No. 627 March 2020

Optimal control of perfect plasticity
Part II: Displacement tracking

C. Meyer, S. Walther

ISSN: 2190-1767

OPTIMAL CONTROL OF PERFECT PLASTICITY PART II: DISPLACEMENT TRACKING[*]

CHRISTIAN MEYER[†] AND STEPHAN WALTHER[†]

Abstract. The paper is concerned with an optimal control problem governed by the rate-independent system of quasi-static perfect elasto-plasticity. The objective is to optimize the displacement field in the domain occupied by the body by means of prescribed Dirichlet boundary data, which serve as control variables. The arising optimization problem is nonsmooth for several reasons, in particular, since the control-to-state mapping is not single-valued. We therefore apply a Yosida regularization to obtain a single-valued control-to-state operator. Beside the existence of optimal solutions, their approximation by means of this regularization approach is the main subject of this work. It turns out that a so-called reverse approximation guaranteeing the existence of a suitable recovery sequence can only be shown under an additional smoothness assumption on at least one optimal solution.

Key words. Optimal control of variational inequalities, perfect plasticity, rate-independent systems, Yosida regularization, reverse approximation

AMS subject classifications. 49J20, 49J40, 74C05

1. Introduction. In this paper, we investigate the following optimal control problem governed by the equations of *quasi-static perfect plasticity* at small strain:

$$
\text{(P)} \qquad \left\{
\begin{aligned}
&\min \quad J(u, u_D) := \Psi(u) + \frac{\alpha}{2}\|u_D\|^2_{H^1(0,T;H^2(\Omega;\mathbb{R}^n))} \\
&\text{s.t.} \qquad -\operatorname{div}\sigma = 0 && \text{in } \Omega,\\
&\qquad\qquad \sigma = \mathbb{C}(\nabla^s u - z) && \text{in } \Omega,\\
&\qquad\qquad \dot z \in \partial I_{\mathcal{K}(\Omega)}(\sigma) && \text{in } \Omega,\\
&\qquad\qquad u = u_D && \text{on } \Gamma_D,\\
&\qquad\qquad \sigma\nu = 0 && \text{on } \Gamma_N,\\
&\qquad\qquad u(0) = u_0, \quad \sigma(0) = \sigma_0 && \text{in } \Omega,\\
&\text{and} \qquad u_D(0) = u_0 && \text{on } \Gamma_D.
\end{aligned}
\right.
$$

Herein, $u : (0,T)\times\Omega \to \mathbb{R}^n$, $n = 2, 3$, is the displacement field, while $\sigma, z : (0,T)\times\Omega \to \mathbb{R}^{n\times n}$ are the stress tensor and the plastic strain. The boundary of Ω is split in two disjoint parts Γ_D and Γ_N with outward unit normal ν. Moreover, $\mathbb{C}$ is the elasticity tensor and $\mathcal{K}(\Omega)$ denotes the set of feasible stresses. The initial data u_0 and σ_0 are given and fixed. The Dirichlet data u_D represent the control variable and $\alpha > 0$ a fixed Tikhonov regularization parameter. The objective Ψ only contains the displacement field. Objectives involving the stress are considered in a companion paper [21]. This is the reason for calling (P) *displacement tracking problem*. A mathematically rigorous version of (P) involving the function spaces and a rigorous notion of solutions for the state equation will be formulated in section 3 and 4 below. The precise assumptions

[*]Submitted to the editors DATE.

Funding: This research was supported by the German Research Foundation (DFG) under grant number ME 3281/9-1 within the priority program Non-smooth and Complementarity-based Distributed Parameter Systems: Simulation and Hierarchical Optimization (SPP 1962).

[†]TU Dortmund, Faculty of Mathematics, Vogelpothsweg 87, 44227 Dortmund, Germany (christian2.meyer@tu-dortmund.de, http://www.mathematik.tu-dortmund.de/de/personen/person/Christian+Meyer.html, stephan.walther@tu-dortmund.de, http://www.mathematik.tu-dortmund.de/de/personen/person/Stephan+Walther.html).

on the data are given in section 2. Regarding to a more detailed description of the plasticity model, we refer to [25] and the references therein.

Some words concerning our choice of the control variable are in order: In general, Dirichlet control problems provide particular difficulties due to regularity issues, when control functions in $L^2(\partial\Omega)$ are considered, see e.g. [18]. Nonetheless, we consider the Dirichlet displacement as control variables instead of distributed loads or forces on the Neumann boundary due to the *safe load condition.* It is well known that the existence of solutions for the perfect plasticity system can only be shown under this additional condition (see e.g. [30, 8]), which would lead to rather complex control constraints and it is a completely open question how to incorporate these constraints in the analysis of (P). For this reason, we focus on the Dirichlet control problem. A possible realization of these controls by means of an additional linear elasticity equation avoiding the H^2-norm in the objective is elaborated in the companion paper [21].

Beside the safe-load condition, problem (P) exhibits several additional particular challenges. First of all, it is obviously nonsmooth due to the convex subdifferential appearing in the state equation. Moreover, the state equation is in general not uniquely solvable and its solutions significantly lack regularity, see [30, 8]. Therefore, there is no single-valued control-to-state mapping and (P) should rather be regarded as an optimization problem in Banach space rather than an optimal control problem. Beside the existence of optimal solutions, our main goal is to approximate (P) via replacing $\partial I_{\mathcal{K}(\Omega)}$ by its Yosida regularization. This is of course a classical procedure and, in order to show that the approximation works, i.e., that optimal solutions of the regularized problems converge to solutions of (P) (in a certain topology), the following steps have to be performed:

1. The existence of (weak) accumulation points of sequences of optimal solutions of the regularized problems have to be verified.
2. Weak limits have to be feasible for the original problem (P).
3. In order to show the optimality of the weak limit, one has to construct a *recovery sequence* for at least one optimal solution of the original problem.

The last item is also known as *reverse approximation* and might become a challenging task in the context of optimization of rate-independent systems, see [22]. This also happens to be the case here: In contrast to the perfect plasticity system, its regularized counterpart admits a unique solution with full regularity. It is therefore very unlikely that one can approximate *every* solution of the perfect plasticity system by means of regularization and indeed, as classical examples demonstrate, this is in fact not true, see e.g. [30] and Example 3.10 below. However, in the context of optimal control and optimization, respectively, we have the control as an additional variable at hand and, in order to construct a recovery sequence, we have to find a sequence of *tuples of state and control* feasible for the regularized problems so that the associated objective function values converge to the optimal value of (P). This leads to much more flexibility in the construction of recovery sequences, provided that the set of controls is sufficiently rich. Unfortunately, this is not the case for our Dirichlet control and we need an additional control variable in terms of *distributed loads* for the construction of a recovery sequence. The idea is thus to introduce an additional load in the balance of momentum of the regularized problems and to drive this load to zero for vanishing regularization parameter. Our regularization procedure therefore does not only replace the convex subdifferential by its Yosida regularization, but also introduces a new additional control variable. To the best of our knowledge, this is a completely new idea.

Nevertheless, even with this additional control variable, we are only able to construct a recovery sequence under a fairly restrictive assumption. This assumption is caused by additional smoothness constraints as part of the regularized optimal control problems, which in turn are needed to pass to the limit in the regularized plasticity system, when the regularization parameter is driven to zero. If we assume that at least one optimal solution of the original (i.e., unregularized) optimization problem admits an admittedly high regularity, then we are able to construct a recovery sequence for this particular solution, which meets the smoothness constraints and is therefore feasible for the regularized optimal control problems. We thus obtain the desired approximation result under the assumption that there exists at least one "smooth" solution of (P).

Let us put our work into perspective: Quasi-static perfect plasticity is a rate-independent system. Optimization and optimal control of such systems have been considered by various authors and we only refer to [4, 5, 1, 6, 7, 29, 24, 2, 14] and the references therein. Albeit still nonsmooth, optimization problems of this type substantially simplify, if the energy underlying the rate-independent system is uniformly convex. In quasi-static plasticity, this is the case, if hardening is present. In this case, the plasticity system admits a unique solution in the energy space, which makes the construction of recovery sequences almost trivial. Nevertheless, the derivation of optimality conditions is still an intricate issue, see [32, 33, 34]. While all contributions mentioned so far deal with uniformly convex energies, the literature becomes rather scarce, when it comes to energies that lack strict convexity. In [26, 28, 11, 10] the existence of optimal solutions for problems with non-convex energies are shown. To the best of our knowledge, the approximation of such problems has only been investigated in [22, 27], where a time-discretization instead of a regularization is considered. The approximation via discretization can however be hardly compared to our situation, since the discrete rate-independent systems are still not uniquely solvable so that there is still no (discrete) control-to-state map in contrast to the regularized setting. Therefore, the discrete optimization problems are still all but straight forward to solve, whereas the regularized optimal control problems are amenable for standard adjoint-based optimization methods.

The paper is organized as follows: After introducing our notation and standard assumptions in section 2, we introduce a rigorous notion of solution to the perfect plasticity system and recall the known results concerning the existence of solutions and the lack of uniqueness in section 3. Then, section 4 is devoted to the existence of at least one (globally) optimal solution of (P). In section 5, we lay the foundations for our reverse approximation argument for the construction of a recovery sequence, which is a basic ingredient for our main result in Theorem 6.3. The last section 6 covers this result and shows that solutions of (P) can indeed be approximated via Yosida regularization provided the mentioned regularity assumption is fulfilled.

2. Notation and Standing Assumptions.

We start with a short introduction in the notation used throughout the paper and in parallel list our standing assumptions. The latter are tacitly assumed for the rest of the paper without mentioning them every time.

General notation. Given two vector spaces X and Y, we denote the space of linear and continuous functions from X into Y by $\mathcal{L}(X, Y)$. If $X = Y$, we simply write $\mathcal{L}(X)$. The dual space of X is denoted by $X^* = \mathcal{L}(X, \mathbb{R})$. If H is a Hilbert space, we denote its scalarproduct by $(\cdot, \cdot)_H$. For the whole paper, we fix the final time $T > 0$. To shorten the notation, Bochner-spaces are abbreviated by $L^p(X) :=$

$L^p(0,T;X)$, $W^{1,p}(X) := W^{1,p}(0,T;X)$ ($p \in [1,\infty]$), and $C(X) := C([0,T];X)$. Note that functions in $C(X)$ are continuous on the whole time interval. When $G \in \mathcal{L}(X;Y)$ is a linear and continuous operator, we can define an operator in $\mathcal{L}(L^p(X);L^p(Y))$ by $G(u)(t) := G(u(t))$ for all $u \in L^p(X)$ and for almost all $t \in [0,T]$, we denote this operator also by G, that is, $G \in \mathcal{L}(L^p(X);L^p(Y))$, and analog for Bochner-Sobolev spaces, i.e., $G \in \mathcal{L}(W^{1,p}(X);W^{1,p}(Y))$.

Given a coercive operator $G \in \mathcal{L}(H)$ in a Hilbert space H, we denote its coercivity constant by γ_G, i.e., $(Gh,h)_H \geq \gamma_G \|h\|_H^2$ for all $h \in H$. With this operator we can define a new scalar product, which induces an equivalent norm, by $H \times H \ni (h_1,h_2) \mapsto (Gh_1,h_2)_H \in \mathbb{R}$. We denote the Hilbert space equipped with this scalar product by H_G, that is $(h_1,h_2)_{H_G} = (Gh_1,h_2)_H$ for all $h_1,h_2 \in H$.

If $p \in [1,\infty]$, then we denote its conjugate exponent by p', that is $\frac{1}{p} + \frac{1}{p'} = 1$. Furthermore, $c, C > 0$ are generic constants.

Matrices. Given a matrix $\tau \in \mathbb{R}^{n\times n}$, we define its deviatoric (i.e., trace-free) part as

$$\tau^D := \tau - \tfrac{1}{n}\,\mathrm{tr}(\tau) I$$

and use the same notation for matrix-valued functions. The Frobenius norm is denoted by $|A|_F^2 = \sum_{i,j=1}^n A_{ij}^2$ for $A \in \mathbb{R}^{n\times n}$ and for the associated scalar product, we write $A : B = \sum_{i,j=1}^n A_{ij}B_{ij}$, $A, B \in \mathbb{R}^{n\times n}$. By $\mathbb{R}^{n\times n}_{\mathrm{sym}}$, we denote the space of symmetric matrices.

Domain. The domain $\Omega \subset \mathbb{R}^n$, $n \in \mathbb{N}$, $n \geq 2$, is bounded of class C^1. The boundary consists of two disjoint measurable parts Γ_N and Γ_D such that $\Gamma = \Gamma_N \cup \Gamma_D$. While Γ_N is a relatively open subset, Γ_D is a relatively closed. We moreover suppose that Γ_D has a nonempty relative interior. In addition, the set $\Omega \cup \Gamma_N$ is regular in the sense of Gröger, cf. [15]. Throughout the article, $\nu : \partial\Omega \to \mathbb{R}^n$ denotes the outward unit normal vector.

Thanks to the regularity of Ω, the harmonic extension $\mathfrak{E}$ maps $C^1(\Gamma)$ to $W^{1,p}(\Omega)$ for some $p > n$. Moreover, the maximum principle implies that

$$\|\mathfrak{E}\varphi\|_{L^\infty(\Omega)} \leq \|\varphi\|_{L^\infty(\Gamma)} \quad \forall\, \varphi \in C^1(\Gamma). \tag{2.1}$$

Remark 2.1. The C^1-regularity of Ω and its boundary, respectively, is required for the trace theorem and the formula of integration by parts for BD-functions in [31, Chap. II, Theorem 2.1], which will be used several times throughout the paper. In [12, Section 6], it is claimed that this formula integration by parts also holds in Lipschitz domains, but no proof is provided. Since the minimal regularity of the boundary is not in the focus of this paper and would go beyond the scope of our work, we restrict to domains of class C^1.

Spaces. Throughout the paper, by $L^p(\Omega;M)$ we denote Lebesgue spaces with values in M, where $p \in [1,\infty]$ and M is a finite dimensional space. To shorten notation, we abbreviate

$$\mathbf{L}^p(\Omega) := L^p(\Omega;\mathbb{R}^n) \quad \text{and} \quad \mathbb{L}^p(\Omega) := L^p(\Omega;\mathbb{R}^{n\times n}_{\mathrm{sym}}).$$

Given $s \in \mathbb{N}$ and $p \in [1,\infty]$, the Sobolev spaces of vector- resp. tensor-valued functions are denoted by

$$\begin{aligned} \mathbf{W}^{s,p}(\Omega) &:= W^{s,p}(\Omega;\mathbb{R}^n), & \mathbf{H}^s(\Omega) &:= \mathbf{W}^{s,2}(\Omega),\\ \mathbb{W}^{s,p}(\Omega) &:= W^{s,p}(\Omega;\mathbb{R}^{n\times n}_{\mathrm{sym}}), & \mathbb{H}^s(\Omega) &:= \mathbb{W}^{s,2}(\Omega). \end{aligned}$$

Furthermore, set

$$\mathbf{W}^{1,p}_D(\Omega) := \overline{\{\psi|_\Omega : \psi \in C^\infty_c(\mathbb{R}^n;\mathbb{R}^n),\ \operatorname{supp}(\psi)\cap\Gamma_D = \emptyset\}}^{\mathbf{W}^{1,p}(\Omega)} \tag{2.2}$$

and define $\mathbf{H}^1_D(\Omega)$ analogously. The dual of $\mathbf{H}^1_D(\Omega)$ is denoted by $\mathbf{H}^{-1}_D(\Omega)$. The space of *bounded deformation* is abbreviated by

$$\mathrm{BD}(\Omega) := \{u \in \mathbf{L}^1(\Omega) : \tfrac{1}{2}(\partial_i u_j + \partial_j u_i) \in \mathfrak{M}(\Omega)\ \forall\, i,j = 1,...,n\},$$

where $\mathfrak{M}(\Omega)$ denotes the space of regular Borel measures on Ω and the (partial) derivatives are of course understood in a distributional sense. Equipped with the norm

$$\|u\|_{\mathrm{BD}(\Omega)} := \|u\|_{\mathbf{L}^1(\Omega)} + \sum_{i,j=1}^n \tfrac{1}{2}\|\partial_i u_j + \partial_j u_i\|_{\mathfrak{M}(\Omega)},$$

it becomes a Banach space.

