SIAM J. MATH. ANAL.
Vol. 46, No. 3, pp. 1922–1945

DOUBLY NONLINEAR EQUATIONS AS CONVEX MINIMIZATION*

GORO AKAGI† AND ULISSE STEFANELLI‡

Abstract. We present a variational reformulation of a class of doubly nonlinear parabolic equations as (limits of) constrained convex minimization problems. In particular, an ε-dependent family of *weighted energy-dissipation* (WED) functionals on entire trajectories is introduced and proved to admit minimizers. These minimizers converge to solutions of the original doubly nonlinear equation as $\varepsilon \to 0$. The argument relies on the suitable *dualization* of the former analysis of [G. Akagi and U. Stefanelli, *J. Funct. Anal.*, 260 (2011), pp. 2541–2578] and results in a considerable extension of the possible application range of the WED functional approach to nonlinear diffusion phenomena, including the Stefan problem and the porous media equation.

Key words. doubly nonlinear equations, convex minimization, duality

AMS subject classifications. 58E30, 47J35, 35K92

DOI. 10.1137/13091909X

1. Introduction.

This note is concerned with the description of a global variational approach to doubly nonlinear evolution equations. In particular, our discussion covers the case of the doubly nonlinear PDE

$$(\alpha(u))_t - \operatorname{div}\big(\beta(\nabla u)\big) \ni f, \tag{1.1}$$

where $\alpha \subset \mathbb{R} \times \mathbb{R}$ and $\beta \subset \mathbb{R}^d \times \mathbb{R}^d$ are (possibly multivalued) maximal monotone graphs (e.g., $\beta(\nabla u) = |\nabla u|^{m-2}\nabla u$) and $f = f(x,t)$ is given. The two monotone graphs α and β are assumed to show some polynomial growth of power $p > 1$ and $m > 1$, respectively, and (1.1) is posed in the space-time domain $\Omega \times (0,T)$, where $\Omega \subset \mathbb{R}^d\,(d \geq 1)$ is a bounded and open set with smooth boundary $\partial\Omega$, and complemented with homogeneous Dirichlet conditions (for definiteness) and some initial condition $\alpha(u)(0) \ni v_0$.

The differential problem is classically related to nonlinear diffusion phenomena. In particular, owing to the choice of the graph α, (1.1) may arise in a variety of different situations connected, for instance, with the Stefan problem, the porous media equation, or the Hele-Shaw model. By letting H denote the Heaviside graph, the choices $\alpha = \mathrm{id} + H$, $\alpha(u) = |u|^{p-2}u$ for some $p \in (1,2)$, and $\alpha = H$ correspond to the above mentioned models, respectively. The reader is referred to Visintin [50] for a detailed discussion on the relevance of relation (1.1) in the framework of phase transitions. As for the analytical treatment of (1.1) we limit ourselves to mentioning the classical references of Grange and Mignot [22], Barbu [9], Di Benedetto and Showalter [18], Alt and Luckhaus [5], and Bernis [11], and we further refer the reader to the contributions [1, 2, 19, 24, 25, 26, 33, 46, 47, 48], among many others.

*Received by the editors April 30, 2013; accepted for publication (in revised form) January 27, 2014; published electronically June 5, 2014. This research was supported by the joint CNR–JSPS project VarEvol: Innovative Variational Methods for Evolution PDEs.
http://www.siam.org/journals/sima/46-3/91909.html

†Graduate School of System Informatics, Kobe University, 1-1 Rokkodai-cho, Nada-ku, Kobe 657-8501 Japan (akagi@port.kobe-u.ac.jp). This author's research was supported by JSPS KAKENHI grant 25400163, by Hyogo Science and Technology Association, and by Nikko Co. Ltd.

‡Istituto di Matematica Applicata e Tecnologie Informatiche E. Magenes–CNR, v. Ferrata 1, I-27100 Pavia, Italy, and Faculty of Mathematics, University of Vienna, Oskar-Morgenstern-Platz 1, A-1090 Vienna, Austria (ulisse.stefanelli@imati.cnr.it). This author's research was supported by FP7-IDEAS-ERC-StG grant 200947 BioSMA.

The aim of this paper is to draw a connection between the differential problem (1.1) and a family of constrained convex minimization problems. This will be done in two steps. At first, we resort to *dualizing* (1.1), namely, we transform it into an equivalent problem in the unknown $v = \alpha(u)$. This reads $-\mathrm{div}\big(\beta(\nabla(\alpha^{-1}(v)))\big) \ni f - v_t$. Let now $B^* : (W_0^{1,m}(\Omega))^* \to W_0^{1,m}(\Omega)$; $g \mapsto z$ denote the solution operator related to the nonlinear elliptic problem $-\mathrm{div}\big(\beta(\nabla z)\big) \ni g$ along with the Dirichlet conditions $z|_{\partial\Omega} = 0$. The equation is hence rewritten as $-B^*(f - v_t) + \alpha^{-1}(v) \ni 0$. In particular, if $\beta = \mathrm{id}$, then $B^* = (-\Delta)^{-1}$, and the latter is nothing but the classical dual formulation of (1.1) in $H^{-1}(\Omega)$ [13].

Second, we introduce the so-called weighted energy-dissipation (WED) functionals W_ε defined on entire trajectories $v = v(x,t)$ as

(1.2)
$$W_\varepsilon(v) := \begin{cases} \displaystyle\int_0^T \mathrm{e}^{-t/\varepsilon}\left(\varepsilon\varphi^*(f - v_t) + \int_\Omega \widehat{\alpha}^*(v)\mathrm{d}x\right)\mathrm{d}t \\ \text{if} \quad v \in W^{1,m'}(0,T;(W_0^{1,m}(\Omega))^*) \cap L^{p'}(\Omega\times(0,T)) \ \text{ and } \ v(0) = v_0, \\ \infty \quad \text{otherwise.} \end{cases}$$

Here, $p' = p/(p-1)$, φ^* indicates a potential of the cyclic monotone operator B^*, and $\widehat{\alpha}$ denotes a primitive of α. Namely, $\alpha = \partial\widehat{\alpha}$, where the symbol ∂ denotes the subdifferential in the sense of convex analysis [13] (hence, $\partial\widehat{\alpha}^* \equiv \alpha^{-1}$). For instance, the choice $\beta = \mathrm{id}$ gives back $\varphi^*(\cdot) = (1/2)\|\cdot\|^2_{H^{-1}(\Omega)}$.

The WED functional approach for (1.1) consists in considering the minimizer v_ε of the WED functional W_ε, computing the limit $v_\varepsilon \to v$, as $\varepsilon \to 0$, and checking that indeed $u = \alpha^{-1}(v)$ is a solution of (1.1). The implementation of this strategy entails the possibility of recasting the doubly nonlinear differential problem (1.1) in the form of a (family of) constrained convex minimization problems (followed by the limit $\varepsilon \to 0$). In particular, by providing a global variational formulation for (1.1) we are entitled to directly use on the differential problem the general tools of the calculus of variations such as the direct method, Γ-convergence, and relaxation. This new variational approach provides a novel strategy in order to tackle the doubly nonlinear problem (1.1). In particular, this new perspective allows for some extension of the known existence theory for doubly nonlinear equations, as addressed at the end of section 4.

Apart from former contributions in the linear case (see the classical monograph by Lions and Magenes [30]), the WED formalism was first considered by Ilmanen [27] in the context of mean curvature flow (see also [44]). See also [10] and [23], where similar variational formulations are introduced to construct approximate solutions for some nonlinear evolution equations. An application in mechanics is proposed by Conti and Ortiz [17] and Larsen, Ortiz, and Richardson [29]. The WED approach for abstract gradient flows in Hilbert and metric spaces has been provided in [38] and [42, 39], respectively. The case of rate-independent evolutions is treated by Mielke and Ortiz [36] and further detailed in [37]. Let us mention that the WED formalism has a quite natural counterpart in the hyperbolic case. In particular, a variational approach to Lagrangian mechanics is presented in [32], whereas semilinear wave equations and mixed hyperbolic-parabolic equations are treated in [45, 43] and [31], respectively.

Within the WED literature, the papers [3, 4] are specifically related to the present contribution as they are focused on a *different class* of doubly nonlinear equation which is exemplified by the following nonlinear PDE:

$$\alpha(u_t) - \operatorname{div}\big(\beta(\nabla u)\big) \ni g. \tag{1.3}$$

In particular, the WED approach to the latter is based on the minimization of the functionals $F_\varepsilon : W^{1,p}(0,T;L^p(\Omega)) \cap L^m(0,T;W_0^{1,m}(\Omega)) \to [0,\infty]$ given by

$$F_\varepsilon(u) := \int_0^T \int_\Omega \mathrm{e}^{-t/\varepsilon} \left(\varepsilon\widehat{\alpha}(u_t) + \widehat{\beta}(\nabla u)\right) \mathrm{d}x\,\mathrm{d}t,$$

where $\widehat{\beta}$ is a primitive of β. In fact, the WED variational approach to (1.1) consists in equivalently reformulating our original problem in the form of (1.3) and exploiting (after some suitable extension) some ideas from [3, 4].

We shall present some preliminary material and state our main results in section 2. Then, section 3 is devoted to the study of the Euler–Lagrange equation for the WED functional W_ε. Finally, proofs of the main results are reported in section 4.

Before closing this introduction, it is worth mentioning the celebrated Brézis–Ekeland–Nayroles (BEN) principle [15, 16, 40, 41], which provides a fairly different approach to nonlinear evolution by global convex minimization. The BEN principle has recently attracted attention and has been extended to the doubly nonlinear equation (1.3) in [49] and to general maximal monotone operators by Visintin [51, 52, 53]. Moreover, it is the basis of the development of the far-reaching theory of self-dual lagrangians by Ghoussoub [20, 21]. See also [34, 35] for an application of this concept in a time-dependent setting. The BEN principle is based on the reformulation of the evolution equation as a *null*-convex minimization problem. In particular, no limiting procedure is involved and the difficulty resides in the proof that the minimum value zero is actually achieved. The WED formalism differs from the BEN principle as it consists in a genuine minimization problem combined with a limiting procedure. As such, it appears to be better suited for proving existence results.

2. Assumptions and main results.

2.1. Convex analysis preliminaries. Let us start by introducing some notation and recalling some classical facts from convex analysis. Let E be a Banach space and denote by $\langle\cdot,\cdot\rangle$ the duality pairing between E^* (dual) and E. Let $\phi : E \to (-\infty,\infty]$ be a proper (i.e., $\phi \not\equiv \infty$), convex, and lower semicontinuous functional. We denote by $\partial\phi : E \to 2^{E^*}$ (or $\partial_E\phi$ if any confusion may arise) its subdifferential in the sense of convex analysis, namely,

$$\partial\phi(u) := \{\xi \in E^* \colon \phi(v) - \phi(u) \geq \langle \xi, v-u\rangle \quad \text{for all} \quad v \in D(\phi)\}$$

for $u \in D(\phi) := \{v \in E \colon \phi(v) < \infty\}$ (*effective domain* of ϕ) with the domain $D(\partial\phi) := \{v \in D(\phi) \colon \partial\phi(v) \neq \emptyset\}$. Then, $\mathrm{D}\phi : E \to E^*$ (or $\mathrm{D}_E\phi$) denotes the Gâteaux or Fréchet derivative of ϕ, when ϕ is Gâteaux or Fréchet differentiable (then $\mathrm{D}\phi = \partial\phi$), and $\phi^* : E^* \to (-\infty,\infty]$ is its Legendre–Fenchel conjugate $\phi^*(\xi) := \sup_{u\in E}\{\langle\xi,u\rangle - \phi(u)\}$ for $\xi \in E^*$. In particular, ϕ^* is proper, lower semicontinuous, and convex in E^* and $\phi^*(f) = \langle f,u\rangle - \phi(u)$ iff $f \in \partial\phi(u)$. Eventually,

$$u \in \partial\phi^*(f) \text{ iff } f \in \partial\phi(u). \tag{2.1}$$

Clearly, whenever $0 \in D(\phi)$, one finds that $\phi^*(\xi) \geq -\phi(0)$ for all $\xi \in E^*$. If ϕ is nonnegative, then $\phi^*(0) = -\inf\phi \leq 0$.

