# Conditional extragradient algorithms for variational inequalities*

J.Y. Bello Cruz[†] R. Díaz Millán[‡] Hung M. Phan[§]

November 17, 2014

**Abstract**

In this paper, we generalize the classical extragradient algorithm for solving variational inequality problems by utilizing non-null normal vectors of the feasible set. In particular, two conceptual algorithms are proposed and each of them has three different variants which are related to modified extragradient algorithms.

Our analysis contains two main parts: The first part contains two different linesearches, one on the boundary of the feasible set and the other one along the feasible direction. The linesearches allow us to find suitable halfspaces containing the solution set of the problem. By using non-null normal vectors of the feasible set, these linesearches can potentially accelerate the convergence. If all normal vectors are chosen as zero, then some of these variants reduce to several well-known projection methods proposed in the literature for solving the variational inequality problem. The second part consists of three special projection steps, generating three sequences with different interesting features.

Convergence analysis of both conceptual algorithms is established assuming existence of solutions, continuity and a weaker condition than pseudomonotonicity on the operator. Examples, on each variant, show that the modifications proposed here perform better than previous classical variants. These results suggest that our scheme may significantly improve the extragradient variants.

**Keywords:** Armijo-type linesearch, Extragradient algorithm, Projection algorithms.

**Mathematical Subject Classification (2010):** 58E35; 49J40; 65K15.

## 1 Introduction

In this work, we present a conditional extragradient algorithms for solving constrained variational inequality problems. Given an operator $T : \mathrm{dom}(T) \subseteq \mathbb{R}^n \to \mathbb{R}^n$ and a nonempty closed and convex

*Research for this paper was partially supported by CNPq grant 474160/2013–0 and by projects CAPES-MES-CUBA 226/2012 and UNIVERSAL FAPEG/CNPq

[†]Institute of Mathematics and Statistics, Federal University of Goiás, Goiânia, GO. 74.001-970, Brazil. E-mails: `yunier@impa.br` & `yunier@ufg.br`

[‡]Mathematics, Federal Institute of Education, Science and Technology, Goiânia, GO. 74.055-110, Brazil. E-mail: `rdiazmillan@gmail.com`

[§]Mathematics, University of British Columbia, Kelowna, B.C. V1V 1V7, Canada. E-mail: `hung.phan@ubc.ca`

set $C \subset \mathrm{dom}(T)$, the classical variational inequality problem is to

$$\text{find } \; x_* \in C \; \text{ such that } \; \langle T(x_*), x - x_* \rangle \geq 0, \;\; \forall\, x \in C. \tag{1}$$

The solution set of this problem will be denoted by $S_*$. Problem (1) unifies a broad range of approaches in optimization, equilibrium problem, and related problems, and serves as a useful computational framework in very diverse applications. Indeed, this problem has been well studied and has numerous important applications in physics, engineering, economics and optimization theory (see, *e.g.*, [17, 19, 24] and the references therein).

It is well-known that (1) is closely related with the so-called dual formulation problem of the variational inequalities, written as

$$\text{find } \; x_* \in C \; \text{ such that } \; \langle T(x), x - x_* \rangle \geq 0, \quad \forall\, x \in C. \tag{2}$$

The solution set of problem (2) will be denoted by $S_0$. Throughout this paper, our standing assumptions are the following:

(**A1**) $T$ is continuous on $C$.

(**A2**) Problem (1) has at-least one solution and all solutions of (1) solve the dual problem (2).

Note that assumption (**A1**) implies $S_0 \subseteq S_*$ (see Lemma 2.16 below). So, the existence of solutions of (2) implies that of (1). However, the reverse assertion needs generalized monotonicity assumptions. For example, if $T$ is pseudomonotone then $S_* \subseteq S_0$ (see [27, Lemma 1]). With this results, we note that (**A2**) is strictly weaker than pseudomonotonicity of $T$ (see Example 1.1.3 of [26]). Moreover, the assumptions $S_* \neq \varnothing$ and the continuity of $T$ are natural and classical in the literature for most of methods for solving (1). Assumption (**A2**) has been used in various algorithms for solving problem (1) (see [27, 28]).

### 1.1 Extragradient algorithm

In this paper we focus on projection-type algorithms for solving problem (1). Excellent surveys of projection algorithms for solving variational inequality problems can be found in [16, 18, 26]. One of the most studied projection algorithms is the so-called *Extragradient algorithm*, which first appeared in [29]. The projection methods for solving problem (1) necessarily have to perform two projection steps onto the feasible set because the natural extension of the projected gradient method (one projection and $T = \nabla f$) fails in general for monotone operators (see, *e.g.*, [5]). Thus, an extra projection step is required in order to establish the convergence of the projection methods.

Next we describe a general version of the extragradient algorithm together with important strategies for computing the stepsizes (see, *e.g.*, [16, 26]).

**Algorithm 1.1 (Extragradient Algorithm)** Given $\alpha_k, \beta_k, \gamma_k$.

**Step 0 (Initialization):** Take $x^0 \in C$.

**Step 1 (Iterative Step):** Compute

$$z^k = x^k - \beta_k T(x^k), \tag{3a}$$

$$y^k = \alpha_k P_C(z^k) + (1-\alpha_k)x^k, \tag{3b}$$

$$\text{and} \quad x^{k+1} = P_C\big(x^k - \gamma_k T(y^k)\big). \tag{3c}$$

**Step 2 (Stopping Test):** If $x^{k+1} = x^k$, then stop. Otherwise, set $k \leftarrow k+1$ and go to **Step 1**.

We now describe some possible strategies to choose positive stepsizes $\alpha_k$, $\beta_k$ and $\gamma_k$ in (3b), (3a) and (3c), respectively.

(**a**) Constant stepsizes: $\beta_k = \gamma_k$ where $0 < \check{\beta} \le \beta_k \le \hat{\beta} < +\infty$ and $\alpha_k = 1, \quad \forall k \in \mathbb{N}$.

(**b**) Armijo-type linesearch on the boundary of the feasible set: Set $\sigma > 0$, and $\delta \in (0,1)$. For each $k$, take $\alpha_k = 1$ and $\beta_k = \sigma 2^{-j(k)}$ where

$$\begin{cases} j(k) := \min\left\{ j \in \mathbb{N} : \|T(x^k) - T(P_C(z^{k,j}))\| \le \dfrac{\delta}{\sigma 2^{-j}} \|x^k - P_C(z^{k,j})\|^2 \right\}, \\ \text{and} \quad z^{k,j} = x^k - \sigma 2^{-j} T(x^k). \end{cases} \tag{4}$$

In this approach, we take $\gamma_k = \beta_k$, $y^k = P_C(x^k - \beta_k T(x^k))$ and $\gamma_k = \dfrac{\langle T(y^k), x^k - y^k\rangle}{\|T(y^k)\|^2}$, $\forall k \in \mathbb{N}$.

