ON SOME QUASI-VARIATIONAL INEQUALITIES AND OTHER PROBLEMS WITH MOVING SETS *

JOSÉ-LUIS MENALDI† AND CARLOS N. RAUTENBERG‡

Abstract. Since its introduction over 50 years ago, the concept of Mosco convergence has permeated through diverse areas of mathematics and applied sciences. These include applied analysis, the theory of partial differential equations, numerical analysis, and infinite dimensional constrained optimization, among others. In this paper we explore some of the consequences of Mosco convergence on applied problems that involve moving sets, with some historical accounts, and modern trends and features. In particular, we focus on connections with density of convex intersections, finite element approximations, quasi-variational inequalities, and impulse problems.

On the occasion of the 80th birthday of Umberto Mosco

Key words. Mosco convergence, Variational inequalities, Quasi-variational inequalities.

AMS subject classifications. 35J86, 35J60, 35R35, 65K10, 93E20

1. Introduction. The overwhelming success of Mosco convergence [64, 63] is present in several areas of mathematics. The concept provided the right framework for the study of problems involving *moving* convex sets in reflexive Banach spaces. In fact, it made possible to study perturbation/stability properties for variational inequalities and other nonlinear problems in calculus of variations, provided existence results for quasi-variational inequalities (QVIs), and allowed the study of optimization problems where control/design variables modify constraints.

Historically, set convergence notions in abstract spaces go back to Painlevé in the beginning of the nineteenth century. They were, however, popularized by Kuratowski [52] who remains as part of the name in the set limit names. In between the appearance of the latter, and the one of the seminal paper by Di Giorgi and Franzoni [31] on Γ-convergence (a full analysis of relationships between $K-$ and $\Gamma-$limits can be found in the monograph by Dal Maso [30]), Mosco published his foundational results [64, 63] associated to the convergence of sets. The initial results of Mosco are almost simultaneous to the famous Lions and Stampacchia paper *Variational Inequalities* [55]. This is not coincidence, as in Mosco's own words, it was Stampacchia "who suggested this research". Further, we direct the interested reader to Aubin and Frankowska [12] for an historical account on the notions of set convergence.

There is a vast literature on *Impulse Control Problems* and their connections with QVIs, the reader may check the books by Bensoussan and Lions [22] and Bensoussan [20] for a self-contained account on the subject. In fact, QVIs were initially identified and treated by Bensoussan and Lions [21, 53] through impulse control problems. Hamilton-Jacobi-Bellman equations take the form of QVIs in many applications to (stochastic) control/design problems. Thus, the convergence of sets (which is relevant in the setting of the problems itself) becomes very important in the approximation and implementation of these models. A very short description is given later on, es-

†Department of Mathematics, Wayne State University, Detroit, MI 48202, USA. (menaldi@wayne.edu)

‡Department of Mathematical Sciences and the Center for Mathematics and Artificial Intelligence (CMAI), George Mason University, Fairfax, VA 22030, USA. (crautenb@gmu.edu)

sentially, to alert the reader some other similar (yet different) type of QVIs, where Mosco's convergence has not been completely discussed and explored.

From a general perspective, QVIs are nonlinear, nonconvex, and nonsmooth problems with (in general) non-unique solutions. Specifically, a QVI can be seen as variational problem with an implicit (state dependent) constraint. This leads to the need to approach these problems from the moving set perspective. This level of complexity established QVIs as powerful physical models. It should be noted that QVIs have been successfully applied to the magnetization of superconductors, Maxwell systems, thermohydraulics, image processing, game theory, surface growth of granular (cohensionless) materials, hydrology, and solid and continuum mechanics. For more details, we refer the reader to [37, 48, 32, 40, 51, 54, 68, 71] and to the monographs [15, 50].

In this paper, in addition to providing an account of some basic sufficient conditions for Mosco convergence for several types of sets, we focus on two main consequential aspects associated to Mosco convergence. Initially, we study its relationship with density properties for convex sets, and provide application to finite element discretization of problems involving convex sets. Subsequently, we focus on quasi-variational inequalities, impulse problems, and some stability properties of the set of solutions to the QVI.

As the notion of Mosco convergence can also be described via a functional convergence related to Γ-convergence (this is detailed within the paper), the concept is directly applicable to the study of regularized minimization problems in highly diverse settings. Some of these that are not fully in the scope of this paper include viscosity solutions of minimization problems [6, 7], derivation of variational models for granular material accumulation [47, 27], Tikhonoff regularization for inverse problems [67, 66], and posteriori error estimates for adaptive finite elements [78, 35] (and references within [76]).

The paper is organized as follows. In section 2 we provide some common notation used throughout the paper, and the famous definition of Mosco convergence together with basic results involving general classes of convex sets. A short account on necessary and sufficient conditions of Mosco convergence for unilateral sets are given in section 3. The role of density properties for convex sets in Mosco convergence is explored in section 4, and its subsequent application to finite element discretization is provided on section 4.1. Quasi-Variational Inequalities are considered in section 5, and a short, historical and modern account on impulse control problems is given next on section 5.1. We finalize the paper with a short account on existence of QVIs and stability results for multivalued problems in sections 5.2 and 5.3, respectively.

2. Notations and Preliminary Results. Throughout most of this paper we assume (unless stated otherwise) that V is a reflexive real Banach space of (equivalence) classes of maps of the type $v : \Omega \to \mathbb{R}$ for some Lipschitz domain $\Omega \subset \mathbb{R}^d$ with $d \in \mathbb{N}$. For an arbitrary Banach space V we write $\| \cdot \|_V$ for its associated norm. The topological dual is denoted as V', and by $\langle \cdot, \cdot \rangle_{V',V}$ the associated duality pairing. For a sequence $\{z_n\}_{n \in \mathbb{N}}$ in V we denote its strong convergence to $z \in V$ by "$z_n \to z$" and weak convergence by "$z_n \rightharpoonup z$". Further, for two Banach spaces V_1 and V_2, we write $\mathscr{L}(V_1, V_2)$ for the space of bounded linear operators from V_1 to V_2.

The typical function spaces under consideration are described next. For an open domain Ω in $\mathbb{R}^d$, we denote $H^1(\Omega)$ to be the Sobolev space of $L^2(\Omega)$ functions whose weak gradients belong to $L^2(\Omega)$, and by $H_0^1(\Omega)$ we denote the subset of $H^1(\Omega)$ whose elements are zero in $\partial \Omega$ in the sense of the trace (provided that Ω is regular enough). For functions in $L^p(\Omega)$ with gradients in $L^p(\Omega)$, we utilize $W^{1,p}(\Omega)$, and $W_0^{1,p}(\Omega)$ for

functions vanishing at $\partial\Omega$.

We are now in position to establish the notion of Mosco convergence and some of its basic consequences. The celebrated definition by Mosco [64, 63] is the following

DEFINITION 2.1 (MOSCO CONVERGENCE). *Let $\mathbf{K}$ and $\mathbf{K}_n$, for each $n \in \mathbb{N}$, be non-empty, closed and convex subsets of V. Then the sequence $\{\mathbf{K}_n\}$ is said to converge to $\mathbf{K}$ in the sense of Mosco as $n \to \infty$, denoted by*

$$\mathbf{K}_n \xrightarrow{\ \mathrm{M}\ } \mathbf{K},$$

if the following two conditions are fulfilled:
  (I) *For each $w \in \mathbf{K}$, there exists $\{w_{n'}\}$ such that $w_{n'} \in \mathbf{K}_{n'}$ for $n' \in \mathbb{N}' \subset \mathbb{N}$ and $w_{n'} \to w$ in V.*
 (II) *If $w_n \in \mathbf{K}_n$ and $w_n \rightharpoonup w$ in V along a subsequence, then $w \in \mathbf{K}$.*

In general and in concrete applications, item (II) in Definition 2.1 is significantly simpler to check than (I). In fact, (I) requires clever constructions that leads into problem-tailored approaches.

The relevance of Mosco convergence can be explained by the fact that it provides the right "topology" for the obtention of stability results to solutions of variational inequalities when the constraint sets are perturbed. For this matter, consider $\mathbf{K} \subset V$ non-empty, closed and convex and $f \in V'$. We define $S(f, \mathbf{K})$ as the unique solution to the following variational inequality (VI):

$$\text{Find } y \in \mathbf{K} : \langle Ay - f, v - y \rangle \geq 0, \quad \forall v \in \mathbf{K}, \tag{2.1}$$

where

$$A : V \to V' \text{ is linear, bounded, and strongly monotone}, \tag{2.2}$$

and we assume this about A throughout the paper. For a detailed account of problem (2.1), we refer the author to [49] or [15]. Then, we have the following result by Mosco: If $f_n \to f$ in V', we have that

$$\mathbf{K}_n \xrightarrow{\ \mathrm{M}\ } \mathbf{K} \quad \text{implies} \quad S(f_n, \mathbf{K}_n) \to S(f, \mathbf{K}) \text{ in } V;$$

see [64] or [73].

In this paper we focus on two classes of problems that share similar difficulties with "moving" sets and hence Mosco convergence becomes a crucial tool in their treatment. Initially, we focus on optimization problems and their regularization/discretization and limiting behavior. In particular, we deal with the issue of Mosco convergence via properties of density of convex intersections. Secondly, we focus on some particular classes of quasi-variational inequalities (variational problems with implicit obstacles), and stability properties of the solution set.

We consider a general structure of the sets of interest that is wide enough to include pointwise bounds on function values, their gradient, curl or divergence, and also nonlocal type constraints like the ones arising from linear integral operators. The general structure of the sets of interest are of the form

$$\mathbf{K} = \{w \in V : \psi(Gw) \leq \phi\}, \tag{2.3}$$

where $\phi : \Omega \to \mathbb{R}$ is a nonnegative measurable function and "$v \leq w$" stands for $v(x) \leq w(x)$ for almost all (f.a.a.) $x \in \Omega$, or almost everywhere (a.e.), unless stated

otherwise. We assume that $G \in \mathscr{L}(V, L^p(\Omega)^{\mathrm{m}})$ for some $1 < p < +\infty$ and $\mathrm{m} \in \mathbb{N}$, that is, $G \colon V \to L^p(\Omega)^{\mathrm{m}}$ is linear and bounded. Additionally, we suppose that $\psi \colon \mathbb{R}^{\mathrm{m}} \to \mathbb{R}$ is convex, $\psi(0) = 0$, $\psi(tx) = t\psi(x)$ for all $x \in \mathbb{R}^{\mathrm{m}}$ and $t > 0$, and it is possibly nonsmooth at the origin but smooth everywhere else. Note that the previous implies that $\mathbf{K}$ is convex, and closed. Further, since $\phi \geq 0$ we have that $\mathbf{K}$ is nonempty as well since $0 \in \mathbf{K}$.

A few words are in order to establish the generality of the structure of (2.3). The class of spaces we have in mind are either of Lebesgue or Sobolev type. The possible choice of G is contingent upon the choice of V; for example, if $V = H_0^1(\Omega)$ then G can be considered as the weak gradient ∇, and if $H_0(\mathrm{div}; \Omega)$, we can take $G = \mathrm{div}$. The function ψ commonly refers to a ℓ^p-norm in $\mathbb{R}^{\mathrm{m}}$, and the absolute value if $\mathrm{m} = 1$, or to just the identity, i.e., $\psi(x) = x$. The regularity of ϕ is not an issue for well-posedness of variational problems over $\mathbf{K}$, but additional properties will be required for the obtention of stability results for perturbations of ϕ.