Coefficients. The elasticity tensor satisfies $\mathbb{C} \in \mathcal{L}(\mathbb{R}^{d\times d}_{\mathrm{sym}})$ and is symmetric and coercive. In addition we set $\mathbb{A} := \mathbb{C}^{-1}$ and note that $\mathbb{A}$ is symmetric and coercive, too. Let us note that $\mathbb{C}$ could also depend on space, however, to keep the discussion concise, we restrict ourselves to constant elasticity tensors.

Yield condition. The set defining the yield condition is denoted by $K \subset \mathbb{R}^{n\times n}_{\mathrm{sym}}$ and is closed and convex and there exists $0 < \rho < R$ such that

$$\overline{B_{\mathbb{R}^{n\times n}}(0;\varrho)} \subset K \subset \overline{B_{\mathbb{R}^{n\times n}}(0;R)}. \tag{2.3}$$

Given this set, we define the *set of admissible stresses* as

$$\mathcal{K}(\Omega) := \{\tau \in \mathbb{L}^2(\Omega) : \tau^D(x) \in K \text{ f.a.a. } x \in \Omega\}.$$

Remark 2.2. The boundedness of the set K is not really needed for our analysis. It is only required for the formula of integration by parts in (3.9), which we only need to compare our notion of solution to the one in [8]. Nevertheless, we kept the boundedness assumption on the set K, since it is fulfilled in all practically relevant examples such as e.g. the von Mises or the Tresca yield condition.

Operators. Throughout the paper, $\nabla^s := \frac{1}{2}(\nabla + \nabla^\top) : \mathbf{W}^{1,p}(\Omega) \to \mathbb{L}^p(\Omega)$ denotes the linearized strain. Its restriction to $\mathbf{W}^{1,p}_D(\Omega)$ is denoted by the same symbol and, for the adjoint of this restriction, we write $-\operatorname{div} := (\nabla^s)^* : \mathbb{L}^{p'}(\Omega) \to \mathbf{W}^{1,p}_D(\Omega)^*$.

Let $\mathcal{K} \subset \mathbb{L}^2(\Omega)$ be a closed and convex set. We denote the indicator function by

$$I_\mathcal{K} : \mathbb{L}^2(\Omega) \to \{0,\infty\}, \qquad \tau \mapsto \begin{cases} 0, & \tau \in \mathcal{K}, \\ \infty, & \tau \notin \mathcal{K}. \end{cases}$$

By $\partial I_\mathcal{K} : \mathbb{L}^2(\Omega) \to 2^{\mathbb{L}^2(\Omega)}$ we denote the subdifferential of the indicator function. For $\lambda > 0$, the Yosida regularization is given by

$$I_\lambda : \mathbb{L}^2(\Omega) \to \mathbb{R}, \qquad \tau \mapsto \frac{1}{2\lambda}\|\tau - \pi_\mathcal{K}(\tau)\|^2_{\mathbb{L}^2(\Omega)}, \tag{2.4}$$

where $\pi_{\mathcal{K}}$ is the projection onto $\mathcal{K}$ in $\mathbb{L}^2(\Omega)$, and its Fréchet derivative is

$$\partial I_\lambda(\tau) = \frac{1}{\lambda}(\tau - \pi_{\mathcal{K}}(\tau)).$$

When $\lambda = 0$ we define $I_\lambda = I_0 := I_{\mathcal{K}}$. For a sequence $\{\lambda_n\}_{n\in\mathbb{N}} \subset (0,\infty)$ we abbreviate $I_n := I_{\lambda_n}$.

Initial data. For the initial stress field σ_0, we assume that $\sigma_0 \in \mathbf{W}^{1,p}(\Omega)$ with some $p > n$. Moreover, σ_0 satisfies the equilibrium condition, i.e., $\operatorname{div}\sigma_0 = 0$ a.e. in Ω, and the yield condition, i.e., $\sigma_0 \in \mathcal{K}(\Omega)$. The initial displacement u_0 is supposed to be an element of $\mathbf{H}^2(\Omega)$ and we require $\operatorname{tr}(\nabla^s u_0 - \mathbb{A}\sigma_0) = 0$ a.e. in Ω in order to obtain a purely deviatoric initial plastic strain.

Remark 2.3. The high regularity of u_0 is just needed to ensure that the feasible set of (P) is nonempty. For the mere discussion of the state system, this is not necessary. The same holds for the assumption $\sigma_0 \in \mathbf{W}^{1,p}(\Omega)$, which will be needed to construct a recovery sequence for the optimal control problem.

Optimization Problem. The Tikhonov parameter α is a positive constant and Ψ is a functional that is bounded from below and satisfies a certain lower semicontinuity assumption w.r.t. weak convergence in the displacement space, which will be made precise in section 4 below, see (4.8).

3. State Equation.

We start our investigations with the analysis of the state system and recall some known results concerning quasi-static perfect plasticity. Already since the pioneering work of Suquet [30], it is well known that a precise definition of a solution to the system of perfect plasticity is all but straight forward, since a solution of the system in its "natural" form (below termed strong solution) does in general not exist due to a lack of regularity of the displacement and the plastic strain, respectively. We start with the definition of the function spaces already indicating this lack of regularity:

Definition 3.1 (State spaces).

1. Stress space:

$$\Sigma(\Omega) := \{\tau \in \mathbb{L}^2(\Omega) : \operatorname{div}\tau \in \mathbf{L}^n(\Omega),\ \tau^D \in \mathbb{L}^\infty(\Omega)\}$$

2. Displacement space:

$$\mathcal{U} := \{u \in H^1(\mathbf{L}^{\frac{n}{n-1}}(\Omega)) : \nabla^s \dot{u} \in L^2_w(\mathfrak{M}(\Omega;\mathbb{R}^{n\times n}_{\text{sym}}))\}.$$

Herein, $L^2_w(\mathfrak{M}(\Omega;\mathbb{R}^{n\times n}_{\text{sym}}))$ is the space of weakly measurable functions with values in $\mathfrak{M}(\Omega;\mathbb{R}^{n\times n}_{\text{sym}})$, for which $t \mapsto \|\mu(t)\|_{\mathfrak{M}}$ is an element of $L^2(0,T;\mathbb{R})$. For the definition of weak measurability, we refer to [9, Section 8].

We say that a sequence $\{u_n\} \subset \mathcal{U}$ converges weakly in $\mathcal{U}$ to u and write $u_n \rightharpoonup u$ in $\mathcal{U}$, iff

$$u_n \rightharpoonup u \ \text{ in } H^1(\mathbf{L}^{\frac{n}{n-1}}(\Omega)), \quad \nabla^s \dot{u}_n \rightharpoonup^* \nabla^s \dot{u} \ \text{ in } L^2_w(\mathfrak{M}(\Omega;\mathbb{R}^{n\times n}_{\text{sym}})). \tag{3.1}$$

Note that, by [9, Theorem 8.20.3], $L^2_w(\mathfrak{M}(\Omega;\mathbb{R}^{n\times n}_{\text{sym}})) = L^2(C_0(\Omega;\mathbb{R}^{n\times n}_{\text{sym}}))^*$, which gives a meaning to the weak-$*$ convergence in (3.1).

Remark 3.2. Unfortunately, $\mathrm{BD}(\Omega)$ does not admit the Radon-Nikodým property and therefore weak measurability does not imply Bochner-measurability.

Definition 3.3 (Equilibrium condition). *We define the set of stresses which fulfill the* equilibrium condition *as*

$$\mathcal{E}(\Omega) := \ker(\operatorname{div}) = \{\tau \in \mathbb{L}^2(\Omega) : (\tau, \nabla^s \varphi)_{\mathbb{L}^2(\Omega)} = 0 \ \ \forall \varphi \in \mathbf{H}^1_D(\Omega)\}.$$

Note that $\sigma \in \mathcal{E}(\Omega) \cap \mathcal{K}(\Omega)$ *implies* $\sigma \in \Sigma(\Omega)$.

With the above definitions at hand, we can now define a hierarchy of three different solutions:

Definition 3.4 (Notions of solutions). *Let* $u_D \in H^1(\mathbf{H}^1(\Omega))$ *with* $u_D(0) = u_0$ *a.e. on* Γ_D *be given. Then we define the following notions of solutions to the perfect plasticity system:*

1. Reduced solution: *A function* $\sigma \in H^1(\mathbb{L}^2(\Omega))$ *is called* reduced solution *of the state equation, if, for almost all* $t \in (0, T)$*, the following holds true:*
   - *Equilibrium and yield condition:*
   $$\sigma(t) \in \mathcal{E}(\Omega) \cap \mathcal{K}(\Omega), \tag{3.2a}$$
   - *Reduced flow rule inequality:*
   $$\int_\Omega \big(\mathbb{A}\dot\sigma(t) - \nabla^s \dot u_D(t)\big) : \big(\tau - \sigma(t)\big)\,\mathrm{d}x \ge 0 \quad \forall \tau \in \mathcal{E}(\Omega) \cap \mathcal{K}(\Omega), \tag{3.2b}$$
   - *Initial condition:*
   $$\sigma(0) = \sigma_0. \tag{3.2c}$$
2. Weak solution: *A tuple* $(u, \sigma) \in \mathcal{U} \times H^1(\mathbb{L}^2(\Omega))$ *is called* weak solution *of the state equation, if, for almost all* $t \in (0, T)$*, there holds*
   - *Equilibrium and yield condition:*
   $$\sigma(t) \in \mathcal{E}(\Omega) \cap \mathcal{K}(\Omega), \tag{3.3a}$$
   - *Weak flow rule inequality:*
   $$\begin{aligned} &\int_\Omega \mathbb{A}\dot\sigma(t) : \big(\tau - \sigma(t)\big)\,\mathrm{d}x + \int_\Omega \dot u(t) \cdot \operatorname{div}\big(\tau - \sigma(t)\big)\,\mathrm{d}x \\ &\qquad \ge \int_\Omega \nabla^s \dot u_D(t) : \big(\tau - \sigma(t)\big) + \dot u_D(t) \cdot \operatorname{div}\big(\tau - \sigma(t)\big)\,\mathrm{d}x \\ &\qquad\qquad \forall \tau \in \Sigma(\Omega) \cap \mathcal{K}(\Omega), \end{aligned} \tag{3.3b}$$
   - *Initial condition:*
   $$u(0) = u_0, \quad \sigma(0) = \sigma_0. \tag{3.3c}$$
3. Strong solution: *A tuple* $(u, \sigma) \in H^1(\mathbf{H}^1(\Omega)) \times H^1(\mathbb{L}^2(\Omega))$ *is called* strong solution *of the state equation, if, for almost all* $t \in (0, T)$*, there holds*
   - *Equilibrium and yield condition:*
   $$\sigma(t) \in \mathcal{E}(\Omega) \cap \mathcal{K}(\Omega), \tag{3.4a}$$

- *Strong flow rule inequality:*

$$\int_\Omega \mathbb{A}\dot\sigma(t) : \big(\tau - \sigma(t)\big)\,\mathrm{d}x + \int_\Omega \nabla^s \dot u(t) : \big(\tau - \sigma(t)\big)\,\mathrm{d}x \geq 0 \quad \forall\, \tau \in \mathcal{K}(\Omega), \tag{3.4b}$$

- *Dirichlet boundary condition:*

$$u(t) - u_D(t) \in \mathbf{H}^1_D(\Omega) \tag{3.4c}$$

- *Initial condition:*

$$u(0) = u_0, \quad \sigma(0) = \sigma_0. \tag{3.4d}$$

Some words concerning this definition are in order. First, let us shortly investigate the relationship between the three different solution concepts. By restricting the test functions in (3.3b) to functions in $\mathcal{E}(\Omega)$, one immediately observes that every weak solution is also a reduced solution. Moreover, by integration by parts, it is evident that (3.4b) and (3.4c) imply (3.3b). On the other hand, if a weak solution satisfies $u \in H^1(\mathbf{H}^1(\Omega))$ and the Dirichlet boundary conditions in (3.4c), then integration by parts yields (3.4b), provided that $\Sigma(\Omega) \cap \mathcal{K}(\Omega)$ is dense in $\mathcal{K}(\Omega)$, which is a direct consequence of Lemma A.1 proven in the appendix. Thus, we have the following relations between the three different solution concepts:

Corollary 3.5 (Relations between the solution concepts).
1. *If (u,σ) is a weak solution, then σ is automatically a reduced solution.*
2. *A weak solution (u,σ) is a strong solution, if and only if $u \in H^1(\mathbf{H}^1(\Omega))$ and $(u - u_D)(t) \in \mathbf{H}^1_D(\Omega)$ for all $t \in [0,T]$.*

One may further ask why no Dirichlet boundary conditions appear in the definition of a weak solution. In fact, from a mechanical point of view, it is reasonable that no boundary conditions are imposed, since plastic slips may well develop on the Dirichlet part on the boundary, too, in form of tangential jumps of the displacement perpendicular to the outward normal ν. This observation is implicitly contained in the above definition as demonstrated in [8, Theorem 6.1]. For convenience of the reader, we shortly sketch the underlying arguments. To this end, suppose that a weak solution is given and let us define the *plastic strain* $z \in L^2_w(\mathfrak{M}(\Omega \cup \Gamma_D; \mathbb{R}^{n\times n}_{\mathrm{sym}}))$ by

$$z\lfloor_\Omega := \nabla^s(u) - \mathbb{A}\sigma\,\mathrm{d}x, \quad z\lfloor_{\Gamma_D} := (u - u_D) \odot \nu\mathcal{H}^{n-1}, \tag{3.5}$$

where $\odot$ refers to the symmetrized dyadic product, i.e., $a \odot b = 1/2(a_i b_j + a_j b_i)^n_{i,j=1}$ for $a, b \in \mathbb{R}^n$. Note that functions in $\mathrm{BD}(\Omega)$ admit traces in $L^1(\partial\Omega;\mathbb{R}^n)$ (see e.g. [31, Chap. II, Thm. 2.1]) so that $z\lfloor_{\Gamma_D}$ is well defined. According to [8, Lemma 5.5], these equations carry over to the time derivatives for almost all $t \in (0,T)$, i.e.,

$$\dot z\lfloor_\Omega := \nabla^s(\dot u) - \mathbb{A}\dot\sigma\,\mathrm{d}x, \quad \dot z\lfloor_{\Gamma_D} := (\dot u - \dot u_D) \odot \nu\mathcal{H}^{n-1}. \tag{3.6}$$

Let us prove that the trace of p vanishes. For this purpose, we need the following formula of integration by parts:

Lemma 3.6 ([31, Chap. II, Thm. 2.1]). *For every $v \in \mathrm{BD}(\Omega)$ and every $\varphi \in W^{1,p}(\Omega)$, $p > n$, there holds*

$$\int_\Omega \tfrac12 (v_i \partial_j \varphi + v_j \partial_i \varphi)\,\mathrm{d}x + \int_\Omega \varphi\,\mathrm{d}(\nabla^s v)_{ij} = \int_{\partial\Omega} \varphi\,\tfrac12 (v_i \nu_j + v_j \nu_i)\,\mathrm{d}s \tag{3.7}$$

for all $i, j = 1, ..., n$.

