Weierstraß-Institut

für Angewandte Analysis und Stochastik

Leibniz-Institut im Forschungsverbund Berlin e. V.

Density of convex intersections and applications

Michael Hintermüller[1,2], Carlos N. Rautenberg[2], Simon Rösel[2]

submitted: November 11, 2016

[1] Weierstrass Institute
    Mohrenstr. 39
    10117 Berlin
    Germany
    E-Mail: michael.hintermueller@wias-berlin.de

[2] Department of Mathematics
    Humboldt-Universität zu Berlin
    Unter den Linden 6
    10099 Berlin
    Germany
    E-Mail: hint@math.hu-berlin.de
            carlos.rautenberg@math.hu-berlin.de
            roesel@math.hu-berlin.de

No. 2333

Berlin 2016

ABSTRACT. In this paper we address density properties of intersections of convex sets in several function spaces. Using the concept of Γ-convergence, it is shown in a general framework, how these density issues naturally arise from the regularization, discretization or dualization of constrained optimization problems and from perturbed variational inequalities. A variety of density results (and counterexamples) for pointwise constraints in Sobolev spaces are presented and the corresponding regularity requirements on the upper bound are identified. The results are further discussed in the context of finite element discretizations of sets associated to convex constraints. Finally, two applications are provided, which include elasto-plasticity and image restoration problems.

1. INTRODUCTION

Convex constraint sets K as subsets of an arbitrary Banach space X are common to many fields in mathematics such as calculus of variations, variational inequalities, and control theory. Such constraints are induced by physical limitations of control and/or state variables, but also emerge through Fenchel dualization of convex problems. In this vein, given a set of functions satisfying an arbitrary constraint, density properties of more regular functions satisfying the same restriction are of utmost importance in many instances, e.g., for the study of the limiting behavior of regularized/discretized problems, the closed form determination of Fenchel dual problems, the deduction of a vanishing viscosity limit for variational inequalities, etc. In abstract terms, the density problem under consideration can be stated as follows: Given some dense subspace Y of X, the central point of interest is whether the closure property

$$(1.1) \qquad \overline{K(Y)}^X = K,$$

where $K(Y) = \{u \in Y : u \in K\} = K \cap Y$, is fulfilled.

The paper is organized as follows: Section 2 serves as a motivational framework for the density question under consideration. Here we provide two general environments where the closure property (1.1) emerges as fundamental for their study. In particular, the first setting in section 2.1 involves constrained optimization and the one in section 2.2 is associated with variational inequalities. Within these two settings, we consider regularization, Galerkin approximation, and singular perturbation, and these approaches are treated by methods of Γ-convergence.

In Section 3 we focus on the special setting where $X = X(\Omega)$ is a ($\mathbb{R}^d$-valued) vector space of functions over a bounded domain Ω of $\mathbb{R}^N$ and $K = K(X)$ denotes the subset of elements in $X(\Omega)$ bounded pointwise by a prescribed measurable function $\alpha : \Omega \to \mathbb{R} \cup \{+\infty\}$, i.e.,

$$K(X(\Omega)) = \{w \in X(\Omega) : |w(x)| \leq \alpha(x) \text{ a.e. (almost everywhere) in } \Omega\},$$

with $|\,.\,|$ denoting an $\mathbb{R}^d$-norm. Particularly in this part, $X(\Omega)$ refers to a Lebesgue or Sobolev space and $Y = Y(\Omega)$ refers to the space of continuous or infinitely differentiable functions up to the boundary. We also use the notation

$$K(X(\Omega), |\,.\,|) = \{w \in X(\Omega) : |w(x)| \leq \alpha(x) \text{ a.e. in } \Omega\},$$

whenever it is necessary to explicit the dependence on the specific norm $|\,.\,|$. This becomes useful as several results in this paper depend on specifically chosen $\mathbb{R}^d$-norms. We introduce the problem formulation and give some new density results for continuous obstacles which extents results from [15] relying on the theory of mollification. In the subsequent section 4, we focus on extensions of the density results for discontinuous obstacles. It is first shown in section 4.1

1

that results of the type (1.1) cannot be expected in general if the obstacle is a just a Sobolev function. For this purpose we provide a concrete counterexample. The density results are then extended to discontinuous obstacles which fulfill certain semi-continuity assumptions in the Lebesgue space-case in section 4.2 and in the Sobolev space-case in section 4.3. Subsequently, in section 4.4, a different approach is considered for obstacles which originate from the solution of a PDE (partial differential equation), in which case smooth approximants are constructed by solving a sequence of singularly perturbed elliptic PDEs. In Section 5, we make use of the preceding density results to prove the Mosco convergence of various types of finite element discretizations of K. We finalize the paper with section 6 by providing two concrete applications from elasto-plasticity and image restoration.

2. MOTIVATION

2.1. **Optimization with convex constraints.**
In many variational problems one seeks the solution in a given convex, closed and nonempty subset K of a Banach space $(X, \|.\|)$. To illustrate the problem, let us consider the following abstract class of optimization problems:

$$(2.1) \qquad \begin{cases} \inf & F(u), \quad \text{over } u \in X, \\ \text{s.t.} & u \in K. \end{cases}$$

We assume that $F : X \to \mathbb{R}$ is continuous, coercive and sequentially weakly lower semi-continuous but not necessarily convex. Thus, problem (2.1) admits a solution provided X is reflexive. The problem class (2.1) is ubiquitous, encompassing numerous fields, such as the variational form of partial differential equations, variational inequality problems of potential type, optimal control of partial differential equations with constraints on the state and/or control, and many other. The analysis of (2.1) and the design of suitable solution algorithms often involve the general concepts of perturbation or dualization methods comprising regularization, penalization or discretization approaches or possibly a combination of the latter. The stability properties of (2.1) with regard to a large class of perturbations rely on the closure property (1.1), i.e.,

$$\overline{K(Y)}^X = K,$$

where Y is some dense subspace of X with regard to the norm topology of X and $K(Y)$ is given by

$$K(Y) = \{u \in Y : u \in K\} = K \cap Y.$$

To justify this conjecture, we consider the following abstract perturbation class.

2.1.1. *A class of quasi-monotone perturbations.*
To subsume as many of the above mentioned methods as possible we consider the sequence of perturbed problems

$$(2.2) \qquad \inf \quad F(u) + R_n(u), \quad \text{over } u \in X,$$

defined by a given sequence of functions

$$R_n : X \to \mathbb{R} \cup \{+\infty\}, \quad n \in \mathbb{N},$$

that are perturbations of the indicator function $i_K : X \to \mathbb{R} \cup \{+\infty\}$ in the following sense: there exist functions $\underline{R_n} : X \to \mathbb{R} \cup \{+\infty\}$ and $\overline{R_n} : X \to \mathbb{R} \cup \{+\infty\}$ such that

$$0 \le \underline{R_n} \le R_n \le \overline{R_n} \quad \forall\, n \in \mathbb{N},$$

2

having the additional properties:

$$(2.3) \quad \begin{cases} \underline{R_n} \leq \underline{R_{n+1}} \;\; \forall\, n \in \mathbb{N}, \quad \lim_{n \to +\infty} \underline{R_n}(u) = i_K(u) \;\; \forall\, u \in X, \\ \underline{R_n} \text{ is sequentially weakly lower semicontinuous } \;\; \forall\, n \in \mathbb{N}, \end{cases}$$

and

$$(2.4) \quad \overline{R_n} \geq \overline{R_{n+1}}, \;\; \forall\, n \in \mathbb{N}, \quad \lim_{n \to +\infty} \overline{R_n}(u) = i_{K \cap Y}(u) \;\; \forall\, u \in X.$$

We call mappings (R_n) that share the above features *quasi-monotone perturbations of the indicator function i_K with respect to the (dense) subspace Y*. Note that no additional assumptions are made on R_n itself.

At this point we make the conjecture that the stability of (2.1) with respect to a large class of perturbations hinges on the density condition (1.1). In fact, the following result, which is based on the theory of Γ-convergence [8], substantiates this conjecture.

Under mild assumptions on X, the density property (1.1) ensures that $F + i_K$ is the Γ-limit of $(F + R_n)$ in both, the weak and strong topology.

The proof of this result is referred to the appendix (see Proposition A1). Under the assumptions of Proposition A1, one may infer that, provided each problem (2.2) admits a minimizer u_n, each weak cluster point of the sequence of minimizers (u_n) is a minimizer of (2.1), cf. [8, Corollary 7.20]. We also remark that in case the (sequential) weak and strong Γ-limits coincide, one usually uses the notion Mosco convergence.

In the following, we present a selection of approximation methods that fit into the general class of perturbations defined by (2.2) and which bear high practical relevance. In favor of generality, we do not leave the abstract setting.

Example 2.1 (Tikhonov-Regularization). Let $(Y, \| \, . \, \|_Y)$ be a Banach space which is densely and continuously embedded into X. For a sequence of positive non-decreasing parameters (γ_n) with $\gamma_n \to +\infty$ and fixed $\alpha > 0$, consider in (2.2) the Tikhonov regularization

$$(2.5) \quad R_n(u) = i_K(u) + \frac{1}{2\gamma_n} \|u\|_Y^\alpha,$$

where it is understood that $R_n(u) = +\infty$ if $u \notin Y$. In fact, set $\underline{R_n} := i_K$ for all $n \in \mathbb{N}$ and $\overline{R_n} := R_n$. Obviously, (2.3) and (2.4) are satisfied such that (R_n) fits into the context of quasi-monotone perturbations according to (2.2).

Example 2.2 (Conformal discretization). Let X be a separable Banach space. Suppose (2.1) is approximated by a Galerkin approach using nested and conformal finite-dimensional subspaces X_n, i.e., $X_n \subset X$ and $X_n \subset X_{n+1}$ for all $n \in \mathbb{N}$, such that the Galerkin approximation property

$$\overline{\bigcup_{n \in \mathbb{N}} X_n}^X = X$$

is fulfilled. The resulting discrete counterpart of problem (2.1) is given by (2.2) with $R_n(u) = i_{K \cap X_n}$. Setting $\underline{R_n} = i_K$, (2.3) is clearly fulfilled. Define $Y = \bigcup_{n \in \mathbb{N}} X_n$, then (2.4) is fulfilled with $\overline{R_n} = R_n$.

Example 2.3 (Combined Moreau-Yosida/Tikhonov regularization). Let X be a Hilbert space and $(Y, \| \, . \, \|_Y)$ be a Banach space that is densely and continuously embedded into X. For

3

two sequences of positive non-decreasing parameters $(\gamma_n), (\gamma'_n)$ with $\gamma_n, \gamma'_n \to +\infty$ and fixed $\alpha > 0$, consider the simultaneous Moreau-Yosida and Tikhonov regularization,

$$(2.6) \qquad R_n(u) = \frac{\gamma_n}{2} \inf_{v \in K} \|u - v\|^2 + \frac{1}{2\gamma'_n} \|u\|_Y^\alpha,$$

with $\alpha > 0$ fixed, where it is understood that $R_n(u) = +\infty$ if $u \notin Y$. Setting $\underline{R_n}(u) = \frac{\gamma_n}{2} \inf_{v \in K} \|u - v\|^2$, standard properties of the Moreau-Yosida regularization ensure that $\underline{R_n}$ satisfies (2.3); see, e.g., [3, Prop. 17.2.1]. Defining $\overline{R_n}(u) = i_K(u) + \frac{1}{2\gamma'_n} \|u\|_Y^\alpha$, (2.4) is verified as in the previous example.

Example 2.4 (Conformal discretization and Moreau-Yosida regularization)**.** Let X be a separable Hilbert space and (γ_n) a sequence of positive non-decreasing parameters converging to $+\infty$. The combination of regularization and discretization leads to the definition

$$(2.7) \qquad R_n(u) = \frac{\gamma_n}{2} \inf_{v \in K} \|u - v\|^2 + i_{X_n},$$

where the sequence of spaces (X_n) is defined as in Example 2.2. Setting $\underline{R_n} = \frac{\gamma_n}{2} \inf_{v \in K} \|u - v\|^2$ and $\overline{R_n} = i_{K \cap X_n}$, (2.3) and (2.4) are fulfilled with $Y = \bigcup_{n \in \mathbb{N}} X_n$ and the framework of (2.2) applies.

Consequently, each of these perturbations is stable with respect to (2.1) provided the density result (1.1) is satisfied. It should also be emphasized that these examples only represent an assorted variety of perturbations which fit into the problem class (2.2). Moreover, the density property (1.1) is also a necessary condition for the stability of perturbation schemes in the following sense: First, the Γ-limit of the approximation schemes defined in Example 2.1 and Example 2.2 can be calculated using similar arguments as in the proof of Proposition A1. In fact, under the same conditions on X, one obtains $F + i_{\overline{K \cap Y}}$ as the weak and strong Γ-limit in both cases. Secondly, in the combined approaches of Example 2.3 and Example 2.4, Proposition A1 guarantees that $F + i_K$ is obtained as the weak-strong Γ-limit for *any* coupling of regularization parameter pairs $[\gamma_n, \gamma'_n]$ and $[X_n, \gamma_n]$, respectively. Let us put this statement into a perspective by means of the combined Galerkin-Moreau-Yosida approach (Example 2.4): In this case, it is possible to prove the existence of a suitable combination of n and γ_n to recover $F + i_K$ in the Γ-limit without resorting to the density property (1.1), see [21, Prop. 2.4.6]. On the other hand, the proof is non-constructive and thus not immediately useful for the design of a stable numerical algorithm. Moreover, if (1.1) is violated, one may construct for any $x \in K \setminus \overline{K \cap Y}$ a sequence (γ_n) such that no recovery sequence exists for the element x (see Proposition A2). The analogous statement is valid for the case of combined Moreau-Yosida/Tikhonov regularizations.

2.2. **Elliptic variational inequalities.**

The density of convex intersections of the type (1.1) is also of fundamental importance for the analysis of perturbations of variational inequalities. Assuming X to be a Hilbert space and $K \subset X$ nonempty, closed and convex, we consider the general variational inequality problem of the first kind,

$$(2.8) \qquad \text{find } u \in X : \quad \langle Au, v - u \rangle + i_K(v) - i_K(u) \geq \langle l, v - u \rangle, \quad \forall v \in X.$$

Here, $l \in X^*$ is a linear, bounded operator and $A : X \to X^*$ denotes a, in general, nonlinear operator on X. We further assume A to be Lipschitz continuous and strongly monotone, i.e., there exists $\kappa > 0$ with

$$\langle Av - Au, v - u \rangle \geq \kappa \|v - u\|^2 \quad , \ \forall u, v \in X.$$

In the following, we investigate three main classes of perturbations of (2.8) and their relation to the density properties of convex intersections.

2.2.1. *Quasi-monotone approximation.* Consider the perturbed variational inequality problem,

$$(2.9) \qquad \text{find } u \in X : \quad \langle Au, v - u \rangle + R_n(v) - R_n(u) \geq \langle l, v - u \rangle, \quad \forall\, v \in X,$$

where (R_n) is a quasi-monotone perturbation of i_K with respect to a dense subspace Y of X. The stability of the approximation scheme (2.9) hinges on the density property (1.1). In fact, if the latter condition is fulfilled, then, under mild assumptions on the lower bounds $\underline{R}_n$, the sequence (R_n) Mosco converges to i_K. In this case one may invoke the results from [11, 22] to conclude the consistency of the perturbation scheme with respect to the limit problem (2.8).

