ESAIM: COCV 25 (2019) 36
https://doi.org/10.1051/cocv/2018023

ESAIM: Control, Optimisation and Calculus of Variations
www.esaim-cocv.org

THE WEIGHTED ENERGY-DISSIPATION PRINCIPLE AND EVOLUTIONARY Γ-CONVERGENCE FOR DOUBLY NONLINEAR PROBLEMS

Matthias Liero[1,*] and Stefano Melchionna[2]

Abstract. We consider a family of doubly nonlinear evolution equations that is given by families of convex dissipation potentials, nonconvex energy functionals, and external forces parametrized by a small parameter ε. For each of these problems, we introduce the so-called weighted energy-dissipation (WED) functional, whose minimizers correspond to solutions of an elliptic-in-time regularization of the target problems with regularization parameter δ. We investigate the relation between the Γ-convergence of the WED functionals and evolutionary Γ-convergence of the associated systems. More precisely, we deal with the limits $\delta \to 0$, $\varepsilon \to 0$, as well as $\delta + \varepsilon \to 0$ either in the sense of Γ-convergence of functionals or in the sense of evolutionary Γ-convergence of functional-driven evolution problems, or both. Additionally, we provide some quantitative estimates on the rate of convergence for the limit $\varepsilon \to 0$, in the case of quadratic dissipation potentials and uniformly λ-convex energy functionals. Finally, we discuss a homogenization and a dimension reduction problem as examples of application.

Mathematics Subject Classification. 58E30, 35K55, 47J35.

Received June 26, 2017. Accepted March 22, 2018.

1. Introduction

In this text, we discuss an abstract convergence result for solutions to a family of doubly nonlinear equations depending on a small parameter $\varepsilon > 0$

$$\mathrm{d}\psi_\varepsilon(\dot{u}) + \partial\phi_\varepsilon(u) \ni g_\varepsilon(t), \qquad u(0) = u_\varepsilon^0. \tag{P$_\varepsilon$}$$

Here $\dot{u}$ denotes the time derivative of the unknown $u : [0,T] \to H$, where H is a reflexive Banach space. The evolution is driven by a dissipation functional $\psi_\varepsilon : H \to [0,\infty)$, assumed to be convex and Gâteaux differentiable, by an energy functional $\phi_\varepsilon : H \to (-\infty,\infty]$ with a suitable notion of (sub)differential $\partial\phi_\varepsilon$ (see below), and by time-dependent external forces $t \mapsto g_\varepsilon(t)$.

The abstract system (P$_\varepsilon$) can describe a variety of different dissipative problems in a large number of applications from mechanics to thermodynamics, from population dynamics to finance, just to mention a few.

Keywords and phrases: Doubly nonlinear evolution, weighted-energy-dissipation principle, evolutionary Γ-convergence, variational principle, homogenization, dimension reduction.

[1] Weierstrass Institute Berlin, Mohrenstraße 39, 10117 Berlin, Germany.
[2] Faculty of Mathematics, University of Vienna, Oskar-Morgenstern-Platz 1, A-1090 Vienna, Austria.

[*] Corresponding author: matthias.liero@wias-berlin.de

Problems of this type have been studied by several authors, and we refer to [9, 10, 21] and the references therein for a survey.

The parameter $\varepsilon > 0$ can have different origins: In real world applications often the problem's data are affected by errors or a mathematical model might be too complex such that one is interested in simplifying it without loosing key features. This is the case for example in homogenization problems, where fast oscillations of the coefficient, describing the physical microstructure, can be removed with a limiting procedure to obtain an effective macroscopic model. In this sense, the parameter ε can be understood as a ratio between micro- and macroscopic scales in the system. In other applications, ε might be given by a numerical approximation denoting for example the fineness of the grid.

A natural question is whether solutions to (P_ε) are stable with respect to the these perturbations. In other words, is it possible to prove convergence of the solutions to (P_ε) to solutions of an effective equation assuming convergence of the energy and dissipation functionals and of the data in some proper sense?

We remark that the question of convergence of solutions to problems under suitable assumption on the convergence of operators/functionals involved in the definition of the problem is classical and has been studied by several authors, see, *e.g.* [7] or the monograph [6]. More recently, the above question has been formalized in the notion of *evolutionary* Γ*-convergence* (E-convergence) introduced by Mielke [17] for the case $g_\varepsilon \equiv 0$: We say that a system (ϕ_0, ψ_0, g_0) is the evolutionary Γ-limit of the system associated with (P_ε), if the convergence of the initial data implies that limits of the curves $t \mapsto u_\varepsilon(t)$ are solutions with respect to (ϕ_0, ψ_0, g_0). Sufficient conditions for evolutionary Γ-convergence of dissipative problems have been provided in terms of Γ- and Mosco convergence of the corresponding dissipation and energy functionals, see, *e.g.* [23]. In particular, the convergence results in the survey paper [17] are based on two equivalent formulations of (P_ε): (i) the energy-dissipation principle formulation (EDP formulation) and (ii) the formulation as an evolutionary variational inequality (EVI formulation) in the case that ϕ_ε satisfies a uniform λ-convexity property. The EDP formulation is written in terms of a scalar balance between the final and initial energy and the total dissipation, which in turn is written as sum of the primal and dual dissipation potentials, where the latter is given by the Legendre transform of ψ_ε. The E-convergence result then follows from proving a lower Γ-limit for the involved functionals and establishing that the resulting dissipation functional has again a (ψ_0, ψ_0^*) structure. In contrast, the EVI approach exploits the uniform λ-convexity of the energy functional ϕ_ε, which might not be satisfied in certain applications. In this paper, we are interested in studing E-convergence by means of a variational method. Before illustrating our results let us mention that results in a similar spirit have been obtained by Visintin (see, *e.g.*, [27] and references therein). More precisely, evolution equations are reformulated as null-minimization problems and weak notion of evolutionary Γ-convergence is investigated. See also [20] for a homogenization example.

Our approach to E-convergence is based on a variational formulation of (P_ε) using the so-called *Weighted-Energy-Dissipation* (WED) principle, see, *e.g.* [1, 4, 5]. Given a target evolutionary problem, the WED principle consists of two steps: First, a global parameter-dependent functional $\mathcal{I}_{\varepsilon,\delta}$, defined over entire trajectories, is proved to admit minimizers. In the case of (P_ε) the WED functional has the form

$$\mathcal{I}_{\varepsilon,\delta}(u) = \int_0^T \mathrm{e}^{-t/\delta} \left(\psi_\varepsilon(\dot{u}) + \frac{1}{\delta}\phi_\varepsilon(u) - \frac{1}{\delta} \langle g_\varepsilon(t), u \rangle_H \right) \mathrm{d}t. \tag{1.1}$$

These minimizers solve an elliptic-in-time regularization of the target problem, *i.e.* in the case of (P_ε) as target problem,

$$(\mathrm{P}_{\varepsilon,\delta}) \qquad \begin{gathered} -\delta \frac{\mathrm{d}}{\mathrm{d}t} \mathrm{d}\psi_\varepsilon(\dot{u}) + \mathrm{d}\psi_\varepsilon(\dot{u}) + \partial\phi_\varepsilon(u) \ni g(t) \quad \text{a.e. in } (0,T), \\ u(0) = u_\varepsilon^0, \qquad \delta \mathrm{d}\psi_\varepsilon(\dot{u}(T)) = 0. \end{gathered} \tag{1.2}$$

Second, minimizers (that also solve (1.2)) are proven to converge, up to subsequences, to solutions to the target problem, as the parameter δ goes to 0. Since solutions to (1.2) depend on the future, note in particular the final condition in (1.2), causality is lost for $\delta > 0$. Thus, the limit $\delta \to 0$ is usually referred to as *causal limit*.

The WED principle, originally proposed by Ilmanen [11] (see also [15]), has been brought to new attention by Mielke and Ortiz [18] in the context of rate-independent systems. Later, many authors widely extended the theory, especially concerning the range of applications, *i.e.* the target problem considered. The gradient flow case with λ-convex potentials has been studied by Mielke and Stefanelli [19]. Akagi and Stefanelli have extended the theory to the genuinely nonconvex case for gradient flows [5] and to convex doubly nonlinear systems [4]. Moreover, an analogous approach has been applied to some hyperbolic problems, *e.g.*, the semilinear wave equation [14, 24, 26], and to Lagrangian Mechanics [13]. Recently, nonpotential perturbation problems have also been considered [1, 16].

The interest in such a variational approach lies in the fact that variational methods for evolution equations allow to apply tools and technique of the calculus of variation in the evolutionary setting. This is indeed the spirit of this work, where we want to study *evolutionary* Γ*-convergence* (see below for a discussion) of equations, by looking at Γ*-convergence* of the associated WED functionals.

Throughout this paper, we assume that ϕ_ε can be decomposed into the difference of two functionals, namely $\phi_\varepsilon = \varphi^1_\varepsilon - \varphi^2_\varepsilon$, where $\varphi^1_\varepsilon, \varphi^2_\varepsilon : X \to (-\infty, \infty]$ are proper, lower semicontinuous, and convex functionals on a Banach space $X \subset H$ with φ^1_ε dominating φ^2_ε in a suitable sense. This will allow us to define $\partial\phi_\varepsilon$ as the difference of the convex subdifferentials $\partial\varphi^1_\varepsilon$ and $\partial\varphi^2_\varepsilon$ (see Thm. 2.1 for a more rigorous definition). These assumptions along with suitable uniform growth conditions for φ^1_ε, φ^2_ε, and ψ_ε yield the well-posedness of the WED principle for fixed $\varepsilon > 0$.

A first question, that we answer in this paper, is whether the WED procedure is stable under these perturbations, or, in other words, whether the WED functionals $\mathcal{I}_{\delta,\varepsilon}$ converge to $\mathcal{I}_{\delta,0}$ under some convergence assumptions on the functionals ψ_ε, ϕ_ε, and on the data u^0_ε and g_ε. Indeed, assuming static Γ-convergence of the energy functionals ϕ_ε, and continuous convergence of ψ_ε along strongly converging sequences in H (see Sect. 2.2), as well as the convergence of g_ε, u^0_ε in a proper sense, we prove Γ-convergence of the corresponding WED functionals (for $\delta > 0$ fixed). Note that Γ-convergence is a natural notion of convergence for functionals, as it implies convergence of minimizers. In particular, Γ-convergence of the WED functionals implies evolutionary Γ-convergence for the elliptic-regularized problems (1.2) at least for solutions that are also minimizers of the WED functionals . Our proof is based on using time-discrete approximations of curves $t \mapsto u(t)$ to obtain the lower $\liminf$ estimate and to construct recovery sequences.

Second, we consider the joint limit $\delta + \varepsilon \to 0$ and show that solutions to the elliptic regularized and perturbed problem, *i.e.* (1.2), converge to solutions to the target problem (P_ε) with $\varepsilon = 0$. Here we exploit the uniform growth conditions posed on the functionals to derive the necessary a priori estimates. Let us remark that the Γ-limit of WED functionals for $\delta \to 0$ is highly degenerate (intuitively it is just a constraint on the initial condition, compare Sect. 2.3). Thus, the limit $\delta + \varepsilon \to 0$ is meaningful only at the level of equations and not for functionals.

Finally, we address the question of obtaining explicit convergence rates for the Γ-convergence of the WED functionals. Here we restrict ourselves to the case of quadratic dissipation potentials ψ_ε and uniformly λ-convex energy functionals ϕ_ε. Assuming to have some information on the rates of convergence for a good recovery sequence of the static functionals, we deduce rates of convergence for the minimizers of the corresponding dynamic WED functionals. To the best of our knowledge, the strategy we use in our proof is new. It is based on a simple abstract result (*cf.* Lem. 3.3) which describes sufficient conditions for having quantitative estimates in Γ-convergence problems and time-discrete approximations.

We note here that a related result has been obtained in [4]. More precisely, the authors proved Mosco convergence of the WED functionals in the case $\varphi^2_\varepsilon \equiv 0$ and $g_\varepsilon \equiv 0$ assuming Mosco convergence of φ^1_ε in X and of ψ_ε in H. In contrast to our approach, it is required that a strongly in X converging joint recovery sequence for the static dissipation and energy functionals exists and that the well preparedness of initial data is satisfied. The latter is also necessary in the EDP approach in [17] and means that in additions to $u^0_\varepsilon \to u^0_0$ in H also $\phi_\varepsilon(u^0_\varepsilon) \to \phi_0(u^0_0)$ holds. Note that our result is more flexible since weaker conditions are assumed. In particular, we do not need to assume the existence of a joint recovery sequence for the dissipation and energy functionals. This sequence will be constructed by taking advantage of the coercivity of the energy functional in X and by using the continuous convergence of the dissipation functionals in H. Moreover, weaker assumptions on the Γ-convergence of the energy functionals allow us to deal with a much larger set of applications, namely families of gradient flows driven by Γ-convergent (but not necessary Mosco convergent) energy functionals (*cf.* the homogenization example in Sect. 4.1). The paper is structured as follows: in Section 2 we introduce the abstract setting for the WED principle and the evolutionary Γ-convergence and formulate the main results. In particular, we collect all assumptions on the dissipation potentials, energy functionals, and data that guarantee the well-posedness of the WED principle for fixed parameter $\varepsilon > 0$, hereby relying on the results in [1] (see Sect. 2.1). Next, we fix conditions on the convergence of the "static" functionals ϕ_ε and ψ_ε (Sect. 3.2) that allow us to proof the evolutionary Γ-convergence of (P_ε) in the setting of the WED principle and to obtain explicit convergence rates for the latter, see Theorems 2.2 and 2.5 as well as Theorem 2.7. The proofs of these results are collected in Section 3. Finally, we discuss in Section 4 two examples: (i) establishing the homogenized limit for a nonlinear parabolic equation with rapidly oscillating coefficient functions; (ii) deriving a lower-dimensional effective equation for a doubly nolinear parabolic equation on a thin domain.

2. Preliminaries and main result

In this section, we recall the basic framework of WED functionals and present the main results of this paper whose proofs are postponed to Section 3. We refer to [4, 19] for more details on the WED principle.

2.1. The WED principle

Let H and X be reflexive Banach spaces such that we have the dense and compact embedding $X \subset H$. On H we consider a convex dissipation potential $\psi : H \to [0, \infty)$, which is assumed to be Gâteaux differentiable. In particular, we denote by $\mathrm{d}\psi(v) \in H^*$ the Gâteaux differential for $v \in H$. The energy functional $\phi : H \to \mathbb{R}_\infty := \mathbb{R} \cup \{+\infty\}$ is assumed to be of the form $\phi = \varphi^1 - \varphi^2$, where $\varphi^1, \varphi^2 : H \to [0, \infty]$ are proper, lower semicontinuous, and convex functionals with domains $D(\varphi^i)$, $i = 1, 2$. To include the work of the external forces $g_\varepsilon(t) \in H^*$ we introduce the augmented energy functional $\widetilde{\phi} : [0, T] \times H \to \mathbb{R}_\infty$ by setting

$$\widetilde{\phi}(t, u) = \phi(u) - \langle g(t), u\rangle_H. \tag{2.1}$$

We impose the following growth assumptions on the dissipation and energy functional which are in accordance with [1]. In particular, we refer to Remark 1 of [1] for a discussion of the assumptions. Let $p \in (1, \infty)$ and $m \in (1, \infty)$ be fixed. We assume, that there exists a constant $C > 0$ such that

$$(\mathbf{A1})\ \|v\|_H^p \le C(\psi(v) + 1) \quad \text{for all} \quad v \in H, \tag{2.2a}$$

$$(\mathbf{A2})\ \|u\|_X^m \le C(\varphi^1(u) + 1) \quad \text{for all} \quad u \in D(\varphi^1). \tag{2.2b}$$

In particular, we have that $D(\varphi^1) \subset X$ and we will denote the restriction of φ^1 to X again by φ^1. The (convex) subdifferential with respect to X of φ^i in $u \in X$ is denoted by $\partial_X \varphi^i(u) \subset X^*$ and its domain by $D(\partial_X \varphi^i)$. The element in $\partial_X \varphi^i(u)$ that realizes the minimal X^*-norm is denoted by $(\partial_X \varphi^i(u))^\circ \in X^*$. We assume that

$$(\mathbf{A3}) \qquad \|\eta^1\|_{X^*}^{m'} \le C(\|u\|_X^m + 1) \qquad \text{for all } u \in X \text{ and } \eta^1 \in \partial_X \varphi^1(u), \tag{2.3a}$$

$$(\mathbf{A4}) \qquad \|\mathrm{d}\psi(v)\|_{H^*}^{p'} \le C(\|v\|_H^p + 1) \qquad \text{for all } v \in H, \tag{2.3b}$$

where $m' = m/(m-1)$ and $p' = p/(p-1)$ are the dual exponents.

Additionally, as in [1, 5] we make the assumption that there exist $\kappa \in (0, 1)$ and $c > 0$ such that we have

$$(\mathbf{A5}) \qquad \varphi^2(u) \le \kappa(\varphi^1(u) + c) \qquad \text{for all } u \in D(\varphi^1), \tag{2.4a}$$

(A6) $$\big\|\eta^2\big\|_{H^*}^{p'} \leq c\big(\varphi^1(u)+1\big) \qquad \text{for all } u\in D(\partial_X\varphi^1) \text{ and } \eta^2\in\partial_H\varphi^2(u). \tag{2.4b}$$

Note that due to the domination assumption (2.4a) we can exclude the ambiguous case $\infty-\infty$ in the definition of $\phi=\varphi^1-\varphi^2$ by setting

$$\phi(u)=\begin{cases}\varphi^1(u)-\varphi^2(u) & \text{if } u\in D(\varphi^1),\\ \infty & \text{else.}\end{cases}$$

Finally, for the external forces and the initial data we demand that

$$(\mathbf{A7})\quad g\in L^{p'}(0,T;H^*) \qquad \text{and} \qquad (\mathbf{A8})\quad u^0\in D(\varphi^1). \tag{2.5}$$

For a given time horizon $T>0$ we define the set of admissible trajectories

$$\mathcal{K}(u^0):=\big\{u\in W^{1,p}(0,T;H)\cap L^m(0,T;X): u(0)=u^0\big\}$$

and introduce for a (fixed) constant $\delta>0$ the *Weighted-Energy-Dissipation functional* (WED functional) $\mathcal{I}_\delta: L^p(0,T;H)\to\,]-\infty,\infty]$ *via*

$$\mathcal{I}_\delta(u)=\begin{cases}\int_0^T \mathrm{e}^{-t/\delta}\Big(\psi(\dot u)+\frac{1}{\delta}\widetilde{\phi}(t,u)\Big)\,\mathrm{d}t & \text{if } u\in\mathcal{K}(u^0),\\ \infty & \text{else.}\end{cases} \tag{2.6}$$

The WED principle is concerned with finding a minimizer $u_\delta\in\mathcal{K}(u^0)$ and passing to the limit $\delta\to0$ to recover a solution u of the original doubly nonlinear equation. In particular, in [1] the following result was proven.

Theorem 2.1 (Akagi–Melchionna [1], Thm. 13)**.**

(1) *Assume that* **(A1)**–**(A8)** *are satisfied. Then for every $\delta>0$ sufficiently small the WED functional $\mathcal{I}_\delta$ defined in (2.6) admits at least one global minimizer $u_\delta\in\mathcal{K}(u^0)$. Furthermore, every local minimizer u_δ solves (1.2) in the strong sense, i.e. there exists $\eta_\delta^1\in L^{m'}(0,T;X^*)$ and $\eta_\delta^2\in L^{p'}(0,T;H^*)$ with $\eta_\delta^1\in\partial_X\varphi^1(u_\delta)$ and $\eta_\delta^2\in\partial_H\varphi^2(u_\delta)$ such that*

$$\begin{aligned}-\delta\dot\xi_\delta+\xi_\delta+\eta_\delta^1-\eta_\delta^2=g(t),&\qquad \textit{in } X^* \textit{ a.e. in } (0,T),\\ \textit{and}\quad u_\delta(0)=u^0,&\qquad \delta\xi_\delta(T)=0,\end{aligned} \tag{2.7}$$

where $\xi_\delta=\mathrm{d}\psi(\dot u_\delta)$ satisfies $\xi_\delta\in L^{p'}(0,T;H^*)$ and $\dot\xi_\delta\in L^{m'}(0,T;X^*)+L^{p'}(0,T;H^*)$.

(2) *Moreover, up to subsequences, $u_\delta\to u$ weakly in $W^{1,p}(0,T;H)\cap L^m(0,T;X)$ and strongly in $C([0,T];H)$ to a solution u to*

$$\xi+\eta^1-\eta^2=g(t),\ \textit{in } X^* \textit{ a.e. in } (0,T), \textit{and} \quad u(0)=u^0 \tag{2.8}$$

with $\xi=\mathrm{d}\psi(\dot u)$ satisfying $\xi\in L^{p'}(0,T;H^*)$ and $\eta^1\in L^{m'}(0,T;X^*)$ and $\eta^2\in L^{p'}(0,T;H^*)$ with $\eta^1\in\partial_X\varphi^1(u)$ and $\eta^2\in\partial_H\varphi^2(u)$.

We easily see that limit problem (2.8) can be formally recovered by setting $\delta=0$ in (2.7).

2.2. Γ-convergence of the WED functionals

Let us now consider a family of WED functionals $\mathcal{I}_{\delta,\varepsilon}$ which in turn is induced by families of dissipation and energy functionals ψ_ε and ϕ_ε and external forces g_ε as in (2.6) depending on a small parameter $\varepsilon>0$ describing, *e.g.* the ratio between microscopic and macroscopic length scale. In the following theorem we provide sufficient conditions on the convergence of the dissipation and energy functionals and of the external force to limits ψ_0, ϕ_0, and g_0, respectively, that guarantee the Γ-convergence of the WED functionals $\mathcal{I}_{\delta,\varepsilon}$ to the limiting WED functional $\mathcal{I}_{\delta,0}$.

In particular, we shall assume that the energy functionals $\phi_\varepsilon = \varphi^1_\varepsilon - \varphi^2_\varepsilon$ converge in the following sense:

$$(\mathbf{\Gamma 1})\quad \varphi^1_\varepsilon \xrightarrow{\mathrm{M}} \varphi^1_0 \quad \text{in } H, \qquad (\mathbf{\Gamma 2})\quad \varphi^2_\varepsilon \xrightarrow{\mathrm{C}} \varphi^2_0 \quad \text{weakly in } X. \tag{2.9}$$

Here, "$\xrightarrow{\mathrm{M}}$" denotes Mosco convergence (Γ-convergence with respect to the strong *and* weak topology) and "$\xrightarrow{\mathrm{C}}$" means continuous convergence, *i.e.* $\varphi^2_\varepsilon(u_\varepsilon) \to \varphi^2_0(u^0)$ for all sequences with $u_\varepsilon \to u^0$ weakly in X. In particular, in the case of uniform coercivity of φ^1_ε on X the convergence in ($\mathbf{\Gamma 1}$) is equivalent to $\varphi^1_\varepsilon \xrightarrow{\Gamma} \varphi^1_0$ weakly in X (see, *e.g.* [17], Prop. 2.5).

