

ruhr.paD

UA Ruhr Zentrum für  
partielle Differentialgleichungen

# Finite element discretization of local minimization schemes for rate-independent evolutions

Ch. Meyer and M. Sievers

Preprint 2018-08

# Finite Element Discretization of Local Minimization Schemes for Rate-Independent Evolutions

Christian Meyer · Michael Sievers

Received: date / Accepted: date

**Abstract** This paper is concerned with a space-time discretization of a rate-independent evolution governed by a non-smooth dissipation and a non-convex energy functional. For the time discretization, we apply the local minimization scheme introduced in [EM06], which is known to resolve time discontinuities, which may show up due to the non-convex energy. The spatial discretization is performed by classical linear finite elements. We show that accumulation points of the sequence of discrete solutions for mesh size tending to zero exist and are so-called parametrized solutions of the continuous problem. The discrete problems are solved by means of a mass lumping scheme for the non-smooth dissipation functional in combination with a semi-smooth Newton method. A numerical test indicates the efficiency of this approach. In addition, we compared the local minimization scheme with a time stepping scheme for global energetic solutions, which shows that both schemes yield different solutions with differing time discontinuities.

**Keywords** Rate independent evolutions · parametrized solutions · finite elements · semi-smooth Newton methods

---

This research was supported by the German Research Foundation (DFG) under grant number HE 6077/8-1 within the priority program Non-smooth and Complementarity-based Distributed Parameter Systems: Simulation and Hierarchical Optimization (SPP 1962).

C. Meyer

Technische Universität Dortmund, Fakultät für Mathematik, Vogelpothsweg 87, 44227 Dortmund, Germany

Tel.: +49 (0)231 755 5412

Fax: +49 (0)231 755 5416

E-mail: cmeyer@math.tu-dortmund.de

M. Sievers

Technische Universität Dortmund, Fakultät für Mathematik, Vogelpothsweg 87, 44227 Dortmund, Germany

Tel.: +49 (0)231 755 7243

Fax: +49 (0)231 755 5416

E-mail: michael.sievers@math.tu-dortmund.de

## 1 Introduction

This paper is concerned with the finite element discretization of rate-independent evolutions. We consider an evolutionary system of the form

$$0 \in \partial\mathcal{R}(\dot{z}(t)) + D_z\mathcal{I}(t, z(t)) \quad \text{a.e. in } [0, T], \quad (1.1)$$

where  $\mathcal{I}$  denotes a possibly non-convex energy functional, while  $\mathcal{R}$  is a positive 1-homogeneous dissipation functional. The precise assumptions on  $\mathcal{I}$  and  $\mathcal{R}$  will be specified in Section 2 below. The positive homogeneity of  $\mathcal{R}$  implies that (1.1) is in fact invariant under time rescalings, i.e., the process described by (1.1) is rate-independent. Such a behavior is observed for physical systems driven by external loads ranging on a time-scale that is much slower than the internal scale. Thus, for most of the time, solutions follow the external forces, while forming sudden jumps in between. To be more precise, under appropriate assumptions on the energy, in particular convexity, the solution is continuous in time, whereas discontinuities may show up, if  $\mathcal{I}$  is not convex w.r.t.  $z$ , see [MT04] for details.

There exists a multitude of solution concepts for rate-independent systems accounting for these potential discontinuities. For an overview, we refer to [Mie11, MR15]. In this paper, we focus on the notion of so called *parametrized solutions*. The key idea of this solution concept is that jumps should rather be seen as (infinitely) fast viscous transitions between two stable regimes. One way to derive the existence of parametrized solutions is therefore a viscous regularization of (1.1) accompanied by a subsequent limit analysis for vanishing regularization, cf. e.g. [MZ14, MRS16]. This vanishing viscosity approach leads to the following limit system w.r.t. to an artificial time  $s$  that arises from an arc-length parametrization of the viscous trajectory:

$$\left. \begin{aligned} t(0) = 0, \quad z(0) = z_0, \quad t'(s) + \|z'(s)\| = 1, \\ 0 \in \partial\mathcal{R}(z'(s)) + \lambda(s)z'(s) + D_z\mathcal{I}(t(s), z(s)) \\ \lambda(s) \geq 0, \quad \lambda(s)(1 - \|z'(s)\|) = 0, \end{aligned} \right\} \quad (1.2)$$

This evolutionary system allows to identify three different regimes and ascribe them a physical meaning (see [MR15]):

– *Sticking:*

In this case, the potential forces are too small so that  $z'(s) = 0$  and  $t'(s) = 1$  and the state does not change.

– *Rate-independent slip:*

Here, it holds  $0 < \|z'(s)\| < 1$  and  $0 < t'(s) < 1$  so that the state indeed changes but in such a manner that the dissipation is strong enough to compensate the driving forces.

– *Viscous-jump:*

In this case, we have  $\|z'(s)\| = 1$  and  $t'(s) = 0$  which means that the system may switch into a viscous behavior. Meanwhile the physical time stands still ( $t'(s) = 0$ ) so that this viscous transition is seen as a jump.

A chain-rule for  $\mathcal{I}$  allows the following equivalent reformulation of (1.2) as energy-equality, cf. e.g. [MRS12]:

$$t(0) = 0, \quad z(0) = z_0, \quad t'(s) + \|z'(s)\| = 1, \quad (1.3a)$$

$$\mathcal{I}(t(s), z(s)) + \int_0^s \mathcal{R}(z'(r)) + \lambda(s)\|z'(r)\| dr = \mathcal{I}(0, z_0) + \int_0^s \partial_t \mathcal{I}(t(r), z(r)) dr \quad (1.3b)$$

$$\lambda(s) \geq 0, \quad \lambda(s)(1 - \|z'(s)\|) = 0. \quad (1.3c)$$

The term  $\lambda(s)\|z'(s)\|$  can be interpreted as a viscous contribution, since it is a remnant of the vanishing viscosity approach. This formulation serves as the basis for the precise definition of a parametrized solution including the regularity of the involved functions, cf. Section 2 below.

A natural question concerns the approximation of solutions by discretization. For the class of global energetic solutions, there exists multiple paper on this topic. We only refer to [MR15] and the references therein. In contrast, less is known for parameterized solutions to the best of our knowledge. In [EM06], a particular time-scheme was introduced in order to approximate parametrized solutions. To be more precise, the authors proposed a *time-incremental local minimization scheme* of the following form:

$$z_k \in \arg \min \{ \mathcal{I}(t_{k-1}, z) + \mathcal{R}(z - z_{k-1}) : z \in \mathcal{Z}, \|z - z_{k-1}\|_{\mathbb{V}} \leq \tau \} \quad (1.4a)$$

$$t_k = \min \{ t_{k-1} + \tau - \|z_k - z_{k-1}\|_{\mathbb{V}}, T \}. \quad (1.4b)$$

The motivating background for this scheme is best explained by a comparison with a scheme for the approximation of global energetic solutions: The latter aims at finding global minimizers for the objective in (1.4a) in every time step so that the iteration reads

$$z_k \in \arg \min \{ \mathcal{I}(t_{k-1}, z) + \mathcal{R}(z - z_{k-1}) : z \in \mathcal{Z} \} \quad (1.5a)$$

$$t_k = t_{k-1} + \tau. \quad (1.5b)$$

Thanks to the positive homogeneity of  $\mathcal{R}$ , the stationarity conditions of (1.5a) are given by

$$0 \in \partial \mathcal{R} \left( \frac{z_k - z_{k-1}}{\tau} \right) + \mathcal{I}(t_{k-1}, z_k), \quad (1.6)$$

which, in view of (1.1), motivates the scheme in (1.5). However, in general, the global minimization in (1.5a) may induce unphysical discontinuities. Consider for instance a situation, where the difference between a local minimum of  $\mathcal{I}(t_{k-1}, \cdot)$  in the vicinity of  $z_{k-1}$  to all global minima of  $\mathcal{I}(t_{k-1}, \cdot)$  is so large that it cannot be compensated by the dissipation. Then the iteration will jump to a global minimizer which is certainly not physically meaningful in many applications. This motivates the additional inequality constraint in (1.4a) in order to mimic the search for local minimizers in the neighborhood of the old iterate. If, however, there is no such local minimizer so that the inequality constraint in (1.4a) is active (i.e., fulfilled with equality), then the stationarity condition in (1.6) will (in general) not be fulfilled. In this case, one therefore interrupts the evolution of the physical time (see the update in (1.4b)) until the state  $z$  fulfills the stationarity condition again. This may be seen as discrete analogon to the viscous-jump described above.

In [EM06], a convergence theory for  $\tau \searrow 0$  is developed for the finite dimensional case, where  $\dim(\mathcal{Z}) < \infty$ . The authors prove that a subsequence of discrete solutions (weakly) converges to a parametrized solution. However, it is not shown that the sequence  $\{t_k\}$ , generated by (1.4b), reaches the desired final time in a finite number of iterations. The same holds for a variant of (1.4), which is investigated in [Neg14]. In [NS17], another modification of (1.4) is considered, which does not account for the adaptive time discretization in (1.4b). For this variant, the authors show that the final time is reached in after a finite number of iterations. Recently, the original scheme in (1.4) was investigated in [Kne17] for a certain class of infinite dimensional problems (i.e.,  $\dim(\mathcal{Z}) = \infty$ ), providing a comprehensive convergence analysis. In particular, it is shown that the final time is reached in a finite number of steps and that a subsequence of iterates (weakly) converges to a parametrized solution.

Our paper is concerned with a *full discretization in time and space* of (1.1). The recent work [Kne17] serves as a starting point for our analysis and we include the additional errors induced by the discretization of the infinite dimensional state space  $\mathcal{Z}$  into the convergence analysis of [Kne17]. This leads to a practically implementable numerical scheme, which we tested by means of a numerical example highlighting the difference between (1.4) and (1.5).

Let us shortly outline the paper. In Section 2, we introduce our notation and state the assumptions on the energy and the dissipation functional. Moreover, we introduce the precise notion of parametrized solutions. Section 3 is then devoted to the convergence analysis. We derive suitable a-priori estimates for the discrete solution and establish a discrete version of the energy-equality (1.3b) which is shown to be preserved in the limit yielding in turn our main convergence result in Theorem 3.9. Thereafter, in Section 4, we describe the algorithmic realization of the fully discrete local minimization scheme based on a semismooth Newton-method with particularly chosen Newton-derivative. Moreover, we present an numerical example illustrating our theoretical findings.

## 2 Basic notations and standing assumptions

Let us start with some basic notation used throughout the paper. In the following,  $C > 0$  always is generic constant. Moreover, given two normed linear spaces  $X, Y$ , we denote by  $\langle \cdot, \cdot \rangle_{X^*, X}$  dual pairing and suppress the subscript, if there is no risk for ambiguity. By  $\|\cdot\|_X$ , we denote the norm in  $X$  and  $\mathcal{L}(X, Y)$  is the space of linear and bounded operators from  $X$  to  $Y$ . Furthermore,  $B_X(x, r)$  is the open ball in  $X$  around  $x \in X$  with radius  $r > 0$ . Given a convex functional  $f : X \rightarrow \mathbb{R} \cup \{\infty\}$ , we denote the (convex) subdifferential of  $f$  at  $x$  by  $\partial f(x) \subset X^*$  and its conjugate functional by  $f^* : X^* \rightarrow \mathbb{R} \cup \{\infty\}$ . Finally,  $|\Omega|$  stands for the Lebesgue measure of a set  $\Omega \subset \mathbb{R}^d$ ,  $d \in \mathbb{N}$ .

### 2.1 Assumptions on the data

Let us now introduce the assumptions on the quantities in (1.1). A prototypical example fulfilling all assumptions is given in Section 4.1 below.

#### Spaces

Throughout the paper,  $\mathcal{X}$  is a Banach space and  $\mathcal{Z}, \mathcal{V}$  are Hilbert spaces such that

$$\mathcal{Z} \overset{c,d}{\hookrightarrow} \mathcal{V} \hookrightarrow \mathcal{X},$$

where  $\overset{d}{\hookrightarrow}$  and  $\overset{c}{\hookrightarrow}$  refer to dense and compact embedding, respectively. We equip  $\mathcal{V}$  with the norm  $\|v\|_{\mathbb{V}} := \langle \mathbb{V}v, v \rangle_{\mathcal{V}^*, \mathcal{V}}^{1/2}$ , where  $\mathbb{V} \in \mathcal{L}(\mathcal{V}, \mathcal{V}^*)$  is a self-adjoint and coercive operator, i.e., there exist constants  $\gamma > 0$  such that  $\langle \mathbb{V}v, v \rangle_{\mathcal{V}^*, \mathcal{V}} \geq \gamma \|v\|_{\mathcal{V}}^2$  for all  $v \in \mathcal{V}$ , where  $\|\cdot\|_{\mathcal{V}}$  denotes the natural norm associated with the scalar product in  $\mathcal{V}$ .

#### Energy

The energy functional  $\mathcal{I}$  is supposed to have the following form:

$$\mathcal{I} : [0, T] \times \mathcal{Z} \rightarrow \mathbb{R}, \quad \mathcal{I}(t, z) = \frac{1}{2} \langle Az, z \rangle_{\mathcal{Z}^*, \mathcal{Z}} + \mathcal{F}(z) - \langle \ell(t), z \rangle_{\mathcal{V}^*, \mathcal{V}}. \quad (2.1)$$

Herein,  $A \in \mathcal{L}(\mathcal{Z}, \mathcal{Z}^*)$  is a coercive operator, i.e., there is a constant  $\alpha > 0$  such that  $\langle Az, z \rangle_{\mathcal{Z}^*, \mathcal{Z}} \geq \alpha \|z\|_{\mathcal{Z}}^2$ . The bilinear form  $a : \mathcal{Z} \times \mathcal{Z} \rightarrow \mathbb{R}$  induced by  $a(y, z) = \langle Ay, z \rangle_{\mathcal{Z}^*, \mathcal{Z}}$ ,  $y, z \in \mathcal{Z}$ , is thus bounded and coercive, too. Moreover, we assume that

$$\ell \in C^1([0, T]; \mathcal{V}^*) \text{ and } \mathcal{F} \in C^2(\mathcal{Z}; \mathbb{R}) \text{ with } \mathcal{F} \geq 0. \quad (2.2)$$

such that  $\mathcal{I} \in C^1([0, T] \times \mathcal{Z}; \mathbb{R})$ . The restriction of  $\ell(\cdot)$  to a functional on  $\mathcal{Z}$  is, for convenience, denoted by the same symbol. By setting  $\beta := (c_{\mathcal{Z}} \|\ell\|_{C([0, T]; \mathcal{V}^*)} + 1)/(4\alpha)$ , where  $c_{\mathcal{Z}} > 0$  denotes the embedding constant of  $\mathcal{Z} \hookrightarrow \mathcal{V}$ , we obtain the estimate

$$\mathcal{I}(t, z) \geq \alpha \|z\|_{\mathcal{Z}}^2 - c_{\mathcal{Z}} \|\ell\|_{C([0, T]; \mathcal{V}^*)} \|z\|_{\mathcal{Z}} \geq \|z\|_{\mathcal{Z}} - \beta \quad \forall t \in [0, T], z \in \mathcal{Z},$$

which in turn implies that

$$|\partial_t \mathcal{I}(t, z)| \leq c_{\mathcal{Z}} \|\ell\|_{C^1([0, T]; \mathcal{V}^*)} \|z\|_{\mathcal{Z}} \leq \mu(\mathcal{I}(t, z) + \beta) \quad \forall t \in [0, T], z \in \mathcal{Z},$$

with  $\mu := c_{\mathcal{Z}} \|\ell\|_{C^1([0, T]; \mathcal{V}^*)}$ . Gronwall's lemma thus gives for all  $t, s \in [0, T]$ ,  $z \in \mathcal{Z}$  that

$$\mathcal{I}(t, z) + \beta \leq (\mathcal{I}(s, z) + \beta) \exp(\mu|t - s|) \quad (2.3)$$

$$\text{and } |\partial_t \mathcal{I}(t, z)| \leq \mu(\mathcal{I}(s, z) + \beta) \exp(\mu|t - s|), \quad (2.4)$$

Furthermore, we assume that

$$D_z \mathcal{F} \in C^1(\mathcal{Z}, \mathcal{V}^*), \quad \|D_z^2 \mathcal{F}(z)v\|_{\mathcal{V}^*} \leq C(1 + \|z\|_{\mathcal{Z}}^q) \|v\|_{\mathcal{Z}} \quad (2.5)$$

for some  $q \geq 1$  so that, for every  $z \in \mathcal{Z}$ ,  $D_z \mathcal{F}(z)$  can uniquely be extended to a bounded and linear functional on  $\mathcal{V}$ , which we denote by the same symbol for convenience. Moreover, both  $\mathcal{F} : \mathcal{Z} \rightarrow \mathbb{R}$  and  $D_z \mathcal{F} : \mathcal{Z} \rightarrow \mathcal{Z}^*$  are supposed to be weak-weak continuous. This in particular guarantees that  $\mathcal{I}(t, \cdot)$  is weakly lower-semicontinuous.

### Dissipation

In the following, we denote by  $\mathcal{R}$  the dissipation potential and assume  $\mathcal{R} : \mathcal{V} \rightarrow [0, \infty)$  to be lower semicontinuous (l.s.c.), convex, and positively homogeneous of degree one. Moreover, we require  $\mathcal{R}$  to fulfill

$$\exists \rho, \bar{\rho} > 0 : \quad \rho \|v\|_{\mathcal{X}} \leq \mathcal{R}(v) \leq \bar{\rho} \|v\|_{\mathcal{V}} \quad \forall v \in \mathcal{V}. \quad (2.6)$$

Since  $\mathcal{R}$  is convex and l.s.c., it is locally Lipschitz continuous so that its subdifferential is bounded for every point of the domain.

### Initial state

The initial value  $z_0$  is supposed to satisfy  $z_0 \in \mathcal{Z}$  and  $Az_0 \in \mathcal{V}^*$ .

## 2.2 Definition of parametrized solutions

We now turn to our notion of solutions and give a rigorous definition thereof. As indicated in the introduction, the energy identity serves as a basis therefor.

