GLOBAL WELL-POSEDNESS FOR A TIME-FRACTIONAL DOUBLY NONLINEAR EQUATION

GORO AKAGI, GIACOMO ENRICO SODINI, AND ULISSE STEFANELLI

Abstract. We consider a time-fractional parabolic equation of doubly nonlinear type, featuring nonlinear terms both inside and outside the differential operator in time. The main nonlinearities are maximal monotone graphs, without restrictions on the growth. In addition, a Lipschitz continuous perturbation is considered. The existence of global weak solutions is obtained via a regularization and Galerkin approximation method. Uniqueness is also discussed under some additional assumptions.

1. Introduction

This paper is concerned with the nonlinear time-fractional parabolic equation

$$\partial_t^\theta(\alpha(u)-\alpha(u_0)) - \Delta u + \beta(u) \ni g(x,t,u). \tag{1.1}$$

Here, the symbol ∂_t^θ stands for the *Riemann–Liouville fractional derivative* of order $\theta \in (0,1)$ which is defined by

$$\partial_t^\theta f(t) := \frac{1}{\Gamma(1-\theta)} \frac{\mathrm{d}}{\mathrm{d}s} \int_0^t (t-s)^{-\theta} f(s)\,\mathrm{d}t \quad \forall t > 0$$

where Γ is the Euler Γ-function. Hence if $\alpha(u)|_{t=0} = \alpha(u_0)$, then the Riemann–Liouville derivative $\partial_t^\theta(\alpha(u)-\alpha(u_0))$ coincides of the *Caputo derivative* of $\alpha(u)$ of order $\theta \in (0,1)$ in a weak sense.

Equation (1.1) is posed in the space-time cylinder $Q := \Omega \times (0,T)$, where $T > 0$ is a final time and Ω is a nonempty, open, connected, bounded set of $\mathbb{R}^n$ with Lipschitz boundary $\partial\Omega$. The nonlinearities α, $\beta : \mathbb{R} \to 2^{\mathbb{R}}$ are maximal monotone graphs, with no growth restrictions, whereas $g(x,t,\cdot)$ is a Lipschitz continuous function. The initial value $\alpha(u_0)$ is given and the equation is complemented by homogeneous Dirichlet boundary conditions.

The occurrence of nonlinearities both inside and outside the differential operator in time ∂_t^θ qualifies equation (1.1) as of *doubly nonlinear* type. Doubly nonlinear problems occur frequently in connection with applications, especially in relation to nonlinear diffusion phenomena. In the nonfractional case, for different choices of the map α, equation (1.1) may arise in connection to the two-phase Stefan problem (for $\alpha = \mathrm{id} + H$, where H is the Heaviside graph), the porous-medium equation (for $\alpha(u) = |u|^{p-2}u$ for some $p \in (1,2)$) and the Hele–Shaw model (for $\alpha = H$), see [42] for a detailed discussion.

Also motivated by their relevance in applications, doubly nonlinear problems are a mainstay in evolution-equation theory. In the classical, nonfractional case, existence results for doubly nonlinear equations can be traced back to Grange–Mignot [21]. Besides the classical references by Barbu [12], DiBenedetto–Showalter [17], Alt–Luckhaus [10], and

---

2020 *Mathematics Subject Classification.* Promary: 35K55.

Key words and phrases. Fractional doubly nonlinear equation, approximation, existence, uniqueness.

Bernis [13], one may also refer to [2, 3, 4, 5, 19, 23, 24, 25, 29, 34, 35, 36], among many others. See also [9, 37] for some global variational approaches and [8, 14, 26, 27, 30] for a collection of recent results.

To the best of our knowledge, existence for time-fractional parabolic doubly nonlinear equations has not been proved yet. This note is intended to fill this gap. Our main result is the well-posedness of a variational formulation of equation (1.1), where the elliptic operator is considered in the classical weak sense. Note that, in case α is linear, existence of solutions to equation (1.1) has already been obtained in [6], see also [28, 7]. In fact, the specific case of porous media has already been considered in [31, 32, 33]. In addition, a number of results on long-time behavior and decay is available, see [1, 18, 39, 40], where nonetheless existence of solutions is just assumed. Furthermore, in [15], the asymptotic behavior of solutions for such time-fractional porous medium equations has also been studied in detail. The reader is also referred to [20, 38] for some alternative approaches based on viscosity solutions.

To prove existence of a solution to (1.1) we proceed by regularization and approximation. The maximal monotone graphs α and β are replaced by their Yosida approximations and α is made strongly monotone by adding a multiple of the identity to it. Such a regularized problem is then tackled by a Galerkin discretization method. A suite of a-priori estimates allows the passage to the limit, first in the Galerkin approximation and then in the regularization, bringing to the existence proof. In the classical nonfractional case, a pivotal tool in the analysis of doubly nonlinear problems driven by subdifferential operators is the availability of a local chain-rule equality, which eventually allows for obtaining a-priori estimates. In the time-fractional setting, the local chain-rule equality is replaced by a nonlocal chain-rule inequality, which makes the analysis more challenging.

Uniqueness is generally not to be expected for doubly nonlinear problems, see [17]. We are nonetheless in the position of providing a uniqueness result under some stronger assumptions, namely, for α strongly monotone and β Lipschitz continuous.

The plan of the paper is as follows. We present our setting and state the well-posedness result, i.e., Theorem 2.1, in Section 2. The uniqueness of solutions is proved in Section 3. The existence proof is then presented in Section 4. A regularization and Galerkin approximation of the problem is introduced in Sections 4.1–4.2. We then prove a-priori estimates and pass to the limit in the Galerkin approximation in Sections 4.3–4.4. Additional estimates for the regularized problem are obtained in Section 4.5 and we eventually remove the regularization by a limit passage in Sections 4.6–4.7.

2. MAIN RESULT

We devote this section to the statement of the main well-posedness result, namely Theorem 2.1. Let us start by introducing some notation.

2.1. **Functional analytic framework**

Recall that we ask for

(A0): Let $\Omega \subset \mathbb{R}^n$ be a nonempty bounded Lipschitz domain, and let $T \in (0, \infty)$.

In the following, we set $Q := \Omega \times (0, T)$ and we use the spaces

$$H := L^2(\Omega), \quad V := H^1_0(\Omega), \quad V^* := H^{-1}(\Omega)$$

where V^* is the dual of V, so that (V, H, V^*) is a classical Hilbert triplet. Note that the choice of the space V corresponds to the case of homogeneous Dirichlet boundary conditions. Other boundary conditions can also be treated, by redefining V to be a closed subset of $H^1(\Omega)$ containing $H^1_0(\Omega)$. The symbol $(\cdot,\cdot)$ stands for the standard scalar product in H whereas $\langle\cdot,\cdot\rangle$ is the duality pairing between V^* and V. We denote by $\|\cdot\|$ the norm in H, by $\|\cdot\|_E$ the norm in the generic Banach space E, and by $c_V > 0$ the norm of the injection $V \subset H$, namely, $\|u\| \le c_V\|u\|_V$ for all $u \in V$. We note that c_V is given as $c_V = \lambda_1(\Omega)^{-1/2}$, where $\lambda_1(\Omega)$ stands for the principal eigenvalue of the Dirichlet Laplacian in Ω. We set $A : V \to V^*$ to be the Riesz mapping given by

$$\langle Au, v\rangle := \int_\Omega \nabla u(x)\cdot\nabla v(x)\,\mathrm{d}x \quad \forall u, v \in V.$$

In particular, it holds $\|u\|_V^2 = \langle Au, u\rangle$ for all $u \in V$.

2.2. Maximal monotone operators

Given any maximal monotone graph $\gamma : \mathbb{R} \to 2^{\mathbb{R}}$ with $0 \in \gamma(0)$, for all $r \in D(\gamma) := \{r \in \mathbb{R} : \gamma(r) \neq \emptyset\}$ we define by $\gamma^\circ(r)$ the point in the closed interval $\gamma(r)$ which is closest to 0. One can uniquely find (see e.g. [16, Example 2.8.1, p. 43]) a proper, lower semicontinuous and convex function $\widehat{\gamma} : D(\widehat{\gamma}) \subset \mathbb{R} \to [0,\infty]$ such that $\gamma = \partial\widehat{\gamma}$, where $0 = \widehat{\gamma}(0) = \min\widehat{\gamma}$ and ∂ denotes the subdifferential in the sense of convex analysis [16]. The symbol $D(\widehat{\gamma})$ indicates the *effective domain* of $\widehat{\gamma}$, namely, $D(\widehat{\gamma}) := \{r \in \mathbb{R} : \widehat{\gamma}(r) < \infty\}$. In particular, we have that $\widehat{\gamma} \ge 0$, as well as $\widehat{\gamma}^* \ge 0$, where $\widehat{\gamma}^*$ is the Legendre–Fenchel conjugate of $\widehat{\gamma}$, namely $\widehat{\gamma}^*(r) := \sup_{s\in\mathbb{R}}(rs - \widehat{\gamma}(s))$ for all $r \in \mathbb{R}$. The Fenchel identity $\widehat{\gamma}(s) + \widehat{\gamma}^*(r) = rs$ for all $r \in \gamma(s)$ or, equivalently, all $s \in \partial\widehat{\gamma}^*(r)$ in particular ensures that $\gamma^{-1} = \partial\widehat{\gamma}^*$.

We indicate by $\psi_\gamma : H \to [0,\infty]$ the convex, proper, and lower semicontinuous functional given by

$$\psi_\gamma(v) := \begin{cases} \displaystyle\int_\Omega \widehat{\gamma}(v(x))\,\mathrm{d}x & \text{if } \widehat{\gamma}\circ v \in L^1(\Omega), \\ \infty & \text{else} \end{cases} \tag{2.1}$$

for $v \in H$. Moreover, we set $\Psi_\gamma : L^2(0,T;H) \to [0,\infty]$ to be

$$\Psi_\gamma(v) := \begin{cases} \displaystyle\int_0^T \psi_\gamma(v)\,\mathrm{d}t = \int_0^T\int_\Omega \widehat{\gamma}(v(x,t))\,\mathrm{d}x\,\mathrm{d}t & \text{if } \widehat{\gamma}\circ v \in L^1(Q), \\ \infty & \text{else} \end{cases} \tag{2.2}$$

for $v \in L^2(0,T;H)$.

Let us consider the subdifferential $\gamma_H := \partial\psi_\gamma : H \to 2^H$ and recall that $v \in \gamma_H(u)$ if and only if $u \in D(\psi_\gamma)$ and

$$(v, w - u) \le \psi_\gamma(w) - \psi_\gamma(u) \quad \forall w \in H.$$

In fact, the subdifferential γ_H is characterized (cf. [16, Proposition 2.16, p. 47]) as, for $u, v \in H$,

$$v \in \gamma_H(u) \quad\Leftrightarrow\quad v(x) \in \gamma(u(x)) \ \text{ for a.e. } x \in \Omega$$

with domain $D(\gamma_H) = \{u \in H : \exists v \in H \text{ with } v(x) \in \gamma(u(x)) \text{ for a.e. } x \in \Omega\}$. In the following, for any $u \in D(\gamma_H)$ we use the symbol $\gamma_H^\circ(u)$ to indicate the element of minimal norm in the nonempty, convex, and closed set $\gamma_H(u)$. One can also consider the subdifferential $\partial\Psi_\gamma : L^2(0,T;H) \to 2^{L^2(0,T;H)}$ and check that $D(\partial\Psi_\gamma) = \{u \in L^2(0,T;H) :$

$\exists v \in L^2(0,T;H)$ with $v \in \gamma(u)$ a.e. in $Q\}$ and that, for all $u \in D(\Psi_\gamma)$ one has $\partial\Psi_\gamma(u) = \{v \in L^2(0,T;H) :\ v \in \gamma(u)$ a.e. in $Q\}$.

2.3. **Convergence of maximal monotone operators and convolution**

Given a sequence of maximal monotone maps $(\gamma_h)_h$ for $h > 0$ with $0 \in \gamma_h(0)$, one says that $\partial\Psi_{\gamma_h} \to \partial\Psi_\gamma$ in the *graph sense in* $L^2(0,T;H)$ as $h \to 0$ if, for all u, $v \in L^2(0,T;H)$ with $v \in \partial\Psi_\gamma(u)$, one can find two sequences $(u_h)_h$, $(v_h)_h$ in $L^2(0,T;H)$ with $v_h \in \partial\Psi_{\gamma_h}(u_h)$ such that $u_h \to u$ and $v_h \to v$ in $L^2(0,T;H)$ [11, Definition 3.58, p. 360]. A straightforward extension of [11, Proposition 3.59, p. 361] guarantees that

$$\begin{aligned}
&\partial\Psi_{\gamma_h} \to \partial\Psi_\gamma \text{ in the graph sense in } L^2(0,T;H),\\
&v_h \in \partial\Psi_{\gamma_h}(u_h),\ u_h \rightharpoonup u \text{ in } L^2(0,T;H),\ v_h \rightharpoonup v \text{ in } L^2(0,T;H),\\
&\text{and } \limsup_{h\to 0} \int_0^T (v_h,u_h)\,\mathrm{d}t \le \int_0^T (v,u)\,\mathrm{d}t\\
&\quad \Rightarrow\ v \in \partial\Psi_\gamma(u) \text{ and } \int_0^T (v_h,u_h)\,\mathrm{d}t \to \int_0^T (v,u)\,\mathrm{d}t.
\end{aligned} \tag{2.3}$$

Under the assumption $0 \in \gamma_h(0)$, the graph convergence $\partial\Psi_{\gamma_h} \to \partial\Psi_\gamma$ is equivalent to the *Mosco convergence* $\Psi_{\gamma_h} \to \Psi_\gamma$ in $L^2(0,T;H)$ of the corresponding functionals [11, Theorem 3.66, p. 373], namely, to the following two conditions

$$\Psi_\gamma(u) \le \liminf_{h\to 0} \Psi_{\gamma_h}(u_h) \quad \forall u_h \rightharpoonup u \text{ in } L^2(0,T;H), \text{ and} \tag{2.4}$$

$$\forall y \in L^2(0,T;H)\ \exists y_h \to y \text{ in } L^2(0,T;H) \text{ such that } \Psi_{\gamma_h}(y_h) \to \Psi_\gamma(y). \tag{2.5}$$

Given any $a \in L^p(0,T)$ and $b \in L^q(0,T)$ (or $b \in L^q(0,T;E)$ with E being a Banach space) with $1/p + 1/q = 1 + 1/r$ for some $1 \le p,q,r \le \infty$, in the following, we indicate the standard convolution product on $(0,t)$ as

$$(a * b)(t) := \int_0^t a(t-s)\,b(s)\,\mathrm{d}s$$

and we make use of Young's convolution inequality

$$\|a * b\|_{L^r(0,t)} \le \|a\|_{L^p(0,t)}\,\|b\|_{L^q(0,t)} \quad \forall t \in [0,T]. \tag{2.6}$$

2.4. **Setting of the problem**

Our setting is specified as follows:

(A1): ℓ, $\kappa : (0,T) \to \mathbb{R}$, ℓ, $\kappa \in L^1(0,T)$ are nonnegative and nonincreasing such that $\ell * \kappa \equiv 1$.