For $q > 1$, we denote by $\Phi_q(E)$ the set of all lower semicontinuous convex functionals $\phi : E \to [0, \infty)$ satisfying the following two conditions:

- *q-coercivity of ϕ in E*: There exist some constants $C_1 > 0, C_2 \geq 0$ such that

$$C_1|u|_E^q \leq \phi(u) + C_2 \quad \text{for all} \quad u \in E. \tag{2.2}$$

  Particularly, $D(\phi) = E$.
- *q-boundedness of $\partial\phi$ in E*: There exists a constant $C_3 \geq 0$ such that

$$|\xi|_{E^*}^{q'} \leq C_3\,(|u|_E^q + 1) \quad \text{for all} \quad \xi \in \partial\phi(u) \tag{2.3}$$

  with $q' := q/(q-1)$.

We recall the following well-known facts (see [4, subsection 2.2] for a proof).

PROPOSITION 2.1.

(i) *If q-coercivity* (2.2) *holds, then*

$$c_0|u|_E^q \leq \langle \xi, u\rangle + C_4 \quad \textit{for all} \quad \xi \in \partial\phi(u) \tag{2.4}$$

with constants $c_0 > 0$, $C_4 \geq 0$.

(ii) *If q-boundedness* (2.3) *holds, then*

$$\phi(u) \leq C_5\,(|u|_E^q + 1) \quad \textit{for all} \quad u \in E \tag{2.5}$$

for some $C_5 \geq 0$.

A caveat on notation: Henceforth we shall use the symbol C in order to identify a generic constant depending on data. The reader shall be aware that the value of the constant C may vary from line to line.

PROPOSITION 2.2. *$\phi \in \Phi_q(E)$ iff $\phi^* \in \Phi_{q'}(E^*)$ with $q' = q/(q-1)$.*

Proof. We first prove that the coercivity (2.2) entails that $D(\phi^*) = E^*$. Indeed, for $\xi \in E^*$, we observe that

$$\begin{aligned}\phi^*(\xi) &:= \sup_{w\in E}\{\langle \xi, w\rangle - \phi(w)\}\\ &\overset{(2.2)}{\leq} \sup_{w\in E}\{|\xi|_{E^*}|w|_E - C_1|w|_E^q + C_2\} < \infty\end{aligned}$$

from the fact that $q > 1$. Thus $\phi^*(\xi) < \infty$. We check just the sufficiency part, as the necessity follows by duality. For each $u \in \partial\phi^*(\xi)$ (i.e., $\xi \in \partial\phi(u)$), by the q-coercivity (2.2) of ϕ we have

$$C_1|u|_E^q \leq \phi(u) + C_2 = \langle \xi, u\rangle - \phi^*(\xi) + C_2 \leq \frac{C_1}{2}|u|_E^q + C|\xi|_{E^*}^{q'} + C_2 + \phi(0).$$

Here we also used the fact that $\phi^*(\xi) \geq -\phi(0)$ for any $\xi \in E^*$. Thus,

$$|u|_E^q \leq C\big(|\xi|_{E^*}^{q'} + 1\big) \quad \text{for every} \quad u \in \partial\phi^*(\xi).$$

As for the coercivity of ϕ^*, let $\xi \in E^*$ be arbitrarily fixed. By the q-coercivity of ϕ, $\partial\phi$ is surjective from E into E^*, and hence one can take $u \in D(\partial\phi)$ such that $\xi \in \partial\phi(u)$. Then, we find that

$$\phi^*(\xi) = \sup_{w\in E}\{\langle \xi, w\rangle - \phi(w)\} \geq \langle \xi, \varepsilon u\rangle - \phi(\varepsilon u),$$

where $\varepsilon > 0$ is arbitrarily small. Hence, by (2.5) it follows that

$$\phi^*(\xi) \geq \langle \xi, \varepsilon u \rangle - \phi(\varepsilon u) \geq \varepsilon \langle \xi, u \rangle - \varepsilon^q C_5 |u|_E^q - C_5.$$

Then, we deduce from (2.4) that

$$\phi^*(\xi) \geq \varepsilon c_0 |u|_E^q - \varepsilon C_4 - \varepsilon^q C_5 |u|_E^q - C_5 \geq \frac{\varepsilon}{2} c_0 |u|_E^q - C$$

by taking a constant $\varepsilon > 0$ sufficiently small. Thus, by recalling the q-boundedness (2.3) of $\partial\phi$,

$$|\xi|_{E^*}^{q'} \leq C_3 \left(|u|_E^q + 1 \right),$$

we obtain

$$|\xi|_{E^*}^{q'} \leq C \left(\phi^*(\xi) + 1 \right) \quad \text{for all } \ \xi \in E^*.$$

Combining all these facts, we conclude that $\phi^* \in \Phi_{q'}(E^*)$. $\square$

2.2. Variational formulations. We shall reformulate our doubly nonlinear evolution problem in an abstract Banach-space frame. This reformulation serves the double aim of both generalizing the argument and, to some extent, simplifying notation.

Assume to be given the reflexive Banach spaces $X \subset V$ and let $\psi : V \to [0, \infty)$ and $\varphi : X \to [0, \infty)$ be lower semicontinuous convex functionals. Here we shall be concerned with the abstract version of relation (1.1)

$$\frac{\mathrm{d}v}{\mathrm{d}t} + \partial\varphi(u) \ni f, \quad v \in \partial\psi(u) \quad \text{a.e. in } \ (0, T), \tag{2.6}$$

equipped with the initial condition

$$v(0) = v_0. \tag{2.7}$$

In particular, (2.6) corresponds to a suitable variational reformulation of relation (1.1) along with the choices $X := W_0^{1,m}(\Omega)$, $V := L^p(\Omega)$, and

$$\varphi(u) := \int_\Omega \widehat{\beta}(\nabla u) \mathrm{d}x, \quad \psi(v) := \int_\Omega \widehat{\alpha}(v) \mathrm{d}x,$$

where $\widehat{\alpha}$ and $\widehat{\beta}$ are primitives of α and β, respectively.

We shall here advance a global variational approach for the abstract equation (2.6). As mentioned, this will require a suitable reformulation of (2.6) by *dualization*. In particular, by using property (2.1), relation (2.6) is equivalently transformed into

$$u(t) \in \partial\varphi^*(f(t) - v'(t)) \quad \text{and} \quad u(t) \in \partial\psi^*(v(t)),$$

where $v' = \mathrm{d}v/\mathrm{d}t$. Hence, by focusing on the unknown v we dualize relation (2.6) as

$$-\partial\varphi^*(f(t) - v'(t)) + \partial\psi^*(v(t)) \ni 0. \tag{2.8}$$

In case $f \equiv 0$ and $\varphi^* \in \Phi_{m'}(X^*)$ (equivalently, $\varphi \in \Phi_m(X)$ by Proposition 2.2) for $m' \geq 2$ (equivalently, $m \leq 2$), relation (2.8) falls within the abstract framework

considered in [4], where the corresponding WED approach is developed. Here we are, however, forced to address a problem more general than that of [4]. In particular, we shall allow a forcing term $f \not\equiv 0$, remove the restriction on the power, $m \leq 2$, and weaken the (strict) convexity requirement of either φ^* or ψ^*. Moreover, we add a forcing term $g = g(t)$ to (2.8) (see subsection 5). This slight generalization will allow us to include in our treatment the former case considered in [4], and we obtain here also a refinement of the former analysis. In summary, we shall here be interested in the abstract doubly nonlinear relation

$$-\partial\varphi^*(f - v') + \partial\psi^*(v) \ni g \quad \text{a.e. in } (0,T). \tag{2.9}$$

2.3. Main results. Let us start by listing our assumptions.

(A0) X is densely and compactly embedded in V and the norms $|\cdot|_X$ and $|\cdot|_{X^*}$ are strictly convex and uniformly convex, respectively.

(A1) There exists $m \in (1,\infty)$ such that $\varphi \in \Phi_m(X)$. Moreover, φ^* is Gâteaux differentiable in X^*.

(A2) There exists $p \in (1,\infty)$ such that $\psi \in \Phi_p(V)$.

(A3) $f \in L^{p'}(0,T;V^*) \cap L^{m'}(0,T;X^*)$, $g \in L^m(0,T;X)$, and $v_0 \in V^*$.

Note that the Gâteaux differentiability of φ^* in X^* is ensured, for instance, by assuming that $\varphi \in \Phi_m(X)$ is locally uniformly convex and Fréchet differentiable in X (see Theorem 2.3 of [12]). Moreover, as every reflexive Banach space can be equivalently renormed in such a way that it is strictly convex together with its dual [6, 7], the strict-convexity requirement in (A0) could be dropped.

We are now in the position of introducing the WED functionals W_ε: $L^{m'}(0,T;X^*) \to (-\infty,\infty]$ for relation (2.9) defined as

$$W_\varepsilon(v) = \begin{cases} \displaystyle\int_0^T \frac{e^{-t/\varepsilon}}{\varepsilon}\Big(\varepsilon\varphi^*(f(t) - v'(t)) + \psi^*(v(t)) - \langle v(t), g(t)\rangle\Big)\mathrm{d}t \\ \text{if } \ v \in W^{1,m'}(0,T;X^*) \cap L^{p'}(0,T;V^*) \ \text{ and } \ v(0) = v_0, \\ \infty \quad \text{otherwise.} \end{cases}$$

Then, by (A0)–(A3) the effective domain of W_ε reads

$$D(W_\varepsilon) = \Big\{ v \in W^{1,m'}(0,T;X^*) \cap L^{p'}(0,T;V^*) \colon v(0) = v_0 \Big\}.$$

One can easily check the convexity and the lower semicontinuity of W_ε in $L^{m'}(0,T;X^*)$. In particular, the WED functionals W_ε admit minimizers.

PROPOSITION 2.3 (well-posedness of the minimum problem). *Assume* (A0)–(A3). *The WED functionals W_ε admit minimizers over $L^{m'}(0,T;X^*)$. Moreover, every minimizer v_ε of W_ε is a strong solution* (*in the sense made precise by Definition* 3.1) *of the following system*:

$$\varepsilon\xi_\varepsilon' - \xi_\varepsilon + \eta_\varepsilon = g \quad \textit{in} \ \ (0,T), \tag{2.10}$$

$$\xi_\varepsilon = \mathrm{D}\varphi^*(f - v_\varepsilon'), \quad \eta_\varepsilon \in \partial\psi^*(v_\varepsilon) \quad \textit{in} \ \ (0,T), \tag{2.11}$$

$$v_\varepsilon(0) = v_0, \quad \xi_\varepsilon(T) = 0. \tag{2.12}$$

This proposition will be proved in section 4 by means of an approximation and Γ-convergence argument. Moreover, some energy inequalities will also be established, and they will be used to prove the main result below.

At each $\varepsilon > 0$, the system (2.10)–(2.12) corresponds to an *elliptic in time* regularization of (2.9). In particular, the system is noncausal. Indeed, causality is restored in the limit $\varepsilon \to 0$ so that we refer to this convergence as the *causal limit*.

The main result of the paper is the following.

THEOREM 2.4 (causal limit). *Assume* (A0)–(A3). *Let v_ε minimize W_ε over $L^{m'}(0,T;X^*)$. Then, there exists a sequence $\varepsilon_n \to 0$ such that*

$$v_{\varepsilon_n} \to v \quad \textit{strongly in } \ C([0,T];X^*) \ \textit{ and weakly in } \ W^{1,m'}(0,T;X^*) \cap L^{p'}(0,T;V^*),$$

where v solves (2.9) *and fulfills the initial condition* (2.7).

As mentioned above, the causal-limit convergence result is crucial, for it links the solution of the doubly nonlinear evolution equation (2.9) with the convex minimization of the WED functionals. The result will be proved by establishing suitable ε-independent estimates on the minimizers of W_ε and passing to the causal limit in the Euler–Lagrange equation.

3. Euler–Lagrange equation. This section is devoted to proving Proposition 2.3 and to deriving energy inequalities, which will then be used to prove Theorem 2.4. The Euler–Lagrange equation (2.10)–(2.12) may be justified by a formal computation of the derivative of W_ε. However, the WED functional W_ε is essentially nonsmooth due to a constraint associated with the initial condition, even though ψ^* and φ^* are sufficiently smooth. Hence the equivalence between (2.10)–(2.12) and a usual form of Euler–Lagrange equation, $\partial W_\varepsilon(v_\varepsilon) \ni 0$, is delicate.

Solutions for the system (2.10)–(2.12) are defined as follows.