Additionally, for the dissipation potentials ψ_ε we assume that

$$(\mathbf{\Gamma 3})\quad \psi_\varepsilon \xrightarrow{\mathrm{C}} \psi_0 \quad \text{strongly in } H. \tag{2.10}$$

On the forcing terms we impose

$$(\mathbf{\Gamma 4})\quad g_\varepsilon \to g_0 \quad \text{weakly in } L^{p'}(0,T;H^*). \tag{2.11}$$

Finally, we demand that the initial values u^0_ε yield energies that are finite and uniformly bounded in ε, *i.e.* there exists a constant $C > 0$ independent of ε such that

$$(\mathbf{\Gamma 5})\quad \phi_\varepsilon(u^0_\varepsilon) \le C. \tag{2.12}$$

In particular, given the uniform coercivity of φ^1_ε in X it is natural to assume that

$$u^0_\varepsilon \to u^0 \text{ weakly in } X \text{ (strongly in } H).$$

Theorem 2.2 (Γ-convergence of WED functionals)**.** *Let us assume that φ^1_ε, φ^2_ε, ψ_ε, and g_ε satisfy the assumptions* **(A1)–(A8)** *with C, c, κ, p, m independent of ε as well as the convergence conditions* **(Γ1)–(Γ5)***. Then, the WED functionals*

$$\mathcal{I}_{\delta,\varepsilon}(u) = \begin{cases} \int_0^T \mathrm{e}^{-t/\delta}\left(\psi_\varepsilon(\dot u(t)) + \frac{1}{\delta}\widetilde{\phi}_\varepsilon(t,u(t))\right)\mathrm{d}t & \text{if } u \in \mathcal{K}(u^0_\varepsilon),\\ \infty & \text{else,}\end{cases}$$

Γ-converge to

$$\mathcal{I}_{\delta,0}(u) = \begin{cases} \int_0^T \mathrm{e}^{-t/\delta}\left(\psi_0(\dot u(t)) + \frac{1}{\delta}\widetilde{\phi}_0(t,u(t))\right)\mathrm{d}t & \text{if } u \in \mathcal{K}(u^0),\\ \infty & \text{else,}\end{cases} \tag{2.13}$$

in the weak topology of $W^{1,p}(0,T;H) \cap L^m(0,T;X)$.

Furthermore, let $u^*_{\delta,\varepsilon}$ be a minimizer of $\mathcal{I}_{\delta,\varepsilon}$, then, $u^*_{\delta,\varepsilon} \to u^*_\delta$ weakly in $W^{1,p}(0,T;H) \cap L^m(0,T;X)$ and strongly in $C([0,T];H)$, where u^*_δ is a minimizer of $\mathcal{I}_{\delta,0}$.

The following generalization of the previous result to ε-dependent time intervals and weight functions in the definition of the WED functionals is straightforward.

Corollary 2.3 (ε-dependent weights). *Let the assumption of Theorem 2.2 be satisfied. Moreover, let $T_\varepsilon \nearrow T < \infty$, and let $e_{\varepsilon,\delta} : [0, T_\varepsilon) \to [0, \infty)$ be a family of L^∞ functions uniformly convergent to $t \mapsto \mathrm{e}^{-t/\delta}$ for fixed $\delta > 0$. Then, the WED functionals*

$$\mathcal{I}_{\delta,\varepsilon}(u) = \begin{cases} \int_0^T e_{\varepsilon,\delta}(t)\chi_{[0,T_\varepsilon]}(t)\left(\psi_\varepsilon(\dot{u}(t)) + \frac{1}{\delta}\widetilde{\phi}_\varepsilon(t, u(t))\right)\mathrm{d}t & \text{if } u \in \mathcal{K}_{T_\varepsilon}(u_\varepsilon^0), \\ \infty & \text{else,} \end{cases}$$

where

$$\mathcal{K}_{T_\varepsilon}(u_\varepsilon^0) = \{u \in \mathcal{K}(u_\varepsilon^0) : u \text{ is constant in } [T_\varepsilon, T]\},$$

Γ*-converge to $\mathcal{I}_{\delta,0}$ defined in (2.13) in the weak topology of $W^{1,p}(0,T;H) \cap L^m(0,T;X)$.*

Since minimizers of the WED functionals solve elliptic-in-time regularized problems, namely the associated Euler–Lagrange equations, we immediately deduce evolutionary Γ-convergence in the sense of [17]. More precisely, we have the following.

Corollary 2.4 (Evolutionary Γ-convergence of the elliptic-regularized problems). *Let the assumptions of Theorem 2.2 be satisfied. Then minimizers of the WED functional $\mathcal{I}_{\delta,\varepsilon}$ solve*

$$\begin{aligned} -\delta\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{d}\psi_\varepsilon(\dot{u}) + \mathrm{d}\psi_\varepsilon(\dot{u}) + \partial\phi_\varepsilon(u) \ni g_\varepsilon(t) \quad & a.e. \text{ in } (0,T), \\ u(0) = u_\varepsilon^0, \qquad \delta\mathrm{d}\psi_\varepsilon(\dot{u}(T)) = 0, & \end{aligned} \tag{$\mathrm{P}_{\delta,\varepsilon}$}$$

and converge (up to subsequences) to solutions to

$$\begin{aligned} -\delta\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{d}\psi_0(\dot{u}) + \mathrm{d}\psi_0(\dot{u}) + \partial\phi_0(u) \ni g_0(t) \quad & a.e. \text{ in } (0,T), \\ u(0) = u^0, \qquad \delta\mathrm{d}\psi_0(\dot{u}(T)) = 0. & \end{aligned} \tag{$\mathrm{P}_{\delta,0}$}$$

Note that both equation ($\mathrm{P}_{\delta,\varepsilon}$) and ($\mathrm{P}_{\delta,0}$) admit in general nonunique solutions. Thus, evolutionary Γ-convergence has to be interpreted in the following sense. For all u_ε^0, there exists a solution u_ε to ($\mathrm{P}_{\delta,\varepsilon}$) such that $\{u_\varepsilon\}$ converges, up to subsequences, to a solution u to ($\mathrm{P}_{\delta,0}$). (Of course in Corollary 2.4 we choose $u_\varepsilon = u^*_{\delta,\varepsilon}$ and $u = u^*_{\delta,0}$, *i.e.* the minimizers of the WED functionals). Moreover, solutions to both ($\mathrm{P}_{\delta,\varepsilon}$) and ($\mathrm{P}_{\delta,0}$) are intended in the strong sense defined in Theorem 2.1.

2.3. The joint limit $\varepsilon + \delta \to 0$

The crucial question now is whether it is possible to consider the joint limit $\varepsilon + \delta \to 0$. The main result in this section states that minimizers $u^*_{\delta,\varepsilon}$ of the WED functional $\mathcal{I}_{\delta,\varepsilon}$ (which also solve ($\mathrm{P}_{\delta,\varepsilon}$)), converge (up to subsequences $(\delta_k, \varepsilon_k) \to (0,0)$) to solutions to the target problem (2.8). We remark that the limit $\delta + \varepsilon \to 0$ is meaningless in the sense of Γ-convergence of the WED functionals, since the Γ-limit of $\mathcal{I}_{\delta,\varepsilon}$ for $\delta \to 0$ is highly degenerate (see [18], Sect. 4.3), and provides only little information on the limiting solutions. More precisely, for every $\varepsilon > 0$ fixed the WED functional Γ-converges in the weak topology of $W^{1,p}(0,T;H) \cap L^m(0,T;X)$ to the functional $\mathcal{I}_{0,\varepsilon}$ defined by

$$\mathcal{I}_{0,\varepsilon}(u) = \begin{cases} 0 & \text{if } u \in \mathcal{K}(u^0), \\ \infty & \text{else.} \end{cases}$$

The following theorem combines the convergence results of Theorems 2.1 and 2.2.

Theorem 2.5. *With the same assumptions as in Theorem 2.2 let $u^*_{\varepsilon,\delta}$ denote a solution of ($\mathrm{P}_{\delta,\varepsilon}$) (e.g. a minimizer of the WED functional $\mathcal{I}_{\varepsilon,\delta}$). Assume additionally that for $\varepsilon \geq 0$ we have $g_\varepsilon = g^1_\varepsilon + g^2_\varepsilon$ with*

$$g^1_\varepsilon \in W^{1,p'}(0,T;H^*) \text{ and } g^1_\varepsilon \to g^1_0 \text{ weakly in } W^{1,p'}(0,T;H^*), \tag{2.14}$$

$$g^2_\varepsilon \to g^2_0 \text{ strongly in } L^{p'}(0,T;H^*), \tag{2.15}$$

and well preparedness of initial data, i.e.

$$\phi_\varepsilon(u^0_\varepsilon) \to \phi(u^0). \tag{2.16}$$

Then, there exists a subsequence $(\varepsilon_k, \delta_k)$ such that $\varepsilon_k + \delta_k \to 0$ and $u^*_{\varepsilon_k,\delta_k} \to u$, where u is a solution of the doubly nonlinear equation (2.8) (with $\psi = \psi_0$, $\phi = \phi_0$, and $g = g_0$).

Let us now briefly comment on the assumptions (2.14)–(2.16).

Remark 2.6. Note that the stronger assumptions in Theorem 2.5 are not needed in Theorem 2.2, *i.e.* they are not necessary for the limit passage $\varepsilon \to 0$ with $\delta > 0$ fixed. However, although the assumptions (2.14)–(2.16) refer to the convergence of the data for $\varepsilon \to 0$ they are fundamental for the limit $\delta + \varepsilon \to 0$.

Indeed, the limit $\delta + \varepsilon \to 0$ is computed at the level of the equations ($\mathrm{P}_{\delta,\varepsilon}$) and not for the functionals $\mathcal{I}_{\delta,\varepsilon}$. To compute this limit we first derive uniform estimates on each term in ($\mathrm{P}_{\delta,\varepsilon}$). Note that this can be done just with the assumptions of Theorem 2.1. These uniform estimates suffice to extract converging subsequences and to pass to the limit in each term of equation ($\mathrm{P}_{\delta,\varepsilon}$). The next step is to identify the limits of the nonlinear terms, in particular, to prove $\lim_{\varepsilon\to 0} \mathrm{d}\psi_\varepsilon(\dot{u}^*_{\delta,\varepsilon}) = \mathrm{d}\psi_0(\dot{u})$.

By taking advantage of the convexity of the dissipation potential ψ_0, we only have to prove an upper estimate on $\limsup_{\varepsilon\to 0} \left\langle \mathrm{d}\psi_\varepsilon(\dot{u}^*_{\delta,\varepsilon}), \dot{u}^*_{\delta,\varepsilon} \right\rangle_H$, see (3.25). Since our estimates provide only weak convergence of both $\mathrm{d}\psi_\varepsilon\left(\dot{u}^*_{\delta,\varepsilon}\right)$ and $\dot{u}^*_{\delta,\varepsilon}$, we substitute $\mathrm{d}\psi_\varepsilon\left(\dot{u}^*_{\delta,\varepsilon}\right)$ by using equation ($\mathrm{P}_{\delta,\varepsilon}$). As a consequence we have to handle the two terms $\left\langle \partial\phi_\varepsilon(u^*_{\delta,\varepsilon}), \dot{u}^*_{\delta,\varepsilon} \right\rangle_H$ and $\left\langle g_\varepsilon, \dot{u}^*_{\delta,\varepsilon} \right\rangle_H$. We can pass to the limit in the first by integrating by parts and using the well preparedness of initial data (2.16) and in the second by virtue of (2.14) and (2.15). Note that this procedure is necessary only in the case that the term $\mathrm{d}\psi_\varepsilon$ is nonlinear. In the case of quadratic dissipation potentials there is no need to additionally assume (2.14)–(2.16), and the statement of Theorem 2.5 holds true under the same assumptions of Theorem 2.2, in particular, without assuming well preparedness of initial data.

2.4. Convergence rates

Assuming some quantitative estimates on the rates of the Γ-convergence for the static functionals ϕ_ε and ψ_ε is it possible to derive quantitative estimates on the rate of Γ-convergence for the corresponding WED functionals? In particular, can we estimate the rate of convergence of the minimizers $\left\| u^*_{\delta,\varepsilon} - u^*_{\delta,0} \right\|$ in some norm?

We give a positive answer in the case of quadratic dissipation potentials ψ_ε and λ-convex energy functionals, since stronger estimates on the time derivative of the minimizer of the functional $u^*_{\delta,\varepsilon}$ are available, which provide H^2 regularity (*cf.* [19]). We recall that additional time regularity has been proven also in the non-quadratic case in [2] for a specific nonlinear gradient flow driven by a p-Laplace operator. However, in the general nonquadratic case such results are, to the best of the authors' knowledge, not available. Thus, aiming at clarity, we restrict ourselves to the quadratic (and λ-convex) setting. Moreover, in the case of nonconvex energies minimizers of WED functionals are not unique. Thus, quantitative estimates on convergence rates seem out of reach in this case.

We assume that $p = 2$, H is a Hilbert space with scalar product $(\cdot,\cdot)_H$, and that for all $\varepsilon \geq 0$, ψ_ε is quadratic such that $\psi_\varepsilon(v) = \frac{1}{2}(\mathbb{A}_\varepsilon v, v)_H$ with $\mathbb{A}_\varepsilon \in \mathrm{Lin}(H,H)$ symmetric and positive definite. Moreover, we assume

$\varphi_\varepsilon^2 = 0$, but we relax the convexity assumption on ϕ_ε: here we assume ϕ_ε to be λ-convex, *i.e.*

$$\phi_\varepsilon(\cdot) - \frac{\lambda}{2}\left\|\cdot\right\|_H^2 \text{ is convex for all } \varepsilon \geq 0 \text{ and } \lambda \in \mathbb{R},$$

where λ is independent of ε.

Given a nonconvex functional $\phi : H \to \mathbb{R}$, we recall that its Fréchet subdifferential $\partial_H^{\mathrm{F}}\phi(u) \subset H$ is defined as

$$v \in \partial_H^{\mathrm{F}}\phi(u) \quad \text{iff } u \in D(\phi) \quad \text{and} \quad \liminf_{w\to u} \frac{\phi(w) - \phi(u) - (v, w-u)_H}{\|w-u\|_H} \geq 0.$$

In the case that ϕ is λ-convex, we can identify $\partial_H^{\mathrm{F}}\phi(u) = \partial_H\phi^\lambda(u) + \lambda u$, where $\phi^\lambda(u) = \phi(u) - \lambda\|u\|_H^2/2$ is convex. In particular, we have $D(\phi) = D(\phi^\lambda)$ and $D(\partial_H^{\mathrm{F}}\phi) = D(\partial_H\phi^\lambda)$. Thus, to simplify notation we will also write $\partial_H\phi$ for the Fréchet subdifferential of the λ-convex functional ϕ.

Let us assume that there exists a space B such that $H \subset B \subset X$, and a positive constant C and such that

$$\left\|u_{\delta,\varepsilon}^*\right\|_{H^1(0,T;B)} \leq C \text{ for all } \varepsilon \geq 0. \tag{2.17}$$

Note that in concrete applications the space B is typically an interpolation space between H and X (see, *e.g.* Sect. 4.1).

Let $\mathfrak{R}_\varepsilon : D(\phi_0) \to D(\phi_\varepsilon)$ be the recovery operator for the energy functionals $\phi_\varepsilon \xrightarrow{\mathrm{M}} \phi_0$ in H, namely

$$\forall\, u \in D(\phi_0): \quad \lim_{\varepsilon\to 0} \mathfrak{R}_\varepsilon(u) = u \quad \text{strongly in } H \qquad \text{and} \qquad \lim_{\varepsilon\to 0} \phi_\varepsilon(\mathfrak{R}_\varepsilon(u)) = \phi_0(u).$$

We assume there exists $r_\varepsilon^{\mathfrak{R}}$, $r_\varepsilon^{\mathfrak{R},\phi}$, $r_\varepsilon^{\mathfrak{R},\psi} > 0$ with $\lim_\varepsilon r_\varepsilon^{\mathfrak{R}} = \lim_\varepsilon r_\varepsilon^{\mathfrak{R},\phi} = \lim_\varepsilon r_\varepsilon^{\mathfrak{R},\psi} = 0$, a non-decreasing function $\ell : [0,\infty) \to [0,\infty)$, and a positive constant C, such that

(R1)
$$\|\mathfrak{R}_\varepsilon v - v\|_H \leq r_\varepsilon^{\mathfrak{R}} \left(\ell(\|v\|_H) + \|v\|_X^m\right)^{1/2} \text{ for all } v \in X, \tag{2.18}$$

(R2)
$$\phi_\varepsilon(\mathfrak{R}_\varepsilon v) - \frac{\lambda}{2}\|\mathfrak{R}_\varepsilon v\|_H^2 \leq \phi_0(v) - \frac{\lambda}{2}\|v\|_H^2 + r_\varepsilon^{\mathfrak{R},\phi}\left(\ell(\|v\|_H) + \|v\|_X^m + \|\eta\|_H^2\right)$$
$$\text{for all } v \in D(\phi_0) \text{ and } \eta \in \partial_H(\phi_0(v)), \tag{2.19}$$

(R3)
$$|\psi_\varepsilon(v_\varepsilon) - \psi_0(v)| \leq C(\|v\|_H + \|v_\varepsilon\|_H)\|v_\varepsilon - v\|_H + r_\varepsilon^{\mathfrak{R},\psi}\|v\|_B^2$$
$$\text{for all } v_\varepsilon \in H \text{ and } v \in B. \tag{2.20}$$

Furthermore, we assume that there exists an operator $\mathfrak{S}_\varepsilon : D(\phi_\varepsilon) \to D(\phi_0)$, rates $r_\varepsilon^{\mathfrak{S}}$, $r_\varepsilon^{\mathfrak{S},\phi}$, $r_\varepsilon^{\mathfrak{S},\psi} > 0$ with $\lim_\varepsilon r_\varepsilon^{\mathfrak{S}} = \lim_\varepsilon r_\varepsilon^{\mathfrak{S},\phi} = \lim_\varepsilon r_\varepsilon^{\mathfrak{S},\psi} = 0$, a nondecreasing function, again denoted by $\ell : [0,\infty) \to [0,\infty)$, and a positive constant C, such that

(R4)
$$\|\mathfrak{S}_\varepsilon v_\varepsilon - v_\varepsilon\|_H \leq r_\varepsilon^{\mathfrak{S}}\left(\ell(\|v_\varepsilon\|_H) + \|v_\varepsilon\|_X^m\right)^{1/2} \text{ for all } v_\varepsilon \in D(\phi_\varepsilon). \tag{2.21}$$

(R5)
$$\phi_0(\mathfrak{S}_\varepsilon v_\varepsilon) - \frac{\lambda}{2}\|\mathfrak{S}_\varepsilon v_\varepsilon\|_H^2 \leq \phi_\varepsilon(v_\varepsilon) - \frac{\lambda}{2}\|v_\varepsilon\|_H^2 + r_\varepsilon^{\mathfrak{S},\phi}\left(\ell(\|v_\varepsilon\|_H) + \|v_\varepsilon\|_X^m + \|\eta_\varepsilon\|_H^2\right)$$
$$\text{for all } v_\varepsilon \in D(\phi_\varepsilon) \text{ and } \eta_\varepsilon \in \partial_H\phi_\varepsilon(v_\varepsilon), \tag{2.22}$$

(R6)
$$|\psi_\varepsilon(v_\varepsilon) - \psi_0(v)| \leq C(\|v\|_H + \|v_\varepsilon\|_H)\|v_\varepsilon - v\|_H + r_\varepsilon^{\mathfrak{S},\psi}\|v_\varepsilon\|_B^2$$
$$\text{for all } v_\varepsilon \in B \text{ and } v \in H. \tag{2.23}$$

Finally, we assume that $g_\varepsilon \in H^1(0,T;H)$ and that there exist $C>0$ and rates r_ε^g, $r_\varepsilon^0 > 0$ with $\lim_\varepsilon r_\varepsilon^g = \lim_\varepsilon r_\varepsilon^0 = 0$ such that

$$\textbf{(R7)} \quad \|g_\varepsilon - g_0\|_{L^2(0,T;H)} \le r_\varepsilon^g, \qquad \text{and} \qquad \textbf{(R8)} \quad \left\|u_\varepsilon^0 - u^0\right\|_H \le r_\varepsilon^0. \tag{2.24}$$

In Section 4.1 we will discuss a typical example for the developed theory, namely, the homogenization of a parabolic equation. Since the operator $\mathfrak{S}_\varepsilon$ will be used to smooth out oscillations coming from the microstructure, we will call it *smoothing operator* from now on.

Theorem 2.7 (Convergence rates)**.** *Let **(R1)**–**(R8)** be satisfied and let $u^*_{\delta,\varepsilon}$ and $u^*_{\delta,0}$ be the minimizers of $\mathcal{I}_{\delta,\varepsilon}$ and $\mathcal{I}_{\delta,0}$ respectively. Then, there exists a constant C depending on δ and on the problems data, but independent on ε, such that*

$$\left\|u^*_{\delta,\varepsilon} - u^*_{\delta,0}\right\|_{L^2(0,T;H)} \le C\left(\rho_\varepsilon^1 + \left(\rho_\varepsilon^2 + \rho_\varepsilon^3\right)^{1/2}\right), \tag{2.25}$$

where

$$\begin{aligned}
\rho_\varepsilon^1 &= r_\varepsilon^{\mathfrak{R}} + \rho^{\mathrm{dtc}}_{\tau_\varepsilon^{\mathfrak{R}}} + \left(\tau_\varepsilon^{\mathfrak{R}}\right)^{1/2} r_\varepsilon^0,\\
\rho_\varepsilon^2 &= r_\varepsilon^{\mathfrak{R}} + \rho^{\mathrm{dtc}}_{\tau_\varepsilon^{\mathfrak{R}}} + r_\varepsilon^0 + r_\varepsilon^{\mathfrak{R},\phi} + r_\varepsilon^{\mathfrak{R},\psi} + \tau_\varepsilon^{\mathfrak{R}} + r_\varepsilon^g + r_\varepsilon^{\mathfrak{R}}/\tau_\varepsilon^{\mathfrak{R}} + r_\varepsilon^0,\\
\rho_\varepsilon^3 &= r_\varepsilon^{\mathfrak{S}} + \rho^{\mathrm{dtc}}_{\tau_\varepsilon^{\mathfrak{S}}} + r_\varepsilon^0 + r_\varepsilon^{\mathfrak{S},\phi} + r_\varepsilon^{\mathfrak{S},\psi} + \tau_\varepsilon^{\mathfrak{S}} + r_\varepsilon^g + r_\varepsilon^{\mathfrak{S}}/\tau_\varepsilon^{\mathfrak{S}} + r_\varepsilon^0,
\end{aligned}$$

and

$$\tau_\varepsilon^{\mathfrak{R}} = \arg\min_\tau (r_\varepsilon^{\mathfrak{R}}/\tau + \tau^{1/2} + \tau) \quad \textit{and} \quad \tau_\varepsilon^{\mathfrak{S}} = \arg\min_\tau (r_\varepsilon^{\mathfrak{S}}/\tau + \tau^{1/2} + \tau).$$

Note that, analogously to Theorem 2.2 (where δ is fixed), Theorem 2.7 does not require well preparedness of initial data.

We recall that in the case $\varepsilon > 0$ fixed some estimates on the convergence rates for the limit $\delta \to 0$ have been proven in [19], provided some assumptions on the convergence of initial data are satisfied. However, we remark that these results can not be combined with Theorem 2.7 to obtain estimates on the convergence rates in the case $\delta + \varepsilon \to 0$. This is a consequence of the fact that the constant C in (2.25) depends on δ and can not be guaranteed to be uniformly bounded in δ: In order to prove Theorem 2.7 we use Lemma 3.3 which provides quantitative estimates on the rate of convergence of the minimizers of Γ-converging functionals. One crucial assumption is coercivity condition in (3.29) of such functionals. However, condition (3.29) is satisfied by the WED functionals with a constant $C = C_\delta$ vanishing as $\delta \to 0$ (see Sect. 3.3). Indeed, due to the presence of the exponential weight in the definition of the WED functional, coercivity of $\mathcal{I}_{\delta,\varepsilon}$ cannot be uniform with respect to the parameter δ.

3. Proof of the main results

In this section we collect the proofs of the main results of Section 2.

