

Weierstraß-Institut
für Angewandte Analysis und Stochastik
Leibniz-Institut im Forschungsverbund Berlin e. V.

Preprint

ISSN 2198-5855

**The weighted energy-dissipation principle and evolutionary
 Γ -convergence for doubly nonlinear problems**

Matthias Liero¹, Stefano Melchionna²

submitted: June 26, 2017

¹ Weierstrass Institute
Mohrenstr. 39
10117 Berlin
Germany

E-Mail: matthias.liero@wias-berlin.de

² Faculty of Mathematics
University of Vienna
Oskar-Morgenstern-Platz 1
A-1090 Vienna
Austria

E-Mail: stefano.melchionna@univie.ac.at

No. 2411
Berlin 2017



2010 *Mathematics Subject Classification.* 58E30, 35K55, 47J35.

Key words and phrases. Doubly nonlinear evolution, weighted-energy-dissipation principle, evolutionary Γ -convergence, variational principle.

The work of M.L. is supported by Einstein Center for Mathematics Berlin under project MATHEON SE2. S.M. is supported by the Austrian Science Fund (FWF) project P27052-N25.

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Leibniz-Institut im Forschungsverbund Berlin e. V.
Mohrenstraße 39
10117 Berlin
Germany

Fax: +49 30 20372-303
E-Mail: preprint@wias-berlin.de
World Wide Web: <http://www.wias-berlin.de/>

The weighted energy-dissipation principle and evolutionary Γ -convergence for doubly nonlinear problems

Matthias Liero, Stefano Melchionna

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 minimizer 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 $\delta \rightarrow 0$, $\varepsilon \rightarrow 0$, as well as $\delta + \varepsilon \rightarrow 0$ 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 $\varepsilon \rightarrow 0$, in the case of quadratic dissipation potentials and uniformly λ -convex energy functionals. Finally, we discuss a homogenization problem as an example of application.

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 $\varepsilon > 0$

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

Here \dot{u} denotes the time derivative of the unknown $u : [0, T] \rightarrow H$, where H is a reflexive Banach space. The evolution is driven by a dissipation functional $\psi_\varepsilon : H \rightarrow [0, \infty)$, assumed to be convex and Gâteaux differentiable, by an energy functional $\phi_\varepsilon : H \rightarrow (-\infty, \infty]$ 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 (\mathbf{P}_ε) 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. Problems of this type have been studied by several authors, and we refer to [Col92, CV90, Ô82] 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 losing 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 (\mathbf{P}_ε) are stable with respect to these perturbations. In other words, is it possible to prove convergence of the solutions to (\mathbf{P}_ε) to solutions of an effective

equation assuming convergence of the energy and dissipation functionals and of the data in some proper sense? This question has been formalized in the notion of *evolutionary Γ -convergence* (E-convergence) introduced by Mielke in [Mie16] 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. [SS04]. In particular, the convergence results in the survey paper [Mie16] 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.

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. [AS14, AS16, AM17]. 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

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

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

$$(P_{\varepsilon, \delta}) \quad \begin{aligned} -\delta \frac{d}{dt} d\psi_\varepsilon(\dot{u}) + d\psi_\varepsilon(\dot{u}) + \partial\phi_\varepsilon(u) &\ni g(t) \quad \text{a.e. in } (0, T), \\ u(0) = u_\varepsilon^0, \quad \delta d\psi_\varepsilon(\dot{u}(T)) &= 0. \end{aligned} \quad (1.2)$$

Second, minimizers (i.e. solutions to (1.2)) are proved 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 $\delta \rightarrow 0$ is usually referred to as *causal limit*.

The WED principle, originally proposed by Ilmanen [Ilm94] (see also [LM68]), has been brought to new attention by Mielke and Ortiz [MO08] 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 [MS11]. Akagi and Stefanelli have extended the theory to the genuinely nonconvex case for gradient flows [AS16] and to convex doubly nonlinear systems [AS14]. Moreover, an analogous approach has been applied to some hyperbolic problems, e.g., the semilinear wave equation [LS13b, ST12, Ste11], and to Lagrangian Mechanics [LS13a]. Recently, nonpotential perturbation problems have also been considered [Mel17, AM17].

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 $\phi_\varepsilon = \varphi_\varepsilon^1 - \varphi_\varepsilon^2$, where $\varphi_\varepsilon^1, \varphi_\varepsilon^2 : V \rightarrow (-\infty, \infty]$ are proper, lower semicontinuous, and convex functionals on a Banach space $V \subset H$ with φ_ε^1 dominating φ_ε^2 in a suitable sense. This will allow us to define $\partial\phi_\varepsilon$ as the difference of the convex subdifferentials $\partial\varphi_\varepsilon^1$ and $\partial\varphi_\varepsilon^2$ (see Theorem 2.1 for a more rigorous definition). These assumptions along with suitable uniform growth conditions for $\varphi_\varepsilon^1, \varphi_\varepsilon^2$, 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 $\psi_\varepsilon, \phi_\varepsilon$, and on the data u_ε^0 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 Subsection (2.2)), as well as the convergence of $g_\varepsilon, u_\varepsilon^0$ 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). 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 $\delta + \varepsilon \rightarrow 0$ 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 $\delta \rightarrow 0$ is highly degenerate (intuitively it is just a constraint on the initial condition). Thus, the limit $\delta + \varepsilon \rightarrow 0$ 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. Lemma 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 [AS14]. More precisely, the authors proved Mosco convergence of the WED functionals in the case $\varphi_\varepsilon^2 \equiv 0$ and $g_\varepsilon \equiv 0$ assuming Mosco convergence of φ_ε^1 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 [Mie16] and means that in additions to $u_\varepsilon^0 \rightarrow u_0^0$ in H also $\phi_\varepsilon(u_\varepsilon^0) \rightarrow \phi_0(u_0^0)$ 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., e.g., the homogenization example in Section 4).

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 [AM17] (see Subsection 2.1). Next, we fix conditions on the convergence of the “static” functionals ϕ_ε and ψ_ε (Subsection 3.2) that allow us to prove 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, in Section 4, we discuss the example of deriving the homogenized limit for a nonlinear parabolic equation with rapidly oscillating coefficient functions.

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 [MS11, AS14] 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 $\psi : H \rightarrow [0, \infty)$, which is assumed to be Gâteaux differentiable. In particular, we denote by $d\psi(v) \in H^*$ the Gâteaux differential for $v \in H$. The energy functional $\phi : H \rightarrow \mathbb{R}_\infty := \mathbb{R} \cup \{+\infty\}$ is assumed to be of the form $\phi = \varphi^1 - \varphi^2$, where $\varphi^1, \varphi^2 : H \rightarrow [0, \infty]$ 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 $\tilde{\phi} : [0, T] \times H \rightarrow \mathbb{R}_\infty$ by setting

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

We impose the following growth assumptions on the dissipation and energy functional which are in accordance with [AM17]. In particular, we refer to [AM17, Remark 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

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

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

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

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

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

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

Additionally, as in [AS16, AM17] we make the assumption that there exist $\kappa \in (0, 1)$ and $c > 0$ such that we have

$$(A5) \quad \varphi^2(u) \leq \kappa(\varphi^1(u) + c) \quad \text{for all } u \in D(\varphi^1), \quad (2.4a)$$

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

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

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

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

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

and introduce for a (fixed) constant $\delta > 0$ the *Weighted-Energy-Dissipation functional* (WED functional) via

$$\mathcal{I}_\delta(u) = \begin{cases} \int_0^T e^{-t/\delta} \left(\psi(\dot{u}) + \frac{1}{\delta} \tilde{\phi}(t, u) \right) dt & \text{if } u \in \mathcal{K}(u^0), \\ \infty & \text{else.} \end{cases} \quad (2.6)$$

The WED principle is concerned with finding a minimizer $u_\delta \in \mathcal{K}(u^0)$ and passing to the limit $\delta \rightarrow 0$ to recover a solution u of the original doubly nonlinear equation. In particular, in [AM17] the following result was proven.

Theorem 2.1 (Akagi–Melchionna [AM17, Theorem 13]) *1 Assume that (A1)–(A8) are satisfied.*

Then for every $\delta > 0$ sufficiently small the WED functional \mathcal{I}_δ 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*

$$-\delta \dot{\xi}_\delta + \xi_\delta + \eta_\delta^1 - \eta_\delta^2 = g(t), \quad \text{in } X^* \text{ a.e. in } (0, T), \\ \text{and } u_\delta(0) = u^0, \quad \delta \xi_\delta(T) = 0, \quad (2.7)$$

where $\xi_\delta = d\psi(\dot{u}_\delta)$ 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, $u_\delta \rightarrow u$ 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

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

with $\xi = d\psi(\dot{u})$ 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 $\phi_\varepsilon = \varphi_\varepsilon^1 - \varphi_\varepsilon^2$ converge in the following sense:

$$(\Gamma 1) \quad \varphi_\varepsilon^1 \xrightarrow{M} \varphi_0^1 \quad \text{in } H, \quad (\Gamma 2) \quad \varphi_\varepsilon^2 \xrightarrow{C} \varphi_0^2 \quad \text{weakly in } X. \quad (2.9)$$

Here, “ \xrightarrow{M} ” denotes Mosco convergence (Γ -convergence with respect to the strong *and* weak topology) and “ \xrightarrow{C} ” means continuous convergence, i.e. $\varphi_\varepsilon^2(u_\varepsilon) \rightarrow \varphi_0^2(u^0)$ for all sequences with $u_\varepsilon \rightarrow u^0$ weakly in X . In particular, in the case of uniform coercivity of φ_ε^1 on X the convergence in $(\Gamma 1)$ is equivalent to $\varphi_\varepsilon^1 \xrightarrow{\Gamma} \varphi_0^1$ weakly in X , see e.g. [Mie16, Prop. 2.5].

