J. Evol. Equ. 19 (2019), 479–522

1424-3199/19/020479-44, *published online* January 25, 2019
https://doi.org/10.1007/s00028-019-00484-x

Journal of Evolution Equations

An existence result and evolutionary Γ-convergence for perturbed gradient systems

Aras Bacho, Etienne Emmrich and Alexander Mielke

Abstract. The initial-value problem for the perturbed gradient flow

$$B(t, u(t)) \in \partial\Psi_{u(t)}(u'(t)) + \partial\mathcal{E}_t(u(t)) \text{ for a.a. } t \in (0, T), \qquad u(0) = u_0,$$

with a perturbation B in a Banach space V is investigated, where the dissipation potential $\Psi_u : V \to [0, +\infty)$ and the energy functional $\mathcal{E}_t : V \to (-\infty, +\infty]$ are non-smooth and supposed to be convex and nonconvex, respectively. The perturbation $B : [0, T] \times V \to V^*$, $(t, v) \mapsto B(t, v)$ is assumed to be continuous and satisfies a growth condition. Under suitable assumptions on the dissipation potential and the energy functional, existence of strong solutions is shown by proving convergence of a semi-implicit discretization scheme with a variational approximation technique. Moreover, for perturbed gradient systems $(V, \mathcal{E}^\varepsilon, \Psi^\varepsilon, B^\varepsilon)$ depending on a small parameter $\varepsilon > 0$, we develop a theory of evolutionary Γ-convergence in terms of the suitable convergences of $\mathcal{E}^\varepsilon$, Ψ^ε, and B^ε to the limit system $(V, \mathcal{E}^0, \Psi^0, B^0)$.

1. Introduction

The aim of this paper is to provide existence results for the initial-value problem for the doubly nonlinear evolution inclusion

$$B(t, u(t)) \in \partial\Psi_{u(t)}(u'(t)) + \partial\mathcal{E}_t(u(t)) \text{ in } V^* \text{ for a.a. } t \in (0, T), \qquad u(0) = u_0 \tag{1.1}$$

with a continuous perturbation B in the separable and reflexive real Banach space $(V, \|\cdot\|)$, where $\partial\Psi_u$ and $\partial\mathcal{E}_t$ denote the subdifferential of Ψ_u and $\mathcal{E}_t$, respectively. We will call the quadruple $(V, \mathcal{E}, \Psi, B)$, which defines (1.1), a *perturbed gradient system*, as for $B \equiv 0$ we obtain a gradient system in the sense of [20].

The functional Ψ_u is supposed to be a dissipation potential for all $u \in \mathrm{dom}(\mathcal{E}_t)$, i.e., it is proper, lower semicontinuous and convex with $\Psi_u(0) = 0$ for all $u \in \mathrm{dom}(\mathcal{E}_t)$. If the functionals Ψ_u and $\mathcal{E}_t$ are Fréchet differentiable, the differential inclusion (1.1)

Mathematics Subject Classification: 35A15, 35K50, 35K85, 49Q20
Keywords: Doubly nonlinear equations, Generalized and perturbed gradient flows, Evolutionary Gamma convergence, Homogenization of reaction-diffusion systems.
The research was partially supported by DFG via SFB 910, subprojects A5 and A8.

becomes the abstract evolution equation (also called doubly nonlinear equation in [11,12])

$$\mathrm{D}\Psi_{u(t)}(u'(t)) = -\mathrm{D}\mathcal{E}_t(u(t)) + B(t, u(t)) \quad \text{in } V^* \text{ a.e. in } (0, T),$$

where $\mathrm{D}\Psi_u$ and $\mathrm{D}\mathcal{E}_t$ denote the FRÉCHET derivative of Ψ_u and $\mathcal{E}_t$, respectively.

Perturbed gradient systems in the special case

$$u'(t) + \beta(u(t), u(t)) \ni u^*(t) \quad \text{in } V^* \text{ a.a. in } (0, T)$$

have been studied in [33] assuming $u^* \in \mathrm{L}^{p'}(0, T; V^*)$ and $\beta : V \times V \to 2^{V^*}$ to be weakly lower semicontinuous in the first argument and maximal monotone in the second by using the so-called FITZPATRICK theory and an adapted version of evolutionary Γ-convergence. Also, there are classical results of BRÉZIS [7] for perturbed gradient systems in a HILBERT space using the theory of maximal monotone operators and assuming the perturbation to be LIPSCHITZ-continuous, see also [4,26] and the references therein. However, none of these theories combine the perturbations with the more general doubly nonlinear theory developed in [23], where u' occurs in the form of the subdifferential $\mathrm{D}\Psi_{u(t)}(u'(t))$, which may even depend on the state $u(t)$.

The question arises why it is interesting to study perturbed gradient systems. First, considering perturbed systems is important in order to describe physical systems near or far from equilibrium properly. This is often done by incorporating non-trivial energy-producing perturbations into gradient-flow equations. There are many ways to do so. Second, [19, p. 235] highlights with an example that in some cases it is easier to treat a system with a non-trivial, exact gradient structure $(X, \widetilde{\mathcal{E}}, \widetilde{\Psi})$ as a perturbed gradient system $(V, \mathcal{E}, \Psi, B)$ with simpler energy $\mathcal{E}$ and dissipation potential Ψ_u.

In Sect. 2, we provide the main existence result, see Theorem 2.5, by generalizing the ideas developed in [23]. Section 3 is devoted to the question of evolutionary Γ-convergence of families $(V, \mathcal{E}^\varepsilon, \Psi^\varepsilon, B^\varepsilon)$ of perturbed gradient systems, where $\varepsilon > 0$ is a small parameter in a so-called multiscale problem. The aim is to find a limiting (also called effective) perturbed gradient system $(V, \mathcal{E}^0, \Psi^0, B^0)$ such that the solutions u_ε of

$$\mathrm{D}\Psi^\varepsilon_{u_\varepsilon(t)}(u_\varepsilon'(t)) = -\mathrm{D}\mathcal{E}^\varepsilon_t(u_\varepsilon(t)) + B^\varepsilon(t, u_\varepsilon(t))$$

have subsequences that converge to limiting solutions u that are solutions of the perturbed gradient flow for $\varepsilon = 0$, see the survey [19] in the case of exact gradient systems. By allowing for $B^\varepsilon \neq 0$, our theory provides a generalization of the results developed in [20,30,31] for exact gradient flows where $B^\varepsilon \equiv 0$. The theory developed in SERFATY [31] is quite abstract and general and works even for gradient flows on metric spaces but is restricted to quadratic $\Psi_u(\cdot)$. The core of the approach are the general abstract assumptions that the following liminf estimates hold:

$$\begin{aligned} &\liminf_{\varepsilon\to 0}\int_0^T \Psi^\varepsilon_{u_\varepsilon(t)}(u_\varepsilon'(t))\,\mathrm{d}t \geq \int_0^T \Psi^0_{u(t)}(u'(t))\,\mathrm{d}t \quad \text{and} \\ &\liminf_{\varepsilon\to 0}\Psi^{\varepsilon,*}_{u_\varepsilon(t)}\big(-\partial\mathcal{E}^\varepsilon_t(u_\varepsilon(t))\big) \geq \Psi^{0,*}_{u_0(t)}\big(-\partial\mathcal{E}^0_t(u(t))\big) \text{ for a.a. } t\in(0,T), \end{aligned} \tag{1.2}$$

which remain to be established on a case-by-case analysis. We follow the ideas in [23, Thm. 4.8] working in a BANACH-space setting, where we are able to allow for general convex $\Psi_u(\cdot)$ that may be non-quadratic, which leads to doubly nonlinear evolution equations, see [11,12,18,22,23]. Our Theorem 3.1 shows that under suitable technical assumptions, including uniform semi-convexity of $\mathcal{E}^\varepsilon$, it is enough to establish $\mathcal{E}^\varepsilon_t \xrightarrow{\Gamma} \mathcal{E}^0_t$ (strong Γ-convergence in V) and $\Psi^\varepsilon_{u_\varepsilon} \xrightarrow{\mathrm{M}} \Psi^0_{u_0}$ in V (MOSCO convergence in V). In particular, our conditions guarantee (slightly relaxed) versions of the liminf estimates (1.2) postulated in [30,31].

In Sect. 4, we show that our abstract result on evolutionary Γ-convergence can be used for the homogenization of quasilinear parabolic systems. For this application, the convergence $\Psi^\varepsilon_{u_\varepsilon} \xrightarrow{\mathrm{M}} \Psi^0_{u_0}$ is too restrictive, and it is necessary to generalize it to situations where the strong Γ-convergence $\Psi^\varepsilon_{u_\varepsilon} \xrightarrow{\Gamma} \Psi^0_{u_0}$ is enough, see Corollary 3.3. Using the novel argument from LIERO- REICHELT [17], the weak convergence of $u_\varepsilon \rightharpoonup u_0$ in $\mathrm{W}^{1,1}(0,T;V)$ can be bypassed by exploiting strong convergence of the piecewise affine interpolants $\widehat{u}^\tau_\varepsilon \to \widehat{u}^\tau_0$ in $\mathrm{W}^{1,1}(0,T;V)$ for $\varepsilon \to 0$ and $\tau > 0$ fixed.

The general structure is that we provide a full and detailed proof of the existence result in Sect. 2, where we use DE GIORGI's minimization scheme using variational interpolators. The result on the evolutionary Γ-convergence in Sect. 3 follows the same lines but is considerably simpler as existence of solutions is assumed to be shown. Hence, for getting an overview of the strategy in Sect. 2 it might be helpful to browse through the more compact proof of Theorem 3.1 first. This will facilitate the subsequent reading of the full details in Sect. 2. In particular, the elaborate time discretization using DE GIORGI's variational interpolants is only needed there.

2. The main existence result

Before making all the assumptions concerning the dissipation potential, the energy functional and the perturbation, we need some basic tools from convex analysis.

2.1. Preliminaries and notation

In this section, we collect some important notions and results on convex analysis and Γ-convergence, which we will need later on for the existence result. First of all, we introduce the so-called LEGENDRE- FENCHEL transform (or conjugate) Ψ^* of a proper, lower semicontinuous and convex functional $\Psi : V \to (-\infty, +\infty]$ that is defined by

$$\Psi^*(\xi) := \sup_{u\in V} \{\langle \xi, u\rangle - \Psi(u)\}, \quad \xi \in V^*,$$

where $\langle\cdot,\cdot\rangle$ denotes the duality pairing between the BANACH space V and it's topological dual space V^*. From the definition, the FENCHEL- YOUNG inequality

$$\langle \xi, u\rangle \le \Psi(u) + \Psi^*(\xi), \quad v \in V, \xi \in V^*,$$

immediately follows. The conjugate Ψ^* is again proper, lower semicontinuous and convex, see, e.g., EKELAND and TÉMAM [16]. If, in addition, $\Psi(0) = 0$, then $\Psi^*(0) = 0$ holds too. For a proper functional $F : V \to (-\infty, +\infty]$, the (FRÉCHET)-subdifferential of F is given by the multivalued map $\partial F : V \to 2^{V^*}$ with

$$\partial F(u) := \left\{ \xi \in V^* \;\middle|\; \liminf_{v\to u} \frac{F(v) - F(u) - \langle \xi, v-u\rangle}{\|v-u\|} \ge 0 \right\}$$

for all elements u in the effective domain $\mathrm{dom}(F) := \{v \in V \mid F(v) < +\infty\}$ of F. For convex and proper functions F, it follows by simple calculations that the subdifferential of F is given by

$$\partial F(u) = \left\{ \xi \in V^* \;\middle|\; F(v) \le F(u) + \langle \xi, v-u\rangle \text{ for all } v \in V \right\}.$$

The following lemma gives a relation between the subdifferential of a functional and it's LEGENDRE- FENCHEL transform.

LEMMA 2.1. *Let $\Psi : V \to (-\infty, +\infty]$ be a proper, lower semicontinuous and convex functional and let $\Psi^* : V^* \to (-\infty, +\infty]$ be the* LEGENDRE- FENCHEL *transform of Ψ. Then, for all $(u, \xi) \in V \times V^*$ the following assertions are equivalent:*

(i) $\xi \in \partial\Psi(u)$ *in* V^*;
(ii) $u \in \partial\Psi^*(\xi)$ *in* V;
(iii) $\langle \xi, u\rangle = \Psi(u) + \Psi^*(\xi)$ *in* $\mathbb{R}$.

Proof. EKELAND and TÉMAM [16, Proposition 5.1 and Cor. 5.2 on pp. 21]. □

For the dissipation potentials Ψ_u, we need the notion of Γ-convergence, see [8,9,13] (also called epigraph convergence in [3]). We consider a functional $\Psi : V \to (-\infty, \infty]$ and a sequence $(\Psi_n)_{n\in\mathbb{N}}$ of functionals all of which are lower semicontinuous convex functionals. The (strong) Γ-convergence $\Psi_n \xrightarrow{\Gamma} \Psi$ in V is defined via

$$\Psi_n \xrightarrow{\Gamma} \Psi \iff \begin{cases} \text{(a)} \quad v_n \to v \implies \Psi(v) \le \liminf_{n\to\infty} \Psi_n(v_n), \\ \text{(b)} \quad \forall \widehat{v} \in V\ \exists (\widehat{v}_n)_{n\in\mathbb{N}} :\ \widehat{v}_n \to \widehat{v} \text{ and } \Psi(v) \ge \limsup_{n\to\infty} \Psi_n(v_n). \end{cases}$$

Here, (a) is called the (strong) liminf estimate, while (b) is called the (strong) limsup estimate or the existence of recovery sequences. Similarly, we define the (sequential) weak Γ-convergence $\Psi_n \overset{\Gamma}{\rightharpoonup} \Psi$ in V via

$$\Psi_n \overset{\Gamma}{\rightharpoonup} \Psi \iff \begin{cases} \text{(a)} \quad v_n \rightharpoonup v \implies \Psi(v) \le \liminf_{n\to\infty} \Psi_n(v_n), \\ \text{(b)} \quad \forall \widehat{v} \in V\ \exists\, (\widehat{v}_n)_{n\in\mathbb{N}}: \ \ \widehat{v}_n \rightharpoonup \widehat{v} \text{ and } \Psi(v) \ge \limsup_{n\to\infty} \Psi_n(v_n). \end{cases}$$

If both convergences hold, then we say that Ψ_n Mosco converges to Ψ and write $\Psi_n \overset{\mathrm{M}}{\longrightarrow} \Psi$. In [3, pp. 271], the following fundamental relation between Γ-convergence and the Legendre- Fenchel transform was established:

$$\Psi_n \overset{\Gamma}{\to} \Psi \quad \iff \quad \Psi_n^* \overset{\Gamma}{\rightharpoonup} \Psi^*, \tag{2.1}$$

which always holds on reflexive Banach spaces V if all Ψ_n and Ψ_n^* are nonnegative (as for our dissipation potentials).

2.2. Semi-implicit variational approximation scheme

The idea of showing the existence of strong solutions to (1.1) with an initial condition $u = u_0 \in V$ consists of constructing a solution via a particular discretization scheme, more precisely, with a semi-implicit Euler method. The usual implicit Euler method does not work, since the equation (1.1) does not possess the gradient-flow structure due to the non-potential perturbation. With our approach, it is possible to construct time-discrete solutions via a variational approximation scheme. To illustrate this let for $N \in \mathbb{N}\backslash\{0\}$

$$I_\tau = \{0 = t_0 < t_1 < \cdots < t_n = n\tau < \cdots t_N = T\} \tag{2.2}$$

be an equidistant partition of the time interval $[0, T]$ with step size $\tau := T/N$, where we omit the dependence of t_n on the step size τ for simplicity. The approximation of (1.1) is then given by

$$B(t_{n-1}, U_\tau^{n-1}) \in \partial\Psi_{U_\tau^{n-1}}\left(\frac{U_\tau^n - U_\tau^{n-1}}{\tau}\right) + \partial\mathcal{E}_{t_n}(U_\tau^n), \quad n = 1, \ldots, N, \tag{2.3}$$

where the values $U_\tau^n \approx u(t_n)$, which shall approximate the exact solution of (1.1) at t_n, are to determine. If both the dissipation potential and the energy functional are Fréchet differentiable, the inclusion (2.3) becomes the equation

$$B(t_{n-1}, U_\tau^{n-1}) = \mathrm{D}\Psi_{U_\tau^{n-1}}\left(\frac{U_\tau^n - U_\tau^{n-1}}{\tau}\right) + \mathrm{D}\mathcal{E}_{t_n}(U_\tau^n), \quad n = 1, \ldots, N. \tag{2.4}$$

It is now simple to see that the value U_τ^n can be characterized as a solution of the Euler- Lagrange equation associated with the map

$$v \mapsto \Phi(\tau, t_{n-1}, U_\tau^{n-1}, B(t_{n-1}, U_\tau^{n-1}); v),$$

where

$$\Phi(r, t, u, w; v) = r\Psi_u\left(\frac{v-u}{r}\right) + \mathcal{E}_{t+r}(v) - \langle w, v\rangle \tag{2.5}$$

for $r \in \mathbb{R}^{>0}$ and$t \in [0, T)$ with $r + t \in [0, T]$, $u, v \in V$, and $w \in V^*$. In fact, we determine the value U_τ^n by minimizing the functional Φ in the variable $v \in V$ under suitable conditions on the dissipation potential and the energy functional. To assure that the value U_τ^n satisfies the inclusion (2.3) also in the non-smooth case, which is in general not true, we make an assumption to enforce this property.

2.3. Assumptions for the main existence result

We now collect the assumptions on the perturbed gradient system $\mathrm{PG} = (V, \mathcal{E}, \Psi, B)$ for our existence result. They will be denoted by symbols like (2.En), (2.Ψm), and (2.Bk).

The assumptions on the energy functional are the following.

(2.Ea) **Constant domain.** For all $t \in [0, T]$, the functional $\mathcal{E}_t : V \to (-\infty, +\infty]$ is proper and lower semicontinuous with the time-independent effective domain $D \equiv \mathrm{dom}(\mathcal{E}_t) \subset V$ for all $t \in [0, T]$.

(2.Eb) **Compactness of sublevels.** There exists $t^* \in [0, T]$ such that the functional $\mathcal{E}_{t^*}$ has compact sublevels in V.

(2.Ec) **Energetic control of power.** For all $u \in D$, the power map $t \mapsto \mathcal{E}_t(u)$ is continuous on $[0, T]$ and differentiable in $(0, T)$ and its derivative $\partial_t\mathcal{E}_t$ is controlled by the function $\mathcal{E}_t$, i.e., there exists $C > 0$ such that

$$|\partial_t\mathcal{E}_t(u)| \leq C\mathcal{E}_t(u) \quad \text{for all } t \in (0, T) \text{ and } u \in D.$$

(2.Ed) **Chain rule.** For every absolutely continuous curve $v \in \mathrm{AC}([0, T]; V)$ and every BOCHNER integrable function $\xi \in \mathrm{L}^1(0, T; V^*)$ such that

$$\sup_{t\in[0,T]} |\mathcal{E}_t(u(t))| < +\infty, \quad \xi(t) \in \partial\mathcal{E}_t(u(t)) \quad \text{a.e. in } (0, T),$$

$$\int_0^T \Psi_{u(t)}(u'(t))\,\mathrm{d}t < +\infty \quad \text{and} \quad \int_0^T \Psi^*_{u(t)}(\xi(t))\,\mathrm{d}t < +\infty,$$

the map $t \mapsto \mathcal{E}_t(u(t))$ is absolutely continuous on $[0, T]$ and

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}_t(u(t)) \geq \langle\xi(t), u'(t)\rangle + \partial_t\mathcal{E}_t(u(t)) \quad \text{a.e. in } (0, T).$$

(2.Ee) **Strong-weak closedness.** For all $t \in [0, T]$ and all sequences $(u_n, \xi_n)_{n\in\mathbb{N}} \subset V \times V^*$ with $\xi_n \in \partial\mathcal{E}_t(u_n)$ such that

$$u_n \to u \in V, \quad \xi_n \rightharpoonup \xi \in V^*, \quad \mathcal{E}_t(u_n) \to \mathcal{E} \in \mathbb{R} \quad \text{and} \quad \partial_t\mathcal{E}_t(u_n) \to \mathcal{P} \in \mathbb{R}$$

as $n \to \infty$, we have the relations

$$\xi \in \partial\mathcal{E}_t(u), \quad \mathcal{P} \leq \partial_t\mathcal{E}_t(u) \quad \text{and} \quad \mathcal{E} = \mathcal{E}_t(u).$$

We first give a few relevant comments on these assumptions that will be important below.

REMARK 2.2. (i) From Assumption (2.Ec), we deduce with GRONWALL's lemma the chain of inequalities

$$\mathrm{e}^{-C|t-s|}\mathcal{E}_s(u) \le \mathcal{E}_t(u) \le \mathrm{e}^{C|t-s|}\mathcal{E}_s(u) \quad \text{for all } s, t \in [0, T]. \tag{2.6}$$

In particular, there exists a constant $C_1 > 0$ such that

$$G(u) = \sup_{t\in[0,T]} \mathcal{E}_t(u) \le C_1 \inf_{t\in[0,T]} \mathcal{E}_t(u) \quad \text{for all } u \in D. \tag{2.7}$$

(ii) From Assumptions (2.Eb) and (2.Ec), we deduce the existence of a real number S which bounds the energy functional from below, i.e.,

$$\mathcal{E}_t(u) \ge S \quad \text{for all } u \in V,\ t \in [0, T]. \tag{2.8}$$

(iii) From the strong-weak closedness property of the graph of $\partial\mathcal{E}$ in Assumption (2.Ee) and MORDUKHOVICH [21, Lemma 2.32, p. 214], one can argue as in [23, Proposition 4.2, p. 273], in order to show the following variational sum rule:

If for $u_0 \in V$, $r > 0$, and $t \in [0, T]$ the point $u \in V$ is a global minimizer of $\Phi(\tau, t, u_0, w; \cdot)$, then

$$\exists\, \xi \in \partial\mathcal{E}_t(u): \quad w - \xi \in \partial\Psi_{u_0}\left(\frac{u - u_0}{r}\right) \tag{2.9}$$

or equivalently $w \in \partial\Psi_{u_0}\left(\frac{u-u_0}{r}\right) + \partial\mathcal{E}_{t+r}(u)$.

(iv) Assumption (2.Eb) and point $i)$ in this remark immediately yield that the functional $\mathcal{E}_t$ has compact sublevels for all $t \in [0, T]$.

(v) It is possible to relax Assumption (2.Ec) by assuming not the time differentiability but a kind of LIPSCHITZ continuity and a conditioned one-sided time differentiability of the map $t \mapsto \mathcal{E}_t(u)$, see [23]. We shall confine ourselves to Assumption (2.Ec) just to simplify the proofs.

Now, we collect the assumptions concerning the dissipation potential Ψ.

