Strong Convergence of a Double Projection-type Method for Monotone Variational Inequalities in Hilbert Spaces

Christian Kanzow* and Yekini Shehu†

February 13, 2017

Abstract

We introduce a projection-type algorithm for solving monotone variational inequality problems in real Hilbert spaces. We prove that the whole sequence of iterates converges strongly to a solution of the variational inequality. The method uses only two projections onto the feasible set in each iteration in contrast to other strongly convergent algorithms which either require plenty of projections within a stepsize rule or have to compute projections on possibly more complicated sets. Some numerical results illustrate the practical behaviour of our method.

1 Introduction

In this paper, we consider the following variational inequality (for short, VI(A, C)): find $x \in C$ such that

$$\langle Ax, y - x\rangle \geq 0, \quad \forall y \in C. \tag{1}$$

Let SOL denote the solution set of VI(A, C). It is well-known that x solves the VI(A, C) if and only if x solves the fixed point equation

$$x = P_C(x - \gamma Ax)$$

or, equivalently, x solves the residual equation

$$r_\gamma(x) = 0, \quad \text{where} \quad r_\gamma(x) := x - P_C(x - \gamma Ax) \tag{2}$$

for an arbitrary positive constant γ, see [16] for details. Therefore, the knowledge of fixed-point algorithms (see, for example, [15, 36]) can be used to solve (1).

*University of Würzburg, Institute of Mathematics, Campus Hubland Nord, Emil-Fischer-Str. 30, 97074 Würzburg, Germany; e-mail: kanzow@mathematik.uni-wuerzburg.de

†University of Nigeria, Department of Mathematics, Nsukka, Nigeria; e-mail: yekini.shehu@unn.edu.ng. Current address (May 2016 – April 2018): University of Würzburg, Institute of Mathematics, Campus Hubland Nord, Emil-Fischer-Str. 30, 97074 Würzburg, Germany.
The research of this author is supported by the Alexander von Humboldt-Foundation.

Variational inequality theory is an important tool in studying a wide class of obstacle, unilateral, and equilibrium problems arising in several branches of pure and applied sciences in a unified and general framework (see, for example, [2, 3, 16, 23, 24]). This dynamic field is experiencing an explosive growth in both theory and applications. Several numerical methods have been developed for solving variational inequalities and related optimization problems, see the monographs [15, 24] and references therein.

The extragradient method, introduced in 1976 by Korpelevich [25], which is given by

$$\begin{cases} x_1 \in C, \\ y_n = P_C(x_n - \gamma A x_n) \\ x_{n+1} = P_C(x_n - \gamma A y_n), \quad n \geq 1, \end{cases} \tag{3}$$

where $\gamma \in (0, \frac{1}{L})$ for a finite-dimensional space, provides an iterative process converging to a solution of $\mathrm{VI}(A, C)$ by only assuming that $C \subseteq \mathbb{R}^n$ is nonempty, closed, and convex, and $A : C \to \mathbb{R}^n$ is monotone and L-Lipschitz continuous. The extragradient method was further extended to infinite-dimensional spaces by many authors; see, for instance, [1, 10, 12, 18, 20, 21, 35, 38, 40]. In the setting of Hilbert spaces, this method is only known to be weakly convergent. Note that the extragradient method needs two projections onto the set C and two evaluations of A per iteration.

A crucial feature regarding the design of numerical methods related to the extragradient method is to minimize the number of evaluations of P_C per iteration. So the extragradient method needs to be improved in situations, where a projection onto C is hard to evaluate and therefore computationally expensive. An attempt in this direction was initiated by Censor et al. [13], who modified Korpelevich's method (3) by replacing the second projection onto the closed and convex subset C with the one onto a subgradient half-space. Their method, which therefore uses only one projection onto C, is called the subgradient extragradient method. This subgradient extragradient method is shown to be weakly convergent to a solution of the variational inequality $\mathrm{VI}(A, C)$. Using only a single projection onto C, Maingé and Gobinddass [30] (see also Maingé [29]) also obtained weak convergence results for solving the $\mathrm{VI}(A, C)$ in real Hilbert spaces with a monotone and Lipschitz continuous mapping A, by means of a projected reflected gradient-type method [32] and inertial terms. In fact, in certain situations it is also possible to get rid of any projections onto C by replacing this projection by a (finite) sequence of projections onto suitable halfspaces for which explicit formulas exist, cf. Bello Cruz and Iusem [6]. Several alternatives to the extragradient method or its modifications have also been proposed in the literature by several authors, see, for example, [28, 33, 37, 40].

Hence, the situation is very comfortable if one aims to obtain a weakly convergent projection-type method for (monotone) variational inequalities. Unfortunately, in an infinite-dimensional setting, weak convergence of an iterative scheme is an unsatisfactory property. Typically, one is looking for an algorithm that generates a strongly convergent sequence. Unfortunately, strongly convergent projection-type methods for (monotone) variational inequalities are still rare and usually require stronger assumptions and a higher computational overhead per iteration as their weakly convergent counterparts.

An early attempt is due to Noor [26], where four projections onto C and a couple of function evaluations are needed per iteration in order to obtain a strongly convergent iteration sequence. To the best of our knowledge, Censor et al. [12] were the first to prove strong convergence for monotone and Lipschitz continuous mappings A using essentially only one projection onto C. Under the same assumptions, Kraikaew and Saejung [27] obtained a strong convergence result using a combination of a Halpern-type iterative scheme and the subgradient extragradient method. Recently, also Malitsky and Semenov [34] prove strong convergence of a suitable projection-type method using the method of Haugazeau when A is Lipschitz continuous and monotone. Similar to [12], also [27, 34] require mainly a single projection onto the feasible set C at each iteration.

All these papers assume A to be Lipschitz continuous, and the Lipschitz constant is typically assumed to be known and defines, at least implicitly, a suitable stepsize within the algorithm. Apart from the fact that many operators A are not Lipschitz continuous, it is usually unrealistic to have a good estimate of such a Lipschitz constant; moreover, the stepsize defined by this Lipschitz constant is often very small and deteriorates the convergence rate. In practice, larger stepsizes can often be used and yield better numerical results.

The only paper which we are currently aware of and which proves strong convergence of a projection-type method for monotone variational inequalities without assuming A to be Lipschitzian is due to Bello Cruz and Iusem [5]. They include a stepsize procedure and require two projections per iteration, one onto C and the second one onto a more complicated set which also changes from iteration to iteration. The latter is undesirable for problems where projections onto C itself can be computed in a relatively efficient way.

Our aim in this paper is therefore to prove strong convergence of a double projection method for monotone variational inequality problems in a real Hilbert space which, at each iteration, needs only two projections onto C itself. It involves a stepsize rule which might need some evaluations of A in the inner iteration, but no additional projections, in contrast to some other exsiting stepsize rules like those from [14, 22, 40].

Note that, in the finite-dimensional case, many of the extragradient-like schemes actually work for the larger class of pseudomonotone mappings A. The technique of proof is usually almost identical as the one for monotone problems. In the infinite-dimensional setting, however, it seems to be much more difficult to generalize existing methods to pseudomonotone mappings. Two recent contributions in this direction can be found in [11, 41], but the authors require additional and very strong assumptions regarding the operator A. We comment on this in some more detail at the end of our convergence analysis.

The paper is organized as follows: We first recall some basic definitions and results in Section 2. Some discussions about our projection-type method used in this paper are given in Section 3. The strong convergence of our double projection algorithm is then investigated in Section 4. Some numerical experiments can be found in Section 5. We conclude with some final remarks in Section 6.

2 Preliminaries

This section contains some definitions and basic results that will be used in our subsequent analysis. The letter H always denotes a real Hilbert space.

We first state the formal definition of some classes of functions that play an essential role in our analysis.

Definition 2.1. *Let $X \subseteq H$ be a nonempty subset. Then a mapping $A : X \to H$ is called*

(a) monotone *on X if $\langle Ax - Ay, x - y\rangle \geq 0$ for all $x, y \in X$;*

(b) pseudomonotone *on X if, for all $x, y \in X$, $\langle Ax, y - x\rangle \geq 0 \Rightarrow \langle Ay, y - x\rangle \geq 0$;*

(c) Lipschitz continuous *on X if there exists a constant $L > 0$ such that $\|Ax - Ay\| \leq L\|x - y\|$ for all $x, y \in X$.*

We next recall some properties of the projection, cf. [4] for more details. To this end, let $C \subseteq H$ be a nonempty, closed, and convex subset of a real Hilbert space H. For any point $u \in H$, there exists a unique point $P_C u \in C$ such that

$$\|u - P_C u\| \leq \|u - y\| \quad \forall y \in C.$$

P_C is called the *metric projection* of H onto C. We know that P_C is a nonexpansive mapping of H onto C. It is also known that P_C satisfies

$$\langle x - y, P_C x - P_C y\rangle \geq \|P_C x - P_C y\|^2 \quad \forall x, y \in H. \tag{4}$$

Furthermore, $P_C x$ is characterized by the properties

$$P_C x \in C \quad \text{and} \quad \langle x - P_C x, P_C x - y\rangle \geq 0 \quad \forall y \in C. \tag{5}$$

This characterization implies that

$$\|x - y\|^2 \geq \|x - P_C x\|^2 + \|y - P_C x\|^2 \quad \forall x \in H, \forall y \in C. \tag{6}$$

The following elementary lemma will be used in our convergence analysis.