(**c**) Armijo-type linesearch along the feasible direction: Set $\delta \in (0,1)$, $z^k := x^k - \beta_k T(x^k)$ with, $(\beta_k)_{k\in\mathbb{N}} \subset [\check{\beta}, \hat{\beta}]$ such that $0 < \check{\beta} \le \hat{\beta} < +\infty$, and $\alpha_k = 2^{-\ell(k)}$ where

$$\begin{cases} \ell(k) := \min\left\{ \ell \in \mathbb{N} : \langle T(z^{k,\ell}), x^k - P_C(z^k)\rangle \ge \dfrac{\delta}{\beta_k} \|x^k - P_C(z^k)\|^2 \right\}, \\ \text{and} \quad z^{k,\ell} = 2^{-\ell} P_C(z^k) + (1 - 2^{-\ell}) x^k. \end{cases} \tag{5}$$

Then, define $y^k = \alpha_k P_C(z^k) + (1-\alpha_k)x^k$ and $\gamma_k = \frac{\langle T(y^k), x^k - y^k\rangle}{\|T(y^k)\|^2}$, $\forall k \in \mathbb{N}$.

Below follow several comments explaining the differences between these strategies.

It has been proved in [29] that the extragradient algorithm with Strategy (**a**) is globally convergent if $T$ is monotone and Lipschitz continuous on $C$. The main difficulty of this strategy is the necessity of choosing $\beta_k$ in (3a) satisfying $0 < \beta_k \le \beta < 1/L$ where $L$ is the Lipschitz constant of $T$ and when $L$ is not available, the stepsizes have be take sufficiently small to ensures the convergence.

Strategy (**b**) was first studied in [25] under monotonicity and Lipschitz continuity of $T$. The Lipschitz continuity assumption was removed later in [21]. Note that this strategy requires computing the projection onto $C$ inside the inner loop of the Armijo-type linesearch (4). So, the possibility of computing many projections for each iteration $k$ makes Strategy (**b**) inefficient when explicit formula for $P_C$ is not available.

Strategy (**c**) was presented in [22]. This strategy guarantees convergence by assuming only the monotonicity of $T$ and the existence of solutions of (1), and without assuming Lipschitz continuity on $T$. This approach demands only one projection for each outer step $k$. In Strategies (**b**) and (**c**), $T$ is evaluated at least twice and the projection is computed at least twice per iteration. The resulting algorithm is applicable to the whole class of monotone variational inequalities. It has the advantage of not requiring exogenous parameters. Furthermore, Strategies (**b**) and (**c**) occasionally allow long steplength because they both exploit much the information available at each iteration. Extragradient-type algorithms is currently a subject of intense research (see, *e.g.*, [1,4,5,7,14,33,35]). A special modification on Strategy (**c**) was presented in [28] where the monotonicity was replaced by (**A2**). The main difference is that it performs

$$\begin{cases} \ell(k) := \min\left\{ \ell \in \mathbb{N} : \langle T(z^{k,\ell}), x^k - P_C(z^k)\rangle \geq \delta \langle T(x^k), x^k - P_C(z^k)\rangle \right\}, \\ \text{and} \quad z^{k,\ell} = 2^{-\ell} P_C(z^k) + (1-2^{-\ell})x^k, \end{cases} \tag{6}$$

instead of (5).

### 1.2 Proposed schemes

The main part of this work contains two conceptual algorithms, each of them with three variants. Convergence analysis of both conceptual algorithms is established assuming weaker assumptions than previous extragradient algorithms [8, 30].

The approach presented here is closely related to the extragradient algorithm in the above subsection and is based on combining, modifying and generalizing of several ideas contained in various classical extragradient variants. Our scheme was inspired by the conditional subgradient method proposed in [30], and it uses a similar idea of **Algorithm 1.1** over Strategies (**a**), (**b**) and (**c**). The example presented in Section 3 motives our scheme and shows that the variants proposed here may perform better than previous classical variants.

Basically, our two conceptual algorithms contain two parts: The first has two different linesearches: one on the boundary of the feasible set and the other along the feasible direction. These linesearches allow us to find a suitable halfspace separating the current iteration and the solution set. The second has three projection steps allowing several variants with different and interesting features on the generated sequence. In this setting some of the proposed variants on the conceptual algorithms are related to the algorithms presented in [4, 22, 33]. An essential characteristic of the conceptual algorithms is the convergence under very mild assumptions, like continuity of the operator $T$ (see (**A1**)), existence of solutions of (1), and assuming that all such solutions also solve the dual variational inequality (2) (see (**A2**)). We would like to emphasize that this concept is less restrictive than pseudomonotonicity of $T$ and plays a central role in the convergence analysis of our algorithms.

This work is organized as follows: The next section provides and reviews some preliminary results and notations. In Section 3, we present an example where extragradient method with non-null normal vectors can yield some advantages over the original version without normal vectors. In Section 4, we provide the convergence analysis for extragradient method with non-null normal vectors. Section 5 states two different linesearches, which will be used in the Conceptual Algorithms presented in Sections 6 and 7. Each next section (6 and 7) contains three subsection with the convergence analysis of the variants on the two proposed conceptual algorithms. Moreover, we show some comparisons and advantages of our variants with examples in several part of this paper. Finally, Section 8 gives some concluding remarks.

## 2 Preliminary results

In this section, we present some notations, definitions and results needed for the convergence analysis of the proposed methods. We begin reminded some basic notation and definitions used in this paper, which is standard and follows [3]. Throughout this paper, we write $p := q$ to indicate that $p$ is defined to be equal to $q$. The inner product in $\mathbb{R}^n$ is denoted by $\langle \cdot, \cdot \rangle$, and the norm induced by this inner product by $\| \cdot \|$, *i.e.*, $\|x\| := \sqrt{\langle x, x \rangle}$ for all $x \in \mathbb{R}^n$. We write $\mathbb{N}$ for the nonnegative integers $\{0, 1, 2, \ldots\}$ and the extended-real number system is $\bar{\mathbb{R}} := \mathbb{R} \cup \{+\infty\}$. The closed ball centered at $x \in \mathbb{R}^n$ with radius $\rho > 0$ will be denoted by $\mathbb{B}[x, \rho]$, *i.e.*, $\mathbb{B}[x, \rho] := \{y \in \mathbb{R}^n \colon \|y - x\| \leq \rho\}$. The domain of any function $f : \mathbb{R}^n \to \bar{\mathbb{R}}$, denoted by $\mathrm{dom}(f)$, is defined as $\mathrm{dom}(f) := \{x \in \mathbb{R}^n : f(x) < +\infty\}$ and we say $f$ is proper if $\mathrm{dom}(f) \neq \varnothing$. For any set $G$ we denote by cl(G) and $\mathrm{cone}(G)$, respectively the closure topological and the conic hull of $G$. Finally, let $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ be an operator. Then, the domain and the graph of $T$ are $\mathrm{dom}(T) = \{x \in \mathbb{R}^n : T(x) \neq \varnothing\}$ and $\mathrm{Gph}(T) = \{(x, u) \in \mathbb{R}^n \times \mathbb{R}^n : u \in T(x)\}$. We continue with the definition and some results on normal cone.