(2.Ψa) **Dissipation potential.** For all $u \in V$, the functional $\Psi_u : V \to [0, +\infty)$ is lower semicontinuous and convex with $\Psi(0) = 0$. Furthermore, if $w_1, w_2 \in \partial\Psi_u(v)$ for any $v \in V$, then $\Psi_u^*(w_1) = \Psi_u^*(w_2)$.

(2.Ψb) **Superlinearity.** The functionals Ψ_u and Ψ_u^* are coercive uniformly with respect to $u \in V$ in sublevels of $\mathcal{E}$, i.e., for all $R > 0$ it holds

$$\lim_{\|\xi\|_*\to+\infty} \frac{1}{\|\xi\|_*}\Big(\inf_{\substack{u\in V\\ G(u)\le R}} \Psi_u^*(\xi)\Big) = \infty, \qquad \lim_{\|v\|\to+\infty} \frac{1}{\|v\|}\Big(\inf_{\substack{u\in V\\ G(u)\le R}} \Psi_u(v)\Big) = \infty,$$

where $G(u) := \sup_{t\in[0,T]} \mathcal{E}_t(u)$ for all $u \in V$.

(2.Ψc) **State-dependence is Mosco continuous.** The functional Ψ is continuous in the sense of Mosco-convergence, i.e., for all $R > 0$ and sequences $(u_n)_{n\in\mathbb{N}} \subset V$ with $u_n \to u \in V$ as $n \to \infty$ and $\sup_{n\in\mathbb{N}} G(u_n) \le R$, we have $\Psi_{u_n} \xrightarrow{\mathrm{M}} \Psi_u$.

REMARK 2.3. (i) Since $\mathrm{dom}(\Psi_u) = V$ for all $u \in V$, the lower semicontinuity and convexity of Ψ_u yield continuity of Ψ_u on finite-dimensional subspaces and $\partial\Psi_u(v) \ne \emptyset$ for all $u \in V, u \in D$. Together with Assumption (2.Ψb), this implies that the Legendre- Fenchel conjugate Ψ^* is everywhere finite, i.e., $\mathrm{dom}(\Psi^*) = V^*$, and the operator $\partial\Psi_u : V \to 2^{V^*}$ is for all $u \in D$ bounded, i.e., it maps bounded subsets of V to bounded subsets of V^*. The former in turn entail the same properties for Ψ_u^* for all $u \in V$.

(ii) The Mosco convergence of $\Psi_{u_n} \xrightarrow{\mathrm{M}} \Psi_u$ from Assumption (2.Ψc) implies Mosco convergence of the dual potentials, namely $\Psi_{u_n}^* \xrightarrow{\mathrm{M}} \Psi_n^*$, see (2.1). In particular, this implies that for all $R > 0$, all sequences $(u_n)_{n\in\mathbb{N}} \subset V$ with $u_n \to u \in V$ and $\sup_{n\in\mathbb{N}} G(u_n) \le R$, and all sequences $(\xi_n)_{n\in\mathbb{N}} \subset V^*$ with $\xi_n \rightharpoonup \xi \in V^*$ we have

$$\Psi_u^*(\xi) \le \liminf_{n\to\infty} \Psi_{u_n}^*(\xi_n). \tag{2.10}$$

Finally, we make the following assumptions on the non-variational perturbation B.

(2.Ba) **Continuity.** The map $(t, u) \mapsto B(t, u) : [0, T] \times V \to V^*$ is continuous on sublevels of G, i.e., $(t_n, u_n) \to (t, u)$ in $[0, T]\times V$ and $\sup_{n\in\mathbb{N}} G(u_n) \le R$ implies $B(t_n, u_n) \to B(t, u)$ in V^*.

(2.Bb) **Control of B by the energy.** There exist $\beta > 0$ and $c \in (0, 1)$ such that

$$c\,\Psi_u^*\left(\frac{1}{c}B(t,u)\right) \le \beta\big(1 + \mathcal{E}_t(u)\big) \quad \text{for all } u \in D,\ t \in [0, T].$$

REMARK 2.4. We note that Assumption (2.Ba) ensures that the Nemytskij operator associated with B maps strongly measurable functions contained in sublevels of G to strongly measurable functions, i.e., for all strongly measurable functions u with $\sup_{t\in[0,T]} G(u(t)) \le R$, the map $t \mapsto B(t, u(t))$ is strongly measurable.

2.4. Statement of the existence result

Before we state the main result, we say that $u \in \mathrm{AC}([0, T]; V)$ is a solution to (1.1) with the initial datum $u_0 \in D$ if u satisfies the differential inclusion (1.1) with $u(0) = u_0$.

THEOREM 2.5. (Main existence result for $\mathrm{PG} = (V, \mathcal{E}, \Psi, B)$) *Let the perturbed gradient system* $(V, \mathcal{E}, \Psi, B)$ *satisfy Assumptions* (2.E), (2.Ψ)*, and* (2.B)*. Then, for every* $u_0 \in D$ *there exists a solution* $u \in \mathrm{AC}([0, T]; V)$ *to* (1.1) *with* $u(0) = u_0$ *and an integrable function* $\xi \in \mathrm{L}^1(0, T; V)$ *with* $\xi(t) \in \partial\mathcal{E}_t(u(t))$ *for a.a.* $t \in (0, T)$ *such* *that the following energy-dissipation balance holds:*

$$\mathcal{E}_t(u(t)) + \int_s^t \Big(\Psi_{u(r)}(u'(r)) + \Psi^*_{u(r)}\big(B(r,u(r)) - \xi(r)\big)\Big)\,\mathrm{d}r \\ = \mathcal{E}_s(u(s)) + \int_s^t \partial_r \mathcal{E}_r(u(r))\,\mathrm{d}r + \int_s^t \langle B(r,u(r)), u'(r)\rangle\,\mathrm{d}r \quad \textit{for all } s,t \in [0,T]. \tag{2.11}$$

It is clear that every solution of (2.11) is already a solution for the perturbed gradient system PG $= (V, \mathcal{E}, \Psi, B)$, since by the chain rule can and the LEGENDRE- FENCHEL theory we easily recover (1.1), see, e.g., [1,29].

Our proof will be done by time discretization and solving variational problems for each time interval $(t_n, t_{n+1}]$. To obtain a useful discrete counterpart of the energy-dissipation balance, we employ DE GIORGI's variational interpolant, see [2, Lemma 2.5] or [29, Sect. 4.2]. We then follow the ideas in [23], but need to generalize these to the case of a non-trivial perturbation B, which only satisfies our mild assumptions (2.Ba) and (2.Bb). The proof will be completed in Sect. 2.7.

2.5. Estimates on the MOREAU- YOSIDA regularization

In order to prove the existence result, we need to show some properties of the Ψ-MOREAU- YOSIDA regularization

$$\Phi_{r,t}(w;u) := \inf_{v\in V} \Phi(r,t,u,w;v)$$

for $r > 0$, $t \in [0,T)$ with $r + t \in [0,T]$ and $u \in D$ as well as $w \in V^*$. Therefore, we have to ensure that the resolvent set $J_{r,t}(w;u) := \arg\min_{v\in V} \Phi(r,t,u,w;v)$ is not empty.

LEMMA 2.6. *Let the perturbed gradient system* $(V, \mathcal{E}, \Psi, B)$ *satisfy Assumptions* (2.Ea)-(2.Eb) *and* (2.Ψa). *Then, for all* $r > 0$*,* $t \in [0,T)$ *with* $t + r \le T$*,* $u \in D$*, and* $w \in V^*$*, the resolvent set* $J_{r,t}(w;u)$ *is non-empty.*

Proof. Let $u \in D$, $w \in V^*$ and $r > 0$, $t \in [0,T)$ with $r + t \in [0,T]$ be given. First of all, we see with the FENCHEL- YOUNG inequality and with (2.8) that

$$\begin{aligned}\Phi(r,t,u,w;v) &= r\Psi_u\left(\frac{v-u}{r}\right) + \mathcal{E}_{t+r}(v) - \langle w, v\rangle \\ &\ge -r\Psi^*_u(w) + \mathcal{E}_{t+r}(v) - \langle w, u\rangle \\ &\ge -r\Psi^*_u(w) + S - \langle w, u\rangle.\end{aligned} \tag{2.12}$$

This implies $\Phi_{r,t}(w;u) > -\infty$. On the other hand, we observe that

$$\inf_{v\in V}\left\{ r\Psi_u\left(\frac{v-u}{r}\right) + \mathcal{E}_{t+r}(v) - \langle w, v\rangle\right\} \le \mathcal{E}_{t+r}(u) - \langle w, u\rangle, \tag{2.13}$$

so that we also have $\Phi_{r,t}(w;u) < +\infty$. Let now $(v_n)_{n\in\mathbb{N}} \subset V$ be a minimizing sequence for $\Phi(r,t,u,w;\cdot)$. From (2.12), we deduce with (2.6) that $(v_n)_{n\in\mathbb{N}} \subset V$ is contained in a sublevel set of $\mathcal{E}_t$. Thus, by Assumption (2.Eb) and Remark 2.2 *iv*) there exists a subsequence (not relabeled) which converges strongly in V toward a limit $v \in V$. Together with the lower semicontinuity of the map $v \mapsto \Phi(r,t,u,w;v)$, we have

$$\Phi(r,t,u,w;v) \le \liminf_{n\to\infty} \Phi(r,t,u,w;v_n) = \inf_{\tilde{v}\in V} \Phi(r,t,u,w;\tilde{v})$$

and therefore $v \in J_{r,t}(w;u) \neq \emptyset$ from which $v \in D$ follows. □

Lemma 2.6 is important for justifying the existence of a sequence of approximate values $(U_\tau^n)_{n=1}^N \subset D$ that complies with

$$U_\tau^n \in J_{\tau,t_{n-1}}(B(t_{n-1},U_\tau^{n-1}),U_\tau^{n-1}) \quad \text{for all } n=1,\dots,N, \tag{2.14}$$

in order to construct discrete solutions to (2.3), where $U_\tau^0 := u_0$ and the time $t \in [0,T)$ as well as the time step $\tau \in (0, T-t)$ are fixed.

The following lemma is crucial in order to proof the existence result and in particular to derive a priori estimates for the interpolation functions we define later on. The result is an adaptation to the case $w \neq 0$ of [29, Lemma 4.2] and [23, Lem. 6.1].

LEMMA 2.7. *Let the perturbed gradient system* $(V, \mathcal{E}, \Psi, B)$ *satisfy Assumptions* (2.E), (2.Ψ), *and* (2.B). *Then, for every* $\mathrm{t} \in [0,T)$, $u \in D$ *and* $w \in V^*$ *there exists a measurable selection* $r \mapsto u_r : (0, T-\mathrm{t}) \to J_{r,\mathrm{t}}(w;u)$ *such that*

$$w \in \partial\Psi_u\left(\frac{u_r-u}{r}\right) + \partial\mathcal{E}_{\mathrm{t}+r}(u) \tag{2.15}$$

and there exists a constant $\widetilde{C} > 0$ *such that*

$$G(u_r) \le \widetilde{C}(G(u) + r\Psi_u^*(w)) \quad \textit{for all } r \in (0, T-\mathrm{t}). \tag{2.16}$$

Furthermore, it holds

$$\lim_{r\to 0} \sup_{u_r\in J_{r,\mathrm{t}}(w;u)} \|u_r-u\| = 0 \quad \textit{and} \quad \lim_{r\to 0} \Phi_{r,\mathrm{t}}(w;u) = \mathcal{E}_{\mathrm{t}}(u) - \langle w,u\rangle \tag{2.17}$$

for all $\mathrm{t} \in [0,T)$, $u \in D$ *and* $w \in V^*$. *Finally, the map* $r \mapsto \Phi_{r,\mathrm{t}}(w;u)$ *is almost everywhere differentiable in* $(0, T-\mathrm{t})$*; and for every* $r_0 \in (0, T-\mathrm{t})$ *and every measurable selection* $r \mapsto u_r : (0,r_0) \to J_{r,\mathrm{t}}(w;u)$, *there exists a measurable selection* $r \mapsto \xi_r : (0, T-\mathrm{t}) \to \partial\mathcal{E}_{\mathrm{t}+r}(u)$ *with* $w - \xi_r \in \partial\Psi_u\left(\frac{u_r-u}{r}\right)$ *such that*

$$\mathcal{E}_{\mathrm{t}+r_0}(u_{r_0}) + r_0\Psi_u\left(\frac{u_{r_0}-u}{r_0}\right) + \int_0^{r_0} \Psi_u^*(w-\xi_r)\,\mathrm{d}r \le \mathcal{E}_{\mathrm{t}}(u) + \int_0^{r_0} \partial_r\mathcal{E}_{\mathrm{t}+r}(u_r)\,\mathrm{d}r + \langle w, u_{r_0}-u\rangle. \tag{2.18}$$

Proof. Let $\mathrm{t} \in [0, T)$, $u \in D$ and $w \in V^*$ be given. The non-emptiness of the resolvent set $J_{r,\mathrm{t}}(w; u)$ for all $r \in (0, T - \mathrm{t})$ is guaranteed by Lemma 2.7. The existence of a measurable selection $r \mapsto u_r : (0, T - \mathrm{t}) \to J_{r,\mathrm{t}}(w; u)$ is provided by Castaing and Valadier [10, Cor. III.3, Prop. III.4, Thm. III.6, pp. 63]. The inclusion (2.15) follows then by the variational sum rule (2.9). Further, we obtain from (2.12) for $v = u_r$, $r \in (0, T - \mathrm{t})$ and (2.13) the inequality

$$\mathcal{E}_{\mathrm{t}+r}(u_r) \leq \mathcal{E}_{\mathrm{t}+r}(u) + r\Psi_u^*(w),$$

so that together with the estimate (2.7) the inequality (2.16) with $\widetilde{C} = C_1$ follows, where $C_1 > 0$ is the constant from (2.7). In order to show the convergences in (2.17), we note that Assumption (2.Ψb) implies: For all $R > 0$ and $\gamma > 0$, there exists $K > 0$ such that

$$\Psi_u(v) \geq \gamma \|v\|$$

for all $u \in D$ with $G(u) \leq R$ and all $v \in V$ with $\|v\| \leq K$. Based on this fact, we infer

$$\gamma \left\| \frac{u_r - u}{r} \right\| \leq \Psi_u\left(\frac{u_r - u}{r}\right) + \gamma K \quad \text{for every } r > 0. \tag{2.19}$$

Together with (2.8), (2.12) and (2.13), we obtain

$$\begin{aligned}\gamma \|u_r - u\| &\leq \langle w, u_r - u\rangle + \mathcal{E}_{\mathrm{t}+r}(u) - \mathcal{E}_{\mathrm{t}+r}(u_r) + r\gamma K \\ &\leq \|w\|_* \|u_r - u\| + \mathcal{E}_{\mathrm{t}+r}(u) - S + r\gamma K.\end{aligned}$$

This implies the estimate

$$(\gamma - \|w\|_*)\|u_r - u\| \leq \mathcal{E}_{\mathrm{t}+r}(u) - S + r\gamma K \leq \mathrm{e}^{CT}\mathcal{E}_0(u) - S + r\gamma K$$

for all $r \in (0, T - \mathrm{t})$ and $u_r \in J_{r,\mathrm{t}}(w; u)$, where we used (2.6) again. By taking the supremum over all $u_r \in J_{r,\mathrm{t}}(w; u)$ and taking the limes superior as $r \to 0$, we finally obtain

$$(\gamma - \|w\|_*) \limsup_{r\to 0} \sup_{u_r \in J_{r,\mathrm{t}}(w;u)} \|u_r - u\| \leq \mathrm{e}^{CT}\mathcal{E}_0(u) - S \quad \text{for every } \gamma > \|w\|_*.$$

By choosing $\gamma > 0$ sufficiently large, we conclude

$$\limsup_{r\to 0} \sup_{u_r \in J_{r,\mathrm{t}}(w;u)} \|u_r - u\| = 0,$$

which shows the first convergence in (2.17). We now use the lower semicontinuity, the time continuity of the energy functional, the estimate

$$\begin{aligned}&\mathcal{E}_{t+r}(u_r) - \langle w, u_r\rangle \le \Phi_{r,t}(w; u)\\&= r\Psi_u\left(\frac{u_r - u}{r}\right) + \mathcal{E}_{t+r}(u_r) - \langle w, u_r\rangle \le \mathcal{E}_{t+r}(u) - \langle w, u\rangle,\end{aligned}$$

and the fact that $\liminf_{r\to 0} \mathcal{E}_{t+r}(u_r) = \liminf_{r\to 0} \mathcal{E}_t(u_r)$, which follows from (2.6). Hence, the second convergence in (2.17) follows from the estimate

$$\begin{aligned}\mathcal{E}_t(u) - \langle w, u\rangle &\le \liminf_{r\to 0} \left(\mathcal{E}_{t+r}(u_r) - \langle w, u_r\rangle\right)\\&\le \liminf_{r\to 0} \Phi_{r,t}(w; u) \le \limsup_{r\to 0} \Phi_{r,t}(w; u)\\&\le \limsup_{r\to 0} \left(\mathcal{E}_{t+r}(u) - \langle w, u\rangle\right) = \mathcal{E}_t(u) - \langle w, u\rangle.\end{aligned}$$

In order to show the last assertion of this lemma, let $u_{r_i} \in J_{r,t}(w; u)$, $i = 1, 2$, with $0 < r_1 < r_2 < T - t$. Then, we have

$$\begin{aligned}&\Phi_{r_2,t}(w; u) - \Phi_{r_1,t}(w; u) - \left(\mathcal{E}_{t+r_2}(u_{r_1}) - \mathcal{E}_{t+r_1}(u_{r_1})\right)\\&\le r_2\Psi_u\left(\frac{u_{r_1} - u}{r_2}\right) - r_1\Psi_u\left(\frac{u_{r_1} - u}{r_1}\right)\\&= (r_2 - r_1)\Psi_u\left(\frac{u_{r_1} - u}{r_2}\right) + r_1\left(\Psi_u\left(\frac{u_{r_1} - u}{r_2}\right) - \Psi_u\left(\frac{u_{r_1} - u}{r_1}\right)\right)\\&\le (r_2 - r_1)\left(\Psi_u\left(\frac{u_{r_1} - u}{r_2}\right) - \left\langle w_2^1, \frac{u_{r_1} - u}{r_2}\right\rangle\right) \qquad (2.20)\\&= -(r_2 - r_1)\Psi_u^*(w_2^1) \le 0, \qquad (2.21)\end{aligned}$$

where in (2.20) we used Remark 2.3 *i*), which states $w_2^1 \in \partial\Psi_u\left(\frac{u_{r_1}-u}{r_2}\right) \ne \emptyset$, in (2.21) we used the statement of Lemma 2.1, and for the last inequality we used that by the FENCHEL- YOUNG inequality we have $\Psi_u^*(w) \ge 0$ for all $w \in V^*$. Further, we deduce with the aid of (2.Ec), (2.7), and the already proven inequality (2.16) that

$$\begin{aligned}\Phi_{r_2,t}(w; u) &\le \Phi_{r_1,t}(w; u) + \left(\mathcal{E}_{t+r_2}(u_{r_1}) - \mathcal{E}_{t+r_1}(u_{r_1})\right)\\&= \Phi_{r_1,t}(w; u) - \int_{r_1}^{r_2} \partial_r\mathcal{E}_{t+r}(u_{r_1})\,\mathrm{d}r\\&\le \Phi_{r_1,t}(w; u) + (r_2 - r_1)\, CC_1 G(u_{r_1})\\&\le \Phi_{r_1,t}(w; u) + (r_2 - r_1)\, CC_1(G(u) + r_1\Psi^*(w))\\&\le \Phi_{r_1,t}(w; u) + (r_2 - r_1)\, CC_1(G(u) + T\Psi^*(w)). \qquad (2.22)\end{aligned}$$

We conclude that the map $r \mapsto \Phi_{r,t}(w; u) - rCC_1(G(u) + T\Psi^*(w))$ is non-increasing on $(0, T - t)$ and therefore, as a real-valued function, is almost everywhere differentiable. Since the map $r \mapsto \Phi_{r,t}(w; u)$ is a linear perturbation of a monotone function, it is also almost everywhere differentiable in $(0, T - t)$. Thus, there exists a negligible set $\mathcal{N} \subset (0, T - t)$, such that the map $r \mapsto \Phi_{r,t}(w; u)$ is differentiable on $(0, T - t)\backslash\mathcal{N}$. We remark that the negligible set depends on u and w, that is $\mathcal{N} = \mathcal{N}_{u,w}$. Now, to conclude, we want to use the inequality (2.21). For this, let $r \in (0, T-\mathrm{t})\backslash\mathcal{N}$ be fixed. Additionally let $(h_n)_{n\in\mathbb{N}} \in \mathbb{R}^{>0}$ be a sequence which converges from above toward zero and whose elements are sufficiently small. Let also the sequence $(w_n^r)_{n\in\mathbb{N}} \subset V^*$ be given by $w_n^r \in \partial\Psi_u\left(\frac{u_r-u}{r+h_n}\right)$ for all $n \in \mathbb{N}$. The boundedness of the operator $\partial\Psi_u$ according to Remark 2.3 i) implies that the sequence $(w_n^r)_{n\in\mathbb{N}} \subset V^*$ is bounded in V^*. Thus, there exists a subsequence, labeled as before, and an element $w_r \in V^*$ such that $w_n^r \rightharpoonup w_r$ weakly in V^*. From the strong-weak closedness of the graph of $\partial\Psi_u$ in $V \times V^*$, it follows that $w_r \in \partial\Psi_u\left(\frac{u_r-u}{r}\right)$. Since the conjugate Ψ_u^* is convex and lower semicontinuous, it is also weakly lower semicontinuous. Then, with Lemma 2.1 and the continuity of Ψ_u on finite-dimensional subspaces (Remark (2.3)) we find that

$$\begin{aligned}\Psi_u^*(w_r) &\le \liminf_{n\to\infty} \Psi_u^*(w_n^r) \le \limsup_{n\to\infty} \Psi_u^*(w_n^r)\\ &= \limsup_{n\to\infty}\left(\left\langle w_n^r, \frac{u_r-u}{r+h_n}\right\rangle - \Psi_u\left(\frac{u_r-u}{r+h_n}\right)\right)\\ &= \left\langle w_r, \frac{u_r-u}{r}\right\rangle - \Psi_u\left(\frac{u_r-u}{r}\right) = \Psi_u^*(w_r)\end{aligned}$$

and thus $\lim_{n\to\infty}\Psi_u^*(w_n^r) = \Psi_u^*(w_r)$. Due to the inclusion (2.15), there exists $\xi_r \in \partial\mathcal{E}_{\mathrm{t}+r}(u)$ such that $w-\xi_r \in \partial\Psi_u\left(\frac{u_r-u}{r}\right)$. By Aubin and Frankowska [5, Thm. 8.2.9, p. 315], the selection $r \mapsto \xi_r : (0, T-\mathrm{t}) \to \partial\mathcal{E}_{\mathrm{t}+r}(u)$ can be chosen to be measurable. Further, from Assumption (2.Ψa) we get $\Psi_u^*(w_r) = \Psi_u^*(w-\xi_r)$. By the differentiability of the map $r \mapsto \Phi_{r,\mathrm{t}}(w; u)$ in r, we obtain with (2.21)

$$\begin{aligned}&\frac{\mathrm{d}}{\mathrm{d}r}\Phi_{r,\mathrm{t}}(w;u)|_{r=r} + \Psi_u^*(w-\xi_r) = \lim_{n\to\infty}\left(\frac{\Phi_{r+h_n,\mathrm{t}}(w;u) - \Phi_{r,\mathrm{t}}(w;u)}{h_n} + \Psi_u^*(w_n^r)\right)\\ &\le \liminf_{n\to\infty}\left(\frac{\mathcal{E}_{\mathrm{t}+r+h_n}(u_r) - \mathcal{E}_{\mathrm{t}+r}(u_r)}{h_n}\right) = \partial_t\mathcal{E}_{\mathrm{t}+r}(u_r) \quad \text{for a.a. } r \in (0, T-\mathrm{t}),\end{aligned} \tag{2.23}$$

where we also used the fact that the map $t \mapsto \mathcal{E}_t$ is differentiable. The claim finally follows by integrating (2.23) from $r = 0$ to $r = r_0$ and by using (2.17). □