2.2.2. *Galerkin approximation of variational inequalities.* In general, finite-dimensional approximations of K are neither conformal nor nested as it was the case in Example 2.2 and Example 2.4, where K was 'discretized' by $K \cap X_n$, which is numerically realizable only in special cases. Instead, it is often more favorable to consider non-nested approximations $K_n \subset X_n$ that may contain infeasible elements, such that $K_n \subset K$ does not hold true in general. The resulting finite-dimensional variational inequality problems,

$$(2.10) \qquad \text{find } u \in X : \quad \langle Au, v - u \rangle + i_{K_n}(v) - i_{K_n}(u) \geq \langle l, v - u \rangle, \quad \forall\, v \in X,$$

do not fit into the framework of quasi-monotone perturbations from (2.2). The Mosco convergence of (K_n) to K, or equivalently, the weak and strong sequential Γ-convergence of (i_{K_n}) to i_K, suffices to ensure that the approximation (2.10) is stable with respect to the limit problem (2.8); cf. [11, I, Theorem 6.2]. This property is maintained in a very general context, that is, appropriate perturbations A and l may be incorporated, and under weak monotonicity assumptions on A and its possible perturbations one may even derive strong convergence for the discrete solutions (u_n); see [22] for details. However, Mosco convergence requires the existence of a recovery sequence for any element $u \in K$. To construct this sequence in the context of finite element methods, one typically uses an interpolation procedure which is only defined on the (supposedly) dense subset $K \cap Y$ of K, where typically $Y = C^\infty(\overline{\Omega})$ or $Y = C(\overline{\Omega})$, cf. [11] and Section 5. This leads again to problem (1.1).

2.2.3. *Singular perturbations.* In the context of variational inequalities, the closure property (1.1) also plays a role in the limiting behavior of singular perturbations. In fact, let $A_1 : Y \to Y^*$ be a Lipschitz continuous and strongly monotone operator on a Hilbert space $(Y, \|\,.\,\|_Y)$ that embeds densely and continuously into X. For a sequence of regularization parameters (γ_n) with $\gamma_n \to +\infty$ consider the perturbed problems,

$$(2.11) \qquad \text{find } u_n \in K \cap Y : \quad \langle (A + \tfrac{1}{\gamma_n} A_1) u_n, v - u_n \rangle \geq \langle l, v - u_n \rangle, \quad \forall\, v \in K \cap Y.$$

Observe that problem (2.11) admits a unique solution $u_n \in K \cap Y$ provided that $K \cap Y$ is closed in Y. The appropriate limit problem is then given by,

$$(2.12) \qquad \text{find } u \in \overline{K \cap Y}^{X} : \quad \langle Au, v - u \rangle \geq \langle l, v - u \rangle, \quad \forall\, v \in \overline{K \cap Y}^{X}.$$

Note that (2.12) corresponds to the initial variational inequality problem if the density property (1.1) holds true. In this case, the sequence (u_n) converges strongly in X to the solution of (2.8). Here, the assumptions on A_1 may be alleviated. This type of application also plays a role in the analysis and the design of algorithms for hyperbolic variational inequalities through the vanishing viscosity approach. For details, [19, Section 4.9] may be consulted.

3. DENSITY RESULTS FOR CONTINUOUS OBSTACLES

We first fix some notation. In this section, $\Omega \subset \mathbb{R}^N$ denotes a bounded Lipschitz domain. The space of functions that are restrictions to Ω of smooth functions with compact support on $\mathbb{R}^N$ is denoted by $\mathcal{D}(\overline{\Omega})$,

$$\mathcal{D}(\overline{\Omega}) = \{\varphi|_\Omega : \varphi \in C_c^\infty(\mathbb{R}^N)\}.$$

The standard Lebesgue and Sobolev spaces over Ω are denoted by $L^p(\Omega), W^{1,p}(\Omega)$ and $W_0^{1,p}(\Omega)$, and we also employ the spaces

$$H(\mathrm{div};\Omega) = \{u \in L^2(\Omega;\mathbb{R}^N) : \mathrm{div}\, u \in L^2(\Omega)\},$$

and

$$H_0(\mathrm{div};\Omega) = \overline{C_c^\infty(\Omega;\mathbb{R}^N)}^{H(\mathrm{div};\Omega)} = \{u \in H(\mathrm{div};\Omega) : u \cdot \nu = 0 \text{ on } \partial\Omega\}.$$

In the recent paper [15], it has been shown that for any $\alpha \in C(\overline{\Omega})$ with

$$\tag{3.1} \operatorname*{ess\,inf}_{x\in\Omega} \alpha(x) > 0,$$

the following density result for the spaces $X(\Omega) \in \{L^p(\Omega)^d, W_0^{1,p}(\Omega)^d, H_0(\mathrm{div};\Omega)\}$, and $1 \le p < +\infty$, holds true:

$$\tag{3.2} \overline{K(C_c^\infty(\Omega)^d)}^{X(\Omega)} = K(X(\Omega)),$$

where the constraint set $K(X(\Omega))$ with respect to a given subspace

$$X(\Omega) \subset L^1(\Omega)^d$$

is defined by a pointwise constraint on an arbitrary norm $|\,.\,|$ on $\mathbb{R}^d$, i.e.,

$$K(X(\Omega)) := \{w \in X(\Omega) : |w(x)| \le \alpha(x) \text{ a.e. in } \Omega\}.$$

Here, $\alpha : \Omega \to \mathbb{R} \cup \{+\infty\}$ is a given nonnegative Lebesgue measurable function. It is further understood that $d = N$ in (3.2) if $X(\Omega) = H_0(\mathrm{div};\Omega)$.

To analyze the case without homogeneous Dirichlet boundary conditions, a small modification of the approximating sequence constructed in [15] is sufficient in order to arrive at the following statement.

Theorem 3.1. *Let $\alpha \in C(\overline{\Omega})$ fulfill (3.1) and $1 \le p < +\infty$. Then it holds that*

$$\tag{3.3} \overline{K(\mathcal{D}(\overline{\Omega})^d)}^{W^{1,p}(\Omega)^d} = K(W^{1,p}(\Omega)^d),$$

i.e., $K(\mathcal{D}(\overline{\Omega})^d)$ is dense in $K(W^{1,p}(\Omega)^d)$ with respect to the norm topology in $W^{1,p}(\Omega)^d$.

Proof. Let $w \in K(W^{1,p}(\Omega)^d)$. Since Ω is a bounded Lipschitz domain we may extend w to a function in $W^{1,p}(\mathbb{R}^N)^d$ using for each component the extension-by-reflection operator. The resulting operator

$$\tag{3.4} E : W^{1,p}(\Omega)^d \to W^{1,p}(\mathbb{R}^N)^d$$

has the properties $Ew|_\Omega = w$ for all $w \in W^{1,p}(\Omega)^d$ and $E \in \mathcal{L}(W^{1,p}(\Omega)^d, W^{1,p}(\mathbb{R}^N)^d)$; see, for instance, [2]. Since E is obtained by a partition of unity argument using local reflection with respect to the Lipschitz graphs into which $\partial\Omega$ can be decomposed, the property $|w(x)| \le \alpha(x)$ in Ω is preserved by the extension in that

$$\tag{3.5} |(Ew)(x)| \le E_{C(\overline{\Omega})}\alpha(x), \quad \text{a.e. } x \in \mathbb{R}^N,$$

where $E_{C(\overline{\Omega})} : C(\overline{\Omega}) \to C(\mathbb{R}^N)$ denotes the application of the extension by reflection procedure to bounded uniformly continuous functions, i.e., $(E_{C(\overline{\Omega})}\alpha)|_\Omega = \alpha$. Further inspecting the construction of E, it may also be observed that the support of Ew is compactly contained in $\mathbb{R}^N$. Analogously, we obtain $E_{C(\overline{\Omega})}\alpha \in C_c(\mathbb{R}^N)$. For a sequence (ρ_n) of smooth mollifiers

(3.6)
$$\rho_n(x) = n^N \rho(nx),$$

where

$$\rho \in \mathcal{D}(\mathbb{R}^N), \ \rho \geq 0, \ \rho(x) = 0 \ \text{if } |x| \geq 1, \ \int_\Omega \rho \, dx = 1,$$

we define the approximating sequence $S_n(w, \Omega)$ to w by

(3.7)
$$S_n(w, \Omega)(x) := (\rho_n * Ew)(x) = \int_{\mathbb{R}^N} Ew(y) \, \rho_n(x - y) \, dy, \quad x \in \mathbb{R}^N.$$

It is well known that

(3.8)
$$S_n(w, \Omega)|_\Omega \to w \text{ in } W^{1,p}(\Omega)^d \quad \text{as } n \to \infty,$$

and, since Ew has compact support in $\mathbb{R}^N$, it holds that $S_n(w, \Omega)|_\Omega \in \mathcal{D}(\overline{\Omega})^d$. In order to achieve feasibility, we use the scaling sequence

$$\beta_n := \left(1 + \frac{\sup_{x \in \mathbb{R}^N} |\alpha_n(x) - E_{C(\overline{\Omega})}\alpha(x)|}{\min_{x \in \overline{\Omega}} \alpha(x)}\right)^{-1},$$

where $\alpha_n(x) := ((E_{C(\overline{\Omega})}\alpha) * \rho_n)(x), \ x \in \mathbb{R}^N$. Since $E_{C(\overline{\Omega})}\alpha \in C_c(\mathbb{R}^N)$, α_n converges to $E_{C(\overline{\Omega})}\alpha$ uniformly in $\mathbb{R}^N$ and thus $\beta_n \to 1$ as $n \to \infty$. In addition, (3.5) together with (3.7) yields $|S_n(w, \Omega)| \leq \alpha_n(x)$ for $x \in \mathbb{R}^N$ such that

(3.9)
$$\beta_n^{-1}\alpha(x) = \alpha(x) + \frac{\sup_{x \in \mathbb{R}^N} |\alpha_n(x) - E_{C(\overline{\Omega})}\alpha(x)|}{\min_{x \in \overline{\Omega}} \alpha(x)}\alpha(x) \geq \alpha_n(x) \geq |S_n(w, \Omega)|,$$

for all $x \in \Omega$. As a result, $\beta_n S_n(w, \Omega) \in K(\mathcal{D}(\overline{\Omega})^d)$ and, taking account of (3.8), the proof is accomplished. $\square$

Remark 3.2. In order to incorporate a homogeneous Dirichlet boundary condition in the context of Theorem 3.1, one may use an additional reparametrization to construct a suitable approximating sequence; see [15].

4. Density results for discontinuous obstacles

4.1. **A counterexample for obstacles in Sobolev spaces.** Note that Theorem 3.1 requires continuous obstacles. In some applications, such as in the regularization and discretization of elasto-plastic contact problems or image restoration problems (see section 6), it may be useful to consider obstacles that are not continuous. Under such circumstances, the following example shows that density properties of the type (3.2) or (3.3) cannot be expected if the obstacle is just a Sobolev function: Without loss of generality, assume that $0 \in \Omega \subset \mathbb{R}^N$ with $N \geq 2$ and denote by

$$B_\varepsilon(x) := \{y \in \mathbb{R}^N : |x - y|_2 \leq \varepsilon\},$$

the open ball with center $x \in \mathbb{R}^N$ and radius $\varepsilon > 0$ with respect to the Euclidean norm $|\,.\,|_2$ in $\mathbb{R}^N$. Let $\{x_k : k \in \mathbb{N}\}$ be a countable dense subset, i.e.,

$$\overline{\{x_k : k \in \mathbb{N}\}} = \overline{\Omega},$$

and $r > 0$ such that $B_r(0) \subset \Omega$. Consider the function

$$\varphi(x) := \tilde{\varphi}(x) \cdot \ln(\ln(c\,|x|_2^{-1})), \quad c \geq \mathrm{e}r \text{ fixed}, \tag{4.1}$$

where $\tilde{\varphi} \in C_c^\infty(B_r(0))$ is a smooth cut-off function with $\tilde{\varphi}(x) \geq 0$ for all $x \in B_r(0)$ and $\tilde{\varphi} \equiv 1$ on $B_{r/2}(0)$. We note that φ is nonnegative with a singularity at the origin, and its zero extension belongs to $W^{1,N}(\mathbb{R}^N)$; cf. [1, Example 4.43]. Further set

$$g(x) := \sum_{k=1}^\infty k^{-2}\varphi(x - x_k), \quad x \in \Omega, \tag{4.2}$$

and note that $g \in W^{1,N}(\Omega)$ with g being unbounded at each x_k; see [10, p.247, Example 4]. Further take a function $\phi \in C^1(\mathbb{R})$ with $0 \leq \phi(t) < 1$, $\phi(t) \to 1$ for $t \to +\infty$ and ϕ' uniformly bounded in $\mathbb{R}$. By the chain rule for Sobolev functions, the obstacle

$$\alpha := 2 - \phi \circ g \tag{4.3}$$

belongs to $W^{1,N}(\Omega)$; see, e.g., [20, Lemma A.3]. Notice also that α is bounded away from zero and that it is basically equal to 1 on the dense set $\{x_k : k \in \mathbb{N}\}$. Consequently, any continuous function w with $w \leq \alpha$ a.e. in Ω fulfills $w \leq 1$ on Ω:

Assume that the latter implication is false. Then there exist $k_0 \in \mathbb{N}$ as well as $\mu > 0, \delta > 0$ such that

$$w(x) \geq 1 + \mu \quad \forall\, x \in B_\delta(x_{k_0}). \tag{4.4}$$

Let $R > 0$ be such that $\phi(t) \geq 1 - \frac{\mu}{2}$ for all $t \geq R$. By continuity, there also exists $\delta' > 0$ such that $\varphi(x - x_{k_0}) \geq Rk_0^2$ a.e. in $B_{\delta'}(x_{k_0})$ such that

$$g(x) \geq k_0^{-2}\varphi(x - x_{k_0}) \geq R, \quad \text{a.e. } x \in B_{\delta'}(x_{k_0}),$$

which implies

$$w(x) \leq \alpha(x) = 2 - \phi(g(x)) \leq 1 + \tfrac{\mu}{2}, \quad \text{a.e. } x \in B_{\delta'}(x_{k_0}),$$

contradicting (4.4). Hence, any sequence of continuous functions approximating α from below is bounded above by 1. However, as $\alpha(x) > 1$ for a.e. $x \in \Omega$ by definition, and convergence in the norm topology of $L^p(\Omega)$ implies convergence pointwise a.e. (along a subsequence), we obtain that

$$\alpha \in K(L^p(\Omega)) \setminus \overline{K(C(\Omega) \cap L^p(\Omega))}^{L^p(\Omega)}, \tag{4.5}$$

for any $1 \leq p \leq +\infty$, and

$$\alpha \in K(W^{1,p}(\Omega)) \setminus \overline{K(C(\Omega) \cap W^{1,p}(\Omega))}^{W^{1,p}(\Omega)}, \tag{4.6}$$

for all $p \leq N$, where α is defined by (4.3).