3.1. Proof of the Γ-convergence of the WED functionals

As usual, we divide the proof of the Γ-convergence into two parts: first, we prove the liminf estimate and then the existence of recovery sequences. Since $\delta > 0$ is fixed throughout this subsection, we will omit it in the indices of the WED functionals, etc.

Proposition 3.1. *Let $u_\varepsilon \to u$ weakly in $W^{1,p}(0,T;H) \cap L^m(0,T;X)$, then ,*

$$\liminf_{\varepsilon\to 0} \mathcal{I}_\varepsilon(u_\varepsilon) \geq \mathcal{I}_0(u).$$

Proof. Note that, as a consequence of the Mosco and continuous convergence of φ^1_ε and φ^2_ε, respectively, in (2.9), and of the coercivity condition in (2.2b), we have the Mosco convergence $\phi_\varepsilon \xrightarrow{\mathrm{M}} \phi$ in H. Consider now a sequence satisfying $u_\varepsilon \to u$ weakly in $W^{1,p}(0,T;H) \cap L^m(0,T;X)$. In particular, by standard embedding results we also have

$$u_\varepsilon \to u \text{ strongly in } C([0,T];H). \tag{3.1}$$

Thus, from the Mosco convergence of ϕ_ε and the Fatou Lemma it follows that

$$\liminf_{\varepsilon\to 0} \int_0^T \mathrm{e}^{-t/\delta}\phi_\varepsilon(u_\varepsilon(t))\,\mathrm{d}t \geq \int_0^T \mathrm{e}^{-t/\delta}\phi_0(u(t))\,\mathrm{d}t,$$

see also [25]. Moreover, thanks to (2.11) and to the strong convergence (3.1), we have

$$\lim_{\varepsilon\to 0} \int_0^T \mathrm{e}^{-t/\delta}\left\langle g_\varepsilon(t), u_\varepsilon(t)\right\rangle_H \,\mathrm{d}t = \int_0^T \mathrm{e}^{-t/\delta}\left\langle g_0(t), u(t)\right\rangle_H \,\mathrm{d}t.$$

We focus now on the dissipation part. Fix $N \in \mathbb{N}$, $N \geq 2$ and set $\tau := T/N$. By using convexity of ψ_ε, Jensen's inequality yields

$$\begin{aligned}\liminf_{\varepsilon\to 0} \int_0^T \mathrm{e}^{-t/\delta}\psi_\varepsilon(\dot u_\varepsilon)\,\mathrm{d}t &\geq \liminf_{\varepsilon\to 0} \sum_{i=1}^N \mathrm{e}^{-i\tau/\delta}\tau \int_{(i-1)\tau}^{i\tau} \psi_\varepsilon(\dot u_\varepsilon)\frac{\mathrm{d}t}{\tau} \\ &\geq \liminf_{\varepsilon\to 0} \sum_{i=1}^N \mathrm{e}^{-i\tau/\delta}\tau\psi_\varepsilon\left(\int_{(i-1)\tau}^{i\tau} \dot u_\varepsilon \frac{\mathrm{d}t}{\tau}\right),\end{aligned} \tag{3.2}$$

where we used $\mathrm{e}^{-i\tau/\delta} \leq \mathrm{e}^{-t/\delta}$ for $t \leq i\tau$ in the estimate. Thus, exploiting the continuous convergence of ψ_ε in (2.10) and to (3.1) we can pass to the limit $\varepsilon \to 0$ to arrive at

$$\lim_{\varepsilon\to 0} \sum_{i=1}^N \mathrm{e}^{-i\tau/\delta}\tau\psi_\varepsilon\left(\frac{u_\varepsilon(i\tau) - u_\varepsilon((i-1)\tau)}{\tau}\right) = \sum_{i=1}^N \mathrm{e}^{-i\tau/\delta}\tau\psi_0\left(\frac{u(i\tau) - u((i-1)\tau)}{\tau}\right). \tag{3.3}$$

Let us denote by $\widehat{u}_\tau$ the piecewise affine interpolant of the nodes $\{u(i\tau)\}_{k=0}^N$ (compare (3.4)). We easily check that $\frac{\mathrm{d}}{\mathrm{d}t}\widehat{u}_\tau \to \dot u$ weakly in $L^p(0,T;H)$ for $\tau \to 0$. Moreover, for $(i-1)\tau \leq t \leq i\tau$ we can estimate $\exp(t - i\tau) \geq \exp(-\tau)$ hence

$$\sum_{i=1}^N \mathrm{e}^{-i\tau/\delta}\tau\psi_0\left(\frac{u(i\tau) - u((i-1)\tau)}{\tau}\right)\mathrm{d}t \geq \mathrm{e}^{-\tau/\delta}\int_0^T \mathrm{e}^{-t/\delta}\psi_0\left(\frac{\mathrm{d}}{\mathrm{d}t}\widehat{u}_\tau\right)\mathrm{d}t \quad \text{for all } \tau.$$

Since ψ_0 is lower semicontinuous and convex, we can pass to the limit $\tau \to 0$ to arrive with (3.2) and (3.3) at the liminf inequality. □

Next, we construct recovery sequences for $\mathcal{I}_{\varepsilon,\delta}$. Here, we use a density argument.

Proposition 3.2. *For every $u \in \mathcal{K}(u^0)$ there exists a subsequence $\varepsilon_k \to 0$ and a sequence $u_{\varepsilon_k} \to u$ strongly in $W^{1,p}(0,T;H)$ and weakly in $L^m(0,T;X)$ such that*

$$\lim_{\varepsilon_k \to 0} \mathcal{I}_{\varepsilon_k}(u_{\varepsilon_k}) = \mathcal{I}_0(u).$$

Proof. Our proof follows the lines of Theorem 7.2 of [3]. Let us start by assuming that $u \in \mathcal{K}(u^0) \cap C^1([0,T];X)$ (recall that $u^0 \in X$). We fix $N \in \mathbb{N}$, set $\tau = T/N$, and define the nodal values $u_\tau^i = u(i\tau) \in X$, for $i = 0, ..., N$. Let $\widehat{u}_\tau$, and $\overline{u}_\tau$ be the piecewise affine and piecewise forward constant interpolants of $\{u_\tau^i\}_{i=0}^N$ defined *via*

$$\left.\begin{aligned} \widehat{u}_\tau(t) &= \alpha_\tau^i(t) u_\tau^i + (1-\alpha_\tau^i(t)) u_\tau^{i+1} \\ \overline{u}_\tau(t) &= u_\tau^{i+1} \end{aligned}\right\} \quad \text{for } t \in [i\tau, (i+1)\tau)\,, \tag{3.4}$$

where $\alpha_\tau^i(t) = 1 - (i\tau - t)/\tau$.

Since $u \in C^1([0,T];X)$ it follows that

$$\widehat{u}_\tau \to u \text{ strongly in } W^{1,\infty}(0,T;H), \tag{3.5a}$$

$$\overline{u}_\tau \to u \text{ strongly in } L^\infty(0,T;X). \tag{3.5b}$$

In particular, by the continuity properties of the functionals ϕ_0 and ψ_0 we immediately obtain the convergence of the limiting WED functional

$$\mathcal{I}_0(\widehat{u}_\tau) \to \mathcal{I}_0(u) \text{ for } \tau \to 0 \tag{3.6}$$

and (see [3], Sect. 6.1.2)

$$\mathcal{I}_0(\widehat{u}_\tau) = \sum_{i=0}^{N-1} \int_{i\tau}^{(i+1)\tau} \mathrm{e}^{-t/\delta} \left(\psi_0 \left(\frac{u_\tau^{i+1} - u_\tau^i}{\tau} \right) + \frac{1}{\delta} \phi_0(u_\tau^{i+1}) - \frac{1}{\delta} \langle g_0(t), \widehat{u}_\tau(t) \rangle_H \right) \mathrm{d}t + h_\tau^1,$$

$$\text{where} \quad h_\tau^1 := \int_0^T \frac{\mathrm{e}^{-t/\delta}}{\delta} \big(\phi_0(\widehat{u}_\tau) - \phi_0(\overline{u}_\tau) \big) \, \mathrm{d}t.$$

Indeed, using the convexity of φ_0^1 and choosing $\eta_\tau^{1,i} \in \partial_X \varphi_X^1(u_\tau^i)$ and $\eta_\tau^{2,i} \in \partial_X \varphi_X^2(u_\tau^i)$ we arrive at the estimate

$$\begin{aligned} h_\tau^1 &\le \sum_{i=0}^{N-1} \int_{i\tau}^{(i+1)\tau} \frac{\mathrm{e}^{-t/\delta}}{\delta} \Big(\alpha_\tau^i(t) (\varphi_0^1(u_\tau^i) - \varphi_0^1(u_\tau^{i+1})) + \varphi_0^2(u_\tau^{i+1}) - \varphi_0^2(\widehat{u}_\tau(t)) \Big) \, \mathrm{d}t \\ &\le C \sum_{i=0}^{N-1} \tau \big(\|\eta_\tau^{1,i}\|_{X^*} + \|\eta_\tau^{2,i+1}\|_{X^*} \big) \|u_\tau^{i+1} - u_\tau^i\|_X. \end{aligned}$$

Hereafter, the symbol C will denote a positive constant independent of τ and ε. By using (2.3a) and (2.4b), and convergence (3.5b), we estimate

$$\|\eta_\tau^{1,i}\|_{X^*} + \|\eta_\tau^{2,i+1}\|_{X^*} \le C.$$

Thus, we obtain $h_\tau^1 \to 0$ for $\tau \to 0$.

We now build a recovery sequence for the WED functional $\mathcal{I}_{\varepsilon,\delta}$ and $u \in \mathcal{K}(u^0) \cap C^1([0,T];X)$. The strategy is to discretize the time interval and to interpolate between the recovery sequences for the energy functional at each node. To unify notation, we define $u^0_{\tau,\varepsilon} := u^0_\varepsilon$. Due to (2.9) and the coercivity of ϕ_ε in X we find sequences $u^i_{\tau,\varepsilon}$ for every $i \in \{1,...,N\}$ such that

$$u^i_{\tau,\varepsilon} \to u^i_\tau \text{ weakly in } X \qquad \text{and} \qquad \varphi^1_\varepsilon(u^i_{\tau,\varepsilon}) \to \varphi^1_0(u^i_\tau).$$

Note that $u^i_{\tau,\varepsilon} \to u^i_\tau$ strongly in H. Thus, thanks to (2.9) and (2.10), we have

$$\psi_\varepsilon\left(\frac{u^i_{\tau,\varepsilon} - u^{i-1}_{\tau,\varepsilon}}{\tau}\right) \to \psi_0\left(\frac{u^i_\tau - u^{i-1}_\tau}{\tau}\right) \text{ for all } i \in \{1,...,N\},$$
$$\varphi^2_\varepsilon\left(u^i_{\tau,\varepsilon}\right) \to \varphi^2_0\left(u^i_\tau\right) \text{ for all } i \in \{0,...,N\}.$$

Let $\widehat{u}_{\tau,\varepsilon}$ and $\overline{u}_{\tau,\varepsilon}$ be the correspondent piecewise affine and piecewise constant interpolants as in (3.4). Thus, for τ still fixed we obtain the convergences

$$\widehat{u}_{\tau,\varepsilon} \to \widehat{u}_\tau \text{ weakly in } L^m(0,T;X), \tag{3.7}$$
$$\widehat{u}_{\tau,\varepsilon} \to \widehat{u}_\tau \text{ strongly in } W^{1,p}(0,T;H), \tag{3.8}$$
$$\overline{u}_{\tau,\varepsilon} \to \overline{u}_\tau \text{ weakly in } L^m(0,T;X). \tag{3.9}$$

We now claim that the following upper estimate holds

$$\mathcal{I}_\varepsilon(\widehat{u}_{\tau,\varepsilon}) \le \mathcal{I}_0(\widehat{u}_\tau) + r^1_{\tau,\varepsilon} + r^2_\tau, \tag{3.10}$$

where

$$\lim_{\tau\to 0}\lim_{\varepsilon\to 0} r^1_{\tau,\varepsilon} = 0 \quad \text{and} \quad \lim_{\tau\to 0} r^2_\tau = 0.$$

Indeed, we compute

$$\begin{aligned}
\mathcal{I}_\varepsilon(\widehat{u}_{\tau,\varepsilon}) &= \sum_{i=0}^{N-1}\int_{i\tau}^{(i+1)\tau} \mathrm{e}^{-t/\delta}\left(\psi_\varepsilon\left(\frac{u^{i+1}_{\tau,\varepsilon} - u^i_{\tau,\varepsilon}}{\tau}\right) + \frac{1}{\delta}\phi_\varepsilon(u^{i+1}_{\tau,\varepsilon}) - \frac{1}{\delta}\left\langle g_\varepsilon, \widehat{u}_{\tau,\varepsilon}\right\rangle_H\right)\mathrm{d}t \\
&\quad + \frac{1}{\delta}\sum_{i=0}^{N-1}\int_{i\tau}^{(i+1)\tau} \mathrm{e}^{-t/\delta}\left(\varphi^1_\varepsilon\left(\alpha^i_\tau(t)u^i_{\tau,\varepsilon} + (1-\alpha^i_\tau(t))u^{i+1}_{\tau,\varepsilon}\right) - \varphi^1_\varepsilon(u^{i+1}_{\tau,\varepsilon})\right)\mathrm{d}t \\
&\quad - \frac{1}{\delta}\sum_{i=0}^{N-1}\int_{i\tau}^{(i+1)\tau} \mathrm{e}^{-t/\delta}\left(\varphi^2_\varepsilon\left(\widehat{u}_{\tau,\varepsilon}\right) - \varphi^2_\varepsilon(u^{i+1}_{\tau,\varepsilon})\right)\mathrm{d}t =: I_{\tau,\varepsilon} + h^2_{\tau,\varepsilon} + h^3_{\tau,\varepsilon}.
\end{aligned}$$

Clearly, we have $I_{\tau,\varepsilon} \to \mathcal{I}_0(\widehat{u}_\tau) + h^1_\tau$ for $\varepsilon \to 0$. Moreover, using the convexity of φ^1_ε and the definition of the sequences $u^i_{\tau,\varepsilon}$ we arrive at

$$\begin{aligned}
\lim_{\varepsilon\to 0} h^2_{\tau,\varepsilon} &\le \lim_{\varepsilon\to 0}\sum_{i=1}^{N-1}\int_{i\tau}^{(i+1)\tau}\frac{\mathrm{e}^{-t/\delta}}{\delta}\alpha^i_\tau(t)\left(\varphi^1_\varepsilon\left(u^i_{\tau,\varepsilon}\right) - \varphi^1_\varepsilon(u^{i+1}_{\tau,\varepsilon})\right)\mathrm{d}t \\
&\quad + \lim_{\varepsilon\to 0}\int_0^\tau\left(\mathrm{e}^{-t/\delta}\frac{\tau - t}{\tau}\right)\left(\varphi^1_\varepsilon\left(u^0_\varepsilon\right) - \varphi^1_\varepsilon(u^1_{\tau,\varepsilon})\right)\mathrm{d}t \\
&= \sum_{i=1}^{N-1}\int_{i\tau}^{(i+1)\tau}\frac{\mathrm{e}^{-t/\delta}}{\delta}\alpha^i_\tau(t)\big(\varphi^1_0(u^i_\tau) - \varphi^1_0(u^{i+1}_\tau)\big)\,\mathrm{d}t + \lim_{\varepsilon\to 0} h^{2,2}_{\tau,\varepsilon}.
\end{aligned}$$

Note that we have to treat the case $i = 0$ separately since u_ε^0 is not a recovery sequence.

Choosing $\eta_\tau^i \in \partial_X \varphi_0^1(u_\tau^i)$ we can argue as above to obtain

$$\lim_{\varepsilon\to 0} h_{\tau,\varepsilon}^2 \le \sum_{i=0}^{N-1} C\tau \left\|\eta_\tau^i\right\|_{X^*} \left\|u_\tau^i - u_\tau^{i+1}\right\|_X + \lim_{\varepsilon\to 0} h_{\tau,\varepsilon}^{2,2},$$

whose right-hand side tends to zero as $\tau \to 0$. Moreover, for $i = 0$, thanks to (2.12) and to the fact that $\varphi_0^1(\overline{u}_\tau) \in L^1(0,T)$, we get

$$h_{\tau,\varepsilon}^{2,2} = \int_0^\tau \left(\mathrm{e}^{-t/\delta}\frac{\tau - t}{\tau}\right)\left(\varphi_\varepsilon^1\left(u_\varepsilon^0\right) - \varphi_\varepsilon^1(u_{\tau,\varepsilon}^1)\right)\,\mathrm{d}t \le C\tau.$$

Here we used (2.3a) and strong convergence (3.5b).

It remains to show that also $h_{\tau,\varepsilon}^3 \to 0$ as $\varepsilon \to 0$ and $\tau \to 0$. For this, note that for all $t \in [0,T]$ one has $\widehat{u}_{\tau,\varepsilon}(t) \to \widehat{u}_\tau(t)$ weakly in X. Thus, as a consequence of the continuous convergence in (2.9) we have $\lim_{\varepsilon\to 0}\varphi_\varepsilon^2\left(\widehat{u}_{\tau,\varepsilon}(t)\right) = \varphi_0^2\left(\widehat{u}_\tau(t)\right)$. Moreover, since $t \mapsto \varphi_\varepsilon^2(\widehat{u}_{\tau,\varepsilon}(t))$ are uniformly integrable due to (2.4a), we can apply the Vitali convergence Theorem ([22], p. 133) to deduce

$$\lim_{\varepsilon\to 0} h_{\tau,\varepsilon}^3 = \int_0^T \frac{\mathrm{e}^{-t/\delta}}{\delta}\left(\varphi_0^2\left(\widehat{u}_\tau\right) - \varphi_0^2(\overline{u}_\tau)\right)\,\mathrm{d}t.$$

With the convergences in (3.5a) and (3.5b) as well as estimate (2.4b), we conclude that the right-hand side vanishes as $\tau \to 0$.

Finally, the convergences (3.8) and (2.11) give us

$$\lim_{\varepsilon\to 0}\int_0^T \mathrm{e}^{-t/\delta}\,\langle g_\varepsilon, \widehat{u}_{\tau,\varepsilon}\rangle_H\,\mathrm{d}t = \int_0^T \mathrm{e}^{-t/\delta}\,\langle g_0, \widehat{u}_\tau\rangle_H\,\mathrm{d}t,$$

which proves the claim in (3.10).

Note additionally that the sequence $\widehat{u}_{\tau,\varepsilon}$ is bounded in $L^m(0,T;X)$ independently of τ and ε. Indeed, without loss of generality, we can assume that $u_{\tau,\varepsilon}^i$ and u_τ^i are such that

$$\varphi_\varepsilon^1(u_{\tau,\varepsilon}^i) \le \varphi_0^1(u_\tau^i) + 1 = \varphi_0^1(u(i\tau)) + 1 \text{ for all } i = \{0, ..., N\}.$$

As the functionals φ_ε^1 and φ_0^1 are uniformly coercive in X, we have that

$$\left\|u_{\tau,\varepsilon}^i\right\|_X^m \le C\left(\varphi_\varepsilon^1(u_{\tau,\varepsilon}^i) + 1\right) \le C\left(\varphi_0^1(u(i\tau)) + 2\right)$$

and hence, by convexity,

$$\begin{aligned}\|\widehat{u}_{\tau,\varepsilon}\|_{L^m(0,T;X)}^m &\le \sum_{i=0}^{N-1}\int_{i\tau}^{(i+1)\tau}\left\{\alpha_\tau^i(t)\left\|u_{\tau,\varepsilon}^i\right\|_X^m + \left(1-\alpha_\tau^i(t)\right)\left\|u_{\tau,\varepsilon}^{i+1}\right\|_X^m\right\}\mathrm{d}t\\ &\le C\sum_{i=0}^{N-1}\int_{i\tau}^{(i+1)\tau}\left\{\alpha_\tau^i(t)\varphi_0^1(u(i\tau)) + \left(1-\alpha_\tau^i(t)\right)\varphi_0^1(u((i+1)\,\tau))\right\}\mathrm{d}t + 2CT.\end{aligned}$$

Therefore, without loss of generality, we can assume

$$\|\widehat{u}_{\tau,\varepsilon}\|^m_{L^m(0,T;X)} \leq \widetilde{C}\Big(\int_0^T \varphi_0^1(u)\,\mathrm{d}t + T\Big) \tag{3.11}$$

for some constant $\widetilde{C}$ independent of τ, ε, and u.

We now show that we can extract subsequences ε_k and τ_k such that the sequence defined by $u_{\varepsilon_k} := \widehat{u}_{\varepsilon_k,\tau_k}$ satisfies the statement of the proposition. Indeed, note that

$$\|\widehat{u}_{\tau,\varepsilon} - u\|_{W^{1,p}(0,T;H)} \leq \|\widehat{u}_{\tau,\varepsilon} - \widehat{u}_\tau\|_{W^{1,p}(0,T;H)} + \|\widehat{u}_\tau - u\|_{W^{1,p}(0,T;H)}\,, \tag{3.12}$$

and for all for all $\zeta \in L^{m'}(0,T;X^*)$ we find

$$|\,\langle \zeta, \widehat{u}_{\tau,\varepsilon} - u\rangle_{L^m(0,T;X)}\,| \leq |\,\langle \zeta, \widehat{u}_{\tau,\varepsilon} - \widehat{u}_\tau\rangle_{L^m(0,T;X)}\,| + |\,\langle \zeta, \widehat{u}_\tau - u\rangle_{L^m(0,T;X)}\,|. \tag{3.13}$$

Let $\{\zeta_j\}_{j=0}^\infty$ be a dense subset of $L^{m'}(0,T;X^*)$. Thanks to the convergences in (3.7)–(3.9) and estimate (3.10), we can choose $\varepsilon = \varepsilon_\tau$ such that

$$\|\widehat{u}_{\tau,\varepsilon_\tau} - \widehat{u}_\tau\|_{W^{1,p}(0,T;H)} \leq \tau \quad \text{and} \quad |\,\langle \zeta_j, \widehat{u}_{\tau,\varepsilon_\tau} - \widehat{u}_\tau\rangle_{L^m(0,T;X)}\,| \leq \tau \text{ for all } j \leq \frac{1}{\tau}. \tag{3.14}$$

Moreover, we can assume that $r^1_{\tau,\varepsilon_\tau} \leq \tau$ with $r^1_{\tau,\varepsilon}$ from (3.10). By using convergences (3.5a) and (3.5b) and estimates (3.12) and (3.13), we obtain that

$$\lim_{\tau\to 0} \|\widehat{u}_{\tau,\varepsilon_\tau} - u\|_{W^{1,p}(0,T;H)} = 0 \quad \text{and} \quad \limsup_{\tau\to 0} \mathcal{I}_{\varepsilon_\tau}(\widehat{u}_{\tau,\varepsilon_\tau}) \leq \mathcal{I}_0(u).$$

In particular, $\widehat{u}_{\tau,\varepsilon_\tau} \to u$ strongly in $W^{1,p}(0,T;H)$.