2.6. Time discretization and discrete energy-dissipation estimate

With the help of the preceding lemma, we derive in the forthcoming result a priori estimates for the approximate solutions, more precisely for both the piecewise constant interpolation functions $\overline{U}_\tau$ and $\underline{U}_\tau$, and for the piecewise linear interpolation function $\widehat{U}_\tau$ as well as for the so-called De Giorgi interpolation function $\widetilde{U}_\tau$. In order to define the interpolation functions, let the initial value $u_0 \in D$ and the time step $\tau > 0$ be fixed. Further let $(U_\tau^n)_{n=1}^N \subset D$ be the sequence of approximate values, which are defined by the variational approximation scheme

$$\begin{cases} U_\tau^0 = u_0,\\ U_\tau^n \in J_\tau(B(t_{n-1}, U_\tau^{n-1}); U_\tau^{n-1})), \quad n = 1, 2, \ldots, N.\end{cases} \tag{2.24}$$

We define the piecewise constant and linear interpolation functions by

$$\begin{aligned}
&\overline{U}_\tau(0) = \underline{U}_\tau(0) = \widehat{U}_\tau(0) := U_\tau^0 \quad \text{and} \\
&\underline{U}_\tau(t) := U^{n-1}, \quad \widehat{U}_\tau(t) := \frac{t_n - t}{\tau} U_\tau^{n-1} + \frac{t - t_{n-1}}{\tau} U_\tau^n \quad \text{for } t \in [t_{n-1}, t_n), \\
&\overline{U}_\tau(t) := U_\tau^n \quad \text{for } t \in (t_{n-1}, t_n] \quad \text{and all } n = 1, \ldots, N.
\end{aligned} \tag{2.25}$$

Furthermore, we define by the approximation scheme

$$\begin{cases} \widetilde{U}_\tau(0) := U_\tau^0, \\ \widetilde{U}_\tau(t) \in J_r(B(t_{n-1}, U_\tau^{n-1}); U_\tau^{n-1})) \ \text{for } t = t_{n-1}+r \in (t_{n-1}, t_n], \end{cases} \tag{2.26}$$

$n = 1, 2, \ldots, N$, the DE GIORGI interpolation $\widetilde{U}_\tau$. We note that we can assume the measurability of the function $\widetilde{U}_\tau$, since by Lemma 2.7 there always exists a measurable selection of the DE GIORGI interpolation. Due to the fact that for all $t \in I_\tau$ the approximation scheme (2.26) yields the usual scheme (2.24), we can assume without loss of generality that all interpolation functions coincide on the nodes t_n, i.e.,

$$\widetilde{U}_\tau(t_n) = \overline{U}_\tau(t_n) = \underline{U}_\tau(t_n) = \widehat{U}_\tau(t_n) = U_\tau^n \quad \text{for all } n = 1, \ldots, N.$$

Moreover, we denote by $\widetilde{\xi}_\tau$ the interpolation function obtained from Remark 2.2 *iii*) with the variational sum rule by choosing $t = t_{n-1}, u_0 = \widetilde{U}_\tau(t), u = U_\tau^{n-1}$ and $w = B(t_{n-1}, U_\tau^{n-1})$, and which satisfies

$$\widetilde{\xi}_\tau(t) \in \partial \mathcal{E}_{t_{n-1}+r}(\widetilde{U}_\tau(t)) \quad \text{for } t = t_{n-1} + r \in (t_{n-1}, t_n], \tag{2.27}$$

as well as

$$B(t_{n-1}, U_\tau^{n-1}) - \widetilde{\xi}_\tau(t) \in \partial \Psi_{U_\tau^{n-1}} \left(\frac{\widetilde{U}_\tau(t) - U_\tau^{n-1}}{t - t_{n-1}} \right) \ \text{for } t = t_{n-1}+r \in (t_{n-1}, t_n]$$

for all $n = 1, \ldots, N$. The measurability of the function $\widetilde{\xi}_\tau : (0, T) \to V^*$, again, follows from Lemma 2.7.

For notational convenience, we also introduce the piecewise constant interpolation functions $\bar{\mathbf{t}}_\tau : [0, T] \to [0, T]$ and $\underline{\mathbf{t}}_\tau : [0, T] \to [0, T]$ given by

$$\begin{aligned}
&\bar{\mathbf{t}}_\tau(0) := 0 \quad \text{and } \bar{\mathbf{t}}_\tau(t) := t_n \qquad \text{for } t \in (t_{n-1}, t_n], \quad n = 1, \ldots, N, \\
&\underline{\mathbf{t}}_\tau(T) := T \ \text{ and } \underline{\mathbf{t}}_\tau(t) := t_{n-1} \quad \text{for } t \in [t_{n-1}, t_n), \quad n = 1, \ldots, N.
\end{aligned}$$

Obviously, it holds $\bar{\mathbf{t}}_\tau(t) \to t$ and $\underline{\mathbf{t}}_\tau(t) \to t$ as $\tau \to 0$.

We are now in the position to show a priori estimates for the approximate solutions.

LEMMA 2.8. *Let the perturbed gradient system* $(V, \mathcal{E}, \Psi, B)$ *satisfy Assumptions* (2.E), (2.Ψ), *and* (2.B). *Furthermore, let* $\widetilde{U}_\tau, \overline{U}_\tau, \underline{U}_\tau, \widehat{U}_\tau$ *and* $\widetilde{\xi}_\tau$ *be the interpolation* *functions defined in* (2.25)–(2.27) *associated with a fixed initial datum* $u_0 \in D$ *and a step size* $\tau > 0$. *Then, the discrete upper energy estimate*

$$\mathcal{E}_{\bar{\mathbf{t}}_\tau(t)}(\overline{U}_\tau(t)) + \int_{\bar{\mathbf{t}}_\tau(s)}^{\bar{\mathbf{t}}_\tau(t)} \Big(\Psi_{\underline{U}_\tau(r)}\big(\widehat{U}'_\tau(r)\big) + \Psi^*_{\underline{U}_\tau(r)}\big(B(\underline{\mathbf{t}}_\tau(r), \underline{U}_\tau(r)) - \widetilde{\xi}_\tau(r)\big)\Big)\,\mathrm{d}r$$
$$\leq \mathcal{E}_{\bar{\mathbf{t}}_\tau(s)}(\overline{U}_\tau(s)) + \int_{\bar{\mathbf{t}}_\tau(s)}^{\bar{\mathbf{t}}_\tau(t)} \partial_r\Big(\mathcal{E}_r(\widetilde{U}_\tau(r)) + \big\langle B(\underline{\mathbf{t}}_\tau(r), \underline{U}_\tau(r)), \widehat{U}'_\tau(r)\big\rangle\Big)\,\mathrm{d}r \tag{2.28}$$

holds for all $0 \leq s < t \leq T$. *Moreover, there exist positive constants* $M, \tau^* > 0$ *such that the estimates*

$$\sup_{t\in[0,T]} \mathcal{E}_t(\overline{U}_\tau(t)) \leq M, \quad \sup_{t\in[0,T]} \mathcal{E}_t(\widetilde{U}_\tau(t)) \leq M, \quad \sup_{t\in[0,T]} |\partial_t\mathcal{E}_t(\widetilde{U}_\tau(t))| \leq M \tag{2.29}$$

$$\int_0^T \Big(\Psi_{\underline{U}_\tau(r)}\big(\widehat{U}'_\tau(r)\big) + \Psi^*_{\underline{U}_\tau(r)}\big(B(\underline{\mathbf{t}}_\tau(r), \underline{U}_\tau(r)) - \widetilde{\xi}_\tau(r)\big)\Big)\,\mathrm{d}r \leq M \tag{2.30}$$

hold for all $0 < \tau \leq \tau^*$. *Besides, the families* $(\widehat{U}'_\tau)_{0<\tau\leq\tau^*} \subset \mathrm{L}^1(0,T;V)$ *as well as* $(B(\underline{\mathbf{t}}_\tau, \underline{U}_\tau))_{0<\tau\leq\tau^*} \subset \mathrm{L}^1(0,T;V^*)$ *and* $(\widetilde{\xi}_\tau)_{0<\tau\leq\tau^*} \subset \mathrm{L}^1(0,T;V^*)$ *are uniformly integrable with respect to* τ *in the respective spaces. Finally, it holds*

$$\|\underline{U}_\tau - \overline{U}_\tau\|_{\mathrm{L}^\infty(0,T;V)} + \|\widehat{U}_\tau - \overline{U}_\tau\|_{\mathrm{L}^\infty(0,T;V)} + \|\widetilde{U}_\tau - \underline{U}_\tau\|_{\mathrm{L}^\infty(0,T;V)} \to 0 \tag{2.31}$$

as $\tau \to 0$.

Proof. In order to show the discrete upper energy estimate (2.28), it is sufficient to restrict ourselves to the case $s = t_{n-1}$ and $t = t_n$ for $n \in n = 1, \ldots, N$. The general case follows by summing up the particular inequalities on the subintervals. The special case follows from (2.18) in Lemma 2.7 by choosing $\mathrm{t} = t_{n-1}, u = U_\tau^{n-1}, r_0 = t - t_{n-1}, u_{r_0} = \widetilde{U}_\tau(t), u_r = \widetilde{U}_\tau(t_{n-1}+r)$ and $\xi_r = \widetilde{\xi}_\tau(t_{n-1}+r)$, where we chose $t \in (t_{n-1}, t_n]$ to be fixed. Then, we find

$$(t-t_{n-1})\Psi_{U_\tau^{n-1}}\left(\frac{\widetilde{U}_\tau(t)-U_\tau^{n-1}}{t-t_{n-1}}\right) + \int_{t_{n-1}}^t \Psi_{U_\tau^{n-1}}\Big(B(t_{n-1}, U_\tau^{n-1}) - \widetilde{\xi}_\tau(r)\Big)\,\mathrm{d}r$$
$$+ \mathcal{E}_t(\widetilde{U}_\tau(t))$$
$$\leq \mathcal{E}_{t_{n-1}}(U_\tau^{n-1}) + \int_{t_{n-1}}^t \partial_r\mathcal{E}_r(\widetilde{U}_\tau(r))\,\mathrm{d}r + \langle B(t_{n-1}, U_\tau^{n-1}), U_\tau^n - U_\tau^{n-1}\rangle. \tag{2.32}$$

By choosing $t = t_n$, we obtain

$$\int_{t_{n-1}}^{t_n} \left(\Psi_{\underline{U}_\tau(r)} \left(\widehat{U}_\tau'(r) \right) + \Psi^*_{\underline{U}_\tau(r)} \left(B(t_{n-1}, \underline{U}_\tau(r)) - \widetilde{\xi}_\tau(r) \right) \right) \mathrm{d}r + \mathcal{E}_{t_n}(\overline{U}_\tau(t_n))$$
$$\leq \mathcal{E}_{t_{n-1}}(\underline{U}_\tau(t_{n-1})) + \int_{t_{n-1}}^{t_n} \partial_r \mathcal{E}_r(\widetilde{U}_\tau(r))\,\mathrm{d}r + \int_{t_{n-1}}^{t_n} \langle B(t_{n-1}, \underline{U}_\tau(r)), \widehat{U}_\tau'(r) \rangle\,\mathrm{d}r \tag{2.33}$$

for all $n = 1, \ldots, N$, which yields the discrete upper energy estimate. Further, we notice that from Assumption (2.Bb), we obtain the estimation

$$\begin{aligned}
&\int_{t_{n-1}}^{t_n} \langle B(t_{n-1}, \underline{U}_\tau(r)), \widehat{U}_\tau'(r) \rangle\,\mathrm{d}r \\
&\quad \leq c \int_{t_{n-1}}^{t_n} \Psi_{\underline{U}_\tau(r)} \left(\widehat{U}_\tau'(r) \right) \mathrm{d}r + c \int_{t_{n-1}}^{t_n} \Psi^*_{\underline{U}_\tau(r)} \left(\frac{B(t_{n-1}, \underline{U}_\tau(r))}{c} \right) \mathrm{d}r \\
&\quad \leq c \int_{t_{n-1}}^{t_n} \Psi_{\underline{U}_\tau(r)} \left(\widehat{U}_\tau'(r) \right) \mathrm{d}r + \tau\beta(1 + \mathcal{E}_{t_{n-1}}(U_\tau^{n-1})) \\
&\quad \leq c \int_{t_{n-1}}^{t_n} \Psi_{\underline{U}_\tau(r)} \left(\widehat{U}_\tau'(r) \right) \mathrm{d}r + \tau\beta(1 + G(U_\tau^{n-1})),
\end{aligned} \tag{2.34}$$

where we also used the FENCHEL- YOUNG inequality. Since $c \in (0, 1)$, inequality (2.33) and (2.34) together yield the estimation

$$\begin{aligned}
\mathcal{E}_{t_n}(U_\tau^n) &\leq \mathcal{E}_{t_{n-1}}(U_\tau^{n-1}) + \int_{t_{n-1}}^{t_n} \partial_r \mathcal{E}_r(\widetilde{U}_\tau(r))\,\mathrm{d}r + \tau\beta(1 + G(U_\tau^{n-1})) \\
&\leq \mathcal{E}_{t_{n-1}}(U_\tau^{n-1}) + \tau\beta(1 + G(U_\tau^{n-1})) + C\widetilde{C} \int_{t_{n-1}}^{t_n} G(U_\tau^{n-1})\,\mathrm{d}r \\
&\quad + \int_{t_{n-1}}^{t_n} (r - t_{n-1}) \Psi^*_{U_\tau^{n-1}}(B(t_{n-1}, U_\tau^{n-1}))\,\mathrm{d}r \qquad (2.35) \\
&\leq \mathcal{E}_{t_{n-1}}(U_\tau^{n-1}) + \tau\beta(1 + G(U_\tau^{n-1})) + C\widetilde{C} \int_{t_{n-1}}^{t_n} G(U_\tau^{n-1})\,\mathrm{d}r \\
&\quad + \int_{t_{n-1}}^{t_n} c\tau \Psi^*_{\underline{U}_\tau(r)} \left(\frac{B(t_{n-1}, \underline{U}_\tau(r))}{c} \right) \mathrm{d}r \qquad (2.36) \\
&\leq \mathcal{E}_{t_{n-1}}(U_\tau^{n-1}) + \tau\beta(1 + G(U_\tau^{n-1})) + C\widetilde{C}\tau G(U_\tau^{n-1}) \\
&\quad + \tau\beta(1 + G(U_\tau^{n-1})) \\
&= \mathcal{E}_{t_{n-1}}(U_\tau^{n-1}) + \tau(2\beta + C\widetilde{C})G(U_\tau^{n-1}) + 2\tau\beta \qquad (2.37)
\end{aligned}$$

for all $n = 1, \ldots, N$ and $0 < \tau \leq 1$, where in (2.35) we used the inequality

$$G(\widetilde{U}_\tau(t)) \leq \widetilde{C}(G(U_\tau^{n-1}) + (t - t_{n-1})\Psi^*_{U_\tau^{n-1}}(B(t_{n-1}, U_\tau^{n-1}))), \quad t \in (t_{n-1}, t_n],$$

from Lemma 2.7 and in (2.36) the fact that the map $r \mapsto r\Psi_u^*\left(\frac{\xi}{r}\right)$ is non-decreasing on $(0, +\infty)$ for every $\xi \in V^*$. Defining $A := (2\beta + C\widetilde{C})$ and summing up the inequalities (2.37), we obtain

$$G(U_\tau^n) \leq C_1 \mathcal{E}_{t_n}(U_\tau^n) \leq C_1 \mathcal{E}_0(u_0) + 2C_1 T\beta + \tau C_1 A \sum_{k=1}^{n} G(U_\tau^{k-1}) \tag{2.38}$$

for all $n = 1, \ldots, N$ and $0 < \tau \leq 1$. Then, applying the discrete version of the Gronwall Lemma to (2.38) yields the uniform boundedness of $G(U_\tau^n)$ for all $n = 1, \ldots, N$ and $0 < \tau < \min\{1, 1/(2C_1 A)\} =: \tau^*$, from which we deduce

$$\sup_{t\in(0,T)} \mathcal{E}_t(\overline{U}_\tau(t)) \leq C_1 \quad \text{for all } 0 < \tau < \tau^* \tag{2.39}$$

for a positive constant $C_1 > 0$ independent from τ. Taking into account the inequality (2.37) and Assumptions (2.Bb) and (2.Ec), we also obtain the last two inequalities in (2.29). By employing (2.33) and (2.34), and arguing as before, we also get (2.30). The constant M can be chosen as the sum of all constants obtained from the shown inequalities of this lemma. Further, the uniform integrability of $(\widehat{U}_\tau')_{0<\tau\leq\tau^*}$ as well as $(B(\underline{\mathbf{t}}_\tau, \underline{U}_\tau))_{0<\tau\leq\tau^*}$ and $(\widetilde{\xi}_\tau)_{0<\tau\leq\tau^*}$ in $\mathrm{L}^1(0, T; V)$ and $\mathrm{L}^1(0, T; V^*)$, respectively, follows from the superlinear growth of Ψ_u and Ψ_u^* (Assumption (2.Ψb)), inequality (2.30), and the growth condition (2.Bb). To clarify this, let $\varepsilon > 0$ and $\widetilde{M} := \max\{\beta(1+M), M\}$ be given, where M is the constant obtained from the boundedness (2.29) and (2.30). Then, by Assumption (2.Ψb) for M and $\widetilde{M}/\varepsilon$ there exist positive numbers K_1, K_2, such that

$$\Psi_u(v) \geq \frac{\widetilde{M}}{\varepsilon}\|v\| \quad \text{and} \quad \Psi_u^*(\eta) \geq \frac{\widetilde{M}}{\varepsilon}\|\eta\|_* \tag{2.40}$$

for all $v \in V$ with $\|v\| \geq K_1$, all $\eta \in V^*$ with $\|\eta\|_* \geq K_2$ and all $u \in D$ with $G(u) \leq M$. For notational convenience, we define $f_\tau : [0, T] \to V$, $g_\tau : [0, T] \to V^*$ and $h_\tau : [0, T] \to V^*$ by $f_\tau(t) := \widehat{U}_\tau'(t)$, $g_\tau(t) := B(\underline{\mathbf{t}}_\tau(t), \underline{U}_\tau(t))$ and $h_\tau(t) := (B(\underline{\mathbf{t}}_\tau(t), \underline{U}_\tau(t)) - \widetilde{\xi}_\tau(t))$ for all $t \in [0, T]$. Then, by (2.40), (2.29) and (2.30) it holds

$$\int_{\{t\in[0,T]: f_\tau(t)\geq K_1\}} \|f_\tau(t)\|\,\mathrm{d}t \leq \frac{\varepsilon}{\widetilde{M}} \int_{\{t\in[0,T]: f_\tau(t)\geq K_1\}} \Psi_{\underline{U}_\tau(t)}(f_\tau(t))\,\mathrm{d}t \leq \varepsilon$$

$$\int_{\{t\in[0,T]: g_\tau(t)\geq K_2\}} \|g_\tau(t)\|_*\,\mathrm{d}t \leq \frac{\varepsilon}{\widetilde{M}} \int_{\{t\in[0,T]: g_\tau(t)\geq K_2\}} \Psi^*_{\underline{U}_\tau(t)}(g_\tau(t))\,\mathrm{d}t \leq \varepsilon$$

$$\int_{\{t\in[0,T]: h_\tau(t)\geq K_2\}} \|h_\tau(t)\|_*\,\mathrm{d}t \leq \frac{\varepsilon}{\widetilde{M}} \int_{\{t\in[0,T]: h_\tau(t)\geq K_2\}} \Psi^*_{\underline{U}_\tau(t)}(h_\tau(t))\,\mathrm{d}t \leq \varepsilon$$

for all $0 < \tau \leq \tau^*$, which yields the uniform integrability. Since the sum of two uniformly integrable functions is again uniformly integrable, it follows that $(\widetilde{\xi}_\tau)_{0<\tau\leq\tau^*}$ is also uniformly integrable in $\mathrm{L}^1(0, T; V^*)$ with respect to $\tau > 0$. For the last assertion, we first notice that inequality (2.32), considering (2.29) and (2.30), implies

$$\sup_{t\in[0,T]} (t - \underline{\mathbf{t}}_\tau(t))\Psi_{\underline{U}_\tau(t)}\left(\frac{\widetilde{U}_\tau(t) - \underline{U}_\tau(t)}{t - \underline{\mathbf{t}}_\tau(t)}\right) \leq C_2$$

for a constant $C_2 > 0$. Then, again Assumption (2.Ψb) implies that for every $R > 0$ and $\gamma > 0$ there exists $K > 0$ such that

$$\gamma \|\widetilde{U}_\tau(t) - \underline{U}_\tau(t)\| \le (t - \underline{\mathbf{t}}_\tau(t))\Psi_{\underline{U}_\tau(t)}\left(\frac{\widetilde{U}_\tau(t) - \underline{U}_\tau(t)}{t - \underline{\mathbf{t}}_\tau(t)}\right) + (t - \underline{\mathbf{t}}_\tau(t))\gamma K$$
$$\le M + \tau\gamma K \quad \text{for all } t \in [0, T] \text{ and all } 0 < \tau < \tau^*. \tag{2.41}$$

Taking the supremum of the left-hand side over all $t \in [0, T]$ and then taking the limes superior as $\tau \to 0$, we obtain

$$\gamma \limsup_{\tau \to 0} \sup_{t \in [0,T]} \|\widetilde{U}_\tau(t) - \underline{U}_\tau(t)\| \le M \tag{2.42}$$

for all $\gamma > 0$, which implies necessarily $\lim_{\tau\to 0} \sup_{t\in[0,T]} \|\widetilde{U}_\tau(t) - \underline{U}_\tau(t)\| = 0$. Since inequality (2.42) holds for every $t \in [0, T]$, it is particularly satisfied for $t = t_n$, $n = 1, \ldots, N$, so that we also obtain $\lim_{\tau\to 0} \sup_{t\in[0,T]} \|\overline{U}_\tau(t) - \underline{U}_\tau(t)\| = 0$. The latter convergence in turn implies $\lim_{\tau\to 0} \sup_{t\in[0,T]} \|\widehat{U}_\tau(t) - \overline{U}_\tau(t)\| = 0$, which completes the proof. □

2.7. Passage to the limit and completion of the proof

The next step in constructing a solution to our CAUCHY problem relies on compactness arguments in order to show the existence of a limit function, which obeys the differential inclusion (1.1) and satisfies the initial condition. For this, it is natural to make use of the fact that the interpolation functions are contained in a sublevel set of the energy functional, which by assumption is compact. We elaborate on this in the following result, which also provides the characterization of the limit function by YOUNG measures.

