitz continuous.

Remark 2.6. The local uniform convergence on $\mathbb{R}$ of ν_n to ν holds if we assume that $\beta_n' \to \beta'$ almost everywhere on $\mathbb{R}$, or $\zeta_n' \to \zeta'$ almost everywhere on $\mathbb{R}$. Indeed, suppose that the latter pointwise convergence holds. Since $(\beta_n')_{n\in\mathbb{N}}$ is bounded by L_β, up to a subsequence, $\beta_n' \rightharpoonup \chi$ weak-$*$ in $L^\infty(\mathbb{R})$ for some bounded $\chi : \mathbb{R} \to \mathbb{R}$. Then as $n \to \infty$,

$$\beta_n(s) = \int_0^s \beta_n'(q)\,\mathrm{d}q \to \int_0^s \chi(q)\,\mathrm{d}q.$$

But $\beta_n(s) \to \beta(s)$ for every $s \in \mathbb{R}$, so it must be that $\chi = \beta'$ almost everywhere on $\mathbb{R}$ and therefore that $\beta_n' \rightharpoonup \beta'$ weak-$*$ in $L^\infty(\mathbb{R})$. One can then pass to the limit as $n \to \infty$ in the definition of ν_n, using dominated convergence on the sequence $(\zeta_n')_{n\in\mathbb{N}}$, to obtain the local uniform convergence towards ν.

Remark 2.7. Observe that in the case $1 < p \leq \frac{2d}{d+2}$ in (2.2), we do not need the strict monotonicity of each β_n; we only require that the limit β does not have any plateaux.

As a by-product of this convergence result, we obtain existence for (P).

Corollary 2.8. *Under Hypotheses* (2.1) *and* (2.2), *there exists a solution to* (P).

Proof. Theorem 2.3 shows that we only need to establish the existence of a solution for perturbed problems (P). Upon replacing β and ζ by $\beta + \delta\,\mathrm{Id}$ and $\zeta + \delta\,\mathrm{Id}$ for some small $\delta > 0$, we can therefore assume that

$$\beta' \geq \delta \quad \text{and} \quad \zeta' \geq \delta \quad \text{on } \mathbb{R}.$$

In particular, these perturbed β and ζ are bi-Lipschitz homeomorphisms, and we define

$$a_0(x,s,\xi) = a(x, \nu(\beta^{-1}(s)), (\zeta \circ \beta^{-1})'(s)\xi), \tag{2.9}$$

where for some $\rho > 0$, $(\zeta \circ \beta^{-1})'(s) \in [\rho, \rho^{-1}]$ for all $s \in \mathbb{R}$. The function a_0 satisfies (2.1d)–(2.1f). J.-L. Lions showed [23] that there exists a solution to

$$\begin{cases} \partial_t v - \operatorname{div}(a_0(x,v,\nabla v)) = f & \text{in } \Omega \times (0,T), \\ v(x,0) = \beta(u^{\text{ini}})(x) & \text{in } \Omega, \\ v = 0 & \text{on } \partial\Omega \times (0,T) \end{cases} \tag{2.10}$$

in the sense that $v \in L^p(0,T;W_0^{1,p}(\Omega)) \cap C([0,T];L^2(\Omega))$, $\partial_t v \in L^{p'}(0,T;W^{-1,p'}(\Omega))$, $v(\cdot,0) = \beta(u^{\text{ini}})$, and the equation is satisfied against any test function in $L^p(0,T;W_0^{1,p}(\Omega))$.

We then set $v = \beta(u)$. Then $\zeta(u) = (\zeta \circ \beta^{-1})(v)$ with $\zeta \circ \beta^{-1}$ Lipschitz continuous, and thus $\zeta(u) \in L^p(0,T;W_0^{1,p}(\Omega))$. We have $\beta(u) = v \in C([0,T];L^2(\Omega))$, $\beta(u)(\cdot,0) = v(\cdot,0) = \beta(u^{\text{ini}})$, and $\partial_t \beta(u) = \partial_t v \in L^{p'}(0,T;W^{-1,p'}(\Omega))$. The definition (2.9) of a_0 shows that

$$a_0(x,v,\nabla v) = a(x,\nu(\beta^{-1}(v)), (\zeta \circ \beta^{-1})'(v)\nabla v) = a(x,\nu(u),\nabla\zeta(u))$$

and thus the integral equation in (2.5) follows from writing the equation (2.10) against test functions in $L^p(0,T;W_0^{1,p}(\Omega))$. Finally, since $B \circ \beta$ grows quadratically (see (4.1d) below) and $u = \beta^{-1}(v) \in C([0,T];L^2(\Omega))$, we have $B(\beta(u)) \in L^\infty(0,T;L^1(\Omega))$. Thus u is a solution to (P). □

3. A maximal monotone operator viewpoint

This section demonstrates that our setting covers problems defined by sublinear maximal monotone operators. We begin with a lemma.

Lemma 3.1 *(Maximal monotone operator). Let $\mathcal{T} : \mathbb{R} \to \mathcal{P}(\mathbb{R})$ be a multi-valued operator. Then the following are equivalent:*

- (i) *$\mathcal{T}$ is a maximal monotone operator with domain $\mathbb{R}$, $0 \in \mathcal{T}(0)$ and $\mathcal{T}$ is sublinear in the sense that there exist $T_1, T_2 \geq 0$ such that, for all $x \in \mathbb{R}$ and all $y \in \mathcal{T}(x)$, $|y| \leq T_1|x| + T_2$;*
- (ii) *There exist ζ and β satisfying (2.1b) and (2.1a) such that the graph of $\mathcal{T}$ is given by* $\text{Gr}(\mathcal{T}) = \{(\zeta(s),\beta(s)), s \in \mathbb{R}\}$.

Proof. (ii) ⇒ (i). Clearly $0 = (\zeta(0),\beta(0)) \in \mathcal{T}(0)$. The monotonicity of $\mathcal{T}$ follows from the fact that ζ and β are nondecreasing. We prove that $\mathcal{T}$ is maximal, that is if x, y satisfy $(\zeta(s) - x)(\beta(s) - y) \geq 0$ for all $s \in \mathbb{R}$ then $(x,y) \in \text{Gr}(\mathcal{T})$. By (2.1a) and (2.1b) the mapping $\beta + \zeta : \mathbb{R} \to \mathbb{R}$ is surjective, so there exists $w \in \mathbb{R}$ such that

$$\beta(w) + \zeta(w) = x + y. \tag{3.1}$$

Then $\zeta(w) - x = y - \beta(w)$ and therefore $0 \le (\zeta(w) - x)(\beta(w) - y) = -(\beta(w) - y)^2$. This implies $\beta(w) = y$ and, combined with (3.1), $\zeta(w) = x$. Hence $(x, y) \in \mathrm{Gr}(\mathcal{T})$. The sub-linearity of T follows from $|\beta(w)| \le L_\beta |w| \le L_\beta(|\zeta(w)| + M_2)/M_1$.

(i) ⇒ (ii). Recall that the resolvent $\mathcal{R}(\mathcal{T}) = (\mathrm{Id} + \mathcal{T})^{-1}$ of the maximal monotone operator $\mathcal{T}$ is a single-valued function $\mathbb{R} \to \mathbb{R}$ that is nondecreasing and Lipschitz continuous with Lipschitz constant 1. Set $\zeta = \mathcal{R}(\mathcal{T})$ and $\beta = \mathrm{Id} - \zeta$. These functions are nondecreasing and Lipschitz continuous with constant 1. By definition of the resolvent,

$$(x, y) \in \mathrm{Gr}(\mathcal{T}) \Leftrightarrow (x, x + y) \in \mathrm{Gr}(\mathrm{Id} + \mathcal{T}) \Leftrightarrow (x + y, x) \in \mathrm{Gr}(\zeta) \Leftrightarrow x = \zeta(x + y).$$

Since $\beta = \mathrm{Id} - \zeta$, setting $s = x + y$ shows that $(x, y) \in \mathrm{Gr}(\mathcal{T})$ is equivalent to $(x, y) = (\zeta(s), \beta(s))$. Since $0 \in T(0)$ this gives $\beta(0) = \zeta(0) = 0$. Finally, the existence of M_1 and M_2 in (2.1b) follows from the sublinearity of $\mathcal{T}$. If $(x, y) \in \mathrm{Gr}(\mathcal{T})$ then $|y| \le T_1|x| + T_2$ and $x = \zeta(x + y)$, which gives $|x + y| \le ((1 + T_1)|\zeta(x + y)| + T_2)$. □

Using this lemma, we recast (P) as

$$\begin{cases} \partial_t \mathcal{T}(z) - \operatorname{div}(a(x, \nu(z + \mathcal{T}(z)), \nabla z)) = f & \text{in } \Omega \times (0, T), \\ \mathcal{T}(z)(\cdot, 0) = b^{\text{ini}} & \text{in } \Omega, \\ z = 0 & \text{on } \partial\Omega \times (0, T). \end{cases} \tag{PM}$$

Hypotheses (2.1a) and (2.1b) translate into

$$\begin{gathered} \mathcal{T} \text{ is a maximal monotone operator with domain } \mathbb{R},\ 0 \in \mathcal{T}(0) \\ \text{and } \mathcal{T} \text{ is sublinear in the sense that there exist } T_1, T_2 \ge 0 \text{ such that,} \\ \text{for all } x \in \mathbb{R} \text{ and all } y \in \mathcal{T}(x),\ |y| \le T_1|x| + T_2. \end{gathered} \tag{3.2}$$

Hypothesis (2.2) becomes

$$\begin{cases} \text{(I)} & p \ge 2, \\ \text{or} & \\ \text{(II)} & \dfrac{2d}{d+2} < p < 2 \text{ and there are positive constants } T_3, T_4 \text{ such that} \\ & \text{for all } (x, y) \in \mathrm{Gr}(\mathcal{T}),\ |y| \ge T_3|x| - T_4, \\ \text{or} & \\ \text{(III)} & 1 < p \le \dfrac{2d}{d+2},\ \text{there are positive constants } T_3, T_4 \text{ such that} \\ & \text{for all } (x, y) \in \mathrm{Gr}(\mathcal{T}),\ |y| \ge T_3|x| - T_4, \text{ and } \mathcal{T} \text{ is strictly monotone.} \end{cases} \tag{3.3}$$

In (PM), ν is defined as the anti-derivative of $\zeta'\beta'$, where $\zeta = \mathcal{R}(\mathcal{T})$ and $\beta = \mathrm{Id} - \zeta$. The reciprocal $\mathcal{T}^{-1}$ of $\mathcal{T}$ is itself a maximal monotone operator, and the function $\zeta(\beta^r(s))$ in (2.4) can be computed in terms of $\mathcal{T}^{-1}$: $\zeta(\beta^r(s)) = \inf \mathcal{T}^{-1}(s)$ if $s > 0$, $\zeta(\beta^r(0)) = 0$, and $\zeta(\beta^r(s)) = \sup \mathcal{T}^{-1}(s)$ if $s < 0$. We then see that, for all s in the domain of $\mathcal{T}^{-1}$, $\mathcal{T}^{-1}(s)$ is the convex sub-differential $\partial B(s)$ of B at s.

Definition 3.2. Under Hypotheses (3.2) and (2.1d)–(2.1g), take a measurable function b^{ini} satisfying $b^{\mathrm{ini}}(x) \in \mathcal{T}(u^{\mathrm{ini}}(x))$ for a.e. $x \in \Omega$. A *solution* to (PM) is a pair of functions (z, b) satisfying

$$\begin{cases} z \in L^p(0,T;W_0^{1,p}(\Omega)),\ b(x,t) \in \mathcal{T}(z(x,t)) \text{ for a.e. } (x,t) \in \Omega \times (0,T), \\ B(b) \in L^\infty(0,T;L^1(\Omega)), \quad b \in C([0,T];L^2(\Omega)\text{-w}), \\ \partial_t b \in L^{p'}(0,T;W^{-1,p'}(\Omega)), \quad b(\cdot,0) = b^{\mathrm{ini}} \text{ in } L^2(\Omega), \\ \displaystyle\int_0^T \langle \partial_t b(\cdot,t), v(\cdot,t)\rangle_{W^{-1,p'},W_0^{1,p}}\,\mathrm{d}t \\ \qquad \displaystyle + \int_0^T \int_\Omega a(x,\nu((b+z)(x,t)),\nabla z(x,t)) \cdot \nabla v(x,t)\,\mathrm{d}x\,\mathrm{d}t \\ \quad \displaystyle = \int_0^T \langle f(\cdot,t), v(\cdot,t)\rangle_{W^{-1,p'},W_0^{1,p}}\,\mathrm{d}t \quad \forall v \in L^p(0,T;W_0^{1,p}(\Omega)). \end{cases}$$

Remark 3.3. The sublinearity of $\mathcal{T}$ ensures that $b^{\mathrm{ini}} \in L^2(\Omega)$ since $u^{\mathrm{ini}} \in L^2(\Omega)$.

The following theorem is an immediate consequence of Theorem 2.3 and Corollary 2.8. We simply take $u = b + z$, which implies $\beta(u) = b$ and $\zeta(u) = z$ since $(z, b) \in \mathrm{Gr}(\mathcal{T})$.

Theorem 3.4. *Under Hypotheses (2.1d)–(2.1g), (3.2) and (3.3), (PM) has at least one solution.*

Moreover, let $(\mathcal{T}_n, a_n, f_n, u_n^{\mathrm{ini}})_{n\in\mathbb{N}}$ be a sequence that converges to $(\mathcal{T}, a, f, u^{\mathrm{ini}})$ in the following sense:

$$\begin{cases} \mathcal{R}(\mathcal{T}_n) \textit{ and } \nu_n \textit{ converge locally uniformly on } \mathbb{R} \textit{ to } \mathcal{R}(\mathcal{T}) \textit{ and } \nu \textit{ respectively;} \\ \textit{for almost every } x \in \Omega,\ a_n(x,\cdot,\cdot) \to a(x,\cdot,\cdot) \textit{ locally uniformly on } \mathbb{R}\times\mathbb{R}^d; \\ f_n \to f \textit{ in } L^{p'}(0,T;W^{-1,p'}(\Omega)) \textit{ and } u_n^{\mathrm{ini}} \to u^{\mathrm{ini}} \textit{ in } L^2(\Omega). \end{cases}$$

Assume that $(\mathcal{T}, a, f, u^{\mathrm{ini}})$ and $(\mathcal{T}_n, a_n, f_n, u_n^{\mathrm{ini}})$ (for every $n \in \mathbb{N}$) satisfy (3.2) and (2.1d)–(2.1g), and that the constants T_1, T_2, T_3, T_4, $\underline{a}$, μ and the function $\overline{a}$ are independent of n. Let (z_n, b_n) be a solution of (PM) with $(\mathcal{T}, a, f, u^{\mathrm{ini}})$ replaced with $(\mathcal{T}_n, a_n, f_n, u_n^{\mathrm{ini}})$. Then there is a solution (z, b) of (PM) such that, up to a subsequence,

$$\begin{cases} b_n \to b \quad \textit{in } C([0,T];L^2(\Omega)\text{-w}), \\ \nu_n(b_n + z_n) \to \nu(b+z) \quad \textit{in } C([0,T];L^2(\Omega)), \textit{ and} \\ z_n \rightharpoonup z \quad \textit{weakly in } L^p(0,T;W_0^{1,p}(\Omega)). \end{cases}$$

If in addition we assume that a is strictly monotone, that is, the inequality in (2.1f) with $\chi \neq \xi$ is strict, then $z_n \to z$ strongly in $L^p(0,T;W_0^{1,p}(\Omega))$.

Remark 3.5. Blanchard and Porretta [8] prove the $L^\infty(0,T;L^1(\Omega))$-stability of renormalised solutions to

$$\begin{cases} \partial_t \mathcal{T}(u) - \operatorname{div}(a(x,u,\nabla u)) + \operatorname{div}(\Phi(u)) = f & \text{in } \Omega \times (0,T), \\ \mathcal{T}(u)(x,0) = b_0(x) & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega \times (0,T), \end{cases}$$

with $f \in L^1(\Omega \times (0,T))$ and $b_0 \in L^1(\Omega)$. They assume that $\mathcal{T}$ is a maximal monotone graph on $\mathbb{R}$, $\mathcal{T}^{-1} \in C(\mathbb{R})$ and a is a Leray–Lions operator. Although the continuity assumption on $\mathcal{T}^{-1}$ – which prevents $\mathcal{T}$ from having plateaux – is not required for the stability result, it is necessary in their existence theorem for identifying u as the pointwise limit of solutions to regularised problems using compactness arguments. If $p > \frac{2d}{d+2}$ we overcome this assumption on $\mathcal{T}$ in the variational setting by using monotonicity and compensated compactness arguments; see Section 5.2. Indeed, it may be interesting to determine whether similar arguments may be used in the setting of renormalised solutions in [8]. If p is 'too small' — that is, in case (III) of Hypothesis (3.3) — we must also assume that $\mathcal{T}$ (respectively β outside the present section) does not have any plateaux, but we still identify weak limits by monotonicity and compensated compactness arguments rather than by pointwise convergence.

4. Preliminaries

4.1. Properties of B

We recall here two lemmas proved in [14]. Lemma 4.1 states some properties of the functions ν and B. The uniform convexity property (4.1e) plays a critical role in our proof of the uniform temporal convergence of $\nu_n(u_n)$. Lemma 4.2 brings together two identities — an integration-by-parts formula and an energy equality — and some continuity properties of the solution. Although the integration-by-parts formula (4.2) apparently follows from the formal relation $\zeta(u)\partial_t(\beta(u)) = \zeta(u)\beta'(u)\partial_t u = (B \circ \beta)'(u)\partial_t u = \partial_t(B(\beta(u))$, its rigorous justification is quite technical, owing to the lack of regularity of u.