(A2): $\alpha = \partial\widehat{\alpha} : \mathbb{R} \to 2^{\mathbb{R}}$ is maximal monotone with a strictly convex potential $\widehat{\alpha}$.

(A3): $\beta : \mathbb{R} \to 2^{\mathbb{R}}$ is maximal monotone.

(A4): $g:\Omega\times(0,T)\times\mathbb{R}\to\mathbb{R}$ is a Carathéodory function, and moreover, there exist $q>2$ and $\Lambda_g>0$ such that

$$|g(x,t,u_1)-g(x,t,u_2)|\le\Lambda_g|u_1-u_2|\quad\forall u_1,\,u_2\in\mathbb{R},\ \text{a.e. }(x,t)\in Q,$$
$$|g(x,t,u)|\le\Lambda_g(1+u^{2/q})\quad\forall u\in\mathbb{R},\ \text{a.e. }(x,t)\in Q.$$

(A5): $v_0\in\alpha_H(u_0)\cap L^{2q-2}(\Omega)$ for some $u_0\in D(\alpha_H)\cap D(\beta_H)$ with $\beta_H^\circ(u_0)\in L^{2q-2}(\Omega)$, where $q>2$ is given in (A4).

Assumption (A1) covers the case of the Riemann–Liouville fractional derivative of order $\theta\in(0,1)$, corresponding to the choices

$$\kappa(t)=\frac{t^{-\theta}}{\Gamma(1-\theta)}\quad\text{and}\quad\ell(t)=\frac{t^{\theta-1}}{\Gamma(\theta)}\quad\forall t\in(0,T).$$

Then κ is also convex.

Note that no growth conditions are imposed on α and β in (A2)–(A3). Correspondingly, both maximal monotone operators α_H and β_H are possibly unbounded. Moreover, β_H can be degenerate.

Assumption (A5) entails that $D(\alpha_H)\cap D(\beta_H)\neq\emptyset$. Without loss of generality, in the following, we additionally assume that

$$0\in\alpha(0)\quad\text{and}\quad0\in\beta(0).$$

Indeed, this can be achieved by considering the shifted graphs $\tilde\alpha(r)=\alpha(r+r^0)-\alpha^0$ and $\tilde\beta(r)=\beta(r+r_0)-\beta^0$ for some $r^0\in D(\alpha)\cap D(\beta)$, $\alpha^0\in\alpha(r^0)$, and $\beta^0\in\beta(r^0)$, as well as the shifted Lipschitz function $\tilde g(x,t,r)=g(x,t,r+r_0)-\beta^0$.

Given assumption (A4) we can define a mapping $G:L^2(0,T;H)\to L^2(0,T;H)$ setting $G(u)(x,t):=g(x,t,u(x,t))$ for all $u\in L^2(0,T;H)$ and a.e. $(x,t)\in Q$. Indeed, for all such u the function $g(x,t,u(x,t))$ is measurable in Q and

$$|g(x,t,u(x,t))|\le|g(x,t,0)|+\Lambda_g|u(x,t)|\le\Lambda_g(1+|u(x,t)|)\quad\text{for a.e. }(x,t)\in Q$$

where we also used the sublinearity bound in (A4) to control $|g(x,t,0)|\le\Lambda_g$. This ensures that $G(u)\in L^2(0,T;H)$, as well as

$$\|G(u)(t)\|\le\Lambda_g(|\Omega|^{1/2}+\|u(t)\|)\quad\forall u\in L^2(0,T;H)\quad\text{for a.e. }t\in(0,T).\tag{2.7}$$

Moreover, for all u_1, $u_2\in L^2(0,T,H)$ we have that

$$\|G(u_1)(t)-G(u_2)(t)\|\le\Lambda_g\|u_1(t)-u_2(t)\|\quad\text{for a.e. }t\in(0,T)\tag{2.8}$$

and the sublinearity assumption in (A4) ensures, with an application of Young's inequality, that

$$(G(u),u)\le\frac{1}{2c_V^2}\|u\|^2+C_G\quad\forall u\in L^2(0,T;H),\ \text{a.e. in }(0,T)\tag{2.9}$$

for $C_G>0$ given by

$$C_G:=c_V^2\Lambda_g^2|\Omega|+\frac{q-2}{2q}\Lambda_g^{\frac{2q}{q-2}}\left(\frac{q}{2(2+q)c_V^2}\right)^{-\frac{q+2}{q-2}}|\Omega|.$$

As $(x,t)\mapsto g(x,t,0)\in L^\infty(Q)$, one can find a measurable set $J\subset(0,T)$ of full measure, i.e., $|(0,T)\setminus J|=0$, such that $x\mapsto g(x,t,0)\in L^\infty(\Omega)$ for all $t\in J$. We may now define $\mathfrak{g}:(0,T)\times H\to H$ by $\mathfrak{g}(t,v)(x)=g(x,t,v(x))$ for $t\in J$ and $\mathfrak{g}(t,v)=0$ for $t\in(0,T)\setminus J$ for

$v \in H$. Moreover, since $(x,t) \mapsto g(x,t,v(x)) \in L^2(Q)$, one observes that $x \mapsto g(x,t,v(x))$ is measurable in Ω, for all $t \in (0,T)$. Moreover,

$$|\mathfrak{g}(t,v)(x)| \le |g(x,t,0)| + \Lambda_g |v(x)| \quad \text{for a.e. } x \in \Omega, \ \forall t \in (0,T)$$

which implies that $\mathfrak{g}(t,v) \in H$ for all $t \in (0,T)$ and that the map $t \in (0,T) \mapsto \mathfrak{g}(t,v) \in H$ is strongly measurable. Eventually,

$$\|\mathfrak{g}(t,v_1) - \mathfrak{g}(t,v_2)\| \le \Lambda_g \|v_1 - v_2\| \quad \forall v_1,\, v_2 \in H, \ \forall t \in (0,T). \tag{2.10}$$

Let X be a Banach space and let $1 \le p \le \infty$. Define a map $\mathcal{B} : D(\mathcal{B}) \subset L^p(0,T;X) \to L^p(0,T;X)$ by $u \mapsto \mathcal{B}(u) := \partial_t(\kappa * u)$ for $u \in D(\mathcal{B}) := \{u \in L^p(0,T;X) \colon \kappa * u \in W^{1,p}(0,T;X),\ (\kappa * u)(0) = 0\}$. Then $\mathcal{B}$ is linear and m-accretive in $L^p(0,T;X)$ (see, e.g., [43, 6]). In the following, we use the nonlocal-in-time chain-rule inequality from [7, Proposition 4.2. ii]: if $\gamma : \mathbb{R} \to 2^{\mathbb{R}}$ is a maximal monotone map, for all $u_0 \in D(\psi_\gamma)$, $u \in L^2(0,T,H)$ with $\psi_\gamma(u) \in L^1(0,T)$ and $\kappa * (u - u_0) \in W^{1,2}(0,T;H)$ satisfying $(\kappa * (u - u_0))(0) = 0$, and $z \in L^2(0,T,H)$ with $z \in \gamma_H(u)$ a.e. in $(0,T)$, one has that

$$\Big[\ell * \big(\partial_t(\kappa * (u - u_0)), z\big)\Big](t) \ge \psi_\gamma(u(t)) - \psi_\gamma(u_0) \quad \text{for a.e. } t \in (0,T). \tag{2.11}$$

If γ is (single-valued and) Lipschitz continuous, a nonlocal chain-rule inequality holds also for $u \in L^2(0,T;V)$ satisfying $\kappa * (u - u_0) \in W^{1,2}(0,T;V^*)$ and $(\kappa * (u - u_0))(0) = 0$ only. Then we see $\gamma(u) \in L^2(0,T;V)$. We argue by approximation: Let $u_\delta \in W^{1,2}(0,T;V)$ be such that $u_\delta \to u$ in $L^2(0,T;V)$ and $\kappa * (u_\delta - u_0) \to \kappa * (u - u_0)$ in $W^{1,2}(0,T;V^*)$ as $\delta \to 0$. Moreover, we find that $(\kappa * (u_\delta - u_0))(0) = 0$. By passing to the $\liminf$ as $\delta \to 0$ in inequality (2.11) written for u_δ and $z_\delta = \gamma_H(u_\delta)$ a.e. in $(0,T)$ we get

$$\Big[\ell * \big\langle \partial_t(\kappa * (u - u_0)), z\big\rangle\Big](t) \ge \psi_\gamma(u(t)) - \psi_\gamma(u_0) \quad \text{for a.e. } t \in (0,T). \tag{2.12}$$

We are now in the position of stating our main result.

Theorem 2.1 (Well-posedness)**.** *Under assumptions* (A0)–(A5) *there exists a triplet* $(u,v,w) \in L^2(0,T;V \times H \times H)$ *such that* $\kappa * (v - v_0) \in W^{1,2}(0,T;V^*)$, $(\kappa * (v - v_0))(0) = 0$ *and*

$$\partial_t(\kappa * (v - v_0)) + Au + w = G(u) \quad \text{in } V^*, \text{ a.e. in } (0,T), \tag{2.13}$$

$$v \in \alpha(u) \quad \text{a.e. in } Q, \tag{2.14}$$

$$w \in \beta(u) \quad \text{a.e. in } Q. \tag{2.15}$$

In addition, if κ is convex, α is strongly monotone and β is (*single-valued and*) *Lipschitz continuous, such u is unique.*

The remainder of the paper is devoted to giving a proof of Theorem 2.1. At first, we check uniqueness in Section 3. Existence is then proved in Section 4 via an approximation and Galerkin discretization approach.

3. PROOF OF THEOREM 2.1: UNIQUENESS

For the purpose of proving uniqueness, κ is assumed to be convex and α and β are assumed to be strongly monotone and (single-valued and) Lipschitz continuous, respectively.

That is, let $C_\alpha > 0$ and $\Lambda_\beta > 0$ be such that

$$C_\alpha |r_1 - r_2|^2 \le (y_1 - y_2)(r_1 - r_2) \quad \forall y_i \in \alpha(r_i),\ i = 1, 2,$$
$$|z_1 - z_2| \le \Lambda_\beta |r_1 - r_2| \quad \forall z_i = \beta(r_i),\ i = 1, 2.$$

As $\ell \in L^1(0, T)$, we can find $m \in \mathbb{N}$ large enough so that the number $\tau := T/m$ satisfies

$$\|\ell\|_{L^1(0,\tau)} \le \frac{C_\alpha}{\sqrt{2}(\Lambda_\beta + \Lambda_g)}. \tag{3.1}$$

Let (u_i, v_i, w_i) for $i = 1, 2$ solve (2.13)–(2.15) in the sense of Theorem 2.1 and define $\tilde{u} := u_1 - u_2$, $\tilde{v} := v_1 - v_2$, $\tilde{w} := w_1 - w_2$, and $\tilde{G} := G(u_1) - G(u_2)$. Writing (2.13) for (u_i, v_i, w_i), for $i = 1, 2$, in place of (u, v, w) and taking the difference of the two resulting equations we get

$$\partial_t(\kappa * \tilde{v}) + A\tilde{u} + \tilde{w} = \tilde{G} \quad \text{in } V^*, \text{ a.e. in } (0, T).$$

Convolving with ℓ and using that $\ell * \kappa \equiv 1$, see (A1), one obtains

$$\tilde{v} + \ell * A\tilde{u} + \ell * \tilde{w} = \ell * \tilde{G} \quad \text{in } V^*, \text{ a.e. in } (0, T). \tag{3.2}$$

Test (3.2) by $\tilde{u}$ and integrate on $(0, t)$ for $t \in [0, T]$. Using the strong monotonicity of α we get

$$\begin{aligned} C_\alpha \int_0^t \|\tilde{u}\|^2 \, \mathrm{d}s &\le \int_0^t (\tilde{v}, \tilde{u}) \, \mathrm{d}s = \int_0^t \langle \tilde{v}, \tilde{u} \rangle \, \mathrm{d}s \\ &\overset{(3.2)}{=} - \int_0^t \langle \ell * A\tilde{u}, \tilde{u} \rangle \, \mathrm{d}s - \int_0^t (\ell * \tilde{w}, \tilde{u}) \, \mathrm{d}s + \int_0^t (\ell * \tilde{G}, \tilde{u}) \, \mathrm{d}s. \end{aligned} \tag{3.3}$$

Since, in addition to (A0), we assumed that κ is convex, then, as ℓ is of positive type (see [22, p. 500, Proposition 3.1]), the first term in the above right-hand side is nonnegative, namely,

$$\int_0^t \langle \ell * A\tilde{u}, \tilde{u} \rangle \, \mathrm{d}s = \int_0^t \int_\Omega (\ell * \nabla\tilde{u}) \cdot \nabla\tilde{u} \, \mathrm{d}x \, \mathrm{d}s \ge 0.$$