Note that the expression (2.3), for given functions ψ, G, and ϕ, determines a fixed closed and convex set $\mathbf{K}$. However, for problems like quasi-variational inequalities (QVIs), the set $\mathbf{K}$ is actually a state-dependent quantity: This would lead to a problem like 2.1 where $y \mapsto \mathbf{K}(y)$ is not constant. In terms of (2.3) and this setting, the dependence of $\mathbf{K}(y)$ on y is determined by assuming that $\phi = \Phi(y)$; this is discussed later on. Based on the structure of (2.3), we have a general result under relatively weak conditions for (II) in Definition 2.1 to hold.

PROPOSITION 2.2. *Suppose that $\phi_n \to \phi$ in $L^1(\Omega)$. Then* (II) *in Definition 2.1 holds true for*

$$\mathbf{K}_n = \{w \in V : \psi(Gw) \leq \phi_n\}. \tag{2.4}$$

Proof. For $w_n \in \mathbf{K}_n$, we have $\psi(Gw_n) \leq \phi_n$, and if $w_n \rightharpoonup w$ in V, it follows that $Gw_n \rightharpoonup Gw$ in $L^p(\Omega)^d$. By Mazur's lemma, there exists $z_n = \sum_{k=n}^{N(n)} \alpha_k(n) Gw_k$ where $\sum_{k=n}^{N(n)} \alpha_k(n) = 1$ and $\alpha_k(n) \geq 0$ such that $z_n \to Gw$ in $L^p(\Omega)^d$. Since $\psi \colon \mathbb{R}^d \to \mathbb{R}$ is convex,

$$\psi(z_n) \leq \sum_{k=n}^{N(n)} \alpha_k(n)\psi(Gw_k) \leq \phi + \underbrace{\sum_{k=n}^{N(n)} \alpha_k(n)|\phi_k - \phi|}_{\epsilon_n}.$$

Since $\phi_n \to \phi$ in $L^1(\Omega)$, then $\epsilon_n \to 0$ in $L^1(\Omega)$. Therefore, we obtain $w \in \mathbf{K}$ by taking the limit above (over some subsequence converging in the pointwise almost everywhere sense). □

On the other hand, the existence of the subsequence in (I) of Definition 2.1 requires problem-specific constructions rendering it (in general) much harder to prove than the condition in (II) . Perhaps the simplest situation in which (I) holds is the obstacle case with $\phi_n \to \phi$ in V and where $V \ni z \mapsto \min(0, z) \in V$ is continuous: Let $w \leq \phi$ be arbitrary and define $w_n := \min(w, \phi_n)$ so that $w_n \leq \phi_n$. Since $\phi_n \to \phi$ in V, it follows that $w_n \to w$ in V. Consequently (I) holds true. The relaxation of " $\phi_n \to \phi$ in V" is a complex task that we tackle in some simple cases. We provide now some general constructions for (I) and for Mosco convergence.

PROPOSITION 2.3. *Let V be either $W_0^{1,p}(\Omega)$ or $W^{1,p}(\Omega)$ with $1 \leq p < +\infty$, and $\phi, \phi_n \in L^\infty(\Omega)$ for $n \in \mathbb{N}$. Suppose that for some $\nu \in [0, \infty)$, $\phi_n \geq \nu$ a.e. for all $n \in \mathbb{N}$ and*

$$\phi_n \to \phi \quad in \quad L^\infty(\Omega).$$

Then, if either one of the two conditions

> (i) $\nu = 0$, $G = \mathrm{id}$ *and* $\psi(x) = |x|$ *or* $\psi(x) = x$
>
> (ii) $\nu > 0$, $G = \nabla$ *and* $\psi(x) = \|x\|_{\ell^q}$ *with* $1 \leq q \leq +\infty$

hold true, we observe that

$$\{w \in V : \psi(Gw) \leq \phi_n\} \xrightarrow{\mathrm{M}} \{w \in V : \psi(Gw) \leq \phi\}.$$

Proof. Suppose that $\nu > 0$. Note that since $\phi_n \to \phi$ in $L^\infty(\Omega)$ by Proposition 2.2, (II) in Definition 2.1 holds true. In order to prove (I), let $w^* \in V$ and $\psi(Gw^*) \leq \phi$ be arbitrary. Define $w_n := \beta_n w^*$ where

$$\beta_n := \left(1 + \frac{\|\phi_n - \phi\|_{L^\infty(\Omega)}}{\nu}\right)^{-1}.$$

It follows that $w_n \to w^*$ in $W_0^{1,p}(\Omega)$ and $\psi(Gw_n) \leq \phi_n$ (cf. Hintermüller and Rautenberg [41]) which finishes the proof.

In the case $\nu = 0$ for (i), consider

$$w_n = T_n(w^*) := \begin{cases} (\psi(w^*) - \|\phi - \phi_n\|_{L^\infty(\Omega)})^+ \frac{w^*}{\psi(w^*)}, & \text{if } \psi(w^*) \neq 0; \\ 0, & \text{if } \psi(w^*) = 0. \end{cases}$$

Note first that $w_n \in L^p(\Omega)$ and clearly $w_n \to w^*$ in $L^p(\Omega)$. Further, note that $T_n(w^*)(x) = h_n(w^*(x))$ for some $h_n : \mathbb{R} \to \mathbb{R}$ and for each $n \in \mathbb{N}$. In this case, we have that $x \mapsto h_n(x)$ is globally Lipschitz, $h_n(0) = 0$, then it follows that $T_n : V \to V$ is continuous (see Marcus and Mizel [56]). Hence $\nabla w_n \in L^p(\Omega)$ and further

$$\nabla w_n = \nabla w^* \, T_n'(w^*) \qquad \text{with} \qquad T_n'(x) := \begin{cases} x/\psi(x), & \text{if } \psi(x) \geq \|\phi - \phi_n\|_{L^\infty(\Omega)}; \\ 0, & \text{otherwise.} \end{cases}$$

where we have used that $\psi(x) = x$ or $\psi(x) = |x|$. Hence, $\nabla w_n \to \nabla w^*$ in $L^p(\Omega)$ given that $T_n'(w) \to x/\psi(x)$ in $L^q(\Omega)$ for any $1 \leq q < \infty$, i.e., $w_n \to w^*$ in V. Finally, for $x \in \Omega$ we observe

$$\psi(w_n(x)) = |(\psi(w(x)) - \|\phi - \phi_n\|_{L^\infty(\Omega)})^+| \leq |(\psi(w(x)) - \phi(x) + \phi_n(x))^+| \leq \phi_n(x),$$

which completes the proof. $\square$

The case $\nu = 0$ can also be handled in the gradient constraint case under additional assumptions on the regularity of the domain Ω. Indeed, if Ω is bounded with $\partial\Omega$ of class C^2, then the result of the above theorem holds true in the (i) case, e.g. see Azevedo and Santos [13].

3. Further sufficient and necessary conditions for unilateral sets. In applications, it is common to encounter obstacle-type (or unilateral) constraints, i.e.,

$$\mathbf{K} = \{w \in W_0^{1,p}(\Omega) : w \leq \phi\}, \quad \text{and} \quad \mathbf{K}_n = \{w \in W_0^{1,p}(\Omega) : w \leq \phi_n\}, \qquad (3.1)$$

for $1 < p < +\infty$, and Ω open and bounded. As we have shown before, it is simple to observe that if $\phi_n \to \phi$ in $W_0^{1,p}(\Omega)$, then $\mathbf{K}_n \xrightarrow{\mathrm{M}} \mathbf{K}$. This, however, can be relaxed significantly and still preserve the Mosco convergence as we briefly discuss next.

The study of sufficient and of necessary conditions for convergence in the sense of Mosco has been an active area of research for several decades. In the case of unilateral sets and in $W_0^{1,p}(\Omega)$, a complete answer was given by Dal Maso [29] where the condition involves properties on the $p-$capacities of the sets $\{x \in \Omega : \phi_n(x) < t\}$ and $\{x \in \Omega : \min(\phi(x), \phi_n(x)) < t\}$. A similar capacitary approach was taken by Attouch and Picard [9, 10, 8] and sufficient conditions were obtained under stronger conditions than Dal Maso.

As a sufficient condition for Mosco convergence, $\phi_n \to \phi$ in $W_0^{1,p}(\Omega)$ can be relaxed also substantially by means of the compactness result in Murat [65] for Lipschitz domains that states: If $F_n \rightharpoonup F$ in $H^{-1}(\Omega)$ with $F_n \geq 0$ for all $n \in \mathbb{N}$, then $F_n \to F$ in $W^{-1,q}(\Omega)$ with $q < 2$. Here, $F_n \geq 0$ refers to $\langle F_n, \sigma \rangle \geq 0$ for all $\sigma \in H_0^1(\Omega)$ with $\sigma \geq 0$. Moreover, the Lipschitz regularity of $\partial\Omega$ can be dropped and the result still remains intact; see Brézis [28]. In our setting, this result leads to the following useful assertion; see Boccardo and Murat [26, 25]: If $\phi_n \rightharpoonup \phi$ in $W^{1,q}(\Omega)$ or $W_0^{1,q}(\Omega)$ for some $q > p$, then $\mathbf{K}_n \xrightarrow{\mathrm{M}} \mathbf{K}$. In summary, Mosco convergence is maintained when switching from strong into weak convergence of the obstacles, provided that the gradients of the obstacles possess an $\epsilon > 0$ extra amount of integral regularity.

Analogous results to the one of Boccardo and Murat and of Dal Maso [29] were unknown for fractional spaces $W_0^{s,p}(\Omega)$ for $s \in (0, 1)$ until recently (see [4]). Applications for these kind of problems can be seen in Antil and Rautenberg [3]. In the same vein, it is an open question whether it is possible to extend the above result of Boccardo and Murat [26, 25] to weighted Sobolev spaces $W_0^{1,p}(\Omega; w)$ for some w in a Muckenhoupt class.

4. The role of density in Mosco convergence. This section entails a discussion on how density properties are related to Mosco convergence in regularization/discretization of optimization problems; we follow closely [45].

In variational problems with constraints, one seeks the solution in a given convex, closed and nonempty feasible set $\mathbf{K}$ of a certain Banach space $(V, \|\cdot\|)$ not necessarily reflexive. To start the discussion in this section, let us consider the following abstract class of optimization problems:

$$\min F(u) \quad \text{over} \quad u \in \mathbf{K}, \tag{4.1}$$

where assume that $F : V \to \mathbb{R}$ is continuous, coercive and sequentially weakly lower semicontinuous (not necessarily convex).

Problem (4.1) admits a solution provided V is reflexive: Let $\{u_n\}$ be an infimizing sequence. Since F is coercive and there is a feasible point, $\{u_n\}$ is bounded. Since V is reflexive, $u_n \rightharpoonup u^*$ along a subsequence for some $u^* \in V$. Since $\mathbf{K}$ is convex and closed, it is weakly closed, and hence $u^* \in \mathbf{K}$. Finally, since F is sequentially weakly lower semicontinuous, we have

$$F(u^*) \leq \liminf F(u_n) = \inf_{u \in \mathbf{K}} F(u),$$

i.e., u^* is a minimizer, and a subsequence of $\{u_n\}$ is not only a infimizing sequence but a minimizing one as well.

The problem class (4.1) is general enough to encompass numerous fields, such as variational inequality problems of potential type, and optimal control of partial differential equations with constraints on the state and/or control among others. The study of (4.1) and the design of solution algorithms involve concepts of perturbation

or dualization methods comprising regularization, penalization or discretization approaches (or a combination thereof). The stability properties of (4.1) with respect to a large class of perturbations is contingent upon the following density property: For a particular dense subspace Y of V, it holds true that

$$\overline{\{u \in Y : u \in \mathbf{K}\}}^V = \mathbf{K}, \tag{4.2}$$

or in short $\overline{\mathbf{K} \cap Y}^V = \mathbf{K}$. Note that $\overline{\mathbf{K} \cap Y}^V$ refers to the closure in the V-norm of the set $\mathbf{K} \cap Y$. In order to prove this, we start with the definition of Γ-convergence and its relation to Mosco convergence.