**Definition 2.1** *Let an initial value  $z_0 \in \mathcal{Z}$  be given. We call a tuple  $(\hat{t}, \hat{z})$  parametrized solution of (1.1), if there exists an artificial end time  $S \geq T$  such that the following conditions are satisfied:*

(i) *Regularity:*

$$\hat{t} \in W^{1,\infty}(0, S), \quad \hat{z} \in W^{1,\infty}(0, S; \mathcal{V}) \cap L^\infty(0, S; \mathcal{Z}), \quad (2.7)$$

$$D_z \mathcal{I}(\hat{t}(s), \hat{z}(s)) \in \mathcal{V}^* \quad \text{f.a.a. } s \in (0, S). \quad (2.8)$$

(ii) *Initial and end time condition:*

$$\hat{t}(0) = 0, \quad \hat{z}(0) = z_0, \quad \hat{t}(S) = T. \quad (2.9)$$

(iii) *Complementarity-like relations:*

$$\hat{t}'(s) \geq 0, \quad \hat{t}'(s) + \|\hat{z}'(s)\|_{\mathbb{V}} \leq 1, \quad (2.10a)$$

$$\hat{t}'(s) \operatorname{dist}_{\mathcal{V}^*} \{-D_z \mathcal{I}(\hat{t}(s), \hat{z}(s)), \partial \mathcal{R}(0)\} = 0 \quad \text{f.a.a. } s \in (0, S), \quad (2.10b)$$

where  $\operatorname{dist}_{\mathcal{V}^*} \{\eta, \partial \mathcal{R}(0)\} = \inf \{\|\eta - w\|_{\mathbb{V}^{-1}} : w \in \partial \mathcal{R}(0)\}$  and  $\|\eta\|_{\mathbb{V}^{-1}}^2 = \langle \eta, \mathbb{V}^{-1} \eta \rangle$ .

(iv) *Energy identity:*

$$\begin{aligned} \mathcal{I}(\hat{t}(s), \hat{z}(s)) + \int_0^s \mathcal{R}(\hat{z}'(\sigma)) + \|\hat{z}'(\sigma)\|_{\mathbb{V}} \operatorname{dist}_{\mathcal{V}^*} \{-D_z \mathcal{I}(\hat{t}(\sigma), \hat{z}(\sigma)), \partial \mathcal{R}(0)\} d\sigma \\ = \mathcal{I}(0, z_0) + \int_0^s \partial_t \mathcal{I}(\hat{t}(\sigma), \hat{z}(\sigma)) \hat{t}'(\sigma) d\sigma \quad \forall s \in [0, S]. \end{aligned} \quad (2.11)$$

If, in addition to the second inequality in (2.10a), there is a constant  $\delta > 0$  such that  $\hat{t}'(s) + \|\hat{z}'(s)\|_{\mathbb{V}} > \delta$  f.a.a.  $s \in (0, S)$ , then the solution is called *non-degenerate parametrized solution*, otherwise we call it *degenerate parametrized solution*.

We point out that, in case of a non-degenerate parametrized solution, it is always possible to rescale the artificial time in order to obtain a *normalized parametrized solution*, where  $\hat{t}'(s) + \|\hat{z}'(s)\|_{\mathbb{V}} = 1$  f.a.a.  $s \in (0, S)$ .

## 2.3 Assumptions on the discretization

Our assumptions on the (spatial) discretization are as follows: let  $\mathcal{Z}_h \subset \mathcal{Z}$  be a finite dimensional subspace, where  $h > 0$  indicates the fineness of the approximation, and denote by  $\Pi_h : \mathcal{V} \rightarrow \mathcal{Z}_h$  the associated orthogonal projection. Then we assume that  $\Pi_h$  is stable w.r.t. the  $\mathcal{Z}$ -norm, i.e.,

$$\|\Pi_h(z)\|_{\mathcal{Z}} \leq C \|z\|_{\mathcal{Z}} \quad \forall z \in \mathcal{Z} \quad (2.12)$$

with a constant  $C > 0$  independent of  $h$ . Note that, as an immediate consequence, we obtain the best approximation property of the orthogonal projection, i.e.,

$$\|z - \Pi_h(z)\|_{\mathcal{Z}} \leq \inf_{z_h \in \mathcal{Z}_h} (1 + C) \|z - z_h\|_{\mathcal{Z}}. \quad (2.13)$$

The stability assumption in (2.12) is fulfilled in prominent cases such as finite element discretizations based on shape-regular triangulations, as we will see in Section 4.2 below. In the following, we will frequently consider  $\Pi_h$  as an operator in  $\mathcal{V}$  and  $\mathcal{Z}$ , respectively, denoted for simplicity by the same symbol.

We further introduce the Ritz-projection  $P_h : \mathcal{Z} \rightarrow \mathcal{Z}_h$  as unique solution of

$$P_h(u) \in \mathcal{Z}_h, \quad a(P_h(u), v) = a(u, v) \quad \forall v \in \mathcal{Z}_h,$$

where  $a$  is the bilinear form induced by  $A$ . For the initial value of the algorithm, we set  $z_0^{\tau, h} := P_h(z_0) \in \mathcal{Z}_h$ .

Furthermore, it is assumed that, for all  $v \in \mathcal{V}$  and all  $z \in \mathcal{Z}$ , respectively, it holds

$$\Pi_h(v) \rightarrow v \quad \text{in } \mathcal{V} \quad \text{and} \quad P_h(z) \rightarrow z \quad \text{in } \mathcal{Z} \quad (2.14)$$

as  $h \searrow 0$ . Note that the stability property in (2.13) then automatically yields for all  $z \in \mathcal{Z}$  that

$$\Pi_h(z) \rightarrow z \quad \text{in } \mathcal{Z}. \quad (2.15)$$

In addition, the dissipation functional  $\mathcal{R}$  is approximated by a functional  $\mathcal{R}_h : \mathcal{Z}_h \rightarrow [0, \infty)$  satisfying the following conditions:

- (a) Analogously to  $\mathcal{R}$ , its approximation  $\mathcal{R}_h$  is convex, lower semicontinuous, positively homogeneous.
- (b) Furthermore, it satisfies

$$0 \leq \mathcal{R}_h(v_h) \leq \tilde{\rho} \|v_h\|_{\mathcal{V}} \quad \forall v_h \in \mathcal{Z}_h \quad (2.16)$$

with some constant  $\tilde{\rho} > 0$  independent of  $h$ .

- (c) There is a dense subset  $\mathcal{U}$  of  $\mathcal{V}$  such that, for every  $v \in \mathcal{U}$ , there holds  $\mathcal{R}_h(\Pi_h v) \rightarrow \mathcal{R}(v)$  as  $h \searrow 0$ .
- (d) For every  $v_h \in \mathcal{Z}_h$ , it holds  $\mathcal{R}(v_h) \leq \mathcal{R}_h(v_h)$ .

Note that the choice  $\mathcal{R}_h = \mathcal{R}$  (i.e., no additional approximation of  $\mathcal{R}$ ) fulfills all these assumptions. Another example fulfilling all conditions is given in Section 4.2 below.

### 3 Local minimization algorithm

Given a time-discretization parameter  $\tau > 0$ , our fully discrete counterpart of the local minimization algorithm in (1.4) reads as follows:

#### Algorithm 1 (Fully discrete local minimization)

- 1: Set  $z_0^{\tau, h} = P_h(z_0)$ ,  $t_0 = 0$ , and  $k = 1$
- 2: **while**  $t_k < T$  **do**
- 3:   Compute  $z_k^{\tau, h}$  as solution of

$$z_k^{\tau, h} \in \arg \min \{ \mathcal{I}(t_{k-1}^{\tau, h}, z) + \mathcal{R}_h(z - z_{k-1}^{\tau, h}) : z \in \mathcal{Z}_h, \|z - z_{k-1}^{\tau, h}\|_{\mathcal{V}} \leq \tau \} \quad (3.1)$$

- 4:   Time update:

$$t_k^{\tau, h} = \min \{ t_{k-1}^{\tau, h} + \tau - \|z_k^{\tau, h} - z_{k-1}^{\tau, h}\|_{\mathcal{V}}, T \} \quad (3.2)$$

- 5: Set  $k = k + 1$ .  
 6: **end while**

Note that, thanks to our assumptions on  $\mathcal{I}$  and  $\mathcal{R}_h$ , in particular weak lower semicontinuity, the existence of minimizers of (3.1) follows immediately from the direct method in the calculus of variations.

### 3.1 Approximate discrete parameterized solution

From the characterization of  $z_k^{\tau,h}$  as a minimizer of (3.1), we obtain the following necessary optimality conditions, which will serve as foundation for the derivation of a discrete analogon to the energy identity in (2.11):

**Lemma 3.1 (Discrete optimality System)** *Let  $k \geq 1$  and  $z_k^{\tau,h}$  be an arbitrary solution of (3.1) with associated  $t_k^{\tau,h}$  given by (3.2). Then the following optimality properties are satisfied: There exists a Lagrange multiplier  $\lambda_k^{\tau,h} \geq 0$  such that*

$$\lambda_k^{\tau,h} (\|z_k^{\tau,h} - z_{k-1}^{\tau,h}\|_{\mathbb{V}} - \tau) = 0 \quad (3.3)$$

$$\tau \operatorname{dist}_{\mathcal{V}^*} \{-\Pi_h^* D_z \mathcal{I}(t_{k-1}^{\tau,h}, z_k^{\tau,h}), \partial(\mathcal{R}_h \circ \Pi_h)(0)\} = \lambda_k^{\tau,h} \|z_k^{\tau,h} - z_{k-1}^{\tau,h}\|_{\mathbb{V}}^2 \quad (3.4)$$

$$\left. \begin{aligned} \mathcal{R}_h(z_k^{\tau,h} - z_{k-1}^{\tau,h}) + \tau \operatorname{dist}_{\mathcal{V}^*} \{-\Pi_h^* D_z \mathcal{I}(t_{k-1}^{\tau,h}, z_k^{\tau,h}), \partial(\mathcal{R}_h \circ \Pi_h)(0)\} \\ = \langle -D_z \mathcal{I}(t_{k-1}^{\tau,h}, z_k^{\tau,h}), z_k^{\tau,h} - z_{k-1}^{\tau,h} \rangle_{\mathcal{Z}^*, \mathcal{Z}} \end{aligned} \right\} \quad (3.5)$$

$$\mathcal{R}_h(\Pi_h(v)) \geq -\langle \lambda_k^{\tau,h} \mathbb{V}(z_k^{\tau,h} - z_{k-1}^{\tau,h}) + \Pi_h^* D_z \mathcal{I}(t_{k-1}^{\tau,h}, z_k^{\tau,h}), v \rangle_{\mathcal{V}^*, \mathcal{V}} \quad \forall v \in \mathcal{V}. \quad (3.6)$$

*Proof* To shorten the notation, we suppress the superscripts  $\tau, h$  throughout the proof. We first note that the solution sets of the two problems

$$(P_1) \quad z_k \in \arg \min \{\mathcal{I}(t_{k-1}, z) + \mathcal{R}_h(z - z_{k-1}) : z \in \mathcal{Z}_h, \|z - z_{k-1}\|_{\mathbb{V}} \leq \tau\}$$

$$(P_2) \quad \tilde{z}_k \in \Pi_h \left( \arg \min \{ \mathcal{I}(t_{k-1}, \Pi_h(z)) + \mathcal{R}_h(\Pi_h(z) - z_{k-1}) : z \in \mathcal{V}, \|z - z_{k-1}\|_{\mathbb{V}} \leq \tau \} \right)$$

are in fact the same. For this reason, observe that, by conformity  $\mathcal{Z}_h \subset \mathcal{Z} \subset \mathcal{V}$ , any  $z \in \mathcal{Z}_h$  that is admissible in (P<sub>1</sub>) is also feasible for (P<sub>2</sub>). On the other hand, the non-expansivity of the projection implies that a solution  $\tilde{z}_k$  of (P<sub>2</sub>) is also feasible for (P<sub>1</sub>). Moreover, since the restriction of  $\Pi_h$  to  $\mathcal{Z}_h$  just equals the identity, the objective functionals in (P<sub>1</sub>) and (P<sub>2</sub>) coincide on  $\mathcal{Z}_h$ , which finally guarantees that the two problems provide the same solutions sets. With this result at hand, the proof now follows the lines of [Kne17], where the local minimization algorithm without additional spatial discretization is discussed. For convenience of the reader, we present the arguments in detail.

Thus, an arbitrary solution  $z_k$  of (P<sub>1</sub>) also satisfies the necessary optimality conditions of (P<sub>2</sub>), which read

$$0 \in \partial \mathcal{R}_{\tau,h}(z_k - z_{k-1}) + \Pi_h^* D_z \mathcal{I}(t_{k-1}, z_k). \quad (3.7)$$

Herein we set  $\mathcal{R}_{\tau,h} := \mathcal{R}_h \circ \Pi_h + I_\tau$ , where  $I_\tau : \mathcal{V} \rightarrow [0, \infty]$  denotes the indicator functional associated with the constraints in (P<sub>2</sub>), i.e.,

$$I_\tau(v) := \begin{cases} 0, & \text{if } \langle \mathbb{V}v, v \rangle \leq \tau^2 \\ \infty, & \text{else.} \end{cases} \quad (3.8)$$

Note that (3.7) is an equation in  $\mathcal{V}^*$  and not in  $\mathcal{Z}_h^*$ , which is the essential advantage of considering (P<sub>2</sub>) instead of (P<sub>1</sub>). Now, thanks to a classical result of convex analysis, (3.7) is equivalent to

$$\begin{aligned} & \mathcal{R}_{\tau,h}(z_k - z_{k-1}) + \mathcal{R}_{\tau,h}^*(-\Pi_h^* D_z \mathcal{I}(t_{k-1}, z_k)) \\ &= \langle -\Pi_h^* D_z \mathcal{I}(t_{k-1}, z_k), z_k - z_{k-1} \rangle_{\mathcal{V}^*, \mathcal{V}} \\ &= \langle -D_z \mathcal{I}(t_{k-1}, z_k), \Pi_h(z_k - z_{k-1}) \rangle_{\mathcal{Z}^*, \mathcal{Z}} \\ &= \langle -D_z \mathcal{I}(t_{k-1}, z_k), z_k - z_{k-1} \rangle_{\mathcal{Z}^*, \mathcal{Z}}. \end{aligned} \quad (3.9)$$

Since  $\Pi_h(z_k - z_{k-1}) = z_k - z_{k-1}$  and  $\|z_k - z_{k-1}\|_{\mathbb{V}} \leq \tau$ , we obtain

$$\mathcal{R}_{\tau,h}(z_k - z_{k-1}) = \mathcal{R}_h(z_k - z_{k-1}). \quad (3.10)$$

From Lemma C.2 we moreover infer

$$\mathcal{R}_{\tau,h}^*(-\Pi_h^* D_z \mathcal{I}(t_{k-1}^{\tau,h}, z_k^{\tau,h})) = \tau \operatorname{dist}_{\mathcal{V}^*} \{-\Pi_h^* D_z \mathcal{I}(t_{k-1}^{\tau,h}, z_k^{\tau,h}), \Pi_h^* \partial \mathcal{R}_h(0)\}.$$

Inserting this together with (3.10) in (3.9) gives (3.5).

To prove (3.3), we consider (3.7) again. Since  $\mathcal{R}_h \circ \Pi_h$  is continuous, we are allowed to apply the sum rule for convex subdifferentials giving the existence of a  $\zeta_k \in \partial I_{\tau}(z_k - z_{k-1})$  such that

$$0 \in \partial(\mathcal{R}_h \circ \Pi_h)(z_k - z_{k-1}) + \zeta_k + \Pi_h^* D_z \mathcal{I}(t_{k-1}, z_k) \quad (3.11)$$

and thus

$$\begin{aligned} & \mathcal{R}_h(z_k - z_{k-1}) + (\mathcal{R}_h \circ \Pi_h)^*(-\zeta_k - \Pi_h^* D_z \mathcal{I}(t_{k-1}, z_k)) \\ &= -\langle \zeta_k + \Pi_h^* D_z \mathcal{I}(t_{k-1}, z_k), z_k - z_{k-1} \rangle_{\mathcal{V}^*, \mathcal{V}} \\ &= -\langle \zeta_k, z_k - z_{k-1} \rangle_{\mathcal{V}^*, \mathcal{V}} - \langle D_z \mathcal{I}(t_{k-1}, z_k), z_k - z_{k-1} \rangle_{\mathcal{Z}^*, \mathcal{Z}}. \end{aligned}$$

by comparing this with (3.5), we arrive at

$$\begin{aligned} & (\mathcal{R}_h \circ \Pi_h)^*(-\zeta_k - \Pi_h^* D_z \mathcal{I}(t_{k-1}, z_k)) \\ &= \tau \operatorname{dist}_{\mathcal{V}^*} \{-\Pi_h^* D_z \mathcal{I}(t_{k-1}, z_k), \Pi_h^* \partial \mathcal{R}_h(0)\} - \langle \zeta_k, z_k - z_{k-1} \rangle_{\mathcal{V}^*, \mathcal{V}}. \end{aligned} \quad (3.12)$$

Now, since  $\zeta_k \in \partial I_{\tau}(z_k - z_{k-1})$ , Lemma C.3 implies the existence of a multiplier  $\lambda_k \in \mathbb{R}$  so that

$$\lambda_k \geq 0, \quad \zeta_k = \lambda_k \mathbb{V}(z_k - z_{k-1}), \quad \lambda_k (\|z_k - z_{k-1}\|_{\mathbb{V}} - \tau) = 0. \quad (3.13)$$

which is just (3.3).

Next we verify (3.4). For this purpose, first observe that Lemma C.1 applied to  $\mathcal{J} := \mathcal{R}_h \circ \Pi_h$  (cf. Remark 4) in combination with (3.11) gives

$$-\zeta_k - \Pi_h^* D_z \mathcal{I}(t_{k-1}, z_k) \in \partial(\mathcal{R}_h \circ \Pi_h)(z_k - z_{k-1}) \subset \partial(\mathcal{R}_h \circ \Pi_h)(0) \quad (3.14)$$

$$\implies (\mathcal{R}_h \circ \Pi_h)^*(-\zeta_k - \Pi_h^* D_z \mathcal{I}(t_{k-1}, z_k)) = 0. \quad (3.15)$$

Inserting this and the second equation in (3.13) into (3.12) then yields (3.4).

