Research

Author for correspondence:
M. Hintermüller
e-mail: hint@math.hu-berlin.de

Density of convex intersections and applications

M. Hintermüller[1,2], C. N. Rautenberg[1] and S. Rösel[1]

[1]Department of Mathematics, Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany
[2]Weierstrass Institute, Mohrenstrasse 39, 10117 Berlin, Germany

MH, 0000-0001-9471-2479

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 with 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 infinite-dimensional 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; e.g. [1–3] for fundamental concepts in variational analysis. In this vein, given a set of functions satisfying an arbitrary convex constraint, density properties of more regular functions satisfying the same restriction are of utmost importance. In abstract terms, given some dense subspace Y of X, the central point of interest is whether the closure property

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

with $K(Y) = \{u \in Y : u \in K\} = K \cap Y$, is fulfilled, and how this problem is intimately linked to the solution of constrained optimization and variational inequality problems.

In the literature, problems of dense intersections appear in connection with the discretization of variational inequality problems in Sobolev spaces and the convergence analysis for finite-element methods under minimal regularity (e.g. [4–6]). Moreover, the limiting behaviour of singular perturbations of elliptic variational inequalities can be traced back to the density issue (see [7] and references therein). This also pertains to the deduction of a vanishing viscosity limit for hyperbolic variational inequalities with an obstacle constraint [8]. In the context of plasticity problems, certain density properties represent an important step towards the determination of appropriate relaxed formulations (cf. [9,10]). However, to the best of our knowledge, the investigation of problem (1.1) is restricted to special cases and the literature still lacks a general and systematic treatment of the density issue.

To motivate the study of the abstract problem (1.1), §2 provides a novel unifying framework for various perturbation approaches to non-smooth constrained optimization and variational inequality problems. The general setting includes regularization, Galerkin approximation and singular perturbations, and, most remarkably, it allows to reduce the study of the corresponding limit problems for a wide range of practically relevant perturbations to the study of the density property (1.1). In particular, we prove that the dense intersection (1.1) is a necessary and sufficient stability condition for the retrieval of the original problem in the (joint) limit of vanishing regularization and/or discretization parameters.

Starting from §3 we focus on the 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.

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

with $|\cdot|$ 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 $K(X(\Omega); |\cdot|)$ whenever it is necessary to make the dependence on the norm $|\cdot|$ explicit. Despite the fact that a small number of specific density results for very regular bounds α are available [4,9,11], a systematic investigation of density properties in terms of the regularity of α seems not available in the literature.

In order to close this gap, we prove new density results for continuous obstacles (§4), and we also consider different classes of discontinuous obstacles. In fact, in §4a, the density issue is studied in the context of the regularity of the obstacle as a Sobolev function. More precisely, we prove that results of the type (1.1) cannot be expected if the obstacle is just a Sobolev function by providing a counterexample. The density results are then proved to be valid even for certain classes of lower semicontinuous obstacles; see §4b,c. Subsequently, in §4d, a different approach is considered for obstacles that originate from the solution of a partial differential equation (PDE).

In §5, we focus on the Mosco convergence of finite-element discretized convex sets, which, in general, is a delicate matter, and only a limited number of results for more regular obstacles are known (e.g. [4,5]). In this respect, the construction of a recovery sequence essentially reduces to the verification of density properties of the type (1.1). Making use of the density results provided by the preceding sections, we prove several new Mosco convergence results in the Hilbert spaces L^2, H^1 and $H(\mathrm{div})$ for different types of finite-element discretizations of K, even for discontinuous obstacles α. The results are extended to a more general constraint setting involving pointwise restrictions on partial derivatives. We conclude the paper by presenting two important applications that further highlight the paramount significance of dense intersections. First, we consider the regularization of an elasto-plastic contact problem, where the closure property turns out to be fundamental for the efficient solution by a semismooth Newton method. Secondly, we discuss an example from total variation-based image restoration with a distributed non-smooth regularization parameter. Here, the density property arises as an essential condition for the equivalent reformulation of the problem in the Hilbert space $H(\mathrm{div})$ by means of Fenchel duality.

2. Motivation

(a) Optimization with convex constraints

In many variational problems, one seeks the solution in a given convex, closed and non-empty subset K of an infinite-dimensional Banach space $(X, \|.\|)$. To illustrate the problem, let us consider the following abstract class of optimization problems:

$$\left. \begin{aligned} \inf \quad & F(u), \quad \text{over } u \in X, \\ \text{s.t.} \quad & u \in K. \end{aligned} \right\} \tag{2.1}$$

We assume that $F : X \to \mathbb{R}$ is continuous, coercive and sequentially weakly lower semicontinuous, 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 PDEs, variational inequality problems of potential type, optimal control of PDEs 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 (e.g. [1–5] and references therein). The central result of this section is that the stability of (2.1) with respect to a large class of perturbations can be characterized by the closure property (1.1), i.e.

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

where Y is some dense subspace of X (in the norm topology of X), and $K(Y)$ is given by

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

In what follows, we will identify a very general class of perturbations for which the stability analysis effectively reduces to the study of the density property (1.1).

(i) A class of quasi-monotone perturbations

To subsume as many of the above-mentioned methods as possible, we consider the sequence of perturbed problems

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

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

$$0 \leq \underline{R}_n \leq R_n \leq \overline{R}_n \quad \forall n \in \mathbb{N},$$

having the additional properties

$$\left. \begin{aligned} \underline{R}_n \leq \underline{R}_{n+1} \ \forall n \in \mathbb{N}, \quad \lim_{n \to +\infty} \underline{R}_n(u) = i_K(u) \ \forall u \in X \\ \underline{R}_n \text{ is sequentially weakly lower semicontinuous } \forall n \in \mathbb{N} \end{aligned} \right\} \tag{2.3}$$

and

and

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

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.

The first main result of this section states that the dense intersection property implies the stability of *any* quasi-monotone perturbation scheme. The proof is deferred to appendix A.

Theorem 2.1 (Sufficient condition). *Let the Banach space X be reflexive or assume that the dual space X^* is separable. For a closed, convex and non-empty 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.*

Under the assumptions of theorem 2.1, one may infer that, provided each problem (2.2) admits a global minimizer u_n, each weak cluster point of the sequence of minimizers (u_n) is a global minimizer of (2.1); see [12] for an introduction to Γ-convergence. At the end of this section, it is further clarified that theorem 2.1 is sharp in the sense that the stability result in general fails if (1.1) does not hold. 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), which bear high practical relevance. In favour of generality, we do not leave the abstract setting.