DEFINITION 4.1. *Let $G_n : V \to \mathbb{R} \cup \{+\infty\}$ for $n \in \mathbb{N}$. We define the Γ-upper and -lower limit at u of G_n as*

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

and

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

respectively, where $\mathcal{N}(u)$ denotes the set of all open neighborhoods in the norm of V. Analogously, we denote the weak versions of the above $\Gamma_w\text{-}\limsup G_n(u)$ and $\Gamma_w\text{-}\liminf G_n(u)$ where open neighborhoods are considered in the weak topology. Provided the limits exists and are identical, we write

$$\Gamma\text{-}\lim_{n \to +\infty} G_n(u) := \Gamma\text{-}\limsup_{n \to +\infty} G_n(u) = \Gamma\text{-}\liminf_{n \to +\infty} G_n(u),$$

and say the quantity above is the (norm) Γ-limit of G_n at u. Similarly, in the weak topology case, we define

$$\Gamma_w\text{-}\lim_{n \to +\infty} G_n(u) := \Gamma_w\text{-}\limsup_{n \to +\infty} G_n(u) = \Gamma_w\text{-}\liminf_{n \to +\infty} G_n(u),$$

provided the limits exist and are equal.

The connection of Γ-convergence and Mosco convergence is immediate. Consider the sequence $\{i_{\mathbf{K}_n}\}$ of indicator functions $i_{\mathbf{K}} : V \to \mathbb{R} \cup \{+\infty\}$ for the sequence of convex closed and non-empty sets $\{\mathbf{K}_n\}$. Then, $\mathbf{K}_n \xrightarrow{\mathrm{M}} \mathbf{K}$ if and only if for each sequence $u_n \rightharpoonup u$ in V with $u_n \in \mathbf{K}_n$, we have

$$\liminf_{n \to +\infty} i_{\mathbf{K}_n}(u_n) \geq i_{\mathbf{K}}(u),$$

and for each $u \in \mathbf{K}$, there exists a sequence such $u_n \in \mathbf{K}_n$ such that $u_n \to u$ and

$$\limsup_{n \to +\infty} i_{\mathbf{K}_n}(u_n) \leq i_{\mathbf{K}}(u).$$

We consider the above concepts applied to a general class of problems. For this matter, we define the sequence of perturbed problems

$$\inf \quad F(u) + R_n(u), \quad \text{over } u \in V, \tag{4.3}$$

defined by given perturbations $R_n : V \to \mathbb{R} \cup \{+\infty\}$ of the indicator function $i_{\mathbf{K}} : V \to \mathbb{R} \cup \{+\infty\}$ such that there exist functions $\underline{R}_n : X \to \mathbb{R} \cup \{+\infty\}$ and $\overline{R}_n : X \to \mathbb{R} \cup \{+\infty\}$ where

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

and the additional properties hold

$$\underline{R}_n \le \underline{R}_{n+1} \ \forall n \in \mathbb{N}, \quad \lim_{n \to +\infty} \underline{R}_n(u) = i_{\mathbf{K}}(u) \ \forall u \in V,$$

$$\underline{R}_n \text{ sequentially weakly lower semicontinuous } \forall n \in \mathbb{N},$$

(4.4)

i.e., if $v_k \rightharpoonup v$ then $\underline{R}_n(v) \le \liminf_k \underline{R}_n(v_k)$, and

$$\overline{R}_n \ge \overline{R}_{n+1}, \ \forall n \in \mathbb{N}, \quad \lim_{n \to +\infty} \overline{R}_n(u) = i_{\mathbf{K} \cap Y}(u) \quad \forall u \in V. \tag{4.5}$$

Mappings R_n that share the above features are usually called *quasi-monotone perturbations of the indicator function* $i_{\mathbf{K}}$ *with respect to the (dense) subspace* Y. We assume no additional assumptions for R_n itself. In the stability analysis of (4.3), the condition (4.2) appears immediately if using the theory of Γ-convergence ([30]): Under mild assumptions on V, the density property (4.2) ensures that $F + i_{\mathbf{K}}$ is the Γ-limit of $F + R_n$ in both, the weak and strong topology. In this setting, the problem (4.3) admits a minimizer u, and each weak cluster point of any sequence of minimizers $\{u_n\}$ is a minimizer of (4.1); see Dal Maso [30, Corollary 7.20].

We are now in position to establish the relation between Γ-convergence (and Mosco convergence) to the density property (4.2)

THEOREM 4.2. *Let* $\{R_n\}$ *be a sequence of quasi-monotone perturbations of* $i_{\mathbf{K}}$ *with respect to the dense subspace* Y. *Let the Banach space* V *be reflexive or assume that* V^* *is separable. If the density property* (4.2) *holds true, then* $F + i_{\mathbf{K}}$ *is the* Γ-*limit of* $F + R_n$ *in both, the weak and strong topology.*

Proof. Denote by $\mathrm{sc}^- G$ to the lower semicontinuous envelope of $G : V \to \mathbb{R} \cup \{+\infty\}$. From the relationship between Γ- and pointwise convergence [30, Chapter 5], with (4.5) and the continuity of F, we observe

$$\begin{aligned} \Gamma_w\text{-}\limsup(F + R_n) &\le \Gamma\text{-}\limsup(F + R_n) \\ &\le \Gamma\text{-}\limsup(F + \overline{R}_n) \\ &= \mathrm{sc}^-(F + i_{\mathbf{K} \cap Y}) \\ &= F + i_{\overline{\mathbf{K} \cap Y}}, \end{aligned}$$

where we use [30, Prop. 6.3, Prop. 6.7, Prop. 5.7, Prop. 3.7].

Analogously, (4.4) together with [30, Prop. 6.7, Prop. 5.4] leads to

$$\Gamma_w \liminf(F + R_n) \ge \Gamma_w \liminf(F + \underline{R}_n) = \lim_{n \to +\infty} \mathrm{sc}_w^-(F + \underline{R}_n) \tag{4.6}$$

where $\mathrm{sc}_w^-(F + \underline{R}_n)$ denotes the lower semicontinuous envelope of $F + \underline{R}_n$ in the weak topology of V. In addition, note that the coercivity and the sequential weak lower semicontinuity of $F + \underline{R}_n$ imply that the level sets $\{u \in V : F(u) + R_n(u) \le t\}$ are bounded and sequentially weakly closed. Since V is reflexive or it has a separable dual V^*, then the sequential weak closure of bounded subsets coincides with the weak closure, see [30, Prop. 8.7, Prop. 8.14]. Further, $F + \underline{R}_n$ is weakly lower semicontinuous which determines

$$\Gamma_w\text{-}\liminf(F + R_n) \ge \lim_{n \to +\infty}(F + \underline{R}_n) = F + i_{\mathbf{K}},$$

by (4.6). Therefore, we observe that

$$
\begin{aligned}
F + i_{\mathbf{K}} &\leq \Gamma_w\text{-}\liminf(F + R_n) \\
&\leq \Gamma_w\text{-}\limsup(F + R_n) \\
&\leq \Gamma\text{-}\limsup(F + R_n) \\
&\leq F + i_{\overline{\mathbf{K} \cap Y}},
\end{aligned}
$$

such that $\Gamma\text{-}\lim(F + R_n) = \Gamma_w\text{-}\lim(F + R_n) = F + i_{\mathbf{K}}$ if (4.2) holds true. $\square$

In what follows, we provide a selection of approximation/regularization methods which fit into the general class of perturbations given by (4.3) and which are used very frequently in practice.

EXAMPLE 4.1 (Tikhonov-Regularization). *Let $(Y, \|\cdot\|_Y)$ be a Banach space, and suppose that Y is densely and continuously embedded into V. For a sequence of positive non-decreasing numbers $\{\gamma_n\}$ with $\gamma_n \to +\infty$ and fixed $\alpha > 0$, consider in (4.3) the Tikhonov regularization*

$$
R_n(u) = i_{\mathbf{K}}(u) + \tfrac{1}{2\gamma_n}\|u\|_Y^\alpha. \tag{4.7}
$$

We assume that $R_n(u) = +\infty$ if $u \notin Y$. Then, set

$$
\underline{R}_n := i_{\mathbf{K}}, \qquad and \qquad \overline{R}_n := R_n,
$$

for all $n \in \mathbb{N}$, and (4.4) and (4.5) are satisfied so that R_n is in the context of (4.3).

EXAMPLE 4.2 (Conformal discretization). *Let V be a separable Banach space. Assume that (4.1) is approximated by a Galerkin approach using nested and conformal finite-dimensional subspaces V_n, i.e., we have $V_n \subset V$ and $V_n \subset V_{n+1}$ for all $n \in \mathbb{N}$ with the Galerkin approximation property:*

$$
\overline{\bigcup_{n \in \mathbb{N}} V_n}^V = V.
$$

Therefore, problem (4.1) is replaced by (4.3) by the discretized counterpart defined by $R_n(u) = i_{\mathbf{K} \cap V_n}$. In this setting, define

$$
\underline{R}_n := i_{\mathbf{K}}, \qquad and \qquad \overline{R}_n = R_n.
$$

It follows that (4.4) is satisfied, and if $Y := \bigcup_{n \in \mathbb{N}} V_n$, then (4.5) is fulfilled as well.

EXAMPLE 4.3 (Combined Moreau-Yosida-Tikhonov-Regularization). *Let V be a Hilbert space and $(Y, \|\cdot\|_Y)$ a Banach space with Y densely and continuously embedded into V. For two sequences of positive non-decreasing numbers $\{\gamma_n\}, \{\gamma_n'\}$ with $\gamma_n, \gamma_n' \to +\infty$ and fixed $\alpha > 0$, consider the simultaneous Moreau-Yosida and Tikhonov regularization:*

$$
R_n(u) = \tfrac{\gamma_n}{2} \inf_{v \in \mathbf{K}} \|u - v\|^2 + \tfrac{1}{2\gamma_n'}\|u\|_Y^\alpha, \tag{4.8}
$$

with $\alpha > 0$ fixed. We assume that $R_n(u) = +\infty$ if $u \notin Y$, and define

$$
\underline{R}_n(u) = \tfrac{\gamma_n}{2} \inf_{v \in \mathbf{K}} \|u - v\|^2 \qquad and \qquad \overline{R}_n(u) = i_{\mathbf{K}}(u) + \tfrac{1}{2\gamma_n}\|u\|_Y^\alpha.
$$

It is well-known from the theory of Moreau-Yosida regularizations that $\underline{R}_n$ satisfies (4.4), and (4.5) it is also directly verified.

EXAMPLE 4.4 (Conformal discretization and Moreau-Yosida regularization). *Let V be a separable Hilbert space and $\{\gamma_n\}$ a sequence of positive non-decreasing numbers with $\gamma_n \to +\infty$. The simultaneous regularization and discretization leads to*

$$R_n(u) = \tfrac{\gamma_n}{2} \inf_{v \in \mathbf{K}} \|u - v\|^2 + i_{V_n}(u), \tag{4.9}$$

where the sequence of spaces $\{V_n\}$ is defined as in the previous examples. Defining

$$\underline{R}_n = \tfrac{\gamma_n}{2} \inf_{v \in \mathbf{K}} \|u - v\|^2 \qquad and \qquad \overline{R}_n = i_{\mathbf{K} \cap V_n},$$

(4.4) and (4.5) are fulfilled with $Y = \bigcup_{n \in \mathbb{N}} V_n$ and the framework of (4.3) applies.