Finally, (3.6) directly follows from (3.14), i.e.,

$$\mathcal{R}_h(\Pi_h(v)) \geq -\langle \zeta_k + \Pi_h^* D_z \mathcal{I}(t_{k-1}, z_k), v \rangle \quad \forall v \in \mathcal{V}$$

and the characterization of  $\zeta_k$  in (3.13).  $\square$

Let us take a further look at (3.11). In combination with the characterization of  $\zeta_k^{\tau,h}$  in (3.13), this yields

$$0 \in \partial(\mathcal{R}_h \circ \Pi_h)(z_k^{\tau,h} - z_{k-1}^{\tau,h}) + \lambda_k^{\tau,h} \mathbb{V}(z_k^{\tau,h} - z_{k-1}^{\tau,h}) + \Pi_h^* D_z \mathcal{I}(t_{k-1}^{\tau,h}, z_k^{\tau,h}).$$

Since  $\lambda_k^{\tau,h} > 0$  only if the local stability  $0 \in \partial(\mathcal{R}_h \circ \Pi_h)(z_k^{\tau,h} - z_{k-1}^{\tau,h}) + \Pi_h^* D_z \mathcal{I}(t_{k-1}^{\tau,h}, z_k^{\tau,h})$  is violated, we can interpret this inclusion as a discrete version of (1.2). This observation will be taken up in Section 3.3 to obtain, similarly to the continuous case (see [MR15, MRS12]), a discrete version of the energy-equality in (1.3).

### 3.2 A-priori estimates

In the following, we derive several a-priori estimates that will allow a passage to the limit in the discrete energy identity in Section 3.3 and 3.4, respectively. Furthermore, we show that the discrete physical time given the time update in (3.2) reaches the final time  $T$  in a finite number of iterations, see Proposition 3.6 below. We start with the following result, whose proof is completely analogous to the continuous case, cf. e.g. [MR15], and therefore postponed to Appendix A.

**Lemma 3.2 (Boundedness for energy and dissipation)** *For all  $h, \tau > 0$  and all  $k \in \mathbb{N}$ , it holds*

$$\mathcal{I}(t_k^{\tau,h}, z_k^{\tau,h}) + \sum_{i=1}^k \mathcal{R}_h(z_i^{\tau,h} - z_{i-1}^{\tau,h}) \leq (\beta + \mathcal{I}(0, z_0^{\tau,h})) \exp(\mu T), \quad (3.16)$$

where  $\beta$  and  $\mu$  are the constants from Section 2.

**Lemma 3.3 (Uniform a-priori estimate for iterates)** *The iterates of Algorithm 1 fulfill*

$$\sup_{h, \tau > 0, k \in \mathbb{N}} \|z_k^{\tau,h}\|_{\mathcal{Z}} < \infty. \quad (3.17)$$

*Proof* Using  $\mathcal{F} \geq 0$  and the coercivity of  $A$ , Young's inequality implies for every  $t \in [0, T]$  and every  $z \in \mathcal{Z}$  that

$$\begin{aligned} \mathcal{I}(t, z) &= \frac{1}{2} \langle Az, z \rangle + \mathcal{F}(z) - \langle \ell(t), z \rangle \\ &\geq \frac{\alpha}{2} \|z\|_{\mathcal{Z}}^2 - \langle \ell(t), z \rangle \geq \frac{\alpha}{4} \|z\|_{\mathcal{Z}}^2 - c_\alpha \|\ell\|_{C^0([0, T]; \mathcal{V}^*)} \end{aligned}$$

with a constant  $c_\alpha$ , depending only on  $\alpha$  and the embedding constant of  $\mathcal{Z} \hookrightarrow \mathcal{V}$ . Combining this with (3.16) and using  $\mathcal{R}_h \geq 0$ , we arrive at

$$\begin{aligned} \frac{\alpha}{4} \|z_k^{\tau,h}\|_{\mathcal{Z}}^2 &\leq c_\alpha \|\ell\|_{C^0([0, T]; \mathcal{V}^*)}^2 + \mathcal{I}(t_k^{\tau,h}, z_k^{\tau,h}) \\ &\leq c_\alpha \|\ell\|_{C^0([0, T]; \mathcal{V}^*)}^2 + (\mathcal{I}(0, z_0^{\tau,h}) + \beta) \exp(\mu T). \end{aligned}$$

Because of (2.14) and the continuity of  $\mathcal{I}$  by assumption,  $\mathcal{I}(0, z_0^{\tau,h})$  converges to  $\mathcal{I}(0, z_0)$  and is thus bounded, which gives the assertion.  $\square$

**Lemma 3.4** *Let  $z_0 \in \mathcal{Z}$  be such that  $Az_0 \in \mathcal{V}^*$ . Then it holds*

$$\Pi_h^* D_z \mathcal{I}(0, z_0^{\tau,h}) \rightharpoonup D_z \mathcal{I}(0, z_0) \quad \text{in } \mathcal{V}^*, \text{ as } h \searrow 0.$$

*Proof* By assumption, the energy functional is continuously differentiable in  $\mathcal{Z}$  with

$$D_z \mathcal{I}(0, z_0^{\tau, h}) = Az_0^{\tau, h} + D_z \mathcal{F}(z_0^{\tau, h}) + \ell(0) \in \mathcal{Z}^*.$$

We moreover assumed that  $\ell(0) \in \mathcal{V}^*$  and  $D_z \mathcal{F} \in C(\mathcal{Z}; \mathcal{V}^*)$ , cf. (2.2) and (2.5), so that, for every  $v \in \mathcal{V}$ ,

$$\begin{aligned} & \langle \Pi_h^*(D_z \mathcal{F}(z_0^{\tau, h}) + \ell(0)), v \rangle_{\mathcal{V}^*, \mathcal{V}} \\ &= \langle D_z \mathcal{F}(z_0^{\tau, h}) + \ell(0), \Pi_h(v) \rangle_{\mathcal{Z}^*, \mathcal{Z}} \\ &= \langle D_z \mathcal{F}(z_0^{\tau, h}) + \ell(0), \Pi_h(v) \rangle_{\mathcal{V}^*, \mathcal{V}} \rightarrow \langle D_z \mathcal{F}(z_0) + \ell(0), v \rangle_{\mathcal{V}^*, \mathcal{V}}, \end{aligned} \quad (3.18)$$

where we used that  $z_0^{\tau, h} = P_h(z_0) \rightarrow z_0$  in  $\mathcal{Z}$  and  $\Pi_h(v) \rightarrow v$  in  $\mathcal{V}$  by (2.14). Moreover, the definition of the Ritz-projection and  $Az_0 \in \mathcal{V}^*$  by assumption imply for every  $v \in \mathcal{V}$  that

$$\begin{aligned} \langle \Pi_h^*(Az_0^{\tau, h}), v \rangle_{\mathcal{V}^*, \mathcal{V}} &= \langle AP_h(z_0), \Pi_h v \rangle_{\mathcal{Z}^*, \mathcal{Z}} = \langle Az_0, \Pi_h v \rangle_{\mathcal{Z}^*, \mathcal{Z}} \\ &= \langle Az_0, \Pi_h v \rangle_{\mathcal{V}^*, \mathcal{V}} \rightarrow \langle Az_0, v \rangle_{\mathcal{V}^*, \mathcal{V}}. \end{aligned}$$

Together with (3.18), this yields the assertion.  $\square$

As indicated in the introduction, one major issue in the convergence analysis for parameterized solutions concerns the boundedness of the artificial time, even in the continuous setting. In terms of the local minimization algorithms, this means one has to show that the physical end time  $T$  is reached after a finite number of iterations. For this purpose, we need the following

**Lemma 3.5** *For every  $r > 0$  and  $\epsilon > 0$ , there exists  $C_{r, \epsilon} > 0$ , independent of  $h$ , such that*

$$|\langle D_z \mathcal{F}(z_1^h) - D_z \mathcal{F}(z_2^h), z_1^h - z_2^h \rangle_{\mathcal{V}^*, \mathcal{V}}| \leq \epsilon \|z_1^h - z_2^h\|_{\mathcal{Z}}^2 + C_{r, \epsilon} \mathcal{R}_h(z_1^h - z_2^h) \|z_1^h - z_2^h\|_{\mathcal{V}}$$

for all  $z_1^h, z_2^h \in \mathcal{Z}_h \cap B_{\mathcal{Z}}(0, r)$ .

*Proof* The proof is analogous to the infinite-dimensional case, one just has to employ Assumption (d) on  $\mathcal{R}_h$  from Section 2.3 at the end. For convenience of the reader, we explain the arguments in detail. According to Ehrling's lemma, for every  $\delta > 0$ , there exists a constant  $C_\delta$  (obviously independent of  $h$ ) such that

$$\|z\|_{\mathcal{V}} \leq \delta \|z\|_{\mathcal{Z}} + C_\delta \|z\|_{\mathcal{X}} \quad \forall z \in \mathcal{Z} \quad (3.19)$$

Now let  $z_1^h, z_2^h \in B_{\mathcal{Z}}(0, r) \cap \mathcal{Z}_h$  be arbitrary. Using the growth condition on  $D_z^2 \mathcal{F}$  in (2.5) and the above inequality for  $\delta = \epsilon / (2C(1 + r^q))$  in combination with Young's inequality gives

$$\begin{aligned} & |\langle D_z \mathcal{F}(z_1^h) - D_z \mathcal{F}(z_2^h), z_1^h - z_2^h \rangle_{\mathcal{V}^*, \mathcal{V}}| \\ & \leq \|D_z \mathcal{F}(z_1^h) - D_z \mathcal{F}(z_2^h)\|_{\mathcal{V}^*} \|z_1^h - z_2^h\|_{\mathcal{V}} \\ & \leq C(1 + r^q) \|z_1^h - z_2^h\|_{\mathcal{Z}} (\delta \|z_1^h - z_2^h\|_{\mathcal{Z}} + C_\delta \|z_1^h - z_2^h\|_{\mathcal{X}}) \\ & \leq \epsilon \|z_1^h - z_2^h\|_{\mathcal{Z}}^2 + \tilde{C}_{r, \epsilon} \|z_1^h - z_2^h\|_{\mathcal{X}}^2 \end{aligned}$$

with a constant  $\tilde{C}_{r, \epsilon}$  depending only on  $\epsilon$  and  $r$ . Now, (2.6) and Assumption (d) on the discretization of  $\mathcal{R}$  result in

$$\|z_1^h - z_2^h\|_{\mathcal{X}} \leq \frac{1}{\rho} \mathcal{R}(z_1^h - z_2^h) \leq \frac{1}{\rho} \mathcal{R}_h(z_1^h - z_2^h),$$

which, together with the embedding  $\mathcal{V} \hookrightarrow \mathcal{X}$ , which completes the proof.  $\square$

Now, we are in the position to prove that the final time  $T$  is reached in a finite number of iterations.

**Proposition 3.6 (Bound on artificial time)** *For every  $h, \tau > 0$ , there exists an index  $N(\tau, h) \in \mathbb{N}$  such that  $t_{N(\tau, h)}^{\tau, h} = T$ . Moreover, there are constants  $C_1, C_2, C_3 > 0$  such that, for all  $h, \tau > 0$ , it holds*

$$\sum_{i=1}^{N(\tau, h)} \|z_i^{\tau, h} - z_{i-1}^{\tau, h}\|_{\mathcal{Z}} \leq C_1, \quad (3.20)$$

$$\lambda_{k+1}^{\tau, h} \|z_{k+1}^{\tau, h} - z_k^{\tau, h}\|_{\mathbb{V}} \leq C_2, \quad \forall k = 0, \dots, N(\tau, h) - 1, \quad (3.21)$$

$$\text{and } \|\Pi_h^* D_z \mathcal{I}(t_k^{\tau, h}, z_{k+1}^{\tau, h})\|_{\mathcal{V}^*} \leq C_3 \quad \forall k = 0, \dots, N(\tau, h) - 1. \quad (3.22)$$

*Proof* The arguments are similar to [Kne17], but, since one has to take account of the discretization at several points, we depict the proof in detail. Let  $k \in \mathbb{N}$  be arbitrary. For convenience, we again suppress the superscript  $\tau, h$  throughout the proof, except for  $z_0^{\tau, h}$  in order to avoid confusion with the initial data. Testing (3.6) with  $v = z_{k+1} - z_k$  yields

$$\begin{aligned} \mathcal{R}_h(z_{k+1} - z_k) &\geq -\langle \lambda_k \mathbb{V}(z_k - z_{k-1}) + \Pi_h^* D_z \mathcal{I}(t_{k-1}, z_k), z_{k+1} - z_k \rangle_{\mathcal{V}^*, \mathcal{V}} \\ &= -\langle \lambda_k \mathbb{V}(z_k - z_{k-1}), z_{k+1} - z_k \rangle_{\mathcal{V}^*, \mathcal{V}} - \langle D_z \mathcal{I}(t_{k-1}, z_k), z_{k+1} - z_k \rangle_{\mathcal{Z}^*, \mathcal{Z}}. \end{aligned} \quad (3.23)$$

Inserting (3.4) into (3.5) and rewriting this identity for the index  $k+1$  (instead of  $k$ ) gives

$$\mathcal{R}_h(z_{k+1} - z_k) + \lambda_{k+1} \|z_{k+1} - z_k\|^2 = \langle -D_z \mathcal{I}(t_k, z_{k+1}), z_{k+1} - z_k \rangle_{\mathcal{Z}^*, \mathcal{Z}}.$$

Subtracting this from (3.23) leads to

$$\begin{aligned} 0 &\geq \lambda_{k+1} \|z_{k+1} - z_k\|_{\mathbb{V}}^2 - \lambda_k \langle \mathbb{V}(z_k - z_{k-1}), z_{k+1} - z_k \rangle_{\mathcal{V}^*, \mathcal{V}} \\ &\quad + \langle D_z \mathcal{I}(t_k, z_{k+1}) - D_z \mathcal{I}(t_{k-1}, z_k), z_{k+1} - z_k \rangle_{\mathcal{Z}^*, \mathcal{Z}}. \end{aligned}$$

With this inequality at hand, we can now follow the lines of [Kne17, Proposition 2.3]: On account of  $\ell(\cdot) \in \mathcal{V}^*$  and  $D_z \mathcal{F}(\cdot) \in \mathcal{V}^*$  by assumption, inserting the definition of  $\mathcal{I}$  into this inequality gives

$$\begin{aligned} &\langle D_z \mathcal{F}(z_k) - D_z \mathcal{F}(z_{k+1}), z_{k+1} - z_k \rangle_{\mathcal{V}^*, \mathcal{V}} + \langle \ell(t_{k-1}) - \ell(t_k), z_{k+1} - z_k \rangle_{\mathcal{V}^*, \mathcal{V}} \\ &\geq \lambda_{k+1} \|z_{k+1} - z_k\|_{\mathbb{V}}^2 - \lambda_k \langle \mathbb{V}(z_k - z_{k-1}), z_{k+1} - z_k \rangle_{\mathcal{V}^*, \mathcal{V}} \\ &\quad + \langle A(z_{k+1} - z_k), z_{k+1} - z_k \rangle_{\mathcal{Z}^*, \mathcal{Z}} \\ &\geq \lambda_{k+1} \|z_{k+1} - z_k\|_{\mathbb{V}}^2 - \lambda_k \|z_k - z_{k-1}\|_{\mathbb{V}} \|z_{k+1} - z_k\|_{\mathbb{V}} + \alpha \|z_{k+1} - z_k\|_{\mathcal{Z}}^2, \end{aligned} \quad (3.24)$$

where we used the coercivity of  $A$  for the last estimate. We now estimate the terms on the left hand side of (3.24) separately. For the first term, we apply Lemma 3.5 with  $\varepsilon = \alpha/2$ . Note that, by Lemma 3.3, there is an  $r > 0$  so that  $z_k \in B_{\mathcal{Z}}(0, r)$  for all  $k \in \mathbb{N}$  giving in turn that Lemma 3.5 is indeed applicable and yields

$$\begin{aligned} &\langle D_z \mathcal{F}(z_k) - D_z \mathcal{F}(z_{k+1}), z_{k+1} - z_k \rangle_{\mathcal{V}^*, \mathcal{V}} \\ &\leq \frac{\alpha}{2} \|z_{k+1} - z_k\|_{\mathcal{Z}}^2 + C_\alpha \|z_{k+1} - z_k\|_{\mathbb{V}} \mathcal{R}_h(z_{k+1} - z_k), \end{aligned} \quad (3.25)$$

with a constant  $C_\alpha > 0$ , which is independent of  $\tau, h$ , and  $k$ . The second term on the left hand side of (3.24) is estimated by

$$\langle \ell(t_{k-1}) - \ell(t_k), z_{k+1} - z_k \rangle \leq \|\ell\|_{C^1([0, T]; \mathcal{V}^*)} (t_k - t_{k-1}) \|z_{k+1} - z_k\|_{\mathcal{V}}. \quad (3.26)$$

Inserting (3.25) and (3.26) in (3.24) gives

$$\begin{aligned} & \lambda_{k+1} \|z_{k+1} - z_k\|_{\mathbb{V}}^2 - \lambda_k \|z_k - z_{k-1}\|_{\mathbb{V}} \|z_{k+1} - z_k\|_{\mathbb{V}} + \frac{\alpha}{2} \|z_{k+1} - z_k\|_{\mathcal{Z}}^2 \\ & \leq C \|z_{k+1} - z_k\|_{\mathbb{V}} \mathcal{R}_h(z_{k+1} - z_k) + \|\ell\|_{C^1([0,T];\mathcal{V}^*)} (t_k - t_{k-1}) \|z_{k+1} - z_k\|_{\mathbb{V}}, \end{aligned}$$

which, thanks to the continuous embedding  $\mathcal{Z} \hookrightarrow \mathcal{V}$  and the norm equivalence of  $\|\cdot\|_{\mathbb{V}}$  and  $\|\cdot\|_{\mathcal{Z}}$ , in turn implies

$$\begin{aligned} & \lambda_{k+1}^{\tau,h} \|z_{k+1} - z_k\|_{\mathbb{V}} - \lambda_k^{\tau,h} \|z_k - z_{k-1}\|_{\mathbb{V}} + c \|z_{k+1} - z_k\|_{\mathcal{Z}} \\ & \leq C (\mathcal{R}_h(z_{k+1} - z_k) + (t_k - t_{k-1})). \end{aligned}$$

Summing up this estimate with respect to  $k$  then yields

$$\begin{aligned} & \lambda_{k+1} \|z_{k+1} - z_k\|_{\mathbb{V}} + c \sum_{i=1}^k \|z_{i+1} - z_i\|_{\mathcal{Z}} \\ & \leq \lambda_1 \|z_1 - z_0^{\tau,h}\|_{\mathbb{V}} + C \left( t_k + \sum_{i=1}^k \mathcal{R}_h(z_{i+1} - z_i) \right). \end{aligned} \quad (3.27)$$