Remark 3.7. The result in [31] is only stated for test functions in $C^1(\bar\Omega)$. However, the embeddings $\mathrm{BD}(\Omega) \hookrightarrow \mathbf{L}^{\frac{n}{n-1}}(\Omega)$ and $W^{1,p}(\Omega) \hookrightarrow C(\bar\Omega)$, $p > n$, along with the trace theorem for BD-functions and the density of $C^1(\bar\Omega)$ in $W^{1,p}(\Omega)$ imply that the integration by parts also holds for test functions in $W^{1,p}(\Omega)$.

Now, let $\varphi \in C_c^\infty(\Omega)$ be arbitrary. Then, since $\mathcal{K}(\Omega)$ just acts on the deviatoric part, $\varphi\,\delta_{ij} + \sigma_{ij}(t) \in \Sigma(\Omega) \cap \mathcal{K}(\Omega)$ for all $t \in [0, T]$ and therefore (3.3b) and the above formula of integration by parts give

$$\sum_i \Big(\int_\Omega (\mathbb{A}\dot\sigma)_{ii}\,\varphi\,\mathrm{d}x + \int_\Omega \varphi\,\mathrm{d}(\nabla^s \dot u)_{ii} \Big) = 0 \quad \forall\,\varphi \in C_c^\infty(\Omega)$$

and therefore $\operatorname{tr}\dot z\lfloor_\Omega = 0$ f.a.a. $t \in (0, T)$. Since $\operatorname{tr}(\nabla^s(u_0) - \mathbb{A}\sigma_0) = 0$, [8, Theorem 7.1] yields $\operatorname{tr} z\lfloor_\Omega = 0$ for all $t \in [0, T]$. Similarly, we choose an arbitrary test function $\psi \in C^\infty(\Gamma)$ with $\operatorname{supp}(\psi) \subset \Gamma_D$ and test (3.3b) with $\mathfrak{E}\psi\,\delta_{ij} + \sigma_{ij}(t) \in \Sigma(\Omega) \cap \mathcal{K}(\Omega)$. Note that $\mathfrak{E}\psi\,\delta_{ij} \in \Sigma(\Omega)$, since the harmonic extension maps into $W^{1,p}(\Omega)$ with $p > n$. Applying then again the formula of integration by parts implies, in view of $\operatorname{tr}\dot z\lfloor_\Omega = 0$, that

$$(\dot u - \dot u_D) \cdot \nu = 0 \quad \text{a.e. on } \Gamma_D. \tag{3.8}$$

As $u_0 = u_D(0)$ a.e. on Γ_D, this yields $(u - u_D) \cdot \nu = 0$ a.e. on Γ_D, giving in turn $\operatorname{tr} z\lfloor_{\Gamma_D} = 0$ for all $t \in [0, T]$. Now that we know that z is deviatoric, the formula of integration by parts from [8, Proposition 2.2] is applicable, which yields

$$\langle \tau^D, \dot z(t)\rangle + \int_\Omega \tau : \big(\mathbb{A}\dot\sigma(t) - \nabla^s(\dot u_D(t))\big)\,\mathrm{d}x = \int_\Omega \operatorname{div}\tau \cdot (\dot u(t) - \dot u_D(t))\,\mathrm{d}x \tag{3.9}$$

for almost all $t \in (0, T)$ and all $\tau \in \Sigma(\Omega)$. It is to be noted that the duality product $\langle \tau^D, \dot z\rangle$ has to be treated with care, since, in general, $\tau^D \notin C(\bar\Omega; \mathbb{R}^{n\times n}_{\mathrm{sym}})$, but $\dot z$ is only a measure. For a detailed and rigorous discussion of this issue, we refer to [8, Section 2.3]. Inserting (3.9) in the flow rule inequality (3.3b) then results in

$$\langle \tau^D - \sigma^D(t), \dot z(t)\rangle \geq 0 \quad \forall\,\tau \in \Sigma(\Omega) \cap \mathcal{K}(\Omega), \tag{3.10}$$

which is just the maximum plastic work inequality illustrating that z as defined in (3.5) is indeed the correct object for the plastic strain. As a byproduct, we obtain the second equation in (3.5) as boundary condition on Γ_D indicating that the Dirichlet boundary condition in (3.4c) as part of the definition of a strong solution is in general too restrictive as already mentioned above. Accordingly, a strong solution does in general not exist, while we have the following result for a weak solution:

Proposition 3.8 (Existence of weak solutions, [30, Résultat 2]). *For all $u_D \in H^1(\mathbf{H}^1(\Omega))$, there exists a weak solution in the sense of Definition 3.4.*

Proof. Using the Yosida regularization, Suquet showed in [30] the existence of a functions $\sigma \in H^1(\mathbb{L}^2(\Omega))$ and $v \in L^2_w(\mathrm{BD}(\Omega))$ so that, for almost all $t \in (0, T)$,

$$\begin{aligned} &-\operatorname{div}\sigma(t) \in \mathcal{E}(\Omega) \cap \mathcal{K}(\Omega),\\ &\int_\Omega \mathbb{A}\dot\sigma(t) : \big(\tau - \sigma(t)\big)\,\mathrm{d}x + \int_\Omega v(t) \cdot \operatorname{div}\big(\tau - \sigma(t)\big)\,\mathrm{d}x\\ &\qquad\qquad \geq \langle \dot u_D(t), (\tau - \sigma(t))\nu\rangle_{H^{1/2}(\Gamma_D), H^{-1/2}(\Gamma_D)} \quad \forall\,\tau \in \Sigma(\Omega) \cap \mathcal{K}(\Omega),\\ &\quad \sigma(0) = \sigma_0. \end{aligned} \tag{3.11}$$

Due to the continuous embedding $\mathrm{BD}(\Omega) \hookrightarrow \mathbf{L}^{\frac{n}{n-1}}(\Omega)$ (see e.g. [31, Chap. II, Theorem 2.2]) and the Radon-Nikodym property of $\mathbf{L}^{\frac{n}{n-1}}(\Omega)$, we have that $v \in L^2(\mathbf{L}^{\frac{n}{n-1}}(\Omega))$. Therefore,

$$u(t) := u_0 + \int_0^t v(r)\,\mathrm{d}r$$

is an element of $H^1(\mathbf{L}^{\frac{n}{n-1}}(\Omega))$ and satisfies the initial condition in (3.3c). Inserting this in (3.11) and integrating the right hand side by parts gives the desired flow rule inequality (3.3b). The claimed regularity of u directly follows from the regularity of $v = \dot{u}$. ∎

Remark 3.9 (Other equivalent notions of solutions). Beside the reformulation of the flow rule in terms of the maximum plastic work inequality (3.10), there are other solutions concepts, which are equivalent to the definition of a weak solution, such as the notion of a *quasi-static evolution*, which in essence corresponds to a global energetic solution in the sense of [23]. For an overview over the various notions of solutions and a rigorous proof of their equivalence, we refer to [8, Section 6]. A slightly sloppy, but very illustrating derivation of the flow rule out of the quasi-static evolution can also be found in [13].

Unfortunately, the weak solution is not unique, as the following example shows:

EXAMPLE 3.10 ([30, Section 2.1]). *We choose $\Omega = (0,1)$, $\Gamma_D = \partial\Omega = \{0,1\}$, $T = 1$, $K = [-1,1]$, $\mathbb{C} = 1$, $(\sigma_0, u_0) = 0$, and $u_D(t,x) := 2tx$. One easily verifies that the stress does only depend on the time with $\sigma(t) = 2t$ for $t \in (0, \frac{1}{2})$ and $\sigma(t) = 1$ for $t \in (\frac{1}{2}, 1)$. For the displacement one obtains $u(t,x) = 2tx$ for $(t,x) \in (0,\frac{1}{2}) \times (0,1)$ so that it is unique for $t \in (0,\frac{1}{2})$. For $t \in (\frac{1}{2}, 0)$ there are more than one solution, for example*

$$u(t,x) = 2tx, \quad \text{if } (t,x) \in (\tfrac{1}{2},1)\times(0,1),$$

$$u(t,x) = \begin{cases} \frac{2tx}{\beta} + x - \frac{x}{\beta}, & \text{if } (t,x) \in (\frac{1}{2},1)\times[0,\beta], \\ 2t + x - 1, & \text{if } (t,x) \in (\frac{1}{2},1)\times[\beta,1], \end{cases}$$

$$u(t,x) = \begin{cases} x, & \text{if } (t,x) \in (\frac{1}{2},1)\times[0,\beta], \\ \alpha t + x - \frac{\alpha}{2}, & \text{if } (t,x) \in (\frac{1}{2},1)\times[\beta,1], \end{cases}$$

where $\alpha \in [0,2]$ and $\beta \in [0,1]$ can be freely chosen. Note that the last solution just provides the minimal regularity, i.e., $\partial_x \dot{u}(t) \in \mathfrak{M}(\Omega)$.

The uniqueness of the stress field observed in the above example is a general result:

LEMMA 3.11 (Uniqueness of the stress, [17, Theorem 1], [21, Lemma 3.5]). *Assume that $\sigma_1, \sigma_2 \in H^1(\mathbb{L}^2(\Omega))$ are two reduced solutions. Then $\sigma_1 = \sigma_2$.*

Remark 3.12 (Optimal control vs. optimization). Since the displacement field as part of a weak solution is not unique in general, there is no (single-valued) control-to-state operator mapping u_D to u. Therefore, one might argue that (P) is actually no real optimal control problem. Strictly speaking, one should rather regard it as an *optimization problem* with the triple (u, σ, u_D) as optimization variables.

4. Existence of Optimal Solutions. Before we come to the main point of our analysis, which concerns the approximation of (P) by means of regularized optimal control problems, let us address the existence of optimal solutions. The proof in principle follows the classical direct method, for which we need the following boundedness and continuity results:

LEMMA 4.1 ([21, Lemma 3.6]). *Let $u_D \in H^1(\mathbf{H}^1(\Omega))$ be given and σ be the associated reduced solution. Then there holds*

$$\|\dot\sigma\|_{L^2(\mathbb{L}^2(\Omega))} \le \gamma_{\mathbb{A}}^{-1}\,\|u_D\|_{H^1(\mathbf{H}^1(\Omega))}, \tag{4.1}$$

where $\gamma_{\mathbb{A}}$ is the coercivity constant of $\mathbb{A}$. Consequently, there is a constant $C > 0$ such that $\|\sigma\|_{H^1(\mathbb{L}^2(\Omega))} \le C(\|\sigma_0\|_{\mathbb{L}^2(\Omega)} + \|u_D\|_{H^1(\mathbf{H}^1(\Omega))})$.

LEMMA 4.2 (Continuity of reduced solutions, [21, Proposition 3.10]). *Let $\{u_{D,n}\} \subset H^1(\mathbf{H}^1(\Omega))$ be a sequence such that*

$$\begin{aligned} u_{D,n} \rightharpoonup u_D \quad &\text{in } H^1(\mathbf{H}^1(\Omega)), \quad u_{D,n} \to u_D \quad \text{in } L^2(\mathbf{H}^1(\Omega)),\\ &u_{D,n}(T) \to u_D(T) \quad \text{in } \mathbf{H}^1(\Omega) \end{aligned} \tag{4.2}$$

and denote the (unique) reduced solution associated with $u_{D,n}$ by σ_n. Then $\sigma_n \rightharpoonup \sigma$ in $H^1(\mathbb{L}^2(\Omega))$, where σ is the reduced solution w.r.t. u_D.

LEMMA 4.3. *There is a constant $C > 0$, independent of u_D, such that every weak solution w.r.t. u_D fulfills*

$$\Big(\int_0^T \|\dot u(t)\|^2_{\mathrm{BD}(\Omega)}\,\mathrm{d}t\Big)^{1/2} \le C\,\|u_D\|_{H^1(\mathbf{H}^1(\Omega))}\big(1 + \|u_D\|_{H^1(\mathbf{H}^1(\Omega))}\big).$$

Proof. Let $\varphi \in C_c^\infty(\Omega)$ with $\|\varphi\|_{L^\infty(\Omega)} \le 1$ and $i, j \in \{1, ..., n\}$ be arbitrary. According to (2.3), the test function

$$(\tau_\varphi)_{ij} = (\tau_\varphi)_{ji} := -\frac{\varrho}{\sqrt{2}}\,\varphi, \quad (\tau_\varphi)_{kl} = 0 \quad \forall\,(k,l) \notin \{(i,j),(j,i)\}$$

is admissible for (3.3b). Using $\operatorname{div}\sigma = 0$, we deduce

$$\int_\Omega \varphi\,\mathrm{d}(\nabla^s \dot u)_{ij} \le \frac{\sqrt{2}}{\varrho}\Big(\int_\Omega \nabla^s \dot u_D : \sigma\,\mathrm{d}x - \int_\Omega \mathbb{A}\dot\sigma : (\tau_\varphi - \sigma)\,\mathrm{d}x\Big)$$

and consequently, since $\varphi \in C_c^\infty(\Omega)$ with $\|\varphi\|_{L^\infty(\Omega)} \le 1$ was arbitrary,

$$\begin{aligned} \|\nabla^s \dot u\|_{L^2_w(\mathfrak{M}(\Omega;\mathbb{R}^{n\times n}_{\mathrm{sym}}))} &\le C\big(\|u_D\|_{H^1(\mathbf{H}^1(\Omega))}\,\|\sigma\|_{L^\infty(\mathbb{L}^2(\Omega))}\\ &\qquad + \|\dot\sigma\|_{L^2(\mathbb{L}^2(\Omega))} + \|\dot\sigma\|_{L^2(\mathbb{L}^2(\Omega))}\|\sigma\|_{L^\infty(\mathbb{L}^2(\Omega))}\big)\\ &\le C\,\|u_D\|_{H^1(\mathbf{H}^1(\Omega))}\big(1 + \|u_D\|_{H^1(\mathbf{H}^1(\Omega))}\big), \end{aligned}$$

where we used Lemma 4.1.

Since Γ_D is assumed to have a nonempty relative interior, there is a set $\Lambda \subset \Gamma_D$ and a constant $\delta > 0$ such that Λ has positive boundary measure and $\operatorname{dist}(\Lambda, \partial\Gamma_D) \ge \delta$. By [31, Chap. II, Theorem 2.1], $\dot u(t)$ admits a trace in $\mathbf{L}^1(\Gamma)$ for almost all $t \in (0, T)$. In the following, we neglect the variable t for the sake of readability. The restriction of this trace to Λ is denoted by $\dot u|_\Lambda$. We extend $\operatorname{sign}(\dot u|_\Lambda)$ (where the sign is to be understood componentwise) to the whole boundary Γ by zero and apply convolution with a smoothing kernel to obtain a sequence of functions $\{\varphi_n\} \subset C^\infty(\Gamma; \mathbb{R}^n)$ with

$\operatorname{supp}(\varphi_n) \subset \Gamma_D$ (thanks to $\operatorname{dist}(\Lambda, \partial\Gamma_D) \geq \delta$) and $\|\varphi_n\|_{L^\infty(\Gamma;\mathbb{R}^n)} \leq 1$ for all $n \in \mathbb{N}$. Given these functions, let us define

$$(\tau_n)_{ij} = \frac{\varrho}{\sqrt{2}}\, \mathfrak{E}(\varphi_{n,i}\,\nu_j + \varphi_{n,j}\,\nu_i),$$

where $\mathfrak{E}$ denoted the harmonic extension and ν is the outward normal. Then, (2.1) implies $\|\tau_n\|_{\mathbb{L}^\infty(\Omega)} \leq \varrho$ and, since in addition τ_n vanishes on Γ_N by construction, we have $\tau_n \in \Sigma(\Omega) \cap \mathcal{K}(\Omega)$. Note that, by the mapping properties of $\mathfrak{E}$, $\tau_n \in \mathbb{W}^{1,p}(\Omega) \hookrightarrow \Sigma(\Omega)$. If we insert this as test function in (3.3b) and apply again the integration by parts from Lemma 3.6, then $\operatorname{div}\sigma = 0$ and (3.8) imply

$$\int_{\Gamma_D} \varphi_n \cdot \dot u \,\mathrm{d}s \leq \frac{\sqrt{2}}{\varrho}\Big(\int_\Omega \tau_n : \mathrm{d}\nabla^s(\dot u) - \int_\Omega \nabla^s \dot u_D : \sigma \,\mathrm{d}x + \int_\Omega \mathbb{A}\dot\sigma : (\tau_n - \sigma)\,\mathrm{d}x\Big).$$

Now, since $\varphi_n \to \operatorname{sign}(\dot u)$ a.e. in Λ, $\varphi_n \to 0$ a.e. in $\Gamma_D \setminus \Lambda$ and $|\varphi_n \cdot \dot u| \leq |\dot u|$ a.e. on Γ_D, Lebesgue's dominated convergence theorem along with our previous estimate gives

$$\|\dot u\|_{L^2(\mathbf{L}^1(\Lambda))} \leq C\,\|u_D\|_{H^1(\mathbf{H}^1(\Omega))}\big(1 + \|u_D\|_{H^1(\mathbf{H}^1(\Omega))}\big).$$

Thanks to [31, Chap. II, Proposition 2.4], this completes the proof. ∎

Remark 4.4. A priori estimates for quasistatic evolutions (which is an equivalent notion of solution as mentioned above) are already proven in [8, Thm. 5.2] in a slightly different setting.