DEFINITION 3.1 (strong solution). *A function $v : [0,T] \to X^*$ is said to be a* strong solution *of the system* (2.10)–(2.12) *if the following conditions are satisfied*:

$$v \in L^{p'}(0,T;V^*) \cap W^{1,m'}(0,T;X^*), \tag{3.1}$$

$$\xi := \mathrm{D}\varphi^*(f - v') \in L^m(0,T;X), \tag{3.2}$$

$$\xi' \in L^m(0,T;X) + L^p(0,T;V), \tag{3.3}$$

and there exists $\eta \in L^p(0,T;V)$ such that

$$\eta(t) \in \partial\psi^*(v(t)), \quad \varepsilon\xi'(t) - \xi(t) + \eta(t) = g(t) \ \ \textit{for a.a. } t \in (0,T), \tag{3.4}$$

$$v(0) = v_0 \ \ \textit{and} \ \ \xi(T) = 0. \tag{3.5}$$

The main result of this section reads as follows.

THEOREM 3.2 (minimizers solve the Euler–Lagrange equation). *Under assumptions* (A0)–(A3), *every minimizer v_ε of the WED functional W_ε is a strong solution of the Euler–Lagrange equation* (2.10)–(2.12) *such that the following energy inequalities hold with $\xi_\varepsilon = \mathrm{d}\varphi^*(f - v_\varepsilon')$ and $\eta_\varepsilon \in \partial\psi^*(v_\varepsilon)$ as in Definition* 3.1:

$$\begin{aligned}\int_0^T \psi^*(v_\varepsilon(t))\mathrm{d}t \le{}& \varepsilon T\varphi(0) + CT\left(\int_0^T |g(\tau)|_X^m \mathrm{d}\tau + T\right) + T\psi^*(v_0)\\ &+ \int_0^T\int_0^t \langle f(\tau), \eta_\varepsilon(\tau)\rangle\,\mathrm{d}\tau\mathrm{d}t + \varepsilon\int_0^T \langle f(t) - v_\varepsilon'(t), \xi_\varepsilon(t)\rangle\,\mathrm{d}t,\end{aligned} \tag{3.6}$$

$$\int_0^T \langle v_\varepsilon(t), \eta_\varepsilon(t)\rangle\mathrm{d}t = \langle v_0, \varepsilon\xi_\varepsilon(0)\rangle + \int_0^T \langle v_\varepsilon'(t), \varepsilon\xi_\varepsilon(t)\rangle\mathrm{d}t + \int_0^T \langle v_\varepsilon(t), \xi_\varepsilon(t) + g(t)\rangle\mathrm{d}t, \tag{3.7}$$

$$\int_0^T \langle f(t) - v_\varepsilon'(t), \xi_\varepsilon(t)\rangle \,\mathrm{d}t + \psi^*(v_\varepsilon(T)) \tag{3.8}$$
$$\leq \varepsilon\varphi(0) - \int_0^T \langle f(t) - v_\varepsilon'(t), g(t)\rangle \,\mathrm{d}t + \psi^*(v_0) + \int_0^T \langle f(t), \eta_\varepsilon(t)\rangle \,\mathrm{d}t$$

with some constant $C \geq 0$ independent of v_0, f, g, T, and ε.

Our strategy of proof is the following. We first introduce some auxiliary problems whose solutions approximate both solutions of the Euler–Lagrange equation (2.10)–(2.12) as well as minimizers of the WED functional W_ε. Next, we verify the convergence of such approximate solutions to a limit and check that this is a strong solution of (2.10)–(2.12) and a minimizer of W_ε. In case the minimizer of W_ε is unique (for instance, under suitable strict convexity assumptions), then Theorem 2.3 follows. In case the minimizer is not unique we shall proceed with some penalization technique. More precisely, let $\hat{v}_\varepsilon$ be a minimizer of W_ε over $L^{m'}(0,T;X^*)$ and introduce

$$\widehat{W}_\varepsilon(v) := W_\varepsilon(v) + \frac{c}{m'}\int_0^T \frac{e^{-t/\varepsilon}}{\varepsilon}|v(t) - \hat{v}_\varepsilon(t)|_{X^*}^{m'}\,\mathrm{d}t \quad \text{for } \ v \in L^{m'}(0,T;X^*)$$

with a constant $c \geq 0$. Then $\hat{v}_\varepsilon$ becomes the unique minimizer of $\widehat{W}_\varepsilon$ whenever $c > 0$ (see also section 3.4). Taking account of this penalization, we shall consider the functionals $\widehat{W}_\varepsilon$ instead of the original ones W_ε in the following subsections.

3.1. Approximation. We introduce the following approximated WED functionals $\widehat{W}_{\varepsilon,\lambda}$ for $\lambda > 0$ on the smaller domain $\mathcal{V} := L^\sigma(0,T;X^*)$ with $\sigma := \max\{2, m'\}$ (cf. [4], where $\sigma = m' \geq 2$ was chosen instead):

$$\widehat{W}_{\varepsilon,\lambda}(v) = \begin{cases} \displaystyle\int_0^T \frac{e^{-t/\varepsilon}}{\varepsilon}\Big(\varepsilon\varphi^*(f(t) - v'(t)) + (\overline{\psi^*})_\lambda(v(t)) - \langle v(t), g(t)\rangle\Big)\,\mathrm{d}t \\ \displaystyle\quad + \frac{c}{m'}\int_0^T \frac{e^{-t/\varepsilon}}{\varepsilon}|v(t) - \hat{v}_\varepsilon(t)|_{X^*}^{m'}\,\mathrm{d}t \\ \qquad\qquad\qquad\qquad \text{if } \ v \in W^{1,m'}(0,T;X^*) \ \text{ and } \ v(0) = v_0, \\ \infty \quad \text{otherwise,} \end{cases}$$

where $\overline{\psi^*}$ denotes the trivial extension to X^* of $\psi^* : V^* \to \mathbb{R}$ given by

$$\overline{\psi^*}(u) := \begin{cases} \psi^*(u) & \text{if } u \in V^*, \\ \infty & \text{if } u \in X^* \setminus V^*. \end{cases}$$

(Then $\overline{\psi^*} : X^* \to (-\infty, \infty]$ is lower semicontinuous and convex in X^* by $\psi^* \in \Phi_{p'}(V^*)$ and the continuous embedding $V^* \hookrightarrow X^*$.) Moreover, $(\overline{\psi^*})_\lambda$ denotes the Moreau–Yosida regularization of $\overline{\psi^*}$ in X^*, i.e.,

$$(\overline{\psi^*})_\lambda(u) = \inf_{w \in X^*}\left\{\frac{1}{2\lambda}|u - w|_{X^*}^2 + \overline{\psi^*}(w)\right\}, \tag{3.9}$$

which is Gâteaux differentiable in X^*. Note that the Gâteaux derivative $\mathrm{D}(\overline{\psi^*})_\lambda$ coincides with the Yosida approximation of the subdifferential operator $\partial(\overline{\psi^*}) : X^* \to X$. In particular, given a maximal monotone operator $A : X^* \to X$ and $\lambda > 0$, one can consider for each $u \in X^*$ the problem

$$F^*(v - u) + \lambda A v \ni 0, \tag{3.10}$$

where $F^* : X^* \to X$ is the standard duality mapping (which is strictly monotone as X^* is strictly convex). Equation (3.10) has the unique solution thanks to [8, Corollary 1.1, p. 39]. By letting $J_\lambda u$ be the solution v, one usually defines the *Yosida approximation* A_λ of A at level λ as $A_\lambda u := -F^*(J_\lambda u - u)/\lambda \in A(J_\lambda u)$ for all $u \in X^*$.

By using the direct method of the calculus of variations, one can obtain a global minimizer $v_{\varepsilon,\lambda}$ of $W_{\varepsilon,\lambda}$, i.e., $\partial_{\mathcal{V}} W_{\varepsilon,\lambda}(v_{\varepsilon,\lambda}) \ni 0$. Then, we have the following.

LEMMA 3.3. *Minimizers $v_{\varepsilon,\lambda}$ are strong solutions of*

$$\varepsilon \xi'_{\varepsilon,\lambda}(t) - \xi_{\varepsilon,\lambda}(t) + \eta_{\varepsilon,\lambda}(t) + \zeta_{\varepsilon,\lambda}(t) = g(t), \quad 0 < t < T, \tag{3.11}$$

$$\xi_{\varepsilon,\lambda}(t) := \mathrm{D}\varphi^*(f(t) - v'_{\varepsilon,\lambda}(t)), \quad \eta_{\varepsilon,\lambda}(t) := \mathrm{D}(\overline{\psi^*})_\lambda(v_{\varepsilon,\lambda}(t)), \tag{3.12}$$

$$\zeta_{\varepsilon,\lambda}(t) := c|v_{\varepsilon,\lambda}(t) - \hat{v}_\varepsilon(t)|_{X^*}^{m'-2} F_{X^*}(v_{\varepsilon,\lambda}(t) - \hat{v}_\varepsilon(t)), \tag{3.13}$$

$$v_{\varepsilon,\lambda}(0) = v_0, \quad \xi_{\varepsilon,\lambda}(T) = 0, \tag{3.14}$$

where $\mathrm{D}\varphi^*$ and $\mathrm{D}(\overline{\psi^*})_\lambda$ denote the Gâteaux derivatives of φ^* and $(\overline{\psi^*})_\lambda$, respectively, and $F_{X^*} : X^* \to X$ stands for a duality mapping between X^* and X. Moreover, $\xi'_{\varepsilon,\lambda}$ belongs to $L^m(0,T;X)$.

Proof. In order to derive a representation of $\partial_{\mathcal{V}} W_{\varepsilon,\lambda}$, we decompose $W_{\varepsilon,\lambda}$ into $W_{\varepsilon,\lambda} = I^1_\varepsilon + I^2_{\varepsilon,\lambda} + I^3_\varepsilon$.

First, define a functional $I^1_\varepsilon : \mathcal{V} \to (-\infty, \infty]$ by

$$I^1_\varepsilon(v) = \begin{cases} \displaystyle\int_0^T e^{-t/\varepsilon} \varphi^*(f(t) - v'(t))\, \mathrm{d}t \\ \qquad \text{if } \ v \in W^{1,m'}(0,T;X^*) \ \text{ and } \ v(0) = v_0, \\ \infty \quad \text{otherwise.} \end{cases}$$

As in [4], we can prove that

$$\partial_{\mathcal{V}} I^1_\varepsilon(v)(t) = \mathcal{A}(v) := \frac{\mathrm{d}}{\mathrm{d}t}\Big(e^{-t/\varepsilon} \mathrm{D}\varphi^*(f(t) - v'(t))\Big)$$

and

$$\begin{aligned} D(\partial_{\mathcal{V}} I^1_\varepsilon) = D(\mathcal{A}) := \big\{ v \in D(I^1_\varepsilon) \colon \mathrm{D}\varphi^*(f(\cdot) - v'(\cdot)) \in W^{1,\sigma'}(0,T;X) \\ \text{and } \ \mathrm{D}\varphi^*(f(T) - v'(T)) = 0 \big\}. \end{aligned}$$

Indeed, one can easily check $\mathcal{A} \subset \partial_{\mathcal{V}} I^1_\varepsilon$. Concerning the inverse inclusion, define a functional $J : \mathcal{W} := W^{1,m'}(0,T;X^*) \to \mathbb{R}$ by

$$J(v) := \int_0^T e^{-t/\varepsilon} \varphi^*(f(t) - v'(t)) \mathrm{d}t \quad \text{ for } \ v \in \mathcal{W}.$$

By using Lebesgue's dominated convergence theorem and the Gâteaux differentiability of φ^* in X^*, we deduce

$$\frac{J(v + hw) - J(v)}{h} \to \int_0^T e^{-t/\varepsilon} \left\langle -w'(t), \mathrm{D}\varphi^*(f(t) - v'(t)) \right\rangle_X \mathrm{d}t$$

as $h \to 0$ for each $w \in \mathcal{W}$. Since $\partial_{\mathcal{W}} I^1_\varepsilon = \mathrm{D}_{\mathcal{W}} J + \partial_{\mathcal{W}} I_K$ on $\mathcal{W}$, where I_K is the indicator function of the set, $K := \{v \in \mathcal{W} \colon v(0) = v_0\}$ (see [14]), we have, for each $\xi \in \partial_{\mathcal{W}} I^1_\varepsilon(v)$,

$$\langle \xi, w \rangle_{\mathcal{W}} = \int_0^T e^{-t/\varepsilon} \left\langle -w'(t), \mathrm{D}\varphi^*(f(t) - v'(t)) \right\rangle_X \mathrm{d}t$$

for any $w \in \mathcal{W}$ satisfying $w(0) = 0$. Furthermore, noting that $\partial_{\mathcal{V}} I^1_\varepsilon \subset \partial_{\mathcal{W}} I^1_\varepsilon$ by the continuous embedding $\mathcal{W} \hookrightarrow \mathcal{V}$ and $D(I^1_\varepsilon) \subset \mathcal{W}$, for any $[v, \xi] \in \partial_{\mathcal{V}} I^1_\varepsilon$, we see that

$$\int_0^T \langle w(t), \xi(t)\rangle_X \,\mathrm{d}t = \int_0^T e^{-t/\varepsilon} \left\langle -w'(t), \mathrm{D}\varphi^*(f(t) - v'(t))\right\rangle_X \mathrm{d}t \tag{3.15}$$

for any $w \in \mathcal{W}$ satisfying $w(0) = 0$; thus we obtain $t \mapsto e^{-t/\varepsilon}\mathrm{D}\varphi^*(f(t) - v'(t)) \in W^{1,\sigma'}(0, T; X)$ by $\xi \in L^{\sigma'}(0, T; X)$, and

$$\xi = \frac{\mathrm{d}}{\mathrm{d}t}\big(e^{-t/\varepsilon}\mathrm{D}\varphi^*(f(t) - v'(t))\big)$$

by integration by parts. Here, the final condition $\mathrm{D}\varphi^*(f(T) - v'(T)) = 0$ also follows from the arbitrariness of $w(T)$ in (3.15). Therefore $\partial_{\mathcal{V}} I^1_\varepsilon$ coincides with $\mathcal{A}$.