The second and third terms in the right-hand side of (3.3) can be controlled as follows

$$\begin{aligned} &- \int_0^t (\ell * \tilde{w}, \tilde{u}) \, \mathrm{d}s + \int_0^t (\ell * \tilde{G}, \tilde{u}) \, \mathrm{d}s \le \int_0^t \|\ell * \tilde{w}\| \|\tilde{u}\| \, \mathrm{d}s + \int_0^t \|\ell * \tilde{G}\| \|\tilde{u}\| \, \mathrm{d}s \\ &\le \int_0^t (\ell * \|\tilde{w}\|) \|\tilde{u}\| \, \mathrm{d}s + \int_0^t (\ell * \|\tilde{G}\|) \|\tilde{u}\| \, \mathrm{d}s \le (\Lambda_\beta + \Lambda_g) \int_0^t (\ell * \|\tilde{u}\|) \|\tilde{u}\| \, \mathrm{d}s \\ &\le \frac{C_\alpha}{2} \int_0^t \|\tilde{u}\|^2 \, \mathrm{d}s + \frac{(\Lambda_\beta + \Lambda_g)^2}{2C_\alpha} \int_0^t (\ell * \|\tilde{u}\|)^2 \, \mathrm{d}s, \end{aligned}$$

where we have used the Lipschitz continuity of β and G. Therefore (3.3) reads,

$$\frac{C_\alpha}{2} \int_0^t \|\tilde{u}\|^2 \, \mathrm{d}s \le \frac{(\Lambda_\beta + \Lambda_g)^2}{2C_\alpha} \int_0^t (\ell * \|\tilde{u}\|)^2 \, \mathrm{d}s. \tag{3.4}$$

Now, if $t = \tau$, we can use the Young convolution inequality (2.6) and inequality (3.1) in order to obtain

$$\frac{C_\alpha}{2} \int_0^t \|\tilde{u}\|^2 \, \mathrm{d}s \le \frac{(\Lambda_\beta + \Lambda_g)^2}{2C_\alpha} \|\ell\|^2_{L^1(0,\tau)} \int_0^t \|\tilde{u}\|^2 \mathrm{d}s \overset{(3.1)}{\le} \frac{C_\alpha}{4} \int_0^t \|\tilde{u}\|^2 \mathrm{d}s.$$

We deduce that $\tilde{u} = 0$ a.e. in $\Omega \times (0, \tau)$.

We now prove that, if $\tilde{u} = 0$ a.e. in $\Omega \times (0, j\tau)$ for some $j = 1, \ldots, m-1$, one also has that $\tilde{u} = 0$ a.e. in $\Omega \times (0, (j+1)\tau)$. This will imply that $\tilde{u} = 0$ a.e. in Q. From (3.4) with $t \in [j\tau, (j+1)\tau]$ and the fact that $\tilde{u} = 0$ a.e. in $\Omega \times (0, j\tau)$ we get that

$$\frac{C_\alpha}{2} \int_{j\tau}^{t} \|\tilde{u}\|^2 \, \mathrm{d}s \le \frac{(\Lambda_\beta + \Lambda_g)^2}{2C_\alpha} \int_{j\tau}^{t} (\ell * \|\tilde{u}\|)^2 \, \mathrm{d}s. \tag{3.5}$$

Let $\tilde{\ell}_t : (0, t) \to \mathbb{R}$ be defined as

$$\tilde{\ell}_t(\sigma) := \begin{cases} \ell(\sigma) & \text{for } \ 0 < \sigma \le t - j\tau, \\ 0 & \text{for } \ t - j\tau < \sigma < t. \end{cases}$$

Note that, for $j\tau < r < s \le t$ we have $0 < s - r < t - j\tau \le \tau$. This in particular entails that

$$\ell(s-r)\|\tilde{u}(r)\| = \tilde{\ell}_t(s-r)\|\tilde{u}(r)\| \quad \text{for } j\tau < r < s \le t. \tag{3.6}$$

One can use this fact in order to control the right-hand side of (3.5) as follows:

$$\begin{aligned}
\frac{(\Lambda_\beta + \Lambda_g)^2}{2C_\alpha} \int_{j\tau}^{t} (\ell * \|u\|)^2 \, \mathrm{d}s &= \frac{(\Lambda_\beta + \Lambda_g)^2}{2C_\alpha} \int_{j\tau}^{t} \left(\int_0^s \ell(s-r)\|\tilde{u}(r)\| \, \mathrm{d}r \right)^2 \mathrm{d}s \\
&= \frac{(\Lambda_\beta + \Lambda_g)^2}{2C_\alpha} \int_{j\tau}^{t} \left(\int_{j\tau}^s \ell(s-r)\|\tilde{u}(r)\| \, \mathrm{d}r \right)^2 \mathrm{d}s \\
&\overset{(3.6)}{=} \frac{(\Lambda_\beta + \Lambda_g)^2}{2C_\alpha} \int_{j\tau}^{t} \left(\int_{j\tau}^s \tilde{\ell}_t(s-r)\|\tilde{u}(r)\| \, \mathrm{d}r \right)^2 \mathrm{d}s \\
&= \frac{(\Lambda_\beta + \Lambda_g)^2}{2C_\alpha} \int_{0}^{t} \left(\int_{0}^s \tilde{\ell}_t(s-r)\|\tilde{u}(r)\| \, \mathrm{d}r \right)^2 \mathrm{d}s = \frac{(\Lambda_\beta + \Lambda_g)^2}{2C_\alpha} \left\| \tilde{\ell}_t * \|\tilde{u}\| \right\|_{L^2(0,t)}^2 \\
&\overset{(2.6)}{\le} \frac{(\Lambda_\beta + \Lambda_g)^2}{2C_\alpha} \|\tilde{\ell}_t\|_{L^1(0,t)}^2 \int_0^t \|\tilde{u}\|^2 \, \mathrm{d}s = \frac{(\Lambda_\beta + \Lambda_g)^2}{2C_\alpha} \|\ell\|_{L^1(0,t-j\tau)}^2 \int_{j\tau}^t \|\tilde{u}\|^2 \, \mathrm{d}s \\
&\le \frac{(\Lambda_\beta + \Lambda_g)^2}{2C_\alpha} \|\ell\|_{L^1(0,\tau)}^2 \int_{j\tau}^t \|\tilde{u}\|^2 \, \mathrm{d}s \overset{(3.1)}{\le} \frac{C_\alpha}{4} \int_{j\tau}^t \|\tilde{u}\|^2 \, \mathrm{d}s
\end{aligned}$$

for $t \in [j\tau, (j+1)\tau]$. In combination with (3.5), this proves that $\tilde{u} = 0$ a.e. in $\Omega \times (j\tau, t)$ for all $t \in [j\tau, (j+1)\tau]$. We have hence checked that $\tilde{u} = 0$ a.e. in $\Omega \times (0, (j+1)\tau)$ and the uniqueness assertion follows.

4. PROOF OF THEOREM 2.1: EXISTENCE

As mentioned in Section 1, the proof of the existence of solutions to (2.13)–(2.15) follows from regularization and Galerkin approximation procedures.

For the reader's convenience, we split the argument into subsections. At first, we regularize the maps α and β by passing to their Yosida approximations and by adding a multiple of the identity to α (Section 4.1). The ensuing regularized problem is then tackled via a Galerkin approximation (Section 4.2). After establishing some a-priori estimates (Section 4.3), we pass to the limit in the Galerkin approximation (Section 4.4) and prove the well-posedness of the regularized problem, namely, Proposition 4.2 below. We then obtain additional a-priori estimates for the solution to the regularized problem (Section

4.5) and pass to the limit in the regularizations (Sections 4.6–4.7), eventually proving Theorem 2.1.

4.1. **Regularization**

Given a parameter $\varepsilon \in (0,1)$, we denote by

$$\alpha_\varepsilon := \frac{\mathrm{id}_{\mathbb{R}} - (\mathrm{id}_{\mathbb{R}} + \varepsilon\alpha)^{-1}}{\varepsilon}, \quad \tilde{\beta}_\varepsilon := \frac{\mathrm{id}_{\mathbb{R}} - (\mathrm{id}_{\mathbb{R}} + \varepsilon\beta)^{-1}}{\varepsilon},$$

the Yosida approximations of the maximal monotone graphs α and β at level ε, respectively. Here, $\mathrm{id}_{\mathbb{R}}$ stands for the identity in $\mathbb{R}$. Moreover, we define the truncated one by

$$\beta_\varepsilon(r) = \min\{\max\{\tilde{\beta}_\varepsilon(r), -\varepsilon^{-1}\}, \varepsilon^{-1}\} \quad \forall r \in \mathbb{R}.$$

Recall that $\alpha_{\varepsilon H}$ and $\beta_{\varepsilon H}$ are single-valued, ε^{-1}-Lipschitz continuous in H, and that $\alpha_{\varepsilon H}(0) = \beta_{\varepsilon H}(0) = 0$ (cf. [16, Proposition 2.6 (i), p. 28]).

We additionally define

$$\alpha_\nu := \nu\,\mathrm{id}_{\mathbb{R}} + \alpha, \quad \alpha_{\nu\varepsilon} := \nu\,\mathrm{id}_{\mathbb{R}} + \alpha_\varepsilon, \quad \eta_{\nu\varepsilon} := \alpha_{\nu\varepsilon}^{-1}$$

for all $\nu \in (0,1)$. Note that $\alpha_{\nu\varepsilon H}$ is monotone, $(\nu+\varepsilon^{-1})$-Lipschitz continuous, and coercive from H to H. Hence, it is maximal monotone and onto. The mapping $\eta_{\nu\varepsilon H}$ is ν^{-1}-Lipschitz continuous in H. Moreover, we have $\alpha_{\nu\varepsilon H}(0) = \eta_{\nu\varepsilon H}(0) = 0$.

Recall that $\alpha_\varepsilon(r) \to \alpha^\circ(r)$ and $|\alpha_\varepsilon(r)| \le |\alpha^\circ(r)|$ for all $r \in D(\alpha)$ and $\beta_\varepsilon(r) \to \beta^\circ(r)$ and $|\beta_\varepsilon(r)| \le |\beta^\circ(r)|$ for all $r \in D(\beta)$, where we recall that $\alpha^\circ(r)$ and $\beta^\circ(r)$ denote the elements of smallest absolute value in the closed intervals $\alpha(r)$ and $\beta(r)$, respectively (cf. [16, Proposition 2.6(iii)]). Moreover, we see that $\widehat{\alpha_\varepsilon} \nearrow \widehat{\alpha}$, $\widehat{\alpha_{\nu\varepsilon}} \nearrow \widehat{\alpha}_\nu$, and $\widehat{\beta}_\varepsilon \nearrow \widehat{\beta}$ pointwisely. Owing to [11, Theorem 3.20, p. 298] we have that $\Psi_{\alpha_\varepsilon} \to \Psi_\alpha$, $\Psi_{\alpha_{\nu\varepsilon}} \to \Psi_{\alpha_\nu}$, and $\Psi_{\beta_\varepsilon} \to \Psi_\beta$ in the Mosco sense in $L^2(0,T;H)$. Hence, [11, Theorem 3.66, p. 373] implies that

$$\partial\Psi_{\alpha_{\nu\varepsilon}} \to \partial\Psi_{\alpha_\nu} \quad \text{and} \quad \partial\Psi_{\beta_\varepsilon} \to \partial\Psi_\beta \quad \text{in the graph sense in } L^2(0,T,H) \tag{4.1}$$

as $\varepsilon \to 0$. Moreover, one can easily check that

$$\partial\Psi_{\alpha_\nu} \to \partial\Psi_\alpha \quad \text{in the graph sense in } L^2(0,T,H) \tag{4.2}$$

as $\nu \to 0$.

We next introduce some approximation $(u_{0\varepsilon}, v_{0\varepsilon})$ of initial data (u_0, v_0) satisfying (A5), for which we can check not only that $\alpha_{\varepsilon H}(u_{0\varepsilon}) \to v_0$ in H but also that $\beta_{\varepsilon H}(u_{0\varepsilon})$ is bounded in H as $\varepsilon \to 0$. This approximation will play a crucial role when α is not supposed to be single-valued (in other words, when α is single-valued, i.e., $\alpha_H(u_0) = v_0$, the following construction is not necessary). Let $(s_j)_{j\in J}$ indicate all values in $D(\alpha)$ where the value of α is not a singleton. As α is monotone in $\mathbb{R}$, the index set J is at most countable. Starting from $u_0 \in D(\alpha_H)$ as in (A5), for all $j \in J$ we define the measurable sets $\Omega_j := u_0^{-1}(s_j)$ and we set $\hat{\Omega} := \Omega \setminus \cup_{j\in J}\Omega_j$. Note that these sets form a disjoint partition of Ω, up to null sets. We set

$$u_{0\varepsilon} := \begin{cases} u_0 & \text{on } \hat{\Omega}, \\ s_j + \varepsilon v_0 & \text{on } \Omega_j,\ \forall j \in J, \end{cases} \tag{4.3}$$

$$v_{0\varepsilon} := \alpha_{\varepsilon H}(u_{0\varepsilon}), \tag{4.4}$$

where v_0 is as in (A5). We can now state several properties of the approximation of the initial data.

Lemma 4.1 (Properties of the approximation of the initial data)**.** *Under assumptions* (A0), (A2)–(A5), *and positions* (4.3)–(4.4) *we have the following*:

(i) *It holds that*

$$v_{0\varepsilon} \to v_0 \quad \textit{in } H \quad \textit{and} \quad \|v_{0\varepsilon}\| \le \|v_0\| \tag{4.5}$$

as $\varepsilon \to 0$.

(ii) *It holds that*

$$|u_{0\varepsilon} - u_0| \le \varepsilon |v_0| \quad \textit{a.e. in } \Omega, \tag{4.6}$$

which in particular yields

$$\|u_{0\varepsilon} - u_0\| \le \varepsilon \|v_0\|, \tag{4.7}$$

$$|\beta_{\varepsilon H}(u_{0\varepsilon})| \le |\beta_H^\circ(u_0)| + |v_0| \quad \textit{a.e. in } \Omega, \tag{4.8}$$

for any $\varepsilon \in (0,1)$.