It remains to show that also $\widehat{u}_{\tau,\varepsilon_\tau} \to u$ weakly in $L^m(0,T;X)$. Given $\zeta \in L^{m'}(0,T;X^*)$, for all $\alpha > 0$ there exists $j \in \mathbb{N}$ such that $\|\zeta - \zeta_j\|_{L^{m'}(0,T;X^*)} \leq \alpha$ and $\tau = \tau(j)$ with $j \leq 1/\tau$, $\tau \leq \alpha$ and satisfying

$$\big|\langle \zeta_l, \widehat{u}_\tau - u\rangle_{L^m(0,T;X)}\big| \leq \alpha \text{ for all } l \leq \frac{1}{\tau}.$$

Thus, by using (3.14) and (3.11), we have that

$$\begin{aligned} |\,\langle \zeta, \widehat{u}_{\tau,\varepsilon_\tau} - u\rangle_{L^m(0,T;X)}\,| &= |\,\langle \zeta - \zeta_j, \widehat{u}_{\tau,\varepsilon_\tau} - u\rangle_{L^m(0,T;X)}\,| \\ &\quad + |\,\langle \zeta_j, \widehat{u}_{\tau,\varepsilon_\tau} - \widehat{u}_\tau\rangle_{L^m(0,T;X)}\,| + |\,\langle \zeta_j, \widehat{u}_\tau - u\rangle_{L^m(0,T;X)}\,| \\ &\leq C\alpha + \alpha + \alpha. \end{aligned}$$

In particular, $\widehat{u}_{\tau,\varepsilon_\tau} \to u$ weakly in $L^m(0,T;X)$. This proves the existence of a recovery sequence under the additional assumption $u \in C^1([0,T];X)$.

Let us now consider the general case with $u \in W^{1,p}(0,T;H) \cap L^m(0,T;X)$. We can find an approximating sequence $u_n \in C^1([0,T],X)$ with

$$u_n \to u \text{ strongly in } W^{1,p}(0,T;H) \cap L^m(0,T;X).$$

Hence, $\mathcal{I}_0(u_n) \to \mathcal{I}(u)$. For every n let $u_{n,\varepsilon}$ be a recovery sequence for u_n as constructed in the first step, *i.e.* $u_{n,\varepsilon} \to u_n$ strongly in $W^{1,p}(0,T;H)$ and weakly in $L^m(0,T;X)$ and such that $\limsup_{\epsilon\to 0} \mathcal{I}_\varepsilon(u_{n,\varepsilon}) \leq \mathcal{I}_0(u_n)$.

Moreover, we can assume that $u_{n,\varepsilon}$ satisfies the additional requirement

$$\|u_{n,\varepsilon}\|^m_{L^m(0,T;X)} \le \widetilde{C}(\varphi^1(u) + T + 1).$$

Thus, it is possible to apply Lemma A.1 and extract a subsequence u_{ε_k,n_k} such that $u_{n_k,\varepsilon_k} \to u$ strongly in $W^{1,p}(0,T;H)$ and weakly in $L^m(0,T;X)$ and such that the lim sup-inequality is satisfied. As a consequence of the liminf-estimate obtained in Proposition 3.1, we conclude $\lim_{\varepsilon_k\to 0} \mathcal{I}_{\varepsilon_k}(u_{n_k,\varepsilon_k}) = \mathcal{I}_0(u)$. ☐

Theorem 2.2 can be now proved by simply combining the two propositions above, by using the coercivity of the WED functionals on the space $W^{1,p}(0,T;H) \cap L^m(0,T;X)$ and by recalling the compact embedding of the latter in $C([0,T];H)$.

Proof of Corollary 2.3. For all $u \in \mathcal{K}_{T_\varepsilon}(u^0_\varepsilon)$ we decompose

$$\begin{aligned}\mathcal{I}_\varepsilon(u) &= \int_0^T e_{\varepsilon,\delta}(t)\chi_{[0,T_\varepsilon]}(t)\left(\psi_\varepsilon(\dot{u}) + \frac{1}{\delta}\widetilde{\phi}_\varepsilon(t,u)\right)\,\mathrm{d}t \\ &= \int_0^T \mathrm{e}^{-t/\delta}\left(\psi_\varepsilon(\dot{u}) + \frac{1}{\delta}\widetilde{\phi}_\varepsilon(t,u)\right)\,\mathrm{d}t - \int_{T_\varepsilon}^T \mathrm{e}^{-t/\delta}\left(\psi_\varepsilon(0) + \frac{1}{\delta}\widetilde{\phi}_\varepsilon(t,u)\right)\,\mathrm{d}t \\ &\quad + \int_0^{T_\varepsilon}\left(e_{\varepsilon,\delta}(t) - \mathrm{e}^{-t/\delta}\right)\left(\psi_\varepsilon(\dot{u}) + \frac{1}{\delta}\widetilde{\phi}_\varepsilon(t,u)\right)\,\mathrm{d}t. \end{aligned} \tag{3.15}$$

Note that, since $t \mapsto \psi_\varepsilon(\dot{u}(t)) + \frac{1}{\delta}\widetilde{\phi}_\varepsilon(t,u(t))$ is uniformly bounded in $L^1(0,T)$, thanks to the uniform convergence of $e_{\varepsilon,\delta}$ to $t \mapsto \mathrm{e}^{-t/\delta}$ and to the convergence of T_ε, we have that the second and the third terms in (3.15) vanish as $\varepsilon \to 0$ and independently of u as long as u lies in a bounded subset of $W^{1,p}(0,T;H) \cap L^m(0,T;X)$. By applying Theorem 2.2 we have that the first term Γ-converges to $\mathcal{I}_0$. By recalling the equicoercivity of the functionals $\mathcal{I}_\varepsilon$ in $W^{1,p}(0,T;H) \cap L^m(0,T;X)$ we then conclude the proof of Corollary 2.3. ☐

3.2. Proof of the joint limit

Let $u_{\varepsilon,\delta}$ be any solution to

$$\begin{gathered} -\delta\dot{\xi}_{\varepsilon,\delta} + \xi_{\varepsilon,\delta} + \eta^1_{\varepsilon,\delta} - \eta^2_{\varepsilon,\delta} = g_\varepsilon \text{ in } X^* \text{ a.e. in } (0,T), \\ \text{with } \delta\xi_{\varepsilon,\delta}(T) = 0 \quad \text{and} \quad u_{\varepsilon,\delta}(0) = u^0_\varepsilon, \end{gathered} \tag{3.16}$$

where $\xi_{\varepsilon,\delta} = \mathrm{d}\psi_\varepsilon(\dot{u}_{\varepsilon,\delta}) \in L^{p'}(0,T;H^*)$, $\eta^1_{\varepsilon,\delta} \in L^{m'}(0,T;X^*)$, $\eta^2_{\varepsilon,\delta} \in L^{p'}(0,T;H^*)$ with $\eta^1_{\varepsilon,\delta}(t) \in \partial_X\varphi^1_\varepsilon(u_{\varepsilon,\delta}(t))$, and $\eta^2_{\varepsilon,\delta}(t) \in \partial_H\varphi^2_\varepsilon(u_{\varepsilon,\delta}(t))$ for almost every $t \in (0,T)$. We recall that solutions to this problem exist. Indeed, as stated in Theorem 2.1 (see also [1, 3]), every minimizer of the WED functional $\mathcal{I}_{\varepsilon,\delta}$, solves (3.16).

In order to prove Theorem 2.5, we want to pass to the limit $\delta + \varepsilon \to 0$ in (3.16). To this aim we recall that $u_{\varepsilon,\delta}$ satisfies the following uniform estimate (*cf.* [1], Sect. 4.4, see also [3] for the convex energy case)

$$\begin{aligned}\left\|\xi_{\varepsilon,\delta}\right\|_{L^{p'}(0,T;H^*)} + \left\|u_{\varepsilon,\delta}\right\|_{W^{1,p}(0,T;H)\cap L^m(0,T;X)} + \left\|\eta^1_{\varepsilon,\delta}\right\|_{L^{m'}(0,T;X^*)} \\ + \left\|\eta^2_{\varepsilon,\delta}\right\|_{L^{p'}(0,T;H^*)} + \left\|\xi_{\varepsilon,\delta}\right\|_{C([0,T];X^*)} + \left\|\delta\dot{\xi}_{\varepsilon,\delta}\right\|_{L^{p'}(0,T;H^*)+L^{m'}(0,T;X^*)} \le C, \end{aligned} \tag{3.17}$$

which implies, up to not-relabeled subsequences, the following convergence results for the joint limit $\delta + \varepsilon \to 0$

$$u_{\varepsilon,\delta} \to u \quad \text{weakly in } L^m(0,T;X) \cap W^{1,p}(0,T;H), \tag{3.18a}$$

$$u_{\varepsilon,\delta} \to u \quad \text{strongly in } C\left([0,T];H\right), \tag{3.18b}$$

$$\xi_{\varepsilon,\delta} \to \xi \quad \text{weakly in } L^{p'}(0,T;H^*), \tag{3.18c}$$
$$\delta\xi_{\varepsilon,\delta} \to 0 \quad \text{strongly in } L^{p'}(0,T;H^*), \tag{3.18d}$$
$$\eta^1_{\varepsilon,\delta} \to \eta^1 \text{ weakly in } L^{m'}(0,T;X^*), \tag{3.18e}$$
$$\eta^2_{\varepsilon,\delta} \to \eta^2 \text{ weakly in } L^{p'}(0,T;H^*), \tag{3.18f}$$
$$\delta\dot{\xi}_{\varepsilon,\delta} \to 0 \quad \text{weakly in } L^{p'}(0,T;H^*) + L^{m'}(0,T;X^*). \tag{3.18g}$$

for some limits $\eta^1 \in L^{m'}(0,T;X^*)$, $\eta^2 \in L^{p'}(0,T;H^*)$ and $\xi \in L^{p'}(0,T;H^*)$. Furthermore, we recall that $g_\varepsilon \to g_0$ weakly in $L^{p'}(0,T;H^*)$ by assumption (2.11). Thus, we can pass to the limit $\delta+\varepsilon \to 0$ to obtain the limiting equation

$$\xi + \eta^1 - \eta^2 = g_0 \text{ in } X^* \text{ a.e. in } (0,T), \tag{3.19}$$
$$u(0) = u^0. \tag{3.20}$$

Note that, since η^2, g_0, and ξ belong to the space $L^{p'}(0,T;H^*)$, then, by comparison in (3.19) we also have $\eta^1 \in L^{p'}(0,T;H^*)$. In particular, equation (3.19) can be equivalently rewritten in the stronger form

$$\xi + \eta^1 - \eta^2 = g_0 \text{ in } H^* \text{ a.e. in } (0,T).$$

It remains to identify the limits ξ, η^1, and η^2.

3.2.1. Identification of $\eta^2 \in \partial_H \varphi^2_0(u)$

Let $w \in D(\varphi^2_0)$ and let $\rho \in C^\infty_c(0,T)$, $\rho \ge 0$. By the definition of the subdifferential, we have

$$\int_0^T \varphi^2_\varepsilon(w)\rho(t)\mathrm{d}t + \int_0^T \left\langle \eta^2_{\varepsilon,\delta}(t), u_{\varepsilon,\delta}(t) - w \right\rangle_H \rho(t)\mathrm{d}t \ge \int_0^T \varphi^2_\varepsilon(u_{\varepsilon,\delta}(t))\rho(t)\mathrm{d}t. \tag{3.21}$$

As a consequence of convergence (3.18b) and (3.18f), we get

$$\lim_{\varepsilon\to 0} \int_0^T \left\langle \eta^2_{\varepsilon,\delta}(t), u_{\varepsilon,\delta}(t) - w \right\rangle_H \rho(t)\mathrm{d}t = \int_0^T \left\langle \eta^2(t), u(t) - w \right\rangle_H \rho(t)\mathrm{d}t.$$

Thanks to continuous convergence of φ^2_ε in (2.9) and the convergence in (3.18b), we have

$$\int_0^T \varphi^2_\varepsilon(w)\rho(t)\mathrm{d}t \to \int_0^T \varphi^2_0(w)\rho(t)\mathrm{d}t \quad \text{and} \int_0^T \varphi^2_\varepsilon(u_{\varepsilon,\delta}(t))\rho(t)\mathrm{d}t \to \int_0^T \varphi^2_0(u(t))\rho(t)\mathrm{d}t.$$

Thus, passing to the limit in (3.21), we obtain

$$\int_0^T \varphi^2_0(w)\rho(t)\mathrm{d}t + \int_0^T \left\langle \eta^2(t), u(t) - w \right\rangle_H \rho(t)\mathrm{d}t \ge \int_0^T \varphi^2_0(u(t))\rho(t)\mathrm{d}t.$$

As ρ and w are arbitrary, we deduce $\eta^2(t) \in \partial_H \varphi^2_0(u(t))$ for a.a. $t \in (0,T)$.

3.2.2. Identification of $\eta^1 \in \partial_X \varphi^1_X(u)$

As before, let $w \in D(\varphi^1_0)$ and let $\rho \in C^\infty_c(0,T)$ with $\rho \geq 0$. Let now $\{w_\varepsilon\} \subset X$ be a recovery sequence for φ^1_0 at w, *i.e.*

$$w_\varepsilon \to w \text{ weakly in } X \text{ (strongly in } H) \tag{3.22}$$

and $\varphi^1_\varepsilon(w_\varepsilon) \to \varphi^1_0(w)$. By the definition of the subdifferential, we have

$$\int_0^T \varphi^1_\varepsilon(w_\varepsilon)\rho(t)\mathrm{d}t + \int_0^T \left\langle \eta^1_{\varepsilon,\delta}(t), u_{\varepsilon,\delta}(t) - w_\varepsilon \right\rangle_X \rho(t)\mathrm{d}t \geq \int_0^T \varphi^1_\varepsilon(u_{\varepsilon,\delta}(t))\rho(t)\mathrm{d}t. \tag{3.23}$$

As w_ε is a recovery sequence, we have convergence

$$\int_0^T \varphi^1_\varepsilon(w_\varepsilon)\rho(t)\mathrm{d}t \to \int_0^T \varphi^1_0(w)\rho(t)\mathrm{d}t.$$

Testing equation (3.16) with $(u_{\varepsilon,\delta} - w_\varepsilon)\rho$ leads to

$$\begin{aligned}
&\int_0^T \left\langle \eta^1_{\varepsilon,\delta}(t), u_{\varepsilon,\delta}(t) - w_\varepsilon \right\rangle_X \rho(t)\mathrm{d}t \\
&= \int_0^T \left\langle \eta^2_{\varepsilon,\delta}(t), u_{\varepsilon,\delta}(t) - w_\varepsilon \right\rangle_H \rho(t)\mathrm{d}t + \int_0^T \left\langle g_\varepsilon(t), u_{\varepsilon,\delta}(t) - w_\varepsilon \right\rangle_H \rho(t)\mathrm{d}t \\
&\quad + \int_0^T \left\langle \delta\xi'_{\varepsilon,\delta}(t), u_{\varepsilon,\delta}(t) - w_\varepsilon \right\rangle_X \rho(t)\mathrm{d}t - \int_0^T \left\langle \xi_{\varepsilon,\delta}(t), u_{\varepsilon,\delta}(t) - w_\varepsilon \right\rangle_H \rho(t)\mathrm{d}t.
\end{aligned} \tag{3.24}$$

Note that, as a consequence of convergences (3.18b), (3.18f), (3.22), we obtain

$$\int_0^T \left\langle \eta^2_{\varepsilon,\delta}(t), u_{\varepsilon,\delta}(t) - w_\varepsilon \right\rangle_H \rho(t)\mathrm{d}t \to \int_0^T \left\langle \eta^2(t), u(t) - w \right\rangle_H \rho(t)\mathrm{d}t.$$

Integration by parts, together with the final condition for $\xi_{\varepsilon,\delta}$ in (3.16) and estimate (3.17) yield (see [3] for a rigorous derivation of the integration by parts formula)

$$\begin{aligned}
&\int_0^T \left\langle \delta\xi'_{\varepsilon,\delta}(t), u_{\varepsilon,\delta}(t) - w_\varepsilon \right\rangle_X \rho(t)\mathrm{d}t \\
&= -\delta \int_0^T \left\langle \xi_{\varepsilon,\delta}(t), \dot{u}_{\varepsilon,\delta}(t) \right\rangle_H \rho(t)\mathrm{d}t - \delta \int_0^T \left\langle \xi_{\varepsilon,\delta}(t), u_{\varepsilon,\delta}(t) - w_\varepsilon \right\rangle_H \dot{\rho}(t)\mathrm{d}t \\
&\leq \delta \|\rho\|_{W^{1,\infty}} \|\xi_{\varepsilon,\delta}\|_{L^{p'}(0,T;H^*)} \Big(\|\dot{u}_{\varepsilon,\delta}\|_{L^p(0,T;H)} + \|u_{\varepsilon,\delta} - w_\varepsilon\|_{L^p(0,T;H)} \Big) \to 0.
\end{aligned}$$

Furthermore, thanks to assumption (2.11) and strong convergences (3.18b) and (3.22),

$$\int_0^T \left\langle g_\varepsilon(t), u_{\varepsilon,\delta}(t) - w_\varepsilon \right\rangle_H \rho(t)\mathrm{d}t \to \int_0^T \left\langle g_0(t), u(t) - w \right\rangle_H \rho(t)\mathrm{d}t.$$

Finally, from (3.18b), (3.18c), (3.22) it follows that

$$\int_0^T \langle \xi_{\varepsilon,\delta}(t), u_{\varepsilon,\delta}(t) - w_\varepsilon \rangle_H \rho(t)\mathrm{d}t \to \int_0^T \langle \xi(t), u(t) - w \rangle_H \rho(t)\mathrm{d}t.$$

Substituting into (3.24) and using identity (3.19), one gets

$$\begin{aligned}\int_0^T \left\langle \eta^1_{\varepsilon,\delta}(t), u_{\varepsilon,\delta}(t) - w_\varepsilon \right\rangle_X \rho(t)\mathrm{d}t &\to \int_0^T \left\langle \eta^2(t) + \xi(t) + g_0(t), u(t) - w \right\rangle_H \rho(t)\mathrm{d}t \\ &= \int_0^T \left\langle \eta^1(t), u(t) - w \right\rangle_X \rho(t)\mathrm{d}t.\end{aligned}$$

The Mosco convergence of φ^1_ε, (3.18b), and the Fatou Lemma yield

$$\liminf_{\delta+\varepsilon\to 0} \int_0^T \varphi^1_\varepsilon(u_{\varepsilon,\delta}(t))\rho(t)\mathrm{d}t \geq \int_0^T \varphi^1_0(u(t))\rho(t)\mathrm{d}t.$$

Combining all these facts from (3.23) we deduce

$$\int_0^T \varphi^1_0(w)\rho(t)\mathrm{d}t + \int_0^T \left\langle \eta^1(t), u(t) - w \right\rangle_X \rho(t)\mathrm{d}t \geq \int_0^T \varphi^1_0(u(t))\rho(t)\mathrm{d}t.$$

Hence, $\eta^1 \in \partial_X \varphi^1_0(u)$ a.e. in $(0,T)$. Moreover, since $\eta^1(t) \in H^*$ for a.a. $t \in (0,T)$, then $\eta^1(t) \in \partial_H \varphi^1_0(u(t))$ for a.a. $t \in (0,T)$ (*cf.* [3] for details).

3.2.3. Identification of $\xi = \mathrm{d}_H \psi_0(\dot{u})$

Let $v \in L^p(0,T;H)$. Again by the definition of the subdifferential, we have

$$\int_0^T \psi_\varepsilon(v(t))\mathrm{d}t + \int_0^T \langle \xi_{\varepsilon,\delta}(t), \dot{u}_{\varepsilon,\delta}(t) - v(t) \rangle_H \mathrm{d}t \geq \int_0^T \psi_\varepsilon(\dot{u}_{\varepsilon,\delta}(t))\mathrm{d}t. \tag{3.25}$$

Note that, as $\psi_\varepsilon \xrightarrow{C} \psi_0$ strongly in H, we have

$$\int_0^T \psi_\varepsilon(v(t))\mathrm{d}t \to \int_0^T \psi_0(v)\mathrm{d}t.$$

Arguing as in Proposition 3.1, the following inequality follows from Jensen's inequality

$$\liminf_{\varepsilon+\delta\to 0} \int_0^T \psi_\varepsilon(\dot{u}_{\varepsilon,\delta}(t))\mathrm{d}t \geq \int_0^T \psi_0(\dot{u}(t))\mathrm{d}t.$$

The weak convergence in (3.18c) implies

$$\int_0^T \langle \xi_{\varepsilon,\delta}(t), v(t) \rangle_H \mathrm{d}t \to \int_0^T \langle \xi(t), v(t) \rangle_H \mathrm{d}t.$$

Using the Euler–Lagrange equation (3.16), we obtain

$$\limsup_{\varepsilon+\delta\to 0}\int_0^T \langle \xi_{\varepsilon,\delta}(t), \dot u_{\varepsilon,\delta}(t)\rangle_H \,\mathrm{d}t \le \limsup_{\varepsilon+\delta\to 0}\int_0^T \Big\langle \delta\dot\xi_{\varepsilon,\delta}(t) - \eta^1_{\varepsilon,\delta}(t), \dot u_{\varepsilon,\delta}(t)\Big\rangle_H \,\mathrm{d}t$$
$$+ \limsup_{\varepsilon+\delta\to 0}\int_0^T \big\langle \eta^2_{\varepsilon,\delta}(t), \dot u_{\varepsilon,\delta}(t)\big\rangle_H + \langle g_\varepsilon(t), \dot u_{\varepsilon,\delta}(t)\rangle_H \,\mathrm{d}t.$$

Note that, as $\xi_{\varepsilon,\delta} = \mathrm{d}\psi_\varepsilon(\dot u_{\varepsilon,\delta}) \in \partial_H \psi_\varepsilon(\dot u_{\varepsilon,\delta})$, then $\dot u_{\varepsilon,\delta} \in \partial_{H^*}\psi^*_\varepsilon(\xi_{\varepsilon,\delta})$, where $\psi^*_\varepsilon : H^* \to [0,\infty)$ is the Fenchel conjugate of ψ_ε, *i.e.* $\psi^*_\varepsilon(\xi) = \sup_v\{\langle \xi, v\rangle - \psi_\varepsilon(v)\}$. Using the chain rule and recalling the final condition $\xi_{\varepsilon,\delta}(T) = 0$ (*cf.* [3] for a rigorous proof and more details), we get

$$\begin{aligned}\int_0^T \Big\langle \delta\dot\xi_{\varepsilon,\delta}(t) - \eta^1_{\varepsilon,\delta}(t), \dot u_{\varepsilon,\delta}(t)\Big\rangle_H \,\mathrm{d}t &\le \int_0^T \Big(\delta \frac{\mathrm{d}}{\mathrm{d}t}\psi^*_\varepsilon(\xi_{\varepsilon,\delta}(t)) - \frac{\mathrm{d}}{\mathrm{d}t}\varphi^1_\varepsilon(u_{\varepsilon,\delta}(t))\Big)\,\mathrm{d}t\\ &\le -\delta\psi^*_\varepsilon(\xi_{\varepsilon,\delta}(0)) - \varphi^1_\varepsilon(u_{\varepsilon,\delta}(T)) + \varphi^1_\varepsilon(u^0_\varepsilon)\\ &\le -\varphi^1_\varepsilon(u_{\varepsilon,\delta}(T)) + \varphi^1_\varepsilon(u^0_\varepsilon).\end{aligned}$$

By using the well preparedness of the initial data, the Mosco convergence of φ^1_ε and (3.18b), we can pass to the limit above to arrive at

$$\begin{aligned}\limsup_{\varepsilon+\delta\to 0}\int_0^T \Big\langle \delta\dot\xi_{\varepsilon,\delta}(t) - \eta^1_{\varepsilon,\delta}(t), \dot u_{\varepsilon,\delta}(t)\Big\rangle_H \,\mathrm{d}t &\le -\liminf_{\varepsilon+\delta\to 0}\varphi^1_\varepsilon(u_{\varepsilon,\delta}(T)) + \lim_{\varepsilon+\delta\to 0}\varphi^1_\varepsilon(u^0_\varepsilon)\\ &= -\varphi^1_0(u(T)) + \varphi^1_0(u^0).\end{aligned}$$

As already observed $\eta^1(t) \in H^*$ for a.a. $t \in (0,T)$, as a consequence of identity (3.19), thus, we obtain

$$-\varphi^1_0(u(T)) + \varphi^1_0(u^0) = \int_0^T \big\langle \eta^1(t), \dot u(t)\big\rangle_H \,\mathrm{d}t.$$