Additionally, for the dissipation potentials ψ_ε we assume that

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

On the forcing terms we impose

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

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

$$(\Gamma 5) \quad \phi_\varepsilon(u_\varepsilon^0) \leq C. \quad (2.12)$$

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

$$u_\varepsilon^0 \rightarrow u^0 \quad \text{weakly in } X \text{ (strongly in } H).$$

Theorem 2.2 (Γ -convergence of WED functionals) *Let us assume that φ_ε^1 , φ_ε^2 , 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*

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

Γ -converge to

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

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

In particular, let $u_{\delta,\varepsilon}^*$ be a minimizer of $\mathcal{I}_{\delta,\varepsilon}$, then, $u_{\delta,\varepsilon}^* \rightarrow u_\delta^*$ 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 $e_{\varepsilon,\delta} : [0, T_\varepsilon] \rightarrow [0, \infty)$ be a family of L^∞ functions uniformly convergent to $t \mapsto e^{-t/\delta}$ for fixed $\delta > 0$. Then, the WED functionals*

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

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 [Mie16]. 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 solutions to the problem*

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

converge (up to subsequences) to solutions to

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

Note that both equation (P_{δ,ε}) and (P_{δ,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 (P_{δ,ε}) such that $\{u_\varepsilon\}$ converges, up to subsequences, to a solution u to (P_{δ,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 (P_{δ,ε}) and (P_{δ,0}) are intended in the strong sense defined in Theorem 2.1.

2.3 The joint limit $\varepsilon + \delta \rightarrow 0$

The crucial question now is whether it is possible to consider the joint limit $\varepsilon + \delta \rightarrow 0$. The main result in this section states that minimizers $u_{\delta,\varepsilon}^*$ of the WED functional $\mathcal{I}_{\delta,\varepsilon}$, i.e. strong solutions to (P_{δ,ε}), converge (up to subsequences $(\delta_k, \varepsilon_k) \rightarrow (0, 0)$) to solutions to the target problem (2.8). We remark that the limit $\delta + \varepsilon \rightarrow 0$ is meaningless in the sense of Γ -convergence of the WED functionals, since the Γ -limit of $\mathcal{I}_{\delta,\varepsilon}$ for $\delta \rightarrow 0$ is highly degenerate, see [MO08, Sect. 4.3], and provides only little information on the limiting solutions.

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 minimizer of the WED functional $\mathcal{I}_{\varepsilon,\delta}$. Assume additionally that for $\varepsilon \geq 0$ we have $g_\varepsilon = g_\varepsilon^1 + g_\varepsilon^2$ with*

$$g_\varepsilon^1 \in W^{1,p'}(0, T; H^*) \text{ and } g_\varepsilon^1 \rightarrow g_0^1 \text{ weakly in } W^{1,p'}(0, T; H^*), \quad (2.14)$$

$$g_\varepsilon^2 \rightarrow g_0^2 \text{ strongly in } L^{p'}(0, T; H^*), \quad (2.15)$$

and well preparedness of initial data, i.e.

$$\phi_\varepsilon(u_\varepsilon^0) \rightarrow \phi(u^0). \quad (2.16)$$

Then, there exists a subsequence $(\varepsilon_k, \delta_k)$ such that $\varepsilon_k + \delta_k \rightarrow 0$ and $u_{\varepsilon_k, \delta_k}^ \rightarrow u$, 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 $\varepsilon \rightarrow 0$ with $\delta > 0$ fixed. However, although the assumptions (2.14)-(2.16) refer to the convergence of the data for $\varepsilon \rightarrow 0$ they are fundamental for the limit $\delta + \varepsilon \rightarrow 0$.*

Indeed, the limit $\delta + \varepsilon \rightarrow 0$ is computed at the level of the equations $(P_{\delta,\varepsilon})$ and not for the functionals $\mathcal{I}_{\delta,\varepsilon}$. To compute this limit we first derive uniform estimates on each term in $(P_{\delta,\varepsilon})$. 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 $(P_{\delta,\varepsilon})$. The next step is to identify the limits of the nonlinear terms, in particular, to prove $\lim_{\varepsilon \rightarrow 0} d\psi_\varepsilon(\dot{u}_{\delta,\varepsilon}^) = d\psi_0(\dot{u})$.*

By taking advantage of the convexity of the dissipation potential ψ_0 , we only have to prove an upper estimate on $\limsup_{\varepsilon \rightarrow 0} \langle d\psi_\varepsilon(\dot{u}_{\delta,\varepsilon}^), \dot{u}_{\delta,\varepsilon}^* \rangle_H$, see (3.25). Since our estimates provide only weak convergence of both $d\psi_\varepsilon(\dot{u}_{\delta,\varepsilon}^*)$ and $\dot{u}_{\delta,\varepsilon}^*$, we substitute $d\psi_\varepsilon(\dot{u}_{\delta,\varepsilon}^*)$ by using equation $(P_{\delta,\varepsilon})$. As a consequence we have to handle the two terms $\langle \partial\phi_\varepsilon(u_{\delta,\varepsilon}^*), \dot{u}_{\delta,\varepsilon}^* \rangle_H$ and $\langle g_\varepsilon, \dot{u}_{\delta,\varepsilon}^* \rangle_H$. 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)-(2.15). Note that this procedure is necessary only in the case that the term $d\psi_\varepsilon$ 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 $\|u_{\delta,\varepsilon}^* - u_{\delta,0}^*\|$ 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. [MS11]). 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 $\mathbb{A}_\varepsilon \in \text{Lin}(H, H)$ 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.

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

where λ is independent of ε .

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

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

In the case that ϕ is λ -convex, we can identify $\partial_H^F \phi(u) = \partial_H \phi^\lambda(u) + \lambda u$, where $\phi^\lambda(u) = \phi(u) - \lambda \|u\|_H^2/2$ is convex. In particular, we have $D(\phi) = D(\phi^\lambda)$ and $D(\partial_H^F \phi) = D(\partial_H \phi^\lambda)$. 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

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

Note that in concrete applications the space B is typically an interpolation space between H and X (see e.g. Section 4).

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

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

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 $\ell : [0, \infty) \rightarrow [0, \infty)$, and a positive constant C , such that

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

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

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

Furthermore, we assume that there exists an operator $\mathfrak{S}_\varepsilon : D(\phi_\varepsilon) \rightarrow D(\phi_0)$, 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 $\ell : [0, \infty) \rightarrow [0, \infty)$, and a positive constant C , such that

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

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

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

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

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

In Section 4 we will discuss a typical example for the developed theory, namely, the homogenization of a parabolic equation. Since the operator \mathfrak{S}_ε 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*

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

where

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

and

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

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

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

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

Proof. Note that, as a consequence of the Mosco and continuous convergence of φ_ε^1 and φ_ε^2 , respectively, in (2.9), and of the coercivity condition in (2.2b), we have the Mosco convergence $\phi_\varepsilon \xrightarrow{M} \phi$

in H . Consider now a sequence satisfying $u_\varepsilon \rightarrow u$ weakly in $W^{1,p}(0, T; H) \cap L^m(0, T; X)$. In particular, by standard embedding results we also have

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

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

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

see also [Ste08]. Moreover, thanks to (2.11) and to the strong convergence (3.1), we have

$$\lim_{\varepsilon \rightarrow 0} \int_0^T e^{-t/\delta} \langle g_\varepsilon(t), u_\varepsilon(t) \rangle_H dt = \int_0^T e^{-t/\delta} \langle g_0(t), u(t) \rangle_H dt.$$

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

$$\begin{aligned} \liminf_{\varepsilon \rightarrow 0} \int_0^T e^{-t/\delta} \psi_\varepsilon(\dot{u}_\varepsilon) dt &\geq \liminf_{\varepsilon \rightarrow 0} \sum_{i=1}^N e^{-i\tau/\delta} \tau \int_{(i-1)\tau}^{i\tau} \psi_\varepsilon(\dot{u}_\varepsilon) \frac{dt}{\tau} \\ &\geq \liminf_{\varepsilon \rightarrow 0} \sum_{i=1}^N e^{-i\tau/\delta} \tau \psi_\varepsilon \left(\int_{(i-1)\tau}^{i\tau} \dot{u}_\varepsilon \frac{dt}{\tau} \right), \end{aligned} \quad (3.2)$$

where we used $e^{-i\tau/\delta} \leq e^{-t/\delta}$ 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 $\varepsilon \rightarrow 0$ to arrive at