Remark 4.1 (Complements on the counterexample)**.** An interesting point in the preceding counterexample is the structure of the set of singularities $\mathcal{S}$ where $g(x)$ is not well-defined as a real number by the infinite sum (4.2) if φ from (4.1) is understood as a function in $C(\Omega \setminus \{0\})$. Extending φ to Ω by setting $\varphi(0) := +\infty$, we obtain $g(x_k) = +\infty$ for all $k \in \mathbb{N}$ and, understanding $g : \Omega \to \mathbb{R} \cup \{+\infty\}$ as an extended real-valued function, we arrive at the following definition:

$$\mathcal{S} := \{x \in \Omega : g(x) = +\infty \text{ with } g(x) \text{ defined by (4.2) where } \varphi(0) = +\infty\}.$$

By definition, the set $\{x_k : k \in \mathbb{N}\}$ is contained in $\mathcal{S}$. Besides, it is certain that $\mathcal{S}$, and then the points where the infinite series does not converge, must have measure zero. On the other hand, $\mathcal{S}$ is in a certain sense much "bigger" than $\{x_k : k \in \mathbb{N}\}$. First observe that the set $\{x_k : k \in \mathbb{N}\}$ is strictly contained in $\mathcal{S}$. Otherwise, the concrete representative of α from (4.3) given by

$$\alpha(x) = \begin{cases} 1, & \text{on } \{x_k : k \in \mathbb{N}\} \\ 2 - \phi(g(x)), & \text{on } \Omega \setminus \{x_k : k \in \mathbb{N}\} \end{cases}$$

would define a real-valued function that is continuous on $\{x_k : k \in \mathbb{N}\}$. By the density property of $\{x_k : k \in \mathbb{N}\}$ in Ω and the fact that $\alpha(x) > 1$ for all $x \notin \{x_k : k \in \mathbb{N}\}$, it is discontinuous on the complement $\Omega \setminus \{x_k : k \in \mathbb{N}\}$. This represents a contradiction to the Baire category theorem. In the same way, one can show that the set $\mathcal{S}$ is nonmeager, i.e., it cannot be expressed as the countable union of nowhere dense subsets of $\mathbb{R}^N$. In the literature that relates to the Baire category theorem, a nonmeager set is often called of second category. To substantiate this claim, we define the nested sets

$$\mathcal{S}_n := \bigcup_{k \in \mathbb{N}} B_{ce^{-e^{nk^2}}}(x_k) \cap \Omega,$$

which consist of the union of open balls with diminishing radius around the points x_k. It can be verified that

$$\bigcap_{n \in \mathbb{N}} \mathcal{S}_n \subset \mathcal{S}.$$

In fact, let $x \in \bigcap_{n \in \mathbb{N}} \mathcal{S}_n$ and $n \in \mathbb{N}$ arbitrary. By definition, there exists an index k_0 with $x \in B_{ce^{-e^{nk_0^2}}}(x_{k_0}) \cap \Omega$. Hence,

$$g(x) = \sum_{k=1}^{\infty} k^{-2} \varphi(x - x_k) \geq k_0^{-2} \ln \ln(c|x - x_{k_0}|_2^{-1}|) \geq n.$$

Letting $n \to \infty$ yields $g(x) = +\infty$ and thus $x \in \mathcal{S}$. On the other hand, we observe that any complement $\mathcal{S}_n^c$ is closed in $\mathbb{R}^N$. Moreover, writing Ω as a countable union of closed sets $G_j \subset \mathbb{R}^N$, one obtains that

$$\left(\bigcap_{n \in \mathbb{N}} \mathcal{S}_n \right)^c \cap \Omega = \bigcup_{n \in \mathbb{N}} (\mathcal{S}_n^c \cap \Omega) = \bigcup_{j,n \in \mathbb{N}} (\mathcal{S}_n^c \cap G_j).$$

Since all sets $\mathcal{S}_n$ contain the dense set $\{x_k : k \in \mathbb{N}\}$, $\mathcal{S}_n^c \cap G_j$ also has empty interior. Therefore the complement of $\bigcap_{n \in \mathbb{N}} \mathcal{S}_n$ in Ω is meager, or, in other words, of first category. The Baire category theorem implies that $\bigcap_{n \in \mathbb{N}} \mathcal{S}_n$, and thus $\mathcal{S}$, is nonmeager.

We summarize the preceding results on general discontinuous obstacles in the following theorem.

Theorem 4.2. *Let $\Omega \subset \mathbb{R}^N$ be a bounded Lipschitz domain. The following density results hold true:*

(i) Let $N \geq 2$ and $1 \leq p \leq +\infty$. Then there exists an obstacle $\alpha \in W^{1,N}(\Omega) \cap L^\infty(\Omega)$ satisfying (3.1) such that

$$\overline{K(C(\Omega) \cap L^p(\Omega))}^{L^p(\Omega)} \subsetneq K(L^p(\Omega)),$$

the inclusion being strict.

(ii) Let $N \geq 2$ and $1 \leq p \leq N$. Then there exists an obstacle $\alpha \in W^{1,N}(\Omega) \cap L^{\infty}(\Omega)$ satisfying (3.1) such that

$$\overline{K(C(\Omega) \cap W^{1,p}(\Omega))}^{W^{1,p}(\Omega)} \subsetneq K(W^{1,p}(\Omega)),$$

the inclusion being strict.

(iii) Let $N < p < +\infty$ or $p = N = 1$. For any measurable obstacle function $\alpha : \Omega \to \mathbb{R} \cup \{+\infty\}$ which satisfies (3.1), it holds that

$$\overline{K(\mathcal{D}(\overline{\Omega})^d)}^{W^{1,p}(\Omega)^d} = K(W^{1,p}(\Omega)^d).$$

Proof. We only prove assertion (iii) since (i) and (ii) follow immediately from (4.5) and (4.6). As a consequence of the Sobolev imbedding theorem, any $w \in K(W^{1,p}(\Omega)^d)$ is contained in $C(\overline{\Omega})^d$. Let $w \in K(W^{1,p}(\Omega)^d)$. Setting

$$\hat{\alpha}(x) = \max(|w(x)|, \operatorname*{ess\,inf}_{x \in \Omega} \alpha(x)),$$

it follows that $|w(x)| \leq \hat{\alpha}(x)$ a.e. in Ω. Since $\hat{\alpha} \in C(\overline{\Omega})$ and (3.1) holds with $\hat{\alpha}$ instead of α, we may invoke Theorem 3.1 to infer that there exists a sequence (w_n) with $w_n \in \mathcal{D}(\overline{\Omega})^d$, $w_n \to w$ in $W^{1,p}(\Omega)^d$ and $|w_n(x)| \leq \hat{\alpha}(x) \leq \alpha(x)$ a.e. in Ω. This entails that $w_n \in K(\mathcal{D}(\overline{\Omega})^d)$ for all $n \in \mathbb{N}$, which accomplishes the proof. $\qquad \square$

We immediately infer the corresponding statements for Sobolev spaces incorporating homogeneous Dirichlet boundary conditions.

Corollary 4.3. *Let $\Omega \subset \mathbb{R}^N$ be a bounded Lipschitz domain. The following density results hold true:*

(i) *Let $N \geq 2$ and $p \leq N$. Then there exists an obstacle $\alpha \in W^{1,N}(\Omega) \cap L^{\infty}(\Omega)$ satisfying (3.1) such that*

$$\overline{K(C(\Omega) \cap W_0^{1,p}(\Omega))}^{W_0^{1,p}(\Omega)} \subsetneq K(W_0^{1,p}(\Omega)),$$

the inclusion being strict.

(ii) *Let $N < p < +\infty$ or $p = N = 1$. For any measurable obstacle function $\alpha : \Omega \to \mathbb{R} \cup \{+\infty\}$ which satisfies (3.1) it holds that*

$$\overline{K(C_c^{\infty}(\Omega)^d)}^{W_0^{1,p}(\Omega)^d} = K(W_0^{1,p}(\Omega)^d).$$

Proof. (i) Define the upper bound α by (4.3). Let $\hat{\varphi} \in C_c^{\infty}(\Omega)$ be a smooth cut-off function with $0 \leq \hat{\varphi} \leq 1$ a.e. on Ω and $\hat{\varphi} \equiv 1$ except on a sufficiently small neighborhood of $\partial\Omega$. Then it holds that $\alpha \cdot \hat{\varphi} \in K(W_0^{1,p}(\Omega))$ and the assertion now follows directly from the discussion preceding Remark 4.1.

(ii) Taking account of (3.2), statement (ii) can be proven as Theorem 4.2(iii).

$\qquad \square$

4.2. **Lower semicontinuous obstacles and L^p-spaces.** The preceding counterexample provides a regularity limit in terms of the upper bound α for which the density property (3.2) in the space $X(\Omega) = L^p(\Omega)^d$ can be expected to hold. In this regard, however, uniform continuity is far from being a necessary condition. In order to enlarge the space of obstacles compatible with (3.2), we first consider a generalized lower semicontinuity condition.

Definition 4.4. The set of functions $\mathbb{LC}(\Omega)$ comprises all measurable functions $\alpha : \Omega \to \mathbb{R} \cup \{+\infty\}$ for which there exists a sequence of functions $\alpha_n : \Omega \to \mathbb{R}$ with

$$(4.7) \qquad \begin{cases} \alpha_n \in C(\overline{\Omega}), \quad \inf_{x\in\Omega} \alpha_n(x) > 0, \quad \alpha_n \leq \alpha, \quad \forall\, n \in \mathbb{N}, \\ \lim_{n\to\infty} \alpha_n(x) \to \alpha(x) \text{ for a.e. } x \in \Omega. \end{cases}$$

This property is more general than lower semicontinuity in the following sense: Consider a lower semicontinuous function $\alpha : \Omega \to \mathbb{R} \cup \{+\infty\}$ that fulfills (3.1). Without loss of generality, we may assume that $\inf_{x\in\Omega} \alpha(x) > 0$. Denote by $\tilde{\alpha}$ the extension by zero of α, i.e., $\tilde{\alpha}(x) = \alpha(x), x \in \Omega, \tilde{\alpha}(x) = 0$ on $\mathbb{R}^N \setminus \Omega$, and note that $\tilde{\alpha}$ is lower semicontinuous (l.s.c.) on $\mathbb{R}^N$. The Lipschitz regularization of $\tilde{a}$,

$$\alpha_n(x) = \inf_{y\in\mathbb{R}^N}\{\tilde{a}(y) + n\|x - y\|\},$$

yields the desired sequence (α_n) that complies with the requirements of Definition 4.4 (see, e.g., [3, Theorem 9.2.1]), such that $\alpha \in \mathbb{LC}(\Omega)$.

Theorem 4.5. *Let $1 \leq p < +\infty$. If $\alpha \in \mathbb{LC}(\Omega)$, then it holds that*

$$\overline{K(C_c^\infty(\Omega)^d)}^{L^p(\Omega)^d} = K(L^p(\Omega)^d).$$

Proof. Let $w \in K(L^p(\Omega)^d)$ for $\alpha \in \mathbb{LC}(\Omega)$. For a sequence (α_n) given by Definition 4.4 consider the functions

$$w_n(x) := \min\{|w(x)|, \alpha_n(x)\}\frac{w(x)}{|w(x)|},$$

where it is understood that $w_n(x) := 0$ if $w(x) = 0$. It follows from Lebesgue's theorem on dominated convergence that $w_n \to w$ in $L^p(\Omega)^d$. Further observe that $w_n \in K_n(L^p(\Omega)^d)$ where

$$K_n(X(\Omega)) := \{w \in X(\Omega) : |w(x)| \leq \alpha_n(x) \text{ a.e. on } \Omega\}.$$

Let $\varepsilon > 0$. According to (3.2), for each $n \in \mathbb{N}$, w_n can be approximated by a smooth function $\tilde{w}_n \in K_n(C_c^\infty(\Omega)^d) \subset K(C_c^\infty(\Omega)^d)$ such that

$$\|w_n - \tilde{w}_n\|_{L^p(\Omega)^d} < \varepsilon/2.$$

For sufficiently large n, we conclude that

$$(4.8) \qquad \|w - \tilde{w}_n\|_{L^p(\Omega)^d} \leq \|w - w_n\|_{L^p(\Omega)^d} + \|w_n - \tilde{w}_n\|_{L^p(\Omega)^d} < \varepsilon/2 + \varepsilon/2 = \varepsilon,$$

which concludes the proof. $\qquad\square$

We proceed by considering the important special case of a piecewise continuous upper bound; suppose there exists a partition of Ω into open subsets $\Omega_l \subset \Omega$ with Lipschitz boundary such that $\overline{\Omega} = \cup_{l=1}^L \overline{\Omega}_l, \Omega_i \cap \Omega_j = \emptyset$ for $i \neq j$ and

$$(4.9) \qquad \alpha|_{\Omega_l} \in C(\overline{\Omega}_l), \quad \inf_{x\in\Omega_l} \alpha|_{\Omega_l}(x) > 0, \quad l = 1, \ldots, L.$$

Theorem 4.5 ensures that for obstacles of this class the density result in the norm topology of the L^p−spaces holds true.

11

4.3. **Lower semicontinuous obstacles and Sobolev spaces.** Conditions on the obstacle α so that the density results for Sobolev spaces hold can be relaxed from assuming that $\alpha \in C(\overline{\Omega})$ to lower regularity requirements with the aid of Mosco convergence of closed and convex sets. The following definition goes back to [22].

Definition 4.6 (Mosco convergence)**.** Let X be a reflexive Banach space and (K_n) a sequence of closed convex subsets with $K_n \subset X$ for all $n \in \mathbb{N}$. Then $K_n \xrightarrow{M} K$ as $n \to +\infty$, i.e., K_n is said to Mosco converge to the set $K \subset X$ if and only if

(M1) $\qquad K \supset \{v \in X : (\exists (v_k) \subset X : v_k \in K_{n_k} \forall k \in \mathbb{N}, v_k \rightharpoonup v)\},$

(M2) $\qquad K \subset \{v \in X : (\exists (v_n) \subset X, \ \exists N \in \mathbb{N} : v_n \in K_n \forall n \geq N, v_n \to v)\}.$

Here, (K_{n_k}) denotes an arbitrary subsequence of (K_n) and the subset notation $(v_k) \subset X$ has to be understood in the sense that $\{v_k\} \subset X$.

The following class of obstacles encompasses functions $W^{1,q}(\Omega)$ that fulfill a generalized lower semicontinuity condition, which is slightly stronger than Definition 4.4.

Definition 4.7. We denote by $\mathbb{W}^q(\Omega)$ for $q \geq 1$ the set of functions $\alpha \in W^{1,q}(\Omega)$ for which there exists a sequence of functions (α_n) with α_n satisfying (3.1), $\alpha_n \leq \alpha$ a.e. in Ω and $\alpha_n \in C(\overline{\Omega}) \cap W^{1,q}(\Omega)$ for all $n \in \mathbb{N}$ such that $\alpha_n \rightharpoonup \alpha$ in $W^{1,q}(\Omega)$.

Note that the class $\mathbb{W}^q(\Omega)$ is strictly contained in $W^{1,q}(\Omega)$. Additionally, if the sequence (α_n) is non-decreasing, then the obstacle α is lower semicontinuous for being the pointwise limit of a non-decreasing sequence of continuous functions: note that $W^{1,q}(\Omega)$ embeds compactly in $L^1(\Omega)$ and hence there exists a pointwise converging subsequence $\alpha_{n_j}(x) \to \alpha(x)$ for $j \to \infty$, where we consider α as an extended-real valued function. However, the functions in $\mathbb{W}^q$ are not necessarily continuous: it suffices to consider the example from (4.1) for $\Omega = B_r(0)$, $N > 1$ and

(4.10) $$\alpha(x) = \ln(\ln(c|x|^{-1})), \quad c \geq \mathrm{e}r \text{ fixed.}$$

It follows that $\alpha \in W^{1,q}(\Omega)$ for all $q \leq N$ (see [1, 4.43]), $\alpha \notin C(\overline{\Omega})$, and the sequence (α_n) defined as $\alpha_n(x) = \min(\alpha(x), n)$ for $n \in \mathbb{N}$ satisfies the requirements of the definition of $\mathbb{W}^q(\Omega)$.

We now can establish the density result involving the class $\mathbb{W}^q(\Omega)$ for $q \geq 1$ with the aid of the results of Boccardo and Murat [5, 6].