Example 2.2 (Tikhonov regularization). Let $(Y, \| \dots \|_Y)$ 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

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

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 2.3 (Conforming discretization). Let X be a separable Banach space. Suppose (2.1) is approximated by a Galerkin approach using nested and conforming 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 2.4 (Combined Moreau–Yosida/Tikhonov regularization). Let X be a Hilbert space and $(Y, \| \dots \|_Y)$ be a Banach space that is densely and continuously embedded into X. For 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

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

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

Example 2.5 (Conforming 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

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

where the sequence of spaces (X_n) is defined as in example 2.3. Setting $\underline{R}_n(u) = (\gamma_n/2) \inf_{v \in K} \|u - v\|^2$ and $\overline{R}_n(u) = i_{K \cap X_n}(u)$, (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 that 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 examples 2.2 and 2.3 can be calculated using similar arguments as in the proof of theorem 2.1. 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 examples 2.4 and 2.5, theorem 2.1 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 2.5). 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 [13, Prop. 2.4.6]. However, the proof is non-constructive and thus not immediately useful for the design of a stable numerical algorithm. On the other hand, if (1.1) is violated, the Γ-convergence to the original problem (2.1) cannot be guaranteed independently from the choice of the regularization/discretization parameter pair. In fact, the following result, which we prove in appendix A, holds true.

Proposition 2.6 (Necessary condition). *Consider example 2.5 with the corresponding definitions of Y and (R_n). 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*

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

for all $(y_n) \subset X$ with $y_n \to x$, i.e. there exists no recovery sequence at x in the norm topology.

The analogous statement is valid in the case of combined Moreau–Yosida/Tikhonov regularizations given a fixed sequence (γ_n'); cf. example 2.4. In conclusion, theorem 2.1 is sharp with respect to condition (1.1) in the sense of proposition 2.6 and the preceding discussion.

(b) 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 $K \subset X$ non-empty, closed and convex, we consider the general variational inequality problem of the first kind,

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

e.g. [7,14] for an introduction. 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.

(i) Quasi-monotone perturbation

Consider the perturbed variational inequality problem,

$$\text{find } u_n \in X : \quad \langle A_n u_n, v - u_n \rangle + R_n(v) - R_n(u_n) \geq \langle l_n, v - u_n \rangle, \quad \forall v \in X, \tag{2.9}$$

where A_n and l_n are appropriate perturbations of A and l, respectively, and (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 the sequence (R_n) Mosco converges to i_K provided $\underline{R}_n$ is weakly lower semicontinuous. Under mild assumptions on (A_n) and (l_n) one may then invoke known stability

results, cf. [7, p.99, 15], to conclude the consistency of the perturbation scheme with respect to the limit problem (2.8).

(ii) Galerkin approximation of variational inequalities

In general, finite-dimensional approximations of K are neither conforming nor nested as it was the case in examples 2.3 and 2.5, where K was 'discretized' by $K \cap X_n$, which is numerically realizable only in special cases. Instead, it is often more favourable 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 [4,5]. As a result, the finite-dimensional variational inequality problems,

$$\text{find } u_n \in X_n: \quad \langle A_n u_n, v - u_n \rangle + i_{K_n}(v) - i_{K_n}(u_n) \geq \langle l_n, v - u_n \rangle, \quad \forall v \in X_n, \tag{2.10}$$

do not fit into the framework of (2.9). Again, under mild assumptions on (A_n) and (l_n), the Mosco convergence of (K_n) to K ensures that the approximation (2.10) is stable with respect to the limit problem (2.8). However, Mosco convergence requires the existence of a recovery sequence (see definition 4.5) for any element $u \in K$. To construct this sequence in the context of finite-element methods, one may use an interpolation procedure which typically is only defined on the (supposedly) dense subset $K \cap Y$ of K, where $Y = C^k(\bar{\Omega})$ for some $k \in \mathbb{N}_0$ (cf. [4, II, Theorem 2.3] and §5). This leads again to problem (1.1).

(iii) Singular perturbations

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

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

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

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

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, [7, section 4.9, 8] 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 $\mathcal{D}(\bar{\Omega})$,

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

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

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

and

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

In the recent paper [11], it has been shown that for any $\alpha \in C(\bar{\Omega})$ with

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

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:

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

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 $|\cdot|$ on $\mathbb{R}^d$, i.e.

$$K(X(\Omega)) := \{w \in X(\Omega) : |w(x)| \le \alpha(x) \text{ a.e. in } \Omega\}. \tag{3.3}$$

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

When considering the case $X = W^{1,p}$ instead of $W_0^{1,p}$ in (3.2), the choice of the approximating sequence from [11, Theorem 1], which relies on the trivial extension of Sobolev functions, fails. As a result, a different extension operator has to be employed.

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

$$\overline{K(\mathcal{D}(\bar{\Omega})^d)}^{W^{1,p}(\Omega)^d} = K(W^{1,p}(\Omega)^d), \tag{3.4}$$

i.e. $K(\mathcal{D}(\bar{\Omega})^d)$ 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

$$E : W^{1,p}(\Omega)^d \to W^{1,p}(\mathbb{R}^N)^d \tag{3.5}$$

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, [16]. 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

$$|(Ew)(x)| \le E_{C(\bar{\Omega})}\alpha(x), \quad \text{a.e. } x \in \mathbb{R}^N, \tag{3.6}$$

where $E_{C(\bar{\Omega})} : C(\bar{\Omega}) \to C(\mathbb{R}^N)$ denotes the application of the extension by reflection procedure to bounded uniformly continuous functions, i.e. $(E_{C(\bar{\Omega})}\alpha)|_\Omega = \alpha$. 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 $E_{C(\bar{\Omega})}\alpha \in C_c(\mathbb{R}^N)$. For a sequence (ρ_n) of smooth mollifiers

$$\rho_n(x) = n^N \rho(nx), \tag{3.7}$$

where

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

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

$$S_n(w, \Omega)(x) := (\rho_n * Ew)(x) = \int_{\mathbb{R}^N} Ew(y)\rho_n(x - y) \, dy, \quad x \in \mathbb{R}^N. \tag{3.8}$$

It is well known that

$$S_n(w, \Omega)|_\Omega \to w \text{ in } W^{1,p}(\Omega)^d \quad \text{as } n \to \infty, \tag{3.9}$$

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

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

where $\alpha_n(x) := ((E_{C(\bar{\Omega})}\alpha) * \rho_n)(x)$, $x \in \mathbb{R}^N$. Since $E_{C(\bar{\Omega})}\alpha \in C_c(\mathbb{R}^N)$, α_n converges to $E_{C(\bar{\Omega})}\alpha$ uniformly in $\mathbb{R}^N$ and thus $\beta_n \to 1$ as $n \to \infty$. In addition, (3.6) together with (3.8) yields $|S_n(w, \Omega)| \leq \alpha_n(x)$ for $x \in \mathbb{R}^N$ such that

$$\beta_n^{-1}\alpha(x) = \alpha(x) + \frac{\sup_{x \in \mathbb{R}^N} |\alpha_n(x) - E_{C(\bar{\Omega})}\alpha(x)|}{\min_{x \in \bar{\Omega}} \alpha(x)} \alpha(x) \geq \alpha_n(x) \geq |S_n(w, \Omega)|, \tag{3.10}$$

for all $x \in \Omega$. As a result, $\beta_n S_n(w, \Omega) \in K(\mathcal{D}(\bar{\Omega})^d)$ and, taking account of (3.9), the proof is accomplished. $\blacksquare$

Remark 3.2 (boundary conditions).