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

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

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

Since ψ_0 is lower semicontinuous and convex, we can pass to the limit $\tau \rightarrow 0$ 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 $u \in \mathcal{K}(u^0)$ and there exists a subsequence $\varepsilon_k \rightarrow 0$ and a sequence $u_{\varepsilon_k} \rightarrow u$ strongly in $W^{1,p}(0, T; H)$ and weakly in $L^m(0, T; X)$ such that*

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

Proof. Our proof follows the lines of [AS11, Thm. 7.2]. 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 $i = 0, \dots, N$. Let \widehat{u}_τ , and \bar{u}_τ be the piecewise affine and piecewise forward constant interpolants of $\{u_\tau^i\}_{i=0}^N$ defined via

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

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

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

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

$$\bar{u}_\tau \rightarrow u \text{ strongly in } L^\infty(0, T; X). \quad (3.5b)$$

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

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

and (see [AS11, Sect. 6.1.2])

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

where $h_\tau^1 := \int_0^T \frac{e^{-t/\delta}}{\delta} (\phi_0(\widehat{u}_\tau) - \phi_0(\bar{u}_\tau)) dt$.

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

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

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

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

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

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 $u_{\tau,\varepsilon}^0 := u_\varepsilon^0$. Due to (2.9) and the coercivity of ϕ_ε in X we find sequences $u_{\tau,\varepsilon}^i$ for every $i \in \{1, \dots, N\}$ such that

$$u_{\tau,\varepsilon}^i \rightarrow u_\tau^i \text{ weakly in } X \quad \text{and} \quad \varphi_\varepsilon^1(u_{\tau,\varepsilon}^i) \rightarrow \varphi_0^1(u_\tau^i).$$

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

$$\begin{aligned} \psi_\varepsilon \left(\frac{u_{\tau,\varepsilon}^i - u_{\tau,\varepsilon}^{i-1}}{\tau} \right) &\rightarrow \psi_0 \left(\frac{u_\tau^i - u_\tau^{i-1}}{\tau} \right) \text{ for all } i \in \{1, \dots, N\}, \\ \varphi_\varepsilon^2(u_{\tau,\varepsilon}^i) &\rightarrow \varphi_0^2(u_\tau^i) \text{ for all } i \in \{0, \dots, N\}. \end{aligned}$$

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

$$\widehat{u}_{\tau,\varepsilon} \rightarrow \widehat{u}_\tau \text{ weakly in } L^m(0, T; X), \quad (3.7)$$

$$\widehat{u}_{\tau,\varepsilon} \rightarrow \widehat{u}_\tau \text{ strongly in } W^{1,p}(0, T; H), \quad (3.8)$$

$$\overline{u}_{\tau,\varepsilon} \rightarrow \overline{u}_\tau \text{ weakly in } L^m(0, T; X). \quad (3.9)$$

We now claim that the following upper estimate holds

$$\mathcal{I}_\varepsilon(\widehat{u}_{\tau,\varepsilon}) \leq \mathcal{I}_0(\widehat{u}_\tau) + r_{\tau,\varepsilon}^1 + r_\tau^2, \quad (3.10)$$

where

$$\lim_{\tau \rightarrow 0} \lim_{\varepsilon \rightarrow 0} r_{\tau,\varepsilon}^1 = 0 \quad \text{and} \quad \lim_{\tau \rightarrow 0} r_\tau^2 = 0.$$

Indeed, we compute

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

Clearly, we have $I_{\tau,\varepsilon} \rightarrow \mathcal{I}_0(\widehat{u}_\tau) + h_\tau^1$ for $\varepsilon \rightarrow 0$. Moreover, using the convexity of φ_ε^1 and the definition of the sequences $u_{\tau,\varepsilon}^i$ we arrive at

$$\begin{aligned} \lim_{\varepsilon \rightarrow 0} h_{\tau,\varepsilon}^2 &\leq \lim_{\varepsilon \rightarrow 0} \sum_{i=1}^{N-1} \int_{i\tau}^{(i+1)\tau} e^{-t/\delta} \frac{\alpha_\tau^i(t)}{\delta} \left(\varphi_\varepsilon^1(u_{\tau,\varepsilon}^i) - \varphi_\varepsilon^1(u_{\tau,\varepsilon}^{i+1}) \right) dt \\ &\quad + \lim_{\varepsilon \rightarrow 0} \int_0^\tau \left(e^{-t/\delta} \frac{\tau-t}{\tau} \right) \left(\varphi_\varepsilon^1(u_\varepsilon^0) - \varphi_\varepsilon^1(u_{\tau,\varepsilon}^1) \right) dt \\ &= \sum_{i=1}^{N-1} \int_{i\tau}^{(i+1)\tau} e^{-t/\delta} \frac{\alpha_\tau^i(t)}{\delta} \left(\varphi_0^1(u_\tau^i) - \varphi_0^1(u_\tau^{i+1}) \right) dt + \lim_{\varepsilon \rightarrow 0} h_{\tau,\varepsilon}^{2,2}. \end{aligned}$$

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

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

whose right-hand side tends to zero as $\tau \rightarrow 0$. 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

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

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

It remains to show that also $h_{\tau,\varepsilon}^3 \rightarrow 0$ as $\varepsilon \rightarrow 0$ and $\tau \rightarrow 0$. For this, note that for all $t \in [0, T]$ one has $\widehat{u}_{\tau,\varepsilon}(t) \rightarrow \widehat{u}_\tau(t)$ weakly in X . Thus, as a consequence of the continuous convergence in (2.9) we have $\lim_{\varepsilon \rightarrow 0} \varphi_\varepsilon^2(\widehat{u}_{\tau,\varepsilon}(t)) = \varphi_0^2(\widehat{u}_\tau(t))$. 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 [Rud87, Pag. 133] to deduce

$$\lim_{\varepsilon \rightarrow 0} h_{\tau,\varepsilon}^3 = \int_0^T \frac{e^{-t/\delta}}{\delta} (\varphi_0^2(\widehat{u}_\tau) - \varphi_0^2(\bar{u}_\tau)) dt.$$

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

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

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

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

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

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

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

and hence, by convexity,

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

Therefore, without loss of generality, we can assume

$$\|\widehat{u}_{\tau,\varepsilon}\|_{L^m(0,T;X)}^m \leq \widetilde{C} \left(\int_0^T \varphi_0^1(u) dt + T \right) \quad (3.11)$$

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

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

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

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

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

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

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

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

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

It remains to show that also $\widehat{u}_{\tau, \varepsilon_\tau} \rightarrow u$ 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

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

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, $\widehat{u}_{\tau, \varepsilon_\tau} \rightarrow u$ 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

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

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

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

Thus, it is possible to apply Lemma 5.1 and extract a subsequence u_{ε_k, n_k} such that $u_{\varepsilon_k, n_k} \rightarrow u$ strongly in $W^{1,p}(0, T; H)$ and weakly in $L^m(0, T; X)$ and such that the limsup-inequality is satisfied. As a consequence of the liminf-estimate obtained in Proposition 3.1, we conclude that $\lim_{\varepsilon_k \rightarrow 0} \mathcal{I}_{\varepsilon_k}(u_{\varepsilon_k, n_k}) = \mathcal{I}_0(u)$. ■

Theorem 2.2 can be now proved by simply combining the two propositions above.

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

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

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 $t \mapsto e^{-t/\delta}$ and to the convergence of T_ε , we have that the second and the third terms in (3.15) vanish as $\varepsilon \rightarrow 0$ 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}_ε 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 $u_{\varepsilon,\delta} := u_{\varepsilon,\delta}^*$ denote the minimizer of the WED functional $\mathcal{I}_{\varepsilon,\delta}$, which solves the Euler–Lagrange system as stated in Theorem 2.1 (see also [AS11, AM17])

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

where $\xi_{\varepsilon,\delta} = d\psi_\varepsilon(\dot{u}_{\varepsilon,\delta}) \in L^{p'}(0, T; H^*)$, $\eta_{\varepsilon,\delta}^1 \in L^{m'}(0, T; X^*)$, $\eta_{\varepsilon,\delta}^2 \in L^{p'}(0, T; H^*)$ with $\eta_{\varepsilon,\delta}^1(t) \in \partial_X \varphi_\varepsilon^1(u_{\varepsilon,\delta}(t))$, and $\eta_{\varepsilon,\delta}^2(t) \in \partial_H \varphi_\varepsilon^2(u_{\varepsilon,\delta}(t))$ for almost every $t \in (0, T)$.