Thanks to (3.16), this inequality already nearly gives (3.21), provided that  $\lambda_1 \|z_1 - z_0^{\tau,h}\|_{\mathbb{V}}$  is bounded independent of  $\tau$  and  $h$ , which is shown next. For this purpose, we again insert (3.4) into (3.5) to obtain for  $k = 1$ :

$$\mathcal{R}_h(z_1 - z_0) + \lambda_1 \|z_1 - z_0^{\tau,h}\|_{\mathbb{V}}^2 = \langle -D_z \mathcal{I}(0, z_1), z_1 - z_0^{\tau,h} \rangle_{\mathcal{Z}^*, \mathcal{Z}}.$$

Adding a zero, using  $\mathcal{R}_h \geq 0$ , and rearranging terms yield

$$\begin{aligned} & \langle D_z \mathcal{I}(0, z_1) - D_z \mathcal{I}(0, z_0^{\tau,h}), z_1 - z_0^{\tau,h} \rangle_{\mathcal{Z}^*, \mathcal{Z}} + \lambda_1 \|z_1 - z_0^{\tau,h}\|_{\mathbb{V}}^2 \\ & \leq \langle -D_z \mathcal{I}(0, z_0^{\tau,h}), z_1 - z_0^{\tau,h} \rangle_{\mathcal{Z}^*, \mathcal{Z}} = \langle -\Pi_h^* D_z \mathcal{I}(0, z_0^{\tau,h}), z_1 - z_0^{\tau,h} \rangle_{\mathcal{V}^*, \mathcal{V}}. \end{aligned}$$

The first term on the left hand side is treated completely analogously to above by employing Lemma 3.5 resulting in

$$\begin{aligned} & \frac{\alpha}{2} \|z_1 - z_0^{\tau,h}\|_{\mathcal{Z}}^2 + \lambda_1 \|z_1 - z_0^{\tau,h}\|_{\mathbb{V}}^2 \\ & \leq C \mathcal{R}_h(z_1 - z_0^{\tau,h}) \|z_1 - z_0^{\tau,h}\|_{\mathbb{V}} + \langle -\Pi_h^* D_z \mathcal{I}(0, z_0^{\tau,h}), z_1 - z_0^{\tau,h} \rangle_{\mathcal{V}^*, \mathcal{V}}. \end{aligned}$$

Using again  $\mathcal{Z} \hookrightarrow \mathcal{V}$  and the norm equivalence of  $\|\cdot\|_{\mathcal{Z}}$  and  $\|\cdot\|_{\mathbb{V}}$ , it follows

$$\lambda_1 \|z_1 - z_0^{\tau,h}\|_{\mathbb{V}} + c \|z_1 - z_0^{\tau,h}\|_{\mathcal{Z}} \leq C (\mathcal{R}_h(z_1 - z_0^{\tau,h}) + \|\Pi_h^* D_z \mathcal{I}(0, z_0^{\tau,h})\|_{\mathcal{V}^*}). \quad (3.28)$$

From Lemma 3.4, we know that  $\Pi_h^* D_z \mathcal{I}(0, z_0^{\tau,h})$  converges weakly in  $\mathcal{V}^*$  and is thus bounded. Together with (3.16), this gives the desired boundedness of  $\lambda_1 \|z_1 - z_0^{\tau,h}\|_{\mathbb{V}}$ , i.e., (3.21) for  $k = 0$ . By adding (3.28) to (3.27), we obtain

$$\begin{aligned} & \lambda_{k+1} \|z_{k+1} - z_k\|_{\mathbb{V}} + c \sum_{i=0}^k \|z_{i+1} - z_i\|_{\mathcal{Z}} \\ & \leq C \left( t_k + \sum_{i=0}^k \mathcal{R}_h(z_{i+1} - z_i) + \|\Pi_h^* D_z \mathcal{I}(0, z_0^{\tau,h})\|_{\mathcal{V}^*} \right) \\ & \leq C \left( T + (\mathcal{I}(0, z_0^{\tau,h}) + \beta) \exp(\mu T) + \|\Pi_h^* D_z \mathcal{I}(0, z_0^{\tau,h})\|_{\mathcal{V}^*} \right), \end{aligned}$$

where we used (3.2) and (3.16) for the last estimate. As already seen at the end of the proof of Lemma 3.3,  $\mathcal{I}(0, z_0^{\tau, h})$  is bounded independent of  $h$ . Thus, employing again the boundedness of  $\|II_h^* D_z \mathcal{I}(0, z_0^{\tau, h})\|_{\mathcal{V}^*}$  gives

$$\lambda_{k+1} \|z_{k+1} - z_k\|_{\mathbb{V}} + c \sum_{i=0}^k \|z_{i+1} - z_i\|_{\mathcal{Z}} \leq C, \quad (3.29)$$

i.e., (3.21) for  $k \geq 1$ . Note that the constant  $C$  is independent of  $\tau$ ,  $h$ , and  $k$ .

We proceed with showing (3.22). To this end, we first note that, since  $\|z_{k+1} - z_k\|_{\mathbb{V}} \leq \tau$  by (3.1), the identity (3.4) implies

$$\text{dist}_{\mathcal{V}^*} \{-II_h^* D_z \mathcal{I}(t_{k-1}, z_k), \partial(\mathcal{R}_h \circ \Pi_h)(0)\} \leq \lambda_k \|z_k - z_{k-1}\|_{\mathbb{V}}. \quad (3.30)$$

Because of

$$\begin{aligned} \xi \in \partial(\mathcal{R}_h \circ \Pi_h)(0) &\implies \langle \xi, v \rangle \leq \mathcal{R}_h(\Pi_h(v)) \leq \tilde{\rho} \|v\|_{\mathcal{V}} \quad \forall v \in \mathcal{V} \\ &\implies \|\xi\|_{\mathcal{V}^*} \leq \tilde{\rho}, \end{aligned} \quad (3.31)$$

the subdifferential  $\partial(\mathcal{R}_h \circ \Pi_h)(0)$  is bounded by  $\tilde{\rho}$  so that

$$\|II_h^* D_z \mathcal{I}(t_{k-1}, z_k)\|_{\mathcal{V}^*} \leq \|II_h^* D_z \mathcal{I}(t_{k-1}, z_k) - w\|_{\mathcal{V}^*} + \tilde{\rho} \quad \forall w \in \partial(\mathcal{R}_h \circ \Pi_h)(0).$$

Hence, (3.30) yields

$$\begin{aligned} \|II_h^* D_z \mathcal{I}(t_{k-1}, z_k)\|_{\mathcal{V}^*} &\leq \text{dist}_{\mathcal{V}^*} \{-II_h^* D_z \mathcal{I}(t_{k-1}, z_k), \partial(\mathcal{R}_h \circ \Pi_h)(0)\} + \tilde{\rho} \\ &\leq \lambda_k \|z_k - z_{k-1}\|_{\mathbb{V}} + \tilde{\rho} \end{aligned}$$

and consequently (3.22) follows from (3.21).

Finally, assume that  $T$  is not reached after a finite number of steps. Thanks to (3.29), the series  $\sum_{i=0}^{\infty} \|z_{i+1} - z_i\|_{\mathbb{V}}$  then converges and, by definition of  $t_k$  in (3.2), there exists  $t_* \leq T$  such that  $\lim_{k \rightarrow \infty} t_k = t_*$ . In particular, the sequence  $(t_{k+1} - t_k)_{k \in \mathbb{N}}$  converges to zero. But, in view of (3.2), this implies that  $(\|z_{k+1} - z_k\|_{\mathbb{V}})_{k \in \mathbb{N}}$  tends to  $\tau$ , in contradiction to the convergence of the series  $\sum_{k=0}^{\infty} \|z_{k+1} - z_k\|_{\mathbb{V}}$ . Lastly, since (3.29) holds for every  $k$ , we obtain (3.20).  $\square$

In the following we will abbreviate the index  $N(\tau, h)$  simply by  $N$  having mind that the number  $N$  of timesteps always depends on  $\tau$  and  $h$

### 3.3 Discrete energy-equality

The goal of this section is to derive a discrete analogon to the energy identity (2.11). For this purpose, we introduce the piecewise affine and the left and right continuous piecewise constant interpolants associated with the iterates  $z_k^{\tau, h}$ . As depicted in Sections 1 and 2, in the setting of parametrized solutions, the potential discontinuities of the solution are resolved by introducing an artificial time and interpret the physical time as a function thereof. This is also reflected by the time-incremental local minimization scheme, where the artificial time is simply defined by means of the time step size  $\tau$ . To be more precise, we set

$$s_k^{\tau, h} := \min\{k\tau, S_{\tau, h}\}, \quad \text{where} \quad S_{\tau, h} := T + \sum_{i=1}^N \|z_i^{\tau, h} - z_{i-1}^{\tau, h}\|_{\mathbb{V}} \quad (3.32)$$

From (3.2) and Proposition 3.6, it follows that  $s_k^{\tau,h} = k\tau$  for  $k < N$ , while  $s_N^{\tau,h} = S_{\tau,h}$ . Moreover, (3.20) and  $\mathcal{Z} \hookrightarrow \mathcal{V}$  immediately give that

$$S_{\tau,h} \leq C_s \quad (3.33)$$

with a constant  $C_s > 0$ , independent of  $\tau$  and  $h$  so that the artificial time interval is indeed bounded independent of the discretization. Now given the artificial time, we define the interpolants as follows: For  $s \in [s_{k-1}^{\tau,h}, s_k^{\tau,h}) \subset [0, S_{\tau,h})$ , the continuous and piecewise affine interpolants are defined through

$$\begin{aligned} \hat{z}_{\tau,h}(s) &:= z_{k-1}^{\tau,h} + \frac{(s - s_{k-1}^{\tau,h})}{s_k^{\tau,h} - s_{k-1}^{\tau,h}} (z_k^{\tau,h} - z_{k-1}^{\tau,h}), \\ \hat{t}_{\tau,h}(s) &:= t_{k-1}^{\tau,h} + \frac{(s - s_{k-1}^{\tau,h})}{s_k^{\tau,h} - s_{k-1}^{\tau,h}} (t_k^{\tau,h} - t_{k-1}^{\tau,h}), \end{aligned} \quad (3.34)$$

while the piecewise constant interpolants are given by

$$\bar{z}_{\tau,h}(s) := z_k^{\tau,h}, \quad \bar{t}_{\tau,h}(s) := t_k^{\tau,h}, \quad \underline{z}_{\tau,h}(s) := z_{k-1}^{\tau,h}, \quad \underline{t}_{\tau,h}(s) := t_{k-1}^{\tau,h}. \quad (3.35)$$

Since the artificial final time  $S_{\tau,h}$  depends on the chosen discretization level, we extend all interpolants constantly onto  $[0, \tilde{S}]$  with  $\tilde{S} := \sup_{\tau,h} S_{\tau,h}$  by

$$\left. \begin{aligned} \bar{z}_{\tau,h}(s) = \underline{z}_{\tau,h}(s) = \hat{z}_{\tau,h}(s) &:= z_N^{\tau,h} \\ \text{and } \bar{t}_{\tau,h}(s) = \underline{t}_{\tau,h}(s) = \hat{t}_{\tau,h}(s) &:= T \end{aligned} \right\} \quad \forall s \in [S_{\tau,h}, \tilde{S}]. \quad (3.36)$$

Observe that  $\tilde{S} \leq C_s$  by (3.33). By construction, we immediately have  $(\hat{t}_{\tau,h}, \hat{z}_{\tau,h}) \in W^{1,\infty}([0, \tilde{S}], \mathbb{R} \times \mathcal{V})$ , but we even obtain pointwise bounds for the time derivatives independent of the discretization:

**Lemma 3.7 (Properties of affine interpolants)** *For almost all  $s \in [0, S_{\tau,h}]$ , the affine interpolants from (3.34) fulfill*

$$\hat{t}'_{\tau,h}(s) \geq 0, \quad \hat{t}'_{\tau,h}(s) + \|\hat{z}'_{\tau,h}(s)\|_{\mathcal{V}} = 1 \quad (3.37)$$

$$\lambda_k^{\tau,h} (1 - \|\hat{z}'_{\tau,h}(s)\|_{\mathcal{V}}) = 0 \quad (3.38)$$

for almost every  $s \in [0, S_{\tau,h}]$ .

*Proof* The first statement in (3.37) is a direct consequence of (3.2), which immediately implies  $t_k^{\tau,h} - t_{k-1}^{\tau,h} \geq 0$ . To prove the second one, first consider the case  $k < N$ . Then  $s_k^{\tau,h} = k\tau$  and therefore, (3.2) implies for every  $s \in [s_{k-1}^{\tau,h}, s_k^{\tau,h})$  that

$$\hat{t}'_{\tau,h}(s) = \frac{(t_k^{\tau,h} - t_{k-1}^{\tau,h})}{s_k^{\tau,h} - s_{k-1}^{\tau,h}} = 1 - \frac{\|z_k^{\tau,h} - z_{k-1}^{\tau,h}\|_{\mathcal{V}}}{\tau} = 1 - \|\hat{z}'_{\tau,h}(s)\|_{\mathcal{V}},$$

which is the assertion for  $k < N$ . In case of  $s \in [s_{N-1}^{\tau,h}, S_{\tau,h})$ , we obtain in view of (3.32) and  $t_{N-1}^{\tau,h} = s_{N-1}^{\tau,h} - \sum_{i=1}^{N-1} \|z_i^{\tau,h} - z_{i-1}^{\tau,h}\|_{\mathcal{V}}$ , which follows from (3.2) and  $s_{N-1}^{\tau,h} = (N-1)\tau$ , that

$$\hat{t}'_{\tau,h}(s) + \|\hat{z}'_{\tau,h}(s)\|_{\mathcal{V}} = \frac{T - t_{N-1}^{\tau,h}}{S_{\tau,h} - s_{N-1}^{\tau,h}} + \frac{\|z_N^{\tau,h} - z_{N-1}^{\tau,h}\|_{\mathcal{V}}}{S_{\tau,h} - s_{N-1}^{\tau,h}} = 1.$$

Together with the above, this confirms (3.37).

For  $s \in [s_{k-1}^{\tau,h}, s_k^{\tau,h})$  with  $k < N$ , the complementarity in (3.38) is a direct consequence of (3.3), because

$$0 = \lambda_k^{\tau,h} (\|z_k^{\tau,h} - z_{k-1}^{\tau,h}\|_{\mathbb{V}} - \tau) = \tau \lambda_k^{\tau,h} (1 - \|\hat{z}'_{\tau,h}(s)\|_{\mathbb{V}}).$$

For  $s \in [s_{N-1}^{\tau,h}, S_{\tau,h})$ , this complementarity follows from  $\lambda_N^{\tau,h} = 0$ , which is valid, since otherwise, by (3.3),  $\|z_N^{\tau,h} - z_{N-1}^{\tau,h}\|_{\mathbb{V}} = \tau$  so that (3.2) gives  $t_N^{\tau,h} = t_{N-1}^{\tau,h} < T$  in contradiction to the first assertion of Proposition 3.6.  $\square$

*Remark 1* From Lemma 3.7 and Lemma 3.3, we deduce the existence of a constant  $C > 0$ , independent of  $\tau$  and  $h$ , so that

$$\|\hat{t}_{\tau,h}\|_{W^{1,\infty}((0,\bar{S}))} + \|\hat{z}_{\tau,h}\|_{W^{1,\infty}((0,\bar{S});\mathbb{V})} + \|\hat{z}_{\tau,h}\|_{L^\infty((0,\bar{S});\mathbb{Z})} \leq C.$$

These a-priori bounds will be essential to pass to the limit in the discrete energy identity, which is derived next.

We are now in the position to show a discrete version of the energy-equality. Its proof is based on Lemma 3.7, the a-priori estimates derived in Section 3.2, and Assumption (a) on the discretization of  $\mathcal{R}$ , which essentially ensures that  $\mathcal{R}_h$  has the same properties as  $\mathcal{R}$ . Therefore, the proof is completely analogous to the continuous case and thus postponed to Appendix B.

**Lemma 3.8 (Discrete energy-equality)** *For all  $\sigma_1, \sigma_2 \in [0, S_{\tau,h}]$  with  $\sigma_1 \leq \sigma_2$ , it holds*

$$\begin{aligned} & \mathcal{I}(\hat{t}_{\tau,h}(\sigma_2), \hat{z}_{\tau,h}(\sigma_2)) \\ & + \int_{\sigma_1}^{\sigma_2} \mathcal{R}_h(\hat{z}'_{\tau,h}) + \text{dist}_{\mathbb{V}^*} \{-\Pi_h^* D_z \mathcal{I}(t_{\tau,h}(s), \bar{z}_{\tau,h}(s)), \partial(\mathcal{R}_h \circ \Pi_h)(0)\} ds \\ & = \mathcal{I}(\hat{t}_{\tau,h}(\sigma_1), \hat{z}_{\tau,h}(\sigma_1)) \\ & + \int_{\sigma_1}^{\sigma_2} \partial_t \mathcal{I}(\hat{t}_{\tau,h}(s), \hat{z}_{\tau,h}(s)) \hat{t}'_{\tau,h}(s) ds + \int_{\sigma_1}^{\sigma_2} r_{\tau,h}(s) ds, \end{aligned} \quad (3.39)$$

where

$$r_{\tau,h}(s) := \langle D_z \mathcal{I}(\hat{t}_{\tau,h}(s), \hat{z}_{\tau,h}(s)) - D_z \mathcal{I}(t_{\tau,h}(s), \bar{z}_{\tau,h}(s)), \hat{z}'_{\tau,h}(s) \rangle_{\mathbb{Z}^*, \mathbb{Z}}. \quad (3.40)$$

Moreover, the complementarity condition

$$\hat{t}'_{\tau,h}(s) \text{dist}_{\mathbb{V}^*} \{-\Pi_h^* D_z \mathcal{I}(t_{\tau,h}(s), \bar{z}_{\tau,h}(s)), \partial(\mathcal{R}_h \circ \Pi_h)(0)\} = 0 \quad (3.41)$$

is fulfilled f.a.a.  $s \in (0, S_{\tau,h})$ , and there exists a constant  $C > 0$  such that the remainder  $r_{\tau,h}$  satisfies for all  $h, \tau > 0$  and all  $\sigma_1 \leq \sigma_2 \in [0, S_{\tau,h}]$

$$\int_{\sigma_1}^{\sigma_2} r_{\tau,h}(s) ds \leq C\tau. \quad (3.42)$$

*Remark 2* A comparison of the discrete energy identity in (3.39) and the continuous one in (2.11) shows that the coefficient  $\|\hat{z}'_{\tau,h}\|$  is missing in front of the distance. It would be possible to reformulate the optimality conditions in Lemma 3.1 in a way such that this coefficient would arise in (3.39). This however would complicate the passage to the limit in the next section. As we will see at the end of the proof of Theorem 3.9, (3.39) is sufficient to obtain the desired energy identity in (2.11).