Lemma 4.1. *Assume (2.1). Then for every $a, b \in \mathbb{R}$,*

$$|\nu(a) - \nu(b)| \le L_\beta |\zeta(a) - \zeta(b)|, \quad \textit{and} \tag{4.1a}$$

$$[\nu(a) - \nu(b)]^2 \le L_\beta L_\zeta [\zeta(a) - \zeta(b)][\beta(a) - \beta(b)]. \tag{4.1b}$$

The functions $B : \overline{R_\beta} \to [0,\infty]$ and $B \circ \beta : \mathbb{R} \to [0,\infty)$ are continuous, and for all $s \in \mathbb{R}$,

$$B(\beta(s)) = \int_0^s \zeta(q)\beta'(q)\,\mathrm{d}q. \tag{4.1c}$$

There are positive constants K_1, K_2 and K_3, depending only upon L_β, L_ζ and the constants M_1, M_2 in (2.1b), such that for all $s \in \mathbb{R}$,

$$K_1 \beta(s)^2 - K_2 \le B(\beta(s)) \le K_3 s^2. \tag{4.1d}$$

Finally, for every $a, b \in \mathbb{R}$,

$$[\nu(a) - \nu(b)]^2 \le 4 L_\beta L_\zeta \left[B(\beta(a)) + B(\beta(b)) - 2B\left(\frac{\beta(a)+\beta(b)}{2}\right)\right]. \tag{4.1e}$$

Before stating the next lemma, a few remarks on notation are necessary. The discussion following Definition 2.2 concerning the continuity of $\beta(u)$ neglects a subtlety that one must account for in order to give meaning to the convergences (2.7). Indeed, by the statement $\beta(u) \in C([0,T]; L^2(\Omega)\text{-w})$, one understands that the mapping $(x,t) \mapsto \beta(u(x,t))$ is equal almost everywhere on $\Omega \times (0,T)$ to a function Z that is continuous as a map from $[0,T]$ to $L^2(\Omega)$-w. We henceforth write $\overline{\beta(u)}$ for Z; similarly $\overline{\nu(u)}$ for the continuous in time almost-everywhere representative of $(x,t) \mapsto \nu(u(x,t))$ (see part (ii) of the following lemma). This distinction is essential in the present context, where we are frequently concerned with the values of these functions at a particular point in time. The composition $\beta(u(\cdot,\cdot))$ is only defined up to null sets in $\Omega \times (0,T)$, so for a particular $t \in [0,T]$ the expression $\beta(u(\cdot,t))$ is ill-defined. The expression $\overline{\beta(u)}(\cdot,t)$ is, however, well-defined, and we take care to use the notation $\beta(u)$ (without the bar) only when this quantity is used in an average sense. Nonetheless, for the sake of clarity Theorem 2.3 is stated without this distinction.

Lemma 4.2. *Let (2.1) hold.*

(i) *If v is a measurable function on $\Omega \times (0,T)$ 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 the mapping $[0,T] \ni t \mapsto \int_\Omega B(\overline{\beta(v)}(x,t))\,\mathrm{d}x \in [0,\infty)$ is continuous and bounded, and for all $T_0 \in [0,T]$,*

$$\int_0^{T_0} \langle \partial_t\beta(v)(\cdot,t), \zeta(v(\cdot,t))\rangle_{W^{-1,p'},W_0^{1,p}}\,\mathrm{d}t \\ = \int_\Omega B(\overline{\beta(v)}(x,T_0))\,\mathrm{d}x - \int_\Omega B(\overline{\beta(v)}(x,0))\,\mathrm{d}x. \tag{4.2}$$

(ii) *If u is a solution to (P) then for all $T_0 \in [0,T]$,*

$$\int_\Omega B(\overline{\beta(u)}(x,T_0))\,\mathrm{d}x + \int_0^{T_0}\int_\Omega a(x,\nu(u),\nabla\zeta(u))\cdot\nabla\zeta(u)\,\mathrm{d}x\,\mathrm{d}t \\ = \int_\Omega B(\beta(u^{\mathrm{ini}}(x)))\,\mathrm{d}x + \int_0^{T_0}\langle f(\cdot,t), \zeta(u)(\cdot,t)\rangle_{W^{-1,p'},W_0^{1,p}}\,\mathrm{d}t \tag{4.3}$$

and the function $\nu(u)$ is continuous from $[0,T]$ into $L^2(\Omega)$.

Since B plays such a critical role to our main result, we highlight its stability properties in the following lemma.

Lemma 4.3. *Assume (2.1) and (2.6). Define $B_n : \overline{R_{\beta_n}} \to [0,\infty]$ from ζ_n, β_n analogously to (2.4), and extend B_n to $\mathbb{R}$ by setting $B_n = +\infty$ outside $\overline{R_{\beta_n}}$. Then*

(i) *B and all $(B_n)_{n\in\mathbb{N}}$ are convex lower semi-continuous on $\mathbb{R}$;*
(ii) *$B_n \circ \beta_n \to B \circ \beta$ locally uniformly on $\mathbb{R}$ as $n \to \infty$;*
(iii) *For any $z \in \mathbb{R}$ and any sequence $(z_n)_{n\in\mathbb{N}}$ that converges to z, $B(z) \le \liminf_{n\to\infty} B_n(z_n)$.*

Proof. (i) The convexity has already been noted. Since B and B_n are continuous on $\overline{R_\beta}$ and $\overline{R_{\beta_n}}$ respectively by Lemma 4.1, their extension by $+\infty$ outside their initial domain ensures their lower semi-continuity.

(ii) Let $M > 0$. By (4.1c) applied to B_n, $B_n(\beta_n)$ is Lipschitz continuous on $[-M, M]$ with Lipschitz constant $L_\beta \sup_{|s|\le M} |\zeta_n(s)|$. This quantity is bounded with respect to n since $(\zeta_n)_{n\in\mathbb{N}}$ converges uniformly on $[-M, M]$. Hence, the local uniform convergence of $(B_n(\beta_n))_{n\in\mathbb{N}}$ follows from the Arzelà–Ascoli theorem if we can prove that $B_n(\beta_n) \to B(\beta)$ pointwise. The reasoning in Remark 2.6 shows that $\beta_n' \rightharpoonup \beta'$ weak-$*$ in $L^\infty(\mathbb{R})$. Hence, for any $s \in \mathbb{R}$, since $\zeta_n \to \zeta$ uniformly on $[0, s]$,

$$B_n(\beta_n(s)) = \int_0^s \zeta_n(q)\beta_n'(q)\,\mathrm{d}q \to \int_0^s \zeta(q)\beta'(q)\,\mathrm{d}q = B(\beta(s)) \text{ as } n \to \infty,$$

and the proof of (ii) is complete.

(iii) Without loss of generality, we can assume that $(B_n(z_n))_{n\in\mathbb{N}}$ converges in $[0, \infty]$, otherwise we extract a subsequence that converges to the inferior limit. We study four distinct cases.

Case A: $z_n \notin \overline{R_{\beta_n}}$ for an infinite number of n. Then the corresponding $B_n(z_n)$ are equal to $+\infty$ and therefore $\lim_{n\to\infty} B_n(z_n) = +\infty \ge B(z)$.

Case B: $z_n \in \overline{R_{\beta_n}}$ for n large, and $z \notin \overline{R_\beta}$. Assume that $z > \sup R_\beta$ (the case $z < \inf R_\beta$ is similar). Take $Z \in (\sup R_\beta, z) \subset (0, \infty)$. For n sufficiently large, $z_n > Z$ and $z_n \in \overline{R_{\beta_n}}$. Then use the definition (2.4) of B_n, Hypothesis (2.1b) and the fact that β_n^r is nondecreasing to see that

$$\begin{aligned} B_n(z_n) &= \int_0^{z_n} \zeta_n(\beta_n^r(s))\,\mathrm{d}s \ge \int_Z^{z_n} \zeta_n(\beta_n^r(s))\,\mathrm{d}s \\ &\ge \int_Z^{z_n} (M_1\beta_n^r(s) - M_2)\,\mathrm{d}s \ge (z_n - Z)(M_1\beta_n^r(Z) - M_2). \end{aligned} \tag{4.4}$$

We prove by contradiction that $(\beta_n^r(Z))_{n\in\mathbb{N}}$ is not bounded. Otherwise, upon extraction of a subsequence it converges to some $m \in \mathbb{R}$. Then, by local uniform convergence of β_n, $Z = \beta_n(\beta_n^r(Z)) \to \beta(m) \in R_\beta$. But $Z > \sup R_\beta$, which is a contradiction. Hence, $\beta_n^r(Z) \to +\infty$ as $n \to \infty$. Since $z_n - Z \to z - Z > 0$, passing to the limit in (4.4) gives $\lim_{n\to\infty} B_n(z_n) = +\infty \ge B(z)$.

Case C: $z_n \in \overline{R_{\beta_n}}$ for n large, $z \in \overline{R_\beta}$ and $(\beta_n^r(z_n))_{n\in\mathbb{N}}$ is bounded in $\mathbb{R}$. Let $s_n = \beta_n^r(z_n)$, which gives $z_n = \beta_n(s_n)$. Since $(s_n)_{n\in\mathbb{N}}$ is bounded, up to extraction of a subsequence we have $s_n \to s \in \mathbb{R}$ and thus, by (ii), $B_n(z_n) = B_n \circ \beta_n(s_n) \to B \circ \beta(s)$. The local uniform convergence of $(\beta_n)_{n\in\mathbb{N}}$ gives $z_n = \beta_n(s_n) \to \beta(s)$, which means that $\beta(s) = z$. Hence $B_n(z_n) \to B(\beta(s)) = B(z)$ and the proof is complete.

Case D: $z_n \in \overline{R_{\beta_n}}$ for n large, $z \in \overline{R_\beta}$ and $(\beta_n^r(z_n))_{n\in\mathbb{N}}$ is unbounded. Again, let $s_n = \beta_n^r(z_n) \in [-\infty, +\infty]$. The function B_n is continuous (with values in $[0, +\infty]$) at the endpoints of R_{β_n}. Since these endpoints correspond to $\lim_{s\to\pm\infty} \beta_n(s)$, applying the monotone convergence theorem to (4.1c) then shows that this formula also holds if $s = \pm\infty$. Hence, for any n,

$$B_n(z_n) = B_n(\beta_n(s_n)) = \int_0^{s_n} \zeta_n(q)\beta_n'(q)\,\mathrm{d}q.$$

The sequence $(s_n)_{n\in\mathbb{N}}$ contains a subsequence that goes to $\pm\infty$. Say, without explicitly denoting the subsequence, that $s_n \to +\infty$ (the case $s_n \to -\infty$ is similar). Let $M \geq 0$ and for n sufficiently large, since $\zeta_n \geq 0$ on $\mathbb{R}^+$ and $\beta_n' \geq 0$, write

$$B_n(z_n) = \int_0^{s_n} \zeta_n(q)\beta_n'(q)\,\mathrm{d}q \geq \int_{\mathbb{R}} \mathbf{1}_{[0,M]}(q)\zeta_n(q)\beta_n'(q)\,\mathrm{d}q.$$

By the reasoning in Remark 2.6, $\beta_n' \to \beta'$ in $L^\infty(\mathbb{R})$ weak-$*$. Since $\zeta_n \to \zeta$ uniformly on $[0, M]$, we can conclude that

$$\varliminf_{n\to\infty} B_n(z_n) \geq \int_{\mathbb{R}} \mathbf{1}_{[0,M]}(q)\zeta(q)\beta'(q)\,\mathrm{d}q.$$

Take the limit inferior as $M \to \infty$ using Fatou's lemma to deduce that

$$\varliminf_{n\to\infty} B_n(z_n) \geq \int_0^{\infty} \zeta(q)\beta'(q)\,\mathrm{d}q.$$

Since $z \geq 0$ (because for n large enough, each $z_n = \beta_n(s_n)$ is nonnegative), $s = \beta^r(z) \in [0, \infty]$ and thus

$$\varliminf_{n\to\infty} B_n(z_n) \geq \int_0^{s} \zeta(q)\beta'(q)\,\mathrm{d}q.$$

We already saw that (4.1c) is valid for any $s \in [-\infty, \infty]$, and we infer that $\varliminf_{n\to\infty} B_n(z_n) \geq B(\beta(s)) = B(z)$ as required. □

4.2. Estimates

The results of the previous section enable energy estimates, the subject of our next lemma. Note that none of the estimates we prove in this section require Hypothesis (2.2).

Lemma 4.4. *Let $(\beta_n, \zeta_n, \nu_n, a_n, f_n, u_n^{\mathrm{ini}})_{n\in\mathbb{N}}$ be a sequence of data that satisfies the hypotheses of Theorem 2.3, and let u_n be a solution to (P_n). Then there exists $C_1 > 0$ independent of n such that the following quantities are bounded above by C_1:*

$$\begin{aligned}
&\sup_{t\in[0,T]} \left\| B_n(\overline{\beta_n(u_n)}(\cdot,t)) \right\|_{L^1(\Omega)}, \qquad \|\zeta_n(u_n)\|_{L^p(0,T;W_0^{1,p}(\Omega))},\\
&\sup_{t\in[0,T]} \left\| \overline{\beta_n(u_n)}(\cdot,t) \right\|_{L^2(\Omega)}, \qquad \|\partial_t\beta_n(u_n)\|_{L^{p'}(0,T;W^{-1,p'}(\Omega))}.
\end{aligned} \tag{4.5}$$

Proof. By hypothesis, $(\beta_n, \zeta_n, \nu_n, a_n, f_n, u_n^{\mathrm{ini}})$ (for every $n \in \mathbb{N}$) satisfies an identity analogous to (4.3). From this identity, the quadratic growth (4.1d) of $B_n \circ \beta_n$, the uniform coercivity of $(a_n)_{n\in\mathbb{N}}$ and Young's inequality,

$$\int_\Omega B_n(\overline{\beta_n(u_n)}(x,T_0))\,\mathrm{d}x + \underline{a}\int_0^{T_0}\int_\Omega |\nabla\zeta_n(u_n)|^p\,\mathrm{d}x\,\mathrm{d}t$$
$$\le K_3\left\|u_n^{\mathrm{ini}}\right\|_{L^2(\Omega)}^2 + \frac{\underline{a}}{2}\|\nabla\zeta_n(u_n)\|_{L^p(0,T_0;L^p(\Omega)^d)}^p + \frac{1}{p'}\left(\frac{2}{\underline{a}p}\right)^{p'/p}\|f_n\|_{L^{p'}(0,T_0;W^{-1,p'}(\Omega))}^{p'}. \tag{4.6}$$

Taking $T_0 = T$ shows that

$$\|\nabla\zeta_n(u_n)\|_{L^p(0,T;L^p(\Omega)^d)}^p \le \frac{2}{\underline{a}}\left(K_3\left\|u_n^{\mathrm{ini}}\right\|_{L^2(\Omega)}^2 + \frac{1}{p'}\left(\frac{2}{\underline{a}p}\right)^{p'/p}\|f_n\|_{L^{p'}(0,T;W^{-1,p'}(\Omega))}^{p'}\right).$$

With the assumed convergence properties of $(u_n^{\mathrm{ini}})_{n\in\mathbb{N}}$ and $(f_n)_{n\in\mathbb{N}}$, substituting the previous inequality into (4.6) gives the first two estimates in (4.5). The estimate on $(\overline{\beta_n(u_n)})_{n\in\mathbb{N}}$ follows from that on $(B_n(\overline{\beta_n(u_n)}))_{n\in\mathbb{N}}$ and (4.1d). To prove the estimate on $\partial_t\beta_n(u_n)$, let $v \in L^p(0,T;W_0^{1,p}(\Omega))$ and deduce from (2.5) that

$$\left|\int_0^T \langle\partial_t\beta_n(u_n)(\cdot,t), v(\cdot,t)\rangle_{W^{-1,p'},W_0^{1,p}}\,\mathrm{d}t\right|$$
$$\le \|v\|_{L^p(0,T;W_0^{1,p}(\Omega))}\left(\|\overline{a}\|_{L^{p'}(\Omega)} + \mu\,\|\nabla\zeta_n(u_n)\|_{L^p(0,T;L^p(\Omega)^d)}^{p-1} + \|f_n\|_{L^{p'}(0,T;W^{-1,p'}(\Omega))}\right).$$

Take the supremum over v in the unit ball of $L^p(0,T;W_0^{1,p}(\Omega))$ and use the bound on $(\zeta_n(u_n))_{n\in\mathbb{N}}$ in $L^p(0,T;W_0^{1,p}(\Omega))$ to complete the proof. □

The following lemma, applied to $F_n = \zeta_n$, $G_n = \beta_n$ and u_n the solution to (P_n), provides us with crucial estimates of the time translates of $\nu_n(u_n)$. Nevertheless we state it in a generic setting, as it will also be applied with different functions.

Lemma 4.5. *For every $n \in \mathbb{N}$, let $F_n : \mathbb{R} \to \mathbb{R}$ and $G_n : \mathbb{R} \to \mathbb{R}$ be nondecreasing and Lipschitz continuous, uniformly with respect to n. Suppose also that $F_n(0) = 0$. Define $H_n(s) := \int_0^s F_n'(q)G_n'(q)\,\mathrm{d}q$. Take $p \ge 1$ and $(u_n)_{n\in\mathbb{N}}$ a sequence of measurable functions on $\Omega\times(0,T)$ such that $(F_n(u_n))_{n\in\mathbb{N}}$ is bounded in $L^p(0,T;W_0^{1,p}(\Omega))$, $(G_n(u_n))_{n\in\mathbb{N}}$ is bounded in $L^\infty(0,T;L^2(\Omega))$ and $(\partial_t(G_n(u_n)))_{n\in\mathbb{N}}$ is bounded in $L^{p'}(0,T;W^{-1,p'}(\Omega))$. Then there exists $C_2 > 0$ independent of n such that, for all $2 \le r < \infty$ and all $\tau > 0$,*

$$\|H_n(u_n)(\cdot,\cdot+\tau) - H_n(u_n)\|_{L^r(\mathbb{R};L^2(\Omega)))} \le C_2\tau^{1/r}, \tag{4.7}$$

where $H_n(u_n)$ is extended by zero outside $\Omega\times(0,T)$.

