Weierstraß-Institut

für Angewandte Analysis und Stochastik

Leibniz-Institut im Forschungsverbund Berlin e. V.

Preprint ISSN 2198-5855

Fully discrete approximation of rate-independent damage models with gradient regularization

Sören Bartels[1], Marijo Milicevic[1], Marita Thomas[2], Nico Weber[1]

submitted: March 30, 2020

[1] Department of Applied Mathematics
Mathematical Institute
University of Freiburg
Hermann-Herder-Str. 9
79104 Freiburg i. Br.
Germany
E-Mail: bartels@mathematik.uni-freiburg.de
marijo.milicevic@mathematik.uni-freiburg.de
nico.weber@mathematik.uni-freiburg.de

[2] Weierstrass Institute
Mohrenstr. 39
10117 Berlin
Germany
E-Mail: marita.thomas@wias-berlin.de

No. 2707
Berlin 2020

2010 *Mathematics Subject Classification.* 35K86, 74R05, 49J45, 49S05, 65M60, 65M12.

Key words and phrases. Partial damage, damage evolution with gradient regularization, semistable energetic solutions, numerical approximation, iterative solution.

This work was carried out within the project *Finite element approximation of functions of bounded variation and application to models of damage, fracture, and plasticity* within the DFG Priority Programme SPP 1748 "Reliable Simulation Techniques in Solid Mechanics. Development of Non-standard Discretisation Methods, Mechanical and Mathematical Analysis.".

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

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

Fully discrete approximation of rate-independent damage models with gradient regularization

Sören Bartels, Marijo Milicevic, Marita Thomas, Nico Weber

Abstract

This work provides a convergence analysis of a time-discrete scheme coupled with a finite-element approximation in space for a model for partial, rate-independent damage featuring a gradient regularization as well as a non-smooth constraint to account for the unidirectionality of the damage evolution. The numerical algorithm to solve the coupled problem of quasistatic small strain linear elasticity with rate-independent gradient damage is based on a Variable ADMM-method to approximate the nonsmooth contribution. Space-discretization is based on P1 finite elements and the algorithm directly couples the time-step size with the spatial grid size h. For a wide class of gradient regularizations, which allows both for Sobolev functions of integrability exponent $r \in (1,\infty)$ and for BV-functions, it is shown that solutions obtained with the algorithm approximate as $h \to 0$ a semistable energetic solution of the original problem. The latter is characterized by a minimality property for the displacements, a semistability inequality for the damage variable and an energy dissipation estimate. Numerical benchmark experiments confirm the stability of the method.

1 Introduction

This paper investigates the convergence of a numerical algorithm for the computation of solutions to a class of models for quasistatic, rate-independent, partial damage processes in elastically deformable solids at small strains. Along a fixed time-interval $[0,\mathsf{T}]$ and with a deformable and damageable body occupying the bounded domain $\Omega \subset \mathbb{R}^d$, the process is modeled as a rate-independent system $(\mathbf{U}\times\mathbf{Z},\mathcal{E},\mathcal{R})$ governed by an energy functional $\mathcal{E} : [0,\mathsf{T}]\times\mathbf{U}\times\mathbf{Z}\to\mathbb{R}\cup\{\infty\}$, a positively 1-homogeneous dissipation potential $\mathcal{R}$ with a suitable state-space $\mathbf{U}$ for the displacement field $u : [0,T]\times\Omega\to\mathbb{R}^2$ and the damage variable $z : [0,T]\times\Omega\to\mathbb{R}$. More precisely, we will consider $\mathcal{R} : \mathbf{Z}\to\mathbb{R}\cup\{\infty\}$,

$$\mathcal{R}(v) := \int_\Omega \mathrm{R}(v)\,\mathrm{d}x, \quad \text{with } \mathrm{R}(v) := \begin{cases} \varrho|v|\,\mathrm{d}x & \text{if } v\in(-\infty,0],\\ +\infty & \text{if } v>0 \end{cases} \tag{1.1a}$$

$$\text{with } \mathbf{Z} := L^1(\Omega)\,, \tag{1.1b}$$

and with a constant dissipation rate $\varrho>0$. Having in mind the convention of the damage variable to describe the volume fraction of undamaged material with $z=1$ for the unbroken and $z=0$ for the broken state of the material, the dissipation potential enforces z to decrease in time and thus ensures the unidirectionality of the process by preventing healing of the material.

Moreover, we shall consider the function spaces

$$\mathbf{U} := \{v\in H^1(\Omega,\mathbb{R}^d),\ v=0 \text{ on } \Gamma_D \text{ in trace sense}\}\,, \tag{1.2a}$$

$$\mathbf{X} := W^{1,r}(\Omega) \text{ for } r\in(1,\infty), \text{ resp. } \mathbf{X} := \mathrm{BV}(\Omega) \text{ for } r=1\,, \tag{1.2b}$$

and the energy functional $\mathcal{E} : [0,\mathsf{T}]\times\mathbf{U}\times\mathbf{Z}\to\mathbb{R}\cup\{\infty\}$

$$\mathcal{E}(t,u,z) := \begin{cases} \widehat{\mathcal{E}}(t,u,z) & \text{if } (u,z)\in\mathbf{U}\times\mathbf{X},\\ \infty & \text{otherwise,} \end{cases} \tag{1.3a}$$

DOI 10.20347/WIAS.PREPRINT.2707 Berlin 2020

with $\widehat{\mathcal{E}} : [0,\mathsf{T}] \times \mathbf{U} \times \mathbf{X} \to \mathbb{R} \cup \{\infty\}$ of the form

$$\begin{aligned}
\widehat{\mathcal{E}}(t,u,z) :=& \mathcal{W}(u,z) + \mathcal{G}(z) + \mathcal{L}_{\mathrm{Neu}}(t,u)\,,\\
\text{with } \mathcal{W}(u,z) :=& \frac{1}{2}\int_\Omega f(z)\big(\lambda\big|\operatorname{tr} e(u+u_{\mathrm{D}}(t))\big|^2 + 2\mu|e(u+u_{\mathrm{D}}(t))|^2\big)\,\mathrm{d}x\,,\\
\mathcal{L}_{\mathrm{Neu}}(t,u) :=& -\int_{\Gamma_{\mathrm{Neu}}} l_{\mathrm{Neu}}(t)\cdot(u+u_{\mathrm{D}}(t))\,\mathrm{d}s\,,\\
\text{and with } \mathcal{G}(z) :=& \int_\Omega \big(\tfrac{\kappa_1}{r}|\nabla z|^r + \tfrac{\kappa_2}{r}|z|^r\big)
\end{aligned} \tag{1.3b}$$

a regularization of the damage variable with $r \in [1,\infty)$. Let us mention that we also allow for $r = 1$, which sets solutions z of the model to be elements of $\mathrm{BV}(\Omega)$ and the L^1-gradient contributing to $\mathcal{G}$ is then to be replaced by the total variation $|\mathrm{D}z|(\Omega)$ of z in Ω.

Additionally, $\mathcal{W}$ in (1.3b) features the Lamé constants $\lambda, \mu > 0$, $e(u) := \frac{1}{2}(\nabla u + \nabla u^\top)$ the linearized strain tensor, $u_{\mathrm{D}} : [0,\mathsf{T}] \times \Omega \to \mathbb{R}^d$ a suitable extension of a given Dirichlet datum into the domain Ω and $\mathcal{L}_{\mathrm{Neu}}$ with $l_{\mathrm{Neu}} : [0,\mathsf{T}] \times \Gamma_{\mathrm{Neu}} \to \mathbb{R}^d$ represents a given surface loading acting along the Neumann-boundary Γ_{Neu}. Due to the mapping properties of the convex, monotonously increasing function $f : \mathbb{R} \to [a,\infty)$ featured in $\mathcal{W}$ with a constant $a > 0$ the model will capture partial damage only: It is $f(0) \geq a$ and $f(z) = \mathrm{const} \geq a$ for $z \leq 0$. Hence, even in the state of maximal damage the solid has the ability to counteract external loadings with suitable stresses and displacements. Note that the energy functional $\mathcal{E}$ does not impose any constraint enforcing the damage variables to take values in $[0,1]$ only. Instead, we will show that there is a solution of the evolution process with this property, provided the initial datum z_0 takes its values in $[0,1]$. This is based on the unidirectionality constraint featured in $\mathcal{R}$ and on growth properties imposed on the degradation function f.

Remark 1.1 (Convexity properties of the energy functional)**.** Let us point out here as a characteristic feature in the modeling of damage processes that the energy functional $\mathcal{E}(t,\cdot,\cdot)$ is separately convex in the variables u and z but not (jointly) convex in the pair (u,z). A counterexample is, e.g., provided in case of $u_{\mathrm{D}}(t) = 0$ and $z_0 \equiv 1$ by the functions $(u_1,z_1) := (\mathrm{id},0)$ and $(u_2,z_2) := (0,1)$. With this choice one finds $\mathcal{W}(\alpha(u_1,z_1)+(1-\alpha)(u_2,z_2)) > \alpha\mathcal{W}(u_1,z_1) = \alpha\mathcal{W}(u_1,z_1) + (1-\alpha)\mathcal{W}(\alpha(u_2,z_2)$ for any $\alpha \in (0,1)$ in contradiction to convexity.

It is the aim of this paper to study the existence of solutions for the rate-independent system $(\mathbf{U}\times\mathbf{Z},\mathcal{E},\mathcal{R})$ given by (1.1), (1.2), (1.3) by proving the convergence of a numerical method. For this, we will impose a partition $\Pi_N := \{t_N^k,\, k \in \{0,1,\dots,N\}, 0 = t_N^0 < \ldots < t_N^N = \mathsf{T}\}$ of the time-interval $[0,\mathsf{T}]$ and a space discretization in terms of $P1$ finite elements, yielding finite-element spaces $\mathbf{U}_h, \mathbf{X}_h$. At each time-step $t_N^k \in \Pi_N$, we will determine approximate solutions in $\mathbf{U}_h \times \mathbf{X}_h$ via an alternating minimization scheme, i.e., starting from an approximation $(u_{0h}, z_{0h}) \in \mathbf{U}_h \times \mathbf{X}_h$ of the initial datum (u_0,z_0) at t_N^0, we alternatingly compute for given $(u_{Nh}^0, z_{Nh}^0) = (u_{0h}, z_{0h})$

$$u_{Nh}^k = \mathrm{argmin}_{u\in\mathbf{U}_h}\, \mathcal{E}(t_N^k, u, z_{Nh}^{k-1}), \tag{1.4a}$$

$$z_{Nh}^k \in \mathrm{argmin}_{z\in\mathbf{X}_h}\, \mathcal{E}(t_N^k, u_{Nh}^k, z) + \mathcal{R}(z - z_{Nh}^{k-1})\,. \tag{1.4b}$$

While the computation of u_{Nh}^k reduces to the solution of a linear system of equations, the computation of z_{Nh}^k requires the solution of a constrained nonsmooth minimization problem. This problem is qualitatively of the form of the Rudin-Osher-Fatemi (ROF) problem [ROF92] for which various numerical schemes have been proposed for its iterative solution, cf., e.g., [Bar12, BT09, Cha04, CP11, GOSB14, GO09, HRH14, LM79, Roc76, WT10]. We approximate a minimizer z_{Nh}^k by converting the minimization problem into a saddle-point problem and use a variant of the alternate direction method of multipliers (ADMM) [FG83, GM76, Glo84, GLT89, GM75] recently introduced in [BM17] as *Variable-ADMM* for the approximate solution of the saddle-point problem.

We show stability of the alternate minimization scheme and prove that suitable interpolants constructed from (1.4) satisfy a discrete version of the notion of solution under consideration, which will provide uniform a priori estimates for the interpolants independent of N and h. As $h = h(N) \to 0$ for $N \to \infty$ we prove that (a subsequence of the) discrete solutions approximates a limit pair $(u,z) : [0,\mathsf{T}] \to \mathbf{U}\times\mathbf{X}$ that provides a solution to semistable energetic formulation of the system $(\mathbf{U}\times\mathbf{Z},\mathcal{E},\mathcal{R})$:

Definition 1.2 (semistable energetic solution)**.** *A function $q = (u,z) : [0,\mathsf{T}] \to \mathbf{U}\times\mathbf{Z}$ is called semistable energetic solution for the system $(\mathbf{U}\times\mathbf{Z},\mathcal{E},\mathcal{R})$, if $t \to \partial_t\mathcal{E}(t,q) \in L^1((0,\mathsf{T}))$ and if for all $s,t \in [0,\mathsf{T}]$ we have $\mathcal{E}(t,q(t)) < \infty$,*

if for a.a. $t \in (0,\mathsf{T})$ *minimality condition* (1.5a) *is satisfied and if for all* $t \in [0,\mathsf{T}]$ *semistability* (1.5b) *as well as the upper energy-dissipation estimate* (1.5c) *hold true, i.e.:*

$$\text{for all } \tilde{u} \in \mathbf{U}: \quad \mathcal{E}(t,u(t),z(t)) \le \mathcal{E}(t,\tilde{u},z(t))\,, \tag{1.5a}$$

$$\text{for all } \tilde{z} \in \mathbf{X}: \quad \mathcal{E}(t,u(t),z(t)) \le \mathcal{E}(t,u(t),\tilde{z}) + \mathcal{R}(\tilde{z}-z(t)), \tag{1.5b}$$

$$\mathcal{E}(t,q(t)) + \mathcal{R}(z(t)-z(0)) \le \mathcal{E}(0,q(0)) + \int_0^t \partial_\xi \mathcal{E}(\xi,q(\xi))\,\mathrm{d}\xi\,, \tag{1.5c}$$

where the dissipated energy up to time t *is given by the total variation induced by the dissipation potential* $\mathcal{R}$ *with unidirectionality constraint and, by the induced monotonicity of* $z:[0,\mathsf{T}]\to\mathbf{Z}$*, takes the form* $\mathcal{R}(z(t)-z(0))$*.*

In fact, we observe in our convergence result of Theorem 4.2 that the approximation of the semistability inequality depends on the properties of the meshes used. More precisely, we find that the limit pair (u,z) satisfies semistability only up to an error term due to the mesh quality:

$$\text{for all } \tilde{z} \in \mathbf{X}: \quad \mathcal{E}(t,u(t),z(t)) \le \mathcal{E}(t,u(t),\tilde{z}) + \mathcal{R}(\tilde{z}-z(t)) + \mathrm{ERR} \tag{1.6}$$

and ERR depends on the smallest and largest angles appearing in the approximating triangulations $\mathcal{T}_{h(N)}$. In the case that $\mathcal{T}_{h(N)}$ tends to right-angled triangulation as $N\to\infty$ our result proves $\mathrm{ERR}=0$. The error term ERR originates in the analysis from the interplay of the gradient regularization $\mathcal{G}$ in (1.3b) with general growth $r\in[1,\infty)$ and the projection into the FE-spaces of the non-smooth unidirectionality constraint featured in $\mathcal{R}$, which is also enforced by the numerical algorithm in the discrete FE-setting. In the limit passage from the discrete to the continuous formulation the presence of the non-smooth constraint results in the weak convergence of the approximating semistable damage variables $(\overline{z}_N)_N$ only and thus leads to the emergence of ERR.

Remark 1.3 (Comparison with other results in literature)**.** With the devised algorithm one determines approximate solutions of the first-order optimality conditions corresponding to (1.4), i.e.,

$$\mathrm{D}_u\mathcal{E}(t_N^k,u_{Nh}^k,z_{Nh}^{k-1}) \;=\; 0\,, \tag{1.7a}$$

$$\partial_z\mathcal{R}(z_{Nh}^k - z_{Nh}^{k-1}) + \mathrm{D}_z\mathcal{E}(t_N^k,u_{Nh}^k,z_{Nh}^k) \;\ni\; 0\,. \tag{1.7b}$$

Exploiting the convexity of $\mathcal{R}$ and the separate convexity of $\mathcal{E}(t_N^k,\cdot,\cdot)$ results in discrete versions of (1.5a) and (1.5b).

In case of $\mathcal{E}(t,\cdot,\cdot)$ being convex in the pair (u,z) the first-order optimality conditions lead to the stability of the pair (u,z):

$$\text{for all } (\tilde{u},\tilde{z}) \in \mathbf{U}\times\mathbf{X}: \quad \mathcal{E}(t,u(t),z(t)) \le \mathcal{E}(t,\tilde{u},\tilde{z}) + \mathcal{R}(\tilde{z}-z(t)), \tag{1.8a}$$

The stability inequality (1.8a) allows it to deduce an energy-dissipation estimate opposite to (1.5c) and thus leads to the energy-dissipation balance

$$\mathcal{E}(t,q(t)) + \mathcal{R}(z(t)-z(0)) = \mathcal{E}(0,q(0)) + \int_0^t \partial_\xi \mathcal{E}(\xi,q(\xi))\,\mathrm{d}\xi\,. \tag{1.8b}$$

A pair $q=(u,z):[0,\mathsf{T}]\to\mathbf{U}\times\mathbf{Z}$ that satisfies (1.8) for all times $t\in[0,T]$ is called an *energetic solution* of the rate-independent system $(\mathbf{U}\times\mathbf{Z},\mathcal{E},\mathcal{R})$; see, e.g., the monograph [MR15] for more details on the theory of energetic solutions. In general, in the setting of infinite-dimensional Banach spaces $\mathcal{U}\times\mathcal{Z}$, energetic solutions to rate-independent systems can be obtained by the approximation with solutions of a time-discrete scheme that *simultaneously* minimizes with respect to the pair (u,z), i.e., for all $k\in\{1,\dots,N\}$ it is

$$(u_N^k,z_N^k) \in \operatorname{argmin}_{(\tilde{u},\tilde{z})\in\mathbf{U}\times\mathbf{Z}}\big(\mathcal{E}(t_N^k,\tilde{u},\tilde{z}) + \mathcal{R}(\tilde{z}-z_N^{k-1})\big)\,. \tag{1.9}$$

On the level of minimality this approximation procedure has been successfully carried out also for energy functionals with weaker convexity properties. However, as soon as the first-order optimality condition for (1.9) is employed to determine a solution, which is the case in algorithms based on a FE-discretization in space, then convexity of $\mathcal{E}(t,\cdot,\cdot)$ in the pair (u,z) is needed to find energetic solutions, whereas separate convexity will in general result in semistable energetic solutions, only. A semistable energetic solution that satisfies (1.5c) as an *equality* is called *enhanced*.

Recently the concept of *Balanced Viscosity solutions* for rate-independent systems has gained attention, see, e.g., [MRS12]. This notion of solution can be obtained by introducing an additional viscous dissipation for z, weighted with a parameter ε.

As $\varepsilon \to 0$ a selection of solution for the resulting rate-independent system is made, which is characterized by a local stability condition and an energy-dissipation balance that features in comparison to (1.5c) additional dissipative terms which become active in particular in jump regimes of the solutions, see, e.g., also [KRZ13] in the setting of damage models. Using an arc-length reparametrization of the problems as $\varepsilon \to 0$ one finds that solutions of the limit problem satisfy an equilibrium criterion (1.10a) in terms of the partial weighted slopes of $\mathcal{F}(t,\cdot,\cdot) := \mathcal{E}(t,\cdot,\cdot) + \mathcal{R}(\cdot)$ with respect to u and z as well as an energy balance (1.10b), both expressed in a parametrized setting with an artificial parameter $s \in [0,S]$. More precisely, such a *parametrized Balanced Viscosity solution* of the $(\mathbf{U}\times\mathbf{Z},\mathcal{E},\mathcal{R})$ is a Lipschitz-continuous curve $(t,u,z):[0,S]\to[0,T]\times\mathbf{U}\times\mathbf{Z}$ such that $t(0)=0$, $t(S)=T$ and

$$\begin{aligned}&\text{for all } s\in[0,S] \text{ with } t'(s)>0:\\ &\qquad |\partial_u\mathcal{F}(t(s),u(s),z(s))|_{z(s)}=0, \quad |\partial_z\mathcal{F}(t(s),u(s),z(s))|_{u(s)}=0\,,\end{aligned} \tag{1.10a}$$

$$\begin{aligned}\text{for all } s\in[0,S]\colon \mathcal{F}(t(s),u(s),z(s)) &+ \int_0^s \big|\partial_u\mathcal{F}(t(\xi),u(\xi),z(\xi))\big|_{z(\xi)}\,\|u'(\xi)\|_{z(\xi)}\,\mathrm{d}\xi\\ &+\int_0^s \big|\partial_z\mathcal{F}(t(\xi),u(\xi),z(\xi))\big|_{u(\xi)}\,\|z'(\xi)\|_{u(\xi)}\,\mathrm{d}\xi\\ &=\mathcal{F}(t(0),u(0),z(0))+\int_0^s \partial_t\mathcal{F}(t(\xi),u(\xi),z(\xi))\,t'(\xi)\,\mathrm{d}\xi\end{aligned} \tag{1.10b}$$

Here, $|\partial_u\mathcal{F}(t,u,z)|_z$ and $|\partial_z\mathcal{F}(t,u,z)|_u$ are the partial weighted slopes and $\|u'\|_z$ and $\|z'\|_u$ are suitable weighted norms. It has been shown in [KN17, AN19] that solutions of the type (1.10) can be obtained for phase-field fracture problems with the aid of an alternating multi-step algorithm in time, cf. Remark 3.6 for some more details on this algorithm. In [ABN18] the convergence of alternating single- and multi-step algorithms in combination with FE-discretization is analyzed in the setting of L^2-gradient flows for the Ambrosio-Tortorelli phase-field fracture model with (non-vanishing) viscous regularization of the damage variable. For this model the authors show that solutions of the limit problem that satisfy the unidirectionality constraint $z(s)\geq z(t)$ for all $s\leq t\in[0,T]$ can be approximated by a posteriori truncated solutions of the discrete, unconstrained problems. Finally, we also refer to [AB19], where the P1 FE-approximation of the quasistatic evolution in terms of semistable energetic solutions is analyzed for the Ambrosio-Tortorelli functional, i.e., in the special case of $r=2$ in the damage regularization $\mathcal{G}$ in (1.3b), see also Remark 4.11.

Plan of the paper: Section 2 collects the notation and the setup used throughout our studies. Section 3 presents the numerical algorithms used to compute the discrete solutions and provides their immediate properties. Based on this, in Section 4, we prove the convergence of the discrete solutions to a semistable energetic solution in the sense of (1.5a), (1.5c) & (1.6) for the rate-independent system $(\mathbf{U}\times\mathbf{Z},\mathcal{E},\mathcal{R})$. Here, the limit passage from discrete to continuous in the semistability inequality, resulting in (1.6) with the emergence of the error term ERR for the limit, is the main part of the proof of Thm. 4.2. Due to the non-trivial interplay of the general gradient regularization $\mathcal{G}$ and the non-smooth dissipation potential $\mathcal{R}$ this step requires some detailed estimates for the gradients of the test functions used in the semistability inequality of the approximating problems set in P1 finite element spaces. These estimates are given in Lemmata 4.3 and 4.4; in particular the proof of the perturbed monotonicity estimate of Lemma 4.3 is of very technical nature and makes the main reason why our analysis focuses on space dimension $d=2$. Finally, in Section 5 we illustrate features of the rate-independent model problem via numerical experiments. We consider a benchmark problem often used within elastoplasticity in which a square membrane with hole is loaded along certain sides. The geometry and data of the problem specification lead to nontrivial damage regions. We compare different Sobolev regularizations of the damage variable and observe that this strongly affects the width of the transition regions between damaged and undamaged zones. The second class of experiments considers a square shaped domain which allows us to investigate the influence of symmetry properties of triangulations on the convergence behavior of numerical approximations. The results of the simulations indicate that no significant influence occurs so that the devised numerical scheme enjoys good practical stability and consistency properties and that the mesh conditions needed in the convergence theory are rather of technical nature. The extremal case $r=1$ is covered by our analysis but requires a slightly different numerical treatment and leads to longer computing times. A suitable algorithm for the treatment of the model in the case $r=1$ leading to a BV-regularization and results on the properties of the discrete solutions are discussed in [BMT18]. Since we expect that solutions do not significantly differ for r close to 1, we decided not to include corresponding simulations in the present work. The comparison provided in the numerical experiments of [BMT18] considers the cases $r=1$ and $r=2$ and reveals similar differences as the choices $r=1.1$ and $r=2$ presented here.