(i) 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 [11].

(ii) If the set $K(W^{1,p}(\Omega)^d)$ is additionally restricted by an inhomogeneous Dirichlet boundary condition given by a function $g \in W^{1-1/p,p}(\partial\Omega)^d$ with $|g(x)| \leq \alpha(x)$ on $\partial\Omega$, the proof of theorem 3.1 fails. In fact, the sequence (3.8), which is based on the standard mollifier, does not preserve a given trace condition. In any case, the regularity of g (and $\partial\Omega$) determines an *a priori* regularity limitation for the functions in Y in order to be compatible with a closure property analogous to (3.4), e.g. if $Y = C(\bar{\Omega})$ and $g \notin C(\partial\Omega)$, then $K \cap Y = \emptyset$. In this case, a different mollification approach needs to be pursued; cf. also §7 for an outlook on this matter.

4. Density results for discontinuous obstacles

(a) 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 §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 (3.4) 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 centre $x \in \mathbb{R}^N$ and radius $\varepsilon > 0$ with respect to the Euclidean norm $|\cdot|_2$ in $\mathbb{R}^N$. Let $\{x_k : k \in \mathbb{N}\}$ be a countable dense subset, i.e.

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

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

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

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 non-negative with a singularity at the origin, and its zero extension belongs to $W^{1,N}(\mathbb{R}^N)$; cf. [17, 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 [18, 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 $W^{1,N}(\Omega)$; e.g. [14, 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 fulfils $w \leq 1$ 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 $\phi(t) \geq 1 - \mu/2$ for all $t \geq R$. By continuity, there also exists $\delta' > 0$ such that $\varphi(x - x_{k_0}) \geq R k_0^2$ 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

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

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:

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

Using the Baire category theorem, one may show that the set $\mathcal{S}$ is a non-meagre set with vanishing Lebesgue measure [19].

The previous construction of the counterexample is the basis for the following result.

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*

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

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 $C(\bar{\Omega})^d$. Let $w \in K(W^{1,p}(\Omega)^d)$. Setting

$$\hat{\alpha}(x) = \max \left(|w(x)|, \underset{x \in \Omega}{\mathrm{ess\,inf}\,} \alpha(x) \right),$$

it follows that $|w(x)| \leq \hat{\alpha}(x)$ a.e. in Ω. Since $\hat{\alpha} \in C(\bar{\Omega})$ 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 $w_n \in \mathcal{D}(\bar{\Omega})^d$, $w_n \to w$ in $W^{1,p}(\Omega)^d$ and $|w_n(x)| \leq \hat{\alpha}(x) \leq \alpha(x)$ a.e. in Ω. This entails that $w_n \in K(\mathcal{D}(\bar{\Omega})^d)$ for all $n \in \mathbb{N}$, which accomplishes the proof. ∎

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 neighbourhood 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). ∎

(b) Lower semicontinuous obstacles and Lebesgue 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 upper bounds that allow for a lower semicontinuous representative, i.e. there exists a lower semicontinuous function in the equivalence class of functions that are Lebesgue-almost everywhere equal to α.

Theorem 4.4. *Let $\Omega \subset \mathbb{R}^N$ be a bounded Lipschitz domain and $1 \leq p < +\infty$. If $\alpha : \Omega \to \mathbb{R} \cup \{+\infty\}$ has a lower semicontinuous representative that fulfils (3.1), then it holds that*

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

Proof. Let $w \in K(L^p(\Omega)^d)$. Consider a lower semicontinuous function $\alpha : \Omega \to \mathbb{R} \cup \{+\infty\}$ that fulfils (3.1). Without loss of generality, we may assume that $\inf_{x \in \Omega} \alpha(x) > 0$. Denote by $\tilde{\alpha}$ the extension of α given by $\tilde{\alpha}(x) := \alpha(x), x \in \Omega$, $\tilde{\alpha}(x) := \inf_{x \in \Omega} \alpha(x)$ on $\mathbb{R}^N \setminus \Omega$, and note that $\tilde{\alpha}$ is lower

10

rspa.royalsocietypublishing.org *Proc. R. Soc. A* **473**: 20160919

semicontinuous (l.s.c.) on $\mathbb{R}^N$. The Lipschitz regularization of $\tilde{\alpha}$,

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

yields a sequence (α_n) with $\alpha_n \in C(\bar{\Omega})$, $\inf_{x \in \Omega}\alpha(x) \le \alpha_n(x) \le \alpha(x)$ for all $x \in \Omega, n \in \mathbb{N}$ and $\alpha_n(x) \to \alpha(x)$ a.e. in Ω; see, e.g. [2, Theorem 9.2.1]. Now 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

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

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

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

For sufficiently large n, we conclude that

$$\|w - \tilde{w}_n\|_{L^p(\Omega)^d} \le \|w - w_n\|_{L^p(\Omega)^d} + \|w_n - \tilde{w}_n\|_{L^p(\Omega)^d} < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon, \tag{4.7}$$

which concludes the proof. ∎

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 $\bar{\Omega} = \cup_{l=1}^L \bar{\Omega}_l$, $\Omega_i \cap \Omega_j = \emptyset$ for $i \ne j$ and

$$\alpha|_{\Omega_l} \in C(\bar{\Omega}_l), \quad \inf_{x \in \Omega_l} \alpha|_{\Omega_l}(x) > 0, \quad l = 1, \dots, L. \tag{4.8}$$

Theorem 4.4 ensures that for obstacles of this class the density result in the norm topology of the L^p-spaces holds true.

(c) 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 $\alpha \in C(\bar{\Omega})$ to lower regularity requirements with the aid of Mosco convergence of closed and convex sets. The following definition goes back to [15].

Definition 4.5 (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

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

and

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

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 in $W^{1,q}(\Omega)$ that fulfil a generalized lower semicontinuity condition.

Definition 4.6. We denote by $\mathbb{W}^q(\Omega)$ for $q \ge 1$ 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), $\alpha_n \le \alpha$ 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, any obstacle $\alpha \in \mathbb{W}^q(\Omega)$ has a lower semicontinuous representative, which follows easily from definition 4.6 and by extraction of a pointwise almost everywhere converging subsequence. 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

$$\alpha(x) = \ln(\ln(c|x|^{-1})), \quad c \geq er \text{ fixed.} \tag{4.9}$$

It follows that $\alpha \in W^{1,q}(\Omega)$ for all $q \leq N$ (see [17, Example 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 $\mathbb{W}^q(\Omega)$.