From Theorem 4.2, the perturbations defined in the above examples are stable with respect to (4.1) provided the density result (4.2) holds. Moreover, the density property (4.2) is also a necessary condition for the stability of perturbation schemes in the following sense: Firstly, note that the Γ-limit of the approximation schemes defined in Example 4.1 and Example 4.2 can be calculated using similar arguments as in the proof of Theorem 4.2. Under the same conditions on V, namely that V is reflexive or with separable dual, one infers that $F + i_{\overline{\mathbf{K} \cap Y}}$ is the weak and strong Γ-limit in both examples. Secondly, in the approaches of Example 4.3 and Example 4.4, Theorem 4.2 guarantees that $F + i_{\mathbf{K}}$ is obtained as the weak-strong Γ-limit for *any* coupling of regularization (parameter) pairs $[\gamma_n, \gamma'_n]$ and $[V_n, \gamma_n]$, respectively. Further, in the combined Galerkin-Moreau-Yosida approach (Example 4.4), it is possible to prove the existence of a combination of n and γ_n to recover $F + i_{\mathbf{K}}$ in the Γ-limit without resorting to the density property (4.2), see [62, Prop. 2.46]. However, the proof is non-constructive! Hence, it is not applicable for the design of solvers. Moreover, if (4.2) is violated, one may construct for any $w \in \mathbf{K} \setminus \overline{\mathbf{K} \cap Y}$ a sequence γ_n such that no recovery sequence exists for the element w. The analogous statement is valid for the case of combined Moreau-Yosida-Tikhonov regularizations. Let us now rigorously establish the preceding statements.

THEOREM 4.3. *Let the assumptions of* Example 4.4 *be satisfied. Further suppose that*

$$\overline{\mathbf{K} \cap Y} \subsetneq \mathbf{K}.$$

Then for each $w \in \mathbf{K} \setminus \overline{\mathbf{K} \cap Y}$ there exists a strictly increasing sequence of numbers $\{\gamma_n\}$ with $\gamma_n \to \infty$ such that there exists no strong recovery sequence at w, i.e.,

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

for all $\{y_n\}$ in V with $y_n \to w$, where R_n is given by (4.9).

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

(a) *We first prove the following result:*

$$\forall n \in \mathbb{N}, \exists \gamma_n > 0 : \left[y \in Y \wedge \operatorname{dist}(y, \mathbf{K} \cap \overline{B_\rho(w)})^2 < \tfrac{1}{\gamma_n} \implies y \notin V_n \right]. \tag{4.10}$$

Assume the opposite, i.e.,

$$\exists n_0 \in \mathbb{N} \quad : \quad \left[\forall n \in \mathbb{N}, \exists w_n \in V_{n_0}, v_n \in (\mathbf{K} \cap \overline{B_\rho(w)}) \quad : \quad \|w_n - v_n\|^2 \leq \tfrac{1}{n} \right].$$

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

(b) *Non-existence of a strong recovery sequence:* Choose γ_n according to (4.10) and suppose there exists a recovery sequence y_n to w, i.e., $y_n \to w$ and $F(y_n) + \frac{\gamma_n}{2} \mathrm{dist}(y_n, \mathbf{K})^2 + i_{V_n}(y_n) \to F(w)$. The continuity of F implies that $y_n \in V_n$ and $\frac{\gamma_n}{2} \mathrm{dist}(y_n, \mathbf{K})^2 \to 0$. Consequently, using $y_n \to w$ and $w \in \mathbf{K}$, there exists $n_1 \in \mathbb{N}$ such that

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

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

4.1. Applications to Finite Element approximations. We now concentrate efforts in how the previously described ideas permeate through their finite dimensional approximation. In this section we assume that $\Omega \subset \mathbb{R}^d$ is a Lipschitz polyhedral domain. We start with a small generalization of item (I) in Definition 2.1 for the finite dimensional case. In some textbooks on finite-dimensional approximations of variational inequalities, cf., e.g. Glowinski [38], Han and Reddy [39], condition (I) is commonly replaced by the following criterion:

(i) There exists a dense subset $\tilde{\mathbf{K}} \subset \mathbf{K}$ and an operator $r_n : \tilde{\mathbf{K}} \to V$, such that for all $v \in \tilde{\mathbf{K}}$ it holds $r_n v \to v$ in V and there exists $n_0 \in \mathbb{N}$ such that $r_n v \in \mathbf{K}_n$ for all $n \geq n_0$.

It is easy to show that (i) implies (I) in Definition 2.1. In fact, let $v \in \mathbf{K}$ and denote by $\pi_{\mathbf{K}_n} v$ its (not necessarily uniquely determined) projection onto $\mathbf{K}_n$. By density, for $\epsilon > 0$, there exists $v^\epsilon \in \tilde{\mathbf{K}}$ such that $\|v^\epsilon - v\| \leq \epsilon$. Thus, we have

$$\|v - \pi_{\mathbf{K}_n} v\| = \inf_{v^n \in \mathbf{K}_n} \|v - v^n\| \leq \|v - r_n v^\epsilon\| \leq \epsilon + \|v^\epsilon - r_n v^\epsilon\|$$

for sufficiently large n such that $\lim_{n \to \infty} \|v - \pi_{\mathbf{K}_n} v\| \leq \epsilon$ where ϵ was arbitrary.

Condition (i) is more convenient in the context of finite-dimensional approximations, where r_n is given by interpolation operators that are only defined on a dense subset of V. Thus, giving rise to sets $\tilde{\mathbf{K}}$ of the type $\mathbf{K} \cap C_c^\infty(\Omega)$. in fact, this is precisely where the density results of the previous sections are required. For practical relevance, we consider the perturbation of variational inequalities.

In what follows, the sequence of approximating sets is assumed to be originating from a finite-dimensional approximation $\mathbf{K}_n = \mathbf{K}_{h_n}$ of the set $\mathbf{K}$ in the framework of classical Finite Element methods: The parameter n is associated with a sequence of mesh sizes $\{h_n\}$ converging to zero. Concerning the literature and in the context of approximation of variational inequalities, Falk [34]'s a priori estimate for elliptic variational inequalities shows that it is sufficient to tailor the sets $\mathbf{K}_n$ with respect to the VI solution u: This gives rise to the class of adaptive Finite Elements methods. Rigorous convergence proofs for adaptive discretizations of variational inequalities are restricted to special cases and usually require strong assumptions. See for example, in the case of the obstacle problem with a piecewise affine obstacle, the article Siebert and Veeser [78]. Furthermore, density results may still be useful in the analysis of adaptive schemes utilizing interpolation operators, cf. Siebert [77].

Consider a sequence of geometrically conformal affine simplicial meshes $\{\mathcal{T}_h\}_{h>0}$ of Ω of mesh size h, i.e.,

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

where $\operatorname{diam}(T) = \max_{x,y \in T} |x - y|$ denotes the diameter of T. We call $\mathcal{T}_h$, a triangulation of Ω. The Lebesgue measure of an element $T \in \mathcal{T}_h$ is denoted by $\lambda(T)$. We further assume that the sequence $\{\mathcal{T}_h\}$ is shape-regular, that is

$$\exists c > 0 : \quad \frac{\operatorname{diam}(T)}{\rho_T} \leq c \quad \forall h, \quad \forall T \in \mathcal{T}_h, \tag{4.11}$$

where ρ_T is the diameter of the largest ball that is contained in T. Additionally, we 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. Abusing notation, we write $|\mathcal{M}_h|$ and $|\mathcal{N}_h|$ for the cardinality of the respective sets. Let $\chi_T : \Omega \to \mathbb{R}$ be the characteristic function of T:

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

The standard $H^1(\Omega)$-conformal Finite Element space of globally continuous piecewise affine functions associated to $\mathcal{T}_h$ is given by

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

Here, $\mathbb{P}_1$ denotes the space of polynomials of degree less than or equal to one. Associated to H_h and its standard nodal basis $\{\phi_x : x \in \mathcal{N}_h\}$, we define the global interpolation operator

$$I_h : C(\overline{\Omega}) \to H_h, \quad I_h u := \sum_{x \in \mathcal{N}_h} u(x)\phi_x. \tag{4.12}$$

Note that I_h is only defined on a dense subspace of $H^1(\Omega)$.

We define the Hilbert space $H(\operatorname{div}, \Omega) := \{v \in L^2(\Omega)^N : \operatorname{div} v \in L^2(\Omega)\}$ endowed with the inner product

$$(v, w)_{H(\operatorname{div})} := (v, w)_{L^2(\Omega)^N} + (\operatorname{div} v, \operatorname{div} w)_{L^2(\Omega)}.$$

The closure of $C_c^\infty(\Omega)^N$ with respect to the $H(\operatorname{div}, \Omega)$-norm is denoted by $H_0(\operatorname{div}, \Omega)$ and in the case Ω has a Lipschitz boundary it is equivalent to

$$H_0(\operatorname{div}, \Omega) = \{v \in H(\operatorname{div}, \Omega) : \quad \gamma v := v \cdot \nu|_{\partial\Omega} = 0\}, \tag{4.13}$$

where ν denotes the outer normal vector. The operator γ can be proven to be continuous from $H(\operatorname{div}, \Omega)$ to $H^{1/2}(\partial\Omega)$. For the discretization of variational problems in $H(\Omega; \operatorname{div})$, it is usual to consider the $H(\Omega; \operatorname{div})$-conforming space of Raviart-Thomas Finite Elements of lowest order:

$$RT_h = \{w \in L^2(\Omega)^d : w|_T \in \mathbb{RT} \ \forall T \in \mathcal{T}_h, \ [w \cdot \nu]|_E = 0 \ \forall E \in \mathcal{E}_h \cap \Omega\}, \tag{4.14}$$

where $\mathbb{RT} = \{w \in \mathbb{P}_1^d : \exists a \in \mathbb{R}^d, b \in \mathbb{R}, \text{ for which } w(x) = a + bx\}$ and ν denotes the unit outer normal to T. The incorporation of zero boundary conditions in the normal direction requires the use of the $H_0(\operatorname{div}, \Omega)$-conforming subspace

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

Suitable edge-based basis functions $\{\phi_E : E \in \mathcal{E}_h\}$ can be found in the literature, cf., for instance, Bahriawati and Carstensen [14]. Finally, the global Raviart-Thomas interpolation operator is given by

$$I_h^{RT} : W^{1,1}(\Omega)^{\mathrm{d}} \to RT_h, \quad I_h^{RT} w := \sum_{E \in \mathcal{E}_h} \left(\int_E w \cdot \nu \, \mathrm{d}\mathcal{H}^{\mathrm{d}-1} \right) \phi_E. \tag{4.15}$$

We are now in shape to present the pertinent Mosco convergence results associated to finite element discretizations.

THEOREM 1. *Suppose that $\alpha \in C(\overline{\Omega})$ satisfies $\inf_{x \in \Omega} \alpha(x) > 0$, and that $N \in \mathbb{N}$ is given. Then the sets*

$$\mathbf{K}_h^1 := \{w \in (H_h)^N : |w(x_T)|_{\ell^p} \leq \alpha(x_T) \text{ for all } T \in \mathcal{T}_h\},$$
$$\mathbf{K}_h^2 := \{w \in (H_h)^N : |w(x)|_{\ell^p} \leq \alpha(x) \text{ for all } x \in \mathcal{N}_h\},$$

for $1 \leq p \leq +\infty$, Mosco-converge for $h \to 0$ to the set

$$\mathbf{K} = \{w \in H^1(\Omega)^N : |w|_{\ell^p} \leq \alpha\}$$

in $H^1(\Omega)$.

Proof. We concentrate on $\mathbf{K}_h^1$ as the proof for $\mathbf{K}_h^2$ follows analogously, and we separate the proof into two steps.