2 Setup and notation

Throughout this work, we consider the time interval $[0,\mathsf{T}]$ for some time horizon $\mathsf{T}>0$ and an open bounded Lipschitz domain $\Omega\subset\mathbb{R}^d$, $d=2$, with Dirichlet boundary $\Gamma_D\subset\partial\Omega$ with $(d-1)$-dimensional Hausdorff-measure $\mathcal{H}^{d-1}(\Gamma_D)>0$. We denote by $(\cdot,\cdot)$ the L^2-inner product, by $\|\cdot\|$ the L^2-norm, and by $|\cdot|$ the Euclidean norm on $\mathbb{R}^d$. Moreover, by $\mathrm{B}([0,\mathsf{T}],\bullet)$ we denote the space of functions f mapping time into a space $\bullet$, which are bounded and defined everywhere in $[0,\mathsf{T}]$.

Regarding the given data appearing in (1.3b) we make the following assumptions:

Assumption 2.1 (Assumptions on the given data)**.** *1 The function $f:\mathbb{R}\to\mathbb{R}$ is continuously differentiable and convex and such that $f|_{[0,1]}:[0,1]\to[a,b]$ is monotonically increasing and $f\equiv a$ on $(-\infty,0]$.*

2 The Lamé constants satisfy $\lambda,\mu>0$.

3 The extension of the Dirichlet datum is of regularity $u_{\mathrm{D}}\in C^1([0,\mathsf{T}],H^1(\Omega;\mathbb{R}^d))$ with $C_{u_{\mathrm{D}}}:=\|u_{\mathrm{D}}\|_{C^1([0,\mathsf{T}],H^1(\Omega;\mathbb{R}^d))}$.

4 The Neumann datum l_{Neu} is of regularity $l_{\mathrm{Neu}}\in C^1([0,\mathsf{T}],L^2(\Gamma_{\mathrm{Neu}};\mathbb{R}^d))$ with $C_{l_{\mathrm{Neu}}}:=\|l_{\mathrm{Neu}}\|_{C^1([0,\mathsf{T}],L^2(\Gamma_{\mathrm{Neu}};\mathbb{R}^d))}$.

As indicated in (1.2b) we consider the state space $\mathbf{X}$ for the damage variable to be either $\mathbf{X}:=W^{1,r}(\Omega)$ with $r\in(1,\infty)$ or $\mathbf{X}:=\mathrm{BV}(\Omega)$. For sequences $(v_N)_N\subset\mathbf{X}$ and their limit v we will always write

$$v_N\rightharpoonup v \text{ in } \mathbf{X}, \text{ with the meaning } \begin{cases} v_N\rightharpoonup v & \text{in } W^{1,r}(\Omega) \text{ for } r\in(1,\infty),\\ v_N\overset{*}{\rightharpoonup} v & \text{in } \mathrm{BV}(\Omega).\end{cases} \tag{2.1}$$

For the space discretization we will use the following notation related to **finite element spaces**: Let $(\mathcal{T}_h)_{h>0}$ be a family of triangulations of Ω where the index h denotes the mesh size $h=\max_{T\in\mathcal{T}_h}h_T$ with h_T being the diameter of the simplex T. The minimal diameter is given by $h_{\min}=\min_{T\in\mathcal{T}_h}h_T$. The sets $\mathcal{N}_h$ and E_h contain all nodes and edges, respectively, of the triangulation $\mathcal{T}_h$. We will use the finite element space of continuous, piecewise affine functions ($\gamma=1$) or vector fields ($\gamma=d$), denoted by $\mathcal{S}^1(\mathcal{T}_h)^\gamma$ and of elementwise constant vector fields $\mathcal{L}^0(\mathcal{T}_h)^d$, i.e.,

$$\mathcal{S}^1(\mathcal{T}_h)^\gamma := \{v_h\in C(\Omega;\mathbb{R}^\gamma):\ v_h|_T \text{ affine for all } T\in\mathcal{T}_h\}, \tag{2.2a}$$

$$\mathcal{L}^0(\mathcal{T}_h)^d := \{\tilde{p}_h\in L^\infty(\Omega;\mathbb{R}^d):\ \tilde{p}_h|_T \text{ constant for all } T\in\mathcal{T}_h\}. \tag{2.2b}$$

Moreover, denoting by $\mathcal{I}_h:C^0(\overline{\Omega})\to\mathcal{S}^1(\mathcal{T}_h)$ the standard nodal interpolation operator we will consider the discrete inner products

$$(v_h,w_h)_h:=\int_\Omega\mathcal{I}_h[v_hw_h]\,\mathrm{d}x=\sum_{y\in\mathcal{N}_h}\beta_yv_h(y)w_h(y)\quad\text{on }\mathcal{S}^1(\mathcal{T}_h),$$

$$(p_h,\tilde{p}_h)_{w_r}:=h_{\min}^{\max\{0,d(2/r-1)\}}(p_h,\tilde{p}_h)\quad\text{on }\mathcal{L}^0(\mathcal{T}_h)^d,$$

where $\beta_y=\int_\Omega\varphi_y\,\mathrm{d}x$ with φ_y the nodal basis function associated to $y\in\mathcal{N}_h$. We have the relations

$$\|v_h\|\le\|v_h\|_h\le(d+2)^{1/2}\|v_h\|,\quad\text{and}\quad\|\tilde{p}_h\|_{w_r}\le c\|\tilde{p}_h\|_{L^r(\Omega)},$$

for all $v_h\in\mathcal{S}^1(\mathcal{T}_h)$ and $\tilde{p}_h\in\mathcal{L}^0(\mathcal{T}_h)^d$, see [Bar15, Lemma 3.9] and [BS08, Thm. 4.5.11]. Finally, for a sequence of step sizes $(\tau_j)_{j\in\mathbb{N}}$ and functions $(a^j)_{j\in\mathbb{N}}$ we will denote the backward difference quotient by

$$d_ta^j=\frac{a^j-a^{j-1}}{\tau_j}.$$

3 Numerical Method

We now discuss the numerical algorithms used to solve the alternate minimization problem (1.4) on the discrete level. With $\mathcal{S}^1(\mathcal{T}_h)^d$ and $\mathcal{S}^1(\mathcal{T}_h)$ from (2.2) we set $\mathbf{U}_h:=\mathcal{S}^1(\mathcal{T}_h)^d\cap\{v\in C(\overline{\Omega};\mathbb{R}^d),v=0\text{ on }\Gamma_{\mathrm{D}}\}\subset\mathbf{U}=H^1_{\mathrm{D}}(\Omega;\mathbb{R}^d)$ in (1.4a) and $\mathbf{X}_h:=\mathcal{S}^1(\mathcal{T}_h)\subset\mathbf{X}=W^{1,r}(\Omega)$, resp. $\mathrm{BV}(\Omega)$, in (1.4b). While the minimization problem (1.4a) to determine u_{Nh}^k reduces to the solution of a linear system of equations, the minimization problem (1.4b) to find z_{Nh}^k is more difficult due to the nonlinearity of the gradient-regularization for z_{Nh}^k for $r\neq2$, being even non-differentiable in case of the BV-seminorm for $r=1$, and the occurence of non-smooth constraints in $\mathcal{E}$ and $\mathcal{R}$. We will deal with the minimization problem (1.4b) in Subsection 3.1 and subsequently explain the algorithm for the full alternate minimization problem in Subsection 3.2.

3.1 Minimization with respect to z in (1.4b)

For the following discussion we consider a partition Π_N of $[0,\mathsf{T}]$ with $N \in \mathbb{N}$ fixed. We also keep $t_N^k \in \Pi_N$ and u_{Nh}^k the solution of (1.4a) fixed. For simpler notation we here write $t_k = t_N^k$, $u_h^k = u_{Nh}^k$, and $z_h^k = z_{Nh}^k$, i.e., we do not indicate the dependence of these quantities on $N \in \mathbb{N}$ fixed. We first of all note that a minimizer $z_h^k = z_{Nh}^k$ obtained in (1.4b) is required to satisfy $z_h^k - z_h^{k-1} \leq 0$ almost everywhere in Ω since otherwise $\mathcal{R}(z_h^k - z_h^{k-1})$ is infinite. Since $z_h^k, z_h^{k-1} \in \mathbf{X}_h = \mathcal{S}^1(\mathcal{T}_h)$ are globally continuous and piecewise affine this is equivalent to $z_h^k(x) \leq z_h^{k-1}(x)$ for all $x \in \mathcal{N}_h$. Particularly, $|z_h^k(x) - z_h^{k-1}(x)| = z_h^{k-1}(x) - z_h^k(x)$. Hence, letting for $k \geq 1$

$$K_k := \{v_h \in \mathcal{S}^1(\mathcal{T}_h) : \ v_h(x) \leq z_h^{k-1}(x) \ \forall\, x \in \mathcal{N}_h\} \tag{3.1}$$

we define the auxiliary functional $\mathcal{F}(t_k,\cdot,\cdot) : \mathbf{U}_h \times \mathbf{X}_h \to \mathbb{R} \cup \{\infty\}$,

$$\begin{aligned}\mathcal{F}(t_k,u_h,z_h) := \frac{1}{2}\int_\Omega f(z_h)\big(\lambda|\operatorname{tr} e(u_h + u_{\mathrm{D}}(t_k))|^2 + 2\mu|e(u_h+u_{\mathrm{D}}(t_k))|^2\big)\,\mathrm{d}x \\ - \int_{\Gamma_{\mathrm{Neu}}} l_{\mathrm{Neu}}(t_k)\cdot(u_h + u_{\mathrm{D}}(t_k))\,\mathrm{d}s + \mathcal{G}(z_h) + \int_\Omega \varrho(-z_h)\,\mathrm{d}x + I_{K_k}(z_h).\end{aligned}$$

We obtain that minimality property (1.4b) is equivalent to

$$z_h^k \in \operatorname{argmin}_{z_h \in \mathbf{X}_h} \mathcal{F}(t_k,u_h,z_h)\,.$$

In order to approximate a minimizer z_h^k we consider for $\tau_j > 0$ and $\mathbb{C}A = \lambda \operatorname{tr}(A) I + 2\mu A$ for $A \in \mathbb{R}^{d\times d}$ the augmented Lagrangian functional

$$\begin{aligned}L_h^k(z_h,p_h,s_h;\eta_h,\zeta_h) := \frac{1}{2}\int_\Omega \big(f(z_h)e(u_h^k + u_{\mathrm{D}}(t_k)) : \mathbb{C}e(u_h^k + (t_k)) + \varrho(-z_h)\big)\,\mathrm{d}x \\ + \int_\Omega \big(\tfrac{\kappa_1}{r}|p_h|^r + \tfrac{\kappa_2}{r}|s_h|^r\big)\,\mathrm{d}x + (\eta_h, \nabla z_h - p_h)_{w_r} + \frac{\tau_j}{2}\|\nabla z_h - p_h\|_{w_r}^2 \\ + I_{K_k}(s_h) + (\zeta_h, z_h - s_h)_h + \frac{\tau_j}{2}\|z_h - s_h\|_h^2.\end{aligned} \tag{3.2}$$

For the approximation of a minimizer z_h^k we use the following algorithm [BM17] which generalizes the alternating direction method of multipliers (ADMM) established and analyzed, e.g., in [FG83, GM76, Glo84, GLT89, GM75] by using variable step sizes.

Algorithm 3.1 (Variable-ADMM). *Choose $z_h^0 = z_h^{k-1}$, $\eta_h^0 = 0$ and $\zeta_h^0 = 0$. Choose $\underline{\tau}, \overline{\tau} > 0$ with $\underline{\tau} \leq \overline{\tau}$, $\delta \in (0,1)$, $\underline{\gamma}, \overline{\gamma} \in (0,1)$ with $\underline{\gamma} \leq \overline{\gamma}$, and $\overline{R} \gg 1$. Set $j = 1$.*

(1) Set $\gamma_1 = \underline{\gamma}$, $\tau_1 = \overline{\tau}$ and $R_0 = \overline{R}$.

(2) Compute a minimizer $(p_h^j, s_h^j) \in \mathcal{L}^0(\mathcal{T}_h)^d \times \mathcal{S}^1(\mathcal{T}_h)$ of

$$(p_h,s_h) \mapsto L_h^k(z_h^{j-1},p_h,s_h;\eta_h^{j-1},\zeta_h^{j-1}).$$

(3) Compute a minimizer $z_h^j \in \mathcal{S}^1(\mathcal{T}_h)$ of

$$z_h \mapsto L_h^k(z_h,p_h^j,s_h^j;\eta_h^{j-1},\zeta_h^{j-1}).$$

(4) Update

$$\begin{aligned}\eta_h^j &= \eta_h^{j-1} + \tau_j(\nabla z_h^j - p_h^j),\\ \zeta_h^j &= \zeta_h^{j-1} + \tau_j(z_h^j - s_h^j).\end{aligned}$$

(5) Define

$$R_j = \big(\|\eta_h^j - \eta_h^{j-1}\|_{w_r}^2 + \tau_j^2\|\nabla(z_h^j - z_h^{j-1})\|_{w_r}^2 + \|\zeta_h^j - \zeta_h^{j-1}\|_h^2 + \tau_j^2\|z_h^j - z_h^{j-1}\|_h^2\big)^{1/2}.$$

(6) Stop if R_j is sufficiently small.

(7) Define $(\tau_{j+1},\gamma_{j+1})$ as follows:

- *If* $R_j \le \gamma_j R_{j-1}$ *or if* $\tau_j = \underline{\tau}$ *and* $\gamma_j = \overline{\gamma}$ *set*

$$\tau_{j+1} = \tau_j \quad \textit{and} \quad \gamma_{j+1} = \gamma_j.$$

- *If* $R_j > \gamma_j R_{j-1}$ *and* $\tau_j > \underline{\tau}$ *set*

$$\tau_{j+1} = \max\{\delta\tau_j, \underline{\tau}\} \quad \textit{and} \quad \gamma_{j+1} = \gamma_j.$$

- *If* $R_j > \gamma_j R_{j-1}, \tau_j = \underline{\tau}$ *and* $\gamma_j < \overline{\gamma}$ *set*

$$\tau_{j+1} = \overline{\tau},\ \gamma_{j+1} = \min\Big\{\frac{\gamma_j + 1}{2}, \overline{\gamma}\Big\},\ u^j = u^0 \textit{ and } \lambda^j = \lambda^0.$$

(8) Set $j = j + 1$ *and continue with (2).*

In the following proposition we prove that the iterates are bounded, that the algorithm terminates and that the residuals R_j are monotonically decreasing.

Proposition 3.2 (Termination of Alg. 3.1 and monotonicity of residuals)**.** *Let* $(z_h, p_h, s_h; \eta_h, \zeta_h)$ *be a saddle-point for* L_h^k*. For the iterates* $(z_h^j, p_h^j, s_h^j; \eta_h^j, \zeta_h^j)$*,* $j \ge 0$*, of Algorithm 3.1 and the distance*

$$D_j^2 = \|\eta_h - \eta_h^j\|_{w_r}^2 + \|\zeta_h - \zeta_h^j\|_h^2 + \tau_j^2\|\nabla(z_h - z_h^j)\|_{w_r}^2 + \tau_j^2\|z_h - z_h^j\|_h^2,$$

we have that for every $J \ge 1$ *it holds*

$$D_J^2 + \sum_{j=1}^{J} R_j^2 \le D_0^2.$$

In particular, $R_j \to 0$ *as* $j \to \infty$ *and Algorithm 3.1 terminates. Moreover, we have*

$$R_{j+1}^2 \le R_j^2,$$

i.e., the residual is non-increasing.

Proof. For shorter notation we define the convex functionals

$$\begin{aligned}
\widetilde{\mathcal{G}}(p_h, s_h) &= \int_\Omega \big(\tfrac{\kappa_1}{r}|p_h|^r + \tfrac{\kappa_2}{r}|s_h|^r\big)\,\mathrm{d}x + I_{K_k}(s_h),\\
\widetilde{\mathcal{W}}(z_h) &= \frac{1}{2}\int_\Omega \big(f(z_h)e(u_h^k + u_{\mathrm{D}}(t_k)) : \mathbb{C}e(u_h^k + u_{\mathrm{D}}(t_k))\big)\,\mathrm{d}x + \int_\Omega \varrho(-z_h)\,\mathrm{d}x,
\end{aligned}$$

so that the augmented Lagrangian from (3.2) now equivalently rewrites as

$$\begin{aligned}
L_h^k(z_h, p_h, s_h; \eta_h, \zeta_h) =& \widetilde{\mathcal{W}}(z_h) + \widetilde{\mathcal{G}}(p_h, s_h)\\
&+ (\eta_h, \nabla z_h - p_h)_{w_r} + \frac{\tau_j}{2}\|\nabla z_h - p_h\|_{w_r}^2 + (\zeta_h, z_h - s_h)_h + \frac{\tau_j}{2}\|z_h - s_h\|_h^2.
\end{aligned}$$

The optimality conditions for a saddle-point $(z_h, p_h, s_h; \eta_h, \zeta_h)$ of L_h^k are given by

$$\begin{aligned}
(\eta_h, \tilde{p}_h - p_h)_{w_r} + (\zeta_h, \tilde{s}_h - s_h)_h + \widetilde{\mathcal{G}}(p_h, s_h) \le \widetilde{\mathcal{G}}(\tilde{p}_h, \tilde{s}_h) \quad &\forall\ (\tilde{p}_h, \tilde{s}_h) \in \mathcal{L}^0(\mathcal{T}_h)^d \times \mathcal{S}^1(\mathcal{T}_h),\\
-(\eta_h, \nabla(\tilde{z}_h - z_h))_{w_r} - (\zeta_h, \tilde{z}_h - z_h)_h + \widetilde{\mathcal{W}}(z_h) \le \widetilde{\mathcal{W}}(\tilde{z}_h) \quad &\forall\ \tilde{z}_h \in \mathcal{S}^1(\mathcal{T}_h),
\end{aligned} \tag{3.3}$$

with $p_h = \nabla z_h$ and $s_h = z_h$. On the other hand, for the iterates $(z_h^j, p_h^j, s_h^j; \eta_h^j, \zeta_h^j)$ of Algorithm 3.1, with $\tilde{\eta}_h^j = \eta_h^{j-1} + \tau_j(\nabla z_h^{j-1} - p_h^j)$ and $\tilde{\zeta}_h^j = \zeta_h^{j-1} + \tau_j(z_h^{j-1} - s_h^j)$, the optimality conditions read

$$\begin{aligned}
(\tilde{\eta}_h^j, \tilde{p}_h - p_h^j)_{w_r} + (\tilde{\zeta}_h^j, \tilde{s}_h - z_h^j)_h + \widetilde{\mathcal{G}}(p_h^j, s_h^j) \le \widetilde{\mathcal{G}}(\tilde{p}_h, \tilde{s}_h) \quad &\forall\ (\tilde{p}_h, \tilde{s}_h) \in \mathcal{L}^0(\mathcal{T}_h)^d \times \mathcal{S}^1(\mathcal{T}_h),\\
-(\eta_h^j, \nabla(\tilde{z}_h - z_h^j))_{w_r} - (\zeta_h^j, \tilde{z}_h - z_h^j)_h + \widetilde{\mathcal{W}}(z_h^j) \le \widetilde{\mathcal{W}}(\tilde{z}_h) \quad &\forall\ \tilde{z}_h \in \mathcal{S}^1(\mathcal{T}_h).
\end{aligned} \tag{3.4}$$

Testing (3.3) and (3.4) with $(\tilde{p}_h, \tilde{s}_h, \tilde{z}_h) = (p_h^j, s_h^j, z_h^j)$ and $(\tilde{p}_h, \tilde{s}_h, \tilde{z}_h) = (p_h, s_h, z_h)$, respectively, and adding corresponding inequalities gives

$$\begin{aligned}(\tilde{\eta}_h^j - \eta_h, p_h - p_h^j)_{w_r} + (\tilde{\zeta}_h^j - \zeta_h, s_h - s_h^j)_h \leq 0,\\ (\eta_h - \eta_h^j, \nabla(z_h - z_h^j))_{w_r} + (\zeta_h - \zeta_h^j, z_h - z_h^j)_h \leq 0.\end{aligned}$$

The rest of the proof of the first estimate is analogous to the proof of [BM17, Thm. 3.7].

The proof of the monotonicity follows by testing (3.4) at iterations j and $j+1$ with $(\tilde{p}_h, \tilde{s}_h, \tilde{z}_h) = (p_h^{j+1}, s_h^{j+1}, z_h^{j+1})$ and $(\tilde{p}_h, \tilde{s}_h, \tilde{z}_h) = (p_h^j, s_h^j, z_h^j)$, respectively, and adding the inequalities, which gives

$$\begin{aligned}0 \leq\ &-(\tilde{\eta}_h^{j+1} - \tilde{\eta}_h^j, p_h^j - p_h^{j+1})_{w_r} - (\eta_h^j - \eta_h^{j+1}, \nabla(z_h^j - z_h^{j+1}))_{w_r}\\ &-(\tilde{\zeta}_h^{j+1} - \tilde{\zeta}_h^j, s_h^j - s_h^{j+1})_h - (\zeta_h^j - \zeta_h^{j+1}, z_h^j - z_h^{j+1})_h.\end{aligned}$$

The monotonicity then follows as in the proof of [BM17, Prop. 3.11]. □

In the next step, we show that the residual R_j controls the difference in the objective values.