Lemma 2.2. *The following statements hold in any real Hilbert space H:*

(a) $\|x + y\|^2 = \|x\|^2 + 2\langle x, y\rangle + \|y\|^2$ for all $x, y \in H$;

(b) $2\langle x - y, x - z\rangle = \|x - y\|^2 + \|x - z\|^2 - \|y - z\|^2$ for all $x, y, z \in H$.

We next recall some existing results from the literature to facilitate our proof of strong convergence. The first of these results is taken from [31, Lem. 4.3]. Note that the sequence $\{r_n\}$ occuring in this result is assumed to be bounded in the original reference [31], but that the proof goes through under the weaker assumption where the real sequence $\{r_n\}$ is only bounded from above. For the sake of completeness, we provide a complete proof of this slightly stronger result, since this variant will actually be exploited in our main convergence theorem.

Lemma 2.3. *Let $\{a_n\}$ be a sequence of non-negative numbers such that*

$$a_{n+1} \le (1-\gamma_n)a_n + \gamma_n r_n, \tag{7}$$

where $\{r_n\}$ is a sequence of real numbers bounded from above and $\{\gamma_n\} \subset [0,1]$ satisfies $\sum \gamma_n = \infty$. Then it holds that $\limsup\limits_{n\to\infty} a_n \le \limsup\limits_{n\to\infty} r_n$.

Proof. Since $\{r_n\}$ is bounded from above, there exists $M \in \mathbb{R}$ such that $r_n \le M, \ \forall n \ge 1$. Using (7), we get $a_{n+1} \le (1-\gamma_n)a_n + (1-(1-\gamma_n))M$. This implies $a_2 \le (1-\gamma_1)a_1 + (1-(1-\gamma_1))M$, from which we obtain

$$\begin{aligned}
a_3 &\le (1-\gamma_2)a_2 + (1-(1-\gamma_2))M \\
&\le (1-\gamma_2)\Big((1-\gamma_1)a_1 + (1-(1-\gamma_1))M\Big) + (1-(1-\gamma_2))M \\
&= (1-\gamma_1)(1-\gamma_2)a_1 + \Big(1-(1-\gamma_1)(1-\gamma_2)\Big)M.
\end{aligned}$$

Using induction, we get for $n \ge 2$,

$$a_n \le \prod_{k=1}^{n-1}(1-\gamma_k)a_1 + \Big(1-\prod_{k=1}^{n-1}(1-\gamma_k)\Big)M. \tag{8}$$

Since $\sum \gamma_n = \infty$, it follows that $\prod(1-\gamma_n)$ converges to zero. Hence the sequence on the right-hand side of (8) converges to M. Hence, taking the limit superior on both sides of (8), we obtain $\limsup_{n\to\infty} a_n \le M$. Note that this holds for any upper bound M of the sequence $\{r_n\}$. In particular, if we define $\beta := \limsup_{n\to\infty} r_n$, we can find, for any $\varepsilon > 0$, an index $N \in \mathbb{N}$ such that $r_n \le M_\varepsilon := \beta + \varepsilon$. Since $\varepsilon > 0$ is arbitrary, the statement follows. $\square$

The following result is a special case of [19, Lem. 1].

Lemma 2.4. *For all $0 \ne v \in H$, $\tilde{y} \in H$, $x \in D^+$ and $\bar{x} \in D^-$, we have that $\|\bar{x} - x\|^2 \ge \|\bar{x} - z\|^2 + \|z - x\|^2$, where $z := P_D x$ is the projection of x onto the set $D := \{y \in H : \langle v, y - \tilde{y}\rangle = 0\}$, whereas D^+ and D^- are defined by $D^+ := \{y \in H : \langle v, y - \tilde{y}\rangle \ge 0\}$ and $D^- := \{y \in H : \langle v, y - \tilde{y}\rangle \le 0\}$, respectively.*

The following lemma was stated in [20, Prop. 2.11], see also [19, Prop. 4].

Lemma 2.5. *Let H_1 and H_2 be two real Hilbert spaces. Suppose $A : H_1 \to H_2$ is uniformly continuous on bounded subsets of H_1 and M is a bounded subset of H_1. Then $A(M)$ is bounded.*

The following lemma was proved, e.g., in [8, 17] and justifies why we are allowed to take metric projections onto the solution set of a continuous and monotone variational inequality.

Lemma 2.6. *Suppose A is a continuous monotone operator on a nonempty, closed, and convex subset C of a real Hilbert space H. Then the set of solutions to the variational inequality* (1) *is closed and convex.*

We finally restate a result which essentially states the equivalence between a primal and a dual variational inequality for continuous, monotone operators. One direction follows immediately from the monotonicity, whereas the other direction can be found, e.g, in [39, Lem. 7.1.7].

Lemma 2.7. *Let C be a nonempty, closed, and convex subset of a Hilbert space H. Let $A : C \to H$ be a continuous, monotone mapping and $z \in C$. Then*

$$z \in \mathrm{SOL} \Longleftrightarrow \langle Ax, x - z\rangle \geq 0 \quad \textit{for all } x \in C.$$

3 Projection-type Method

In this section, we give a precise statement of our projection-type method and discuss some of its elementary properties. Its convergence analysis is postponed to the next section. We first state the assumptions that we will assume to hold through the rest of this paper.

Assumption 3.1.
(a) The feasible set C is a nonempty, closed, and convex subset of the real Hilbert space H.

(b) $A : C \to H$ is a monotone and uniformly continuous on bounded subsets of C.

(c) The solution set SOL of $\mathrm{VI}(A, C)$ is nonempty.

Assumption (a) implies that projections onto C are well-defined. Condition (b) is slightly stronger than continuity of A; the same (or very similar) condition is also used, e.g., in [20, 6]. Note that this assumption is automatically satisfied for continuous operators defined on finite-dimensional Hilbert spaces $H = \mathbb{R}^n$. It also holds for the large class of bounded linear operators A on a general Hilbert space H.

Since our method depends on the choice of some sequences of parameters, we next summarize the conditions regarding these sequences. There is some freedom for the user to choose these parameters, but they have to be chosen with some care such that the conditions from the following assumption hold.

Assumption 3.2. Suppose the real sequences $\{\alpha_n\}$ and $\{\beta_n\}$ satisfy the following conditions:

(a) $\{\alpha_n\} \subset (0, 1)$ with $\lim_{n\to\infty} \alpha_n = 0$ and $\sum_{n=1}^{\infty} \alpha_n = \infty$.

(b) There is a constant $b < 1$ such that $0 < \beta_n \leq b$ for all $n \in \mathbb{N}$.

(c) $\lim_{n\to\infty} \alpha_n/\beta_n = 0$.

These conditions are satisfied, e.g., for $\alpha_n = 1/(n + 1)$ and $\beta_n = 1/\sqrt{2(n + 1)}$ or $\beta_n = \beta$ for all $n \in \mathbb{N}$, where $\beta \in (0, 1)$ is a given constant.

We next give a precise statement of our projection-type method. To this end, we use the abbreviation

$$r(x) := r_1(x) = x - P_C(x - Ax)$$

for the residual from (2) with $\gamma = 1$.

Algorithm 3.3. *(Projection-type Method)*

(S.0) *Choose sequences* $\{\alpha_n\}$ *and* $\{\beta_n\}$ *such that the conditions from Assumption 3.2 hold, and take* $\gamma, \sigma \in (0,1), s > 0$. *Let* $x_1 \in C$ *be a given starting point. Set* $n := 1$.

(S.1) *Set*

$$w_n := (1-\alpha_n)x_n + \alpha_n x_1.$$

If $r(w_n) = 0$*: STOP.*

(S.2) *Let* $y_n(\eta) := (1-\eta)w_n + \eta P_C(w_n - Aw_n)$ *for* $\eta \in \mathbb{R}$. *Compute* η_n *as the maximum of the numbers* $s, s\gamma, s\gamma^2, \ldots$ *such that*

$$\big\langle Ay_n(\eta_n), r(w_n)\big\rangle \geq \frac{\sigma}{2}\|r(w_n)\|^2,$$

and define $y_n := y_n(\eta_n)$.