LEMMA 4.5. *Let $\{u_n\} \subset \mathcal{U}$ be a sequence such that, for all $n \in \mathbb{N}$,*

$$u_n(0) = u_0 \quad \text{and} \quad \int_0^T \|\dot u_n(t)\|^2_{\mathrm{BD}(\Omega)}\,\mathrm{d}t \leq C \tag{4.3}$$

with a constant $C > 0$. Then there exists a subsequence converging weakly in $\mathcal{U}$ as defined in (3.1).

Proof. Owing to (4.3), $\{\nabla^s \dot u_n\}$ is bounded in $L^2_w(\mathfrak{M}(\Omega;\mathbb{R}^{n\times n}_{\mathrm{sym}}))$, which, according to [9, Theorem 8.20.3], is the dual of $L^2(C_0(\Omega;\mathbb{R}^{n\times n}_{\mathrm{sym}}))^*$. Thus, there exists a subsequence such that

$$\nabla^s \dot u_{n_k} \rightharpoonup^* w \quad \text{in } L^2_w(\mathfrak{M}(\Omega;\mathbb{R}^{n\times n}_{\mathrm{sym}})). \tag{4.4}$$

Due to $\mathrm{BD}(\Omega) \hookrightarrow \mathbf{L}^{\frac{n}{n-1}}(\Omega)$, $\{\dot u_{n_k}\}$ is bounded in $L^2(\mathbf{L}^{\frac{n}{n-1}}(\Omega))$ and, since all u_n share the same initial value, $\{u_{n_k}\}$ is bounded in $H^1(\mathbf{L}^{\frac{n}{n-1}}(\Omega))$ so that, by reflexivity, there is another subsequence (denoted w.l.o.g. by the same symbol) such that

$$u_{n_k} \rightharpoonup u \quad \text{in } H^1(\mathbf{L}^{\frac{n}{n-1}}(\Omega)). \tag{4.5}$$

Now, for every $\tau \in C_c^\infty(\Omega;\mathbb{R}^{n\times n}_{\mathrm{sym}})$ and every $\varphi \in C_c^\infty(0,T)$, (4.4) and (4.5) imply

$$\begin{aligned}
\int_0^T \langle w(t), \tau\rangle \varphi(t)\,\mathrm{d}t &= \lim_{k\to\infty}\int_0^T \langle \nabla^s \dot u_{n_k}(t), \tau\rangle\varphi(t)\,\mathrm{d}t \\
&= \lim_{k\to\infty}\int_0^T\int_\Omega \dot u_{n_k}(t)\cdot \operatorname{div}\tau\,\mathrm{d}x\,\varphi(t)\,\mathrm{d}t \\
&= \int_0^T\int_\Omega \dot u(t)\cdot\operatorname{div}\tau\,\mathrm{d}x\,\varphi(t)\,\mathrm{d}t
\end{aligned}$$

and hence $w(t) = \nabla^s \dot u(t)$ a.e. in $(0,T)$. ∎

PROPOSITION 4.6 (Continuity properties of weak solutions). *Let* $\{u_{D,n}\}_{n\in\mathbb{N}} \subset H^1(\mathbf{H}^1(\Omega))$ *be a sequence fulfilling* (4.2)*. Then, there is a subsequence of weak solutions* $\{u_{n_k}, \sigma_{n_k}\}_{k\in\mathbb{N}}$ *associated with* $\{u_{D,n_k}\}$ *such that*

$$\sigma_{n_k} \rightharpoonup \sigma \;\text{ in } H^1(\mathbb{L}^2(\Omega)), \quad u_{n_k} \rightharpoonup u \;\text{ in } \mathcal{U},$$

and the weak limit (u, σ) *is a weak solution associated with the limit* u_D*.*

Proof. Since we already know that the stress component of every weak solution is also a reduced one and the latter is unique by Lemma 3.11, the convergence of the stresses follows from Lemma 4.2 (even for the whole sequence).

Owing to Lemma 4.3, $\{\dot u_n\}$ fulfills the boundedness assumption in (4.3) so that, by Lemma 4.5, there is a subsequence $\{u_{n_k}\}$ converging weakly in $\mathcal{U}$ to some limit $u \in \mathcal{U}$. Due to $H^1(\mathbf{L}^1(\Omega)) \hookrightarrow C(\mathbf{L}^1(\Omega))$, the weak limit u also satisfies the initial condition.

It remains to prove that (u, σ) fulfills the flow rule inequality (3.3b). To this end, choose an arbitrary $\tau \in L^2(\mathbb{L}^2(\Omega))$ with $\tau(t) \in \Sigma(\Omega) \cap \mathcal{K}(\Omega)$ for almost all $t \in [0, T]$. Then, the flow rule inequality for (u_{n_k}, σ_{n_k}) along with $\operatorname{div} \sigma_{n_k} = 0$ and the (weak) convergences of u_{D,n_k}, u_{n_k}, and σ_{n_k} yields

$$\begin{aligned}
&\liminf_{k\to\infty} \big(\mathbb{A}\dot\sigma_{n_k}, \sigma_{n_k}\big)_{L^2(\mathbb{L}^2(\Omega))} \\
&\le \lim_{k\to\infty} \Big[\big(\mathbb{A}\dot\sigma_{n_k} - \nabla^s \dot u_{D,n_k}, \tau\big)_{L^2(\mathbb{L}^2(\Omega))} \\
&\qquad + \int_0^T \int_\Omega (\dot u_{n_k} - \dot u_{D,n_k}) \operatorname{div} \tau \,\mathrm{d}x\mathrm{d}t - \big(\nabla^s \dot u_{D,n_k}, \sigma_{n_k}\big)_{L^2(\mathbb{L}^2(\Omega))} \Big] \\
&= \big(\mathbb{A}\dot\sigma - \nabla^s \dot u_D, \tau\big)_{L^2(\mathbb{L}^2(\Omega))} + \int_0^T \int_\Omega (\dot u - \dot u_D) \operatorname{div} \tau \,\mathrm{d}x\mathrm{d}t - \big(\nabla^s \dot u_D, \sigma\big)_{L^2(\mathbb{L}^2(\Omega))}
\end{aligned} \tag{4.6}$$

where we used Lemma 3.9 in our companion paper [21] for the convergence of the last term. On the other hand, the weak lower semicontinuity of $\|\cdot\|_{\mathbb{L}^2(\Omega)_\mathbb{A}}$ together with $H^1(\mathbb{L}^2(\Omega)) \hookrightarrow C(\mathbb{L}^2(\Omega))$ gives

$$\begin{aligned}
\liminf_{k\to\infty} &\big(\mathbb{A}\dot\sigma_{n_k}, \sigma_{n_k}\big)_{L^2(\mathbb{L}^2(\Omega))} \\
&= \frac{1}{2} \liminf_{k\to\infty} \|\sigma_{n_k}(T)\|^2_{\mathbb{L}^2(\Omega)_\mathbb{A}} - \frac{1}{2} \|\sigma_0\|^2_{\mathbb{L}^2(\Omega)_\mathbb{A}} \\
&\ge \frac{1}{2} \|\sigma(T)\|^2_{\mathbb{L}^2(\Omega)_\mathbb{A}} - \frac{1}{2} \|\sigma_0\|^2_{\mathbb{L}^2(\Omega)_\mathbb{A}} = \big(\mathbb{A}\dot\sigma, \sigma\big)_{L^2(\mathbb{L}^2(\Omega))}.
\end{aligned} \tag{4.7}$$

Together with (4.6) and $\operatorname{div} \sigma = 0$, this implies the flow rule inequality for the weak limit. ∎

Given these boundedness and continuity results, we can now establish the existence of at least one optimal solution. Before we do so, let us recall our optimization problem and state it in a rigorous manner:

$$\text{(P)} \qquad \begin{cases} \min & J(u, u_D) := \Psi(u) + \dfrac{\alpha}{2} \|u_D\|^2_{H^1(\mathbf{H}^2(\Omega))} \\ \text{s.t.} & u_D \in H^1(\mathbf{H}^2(\Omega)), \quad (u, \sigma) \in \mathcal{U} \times H^1(\mathbb{L}^2(\Omega)), \\ & (u, \sigma) \text{ is a weak solution w.r.t. } u_D, \quad \text{and} \quad u_D(0) - u_0 \in \mathbf{H}^1_D(\Omega), \end{cases}$$

where $\Psi : \mathcal{U} \to \mathbb{R}$ is bounded from below and lower semicontinous w.r.t. weak convergence in $\mathcal{U}$ as defined in (3.1), i.e.,

$$u_n \rightharpoonup u \text{ in } \mathcal{U} \quad \Longrightarrow \quad \liminf_{n\to\infty} \Psi(u_n) \ge \Psi(u). \tag{4.8}$$

An example for such a functional Ψ will be given in section 6 below.

THEOREM 4.7 (Existence of optimal solutions). *There exists a globally optimal solution of* (P)*.*

Proof. Based on our above findings, the assertion immediately follows from the standard direct method of calculus of variations. Nevertheless, let us shortly sketch the arguments. First, we observe that the triple $(u, \sigma, u_D) \equiv (u_0, \sigma_0, u_0)$ (constant in time) satisfies the constraints in (P) so that the feasible set is nonempty. (At this point, we need the additional regularity $u_0 \in \mathbf{H}^2(\Omega)$.) Let $(u_n, \sigma_n, u_{D,n})$ be a minimizing sequence. Then either (u_0, u_0) is already optimal or $J(u_n, u_{D,n}) \leq J(u_0, u_0) < \infty$ for $n \in \mathbb{N}$ sufficiently large. Thus, since Ψ is bounded from below, $\{u_{D,n}\}$ is bounded in $H^1(\mathbf{H}^2(\Omega))$. Via continuous and compact embedding, there is thus a subsequence satisfying (4.2). Clearly, the associated limit satisfies the conditions on the initial value in (P). Moreover, according to Proposition 4.6, a subsequence of weak solutions converges weakly in $\mathcal{U} \times H^1(\mathbb{L}^2(\Omega))$ to a weak solution. Thus the weak limit is feasible and the weak lower semicontinuity of norms and of Ψ implies its optimality. ∎

Remark 4.8 (More general objectives). The proof of existence readily transfers to slightly more general objectives than the one in (P). For instance, one could add a term of the form $\Phi(\sigma)$ with a function $\Phi : H^1(\mathbb{L}^2(\Omega)) \to \mathbb{R}$, which weakly lower semicontinuous and bounded from below. Since objectives of this form have already been discussed in the companion paper, we restrict ourselves to objectives just depending on u in order to keep the discussion concise. Moreover, one could use other Tikhonov terms different from the $H^1(\mathbf{H}^2(\Omega))$-norm to ensure the convergence properties in (4.2) required for Proposition 4.6. For example, thanks to the Aubin-Lions lemma, a Tikhonov term of the form

$$\frac{\alpha}{2}\left(\|u_D\|^2_{H^1(\mathbf{H}^1(\Omega))} + \|u_D\|^2_{L^2(X)}\right)$$

with any Banach space X embedding compactly in $\mathbf{H}^1(\Omega)$ (such as e.g. $\mathbf{H}^2(\Omega)$) is sufficient to guarantee (4.2) for (a subsequence of) a minimizing sequence. However, in order to shorten presentation, we just consider the $H^1(\mathbf{H}^2(\Omega))$-norm.

5. Yosida Regularization and Reverse Approximation.

As already mentioned above, the ultimate goal of our analysis is to establish conditions that guarantee that optimal solutions to the optimization problem (P) governed by perfect plasticity can be approximated via Yosida regularization. The most crucial point in this respect is the so-called *reverse approximation*, which essentially means to construct a *recovery sequence* for a given perfect plastic solution. This is a rather challenging task, as Example 3.10 illustrates: one easily verifies that every sequence of regularized solutions tends to the linear solution $u(t, x) = 2\,t\,x$ for regularization parameter tending to zero, although there are infinitely many other solutions. There is thus no hope that every perfect plastic solution can be approximated via Yosida regularization! However, when it comes to optimization, there is not only the state (i.e., the solution of the perfect plasticity system), but also the *control* variables, which can be used to construct a recovery sequence. Unfortunately, the Dirichlet data u_D, which serve as control variables in our case, are not sufficient for this purpose. Instead we need a set of control variables that is rich enough to generate a sufficiently large set of regularized solutions. For this purpose, we introduce an *additional control variable in form of distributed loads* and end up with the following regularized version of the state equation:

$$-\operatorname{div}\sigma_\lambda(t) = \ell(t) \quad \text{in } \mathbf{H}_D^{-1}(\Omega), \tag{5.1a}$$
$$\sigma_\lambda(t) = \mathbb{C}(\nabla^s u_\lambda(t) - z_\lambda(t)) \quad \text{in } \mathbb{L}^2(\Omega), \tag{5.1b}$$
$$\dot z_\lambda(t) = \partial I_\lambda(\sigma_\lambda(t)) \quad \text{in } \mathbb{L}^2(\Omega), \tag{5.1c}$$
$$u_\lambda(t) - u_D(t) \in \mathbf{H}_D^1(\Omega), \tag{5.1d}$$
$$(u_\lambda, \sigma_\lambda)(0) = (u_0, \sigma_0) \quad \text{in } \mathbf{H}^1(\Omega) \times \mathbb{L}^2(\Omega). \tag{5.1e}$$

where $\lambda > 0$ is the regularization parameter, I_λ is the Yosida regularization of the indicator functional, see (2.4), and $\ell \in H^1(\mathbf{H}_D^{-1}(\Omega))$ is the mentioned load. Existence and uniqueness of a solution to the regularized state equation (5.1) follows from Banach's fixed point theorem and can be proven by a reduction of the system to an equation in the variable z only, cf. e.g. [21, Proposition 3.15]. This gives rise to the following

Lemma 5.1 (Existence of solutions to the regularized state system, [21, Corollary 3.16]). *For every $\lambda > 0$, $\ell \in H^1(\mathbf{H}_D^{-1}(\Omega))$, and $u_D \in H^1(\mathbf{H}^1(\Omega))$ with $\ell(0) = 0$ and $u_D(0)|_{\Gamma_D} = u_0|_{\Gamma_D}$, there exists a unique solution $(u_\lambda, \sigma_\lambda, z_\lambda) \in H^1(\mathbf{H}^1(\Omega)) \times H^1(\mathbb{L}^2(\Omega)) \times H^1(\mathbb{L}^2(\Omega))$ of (5.1).*

The associated solution operator is globally Lipschitz continuous with a Lipschitz constant proportional to λ^{-1}.