Step 1: We prove first that (II) *in Definition* (2.1) *holds true.* That is, suppose $w_h \in \mathbf{K}_h^1$ and $w_h \rightharpoonup w$ in $H^1(\Omega)$ along a subsequence, then we prove that $w \in \mathbf{K}$. It suffices to show that $i_{\mathbf{K}}(w) = 0$. Furthermore, it holds $i_{\mathbf{K}} = j^*$ where j^* denotes the Fenchel-Legendre conjugate

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

of the map $j : L^2(\Omega) \to \mathbb{R}$, defined as

$$j(v) := \int_\Omega \alpha |v|_{\ell^q} \, \mathrm{d}x,$$

with $1/p + 1/q = 1$, and where we use the duality relation between ℓ^p and ℓ^q norms, i.e.,

$$|v^*|_{\ell^p} = \sup_{v \in \mathbb{R} \setminus \{0\}} v^* \cdot v / |v|_{\ell^q}.$$

From the definition of j^*, we obtain that $i_{\mathbf{K}}(w) = 0$ is equivalent to

$$(w, v) \leq \int_\Omega \alpha |v|_{\ell^q} \quad \forall \, v \in L^2(\Omega). \tag{4.16}$$

Via density, it is enough to prove this result for all $v \in C_c(\Omega)$. Define

$$\alpha_h := \sum_{T \in \mathcal{T}_h} \alpha(x_T) \chi_T, \quad \text{and} \quad v_h := \sum_{T \in \mathcal{T}_h} v(x_T) \chi_T \tag{4.17}$$

the piecewise constant interpolants of α and v, respectively. Since α and v are uniformly continuous, then $\alpha_h \to \alpha$ and $v_h \to v$ in $L^\infty(\Omega)$. By the weak convergence of

w_h, and the strong convergence of α_h and v_h, we have

$$\int_\Omega w_h \cdot v_h \, \mathrm{d}x \xrightarrow{h\downarrow 0} \int_\Omega w \cdot v \, \mathrm{d}x, \tag{4.18}$$

$$\int_\Omega \alpha_h |v_h|_{\ell^q} \, \mathrm{d}x \xrightarrow{h\downarrow 0} \int_\Omega \alpha |v|_{\ell^q} \, \mathrm{d}x. \tag{4.19}$$

Further, by the midpoint quadrature rule, and that $|w_h(x_T)|_{\ell^p} \le \alpha(x_T)$, we observe that

$$\begin{aligned}
\int_\Omega w_h \cdot v_h \, \mathrm{d}x &= \sum_{T \in \mathcal{T}_h} \int_T w_h \cdot v_h \, \mathrm{d}x \\
&= \sum_{T \in \mathcal{T}_h} \lambda(T) \, w_h(x_T) \cdot v_h|_T \, \mathrm{d}x \\
&\le \sum_{T \in \mathcal{T}_h} \lambda(T) \, |w_h(x_T)|_{\ell^p} |v_h|_T|_{\ell^q} \, \mathrm{d}x \\
&\le \sum_{T \in \mathcal{T}_h} \lambda(T) \, \alpha(x_T) \, |v_h|_T|_{\ell^q} \, \mathrm{d}x \\
&= \int_\Omega \alpha_h |v_h|_{\ell^q} \, \mathrm{d}x,
\end{aligned}$$

which by (4.18) and (4.19) proves (4.16).

Step 2: We prove that (I) *in Definition* (2.1) *holds true.* Note that the assumptions on α imply that

$$\overline{\mathbf{K} \cap C_c^\infty(\Omega)^N}^{H^1(\Omega)^N} = \mathbf{K},$$

that is, the set $\mathbf{K} \cap C_c^\infty(\Omega)^N$ is dense, with respect to the $H^1(\Omega)^N$-norm, in $\mathbf{K}$; see Hintermüller and Rautenberg [43]. This further implies that the set

$$\tilde{\mathbf{K}} := \{\phi \in C^\infty(\overline{\Omega})^N : |\phi(x)|_{\ell^p} < \alpha(x) \text{ for all } x \in \Omega\}, \tag{4.20}$$

is also dense in $\mathbf{K}$ w.r.t. the $H^1(\Omega)^N$-norm. For the global interpolation operator I_h defined in (4.12) we have the classical estimate,

$$\|u - I_h u\|_{L^\infty(\Omega)} \le ch^2 \|u\|_{W^{2,\infty}(\Omega)} \quad \forall\, u \in W^{2,\infty}(\Omega). \tag{4.21}$$

Here, c denotes a constant independent of h on account of the shape-regularity of the triangulation (4.11); see Ern and Guermond [33].

We set $r_h : \tilde{\mathbf{K}} \to (H_h)^N$ to be defined by $r_h w = \{I_h w_i\}_{i=1}^N$ and it follows that $r_h w \to w$ as $h \to 0$ in $H^1(\Omega)^N$ for all $w \in \tilde{\mathbf{K}}$. Hence,

$$\| \, |w - r_h w|_{\ell^p} \, \|_{L^\infty(\Omega)} \le \tilde{c} h^2 \|w\|_{W^{2,\infty}(\Omega)^N}, \tag{4.22}$$

for some $\tilde{c} > 0$, which implies

$$|r_h w(x)|_{\ell^p} \le |w(x)|_{\ell^p} + ch^2 \|w\|_{W^{2,\infty}(\Omega)^N} \quad \forall\, x \in \Omega. \tag{4.23}$$

Thus for $w \in \tilde{\mathbf{K}}$, there exists $h_0 = h_0(w)$ such that $r_h w \in \mathbf{K}_h^1$ for all $h \le h_0$ which implies (i). □

The role of density properties can also be seen in the following result involving other kinds of constraints. The proof carries over mutandis mutatis from the previous proof.

THEOREM 2. *Let $1 \leq p \leq +\infty$, and assume that $\alpha \in C(\overline{\Omega})$ satisfies $\inf_{x \in \Omega} \alpha(x) > 0$. Then the set*

$$\mathbf{K}_h^i := \{w \in H_h : |\nabla w|_T|_{\ell^p} \leq \alpha(x_T) \text{ for all } T \in \mathcal{T}_h\},$$

Mosco-converges for $h \to 0$ to the set

$$\mathbf{K}^i = \{w \in H^1(\Omega) : |\nabla w|_{\ell^p} \leq \alpha\},$$

in $H^1(\Omega)$. Further, the sets

$$\mathbf{K}_h^{ii} := \{w \in RT_{0,h} : |w(x_T)|_{\ell^p} \leq \alpha(x_T) \text{ for all } T \in \mathcal{T}_h\},$$
$$\mathbf{K}_h^{iii} := \{w \in RT_{0,h} : |\text{div } w|_T| \leq \alpha(x_T) \text{ for all } T \in \mathcal{T}_h\},$$

Mosco-converge for $h \to 0$ to the sets

$$\mathbf{K}^{ii} = \{w \in H_0(\text{div}, \Omega) : |w|_{\ell^p} \leq \alpha\},$$
$$\mathbf{K}^{iii} = \{w \in H_0(\text{div}, \Omega) : |\text{div } w| \leq \alpha\},$$

in $H_0(\text{div}, \Omega)$. □

5. Quasi-variational inequalities. The structure (2.3) of $\mathbf{K}$ is adapted to VIs, where the convex set $\mathbf{K}$ is part of the fixed data. However, to treat Quasi-Variational Inequalities (QVIs) we need to consider the convex set as unknown a-priori. Therefore, instead of a convex $\mathbf{K}$ we have a map $\mathbf{K} : V \to 2^V$ written as $v \mapsto \mathbf{K}(v)$ with the following difference with respect to (2.3), the function ϕ is contingent upon the state y as well. Indeed, in the unilateral case, there is an operator Φ such that $\Phi(v) : \Omega \to \mathbb{R}$ is measurable function for each $v \in V$, and

$$\mathbf{K}(v) := \{w \in V : w \leq \Phi(v)\}. \tag{5.1}$$

Thus, if $\mathbf{K}(\cdot)$ is as above and $f \in V'$ is given, then

$$\text{Find } y \in \mathbf{K}(y) : \langle Ay - f, v - y \rangle \geq 0, \quad \forall v \in \mathbf{K}(y), \tag{P_{QVI}}$$

is referred to as a QVI. This kind of problems arose initially from the work of Bensoussan and Lions [21, 53] (see also the monographs [20, 22]) on *impulse control problems*, and later found application modeling a wide variety of non-convex and non-smooth phenomena in applied sciences. Specifically, areas including superconductivity (Kunze and Rodrigues [51], Rodrigues and Santos [74, 75], Barrett and Prigozhin [16, 19], Prigozhin [71], Hintermüller and Rautenberg [44, 42, 41, 46]), continuum mechanics (Friedman [36]), growth of sandpiles (Barrett and Prigozhin [17, 18, 19], Prigozhin [69, 70, 72]), and the determination of rivers/lakes networks (Barrett and Prigozhin [18], Prigozhin [70, 72]), among others. For a complete and classical account on QVIs, we refer the reader to the text of Baiocchi and Capelo [15].

5.1. Impulse Control Problems and QVIs. Because this is an application in stochastic control problems, the proper description of the impulse control involves diffusion processes as the state of the system to be controlled, and a complete setting can be found, for instance, in most of the quoted references below.

In the simplest case, impulse control (or control by interventions) refers to a sequential choices of parameters that modify the free evolution of the system, e.g., beginning at time $\theta_0 = 0$ and a state $x_0 \in \mathbb{R}^d$, the state $\{x(t) : t \geq \theta_0\}$ system is allowed to evolve with a running cost given by $f(x(t))e^{-\alpha t}$ (assuming $a_0(x) = \alpha$, constant) until a time θ_1, where the controller intervenes and changes the state and/or evolution of the system, e.g., if the current state is $x(\theta_1)$ then immediately, the state is moved to the state $x(\theta_1) + \xi_1$, $\xi_1 \geq 0$ and the evolution continues with a similar law. For instance, in finance, the state $x(t)$ may represent the inventory at time t and ξ_1 the order placed at time $t = \theta_1$. Iterating this, a control policy $\{(\theta_i, \xi_i) : i \geq 1\}$ is obtained, and the control problem could be properly defined.

In the context described above and subsequently, the dynamic programming is applied to obtain the so-called Hamilton-Jacobi-Bellman equation, which takes the form of a QVI. In particular, if y is the optimal cost, then at any given time the controller has to decide whether to continue the (free) evolution, i.e., following the equation $Ay = f$, or to make an impulse (intervention), which has a cost (and changes y into My). This can be accounted as

$$\text{(a)} \ \ Ay \leq f, \qquad \text{(b)} \ \ y \leq My, \qquad \text{(c)} \ \ (Ay - f)(y - My) = 0,$$

where A is a second order elliptic operator with Lipschitz continuous and bounded coefficients in a smooth domain Ω of $\mathbb{R}^d$, i.e.,

$$Ay = -\sum_{i,j=1}^{d} a_{ij}(x)\partial_{ij}y(x) + \sum_{i=1}^{d} a_i(x)\partial_i y(x) + a_0(x)y(x),$$

where (a_{ij}) and (a_i) are related with the diffusion and drift terms and the operator M takes the form

$$My(x) = \inf \{y(x + \xi) + k(\xi) : \xi \geq 0\},$$

for a suitable function $k(\xi) \geq k_0 > 0$ representing the cost-per-impulse. Usually, there may be more that one solution of these inequalities, even the *complementary* condition (c) is not enough to ensure uniqueness in a general setting. Moreover, adding those other conditions, a minimum (minimal or maximal, depending on the setting) solution satisfying (a) and (b) is found. In variational form, this is equivalent to $(\mathrm{P_{QVI}})$, and the perturbation of extremal solutions thereof is treated on §5.3.