LEMMA 2.9. *Under the same assumptions of Lemma 2.7, let $u_0 \in D$ and $(\tau_n)_{n\in\mathbb{N}}$ be a vanishing sequence of positive real numbers. Then, there exists a subsequence $(\tau_{n_k})_{k\in\mathbb{N}}$, an absolutely continuous curve $u \in \mathrm{AC}([0, T]; V)$ with $u(0) = u_0$, an integrable function $\widetilde{\xi} \in \mathrm{L}^1(0, T; V^*)$, a function $\mathscr{E} : [0, T] \to \mathbb{R}$ of bounded variation, an essentially bounded function $\mathscr{P} \in \mathrm{L}^\infty(0, T)$, and a time-depended YOUNG measure $\boldsymbol{\mu} = (\mu_t)_{t\in[0,T]} \in \mathscr{Y}(0, T; V \times V^* \times \mathbb{R})$, such that*

$$\overline{U}_{\tau_{n_k}}, \underline{U}_{\tau_{n_k}}, \widetilde{U}_{\tau_{n_k}}, \widehat{U}_{\tau_{n_k}} \to u \quad \text{in } \mathrm{L}^\infty(0, T; V), \tag{2.43a}$$
$$\widehat{U}'_{\tau_{n_k}} \rightharpoonup u' \quad \text{in } \mathrm{L}^1(0, T; V), \tag{2.43b}$$
$$\widetilde{\xi}_{\tau_{n_k}} \rightharpoonup \widetilde{\xi} \quad \text{in } \mathrm{L}^1(0, T; V^*), \tag{2.43c}$$
$$B(\underline{\mathbf{t}}_{\tau_{n_k}}, \underline{U}_{\tau_{n_k}}) \to B(\cdot, u(\cdot)) \quad \text{in } \mathrm{L}^\infty(0, T; V^*), \tag{2.43d}$$
$$\partial_t \mathcal{E}_t(\widetilde{U}_{\tau_{n_k}}(t)) \rightharpoonup^* \mathscr{P} \quad \text{in } \mathrm{L}^\infty(0, T), \tag{2.43e}$$

and

$$\begin{cases} \mathcal{E}_t(\overline{U}_{\tau_{n_k}}(t)) \to \mathscr{E}(t) & \text{for all } t \in [0, T], \quad \mathcal{E}_0(u_0) = \mathscr{E}(0), \\ \mathcal{E}_t(u(t)) \le \mathscr{E}(t) & \text{for all } t \in [0, T] \\ \mathcal{E}_t(u(t)) = \mathscr{E}(t) & \text{for a.a. } t \in (0, T), \end{cases} \tag{2.44}$$

as $k \to \infty$. *Furthermore, it holds*

$$u'(t) = \int_{V\times V^*\times\mathbb{R}} v \,\mathrm{d}\mu_t(v,\zeta,p) \quad \text{for a.a. } t\in[0,T], \tag{2.45a}$$

$$\widetilde{\xi}(t) = \int_{V\times V^*\times\mathbb{R}} \zeta \,\mathrm{d}\mu_t(v,\zeta,p) \quad \text{for a.a. } t\in[0,T], \tag{2.45b}$$

$$\mathscr{P}(t) = \int_{V\times V^*\times\mathbb{R}} p \,\mathrm{d}\mu_t(v,\zeta,p) \le \partial_t\mathcal{E}_t(u(t)) \quad \text{for a.a. } t\in[0,T]. \tag{2.45c}$$

and the following energy inequality

$$\begin{aligned}
&\int_s^t \Big(\Psi_{u(r)}(u'(r)) + \Psi^*_{u(r)}(B(r,u(r)) - \widetilde{\xi}(r))\Big)\,\mathrm{d}r + \mathscr{E}(t)\\
&\le \int_s^t \int_{V\times V^*\times\mathbb{R}} \Big(\Psi_{u(r)}(v) + \Psi^*_{u(r)}(B(r,u(r)) - \zeta)\Big)\,\mathrm{d}\mu_r(v,\zeta,p)\,\mathrm{d}r + \mathscr{E}(t)\\
&\le \mathscr{E}(s) + \int_s^t \mathscr{P}(r)\,\mathrm{d}r + \int_s^t \langle B(r,u(r)), u'(r)\rangle\,\mathrm{d}r\\
&\le \mathscr{E}(s) + \int_s^t \partial_r\mathcal{E}_r(u(r))\,\mathrm{d}r + \int_s^t \langle B(r,u(r)), u'(r)\rangle\,\mathrm{d}r
\end{aligned} \tag{2.46}$$

for all s *and* t *with* $0 \le s < t \le T$.

Proof. Let the initial datum $u_0 \in D$ and the sequence $(\tau_n)_{n\in\mathbb{N}}$ of vanishing time steps be given, such that $\tau_n < \tau^*$ for all $n \in \mathbb{N}$. In order to show the existence of an absolutely continuous function, we employ the Arzelà- Ascoli theorem on the family of continuous functions $(\widehat{U}_{\tau_n})_{n\in\mathbb{N}} \subset C([0,T];V)$. First, we notice that the uniform integrability of $(\widehat{U}'_{\tau_n})_{n\in\mathbb{N}}$ leads to the equicontinuity of $(\widehat{U}_{\tau_n})_{n\in\mathbb{N}}$. Second, the fact that the set $\{\overline{U}_{\tau_n}(t)\}_{t\in[0,T]}$ belongs to a sublevel set of the energy functional $\mathcal{E}_t$ for all $n\in\mathbb{N}$, which by Assumption (2.Eb) are compact, implies by Mazur's lemma that the set $\{\overline{U}_{\tau_n}(t)\}_{t\in[0,T]}$ also belongs for all $n\in\mathbb{N}$ to a compact subset of V. Therefore, by Arzelà- Ascoli, there exists a continuous function $u \in C([0,T];V)$ such that $\|\widehat{U}_{\tau_n} - u\|_{C([0,T];V)} \to 0$ as $k\to\infty$ so that in particular $u(0) = u_0$. Then, the convergences in (2.43a) follow from those in (2.31).

Further, from the Dunford- Pettis theorem, see, e.g., Dunford and Schwartz [15, Cor. 11, p. 294], which can be applied since both V and V^* are reflexive Banach spaces, we obtain with the uniform integrability of $(\widehat{U}'_{\tau_n})_{n\in\mathbb{N}}$ and $(\widetilde{\xi}_{\tau_n})_{n\in\mathbb{N}}$ in $\mathrm{L}^1(0,T;V)$ and $\mathrm{L}^1(0,T;V^*)$, respectively, the existence of a subsequence (labeled as before) and weak limits $v \in \mathrm{L}^1(0,T;V)$ and $\widetilde{\xi} \in \mathrm{L}^1(0,T;V^*)$ such that $\widehat{U}'_{\tau_n} \rightharpoonup v$ weakly in $\mathrm{L}^1(0,T;V)$ and $\widetilde{\xi}_{\tau_n} \rightharpoonup \widetilde{\xi}$ weakly in $\mathrm{L}^1(0,T,V^*)$ as $n\to\infty$. From a well-known argument, one can identify v as weak derivative of u, i.e., $u' = v$ in the weak sense which yields $u \in \mathrm{W}^{1,1}(0,T;V)$ and due to continuity of u, $u \in \mathrm{AC}([0,T];V)$.

Now, we shall prove the convergence (2.43d) of the perturbation. We first note that the functions $t \mapsto B(t,u(t))$ and $t \mapsto B(\underline{\mathbf{t}}_{\tau_{n_k}}(t), \underline{U}_{\tau_{n_k}}(t))$ both belong to the space

$L^\infty(0,T;V)$, where the measurability follows from the continuity of u and B, and Assumptions (2.Ba) and (2.49), respectively, whereas the (essential) boundedness is a consequence of Assumptions (2.Bb) and (2.Ψb) as well as the a priori estimates. Now, since the interpolation functions are contained in a sublevel set of the energy functional, uniformly in $\tau > 0$ and for all $t \in (0,T)$, they also take values in a compact subset of V, uniformly in $\tau > 0$ and for all $t \in (0,T)$. Therefore, there exists a compact set $\mathcal{K} \subset V$ such that by TYCHONOV'S theorem the set $[0,T] \times \mathcal{K}$ is compact with respect to the product topology of $[0,T] \times V$. This in turn implies with Assumption (2.Ba) the uniform continuity of the map $(t,u) \mapsto B(t,u)$ on $[0,T] \times \mathcal{K}$. Together with the convergence of $(\underline{\mathbf{t}}_{\tau_{n_k}}(t), \underline{U}_{\tau_{n_k}}(t))) \to (t, u(t))$ uniformly in $t \in (0,T)$, we obtain

$$\lim_{n\to\infty} \sup_{t\in(0,T)} \|B(\underline{\mathbf{t}}_{\tau_{n_k}}(t), \underline{U}_{\tau_{n_k}}(t)) - B(t,u(t))\|_* \quad \text{as } n \to \infty. \tag{2.47}$$

In order to show the convergence in (2.43e), we notice that due to (2.29) it holds $(\partial_t \mathcal{E}_t(\widetilde{U}_{\tau_{n_k}}))_{k\in\mathbb{N}} \subset L^\infty(0,T)$. Since the LEBESGUE space $L^\infty(0,T)$ is the dual space of the separable BANACH space $L^1(0,T)$, there exists a limit $\mathscr{P} \in L^\infty(0,T)$ such that (up to a subsequence) $\partial_t \mathcal{E}_t(\widetilde{U}_{\tau_{n_k}}) \rightharpoonup^* \mathscr{P}$ weakly* in $L^\infty(0,T)$ as $k \to \infty$.

Now, we shall prove (2.44). For this, we define

$$\eta_\tau(t) := \mathcal{E}_{\bar{\mathbf{t}}_\tau(t)}(\overline{U}_\tau(t)) - \int_0^{\bar{\mathbf{t}}_\tau(t)} \partial_r \mathcal{E}_r(\widetilde{U}_\tau(r))\,\mathrm{d}r - \int_0^{\bar{\mathbf{t}}_\tau(t)} \langle B(\underline{\mathbf{t}}_\tau(r), \underline{U}_\tau(r)), \widehat{U}_\tau'(r)\rangle\,\mathrm{d}r$$

for $t \in [0,T]$ and we deduce from the discrete upper energy estimate (2.28) that the map $t \mapsto \eta_\tau(t) : [0,T] \to \mathbb{R}$ is non-increasing. Then, by HELLY's theorem there exists a non-increasing function $\eta : [0,T] \to \mathbb{R}$ and a subsequence (labeled as before) such that $\eta_{\tau_{n_k}}(t) \to \eta(t)$ as $k \to \infty$ for all $t \in [0,T]$. Moreover, we define

$$\psi_\tau(t) := \int_0^{\bar{\mathbf{t}}_\tau(t)} \langle B(\underline{\mathbf{t}}_\tau(r), \underline{U}_\tau(r)), \widehat{U}_\tau'(r)\rangle\,\mathrm{d}r \quad \text{for } t \in [0,T].$$

Since we have strong convergence of the perturbation $B(\underline{\mathbf{t}}_\tau, \underline{U}_{\tau_{n_k}})$ in $L^\infty(0,T;V^*)$ and weak convergence of the derivative $\widehat{U}_{\tau_{n_k}}'$ in $L^1(0,T;V)$ as $k \to \infty$, it holds

$$\psi_{\tau_{n_k}}(t) \to \psi(t) := \int_0^t \langle B(r,u(r)), u'(r)\rangle\,\mathrm{d}r \quad \text{as } k \to \infty \tag{2.48}$$

for all $t \in [0,T]$. Considering convergence (2.43e), we obtain

$$\mathcal{E}_{\bar{\mathbf{t}}_{\tau_{n_k}}(t)}(\overline{U}_{\tau_{n_k}}(t)) \to \mathscr{E}(t) := \eta(t) + \int_0^t \mathscr{P}(r)\,\mathrm{d}r + \psi(t) \quad \text{for all } t \in [0,T]$$

as $k \to \infty$. Since the function η is monotone and both the function ψ and the map $t \mapsto \int_0^t \mathscr{P}(r)\,\mathrm{d}r$ are absolutely continuous, it follows that the function $\mathscr{E}$ is of bounded variation. In order to conclude the convergence in (2.44), we notice that

$$\left|\mathcal{E}_{\bar{\mathbf{t}}_{\tau_{n_k}}(t)}(\overline{U}_{\tau_{n_k}}(t)) - \mathcal{E}_t(\overline{U}_{\tau_{n_k}}(t))\right| \to 0 \quad \text{as } k \to \infty,$$

which follows from (2.6), (2.29), and the fact that $\overline{\mathbf{t}}_{\tau_{n_k}}(t) \to t$ as $k \to \infty$ for all $t \in [0, T]$. Further, by the lower semicontinuity of the energy functional, we obtain due to the convergence (2.31)

$$\mathcal{E}_t(u(t)) \le \liminf \mathcal{E}_t(\overline{U}_{\tau_{n_k}}(t)) = \mathscr{E}(t) \le M \quad \text{for all } t \in [0, T], \tag{2.49}$$

where the last inequality follows from (2.29). The last assertion in (2.44) follows from Assumption (2.Ee).

We continue by showing (2.45). For this purpose, we define the (reflexive) BANACH space $\mathcal{V} := V \times V^* \times \mathbb{R}$ endowed with the product topology and employ the fundamental theorem of weak topologies (Theorem A.2) applied to the sequence $w_k := (\widehat{U}'_{\tau_{n_k}}, \widetilde{\xi}_{\tau_{n_k}}, \partial_t \mathcal{E}_t(\widetilde{U}_{\tau_{n_k}}))_{k\in\mathbb{N}}$ which belongs to $\mathrm{L}^1(0, T; \mathcal{V})$ by the a priori estimates, and is uniformly integrable in $\mathrm{L}^1(0, T; \mathcal{V})$ since every component is in the respective space. Thus, there exists a YOUNG measure $\boldsymbol{\mu} = (\mu_t)_{t\in[0,T]} \in \mathscr{Y}(0, T; V \times V^* \times \mathbb{R})$ such that μ_t is for almost all $t \in (0, T)$ concentrated on the set

$$\mathrm{Li}(t) := \bigcap_{p=1}^{\infty} \mathrm{clos}_{\mathrm{weak}}\big(\{w_k(t) : k \ge p\}\big)$$

of all limit points of $w_k(t)$ with respect to the weak-weak-strong topology of $V \times V^* \times \mathbb{R}$, i.e., $\mathrm{sppt}(\mu_t) \subset \mathrm{Li}(t)$. Since the weak limits in (2.43b), (2.43c) and (2.43e) are unique, the identities in (2.45a) and (2.45b) are direct consequences of the fundamental theorem of weak topologies, whereas the inequality in (2.45c) is true due to the fact that for a.a. $t \in (0, T)$, it holds

$$\zeta \in \partial \mathcal{E}_t(u(t)) \quad \text{and} \quad p \le \partial_t \mathcal{E}_t(u(t)) \quad \text{for all } (v, \zeta, p) \in \mathrm{Li}(t). \tag{2.50}$$

Property (2.50) in turn follows from Assumption (2.Ee) with the convergences in (2.43) and (2.44) as well as the inclusion (2.27): Let $\mathcal{N} \subset (0, T)$ be a negligible set such that for all $t \in (0, T)\backslash\mathcal{N}$ the set $\mathrm{Li}(t)$ is non-empty. Now let be $t \in (0, T)\backslash\mathcal{N}$ and $(v, \zeta, p) \in \mathrm{Li}(t)$, then there exists a subsequence $(k_l)_{l\in\mathbb{N}}$ such that $\widehat{U}'_{\tau_{n_{k_l}}}(t) \rightharpoonup v$, $\widetilde{\xi}_{\tau_{n_{k_l}}}(t) \rightharpoonup^* \zeta$ and $\partial_t \mathcal{E}_t(\widetilde{U}_{\tau_{n_{k_l}}}(t)) \to p$ as $l \to \infty$, where the latter convergence follows from the fact that in finite-dimensional spaces the weak topology coincides with the strong topology. In view of convergence (2.43a) and the inclusion (2.27), (2.50) follows by Assumption (2.Ee). Integrating the inequality in (2.50) with respect to the BOREL probability measure yields (2.45c). In order to show the energy inequality (2.46), we notice first of all that from JENSEN's inequality, we obtain for a.a. $t \in (0, T)$

$$\Psi_{u(t)}(u'(t)) \le \int_{V\times V^*\times\mathbb{R}} \Psi_{u(t)}(v)\,\mathrm{d}\mu_t(v, \zeta, p), \tag{2.51}$$

$$\Psi^*_{u(t)}(B(t, u(t)) - \widetilde{\xi}(t))) \le \int_{V\times V^*\times\mathbb{R}} \Psi^*_{u(t)}(B(t, u(t)) - \zeta)\,\mathrm{d}\mu_t(v, \zeta, p). \tag{2.52}$$

This can also be obtained by integrating the inequalities

$$\Psi_{u(t)}(u'(t)) \le \Psi_{u(t)}(v) + \langle w^*, u'(t) - v\rangle \quad \text{for all } v \in V$$

$$\Psi^*_{u(t)}(B(t,u(t)) - \widetilde{\xi}(t)) \leq \Psi^*_{u(t)}(B(t,u(t)) - \zeta) + \langle \zeta - \widetilde{\xi}(t), w\rangle \quad \text{for all } \zeta \in V^*$$

using the identities in (2.45) as well as the fact that $w^* \in \partial\Psi_{u(t)}(u'(t)) \neq \emptyset$ and $w \in \partial\Psi^*_{u(t)}(B(t,u(t)) - \widetilde{\xi}(t)) \neq \emptyset$, see Remark 2.3 i).

Defining $\mathcal{H}_k : [0,T] \times \mathcal{V} \to \mathbb{R}$ by

$$\mathcal{H}_k(r,w) := \chi_{[\bar{\mathbf{t}}_{\tau_{n_k}}(s),\bar{\mathbf{t}}_{\tau_{n_k}}(t)]}\Psi_{\underline{U}_{\tau_{n_k}}(r)}(v), \quad (r,v,\zeta,p) \in [0,T] \times \mathcal{V},$$

together with (2.29) and (2.43a), the MOSCO continuity (2.Ψc) leads to

$$\mathcal{H}(r,w) := \chi_{[s,t]}\Psi_{u(r)}(v) \leq \liminf_{k\to\infty} \mathcal{H}_k(r,w_n) \tag{2.53}$$

for all $(r,w) = (r,v,\zeta,p) \in [0,T] \times \mathcal{V}$ and all weakly convergent sequences $w_k \rightharpoonup w \in \mathcal{V}$, where $s,t \in [0,T]$ with $s \leq t$ are chosen to be fixed. As the BANACH space $\mathcal{V}$ is reflexive, the map

$$(v,\zeta,p) \mapsto (\|v\| + \|\zeta\|_* + |p|)$$

has compact sublevel sets with respect to the weak topology of $\mathcal{V}$. Together with the boundedness of the afore-defined sequence $(w_k)_{k\in\mathbb{N}}$, which follows from (2.43), we obtain the weak-tightness of $(w_k)_{k\in\mathbb{N}}$. Therefore, for a subsequence of $(n_k)_{k\in\mathbb{N}}$ (not relabeled), Theorem A.1 provides the inequality

$$\int_0^T \int_{\mathcal{V}} \mathcal{H}(r,w)\,\mathrm{d}\mu_r(w)\,\mathrm{d}r \leq \liminf_{k\to\infty} \int_0^T \mathcal{H}_k(r,w_k)\,\mathrm{d}r,$$

which means

$$\int_s^t \int_{\mathcal{V}} \Psi_{u(r)}(v)\,\mathrm{d}\mu(v,\zeta,p)\,\mathrm{d}r \leq \liminf_{k\to\infty} \int_{\bar{\mathbf{t}}_{\tau_{n_k}}(s)}^{\bar{\mathbf{t}}_{\tau_{n_k}}(t)} \Psi_{\underline{U}_{\tau_{n_k}}(r)}(\widehat{U}'_{\tau_{n_k}}(r))\,\mathrm{d}r < +\infty, \tag{2.54}$$

where the boundedness follows from the a priori estimate (2.30). Taking into account Remark 2.3 (iii), Theorem A.1 applied to the function

$$\mathcal{H}^*_k(r,w) := \chi_{[\bar{\mathbf{t}}_{\tau_{n_k}}(s),\bar{\mathbf{t}}_{\tau_{n_k}}(t)]}\Psi^*_{\underline{U}_{\tau_{n_k}}(r)}(B(\underline{\mathbf{t}}_{\tau_{n_k}}(r),\underline{U}_{\tau_{n_k}}(r)) - \zeta),$$
$$(r,v,\zeta,p) \in [0,T] \times \mathcal{V},$$

yields

$$\begin{aligned}&\int_s^t \int_{\mathcal{V}} \Psi^*_{u(t)}(B(r,u(r)) - \zeta)\,\mathrm{d}\mu(v,\zeta,p)\,\mathrm{d}r \\ &\leq \liminf_{k\to\infty} \int_{\bar{\mathbf{t}}_{\tau_{n_k}}(s)}^{\bar{\mathbf{t}}_{\tau_{n_k}}(t)} \Psi^*_{\underline{U}_{\tau_{n_k}}(r)}(B(\underline{\mathbf{t}}_{\tau_{n_k}}(r),\underline{U}_{\tau_{n_k}}(r)) - \widetilde{\xi}_{\tau_{n_k}}(r))\,\mathrm{d}r < +\infty,\end{aligned} \tag{2.55}$$

where again the boundedness follows from (2.30). Integrating (2.51) and (2.52) with respect to t yields the first inequality in (2.46). The second and third inequality follow by passing to the limit in the discrete upper energy estimate (2.28) as $k \to \infty$ and considering (2.43e), (2.44), (2.45c), (2.48), (2.50) as well as (2.54) and (2.55). This proves Lemma 2.9. □

We are now ready to complete the proof of our main existence Theorem 2.5.