Lemma 3.3. *Let $(z_h, p_h, s_h; \eta_h, \zeta_h)$ be a saddle-point of L_h^k. Then there exists a constant $C_0 > 0$ such that we have for any $j \geq 1$*

$$|\mathcal{E}(t_k, u_h^k, s_h^j) + \mathcal{R}(s_h^j - z_h^{k-1}) - \mathcal{E}(t_k, u_h^k, z_h) - \mathcal{R}(z_h - z_h^{k-1})| \leq C_0 R_j. \tag{3.5}$$

Proof. We use the short notation $\delta_\eta^j = \eta_h - \eta_h^j$, $\delta_\zeta^j = \zeta_h - \zeta_h^j$, $\delta_p^j = p_h - p_h^j$, $\delta_s^j = s_h - s_h^j$ and $\delta_z^j = z_h - z_h^j$. Testing (3.4) with $(\tilde{p}_h, \tilde{s}_h, \tilde{z}_h) = (p_h, s_h, z_h)$, adding the inequalities, noting that $p_h = \nabla z_h$ and $s_h = z_h$ and using $\eta_h^j - \tilde{\eta}_h^j = \tau_j \nabla(z_h^j - z_h^{j-1})$, $\zeta_h^j - \tilde{\zeta}_h^j = \tau_j(z_h^j - z_h^{j-1})$ we obtain

$$\begin{aligned}&\widetilde{\mathcal{G}}(p_h^j, s_h^j) + \widetilde{\mathcal{W}}(z_h^j) - \widetilde{\mathcal{G}}(p_h, s_h) - \widetilde{\mathcal{W}}(z_h)\\ &\leq\ -(\tilde{\eta}_h^j, \delta_p^j)_{w_r} + (\eta_h^j, \nabla\delta_z^j)_{w_r} - (\tilde{\zeta}_h^j, \delta_s^j)_h + (\zeta_h^j, \delta_z^j)_h\\ &=\ -(\eta_h^j, d_t\eta_h^j)_{w_r} - \tau_j^2(\nabla d_t\delta_z^j, \delta_p^j)_{w_r} - (\zeta_h^j, d_t\zeta_h^j)_h - \tau_j^2(d_t\delta_z^j, \delta_s^j)_h.\end{aligned} \tag{3.6}$$

Testing the optimality conditions of z_h^j and z_h^{j-1} with $\tilde{z}_h = z_h^{j-1}$ and $\tilde{z}_h = z_h^j$, respectively, and adding the corresponding inequalities gives

$$0 \leq -\tau_j^2(d_t\eta_h^j, \nabla d_t z_h^j)_{w_r} - \tau_j^2(d_t\zeta_h^j, d_t z_h^j)_h.$$

Using $d_t\eta_h^j = \nabla z_h^j - p_h^j$ and $d_t\zeta_h^j = z_h^j - s_h^j$ and inserting $p_h = \nabla z_h$ and $s_h = z_h$ on the right-hand side gives

$$0 \leq -\tau_j^2(\nabla\delta_z^j, \nabla d_t\delta_z^j)_{w_r} + \tau_j^2(\delta_p^j, \nabla d_t\delta_z^j)_{w_r} - \tau_j^2(\delta_z^j, d_t\delta_z^j)_h + \tau_j^2(\delta_s^j, d_t\delta_z^j)_h. \tag{3.7}$$

Adding (3.6) and (3.7) we get

$$\begin{aligned}&\widetilde{\mathcal{G}}(p_h^j, s_h^j) + \widetilde{\mathcal{W}}(z_h^j) - \widetilde{\mathcal{G}}(p_h, s_h) - \widetilde{\mathcal{W}}(z_h)\\ &\leq\ -(\eta_h^j, d_t\eta_h^j)_{w_r} + \tau_j^2(\nabla\delta_z^j, \nabla d_t z_h^j)_{w_r} - (\zeta_h^j, d_t\zeta_h^j)_h + \tau_j^2(\delta_z^j, d_t z_h^j)_h\\ &\leq \|\eta_h^j\|_{w_r}\|d_t\eta_h^j\|_{w_r} + \tau_j^2\|\nabla\delta_z^j\|_{w_r}\|\nabla d_t z_h^j\|_{w_r} + \|\zeta_h^j\|_h\|d_t\zeta_h^j\|_h + \tau_j^2\|\delta_z^j\|_h\|d_t z_h^j\|_h \leq C_0 R_j,\end{aligned}$$

with C_0 being bounded due to Proposition 3.2.

Let us furthermore note that by Proposition 3.2 we have that s_h^j and z_h^j are bounded, particularly $s_h^j \leq z_h^{k-1}$ for all $j \geq 0$. Since f is Lipschitz continuous on bounded intervals, the Hölder inequality, the Lipschitz continuity of f and the inverse estimate $\|w_h\|_{L^\infty(\Omega)} \leq h^{-d/2}\|w_h\|$ (cf. [BS08, Thm. 4.5.11]) yield

$$\frac{1}{2}\int_\Omega (f(s_h^j) - f(z_h^j))e(u_h^k + u_{\mathrm{D}}(t_k)) : \mathbb{C}e(u_h^k + u_{\mathrm{D}}(t_k))\,\mathrm{d}x \leq ch^{-d/2}\|s_h^j - z_h^j\|.$$

We finally observe that using $s_h^j \leq z_h^{k-1}$, $z_h \leq z_h^{k-1}$, the triangle inequality, the inverse estimates $\|\nabla w_h\|_{L^r(\Omega)} \leq$

$ch^{-1}\|w_h\|_{L^r(\Omega)}, \|w_h\|_{L^r(\Omega)} \le ch^{\min\{0,d/r-d/2\}}\|w_h\|_{L^2(\Omega)}$ and the equivalence of $\|\cdot\|$ and $\|\cdot\|_h$ we have

$$\begin{aligned}
&\mathcal{E}(t_k,u_h^k,s_h^j)+\mathcal{R}(s_h^j-z_h^{k-1})-\mathcal{E}(t_k,u_h^k,z_h)-\mathcal{R}(z_h-z_h^{k-1})\\
&=\widetilde{\mathcal{G}}(p_h^j,s_h^j)+\widetilde{\mathcal{W}}(z_h^j)-\widetilde{\mathcal{G}}(p_h,s_h)-\widetilde{\mathcal{W}}(z_h)+\frac{\kappa_1}{r}\int_\Omega\left(|\nabla s_h^j|^r-|p_h^j|^r\right)\mathrm{d}x\\
&\quad+\frac{1}{2}\int_\Omega(f(s_h^j)-f(z_h^j))e(u_h^k+u_{\mathrm{D}}(t_k)):\mathbb{C}e(u_h^k+u_{\mathrm{D}}(t_k))\,\mathrm{d}x+\rho\int_\Omega\left(z_h^j-s_h^j\right)\mathrm{d}x\\
&\le C_0R_j+c\kappa h^{-|d/2-d/r|}\|\nabla z_h^j-p_h^j\|_{w_r}+c\kappa h^{-1}h^{\min\{0,d/r-d/2\}}\|s_h^j-z_h^j\|_h+c(\rho+h^{-d/2})\|z_h^j-s_h^j\|_h\\
&\le C_0R_j.
\end{aligned}$$

This proves the assertion, taking into account that z_h as a minimizer of $\left(\mathcal{E}(t_k,u_h^k,\cdot)+\mathcal{R}(\cdot-z_h^{k-1})\right)$ in particular satisfies $\mathcal{E}(t_k,u_h^k,z_h)+\mathcal{R}(z_h-z_h^{k-1})\le\mathcal{E}(t_k,u_h^k,s_h^j)+\mathcal{R}(s_h^j-z_h^{k-1})$. □

Remark 3.4. *In general, the iterates $(z_h^j)_{j\ge0}$ of Algorithm 3.1 may penetrate the obstacles, i.e., $z_h^j\notin K_k$ for some $j\in\mathbb{N}$, cf.* (3.1)*. Therefore, if $(z_h^{stop},p_h^{stop},s_h^{stop};\eta_h^{stop},\zeta_h^{stop})$ is the output of the algorithm, we set $z_h^k=s_h^{stop}\in K_k$ to ensure the coercivity of the bulk energy.*

3.2 Alternate minimization (1.4)

In order to solve the full problem (1.4) we apply the following scheme:

Algorithm 3.5 (Alternate Minimization)**.** *For $N\in\mathbb{N}$ choose $h=h(N)$ as well as a stable initial pair $(u_h^0,z_h^0)\in\mathcal{S}^1(\mathcal{T}_h)^d\times\mathcal{S}^1(\mathcal{T}_h)$ and a partition $0=t=0<\ldots<t_N=\mathsf{T}$ of the time interval and set $k=1$.*
(1) Compute the unique minimizer u_h^k of

$$u_h\mapsto\mathcal{F}(t_k,u_h,z_h^{k-1}).$$

(2) Compute an approximate minimizer z_h^k of

$$z_h\mapsto\mathcal{F}(t_k,u_h^k,z_h)$$

by using Algorithm 3.1, i.e., set $z_h^k=s_h^{stop}$ with s_h^{stop} computed by Algorithm 3.1.
(3) Stop if $k=N$. Otherwise, increase $k\to k+1$ and continue with (1).

The optimality condition for u_h^k in step (1) of the algorithm reads

$$\int_\Omega f(z_h^{k-1})e(u_h^k):\mathbb{C}e(v_h)\,\mathrm{d}x=-\int_\Omega f(z_h^{k-1})e(u_{\mathrm{D}}(t_k)):\mathbb{C}e(v_h)\,\mathrm{d}x+\int_{\Gamma_{\mathrm{Neu}}}l_{\mathrm{Neu}}(t_k)\cdot v_h\,\mathrm{d}s$$

for all $v_h\in\mathbf{U}_h$. In our computation we replace u_{D} by $u_{\mathrm{D}h}=\mathcal{I}_hu_{\mathrm{D}}$ on the right-hand side with $\mathcal{I}_h$ being the nodal interpolant and u_{D} sufficiently smooth. We further use the midpoint rule to compute for $T\in\mathcal{T}_h$ and $e\in E_h$ the integrals

$$\int_T f(z_h^{k-1})\,\mathrm{d}x,\quad\text{and}\quad\int_e l_{\mathrm{Neu}}(t_k)\cdot v_h\,\mathrm{d}s.$$

The computation of u_h^k then amounts to solving a linear system of equations with a weighted stiffness matrix.

Remark 3.6 (Comparison of Algorithm 3.5 with a multi-step algorithm from [KN17, AN19])**.** For each time-step size $N\in\mathbb{N}$ and mesh-size $h=h(N)$ fixed, for each $k\in\{1,\ldots,N\}$ Algorithm 3.5 first determines u_h^k by solving the linear system corresponding to the minimization of $\mathbf{U}_h\ni\tilde{u}\mapsto\mathcal{F}(t_k,\tilde{u},z_h^{k-1})$ and secondly determines an approximate minimizer z_h^k of the non-smooth functional $\mathbf{X}\ni\tilde{z}\mapsto\mathcal{F}(t_k,u_h^k,\tilde{z})$ using the Variable ADMM Algorithm 3.1. In other words, when solving for u_h^k the approximate solution z_h^{k-1} determined in the previous time step is kept fixed and, in turn, u_h^k is kept fixed when determining an approximate solution z_h^k. Thus, Algorithm 3.5 carries out alternate minimization in a single step at each time step. In comparison, without discretization in space, the works [KN17, AN19] propose a multi-step algorithm for the approximation procedure in time: Consider partitions $\Pi_N=\{t_N^k,\,k=1,\ldots,N\}$ of $[0,T]$. Let (u_N^{k-1},z_N^{k-1}) be a solution determined by the algorithm at time t_N^{k-1}. For every $k\in\{1,\ldots,N\}$, determine a solution (u_N^k,z_N^k) at time t_N^k as follows:

- Set $(u_N^{k,0},z_N^{k,0}):=(u_N^{k-1},z_N^{k-1})$.

- For every $i \in \{1, \dots, i_{\max}\}$ (with $i_{\max}$ imposed by some stopping criterion) determine

$$\begin{aligned} u_N^{k,i} &= \operatorname*{Argmin}_{u \in \mathbf{U}} \mathcal{F}(t_N^k, u, z_N^{k,i-1}) \,, \\ z_N^{k,i} &\in \operatorname*{argmin}_{z \in \mathbf{Z}} \{\mathcal{F}(t_N^k, u_N^{k,i}, z),\ z \le z_N^{k,i-1}\} \,. \end{aligned}$$

- Set $(u_N^k, z_N^k) := (u_N^{k,i_{\max}}, z_N^{k,i_{\max}})$.

In [KN17, AN19] the functional $\mathcal{F}$ is given by the Ambrosio-Tortorelli functional for phase-field fracture. In this setting the authors show that the piecewise constant interpolants in time obtained with the solutions of the multi-step algorithm approximate a *parametrized Balanced Viscosity solution* of the time-continous limit system, see Remark 1.3.

4 Existence of semistable energetic solutions for $(\mathbf{U} \times \mathbf{Z}, \mathcal{E}, \mathcal{R})$

In this section we show that suitable time-interpolants of the solutions $(u_{Nh}^k, z_{Nh}^k)_{Nh}$ obtained at each time step t_N^k via the alternate minimization problem (1.4) can be used to approximate a semistable energetic solution of system $(\mathbf{U} \times \mathbf{Z}, \mathcal{E}, \mathcal{R})$. To this end, with $\mathcal{S}^1(\mathcal{T}_h)^d$ and $\mathcal{S}^1(\mathcal{T}_h)$ from (2.2), we set in (1.4)

$$\mathbf{U}_h := \mathcal{S}^1(\mathcal{T}_h)^d \cap \{v \in C(\overline{\Omega}; \mathbb{R}^d), v = 0 \text{ on } \Gamma_D\} \text{ and } \mathbf{X}_h := \mathcal{S}^1(\mathcal{T}_h) \,. \tag{4.1}$$

We recall that $\mathbf{U}_h \subset H^1_D(\Omega; \mathbb{R}^d)$ and $\mathbf{X}_h \subset BV(\Omega)$ for all $h > 0$ and

$$\bigcup_h \mathbf{U}_h \subset H^1_D(\Omega; \mathbb{R}^d) \text{ densely and } \bigcup_h \mathbf{X}_h \subset \mathbf{X} \text{ densely.} \tag{4.2}$$

We now choose a sequence $(h(N))_{N \in \mathbb{N}}$ such that $h(N) \to 0$ as $N \to \infty$ and consider a sequence of partitions $(\Pi_N)_N$ of $[0, \mathsf{T}]$ such that the time-step size $\Delta_N \to 0$ as $N \to \infty$. With $\mathcal{E}$ from (1.3) we introduce the energy functionals $\mathcal{E}_N : [0, \mathsf{T}] \times \mathbf{U} \times \mathbf{Z} \to \mathbb{R} \cup \{\infty\}$,

$$\mathcal{E}_N(t, u, z) := \begin{cases} \mathcal{E}(t, u, z) & \text{if } (u, z) \in \mathbf{U}_{h(N)} \times \mathbf{X}_{h(N)}, \\ \infty & \text{otherwise,} \end{cases} \tag{4.3}$$

where the given data $u_{\mathrm{D}}(t)$ and $l_{\mathrm{Neu}}(t)$ are replaced by suitably interpolated versions $u_{\mathrm{D}N}(t)$ and $l_{\mathrm{Neu}N}(t)$ in the discrete spaces, which are uniformly bounded and converge strongly to the original datum. We thus compute for every $N \in \mathbb{N}$ and $h(N) > 0$, for each $t_N^k \in \Pi_N$ a solution $(u_N^k, z_N^k) = (u_{Nh(N)}^k, z_{Nh(N)}^k)$ to (1.4) using Algorithm 3.5. In particular, according to Algorithm 3.1 the pair $(u_N^k, z_N^k) = (u_{Nh(N)}^k, z_{Nh(N)}^k)$ satisfies

$$\forall\, u \in \mathbf{U} : \ \mathcal{E}_N(t_N^k, u_N^k, z_N^{k-1}) \le \mathcal{E}_N(t_N^k, u, z_N^{k-1}) \,, \tag{4.4a}$$

$$\forall\, z \in \mathbf{X} : \\ \mathcal{E}_N(t_N^k, u_N^k, z_N^k) + \mathcal{R}(z_N^k - z_N^{k-1}) \le \mathcal{E}_N(t_N^k, u_N^k, z) + \mathcal{R}(z - z_N^{k-1}) + \mathrm{TOL}(N) \tag{4.4b}$$

with some $h(N)$-dependent tolerance $\mathrm{TOL}(N)$, which bounds the residual R_j^h, cf. Algorithm 3.1, Step (5). In view of Lemma 3.3 a sequence $(\mathrm{TOL}(N))_N$ can be chosen such that

$$\mathrm{TOL}(N) N \to 0 \text{ as } N \to \infty \,. \tag{4.5}$$

We evaluate the given data in the partition $\{t_N^0, \dots, t_N^N\}$ which results in an $(N+1)$-tupel. Moreover, for any tupel $(v_N^0, \dots, v_N^N)$ we introduce the piecewise constant left-continuous (right-continuous) interpolant $\overline{v}_N$ ($\underline{v}_N$):

$$\overline{v}_N(t) := v_N^{k+1} \text{ for all } t \in (t_N^k, t_N^{k+1}] \,, \tag{4.6a}$$

$$\underline{v}_N(t) := v_N^k \text{ for all } t \in [t_N^k, t_N^{k+1}) \,. \tag{4.6b}$$

Accordingly, $\overline{\mathcal{E}}$, resp. $\underline{\mathcal{E}}$, indicates that the interpolants $\overline{u_{\mathrm{D}}}_N$ and $\overline{l_{\mathrm{Neu}}}_N$, resp. $\underline{u_{\mathrm{D}}}_N$ and $\underline{l_{\mathrm{Neu}}}_N$ are used. In particular, thanks to Assumptions 2.1 we have for all $t \in [0, \mathsf{T}]$

$$\overline{u_{\mathrm{D}}}_N(t) \to u_{\mathrm{D}}(t) \text{ in } \mathbf{U} \ \ \& \ \ \overline{l_{\mathrm{Neu}}}_N(t) \to l_{\mathrm{Neu}}(t) \text{ in } L^2(\Gamma_{\mathrm{Neu}}; \mathbb{R}^d) \,. \tag{4.7}$$

This puts us in the position to find the following properties of the interpolants $(\overline{u}_N, \underline{u}_N, \overline{z}_N, \underline{z}_N)$ constructed from $(u_N^k, z_N^k)_{k=0}^N$ via (4.6):

Proposition 4.1 (Discrete version of (1.5) and apriori estimates)**.** *Let the assumptions of Section 2 hold true and keep $N \in \mathbb{N}$ fixed. For each $k \in \{0, 1, \dots, N\}$ let (u_N^k, z_N^k) satisfy* (4.4)*. Then the corresponding interpolants $(\overline{u}_N, \underline{u}_N, \overline{z}_N, \underline{z}_N)$ obtained via* (4.6)*, fulfill the following discrete version of* (1.5) *for all $t \in [0, \mathsf{T}]$:*

$$\text{for all } \tilde{u} \in \mathbf{U}: \quad \overline{\mathcal{E}}_N(t, \overline{u}_N(t), \underline{z}_N(t)) \le \overline{\mathcal{E}}_N(t, \tilde{u}, \underline{z}_N(t)), \tag{4.8a}$$

$$\text{for all } \tilde{z} \in \mathbf{X}: \quad \overline{\mathcal{E}}_N(t, \overline{u}_N(t), \overline{z}_N(t)) \le \overline{\mathcal{E}}_N(t, \overline{u}_N(t), \tilde{z}) + \mathcal{R}(\tilde{z} - \overline{z}_N(t)) + \mathrm{TOL}(N), \tag{4.8b}$$

$$\overline{\mathcal{E}}_N(t, \overline{q}_N(t)) + \mathrm{Diss}_{\mathcal{R}}(\overline{z}_N, [0,t]) \le \overline{\mathcal{E}}_N(0, q_N^0) + \int_0^t \partial_\xi \mathcal{E}_N(\xi, \underline{q}_N(\xi)) \,\mathrm{d}\xi + \mathrm{TOL}(N)N. \tag{4.8c}$$

In particular, there is a constant $C > 0$ such that the following bounds hold true uniformly for all $N \in \mathbb{N}$:

$$\text{for all } t \in [0, \mathsf{T}]: \ \|u_N(t)\|_{\mathbf{U}} \le C, \tag{4.9a}$$

$$\text{for all } t \in [0, \mathsf{T}]: \ \|z_N(t)\|_{\mathbf{X}} \le C, \tag{4.9b}$$

$$\mathcal{R}(\overline{z}_N(\mathsf{T}) - z_N^0) \le C \ \ \& \ \ \|\overline{z}_N\|_{BV(0,\mathsf{T};\mathbf{Z})} \le C, \tag{4.9c}$$

where (u_N, z_N) in (4.9a) *&* (4.9b) *stands for both $(\overline{u}_N, \overline{z}_N)$ and $(\underline{u}_N, \underline{z}_N)$.*

Based on these properties we will establish the proof of the following convergence result in Sec. 4.3:

Theorem 4.2 (Convergence of $(\mathbf{U} \times \mathbf{Z}, \mathcal{E}_N, \mathcal{R})$ to $(\mathbf{U} \times \mathbf{Z}, \mathcal{E}, \mathcal{R})$ in the sense of (1.5))**.** *Let the assumptions of Prop. 4.1 as well as the sharpened minimal angle condition* (4.13) *below for the triangulations $(\mathcal{T}_{h(N)})_N$ hold true. Then there exists a not relabeled subsequence $(\overline{u}_N, \underline{u}_N, \overline{z}_N, \underline{z}_N)_N$ of discrete solutions fulfilling* (4.8) *&* (4.9) *for each $N \in \mathbb{N}$ and a limit pair $(u, z) \in \mathrm{B}(0, \mathsf{T}; \mathbf{U}) \times \big(\mathrm{B}(0, \mathsf{T}; \mathbf{X}) \cap BV(0, \mathsf{T}; \mathbf{Z})\big)$ such that the following convergences hold true:*

$$\text{for all } t \in [0, \mathsf{T}]: \ \overline{u}_N(t) \rightharpoonup u(t) \text{ in } \mathbf{U} \text{ and } \overline{z}_N(t) \overset{*}{\rightharpoonup} z(t), \ \underline{z}_N(t) \overset{*}{\rightharpoonup} \underline{z}(t) \text{ in } \mathbf{X}, \tag{4.10a}$$

$$\text{for all } t \in [0, \mathsf{T}] \backslash \mathrm{J}: \ \underline{u}_N(t) \rightharpoonup u(t) \text{ in } \mathbf{U} \text{ and } z(t) = \underline{z}(t), \tag{4.10b}$$

$$\text{for all } t \in [0, \mathsf{T}]: \ \overline{u}_N(t) \to u(t) \text{ strongly in } \mathbf{U}, \tag{4.10c}$$

where J *in* (4.10b) *denotes the union of the jump times of $z, \underline{z} \in BV(0, \mathsf{T}; \mathbf{Z})$. The functions $z, \underline{z}$ are monotonously decreasing in time. In particular, (u, z) is a semistable energetic solution of the system $(\mathbf{U} \times \mathbf{Z}, \mathcal{E}, \mathcal{R})$ up to an error due to the mesh quality, i.e., the limit pair (u, z) satisfies minimality condition* (1.5a) *and the upper energy dissipation estimate* (1.5c)*, but the semistability inequality only in the following sense*

$$\text{for all } \tilde{z} \in \mathbf{X}: \quad \mathcal{E}(t, u(t), z(t)) \le \mathcal{E}(t, u(t), \tilde{z}) + \mathcal{R}(\tilde{z} - z(t)) + \mathrm{ERR}, \tag{4.11}$$

with $\mathrm{ERR} = \limsup_{N\to\infty} \big(C(\mathcal{T}_{h(N)})^r - 1\big)^{\frac{\kappa_2}{r}} C$; here, $C(\mathcal{T}_{h(N)}) := \max_{T \in \mathcal{T}_{h(N)}} C(\alpha_T, \delta_T)$ is a constant that depends on the smallest angle α_T and the largest angle $\frac{\pi}{2} + \delta_T$, $\delta_T \ge 0$, occurring the triangles $T \in \mathcal{T}_{h(N)}$ of the triangulation $\mathcal{T}_{h(N)}$, cf. Fig. 4.1 below, and C is the uniform bound from (4.9)*. In particular, if the triangulations $(\mathcal{T}_{h(N)})_N$ tend to a right-angled triangulation as $N \to \infty$, i.e., $\delta_N \to 0$, then $\mathrm{ERR} = 0$.*

Finally, also the pair $(u, \hat{z})$ with $\hat{z} := \max\{0, z\}$ provides a semistable energetic solution for the system $(\mathbf{U} \times \mathbf{Z}, \mathcal{E}, \mathcal{R})$ up to an error due to the mesh quality, which, in addition to (4.11) *and* (1.5c) *also has the property $z(t) \in [0, z_0]$ for all $t \in [0, T]$.*

The error term ERR in (4.11) arises from the projection of the nonsmooth unidirectionality constraint into the finite element space $\mathbf{X}_{h(N)}$ based on a given triangulation $\mathcal{T}_{h(N)}$ in the interplay with the gradient regularization $\mathcal{G}$ of general integrability $r \in [1, \infty)$ (see also Remark 4.11 for the case $r = 2$). Because of the nonsmooth constraint, which is induced by the functions z_N themselves one cannot establish the approximation property for the semistability inequality by arguing via the density of smooth functions in the finite-element space intersected with the constraint. Indeed, this statement would be false in general, cf. [HRR17]. Instead, for any smooth approximant of a test function $\tilde{z} \in \mathbf{X}$, here also denoted by $\tilde{z}$, we consider on any triangle $T \in \mathcal{T}_{h(N)}$ in each of its three nodes $\boldsymbol{x}_i$ the values $f^i = f(\boldsymbol{x}_i) = \min\{\tilde{z}^i, z_N^i\}$, $i = 1, 2, 3$. The gradient of the corresponding affine function f on T can then be estimated by trigonometric arguments in dependence of the smallest and largest angle of T as follows:

Lemma 4.3 (Perturbed discrete monotonicity estimate)**.** *Let $T \in \mathcal{T}_{h(N)}$ a triangle given by the nodes $\boldsymbol{x}_i$, $i = 1, 2, 3$, i.e., $T = \mathrm{conv}\{\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3\}$ with its smallest angle $\alpha \in (0, \frac{\pi}{4}]$, its largest angle $\gamma = \frac{\pi}{2} + \delta$ with $\delta \in [0, \frac{\pi}{4})$, and*

its longest edge h, see Fig. 4.1 below. Consider f the affine function on T defined through the values in the nodes as $f^i = f(x^i) = \min\{\tilde{z}^i, z_N^i\}$, $i = 1, 2, 3$. Then

$$\begin{aligned} |\nabla_{\boldsymbol{x}} f(\boldsymbol{x})| &= C(\alpha,\delta)\big(|\nabla z_N| + |\nabla \tilde{z}|\big)\,, \\ \text{where } C(\alpha,\delta) &:= \big(1+\tfrac{\tan\delta}{\cos\alpha(1-\tan\alpha\tan\delta)}\big)\big(1+\sqrt{(\tan\alpha\tan\delta)^2+(\tan\delta)^2}\big)\,. \end{aligned} \tag{4.12}$$

In particular, $C(\alpha,\delta) = 1$ for $\delta = 0$, i.e., for $\gamma = \pi/2$.