### 3.4 Main Convergence Theorem

We now have everything at hand to prove our main convergence result.

**Theorem 3.9 (Convergence towards parametrized solutions)** *There exists a sequence  $\{\tau_n, h_n\}_{n \in \mathbb{N}} \subset \mathbb{R}_+ \times \mathbb{R}_+$  converging to zero so that the affine interpolants generated by the fully discrete local minimization algorithm 1 and the artificial end time defined in (3.32) satisfy*

$$S_{\tau_n, h_n} \rightarrow S, \quad (3.43)$$

$$\hat{t}_{\tau_n, h_n} \xrightarrow{*} \hat{t} \quad \text{in } W^{1, \infty}((0, S); \mathbb{R}), \quad (3.44)$$

$$\hat{z}_{\tau_n, h_n} \xrightarrow{*} \hat{z} \quad \text{in } W^{1, \infty}((0, S); \mathcal{V}) \cap L^\infty((0, S); \mathcal{Z}), \quad (3.45)$$

$$\hat{z}_{\tau_n, h_n}(s) \rightarrow \hat{z}(s) \quad \text{in } \mathcal{Z} \text{ for every } s \in [0, S] \quad (3.46)$$

and the limit  $(\hat{t}, \hat{z})$  is a parametrized solution in the sense of Definition 2.1.

Moreover, every accumulation point  $(\hat{t}, \hat{z})$  of time incremental sequences in the sense of (3.43)–(3.46) is a parametrized solution.

*Proof* The arguments are similar to the semi-discrete case without a spatial discretization, which is discussed in [Kne17]. However, as we have to include the passage to limit  $h \searrow 0$ , we present the proof in detail.

The existence of a (sub-)sequence satisfying (3.43)–(3.45) is an immediate consequence of the uniform estimates in Lemma 3.3, Lemma 3.7, and (3.33), cf. also Remark 1. By the Aubin-Lions lemma  $W^{1, \infty}((0, S); \mathcal{V}) \cap L^\infty((0, S); \mathcal{Z})$  compactly embeds in  $C([0, S]; \mathcal{V})$  so that  $\hat{z}_{\tau_n, h_n}$  uniformly converges in  $\mathcal{V}$  to  $\hat{z}$ . However, Lemma 3.3 tells us that, for every  $s \in [0, S]$ ,  $\{\hat{z}_{\tau_n, h_n}(s)\}$  is bounded in  $\mathcal{Z}$  and therefore there is a weakly convergent subsequence. Due to the uniform convergence in  $\mathcal{V}$ , the pointwise limit equals  $\hat{z}(s)$ , which implies (3.46).

It remains to show that every (weak) limit is a parametrized solution. For this purpose, let  $\{\tau_n, h_n\}$  be an arbitrary null sequence and assume that the convergence in (3.43)–(3.46) holds. In order to simplify the notation, we indicate by  $\{\cdot\}_n$  the sequence of  $\{\cdot\}_{\tau, h}$  corresponding to  $\{\tau_n, h_n\}$ . Analogously, we abbreviate the index  $h_n$  simply by  $n$ . We proceed in several steps and start with the following:

*Convergence of piecewise constant interpolants* First we show that the piecewise constant interpolants converge pointwisely to the same limit. We exemplarily consider  $\bar{z}_{\tau, h}$ . Because of (3.17), there is a subsequence, for convenience also denoted by  $\bar{z}_n$ , converging in every  $s \in [0, S]$  weakly in  $\mathcal{Z}$  to some  $\tilde{z}(s)$ . Hence, the compact embedding of  $\mathcal{Z}$  in  $\mathcal{V}$  implies  $\bar{z}_n(s) \rightarrow \tilde{z}(s)$  in  $\mathcal{V}$  for all  $s \in [0, S]$ . Moreover, by (3.34) and (3.35), we have for all  $k \in \{1, \dots, N\}$  and all  $s \in [s_{k-1}^n, s_k^n]$  that

$$\|\hat{z}_n(s) - \bar{z}_n(s)\|_{\mathcal{V}} = |s - s_k^n| \|\hat{z}'_n(s)\|_{\mathcal{V}} \leq \tau \rightarrow 0,$$

where we used (3.37) and (3.33) for the last estimate. Hence, we obtain  $\tilde{z}(s) = \hat{z}(s)$  for all  $s \in [0, S]$  and the uniqueness of the weak limit implies the weak convergence of the whole sequence  $\{\bar{z}_n\}$ . For the other piecewise constant interpolants, one argues completely analogously so that

$$\underline{t}_n(s), \bar{t}_n(s) \rightarrow \hat{t}(s), \quad \underline{z}_n(s), \bar{z}_n(s) \rightarrow \hat{z}(s) \quad \text{in } \mathcal{Z} \quad \forall s \in [0, S] \quad (3.47)$$

is obtained, as desired.

*Initial and end time conditions* Since the Ritz projection trivially fulfills

$$\hat{z}_n(0) = z_0^{\tau, h} = P_n(z_0) \rightarrow z_0 \quad \text{in } \mathcal{Z}, \quad (3.48)$$

the pointwise convergence in (3.46) implies  $\hat{z}(0) = z_0$  as desired. Moreover, thanks to (3.44),  $\hat{t}_n$  converges uniformly to  $\hat{t}$  so that

$$0 = \hat{t}_n(0) \rightarrow \hat{t}(0) \quad \text{and} \quad T = \hat{t}_n(S_n) \rightarrow \hat{t}(S),$$

where we also used (3.43).

*Complementarity relations* We continue with the complementarity-like relations in (2.10). First, the set

$$\{(\tau, \zeta) \in L^2((0, S)) \times L^2((0, S); \mathcal{V}) : \tau(s) \geq 0, \tau(s) + \|\zeta(s)\|_{\mathcal{V}} \leq 1 \text{ f.a.a. } s \in (0, S)\}$$

is clearly convex and closed, thus weakly closed and consequently, we obtain that the weak limit  $(\hat{t}, \hat{z})$  satisfies the inequalities in (2.10a).

Next we turn to (2.10b), whose derivation is by far more involved. Since  $D_z \mathcal{F}$  is supposed to be weakly continuous as a mapping from  $\mathcal{Z}$  to  $\mathcal{Z}^*$ , it follows from (3.47) that

$$D_z \mathcal{I}(\underline{t}_n(s), \bar{z}_n(s)) \rightarrow D_z \mathcal{I}(\hat{t}(s), \hat{z}(s)) \quad \text{in } \mathcal{Z}^* \quad \forall s \in [0, S].$$

Thanks to (2.15), i.e.,  $\Pi_n(z) \rightarrow z$  in  $\mathcal{Z}$  for every  $z \in \mathcal{Z}$ , this gives

$$\begin{aligned} & \langle \Pi_n^* D_z \mathcal{I}(\underline{t}_n(s), \bar{z}_n(s)), z \rangle_{\mathcal{V}^*, \mathcal{V}} \\ &= \langle D_z \mathcal{I}(\underline{t}_n(s), \bar{z}_n(s)), \Pi_n z \rangle_{\mathcal{Z}^*, \mathcal{Z}} \rightarrow \langle D_z \mathcal{I}(\hat{t}(s), \hat{z}(s)), z \rangle_{\mathcal{Z}^*, \mathcal{Z}} \quad \forall z \in \mathcal{Z} \end{aligned} \quad (3.49)$$

i.e., weak convergence of  $\Pi_n^* D_z \mathcal{I}(\underline{t}_n(s), \bar{z}_n(s))$  to  $D_z \mathcal{I}(\hat{t}(s), \hat{z}(s))$  in  $\mathcal{Z}^*$ . Since, by (3.22),  $\Pi_n^* D_z \mathcal{I}(\underline{t}_n(s), \bar{z}_n(s))$  is uniformly bounded in  $\mathcal{V}^*$ , there exists a subsequence converging weakly to a  $g \in \mathcal{V}^*$ . From (3.49), we infer that  $g = D_z \mathcal{I}(\hat{t}(s), \hat{z}(s))$  holds in  $\mathcal{Z}^*$  and by density also in  $\mathcal{V}^*$ . Thus the weak limit in  $\mathcal{V}^*$  is unique and we deduce weak convergence of  $\Pi_n^* D_z \mathcal{I}(\underline{t}_n(s), \bar{z}_n(s))$  to  $D_z \mathcal{I}(\hat{t}(s), \hat{z}(s))$  in  $\mathcal{V}^*$  of the whole sequence.

Now we can take a closer look at the distance in (2.10b). Since  $\mathcal{R}_h : \mathcal{Z}_h \rightarrow \mathbb{R}$  is assumed to be l.s.c. and convex and  $\Pi_h : \mathcal{V} \rightarrow \mathcal{Z}_h$  is linear and continuous, the subdifferential of  $\mathcal{R}_h \circ \Pi_h$  is convex, closed, and bounded. Therefore,

$$\begin{aligned} & \text{dist}_{\mathcal{V}^*} \{-\Pi_n^* D_z \mathcal{I}(\underline{t}_n(s), \bar{z}_n(s)), \partial(\mathcal{R}_n \circ \Pi_n)(0)\} \\ &= \inf_{w \in \partial(\mathcal{R}_n \circ \Pi_n)(0)} \|-\Pi_n^* D_z \mathcal{I}(\underline{t}_n(s), \bar{z}_n(s)) - w\|_{\mathcal{V}^{-1}} \end{aligned}$$

admits a solution  $\eta_n \in \partial(\mathcal{R}_n \circ \Pi_n)(0)$ . Because of (3.31), there exists a subsequence, again denoted by  $\eta_n$ , converging weakly to a  $\tilde{\eta}$  in  $\mathcal{V}^*$ . Now,  $\eta_n \in \partial(\mathcal{R}_n \circ \Pi_n)(0)$  is equivalent to

$$\mathcal{R}_n(\Pi_n v) \geq \langle \eta_n, v \rangle \quad \forall v \in \mathcal{V}.$$

By weak convergence, the right hand side converges to  $\langle \tilde{\eta}, v \rangle$ . The left hand side converges to  $\mathcal{R}(v)$  on a dense subset  $\mathcal{U}$  by Assumption (c) on the approximation of the dissipation potential. By density of  $\mathcal{U} \subset \mathcal{V}$  and continuity of  $\mathcal{R}$ , we thus obtain  $\mathcal{R}(v) \geq \langle \tilde{\eta}, v \rangle$  for all

$v \in \mathcal{V}$  so that  $\eta \in \partial\mathcal{R}(0)$ . Together with the weak lower semicontinuity of the norm and the weak convergence of  $\Pi_n^* D_z \mathcal{I}(t_n(s), \bar{z}_n(s))$  established above, this implies

$$\begin{aligned} \liminf_{n \rightarrow \infty} \text{dist}_{\mathcal{V}^*} \{ -\Pi_n^* D_z \mathcal{I}(t_n(s), \bar{z}_n(s)), \partial(\mathcal{R}_n \circ \Pi_n)(0) \} \\ = \liminf_{n \rightarrow \infty} \| \Pi_n^* D_z \mathcal{I}(t_n(s), \bar{z}_n(s)) - \eta_n \|_{\mathcal{V}^{-1}} \\ \geq \| -D_z \mathcal{I}(\hat{t}(s), \hat{z}(s)) - \tilde{\eta} \|_{\mathcal{V}^{-1}} \geq \text{dist}_{\mathcal{V}^*} \{ -D_z \mathcal{I}(\hat{t}(s), \hat{z}(s)), \partial\mathcal{R}(0) \}. \end{aligned}$$

As this is the case for all weakly converging subsequences of  $\eta_n$ , we obtain

$$\begin{aligned} \liminf_{n \rightarrow \infty} \text{dist}_{\mathcal{V}^*} \{ -\Pi_n^* D_z \mathcal{I}(t_n(s), \bar{z}_n(s)), \partial(\mathcal{R}_n \circ \Pi_n)(0) \} \\ \geq \text{dist}_{\mathcal{V}^*} \{ -D_z \mathcal{I}(\hat{t}(s), \hat{z}(s)), \partial\mathcal{R}(0) \} \end{aligned} \quad (3.50)$$

for the whole sequence. To show (2.10b), let us abbreviate

$$\begin{aligned} \xi_n(s) &:= \text{dist}_{\mathcal{V}^*} \{ -\Pi_n^* D_z \mathcal{I}(t_n(s), \bar{z}_n(s)), \partial(\mathcal{R}_n \circ \Pi_n)(0) \}, \\ \xi(s) &:= \text{dist}_{\mathcal{V}^*} \{ -D_z \mathcal{I}(\hat{t}(s), \hat{z}(s)), \partial\mathcal{R}(0) \} \end{aligned}$$

so that (3.50) reads

$$\liminf_{n \rightarrow \infty} \xi_n(s) \geq \xi(s) \geq 0 \quad \forall s \in [0, S]. \quad (3.51)$$

We next address the measurability of  $\xi$ . By (3.22),  $\Pi_n^* D_z \mathcal{I}(t_n(\cdot), \bar{z}_n(\cdot))$  is uniformly bounded in time with values in  $\mathcal{V}^*$  and therefore weakly-\* converging in  $L^\infty(0, S; \mathcal{V}^*)$ . Since weak and pointwise limit coincide almost everywhere, we find that  $D_z \mathcal{I}(\hat{t}(\cdot), \hat{z}(\cdot))$  is an element of  $L^\infty(0, S; \mathcal{V}^*)$  and thus Bochner-measurable. Since  $\mathcal{V}^* \ni g \mapsto \text{dist}_{\mathcal{V}^*} \{ g, \partial\mathcal{R}(0) \} \in \mathbb{R}$  is continuous, this implies the desired measurability of  $\xi$ .

Now, consider an arbitrary  $\kappa \geq 0$  and define  $\xi_{n,\kappa}(s) := \min\{\xi_n(s), \kappa\}$  such that  $\xi_{n,\kappa}(s)$  converges to  $\xi_\kappa(s) := \min\{\xi(s), \kappa\}$  almost everywhere in  $(0, S)$ . Since  $\xi_\kappa$  is measurable (as  $\xi$  is so) and  $\kappa \geq \xi_{n,\kappa}(s)$ , Lebesgue's dominated convergence theorem gives

$$\xi_{n,\kappa} \rightarrow \xi_\kappa \text{ in } L^1(0, S).$$

Thus, thanks to  $\xi_n(s) \geq \xi_{n,\kappa}(s)$  and the weak-\* convergence of  $\hat{t}'$ , we obtain from (3.41) that

$$0 = \liminf_{n \rightarrow \infty} \int_0^S \hat{t}'_n(s) \xi_n(s) ds \geq \liminf_{n \rightarrow \infty} \int_0^S \hat{t}'_n(s) \xi_{n,\kappa}(s) ds = \int_0^S \hat{t}'(s) \xi_\kappa(s) ds.$$

Since  $\kappa \geq 0$  was arbitrary, this inequality holds for every  $\kappa$  so that Fatou's lemma yields

$$0 \geq \liminf_{\kappa \rightarrow \infty} \int_0^S \hat{t}'(s) \xi_\kappa(s) ds \geq \int_0^S \hat{t}'(s) \xi(s) ds \geq 0.$$

Because of  $\xi \geq 0$  and  $\hat{t}' \geq 0$  a.e. in  $(0, S)$ , cf. (2.10a), this gives (2.10b).

*Energy identity* Let  $s \in [0, S]$  be arbitrary. Thanks to the weak continuity of  $\mathcal{I}(t, \cdot)$  and the continuity of  $\ell$  by assumption, (3.44) and (3.46) yield

$$\mathcal{I}(\hat{t}(s), \hat{z}(s)) \leq \liminf_{n \rightarrow \infty} \mathcal{I}(\hat{t}_n(s), \hat{z}_n(s)).$$

Moreover, the lower semicontinuity and convexity of  $\mathcal{R}$  and Assumption (d) on the discretization thereof imply in view of (3.45) that

$$\int_0^s \mathcal{R}(\hat{z}'(\sigma)) d\sigma \leq \liminf_{n \rightarrow \infty} \int_0^s \mathcal{R}(\hat{z}'_n(\sigma)) d\sigma \leq \liminf_{n \rightarrow \infty} \int_0^s \mathcal{R}_n(\hat{z}'_n(\sigma)) d\sigma.$$

These estimates in combination with  $\|\hat{z}'(s)\|_{\mathbb{V}} \leq 1$ , (3.50), and Lemma 3.8 yield

$$\begin{aligned} & \mathcal{I}(\hat{t}(s), \hat{z}(s)) + \int_0^s \mathcal{R}(\hat{z}'(\sigma)) + \|\hat{z}'(\sigma)\|_{\mathbb{V}} \operatorname{dist}_{\mathcal{V}^*} \{-D_z \mathcal{I}(\hat{t}(\sigma), \hat{z}(\sigma)), \partial \mathcal{R}(0)\} d\sigma \\ & \leq \liminf_{n \rightarrow \infty} \left( \mathcal{I}(\hat{t}_n(s), \hat{z}_n(s)) \right. \\ & \quad \left. + \int_0^s \mathcal{R}_n(\hat{z}'_n(\sigma)) + \operatorname{dist}_{\mathcal{V}^*} \{-\Pi_n^* D_z \mathcal{I}(\hat{t}_n(\sigma), \hat{z}_n(\sigma)), \partial(\mathcal{R}_n \circ \Pi_n)(0)\} d\sigma \right) \\ & = \liminf_{n \rightarrow \infty} \left( \mathcal{I}(\hat{t}_n(0), \hat{z}_n(0)) + \int_0^s \partial_t \mathcal{I}(\hat{t}_n(\sigma), \hat{z}_n(\sigma)) \hat{t}'_n(\sigma) d\sigma + \int_0^s r_n(\sigma) d\sigma \right). \end{aligned}$$

Due to the continuous differentiability of  $\ell$ , the strong convergence of  $\hat{z}_n$  in  $L^1((0, S); \mathcal{Z})$  (by (3.45) and (3.46)), the weak-\* convergence of  $\hat{t}'_n$  in  $L^\infty(0, S)$ , (3.42), and (3.48), the last bracket converges to  $\mathcal{I}(0, z_0) + \int_0^s \partial_t \mathcal{I}(\hat{t}(\sigma), \hat{z}(\sigma)) \hat{t}'(\sigma) d\sigma$  so that

$$\begin{aligned} & \mathcal{I}(\hat{t}(s), \hat{z}(s)) + \int_0^s \mathcal{R}(\hat{z}'(\sigma)) + \|\hat{z}'(\sigma)\|_{\mathbb{V}} \operatorname{dist}_{\mathcal{V}^*} \{-D_z \mathcal{I}(\hat{t}(\sigma), \hat{z}(\sigma)), \partial \mathcal{R}(0)\} d\sigma \\ & \leq \mathcal{I}(0, z_0) + \int_0^s \partial_t \mathcal{I}(\hat{t}(\sigma), \hat{z}(\sigma)) \hat{t}'(\sigma) d\sigma \end{aligned}$$

is obtained. The reverse inequality can be shown completely analogously to [KRZ13, Lemma 5.2] by using the chain rule and the characterization of  $\partial \mathcal{R}(0)$  in Lemma C.1. We point out that the arguments in [KRZ13] do not rely on any regularization or discretization and can therefore directly be applied to the limit  $(\hat{t}, \hat{z})$ . All in all, we have verified the energy identity in (2.11), which completes the proof.  $\square$

## 4 Algorithmic realization and numerical tests

### 4.1 Test problem

To test the fully discrete local minimization algorithm numerically, we choose the following setting for the problem data in Section 2:

- In the following,  $\Omega \subset \mathbb{R}^d$ ,  $d = 2, 3$ , is a bounded domain.
- The spaces are chosen to be  $\mathcal{Z} = H_0^1(\Omega)$ ,  $\mathcal{V} = L^2(\Omega)$ , and  $\mathcal{X} = L^1(\Omega)$ .
- For the operator  $\mathbb{V} : L^2(\Omega) \rightarrow L^2(\Omega)^*$ , we just choose the Riesz isomorphism.
- The operator  $A$  within the energy functional is set to  $A = -\Delta : H_0^1(\Omega) \rightarrow H^{-1}(\Omega)$  such that the coercivity constant  $\alpha$  equals Poincaré's constant.