(S.3) *Compute*

$$\begin{aligned} \lambda_n &:= \frac{\langle Ay_n, w_n - y_n\rangle}{\|Ay_n\|^2}, \\ x_{n+1} &:= (1-\beta_n)w_n + \beta_n P_C(w_n - \lambda_n Ay_n). \end{aligned}$$

(S.4) *Set* $n \leftarrow n+1$, *and go to (S.1).*

Recall that $r(w_n) = 0$ implies that we are at a solution of the variational inequality. In our convergence theory, we will implicitly assume that this does not occur after finitely many iterations, so that Algorithm 3.3 generates an infinite sequence satisfying, in particular, $r(w_n) \neq 0$ for all $n \in \mathbb{N}$. We will see that this property implies that Algorithm 3.3 is well-defined. From a practical point of view, this termination criterion has to be replaced by something like $\|r(w_n)\| \leq \varepsilon$ for some small $\varepsilon > 0$. In addition, a suitable implementation might also use an additional termination check like $\|r(x_{n+1})\| \leq \varepsilon$ in (S.3).

Remark 3.4. (a) Since C is convex, it is easy to see by a simple induction argument that all iterates x_n, y_n, w_n generated by Algorithm 3.3 belong to C. Consequently, the operator A needs to be defined on C only, not necessarily on the entire Hilbert space H.

(b) Algorithm 3.3 requires, at each iteration, only two projections onto the feasible set C. In particular, no projections onto sets like $C \cap H_k$ for some half-space H_k are needed. On the other hand, the stepsize rule in (S.2) involves a couple of evaluations of A, but these are often much less expansive than projections onto C.

(c) The scaling parameter $s > 0$ within the stepsize rule in (S.2) allows to start with a trial step $\eta \neq 1$ in each outer iteration. This plays some role from a numerical point of view since there exist both examples where larger stepsizes $\eta_n > 1$ can be accepted (which then, typically, yields faster convergence) and examples where $\eta_n < 1$ might be small (in which case a choice of $s < 1$ is useful to avoid unnecessarily many evaluations of A within the inner loop). $\Diamond$

We next want to show that Algorithm 3.3 is well-defined. To this end, we have to show that the inner loop in the stepsize calculation in (S.2) is always finite, and that the denominator in the definition of λ_n is nonzero.

Lemma 3.5. *The stepsize procedure in (S.2) of Algorithm 3.3 is well-defined, i.e. it terminates after finitely many inner loops.*

Proof. Consider an arbitrary index $n \in \mathbb{N}$ and recall that we always assume implicitly that $r(w_n) \neq 0$. Assume that the stepsize rule does not terminate finitely at this iteration n. Then we have

$$\big\langle A((1 - s\gamma^m)w_n + s\gamma^m P_C(w_n - Aw_n)), r(w_n)\big\rangle < \frac{\sigma}{2}\|r(w_n)\|^2, \quad \forall m \geq 0.$$

Since A is continuous, we obtain for $m \to \infty$ that

$$\big\langle Aw_n, w_n - P_C(w_n - Aw_n)\big\rangle \leq \frac{\sigma}{2}\|w_n - P_C(w_n - Aw_n)\|^2.$$

Let $z_n := w_n - Aw_n$. Then we get

$$2\big\langle w_n - z_n, w_n - P_C(w_n - Aw_n)\big\rangle \leq \sigma\|w_n - P_C(w_n - Aw_n)\|^2.$$

Using Lemma 2.2 (b), we obtain from the previous inequality

$$\begin{aligned}\|P_C(w_n - Aw_n) - w_n\|^2 + \|w_n - z_n\|^2 - \|P_C(w_n - Aw_n) - z_n\|^2 \\ \leq \sigma\|P_C(w_n - Aw_n) - w_n\|^2.\end{aligned}$$

Since $\|P_C(w_n - Aw_n) - w_n\| = \|r(w_n)\| > 0$ and $\sigma \in (0, 1)$, we obtain

$$\begin{aligned}\|P_C(w_n - Aw_n) - w_n\|^2 + \|w_n - z_n\|^2 - \|P_C(w_n - Aw_n) - z_n\|^2 \\ < \|P_C(w_n - Aw_n) - w_n\|^2.\end{aligned}$$

Hence $\|w_n - z_n\| < \|P_C(w_n - Aw_n) - z_n\|$. Since $z_n = w_n - Aw_n$ by definition and $w_n \in C$ in view of Remark 3.4 (a), this contradicts the definition of a metric projection. $\square$

A direct consequence of the previous result is that the scalar λ_n in (S.3) of Algorithm 3.3 is also well-defined.

Corollary 3.6. *We have $\langle Ay_n, w_n - y_n\rangle > 0$; in particular, $Ay_n \neq 0$ and, therefore λ_n is well-defined and positive.*

Proof. Consider once again a fixed iteration index $n \in \mathbb{N}$, and recall that $\|w_n - P_C(w_n - Aw_n)\| = \|r(w_n)\| > 0$ holds due to our implicit assumption regarding termination of the algorithm. Since the stepsize rule in (S.2) is well-defined by Lemma 3.5, the definition of y_n yields

$$\langle Ay_n, w_n - y_n\rangle = \eta_n\big\langle Ay_n, w_n - P_C(w_n - Aw_n)\big\rangle \geq \frac{\sigma\eta_n}{2}\|w_n - P_C(w_n - Aw_n)\|^2 > 0,$$

so the statements follow. $\square$

Observe that, in finding η_n, the operator A is evaluated (possibly) many times, but no extra projections onto the set C are needed. This is in contrast to a couple of related algorithms for the solution of monotone variational inequalities where the calculation of a suitable stepsize requires (possibly) many projections onto C, see, e.g., [14, 22, 40].

4 Convergence Analysis

Here we show that Algorithm 3.3 generates a sequence $\{x_n\}$ which converges strongly to a solution of the underlying variational inequality under Assumptions 3.1 and 3.2. To this end, we begin with a result that shows that the sequence $\{x_n\}$ generated by Algorithm 3.3 is bounded under the given assumptions.

Proposition 4.1. *Let Assumptions 3.1 and 3.2 hold. Then the sequence $\{x_n\}$ generated by Algorithm 3.3 is bounded.*

Proof. Let us define, for each n, the three sets

$$\begin{array}{rcl} D_n^- & := & \{x \in H : \langle Ay_n, x - y_n \rangle \leq 0\}, \\ D_n & := & \{x \in H : \langle Ay_n, x - y_n \rangle = 0\}, \quad \text{and} \\ D_n^+ & := & \{x \in H : \langle Ay_n, x - y_n \rangle \geq 0\}, \end{array}$$

where $\{y_n\}$ denotes the sequence generated by Algorithm 3.3. Recall also from Corollary 3.6 that $Ay_n \neq 0$ for all $n \in \mathbb{N}$.

Let $x^* \in \mathrm{SOL}$ be an arbitrary solution whose existence is guaranteed by Assumption 3.1 (c). Since A is monotone, we have

$$\langle Ax, x - x^* \rangle \geq 0, \quad \forall x \in C.$$

This implies $x^* \in D_n^-$ for all $n \in \mathbb{N}$ since $y_n \in C$ in view of Remark 3.4 (a). Furthermore, since we implicitly assume that Algorithm 3.3 does not terminate after finitely many steps with an exact solution, we have $\langle Ay_n, w_n - y_n \rangle > 0$ in view of Corollary 3.6. Therefore, $w_n \in D_n^+$ and $w_n \notin D_n^-$ for all $n \in \mathbb{N}$. Let $u_n := w_n - \lambda_n Ay_n$. Using the definition of λ_n, we have

$$u_n = w_n - \lambda_n Ay_n = w_n - \frac{\langle Ay_n, w_n - y_n \rangle}{\|Ay_n\|^2} Ay_n = P_{D_n}(w_n),$$

cf. [9, p. 130] for the last equation. Hence u_n is the metric projection of w_n onto the set D_n; in particular, we therefore have $u_n \in D_n$. Consequently, we obtain from Lemma 2.4 that

$$\|w_n - x^*\|^2 \geq \|u_n - x^*\|^2 + \|u_n - w_n\|^2. \tag{9}$$

Using Lemma 2.2 (b), (5), and setting $v_n := P_C(w_n - \lambda_n Ay_n) = P_C(u_n)$, we obtain

$$\|v_n - x^*\|^2 + \|v_n - u_n\|^2 - \|u_n - x^*\|^2 = 2\Big\langle v_n - u_n, v_n - x^* \Big\rangle \leq 0.$$

This implies

$$\|u_n - x^*\|^2 \geq \|v_n - x^*\|^2 + \|v_n - u_n\|^2. \tag{10}$$

It then follows from (9) and (10) that

$$\|w_n - x^*\|^2 \geq \|v_n - x^*\|^2 + \|v_n - u_n\|^2 + \|u_n - w_n\|^2.$$

Therefore,

$$\|v_n - x^*\|^2 \leq \|w_n - x^*\|^2 - \|v_n - u_n\|^2 - \|u_n - w_n\|^2. \tag{11}$$