(iii) *For all* $\varepsilon \in (0,1)$ *and* $\nu \in (0,1)$, *set*

$$v_{0\nu\varepsilon} := \alpha_{\nu\varepsilon H}(u_{0\varepsilon}) = \nu u_{0\varepsilon} + v_{0\varepsilon},$$

where $u_{0\varepsilon}$ *and* $v_{0\varepsilon}$ *are defined in* (4.3)–(4.4). *Then, there exists a constant* $C > 0$ *depending on* $\|u_0\|$ *and* $\|v_0\|$ *but independent of* ε *and* ν *such that*

$$\int_\Omega \widehat{\eta_{\nu\varepsilon}}(v_{0\nu\varepsilon})\, \mathrm{d}x \le \nu \|u_{0\varepsilon}\|^2 + |(u_{0\varepsilon}, v_{0\varepsilon})| \le C,$$

where $\widehat{\eta_{\nu\varepsilon}} = \widehat{\alpha_{\nu\varepsilon}}^*$. *This in particular implies that* $v_{0\nu\varepsilon} \in D(\psi_{\eta_{\nu\varepsilon}})$, *for any* $\varepsilon, \nu \in (0,1)$.

(iv) *Let* $q > 2$ *be given as in* (A4). *For all* $\varepsilon \in (0,1)$, *setting* $\beta_{\varepsilon q}(r) := |\beta_\varepsilon(r)|^{q-2}\beta_\varepsilon(r)$ *for all* $r \in \mathbb{R}$, *it holds that*

$$\int_\Omega \widehat{\beta_{\varepsilon q} \circ \alpha_{\nu\varepsilon}^{-1}}(v_{0\nu\varepsilon})\, \mathrm{d}x \le \tilde{C}(\nu \|u_{0\varepsilon}\| + \|v_{0\varepsilon}\|),$$

where $\tilde{C} > 0$ *depends on* q, $\|\beta_H^\circ(u_0)\|_{L^{2q-2}(\Omega)}$, *and* $\|v_0\|_{L^{2q-2}(\Omega)}$ *but is independent of* ε *and* ν. *This also yields that* $v_{0\nu\varepsilon} \in D(\psi_{\beta_{\varepsilon q} \circ \alpha_{\nu\varepsilon}^{-1}})$.

Proof. We prove (i). As $v_{0\varepsilon} = \alpha_\varepsilon(u_0)$ a.e. on $\hat{\Omega}$, since there $\alpha(u_0)$ is singleton, i.e., $\alpha(u_0) = \{v_0\} = \{\alpha^\circ(u_0)\}$, one has that $v_{0\varepsilon} \to \alpha^\circ(u_0) = v_0$ and $|v_{0\varepsilon}| \le |\alpha^\circ(u_0)| = |v_0|$ a.e. on $\hat{\Omega}$. On the other hand, for all $j \in J$ one has

$$v_{0\varepsilon} = \alpha_\varepsilon(s_j + \varepsilon v_0) = \frac{1}{\varepsilon}\big((s_j + \varepsilon v_0) - (\mathrm{id} + \varepsilon\alpha)^{-1}(s_j + \varepsilon v_0)\big) = v_0 \quad \text{a.e. on } \Omega_j,$$

where we used that $s_j + \varepsilon\alpha(s_j) \ni s_j + \varepsilon v_0$, and hence, $(\mathrm{id} + \varepsilon\alpha)^{-1}(s_j + \varepsilon v_0) = s_j$. We have hence proved that $v_{0\varepsilon} \to v_0$ and $|v_{0\varepsilon}| \le |v_0|$ a.e. in Ω, whence (4.5) follows by the dominated convergence.

As for (ii), as we have that

$$u_{0\varepsilon} - u_0 := \begin{cases} 0 & \text{on } \hat{\Omega}, \\ \varepsilon v_0 & \text{on } \Omega_j, \ \forall j \in J, \end{cases}$$

we readily obtain (4.6) and (4.7). This in particular implies that

$$\begin{aligned} |\beta_{\varepsilon H}(u_{0\varepsilon})| &\le |\beta_{\varepsilon H}(u_0)| + |\beta_{\varepsilon H}(u_{0\varepsilon}) - \beta_{\varepsilon H}(u_0)| \\ &\le |\beta_H^\circ(u_0)| + \frac{1}{\varepsilon}|u_{0\varepsilon} - u_0| \overset{(4.6)}{\le} |\beta_H^\circ(u_0)| + |v_0| \quad \text{a.e. in } \Omega. \end{aligned}$$

Thus (4.8) follows.

We next prove (iii). Since $v_{0\nu\varepsilon} = \alpha_{\nu\varepsilon H}(u_{0\varepsilon})$ we have

$$\begin{aligned} \int_\Omega \widehat{\eta_{\nu\varepsilon}}(v_{0\nu\varepsilon})\,\mathrm{d}x = \int_\Omega \widehat{\alpha_{\nu\varepsilon}}^*(v_{0\nu\varepsilon})\,\mathrm{d}x &= -\int_\Omega \widehat{\alpha_{\nu\varepsilon}}(u_{0\varepsilon})\,\mathrm{d}x + (u_{0\varepsilon}, v_{0\nu\varepsilon}) \\ &\le |(u_{0\varepsilon}, v_{0\nu\varepsilon})| \le \nu\|u_{0\varepsilon}\|^2 + |(u_{0\varepsilon}, v_{0\varepsilon})| \\ &\le 2\nu\|u_0\|^2 + 2\nu\|u_{0\varepsilon} - u_0\|^2 + (\|u_0\| + \|u_{0\varepsilon} - u_0\|)\|v_0\| \\ &\le 2\nu\|u_0\|^2 + \|u_0\|\,\|v_0\| + (2\nu\varepsilon + 1)\varepsilon\|v_0\|^2. \end{aligned}$$

where we also used (4.5) and (4.7).

We finally prove (iv). Recalling that $u_{0\varepsilon} = \alpha_{\nu\varepsilon}^{-1}(v_{0\nu\varepsilon})$ a.e. in Ω, we see that

$$\begin{aligned} &\int_\Omega \widehat{\beta_{\varepsilon q} \circ \alpha_{\nu\varepsilon}^{-1}}(v_{0\nu\varepsilon})\,\mathrm{d}x \le ((\beta_{\varepsilon qH} \circ \alpha_{\nu\varepsilon H}^{-1})(v_{0\nu\varepsilon}), v_{0\nu\varepsilon}) = (\beta_{\varepsilon qH}(u_{0\varepsilon}), v_{0\nu\varepsilon}) \\ &\quad \le \|\beta_{\varepsilon qH}(u_{0\varepsilon})\|(\nu\|u_{0\varepsilon}\| + \|v_{0\varepsilon}\|). \end{aligned}$$

The assertion follows by using (4.8) in order to check that

$$\begin{aligned} &\|\beta_{\varepsilon qH}(u_{0\varepsilon})\| = \left(\int_\Omega |\beta_{\varepsilon H}(u_{0\varepsilon})|^{2q-2}\,\mathrm{d}x\right)^{1/2} \overset{(4.8)}{\le} \left(\int_\Omega (|\beta_H^\circ(u_0)| + |v_0|)^{2q-2}\,\mathrm{d}x\right)^{1/2} \\ &\quad \le 2^{q-3/2}(\|\beta_H^\circ(u_0)\|_{L^{2q-2}(\Omega)}^{q-1} + \|v_0\|_{L^{2q-2}(\Omega)}^{q-1}). \end{aligned}$$

This completes the proof. □

4.2. **Galerkin approximation**

Let $(e_i)_i \subset V$ denote normalized eigenfunctions of the Laplacian with the homogeneous Dirichlet boundary condition, i.e., solutions $e_i \in V$ to $Ae_i = \lambda_i e_i$ with $\|e_i\|_V = 1$ for every $i \in \mathbb{N}$ and increasing eigenvalues $\lambda_i > 0$. Such a set is orthogonal in H and orthonormal in V. By setting

$$H_n := \operatorname{span}\{e_1, \dots, e_n\} \simeq \mathbb{R}^n$$

for all $n \in \mathbb{N}$ one has that the closures of $\cup_n H_n$ in H and V coincide with H and V, respectively. Let $\pi_n : H \to H_n$ denote the orthogonal projection on H_n and recall that $\pi_n(u) = \sum_{i=1}^n u_i e_i$ for all $u \in H$ where $u_i := (u, e_i)$ for $i = 1, \dots, n$. We also introduce the mapping $r_n : \mathbb{R}^n \to H_n$ given by

$$r_n(\boldsymbol{x}) := \sum_{i=1}^n x_i e_i \quad \forall \boldsymbol{x} = (x_1, \dots, x_n) \in \mathbb{R}^n.$$

We claim that

$$\begin{aligned} (\pi_n \circ \alpha_{\nu\varepsilon H})|_{H_n} &= (\pi_n \circ (\nu\,\mathrm{id}_H + \alpha_{\varepsilon H}))|_{H_n} \\ &= \nu\,\mathrm{id}_{H_n} + (\pi_n \circ \alpha_{\varepsilon H})|_{H_n} : H_n \to H_n \quad \text{is onto} \qquad (4.9) \end{aligned}$$

where id_H and id_{H_n} denote the identities in H and H_n, respectively. Indeed, (4.9) follows from the fact that $\nu\,\mathrm{id}_{H_n} + (\pi_n \circ \alpha_{\varepsilon H})|_{H_n}$ is well defined on H_n, strongly monotone and Lipschitz continuous on H_n, hence maximal monotone and onto. In particular, its inverse

$$\eta_{\nu\varepsilon nH} := ((\pi_n \circ \alpha_{\nu\varepsilon H})|_{H_n})^{-1} = (\nu\,\mathrm{id}_{H_n} + (\pi_n \circ \alpha_{\varepsilon H})|_{H_n})^{-1} : H_n \to H_n$$

is well defined, monotone, and Lipschitz continuous (and therefore, maximal monotone).

We now proceed to the Galerkin approximation. To start with, for all ε, $\nu \in (0,1)$, and $n \in \mathbb{N}$ we set

$$z_{0\nu\varepsilon n} := \pi_n(v_{0\nu\varepsilon}) = (\pi_n \circ \alpha_{\nu\varepsilon H})(u_{0\varepsilon}).$$

Here $v_{0\nu\varepsilon}$ is given by Lemma 4.1. We assume $\boldsymbol{z}_{0\nu\varepsilon n} \in \mathbb{R}^n$ to be given by $(\boldsymbol{z}_{0\nu\varepsilon n})_i := (z_{0\nu\varepsilon n}, e_i)$, for $i = 1, \ldots, n$, and we define the mapping $\boldsymbol{\delta}_{\nu\varepsilon n} : [0,T] \times \mathbb{R}^n \to \mathbb{R}^n$ as

$$\begin{aligned}(\boldsymbol{\delta}_{\nu\varepsilon n}(t,\boldsymbol{x}))_i &:= \lambda_i((\eta_{\nu\varepsilon nH} \circ r_n)(\boldsymbol{x}), e_i) + ((\beta_\varepsilon \circ \eta_{\nu\varepsilon nH} \circ r_n)(\boldsymbol{x}), e_i)\\ &\quad - ((\mathfrak{g}(t,\cdot) \circ \eta_{\nu\varepsilon nH} \circ r_n)(\boldsymbol{x}), e_i) \quad \text{for } i = 1, \ldots, n.\end{aligned}$$

Notice that $\boldsymbol{\delta}_{\nu\varepsilon n}(t,\cdot)$ is Lipschitz continuous in $\mathbb{R}^n$ uniformly for $t \in [0,T]$, as it is a composition of Lipschitz continuous maps (see in particular (2.10)) and $H_n \simeq \mathbb{R}^n$. An application of [6, Theorem 5.1] then ensures that there exists a unique $\boldsymbol{z}_{\nu\varepsilon n} \in L^2(0,T;\mathbb{R}^n)$ such that $\kappa * (\boldsymbol{z}_{\nu\varepsilon n} - \boldsymbol{z}_{0\nu\varepsilon n}) \in W^{1,2}(0,T;\mathbb{R}^n)$, $(\kappa * (\boldsymbol{z}_{\nu\varepsilon n} - \boldsymbol{z}_{0\nu\varepsilon n}))(0) = 0$, and

$$\partial_t(\kappa * (\boldsymbol{z}_{\nu\varepsilon n} - \boldsymbol{z}_{0\nu\varepsilon n}))(t) + \boldsymbol{\delta}_{\nu\varepsilon n}(t, \boldsymbol{z}_{\nu\varepsilon n}(t)) = 0 \quad \text{in } \mathbb{R}^n, \text{ for a.e. } t \in (0,T). \tag{4.10}$$

Indeed, in the notation of [6, Theorem 5.1] it suffices to choose $H := \mathbb{R}^n$, $\varphi := 0$, $f := 0$, and $F := \boldsymbol{\delta}_{\nu\varepsilon n}$.

Set now $z_{\nu\varepsilon n} := r_n(\boldsymbol{z}_{\nu\varepsilon n}) \in L^2(0,T;H_n)$ and $u_{\nu\varepsilon n} := \eta_{\nu\varepsilon nH}(z_{\nu\varepsilon n}) \in L^2(0,T;H_n)$, so that

$$z_{\nu\varepsilon n} = (\pi_n \circ \alpha_{\nu\varepsilon H})(u_{\nu\varepsilon n}).$$

By multiplying the i-th component of (4.10) by e_i and summing from $i = 1$ to $i = n$ we find that

$$\begin{aligned}&\partial_t(\kappa * (z_{\nu\varepsilon n} - z_{0\nu\varepsilon n})) + Au_{\nu\varepsilon n} + (\pi_n \circ \beta_{\varepsilon H})(u_{\nu\varepsilon n})\\ &\quad = (\pi_n \circ G)(u_{\nu\varepsilon n}) \quad \text{in } H_n, \text{ a.e. in } (0,T),\end{aligned} \tag{4.11}$$

where $Au_{\nu\varepsilon n}$ also denotes its realization in H, i.e., $Au_{\nu\varepsilon n} = \sum_{i=1}^n \lambda_i(u_{\nu\varepsilon n}, e_i)e_i \in H_n$. In the following, we shall use the notation

$$v_{\nu\varepsilon n} := \alpha_{\nu\varepsilon H}(u_{\nu\varepsilon n}) \in L^2(0,T;H), \tag{4.12}$$

so that $z_{\nu\varepsilon n} = \pi_n(v_{\nu\varepsilon n}) \in L^2(0,T;H_n)$.