The proof of existence is a direct consequence of the Lipschitz continuity of ∂I_λ and Banach's contraction principle. In [21], the external loads are set to zero, but it is straightforward to incorporate them into the existence theory. The Lipschitz continuity of the solution mapping directly follows from the Lipschitz estimate for the Yosida approximation, see e.g. [3, Proposition 55.2(b)].

Before we address the approximation properties of this regularization approach and its convergence behavior for λ tending to zero in section 6 below, see Proposition 6.2, we first lay the foundations for the construction of a recovery sequence in the upcoming three lemmas. Unfortunately, as already indicated in the introduction, the passage to the limit in the regularized state equation in Proposition 6.2 below requires a rather high regularity of the stress field, and the recovery sequence has to fulfill this regularity, too, as it is a constraint in the regularized optimal control problem (P_λ). The key issue for our reverse approximation argument is therefore to improve the regularity of the stress field provided a displacement field with higher regularity is given. To this end, we first need an auxiliary result on the derivative of the Yosida regularization. Since the set of admissible stresses admits a pointwise representation by the set K, the Fréchet-derivative of the Yosida regularization does the same, i.e., given an arbitrary $\tau \in \mathbb{L}^2(\Omega)$, it holds

$$\partial I_\lambda(\tau)(x) = \frac{1}{\lambda}\big[\tau(x) - \pi_K(\tau(x))\big] \quad \text{f.a.a. } x \in \Omega, \tag{5.2}$$

where $\pi_K : \mathbb{R}^{n\times n}_{\text{sym}} \to \mathbb{R}^{n\times n}_{\text{sym}}$ is the projection on K. This pointwise representation allows to derive the following

Lemma 5.2. *Let $\lambda > 0$, $p > 2$, and $\tau \in \mathbb{W}^{1,p}(\Omega)$ be arbitrary. Then $\partial I_\lambda(\tau) \in \mathbb{W}^{1,p}(\Omega)$ and there holds*

$$\|\partial I_\lambda(\tau)\|_{\mathbb{W}^{1,p}(\Omega)} \le \frac{1}{\lambda}\,\|\tau\|_{\mathbb{W}^{1,p}(\Omega)} \tag{5.3}$$

and

$$(5.4)\qquad \big(\partial_i(\partial I_\lambda(\tau))\colon \partial_i\tau\big)(x) \geq 0 \quad \textit{a.e. in } \Omega, \quad \forall\, i = 1, ..., n.$$

Proof. As a projection, $\pi_k : \mathbb{R}^{n\times n}_{\text{sym}} \to \mathbb{R}^{n\times n}_{\text{sym}}$ is globally Lipschitz continuous. Thus, the chain rule for Sobolev functions (see e.g. [36, Thm 2.1.11]) implies that $\partial I_\lambda(\tau) \in \mathbb{W}^{1,p}(\Omega)$ with

$$(5.5)\qquad \frac{\partial}{\partial x_m}[\partial I_\lambda(\tau)]_{ij} = \frac{1}{\lambda}\Big(\frac{\partial \tau_{ij}}{\partial x_m} - \sum_{kl} \frac{\partial}{\partial \tau_{kl}}[\pi_K(\tau)]_{ij}\frac{\partial \tau_{kl}}{\partial x_m}\Big).$$

Since the Lipschitz constant of the projection equals one, its directional derivative clearly satisfies $|\pi_K'(A;B)|_F \leq |B|_F$ for all $A, B \in \mathbb{R}^{n\times n}_{\text{sym}}$ and, consequently,

$$\partial_m(\partial I_\lambda(\tau))\colon \partial_m\tau = \frac{1}{\lambda}\big(|\partial_m\tau|_F^2 - \pi_K'(\tau;\partial_m\tau) : \partial_m\tau\big) \geq 0,$$

which is (5.4). It is moreover easily seen that $\operatorname{Id} - \pi_K : \mathbb{R}^{n\times n}_{\text{sym}} \to \mathbb{R}^{n\times n}_{\text{sym}}$ is globally Lipschitz with Lipschitz constant 1, too. Thus, for every $A, B \in \mathbb{R}^{n\times n}_{\text{sym}}$, there holds $|(\operatorname{Id} - \pi_K)'(A;B)|_F \leq |B|_F$. Since $(\operatorname{Id} - \pi_k)(0) = 0$, the Lipschitz continuity moreover entails $|(\operatorname{Id} - \pi_k)(A)|_F \leq |A|_F$ for all $A \in \mathbb{R}^{n\times n}_{\text{sym}}$. In view of (5.5), this yields (5.3). □

The next lemma addresses the crucial regularity result for the stress field σ_λ as solution of

$$(5.6)\qquad w - \mathbb{A}\dot\sigma_\lambda = \partial I_\lambda(\sigma_\lambda), \qquad \sigma_\lambda(0) = \sigma_0.$$

In the proof of our main result in Theorem 6.3, an optimal strain rate will play the role of w and the following regularity result will be essential for the construction of a recovery sequence associated with that strain rate. The required regularity of w will carry over to this optimal strain rate and represents the most restrictive assumption of our reverse approximation approach.

LEMMA 5.3 (Higher regularity of the stress field). *Let $\lambda > 0$ be arbitrary and $w \in L^2(\mathbb{L}^2(\Omega)) \cap L^1(\mathbb{W}^{1,p}(\Omega))$ with $p \geq 2$ be given. Then (5.6) admits a unique solution $\sigma_\lambda \in H^1(\mathbb{L}^2(\Omega)) \cap L^\infty(\mathbb{W}^{1,p}(\Omega))$ and there holds*

$$(5.7)\qquad \|\sigma_\lambda\|_{L^\infty(\mathbb{W}^{1,p}(\Omega))} \leq C_p\Big(\|w\|_{L^1(\mathbb{W}^{1,p}(\Omega))} + \|\sigma_0\|^p_{\mathbb{W}^{1,p}(\Omega)}\Big)$$

with $C_p := p\,\|\mathbb{A}\|^{p/2-1}$.

Proof. Step 1. Existence of solutions in $H^1(\mathbb{L}^2(\Omega))$: First we note that (5.6) is just an ODE in $\mathbb{L}^2(\Omega)$ and ∂I_λ is globally Lipschitz in $\mathbb{L}^2(\Omega)$. Thus, the existence and uniqueness of solutions in $H^1(\mathbb{L}^2(\Omega))$ follows from the generalized Picard-Lindelöf theorem in Banach spaces. However, a pointwise projection is in general not Lipschitz continuous in Sobolev spaces. Therefore, we cannot apply this simple argument to show that the solution is an element of $W^{1,1}(\mathbb{W}^{1,p}(\Omega))$.

Step 2. Higher regularity in case of smooth data: To prove this, we apply a time discretization scheme, namely the explicit Euler method. At first we consider the case $w \in C(\mathbb{W}^{1,p}(\Omega))$. For $N \in \mathbb{N}$ and $n \in \{0, ..., N\}$, we set $d_t^N := \frac{T}{N}$ and $t_n^N := n d_t^N$ such that $0 = t_0^N < t_1^N < ... < t_N^N = T$. Now define $\sigma_0^N := \sigma_0 \in \mathbb{W}^{1,p}(\Omega)$ and

$$\sigma_n^N := \sigma_{n-1}^N + d_t^N \mathbb{C}\big(w(t_{n-1}^N) - \partial I_\lambda(\sigma_{n-1}^N)\big) \in \mathbb{W}^{1,p}(\Omega) \quad \text{(by Lemma 5.2)}$$

such that

$$\mathbb{A}\frac{\sigma_n^N - \sigma_{n-1}^N}{d_t^N} + \partial I_\lambda(\sigma_{n-1}^N) = w(t_{n-1}^N) \tag{5.8}$$

for all $N \in \mathbb{N}$ and $n \in \{1, ..., N\}$. We define the piecewise linear approximation $\sigma^N \in W^{1,\infty}(\mathbb{W}^{1,p}(\Omega))$ by

$$\sigma^N(t) := \sigma_{n-1}^N + \frac{t - t_{n-1}^N}{d_t^N}(\sigma_n^N - \sigma_{n-1}^N)$$

and the piecewise constant approximation $\tilde{\sigma}^N \in L^\infty(\mathbb{W}^{1,p}(\Omega))$ by $\tilde{\sigma}^N(t) := \sigma_{n-1}^N$ for $t \in [t_{n-1}^N, t_n^N)$. Using (5.3), we deduce from (5.8) that

$$\|\sigma_n^N\|_{\mathbb{W}^{1,p}(\Omega)} \leq \|\sigma_0\|_{\mathbb{W}^{1,p}(\Omega)} + d_t^N C\Big(\sum_{i=0}^{n-1} \|\sigma_i^N\|_{\mathbb{W}^{1,p}(\Omega)}\Big) + C\|w\|_{C(\mathbb{W}^{1,p}(\Omega)},$$

which, together with the discrete Gronwall lemma (cf. [16, Lemma 5.1 and the following remark]), shows that $\sigma^N$is bounded in $L^\infty(\mathbb{W}^{1,p}(\Omega))$ by a constant independent of d_t^N. Thus, again owing to (5.8) and (5.3), $\dot{\sigma}^N(t) = \frac{\sigma_n^N - \sigma_{n-1}^N}{d_t^N}$, $t \in (t_{n-1}^N, t_n^N)$ is also bounded in $L^\infty(\mathbb{W}^{1,p}(\Omega))$. Therefore, σ^N is bounded in $H^1(\mathbb{L}^2(\Omega))$ and consequently, there is a weakly converging subsequence, for simplicity also denoted by σ^N, such that $\sigma^N \rightharpoonup \sigma$ in $H^1(\mathbb{L}^2(\Omega))$ and $\sigma^N \rightharpoonup^* \sigma$ in $L^\infty(\mathbb{W}^{1,p}(\Omega))$ as $N \to \infty$. Note that, due to the reflexivity of $\mathbb{W}^{1,p}(\Omega)$, $L^\infty(\mathbb{W}^{1,p}(\Omega))$ can be identified with the dual of $L^1(\mathbb{W}^{1,p'}(\Omega))$ so there is a weakly-$*$ converging subsequence. It remains to show that σ solves (5.6). Since σ^N is bounded in $W^{1,\infty}(\mathbb{W}^{1,p}(\Omega))$ as seen above, we have by compact embeddings that $\sigma^N \to \sigma$ in $C(\mathbb{L}^2(\Omega))$. Thus, we find for the piecewise constant interpolation that, for every $t \in [t_{n-1}^N, t_n^N)$,

$$\|\tilde{\sigma}^N(t) - \sigma(t)\|_{\mathbb{L}^2(\Omega)} \leq \|\sigma^N(t_{n-1}^N) - \sigma(t)\|^2_{\mathbb{L}}(\Omega) \to 0 \quad \text{as } N \to \infty.$$

Therefore, (5.8) and the Lipschitz continuity of ∂I_λ in $\mathbb{L}^2(\Omega)$ give

$$\begin{aligned}
&\|\mathbb{A}\dot{\sigma}^N + \partial I_\lambda(\sigma^N) - w\|_{L^2(\mathbb{L}^2(\Omega))} \\
&\leq \|\mathbb{A}\dot{\sigma}^N + \partial I_\lambda(\tilde{\sigma}^N) - \tilde{w}^N\|_{L^2(\mathbb{L}^2(\Omega))} + \|\partial I_\lambda(\sigma^N) - \partial I_\lambda(\tilde{\sigma}^N)\|_{L^2(\mathbb{L}^2(\Omega))} + \|\tilde{w}^N - w\|_{L^2(\mathbb{L}^2(\Omega))} \\
&\leq \frac{1}{\lambda}\|\sigma^N - \tilde{\sigma}^N\|_{L^2(\mathbb{L}^2(\Omega))} + \|\tilde{w}^N - w\|_{L^2(\mathbb{L}^2(\Omega))} \to 0 \quad \text{as } N \to \infty,
\end{aligned}$$

where $\tilde{w}^N$ denotes the piecewise constant interpolation of w, which converges strongly in $C(\mathbb{L}^2(\Omega))$ to w thanks to the assumed regularity of w. Therefore, by the weak lower semicontinuity of the $L^2(\mathbb{L}^2(\Omega))$-norm, we see that the limit satisfies (5.6).

Step 3. Higher regularity for nonsmooth data: Let now $w \in L^2(\mathbb{L}^2(\Omega)) \cap L^1(\mathbb{W}^{1,p}(\Omega))$ be arbitrary and take a sequence $\{w_n\} \subset C(\mathbb{W}^{1,p}(\Omega))$ such that $w_n \to w$ in $L^1(\mathbb{W}^{1,p}(\Omega))$. Let $\sigma_\lambda \in H^1(\mathbb{L}^2(\Omega))$ be the solution of (5.6) and denote by $\sigma_{\lambda,n} \in H^1(\mathbb{L}^2(\Omega)) \cap L^\infty(\mathbb{W}^{1,p}(\Omega))$ the solution of

$$w_n - \mathbb{A}\dot{\sigma}_{\lambda,n} = \partial I_\lambda(\sigma_{\lambda,n}), \qquad \sigma_{\lambda,n}(0) = \sigma_0. \tag{5.9}$$

Since $\partial I_\lambda : \mathbb{L}^2(\Omega) \to \mathbb{L}^2(\Omega)$ is monotone, one obtains $\sigma_{\lambda,n} \to \sigma_\lambda$ in $H^1(\mathbb{L}^2(\Omega))$ by standard arguments. Moreover, (5.9) holds almost everywhere in time and space and so that, f.a.a. $t \in [0,T]$,

$$\partial_j w_n(t) - \mathbb{A}\partial_j \dot\sigma_{\lambda,n}(t) = \partial_j \partial I_\lambda(\sigma_{\lambda,n})(t) \quad \text{a.e. in } \Omega.$$

follows. Testing this equation with $((\mathbb{A}\partial_j\sigma_{\lambda,n} : \partial_j\sigma_{\lambda,n})^{p/2-1}\partial_j\sigma_{\lambda,n})(t) \in \mathbb{W}^{1,p'}(\Omega)$ and using (5.4) leads to

$$\begin{aligned}
&\frac{d}{dt}\int_\Omega (\mathbb{A}\partial_j\sigma_{\lambda,n} : \partial_j\sigma_{\lambda,n})^{p/2}\mathrm{d}x \\
&\leq p\int_\Omega (\mathbb{A}\partial_j\sigma_{\lambda,n} : \partial_j\sigma_{\lambda,n})^{p/2-1}\big(\mathbb{A}\partial_j\sigma_{\lambda,n} : \partial_j\dot\sigma_{\lambda,n} + \partial_j\partial I_\lambda(\sigma_{\lambda,n}) : \partial_j\sigma_{\lambda,n}\big)\mathrm{d}x \\
&= p\int_\Omega (\mathbb{A}\partial_j\sigma_{\lambda,n} : \partial_j\sigma_{\lambda,n})^{p/2-1}\,\partial_j w_n : \partial\sigma_n\,\mathrm{d}x \\
&\leq C_p\int_\Omega |\partial_j w|_F\,|\partial_j\sigma_{\lambda,n}|_F^{p-1}\,\mathrm{d}x \leq C_p\,\|w\|_{\mathbb{W}^{1,p}(\Omega)}\,\|\sigma_{\lambda,n}\|_{\mathbb{W}^{1,p}(\Omega)}^{p-1},
\end{aligned} \tag{5.10}$$

with C_p as defined in the statement of the lemma. Integrating this inequality in time and taking the coercivity of $\mathbb{A}$ into account gives

$$\|\sigma_{\lambda,n}\|_{L^\infty(\mathbb{W}^{1,p}(\Omega))} \leq C_p\Big(\|w_n\|_{L^1(\mathbb{W}^{1,p}(\Omega))} + \|\sigma_0\|^p_{\mathbb{W}^{1,p}(\Omega)}\Big).$$

Therefore, $\sigma_{\lambda,n}$ is bounded in $L^\infty(\mathbb{W}^{1,p}(\Omega))$ and we can select a weakly-$*$ converging subsequence. The uniqueness of the weak limit then gives $\sigma \in L^\infty(\mathbb{W}^{1,p}(\Omega))$ as claimed. The estimate in (5.7) finally follows from the above inequality and the lower semicontinuity of the norm w.r.t. weak-$*$ convergence. ∎

Remark 5.4. We observe that (5.3), (5.6), and the proven regularity of σ_λ even imply that $\sigma_\lambda \in W^{1,1}(\mathbb{W}^{1,p}(\Omega))$. However, we do not obtain an estimate independent of λ in this norm (in contrast to (5.7)) and therefore, this additional regularity is not useful for us.