Theorem 4.8. *Let* $1 < p < \infty$ *and suppose that* $\alpha \in \mathbb{W}^q(\Omega)$ *with* $p < q < +\infty$*. Then, the following density result holds true*

(4.11) $$\overline{K(\mathcal{D}(\Omega)^d; |\,.\,|_\infty)}^{W_0^{1,p}(\Omega)^d} = K(W_0^{1,p}(\Omega)^d; |\,.\,|_\infty),$$

where $K(X(\Omega); |\,.\,|_\infty) = \{w \in X(\Omega) : |w(x)|_\infty \leq \alpha(x)$ *a.e.* $x \in \Omega\}$*.*

Proof. Without loss of generality, consider the one-dimensional case $d = 1$.
Let $w \in K(W_0^{1,p}(\Omega); |\,.\,|_\infty)$. Since $\alpha_n \rightharpoonup \alpha$ in $W^{1,q}(\Omega)$ with $q > p > 1$, one obtains the Mosco convergence result

$$K_n^{\pm}(W_0^{1,p}(\Omega)) \xrightarrow{M} K^{\pm}(W_0^{1,p}(\Omega))$$

for the unilateral constraint sets

$$K_n^-(X(\Omega)) := \{w \in X(\Omega) : w(x) \geq -\alpha_n \text{ a.e. in } \Omega\},$$
$$K_n^+(X(\Omega)) := \{w \in X(\Omega) : w(x) \leq \alpha_n \text{ a.e. in } \Omega\},$$
$$K_-(X(\Omega)) := \{w \in X(\Omega) : w(x) \geq -\alpha \text{ a.e. in } \Omega\},$$
$$K_+(X(\Omega)) := \{w \in X(\Omega) : w(x) \leq \alpha \text{ a.e. in } \Omega\},$$

from [5, p.87]. Consequently, there exist two recovery sequences,

$$(4.12) \qquad\qquad w_n^\pm \in K_n^\pm(W_0^{1,p}(\Omega)),$$

with $w_n^\pm \to w$ in $W_0^{1,p}(\Omega)$. Using the continuity of

$$\max(\,.\,,0), \min(\,.\,,0) : \ W_0^{1,p}(\Omega) \to W_0^{1,p}(\Omega),$$

it follows that the sequence

$$w_n = \max(w_n^+, 0) + \min(w_n^-, 0),$$

converges to w in $W_0^{1,p}(\Omega)$. Moreover, it holds that $|w_n| \leq \alpha_n$ for all $n \in \mathbb{N}$. For each $n \in \mathbb{N}$, the assumptions on α_n allow to use (3.2) to infer the existence of a smooth function $\tilde{w}_n \in C_c^\infty(\Omega)$ with $|\tilde{w}_n| \leq \alpha_n \leq \alpha$ a.e. in Ω that approximates w_n arbitrarily well. Using $w_n \to w$ in $W_0^{1,p}(\Omega)^d$, the assertion follows by an $\varepsilon/2$-argument analogously to (4.8). $\qquad\square$

For piecewise continuous obstacles $\alpha : \Omega \to \mathbb{R}$ according to (4.9) only the singularities on the interfaces play a role. We define for $\eta > 0$ the enlarged interior boundaries of $\mathcal{I} = \cup_{k=1}^M \partial\Omega_k \backslash \partial\Omega$ as

$$\mathcal{I}_\eta := \{x \in \Omega : \text{dist}(x, \mathcal{I}) \leq \eta\},$$

and we consider the space of functions $C(\mathcal{I}; \Omega)$ which are uniformly continuous across $\mathcal{I}$,

$$C(\mathcal{I}; \Omega) := \{f : \Omega \to \mathbb{R} : f|_{\mathcal{I}_\eta} \in C(\overline{\mathcal{I}_\eta}) \text{ for some } \eta > 0\}.$$

The corresponding approximation result reads as follows.

Theorem 4.9. *Let $1 \leq p < \infty$. Let α be piecewise continuous in the sense of (4.9) and assume that (3.1) is fulfilled. Then the following density result holds true:*

$$\overline{K(\mathcal{D}(\overline{\Omega})^d)}^{W^{1,p}(\Omega)^d} = \overline{K(W^{1,p}(\Omega)^d \cap C(\mathcal{I};\Omega)^d)}^{W^{1,p}(\Omega)^d}.$$

Proof. Let $w \in K(W^{1,p}(\Omega)^d)$ so that $|w| \leq \alpha$ a.e. on Ω and assume that w is uniformly continuous on $\mathcal{I}_\eta$ for some fixed $\eta > 0$. Consider $Ew \in W^{1,p}(\mathbb{R}^N)^d$, the extension of w to the entire $\mathbb{R}^N$ via the extension-by-reflection operator E defined previously in (3.4). Let $E\alpha : \mathbb{R}^N \to \mathbb{R}$ be the analogous extension of α. As shown in the proof of Theorem 3.1, this extension is bound preserving:

$$|Ew| \leq E\alpha \text{ a.e. in } \mathbb{R}^N.$$

Denote by $S_n(w, \Omega) := \rho_n * Ew$ and $\alpha_n = \rho_n * E\alpha$ the mollifications of Ew and $E\alpha$ from (3.7), respectively. Since α is continuous on $\overline{E_\eta}$ where $E_\eta := (\mathcal{I}_\eta)^c \cap \Omega$, it follows that $\alpha_n \to \alpha$ uniformly on E_η. Further define

$$\beta_n := \left(1 + \frac{\sup_{x \in E_\eta} |\alpha(x) - \alpha_n(x)|}{\text{ess}\inf_{x \in \Omega} \alpha(x)}\right)^{-1},$$

13

where we use that $\operatorname{ess\,inf}_{x\in\Omega}\alpha(x) > 0$. It follows that $\beta_n \uparrow 1$ as $n \to \infty$ and $\beta_n\alpha_n(x) \le \alpha(x)$ for all $x \in E_\eta$. Since $|Ew(x)| \le E\alpha(x)$ a.e. on $\mathbb{R}^N$, one obtains $|S_n(w,\Omega)(x)| \le \alpha_n(x)$, which implies

$$(4.13) \qquad \beta_n|S_n(w,\Omega)(x)| \le \alpha(x), \quad \forall\, x \in E_\eta.$$

To enforce the feasibility on the enlarged interface set $\mathcal{I}_\eta$, we decompose $\mathcal{I}_\eta$ as $\mathcal{I}_\eta = A^+ \cup A^-$ where $A^+ := \{x \in \mathcal{I}_\eta : |w(x)| \ge s\}$ for fixed $s > 0$ with $s < \operatorname{ess\,inf}_{x\in\Omega}\alpha(x)$, and $A^- := \mathcal{I}_\eta \setminus A^+$. Define

$$\gamma_n := \left(1 + \frac{\sup_{x\in A^+}|w(x) - S_n(w,\Omega)(x)|}{s}\right)^{-1}.$$

Since $S_n(w,\Omega) \to w$ uniformly on $\mathcal{I}_\eta$, one obtains that $\gamma_n \uparrow 1$ as $n \to \infty$. We further have that

$$(4.14) \qquad \gamma_n|S_n(w,\Omega)(x)| \le |w(x)| \le \alpha(x), \quad \forall\, x \in A^+.$$

By definition, $|w(x)| < s < \operatorname{ess\,inf}_{x\in\Omega}\alpha(x)$ for all $x \in A^-$. Using once again the uniform convergence of $S_n(w,\Omega)$ to w on $\mathcal{I}_\eta$, one observes that, for sufficiently large n,

$$(4.15) \qquad |S_n(w,\Omega)| \le \operatorname*{ess\,inf}_{x\in\Omega}\alpha(x), \quad \forall\, x \in A^-.$$

Finally, the sequence

$$w_n(x) := \gamma_n\beta_n S_n(w,\Omega)(x)$$

satisfies $w_n \in \mathcal{D}(\overline{\Omega})$ for all $n \in \mathbb{N}$ and

$$w_n \to w \quad \text{in } W^{1,p}(\Omega)^d, \quad |w_n(x)| \le \alpha(x), \text{ a.e. in } \Omega,$$

for sufficiently large n; where we have used (4.13), (4.14) and (4.15). This completes the proof. $\qquad\square$

4.4. Supersolutions of elliptic PDEs.
By now, density properties for pointwise constraints in Sobolev spaces of the type

$$\overline{K(C_c^\infty(\Omega)^d)}^{W_0^{1,p}(\Omega)^d} = K(W_0^{1,p}(\Omega)^d), \text{ or } \overline{K(\mathcal{D}(\overline{\Omega})^d)}^{W^{1,p}(\Omega)^d} = K(W^{1,p}(\Omega)^d),$$

have been obtained on the basis of mollification and a subsequent procedure to enforce feasibility. An alternative approach is the approximation of a function via the solution of an appropriate sequence of elliptic PDEs. Using standard regularity theory, one may prove higher regularity of the approximating sequence and one is left to prove feasibility. In this section we focus on obstacles which are solutions of an elliptic PDE. Therefore consider a general second order differential operator A in divergence form;

$$(4.16) \qquad A = \sum_{i,j=1}^{N} -\frac{\partial}{\partial x_i}a_{ij}(x)\frac{\partial}{\partial x_j} + \sum_{i=1}^{n} b_i(x)\frac{\partial}{\partial x_i} + c(x)$$

where $a_{ij}, b_i, c \in L^\infty(\Omega)$ for $1 \le i,j \le N$. Here, the matrix $[a_{ij}(x)]$ is symmetric a.e. and uniformly elliptic, i.e., there exists a $\kappa_a > 0$ such that

$$\sum_{i,j=1}^{N} a_{ij}(x)\xi_i\xi_j \ge \kappa_a|\xi|^2, \qquad \forall\, \xi \in \mathbb{R}^N,$$

for a.e. $x \in \Omega$. It is further assumed that a_{ij}, b_i, c are such that $A : H_0^1(\Omega) \to H^{-1}(\Omega)$ is strongly monotone, i.e., there exists $\kappa > 0$ such that

$$\langle Au, u \rangle \geq \kappa \|u\|_{H_0^1(\Omega)}^2, \quad \forall u \in H_0^1(\Omega),$$

where $\langle \,.\,,\,.\, \rangle$ denotes the duality pairing in $H^{-1}(\Omega)$. For example, this is the case if $b_i \equiv 0$ for $1 \leq i \leq N$ and $c(x) \geq 0$ a.e. in Ω. We call a function $\alpha \in H^1(\Omega)$ weak supersolution with respect to the elliptic operator A, if $A\alpha \geq 0$ in the $H^{-1}(\Omega)$-sense, that is,

(4.17) $$\langle A\alpha, v \rangle \geq 0, \quad \forall\, v \in H_0^1(\Omega), v \geq 0 \text{ a.e. in } \Omega.$$

The subsequent theorem covers density properties for obstacles that are weak supersolutions of an elliptic PDE of type (4.16).

Theorem 4.10. *Let Ω be a bounded domain. Suppose that $\alpha \in H^1(\Omega)$ is a weak super-solution for some A as in (4.16) in the sense of (4.17) with $\alpha \geq 0$ on $\partial\Omega$. For $X(\Omega) \in \{L^2(\Omega)^d, H_0^1(\Omega)^d\}$ it holds that*

$$\overline{K(Y(\Omega), |\,.\,|_\infty)}^{X(\Omega)} = K(X(\Omega), |\,.\,|_\infty),$$

in the following cases.

(i) $a_{ij} \in C^{0,1}(\overline{\Omega})$ or $a_{ij} \in C^1(\Omega)$: $Y(\Omega) = (H_{loc}^2(\Omega) \cap H_0^1(\Omega))^d$,
(ii) $\partial\Omega \in C^{1,1}$ or Ω convex, $a_{ij} \in C^{0,1}(\overline{\Omega})$: $Y(\Omega) = (H^2(\Omega) \cap H_0^1(\Omega))^d$,
(iii) $a_{ij}, b_i, c \in C^{m+1}(\Omega), m \in \mathbb{N}_0$: $Y(\Omega) = (H_{loc}^{m+2}(\Omega) \cap H_0^1(\Omega))^d$,
(iv) $\partial\Omega \in C^{m+2}, a_{ij}, b_i, c \in C^{m+1}(\overline{\Omega}), m \in \mathbb{N}_0$: $Y(\Omega) = (H^{m+2}(\Omega) \cap H_0^1(\Omega))^d$.

Proof. Without loss of generality, assume $d = 1$. First observe that the maximum principle implies $\alpha(x) \geq 0$ a.e. in Ω. Let $w \in K(X(\Omega))$ be arbitrary. Consider the sequence (w_n), where w_n is defined as the unique solution to the problem,

(4.18) $$\text{find } y \in H_0^1(\Omega) : \quad \frac{1}{n}Ay + y = w \quad \text{in } H^{-1}(\Omega).$$

We denote by T_n the solution mapping to (4.18), i.e., $w_n = T_n(w)$.

Step 1: T_n-invariance of $K(H_0^1(\Omega))$: We now prove that for any $n \in \mathbb{N}$, we have that $-\alpha \leq w_n \leq \alpha$ a.e., i.e.,

(4.19) $$T_n : K(L^2(\Omega)) \to K(H_0^1(\Omega)),$$

given that $A\alpha \geq 0$ in the $H^{-1}(\Omega)$. Proceeding as in [27], we consider $(w_n - \alpha)^+$ as a test function on (4.18) and add to both sides $-\langle \frac{1}{n}A\alpha + \alpha, (w_n - \alpha)^+ \rangle$. Then,

$$\frac{\kappa}{n}\|(w_n - \alpha)^+\|_{H_0^1(\Omega)}^2 + \|(w_n - \alpha)^+\|_{L^2(\Omega)}^2 \leq \langle (\frac{1}{n}A + I)(w_n - \alpha), (w_n - \alpha)^+ \rangle$$

$$\leq \langle w - \alpha - \frac{1}{n}A\alpha, (w_n - \alpha)^+ \rangle$$

$$\leq -\frac{1}{n}\langle A\alpha, (w_n - \alpha)^+ \rangle \leq 0,$$

where we have used that $w - \alpha \leq 0$ a.e. in Ω. Therefore, $w_n \leq \alpha$ a.e. in Ω. Analogously, we obtain that $w_n \geq -\alpha$ a.e., by considering $(-\alpha - w_n)^+$ as a test function and by adding to both sides $-\langle \frac{1}{n}A\alpha + \alpha, (-\alpha - w_n)^+ \rangle$. This proves (4.19), i.e., $w_n \in K(H_0^1(\Omega))$.

Step 2: Some convergence results for singular perturbations.

15

The desired convergence modes of the approximating sequences rely on standard arguments for singular perturbations, cf. [19, Theorem 9.1, Theorem 9.4] for the case of singularly perturbed variational inequalities. First, for $y \in L^2(\Omega)$ it holds

(4.20) $$\lim_{n \to \infty} y_n = y \text{ in } L^2(\Omega) \implies \hat{y}_n := T_n(y_n) \to y \text{ in } L^2(\Omega).$$

Secondly, for $y \in H_0^1(\Omega)$, we prove that

(4.21) $$\lim_{n \to \infty} y_n = y \text{ in } H_0^1(\Omega) \implies \lim_{n \to \infty} \hat{y}_n = y \text{ in } H_0^1(\Omega).$$

In fact, since $y_n \in H_0^1(\Omega)$ and A is strongly monotone, we observe that

$$\frac{\kappa}{n} \|\hat{y}_n - y_n\|_{H_0^1(\Omega)}^2 + \|\hat{y}_n - y_n\|_{L^2(\Omega)}^2 \leq \left\langle \left(\frac{1}{n} A + I \right) (\hat{y}_n - y_n), \hat{y}_n - y_n \right\rangle$$

$$= \frac{1}{n} \langle Ay_n, y_n - \hat{y}_n \rangle$$

$$\leq \frac{1}{n} \|Ay_n\|_{H^{-1}(\Omega)} \|y_n - \hat{y}_n\|_{H_0^1(\Omega)},$$

where we have used that $\hat{y}_n$ solves (4.18) with y_n as right hand side. Hence $(\hat{y}_n)$ is bounded in $H_0^1(\Omega)$. Employing (4.20) one obtains that $\hat{y}_n \rightharpoonup y$ in $H_0^1(\Omega)$ along a subsequence, and by uniqueness, it holds $\hat{y}_n \rightharpoonup y$ for the entire sequence $(\hat{y}_n)$. Finally, from the inequalities above, we have

$$\kappa \limsup_{n \to \infty} |\hat{y}_n - y_n|_{H_0^1(\Omega)}^2 \leq \limsup_{n \to \infty} \langle Ay_n, y_n - \hat{y}_n \rangle = 0,$$

so that $\hat{y}_n = T_n(y_n) \to y$ in $H_0^1(\Omega)$ and thus (4.21) is proven.