- The nonlinearity  $\mathcal{F}$  in the energy is defined as the well-known double well potential

$$\mathcal{F}(z) := 48 \int_{\Omega} (1 - z(x)^2)^2 dx,$$

which is non-negative and twice Fréchet-differentiable as a functional in  $L^4(\Omega)$  and, via Sobolev embeddings, also in  $H_0^1(\Omega)$  with  $D_z \mathcal{F}(z)h = -192 \int_{\Omega} (1 - z(x)^2)z(x)h(x)dx$ . Due to  $H_0^1(\Omega) \hookrightarrow L^6(\Omega)$ , the derivative can be extended to a linear functional on  $L^2(\Omega)$  and is continuous in these spaces. Moreover, using once again Sobolev embeddings and Hölder's inequality, one finds

$$|D_z^2 \mathcal{F}(z)[v, h]| \leq C(1 + \|z\|_{H^1(\Omega)}^2) \|v\|_{H^1(\Omega)} \|h\|_{L^2(\Omega)} \quad \forall z, v \in H_0^1(\Omega), h \in L^2(\Omega),$$

which is (2.5) with  $q = 2$ .

- The external loads are only depending on  $t$  and given by

$$\ell(t, x) = \ell(t) := -48 \sin(2\pi t), \quad (t, x) \in [0, T] \times \Omega$$

and thus satisfies the regularity requirements in (2.2).

- The dissipation functional is given by the  $L^1$ -norm, i.e.,  $\mathcal{R}(v) = \|v\|_{L^1(\Omega)}$  so that (2.6) is fulfilled with  $\underline{\rho} = 1$  and  $\bar{\rho} = |\Omega|^{1/2}$ .

Hence, all assumptions on the data are fulfilled by this concrete setting and we can turn to the spatial discretization of the problem under consideration.

#### 4.2 Finite Element discretization

We employ classical linear finite elements (FE) to discretize the energy and the dissipation functional. For this purpose, assume that a family  $\{\mathcal{T}_h\}_{h>0}$  of shape-regular triangulations of the domain  $\Omega$  be given. Herein,  $h$  denotes the mesh size defined by  $h := \max_{T \in \mathcal{T}_h} \text{diam}(T)$ . To keep the discussion concise, we assume that  $\Omega$  is a polygon and polyhedron, respectively, and that the triangulations exactly fit the boundary. For the discrete space, we choose the space of piecewise linear and continuous test functions, i.e.,

$$\mathcal{Z}_h := \{v \in C(\bar{\Omega}) \cap H_0^1(\Omega) : v|_T \in \mathcal{P}_1 \forall T \in \mathcal{T}_h\}.$$

By classical results on Lagrange and quasi-interpolation, respectively, the best approximation properties of the orthogonal and the Ritz projection show that the approximation assumptions in (2.14) are fulfilled. Moreover, as shown in [GHS16], the shape-regularity of the triangulation guarantees that the stability assumption in (2.12) is satisfied.

The discretization of the dissipation potential in form of the  $L^1$ -norm is performed by a mass lumping scheme, which turns out to be advantageous for the numerical solution of the local minimization problems in (3.1), as we will see in Section 4.3. Let us denote the nodes of the triangulation  $\mathcal{T}_h$  and the associated nodal basis by  $x_i$  and  $\varphi_i$ ,  $i = 1, \dots, N_h$ . Moreover, given a function  $z_h \in \mathcal{Z}_h$ , we denote the coefficient vector of  $z_h$  w.r.t. the nodal basis by  $\mathbf{z} = (z_1, \dots, z_{N_h}) \in \mathbb{R}^{N_h}$ , i.e.,  $z_h(x) = \sum_{i=1}^{N_h} z_i \varphi_i(x)$ . Then, the discrete dissipation potential  $\mathcal{R}_h : \mathcal{Z}_h \rightarrow \mathbb{R}$  is defined by

$$\mathcal{R}_h(z_h) := \int_{\Omega} \sum_{i=1}^{N_h} |z_i| \varphi_i(x) dx. \quad (4.1)$$

It remains to verify the Assumptions (a)–(d) on  $\mathcal{R}_h$ , which is done next.

**Proposition 4.1** *The discrete dissipation potential defined in (4.1) satisfies the conditions (a)–(d) from Section 2.3.*

*Proof* Due to the positivity of the nodal basis,  $\mathcal{R}_h$  is only a scaled version of the  $|\cdot|_1$ -norm on  $\mathbb{R}^N$  and consequently, it fulfills assumption (a). Moreover, the first inequality in (2.16) is obvious. For the second one, we use the affine transformation  $F_T : \hat{T} \rightarrow T$  to the reference element  $\hat{T} = \text{conv}((0, 0), (1, 0), (0, 1))$  to obtain

$$\mathcal{R}_h(z_h) = \sum_{T \in \mathcal{T}_h} \int_{\hat{T}} \sum_{x_i \in \bar{\hat{T}}} |z_i| (\varphi_i \circ F_T)(\hat{x}) |\det(DF_T(\hat{x}))| d\hat{x}$$

Let us denote the transformed basis functions by  $\hat{\varphi}_j$ ,  $j = 1, \dots, d$ . Due to the non-negativity of the nodal basis, each of the mappings

$$\mathbb{R}^d \ni (v_j)_{j=1}^d \mapsto \int_{\hat{T}} \sum_{j=1}^d |v_j| \hat{\varphi}_j(\hat{x}) d\hat{x} \quad \text{and} \quad \mathbb{R}^d \ni (v_j)_{j=1}^d \mapsto \int_{\hat{T}} \left| \sum_{j=1}^d v_j \hat{\varphi}_j(\hat{x}) \right| d\hat{x}$$

forms a norm on  $\mathbb{R}^d$ . Thus, by the norm-equivalence in finite dimensions, there exists a constant  $c > 0$ , only depending on  $d = \dim(\Omega)$ , such that

$$\mathcal{R}_h(z_h) \leq \sum_{T \in \mathcal{T}_h} c \int_{\hat{T}} \left| \sum_{x_i \in \bar{\hat{T}}} z_i (\varphi_i \circ F_T)(\hat{x}) |\det(DF_T(\hat{x}))| \right| d\hat{x} = c \|z_h\|_{L^1(\Omega)}, \quad (4.2)$$

which is (2.16).

For the convergence in (c), we set  $\mathcal{U} = C_c^\infty(\Omega)$  and estimate for all  $u \in \mathcal{U}$ :

$$\begin{aligned} & |\mathcal{R}_h(\Pi_h u) - \mathcal{R}(u)| \\ & \leq |\mathcal{R}_h(\Pi_h(u)) - \mathcal{R}_h(I_h(u))| + |\mathcal{R}_h(I_h(u)) - \mathcal{R}(I_h(u))| + |\mathcal{R}(I_h u) - \mathcal{R}(u)| \\ & =: e_1 + e_2 + e_3, \end{aligned}$$

where  $I_h : C(\bar{\Omega}) \rightarrow \mathcal{Z}_h$  denotes the Lagrange interpolation operator. Using the reverse triangle inequality in combination with (4.2), the first difference can be estimated by

$$\begin{aligned} e_1 &= \left| \int_{\Omega} \sum_{i=1}^{N_h} (|\Pi_h u(x_i)| - |I_h u(x_i)|) \varphi_i(x) dx \right| \\ &\leq \mathcal{R}_h(\Pi_h(u) - I_h(u)) \\ &\leq c \|\Pi_h(u) - I_h(u)\|_{L^1(\Omega)} \leq C \|\Pi_h\|_{\mathcal{L}(L^2(\Omega), L^1(\Omega))} \|u - I_h(u)\|_{L^2(\Omega)} \rightarrow 0, \quad \text{as } h \searrow 0. \end{aligned}$$

Thanks to the partition-of-unity property of the nodal basis, we obtain for the second difference by applying the reverse triangle inequality once again

$$\begin{aligned} e_2 &\leq \sum_{T \in \mathcal{T}_h} \int_T \sum_{i=1}^{N_h} |(I_h u)(x_i) - (I_h u)(x)| \varphi_i(x) dx \\ &\leq \sum_{T \in \mathcal{T}_h} h \|I_h u\|_{W^{1,\infty}(T)} \int_T \sum_{i=1}^{N_h} \varphi_i(x) dx \leq C h \|u\|_{W^{1,\infty}(\Omega)} \rightarrow 0, \quad \text{as } h \searrow 0. \end{aligned}$$

Using the reverse triangle inequality a third time, we estimate the last difference by

$$e_3 \leq \mathcal{R}(I_h u - u) \leq \bar{\rho} \|I_h u - u\|_{L^2(\Omega)} \rightarrow 0, \quad \text{as } h \searrow 0.$$

Overall, we thus obtain Assumption (c) with  $\mathcal{U} = C_c^\infty(\Omega)$ , which is clearly dense in  $L^2(\Omega)$ .

Finally, the non-negativity of the nodal basis directly implies for every  $v_h \in \mathcal{Z}_h$  that

$$\mathcal{R}(v_h) = \int_{\Omega} \left| \sum_{i=1}^{N_h} v_i \varphi_i(x) \right| dx \leq \int_{\Omega} \sum_{i=1}^{N_h} |v_i| \varphi_i(x) dx = \mathcal{R}_h(v_h),$$

which is the last Assumption (d).  $\square$

### 4.3 Numerical solution of the local minimization problems

The essential advantage of the discretization of  $\mathcal{R} = \|\cdot\|_{L^1(\Omega)}$  in (4.1) is that its subdifferential admits a *component-wise* characterization. This allows to rewrite the first-order optimality conditions associated with (3.1) as a system of non-smooth equations, which is amenable to semi-smooth Newton methods. To see this, note that the discrete dissipation potential can equivalently be rewritten as

$$\mathcal{R}_h(z_h) = \mathbf{R}(z) := \mathbf{m}^\top |z| \quad \text{with} \quad \mathbf{m} = (m_1, \dots, m_{N_h}) := M\mathbf{1}, \quad (4.3)$$

where  $M_{ij} = \int_{\Omega} \varphi_i \varphi_j dx \in \mathbb{R}^{N_h \times N_h}$  is the mass matrix,  $\mathbf{1} = (1, \dots, 1) \in \mathbb{R}^{N_h}$ , and  $|z| = (|z_1|, \dots, |z_{N_h}|)$ . Therefore, the convex subdifferential of  $\mathcal{R}_h$  can be expressed as follows:

$$\mathbf{q} = (q_1, \dots, q_{N_h}) \in \partial \mathbf{R}(z_h) \iff |q_i| \leq m_i, \quad q_i z_i = m_i |z_i| \quad \forall i = 1, \dots, N_h,$$

which can equivalently be formulated as

$$\max\{|q_i| - m_i, m_i |z_i| - q_i z_i\} = 0 \quad \forall i = 1, \dots, N_h.$$

To reformulate the optimality conditions of (3.1) as non-smooth equation, let us abbreviate the coefficient vector associated with  $z_k^{\tau, h}$  by  $z^k$ . Moreover, we denote the energy functional considered as mapping acting on the coefficient vector by  $\mathbf{I} : \mathbb{R}^{N_h} \rightarrow \mathbb{R}$ . Then, (3.1) is equivalent to the following minimization problem for the coefficient vector  $z^k$ :

$$(4.4) \quad \begin{cases} \min_{z \in \mathbb{R}^{N_h}} & \mathbf{I}(t_{k-1}, z) + \mathbf{R}(z - z^{k-1}) \\ \text{s.t.} & G(z) \leq 0, \end{cases}$$

with  $G(z) = \frac{1}{2}((z - z^{k-1})^\top M(z - z^{k-1}) - \tau^2)$ . Here and for the rest of this section, we abbreviate  $t_{k-1}^{\tau, h}$  simply by  $t_{k-1}$ . Based on the above description of the convex subdifferential of  $\mathbf{R}$ , standard arguments, such as e.g. Fenchel duality subsequent to linearization, can be applied to establish the following first-order necessary optimality conditions for this finite dimensional problem:

**Lemma 4.2** *If  $z^k \in \mathbb{R}^{N_h}$  is a local solution of (4.4), then there exists multipliers  $\mathbf{q} \in \mathbb{R}^{N_h}$  and  $\lambda \in \mathbb{R}$  such that*

$$D_z \mathbf{I}(t_{k-1}, z^k) + \lambda G'(z^k) + \mathbf{q} = 0, \quad (4.5a)$$

$$\max\{|q_i| - m_i, m_i |z_i^k - z_i^{k-1}| - q_i(z_i^k - z_i^{k-1})\} = 0 \quad \forall i = 1, \dots, N_h, \quad (4.5b)$$

$$\max\{-\lambda, G(z)\} = 0. \quad (4.5c)$$

*Remark 3* Denoting (with a little abuse of notation) the stiffness matrix associated with the FE discretization of the Laplacian by  $A \in \mathbb{R}^{N_h \times N_h}$ , (4.5a) is equivalent to

$$Az^k - 192 \left( \int_{\Omega} (1 - z_k^{\tau, h}(x))^2 z_k^{\tau, h}(x) \varphi_i(x) dx \right)_{i=1}^{N_h} + \ell(t_{k-1}) \mathbf{m} + \lambda M(\mathbf{z}^k - \mathbf{z}^{k-1}) + \mathbf{q} = 0.$$

Thus, by employing an appropriate quadrature rule, equation (4.5a) can be evaluated without any additional discretization error.

The optimality system in (4.5) is solved numerically by a semismooth Newton-algorithm, see e.g. [HPUU08]. To describe this in detail, let us denote the left hand side of (4.5) by  $F : \mathbb{R}^{2N_h+1} \rightarrow \mathbb{R}^{2N_h+1}$  so that (4.5) becomes  $F(\mathbf{z}, \mathbf{q}, \lambda) = 0$ . Of course,  $F$  depends on time discretization level  $k$ , but we suppress this dependency for the time being to shorten the notation. Now, given an iterate  $\mathbf{x}^n = (\mathbf{z}^n, \mathbf{q}^n, \lambda^n)$ , we compute the next one by solving the following semi-smooth Newton equation

$$H_n(\mathbf{x}^{n+1} - \mathbf{x}^n) = -F(\mathbf{x}^n) \quad \text{with} \quad H_n \in \partial^N F(\mathbf{x}^n),$$

where  $\partial^N F$  denotes the Newton-derivative according to [IK08]. As a composition of Newton-differentiable functions,  $F$  itself is Newton-differentiable, see [HPUU08, Thm. 2.10]. For our implementation, we choose

$$H_n := \begin{pmatrix} D_{zz}^2 \mathbf{I}(t_{k-1}, \mathbf{z}^n) + \lambda_n M & \text{Id}_{N_h \times N_h} & M(\mathbf{z}^n - \mathbf{z}^{k-1}) \\ \text{diag}(\boldsymbol{\alpha}^n) & \text{diag}(\boldsymbol{\beta}^n) & 0_{N_h} \\ \chi_n (\mathbf{z}^n - \mathbf{z}^{k-1})^\top M & 0_{N_h}^\top & -1 + \chi_n \end{pmatrix} \quad (4.6)$$

with

$$\boldsymbol{\alpha}_i^n := \begin{cases} 0, & m_i |z_i^n - z_i^{k-1}| - q_i^n (z_i^n - z_i^{k-1}) - |q_i^n| + m_i \leq 0, \\ m_i \widetilde{\text{sgn}}(z_i^n - z_i^{k-1}) - q_i^n, & m_i |z_i^n - z_i^{k-1}| - q_i^n (z_i^n - z_i^{k-1}) - |q_i^n| + m_i > 0, \end{cases}$$

$$\boldsymbol{\beta}_i^n := \begin{cases} \widetilde{\text{sgn}}(q_i^n), & m_i |z_i^n - z_i^{k-1}| - q_i^n (z_i^n - z_i^{k-1}) - |q_i^n| + m_i \leq 0, \\ z_i^{k-1} - z_i^n, & m_i |z_i^n - z_i^{k-1}| - q_i^n (z_i^n - z_i^{k-1}) - |q_i^n| + m_i > 0, \end{cases}$$

and

$$\chi_n := \begin{cases} 1, & G(\mathbf{z}^n) > -\lambda^n, \\ 0, & G(\mathbf{z}^n) \leq -\lambda^n, \end{cases} \quad \text{and} \quad \widetilde{\text{sgn}}(x) := \begin{cases} 1, & x \geq 0, \\ -1, & x < 0. \end{cases}$$

According to [IK08],  $\widetilde{\text{sgn}}$  constitutes an element of the Newton-derivative of the absolute value function. We choose this particular element instead of the  $\text{sgn}$ -function satisfying  $\text{sgn}(0) = 0$  in order to avoid the appearance of zero rows in  $H_n$ . With this choice, all matrices  $H_n$  appearing in our numerical test have shown to be invertible and the semismooth Newton method performed well with respect to both, robustness and efficiency. In particular, no globalization efforts are needed to ensure convergence of the method. A rigorous convergence analysis of the method however would go beyond the scope of this paper and is subject to future research.