**Definition 2.1** *Let $C$ be a subset of $\mathbb{R}^n$ and let $u \in C$. A vector $u \in \mathbb{R}^n$ is called a* normal *to $C$ at $x$ if for all $y \in C$, $\langle u, y - x \rangle \leq 0$. The collection of all such normal $u$ is called the* normal cone *of $C$ at $x$ and is denoted by $\mathcal{N}_C(x)$. If $x \notin C$, we define $\mathcal{N}_C(x) = \varnothing$.*

Note that in some special cases, the normal cone can be computed explicitly as showed below.

**Example 2.2** (See [30]) If $C$ is a polyhedral set, *i.e.*, $C = \{x \in \mathbb{R}^n : \langle a_i, x \rangle \leq b_i, \ i = 1, \ldots, m\}$, then

$$\mathcal{N}_C(x) = \{d \in \mathbb{R}^n : d = \sum_{i=1}^{m} \lambda_i a_i, ; \lambda_i(\langle a_i, x \rangle - b_i) = 0, \ \lambda_i \geq 0, \ i = 1, \ldots, m\}$$

for $x \in C$, and $\mathcal{N}_C(x) = \varnothing$ for $x \notin C$.

**Example 2.3** (See [10, Example 2.62]) Let $C$ be a closed convex cone in $\mathbb{R}^n$. Define $C^{\ominus} := \{d \in \mathbb{R}^n : \langle d, x \rangle \leq 0, \ \forall x \in C\}$. Then, $\mathcal{N}_C(x) = C^{\ominus} \cap \{x\}^{\perp} = \{d \in C^{\ominus} : \langle d, x \rangle = 0\}$, if $x \in C$ and $\varnothing$ if $x \notin C$.

**Example 2.4** (See [32, Theorem 6.14]) Let $C = \{x \in X : F(x) \in D\}$ for closed sets $X \subset \mathbb{R}^m$ and $D \in \mathbb{R}^m$ and $\mathcal{C}^1$ mapping $F : \mathbb{R}^n \to \mathbb{R}^m$ such that $F(x) = (f_1(x), f_2(x), \ldots, f_m(x))$. Assume that $x \in C$ satisfies the constraint qualification that:

$$\begin{cases} \text{the unique vector } y = (y_1, \ldots, y_n) \in \mathcal{N}_D(F(x)) \text{ for which} \\ \qquad -\displaystyle\sum_{i=1}^{m} y_i \nabla f_i(x) \in \mathcal{N}_X(x) \text{ is } y = (0, \ldots, 0). \end{cases}$$

Then, $\mathcal{N}_C(x) = \left\{ \displaystyle\sum_{i=1}^{m} y_i \nabla f_i(x) + z : y \in \mathcal{N}_D(F(x)), \ z \in \mathcal{N}_X(x) \right\}$.

The normal cone can be seen as an operator, *i.e.*, $\mathcal{N}_C : C \subset \mathbb{R}^n \rightrightarrows \mathbb{R}^n : x \mapsto \mathcal{N}_C(x)$. Indeed, recall that the indicator function of $C$ is defined by $\delta_C(y) := 0$, if $y \in C$ and $+\infty$, otherwise, and the classical convex subdifferential operator for a proper function $f : \mathbb{R}^n \to \bar{\mathbb{R}}$ is defined by $\partial f : \mathbb{R}^n \rightrightarrows \mathbb{R}^n : x \mapsto \partial f(x) := \{u \in \mathbb{R}^n : f(y) \geq f(x) + \langle u, y - x\rangle, \ \forall\, y \in \mathbb{R}^n\}$. Then, it is well-known that the normal cone operator can be expressed as $\mathcal{N}_C = \partial \delta_C$.

**Example 2.5** (See [31, Theorem 23.7] and [10, Proposition 2.61]) Let $f : \mathbb{R}^n \to \bar{\mathbb{R}}$ be a proper and convex function. Consider $C = \{x \in \mathbb{R}^n : f(x) \leq 0\}$. Suppose that there exists $\overline{x}$ such that $f(\overline{x}) < 0$ *(Slater condition)*. Then,

$$\mathcal{N}_C(x) = \begin{cases} \mathrm{cl}(\mathrm{cone}(\partial f(x))), & \text{if } f(x) = 0; \\ 0, & \text{if } f(x) < 0; \\ \varnothing, & \text{if } f(x) > 0. \end{cases}$$

**Fact 2.6** The normal cone operator for $C$, $\mathcal{N}_C$, is maximal monotone operator and its graph, $\mathrm{Gph}(\mathcal{N}_C)$, is closed, *i.e.*, for all sequences $(x^k, u^k)_{k\in\mathbb{N}} \subset \mathrm{Gph}(\mathcal{N}_C)$ that converges to $(x, u)$, we have $(x, u) \in \mathrm{Gph}(\mathcal{N}_C)$.

*Proof.* See [12, Proposition 4.2.1(ii)]. ■

Recall that the orthogonal projection of $x$ onto $C$, $P_C(x)$, is the unique point in $C$, such that $\|P_C(x)-y\| \leq \|x-y\|$ for all $y \in C$. Now, we state some well-known facts on orthogonal projections.

**Fact 2.7** For all $x, y \in \mathbb{R}^n$ and all $z \in C$ the following hold:

(i) $\|P_C(x) - P_C(y)\|^2 \leq \|x - y\|^2 - \|(x - P_C(x)) - \big(y - P_C(y)\big)\|^2$. (a.k.a. firm nonexpansiveness).

(ii) $\langle x - P_C(x), z - P_C(x)\rangle \leq 0$.

(iii) Let $x \in C$, $y \in \mathbb{R}^n$ and $z = P_C(y)$, then $\langle x - y, x - z\rangle \geq \|x - z\|^2$.

*Proof.* (i) & (ii): See [37, Lemmas 1.1 and 1.2].
(iii): Using (ii), we have $\langle x - y, x - z\rangle = \langle x - z, x - z\rangle + \langle x - z, z - y\rangle \geq \|x - z\|^2$. ■

**Corollary 2.8** *For all $x, p \in \mathbb{R}^n$ and $\alpha > 0$, we have*

$$\frac{x - P_C(x - \alpha p)}{\alpha} \in p + \mathcal{N}_C(P_C(x - \alpha p)).$$

*Proof.* Let $z = x - \alpha p$, then the conclusion follows from $z - P_C(z) \in \mathcal{N}_C(P_C(z))$. ■

Next, we present some lemmas that are useful in the sequel.