4.3. **A-priori estimates for the Galerkin approximation**

For the sake of notational simplicity, henceforth we use the same symbol C to denote a generic positive constant, possibly depending on data but independent of ε, ν, and n. The actual value of the constant C may vary from line to line. We will denote by $C(\varepsilon)$ a generic positive constant depending on data and on ε, but not on neither ν nor n.

Let us start by testing equation (4.11) by $u_{\nu\varepsilon n}$ and getting

$$\begin{aligned}&(\partial_t(\kappa * (z_{\nu\varepsilon n} - z_{0\nu\varepsilon n})), u_{\nu\varepsilon n}) + (Au_{\nu\varepsilon n}, u_{\nu\varepsilon n})\\ &\quad + ((\pi_n \circ \beta_{\varepsilon H})(u_{\nu\varepsilon n}), u_{\nu\varepsilon n}) = ((\pi_n \circ G)(u_{\nu\varepsilon n}), u_{\nu\varepsilon n}) \quad \text{a.e. in } (0,T).\end{aligned}$$

Since $u_{\nu\varepsilon n}$, $z_{\nu\varepsilon n} \in H_n$ a.e. in $(0,T)$, we drop π_n in the above equation, and replace $z_{\nu\varepsilon n}$ by $v_{\nu\varepsilon n}$ and $z_{0\nu\varepsilon n}$ by $v_{0\nu\varepsilon}$. This gives

$$(\partial_t(\kappa * (v_{\nu\varepsilon n} - v_{0\nu\varepsilon})), u_{\nu\varepsilon n}) + \|u_{\nu\varepsilon n}\|_V^2 + (\beta_{\varepsilon H}(u_{\nu\varepsilon n}), u_{\nu\varepsilon n}) = (G(u_{\nu\varepsilon n}), u_{\nu\varepsilon n}) \tag{4.13}$$

a.e. in $(0,T)$. As $v_{0\nu\varepsilon} \in D(\psi_{\eta_{\nu\varepsilon}})$ (see Lemma 4.1, (iii)), $v_{\nu\varepsilon n}, u_{\nu\varepsilon n} \in L^2(0,T;H)$, $\kappa * (v_{\nu\varepsilon n} - v_{0\nu\varepsilon}) \in W^{1,2}(0,T;H)$, $(\kappa * (v_{\nu\varepsilon n} - v_{0\nu\varepsilon}))(0) = 0$, $\psi_{\eta_{\nu\varepsilon}}(v_{\nu\varepsilon n}) \in L^1(0,T)$ (indeed, one can prove it as in the proof of Lemma 4.1, (iii)), and $u_{\nu\varepsilon n} \in \eta_{\nu\varepsilon H}(v_{\nu\varepsilon n})$ a.e. in $(0,T)$ (see (4.12)), by applying the nonlocal chain-rule inequality (2.11) we get

$$\ell * \big(\partial_t(\kappa * (v_{\nu\varepsilon n} - v_{0\nu\varepsilon})), u_{\nu\varepsilon n}\big) \geq \int_\Omega \widehat{\eta_{\nu\varepsilon}}(v_{\nu\varepsilon n})\,\mathrm{d}x - \int_\Omega \widehat{\eta_{\nu\varepsilon}}(v_{0\nu\varepsilon})\,\mathrm{d}x \quad \text{a.e. in } (0,T).$$

Hence, convolving (4.13) with ℓ we get

$$\begin{aligned}&\int_\Omega \widehat{\eta_{\nu\varepsilon}}(v_{\nu\varepsilon n})\,\mathrm{d}x - \int_\Omega \widehat{\eta_{\nu\varepsilon}}(v_{0\nu\varepsilon})\,\mathrm{d}x + \ell * \|u_{\nu\varepsilon n}\|_V^2 + \ell * (\beta_{\varepsilon H}(u_{\nu\varepsilon n}), u_{\nu\varepsilon n})\\ &\quad\leq \ell * (G(u_{\nu\varepsilon n}), u_{\nu\varepsilon n}) \quad \text{a.e. in } (0,T).\end{aligned}$$

The above right-hand side can be controlled via (2.9) as

$$\ell * (G(u_{\nu\varepsilon n}), u_{\nu\varepsilon n}) \overset{(2.9)}{\leq} \ell * \frac{1}{2c_V^2}\|u_{\nu\varepsilon n}\|^2 + \ell * C_G \leq \ell * \frac{1}{2}\|u_{\nu\varepsilon n}\|_V^2 + C \quad \text{a.e. in } (0,T).$$

Accordingly, we get

$$\begin{aligned}\int_\Omega \widehat{\eta_{\nu\varepsilon}}(v_{\nu\varepsilon n})\,\mathrm{d}x + \ell * \frac{1}{2}\|u_{\nu\varepsilon n}\|_V^2 + \ell * (\beta_{\varepsilon H}(u_{\nu\varepsilon n}), u_{\nu\varepsilon n})&\\ \leq \int_\Omega \widehat{\eta_{\nu\varepsilon}}(v_{0\nu\varepsilon})\,\mathrm{d}x + C&\end{aligned} \tag{4.14}$$

a.e. in $(0,T)$. Recalling (iii) of Lemma 4.1, we also find that

$$\int_\Omega \widehat{\eta_{\nu\varepsilon}}(v_{0\nu\varepsilon})\,\mathrm{d}x \leq C.$$

Since all the terms in the left-hand side of (4.14) are nonnegative, each of them is uniformly bounded. In particular, we have that $\ell * \|u_{\nu\varepsilon n}\|_V^2 \leq C$ a.e. in $(0,T)$. By convolving this with κ we get that

$$\|u_{\nu\varepsilon n}\|_{L^2(0,T;V)} \leq C. \tag{4.15}$$

As $A : V \to V^*$ is bounded we also get

$$\|Au_{\nu\varepsilon n}\|_{L^2(0,T;V^*)} \leq C. \tag{4.16}$$

Bound (4.15) and the ε^{-1}-Lipschitz continuity of α_ε and β_ε entail that

$$\|\alpha_{\varepsilon H}(u_{\nu\varepsilon n})\|_{L^2(0,T;V)} + \|\beta_{\varepsilon H}(u_{\nu\varepsilon n})\|_{L^2(0,T;V)} \leq C(\varepsilon), \tag{4.17}$$

which in particular also gives

$$\|(\pi_n \circ \beta_{\varepsilon H})(u_{\nu\varepsilon n})\|_{L^2(0,T;H)} \leq C(\varepsilon). \tag{4.18}$$

Moreover, the $(\nu + \varepsilon^{-1})$-Lipschitz continuity of $\alpha_{\nu\varepsilon H}$ and $\alpha_{\nu\varepsilon H}(0) = 0$, gives that

$$\|v_{\nu\varepsilon n}\|_{L^2(0,T;V)} + \|z_{\nu\varepsilon n}\|_{L^2(0,T;H)} \leq C(\varepsilon) \tag{4.19}$$

uniformly for $n \in \mathbb{N}$ and $\nu \in (0,1)$ (see (4.12)). Bound (2.7) implies that

$$\begin{aligned} \|(\pi_n \circ G)(u_{\nu\varepsilon n})\|_{L^2(0,T;H)} &\leq \|G(u_{\nu\varepsilon n})\|_{L^2(0,T;H)} \\ &\leq \Lambda_g|\Omega|^{1/2}T^{1/2} + \Lambda_g\|u_{\nu\varepsilon n}\|_{L^2(0,T;H)} \leq C. \end{aligned} \tag{4.20}$$

Finally, a comparison in (4.11) and bounds (4.16), (4.18), and (4.20) entail

$$\|\partial_t(\kappa * (z_{\nu\varepsilon n} - z_{0\nu\varepsilon n}))\|_{L^2(0,T;V^*)} \leq C(\varepsilon). \tag{4.21}$$

4.4. **Passage to the limit as $n \to \infty$**

The a-priori estimates from Section 4.3 allow us to pass to the limit as $n \to \infty$, for $\varepsilon > 0$ and $\nu \in (0,1)$ fixed. We start by controlling the time increments of $z_{\nu\varepsilon n}$. Using the short-hand notation

$$g_{\nu\varepsilon n}(t) := \partial_t(\kappa * (z_{\nu\varepsilon n} - z_{0\nu\varepsilon n}))(t), \quad t \in (0,T),$$

for every $h \in (0,T)$, we can estimate

$$\begin{aligned} &\int_0^{T-h} \|z_{\nu\varepsilon n}(t+h) - z_{\nu\varepsilon n}(t)\|_{V^*}^2 \,\mathrm{d}t \\ &\quad = \int_0^{T-h} \|(\ell * g_{\nu\varepsilon n})(t+h) - (\ell * g_{\nu\varepsilon n})(t)\|_{V^*}^2 \,\mathrm{d}t \\ &\quad = \left\| (\ell(\cdot + h) - \ell) * g_{\nu\varepsilon n} + \int_{\cdot}^{\cdot + h} \ell(\cdot + h - s) g_{\nu\varepsilon n}(s)\,\mathrm{d}s \right\|_{L^2(0,T-h;V^*)}^2 \\ &\quad \leq 2\,\|(\ell(\cdot + h) - \ell) * g_{\nu\varepsilon n}\|_{L^2(0,T-h;V^*)}^2 \\ &\qquad + 2\left\| \int_{\cdot}^{\cdot + h} \ell(\cdot + h - s) g_{\nu\varepsilon n}(s)\,\mathrm{d}s \right\|_{L^2(0,T-h;V^*)}^2 . \end{aligned} \tag{4.22}$$

The first term in the above right-hand side can be bounded via the Young convolution inequality (2.6) as follows:

$$\begin{aligned} 2\|(\ell(\cdot + t) - \ell) * g_{\nu\varepsilon n}\|_{L^2(0,T-h;V^*)}^2 &\overset{(2.6)}{\leq} 2\|\ell(\cdot + h) - \ell\|_{L^1(0,T)}^2 \|g_{\nu\varepsilon n}\|_{L^2(0,T;V^*)}^2 \\ &\leq C(\varepsilon)\|\ell(\cdot + h) - \ell\|_{L^1(0,T)}^2, \end{aligned}$$

where we used (4.21) in the last inequality. By defining

$$\ell_h(\sigma) := \begin{cases} \ell(\sigma) & \text{for } \sigma \in (0,h), \\ 0 & \text{for } \sigma \in [h,T), \end{cases}$$

we can control the second term in the right-hand side of (4.22) as follows:

$$
\begin{aligned}
&2\left\|\int_{\cdot}^{\cdot+h} \ell(\cdot+h-s) g_{\nu\varepsilon n}(s)\,\mathrm{d}s\right\|^2_{L^2(0,T-h;V^*)} \\
&\quad\le 2\int_0^{T-h}\left(\int_t^{t+h} \ell(t+h-s)\|g_{\nu\varepsilon n}(s)\|_{V^*}\,\mathrm{d}s\right)^2 \mathrm{d}t \\
&\quad= 2\int_0^{T-h}\left(\int_0^{t+h} \ell_h(t+h-s)\|g_{\nu\varepsilon n}(s)\|_{V^*}\,\mathrm{d}s\right)^2 \mathrm{d}t \\
&\quad= 2\big\|\ell_h * \|g_{\nu\varepsilon n}\|_{V^*}\big\|^2_{L^2(h,T)} \le 2\|\ell_h\|^2_{L^1(0,T)}\|g_{\nu\varepsilon n}\|^2_{L^2(0,T;V^*)} \overset{(4.21)}{\le} C(\varepsilon)\|\ell\|^2_{L^1(0,h)}.
\end{aligned}
$$

All in all, we have proved that

$$
\int_0^{T-h} \|z_{\nu\varepsilon n}(t+h)-z_{\nu\varepsilon n}(t)\|^2_{V^*}\,\mathrm{d}t \le C(\varepsilon)\left(\|\ell(\cdot+h)-\ell\|^2_{L^1(0,T)} + \|\ell\|^2_{L^1(0,h)}\right).
$$

As $\ell \in L^1(0,T)$ this entails that

$$
\limsup_{h\to 0}\sup_{n\in\mathbb{N}} \int_0^{T-h} \|z_{\nu\varepsilon n}(t+h)-z_{\nu\varepsilon n}(t)\|^2_{V^*}\,\mathrm{d}t = 0. \tag{4.23}
$$

As a consequence of bounds (4.15)–(4.19) and of (4.23), an application of the Aubin–Lions lemma entails that, up to a subsequence, as $n\to\infty$,

$$
\begin{aligned}
&u_{\nu\varepsilon n} \rightharpoonup u_{\nu\varepsilon} \quad \text{in } L^2(0,T;V), && (4.24)\\
&v_{\nu\varepsilon n} \rightharpoonup v_{\nu\varepsilon} \quad \text{in } L^2(0,T;V), && (4.25)\\
&z_{\nu\varepsilon n} \to \bar{v}_{\nu\varepsilon} \quad \text{in } L^2(0,T;V^*), && (4.26)\\
&z_{\nu\varepsilon n} \rightharpoonup \bar{v}_{\nu\varepsilon} \quad \text{in } L^2(0,T;H), && (4.27)\\
&\partial_t(\kappa*(z_{\nu\varepsilon n}-z_{0\nu\varepsilon n})) \rightharpoonup \xi_{\nu\varepsilon} \quad \text{in } \ L^2(0,T;V^*). && (4.28)
\end{aligned}
$$

