AN ADAPTIVE TIME STEPPING SCHEME FOR RATE-INDEPENDENT SYSTEMS WITH NON-CONVEX ENERGY

M. Andreia, C. Meyer

AN ADAPTIVE TIME STEPPING SCHEME FOR RATE-INDEPENDENT SYSTEMS WITH NON-CONVEX ENERGY

MERLIN ANDREIA AND CHRISTIAN MEYER

ABSTRACT. We investigate a local incremental stationary scheme for the numerical solution of rate-independent systems. Such systems are characterized by a (possibly) non-convex energy and a dissipation potential, which is positively homogeneous of degree one. Due to the non-convexity of the energy, the system does in general not admit a time-continuous solution. In order to resolve these potential discontinuities, the algorithm produces a sequence of state variables and physical time points as functions of a curve parameter. The main novelty of our approach in comparison to existing methods is an adaptive choice of the step size for the update of the curve parameter depending on a prescribed tolerance for the residua in the energy-dissipation balance and in a complementarity relation concerning the so-called local stability condition. It is proven that, for tolerance tending to zero, the piecewise affine approximations generated by the algorithm converge (weakly) to a so-called $\mathbb{V}$-parametrized balanced viscosity solution. Numerical experiments illustrate the theoretical findings and show that an adaptive choice of the step size indeed pays off as they lead to a significant increase of the step size during sticking and in viscous jumps.

1. INTRODUCTION

This paper is concerned with the design and analysis of an adaptive time stepping scheme for rate-independent systems of the form

$$(1.1) \qquad 0 \in \partial \mathcal{R}(\dot{z}(t)) + D_z \mathcal{I}(t, z(z)) \quad \text{f.a.a. } t \in (0, T), \quad z(0) = z_0,$$

where $\mathcal{R} : \mathcal{X} \to [0, \infty)$ denotes an 1-homogeneous and convex dissipation functional, while $\mathcal{I} : \mathcal{Z} \to \mathbb{R}$ is a non-convex, but smooth energy functional. The precise assumptions on the data and the involved spaces will be given in Section 1.3 below. Systems of the form (1.1) appear in various applications such as elastoplasticity, phase field models for damage evolution, or shape-memory alloys, see e.g. [3, 7] and the references therein.

It is well known that, caused by the interplay between the non-smooth, but convex dissipation and the smooth, but non-convex energy, there is in general no solution that fulfills (1.1) a.e. in time. We refer to the counterexample in [19, Section 2.3]. For this reason, several generalized notions of solutions have been introduced in the recent past, among them local solutions, global energetic solutions, and balanced viscosity (BV) solutions. We refer to [12] for a comprehensive survey on this matter. All these notions have in common that they allow for discontinuous solutions, even if the data, in particular the force driving the system, are smooth. This is essential to guarantee the existence of solutions as the aforementioned counterexample illustrates.

In this paper, we focus on a solution concept entitled $\mathbb{V}$-*parametrized BV solution*, which will be defined precisely in Section 2. This concept has two major advantages in comparison

---

to other notions of solution. First, it arises as limit in a vanishing viscosity approach and as such, solutions of this type will only provide jumps, if they are really necessary. Thus, solutions will be continuous in time as long as possible, which may be seen physically reasonable in many applications. Moreover, since the solution trajectory is parametrized by a curve parameter as the title indicates, $\mathbb{V}$-parametrized BV solutions enable a refined resolution of the jumps. As a result, a solution that was discontinuous as a function of time, becomes smooth in terms of the curve parameter.

Despite this gain of smoothness, the numerical approximation of solutions of this type is still a challenging issue. This is mainly due to the non-uniqueness of $\mathbb{V}$-parametrized BV solutions, see e.g. the counterexample in [10, Section 2.4]. We underline that a lack of uniqueness is typical for rate-independent systems with non-convex energies and also concerns the other notions of solutions, see [12, 17, 19] for various other counterexamples. One can therefore not expect that an approximation will converge to a solution unless the energy is uniformly convex (at least locally around a given solution trajectory), which guarantees uniqueness of the solution (at least locally). Nevertheless, in case of a uniform refinement of the interval for the curve parameter, discretization schemes have been developed that provide subsequences which converge weakly to $\mathbb{V}$-parametrized BV solutions. Such a scheme has first been introduced and analyzed in [4] in the finite dimensional setting, where $\mathcal{Z} = \mathbb{R}^n$. It is known as *local incremental minimization scheme*. In [6, 16, 8, 18], this approach has been transferred to infinite dimensions along with several modifications thereof. A combination of such a time stepping scheme with a spatial discretization is investigated in [9]. The aforementioned strong convergence in the uniformly convex case including convergence rates is established in [10]. Let us finally mention that incremental minimization schemes are also employed to approximate other types of solutions, such as global energetic solutions, we exemplarily refer to [2, 1] and the references therein.

All aforementioned contributions employ a uniform discretization of the interval for the curve parameter. While the discretization of the physical time interval is intrinsically adaptive in context of local incremental minimization schemes, the curve parameter was so far only updated by a fixed step size τ in each iteration. One may however essentially distinguish three different regimes characterizing a $\mathbb{V}$-parametrized BV solution, that is *rate-independent slip*, *viscous jump*, and *sticking*. In the last two regimes, one component of the solution curve $s \mapsto (t(s), z(s))$ is constant (t in a viscous jump and z during sticking), which calls for coarser step sizes for the curve parameter especially in this parts of the curve. On the other hand, the switching points between two different regimes should be resolved properly, which may require local refinement. There is thus a clear intuition that an adaptive choice of the step size for the curve parameter makes sense in the context of local incremental minimization schemes. Nevertheless, to the best of our knowledge, this has not been investigated so far. With this work, we aim to develop such an adaptive algorithm, which at the same time provides the same convergence properties as the existing methods with uniform step size, i.e., weak convergence of subsequences to $\mathbb{V}$-parametrized BV solutions, which is all one can expect in case of a non-convex energy as explained above.

1.1. **Outline of the Paper.** Let us give a brief overview over the paper. After introducing the notation and our standing assumptions on the data in (1.1) in the rest of this introduction, we will precisely define $\mathbb{V}$-parametrized BV solutions and motivate their definition in more details in Section 2. Afterwards, in Section 3, we present the basic version of our adaptive algorithm. Section 4 is then devoted to several results known from the uniform version of the algorithm with constant step size. If a proof is entirely analogous to the uniform case, we just state the associated result from the existing literature, otherwise we give a proof in the appendix. In Section 5, we then investigate the convergence of the algorithm in case $\mathcal{Z} = \mathbb{R}^n$. As indicated above, we show in Theorem 5.4 that subsequence of iterates converge weakly to $\mathbb{V}$-parametrized BV solutions, if the tolerance in the step size control tends to zero. Unfortunately, an analogous result for the infinite dimensional case where $\mathcal{Z} \neq \mathbb{R}^n$ can only be proven under the additional assumption that the

meshes generated by the adaptive algorithm are nested. In Section 6, we give an example for an algorithm ensuring this assumption and adapt the convergence result to this case in Theorem 6.4. Finally, we illustrate our theoretical findings with two numerical tests in Section 7 showing that the adaptive step size control behaves as desired and refines the step size close to switching points.

1.2. **Notation.** If a normed linear space X is continuously embedded in another space Y, we write $X \hookrightarrow Y$. Compact embeddings are denoted by $X \hookrightarrow^c Y$. The dual pairing between X and its dual X^* is denoted by $\langle \cdot, \cdot \rangle_{X^*, X}$, where the subscript is sometimes neglected, if there is no risk for ambiguity. Throughout the paper, c and C denote positive generic constants. Given a Lebesgue measurable set $M \subset \mathbb{R}$, we denote its Lebesgue measure by $|M|$.

1.3. **Standing Assumptions.** We now introduce the precise assumptions on the data in (1.1).

Spaces. The spaces underlying (1.1) are $\mathcal{X}$, $\mathcal{V}$, $\mathcal{W}$, and $\mathcal{Z}$, where $\mathcal{Z}$, $\mathcal{V}$ are Hilbert and $\mathcal{X}$, $\mathcal{W}$ Banach spaces such that

$$\mathcal{Z} \hookrightarrow^c \mathcal{W} \hookrightarrow \mathcal{V} \hookrightarrow \mathcal{X}$$

and $\mathcal{Z}$ is dense in $\mathcal{V}$. Furthermore, $\mathbb{V} \in \mathcal{L}(\mathcal{V}, \mathcal{V}^*)$ is a linear, coercive, and self-adjoint operator, i.e., there is a constant $\gamma > 0$ such that

$$\langle \mathbb{V}v, v \rangle_{\mathcal{V}^*, \mathcal{V}} \geq \gamma \|v\|_{\mathcal{V}}^2 \quad \forall\, v \in \mathcal{V}.$$

We sometimes equip $\mathcal{V}$ with the equivalent norm $\|.\|_{\mathbb{V}} : \mathcal{V} \to \mathbb{R}$ defined by $\|v\|_{\mathbb{V}} = \langle \mathbb{V}v, v \rangle_{\mathcal{V}^*, \mathcal{V}}^{\frac{1}{2}}$. Clearly, $\mathbb{V}$ is invertible and by $\eta \ni \mathcal{V}^* \mapsto \langle \eta, \mathbb{V}^{-1}\eta \rangle_{\mathcal{V}^*, \mathcal{V}}^{\frac{1}{2}} \in \mathbb{R}$ we obtain a norm on $\mathcal{V}^*$, which we denote by $\|\eta\|_{\mathbb{V}^{-1}}$.

Dissipation. For the dissipation $\mathcal{R} : \mathcal{X} \to [0, \infty)$, we assume

(1.2) $\qquad\qquad \mathcal{R}$ is proper, convex, and lower semicontinuous,

(1.3) $\qquad\qquad \mathcal{R}$ is positive 1-homogeneous, i.e., $\mathcal{R}(\lambda v) = \lambda \mathcal{R}(v) \ \forall\, v \in \mathcal{X}, \ \lambda > 0,$

(1.4) $\qquad\qquad \exists\, \rho, R > 0, \ \text{such that } \rho \|v\|_{\mathcal{X}} \leq \mathcal{R}(v) \leq R \|v\|_{\mathcal{V}} \ \forall\, v \in \mathcal{V}.$

In all what follows, we will consider $\mathcal{R}$ as a mapping from $\mathcal{Z}$ to $\mathbb{R}$. Accordingly, its Fenchel conjugate and its convex subdifferential are considered as a mapping and a subset, respectively, in $\mathcal{Z}^*$.

Initial data. The initial state z_0 satisfies $z_0 \in \mathcal{Z}$ and $D_z \mathcal{I}(0, z_0) \in \mathcal{V}^*$.

Energy. The energy functional $\mathcal{I} : [0, T] \times \mathcal{Z} \to \mathbb{R}$ is of the form

(1.5) $$\mathcal{I}(t, z) = \frac{1}{2}\langle Az, z \rangle_{\mathcal{Z}^*, \mathcal{Z}} + \mathcal{F}(z) - f(t, z),$$

where $A \in \mathcal{L}(\mathcal{Z}, \mathcal{Z}^*)$ is coercive and self-adjoint such that there is a constant $\alpha > 0$ so that

$$\langle Az, z \rangle_{\mathcal{Z}^*, \mathcal{Z}} \geq \alpha \|z\|_{\mathcal{Z}}^2 \quad \forall\, z \in \mathcal{Z}.$$

The functional $\mathcal{F} : \mathcal{Z} \to \mathbb{R}$ fulfills

(1.6) $\qquad\qquad \mathcal{F} \in C^2(\mathcal{Z}; \mathbb{R}), \ \mathcal{F} \geq 0$

(1.7) $\qquad\qquad D_z\mathcal{F} \in C^1(\mathcal{Z}, \mathcal{V}^*) \text{ and } \|D_z^2\mathcal{F}(z)v\|_{\mathcal{V}^*} \leq C(1 + \|z\|_{\mathcal{Z}}^q)\|v\|_{\mathcal{Z}} \ \forall\, z, v \in \mathcal{Z},$

with constants $C > 0$ and $q \geq 1$. Moreover, we require the derivate $D_z\mathcal{F}$ to be continuous in the following sense

(1.8) $\qquad\qquad z_k \rightharpoonup z \text{ in } \mathcal{Z} \implies D_z\mathcal{F}(z_k) \rightharpoonup D_z\mathcal{F}(z) \text{ in } \mathcal{Z}^* \text{ for } k \to \infty.$

The time dependent part $f : [0, T] \times \mathcal{Z} \to \mathbb{R}$ is supposed to satisfy the following properties

(1.9) $\qquad\qquad f \in C^1([0, T] \times \mathcal{Z}; \mathbb{R}),$

(1.10) $f(t, z) \leq c(\|z\|_{\mathcal{Z}} + 1), \ |\partial_t f(t, z)| \leq \mu(\|z\|_{\mathcal{Z}} + 1)$

(1.11) $|\langle D_z f(t_1, z_1) - D_z f(t_2, z_2), v \rangle_{\mathcal{V}^*, \mathcal{V}}| \leq \nu(|t_1 - t_2| + \|z_1 - z_2\|_{\mathcal{W}})\|v\|_{\mathcal{V}}$

for all $t, t_1, t_2 \in [0, T]$, $z, z_1, z_2 \in \mathcal{Z}$, and $v \in \mathcal{V}$ with constants $c, \mu, \nu > 0$. Moreover, we require that

(1.12) $f(t, .)$ is weakly continuous for all $t \in [0, T]$,

(1.13) $\partial_t f(t_k, z_k) \to \partial_t f(t, z)$ holds for all sequences $t_k \to t$ in $\mathbb{R}$ and $z_k \rightharpoonup z$ in $\mathcal{Z}$ for $k \to \infty$.

We end this section with several useful results that immediately follow from from the above assumptions on the energy.

Lemma 1.1. *The functionals $\mathcal{F}$ and f are weakly continuous in the following sense:*

$$z_k \rightharpoonup z \ in \ \mathcal{Z} \implies \mathcal{F}(z_k) \to \mathcal{F}(z),$$
$$t_k \to t \ in \ \mathbb{R}, \ z_k \rightharpoonup z \ in \ \mathcal{Z} \implies f(t_k, z_k) \to f(t, z),$$

as $k \to \infty$.

Proof. In view of the required regularity of $\mathcal{F}$, we can apply the mean value theorem, which yields the existence of $\zeta_k \in (0, 1)$ such that

$$\|\mathcal{F}(z_k) - \mathcal{F}(z)\|_{\mathcal{Z}} \leq |\langle D_z \mathcal{F}(z), z_k - z \rangle_{\mathcal{V}^*, \mathcal{V}}| + \frac{1}{2}|D_z^2 \mathcal{F}(z + \zeta_k(z_k - z))(z_k - z)^2|$$
$$\leq \|D_z \mathcal{F}(z)\|_{\mathcal{V}^*}\|z_k - z\|_{\mathcal{V}} + \frac{C}{2}(1 + \|z + \zeta_k(z_k - z)\|_{\mathcal{Z}}^q)\|z_k - z\|_{\mathcal{Z}}\|z_k - z\|_{\mathcal{V}}$$
$$\to 0, \ k \to \infty,$$

where we exploited (1.7) and the weak convergence of $\{z_k\}$ in $\mathcal{Z}$, which gives the boundedness of $\{z_k\}$ in $\mathcal{Z}$ and, due to the compactness of $\mathcal{Z} \hookrightarrow \mathcal{V}$ by assumption, the strong convergence $z_k \to z$ in $\mathcal{V}$.

For the second statement, we argue similarly. Again the mean value theorem implies the existence of an intermediate value ζ_k such that $f(t_k, z_k) = \partial_t f(\zeta_k, z_k)(t_k - t) + f(t, z_k)$. Then, (1.12) and (1.13) immediately imply the assertion. $\square$

Lemma 1.2. *The energy functional $\mathcal{I}$ satisfies $\mathcal{I} \in C^1([0, T] \times \mathcal{Z}; \mathbb{R})$ and*

(1.14) $t_k \to t \ in \ \mathbb{R}, \ z_k \rightharpoonup z \ in \ \mathcal{Z} \implies \mathcal{I}(t, z) \leq \liminf_{k \to \infty} \mathcal{I}(t_k, z_k),$

(1.15) $t_k \to t \ in \ \mathbb{R}, \ z_k \rightharpoonup z \ in \ \mathcal{Z} \implies D_z \mathcal{I}(t_k, z_k) \rightharpoonup D_z \mathcal{I}(t, z) \ in \ \mathcal{Z}^* \ for \ k \to \infty.$

Proof. The differentiability of $\mathcal{I}$ follows immediately from the required regularity of f and $\mathcal{F}$. The lower semicontinuity in (1.14) is an immediate consequence of Lemma 1.1 and the coercivity of A. To prove (1.15), observe that (1.11) and the compactness of $\mathcal{Z} \hookrightarrow \mathcal{W}$ imply that $D_z f(t_k, z_k)$ converge strongly in $\mathcal{V}^*$ and thus even more in $\mathcal{Z}^*$. Together with (1.8) and the linearity of A, this gives the assertion. $\square$

Moreover, exploiting the coercivity of A together with (1.6) and (1.10) gives the lower bound estimate

(1.16) $\mathcal{I}(t, z) \geq \frac{\alpha}{2}\|z\|_{\mathcal{Z}}^2 - c(\|z\|_{\mathcal{Z}} + 1) \geq \|z\|_{\mathcal{Z}} - \tilde{c} \quad \forall t \in [0, T], \ z \in \mathcal{Z},$

where $\tilde{c} = \frac{(c+1)^2}{2\alpha} + c$. Hence, combining this and (1.10) results in

$$|\partial_t \mathcal{I}(t, z)| \leq \mu(\|z\|_{\mathcal{Z}} + 1) \leq \mu(\mathcal{I}(t, z) + \beta),$$

where $\beta = \tilde{c} + 1$. Now integrating over (s, t) and (t, s), respectively, and using Gronwall's lemma leads to

Lemma 1.3. *For all $z \in \mathcal{Z}$ and and all $s, t \in [0, T]$, there holds*

$$(1.17) \qquad \mathcal{I}(t, z) + \beta \leq (\mathcal{I}(s, z) + \beta) \exp(\mu|t - s|),$$

$$(1.18) \qquad |\partial_t \mathcal{I}(t, z)| \leq \mu(\mathcal{I}(s, z) + \beta) \exp(\mu|t - s|).$$

Example 1.4. Let us give an example for a rate-independent system fulfilling the above assumptions. For this purpose, let $\Omega \subset \mathbb{R}^d$, $d \in \mathbb{N}$, be a bounded Lipschitz-domain in the sense of [5, Def. 1.2.1.2]. Then we choose for the spaces

$$\mathcal{Z} := H_0^1(\Omega), \quad \mathcal{W} = \mathcal{V} := L^2(\Omega), \quad \mathcal{X} := L^1(\Omega),$$

for the dissipation functional

$$\mathcal{R}(z) := \|z\|_{L^1(\Omega)},$$

and for the energy

$$(1.19) \qquad A := -\triangle, \quad \mathcal{F}(z) := \int_\Omega g(z)(x)\, \mathrm{d}x, \quad f(t, z) := \langle \ell(t), z \rangle_{\mathcal{V}^*, \mathcal{V}},$$

where $g : L^p(\Omega) \to L^2(\Omega)$, $p < 2d/(d-2)$, is a Nemyzkii operator that is twice continuously Fréchet-differentiable and $\ell : [0, T] \to \mathcal{V}^*$ is continuously Fréchet-differentiable, too. Then, it is easily seen by means of Sobolev embeddings that the example fits into the above setting, provided that $\frac{\mathrm{d}}{\mathrm{d}t}\ell$ is Lipschitz continuous and g is non-negative and satisfies

$$(1.20) \qquad \|g''(z)v\|_{L^2(\Omega)} \leq C(1 + \|z\|_{L^p(\Omega)}^q)\|v\|_{L^p(\Omega)}.$$

We will come back to an example of this type in our numerical experiments in Section 7 below.