Theorem 4.7. *Let $\Omega \subset \mathbb{R}^N$ be a bounded Lipschitz domain. Let $1 \leq p < \infty$ and $\alpha \in \mathbb{W}^p(\Omega)$. Then the following density results hold true*:

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

and
$$\overline{K(\mathcal{D}(\bar{\Omega})^d; |\cdot|_\infty)}^{W^{1,p}(\Omega)^d} = K(W^{1,p}(\Omega)^d; |\cdot|_\infty),$$

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

Proof. Without loss of generality, consider the one-dimensional case $d = 1$. Let $w \in K(W_0^{1,p}(\Omega); |\cdot|_\infty)$ and $(\alpha_n) \subset W^{1,p}(\Omega)$ according to definition 4.6. By Mazur's lemma, we may as well assume that (α_n) converges strongly to α in $W^{1,p}(\Omega)$ since convex combinations preserve order and continuity. Hence, 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

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

and
$$K_+(X(\Omega)) := \{w \in X(\Omega) : w(x) \leq \alpha \text{ a.e. in } \Omega\}.$$

Consequently, there exist two recovery sequences,

$$w_n^\pm \in K_n^\pm(W_0^{1,p}(\Omega)), \tag{4.10}$$

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

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

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 as in (4.7). The proof for the case $X(\Omega) = W^{1,p}(\Omega)^d$ follows analogously by invoking theorem 3.1. ∎

(d) Supersolutions of elliptic partial differential equations

By now, density properties for pointwise constraints in Sobolev spaces of the type

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

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;

$$A = \sum_{i,j=1}^{N} -\frac{\partial}{\partial x_i} a_{ij}(x) \frac{\partial}{\partial x_j} + \sum_{i=1}^{n} b_i(x) \frac{\partial}{\partial x_i} + c(x), \tag{4.11}$$

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

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

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

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

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

$$\langle A\alpha, v \rangle \ge 0, \quad \forall v \in H_0^1(\Omega), v \ge 0 \text{ a.e. in } \Omega. \tag{4.12}$$

The subsequent theorem covers density properties for obstacles that are weak supersolutions of an elliptic PDE of type (4.11).

Theorem 4.8. *Let Ω be a bounded domain. Let $\alpha \in H^1(\Omega)$ be a weak supersolution for some A as in (4.11) in the sense of (4.12), with $\alpha \ge 0$ on $\partial\Omega$. For $X(\Omega) \in \{L^2(\Omega)^d, H_0^1(\Omega)^d\}$, it holds that*

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

in the following cases:

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

Proof. Without loss of generality, assume $d = 1$. First observe that the maximum principle implies $\alpha(x) \ge 0$ 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,

$$\text{find } y \in H_0^1(\Omega): \frac{1}{n}Ay + y = w \quad \text{in } H^{-1}(\Omega). \tag{4.13}$$

We denote by T_n the solution mapping to (4.13), 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 $-\alpha \le w_n \le \alpha$ a.e., i.e.

$$T_n : K(L^2(\Omega)) \to K(H_0^1(\Omega)), \tag{4.14}$$

given that $A\alpha \ge 0$ in the $H^{-1}(\Omega)$. Proceeding as in [20], we consider $(w_n - \alpha)^+$ as a test function on (4.13) and add to both sides $-\langle (1/n)A\alpha + \alpha, (w_n - \alpha)^+ \rangle$. Then,

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

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

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

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 $-\langle (1/n)A\alpha + \alpha, (-\alpha - w_n)^+ \rangle$. This proves (4.14), i.e. $w_n \in K(H_0^1(\Omega))$.

Step 2: Some convergence results for singular perturbations.

The desired convergence modes of the approximating sequences rely on standard arguments for singular perturbations, cf. [7, Theorems 9.1 and 9.4] for the case of singularly perturbed variational inequalities. First, for $y \in L^2(\Omega)$ it holds

$$\lim_{n \to \infty} y_n = y \text{ in } L^2(\Omega) \implies \hat{y}_n := T_n(y_n) \to y \text{ in } L^2(\Omega). \tag{4.15}$$

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

$$\lim_{n \to \infty} y_n = y \text{ in } H_0^1(\Omega) \implies \lim_{n \to \infty} \hat{y}_n = y \text{ in } H_0^1(\Omega). \tag{4.16}$$

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

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

where we have used that $\hat{y}_n$ solves (4.13) with y_n as right-hand side. Hence $(\hat{y}_n)$ is bounded in $H_0^1(\Omega)$. Employing (4.15) 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 (4.16) 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 (4.15) and (4.16) by induction that

$$\lim_{n \to \infty} w_n^q = w \quad \text{in } L^2(\Omega), \quad \forall q \in \mathbb{N} \cup \{0\}, \tag{4.17}$$

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

$$\lim_{n \to \infty} w_n^q = w \quad \text{in } H_0^1(\Omega), \quad \forall q \in \mathbb{N} \cup \{0\}, \tag{4.18}$$

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 (4.13) is different with respect to the statement cases: if $a_{ij} \in C^{0,1}(\bar{\Omega})$ or $a_{ij} \in C^1(\Omega)$ for $1 \leq i,j \leq N$, the solution $T_n(w)$ belongs to $H_0^1(\Omega) \cap H_{\text{loc}}^2(\Omega)$ (see [21] for the first case and [18] 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 [21] or when Ω is convex [22].

In case $w \in K(L^2(\Omega))$, (4.15) with $y_n \equiv w$ ensures that $w_n \to w$ in $L^2(\Omega)$. In conjunction 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 (4.16) 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 [18],

$$w \in H^k_{\mathrm{loc}}(\Omega) \Longrightarrow T_n(w) \in H^{k+2}_{\mathrm{loc}}(\Omega) \cap H^1_0(\Omega), \quad 0 \le k \le m;$$

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

$$w \in H^k(\Omega) \Longrightarrow T_n(w) \in H^{k+2}(\Omega) \cap H^1_0(\Omega), \quad 0 \le k \le m.$$

Finally, this proves (iii) given that $w_n^q \in H^{m+2}_{\mathrm{loc}}(\Omega) \cap H^1_0(\Omega)$ for $2q \ge m + 2$, $w_n^q \in K(H^1_0(\Omega))$, and $w_n^q \to w$ as $n \to \infty$ in $L^2(\Omega)$ or $H^1_0(\Omega)$ depending on the regularity of w, cf. (4.17) and (4.18). The analogous reasoning applies to (iv). ∎

Let us briefly comment on the relation to the density results from theorem 4.4 and theorem 4.7. First, note that we do not require the obstacle to be bounded away from zero as we did in the preceding paragraphs. Secondly, the maximal regularity of the feasible approximation hinges on the coefficients of the elliptic operator associated with the obstacle and the smoothness of the boundary. Concerning the semicontinuity requirements of the upper bound, a classical result from Trudinger [23, 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 (4.12) is upper semicontinuous. 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 non-negative on Ω. For example, this is the case if a weak subsolution satisfies a Dirichlet boundary condition.