Second, define $I^2_{\varepsilon,\lambda} : \mathcal{V} \to (-\infty, \infty]$ by

$$I^2_{\varepsilon,\lambda}(v) := \int_0^T \frac{e^{-t/\varepsilon}}{\varepsilon}\big((\overline{\psi^*})_\lambda(v(t)) - \langle v(t), g(t)\rangle\big)\,\mathrm{d}t \quad \text{for} \quad v \in \mathcal{V}.$$

Then $D(I^2_{\varepsilon,\lambda}) = \mathcal{V}$, since for each $\lambda > 0$ there exists $C_\lambda \geq 0$ such that $(\overline{\psi^*})_\lambda(v) \leq C_\lambda(|v|^2_{X^*} + 1)$ for any $v \in X^*$ by definition (3.9) and $v \in \mathcal{V} \subset L^2(0, T; X^*)$. Moreover, for $v \in \mathcal{V}$, $\eta \in \mathcal{V}^*$, it holds that

$$\eta \in \partial_{\mathcal{V}} I^2_{\varepsilon,\lambda}(v) \quad \text{iff} \quad \eta(t) = \frac{e^{-t/\varepsilon}}{\varepsilon}\big(\mathrm{D}(\overline{\psi^*})_\lambda(v(t)) - g(t)\big) \quad \text{for a.e. } t \in (0, T).$$

Eventually, set the functional I^3_ε on $\mathcal{V}$ to be

$$I^3_\varepsilon(v) := \frac{c}{m'}\int_0^T \frac{e^{-t/\varepsilon}}{\varepsilon}|v(t) - \hat{v}_\varepsilon(t)|^{m'}_{X^*}\mathrm{d}t \quad \text{for} \quad v \in \mathcal{V}.$$

Then I^3_ε is Gâteaux differentiable in $L^{m'}(0, T; X^*)$ (hence in $\mathcal{V}$), and its derivative has the following representation:

$$\mathrm{D}_{\mathcal{V}} I^3_\varepsilon(v)(t) = \frac{ce^{-t/\varepsilon}}{\varepsilon}|v(t) - \hat{v}_\varepsilon(t)|^{m'-2}_{X^*} F_{X^*}(v(t) - \hat{v}_\varepsilon(t));$$

see Remark A.2 in the appendix.

Hence since both $D(I^2_{\varepsilon,\lambda})$ and $D(I^3_\varepsilon)$ coincide with $\mathcal{V}$, it follows by Theorem 2.2 of [14] that

$$\begin{aligned}
\partial_{\mathcal{V}} W_{\varepsilon,\lambda}(v)(t) &= \partial_{\mathcal{V}} I^1_\varepsilon(v)(t) + \partial_{\mathcal{V}} I^2_{\varepsilon,\lambda}(v)(t) + \mathrm{D}_{\mathcal{V}} I^3_\varepsilon(v)(t) \\
&= \frac{\mathrm{d}}{\mathrm{d}t}\big(e^{-t/\varepsilon}\mathrm{D}\varphi^*(f(t) - v'(t))\big) + \frac{e^{-t/\varepsilon}}{\varepsilon}\big(\mathrm{D}(\overline{\psi^*})_\lambda(v(t)) - g(t)\big) \\
&\quad + \frac{ce^{-t/\varepsilon}}{\varepsilon}|v(t) - \hat{v}_\varepsilon(t)|^{m'-2}_{X^*} F_{X^*}(v(t) - \hat{v}_\varepsilon(t)) \\
&= \frac{e^{-t/\varepsilon}}{\varepsilon}\bigg(\varepsilon\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{D}\varphi^*(f(t) - v'(t)) - \mathrm{D}\varphi^*(f(t) - v'(t)) \\
&\quad + \mathrm{D}(\overline{\psi^*})_\lambda(v(t)) - g(t) + c|v(t) - \hat{v}_\varepsilon(t)|^{m'-2}_{X^*} F_{X^*}(v(t) - \hat{v}_\varepsilon(t))\bigg)
\end{aligned}$$

for any $v \in D(\partial_{\mathcal{V}} W_{\varepsilon,\lambda}) = D(\partial_{\mathcal{V}} I^1_\varepsilon) \cap D(\partial_{\mathcal{V}} I^2_{\varepsilon,\lambda})$. Moreover, each term in the last line above belongs to $\mathcal{V}^* = L^{\sigma'}(0,T;X)$ with $\sigma' = \sigma/(\sigma-1) \le m$. Therefore, solutions $v_{\varepsilon,\lambda}$ of the Euler–Lagrange equation, $\partial_{\mathcal{V}} W_{\varepsilon,\lambda}(v_{\varepsilon,\lambda}) \ni 0$, satisfy the system (3.11)–(3.14).

Finally, since $\varphi^* \in \Phi_{m'}(X^*)$ and $v_{\varepsilon,\lambda} \in D(\partial_{\mathcal{V}} W_{\varepsilon,\lambda}) \subset D(I^1_\varepsilon) \subset W^{1,m'}(0,T;X^*)$, we obtain $\xi_{\varepsilon,\lambda} := \mathrm{D}\varphi^*(f - v'_{\varepsilon,\lambda}) \in L^m(0,T;X)$. We see that $\zeta_{\varepsilon,\lambda} := c|v_{\varepsilon,\lambda} - \hat{v}_\varepsilon|^{m'-2}_{X^*} F_{X^*}(v_{\varepsilon,\lambda} - \hat{v}_\varepsilon) \in L^m(0,T;X)$ since

$$|\zeta_{\varepsilon,\lambda}(t)|^m_X = c^m |v_{\varepsilon,\lambda}(t) - \hat{v}_\varepsilon(t)|^{m'}_{X^*}. \tag{3.16}$$

Moreover, since $\mathrm{D}(\overline{\psi^*})_\lambda$ is bounded from X^* to X, we deduce that $\eta_{\varepsilon,\lambda} := \mathrm{D}(\overline{\psi^*})_\lambda(v_{\varepsilon,\lambda})$ belongs to $L^\infty(0,T;X)$. By comparison we then have $\varepsilon \xi'_{\varepsilon,\lambda} \in L^m(0,T;X)$. ∎

3.2. A priori estimates. For the aim of checking the causal limit $\varepsilon \to 0$ we shall now turn to proving a priori estimates independent of ε. For notational simplicity, we shall systematically omit the subscript ε in the remainder of this section. Moreover, we use Gâteaux differentiable functionals $P : X^* \to \mathbb{R}$ and $\mathcal{P} : L^{m'}(0,T;X^*) \to \mathbb{R}$ given by

$$P(v) := \frac{c}{m'}|v|^{m'}_{X^*} \quad \text{for } v \in X^*, \quad \mathcal{P}(v) := \int_0^T P(v(t) - \hat{v}(t))\mathrm{d}t \quad \text{for } v \in L^{m'}(0,T;X^*).$$

Here we further notice that $\zeta_\lambda = \mathrm{D}\mathcal{P}(v_\lambda)$ and $\zeta_\lambda(t) = \mathrm{D}P(v_\lambda(t) - \hat{v}(t))$ (see Remark A.2 in the appendix).

Test (3.11) by $f(t) - v'_\lambda(t)$ and integrate in time over $(0,T)$ to get

$$\begin{aligned}
\varepsilon \int_0^T \langle f(t) - v'_\lambda(t), \xi'_\lambda(t)\rangle \,\mathrm{d}t &- \int_0^T \langle f(t) - v'_\lambda(t), \xi_\lambda(t)\rangle \,\mathrm{d}t \\
&+ \int_0^T \langle f(t), \eta_\lambda(t)\rangle \,\mathrm{d}t - (\overline{\psi^*})_\lambda(v_\lambda(T)) + (\overline{\psi^*})_\lambda(v_0) \\
&+ \int_0^T \langle f(t) - \hat{v}'(t), \zeta_\lambda(t)\rangle \,\mathrm{d}t - P(v_\lambda(T) - \hat{v}(T)) \\
&= \int_0^T \langle f(t) - v'_\lambda(t), g(t)\rangle \,\mathrm{d}t.
\end{aligned}$$

Here we used the fact that $P(v_\lambda(0) - \hat{v}(0)) = 0$ by the initial constraint $v_\lambda(0) = \hat{v}(0) = v_0$. Then by the definition of convex conjugate, since $\partial\varphi(\xi_\lambda) \ni f - v'_\lambda \in L^{m'}(0,T;X^*)$ and $\xi_\lambda \in W^{1,m}(0,T;X)$, we have

$$\begin{aligned}
\int_0^T \langle f(t) - v'_\lambda(t), \xi'_\lambda(t)\rangle \,\mathrm{d}t &= \varphi(\xi_\lambda(T)) - \varphi(\xi_\lambda(0)) \\
&\le \langle f(T) - v'_\lambda(T), \xi_\lambda(T)\rangle - \varphi^*(\xi_\lambda(T)) \le \varphi(0).
\end{aligned}$$

Here, we have also used the final condition $\xi_\lambda(T) = 0$ and the fact that φ is nonnegative (see also section 2.1). Hence, we obtain

$$
\begin{aligned}
(3.17)\qquad & \int_0^T \langle f(t) - v_\lambda'(t), \xi_\lambda(t)\rangle \,\mathrm{d}t + (\overline{\psi^*})_\lambda(v_\lambda(T)) + P(v_\lambda(T) - \hat{v}(T)) \\
&\le \varepsilon\varphi(0) - \int_0^T \langle f(t) - v_\lambda'(t), g(t)\rangle \,\mathrm{d}t + \psi^*(v_0) + \int_0^T \langle f(t), \eta_\lambda(t)\rangle \,\mathrm{d}t \\
&\quad + \int_0^T \langle f(t) - \hat{v}'(t), \zeta_\lambda(t)\rangle \,\mathrm{d}t,
\end{aligned}
$$

where we have also exploited the fact that $(\overline{\psi^*})_\lambda \le \overline{\psi^*} = \psi^*$ on V^*. Therefore, by using assumption (2.4) for $\varphi^* \in \Phi_{m'}(X^*)$, we get

$$
\begin{aligned}
(3.18)\qquad & \frac{c_0}{2}\int_0^T |f(t) - v_\lambda'(t)|_{X^*}^{m'} \mathrm{d}t + (\overline{\psi^*})_\lambda(v_\lambda(T)) + P(v_\lambda(T) - \hat{v}(T)) \\
&\le \varepsilon\varphi(0) + C\left(\int_0^T |g(t)|_X^m \mathrm{d}t + T\right) + \psi^*(v_0) + \int_0^T \langle f(t), \eta_\lambda(t)\rangle \,\mathrm{d}t \\
&\quad + \int_0^T \langle f(t) - \hat{v}'(t), \zeta_\lambda(t)\rangle \,\mathrm{d}t.
\end{aligned}
$$

Note that the above right-hand side contains terms involving η_λ and ζ_λ which arise due to the presence of the external force f (cf. [4] for the case $f \equiv 0$ and $g \equiv 0$) as well as due to the penalization.