The explicit dependence of the constant $C(\alpha_T, \delta_T)$ in (4.12) on the angles of a triangle $T \in \mathcal{T}_{h(N)}$ shows that the mesh structure of a sequence of triangulations $(\mathcal{T}_{h(N)})_N$ needs to be controlled uniformly in terms of the following

$$\text{sharpened minimal angle condition: } \exists\, C_{\mathcal{T}} > 0\, \forall\, N \in \mathbb{N},\, \forall\, T \in \mathcal{T}_{h(N)} :\ C(\alpha_T, \delta_T) \le C_{\mathcal{T}}\,. \tag{4.13}$$

This assumption will allow it to establish suitable compactness properties for sequences constructed by the operation $\min\{\tilde{z}, z_N\}$, but this step will additionally also require an estimate on the gradient from below.

Lemma 4.4 (Lower bound on the gradient)**.** *Let the assumptions and notation of Lemma 4.3 hold true. Let f_N be an affine function on the triangle $T = \operatorname{conv}\{\boldsymbol{y}_1, \boldsymbol{y}_2, \boldsymbol{y}_3\} \in \mathcal{T}_{h(N)}$ with the following properties of its nodal values*

$$f_N(\boldsymbol{y}_1) \ge \omega > 0,\ f_N(\boldsymbol{y}_2) \in [f_N(\boldsymbol{y}_3), f_N(\boldsymbol{y}_1)],\ \text{and}\ f_N(\boldsymbol{y}_3) < 0\,. \tag{4.14}$$

With $\alpha \in (0, \frac{\pi}{4}]$ and $\delta \in [0, \frac{\pi}{4})$ there is a constant $c \in (0,1)$ such that

$$|\nabla_{\boldsymbol{x}} f_N(\boldsymbol{x})| \ge \frac{c\,\omega}{h(N)}\,. \tag{4.15}$$

Proof. We use the notation and geometric settings of Lemma 4.4 and Fig. 4.1. This means that the set of nodes $\{\boldsymbol{y}_1, \boldsymbol{y}_2, \boldsymbol{y}_3\}$ with the properties (4.14) specified in Lemma 4.4 coincides with the set of nodes $\{\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3\}$ from Fig. 4.1. We now note that the norm $|\nabla_{\boldsymbol{x}} f_N(\boldsymbol{x})|$ of the elementwise constant gradient $\nabla_{\boldsymbol{x}} f_N$ is an upper bound for any directional derivative of f_N. Computing the directional derivative of f_N along the edge that connects the vertices $\boldsymbol{y}_1$ and $\boldsymbol{y}_3$ thus proves the estimate. □

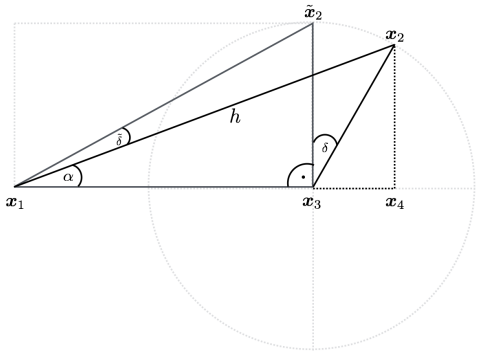


Figure 4.1: Triangle $T = \operatorname{conv}\{\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3\}$ with smallest angle $\alpha \in (0, \frac{\pi}{4}]$ and largest angle $\gamma = \frac{\pi}{2} + \delta$, for $\delta \in [0, \frac{\pi}{4})$. T is linked with the right-angled triangle $\widetilde{T} := \operatorname{conv}\{\boldsymbol{x}_1, \tilde{\boldsymbol{x}}_2, \boldsymbol{x}_3\}$ through the linear map $A : T \to \widetilde{T}$, $\boldsymbol{x} \mapsto \widetilde{\boldsymbol{x}}$ defined by (4.16), resp (4.18).

4.1 Proof of Lemma 4.3

In the following remark we first outline the proof of Lemma 4.3 in the case of a right-angled triangulation.

Remark 4.5 (Right-angled triangulation)**.** For the subsequent considerations we use the mesh size h as an index instead of $h(N)$ or N. In the case of a triangulation consisting of right-angled triangles or tetrahedra that have three mutually orthogonal edges the error term ERR vanishes. For such triangulations we have for two continuous, piecewise affine functions z_h and $\tilde{z}_h$ that the function

$$f_h = \mathcal{I}_h \max\{z_h, \tilde{z}_h\}$$

with the nodal interpolation operator $\mathcal{I}_h$ satisfies almost everywhere the estimate

$$|\nabla_{\boldsymbol{x}} f_h| \le |\nabla_{\boldsymbol{x}} z_h| + |\nabla_{\boldsymbol{x}} \tilde{z}_h|.$$

To see this we note that the norm of the gradient equals on every element the square root of the sum of the directional derivatives along the orthogonal edges $E_1, E_2, \dots, E_d$. The latter are given by difference quotients along these edges. With the auxiliary Lipschitz continuous function $f = \max\{z_h, \tilde{z}_h\}$ we have for an element T with edge $E_i = [\boldsymbol{x}_{i_1}, \mathbf{x}_{i_2}]$ of length h_{E_i} and with tangent vector $\boldsymbol{t}_i$ that

$$h_{E_i}^{-1} |f_h(\boldsymbol{x}_{i_2}) - f_h(\boldsymbol{x}_{i_1})| = h_{E_i}^{-1} |f(\boldsymbol{x}_{i_2}) - f(\boldsymbol{x}_{i_1})| \le \|\nabla_{\boldsymbol{x}} f \cdot \boldsymbol{t}_i\|_{L^\infty(T)}.$$

Noting that $\nabla_{\boldsymbol{x}} f(\boldsymbol{x})$ is for almost every $\boldsymbol{x} \in T$ given by $\nabla z_h|_T$ or $\nabla \tilde{z}_h$ we deduce the estimate. Provided an initial triangulation with this property exists, then red-refinement or yellow-refinement procedures in the cases $d = 2$ and $d = 3$, respectively, preserve this property. Red refinement subdivides every triangle into four subtriangles by connecting the midpoints of the sites; for details on yellow refinement we refer the reader to [BKKŠ09].

We now carry out the proof for a general triangle. For this, let $T \in \mathcal{T}_{h(N)}$ a triangle given by the nodes $\boldsymbol{x}_i$, $i = 1, 2, 3$, i.e., $T = \operatorname{conv}\{\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3\}$ with its smallest angle $\alpha \in (0, \frac{\pi}{4}]$, its largest angle $\gamma = \frac{\pi}{2} + \delta$ with $\delta \in [0, \frac{\pi}{4})$, and its longest edge h, see Fig. 4.1 below. Denote by $\widetilde{T} := \operatorname{conv}\{\boldsymbol{x}_1, \tilde{\boldsymbol{x}}_2, \boldsymbol{x}_3\}$ the right-angled triangle that shares with T the edge $\overline{\boldsymbol{x}_1\boldsymbol{x}_3}$ and for which the edges $\overline{\boldsymbol{x}_3\boldsymbol{x}_2}$ and $\overline{\boldsymbol{x}_3\tilde{\boldsymbol{x}}_2}$ have the same length. We denote by $\boldsymbol{x}$ the coordinates of points of T and, correspondingly, by $\tilde{\boldsymbol{x}}$ the coordinates of points of $\widetilde{T}$. We introduce the linear map $A : T \to \widetilde{T}$, $\boldsymbol{x} \mapsto \widetilde{\boldsymbol{x}}$ through the following relations

$$A : T \to \widetilde{T},\ A\overline{\boldsymbol{x}_1\boldsymbol{x}_3} = \overline{\boldsymbol{x}_1\boldsymbol{x}_3},\ A\overline{\boldsymbol{x}_1\boldsymbol{x}_2} = \overline{\boldsymbol{x}_1\tilde{\boldsymbol{x}}_2} \tag{4.16}$$

For any function $\tilde{f} : \widetilde{T} \to \mathbb{R}$ and $f = \tilde{f} \circ A : T \to \mathbb{R}$ we thus have the following transformation relation for their gradients

$$\nabla_{\boldsymbol{x}} f(\boldsymbol{x}) = A^\top \nabla_{\tilde{\boldsymbol{x}}} \tilde{f}(\tilde{\boldsymbol{x}})\,. \tag{4.17}$$

Relation (4.17) will be used to deduce the gradient estimate (4.12) for any affine function f on T defined through the values in the nodes as $f^i = f(\boldsymbol{x}_i) = \min\{\tilde{z}^i, z_N^i\}$, $i = 1, 2, 3$. For this, we proceed along the following steps:

Step 1: Express the map A from (4.16) in terms of the angles $\alpha \in (0, \frac{\pi}{4}]$ and $\delta \in [0, \frac{\pi}{4}]$, i.e., show that

$$A := \begin{pmatrix} 1 & -\tan\delta \\ 0 & \frac{1}{\cos\delta} \end{pmatrix} \tag{4.18}$$

Step 2: Calculate $\nabla_{\tilde{\boldsymbol{x}}} \tilde{f}(\tilde{\boldsymbol{x}})$ on the right-angled triangle $\widetilde{T}$ of the form

$$\nabla_{\tilde{\boldsymbol{x}}} \tilde{f}(\tilde{\boldsymbol{x}}) = B\big(\frac{\tilde{f}^3 - \tilde{f}^1}{h\cos\alpha}, \frac{\tilde{f}^2 - \tilde{f}^3}{h\sin\alpha}\big)^\top,\ \text{where}\ B = \begin{pmatrix} \frac{1}{1-\tan\alpha\tan\delta} & 0 \\ 0 & \frac{1}{\cos\delta} \end{pmatrix}. \tag{4.19}$$

Step 3: Verify for $\tilde{f}(\boldsymbol{x}_i) = \tilde{f}^i = \min\{z_N^i, \tilde{z}^i\}$ the estimate

$$\Big|\big(\tfrac{\tilde{f}^3 - \tilde{f}^1}{h\cos\alpha}, \tfrac{\tilde{f}^2 - \tilde{f}^3}{h\sin\alpha}\big)^\top\Big| \le \big(1 + |B^{-1}A^{-\top} - \mathrm{Id}|\big)\big(|\nabla_{\boldsymbol{x}} z_N| + |\nabla_{\boldsymbol{x}} \tilde{z}|\big)\,. \tag{4.20}$$

Step 4: Conclude (4.12) as a combination of (4.17)–(4.20).

To Step 1: In accordance with Fig. 4.1 we set

$$\begin{aligned} \boldsymbol{x}_1 &= (0,0)^\top, \qquad \boldsymbol{x}_2 = (h\cos\alpha, h\sin\alpha)^\top, \qquad \boldsymbol{x}_4 = (h\cos\alpha, 0)^\top, \\ \boldsymbol{x}_3 &= \boldsymbol{x}_4 - (h\sin\alpha\tan\delta, 0)^\top = (h\cos\alpha(1 - \tan\alpha\tan\delta), 0)^\top, \\ \tilde{\boldsymbol{x}}_2 &= (h\cos\alpha(1 - \tan\alpha\tan\delta), \tan(\alpha + \tilde{\delta})h\cos\alpha(1 - \tan\alpha\tan\delta))^\top, \end{aligned}$$

In this way the mapping relations (4.16) lead to the following entries for the matrix A

$$A_{11} = 1, \quad A_{12} = -\tan\delta, \quad A_{21} = 0, \quad A_{22} = \tan(\alpha + \tilde{\delta})\cot\alpha(1 - \tan\alpha\tan\delta)$$

and now the angle $\tilde{\delta}$ appearing in A_{22} has to be expressed through the angles α and δ. For this, we deduce that

$$\tan(\alpha + \tilde{\delta}) \quad = \quad \frac{|\tilde{\boldsymbol{x}}_2 - \boldsymbol{x}_3|}{|\boldsymbol{x}_3|} = \frac{\tan\alpha\sqrt{1 + (\tan\delta)^2}}{|1 - \tan\alpha\tan\delta|}\,,$$

which yields that

$$A_{22} = \frac{(1-\tan\alpha\tan\delta)\sqrt{1+(\tan\delta)^2}}{|1-\tan\alpha\tan\delta|} = \frac{1}{\cos\delta}\,,$$

where the last equality holds true because of $(1-\tan\alpha\tan\delta) > 0$ thanks to $\alpha \in (0, \frac{\pi}{4}]$ and $\delta \in [0, \frac{\pi}{4})$. This proves (4.18) and thus concludes Step 1.

To Step 2: Let $\tilde f$ be an affine function on the triangle $\widetilde T = \mathrm{conv}\{\boldsymbol{x}_1, \tilde{\boldsymbol{x}}_2, \boldsymbol{x}_3\}$. Then $\nabla_{\tilde{\boldsymbol{x}}}\tilde f(\tilde{\boldsymbol{x}}) = (\tilde m_1, \tilde m_2)^\top$. We set $\tilde f^i := \tilde f(\tilde{\boldsymbol{x}}_i)$ and then the two constants $\tilde m_1, \tilde m_2$ can be expressed as follows

$$\tilde m_1 = \frac{\tilde f^3 - \tilde f^1}{\boldsymbol{x}_{31} - \boldsymbol{x}_{11}} = \frac{\tilde f^3 - \tilde f^1}{h\cos\alpha(1-\tan\alpha\tan\delta)} = \frac{1}{(1-\tan\alpha\tan\delta)}\cdot\frac{\tilde f^3 - \tilde f^1}{h\cos\alpha}\,, \tag{4.21a}$$

$$\tilde m_2 = \frac{\tilde f^2 - \tilde f^1 - \tilde m_1(\tilde{\boldsymbol{x}}_{2_1} - \boldsymbol{x}_{1_1})}{\tilde{\boldsymbol{x}}_{22} - \boldsymbol{x}_{12}} = \frac{\tilde f^2 - \tilde f^3}{\tan(\alpha+\tilde\delta)h\cos\alpha(1-\tan\alpha\tan\delta)} = \cos\delta\cdot\frac{\tilde f^2 - \tilde f^3}{h\sin\alpha}\,. \tag{4.21b}$$

This proves (4.19).

To Step 3: Let now $\tilde f(\boldsymbol{x}_i) = \tilde f^i = \min\{z_N^i, \tilde z^i\}$. Using the short-hand notation $\tilde{\boldsymbol{m}}_f = \big(\frac{\tilde f^3-\tilde f^1}{h\cos\alpha}, \frac{\tilde f^2-\tilde f^3}{h\sin\alpha}\big)^\top$ and $\tilde{\boldsymbol{m}}_{z_N}$, $\tilde{\boldsymbol{m}}_{\tilde z}$ for the corresponding expressions for the affine functions $z_N, \tilde z$ we observe that

$$|\tilde{\boldsymbol{m}}_f| = \left|\Big(\tfrac{\min\{\tilde z^3, z_N^3\}-\min\{\tilde z^1, z_N^1\}}{h\cos\alpha}, \tfrac{\min\{\tilde z^2, z_N^2\}-\min\{\tilde z^3, z_N^3\}}{h\sin\alpha}\Big)^\top\right| \le |\tilde{\boldsymbol{m}}_{z_N}| + |\tilde{\boldsymbol{m}}_{\tilde z}|\,. \tag{4.22}$$

To see this, exemplarily consider the case that $\min\{\tilde z^3, z_N^3\} - \min\{\tilde z^1, z_N^1\} = \tilde z^3 - z_N^1$, i.e., $\tilde z^3 \le z_N^3$ and $\tilde z^1 \ge z_N^1$ We have to distinguish the two cases $\tilde z^3 - z_N^1 \ge 0$ and $\tilde z^3 - z_N^1 < 0$. On the one hand, if $\tilde z^3 - z_N^1 \ge 0$, then $0 \le \min\{\tilde z^3, z_N^3\} - \min\{\tilde z^1, z_N^1\} = \tilde z^3 - z_N^1 \le z_N^3 - z_N^1$. On the other hand, if $\tilde z^3 - z_N^1 < 0$, then $0 < \min\{\tilde z^1, z_N^1\} - \min\{\tilde z^3, z_N^3\} = z_N^1 - \tilde z^3 \le \tilde z^1 - \tilde z^3$. Combining the two results we conclude that $|\min\{\tilde z^3, z_N^3\} - \min\{\tilde z^1, z_N^1\}| = |\tilde z^3 - z_N^1| \le |z_N^3 - z_N^1| + |\tilde z^1 - \tilde z^3|$. Analogous estimates for the remaining cases can be obtained with similar arguments.

From (4.22), estimate (4.20) is then concluded as follows

$$\begin{aligned}
|\tilde{\boldsymbol{m}}_f| &\le |\tilde{\boldsymbol{m}}_{z_N}| + |\tilde{\boldsymbol{m}}_{\tilde z}| = |B^{-1}A^{-\top}A^\top B\,\tilde{\boldsymbol{m}}_{z_N}| + |B^{-1}A^{-\top}A^\top B\,\tilde{\boldsymbol{m}}_{\tilde z}| \\
&\le |A^\top B\,\tilde{\boldsymbol{m}}_{z_N} + (B^{-1}A^{-\top} - \mathrm{Id})A^\top B\,\tilde{\boldsymbol{m}}_{z_N}| + |A^\top B\,\tilde{\boldsymbol{m}}_{\tilde z} + (B^{-1}A^{-\top} - \mathrm{Id})A^\top B\,\tilde{\boldsymbol{m}}_{\tilde z}| \\
&\le \big(1 + |B^{-1}A^{-\top} - \mathrm{Id}|\big)\big(|A^\top B\,\tilde{\boldsymbol{m}}_{z_N}| + |A^\top B\,\tilde{\boldsymbol{m}}_{\tilde z}|\big) \\
&= \big(1 + |B^{-1}A^{-\top} - \mathrm{Id}|\big)\big(|\nabla_{\boldsymbol{x}} z_N| + |\nabla_{\boldsymbol{x}}\tilde z|\big)\,.
\end{aligned}$$

To Step 4: From (4.17) and (4.19) we gather that $\nabla_{\boldsymbol{x}} f(\boldsymbol{x}) = A^\top B\big(\frac{\tilde f^3-\tilde f^1}{h\cos\alpha}, \frac{\tilde f^2-\tilde f^3}{h\sin\alpha}\big)^\top$ and further estimate

$$\begin{aligned}
|\nabla_{\boldsymbol{x}} f(\boldsymbol{x})| &= |A^\top B\,\tilde{\boldsymbol{m}}_f| \le |\mathrm{Id}\,\tilde{\boldsymbol{m}}_f| + |(A^\top B - \mathrm{Id})\tilde{\boldsymbol{m}}_f| \le \big(1 + |A^\top B - \mathrm{Id}|\big)|\tilde{\boldsymbol{m}}_f| \\
&\le \big(1 + |A^\top B - \mathrm{Id}|\big)\big(1 + |B^{-1}A^{-\top} - \mathrm{Id}|\big)\big(|\nabla_{\boldsymbol{x}} z_N| + |\nabla_{\boldsymbol{x}}\tilde z|\big)\,,
\end{aligned}$$

where the last estimate stems from (4.20). Calculating that $|A^\top B - \mathrm{Id}| = \frac{\tan\delta}{\cos\alpha(1-\tan\alpha\tan\delta)}$ as well as $|B^{-1}A^{-\top} - \mathrm{Id}| = \sqrt{(\tan\alpha\tan\delta)^2 + (\tan\delta)^2}$ ultimately proves the desired estimate (4.12) with

$$C(\alpha,\delta) := \big(1 + |A^\top B - \mathrm{Id}|\big)\big(1 + |B^{-1}A^{-\top} - \mathrm{Id}|\big) = \big(1 + \tfrac{\tan\delta}{\cos\alpha(1-\tan\alpha\tan\delta)}\big)\big(1 + \sqrt{(\tan\alpha\tan\delta)^2 + (\tan\delta)^2}\big)\,.$$

This concludes the proof of Lemma 4.3. ■

4.2 Proof of Prop. 4.1

Proof of properties (4.8)**:** Taking into account the definition (4.6) of the interpolants $(\overline u_N, \underline u_N, \overline z_N, \underline z_N)$ we see that minimality properties (4.4) can be directly translated into (4.8a) & (4.8b). To find the discrete upper energy-dissipation estimate (4.8c) we test the minimality of u_N^k in (4.4a) by u_N^{k-1} and the minimality of z_N^k in (4.4b) by z_N^{k-1}. This results in

$$\begin{aligned}
\mathcal{E}_N(t_N^k, u_N^k, z_N^{k-1}) &\le \mathcal{E}_N(t_N^k, u_N^{k-1}, z_N^{k-1}) \\
\mathcal{E}_N(t_N^k, u_N^k, z_N^k) + \mathcal{R}(z_N^k - z_N^{k-1}) &\le \mathcal{E}_N(t_N^k, u_N^k, z_N^{k-1}) + \mathrm{TOL}(N)\,.
\end{aligned}$$

Let now $t \in (0, t_N^n]$ for some $n \leq N$. Adding the above two inequalities, adding and subtracting $\mathcal{E}_N(t_N^{k-1}, u_N^{k-1}, z_N^{k-1})$, and summing over $k \in \{1, \dots, n\}$ we find

$$\begin{aligned}
&\mathcal{E}_N(t_N^n, u_N^n, z_N^n) + \mathcal{R}(z_N^n - z_N^0) \\
&\leq \mathcal{E}_N(t_N^0, u_N^0, z_N^0) + \sum_{k=1}^{n} \mathcal{E}_N(t_N^k, u_N^{k-1}, z_N^{k-1}) - \mathcal{E}_N(t_N^{k-1}, u_N^{k-1}, z_N^{k-1}) + n\mathrm{TOL}(N) \\
&= \mathcal{E}_N(t_N^0, u_N^0, z_N^0) + \sum_{k=1}^{n} \int_{t_N^{k-1}}^{t_N^k} \partial_\xi \mathcal{E}_N(\xi, u_N^{k-1}, z_N^{k-1}) \,\mathrm{d}\xi + n\mathrm{TOL}(N)\,,
\end{aligned} \tag{4.23}$$

which yields (4.8c) for all $t \in (0, t_N^n]$ and integers $n \leq N$.