Proof. Denote by L_F and L_G the uniform Lipschitz constants of $(F_n)_{n\in\mathbb{N}}$ and $(G_n)_{n\in\mathbb{N}}$, respectively. We introduce the truncations $\mathbb{T}_k : \mathbb{R} \to \mathbb{R}$ at level $k > 0$, defined by $\mathbb{T}_k(s) := \max(-k, \min(s,k))$, and the functions

$$F_n^k(s) := \mathbb{T}_k(F_n(s)) \quad\text{and}\quad H_n^k(s) := \int_0^s (F_n^k)'(q)G_n'(q)\,\mathrm{d}q.$$

Then $F_n^k(u_n), H_n^k(u_n) \in L^2(0,T;L^2(\Omega))$, the latter coming from

$$|H_n^k(u_n)| \le L_F |G_n(u_n)|. \tag{4.8}$$

Now let $\tau \in (0,T)$. Inequality (4.1b) with (F_n^k, G_n, H_n^k) in place of (ζ, β, ν) yields

$$\begin{aligned}
&\int_0^{T-\tau}\int_\Omega \Big(H_n^k(u_n)(x,t+\tau) - H_n^k(u_n)(x,t)\Big)^2 \,\mathrm{d}x\,\mathrm{d}t \\
&\quad \le L_F L_G \int_0^{T-\tau}\int_\Omega (G_n(u_n)(x,t+\tau) - G_n(u_n)(x,t)) \\
&\qquad \times \Big(F_n^k(u_n)(x,t+\tau) - F_n^k(u_n)(x,t)\Big)\,\mathrm{d}x\,\mathrm{d}t \\
&\quad = L_F L_G \int_0^{T-\tau} \Big\langle G_n(u_n)(\cdot,t+\tau) - G_n(u_n)(\cdot,t), F_n^k(u_n)(\cdot,t+\tau) \\
&\qquad - F_n^k(u_n)(\cdot,t)\Big\rangle_{W^{-1,p'},W_0^{1,p}} \mathrm{d}t \\
&\quad = L_F L_G \int_0^{T-\tau} \Big\langle \int_t^{t+\tau} \partial_t G_n(u_n)(\cdot,s)\,\mathrm{d}s, F_n^k(u_n)(\cdot,t+\tau) - F_n^k(u_n)(\cdot,t)\Big\rangle_{W^{-1,p'},W_0^{1,p}} \mathrm{d}t \\
&\quad = L_F L_G \int_0^{T-\tau}\int_t^{t+\tau} \Big\langle \partial_t G_n(u_n)(\cdot,s), F_n^k(u_n)(\cdot,t+\tau) - F_n^k(u_n)(\cdot,t)\Big\rangle_{W^{-1,p'},W_0^{1,p}} \mathrm{d}s\,\mathrm{d}t,
\end{aligned}$$

where the first equality holds since $G_n(u_n)(\cdot,t) \in L^2(\Omega) \cap W^{-1,p'}(\Omega)$ and $F_n^k(u_n)(\cdot,t) \in L^2(\Omega) \cap W_0^{1,p}(\Omega)$ for a.e. $t \in (0,T)$. Note that obtaining this L^2 integrability of $F_n^k(u_n)$ is the only reason for introducing the truncations; if $p \ge 2$ then the truncations are redundant. As $k \to \infty$, $H_n^k(u_n) \to H_n(u_n)$ almost everywhere on $\Omega \times (0,T)$ and therefore also in $L^2(0,T;L^2(\Omega))$ by dominated convergence with (4.8). Thanks to G. Stampacchia's important result [31], we can write $\nabla F_n^k(u_n) = \mathbb{T}_k'(F_n(u_n))\nabla F_n(u_n) = \mathbf{1}_{\{|F_n(u_n)|\le k\}}\nabla F_n(u_n)$, which converges in $L^p(0,T;L^p(\Omega)^d)$ to $\nabla F_n(u_n)$ as $k \to \infty$. So $F_n^k(u_n) \to F_n(u_n)$ in $L^p(0,T;W_0^{1,p}(\Omega))$. Let $k \to \infty$ on both sides of the above inequality to obtain

$$\begin{aligned}
&\int_0^{T-\tau}\int_\Omega (H_n(u_n)(x,t+\tau) - H_n(u_n)(x,t))^2 \,\mathrm{d}x\,\mathrm{d}t \\
&\quad \le L_F L_G \int_0^{T-\tau}\int_t^{t+\tau} \langle \partial_t G_n(u_n)(\cdot,s), F_n(u_n)(\cdot,t+\tau) - F_n(u_n)(\cdot,t)\rangle_{W^{-1,p'},W_0^{1,p}} \mathrm{d}s\,\mathrm{d}t \\
&\quad \le L_F L_G \int_0^{T-\tau}\int_t^{t+\tau} \|\partial_t G_n(u_n)(\cdot,s)\|_{W^{-1,p'}(\Omega)} \|F_n(u_n)(\cdot,t+\tau) - F_n(u_n)(\cdot,t)\|_{W_0^{1,p}} \mathrm{d}s\,\mathrm{d}t
\end{aligned}$$

Apply Young's inequality and interchange the order of integration in s and t where appropriate to obtain

$$\begin{aligned}&\|H_n(u_n)(\cdot,\cdot+\tau)-H_n(u_n)(\cdot,\cdot)\|^2_{L^2(\Omega\times(0,T-\tau))}\\&\quad\le L_F L_G\tau\left(\frac{1}{p'}\,\|\partial_t G_n(u_n)\|^{p'}_{L^{p'}(0,T;W^{-1,p'}(\Omega))}+\frac{2}{p}\,\|F_n(u_n)\|^{p}_{L^{p}(0,T;W^{1,p}_0(\Omega))}\right)\le C_3\tau\end{aligned}\tag{4.9}$$

where C_3 does not depend on n or τ. From the definition of H_n we have $|H_n(u_n)|\le L_F|G_n(u_n)|$. Hence, $(H_n(u_n))_{n\in\mathbb{N}}$ is bounded in $L^\infty(0,T;L^2(\Omega))$. This enables us to estimate the time translates on $(0,\tau)$ and $(T-\tau,T)$ and, combined with (4.9) we deduce that (4.7) holds for $r=2$. The conclusion for a generic $r\in[2,\infty)$ follows by interpolation (Hölder's inequality), using the bound of $(H_n(u_n))_{n\in\mathbb{N}}$ in $L^\infty(0,T;L^2(\Omega))$. □

4.3. Two lemmas: convexity and monotonicity

Lemma 4.6 is a general result on the uniform weak lower semi-continuity of sequences of convex functions.

Lemma 4.6. *Let $\Psi,\Psi_n:\mathbb{R}\to[0,\infty]$ be convex lower semi-continuous functions such that for every $n\in\mathbb{N}$, $\Psi_n(0)=\Psi(0)=0$. Assume that for any $z\in\mathbb{R}$ and any sequence $(z_n)_{n\in\mathbb{N}}$ converging to z, $\Psi(z)\le\liminf_{n\to\infty}\Psi_n(z_n)$. If $(v_n)_{n\in\mathbb{N}}\subset L^2(\Omega)$ converges weakly to v in $L^2(\Omega)$ then*

$$\int_\Omega \Psi(v(x))\,\mathrm{d}x\le\liminf_{n\to\infty}\int_\Omega\Psi_n(v_n(x))\,\mathrm{d}x.$$

Proof. For $x\in\Omega$, $r>0$ and $n\in\mathbb{N}$, extend v_n by zero outside Ω and write

$$[v_n]_r(x):=\fint_{B(x,r)} v_n(y)\,\mathrm{d}y=\frac{1}{|B(x,r)|}\int_{B(x,r)}v_n(y)\,\mathrm{d}y$$

for the mean value of v_n over the closed ball of radius r centred at x. Since $v_n\rightharpoonup v$ in $L^2(\Omega)$ as $n\to\infty$, for every $x\in\Omega$,

$$[v_n]_r(x)\to\fint_{B(x,r)}v(y)\,\mathrm{d}y=:[v]_r(x).$$

We have extended v by 0 outside Ω. Hence, $\forall x\in\Omega$, $\Psi([v]_r(x))\le\liminf_{n\to\infty}\Psi_n([v_n]_r(x))$. We can apply Fatou's lemma and Jensen's inequality to obtain

$$\int_\Omega\Psi([v]_r(x))\,\mathrm{d}x\le\liminf_{n\to\infty}\int_\Omega\Psi_n([v_n]_r(x))\,\mathrm{d}x\le\liminf_{n\to\infty}\int_\Omega\fint_{B(x,r)}\Psi_n(v_n(y))\,\mathrm{d}y\,\mathrm{d}x.$$

Use Fubini–Tonelli and the fact that $\Psi_n(v_n)=0$ outside Ω to write

$$\int_\Omega\fint_{B(x,r)}\Psi_n(v_n(y))\,\mathrm{d}y\,\mathrm{d}x=\int_\Omega\Psi_n(v_n(y))\,\mathrm{d}y.$$

Thus

$$\int_\Omega \Psi([v]_r(x))\,\mathrm{d}x \le \liminf_{n\to\infty}\int_\Omega \Psi_n(v_n(y))\,\mathrm{d}y. \tag{4.10}$$

Almost every $x \in \Omega$ is a Lebesgue point of v and for those we have $\lim_{r\to 0}[v]_r(x) = v(x)$. Then from the lower semi-continuity of Ψ, another application of Fatou's lemma and (4.10), we deduce that

$$\begin{aligned}\int_\Omega \Psi(v(x))\,\mathrm{d}x &\le \int_\Omega \liminf_{r\to 0}\Psi([v]_r(x))\,\mathrm{d}x \le \liminf_{r\to 0}\int_\Omega \Psi([v]_r(x))\,\mathrm{d}x\\ &\le \liminf_{r\to 0}\liminf_{n\to\infty}\int_\Omega \Psi_n(v_n(y))\,\mathrm{d}y = \liminf_{n\to\infty}\int_\Omega \Psi_n(v_n(x))\,\mathrm{d}x. \qquad \square\end{aligned}$$

We employ the next result to identify weak nonlinear limits in Section 5.2.

Lemma 4.7. *Let V be a measurable subset of $\mathbb{R}^d$. Take sequences $(\chi_n)_{n\in\mathbb{N}}$, $(\psi_n)_{n\in\mathbb{N}} \subset C(\mathbb{R})$ of nondecreasing functions satisfying $\chi_n(0) = \psi_n(0) = 0$ for every $n \in \mathbb{N}$ and such that $\chi_n \to \chi$ and $\psi_n \to \psi$, both pointwise on $\mathbb{R}$. Assume there is a sequence $(v_n)_{n\in\mathbb{N}}$ of measurable functions on V and two functions $\widetilde{\chi}$, $\widetilde{\psi} \in L^2(V)$ such that*

(i) *$\chi_n(v_n) \rightharpoonup \widetilde{\chi}$ and $\psi_n(v_n) \rightharpoonup \widetilde{\psi}$, both weakly in $L^2(V)$;*
(ii) *there exists an almost-everywhere strictly positive function $\varphi \in L^\infty(V)$ such that*

$$\lim_{n\to\infty}\int_V \varphi(z)\chi_n(v_n(z))\psi_n(v_n(z))\,\mathrm{d}z = \int_V \varphi(z)\widetilde{\chi}(z)\widetilde{\psi}(z)\,\mathrm{d}z.$$

Then for all measurable functions v satisfying $(\chi + \psi)(v) = \widetilde{\chi} + \widetilde{\psi}$ almost everywhere in V,

$$\widetilde{\chi} = \chi(v) \quad \textit{and} \quad \widetilde{\psi} = \psi(v) \quad \textit{almost everywhere in } V.$$

Proof. Observe that $\chi(v), \psi(v) \in L^2(V)$ since by hypothesis they have the same sign, so that $|\chi(v)| + |\psi(v)| = |(\chi + \psi)(v)| = |\widetilde{\chi} + \widetilde{\psi}| \in L^2(V)$. Let $\mathbb{T}_k(s) = \min(k, \max(-k, s))$ be the truncation at level k. Since χ_n and ψ_n are nondecreasing, for the function φ in (ii), write

$$\int_V \varphi(z)\big[\chi_n(v_n(z)) - \chi_n(\mathbb{T}_k(v(z)))\big]\big[\psi_n(v_n(z)) - \psi_n(\mathbb{T}_k(v(z)))\big]\,\mathrm{d}z \ge 0. \tag{4.11}$$

By their monotonicity and sign properties, the functions χ_n and ψ_n are bounded on $[-k, k]$ by $\max(|\chi_n(\pm k)|, |\psi_n(\pm k)|)$, which is uniformly bounded with respect to n since χ_n and ψ_n converge pointwise. Hence, by the dominated convergence theorem, $\chi_n(\mathbb{T}_k(v)) \to \chi(\mathbb{T}_k(v))$ and $\psi_n(\mathbb{T}_k(v)) \to \psi(\mathbb{T}_k(v))$ in $L^2(V)$ as $n \to \infty$. Using (i) and (ii), we can therefore pass to the limit $n \to \infty$ in (4.11) and we find

$$\int_V \varphi(z)\big[\widetilde{\chi}(z) - \chi(\mathbb{T}_k(v(z)))\big]\big[\widetilde{\psi}(z) - \psi(\mathbb{T}_k(v(z)))\big]\,\mathrm{d}z \ge 0. \tag{4.12}$$

The monotonicity and sign properties of χ and ψ ensure that $|\chi(\mathbb{T}_k(v))| \le |\chi(v)|$ and $|\psi(\mathbb{T}_k(v))| \le |\psi(v)|$. Since $\chi(v)$ and $\psi(v)$ belong to $L^2(V)$, we deduce that as $k \to \infty$, $\chi(\mathbb{T}_k(v)) \to \chi(v)$ and $\psi(\mathbb{T}_k(v)) \to \psi(v)$, both in $L^2(V)$. Passing to the limit in (4.12), we obtain

$$\int_V \varphi(z)[\widetilde{\chi}(z) - \chi(v(z))][\widetilde{\psi}(z) - \psi(v(z))]\,\mathrm{d}z \ge 0. \tag{4.13}$$

The identity $\chi(v) + \psi(v) = \widetilde{\chi} + \widetilde{\psi}$ gives $\widetilde{\chi}(z) - \chi(v(z)) = -(\widetilde{\psi}(z) - \psi(v(z)))$, which after substitution into (4.13) yields

$$-\int_V \varphi(z)\left[\widetilde{\chi}(z) - \chi(v(z))\right]^2 \mathrm{d}z = -\int_V \varphi(z)\left[\widetilde{\psi}(z) - \psi(v(z))\right]^2 \mathrm{d}z \ge 0.$$

From the positivity of φ we conclude that $\widetilde{\chi}(z) = \chi(v(z))$ and $\widetilde{\psi}(z) = \psi(v(z))$ for almost every $z \in V$. □

4.4. *Uniform-temporal, weak-spatial compactness*

The classical Aubin–Simon compactness theorem — an amalgamation of the work of J.-P. Aubin [4] and J. Simon [30] — does ensure uniform temporal compactness in Lebesgue spaces (for the norm topology), provided that a spatial compactness estimate is available in such spaces. This usually requires control of the gradients in Lebesgue spaces. Since we lack such estimates on the gradient of $\beta(u)$, we must forfeit (at least initially) strong compactness in the spatial variable. We first recall a basic definition.

Definition 4.8. A sequence of continuous functions $v_n : [0,T] \to L^2(\Omega)$-w converges in the space $C([0,T]; L^2(\Omega)\text{-w})$ to a function $v : [0,T] \to L^2(\Omega)$ if for all $\varphi \in L^2(\Omega)$, the sequence of functions $[0,T] \ni t \mapsto \langle v_n(t), \varphi\rangle_{L^2(\Omega)}$ converges uniformly on $[0,T]$ to $[0,T] \ni t \mapsto \langle v(t), \varphi\rangle_{L^2(\Omega)}$ as $n \to \infty$.

Note that v is then necessarily an element of $C([0,T]; L^2(\Omega)\text{-w})$.

Proposition 4.9. *Let $(v_n)_{n\in\mathbb{N}}$ be a sequence of real-valued measurable functions on $\Omega \times (0,T)$. Suppose that there exists $q > 1$ and $R > 0$ such that for every $n \in \mathbb{N}$,*

$$\sup_{t\in[0,T]} \|v_n(\cdot,t)\|_{L^2(\Omega)} \le R, \qquad \|\partial_t v_n\|_{L^q(0,T;W^{-1,1}(\Omega))} \le R. \tag{4.14}$$

Then $(v_n)_{n\in\mathbb{N}}$ is relatively compact in $C([0,T]; L^2(\Omega)\text{-w})$; that is, there is a subsequence of $(v_n)_{n\in\mathbb{N}}$ that converges in the sense of Definition 4.8.

Remark 4.10. The space $W^{-1,1}(\Omega)$ has been chosen by convenience, but it could be replaced with the dual space of any Banach space in which $C_c^\infty(\Omega)$ is dense.