LEMMA 5.5 ([20, Section 3]). *Let $w \in L^2(\mathbb{L}^2(\Omega))$ be given and $\lambda \searrow 0$. Then $\sigma_\lambda \to \sigma$ in $H^1(\mathbb{L}^2(\Omega))$, where σ is the solution of*

$$w - \mathbb{A}\dot\sigma \in \partial I_{\mathcal{K}(\Omega)}(\sigma), \qquad \sigma(0) = \sigma_0. \tag{5.11}$$

Moreover, there holds

$$\|\sigma_\lambda - \sigma\|^2_{C(\mathbb{L}^2(\Omega))} \leq \lambda\,\frac{\|\mathbb{C}\|^2}{\gamma_{\mathbb{C}}}\,\|w - \mathbb{A}\dot\sigma\|^2_{L^2(\mathbb{L}^2(\Omega))}, \tag{5.12}$$

where $\gamma_{\mathbb{C}} > 0$ is the coercivity constant of $\mathbb{C}$.

Proof. The assertion is proven in [20], but, for convenience of the reader, we shortly sketch the arguments. First, observe that $\sigma_\lambda \in H^1(\mathbb{L}^2(\Omega))$ and $\sigma \in H^1(\mathbb{L}^2(\Omega))$ solve (5.6) and (5.11), respectively, if and only if $z_\lambda := W - \mathbb{A}\sigma_\lambda$ and $z := W - \mathbb{A}\sigma$ with $W(t) := \int_0^t w(s)\mathrm{d}s$ solve

$$\dot z_\lambda = \partial I_\lambda(\mathbb{C}W - \mathbb{C}z_\lambda), \qquad z_\lambda(0) = z_0 := -\mathbb{A}\sigma_0 \tag{5.13}$$

and

$$\dot z \in \partial I_{\mathcal{K}(\Omega)}(\mathbb{C}W - \mathbb{C}z), \qquad z(0) = z_0, \tag{5.14}$$

respectively. These equations are exactly of the form studied in [20, Section 3] with the setting $A := \partial I_{\mathcal{K}(\Omega)}$, $Q = R := \mathbb{C}$, and $\ell := W$. The existence of σ in $H^1(\mathbb{L}^2(\Omega))$ then follows from [20, Theorem 3.3], while the convergence $\sigma_\lambda \to \sigma$ in $H^1(\mathbb{L}^2(\Omega))$ as well as the estimate

$$\|\mathbb{A}(\sigma_\lambda - \sigma)\|^2_{C(\mathbb{L}^2(\Omega))} \leq \frac{\lambda}{\gamma_{\mathbb{C}}} \|w - \mathbb{A}\dot{\sigma}\|^2_{L^2(\mathbb{L}^2(\Omega))}$$

are consequences of [20, Proposition 3.5]. (Note that $D(A) = \mathcal{K}(\Omega)$ is closed and $A^0 \equiv 0$ in this case, hence, the assumptions in [20, Section 2] are fulfilled). The inequality in (5.12) now follows easily using $\|\sigma_\lambda - \sigma\|_{\mathbb{L}^2(\Omega)} = \|\mathbb{C}\mathbb{A}(\sigma_\lambda - \sigma)\|_{\mathbb{L}^2(\Omega)} \leq \|\mathbb{C}\| \, \|\mathbb{A}(\sigma_\lambda - \sigma)\|_{\mathbb{L}^2(\Omega)}$. ∎

Remark 5.6. As a consequence of (5.7), the solution of (5.11) is an element of $L^\infty(\mathbb{W}^{1,p}(\Omega))$, provided that $w \in L^1(\mathbb{W}^{1,p}(\Omega))$. However, we do not need this regularity result for the upcoming analysis.

As already mentioned, in the proof of our final convergence result in Theorem 6.3, $\nabla^s \dot{\overline{u}}$ will play the role of the function w, where $\overline{u}$ is an optimal solution of (P). This already indicates our most restrictive assumption, namely the existence of an optimal solution providing the high regularity required for w. We will come back to this point in Remark 6.4.

6. Convergence of Minimizers.

We are now in the position to state the regularized optimal control problems. Beside the additional control variable ℓ required for the reverse approximation, they differ from (P) in an additional inequality constraint on the stress field, which is needed to improve the regularity of the stress in order to pass to the limit in the regularized state equation, see the proof of Proposition 6.2 below. This additional regularity of the stresses is unfortunately not enough to pass to the limit in the state system. We additionally need to bound the displacement in $\mathcal{U}$, since this is not guaranteed a priori by the regularized state system itself, unless the loads fulfill a safe load condition. This however cannot be ensured for the loads arising in the construction of the recovery sequence in the proof of our main Theorem 6.3 (at least, we were not able to verify it). Therefore, we directly enforce this boundedness by a special choice of the objective functional as a tracking type objective of the following form:

$$\Psi(u) := \int_0^T \|\nabla^s \dot{u}(t) - \mu(t)\|^2_{\mathfrak{M}(\Omega;\mathbb{R}^{n\times n}_{\mathrm{sym}})} + \|\dot{u}(t) - v(t)\|^2_{\mathbf{L}^1(\Omega)} \,\mathrm{d}t \tag{6.1}$$

with a given desired strain rate $\mu \in L^2(\mathbb{L}^1(\Omega))$ and a desired displacement rate $v \in L^2(\mathbf{L}^1(\Omega))$. Note that this objective trivially fulfills the lower semicontinuity assumption in (4.8). One could even allow for less regular desired strain rates (in the space of measures), but for convenience, we restrict to functions in $L^2(\mathbb{L}^1(\Omega))$. The regularized counterpart of (P) now reads as follows:

$$(\mathrm{P}_\lambda) \quad \begin{cases} \min & J_\lambda(u, u_D, \ell) := \|\nabla^s \dot{u} - \mu\|^2_{L^2(\mathbb{L}^1(\Omega))} + \|\dot{u} - v\|^2_{L^2(\mathbf{L}^1(\Omega))} \\ & \qquad + \dfrac{\alpha}{2}\|u_D\|^2_{H^1(\mathbf{H}^2(\Omega))} + \lambda^{-\theta}\|\ell\|^2_{L^2(\mathbf{H}_D^{-1}(\Omega))} + \|\dot{\ell}\|^2_{L^2(\mathbf{H}_D^{-1}(\Omega))} \\ \text{s.t.} & u_D \in H^1(\mathbf{H}^2(\Omega)), \quad \ell \in H^1(\mathbf{H}_D^{-1}(\Omega)), \\ & u_D(0) - u_0 \in \mathbf{H}_D^1(\Omega), \quad \ell(0) = 0, \\ & (u, \sigma, z) \in \mathcal{U} \cap L^2(\mathbf{H}^1(\Omega)) \times L^2(\mathbb{L}^2(\Omega)) \times H^1(\mathbb{L}^2(\Omega)), \\ & (u, \sigma, z) \text{ is the solution of (5.1) w.r.t. } u_D \text{ and } \ell, \\ & \|\dot{\sigma}\|_{L^2(\mathbb{L}^2(\Omega))} + \|\sigma\|_{L^s(\mathbb{W}^{1,p}(\Omega))} \le R \end{cases}$$

with $0 < \theta < 1$ and

$$(6.2) \qquad p > n \quad \text{and} \quad s > \max\Big\{1, \frac{2np}{np + 2(p-n)}\Big\}$$

and $R \ge \|\sigma_0\|_{\mathbb{W}^{1,p}(\Omega)}$ to be specified later, see (6.6) below. With the exponents in (6.2), [19, Lemma 4.2(i)] is applicable and tells us that $H^1(\mathbb{L}^2(\Omega)) \cap L^s(\mathbb{W}^{1,p}(\Omega))$ embeds compactly in $L^2(C(\bar{\Omega}; \mathbb{R}^{n\times n}_{\text{sym}}))$, which will be useful at several places in the upcoming proofs. The term in the objective associated with θ will be used to force the additional loads to zero in the limit.

Proposition 6.1. *For every $\lambda > 0$, there exists a globally optimal solution of* (P_λ).

Proof. The proof is almost standard, except for a lack of compactness with regard to the control space. Let $(u_n, \sigma_n, z_n, u_{D,n}, \ell_n)$ be a minimizing sequence. As in the proof of Theorem 4.7, $(u, \sigma, z, u_D, \ell) \equiv (u_0, \sigma_0, \nabla^s u_0 - \mathbb{A}\sigma_0, u_0, 0)$ is feasible for (P_λ). Thus, $\{u_{D,n}, \ell_n\}$ is bounded in $H^1(\mathbf{H}^2(\Omega)) \times H^1(\mathbf{H}_D^{-1}(\Omega))$. Hence the Lipschitz continuity of the solution operator associated with (5.1) implies that $\{(u_n, \sigma_n, z_n)\}$ is bounded in $H^1(\mathbf{H}^1(\Omega) \times \mathbb{L}^2(\Omega) \times \mathbb{L}^2(\Omega))$. Therefore, there exist weakly convergent subsequences and we can pass to the limit in (5.1) except for the nonlinearity in ∂I_λ. However, the additional constraint on the stress implies that σ_n also converges weakly in $H^1(\mathbb{L}^2(\Omega)) \cap L^s(\mathbb{W}^{1,p}(\Omega))$, which is compactly embedded in $L^2(C(\bar{\Omega}; \mathbb{R}^{n\times n}_{\text{sym}}))$ as mentioned above. Thus $\{\sigma_n\}$ converges strongly in $L^2(C(\bar{\Omega}; \mathbb{R}^{n\times n}_{\text{sym}}))$, which allows to pass to the limit in $\partial I_\lambda(\sigma_n)$ so that the weak limit solves (5.1). Moreover, the inequality constraint on σ_n is clearly weakly closed so that the weak limit is indeed feasible for (P_λ). Since the objective is convex and continuous and thus weakly lower semicontinuous, the weak limit is also optimal. ∎

Proposition 6.2 (Convergence of the Yosida regularization with varying loads). *Let $\{\lambda_n\}_{n\in\mathbb{N}}$ be a sequence converging to zero. Suppose moreover that two sequences $\{\ell_n\} \subset H^1(\mathbf{H}_D^{-1}(\Omega))$ and $\{u_{D,n}\} \subset H^1(\mathbf{H}^1(\Omega))$ are given and denote the solution of* (5.1) *associated with λ_n, ℓ_n, and $u_{D,n}$ by (u_n, σ_n, z_n). Furthermore, we assume that $\{u_{D,n}\}$ satisfies the convergence properties in* (4.2)*, i.e.,*

$$(6.3) \qquad \begin{aligned} u_{D,n} \rightharpoonup u_D \ \text{ in } H^1(\mathbf{H}^1(\Omega)), \quad u_{D,n} \to u_D \ \text{ in } L^2(\mathbf{H}^1(\Omega)), \\ u_{D,n}(T) \to u_D(T) \ \text{ in } \mathbf{H}^1(\Omega), \end{aligned}$$

and that

$$(6.4) \qquad \ell_n \rightharpoonup 0 \ \text{ in } L^2(\mathbf{H}_D^{-1}(\Omega)), \quad u_n \rightharpoonup u \ \text{ in } \mathcal{U},$$

$$(6.5) \qquad \sigma_n \rightharpoonup \sigma \ \text{ in } H^1(\mathbb{L}^2(\Omega)), \qquad \sigma_n \to \sigma \ \text{ in } L^2(C(\bar{\Omega}; \mathbb{R}^{n\times n}_{\text{sym}}).$$

Then (u,σ) is a weak solution associated with u_D.

Proof. The arguments are similar to the proof of Proposition 4.6. First, since $\sigma_n \rightharpoonup \sigma$ in $H^1(\mathbb{L}^2(\Omega))$, $\ell_n \rightharpoonup 0$ in $L^2(\mathbf{H}_D^{-1}(\Omega))$, and $-\operatorname{div}\sigma_n = \ell_n$ for all $n \in \mathbb{N}$, it follows that $\sigma(t) \in \mathcal{E}(\Omega)$ f.a.a. $t \in (0,T)$. Moreover, from Lemma 3.20 and 3.21 in our compaion paper [21], we deduce that $\sigma(t) \in \mathcal{K}(\Omega)$ a.e. in $(0,T)$, cf. also the first part of the proof of [21, Theorem 3.22]. Moreover, due to $H^1(\mathbf{L}^{\frac{n}{n-1}}(\Omega)) \hookrightarrow C(\mathbf{L}^{\frac{n}{n-1}}(\Omega))$ and $H^1(\mathbb{L}^2(\Omega)) \hookrightarrow C(\mathbb{L}^2(\Omega))$, the weak limit satisfies the initial conditions.