Now, let us test (3.11) by $f(t) - v_\lambda'(t)$ and integrate just over $(0, t)$ instead of $(0, T)$. Then, an additional term appears (cf. [4]) as

$$
\begin{aligned}
& \frac{c_0}{2}\int_0^t |f(\tau) - v_\lambda'(\tau)|_{X^*}^{m'} \mathrm{d}\tau + (\overline{\psi^*})_\lambda(v_\lambda(t)) + P(v_\lambda(t) - \hat{v}(t)) \\
&\le \varepsilon\varphi(0) + C\left(\int_0^T |g(\tau)|_X^m \mathrm{d}\tau + T\right) + \psi^*(v_0) + \int_0^t \langle f(\tau), \eta_\lambda(\tau)\rangle \,\mathrm{d}\tau \\
&\quad + \varepsilon\langle f(t) - v_\lambda'(t), \xi_\lambda(t)\rangle + \int_0^t \langle f(\tau) - \hat{v}'(\tau), \zeta_\lambda(\tau)\rangle \,\mathrm{d}\tau.
\end{aligned}
$$

Integrating both sides over $(0, T)$ again, we obtain

(3.19)

$$
\begin{aligned}
& \int_0^T (\overline{\psi^*})_\lambda(v_\lambda(t))\mathrm{d}t + \int_0^T P(v_\lambda(t) - \hat{v}(t))\mathrm{d}t \\
&\le \varepsilon T\varphi(0) + CT\left(\int_0^T |g(\tau)|_X^m \mathrm{d}\tau + T\right) + T\psi^*(v_0) + \int_0^T\int_0^t \langle f(\tau), \eta_\lambda(\tau)\rangle \,\mathrm{d}\tau\mathrm{d}t \\
&\quad + \varepsilon\int_0^T \langle f(t) - v_\lambda'(t), \xi_\lambda(t)\rangle \mathrm{d}t + \int_0^T\int_0^t \langle f(\tau) - \hat{v}'(\tau), \zeta_\lambda(\tau)\rangle \,\mathrm{d}\tau\mathrm{d}t.
\end{aligned}
$$

Here we note by (3.16) that

$$
\begin{aligned}
& \int_0^T\int_0^t \langle f(\tau) - \hat{v}'(\tau), \zeta_\lambda(\tau)\rangle \,\mathrm{d}\tau\mathrm{d}t \\
&\le \frac{1}{2}\int_0^T P(v_\lambda(\tau) - \hat{v}(\tau))\mathrm{d}\tau + CT\int_0^T |f(\tau) - \hat{v}'(\tau)|_{X^*}^{m'} \mathrm{d}\tau.
\end{aligned}
$$

Employing the m'-boundedness of $\mathrm{D}\varphi^*$ and recalling estimate (3.18), we deduce

$$\begin{aligned}
&\int_0^T (\overline{\psi^*})_\lambda(v_\lambda(t))\mathrm{d}t + \frac{1}{2}\int_0^T P(v_\lambda(t)-\hat{v}(t))\mathrm{d}t \\
&\quad \le \varepsilon T\varphi(0) + CT\left(\int_0^T |g(\tau)|_X^m \mathrm{d}\tau + \int_0^T |f(\tau)-\hat{v}'(\tau)|_{X^*}^{m'}\mathrm{d}\tau + T\right) \\
&\qquad + T\psi^*(v_0) + \int_0^T\int_0^t \langle f(\tau), \eta_\lambda(\tau)\rangle \,\mathrm{d}\tau\mathrm{d}t \\
&\qquad + \varepsilon C\Bigg(\varepsilon\varphi(0) + \int_0^T |g(t)|_X^m \mathrm{d}t + T + \psi^*(v_0) + \int_0^T \langle f(t), \eta_\lambda(t)\rangle \,\mathrm{d}t \\
&\qquad + \int_0^T \langle f(t)-\hat{v}'(t), \zeta_\lambda(t)\rangle \,\mathrm{d}t + \psi(0)\Bigg).
\end{aligned}$$

Then, we use $\psi^* \in \Phi_{p'}(V^*)$ in order to conclude that

$$\int_0^T |\eta_\lambda(t)|_V^p \mathrm{d}t \le C\left(\int_0^T (\overline{\psi^*})_\lambda(v_\lambda(t))\mathrm{d}t + 1\right). \tag{3.20}$$

Indeed, denote by $J_\lambda : X^* \to D(\partial(\overline{\psi^*}))$ the *resolvent* of the subdifferential operator $\partial(\overline{\psi^*}) : X^* \to X$ and note that $J_\lambda v_\lambda(t) \in D(\partial(\overline{\psi^*})) \subset D(\overline{\psi^*}) = D(\psi^*) = V^*$ and $\eta_\lambda(t) = \mathrm{D}(\overline{\psi^*})_\lambda(v_\lambda(t)) \in \partial(\overline{\psi^*})(J_\lambda v_\lambda(t)) \subset \partial\psi^*(J_\lambda v_\lambda(t))$. It then follows that

$$\begin{aligned}
|\eta_\lambda(t)|_V^p &\le C\left(|J_\lambda v_\lambda(t)|_{V^*}^{p'} + 1\right) \le C\left(\psi^*(J_\lambda v_\lambda(t)) + 1\right) \\
&= C\left(\overline{\psi^*}(J_\lambda v_\lambda(t)) + 1\right) \le C\left((\overline{\psi^*})_\lambda(v_\lambda(t)) + 1\right).
\end{aligned}$$

Hence, by applying the Cauchy–Schwartz and Young inequalities together with (3.16) as well, we obtain

$$\begin{aligned}
&\int_0^T (\overline{\psi^*})_\lambda(v_\lambda(t))\mathrm{d}t + \int_0^T P(v_\lambda(t)-\hat{v}(t))\mathrm{d}t \\
&\quad \le C\Bigg(\varphi(0) + \psi(0) + \int_0^T |g(\tau)|_X^m \mathrm{d}\tau + \int_0^T |f(\tau)-\hat{v}'(\tau)|_{X^*}^{m'}\mathrm{d}\tau + \psi^*(v_0) \\
&\qquad + \int_0^T |f(\tau)|_{V^*}^{p'}\mathrm{d}\tau + 1\Bigg),
\end{aligned} \tag{3.21}$$

which, together with estimates (3.16) and (3.20), gives

$$\int_0^T |\eta_\lambda(t)|_V^p \mathrm{d}t \le C, \quad \int_0^T |\zeta_\lambda(t)|_X^m \mathrm{d}t \le C.$$

Furthermore, going back to (3.18), we deduce

$$\int_0^T |v_\lambda'(t)|_{X^*}^{m'}\mathrm{d}t \le C,$$

which also implies uniform estimates for v_λ and $J_\lambda v_\lambda$ in $C([0,T];X^*)$ by $|J_\lambda v_\lambda(t)|_{X^*} \le C(|v_\lambda(t)|_{X^*}+1)$. From the m'-boundedness of $\mathrm{D}\varphi^*$ we deduce that

$$\int_0^T |\xi_\lambda(t)|_X^m \mathrm{d}t \le C$$

and, by comparison,

$$\|\xi_\lambda'\|_{L^m(0,T;X)+L^p(0,T;V)} \le C.$$

Moreover, by standard properties of the Moreau–Yosida regularization and the p'-coercivity of ψ^*, one can derive from (3.21) that

$$\int_0^T |J_\lambda v_\lambda(t)|_{V^*}^{p'} \mathrm{d}t \le C.$$

3.3. Convergence as $\lambda \to 0$. The limit $\lambda \to 0$ can be ascertained by arguing as in [4] with appropriate modifications in places due to the presence of f, g, and ζ_λ.

From the a priori estimates of subsection 3.2, one can obtain the following limits by taking a sequence $\lambda_n \to 0$ (but we simply write λ instead of λ_n):

$$v_\lambda \to v \quad \text{weakly in } W^{1,m'}(0,T;X^*), \tag{3.22}$$

$$J_\lambda v_\lambda \to v \quad \text{weakly in } L^{p'}(0,T;V^*), \tag{3.23}$$

$$\eta_\lambda \to \eta \quad \text{weakly in } L^p(0,T;V), \tag{3.24}$$

$$\zeta_\lambda \to \zeta \quad \text{weakly in } L^m(0,T;X), \tag{3.25}$$

$$\xi_\lambda \to \xi \quad \text{weakly in } L^m(0,T;X), \tag{3.26}$$

$$\xi_\lambda' \to \xi' \quad \text{weakly in } L^m(0,T;X)+L^p(0,T;V). \tag{3.27}$$

Thus we have $\varepsilon\xi' - \xi + \eta + \zeta = g$. Moreover, as in subsection 3.3 of [4], we deduce from the compact embedding $X \hookrightarrow V$ (equivalently, $V^* \hookrightarrow X^*$) that

$$J_\lambda v_\lambda \to v \quad \text{strongly in } C([0,T];X^*), \tag{3.28}$$

$$v_\lambda \to v \quad \text{strongly in } L^q(0,T;X^*), \tag{3.29}$$

$$\xi_\lambda \to \xi \quad \text{strongly in } C([0,T];V) \tag{3.30}$$

for any $q \in [1,\infty)$, and hence, $v(0) = v_0$ (by $J_\lambda v_0 \to v_0$) and $\xi(T) = 0$. Since $\zeta_\lambda = \mathrm{D}\mathcal{P}(v_\lambda)$, by virtue of the maximal monotonicity of $\mathrm{D}\mathcal{P}$ in $L^{m'}(0,T;X^*)\times L^m(0,T;X)$, we deduce that $\zeta = \mathrm{D}\mathcal{P}(v)$, i.e., $\zeta(t) = \mathrm{D}P(v(t)-\hat{v}(t))$ for a.e. $t \in (0,T)$.

In order to prove $\eta(t) \in \partial\psi^*(v(t))$ and $\xi(t) = \mathrm{D}\varphi^*(f(t)-v'(t))$, we need an additional lower semicontinuity argument. Let $0 \le t_1 < t_2 \le T$ and note that

$$\int_{t_1}^{t_2} \langle v_\lambda'(t), \xi_\lambda(t)\rangle \mathrm{d}t = -\int_{t_1}^{t_2} \langle f(t)-v_\lambda'(t), \xi_\lambda(t)\rangle \mathrm{d}t + \int_{t_1}^{t_2} \langle f(t), \xi_\lambda(t)\rangle \mathrm{d}t.$$

Then, since $\xi_\lambda(t) = \mathrm{D}\varphi^*(f(t)-v_\lambda'(t))$, it follows from (3.22) and (3.26) that

$$\begin{aligned} &\limsup_{\lambda\to 0} \int_{t_1}^{t_2} \langle v_\lambda'(t), \xi_\lambda(t)\rangle \mathrm{d}t \\ &\quad = -\liminf_{\lambda\to 0} \int_{t_1}^{t_2} \langle f(t)-v_\lambda'(t), \xi_\lambda(t)\rangle \mathrm{d}t + \lim_{\lambda\to 0} \int_{t_1}^{t_2} \langle f(t), \xi_\lambda(t)\rangle \mathrm{d}t \\ &\quad \le -\int_{t_1}^{t_2} \langle f(t)-v'(t), \xi(t)\rangle \mathrm{d}t + \int_{t_1}^{t_2} \langle f(t), \xi(t)\rangle \mathrm{d}t = \int_{t_1}^{t_2} \langle v'(t), \xi(t)\rangle \mathrm{d}t. \end{aligned} \tag{3.31}$$

By following closely the argument of [4, p. 2556] let us take

$$\begin{aligned}\mathcal{L} := \big\{ t \in (0,T) \colon t \text{ is a Lebesgue point of the function } t \mapsto \langle v(t), \xi(t)\rangle \\ \text{such that for any sequence } \lambda_n \to 0, \text{ there exists a subsequence} \\ \lambda_{n'} \to 0 \ \text{ with } \ \langle v_{\lambda_{n'}}(t), \xi_{\lambda_{n'}}(t)\rangle \to \langle v(t), \xi(t)\rangle \big\},\end{aligned}$$

which has full Lebesgue measure. Recall that $\eta_\lambda(t) \in \partial\psi^*(J_\lambda v_\lambda(t))$ for almost every $t \in (0,T)$. For each $t_1, t_2 \in \mathcal{L}$ with $t_1 \le t_2$, by virtue of (3.11) and the lim sup inequality (3.31), we have