Hence $\|v_n - x^*\| \leq \|w_n - x^*\|$, and this implies

$$\begin{aligned}\|v_n - w_n\|^2 &= \|v_n - x^* + x^* - w_n\|^2 \\ &= \|v_n - x^*\|^2 + \|w_n - x^*\|^2 + 2\langle v_n - x^*, x^* - w_n\rangle \\ &\leq 2\|w_n - x^*\|^2 + 2\langle w_n - x^*, x^* - v_n\rangle \\ &= 2\langle w_n - x^*, w_n - v_n\rangle. \end{aligned} \tag{12}$$

Now, we obtain from Algorithm 3.3 and (12) that

$$\begin{aligned}\|x_{n+1} - x^*\|^2 &= \|(w_n - x^*) - \beta_n(w_n - v_n)\|^2 \\ &= \|w_n - x^*\|^2 + \beta_n^2\|w_n - v_n\|^2 - 2\beta_n\langle w_n - x^*, w_n - v_n\rangle \\ &\leq \|w_n - x^*\|^2 - \beta_n(1 - \beta_n)\|w_n - v_n\|^2. \end{aligned} \tag{13}$$

Since $v_n - w_n = \frac{1}{\beta_n}(x_{n+1} - w_n)$, it follows from (13) that

$$\|x_{n+1} - x^*\|^2 \leq \|w_n - x^*\|^2 - \frac{1}{\beta_n}(1 - \beta_n)\|x_{n+1} - w_n\|^2. \tag{14}$$

Using Algorithm 3.3 and (13), we inductively obtain

$$\begin{aligned}\|x_{n+1} - x^*\| &\leq \|w_n - x^*\| \\ &\leq \alpha_n\|x_1 - x^*\| + (1 - \alpha_n)\|x_n - x^*\| \\ &\leq \max\left\{\|x_n - x^*\|, \|x_1 - x^*\|\right\} \\ &\vdots \\ &\leq \max\left\{\|x_1 - x^*\|, \|x_1 - x^*\|\right\} \\ &= \|x_1 - x^*\|. \end{aligned}$$

This shows that $\{x_n\}$ is bounded. □

Note that the previous proof does not need the uniform continuity assumption of A on bounded subsets of C. Furthermore, it does not require all conditions from Assumption 3.2. In particular, conditions (a) and (c) of Assumption 3.2 are not used in order to verify the boundedness of the sequence $\{x_n\}$ generated by Algorithm 3.3. Also note that the inequality (14) has not been used within the previous proof, but has been stated there since we will refer to it later.

As a simple consequence of the previous result, we also obtain the boundedness of several other sequences.

Corollary 4.2. *Let Assumptions 3.1 and 3.2 hold. Then the sequences $\{w_n\}, \{y_n\}$, and $\{Ay_n\}$ are also bounded.*

Proof. Since $\{x_n\}$ is bounded by Proposition 4.1, it follows immediately from the definition in Algorithm 3.3 that the sequence $\{w_n\}$ is also bounded. Using the fact that A is uniformly continuous on bounded subsets of C by Assumption 3.1 (b), we therefore obtain from Lemma 2.5 that the sequence $\{Aw_n\}$ is bounded. Consequently, $\{w_n - Aw_n\}$ is bounded, and the nonexpansiveness of the projection operator then implies that the sequence $\{P_C(w_n - Aw_n)\}$ is bounded. This, in turn, yields the boundedness of $\{y_n\}$. Using once more the uniform continuity of A on bounded subsets of C, we finally obtain the boundedness of the sequence $\{Ay_n\}$. □

The previous results allow us to verify strong convergence of any sequence $\{x_n\}$ generated by Algorithm 3.3. Note that, within the proof of this strong convergence result, we define some additional auxiliary sequences whose boundedness is stated without an explicit proof, but that the corresponding proofs are more or less identical to the one given for the sequences from Corollary 4.2.

Theorem 4.3. *Let Assumptions 3.1 and 3.2 hold. Then the sequence $\{x_n\}$ generated by Algorithm 3.3 strongly converges to a solution $z \in SOL$, where $z := P_{SOL}x_1$.*

Proof. As in the proof of Proposition 4.1, we use the abbreviations

$$u_n := w_n - \lambda_n A y_n \quad \text{and} \quad v_n := P_C(u_n) = P_C(w_n - \lambda A y_n).$$

We now divide the proof into four steps.

Step 1: Here we show that there is a subsequence such that $\lim_{k\to\infty}\langle Ay_{n_k}, w_{n_k} - y_{n_k}\rangle = 0$ holds. To this end, observe that Algorithm 3.3 yields

$$\|v_n - w_n\|^2 = \Big\|\frac{1}{\beta_n}(x_{n+1} - w_n)\Big\|^2 = \frac{\alpha_n}{\beta_n}\left(\frac{\|x_{n+1} - w_n\|^2}{\alpha_n\beta_n}\right). \tag{15}$$

Using Algorithm 3.3 and noting that $\alpha_n \in (0,1)$, we have

$$\begin{aligned}
\|w_n - z\|^2 &= \|\alpha_n(x_1 - z) + (1-\alpha_n)(x_n - z)\|^2 \\
&= \alpha_n^2\|x_1 - z\|^2 + 2\alpha_n(1-\alpha_n)\langle x_1 - z, x_n - z\rangle + (1-\alpha_n)^2\|x_n - z\|^2 \\
&\leq \alpha_n^2\|x_1 - z\|^2 - 2\alpha_n(1-\alpha_n)\langle x_1 - z, z - x_n\rangle + (1-\alpha_n)\|x_n - z\|^2.
\end{aligned}$$

Exploiting this inequality in (14) (with the general solution x^* replaced by the particular solution z), we obtain

$$\begin{aligned}
\|x_{n+1} - z\|^2 &\leq \alpha_n^2\|x_1 - z\|^2 - 2\alpha_n(1-\alpha_n)\langle x_1 - z, z - x_n\rangle \\
&\quad +(1-\alpha_n)\|x_n - z\|^2 - \frac{1}{\beta_n}(1-\beta_n)\|x_{n+1} - w_n\|^2 \\
&= (1-\alpha_n)\|x_n - z\|^2 - \alpha_n\Big(-\alpha_n\|x_1 - z\|^2 + \\
&\quad 2(1-\alpha_n)\langle x_1 - z, z - x_n\rangle + \frac{1}{\alpha_n\beta_n}(1-\beta_n)\|x_{n+1} - w_n\|^2\Big).
\end{aligned} \tag{16}$$

Using the abbreviation

$$\Gamma_n := -\alpha_n\|x_1 - z\|^2 + 2(1-\alpha_n)\langle x_1 - z, z - x_n\rangle + \frac{1}{\alpha_n\beta_n}(1-\beta_n)\|x_{n+1} - w_n\|^2 \tag{17}$$

for the term in parantheses, we can rewrite (16) as

$$\|x_{n+1} - z\|^2 \leq (1-\alpha_n)\|x_n - z\|^2 - \alpha_n\Gamma_n. \tag{18}$$

Since

$$\begin{aligned}
\Gamma_n &\geq -\alpha_n\|x_1 - z\|^2 + 2(1-\alpha_n)\langle x_1 - z, z - x_n\rangle \\
&\geq -\alpha_n\|x_1 - z\|^2 - 2(1-\alpha_n)\cdot\|x_1 - z\|\cdot\|z - x_n\|
\end{aligned}$$

and the sequence $\{x_n\}$ is bounded by Proposition 4.1, it follows that $\{\Gamma_n\}$ is bounded from below. Consequently, $\liminf_{n\to\infty} \Gamma_n$ is a finite real number, and by Assumption 3.2, we have from (17) that

$$\liminf_{n\to\infty} \Gamma_n = \liminf_{n\to\infty} \Big(2\langle x_1 - z, z - x_n\rangle + \frac{1}{\alpha_n \beta_n}(1-\beta_n)\|x_{n+1} - w_n\|^2\Big).$$

Exploiting the boundedness of $\{x_n\}$ once more, there exists a subsequence $\{x_{n_k}\}$ of $\{x_n\}$ such that

$$\liminf_{n\to\infty} \Gamma_n = \lim_{k\to\infty} \Big(2\langle x_1 - z, z - x_{n_k}\rangle + \frac{1}{\alpha_{n_k} \beta_{n_k}}(1-\beta_{n_k})\|x_{n_k+1} - w_{n_k}\|^2\Big) \tag{19}$$

and $x_{n_k} \rightharpoonup p$ for some element $p \in H$. Furthermore, since $x_n \in C$ for all $n \in \mathbb{N}$ in view of Remark 3.4 (a) and the closed set C is also weakly sequentially closed, cf. [4, Thm. 3.32], we have $p \in C$.