It is a standard matter to check that $\bar{v}_{\nu\varepsilon} = v_{\nu\varepsilon}$. Indeed, for all $y \in L^2(0,T,H)$ one has

$$
\begin{aligned}
\int_0^T (\bar{v}_{\nu\varepsilon}, y)\,\mathrm{d}t &\overset{(4.27)}{=} \lim_{n\to\infty}\int_0^T (z_{\nu\varepsilon n}, y)\,\mathrm{d}t = \lim_{n\to\infty}\int_0^T (v_{\nu\varepsilon n}, \pi_n(y))\,\mathrm{d}t \\
&= \lim_{n\to\infty}\left(\int_0^T (v_{\nu\varepsilon n}, y)\,\mathrm{d}t + \int_0^T (v_{\nu\varepsilon n}, \pi_n(y)-y)\,\mathrm{d}t\right) \overset{(4.25)}{=} \int_0^T (v_{\nu\varepsilon}, y)\,\mathrm{d}t
\end{aligned}
$$

where we used bound (4.19) and the fact that $\pi_n(y) \to y$ in $L^2(0,T;H)$ as $n\to\infty$. Moreover, from the linearity and maximal monotonicity of the map $\mathcal{B}$ defined with $X = V^*$ and $p = 2$ (see §2.4), from the fact that $z_{0\nu\varepsilon n} = \pi_n(v_{0\nu\varepsilon}) \to v_{0\nu\varepsilon}$ in H as $n\to\infty$, we readily obtain $\kappa*(v_{\nu\varepsilon}-v_{0\nu\varepsilon}) \in W^{1,2}(0,T;V^*)$ with $(\kappa*(v_{\nu\varepsilon}-v_{0\nu\varepsilon}))(0) = 0$ and $\xi_{\nu\varepsilon} = \partial_t(\kappa*(v_{\nu\varepsilon}-v_{0\nu\varepsilon}))$. Furthermore, we note that

$$
Au_{\nu\varepsilon n} \rightharpoonup Au_{\nu\varepsilon} \quad \text{in } \ L^2(0,T;V^*). \tag{4.29}
$$

Owing to convergences (4.24) and (4.26) we have that

$$\begin{aligned}\int_0^T (v_{\nu\varepsilon n}, u_{\nu\varepsilon n})\,\mathrm{d}t &= \int_0^T (z_{\nu\varepsilon n}, u_{\nu\varepsilon n})\,\mathrm{d}t = \int_0^T \langle z_{\nu\varepsilon n}, u_{\nu\varepsilon n}\rangle\,\mathrm{d}t \\ &\to \int_0^T \langle v_{\nu\varepsilon}, u_{\nu\varepsilon}\rangle\,\mathrm{d}t = \int_0^T (v_{\nu\varepsilon}, u_{\nu\varepsilon})\,\mathrm{d}t. \end{aligned} \tag{4.30}$$

By the classical tool [16, Proposition 2.5, p. 27], it follows from (4.12) that $v_{\nu\varepsilon} \in \partial\Psi_{\alpha_{\nu\varepsilon}}(u_{\nu\varepsilon})$ or, equivalently, $v_{\nu\varepsilon} = \alpha_{\nu\varepsilon}(u_{\nu\varepsilon})$ a.e. in Q.

The lower semicontinuity and convexity of $\Psi_{\alpha_{\nu\varepsilon}}$, convergences (4.24)–(4.25), and (4.30) ensure that

$$\begin{aligned}\Psi_{\alpha_{\nu\varepsilon}}(u_{\nu\varepsilon}) &\le \liminf_{n\to\infty} \Psi_{\alpha_{\nu\varepsilon}}(u_{\nu\varepsilon n}) \le \limsup_{n\to\infty} \Psi_{\alpha_{\nu\varepsilon}}(u_{\nu\varepsilon n}) \\ &\le \limsup_{n\to\infty} \left(\Psi_{\alpha_{\nu\varepsilon}}(u_{\nu\varepsilon}) + \int_0^T (v_{\nu\varepsilon n}, u_{\nu\varepsilon n} - u_{\nu\varepsilon})\,\mathrm{d}t \right) = \Psi_{\alpha_{\nu\varepsilon}}(u_{\nu\varepsilon}).\end{aligned}$$

This in particular proves that

$$\Psi_{\alpha_{\nu\varepsilon}}(u_{\nu\varepsilon n}) \to \Psi_{\alpha_{\nu\varepsilon}}(u_{\nu\varepsilon}).$$

Since the functional $\Psi_{\alpha_{\nu\varepsilon}}$ is strictly convex, we deduce from convergence (4.24) and [41, Theorem 3 (ii)] that indeed

$$u_{\nu\varepsilon n} \to u_{\nu\varepsilon} \quad \text{in } L^2(0,T;H). \tag{4.31}$$

Since $\beta_{\varepsilon H}$ is ε^{-1}-Lipschitz continuous in H, we also find that

$$\beta_{\varepsilon H}(u_{\nu\varepsilon n}) \to \beta_{\varepsilon H}(u_{\nu\varepsilon}) \quad \text{in } L^2(0,T;H). \tag{4.32}$$

Here we note that $\beta_{\varepsilon H}(u_{\nu\varepsilon}) \in L^2(0,T;V)$ as β_ε is ε^{-1}-Lipschitz continuous in $\mathbb{R}$ and $u_{\nu\varepsilon} \in L^2(0,T;V)$. The strong convergence (4.31) and the Lipschitz continuity in (2.10) allow us to check that

$$\begin{aligned}&\|(\pi_n \circ G)(u_{\nu\varepsilon n}) - G(u_{\nu\varepsilon})\|_{L^2(0,T;H)} \\ &\quad \le \|(\pi_n \circ G)(u_{\nu\varepsilon n}) - (\pi_n \circ G)(u_{\nu\varepsilon}))\|_{L^2(0,T;H)} + \|(\pi_n \circ G)(u_{\nu\varepsilon}) - G(u_{\nu\varepsilon})\|_{L^2(0,T;H)} \\ &\quad \le \Lambda_g \|u_{\nu\varepsilon n} - u_{\nu\varepsilon}\|_{L^2(0,T;H)} + \|(\pi_n \circ G)(u_{\nu\varepsilon}) - G(u_{\nu\varepsilon})\|_{L^2(0,T;H)} \to 0,\end{aligned}$$

so that

$$(\pi_n \circ G)(u_{\nu\varepsilon n}) \to G(u_{\nu\varepsilon}) \quad \text{in } L^2(0,T;H). \tag{4.33}$$

Moreover, one can similarly verify that

$$(\pi_n \circ \beta_{\varepsilon H})(u_{\nu\varepsilon n}) \to \beta_{\varepsilon H}(u_{\nu\varepsilon}) \quad \text{in } L^2(0,T;H). \tag{4.34}$$

Owing to convergences (4.28), (4.29), (4.33), and (4.34) we can pass to the limit in equation (4.11) and in bound (4.15). Taking also into account the uniqueness argument of Section 3, we have proved the following proposition.

Proposition 4.2 (Well-posedness of the regularized problem)**.** *Under assumptions* (A0)–(A5)*, for all* $\varepsilon > 0$ *and* $\nu \in (0,1)$*, set* $v_{0\nu\varepsilon} = \nu u_{0\varepsilon} + v_{0\varepsilon}$ *with* $u_{0\varepsilon}$ *and* $v_{0\varepsilon}$ *defined in*

(4.3)–(4.4). *Then there exists a unique triplet* $(u_{\nu\varepsilon}, v_{\nu\varepsilon}, w_{\nu\varepsilon}) \in L^2(0,T;V\times V\times V)$ *such that* $\kappa*(v_{\nu\varepsilon}-v_{0\nu\varepsilon}) \in W^{1,2}(0,T;V^*)$ *with* $(\kappa*(v_{\nu\varepsilon}-v_{0\nu\varepsilon}))(0)=0$ *and*

$$\partial_t(\kappa*(v_{\nu\varepsilon}-v_{0\nu\varepsilon})) + Au_{\nu\varepsilon} + w_{\nu\varepsilon} = G(u_{\nu\varepsilon}) \quad \textit{in } V^*\textit{, a.e. in } (0,T), \tag{4.35}$$

$$v_{\nu\varepsilon} = \alpha_{\nu\varepsilon}(u_{\nu\varepsilon}) \quad \textit{a.e. in } Q, \tag{4.36}$$

$$w_{\nu\varepsilon} = \beta_\varepsilon(u_{\nu\varepsilon}) \quad \textit{a.e. in } Q. \tag{4.37}$$

Moreover, there exists a constant $C_1>0$ *independent of* $\nu\in(0,1)$ *and* $\varepsilon\in(0,1)$ *such that*

$$\|u_{\nu\varepsilon}\|_{L^2(0,T;V)} \le C_1. \tag{4.38}$$

4.5. **Additional a-priori estimates for the regularized problem**

With the aim of passing to the limit in the regularizations, we complement bound (4.38) by proving some additional estimates for the solution $(u_{\nu\varepsilon}, v_{\nu\varepsilon}, w_{\nu\varepsilon})$ of the regularized problem (4.35)–(4.37).

Recalling that $w_{\nu\varepsilon} = \beta_{\varepsilon H}(u_{\nu\varepsilon}) \in L^2(0,T;V)$ and that $|\beta_\varepsilon(r)| \le 1/\varepsilon$ for all $r\in\mathbb{R}$, we readily check that $y_{\nu\varepsilon} := |w_{\nu\varepsilon}|^{q-2}w_{\nu\varepsilon} \in L^2(0,T;V)$ with $q>2$ given in (A4), as well. We can hence test (4.35) by $y_{\nu\varepsilon}$ and obtain

$$\langle\partial_t(\kappa*(v_{\nu\varepsilon}-v_{0\nu\varepsilon})), y_{\nu\varepsilon}\rangle + \langle Au_{\nu\varepsilon}, y_{\nu\varepsilon}\rangle + \|w_{\nu\varepsilon}\|^q_{L^q(\Omega)} = (G(u_{\nu\varepsilon}), y_{\nu\varepsilon}) \quad \text{a.e. in } (0,T). \tag{4.39}$$

The second term in the left-hand side of (4.39) is nonnegative, since we see that

$$\begin{aligned}\langle Au_{\nu\varepsilon}, y_{\nu\varepsilon}\rangle &= \int_\Omega \nabla u_{\nu\varepsilon}\cdot\nabla(|\beta_\varepsilon(u_{\nu\varepsilon})|^{q-2}\beta_\varepsilon(u_{\nu\varepsilon}))\,\mathrm{d}x\\ &= (q-1)\int_\Omega |\beta_\varepsilon(u_{\nu\varepsilon})|^{q-2}\beta_\varepsilon'(u_{\nu\varepsilon})|\nabla u_{\nu\varepsilon}|^2\,\mathrm{d}x \ge 0.\end{aligned} \tag{4.40}$$

The right-hand side of (4.39) can be bounded from the sublinearity assumption in (A4) as

$$\begin{aligned}(G(u_{\nu\varepsilon}), y_{\nu\varepsilon}) &\le \frac{1}{q'}\|w_{\nu\varepsilon}\|^q_{L^q(\Omega)} + \frac{1}{q}\|G(u_{\nu\varepsilon})\|^q_{L^q(\Omega)}\\ &\le \frac{1}{q'}\|w_{\nu\varepsilon}\|^q_{L^q(\Omega)} + C(1+\|u_{\nu\varepsilon}\|^2) \quad \text{a.e. in } (0,T).\end{aligned} \tag{4.41}$$

Using the fact that $\alpha_{\nu\varepsilon H}$ is invertible and recalling that $\beta_{\varepsilon q}(r) := |\beta_\varepsilon(r)|^{q-2}\beta_\varepsilon(r)$ for all $r\in\mathbb{R}$ as in Lemma 4.1(iv), one can write

$$y_{\nu\varepsilon} = \beta_{\varepsilon qH}(u_{\nu\varepsilon}) = (\beta_{\varepsilon qH}\circ\alpha_{\nu\varepsilon H}^{-1})(\alpha_{\nu\varepsilon H}(u_{\nu\varepsilon})) = (\beta_{\varepsilon qH}\circ\alpha_{\nu\varepsilon H}^{-1})(v_{\nu\varepsilon}).$$

As $v_{0\nu\varepsilon} \in D(\psi_{\beta_{\varepsilon q}\circ\alpha_{\nu\varepsilon}^{-1}})$ (see Lemma 4.1 (iv)), $v_{\nu\varepsilon}$, $y_{\nu\varepsilon} \in L^2(0,T;V)$, $\kappa*(v_{\nu\varepsilon}-v_{0\nu\varepsilon}) \in W^{1,2}(0,T;V^*)$ with $(\kappa*(v_{\nu\varepsilon}-v_{0\nu\varepsilon}))(0)=0$, $\psi_{\beta_{\varepsilon q}\circ\alpha_{\nu\varepsilon}^{-1}}(v_{\nu\varepsilon}) \in L^1(0,T)$ (indeed, one can prove it as in the proof of Lemma 4.1 (iv))), and $y_{\nu\varepsilon} = \beta_{\varepsilon qH}(u_{\nu\varepsilon}) = (\beta_{\varepsilon qH}\circ\alpha_{\nu\varepsilon H}^{-1})(v_{\nu\varepsilon})$ a.e. in $(0,T)$, by applying the nonlocal chain-rule inequality (2.12) we get

$$\ell*\langle\partial_t(\kappa*(v_{\nu\varepsilon}-v_{0\nu\varepsilon})), y_{\nu\varepsilon}\rangle \ge \int_\Omega \widehat{\beta_{\varepsilon q}\circ\alpha_{\nu\varepsilon}^{-1}}(v_{\nu\varepsilon})\,\mathrm{d}x - \int_\Omega \widehat{\beta_{\varepsilon q}\circ\alpha_{\nu\varepsilon}^{-1}}(v_{0\nu\varepsilon})\,\mathrm{d}x$$

a.e. in $(0,T)$. Hence, by convolving (4.39) with ℓ, also using (4.41) and Lemma 4.1(iv), we deduce that

$$\int_\Omega \widehat{\beta_{\varepsilon q}\circ\alpha_{\nu\varepsilon}^{-1}}(v_{\nu\varepsilon})\,\mathrm{d}x + \ell*\frac{1}{q}\|w_{\nu\varepsilon}\|^q_{L^q(\Omega)} \le C + \ell*C(1+\|u_{\nu\varepsilon}\|^2).$$