Proof. Denote by E the ball of radius R in $L^2(\Omega)$, endowed with the weak topology. Take $(\varphi_l)_{l\in\mathbb{N}} \subset C_c^\infty(\Omega)$ a dense sequence in $L^2(\Omega)$ and equip E with the metric

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

The $L^2(\Omega)$ weak topology on E is the topology induced by this metric. The set E is metric compact and therefore complete. The first bound in (4.14) ensures that every v_n takes values in E. It remains to estimate $d_E(v_n(s), v_n(s'))$. To this end,

$$\begin{aligned}
\left|\langle v_n(s') - v_n(s), \varphi_l\rangle_{L^2(\Omega)}\right| &= \left|\int_\Omega (v_n(x,s') - v_n(x,s))\varphi_l(x)\,\mathrm{d}x\right| \\
&= \left|\langle v_n(\cdot,s') - v_n(\cdot,s), \varphi_l\rangle_{W^{-1,1},W_0^{1,\infty}}\right| \\
&= \left|\int_s^{s'} \langle \partial_t v_n(\cdot,t), \varphi_l\rangle_{W^{-1,1},W_0^{1,\infty}}\,\mathrm{d}t\right| \\
&\le \|\partial_t v_n\|_{L^q(0,T;W^{-1,1}(\Omega))} \|\mathbf{1}\|_{L^{q'}(s,s')} \|\varphi_l\|_{W_0^{1,\infty}(\Omega)} \\
&\le R|s-s'|^{1/q'} \|\varphi_l\|_{W_0^{1,\infty}(\Omega)}.
\end{aligned}$$

Then

$$d_E(v_n(s), v_n(s')) \le \sum_{l\in\mathbb{N}} 2^{-l} \min\left(1, R|s-s'|^{1/q'} \|\varphi_l\|_{W_0^{1,\infty}(\Omega)}\right) =: \omega(s,s').$$

Dominated convergence for series then implies that $\omega(s,s') \to 0$ as $|s-s'| \to 0$. Hence, $(v_n)_{n\in\mathbb{N}}$ belongs to $C([0,T];E)$ and is equi-continuous in that space. Invoking the Arzelà–Ascoli theorem and the compactness of E in $L^2(\Omega)$-w completes the proof. □

5. Proof of the main result

We prove Theorem 2.3 in five steps. In Step 1 we obtain compactness of the sequences of interest, and in Step 2 we identify the limits of these sequences. In Step 3 we pass to the limit in (2.5). Step 4 improves the temporal convergence of $(\nu_n(u_n))_{n\in\mathbb{N}}$ to establish (2.7). We conclude by establishing the strong convergence (2.8) in Step 5.

5.1. Step 1: compactness results

Apply Proposition 4.9 using Estimates (4.5) on $(\beta_n(u_n))_{n\in\mathbb{N}}$ and $(\partial_t\beta_n(u_n))_{n\in\mathbb{N}}$, and Lemma A.1 with $H_n = \beta_n$, $v_n = u_n^{\mathrm{ini}}$ to deduce the existence of $\widetilde{\beta} \in C([0,T]; L^2(\Omega)\text{-w})$ satisfying $\widetilde{\beta}(\cdot,0) = \beta(u^{\mathrm{ini}})$ in $L^2(\Omega)$ and such that up to a subsequence,

$$\overline{\beta_n(u_n)} \rightharpoonup \widetilde{\beta} \quad \text{in } C([0,T]; L^2(\Omega)\text{-w}). \tag{5.1}$$

From (4.5), up to a subsequence,

$$\zeta_n(u_n) \rightharpoonup \widetilde{\zeta} \quad \text{weakly in } L^p(0,T;W_0^{1,p}(\Omega)) \tag{5.2}$$

for some function $\widetilde{\zeta} \in L^p(0,T;W_0^{1,p}(\Omega))$. Next we obtain strong compactness of the sequence $(\nu_n(u_n))_{n\in\mathbb{N}}$ by demonstrating that the translates in space and time vanish. Recalling (4.1a) and using a classical translate estimate in $W_0^{1,p}(\Omega)$, for $\xi \in \mathbb{R}^d$ and $q < p^*$,

$$\begin{aligned}\|\nu_n(u_n)(\cdot+\xi,\cdot) - \nu_n(u_n)\|_{L^p(0,T;L^q(\Omega))} &\le L_\beta \|\zeta_n(u_n)(\cdot+\xi,\cdot) - \zeta_n(u_n)\|_{L^p(0,T;L^q(\Omega))} \\ &\le C_4 \|\nabla\zeta_n(u_n)\|_{L^p(\Omega\times(0,T))^d} |\xi|^\theta \le C_4C_1|\xi|^\theta,\end{aligned} \tag{5.3}$$

where $\theta > 0$ and C_4 do not depend on ξ or n, and $\nu_n(u_n)$ and $\zeta_n(u_n)$ are extended by zero on the complement of Ω. But $|\nu_n(u_n)| \le L_\zeta|\beta_n(u_n)|$ and $(\nu_n(u_n))_{n\in\mathbb{N}}$ is therefore bounded in $L^\infty(0,T;L^2(\Omega))$. Interpolated with (5.3), this shows that, for all $r < +\infty$,

$$\|\nu_n(u_n)(\cdot+\xi,\cdot) - \nu_n(u_n)\|_{L^r(0,T;L^{\min(2,q)}(\Omega))} \le C_5|\xi|^{\theta_r}, \tag{5.4}$$

where $\theta_r > 0$ and C_5 do not depend on ξ or n. By the energy estimates (4.5), Lemma 4.5 applied with $F_n = \zeta_n$ and $G_n = \beta_n$ shows that the time translates of $\nu_n(u_n)$ converge uniformly to zero in $L^r(0,T;L^2(\Omega))$ for all $r < +\infty$. Combined with (5.4) and the Kolmogorov–M. Riesz–Fréchet compactness theorem, this establishes that, up to a subsequence,

$$\nu_n(u_n) \to \widetilde{\nu} \quad \text{in } L^r(0,T;L^{\min(2,q)}(\Omega)) \text{ for all } r < +\infty \text{ and all } q < p^*. \tag{5.5}$$

From the uniform growth of the sequence $(a_n)_{n\in\mathbb{N}}$ and (4.5), we assert the existence of $\widetilde{a} \in L^{p'}(\Omega\times(0,T))^d$ such that, up to a subsequence,

$$a_n(\cdot,\nu_n(u_n),\nabla\zeta_n(u_n)) \rightharpoonup \widetilde{a} \quad \text{weakly in } L^{p'}(\Omega\times(0,T))^d. \tag{5.6}$$

5.2. *Step 2: identifying nonlinear weak limits*

We show that there exists a measurable u such that $\widetilde{\beta} = \beta(u)$, $\widetilde{\zeta} = \zeta(u)$ and $\widetilde{\nu} = \nu(u)$. Three separate analyses are required, depending on the case in Hypothesis (2.2).

5.2.1. *Case* (I)*: $p \ge 2$*

Define $\mu = \beta + \zeta$, $\mu_n = \beta_n + \zeta_n$ and $\widetilde{\mu} = \widetilde{\beta} + \widetilde{\zeta}$. Fix a measurable function u such that $(\mu+\nu)(u) = \widetilde{\mu} + \widetilde{\nu}$. Such a u exists since the hypotheses on β and ζ ensure that the range of $\mu + \nu$ is all of $\mathbb{R}$ and therefore the domain of the right inverse $(\mu+\nu)^r$ of $(\mu+\nu)$ (defined analogously to (2.3)) is $\mathbb{R}$. One possible choice for u is then $u = (\mu+\nu)^r(\widetilde{\mu}+\widetilde{\nu})$. We now demonstrate that for such a u, $\widetilde{\beta} = \beta(u)$, $\widetilde{\zeta} = \zeta(u)$ and $\widetilde{\nu} = \nu(u)$.

Using $p \ge 2$, the convergences (5.2) and (5.5) ensure that $\zeta_n(u_n) \rightharpoonup \widetilde{\zeta}$ weakly in $L^2(\Omega\times(0,T))$, and that $\nu_n(u_n) \to \widetilde{\nu}$ strongly in $L^2(\Omega\times(0,T))$. We deduce that $\mu_n(u_n) = \beta_n(u_n) + \zeta_n(u_n) \rightharpoonup \widetilde{\beta} + \widetilde{\zeta} = \widetilde{\mu}$ weakly in $L^2(\Omega\times(0,T))$ and that

$$\int_{\Omega\times(0,T)} \mu_n(u_n)(x,t)\nu_n(u_n)(x,t)\,\mathrm{d}x\,\mathrm{d}t \to \int_{\Omega\times(0,T)} \widetilde{\mu}(x,t)\widetilde{\nu}(x,t)\,\mathrm{d}x\,\mathrm{d}t.$$

We can thus apply Lemma 4.7 with $\varphi \equiv 1$, $v_n = u_n$, $v = u$, $\chi_n = \mu_n$ and $\psi_n = \nu_n$ to deduce that $\widetilde{\nu} = \nu(u)$ and $\widetilde{\mu} = \mu(u)$ almost everywhere on $\Omega \times (0, T)$, the latter of which states that $(\beta + \zeta)(u) = \widetilde{\beta} + \widetilde{\zeta}$.

Since $p \geq 2$, Estimates (4.5) ensure that $(\zeta_n(u_n))_{n\in\mathbb{N}}$ and $(\beta_n(u_n))_{n\in\mathbb{N}}$ satisfy the hypotheses of Lemma B.1, and so $\beta_n(u_n)\zeta_n(u_n) \rightharpoonup \widetilde{\beta}\widetilde{\zeta}$ in $(C(\overline{\Omega} \times [0, T]))'$. Now as $(\beta + \zeta)(u) = \widetilde{\beta} + \widetilde{\zeta}$, we apply Lemma 4.7 again with $\varphi \equiv 1$, $v_n = u_n$, $v = u$, $\chi_n = \beta_n$ and $\psi_n = \zeta_n$ to conclude that $\widetilde{\beta} = \beta(u)$ and $\widetilde{\zeta} = \zeta(u)$ almost everywhere on $\Omega \times (0, T)$.

5.2.2. *Case* (II)*: $\frac{2d}{d+2} < p < 2$ and $|\beta_n(s)| \geq M_3|s| - M_4$*

Since $(\beta_n(u_n))_{n\in\mathbb{N}}$ is bounded in $L^\infty(0, T; L^2(\Omega))$, the assumption on β_n shows that $(u_n)_{n\in\mathbb{N}}$ is bounded in the same space. By the uniform Lipschitz continuity of ζ_n, we infer that $(\zeta_n(u_n))_{n\in\mathbb{N}}$ is also bounded in $L^\infty(0, T; L^2(\Omega))$. Hence, as in the previous case the convergence (5.2) also holds weakly in $L^2(\Omega \times (0, T))$. Since $p^* > 2$, (5.5) gives the strong convergence of $\nu_n(u_n)$ in $L^2(\Omega \times (0, T))$.

We proceed as in the previous case to see that with $u = (\mu + \nu)^r(\widetilde{\mu} + \widetilde{\nu})$, $\nu(u) = \widetilde{\nu}$ and $\beta(u) + \zeta(u) = \widetilde{\beta} + \widetilde{\zeta}$. Now apply Lemma B.1 to $(\zeta_n(u_n))_{n\in\mathbb{N}}$ and $(\beta_n(u_n))_{n\in\mathbb{N}}$. As in Case (I), this gives $\widetilde{\beta} = \beta(u)$ and $\widetilde{\zeta} = \zeta(u)$.

5.2.3. *Case* (III)*: $1 < p \leq \frac{2d}{d+2}$, $|\beta_n(s)| \geq M_3|s| - M_4$ and β is strictly increasing*

As in Case (II), the coercivity assumption on β_n ensures that $(\zeta_n(u_n))_{n\in\mathbb{N}}$ converges weakly in $L^2(\Omega \times (0, T))$. However, we can no longer ensure the strong convergence of $\nu_n(u_n)$ in L^2. We must therefore truncate ζ_n first. Let $\zeta_n^k = \mathbb{T}_k(\zeta_n)$, where $\mathbb{T}_k(s) = \min(k, \max(-k, s))$ is the truncation at level k. Up to a subsequence, for some $\widetilde{\zeta}^k \in L^2(\Omega \times (0, T))$, $\zeta_n^k(u_n) \rightharpoonup \widetilde{\zeta}^k$ weakly in $L^2(\Omega \times (0, T))$. Set

$$\nu_n^k(s) = \int_0^s \beta_n'(q)(\zeta_n^k)'(q)\,\mathrm{d}q.$$

Note that $(\nabla\zeta_n^k(u_n))_{n\in\mathbb{N}} = (\mathbf{1}_{\{|\zeta_n(u_n)|\leq k\}}\nabla\zeta_n(u_n))_{n\in\mathbb{N}}$ is bounded in $L^p(\Omega \times (0, T))^d$. Hence, following the reasoning in (5.3) and using an interpolation in space between p and ∞ (we have $|\zeta_n^k| \leq k$),

$$\begin{aligned}
\left\|\nu_n^k(u_n)(\cdot+\xi,\cdot) - \nu_n^k(u_n)\right\|_{L^p(0,T;L^2(\Omega))} &\leq L_\beta \left\|\zeta_n^k(u_n)(\cdot+\xi,\cdot) - \zeta_n^k(u_n)\right\|_{L^p(0,T;L^2(\Omega))} \\
&\leq L_\beta(2k)^{1-\frac{p}{2}} \left\|\zeta_n^k(u_n)(\cdot+\xi,\cdot) - \zeta_n^k(u_n)\right\|_{L^p(0,T;L^p(\Omega))}^{\frac{p}{2}} \\
&\leq C_6 \left\|\nabla\zeta_n^k(u_n)\right\|_{L^p(\Omega\times(0,T))^d}^{\frac{p}{2}} |\xi|^{\frac{p}{2}} \leq C_7|\xi|^{\frac{p}{2}},
\end{aligned}$$

where C_6 and C_7 depend on k but not on n or ξ. Use the bound on $(\nu_n^k(u_n))_{n\in\mathbb{N}}$ in $L^\infty(0, T; L^2(\Omega))$ to infer that the space translates of these functions vanish uniformly with respect to n in $L^r(0, T; L^2(\Omega))$ for all $r < +\infty$. Lemma 4.5 applied to $F_n = \zeta_n^k$ and $G_n = \beta_n$ shows that the time translates of $\nu_n^k(u_n)$ vanish uniformly with respect to n in $L^r(0, T; L^2(\Omega))$ for all $r < +\infty$. Hence, $(\nu_n^k(u_n))_{n\in\mathbb{N}}$ strongly converges, up to a subsequence, to some $\widetilde{\nu}^k$ in $L^2(\Omega \times (0, T))$.

We can then work as in the previous cases with β_n, ζ_n^k and ν_n^k. We define $\zeta^k = \mathbb{T}_k(\zeta)$ and $\nu^k(s) = \int_0^s \beta'(q)(\zeta^k)'(q)\,\mathrm{d}q$, and we let $\mu^k = \beta + \zeta^k$. By coercivity of β, the mapping $\mu^k + \nu^k$ is onto and we can define $u^k = (\mu^k + \nu^k)^r(\widetilde{\mu}^k + \widetilde{\nu}^k)$, where $\widetilde{\mu}^k = \widetilde{\beta} + \widetilde{\zeta}^k$ is the weak limit in $L^2(\Omega \times (0,T))$ of $\beta_n + \zeta_n^k$. By strong convergence in $L^2(\Omega \times (0,T))$ of $(\nu_n^k(u_n))_{n\in\mathbb{N}}$ we can apply Lemma 4.7 to see that $\widetilde{\nu}^k = \nu^k(u^k)$ and

$$\widetilde{\beta} + \widetilde{\zeta}^k = \beta(u^k) + \zeta^k(u^k). \tag{5.7}$$

We now apply Lemma B.1 to $(\zeta_n^k(u_n))_{n\in\mathbb{N}}$ and $(\beta_n(u_n))_{n\in\mathbb{N}}$. Indeed, $(\zeta_n^k(u_n))_{n\in\mathbb{N}}$ is bounded in $L^\infty(\Omega \times (0,T))$. We therefore obtain $\beta_n(u_n)\zeta_n^k(u_n) \rightharpoonup \widetilde{\beta}\,\widetilde{\zeta}^k$ weakly in $C(\overline{\Omega} \times [0,T])'$. Use Lemma 4.7 and (5.7) to deduce that $\widetilde{\zeta}^k = \zeta^k(u^k)$ and $\widetilde{\beta} = \beta(u^k)$.

Since β does not have any plateaux and $\widetilde{\beta}$ does not depend on k, the latter relation shows that u^k does not depend on k. Write $u = u^k$. Then $\widetilde{\beta} = \beta(u)$, $\widetilde{\zeta}^k = \mathbb{T}_k(\zeta(u))$ and $\widetilde{\nu}^k = \nu^k(u)$. If we can show that $\widetilde{\zeta}^k \to \widetilde{\zeta}$ and $\widetilde{\nu}^k \to \widetilde{\nu}$ in $\mathcal{D}'(\Omega \times (0,T))$ as $k \to \infty$, then we can pass to the limit in the previous equalities to get $\widetilde{\zeta} = \zeta(u)$ and $\widetilde{\nu} = \nu(u)$, as required.

Begin with the convergence of $\widetilde{\zeta}^k$. This function is the weak limit in $L^2(\Omega \times (0,T))$ of $(\zeta_n^k(u_n))_{n\in\mathbb{N}}$. By Tchebycheff's inequality, uniform Lipschitz continuity of ζ_n^k and the bound of $(u_n)_{n\in\mathbb{N}}$ in $L^\infty(0,T;L^2(\Omega))$ (from the coercivity of β_n),

$$\mathrm{meas}(\{|\zeta_n^k(u_n)| \geq k\}) \leq \frac{C_8}{k} \tag{5.8}$$

with C_8 not depending on k or n. Let $\varphi \in C_c^\infty(\Omega \times (0,T))$. Then

$$\begin{aligned}
&\left| \int_{\Omega\times(0,T)} [\widetilde{\zeta}^k - \widetilde{\zeta}(x,t)]\varphi(x,t)\,\mathrm{d}x\,\mathrm{d}t \right| \\
&\quad \leq \left| \int_{\Omega\times(0,T)} [\widetilde{\zeta}^k - \zeta_n^k(u_n)](x,t)\varphi(x,t)\,\mathrm{d}x\,\mathrm{d}t \right| \\
&\qquad + \left| \int_{\Omega\times(0,T)} [\zeta_n^k(u_n) - \zeta_n(u_n)](x,t)\varphi(x,t)\,\mathrm{d}x\,\mathrm{d}t \right| \\
&\qquad + \left| \int_{\Omega\times(0,T)} [\zeta_n(u_n) - \widetilde{\zeta}](x,t)\varphi(x,t)\,\mathrm{d}x\,\mathrm{d}t \right|.
\end{aligned} \tag{5.9}$$

By (5.2), the last term tends to 0 as $n \to \infty$. The first term also vanishes as $n \to \infty$. Estimate the second term using $|\zeta_n^k(u_n) - \zeta_n(u_n)| \leq \mathbf{1}_{\{|\zeta_n(u_n)|\geq k\}}|\zeta_n(u_n)|$, Hölder's inequality, the energy estimate (4.5) and the inequality (5.8). Taking the limit superior as $n \to \infty$ of (5.9) yields

$$\left| \int_{\Omega\times(0,T)} [\widetilde{\zeta}^k - \widetilde{\zeta}(x,t)]\varphi(x,t)\,\mathrm{d}x\,\mathrm{d}t \right| \leq C_1 ||\varphi||_{L^\infty(\Omega\times(0,T))} \left(\frac{C_8}{k}\right)^{1/p'}.$$

Letting $k \to \infty$ concludes the proof that $\widetilde{\zeta}^k \to \widetilde{\zeta}$ in the sense of distributions.