2. Solution Concepts in Brief

Before we present our adaptive time stepping scheme, let us introduce the notion of solution of (1.1) underlying this scheme and give a brief overview of the various solution concepts for (1.1) without claiming to be exhaustive. For a more detailed survey of the various notions of solutions to rate independent systems, we refer to [12] and the references therein.

The most natural solution concept for (1.1) is probably the notion of a *differential solution*. Here, one seeks for a function $z \in W^{1,1}(0, T; \mathcal{Z})$ with $z(0) = z_0$ satisfying the differential inclusion in (1.1) almost everywhere. It however turns out that, if the energy is not uniformly convex, the existence of such a solution can in general not be expected, as e.g. the counterexample in [12, Example 1.8.1] illustrates. For this reason, Alexander Mielke and his co-authors came up with the meanwhile classical concept of *global energetic solutions*, see [13, 15, 14]. This concept involves a global stability condition together with an energy balance identity, and existence of solutions of this type can be shown under rather mild assumptions on dissipation and energy, in particular without assuming the energy functional to be convex, we refer to [12, Thm. 2.1.6]. However, caused by the global stability condition, global energetic solutions have the tendency to jump as early as possible to global minimizers in the energy landscape and in this way may cross energy barriers, which can be seen unphysical in applications.

A solution concept that overcomes this drawback is the concept of *parametrized balanced viscosity (BV) solutions*, which also forms the basis for our adaptive time stepping scheme. To the best of our knowledge, it was first introduced in [4], but has by now been analyzed by various authors in multiple aspects, we only refer to [11] and the reference therein. As the denomination indicates, we focus on BV solutions in parametrized form, where physical time and state variable are given as functions of a parameter s. For our particular setting, the precise definition of a parametrized BV solution reads as follows:

Definition 2.1 ($\mathbb{V}$-parametrized BV solution)**.** We call a tuple $(\hat{t}, \hat{z})$ $\mathbb{V}$-parametrized BV solution of the rate-independent system (1.1), if there exists $S \geq T$ such that the following is fulfilled:

- Regularity:

$$(2.1) \qquad \hat{t} \in W^{1,\infty}(0,S), \quad \hat{z} \in W^{1,\infty}(0,S;\mathcal{V}) \cap L^\infty(0,S;\mathcal{Z}),$$

- Initial and end time condition:

$$(2.2) \qquad \hat{z}(0) = z_0, \ \hat{t}(0) = 0, \ \hat{t}(S) = T,$$

- Complementary condition:

$$(2.3a) \qquad \hat{t}'(s) \geq 0, \ \hat{t}'(s) + \|\hat{z}'(s)\|_\mathbb{V} \leq 1 \qquad\qquad \text{f.a.a. } s \in (0,S),$$

$$(2.3b) \qquad \hat{t}'(s)\,\overline{\text{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(\hat{t}(s),\hat{z}(s)),\partial\mathcal{R}(0)\} = 0 \qquad\qquad \text{f.a.a. } s \in (0,S),$$

where $\overline{\text{dist}}_{\mathcal{V}^*}\{\cdot,\partial\mathcal{R}(0)\} : \mathcal{Z}^* \to [0,\infty]$ is defined by

$$(2.4) \qquad \overline{\text{dist}}_{\mathcal{V}^*}\{\eta,\partial\mathcal{R}(0)\} := \inf\{\|w\|_{\mathbb{V}^{-1}} : w \in \mathcal{V}^* \cap (\partial\mathcal{R}(0) - \eta)\}$$

(with the usual convention that $\inf \emptyset = \infty$),

- Energy equality:

$$(2.5) \qquad \begin{aligned} \mathcal{I}(\hat{t}(s),\hat{z}(s)) &+ \int_0^s \mathcal{R}(\hat{z}'(\sigma)) + \|\hat{z}'(\sigma)\|_\mathbb{V}\,\overline{\text{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(\hat{t}(\sigma),\hat{z}(\sigma)),\partial\mathcal{R}(0)\}\,d\sigma \\ &= \mathcal{I}(0,z_0) + \int_0^s \partial_t\mathcal{I}(\hat{t}(\sigma),\hat{z}(\sigma))\hat{t}'(\sigma)d\sigma \quad \forall s \in [0,S]. \end{aligned}$$

We call a $\mathbb{V}$-parametrized solution non-degenerate, if there is a constant $\delta > 0$ such that $\hat{t}'(s) + \|\hat{z}'(s)\|_\mathbb{V} \geq \delta$ a.e. in $(0,S)$. If $\delta = 1$, then the solution is called normalized $\mathbb{V}$-parametrized BV solution.

Some words concerning this definition are in order. The interpretation of Definition 2.1 is probably best understood, if one introduces the function $\lambda : [0,S] \to [0,\infty]$ by

$$(2.6) \qquad \lambda(s) := \begin{cases} \frac{1}{\|\hat{z}'(s)\|_\mathbb{V}}\overline{\text{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(\hat{t}(s),\hat{z}(s)),\partial\mathcal{R}(0)\}, & \text{if } \hat{z}'(s) \neq 0, \\ = 0, & \text{else} \end{cases}$$

and applies the chain rule to (2.5) to obtain the pointwise relation

$$(2.7) \qquad 0 \in \partial\mathcal{R}(\hat{z}'(s)) + \lambda(s)\hat{z}'(s) + D_z\mathcal{I}(\hat{t}(s),\hat{z}(s)),$$

cf. [11], which is (1.1) enriched with the additional viscous term $\lambda(s)\hat{z}'(s)$. We observe that, thanks to the complementarity relation in (2.3b), this additional term only appears, if $\hat{t}'(s) = 0$, i.e., if the physical time stands still, which corresponds to a discontinuous jump of the system. In case of a jump, the above equation thus describes the viscous transition of the state variable $\hat{z}$ through the complement of the set of local stability $\{(t,z) \in \mathbb{R} \times \mathcal{Z} : -D_z\mathcal{I}(t,z) \in \partial\mathcal{R}(0)\}$. Of course, one may re-parametrize the curve $s \mapsto (\hat{t}(s),\hat{z}(s))$ such that a parametrized BV solution is intrinsically non-unique. Here, we have chosen a parametrization by arc length w.r.t. the $\mathbb{V}$-norm, another parametrization that is often used involves the so-called vanishing viscosity contact potential. One may also define BV solutions in the physical time regime without parametrization, leading to a modified energy balance. For more details on (parametrized) BV solutions, we refer to [11] and the references therein.

Remark 2.2 ([7, Lemma 5.6], [17, Lemma 2.4.6]). It is to be noted that, provided a chain rule can be applied, an energy inequality is sufficient for (2.5) to hold. To be more precise, if $(\hat{t},\hat{z}) \in W^{1,\infty}(0,S) \times W^{1,1}(0,S;\mathcal{Z})$ with $\overline{\text{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(\hat{t}(s),\hat{z}(s)),\partial\mathcal{R}(0)\} < \infty$ f.a.a. $s \in (0,S)$, then the inequality

$$(2.8) \qquad \begin{aligned} 0 \leq \mathcal{I}(\hat{t}(s_2),\hat{z}(s_2)) &- \mathcal{I}(\hat{t}(s_1),\hat{z}(s_1)) - \int_{s_1}^{s_2} \partial_t\mathcal{I}(\hat{t}(\sigma),\hat{z}(\sigma))\hat{t}'(\sigma)d\sigma \\ &+ \int_{s_1}^{s_2} \mathcal{R}(\hat{z}'(\sigma)) + \|\hat{z}'(\sigma)\|_\mathbb{V}\,\overline{\text{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(\hat{t}(\sigma),\hat{z}(\sigma)),\partial\mathcal{R}(0)\}d\sigma \end{aligned}$$

is automatically fulfilled for all $s_1, s_2 \in [0, S]$ with $s_1 \leq s_2$. The required chain rule for the energy $\mathcal{I}$ is applicable here thanks to our standing assumptions, cf., e.g., [17, Lemma A.2.5].

While the existence of parametrized BV solutions is frequently established by a vanishing viscosity approach (as the name indicates), it is also possible to prove their existence by means of tailored time stepping schemes and a limit analysis for time step size tending to zero. The corresponding schemes are known as *local incremental minimization* or *local stationary scheme*. They have been analyzed in [4] for the finite dimensional case and in [6, 9, 18, 8] for the infinite dimensional setting. Moreover, as the numerical experiments in [9, 18] demonstrate, these schemes can also be realized in practice and used for numerical computations. So far however, all these schemes employ a uniform step size for the curve parameter s. In the following, we present a method that chooses an adaptive step size, but still can be shown to converge to a $\mathbb{V}$-parametrized BV solution in the sense of Definition 2.1.

3. An Adaptive Incremental Stationary Scheme

For the construction of our adaptive algorithm, we need the following

Definition 3.1. By $I_{\tau_k} : \mathcal{Z} \to \mathbb{R}$, we denote the indicator functional of $B_{\mathbb{V}}(0, \tau_k)$, i.e.,

$$I_{\tau_k}(z) = \begin{cases} 0, & \text{if } \|z\|_{\mathbb{V}} \leq \tau_k \\ \infty, & \text{else.} \end{cases}$$

With I_{τ_k} at hand, the algorithm then reads as follows:

Algorithm 3.2 (Adaptive Incremental Stationary Scheme).

1: Set $t_0 = 0$, $s_0 = 0$, $k = 1$ and choose $\texttt{tol} > 0$, $\tau_k > 0$.
2: **while** $t_{k-1} < T$ **do**
3:　　Calculate a solution z_k of

$$(3.1) \qquad 0 \in \partial(\mathcal{R} + I_{\tau_k})(z_k - z_{k-1}) + D_z\mathcal{I}(t_{k-1}, z_k),$$

　　　　which, moreover, fulfills

$$(3.2) \qquad \mathcal{I}(t_{k-1}, z_k) + \mathcal{R}(z_k - z_{k-1}) \leq \mathcal{I}(t_{k-1}, z_{k-1}).$$

4:　　Time update:

$$t_k = t_{k-1} + \tau_k - \|z_k - z_{k-1}\|_{\mathbb{V}}, \quad s_k = \sum_{i=1}^{k} \tau_i.$$

5:　　Define for $s \in [s_{k-1}, s_k)$ the affine and constant interpolants

$$\hat{z}(s) = z_{k-1} + \frac{s - s_{k-1}}{\tau_k}(z_k - z_{k-1}), \quad \hat{t}(s) = t_{k-1} + \frac{s - s_{k-1}}{\tau_k}(t_k - t_{k-1}),$$

$$\bar{z}(s) = z_k, \qquad\qquad\qquad\qquad \underline{t}(s) = t_{k-1}.$$

6:　　Compute the integrals

$$(3.3) \qquad I_1^k = \int_{s_{k-1}}^{s_k} \hat{t}'(s)\overline{\mathrm{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(\hat{t}(s), \hat{z}(s)), \partial\mathcal{R}(0)\}ds$$

$$(3.4) \qquad I_2^k = \int_{s_{k-1}}^{s_k} \Big[\langle D_z\mathcal{I}(\hat{t}(s), \hat{z}(s)) - D_z\mathcal{I}(\underline{t}(s), \bar{z}(s)), \hat{z}'(s) \rangle_{\mathcal{Z}^*, \mathcal{Z}}$$
$$+ \|\hat{z}'(s)\|_{\mathbb{V}}\Big(\overline{\mathrm{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(\hat{t}(s), \hat{z}(s)), \partial\mathcal{R}(0)\}$$
$$- \overline{\mathrm{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(\underline{t}(s), \bar{z}(s)), \partial\mathcal{R}(0)\} \Big) \Big] ds.$$

7: **if** $I_1^k < \texttt{tol}$ and $I_2^k < \texttt{tol}$ **then**
8:      Set $k = k + 1$.
9:      **if** $I_1^k < \frac{\texttt{tol}}{2}$ and $I_2^k < \frac{\texttt{tol}}{2}$ **then**
10:         Set $\tau_{k+1} = 2\,\tau_k$.
11:     **end if**
12: **else**
13:     Set $\tau_k = \frac{\tau_k}{2}$ and **go to** Step 3.
14: **end if**
15: **end while**

It is to be noted that the above algorithm coincides with the local stationary scheme from [17, 18] except for the adaptive choice of the step size τ_k in Steps 6–14. As mentioned before, in contrast to Algorithm 3.2, the step size in [17, 18] is fixed throughout the whole iteration.

Let us first show that each iteration of the above algorithm is well posed.

Lemma 3.3. *Given* $t_k \geq 0$ *and* $z_{k-1} \in \mathcal{Z}$, *there exists at least one* $z_k \in \mathcal{Z}$ *fulfilling* (3.1) *and* (3.2).

Proof. Owing to the differentiability of $\mathcal{I}$ and the convexity of $\mathcal{R}$, (3.1) is nothing else than the necessary optimality condition of

$$(3.5) \qquad z_k \in \mathrm{argmin}\{\mathcal{I}(t_{k-1}, z) + \mathcal{R}(z - z_{k-1}) : z \in \mathcal{Z}, \|z - z_{k-1}\|_\mathbb{V} \leq \tau\}.$$

Now, as $\mathcal{Z}$ is reflexive, $\mathcal{I}(t_{k-1}, \cdot)$ is weakly lower semicontinuous by (1.14) and the same holds for $\mathcal{R}$, since it is convex and continuous, and $\mathcal{I}$ is radially unbounded by (1.16), the direct method gives the existence of a solution to (3.5) and thus to (3.1). As it is a global minimum and z_{k-1} is feasible, (3.2) follows immediately. $\qquad\square$

Remark 3.4. The original incremental minimization scheme from [4] is based on (3.5) instead of (3.1)–(3.2). However, any optimization algorithm can hardly be guaranteed to solve a non-convex problem up to (global) optimality. What can be shown instead is convergence to stationary points, at least under suitable assumptions. From an algorithmic point of view, it therefore makes more sense to require the stationarity condition in (3.1) instead of solving (3.5). The second condition in (3.2) is not restrictive in this context, provided that one employs a descent method, e.g., via backtracking or other line search methods.

Remark 3.5. It is to be noted that, due to the indicator functional I_{τ_k} involved in (3.1), there always holds $\|z_k - z_{k-1}\|_\mathbb{V} \leq \tau_k$ such that $t_k \geq t_{k-1}$ by the time update in Step 4, i.e., the algorithm does not go backwards in time.

As we will see below, the integrals in Step 6 are well defined, see Lemma 4.9 below. Therefore, each iteration of the algorithm is well defined, too. Moreover, the overall iteration is well posed in the sense that the end time T is reached after finitely many iterations, as we will see in Proposition 5.1 below. Let us shortly comment on the update of the step size in Steps 6–14. As we shall see in the next section, the integrals I_1^k, I_2^k calculated in Step 6 measure the residuum that arises, if one inserts the affine interpolants from Step 5 in the complementarity relation in (2.3b) and in the energy identity in (2.5), see Lemma 4.11 below. The aim of the algorithm is thus to drive this residuum below a given tolerance $\texttt{tol}$. However, given a sequence $\{\texttt{tol}\}_{n\in\mathbb{N}}$ tending to zero with asscoiated affine interpolants $\{(\hat{t}_n, \hat{z}_n)\}_{n\in\mathbb{N}}$ resulting from Algorithm 3.2, there is in general no hope that $(\hat{t}_n, \hat{z}_n)$ will converge (strongly) to a $\mathbb{V}$-parametrized BV solution, although the residuum tends to zero. This is due to the non-uniqueness of $\mathbb{V}$-parametrized BV solutions and is even the case for a uniform refinement of the step size, as demonstrated by the counterexample in [10, Section 2.4]. Nevertheless, for a uniform step size, one can show that subsequences converge weakly to a $\mathbb{V}$-parametrized BV solution and, as we will see in the following, the same holds true for our adaptive algorithm.

4. Preliminaries and Known Results

For the final proof of our convergence results in Sections 5 and 6, we need several auxiliary results that are established in the following. We point out that most of the proofs are very similar to the case with a uniform step size as discussed in [6, 9]. Therefore, we omit a proof, when it is entirely analogous, or postpone it to the appendix.

Moreover, it is important to note that all proofs of this section do *neither* depend on the choice of the step size *nor* on the tolerance, i.e., if we choose an update rule for the steps size, that differs from the one in Steps 7–14, the results of this section still apply. In fact, an inspection of the proofs in this section shows that the arguments are only based on the construction of z_k and t_k in Step 3 and 4, see also Remark 4.13 below.

We start with a reformulation of (3.1), which follows from classical results of convex analysis.

Lemma 4.1 ([6, Proposition 2.2], [9, Lemma 3.1], [17, Lemma 3.2.1]). *Given $\tau_k > 0$, $t_{k-1} \geq 0$, and $z_k \in \mathcal{Z}$, the differential inclusion in* (3.1) *is equivalent to the existence of a Lagrange multiplier $\lambda_k \geq 0$ such that*

$$(4.1) \qquad \lambda_k(\|z_k - z_{k-1}\|_{\mathbb{V}} - \tau_k) = 0$$

$$(4.2) \qquad \tau_k \overline{\mathrm{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(t_{k-1}, z_k), \partial\mathcal{R}(0)\} = \lambda_k\|z_k - z_{k-1}\|_{\mathbb{V}}^2$$

$$(4.3) \qquad \mathcal{R}(z_k - z_{k-1}) + \tau_k \overline{\mathrm{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(t_{k-1}, z_k), \partial\mathcal{R}(0)\} = \langle -D_z\mathcal{I}(t_{k-1}, z_k), z_k - z_{k-1}\rangle_{\mathcal{Z}^*, \mathcal{Z}}$$

$$(4.4) \qquad \mathcal{R}(v) \geq -\langle\lambda_k\mathbb{V}(z_k - z_{k-1}) + D_z\mathcal{I}(t_{k-1}, z_k), v\rangle_{\mathcal{Z}^*, \mathcal{Z}} \quad \forall v \in \mathcal{Z}.$$

The next lemma has been proven in various publications for the case of a uniform step size. Its proof is only based on the condition (3.2) and on the standing assumptions on $\mathcal{R}$ and $\mathcal{I}$, in particular Lemma 1.3, and therefore readily carries over to an adaptive choice of τ_k as in Algorithm 3.2.

Lemma 4.2 ([12, Theorem 2.1.5], [6, Proposition 2.1], [9, Lemma 3.2], [17, Lemma 3.2.4]). *For all $k \in \mathbb{N}$, the iterates (t_k, z_k) of Algorithm 3.2 satisfy*

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

If we combine the estimate (1.16) with the above lemma and the non-negativity of $\mathcal{R}$, we obtain

$$\|z_k\|_{\mathcal{Z}} \leq \mathcal{I}(t_k, z_k) + \tilde{c} \leq (\beta + \mathcal{I}(0, z_0))\exp(\mu T) + \tilde{c},$$

i.e., we have the following

Corollary 4.3. *For every choice of step sizes $\tau_k > 0$, the iterates of Algorithm 3.2 satisfy*

$$(4.5) \qquad\qquad\qquad\qquad \|z_k\|_{\mathcal{Z}} \leq C \quad \forall k \in \mathbb{N}$$

with a constant $C > 0$ independent of τ_k.