As $\widehat{\beta_{\varepsilon q} \circ \alpha_{\nu\varepsilon}^{-1}} \geq 0$, one has that

$$\ell * \frac{1}{q}\|w_{\nu\varepsilon}\|^q_{L^q(\Omega)} \leq C + \ell * C(1 + \|u_{\nu\varepsilon}\|^2).$$

By convolving with κ and using bound (4.38) we deduce that

$$\|w_{\nu\varepsilon}\|_{L^q(Q)} \leq C. \tag{4.42}$$

Testing (4.35) by $v_{\nu\varepsilon} \in L^2(0,T;V)$, we have

$$\langle \partial_t(\kappa * (v_{\nu\varepsilon} - v_{0\nu\varepsilon})), v_{\nu\varepsilon}\rangle + \langle A u_{\nu\varepsilon}, v_{\nu\varepsilon}\rangle + (w_{\nu\varepsilon}, v_{\nu\varepsilon}) = (G(u_{\nu\varepsilon}), v_{\nu\varepsilon}) \quad \text{a.e. in } (0,T). \tag{4.43}$$

The second and the third terms in the left-hand side of (4.43) are nonnegative since $\langle A u_{\nu\varepsilon}, v_{\nu\varepsilon}\rangle = \langle A u_{\nu\varepsilon}, \alpha_{\nu\varepsilon}(u_{\nu\varepsilon})\rangle \geq 0$, see (4.40), and

$$(w_{\nu\varepsilon}, v_{\nu\varepsilon}) = (\beta_{\varepsilon H}(u_{\nu\varepsilon}), \alpha_{\nu\varepsilon H}(u_{\nu\varepsilon})) = \int_\Omega \beta_\varepsilon(u_{\nu\varepsilon})\, \alpha_{\nu\varepsilon}(u_{\nu\varepsilon})\, \mathrm{d}x \geq 0$$

as $\beta_\varepsilon(r)\,\alpha_{\nu\varepsilon}(r) \geq 0$ for all $r \in \mathbb{R}$. Moreover, the right-hand side of (4.43) can be bounded as in (4.41) with the aid of (2.7), namely,

$$\begin{aligned}(G(u_{\nu\varepsilon}), v_{\nu\varepsilon}) &\leq \frac{1}{4\|\ell\|_{L^1(0,T)}}\|v_{\nu\varepsilon}\|^2 + \|\ell\|_{L^1(0,T)}\|G(u_{\nu\varepsilon})\|^2 \\ &\leq \frac{1}{4\|\ell\|_{L^1(0,T)}}\|v_{\nu\varepsilon}\|^2 + 2\|\ell\|_{L^1(0,T)}\left(\Lambda_g^2|\Omega| + \|u_{\nu\varepsilon}\|^2\right).\end{aligned}$$

As $v_{\nu\varepsilon} \in L^2(0,T;V)$ and $\kappa*(v_{\nu\varepsilon} - v_{0\nu\varepsilon}) \in W^{1,2}(0,T;V^*)$ along with $(\kappa*(v_{\nu\varepsilon} - v_{0\nu\varepsilon}))(0) = 0$, an application of the nonlocal chain-rule (2.12) for the functional $v \mapsto \psi_{\mathrm{id}}(v) = \|v\|^2/2$ ensures that

$$\ell * \langle \partial_t(\kappa * (v_{\nu\varepsilon} - v_{0\nu\varepsilon})), v_{\nu\varepsilon}\rangle \geq \frac{1}{2}\|v_{\nu\varepsilon}\|^2 - \frac{1}{2}\|v_{0\nu\varepsilon}\|^2 \quad \text{a.e. in } (0,T).$$

Taking the convolution of (4.43) with ℓ and using the above bounds we hence have that

$$\frac{1}{2}\|v_{\nu\varepsilon}\|^2 - \frac{1}{2}\|v_{0\nu\varepsilon}\|^2 \leq \ell * \frac{1}{4\|\ell\|_{L^1(0,T)}}\|v_{\nu\varepsilon}\|^2 + \ell * 2\|\ell\|_{L^1(0,T)}\left(\Lambda_g^2|\Omega| + \|u_{\nu\varepsilon}\|^2\right).$$

Recalling that $\|v_{0\nu\varepsilon}\| \leq \nu\|u_{0\varepsilon}\| + \|v_{0\varepsilon}\| \leq C$ by (4.5)–(4.6) and using the Young convolution inequality (2.6) along with bound (4.38) we conclude that

$$\|v_{\nu\varepsilon}\|_{L^2(0,T;H)} \leq C. \tag{4.44}$$

Finally, a comparison in (4.35), bounds (2.7), (4.38), (4.42), and the boundedness of $A : V \to V^*$ ensure that

$$\|\partial_t(\kappa * (v_{\nu\varepsilon} - v_{0\nu\varepsilon}))\|_{L^2(0,T;V^*)} \leq C. \tag{4.45}$$

4.6. **Passage to the limit as $\varepsilon \to 0$**

Bound (4.38) and the a-priori estimates from Section 4.5 allow us to pass to the limit as $\varepsilon \to 0$ and $\nu \to 0$. In fact, these limits can be taken simultaneously. Still, we prefer to pass to the limit as $\varepsilon \to 0$ first, in order to obtain an intermediate existence result, which could be of independent interest. The parameter $\nu \in (0,1)$ is hence kept fixed in this section. The limit as $\nu \to 0$ is discussed in Section 4.7 below.

From bounds (4.38), (4.42), and (4.44)–(4.45), by passing to the limit (up to a subsequence) as $\varepsilon \to 0$, we get

$$u_{\nu\varepsilon} \rightharpoonup u_\nu \quad \text{in } L^2(0,T;V), \tag{4.46}$$

$$v_{\nu\varepsilon} \rightharpoonup v_\nu \quad \text{in } L^2(0,T;H), \tag{4.47}$$

$$w_{\nu\varepsilon} \rightharpoonup w_\nu \quad \text{in } L^q(\Omega \times (0,T)), \tag{4.48}$$

$$\partial_t(\kappa * (v_{\nu\varepsilon} - v_{0\nu\varepsilon})) \rightharpoonup \xi_\nu \quad \text{in } \ L^2(0,T;V^*). \tag{4.49}$$

Arguing as in Section 4.4, bound (4.45) guarantees that the time increments of $v_{\nu\varepsilon}$ can be controlled as

$$\limsup_{h\to 0} \sup_{\varepsilon\in(0,1)} \int_0^{T-h} \|v_{\nu\varepsilon}(t+h) - v_{\nu\varepsilon}(t)\|_{V^*}^2 \,\mathrm{d}t = 0. \tag{4.50}$$

The Aubin–Lions lemma then gives

$$v_{\nu\varepsilon} \to v_\nu \quad \text{in } L^2(0,T;V^*). \tag{4.51}$$

Together with convergence (4.46), this ensures that

$$\int_0^T (v_{\nu\varepsilon}, u_{\nu\varepsilon}) \,\mathrm{d}t = \int_0^T \langle v_{\nu\varepsilon}, u_{\nu\varepsilon}\rangle \,\mathrm{d}t \to \int_0^T \langle v_\nu, u_\nu\rangle \,\mathrm{d}t = \int_0^T (v_\nu, u_\nu) \,\mathrm{d}t. \tag{4.52}$$

Recall that we have the convergence (4.1), in particular, $\partial\Psi_{\alpha_{\nu\varepsilon}} \to \partial\Psi_{\alpha_\nu}$ in the graph sense in $L^2(0,T;H)$. Given the convergences (4.46)–(4.47) and (4.52), the extension (2.3) of [11, Proposition 3.59, p. 361] guarantees that $v_\nu \in \partial\Psi_{\alpha_\nu}(u_\nu)$, namely, $v_\nu \in \alpha_\nu(u_\nu)$ a.e. in Q.

At the same time, $\Psi_{\alpha_{\nu\varepsilon}} \to \Psi_{\alpha_\nu}$ in the Mosco sense in $L^2(0,T;H)$, so that convergence (4.46) and the lim inf inequality (2.4) imply that

$$\Psi_{\alpha_\nu}(u_\nu) \le \liminf_{\varepsilon\to 0} \Psi_{\alpha_{\nu\varepsilon}}(u_{\nu\varepsilon}). \tag{4.53}$$

Hence, we can use the pointwise upper bound $\Psi_{\alpha_{\nu\varepsilon}}(y) \le \Psi_{\alpha_\nu}(y)$ for all $y \in L^2(0,T;H)$ and all $\varepsilon > 0$ (indeed, we have $\widehat{\alpha_{\nu\varepsilon}} \le \widehat{\alpha}_\nu$ in $\mathbb{R}$), and limit (4.52) in order to get

$$\begin{aligned}\Psi_{\alpha_\nu}(u_\nu) &\overset{(4.53)}{\le} \liminf_{\varepsilon\to 0} \Psi_{\alpha_{\nu\varepsilon}}(u_{\nu\varepsilon}) \le \limsup_{\varepsilon\to 0} \Psi_{\alpha_{\nu\varepsilon}}(u_{\nu\varepsilon}) \\ &\le \limsup_{\varepsilon\to 0} \left(\Psi_{\alpha_{\nu\varepsilon}}(u_\nu) + \int_0^T (v_{\nu\varepsilon}, u_{\nu\varepsilon} - u_\nu) \,\mathrm{d}t \right) \le \Psi_{\alpha_\nu}(u_\nu),\end{aligned}$$

which gives the convergence $\Psi_{\alpha_{\nu\varepsilon}}(u_{\nu\varepsilon}) \to \Psi_{\alpha_\nu}(u_\nu)$ as $\varepsilon \to 0$. We can similarly verify that $\Psi_{\alpha_\varepsilon}(u_{\nu\varepsilon}) \to \Psi_\alpha(u_\nu)$ as $\varepsilon \to 0$. Indeed, noting that $\alpha_\varepsilon(u_{\nu\varepsilon}) = v_{\nu\varepsilon} - \nu u_{\nu\varepsilon} \rightharpoonup v_\nu - \nu u_\nu =: a_\nu$ weakly in $L^2(0,T;H)$, we have

$$\begin{aligned}&\limsup_{\varepsilon\to 0} (\alpha_\varepsilon(u_{\nu\varepsilon}), u_{\nu\varepsilon}) \\ &\le \limsup_{\varepsilon\to 0} (\alpha_{\nu\varepsilon}(u_{\nu\varepsilon}), u_{\nu\varepsilon}) - \liminf_{\varepsilon\to 0} \nu\|u_{\nu\varepsilon}\|^2 \\ &\le (v_\nu, u_\nu) - \nu\|u_\nu\|_H^2 = (a_\nu + \nu u_\nu, u_\nu) - \nu\|u_\nu\|_H^2 = (a_\nu, u_\nu),\end{aligned}$$

which also enables us to check $\Psi_{\alpha_\varepsilon}(u_{\nu\varepsilon}) \to \Psi_\alpha(u_\nu)$ as $\varepsilon \to 0$ in a similar fashion. Therefore it follows that

$$\begin{aligned}\limsup_{\varepsilon\to0} \frac{\nu}{2}\|u_{\nu\varepsilon}\|^2_{L^2(0,T;H)} &= \limsup_{\varepsilon\to0}\left(\Psi_{\alpha_{\nu\varepsilon}}(u_{\nu\varepsilon}) - \Psi_{\alpha_\varepsilon}(u_{\nu\varepsilon})\right)\\ &\le \limsup_{\varepsilon\to0}\Psi_{\alpha_{\nu\varepsilon}}(u_{\nu\varepsilon}) - \liminf_{\varepsilon\to0}\Psi_{\alpha_\varepsilon}(u_{\nu\varepsilon})\\ &\le \Psi_{\alpha_\nu}(u_\nu) - \Psi_\alpha(u_\nu) = \frac{\nu}{2}\|u_\nu\|^2_{L^2(0,T;H)},\end{aligned}$$

which, along with convergence (4.46) and the uniform convexity of $\|\cdot\|_{L^2(0,T;H)}$, implies that

$$u_{\nu\varepsilon} \to u_\nu \quad \text{in } L^2(0,T;H). \tag{4.54}$$

Moving from this strong convergence, the continuity of G in (2.8) gives that $G(u_{\nu\varepsilon}) \to G(u_\nu)$ in $L^2(0,T;H)$ and we also have

$$\int_0^T (w_{\nu\varepsilon}, u_{\nu\varepsilon})\,\mathrm{d}t \to \int_0^T (w_\nu, u_\nu)\,\mathrm{d}t.$$

As $\partial\Psi_{\beta_\varepsilon} \to \partial\Psi_\beta$ in the graph sense in $L^2(0,T;H)$, see (4.1), this last convergence and convergences (4.46) and (4.48) allow us to apply again the extension (2.3) of [11, Proposition 3.59, p. 361] and obtain $w_\nu \in \partial\Psi_\beta(u_\nu)$, that is, $w_\nu \in \beta(u_\nu)$ a.e. in Q. Using the linear maximal monotonicity and m-accretivity of A and $\mathcal{B}$ with $X = V^*$ and $p = 2$ (see §2.4), respectively, one can pass to the limit as $\varepsilon \to 0$ in equation (4.35) as well as in bounds (4.38), (4.42), and (4.44)–(4.45), eventually obtaining the following.