The expression of the operator M can be modified to deal with more complex settings, e.g., if a fixed time delay τ is imposed (i.e., $\theta_{i+1} \geq \theta_i + \tau$) then

$$My(x) = \inf \{\mathbb{E}_x\{y(x(\tau) + \xi) + k(\xi)\} : \xi \geq 0\},$$

where $\mathbb{E}_x$ is the expectation given $x(0) = x$. In general, the region $\{\xi \geq 0\}$ may be replaced by a subset $\Gamma(x) \subset \Omega$ depending on the given x. Moreover, the whole state space Ω can be divided into three regions, where (1) impulses are not allowed, (2) impulses are allowed, and (3) impulses are required; e.g., (3) is a piece Γ_0 of the boundary of Ω, (2) is the interior of Ω and (1) is the complement of Γ_0 (or empty). In this case, the expression of M changes considerable, but some of the essential properties (e.g., like its monotone character) are retained. This last example is included in the so-called *hybrid models*, where discrete and continuous type variables are used, e.g., examples of this situation can be found in Bensoussan and Menaldi [23, 24], as well as particular cases in more recent papers Menaldi and Robin [58, 61],

among others. For degenerate problems the reader may check [57], and applications to Navier-Stokes are considered by Menaldi and Sritharan [60].

A vector form goes under the name of *switching control*, the coefficients of the diffusion depend on a parameter $i = 1, 2, \ldots, \bar{n}$, i.e., the operator A becomes A_i and a simple expression for M takes the form

$$Mv(x, i) = \inf \{v(x, j) + k(i, j)\} : j \neq i\}, \quad k(i, j) \geq 0,$$

which can be combined with previous forms of M. There is a vast literature on these problems, as recent books, the reader may consult Arapostathis et al. [5], Yin and Zhu [80], among others; and for instance, a relative complex situations is discussed in Menaldi and Robin [59].

5.2. Elementary Existence Theory. For the study of existence of solutions, we define the map

$$T(v) := S(f, \mathbf{K}(v)), \tag{5.2}$$

where S is the solution map associated to the variational inequality (2.1), relative to f and $\mathbf{K}(v)$. Thus, solutions to $(\mathrm{P_{QVI}})$ are equivalently defined as fixed points of the map T, i.e., v solves $(\mathrm{P_{QVI}})$ iff

$$T(v) = v.$$

A direct approach to determine existence of fixed points is the following. The coercivity of the operator A implies that $T(V) \subset B_R(0; V)$ for some $R > 0$. Hence, any sequence $\{v_n\}$ in $B_R(0; V)$ contains a subsequence such that $v_n \rightharpoonup v$ and $T(v_n) \rightharpoonup z$ in V for some v and z. Hence, provided that $\mathbf{K}(v_n) \xrightarrow{\mathrm{M}} \mathbf{K}(v)$ then $T(v_n) \to T(v)$ in V, i.e., the map T is compact and a fixed point exists due to the theorem of Schauder. In summary, a sufficient condition for the existence of solutions to $(\mathrm{P_{QVI}})$ is that $v_n \rightharpoonup v$ in V implies that $\mathbf{K}(v_n) \xrightarrow{\mathrm{M}} \mathbf{K}(v)$.

While the above is suitable to understand the problem of existence, it is not enough to understand the behavior of the set of all solutions $\mathbf{Q}(f)$ to $(\mathrm{P_{QVI}})$ with respect to perturbations of f. For this, we consider an ordering approach.

5.3. Exploiting order and cone structure. We consider an approach based on order that was pioneered by Tartar; see [79], [11, Chapter 15, §15.2], and we follow closely a simplified version of [2, 1]. In particular, we focus on existence and stability properties of the solution set.

Let (V, H, V') be a Gelfand triple of Hilbert spaces, that is, we have $V \hookrightarrow H \hookrightarrow V'$, where the embedding $V \hookrightarrow H$ is dense and continuous, and H is identified with its topological dual H' so that the embedding $H \hookrightarrow V'$ is also dense and continuous. Within this section, $(\cdot, \cdot)$ denotes the inner product in H.

We assume that $H^+ \subset H$ is a convex cone satisfying

$$H^+ = \{v \in H : (v, y) \geq 0 \text{ for all } y \in H^+\}.$$

Note that H^+ defines the cone of non-negative elements inducing the vector ordering:

$$x \leq y \quad \text{if and only if} \quad y - x \in H^+.$$

Given $x \in H$, let x^+ denote the orthogonal projection of x onto H^+, and define $x^- := x^+ - x$. Clearly, one has the decomposition $x = x^+ - x^- \in H^+ - H^+$ for

every $x \in H$, and $(x^+, x^-) = 0$. Further, the infimum and supremum of two elements $x, y \in H$ are defined as $\sup(x, y) := x + (y - x)^+$ and $\inf(x, y) := x - (x - y)^+$, respectively. The supremum of an arbitrary completely ordered subset R of H that is bounded (in the order) above is also properly defined: R can be written as $\{x_i\}_{i \in J}$, where J is completely ordered, and it follows that $\{x_i\}_{i \in J}$ is a generalized Cauchy sequence in H (e.g., see Aubin [11, Chapter 15, §15.2, Proposition 1]); its limit is the upper bound of the original set. Additionally, we have that that norm convergence preserves order, i.e., if $z_n \to z$ and $y_n \to y$ in H, then $z_n \leq y_n$ $(y_n - z_n \in H^+)$ implies $z \leq y$, since H^+ is closed.

We further assume that

$$y \in V \Rightarrow y^+ \in V \qquad \text{and} \qquad \exists \mu > 0 : \|y^+\|_V \leq \mu \|y\|_V, \ \forall y \in V.$$

Then the order in H induces one in V', as well. In fact, for $f, g \in V'$, we write $f \leq g$ if $\langle f, \phi \rangle \leq \langle g, \phi \rangle$ for all $\phi \in V^+ := V \cap H^+$.

Finally, V and H are assumed to be spaces of maps $h : \Omega \to \mathbb{R}$ over some open set $\Omega \subset \mathbb{R}^N$ with the following dense and continuous embedding: $L^\infty(\Omega) \hookrightarrow H$ such that $L^\infty(\Omega) \hookrightarrow V'$, as well. Additionally, we assume that $H \hookrightarrow L^1(\Omega)$.

A common example of Gelfand triple (V, H, V') and cone H^+ that satisfies all conditions is given by $(V, H, V') = (H_0^1(\Omega), L^2(\Omega), H^{-1}(\Omega))$ with $H^+ = L^2(\Omega)^+$, the set of almost everywhere (a.e.) non-negative functions, and $v \leq w$ in the a.e. sense.

5.3.1. Minimal and Maximal Solutions. We start this section with the definition of an increasing map, and existence of fixed points thereof under rather weak conditions. Subsequently, we provide conditions for the map T to be increasing.

DEFINITION 5.1. *A map $R \colon H \to H$ is said to be increasing if for $y, z \in V$ we have that*

$$y \leq z \quad implies \quad R(y) \leq R(z).$$

A general result concerning existence of fixed points for increasing maps is available as we see next (its proof can be found on [11]). This provides a fundamental tool to prove existence of solutions to problem (P_{QVI}) under very weak assumptions.

THEOREM 3 (BIRKHOFF-TARTAR). *Suppose $R : H \to H$ is an increasing map and let $\underline{y}$ be a sub-solution and $\overline{y}$ be a super-solution of the map R, that is:*

$$\underline{y} \leq R(\underline{y}) \quad and \quad R(\overline{y}) \leq \overline{y}.$$

If $\underline{y} \leq \overline{y}$, then the set of fixed points of the map R in the interval $[\underline{y}, \overline{y}]$ is non-empty and has a smallest and a largest element. □

The above theorem mainly states that if a map is increasing, has a subsolution y_1 and a supersolution y_2, then it has a fixed point between (with respect to the order induced in H) y_1 and y_2. Moreover, there are minimal and maximal fixed points in $[y_1, y_2]$.

For the map $T \colon H \to H$ defined as $T(v) = S(f, \mathbf{K}(v))$ to be increasing, some assumptions are required on the structure of $\mathbf{K}$ and on the operator A. For this purpose, in addition to $A : V \to V'$ satisfying (2.2) (i.e., A is linear, continuous, and strongly monotone), we assume it is strictly T-monotone, i.e.,

$$\langle Ay^-, y^+ \rangle \leq 0, \qquad \forall y \in V. \tag{5.3}$$

Further, we assume that

$$\Phi \colon H \to H \cap [\nu, +\infty) \quad \text{is increasing,} \tag{5.4}$$

for some $\nu > 0$, and that $0 \le f \le f_{\max}$ for some $f_{\max} \in V'$. Then, it follows that

$$\underline{y} = A^{-1}0 = 0 \quad \text{and} \quad \overline{y} = A^{-1}f_{\max}$$

are sub- and supersolutions, respectively, of T, and all assumptions of the previous theorem are satisfied: In fact, we have that

$$(f, v) \mapsto S(f, \mathbf{K}(v)) \quad \text{is increasing,}$$

see Rodrigues [73, Section 4:5, Theorem 5.1]. Hence, defining $\mathbf{A}_{\mathrm{ad}} = \{g \in V' : 0 \le g \le f_{\max}\}$, we have the operators

$$\mathsf{m} \colon \mathbf{A}_{\mathrm{ad}} \to V \quad \text{and} \quad \mathsf{M} \colon \mathbf{A}_{\mathrm{ad}} \to V$$

that take elements of $\mathbf{A}_{\mathrm{ad}}$ to minimal and maximal solutions to $(\mathrm{P}_{\mathrm{QVI}})$ in the interval $[\underline{y}, \overline{y}] = [0, A^{-1}f_{\max}]$.

5.3.2. A class of QVIs . Consider the following class of compliant obstacle problems where the obstacle is given implicitly by solving a PDE, thus coupling a VI and a PDE. It consists in finding $(y, \Phi, z) \in V \times H \times W$ such that

$$y \le \Phi, \quad \langle A(y) - f, y - v \rangle \le 0, \qquad \forall v \in V : v \le \Phi, \tag{5.5}$$
$$\langle Bz + G(\Phi, y) - g, w \rangle = 0 \qquad \forall w \in W, \tag{5.6}$$
$$\Phi = Lz, \qquad \text{in } H. \tag{5.7}$$

Here, $V \hookrightarrow W \hookrightarrow H \hookrightarrow W' \hookrightarrow V'$, $f, g \in H^+$, $G : H \times H \to H$ is continuous and bounded, i.e., for some $M_G > 0$, $\|G(\Phi, y)\|_H \le M_G(\|\Phi\|_H + \|y\|_H)$, for all $(\Phi, y) \in H \times H$. Further, $L \colon W \to H$ is an increasing affine linear continuous map with $L(0) \ge \nu > 0$. Additionally, $B \in \mathcal{L}(W, W')$ is coercive and satisfies $\langle Bz^-, z^+ \rangle = 0$ for all $z \in W$ (i.e., B is T-monotone).