Since $\{x_n\}$ is bounded and $\liminf_{n\to\infty} \Gamma_n$ is finite, we obtain from (19) that the subsequence $\frac{1}{\alpha_{n_k}\beta_{n_k}}(1-\beta_{n_k})\|x_{n_k+1} - w_{n_k}\|^2$ is bounded. Furthermore, by Assumption 3.2, there exists $b \in (0,1)$ such that $\beta_n \leq b < 1$, and this implies that $\frac{1}{\alpha_n\beta_n}(1-\beta_n) \geq \frac{1}{\alpha_n\beta_n}(1-b) > 0$, so we have that $\frac{1}{\alpha_{n_k}\beta_{n_k}}\|x_{n_k+1} - w_{n_k}\|^2$ is bounded, too. Hence we obtain from (15) and $\frac{\alpha_{n_k}}{\beta_{n_k}} \to 0,\ k \to \infty$ that $\|v_{n_k} - w_{n_k}\| \to 0,\ k \to \infty$. Using (11) with x^* replaced by z, we therefore obtain

$$\begin{aligned}
\|u_{n_k} - w_{n_k}\|^2 &\leq \|w_{n_k} - z\|^2 - \|v_{n_k} - z\|^2 \\
&= \big(\|w_{n_k} - z\| - \|v_{n_k} - z\|\big)\big(\|w_{n_k} - z\| + \|v_{n_k} - z\|\big) \\
&\leq \|w_{n_k} - v_{n_k}\|\big(\|w_{n_k} - z\| + \|v_{n_k} - z\|\big) \\
&\leq \|w_{n_k} - v_{n_k}\| M \\
&\to 0, \quad k \to \infty,
\end{aligned}$$

for some constant $M > 0$ whose existence follows from the boundedness of $\{w_n\}$ and $\{v_n\}$, cf. Corollary 4.2 and the comments after its proof. Since $u_n \in D_n$, with D_n defined as in the proof of Proposition 4.1, we have

$$0 = \langle Ay_n, u_n - y_n\rangle = \langle Ay_n, u_n - w_n\rangle + \langle Ay_n, w_n - y_n\rangle.$$

Hence, using the boundedness of $\{\|Ay_n\|\}$ by Corollary 4.2, we get

$$\begin{aligned}
|\langle Ay_{n_k}, w_{n_k} - y_{n_k}\rangle| &= |\langle Ay_{n_k}, w_{n_k} - u_{n_k}\rangle| \\
&\leq \|Ay_{n_k}\|\|w_{n_k} - u_{n_k}\| \to 0, \quad k \to \infty.
\end{aligned}$$

We therefore have $\lim_{k\to\infty}\langle Ay_{n_k}, w_{n_k} - y_{n_k}\rangle = 0$.

Step 2: We claim that there exists a subsequence of $\{w_{n_k}\}$ such that, for all $x \in C$, it holds that $0 \leq \liminf_{k\in K}\langle Aw_{n_k}, x - w_{n_k}\rangle$, where $K \subseteq \mathbb{N}$ defines the corresponding subsequence; here, $\{w_{n_k}\}$ denotes the subsequence already given from Step 1.

To achieve this, let us define $s_{n_k} := P_C(w_{n_k} - Aw_{n_k})$ for all $k \in \mathbb{N}$. We distinguish two cases depending on the behaviour of (the bounded) sequence of stepsizes $\{\eta_{n_k}\}$.

Case 1: Suppose that $\liminf_{k\to\infty}\{\eta_{n_k}\} = 0$. Subsequencing if necessary, we may assume without loss of generality that $\lim_{k\to\infty}\{\eta_{n_k}\} = 0$. We may therefore assume that $\eta_{n_k} < s$ for all $k \in \mathbb{N}$ so that the stepsize gets reduced at least once for all iterations belonging to this subsequence. In other words, this means that the trial stepsize $\frac{1}{\gamma}\eta_{n_k}$ did not satisfy the test from (S.2) of Algorithm 3.3.

We first show that this implies $\lim_{k\to\infty}\|s_{n_k} - w_{n_k}\| = 0$. To this end, it is obviously enough to show that $\limsup_{k\to\infty}\|s_{n_k} - w_{n_k}\| = 0$ holds. Assume the contrary that $\limsup_{k\to\infty}\|s_{n_k} - w_{n_k}\| = \delta$ for some (finite) constant $\delta > 0$. Let $\bar{y}_k := \frac{1}{\gamma}\eta_{n_k}s_{n_k} + (1 - \frac{1}{\gamma}\eta_{n_k})w_{n_k}$ or, equivalently, $\bar{y}_k - w_{n_k} = \frac{1}{\gamma}\eta_{n_k}(s_{n_k} - w_{n_k})$. Since $\{s_{n_k} - w_{n_k}\}$ is bounded in view of the definition of s_{n_k} and Corollary 4.2, and since $\lim_{k\to\infty}\eta_{n_k} = 0$ holds, it follows that

$$\lim_{k\to\infty}\|\bar{y}_k - w_{n_k}\| = 0. \tag{20}$$

From the stepsize rule and the definition of $\bar{y}_k$, we have

$$\langle A\bar{y}_k, w_{n_k} - s_{n_k}\rangle < \frac{\sigma}{2}\|w_{n_k} - s_{n_k}\|^2, \quad \forall k \in \mathbb{N}.$$

Since A is uniformly continuous on bounded subsets of C, $\sigma \in (0,1)$, and the right-hand side is bounded from below by a positive constant in view of our assumption, we obtain from (20) that there exists $N \in \mathbb{N}$ such that

$$2\langle Aw_{n_k}, w_{n_k} - s_{n_k}\rangle < \|w_{n_k} - s_{n_k}\|^2, \quad \forall k \in \mathbb{N}, k \geq N.$$

Therefore,

$$2\langle w_{n_k} - t_{n_k}, w_{n_k} - s_{n_k}\rangle < \|w_{n_k} - s_{n_k}\|^2, \quad \forall k \in \mathbb{N}, k \geq N,$$

where $t_{n_k} := w_{n_k} - Aw_{n_k}$. Using Lemma 2.2 (b) in the last inequality, we obtain

$$\|w_{n_k} - s_{n_k}\|^2 + \|w_{n_k} - t_{n_k}\|^2 - \|s_{n_k} - t_{n_k}\|^2 < \|w_{n_k} - s_{n_k}\|^2 \quad \forall k \in \mathbb{N}, k \geq N.$$

Hence $\|w_{n_k} - t_{n_k}\| < \|s_{n_k} - t_{n_k}\|$ which is a contradiction to the definition of $s_{n_k} = P_C(t_{n_k})$. Therefore $\lim_{k\to\infty}\|s_{n_k} - w_{n_k}\| = \limsup_{k\to\infty}\|s_{n_k} - w_{n_k}\| = 0$.

Furthermore, the definition of s_{n_k} together with (5) yields

$$\langle w_{n_k} - Aw_{n_k} - s_{n_k}, x - s_{n_k}\rangle \leq 0, \quad \forall x \in C,$$

which implies that

$$\langle w_{n_k} - s_{n_k}, x - s_{n_k}\rangle \leq \langle Aw_{n_k}, x - s_{n_k}\rangle, \quad \forall x \in C.$$

Hence,

$$\langle w_{n_k} - s_{n_k}, x - s_{n_k}\rangle + \langle Aw_{n_k}, s_{n_k} - w_{n_k}\rangle \leq \langle Aw_{n_k}, x - w_{n_k}\rangle, \quad \forall x \in C. \tag{21}$$

Fix $x \in C$ and let $k \to \infty$ in (21). Since $\lim_{k\to\infty}\|s_{n_k} - w_{n_k}\| = 0$, we have $0 \leq \liminf_{k\to\infty}\langle Aw_{n_k}, x - w_{n_k}\rangle$ for all $x \in C$ and therefore the desired statement.

Case 2: Suppose that $\liminf_{k\to\infty}\{\eta_{n_k}\} > 0$. Then there is a constant $\mu > 0$ such that $\eta_{n_k} \geq \mu > 0$ holds for all $k \in \mathbb{N}$. It follows from the stepsize rule in Algorithm 3.3 that

$$\langle Ay_{n_k}, w_{n_k} - y_{n_k}\rangle \geq \frac{\sigma}{2}\eta_{n_k}\|w_{n_k} - s_{n_k}\|^2.$$

Therefore, Step 1 implies $\lim_{k\to\infty}\|s_{n_k} - w_{n_k}\| = 0$. Following the same line of arguments in (21) above, we obtain $0 \leq \liminf_{k\to\infty}\langle Aw_{n_k}, x - w_{n_k}\rangle$ for all $x \in C$, hence the statement from Step 2 also holds in the second case.