Proof. (Proof of Theorem 2.5) In order to show that the curve $u \in \mathrm{AC}([0,T];V)$ obtained from Lemma 2.9 is a solution to the differential inclusion (1.1), we make use of the chain rule for YOUNG measures in Lemma A.3, which is justified by (2.43e), (2.45a), (2.50), (2.54) and (2.55), where $\boldsymbol{\mu} = (\mu_t)_{t\in[0,T]} \in \mathscr{Y}(0,T; V\times V^*\times\mathbb{R})$ is to be chosen as in Lemma 2.9. Hence, by the chain rule condition, the map $t \mapsto \mathcal{E}_t(u(t))$ is absolutely continuous on $(0,T)$ and there holds

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}_t(u(t)) \geq \int_{V\times V^*\times\mathbb{R}} \langle \zeta, u'(t)\rangle \,\mathrm{d}\mu_t(v,\zeta,p) + \partial_t\mathcal{E}_t(u(t)) \quad \text{for a.a. } t\in(0,T).$$

Thus, together with (2.44), (2.45c) and (2.46), we obtain with $s=0$

$$\begin{aligned}
&\int_0^t \int_{V\times V^*\times\mathbb{R}} \Big(\Psi_{u(r)}(u'(r)) + \Psi^*_{u(r)}(B(r,u(r))-\zeta)\Big)\,\mathrm{d}\mu_r(v,\zeta,p)\,\mathrm{d}r + \mathcal{E}_t(u(t))\\
&\quad\leq \mathcal{E}_0(u_0) + \int_0^t \partial_r\mathcal{E}_r(u(r))\,\mathrm{d}r + \int_0^t \langle B(r,u(r)), u'(r)\rangle\,\mathrm{d}r\\
&\quad\leq \mathcal{E}_t(u(t)) - \int_0^t\int_{V\times V^*\times\mathbb{R}} \langle\zeta, u'(r)\rangle\,\mathrm{d}\mu_r(v,\zeta,p)\,\mathrm{d}r + \int_0^t \langle B(r,u(r)), u'(r)\rangle\,\mathrm{d}r\\
&\quad= \mathcal{E}_t(u(t)) + \int_0^t\int_{V\times V^*\times\mathbb{R}} \big\langle B(r,u(r))-\zeta, u'(r)\big\rangle\,\mathrm{d}\mu_r(v,\zeta,p)\,\mathrm{d}r
\end{aligned} \tag{2.56}$$

for all $t\in[0,T]$. Therefore, it holds

$$\begin{aligned}
&\int_0^t \int_{V\times V^*\times\mathbb{R}} \Big(\Psi_{u(r)}(u'(r)) + \Psi^*_{u(r)}(B(r,u(r))-\zeta)\\
&\quad - \big\langle B(r,u(r))-\zeta, u'(r)\big\rangle\Big)\,\mathrm{d}\mu_r(v,\zeta,p)\,\mathrm{d}r \leq 0 \quad \text{for all } t\in[0,T].
\end{aligned} \tag{2.57}$$

Then, from the FENCHEL- YOUNG inequality we deduce the non-negativity of the integrand in (2.57) and infer therefore

$$\begin{aligned}
&\int_{V\times V^*\times\mathbb{R}} \Big(\Psi_{u(t)}(u'(t)) + \Psi^*_{u(t)}\big(B(t,u(t))-\zeta\big) - \big\langle B(t,u(t))-\zeta, u'(t)\big\rangle\Big)\,\mathrm{d}\mu_t(v,\zeta,p)\\
&\quad = 0 \quad \text{for a.a. } t\in(0,T).
\end{aligned} \tag{2.58}$$

It follows that all inequalities in (2.56) become equalities for all $t\in[0,T]$, so that we obtain the equation

$$\begin{aligned}
&\int_s^t \int_{V\times V^*\times\mathbb{R}} \Big(\Psi_{u(r)}(u'(r)) + \Psi^*_{u(r)}\big(B(r,u(r))-\zeta\big)\Big)\,\mathrm{d}\mu_r(v,\zeta,p)\,\mathrm{d}r + \mathcal{E}_t(u(t))\\
&\quad = \mathcal{E}_s(u(s)) + \int_s^t \partial_r\mathcal{E}_r(u(r))\,\mathrm{d}r + \int_s^t \langle B(r,u(r)), u'(r)\rangle\,\mathrm{d}r
\end{aligned} \tag{2.59}$$

for all $0 \le s, t \le T$. Defining the marginal $\boldsymbol{\nu} = (\nu_t)_{t\in[0,T]} := \pi^{2,3}_{\#}\boldsymbol{\mu}$ of $\boldsymbol{\mu}$ by $\nu_t(B) := \mu_t((\pi^{2,3})^{-1}(B))$ for all $B \in \mathscr{B}(V^* \times \mathbb{R})$, where $\pi^{2,3} : V \times V^* \times \mathbb{R} \to V^* \times \mathbb{R}$ denotes the canonical projection and $\mathscr{B}(V^* \times \mathbb{R})$ the BOREL σ-algebra of $V^* \times \mathbb{R}$. Setting

$$\mathcal{S}(t, u(t), u'(t)) := \Big\{(\zeta, p) \in V^*\times\mathbb{R} \,\Big|\, \zeta \in \partial\mathcal{E}_t(u(t)) \cap \big(B(t, u(t)) - \partial\Psi_{u(t)}(u'(t))\big) \text{ and } p \le \partial_t\mathcal{E}_t(u(t))\Big\} \tag{2.60}$$

we notice that by (2.50) and (2.58) it follows that $\nu_t(\mathcal{S}(t, u(t), u'(t))) = 1$ for a.a. $t \in (0, T)$ and assumption (A.7) is fulfilled. Therefore, by Lemma A.4 there exists a measurable selection $\xi : [0, T] \to V^*$ and $p : [0, T] \to \mathbb{R}$ with

$$\int_0^T \Psi^*_{u(t)}(B(t, u(t)) - \xi(t))\,\mathrm{d}t < +\infty, \tag{2.61}$$

such that $(\xi(t), p(t)) \in \mathcal{S}(t, u(t), u'(t))$ and it holds

$$\Psi^*_{u(t)}(B(t, u(t)) - \xi(t)) - p(t) = \min_{(\hat{\zeta},\hat{p})\in\mathcal{S}(t,u(t),u'(t))} \Psi^*_{u(t)}(B(t, u(t)) - \hat{\zeta}) - \hat{p}. \tag{2.62}$$

Since (2.61) holds and $B(\cdot, u(\cdot)) \in \mathrm{L}^\infty(0, T; V^*)$, we deduce from the superlinearity of Ψ^*_u that $\xi \in \mathrm{L}^1(0, T; V^*)$, so that the pair (u, ξ) solves the differential inclusion (1.1) and u satisfies the initial condition $u(0) = u_0$, where the former follows from (2.62) and the latter by Lemma 2.9.

Furthermore, taking into account property (2.50) and equation (2.58), Lemma 2.1 yields $\nu_t(\mathcal{S}(t, u(t), u'(t)) = 1$ for a.a. $t \in (0, T)$. Thus, from equality (2.62) and the definition of $\mathcal{S}(\cdot, u(\cdot).u'(\cdot))$, it follows that

$$\int_s^t \Psi^*_{u(r)}(B(r, u(r) - \xi(r))\,\mathrm{d}r - \int_s^t p(r)\,\mathrm{d}r \le \int_s^t \int_{V\times V^*\times\mathbb{R}} \Psi^*_{u(r)}(B(r, u(r)) - \zeta)\,\mathrm{d}\mu_r(v, \zeta, p)\,\mathrm{d}r - \int_s^t p(r)\,\mathrm{d}r.$$

Now, by comparison with equation (2.59), we infer

$$\int_s^t \Big(\Psi_{u(r)}(u'(r)) + \Psi^*_{u(r)}(B(r, u(r)) - \xi(r))\Big)\,\mathrm{d}r + \mathcal{E}_t(u(t)) \le \mathcal{E}_s(u(s)) + \int_s^t \partial_r\mathcal{E}_r(u(r))\,\mathrm{d}r + \int_s^t \langle B(r, u(r)), u'(r)\rangle\,\mathrm{d}r$$

for all $0 \le s \le t \le T$. On the other hand, applying the chain rule condition (2.Ed) to the pair (u, ξ) yields

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}_t(u(t)) \ge \langle \xi(t), u'(t)\rangle + \partial_t\mathcal{E}_t(u(t)) \quad \text{for a.e. } t \in (0, T).$$

Together with the identity

$$\Psi_{u(r)}(u'(r)) + \Psi^*_{u(r)}(B(r,u(r)) - \xi(r)) = \langle B(r,u(r)) - \xi(r), u'(r)\rangle \quad \text{a.e. in}(0,T),$$

which again follows from Lemma 2.1, and the definition of $\mathcal{S}(\cdot, u(\cdot).u'(\cdot))$, we conclude the energy-dissipation balance (2.11). □

REMARK 2.10. It is not difficult to prove that for every sequence $(\tau_n)_{n\in\mathbb{N}}$ there exists a subsequence (not relabeled) such that the following convergences hold:

$$\begin{aligned}
\mathcal{E}_t(\overline{U}_{\tau_n}(t)) &\to \mathcal{E}_t(u(t)) \quad \text{for all } t\in[0,T],\\
\int_s^t \Psi_{\underline{U}_{\tau_n}(r)}(\widehat{U}'_{\tau_n}(r))\,\mathrm{d}r &\to \int_s^t \Psi_{u(r)}(u'(r))\,\mathrm{d}r \quad \text{and}\\
\int_s^t \Psi^*_{\underline{U}_{\tau_n}(r)}(B(\underline{\mathbf{t}}_{\tau_n}(r), \underline{U}_{\tau_n}(r)) - \widetilde{\xi}_{\tau_n}(r))\,\mathrm{d}r &\to \int_s^t \Psi^*_{u(r)}(B(r,u(r)) - \xi(r))\,\mathrm{d}r
\end{aligned}$$

for all $0 \le s \le t \le T$ as $n \to \infty$. Furthermore, if we additionally assume that the dissipation potential Ψ_u and its conjugate Ψ^*_u are strictly convex for all $u \in V$, then it holds $\pi^1_\#\boldsymbol{\mu} = \delta_{u'(t)}$ and $\pi^2_\#\boldsymbol{\mu} = \delta_{\xi(t)}$, respectively, and we obtain the pointwise convergences

$$\widehat{U}'_{\tau_n}(t) \rightharpoonup u'(t) \quad \text{and} \quad \widetilde{\xi}_{\tau_n}(t) \rightharpoonup \xi(t) \quad \text{for a.a. } t \in (0,T).$$

as well as $\widetilde{\xi}_{\tau_n} \rightharpoonup \xi$ in $\mathrm{L}^1(0,T;V^*)$ as $n\to\infty$.

3. A result for evolutionary Γ-convergence

In this section, we consider a family of perturbed gradients systems $\mathrm{PG}^\varepsilon := (V, \mathcal{E}^\varepsilon, \Psi^\varepsilon, B^\varepsilon)$, where $\varepsilon\in[0,1]$ is a small parameter. Here the case $\varepsilon = 0$ is the supposed limit equation, also called effective equation. The major question is what type of convergence of $\mathcal{E}^\varepsilon$, Ψ^ε, and B^ε is sufficient to conclude that solutions $u_\varepsilon : [0,T] \to V$ for PG^ε with $\varepsilon > 0$ have subsequences $\varepsilon_k \to 0$ that converge pointwise in $t \in [0,T]$ to a limit function $u_0 : [0,T] \to V$ and that u_0 is indeed a solution for PG^0.

The theory developed here follows [23, Thm. 4.8], where the case of pure gradient systems (i.e., $B_\varepsilon \equiv 0$) was considered.

3.1. Assumptions and results

Our assumptions follow closely the assumptions for the existence theory in Sect. 2.3, where we need uniformity with respect to $\varepsilon\in[0,1]$. For definiteness, we now list the precise assumptions on PG^ε. For describing the energy functionals $\mathcal{E}^\varepsilon$, we define

$$\begin{aligned}
G^\varepsilon(u) &= \sup\left\{ \mathcal{E}^\varepsilon_t(u) \,\middle|\, t\in[0,T] \right\},\\
\underline{G}(u) &:= \inf\left\{ \mathcal{E}^\varepsilon_t(u) \,\middle|\, t\in[0,T],\ \varepsilon\in[0,1] \right\}.
\end{aligned}$$

Without loss of generality, we may assume that $\underline{G}$ is bounded from below by a positive constant $\gamma > 0$.

$$\left.\begin{array}{l}\textbf{Constant domains.}\quad \forall\, t \in [0,T]\ \forall\, \varepsilon \in [0,1]:\\ \mathcal{E}_t^\varepsilon : V \to (0,\infty] \text{ is proper and lower semicontinuous with}\\ \text{time-independent domain } D^\varepsilon := \mathrm{dom}(\mathcal{E}_t^\varepsilon) \subset V \text{ for all } t \in [0,T].\end{array}\right\} \quad (3.\mathrm{E}^\varepsilon\mathrm{a})$$

Equicompactness of sublevels.

The sublevels of $\underline{G}$ have compact closure in V. (3.E$^\varepsilon$b)

$$\left.\begin{array}{l}\textbf{Uniform energetic control of power.}\\ \forall\, \varepsilon \in [0,1]\ \forall\, u \in D^\varepsilon: \quad t \mapsto \mathcal{E}_t^\varepsilon(u) \text{ is differentiable on } (0,T) \text{ and}\\ \exists\, C_T > 0\ \forall\, \varepsilon \in [0,1]\ \forall\, t \in (0,T)\ \forall\, u \in D^\varepsilon: \quad |\partial_t \mathcal{E}_t^\varepsilon(u)| \le C_T \mathcal{E}_t^\varepsilon(u).\end{array}\right\}$$

(3.E$^\varepsilon$c)

Chain rule. $\forall\, \varepsilon \in [0,1]:$ the chain rule of (2.Ed) holds for $(V, \mathcal{E}^\varepsilon, \Psi^\varepsilon)$.

(3.E$^\varepsilon$d)

Liminf estimate. $(\varepsilon_k, u_k) \to (0,u)$ implies $\mathcal{E}_t^0(u) \le \liminf_{k\to\infty} \mathcal{E}_t^{\varepsilon_k}(u_k)$. (3.E$^\varepsilon$e)

$$\left.\begin{array}{l}\textbf{Strong-weak closedness in the limit } \varepsilon \to 0.\quad \text{For all } t \in [0,T] \text{ and}\\ \text{all sequences } (\varepsilon_n, u_n, \xi_n)_{n\in\mathbb{N}} \subset [0,1]\times V \times V^* \text{ with } \xi_n \in \partial\mathcal{E}_t^{\varepsilon_n}(u_n) \text{ and}\\ \varepsilon_n \to 0,\ \ u_n \to u \in V,\ \ \xi_n \rightharpoonup \xi \in V^*,\ \ \mathcal{E}_t^{\varepsilon_n}(u_n) \to \mathcal{E}_0,\ \ \partial_t\mathcal{E}_t^{\varepsilon_n}(u_n) \to \mathcal{P}\\ \text{for } n \to \infty, \text{ we have the relations}\\ \quad \xi \in \partial\mathcal{E}_t^0(u), \quad \mathcal{E}_t^0(u) = \mathcal{E}_0, \quad \text{and} \quad \partial_t\mathcal{E}_t^0(u) \ge \mathcal{P}.\end{array}\right\}$$

(3.E$^\varepsilon$f)

As in the existence theory, we use a control of the time-derivative, see (3.E$^\varepsilon$c), which gives $\mathcal{E}_t^\varepsilon(u) \ge \mathrm{e}^{-C_T|t-s|}\mathcal{E}_s^\varepsilon(u)$. Thus, for all $\varepsilon \in [0,1]$ and $t \in [0,T]$ we have the relations

$$\underline{G}(u) \le G^\varepsilon(u) \le \mathrm{e}^{C_T T}\mathcal{E}_t^\varepsilon(u) \le \mathrm{e}^{C_T T} G^\varepsilon(u).$$

Note that we cannot use a uniform upper bound $G^\varepsilon(u) \le \overline{G}(u)$, as this would exclude many useful results on Γ-convergence.

In the present form of condition (3.E$^\varepsilon$f), we do not ask for the strong-weak closedness for $\mathcal{E}_t^\varepsilon$ with a given positive ε. However, in our main result we simply assume the existence of solutions $u_\varepsilon : [0,T] \to V$ for PG$^\varepsilon$. If we want to show this with the theory of Sect. 2, then one has to impose (2.Ee) for all $\varepsilon > 0$ as well (which is the same as allowing constant sequences $\varepsilon_n = \varepsilon$ in (3.E$^\varepsilon$f)).

The closedness condition (3.E$^\varepsilon$f) looks rather strong, however in Remark 3.2, see after the statement of the main convergence result, we will show that convexity of $\mathcal{E}_t^\varepsilon(\cdot)$ and strong Γ-convergence to $\mathcal{E}_t^0$ already imply the desired closedness.

The conditions of the dissipation potentials $\Psi_u^\varepsilon : V \to [0,\infty)$ are the following.

Dissipation potentials. $\forall\, \varepsilon \in [0,1]\ \forall\, u \in V$:

$$\Psi_u^\varepsilon : V \to [0,\infty) \text{ is lower semicontinuous and convex with } \Psi_u^\varepsilon(0)=0. \tag{3.Ψ^εa}$$

Superlinearity. $\forall\, R>0\, \exists\, g_R : [0,\infty)\to[0,\infty)$ superlinear:
$\forall\, \varepsilon\in[0,1]\, \forall\,(v,\xi)\in V\times V^*\, \forall\, u\in V$ with $\underline{G}(u)<R$:

$$\Psi_u^\varepsilon(v)\ge g_R(\|v\|) \text{ and } \Psi_u^{\varepsilon,*}(\xi)\ge g_R(\|\xi\|). \tag{3.Ψ^εb}$$

Mosco convergence. For all $R>0$ and sequences $(\varepsilon_n,u_n)_{n\in\mathbb{N}}\subset[0,1]\times V$

$$\text{with } \underline{G}(u_n)\le R \text{ and } (\varepsilon_n,v_n)\to(0,v): \quad \Psi_{u_n}^{\varepsilon_n}\xrightarrow{\mathrm{M}}\Psi_u^0. \tag{3.Ψ^εc}$$

Again we have formulated the Mosco convergence of the dissipation potentials only with the limit $\varepsilon_n\to 0$, which is sufficient for the passage to the limit when solutions $u_\varepsilon:[0,T]\to V$ are given. To show the existence of solutions, we need (2.Ψc) for all $\varepsilon\in(0,1]$ as well.

Finally, we impose the conditions on the non-variational perturbation B^ε, namely

$$\textbf{Continuity.} \text{ The map } \begin{cases} [0,1]\times[0,T]\times V \to & V^*, \\ \quad (\varepsilon,t,u) \mapsto & B^\varepsilon(t,u), \end{cases} \text{ is continuous.} \tag{3.B$^\varepsilon$a}$$

Control of B^ε by the energy. $\exists\, C_B>0\,\forall\,(\varepsilon,t)\in[0,1]\times[0,T]$

$$\forall\, u\in D^\varepsilon: \quad \Psi_u^{\varepsilon,*}\big(B^\varepsilon(t,u)\big)\le C_B\mathcal{E}_t^\varepsilon(u). \tag{3.B$^\varepsilon$b}$$

We are now ready to formulate our result of evolutionary Γ-convergence. In [20] the convergence which we will establish is called “pE-convergence” as we have to impose the well-preparedness of the initial conditions u_ε^0, viz.

$$u_\varepsilon^0\to u^0 \text{ in } V \text{ and } \mathcal{E}_0^\varepsilon(u_\varepsilon^0)\to\mathcal{E}_0^0(u^0)<\infty \quad \text{for } \varepsilon\to 0. \tag{3.4}$$

Moreover, in the sense of [14,18] we even have the much stronger notion of EDP convergence, which means convergence in the sense of the energy-dissipation balance. Indeed, as for the existence result in Sect. 2 we will also strongly rely on the energy-dissipation principle and perform the limit $\varepsilon\to 0$ in the energy-dissipation balance (2.11). Our proof will be an adaptation of [23, Thm. 4.8].

THEOREM 3.1. (Evolutionary Γ-convergence) *Assume that the perturbed gradient systems* $\mathrm{PG}^\varepsilon=(V,\mathcal{E}^\varepsilon,\Psi^\varepsilon,B^\varepsilon)$, $\varepsilon\in[0,1]$ *satisfy the assumptions* (3.E$^\varepsilon$), (3.Ψ^ε), *and* (3.B$^\varepsilon$). *Moreover, assume that for* $\varepsilon>0$ *we have solutions* $u_\varepsilon:[0,T]\to V$ *of* PG^ε *such that the initial conditions* $u_\varepsilon(0)=u_\varepsilon^0$ *satisfy* (3.4). *Then, there exists a subsequence* $\varepsilon_k\to 0$ *and a solution* $u:[0,T]\to V$ *of the limit system* PG^0 *with* $u(0)=u^0$ *such that the following convergences hold:*

$$u_{\varepsilon_k}(t) \to u(t) \text{ in } \mathrm{C}^0([0,T];V); \tag{3.5a}$$

$$\forall\, t\in[0,T]: \quad \mathcal{E}_t^{\varepsilon_k}(u_{\varepsilon_k}(t)) \to \mathcal{E}_t^0(u(t)); \tag{3.5b}$$

$$u'_{\varepsilon_k} \rightharpoonup u' \text{ in } \mathrm{L}^1(0,T;V); \tag{3.5c}$$

$$\forall r < s: \int_r^s \Psi^{\varepsilon_k}_{u_{\varepsilon_k}(t)}(u'_{\varepsilon_k}(t))\,\mathrm{d}t \;\to\; \int_r^s \Psi^0_{u(t)}(u'(t))\,\mathrm{d}t; \tag{3.5d}$$

$$\forall r < s: \int_r^s \Psi^{\varepsilon_k,*}_{u_{\varepsilon_k}(t)}\big(B^{\varepsilon_k}(t,u_{\varepsilon_k}(t)){-}\xi_{\varepsilon_k}(t)\big)\,\mathrm{d}t \;\to\; \int_r^s \Psi^{0,*}_{u(t)}\big(B^0(t,u_0(t)){-}\xi_0(t)\big)\,\mathrm{d}t, \tag{3.5e}$$

where $\xi_\varepsilon(t) \in \partial\mathcal{E}^\varepsilon_t(u_\varepsilon(t))$ *for* $\varepsilon \in [0,1]$ *and a.a.* $t \in [0,T]$.

The proof of this result is contained in the following two Sects. 3.3 and 3.4. However, we do not give all the details and refer to the full proof of Theorem 2.5 in Sect. 2 for the details.