Proof of estimates (4.9)**:** Observe that there are constants $c_0, c_1 > 0$, such that for all $(t, u, z) \in [0, \mathsf{T}] \times \mathbf{U} \times \mathbf{Z}$ with $\mathcal{E}_N(t, u, z) < \infty$ it holds $|\partial_t \mathcal{E}_N(t, u, z)| \leq c_1(c_0 + \mathcal{E}_N(t, u, z))$. This entitles us to apply a Gronwall estimate under the time-integral in (4.23). Classical arguments for energy-dissipation inequalities in the rate-independent setting, cf., e.g., [MR15, Prop. 2.1.4], result in the estimates

$$c_0 + \overline{\mathcal{E}}_N(t_N^k, u_N^k, z_N^k) \quad \leq \quad (c_0 + \overline{\mathcal{E}}_N(0, u_N^0, z_N^0))\exp(c_1 T) \leq C\,, \tag{4.24a}$$

$$\mathcal{R}(z_N^k - z_N^0) \quad \leq \quad (c_0 + \overline{\mathcal{E}}_N(0, u_N^0, z_N^0))\exp(c_1 T) \leq C\,, \tag{4.24b}$$

where the uniform boundedness by $C > 0$ is due to (4.7) and Assumption 2.1. The estimate (4.9a) is then standardly obtained from the bound (4.24a), exploiting that $f(0) \geq a > 0$ and $\mu > 0$ by Assumption 2.1, as well as Korn's and Young's inequality. The estimate (4.9b) follows from the uniform boundedness of the damage regularization $\mathcal{G}(z_N^k) \leq C$, ensured by (4.24a), whereas the first estimate in (4.9c) is due to (4.24b) and the second is a direct consequence taking into account the form of $\mathcal{R}$, see (1.1a). This concludes the proof of Prop. 4.1. ■

4.3 Proof of Thm. 4.2

Proof of convergences (4.10a)**–**(4.10b)**:** To obtain the convergence result for the damage variables in (4.10a) we make use of the uniform bound in (4.9b). This, together with the fact that $\mathcal{R} : \mathbf{Z} \times \mathbf{Z} \to [0, \infty]$ is a weakly sequentially lower semicontinuous dissipation distance, allows us to apply a generalized version of Helly's selection principle, see e.g. [MR15, Thm. 2.1.24], and hence to find a (not relabeled) subsequence as well as limit functions $z, \underline{z} \in \mathrm{BV}([0, \mathsf{T}], \mathbf{Z})$, such that for all $t \in [0, \mathsf{T}]$:

$$\mathcal{R}(\overline{z}_N(t) - z_N^0) \to \mathcal{R}(z(t) - z_0)\,,\ \overline{z}_N(t) \rightharpoonup z(t)\ \&\ \underline{z}_N(t) \rightharpoonup \underline{z}(t)\ \text{ in } \mathbf{X}\,. \tag{4.25}$$

For some $t \in [0, \mathsf{T}]$ fixed, select a further subsequence such that $\overline{u}_N(t) \rightharpoonup u(t)$ in $\mathbf{U}$. Exploiting the minimality (4.8a) of $\overline{u}_N(t)$ for $\overline{\mathcal{E}}_N(t, \cdot, \underline{z}_N(t))$ as well as cancellations and the weak sequential lower semicontinuity properties, we find

$$0 \leq \limsup_{N\to\infty} \big(\overline{\mathcal{E}}_N(t, \tilde{u}, \underline{z}_N(t)) - \overline{\mathcal{E}}_N(t, \overline{u}_N(t), \underline{z}_N(t))\big) \leq \mathcal{E}(t, \tilde{u}, \underline{z}(t)) - \mathcal{E}(t, u(t), \underline{z}(t))\big) \tag{4.26}$$

for all $\tilde{u} \in \mathbf{U}$. In other words, $u(t)$ is the unique minimizer of the strictly convex functional $\mathcal{E}(t, \cdot, \underline{z}(t)) : \mathbf{U} \to \mathbb{R} \cup \{\infty\}$. Thus, the above selection of a subsequence of $(\overline{u}_N(t))_N$ was unnecessary. This observation holds for all $t \in [0, \mathsf{T}]$. Moreover, since $\underline{z}$ and and the given data are measurable with respect to time, we also have that $u : [0, \mathsf{T}] \to \mathbf{U}$ is measurable. This concludes the proof of statement (4.10a).

To find convergence (4.10b), let $\mathrm{J} \subset [0, \mathsf{T}]$ denote the union of the jump times of $z, \underline{z} \in \mathrm{BV}(0, \mathsf{T}; \mathbf{Z})$. By the properties of BV-functions, J is at most countable. Consider $t \in [0, \mathsf{T}]\backslash\mathrm{J}$ and a sequence $t_N^{l_N} \to t$ as $N \to \infty$ with $t_N^{l_N} \in \Pi_N$ for all $N \in \mathbb{N}$. With $z_N^{l_N}$ obtained by (1.4b), it holds that $z_N^{l_N} = \underline{z}_N(t_{l_N}^1) = \overline{z}_k(t_{l_N}^2)$ for all $t_N^{l_N - 1} \leq t_{l_N}^2 \leq t_N^{l_N} \leq t_{l_N}^1 \leq t_N^{l_N+1}$. For $t_{l_N}^1 \to t$ and $t_{l_N}^2 \to t$ as $N \to \infty$ we thus conclude $z(t) = \underline{z}(t)$ for all $t \in [0, \mathsf{T}]\backslash\mathrm{J}$. Since $u_N^{l_N}$ is the unique minimizer of $\mathcal{E}_N(t_N^{l_N}, \cdot\,, z_N^{l_N - 1})$, we find the convergence result for $(\underline{u}_N(t))_N$ in (4.10b) with similar arguments.

Limit passage in the discrete notion of solution (4.8)**:** Thanks to convergences (4.10) and (4.7) the **limit passage in the upper energy-dissipation estimate** (4.8c) can be carried out by means of lower semicontinuity arguments on the left-hand side of (4.8c). To pass to the limit on the right-hand side of (4.8c) we make use of the strong convergence $(u_N^0, z_N^0) \to (u_0, z_0)$ in $\mathbf{U} \times \mathbf{X}$ for the energy at initial time and of the fact that $\mathrm{TOL}(N)N \to 0$ as $N \to \infty$. Convergence of the power of the external loadings follows via weak-strong convergence arguments from (4.10) and (4.7).

In order to pass to the **limit in minimality condition** (4.8a) we have to argue via a suitable recovery sequence. More precisely, for any $\tilde{u} \in \mathbf{U}$ we construct a recovery sequence $(\tilde{u}_N)_N$ such that $u_N \in \mathbf{U}_{h(N)}$ for each $N \in \mathbb{N}$. For this, thanks to the density of smooth functions in $\mathbf{U}$, we first find a sequence $(\hat{u}_N)_N \in C^\infty(\Omega;\mathbb{R}^d) \cap \mathbf{U}$ with $\|\tilde{u}-\hat{u}_N\|_{\mathbf{U}} \to 0$. We then set

$$\tilde{u}_N := \mathcal{I}_{h(N)} \hat{u}_N \,, \tag{4.27}$$

which thus ensures that $\tilde{u}_N \to \tilde{u}$ *strongly* in $\mathbf{U}$, cf. [Cia02, Thm. 3.2.3].

Let $\mathcal{W}$ and $\mathcal{L}_{\mathrm{Neu}}$ as in (1.3b) and set now $W(e,z) := f(z)\big(\lambda|\operatorname{tr} e(u)|^2 + 2\mu|e(u)|^2\big)$. Observing that, for each $N \in \mathbb{N}$, the term $\mathcal{G}(z_N(t))$ cancels out in (4.8a), we find

$$\begin{aligned}
&\mathcal{W}(u(t),\underline{z}(t)) + \mathcal{L}_{\mathrm{Neu}}(t,u(t)) \\
&= \int_\Omega f(\underline{z}(t))\big(\lambda|\operatorname{tr} e(u(t)+u_{\mathrm{D}}(t))|^2 + 2\mu|e(u(t)+u_{\mathrm{D}}(t))|^2\big)\,\mathrm{d}x + \int_{\Gamma_{\mathrm{Neu}}} l_{\mathrm{Neu}}(t)\cdot(u(t)+u_{\mathrm{D}}(t))\,\mathrm{d}\mathcal{H}^{d-1} \\
&\le \liminf_{N\to\infty} \Big(\mathcal{W}(\overline{u}_N(t),\underline{z}_N(t)) + \overline{\mathcal{L}}^N_{\mathrm{Neu}}(t,\overline{u}_N(t))\Big) \\
&= \liminf_{N\to\infty} \Big(\int_\Omega W(e(\overline{u}_N(t)+\overline{u_{\mathrm{D}}}_N(t)),\underline{z}_N(t))\,\mathrm{d}x + \int_{\Gamma_{\mathrm{Neu}}} \overline{l_{\mathrm{Neu}}}(t)\cdot(\overline{u}_N(t)+\overline{u_{\mathrm{D}}}_N(t))\,\mathrm{d}\mathcal{H}^{d-1}\Big) \\
&\le \lim_{N\to\infty} \Big(\mathcal{W}(\tilde{u}_N,\underline{z}_N(t)) + \overline{\mathcal{L}}^N_{\mathrm{Neu}}(t,\tilde{u}_N)\Big) \\
&= \lim_{N\to\infty} \Big(\int_\Omega W(e(\tilde{u}_N+\overline{u_{\mathrm{D}}}_N(t)),\underline{z}_N(t))\,\mathrm{d}x + \int_{\Gamma_{\mathrm{Neu}}} \overline{l_{\mathrm{Neu}}}(t)\cdot(\tilde{u}_N+\overline{u_{\mathrm{D}}}_N(t))\,\mathrm{d}\mathcal{H}^{d-1}\Big) \\
&= \int_\Omega f(\underline{z}(t))\big(\lambda|\operatorname{tr} e(\tilde{u}+u_{\mathrm{D}}(t))|^2 + 2\mu|e(\tilde{u}+u_{\mathrm{D}}(t))|^2\big)\,\mathrm{d}x + \int_{\Gamma_{\mathrm{Neu}}} l_{\mathrm{Neu}}(t)\cdot(\tilde{u}+u_{\mathrm{D}}(t))\,\mathrm{d}\mathcal{H}^{d-1} \\
&= \mathcal{W}(\tilde{u},\underline{z}(t)) + \mathcal{L}_{\mathrm{Neu}}(t,\tilde{u})\,,
\end{aligned}$$

for all $t \in [0,T]$. Taking into account that $\underline{z}(t) = z(t)$ for all $t \in [0,T]\backslash J$ this amounts to the validity of (1.5a) for a.e. $t \in (0,T)$, also upon adding $\mathcal{G}(z(t))$ on both sides of the above inequality.

In order to pass to the **limit in the semistability inequality** (4.8b) we construct for each $\tilde{z} \in \mathbf{X}$ a *mutual recovery sequence* $(\tilde{z}_N)_N$ such that $\tilde{z}_N \in \mathbf{X}_{h(N)}$ for each $n \in \mathbb{N}$ and such that

$$\begin{aligned}
&\limsup_{N\to\infty} \big(\overline{\mathcal{E}}_N(t,\overline{u}_N(t),\tilde{z}_N) + \mathcal{R}(\tilde{z}_N - \overline{z}_N(t)) - \overline{\mathcal{E}}_N(t,\overline{u}_N(t),\overline{z}_N(t))\big) \\
&\le \mathcal{E}(t,u(t),\tilde{z}) + \mathcal{R}(\tilde{z}-z(t)) - \mathcal{E}(t,u(t),z(t))
\end{aligned} \tag{4.28}$$

Clearly, since the term on the left-hand side of (4.28) is nonnegative for all $N \in \mathbb{N}$ a successful limit passage implies the semistability (1.5b) of the limit.

Consider $\tilde{z} \in \mathbf{X}$ such that $0 \le \tilde{z} \le 1$ a.e. in Ω. By the density of smooth functions in $\mathbf{X}$, we find a sequence $(z_N^\circ)_N \subset C^\infty(\Omega) \cap \mathbf{X}$ such that

$$\begin{aligned}
&\|\nabla z_N^\circ\|_{\mathbf{X}} \le \|\tilde{z}\|_{\mathbf{X}} + \omega(N) \text{ with } \omega(N) \to 0 \text{ as } N \to \infty \text{ and} \\
&\|\tilde{z} - z_N^\circ\|_{L^r(\Omega)} \le \nu_N/2,
\end{aligned} \tag{4.29}$$

with the sequence $\nu_N \to 0$ specified in (4.32) below. To obtain an element of the space $\mathbf{X}_{h(N)}$ we set

$$\hat{z}_N := \mathcal{I}_{h(N)} z_N^\circ \in \mathbf{X}_{h(N)} \tag{4.30}$$

with $\mathcal{J}_{h(N)}$ a quasi-interpolation operator, cf. [Clé75]. This gives

$$\hat{z}_N \to \tilde{z} \quad \text{in } \mathbf{X} \tag{4.31}$$

strongly if $\mathbf{X} = W^{1,r}(\Omega)$, $r > 1$ and intermediately if $\mathbf{X} = \mathrm{BV}(\Omega)$, cf. [Bar12, Thm. 3.1] and [Bar15, Lemma 10.1].

If $\mathcal{R}(\tilde{z}-z(t)) = \infty$ then it is sufficient to choose $\tilde{z}_N := \hat{z}_N$ from (4.30) and (4.28) follows.

Instead, if $\mathcal{R}(\tilde{z}-z(t)) < \infty$, we use $\hat{z}_N$ from (4.30) obtained by mollification and interpolation, and define $\tilde{z}_N \in \mathbf{X}_{h(N)}$ as follows

$$\tilde{z}_N(\boldsymbol{x}) := \min\{\hat{z}_N(\boldsymbol{x}) - \nu_N, \overline{z}_N(\boldsymbol{x})\} \text{ with } \nu_N := \max\{h(N)^{1/2}, \|\overline{z}_N(t) - z(t)\|^{1/2}_{L^2(\Omega)}\} \tag{4.32}$$

for all $\boldsymbol{x} \in \mathcal{N}_{h(N)}$, i.e., $\tilde{z}_N = \mathcal{I}_{h(N)} \min\{\hat{z}_N - \nu_N, \overline{z}_N\}$. This construction ensures $\mathcal{R}(\tilde{z}_N - \overline{z}_N(t)) < \infty$.

In the context of energetic solutions for system $(\mathbf{U} \times \mathbf{Z}, \mathcal{E}, \mathcal{R})$ it was shown that a construction alike (4.32) satisfies the analogon of the mutual recovery condition (4.28), in [Tho13, Sec. 2.2] for $\hat{z}_N = \tilde{z} \in BV(\Omega)$ and in [TM10, Sec. 3.2.5] for $\hat{z}_N = \tilde{z} \in \mathbf{X} = W^{1,r}(\Omega)$ with $r \in (1, \infty)$. Moreover, in a thermo-viscoelastic setting, [RT15, Sec. 4.2] handles the limit passage in the semistability inequality from an adhesive delamination model with a regularization of Modica-Mortola-type to a delamination model, where the delamination variable is the characteristic function of a set of finite perimeter and thus only accounts for the sound and the broken state of the glue. Many of the arguments developed in the context of [TM10, Tho13, RT15] can be used also in the present situation. In particular, as in [TM10, Sec. 3.2.5] we introduce the sets

$$A_N := [\hat{z}_N - \nu_N \leq \overline{z}_N(t)] \text{ and } B_N := [\overline{z}_N(t) < \hat{z}_N - \nu_N]\,, \tag{4.33}$$

where we used the short-hand $[\ldots] := \{x \in \Omega \text{ s.th. } \ldots\}$. Based on this, we decompose the underlying triangulation $\mathcal{T}_{h(N)}$ as follows:

$$\mathcal{T}_{h(N)} = \mathcal{A}_{h(N)} \cup \mathcal{B}_{h(N)} \cup \mathcal{C}_{h(N)}, \tag{4.34}$$

where $\mathcal{A}_{h(N)} := \{T \in \mathcal{T}_{h(N)},\, T \subset A_N\}$, $\mathcal{B}_{h(N)} := \{T \in \mathcal{T}_{h(N)},\, T \subset B_N\}$, and $\mathcal{C}_{h(N)} := \mathcal{T}_{h(N)} \backslash (\mathcal{A}_{h(N)} \cup \mathcal{B}_{h(N)})$ contains those triangles which intersect both with A_N and with B_N. With a slight abuse of notation we will also write $\mathcal{A}_{h(N)} = \cup_{T \in \mathcal{A}_{h(N)}} T$, and analogously for $\mathcal{B}_{h(N)}$ and $\mathcal{C}_{h(N)}$.

Along the lines of [TM10, Tho13, RT15], to verify condition (4.28), we will now argue with the sequence $(\tilde{z}_N)_N$ from (4.32) by proceeding in the following way:

Step 1: Show that $\mathcal{L}^2(B_N) \to 0$ as well as $\mathcal{L}^2(\mathcal{B}_{h(N)}) \to 0$. The first is equivalent to $\mathcal{L}^2(\Omega \backslash A_N) \to 0$.

Step 2: Show that $\mathcal{L}^2(\mathcal{C}_{h(N)}) \to 0$, hence $\mathcal{L}^2(\mathcal{C}_{h(N)} \cup \mathcal{B}_{h(N)}) \to 0$, i.e., $\mathcal{L}^2(\Omega \backslash \mathcal{A}_{h(N)}) \to 0$. For this we will consider differences $f_N := \overline{z}_N - \hat{z}_N$, cf. (4.30), and argue by contradiction using Lemma 4.4.

Step 3: Show that $\tilde{z}_N \rightharpoonup \tilde{z}$ in $\mathbf{X}$. This requires an estimate on the gradients of $\tilde{z}_N$ in terms of the functions involved in the interpolation along the triangulation, which can be established thanks to the sharpened minimal angle condition (4.13). Due to $\tilde{z}_N \leq \overline{z}_N(t)$ a.e. in Ω by construction and convergence (4.10a), the here obtained weak convergence result also implies that $\mathcal{R}(\tilde{z}_N - \overline{z}_N(t)) \to \mathcal{R}(\tilde{z} - z(t))$.

Step 4: Conclude the mutual recovery condition (4.28). Since only weak convergence of $(\tilde{z}_N)_N$ to $\tilde{z}$ is available, our proof will explicitely make use of cancellation arguments for the differences $|\nabla \tilde{z}_N|^r - |\nabla \overline{z}_N(t)|^r$ on $\mathcal{B}_{h(N)} \cup \mathcal{C}_{h(N)}$ from (4.34).

We now carry out in detail the four steps outlined above.

To Step 1: Show that $\mathcal{L}^2(B_N) \to 0$ and $\mathcal{L}^2(\mathcal{B}_{h(N)}) \to 0$: In view of (4.32) we find for B_N from (4.33) that

$$\begin{aligned}
B_N &= [\nu_N < \hat{z}_N - \overline{z}_N(t)] = [\nu_N - \hat{z}_N + \tilde{z} < \tilde{z} - \overline{z}_N(t)] \\
&\subseteq \big([-\nu_N/2 \leq \tilde{z} - \hat{z}_N] \cap [\nu_N - \nu_N/2 < \tilde{z} - \overline{z}_N(t)]\big) \\
&\qquad\qquad \cup \big([-\nu_N/2 > \tilde{z} - \hat{z}_N] \cap [\nu_N - \hat{z}_N + \tilde{z} < \tilde{z} - \overline{z}_N(t)]\big) \\
&\subseteq [\nu_N/2 < \tilde{z} - \overline{z}_N(t)] \cup [\nu_N/2 \leq |\tilde{z} - \hat{z}_N|] \\
&\subseteq [\nu_N/2 < z(t) - \overline{z}_N(t)] \cup [\nu_N/2 \leq |\tilde{z} - \hat{z}_N|]\,,
\end{aligned}$$

where the last set inclusion follows from the relation $\tilde{z} \leq z(t)$ a.e. in Ω. Thanks to this chain of inclusions it can be shown as in [TM10, Sec. 3.2.5] that

$$\mathcal{L}^2(B_N) = \mathcal{L}^2(\Omega \backslash A_N) \leq \mathcal{L}^2([\nu_N/2 < |z(t) - \overline{z}_N(t)|]) + \mathcal{L}^2([\nu_N/2 \leq |\tilde{z} - \hat{z}_N|]) \to 0 \tag{4.35}$$

by exploiting the convergences (4.10a) and (4.29) in a Markov-type inequality.

In view of the definition of the sets $\mathcal{B}_{h(N)}$ and B_N, cf. (4.34) and (4.35), we also infer that

$$\mathcal{L}^2(\mathcal{B}_{h(N)}) \leq \mathcal{L}^2(B_N) \to 0 \quad \text{as } h(N) \to 0 \text{ for } N \to \infty\,. \tag{4.36}$$

To Step 2: Show that

$$\mathcal{L}^2(\mathcal{C}_{h(N)}) \to 0\,. \tag{4.37}$$

Consider the differences $f := z - \tilde{z} \geq 0$ a.e. in Ω and $f_N := \overline{z}_N - \hat{z}_N$, cf. (4.30). The functions f_N are piecewise affine with

$$\|\nabla f_N\|_{L^r(\Omega)} \leq \|\nabla \overline{z}_N\|_{L^r(\Omega)} + \|\nabla \hat{z}_N\|_{L^r(\Omega)} \leq C\,. \tag{4.38}$$

Moreover, each function f_N changes its sign on every $T \in \mathcal{C}_{h(N)}$. In order to show that $\mathcal{L}^2(\mathcal{C}_{h(N)}) \to 0$ as $N \to \infty$ we argue by contradiction and assume instead that there is a constant $\kappa > 0$ such that

$$\mathcal{L}^2(\mathcal{C}_{h(N)}) \geq \kappa > 0 \quad \text{for all } N \in \mathbb{N}\,. \tag{4.39}$$

Since $f_N \to f$ in $L^r(\Omega)$ by construction we find for any $\varepsilon \in (0,\kappa)$ an index $N_\varepsilon \in \mathbb{N}$ such that for all $N \geq N_\varepsilon$ there holds $\mathcal{L}^2([\frac{\nu_N}{3} < |f_N - f|]) < \varepsilon \ll \kappa$. Thus, for all $N \geq N_\varepsilon$ it is $\mathcal{L}^2(\mathcal{C}_{h(N)}\backslash[\frac{\nu_N}{3} < |f_N - f|]) > \kappa - \varepsilon$. Moreover, for all $T \in \mathcal{T}_{h(N)}$ such that $\mathcal{L}^2(T \cap \mathcal{C}_{h(N)}\backslash[\frac{\nu_N}{3} < |f_N - f|]) > 0$ there is one node $\boldsymbol{y}_1$ such that $f_N(\boldsymbol{y}_1) > \frac{2\nu_N}{3} =: \omega$ and one node $\boldsymbol{y}_3$ such that $f_N(\boldsymbol{y}_3) \leq \omega_- < 0$. Thus Lemma 4.4 implies that $|\nabla f_N| \geq \frac{c\omega}{h(N)}$. Together with (4.38) this establishes a contradiction as

$$C \geq \sum_{T\in\mathcal{C}_{h(N)}} \int_T |\nabla f_N|^r \,\mathrm{d}x > \left[\!\!\left[\tfrac{\kappa-\varepsilon}{\mathcal{L}^2(T)} \right]\!\!\right] \mathcal{L}^2(T) \left| \tfrac{2c\nu_N}{3h(N)} \right|^r \geq (\kappa-\varepsilon) \left| \tfrac{2c}{3(h(N))^{1/2}} \right|^r \to \infty \quad \text{as } N \to \infty\,,$$

with $[\![\cdot]\!]$ the Gauß-bracket. Thus assumption (4.39) is false and we conclude that $\mathcal{L}^2(\mathcal{C}_{h(N)}) \to 0$. By (4.36) we also find

$$\mathcal{L}^2(\mathcal{B}_{h(N)} \cup \mathcal{C}_{h(N)}) \to 0 \tag{4.40}$$

and consequently $\mathcal{L}^2(\Omega\backslash\mathcal{A}_{h(N)}) = \mathcal{L}^2(\Omega\backslash(\mathcal{B}_{h(N)} \cup \mathcal{C}_{h(N)})) \to 0$.