In order to prove Theorem 2.5, we want to pass to the limit $\delta + \varepsilon \rightarrow 0$ in (3.16). To this aim we recall that $u_{\varepsilon,\delta}$ satisfies the following uniform estimate (cf. [AM17, Sec. 4.4], see also [AS11] for the convex energy case)

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

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

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

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

$$\xi_{\varepsilon,\delta} \rightarrow \xi \text{ weakly in } L^{p'}(0, T; H^*), \quad (3.18c)$$

$$\delta \xi_{\varepsilon,\delta} \rightarrow 0 \text{ strongly in } L^{p'}(0, T; H^*), \quad (3.18d)$$

$$\eta_{\varepsilon,\delta}^1 \rightarrow \eta^1 \text{ weakly in } L^{m'}(0, T; X^*), \quad (3.18e)$$

$$\eta_{\varepsilon,\delta}^2 \rightarrow \eta^2 \text{ weakly in } L^{p'}(0, T; H^*), \quad (3.18f)$$

$$\delta \dot{\xi}_{\varepsilon,\delta} \rightarrow 0 \text{ weakly in } L^{p'}(0, T; H^*) + L^{m'}(0, T; X^*). \quad (3.18g)$$

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 $g_\varepsilon \rightarrow g_0$ weakly in $L^{p'}(0, T; H^*)$ by assumption (2.11). Thus, we can pass to the limit $\delta + \varepsilon \rightarrow 0$ to obtain the limiting equation

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

$$u(0) = u^0. \quad (3.20)$$

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 .

Identification of $\eta^2 \in \partial_H \varphi_0^2(u)$. Let $w \in D(\varphi_0^2)$ and let $\rho \in C_c^\infty(0, T)$, $\rho \geq 0$. By the definition of the subdifferential, we have

$$\int_0^T \varphi_\varepsilon^2(w) \rho(t) dt + \int_0^T \langle \eta_{\varepsilon,\delta}^2(t), u_{\varepsilon,\delta}(t) - w \rangle_H \rho(t) dt \geq \int_0^T \varphi_\varepsilon^2(u_{\varepsilon,\delta}(t)) \rho(t) dt \quad (3.21)$$

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

$$\lim_{\varepsilon \rightarrow 0} \int_0^T \langle \eta_{\varepsilon,\delta}^2(t), u_{\varepsilon,\delta}(t) - w \rangle_H \rho(t) dt = \int_0^T \langle \eta^2(t), u(t) - w \rangle_H \rho(t) dt.$$

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

$$\int_0^T \varphi_\varepsilon^2(w) \rho(t) dt \rightarrow \int_0^T \varphi_0^2(w) \rho(t) dt \quad \text{and} \quad \int_0^T \varphi_\varepsilon^2(u_{\varepsilon,\delta}(t)) \rho(t) dt \rightarrow \int_0^T \varphi_0^2(u(t)) \rho(t) dt.$$

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

$$\int_0^T \varphi_0^2(w) \rho(t) dt + \int_0^T \langle \eta^2(t), u(t) - w \rangle_H \rho(t) dt \geq \int_0^T \varphi_0^2(u(t)) \rho(t) dt.$$

As ρ and w are arbitrary, we deduce $\eta^2(t) \in \partial_H \varphi_0^2(u(t))$ for a.a. $t \in (0, T)$.

Identification of $\eta^1 \in \partial_X \varphi_X^1(u)$. As before, let $w \in D(\varphi_0^1)$ and let $\rho \in C_c^\infty(0, T)$ with $\rho \geq 0$. Let now $\{w_\varepsilon\} \subset X$ be a recovery sequence for φ_0^1 at w , i.e.

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

and $\varphi_\varepsilon^1(w_\varepsilon) \rightarrow \varphi_0^1(w)$. By the definition of the subdifferential, we have

$$\int_0^T \varphi_\varepsilon^1(w_\varepsilon) \rho(t) dt + \int_0^T \langle \eta_{\varepsilon,\delta}^1(t), u_{\varepsilon,\delta}(t) - w_\varepsilon \rangle_X \rho(t) dt \geq \int_0^T \varphi_\varepsilon^1(u_{\varepsilon,\delta}(t)) \rho(t) dt. \tag{3.23}$$

As w_ε is a recovery sequence, we have convergence

$$\int_0^T \varphi_\varepsilon^1(w_\varepsilon) \rho(t) dt \rightarrow \int_0^T \varphi_0^1(w) \rho(t) dt.$$

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

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

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

$$\int_0^T \langle \eta_{\varepsilon,\delta}^2(t), u_{\varepsilon,\delta}(t) - w_\varepsilon \rangle_H \rho(t) dt \rightarrow \int_0^T \langle \eta^2(t), u(t) - w \rangle_H \rho(t) dt.$$

Integration by parts, together with the final condition for $\xi_{\varepsilon,\delta}$ in (3.16) and estimate (3.17) yield (see [AS11] for a rigorous derivation of the integration by parts formula)

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

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

$$\int_0^T \langle g_\varepsilon(t), u_{\varepsilon,\delta}(t) - w_\varepsilon \rangle_H \rho(t) dt \rightarrow \int_0^T \langle g_0(t), u(t) - w \rangle_H \rho(t) dt$$

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

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

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

$$\begin{aligned} \int_0^T \langle \eta_{\varepsilon,\delta}^1(t), u_{\varepsilon,\delta}(t) - w_\varepsilon \rangle_X \rho(t) dt &\rightarrow \int_0^T \langle \eta^2(t) + \xi(t) + g_0(t), u(t) - w \rangle_H \rho(t) dt \\ &= \int_0^T \langle \eta^1(t), u(t) - w \rangle_X \rho(t) dt. \end{aligned}$$

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

$$\liminf_{\delta+\varepsilon \rightarrow 0} \int_0^T \varphi_\varepsilon^1(u_{\varepsilon,\delta}(t)) \rho(t) dt \geq \int_0^T \varphi_0^1(u(t)) \rho(t) dt.$$

Combining all these facts from (3.23) we deduce

$$\int_0^T \varphi_0^1(w) \rho(t) dt + \int_0^T \langle \eta^1(t), u(t) - w \rangle_X \rho(t) dt \geq \int_0^T \varphi_0^1(u(t)) \rho(t) dt.$$

Hence, $\eta^1 \in \partial_X \varphi_0^1(u)$ a.e. in $(0, T)$. Moreover, since $\eta^1(t) \in H^*$ for a.a. $t \in (0, T)$, then $\eta^1(t) \in \partial_H \varphi_0^1(u(t))$ for a.a. $t \in (0, T)$ (cf. [AS11] for details).

Identification of $\xi = \mathbf{d}_H \psi_0(u')$. Let $v \in L^p(0, T; H)$. Again by the definition of the subdifferential, we have

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

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

$$\int_0^T \psi_\varepsilon(v(t)) dt \rightarrow \int_0^T \psi_0(v) dt.$$

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

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

The weak convergence in (3.18c) implies

$$\int_0^T \langle \xi_{\varepsilon,\delta}(t), v(t) \rangle_H dt \rightarrow \int_0^T \langle \xi(t), v(t) \rangle_H dt.$$

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

$$\begin{aligned} \limsup_{\varepsilon+\delta \rightarrow 0} \int_0^T \langle \xi_{\varepsilon,\delta}(t), \dot{u}_{\varepsilon,\delta}(t) \rangle_H dt &\leq \limsup_{\varepsilon+\delta \rightarrow 0} \int_0^T \langle \delta \dot{\xi}_{\varepsilon,\delta}(t) - \eta_{\varepsilon,\delta}^1(t), \dot{u}_{\varepsilon,\delta}(t) \rangle_H dt \\ &\quad + \limsup_{\varepsilon+\delta \rightarrow 0} \int_0^T \langle \eta_{\varepsilon,\delta}^2(t), \dot{u}_{\varepsilon,\delta}(t) \rangle_H + \langle g_\varepsilon(t), \dot{u}_{\varepsilon,\delta}(t) \rangle_H dt. \end{aligned}$$

Note that, as $\xi_{\varepsilon,\delta} = d\psi_\varepsilon(\dot{u}_{\varepsilon,\delta}) \in \partial_H \psi_\varepsilon(\dot{u}_{\varepsilon,\delta})$, then $\dot{u}_{\varepsilon,\delta} \in \partial_{H^*} \psi_\varepsilon^*(\xi_{\varepsilon,\delta})$, where $\psi_\varepsilon^* : H^* \rightarrow [0, \infty)$ 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. [AS11] for a rigorous proof and more details), we get

$$\begin{aligned} \int_0^T \langle \delta \dot{\xi}_{\varepsilon,\delta}(t) - \eta_{\varepsilon,\delta}^1(t), \dot{u}_{\varepsilon,\delta}(t) \rangle_H dt &\leq \int_0^T \left(\delta \frac{d}{dt} \psi_\varepsilon^*(\xi_{\varepsilon,\delta}(t)) - \frac{d}{dt} \varphi_\varepsilon^1(u_{\varepsilon,\delta}(t)) \right) dt \\ &\leq -\delta \psi_\varepsilon^*(\xi_{\varepsilon,\delta}(0)) - \varphi_\varepsilon^1(u_{\varepsilon,\delta}(T)) + \varphi_\varepsilon^1(u_\varepsilon^0) \\ &\leq -\varphi_\varepsilon^1(u_{\varepsilon,\delta}(T)) + \varphi_\varepsilon^1(u_\varepsilon^0). \end{aligned}$$