Using our assumptions (1.7) and (1.11) on the components $\mathcal{F}$ and f of the energy functional in combination with the lower bound of the dissipation functional from (1.4) and Ehrling's lemma, one finds the following

Lemma 4.4 ([9, Lemma 3.5], [17, Lemma 3.2.10]). *Let $\epsilon > 0$ be arbitrary. Then, for all $r > 0$, there exists a constant $C_{\epsilon,r} > 0$ such that*

$$(4.6) \qquad |\langle D_z\mathcal{F}(z_1) - D_z\mathcal{F}(z_2), z_1 - z_2\rangle_{\mathcal{V}^*, \mathcal{V}}| \leq \epsilon\|z_1 - z_2\|_{\mathcal{Z}}^2 + C_{\epsilon,r}\,\mathcal{R}(z_1 - z_2)\|z_1 - z_2\|_{\mathcal{V}}$$

holds true for all $z_1, z_2 \in B_{\mathcal{Z}}(0, r)$. Moreover, there is a constant $c_\epsilon > 0$ such that

$$(4.7) \qquad \begin{aligned} |\langle D_zf(t_1, z_1) &- D_zf(t_2, z_2), v\rangle_{\mathcal{V}^*, \mathcal{V}}| \\ &\leq \nu\,|t_1 - t_2|\,\|v\|_{\mathcal{V}} + \epsilon\|z_1 - z_2\|_{\mathcal{Z}}\,\|v\|_{\mathcal{Z}} + c_\epsilon\,\mathcal{R}(z_1 - z_2)\|v\|_{\mathcal{V}} \end{aligned}$$

holds for all $t_1, t_2 \in [0, T]$, $z_1, z_2, v \in \mathcal{Z}$.

Lemma 4.5. *There exists a constant $C > 0$ such that, for every choice of the step size τ_k, $k \in \mathbb{N}$, there holds*

$$(4.8) \qquad \lambda_{k+1} \|z_{k+1} - z_k\|_{\mathbb{V}} + \sum_{i=0}^{k} \|z_{i+1} - z_i\|_{\mathcal{Z}} \leq C$$

for all $k \geq 0$.

Proof. The proof is similar to the ones of [9, Proposition 3.6] and [6, Proposition 2.3], but, since we need some parts of the proof for later reference and, in particular, the estimate of $\sum_{i=0}^{k} \|z_{i+1} - z_i\|_{\mathcal{Z}}$ can be seen as the basis for the overall convergence analysis, we present the arguments in detail for convenience of the reader.

Let $k \in \mathbb{N}$ be arbitrary. We start with inserting $v = z_{k+1} - z_k$ in (4.4) in order to obtain

$$(4.9) \qquad \mathcal{R}(z_{k+1} - z_k) \geq -\langle \lambda_k \mathbb{V}(z_k - z_{k-1}), z_{k+1} - z_k \rangle_{\mathcal{V}^*, \mathcal{V}} - \langle D_z \mathcal{I}(t_{k-1}, z_k), z_{k+1} - z_k \rangle_{\mathcal{Z}^*, \mathcal{Z}}$$

Rewriting the statements (4.2) and (4.3) for $k = k + 1$ and inserting them into each other yields

$$(4.10) \qquad \mathcal{R}(z_{k+1} - z_k) + \lambda_{k+1} \|z_{k+1} - z_k\|_{\mathbb{V}}^2 = \langle -D_z \mathcal{I}(t_k, z_{k+1}), z_{k+1} - z_k \rangle_{\mathcal{Z}^*, \mathcal{Z}}.$$

Combining (4.9) and (4.10) leads to

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

By inserting the definition of $\mathcal{I}$ from (1.5) and using the coercivity of A by assumption, we obtain the inequality

$$(4.11) \qquad \begin{aligned} \lambda_{k+1} \|z_{k+1} - z_k\|_{\mathbb{V}}^2 &- \lambda_k \|z_k - z_{k-1}\|_{\mathbb{V}} \|z_{k+1} - z_k\|_{\mathbb{V}} + \alpha \|z_{k+1} - z_k\|_{\mathcal{Z}}^2 \\ &\leq \langle D_z \mathcal{F}(z_k) - D_z \mathcal{F}(z_{k+1}), z_{k+1} - z_k \rangle_{\mathcal{V}^*, \mathcal{V}} \\ &\quad + \langle D_z f(t_{k-1}, z_k) - D_z f(t_k, z_{k+1}), z_{k+1} - z_k \rangle_{\mathcal{V}^*, \mathcal{V}} \end{aligned}$$

Now we apply (4.6) and (4.7), both with $\epsilon = \frac{\alpha}{4}$, to estimate the right hand side. Note at this point that, due to (4.5), the iterates $(z_i)_{i \in \mathbb{N}}$ are bounded in $\mathcal{Z}$ so that the requirements for (4.6) are satisfied. This leads to the estimate

$$(4.12) \qquad \begin{aligned} \lambda_{k+1} \|z_{k+1} - z_k\|_{\mathbb{V}}^2 &- \lambda_k \|z_k - z_{k-1}\|_{\mathbb{V}} \|z_{k+1} - z_k\|_{\mathbb{V}} + \frac{\alpha}{2} \|z_{k+1} - z_k\|_{\mathcal{Z}}^2 \\ &\leq (C_{\alpha/4} + c_{\alpha/4}) \mathcal{R}(z_{k+1} - z_k) \|z_{k+1} - z_k\|_{\mathcal{V}} + \nu \, (t_k - t_{k-1}) \|z_{k+1} - z_k\|_{\mathcal{V}}. \end{aligned}$$

By dividing this by $\|z_{k+1} - z_k\|_{\mathbb{V}}$ and using the continuous embedding $\mathcal{Z} \hookrightarrow \mathcal{V}$, we end up with

$$\lambda_{k+1} \|z_{k+1} - z_k\|_{\mathbb{V}} - \lambda_k \|z_k - z_{k-1}\|_{\mathbb{V}} + c \|z_{k+1} - z_k\|_{\mathcal{Z}} \leq C(\mathcal{R}(z_{k+1} - z_k) + t_k - t_{k-1}).$$

Now let us sum up this inequality with respect to k in order to conclude

$$(4.13) \qquad \lambda_{k+1} \|z_{k+1} - z_k\|_{\mathbb{V}} + c \sum_{i=1}^{k} \|z_{i+1} - z_i\|_{\mathcal{Z}} \leq \lambda_1 \|z_1 - z_0\|_{\mathbb{V}} + C \Big(t_k + \sum_{i=1}^{k} \mathcal{R}(z_{i+1} - z_i) \Big).$$

Next for estimating the term $\lambda_1 \|z_1 - z_0\|_{\mathbb{V}}$ we consider (4.10) for $k = 0$, which reads

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

which, in view of the non-negativity of $\mathcal{R}$, in turn yields

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

For the first part on the left-hand side, we argue in the same way as above, i.e., we split the derivative of the energy into its individuals parts, rearrange terms and use the coercivity of A, as well as (4.6) and (4.7) with $\epsilon = \alpha/4$ as above in order to obtain

$$(4.14) \qquad \frac{\alpha}{2}\|z_1 - z_0\|_{\mathcal{Z}}^2 + \lambda_1 \|z_1 - z_0\|_{\mathbb{V}}^2 \leq C(\mathcal{R}(z_1 - z_0) + \|D_z\mathcal{I}(0, z_0)\|_{\mathcal{V}^*})\|z_1 - z_0\|_{\mathcal{V}}.$$

By dividing by $\|z_1 - z_0\|_{\mathbb{V}}$ and exploiting the continuous embedding $\mathcal{V} \hookrightarrow \mathcal{Z}$, it follows that

$$(4.15) \qquad c\|z_1 - z_0\|_{\mathcal{Z}} + \lambda_1 \|z_1 - z_0\|_{\mathbb{V}} \leq C(\mathcal{R}(z_1 - z_0) + \|D_z\mathcal{I}(0, z_0)\|_{\mathcal{V}^*}).$$

Now inserting this into (4.13) yields

$$\lambda_{k+1}\|z_{k+1} - z_k\|_{\mathbb{V}} + c\sum_{i=0}^k \|z_{i+1} - z_i\|_{\mathcal{Z}} \leq C\Big(t_k + \sum_{i=0}^k \mathcal{R}(z_{i+1} - z_i) + \|D_z\mathcal{I}(0, z_0)\|_{\mathcal{V}^*}\Big).$$

From this, in combination with Lemma 4.2, the lower bound (1.16), the termination criterion of the algorithm, and our standing assumptions on the initial state, we deduce the claim. Note that the case $k = 0$ is covered by (4.15) by the same arguments. $\qquad \square$

Lemma 4.6. *There exists a constant $\tau_{\max} > 0$, independent of the chosen tolerance, such that $\max_{k\in\mathbb{N}} \tau_k \leq \tau_{\max}$.*

Proof. According to the termination criterion and the time update in Step 4 of the algorithm, all but the last iterations fulfill

$$(4.16) \qquad T \geq t_k = t_0 + \sum_{i=1}^k \tau_i - \sum_{i=1}^k \|z_i - z_{i-1}\|_{\mathbb{V}},$$

which, together with (4.8), yields the assertion for all k except for the last iteration (if there is one at all, which we will see in fact in Proposition 5.1 below). Since the last step size can just be twice as large as the second to last one by Step 10, the bound carries over to all step sizes. $\qquad \square$

Lemma 4.7. *There is a constant $C > 0$ such that, for all $k \in \mathbb{N}$, there holds*

$$(4.17) \qquad \overline{\mathrm{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(t_{k-1}, z_k), \partial\mathcal{R}(0)\} \leq C,$$

$$(4.18) \qquad \overline{\mathrm{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(\hat{t}(s), \hat{z}(s)), \partial\mathcal{R}(0)\} \leq C \quad \forall s \in [s_{k-1}, s_k).$$

Proof. For the first claim we make use of (4.1) and (4.2), which gives

$$\overline{\mathrm{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(t_{k-1}, z_k), \partial\mathcal{R}(0)\} = \lambda_k\|z_k - z_{k-1}\|_{\mathbb{V}}$$

such that (4.17) follows from (4.8).

To prove (4.18), we first note that the structure of $\mathcal{I}$ and the assumption in (1.11) imply

$$(4.19) \qquad \begin{aligned} &\overline{\mathrm{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(t_k, z_k), \partial\mathcal{R}(0)\} \\ &= \overline{\mathrm{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(t_{k-1}, z_k) + D_z f(t_k, z_k) - D_z f(t_{k-1}, z_k), \partial\mathcal{R}(0)\} \\ &\leq \overline{\mathrm{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(t_{k-1}, z_k), \partial\mathcal{R}(0)\} + \|D_z f(t_k, z_k) - D_z f(t_{k-1}, z_k)\|_{\mathbb{V}^{-1}} \\ &\leq \overline{\mathrm{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(t_{k-1}, z_k), \partial\mathcal{R}(0)\} + C\tau_k. \end{aligned}$$

Hence, thanks to Lemma 4.6, $\overline{\mathrm{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(t_k, z_k), \partial\mathcal{R}(0)\}$ is bounded for all $k \geq 1$. Due to $D_z\mathcal{I}(t_0, z_0) = D_z\mathcal{I}(0, z_0) \in \mathcal{V}^*$ by our standing assumption, the claim also holds for $k = 0$. Next we rewrite the affine interpolants as

$$\hat{z}(s) = \lambda(s)z_{k+1} + (1 - \lambda(s))z_k \quad \text{and} \quad \hat{t}(s) = \lambda(s)t_{k+1} + (1 - \lambda(s))t_k,$$

where λ is given by $\lambda(s) = \frac{s-s_k}{\tau_{k+1}} \in [0,1)$ for $s \in [s_k, s_{k+1})$. The convexity of the distance then gives

$$\overline{\operatorname{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(\hat{t}(s), \hat{z}(s)), \partial\mathcal{R}(0)\}$$

$$\leq \lambda(s)\overline{\operatorname{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(t_{k+1}, z_{k+1}), \partial\mathcal{R}(0)\} + (1-\lambda(s))\overline{\operatorname{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(t_k, z_k), \partial\mathcal{R}(0)\}$$

$$+ \lambda(s)\|D_z\mathcal{F}(z_{k+1}) - D_z\mathcal{F}(\hat{z}(s))\|_{\mathbb{V}^{-1}} + (1-\lambda(s))\|D_z\mathcal{F}(z_k) - D_z\mathcal{F}(\hat{z}(s))\|_{\mathbb{V}^{-1}}$$

$$+ \lambda(s)\|D_zf(t_{k+1}, z_{k+1}) - D_zf(\hat{t}(s), \hat{z}(s))\|_{\mathbb{V}^{-1}} + (1-\lambda(s))\|D_zf(t_k, z_k) - D_zf(\hat{t}(s), \hat{z}(s))\|_{\mathbb{V}^{-1}}.$$

While the first two addends on the right hand side are bounded by (4.19), we use (1.7) and (4.5) to estimate the third one by

$$\|D_z\mathcal{F}(z_{k+1}) - D_z\mathcal{F}(\hat{z}(s))\|_{\mathbb{V}^{-1}} = \left\|\int_0^1 D_z^2\mathcal{F}(\hat{z}(s) + r(z_{k+1} - \hat{z}(s)))[z_{k+1} - \hat{z}(s)]dr\right\|_{\mathbb{V}^{-1}}$$

$$\leq C\int_0^1 (1 + \|\hat{z}(s) + r(z_{k+1} - \hat{z}(s))\|_{\mathcal{Z}}^q)\|z_{k+1} - \hat{z}(s)\|_{\mathcal{Z}}dr \leq C.$$

Similarly, we deduce from (1.11), (4.5), and $t_{k+1} - t_k \leq \tau_k \leq C$ for all $k \in \mathbb{N}$ that

$$\|D_zf(t_{k+1}, z_{k+1}) - D_zf(\hat{t}(s), \hat{z}(s))\|_{\mathbb{V}^{-1}} \leq C(|t_{k+1} - \hat{t}(s)| + \|z_{k+1} - \hat{z}(s)\|_{\mathcal{Z}}) \leq C.$$

For the terms including t_k and z_k one can argue in the same way. This finally gives the claim. $\square$

We proceed with the complementarity relation in (2.3a). As an immediate consequence of the time update in Step 4, Remark 3.5, and the construction of the affine interpolants in Step 5 of the algorithm, we deduce the following

Lemma 4.8. *The affine interpolants satisfy*

$$(4.20) \qquad\qquad \hat{t}'(s) \geq 0, \quad \hat{t}'(s) + \|\hat{z}'(s)\|_{\mathbb{V}} = 1$$

for all $s \neq s_k$, $k \in \mathbb{N}$.

The next result essentially allows us to prove that the end time is reached and the algorithm will stop after finitely many iterations, see Proposition 5.1 below. Since the associated proof is essentially analogous to the one of [6, Proposition 2.4], we postpone it to Appendix A.

Lemma 4.9. *Define the function $r : \mathbb{R}_+ \setminus \{s_k\}_{k\in\mathbb{N}} \to \mathbb{R}$ by*

$$(4.21) \qquad\qquad r(s) := \langle D_z\mathcal{I}(\hat{t}(s), \hat{z}(s)) - D_z\mathcal{I}(\underline{t}(s), \bar{z}(s)), \hat{z}'(s)\rangle_{\mathcal{Z}^*, \mathcal{Z}}.$$

Then, for all $k \in \mathbb{N}$, there holds

$$(4.22) \qquad\qquad r(s) \leq \bar{c}\,\tau_k \quad \forall s \in (s_{k-1}, s_k)$$

and

$$(4.23) \qquad I_1^k = \int_{s_{k-1}}^{s_k} \hat{t}'(s)\,\overline{\operatorname{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(\hat{t}(s), \hat{z}(s)), \partial\mathcal{R}(0)\}ds \leq \bar{c}\,\tau_k,$$

$$(4.24) \qquad I_2^k = \int_{s_{k-1}}^{s_k} r(s) + \|\hat{z}'(s)\|_{\mathbb{V}}\Big(\overline{\operatorname{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(\hat{t}(s), \hat{z}(s)), \partial\mathcal{R}(0)\}$$
$$- \overline{\operatorname{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(\underline{t}(s), \bar{z}(s)), \partial\mathcal{R}(0)\}\Big)ds \leq \bar{c}\,\tau_k,$$

with a constant $\bar{c} > 0$, independent of τ_k and the chosen tolerance.

Lemma 4.8 already provides a bound for $\hat{z}'$, but only in $\mathcal{V}$. The next lemma addresses a bound w.r.t. the $\mathcal{Z}$-norm. Its proof combines arguments from [18, Proposition 3.7 and Lemma 3.9] and is given in Appendix B for convenience of the reader.

Lemma 4.10. *There exists a constant $C > 0$, independent of the tolerance and the choice of the step size, such that, for all $k \in \mathbb{N}$, there holds $\|\hat{z}\|_{H^1(0,s_k;\mathcal{Z})} \leq C$.*

The next auxiliary result addresses the complementarity relation and the energy balance in (2.3b) and (2.5). Its proof is only based on the stationarity conditions from Lemma 4.1 in combination with a chain rule for the energy. It is therefore completely analogous to the case with uniform step size and hence omitted.

Lemma 4.11 ([6, Proposition 2.4], [9, Lemma 3.8]). *Let $k \in \mathbb{N}$ be arbitrary. For every choice of the step sizes, the affine and constant interpolants generated by Algorithm 3.2 fulfill*

$$(4.25) \qquad \hat{t}'(s)\,\overline{\mathrm{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(\underline{t}(s), \bar{z}(s)), \partial\mathcal{R}(0)\} = 0 \quad a.e. \ in \ (0, s_k).$$

and, for all $0 \le a \le b \le s_k$, the following discrete energy identity holds true

$$(4.26) \qquad \mathcal{I}(\hat{t}(b), \hat{z}(b)) + \int_a^b \mathcal{R}(\hat{z}'(\sigma)) + \overline{\mathrm{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(\underline{t}(\sigma), \bar{z}(\sigma)), \partial\mathcal{R}(0)\}d\sigma$$

$$-\mathcal{I}(\hat{t}(a), \hat{z}(a)) - \int_a^b \partial_t\mathcal{I}(\hat{t}(\sigma), \hat{z}(\sigma))\,\hat{t}'(\sigma)d\sigma = \int_a^b r(\sigma)d\sigma,$$

where r is as defined in (4.21).

Lemma 4.12. *For all $k \in \mathbb{N}$, there holds $I_2^k \ge 0$.*

Proof. Using $\hat{t}' + \|\hat{z}'\|_{\mathbb{V}} = 1$ a.e. by Lemma 4.8 and the discrete energy identity from (4.26), we find

$$I_2^k = \int_{s_{k-1}}^{s_k} r(s) + \|\hat{z}'(s)\|_{\mathbb{V}}\Big(\overline{\mathrm{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(\hat{t}(s), \hat{z}(s)), \partial\mathcal{R}(0)\}$$
$$- \overline{\mathrm{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(\underline{t}(s), \bar{z}(s)), \partial\mathcal{R}(0)\}\Big)ds$$

$$= \mathcal{I}(\hat{t}(s_k), \hat{z}(s_k)) + \int_{s_{k-1}}^{s_k} \mathcal{R}(\hat{z}'(s)) + \|\hat{z}'(s)\|_{\mathbb{V}}\overline{\mathrm{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(\hat{t}(s), \hat{z}(s)), \partial\mathcal{R}(0)\}ds$$

$$- \mathcal{I}(\hat{t}(s_{k-1}), \hat{z}(s_{k-1})) - \int_{s_{k-1}}^{s_k} \partial_t\mathcal{I}(\hat{t}(s), \hat{z}(s))\,\hat{t}'(s)ds$$

$$\ge 0,$$

where the last inequality follows from Remark 2.2. $\qquad\square$

Lemma 4.11 demonstrates that the complementarity relation (2.3b) and the energy identity (2.5) are not fulfilled by the affine interpolants (as expected). The mismatch between (2.3b) and its discrete counterpart (4.25) arises, since the constant instead of the affine interpolant appears in the distance in (4.25). This gives rise to our first residuum in (3.3), which measures the defect, if we insert $(\hat{t}, \hat{z})$ in the complementarity condition instead of $(\underline{t}, \bar{z})$. Similarly, the "wrong" interpolant appears in the distance in (4.26), too. Note in this context that the missing $\mathbb{V}$-norm of $\hat{z}'$ in front of the distance in comparison to (2.5) can be inserted here, since, by (4.20) and (4.25),

$$\overline{\mathrm{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(\underline{t}(s), \bar{z}(s)), \partial\mathcal{R}(0)\} = \|\hat{z}'(s)\|_{\mathbb{V}}\,\overline{\mathrm{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(\underline{t}(s), \bar{z}(s)), \partial\mathcal{R}(0)\} \quad \text{a.e. in } (0, s_k).$$

However, the discrete energy identity in form of (4.26) without this norm is better suited for the convergence analysis in the upcoming sections. Moreover, an additional term involving r emerges on the right hand side of the discrete energy identity. Comparing (2.5) with (4.26) thus results in the second residuum in (3.4), cf. also (4.24).