(3.32)
$$\begin{aligned}&\limsup_{\lambda\to 0} \int_{t_1}^{t_2} \langle J_\lambda v_\lambda(t), \eta_\lambda(t)\rangle \,\mathrm{d}t \\ &\quad \le -\lim_{\lambda\to 0} \left(\langle v_\lambda(t_2), \varepsilon\xi_\lambda(t_2)\rangle - \langle v_\lambda(t_1), \varepsilon\xi_\lambda(t_1)\rangle\right) + \limsup_{\lambda\to 0} \int_{t_1}^{t_2} \langle v_\lambda'(t), \varepsilon\xi_\lambda(t)\rangle \mathrm{d}t \\ &\qquad + \lim_{\lambda\to 0} \int_{t_1}^{t_2} \langle v_\lambda(t), \xi_\lambda(t) - \zeta_\lambda(t) + g(t)\rangle \mathrm{d}t \\ &\quad \le \int_{t_1}^{t_2} \langle v(t), \eta(t)\rangle \mathrm{d}t,\end{aligned}$$

where we have also used the fact that $\langle J_\lambda v_\lambda(t), \eta_\lambda(t)\rangle = \langle v_\lambda(t), \eta_\lambda(t)\rangle - \lambda|\eta_\lambda(t)|_X^2 \le \langle v_\lambda(t), \eta_\lambda(t)\rangle$ and exploited the integration by parts formula from [4, Proposition 2.3]. Hence, by using standard properties of maximal monotone operators (see Lemma 1.2 of [14] and Proposition 1.1 of [28]) and the arbitrariness of $t_1, t_2 \in \mathcal{L}$, we deduce from (3.23)–(3.24) that $\eta(t) \in \partial\psi^*(v(t))$ for almost every $t \in (0,T)$ and

$$\lim_{\lambda\to 0} \int_{t_1}^{t_2} \langle J_\lambda v_\lambda(t), \eta_\lambda(t)\rangle \mathrm{d}t = \int_{t_1}^{t_2} \langle v(t), \eta(t)\rangle \mathrm{d}t. \tag{3.33}$$

As for the limit of $\xi_\lambda = \mathrm{D}\varphi^*(f(\cdot) - v_\lambda'(\cdot))$ we check

$$\begin{aligned}\int_{t_1}^{t_2} \langle f(t) - v_\lambda'(t), \varepsilon\xi_\lambda(t)\rangle \mathrm{d}t \le &\int_{t_1}^{t_2} \langle f(t), \varepsilon\xi_\lambda(t)\rangle \mathrm{d}t - \langle v_\lambda(t_2), \varepsilon\xi_\lambda(t_2)\rangle + \langle v_\lambda(t_1), \varepsilon\xi_\lambda(t_1)\rangle \\ &+ \int_{t_1}^{t_2} \langle v_\lambda(t), \xi_\lambda(t)\rangle \mathrm{d}t - \int_{t_1}^{t_2} \langle J_\lambda v_\lambda(t), \eta_\lambda(t)\rangle \mathrm{d}t \\ &- \int_{t_1}^{t_2} \langle v_\lambda(t), \zeta_\lambda(t)\rangle \mathrm{d}t + \int_{t_1}^{t_2} \langle v_\lambda(t), g(t)\rangle \mathrm{d}t \\ &\to \int_{t_1}^{t_2} \langle f(t) - v'(t), \varepsilon\xi(t)\rangle \mathrm{d}t \quad \text{ as } \ \lambda \to 0.\end{aligned}$$

Hence, since $\mathrm{D}\varphi^*$ is maximal monotone in $X^* \times X$, we obtain $\xi(t) = \mathrm{D}\varphi^*(f(t) - v'(t))$ for almost every $t \in (0,T)$ and

$$\lim_{\lambda\to 0} \int_{t_1}^{t_2} \langle f(t) - v_\lambda'(t), \xi_\lambda(t)\rangle \mathrm{d}t = \int_{t_1}^{t_2} \langle f(t) - v'(t), \xi(t)\rangle \mathrm{d}t. \tag{3.34}$$

Now, we have proved that the limit v of v_λ is a strong solution of

$$\varepsilon\xi'(t) - \xi(t) + \eta(t) + \zeta(t) = g(t), \quad 0 < t < T, \tag{3.35}$$

$$\xi(t) = \mathrm{D}\varphi^*(f(t) - v'(t)), \quad \eta(t) \in \partial\psi^*(v(t)), \tag{3.36}$$

$$\zeta(t) = c|v(t) - \hat{v}(t)|_{X^*}^{m'-2} F_{X^*}(v(t) - \hat{v}(t)), \tag{3.37}$$

$$v(0) = v_0, \quad \xi(T) = 0. \tag{3.38}$$

3.4. Well-posedness of the minimum problem. Now we are ready to prove Theorem 3.2. Let $\hat{v}_\varepsilon$ be a (prescribed) minimizer of W_ε over $L^{m'}(0,T;X^*)$. Then one can verify that $\hat{v}_\varepsilon$ is the unique minimizer of the penalized functional $\widehat{W}_\varepsilon$ over $L^{m'}(0,T;X^*)$ from the following relation:

$$\widehat{W}_\varepsilon(\hat{v}_\varepsilon) = W_\varepsilon(\hat{v}_\varepsilon) \le W_\varepsilon(v) < \widehat{W}_\varepsilon(v) \quad \text{for all } \ v \in L^{m'}(0,T;X^*) \setminus \{\hat{v}_\varepsilon\}.$$

Recall that the solutions $v_{\varepsilon,\lambda}$ of (3.11)–(3.14) minimize the approximated functionals $\widehat{W}_{\varepsilon,\lambda}$ over $\mathcal{V} = L^\sigma(0,T;X^*)$ with $\sigma = \max\{2, m'\}$. Hence, by letting $\lambda \to 0$, one can exploit the Γ-convergence $\widehat{W}_{\varepsilon,\lambda} \to \widehat{W}_\varepsilon$ along with (3.22) and (3.28) to obtain

$$\widehat{W}_\varepsilon(v_\varepsilon) \le \liminf_{\lambda\to 0} \widehat{W}_{\varepsilon,\lambda}(v_{\varepsilon,\lambda}) \le \lim_{\lambda\to 0} \widehat{W}_{\varepsilon,\lambda}(w) = \widehat{W}_\varepsilon(w) \quad \text{for all } \ w \in \mathcal{V}.$$

(See [4, Proof of Lemma 4.3] for more details.) From the fact that $D(\widehat{W}_\varepsilon) \subset W^{1,m'}(0,T;X^*) \subset \mathcal{V}$, the limit v_ε minimizes $\widehat{W}_\varepsilon$ over $L^{m'}(0,T;X^*)$ as well. Since the minimizer of $\widehat{W}_\varepsilon$ is unique, v_ε must coincide with $\hat{v}_\varepsilon$, a prescribed minimizer of W_ε over $L^{m'}(0,T;X^*)$. Therefore from the above arguments, $\hat{v}_\varepsilon$ $(= v_\varepsilon)$ turns out to be a strong solution of (3.35)–(3.38), and moreover, $\zeta_\varepsilon(t) = \mathrm{D}P(v_\varepsilon(t) - \hat{v}_\varepsilon(t)) \equiv 0$. Thus $\hat{v}_\varepsilon$ becomes a strong solution of (2.10)–(2.12).

3.5. Energy inequalities. It still remains to derive energy inequalities (3.6)–(3.8). To this end, let us recall approximate problems (3.11)–(3.14) again. By repeating the computation leading to estimate (3.32) with $t_1 = 0$ and $t_2 = T$ and using $\xi_\lambda(T) = 0$ and $\zeta = \lim_{\lambda\to 0} \zeta_\lambda = 0$, one can derive the energy inequality:

$$\begin{aligned}\int_0^T \langle v(t), \eta(t)\rangle \mathrm{d}t &\le \langle v_0, \varepsilon\xi(0)\rangle + \int_0^T \langle v'(t), \varepsilon\xi(t)\rangle \mathrm{d}t + \int_0^T \langle v(t), \xi(t) + g(t)\rangle \mathrm{d}t \\ &\le \int_0^T \langle v(t), \eta(t)\rangle \mathrm{d}t,\end{aligned}$$

which implies (3.7). Furthermore, by combining estimates (3.17) and (3.19) with the above convergences (particularly, (3.34)) and by using $\zeta = 0$, we obtain estimates (3.8) and (3.6), respectively. Here, we also used the fact that

$$\int_0^T \int_0^t \langle f(\tau), \eta_\lambda(\tau) - \eta(\tau)\rangle \mathrm{d}\tau \mathrm{d}t \to 0. \tag{3.39}$$

Indeed, since $\eta_\lambda \to \eta$ weakly in $L^p(0,T;V)$, it follows that

$$\int_0^t \langle f(\tau), \eta_\lambda(\tau) - \eta(\tau)\rangle \mathrm{d}\tau = \int_0^T \langle \chi_{(0,t)}(\tau) f(\tau), \eta_\lambda(\tau) - \eta(\tau)\rangle \mathrm{d}\tau \to 0,$$

where $\chi_{(0,t)}$ stands for the characteristic function over $(0,t)$, at each $t \in (0,T)$. Moreover, we observe that

$$\left| \int_0^t \langle f(\tau), \eta_\lambda(\tau) - \eta(\tau) \rangle \mathrm{d}\tau \right| \le C$$

uniformly for all $t \in (0,T)$ and λ. Thus, convergence (3.39) follows.

4. Causal limit. Let us now come to the proof of our main Theorem 2.4. The argument here differs from that of [4] as regards a priori estimation. Indeed, in contrast to [4], no uniform estimate directly follows from energy inequalities obtained in Theorem 3.2.

By using assumption (2.4) for φ^* into estimate (3.8) we deduce

$$\begin{aligned} &\frac{c_0}{2} \int_0^T |f(t) - v_\varepsilon'(t)|_{X^*}^{m'} \mathrm{d}t + \psi^*(v_\varepsilon(T)) \\ &\qquad \le \varepsilon \varphi(0) + C \left(\int_0^T |g(t)|_X^m \mathrm{d}t + T \right) + \psi^*(v_0) + \int_0^T \langle f(t), \eta_\varepsilon(t) \rangle \, \mathrm{d}t. \end{aligned} \tag{4.1}$$

Hence, it follows from estimate (3.6) with the m'-boundedness of $\mathrm{D}\varphi^*$ that

$$\begin{aligned} &\int_0^T \psi^*(v_\varepsilon(t)) \mathrm{d}t \\ &\quad \le \varepsilon T \varphi(0) + CT \left(\int_0^T |g(\tau)|_X^m \mathrm{d}\tau + T \right) + T\psi^*(v_0) + \int_0^T \int_0^t \langle f(\tau), \eta_\varepsilon(\tau) \rangle \, \mathrm{d}\tau \mathrm{d}t \\ &\qquad + \varepsilon C \left(\varepsilon \varphi(0) + \int_0^T |g(t)|_X^m \mathrm{d}t + T + \psi^*(v_0) + \int_0^T \langle f(t), \eta_\varepsilon(t) \rangle \, \mathrm{d}t + \psi(0) \right). \end{aligned}$$

From the fact that $\psi^* \in \Phi_{p'}(V^*)$, we have

$$|\eta_\varepsilon(t)|_V^p \le C \left(\psi^*(v_\varepsilon(t)) + 1 \right). \tag{4.2}$$

Hence, we obtain

$$\int_0^T \psi^*(v_\varepsilon(t)) \mathrm{d}t \le C_T \left(\varphi(0) + \psi(0) + \int_0^T |g(\tau)|_X^m \mathrm{d}\tau + \psi^*(v_0) + \int_0^T |f(\tau)|_{V^*}^{p'} \mathrm{d}\tau + 1 \right). \tag{4.3}$$

Therefore, the estimates

$$\int_0^T |\eta_\varepsilon(t)|_V^p \mathrm{d}t + \int_0^T |v_\varepsilon(t)|_{V^*}^{p'} \mathrm{d}t + \int_0^T |v_\varepsilon'(t)|_{X^*}^{m'} \mathrm{d}t + \int_0^T |\xi_\varepsilon(t)|_X^m \mathrm{d}t \le C \tag{4.4}$$

follow from (4.2), the p'-coercivity of ψ^*, bound (4.1), and the m'-boundedness of $\mathrm{D}\varphi^*$, respectively. Moreover, by comparison in (2.10) we have

$$\|\varepsilon \xi_\varepsilon'\|_{L^p(0,T;V) + L^m(0,T;X)} \le C.$$

Therefore, we obtain, up to some not relabeled subsequence,

$$v_\varepsilon \to v \quad \text{weakly in } W^{1,m'}(0,T;X^*) \cap L^{p'}(0,T;V^*), \tag{4.5}$$

$$\eta_\varepsilon \to \eta \quad \text{weakly in } L^p(0,T;V), \tag{4.6}$$

$$\xi_\varepsilon \to \xi \quad \text{weakly in } L^m(0,T;X), \tag{4.7}$$

$$\varepsilon \xi_\varepsilon' \to 0 \quad \text{weakly in } L^p(0,T;V) + L^m(0,T;X), \tag{4.8}$$

which implies $-\xi + \eta = g$ by (2.10) and

$$\varepsilon\xi_\varepsilon(t) \to 0 \quad \text{weakly in } V \quad \text{for all } t \in [0,T]$$

because $\xi_\varepsilon(T) = 0$. Furthermore, since V^* is compactly embedded in X^*, we have

$$v_\varepsilon \to v \quad \text{strongly in } C([0,T];X^*), \tag{4.9}$$

which yields $v(0) = v_0$. Passing to the limit as $\varepsilon \to 0$ in (3.7), we have

$$\limsup_{\varepsilon\to 0} \int_0^T \langle v_\varepsilon(t), \eta_\varepsilon(t)\rangle \mathrm{d}t \le \int_0^T \langle v(t), \xi(t) + g(t)\rangle \mathrm{d}t,$$

which together with convergences (4.5)–(4.6) gives $\eta(t) \in \partial\psi^*(v(t))$ for almost every $t \in (0,T)$ by the maximal monotonicity of $\partial\psi^*$.