**Lemma 2.9** *Let $H \subseteq \mathbb{R}^n$ be a closed halfspace and $C \subseteq \mathbb{R}^n$ such that $H \cap C \neq \varnothing$. Then, for every $x \in C$, we have $P_{H\cap C}(x) = P_{H\cap C}(P_H(x))$.*

*Proof.* If $x \in H$, then $x = P_{H\cap C}(x) = P_{H\cap C}(P_H(x))$. Suppose that $x \notin H$. Fix any $y \in C \cap H$. Since $x \in C$ but $x \notin H$, there exists $\gamma \in [0, 1)$, such that $\tilde{x} = \gamma x+(1-\gamma)y \in C\cap \mathrm{bd}\, H$, where $\mathrm{bd}\, H$ is the hyperplane boundary of $H$. So, $(\tilde{x}-P_H(x))\perp(x-P_H(x))$ and $(P_{H\cap C}(x)-P_H(x))\perp(x-P_H(x))$, then

$$\|\tilde{x} - x\|^2 = \|\tilde{x} - P_H(x)\|^2 + \|x - P_H(x)\|^2, \tag{7}$$

and

$$\|P_{H\cap C}(x) - x\|^2 = \|P_{H\cap C}(x) - P_H(x)\|^2 + \|x - P_H(x)\|^2, \tag{8}$$

respectively. Using (7) and (8), we get

$$\begin{aligned}
&\|y - P_H(x)\|^2 \geq \|\tilde{x} - x\|^2 = \|\tilde{x} - P_H(x)\|^2 + \|P_H(x) - x\|^2 \geq \|\tilde{x} - P_H(x)\|^2. \\
&= \|\tilde{x} - x\|^2 - \|x - P_H(x)\|^2 \geq \|P_{H\cap C}(x) - x\|^2 - \|x - P_H(x)\|^2 = \|P_{H\cap C}(x) - P_H(x)\|^2.
\end{aligned}$$

Thus, $\|y - P_H(x)\| \geq \|P_{H\cap C}(x) - P_H(x)\|$ for all $y \in C \cap H$ and as consequence, $P_{H\cap C}(x) = P_{C\cap H}(P_H(x))$. ■

**Lemma 2.10** *Let $S$ be a nonempty, closed and convex set. Let $x^0, x \in \mathbb{R}^n$. Assume that $x^0 \notin S$ and that $S \subseteq W(x) = \{y \in \mathbb{R}^n : \langle y - x, x^0 - x\rangle \leq 0\}$. Then, $x \in B[\frac{1}{2}(x^0 + \overline{x}), \frac{1}{2}\rho]$, where $\overline{x} = P_S(x^0)$ and $\rho = \operatorname{dist}(x^0, S)$.*

*Proof.* Since $S$ is convex and closed, $\overline{x} = P_S(x^0)$ and $\rho = \operatorname{dist}(x^0, S)$ are well-defined. $S \subseteq W(x)$ implies that $\overline{x} = P_S(x^0) \in W(x)$. Define $v := \frac{1}{2}(x_0 + \overline{x})$ and $r := x^0 - v = \frac{1}{2}(x^0 - \overline{x})$, then $\overline{x} - v = -r$ and $\|r\| = \frac{1}{2}\|x^0 - \overline{x}\| = \frac{1}{2}\rho$. It follows that

$$\begin{aligned}
0 &\geq \langle \overline{x} - x, x^0 - x\rangle = \big\langle \overline{x} - v + v - x, x^0 - v + v - x\big\rangle \\
&= \langle -r + (v - x), r + (v - x)\rangle = \|v - x\|^2 - \|r\|^2.
\end{aligned}$$

The result then follows. ■

We will continue defining the so-called Fejér convergence.

**Definition 2.11** *Let $S$ be a nonempty subset of $\mathbb{R}^n$. A sequence $(x^k)_{k\in\mathbb{N}} \subset \mathbb{R}^n$ is said to be Fejér convergent to $S$ if and only if for all $x \in S$ there exists $k_0 \in \mathbb{N}$ such that $\|x^{k+1} - x\| \leq \|x^k - x\|$ for all $k \geq k_0$.*

This definition was introduced in [11] and has been elaborated further in [2, 23]. The following are useful properties of Fejér sequences.

**Fact 2.12** If $(x^k)_{k\in\mathbb{N}}$ is Fejér convergent to $S$, then the following hold

(i) The sequence $(x^k)_{k\in\mathbb{N}}$ is bounded.

(ii) The sequence $\big(\|x^k - x\|\big)_{k\in\mathbb{N}}$ converges for all $x \in S$.

(iii) If an accumulation point $x_*$ belongs to $S$, then the sequence $(x^k)_{k\in\mathbb{N}}$ converges to $x_*$.

*Proof.* (i) and (ii): See [3, Proposition 5.4]. (iii): See [3, Theorem 5.5]. ■

We recall the following well-known characterization of $S_*$ which will be used repeatedly.

**Fact 2.13** The following statements are equivalent:

(i) $x \in S_*$.

(ii) $-T(x) \in \mathcal{N}_C(x)$.

(iii) There exists $\beta > 0$ such that $x = P_C(x - \beta T(x))$.

*Proof.* See Proposition 1.5.8 of [16]. ∎

**Proposition 2.14** *Given $T : \mathrm{dom}(T) \subseteq \mathbb{R}^n \to \mathbb{R}^n$ and $\alpha > 0$. If $x = P_C(x - \alpha(T(x) + \alpha u))$, with $u \in \mathcal{N}_C(x)$, then $x \in S_*$, or equivalently, $x = P_C(x - \beta T(x))$ for all $\beta > 0$.*

*Proof.* It follows from Corollary 2.8 that $0 \in T(x) + \alpha u + \mathcal{N}_C(x)$, which is implies that $-T(x) \in \mathcal{N}_C(x)$. The conclusion is now immediate from Fact 2.13. ∎

**Remark 2.15** It is quite easy to see that the reverse of Proposition 2.14 is not true in general.

**Lemma 2.16** *If $T : \mathrm{dom}(T) \subseteq \mathbb{R}^n \to \mathbb{R}^n$ is continuous, then $S_0 \subseteq S_*$.*

*Proof.* Assume that $x \in \{x \in C : \langle T(y), y - x\rangle \geq 0 \; \forall\, y \in C\}$. Take $y(\alpha) = (1-\alpha)x + \alpha y$, $y \in C$ with $\alpha \in (0, 1)$. Since $y(\alpha) \in C$ and hence

$$\begin{aligned} 0 \leq \langle T(y(\alpha)), y(\alpha) - x\rangle &= \langle T((1-\alpha)x + \alpha y), (1-\alpha)x + \alpha y - x\rangle \\ &= \alpha\langle T((1-\alpha)x + \alpha y), y - x\rangle. \end{aligned}$$