The proof that $\widetilde{\nu}^k$ converges as $k \to \infty$ to $\widetilde{\nu}$ in the sense of distributions is similar. The functions $\widetilde{\nu}^k$ and $\widetilde{\nu}$ are the weak limits in $L^2(\Omega \times (0,T))$ of $(\nu_n^k(u_n))_{n\in\mathbb{N}}$ and $(\nu_n(u_n))_{n\in\mathbb{N}}$ (note that since the latter sequence is bounded in $L^\infty(0,T;L^2(\Omega))$, the convergence (5.5) also holds weakly in this space). Moreover

$$\nu_n^k(u_n)-\nu_n(u_n)=\int_0^{u_n}\beta_n'(q)\,(\mathbb{T}_k(\zeta_n)-\zeta_n)'(q)\,\mathrm{d}q=0 \quad \text{if } |u_n|\le k.$$

We can therefore reproduce the same reasoning as for the convergence of $(\widetilde{\zeta}^k)_{k\to\infty}$ to see that $\widetilde{\nu}^k\to\widetilde{\nu}$ in the sense of distributions as $k\to\infty$.

Remark 5.1. If $\beta_n=\mathrm{Id}$ (resp. $\zeta_n=\mathrm{Id}$), then $\nu_n=\zeta_n$ (resp. $\nu_n=\beta_n$) and the strong convergence of $\nu_n(u_n)$ enables us to pass to the limit in $\int_{\Omega\times(0,T)}\beta_n(u_n)\zeta_n(u_n)$ (or the truncated version if p is small). We only need the compensated compactness lemma to identify this limit in the case of two genuine degeneracies, that is $\beta_n\neq\mathrm{Id}$ and $\zeta_n\neq\mathrm{Id}$.

5.3. *Step 3: the function u is a solution to (P)*

We know that $\zeta(u)=\widetilde{\zeta}\in L^p(0,T;W_0^{1,p}(\Omega))$, $\beta(u)=\widetilde{\beta}\in C([0,T];L^2(\Omega)\text{-w})$ (with an abuse of notation), $\overline{\beta(u)}(\cdot,0)=\widetilde{\beta}(\cdot,0)=\beta(u^{\mathrm{ini}})$. Since $(\partial_t\beta_n(u_n))_{n\in\mathbb{N}}$ is bounded in $L^{p'}(0,T;W^{-1,p'}(\Omega))$, we infer that $\partial_t\beta_n(u_n)\rightharpoonup\partial_t\beta(u)$ weakly in this space. Lemma 4.3 shows that $\Psi=B$ and $\Psi_n=B_n$ satisfy the assumptions of Lemma 4.6. Let $T_0\in[0,T]$. By (5.1), $\overline{\beta_n(u_n)}(\cdot,T_0)\rightharpoonup\overline{\beta(u)}(\cdot,T_0)$ weakly in $L^2(\Omega)$. Hence by Lemma 4.6,

$$\int_\Omega B(\overline{\beta(u)}(x,T_0))\,\mathrm{d}x\le\liminf_{n\to\infty}\int_\Omega B_n(\overline{\beta_n(u_n)}(x,T_0))\,\mathrm{d}x. \tag{5.10}$$

Combined with (4.5), this shows that $B(\beta(u))\in L^\infty(0,T;L^1(\Omega))$.

Passing to the limit as $n\to\infty$ in (2.5) is then possible thanks to the convergence properties of $\partial_t\beta_n(u_n)$ and $a_n(\cdot,\nu_n(u_n),\nabla\zeta_n(u_n))$. We obtain

$$\begin{aligned}&\int_0^T\langle\partial_t\beta(u)(\cdot,t),v(\cdot,t)\rangle_{W^{-1,p'},W_0^{1,p}}\,\mathrm{d}t+\int_0^T\int_\Omega\widetilde{a}(x,t)\cdot\nabla v(x,t)\,\mathrm{d}x\,\mathrm{d}t\\&=\int_0^T\langle f(\cdot,t),v(\cdot,t)\rangle_{W^{-1,p'},W_0^{1,p}}\,\mathrm{d}t\quad\forall v\in L^p(0,T;W_0^{1,p}(\Omega)).\end{aligned} \tag{5.11}$$

To complete Step 3, it remains to demonstrate that

$$\widetilde{a}(x,t)=a(x,\nu(u),\nabla\zeta(u))(x,t)\quad\text{for a.e. }(x,t)\in\Omega\times(0,T). \tag{5.12}$$

Let $T_0\in[0,T]$ and consider the identity (4.3) with data $(\beta_n,\zeta_n,\nu_n,a_n,f_n,u_n^{\mathrm{ini}})$. Take the limit superior and use (5.2) (recall that $\widetilde{\zeta}=\zeta(u)$) to obtain

$$\begin{aligned}&\limsup_{n\to\infty}\int_0^{T_0}\int_\Omega a_n(x,\nu_n(u_n),\nabla\zeta_n(u_n))\cdot\nabla\zeta_n(u_n)\,\mathrm{d}x\,\mathrm{d}t\\&\le\limsup_{n\to\infty}\int_\Omega B_n(\beta_n(u_n^{\mathrm{ini}}(x)))\,\mathrm{d}x+\int_0^{T_0}\langle f(\cdot,t),\zeta(u)(\cdot,t)\rangle_{W^{-1,p'},W_0^{1,p}}\,\mathrm{d}t\\&\quad-\liminf_{n\to\infty}\int_\Omega B_n(\overline{\beta_n(u_n)}(x,T_0))\,\mathrm{d}x.\end{aligned} \tag{5.13}$$

Part (ii) of Lemma 4.3 and (4.1d) show that $B_n \circ \beta_n$ converges uniformly and has uniform quadratic growth. By applying Lemma A.1, the convergence $u_n^{\mathrm{ini}} \to u^{\mathrm{ini}}$ in $L^2(\Omega)$ shows that $(B_n \circ \beta_n)(u_n^{\mathrm{ini}}) \to (B \circ \beta)(u^{\mathrm{ini}})$ in $L^1(\Omega)$. Together with the inequality (5.10), this gives

$$\begin{aligned}\limsup_{n\to\infty} \int_0^{T_0} \int_\Omega & a_n(x, \nu_n(u_n), \nabla\zeta_n(u_n)) \cdot \nabla\zeta_n(u_n)\,\mathrm{d}x\,\mathrm{d}t \\ &\leq \int_\Omega B(\beta(u^{\mathrm{ini}}(x)))\,\mathrm{d}x + \int_0^{T_0} \langle f(\cdot,t), \zeta(u)(\cdot,t)\rangle_{W^{-1,p'},W_0^{1,p}}\,\mathrm{d}t \\ &\quad - \int_\Omega B(\overline{\beta(u)}(x,T_0))\,\mathrm{d}x = \int_0^{T_0} \int_\Omega \widetilde{a}(x,t) \cdot \nabla\zeta(u)\,\mathrm{d}x\,\mathrm{d}t, \end{aligned} \tag{5.14}$$

using the identities (4.2) (with $v = u$) and (5.11) (with $v = \zeta(u)$).

We now employ the classical Minty–Browder argument. For $G \in L^p(0,T; L^p(\Omega)^d)$, the monotonicity of $(a_n)_{n\in\mathbb{N}}$ gives

$$\int_0^{T_0} \int_\Omega [a_n(x, \nu_n(u_n), \nabla\zeta_n(u_n)) - a_n(x, \nu_n(u_n), G)] \cdot [\nabla\zeta_n(u_n) - G]\,\mathrm{d}x\,\mathrm{d}t \geq 0. \tag{5.15}$$

Together with the strong convergence in $L^1(\Omega \times (0,T))$ of $\nu_n(u_n)$ to $\nu(u)$, the assumptions on the sequence $(a_n)_{n\in\mathbb{N}}$ ensure that $a_n(\cdot, \nu_n(u_n), G)$ converges in $L^{p'}(\Omega \times (0,T))^d$ to $a(\cdot, \nu(u), G)$. Using this, (5.14) and the weak convergence (5.6), we pass to the limit superior on the expanded form of (5.15) with $T_0 = T$ to see that

$$\int_0^{T_0} \int_\Omega [\widetilde{a}(x,t) - a(x, \nu(u(x,t)), G(x,t))] \cdot [\nabla\zeta(u)(x,t) - G(x,t)]\,\mathrm{d}x\,\mathrm{d}t \geq 0.$$

Following G.J. Minty [26], take $G = \nabla\zeta(u) \pm r\varphi$ for $\varphi \in L^p(0,T; L^p(\Omega)^d)$, divide by $r > 0$ and let $r \to 0$ to obtain (5.12).

5.4. Step 4: uniform temporal convergence of $\nu_n(u_n)$ to $\nu(u)$

Take $T_\infty \in [0,T]$ and $(T_n)_{n\in\mathbb{N}} \subset [0,T]$ a sequence converging to T_∞. Thanks to Lemma A.2, the convergence of $(\overline{\nu_n(u_n)})_{n\in\mathbb{N}}$ in $C([0,T]; L^2(\Omega))$ follows if we can demonstrate that

$$\lim_{n\to\infty} \left\| \overline{\nu_n(u_n)}(\cdot, T_n) - \overline{\nu(u)}(\cdot, T_\infty) \right\|_{L^2(\Omega)} = 0. \tag{5.16}$$

Note the use of the continuous representatives $[0,T] \to L^2(\Omega)$ of $\nu_n(u_n)$ and $\nu(u)$ (whose existence is ensured by Lemma 4.2). Without loss of generality, we can assume that T_n is such that

$$\beta_n(u_n(\cdot, T_n)) = \overline{\beta_n(u_n)}(\cdot, T_n) \text{ and } \nu_n(u_n(\cdot, T_n)) = \overline{\nu_n(u_n)}(\cdot, T_n) \text{ a.e. on } \Omega. \tag{5.17}$$

Indeed, by definition of the continuous representatives, there is $T_n' \in (T_n - 1/n, T_n + 1/n) \cap [0,T]$ such that (5.17) holds at T_n' and such that

$$\left\| \overline{\nu_n(u_n)}(\cdot, T_n) - \overline{\nu_n(u_n)}(\cdot, T_n') \right\|_{L^2(\Omega)} \leq \frac{1}{n},$$

using $\overline{\nu_n(u_n)} \in C([0,T]; L^2(\Omega))$. Proving (5.16) with T_n' instead of T_n establishes it for T_n also.

To estimate the quantity in (5.16), which involves a variation of ν_n and u_n with respect to n, our strategy is to freeze one of these variations using the triangle inequality with $\nu_n(u)(\cdot, T_\infty)$ as an intermediate point. But $\nu_n(u)$ may not be continuous in time, so its value at T_∞ is not well-defined. Instead we use $\nu_n(u)(\cdot, s)$ and average over a small interval around T_∞. To this end, let $\varepsilon > 0$ and define $I_\varepsilon := [T_\infty - \varepsilon, T_\infty + \varepsilon] \cap [0, T]$. Using (5.17) and (4.1e) with ν_n, B_n and β_n, write

$$\begin{aligned}
&\left\| \overline{\nu_n(u_n)}(\cdot, T_n) - \overline{\nu(u)}(\cdot, T_\infty) \right\|_{L^2(\Omega)}^2 \\
&\quad \leq 2 \fint_{I_\varepsilon} \| \nu_n(u_n(\cdot, T_n)) - \nu_n(u(\cdot, s)) \|_{L^2(\Omega)}^2 \, \mathrm{d}s + 2 \fint_{I_\varepsilon} \left\| \nu_n(u(\cdot, s)) - \overline{\nu(u)}(\cdot, T_\infty) \right\|_{L^2(\Omega)}^2 \mathrm{d}s \\
&\quad \leq 8 L_\beta L_\zeta \Bigg(\int_\Omega B_n(\beta_n(u_n(x, T_n))) \, \mathrm{d}x + \fint_{I_\varepsilon} \int_\Omega B_n(\beta_n(u(x,s))) \, \mathrm{d}x \, \mathrm{d}s \\
&\qquad - 2 \fint_{I_\varepsilon} \int_\Omega B_n\left(\frac{\beta_n(u_n(x, T_n)) + \beta_n(u(x,s))}{2} \right) \mathrm{d}x \, \mathrm{d}s \Bigg) \\
&\qquad + 2 \fint_{I_\varepsilon} \left\| \nu_n(u(\cdot, s)) - \overline{\nu(u)}(\cdot, T_\infty) \right\|_{L^2(\Omega)}^2 \mathrm{d}s \\
&\quad =: 8 L_\beta L_\zeta [\mathcal{I}_1(n) + \mathcal{I}_2(n, \varepsilon) - 2\mathcal{I}_3(n, \varepsilon)] + 2\mathcal{I}_4(n, \varepsilon).
\end{aligned} \tag{5.18}$$

To determine the convergence of $\mathcal{I}_1$, expand (5.15) with $T_0 = T_n$, $G = \nabla\zeta(u)$ and take the limit inferior of the resulting expression. Noting the identity (5.12), we obtain

$$\begin{aligned}
&\liminf_{n\to\infty} \int_0^{T_n} \int_\Omega a_n(x, \nu_n(u_n), \nabla\zeta_n(u_n)) \cdot \nabla\zeta_n(u_n) \, \mathrm{d}x \, \mathrm{d}t \\
&\qquad \geq \int_0^{T_\infty} \int_\Omega a(x, \nu(u), \nabla\zeta(u)) \cdot \nabla\zeta(u) \, \mathrm{d}x \, \mathrm{d}t.
\end{aligned} \tag{5.19}$$

Now in (4.3), replace $(\beta, \zeta, \nu, a, f, u^{\mathrm{ini}}, T_0)$ with $(\beta_n, \zeta_n, \nu_n, a_n, f_n, u_n^{\mathrm{ini}}, T_n)$ and using (5.17), (5.19) and the fact that u satisfies the energy equality (4.3), take the limit superior as $n \to \infty$ to deduce that

$$\limsup_{n\to\infty} \mathcal{I}_1(n) \leq \int_\Omega B(\overline{\beta(u)}(x, T_\infty)) \, \mathrm{d}x < +\infty. \tag{5.20}$$

To handle $\mathcal{I}_2$, recall that $B_n \circ \beta_n$ converges locally uniformly on $\mathbb{R}$ to $B \circ \beta$ (Lemma 4.3). By Hypothesis (2.2), $u \in L^2(\Omega \times (0,T))$. Hence by the dominated convergence theorem, the quadratic growth (4.1d) of B_n ensures that $B_n(\beta_n(u)) \to B(\beta(u)) = B(\overline{\beta(u)})$ in $L^1(\Omega \times (0,T))$ and so

$$\lim_{n\to\infty} \mathcal{I}_2(n, \varepsilon) = \fint_{I_\varepsilon} \int_\Omega B(\overline{\beta(u)}(x,s)) \, \mathrm{d}x \, \mathrm{d}s. \tag{5.21}$$

Now the convexity of B_n enables the application of Jensen's inequality to $\mathcal{I}_3$, yielding

$$\begin{aligned}\mathcal{I}_3(n,\varepsilon) &= \fint_{I_\varepsilon}\int_\Omega B_n\left(\frac{\beta_n(u_n(x,T_n))+\beta_n(u(x,s))}{2}\right)\mathrm{d}x\,\mathrm{d}s\\ &\geq \int_\Omega B_n\left(\frac{\beta_n(u_n(x,T_n))+\fint_{I_\varepsilon}\beta_n(u(x,s))\,\mathrm{d}s}{2}\right)\mathrm{d}x.\end{aligned}$$

The convergence in $C([0,T];L^2(\Omega)\text{-w})$ of $\overline{\beta_n(u_n)}$ to $\overline{\beta(u)}$ and the continuity of the latter imply by Lemma A.2 that $\beta_n(u_n(\cdot,T_n)) = \overline{\beta_n(u_n)}(\cdot,T_n) \rightharpoonup \overline{\beta(u)}(\cdot,T_\infty)$ weakly in $L^2(\Omega)$. Since $u\in L^2(\Omega\times(0,T))$ the assumptions on β_n give $\beta_n(u)\to\beta(u)=\overline{\beta(u)}$ in $L^2(0,T;L^2(\Omega))$, and so

$$\fint_{I_\varepsilon}\beta_n(u(\cdot,s))\,\mathrm{d}s \to \fint_{I_\varepsilon}\overline{\beta(u)}(\cdot,s)\,\mathrm{d}s \quad \text{in } L^2(\Omega).$$

Thus $\frac{1}{2}(\beta_n(u_n(,\cdot,T_n)) + \fint_{I_\varepsilon}\beta_n(u(\cdot,s))\,\mathrm{d}s) \rightharpoonup \frac{1}{2}(\overline{\beta(u)}(\cdot,T_\infty) + \fint_{I_\varepsilon}\overline{\beta(u)}(\cdot,s)\,\mathrm{d}s)$ weakly in $L^2(\Omega)$ and Lemma 4.6 gives

$$\int_\Omega B\left(\frac{\overline{\beta(u)}(x,T_\infty)+\fint_{I_\varepsilon}\overline{\beta(u)}(x,s)\,\mathrm{d}s}{2}\right)\mathrm{d}x \leq \liminf_{n\to\infty}\mathcal{I}_3(n,\varepsilon). \tag{5.22}$$