Remark 4.13. As indicated at the beginning of this section, the aforementioned results are completely independent of the choice of the step size, since their proofs only employ the stationarity conditions from Lemma 4.1, the decay condition in (3.2), the time update formula from Step 4, and our standing assumptions on energy and dissipation. This observation will be of major importance in Section 6.

As it will turn out, the trouble maker in the convergence analysis is not the mismatch induced by r, since we already have a fairly rigorous estimate thereof in form of (4.22). The delicate issue is to pass to the limit with the distance terms. If however the involved spaces and their respective norms happen to be equivalent, the analysis simplifies significantly and one obtains much sharper estimates. Moreover, an additional assumption on the algorithm leading to a modification of the update of the step size becomes superfluous in this case (see Assumption 6.1 below). For this reason, we will treat this case separately in the next section.

5. CONVERGENCE ANALYSIS IN $\mathbb{R}^n$

As indicated above, we restrict ourselves to the case, where $\mathcal{Z} = \mathcal{V}$. We denote this space by $\mathcal{V}$ and its norm by $\|\cdot\|$ in the following. A natural example for this setting is of course the finite dimensional case, i.e., $\mathcal{V} = \mathbb{R}^n$, but infinite dimensional settings are conceivable as well. Consider for instance a measure space $(\Omega, \mathcal{A}, \mu)$ with a finite measure μ and take $\mathcal{V} = L^2(\mu)$ and $\mathcal{X} = L^1(\mu)$. Then, $\mathcal{R}$ and f can be chosen as in Example 1.4 and A is set to be a weighted $L^2(\mu)$-norm squared. If $\mathcal{F}$ is chosen as $\mathcal{F}(z) := \int_\Omega g(Tz)d\mu$, where $T \in L(L^2(\mu), L^p(\mu))$, $p > 2$, is a compact operator (e.g., a convolution with a smoothing kernel or another integral operator) and $g : L^p(\mu) \to L^2(\mu)$ is a suitable smooth Nemyzkii operator, then the assumptions in (1.6)–(1.8) can be fulfilled so that the example fits into our general setting. Nevertheless, the most relevant example is probably $\mathcal{V} = \mathbb{R}^n$, which also motivates the title of this section.

So far, we do not know whether Algorithm 3.2 is well posed in the sense that the termination criterion in Step 2 is fulfilled for k sufficiently large and the end time is reached. This will be shown next under the additional assumption that

$$(5.1) \qquad \frac{\texttt{tol}}{2\,\bar{c}} \leq \tau_1.$$

Of course, since we drive $\texttt{tol}$ to zero anyway, this assumption is not restrictive at all. Therefore, we tacitly take it for granted for the rest of this section without mentioning it every time.

Proposition 5.1. *Algorithm 3.2 terminates after a finite number of iterations, i.e., there exists $N_{\texttt{tol}} \in \mathbb{N}$, depending on the chosen tolerance, so that $t_{N_{\texttt{tol}}} \geq T$.*

Proof. Due to Steps 7 and 13 and (4.23), (4.24) and thanks to the assumption in (5.1), it holds for each $k \in \mathbb{N}$ that

$$(5.2) \qquad \tau_k \geq \frac{\texttt{tol}}{2\,\bar{c}} =: \tau_{\min},$$

where $\bar{c} > 0$ is the constant from Lemma 4.9, which is independent of the tolerance. On account of the time update in Step 4 and estimate (4.8), we thus obtain

$$(5.3) \qquad t_k = t_0 + \sum_{i=1}^k \tau_i - \sum_{i=1}^k \|z_i - z_{i-1}\|_{\mathbb{V}} \geq k\,\tau_{\min} - \sum_{i=1}^k \|z_i - z_{i-1}\|_{\mathbb{V}} \geq k\,\tau_{\min} - C \to \infty,$$

as $k \to \infty$. $\qquad\square$

The above result allows us to introduce the maximal value for the curve parameter by setting

$$(5.4) \qquad S := \min\{s \in [0, \infty) : \hat{t}(s) = T\}.$$

Remark 5.2. It is to be noted that the proof of Proposition 5.1 does not require that $\mathcal{Z}$ and $\mathcal{V}$ coincide. The assertion of Proposition 5.1 thus also holds for $\mathcal{Z} \neq \mathcal{V}$.

In the case that all spaces are equal, we can substantially sharpen the estimates from Lemma 4.9, as the next lemma shows.

Lemma 5.3. *For all $k = 1, ..., N_{\mathtt{tol}}$, there holds*

$$I_1^k \leq C \, \tau_k^2 \quad and \quad I_2^k \leq C \, \tau_k^2$$

with a constant $C > 0$ independent of τ_k and the chosen tolerance.

Proof. First note that, thanks to Remark 3.5, the construction of the affine and constant interpolants, and the time update in Step 4,

$$\|\hat{z}(s) - \bar{z}(s)\| \leq C \, \|z_k - z_{k-1}\|_{\mathbb{V}} \leq C \, \tau_k \quad and \quad |\hat{t}(s) - \underline{t}(s)| \leq |t_k - t_{k-1}| \leq \tau_k,$$

holds true for all $s \in [s_{k-1}, s_k)$, whereby it follows

$$(5.5) \quad \begin{aligned} &\|D_z\mathcal{I}(\hat{t}(s), \hat{z}(s)) - D_z\mathcal{I}(\underline{t}(s), \bar{z}(s))\|_{\mathcal{V}^*} \\ &\leq \|A(\hat{z}(s) - \bar{z}(s))\|_{\mathcal{V}^*} + \|D_z\mathcal{F}(\hat{z}(s)) - D_z\mathcal{F}(\bar{z}(s))\|_{\mathcal{V}^*} \\ &\qquad + \|D_z f(\hat{t}(s), \hat{z}(s)) - D_z f(\underline{t}(s), \bar{z}(s))\|_{\mathcal{V}^*} \leq C \, \tau_k, \end{aligned}$$

where we used (1.11) and (1.7), the latter in combination with the boundedness of the iterates by Corollary 4.3. If $\mathcal{Z} = \mathcal{V}$, then $\overline{\mathrm{dist}}_{\mathcal{V}^*}\{\eta, \partial\mathcal{R}(0)\}$ equals the usual distance of η to the bounded set $\partial\mathcal{R}(0) \subset \mathcal{V}^*$ and is always finite. We denote this distance therefore just by dist. Exploiting the Lipschitz-continuity of the distance then results in

$$(5.6) \quad \begin{aligned} &|\,\mathrm{dist}\{-D_z\mathcal{I}(\hat{t}(s), \hat{z}(s)), \partial\mathcal{R}(0)\} - \mathrm{dist}\{-D_z\mathcal{I}(\underline{t}(s), \bar{z}(s)), \partial\mathcal{R}(0)\}| \\ &\qquad\qquad \leq L \, \|D_z\mathcal{I}(\hat{t}(s), \hat{z}(s)) - D_z\mathcal{I}(\underline{t}(s), \bar{z}(s))\|_{\mathcal{V}^*} \leq C \, \tau_k. \end{aligned}$$

With this estimate and (3.3), (4.20), and (4.25) at hand, we derive

$$\begin{aligned} I_1^k &= \int_{s_{k-1}}^{s_k} \hat{t}'(s)\big(\mathrm{dist}\{-D_z\mathcal{I}(\hat{t}(s), \hat{z}(s)), \partial\mathcal{R}(0)\} - \mathrm{dist}\{-D_z\mathcal{I}(\underline{t}(s), \bar{z}(s)), \partial\mathcal{R}(0)\}\big)ds \\ &\leq \int_{s_{k-1}}^{s_k} \big|\,\mathrm{dist}\{-D_z\mathcal{I}(\hat{t}(s), \hat{z}(s)), \partial\mathcal{R}(0)\} - \mathrm{dist}\{-D_z\mathcal{I}(\underline{t}(s), \bar{z}(s)), \partial\mathcal{R}(0)\}\big|ds \leq C \, \tau_k^2. \end{aligned}$$

Concerning I_2^k, we use (4.22) in combination with (5.6) and (4.25) to obtain

$$\begin{aligned} I_2^k &= \int_{s_{k-1}}^{s_k} r(s) + \|\hat{z}'(s)\|_{\mathbb{V}}\big(\mathrm{dist}\{-D_z\mathcal{I}(\hat{t}(s), \hat{z}(s)), \partial\mathcal{R}(0)\} - \mathrm{dist}\{-D_z\mathcal{I}(\underline{t}(s), \bar{z}(s)), \partial\mathcal{R}(0)\}\big)ds \\ &\leq C \, \tau_k^2, \end{aligned}$$

which completes the proof. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

These improved a priori estimates allow us to give a sharper estimate of $N_{\mathtt{tol}}$, the maximum number of steps needed to reach the final time T. Arguing as in the proof of Proposition 5.1 with Lemma 5.3 instead of (4.23) and (4.24), we find that, for each iteration,

$$(5.7) \qquad\qquad\qquad\qquad \tau_k \geq c \, \sqrt{\mathtt{tol}} =: \tau_{min}.$$

with a constant $c > 0$ independent of the chosen tolerance. Analogously to (5.3), we then find for the maximum number of iterations

$$(5.8) \qquad\qquad\qquad\qquad N_{\mathtt{tol}} \leq \left\lceil \frac{T + C}{\tau_{min}} \right\rceil = \left\lceil \frac{T + C}{c\sqrt{\mathtt{tol}}} \right\rceil,$$

where $C > 0$ is the constant from Lemma 4.5. With this estimate at hand, we are now in the position to prove the weak* convergence of a subsequence of affine interpolants generated by Algorithm 3.2 to a $\mathbb{V}$-parametrized BV solution, provided the tolerance tends to zero. For this purpose, let $\{\mathtt{tol}_n\}_{n \in \mathbb{N}}$ be an arbitrary positive sequence converging monotonously to zero. For each $n \in \mathbb{N}$, we denote the iterates of the algorithm from Step 3 and 4 by (t_k^n, z_k^n) with associated step size τ_k^n, $k = 0, \ldots, N_{\mathtt{tol}_n}$, and the corresponding affine and constant interpolants by $\hat{z}_n$, $\hat{t}_n$,

$\bar{z}_n$, and $\underline{t}_n$. Moreover, we abbreviate $N_n := N_{\mathtt{tol}_n}$. Finally, the maximum value for the curve parameter from (5.4) associated with $\mathtt{tol}_n$ is denoted by S_n and we define

$$\tilde{S} := \sup_{n \in \mathbb{N}} S_n. \tag{5.9}$$

Due to (4.8), we obtain

$$S_n \le s_{N_n}^n = \sum_{i=1}^{N_n} \tau_i^n = \sum_{i=1}^{N_n} t_i^n - t_{i-1}^n + \|z_i^n - z_{i-1}^n\|_{\mathbb{V}} = t_{N_n}^n + \sum_{i=1}^{N_n} \|z_i^n - z_{i-1}^n\|_{\mathbb{V}} \le C, \tag{5.10}$$

which implies the boundedness of the sequence $\{S_n\}_{n \in \mathbb{N}}$ and therefore $\tilde{S}$ is finite. Finally, we constantly extend $\hat{z}_n, \hat{t}_n, \bar{z}_n$, and $\underline{t}_n$ from $[0, s_{N_n}^n]$ to the interval $[0, \tilde{S}]$ and denote these constant continuations by the same symbolds.

Theorem 5.4. *Assume that $\mathcal{Z} = \mathcal{V}$. Then every sequence $\{\mathtt{tol}_n\}_{n \in \mathbb{N}} \subset \mathbb{R}^+$ converging to zero admits a subsequence, denoted by the same symbol to ease notation, such that*

$$S_n \to S \text{ in } \mathbb{R}, \quad \hat{t}_n \rightharpoonup^* \hat{t} \text{ in } W^{1,\infty}(0, S; \mathbb{R}), \quad \hat{z}_n \rightharpoonup^* \hat{z} \text{ in } W^{1,\infty}(0, S; \mathcal{V}), \tag{5.11}$$

and every weak limit in the above sense is a $\mathbb{V}$-parametrized BV solution of the rate-independent system.

Proof. To some extend, the proof follows the lines of [6, Theorem 2.5] and [9, Theorem 3.9], but, since we have to modify the arguments at some distinct places, we present the proof in detail.

First, due to (5.10) and (4.20), which hold for every $n \in \mathbb{N}$, we observe that $\{S_n\}, \{\hat{t}_n\}$, and $\{\hat{z}_n\}$ are bounded in $\mathbb{R}, W^{1,\infty}(0, \tilde{S})$, and $W^{1,\infty}(0, \tilde{S}; \mathcal{V})$, respectively, and the existence of a weakly* converging subsequence as in (5.11) follows immediately. Note that $S \le \tilde{S}$ by definition of $\tilde{S}$.

(i) Initial and end time condition

To show that every weak* accumulation point is a $\mathbb{V}$-parametrized BV solution, assume that $\{S_n, \hat{t}_n, \hat{z}_n\}_{n \in \mathbb{N}}$ is a sequence generated by Algorithm 3.2 associated with $\mathtt{tol}_n$ such that (5.11) holds. Then, by compact resp. continuous embedding, we know that

$$\hat{t}_n \to \hat{t} \text{ in } C([0, \tilde{S}]), \quad \hat{z}_n(s) \rightharpoonup \hat{z}(s) \text{ in } \mathcal{V} \ \forall s \in [0, \tilde{S}] \tag{5.12}$$

and consequently,

$$z_0 = \hat{z}_n(0) \rightharpoonup \hat{z}(0) \text{ in } \mathcal{V}, \quad 0 = \hat{t}_n(0) \to \hat{t}(0), \quad T = \hat{t}_n(S_n) \to \hat{t}(S) \quad \text{for } n \to \infty, \tag{5.13}$$

so that the initial and end time condition in (2.2) are fulfilled.

(ii) Complementarity relation

First, thanks to Lemma 4.8, for every $n \in \mathbb{N}$ and almost every $s \in (0, \tilde{S})$, there holds $0 \le \hat{t}_n'(s)$ and $\hat{t}_n'(s) + \|\hat{z}_n'(s)\|_{\mathbb{V}} \le 1$. Note that this also holds f.a.a. $s \ge s_{N_n}^n$ because of the constant continuation we have chosen. To prove that this complementarity relation transfers to the weak limit, let us define the set

$$M := \{(\tau, \xi) \in L^2(0, S; \mathbb{R}) \times L^2(0, S; \mathcal{V}) : \tau(s) \ge 0, \ \tau(s) + \|\xi(s)\|_{\mathbb{V}} \le 1 \text{ a.e. in } (0, \tilde{S})\}. \tag{5.14}$$

Clearly M is convex and closed and therefore weakly closed such that $(\hat{t}_n', \hat{z}_n') \in M$ for all $n \in \mathbb{N}$ yields $(\hat{t}', \hat{z}') \in M$, i.e., (2.3a) for the weak limit.

To shorten the notation, we use the following abbreviations in the rest of the proof:

$$\mathrm{dist}(s) := \mathrm{dist}\{-D_z \mathcal{I}(\hat{t}(s), \hat{z}(s)), \partial \mathcal{R}(0)\}, \ \mathrm{dist}_n(s) := \mathrm{dist}\{-D_z \mathcal{I}(\hat{t}_n(s), \hat{z}_n(s)), \partial \mathcal{R}(0)\}, \ s \in [0, S].$$

To prove the second complementarity relation in (2.3b), we first note that the weak* convergence of $\hat{t}_n'$ yields

$$0 \le \int_0^S \hat{t}'(\sigma) \ \mathrm{dist}(\sigma) \, d\sigma$$

$$= \lim_{n\to\infty} \int_0^S \hat{t}'_n(\sigma) \operatorname{dist}(\sigma) \, d\sigma$$

$$\leq \limsup_{n\to\infty} \int_0^S \hat{t}'_n(\sigma) \big(\operatorname{dist}(\sigma) - \operatorname{dist}_n(\sigma) \big) d\sigma + \limsup_{n\to\infty} \int_0^S \hat{t}'_n(\sigma) \operatorname{dist}_n(\sigma) \, d\sigma.$$

Now, thanks to the pointwise (weak) convergence in (5.12), the weak continuity of $D_z\mathcal{I}$ by (1.15), the weak lower semicontinuity of the distance implies

(5.15)
$$\liminf_{n\to\infty} \operatorname{dist}_n(s) = \liminf_{n\to\infty} \operatorname{dist}\{-D_z\mathcal{I}(\hat{t}_n(s), \hat{z}_n(s)), \partial\mathcal{R}(0)\}$$
$$\geq \operatorname{dist}\{-D_z\mathcal{I}(\hat{t}(s), \hat{z}(s)), \partial\mathcal{R}(0)\} = \operatorname{dist}(s) \quad \forall\, s \in [0, S],$$

which, due to $0 \leq \hat{t}'_n(s) \leq 1$ a.e. in $(0, S)$, in turn gives

(5.16)
$$\limsup_{n\to\infty} \hat{t}'_n(s)\big(\operatorname{dist}(s) - \operatorname{dist}_n(s) \big) \leq 0$$

f.a.a. $s \in (0, S)$. Thus Fatou's lemma implies

$$\limsup_{n\to\infty} \int_0^S \hat{t}'_n(\sigma)\big(\operatorname{dist}(\sigma) - \operatorname{dist}_n(\sigma) \big) d\sigma \leq \int_0^S \limsup_{n\to\infty} \big(\hat{t}'_n(\sigma)\big(\operatorname{dist}(\sigma) - \operatorname{dist}_n(\sigma) \big) d\sigma \big) d\sigma \leq 0.$$

For the second integral we obtain by means of our refinement criterion in Step 7 that

$$\limsup_{n\to\infty} \int_0^S \hat{t}'_n(\sigma) \operatorname{dist}_n(\sigma) \, d\sigma \leq \limsup_{n\to\infty} \sum_{k=1}^{N_n} I_{1,n}^k \leq \limsup_{n\to\infty} N_n\mathtt{tol}_n \leq \lim_{n\to\infty} \left\lceil \frac{(T+C)}{c\sqrt{\mathtt{tol}_n}} \right\rceil \mathtt{tol}_n = 0,$$

where we used (5.8) for the last estimate. Note that the integrand on the left hand side vanishes a.e. in $(s_{N_n}^n, S)$ due to constant continuation, which justifies the first inequality. Hence, we have proven

$$\int_0^S \hat{t}'(\sigma) \operatorname{dist}\{-D_z\mathcal{I}(\hat{t}(\sigma), \hat{z}(\sigma)), \partial\mathcal{R}(0)\} d\sigma = 0.$$

and the non-negativity of the integrand yields that the weak limit $(\hat{t}, \hat{z})$ indeed fulfills (2.3b).