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

$$\begin{aligned} \limsup_{\varepsilon+\delta \rightarrow 0} \int_0^T \langle \delta \dot{\xi}_{\varepsilon,\delta}(t) - \eta_{\varepsilon,\delta}^1(t), \dot{u}_{\varepsilon,\delta}(t) \rangle_H dt &\leq -\liminf_{\varepsilon+\delta \rightarrow 0} \varphi_\varepsilon^1(u_{\varepsilon,\delta}(T)) + \lim_{\varepsilon+\delta \rightarrow 0} \varphi_\varepsilon^1(u_\varepsilon^0) \\ &= -\varphi_0^1(u(T)) + \varphi_0^1(u^0). \end{aligned}$$

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

$$-\varphi_0^1(u(T)) + \varphi_0^1(u^0) = \int_0^T \langle \eta^1(t), \dot{u}(t) \rangle_H dt.$$

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

$$\begin{aligned} \limsup_{\varepsilon+\delta \rightarrow 0} \int_0^T \langle \eta_{\varepsilon,\delta}^2(t), \dot{u}_{\varepsilon,\delta}(t) \rangle_H dt \\ = \lim_{\varepsilon+\delta \rightarrow 0} \varphi_\varepsilon^2(u_{\varepsilon,\delta}(T)) - \lim_{\varepsilon+\delta \rightarrow 0} \varphi_\varepsilon^2(u_\varepsilon^0) = \varphi_0^2(u(T)) - \varphi_0^2(u^0) = \int_0^T \langle \eta^2(t), \dot{u}(t) \rangle_H dt. \end{aligned}$$

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

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

Combining the above estimates and (3.25), we conclude

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

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

$$\langle d_{L^p(0,T;H)} \bar{\psi}_0(v), w \rangle_{L^p(0,T;H)} \leq \int_0^T \langle d_H \psi_0(v(t)), w(t) \rangle_H dt \text{ for all } v, w \in L^p(0, T; H).$$

In particular, $\xi \in \partial_{L^p(0,T;H)} \bar{\psi}_0(\dot{u})$ implies $\xi(t) = d_H \psi_0(\dot{u}(t))$ 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 $I_\varepsilon, I_0 : \mathcal{V} \rightarrow (-\infty, \infty]$ 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 $\varepsilon \rightarrow 0$ 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*

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

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

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

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

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

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

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

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

Proof. We estimate

$$\begin{aligned} \|u_\varepsilon^* - u^*\|_{\mathcal{V}} &\leq \|u_\varepsilon^R - u^*\|_{\mathcal{V}} + \|u_\varepsilon^* - u_\varepsilon^R\|_{\mathcal{V}} \\ &\stackrel{(3.26)+(3.29)}{\leq} \rho_\varepsilon^1 + \bar{C} (I_\varepsilon(u_\varepsilon^R) - I_\varepsilon(u_\varepsilon^*) + \rho_\varepsilon^4)^{1/\gamma} \\ &= \rho_\varepsilon^1 + \bar{C} (I_\varepsilon(u_\varepsilon^R) - I(u^*) + I(u^*) - I_\varepsilon(u_\varepsilon^*) + \rho_\varepsilon^4)^{1/\gamma} \\ &\stackrel{(3.27)}{\leq} \rho_\varepsilon^1 + \bar{C} (\rho_\varepsilon^2 + I(u_\varepsilon^S) - I_\varepsilon(u_\varepsilon^*) + \rho_\varepsilon^4)^{1/\gamma} \\ &\stackrel{(3.28)}{\leq} \rho_\varepsilon^1 + \bar{C} (\rho_\varepsilon^2 + \rho_\varepsilon^3 + \rho_\varepsilon^4)^{1/\gamma}, \end{aligned}$$

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. [MS11, Sect. 2.6]. 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}_δ is defined as in (2.6) while its time-discrete counterpart $\mathcal{I}_\delta^\tau : \mathcal{K}_\tau(u^0) \rightarrow \mathbb{R} \cup \{+\infty\}$ where $\mathcal{K}_\tau(u^0) := \{u^0\} \times H^N$ is defined via

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

where \bar{u}_τ and \widehat{u}_τ denote the piecewise constant and piecewise affine interpolants of the nodes (u^0, \dots, u^N) (analogously for \bar{g}_τ and $\bar{\rho}_\tau$). 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 $\|\bar{g}_\tau - g\|_{L^2(0, T; H)} \leq C\tau$. Note that $\mathcal{K}_\tau(u^0)$ is in bijective correspondence with the set

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

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

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

where we additionally have the final condition $v_\tau^{*,N} = 0$, see [MS11, Eqn. 6.6]. Moreover, arguing as in [MS11, Sect. 6.2] we obtain the following estimate with a constant C independent of $\varepsilon \geq 0$ for the time-discrete solution

$$\begin{aligned} \delta^{1/2} \|\bar{w}_\tau^*\|_{L^2(0, T; H)} + \delta^{1/2} \|\bar{v}_\tau^*\|_{L^\infty(0, T; H)} + \|\bar{u}_\tau^*\|_{L^2(0, T; H)} \\ + \|u_\tau^{\text{disc}}\|_{H^1(0, T; H)} + \|\bar{\eta}_\tau^*\|_{L^2(0, T; H)} + \int_0^T \phi(\bar{u}_\tau^*) dt \leq C, \end{aligned} \quad (3.32)$$

where \bar{w}_τ^* 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

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

see [MS11, Sect. 4].

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

Proposition 3.5 *Let u_τ^{disc} and u^* be the minimizers of the time-discrete and -continuous WED functional \mathcal{I}_δ^τ and \mathcal{I}_δ , respectively. Then, we have*

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

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 $\mathbb{A} = \text{Id}$. The general case follows from 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,

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

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

$$\begin{aligned} & -\delta \sum_{i=0}^{K-1} \int_{i\tau}^{(i+1)\tau} \left(\frac{v_\tau^{*,i+1} - v_\tau^{*,i}}{\tau}, U_\tau^* \right)_H ds + \delta \int_0^t (\partial_{tt}u^*, U_\tau^*)_H ds \\ & + \int_0^t (\bar{v}_\tau^* - \partial_t u^*, U_\tau^*)_H ds + \int_0^t (\bar{\eta}_\tau^* - \eta^*, U_\tau^*)_H ds = \int_0^t (\bar{g}_\tau - g, U_\tau^*)_H ds. \end{aligned} \quad (3.35)$$

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

$$\begin{aligned} & \frac{\delta}{\tau} \int_0^\tau (v_\tau^{*,0}, U_\tau^*)_H ds - \frac{\delta}{\tau} \int_{(K-1)\tau}^{K\tau} (v_\tau^{*,K}, U_\tau^*)_H ds + \delta(\partial_t u^*(t), U_\tau^*(t))_H \\ & + \delta \sum_{i=1}^{K-1} \int_{i\tau}^{(i+1)\tau} \left(v_\tau^{*,i}, \frac{U_\tau^*(s+\tau) - U_\tau^*(s)}{\tau} \right)_H ds - \delta \int_0^t (\partial_t u^*, \partial_t U_\tau^*)_H ds + \frac{1}{2} \int_0^t \frac{d}{dt} \|U_\tau^*\|_H^2 ds \\ & + \int_0^t (\bar{\eta}_\tau^* - \eta^*, \bar{u}_\tau^{\text{disc}} - u^*)_H ds + \int_0^t (\bar{\eta}_\tau^* - \eta^*, u_\tau^{\text{disc}} - \bar{u}_\tau^{\text{disc}})_H ds \leq \int_0^t \|\bar{g}_\tau - g\|_H \|U_\tau^*\|_H ds, \end{aligned}$$

where $\bar{u}_\tau^{\text{disc}}$ 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.

$$\int_0^t (\bar{\eta}_\tau^* - \eta^*, \bar{u}_\tau^{\text{disc}} - u^*)_H ds \geq 0.$$

Moreover, thanks to the uniform bounds on $\bar{\eta}_\tau^*$, on η^* and on the time derivative of u_τ^{disc} , we have that $|\int_0^t (\bar{\eta}_\tau^* - \eta^*, u_\tau^{\text{disc}} - \bar{u}_\tau^{\text{disc}})_H ds| \leq C\tau$. 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 $\tau \rightarrow 0$ with an order of at least τ .

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

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

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

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

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

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

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

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

and, by definition of u_τ^{disc} ,

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

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 $t = K\tau$, $K \in \{1, \dots, N\}$,

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

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 [Mel17, Section 3.2], we deduce the estimate $\|U_\tau^*\|_{L^2(0,T;H)} \leq C\tau^{1/2}$. 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 [MS11]

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

for some C_δ depending on δ , but independent of ε , where η_ε^* (and η^* are the selection of the sub-differentials $\partial_H \phi_\varepsilon(u_\varepsilon^*)$ 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 [MS11].