Then, recall estimate (3.8) in order to get

$$\begin{aligned}\limsup_{\varepsilon\to 0} &\int_0^T \langle f(t) - v_\varepsilon'(t), \xi_\varepsilon(t)\rangle \,\mathrm{d}t \\ &\le -\psi^*(v(T)) - \int_0^T \langle f(t) - v'(t), g(t)\rangle \,\mathrm{d}t + \psi^*(v_0) + \int_0^T \langle f(t), \eta(t)\rangle \,\mathrm{d}t,\end{aligned}$$

which ensures $\xi(t) = \mathrm{D}\varphi^*(f(t) - v'(t))$ for almost every $t \in (0,T)$.

Let us close this section by recording the following.

COROLLARY 4.1 (existence of solutions for (2.6)). *Assume* (A0)–(A3) *with* $g \equiv 0$. *Then the Cauchy problem* (2.6), (2.7) *admits at least one strong solution.*

Before closing this section let us explicitly record that this new variational technique provides a novel development of the existence theory for relation (2.6). Indeed, existence of strong solutions for relation (2.6) has already been established by Barbu for a time-differentiable forcing term. More precisely, in [9] it is assumed that $f \in W^{1,2}(0,T;X^*) \cap L^2(0,T;H)$ with some pivot Hilbert space H in a Gel'fand triplet setting, $X \hookrightarrow H \equiv H^* \hookrightarrow X^*$. Here we do not assume the differentiability of the forcing term g in time and we do not rely on such a triplet setting.

An existence result under weaker regularity assumptions for f has also been obtained by Maitre and Witomski in [33]. There f is just required to belong to $L^{m'}(0,T;X^*)$, no coercivity of ψ is imposed, but $p = m$. The present existence requires no restriction on p and m instead.

5. Improvements of the abstract framework in [4]. The present analysis can be shown to improve the former in [4] to the case $p < 2$ and WED functionals possess multiple minimizers. Indeed, let V and X be reflexive Banach spaces such that X is compactly and densely embedded in V and $|\cdot|_X$ and $|\cdot|_{X^*}$ are uniformly convex and strictly convex norms of X and X^*, respectively. Let ψ (respectively, ϕ) be a Gâteaux differentiable (respectively, proper lower semicontinuous) convex functional defined in V. We consider the abstract doubly nonlinear evolution equation

$$\mathrm{D}\psi(u') + \partial\phi(u) \ni g \quad \text{in } (0,T), \quad u(0) = u_0, \tag{5.1}$$

for $g : (0,T) \to V^*$ and $v_0 \in D(\phi)$. Then under the assumptions that $\psi \in \Phi_p(V)$ and $\phi_X \in \Phi_m(X)$, where ϕ_X denotes the restriction of ϕ to X, the corresponding WED functional is given by

$$I_\varepsilon(u) := \int_0^T \frac{e^{-t/\varepsilon}}{\varepsilon}\big(\varepsilon\psi(u'(t)) + \phi(u(t)) - \langle g(t), u(t)\rangle\big)\mathrm{d}t$$

along with the effective domain

$$D(I_\varepsilon) = \left\{ u \in W^{1,p}(0,T;V) \cap L^m(0,T;X) \colon u(0) = u_0 \right\}.$$

The corresponding Euler–Lagrange equation reads

$$-\varepsilon \frac{\mathrm{d}}{\mathrm{d}t} \mathrm{D}\psi(u') + \mathrm{D}\psi(u') + \partial\phi_X(u) \ni g \quad \text{in} \quad (0,T), \tag{5.2}$$

$$u(0) = u_0, \quad \mathrm{D}\psi(u'(T)) = 0. \tag{5.3}$$

Applying the present abstract theory, we deduce the following (cf. [4]; it is assumed that $p \geq 2$, $g \equiv 0$, and either ψ or ϕ is strictly convex).

THEOREM 5.1. *Assume that $\psi \in \Phi_p(V)$ and $\phi_X \in \Phi_m(X)$ with some $1 < m, p < \infty$, and that ψ is Gâteaux differentiable in V. Then, for each $g \in L^{p'}(0,T;V^*)$ and $u_0 \in X$,*

(i) *for all $\varepsilon > 0$, the WED functionals I_ε admit minimizers over $L^p(0,T;V)$, and moreover, every minimizer u_ε of I_ε becomes a strong solution of* (5.2)–(5.3)*;*

(ii) *up to some not relabeled subsequence, as $\varepsilon \to 0$ the trajectory u_ε converges to a strong solution u of* (5.1) *in the following sense*:

$$u_\varepsilon \to u \quad \begin{array}{l} \textit{strongly in} \ \ C([0,T];V), \\ \textit{weakly in} \ \ W^{1,p}(0,T;V) \cap L^m(0,T;X). \end{array}$$

Proof. Define $\tilde{\psi} : V \to [0,\infty)$ by $\tilde{\psi}(u) := \psi(-u)$ for $u \in V$, so that $\mathrm{D}\psi(u) = -\mathrm{D}\tilde{\psi}(-u)$. Hence, (5.1) is equivalently rewritten as

$$-\mathrm{D}\tilde{\psi}(-u') + \partial\varphi(u) \ni g \quad \text{in} \quad (0,T), \quad u(0) = u_0.$$

Finally, we apply Theorems 3.2 and 2.4 and Proposition 2.3 with $f \equiv 0$ and by replacing v, X^*, V^*, φ^*, ψ^* (or $\overline{\psi^*}$), m', and p' with u, V, X, $\tilde{\psi}$, ϕ_X (or ϕ), p, and m, respectively. ∎

6. Applications to nonlinear diffusion. As mentioned in the introduction, the present abstract framework is designed to encompass nonlinear diffusion equations such as the Stefan problem and the porous medium equation. In this section, we provide some detail in this direction. Let us consider the doubly nonlinear diffusion equation of the form

$$v_t - \Delta_m u = f, \quad v \in \alpha(u) \quad \text{in} \ \ \Omega \times (0,T), \tag{6.1}$$

$$u = 0 \quad \text{on} \ \ \partial\Omega \times (0,T), \tag{6.2}$$

$$v(\cdot,0) = v_0 \quad \text{in} \ \ \Omega, \tag{6.3}$$

where Ω is a bounded domain of $\mathbb{R}^d$ with smooth boundary $\partial\Omega$ and Δ_m stands for the m-Laplacian given by $\Delta_m u = \mathrm{div}(|\nabla u|^{m-2}\nabla u)$ with $1 < m < \infty$. (In particular, if $m = 2$, then Δ_m is the usual linear Laplacian.) Moreover, α is a (possibly multivalued) maximal monotone graph satisfying the following hypotheses with $1 < p < \infty$:

$$c_1|u|^p \leq \widehat{\alpha}(u) + c_2 \ \ \text{for all} \ \ u \in \mathbb{R}, \quad |\eta|^{p'} \leq c_3(|u|^p + 1) \ \ \text{for all} \ \ \eta \in \alpha(u) \ \ \text{and} \ \ u \in \mathbb{R},$$

where $\widehat{\alpha}$ denotes a primitive of α (i.e., $\partial\widehat{\alpha} = \alpha$). In order to variationally reformulate the problem into the doubly nonlinear evolution equation (2.6), we set

$$\psi(u) := \int_\Omega \widehat{\alpha}(u(x))\mathrm{d}x \ \ \text{for} \ \ u \in V, \quad \varphi(u) := \frac{1}{m}\int_\Omega |\nabla u(x)|^m \mathrm{d}x \ \ \text{for} \ \ u \in X$$

with $V := L^p(\Omega)$ and $X := W_0^{1,m}(\Omega)$. Then (A0) holds under the assumption that

$$1 < p < m^* := \begin{cases} \dfrac{dm}{d-m} & \text{if } m < d, \\ \infty & \text{otherwise,} \end{cases} \tag{6.4}$$

which particularly entails the compact embedding $X \hookrightarrow V$. Moreover, one can easily observe that $\phi \in \Phi_m(X)$, and furthermore, ϕ is uniformly convex and Fréchet differentiable in X. Therefore by [12], we have that φ^* is Fréchet differentiable in X^*. Thus (A1) holds.

Now, let us check assumptions for the two prototypical examples of α.

- *Stefan problem*: $\alpha(u) = u + H(u)$, where $H(\cdot)$ is defined by

$$H(u) = \begin{cases} 1 & \text{if } u > 0, \\ [0,1] & \text{if } u = 0, \\ 0 & \text{if } u < 0. \end{cases}$$

  Then $\widehat{\alpha}(u) = (1/2)u^2 + H(u)u$, and therefore, by setting $p = 2$, one has (A2).
- *Porous medium and fast diffusion equations*: $\alpha(u) = |u|^{p-2}u$ with $1 < p < 2$ for the porous medium equation and $p > 2$ for the fast diffusion equation. Then, (A2) is satisfied for each p.

Then for all $f \in L^{p'}(0,T;L^{p'}(\Omega)) \cap L^{m'}(0,T;W^{-1,m'}(\Omega))$ and $v_0 \in L^{p'}(\Omega)$, we can apply the main results to (6.1). In particular, the corresponding WED functional W_ε is given by

$$W_\varepsilon(v) = \int_0^T e^{-t/\varepsilon} \left(\frac{\varepsilon}{m'} \left\| (-\Delta_m)^{-1}(f - v_t) \right\|_{L^m(\Omega)}^m + \int_\Omega \widehat{\alpha}^*(v) \mathrm{d}x \right) \mathrm{d}t$$

for $v \in W^{1,m'}(0,T;W^{-1,m'}(\Omega)) \cap L^p(\Omega \times (0,T))$ satisfying $v(\cdot,0) = v_0$. Here $(-\Delta_m)^{-1} : W^{-1,m'}(\Omega) \to W_0^{1,m}(\Omega)$; $\xi \mapsto u$ denotes the solution operator

$$-\Delta_m u = \xi \ \text{ in } \Omega, \quad u = 0 \ \text{ on } \partial\Omega.$$

Furthermore, $\widehat{\alpha}^*(v)$ is given by

$$\widehat{\alpha}^*(v) = \frac{1}{2}((-u)^+)^2 + \frac{1}{2}((u-1)^+)^2$$

for the Stefan problem (where $r^+ := \max\{0, r\}$) and $\widehat{\alpha}^*(u) = (1/p')|u|^{p'}$ with $p' = p/(p-1)$ for the porous media and the fast diffusion equation.

COROLLARY 6.1 (WED formulation for nonlinear diffusion). *Let Ω be a bounded domain of $\mathbb{R}^d$ with smooth boundary $\partial\Omega$, $1 < m, p < \infty$, and let $f \in L^{p'}(0,T;L^{p'}(\Omega)) \cap L^{m'}(0,T;W^{-1,m'}(\Omega))$ and $v_0 \in L^{p'}(\Omega)$. Nonlinear diffusion equations of the form* (6.1) *involving the Stefan problem as well as porous medium/fast diffusion equations can be formulated as minimization problems of the WED functionals W_ε above $L^{m'}(0,T;W^{-1,m'}(\Omega))$, provided that* (6.4) *holds for the fast diffusion case. Namely, let v_ε be a minimizer of W_ε for each $\varepsilon > 0$. Then, up to a subsequence, v_ε converges to a limit v strongly in $C([0,T];W^{-1,m'}(\Omega))$ and weakly in $W^{1,m'}(0,T;W^{-1,m'}(\Omega)) \cap L^{p'}(\Omega \times (0,T))$ as $\varepsilon \to 0$, and moreover, the limit v solves* (6.1)–(6.3).