(iii) Energy identity

It remains to verify the energy identity. For this purpose, let $s \in (0, S)$ be arbitrary. First of all, the pointwise (weak) convergence in (5.12) together with the weak lower semicontinuity of $\mathcal{I}$ by (1.14) imply

(5.17)
$$\mathcal{I}(\hat{t}(s), \hat{z}(s)) \leq \liminf_{n\to\infty} \mathcal{I}(\hat{t}_n(s), \hat{z}_n(s)).$$

For the energy at initial time, we have

(5.18)
$$\mathcal{I}(\hat{t}_n(0), \hat{z}_n(0)) = \mathcal{I}(0, z_0) = \mathcal{I}(\hat{t}(0), \hat{z}(0)) \quad \forall\, n \in \mathbb{N}.$$

Furthermore, the convexity and continuity of $\mathcal{R}$ by assumption together with the weak* convergence of $\hat{z}'$ in $L^\infty(0, S; \mathcal{V})$ yields

(5.19)
$$\int_0^s \mathcal{R}(\hat{z}'(\sigma)) d\sigma \leq \liminf_{n\to\infty} \int_0^s \mathcal{R}(\hat{z}'_n(\sigma)) d\sigma.$$

Moreover, the pointwise (weak) convergence in (5.12) and the assumption in (1.13) yield

$$\partial_t f(\hat{t}_n(\sigma), \hat{z}_n(\sigma)) \to \partial_t f(\hat{t}(\sigma), \hat{z}(\sigma)) \quad \forall\, \sigma \in [0, S],$$

and, since this expression is bounded by (1.10) and (4.5), Lebesgue's dominated convergence theorem implies that it converges strongly in $L^2(0, S)$. Together with the weak convergence of $\hat{t}'_n$ in $L^2(0, T)$, it follows

(5.20)
$$\int_0^s \partial_t\mathcal{I}(\hat{t}_n(\sigma), \hat{z}_n(\sigma))\hat{t}'_n(\sigma) d\sigma \to \int_0^s \partial_t\mathcal{I}(\hat{t}(\sigma), \hat{z}(\sigma))\hat{t}'(\sigma) d\sigma.$$

In addition, as $\|\hat{z}'(\sigma)\|_{\mathbb{V}} \leq 1$ a.e. in $(0, S)$ by (2.3a), Fatou's lemma together with (5.15) gives

$$\int_0^s \|\hat{z}'(\sigma)\|_{\mathbb{V}} \, \mathrm{dist}(s) \, d\sigma \leq \int_0^s \mathrm{dist}(s) \, d\sigma \leq \liminf_{n\to\infty} \int_0^s \mathrm{dist}_n(s) \, d\sigma.$$

Since $s < S$ and $S_n \to S$, there holds $s < S_n \leq s_{N_n}^n$ for sufficiently large $n \in \mathbb{N}$ and thus, for those n, the discrete energy identity from (4.26) holds with $a = 0$ and $b = s$. Combining the above convergence results with the discrete energy identity leads to

$$\mathcal{I}(\hat{t}(s), \hat{z}(s)) - \mathcal{I}(\hat{t}(0), \hat{z}(0)) - \int_0^s \partial_t \mathcal{I}(\hat{t}(\sigma), \hat{z}(\sigma))\hat{t}'(\sigma)d\sigma + \int_0^s \mathcal{R}(\hat{z}'(\sigma)) + \|\hat{z}'(\sigma)\|_{\mathbb{V}} \, \mathrm{dist}(\sigma) \, d\sigma$$

$$\leq \liminf_{n\to\infty} \mathcal{I}(\hat{t}_n(s), \hat{z}_n(s)) - \lim_{n\to\infty} \left(\mathcal{I}(\hat{t}_n(0), \hat{z}_n(0)) + \int_0^s \partial_t \mathcal{I}(\hat{t}_n(\sigma), \hat{z}_n(\sigma))\hat{t}_n'(\sigma)d\sigma \right)$$

$$+ \liminf_{n\to\infty} \int_0^s \mathcal{R}(\hat{z}_n'(\sigma))d\sigma + \liminf_{n\to\infty} \int_0^s \mathrm{dist}_n(\sigma) \, d\sigma$$

$$\leq \limsup_{n\to\infty} \left(\mathcal{I}(\hat{t}_n(s), \hat{z}_n(s)) - \mathcal{I}(\hat{t}_n(0), \hat{z}_n(0)) + \int_0^s \partial_t \mathcal{I}(\hat{t}_n(\sigma), \hat{z}_n(\sigma))\hat{t}_n'(\sigma)d\sigma \right.$$

$$\left. + \int_0^s \mathcal{R}(\hat{z}_n'(\sigma)) + \mathrm{dist}_n(\sigma) \, d\sigma \right),$$

$$= \limsup_{n\to\infty} \int_0^s r_n(\sigma) + \mathrm{dist}\{-D_z\mathcal{I}(\hat{t}_n(\sigma), \hat{z}_n(\sigma)), \partial\mathcal{R}(0)\} - \mathrm{dist}\{-D_z\mathcal{I}(\underline{t}_n(\sigma), \bar{z}_n(\sigma)), \partial\mathcal{R}(0)\}d\sigma$$

$$= \limsup_{n\to\infty} \sum_{k=1}^{N_n} \left(I_{1,n}^k + I_{2,n}^k \right) \leq \limsup_{n\to\infty} 2 \cdot N_n \mathtt{tol}_n \leq \limsup_{n\to\infty} \left\lceil \frac{2(T+C)}{c\sqrt{\mathtt{tol}_n}} \right\rceil \mathtt{tol}_n = 0,$$

where we used the definition of the residua and their positivity by Lemma 4.12 along with (4.20) and again (5.8). Thus, the limit $(\hat{t}, \hat{z})$ satisfies the energy inequality in $(0, S)$, which, by Remark 2.2, is equivalent to the energy identity (2.5). By continuity, the energy identity holds on the whole interval $[0, S]$. Consequently, $(\hat{t}, \hat{z})$ is indeed a $\mathbb{V}$-parametrized BV solution, as claimed. $\qquad\square$

We point out that one looses the normalization condition in (4.20) by passing to the limit and it might happen that the weak limit is a degenerate $\mathbb{V}$-parametrized BV solution. Nevertheless, one may cut out sets of positive measure from $(0, S)$, where $\hat{t}'(s) + \|\hat{z}'(s)\|_{\mathbb{V}} = 0$, without changing the curve and convert the remaining curve into a normalized solution by means of a suitable re-parametrization, see [17, Lemma A.4.3] for details.

6. Convergence for Systems in Infinite Dimensional Spaces

In this section, we turn to the case $\mathcal{Z} \neq \mathcal{V}$ such that Example 1.4 is covered by the convergence analysis, too. Unfortunately, as indicated above, we need an additional assumption on the algorithm to treat this case.

Assumption 6.1. *Let a sequence of tolerances* $\{\mathtt{tol}_n\}_{n\in\mathbb{N}}$ *tending monotonically to zero be given. Then we assume that, for each* $n \in \mathbb{N}$, *the update of the step size in Algorithm 3.2 is such that the following holds true:*

(1) *For all* $n \in \mathbb{N}$, *the end time is reached, i.e., there is an index* $N_n \in \mathbb{N}$ *such that* $t_{N_n}^n \geq T$.
(2) *After each iteration, the estimates of Step 7 are fulfilled, i.e.,* $I_{1,n}^k < \mathtt{tol}$ *and* $I_{2,n}^k < \mathtt{tol}$ *for all* $n \in \mathbb{N}$ *and all* $k \leq N_n$.

Thanks to (1), *we can again construct* S_n *and* $\tilde{S}$ *as in* (5.4) *and* (5.9), *respectively. We then recursively define the family* $\{\bar{\tau}_n\}_{n\in\mathbb{N}}$ *of step size functions by*

$$\bar{\tau}_1 : [0, \tilde{S}] \to [0, \infty), \qquad \bar{\tau}_1(s) := \begin{cases} \tau_{k+1}^1, & s \in [s_k^1, s_{k+1}^1), \ k = 0, \ldots, N_1 - 1, \\ \tau_{\max}, & s \geq s_{N_1}^1, \end{cases}$$

$$\bar{\tau}_n : [0, \tilde{S}] \to [0, \infty), \ n > 1, \quad \bar{\tau}_n(s) := \begin{cases} \tau_{k+1}^n, & s \in [s_k^n, s_{n+1}^1), \ k = 0, \ldots, N_n - 1, \\ \min_{1 \leq j \leq n-1} \bar{\tau}_j(s), & s \geq s_{N_n}^n, \end{cases}$$

where $\tau_{\max}$ *is the constant from Lemma* 4.6. *Given the step size functions, we additionally require that*

(3) *The sequence of step size functions is pointwise monotonically decreasing, i.e.,* $\bar{\tau}_{n+1}(s) \leq \bar{\tau}_n(s)$ *for all* $s \in [0, \tilde{S}]$ *and all* $n \in \mathbb{N}$.

The rather intricate definition of $\bar{\tau}_n$ is due to the fact that all iterates "live" on their own interval $[0, s_{N_n}^n]$ and, potentially, we need to extend them to the fixed interval $[0, \tilde{S}]$. Note that the construction of the extension guarantees the monotony of $\{\bar{\tau}_n\}_{n\in\mathbb{N}}$ in $[s_{N_n}^n, \tilde{S}]$.

While Assumption 6.1(2) and (1) are fulfilled by the basic version of our adaptive algorithm, cf. Proposition 5.1 and Remark 5.2, the third assumption on the step size function need not be satisfied in general. For this reason, the update rule for the step size from Algorithm 3.2 has to be modified, if $\mathcal{Z} \neq \mathcal{V}$, in order to ensure Assumption 6.1(3). We again underline that the results of Section 4 remain valid for any other choice of the step size, as already noted in Remark 4.13, so we still have them at our disposal, no matter how the step size is chosen.

In the following, we present one possible modification of our basic algorithm that fulfills Assumption 6.1. The principle idea is to use nested grids, where the set of grid points $\{s_k^n\}_{0=1,\ldots,N_n}$ for the tolerance $\texttt{tol}_n$ contains all grid points s_k^j, $k = 0, \ldots, N_j$, corresponding to all previous tolerances $\texttt{tol}_j$, $j < n$. This is realized as follows:

Algorithm 6.2 (Adaptive Scheme Producing Nested Grids)**.**
1: Choose $\texttt{tol}_1 > 0$ and $\tau_1 > 0$, set $n = 1$.
2: **loop**
3:    Set $t_0^n = 0$, $s_0^n = 0$, $k = 1$, $\sigma_1^n = \tau_1$.
4:    **while** $t_{k-1}^n < T$ **do**
5:      Set $\tau_k^n = \sigma_k^n$
6:      **if** $n > 1$ **then**
7:        Define

$$(6.1) \qquad \bar{s}_k^n := \min_{1 \leq j \leq n-1} \min_{1 \leq \ell \leq N_j} \{s_\ell^j : s_\ell^j > s_{k-1}^n\}$$

        (with the convention $\min \emptyset = \infty$)
8:        **if** $s_{k-1}^n + \sigma_k^n > \bar{s}_k^n$ **then**
9:          Set $\tau_k^n = \bar{s}_k^n - s_{k-1}^n$
10:        **end if**
11:      **end if**
12:      Perform Steps 3–6 from Algorithm 3.2 to compute the constant and affine interpolants and the residua $I_{1,n}^k$ and $I_{2,n}^k$
13:      **if** $I_{1,n}^k < \texttt{tol}_n$ and $I_{2,n}^k < \texttt{tol}_n$ **then**
14:        Set $k = k + 1$.
15:        **if** $I_{1,n}^k < \frac{\texttt{tol}_n}{2}$ and $I_{2,n}^k < \frac{\texttt{tol}_n}{2}$ **then**
16:          Set $\sigma_{k+1}^n = 2\sigma_k^n$.
17:        **end if**
18:      **else**

19:        Set $\sigma_k^n = \frac{1}{2}\,\tau_k^n$ and **go to** Step 5.
20:     **end if**
21:   **end while**
22:   Choose $\texttt{tol}_{n+1} < \texttt{tol}_n$, set $n = n + 1$.
23: **end loop**

It is to be noted that, as long as $s_{N_{n-1}}^{n-1} > s_{k-1}^n$, it suffices to consider $j = n - 1$ in (6.1), because the grids are nested. Although it is rather obvious that Algorithm 6.2 indeed satisfies Assumption 6.1, it is fairly technical to prove. For this reason, we postpone the proof of the following proposition to Appendix C.

Proposition 6.3. *Algorithm 6.2 satisfies Assumption 6.1.*

Now, we are in the position to turn to our main convergence result. As in Section 5, we again consider a sequence $\{\texttt{tol}_n\}_{n\in\mathbb{N}}$ tending monotonically to zero and run an adaptive algorithm with that sequence, that fulfills Assumption 6.1 (not necessarily Algorithm 6.2). We use the same notation for the associated sequence of interpolants etc. as in Section 5, i.e., e.g., $\hat{z}_n$, $\hat{t}_n$, and so on. Moreover, we again extend the affine and constant interpolants beyond $s_{N_n}^n$ by constant continuation and denoted these extensions by the same symbols.

Theorem 6.4. *Let $\{\texttt{tol}_n\}_{n\in\mathbb{N}}$ be a sequence converging monotonically to zero and assume that a step size update is used such that Assumption 6.1 is fulfilled. Then there exists a subsequence such that*

$$(6.2) \qquad\qquad S_{n_k} \to S \quad \text{in } \mathbb{R}$$

$$(6.3) \qquad\qquad \hat{t}_{n_k} \rightharpoonup^* \hat{t} \quad \text{in } W^{1,\infty}(0,S;\mathbb{R})$$

$$(6.4) \qquad\qquad \hat{z}_{n_k} \rightharpoonup^* \hat{z} \quad \text{in } W^{1,\infty}(0,S;\mathcal{V}) \cap H^1(0,S;\mathcal{Z})$$

and the limit $(\hat{t}, \hat{z})$ is a $\mathbb{V}$-parametrized BV solution of the rate-independent system. Furthermore, every accumulation point of $\{(S_n, \hat{t}_n, \hat{z}_n)\}_{n\in\mathbb{N}}$ with respect to the above convergence is a $\mathbb{V}$-parametrized BV solution.

Proof. The proof is in principle similar to the one of Theorem 5.4, but, since we do not have the tightened bounds from Lemma 5.3 at hand if $\mathcal{Z} \neq \mathcal{V}$, we need to substantially modify the derivation of the complementarity relation and the energy identity for the weak limit. We again divide the proof in several steps.

(i) Boundedness and existence of converging subsequences

Since Lemma 4.5 does not depend on the choice of the step size, (5.10) remains true for the modified algorithm, too, such that $\tilde{S}$ is finite. Owing to the constant continuation we have chosen, the bounds from Lemmas 4.8 and 4.10 transfer to $(0, \tilde{S})$ and hence, there is a constant $C > 0$ with

$$S_n \leq \tilde{S}, \quad \|\hat{t}'\|_{L^\infty(0,\tilde{S})} \leq C, \quad \|\hat{z}'\|_{L^\infty(0,\tilde{S};\mathcal{V})} \leq C, \quad \|\hat{z}\|_{H^1(0,\tilde{S};\mathcal{Z})} \leq C.$$

Consequently there is a subsequence such that (6.2)–(6.4) holds true.

(ii) Initial and end time condition

In order to show that every accumulation point in the sense of (6.2)–(6.4) is a $\mathbb{V}$-parametrized BV solution, let us consider an arbitrary sequence $\{(S_n, \hat{t}_n, \hat{z}_n)\}_{n\in\mathbb{N}}$ such that

$$(6.5) \quad S_n \to S \text{ in } \mathbb{R}, \quad \hat{t}_n \rightharpoonup^* \hat{t} \text{ in } W^{1,\infty}(0,S;\mathbb{R}), \quad \hat{z}_n \rightharpoonup^* \hat{z} \text{ in } W^{1,\infty}(0,S;\mathcal{V}) \cap H^1(0,S;\mathcal{Z}).$$

First we note that the linearity and continuity of $C([0,S];\mathcal{Z}) \ni z \mapsto z(s) \in \mathcal{Z}$, $s \in [0,S]$, along with the continuous embedding $H^1(0,S;\mathcal{Z}) \hookrightarrow C([0,S];\mathcal{Z})$ yields the following pointwise convergence

$$(6.6) \qquad\qquad \hat{z}_n(s) \rightharpoonup \hat{z}(s) \quad \text{in } \mathcal{Z} \text{ for all } s \in [0,S].$$

In addition, weak* convergence in $W^{1,\infty}(0,S)$ implies that $\hat{t}_n$ converges uniformly in $[0,S]$. With these pointwise convergences at hand, we can argue as in part (i) of the proof of Theorem 5.4 to see that the weak limit satisfies the initial and end time condition.

(iii) Complementarity relation

Concerning the first complementarity relation in (2.3a), we again make use of the set M as defined in (5.14). Since its weak closedness is completely independent of the space $\mathcal{Z}$, the weak* convergence in (6.5) together with Lemma 4.8 yields $(\hat{t}', \hat{z}') \in M$ so that the weak limit indeed satisfies (2.3a).

For convenience, we use the following abbreviations in the rest of the proof:

$$\widehat{\mathrm{dist}}(s) := \overline{\mathrm{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(\hat{t}(s), \hat{z}(s)), \partial\mathcal{R}(0)\},$$

$$\widehat{\mathrm{dist}}_n(s) := \overline{\mathrm{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(\hat{t}_n(s), \hat{z}_n(s)), \partial\mathcal{R}(0)\},$$

$$\overline{\mathrm{dist}}_n(s) := \overline{\mathrm{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(\underline{t}_n(s), \bar{z}_n(s)), \partial\mathcal{R}(0)\},$$

where $s \in [0, S]$ and $\underline{t}_n$ and $\bar{z}_n$ again denote the constant interpolants.