5. Application to finite elements

(a) Finite-element discretized convex sets

In the following, we investigate the issue of the Mosco convergence (definition 4.5) of finite-dimensional approximations K_n of a convex constraint set $K(X(\Omega))$ of the type (3.3); see §2b(ii) for a general motivation in the context of variational inequality problems. In this section, it is assumed that the sets (K_n) result from a suitable finite-element discretization such that the parameter n is associated with a sequence of mesh widths (h_n) tending to zero. The convergence of (K_n) in the sense of definition 4.5 ensures that the solutions of the discrete problems converge to the solution of the original infinite-dimensional problem irrespectively of the regularity of the data or the obstacle defining $K(X(\Omega))$; see [7, ch. 4, Theorem 4.1]. Mosco convergence results of this type are rarely found in the literature and are typically confined to simpler constraint sets and higher regularity assumptions on the obstacle; see, for instance, [4] for the case of an $H^1(\Omega) \cap C(\bar{\Omega})$-bound in the context of the obstacle problem. The density results from the preceding sections provide the basis for new Mosco convergence results under minimal regularity (of the solution) and under weaker assumptions on the regularity of the obstacle α. We further provide novel Mosco convergence results for discretized constraints on partial derivatives, including Raviart–Thomas finite-element approaches for problems in $H(\mathrm{div})$. As a general rule, density results of the type (1.1) represent a powerful means to verify the convergence of finite-element methods for convex constrained problems under minimal regularity. Applications involving constraint sets of the type (3.3) with low regularity of α are manifold and comprise, for instance, the discretization of variational problems in mechanics, such as in elasto-plasticity with hardening [24], and in image restoration, with regard to the predual problem of TV-regularization [25]. Moreover, the issue occurs in fixed point-based approaches to the solution of quasi-variational inequalities through the implicit definition of obstacles.

Remark 5.1. In some textbooks on finite-dimensional approximations of variational inequalities, cf. e.g. [4,6], condition (M2) is replaced by the following criterion:

$$\left.\begin{array}{l} \text{there exists a dense subset } \tilde{K} \subset K \text{ and an operator } r_n : \tilde{K} \to X \\[4pt] \text{such that for all } v \in \tilde{K} \text{ it holds } r_n v \to v \text{ in } X \text{ and there exists } n_0 \in \mathbb{N} \\[4pt] \text{such that } r_n v \in K_n \text{ for all } n \geq n_0. \end{array}\right\} \quad \text{(M2$'$)}$$

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 $K(Y(\Omega))$. This is precisely the point where the density results of §3 are needed.

Note that 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*.

By contrast, *adaptive* finite-element methods intend to design the sets K_n in order to approximate *the solution of a specific problem*. 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. [27]).

(b) Finite-element spaces and interpolation operators

In this section, we assume that $\Omega \subset \mathbb{R}^N$ is polyhedral. Together with Ω, a sequence of geometrically conforming 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 [28]. 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.

$$\exists c > 0 : \quad \frac{\operatorname{diam}(T)}{\rho_T} \leq c \quad \forall h \, \forall T \in \mathcal{T}_h, \tag{5.1}$$

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

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

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

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

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

$$I_h : C(\bar{\Omega}) \to P_{1,h}(\Omega) \quad \text{and} \quad I_h u := \sum_{x \in \mathcal{N}_h} u(x) \varphi_x. \tag{5.2}$$

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

$$\text{RT}_h(\Omega) := \{w \in L^2(\Omega)^N : w|_T \in \mathbb{RT} \; \forall T \in \mathcal{T}_h, [w \cdot v]|_{E \cap \Omega} = 0 \; \forall E \in \mathcal{E}_h\}, \tag{5.3}$$

where $\mathbb{RT} := \{w \in \mathbb{P}_1^d : \exists a \in \mathbb{R}^d, b \in \mathbb{R} : w(x) = a + bx\}$ and v denotes the unit outer normal to T. To incorporate homogeneous Neumann boundary conditions, one uses the $H_0(\text{div}; \Omega)$-conforming subspace

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

The construction of suitable edge-based basis functions $\{\varphi_E : E \in \mathcal{E}_h\}$ can be found in the literature, cf., for instance, [29], such that the boundary condition in the definition of $\text{RT}_{0,h}(\Omega)$ can be easily accounted for. The global Raviart–Thomas interpolation operator is given by

$$I_h^{\text{RT}} : W^{1,1}(\Omega)^N \to \text{RT}_h(\Omega), \quad I_h^{\text{RT}} w := \sum_{E \in \mathcal{E}_h} \left(\int_E w \cdot v \, d\mathcal{H}^{N-1} \right) \varphi_E. \tag{5.4}$$

(c) Mosco convergence results under minimal regularity

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, e.g. (2.8), 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. [30]). 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. [4]).

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 5.2. *Let $\Omega \subset \mathbb{R}^N$ be a polyhedral domain and $\alpha \in C(\bar{\Omega})$ with $\alpha(x) \geq 0$ in Ω. Further let (w_h) be a sequence that fulfils 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 $|\cdot|$. From the definition of j^*, we obtain that $i_K(w) = 0$ is equivalent to

$$(w, v) \leq \int_\Omega \alpha |v|_* \quad \forall v \in L^2(\Omega)^d. \tag{5.5}$$

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

$$\alpha_h := \sum_{T \in \mathcal{T}_h} \alpha(x_T)\chi_T \quad \text{and} \quad v_h := \sum_{T \in \mathcal{T}_h} v(x_T)\chi_T, \tag{5.6}$$

the piecewise constant interpolants of α and v, respectively. By definition of (a_h) and (v_h) as well as the uniform continuity of α and v it follows that $\alpha_h \to \alpha$ and $v_h \to v$, both 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

$$\int_\Omega w \cdot v \, dx \leftarrow \int_\Omega w_h \cdot v_h \, dx = \sum_{T \in \mathcal{T}_h} \int_T w_h \cdot v_h \, dx$$

$$= \sum_{T \in \mathcal{T}_h} \lambda(T) \, w_h(x_T) \cdot v_h|_T \, dx$$

$$\leq \sum_{T \in \mathcal{T}_h} \lambda(T)\alpha(x_T)|v_h|_T|_* \, dx$$

$$= \int_\Omega \alpha_h |v_h|_* \, dx \to \int_\Omega \alpha |v|_* \, dx, \tag{5.7}$$

which proves (5.5). ∎

Lemma 5.3. *Let $\Omega \subset \mathbb{R}^N$ be a polyhedral domain and $\alpha \in C(\bar\Omega)$ with $\alpha(x) \geq 0$ in Ω. Let (w_h) be a sequence that fulfils 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 5.2. 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 fulfils $\alpha(x) = (I_h\alpha)(x)$ for all $x \in \mathcal{N}_h$ and $\alpha_h \to \alpha$ strongly in $L^\infty(\Omega)^d$. By (5.7), we obtain

$$\int_\Omega w \cdot v \, dx \leftarrow \int_\Omega w_h \cdot v_h \, dx = \sum_{T \in \mathcal{T}_h} \frac{\lambda(T)}{N+1} \sum_{x \in \mathcal{N}_h \cap T} w_h(x) \cdot v_h|_T \, dx$$

$$\leq \sum_{T \in \mathcal{T}_h} \frac{\lambda(T)}{N+1} \sum_{x \in \mathcal{N}_h \cap T} |w_h(x)||v_h|_T|_*$$

$$\leq \sum_{T \in \mathcal{T}_h} \frac{\lambda(T)}{N+1} \sum_{x \in \mathcal{N}_h \cap T} \alpha(x)|v_h|_T|_*$$

$$= \int_\Omega \alpha_h |v_h|_* \, dx \to \int_\Omega \alpha |v|_* \, dx. \qquad \blacksquare$$