Step 3: We show that $p \in \mathrm{SOL}$, where $p \in C$ denotes the weak limit of the subsequence $\{x_{n_k}\}$ from Step 1 of this proof. Since A is monotone, we have for an arbitrary $x \in C$ that

$$\langle Ax, x - w_{n_k}\rangle \geq \langle Aw_{n_k}, x - w_{n_k}\rangle, \quad \forall k \in \mathbb{N}. \tag{22}$$

Taking the lim inf on both sides of (22), and using Step 2 above, we get

$$\liminf_{k\to\infty}\langle Ax, x - w_{n_k}\rangle \geq \liminf_{k\to\infty}\langle Aw_{n_k}, x - w_{n_k}\rangle \geq 0$$

for all $x \in C$. Since $\|w_{n_k} - x_{n_k}\| \leq \alpha_{n_k}\|x_1 - x_{n_k}\| \to 0$ and $x_{n_k} \rightharpoonup p$, it follows that $w_{n_k} \rightharpoonup p$. We therefore have for all $x \in C$ that

$$\langle Ax, x - p\rangle = \lim_{k\to\infty}\langle Ax, x - w_{n_k}\rangle = \liminf_{k\to\infty}\langle Ax, x - w_{n_k}\rangle \geq 0.$$

In view of Lemma 2.7, this implies $p \in \mathrm{SOL}$.

Step 4: We finally show that $\lim_{k\to\infty} x_n = z$. Using (19), (5), and $p \in \mathrm{SOL}$ in view of Step 3, we obtain

$$\begin{aligned}\liminf_{n\to\infty}\Gamma_n &= \lim_{k\to\infty}\Big(2\langle x_1 - z, z - x_{n_k}\rangle + \frac{1}{\alpha_{n_k}\beta_{n_k}}(1-\beta_{n_k})\|x_{n_k+1} - w_{n_k}\|^2\Big)\\ &\geq 2\lim_{k\to\infty}\langle x_1 - z, z - x_{n_k}\rangle = \langle x_1 - z, z - p\rangle \geq 0.\end{aligned}$$

On the other hand, applying Lemma 2.3 with $r_n := -\Gamma_n$ and recalling that this sequence $\{r_n\}$ is bounded from above, see the proof of Step 1, we obtain from Assumption 3.2 and (18) that

$$\limsup_{n\to\infty}\|x_n - z\|^2 \leq \limsup_{n\to\infty}(-\Gamma_n) = -\liminf_{n\to\infty}\Gamma_n.$$

Consequently, we get $\limsup_{n\to\infty}\|x_n - z\|^2 \leq -\liminf_{n\to\infty}\Gamma_n \leq 0$. Therefore, $\lim_{n\to\infty}\|x_n - z\| = 0$, and this means that $\{x_n\}$ converges strongly to z. □

Note that the convergence of the iterates to the projection of the starting point onto the solution set, as guaranteed by Theorem 4.3, is an interesting property which is different, for example, from the class of Tikhonov-type regularization approaches where the corresponding sequence always converges to the same solution. Hence, by a suitable choice of the starting point, we have the chance to compute different solutions, and using many starting points, one might even get an idea of the geometric shape of the whole solution set of the underlying variational inequality. Furthermore, if one has some a priori knowledge regarding the location of a solution and is therefore interested in computing a particular solution which is as close as possible to this a priori knowledge, Algorithm 3.3 allows to take this knowledge into account by a suitable choice of x_1.

We next discuss some of the assumptions used in Theorem 4.3.

Remark 4.4. (a) In the case when the operator A is pseudomonotone and uniformly continuous on bounded subsets of C, we can still obtain a strong convergence result using our iterative method by assuming additionally that A is (w,s) sequential continuous on C (i.e., that A maps weakly convergent sequences into strongly convergent ones). The (w,s) sequential continuity of a pseudomonotone operator A was also assumed by Ceng et al. [11] and Yao and Postolache [41] in order to obtain weak and strong convergence results for variational inequality problems involving a Lipschitz continuous, pseudomonotone operator A in an infinite dimensional Hilbert space. In our own case, different from [11, 41], we can obtain the strong convergence result for a pseudomonotone operator A in an infinite-dimensional Hilbert space without assuming that our sequence of iterates satisfies $0 \leq \liminf_{n\to\infty}\langle Ax_n, z-x_n\rangle, \forall z \in C$. In this respect, we are also able to improve the results from [11, 41]. However, since we feel that the (w,s) sequentional continuity of A is a rather strong assumption, we eventually decided to work with monotone operators only, where this condition can be avoided completely.

(b) In finite-dimensional spaces, the assumption that A is uniformly continuous on bounded subsets of C, automatically holds. Moreover, in this case, for pseudomonotone operators A, there is also no need to assume the (w,s) sequential continuity of the operator A, cf. comment (a), only continuity of A is required. On the other hand, as discussed in the introduction, and taking into account that strong and weak convergence coincide in finite dimensions, there exist suitable methods whose computational overhead is less per iteration than for Algorithm 3.3.

(c) It is difficult to construct suitable couterexamples which show analytically that the conditions from Assumption 3.2 regarding the choice of the two sequences $\{\alpha_n\}$ and $\{\beta_n\}$ cannot be relaxed. The crucial conditions are Assumptions 3.2 (b) and (c), whereas (a) is standard and well-accepted in many fixed-point methods. Regarding (c), there is at least a numerical indication that this condition is strict. To this end, consider an example with $H := C := \mathbb{R}^2$, so the variational inequality reduces to the system of equations $Ax = 0$. We consider the monotone operator $Ax := \max\big\{0, (\langle a, x\rangle - \delta)/\|a\|_2^2\big\}a$ with the data $a := (-2,-1)^T, \delta := -4$. The corresponding solution set is $\mathrm{SOL} = \{x \mid 2x_1 + x_2 \geq 4\}$. We use the parameters $\gamma := 1/4, \sigma := 1/2, s := 1$ as well as the sequences

$$\alpha_n := \frac{1}{n+1}, \quad \beta_n := \begin{cases} \frac{1}{6(n+1)}, & \text{for } n = 1,3,5,7,\dots, \\ \frac{\log(n+1)}{n+1}, & \text{for } n = 2,4,6,8,\dots \end{cases}$$

which satisfy $\liminf_{n\to\infty}\alpha_n/\beta_n = 0$ and $\limsup_{n\to\infty}\alpha_n/\beta_n > 0$, hence condition (c) is violated. Applying Algorithm 3.3 to this example using the starting point $x_1 := (1,1)^T$, the residual $\|r(w_n)\|_2$ after one million iterations is around 0.1481, so the method does not seem to converge. $\Diamond$

5 Numerical Experiments

We consider a few examples in order to illustrate different properties of Algorithm 3.3. To this end, we begin with a class of optimal control problems.

Let $\Omega \subseteq \mathbb{R}^n$ be a given domain, and consider the optimal control problem

$$\begin{array}{ll} \min & J(u) := \frac{1}{2}\|y(u) - y_d\|^2_{L^2(\Omega)} + \frac{\alpha}{2}\|u - u_d\|^2_{L^2(\Omega)} \\ \text{s.t.} & u \in \mathcal{F} := \{u \in L^2(\Omega) \,|\, u \le \psi \text{ a.e. in } \Omega\}, \end{array} \tag{23}$$

where $\alpha > 0$ denotes the regularization parameter, $y = y(u) \in H^1_0(\Omega)$ is the weak solution of

$$-\Delta y = u \quad \text{on } \Omega$$

and

$$y_d, u_d, \psi \in L^2(\Omega)$$

are given functions. Here u denotes the control and y the state variable.

To be more specific, let $\Omega = (0,1) \times (0,1) \subseteq \mathbb{R}^2$ and let A denote the standard five-point finite difference approximation to the negative Laplacian with uniform stepsize $h := 1/(N+1)$ for some $N \in \mathbb{N}$. Then the discretized optimal control problem becomes

$$\min_{u,y} \frac{1}{2}\|y - y_d\|_2^2 + \frac{\alpha}{2}\|u - u_d\|_2^2 \quad \text{s.t.} \quad Ay = u, \;\; \psi - u \ge 0,$$

where the discretized functions u, y etc. are denoted by the same letters as their continuous counterparts. Using $u = Ay$ in order to remove the control variable, we obtain

$$\min_{y} \frac{1}{2}\|y - y_d\|_2^2 + \frac{\alpha}{2}\|Ay - u_d\|_2^2 \quad \text{s.t.} \quad \psi - Ay \ge 0.$$

Setting $v := \psi - Ay$ then gives

$$\min_{v} \frac{1}{2}\|A^{-1}(\psi - v) - y_d\|_2^2 + \frac{\alpha}{2}\|\psi - v - u_d\|_2^2 \quad \text{s.t.} \quad v \ge 0.$$

Defining $v_d := y_d - A^{-1}\psi$ and $\psi_d := u_d - \psi$, we finally obtain the problem

$$\min_{v} f(v) := \frac{1}{2}\|A^{-1}v + v_d\|_2^2 + \frac{\alpha}{2}\|v + \psi_d\|_2^2 \quad \text{s.t.} \quad v \ge 0$$

which is obviously equivalent to the linear complementarity problem

$$v \ge 0, \;\; F(v) \ge 0, \;\; v^T F(v) = 0$$

with

$$F(v) := \nabla f(v) := \underbrace{\left(A^{-1}A^{-1} + \alpha I\right)}_{=:M} v + \underbrace{A^{-1}v_d + \alpha\psi_d}_{=:q}.$$

In other words, we have the variational inequality $\mathrm{VI}(F, C)$ with the nonnegative orthant $C := \mathbb{R}^n_+$, so that projections onto C are easy to compute.