Thus, thanks to the assumption (2.9) and the convergence in (3.18b), we can use the chain rule for $t \mapsto \varphi^2_\varepsilon(u_{\varepsilon,\delta})$ to get

$$\begin{aligned}&\limsup_{\varepsilon+\delta\to 0}\int_0^T \big\langle \eta^2_{\varepsilon,\delta}(t), \dot u_{\varepsilon,\delta}(t)\big\rangle_H \,\mathrm{d}t\\ &\qquad = \lim_{\varepsilon+\delta\to 0}\varphi^2_\varepsilon(u_{\varepsilon,\delta}(T)) - \lim_{\varepsilon+\delta\to 0}\varphi^2_\varepsilon(u^0_\varepsilon) = \varphi^2_0(u(T)) - \varphi^2_0(u^0) = \int_0^T \big\langle \eta^2(t), \dot u(t)\big\rangle_H \,\mathrm{d}t.\end{aligned}$$

Finally, by using convergences (2.14)-(2.15) and (3.18a)-(3.18b), by integrating by parts, we get

$$\begin{aligned}&\int_0^T \langle g_\varepsilon(t), \dot u_{\varepsilon,\delta}(t)\rangle_H \,\mathrm{d}t\\ &= \big\langle g^1_\varepsilon(T), u_{\varepsilon,\delta}(T)\big\rangle_H - \big\langle g^1_\varepsilon(0), u^0_\varepsilon\big\rangle_H - \int_0^T \big\langle \dot g^1_\varepsilon(t), u_{\varepsilon,\delta}(t)\big\rangle_H \,\mathrm{d}t + \int_0^T \big\langle g^2_\varepsilon(t), \dot u_{\varepsilon,\delta}(t)\big\rangle_H \,\mathrm{d}t\\ &\quad \to \big\langle g^1_0(T), u(T)\big\rangle_H - \big\langle \dot g^1_0(0), u^0\big\rangle_H - \int_0^T \big\langle \dot g^1_0(t), u(t)\big\rangle_H \,\mathrm{d}t + \int_0^T \big\langle g^2_0(t), \dot u_0(t)\big\rangle_H \,\mathrm{d}t\\ &= \int_0^T \langle g_0(t), \dot u(t)\rangle_H \,\mathrm{d}t.\end{aligned}$$

Combining the above estimates and (3.25), we conclude

$$\int_0^T \psi_0(v(t))\mathrm{d}t + \int_0^T \langle \xi(t), \dot{u}(t) - v(t)\rangle_H \,\mathrm{d}t \geq \int_0^T \psi_0(\dot{u}(t))\mathrm{d}t,$$

i.e. $\xi \in \partial_{L^p(0,T;H)}\overline{\psi}_0(\dot{u})$, where $\overline{\psi}_0(v) = \int_0^T \psi_0(v(t))\mathrm{d}t$ for all $v \in L^p(0,T;H)$. As ψ_0 is Gâteaux differentiable in H, then $\overline{\psi}_0$ is Gâteaux differentiable in $L^p(0,T;H)$. Moreover,

$$\left\langle \mathrm{d}_{L^p(0,T;H)}\overline{\psi}_0(v), w\right\rangle_{L^p(0,T;H)} \leq \int_0^T \langle \mathrm{d}_H\psi_0(v(t)), w(t)\rangle_H \,\mathrm{d}t \text{ for all } v, w \in L^p(0,T;H).$$

In particular, $\xi \in \partial_{L^p(0,T;H)}\overline{\psi}_0(\dot{u})$ implies $\xi(t) = \mathrm{d}_H\psi_0(\dot{u}(t))$ for a.a. $t \in (0,T)$. This concludes the proof of Theorem 2.5.

3.3. Proof of the convergence rates

We now prove Theorem 2.7, *i.e.* quantitative estimates for the convergence of the WED minimizers. To this aim, we take advantage of the following abstract result for Γ-converging functionals.

Lemma 3.3. *Let $I_\varepsilon, I_0 : \mathcal{V} \to (-\infty, \infty]$ be functionals defined on some Banach space $\mathcal{V}$. Assume that each functional admits a unique minimizer, denoted by $u_\varepsilon^* = \operatorname{argmin} I_\varepsilon$, $u^* = \operatorname{argmin} I$, respectively. Let $I_\varepsilon \xrightarrow{\Gamma} I$ as $\varepsilon \to 0$ in the strong topology of $\mathcal{V}$. Finally, assume that there exist sequences $u_\varepsilon^R \in \mathcal{V}$ and $u_\varepsilon^S \in \mathcal{V}$ satisfying*

(i) *there exists $\rho_\varepsilon^1 > 0$, such that*

$$\|u_\varepsilon^R - u^*\|_{\mathcal{V}} \leq \rho_\varepsilon^1; \tag{3.26}$$

(ii) *there exists $\rho_\varepsilon^2 \in \mathbb{R}$ such that*

$$I_\varepsilon(u_\varepsilon^R) - I_0(u^*) \leq \rho_\varepsilon^2; \tag{3.27}$$

(iii) *there exist $\rho_\varepsilon^3 > 0$ such that*

$$I_0(u_\varepsilon^S) - I_\varepsilon(u_\varepsilon^*) \leq \rho_\varepsilon^3; \tag{3.28}$$

(iv) *there exist $C > 0$, $\Gamma > 1$, and $\rho_\varepsilon^4 > 0$ such that*

$$C\left\|u_\varepsilon^R - u_\varepsilon^*\right\|_{\mathcal{V}}^\Gamma \leq I_\varepsilon(u_\varepsilon^R) - I_\varepsilon(u_\varepsilon^*) + \rho_\varepsilon^4. \tag{3.29}$$

Then, $\|u_\varepsilon^* - u^*\|_{\mathcal{V}} \leq \rho_\varepsilon^1 + \overline{C}\left(\rho_\varepsilon^2 + \rho_\varepsilon^3 + \rho_\varepsilon^4\right)^{1/\Gamma}$ where $\overline{C} = (1/C)^{1/\Gamma}$.

Proof. We estimate

$$\begin{aligned}
\|u_\varepsilon^* - u^*\|_{\mathcal{V}} &\leq \left\|u_\varepsilon^R - u^*\right\|_{\mathcal{V}} + \left\|u_\varepsilon^* - u_\varepsilon^R\right\|_{\mathcal{V}} \\
&\underset{(3.26)+(3.29)}{\leq} \rho_\varepsilon^1 + \overline{C}\left(I_\varepsilon(u_\varepsilon^R) - I_\varepsilon(u_\varepsilon^*) + \rho_\varepsilon^4\right)^{1/\Gamma} \\
&= \rho_\varepsilon^1 + \overline{C}\left(I_\varepsilon(u_\varepsilon^R) - I(u^*) + I(u^*) - I_\varepsilon(u_\varepsilon^*) + \rho_\varepsilon^4\right)^{1/\Gamma}
\end{aligned}$$

$$\underset{(3.27)}{\leq} \rho_\varepsilon^1 + \overline{C}\left(\rho_\varepsilon^2 + I(u_\varepsilon^S) - I_\varepsilon(u_\varepsilon^*) + \rho_\varepsilon^4\right)^{1/\Gamma}$$
$$\underset{(3.28)}{\leq} \rho_\varepsilon^1 + \overline{C}\left(\rho_\varepsilon^2 + \rho_\varepsilon^3 + \rho_\varepsilon^4\right)^{1/\Gamma},$$

which completes the proof. □

Remark 3.4. Note that we need the sequences u_ε^R and u_ε^S to be defined just for the minimizers u^* and u_ε^* and not for every element in the domain of the functionals.

3.3.1. Time-discrete to time-continuous WED functionals

To prove the result in Theorem 2.7, we will use a time-discrete version of the WED functional. It is known that minimizers of the time-discrete WED functional (*i.e.* solutions to the time-discrete Euler–Lagrange equation) converge to their time-continuous counterpart as the discretization parameter τ goes to zero, see, *e.g.* Section 2.6 of [19]. In this subsection we estimate the rate of convergence of the time-discrete minimizers with respect to the time step parameter $\tau > 0$. This result constitutes a first step in the proof of Theorem 2.7. Recall that we restrict ourselves to the setting of a quadratic dissipation potential $\psi_\varepsilon(v) = \frac{1}{2}(\mathbb{A}_\varepsilon v, v)_H$ and of uniformly λ-convex energy functionals ϕ_ε as well as $g_\varepsilon \in H^1(0,T;H)$.

We fix $\delta > 0$, $T > 0$, as well as a time step $\tau = T/N$ for $N \in \mathbb{N}$. The parameter $\varepsilon > 0$ will be fixed in this subsection, hence, we omit the index.

The time-continuous WED functional $\mathcal{I}_\delta$ is defined as in (2.6) while its time-discrete counterpart $\mathcal{I}_\delta^\tau$: $\mathcal{K}_\tau(u^0) \to \mathbb{R} \cup \{+\infty\}$ is defined on $\mathcal{K}_\tau(u^0) := \{u^0\} \times H^N$ *via*

$$\begin{aligned}\mathcal{I}_\delta^\tau(u^0,\dots,u^N) &= \sum_{i=0}^{N-1} \rho_\tau^i \tau \left(\psi\left(\frac{u^{i+1}-u^i}{\tau}\right) + \frac{1}{\delta}\phi(u^i) - \frac{1}{\delta}(g^i,u^i)_H\right) \\ &= \int_0^T \overline{\rho}_\tau(t)\left(\psi(\widehat{u}_\tau') + \frac{1}{\delta}\phi(\overline{u}_\tau) - \frac{1}{\delta}(\overline{g}_\tau(t),\overline{u}_\tau)_H\right)\mathrm{d}t, \end{aligned} \tag{3.30}$$

where $\overline{u}_\tau$ and $\widehat{u}_\tau$ denote the piecewise constant and piecewise affine interpolants of the nodes $(u^0,\dots,u^N)$ (analogously for $\overline{g}_\tau$ and $\overline{\rho}_\tau$). The discrete weights are given *via* $\rho_\tau^i = (\delta/(\delta+\tau))^i$ and $(g_\tau^1,\dots,g_\tau^N) \in H^N$ is a suitable approximation of g so that $\|\overline{g}_\tau - g\|_{L^2(0,T;H)} \leq C\tau$. Note that $\mathcal{K}_\tau(u^0)$ is in bijective correspondence with the set

$$\big\{u \in \mathcal{K}(u^0) \,:\, u \text{ is piecewise affine with respect to } [i\tau,(i{+}1)\tau),\ i=0,\dots,N{-}1\big\}.$$

In particular, from now on, we will identify the tuple $(u^{*,0},\dots,u^{*,N})$ that minimizes the time-discrete WED functional with the piecewise affine interpolant, denoted by u_τ^{disc}. The former solves the discrete Euler–Lagrange equation

$$\begin{aligned} &-\delta\mathbb{A}\frac{v_\tau^{*,i+1} - v_\tau^{*,i}}{\tau} + \mathbb{A}v_\tau^{*,i} + \eta_\tau^{*,i} = g_\tau^i, \\ &\text{where}\quad v_\tau^{*,i} = \frac{u_\tau^{*,i+1} - u_\tau^{*,i}}{\tau} \quad\text{and}\quad \eta_\tau^{*,i} \in \partial_H\phi(u_\tau^{*,i}), \end{aligned} \tag{3.31}$$

where we additionally have the final condition $v_\tau^{*,N} = 0$ (see [19], Eq. 6.6). Moreover, arguing as in Section 6.2 of [19] we obtain the following estimate with a constant C independent of $\varepsilon \geq 0$ for the time-discrete solution

$$\delta^{1/2}\,\|\overline{w}_\tau^*\|_{L^2(0,T;H)} + \delta^{1/2}\,\|\overline{v}_\tau^*\|_{L^\infty(0,T;H)} + \|\overline{u}_\tau^*\|_{L^2(0,T;H)}$$

$$+ \left\|u_\tau^{\rm disc}\right\|_{H^1(0,T;H)} + \|\overline{\eta}_\tau^*\|_{L^2(0,T;H)} + \int_0^T \phi(\overline{u}_\tau^*)\,{\rm d}t \leq C, \quad (3.32)$$

where $\overline{w}_\tau^*$ is the piecewise constant interpolant of the second order difference quotient of $u_\tau^{*,i}$, *i.e.* $w_\tau^{*,i} = (v_\tau^{*,i+1} - v_\tau^{*,i})/\tau$. Note, that this estimate mimics the time-continuous estimate

$$\delta^{1/2}\,\|\partial_{tt}u^*\|_{L^2(0,T;H)} + \delta^{1/2}\,\|\partial_t u^*\|_{C([0,T];H)} + \|u^*\|_{H^1(0,T;H)} + \|\eta^*\|_{L^2(0,T;H)} + \int_0^T \phi(u)\,{\rm d}t \leq C, \quad (3.33)$$

(see [19], Sect. 4).

The following proposition gives a quantitative estimate for the time-discrete solutions.

Proposition 3.5. *Let $u_\tau^{\rm disc}$ and u^* be the minimizers of the time-discrete and -continuous WED functional $\mathcal{I}_\delta^\tau$ and $\mathcal{I}_\delta$, respectively. Then, we have*

$$\left\|u_\tau^{\rm disc} - u^*\right\|_{L^2(0,T;H)} \leq C\tau^{1/2} =: \rho_\tau^{\rm dtc}. \quad (3.34)$$

Proof. In what follows the symbol C will denote a constant independent of τ, δ, and ε possibly varying from line to line.

Without loss of generality, we can assume that $\mathbb{A} = {\rm Id}$. The general case follows from an analogous computation. We start by considering the difference between the time-discrete and time-continuous Euler–Lagrange equation, *i.e.* (3.31) and ($P_{\delta,\varepsilon}$), respectively,

$$-\delta\left(\overline{w}_\tau^* - \partial_{tt}u^*\right) + \left(\overline{v}_\tau^* - \partial_t u^*\right) + \left(\overline{\eta}_\tau^* - \eta^*\right) = \left(\overline{g}_\tau - g\right).$$

We define the auxiliary function $U_\tau^* := u_\tau^{\rm disc} - u^*$, use it as test function in the equation above, and obtain after integration over $[0,t]$ for $t = K\tau$ for some $K \in \{1,\dots,N\}$

$$-\,\delta\sum_{i=0}^{K-1}\int_{i\tau}^{(i+1)\tau}\left(\frac{v_\tau^{*,i+1} - v_\tau^{*,i}}{\tau}, U_\tau^*\right)_H {\rm d}s + \delta\int_0^t (\partial_{tt}u^*, U_\tau^*)_H\,{\rm d}s$$
$$+ \int_0^t \left(\overline{v}_\tau^* - \partial_t u^*, U_\tau^*\right)_H\,{\rm d}s + \int_0^t \left(\overline{\eta}_\tau^* - \eta^*, U_\tau^*\right)_H\,{\rm d}s = \int_0^t \left(\overline{g}_\tau - g, U_\tau^*\right)_H\,{\rm d}s. \quad (3.35)$$

For the first and second term we employ a discrete and continuous integration by parts formula (recall that $u_\tau^{\rm disc}(0) - u^*(0) = U_\tau^*(0) = 0$ and $\overline{v}_\tau^* = \partial_t u_\tau^{\rm disc}$) to arrive at

$$\frac{\delta}{\tau}\int_0^\tau \left(v_\tau^{*,0}, U_\tau^*\right)_H {\rm d}s - \frac{\delta}{\tau}\int_{(K-1)\tau}^{K\tau}\left(v_\tau^{*,K}, U_\tau^*\right)_H {\rm d}s + \delta\big(\partial_t u^*(t), U_\tau^*(t)\big)_H$$
$$+ \delta\sum_{i=1}^{K-1}\int_{i\tau}^{(i+1)\tau}\left(v_\tau^{*,i}, \frac{U_\tau^*(s+\tau) - U_\tau^*(s)}{\tau}\right)_H {\rm d}s - \delta\int_0^t (\partial_t u^*, \partial_t U_\tau^*)_H\,{\rm d}s + \frac{1}{2}\int_0^t \frac{\rm d}{{\rm d}t}\|U_\tau^*\|_H^2\,{\rm d}s$$
$$+ \int_0^t \left(\overline{\eta}_\tau^* - \eta^*, \overline{u}_\tau^{\rm disc} - u^*\right)_H {\rm d}s + \int_0^t \left(\overline{\eta}_\tau^* - \eta^*, u_\tau^{\rm disc} - \overline{u}_\tau^{\rm disc}\right)_H {\rm d}s \leq \int_0^t \|\overline{g}_\tau - g\|_H\,\|U_\tau^*\|_H\,{\rm d}s,$$

where $\overline{u}_\tau^{\rm disc}$ denotes the piecewise constant interpolant associated with the minimizer of the time-discrete WED functional. By monotonicity, the first term on the last line is nonnegative, viz. $\int_0^t \left(\overline{\eta}_\tau^* - \eta^*, \overline{u}_\tau^{\rm disc} - u^*\right)_H {\rm d}s \geq$

0. Moreover, thanks to the uniform bounds on $\overline{\eta}_\tau^*$, on η^* and on the time derivative of $u_\tau^{\rm disc}$, we have that $|\int_0^t \big(\overline{\eta}_\tau^* - \eta^*, u_\tau^{\rm disc} - \overline{u}_\tau^{\rm disc}\big)_H \,\mathrm{d}s| \le C\tau$. Furthermore, it is not hard to see that due to the estimate in (3.33) and (3.32) the first term on the first line vanishes for $\tau \to 0$ with an order of at least τ.

Next, we consider the second and third term on the first line, namely,

$$
\begin{aligned}
-\frac{\delta}{\tau}\int_{(K-1)\tau}^{K\tau} \big(v_\tau^{*,K}, U_\tau^*(s)\big)_H \,\mathrm{d}s + \delta\big(\partial_t u^*(t), U_\tau^*(t)\big)_H &= -\delta\big(\partial_t U_\tau^*(t), U_\tau^*(t)\big)_H - \delta\left(v_\tau^{*,K}, \frac{1}{\tau}\int_{t-\tau}^t U_\tau^*(s)\mathrm{d}s - U_\tau^*(t)\right)_H \\
&=: -\delta\big(\partial_t U_\tau^*(t), U_\tau^*(t)\big)_H + \Delta_\tau^*
\end{aligned}
$$

Here, thanks to the uniform bounds in (3.33) and (3.32), we can estimate

$$
\begin{aligned}
|\Delta_\tau^*| &\le \delta\,\|\overline{v}_\tau^*\|_{L^\infty(0,T;H)} \left\| \frac{1}{\tau}\int_{t-\tau}^t U_\tau^*(s)\mathrm{d}s - U_\tau^*(t)\right\|_H \\
&\le C\tau\delta^{1/2}(\|\overline{v}_\tau^*\|_{L^\infty(0,T;H)} + \|\partial_t u^*\|_{L^\infty(0,T;H)}) \le C\tau.
\end{aligned}
$$

Next, we treat the first and second term on the second line in the above estimate. We write

$$
\begin{aligned}
&\delta\sum_{i=1}^{K-1}\int_{i\tau}^{(i+1)\tau}\left(v_\tau^{*,i}, \frac{U_\tau^*(s+\tau) - U_\tau^*(s)}{\tau}\right)_H \mathrm{d}s - \delta\int_0^t (\partial_t u^*, \partial_t U_\tau^*)_H \,\mathrm{d}s \\
&= \delta\int_0^t \|\partial_t U_\tau^*\|_H^2 \,\mathrm{d}s + \delta\sum_{i=1}^{K-1}\int_{i\tau}^{(i+1)\tau}\left(\partial_t u_\tau^{\rm disc}, \frac{U_\tau^*(s+\tau) - U_\tau^*(s)}{\tau} - \partial_t U_\tau^*\right)_H \mathrm{d}s.
\end{aligned}
$$

We observe that the last term on the right-hand side is of order τ. Indeed, we have that

$$
\begin{aligned}
&\left|\delta\sum_{i=1}^{K-1}\int_{i\tau}^{(i+1)\tau}\left(\partial_t u_\tau^{\rm disc}, \frac{u^*(s+\tau) - u^*(s)}{\tau} - \partial_t u^*\right)_H \mathrm{d}s\right| \\
&\le \delta\,\|\overline{v}_\tau^*\|_{L^\infty(0,T;H)}\left(\sum_{i=1}^{K-1}\int_{i\tau}^{(i+1)\tau}\left\|\frac{u^*(s+\tau) - u^*(s)}{\tau} - \partial_t u^*\right\|_H^2 \mathrm{d}s\right)^{1/2} \\
&\le C\tau\delta^{1/2}\,\|u^*\|_{H^2(0,T;H)} \le C\tau
\end{aligned}
$$

and, by definition of $u_\tau^{\rm disc}$,

$$
\begin{aligned}
&\delta\sum_{i=1}^{K-1}\int_{i\tau}^{(i+1)\tau}\left(\partial_t u_\tau^{\rm disc}, \frac{u_\tau^{\rm disc}(s+\tau) - u_\tau^{\rm disc}(s)}{\tau} - \partial_t u_\tau^{\rm disc}\right)_H \mathrm{d}s \\
&= \delta\sum_{i=1}^{K-1}\int_{i\tau}^{(i+1)\tau}\left(v_\tau^{*,i}, \frac{1}{\tau}\left(v_\tau^{*,i+1}(s-i\tau) + u_\tau^{i+1} - v_\tau^{*,i}(s-i\tau) - u_\tau^i\right) - v_\tau^{*,i}\right)_H \mathrm{d}s \\
&= \delta\sum_{i=1}^{K-1}\int_{i\tau}^{(i+1)\tau}\left(v_\tau^{*,i}, \frac{(s-i\tau)}{\tau}v_\tau^{*,i+1} - v_\tau^{*,i}\right)_H \mathrm{d}s \\
&\le C\delta\tau\,\|\overline{v}_\tau^*\|_{L^\infty(0,T;H)}\,\|\overline{w}_\tau^*\|_{L^2(0,T;H)} \le C\tau.
\end{aligned}
$$

Finally, the right-hand side in the above estimate, containing the external forces, vanishes with an order of at least τ due to the assumptions made on g_τ^i.

Combining the above estimates, we have for all $t = K\tau, K \in \{1, ..., N\}$,

$$\delta \|\partial_t U_\tau^*\|^2_{L^2(0,t;H)} + \frac{1}{2}\|U_\tau^*(t)\|^2_H \le C\tau\big(1 + \|U_\tau^*\|_{L^2(0,t;H)}\big) + \delta\big(\partial_t U_\tau^*(t), U_\tau^*(t)\big)_H. \tag{3.36}$$

Note that the same formula holds true for every $t < T$ (after suitably modifying the constant C). Indeed, for every $t < T$, let K be the biggest integer smaller than t/τ. Thanks to the uniform estimates (3.32)-(3.33), each term in (3.36) evaluated at time t can be replaced by the same term at time $K\tau$ plus an error of order τ.

Starting from the above estimate and arguing as in Section 3.2 of [16], we can deduce $\|U_\tau^*\|_{L^2(0,T;H)} \le C\tau^{1/2}$. Substituting in (3.36) with $t = T$, and recalling that $\partial_t U_\tau^*(T) = 0$, we also have $\delta^{1/2}\|\partial_t U_\tau^*\|_{L^2(0,T;H)} = O(\tau^{1/2})$. □

3.3.2. Proof of the convergence rates

Since $\delta > 0$ is fixed throughout this section we will omit this index of the WED functionals and of their minimizer.