REMARK 3.2. (Strong–weak closedness and Γ-convergence) It is a well-know fact that the strong-weak closedness in the limit $\varepsilon \to 0$ as assumed in (3.E$^\varepsilon$f) often follows from the Γ-convergence $\mathcal{E}^\varepsilon_t \xrightarrow{\Gamma} \mathcal{E}^0_t$. For the readers convenience, we give the argument for the convex case where $\partial\mathcal{E}^\varepsilon_t(u)$ is simply the convex subdifferential, i.e.,

$$\partial\mathcal{E}^\varepsilon_t(u) = \big\{\, \xi \in V^* \,\big|\, \forall\, w \in V:\ \mathcal{E}^\varepsilon_t(w) \geq \mathcal{E}^\varepsilon_t(u) + \langle \xi, w{-}v\rangle \,\big\}.$$

Thus, having a sequences $u_\varepsilon \to u$ and $\xi_\varepsilon \rightharpoonup \xi_0$ with $\xi_\varepsilon \in \partial\mathcal{E}^\varepsilon_t(u_\varepsilon)$ for $\varepsilon > 0$ and $\mathcal{E}^\varepsilon_t(u_\varepsilon) \to \overline{e}$, we can find, for each $w \in W$, a recovery sequence $w_\varepsilon \to w$ with $\mathcal{E}^\varepsilon_t(w_\varepsilon) \to \mathcal{E}^0_t(w)$. Hence, we obtain

$$\mathcal{E}^\varepsilon_t(w_\varepsilon) \geq \mathcal{E}^\varepsilon_t(u_\varepsilon) + \langle \xi_\varepsilon, w_\varepsilon{-}u_\varepsilon\rangle \quad \text{for } \varepsilon > 0.$$

Passing to the limit $\varepsilon \to 0$, we obtain

$$\mathcal{E}^0_t(w) \geq \overline{e} + \langle \xi_0, w{-}u_0\rangle, \tag{3.6}$$

where we used the strong convergence $w_\varepsilon{-}u_\varepsilon \to w{-}u_0$. By $\mathcal{E}^\varepsilon_t \xrightarrow{\Gamma} \mathcal{E}^0_t$, we already know $\mathcal{E}^0_t(u_0) \leq \overline{e}$, but choosing $w = u_0$ in (3.6) gives $\mathcal{E}^0_t(u_0) = \overline{e}$ as desired. With this, (3.6) immediately gives $\xi_0 \in \partial\mathcal{E}^0_t(u_0)$.

The above result is only one of many possible versions and several generalizations are feasible. For instance, we may combine time discretization with time step $\tau \to 0$ with the limit $\varepsilon \to 0$. More precisely, if we solve the time discretized problem (see Sect. 2.6) for PG$^\varepsilon$ with time step τ, we obtain an approximation $\widehat{U}_{\tau_\varepsilon}$. Then, it can be shown that these approximations satisfy good a priori estimates and hence for every sequence $(\tau_n, \varepsilon_n) \to (0,0)$ there exists a subsequence and a solution of PG^0 such that the above convergences hold. We refer to [22, Thm. 4.1] or [24, Thm. 3.12] for results of this type.

3.2. A priori estimates

The energy-dissipation principle states that every solution $u_\varepsilon \in \mathrm{AC}([0,T];V)$ of PG$^\varepsilon$, i.e., (1.1) is satisfied, also satisfies the energy-dissipation balance in the sense that there exists a measurable selection $\xi_\varepsilon : (0,T) \to V^*$ such that $\xi_\varepsilon(t) \in \partial\mathcal{E}_t^\varepsilon(u_\varepsilon(t))$ a.e. in $(0,T)$ and that

$$\begin{aligned}\mathcal{E}_T^\varepsilon(u_\varepsilon(T)) + \int_0^T \Big(\Psi^\varepsilon_{u_\varepsilon(r)}(u_\varepsilon'(r)) + \Psi^{\varepsilon,*}_{u_\varepsilon(r)}\big(B^\varepsilon(r,u_\varepsilon(r)) - \xi_\varepsilon(r)\big)\Big)\,\mathrm{d}r \\ = \mathcal{E}_0^\varepsilon(u_\varepsilon(0)) + \int_0^T \Big(\partial_t\mathcal{E}_r^\varepsilon(u_\varepsilon(r)) + \big\langle B^\varepsilon(t,u_\varepsilon(r)), u_\varepsilon'(t)\big\rangle\Big)\,\mathrm{d}r. \qquad (3.7)\end{aligned}$$

Estimating the last term via the Young- Fenchel inequality and (3.B$^\varepsilon$b), we obtain

$$\begin{aligned}\big\langle B^\varepsilon(r,u_\varepsilon(r)), u_\varepsilon'(r)\big\rangle &\le \Psi^\varepsilon_{u_\varepsilon(r)}(u_\varepsilon'(r)) + \Psi^{\varepsilon,*}_{u_\varepsilon(r)}\big(B^\varepsilon(r,u_\varepsilon(r))\big) \\ &\le \Psi^\varepsilon_{u_\varepsilon(r)}(u_\varepsilon'(r)) + C_B\mathcal{E}_r^\varepsilon(u_\varepsilon(t))\end{aligned}$$

for the last term. Thus, using the terms involving $\Psi^\varepsilon_{u_\varepsilon(r)}(u_\varepsilon'(r))$ and employing $\Psi_u^{\varepsilon,*} \ge 0$ as well as (3.E$^\varepsilon$b), we arrive at

$$\mathcal{E}_T^\varepsilon(u_\varepsilon(T)) \le \mathcal{E}_0^\varepsilon(u_\varepsilon(0)) + \int_0^T \big(C_T + C_B\big)\mathcal{E}_r^\varepsilon(u_\varepsilon(r))\,\mathrm{d}r.$$

With $u_\varepsilon(0) = u_\varepsilon^0$ and the well-preparedness (3.4) the Gronwall lemma yields

$$G^\varepsilon(u_\varepsilon(t)) \;\le\; \mathcal{E}_t^\varepsilon(u_\varepsilon(t)) \le 2\mathcal{E}_0^0(u_0^0)\,\mathrm{e}^{(C_T+C_B)t} \le \overline{E} := 2\mathcal{E}_0^0(u_0^0)\,\mathrm{e}^{(C_T+C_B)T}.$$

Thus, assumption (3.E$^\varepsilon$b) guarantees that there exists a compact set $K \subset V$ such that $u_\varepsilon(t) \in K$ for all $(\varepsilon,t) \in (0,1)\times[0,T]$. As $K \subset B_R(0) \subset V$, we can apply the superlinearity (3.Ψ^εb) and the control (3.B$^\varepsilon$b) of B^ε to estimate

$$g_R\big(B^\varepsilon(t,u_\varepsilon(t))\big) \le \Psi^{\varepsilon,*}_{u_\varepsilon(t)}\big(B^\varepsilon(t,u_\varepsilon(t))\big) \le C_B\mathcal{E}_t^\varepsilon(u_\varepsilon(t)) \le C_B\overline{E}.$$

This implies the boundedness of the non-variational perturbation, viz.

$$\exists\, R_B^* > 0\;\forall\,(\varepsilon,t) \in (0,1)\times[0,T]: \quad \|B^\varepsilon(t,u_\varepsilon(t))\|_{V^*} \le R_B^*. \qquad (3.8)$$

Inserting the bounds for $\mathcal{E}_t^\varepsilon(u_\varepsilon(t))$ (and hence for $\partial_t\mathcal{E}_t^\varepsilon(u_\varepsilon(t))$) and for $B^\varepsilon(t,u_\varepsilon(t))$ into (3.7) we obtain

$$\int_0^T \Big(\Psi^\varepsilon_{u_\varepsilon(r)}(u_\varepsilon'(r)) - R_B^*\|u_\varepsilon'(r)\|_V + \Psi^{\varepsilon,*}_{u_\varepsilon(r)}\big(B^\varepsilon(r,u_\varepsilon(r)) - \xi_\varepsilon(r)\big)\Big)\,\mathrm{d}r \le C_E.$$

Using that Ψ^ε and $\Psi^{\varepsilon,*}$ are bounded from below by the superlinear function g_R (cf. (3.Ψ^εb)) and using (3.8) again, we arrive at

$$\exists\, C_\Psi > 0\;\forall\,\varepsilon \in (0,1]: \quad \int_0^T \big(g_R(\|u_\varepsilon'(t)\|_V) + g_R(\|\xi_\varepsilon\|_{V^*})\big)\,\mathrm{d}t \le C_\Psi. \qquad (3.9)$$

3.3. Convergent subsequences

By (3.9) and the criterion of DE LA VALLÉE- POUSSIN for uniform integrability, the family $u_\varepsilon : [0,T] \to V$ is equicontinuous. As all values $u_\varepsilon(t)$ lie in the compact set K, the ARZELÀ- ASCOLI theorem (e.g., [1, Proposition 3.3.1]) gives a subsequence $\varepsilon_k \to 0$ such that the uniform convergence (3.5a) holds. Moreover, (3.9) also implies weak compactness, hence we may also assume $u'_{\varepsilon_k} \rightharpoonup u'_0$ in $\mathrm{L}^1(0,T;V)$, which is (3.5c).

By the continuity (3.B$^\varepsilon$a), we obtain convergence of the non-variational terms, namely

$$\forall\, t \in [0,T]: \quad B^{\varepsilon_k}(t, u_{\varepsilon_k}(t)) \to B^0(t, u_0(t)) \text{ uniformly in } V^*. \tag{3.10}$$

Using the positivity of Ψ^ε and $\Psi^{\varepsilon,*}$, we then obtain that $\overline{e}^\varepsilon : t \mapsto \mathcal{E}^\varepsilon_t(u_\varepsilon(t))$ are uniformly bounded in $\mathrm{BV}([0,T])$, such that HELLY's selection principle allows to extract a subsequence (not relabeled) such that

$$\forall\, t \in [0,T]: \quad \overline{e}^{\varepsilon_k}(t) \to \overline{e}^0(t) \geq \mathcal{E}^0_t(u_0(t)), \tag{3.11}$$

where the last estimate follows from (3.E$^\varepsilon$e).

Again, based on the superlinear bounds (3.9), we can find a further subsequence (not relabeled) such that $t \mapsto (u'_{\varepsilon_k}(t), \xi_{\varepsilon_k}(t), \partial_t \mathcal{E}^{\varepsilon_k}_t(u_{\varepsilon_k}(t)))$ generates a YOUNG measure $\boldsymbol{\mu} = (\mu_t)_{t\in[0,T]} \in \mathcal{Y}([0,T]; V\times V^*\times\mathbb{R})$ in the sense that

$$\begin{aligned} &\int_0^T F\big(t, u'_{\varepsilon_k}(t), \xi_{\varepsilon_k}(t), \partial_t \mathcal{E}^\varepsilon_t(u_\varepsilon(t))\big)\,\mathrm{d}t \\ &\quad\to \int_0^T \int_{V\times V^*\times\mathbb{R}} F(t,v,\zeta,p)\,\mathrm{d}\mu_t(v,\zeta,p)\,\mathrm{d}t, \end{aligned} \tag{3.12}$$

for all continuous functions $F : [0,T]\times V\times V^*\times\mathbb{R} \to \mathbb{R}$ (where $V\times V^*$ is equipped with the weak topology) with $F(t,v,\zeta,p) \leq C(1+\|v\| + \|\zeta\|_*)$. We refer to Appendix A.

3.4. Passage to the limit and conclusion of the proof of Theorem 3.1

We can now go back to the energy-dissipation balance (3.7) and pass to the limit $\varepsilon_k \to 0$, where we employ BALDER's lower semicontinuity result [6] for weakly normal integrands in the form of [32, Thm. 4.3], see Theorem A.1. The main point here is that for $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \alpha_3) \in [0,\infty)^3$ the mappings

$$F_k^{\boldsymbol{\alpha}} : [0,T]\times V\times V^*\times\mathbb{R} \to \mathbb{R};\ (t,v,\zeta,p) \mapsto \alpha_1 \Psi_{u_{\varepsilon_k}(t)}(v) + \alpha_2 \Psi^{\varepsilon,*}_{u_{\varepsilon_k}(t)}(\zeta) + \alpha_3 p,$$

satisfy a liminf estimate, namely

$$(v_k, \zeta_k, p_k) \rightharpoonup (v, \zeta, p) \text{ in } V\times V^*\times\mathbb{R}$$
$$\Longrightarrow \quad \liminf_{k\to\infty} F_k^{\boldsymbol{\alpha}}(t, v_k, \zeta_k, p_k) \geq F_\infty^{\boldsymbol{\alpha}}(t, v, \zeta, p),$$

where $F_\infty^{\boldsymbol{\alpha}}(t, v, \zeta, p) = \alpha_1\Psi^0_{u_0(t)}(v)+\alpha_2\Psi^{0,*}_{u_0(t)}(\zeta)+\alpha_3 p$. But the latter liminf estimate follows easily from the Mosco convergence condition (3.Ψ^εc), because we already know $u_{\varepsilon_k} \to u_0(t)$ and $\mathcal{E}_t^{\varepsilon_k}(u_{\varepsilon_k}(t)) \leq \overline{E}$. In particular, we obtain the three liminf estimates

$$\int_r^s \Psi^0_{u_0(t)}(u_0'(t))\,\mathrm{d}t \leq \liminf_{k\to\infty}\int_r^s \Psi^{\varepsilon_K}_{u_{\varepsilon_K}(t)}(u'_{\varepsilon_k}(t))\,\mathrm{d}t, \tag{3.13a}$$

$$\int_r^s \Psi^{0,*}_{u_0(t)}\big(B^{0_k}(t, u_0(t)) - \xi_0(t)\big)\,\mathrm{d}t \leq \liminf_{k\to\infty}\int_r^s \Psi^{\varepsilon_k,*}_{u_{\varepsilon_k}(t)}\big(B^{\varepsilon_k}(t, u_{\varepsilon_k}(t)) - \xi_{\varepsilon_k}(t)\big)\,\mathrm{d}t, \tag{3.13b}$$

$$\int_r^s \partial_t\mathcal{E}_t^0(u_0(t))\,\mathrm{d}t \leq \liminf_{k\to\infty}\int_r^s \partial_t\mathcal{E}_t^{\varepsilon_k}(u_{\varepsilon_k}(t))\,\mathrm{d}t, \tag{3.13c}$$

where $0 \leq r < s \leq T$ are arbitrary.

Adding the three inequalities in (3.13) and using the limit $\overline{e}^0$ in (3.11), we arrive at

$$\overline{e}^0(T) + \int_0^T\int_{V\times V^*\times\mathbb{R}}\Big(\Psi^0_{u_0(t)}(v) + \Psi^{0,*}_{u_0(r)}\big(B^0(r, u_0(r)) - \zeta\big) - p\Big)\,\mathrm{d}\mu_t(v, \zeta, p)\,\mathrm{d}t$$
$$\leq \mathcal{E}_0^0(u_0^0) + \int_0^T \big\langle B^0(r, u_0(r)), u_0'(r)\big\rangle\,\mathrm{d}r. \tag{3.14}$$

Here the convergence of the right-hand side follows from the well-preparedness (3.4) and the fact that the strong L^∞ convergence (3.10) and the weak convergence (3.5c) imply the convergence of the integral.

Now, we exploit the main structural property of the Young measure $\boldsymbol{\mu}$ which states that for a.a. $t \in [0, T]$ the supports of μ_t lie in the set of accumulation points of defining sequences. More, there is a set $\mathscr{N} \subset [0, T]$ with $|\mathscr{N}| = 0$ such that for $t \in [0, T]\backslash\mathscr{N}$:

$$\mathrm{sppt}(\mu_t) \subset \mathrm{Li}(t) := \bigcap_{m=1}^\infty \mathrm{clos}_{\mathrm{weak}}\Big(\big\{\, (u'_{\varepsilon_k}(t), \xi_{\varepsilon_k}(t), \partial_t\mathcal{E}_t^{\varepsilon_k}(u_{\varepsilon_k}(t))) \;\big|\; k \geq m \,\big\}\Big).$$

Hence, the closedness condition (3.E^εf) guarantees that for all $t \in [0, T]\backslash\mathscr{N}$ we have

$$\forall\,(v, \zeta, p) \in \mathrm{sppt}(\mu_t):\ \zeta \in \partial\mathcal{E}_t^0(u_0(t)),\ \ p \leq \partial_t\mathcal{E}_t^0(u_0(t)),\ \ \overline{e}^0(t) = \mathcal{E}_t^0(u_0(t)).$$

We can now estimate further in (3.14). By (3.11) the first term $\overline{e}^0(T)$ is estimated from below by $\mathcal{E}_T^0(u_0(T))$. The term involving $\Psi_u^0(v)$ can be estimated by the convexity of $\Psi_u^0(\cdot)$ and the fact that μ_t is a probability measure with v-expectation u_0', i.e.,

$$u_0'(t) = \int_{V\times V^*\times\mathbb{R}} v\,\mathrm{d}\mu_t(v, \zeta, p).$$

This follows simply by testing (3.12) with $F(t, v, \zeta, p) = \langle \eta(t), v\rangle$ for all $\eta \in \mathrm{L}^\infty(0, T; V^*)$. Thus,

$$\int_0^T \Psi_{u_0(t)}(u_0'(t))\,\mathrm{d}t \le \int_0^T \int_{V\times V^*\times\mathbb{R}} \Psi_{u_0(t)}(v)\,\mathrm{d}\mu_t(v, \zeta, p)\,\mathrm{d}t.$$

For the term involving $\Psi_u^{0,*}(v)$, we cannot apply JENSEN's inequality, as $\partial\mathcal{E}_t^0(u)$ may not be convex. Thus, for $t \in [0, T]\backslash\mathscr{N}$ we select $\xi_0(t) \in \partial\mathcal{E}_t^0(u_0(t))$ with

$$\Psi_{u_0(r)}^{0,*}\big(B^0(r, u_0(r)) - \xi_0(t)\big) = \min\big\{\, \Psi_{u_0(r)}^{0,*}\big(B^0(r, u_0(r)) - \zeta\big) \,\big|\, \zeta \in \partial\mathcal{E}_t^0(u_0(t)) \,\big\}.$$

Such a measurable selection exists, see Lemma A.4 in Appendix A.

Finally using $p \le \partial_t\mathcal{E}_t^0(u_0(t))$ on $\mathrm{Li}(t)$ the estimate (3.14) yields, for all $s \in (0, T]$,

$$\begin{aligned}
&\mathcal{E}_s^0(u_0(s)) + \int_0^s \Big(\Psi_{u_0(t)}^0(u_0'(t)) + \Psi_{u_0(r)}^{0,*}\big(B^0(r, u_0(r)) - \xi_0(t)\big) - \partial_t\mathcal{E}_t^0(u_0(t))\Big)\,\mathrm{d}t \\
&\quad \le \mathcal{E}_0^0(u_0^0) + \int_0^s \big\langle B^0(t, u_0(t)), u_0'(t)\big\rangle \mathrm{d}t.
\end{aligned} \tag{3.15}$$

Moreover, by the FENCHEL- YOUNG inequality and the chain-rule inequality (3.E$^\varepsilon$d), which is used for $\varepsilon = 0$ only, the left-hand side can be estimated from below via

$$\begin{aligned}
&\mathcal{E}_s^0(u_0(s)) + \int_0^s \Big(\Psi_{u_0(t)}^0(u_0'(t)) + \Psi_{u_0(t)}^{0,*}\big(B^0(t, u_0(t)) - \xi_0(t)\big) - \partial_t\mathcal{E}_t^0(u_0(t))\Big)\,\mathrm{d}t \\
&\quad \overset{\text{FY}}{\ge} \mathcal{E}_s^0(u_0(s)) + \int_0^s \Big(\langle B^0(t, u_0(t)) - \xi_0(t), u_0'(t)\rangle - \partial_t\mathcal{E}_t^0(u_0(t))\Big)\,\mathrm{d}t \\
&\quad \overset{\text{chain}}{\ge} \mathcal{E}_s^0(u_0(s)) + \int_0^s \Big(\langle B^0(t, u_0(t)), u_0'(t)\rangle - \frac{\mathrm{d}}{\mathrm{d}t}\big(\mathcal{E}_t^0(u_0(t))\big)\Big)\,\mathrm{d}t \\
&\quad = \mathcal{E}_0^0(u_0(0)) + \int_0^s \big\langle B^0(t, u_0(t)), u_0'(t)\big\rangle \mathrm{d}t.
\end{aligned} \tag{3.16}$$

Thus, we conclude that all inequalities in (3.15) and (3.16) are equalities, which means that the FENCHEL- YOUNG estimate has to hold with equality a.e. in $[0, T]$, which gives the desired differential inclusion $B^0(t, u_0(t)) - \xi_0(t) \in \partial\Psi_{u_0(t)}^0(u_0'(t))$ or

$$B^0(t, u_0(t)) \in \partial\Psi_{u_0(t)}^0(u_0'(t)) + \partial\mathcal{E}_t^0(u_0(t)) \qquad \text{a.e. in } [0, T].$$

Additionally, we observe that the liminf estimates

$$\mathcal{E}_t^0(u_0(t)) \le \overline{e}^\infty(t) = \lim_{\varepsilon_k\to 0} \mathcal{E}_t^{\varepsilon_k}(u_{\varepsilon_k}(t))$$

as well as the liminf estimates in (3.13) are indeed equalities as well. Thus, (3.5b), (3.5d), and (3.5b) are established and the proof of Theorem 3.1 is complete.

3.5. Improved result for state-independent dissipation

The result on evolutionary Γ-convergence given in Theorem 3.1 has a rather strong assumption, namely the MOSCO convergence of $(\varepsilon, u) \mapsto \Phi_u^\varepsilon(\cdot)$ in the space V. This assumption is too strong for a number of important applications. For instance, for the parabolic equation

$$\big(2+\cos(x_1/\varepsilon)\big)u' = \operatorname{div}\big(A(\tfrac{1}{\varepsilon}x)\nabla u\big) \text{ in } \Omega \subset \mathbb{R}^d, \qquad u = 0 \text{ on } \partial\Omega,$$

we may choose the gradient structure $(\mathbf{Q}, \mathcal{E}^\varepsilon, \Psi^\varepsilon)$ with

$$V = \mathrm{L}^2(\Omega), \quad \mathcal{E}^\varepsilon(u) = \int_\Omega \frac{1}{2}\nabla u \cdot A(\tfrac{1}{\varepsilon}x)\nabla u \,\mathrm{d}x, \quad \Psi^\varepsilon(v) = \int_\Omega \frac{2+\cos(x_1/\varepsilon)}{2}\, v(x)^2 \,\mathrm{d}x.$$

However, Ψ^ε Γ-converges to Ψ_{harm} in the weak topology of $\mathrm{L}^2(\Omega)$, while it Γ-converges to Ψ_{arith} in the strong topology.

Here we want present a generalized version of [17] where evolutionary Γ-convergence was established under the weaker assumption $\Psi^\varepsilon \xrightarrow{\Gamma} \Psi^0$, i.e., Γ-convergence in the strong topology only, see (4.6) and [20, Lem. 3.2.2] for a simple proof.