Note that there is no need to compute the inverse A^{-1} explicitly. In fact, this can be avoided by computing a few vectors of the form $A^{-1}b$ for some right-hand

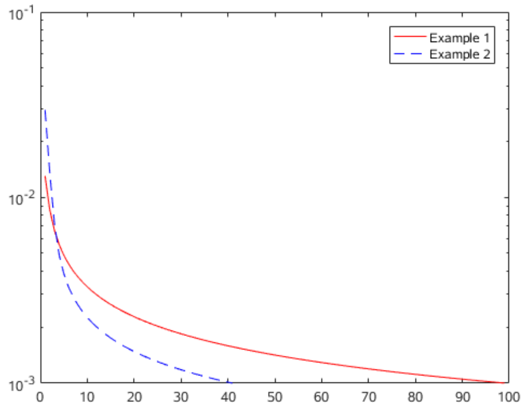

Figure 1: Behaviour of the residuals $\|r(w_n)\|_\infty$ for Examples (24) (Example 1) and (25) (Example 2)

sides b. This means that we have to solve a linear system of equations with the coefficient matrix A. Fortunately, since A corresponds to the five-point finite difference approximation of the negative Laplacian, these systems can be solved, e.g., by a fast sine transform, in only $O(N^2 \log_2 N)$ arithmetic operations. Altogether, it follows that one outer iteration of the double projection method applied to the discretized optimal control problem is reasonably cheap.

Taking into account these considerations, let us apply our method to an example from [7] with the following data:

$$y_d(x_1,x_2) := \frac{1}{6}\sin(2\pi x_1)\sin(2\pi x_2)\exp(2x_1), \quad u_d \equiv 0, \quad \psi \equiv 0, \quad \alpha := 10^{-2}. \tag{24}$$

We use the discretization parameter $N = 128$. Note that the dimension of the corresponding variational inequality is $n = N^2$. The parameters in Algorithm 3.3 were chosen as $\gamma = 0.5, \sigma = 10^{-4}, s = 50$, and the sequences from Assumption 3.2 were taken as $\alpha_n := 10^{-4}/n$ and $\beta_n := 0.99$ for all n. We terminate the iteration if $\|r(w_n)\|_\infty \le \varepsilon$. Since projection-type methods typically have nice global convergence properties and often get close to the solution relatively fast, but have only a poor local rate of convergence, we use $\varepsilon := 10^{-3}$. Algorithm 3.3 terminates successfully after 99 iterations. The iteration history is given in Figure 1. The optimal control and optimal state corresponding to the computed solution are shown in Figure 2.

As a second example, also taken from [7], consider the data

$$\begin{aligned} y_d(x_1,x_2) &:= \begin{cases} 200x_1x_2(x_1-\frac{1}{2})^2(1-x_2), & \text{if } 0 < x_1 \le \frac{1}{2}, \\ 200x_2(x_1-1)(x_1-\frac{1}{2})^2(1-x_2), & \text{if } \frac{1}{2} < x_1 \le 1, \end{cases} \\ u_d &\equiv 0,\ \psi \equiv 1,\ \alpha := 10^{-2}. \end{aligned} \tag{25}$$

Here, our method terminates after 40 iterations. The iteration history is also given in Figure 1, whereas the resulting optimal control and state are depicted in Figure 3.

Finally, we consider an artifical two-dimensional example given by $C := \mathbb{R}^2_+$ and $F = (F_1, F_2)$ with

$$F_i(x) := \max\left\{l, \min\{x-b, u\}\right\} \tag{26}$$

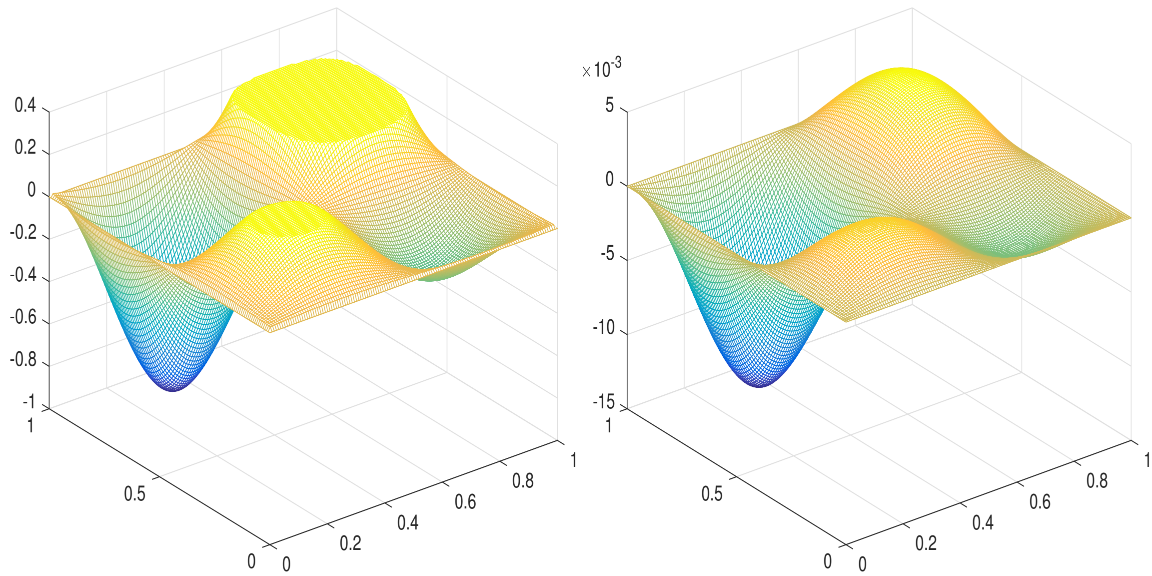

Figure 2: Optimal control (left) and optimal state (right) for the example from (24)

using lower bounds $l := (0,0)^T$, upper bounds $u := (1,1)^T$, and the shift $b := (1,1)^T$. It is not difficult to see that this example is monotone (but not strongly monotone) with the (non-singleton) solution set $\mathrm{SOL} = [0,1]^2$. We apply Algorithm 3.3 with the same parameters as before (except that $s = 1$) and illustrate the behaviour of this method using different starting points from the feasible set C in Figure 4. This figure clearly shows that the sequences really converge to the particular solution given by the projection of the starting point x_1 onto SOL, as indicated by Theorem 4.3.

6 Final Remarks

This paper presents a strong convergence result for monotone variational inequality problems in real Hilbert spaces. The method requires, at each iteration, two projections onto the feasible set of the variational inequality. Part of our future research concentrates on the development of a suitable method where only a single projection is needed, so the computational overhead becomes comparable to some of those methods for which, so far, only weak convergence is known.

References

[1] R.Y. Apostol, A.A. Grynenko, and V.V. Semenov; Iterative algorithms for monotone bilevel variational inequalities, *J. Comp. Appl. Math.* **107** (2012), 3-14.

[2] J.-P. Aubin and I. Ekeland; *Applied Nonlinear Analysis*, Wiley, New York, 1984.

[3] C. Baiocchi and A. Capelo; *Variational and Quasivariational Inequalities; Applications to Free Boundary Problems*, Wiley, New York (1984).