We start by recalling that in the case of λ-convex energies and quadratic dissipation potentials the functionals $\mathcal{I}_{\delta,\varepsilon}$ and $\mathcal{I}_{\delta,0}$ admit unique minimizers u_ε^* and u^* over $\mathcal{K}(u_\varepsilon^0)$ and $\mathcal{K}(u^0)$ respectively for all δ sufficiently small. Moreover, they satisfy the bounds [19]

$$\begin{aligned} &\|u_\varepsilon^*\|_{H^2(0,T;H)\cap L^m(0,T;X)} + \|\eta_\varepsilon^*\|_{L^2(0,T;H)} \le C_\delta, \\ \text{and } &\|u^*\|_{H^2(0,T;H)\cap L^m(0,T;X)} + \|\eta^*\|_{L^2(0,T;H)} \le C_\delta, \end{aligned} \tag{3.37}$$

for some C_δ depending on δ, but independent of ε, where η_ε^* and η^* are the selection of the subdifferentials $\partial_H \phi_\varepsilon(u_\varepsilon^*)$ and $\partial_H \phi_0(u^*)$, respectively, such that u_ε^* and u^* solve the respective Euler–Lagrange equations a.e. in $(0, T)$, see [19].

Our strategy to prove Theorem 2.7 consists in checking that $\mathcal{I}_\varepsilon$ and $\mathcal{I}_0$ satisfy the assumptions (3.26)–(3.29) of Lemma 3.3. To do this, we explicitly build the sequences u_ε^R and u_ε^S starting from the recovery and smoothing operators $\mathfrak{R}_\varepsilon$ and $\mathfrak{S}_\varepsilon$ as well as the minimizers u^* and u_ε^* of the time-discrete WED functionals $\mathcal{I}_0^\tau$ and $\mathcal{I}_\varepsilon^\tau$ defined above.

It is worth noting that although we need to build the sequences u_ε^R and u_ε^S in Lemma 3.3 just for the minimizer of the functionals, we however need informations on the convergence rates for the operators $\mathfrak{R}_\varepsilon$ and $\mathfrak{S}_\varepsilon$ for the static functionals at *any* point (*cf.* (R1)–(R6)). This is due to the fact that the minimizers of the static functionals are not related in any obvious way to the minimizers of the corresponding dynamic WED functionals.

Proof of (3.26) and (3.27). We start by additionally assuming ϕ_ε to be convex for all $\varepsilon \ge 0$, *i.e.* $\lambda = 0$. Let $u^* \in H^2(0, T; H) \cap L^m(0, T; X)$ be the minimizer of the limiting WED functional $\mathcal{I}_{\delta,0}$ over $\mathcal{K}(u^0)$. We fix a time step $\tau = T/N$, $N \in \mathbb{N}$ and define the nodal values $u_\tau^{*,i} = u^*(i\tau)$. As before, we consider the piecewise affine and piecewise constant interpolants $\widehat{u}_\tau^*$ and $\overline{u}_\tau^*$ for the nodes $\{u^*(i\tau)\}$. Let u_τ^{disc} denote the (piecewise affine interpolant of the) minimizer of the discrete WED functional $\mathcal{I}_{\delta,0}^\tau$.

We introduce the following sequence

$$u_{\tau,\varepsilon}^R(t) = \begin{cases} \alpha_\tau^0(t)u_\varepsilon^0 + (1-\alpha_\tau^0(t))\mathfrak{R}_\varepsilon u_\tau^{\mathrm{disc},1} & t \in [0,\tau), \\ \alpha_\tau^i(t)\mathfrak{R}_\varepsilon u_\tau^{\mathrm{disc},i} + (1-\alpha_\tau^i(t))\mathfrak{R}_\varepsilon u_\tau^{\mathrm{disc},i+1} & t \in [i\tau,(i+1)\tau), \quad i = 1,\dots,N-1, \end{cases}$$

where $\alpha_\tau^i(t) = 1 - (t - i\tau)/\tau$. □

Remark 3.6. The idea of the proof of (3.26) and (3.27) is to construct the recovering sequence $\mathcal{R}_\varepsilon u^*$ by using the good properties of the recovering operator $\mathfrak{R}_\varepsilon$ for the energy functional. The naive idea would be to define

$(\mathcal{R}_\varepsilon u^*)(t) = \mathfrak{R}_\varepsilon(u^*(t))$. However, note that this cannot be done rigorously, since $\mathfrak{R}_\varepsilon$ is defined only over X and u^* is not pointwise well defined in X (but just in H) (see also Sect. 4.1).

Moreover, since in general $\mathfrak{R}_\varepsilon$ is not differentiable in time, we would have no informations on the time derivative of $(\mathcal{R}_\varepsilon u^*)(t)$. Thus, a natural idea would be then to take a time discretization of u^*, where nodal values are average on small intervals (this way nodal values are well defined in X) compute $\mathfrak{R}_\varepsilon u^*$ on the nodal values and finally interpolate. This, would provide a good estimate in condition (3.26). However, in order to get some good estimates in (3.27) we need a control of $\mathcal{I}_\varepsilon(\mathcal{R}_\varepsilon u^*) - \mathcal{I}(u^*)$. If we define $\mathcal{R}_\varepsilon u^*$ in terms of the time piecewise affine interpolant $\widehat{u}^*_\tau$ of u^*, since we cannot estimate how good the approximation $\mathcal{I}(\widehat{u}^*_\tau)$ of $\mathcal{I}(u^*)$ is (more precisely, we have no control on the $\partial_H \phi(\widehat{u}^*_\tau)$ as it does not solve any equation) as estimates of $\mathcal{I}_\varepsilon(\mathcal{R}_\varepsilon u^*) - \mathcal{I}(u^*)$ seem hard to obtain.

Then, using the assumptions in (2.18), (2.24), and the bound in (3.34), we estimate

$$\begin{aligned}
&\frac{1}{2}\left\|u^* - u^R_{\tau,\varepsilon}\right\|^2_{L^2(0,T;H)} \le \left\|u^* - u^{\mathrm{disc}}_\tau\right\|^2_{L^2(0,T;H)} + \left\|u^{\mathrm{disc}}_\tau - u^R_{\tau,\varepsilon}\right\|^2_{L^2(0,T;H)}\\
&\le \left(\rho^{\mathrm{dtc}}_\tau\right)^2 + 2\sum_{i=1}^N \tau \left\|\mathfrak{R}_\varepsilon u^{\mathrm{disc},i}_\tau - u^{\mathrm{disc},i}_\tau\right\|^2_H + \tau\left\|u^0_\varepsilon - u^0\right\|^2_H\\
&\le \left(\rho^{\mathrm{dtc}}_\tau\right)^2 + 2\left(r^{\mathfrak{R}}_\varepsilon\right)^2 \sup_i \ell\big(\left\|u^{\mathrm{disc},i}_\tau\right\|_H\big) + 2\left(r^{\mathfrak{R}}_\varepsilon\right)^2 \sum_{i=1}^N \tau\left\|u^{\mathrm{disc},i}_\tau\right\|^m_X + \tau\left(r^0_\varepsilon\right)^2\\
&\le C\left(r^{\mathfrak{R}}_\varepsilon + \tau^{1/2} r^0_\varepsilon + \rho^{\mathrm{dtc}}_\tau\right)^2.
\end{aligned} \tag{3.38}$$

The rate ρ^{dtc}_τ is the rate of convergence of the discrete to continuous approximation, computed in the previous subsection (see (3.34)). We now decompose

$$\begin{aligned}
\mathcal{I}_\varepsilon(u^R_{\tau,\varepsilon}) - \mathcal{I}_0(u^*) &= \int_0^T \mathrm{e}^{-t/\delta}\left(\psi_\varepsilon\left(\partial_t u^R_{\tau,\varepsilon}\right) - \psi_0\left(\partial_t u^*\right)\right)\,\mathrm{d}t + \frac{1}{\delta}\int_0^T \mathrm{e}^{-t/\delta}\left(\phi_\varepsilon\left(u^R_{\tau,\varepsilon}\right) - \phi_0\left(u^*\right)\right)\,\mathrm{d}t\\
&\quad + \frac{1}{\delta}\int_0^T \mathrm{e}^{-t/\delta}\big((g_\varepsilon, u^R_{\tau,\varepsilon})_H - (g, u^*)_H\big)\,\mathrm{d}t =: \int_0^T \left[\Delta^\psi_{\tau,\varepsilon} + \Delta^\phi_{\tau,\varepsilon} + \Delta^g_{\tau,\varepsilon}\right]\mathrm{d}t.
\end{aligned}$$

We will treat each term on the right-hand side separately. Concerning $\Delta^\psi_{\tau,\varepsilon}$ we introduce the auxiliary variable $U_{\varepsilon,\tau}$ *via*

$$\partial_t u^R_{\tau,\varepsilon} = \frac{\mathfrak{R}_\varepsilon u^{\mathrm{disc},i+1} - \mathfrak{R}_\varepsilon u^{\mathrm{disc},i}}{\tau} = \partial_t u^{\mathrm{disc}}_\tau + U_{\varepsilon,\tau}.$$

Using assumption (2.18) for the recovery operator we obtain for all $t \in [0,\tau)$

$$\|U_{\varepsilon,\tau}(t)\|_H \le \frac{r^{\mathfrak{R}}_\varepsilon}{\tau}\left(\ell\left(\left\|u^{\mathrm{disc},1}_\tau\right\|_H\right) + \left\|u^{\mathrm{disc},1}_\tau\right\|^m_X\right)^{1/2} + \frac{r^0_\varepsilon}{\tau},$$

and for all $t \in [i\tau,(i+1)\tau)$, $i = 1,\dots,N-1$,

$$\|U_{\varepsilon,\tau}(t)\|_H \le \frac{r^{\mathfrak{R}}_\varepsilon}{\tau}\left(\ell\left(\left\|u^{\mathrm{disc},i+1}_\tau\right\|_H\right) + \ell\left(\left\|u^{\mathrm{disc},i}_\tau\right\|_H\right) + \left\|u^{\mathrm{disc},i+1}_\tau\right\|^m_X + \left\|u^{\mathrm{disc},i}_\tau\right\|^m_X\right)^{1/2}.$$

By using these estimates, as well as (2.20), (3.34), we get

$$\begin{aligned}\int_0^T \Delta^\psi_{\tau,\varepsilon}\,\mathrm{d}t &\le \int_0^T C\left(\left\|\partial_t u^R_{\tau,\varepsilon}\right\|_H + \|\partial_t u^*\|_H\right)\cdot\left\|\partial_t u^{\mathrm{disc}}_\tau - \partial_t u^* + U_{\tau,\varepsilon}\right\|_H + r^{\mathfrak{R},\psi}_\varepsilon \|\partial_t u^*\|^2_B\,\mathrm{d}t\\ &\le C\left(\frac{r^{\mathfrak{R}}_\varepsilon}{\tau} + r^{\mathfrak{R},\psi}_\varepsilon + \int_0^\tau \frac{r^0_\varepsilon}{\tau}\,\mathrm{d}t\right) + C\left\|\partial_t u^{\mathrm{disc}}_\tau - \partial_t u^*\right\|_{L^2(0,T;H)}\\ &\le C\left(\frac{r^{\mathfrak{R}}_\varepsilon}{\tau} + r^{\mathfrak{R},\psi}_\varepsilon + \rho^{\mathrm{dtc}}_\tau + r^0_\varepsilon\right).\end{aligned} \tag{3.39}$$

Here we used that $\partial_t u^*$ is bounded in $L^2(0,T;B)$ as a consequence of (2.17).

Next, we treat the energetic part. Note that as before, we have to treat the case $i=0$ differently. By using the convexity of ϕ_ε and assumption (2.19), we obtain

$$\begin{aligned}\int_\tau^T \Delta^\phi_{\tau,\varepsilon}\,\mathrm{d}t &\le C\sum_{i=1}^{N-1}\int_{i\tau}^{(i+1)\tau} \mathrm{e}^{-t/\delta}\Big[\alpha^i_\tau(t)\phi_\varepsilon(\mathfrak{R}_\varepsilon u^{\mathrm{disc},i}_\tau) + (1-\alpha^i_\tau(t))\phi_\varepsilon(\mathfrak{R}_\varepsilon u^{\mathrm{disc},i+1}_\tau) - \phi_0(u^*)\Big]\,\mathrm{d}t\\ &\le C r^{\mathfrak{R},\phi}_\varepsilon \sum_{i=1}^{N}\tau\left(\ell\left(\left\|u^{\mathrm{disc},i}_\tau\right\|_H\right) + \left\|u^{\mathrm{disc},i}_\tau\right\|^m_X + \left\|\eta^{\mathrm{disc},i}_\tau\right\|^2_H\right)\\ &\quad + C\sum_{i=1}^{N-1}\int_{i\tau}^{(i+1)\tau} \mathrm{e}^{-t/\delta}\Big[\alpha^i_\tau(t)\phi_0(u^{\mathrm{disc},i}_\tau) + (1-\alpha^i_\tau(t))\phi_0(u^{\mathrm{disc},i+1}_\tau) - \phi_0(u^*)\Big]\,\mathrm{d}t.\end{aligned}$$

The last term on the right-hand side can be estimated as follows: consider $\eta^{\mathrm{disc},i}_\tau \in \partial_H\phi_0(u^{\mathrm{disc},i}_\tau)$, then

$$\begin{aligned}\int_\tau^T \Delta^\phi_{\tau,\varepsilon}\,\mathrm{d}t &\le C r^{\mathfrak{R},\phi}_\varepsilon + C\sum_{i=1}^{N-1}\int_{i\tau}^{(i+1)\tau} \mathrm{e}^{-t/\delta}\Big[\phi_0(u^{\mathrm{disc},i+1}_\tau) - \phi_0(u^*) + \alpha^i_\tau(t)\left\|\eta^{\mathrm{disc},i}_\tau\right\|_H \left\|u^{\mathrm{disc},i}_\tau - u^{\mathrm{disc},i+1}_\tau\right\|_H\Big]\,\mathrm{d}t\\ &\le C\left(r^{\mathfrak{R},\phi}_\varepsilon + \tau\right) + C\sum_{i=0}^{N-1}\int_{i\tau}^{(i+1)\tau} \left\|\eta^{\mathrm{disc},i+1}_\tau\right\|_H \left\|u^{\mathrm{disc},i+1}_\tau - u^*\right\|_H\,\mathrm{d}t\\ &\le C\left(r^{\mathfrak{R},\phi}_\varepsilon + \tau\right) + C\sum_{i=1}^{N-1}\int_{i\tau}^{(i+1)\tau} \left\|\eta^{\mathrm{disc},i+1}_\tau\right\|_H \left(\left\|u^{\mathrm{disc}} - u^*\right\|_H + \left\|u^{\mathrm{disc},i+1}_\tau - u^{\mathrm{disc}}\right\|_H\right)\mathrm{d}t\\ &\le C\left(r^{\mathfrak{R},\phi}_\varepsilon + \tau\right) + C\left\|u^{\mathrm{disc}} - u^*\right\|_{L^2(0,T;H)}\\ &\le C\left(r^{\mathfrak{R},\phi}_\varepsilon + \tau + \rho^{\mathrm{dtc}}_\tau\right).\end{aligned}$$

Here, we used the estimates for $\eta^{\mathrm{disc},i}_\tau$ and the fact that $\sum_{i=0}^{N-1}\int_{i\tau}^{(i+1)\tau}\left\|u^{\mathrm{disc},i}_\tau - u^{\mathrm{disc},i+1}_\tau\right\|^2_H\,\mathrm{d}t \le C\tau^2$ and $\sum_{i=0}^{N-1}\int_{i\tau}^{(i+1)\tau}\left\|u^{\mathrm{disc},i+1}_\tau - u^d\right\|^2_H \le C\tau^2$. Moreover, since $\phi_\varepsilon(u^0_\varepsilon)$ is uniformly bounded, then $\left(\phi_\varepsilon\left(u^R_{\tau,\varepsilon}\right) - \phi_0\left(u^*\right)\right)$ is bounded in $L^1(0,T)$ independently of ε, thus

$$\int_0^\tau \Delta^\phi_{\tau,\varepsilon}\,\mathrm{d}t = \int_0^\tau \mathrm{e}^{-t/\delta}\left(\phi_\varepsilon\left(u^R_{\tau,\varepsilon}\right) - \phi_0\left(u^*\right)\right)\,\mathrm{d}t \le C\tau.$$

Finally, concerning the external forces we use (2.24) and estimate

$$\Delta^g_{\tau,\varepsilon} \le \int_0^T \mathrm{e}^{-t/\delta} \left(g_\varepsilon - g, u^R_{\tau,\varepsilon}\right)_H + (g, u^R_{\tau,\varepsilon} - u^*)_H \tag{3.40}$$

$$\le C \left\|g_\varepsilon - g\right\|_{L^2(0,T;H)} + C \left\|u^R_{\tau,\varepsilon} - u^*\right\|_{L^2(0,T;H)} = C(r^g_\varepsilon + \rho^{\mathrm{dtc}}_\tau + r^{\mathfrak{R}}_\varepsilon + \tau^{1/2} r^0_\varepsilon). \tag{3.41}$$

The case of λ-convex energies can be handled similarly by repeating the above procedure with $\phi_\varepsilon(\cdot)$ replaced by $\phi_\varepsilon(\cdot) - \lambda \left\|\cdot\right\|^2_H/2$ and by estimating additionally

$$\int_0^T \mathrm{e}^{-t/\delta} \left(-\lambda \left\|u^R_{\tau,\varepsilon}\right\|^2_H + \lambda \left\|u^*\right\|^2_H\right) \mathrm{d}t \le C \int_0^T \left\|u^R_{\tau,\varepsilon} + u^*\right\|_H \left\|u^R_{\tau,\varepsilon} - u^*\right\|_H \mathrm{d}t$$
$$\le C \left(r^{\mathfrak{R}}_\varepsilon + \rho^{\mathrm{dct}}_\tau + \tau^{1/2} r^0_\varepsilon\right).$$

Summarizing we have

$$\left\|u^R_{\tau,\varepsilon} - u^*\right\|_{L^2(0,T;H)} \le C(r^{\mathfrak{R}}_\varepsilon + \rho^{\mathrm{dtc}}_\tau + \tau^{1/2} r^0_\varepsilon),$$
$$\mathcal{I}_\varepsilon(u^R_{\tau,\varepsilon}) - \mathcal{I}(u^*) \le C \left(r^{\mathfrak{R}}_\varepsilon/\tau + \rho^{\mathrm{dtc}}_\tau + r^{\mathfrak{R},\phi}_\varepsilon + r^{\mathfrak{R},\psi}_\varepsilon + \tau + r^g_\varepsilon + r^{\mathfrak{R}}_\varepsilon + r^0_\varepsilon + \tau^{1/2} r^0_\varepsilon\right).$$

We now choose $\tau^{\mathfrak{R}}_\varepsilon$ as the minimizer of $\tau \longmapsto r^1_\varepsilon/\tau + \rho^{\mathrm{dtc}}_\tau + \tau + \tau^{1/2} r^0_\varepsilon$, and we easily check that $\tau^{\mathfrak{R}}_\varepsilon \to 0$ as $\varepsilon \to 0$. Note that this is the choice of τ (as a function of ε) that optimizes the convergence rates. In particular, by defining $u^R_\varepsilon := u^R_{\tau^{\mathfrak{R}}_\varepsilon,\varepsilon}$ we see that conditions (3.26) and (3.27) follow from the above estimates where

$$\rho^1_\varepsilon = r^{\mathfrak{R}}_\varepsilon + \rho^{\mathrm{dtc}}_{\tau^{\mathfrak{R}}_\varepsilon} + \left(\tau^{\mathfrak{R}}_\varepsilon\right)^{1/2} r^0_\varepsilon,$$
$$\rho^2_\varepsilon = r^{\mathfrak{R}}_\varepsilon/\tau^{\mathfrak{R}}_\varepsilon + \rho^{\mathrm{dtc}}_{\tau^{\mathfrak{R}}_\varepsilon} + r^{\mathfrak{R},\phi}_\varepsilon + r^{\mathfrak{R},\psi}_\varepsilon + \tau^{\mathfrak{R}}_\varepsilon + r^g_\varepsilon + r^{\mathfrak{R}}_\varepsilon + r^0_\varepsilon + \left(\tau^{\mathfrak{R}}_\varepsilon\right)^{1/2} r^0_\varepsilon.$$

Proof of (3.28). A proof of (3.28) can be obtained by following the proofs of (3.26) and (3.27) described above and exchanging the role of ϕ_0 and ϕ_ε, of ψ_0 and ψ_ε, of u^{disc} and $u^{\mathrm{disc}}_\varepsilon$, and of u^* and u^*_ε. More precisely, let $u^{\mathrm{disc}}_\varepsilon$ be the minimizer of the discrete WED functional $\mathcal{I}^\tau_\varepsilon$ associated with ϕ_ε, ψ_ε, and g_ε, and defined with respect to the partition $\{i\tau\}_{i=0}^N$. Let u^*_ε be the minimizer of the WED functional $\mathcal{I}_\varepsilon$.

For all τ, let us define $u^S_{\tau,\varepsilon}$ as the piecewise affine interpolant with respect to the nodes $\{u^0, \mathfrak{S}u^{\mathrm{disc},1}_\varepsilon, \ldots, \mathfrak{S}u^{\mathrm{disc},N}_\varepsilon\}$. Note that, by using assumptions (2.21), and arguing as for (3.38), we can prove that

$$\left\|u^S_{\tau,\varepsilon} - u^*_\varepsilon\right\|_{L^2(0,T;H)} \le C(r^{\mathfrak{S}}_\varepsilon + \rho^{\mathrm{dtc}}_\tau + \tau^{1/2} r^0_\varepsilon).$$

Computations analogous to the ones used in (3.39) and (3.41) combined with assumptions (2.21)–(2.24) give us

$$\int_0^T \mathrm{e}^{-t/\delta} \left(\psi_0\left(\partial_t u^S_{\tau,\varepsilon}\right) - \psi_\varepsilon(\partial_t u^*_\varepsilon)\right) \mathrm{d}t \le C \left(\frac{r^{\mathfrak{S}}_\varepsilon}{\tau} + r^{\mathfrak{S},\psi}_\varepsilon + \rho^{\mathrm{dtc}}_\tau + r^0_\varepsilon\right),$$
$$\frac{1}{\delta} \int_0^T \mathrm{e}^{-t/\delta} \left(\phi_0(u^S_{\tau,\varepsilon}) - \phi_\varepsilon(u^*_\varepsilon)\right) \mathrm{d}t \le C \left(r^{\mathfrak{S},\phi}_\varepsilon + \tau + \rho^{\mathrm{dtc}}_\tau\right),$$
$$\frac{1}{\delta} \int_0^T \mathrm{e}^{-t/\delta} \left((g(t), u^S_{\tau,\varepsilon})_H - (g_\varepsilon(t), u^*_\varepsilon)_H\right) \mathrm{d}t \le C(r^g_\varepsilon + \rho^{\mathrm{dtc}}_\tau + r^{\mathfrak{S}}_\varepsilon + \tau^{1/2} r^0_\varepsilon).$$

Here, we also used the uniform coercivity and growth assumptions such that all the bounds on u_ε^* and $u_\varepsilon^{\rm disc}$ and the convergence rates for $u_\varepsilon^{\rm disc} \to u_\varepsilon^*$ as $\tau \to 0$ are uniform in ε.