Dividing by $\alpha > 0$, we get $0 \leq \langle T((1-\alpha)x + \alpha y), y - x\rangle$, and taking limits, when $\alpha$ goes to 0, we obtain from the continuity of $T$ that $\langle T(x), y - x\rangle \geq 0$, for all $y \in C$, *i.e.*, $x \in S_*$. ∎

**Lemma 2.17** *For any $(z, v) \in \mathrm{Gph}(\mathcal{N}_C)$ define $H(z, v) := \big\{y \in \mathbb{R}^n : \langle T(z) + v, y - z\rangle \leq 0\big\}$. Then, $S_* = S_0 \subseteq H(z, v)$.*

*Proof.* $S_* = S_0$ by Assumption (**A2**) and Lemma 2.16. Take $x_* \in S_0$, then $\langle T(z), x_* - z\rangle \leq 0$ for all $z \in C$. Since $(z, v) \in \mathrm{Gph}(\mathcal{N}_C)$, we have $\langle v, x_* - z\rangle \leq 0$. Summing up these inequalities, we get $\langle T(z) + v, x_* - z\rangle \leq 0$. Then, $x_* \in H(z, v)$. ∎

Now follow a direct consequence of the definition of $S_0$ the solution set of the dual problem (2), which from Lemma 2.17 coincides with $S_*$.

**Lemma 2.18** *If $T : \mathrm{dom}(T) \subseteq \mathbb{R}^n \to \mathbb{R}^n$ is continuous and Assumption (**A2**) holds, then $S_*$ is a closed and convex set.*

## 3 Motivation: An example

In this section, we present an elementary instance of problem (1), in which normal vectors of the feasible set is beneficial.

**Example 3.1** Let $B := (b_1, b_2) \in \mathbb{R}^2$, recall that the rotation with angle $\gamma \in (0, \pi/2)$ around $B$, given by

$$\mathcal{R}_{\gamma,B} : \mathbb{R}^2 \to \mathbb{R}^2 : x \mapsto \begin{bmatrix} \cos\gamma & \sin\gamma \\ -\sin\gamma & \cos\gamma \end{bmatrix}(x - B) + B,$$

is monotone and Lipschitz continuous with constant 1. We consider problem (1) in $\mathbb{R}^2$ with the operator $T := \mathcal{R}_{\frac{\pi}{2},B} - \mathrm{Id}$, where $B := (\frac{1}{2}, 1)$, and the feasible set

$$C := \left\{(x_1, x_2) \in \mathbb{R}^2 : \quad x_1^2 + x_2^2 \leq 1,\ x_1 \leq 0,\ x_2 \geq 0\right\}.$$

The unique solution of this problem is $\overline{x} \approx (-0.935, 0.355)$ (see Figure 1).

We set a starting point $x^0 = y^0 = (0, 1)$ and a scalar $\beta = 0.4$. Now, we observe the performance of the following two algorithms:

**(a1)** **Algorithm 1.1** with Strategy (**a**) (constant stepsizes, *i.e.*, $\forall k \in \mathbb{N}$, $\gamma_k = \beta_k = \beta$ and $\alpha_k = 1$), which generates the following sequences as follows:

$$\forall k \in \mathbb{N} : \quad \begin{cases} z^k = P_C(y^k - \beta T(y^k)), \\ y^{k+1} = P_C(y^k - \beta T(z^k)). \end{cases}$$

**(a2)** A modified **Algorithm 1.1** with Strategy (**a**) involving *unit* normal vectors in both projection steps, which generates the following sequences as follows:

$$\forall k \in \mathbb{N} : \quad \begin{cases} z^k = P_C(x^k - \beta(T(x^k) + u^k)), \quad \text{where} \quad u^k \in \mathcal{N}_C(x^k),\ \|u^k\| = 1, \\ x^{k+1} = P_C(x^k - \beta(T(z^k) + v^k)), \quad \text{where} \quad v^k \in \mathcal{N}_C(z^k),\ \|v^k\| = 1. \end{cases}$$

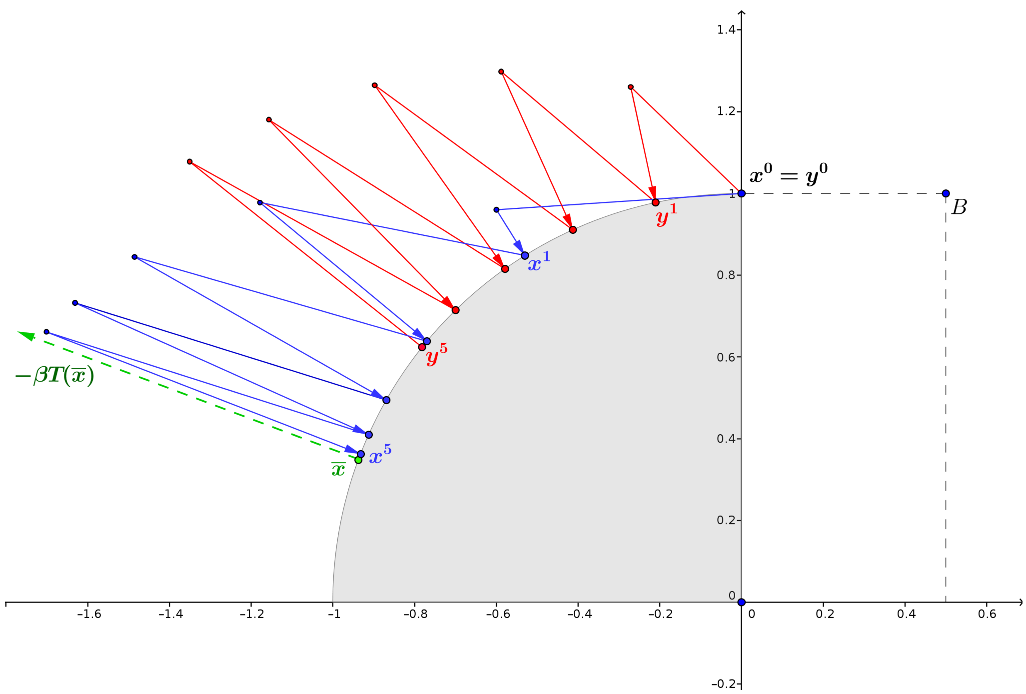


Figure 1: Advantage by using non-null normal vectors.

The first *five* iterations of each algorithm are showed in Figure 1: Recall that $(y^k)_{k\in\mathbb{N}}$ (red) and $(x^k)_{k\in\mathbb{N}}$ (blue) are the sequences generated by the algorithms presented above in (**a1**) and (**a2**), respectively. The comparison suggests that normal vectors of the feasible set can help the extragradient algorithm to improve considerably the convergence speed.