To Step 3: Show that $\tilde{z}_N \rightharpoonup \tilde{z}$ in $\mathbf{X}$: By construction the elements of $\mathbf{X}_{h(N)}$ are continuous. Hence, B_N is an open set and, thanks to (4.35) also A_N has a non-empty interior (from a particular index on), which tends to Ω in measure. In view of construction (4.32) and notation (4.34) we have

$$\tilde{z}_N = \mathcal{I}_{h(N)}(\hat{z}_N - \nu_N) \text{ on } \mathcal{A}_{h(N)} \quad \text{ and } \tilde{z}_N = \overline{z}_N \text{ on } \mathcal{B}_{h(N)}\,, \tag{4.41}$$

while for $T \in \mathcal{C}_{h(N)}$ at least one node lies in A_N and at least one node lies in B_N, which requires $\tilde{z}_N = \mathcal{I}_{h(N)} \min\{\hat{z}_N - \nu_N, \overline{z}_N(t)\}$ to be an interpolation between the two functions.

By construction there holds $|\tilde{z}_N| \leq |\hat{z}_N - \nu_N| + |\overline{z}_N|$ a.e. in Ω, which gives

$$\|\tilde{z}_N\|_{L^r(\Omega)} \leq C\,. \tag{4.42}$$

To deduce compactness in $\mathbf{X}$ it thus remains to establish a uniform bound on the sequence of gradients $(\nabla\tilde{z}_N)_N$. Thanks to the above observations we have that

$$\nabla\tilde{z}_N = \nabla\overline{z}_N(t) \quad \text{for } T \in \mathcal{B}_{h(N)} \text{ and } \nabla\tilde{z}_N = \nabla\hat{z}_N \quad \text{for } T \in \mathcal{A}_{h(N)} \tag{4.43a}$$

Moreover, thanks to Lemma 4.3 we also find that

$$|\nabla\tilde{z}_N| \leq C(\alpha_T,\delta_T)\big(|\nabla\hat{z}_N| + |\nabla\overline{z}_N|\big) \quad \text{for } T \in \mathcal{C}_{h(N)}\,. \tag{4.43b}$$

Hence, putting together the estimates of (4.43) we obtain

$$\begin{aligned}
\|\nabla\tilde{z}_N\|_{L^r(\Omega)}^r &= \int_{\mathcal{A}_{h(N)}} |\nabla\mathcal{I}_{h(N)}(\hat{z}_N - \nu_N)|^r \,\mathrm{d}x + \int_{\mathcal{B}_{h(N)}} |\nabla\overline{z}_N(t)|^r \,\mathrm{d}x + \int_{\mathcal{C}_{h(N)}} |\nabla\tilde{z}_N|^r \,\mathrm{d}x \\
&\leq \int_{\mathcal{A}_{h(N)}} |\nabla\hat{z}_N|^r \,\mathrm{d}x + \int_{\mathcal{B}_{h(N)}} |\nabla\overline{z}_N(t)|^r \,\mathrm{d}x \\
&\qquad + \max_{T\in\mathcal{C}_{h(N)}} C(\alpha_T,\delta_T)^r \left(\int_{\mathcal{C}_{h(N)}} \big(|\nabla\hat{z}_N|^r + |\nabla\overline{z}_N|^r\big)\,\mathrm{d}x + \mathrm{LOT}(\mathcal{C}_{h(N)}) \right) \\
&\leq \int_{\mathcal{A}_{h(N)}} |\nabla\hat{z}_N|^r \,\mathrm{d}x + \int_{\mathcal{B}_{h(N)}} |\nabla\overline{z}_N(t)|^r \,\mathrm{d}x \\
&\qquad + C(\mathcal{T}_{h(N)})^r \left(\int_{\mathcal{C}_{h(N)}} \big(|\nabla\hat{z}_N|^r + |\nabla\overline{z}_N|^r\big)\,\mathrm{d}x + \mathrm{LOT}(\mathcal{C}_{h(N)}) \right) \\
&\leq C(\mathcal{T}_{h(N)})^r \left(\|\nabla\hat{z}_N\|_{L^r(\Omega)}^r + \|\nabla\overline{z}_N(t)\|_{L^r(\Omega)}^r + \mathrm{LOT}(\mathcal{C}_{h(N)}) \right)\,.
\end{aligned} \tag{4.44}$$

Here, we estimated $\max_{T\in\mathcal{C}_{h(N)}} C(\alpha_T,\delta_T) \leq \max_{T\in\mathcal{T}_{h(N)}} C(\alpha_T,\delta_T) =: C(\mathcal{T}_{h(N)})$. Moreover, the lower order terms gathered in $\mathrm{LOT}(\mathcal{C}_{h(N)})$ arise for any exponent $r \geq 1$ such that $r = [\![r]\!] + \tilde{r}$ with $[\![r]\!] \in \mathbb{N}$ and $\tilde{r} \in [0,1)$ by

applying a binonial expansion for the exponent $[\![r]\!] \in \mathbb{N}$ and subsequently Hölder's inequality, see Lemma 4.6 below. More precisely, by Lemma 4.6, $\mathrm{LOT}(\mathcal{C}_{h(N)})$ takes the form

$$\begin{aligned}\mathrm{LOT}(\mathcal{C}_{h(N)}) &= \|\nabla \hat{z}_N\|^{\tilde{r}}_{L^r(\mathcal{C}_N)} \|\nabla \overline{z}_N\|^{[\![r]\!]}_{L^r(\mathcal{C}_N)} + \|\nabla \hat{z}_N\|^{[\![r]\!]}_{L^r(\mathcal{C}_N)} \|\nabla \overline{z}_N\|^{\tilde{r}}_{L^r(\mathcal{C}_N)} \\ &\quad + \sum_{k=1}^{[\![r]\!]-1} C_{[\![r]\!]k} \big(\|\nabla \hat{z}_N\|^{r-k}_{L^r(\mathcal{C}_N)} \|\nabla \overline{z}_N\|^{k}_{L^r(\mathcal{C}_N)} + \|\nabla \hat{z}_N\|^{[\![r]\!]-k}_{L^r(\mathcal{C}_N)} \|\nabla \overline{z}_N\|^{\tilde{r}+k}_{L^r(\mathcal{C}_N)} \big)\end{aligned} \tag{4.45}$$

with binomial coefficients $C_{[\![r]\!]k}$. Here, $\|\nabla \overline{z}_N\|_{L^r(\mathcal{C}_N)} \le \|\nabla \overline{z}_N\|_{L^r(\Omega)} \le C$ by (4.9b) and $\|\nabla \hat{z}_N\|_{L^r(\mathcal{C}_N)} \to 0$ by (4.31) and (4.37), which provides that

$$\mathrm{LOT}(\mathcal{C}_{h(N)}) \to 0 \quad \text{as } N \to \infty\,. \tag{4.46}$$

Additionally, structure-assumption (4.13) also prescribes a uniform bound on $C(\alpha_T, \delta_T)$ in (4.44), so that (4.44) together with (4.42) yields that $\|\tilde{z}_N\|_{\mathbf{X}} \le \tilde{C}$ uniformly for all $N \in \mathbb{N}$, and by compactness, this implies that $\tilde{z}_N \rightharpoonup w$ in $\mathbf{X}$ and $\tilde{z}_N \to w$ in $L^r(\Omega)$ for some $w \in \mathbf{X}$. Moreover, by construction (4.32) combined with (4.40), we find

$$\begin{aligned}\|\tilde{z}_N - \tilde{z}\|^r_{L^r(\Omega)} &= \int_{\mathcal{A}_{h(N)}} |\hat{z}_N - \nu_N - \tilde{z}|^r \,\mathrm{d}x + \int_{\mathcal{B}_{h(N)} \cup \mathcal{C}_{h(N)}} |\tilde{z}_N - \tilde{z}|^r \,\mathrm{d}x \\ &\le \int_\Omega |\hat{z}_N - \nu_N - \tilde{z}|^r \,\mathrm{d}x + 2^r \int_{\mathcal{B}_{h(N)} \cup \mathcal{C}_{h(N)}} \big(|\tilde{z}_N|^r + |\tilde{z}|^r \big) \,\mathrm{d}x\,,\end{aligned} \tag{4.47}$$

where the first term on the right-hand side tends to 0 by the approximation properties of construction (4.29) & (4.30) and the second term vanishes by the strong convergence $\tilde{z}_N \to w$ in $L^r(\Omega)$ and (4.40). Thus, by (4.47) we are entitled to conclude that $w = \tilde{z}$ as well as

$$\tilde{z}_N \rightharpoonup \tilde{z} \text{ in } \mathbf{X}\,. \tag{4.48}$$

To Step 4: Conclude the mutual recovery condition (4.28)**:** For this we observe that

$$\begin{aligned}&\limsup_{N\to\infty} \big(\overline{\mathcal{E}}_N(t, \overline{u}_N(t), \tilde{z}_N) + \mathcal{R}(\tilde{z}_N - \overline{z}_N(t)) - \overline{\mathcal{E}}_N(t, \overline{u}_N(t), \overline{z}_N(t)) \big) \\ &\le \limsup_{N\to\infty} \tfrac{1}{2} \int_\Omega (f(\tilde{z}_N) - f(\overline{z}_N(t))) \big(\lambda |\operatorname{tr} e(\overline{u}_N + \overline{u_{\mathrm{D}N}}(t))|^2 + 2\mu |e(\overline{u}_N + \overline{u_{\mathrm{D}N}}(t)|^2 \big) \,\mathrm{d}x \\ &\quad + \limsup_{N\to\infty} \big(\mathcal{G}(\tilde{z}_N) - \mathcal{G}(\overline{z}_N(t)) \big) + \limsup_{N\to\infty} \mathcal{R}(\tilde{z}_N - \overline{z}_N(t))\,,\end{aligned}$$

where we used that the energy terms involving the Neumann boundary condition cancel out. Since $\tilde{z} \le z(t)$ by assumption and $\tilde{z}_N \le \overline{z}_N(t)$ by construction, we observe that $(f(\tilde{z}) - f(z(t)))$, and $(f(\tilde{z}_N) - f(\overline{z}_N(t))) \le 0$. Hence we can pass to the limit in the quadratic bulk term via weak lower semicontinuity, exploiting convergences (4.10) and (4.7). Furthermore, we have that $\mathcal{R}(\tilde{z}_N - \overline{z}_N(t)) \to \mathcal{R}(\tilde{z} - z(t))$ as well as $\int_\Omega \frac{\kappa_1}{r} \big(|\tilde{z}_N|^r - |\overline{z}_N(t)|^r \big) \,\mathrm{d}x \to \int_\Omega \frac{\kappa_1}{r} \big(|\tilde{z}|^r - |z(t)|^r \big) \,\mathrm{d}x$ thanks to convergences (4.10a) and (4.48). It remains to handle the difference of the damage gradients. From the second estimate in (4.44) we find

$$\begin{aligned}&\tfrac{\kappa_2}{r} \|\nabla \tilde{z}_N\|^r_{L^r(\Omega)} - \tfrac{\kappa_2}{r} \|\nabla \overline{z}_N(t)\|^r_{L^r(\Omega)} \\ \le\ & \int_{\mathcal{A}_{h(N)}} \tfrac{\kappa_2}{r} |\nabla \hat{z}_N|^r \,\mathrm{d}x - \int_{\mathcal{A}_{h(N)}} \tfrac{\kappa_2}{r} |\nabla \overline{z}_N(t)|^r \,\mathrm{d}x - \int_{\mathcal{C}_{h(N)}} \tfrac{\kappa_2}{r} |\nabla \overline{z}_N(t)|^r \,\mathrm{d}x \\ &\qquad + \tfrac{\kappa_2}{r} C(\mathcal{T}_{h(N)})^r \left(\int_{\mathcal{C}_{h(N)}} \tfrac{\kappa_2}{r} \big(|\nabla \hat{z}_N|^r + |\nabla \overline{z}_N|^r \big) \,\mathrm{d}x + \tfrac{\kappa_2}{r} \mathrm{LOT}(\mathcal{C}_{h(N)}) \right) \\ \le\ & \int_\Omega \tfrac{\kappa_2}{r} |\nabla \hat{z}_N|^r \,\mathrm{d}x - \int_{\mathcal{A}_{h(N)}} \tfrac{\kappa_2}{r} |\nabla \overline{z}_N(t)|^r \,\mathrm{d}x \\ &\qquad + \int_{\mathcal{C}_{h(N)}} \tfrac{\kappa_2}{r} \big(C(\mathcal{T}_{h(N)})^r |\nabla \hat{z}_N|^r + \big(C(\mathcal{T}_{h(N)})^r - 1 \big) |\nabla \overline{z}_N(t)|^r \big) \,\mathrm{d}x + \tfrac{\kappa_2}{r} C(\mathcal{T}_{h(N)})^r \mathrm{LOT}(\mathcal{C}_{h(N)})\end{aligned}$$

where $\|\nabla \hat{z}_N\|^r_{L^r(\Omega)} \to \|\nabla \tilde{z}\|^r_{L^r(\Omega)}$ by convergence (4.31). Moreover,

$$\int_{\mathcal{C}_{h(N)}} \tfrac{\kappa_2}{r} C(\mathcal{T}_{h(N)})^r |\nabla \hat{z}_N|^r \,\mathrm{d}x + \tfrac{\kappa_2}{r} C(\mathcal{T}_{h(N)})^r \mathrm{LOT}(\mathcal{C}_{h(N)}) \to 0$$

again by (4.31) together with (4.37), (4.46), and the uniform bound $C(\mathfrak{T}_{h(N)}) \leq C_{\mathfrak{T}}$ by structure-assumption (4.13). In addition, a priori estimate (4.9b) provides that $\|\nabla \overline{z}_N(t)\|^r_{L^r(\mathfrak{C}_{h(N)})} \leq C$, so that we deduce

$$\int_{\mathfrak{C}_{h(N)}} \tfrac{\kappa_2}{r}\big(C(\mathfrak{T}_{h(N)})^r - 1\big)|\nabla \overline{z}_N(t)|^r \,\mathrm{d}x \leq \tfrac{\kappa_2}{r}\big(C(\mathfrak{T}_{h(N)})^r - 1\big)C \tag{4.49}$$

and hence, this term results in the error ERR due to the mesh quality appearing in (4.11). Indeed, from Lemma 4.3 we gather that $\big(C(\mathfrak{T}_{h(N)})^r - 1\big) \to 0$ if $(\mathfrak{T}_{h(N)})_N$ tends to a right-angled triangulation as $N \to \infty$.

It remains to verify that also $\limsup_{N\to\infty} -\|\nabla \overline{z}_N(t)\|^r_{L^r(\mathcal{A}_{h(N)})} \leq -\|\mathrm{D}z(t)\|^r_{L^r(\Omega)}$. For this we may now repeat the arguments developed in [Tho13, Sec. 2.2]. More precisely, in view of (4.40) we may choose a further (not relabeled) subsequence such that $\mathcal{L}^2(\cup_{N\in\mathbb{N}}\mathcal{B}_{h(N)} \cup \mathcal{C}_{h(N)}) < \infty$. We then set $U_n := \cup_{N=n}^{\infty}\mathcal{B}_{h(N)} \cup \mathcal{C}_{h(N)}$ and $\Omega_n := \Omega\backslash U_n$. Since $B_N \subset U_n$ for all $N \geq n$, it is $\Omega_n \subset \mathcal{A}_{h(N)}$ for all $N \geq n$. Moreover, for all $n \in \mathbb{N}$, the set Ω_n is an open set and thus coincides with its measure-theoretic interior Ω_n^1, and by construction it holds

$$\Omega_n = \Omega_n^1 \subset \Omega_{n+1}^1 = \Omega_{n+1} \;\text{ for all } n \in \mathbb{N} \quad \text{and } \Omega_n \to \Omega\,. \tag{4.50}$$

Hence, along this subsequence, there holds $-\|\nabla \overline{z}_N(t)\|^r_{L^r(\mathcal{A}_{h(N)})} \leq -\|\nabla \overline{z}_N(t)\|^r_{L^r(\Omega_n)}$ for every $N \geq n$, and thus, by lower semicontinuity of the variation

$$\limsup_{N\to\infty} -\|\nabla \overline{z}_N(t)\|^r_{L^r(\mathcal{A}_{h(N)})} \leq -\liminf_{N\to\infty} \|\nabla \overline{z}_N(t)\|^r_{L^r(\Omega_n)} \leq -\|\nabla z(t)\|^r_{L^r(\Omega_n)}.$$

By (4.50) we now conclude for $N \geq n$ that

$$\limsup_{N\to\infty} -\|\nabla \overline{z}_N(t)\|^r_{L^r(\mathcal{A}_{h(N)})} \leq -\|\mathrm{D}z(t)\|^r_{L^r(\Omega_n)} \to -\|\nabla z(t)\|^r_{L^r(\Omega)} \quad \text{as } n \to \infty\,.$$

This concludes the proof of (4.28) and hence the limit passage in the discrete notion of solution (4.8).

Proof of the strong convergence (4.10c)**:** For this we verify the convergence of the stored elastic energy as a first step, i.e.,

$$\text{for all } t \in [0,T]:\ \mathcal{W}(\overline{u}_N(t), \underline{z}_N(t)) \to \mathcal{W}(u(t), \underline{z}(t)) \quad \text{as } N \to \infty\,. \tag{4.51}$$

To do so, we apply the construction of the recovery sequence (4.27) to the limit $u(t)$ itself, which provides the sequence $\tilde{u}_N(t) \to u(t)$ in $\mathbf{U}$. This together with convergence (4.10a) allows us to deduce the following chain of inequalities

$$\begin{aligned}\mathcal{W}(u(t), \underline{z}(t)) \leq \liminf_{N\to\infty} \mathcal{W}(\overline{u}_N(t), \underline{z}_N(t)) &\leq \limsup_{N\to\infty} \mathcal{W}(\overline{u}_N(t), \underline{z}_N(t)) \\ &\leq \limsup_{N\to\infty} \mathcal{W}(\tilde{u}_N, \underline{z}_N(t)) = \mathcal{W}(u(t), \underline{z}(t))\,,\end{aligned} \tag{4.52}$$

which proves (4.51). From this we now obtain that

$$\begin{aligned}&2\mu a\big|\|e(\overline{u}_N(t) + \overline{u_{\mathrm{D}}}_N(t))\|_{L^2(\Omega,\mathbb{R}^d)} - \|e(u(t) + u_{\mathrm{D}}(t))\|_{L^2(\Omega,\mathbb{R}^d)}\big| \\ &\leq \big|\mathcal{W}(\overline{u}_N(t), \underline{z}_N(t)) - \mathcal{W}(u(t), \underline{z}_N(t))\big| \\ &\leq \big|\mathcal{W}(\overline{u}_N(t), \underline{z}_N(t)) - \mathcal{W}(u(t), \underline{z}(t))\big| + \big|\mathcal{W}(u(t), \underline{z}(t)) - \mathcal{W}(u(t), \underline{z}_N(t))\big| \;\to 0\end{aligned} \tag{4.53}$$

where we exploited the coercivity of $\mathcal{W}$ provided by Assumption 2.1 as well as convergences (4.10a) and (4.51). Because of $\overline{u_{\mathrm{D}}}_N(t) \to u_{\mathrm{D}}(t)$ in $H^1(\Omega,\mathbb{R}^d)$ by (4.7) above convergence (4.53) implies that $e(\overline{u}_N(t)) \to e(u(t))$ in the locally uniformly convex space $L^2(\Omega,\mathbb{R}^d)$. With Korn's inequality we thus conclude $\overline{u}_N(t) \to u(t)$ in $\mathbf{U}$, i.e., (4.10c).

Proof that the pair $(u,\hat{z})$ is also a semistable energetic solution: We conclude the proof of Thm. 4.2 by verifying that also the pair $(u,\hat{z})$ with $\hat{z} := \max\{0, z\}$ provides a semistable energetic solution up to an error in the sense of (4.11). First of all, for any $t \in [0,T]$ fixed, we note that $\hat{z}(t) \in W^{1,r}(\Omega)$ according to [MM72, MM79], resp. $\hat{z} \in \mathrm{BV}(\Omega)$ according to the chain rule for real-valued BV-functions composed with a Lipschitz function, cf. [AFP05, Thm. 3.99], which makes $\hat{z}$ an admissible candidate for a semistable energetic solution.

Since $f(z) = a$ whenever $z \leq 0$ in Ω we have $\mathcal{E}(t,u(t),z(t)) - \mathcal{G}(z(t)) = \mathcal{E}(t,u(t),\hat{z}(t)) - \mathcal{G}(\hat{z}(t))$. Hence, for all $t \in [0,T]$ the minimality property (1.5a) of $u(t)$ is also satisfied when replacing $z(t)$ by $\hat{z}(t)$. In order to verify that $(u,\hat{z})$ also satisfies the semistability condition (1.5b), resp. (4.11), and the energy-dissipation estimate (1.5c), we additionally note that

$$\mathcal{G}(\hat{z}(t)) \;\leq\; \mathcal{G}(z(t)) \quad \text{as well as} \tag{4.54a}$$

$$\mathcal{R}(\hat{z}(t) - z_0) \;\leq\; \mathcal{R}(z(t) - z_0)\,. \tag{4.54b}$$

Clearly, $\hat z(t) = \max\{0, z(t)\} \geq z(t)$, hence $z_0 - \hat z(t) \leq z_0 - z(t)$, which yields (4.54b). Similarly, $|\max\{0, z(t)\}| \leq |z(t)|$, so that $\|\hat z(t)\|^r_{L^r(\Omega)} \leq \|z(t)\|^r_{L^r(\Omega)}$. Moreover, according to the chain rule for real-valued BV-functions composed with a Lipschitz function $g : \mathbb{R} \to \mathbb{R}$, cf. [AFP05, Thm. 3.99], it holds $\hat z(t) = g \circ z(t) \in BV(\Omega)$, here with $g(\cdot) := \max\{0, \cdot\}$, and for the gradient there holds

$$\mathrm{D}\hat z(t) = g'(z(t))\nabla z(t)\mathcal{L}^2 + \big(g(z^+(t)) - g(z^-(t))\big)\nu_{z(t)}\mathcal{H}^1\big|_{J_{z(t)}} + u_{\mathrm{D}}{}'(\widetilde z(t))\mathrm{D}^c z(t)\,, \tag{4.55}$$

where $\widetilde z(t)$ denotes the approximate limit of $z(t)$, $z^\pm$ are the approximate limits from the right and from the left at the jump set $J_{z(t)}$. More precisely, for $g(\cdot) := \max\{0, \cdot\}$ we have $g'(z) = 0$ for $z < 0$, $g'(z) = 1$ for $z > 0$, and $g'(z) \in [0, 1]$ for $z = 0$. Hence, by (4.55) we find that

$$|\mathrm{D}\hat z(t)|(\Omega) \leq |\mathrm{D}z(t)|(\Omega), \text{ resp. } \|\nabla \hat z(t)\|^r_{L^r(\Omega)} \leq \|\nabla z(t)\|^r_{L^r(\Omega)}\,, \tag{4.56}$$

which proves (4.54a). Estimates (4.54) now allow us to deduce for all $t \in [0, T]$ that

$$\mathcal{E}(t, u(t), \hat z(t)) + \mathcal{R}(\hat z(t) - z_0) \leq \mathcal{E}(t, u(t), z(t)) + \mathcal{R}(z(t) - z_0),$$

which suffices to conclude the energy-dissipation estimate (1.5c) for $(u(t), \hat z(t))$, again, also since $f(z) = a$ for all $z \leq 0$.