Thirdly, in addition to $w_n = T_n(w)$, we define $w_n^q = T_n^q(w)$ where $T_n^q(w) := T_n(T_n^{q-1}(w))$ for $q \in \mathbb{N}, q \geq 2, T_n^1(w) := T_n(w) = w_n$ and $w_n^0 := w$. It can be deduced from (4.20) and (4.21) by induction that

(4.22) $$\lim_{n \to \infty} w_n^q = w \text{ in } L^2(\Omega), \qquad \forall q \in \mathbb{N} \cup \{0\},$$

for $w \in L^2(\Omega)$, and

(4.23) $$\lim_{n \to \infty} w_n^q = w \text{ in } H_0^1(\Omega), \qquad \forall q \in \mathbb{N} \cup \{0\},$$

for $w \in H_0^1(\Omega)$, respectively.

Step 3: Regularity and convergence of the approximating sequences

The extra regularity of the $H_0^1(\Omega)$-solution $T_n(w)$ to (4.18) is different with respect to the statement cases: If $a_{ij} \in C^{0,1}(\overline{\Omega})$ or $a_{ij} \in C^1(\Omega)$ for $1 \leq i, j \leq N$, the solution $T_n(w)$ belongs to $H_0^1(\Omega) \cap H_{loc}^2(\Omega)$ (see [23] for the first case and [10] for the second one). The solution $T_n(w)$ belongs to $H_0^1(\Omega) \cap H^2(\Omega)$ if $\partial\Omega$ is $C^{1,1}$-smooth [23] or when Ω is convex [12].

In case $w \in K(L^2(\Omega))$, (4.20) with $y_n \equiv w$ ensures that $w_n \to w$ in $L^2(\Omega)$. In conjuction with the regularity and the feasibility of $w_n = T_n(w)$ described above, we have then established (i) and (ii) for $X(\Omega) = L^2(\Omega)$. Secondly, note that if $w \in K(H_0^1(\Omega))$ then $w_n \to w$ in $H_0^1(\Omega)$ by (4.21) with $y_n \equiv w$, and as seen above, $w_n \in K(H_0^1(\Omega))$. This, together with the regularity of $w_n = T_n(w)$ established above, proves in turn (i) and (ii) for $X(\Omega) = H_0^1(\Omega)$.

It is left to argue for (iii) and (iv) as follows. If $a_{ij}, b_i, c \in C^{m+1}(\Omega)$ for $1 \le i, j \le N$, then for each $n \in \mathbb{N}$, the operator T_n has the following increasing regularity properties (see [10]),

$$w \in H^k(\Omega) \implies T_n(w) \in H_{\text{loc}}^{k+2}(\Omega) \cap H_0^1(\Omega), \quad 0 \le k \le m;$$

and if $a_{ij}, b_i, c \in C^{m+1}(\overline{\Omega})$ for $1 \le i, j \le N$ and $\partial\Omega$ is of class C^{m+2}, for each $n \in \mathbb{N}$,

$$w \in H^k(\Omega) \implies T_n(w) \in H^{k+2}(\Omega) \cap H_0^1(\Omega), \quad 0 \le k \le m.$$

Finally, this proves (iii) given that $w_n^q \in H_{\text{loc}}^{m+2}(\Omega) \cap H_0^1(\Omega)$ for $2q \ge m+2$, $w_n^q \in K(H_0^1(\Omega))$, and $w_n^q \to w$ as $n \to \infty$ in $L^2(\Omega)$ or $H_0^1(\Omega)$ depending on the regularity of w, cf. (4.22) and (4.23). The analogous reasoning applies to (iv).

$\square$

Let us briefly comment on the relation to the density results from Theorem 4.5 and Theorem 4.8. First, note that we do not require the obstacle to be bounded away from zero as we did in the preceding paragraphs. On the other hand, the maximal regularity of the feasible approximation hinges on the coefficients of the elliptic operator associated to the obstacle and the smoothness of the boundary.

Concerning the semicontinuity requirements of the upper bound, a classical result from Trudinger [28, Cor. 5.3] for the case without lower order terms ($b_i \equiv 0, c \equiv 0$) states that any weak supersolution in the sense of (4.17) is upper semicontinuous. Therefore, the class of obstacles considered in Theorem 4.10 differs from the one of Theorem 4.8. By contrast, the consideration of upper bounds that are weak subsolutions of an elliptic PDE is not useful as these functions may easily fail to be nonnegative on Ω. For example, this is the case if a weak subsolution satisfies a Dirichlet boundary condition.

5. Application to Finite Elements

In this section we want to show how the density results (3.2) and (3.3) can be used to derive the Mosco convergence of certain discretized versions K_h of $K(X(\Omega))$ associated with standard finite element spaces suitable for an approximation of $X(\Omega)$. The very general concept of Mosco-convergence is typically useful for investigating the stability of variational inequality problems which involve convex constraint sets, e.g, those of the type $K(X(\Omega))$, with regard to a suitable class of perturbations. In this context, the discretization of $K(X(\Omega))$ can be seen as a special type of perturbation. Applications are manifold and comprise, for instance, the discretization of variational problems in mechanics, such as in elasto-plasticity with hardening [17], or in image restoration, with regard to the predual problem of TV-regularization [14].

5.1. **Mosco convergence of sets and approximation of variational inequalities.** For the sake of convenience, we repeat at this point the definition of Mosco convergence from Section 4.

Definition 5.1 (Mosco convergence)**.** Let X be a reflexive Banach space and (K_n) a sequence of closed convex subsets with $K_n \subset X$. Then K_n is said to Mosco converge to the set $K \subset X$ if and only if

(M1) $\quad K \supset \{v \in X : (\exists (v_k) \subset X : v_k \in K_{n_k} \ \forall k \in \mathbb{N}, v_k \rightharpoonup v)\},$

(M2) $\quad K \subset \{v \in X : (\exists (v_n) \subset X, \ \exists N \in \mathbb{N} : v_n \in K_n \ \forall n \ge N, v_n \to v)\}$

Note that if (K_n) converges to K in the sense of Mosco, then K is necessarily closed and convex, too.

Remark 5.2. In some textbooks on finite-dimensional approximations of variational inequalities, cf., e.g., [11, 13], condition (M2) is replaced by the following criterion:

(M2')   There exists a dense subset $\tilde{K} \subset K$ and an operator $r_n : \tilde{K} \to X$
          such that for all $v \in \tilde{K}$ it holds $r_n v \to v$ in X and there exists $n_0 \in \mathbb{N}$
          such that $r_n v \in K_n$ for all $n \geq n_0$.

It is easy to show that (M2') implies (M2). In fact, let $v \in K$ and denote by $\pi_{K_n} v$ its (not necessarily uniquely determined) projection onto K_n. By density, for $\varepsilon > 0$, there exists $v^\varepsilon \in \tilde{K}$ such that $||v^\varepsilon - v|| \leq \varepsilon$. Thus it holds

$$||v - \pi_{K_n} v|| = \inf_{v^n \in K_n} ||v - v^n|| \leq ||v - r_n v^\varepsilon|| \leq \varepsilon + ||v^\varepsilon - r_n v^\varepsilon||$$

for sufficiently large n such that $\lim_{n \to \infty} ||v - \pi_{K_n} v|| \leq \varepsilon$, where ε was arbitrary.

The condition (M2') turns out to be convenient especially in the context of finite-dimensional approximations, where r_n is given by suitable interpolation operators which typically are only well-defined on a dense subset $Y(\Omega)$ of $X(\Omega)$ giving rise to sets $\tilde{K}$ of the type $K(Y(\Omega))$. In this respect, this is precisely the point where the density results of section 3 are needed. In view of practical relevance, we pick up on the issue of perturbations of variational inequality problems. To motivate the notion of Mosco convergence, we mention the following well known result from [19, p.99], which is a special case of the general results in [22].

Theorem 5.3. *Let X be a real Hilbert space. For each $n \in \mathbb{N}$, let $K_n \subset X$ be a nonempty, closed and convex subset. Assume $A_n : K_n \to X^*$ to be uniformly Lipschitz and strongly monotone operators that fulfill*

$$\lim_{n \to \infty} A_n v_n = A v \quad \text{in } V^*,$$

for all $(v_n) \subset X$ with $v_n \to v$ as $n \to \infty$, and $v_n \in K_n$ for all $n \in \mathbb{N}$. Further let $(l_n) \subset X^*$ with $l_n \to l$ in X^*, and assume that (K_n) converges to K in X in the sense of Mosco, cf. (M1),(M2). Then the sequence of unique solutions u_n of the problems,

$$\text{find } u_n \in K_n : \quad \langle A_n u_n, v - u_n \rangle \geq \langle l_n, v_n - u_n \rangle, \ \ \forall v_n \in K_n$$

converges strongly to the solution u of the limit problem

(5.1)   *find $u \in K :$ $\langle Au, v - u \rangle \geq \langle l, v - u \rangle, \ \ \forall v \in K$.*

In the following, the perturbation is assumed to be originating from a finite-dimensional approximation $K_n = K_{h_n}$ of the set $K(X(\Omega))$ in the framework of classical finite element methods such that the parameter n is associated with a sequence of mesh sizes (h_n) tending to zero. In this case, Mosco convergence requires that any element of the set $K(X(\Omega))$ can be approximated by discrete feasible elements. Under this condition, Theorem 5.3 ensures that the solutions of the discrete problems converge to the solution of the original infinite-dimensional problem irrespective of the regularity of the data or the obstacle defining $K(X(\Omega))$.

In this sense, Mosco-convergence is a powerful tool whenever the discrete spaces are fixed *a priori*, i.e., regardless of the data of the specific problem. The resulting sequence of finite-dimensional problems can be understood as an approximation of *any problem in a given problem class*. This applies, for example, to classical finite element methods.

In contrast, *adaptive* finite element methods intend to design the sets K_{h_n} in order to approximate *the solution of a specific problem*. In fact, the sets K_{h_n} are successively determined during the course of the adaptive algorithm building upon information on the preceding solution u_{n-1} and the specific data. With the help of suitable a posteriori error estimators, which consecutively exploit information from discrete solutions, adaptive methods aim at a reduction of the discretization error whilst enlarging the dimension of the discrete space as economically as possible. However, rigorous convergence proofs with regard to adaptive discretizations of variational inequalities are restricted to special cases and usually rely on rather strong assumptions. For instance, in the case of the obstacle problem with a piecewise affine obstacle, we mention the article [26]. Moreover, density results may still be useful in the convergence analysis of adaptive schemes which require interpolation operators, cf. [25].

5.2. Finite element discretized convex sets. In this section we assume that $\Omega \subset \mathbb{R}^N$ is polyhedral. Together with Ω, a sequence of geometrically conformal affine simplicial meshes $(\mathcal{T}_h)_{h>0}$ of Ω with mesh size

$$h := \max_{T \in \mathcal{T}_h} \operatorname{diam} T$$

is assumed to be given. For details, we refer to [9]. In analogy to the case $N = 2$, we refer to each $\mathcal{T}_h$ as a triangulation. The (N-dimensional) Lebesgue measure of an element $T \in \mathcal{T}_h$ is denoted by $\lambda(T)$. We also admit the standard assumption that the sequence $(\mathcal{T}_h)$ is shape-regular, i.e.,

(5.2) $$\exists\, c > 0 : \ \frac{\operatorname{diam}(T)}{\rho_T} \le c \quad \forall\, h \ \forall\, T \in \mathcal{T}_h,$$

where $\operatorname{diam}(T) = \max_{x,y \in T} |x-y|$ denotes the diameter of T and ρ_T designates the diameter of the largest ball that is contained in T. We further write x_T for the (barycentric) midpoint of an element T, and $\mathcal{M}_h = \{x_T : T \in \mathcal{T}_h\}$, $\mathcal{N}_h$ and $\mathcal{E}_h$ for the set of element midpoints, triangulation nodes and edges with respect to $\mathcal{T}_h$, respectively. By abuse of notation, we write $|\mathcal{M}_h|$ and $|\mathcal{N}_h|$ for the cardinality of the respective set. Let $\chi_T : \Omega \to \mathbb{R}$ designate the characteristic function of T with respect to Ω, that is,

$$\chi_T(x) = 0, \ \ \forall\, x \notin T, \ \ \chi_T(x) = 1, \ \ \forall\, x \in T.$$

We further make use of the standard $H^1(\Omega)$-conformal finite element space of globally continuous, piecewise affine functions denoted by

$$P_{1,h}(\Omega) := \{u \in C(\overline{\Omega}) : u|_T \in \mathbb{P}_1 \ \ \forall\, T \in \mathcal{T}_h\}.$$

Here, $\mathbb{P}_1$ denotes the space of polynomials of degree less than or equal to one. Together with the finite-dimensional subspace $P_{1,h}(\Omega)$ and its standard nodal basis $\{\varphi_x : x \in \mathcal{N}_h\}$ we consider the global interpolation operator

(5.3) $$I_h : C(\overline{\Omega}) \to P_{1,h}(\Omega), \quad I_h u := \sum_{x \in \mathcal{N}_h} u(x)\varphi_x.$$

Note that I_h is only defined on a dense subspace of $H^1(\Omega)$. For the discretization of variational problems in $H(\operatorname{div}; \Omega)$, it is customary to use the conforming space of Raviart-Thomas finite elements of lowest order,

(5.4) $$RT_h(\Omega) := \{w \in L^2(\Omega)^N : w|_T \in \mathbb{RT} \ \ \forall\, T \in \mathcal{T}_h, [w \cdot \nu]|_{E \cap \Omega} = 0 \ \ \forall\, E \in \mathcal{E}_h\},$$

where $\mathbb{RT} := \{w \in \mathbb{P}_1^d : \ \exists\, a \in \mathbb{R}^d, b \in \mathbb{R} : w(x) = a + bx\}$ and ν denotes the unit outer normal to T. To incorporate homogeneous Neumann boundary conditions, one uses the

$H_0(\mathrm{div}; \Omega)$-conforming subspace

$$RT_{0,h}(\Omega) := RT_h(\Omega) \cap H_0(\mathrm{div}; \Omega).$$

The construction of suitable edge-based basis functions $\{\varphi_E : E \in \mathcal{E}_h\}$ can be found in the literature, cf., for instance, [4], such that the boundary condition in the definition of $RT_{0,h}(\Omega)$ can be easily accounted for. The global Raviart-Thomas interpolation operator is given by

$$(5.5) \qquad I_h^{RT} : W^{1,1}(\Omega)^N \to RT_h(\Omega), \quad I_h^{RT} w := \sum_{E \in \mathcal{E}_h} \left(\int_E w \cdot \nu \, d\mathcal{H}^{N-1} \right) \varphi_E.$$

We emphasize that the subsequent results may be extended to finite elements of higher order, which are typically useful when the solution to the variational problem (5.1) displays a higher regularity. In this regard, higher regularity assumptions on the data and the obstacle are required and the concept of Mosco convergence is not binding to prove the convergence of the finite element method, and a priori error estimates with a rate can be derived, cf., e.g., [7]. However, we do not want to deviate from minimal regularity assumptions on the data. Further, even for simple variational problems such as the classical elasto-plastic torsion problem, there is a regularity limitation for the solution regardless of the smoothness of the data, cf. [11].