Summarizing we have

$$\mathcal{I}(u_{\tau,\varepsilon}^S) - \mathcal{I}_\varepsilon(u_\varepsilon^*) \le C(r_\varepsilon^{\mathfrak{S}}/\tau + \rho_\tau^{\rm dtc} + r_\varepsilon^{\mathfrak{S}} + r_\varepsilon^{\mathfrak{S},\phi} + r_\varepsilon^{\mathfrak{S},\psi} + \tau + r_\varepsilon^g + r_\varepsilon^0 + \tau^{1/2} r_\varepsilon^0).$$

We can now choose $\tau = \tau_\varepsilon^{\mathfrak{S}} := \arg\min_\tau (r_\varepsilon^{\mathfrak{S}}/\tau + \rho_\tau^{\rm dtc} + \tau + \tau^{1/2} r_\varepsilon^0)$, and define $u_\varepsilon^S := u^S_{\tau_\varepsilon^{\mathfrak{S}},\varepsilon}$. Then, condition (3.28) is satisfied with

$$\rho_\varepsilon^3 = r_\varepsilon^{\mathfrak{S}}/\tau_\varepsilon^{\mathfrak{S}} + \rho_{\tau_\varepsilon^{\mathfrak{S}}}^{\rm dtc} + r_\varepsilon^{\mathfrak{S},\phi} + r_\varepsilon^{\mathfrak{S},\psi} + \tau_\varepsilon^{\mathfrak{S}} + r_\varepsilon^g + r_\varepsilon^{\mathfrak{S}} + r_\varepsilon^0 + \left(\tau_\varepsilon^{\mathfrak{S}}\right)^{1/2} r_\varepsilon^0.$$

□

Proof of (3.29). Finally, we want to check that (3.29) holds true for $\Gamma = 2$ and $\rho_\varepsilon^4 = 0$. We first recall that the minimizer u_ε^* of $\mathcal{I}_\varepsilon$ solves the elliptic equation ($\mathrm{P}_{\delta,\varepsilon}$). By testing the Euler–Lagrange equation for $\mathcal{I}_\varepsilon$ with $v = u_\varepsilon^* - u_\varepsilon^R$ we get

$$0 = \int_0^T \mathrm{e}^{-t/\delta} \Big[\big(\mathbb{A}_\varepsilon \partial_t u_\varepsilon^*, \partial_t(u_\varepsilon^* - u_\varepsilon^R)\big)_H + \frac{1}{\delta}\left(\eta_\varepsilon^* - g_\varepsilon, u_\varepsilon^* - u_\varepsilon^R\right)_H \Big] \mathrm{d}t.$$

where $\eta_\varepsilon^* \in \partial_H \phi_\varepsilon(u_\varepsilon^*)$ a.e. in $[0,T]$. We use the equation above to calculate

$$\begin{aligned}
\mathcal{I}_\varepsilon(u_\varepsilon^R) - \mathcal{I}_\varepsilon(u_\varepsilon^*) &= \int_0^T \mathrm{e}^{-t/\delta} \Big[\psi_\varepsilon(\partial_t u_\varepsilon^R) - \psi_\varepsilon(\partial_t u_\varepsilon^*) + \frac{1}{\delta}\left(\phi_\varepsilon(u_\varepsilon^R) - \phi_\varepsilon(u_\varepsilon^*)\right) - \frac{1}{\delta}\left(g_\varepsilon, u_\varepsilon^R - u_\varepsilon^*\right)_H\Big]\,\mathrm{d}t \\
&= \int_0^T \mathrm{e}^{-t/\delta} \left(\psi_\varepsilon(\partial_t u_\varepsilon^R) - \psi_\varepsilon(\partial_t u_\varepsilon^*) + \big(\mathbb{A}_\varepsilon \partial_t u_\varepsilon^*, \partial_t(u_\varepsilon^* - u_\varepsilon^R)\big)_H\right)\,\mathrm{d}t \\
&\quad + \frac{1}{\delta}\int_0^T \mathrm{e}^{-t/\delta}\left(\phi_\varepsilon(u_\varepsilon^R) - \phi_\varepsilon(u_\varepsilon^*) + \left(\eta_\varepsilon^*, u_\varepsilon^* - u_\varepsilon^R\right)_H\right)\,\mathrm{d}t.
\end{aligned}$$

By using λ-convexity of ϕ_ε, we get

$$\frac{1}{\delta}\int_0^T \mathrm{e}^{-t/\delta}\left(\phi_\varepsilon(u_\varepsilon^R) - \phi_\varepsilon(u_\varepsilon^*) + \left(\eta_\varepsilon^*, u_\varepsilon^* - u_\varepsilon^R\right)_H\right)\,\mathrm{d}t \ge \frac{\lambda}{2\delta}\int_0^T \mathrm{e}^{-t/\delta}\left\|u_\varepsilon^R - u_\varepsilon^*\right\|_H^2\,\mathrm{d}t.$$

We now exploit that ψ_ε is quadratic to obtain

$$\begin{aligned}
\mathcal{I}_\varepsilon(u_\varepsilon^R) - \mathcal{I}_\varepsilon(u_\varepsilon^*) &\ge \int_0^T \mathrm{e}^{-t/\delta}\left(\psi_\varepsilon(\partial_t u_\varepsilon^R) + \psi_\varepsilon(\partial_t u_\varepsilon^*) - \left(\mathbb{A}_\varepsilon \partial_t u_\varepsilon^*, \partial_t u_\varepsilon^R\right)_H\right)\,\mathrm{d}t \\
&= \int_0^T \mathrm{e}^{-t/\delta}\left(\psi_\varepsilon(\partial_t u_\varepsilon^R - \partial_t u_\varepsilon^*)\right)\,\mathrm{d}t \\
&\ge c\int_0^T \mathrm{e}^{-t/\delta}\left\|\partial_t u_\varepsilon^R - \partial_t u_\varepsilon^*\right\|_H^2\,\mathrm{d}t + \frac{\lambda}{2\delta}\int_0^T \mathrm{e}^{-t/\delta}\left\|u_\varepsilon^R - u_\varepsilon^*\right\|_H^2\,\mathrm{d}t.
\end{aligned}$$

By arguing as in Proposition 2.1 of [19], recalling that $v(0) = 0$, for all $\delta > 0$ sufficiently small we obtain, after suitably renaming the constant c,

$$\mathcal{I}_\varepsilon(u_\varepsilon^R) - \mathcal{I}_\varepsilon(u_\varepsilon^*) \ge c\int_0^T \mathrm{e}^{-t/\delta}\left\|u_\varepsilon^R - u_\varepsilon^*\right\|_H^2\,\mathrm{d}t,$$

which implies (3.29). We conclude the proof of Theorem 2.7 by applying the abstract Lemma 3.3 with $\mathcal{V} = L^2(0,T;H)$. □

4. Applications

In this section we provide two examples of application of the above abstract theory. The first one is related to homogenization problems while the second deals with dimension reduction.

4.1. Homogenization of a 1D parabolic equation

In the 1-dimensional domain $\Omega = (0,1)$ we consider the equation

$$a_\varepsilon(x)\dot{u} - \partial_x\left(D_\varepsilon(x)\partial_x u\right) + b_\varepsilon(x)u - c_\varepsilon(x)|u|^{q-2}u = g_\varepsilon(t,x) \quad \text{in } \Omega\times(0,T), \tag{4.1}$$

equipped with homogeneous Neumann boundary conditions $\partial_x u(0,t) = \partial_x u(1,t) = 0$. Here, $1 < q < 2$ and $a_\varepsilon(x) = a(x/\varepsilon)$, $b_\varepsilon(x) = b(x/\varepsilon)$, $D_\varepsilon(x) = D(x/\varepsilon)$, $c_\varepsilon(x) = c(x/\varepsilon)$ for some 1-periodic continuous functions $a, b, D, c : [0,1] \to [1/M, M]$ with $M > 0$. The initial data u_ε^0 is assumed to be uniformly bounded in $H^1(\Omega)$ and the external forces satisfy $g_\varepsilon \in L^2([0,T]\times\Omega)$.

Setting $H = L^2(\Omega)$ and $X = H^1(\Omega)$ the equation can be rewritten in the form

$$\mathrm{d}\psi_\varepsilon(\dot{u}) + \partial_H\varphi_\varepsilon^1(u) - \partial_H\varphi_\varepsilon^2(u) = g_\varepsilon(t) \text{ in } H^*,$$

where

$$\psi_\varepsilon(v) = \frac{1}{2}\int_\Omega a_\varepsilon(x)|v|^2\,\mathrm{d}x \qquad \text{and} \qquad \phi_\varepsilon(u) = \varphi_\varepsilon^1(u) - \varphi_\varepsilon^2(u) \quad \text{with}$$

$$\varphi_\varepsilon^1(u) = \begin{cases} \frac{1}{2}\int_\Omega D_\varepsilon(x)|\partial_x u|^2 + b_\varepsilon(x)|u|^2\,\mathrm{d}x & \text{if } u \in H^1(\Omega), \\ \infty & \text{else,} \end{cases} \qquad \text{and}$$

$$\varphi_\varepsilon^2(u) = \frac{1}{q}\int_\Omega c_\varepsilon(x)|u|^q\,\mathrm{d}x.$$

It is not hard to check that assumptions **(A1)–(A7)** are satisfied with $p = m = 2$ and that the elliptic-in-time regularization given by

$$-\delta a_\varepsilon(x)\ddot{u} + a_\varepsilon(x)\dot{u} - \partial_x\left(D_\varepsilon(x)\partial_x u\right) + b_\varepsilon(x)u - c_\varepsilon(x)|u|^{q-2}u = g_\varepsilon(t,x) \text{ in } \Omega\times(0,T), \tag{4.2}$$

with additional final condition $\delta\dot{u}(x,T) = 0$ for a.e. $x \in \Omega$, corresponds to the Euler–Lagrange equation of the WED functional (*cf.* Thm. 2.1)

$$\mathcal{I}_{\delta,\varepsilon}(u) := \begin{cases} \displaystyle\int_0^T\int_\Omega \mathrm{e}^{-t/\delta}\left(\frac{a_\varepsilon}{2}|\dot{u}|^2 + \frac{D_\varepsilon}{2\delta}|\partial_x u|^2 + \frac{b_\varepsilon}{2\delta}|u|^2 - \frac{c_\varepsilon}{q\delta}|u|^q - \frac{1}{\delta}g_\varepsilon(t)u\right)\mathrm{d}x\,\mathrm{d}t & \text{if } u \in \mathcal{K}(u_\varepsilon^0), \\ \infty & \text{else.} \end{cases}$$

We will apply the theory developed in the previous sections to show that the functionals $\mathcal{I}_{\delta,\varepsilon}$ Γ-converge to the limiting WED functional $\mathcal{I}_\delta$ defined by the static Γ-limits ϕ_0 and ψ_0. The latter are given by

$$\psi_0(v) = \frac{1}{2}\int_\Omega a_{\mathrm{arith}}|v|^2\,\mathrm{d}x \quad \text{and} \quad \phi_0(u) = \varphi_0^1(u) - \varphi_0^2(u) \quad \text{with}$$

$$\varphi_0^1(u) = \begin{cases} \frac{1}{2}\int_\Omega D_{\text{harm}}|\partial_x u|^2 + b_{\text{arith}}|u|^2 \,\mathrm{d}x & \text{if } u \in H^1(\Omega), \\ \infty & \text{else,} \end{cases}$$
$$\varphi_0^2(u) = \frac{1}{q}\int_\Omega c_{\text{arith}}|u|^q \,\mathrm{d}x,$$

where a_{arith}, b_{arith}, c_{arith} denote the arithmetic means of a, b, c respectively and D_{harm} denotes the harmonic mean of D, namely $D_{\text{harm}} = (\int_0^1 1/D(y)\,\mathrm{d}y)^{-1}$. We refer the reader to [17] to check that $\varphi_\varepsilon^1 \xrightarrow{M} \varphi_0^1$ in $L^2(\Omega)$ (in particular, $\varphi_\varepsilon^1 \xrightarrow{\Gamma} \varphi_0^1$ weakly in $H^1(\Omega)$), and that $\psi_\varepsilon \xrightarrow{C} \psi_0$ with respect to strong convergence in $L^2(\Omega)$. Moreover, note that for every sequence u_ε with $\varphi_\varepsilon^1(u_\varepsilon) \le C$, we have $u_\varepsilon \to u$ weakly in $H^1(\Omega)$, $u_\varepsilon \to u$ strongly in $L^q(\Omega)$ and, as a consequence of the Sobolev embeddings,

$$|u_\varepsilon(x)|^q \le \|u_\varepsilon\|_{L^\infty(\Omega)}^q \le C\,\|u_\varepsilon\|_{H^1(\Omega)}^q \le C\left(\varphi_\varepsilon^1(u_\varepsilon)\right)^{q/2} \le C \text{ for a.a. } x \in \Omega.$$

Thus, by using the dominated convergence theorem and the convergence $c_\varepsilon \to c_{\text{arith}}$ weakly* in $L^\infty(\Omega)$ we get

$$\varphi_\varepsilon^2(u_\varepsilon) - \varphi_0^2(u) = \frac{1}{q}\int_\Omega c_\varepsilon|u_\varepsilon|^q - c_{\text{arith}}|u|^q \,\mathrm{d}x = \frac{1}{q}\int_\Omega c_\varepsilon\left(|u_\varepsilon|^q - |u|^q\right) + \left(c_\varepsilon - c_{\text{arith}}\right)|u|^p \,\mathrm{d}x \to 0.$$

In particular, it holds that $\phi_\varepsilon \xrightarrow{M} \phi_0$ in $L^2(\Omega)$ (or, equivalently, $\phi_\varepsilon \xrightarrow{\Gamma} \phi_0$ weakly in $H^1(\Omega)$). Finally, we assume that $u_\varepsilon^0 \to u^0$ strongly in $L^2(\Omega)$ and that the external force terms satisfy $g_\varepsilon \to g_0$ weakly in $L^2(0,T;H)$.

By applying Theorems 2.2 and 2.5, we arrive at the following homogenization result.

Theorem 4.1 (Homogenization)**.** *Let $u_{\delta,\varepsilon}^*$ be a minimizer of the WED functional $\mathcal{I}_{\delta,\varepsilon}$. Then, $u_{\delta,\varepsilon}^*$ solves (4.2) and converges to u_δ^*, a minimizer of the WED functional $\mathcal{I}_{\delta,0}$ which is in turn a solution to the homogenized equation*

$$-\delta a_{\text{arith}}\ddot{u} + a_{\text{arith}}\dot{u} - \partial_x\left(D_{\text{harm}}\partial_x u\right) + b_{\text{arith}}u - c_{\text{arith}}|u|^{q-2}u = g_0(t,x) \text{ in } \Omega, \tag{4.3}$$

for $\varepsilon \to 0$. Moreover, if additionally g_ε satisfies (2.14)-(2.15) and the initial data are well prepared, i.e. $\phi_\varepsilon(u_\varepsilon^0) \to \phi(u^0)$, then, up to (not relabeled) subsequences the following convergences in the weak topology of $H^1(0,T;L^2(\Omega)) \cap L^2(0,T;H^1(\Omega))$ and in the strong topology of $C\left([0,T];L^2(\Omega)\right)$ hold true

(i) $\lim_{\varepsilon\to 0} u_{\delta,\varepsilon}^* = u_\delta^*$ *where $u_\delta^* \in \arg\min \mathcal{I}_\delta$ and solves (4.3),*
(ii) $\lim_{\delta\to 0} u_{\delta,\varepsilon}^* = u_\varepsilon$ *where u_ε solves (4.1),*
(iii) $\lim_{\delta+\varepsilon\to 0} u_{\delta,\varepsilon}^* = u$ *where u solves the following homogenized equation*

$$a_{\text{arith}}\dot{u} - \partial_x\left(D_{\text{harm}}\partial_x u\right) + b_{\text{arith}}u - c_{\text{arith}}|u|^{q-2}u = g_0 \text{ in } \Omega \times [0,T].$$

We are now interested to establish some convergence rates. More precisely, we want to estimate the rate of convergence of $\left\|u_{\delta,\varepsilon}^* - u_\delta^*\right\|_{L^2(0,T;L^2(\Omega))}$, where $u_{\delta,\varepsilon}^*$, u_δ^* are defined in Theorem 4.1 (other norms can be considered as well due to interpolation). We obtain these rates of convergence by applying Theorem 2.7. To do this, we restrict ourselves to the case of a convex energy potential, *i.e.* $c(x) = 0$ for all $x \in \Omega$ (*i.e.* $\varphi_\varepsilon^2 = 0$, $\varphi_\varepsilon^1 = \phi_\varepsilon$ for all $\varepsilon \ge 0$).

Let us start to check the assumption of Theorem 2.7. Let us define the recovery operator $\mathfrak{R}_\varepsilon : H^1(\Omega) \to H^1(\Omega)$ for the functionals ϕ_ε as

$$\mathfrak{R}_\varepsilon u(x) = \int_0^x D_\varepsilon^{-1}(y) D_{\text{harm}}\partial_x u(y)\mathrm{d}y + u(0) \qquad \text{for a.a. } x \in \Omega.$$

We start by observing that

$$\|\mathfrak{R}_\varepsilon u - u\|_{L^2(\Omega)} = \varepsilon C \left(1 + \|u\|_{H^1(\Omega)}\right).$$

Moreover, we write

$$\begin{aligned}\phi_\varepsilon(\mathfrak{R}_\varepsilon u) - \phi_0(u) = &\int_\Omega \left(D_\varepsilon(x)\left(\partial_x \mathfrak{R}_\varepsilon u(x)\right)^2 - D_{\text{harm}}\left(\partial_x u(x)\right)^2\right)\mathrm{d}x \\ &+ \int_\Omega \left(b_\varepsilon(x)\left(\mathfrak{R}_\varepsilon u(x)\right)^2 - b_{\text{arith}} u^2(x)\right)\mathrm{d}x.\end{aligned}$$

Recalling that $\partial_x \mathfrak{R}_\varepsilon u = D_\varepsilon^{-1} D_{\text{harm}} \partial_x u$, we get

$$\begin{aligned}\int_\Omega \left(D_\varepsilon(x)\left(\partial_x \mathfrak{R}_\varepsilon u(x)\right)^2 - D_{\text{harm}}\left(\partial_x u(x)\right)^2\right)\mathrm{d}x &\le C \left\|D_\varepsilon^{-1} D_{\text{harm}} - 1\right\|_{(H^1(\Omega))^*} \left\|(\partial_x u)^2\right\|_{H^1(\Omega)} \\ &\le \varepsilon C \|u\|^2_{H^2(\Omega)} \\ &\le \varepsilon C \left(\|\eta\|^2_{L^2(\Omega)} + \|u\|^2_{H^1(\Omega)}\right) \text{ for } \eta \in \partial_H \phi_0(u).\end{aligned}$$

Here we used that, since Ω is 1-dimensional, $\partial_x u \in H^1(\Omega)$ implies $\partial_x u \in L^\infty(\Omega)$ and hence $(\partial_x u)^2 \in H^1(\Omega)$. Moreover, we have that

$$\begin{aligned}\int_\Omega \left(b_\varepsilon(x)\left(\mathfrak{R}_\varepsilon u(x)\right)^2 - b_{\text{arith}} u^2(x)\right)\mathrm{d}x &= \int_\Omega \left(b_\varepsilon(x) - b_{\text{arith}}\right)\left(\mathfrak{R}_\varepsilon u(x)\right)^2 \mathrm{d}x + \int_\Omega b_{\text{arith}}\left(\mathfrak{R}_\varepsilon u(x) - u(x)\right)\left(\mathfrak{R}_\varepsilon u(x) + u(x)\right)\mathrm{d}x \\ &\le \|b_\varepsilon - b_{\text{arith}}\|_{H^{-1}(\Omega)} \left\|(\mathfrak{R}_\varepsilon u)^2\right\|_{H^1(\Omega)} + C \|\mathfrak{R}_\varepsilon u - u\|_{L^2(\Omega)} \|\mathfrak{R}_\varepsilon u + u\|_{L^2(\Omega)} \\ &\le C\varepsilon \|u\|^2_{H^1(\Omega)}.\end{aligned}$$

Combining these estimates, we get

$$\phi_\varepsilon(\mathfrak{R}_\varepsilon u) - \phi_0(u) \le \varepsilon C \left(\|u\|^2_{H^1(\Omega)} + \|\eta\|^2_{L^2(\Omega)}\right) \text{ for } \eta \in \partial_{L^2(\Omega)} \phi_0(u).$$

Let us define $B := H^1(\Omega)$ and prove that it satisfies condition (2.17). Thanks to the first estimate in (3.37), we have that $u^*_{\delta,\varepsilon}$ is uniformly bounded in $H^2(0,T;L^2(\Omega)) \cap L^2(0,T;H^1(\Omega))$. Moreover, since $\partial_H \phi_\varepsilon$, $\varepsilon \ge 0$, are second-order uniformly coercive elliptic operators, we also have that $u^*_{\delta,\varepsilon}$ is uniformly bounded in $L^2(0,T;H^2(\Omega))$. By interpolation we have that $u^*_{\delta,\varepsilon}$ is uniformly bounded in $H^1(0,T;H^1(\Omega))$. We now estimate

$$\begin{aligned}|\psi_\varepsilon(v) - \psi_0(w)| &= \frac{1}{2}\int_\Omega |a_{\text{arith}} w^2(x) - a_\varepsilon(x) v^2(x)|\mathrm{d}x \\ &\le \frac{1}{2}\int_\Omega |a_{\text{arith}} - a_\varepsilon(x)|\, w^2(x) + |w^2(x) - v^2(x)|\, a_\varepsilon(x)\mathrm{d}x \\ &\le \frac{1}{2}\|a_{\text{arith}} - a_\varepsilon(x)\|_{(H^1(\Omega))^*} \left\|w^2(x)\right\|_{H^1(\Omega)} \\ &\quad + \frac{1}{2}\|a_\varepsilon\|_{L^\infty(\Omega)} \|w - v\|_{L^2(\Omega)} \|w + v\|_{L^2(\Omega)} \\ &\le C_\omega \varepsilon \|w\|^2_{H^1(\Omega)} + C\|v - w\|_{L^2(\Omega)}.\end{aligned} \tag{4.4}$$

This proves that $\mathfrak{R}_\varepsilon$ satisfies (2.18)-(2.20) with

$$r_\varepsilon^{\mathfrak{R}} = r_\varepsilon^{\mathfrak{R},\phi} = r_\varepsilon^{\mathfrak{R},\psi} = \varepsilon.$$

We now define the operator $\mathfrak{S}_\varepsilon : H^1(\Omega) \to H^1(\Omega)$ by setting it to the piecewise affine interpolant of the nodes $\{u_\varepsilon^i = u(\varepsilon i)\}$ over the partition $[0,1] = \cup_{i=0}^{N-1}[i\varepsilon,(i+1)\varepsilon]$. Note that for all $u \in H^1(\Omega)$ we have

$$\|\mathfrak{S}_\varepsilon u - u\|_{L^2(\Omega)} \leq \varepsilon C \left(1 + \|u\|_{H^1(\Omega)}\right).$$

Furthermore,

$$\phi_0(\mathfrak{S}_\varepsilon u) \leq \phi_\varepsilon(u) + \varepsilon C \left(1 + \|u\|_{H^1(\Omega)}^2 + \|\eta\|_{L^2(\Omega)}^2\right). \tag{4.5}$$