To prove (2.3b), let $s \in (0, S)$ and $\delta > 0$ be arbitrary. Since $s < S$ and $S_n \to S$, there holds $s < S_n \le s_{N_n}^n$ for $n \in \mathbb{N}$ sufficiently large. Thus, for all n large enough, the results of Section 4 hold true and we do not have to care about the behavior of the discrete solution and its constant continuation, respectively, beyond $s_{N_n}^n$, which will be frequently used in the following. Since we are interested in the passage to the limit $n \to \infty$ anyway, we thus tacitly assume that n is sufficiently large such that $s < s_{N_n}^n$ for the rest of the proof. By Assumption 6.1(3), the step size function $\bar{\tau}_n$ converges pointwise to a measurable function $\bar{\tau} : [0, S] \to [0, \infty)$. For this limit function, we define the sets

(6.7)
$$\begin{aligned} G_0 &:= \{\sigma \in (0, s) : \bar{\tau}(\sigma) = 0\}, \\ G_{<\delta} &:= \{\sigma \in (0, s) : 0 < \bar{\tau}(\sigma) < \delta\}, \\ G_{\ge\delta} &:= \{\sigma \in (0, s) : \bar{\tau}(\sigma) \ge \delta\}. \end{aligned}$$

Note that $G_0 \cup G_{<\delta} \cup G_{\ge\delta} = (0, s)$. From the weak*-convergence of $\hat{t}_n'$ we deduce

(6.8)
$$\begin{aligned} 0 &\le \int_0^s \hat{t}'(\sigma)\, \widehat{\mathrm{dist}}(\sigma)\, d\sigma \\ &= \lim_{n\to\infty} \int_{G_0} \hat{t}_n'(\sigma)\, \widehat{\mathrm{dist}}(\sigma)\, d\sigma + \lim_{n\to\infty} \int_{G_{<\delta}} \hat{t}_n'(\sigma)\, \widehat{\mathrm{dist}}(\sigma)\, d\sigma + \lim_{n\to\infty} \int_{G_{\ge\delta}} \hat{t}_n'(\sigma)\, \widehat{\mathrm{dist}}(\sigma)\, d\sigma \\ &=: I_0 + I_{<\delta} + I_{\ge\delta}. \end{aligned}$$

Next we will pass to the limit with the three integrals separately and start with I_0. For this reason, let $\sigma \in G_0$ be arbitrary. For every $n \in \mathbb{N}$, there exists an interval $[s_k^n, s_{k+1}^n)$ such that $\sigma \in [s_k^n, s_{k+1}^n)$, where the respective k of course depends on n, but we suppress this dependency to ease notation. The pointwise convergence $\bar{\tau}_n(\sigma) \to 0$ by the definition of G_0 implies that the associated step size $\tau_{k+1}^n = s_{k+1}^n - s_k^n$ equals zero in the limit $n \to \infty$, too. The construction of the interpolants along with Lemma 4.8 therefore implies

(6.9) $$\|\hat{z}_n(\sigma) - \bar{z}_n(\sigma)\|_{\mathbb{V}} \le |\sigma - s_k^n|\, \|\hat{z}_n'(\sigma)\|_{\mathbb{V}} \le \tau_{k+1}^n \to 0,$$

(6.10) $$|\hat{t}_n(\sigma) - \underline{t}_n(\sigma)| \le |\sigma - s_{k+1}^n|\, \hat{t}_n'(\sigma) \le \tau_{k+1}^n \to 0$$

for $n \to \infty$. Now (6.6) together with the compactness of the embedding $\mathcal{Z} \hookrightarrow \mathcal{V}$ implies the strong convergence $\hat{z}_n(\sigma) \to \hat{z}(\sigma)$ in $\mathcal{V}$. Because of (4.5), the constant interpolant $\bar{z}_n(\sigma)$ is bounded in $\mathcal{Z}$, which yields the existence of a weak limit in $\mathcal{Z}$ for a subsequence. Again, the compactness $\mathcal{Z} \hookrightarrow \mathcal{V}$ implies strong convergence in $\mathcal{V}$ and (6.9) gives that the weak limit of $\bar{z}_n(\sigma)$ equals $\hat{z}(\sigma)$. Since $\sigma \in G_0$ was arbitrary, we have proven that the constant and affine interpolants have the

same pointwise weak limit in G_0. For the time interpolants one can argue in exactly the same way to show that $\underline{t}(\sigma) \to \hat{t}(\sigma)$ for all $\sigma \in G_0$. Therefore, we deduce from (1.15)

$$D_z\mathcal{I}(\underline{t}_n(\sigma), \bar{z}_n(\sigma)) \rightharpoonup D_z\mathcal{I}(\hat{t}(\sigma), \hat{z}(\sigma)) \quad \text{in } \mathcal{Z}^* \quad \text{for all } \sigma \in G_0,$$

which together with the weak lower semicontinuity of $\overline{\text{dist}}$ by Lemma D.1 in the appendix gives

$$
\begin{aligned}
(6.11) \quad \widehat{\text{dist}}(\sigma) &= \overline{\text{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(\hat{t}(\sigma), \hat{z}(\sigma)), \partial\mathcal{R}(0)\} \\
&\leq \liminf_{n\to\infty} \overline{\text{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(\underline{t}_n(\sigma), \bar{z}_n(s)), \partial\mathcal{R}(0)\} = \liminf_{n\to\infty} \overline{\text{dist}}_n(\sigma)
\end{aligned}
$$

for all $\sigma \in G_0$. Due to $0 \leq \hat{t}'_n(\sigma) \leq 1$ a.e. in $(0, \tilde{S})$ and (4.25), this in turn implies

$$\limsup_{n\to\infty} \hat{t}'_n(\sigma)\,\widehat{\text{dist}}(\sigma) = \limsup_{n\to\infty} \hat{t}'_n(\sigma)\big(\widehat{\text{dist}}(\sigma) - \overline{\text{dist}}_n(\sigma)\big) \leq 0 \quad \text{a.e. in } G_0.$$

Hence, Fatou's lemma implies for the first limit

$$(6.12) \qquad I_0 \leq \int_{G_0} \limsup_{n\to\infty} \hat{t}'_n(\sigma)\,\widehat{\text{dist}}(\sigma)\,d\sigma \leq 0.$$

Next, we turn to the second integral. Similarly to (6.11), the weak lower semicontinuity of $\overline{\text{dist}}_{\mathcal{V}^*}$ in combination with (6.5) and (1.15) yields

$$
\begin{aligned}
(6.13) \quad \widehat{\text{dist}}(\sigma) &= \overline{\text{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(\hat{t}(\sigma), \hat{z}(\sigma)), \partial\mathcal{R}(0)\} \\
&\leq \liminf_{n\to\infty} \overline{\text{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(\hat{t}_n(\sigma), \hat{z}_n(\sigma)), \partial\mathcal{R}(0)\} = \liminf_{n\to\infty} \widehat{\text{dist}}_n(\sigma)
\end{aligned}
$$

for all $\sigma \in G_{<\delta} \cup G_{\geq\delta}$. Therefore, thanks to (4.20) and (4.18), the integrands in $I_{<\delta}$ are bounded by a constant $C > 0$ independent of n and δ and we obtain

$$(6.14) \qquad I_{<\delta} = \lim_{n\to\infty} \int_{G_{<\delta}} \hat{t}'_n(\sigma)\,\widehat{\text{dist}}(\sigma)\,d\sigma \leq C\,|G_{<\delta}|.$$

For the last limit we obtain

$$(6.15) \qquad I_{\geq\delta} \leq \limsup_{n\to\infty} \int_{G_{\geq\delta}} \hat{t}'_n(\sigma)\big(\widehat{\text{dist}}(\sigma) - \widehat{\text{dist}}_n(\sigma)\big)d\sigma + \limsup_{n\to\infty} \int_{G_{\geq\delta}} \hat{t}'_n(\sigma)\,\widehat{\text{dist}}_n(\sigma)\,d\sigma.$$

Similarly to above, it follows from (6.13) and $\hat{t}'_n \in [0, 1]$ a.e. in $(0, \tilde{S})$ that, for all $\sigma \in G_{\geq\delta}$,

$$\limsup_{n\to\infty} \hat{t}'_n(\sigma)\big(\widehat{\text{dist}}(\sigma) - \widehat{\text{dist}}_n(\sigma)\big) \leq 0$$

so that, again by Fatou's lemma, we can estimate the first integral by

$$(6.16) \qquad \limsup_{n\to\infty} \int_{G_{\geq\delta}} \hat{t}'_n(\sigma)\big(\widehat{\text{dist}}(\sigma) - \widehat{\text{dist}}_n(\sigma)\big)d\sigma \leq 0.$$

To estimate the second integral, we introduce the set

$$(6.17) \qquad M_\delta^n := \big\{k \in \{0, ..., N_n - 1\} : \tau_{k+1}^n = s_{k+1}^n - s_k^n \geq \delta\big\}, \quad n \in \mathbb{N}.$$

It describes the parts, where the algorithm uses step sizes larger than $\delta > 0$. Since S_n is bounded by $\tilde{S}$, M_δ^n has less than $\lceil\frac{\tilde{S}}{\delta}\rceil$ elements for every $n \in \mathbb{N}$. This together with our first residuum $I_{1,n}^k$ and the condition in Step 13 allows us to estimate

$$
\begin{aligned}
(6.18) \quad \limsup_{n\to\infty} \int_{G_{\geq\delta}} \hat{t}'_n(\sigma)\,\widehat{\text{dist}}_n(\sigma)\,d\sigma &\leq \limsup_{n\to\infty} \sum_{k\in M_\delta^n} \int_{s_k^n}^{s_{k+1}^n} \hat{t}'_n(\sigma)\,\widehat{\text{dist}}_n(\sigma)\,d\sigma \\
&\leq \limsup_{n\to\infty} \left\lceil\frac{\tilde{S}}{\delta}\right\rceil \texttt{tol}_n = 0,
\end{aligned}
$$

where we made use of the non-negativity of the above integrand and the pointwise monotony of the meshsize functions $\bar{\tau}_n$, which leads to $G_{\geq\delta} \subset \bigcup_{k\in M_\delta^n} [s_k^n, s_{k+1}^n]$ for every $n \in \mathbb{N}$.

Finally combining (6.12), and (6.14)–(6.18) with (6.8) leads to

$$0 \le \int_0^s \hat{t}'(\sigma) \widehat{\mathrm{dist}}(\sigma) \, d\sigma \le C \, |G_{<\delta}|.$$

Since δ was arbitrary, this holds for every $\delta > 0$. From the definition of the set $G_{<\delta}$ in (6.7), it follows that $|G_{<\delta}| \to 0$ as $\delta \searrow 0$. Together with the non-negativity of the integrand and the arbitrariness of $s \in (0, S)$, this yields

$$\hat{t}'(\sigma)\overline{\mathrm{dist}}_{\mathcal{V}^*}\{-D_z\mathcal{I}(\hat{t}(\sigma), \hat{z}(\sigma)), \partial\mathcal{R}(0)\} = 0 \quad \text{a.e. in } (0, S),$$

which is the desired complementarity relation in (2.3b).

(iv) Energy identity

To prove the energy equality, let again $s \in (0, S)$ and $\delta > 0$ be arbitrary and define the sets G_0, $G_{<\delta}$, and $G_{\ge\delta}$ as above. We again assume w.l.o.g. that $n \in \mathbb{N}$ is so large that $s < s_{N_n}^n$. We consider each part in the energy equality individually and, except for the integral involving the distance, all parts can be discussed in the same way as in the case $\mathcal{V} = \mathcal{Z}$. To be more precise, in view of the uniform convergence of $\hat{t}_n$, the pointwise weak convergence in (6.6) (here in $\mathcal{Z}$ and not only in $\mathcal{V}$), the weak convergence of $\hat{t}'_n$ and $\hat{z}'_n$, and our standing assumptions on energy and dissipation give that (5.17)–(5.20) from the proof of Theorem 5.4 also hold in case $\mathcal{V} \ne \mathcal{Z}$.

Concerning the distance term, we exploit (6.11), (6.13), $\|\hat{z}'(\sigma)\|_{\mathbb{V}} \le 1$ by (2.3a), and Fatou's lemma in order to obtain

$$\int_0^s \|\hat{z}'(\sigma)\|_{\mathbb{V}} \widehat{\mathrm{dist}}(\sigma) \, d\sigma \le \int_{G_0} \widehat{\mathrm{dist}}(\sigma) \, d\sigma + \int_{G_{<\delta}\cup G_{\ge\delta}} \widehat{\mathrm{dist}}(\sigma) \, d\sigma$$

$$\le \int_{G_0} \liminf_{n\to\infty} \overline{\mathrm{dist}}_n(\sigma) \, d\sigma + \int_{G_{<\delta}\cup G_{\ge\delta}} \liminf_{n\to\infty} \widehat{\mathrm{dist}}_n(\sigma) \, d\sigma$$

$$\le \liminf_{n\to\infty} \Big(\int_{G_0} \overline{\mathrm{dist}}_n(\sigma) \, d\sigma + \int_{G_{<\delta}\cup G_{\ge\delta}} \widehat{\mathrm{dist}}_n(\sigma) \, d\sigma \Big).$$

Combining this with (5.17)–(5.20) leads to

$$\mathcal{I}(\hat{t}(s), \hat{z}(s)) - \mathcal{I}(\hat{t}(0), \hat{z}(0)) - \int_0^s \partial_t\mathcal{I}(\hat{t}(\sigma), \hat{z}(\sigma))\hat{t}'(\sigma)d\sigma + \int_0^s \mathcal{R}(\hat{z}'(\sigma)) + \|\hat{z}'(\sigma)\|_{\mathbb{V}} \widehat{\mathrm{dist}}(\sigma) \, d\sigma$$

$$\le \liminf_{n\to\infty} \Big(\mathcal{I}(\hat{t}_n(s), \hat{z}_n(s)) - \mathcal{I}(\hat{t}_n(0), \hat{z}_n(0)) - \int_0^s \partial_t\mathcal{I}(\hat{t}_n(\sigma), \hat{z}_n(\sigma))\hat{t}'_n(\sigma)d\sigma + \int_0^s \mathcal{R}(\hat{z}'_n(\sigma))d\sigma \Big)$$

$$\quad + \liminf_{n\to\infty} \Big(\int_{G_0} \overline{\mathrm{dist}}_n(\sigma) \, d\sigma + \int_{G_{<\delta}\cup G_{\ge\delta}} \widehat{\mathrm{dist}}_n(\sigma) \, d\sigma \Big)$$

$$\le \limsup_{n\to\infty} \Big(\mathcal{I}(\hat{t}_n(s), \hat{z}_n(s)) - \mathcal{I}(\hat{t}_n(0), \hat{z}_n(0)) + \int_0^s \big(\mathcal{R}(\hat{z}'_n(\sigma)) + \overline{\mathrm{dist}}_n(\sigma)\big)d\sigma$$

$$\quad - \int_0^s \partial_t\mathcal{I}(\hat{t}_n(\sigma), \hat{z}_n(\sigma))\hat{t}'_n(\sigma)d\sigma + \int_{G_{<\delta}\cup G_{\ge\delta}} \big(\widehat{\mathrm{dist}}_n(\sigma) - \overline{\mathrm{dist}}_n(\sigma)\big)d\sigma \Big)$$

$$= \limsup_{n\to\infty} \Big(\int_0^s r_n(\sigma)d\sigma + \int_{G_{<\delta}\cup G_{\ge\delta}} \big(\widehat{\mathrm{dist}}_n(\sigma) - \overline{\mathrm{dist}}_n(\sigma)\big)d\sigma \Big)$$

$$\le \limsup_{n\to\infty} \int_{G_0} r_n(\sigma)d\sigma + \limsup_{n\to\infty} \int_{G_{<\delta}} \big(r_n(\sigma) + \widehat{\mathrm{dist}}_n(\sigma) - \overline{\mathrm{dist}}_n(\sigma)\big)d\sigma$$

$$\quad + \limsup_{n\to\infty} \int_{G_{\ge\delta}} \big(r_n(\sigma) + \widehat{\mathrm{dist}}_n(\sigma) - \overline{\mathrm{dist}}_n(\sigma)\big)d\sigma$$

$$=: J_0 + J_{<\delta} + J_{\ge\delta},$$

where we made use of the discrete energy equality (4.26). For the first limit, we obtain due to the definition of G_0, (4.22), and Fatou's lemma

$$(6.19) \qquad J_0 \leq \int_{G_0} \limsup_{n \to \infty} r_n(\sigma) d\sigma \leq 0.$$

Owing to (4.22), (4.17), (4.18), and the constant continuation, we see that the integrand in $J_{<\delta}$ is pointwise bounded from above by a constant $C > 0$ independent of n. Thus, we obtain

$$(6.20) \qquad J_{<\delta} \leq C \, |G_{<\delta}|.$$

For the last limit, we use (4.20), (4.25), and the definition of M_δ^n in (6.17) in order to estimate

$$
\begin{aligned}
J_{\geq\delta} &= \limsup_{n\to\infty} \int_{G_{\geq\delta}} \Big[r_n(\sigma) + \|\hat{z}_n'(\sigma)\|_{\mathbb{V}} \big(\widehat{\mathrm{dist}}_n(\sigma) - \overline{\mathrm{dist}}_n(\sigma) \big) + \hat{t}_n'(\sigma) \, \overline{\mathrm{dist}}_n(\sigma) \Big] d\sigma \\
&\leq \limsup_{n\to\infty} \sum_{k\in M_\delta^n} \int_{s_k^n}^{s_{k+1}^n} \Big[r_n(\sigma) + \|\hat{z}_n'(\sigma)\|_{\mathbb{V}} \big(\widehat{\mathrm{dist}}_n(\sigma) - \overline{\mathrm{dist}}_n(\sigma) \big) + \hat{t}_n'(\sigma) \, \overline{\mathrm{dist}}_n(\sigma) \Big] d\sigma \\
&= \limsup_{n\to\infty} \sum_{k\in M_\delta^n} I_{1,n}^k + I_{2,n}^k \leq \limsup_{n\to\infty} 2 \left\lceil \frac{\tilde{S}}{\delta} \right\rceil \mathtt{tol}_n = 0,
\end{aligned}
$$

where the first inequality follows from the non-negativity of the integrand by Lemma 4.12. The last inequality follows analogously to (6.18) from the conditions on the residua in Step 13 of the algorithm.

All in all, we have proven that

$$
\begin{aligned}
&\mathcal{I}(\hat{t}(s), \hat{z}(s)) - \mathcal{I}(\hat{t}(0), \hat{z}(0)) - \int_0^s \partial_t \mathcal{I}(\hat{t}(\sigma), \hat{z}(\sigma)) \hat{t}'(\sigma) d\sigma \\
&+ \int_0^s \mathcal{R}(\hat{z}'(\sigma)) + \|\hat{z}'(\sigma)\|_{\mathbb{V}} \, \overline{\mathrm{dist}}_{\mathcal{V}^*}\{-D_z \mathcal{I}(\hat{t}(\sigma), \hat{z}(\sigma)), \partial\mathcal{R}(0)\} d\sigma \leq C \, |G_{<\delta}| \to 0 \quad \text{as } \delta \searrow 0
\end{aligned}
$$

such that $(\hat{t}, \hat{z})$ fulfills an energy inequality in $(0, S)$. By continuity, this inequality carries over to the whole interval $[0, S]$. From Remark 2.2, we know that this inequality is equivalent to the energy identity and, since $s \in [0, S]$ was arbitrary, we see that $(\hat{t}, \hat{z})$ indeed satisfies (2.5), which finally gives that the weak limit is a $\mathbb{V}$-parametrized BV solution. $\qquad\square$

Remark 6.5. Let us shortly explain why the additional Assumption 6.1(3) is needed. Unfortunately, we were not able to prove a result analogous to Lemma 5.3 in case of $\mathcal{V} \neq \mathcal{Z}$. In essence, the reason is that an estimate of the form (5.5) does not hold, if the $\mathcal{Z}$-norm and the $\mathcal{V}$-norm are not equivalent. Even an estimate of the form $\|z_k - z_{k-1}\|_{\mathcal{Z}} \leq c \, \tau_k$ would not suffice to prove (5.5), since A maps $\mathcal{Z}$ to $\mathcal{Z}^*$ and not to $\mathcal{V}^*$, which is needed to estimate the distance term in (5.6). At this point, the distance-terms in I_k^1 and I_k^2 behave very differently compared to the r-term from (4.21). In case of r, the critical term involving A has the "right" sign, see (A.2) in Appendix A, so that the lack of regularity does not play a role here. Note that, for this reason, we can only derive an estimate for r and not for $|r|$, which luckily suffices for the convergence analysis. To circumvent these regularity issues, we need the additional assumption of monotonically decreasing step sizes. It essentially allows us to go without the sharpened estimate for the number of steps in (5.8) that is an immediate consequence of Lemma 5.3.

7. Numerical Experiments

In the following we test our adaptive algorithm by means of two examples. The stationarity conditions in (3.1) are solved by means of a damped version of the semi-smooth Newton method from [9, Section 4]. For the evaluation of the residua in (3.3) and (3.4), we use a composite Simpson's rule.

7.1. **A One Dimensonal Example.** We start our investigations with an example, where $\mathcal{V} = \mathcal{Z} = \mathbb{R}$. For the energy in (1.5), we choose

$$\langle Az, z \rangle := z^2, \quad \mathcal{F} \equiv 0, \quad f(t, z) := -\ell(t) z + \begin{cases} 4z + 8, & z \leq -2, \\ 4 - z^2, & |z| < 2, \\ -4z + 8, & z \geq 2 \end{cases}$$

where ℓ is the external driving force, which is set to $\ell(t) := t + 1$. The end time is set to $T = 5$. It is easily seen that the standing assumptions in (1.6)–(1.13) are fulfilled with $c = 6$, $\mu = 1$, and $\nu = 2$. By direct calculations one moreover verifies that

$$(7.1) \qquad \hat{t}(s) := \begin{cases} s, & s \in [0, 2], \\ 2, & s \in (2, 10], \\ (s - 6)/2, & s \in (10, 16] \end{cases} \quad \text{and} \quad \hat{z}(s) := \begin{cases} -2, & s \in [0, 2], \\ s - 4, & s \in (2, 10], \\ (s + 2)/2, & s \in (10, 16] \end{cases}$$

is a $\mathbb{V}$-parametrized BV solution in the sense of Definition 2.1. The solution is shown in Figure 1. We observe that the state $\hat{z}$ is constant until $s = 2$. This *sticking* behavior is typical for rate-
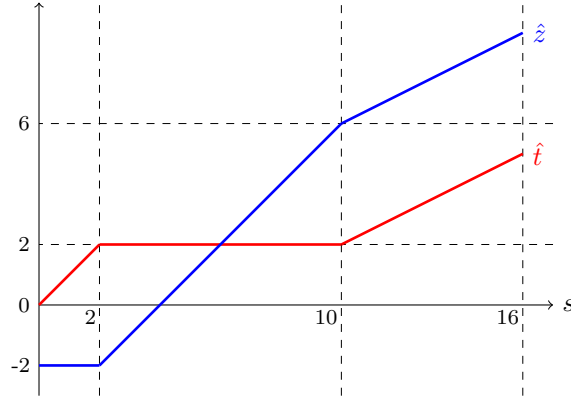
FIGURE 1. $\mathbb{V}$-parametrized BV solution from (7.1)

independent processes and occurs, if the external loading is too small to change the system state such that the dissipation forces the system state to remain constant. At $s = 2$, the system changes to a different regime. Now the physical time remains constant, while the system state starts to change and thus, we observe a discontinuous behavior of the system. In the majority of cases, the state is no more locally stable then, i.e., $-D_z \mathcal{I}(\hat{t}(s), \hat{z}(s)) \notin \partial \mathcal{R}(0)$ and thus λ from (2.6) will be different from zero so that an additional viscous term in (2.7) appears. This gives rise to the notion of a *viscous jump*. There may be discontinuities where the system state is still locally stable, but these situations are rather pathological as [17, Example 2.3.5] demonstrates. The viscous jump lasts until $s = 10$ and afterwards, both $\hat{t}$ and $\hat{z}$ are both changing, a regime which is known as *rate-independent slip*.