Note also that the subsequently covered problems comprise situations where the discrete feasible sets K_h are not necessarily nested and non-conforming in the sense that they are in general not contained in the feasible set $K(X)$. In the following, c denotes a positive constant, which may take different values on different occasions.

Lemma 5.4 (Mosco convergence, first condition). *Let $\Omega \subset \mathbb{R}^N$ be a polyhedral domain and $\alpha \in C(\overline{\Omega})$ with $\alpha(x) \geq 0$ in Ω. Further let (w_h) be a sequence that fulfills for all h, $w_h \in P_{1,h}(\Omega)^d$ and $|w_h(x_T)| \leq \alpha(x_T)$ for all $T \in \mathcal{T}_h$. If $w_h \rightharpoonup w$ for $h \to 0$ in $L^2(\Omega)^d$ then it holds that $|w| \leq \alpha$ a.e. in Ω.*

Proof. It suffices to show that $i_K(w) = 0$, where

$$K := \{w \in L^2(\Omega)^d : |w| \leq \alpha \text{ a.e.}\}.$$

Moreover, it holds that $i_K = j^*$, where j^* denotes the Fenchel conjugate

$$j^*(v^*) := \sup_{v \in L^2(\Omega)^d} \{(v^*, v) - j(v)\}$$

of the mapping $j : L^2(\Omega)^d \to \mathbb{R}$, $j(v) := \int_\Omega \alpha |v|_* \, dx$. Here,

$$|v^*|_* = \sup_{v \in \mathbb{R}^d \setminus \{0\}} v^* \cdot v / |v|$$

denotes the dual norm of $|\,.\,|$. From the definition of j^*, we obtain that $i_K(w) = 0$ is equivalent to

$$(5.6) \qquad (w, v) \leq \int_\Omega \alpha |v|_* \quad \forall \, v \in L^2(\Omega)^d.$$

By a density argument, it suffices to prove this result for all $v \in C_c(\Omega)^d$. Denote by

$$(5.7) \qquad \alpha_h := \sum_{T \in \mathcal{T}_h} \alpha(x_T) \chi_T, \quad v_h := \sum_{T \in \mathcal{T}_h} v(x_T) \chi_T$$

20

the piecewise constant interpolants of α and v, respectively. The uniform continuity of α and v implies $\alpha_h \to \alpha$ and $v_h \to v$ in $L^\infty(\Omega)$. By the weak convergence of (w_h), the strong convergence of (α_h) and (v_h) as well as the midpoint quadrature rule, we obtain

$$
\begin{aligned}
\int_\Omega w \cdot v \, dx \leftarrow \int_\Omega w_h \cdot v_h \, dx &= \sum_{T \in \mathcal{T}_h} \int_T w_h \cdot v_h \, dx \\
&= \sum_{T \in \mathcal{T}_h} \lambda(T) \, w_h(x_T) \cdot v_h|_T \, dx \\
&\leq \sum_{T \in \mathcal{T}_h} \lambda(T) \, \alpha(x_T) \, |v_h|_T|_* \, dx \\
&= \int_\Omega \alpha_h |v_h|_* \, dx \to \int_\Omega \alpha |v|_* \, dx,
\end{aligned}
$$

(5.8)

which proves (5.6). $\qquad\square$

Lemma 5.5. *Let $\Omega \subset \mathbb{R}^N$ be a polyhedral domain and $\alpha \in C(\overline{\Omega})$ with $\alpha(x) \geq 0$ in Ω. Let (w_h) be a sequence that fulfills for all h, $w_h \in P_{1,h}(\Omega)^d$ and $|w_h(x)| \leq \alpha(x)$ for all $x \in \mathcal{N}_h$. If $w_h \rightharpoonup w$ for $h \to 0$ in $L^2(\Omega)^d$ then it holds that $|w| \leq \alpha$ a.e. in Ω.*

Proof. The assertion follows by a slight modification of the proof of Lemma 5.4. Instead of the piecewise constant interpolant we define α_h as the piecewise affine interpolant of α, i.e., $\alpha_h = I_h \alpha$, which fulfills $\alpha(x) = (I_h \alpha)(x)$ for all $x \in \mathcal{N}_h$ and $\alpha_h \to \alpha$ strongly in $L^\infty(\Omega)^d$. By (5.8) we obtain

$$
\begin{aligned}
\int_\Omega w \cdot v \, dx \leftarrow \int_\Omega w_h \cdot v_h \, dx &= \sum_{T \in \mathcal{T}_h} \frac{\lambda(T)}{N+1} \sum_{x \in \mathcal{N}_h \cap T} w_h(x) \cdot v_h|_T \, dx \\
&\leq \sum_{T \in \mathcal{T}_h} \frac{\lambda(T)}{N+1} \sum_{x \in \mathcal{N}_h \cap T} |w_h(x)| \, |v_h|_T|_* \\
&\leq \sum_{T \in \mathcal{T}_h} \frac{\lambda(T)}{N+1} \sum_{x \in \mathcal{N}_h \cap T} \alpha(x) \, |v_h|_T|_* \\
&= \int_\Omega \alpha_h |v_h|_* \, dx \to \int_\Omega \alpha |v|_* \, dx.
\end{aligned}
$$

$\qquad\square$

Theorem 5.6. *Let $\Omega \subset \mathbb{R}^N$ be a polyhedral domain and $\alpha \in C(\overline{\Omega})$ such that (3.1) holds true. Then the sets*

(5.9) $$ K_h = \{ w \in P_{1,h}(\Omega)^d : |w(x_T)| \leq \alpha(x_T) \text{ for all } T \in \mathcal{T}_h \} $$

Mosco-converge for $h \to 0$ to the set $K(H^1(\Omega)^d)$ in $H^1(\Omega)^d$.

Proof. Since weak convergence in $H^1(\Omega)$ implies weak convergence in $L^2(\Omega)$, the preceding Lemma 5.4 shows that (M1) is fulfilled. We now show (M2'). To prove the assertion, we may use a strategy that is similar to the one in [11, Chapter II] and requires (3.3). Note that Theorem 3.1 implies that the set

(5.10) $$ \tilde{K} := \{ \varphi \in C^\infty(\overline{\Omega})^d : |\varphi(x)| < \alpha(x) \text{ for all } x \in \overline{\Omega} \} $$

is also dense in $K(H^1(\Omega)^d)$ with respect to the $H^1(\Omega)^d$-norm. For the global interpolation operator I_h defined in (5.3) we have the classical estimate,

$$(5.11) \qquad ||u - I_h u||_{L^\infty(\Omega)} \le ch^2 ||u||_{W^{2,\infty}(\Omega)} \quad \forall\, u \in W^{2,\infty}(\Omega).$$

Here, c denotes a constant independent of h on account of the shape-regularity of the triangulation (5.2); cf. [9, p.61]. We further define $r_h : \tilde{K} \to P_{1,h}(\Omega)^d$ by

$$r_h w := [I_h w_1, \dots, I_h w_d],$$

and it follows that $r_h w \to w$ as $h \to 0$ in $H^1(\Omega)^d$ for all $w \in \tilde{K}$; see [9, Corollary 1.109]. Applying estimate (5.11) to the components of $w \in \tilde{K}$ and using the equivalence of norms on $\mathbb{R}^d$, one obtains that

$$(5.12) \qquad ||\,|w - r_h w|\,||_{L^\infty(\Omega)} \le ch^2 ||w||_{W^{2,\infty}(\Omega)^d},$$

for a suitable modification of c. This implies

$$(5.13) \qquad |r_h w(x)| \le |w(x)| + ch^2 ||w||_{W^{2,\infty}(\Omega)^d} \quad \forall\, x \in \Omega.$$

Since any $w \in \tilde{K}$ is uniformly bounded away from α, there exists $h_0 = h_0(w)$ such that $r_h w \in K_h \,\forall\, h \le h_0$, which implies (M2'). $\qquad \square$

Corollary 5.7. *Under the conditions of Theorem 5.6, the sequence (K_h) defined in (5.9) Mosco-converges for $h \to 0$ to the set $K(L^2(\Omega)^d)$ in $L^2(\Omega)^d$.*

Proof. Again, Lemma 5.4 implies that (M1) with $X = L^2(\Omega)^d$ holds true. For $\tilde{K}$ defined in (5.10) it holds that $\tilde{K}$ is also dense in $K(L^2(\Omega)^d)$ with respect to the $L^2(\Omega)^d$-norm, cf. (3.2). Thus, (M2') follows analogously to the proof of Theorem 5.6. $\qquad \square$

Corollary 5.8. *Under the conditions of Theorem 5.6 the node-based discrete sets*

$$(5.14) \qquad K_h := \{w \in P_{1,h}(\Omega)^d : |w(x)| \le \alpha(x) \,\forall\, x \in \mathcal{N}_h\},$$

Mosco converge for $h \to 0$ to $K(H^1(\Omega)^d)$ in $H^1(\Omega)^d$.

Proof. The proof is analogous to the proof of Theorem 5.6, noting that (5.13) also implies $r_h w \in K_h \,\forall\, h \le h_0$ with K_h according to the node-based definition (5.14). $\qquad \square$

Remark 5.9. With the help of the density property (3.3) for uniformly continuous upper bounds, the above results on the Mosco convergence of discretized convex sets carry over to spaces involving homogeneous Dirichlet boundary conditions. In this context, the set $P_{1,h}(\Omega)$ in the definitions of the discretized sets K_h in (5.9) and (5.14) has to be replaced by the space

$$P_{1,h}^{\partial\Omega} := \{u \in C(\overline{\Omega}) : u|_T \in \mathbb{P}_1 \,\forall\, T \in \mathcal{T}_h, \, u(x) = 0 \,\forall\, x \in \mathcal{N}_h \cap \partial\Omega\}.$$

The resulting discrete sets K_h incorporate the zero boundary condition and the corresponding results on Mosco convergence for $h \to 0$ remain valid replacing $H^1(\Omega)^d$ by $H_0^1(\Omega)^d$.

With the help of the density result (3.2), one obtains the following result for the discrete approximation of pointwise constraint sets in $H(\text{div}; \Omega)$ by the Raviart-Thomas finite element space $RT_h(\Omega)$; cf. (5.4).

Theorem 5.10. *Let $\Omega \subset \mathbb{R}^N$ be a polyhedral domain. Let $\alpha \in C(\overline{\Omega})$ such that (3.1) is satisfied. Then the sets*

$$K_h := \{w \in RT_{0,h}(\Omega) : |w(x_T)| \le \alpha(x_T) \ \ \forall T \in \mathcal{T}_h\}$$

Mosco-converge to $K(H_0(\operatorname{div};\Omega))$ in $H(\operatorname{div};\Omega)$ and to $K(L^2(\Omega)^N)$ in $L^2(\Omega)^N$.

Proof. Let $w_h \in K_h$ for all h. First observe that if (w_h) weakly converges to w in $H(\operatorname{div};\Omega)$, then it also weakly converges to w in $L^2(\Omega)^N$. Analogously to the proof of Lemma 5.4 one concludes that $|w| \le \alpha$ a.e. in Ω. The continuity of the normal trace mapping

$$H(\operatorname{div};\Omega) \ni w \mapsto \langle w\nu, v\rangle_{H^{-1/2}(\partial\Omega), H^{1/2}(\partial\Omega)} \in \mathbb{R}$$

for fixed $v \in H^1(\Omega)$ implies $w\nu = 0$ in $H^{-1/2}(\partial\Omega)$. We conclude that $w \in K(H_0(\operatorname{div};\Omega))$ whence it follows that (M1) is satisfied. Secondly, note that

$$\overline{K(C_c^\infty(\Omega)^N)}^{H(\operatorname{div};\Omega)} = K(H_0(\operatorname{div};\Omega));$$

cf. (3.2). For the global Raviart-Thomas interpolation operator defined in (5.5), the following interpolation error estimate holds true [9, Corollary 1.115]:

(5.15) $$\|u - I_h^{RT}u\|_{L^\infty(\Omega)^N} + \|\operatorname{div} u - \operatorname{div} I_h^{RT}u\|_{L^\infty(\Omega)} \le ch\|u\|_{W^{1,\infty}(\Omega)^N}$$

for all $u \in W^{2,\infty}(\Omega)^N$. Setting $r_h w := I_h^{RT}w$ for any $w \in \tilde{K}$, where

$$\tilde{K} := \{w \in C_c^\infty(\Omega)^N : |w(x)| < \alpha(x), \ \ \forall x \in \Omega\},$$

and taking account of the fact that $I_h^{RT}w \to w$ in $H(\operatorname{div})$ for all $w \in \tilde{K}$, we may proceed analogously to the proof of Theorem 5.6 to verify (M2'). $\qquad\square$

The previous approach can also be applied to derive approximation results for constraint sets involving pointwise bounds on partial derivatives. To begin with, we consider the gradient-constraint sets

$$K_\nabla(X(\Omega)) = \{w \in X(\Omega) : |\nabla w| \le \alpha \text{ a.e. in } \Omega\},$$

for $X(\Omega) \subset H^1(\Omega)^d$ and an arbitrary norm $|\,.\,|$ on $\mathbb{R}^{N \times d}$.

Theorem 5.11. *Let $\Omega \subset \mathbb{R}^N$ be a polyhedral domain. Let $\alpha \in C(\overline{\Omega})$ such that (3.1) is satisfied. Define*

(5.16) $$K_h := \{w \in P_{1,h}^{\partial\Omega}(\Omega)^d : |\nabla w|_T| \le \alpha(x_T) \ \ \forall T \in \mathcal{T}_h\}.$$

Then the sets K_h Mosco-converge to $K_\nabla(H_0^1(\Omega)^d)$ in $H_0^1(\Omega)^d$.

Proof. To prove (M1) it suffices to notice that if $w_h \rightharpoonup w$ in $H^1(\Omega)^d$ then $\nabla w_h \rightharpoonup \nabla w$ in $L^2(\Omega)^{N \times d}$. Similar to the proof of Lemma 5.4, one obtains for $v \in C_c(\Omega)^{N \times d}$ that

$$\int_\Omega \nabla w : v \, dx \leftarrow \int_\Omega \nabla w_h : v \, dx \le \int_\Omega |\nabla w_h||v|_* \, dx \le \int_\Omega \alpha_h |v|_* \, dx \to \int_\Omega \alpha |v|_* \, dx,$$

using α_h from (5.7). Therefore, (5.6) holds with ∇w in place of w, and (M1) is verified.