Under mild conditions, the above problem can be cast into the form of $(\mathrm{P}_{\mathrm{QVI}})$ as follows. Let $v \in H$, and consider the problem of finding $z \in W$ such that

$$\langle Bz + G(\phi, v) - g, w \rangle = 0, \quad \forall w \in W, \tag{5.8}$$
$$\phi = Lz, \quad \text{in } H. \tag{5.9}$$

Assuming that for each $v \in H$, $z \mapsto G(Lz, v)$ is monotone, one can show the existence of a unique solution $z(v) \in W$ of (5.8)–(5.9). Now set $\Phi(v) := \phi$. Suppose additionally that $(G(Lz, y), z^-) \le 0$ for all $z \in W$ and $y \in H^+$. Hence, $z(v) \ge 0$ and $\Phi(v) = Lz(v) \ge \nu$ for all $v \in H$. In addition, if $v_1 \le v_2$ implies

$$(G(Lv, v_1) - G(Lw, v_2), (v - w)^+) \ge 0,$$

for all w, v, then $z(v_1) \le z(v_2)$ and $\Phi(v_1) \le \Phi(v_2)$, as L is increasing. This finally shows that (5.5)–(5.7) has the form $(\mathrm{P}_{\mathrm{QVI}})$ with Φ as an increasing operator and $\mathbf{K}(v)$ given as $\mathbf{K}(v) := \{v' \in V : v' \le \Phi(v)\}$, where $\Phi(v) = \phi \in W$, and the pair (z, ϕ) is as given by (5.8)–(5.9).

Finally, assuming that $f \in \mathbf{A}_{\mathrm{ad}} = \{h \in V' : 0 \le h \le f_{\max}\}$, we have that the operators $\mathsf{m} \colon \mathbf{A}_{\mathrm{ad}} \to V$ and $\mathsf{M} \colon \mathbf{A}_{\mathrm{ad}} \to V$ are well-defined: They map elements in $\mathbf{A}_{\mathrm{ad}}$ to minimal and maximal solutions to $(\mathrm{P}_{\mathrm{QVI}})$ in the interval $[\underline{y}, \overline{y}] = [0, A^{-1}f_{\max}]$.

5.3.3. A useful Mosco convergence result. The obstacle operator Φ arising from (5.5)-(5.7), can be written as $\Phi(v) = C^{-1}(Lv) + \tilde{g}$ where C is a (nonlinear) partial differential operator, and $\tilde{g}$ is some fixed element in V. In particular, this generates the need to consider Mosco convergence results when obstacles have specific structure. In this vein, we consider the following result.

THEOREM 5.2. *Let $\phi_n, \phi \in V$ for $n \in \mathbb{N}$. Suppose that $\phi_n \to \phi$ in H, and*

$$\mathcal{Q}\phi_n \geq 0 \quad in \ V \quad for \ all \ n \in \mathbb{N},$$

for some strongly monotone $\mathcal{Q} \in \mathscr{L}(V, V')$, such that $\langle \mathcal{Q}v^-, v^+ \rangle \leq 0$ for all $v \in V$. Then,

$$\{w \in V : w \leq \phi_n\} \xrightarrow{\ \mathrm{M}\ } \{w \in V : w \leq \phi\}.$$

holds true.

Proof. First note that since $H \hookrightarrow L^1(\Omega)$, then $\phi_n \to \phi$ in H also implies strong convergence in $L^1(\Omega)$. It follows by Proposition 2.2 that (II) in Definition 2.1 holds true. In order to prove (I) in Definition 2.1 we consider the following construction based on singular perturbations.

Let $w \in V$ such that $w \leq \phi$ be arbitrary and let w_n for $n \in \mathbb{N}$ be defined by

$$\langle r_n \mathcal{Q} w_n + w_n, v \rangle = (\tilde{w}_n, v), \text{ for all } v \in V, \tag{5.10}$$

where $r_n := \|\tilde{w}_n - w\|_H$ and $\tilde{w}_n := \min(w, \phi_n)$, and note that $\tilde{w}_n \to w$ in H and $w \in V$. Then, we can prove that $w_n \to w$ in V. Since $\mathcal{Q}$ is linear, bounded, and $\langle \mathcal{Q}v, v \rangle \geq c\|v\|_V^2$ for all $v \in V$, from the definition of w_n we have

$$
\begin{aligned}
r_n c \|w_n - w\|_V^2 + \|w_n - w\|_H^2 &\leq \langle (r_n \mathcal{Q} + I)(w_n - w), w_n - w \rangle \\
&\leq \langle \tilde{w}_n - w, w_n - w \rangle - r_n \langle \mathcal{Q}w, w_n - w \rangle \tag{5.11} \\
&\leq r_n(C_p + \|\mathcal{Q}w\|_{V'})\|w_n - w\|_V,
\end{aligned}
$$

where C_p is the constant for the embedding $V \hookrightarrow H$. This implies that, $\{w_n\}$ is bounded in V, so that $w_n \rightharpoonup w^*$ (along a subsequence) for some $w^* \in V$. By taking the limit in (5.10), it is shown that $w^* = w$ and that $w_n \rightharpoonup w$ in V not only along a subsequence. It further follows that $w_n \to w$ in H, and since from (5.11) we observe

$$r_n c \|w_n - w\|_V^2 + \|w_n - w\|_H^2 \leq r_n(\|w_n - w\|_H + \langle \mathcal{Q}w, w - w_n \rangle), \tag{5.12}$$

we have that $w_n \to w$ in V.

Next we prove that $w_n \leq \phi_n$. Consider $v = (w_n - \phi_n)^+$ and let us subtract $\langle r_n \mathcal{Q}\phi_n + \phi_n, v \rangle$ from both sides of (5.10). Then, we get

$$
\begin{aligned}
r_n &\left\langle \mathcal{Q}(w_n - \phi_n), (w_n - \phi_n)^+ \right\rangle + \|(w_n - \phi_n)^+\|_H^2 = \\
&- r_n \left\langle \mathcal{Q}\phi_n, (w_n - \phi_n)^+ \right\rangle + (\min(w, \phi_n) - \phi_n, (w_n - \phi_n)^+).
\end{aligned}
$$

Note that $\min(w, \phi_n) - \phi_n \leq 0$ and by assumption $\mathcal{Q}\phi_n \geq 0$. Therefore the right hand side is less or equal to zero. Additionally, since $\mathcal{Q}$ is linear, $\langle \mathcal{Q}v^-, v^+ \rangle \leq 0$, and $\langle \mathcal{Q}v, v \rangle \geq c\|v\|_V^2$ for all $v \in V$, we observe that

$$
\begin{aligned}
r_n c \|(w_n - \phi_n)^+\|_V^2 + \|(w_n - \phi_n)^+\|_H^2 &\leq \\
&\leq r_n \left\langle \mathcal{Q}(w_n - \phi_n)^+, (w_n - \phi_n)^+ \right\rangle + \|(w_n - \phi_n)^+\|_H^2 \leq 0.
\end{aligned}
$$

This yields $w_n \leq \phi_n$, i.e., (I) in Definition 2.1 holds true. This completes the proof. □

In view of the previous result, we assume throughout the rest of this section the following continuity assumption on the obstacle map Φ.

ASSUMPTION 1. *If $v_n \rightharpoonup v$ in V, then $\Phi \colon H \to H \cap [\nu, +\infty)$ satisfies one of the following conditions:*

(a) *$\Phi(v_n) \to \Phi(v)$ in $L^\infty(\Omega)$, or $\Phi(v_n) \to \Phi(v)$ in V.*

(a) *$\Phi(v_n) \to \Phi(v)$ in H and if $v \in V \cap H^+$, then $\Phi(v) \in V$ and $\mathcal{Q}\Phi(v) \geq 0$ in V, for some strongly monotone $\mathcal{Q} \in \mathscr{L}(V, V')$, such that $\langle \mathcal{Q}v^-, v^+\rangle \leq 0$ for all $v \in V$.*

Hence, by Proposition 2.3, and Theorem 5.2, we assume that Φ satisfies conditions to guarantee Mosco convergence of the sets $\mathbf{K}(v_n)$, provided that $v_n \rightharpoonup v$ in V, to $\mathbf{K}(v)$.

5.3.4. Perturbation of minimal and maximal solutions. Existence of solutions to the QVI of interest is established if the following property holds

$$v_n \rightharpoonup v \quad \text{implies that} \quad \mathbf{K}(v_n) \xrightarrow{\mathrm{M}} \mathbf{K}(v).$$

However, we are interested in the stability properties of the maps

$$\mathbf{A}_{\mathrm{ad}} \ni f \mapsto \mathsf{m}(f) \qquad \text{and} \qquad \mathbf{A}_{\mathrm{ad}} \ni f \mapsto \mathsf{M}(f),$$

where

$$\mathbf{A}_{\mathrm{ad}} = \{g \in V' : 0 \leq g \leq f_{\max}\},$$

and hence additional assumptions are needed. In what follows, we establish our fundamental result concerning the behavior of the maps $f \mapsto \mathsf{m}(f)$ and $f \mapsto \mathsf{M}(f)$. As in the previous section we assume that $[\underline{y}, \overline{y}] = [0, A^{-1}f_{\max}]$.

THEOREM 4. *Let $\{f_n\} \subset L^\infty(\Omega) \cap \mathbf{A}_{\mathrm{ad}}$ be such that $f_n \geq c$ for some constant $c > 0$, and $\lim_{n \to \infty} f_n = f^*$ in $L^\infty(\Omega)$. Suppose that the upper bound mapping Φ satisfies Assumption 1 (page 21), and that*

$$\lambda\Phi(y) \geq \Phi(\lambda y) \quad \text{for any} \quad \lambda > 1,\ y \in H. \tag{5.13}$$

Then the assertions

(i) *The sequence of minimal solutions satisfy*

$$\mathsf{m}(f_n) \to \mathsf{m}(f^*) \text{ in } H, \qquad \text{and} \qquad \mathsf{m}(f_n) \rightharpoonup \mathsf{m}(f^*) \text{ in } V. \tag{5.14}$$

(ii) *The sequence of maximal solutions satisfy*

$$\mathsf{M}(f_n) \to \mathsf{M}(f^*) \text{ in } H, \qquad \text{and} \qquad \mathsf{M}(f_n) \rightharpoonup \mathsf{M}(f^*) \text{ in } V. \tag{5.15}$$

hold true.

A few words are in order concerning the previous result. Note that if Φ satisfies Assumption 1 (page 21), but not necessarily (5.13). Then, it is possible to prove that $\mathsf{m}(f_n)$ (and $\mathsf{M}(f_n)$) converge to solutions, elements of $\mathbf{Q}(f^*)$, but not necessarily to $\mathsf{m}(f^*)$ (and $\mathsf{M}(f^*)$). Assumption (5.13) provides the stability of extremal points. Structurally speaking, if Φ is a superposition operator, it states that $\Phi(x) \simeq (x^+)^{1/p}$ for some p.

REFERENCES

[1] A. Alphonse, M. Hintermüller, and C. N. Rautenberg. Recent trends and views on elliptic quasi-variational inequalities. *Topics in Applied Analysis and Optimisation*, pages 1–31, 2019.

[2] A. Alphonse, M. Hintermüller, and C. N. Rautenberg. Stability of the solution set of quasi-variational inequalities and optimal control. *arXiv preprint arXiv:1904.06231*, 2019.

[3] H. Antil and C. N. Rautenberg. Fractional elliptic quasi-variational inequalities: theory and numerics. *Interfaces Free Bound.*, 20(1):1–24, 2018.

[4] H. Antil, C. N. Rautenberg, and A. Schikorra. On a fractional version of a murat compactness result and applications. *SIAM Journal on Mathematical Analysis*, 53(3):3158–3187, 2021.

[5] A. Arapostathis, V. Borkar, and M. Ghosh. *Ergodic control of diffusion processes*. Cambridge University Press, Cambridge, 2012.

[6] H. Attouch. Viscosity solutions of minimization problems. *SIAM Journal on Optimization*, 6(3):769–806, 1996.

[7] H. Attouch and R. Cominetti. A dynamical approach to convex minimization coupling approximation with the steepest descent method. *Journal of Differential Equations*, 128(2):519–540, 1996.

[8] H. Attouch and C. Picard. Inéquations variationnelles avec obstacles et espaces fonctionnels en théorie du potentiel. *Applicable Anal.*, 12(4):287–306, 1981.