If we inspect the proof in the previous section, then we see that the weak Γ-convergence of $\Psi_{u_\varepsilon}^\varepsilon$ was used only once, namely for deriving the liminf estimate (3.13a). The point is that we only derived the weak convergence $u'_{\varepsilon_k} \rightharpoonup u'_0$ in $\mathrm{L}^1(0, T; V)$. However, the "weak" convergence may have two origins, namely first due to oscillations in time and second due to weak convergence of $u'_\varepsilon(t) \rightharpoonup u'_0(t)$ in V. The idea in [17] is to consider piecewise affine interpolants $u_{\varepsilon,\tau}$ of u_ε for fixed time steps $\tau > 0$. This averages potential oscillations in time as $u'_{\varepsilon,\tau}$ is piecewise constant. Moreover, we can use the strong convergence of $u_{\varepsilon_k}(t) \to u_0(t)$ which implies that $u'_{\varepsilon_k,\tau}(t) \to u'_{0,\tau}(t)$ in V for a.a. $t \in [0, T]$. Finally, the limit $\tau \to 0$ is done after the limit $\varepsilon_k \to 0$ is already performed.

Our precise assumptions, which replace (3.Ψ^εc), are the following:

Uniform continuity. For all $R > 0$

$$\exists \text{ modulus of continuity } \omega_R \ \forall \varepsilon \in [0, 1] \ \forall u_1, u_2 \text{ with } G^\varepsilon(u_j) \le R$$
$$\forall v \in V: \quad \big|\Psi_{u_1}^\varepsilon(v) - \Psi_{u_2}^\varepsilon(v)\big| \le \omega_R(\|u_1 - u_2\|_V) g_R(\|v\|_V), \tag{3.17a}$$

Strong Γ-convergence. For all $R > 0$, we have

$$u_\varepsilon \to u_0 \text{ and } \sup \mathcal{E}_t^\varepsilon(u_\varepsilon) \le R \implies \Psi_{u_\varepsilon}^\varepsilon \xrightarrow{\Gamma} \Psi_{u_0}^0, \tag{3.17b}$$

where g_R is the coercivity function defined in (3.Ψ^εb).

COROLLARY 3.3. **(Strong Γ-convergence for $\mathcal{E}^\varepsilon$ and Ψ^ε)** *All results of Theorem 3.1 remain true if assumption* (3.Ψ^εc) *is replaced by* (3.17).

Proof. To start with, we recall that the strong Γ-convergence of (3.17b) implies the weak Γ-convergence of the LEGENDRE- FENCHEL dual, i.e., $\Psi^{\varepsilon,*}_{u_\varepsilon} \xrightarrow{\Gamma} \Psi^{0,*}_{u}$, see (2.1). Thus, the liminf estimate (3.13b) follows exactly as above.

It remains to find a new proof for the liminf estimate (3.13a). Using the notation

$$J^\varepsilon(u,v) := \int_0^T \Psi^\varepsilon_{u(t)}\big(v(t)\big)\,\mathrm{d}t,$$

we have to show $\liminf_{k\to\infty} J^{\varepsilon_k}(u_{\varepsilon_k}, u'_{\varepsilon_k}) \geq J^0(u_0, u'_0)$, where our sequence $(u_{\varepsilon_k})_k$ satisfies

$$\text{(a) } \|u_{\varepsilon_k} - u_0\|_{\mathrm{C}^0([0,T];V)} \to 0,\ \ \text{(b) } \|u'_{\varepsilon_k} - u'_0\|_{\mathrm{L}^1(0,T;V)} \to 0,$$
$$\text{(c) } \int_0^T g_R\big(\|u'_{\varepsilon_k}(t)\|\big)\,\mathrm{d}t \leq C_g,$$

with $R \geq \sup\{\, \|u_{\varepsilon_k}\|_\infty \mid k \in \mathbb{N}\,\}$.

For time steps $\tau = T/N > 0$ with $N \in \mathbb{N}$, we define piecewise constant and piecewise affine interpolants $\overline{u}^\tau_{\varepsilon_k}$ and $\widehat{u}^\tau_{\varepsilon_k}$ as in (2.25). By the uniform convergence (a), we have equicontinuity of the sequence $(u_{\varepsilon_k})_k$, and hence

$$\mu_\tau := \sup\{\, \|u_{\varepsilon_k} - \overline{u}_{\varepsilon_k}\|_{\mathrm{C}^0([0,T];V)} \mid k \in \mathbb{N}\,\} \;\to\; 0 \quad \text{for } \tau \to 0.$$

With (3.17a) and (c), we obtain the lower bound

$$\begin{aligned} J^{\varepsilon_k}(u_{\varepsilon_k}, u'_{\varepsilon_k}) &\geq J^{\varepsilon_k}(\overline{u}^\tau_{\varepsilon_k}, u'_{\varepsilon_k}) - \int_0^T \omega_R\big(\|u_{\varepsilon_k} - \overline{u}^\tau_{\varepsilon_k}\|\big) g_R\big(\|u'_{\varepsilon_k}\|\big)\,\mathrm{d}t \\ &\geq J^{\varepsilon_k}(\overline{u}^\tau_{\varepsilon_k}, u'_{\varepsilon_k}) - \omega_R(\mu_\tau) C_g. \end{aligned}$$

On the intervals $((n-1)\tau, n\tau)$ the integrand $\Psi^{\varepsilon_k}_{\overline{u}^\tau_{\varepsilon_k}(t)}(\cdot)$ is independent of t and convex. Hence, we can apply JENSEN's inequality and replace $v_k(t) = u'_{\varepsilon_k}(\cdot)$ by its average over this interval, which is exactly

$$\frac{1}{\tau}\int_{(n-1)\tau}^{n\tau} u'_{\varepsilon_k}(r)\,\mathrm{d}r = \frac{1}{\tau}\big(u_{\varepsilon_k}(n\tau) - u_{\varepsilon_k}((n-1)\tau)\big) = \widehat{u}^{\tau\,\prime}_{\varepsilon_k}(t) \quad \text{for } t \in ((n-1)\tau, n\tau).$$

Thus, we have the lower bound $J^{\varepsilon_k}(u_{\varepsilon_k}, u'_{\varepsilon_k}) \geq J^{\varepsilon_k}(\overline{u}^\tau_{\varepsilon_k}, \widehat{u}^{\tau\,\prime}_{\varepsilon_k}) - \omega_R(\mu_\tau) C_g$.

For $k \to \infty$, we have $\overline{u}_{\varepsilon_k} \to \overline{u}_0$ in V and $\widehat{u}^{\tau\,\prime}_{\varepsilon_k} \to \widehat{u}^{\tau\,\prime}_0$ in V a.e. in $[0,T]$. Hence, we can exploit the liminf estimate of the strong Γ-convergence $\Psi^\varepsilon_{u_\varepsilon} \xrightarrow{\Gamma} \Psi^0_{u_0}$. FATOU's lemma leads to

$$\begin{aligned} \liminf_{k\to\infty} J^{\varepsilon_k}(u_{\varepsilon_k}, u'_{\varepsilon_k}) &\geq \liminf_{k\to\infty} J^{\varepsilon_k}(\overline{u}^\tau_{\varepsilon_k}, \widehat{u}^{\tau\,\prime}_{\varepsilon_k}) - \omega_R(\mu_\tau) C_g \\ &\overset{\text{Fatou}}{\geq} J^0(\overline{u}^\tau_0, \widehat{u}^{\tau\,\prime}_0) - \omega_R(\mu_\tau) C_g \;\geq\; J^0(u_0, \widehat{u}^{\tau\,\prime}_0) - 2\omega_R(\mu_\tau) C_g, \end{aligned}$$

where we used $\|u_0 - \overline{u}^\tau_0\|_\infty \leq \mu_\tau$ for the last step.

Thus, using $\omega_R(\mu_\tau) \to 0$ for $\tau \to 0$ it remains to show $L := \liminf_{\tau\to 0} J^0(u_0, \widehat{u}_0^{\tau}{}')$ $\geq J^0(u_0, u_0')$. Choose a subsequence τ_m such that $J^0(u_0, \widehat{u}_0^{\tau_m}{}') \to L$. We now use the well-known fact that $\widehat{u}_0^{\tau_m}{}' \to u_0'$ in $\mathrm{L}^1(0, T; V)$, which implies that there exists a further subsequence (not relabeled) such that $\widehat{u}_0^{\tau_m}{}'(t) \to u_0'(t)$ in V a.e. in $[0, T]$. Moreover, $\Psi^0_{u_0(t)}(\cdot) : V \to [0, \infty)$ is continuous, because it is convex and bounded from above by the Legendre- Fenchel dual of $\xi \mapsto g_R(\|\xi\|_{V^*})$. This gives $\Psi^0_{u_0(t)}(\widehat{u}_0^{\tau_m}{}'(t)) \to \Psi^0_{u_0(t)}(u_0'(t))$ a.e. in $[0, T]$, and Fatou's lemma implies $L = \liminf_{m\to\infty} J^0(u_0, \widehat{u}_0^{\tau_m}{}') \geq J^0(u_0, u_0')$.

Altogether we have established $\liminf_{k\to\infty} J^{\varepsilon_k}(u_{\varepsilon_k}, u_{\varepsilon_k}') \geq J^0(u_0, u_0')$, and thus Corollary 3.3 is proved. □

4. Homogenization of reaction-diffusion systems

In this section, we provide a non-trivial example that highlights the applicability of our abstract existence theory as well as the theory of evolutionary Γ-convergence. We refer to [25,27,28] and the references therein for general homogenization results that are typical for semilinear systems, where the leading-order terms are decoupled. Our example of a reaction-diffusion system is a general quasilinear parabolic system, where the leading terms may be coupled but need to have a variational structure.

Our system for the vector $u(t, x) \in \mathbb{R}^I$ reads as follows:

$$\begin{aligned} A^\varepsilon(x, u(t,x))\partial_t u(t,x) &= \operatorname{div}\Big(\partial_{\nabla u} F^\varepsilon\big(x, u(t,x), \nabla u(t,x)\big)\Big) \\ &\quad - \partial_u F^\varepsilon(x, u(t,x), \nabla u(t,x)) + b^\varepsilon(x, t, u(t,x)) \quad \text{in } \Omega, \\ 0 &= \partial_{\nabla u} F^\varepsilon\big(x, u(t,x), \nabla u(t,x)\big)\nu(x) \quad \text{on } \partial\Omega. \end{aligned} \tag{4.1}$$

Generally, we assume that $\Omega \subset \mathbb{R}^d$ is a bounded domain with Lipschitz boundary $\partial\Omega$. For simplicity, we have imposed Neumann boundary conditions only, but more general conditions including Dirichlet or Robin boundary conditions could be used as well.

We first summarize the needed assumptions on the functions A^ε, F^ε, and b^ε, then we show that these assumptions imply the assumptions needed for the existence theory in Sect. 2, and finally discuss under which conditions we have evolutionary Γ-convergence for $\varepsilon \to 0$.

4.1. The existence result

For the matrix $A^\varepsilon(x, u) \in \mathbb{R}^{I\times I}_{\mathrm{sym}} := \{ A \in \mathbb{R}^{I\times I} \mid A = A^\top \}$, we make the assumption

$$\forall \varepsilon \in [0,1]: \quad A^\varepsilon : \Omega \times \mathbb{R}^I \to \mathbb{R}^{I\times I}_{\mathrm{sym}} \text{ is a Carathéodory function,} \tag{4.2a}$$

$$\exists\, C_A > 0\, \forall\, \varepsilon \in [0,1]\, \forall\, x \in \Omega\, \forall\, u, v \in \mathbb{R}^I : \quad \frac{1}{C_A}|v|^2 \leq \langle A^\varepsilon(x,u)v, v\rangle \leq C_A|v|^2. \tag{4.2b}$$

Here $G : \Omega\times\mathbb{R}^M \to \mathbb{R}^N$ is called a Carathéodory function if $x \mapsto G(x,z)$ is measurable for all $z \in \mathbb{R}^m$ and $z \mapsto G(x,z)$ is continuous for a.a. $x \in \Omega$.

For simplicity, we will assume that the functions $F^\varepsilon(x,\cdot,\cdot)$ are convex, but much weaker conditions would be possible (e.g., λ-convexity in u or poly-convexity in $U = \nabla u$).

$$\forall\, \varepsilon \in [0,1] : \quad F^\varepsilon : \Omega \times (\mathbb{R}^I\times\mathbb{R}^{I\times d}) \to \mathbb{R} \text{ is a Carathéodory function,} \tag{4.2c}$$

$$\forall\, \varepsilon \in [0,1]\, \forall_{\text{a.a.}} x \in \Omega : \quad F^\varepsilon(x,\cdot,\cdot) : \mathbb{R}^I\times\mathbb{R}^{I\times d} \to \mathbb{R} \text{ is convex,} \tag{4.2d}$$

$$\exists\, C_F > 0\, \exists\, p, q > 1\, \forall\, \varepsilon \in [0,1]\, \forall\, (x,u,U) \in \Omega\times\mathbb{R}^I\times\mathbb{R}^{I\times d} : \\ F^\varepsilon(x,u,U) \geq C_F\big(1 + |u|^q + |U|^p\big). \tag{4.2e}$$

For the non-gradient terms b^ε, we impose the following conditions:

$$\forall\, \varepsilon \in [0,1] : \quad b^\varepsilon : \Omega \times ([0,T]\times\mathbb{R}^I) \to \mathbb{R}^I \text{ is a Carathéodory function,} \tag{4.2f}$$

$$\exists\, h \in \mathrm{L}^2(\Omega),\ C_B > 0,\ r > 1\, \forall\, (\varepsilon,t,x,u) \in [0,1]\times\Omega\times[0,T]\times\mathbb{R}^I : \\ |b^\varepsilon(x,t,u)| \leq h(x) + C_B|u|^r. \tag{4.2g}$$

We choose the basic space $V = \mathrm{L}^2(\Omega;\mathbb{R}^I)$, the energy functionals

$$\mathcal{E}^\varepsilon(u) = \begin{cases} \displaystyle\int_\Omega F^\varepsilon(x,u(x),\nabla u(x))\,\mathrm{d}x & \text{for } u \in \mathrm{W}^{1,p}(\Omega;\mathbb{R}^I), \\ \infty & \text{otherwise,} \end{cases}$$

and the dissipation potentials

$$\Psi_u^\varepsilon(v) := \int_\Omega \frac{1}{2}\langle A^\varepsilon(x,u(x))v(x), v(x)\rangle\,\mathrm{d}x.$$

Thus, the perturbed gradient systems $\mathrm{PG}^\varepsilon = (V,\mathcal{E}^\varepsilon,\Psi^\varepsilon,b^\varepsilon)$ are fully specified, and we want to apply our abstract theory. Before doing so, we note that in our conditions the exponent q appears three times: (i) the first relation in (4.3) implies $\mathrm{W}^{1,p}(\Omega) \subset \mathrm{L}^q(\Omega)$, (ii) the coercivity (4.2e) of F asks for the lower bound $C_F|u|^q$, and (iii) the second relation in (4.3) says that $B(\cdot,u(\cdot))$ is controlled by $C(1+\|u\|_q^q)$.

PROPOSITION 4.1. *Let the functions A^ε, F^ε, and b^ε satisfy the conditions (4.2), where the coefficients p, q, and r satisfy the relations*

$$1 - \frac{d}{p} > -\frac{d}{q} \quad \textit{and} \quad q \geq 2r. \tag{4.3}$$

Then, for each initial condition $u_\varepsilon^0 \in \mathrm{L}^2(\Omega;\mathbb{R}^I)$ with $\mathcal{E}^\varepsilon(u_\varepsilon^0) < 0$ there is a solution $u_\varepsilon : [0,T] \to \mathrm{L}^2(\Omega;\mathbb{R}^I)$ of (4.1) such that $u_\varepsilon \in \mathrm{H}^1(0,T;\mathrm{L}^2(\Omega)) \cap \mathrm{C}^0_{\mathrm{weak}}([0,T];\mathrm{W}^{1,p}(\Omega))$.

The proof is a consequence of our abstract existence result in Theorem 2.5. We easily find the LEGENDRE- FENCHEL dual $\Psi_u^{\varepsilon,*}(\xi) = \int_\Omega \frac{1}{2}\langle \xi(x), (A^\varepsilon(x,u(x)))^{-1}\xi(x)\rangle\,\mathrm{d}x$. Clearly, (2.Ψa) holds and we have the equicoercivities

$$\Psi_u^\varepsilon(v) \geq \frac{1}{2C_A}\|v\|_V^2 \quad \text{and} \quad \Psi_u^{\varepsilon,*}(\xi) \geq \frac{1}{2C_A}\|\xi\|_{V^*}^2,$$

which imply the desired superlinearities (2.Ψb). Finally, the MOSCO convergence $\Psi_{u_n}^\varepsilon \xrightarrow{\mathrm{M}} \Psi_u^\varepsilon$ (here $\varepsilon > 0$ is still fixed) follows since $u_n \to u$ in V implies that $A^\varepsilon(\cdot, u_n(\cdot)) \to A^\varepsilon(\cdot, u(\cdot))$ a.e. in Ω along suitable subsequences. To see that this is sufficient for MOSCO convergence, we use the MOREAU- YOSIDA regularizations

$$\Psi_u^{\varepsilon,\lambda}(v) := \inf\left\{ \Psi_{u_n}^\varepsilon(w) + \frac{\lambda}{2}\|w-v\|_{\mathrm{L}^2}^2 \;\middle|\; w \in \mathrm{L}^2(\Omega;\mathbb{R}^I) \right\},$$

where $\lambda > 0$. It is easy to see that $\Psi_u^{\varepsilon;\lambda}$ is still quadratic, but now with the matrix $\lambda A^\varepsilon(A^\varepsilon+\lambda I)^{-1}$. By [3, Thm. 3.26], we have $\Psi_{u_n}^\varepsilon \xrightarrow{\mathrm{M}} \Psi_u^\varepsilon$ if and only if for all $v \in V = \mathrm{L}^2(\Omega;\mathbb{R}^I)$ and all $\lambda > 0$ we have the pointwise convergence $\Psi_{u_n}^{\varepsilon,\lambda}(v) \to \Psi_u^{\varepsilon,\lambda}(v)$. But this follows immediately by the boundedness of A^ε and LEBESGUE's dominated convergence theorem. Hence, (2.Ψc) is shown as well.

The energy functionals $\mathcal{E}^\varepsilon$ are convex and independent of time. Hence, (2.Ea) and (2.Ec) hold trivially. By the coercivity of F^ε, we obtain the coercivity of $\mathcal{E}^\varepsilon$, namely

$$\mathcal{E}^\varepsilon(u) \geq \int_\Omega C_F\big(1 + |u|^q + |\nabla u|^p\big)\,\mathrm{d}x \geq \widetilde{c}\|u\|_{\mathrm{W}^{1,p}}^{\min\{p,q\}} - \widetilde{C}, \tag{4.4}$$

such that sublevels are bounded in $\mathrm{W}^{1,p}(\Omega;\mathbb{R}^I)$. Because this space is compactly embedded in $V = \mathrm{L}^2(\Omega;\mathbb{R}^I)$ by assumption (4.3), we conclude that (2.Eb) holds. The chain rule (2.Ed) and the weak-strong closedness of the FRÉCHET subdifferential (which is the same as the convex subdifferential) follows by convexity, see Remark 3.2 or [23].

We now set $B^\varepsilon(t,u)(x) = b^\varepsilon(x,t,u(x))$ and obtain the continuity (2.Ba) simply from the continuity of $b^\varepsilon(x,\cdot,\cdot)$ and $2r \leq q$. The energy control (2.Bb) follows from (4.2e) and the second condition in (4.3). Thus, all the abstract assumptions of Theorem 2.5 are fulfilled, and Proposition 4.1 is established.

4.2. The homogenization result

We want to apply the evolutionary Γ-convergence of Sect. 3 for homogenization, i.e., we assume that the x-dependence of A^ε, F^ε, and b^ε is of oscillatory type, namely

$$A^\varepsilon(x,u) = \mathbb{A}\left(\tfrac{1}{\varepsilon}x, u\right), \quad F^\varepsilon(x,u,U) = \mathbb{F}\left(\tfrac{1}{\varepsilon}x, u, U\right), \quad b^\varepsilon(x,t,u) = \mathbb{B}\left(\tfrac{1}{\varepsilon}x, u\right), \tag{4.5}$$

where the functions $\mathbb{A}$, $\mathbb{F}$, and $\mathbb{B}$ are assumed to be 1-periodic in all directions, i.e., $\mathbb{G}(y+k) = \mathbb{G}(y)$ for all $y \in \mathbb{R}^d$ and $k \in \mathbb{Z}^d$.

For the quadratic dissipation potentials Ψ_u^ε, we have the following Γ-convergences:

$$(\varepsilon_n, u_n) \to (0, u) \in \mathbb{R}\times\mathrm{L}^2(\Omega;\mathbb{R}^n) \implies \Big(\Psi_{u_n}^{\varepsilon_n} \xrightharpoonup{\Gamma} \Psi_u^{\mathrm{harm}} \text{ and } \Psi_{u_n}^{\varepsilon_n} \xrightarrow{\Gamma} \Psi_u^{\mathrm{aver}}\Big), \tag{4.6}$$

where the harmonic-mean Ψ_u^{harm} and the average Ψ_u^{aver} are defined via

$$\Psi_u^{\mathrm{harm}}(v)=\int_\Omega \frac{1}{2}\langle A^{\mathrm{harm}}(u(x))v(x), v(x)\rangle\,\mathrm{d}x \quad \text{with } A^{\mathrm{harm}}(u)^{-1}=\int_{(0,1)^d}\mathbb{A}(y,u)^{-1}\,\mathrm{d}y,$$
$$\Psi_u^{\mathrm{aver}}(v)=\int_\Omega \frac{1}{2}\langle A^{\mathrm{aver}}(u(x))v(x), v(x)\rangle\,\mathrm{d}x \quad \text{with } A^{\mathrm{aver}}(u)=\int_{(0,1)^d}\mathbb{A}(y,u)\,\mathrm{d}y.$$

The strong Γ-convergence $\Psi_{u_n}^{\varepsilon_n} \xrightarrow{\Gamma} \Psi_u^{\mathrm{aver}}$ follows simply from the pointwise convergence $\Psi_{u_n}^{\varepsilon_n}(v) \to \Psi_u^{\mathrm{aver}}(v)$ for all v and the equi-LIPSCHITZ continuity. The weak Γ-convergence $\Psi_{u_n}^{\varepsilon_n} \xrightharpoonup{\Gamma} \Psi_u^{\mathrm{aver}}$ follows by (2.1) and LEGENDRE- FENCHEL transform as $\Psi_u^{\varepsilon,*}$ is given in terms of $(A^\varepsilon)^{-1}$, see also [8, Exa. 2.36].

In particular, we see that MOSCO convergence only holds for the case that the harmonic and the arithmetic mean are equal, which means that $\mathbb{A}(y,u)$ has to be independent of y.