Theorem 5.4. *Let $\Omega \subset \mathbb{R}^N$ be a polyhedral domain and $\alpha \in C(\bar\Omega)$ such that (3.1) holds true. Then the sets*

$$K_h = \{w \in P_{1,h}(\Omega)^d : |w(x_T)| \leq \alpha(x_T) \text{ for all } T \in \mathcal{T}_h\} \tag{5.8}$$

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 5.2 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 [4, ch. II] and requires (3.4). Note that theorem 3.1 implies that the set

$$\tilde{K} := \{\varphi \in C^\infty(\bar\Omega)^d : |\varphi(x)| < \alpha(x) \text{ for all } x \in \bar\Omega\} \tag{5.9}$$

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 (5.2), we have the classical estimate,

$$\|u - I_h u\|_{L^\infty(\Omega)} \leq ch^2 \|u\|_{W^{2,\infty}(\Omega)} \quad \forall u \in W^{2,\infty}(\Omega). \tag{5.10}$$

Here, c denotes a constant independent of h on account of the shape-regularity of the triangulation (5.1) (cf. [28, 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 [28, Corollary 1.109]. Applying estimate (5.10) to the components of $w \in \tilde{K}$ and using the equivalence of norms on $\mathbb{R}^d$, one obtains that

$$\| \, |w - r_h w| \, \|_{L^\infty(\Omega)} \leq ch^2 \|w\|_{W^{2,\infty}(\Omega)^d}, \tag{5.11}$$

for a suitable modification of c. This implies

$$|r_h w(x)| \leq |w(x)| + ch^2 \|w\|_{W^{2,\infty}(\Omega)^d} \quad \forall x \in \Omega. \tag{5.12}$$

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

Corollary 5.5. *Under the conditions of theorem 5.4, the sequence (K_h) defined in (5.8) Mosco converges for $h \to 0$ to the set $K(L^2(\Omega)^d)$ in $L^2(\Omega)^d$.*

Proof. Again, lemma 5.2 implies that (M1) with $X = L^2(\Omega)^d$ holds true. For $\tilde{K}$ defined in (5.9) 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 5.4. ∎

Corollary 5.6. *Under the conditions of theorem 5.4, the node-based discrete sets*

$$K_h := \{w \in P_{1,h}(\Omega)^d : |w(x)| \leq \alpha(x) \, \forall x \in \mathcal{N}_h\}, \tag{5.13}$$

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 5.4, noting that (5.12) also implies $r_h w \in K_h \forall h \leq h_0$ with K_h according to the node-based definition (5.13). ∎

Remark 5.7. With the help of the density property (3.2) 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 (5.8) and (5.13) has to be replaced by the space

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

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. (5.3)).

Theorem 5.8. *Let $\Omega \subset \mathbb{R}^N$ be a polyhedral domain. Let $\alpha \in C(\bar{\Omega})$ such that (3.1) is satisfied. Then the sets*

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

Mosco converge to $K(H_0(\text{div}; \Omega))$ in $H(\text{div}; \Omega)$ 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 $H(\text{div}; \Omega)$, then it also weakly converges to w in $L^2(\Omega)^N$. Analogously to the proof of lemma 5.2 one concludes that

$|w| \leq \alpha$ a.e. in Ω. The continuity of the normal trace mapping

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

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

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

cf. (3.2). For the global Raviart–Thomas interpolation operator defined in (5.4), the following interpolation error estimate holds true [28, Corollary 1.115]:

$$\|u - I_h^{\mathrm{RT}}u\|_{L^\infty(\Omega)^N} + \|\mathrm{div}\,u - \mathrm{div}\,I_h^{\mathrm{RT}}u\|_{L^\infty(\Omega)} \leq ch\|u\|_{W^{1,\infty}(\Omega)^N}, \tag{5.14}$$

for all $u \in W^{2,\infty}(\Omega)^N$. Setting $r_h w := I_h^{\mathrm{RT}}w$ 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 $I_h^{\mathrm{RT}}w \to w$ in $H(\mathrm{div})$ for all $w \in \tilde{K}$, we may proceed analogously to the proof of theorem 5.4 to verify (M2′). ∎

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

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

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

Theorem 5.9. *Let $\Omega \subset \mathbb{R}^N$ be a polyhedral domain. Let $\alpha \in C(\bar{\Omega})$ such that (3.1) is satisfied. Define*

$$K_h := \{w \in P_{1,h}^{\partial\Omega}(\Omega)^d : |\nabla w|_T| \leq \alpha(x_T)\ \forall T \in \mathcal{T}_h\}. \tag{5.15}$$

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 5.2, one obtains for $v \in C_c(\Omega)^{N\times d}$ that

$$\int_\Omega \nabla w : v\,\mathrm{d}x \leftarrow \int_\Omega \nabla w_h : v\,\mathrm{d}x \leq \int_\Omega |\nabla w_h| |v|_*\,\mathrm{d}x \leq \int_\Omega \alpha_h |v|_*\,\mathrm{d}x \to \int_\Omega \alpha |v|_*\,\mathrm{d}x,$$

using α_h from (5.6). Therefore, (5.5) holds with ∇w in place of w, and (M1) is verified.

To prove (M2′), we consider again the global interpolation operator I_h from (5.2). The standard estimate

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

holds true (e.g. [28]). 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 [11, 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 5.4 to deduce (M2′). ∎

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

$$K_{\mathrm{div}}(X(\Omega)) := \{w \in X(\Omega) : |\mathrm{div}\,w| \leq \alpha \text{ a.e. in } \Omega\}. \tag{5.16}$$

Using Raviart–Thomas finite elements, a discrete realization of the inequality constraint in (5.16) 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 5.10. *Let $\Omega \subset \mathbb{R}^N$ be a polyhedral domain. Let $\alpha \in C(\bar{\Omega})$ fulfil (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 (5.16).

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 corollary 5.9. Since $K_{\mathrm{div}}(C_c^\infty(\Omega)^N)$ is dense in $K_{\mathrm{div}}(H_0(\mathrm{div}; \Omega))$ [11, 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 $r_h = I_h^{\mathrm{RT}}$, the estimate (5.14) implies $r_h w \to w$ in $H(\mathrm{div}; \Omega)$ and

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

for all w in $\tilde{K}$. In particular, one may argue as in the proof of theorem 5.4 to verify (M2′). ∎

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

$$\fint_T \alpha \, dx := \frac{\int_T \alpha \, dx}{\lambda(T)}$$

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 (3.4), which represent the main ingredient to prove the consistency of the finite-element approximation, may fail to hold true (e.g. theorem 4.2). On the other hand, the results from §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 non-negative measurable function $\alpha : \Omega \to \mathbb{R} \cup \{+\infty\}$ allows for the density property

$$\overline{K(C(\bar{\Omega}))}^{L^2(\Omega)^d} = K(L^2(\Omega)^d). \tag{5.17}$$

Here, we concentrate on the consistency in the L^2-topology but an extension to the other cases is possible by appropriately modifying assumption (5.17). We stress the fact that assumption (5.17) is fulfilled in relevant situations (cf. e.g. theorem 4.4).