[9] H. Attouch and C. Picard. *Asymptotic analysis of variational problems with constraints of obstacle type*, volume 7 of *Publications Mathématiques d'Orsay 82 [Mathematical Publications of Orsay 82]*. Université de Paris-Sud, Département de Mathématique, Orsay, 1982.

[10] H. Attouch and C. Picard. Variational inequalities with varying obstacles: the general form of the limit problem. *J. Funct. Anal.*, 50(3):329–386, 1983.

[11] J.-P. Aubin. *Mathematical methods of game and economic theory*. North-Holland, 1979.

[12] J.-P. Aubin and H. Frankowska. *Set-Valued Analysis*. Birkhäuser, 2009.

[13] A. Azevedo and L. Santos. Convergence of convex sets with gradient constraint. *Journal of Convex Analysis*, 11(2):285–301, 2004.

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

[15] C. Baiocchi and A. Capelo. *Variational and Quasivariational Inequalities*. John Wiley & Sons Inc., New York, 1984. Applications to free boundary problems, Translated from the Italian by Lakshmi Jayakar.

[16] J. W. Barrett and L. Prigozhin. A quasi-variational inequality problem in superconductivity. *Math. Models Methods Appl. Sci.*, 20(5):679–706, 2010.

[17] J. W. Barrett and L. Prigozhin. A quasi-variational inequality problem arising in the modeling of growing sandpiles. *ESAIM Math. Model. Numer. Anal.*, 47(4):1133–1165, 2013.

[18] J. W. Barrett and L. Prigozhin. Lakes and rivers in the landscape: a quasi-variational inequality approach. *Interfaces Free Bound.*, 16(2):269–296, 2014.

[19] J. W. Barrett and L. Prigozhin. Sandpiles and superconductors: nonconforming linear finite element approximations for mixed formulations of quasi-variational inequalities. *IMA J. Numer. Anal.*, 35(1):1–38, 2015.

[20] A. Bensoussan. *Stochastic Control by Functional Analysis Methods*. North-Holland Publishing Co., Amsterdam, 1982.

[21] A. Bensoussan and J.-L. Lions. Controle impulsionnel et inéquations quasi-variationnelles d'évolutions. *C. R. Acad. Sci. Paris*, 276:1333–1338, 1974.

[22] A. Bensoussan and J.-L. Lions. *Impulse Control and Quasi-Variational Inequalities*. Gauthier-Villars, 1984.

[23] A. Bensoussan and J. Menaldi. Hybrid control and dynamic programming. *Dynam. Contin. Discrete Impuls. Systems*, 3(4):395–442, 1997.

[24] A. Bensoussan and J. Menaldi. Stochastic hybrid control. *J. Math. Anal. Appl.*, 249(1):261–288, 2000. Special issue in honor of Richard Bellman.

[25] L. Boccardo and F. Murat. Nouveaux résultats de convergence dans des problèmes unilatéraux. In *Nonlinear partial differential equations and their applications. Collège de France Seminar, Vol. II (Paris, 1979/1980)*, volume 60 of *Res. Notes in Math.*, pages 64–85, 387–388. Pitman, Boston, Mass.-London, 1982.

[26] L. Boccardo and F. Murat. Homogenization of nonlinear unilateral problems. In *Composite media and homogenization theory (Trieste, 1990)*, volume 5 of *Progr. Nonlinear Differential Equations Appl.*, pages 81–105. Birkhäuser Boston, Boston, MA, 1991.

[27] M. Bocea, M. Mihăilescu, M. Pérez-Llanos, and J. D. Rossi. Models for growth of heterogeneous sandpiles via mosco convergence. *Asymptotic Analysis*, 78(1-2):11–36, 2012.

[28] H. Brezis. Remarque sur l'article précédent de F. Murat. *J. Math. Pures Appl.*, 60:321–322., 1981.

[29] G. Dal Maso. Some necessary and sufficient conditions for the convergence of sequences of unilateral convex sets. *J. Funct. Anal.*, 62(2):119–159, 1985.

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

[31] E. De Giorgi and T. Franzoni. Su un tipo di convergenza variazionale. *Atti della Accademia Nazionale dei Lincei. Classe di Scienze Fisiche, Matematiche e Naturali. Rendiconti*, 58(6):842–850, 1975.

[32] G. Duvaut and J.-P. Lions. *Les Inéquations en Mécanique et en Physique*. Dunod, Paris, 1972.

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

[34] R. Falk. Error estimates for the approximation of a class of variational inequalities. *Mathematics of Computation*, 28(128):963–971, 1974.

[35] M. Feischl, M. Page, and D. Praetorius. Convergence and quasi-optimality of adaptive fem with inhomogeneous dirichlet data. *Journal of computational and applied mathematics*, 255:481–501, 2014.

[36] A. Friedman. *Variational Principles and Free-Boundary Problems*. Wiley-Interscience, 1982.

[37] T. Fukao and N. Kenmochi. A thermohydraulics model with temperature dependent constraint on velocity fields. *Discrete Contin. Dyn. Syst. Ser. S*, 7(1):17–34, 2014.

[38] R. Glowinski. *Numerical Methods for Nonlinear Variational Problems*. Springer, 1982.

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

[40] P. T. Harker. Generalized Nash games and quasi-variational inequalities. *European Journal of Operational Research*, 54:81–94, 1991.

[41] M. Hintermüller and C. N. Rautenberg. A sequential minimization technique for elliptic quasi-variational inequalities with gradient constraints. *SIAM J. Optim.*, 22(4):1224–1257, 2012.

[42] M. Hintermüller and C. N. Rautenberg. Parabolic quasi-variational inequalities with gradient-type constraints. *SIAM J. Optim.*, 23(4):2090–2123, 2013.

[43] M. Hintermüller and C. N. Rautenberg. On the density of classes of closed convex sets with pointwise constraints in Sobolev spaces. *J. Math. Anal. Appl.*, 426(1):585–593, 2015.

[44] M. Hintermüller and C. N. Rautenberg. On the uniqueness and numerical approximation of solutions to certain parabolic quasi-variational inequalities. *Port. Math.*, 74(1):1–35, 2017.

[45] M. Hintermüller, C. N. Rautenberg, and S. Rösel. Density of convex intersections and applications. *Proc. A.*, 473(2205):20160919, 28, 2017.

[46] M. Hintermüller, C. N. Rautenberg, and N. Strogies. Dissipative and non-dissipative evolutionary quasi-variational inequalities with gradient constraints. *Set-Valued and Variational Analysis*, 27(2):433–468, 2019.

[47] N. Igbida, J. Mazón, J. Rossi, and J. Toledo. A monge-kantorovich mass transport problem for a discrete distance. *J. Funct. Anal*, 260(12):3494–3534, 2011.

[48] N. Kenmochi and U. Stefanelli. Existence for a class of nonlocal quasivariational evolution problems. In *Nonlinear phenomena with energy dissipation*, volume 29 of *GAKUTO Internat. Ser. Math. Sci. Appl.*, pages 253–264. Gakkōtosho, Tokyo, 2008.

[49] D. Kinderlehrer and G. Stampacchia. *An introduction to variational inequalities and their applications*, volume 31 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2000. Reprint of the 1980 original.

[50] A. S. Kravchuk and P. J. Neittaanmäki. *Variational and quasi-variational inequalities in mechanics*, volume 147. Springer Science & Business Media, 2007.

[51] M. Kunze and J. Rodrigues. An elliptic quasi-variational inequality with gradient constraints and some of its applications. *Mathematical Methods in the Applied Sciences*, 23:897–908, 2000.

[52] C. Kuratowski. *Topologie I: Espaces métrisables, espaces complets*. Warszawa, 1948.

[53] J.-L. Lions. Sur le côntrole optimal des systemes distribuées. *Enseigne*, 19:125–166, 1973.

[54] J.-L. Lions. Asymptotic behaviour of solutions of variational inequalitites with highly oscillating coefficients. *Applications of Methods of Functional Analysis to Problems in Mechanics, Proc. Joint Symp. IUTAM/IMU. Lecture Notes in Mathematics, Springer, Berlin*, 503, 1975.

[55] J.-P. Lions and G. Stampacchia. Variational inequalities. *Commun. Pure Appl. Math.*, 20:493–519, 1967.

[56] M. Marcus and V. J. Mizel. Every superposition operator mapping one Sobolev space into another is continuous. *J. Functional Analysis*, 33(2):217–229, 1979.

[57] J. Menaldi. Optimal impulse control problems for degenerate diffusions with jumps. *Acta Appl. Math.*, 8(2):165–198, 1987.

[58] J. Menaldi and M. Robin. On some impulse control problems with constraint. *SIAM J. Control*

Optim., 55(5):3204–3225, 2017.

[59] J. Menaldi and M. Robin. Hybrid models and switching control with constraints. *Communications on Stochastic Analysis (Volume in Honor of Pao-Liu Chow, Eds.: H. Kuo and G. Yin)*, 13(3):1–29, Article 3, 2019. (https://doi.org/10.31390/cosa.13.3.03).

[60] J. Menaldi and S. Sritharan. Impulse control of stochastic Navier-Stokes equations. *Nonlinear Anal.*, 52(2):357–381, 2003.

[61] J. L. Menaldi and M. Robin. On Some Ergodic Impulse Control Problems with Constraint. *SIAM J. Control Optim.*, 56(4):2690–2711, 2018.

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

[63] U. Mosco. Approximation of the solutions of some variational inequalities. *Annali della Scuola Normale Superiore di Pisa-Classe di Scienze*, 21(3):373–394, 1967.

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

[65] F. Murat. L'injection du cône positif de H^{-1} dans $W^{-1,\,q}$ est compacte pour tout $q < 2$. *J. Math. Pures Appl. (9)*, 60(3):309–322, 1981.

[66] A. Neubauer. Tikhonov-regularization of ill-posed linear operator equations on closed convex sets. *Journal of Approximation Theory*, 53(3):304–320, 1988.

[67] A. Neubauer and O. Scherzer. Finite-dimensional approximation of tikhonov regularized solutions of non-linear ill-posed problems. *Numerical functional analysis and optimization*, 11(1-2):85–99, 1990.

[68] J.-S. Pang and M. Fukushima. Quasi-variational inequalities, generalized Nash equilibria, and multi-leader-follower games. *Computational Management Science*, 3:373–375, 2009.

[69] L. Prigozhin. Quasivariational inequality describing the shape of a poured pile. *Zhurnal Vichislitel'noy Matematiki i Matematicheskoy Fiziki*, 7:1072–1080, 1986.

[70] L. Prigozhin. Sandpiles and river networks: extended systems with non-local interactions. *Phys. Rev. E*, 49:1161–1167, 1994.

[71] L. Prigozhin. On the Bean critical-state model in superconductivity. *European Journal of Applied Mathematics*, 7:237–247, 1996.

[72] L. Prigozhin. Sandpiles, river networks, and type-ii superconductors. *Free Boundary Problems News*, 10:2–4, 1996.

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

[74] J. F. Rodrigues and L. Santos. A parabolic quasi-variational inequality arising in a superconductivity model. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)*, 29(1):153–169, 2000.

[75] J. F. Rodrigues and L. Santos. Quasivariational solutions for first order quasilinear equations with gradient constraint. *Arch. Ration. Mech. Anal.*, 205(2):493–514, 2012.

[76] S. Rösel. Approximation of nonsmooth optimization problems and elliptic variational inequalities with applications to elasto-plasticity. 2017.

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

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

[79] L. Tartar. Inéquations quasi variationnelles abstraites. *CR Acad. Sci. Paris Sér. A*, 278:1193–1196, 1974.

[80] G. Yin and C. Zhu. *Hybrid switching diffusions*. Springer, New York, 2010. Properties and applications.