Since $u\in L^2(\Omega\times(0,T))$, $\nu_n(u)\to\nu(u)$ in $L^2(\Omega\times(0,T))$ and so

$$\begin{aligned}\mathcal{I}_4(n,\varepsilon) &= \frac{1}{|I_\varepsilon|}\left\|\nu_n(u)-\overline{\nu(u)}(\cdot,T_\infty)\right\|^2_{L^2(\Omega\times I_\varepsilon)} \to \frac{1}{|I_\varepsilon|}\left\|\nu(u)-\overline{\nu(u)}(\cdot,T_\infty)\right\|^2_{L^2(\Omega\times I_\varepsilon)}\\ &= \fint_{I_\varepsilon}\left\|\overline{\nu(u)}(\cdot,s)-\overline{\nu(u)}(\cdot,T_\infty)\right\|^2_{L^2(\Omega)}\mathrm{d}s.\end{aligned} \tag{5.23}$$

Thanks to (5.20), (5.21) and (5.22), we may split the limit superior as $n\to\infty$ of the right-hand side of (5.18), using (5.23) for the remaining term to obtain

$$\begin{aligned}&\limsup_{n\to\infty}\left\|\overline{\nu_n(u_n)}(\cdot,T_n)-\overline{\nu(u)}(\cdot,T_\infty)\right\|^2_{L^2(\Omega)}\\ &\leq 8L_\beta L_\zeta\Bigg(\int_\Omega B(\overline{\beta(u)}(x,T_\infty))\,\mathrm{d}x + \fint_{I_\varepsilon}\int_\Omega B(\overline{\beta(u)}(x,s))\,\mathrm{d}x\,\mathrm{d}s\\ &\quad -2\int_\Omega B\left(\frac{\overline{\beta(u)}(x,T_\infty)+\fint_{I_\varepsilon}\overline{\beta(u)}(x,s)\,\mathrm{d}s}{2}\right)\mathrm{d}x\Bigg)\\ &\quad +2\fint_{I_\varepsilon}\left\|\overline{\nu(u)}(\cdot,s)-\overline{\nu(u)}(\cdot,T_\infty)\right\|^2_{L^2(\Omega)}\mathrm{d}s.\end{aligned} \tag{5.24}$$

To complete the proof it remains to take the superior limit as $\varepsilon \to 0$. By the continuity of the mapping $[0,T] \ni s \mapsto \int_\Omega B(\overline{\beta(u)}(x,s))\,\mathrm{d}x$,

$$\lim_{\varepsilon\to 0} \fint_{I_\varepsilon} \int_\Omega B(\overline{\beta(u)}(x,s))\,\mathrm{d}x\,\mathrm{d}s = \int_\Omega B(\overline{\beta(u)}(x,T_\infty))\,\mathrm{d}x.$$

Using the continuity of $\overline{\nu(u)} : [0,T] \to L^2(\Omega)$,

$$\lim_{\varepsilon\to 0} \fint_{I_\varepsilon} \left\| \overline{\nu(u)}(\cdot,s) - \overline{\nu(u)}(\cdot,T_\infty) \right\|^2_{L^2(\Omega)} \mathrm{d}s = 0.$$

Since B is convex lower semi-continuous and $\frac{1}{2}(\overline{\beta(u)}(\cdot,T_\infty) + \fint_{I_\varepsilon} \overline{\beta(u)}(\cdot,s)\,\mathrm{d}s) \rightharpoonup \overline{\beta(u)}(\cdot,T_\infty)$ weakly in $L^2(\Omega)$ as $\varepsilon \to 0$ (using the continuity of $\overline{\beta(u)} : [0,T] \to L^2(\Omega)$-w), we apply Lemma 4.6 to deduce that

$$\int_\Omega B(\overline{\beta(u)}(x,T_\infty))\,\mathrm{d}x \le \liminf_{\varepsilon\to 0} \int_\Omega B\left(\frac{\overline{\beta(u)}(x,T_\infty) + \fint_{I_\varepsilon} \overline{\beta(u)}(x,s)\,\mathrm{d}s}{2} \right) \mathrm{d}x.$$

Taking the limit supremum as $\varepsilon \to 0$ of (5.24) yields (5.16), hence the result.

Remark 5.2. Since $\overline{\beta_n(u_n)}(\cdot,T_n) \rightharpoonup \overline{\beta(u)}(\cdot,T_\infty)$ weakly in $L^2(\Omega)$ whenever $T_n \to T_\infty$ (see Lemma A.2), (5.10) still holds with T_0 in the left-hand side replaced with T_∞ and T_0 in the right-hand side replaced with T_n. Thus with (5.20) we see that

$$\int_\Omega B_n(\overline{\beta_n(u_n)}(x,T_n))\,\mathrm{d}x \to \int_\Omega B(\overline{\beta(u)}(x,T_\infty))\,\mathrm{d}x \quad \text{as } n \to \infty.$$

Lemma A.2 and Part (i) in Lemma 4.2 then show that $\int_\Omega B_n(\overline{\beta_n(u_n)}(x,\cdot))\,\mathrm{d}x$ converges uniformly to $\int_\Omega B(\overline{\beta(u)}(x,\cdot))\,\mathrm{d}x$ on $[0,T]$.

5.5. Step 5: convergence of $\zeta_n(u_n)$ to $\zeta(u)$ in $L^p(0,T;W_0^{1,p}(\Omega))$

We follow the ideas of J. Leray and J.-L. Lions [22]. Use (5.14) with $T_0 = T$ and (5.12):

$$\limsup_{n\to\infty} \int_0^T \int_\Omega a_n(x,\nu_n(u_n),\nabla\zeta_n(u_n)) \cdot \nabla\zeta_n(u_n)\,\mathrm{d}x\,\mathrm{d}t \\ \le \int_0^T \int_\Omega a(x,\nu(u),\nabla\zeta(u)) \cdot \nabla\zeta(u)\,\mathrm{d}x\,\mathrm{d}t.$$

Together with $T_n = T_\infty = T$ in (5.19), we see that

$$\lim_{n\to\infty} \int_0^T \int_\Omega a_n(x,\nu_n(u_n),\nabla\zeta_n(u_n)) \cdot \nabla\zeta_n(u_n)\,\mathrm{d}x\,\mathrm{d}t \\ = \int_0^T \int_\Omega a(x,\nu(u),\nabla\zeta(u)) \cdot \nabla\zeta(u)\,\mathrm{d}x\,\mathrm{d}t. \tag{5.25}$$

Now define

$$F_n := [a_n(x, \nu_n(u_n), \nabla\zeta_n(u_n)) - a_n(x, \nu_n(u_n), \nabla\zeta(u))] \cdot [\nabla\zeta_n(u_n) - \nabla\zeta(u)] \geq 0,$$

integrate this expression over $\Omega \times (0, T)$ and expand. The convergences (5.2), (5.6), (5.25) and the convergence in $L^{p'}(\Omega \times (0, T))^d$ of $a_n(\cdot, \nu_n(u_n), \nabla\zeta(u))$ to $a(\cdot, \nu(u), \nabla\zeta(u))$ imply that, as $n \to \infty$,

$$\int_0^T \int_\Omega F_n(x, t)\, \mathrm{d}x\, \mathrm{d}t \to 0.$$

The nonnegativity of F_n then ensures that F_n converges to zero in $L^1(\Omega \times (0, T))$ and therefore, upon extraction of a subsequence, almost everywhere on $\Omega \times (0, T)$. Now use the strict monotonicity of a to apply Lemma A.4 with $X = \Omega \times \mathbb{R}$, $b_n(s, \xi) = a_n(x, s, \xi)$, $\chi_n = \nabla\zeta_n(u_n)$ to deduce that, up to a subsequence, $\nabla\zeta_n(u_n) \to \nabla\zeta(u)$ almost everywhere on $\Omega \times (0, T)$. A subsequence of $(\nu_n(u_n))_{n\in\mathbb{N}}$ converges almost everywhere on $\Omega \times (0, T)$ to $\nu(u)$, therefore the local uniform convergence on $\mathbb{R} \times \mathbb{R}^d$ of $(a_n)_{n\in\mathbb{N}}$ ensures that

$$a_n(\cdot, \nu_n(u_n), \nabla\zeta_n(u_n)) \cdot \nabla\zeta_n(u_n) \to a(\cdot, \nu(u), \nabla\zeta(u)) \cdot \nabla\zeta(u) \quad \text{a.e. on } \Omega \times (0, T).$$

Lemma A.3 then guarantees, using (5.25) and the nonnegativity of $a_n(\cdot, \nu_n(u_n), \nabla\zeta_n(u_n)) \cdot \nabla\zeta_n(u_n)$, that

$$a_n(\cdot, \nu_n(u_n), \nabla\zeta_n(u_n)) \cdot \nabla\zeta_n(u_n) \to a(\cdot, \nu(u), \nabla\zeta(u)) \cdot \nabla\zeta(u) \quad \text{in } L^1(\Omega \times (0, T)).$$

Therefore, the sequence $(a_n(\cdot, \nu_n(u_n), \nabla\zeta_n(u_n)) \cdot \nabla\zeta_n(u_n))_{n\in\mathbb{N}}$ is equi-integrable, and so too is $(|\nabla\zeta_n(u_n)|^p)_{n\in\mathbb{N}}$ thanks to the uniform coercivity of $(a_n)_{n\in\mathbb{N}}$. The strong convergence (2.8) then follows from Vitali's theorem. □

Acknowledgments

The authors thank Thierry Gallouët and Clément Cancès for fruitful discussions on compensated compactness results.

Appendix A. Convergence lemmas

We make frequent use of the following lemma, proved in [16].

Lemma A.1. *Let $H_n : \mathbb{R} \to \mathbb{R}$ be a sequence of continuous functions such that*

(i) *there exist positive constants C_9, γ such that for every $s \in \mathbb{R}$, $|H_n(s)| \leq C_9(1 + |s|^\gamma)$;*
(ii) *H_n converges locally uniformly on $\mathbb{R}$ to a continuous function $H : \mathbb{R} \to \mathbb{R}$.*

Let $N \in \mathbb{N}$ and take a bounded subset E of $\mathbb{R}^N$. If $q \in [\gamma, \infty)$ and $(v_n)_{n\in\mathbb{N}} \subset L^q(E)$ is such that $v_n \to v$ in $L^q(E)$, then $H_n(v_n) \to H(v)$ in $L^{q/\gamma}(E)$ as $n \to \infty$.

The next lemma gives an equivalent characterisation of uniform convergence, which is critical to Step 3 of the proof of our main result. For a proof of this lemma, see [14].

Lemma A.2. *Let (K, d_K) be a compact metric space, (E, d_E) a metric space. 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).*

Let $(v_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{F}(K, E)$ and $v : K \to E$ be continuous. Then $v_n \to v$ for $d_{\mathcal{F}}$ if and only if, for any $s \in K$ and any sequence $(s_n)_{n\in\mathbb{N}} \subset K$ converging to s for d_K, $v_n(s_n) \to v(s)$ for d_E.

We employ the final two lemmas of this appendix in Section 5.5 to establish the (strong) convergence in $L^p(0, T; W_0^{1,p}(\Omega))$ of $\zeta_n(u_n)$ to $\zeta(u)$. For a proof of the first of these lemmas, see [15, Lemma 3.3]. The second is a slight modification of [15, Lemma 3.2].

Lemma A.3. *Let $(F_n)_{n\in\mathbb{N}}$ be a sequence of nonnegative functions in $L^1(\Omega)$. Let $F \in L^1(\Omega)$ be such that $F_n \to F$ almost everywhere and*

$$\int_\Omega F_n(x)\,\mathrm{d}x \to \int_\Omega F(x)\,\mathrm{d}x.$$

Then $F_n \to F$ in $L^1(\Omega)$ as $n \to \infty$.

Lemma A.4. *Let X be a metric space and for every $n \in \mathbb{N}$ let $b_n : X \times \mathbb{R}^d \to \mathbb{R}^d$ be continuous and monotone:*

$$(b_n(u, \delta) - b_n(u, \gamma)) \cdot (\delta - \gamma) \geq 0 \quad \forall u \in X,\ \forall \delta, \gamma \in \mathbb{R}^d.$$

Assume that b_n converges locally uniformly on $X \times \mathbb{R}^d$ to a continuous map $b : X \times \mathbb{R}^d \to \mathbb{R}^d$ that is strictly monotone:

$$(b(u, \delta) - b(u, \gamma)) \cdot (\delta - \gamma) > 0 \quad \forall u \in X,\ \forall \delta \neq \gamma \in \mathbb{R}^d.$$

Take a sequence $(u_n, \chi_n) \in X \times \mathbb{R}^d$ and $(u, \chi) \in X \times \mathbb{R}^d$ such that as $n \to \infty$,

$$(b_n(u_n, \chi_n) - b_n(u_n, \chi)) \cdot (\chi_n - \chi) \to 0 \quad \textit{and} \quad u_n \to u.$$

Then $\chi_n \to \chi$.

Proof. Let $\delta \in \mathbb{R}^d \setminus \{0\}$. For $n \in \mathbb{N}$, define $h_{\delta,n} : \mathbb{R} \to \mathbb{R}$ by

$$h_{\delta,n}(s) := (b_n(u_n, \chi + s\delta) - b_n(u_n, \chi)) \cdot \delta.$$

For $s > s'$,

$$(h_{\delta,n}(s) - h_{\delta,n}(s'))(s - s') = (b_n(u_n, \chi + s\delta) - b_n(u_n, \chi + s'\delta)) \cdot \delta(s - s') \geq 0,$$

so $h_{\delta,n}$ is a nondecreasing function. Now assume that χ_n does not converge to χ, so there is some $\varepsilon > 0$ and a subsequence of $(\chi_n)_{n\in\mathbb{N}}$, not relabelled for convenience, such that $s_n := |\chi_n - \chi| \geq \varepsilon$ for all $n \in \mathbb{N}$. Define

$$\delta_n := \frac{\chi_n - \chi}{|\chi_n - \chi|}.$$

There exists $\delta \in \mathbb{R}^d$ with $|\delta| = 1$ such that, upon extraction of a subsequence, $\delta_n \to \delta$. Then

$$\begin{aligned}(b_n(u_n, \chi_n) - b_n(u_n, \chi)) \cdot \frac{\chi_n - \chi}{s_n} \\ = h_{\delta_n,n}(s_n) \geq h_{\delta_n,n}(\varepsilon) = (b_n(u_n, \chi + \varepsilon\delta_n) - b_n(u_n, \chi)) \cdot \delta_n.\end{aligned}$$

Let $n \to \infty$ to see that

$$\begin{aligned}0 &= \lim_{n\to\infty} \frac{1}{s_n}(b_n(u_n, \chi_n) - b_n(u_n, \chi)) \cdot (\chi_n - \chi) \\ &\geq \lim_{n\to\infty} (b_n(u_n, \chi + \varepsilon\delta_n) - b_n(u_n, \chi)) \cdot \delta_n \\ &= (b(u, \chi + \varepsilon\delta) - b(u, \chi)) \cdot \delta > 0,\end{aligned}$$

a contradiction. □

Appendix B. Compensated compactness lemma

Space–time compensated compactness results usually state the convergence of a product $(f_n g_n)_{n\in\mathbb{N}}$ of functions, each one converging only weakly but $(f_n)_{n\in\mathbb{N}}$ having compactness properties in space and $(g_n)_{n\in\mathbb{N}}$ having compactness properties in time. As seen in the work of A.V. Kazhikhov [19], A. Moussa [27] and references therein, the proof of compensated compactness is often a consequence of the Aubin–Simon compactness theorem. The following lemma is no exception.

Lemma B.1. *Let Ω be an open and bounded domain in $\mathbb{R}^d$, $T > 0$, and $p \in (1, \infty)$. Take two sequences of functions $(f_n)_{n\in\mathbb{N}}$ and $(g_n)_{n\in\mathbb{N}}$ in $L^2(\Omega \times (0, T))$ such that*

- *$f_n \rightharpoonup f$ and $g_n \rightharpoonup g$ weakly-$*$ in $L^2(\Omega \times (0, T))$ as $n \to \infty$,*
- *$(f_n)_{n\in\mathbb{N}}$ is bounded in $L^p(0, T; W_0^{1,p}(\Omega))$,*
- *$(\partial_t g_n)_{n\in\mathbb{N}}$ is bounded in $L^{p'}(0, T; W^{-1,p'}(\Omega))$.*

Assume furthermore that one of the following properties holds:

(i) *$p \geq 2$, or*
(ii) *$\frac{2d}{d+2} < p < 2$ and $(g_n)_{n\in\mathbb{N}}$ is bounded in $L^{p'}(0, T; L^2(\Omega))$, or*
(iii) *$1 < p \leq \frac{2d}{d+2}$, $(g_n)_{n\in\mathbb{N}}$ is bounded in $L^{p'}(0, T; L^2(\Omega))$, and $(f_n)_{n\in\mathbb{N}}$ is bounded in $L^\infty(\Omega \times (0, T))$.*

Then $f_n g_n \to fg$ in the sense of measures on $\Omega \times (0, T)$, that is, for all $\varphi \in C(\overline{\Omega} \times [0, T])$,

$$\int_0^T \int_\Omega f_n(x,t) g_n(x,t) \varphi(x,t) \, \mathrm{d}x \, \mathrm{d}t \to \int_0^T \int_\Omega f(x,t) g(x,t) \varphi(x,t) \, \mathrm{d}x \, \mathrm{d}t \text{ as } n \to \infty. \quad \text{(B.1)}$$

Remark B.2. This result is clearly not optimal and the conclusion holds under much weaker assumptions. Using for example the ideas of [14], which consists of reducing the proof to the case where $(f_n)_{n\in\mathbb{N}}$ are tensorial functions, we could establish a convergence result for $(f_n g_n)_{n\in\mathbb{N}}$

under weaker bounds on the functions, and assuming only space-translate estimates of $(f_n)_{n\in\mathbb{N}}$ instead of bounds in a Lebesgue–Sobolev space. We establish only this simpler lemma that is adapted precisely to our setting, and emphasise that we make no claim over the originality of its core idea.