To prove (M2'), we consider again the global interpolation operator I_h from (5.3). The standard estimate

$$\|\nabla u - \nabla I_h u\|_{L^\infty(\Omega)^N} \le ch\|u\|_{W^{2,\infty}(\Omega)}, \quad \forall u \in W^{2,\infty}(\Omega),$$

23

holds true, see e.g. [9]. Note also that $K_\nabla(C_c^\infty(\Omega)^d)$ is dense in $K_\nabla(H_0^1(\Omega)^d)$ for the $H^1(\Omega)^d$-norm [15, Theorem 4]. Thus, the set

$$\tilde{K} := \{w \in C_c^\infty(\Omega)^d : |\nabla w(x)| < \alpha(x) \quad \forall x \in \Omega\}$$

is also dense in $K_\nabla(H^1(\Omega)^d)$. Therefore one may argue as in the proof of Theorem 5.6 to deduce (M2'). $\qquad\square$

Next we consider pointwise constraints on the divergence. For $X(\Omega) \subset H(\mathrm{div}; \Omega)$ let

(5.17) $\qquad\qquad K_{\mathrm{div}}(X(\Omega)) := \{w \in X(\Omega) : |\mathrm{div}\, w| \leq \alpha \text{ a.e. in } \Omega\}.$

Using Raviart-Thomas finite elements, a discrete realization of the inequality constraint in (5.17) can be achieved by imposing the inequality on the midpoints of the triangulation. The following statement ensures that the resulting approach is stable as the mesh width goes to zero.

Theorem 5.12. *Let* $\Omega \subset \mathbb{R}^N$ *be a polyhedral domain. Let* $\alpha \in C(\overline{\Omega})$ *fulfill* (3.1). *Then the sets*

$$K_h := \{w \in RT_{0,h}(\Omega) : |\mathrm{div}\, w|_T| \leq \alpha(x_T), \ \ \forall T \in \mathcal{T}_h\}$$

Mosco-converge in $H(\mathrm{div}; \Omega)$ *to the set* $K_{\mathrm{div}}(H_0(\mathrm{div}; \Omega))$ *as defined in* (5.17).

Proof. Taking account of the fact that $w_h \rightharpoonup w$ in $H(\mathrm{div}; \Omega)$, $w_h \in K_h$, implies $\mathrm{div}\, w_h \rightharpoonup \mathrm{div}\, w$ in $L^2(\Omega)$, (M1) follows analogously to the corresponding part of the proof of Theorem 5.11. Since $K_{\mathrm{div}}(C_c^\infty(\Omega)^N)$ is dense in $K_{\mathrm{div}}(H_0(\mathrm{div}; \Omega))$ [15, Theorem 4], the set

$$\tilde{K} := \{w \in C_c^\infty(\Omega)^d : |\mathrm{div}\, w(x)| < \alpha(x), \ \ \forall x \in \Omega\}$$

is also dense in $K_{\mathrm{div}}(H_0(\mathrm{div}; \Omega))$. Setting $r_h = I_h^{RT}$, the estimate (5.15) implies $r_h w \to w$ in $H(\mathrm{div}; \Omega)$ and

$$\|\mathrm{div}\, w - \mathrm{div}\, r_h w\|_{L^\infty(\Omega)} \leq ch\|w\|_{W^{2,\infty}(\Omega)^N},$$

for all w in $\tilde{K}$. In particular, one may argue as in the proof of Theorem 5.6 to verify (M2'). $\qquad\square$

For a general L^p-function as upper bound, a point-based discretization is obviously not possible. As a remedy, the construction of the discrete sets K_h typically involves some kind of averaging process. For this purpose, we define the integral mean

$$\fint_T \alpha\, dx := \int_T \alpha\, dx / \lambda(T)$$

over some given subset $T \subset \Omega$ (with positive measure).

Now we have to take into account that the density results of the type (3.2) and (3.3), which represent the main ingredient to prove the consistency of the finite element approximation, may fail to hold true (see, e.g., Theorem 4.2). On the other hand, the results from Section 4 indicate that the density property is still guaranteed for a large class of discontinuous obstacles. To maintain the greatest level of generality, we assume that the nonnegative measurable function $\alpha : \Omega \to \mathbb{R} \cup \{+\infty\}$ allows for the density property

(5.18) $\qquad\qquad \overline{K(C(\overline{\Omega}))}^{L^2(\Omega)^d} = K(L^2(\Omega)^d).$

Here, we concentrate on the consistency in the L^2-topology but an extension to the other cases is possible by appropriately modifying assumption (5.18). We stress the fact that assumption (5.18) is fulfilled in relevant situations; cf., e.g., Theorem 4.5.

Lemma 5.13. *Let $\Omega \subset \mathbb{R}^N$ be a polyhedral domain and $\alpha \in L^2(\Omega)$ with $\alpha(x) \geq 0$ a.e. in Ω. Let (w_h) be a sequence that fulfills for all h, $w_h \in P_{1,h}(\Omega)^d$ and $|w_h(x_T)| \leq \fint_T \alpha \, dx$ for all $T \in \mathcal{T}_h$. If $w_h \rightharpoonup w$ for $h \to 0$ in $L^2(\Omega)^d$ then it holds that $|w| \leq \alpha$ a.e. in Ω.*

Proof. The assertion follows analogously to the proof of Lemma 5.4 by a slight modification of the definition of α_h. Instead of the piecewise constant interpolant we consider the piecewise constant quasi-interpolant $\alpha_h := \sum_{T \in \mathcal{T}_h} \chi_T \fint_T \alpha \, dx$. Observe that α_h converges strongly to α in $L^2(\Omega)^d$, which is sufficient for the above argument. $\qquad\square$

Theorem 5.14. *Let $\Omega \subset \mathbb{R}^N$ be a polyhedral domain. Let $\alpha \in L^2(\Omega)$ with (3.1) such that (5.18) holds true. Then the sets*

$$K_h := \{w \in P_{1,h}(\Omega)^d : |w(x_T)| \leq \fint_T \alpha \, dx, \quad \forall T \in \mathcal{T}_h\}$$

Mosco-converge for $h \to 0$ to the set $K(L^2(\Omega)^d)$ in $L^2(\Omega)^d$.

Proof. We only need to prove (M2') since Lemma 5.13 implies (M1). First note that (3.1) and (5.18) imply that the set

$$\tilde{K} := \{w \in C_c^\infty(\Omega)^d : \exists \delta = \delta(w) > 0 \text{ such that } |w(x)| \leq \alpha(x) - \delta \text{ a.e. in } \Omega\},$$

is also dense in $K(L^2(\Omega)^d)$. Furthermore, we set

$$r_h w := [I_h w_1, \ldots, I_h w_d],$$

for $w \in \tilde{K}$ and I_h as in (5.3). Integrating estimate (5.13) yields

$$|\fint_T r_h w \, dx| \leq \fint_T |w| \, dx + ch^2 \|w\|_{W^{2,\infty}(\Omega)^d}, \quad \forall T \in \mathcal{T}_h.$$

Let $w \in \tilde{K}$ be fixed. Since $r_h w$ is affine on each $T \in \mathcal{T}_h$, an application of the midpoint rule shows

$$|r_h w(x_T)| \leq \fint_T |w| \, dx + ch^2 \|w\|_{W^{2,\infty}(\Omega)^d}, \quad \forall T \in \mathcal{T}_h,$$

which implies

(5.19) $$|r_h w(x_T)| \leq \fint_T \alpha \, dx - \delta(w) + ch^2 \|w\|_{W^{2,\infty}(\Omega)^d}, \quad \forall T \in \mathcal{T}_h.$$

This implies $r_h w \in K_h$ for all $w \in \tilde{K}$ and $h \leq h_0(w)$. By (5.11) it holds that $r_h w \to w$ in $L^2(\Omega)^d$ for $h \to 0$, which proves (M2'). $\qquad\square$

6. FURTHER APPLICATIONS

6.1. **Regularization of elasto-plastic contact problems.** In the context of the one time-step problem of quasi-static elasto-plasticity with an associative flow law, the deformation of a material represented by a bounded Lipschitz domain Ω subject to given applied forces is modeled by the evolution of the displacement, the material stress and strain as well as certain internal variables, cf. [13]. An elasto-plastic contact problem arises if the movement of the material is additionally restricted by the presence of a rigid obstacle. From a mathematical point of view, the problem can be equivalently reformulated in terms of the normal stress z^* at the (sufficiently smooth)

contact boundary Γ_c and a variable q that is related to the deviatoric part of the material stress; for details we refer to [17, p.154]:

$$(6.1) \quad \begin{cases} \min & G([z^*, q]) - \langle z^*, \psi \rangle \quad \text{over } [z^*, q] \in H^{1/2}(\Gamma_c)^* \times L^2(\Omega)^d \\ \text{s.t.} & z^* \in H_+^{1/2}(\Gamma_c)^*, \\ & |q|_2 \le \beta \text{ a.e. in } \Omega. \end{cases}$$

Here, $d := \frac{N(N+1)}{2} - 1$ and G is a strongly convex, continuous and coercive functional that models the elasto-plastic material behavior subject to given external loads. Furthermore, a contact constraint on the normal component of the displacement is imposed by a function ψ, which lies in the trace space $H^{1/2}(\Gamma_c)$. The upper bound $\beta \in L^2(\Omega)$ is determined by the hardening law, and it is bounded away from zero by the positive yield stress σ_y, i.e., $\beta(x) \ge \sigma_y$ a.e. in Ω. The normal stress z^* is constrained to lie in the polar cone

$$H_+^{1/2}(\Gamma_c)^* = \{z^* \in H^{1/2}(\Gamma_c)^* : \langle z^*, z \rangle \le 0 \ \forall z \in H_+^{1/2}(\Gamma_c)\},$$

to the cone of nonnegative functions

$$H_+^{1/2}(\Gamma_c) = \{z \in H^{1/2}(\Gamma_c) : z \ge 0 \text{ a.e. on } \Gamma_c\},$$

where $H^{1/2}(\Gamma_c)^*$ designates the topological dual space of $H^{1/2}(\Gamma_c)$. From an algorithmic point of view, it is favorable to replace (6.1) by a combined Moreau-Yosida/Tikhonov regularization given by

$$(6.1_\gamma) \quad \begin{cases} \min & G([z, q]) - (z, \psi)_{L^2(\Gamma_c)} + \frac{\gamma_n}{2}\|z^+\|^2_{L^2(\Gamma_c)} \\ & + \frac{\gamma_n}{2}\|[(|q|_2 - \beta)]^+\|^2_{L^2(\Omega)} + \frac{1}{2\gamma_n'}\|[z, q]\|^2_{H^1(\Gamma_c) \times H^1(\Omega)^d}, \\ \text{over} & [z, q] \in H^1(\Gamma_c) \times H^1(\Omega)^d, \end{cases}$$

where (γ_n) and (γ_n') are sequences with $\gamma_n, \gamma_n' \to +\infty$ as $n \to +\infty$. In contrast to (6.1), (6.1_γ) can be solved efficiently by the semismooth Newton method in the infinite-dimensional setting. As a consequence, the Newton iterates are superlinearly convergent, and the convergence rate is mesh-independent upon discretization. For details, see [17, Section 5]. In order to prove the stability of (6.1_γ) with regard to the limit problem (6.1) in the sense of Proposition A1, we show that the problems (6.1_γ) define a quasi-monotone perturbation of $i_\mathcal{K}$ with respect to the dense subspace $H^1(\Gamma_c) \times H^1(\Omega)^d \subset H^{1/2}(\Gamma_c)^* \times L^2(\Omega)^d$; cf. the definition below (2.2). Here, we write for $\mathcal{X} \subset H^{1/2}(\Gamma_c)^* \times L^2(\Omega)^d$,

$$\mathcal{K}(\mathcal{X}) := \{[z^*, q] \in \mathcal{X} : z^* \in H_+^{1/2}(\Gamma_c)^*, \ |q|_2 \le \beta \text{ a.e. in } \Omega\},$$

and $\mathcal{K} := \mathcal{K}(H^{1/2}(\Gamma_c)^* \times L^2(\Omega)^d)$. In fact, setting

$$R_n([z, q]) := \frac{\gamma_n}{2}\|z^+\|^2_{L^2(\Gamma_c)} + \frac{\gamma_n}{2}\|[(|q|_2 - \beta)]^+\|^2_{L^2(\Omega)} + \frac{1}{2\gamma_n'}\|[z, q]\|^2_{H^1(\Gamma_c) \times H^1(\Omega)^d},$$

where it is understood that $R_n([z^*, q]) = +\infty$, unless $[z^*, q] \in H^1(\Gamma_c) \times H^1(\Omega)^d$, it is easily seen that

$$\overline{R_n}([z, q]) := i_\mathcal{K}([z, q]) + \frac{1}{2\gamma_n'}\|[z, q]\|^2_{H^1(\Gamma_c) \times H^1(\Omega)^d},$$

fulfills (2.4). Moreover, we set

$$\underline{R_n}([z^*, q]) := \frac{\gamma_n}{2}r(z^*) + \frac{\gamma_n}{2}\|[(|q|_2 - \beta)]^+\|^2_{L^2(\Omega)},$$

where

$$r(z^*) := (\max\{ \sup_{\substack{z \in H_+^{1/2}(\Gamma_c), \\ \|z\|_{H^{1/2}(\Gamma_c)}=1}} \langle z^*, z \rangle, 0\})^2.$$

The validity of (2.3) is an immediate consequence of the following lemma.

Lemma 6.1. *The functional* $r : H^{1/2}(\Gamma_c)^* \to \mathbb{R}$ *is weakly l.s.c. and it fulfills*

   (i) $r(z^*) = 0$ *for all* $z^* \in H_+^{1/2}(\Gamma_c)^*$,
   (ii) $r(z^*) > 0$ *for all* $z^* \in H^{1/2}(\Gamma_c)^* \setminus H_+^{1/2}(\Gamma_c)^*$,
   (iii) $r(z) \leq \|z^+\|_{L^2(\Gamma_c)}^2$ *for all* $z \in L^2(\Gamma_c)$.

Proof. As a composition of a convex, continuous and monotone function with a supremum of l.s.c. and convex functions, $r : H^{1/2}(\Gamma_c)^* \to \mathbb{R}$ is weakly l.s.c. Assertions (i) and (ii) are direct consequences of the definition of $H_+^{1/2}(\Gamma_c)^*$. For $z \in L_-^2(\Gamma_c) = \{z \in L^2(\Gamma_c) : z \leq 0 \text{ a.e. in } \Omega\}$, it holds $r(z) = 0$ and (iii) is always satisfied. Now let $z \in L^2(\Gamma_c) \setminus L_-^2(\Gamma_c)$. By the density of $H_+^{1/2}(\Gamma_c)$ in $L_+^2(\Gamma_c)$ it holds that

$$r(z)^{1/2} = \sup_{\substack{z \in H_+^{1/2}(\Gamma_c) \\ \|z\|_{H^{1/2}(\Gamma_c)}=1}} \langle z^*, z \rangle > 0.$$

Moreover, one obtains

$$\|z^+\|_{L^2(\Gamma_c)} = \sup_{\substack{\tilde{z} \in L^2(\Gamma_c) \\ \tilde{z} \neq 0}} \frac{1}{\|\tilde{z}\|_{L^2(\Gamma_c)}} (z^+, \tilde{z})$$

$$\geq \sup_{\substack{\tilde{z} \in L^2(\Gamma_c) \\ \tilde{z} \neq 0, \tilde{z} \geq 0 \text{ a.e.}}} \frac{1}{\|\tilde{z}\|_{L^2(\Gamma_c)}} (z, \tilde{z}) \geq \sup_{\substack{\tilde{z} \in H_+^{1/2}(\Gamma_c) \\ \tilde{z} \neq 0}} \frac{1}{\|\tilde{z}\|_{H^{1/2}(\Gamma_c)}} (z, \tilde{z}) = r(z)^{1/2},$$

which implies (iii). $\qquad\square$

From the discussion in the introduction and Proposition A1, it is known that the consistency of the regularization scheme (6.1_γ) with respect to (6.1) hinges on the density of $K(H^1(\Omega)^d)$ in $K(L^2(\Omega)^d)$, where

$$K(X(\Omega)) := \{q \in X(\Omega) : |q|_2 \leq \beta \text{ a.e. in } \Omega\},$$

in accordance with the notation from the preceding sections. Owing to the results of sections 3 and 4, this is always fulfilled for kinematic hardening, as β is a positive constant in this case. In the same way, it is also fulfilled for large classes of discontinuous obstacles β in the case of combined isotropic-kinematic hardening. Once the density property is ensured, one may use monotonicity properties of G to derive strong convergence properties of regularized (normal) stresses, strains and displacement; cf. [17] for details.