Proposition 4.3 (Well-posedness of the regularized problem for $\nu \in (0,1)$)**.** *Under assumptions (A0)–(A5), for all $\nu \in (0,1)$, set $v_{0\nu} := \nu u_0 + v_0$. Then, there exists a triplet $(u_\nu, v_\nu, w_\nu) \in L^2(0,T; V \times H \times H)$ such that $\kappa * (v_\nu - v_{0\nu}) \in W^{1,2}(0,T;V^*)$ along with $(\kappa * (v_\nu - v_{0\nu}))(0) = 0$ and $w_\nu \in L^q(Q)$, and*

$$\partial_t(\kappa * (v_\nu - v_{0\nu})) + Au_\nu + w_\nu = G(u_\nu) \quad \textit{in } V^*,\ \textit{a.e. in } (0,T), \tag{4.55}$$

$$v_\nu \in \alpha_\nu(u_\nu) \quad \textit{a.e. in } Q, \tag{4.56}$$

$$w_\nu \in \beta(u_\nu) \quad \textit{a.e. in } Q. \tag{4.57}$$

Moreover, there exists a constant $C_2 > 0$ independent of ν such that

$$\begin{aligned}&\|u_\nu\|_{L^2(0,T;V)} + \|v_\nu\|_{L^2(0,T;H)} + \|w_\nu\|_{L^q(Q)}\\ &\quad + \|\partial_t(\kappa * (v_\nu - v_{0\nu}))\|_{L^2(0,T;V^*)} \le C_2.\end{aligned} \tag{4.58}$$

4.7. Passage to the limit as $\nu \to 0$

We now pass to the limit as $\nu \to 0$ for the solutions (u_ν, v_ν, w_ν) to problem (4.55)–(4.57) for proving the existence of a solution to (2.13)–(2.15). Owing to bounds (4.58) we can extract a (not relabeled) subsequence such that

$$u_\nu \rightharpoonup u \quad \text{in } L^2(0,T;V), \tag{4.59}$$

$$v_\nu \rightharpoonup v \quad \text{in } L^2(0,T;H), \tag{4.60}$$

$$w_\nu \rightharpoonup w \quad \text{in } L^q(Q), \tag{4.61}$$

$$\partial_t(\kappa * (v_\nu - v_{0\nu})) \rightharpoonup \xi \quad \text{in } L^2(0,T;V^*). \tag{4.62}$$

Again, thanks to the bound on $\partial_t(\kappa * (v_\nu - v_{0\nu}))$ in $L^2(0,T;V^*)$ from (4.58) we can control the time increments of v_ν as

$$\limsup_{h\to 0} \sup_{\nu\in(0,1)} \int_0^{T-h} \|v_\nu(t+h) - v_\nu(t)\|_{V^*}^2 \, \mathrm{d}t = 0,$$

so that, possibly extracting again without relabeling, the Aubin–Lions lemma gives

$$v_\nu \to v \quad \text{in } L^2(0,T;V^*). \tag{4.63}$$

By reproducing the argument in (4.52) we obtain

$$\int_0^T (v_\nu, u_\nu)\,\mathrm{d}t \to \int_0^T (v,u)\,\mathrm{d}t. \tag{4.64}$$

Using the convergence $\partial\Psi_{\alpha_\nu} \to \partial\Psi_\alpha$ in the graph sense in $L^2(0,T;H)$, see (4.2), the extension (2.3) of [11, Proposition 3.59, p. 361] guarantees that $v \in \partial\Psi_\alpha(u)$, which is nothing but (2.14).

A second consequence of (4.64), together with the bound (4.58) and convergence (4.59), is that

$$\begin{aligned}
\Psi_\alpha(u) &\le \liminf_{\nu\to 0} \Psi_\alpha(u_\nu) \le \limsup_{\nu\to 0} \Psi_\alpha(u_\nu) \le \limsup_{\nu\to 0} \Psi_{\alpha_\nu}(u_\nu) \\
&\le \limsup_{\nu\to 0} \left(\Psi_{\alpha_\nu}(u) + \int_0^T (v_\nu, u_\nu - u)\,\mathrm{d}t \right) \\
&= \limsup_{\nu\to 0} \left(\Psi_\alpha(u) + \frac{\nu}{2}\|u\|_{L^2(0,T;H)}^2 + \int_0^T (v_\nu, u_\nu - u)\ \mathrm{d}t \right) = \Psi_\alpha(u).
\end{aligned}$$

This proves that $\Psi_\alpha(u_\nu) \to \Psi_\alpha(u)$. As $\widehat{\alpha}$ is strictly convex by (A2), [41, Theorem 3 (ii)] and convergence (4.59) ensure that

$$u_\nu \to u \quad \text{in } L^\sigma(0,T;H)$$

for any $\sigma \in [1,2)$. In particular, putting $\sigma = q' \in (1,2)$ (indeed, $q > 2$), we obtain

$$\int_0^T (w_\nu, u_\nu)\,\mathrm{d}t \to \int_0^T (w,u)\,\mathrm{d}t.$$

As convergences (4.59) and (4.61) also imply weak convergence in $L^2(0,T;H)$ the latter allows us to apply [16, Proposition 2.5, p. 27] and obtain the inclusion $w \in \partial\Psi_\beta(u)$, that is (2.15). Moreover, we deduce that $G(u_\nu) \to G(u)$ in $L^\sigma(0,T;H)$ for every $\sigma \in [1,2)$ due to (2.8). Using the linear maximal monotonicity and m-accretivity of A and $\mathcal{B}$ with $X = V^*$ and $p = 2$ (see §2.4), respectively, one can pass to the limit in (4.55) as $\varepsilon \to 0$ and obtain (2.13). This concludes the proof of Theorem 2.1.

Acknowledgments

This research was funded in whole or in part by the Austrian Science Fund (FWF) projects 10.55776/F65, 10.55776/I5149, 10.55776/P32788, as well as by the OeAD-WTZ project CZ 09/2023. For open-access purposes, the authors have applied a CC BY public copyright license to any author-accepted manuscript version arising from this submission. G.A. is supported by JSPS KAKENHI Grant Numbers JP24H00184, JP21KK0044, JP21K18581, JP20H01812 and JP20H00117. Part of this research was conducted during a visit to the Mathematical Institute at Tohoku University, whose warm hospitality is gratefully acknowledged.

References

[1] E. Affili, E. Valdinoci. Decay estimates for evolution equations with classical and fractional time-derivatives. *J. Differential Equations* 266 (2019), 4027–4060.

[2] S. Aizicovici, V.-M. Hokkanen. Doubly nonlinear equations with unbounded operators. *Nonlinear Anal.* 58 (2004), no. 5-6:591–607.

[3] S. Aizicovici, V.-M. Hokkanen. Doubly nonlinear periodic problems with unbounded operators. *J. Math. Anal. Appl.* 292 (2004), no. 2:540–557.

[4] G. Akagi. Doubly nonlinear evolution equations governed by time-dependent subdifferentials in reflexive Banach spaces. *J. Differential Equations* 231 (2006), no. 1, 32–56.

[5] G. Akagi. Doubly nonlinear evolution equations with non-monotone perturbations in reflexive Banach spaces. *J. Evol. Equ.* 11 (2011), no. 1, 1–41.

[6] G. Akagi. Fractional flows driven by subdifferentials in Hilbert spaces. *Israel J. Math.* 234 (2019), no. 2, 809–862.

[7] G. Akagi, Y. Nakajima. Time-fractional gradient flows for nonconvex energies in Hilbert spaces. `arXiv:2501.08059 [math.AP].`

[8] G. Akagi, G. Schimperna. On a class of doubly nonlinear evolution equations in Musielak-Orlicz spaces. *Math. Nachr.* 297 (2024), no. 7, 2686–2729.

[9] G. Akagi, U. Stefanelli. Doubly nonlinear equations as convex minimization. *SIAM J. Math. Anal.* 46 (2014), no. 3, 1922–1945.

[10] H.W. Alt, S. Luckhaus. Quasilinear elliptic-parabolic differential equations. *Math. Z.* 183 (1983), no. 3, 311–341.

[11] H. Attouch. *Variational convergence for functions and operators.* Applicable Mathematics Series. Pitman (Advanced Publishing Program), Boston, MA, 1984.

[12] V. Barbu. Existence for nonlinear Volterra equations in Hilbert spaces. *SIAM J. Math. Anal.* 10 (1979), 552–569.

[13] F. Bernis. Existence results for doubly nonlinear higher order parabolic equations on unbounded domains. *Math. Ann.* 279 (1988), no. 3, 373–394.

[14] V. Bögelein, M. Strunk. A comparison principle for doubly nonlinear parabolic partial differential equations. *Ann. Mat. Pura Appl. (4)* 203 (2024), no. 2, 779–804.

[15] M. Bonforte, M. Gualdani, P. Ibarrondo. Time-fractional porous medium type equations: sharp time decay and regularization. `arXiv:2401.03234 [math.AP]`

[16] H. Brézis. *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert.* North-Holland Publishing Co., Amsterdam, 1973. North-Holland Mathematics Studies, no. 5. Notas de Matemática (50).

[17] E. DiBenedetto, R.E. Showalter. Implicit degenerate evolution equations and applications. *SIAM J. Math. Anal.* 12 (1981), no. 5, 731–751.

[18] S. Dipierro, E. Valdinoci, V. Vespri. Decay estimates for evolutionary equations with fractional time-diffusion. *J. Evol. Equ.* 19 (2019), 435–462.

[19] H. Gajewski, I.V. Skrypnik. To the uniqueness problem for nonlinear parabolic equations. *Discrete Contin. Dyn. Syst.* 10 (2004), no. 1-2, 315–336.

[20] Y. Giga, T. Namba. Well-posedness of Hamilton–Jacobi equations with Caputo's time-fractional derivative. *Comm. Partial Differential Equations*, 42 (2017), 1088–1120.

[21] O. Grange, F. Mignot. Sur la résolution d'une équation et d'une inéquation paraboliques non linéaires. *J. Funct. Anal.* 11 (1972), 77–92.

[22] G. Gripenberg, S.-O. Londen, O. Staffans. *Volterra integral and functional equations.* Encyclopedia Math. Appl., vol. 34, Cambridge University Press, Cambridge, 1990.

[23] V.-M. Hokkanen. An implicit nonlinear time dependent equation has a solution. *J. Math. Anal. Appl.* 161 (1991), no. 1, 117–141.

[24] V.-M. Hokkanen. Existence for nonlinear time dependent Volterra equations in Hilbert spaces. *An. Ştiinţ. Univ. Al. I. Cuza Iaşi Secţ. I a Mat.* 38 (1992), no. 1, 29–49.

[25] V.-M. Hokkanen. On nonlinear Volterra equations in Hilbert spaces. *Differential Integral Equations* 5 (1992), no. 3, 647–669.

[26] G. Huaroto. Doubly nonlinear isotropic fractional parabolic equation. *J. Elliptic Parabol. Equ.* 10 (2024), no. 2, 915–941.

[27] N. Kato, M. Misawa, K. Nakamura, Y. Yamaura. Existence for doubly nonlinear fractional p-Laplacian equations. *Ann. Mat. Pura Appl. (4)* 203 (2024), no. 6, 2481–2527.

[28] W. Li, A.J. Salgado. Time fractional gradient flows: theory and numerics. *Math. Models Methods Appl. Sci.* **33** (2023), no. 2, 377–453.

[29] E. Maitre, P. Witomski. A pseudo-monotonicity adapted to doubly nonlinear elliptic-parabolic equations. *Nonlinear Anal.* 50 (2002), 223–250.

[30] K. Moring, L. Schätzler, C. Scheven. Higher integrability for singular doubly nonlinear systems. *Ann. Mat. Pura Appl. (4)* 203 (2024), no. 5, 2235–2274.

[31] Ł. Płociniczak. Approximation of the Erdélyi–Kober operator with application to the time-fractional porous medium equation. *SIAM J. Appl. Math.* 74 (2014), 1219–1237.

[32] Ł. Płociniczak. Analytical studies of a time-fractional porous medium equation. Derivation, approximation and applications. *Commun. Nonlinear Sci. Numer. Simul.* 24 (2015), 169–183.

[33] Ł. Płociniczak, H. Okrasińska. Approximate self-similar solutions to a nonlinear diffusion equation with time-fractional derivative. *Phys. D* 261 (2013), 85–91.

[34] U. Stefanelli. Well-posedness and time discretization of a nonlinear Volterra integrodifferential equation. *J. Integral Equations Appl.* 13 (2001), no. 3, 273–304.

[35] U. Stefanelli. On a class of doubly nonlinear nonlocal evolution equations. *Differential Integral Equations* 15 (2002), no. 8, 897–922.

[36] U. Stefanelli. On some nonlocal evolution equations in Banach spaces. *J. Evol. Equ.* 4 (2004), no. 1, 1–26.

[37] U. Stefanelli. The Brezis-Ekeland principle for doubly nonlinear equations. *SIAM J. Control Optim.* 47 (2008), 1615–1642.

[38] E. Topp, M. Yangari. Existence and uniqueness for parabolic problems with Caputo time derivative. *J. Differential Equations* 262 (2017), 6018–6046.

[39] V. Vergara, R. Zacher. A priori bounds for degenerate and singular evolutionary partial integro-differential equations. *Nonlinear Anal.* 73 (2010), 3572–3585.

[40] V. Vergara, R. Zacher. Optimal decay estimates for time-fractional and other nonlocal subdiffusion equations via energy methods. *SIAM J. Math. Anal.* 47 (2015), 210–239.

[41] A. Visintin. Strong convergence results related to strict convexity. *Comm. Partial Differential Equations* 9 (1984), no. 5, 439–466.

[42] A. Visintin. *Models of phase transitions*, Progress in Nonlinear Differential Equations and Their Applications, vol. 28, Birkhäuser, 1996.

[43] R. Zacher, Weak solutions of abstract evolutionary integro-differential equations in Hilbert spaces. *Funkcial. Ekvac.* 52 (2009), 1–18.

(Goro Akagi) MATHEMATICAL INSTITUTE AND GRADUATE SCHOOL OF SCIENCE, TOHOKU UNIVERSITY, SENDAI 980-8578, JAPAN

Email address: `goro.akagi@tohoku.ac.jp`

(Giacomo Enrico Sodini) FACULTY OF MATHEMATICS, UNIVERSITY OF VIENNA, OSKAR-MORGENSTERN-PLATZ 1, A-1090 VIENNA, AUSTRIA

Email address: `giacomo.sodini@univie.ac.at`

(Ulisse Stefanelli) FACULTY OF MATHEMATICS, UNIVERSITY OF VIENNA, OSKAR-MORGENSTERN-PLATZ 1, A-1090 VIENNA, AUSTRIA & VIENNA RESEARCH PLATFORM ON ACCELERATING PHOTOREACTION DISCOVERY, UNIVERSITY OF VIENNA, WÄHRINGERSTRASSE 17, A-1090 WIEN, AUSTRIA.

Email address: `ulisse.stefanelli@univie.ac.at`