Remark B.3 *(p small).* If p is too small, then an additional assumption on $(f_n)_{n\in\mathbb{N}}$ is mandatory, as the following example shows.

If $p \le \frac{2d}{d+2}$ then $p^* \le 2$ and $W_0^{1,p}(\Omega)$ is therefore not compactly embedded in $L^2(\Omega)$. Take a sequence $(u_n)_{n\in\mathbb{N}}$ that is bounded in $W_0^{1,p}(\Omega) \cap L^2(\Omega)$ that converges weakly but not strongly to some $u \in L^2(\Omega)$. Set $f_n(x,t) = g_n(x,t) = u_n(x)$ and $f(x,t) = g(x,t) = u(x)$. Then $f_n \rightharpoonup f$ and $g_n \rightharpoonup g$ weakly in $L^\infty(0,T;L^2(\Omega))$, $(f_n)_{n\in\mathbb{N}}$ is bounded in $L^\infty(0,T;W_0^{1,p}(\Omega))$ and $\partial_t g_n = 0$, but the convergence of $\int_{\Omega\times(0,T)} f_n g_n$ to $\int_{\Omega\times(0,T)} fg$ would imply that $\|u_n\|_{L^2(\Omega)} \to \|u\|_{L^2(\Omega)}$. Hence u_n would converge strongly to u in $L^2(\Omega)$, which is a contradiction.

Proof. By density of $C^\infty(\overline{\Omega}\times[0,T])$ in $C(\overline{\Omega}\times[0,T])$, we only need to prove the result for φ smooth. Replacing $(f_n)_{n\in\mathbb{N}}$ with $(\varphi f_n)_{n\in\mathbb{N}}$, which has the same bound and convergence properties as $(f_n)_{n\in\mathbb{N}}$, we can actually assume that $\varphi = 1$ and we only have to prove

$$\int_0^T\int_\Omega f_n(x,t)g_n(x,t)\,\mathrm{d}x\,\mathrm{d}t \to \int_0^T\int_\Omega f(x,t)g(x,t)\,\mathrm{d}x\,\mathrm{d}t \text{ as } n\to\infty. \tag{B.2}$$

We recall a classical consequence of Aubin–Simon's theorem [9,13]: *assume that V, E and F are Banach spaces such that V is compactly embedded in E and E is continuously embedded in F; if $(w_n)_{n\in\mathbb{N}}$ is bounded in $L^r(0,T;V)$ and $(\partial_t w_n)_{n\in\mathbb{N}}$ is bounded in $L^m(0,T;F)$ for some $r,m\in(1,\infty]$, then $(w_n)_{n\in\mathbb{N}}$ is relatively compact in $L^r(0,T;E)$.*

We first consider Cases (i) and (ii). In both cases, $p^* > 2$ and thus $W_0^{1,p}(\Omega)$ is compactly embedded in $L^2(\Omega)$. By duality, we infer that $V = L^2(\Omega)$ is compactly embedded in $E = F = W^{-1,p'}(\Omega)$. Since $(g_n)_{n\in\mathbb{N}}$ is bounded in $L^{p'}(0,T;V)$ (in Case (i), we use the fact that $p' \le 2$), and $(\partial_t g_n)_{n\in\mathbb{N}}$ is bounded in $L^{p'}(0,T;W^{-1,p'}(\Omega))$, the Aubin–Simon theorem shows that $(g_n)_{n\in\mathbb{N}}$ is relatively compact in $L^{p'}(0,T;W^{-1,p'}(\Omega))$, and that its convergence to g is strong in this space. Since $(f_n)_{n\in\mathbb{N}}$ is bounded in $L^p(0,T;W_0^{1,p}(\Omega))$, its convergence to f also holds weakly in this space. Observe that

$$\int_0^T\int_\Omega f_n(x,t)g_n(x,t)\,\mathrm{d}x\,\mathrm{d}t = \int_0^T \langle g_n(t), f_n(t)\rangle_{W^{-1,p'},W_0^{1,p}}\,\mathrm{d}t,$$

so the convergence (B.2) holds by strong/weak convergence.

We now consider Case (iii). Fix $s\in(0,1)$ such that $2s < p$. By the assumptions on $(f_n)_{n\in\mathbb{N}}$ and Lemma B.4 below, the sequence $(f_n)_{n\in\mathbb{N}}$ is bounded in $L^p(0,T;W_0^{s,2}(\Omega))$, and thus converges weakly in this space to f. Since $s > 0$, $W_0^{s,2}(\Omega)$ is compactly embedded in $L^2(\Omega)$ (we use [11, Theorem 7.1] with the extension $W_0^{s,2}(\Omega) \to W^{s,2}(\mathbb{R}^d)$ by 0 outside Ω, which is valid since $W_0^{s,2}(\Omega)$ is the closure in $W^{s,2}(\Omega)$ of compactly supported functions). Dually, $V = L^2(\Omega)$ is compactly embedded in $E = W^{-s,2}(\Omega)$. Set $F = W^{-s,2}(\Omega) + W^{-1,p'}(\Omega)$, and apply the Aubin–Simon theorem to see that $(g_n)_{n\in\mathbb{N}}$ is relatively compact in $L^{p'}(0,T;W^{-s,2}(\Omega))$.

The weak convergence of $(f_n)_{n\in\mathbb{N}}$ in $L^p(0,T;W_0^{s,2}(\Omega))$ therefore allows us to pass to the weak/strong limit in (B.2) as above. □

The following lemma is a simple interpolation result between $W_0^{1,p}(\Omega)$ and $L^\infty(\Omega)$.

Lemma B.4 *(Interpolation estimate). Let Ω be a bounded open subset of $\mathbb{R}^d$ and $p\in(1,\infty)$. If $s\in(0,1)$ and $q\in(p,\infty)$ are such that $sq<p$, then there exists C_{10} such that for all $w\in W_0^{1,p}(\Omega)$*

$$\forall w\in W_0^{1,p}(\Omega)\cap L^\infty(\Omega),\ ||w||_{W_0^{s,q}(\Omega)}\le C_{10}||w||_{L^\infty(\Omega)}^{1-\frac{p}{q}}||w||_{W_0^{1,p}(\Omega)}^{\frac{p}{q}},$$

where $W_0^{s,q}(\Omega)$ is the closure in $W^{s,q}(\Omega)$ (for the norm defined in the proof) of $C_c^\infty(\Omega)$.

Proof. We write, using the change of variable $y=x+\xi$,

$$\begin{aligned}||w||^q_{W_0^{s,q}(\Omega)}&=\int_\Omega|w(x)|^q\,\mathrm{d}x+\int_\Omega\int_\Omega\frac{|w(x)-w(y)|^q}{|x-y|^{d+sq}}\,\mathrm{d}x\,\mathrm{d}y\\&\le||w||^{q-p}_{L^\infty(\Omega)}||w||^p_{L^p(\Omega)}+2||w||^{q-p}_{L^\infty(\Omega)}\int_\Omega\int_\Omega\frac{|w(x)-w(y)|^p}{|x-y|^{d+sq}}\,\mathrm{d}x\,\mathrm{d}y\\&\le||w||^{q-p}_{L^\infty(\Omega)}||w||^p_{L^p(\Omega)}+2||w||^{q-p}_{L^\infty(\Omega)}\int_{\Omega-\Omega}|\xi|^{-d-sq}\\&\quad\times\left(\int_\Omega|w(x+\xi)-w(x)|^p\,\mathrm{d}x\right)\mathrm{d}\xi.\end{aligned}$$

But $\int_\Omega|w(x+\xi)-w(x)|^p\,\mathrm{d}x\le|\xi|^p||\nabla w||^p_{L^p(\Omega)^d}$ and $\Omega-\Omega\subset B(0,D)$ where D is the diameter of Ω. Hence

$$||w||^q_{W_0^{s,q}(\Omega)}\le||w||^{q-p}_{L^\infty(\Omega)}||w||^p_{L^p(\Omega)}+2||w||^{q-p}_{L^\infty(\Omega)}||\nabla w||^p_{L^p(\Omega)^d}\int_{B(0,D)}|\xi|^{p-d-sq}\,\mathrm{d}\xi$$

and the proof is complete since $p-d-sq>-d$. □

Appendix C. A uniqueness result

We state and prove the uniqueness of a solution to (P) when $p=2$ and

$$a(x,\nu(u),\nabla\zeta(u))=\Lambda(x)\nabla\zeta(u)\text{in }\Omega, \tag{C.1}$$

under the hypothesis that

$$\begin{gathered}\Lambda\text{ is a measurable function from }\Omega\text{ to }\mathcal{M}_d(\mathbb{R})\text{ and}\\\text{there exist }\underline{\lambda},\overline{\lambda}>0\text{ such that, for a.e. }x\in\Omega,\\\Lambda(x)\text{ is symmetric with eigenvalues in }[\underline{\lambda},\overline{\lambda}].\end{gathered}\tag{C.2}$$

J. Carrillo [10] gave a proof, based on the doubling variable technique, of the uniqueness of entropy solutions to $\partial_t\beta(u) - \Delta\zeta(u) = f$ (with an additional convective term). Although this could be extended to our framework, we provide here another proof which is shorter and simpler, using the idea due to J. Hadamard [18] of solving the dual problem. This idea has been successfully used in the case of the one-dimensional Stefan problem [7], and subsequently generalised to the higher dimensional case [17].

Note that this uniqueness result applies to the equivalent maximal monotone graph formulation (PM)–(C.1)–(C.2).

Theorem C.1. *Under Hypotheses (2.1), (C.1) and (C.2), let u_1 and u_2 be two solutions to (2.5) in the sense of Definition 2.2. Then $\beta(u_1) = \beta(u_2)$ and $\zeta(u_1) = \zeta(u_2)$.*

Remark C.2. If β and ζ do not have any common plateau, as a corollary of this theorem we see that $u_1 = u_2$. Otherwise, Theorem C.1 is optimal. Indeed, whenever a solution u takes a value in a common plateau of β and ζ, we can change this value into any other value in the same plateau without changing the fact that u is a solution.

Proof. Set $u_d = \beta(u_1) + \zeta(u_1) - \beta(u_2) - \zeta(u_2)$, and for all $(x,t) \in \Omega \times [0,T]$, define

$$q(x,t) = \begin{cases} \frac{\zeta(u_1(x,t))-\zeta(u_2(x,t))}{u_d(x,t)} & \text{if } u_d(x,t) \neq 0, \\ 0 & \text{otherwise.} \end{cases}$$

Take $\psi \in L^2(0,T;H_0^1(\Omega))$ with $\partial_t\psi \in L^2(\Omega\times(0,T))$, $\psi(\cdot,T) = 0$ and $\mathrm{div}(\Lambda\nabla\psi) \in L^2(\Omega\times(0,T))$. Approximating ψ in time by smooth functions, we see that

$$\int_0^T \langle \partial_t\beta(u_i), \psi\rangle = -\int_\Omega \beta(u^{\mathrm{ini}})(x)\psi(x,0)\,\mathrm{d}x - \int_0^T \beta(u_i)(x,t)\partial_t\psi(x,t)\,\mathrm{d}x\,\mathrm{d}t.$$

Then (2.5) gives

$$\int_0^T\int_\Omega u_d(x,t)\Big((1-q(x,t))\partial_t\psi(x,t) + q(x,t)\,\mathrm{div}(\Lambda\nabla\psi)(x,t)\Big)\,\mathrm{d}x\,\mathrm{d}t = 0. \tag{C.3}$$

For $\varepsilon \in (0,1/2)$, denote $q_\varepsilon = (1-2\varepsilon)q + \varepsilon$. Since $0 \le q \le 1$ we have $\varepsilon \le q_\varepsilon \le 1-\varepsilon$ and

$$\frac{(q_\varepsilon - q)^2}{q_\varepsilon} \le \varepsilon \quad \text{and} \quad \frac{(q_\varepsilon - q)^2}{1-q_\varepsilon} \le \varepsilon. \tag{C.4}$$

Let ψ_ε be given by Lemma C.3 below, with $g = q_\varepsilon$ and some $w \in C_c^\infty(\Omega\times(0,T))$. Substituting ψ by ψ_ε in (C.3) and using (C.7),

$$\begin{aligned}&\left|\int_0^T\int_\Omega u_d(x,t)w(x,t)\,\mathrm{d}x\,\mathrm{d}t\right| \\ &\quad\le \left|\int_0^T\int_\Omega u_d(x,t)(q_\varepsilon(x,t)-q(x,t))(\mathrm{div}(\Lambda\nabla\psi_\varepsilon)(x,t) - \partial_t\psi_\varepsilon(x,t))\,\mathrm{d}x\,\mathrm{d}t\right|. \end{aligned}\tag{C.5}$$

The Cauchy–Schwarz inequality, (C.8) and (C.4) imply that

$$\begin{aligned}
&\left[\int_0^T\int_\Omega u_d(x,t)(q_\varepsilon(x,t)-q(x,t))(\operatorname{div}(\Lambda\nabla\psi_\varepsilon)(x,t)-\partial_t\psi_\varepsilon(x,t))\,\mathrm{d}x\,\mathrm{d}t\right]^2\\
&\quad\le 2\left(\int_0^T\int_\Omega u_d(x,t)^2\frac{(q(x,t)-q_\varepsilon(x,t))^2}{q_\varepsilon(x,t)}\,\mathrm{d}x\,\mathrm{d}t\right)\\
&\qquad\times\left(\int_0^T\int_\Omega q_\varepsilon(x,t)\Big(\operatorname{div}(\Lambda\nabla\psi_\varepsilon)(x,t)\Big)^2\,\mathrm{d}x\,\mathrm{d}t\right)\\
&\qquad+2\left(\int_0^T\int_\Omega u_d(x,t)^2\frac{(q(x,t)-q_\varepsilon(x,t))^2}{1-q_\varepsilon(x,t)}\,\mathrm{d}x\,\mathrm{d}t\right)\\
&\qquad\times\left(\int_0^T\int_\Omega (1-q_\varepsilon(x,t))\Big(\partial_t\psi_\varepsilon(x,t)\Big)^2\,\mathrm{d}x\,\mathrm{d}t\right)\\
&\quad\le \varepsilon C_0\left(\|\nabla w\|^2_{L^2(\Omega\times(0,T))}+\|w\|^2_{L^2(\Omega\times(0,T))}+\|\partial_t w\|^2_{L^2(\Omega\times(0,T))}\right)\\
&\qquad\times\int_0^T\int_\Omega u_d(x,t)^2\,\mathrm{d}x\,\mathrm{d}t. \qquad\qquad (C.6)
\end{aligned}$$

The right-hand side of (C.6) vanishes as $\varepsilon\to 0$, and therefore so does left-hand side of (C.5), giving

$$\int_0^T\int_\Omega u_d(x,t)w(x,t)\,\mathrm{d}x\,\mathrm{d}t=0.$$

Since this holds for any function $w\in C_c^\infty(\Omega\times(0,T))$, we get that $u_d(x,t)=0$ for a.e. $(x,t)\in\Omega\times(0,T)$. Hence $\beta(u_1)-\beta(u_2)=-(\zeta(u_1)-\zeta(u_2))$ and, since β and ζ are non-decreasing, the proof of the theorem is complete. □

The following lemma ensures the existence of the function ψ, used in the proof of Theorem C.1.

Lemma C.3. *Let $T>0$, and let Ω be a bounded open subset of $\mathbb{R}^d$ ($d\in\mathbb{N}$). Assume Hypothesis (C.2). Let $w\in C_c^\infty(\Omega\times(0,T))$ and $g\in L^\infty(\Omega\times(0,T))$ such that $g(x,t)\in[g_{\min},1-g_{\min}]$ for a.e. $(x,t)\in\Omega\times(0,T)$, where $g_{\min}$ is a fixed number in $(0,\frac{1}{2})$. Then there exists a function ψ such that:*

(i) $\psi\in L^\infty(0,T;H_0^1(\Omega))$, $\partial_t\psi\in L^2(\Omega\times(0,T))$, $\operatorname{div}(\Lambda\nabla\psi)\in L^2(\Omega\times(0,T))$ *(this implies $\psi\in C([0,T];L^2(\Omega))$),*
(ii) $\psi(\cdot,T)=0$,
(iii) ψ *satisfies*

$$\begin{aligned}
&(1-g(x,t))\partial_t\psi(x,t)+g(x,t)\operatorname{div}(\Lambda\nabla\psi)(x,t)=w(x,t)\\
&\textit{for a.e. } (x,t)\in\Omega\times(0,T), \qquad\qquad (C.7)
\end{aligned}$$

(iv) *there exists* $C_0 > 0$, *depending only on* T, $\mathrm{diam}(\Omega)$, $\underline{\lambda}$ *and* $\overline{\lambda}$ *(and* not *on* $g_{\min}$*), such that*

$$\int_0^T \int_\Omega \left((1 - g(x,t))\Big(\partial_t \psi(x,t)\Big)^2 + g(x,t)\Big(\mathrm{div}(\Lambda\nabla\psi)(x,t)\Big)^2 \right) \mathrm{d}x\,\mathrm{d}t$$
$$\leq C_0 \left(\|\nabla w\|^2_{L^2(\Omega\times(0,T))} + \|w\|^2_{L^2(\Omega\times(0,T))} + \|\partial_t w\|^2_{L^2(\Omega\times(0,T))} \right). \tag{C.8}$$

Proof. After dividing through by g, observe that (C.7) is equivalent to

$$\Phi(x,t)\partial_t\psi(x,t) + \mathrm{div}(\Lambda(x)\nabla\psi(x,t)) = f(x,t), \tag{C.9}$$

where $\Phi \in L^\infty(\Omega \times (0,T))$ satisfies $0 < \phi_* \leq \Phi(x,t) \leq \phi^*$ for a.e. $(x,t) \in \Omega \times (0,T)$ and $f \in L^\infty(\Omega \times (0,T))$. We first show the existence of a solution ψ to (C.9) satisfying (i) and (ii).