In order to verify the semistability of $(u(t), \hat z(t))$ consider now a competitor $\tilde z \in \mathbf{Z}$. If $\tilde z \geq \hat z(t)$ on a subset of Ω with non-zero Lebesgue-measure, then $\mathcal{R}(\tilde z - \hat z(t)) = \infty$ and (4.11) is satisfied with $\hat z(t)$. In case of $\tilde z \leq z(t)$ a.e. in Ω we find $\mathcal{R}(\tilde z - z(t)) \leq \mathcal{R}(\tilde z - \hat z(t))$. Together with (4.54) and the semistability of $z(t)$ we thus see that

$$\begin{aligned}\mathcal{E}(t, u(t), \hat z(t)) \leq \mathcal{E}(t, u(t), z(t)) &\leq \mathcal{E}(t, u(t), \tilde z) + \mathcal{R}(\tilde z - z(t)) + \mathrm{ERR}\\ &\leq \mathcal{E}(t, u(t), \tilde z) + \mathcal{R}(\tilde z - \hat z(t)) + \mathrm{ERR}\,.\end{aligned} \tag{4.57}$$

There remains to discuss the case that $\tilde z \leq \hat z(t)$ but $\tilde z \geq z(t)$ on a set $B \subset \Omega$ of non-zero Lebesgue-measure, so that $\mathcal{R}(\tilde z - \hat z(t))$ is finite but $\mathcal{R}(\tilde z - z(t)) = \infty$. By constuction there holds $\hat z = 0$ a.e. on B, whereas $\tilde z \leq 0$ a.e. on B is possible. At this point we note that the function $\bar z := \min\{\tilde z, z(t)\}$ satisfies $\mathcal{R}(\bar z - z(t)) < \infty$ with $\bar z = z(t)$ on B and $\mathcal{R}(\bar z - z(t)) \leq \mathcal{R}(\tilde z - \hat z(t))$ by construction of $\bar z$ and $\hat z(t)$. Accordingly, with the above reasonings, we are again in the position to verify (4.57). This concludes the proof of Thm. 4.2. ■

We verify now the following statement, which was used to calculate the lower oder terms $\mathrm{LOT}(\mathcal{C}_{h(N)})$, cf. (4.45), in Step 3 of the proof of Thm. 4.2.

Lemma 4.6 (Binomial estimate for the lower order terms)**.** *Let $f, g \in L^r(\Omega)$ be non-negative functions and the exponent $r \geq 1$ such that $r = [\![r]\!] + \tilde r$ with $[\![r]\!] = \max\{x \in \mathbb{N},\, x \leq r\} \in \mathbb{N}$ the Gauß-bracket and $\tilde r \in [0, 1)$. Then*

$$\int_\Omega (f + g)^r \,\mathrm{d}x \leq \int_\Omega \big(f^r + f^{[\![r]\!]}g^{\tilde r} + g^{[\![r]\!]}f^{\tilde r} + g^r\big)\,\mathrm{d}x + \sum_{k=1}^{[\![r]\!]-1} \int_\Omega C_{[\![r]\!]k}\big(f^{r-k}g^k + f^{[\![r]\!]-k}g^{\tilde r+k}\big)\,\mathrm{d}x\,, \tag{4.58}$$

with binomial coefficients $C_{[\![r]\!]k} := \frac{[\![r]\!]!}{k!([\![r]\!]-k)!}$.

Proof. From [HLP34, Thm. 19, p. 28] we pick the following result: *Let $f_j \geq 0$, $j = 1, \dots, n$. If $0 < \tilde r < s$, then*

$$\left(\sum_{j=1}^n f^s\right)^{1/s} \leq \left(\sum_{j=1}^n f^{\tilde r}\right)^{1/\tilde r}\,, \tag{4.59}$$

unless all a_j but one are zero.

In what follows we are going to apply (4.59) with the exponents $s = 1$ and $\tilde r \in (0, 1)$ from the statement of the lemma and using a binomial expansion for the exponent $[\![r]\!]$ we argue as follows:

$$\begin{aligned}\int_\Omega (f + g)^r \,\mathrm{d}x \;&=\; \int_\Omega (f + g)^{[\![r]\!]}(f + g)^{\tilde r}\,\mathrm{d}x = \sum_{k=0}^{[\![r]\!]} C_{[\![r]\!]k} \int_\Omega f^{[\![r]\!]-k}g^k(f + g)^{\tilde r}\,\mathrm{d}x\\
&\overset{(4.59)}{\leq} \sum_{k=0}^{[\![r]\!]} C_{[\![r]\!]k} \int_\Omega f^{[\![r]\!]-k}g^k(f^{\tilde r} + g^{\tilde r})\,\mathrm{d}x\\
&=\; \int_\Omega \big(f^r + f^{[\![r]\!]}g^{\tilde r} + g^{[\![r]\!]}f^{\tilde r} + g^r\big)\,\mathrm{d}x + \sum_{k=1}^{[\![r]\!]-1} \int_\Omega C_{[\![r]\!]k}\big(f^{r-k}g^k + f^{[\![r]\!]-k}g^{\tilde r+k}\big)\,\mathrm{d}x\,.\end{aligned} \tag{4.60}$$

Subsequently, the terms $\int_\Omega f^\alpha g^\beta \,\mathrm{d}x$ are estimated by Hölder's inequality such that a factor of the form $\big(\int_\Omega f^r \,\mathrm{d}x\big)^\gamma$ with $\gamma \in (0,1)$ arises. More precisely, with the exponent $q = r/\alpha > 1$ for $\alpha \in (0,r)$ it is $q' = \frac{q}{q-1} = \frac{r}{r-\alpha}$ and thus Hölder's inequality provides

$$\int_\Omega f^\alpha g^\beta \,\mathrm{d}x \le \left(\int_\Omega f^r\right)^{\alpha/r} \left(\int_\Omega g^{\beta r}\right)^{(r-\alpha)/r} = \|f\|_{L^r(\Omega)}^\alpha \|g\|_{L^{\frac{\beta r}{r-\alpha}}(\Omega)}^\beta . \tag{4.61}$$

It can be further checked for all summands appearing in (4.60) that $\frac{\beta r}{r-\alpha} = r$. Thus, inserting (4.61) into (4.60) gives (4.58). □

We conclude this section with some remarks on the convergence properties of the sequence $(\overline{z}_N)_N$, those of its recovery sequence (4.32), and the semistability (4.11) up to an error due to the mesh quality.

Remark 4.7 (Improved convergence of the approximate solutions $(\overline{z}_N)_N$)**.** As stated in (4.10a), in general, only the weak convergence $\overline{z}_N(t) \rightharpoonup z(t)$ of semistable sequences $(\overline{z}_N)_N$ can be verified. A proof of strong or strict convergence is hampered by the non-smooth unidirectionality constraint featured in $\mathcal{R}$ and by the separate convexity of $\mathcal{E}(t,\cdot,\cdot)$. Yet, suppose that the energy-dissipation estimate holds true as an equality for the limit system, cf. Remark 1.3. Then one finds the improved convergence

$$\mathcal{G}(\overline{z}_N(t)) \to \mathcal{G}(z(t)) \quad \text{as } N \to \infty . \tag{4.62}$$

This can be established from the energy-dissipation estimate at level N and the balance for the limit together with convergences (4.7), (4.10a), and (4.10c) as follows:

$$\begin{aligned} \limsup_{N\to\infty} \mathcal{G}(\overline{z}_N(t)) &\le \limsup_{N\to\infty} \Big(\int_0^t \partial_\xi \overline{\mathcal{L}}^N_{\mathrm{Neu}}(\xi, \underline{u}_N(\xi)) \,\mathrm{d}\xi - \mathcal{W}(\overline{u}_N(t), \overline{z}_N(t)) - \overline{\mathcal{L}}^N_{\mathrm{Neu}}(t, \overline{u}_N(t)) \Big) \\ &= \int_0^t \partial_\xi \mathcal{L}_{\mathrm{Neu}}(\xi, u(\xi)) \,\mathrm{d}\xi - \mathcal{W}(u(t), z(t)) - \mathcal{L}_{\mathrm{Neu}}(t, u(t)) = \mathcal{G}(z(t)) , \end{aligned} \tag{4.63}$$

whereas the opposite estimate follows by convergence (4.10a) from the weak sequential lower semicontinuity $\liminf_{N\to\infty} \mathcal{G}(\overline{z}_N(t)) \ge \mathcal{G}(z(t))$.

Remark 4.8 (Improved convergence of the recovery sequence)**.** Suppose that the improved convergence (4.62) holds true. Then it follows

$$\mathcal{G}(\tilde{z}_N(t)) \to \mathcal{G}(\tilde{z}(t)) \quad \text{as } N \to \infty \tag{4.64}$$

for all $t \in [0,T]$ also for the mutual recovery sequence $(\tilde{z}_N)_N$ constructed by (4.32). Omitting to indicate the dependence on $t \in [0,T]$ and using the notation of (4.28)–(4.50) this can be seen as follows: In view of (4.32) and (4.34) it is

$$\mathcal{G}(\tilde{z}_N) = \int_{\mathcal{A}_{h(N)}} |\nabla \mathcal{I}_{h(N)}(\hat{z}_N - \nu_N)|^r \,\mathrm{d}\boldsymbol{x} + \int_{\mathcal{B}_{h(N)}} |\nabla \overline{z}_N|^r \,\mathrm{d}\boldsymbol{x} + \int_{\mathcal{C}_{h(N)}} |\nabla \tilde{z}_N|^r \,\mathrm{d}\boldsymbol{x} \tag{4.65}$$

Moreover, by (4.43b) we estimate

$$\int_{\mathcal{B}_{h(N)}} |\nabla \overline{z}_N|^r \,\mathrm{d}\boldsymbol{x} + \int_{\mathcal{C}_{h(N)}} |\nabla \tilde{z}_N|^r \,\mathrm{d}\boldsymbol{x} \le C(\alpha,\delta)^r \int_{\mathcal{B}_{h(N)}\cup\mathcal{C}_{h(N)}} \big(|\nabla \hat{z}_N| + |\nabla \overline{z}_N|\big)^r \,\mathrm{d}\boldsymbol{x} , \tag{4.66}$$

where we set $C(\alpha,\delta) := \max_{T\in\mathcal{B}_{h(N)}\cup\mathcal{C}_{h(N)}, N\in\mathbb{N}} C(\alpha_T,\delta_T)$. Because of $\mathcal{L}^2(\mathcal{B}_{h(N)} \cup \mathcal{C}_{h(N)}) \to 0$ by (4.40) and by the strict/strong convergences (4.43b) and (4.62) we now argue that the term on the right-hand side of (4.66) tends to 0. Accordingly, we now argue in (4.65) that

$$\limsup_{N\to\infty} \mathcal{G}(\tilde{z}_N) \le \limsup_{N\to\infty} \int_\Omega |\nabla \mathcal{I}_{h(N)}(\hat{z}_N - \nu_N)|^r \,\mathrm{d}\boldsymbol{x} + \lim_{N\to\infty} C(\alpha,\delta)^r \int_{\mathcal{B}_{h(N)}\cup\mathcal{C}_{h(N)}} \big(|\nabla \hat{z}_N| + |\nabla \overline{z}_N|\big)^r \,\mathrm{d}\boldsymbol{x} = \mathcal{G}(\tilde{z})$$

again by the strict/strong convergences (4.43b), whereas the opposite estimate $\liminf_{N\to\infty} \mathcal{G}(\tilde{z}_N) \ge \mathcal{G}(\tilde{z})$ follows by lower semicontinuity.

Remark 4.9 (Semistability in case of improved convergence)**.** In case that the improved convergences (4.62) and (4.64) hold true the semistability proof for the limit function z does not require the cancellation arguments used in Step 4 of the proof of Thm 4.2. In this way the error term ERR due to the mesh quality arising in (4.49) will be omitted.

Remark 4.10 (Energy-dissipation balance)**.** It has to be stressed that the assumption on the validity of an energy-dissipation *balance* made in Remark 4.7 is rather strong. In the general situation of a separately convex energy density it might result by particular loading scenarios, only, e.g., such that solutions can still evolve smoothly in time. Also in case of a convex energy functional (jointly in the pair (u, z)) the proof an energy-dissipation balance for the limit system is usually based on a Riemann-sum argument based on the (already verified) stability (1.8a) of the limit, see, e.g., [MR15]. If the latter can be only verified upon an error due to the mesh quality as in (4.11), then this error term may also hamper the arguments to find the energy-dissipation balance. One way to obtain an energy-dissipation balance for the limit system is to incorporate an additional viscous dissipation into the approximating systems, so that the damage variable is of better regularity in time provided by an L^2-gradient flow, cf. [ABN18] for the convergence analysis also in combination with a FE-discretization. Using a vanishing viscosity method, which inserts an additional viscous dissipation for the damage variable on the approximating level and applies an arc-length reparametrization for the limit passage one may deduce *Balanced-Viscosity solutions*, cf. e.g. [KRZ13]. The *parametrized Balanced-Viscosity solutions* are characterized by an energy-dissipation balance featuring additional dissipative contributions arising as an artefact of the vanished viscous dissipation which is active at jump times. A similar result may be obtained when using multi-step approximation method of [KN17, AN19]: There, the staggered time-discrete algorithm is not iterated only once per single time-step but multiple times. Arc-length reparametrization also proves the discrete solutions to approximate *parametrized Balanced-Viscosity solutions* as time-step size tends to 0, cf. also Remark 3.6 for more details on the multi-step algorithm.

Remark 4.11 (The case $r = 2$)**.** *Let the assumptions of Thm. 4.2 hold true and $r = 2$. Then semistability inequality* (4.11) *is valid with $ERR = 0$ independently of the approximating triangulations $(\mathcal{T}_{h(N)})_N$.*

This is due to the fact that in the Hilbert space setting $r = 2$ the sequence $\tilde{z}_N := \overline{z}_N + \mathcal{I}_{h(N)}(\tilde{z} - z)$ can be used as a mutual recovery sequence in place of (4.32). We first observe that $\tilde{z}_N \leq \overline{z}_N$ since $\mathcal{I}_{h(N)}(\tilde{z} - z) \leq 0$ for $\tilde{z} \leq z$ a.e. in Ω and that $\nabla\mathcal{I}_{h(N)}(\tilde{z} - z) \to \nabla(\tilde{z} - z)$ strongly in $H^1(\Omega)$ by the properties of the interpolation operator $\mathcal{I}_{h(N)}$. Hence, for the gradient term one finds

$$\begin{aligned}\lim_{N\to\infty}\Big(\mathcal{G}(\tilde{z}_N) - \mathcal{G}(\overline{z}_N)\Big) &= \lim_{N\to\infty}\int_\Omega \big(|\nabla\overline{z}_N|^2 + 2\nabla\overline{z}_N\cdot\nabla\mathcal{I}_{h(N)}(\tilde{z}-z) + |\nabla\mathcal{I}_{h(N)}(\tilde{z}-z)|^2 - |\nabla\overline{z}_N|^2\big)\,\mathrm{d}x\\ &= \int_\Omega \big(|\nabla\tilde{z}|^2 - |\nabla z|^2\big)\big)\,\mathrm{d}x = \mathcal{G}(\tilde{z}) - \mathcal{G}(z)\,.\end{aligned}$$

See also [AB19], where the same argument is used for approximation of the Ambrosio-Tortorelli functional.

5 Numerical experiments

In this section we illustrate typical evolutions for the considered rate-independent model problem and the performance of the devised numerical method via numerical studies for two-dimensional specifications of the general framework. The first one studies a benchmark problem in which a plate with hole is stretched via loads on two opposite sides. The second one uses as square-shaped domain which undergoes a shearing deformation via nonhomogeneous loads on two neighboring sides. In the experiments we investigate the dependence of approximations on the integrability exponent r and validate the energy inequality for the iterative solution. In addition we test the influence of geometrical properties of finite element meshes on the numerical solution.

5.1 Setup

Figure 5.1 displays the geometries of the two considered settings. Using expected symmetry features of solutions the first problem is reduced to one quadrant of the original domain.

We use the same material parameters and initial conditions for both settings which are specified in Table 1. The iteration of Algorithm 3.1 stops if the residual satisfies $R_j \leq 10^{-6}/(2\max\{1, 1/(\tau_j h_{\min})\})$ with the parameters $\overline{\tau} = h_{\min}^{-2}$, $\underline{\tau} = 10^{-4}$, $\delta = 0.5$, $\underline{\gamma} = 0.5$, $\overline{\gamma} = 0.999$.

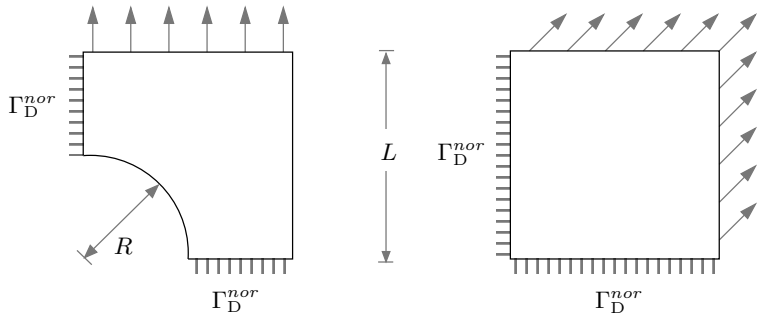


Figure 5.1: Geometries of the domain with hole and square-shaped domain with Dirichlet boundaries $\Gamma_{\mathrm{D}}^{nor}$ along which normal components of displacements vanish. The remaining parts of the boundaries are Neumann boundaries with surface loads as indicated and specified below.

time horizon, domain size	$\mathsf{T} = 1\mathrm{s}$, $L = 1\mathrm{mm}$, $R = \sqrt{2}/3 \cdot L$, cf. Fig. 5.1
initialization	stable state $u_h^0 \equiv 0$, $z_h^0 \equiv 1$
fracture toughness	$\rho = 4 \cdot 10^{-4}\mathrm{N/mm^2}$;
degradation function	$f(z) = a + (b-a)z^2$, $z \geq 0$, $a = 1/2$, $b = 1$
regularization parameters	$\kappa_1 = 10^{-6}\mathrm{N}$, $\kappa_2 = 10^{-9}\mathrm{N/mm^2}$, $r \in \{1.1, 1.5, 2\}$
partition in time	equidistant with step-size $\Delta t = 0.1 \cdot \mathsf{T}h$
partition in space	quasiuniform triangulations with mesh-size h
Lamé constants	$\lambda = 4142.9\mathrm{N/mm^2}$ and $\mu = 1035.7\mathrm{N/mm^2}$

Table 1: Settings used for the numerical experiments.

5.2 Membrane with hole

For the first setting with the reduced domain $\Omega = (0, L)^2 \setminus \{x \in \mathbb{R}^2 : |x| \leq R\}$, which has been used in [ACFK02] in the context of elastoplasticity, depicted in the left plot of Figure 5.1 we use nonhomogeneous, time-dependent Neumann conditions on the upper side, i.e., along $\Gamma_{\mathrm{Neu}}^{top} = (0, L) \times \{L\}$, given by

$$l_{\mathrm{Neu}}(t, x) = t \cdot 0.75 \, \frac{\mathrm{N}}{\mathrm{mm^2 s}} \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

We used four triangulations $\mathcal{T}_h$ with different mesh-sizes $h > 0$ generated with the routine `distmesh` (see [PS04]).

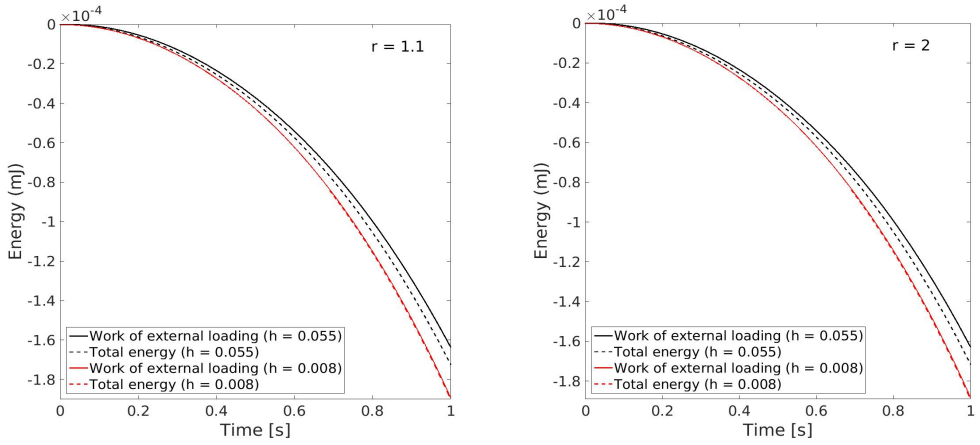


Figure 5.2: Verification of energy estimate (4.8c) as a function of t_N^n for two different mesh-sizes and regularization parameters $r = 1.1$ (left) and $r = 2$ (right). Sum of stored and dissipated energy (= total energy = left-hand side of (4.8c); power of external loading up to time t_N^n (= 'work' = right-hand side of (4.8c) with $\overline{\mathcal{E}}_N(0, q_N^0) = 0$). The sum of the stored and dissipated energy is bounded by the power of the external loading at every time step, their differences decrease with the mesh-size h.