#### 4.4 Numerical results

For our numerical test, we choose the unit square  $\Omega = (0, 1)^2$  as computational domain. Moreover, the initial state is set to  $z_0 \equiv 0$  and the final time is  $T = 1.0$ . Note that the initial state thus satisfies  $Az_0 \equiv 0 \in L^2(\Omega)$  as required by our standing assumptions in Section 2.1. The domain is discretized by a Friedrich-Keller triangulation with mesh size  $h = \sqrt{2}/50$ . For the time step size, we choose  $\tau = 0.01$ . The numerical computations are performed with MATLAB<sup>®</sup> and the linear systems of equations arising in each semi-smooth Newton step are solved by MATLAB's inbuilt direct solver based on UMFPAK.

We compare the local minimization algorithm with the global minimization scheme from (1.5), which is discretized in the same way as (3.1) by using piecewise linear finite elements. The minimization problem in (1.5a) is also solved by means of the semi-smooth Newton method. In order to ensure the convergence to global minimizers, we choose the two global minimizers of the nonlinear function  $\mathcal{F}$  (i.e.,  $z \equiv 1$  and  $z \equiv -1$ ) as starting points for the semi-smooth Newton method.

Let us first comment on the results of the local minimization iteration from Algorithm 1. Since  $\ell(0) \equiv 0$ , the initial state  $z_0 = P_h z_0$  is locally stable, meaning  $-D_z \mathcal{I}(0, z_0) \in \partial \mathcal{R}_h(0)$ . Consequently, the state does not change in the first iteration and, thanks to the time update in (3.2), the physical time proceeds by  $\tau = 0.01$ . However,  $z_0 \equiv 0$  is only a local maximum of the non-linearity  $\mathcal{F}$  and therefore, the external load enforces the state to jump immediately after the first time step into a local minimum in the subsequent iterations. In case of the local minimization algorithm, this jumps evolves as a viscous transition, while the physical time stagnates, see Figure 4.2. The state after this viscous transition is shown in Figure 4.1b. Afterwards the system evolves in a time continuous manner until  $t \approx 0.6724$ . At this time, a second jump occurs and the system switches into a viscous behaviour, which can be observed in Figures 4.1d-4.1h, finally yielding the state in Figure 4.1i. Meanwhile the physical time again stands still (see Figure 4.2), so that the solution in fact changes in a jump-like fashion. The end state is shown in Figure 4.1j.

Let us now turn to the results of the global minimization scheme from (1.5). Just as in case of Algorithm 1, the state jumps to a global minimum immediately after the first time step and evolves continuously afterwards. However, as the time evolves, both solutions show a quite different behaviour. While the second discontinuity of the parametrized solution shows up at  $t \approx 0.6724$ , as depicted above, the global energetic solution already jumps at  $t \approx 0.51$ . In view of the global minimization in (1.5a), it is intuitively expected that the global energetic solution jumps as soon as possible (cf. also [MRS12, Mie03] and the references therein). The difference between the global energetic and the parametrized solution can even be further enhanced. For instance, by choosing  $\ell(t) = -32 \sin(2\pi t)$ , the global energetic solution still provides a discontinuity at  $t \approx 0.51$ , while the parameterized one remains continuous until the end time is reached.

## 5 Conclusions

We presented a full space-time discretization scheme of a rate-independent evolution including a non-convex energy functional. The time discretization is based on the local minimization scheme introduced in [EM06], whereas we employ standard piecewise linear finite elements for the spatial discretization. By adapting the analysis of the recent contribution

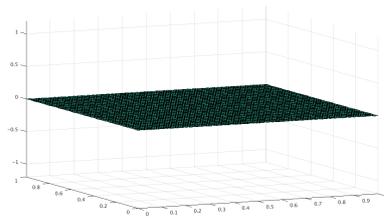
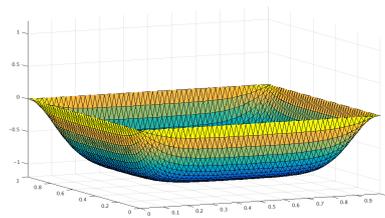
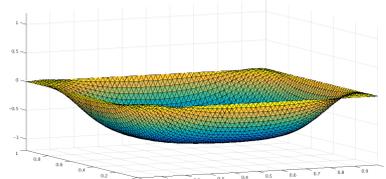
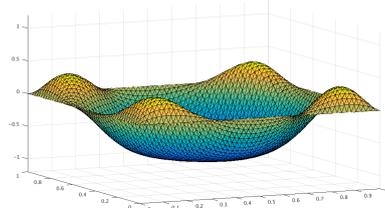
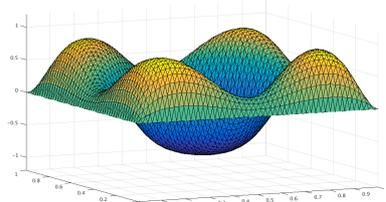
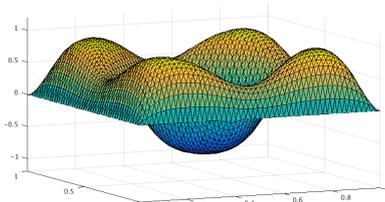
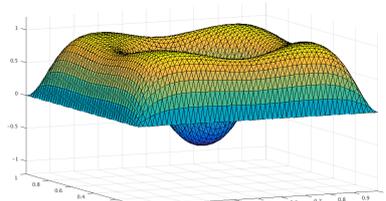
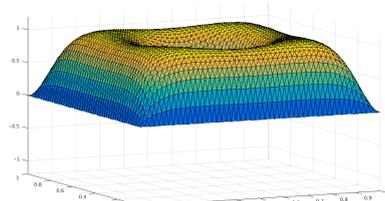
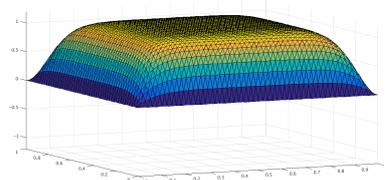
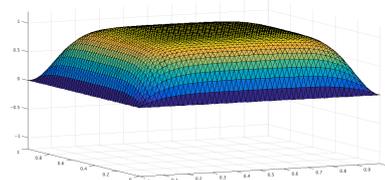
(a)  $t = 0.00$ (b)  $t = 0.01$ (c)  $t = 0.6724$ (d)  $t = 0.6724$ (e)  $t = 0.6724$ (f)  $t = 0.6724$ (g)  $t = 0.6724$ (h)  $t = 0.6724$ (i)  $t = 0.6729$ (j)  $t = 1.00$ 

Fig. 4.1: Computed parameterized solution to the problem described in Section 4. Figures 4.1c–4.1i show the viscous transition corresponding to the discontinuity at time  $t \approx 0.6724$ .

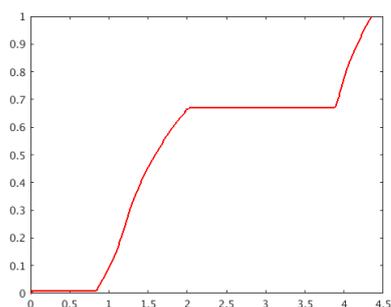


Fig. 4.2: Evolution of the physical time as function of the artificial time. The physical time stands still during the viscous transitions at time  $t \approx 0.01$  and  $t \approx 0.6724$

[Kne17], we proved that (weak) accumulation points of the sequence of discrete solutions for mesh size tending to zero exist and are so-called parametrized solutions of the original evolutionary system. For the discretization of the dissipation functional, we employed a mass lumping scheme, which on the one hand turned out to be advantageous for the numerical realization of the algorithm and on the other hand can be incorporated into the (abstract) convergence analysis. This mass lumping allows a reformulation of the discrete optimality system arising in each step of the local minimization scheme as non-smooth equation, which is amenable for semi-smooth Newton methods. By choosing a particular Newton derivative, the arising algorithm behaves robust and efficient in our numerical test. For the case of a double-well potential, we compared the local minimization scheme with another time discretization scheme known to converge to global energetic solutions. We observe that both schemes indeed provide different solutions, which jump at different points in time.

There are various future research questions arising in the context of a fully discrete local minimization scheme. This especially concerns the derivation of a priori and a posteriori error estimates. In our opinion, there is no hope to derive a priori error estimates without further assumptions, since the solutions are in general not unique and there may be a whole continuum of solutions. In case of uniform convex energies, error estimates for global minimization schemes of the form (1.5) have been derived in [MPPS10], and we expect that parts of the analysis can be adapted to the local minimization scheme, although an estimation of the discrete physical time as function of the artificial time will certainly be an issue in this context. A further question addresses the convergence behavior of the semi-smooth Newton method introduced in Section 4.3. As the method proved to be fairly robust in our numerical computations, it seems that the Newton-matrices in (4.6) are all invertible and that their inverses are bounded so that the classical convergence analysis of semi-smooth Newton methods applies. This however is subject to future research, in particular also in infinite dimensions, i.e., without a spatial discretization.

**Acknowledgements** The authors are very grateful to Dorothee Knees (University of Kassel) for various helpful discussions.

### A Proof of Lemma 3.2

We start the proof of Lemma 3.2 with the following result, which is a direct consequence of the characterization of  $z_k^{\tau,h}$  as a minimizer of (3.1). (This is actually the only point, where one uses that  $z_k^{\tau,h}$  is a minimizer and not only a stationary point satisfying (3.3)–(3.6).)

**Lemma A.1 (Local energy-inequality)** *For all  $h, \tau > 0$  and all  $k \in \mathbb{N}$ , the inequality*

$$\mathcal{I}(t_k^{\tau,h}, z_k^{\tau,h}) + \mathcal{R}_h(z_k^{\tau,h} - z_{k-1}^{\tau,h}) \leq \mathcal{I}(t_{k-1}^{\tau,h}, z_{k-1}^{\tau,h}) + \int_{t_{k-1}^{\tau,h}}^{t_k^{\tau,h}} \partial_t \mathcal{I}(s, z_k^{\tau,h}) ds \quad (\text{A.1})$$

is valid.

*Proof* By optimality of  $z_k^{\tau,h}$  and feasibility of  $z_{k-1}^{\tau,h}$  for the minimization problem in (3.1), we obtain

$$\mathcal{I}(t_{k-1}^{\tau,h}, z_k^{\tau,h}) + \mathcal{R}_h(z_k^{\tau,h} - z_{k-1}^{\tau,h}) \leq \mathcal{I}(t_{k-1}^{\tau,h}, z_{k-1}^{\tau,h}) \quad (\text{A.2})$$

Adding  $\mathcal{I}(t_k^{\tau,h}, z_k^{\tau,h})$  on both sides then gives the result.  $\square$

The local energy inequality is now used to derive the uniform bounds on energy and dissipation in Lemma 3.2:

*Proof (of Lemma 3.2)* Again we suppress the superscript  $\tau, h$  in the proof to shorten the notation, except for  $z_0^{\tau,h}$  in order to avoid confusion with the initial value. Estimating the right hand side in (A.1) by employing (2.4) gives

$$\begin{aligned} & \mathcal{I}(t_k, z_k) + \mathcal{R}_h(z_k - z_{k-1}) \\ & \leq \mathcal{I}(t_{k-1}, z_{k-1}) + \int_{t_{k-1}}^{t_k} \mu(\mathcal{I}(t_{k-1}, z_k) + \beta) \exp(\mu(s - t_{k-1})) ds \\ & = \mathcal{I}(t_{k-1}, z_{k-1}) + (\mathcal{I}(t_{k-1}, z_k) + \beta)(\exp(\mu(t_k - t_{k-1})) - 1). \end{aligned}$$

Using the non-negativity of  $\mathcal{R}_h$  by (2.16) in combination with (A.2) yields  $\mathcal{I}(t_{k-1}, z_k) \leq \mathcal{I}(t_{k-1}, z_{k-1})$  so that

$$\begin{aligned} & \mathcal{I}(t_k, z_k) + \mathcal{R}_h(z_k - z_{k-1}) \\ & \leq \mathcal{I}(t_{k-1}, z_{k-1}) + (\mathcal{I}(t_{k-1}, z_{k-1}) + \beta)(\exp(\mu(t_k - t_{k-1})) - 1) \end{aligned} \quad (\text{A.3})$$

is obtained. By exploiting once again  $\mathcal{R}_h \geq 0$ , this implies

$$\mathcal{I}(t_k, z_k) \leq (\mathcal{I}(t_{k-1}, z_{k-1}) + \beta) \exp(\mu(t_k - t_{k-1})) - \beta$$

such that induction over  $k$  already gives the desired result for the energy:

$$\begin{aligned} \mathcal{I}(t_k, z_k) & \leq (\mathcal{I}(0, z_0^{\tau,h}) + \beta) \prod_{j=1}^k \exp(\mu(t_j - t_{j-1})) - \beta \\ & \leq (\mathcal{I}(0, z_0^{\tau,h}) + \beta) \exp(\mu t_k) - \beta. \end{aligned} \quad (\text{A.4})$$

To include the dissipation in the estimate, we sum up (A.3) to obtain

$$\begin{aligned} & \mathcal{I}(t_k, z_k) + \sum_{j=1}^k \mathcal{R}_h(z_j - z_{j-1}) \\ & \leq \mathcal{I}(0, z_0^{\tau,h}) + \sum_{j=1}^k (\mathcal{I}(t_{j-1}, z_{j-1}) + \beta)(\exp(\mu(t_j - t_{j-1})) - 1). \end{aligned}$$

Inserting (A.4) and adding  $\beta$  on both sides, we finally obtain

$$\begin{aligned} \mathcal{I}(t_k, z_k) &+ \sum_{j=1}^k \mathcal{R}_h(z_j - z_{j-1}) + \beta \\ &\leq (\mathcal{I}(0, z_0^{\tau, h}) + \beta) + \sum_{j=1}^k (\mathcal{I}(0, z_0^{\tau, h}) + \beta) \exp(\mu t_{j-1}) (\exp(\mu(t_j - t_{j-1})) - 1) \\ &= (\mathcal{I}(0, z_0^{\tau, h}) + \beta) \exp(\mu t_k) \leq (\mathcal{I}(0, z_0^{\tau, h}) + \beta) \exp(\mu T), \end{aligned}$$

which is the claimed estimate.  $\square$

## B Proof of the discrete energy identity in Lemma 3.8

We follow the lines of [Kne17] and start with the proof of (3.41). From (3.3), (3.4), and the binomial formula, we infer

$$\text{dist}_{\mathcal{V}^*} \{-\Pi_h^* D_z \mathcal{I}(t_{k-1}^{\tau, h}, z_k^{\tau, h}), \partial(\mathcal{R}_h \circ \Pi_h)(0)\} = \lambda_k^{\tau, h} \tau. \quad (\text{B.1})$$

For arbitrary  $k \in \{1, \dots, N-1\}$ , we thus deduce from (3.3) and (3.37) that

$$\begin{aligned} 0 &= \lambda_k^{\tau, h} (\|z_k^{\tau, h} - z_{k-1}^{\tau, h}\|_{\mathbb{V}} - \tau) \\ &= \lambda_k^{\tau, h} \tau (1 - \|\hat{z}'_{\tau, h}(s)\|_{\mathbb{V}}) \\ &= \hat{t}'_{\tau, h}(s) \text{dist}_{\mathcal{V}^*} \{-\Pi_h^* D_z \mathcal{I}(\hat{t}_{\tau, h}(s), \bar{z}_{\tau, h}(s)), \partial(\mathcal{R}_h \circ \Pi_h)(0)\}, \quad \forall s \in [s_{k-1}^{\tau, h}, s_k^{\tau, h}), \end{aligned}$$

where we employed the definition of the constant interpolants in (3.35). This gives (3.41) for almost all  $s \in (0, s_{N-1}^{\tau, h})$ . As seen at the end of the proof of Lemma 3.7, it holds  $\lambda_N^{\tau, h} = 0$  so that (B.1) implies the assertion for  $s \in (s_{N-1}^{\tau, h}, S_{\tau, h})$ .