To show the flow rule inequality, let $\tau \in L^2(\mathbb{L}^2(\Omega))$ with $\tau(t) \in \Sigma(\Omega)\cap\mathcal{K}(\Omega)$ f.a.a. $t \in (0,T)$ be arbitrary. Then, (5.1b) and (5.1c) along with $I_n(a) = 0$ for $a \in \mathcal{K}(\Omega)$ and $I_n \geq 0$, imply

$$
\begin{aligned}
0 &= \int_0^T I_n(\tau(t))dt \\
&\geq \left(\nabla^s \dot{u}_n - \mathbb{A}\dot{\sigma}_n, \tau - \sigma_n\right)_{L^2(\mathbb{L}^2(\Omega))} \\
&= \left(\nabla^s \dot{u}_{D,n} - \mathbb{A}\dot{\sigma}_n, \tau - \sigma_n\right)_{L^2(\mathbb{L}^2(\Omega))} - \int_0^T \int_\Omega (\dot{u}_n - \dot{u}_{D,n}) \operatorname{div}\tau \, \mathrm{d}x\mathrm{d}t \\
&\quad + \left(\nabla^s \dot{u}_{D,n} - \nabla^s \dot{u}_n, \sigma_n\right)_{L^2(\mathbb{L}^2(\Omega))}.
\end{aligned}
$$

Now one can pass to the limit with the first two terms on the right hand side exactly as described at the end of the proof of Proposition 4.6, see (4.6) and (4.7). Concerning the last term, we argue as follows: Since $\operatorname{div}\sigma = 0$ and u_n satisfies the Dirichlet boundary condition, i.e., $\dot{u}_n = \dot{u}_{D,n}$ on Γ_D, we obtain

$$
\begin{aligned}
&\left| \left(\nabla^s \dot{u}_{D,n} - \nabla^s \dot{u}_n, \sigma_n\right)_{L^2(\mathbb{L}^2(\Omega))} \right| \\
&\qquad = \left| \left(\nabla^s \dot{u}_{D,n} - \nabla^s \dot{u}_n, \sigma_n - \sigma\right)_{L^2(\mathbb{L}^2(\Omega))} \right| \\
&\qquad \leq \|\nabla^s \dot{u}_{D,n} - \nabla^s \dot{u}_n\|_{L^2(\mathbb{L}^1(\Omega))} \|\sigma_n - \sigma\|_{L^2(C(\bar{\Omega};\mathbb{R}^{n\times n}_{\text{sym}})} \to 0,
\end{aligned}
$$

thanks to the boundedness of u_n in $\mathcal{U}$ and the convergence of σ_n. ∎

The last step of the above proof illustrates, where the high regularity of the stress field enforced by the additional inequality constraint in (P_λ) comes into play: we need the strong convergence of the stress in $L^2(C(\bar{\Omega};\mathbb{R}^{n\times n}_{\text{sym}})$ in order to pass to the limit in the flow rule inequality. Unfortunately, the recovery sequence needs to be feasible for (P_λ) and thus has to fulfill this inequality constraint, too. Using our results from section 5, this can be guaranteed, provided that there is at least one optimal solution, whose strain rate admits higher regularity. This is the most severe restriction for our main result:

THEOREM 6.3 (Approximation of global minimizers). *Let the objective in* (P) *be of the form* (6.1). *Assume moreover that there exists a global minimizer $(\overline{u}, \overline{\sigma}, \overline{u}_D)$ of* (P) *such that $\nabla^s\dot{\overline{u}} \in L^2(\mathbb{L}^2(\Omega)) \cap L^1(\mathbb{W}^{1,p}(\Omega))$ and $\overline{u} - \overline{u}_D \in \mathbf{H}^1_D(\Omega)$ for all $t \in (0,T)$. Suppose in addition that R in* (P_λ) *is chosen so large that*

$$
R \geq \frac{1}{\gamma_{\mathbb{A}}} \|\overline{u}_D\|_{H^1(\mathbf{H}^1(\Omega))} + p\,\|\mathbb{A}\|^{p/2-1}\Big(\|\nabla^s\dot{\overline{u}}\|_{L^1(\mathbb{W}^{1,p}(\Omega))} + \|\sigma_0\|^p_{\mathbb{W}^{1,p}(\Omega)}\Big). \tag{6.6}
$$

Furthermore, let $\{\overline{u}_\lambda, \overline{\sigma}_\lambda, \overline{z}_\lambda, \overline{u}_{D,\lambda}, \overline{\ell}_\lambda\}_{\lambda>0}$ be a sequence of global minimizers of (P_λ) *for $\lambda \searrow 0$.*

Then there exists an accumulation point of $\{\overline{u}_\lambda, \overline{\sigma}_\lambda, \overline{u}_{D,\lambda}\}_{\lambda>0}$ w.r.t. weak convergence in $\mathcal{U} \times H^1(\mathbb{L}^2(\Omega)) \cap L^s(\mathbb{W}^{1,p}(\Omega)) \times H^1(\mathbf{H}^2(\Omega))$. Moreover, every such accumulation point is a global minimizer of (P).

Furthermore, if $(\tilde u,\tilde\sigma,\tilde u_D)$ is such an accumulation point and $\{\overline{u}_\lambda,\overline{\sigma}_\lambda,\overline{u}_{D,\lambda}\}_{\lambda>0}$ the associated sequence converging weakly to it, then

$$\overline{u}\to\tilde u \quad in\ H^1(L^1(\Omega;\mathbb{R}^n)),\qquad \overline{u}_{D,\lambda}\to\tilde u_D \quad in\ H^1(\mathbf{H}^2(\Omega)), \tag{6.7}$$

$$\overline{\sigma}_\lambda\to\tilde\sigma \quad in\ L^2(C(\bar\Omega;\mathbb{R}^{n\times n}_{\rm sym})),\qquad \overline{\ell}_\lambda\to 0 \quad in\ H^1(\mathbf{H}_D^{-1}(\Omega)). \tag{6.8}$$

Proof. Step 1. Existence of an accumulation point. Since $\{\overline{u}_\lambda,\overline{\sigma}_\lambda\overline{z}_\lambda,\overline{u}_{D,\lambda},\overline{\ell}_\lambda\}_{\lambda>0}$ is a global solution of (P$_\lambda$) and the constant tuple $(u,\sigma,z,u_D,\ell)\equiv(u_0,\sigma_0,\nabla^s u_0-\mathbb{A}\sigma_0,u_0,0)$ is feasible for (P$_\lambda$), we obtain

$$J_\lambda(\overline{u}_\lambda,\overline{u}_{D,\lambda},\overline{\ell}_\lambda)\le J_\lambda(u_0,u_0,0)=\frac{\alpha}{2}\|u_0\|^2_{L^2(\mathbf{H}^2(\Omega))}=:C<\infty. \tag{6.9}$$

Since all $\overline{u}_\lambda$ share the same initial value and due to the special structure of the objective in (6.1), this implies that $\overline{u}_\lambda$ satisfies the boundedness assumption in (4.3) such that Lemma 4.5 yields the existence of a subsequence converging weakly in $\mathcal{U}$. Moreover, the inequality constraint on the stress and the $H^1(H^2)$-norm in the objective immediately yield the boundedness of $\overline{\sigma}_\lambda$ and $\overline{u}_{D,\lambda}$ in their respective spaces, and the reflexivity of the latter imply the existence of a weakly convergent subsequence.

Step 2. Feasibility of an accumulation point. Let us now assume that a given subsequence of $\{\overline{u}_\lambda,\overline{\sigma}_\lambda,\overline{u}_{D,\lambda}\}_{\lambda>0}$, denoted by the same symbol for simplicity, converges weakly to $(\tilde u,\tilde\sigma,\tilde u_D)$ in $\mathcal{U}\times H^1(\mathbb{L}^2(\Omega))\cap L^s(\mathbb{W}^{1,p}(\Omega))\times H^1(\mathbf{H}^2(\Omega))$. By the compact embedding of $H^1(\mathbf{H}^2(\Omega))$ in $C(\mathbf{H}^1(\Omega))$, this ensures the convergence properties required in (6.3) and in addition $\tilde u_D(0)-u_0\in\mathbf{H}^1_D(\Omega)$. Moreover, the assumptions on p and s in (6.2) guarantee that $H^1(\mathbb{L}^2(\Omega))\cap L^s(\mathbb{W}^{1,p}(\Omega))$ embeds compactly in $L^2(C(\bar\Omega;\mathbb{R}^{n\times n}_{\rm sym}))$, as already mentioned above, so that (6.5) is valid. Furthermore, considering again (6.9), we see that $\lambda^{-\theta}\|\overline{\ell}_\lambda\|_{L^2(\mathbf{H}_D^{-1}(\Omega))}$ is bounded, hence, $\overline{\ell}_\lambda\to\ell=0$ in $L^2(\mathbf{H}_D^{-1}(\Omega))$ (even with strong convergence). Altogether, we observe that the convergence properties in (6.3)–(6.5) are fulfilled such that Proposition 6.2 yields that the weak accumulation point $(\tilde u,\tilde\sigma)$ is a weak solution associated with $\tilde u_D$ and therefore feasible for the original optimization problem (P).

Step 3. Construction of a recovery sequence. First, observe that, since $\overline{u}$ is assumed to be in $H^1(\mathbf{H}^1(\Omega))$ and to satisfy the Dirichlet boundary conditions, Corollary 3.5 gives that $(\overline{\sigma},\overline{u})$ is a strong solution associated with $\overline{u}_D$.

The recovery sequence for $(\overline{u},\overline{\sigma},\overline{u}_D)$ is constructed based on our findings in section 5. To be more precise, we apply Lemma 5.3 and Lemma 5.5 with $w=\nabla^s\dot{\overline{u}}$. According to these lemmas, $\sigma_\lambda\in H^1(\mathbb{L}^2(\Omega))$ defined as unique solution of

$$\nabla^s\dot{\overline{u}}-\mathbb{A}\dot\sigma_\lambda=\partial I_\lambda(\sigma_\lambda),\qquad \sigma_\lambda(0)=\sigma_0,$$

satisfies the bound in (5.7) and converges strongly in $H^1(\mathbb{L}^2(\Omega))$ to σ, which is the solution to

$$\nabla^s\dot{\overline{u}}-\mathbb{A}\dot\sigma\in\partial I_{\mathcal{K}(\Omega)}(\sigma),\qquad \sigma(0)=\sigma_0.$$

This equation is just the strong form of the flow rule in (3.4b). The monotonicity of $\partial I_{\mathcal{K}(\Omega)}$ immediately gives that (3.4b) is uniquely solvable. Therefore, the limit σ coincides with $\overline{\sigma}$, i.e., the stress associated with $\overline{u}_D$. If we now define

$$z_\lambda:=\nabla^s\overline{u}-\mathbb{A}\sigma_\lambda\in H^1(\mathbb{L}^2(\Omega))\quad\text{and}\quad \ell_\lambda:=-\operatorname{div}\sigma_\lambda\in H^1(\mathbf{H}_D^{-1}(\Omega)),$$

then we observe that $(\overline{u}, \sigma_\lambda, z_\lambda)$ is the solution of the regularized plasticity system in (5.1) w.r.t. $\overline{u}_D$ and ℓ_λ. In addition, we have $\ell_\lambda(0) = -\operatorname{div}\sigma_0 = 0$ and $\overline{u}_D(0) - u_0 = \overline{u}_D(0) - \overline{u}(0) \in \mathbf{H}^1_D(\Omega)$. Therefore, since σ_λ satisfies the bounds in (5.7) and (4.1) (by Lemma 4.1), $(\overline{u}, \sigma_\lambda, z_\lambda, \overline{u}_D, \ell_\lambda)$ satisfies all constraints in (P_λ).

Next we show the convergence of the objective functional. As $\overline{\sigma}$ fulfills the equilibrium condition, i.e., $\overline{\sigma} \in \mathcal{E}(\Omega)$, the convergence of σ_λ by Lemma 5.5 implies

$$\ell_\lambda = -\operatorname{div}\sigma_\lambda \to -\operatorname{div}\overline{\sigma} = 0 \quad \text{in } H^1(\mathbf{H}_D^{-1}(\Omega)).$$

Furthermore, (5.12) gives

$$\begin{aligned}
\lambda^{-\theta}\|\ell_\lambda\|^2_{L^2(\mathbf{H}_D^{-1}(\Omega))} &= \lambda^{-\theta}\,\|\operatorname{div}\sigma_\lambda - \operatorname{div}\overline{\sigma}\|^2_{L^2(\mathbf{H}_D^{-1}(\Omega))} \\
&\le C\,\lambda^{-\theta}\,\|\sigma_\lambda - \overline{\sigma}\|^2_{L^2(\mathbb{L}^2(\Omega))} \\
&\le C\,\lambda^{1-\theta}\,\|\nabla^s\dot{\overline{u}} - \mathbb{A}\dot{\overline{\sigma}}\|^2_{L^2(\mathbb{L}^2(\Omega))} \to 0 \quad \text{as } \lambda \searrow 0.
\end{aligned}$$

To summarize, we found that $(\overline{u}, \sigma_\lambda, z_\lambda, \overline{u}_D, \ell_\lambda)$ is feasible for (P_λ) and fulfills

$$J_\lambda(\overline{u}, \overline{u}_D, \ell_\lambda) \to J(\overline{u}, \overline{u}_D). \tag{6.10}$$

Step 4. Strong convergence and global minimizer. The feasibility and the convergence of the recovery sequence and the optimality of $(\overline{u}_\lambda, \overline{u}_{D,\lambda}, \overline{\ell}_\lambda)$ give

$$\begin{aligned}
J(\tilde{u}, \tilde{u}_D) &\le \liminf_{\lambda\searrow 0} J(\overline{u}_\lambda, \overline{u}_{D,\lambda}) \\
&\le \limsup_{\lambda\searrow 0} J(\overline{u}_\lambda, \overline{u}_{D,\lambda}) \\
&\le \limsup_{\lambda\searrow 0} J_\lambda(\overline{u}_\lambda, \overline{u}_{D,\lambda}, \overline{\ell}_\lambda) \le \limsup_{\lambda\searrow 0} J_\lambda(\overline{u}, \overline{u}_D, \ell_\lambda) = J(\overline{u}, \overline{u}_D),
\end{aligned} \tag{6.11}$$

which, together with the feasibility of $(\tilde{u}, \tilde{\sigma}, \tilde{u}_D)$ for (P) shown in step 2, implies that $(\tilde{u}, \tilde{\sigma}, \tilde{u}_D)$ is a global minimizer of (P).

To show the strong convergence in (6.7) and (6.8), we first observe that (6.11) yields $J(\overline{u}_\lambda, \overline{u}_{D,\lambda}) \to J(\tilde{u}, \tilde{u}_D)$, from which we deduce the convergence of the norms $\|\dot{\overline{u}}_\lambda\|_{L^2(\mathbf{L}^1(\Omega))}$ and $\|\overline{u}_{D,\lambda}\|_{H^1(\mathbf{H}^2(\Omega))}$ to $\|\dot{\tilde{u}}\|_{L^2(\mathbf{L}^1(\Omega))}$ and $\|\tilde{u}_D\|_{H^1(\mathbf{H}^2(\Omega))}$, respectively. Since both norms are Kadec norms and we already have weak convergence in the respective spaces, this implies (6.7). Similarly, (6.11) yields $\|\overline{\ell}_\lambda\|_{H^1(\mathbf{H}_D^{-1}(\Omega))} \to 0$. Finally, the strong convergence of the stresses follows from the compact embedding of $H^1(\mathbb{L}^2(\Omega)) \cap L^s(\mathbb{W}^{1,p}(\Omega))$ in $L^2(C(\bar{\Omega}; \mathbb{R}^{n\times n}_{\text{sym}}))$, already used above. ∎

Some comments concerning our approximation result are in order:

Remark 6.4 (Crucial regularity assumption). The assumption of existence of a global minimizer $(\overline{u}, \overline{\sigma}, \overline{u}_D)$ with the properties listed in Theorem 6.3 is admittedly very restrictive. Notice in particular that the regularity assumptions on $\overline{u}$ imply that $(\overline{u}, \overline{\sigma})$ is a *strong solution* w.r.t. $\overline{u}_D$, whose existence can in general not be guaranteed. The regularity assumption however seems to be indispensable, as the above proof demonstrates: In order to pass to the limit in the flow rule inequality to show the feasibility of an accumulation point in step 2 of the proof, we need the additional regularity of the stress ensured by the inequality constraint in (P_λ). The generic regularity of the stress, which is $H^1(\mathbb{L}^2(\Omega))$ (see Lemma 4.1), is by far not sufficient for this passage to the limit. It therefore appears to be unavoidable to enforce the required regularity by additional inequality constraints in (P_λ). The elements of the recovery sequence however have to be feasible for (P_λ) and thus have to fulfill this inequality constraint, too. As the generic regularity of the stress is $H^1(\mathbb{L}^2(\Omega))$, it is not possible to guarantee this constraint to be fulfilled without further hypotheses on the recovery sequence and its limit, respectively. At the end, this leads to the regularity assumption on $\nabla^s\dot{\overline{u}}$.

We however emphasize that we do not require the existence of a strong solution with the addition regularity of the strain rate for every Dirichlet displacement $u_D \in H^1(\mathbf{H}^2(\Omega))$ (which would really be unrealistic), but only for one optimal $\overline{u}_D$. (Of course, there might be many optimal solutions, since (P) is a non-convex problem). Whether an optimal solution fulfilling these regularity assumptions exists or not, clearly depends on the data, especially on the smoothness of the desired strain rate μ in (6.1).

Remark 6.5 (Extensions and modifications of the approximation result).