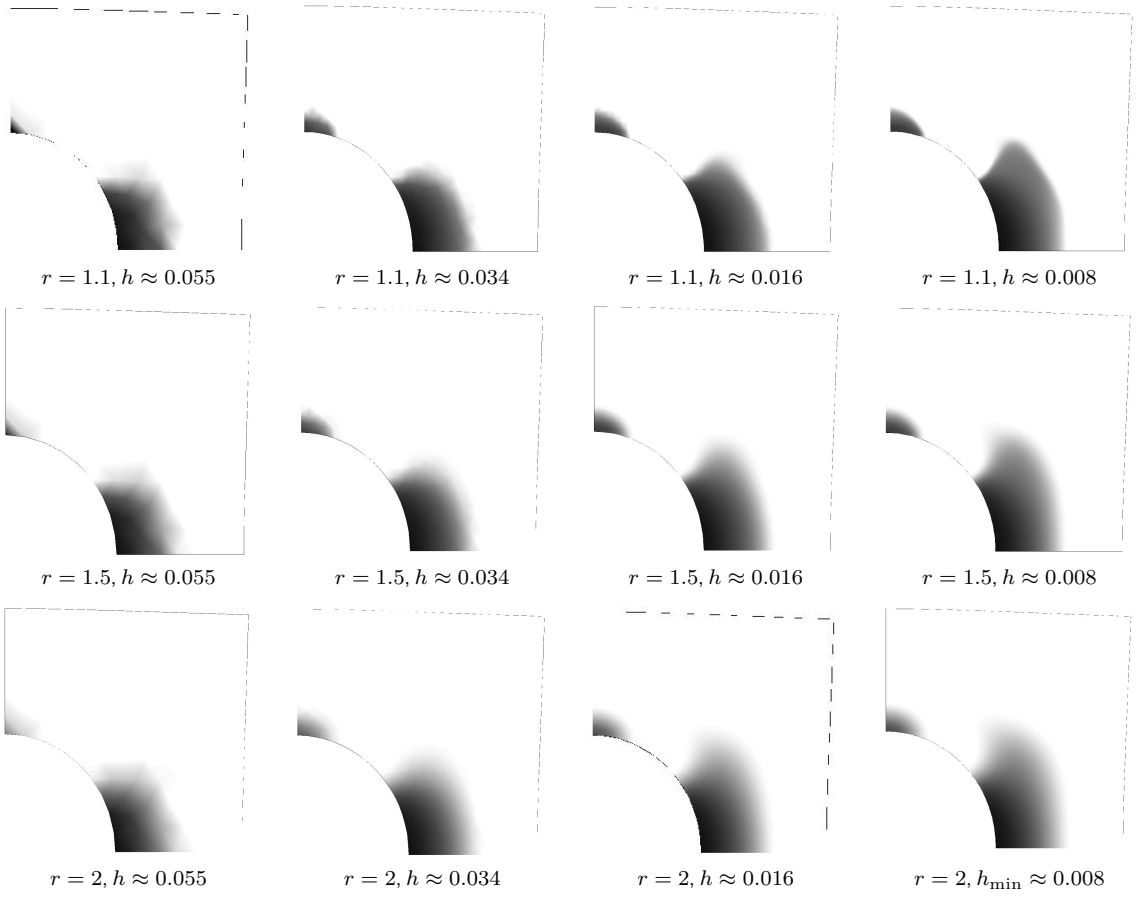


Figure 5.3: Damage variable at final time $\mathsf{T} = 1\mathrm{s}$ for different mesh-sizes and Sobolev exponents; displacements are magnified by a factor of 40. The sharpness of the interface between damaged and undamaged regions depends on the regularization parameter r.

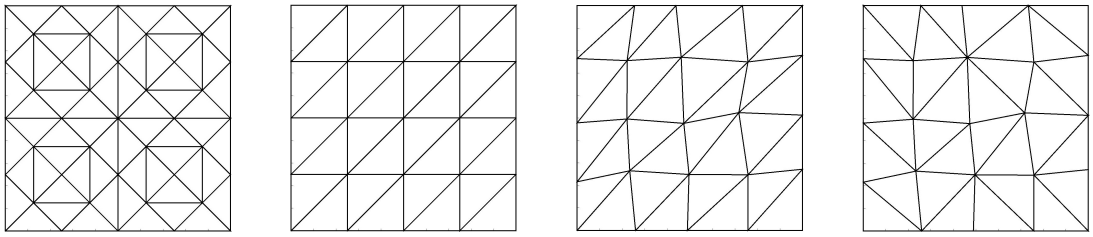
In Figure 5.3 the damaged and deformed material at final time $\mathsf{T} = 1\mathrm{s}$ is depicted for the Sobolev exponents $r = 1.1, 1.5$ and 2 computed by Algorithm 3.5 for $h = 0.055, 0.034, 0.016, 0.008$. The displacements are magnified by a factor of 40. The choice of the exponent r influences the sharpness of the interface between damaged and undamaged regions. One clearly observes that the interface becomes more diffuse as the exponent tends to two. The interface for $r = 1.1$ is localized but has certain smoothness properties, The geometry of the damaged region depends on the mesh size h but appears to converge. We observe slightly larger damaged areas for smaller mesh-sizes h, which is the result of a better localization for the interface. In Figure 5.2 we verify the upper energy-dissipation estimate (4.8c) as a function of $t_N^n, n \leq N$, for two mesh-sizes $h = 0.055, 0.008$. Indeed, (4.8c) holds true, the involved terms decrease with $t \in [0, \mathsf{T}]$, and the difference of the left- and the right-hand side of the become smaller with the mesh-size h; see Subsection 5.4 below for a discussion of the results.

5.3 Sheared square

The geometry of the domain $\Omega = (0, L)^2$ is seen on the right side of Figure 5.1. The material is pulled in northeast direction on the upper and right boundary sides $\Gamma_{\mathrm{Neu}} = \big((0, L) \times \{L\}\big) \cup \big(\{L\} \times (0, L)\big)$ modeled by the surface force

$$l_{\mathrm{Neu}}(t, x) = t \cdot 0.9 \, \frac{\mathrm{N}}{\mathrm{mm^2 s}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

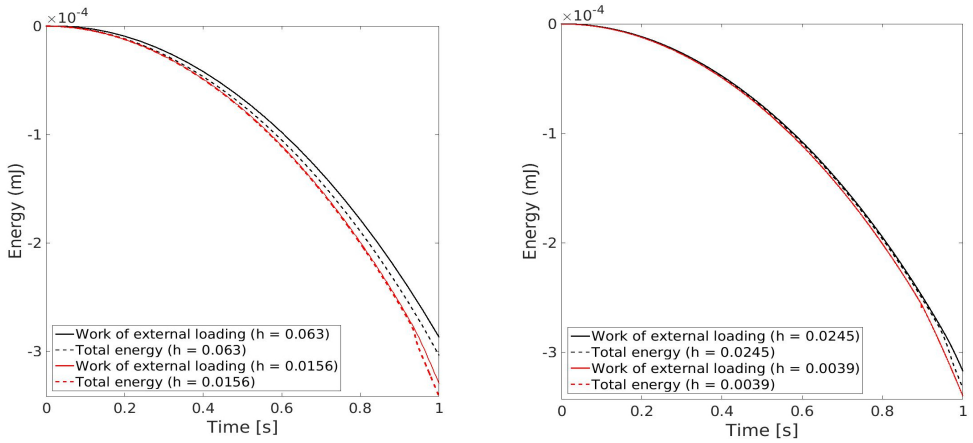
To investigate the influence of symmetry properties of triangulations we test our algorithm on right-angled symmetric, right-angled non-symmetric, perturbed and Delaunay triangulations. The triangulation types are illustrated in Figure 5.4. We obtain the right-angled symmetric and right-angled nonsymmetric triangulations with a direct application of uniform



right-angled symmetric triangulation | right-angled nonsymmetric triangulation | perturbed triangulation | Delaunay triangulation

Figure 5.4: Right-angled symmetric, right-angled nonsymmetric, perturbed and Delaunay triangulations of a square with comparable mesh-sizes.

red refinement procedures. To generate the perturbed triangulations we displaced the inner nodes of the initial right-angled nonsymmetric triangulation after every uniform refinement step by random vectors with mean value zero and standard deviation $h/12$. The Delaunay triangulations are obtained by applying a Delaunay regularization procedure to the perturbed triangulations.



right-angled nonsymmetric triangulation | perturbed triangulation

Figure 5.5: Verification of energy-dissipation estimate (4.8c) as a function of time t_N^n for two different mesh-sizes and regularization parameter $r = 1.1$ using right-angled nonsymmetric (left) and perturbed (right) triangulations. Sum of stored and dissipated energy (= total energy = left-hand side of (4.8c); power of external loading up to time t_N^n (= 'work' = right-hand side of (4.8c) with $\overline{\mathcal{E}}_N(0, q_N^0) = 0$). The sum of stored and dissipated energy is bounded by the power of the external loading for every time step.

In Figure 5.6 the damaged and deformed material at time $\mathsf{T} = 1\mathrm{s}$ is depicted for the Sobolev exponent $r = 1.1$ computed by Algorithm 3.5 for different mesh-sizes h. The displacements are magnified by a factor of 40. As expected the damage is in regions with a high concentration of stresses. For larger mesh sizes h we obtain different damaged regions depending on the triangulation. In the case of the right-angled nonsymmetric, perturbed and Delaunay triangulation we have the same basis triangulation and the same number of red refinements. For small mesh-sizes h we do not observe significant differences. Our overall observation from the numerical experiments is that we find better robustness for varying triangulations than expected from theory; only the iteration numbers of the ADMM iteration in the time steps were smaller for right-angled triangulations. In Figure 5.5 we verify the upper energy-dissipation estimate (4.8c) for the right-angled nonsymmetric and perturbed triangulation as a function of $t_N^n, n \leq N$ for two mesh-sizes h. The chronological development slightly differs for the different triangulations but differences decrease in both cases as discretizations become more accurate; see Subsection 5.4 for a discussion of the results.

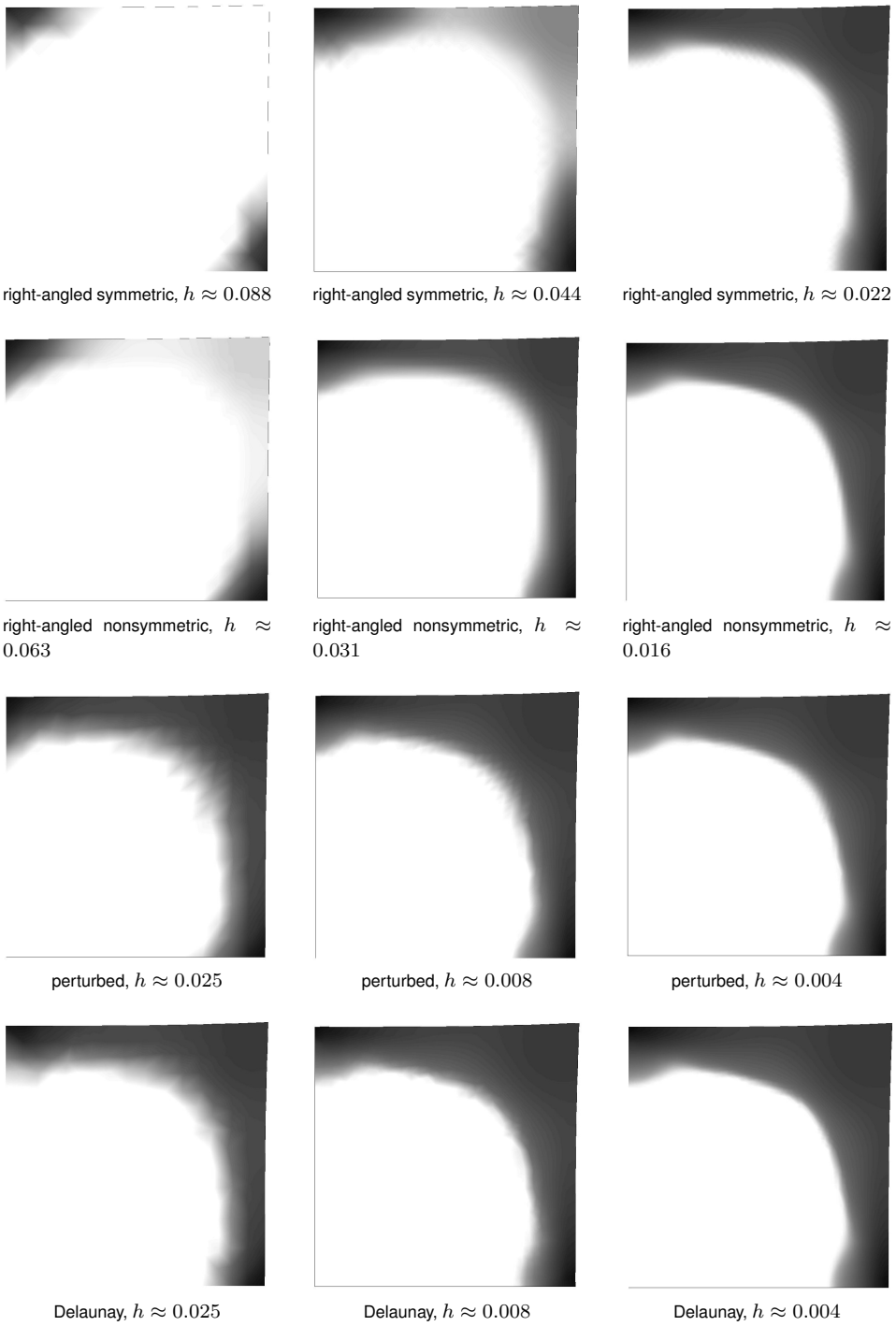


Figure 5.6: Damage variable at $\mathsf{T} = 1\mathrm{s}$ for different triangulations and mesh-sizes and fixed regularization parameter $r = 1.1$; displacements are magnified by a factor of 40. Differences of solutions visible for coarse meshes disappear as mesh-sizes become finer.

5.4 Discussion

In both benchmark examples the results in Fig. 5.2 and 5.5 hint that the upper energy dissipation estimate (4.8c) approaches an equality as $h \to 0$ in the following sense

$$\begin{aligned}
&\int_0^t \partial_\xi \mathcal{E}_N(\xi, \underline{u}_N(\xi))\,\mathrm{d}\xi = \sum_{k=1}^{K_N(t)} \int_{t_N^{k-1}}^{t_N^k} \partial_\xi \mathcal{L}_{\mathrm{Neu}}^N(\xi, u_N^{k-1}(\xi))\,\mathrm{d}\xi \\
&\geq \overline{\mathcal{E}}_N(t, \overline{q}_N(t)) - \overline{\mathcal{E}}_N(0, q_N^0) + \mathrm{Diss}_{\mathcal{R}}(\overline{z}_N, [0,t]) - \mathrm{TOL}(N)N \\
&= \sum_{k=1}^{K_N(t)} \big(\mathcal{W}(z_N^k, u_N^k) - \mathcal{W}(z_N^{k-1}, u_N^{k-1}) + \mathcal{L}_{\mathrm{Neu}}(t_N^k, u_N^k) - \mathcal{L}_{\mathrm{Neu}}(t_N^{k-1}, u_N^{k-1}) + \mathcal{G}(z_N^k) - \mathcal{G}(z_N^{k-1}) + \mathcal{R}(z_N^k - z_N^{k-1})\big) \\
&\qquad\qquad - \mathrm{TOL}(N)N \\
&= \sum_{k=1}^{K_N(t)} \int_{t_N^{k-1}}^{t_N^k} \partial_\xi \mathcal{L}_{\mathrm{Neu}}^N(\xi, u_N^{k-1}(\xi))\,\mathrm{d}\xi + \mathrm{Diff(N)}
\end{aligned}$$

$$\begin{aligned}
\text{with}\quad 0 \geq \mathrm{Diff(N)} := &\sum_{k=1}^{K_N(t)} \big(\mathcal{W}(z_N^k, u_N^k) - \mathcal{W}(z_N^{k-1}, u_N^{k-1}) + \mathcal{L}_{\mathrm{Neu}}(t_N^k, u_N^k) - \mathcal{L}_{\mathrm{Neu}}(t_N^k, u_N^{k-1})\big) \\
&+ \sum_{k=1}^{K_N(t)} \big(\mathcal{G}(z_N^k) - \mathcal{G}(z_N^{k-1}) + \mathcal{R}(z_N^k - z_N^{k-1})\big) - \mathrm{TOL}(N)N\,,
\end{aligned}$$

and $\mathrm{Diff(N)} \to 0$ as $N \to \infty$.

As a first interpretation, this appears to indicate for the two benchmark examples that the upper energy-dissipation estimate (1.5c) holds as an equality along the limit solutions. This is the case if the algorithm indeed approximates here an enhanced semistable energetic solution or even a global energetic solution, cf. Remark 1.3. However, this result cannot be ultimately concluded from $\mathrm{Diff(N)} \to 0$, also because the convergence result

$$\mathcal{G}(\overline{z}_N(t)) \to \mathcal{G}(z(t)) \tag{5.1}$$

is not available in the present case of non-convex and non-smooth energy functionals. On the other hand, as explained in Remark 4.7, an energy-dissipation balance being true for the limit would allow it to conclude the convergence result (5.1). In turn, this would ensure also the strong/strict convergence of the mutual recovery sequence (4.32) and thus lead to a limit semistability inequality (4.11) featuring $\mathrm{ERR} = 0$, see Remark 4.9. The robustness of the method for varying triangulations observed in Fig. 5.4 may be an indication that (5.1) holds true here, but cannot be confirmed in general. A further interpretation of the tendency to a balance observed in Figures 5.2 and 5.5 is that this balance is provided in the limit by that of a *parametrized Balanced Viscosity solution*, cf. also Remark 4.10. In other words, the solutions obtained by Algorithm 3.5 are confirmed in the benchmark experiments to be semistable energetic solutions in the sense of Theorem 4.2, but show additional features better than predicted by this theory. Heuristically, these additional features may be explained within the concept of *parametrized Balanced Viscosity solutions*. In conclusion, the results motivate further investigations whether Algorithm 3.5 in general approximates *parametrized Balanced Viscosity solutions*.

References

[AB19] S. Almi and S. Belz. Consistent finite-dimensional approximation of phase-field models of fracture. *Ann. Mat. Pura Appl.*, 198(4):1191–1225, 2019.

[ABN18] S. Almi, S. Belz, and M. Negri. Convergence of discrete and continuous unilateral flows for Ambrosio-Tortorelli energies and application to mechanics. *ESAIM M2AN*, 53(2):659–699, 2018.

[ACFK02] J. Alberty, C. Carstensen, S. A. Funken, and R. Klose. Matlab implementation of the finite element method in elasticity. *Computing*, 69:239–263, 2002.

[AFP05] L. Ambrosio, N. Fusco, and D. Pallara. *Functions of Bounded Variation and Free Discontinuity Problems*. Oxford University Press, 2005.

[AN19] S. Almi and M. Negri. Analysis of staggered evolutions for nonlinear energies in phase field fracture. *Arch. Ration. Mech. Anal.*, 2019.

[Bar12] S. Bartels. Total variation minimization with finite elements: convergence and iterative solution. *SIAM J. Numer. Anal.*, pages 1162–1180, 2012.

[Bar15] S. Bartels. *Numerical Methods for Nonlinear Partial Differential Equations*. Springer, Heidelberg, 2015.

[BKKŠ09] J. Brandts, S. Korotov, M. Křížek, and J. Šolc. On nonobtuse simplicial partitions. *SIAM Rev.*, 51(2):317–335, 2009.

[BM17] S. Bartels and M. Milicevic. Alternating direction method of multipliers with variable step sizes. 2017.

[BMT18] S. Bartels, M. Milicevic, and M. Thomas. Numerical approach to a model for quasistatic damage with spatial BV-regularization. In E. Rocca, U. Stefanelli, and L. Truskinovsky, editors, *Proceedings of the INdAM-ISIMM Workshop on Trends on Applications of Mathematics to Mechanics, Rome, Italy, September 2016*, volume 27, pages 179–203. Springer International Publishing, Cham, 2018.

[BS08] S. C. Brenner and L. R. Scott. *The Mathematical Theory of Finite Element Methods*. Springer, New York, 2008.

[BT09] A. Beck and M. Teboulle. Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems. *IEEE Transactions on Image Processing*, 18:2419–2434, 2009.

[Cha04] A. Chambolle. An algorithm for total variation minimization and applications. *J. Math. Imaging Vis.*, 20:89–97, 2004.

[Cia02] P. G. Ciarlet. *The Finite Element Method for Elliptic Problems*, volume 40 of *Classics in Applied Mathematics*. SIAM, Philadelphia, 2002.

[Clé75] P. Clément. Approximation by finite element functions using local regularization. *Revue française d'automatique, informatique, recherche opérationelle. Analyse numérique*, 9(2):77–84, 1975.

[CP11] A. Chambolle and T. Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. *J. Math. Imaging Vis.*, 40:120–145, 2011.

[FG83] M. Fortin and R. Glowinski. *Augmented Lagrangian Methods*. North-Holland Publishing Co., Amsterdam, 1983.

[Glo84] R. Glowinski. *Numerical Methods for Nonlinear Variational Problems*. Springer, New York, 1984.

[GLT89] R. Glowinski and P. Le Tallec. *Augmented Lagrangians and Operator-Splitting Methods in Nonlinear Mechanics*. IAM, Philadelphia, 1989.

[GM75] R. Glowinski and A. Marroco. Sur l'approximation par éléments finis d'ordre un, et la résolution, par pénalisation-dualité d'une classe de problèmes de dirichlet non linéaires. *Revue française d'automatique, informatique, recherche opérationelle. Analyse numérique*, 9:41–76, 1975.

[GM76] D. Gabay and B. Mercier. A dual algorithm for the solution of nonlinear variational problems via finite element approximation. *Comp. & Maths. with Appls.*, 2:17–40, 1976.

[GO09] T. Goldstein and S. Osher. The split Bregman method for L1 regularized problems. *SIAM J. Imaging Sci.*, 2:323–343, 2009.

[GOSB14] T. Goldstein, B. O'Donoghue, S. Setzer, and R. Baraniuk. Fast alternating direction optimization methods. *SIAM J. Imaging Sci.*, 7:1588–1623, 2014.

[HLP34] G.H. Hardy, J.E. Littlewood, and G. Pólya. *Inequalities*. Cambridge University Press, 1934.

[HRH14] M. Hintermüller, C. N. Rautenberg, and J. Hahn. Functional-analytic and numerical issues in splitting methods for total variation-based image reconstruction. *Inverse Problems*, 30(5):055014, 2014.

[HRR17] M. Hintermüller, C. N. Rautenberg, and S. Rösel. Density of convex intersections and applications. *Proc. Roy. Soc. Edinburgh Sect. A*, 473:20160919/1–20160919/28, 2017.

[KN17] D. Knees and M. Negri. Convergence of alternate minimization schemes for phase-field fracture and damage. *Math. Models Methods Appl. Sci.*, 27(9):1743–1794, 2017.

[KRZ13] D. Knees, R. Rossi, and C. Zanini. A vanishing viscosity approach to a rate-independent damage model. *Math. Models Methods Appl. Sci.*, 23(04):565–616, 2013.

[LM79] P. L. Lions and B. Mercier. Splitting algorithms for the sum of two nonlinear operators. *SIAM J. Numer. Anal.*, 16:964–979, 1979.

[MM72] M. Marcus and V.J. Mizel. Absolute continuity on tracks and mappings of Sobolev spaces. *Arch. Rational Mech. Anal.*, 45:294–320, 1972.

[MM79] M. Marcus and V.J. Mizel. Every superposition operator mapping one Sobolev space into another is continuous. *J. Functional Analysis*, 33:217–229, 1979.

[MR15] A. Mielke and T. Roubíček. *Rate-independent Systems: Theory and Application*, volume 193 of *Applied Mathematical Sciences*. Springer, 2015.

[MRS12] A. Mielke, R. Rossi, and G. Savaré. BV solutions and viscosity approximations of rate-independent systems. *ESAIM Control Optim. Calc. Var.*, 18(1):36–80, 2012.

[PS04] P.-O. Persson and G. Strang. A simple mesh generator in `matlab`. *SIAM Review*, 42:329–345, 2004.

[Roc76] R. T. Rockafellar. Monotone operators and the proximal point algorithm. *SIAM J. Control. Optim.*, 14:877–898, 1976.

[ROF92] L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60:259–268, 1992.

[RT15] R. Rossi and M. Thomas. From an adhesive to a brittle delamination model in thermo-visco-elasticity. *ESAIM Control Optim. Calc. Var.*, 21:1–59, 2015.

[Tho13] M. Thomas. Quasistatic damage evolution with spatial BV-regularization. *Discrete Contin. Dyn. Syst. Ser. S*, 6:235–255, 2013.

[TM10] M. Thomas and A. Mielke. Damage of nonlinearly elastic materials at small strain: existence and regularity results. *Zeit. angew. Math. Mech.*, 90(2):88–112, 2010.

[WT10] C. Wu and X.-C. Tai. Augmented Lagrangian method, dual methods, and split Bregman iteration for ROF, vectorial TV, and higher order models. *SIAM J. Imaging Sci.*, 3:300–339, 2010.