6.2. **Fenchel duality in image restoration.** Optimization problems with total variation regularization have been successfully considered in the image restoration context. In the denoising setting, an original image u_{true} that belongs to the space of functions of bounded variation

$BV(\Omega)$, $\Omega \subset \mathbb{R}^2$, is sought to be recovered from a noise perturbed measurement $f = u_{\text{true}} + \eta$ with $\eta \in L^2(\Omega)$, $\int \eta = 0$ and $\int |\eta|^2 = \sigma^2$. This motivates the optimization problem

$$\min \ \frac{1}{2} \int_\Omega |u - f|^2 \mathrm{d}x + \alpha \int_\Omega |\mathcal{D}u|_1 \quad \text{over } u \in BV(\Omega),$$

for $\alpha \in \mathbb{R}$ in the seminal work [24] by Rudin, Osher and Fatemi. Here, $\mathcal{D}u$, the distributional gradient of $u \in BV(\Omega)$, is a Borel measure and $|\mathcal{D}u|_1$ is its total variation measure with total mass $\int_\Omega |\mathcal{D}u|_1$, which is characterized via duality as

$$\int_\Omega |\mathcal{D}u|_1 = \sup \left\{ \int_\Omega u \ \mathrm{div}\, v \ \mathrm{d}x : v \in C_c^1(\Omega; \mathbb{R}^2), \ |v(x)|_\infty \leq 1, \ \forall\, x \in \Omega \right\}.$$

The drawback of the above reconstruction scheme is that the choice of the regularization parameter α is global: A good reconstruction locally requires high values of α in some regions of Ω (e.g., flat regions of u_{true}) and low values in other regions (e.g., locations of details of u_{true}). A recent approach in [16, 18] proposes the following alternative: For a function $\alpha : \Omega \to \mathbb{R}$ such that (3.1) holds true, consider the optimization problem

$$(6.2) \qquad \min \ \ \frac{1}{2} \int_\Omega |u - f|^2 \mathrm{d}x + \int_\Omega \alpha(x)|\mathcal{D}u|_1 \quad \text{over } u \in BV(\Omega),$$

where $\int_\Omega \alpha(x)|\mathcal{D}u|_1$ stands for the integral of α on Ω with respect to the measure $|\mathcal{D}u|_1$. Hence, α *needs* to be a $|\mathcal{D}u|_1$-integrable function in order for $\int_\Omega \alpha|\mathcal{D}u|_1$ to be correctly defined. A sufficient condition for this is given by $\alpha \in C(\Omega)$, the space of continuous functions on Ω.

As usual in convex optimization, it is convenient to consider the problem in (6.2) from the point of view of Fenchel duality. In fact, (6.2) can be characterized as the Fenchel dual problem of the following constrained optimization problem

$$(6.3) \qquad \begin{cases} \min & \frac{1}{2}\| \mathrm{div}\, p + f\|_{L^2(\Omega)}^2 \quad \text{over } p \in H_0(\mathrm{div}; \Omega) \\ \text{s.t.} & p \in K(H_0(\mathrm{div}; \Omega), |\,.\,|_\infty), \end{cases}$$

if the following density result holds true:

$$\overline{K(C_c^1(\Omega)^2), |\,.\,|_\infty}^{H_0(\mathrm{div};\Omega)} = K(H_0(\mathrm{div}; \Omega), |\,.\,|_\infty),$$

where, according to the above notational convention,

$$K(H_0(\mathrm{div}; \Omega), |\,.\,|_\infty) = \{q \in H_0(\mathrm{div}; \Omega) : |q(x)|_\infty \leq \alpha(x) \text{ a.e. in } \Omega\}.$$

APPENDIX: PROPERTIES OF QUASI-MONOTONE PERTURBATIONS

Proposition A1. *Let the Banach space X be reflexive or assume that the dual space X^* is separable. For a closed, convex and nonempty set $K \subset X$, let (R_n) be a sequence of quasi-monotone perturbations of i_K with respect to the dense subspace Y according to (2.2). If the density property (1.1) holds true, then $F + i_K$ is the Γ-limit of $(F + R_n)$ in both, the weak and strong topology.*

Proof. Denote by

$$\Gamma\text{-}\limsup_{n \to +\infty} G_n(u) := \sup_{U \in \mathcal{N}(u)} \limsup_{n \to +\infty} \inf_{u \in U} G_n(u)$$

the Γ-upper limit at u of a sequence of functions $G_n : X \to \mathbb{R} \cup \{+\infty\}$ in the norm topology. Here, $\mathcal{N}(u)$ denotes the set of all open neighborhoods of u in the norm of X. By analogy, define $\Gamma_w\text{-}\limsup_{n \to +\infty} G_n$, the Γ-upper limit of G_n in the weak topology of X, as well as the lower limit counterpart $\Gamma_w\text{-}\liminf_{n \to +\infty} G_n$. We write

$$\Gamma_w\text{-}\lim_{n \to +\infty} G_n = \Gamma_w\text{-}\limsup_{n \to +\infty} G_n = \Gamma_w\text{-}\liminf_{n \to +\infty} G_n$$

for the weak Γ-limit of (G_n) provided the latter equality is satisfied. For the corresponding definitions we refer to the monograph [8]. Further denote by $\mathrm{sc}^- G$ the lower semicontinuous envelope of $G : X \to \mathbb{R} \cup \{+\infty\}$. Exploiting the relations between Γ- and pointwise convergence [8, Chapter 5], one obtains with (2.4) and the continuity of F,

$$\Gamma_w\text{-}\limsup_{n \to +\infty}(F + R_n) \leq \Gamma\text{-}\limsup_{n \to +\infty}(F + R_n)$$
$$\leq \Gamma\text{-}\limsup_{n \to +\infty}(F + \overline{R_n}) = \mathrm{sc}^-(F + i_{K \cap Y}) = F + i_{\overline{K \cap Y}},$$

where we use [8, Prop. 6.3, Prop. 6.7, Prop. 5.7, Prop. 3.7]. Similarly, (2.3) together with [8, Prop. 6.7, Prop. 5.4] implies

$$(6.4) \qquad \Gamma_w\text{-}\liminf_{n \to +\infty}(F + R_n) \geq \Gamma_w\text{-}\liminf_{n \to +\infty}(F + \underline{R_n}) = \lim_{n \to +\infty} \mathrm{sc}_w^-(F + \underline{R_n})$$

where $\mathrm{sc}_w^-(F + \underline{R_n})$ denotes the lower semicontinuous envelope of $F + \underline{R_n}$ in the weak topology of X. Further note that the coercivity and the sequential weak lower semicontinuity of $F + \underline{R_n}$ imply that the level sets $\{u \in X : F(u) + R_n(u) \leq t\}$, $t \in \mathbb{R}$, are bounded and sequentially weakly closed. If X is reflexive or if the dual space X^* is separable, then the sequential weak closure of bounded subsets of X coincides with the weak closure, see [8, Prop. 8.7, Prop. 8.14], such that $F + \underline{R_n}$ is weakly lower semicontinuous, which entails

$$\Gamma_w\text{-}\liminf_{n \to +\infty}(F + R_n) \geq \lim_{n \to +\infty}(F + \underline{R_n}) = F + i_K,$$

by (6.4). Eventually, it holds that

$$F + i_K \leq \Gamma_w\text{-}\liminf_{n \to +\infty}(F + R_n)$$
$$\leq \Gamma_w\text{-}\limsup_{n \to +\infty}(F + R_n) \leq \Gamma\text{-}\limsup_{n \to +\infty}(F + R_n) \leq F + i_{\overline{K \cap Y}},$$

such that $\Gamma\text{-}\lim_{n \to +\infty}(F + R_n) = \Gamma_w\text{-}\lim_{n \to +\infty}(F + R_n u) = F + i_K$, if (1.1) holds true. $\quad\square$

Proposition A2. *Let the assumptions of Example 2.4 be satisfied. Further suppose that $\overline{K \cap Y} \subsetneq K$. Then for all $x \in K \setminus \overline{K \cap Y}$ there exists a strictly increasing sequence (γ_n) with $\gamma_n \to \infty$ such that there exists no strong recovery sequence at x, i.e.,*

$$F(y_n) + R_n(y_n) \nrightarrow F(x)$$

for all $(y_n) \subset X$ with $y_n \to x$, where (R_n) is given by (2.7).

Proof. Let $x \in K \setminus \overline{K \cap Y}$ and $\rho > 0$ such that $\overline{B_\rho(x)} \cap \overline{K \cap Y} = \emptyset$, where $B_\rho(x) := \{y \in X : \|x - y\| < \rho\}$.

(a) We first prove the following result:

$$(6.5) \qquad \forall\, n \in \mathbb{N} \ \exists\, \gamma_n > 0 : \quad \left[y \in X \land \mathrm{dist}(y, K \cap \overline{B_\rho(x)})^2 < \tfrac{1}{\gamma_n} \implies y \notin X_n \right].$$

Assume that the opposite holds, i.e.,

$$\exists\, n_0 \in \mathbb{N}: \quad \left[\, \forall\, n \in \mathbb{N}\ \exists\, x_n \in X_{n_0}, v_n \in K \cap \overline{B_\rho(x)}: \quad \|x_n - v_n\|^2 \le \tfrac{1}{n} \right].$$

Since $v_n \in \overline{B_\rho(x)} \cap K$ for all $n \in \mathbb{N}$ and $\overline{B_\rho(x)} \cap K$ is convex, bounded and closed, there exists a subsequence (v_{n_k}) of (v_n) with $v_{n_k} \rightharpoonup v$ and $v \in \overline{B_\rho(x)} \cap K$. As $x_n - v_n \to 0$, one also obtains $x_{n_k} \rightharpoonup v$ and thus $v \in X_{n_0}$. Hence, $v \in X_{n_0} \cap K \cap \overline{B_\rho(x)} = \emptyset$, which is a contradiction.

(b) Non-existence of a strong recovery sequence:

Choose (γ_n) according to (6.5) and assume that there exists a recovery sequence (y_n) to x, which means that $y_n \to x$ and $F(y_n) + \frac{\gamma_n}{2} \operatorname{dist}(y_n, K)^2 + i_{X_n}(y_n) \to F(x)$. The continuity of F implies that $y_n \in X_n$ for sufficiently large n and that $\frac{\gamma_n}{2} \operatorname{dist}(y_n, K)^2 \to 0$. Consequently, using $y_n \to x$ and $x \in K$, there exists $n_1 \in \mathbb{N}$ such that

$$\operatorname{dist}(y_n, K)^2 = \operatorname{dist}(y_n, K \cap B_\rho(x))^2 \le \tfrac{1}{\gamma_n}$$

for all $n \ge n_1$. With the help of part (a), we conclude that $y_n \notin X_n$ for all $n \ge n_1$, which is a contradiction.

$\square$

REFERENCES

[1] R.A. Adams and J.J.F. Fournier. *Sobolev Spaces*. Pure and Applied Mathematics. Elsevier Science, 2003.

[2] H.W. Alt. *Lineare Funktionalanalysis*. Springer, 2012.

[3] H. Attouch, G. Buttazzo, and G. Michaille. *Variational Analysis in Sobolev and BV Spaces: Applications to PDEs and Optimization*. SIAM, 2nd edition, 2014.

[4] C. Bahriawati and C. Carstensen. Three MATLAB implementations of the lowest-order Raviart-Thomas MFEM with a posteriori error control. *Computational Methods in Applied Mathematics*, 5(4):333–361, 2005.

[5] L. Boccardo and F. Murat. Nouveaux résultats de convergence dans des problèmes unilatéraux. In *Non-linear Partial Differential Equations and their Applications (Collège de France Seminar)*, volume II, pages 68–85. Pitman Research Notes in Math., 1982.

[6] L. Boccardo and F. Murat. Homogenization of nonlinear unilateral problems. In *Composite Media and Homogenization Theory*, pages 81–105. Springer, 1991.

[7] F. Brezzi, W.W. Hager, and P.A. Raviart. Error Estimates for the Finite Element Solution of Variational Inequalities. Part I. Primal Theory. *Numerische Mathematik*, 28:431–444, 1977.

[8] G. Dal Maso. *An Introduction to Γ-convergence*. Birkhäuser, Boston, 1993.

[9] A. Ern and J.-L. Guermond. *Theory and Practice of Finite Elements*. Springer, 2004.

[10] L.C. Evans. *Partial Differential Equations*. American Mathematical Society, 1998.

[11] R. Glowinski. *Numerical Methods for Nonlinear Variational Problems*. Springer, 1984.

[12] P. Grisvard. *Elliptic Problems in Nonsmooth Domains*. Pitman Publishing, 1985.

[13] W. Han and B.D. Reddy. *Plasticity: Mathematical Theory and Numerical Analysis*. Springer, New York, 2nd edition, 2013.

[14] M. Hintermüller and K. Kunisch. Total bounded variation regularization as bilaterally constrained optimization problem. *SIAM J. Applied Mathematics*, 64:1311–1333, 2004.

[15] M. Hintermüller and C.N. Rautenberg. On the density of classes of closed convex sets with pointwise constraints in Sobolev spaces. *Journal of Mathematical Analysis and Applications*, 426(1):585 – 593, 2015.

[16] M. Hintermüller and C.N. Rautenberg. Optimal selection of the regularization function in a generalized total variation model. Part I: Modelling and theory. Technical report, WIAS Berlin, 2016.

[17] M. Hintermüller and S. Rösel. A duality-based path-following semismooth Newton method for elasto-plastic contact problems. *Journal of Computational and Applied Mathematics*, 292:150 – 173, 2016.

[18] M. Hintermüller, C. N. Rautenberg, T. Wu, and Langer A. Optimal selection of the regularization function in a generalized total variation model. Part II: Algorithm, its analysis and numerical tests. Technical report, WIAS Berlin, 2016.

[19] J.-F. Rodrigues. *Obstacle Problems in Mathematical Phyisics*. North-Holland, 1987.

[20] D. Kinderlehrer and G. Stampacchia. *An Introduction to Variational Inequalities and their Applications*. Academic Press, Inc., New York, 1980.

[21] A. Mielke and T. Roubíček. *Rate-independent Systems*. Springer, 2015.

[22] U. Mosco. Convergence of convex sets and of solutions of variational inequalities. *Advances in Mathematics*, 3(4):510 – 585, 1969.

[23] J. Nečas. *Les méthodes directes en théorie des équations elliptiques*. Academia, Prague, 1967.

[24] L.I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Phys. D*, 60(1-4):259–268, 1992. Experimental mathematics: computational issues in nonlinear science (Los Alamos, NM, 1991).

[25] K. G. Siebert. A convergence proof for adaptive finite elements without lower bound. *IMA Journal of Numerical Analysis*, 31(3):947–970, 2011.

[26] K. G. Siebert and A. Veeser. A unilaterally constrained quadratic minimization with adaptive finite elements. *SIAM Journal on Optimization*, 18(1):260–289, 2007.

[27] H. Toyoizumi. Continuous dependence on obstacles in variational inequalities. *Funkcialaj Ekvacioj*, 34:103–115, 1991.

[28] N. S. Trudinger. On the regularity of generalized solutions of linear, non-uniformly elliptic equations. *Archive for Rational Mechanics and Analysis*, 42(1):50–62, 1971.