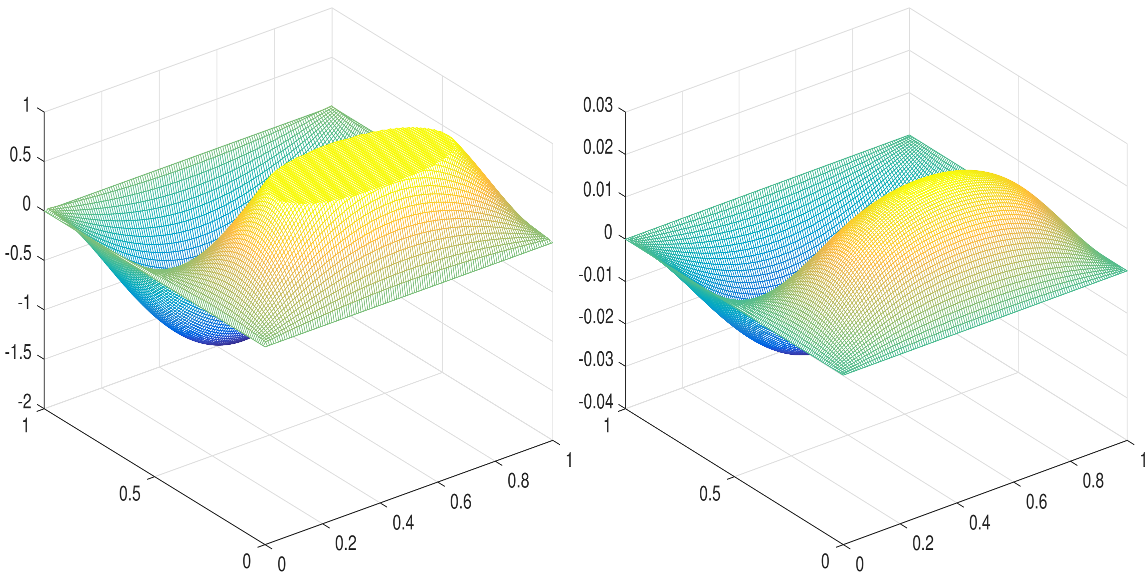

Figure 3: Optimal control (left) and optimal state (right) for the example from (25)

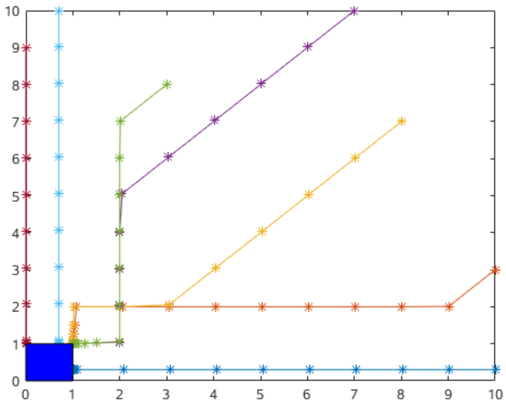

Figure 4: Iterations for different starting points for the artificial example from (26)

[4] H.H. Bauschke and P.L. Combettes; *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, CMS Books in Mathematics, Springer, New York (2011).

[5] J.Y. Bello Cruz and A.N. Iusem; A strongly convergent direct method for monotone variational inequalities in Hilbert spaces, *Num. Funct. Anal. Optim.* **30** (2009), 23-36.

[6] J.Y. Bello Cruz and A.N. Iusem; An explicit algorithm for monotone variational inequalities, *Optim.* **61** (2012), 855-871.

[7] M. Bergounioux, M. Haddou, M. Hintermüller, and K. Kunisch:; A comparison of a Moreau-Yosida based active set strategy and interior point methods for constrained optimal control problems, *SIAM J. Opt.* **11** (2001), 495-521.

[8] F. Browder; Nonlinear monotone operators and convex sets in Banach spaces, *Bull. Amer. Math. Soc.* **71** (1965), 780-785.

[9] A. Cegielski; *Iterative Methods for Fixed Point Problems in Hilbert Spaces*, Lecture Notes in Mathematics 2057, Springer, Berlin, 2012.

[10] L.C. Ceng, N. Hadjisavvas, and N.-C. Wong; Strong convergence theorem by a hybrid extragradient-like approximation method for variational inequalities and fixed point problems, *J. Glob. Optim.* **46** (2010), 635-646.

[11] L.C. Ceng, M. Teboulle, and J.C. Yao; Weak convergence of an iterative method for pseudomonotone variational inequalities and fixed-point problems, *J. Optim. Theory Appl.* **146** (2010), 19-31.

[12] Y. Censor, A. Gibali, and S. Reich; Strong convergence of subgradient extragradient methods for the variational inequality problem in Hilbert space, *Optim. Methods Softw.* **26** (2011), 827-845.

[13] Y. Censor, A. Gibali, and S. Reich; The subgradient extragradient method for solving variational inequalities in Hilbert space, *J. Optim. Theory Appl.* **148** (2011), 318-335.

[14] S. Denisov, V. Semenov, and L. Chabak; Convergence of the modified extragradient method for variational inequalities with non-Lipschitz operators, *Cybernet. Systems Anal.* **51** (2015), 757-765.

[15] F. Facchinei and J.-S. Pang; *Finite-Dimensional Variational Inequalities and Complementarity Problems, Volume II.* Springer Series in Operations Research, Springer, New York, 2003.

[16] R. Glowinski, J.-L. Lions, and R. Trémolières; *Numerical Analysis of Variational Inequalities*, North-Holland, Amsterdam (1981).

[17] Ph. Hartman and G. Stampacchia; On some non linear elliptic differential functional equations, *Acta Math.* **115** (1966), 271-310.

[18] B.-S. He, Z.-H. Yang, and X.-M. Yuan; An approximate proximal-extragradient type method for monotone variational inequalities, *J. Math. Anal. Appl.* **300** (2004), 362-374.

[19] A.N. Iusem and R. Gárciga Otero; Inexact versions of proximal point and augmented Lagrangian algorithms in Banach spaces, *Numer. Funct. Anal. Optim.* **22** (2001), 609-640.

[20] A.N. Iusem and M. Nasri; Korpelevich's method for variational inequality problems in Banach spaces, *J. Glob. Optim.* **50** (2011), 59-76.

[21] A.N. Iusem and B.F. Svaiter, A variant of Korpelevich's method for variational inequalities with a new search strategy, *Optimization* **42** (1997), 309-321.

[22] E.N. Khobotov; Modification of the extragradient method for solving variational inequalities and certain optimization problems, *USSR Comput. Math. Math. Phys.* **27** (1989), 120-127.

[23] D. Kinderlehrer and G. Stampacchia; *An Introduction to Variational Inequalities and Their Applications*, Academic Press, New York (1980).

[24] I.V. Konnov; *Combined Relaxation Methods for Variational Inequalities*, Springer-Verlag, Berlin (2001).

[25] G.M. Korpelevich; The extragradient method for finding saddle points and other problems, *Ekonomika i Mat. Metody* **12** (1976), 747-756.

[26] M.A. Noor: Modified projection method for pseudomonotone variational inequalities, *Appl. Math. Letters* **15** (2002), 315-320.

[27] R. Kraikaew and S. Saejung; Strong convergence of the Halpern subgradient extragradient method for solving variational inequalities in Hilbert spaces, *J. Optim. Theory Appl.* **163** (2014), 399-412.

[28] S.I. Lyashko, V.V. Semenov, and T.A. Voitova; Low-cost modification of Korpelevich's method for monotone equilibrium problems, *Cybernet. Systems Anal.* **47** (2011), 631-639.

[29] P.-E. Maingé; Numerical approach to monotone variational inequalities by a one-step projected reflected gradient method with line-search procedure, *Comp. Math. Appl.* **72** (3) (2016), 720-728.

[30] P.-E. Maingé and M.L. Gobinddass; Convergence of one-step projected gradient methods for variational inequalities, *J. Optim. Theory Appl.* **171** (2016), 146-168.

[31] P.-E. Maingé and S. Măruşter; Convergence in norm of modified Krasnoselki-Mann iterations for fixed points of demicontractive mappings, *Set-Valued Anal.* **15** (2007), 67-79.

[32] Yu.V. Malitsky; Projected reflected gradient methods for monotone variational inequalities, *SIAM J. Optim.* **25** (2015), 502-520.

[33] Yu.V. Malitsky and V.V. Semenov; An extragradient algorithm for monotone variational inequalities, *Cybernet. Systems Anal.* **50** (2014), 271-277.

[34] Yu.V. Malitsky and V.V. Semenov; A hybrid method without extrapolation step for solving variational inequality problems, *J. Global Optim.* **61** (2015), 193-202.

[35] N. Nadezhkina and W. Takahashi; Strong convergence theorem by a hybrid method for nonexpansive mappings and Lipschitz-continuous monotone mappings, *SIAM J. Optim.* **16** (2006), 1230-1241.

[36] A. Nagurney; *Network Economics: A Variational Inequality Approach*, Kluwer Academic Publishers, Dordrecht (1999).

[37] L.D. Popov; A modification of the Arrow-Hurwicz method for finding saddle points, *Math. Notes* **28** (1980), 845-848.

[38] M.V. Solodov and B.F. Svaiter; A new projection method for variational inequality problems, *SIAM J. Control Optim.* **37** (1999), 765-776.

[39] W. Takahashi; *Nonlinear Functional Analysis*, Yokohama Publishers, Yokohama, (2000).

[40] P. Tseng; A modified forward-backward splitting method for maximal monotone mappings, *SIAM J. Control Optim.* **38** (2000), 431-446.

[41] Y. Yao and M. Postolache; Iterative methods for pseudomonotone variational inequalities and fixed-point problems, *J. Optim. Theory Appl.* **155** (2012), 273-287.