Lemma 5.11. *Let $\Omega \subset \mathbb{R}^N$ be a polyhedral domain and $\alpha \in L^2(\Omega)$ with $\alpha(x) \ge 0$ a.e. in Ω. Let (w_h) be a sequence that fulfils for all h, $w_h \in P_{1,h}(\Omega)^d$ and $|w_h(x_T)| \le \fint_T \alpha \, dx$ 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 $|w| \le \alpha$ a.e. in Ω.*

Proof. The assertion follows analogously to the proof of lemma 5.2 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. ∎

Theorem 5.12. *Let $\Omega \subset \mathbb{R}^N$ be a polyhedral domain. Let $\alpha \in L^2(\Omega)$ with (3.1) such that (5.17) holds true. Then the sets*

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

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 corollary 5.11 implies (M1). First note that (3.1) and (5.17) imply that the set

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

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

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

for $w \in \tilde{K}$ and I_h as in (5.2). Integrating estimate (5.12) yields

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

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

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

which implies

$$|r_h w(x_T)| \leq \fint_T \alpha \, \mathrm{d}x - \delta(w) + ch^2 \|w\|_{W^{2,\infty}(\Omega)^d}, \quad \forall T \in \mathcal{T}_h. \tag{5.18}$$

This implies $r_h w \in K_h$ for all $w \in \tilde{K}$ and $h \leq h_0(w)$. By (5.10), it holds that $r_h w \to w$ in $L^2(\Omega)^d$ for $h \to 0$, which proves (M2′). ∎

6. Further applications

(a) 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 modelled by the evolution of the displacement, the material stress and strain as well as certain internal variables (cf. [6]). 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 $\Gamma_c \subset \partial\Omega$, and a variable q that is related to the deviatoric part of the material stress; for details we refer to [24, p.154]:

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

Here, $d := N(N+1)/2 - 1$ and G is a strongly convex, continuous and coercive functional that models the elasto-plastic material behaviour 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. $\beta(x) \geq \sigma_y$ a.e. in Ω. The normal stress z^* is constrained to lie in the polar cone

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

to the cone of non-negative functions

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

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 favourable to replace (6.1) by a combined Moreau–Yosida/Tikhonov regularization given by

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

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 [24, section 5]. In order to prove the stability of (6.1_γ) with regard to the limit problem (6.1) in the sense of theorem 2.1, 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. definition (2.2)). Here, we write for $\mathcal{X} \subset H^{1/2}(\Gamma_c)^* \times L^2(\Omega)^d$,

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

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

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

fulfils (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

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

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 fulfils*

(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) $r(z) \leq \|z^+\|^2_{L^2(\Gamma_c)}$ *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

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

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). ∎

From the discussion in the introduction and theorem 2.1, 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 §§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. [24] for details.

(b) Fenchel duality in image restoration

Optimization problems with total variation regularization have been successfully considered in the image restoration context. In the denoizing setting, an original image u_{true} that belongs to the space of functions of bounded variation $\text{BV}(\Omega)$, $\Omega \subset \mathbb{R}^2$, 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

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

for $\alpha \in \mathbb{R}$ in the seminal work by Rudin *et al.* [31]. Here, $\mathcal{D}u$, the distributional gradient of $u \in \text{BV}(\Omega)$, 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

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

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 [32,33] proposes the following alternative: For a function $\alpha : \Omega \to \mathbb{R}$ such that (3.1) holds true, consider the optimization problem

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

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:

$$\left. \begin{aligned} \min \quad & \tfrac{1}{2} \| \operatorname{div} p + f \|_{L^2(\Omega)}^2 \quad \text{over } p \in H_0(\operatorname{div}; \Omega) \\ \text{s.t.} \quad & p \in K(H_0(\operatorname{div}; \Omega), |\cdot|_\infty), \end{aligned} \right\} \tag{6.3}$$

if the following density result holds true:

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

where, according to the above notational convention,

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

7. Conclusion

We investigate the stability of a large number of perturbation and dualization approaches to variational inequality and constrained optimization problems in the context of density properties of a convex constraint set. If the intersection with certain dense subspaces is dense in the feasible set, one may prove the unconditional consistency of various perturbation schemes including Galerkin approximations. In this regard, the class of quasi-monotone perturbations provides a unified framework.

The abstract motivation leads to the study of density properties of constraint sets in Sobolev spaces with respect to spaces of smooth functions. We focus specifically on sets that are defined by a pointwise constraint on the norm of the function value. In this case, the density property is determined by the regularity of the upper bound. Whereas the case of a uniformly continuous obstacle gives rise to positive density results in various Sobolev spaces, the result fails to be valid in general, if the obstacle is just a Sobolev function. However, a large variety of discontinuous upper bounds still remains compatible with the density property. This includes functions that fulfil a generalized lower semicontinuity condition as well as supersolutions of elliptic PDEs.

Density results further allow to deduce the Mosco convergence of various finite-element discretized constraint sets in Sobolev spaces. Finally, the previous results are applied in the context of the regularization of quasi-static elasto-plastic contact problems and the dualization of total variation-based image restoration problems.

Our future research is concerned with the refined characterization of the class of upper bounds that comply with the density property. Another interesting direction of future research, which we plan to pursue, is related to constraint qualifications (CQs) in the context of Fenchel–Legendre dualization in convex and possibly non-smooth optimization. Here it appears that the density of convex intersections may provide a suitable constraint qualification implying duality without a duality gap. Such a density CQ appears to neither imply nor be contained in currently known constraint qualifications like those used in the work by Hedy Attouch and co-authors (e.g. [34]). Another current research aspect concerns the ramification of remark 3.2(ii). In the presence of an inhomogeneous Dirichlet boundary condition, the construction of suitable trace-preserving mollification operators with variable support appears promising. Those operators are also of utmost interest in the context of non-smooth variational problems in image restoration.

Appendix A. Properties of quasi-monotone perturbations

Proof of theorem 2.1. 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 neighbourhoods 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 [12]. 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 [12, ch. 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 [12, Prop. 6.3, Prop. 6.7, Prop. 5.7, Prop. 3.7]. Similarly, (2.3) together with [12, Prop. 6.7, Prop. 5.4] implies

$$\Gamma_w\text{-}\liminf_{n\to+\infty}(F + R_n) \geq \Gamma_w\text{-}\liminf_{n\to+\infty}(F + \underline{R}_n) = \lim_{n\to+\infty} \mathrm{sc}^-_w(F + \underline{R}_n), \tag{A 1}$$

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 [12, 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 (A 1). 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. ∎

Proof of proposition 2.6. 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:

$$\forall n \in \mathbb{N}\, \exists\, \gamma_n > 0 : \quad \left[y \in X \wedge \mathrm{dist}(y, K \cap \overline{B_\rho(x)})^2 < \frac{1}{\gamma_n} \Longrightarrow y \notin X_n \right]. \tag{A 2}$$

Assume that the opposite holds, i.e.