## 4 The extragradient algorithm with normal vectors

Inspired by the previous section, we investigate the extragradient algorithm with constant stepsizes involving normal vectors of the feasible set. In this section we assume that $T$ is Lipschitz with constant $L$ and (**A2**) holds. The algorithm proposed in this section is related with **Algorithm 1.1** over Strategy (**a**) (constant stepsizes). It is defined as:

**Algorithm 4.1 (Conditional Extragradient Algorithm)** Take $(\beta_k)_{k\in\mathbb{N}} \subset [\check{\beta}, \hat{\beta}]$ such that $0 < \check{\beta} \le \hat{\beta} < 1/(L+1)$ and $\delta \in (0,1)$.

**Step 1 (Initialization):** Take $x^0 \in C$ and set $k \leftarrow 0$.

**Step 1 (Stopping Test 1):** If $x^k = P_C(x^k - \beta_k T(x^k))$, *i.e.*, $x^k \in S_*$, then stop. Otherwise:

**Step 2 (First Projection):** Take $u^k \in \mathcal{N}_C(x^k)$ such that

$$\|u^k\| \le \delta \|x^k - P_C(x^k - \beta_k(T(x^k) + u^k))\|. \tag{9}$$

$$z^k = P_C(x^k - \beta_k(T(x^k) + u^k)). \tag{10}$$

**Step 3 (Second Projection):** Take $v^k \in \mathcal{N}_C(z^k)$ such that

$$\|v^k - u^k\| \le \|x^k - z^k\|. \tag{11}$$

$$x^{k+1} = P_C(x^k - \beta_k(T(z^k) + v^k)). \tag{12}$$

**Step 4 (Stopping Test 2):** If $x^{k+1} = x^k$ then stop. Otherwise, set $k \leftarrow k+1$ and go to **Step 1**.

**Proposition 4.2** *The following hold:*

(i) *If* **Algorithm 4.1** *stops then* $x^k \in S_*$.

(ii) **Algorithm 4.1** *is well-defined.*

*Proof.* (i): If the algorithm stops at **Step 1** then Proposition 2.13 garantes the optimality of $x^k$. If $x^k = P_C(x^k - \beta_k(T(z^k) + v^k))$ then $x^k \in S_*$ by Proposition 2.14.

(ii): It is sufficient to prove that if **Step 1** is not satisfied, *i.e.*,

$$\|x^k - P_C(x^k - \beta_k T(x^k))\| > 0. \tag{13}$$

Then, **Step 2** and **Step 3** are attainable.

**Step 2** *is attainable:* Suppose that (9) does not hold for every $\alpha u^k \in \mathcal{N}_C(x^k)$ with $\alpha > 0$, *i.e.*,

$$\|\alpha u^k\| > \delta \|x^k - P_C(x^k - \beta_k(T(x^k) + \alpha u^k))\| \ge 0.$$

Taking limit when $\alpha$ goes to 0, we get $\|x^k - P_C(x^k - \beta_k T(x^k))\| = 0$, which contradicts (13).

**Step 3** *is attainable:* Suppose that (11) does not hold for every $\alpha v^k \in \mathcal{N}_C(z^k)$ with $\alpha > 0$, *i.e.*,

$$\|\alpha v^k - u^k\| > \|x^k - z^k\|,$$

where $z^k = P_C(x^k - \beta_k(T(x^k) + u^k))$ as (10) and $u^k \in \mathcal{N}_C(x^k)$ satisfying (9). Letting $\alpha$ goes to 0 and using (9), we get $\|x^k - z^k\| \leq \|u^k\| \leq \delta\|x^k - z^k\|$. So, $x^k = z^k$. Then, Proposition 2.14 implies a contradiction to (13). ■

In the rest of this section, we investigate the remaining case that **Algorithm 4.1** does not stop and generates an infinite sequence $(x^k)_{k\in\mathbb{N}}$.

**Lemma 4.3** *Suppose that $T$ is Lipschitz continuous with constant $L$. Let $x_* \in S_*$. Then, for every $k \in \mathbb{N}$, it holds that*

$$\|x^{k+1} - x_*\|^2 \leq \|x^k - x_*\|^2 - (1 - \beta_k^2(L+1)^2)\|z^k - x^k\|^2.$$

*Proof.* Define $w^k = x^k - \beta_k(T(z^k) + v^k)$ with $v^k \in \mathcal{N}_C(z^k)$ as **Step 3**. Then, using (12) and applying Fact 2.7(i), with $x = w^k$ and $y = x^*$, we get

$$\begin{aligned}\|x^{k+1} - x_*\|^2 &\leq \|w^k - x_*\|^2 - \|w^k - P_C(w^k)\|^2 \\ &\leq \|x^k - x_* - \beta_k(T(z^k) + v^k)\|^2 - \|x^k - x^{k+1} - \beta_k(T(z^k) + v^k)\|^2 \\ &= \|x^k - x_*\|^2 - \|x^k - x^{k+1}\|^2 + 2\beta_k\langle T(z^k) + v^k, x_* - x^{k+1}\rangle. \end{aligned} \tag{14}$$

Since $v^k \in \mathcal{N}_C(z^k)$ and (**A2**), we have

$$\begin{aligned}\langle T(z^k) + v^k, x_* - x^{k+1}\rangle &= \langle T(z^k) + v^k, z^k - x^{k+1}\rangle + \langle T(z^k) + v^k, x_* - z^k\rangle \\ &\leq \langle T(z^k) + v^k, z^k - x^{k+1}\rangle + \langle T(z^k), x_* - z^k\rangle \\ &\leq \langle T(z^k) + v^k, z^k - x^{k+1}\rangle.\end{aligned}$$

Substituting the above inequality in (14), we get

$$\begin{aligned}&\|x^{k+1} - x_*\|^2 \leq \|x^k - x_*\|^2 - \|x^k - x^{k+1}\|^2 - 2\beta_k\langle T(z^k) + v^k, x^{k+1} - z^k\rangle \\ &= \|x^k - x_*\|^2 - \|x^k - z^k\|^2 - \|z^k - x^{k+1}\|^2 + 2\langle x^k - \beta_k(T(z^k) + v^k) - z^k, x^{k+1} - z^k\rangle.\end{aligned} \tag{15}$$