Figure 2 shows the result of the adaptive algorithm for two different values of the tolerance `tol`. We observe significantly larger step sizes during sticking and in a viscous jump accompanied with a refined solution of the break points at the transition between different regimes. This is explained as follows:

(i) Sticking:
Assume that Step 3 delivers $z_k = z_{k-1}$, i.e., the driving force in t_{k-1} is so small that it is compensated by the dissipation and the system state does not change. Then, by Step 4, the

(A) $\hat{t}$ and $\hat{z}$ for $\mathtt{tol} = 10^{-2}$

(B) τ_k for $\mathtt{tol} = 10^{-2}$

(C) $\hat{t}$ and $\hat{z}$ for $\mathtt{tol} = 10^{-3}$
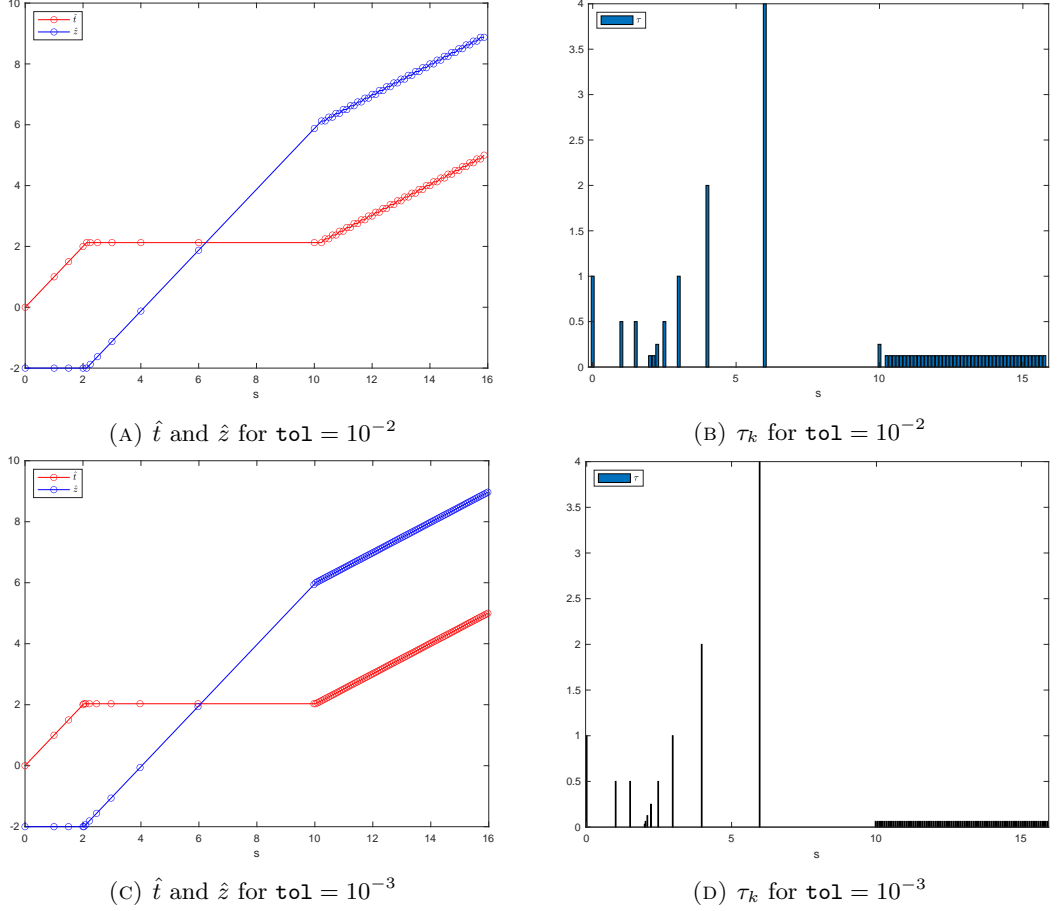
(D) τ_k for $\mathtt{tol} = 10^{-3}$

FIGURE 2. Affine interpolants and step size for different tolerances

physical time is set to $t_k = t_{k-1} + \tau_k$ such that $\hat{t}'(s) = 1$ and $\hat{z}'(s) = 0$ for all $s \in [s_{k-1}, s_k)$. It thus holds

$$I_2^k = 0, \quad I_1^k = \int_{s_{k-1}}^{s_k} \mathrm{dist}\{-D_z \mathcal{I}(\hat{t}(s), z_k), \partial \mathcal{R}(0)\} ds.$$

Therefore, if the variations of the driving force $(s_{k-1}, s_k) \ni s \mapsto \ell(\hat{t}(s))$ are not too large, then the state z_k remains locally stable on the whole interval, i.e., $-D_z \mathcal{I}(\hat{t}(s), z_k) \in \partial \mathcal{R}(0)$ f.a.a. $s \in (s_{k-1}, s_k)$, and I_1^k equals 0, too. The iteration is then accepted and the step size is doubled.

(ii) Viscous jump:
Suppose that, in iteration $k \in \mathbb{N}$, the inequality constraint is active, i.e., $\|z_{k-1} - z_k\|_{\mathbb{V}} = \tau_k$. Then, by Step 4, $\hat{t}'(s) = 0$ and $\|\hat{z}'(s)\|_{\mathbb{V}} = 1$ for all $s \in [s_{k-1}, s_k)$ and hence

$$I_1^k = 0, \quad I_2^k = \int_{s_{k-1}}^{s_k} \big[\langle D_z \mathcal{I}(\hat{t}(s), \hat{z}(s), \hat{z}'(s) \rangle + \mathrm{dist}\{-D_z \mathcal{I}(\hat{t}(s), \hat{z}(s)), \partial \mathcal{R}(0)\} + \mathcal{R}(\hat{z}'(s)) \big] ds,$$

where we used (4.3) along with the positive homogeneity of $\mathcal{R}$. Now, the inf-convolution-formula gives for the distance

$$\mathrm{dist}\{\eta, \partial \mathcal{R}(0)\} = \frac{1}{\tau_k} \left(\mathcal{R} + I_{\tau_k} \right)^*(\eta) = \sup_{\|v\|_{\mathbb{V}} \leq 1} \left(\langle \eta, v \rangle - \mathcal{R}(v) \right),$$

cf., e.g., [9, Lemma C.2]. In one dimension and $\mathbb{V} = 1$, it thus follows

$$\text{dist}\{\eta, \partial\mathcal{R}(0)\} = \max\{\eta - \mathcal{R}(1), -\eta - \mathcal{R}(-1), 0\}.$$

Therefore, $\hat{z}'(s) = \pm 1$ must only have the right sign and I_2^k will vanish, too, such that the iteration is successful and the step size is doubled. In multiple dimensions, this easy argument does of course not work, but, nonetheless, we frequently observe an increase of the step size in a viscous jump there, too, as the next example in Section 7.2 shows.

The aforementioned advantage of the adaptive scheme, i.e., the larger step size during *sticking* and *viscous jumps*, results in less iterations in order to approximate a solution with the same accuracy in comparison to an algorithm with fixed step size. This is illustrated in Figure 3, where the difference between the numerical result (computed with fixed and adaptive step size) and the exact solution from (7.1) measured in the L^2- and the L^∞-norm is shown. We observe that the qualitative behavior of the error is the same, but approximately only a quarter of iterations is needed in the adaptive case.
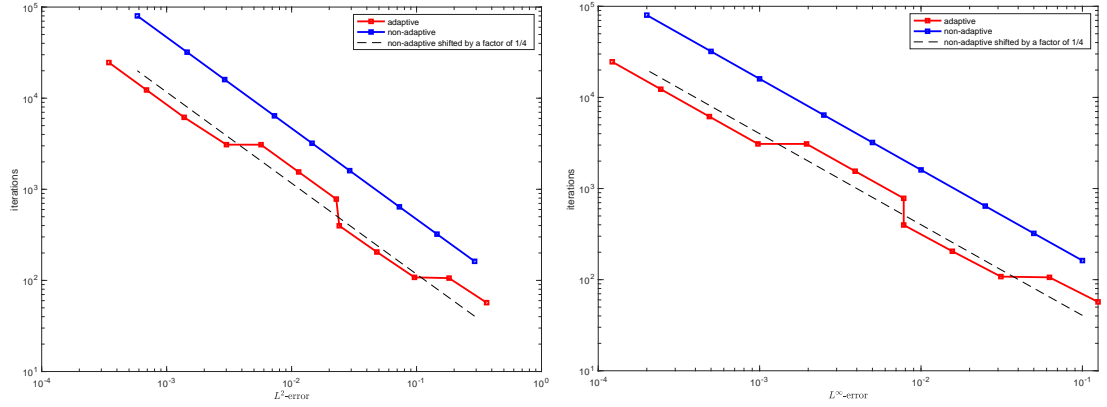

FIGURE 3. Comparison of the number of iterations between the adaptive and the non-adaptive scheme in the ODE example

However, it is to be noted that the error between "the" solution and its approximation is in general no meaningful quantity, since solutions are in general not unique and we can only guarantee that subsequences converge to $\mathbb{V}$-parametrized BV solutions as seen in Theorem 6.4. It may therefore well happen that, for different tolerances, different solutions are approximated as the next example shows.

7.2. **An Example in Sobolev Space.** This example addresses a problem in function space, where $\mathcal{V} \neq \mathcal{Z}$. To be more precise, we consider the setting from Example 1.4 with

$$T = 1, \quad \Omega := (0,1)^2, \quad \mathbb{V} = \text{id} : L^2(\Omega) \to L^2(\Omega).$$

Moreover, we set

(7.2) $$g : \mathbb{R} \to \mathbb{R}, \quad g(z) := 48(1 - z^2)^2$$

(where the nonlinear function and the Nemyzki operator are denoted by the same symbol with a slight abuse of notation). The associated Nemyzki operator is considered as a mapping from $H_0^1(\Omega) \hookrightarrow L^8(\Omega)$ to $L^2(\Omega)$ and it is easily verified that (1.20) is satisfied with $p = 8$ and $q = 2$. Furthermore, the external load from (1.19) is set to

$$\langle \ell(t), z \rangle_{\mathcal{V}^*, \mathcal{V}} = -200 \int_\Omega t^3\, z(x)\, dx, \quad (t, z) \in [0, T] \times H_0^1(\Omega).$$

For the spatial discretization, we use standard linear finite elements with 1681 nodes along with the mass lumping approach for the L^1-norm as described in [9, Section 4].

The result of Algorithm 3.2 for the tolerances $\mathtt{tol} = 10^{-4}$, $\mathtt{tol} = 10^{-5}$ and $\mathtt{tol} = 10^{-6}$ is shown in Figure (4). We first observe that, for $\mathtt{tol} = 10^{-5}$, seemingly another solution with an
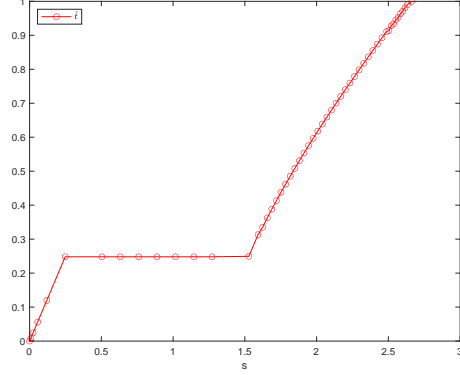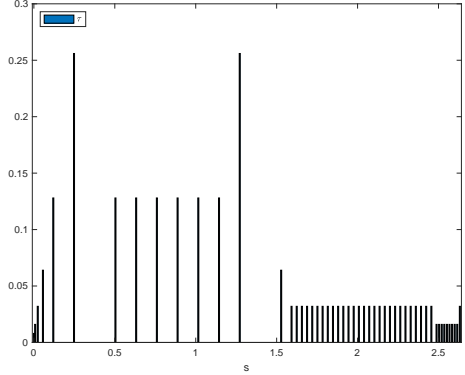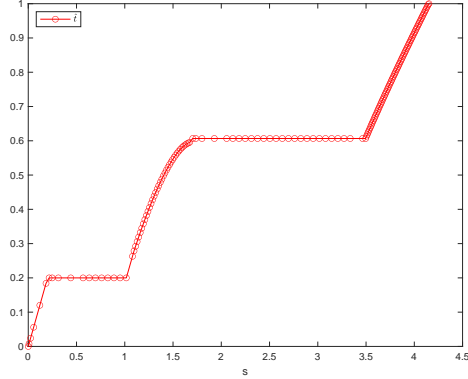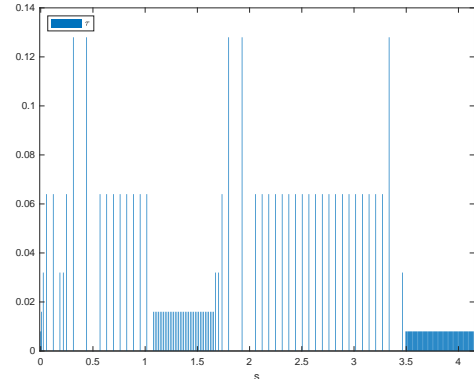

(A) $\hat{t}$ for $\mathtt{tol} = 10^{-4}$

(B) τ_k for $\mathtt{tol} = 10^{-4}$

(C) $\hat{t}$ for $\mathtt{tol} = 10^{-5}$

(D) τ_k for $\mathtt{tol} = 10^{-5}$
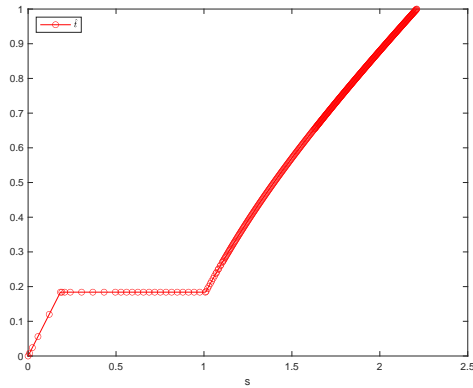
(E) $\hat{t}$ for $\mathtt{tol} = 10^{-6}$
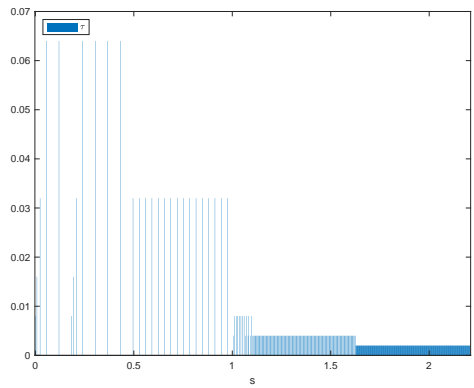
(F) τ_k for $\mathtt{tol} = 10^{-6}$

FIGURE 4. Physical time and step size for $\mathtt{tol} = 10^{-4}$, $\mathtt{tol} = 10^{-5}$ and $\mathtt{tol} = 10^{-6}$

additional viscous jump is approximated than for the other tolerances. Note that, due to the

nonconvexity of the double-well-type potential in (7.2), there may well be multiple stationary points fulfilling (3.1) and (3.2) and it is just a matter of chance, which one is computed by the globalized semi-smooth Newton method. Nevertheless, we again observe an increased step size during sticking and in viscous jumps.

Due to the ambiguity of numerical solutions and since an exact solution is unknown in this case, there is no error to evaluate and we investigate the residua in from (3.3) and (3.4) instead in order to compare the adaptive with the uniform scheme. Figure 5 shows the local and the global residuum defined by

$$\max_{k=1,\dots,N_{\text{tol}}} I_1^k + I_k^2 \quad \text{and} \quad \sum_{k=1}^{N_{\text{tol}}} I_1^k + I_k^2$$

for the adaptive and the uniform scheme. Note that both, I_1^k and I_k^2, are non-negative by Lemma 4.12. Again, we observe the same qualitative behavior, but the adaptive scheme only needs approximately half of the steps to reach a residuum of the same size. The reduction of iterations is smaller compared to the previous one-dimensional example, because the portion of the rate-independent slip regime is substantially larger compared to the example in Section 7.1. The kinks in the curves in Figure 5 are due to the aforementioned ambiguity of numerical solutions with an additional viscous jump, which leads to an enlarged artificial time horizon and in this way increases the number of iterations without reducing the residuum.
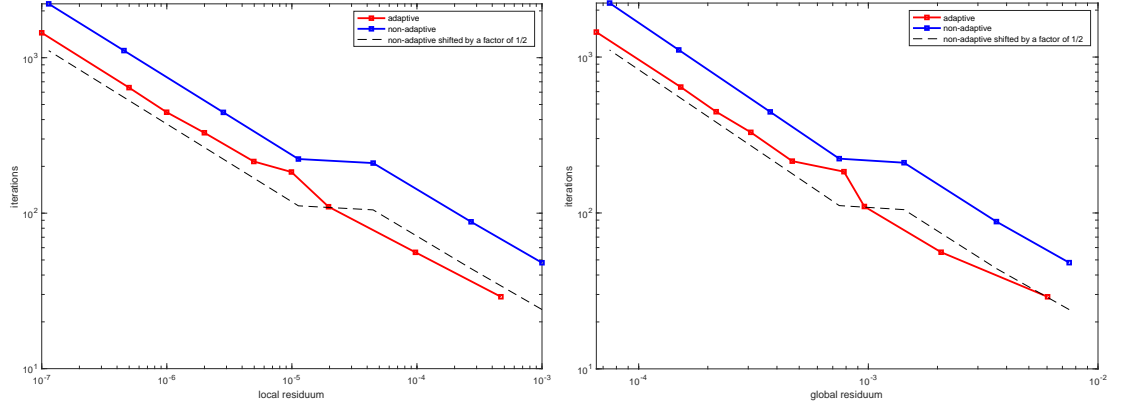

FIGURE 5. Comparison of number of iterations between adaptive and non-adaptive scheme in the PDE example

APPENDIX A. PROOF OF LEMMA 4.9

We start by proving (4.22). For this purpose, let $k \in \mathbb{N}$ and $s \in (s_{k-1}, s_k)$ be arbitrary. We first observe that the definition of the affine and constant interpolants imply for all $s \in (s_{k-1}, s_k)$ that

$$(A.1) \qquad \hat{z}'(s) = \frac{\hat{z}(s) - \bar{z}(s)}{s - s_k} \quad \text{and} \quad \hat{t}'(s) = \frac{\hat{t}(s) - \underline{t}(s)}{s - s_{k-1}}.$$