Our strategy to prove Theorem 2.7 consists in checking that \mathcal{I}_ε 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}_ε and \mathfrak{S}_ε as well as the minimizers u^* and u_ε^* of the time-discrete WED functionals \mathcal{I}_0^τ and $\mathcal{I}_\varepsilon^\tau$ defined above.

It is worth noting that although we need to build the sequences \widehat{u}_ε and $\widetilde{u}_\varepsilon$ in Lemma 3.3 just for the minimizer of the functionals, we however need informations on the convergence rates for the operators \mathfrak{R}_ε and \mathfrak{S}_ε 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)-(3.27) We start by additionally assuming ϕ_ε to be convex for all $\varepsilon \geq 0$, 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}_τ^* and \bar{u}_τ^* 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

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

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

Remark 3.6 *The idea of the proof of (3.26)-(3.27) is to construct the recovering sequence $\mathcal{R}_\varepsilon u^*$ by using the good properties of the recovering operator \mathfrak{R}_ε 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}_ε is defined only over X and u^* is not pointwise well defined in X (but just in H) (see also the Section 4).

Moreover, since in general \mathfrak{R}_ε 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}_τ^* of u^* , since we cannot estimate how good is the approximation $\mathcal{I}(\widehat{u}_\tau^*)$ of $\mathcal{I}(u^*)$ (more precisely, we have no control on the $\partial_H \phi(\widehat{u}_\tau^*)$ since it does not solve any equation) estimates of $\mathcal{I}_\varepsilon(\mathcal{R}_\varepsilon u^*) - \mathcal{I}(u^*)$ seems hard to obtain.

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

$$\begin{aligned} \frac{1}{2} \|u^* - u_{\tau,\varepsilon}^R\|_{L^2(0,T;H)}^2 &\leq \|u^* - u_\tau^{\text{disc}}\|_{L^2(0,T;H)}^2 + \|u_\tau^{\text{disc}} - u_{\tau,\varepsilon}^R\|_{L^2(0,T;H)}^2 \\ &\leq (\rho_\tau^{\text{dtc}})^2 + 2 \sum_{i=1}^N \tau \|\mathfrak{R}_\varepsilon u_\tau^{\text{disc},i} - u_\tau^{\text{disc},i}\|_H^2 + \tau \|u_\varepsilon^0 - u^0\|_H^2 \\ &\leq (\rho_\tau^{\text{dtc}})^2 + 2 (r_\varepsilon^{\mathfrak{R}})^2 \sup_i \ell(\|u_\tau^{\text{disc},i}\|_H) + 2 (r_\varepsilon^{\mathfrak{R}})^2 \sum_{i=1}^N \tau \|u_\tau^{\text{disc},i}\|_X^m + \tau (r_\varepsilon^0)^2 \\ &\leq C (r_\varepsilon^{\mathfrak{R}} + \tau^{1/2} r_\varepsilon^0 + \rho_\tau^{\text{dtc}})^2. \end{aligned} \quad (3.38)$$

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

$$\begin{aligned} \mathcal{I}_\varepsilon(u_{\tau,\varepsilon}^R) - \mathcal{I}_0(u^*) &= \int_0^T e^{-t/\delta} (\psi_\varepsilon(\partial_t u_{\tau,\varepsilon}^R) - \psi_0(\partial_t u^*)) dt + \frac{1}{\delta} \int_0^T e^{-t/\delta} (\phi_\varepsilon(u_{\tau,\varepsilon}^R) - \phi_0(u^*)) dt \\ &\quad + \frac{1}{\delta} \int_0^T e^{-t/\delta} ((g_\varepsilon, u_{\tau,\varepsilon}^R)_H - (g, u^*)_H) dt =: \int_0^T [\Delta_{\tau,\varepsilon}^\psi + \Delta_{\tau,\varepsilon}^\phi + \Delta_{\tau,\varepsilon}^g] dt. \end{aligned}$$

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

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

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

$$\|U_{\varepsilon,\tau}(t)\|_H \leq \frac{r_\varepsilon^{\mathfrak{R}}}{\tau} (\ell(\|u_\tau^{\text{disc},1}\|_H) + \|u_\tau^{\text{disc},1}\|_X^m)^{1/2} + \frac{r_\varepsilon^0}{\tau},$$

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

$$\|U_{\varepsilon,\tau}(t)\|_H \leq \frac{r_\varepsilon^{\mathfrak{R}}}{\tau} (\ell(\|u_\tau^{\text{disc},i+1}\|_H) + \ell(\|u_\tau^{\text{disc},i}\|_H) + \|u_\tau^{\text{disc},i+1}\|_X^m + \|u_\tau^{\text{disc},i}\|_X^m)^{1/2}.$$

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

$$\begin{aligned}
\int_0^T \Delta_{\tau,\varepsilon}^\psi dt &\leq \int_0^T C \left(\|\partial_t u_{\tau,\varepsilon}^R\|_H + \|\partial_t u^*\|_H \right) \cdot \|\partial_t u_\tau^{\text{disc}} - \partial_t u^* + U_{\tau,\varepsilon}\|_H + r_\varepsilon^{\mathfrak{R},\psi} \|\partial_t u^*\|_B^2 dt \\
&\leq C \left(\frac{r_\varepsilon^{\mathfrak{R}}}{\tau} + r_\varepsilon^{\mathfrak{R},\psi} + \int_0^\tau \frac{r_\varepsilon^0}{\tau} dt \right) + C \|\partial_t u_\tau^{\text{disc}} - \partial_t u^*\|_{L^2(0,T;H)} \\
&\leq C \left(\frac{r_\varepsilon^{\mathfrak{R}}}{\tau} + r_\varepsilon^{\mathfrak{R},\psi} + \rho_\tau^{\text{dtc}} + r_\varepsilon^0 \right). \tag{3.39}
\end{aligned}$$

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

$$\begin{aligned}
\int_\tau^T \Delta_{\tau,\varepsilon}^\phi dt &\leq C \sum_{i=1}^{N-1} \int_{i\tau}^{(i+1)\tau} e^{-t/\delta} \left[\alpha_\tau^i(t) \phi_\varepsilon(\mathfrak{R}_\varepsilon u_\tau^{\text{disc},i}) + (1 - \alpha_\tau^i(t)) \phi_\varepsilon(\mathfrak{R}_\varepsilon u_\tau^{\text{disc},i+1}) - \phi_0(u^*) \right] dt \\
&\leq C r_\varepsilon^{\mathfrak{R},\phi} \sum_{i=1}^N \tau \left(\ell (\|u_\tau^{\text{disc},i}\|_H) + \|u_\tau^{\text{disc},i}\|_X^m + \|\eta_\tau^{\text{disc},i}\|_H^2 \right) \\
&\quad + C \sum_{i=1}^{N-1} \int_{i\tau}^{(i+1)\tau} e^{-t/\delta} \left[\alpha_\tau^i(t) \phi_0(u_\tau^{\text{disc},i}) + (1 - \alpha_\tau^i(t)) \phi_0(u_\tau^{\text{disc},i+1}) - \phi_0(u^*) \right] dt.
\end{aligned}$$

The last term on the right-hand side can be estimated as follows: Consider $\eta_\tau^{\text{disc},i} \in \partial_H \phi_0(u_\tau^{\text{disc},i})$, then

$$\begin{aligned}
\int_\tau^T \Delta_{\tau,\varepsilon}^\phi dt &\leq C r_\varepsilon^{\mathfrak{R},\phi} + C \sum_{i=1}^{N-1} \int_{i\tau}^{(i+1)\tau} e^{-t/\delta} \left[\phi_0(u_\tau^{\text{disc},i+1}) - \phi_0(u^*) + \alpha_\tau^i(t) \|\eta_\tau^{\text{disc},i}\|_H \|u_\tau^{\text{disc},i} - u_\tau^{\text{disc},i+1}\|_H \right] dt \\
&\leq C (r_\varepsilon^{\mathfrak{R},\phi} + \tau) + C \sum_{i=0}^{N-1} \int_{i\tau}^{(i+1)\tau} \|\eta_\tau^{\text{disc},i+1}\|_H \|u_\tau^{\text{disc},i+1} - u^*\|_H dt \\
&\leq C (r_\varepsilon^{\mathfrak{R},\phi} + \tau) + C \sum_{i=1}^{N-1} \int_{i\tau}^{(i+1)\tau} \|\eta_\tau^{\text{disc},i+1}\|_H (\|u^{\text{disc}} - u^*\|_H + \|u_\tau^{\text{disc},i+1} - u^{\text{disc}}\|_H) dt \\
&\leq C (r_\varepsilon^{\mathfrak{R},\phi} + \tau) + C \|u^{\text{disc}} - u^*\|_{L^2(0,T;H)} \\
&\leq C (r_\varepsilon^{\mathfrak{R},\phi} + \tau + \rho_\tau^{\text{dtc}}).
\end{aligned}$$