Define $\overline{x}^k = x^k - \beta_k(T(x^k) + u^k)$ with $u^k \in \mathcal{N}_C(x^k)$ as **Step 2** and recall that $z^k = P_C(\bar{x}^k)$ and $x^{k+1} = P_C(w^k) = P_C(x^k - \beta_k(T(z^k) + v^k))$, we have

$$\begin{aligned}&2\langle x^k - \beta_k(T(z^k) + v^k) - z^k, x^{k+1} - z^k\rangle = 2\langle w^k - P_C(\overline{x}^k), P_C(w^k) - P_C(\overline{x}^k)\rangle \\ &= 2\langle \overline{x}^k - P_C(\overline{x}^k), P_C(w^k) - P_C(\overline{x}^k)\rangle + 2\langle w^k - \overline{x}^k, P_C(w^k) - P_C(\overline{x}^k)\rangle \\ &\leq 2\langle w^k - \overline{x}^k, P_C(w^k) - P_C(\overline{x}^k)\rangle \\ &= 2\langle w^k - \overline{x}^k, x^{k+1} - z^k\rangle = 2\beta_k\langle (T(x^k) + u^k) - (T(z^k) + v^k), x^{k+1} - z^k\rangle \\ &\leq 2\beta_k\left(\|T(z^k) - T(x^k)\| + \|v^k - u^k\|\right)\|x^{k+1} - z^k\| \\ &\leq 2\beta_k(L+1)\|z^k - x^k\|\|x^{k+1} - z^k\| \leq \beta_k^2(L+1)^2\|z^k - x^k\|^2 + \|x^{k+1} - z^k\|^2,\end{aligned} \tag{16}$$

using Fact 2.7(ii), with $x = x^k - \beta_k(T(x^k) + u^k)$ and $z = x^{k+1}$, in the first inequality, the Cauchy-Schwarz inequality in the second one and the Lipschitz continuity of $T$ and (11) in the third one. Therefore, (16) together with (15) proves the lemma. ■

**Corollary 4.4** *The sequence $(x^k)_{k\in\mathbb{N}}$ is Fejér convergent to $S_*$ and* $\lim_{k\to\infty}\|z^k - x^k\| = 0$.

*Proof.* It follows directly from Lemma 4.3 and the fact $\beta_k \leq \hat{\beta} < 1/(L+1)$ for all $k \in \mathbb{N}$ that.

$$\|x^{k+1} - x_*\|^2 \leq \|x^k - x_*\|^2 - (1 - \hat{\beta}^2 L^2)\|z^k - x^k\|^2 \leq \|x^k - x_*\|^2.$$

So, $(x^k)_{k\in\mathbb{N}}$ is Fejér convergent to $S_*$ and Fact 2.12(ii) together with the above inequality imply $\lim_{k\to\infty} \|z^k - x^k\| = 0$. ■

**Proposition 4.5** *The sequence $(x^k)_{k\in\mathbb{N}}$ converges to a point in $S_*$.*

*Proof.* The sequence $(x^k)_{k\in\mathbb{N}}$ is bounded by Lemma 4.3 and Fact 2.12(i). Let $\tilde{x}$ be an accumulation point of a subsequence $(x^{i_k})_{k\in\mathbb{N}}$. By Corollary 4.4, $\tilde{x}$ is also an accumulation point of $(z^{i_k})_{k\in\mathbb{N}}$. Without loss of generality, we suppose that the corresponding parameters $(\beta_{i_k})_{k\in\mathbb{N}}$ and $(u^{i_k})_{k\in\mathbb{N}}$ converge to $\tilde{\beta}$ and $\tilde{u}$, respectively. Since $z^k = P_C(x^k - \beta_k(T(x^k) + u^k))$, taking the limit along the subsequence $(i_k)_{k\in\mathbb{N}}$, we obtain $\tilde{x} = P_C(\tilde{x} - \tilde{\beta}(T(\tilde{x}) + \tilde{u}))$. Thus, Fact 2.6 and Proposition 2.14 imply $\tilde{x} \in S_*$. ■

Next, we examine the performance of **Algorithm 4.1** for the variational inequality in Example 3.1 with and without normal vectors.

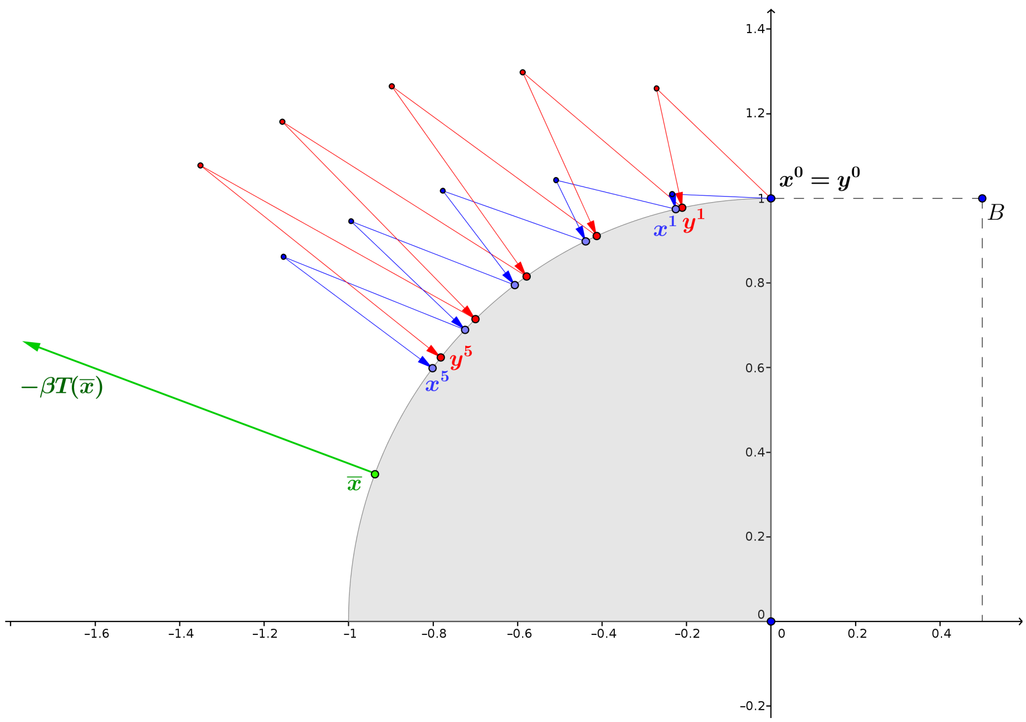


Figure 2: Conditional extragradient method with and without normal vectors.

Figure 2 shows the first *five* elements of sequences $(y^k)_{k\in\mathbb{N}}$ (generated without normal vectors) and $(x^k)_{k\in\mathbb{N}}$ (generated with non-null normal vectors). In practice, using normal vectors with large magnitude can potentially produce significant difference.

## 5 Two linesearches

In this section we present two linesearches, which will be used in our conceptual algorithms in the next two sections. These linesearches are related to Strategies (**b**) and (**c**) in **Algorithm 1.1**. The main difference here is that the new ones utilize normal vectors to the feasible sets.

We begin introducing a linesearch on the boundary of the feasible set, which is closely related with the linesearch in Strategy (**b**) of the extragradient algorithm (see **Algorithm 1.1**). Indeed, if we set the vectors $u \in \mathcal{N}_C(x)$ and $v_\alpha \in \mathcal{N}_C(z_\alpha)$ ($\alpha \in \{\sigma, \sigma\theta, \sigma\theta^2, \ldots, \}$) as the null vector in **Linesearch B** below, then it becomes Strategy (**b**) presented in (4).