For the energy functional $\mathcal{E}^\varepsilon$, we can rely on the general theory of homogenization as surveyed in [9]. Using the uniform coercivity (4.4), we obtain weak Γ-convergence in $\mathrm{W}^{1,p}(\Omega;\mathbb{R}^I)$ and, by the compact embedding, strong Γ-convergence in $V = \mathrm{L}^2(\Omega;\mathbb{R}^I)$ toward the limit

$$\mathcal{E}^0(u) = \int_\Omega F^{\mathrm{hom}}(u(x), \nabla u(x))\,\mathrm{d}x \text{ with}$$
$$F^{\mathrm{hom}}(u,U) := \min\Big\{ \int_{(0,1)^d} \mathbb{F}(y,u,U+\nabla\Phi(y))\,\mathrm{d}y \;\Big|\; \Phi \in \mathrm{W}^{1,p}_{\mathrm{per}}((0,1)^d;\mathbb{R}^I)\Big\},$$

see [9, Thm. 5.1, pp. 135]. Of course, $\mathcal{E}^0 : V \to [0,\infty]$ is a again a convex and lower semicontinuous functional. Finally, setting

$$B^0(t,u) : x \mapsto b^{\mathrm{aver}}(t,u(x)) \quad \text{with } b^{\mathrm{aver}}(t,u) = \int_{(0,1)^d}\mathbb{B}(y,u)\,\mathrm{d}y,$$

we obtain the desired convergence $B^{\varepsilon_n}(t_n,u_n) \to B^0(t,u)$ if $(\varepsilon_n,t_n,u_n) \to (0,t,u)$ in $[0,1]\times[0,T]\times V$.

Hence, we see that Theorem 3.1, which is the main result on evolutionary Γ-convergence, is only applicable if we have the MOSCO convergence $\Psi_{u_k}^{\varepsilon_k} \xrightarrow{\mathrm{M}} \Psi_u^0$, which means $\Psi_u^{\mathrm{harm}} = \Psi_u^{\mathrm{aver}}$. Thus, we need to assume that $\mathbb{A}(y,u)$ does not depend on the microscopic periodicity variable $y \in \mathbb{R}^d/_{\mathbb{Z}^d}$. In summary, we obtain the following result.

THEOREM 4.2. (Homogenization I) *Consider the perturbed gradient system* $\mathrm{PG}^\varepsilon = (\mathrm{L}^2(\Omega;\mathbb{R}^I), \mathcal{E}^\varepsilon, \Psi^\varepsilon, b^\varepsilon)$ *given as above. Assume that* (4.2) *holds and that* (4.5) *holds* *with* $\mathbb{A}$ *independent of the variable* $y = \frac{1}{\varepsilon}x$*, then we have evolutionary* Γ*-convergence in the sense of Theorem 3.1 to the perturbed gradient system* $(\mathrm{L}^2(\Omega;\mathbb{R}^I), \mathcal{E}^0, \Psi^{\mathrm{aver}}, b^{\mathrm{aver}})$*, i.e., solutions* u_ε *of the reaction-diffusion system* (4.1) *converge to solutions of the homogenized system*

$$\begin{aligned} A^{\mathrm{aver}}(u)\partial_t u &= \mathrm{div}\,\big(\partial_{\nabla u} F^{\mathrm{hom}}(u,\nabla u)\big) - \partial_u F^{\mathrm{hom}}(u,\nabla u) + b^{\mathrm{aver}}(t,u) && \textit{in } \Omega, \\ 0 &= \partial_{\nabla u} F^{\mathrm{hom}}(u,\nabla u)\nu && \textit{on } \partial\Omega. \end{aligned} \tag{4.7}$$

The case where $\mathbb{A}(y,u)$ depends on $y \in \mathbb{R}^d/_{\mathbb{Z}^d}$ is more difficult. Under additional assumptions, we will be able to use the improved theory developed in Corollary 3.3, as we can use $\Psi^{\varepsilon_k}_{u_k} \xrightarrow{\Gamma} \Psi^{\mathrm{aver}}_u$, which gives assumption (3.17b). However, we need to establish the uniform continuity (3.17a). For this, we note that $G^\varepsilon(u_j) \le R$ implies $\|u_j\|_{\mathrm{W}^{1,p}} \le C_R$. Now, assuming $p > d$ we first observe $\|u_j\|_{\mathrm{L}^\infty} \le \widetilde{C}_R < \infty$, and a Gagliardo- Nirenberg estimate yields

$$\|u_1 - u_2\|_{\mathrm{L}^\infty} \le C_{\mathrm{GN}} \|u_1 - u_2\|^\theta_{\mathrm{L}^2} \|u_1 - u_2\|^{1-\theta}_{\mathrm{W}^{1,p}} \le C_{\mathrm{GN}} (2C_R)^{1-\theta} \|u_1 - u_2\|^\theta_{\mathrm{L}^2}.$$

Now assuming the uniform continuity

$$\begin{aligned} &\forall\,\rho > 0\ \exists \text{ modulus of contin.}\, \omega_\rho\ \forall\, y \in (0,1)^d\ \forall\, u_j \in B_\rho(0) \subset \mathbb{R}^I : \\ &\quad \big|\mathbb{A}(y,u_1) - \mathbb{A}(y,u_2)\big| \le \omega_\rho(|u_1 - u_2|), \end{aligned} \tag{4.8}$$

we can estimate the difference $\Psi^\varepsilon_{u_1}(v) - \Psi^\varepsilon_{u_2}(v)$ of the dissipation potentials pointwise under the integral and obtain

$$\begin{aligned} &\forall\, u_j \in V \text{ with } G^\varepsilon(u_j) \le R : \ \big|\Psi^\varepsilon_{u_1}(v) - \Psi^\varepsilon_{u_2}(v)\big| \\ &\quad \le \omega_{\widehat{C}_R}\big(C_{\mathrm{GN}}(2C_R)^{1-\theta}\|u_1 - u_2\|^\theta_{\mathrm{L}^2}\big)\|v\|^2_{\mathrm{L}^2}. \end{aligned}$$

This is exactly the desired uniform continuity (3.17a). Thus, Corollary 3.3 is applicable under the additional assumption that $p > d$ and that (4.8) holds, which gives our second homogenization result, where $\mathbb{A}$ now may depend periodically on $y = \frac{1}{\varepsilon}x$.

THEOREM 4.3. (Homogenization II) *Consider perturbed gradient systems* $(V, \mathcal{E}^\varepsilon, \Psi^\varepsilon, b^\varepsilon)$ *with* $V = \mathrm{L}^2(\Omega;\mathbb{R}^I)$ *given as above. Assume that* (4.2) *holds with* $p > d$ *and that* (4.5) *together with* (4.8)*. Then, all the conclusions of Theorem 4.2 remain true.*

Indeed, we conjecture that these two additional conditions (either $\mathbb{A}$ independent of $y = \frac{1}{\varepsilon}x$ or (4.8)) are not really necessary. Using two-scale unfolding as in [25, 27,28] and a suitable version of Ioffe's theorem, it should be possible to prove the fundamental liminf estimate

$$\int_0^T \Psi_{u(t)}^{\text{aver}}(u'(t))\,\mathrm{d}t \le \liminf_{k\to\infty} \int_0^T \Psi_{u_{\varepsilon_k}(t)}^{\varepsilon_k}(u'_{\varepsilon_k}(t))\,\mathrm{d}t$$

in much more general cases.

Acknowledgements

This research has been partially funded by Deutsche Forschungsgemeinschaft (DFG) through the grant SFB 910 Control of self-organizing nonlinear systems, Project A5 "Pattern formation in coupled parabolic systems" (for A.M.) and Project A8 "Nonlinear evolution equations: model hierarchies and complex fluids" (for A.B. and E.E.).

A Appendix

In this section, we provide some tools on parametrized YOUNG measures of which we made use in the last section. First, we give some notions related to YOUNG measures and which were originally introduced by BALDER [6].

In the following, let $\mathcal{V}$ be a reflexive (separable) BANACH space. In fact, we employed the following results to the separable and reflexive BANACH space $\mathcal{V} = V \times V^* \times \mathbb{R}$ endowed with the product topology. Further, let for an interval $\mathscr{L}_{(0,T)}$ be the LEBESGUE σ-algebra of $(0, T)$ and let $\mathscr{B}(\mathcal{V})$ be the BOREL σ-algebra of $\mathcal{V}$.

Then, we say that a $\mathscr{L}_{(0,T)} \otimes \mathscr{B}(\mathcal{V})$-measurable function $\mathcal{H}(0, T) \times \mathcal{V} \to (-\infty, +\infty]$ is a *weakly normal integrand* if for a.a. $t \in (0, T)$ the map $w \mapsto \mathcal{H}(t, w)$ is sequentially lower semicontinuous with respect to the weak topology of $\mathcal{V}$.

Furthermore, we denote by $\mathscr{M}(0, T; \mathcal{V})$ the set of all $\mathscr{L}_{(0,T)}$-measurable functions $y : (0, T) \to \mathcal{V}$. Then, a sequence $(w_n)_{n\in\mathbb{N}} \subset \mathscr{M}(0, T; \mathcal{V})$ is said to be *weakly-tight* if there exists a weakly normal integrand $\mathcal{H} : (0, T) \times \mathcal{V} \to (-\infty, +\infty]$ such that the map $w \mapsto \mathcal{H}(t, w)$ has weakly compact sublevels in $\mathcal{V}$ for a.a. $t \in (0, T)$, and it holds

$$\sup_{n\in\mathbb{N}} \int_0^T \mathcal{H}(t, w_n(t))\,\mathrm{d}r < +\infty. \tag{A.1}$$

Finally, a family $\boldsymbol{\mu} = (\mu_t)_{t\in(0,T)}$ of BOREL probability measures on $\mathcal{V}$ is called YOUNG measure if on $(0, T)$ the map $t \mapsto \mu_t(B)$ is $\mathscr{L}_{(0,T)}$-measurable for all $B \in \mathscr{B}(\mathcal{V})$. With $\mathscr{Y}(0, T; \mathcal{V})$, we denote the set of all YOUNG measures on $\mathcal{V}$.

THEOREM A.1. *For all $n \in \mathbb{N}$, let $\mathcal{H}_n, \mathcal{H} : (0, T) \times \mathcal{V} \to (-\infty, +\infty]$ be a weakly normal integrand such that for all $w \in \mathcal{V}$ and for a.a. $t \in (0, T)$ we have*

$$\mathcal{H}(t, w) \le \inf\{\liminf_{n\to\infty} \mathcal{H}_n(t, w_n) \mid w_n \rightharpoonup w \quad in\ \mathcal{V}\}. \tag{A.2}$$

Let $(w_n)_{n\in\mathbb{N}} \subset \mathscr{M}(0, T; \mathcal{V})$ be a weakly-tight sequence. Then, there exists a subsequence $(w_{n_k})_{k\in\mathbb{N}}$ and a YOUNG *measure $\boldsymbol{\mu} = (\mu_t)_{t\in(0,T)}$ such that for almost every $t \in (0, T)$ we have*

$$\operatorname{sppt}(\mu_t) \subset \operatorname{Li}(t) := \bigcap_{p=1}^{\infty} \operatorname{clos}_{\text{weak}}\big(\{w_{n_k}(t) \mid k \ge p\}\big), \tag{A.3}$$

i.e., μ_t is concentrated on the set of all limit points of the sequence $(w_{n_k}(t))_{k\in\mathbb{N}}$ with respect to the weak topology of $\mathcal{V}$ and if additionally the sequence $t \mapsto \mathcal{H}_n^-(t, w_{n_k}(t)) :=$ $\max\{-\mathcal{H}_n(t, w_{n_k}(t)), 0\}$ *is uniformly integrable, it holds*

$$\int_0^T \int_{\mathcal{V}} \mathcal{H}(t, w)\, \mathrm{d}\mu_t(w)\, \mathrm{d}t \le \liminf_{k\to} \int_0^T \mathcal{H}_{n_k}(t, w_{n_k}(t))\, \mathrm{d}t. \tag{A.4}$$

Proof. This is shown in STEFANELLI [32, Thm. 4.3, pp. 1626]. □

As a corollary of the previous theorem, we obtain a version of the so-called "Fundamental Theorem for YOUNG Measures" in the weak topology. It provides a characterization of limits of weakly converging sequences $(w_n)_{n\in\mathbb{N}}$.

THEOREM A.2. (Fundamental Theorem for Young measures) *Let $1 \le p \le \infty$ and let $(w_n)_{n\in\mathbb{N}} \subset \mathrm{L}^p(0, T; \mathcal{V})$ be a bounded sequence. If $p = 1$, we suppose further that $(w_n)_{n\in\mathbb{N}}$ is uniformly integrable in $\mathrm{L}^1(0, T; \mathcal{V})$. Then, there exists a subsequence $(w_{n_k})_{k\in\mathbb{N}}$ and a* YOUNG *measure $\boldsymbol{\mu} = (\mu_t)_{t\in(0,T)} \in \mathscr{Y}(0, T; \mathcal{V})$ such that for a.a. $t \in (0, T)$ relation (A.3) holds and, setting*

$$\mathrm{w}(t) := \int_{\mathcal{V}} w \, \mathrm{d}\mu_t(w) \quad a.a.\ t \in (0, T). \tag{A.5}$$

it holds

$$w_{n_k} \rightharpoonup \mathrm{w} \quad in\ \mathrm{L}^p(0, T; \mathcal{V}) \quad as\ k \to \infty, \tag{A.6}$$

with $\rightharpoonup$ replaced by $\rightharpoonup^*$ if $p = \infty$.

LEMMA A.3. *Let the perturbed gradient system $(V, \mathcal{E}, \Psi, B)$ satisfy Assumptions* (2.E), (2.Ψ) *as well as* (2.B)*, and let $u \in \mathrm{AC}(0, T; \mathcal{V})$ be an absolutely continuous curve such that*

$$\partial\mathcal{E}_t(u(t)) \ne \emptyset \quad for\ a.a.\ t \in (0, T) \quad and \quad \sup_{t\in(0,T)} \mathcal{E}_t(u(t)) < \infty. \tag{A.7}$$

Let $\boldsymbol{\mu} = (\mu_t)_{t\in(0,T)} \in \mathscr{Y}(0, T; \mathcal{V})$ be a YOUNG *measure such that for a.a. $t \in (0, T)$ we have*

$$\int_0^T \int_{V\times V^*\times\mathbb{R}} \Big(\Psi_{u(t)}(v) + \Psi^*_{u(t)}(B(t,u(t)) - \zeta)\Big)\,\mathrm{d}\mu_t(v,\zeta,p)\,\mathrm{d}t < \infty,$$
$$u'(t) = \int_{V\times V^*\times\mathbb{R}} v\,\mathrm{d}\mu_t(v,\zeta,p),$$
$$\zeta \in \partial\mathcal{E}_t(u(t)) \text{ and } p \le \partial_t\mathcal{E}_t(u(t)) \text{ for all } (v,\zeta,p) \in \mathrm{sppt}(\mu_t).$$

Then,

$$t \mapsto \mathcal{E}_t(u(t)) \text{ is absolutely continuous on } (0,T) \text{ and}$$
$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}_t(u(t)) \ge \int_{V\times V^*\times\mathbb{R}} \big(\langle u'(t),\zeta\rangle + p\big)\,\mathrm{d}\mu_t(v,\zeta,p) \quad \text{for a.a. } t \in (0,T).$$

Proof. This can be proven in exactly the same manner as in [23, Prop. B.1]. □

LEMMA A.4. (Measurable selection) *Let the perturbed gradient system* $(V, \mathcal{E}, \Psi, B)$ *satisfy Assumptions* (2.E), (2.Ψ)*, and* (2.B). *Furthermore, let* $u \in \mathrm{AC}([0,T]; V)$ *be an absolutely continuous curve satisfying* (A.7)*, and assume that for all* $t \in [0,T]$

$$\mathcal{S}(t,u(t),u'(t)) := \Big\{(\zeta,p) \in V^*\times\mathbb{R} \,\Big|\, \zeta \in \partial\mathcal{E}_t(u(t)) \cap \big(B(t,u(t)) - \partial\Psi_{u(t)}(u'(t))\big),$$
$$p \le \partial_t\mathcal{E}_t(u(t))\Big\} \text{ is non-empty.}$$

Then, there exist measurable functions $\xi : (0,T) \to V^*$ *and* $p : (0,T) \to \mathbb{R}$ *such that*

$$(\xi(t), p(t)) \in \arg\min \Big\{ \Psi^*_{u(t)}\big(B(t,u(t)) - \zeta\big) - p \,\Big|\, (\zeta,p) \in \mathcal{S}(t,u(t),u'(t)) \Big\}$$

for a.a. $t \in (0,T)$.

Proof. This can be proven in the same way as [23, Lem. B.2]. □

REFERENCES

[1] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient flows in metric spaces and in the space of probability measures*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 2005.

[2] L. Ambrosio. Minimizing movements. *Rend. Accad. Naz. Sci. XL Mem. Mat. Appl. (5)*, 19, 191–246, 1995.

[3] H. Attouch. *Variational Convergence of Functions and Operators*. Pitman Advanced Publishing Program. Pitman, 1984.

[4] H. Attouch. *On Multivalued Evolution Equations in Hilbert Spaces*. *Israel J. Math.*, 12, 373–390, 1972.

[5] J.-P. Aubin and H. Frankowska. *Set-valued Analysis*. Birkhäuser, Boston, 1990.

[6] E. J. Balder. A general approach to lower semicontinuity and lower closure in optimal control theory. *SIAM J. Control Optim.*, 22, 570–598, 1984.

[7] H. Brézis. Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert. North-Holland Publishing Company, Amsterdam · London, 1973.

[8] A. Braides. *Γ-Convergence for Beginners*. Oxford University Press, 2002.

[9] A. Braides. A handbook of Γ-convergence. In M. Chipot and P. Quittner, editors, *Handbook of Differential Equations. Stationary Partial Differential Equations. Volume 3*, pages 101–213. Elsevier, 2006.

[10] C. Castaing and M. Valadier. *Convex analysis and measurable multifunctions*. Lect. Notes Math. Vol. 580. Springer, Berlin, 1977.

[11] P. Colli. On some doubly nonlinear evolution equations in Banach spaces. *Japan J. Indust. Appl. Math.*, 9, 181–203, 1992.

[12] P. Colli and A. Visintin. On a class of doubly nonlinear evolution equations. *Commun. Partial Differential Equations*, 15(5), 737–756, 1990.

[13] G. Dal Maso. *An Introduction to Γ-Convergence*. Birkhäuser Boston Inc., Boston, MA, 1993.

[14] P. Dondl, T. Frenzel, and A. Mielke. A gradient system with a wiggly energy and relaxed EDP-convergence. *WIAS preprint 2459*, 2017.

[15] N. Dunford and J. T. Schwartz. *Linear Operator. Part I*. Interscience, John Wiley & Sons, 1959.

[16] I. Ekeland and R. Temam. *Analyse Convexe et Problèmes Variationnels*. Dunod, 1974.

[17] M. Liero and S. Reichelt. Homogenization of Cahn–Hilliard-type equations via evolutionary Γ-convergence. *Nonl. Diff. Eqns. Appl. (NoDEA)*, 2018.

[18] M. Liero, A. Mielke, M. A. Peletier, and D. R. M. Renger. On microscopic origins of generalized gradient structures. *Discr. Cont. Dynam. Systems Ser. S*, 10(1), 1–35, 2017.

[19] A. Mielke. Deriving effective models for multiscale systems via evolutionary Γ-convergence. In E. Schöll, S. Klapp, and P. Hövel, editors, *Control of Self-Organizing Nonlinear Systems*, pages 235–251. Springer, 2016.

[20] A. Mielke. On evolutionary Γ-convergence for gradient systems (Ch. 3). In A. Muntean, J. Rademacher, and A. Zagaris, editors, *Macroscopic and Large Scale Phenomena: Coarse Graining, Mean Field Limits and Ergodicity*, Lecture Notes in Applied Mathematics and Mechanics Vol. 3, pages 187–249. Springer, 2016.

[21] B. S. Mordukhovich. *Variational analysis and generalized differentiation I–Basic Theory*. Springer Berlin, 2006.

[22] A. Mielke, T. Roubíček, and U. Stefanelli. Γ-limits and relaxations for rate-independent evolutionary problems. *Calc. Var. Part. Diff. Eqns.*, 31, 387–416, 2008.

[23] A. Mielke, R. Rossi, and G. Savaré. Nonsmooth analysis of doubly nonlinear evolution equations. *Calc. Var. Part. Diff. Eqns.*, 46(1-2), 253–310, 2013.

[24] A. Mielke, R. Rossi, and G. Savaré. Balanced-Viscosity (BV) solutions to infinite-dimensional rate-independent systems. *J. Europ. Math. Soc.*, 18, 2107–2165, 2016.

[25] A. Mielke, S. Reichelt, and M. Thomas. Two-scale homogenization of nonlinear reaction-diffusion systems with slow diffusion. *Networks Heterg. Materials*, 9(2), 353–382, 2014.

[26] M. Ôtani. Nonmonotone Perturbations for Nonlinear Parabolic Equations Associated with Subdifferential Operators, Periodic Problems. *J. Diff. Eqns.*, 54, 248–273, 1982.

[27] S. Reichelt. Error estimates for elliptic equations with not exactly periodic coefficients. *Adv. Math. Sci. Appl.*, 25, 117–131, 2016.

[28] S. Reichelt. Corrector estimates for a class of imperfect transmission problems. *Asymptot. Analysis*, 105, 3–26, 2017.

[29] R. Rossi and G. Savaré. Gradient flows of non convex functionals in Hilbert spaces and applications. *ESAIM Control Optim. Calc. Var.*, 12, 564–614, 2006.

[30] E. Sandier and S. Serfaty. Gamma-convergence of gradient flows with applications to Ginzburg-Landau. *Gamma-convergence of gradient flows with applications to Ginzburg-Landau.*, , 1627–1672, 2004.

[31] S. Serfaty. Gamma-convergence of gradient flows on Hilbert spaces and metric spaces and applications. *Discr. Cont. Dynam. Systems Ser. A*, 31(4), 1427–1451, 2011.

[32] U. Stefanelli. The Brezis-Ekeland principle for doubly nonlinear equations. *SIAM J. Control Optim.*, 47(3), 1615–1642, 2008.

[33] A. Visintin. Structural compactness and stability of semi-monotone flows. *SIAM J. Math. Anal.*, 50(3), 2628–2663, 2018.

Aras Bacho and Etienne Emmrich
Sekretariat MA 5-3, Institut für Mathematik
Technische Universität Berlin
Straße des 17. Juni 136
10623 Berlin
Germany

Alexander Mielke
Weierstraß-Institut für Angewandte Analysis und Stochastik
Mohrenstraße 39
10117 Berlin
Germany

and

Institut für Mathematik
Humboldt-Universität zu Berlin
Rudower Chausee 25
12489 Berlin
Germany
E-mail: alexander.mielke@wias-berlin.de