Here, we used the estimates for $\eta_\tau^{\text{disc},i}$ and the fact that $\sum_{i=0}^{N-1} \int_{i\tau}^{(i+1)\tau} \|u_\tau^{\text{disc},i} - u_\tau^{\text{disc},i+1}\|_H^2 dt \leq C\tau^2$ and $\sum_{i=0}^{N-1} \int_{i\tau}^{(i+1)\tau} \|u_\tau^{\text{disc},i+1} - u^d\|_H^2 \leq C\tau^2$. Moreover, since $\phi_\varepsilon(u_\varepsilon^0)$ is uniformly bounded, then $(\phi_\varepsilon(u_{\tau,\varepsilon}^R) - \phi_0(u^*))$ is bounded in $L^1(0, T)$ independently of ε , thus

$$\int_0^\tau \Delta_{\tau,\varepsilon}^\phi dt = \int_0^\tau e^{-t/\delta} (\phi_\varepsilon(u_{\tau,\varepsilon}^R) - \phi_0(u^*)) dt \leq C\tau.$$

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

$$\begin{aligned}
\Delta_{\tau,\varepsilon}^g &\leq \int_0^T e^{-t/\delta} (g_\varepsilon - g, u_{\tau,\varepsilon}^R)_H + (g, u_{\tau,\varepsilon}^R - u^*)_H \\
&\leq C \|g_\varepsilon - g\|_{L^2(0,T;H)} + C \|u_{\tau,\varepsilon}^R - u^*\|_{L^2(0,T;H)} = C (r_\varepsilon^g + \rho_\tau^{\text{dtc}} + r_\varepsilon^{\mathfrak{R}} + \tau^{1/2} r_\varepsilon^0). \tag{3.40}
\end{aligned}$$

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

$$\begin{aligned} \int_0^T e^{-t/\delta} \left(-\lambda \|u_{\tau,\varepsilon}^R\|_H^2 + \lambda \|u^*\|_H^2 \right) dt &\leq C \int_0^T \|u_{\tau,\varepsilon}^R + u^*\|_H \|u_{\tau,\varepsilon}^R - u^*\|_H dt \\ &\leq C \left(r_\varepsilon^{\mathfrak{R}} + \rho_\tau^{\text{dct}} + \tau^{1/2} r_\varepsilon^0 \right). \end{aligned}$$

Summarizing we have

$$\begin{aligned} \|u_{\tau,\varepsilon}^R - u^*\|_{L^2(0,T;H)} &\leq C \left(r_\varepsilon^{\mathfrak{R}} + \rho_\tau^{\text{dct}} + \tau^{1/2} r_\varepsilon^0 \right), \\ \mathcal{I}_\varepsilon(u_{\tau,\varepsilon}^R) - \mathcal{I}(u^*) &\leq C \left(r_\varepsilon^{\mathfrak{R}}/\tau + \rho_\tau^{\text{dct}} + r_\varepsilon^{\mathfrak{R},\phi} + r_\varepsilon^{\mathfrak{R},\psi} + \tau + r_\varepsilon^g + r_\varepsilon^{\mathfrak{R}} + r_\varepsilon^0 + \tau^{1/2} r_\varepsilon^0 \right). \end{aligned}$$

We now choose $\tau_\varepsilon^{\mathfrak{R}}$ as the minimizer of $\tau \mapsto r_\varepsilon^{\mathfrak{R}}/\tau + \rho_\tau^{\text{dct}} + \tau + \tau^{1/2} r_\varepsilon^0$, and we easily check that $\tau_\varepsilon^{\mathfrak{R}} \rightarrow 0$ as $\varepsilon \rightarrow 0$. Note that this is the choice of τ (as a function of ε) that optimizes the convergence rates. In particular, by defining $u_\varepsilon^R := u_{\tau_\varepsilon^{\mathfrak{R}},\varepsilon}^R$ we see that conditions (3.26)-(3.27) follow from the above estimates where

$$\begin{aligned} \rho_\varepsilon^1 &= r_\varepsilon^{\mathfrak{R}} + \rho_{\tau_\varepsilon^{\mathfrak{R}}}^{\text{dct}} + \left(\tau_\varepsilon^{\mathfrak{R}} \right)^{1/2} r_\varepsilon^0, \\ \rho_\varepsilon^2 &= r_\varepsilon^{\mathfrak{R}}/\tau_\varepsilon^{\mathfrak{R}} + \rho_{\tau_\varepsilon^{\mathfrak{R}}}^{\text{dct}} + r_\varepsilon^{\mathfrak{R},\phi} + r_\varepsilon^{\mathfrak{R},\psi} + \tau_\varepsilon^{\mathfrak{R}} + r_\varepsilon^g + r_\varepsilon^{\mathfrak{R}} + r_\varepsilon^0 + \left(\tau_\varepsilon^{\mathfrak{R}} \right)^{1/2} r_\varepsilon^0. \end{aligned}$$

Proof of (3.28) A proof of (3.28) can be obtained by following the proof 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 $u_\varepsilon^{\text{disc}}$, and of u^* and u_ε^* . More precisely, let $u_\varepsilon^{\text{disc}}$ be the minimizer of the discrete WED functional $\mathcal{I}_\varepsilon^\tau$ 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}_ε .

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

$$\|u_{\tau,\varepsilon}^S - u_\varepsilon^*\|_{L^2(0,T;H)} \leq C \left(r_\varepsilon^{\mathfrak{S}} + \rho_\tau^{\text{dct}} + \tau^{1/2} r_\varepsilon^0 \right).$$

Computations analogous to the ones used in (3.39)-(3.40) combined with assumptions (2.21)-(2.24) give us

$$\begin{aligned} \int_0^T e^{-t/\delta} \left(\psi_0(\partial_t u_{\tau,\varepsilon}^S) - \psi_\varepsilon(\partial_t u_\varepsilon^*) \right) dt &\leq C \left(\frac{r_\varepsilon^{\mathfrak{S}}}{\tau} + r_\varepsilon^{\mathfrak{S},\psi} + \rho_\tau^{\text{dct}} + r_\varepsilon^0 \right), \\ \frac{1}{\delta} \int_0^T e^{-t/\delta} \left(\phi_0(u_{\tau,\varepsilon}^S) - \phi_\varepsilon(u_\varepsilon^*) \right) dt &\leq C \left(r_\varepsilon^{\mathfrak{S},\phi} + \tau + \rho_\tau^{\text{dct}} \right), \\ \frac{1}{\delta} \int_0^T e^{-t/\delta} \left((g(t), u_{\tau,\varepsilon}^S)_H - (g_\varepsilon(t), u_\varepsilon^*)_H \right) dt &\leq C \left(r_\varepsilon^g + \rho_\tau^{\text{dct}} + r_\varepsilon^{\mathfrak{S}} + \tau^{1/2} r_\varepsilon^0 \right). \end{aligned}$$

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

Summarizing we have

$$\mathcal{I}(u_{\tau,\varepsilon}^S) - \mathcal{I}_\varepsilon(u_\varepsilon^*) \leq C \left(r_\varepsilon^{\mathfrak{S}}/\tau + \rho_\tau^{\text{dct}} + r_\varepsilon^{\mathfrak{S}} + r_\varepsilon^{\mathfrak{S},\phi} + r_\varepsilon^{\mathfrak{S},\psi} + \tau + r_\varepsilon^g + r_\varepsilon^0 + \tau^{1/2} r_\varepsilon^0 \right).$$

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

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

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}_ε solves the elliptic equation $(P_{\delta,\varepsilon})$. By testing the Euler–Lagrange equation for \mathcal{I}_ε with $v = u_\varepsilon^* - u_\varepsilon^R$ we get

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

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

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

By using λ -convexity of ϕ_ε , we get

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

We now exploit that ψ_ε is quadratic to obtain

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

By arguing as in [MS11, Prop. 2.1], recalling that $v(0) = 0$, for all $\delta > 0$ sufficiently small we obtain, after suitably renaming the constant c ,

$$\mathcal{I}_\varepsilon(u_\varepsilon^R) - \mathcal{I}_\varepsilon(u_\varepsilon^*) \geq c \int_0^T e^{-t/\delta} \|u_\varepsilon^R - u_\varepsilon^*\|_H^2 dt,$$

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 Homogenization of a 1D parabolic equation

Evolutionary Γ -convergence via WED functionals. In this section we provide an example of application of the above abstract theory. In the 1-dimensional domain $\Omega = (0, 1)$ we consider the equation

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

equipped with homogeneous Neumann boundary conditions $\partial_x u(0, t) = \partial_x u(1, t) = 0$. Here, $1 < q \leq 2$ 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 $a, b, D, c : [0, 1] \rightarrow [1/M, M]$ 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