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

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

$$\mathrm{dist}(y_n, K)^2 = \mathrm{dist}(y_n, K \cap B_\rho(x))^2 \leq \frac{1}{\gamma_n},$$

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

References

1. Ekeland I, Temam R. 1987 *Convex analysis and variational problems*. Classics in applied mathematics, vol. 28. Philadelphia, PA: Society for Industrial and Applied Mathematics.
2. Attouch H, Buttazzo G, Michaille G. 2014 *Variational analysis in Sobolev and BV spaces: applications to PDEs and optimization*, 2nd edn. Philadelphia, PA: SIAM.

3. Clarke FH. 1990 *Optimization and nonsmooth analysis,* 2nd edn. Philadelphia, PA: Society for Industrial and Applied Mathematics.
4. Glowinski R. 1984 *Numerical methods for nonlinear variational problems*. Berlin, Germany: Springer.
5. Lions J-L, Glowinski R, Trémolières R. 1981 *Numerical analysis of variational inequalities*. Amsterdam, The Netherlands: North-Holland.
6. Han W, Reddy BD. 2013 *Plasticity: mathematical theory and numerical analysis*, 2nd edn. New York, NY: Springer.
7. Rodrigues JF. 1987 *Obstacle problems in mathematical physics*. Amsterdam, The Netherlands: North-Holland.
8. Rodrigues JF. 2004 On hyperbolic variational inequalities of first order and some applications. *Monatshefte für Mathematik* **142**, 157–177. (doi:10.1007/s00605-004-0238-3)
9. Dal Maso G, DeSimone A, Mora MG. 2006 Quasistatic evolution problems for linearly elastic-perfectly plastic materials. *Arch. Ration. Mech. Anal.* **180**, 237–291. (doi:10.1007/s00205-005-0407-0)
10. Bartels S, Mielke A, Roubíček T. 2012 Quasi-static small-strain plasticity in the limit of vanishing hardening and its numerical approximation. *SIAM J. Numer. Anal.* **50**, 951–976. (doi:10.1137/100819205)
11. Hintermüller M, Rautenberg CN. 2015 On the density of classes of closed convex sets with pointwise constraints in Sobolev spaces. *J. Math. Anal. Appl.* **426**, 585–593. (doi:10.1016/j.jmaa.2015.01.060)
12. Dal Maso G. 1993 *An introduction to Γ-convergence*. Boston, MA: Birkhäuser.
13. Mielke A, Roubíček T. 2015 *Rate-independent Systems*. Berlin, Germany: Springer.
14. Kinderlehrer D, Stampacchia G. 1980 *An introduction to variational inequalities and their applications*. New York, NY: Academic Press.
15. Mosco U. 1969 Convergence of convex sets and of solutions of variational inequalities. *Adv. Math. (N.Y)* **3**, 510–585. (doi:10.1016/0001-8708(69)90009-7)
16. Alt HW. 2012 *Lineare Funktionalanalysis*. Berlin, Germany: Springer.
17. Adams RA, Fournier JJF. 2003 *Sobolev spaces,* 2nd edn. Pure and Applied Mathematics Series, vol. 140. Amsterdam, The Netherlands: Elsevier Science.
18. Evans LC. 1998 *Partial differential equations*. Providence, RI: American Mathematical Society.
19. Hintermüller M, Rautenberg CN, Rösel S. 2016 *Density of convex intersections and applications*. Berlin, Germany: WIAS. WIAS Preprint No. 2333.
20. Toyoizumi H. 1991 Continuous dependence on obstacles in variational inequalities. *Funkcialaj Ekvacioj* **34**, 103–115.
21. Nečas J. 1967 *Les méthodes directes en théorie des équations elliptiques*. Prague, Czechoslovakia: Academia.
22. Grisvard P. 1985 *Elliptic problems in nonsmooth domains*. London, UK: Pitman Publishing.
23. Trudinger NS. 1971 On the regularity of generalized solutions of linear, non-uniformly elliptic equations. *Arch. Ration. Mech. Anal.* **42**, 50–62. (doi:10.1007/BF00282317)
24. Hintermüller M, Rösel S. 2016 A duality-based path-following semismooth Newton method for elasto-plastic contact problems. *J. Comput. Appl. Math.* **292**, 150–173. (doi:10.1016/j.cam.2015.06.010)
25. Hintermüller M, Kunisch K. 2004 Total bounded variation regularization as bilaterally constrained optimization problem. *SIAM J. Appl. Math.* **64**, 1311–1333. (doi:10.1137/S0036139903422784)
26. Siebert KG, Veeser A. 2007 A unilaterally constrained quadratic minimization with adaptive finite elements. *SIAM. J. Optim.* **18**, 260–289. (doi:10.1137/05064597X)
27. Siebert KG. 2011 A convergence proof for adaptive finite elements without lower bound. *IMA J. Num. Anal.* **31**, 947–970. (doi:10.1093/imanum/drq001)
28. Ern A, Guermond JL. 2004 *Theory and practice of finite elements*. Berlin, Germany: Springer.
29. Bahriawati C, Carstensen C. 2005 Three MATLAB implementations of the lowest-order Raviart-Thomas MFEM with a posteriori error control. *Comput. Methods Appl. Math.* **5**, 333–361. (doi:10.2478/cmam-2005-0016)
30. Brezzi F, Hager WW, Raviart PA. 1977 Error estimates for the finite element solution of variational inequalities. Part I. Primal theory. *Numerische Mathematik* **28**, 431–444. (doi:10.1007/BF01404345)
31. Rudin LI, Osher S, Fatemi E. 1992 Nonlinear total variation based noise removal algorithms. *Phys. D* **60**, 259–268. (doi:10.1016/0167-2789(92)90242-F)

32. Hintermüller M, Rautenberg CN. 2017 Optimal selection of the regularization function in a generalized total variation model. Part I: Modelling and theory. *J. Math. Imaging Vis.* (doi:10.1007/s10851-017-0744-2)

33. Hintermüller M, Rautenberg CN, Wu T, Langer A. 2017 Optimal selection of the regularization function in a generalized total variation model. Part II: Algorithm its analysis and numerical tests. *J. Math. Imaging Vis.* (doi:10.1007/s10851-017-0736-2)

34. Attouch H, Brezis H. 1986 Duality for the sum of convex functions in general Banach spaces. In *Aspects of mathematics and its applications*. North-Holland Mathematical Library, vol. 34, pp. 125–133. Amsterdam, The Netherlands: North-Holland.