As a concluding remark let us mention that the closely related Hele-Shaw equation cannot be covered by our theory. Indeed, in the Hele-Shaw problem the primitive function $\widehat{\alpha}(u) = H(u)u$ of the corresponding maximal monotone graph $\alpha(u) = H(u)$ is no longer coercive. Then, ψ does not belong to $\Phi_p(V)$ for $V = L^p(\Omega)$ and any $1 < p < \infty$, and hence it does not fall within the scope of the main results of the present paper. Likewise, first-order derivatives (advection) terms are not covered here, since the elliptic operator is required to be symmetric. The more general maximal monotone case is covered in [18] but with more severe growth restrictions on the operators.

Appendix A.

PROPOSITION A.1. *Let E be a Banach space and F_E a duality mapping from E to its dual space E^*. Suppose that F_E is single-valued and hemicontinuous (i.e., $F_E(u+he) \to F_E(u)$ weakly in E^* as $h \to 0$ for any $u, e \in E$). Let P be a functional defined on E by*

$$P(u) := \frac{1}{p}|u|_E^p \quad \textit{for} \;\; u \in E.$$

Then P is Gâteaux differentiable in E and

$$\mathrm{D}P(u) = |u|_E^{p-2} F_E(u) \quad \textit{for} \;\; u \in E.$$

In addition, if F_E is continuous from E to E^*, then P is Fréchet differentiable in E.

Proof. It is well known (see, e.g., [8]) that

$$F_E(u) = \partial\left(\frac{1}{2}|u|_E^2\right) \quad \text{for all} \;\; u \in E.$$

For $h > 0$ and $u, e \in E$, we observe

$$\begin{aligned}\frac{P(u+he) - P(u)}{h} &= \frac{1}{ph}\left(|u+he|_E^p - |u|_E^p\right)\\ &\geq \frac{1}{2h}|u|_E^{2\cdot\frac{p-2}{2}}\left(|u+he|_E^2 - |u|_E^2\right)\\ &\geq |u|_E^{p-2}\langle F_E(u), e\rangle,\end{aligned}$$

and moreover, by the hemicontinuity of F_E,

$$\begin{aligned}\frac{P(u+he) - P(u)}{h} &= \frac{1}{ph}\left(|u+he|_E^p - |u|_E^p\right)\\ &\leq \frac{1}{2h}|u+he|_E^{2\cdot\frac{p-2}{2}}\left(|u+he|_E^2 - |u|_E^2\right)\\ &\leq |u+he|_E^{p-2}\langle F_E(u+he), e\rangle \to |u|_E^{p-2}\langle F_E(u), e\rangle.\end{aligned}$$

Thus we deduce that

$$\lim_{h\searrow 0}\frac{P(u+he) - P(u)}{h} = \left\langle |u|_E^{p-2}F_E(u), e\right\rangle.$$

By repeating a similar argument for $h < 0$, one can conclude that P is Gâteaux differentiable in E and

$$\langle \mathrm{D}P(u), e\rangle = \left\langle |u|_E^{p-2} F_E(u), e\right\rangle \quad \text{for all} \quad u, e \in E.$$

As for the case that F_E is continuous in E, the Gâteaux derivative $\mathrm{D}P$ is continuous in E, and therefore, P is Fréchet differentiable in E. $\square$

Remark A.2. We shall mention two facts:

(i) If $|\cdot|_{E^*}$ is strictly convex, then F_E is single-valued and demicontinuous.

(ii) If E^* is uniformly convex, then F_E is locally uniformly continuous from E to E^*.

See, e.g., [8] for more details. Combining (i) above with Proposition A.1, one can conclude in section 3 that P is Gâteaux differentiable in X^*. Moreover, both I_ε^3 and $\mathcal{P}$ are Gâteaux differentiable in $L^{m'}(0,T;X^*)$. We can also reveal the representations of their derivatives.

Acknowledgment. We express our gratitude to the anonymous referees for the careful reading of the manuscript.

REFERENCES

[1] S. Aizicovici and V.-M. Hokkanen, *Doubly nonlinear equations with unbounded operators*, Nonlinear Anal., 58 (2004), pp. 591–607.

[2] S. Aizicovici and V.-M. Hokkanen, *Doubly nonlinear periodic problems with unbounded operators*, J. Math. Anal. Appl., 292 (2004), pp. 540–557.

[3] G. Akagi and U. Stefanelli, *A variational principle for doubly nonlinear evolution*, Appl. Math. Lett., 23 (2010), pp. 1120–1124.

[4] G. Akagi and U. Stefanelli, *Weighted energy-dissipation functionals for doubly nonlinear evolution*, J. Funct. Anal., 260 (2011), pp. 2541–2578.

[5] H. W. Alt and S. Luckhaus, *Quasilinear elliptic-parabolic differential equations*, Math. Z., 183 (1983), pp. 311–341.

[6] E. Asplund, *Averaged norms*, Israel J. Math., 1967, pp. 227–233.

[7] E. Asplund, *Topics in the theory of convex functions*, in Theory and Applications of Monotone Operators (Proceedings of the NATO Advanced Study Institute Venice, 1968), Oderisi, Gubbio, 1969, pp. 1–33.

[8] V. Barbu, *Nonlinear Semigroups and Differential Equations in Banach Spaces*, P. Noordhoff, Groningen, Netherlands, 1976.

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

[10] V. Barbu, *Existence theorems for a class of two point boundary problems*, J. Differential Equations, 17 (1975), pp. 236–257.

[11] F. Bernis, *Existence results for doubly nonlinear higher order parabolic equations on unbounded domains*, Math. Ann., 279 (1988), pp. 373–394.

[12] J. M. Borwein and J. Vanderwerff, *Fréchet-Legendre functions and reflexive Banach spaces*, J. Convex Anal., 17 (2010), pp. 915–924.

[13] H. Brézis, *Monotonicity methods in Hilbert spaces and some applications to non-linear partial differential equations*, in Contributions to Nonlinear Functional Analysis, E. Zarantonello, ed., Academic Press, New York, 1971, pp. 101–156.

[14] H. Brézis, M. G. Crandall, and A. Pazy, *Perturbations of nonlinear maximal monotone sets in Banach space*, Comm. Pure. Appl. Math., 23 (1970), pp. 123–144.

[15] H. Brézis and I. Ekeland, *Un principe variationnel associé à certaines équations paraboliques. Le cas indépendant du temps*, C. R. Acad. Sci. Paris Sér. A-B, 282 (1976), pp. A971–A974.

[16] H. Brézis and I. Ekeland, *Un principe variationnel associé à certaines équations paraboliques. Le cas dépendant du temps*, C. R. Acad. Sci. Paris Sér. A-B, 282 (1976), pp. A1197–A1198.

[17] S. Conti and M. Ortiz, *Minimum principles for the trajectories of systems governed by rate problems*, J. Mech. Phys. Solids, 56 (2008), pp. 1885–1904.

[18] E. Di Benedetto and R. E. Showalter, *Implicit degenerate evolution equations and applications*, SIAM J. Math. Anal., 12 (1981), pp. 731–751.

[19] H. Gajewski and I. V. Skrypnik, *To the uniqueness problem for nonlinear parabolic equations*, Discrete Contin. Dyn. Syst., 10 (2004), pp. 315–336.

[20] N. Ghoussoub, *Selfdual Partial Differential Systems and Their Variational Principles*, Springer, New York, 2009.

[21] N. Ghoussoub and L. Tzou, *A variational principle for gradient flows*, Math. Ann., 330 (2004), pp. 519–549.

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

[23] N. Hirano, *Existence of periodic solutions for nonlinear evolution equations in Hilbert spaces*, Proc. Amer. Math. Soc., 120 (1994), pp. 185–192.

[24] V.-M. Hokkanen, *An implicit nonlinear time dependent equation has a solution*, J. Math. Anal. Appl., 161 (1991), pp. 117–141.

[25] 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), pp. 29–49.

[26] V.-M. Hokkanen, *On nonlinear Volterra equations in Hilbert spaces*, Differential Integral Equations, 5 (1992), pp. 647–669.

[27] T. Ilmanen, *Elliptic regularization and partial regularity for motion by mean curvature*, Mem. Amer. Math. Soc., 108 (1994), pp. x+90.

[28] N. Kenmochi, *Some nonlinear parabolic variational inequalities*, Israel J. Math., 22 (1975), pp. 304–331.

[29] C. J. Larsen, M. Ortiz, and C. L. Richardson, *Fracture paths from front kinetics: Relaxation and rate independence*, Arch. Ration. Mech. Anal., 193 (2009), pp. 539–583.

[30] J.-L. Lions and E. Magenes, *Non-homogeneus Boundary Value Problems and Applications*, vol. 1, Springer, New York, 1972.

[31] M. Liero and U. Stefanelli, *Weighted Inertia-Dissipation-Energy functionals for semilinear equations*, Boll. Unione Mat. Ital. (9), 6 (2013), pp. 1–27.

[32] M. Liero and U. Stefanelli, *A new minimum principle for Lagrangian mechanics*, J. Nonlinear Sci., 23 (2013), pp. 179–204.

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

[34] G. Marinoschi, *A variational approach to nonlinear diffusion equations with time periodic coefficients*, Ann. Univ. Buchar. Math. Ser., 3 (2012), pp. 173–185.

[35] G. Marinoschi, *Existence of solutions to time-dependent nonlinear diffusion equations via convex optimization*, J. Optim. Theory Appl., 154 (2012), pp. 792–817.

[36] A. Mielke and M. Ortiz, *A class of minimum principles for characterizing the trajectories and the relaxation of dissipative systems*, ESAIM Control Optim. Calc. Var., 14 (2008), pp. 494–516.

[37] A. Mielke and U. Stefanelli, *A discrete variational principle for rate-independent evolution*, Adv. Calc. Var., 1 (2008), pp. 399–431.

[38] A. Mielke and U. Stefanelli, *Weighted energy-dissipation functionals for gradient flows*, ESAIM Control Optim. Calc. Var., 17 (2011), pp. 52–85.

[39] R. Rossi, G. Savaré, A. Segatti, and U. Stefanelli, *Weighted energy-dissipation functionals for gradient flows in metric spaces*, in preparation, 2014.

[40] B. Nayroles, *Deux théorèmes de minimum pour certains systèmes dissipatifs*, C. R. Acad. Sci. Paris Sér. A-B, 282 (1976), pp. A1035–A1038.

[41] B. Nayroles, *Un théorème de minimum pour certains systèmes dissipatifs, Variante hilbertienne*, Travaux Sém. Anal. Convexe, 6 (1976), p. 22.

[42] R. Rossi, G. Savaré, A. Segatti, and U. Stefanelli, *A variational principle for gradient flows in metric spaces*, C. R. Math. Acad. Sci. Paris, 349 (2011), pp. 1224–1228.

[43] E. Serra and P. Tilli, *Nonlinear wave equations as limits of convex minimization problems: Proof of a conjecture by De Giorgi*, Ann. of Math. (2), 175 (2012), pp. 1551–1574.

[44] E. N. Spadaro and U. Stefanelli, *A variational view at mean curvature evolution for linear growth functionals*, J. Evol. Equ., 11 (2011), pp. 793–809.

[45] U. Stefanelli, *The De Giorgi conjecture on elliptic regularization*, Math. Models Methods Appl. Sci., 21 (2011), pp. 1377–1394.

[46] U. Stefanelli, *Well-posedness and time discretization of a nonlinear Volterra integrodifferential equation*, J. Integral Equations Appl., 13 (2001), pp. 273–304.

[47] U. Stefanelli, *On a class of doubly nonlinear nonlocal evolution equations*, Differential Integral Equations, 15 (2002), pp. 897–922.

[48] U. Stefanelli, *On some nonlocal evolution equations in Banach spaces*, J. Evol. Equ., 4 (2004), pp. 1–26.

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

[50] A. Visintin, *Models of Phase Transitions*, Progr. Nonlinear Differential Equations Appl. 28, Birkhäuser, Boston, 1996.

[51] A. Visintin, *Extension of the Brezis-Ekeland-Nayroles principle to monotone operators*, Adv. Math. Sci. Appl., 18 (2008), pp. 633–650.

[52] A. Visintin, *Structural stability of doubly-nonlinear flows*, Boll. Unione Mat. Ital. (9), 4 (2011), pp. 363–391.

[53] A. Visintin, *Structural stability of rate-independent nonpotential flows*, Discrete Contin. Dyn. Syst. Ser. S, 6 (2013), pp. 257–275.