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

where

$$\begin{aligned} \psi_\varepsilon(v) &= \frac{1}{2} \int_\Omega a_\varepsilon(x) |v|^2 dx & \text{and} & & \phi_\varepsilon(u) &= \varphi_\varepsilon^1(u) - \varphi_\varepsilon^2(u) & \text{with} \\ \varphi_\varepsilon^1(u) &= \begin{cases} \frac{1}{2} \int_\Omega D_\varepsilon(x) |\partial_x u|^2 + b_\varepsilon(x) |u|^2 dx & \text{if } u \in H^1(\Omega), \\ \infty & \text{else,} \end{cases} & & & & & \text{and} \\ \varphi_\varepsilon^2(u) &= \frac{1}{q} \int_\Omega c_\varepsilon(x) |u|^q dx. \end{aligned}$$

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

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

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. Theorem 2.1)

$$\mathcal{I}_{\delta, \varepsilon}(u) := \begin{cases} \int_0^T e^{-t/\delta} (\psi_\varepsilon(\dot{u}) + \frac{1}{\delta} \phi_\varepsilon(u) - \frac{1}{\delta} (g_\varepsilon(t), u)_H) dt & \text{if } u \in \mathcal{K}(u_\varepsilon^0), \\ \infty & \text{else.} \end{cases}$$

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}_δ defined by the static Γ -limits ϕ_0 and ψ_0 . The latter are given by

$$\begin{aligned} \psi_0(v) &= \frac{1}{2} \int_\Omega a_{\text{arith}} |v|^2 dx & \text{and} & & \phi_0(u) &= \varphi_0^1(u) - \varphi_0^2(u) & \text{with} \\ \varphi_0^1(u) &= \begin{cases} \frac{1}{2} \int_\Omega D_{\text{harm}} |\partial_x u|^2 + b_{\text{arith}} |u|^2 dx & \text{if } u \in H^1(\Omega), \\ \infty & \text{else,} \end{cases} \\ \varphi_0^2(u) &= \frac{1}{q} \int_\Omega c_{\text{arith}} |u|^q dx, \end{aligned}$$

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 $D_{\text{harm}} = (\int_0^1 1/D(y) dy)^{-1}$. We refer the reader to [Mie16] 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 $\varphi_\varepsilon^1(u_\varepsilon) \leq C$, we have $u_\varepsilon \rightarrow u$ weakly in $H^1(\Omega)$, $u_\varepsilon \rightarrow u$ strongly in $L^q(\Omega)$ and, as a consequence of the Sobolev embeddings,

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

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

$$\varphi_\varepsilon^2(u_\varepsilon) - \varphi_0^2(u) = \frac{1}{q} \int_\Omega c_\varepsilon |u_\varepsilon|^q - c_{\text{arith}} |u|^q \, dx = \frac{1}{q} \int_\Omega c_\varepsilon (|u_\varepsilon|^q - |u|^q) + (c_\varepsilon - c_{\text{arith}}) |u|^q \, dx \rightarrow 0.$$

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 $u_\varepsilon^0 \rightarrow u^0$ strongly in $L^2(\Omega)$ and that the external force terms satisfy $g_\varepsilon \rightarrow g_0$ weakly in $L^2(0, T; H)$.

By applying Theorem 2.2 and Theorem 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*

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

for $\varepsilon \rightarrow 0$. Moreover, if additionally g_ε satisfies (2.14)-(2.15) and the initial data are well prepared, i.e. $\phi_\varepsilon(u_\varepsilon^0) \rightarrow \phi(u^0)$, 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 $C([0, T]; L^2(\Omega))$ hold true

(i) $\lim_{\varepsilon \rightarrow 0} u_{\delta,\varepsilon}^* = u_\delta^*$ where $u_\delta^* \in \arg \min \mathcal{I}_\delta$ and solves (4.3),

(ii) $\lim_{\delta \rightarrow 0} u_{\delta,\varepsilon}^* = u_\varepsilon$ where u_ε solves (4.1)

(iii) $\lim_{\delta+\varepsilon \rightarrow 0} u_{\delta,\varepsilon}^* = u$ where u solves the following homogenized equation

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

Rates of convergence. We are now interested to establish some convergence rates. More precisely, we want to estimate the rate of convergence of $\|u_{\delta,\varepsilon}^* - u_\delta^*\|_{L^2(0,T;L^2(\Omega))}$, 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 $\varepsilon \geq 0$).

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

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

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

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

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

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

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

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

Combining these estimates, we get

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

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 $\partial_H \phi_\varepsilon$, $\varepsilon \geq 0$, 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

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

This proves that \mathfrak{R}_ε 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 $\mathfrak{S}_\varepsilon : H^1(\Omega) \rightarrow H^1(\Omega)$ 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,

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

Indeed, by convexity, we have that

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

Here, we have

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

and

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

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

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

Hence, \mathfrak{S}_ε 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

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

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*

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

Moreover, by interpolation one can prove the following

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

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}_δ , respectively. Then, for all $\delta > 0$ sufficiently small*

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

for some positive constant C_δ .

5 Appendix

Lemma 5.1 (Diagonal extraction for weakly converging sequences) *Let $u_n \rightarrow u$ weakly in a separable and reflexive Banach space B and, for all n , let $u_{m,n} \rightarrow u_n$ weakly in B . Moreover, assume that $\|u_{m,n}\|_B \leq C$, where C is a positive constant independent of m and n . Then, there exists a subsequence $u_{m_k,n_k} \rightarrow u$ 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

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

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

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

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

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

■

Acknowledgement

The authors thank Alexander Mielke for stimulating discussions and helpful comments. The work of M.L. is supported by Einstein Center for Mathematics Berlin under project MATHEON SE2. S.M. is supported by the Austrian Science Fund (FWF) project P27052-N25.

References

- [AM17] G. Akagi and S. Melchionna. Elliptic-regularization of nonpotential perturbations of doubly-nonlinear flows of nonconvex energies: a variational approach. *arxiv*, 1702.06755v2, 2017. to appear in *J. Convex Anal.*
- [AS11] G. Akagi and U. Stefanelli. Weighted energy-dissipation functionals for doubly-nonlinear evolution. *J. Funct. Anal.*, 260:2541–2578, 2011.
- [AS14] G. Akagi and U. Stefanelli. Doubly nonlinear equations as convex minimization. *SIAM J. Math. Anal.*, 46:1922–1945, 2014.
- [AS16] Goro Akagi and Ulisse Stefanelli. A variational principle for gradient flows of nonconvex energies. *J. Convex Anal.*, 23(1):53–75, 2016.
- [Col92] P. Colli. On some doubly nonlinear evolution equations in Banach spaces. *Jpn. J. Ind. Appl. Math.*, 9:181–203, 1992.
- [CV90] P. Colli and A. Visintin. On a class of doubly nonlinear evolution equations. *Comm. Partial Differential Equations*, 15:737—756, 1990.
- [Ilm94] T. Ilmanen. Elliptic regularization and partial regularity for motion by mean curvature. *Mem. Amer. Math. Soc.*, 108:520x:+90, 1994.
- [LM68] J.-L. Lions and E. Magenes. *Problèmes aux limites non homogènes et applications*. Travaux et recherches Matématique, Dunod, Paris, 1968.
- [LS13a] M. Liero and U. Stefanelli. A new minimum principle for Lagrangian mechanics. *J. Nonlinear Sci.*, 23:179–204, 2013.
- [LS13b] M. Liero and U. Stefanelli. Weighted inertia-dissipation-energy functionals for semilinear equations. *Boll. Unione Mat. Ital. (9)*, 6:1–27, 2013.
- [Mel17] S. Melchionna. A variational principle for nonpotential perturbations of gradient flows of nonconvex energies. *J. Differential Equations*, 262:3737–3758, 2017.
- [Mie16] A. Mielke. On evolutionary Γ -convergence for gradient systems. *Lecture Notes in Appl. Math. Mech.*, pages 187–249, 2016.
- [MO08] A. Mielke and M. Ortiz. A class of minimum principle for characterizing the trajectories of dissipative systems. *ESAIM Control Optim. Calc. Var.*, 14:494–516, 2008.
- [MS11] A. Mielke and U. Stefanelli. Weighted energy-dissipation functionals for gradient flows. *ESAIM Control Optim. Calc. Var.*, 17:52–85, 2011.
- [Ô82] M. Ôtani. Nonmonotone perturbations associated with subdifferential operators, cauchy problems. *J. Differential Equations*, 46:248–273, 1982.
- [Rud87] W. Rudin. *Real and complex analysis. Third edition*. McGraw Hill, New York, 1987.

- [SS04] E. Sandier and S. Serfaty. Gamma-convergence of gradient flows with applications to Ginzburg-Landau. *Comm. Pure Appl. Math.*, LVII:1627–1672, 2004.
- [ST12] E. Serra and P. Tilli. Nonlinear wave equations as limit of convex minimization problems: proof of a conjecture by De Giorgi. *Ann. of Math. (2)*, 175:1511–1574, 2012.
- [Ste08] U. Stefanelli. The Brezis-Ekeland principle for doubly nonlinear equations. *SIAM J. Control Optim.*, 47:1615–1642, 2008.
- [Ste11] U. Stefanelli. The De Giorgi conjecture on elliptic regularization. *Math. Models Methods Appl. Sci.*, 21:1377–1394, 2011.