With this at hand, we obtain

$$r(s) = \langle D_z \mathcal{I}(\hat{t}(s), \hat{z}(s) - D_z \mathcal{I}(\underline{t}(s), \bar{z}(s), \hat{z}'(s)) \rangle_{\mathcal{Z}^*, \mathcal{Z}}$$

$$= -|s - s_k| \langle A\hat{z}'(s), \hat{z}'(s) \rangle_{\mathcal{Z}^*, \mathcal{Z}} + \left\langle D_z \mathcal{F}(\hat{z}(s)) - D_z \mathcal{F}(\bar{z}(s)), \frac{\hat{z}(s) - \bar{z}(s)}{s - s_k} \right\rangle_{\mathcal{V}^*, \mathcal{V}}$$

(A.2)
$$- \langle D_z f(\hat{t}(s), \hat{z}(s)) - D_z f(\underline{t}(s), \bar{z}(s)), \hat{z}'(s) \rangle_{\mathcal{V}^*, \mathcal{V}}$$

$$\leq -\alpha |s - s_k| \|\hat{z}'(s)\|_{\mathcal{Z}}^2 + \frac{1}{|s - s_k|} |\langle D_z \mathcal{F}(\bar{z}(s)) - D_z \mathcal{F}(\hat{z}(s)), \bar{z}(s) - \hat{z}(s) \rangle_{\mathcal{V}^*, \mathcal{V}}|$$

$$+ |\langle D_z f(\underline{t}(s), \bar{z}(s)) - D_z f(\hat{t}(s), \hat{z}(s)), \hat{z}'(s) \rangle_{\mathcal{V}^*, \mathcal{V}}|.$$

Thanks to (4.6) with $\epsilon = \frac{\alpha}{4}$, which is applicable here, since the interpolants are pointwise bounded on account of (4.5), we conclude

(A.3)
$$\frac{1}{|s - s_k|} |\langle D_z \mathcal{F}(\bar{z}(s)) - D_z \mathcal{F}(\hat{z}(s)), \bar{z}(s) - \hat{z}(s) \rangle_{\mathcal{V}^*, \mathcal{V}}|$$

$$\leq \frac{\alpha}{4} |s - s_k| \|\hat{z}'(s)\|_{\mathcal{Z}}^2 + C_\alpha |s - s_k| \mathcal{R}(\hat{z}'(s)) \|\hat{z}'(s)\|_{\mathcal{V}}.$$

Moreover, (4.7) with $\epsilon = \frac{\alpha}{4}$ and (A.1) yield

(A.4)
$$|\langle D_z f(\underline{t}(s), \bar{z}(s)) - D_z f(\hat{t}(s), \hat{z}(s)), \hat{z}'(s) \rangle_{\mathcal{V}^*, \mathcal{V}}|$$

$$\leq \nu |s - s_{k-1}| \hat{t}'(s) \|\hat{z}'(s)\|_{\mathcal{V}} + c_\alpha |s - s_k| \mathcal{R}(\hat{z}'(s)) \|\hat{z}'(s)\|_{\mathcal{V}} + \frac{\alpha}{4} |s - s_k| \|\hat{z}'(s)\|_{\mathcal{Z}}^2.$$

Inserting (A.3) and (A.4) in (A.2) and using the upper bound for the dissipation potential by assumption (1.4) then gives

$$r(s) \leq C(|s - s_k| \mathcal{R}(\hat{z}'(s)) \|\hat{z}'(s)\|_{\mathbb{V}} + |s - s_{k-1}| \hat{t}'(s) \|\hat{z}'(s)\|_{\mathbb{V}}) \leq C \tau_k \|\hat{z}'(s)\|_{\mathbb{V}} (\|\hat{z}'(s)\|_{\mathbb{V}} + \hat{t}'(s)),$$

which, in view of (4.20), implies (4.22).

Finally, (4.22) together with the boundedness of the distance terms by (4.17) and (4.18) and the boundedness of $\hat{t}'(s)$ and $\|\hat{z}'(s)\|_{\mathbb{V}}$ by (4.20) gives (4.23) and (4.24). $\qquad \square$

Appendix B. Proof of Lemma 4.10

According to (4.5), we already know that $\|\hat{z}\|_{L^\infty(0, s_k; \mathcal{Z})} \leq C$. Hence, all we have to show is the boundedness of $\hat{z}'_n$ in $L^2(0, s_k; \mathcal{Z})$. For this reason, we return to (4.12) to conclude that, for every $i \in \mathbb{N}$,

$$\lambda_{i+1} \|z_{i+1} - z_i\|_{\mathbb{V}}^2 - \lambda_i \|z_i - z_{i-1}\|_{\mathbb{V}} \|z_{i+1} - z_i\|_{\mathbb{V}} + \frac{\alpha}{2} \|z_{i+1} - z_i\|_{\mathcal{Z}}^2$$

$$\leq (C_{\alpha/4} + c_{\alpha/4}) \mathcal{R}(z_{i+1} - z_i) \|z_{i+1} - z_i\|_{\mathbb{V}} + \nu (t_i - t_{i-1}) \|z_{i+1} - z_i\|_{\mathbb{V}}$$

$$\leq C \tau_{i+1} (\mathcal{R}(z_{i+1} - z_i) + t_i - t_{i-1}),$$

where we exploited $\|z_{i+1} - z_i\|_{\mathbb{V}} \leq \tau_{i+1}$, cf. Remark 3.5. Now rearranging the terms and using (4.1) yields

$$\frac{\alpha}{2} \|z_{i+1} - z_i\|_{\mathcal{Z}}^2$$

$$\leq C \tau_{i+1} (\mathcal{R}(z_{i+1} - z_i) + t_i - t_{i-1}) + \lambda_i \tau_{i+1} \|z_i - z_{i-1}\|_{\mathbb{V}} - \lambda_{i+1} \tau_{i+1} \|z_{i+1} - z_i\|_{\mathbb{V}}.$$

From (4.14) it follows by the same arguments

$$\frac{\alpha}{2} \|z_1 - z_0\|_{\mathcal{Z}}^2 \leq C \tau_1 (\mathcal{R}(z_1 - z_0) + \|D_z \mathcal{I}(0, z_0)\|_{\mathcal{V}^*}) - \lambda_1 \tau_1 \|z_1 - z_0\|_{\mathbb{V}}$$

Together with the estimate above, this implies

$$\|\hat{z}'\|_{L^2(0, s_k; \mathcal{Z})}^2 \leq \sum_{i=0}^{k-1} \int_{s_i}^{s_{i+1}} \frac{\|z_{i+1} - z_i\|_{\mathcal{Z}}^2}{\tau_{i+1}^2} ds$$

$$= \sum_{i=0}^{k-1} \frac{\|z_{i+1} - z_i\|_{\mathcal{Z}}^2}{\tau_{i+1}}$$

$$\leq \frac{2}{\alpha} \Big(C\big(\mathcal{R}(z_1 - z_0) + \|D_z \mathcal{I}(0, z_0)\|_{\mathbb{V}^*}\big) - \lambda_1 \|z_1 - z_0\|_{\mathbb{V}}$$

$$+ \sum_{i=1}^{k-1} C\big(\mathcal{R}(z_{i+1} - z_i) + t_i - t_{i-1}\big) + \lambda_i \|z_i - z_{i-1}\|_{\mathbb{V}} - \lambda_{i+1} \|z_{i+1} - z_i\|_{\mathbb{V}} \Big)$$

$$= \frac{2}{\alpha} \Big(C\, t_{k-1} - \lambda_k \|z_k - z_{k-1}\|_{\mathbb{V}} + C\, \|D_z \mathcal{I}(0, z_0)\|_{\mathbb{V}^*} + \sum_{i=0}^{k-1} C\, \mathcal{R}(z_{i+1} - z_i) \Big)$$

$$\leq \frac{2}{\alpha} C \Big(T + \|D_z \mathcal{I}(0, z_0)\|_{\mathbb{V}^*} + \sum_{i=0}^{k-1} \mathcal{R}(z_{i+1} - z_i) \Big) \leq C,$$

where we used $\lambda_k \geq 0$ by Lemma 4.1 and the boundedness of the dissipation by Lemma 4.2. $\square$

APPENDIX C. WELL POSEDNESS OF ALGORITHM 6.2

To prove Proposition 6.3, we have to verify Assumption 6.1 for Algorithm 6.2. First, we immediately observe that Assumption 6.1(2) and (3) are fulfilled by construction of the algorithm.

It thus remains to show that each inner iteration of Algorithm 6.2 (i.e., the iteration w.r.t. the index k) terminates after $N_n \in \mathbb{N}$ steps. Too see this, we argue by induction. For $n = 1$, the inner iteration coincides with the one of Algorithm 3.2 because of Steps 5 and 6. Then, analogously to the proof of Proposition 5.1, there holds

$$(C.1) \qquad \tau_k^1 \geq \min \Big\{ \frac{\mathtt{tol}_1}{2\bar{c}}, \frac{\tau_1}{2} \Big\} := \tau_{\min}^1 \quad \forall\, k \in \mathbb{N}$$

and thus, an estimate as in (5.3) yields the existence of $N_1 \in \mathbb{N}$ such that $t_{N_1}^1 > T$. Now assume that, for some $n \geq 1$, there exists $N_n \in \mathbb{N}$ with $t_{N_n}^n > T$. If we consider the iteration $n + 1$, then an estimate of the form (C.1) (with $\mathtt{tol}_{n+1}$ instead of $\mathtt{tol}_1$) need not necessarily hold for all $k \in \mathbb{N}$ because of the modification of the step size in Step 9. However, if, for some $k \in \mathbb{N}$, the step size τ_k^{n+1} arises from Step 9 and is less then

$$\tau_{\min}^{n+1} := \min \Big\{ \frac{\mathtt{tol}_{n+1}}{2\bar{c}}, \frac{\tau_1}{2} \Big\},$$

then, due to Lemma 4.9, the conditions in Step 13 will be met such that the iteration is accepted and we pass on to iteration $k + 1$. The step size for this iteration is then σ_{k+1}^{n+1} (and *not* τ_k^{n+1} as in the basic version of Algorithm 3.2). By construction, σ_j^{n+1}, $j \in \mathbb{N}$, is however only reduced, if $I_{1,n}^j$ and/or $I_{2,n}^j$ were above $\mathtt{tol}_{n+1}$, see Step 19, and hence, $\sigma_j^{n+1} \geq \tau_{\min}^{n+1}$ for all $j \in \mathbb{N}$, again by Lemma 4.9. These considerations show that a step size less than $\tau_{\min}^{n+1}$ can only appear at most N_n times in the iteration $n + 1$, which allows us to argue similarly to (5.3) as follows

$$(C.2) \qquad \begin{aligned} t_k^{n+1} &= t_0^{n+1} + \sum_{i=1}^{k} \tau_i^{n+1} - \sum_{i=1}^{k} \|z_i^{n+1} - z_{i-1}^{n+1}\|_{\mathbb{V}} \\ &\geq (k - N_n)\, \tau_{\min}^{n+1} - \sum_{i=1}^{k} \|z_i^{n+1} - z_{i-1}^{n+1}\|_{\mathbb{V}} \geq (k - N_n)\, \tau_{\min}^{n+1} - C \;\to\; \infty, \quad \text{as } k \to \infty, \end{aligned}$$

where we again used Lemma 4.5. Note once more that the constant from Lemma 4.5 does neither depend on the tolerance nor on the step sizes. This gives the existence of a constant N_{n+1} such that $t_{N_{n+1}}^{n+1} > T$ as claimed. $\square$

Appendix D. A Lower Semicontinuity Result for the (Generalized) Distance

Lemma D.1. *Let $\{\eta_n\}_{n\in\mathbb{N}} \subset \mathcal{Z}^*$ be a sequence such that $\eta_n \rightharpoonup^* \eta$ in $\mathcal{Z}^*$ as $n \to \infty$ and $\overline{\mathrm{dist}}_{\mathcal{V}^*}\{-\eta_n, \partial\mathcal{R}(0)\} \leq C$ for all $n \in \mathbb{N}$ with a constant $C > 0$. Then the distance is weakly lower semicontinuous, i.e.,*

$$\overline{\mathrm{dist}}_{\mathcal{V}^*}\{-\eta, \partial\mathcal{R}(0)\} \leq \liminf_{n\to\infty} \overline{\mathrm{dist}}_{\mathcal{V}^*}\{-\eta_n, \partial\mathcal{R}(0)\}.$$

Proof. The proof is similar to [9, Theorem 3.9] or [17, Lemma 3.2.18]. First, we show by the direct method that the infimum in the definition of the distance is attained for every $n \in \mathbb{N}$. To this end, let $n \in \mathbb{N}$ be arbitrary and $\{\xi_m^n\}_{m\in\mathbb{N}} \subset \partial\mathcal{R}(0)$ be a minimizing sequence, i.e., $\|\xi_m^n + \eta_n\|_{\mathbb{V}^{-1}} \to \overline{\mathrm{dist}}_{\mathcal{V}^*}\{-\eta_n, \partial\mathcal{R}(0)\}$ as $m \to \infty$. Then, due to the boundedness of the distance by the assumption, there is a subsequence, denoted by the same symbol, such that $\xi_m^n + \eta_n \rightharpoonup^* w$ in $\mathcal{V}^*$ as $m \to \infty$ and therefore $\xi_m^n \rightharpoonup^* w - \eta_n =: \xi_n$ in $\mathcal{Z}^*$. As a subset of $\mathcal{Z}^*$ the convex subdifferential reads

$$\partial\mathcal{R}(0) = \{z^* \in \mathcal{Z}^* : \langle z^*, z \rangle_{\mathcal{Z}^*,\mathcal{Z}} \leq \mathcal{R}(z) \,\forall\, z \in \mathcal{Z}\}$$

(where we used the positive homogeneity of $\mathcal{R}$) such that $\partial\mathcal{R}(0)$ is closed w.r.t. weak* convergence in $\mathcal{Z}^*$, which in turn implies $\xi_n \in \partial\mathcal{R}(0)$. Moreover, the weak lower semicontinuity of $\|\cdot\|_{\mathbb{V}^{-1}}$ gives $\|\xi_n + \eta_n\|_{\mathbb{V}^{-1}} \leq \overline{\mathrm{dist}}_{\mathcal{V}^*}\{-\eta_n, \partial\mathcal{R}(0)\}$, which shows the optimality of ξ_n.

To show the result, we argue similarly. Assume the assertion is wrong such that there exist a subsequence $\{\eta_{n_\ell}\}_{\ell\in\mathbb{N}}$ and $\varepsilon > 0$ such that

$$(\mathrm{D.1}) \qquad \overline{\mathrm{dist}}_{\mathcal{V}^*}\{-\eta_{n_\ell}, \partial\mathcal{R}(0)\} \leq \overline{\mathrm{dist}}_{\mathcal{V}^*}\{-\eta, \partial\mathcal{R}(0)\} - \varepsilon.$$

If we define $w_{n_\ell} := \xi_{n_\ell} + \eta_{n_\ell}$, then, by assumption, w_{n_ℓ} is bounded in $\mathcal{V}^*$ and there exists a subsequence, denoted by the same symbol for simplicity, such that $w_{n_\ell} \rightharpoonup^* w$ in $\mathcal{V}^*$ and therefore

$$\xi_{n_\ell} = w_{n_\ell} - \eta_{n_\ell} \rightharpoonup^* w - \eta =: \xi \quad \text{in } \mathcal{Z}^*.$$

Again the weak closedness of $\partial\mathcal{R}(0)$ yields $\xi \in \partial\mathcal{R}(0)$, which, together with the weak lower semicontinuity of $\|\cdot\|_{\mathbb{V}^{-1}}$, implies

$$\overline{\mathrm{dist}}_{\mathcal{V}^*}\{-\eta, \partial\mathcal{R}(0)\} \leq \|\xi + \eta\|_{\mathbb{V}^{-1}} \leq \liminf_{\ell\to\infty} \|\xi_{n_\ell} + \eta_{n_\ell}\|_{\mathbb{V}^{-1}} = \liminf_{\ell\to\infty} \overline{\mathrm{dist}}_{\mathcal{V}^*}\{-\eta_{n_\ell}, \partial\mathcal{R}(0)\},$$

which contradicts (D.1). $\qquad\square$

References

[1] Bartels, S. Quasi-optimal error estimates for implicit discretizations of rate-independent evolutions. *SIAM Journal on Numerical Analysis 52*, 2 (2014), 708–716.

[2] Bartels, S., Mielke, A., and Roubíček, T. Quasi-static small-strain plasticity in the limit of vanishing hardening and its numerical approximation. *SIAM J. Numer. Anal. 50*, 2 (2012), 951–976.

[3] Dal Maso, G., DeSimone, A., and Mora, M. G. Quasistatic evolution problems for linearly elastic–perfectly plastic materials. *Archive for rational mechanics and analysis 180*, 2 (2006), 237–291.

[4] Efendiev, M. A., and Mielke, A. On the rate-independent limit of systems with dry friction and small viscosity. *Journal of Convex Analysis 13*, 1 (2006), 151–167.

[5] Grisvard, P. *Elliptic Problems in Nonsmooth Domains*. Society for Industrial and Applied Mathematics, Jan. 2011.

[6] Knees, D. Convergence analysis of time-discretisation schemes for rate-independent systems. *ESAIM, Control Optim. Calc. Var. 25* (2019).

[7] Knees, D., Rossi, R., and Zanini, C. A vanishing viscosity approach to a rate-independent damage model. *Mathematical Models and Methods in Applied Sciences 23*, 4 (2013), 565–616.

[8] Knees, D., and Shcherbakov, V. A penalized version of the local minimization scheme for rate-independent systems. *Appl. Math. Lett. 115* (2021), Paper No. 106954, 8.

[9] Meyer, C., and Sievers, M. Finite element discretization of local minimization schemes for rate-independent evolutions. *Calcolo 56*, 1 (2019).

[10] Meyer, C., and Sievers, M. A priori error analysis of local incremental minimization schemes for rate-independent evolutions. *SIAM J. Numer. Anal. 58*, 4 (2020), 2376–2403.

[11] MIELKE, A., ROSSI, R., AND SAVARÉ, G. Bv solutions and viscosity approximations of rate-independent systems. *ESAIM: Control, Optimisation and Calculus of Variations 18*, 1 (2012), 36–80.

[12] MIELKE, A., AND ROUBÍČEK, T. *Rate-Independent Systems: Theorie and Application*. Springer, New York, 2015.

[13] MIELKE, A., AND THEIL, F. A mathematical model for rate-independent phase transformations with hysteresis. In *Proceedings of the workshop "Models of Continuum Mechanics in Analysis and Engineering", TU Darmstadt* (1999), Shaker, pp. 117–129.

[14] MIELKE, A., AND THEIL, F. On rate-independent hysteresis models. *NoDEA: Nonlinear Differential Equations and Applications 11*, 2 (2004), 151–189.

[15] MIELKE, A., THEIL, F., AND LEVITAS, V. I. A variational formulation of rate-independent phase transformations using an extremum principle. *Arch. Rational Mech. Anal. 162*, 2 (2002), 137–177.

[16] NEGRI, M. Quasi-static rate-independent evolutions: characterization, existence, approximation and application to fracture mechanics. *ESAIM Control Optim. Calc. Var. 20* (2014), 983–1008.

[17] SIEVERS, M. *A Numerical Scheme for Rate-Independent Systems*. PhD thesis, TU Dortmund, 2020.

[18] SIEVERS, M. Convergence analysis of a local stationarity scheme for rate-independent systems and application to damage. https://arxiv.org/abs/2104.14997, 2021.

[19] THOMAS, S. *Optimal control of a rate-independent system constrained to parameterized balanced viscosity solutions*. PhD thesis, Universität Kassel, to appear.

TECHNISCHE UNIVERSITÄT DORTMUND, FAKULTÄT FÜR MATHEMATIK, LEHRSTUHL LSX, VOGELPOTHSWEG 87, 44227 DORTMUND, GERMANY

Email address: merlin.andreia@tu-dortmund.de

TECHNISCHE UNIVERSITÄT DORTMUND, FAKULTÄT FÜR MATHEMATIK, LEHRSTUHL LSX, VOGELPOTHSWEG 87, 44227 DORTMUND, GERMANY

Email address: christian2.meyer@tu-dortmund.de