Indeed, by convexity, we have that

$$\begin{aligned}
\phi_0(\mathfrak{S}_\varepsilon u) &= \sum_{i=0}^{N-1} \frac{1}{2} \int_{i\varepsilon}^{(i+1)\varepsilon} \left(D_{\text{harm}} \left(\frac{u_\varepsilon^{i+1} - u_\varepsilon^i}{\varepsilon} \right)^2 + b_{\text{arith}} (\mathfrak{S}_\varepsilon u(x))^2 \right) \mathrm{d}x \\
&\leq \sum_{i=0}^{N-1} \frac{1}{2} \int_{i\varepsilon}^{(i+1)\varepsilon} \left(D_{\text{arith}} \left(\frac{u_\varepsilon^{i+1} - u_\varepsilon^i}{\varepsilon} \right)^2 + b_{\text{arith}} (u(x))^2 \right) \mathrm{d}x + r_\varepsilon^{\mathfrak{S},\phi,1} \\
&\leq \sum_{i=0}^{N-1} \frac{1}{2} \int_{i\varepsilon}^{(i+1)\varepsilon} \left(D_{\text{arith}} \left(\partial_x u(x)\right)^2 + b_{\text{arith}} (u(x))^2 \right) \mathrm{d}x + r_\varepsilon^{\mathfrak{S},\phi,1} \\
&= \sum_{i=0}^{N-1} \frac{1}{2} \int_{i\varepsilon}^{(i+1)\varepsilon} \left(D_\varepsilon(x) \left(\partial_x u(x)\right)^2 + b_\varepsilon(x) (u(x))^2 \right) \mathrm{d}x + r_\varepsilon^{\mathfrak{S},\phi,1} + r_\varepsilon^{\mathfrak{S},\phi,2} = \phi_\varepsilon(u) + r_\varepsilon^{\mathfrak{S},\phi}.
\end{aligned}$$

Here, we have

$$r_\varepsilon^{\mathfrak{S},\phi,1} = \sum_{i=0}^{N-1} \frac{1}{2} \int_{i\varepsilon}^{(i+1)\varepsilon} b_{\text{arith}} \left((\mathfrak{S}_\varepsilon u(x))^2 - u^2(x) \right) \mathrm{d}x \leq \varepsilon C \left(1 + \|u\|_{H^1(\Omega)}^2 \right)$$

and

$$\begin{aligned}
r_\varepsilon^{\mathfrak{S},\phi,2} &= \sum_{i=0}^{N-1} \frac{1}{2} \int_{i\varepsilon}^{(i+1)\varepsilon} \left(\partial_x u(x)\right)^2 \left(D_{\text{arith}} - D_\varepsilon\right) \mathrm{d}x \leq \varepsilon C \left(1 + \|u\|_{H^2(\Omega)}^2 \right) \\
&= \varepsilon C \left(1 + \|\eta\|_{L^2(\Omega)}^2 + \|u\|_{H^1(\Omega)}^2 \right).
\end{aligned}$$

Combining the above estimates, we deduce (4.5). Finally, by arguing as in (4.4), we get

$$\begin{aligned}
|\psi_\varepsilon(v) - \psi_0(w)| &= \frac{1}{2} \int_\Omega \left| a_{\text{arith}} w^2(x) - a_\varepsilon(x) v^2(x) \right| \mathrm{d}x \\
&= \frac{1}{2} \int_\Omega a_{\text{arith}} \left| w^2(x) - v^2(x) \right| + \left| a_{\text{arith}} - a_\varepsilon(x) \right| v^2(x) \, \mathrm{d}x \\
&\leq C_\delta \varepsilon \|v\|_{H^{1/2}(\Omega)}^2 + C \|v - w\|_{L^2(\Omega)}.
\end{aligned}$$

Hence, $\mathfrak{S}_\varepsilon$ satisfies (2.21)–(2.23) with

$$r_\varepsilon^{\mathfrak{S}} = r_\varepsilon^{\mathfrak{S},\phi} = r_\varepsilon^{\mathfrak{S},\psi} = \varepsilon.$$

In order to estimate convergence rates we additionally assume

$$\|g_\varepsilon - g_0\|_{L^2(\Omega\times[0,T])} \le r_\varepsilon^g \quad \text{ and } \quad \left\|u_\varepsilon^0 - u^0\right\|_{L^2(\Omega)} \le r_\varepsilon^0$$

for some $r_\varepsilon^g = o(1)$, $r_\varepsilon^0 = o(1)$. Moreover, we assume that u_ε^0 is uniformly bounded in $H^1(\Omega)$. This guaranties uniform boundedness of the initial energies $\phi_\varepsilon(u_\varepsilon^0)$. Note that, here we do *not* need well preparedness of initial data. With this preparation, a straightforward application of Theorem 2.7 gives us the following estimate on the convergence rates.

Theorem 4.2 (Convergence rates for the homogenization problem)**.** *Let $u_{\delta,\varepsilon}^*$ and u_δ^* be the minimizers of $\mathcal{I}_{\delta,\varepsilon}$ and $\mathcal{I}_{\delta,0}$, respectively. Then, for all $\delta > 0$ sufficiently small there exists a constant C_δ independent of ε such that*

$$\left\|u_{\delta,\varepsilon}^* - u_\delta^*\right\|_{L^2(0,T;L^2(\Omega))} \le C_\delta \left(r_\varepsilon^0 + r_\varepsilon^g + \varepsilon\right)^{1/2}.$$

Moreover, by interpolation one can prove the following

$$\|u\|^2_{L^2(0,T;H^{1/2}(\Omega))} \le C\|u\|_{L^2(0,T;H^1(\Omega))}\|u\|_{L^2(0,T;L^2(\Omega))} \text{ for all } u \in L^2(0,T;H^1(\Omega)),$$
$$\|u\|^2_{H^1(0,T;L^2(\Omega))} \le C\|u\|_{H^2(0,T;L^2(\Omega))}\|u\|_{L^2(0,T;L^2(\Omega))} \text{ for all } u \in H^2(0,T;L^2(\Omega)).$$

As a consequence, we can get convergence rates also in a stronger norm.

Corollary 4.3 (Convergence rates in a stronger norm)**.** *Let $u_{\delta,\varepsilon}^*$ and u_δ^* be the minimizers of $\mathcal{I}_{\delta,\varepsilon}$ and of $\mathcal{I}_\delta$, respectively. Then, for all $\delta > 0$ sufficiently small*

$$\left\|u_{\delta,\varepsilon}^* - u_\delta^*\right\|_{H^1(0,T;L^2(\Omega))\cap L^2(0,T;H^{1/2}(\Omega))} \le C_\delta \left(r_\varepsilon^0 + r_\varepsilon^g + \varepsilon\right)^{1/4},$$

for some positive constant C_δ.

4.2. Dimension reduction

For $\omega \subset \mathbb{R}^{d-1}$, $d \ge 2$, open, bounded, and with sufficiently smooth boundary, we introduce the thin domain $\Omega_\varepsilon := \omega \times]0,\varepsilon[\subset \mathbb{R}^d$. Therein, we consider the doubly nonlinear parabolic equation

$$\alpha(\dot u) - \nabla\cdot\big(E'(\nabla u)\big) + h(u) = \tilde g(t) \quad \text{in } \Omega_\varepsilon, \tag{4.6}$$

subject to homogeneous Dirichlet boundary conditions on $\partial\omega \times [0,\varepsilon]$ and homogeneous Neumann boundary conditions on the remaining part of the boundary.

In order to fit this problem to the abstract setting of Section 2 we make the following assumptions on the data: We assume that α is maximal monotone and there exists $1 < p < \infty$ and constants $c_1, c_2 > 0$ such that for $A(s) = \int_0^s \alpha(\sigma)\,\mathrm{d}\sigma$ we have

$$\forall\, s \in \mathbb{R}: \quad A(s) \ge c_1|s|^p - \frac{1}{c_1}, \qquad |\alpha(s)|^{p'} \le c_2(|s|^p + 1)$$

Moreover, we assume that for $m \in (1,\infty)$ satisfying $1 < p < m^* := dm/(d-m)_+$ there exists $c_3 > 0$ such that the function $E \in C^2(\mathbb{R}^d)$ fulfills, for constants $c_3, c_4 \geq 0$

$$\forall\, z \in \mathbb{R}^d: \quad c_3|z|^m - \frac{1}{c_3} \leq E(z), \quad |E'(z)|^{m'} \leq c_4(|z|^m + 1).$$

The nonlinear function h has the decomposition $h(u) = f_1'(u) - f_2'(u)$ with convex functions $f_1, f_2 \in C^1(\mathbb{R})$ fulfilling the following growth conditions

$$\forall\, u \in \mathbb{R}: \quad |f_1'(u)|^{m'} \leq c_4(|u|^{m^*} + 1), \qquad |f_2'(u)|^{p'} \leq c_5(|u|^{m^*} + 1).$$

Finally, we assume that $\tilde{g} \in L^{p'}((0,T) \times \Omega_\varepsilon)$.

On the spaces $H_\varepsilon = L^p(\Omega_\varepsilon)$ and $X_\varepsilon = W_D^{1,m}(\Omega_\varepsilon) := \{u \in W^{1,m}(\Omega_\varepsilon) \,|\, u = 0 \text{ on } \partial\Omega \times [0,\varepsilon]\}$ we introduce the functionals $\tilde{\psi}_\varepsilon : H_\varepsilon \to [0,\infty[$ and $\tilde{\phi}_\varepsilon : H_\varepsilon \to \mathbb{R}_\infty$ *via*

$$\tilde{\psi}_\varepsilon(v) = \frac{1}{\varepsilon}\int_{\Omega_\varepsilon} A(v)\,\mathrm{d}x, \qquad \tilde{\phi}_\varepsilon(u) = \begin{cases} \frac{1}{\varepsilon}\int_{\Omega_\varepsilon} E(\nabla u) + f_1(u) - f_2(u)\,\mathrm{d}x & \text{if } u \in X_\varepsilon, \\ \infty & \text{otherwise.} \end{cases}$$

The normalization factor $1/\varepsilon$ is a simple scaling proportional to the measure of Ω_ε. Moreover, note that by the Rellich–Kondrachov compact embedding theorem, we obtain the compact embedding $X_\varepsilon \subset V_\varepsilon$.

In order to pass to the limit $\varepsilon \to 0$ we rescale the domain Ω_ε to a domain of fixed size. More precisely, we map a given $x = (x', x_d) \in \Omega_\varepsilon$ to $y = (x', x_d/\varepsilon) \in \Omega_1 := \Omega \times (0,1)$. Then, we associate with a given function $\tilde{u} \in W^{1,m}(\Omega_\varepsilon)$ a function $u \in W^{1,m}(\Omega_1)$ in the obvious way, *i.e.* $\tilde{u}(x', x_d) = u(x', x_d/\varepsilon)$. Thus, we have $\nabla\tilde{u} = (\partial_{x_1}u, \ldots, \partial_{x_{d-1}}u, \partial_{y_d}u/\varepsilon)^\top =: \nabla_\varepsilon u$. The rescaled functionals on the spaces $H = L^p(\Omega_1)$ and $X = W_D^{1,m}(\Omega_1)$ are defined *via* $\psi_\varepsilon(v) = \tilde{\psi}_\varepsilon(\tilde{v})$ and $\phi_\varepsilon(u) = \tilde{\phi}_\varepsilon(\tilde{u})$, where we lifted the function $\tilde{g} \in L^{p'}((0,T) \times \Omega_\varepsilon)$ to a function $g \in L^{p'}((0,T) \times \Omega_1)$ in the natural way. We easily check, that assumptions **(A1)–(A7)** are satisfied.

We introduce the limiting energy density $E_0 : \mathbb{R}^{d-1} \to \mathbb{R}$ *via* $E_0(z') = \min\{E(z', z_d) \,|\, z_d \in \mathbb{R}\}$ and denote by $\xi(z')$ the minimizer in the definition of E_0 for given $z' \in \mathbb{R}^{d-1}$. The minimization of the energy density E with respect to $z_d = \partial_{y_d}u$ is a common feature of linear and nonlinear dimension reduction theories, see, *e.g.* [8, 12]. In particular, it is a well-known result, that ϕ_ε converges in the sense of Mosco to the limiting functional

$$\phi_0(u) = \begin{cases} \int_{\Omega_1} E_0(\nabla' u) + f_1(u) - f_2(u)\,\mathrm{d}y & u \in X^0, \\ +\infty & \text{otherwise.} \end{cases}$$

where the closed subspace $X^0 \subset X$ is given *via*

$$X^0 = \{u \in W_D^{1,m}(\Omega_1) \,|\, \partial_{y_d}u = 0 \text{ a.e. in } \Omega_1\}.$$

Indeed, let $\varphi_\varepsilon^1(u) := \int_{\Omega_\varepsilon} E(\nabla_\varepsilon u) + f_1(u)\,\mathrm{d}y$. For an arbitrary $u \in X$ we have that $\varphi_\varepsilon^1(u) \geq \frac{c}{\varepsilon^m}\|\partial_{y_d}u\|_{L^m}^m - C$ such that $\liminf_{\varepsilon\to 0}\varphi_\varepsilon^1(u_\varepsilon) = \infty$ if $u_\varepsilon \to u$ weakly in X with $u \notin X^0$. On the other hand, we always have that $\varphi_\varepsilon^1(u) \geq \int_{\Omega_1} E_0(\nabla' u) + f_1(u)\,\mathrm{d}y$ due to the definition of E_0. A function $u \in X^0$ is actually independent of the d-th variable and can therefore be identified with a function $U \in W_D^{1,m}(\omega)$.

The construction of a recovery sequence is based on the minimizer in the definition of E_0: Let $u \in X^0$ be given and define $\zeta := \xi(\nabla' u) \in L^m(\Omega_1)$. For a given $u \in X^0$, we set $\mathfrak{R}_\varepsilon u = u + \varepsilon w_\varepsilon$, where $w_\varepsilon(x', y_d) = \int_0^{y_d} \zeta_\varepsilon(x', r)\,\mathrm{d}r$ and ζ_ε is a regularization of ζ defined by the elliptic problem

$$-\varepsilon^\beta \Delta^m \zeta_\varepsilon + |\zeta_\varepsilon|^{m-2}\zeta_\varepsilon = |\zeta|^{m-2}\zeta \quad \text{in } \Omega_1.$$

Here, Δ^m denotes the m-Laplacian and $\beta > 0$ is such that $\beta/m < 1$. By standard elliptic estimates we have that $\varepsilon^\beta \|\nabla\zeta_\varepsilon\|_{L^m}^m + \|\zeta_\varepsilon\|_{L^m}^m \le \|\zeta\|_{L^m}^m$ and $\zeta_\varepsilon \to \zeta$ strongly in $L^m(\Omega_1)$. (The regularization of ζ is necessary to ensure that $w_\varepsilon \in W^{1,m}(\Omega_1)$.) We now have

$$\phi_\varepsilon(\mathfrak{R}_\varepsilon u) = \int_{\Omega_1} \left[E(\nabla' u + \varepsilon \nabla' w_\varepsilon, \zeta_\varepsilon) + f_1(u + \varepsilon w_\varepsilon) - f_2(u + \varepsilon w_\varepsilon)\right] \mathrm{d}y \to \phi_0(u),$$

where we used the convergences $\varepsilon w_\varepsilon \to 0$ strongly in $W^{1,m}(\Omega_1)$ and $\zeta_\varepsilon \to \zeta = \xi(\nabla' u)$ in $L^m(\Omega_1)$ as well as the dominated convergence theorem.

We set $H^0 = \overline{X^0}^H$ and assume that the initial condition satisfies $u^0 \in X^0$ leading to $\sup_{\varepsilon>0} \phi_\varepsilon(u^0) < \infty$. The WED functional in this case is given by

$$\mathcal{I}_{\delta,\varepsilon}(u) := \begin{cases} \displaystyle\int_0^T \int_{\Omega_1} \mathrm{e}^{-t/\delta} \left[A(\dot u) + \frac{1}{\delta}\Big(E(\nabla_\varepsilon u) + f_1(u) - f_2(u) - g(t)u\Big)\right] \mathrm{d}y\, \mathrm{d}t & \text{if } u \in \mathcal{K}(u^0), \\ \infty & \text{otherwise.} \end{cases}$$

The Euler–Lagrange equation in this case reads

$$-\delta \frac{\mathrm{d}}{\mathrm{d}t}\alpha(\dot u) + \alpha(\dot u) - \nabla_\varepsilon \cdot \big(E'(\nabla_\varepsilon u)\big) + h(u) = g(t) \quad \text{in } \Omega_1. \tag{4.7}$$

We confirm that the conditions in **(Γ1)–(Γ5)** are satisfied, thus, applying the machinery developed in the previous sections leads to the following result.

Theorem 4.4 (Dimension reduction)**.** *Let $u^*_{\delta,\varepsilon}$ be a minimizer of the WED functional $\mathcal{I}_{\delta,\varepsilon}$. Then, $u^*_{\delta,\varepsilon}$ solves (4.7) and converges to u^*_δ, a minimizer of the WED functional $\mathcal{I}_{\delta,0}$, which in turn can be identified with a function U^*_δ defined on ω solving the lower-dimensional equation*

$$-\delta \frac{\mathrm{d}}{\mathrm{d}t}\alpha(\dot U) + \alpha(\dot U) - \nabla' \cdot \big(E_0'(\nabla' U)\big) + h(U) = G(t) \qquad \text{in } \omega, \tag{4.8}$$

where $G(t,x') = \int_0^1 g(t,x',r)\,\mathrm{d}r$. Moreover, up to (not relabeled) subsequences the following convergences in the weak topology of $W^{1,p}(0,T;L^p(\Omega_1)) \cap L^m(0,T;W^{1,m}_D(\Omega_1))$ and in the strong topology of $C([0,T];L^p(\Omega_1))$ hold true

(i) $\lim_{\varepsilon\to0} u^*_{\delta,\varepsilon} = u^*_\delta$ *where $u^*_\delta \in \arg\min \mathcal{I}_\delta$ and solves (4.7),*
(ii) $\lim_{\delta\to0} u^*_{\delta,\varepsilon} = u_\varepsilon$ *where u_ε solves (4.6),*
(iii) $\lim_{\delta+\varepsilon\to0} u^*_{\delta,\varepsilon} = u$ *where u can be identified with a function U on ω solving the following lower-dimensional equation*

$$\alpha(\dot U) - \nabla' \cdot \big(E_0'(\nabla' U)\big) + h(U) = G(t) \qquad \text{in } \omega.$$

Appendix A

Lemma A.1 (Diagonal extraction for weakly converging sequences)**.** *Let $u_n \to u$ weakly in a separable and reflexive Banach space B and, for all n, let $u_{m,n} \to u_n$ weakly in B. Moreover, assume that $\|u_{m,n}\|_B \le C$, where C is a positive constant independent of m and n. Then, there exists a subsequence $u_{m_k,n_k} \to u$ weakly in B.*

Proof. As B is separable and reflexive, so is its dual B^*. Let $\{b_j\}_{j=1}^{\infty}$ be a dense subset of B^*. Choose u_{m_k,n_k} such that

$$|\langle b_j, u_{m_k,n_k} - u\rangle_B| \le 1/k \text{ for all } j \le k.$$

Thus, the sequence $u_{m_k,n_k} \to u$ weakly in B. Indeed, let $b \in B^*$, then for all $\varepsilon > 0$ there exists $j \in \mathbb{N}$ such that

$$\|b_j - b\|_{B^*} \le \frac{\varepsilon}{2(C + \|u\|_B)}$$

and $k \in \mathbb{N}$ such that $k \ge j$ and $1/k \le \varepsilon/2$. Thus,

$$\begin{aligned} |\langle b, u_{m_k,n_k} - u\rangle_B| &\le |\langle b_j, u_{m_k,n_k} - u\rangle_B| + |\langle b - b_j, u_{m_k,n_k} - u\rangle_B| \\ &\le \frac{1}{k} + \|b - b_j\|_{B^*} \|u_{m_k,n_k} - u\|_B \\ &\le \frac{\varepsilon}{2} + \frac{\varepsilon}{2(C + \|u\|_B)} \left(\|u_{m_k,n_k}\|_B + \|u\|_B\right) \le \varepsilon. \end{aligned}$$

□

Acknowledgements. The authors thank Alexander Mielke for stimulating discussions and helpful comments. M.L. gratefully acknowledges the support by the Einstein Center for Mathematics Berlin *via* MATHEON projects SE2 and SE18. S.M. is supported by the Austrian Science Fund (FWF) project P27052-N25.

References

[1] G. Akagi and S. Melchionna, Elliptic-regularization of nonpotential perturbations of doubly-nonlinear flows of nonconvex energies: a variational approach. *J. Convex Anal.* **25** (2018) 861–898.
[2] G. Akagi and U. Stefanelli, A variational principle for doubly nonlinear evolution. *Appl. Math. Lett.* **23** (2010) 1120–1124.
[3] G. Akagi and U. Stefanelli, Weighted energy-dissipation functionals for doubly-nonlinear evolution. *J. Funct. Anal.* **260** (2011) 2541–2578.
[4] G. Akagi and U. Stefanelli, Doubly nonlinear equations as convex minimization. *SIAM J. Math. Anal.* **46** (2014) 1922–1945.
[5] G. Akagi and U. Stefanelli, A variational principle for gradient flows of nonconvex energies. *J. Convex Anal.* **23** (2016) 53–75.
[6] H. Attouch, *Variational Convergence for Functions and Operators.* Pitman, Boston (1968).
[7] S. Aizicovici and Q. Yan, Convergence theorems for abstract doubly nonlinear differential equations. *Panamer. Math. J.* **7** (1997) 1–17.
[8] Andrea Braides, A handbook of Γ-convergence, in Vol. 3 of *Handbook of Differential Equations: Stationary Partial Differential Equations*, edited by M. Chipot and P. Quittner. North-Holland (2006) 101–213.
[9] P. Colli, On some doubly nonlinear evolution equations in Banach spaces. *Jpn J. Ind. Appl. Math.* **9** (1992) 181–203.
[10] P. Colli and A. Visintin, On a class of doubly nonlinear evolution equations. *Commun. Partial Differ. Equ.* **15** (1990) 737–756.
[11] T. Ilmanen. Elliptic regularization and partial regularity for motion by mean curvature. Memoirs of the American Mathematical Society (1994) 108:520x:+90
[12] H. Le Dret and A. Raoult, The nonlinear membrane model as variational limit of nonlinear three-dimensional elasticity. *J. Math. Pures Appl.* **74** (1995) 549–578.
[13] M. Liero and U. Stefanelli, A new minimum principle for Lagrangian mechanics. *J. Nonlinear Sci.* **23** (2013) 179–204.
[14] M. Liero and U. Stefanelli, Weighted inertia-dissipation-energy functionals for semilinear equations. *Boll. Unione Mat. Ital.* **6** (2013) 1–27.
[15] J.-L. Lions and E. Magenes, Problèmes aux limites non homogènes et applications. Travaux et Recherches Mathématiques. Dunod, Paris (1968).
[16] S. Melchionna, A variational principle for nonpotential perturbations of gradient flows of nonconvex energies. *J. Differ. Equ.* **262** (2017) 3737–3758.
[17] A. Mielke, On evolutionary Γ-convergence for gradient systems. *Lecture Notes in Applied Mathematics and Mechanics* (2016) 187–249.
[18] A. Mielke and M. Ortiz, A class of minimum principle for characterizing the trajectories of dissipative systems. *ESAIM: COCV* **14** (2008) 494–516.
[19] A. Mielke and U. Stefanelli, Weighted energy-dissipation functionals for gradient flows. *ESAIM: COCV* **17** (2011) 52–85.

[20] A. K. Nandakumaran and A. Visintin, Variational approach to homogenization of doubly-nonlinear flow in a periodic structure. *Nonlinear Analysis Series A* **120** (2015) 14–29.

[21] M. Ôtani, Nonmonotone perturbations for nonlinear parabolic equations associated with subdifferential operators, Cauchy problems. *J. Differ. Equ.* **46** (1982) 268–299.

[22] W. Rudin, Real and Complex Analysis, 3rd edn. McGraw Hill, New York (1987)

[23] E. Sandier and S. Serfaty, Gamma-convergence of gradient flows with applications to Ginzburg-Landau. *Comm. Pure Appl. Math.* LVII (2004) 1627–1672

[24] E. Serra and P. Tilli, Nonlinear wave equations as limit of convex minimization problems: proof of a conjecture by De Giorgi. *Ann. Math.* **175** (2012) 1511–1574.

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

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

[27] A. Visintin, Evolutionary Γ-Convergence of Weak Type. Preprint arXiv:1706.02172 (2017).