Let $W := \{v \in C^0([0,T]; H^1_0(\Omega)) \,|\, \partial_t v \in L^2(\Omega \times (0,T)) \text{ and } v(\cdot,T) = 0\}$. Define $T : W \to W$, where $T(v) = u$ is such that for all $w \in L^2(0,T; H^1_0(\Omega))$,

$$\int_0^T \int_\Omega \big(\phi^* w(x,t)\partial_t u(x,t) - \Lambda(x)\nabla u(x,t)\cdot\nabla w(x,t)\big)\,\mathrm{d}x\,\mathrm{d}t$$
$$= \int_0^T \int_\Omega \big(f(x,t) + (\phi^* - \Phi(x,t))\partial_t v(x,t)\big)\, w(x,t)\,\mathrm{d}x\,\mathrm{d}t. \tag{C.10}$$

Existence of such a $u \in W$ is assured thanks to Lemma C.4 below. Endowing W with the norm

$$\|v\|_W := \left[\sup_{\tau\in[0,T]} \left(\|\partial_t v\|^2_{L^2(\Omega\times(\tau,T))} + \frac{\lambda}{\phi^*}\|\nabla v(\cdot,\tau)\|^2_{L^2(\Omega)^d} \right) \right]^{1/2},$$

dividing (C.10) by ϕ^*, noticing that $\sup_{(x,t)\in\Omega\times(0,T)} \left| \frac{\phi^* - \Phi(x,t)}{\phi^*} \right| \leq \frac{\phi^* - \phi_*}{\phi^*} < 1$ and using (C.15), we see that T is a contraction. It therefore has a unique fixed point $\psi \in W$ that satisfies (i)–(iii).

It remains to verify (C.8). Taking $s, \tau \in [0,T]$, we have

$$\int_s^\tau \int_\Omega w(x,t)\,\mathrm{div}(\Lambda\nabla\psi)(x,t)\,\mathrm{d}x\,\mathrm{d}t = -\int_s^\tau \int_\Omega \Lambda(x)\nabla w(x,t)\cdot\nabla\psi(x,t)\,\mathrm{d}x\,\mathrm{d}t,$$

and

$$\int_s^\tau \int_\Omega w(x,t)\partial_t\psi(x,t)\,\mathrm{d}x\,\mathrm{d}t = \int_\Omega (w(x,\tau)\psi(x,\tau) - w(x,s)\psi(x,s))\,\mathrm{d}x$$
$$- \int_s^\tau \int_\Omega \psi(x,t)\partial_t w(x,t)\,\mathrm{d}x\,\mathrm{d}t.$$

Multiplying (C.7) by $\partial_t\psi(x,t)+\operatorname{div}(\Lambda\nabla\psi)(x,t)$, integrating on $\Omega\times(s,T)$ for $s\in[0,T]$, and using (C.16), $\psi(\cdot,T)=0$ and $\nabla\psi(\cdot,T)=0$, we obtain

$$\begin{aligned}
&\frac{1}{2}\int_\Omega \Lambda(x)\nabla\psi(x,s)\cdot\nabla\psi(x,s)\,\mathrm{d}x\\
&\qquad+\int_s^T\int_\Omega\Big((1-g(x,t))\Big(\partial_t\psi(x,t)\Big)^2+g(x,t)\Big(\operatorname{div}(\Lambda\nabla\psi)(x,t)\Big)^2\Big)\,\mathrm{d}x\,\mathrm{d}t\\
&\quad=-\int_s^T\int_\Omega\Lambda(x)\nabla w(x,t)\cdot\nabla\psi(x,t)\,\mathrm{d}x\,\mathrm{d}t-\int_\Omega w(x,s)\psi(x,s)\,\mathrm{d}x\\
&\qquad-\int_s^T\int_\Omega\psi(x,t)\partial_t w(x,t)\,\mathrm{d}x\,\mathrm{d}t. \qquad\text{(C.11)}
\end{aligned}$$

Integrating (C.11) with respect to $s\in(0,T)$ leads to

$$\begin{aligned}
&\frac{1}{2}\int_0^T\int_\Omega\Lambda(x)\nabla\psi(x,s)\cdot\nabla\psi(x,s)\,\mathrm{d}x\,\mathrm{d}s\le T\int_0^T\int_\Omega|\Lambda(x)\nabla w(x,t)\cdot\nabla\psi(x,t)|\,\mathrm{d}x\,\mathrm{d}t\\
&\quad+\int_0^T\int_\Omega|w(x,s)\psi(x,s)|\,\mathrm{d}x\,\mathrm{d}s+T\int_0^T\int_\Omega|\psi(x,t)\partial_t w(x,t)|\,\mathrm{d}x\,\mathrm{d}t. \qquad\text{(C.12)}
\end{aligned}$$

We then apply the Cauchy–Schwarz and Poincaré inequalities, which leads to

$$\begin{aligned}
&\frac{\underline{\lambda}}{2}\|\nabla\psi\|_{L^2(\Omega\times(0,T))}\\
&\quad\le T\overline{\lambda}\,\|\nabla w\|_{L^2(\Omega\times(0,T))}+\operatorname{diam}(\Omega)\left(\|w\|_{L^2(\Omega\times(0,T))}+T\,\|\partial_t w\|_{L^2(\Omega\times(0,T))}\right). \qquad\text{(C.13)}
\end{aligned}$$

Letting $s=0$ in (C.11), recalling that $w(\cdot,0)=0$, and using (C.13) gives

$$\begin{aligned}
&\int_0^T\int_\Omega\Big((1-g(x,t))(\partial_t\psi(x,t))^2+g(x,t)(\operatorname{div}\Lambda\nabla\psi(x,t))^2\Big)\,\mathrm{d}x\,\mathrm{d}t\\
&\quad\le\frac{2}{\underline{\lambda}}\left(\overline{\lambda}\,\|\nabla w\|_{L^2(\Omega\times(0,T))}+\operatorname{diam}(\Omega)\,\|\partial_t w\|_{L^2(\Omega\times(0,T))}\right)\\
&\qquad\times\left(T\overline{\lambda}\,\|\nabla w\|_{L^2(\Omega\times(0,T))}+\operatorname{diam}(\Omega)(\|w\|_{L^2(\Omega\times(0,T))}+T\,\|\partial_t w\|_{L^2(\Omega\times(0,T))}\right),
\end{aligned}$$

which implies (C.8). □

The following lemma states the time regularity of the solution of a linear backwards parabolic problem with sufficiently regular data. It may be that this lemma can be proved by using the Hille–Yoshida theorem, since the regularity of the solution is coherent with those of the Hille–Yoshida theory, but this exact result, with low regularity assumptions on Ω or Λ, does not seem to exist in the literature. We propose a self-contained proof, which is probably shorter than checking that the Hille–Yoshida framework applies.

Lemma C.4. *Let $T > 0$, and let Ω be a bounded open subset of $\mathbb{R}^d$ ($d \in \mathbb{N}$). Assume Hypothesis (C.2). Let $h \in L^2(\Omega \times (0,T))$, and let $u \in L^2(0,T; H_0^1(\Omega))$ with $\partial_t u \in L^2(0,T; H^{-1}(\Omega))$ such that $u(\cdot, T) = 0$ be the standard weak solution of the backwards problem $\partial_t u + \operatorname{div}(\Lambda \nabla u) = h$, that is*

$$\begin{aligned}&\forall v \in L^2(0,T; H_0^1(\Omega)),\\ &\int_0^T \left(\langle \partial_t u(\cdot,t), v(\cdot,t)\rangle_{H^{-1},H_0^1} - \int_\Omega \Lambda(x)\nabla u(x,t)\cdot\nabla v(x,t)\,\mathrm{d}x \right)\mathrm{d}t\\ &\quad = \int_0^T \int_\Omega h(x,t)v(x,t)\,\mathrm{d}x\,\mathrm{d}t.\end{aligned} \tag{C.14}$$

Then $\partial_t u \in L^2(\Omega \times (0,T))$, $\operatorname{div}(\Lambda\nabla u) \in L^2(\Omega\times(0,T))$, $u \in C^0([0,T]; H_0^1(\Omega))$ and

$$\sup_{t_0 \in [0,T]} \left(\|\partial_t u\|^2_{L^2(\Omega\times(t_0,T))} + \int_\Omega \Lambda(x)\nabla u(x,t_0)\cdot\nabla u(x,t_0)\,\mathrm{d}x \right) \le \|h\|^2_{L^2(\Omega\times(0,T))}. \tag{C.15}$$

Furthermore, for all $s < \tau \in [0,T]$,

$$\begin{aligned}\int_s^\tau \int_\Omega \partial_t u(x,t)\operatorname{div}(\Lambda\nabla u)(x,t)\,\mathrm{d}x\,\mathrm{d}t &= -\frac{1}{2}\int_\Omega \Lambda(x)\nabla u(x,\tau)\cdot\nabla u(x,\tau)\,\mathrm{d}x\\ &\quad + \frac{1}{2}\int_\Omega \Lambda(x)\nabla u(x,s)\cdot\nabla u(x,s)\,\mathrm{d}x.\end{aligned} \tag{C.16}$$

Proof. Let $\rho \in C_c^\infty(\mathbb{R})$ with support in $[-1,0]$, such that $\rho \ge 0$ and $\int_{-1}^0 \rho(s)\,\mathrm{d}s = 1$. For $n \in \mathbb{N}$, define $\rho_n(s) = n\rho(ns)$ and $u_n(x,t) = \int_0^T \rho_n(t-s)u(x,s)\,\mathrm{d}s$. Take

$$v(x,t) = \int_0^T \rho_n(s-t)\partial_t u_n(x,s)\,\mathrm{d}s$$

as test function in (C.14). Since v is a regular function with respect to time and $v(\cdot,0) = 0$ (thanks to the support of ρ), we obtain $T_1 + T_2 = T_3$, with

$$\begin{aligned}T_1 &= \int_0^T \int_\Omega u(x,t)\int_0^T \rho_n'(s-t)\partial_t u_n(x,s)\,\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}t = \int_0^T \int_\Omega (\partial_t u_n(x,s))^2\,\mathrm{d}x\,\mathrm{d}s,\\ T_2 &= -\int_0^T \int_\Omega \Lambda(x)\nabla u(x,t)\cdot\int_0^T \rho_n(s-t)\nabla\partial_t u_n(x,s)\,\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}t\\ &= -\frac{1}{2}\int_0^T \int_\Omega \partial_t(\Lambda(x)\nabla u_n(x,\cdot)\cdot\nabla u_n(x,\cdot))(s)\,\mathrm{d}x\,\mathrm{d}s\\ &= -\frac{1}{2}\left(\int_\Omega \Lambda(x)\nabla u_n(x,T)\cdot\nabla u_n(x,T)\,\mathrm{d}x - \int_\Omega \Lambda(x)\nabla u_n(x,0)\cdot\nabla u_n(x,0)\,\mathrm{d}x \right)\end{aligned}$$

and

$$T_3=\int_0^T\int_\Omega h(x,t)\int_0^T\rho_n(s-t)\partial_t u_n(x,s)\,\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}t=\int_0^T\int_\Omega h_n(x,s)\partial_t u_n(x,s)\,\mathrm{d}s\,\mathrm{d}x,$$

where $h_n(x,s)=\int_0^T h(x,t)\rho_n(s-t)\,\mathrm{d}t$. Observing that $u_n(\cdot,T)=0$ and $\|h_n\|_{L^2(\Omega\times(0,T))}\leq\|h\|_{L^2(\Omega\times(0,T))}$, Young's inequality on the right-hand side of $T_1+T_2=T_3$ yields

$$\int_0^T\int_\Omega(\partial_t u_n(x,s))^2\,\mathrm{d}x\,\mathrm{d}s+\int_\Omega\Lambda(x)\nabla u_n(x,0)\cdot\nabla u_n(x,0)\,\mathrm{d}x\leq\|h\|^2_{L^2(\Omega\times(0,T))}. \qquad \text{(C.17)}$$

Hence, $(\partial_t u_n)_{n\in\mathbb{N}}$ is bounded in $L^2(\Omega\times(0,T))$. Since $u_n\to u$ in $L^2(\Omega\times(0,T))$, this shows that $\partial_t u\in L^2(\Omega\times(0,T))$. The PDE (C.14) gives $\mathrm{div}(\Lambda\nabla u)=\partial_t u-h\in L^2(\Omega\times(0,T))$. Finally, we obtain (C.15) by repeating the reasoning leading to (C.17), starting with an arbitrary time $t_0\in[0,T]$ instead of 0 and by passing to the weak limits in the corresponding inequalities. Equation (C.16) is established by repeating the above computations using the same regularization in time. □

References

[1] H.W. Alt, S. Luckhaus, Quasilinear elliptic–parabolic differential equations, Math. Z. 183 (3) (1983) 311–341.

[2] B. Andreianov, M. Bendahmane, K.H. Karlsen, S. Ouaro, Well-posedness results for triply nonlinear degenerate parabolic equations, J. Differential Equations 247 (1) (2009) 277–302.

[3] B. Andreianov, C. Cancès, A. Moussa, A nonlinear time compactness result and applications to discretization of degenerate parabolic–elliptic PDEs, https://hal.archives-ouvertes.fr/hal-01142499, 2015, arXiv:1504.03891, p. 46p, submitted for publication.

[4] J.-P. Aubin, Un théorème de compacité, C. R. Acad. Sci. Paris 256 (1963) 5042–5044.

[5] P. Bénilan, M.G. Crandall, The continuous dependence on φ of solutions of $u_t-\Delta\varphi(u)=0$, Indiana Univ. Math. J. 30 (2) (1981) 161–177.

[6] M. Bertsch, P. de Mottoni, L.A. Peletier, Degenerate diffusion and the Stefan problem, Nonlinear Anal. 8 (11) (1984) 1311–1336.

[7] M. Bertsch, R. Kersner, L.A. Peletier, Positivity versus localization in degenerate diffusion equations, Nonlinear Anal. 9 (9) (1985) 987–1008.

[8] D. Blanchard, A. Porretta, Stefan problems with nonlinear diffusion and convection, J. Differential Equations 210 (2) (2005) 383–428.

[9] F. Boyer, P. Fabrie, Mathematical Tools for the Study of the Incompressible Navier–Stokes Equations and Related Models, Applied Mathematical Sciences, vol. 183, Springer, New York, 2013.

[10] J. Carrillo, Entropy solutions for nonlinear degenerate problems, Arch. Ration. Mech. Anal. 147 (4) (1999) 269–361.

[11] E. Di Nezza, G. Palatucci, E. Valdinoci, Hitchhiker's guide to the fractional Sobolev spaces, Bull. Sci. Math. 136 (5) (2012) 521–573.

[12] E. DiBenedetto, Degenerate Parabolic Equations, Universitext, Springer-Verlag, New York, 1993.

[13] J. Droniou, Intégration et espaces de Sobolev à valeurs vectorielles, Polycopiés de l'Ecole Doctorale de Mathématiques-Informatique de Marseille, available at http://www-gm3.univ-mrs.fr/polys, 2001.

[14] J. Droniou, R. Eymard, Uniform-in-time convergence result of numerical methods for non-linear parabolic equations, Numer. Math. (2015), http://dx.doi.org/10.1007/s00211-015-0733-6, in press.

[15] J. Droniou, R. Eymard, T. Gallouët, 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) (2013) 2395–2432.

[16] J. Droniou, K.S. Talbot, On a miscible displacement model in porous media flow with measure data, SIAM J. Math. Anal. 46 (5) (2014) 3158–3175.

[17] M. Guedda, D. Hilhorst, M.A. Peletier, Disappearing interfaces in nonlinear diffusion, Adv. Math. Sci. Appl. 7 (2) (1997) 695–710.

[18] J. Hadamard, Lectures on Cauchy's Problem in Linear Partial Differential Equations, Dover Phoenix, New York, NY, USA, 1923.

[19] A.V. Kazhikhov, Recent Developments in the Global Theory of Two-Dimensional Compressible Navier–Stokes Equations, Seminar on Mathematical Sciences, vol. 25, Keio University, Department of Mathematics, Yokohama, 1998.

[20] I.C. Kim, N. Požár, Nonlinear elliptic–parabolic problems, Arch. Ration. Mech. Anal. 210 (3) (2013) 975–1020.

[21] J. Kinnunen, M. Parviainen, Stability for degenerate parabolic equations, Adv. Calc. Var. 3 (1) (2010) 29–48.

[22] J. Leray, J.-L. Lions, Quelques résultats de Višik sur les problèmes elliptiques non linéaires par les méthodes de Minty–Browder, Bull. Soc. Math. France 93 (1965) 97–107.

[23] J.-L. Lions, Sur certaines équations paraboliques non linéaires, Bull. Soc. Math. France 93 (1965) 155–175.

[24] T. Lukkari, M. Parviainen, Stability of degenerate parabolic Cauchy problems, Commun. Pure Appl. Anal. 14 (1) (2015) 201–216.

[25] W. Merz, P. Rybka, Strong solutions to the Richards equation in the unsaturated zone, J. Math. Anal. Appl. 371 (2) (2010) 741–749.

[26] G.J. Minty, On a "monotonicity" method for the solution of non-linear equations in Banach spaces, Proc. Natl. Acad. Sci. USA 50 (6) (1963) 1038.

[27] A. Moussa, Some variants of the classical Aubin–Lions lemma, arXiv:1401.7231, 2014.

[28] F. Otto, L^1-contraction and uniqueness for quasilinear elliptic–parabolic equations, J. Differential Equations 131 (1) (1996) 20–38.

[29] L.A. Richards, Capillary conduction of liquids through porous mediums, J. Appl. Phys. 1 (5) (1931) 318–333.

[30] J. Simon, Compact sets in the space $L^p(0,T;B)$, Ann. Mat. Pura Appl. (4) 146 (1987) 65–96.

[31] G. Stampacchia, Le problème de Dirichlet pour les équations elliptiques du second ordre à coefficients discontinus, Ann. Inst. Fourier (Grenoble) 15 (fasc. 1) (1965) 189–258.