Next we turn to the discrete energy-identity. Since the affine interpolants in (3.34) are by construction elements of  $W^{1, \infty}((0, S_{\tau, h}))$  and  $W^{1, \infty}((0, S_{\tau, h}); \mathcal{Z})$ , respectively, and due to  $\mathcal{I} \in C^1([0, T] \times \mathcal{Z})$  by assumption, the chain rule is applicable and gives for  $s \in [s_{k-1}^{\tau, h}, s_k^{\tau, h})$  that

$$\begin{aligned} &\frac{d}{ds} \mathcal{I}(\hat{t}_{\tau, h}(s), \hat{z}_{\tau, h}(s)) \\ &= \partial_t \mathcal{I}(\hat{t}_{\tau, h}(s), \hat{z}_{\tau, h}(s)) \hat{t}'_{\tau, h}(s) + \langle D_z \mathcal{I}(\hat{t}_{\tau, h}(s), \hat{z}_{\tau, h}(s)), \hat{z}'_{\tau, h}(s) \rangle_{\mathcal{Z}^*, \mathcal{Z}} \\ &= \partial_t \mathcal{I}(\hat{t}_{\tau, h}(s), \hat{z}_{\tau, h}(s)) \hat{t}'_{\tau, h}(s) + \frac{1}{s_k^{\tau, h} - s_{k-1}^{\tau, h}} \langle D_z \mathcal{I}(\hat{t}_{\tau, h}(s), \bar{z}_{\tau, h}(s)), z_k^{\tau, h} - z_{k-1}^{\tau, h} \rangle_{\mathcal{Z}^*, \mathcal{Z}} \\ &\quad + \langle D_z \mathcal{I}(\hat{t}_{\tau, h}(s), \hat{z}_{\tau, h}(s)) - D_z \mathcal{I}(\hat{t}_{\tau, h}(s), \bar{z}_{\tau, h}(s)), \hat{z}'_{\tau, h}(s) \rangle_{\mathcal{Z}^*, \mathcal{Z}}. \end{aligned}$$

From (3.5), we have in combination with the 1-homogeneity of  $\mathcal{R}_h$  that

$$\begin{aligned} &-\frac{1}{s_k^{\tau, h} - s_{k-1}^{\tau, h}} \langle D_z \mathcal{I}(\hat{t}_{\tau, h}(s), \bar{z}_{\tau, h}(s)), z_k^{\tau, h} - z_{k-1}^{\tau, h} \rangle_{\mathcal{Z}^*, \mathcal{Z}} \\ &= \frac{1}{s_k^{\tau, h} - s_{k-1}^{\tau, h}} \left( \mathcal{R}(z_k^{\tau, h} - z_{k-1}^{\tau, h}) + \tau \text{dist}_{\mathcal{V}^*} \{-\Pi_h^* D_z \mathcal{I}(t_{k-1}^{\tau, h}, z_k^{\tau, h}), \partial(\mathcal{R}_h \circ \Pi_h)(0)\} \right) \\ &= \mathcal{R}_h(\hat{z}'_{\tau, h}) + \text{dist}_{\mathcal{V}^*} \{-\Pi_h^* D_z \mathcal{I}(t_{k-1}^{\tau, h}, z_k^{\tau, h}), \partial(\mathcal{R}_h \circ \Pi_h)(0)\}. \end{aligned}$$

Note that, in case of  $k < N$ , the last equation results from  $s_k^{\tau, h} - s_{k-1}^{\tau, h} = \tau$  for  $k < N$ , while, for the case  $k = N$ , it follows from  $\lambda_N^{\tau, h} = 0$  and (B.1), similarly to above. By taking into account the definition of  $r_{\tau, h}$  in (3.40), integration over  $(\sigma_1, \sigma_2)$  then yields (3.39).

It remains to estimate  $r_{\tau, h}$ . To this end, first observe that the definition of the affine and constant interpolants in (3.34) and (3.35) implies for every  $k \in \{1, \dots, N\}$  and every  $s \in [s_{k-1}^{\tau, h}, s_k^{\tau, h})$  that

$$\hat{z}_{\tau, h}(s) - \bar{z}_{\tau, h}(s) = (s - s_k^{\tau, h}) \hat{z}'_{\tau, h}(s) \quad \text{and} \quad \hat{t}_{\tau, h}(s) - \bar{t}_{\tau, h}(s) = (s - s_{k-1}^{\tau, h}) \hat{t}'_{\tau, h}(s),$$

which is frequently used in the following estimates. Now, let  $k \in \{1, \dots, N\}$  and  $s \in [s_{k-1}^{\tau,h}, s_k^{\tau,h})$  be arbitrary. Then, by inserting the concrete form of  $\mathcal{I}$  into the definition of  $r_{\tau,h}$  in (3.40) and employing the coercivity of  $A$  together with  $(s - s_k^{\tau,h}) < 0$ , we arrive at

$$\begin{aligned}
r_{\tau,h}(s) &= (s - s_k^{\tau,h}) \langle A(\hat{z}'_{\tau,h}(s)), \hat{z}'_{\tau,h}(s) \rangle_{\mathcal{Z}^*, \mathcal{Z}} \\
&\quad + \langle D_z \mathcal{F}(\hat{z}_{\tau,h}(s)) - D_z \mathcal{F}(\bar{z}_{\tau,h}(s)), \frac{\hat{z}_{\tau,h}(s) - \bar{z}_{\tau,h}(s)}{(s - s_k^{\tau,h})} \rangle_{\mathcal{V}^*, \mathcal{V}} \\
&\quad - \langle \ell(\hat{t}_{\tau,h}(s)) - \ell(\underline{t}_{\tau,h}(s)), \hat{z}'_{\tau,h}(s) \rangle_{\mathcal{V}^*, \mathcal{V}} \\
&\leq \alpha (s - s_k^{\tau,h}) \|\hat{z}'_{\tau,h}(s)\|_{\mathcal{Z}}^2 \\
&\quad + \frac{1}{|s - s_k^{\tau,h}|} |\langle D_z \mathcal{F}(\bar{z}_{\tau,h}(s)) - D_z \mathcal{F}(\hat{z}_{\tau,h}(s)), \bar{z}_{\tau,h}(s) - \hat{z}_{\tau,h}(s) \rangle_{\mathcal{V}^*, \mathcal{V}}| \\
&\quad + |\hat{t}_{\tau,h}(s) - \underline{t}_{\tau,h}(s)| \|\ell\|_{C^1([0,T], \mathcal{V}^*)} \|\hat{z}'_{\tau,h}(s)\|_{\mathcal{V}}.
\end{aligned} \tag{B.2}$$

We apply Lemma 3.5 with  $\varepsilon = \alpha/2$  to the second term on the right hand side to obtain

$$\begin{aligned}
&\frac{1}{|s - s_k^{\tau,h}|} |\langle D_z \mathcal{F}(\bar{z}_{\tau,h}(s)) - D_z \mathcal{F}(\hat{z}_{\tau,h}(s)), \bar{z}_{\tau,h}(s) - \hat{z}_{\tau,h}(s) \rangle_{\mathcal{V}^*, \mathcal{V}}| \\
&\leq \frac{\alpha}{2} |s - s_k^{\tau,h}| \|\hat{z}'_{\tau,h}(s)\|_{\mathcal{Z}}^2 + C_\alpha |s - s_k^{\tau,h}| \mathcal{R}_h(\hat{z}'_{\tau,h}(s)) \|\hat{z}'_{\tau,h}(s)\|_{\mathcal{V}}
\end{aligned}$$

where we also used the positive homogeneity of  $\mathcal{R}_h$ . By inserting this in (B.2) and using again that  $(s - s_k^{\tau,h}) < 0$ , one deduces

$$r_{\tau,h}(s) \leq C(\mathcal{R}_h(\hat{z}'_{\tau,h}(s)) + \hat{t}'_{\tau,h}(s) \|\ell\|_{C^1([0,T], \mathcal{V}^*)}) (s - s_{k-1}^{\tau,h}) \|\hat{z}'_{\tau,h}(s)\|_{\mathcal{V}}.$$

Integrating and exploiting the definition of  $\hat{z}_{\tau,h}$  and  $\hat{t}_{\tau,h}$ , respectively, then yields

$$\begin{aligned}
&\int_{\sigma_1}^{\sigma_2} r_{\tau,h}(s) ds \\
&\leq \sum_{i=1}^N \int_{s_{i-1}^{\tau,h}}^{s_i^{\tau,h}} C \left( \mathcal{R}_h(z_i^{\tau,h} - z_{i-1}^{\tau,h}) \right. \\
&\quad \left. + (t_i^{\tau,h} - t_{i-1}^{\tau,h}) \|\ell\|_{C^1([0,T], \mathcal{V}^*)} \right) \frac{s_i^{\tau,h} - s}{(s_i^{\tau,h} - s_{i-1}^{\tau,h})^2} \|z_i^{\tau,h} - z_{i-1}^{\tau,h}\|_{\mathcal{V}} ds \\
&= \sum_{i=1}^N \frac{1}{2} C \left( \mathcal{R}_h(z_i^{\tau,h} - z_{i-1}^{\tau,h}) + (t_i^{\tau,h} - t_{i-1}^{\tau,h}) \|\ell\|_{C^1([0,T], \mathcal{V}^*)} \right) \|z_i^{\tau,h} - z_{i-1}^{\tau,h}\|_{\mathcal{V}} \\
&\leq C \tau \left( T + \sum_{i=1}^N \mathcal{R}_h(z_i^{\tau,h} - z_{i-1}^{\tau,h}) \right).
\end{aligned}$$

Thanks to Lemma 3.2, the bracket on the right hand side is bounded independent of  $\tau$  and  $h$  so that (3.42) is proven, too.  $\square$

## C Auxiliary results from convex analysis

In this section, we collect some useful properties of  $\mathcal{R}$  and  $\mathcal{R}_h$ , respectively. We start with the following lemma, whose proof is straight forward and therefore omitted:

**Lemma C.1** *Let  $\mathcal{W}$  be a normed vector space and  $\mathcal{J} : \mathcal{W} \rightarrow \mathbb{R}$  a convex and positive 1-homogeneous functional. Then it holds*

$$\partial \mathcal{J}(v) \subset \partial \mathcal{J}(0) \quad \forall v \in \mathcal{W} \tag{C.1}$$

$$\xi \in \partial \mathcal{J}(0) \iff \mathcal{J}(w) \geq \langle \xi, w \rangle \quad \forall w \in \mathcal{W} \tag{C.2}$$

$$\partial \mathcal{J}(v) = \{\xi \in \partial \mathcal{J}(0) : \mathcal{J}(v) = \langle \xi, v \rangle\} \tag{C.3}$$

$$\mathcal{J}^*(\xi) = I_{\partial \mathcal{J}(0)}(\xi) \quad \forall \xi \in \mathcal{W}^* \tag{C.4}$$

where  $I_{\partial \mathcal{J}(0)}$  denotes the indicator functional of  $\partial \mathcal{J}(0)$ .

*Remark 4* Since only convexity and positive homogeneity is required for Lemma C.1 to hold, we may apply the above results to  $\mathcal{R}$ , its approximation  $\mathcal{R}_h$ , and  $\mathcal{R}_h \circ \Pi_h$ , considered as operators on  $\mathcal{V}$  as well as  $\mathcal{Z}$ .

As in the proof of Lemma 3.1, we abbreviate  $\mathcal{R}_{\tau,h} = \mathcal{R}_h \circ \Pi_h + I_\tau$ , where  $I_\tau$  is as defined in (3.8). Analogously, we set  $\mathcal{R}_\tau^h := \mathcal{R}_h + I_\tau(v)$ . We further denote by  $\partial^{\mathcal{Z}}\mathcal{R}_{\tau,h}$  the subdifferential of  $\mathcal{R}_{\tau,h}$  w.r.t.  $\mathcal{Z} - \mathcal{Z}^*$ -duality, while  $\partial\mathcal{R}_{\tau,h}$  is the subdifferential w.r.t.  $\mathcal{V} - \mathcal{V}^*$ -duality. The same notation is used for  $\mathcal{R}_\tau^h$ .

**Lemma C.2** *For every  $\eta \in \mathcal{V}^*$ , there holds*

$$(\mathcal{R}_{\tau,h})^*(\eta) = \tau \operatorname{dist}_{\mathcal{V}^*}\{\eta, \Pi_h^* \partial\mathcal{R}_h(0)\}, \quad (\text{C.5})$$

where  $\operatorname{dist}_{\mathcal{V}^*}\{\eta, \Pi_h^* \partial\mathcal{R}_h(0)\} = \inf\{\|\eta - \Pi_h^* w\|_{\mathcal{V}^*} : w \in \partial\mathcal{R}_h(0)\}$  and  $\|\eta\|_{\mathcal{V}^*}^2 = \langle \eta, \mathbb{V}^{-1}\eta \rangle$ .

*Proof* We use the inf-convolution formula (see [Att84, Prop. 3.4]), which is applicable, since  $\mathcal{R}_h \circ \Pi_h$  is continuous. This gives

$$(\mathcal{R}_h \circ \Pi_h + I_\tau)^*(\eta) = \inf_{w \in \mathcal{V}^*} ((\mathcal{R}_h \circ \Pi_h)^*(\eta) + I_\tau^*(w - \eta)). \quad (\text{C.6})$$

For  $I_\tau^*$ , direct calculation leads to

$$I_\tau^*(\eta) = \tau \|\eta\|_{\mathcal{V}^*} \quad (\text{C.7})$$

To calculate the conjugate functional of  $(\mathcal{R}_h \circ \Pi_h)^*$ , note that by the linearity of  $\Pi_h$  the composition  $\mathcal{R}_h \circ \Pi_h$  is again convex and 1-homogeneous. Therefore, Lemma C.1 gives  $(\mathcal{R}_h \circ \Pi_h)^*(\eta) = I_{\partial(\mathcal{R}_h \circ \Pi_h)(0)}(\eta)$ . The chain-rule for subdifferentials yields  $\partial(\mathcal{R}_h \circ \Pi_h)(0) = \Pi_h^* \partial\mathcal{R}_h(0)$  so that we obtain

$$(\mathcal{R}_h \circ \Pi_h)^*(\eta) = I_{\partial(\mathcal{R}_h \circ \Pi_h)(0)}(\eta) = I_{\Pi_h^* \partial\mathcal{R}_h(0)}(\eta).$$

Inserting this together with (C.7) in (C.6) finally yields

$$(\mathcal{R}_h \circ \Pi_h + I_\tau)^*(\eta) = \inf_{w \in \Pi_h^* \partial\mathcal{R}_h(0)} \{\tau \|\eta - w\|_{\mathcal{V}^*}\} = \tau \operatorname{dist}_{\mathcal{V}^*}(\eta, \Pi_h^* \partial\mathcal{R}_h(0)),$$

which is (C.5).  $\square$

**Lemma C.3** *Let  $v \in \mathcal{V}$  be arbitrary. Then,  $\xi \in \mathcal{V}^*$  is an element of  $\partial I_\tau(v)$ , iff there exists a multiplier  $\lambda \in \mathbb{R}$  such that  $\xi = \lambda \mathbb{V}v$  and*

$$\|v\|_{\mathcal{V}} \leq \tau, \quad \lambda(\|v\|_{\mathcal{V}} - \tau) = 0, \quad \lambda \geq 0.$$

*Proof* According to a classical result of convex analysis in combination with (C.7), it holds

$$\xi \in \partial I_\tau(v) \iff I_\tau(v) + I_\tau^*(\xi) = \langle \xi, v \rangle \iff \begin{cases} \|v\|_{\mathcal{V}} \leq \tau \\ \tau \|\xi\|_{\mathcal{V}^*} = \langle \xi, v \rangle \end{cases} \quad (\text{C.8})$$

Now, the Cauchy-Schwarz-Inequality implies  $\langle \xi, v \rangle = \langle \mathbb{V}^{-1}\xi, v \rangle \leq \|\xi\|_{\mathcal{V}^*} \|v\|_{\mathcal{V}} \leq \tau \|\xi\|_{\mathcal{V}^*}$  so that the equivalence in (C.8) can only hold if  $\mathbb{V}^{-1}\xi = \lambda v$  for some  $\lambda \in \mathbb{R}$ . Inserting this into (C.8), we conclude that  $\lambda \geq 0$ . Moreover, if  $\|v\|_{\mathcal{V}} < \tau$ , then  $\xi = 0$  so that  $\lambda$  fulfills also  $\lambda(\|v\|_{\mathcal{V}} - \tau) = 0$  as claimed.  $\square$

## References

- Att84. Hedy Attouch. *Variational Convergence for Functions and Operators*. Pitman Advanced Publishing Program, 1984.
- EM06. Messoud A. Efendiev and Alexander Mielke. On the rate-independent limit of systems with dry friction and small viscosity. *Journal of Convex Analysis*, 13(1):151–167, 2006.
- GHS16. Fernando D. Gaspoz, Claus-Justus Heine, and Kunibert G. Siebert. Optimal grading of the newest vertex bisection and  $H^1$ -stability of the  $L^2$ -projection. *IMA Journal of Numerical Analysis*, 36(3):1217–1241, 2016.

- HPUU08. Michael Hinze, René Pinnau, Michael Ulbrich, and Stefan Ulbrich. *Optimization with PDE constraints*, volume 23. Springer Science & Business Media, 2008.
- IK08. Kazufumi Ito and Karl Kunisch. *Lagrange multiplier approach to variational problems and applications*. SIAM, 2008.
- Kne17. Dorothee Knees. Convergence analysis of time-discretization schemes for rate-independent systems. submitted to ESAIM:COCV, 2017. URL: <https://arxiv.org/pdf/1712.06851.pdf>.
- KRZ13. Dorothee Knees, Riccarda Rossi, and Chiara Zanini. A vanishing viscosity approach to a rate-independent damage model. *Mathematical Models and Methods in Applied Sciences*, 23(04):565–616, 2013.
- Mie03. Alexander Mielke. Energetic formulation of multiplicative elasto-plasticity using dissipation distances. *Continuum Mechanics and Thermodynamics*, 15(4):351–382, 2003.
- Mie11. Alexander Mielke. Differential, energetic, and metric formulations for rate-independent processes. *Nonlinear PDE's and Applications*, 2028:87–167, 2011.
- MPPS10. Alexander Mielke, Laetitia Paoli, Adrien Petrov, and Ulisse Stefanelli. Error estimates for space-time discretizations of a rate-independent variational inequality. *SIAM Journal on Numerical Analysis*, 48(5):1625–1646, 2010.
- MR15. Alexander Mielke and Tomáš Roubíček. *Rate-Independent Systems: Theory and Application*. Springer-Verlag, New York, 2015.
- MRS12. Alexander Mielke, Riccarda Rossi, and Giuseppe Savaré. BV solutions and viscosity approximations of rate-independent systems. *ESAIM: Control, Optimization and Calculus of Variations*, 18(1):36–80, 2012.
- MRS16. Alexander Mielke, Riccarda Rossi, and Giuseppe Savaré. Balanced viscosity (BV) solutions to infinite-dimensional rate-independent systems. *Journal of the European Mathematical Society*, 18:2107–2165, 2016.
- MT04. Alexander Mielke and Florian Theil. On rate-independent hysteresis models. *NoDEA: Nonlinear Differential Equations and Applications*, 11(2):151–189, 2004.
- MZ14. Alexander Mielke and Sergey Zelik. On the vanishing-viscosity limit in parabolic systems with rate-independent dissipation terms. *Ann. Sc. Norm. Super. Pisa, Cl. Sci.*, 13(1):67–135, 2014.
- Neg14. Matteo Negri. Quasi-static rate-independent evolutions: characterization, existence, approximation and application to fracture mechanics. *ESAIM Control Optim. Calc. Var.*, 20(4):983–1008, 2014.
- NS17. Matteo Negri and Riccardo Scala. A quasi-static evolution generated by local energy minimizers for an elastic material with a cohesive interface. *Nonlinear Analysis: Real World Applications*, 38:271–305, 2017.