(i) One essential drawback of the approximation result is that the bound R given in (6.6) depends on the unknown solution $(\overline{u}, \overline{u}_D)$ and is therefore in general unknown, too. One could replace the inequality constraints on the stress involving this bound in (P_λ) by an additional tracking term in the objective of the form $\|\sigma - \sigma_d\|^2_{H^1(\mathbb{L}^2(\Omega))} + \|\sigma - \sigma_d\|_X$ with a given desired stress distribution σ_d and a reflexive Banach space X with the following properties: On the one hand, $H^1(\mathbb{L}^2(\Omega)) \cap X$ should compactly embed in $L^2(C(\bar{\Omega}; \mathbb{R}^{n\times n}_{\text{sym}}))$. On the other hand, $H^1(\mathbb{L}^2(\Omega)) \cap L^\infty(\mathbb{W}^{1,p}(\Omega))$ should compactly be embedded in X. Provided these embeddings hold, the steps 2 and 3 of the previous proof can easily be adapted. At this point, one benefits from the strong convergence of the recovery sequence in $H^1(\mathbb{L}^2(\Omega))$ by Lemma 5.5.

(ii) The above analysis is restricted to objectives of the type (6.1) or other types of objectives ensuring the boundedness of $\{\overline{u}_\lambda\}$ in $\mathcal{U}$. This bound cannot be deduced from the regularized plasticity system in (5.1) unless the loads fulfill a *safe load condition*, see [30]. One could thus allow for more general objectives, if a safe load condition would be included in the set of constraints in (P_λ). We were however not able to find a safe load condition that is satisfied by the loads associated with the recovery sequence. This is due to several reasons, among these a lack of regularity of the recovery sequence. This issue is subject to future research.

(iii) By contrast, it is well possible to consider objectives, which give the boundedness of the displacement in more regular spaces such as $H^1(\mathbf{H}^1(\Omega))$. In this case, the inequality constraints on the stress in (P_λ) can be weakened or even be completely left out, since the higher regularity of the displacement enables the passage to the limit at the end of the proof of Lemma 5.5. Such a setting is treated in [35].

(iv) We have chosen the space $H^1(\mathbf{H}^2(\Omega))$ as the control space for the Dirichlet displacement in order to guarantee the compact embeddings in step 2 of the above proof and in the proof of Theorem 4.7. Of course, one might want to avoid the $\mathbf{H}^2(\Omega)$-norm in the objective, which could be achieved by an additional (pseudo-)force-to-Dirichlet-map, for example by solving an additional linear elasticity system. This strategy was employed in [21, Subsection 6.1].

Remark 6.6 (Numerical treatment of (P_λ)). Although they are still nonsmooth optimization problems, the regularized problems in (P_λ) offer ample possibilities for a numerical treatment. A popular strategy is to further regularize the problem by smoothing the Yosida approximation ∂I_λ. This has been used for the numerical computations in the companion paper [21]. Moreover, the non-smooth objective in (P_λ) calls for an additional regularization of the L^1-norms for instance in terms of a Huber-regularization. In this way, one ends up with a smooth optimal control problem, which can be treated by the classical adjoint approach. Our convergence result in Theorem 6.3 implies that, under the certainly restrictive assumptions of this theorem, there is an optimal solution of the original optimization problem governed by the perfect plasticity system that can be approximated by this procedure.

Acknowledgments. We are very grateful to Hannes Meinlschmidt (RICAM Linz) for various fruitful discussions on regularity issues associated with the construction of the recovery sequence.

Appendix A. Auxiliary results.

LEMMA A.1. *Let $\Omega \subset \mathbb{R}^n$ be a bounded domain and X be a Banach space and $A \subset X$ be a convex and closed set with $0 \in A^\circ$. Set $\mathcal{A}(\Omega) := \{v \in L^2(\Omega; X) : v \in A \text{ a.e. in } \Omega\}$. Then $C_c^\infty(\Omega; X) \cap \mathcal{A}(\Omega)$ is dense in $\mathcal{A}(\Omega)$.*

Proof. Let $v \in \mathcal{A}(\Omega)$ and $\varepsilon \in (0,1)$ be arbitrary. By assumption there exists $\delta > 0$ such that $\overline{B_X(0,\delta)} \subset A$. We set $\overline{v} := (1-\varepsilon)v$ and select a sequence $\{v_n\}_{n\in\mathbb{N}} \subset C_c^\infty(\Omega; X)$ such that

$$\|v_n - \overline{v}\|^2_{L^2(\Omega;X)} \leq \frac{\delta^2 \varepsilon^3}{4n} \quad \forall\, n \in \mathbb{N}. \tag{A.1}$$

We moreover define

$$S_n^c := \{x \in \Omega : v_n(x) \in X \setminus A^\circ\}, \quad S_n^o := \{x \in \Omega : v_n(x) \in X \setminus (1 - \tfrac{\varepsilon}{2})A\}.$$

Hence, $S_n^c \subset S_n^o$ and, by continuity and compact support of v_n, S_n^c is compact, while S_n^o is open. Thus, for every $n \in \mathbb{N}$, there is a function $\varphi_n \in C^\infty(\mathbb{R}^n; [0,1])$ with $\varphi_n \equiv 1$ in $\mathbb{R}^n \setminus S_n^o$ and $\varphi_n \equiv 0$ in S_n^c. Furthermore, if $\|v_n(x) - \overline{v}(x)\|_X \leq \frac{\varepsilon}{2}\,\delta$, then the convexity of A and $\overline{B_X(0,\delta)} \subset A$ imply

$$\frac{v_n(x)}{1-\frac{\varepsilon}{2}} = \frac{1-\varepsilon}{1-\frac{\varepsilon}{2}}\, v(x) + \Big(1 - \frac{1-\varepsilon}{1-\frac{\varepsilon}{2}}\Big)\,\frac{2}{\varepsilon}\,\big(v_n(x) - \overline{v}(x)\big) \in A$$

Therefore, we obtain by contraposition that

$$\|v_n - \overline{v}\|^2_{L^2(\Omega;X)} \geq \int_{S_n^o} \|v_n - \overline{v}\|^2_X \mathrm{d}x \geq \frac{\varepsilon^2}{4}\,\delta^2\,|S_n^o|$$

so that (A.1) yields $|S_n^c| \leq |S_n^o| \leq \varepsilon/n$. Thus, due to Lebesgue's dominated convergence theorem, there exists $N = N(\varepsilon) \in \mathbb{N}$ such that

$$\|\overline{v}\|_{L^2(S_N^o;X)} \leq \|v\|_{L^2(S_N^o;X)} \leq \varepsilon. \tag{A.2}$$

Now we define $v_s := \varphi_N\, v_N$. Then, by construction $v_s \in \mathcal{A}(\Omega) \cap C_c^\infty(\Omega; X)$ and, in addition, (A.1) and (A.2) imply

$$\begin{aligned}
\|v - v_s\|_{L^2(\Omega;X)} &\leq \|v - \overline{v}\|_{L^2(\Omega;X)} + \|\overline{v} - v_N\|_{L^2(\Omega;X)} + \|v_N - v_s\|_{L^2(\Omega;X)} \\
&\leq \varepsilon\|v\|_{L^2(\Omega;X)} + \|\overline{v} - v_N\|_{L^2(\Omega;X)} + \|v_N\|_{L^2(S_N^o;X)} \\
&\leq \varepsilon\|v\|_{L^2(\Omega;X)} + 2\,\|\overline{v} - v_N\|_{L^2(\Omega;X)} + \|\overline{v}\|_{L^2(S_N^o;X)} \\
&\leq \varepsilon\Big(\|v\|_{L^2(\Omega;X)} + \frac{\delta\sqrt{\varepsilon}}{\sqrt{N}} + 1\Big).
\end{aligned}$$

Since ε was arbitrary, this finishes the proof. ∎

REFERENCES

[1] L. Adam and J. Outrata, *On optimal control of a sweeping process coupled with an ordinary differential equation*, Discrete Contin. Dyn. Syst. Ser. B, 19 (2014), pp. 2709–2738, https://doi.org/10.3934/dcdsb.2014.19.2709, https://doi.org/10.3934/dcdsb.2014.19.2709.

[2] C. E. Arroud and G. Colombo, *A maximum principle for the controlled sweeping process*, Set-Valued Var. Anal., 26 (2018), pp. 607–629, https://doi.org/10.1007/s11228-017-0400-4, https://doi.org/10.1007/s11228-017-0400-4.

[3] L. Boron and E. Zeidler, *Nonlinear Functional Analysis and its Applications: III: Variational Methods and Optimization*, Springer New York, 1984, https://books.google.de/books?id=jpcnAQAAIAAJ.

[4] M. Brokate, *Optimale Steuerung von gewöhnlichen Differentialgleichungen mit Nichtlinearitäten vom Hysteresis-Typ*, vol. 35 of Methoden und Verfahren der Mathematischen Physik [Methods and Procedures in Mathematical Physics], Verlag Peter D. Lang, Frankfurt am Main, 1987.

[5] M. Brokate and P. Krejčí, *Optimal control of ODE systems involving a rate independent variational inequality*, Discrete Contin. Dyn. Syst. Ser. B, 18 (2013), pp. 331–348, https://doi.org/10.3934/dcdsb.2013.18.331, https://doi.org/10.3934/dcdsb.2013.18.331.

[6] M. Brokate and P. Krejčí, *Weak differentiability of scalar hysteresis operators*, Discrete Contin. Dyn. Syst., 35 (2015), pp. 2405–2421, https://doi.org/10.3934/dcds.2015.35.2405, https://doi.org/10.3934/dcds.2015.35.2405.

[7] G. Colombo, R. Henrion, N. D. Hoang, and B. S. Mordukhovich, *Optimal control of the sweeping process over polyhedral controlled sets*, J. Differential Equations, 260 (2016), pp. 3397–3447, https://doi.org/10.1016/j.jde.2015.10.039, https://doi.org/10.1016/j.jde.2015.10.039.

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

[9] R. Edwards, *Functional analysis: Theory and applications, holt, rinehart and winston, new york, 1965*, MR, 36 (1994), p. 4308.

[10] M. Eleuteri and L. Lussardi, *Thermal control of a rate-independent model for permanent inelastic effects in shape memory materials*, Evol. Equ. Control Theory, 3 (2014), pp. 411–427, https://doi.org/10.3934/eect.2014.3.411, https://doi.org/10.3934/eect.2014.3.411.

[11] M. Eleuteri, L. Lussardi, and U. Stefanelli, *Thermal control of the Souza-Auricchio model for shape memory alloys*, Discrete Contin. Dyn. Syst. Ser. S, 6 (2013), pp. 369–386, https://doi.org/10.3934/dcdss.2013.6.369, https://doi.org/10.3934/dcdss.2013.6.369.

[12] G. A. Francfort and A. Giacomini, *Small-strain heterogeneous elastoplasticity revisited*, Communications on Pure and Applied Mathematics, 65 (2012), pp. 1185–1241.

[13] G. A. Francfort, A. Giacomini, and J.-J. Marigoc, *The elasto-plastic exquisite corpse: A suquet legacy*, Journal of the Mechanics and Physics of Solids, 97 (2016), pp. 125–139.

[14] T. Geiger and D. Wachsmuth, *Optimal control of an evolution equation with non-smooth dissipation*. arXiv:1801.04077, 2018.

[15] K. Gröger, *A $W^{1,p}$-estimate for solutions to mixed boundary value problems for second order elliptic differential equations*, Math. Ann., 283 (1989), pp. 679–687, https://doi.org/10.1007/bf01442860.

[16] J. G. Heywood and R. Rannacher, *Finite-element approximation of the nonstationary navier–stokes problem. part iv: Error analysis for second-order time discretization*, SIAM Journal on Numerical Analysis, 27 (1990), pp. 353–384.

[17] C. Johnson, *Existence theorems for plasticity problems*, Journal de Mathématiques Pures et Appliquées, 55 (1976), pp. 431–444.

[18] S. May, R. Rannacher, and B. Vexler, *Error analysis for a finite element approximation of elliptic Dirichlet boundary control problems*, SIAM Journal on Control and Optimization, 51 (2013), pp. 2585–2611.

[19] H. Meinlschmidt, C. Meyer, and J. Rehberg, *Regularization for optimal control problems associated to nonlinear evolution equations*, Journal of Convex Analysis, 27 (2020), pp. 443–485.

[20] H. Meinlschmidt, C. Meyer, and S. Walther, *Optimal control of an abstract evolution variational inequality with application in homogenized plasticity*. arXiv:1909.13722, 2019.

[21] C. Meyer and S. Walther, *Optimal control of perfect plasticity part I: Stress tracking*. arXiv:2001.02969, 2020.

[22] A. Mielke and F. Rindler, *Reverse approximation of energetic solutions to rate-independent processes*, NoDEA Nonlinear Differential Equations Appl., 16 (2009), pp. 17–40, https:

//doi.org/10.1007/s00030-008-7065-5, https://doi.org/10.1007/s00030-008-7065-5.

[23] A. Mielke and T. Roubíček, *Rate-independent systems. Theory and application.*, vol. 193, New York, NY: Springer, 2015, https://doi.org/10.1007/978-1-4939-2706-7.

[24] C. Münch, *Optimal control of reaction-diffusion systems with hysteresis.* arXiv:1705.11031, 2017.

[25] N. Ottosen and M. Ristinmaa, *The Mechanics of Constitutive Modeling*, Elsevier, Amsterdam, 2005.

[26] F. Rindler, *Optimal control for nonconvex rate-independent evolution processes*, SIAM J. Control Optim., 47 (2008), pp. 2773–2794, https://doi.org/10.1137/080718711, https://doi.org/10.1137/080718711.

[27] F. Rindler, *Approximation of rate-independent optimal control problems*, SIAM J. Numer. Anal., 47 (2009), pp. 3884–3909, https://doi.org/10.1137/080744050, https://doi.org/10.1137/080744050.

[28] U. Stefanelli, *Magnetic control of magnetic shape-memory single crystals*, Physica B, 407 (2012), pp. 1316–1321.

[29] U. Stefanelli, D. Wachsmuth, and G. Wachsmuth, *Optimal control of a rate-independent evolution equation via viscous regularization*, Discrete Contin. Dyn. Syst. Ser. S, 10 (2017), pp. 1467–1485, https://doi.org/10.3934/dcdss.2017076, https://doi.org/10.3934/dcdss.2017076.

[30] P.-M. Suquet, *Sur les équations de la plasticité: existence et régularité des solutions*, J. Mécanique, 20 (1981), pp. 3–39.

[31] R. Temam, *Mathematical problems in plasticity*, Courier Dover Publications, 2018.

[32] G. Wachsmuth, *Optimal control of quasi-static plasticity with linear kinematic hardening, Part I: Existence and discretization in time*, SIAM J. Control Optim., 50 (2012), pp. 2836–2861 + loose erratum, https://doi.org/10.1137/110839187, https://doi.org/10.1137/110839187.

[33] G. Wachsmuth, *Optimal control of quasistatic plasticity with linear kinematic hardening II: Regularization and differentiability*, Z. Anal. Anwend., 34 (2015), pp. 391–418, https://doi.org/10.4171/ZAA/1546, https://doi.org/10.4171/ZAA/1546.

[34] G. Wachsmuth, *Optimal control of quasistatic plasticity with linear kinematic hardening III: Optimality conditions*, Z. Anal. Anwend., 35 (2016), pp. 81–118, https://doi.org/10.4171/ZAA/1556, https://doi.org/10.4171/ZAA/1556.

[35] S. Walther, *Optimal Control of Plasticity Systems*, PhD thesis, TU Dortmund, 2020. in preparation.

[36] W. P. Ziemer, *Weakly Differentiable Functions*, Springer New York, 1989.