**Linesearch B** (Linesearch on the boundary)

**Input:** $(x, u, \sigma, \delta, M)$. Where $x \in C$, $u \in \mathcal{N}_C(x)$, $\sigma > 0$, $\delta \in (0,1)$, and $M > 0$.
Set $\alpha = \sigma$ and $\theta \in (0,1)$ and choose $u \in \mathcal{N}_C(x)$. Denote $z_\alpha = P_C(x - \alpha(T(x) + \alpha u))$ and choose $v_\alpha \in \mathcal{N}_C(z_\alpha)$ with $\|v_\alpha\| \leq M$.

**While** $\alpha\|T(z_\alpha) - T(x) + \alpha v_\alpha - \alpha u\| > \delta\|z_\alpha - x\|$ **do**
$\alpha \leftarrow \theta\alpha$ and choose any $v_\alpha \in \mathcal{N}_C(z_\alpha)$ with $\|v_\alpha\| \leq M$.
**End While**

**Output:** $(\alpha, z_\alpha, v_\alpha)$.

We now show that **Linesearch B** is well-defined assuming only (**A1**), *i.e.*, continuity of $T$ is sufficient to prove the well-definition of **Linesearch B**.

**Lemma 5.1** *If $x \in C$ and $x \notin S_*$, then* **Linesearch B** *stops after finitely many steps.*

*Proof.* Suppose on the contrary that **Linesearch B** does not stop for all $\alpha \in \mathcal{P} := \{\sigma, \sigma\theta, \sigma\theta^2, \ldots\}$ and the chosen vectors

$$v_\alpha \in \mathcal{N}_C(z_\alpha), \quad \|v_\alpha\| \leq M, \tag{17a}$$
$$z_\alpha = P_C(x - \alpha(T(x) + \alpha u)). \tag{17b}$$

We have

$$\alpha\|T(z_\alpha) - T(x) + \alpha v_\alpha - \alpha u\| > \delta\|z_\alpha - x\|. \tag{18}$$

On the one hand, dividing both sides of (18) by $\alpha > 0$ and letting $\alpha$ goes to 0, we obtain by the boundedness of $(v_\alpha)_{\alpha\in\mathcal{P}}$, presented in (17a), and the continuity of $T$ that

$$0 = \liminf_{\alpha\to 0} \|T(z_\alpha) - T(x) + \alpha v_\alpha - \alpha u\| \geq \liminf_{\alpha\to 0} \frac{\|x - z_\alpha\|}{\alpha} \geq 0.$$

Thus, by (17b),

$$\liminf_{\alpha\to 0} \frac{\|x - P_C(x - \alpha(T(x) + \alpha u))\|}{\alpha} = 0. \tag{19}$$

On the other hand, Corollary 2.8 implies

$$\frac{x - P_C(x - \alpha(T(x) + \alpha u))}{\alpha} \in T(x) + \alpha u + \mathcal{N}_C(P_C(x - \alpha(T(x) + \alpha u))).$$

From (19), the continuity of the projection and the closedness of $\mathrm{Gph}(\mathcal{N}_C)$ imply $0 \in T(x)+\mathcal{N}_C(x)$, which is a contradiction since $x \notin S_*$. ■

As mentioned before, the disadvantage of **Linesearch B** is the necessity to compute the projection onto the feasible set inside the inner loop to find the stepsize $\alpha$. To overcome this, we present a linesearch along the feasible direction below, which is closely related to Strategy (**c**) of **Algorithm 1.1**. Indeed, if we set the vectors $u \in \mathcal{N}_C(x)$ and $v_\alpha \in \mathcal{N}_C(z_\alpha)$ ($\alpha \in \{1, \theta, \theta^2, \dots, \}$) as the null vector in **Linesearch F** below, we obtain Strategy (**c**) presented in (6). Furthermore, if we only choose $u = 0$ the projection step is done outside the procedure to find the stepsize $\alpha$.

**Linesearch F** (Linesearch along the feasible direction)

**Input:** $(x, u, \beta, \delta, M)$. Where $x \in C$, $u \in \mathcal{N}_C(x)$, $\beta > 0$, $\delta \in (0,1)$, and $M > 0$.
Set $\alpha \leftarrow 1$ and $\theta \in (0,1)$. Define $z_\alpha = P_C(x - \beta(T(x) + \alpha u))$ and choose $u \in \mathcal{N}_C(x)$, $v_1 \in \mathcal{N}_C(z_1)$ with $\|v_1\| \le M$.
**While** $\langle T\big(\alpha z_\alpha + (1-\alpha)x\big) + v_\alpha, x - z_\alpha\rangle < \delta\langle T(x) + \alpha u, x - z_\alpha\rangle$ **do**
$\alpha \leftarrow \theta\alpha$ and choose any $v_\alpha \in \mathcal{N}_C(\alpha z_\alpha + (1-\alpha)x)$ with $\|v_\alpha\| \le M$.
**End While**
**Output:** $(\alpha, z_\alpha, v_\alpha)$.

In the following we prove that **Linesearch F** is well-defined assuming only (**A1**), *i.e.*, continuity of $T$ is sufficient to prove the well-definition of **Linesearch F**.

**Lemma 5.2** *If $x \in C$ and $x \notin S_*$, then* **Linesearch F** *stops after finitely many steps.*

*Proof.* Suppose on the contrary that **Linesearch F** does not stop for all $\alpha \in \mathcal{P} := \{1, \theta, \theta^2, \dots, \}$ and the chosen

$$v_\alpha \in \mathcal{N}_C\big(\alpha z_\alpha + (1-\alpha)x\big), \quad \|v_\alpha\| \le M, \tag{20a}$$

$$z_\alpha = P_C\big(x - \beta(T(x) + \alpha u)\big). \tag{20b}$$

We have

$$\langle T(\alpha z_\alpha + (1-\alpha)x) + v_\alpha, x - z_\alpha\rangle < \delta\langle T(x) + \alpha u, x - z_\alpha\rangle. \tag{21}$$

By (20a) the sequence $(v_\alpha)_{\alpha\in\mathcal{P}}$ is bounded, thus, without loss of generality, we can assume that it converges to some $v_0 \in \mathcal{N}_C(x)$ (by Fact 2.6). The continuity of the projection operator and (20b) imply that $(z_\alpha)_{\alpha\in\mathcal{P}}$ converges to $z_0 = P_C(x - \beta T(x))$. Taking the limit in (21), when $\alpha$ goes to 0, we get $\langle T(x) + v_0, x - z_0\rangle \le \delta\langle T(x), x - z_0\rangle$. Noticing that $v_0 \in \mathcal{N}_C(x)$, we have

$$0 \ge (1-\delta)\langle T(x), x - z_0\rangle + \langle v_0, x - z_0\rangle \ge (1-\delta)\langle T(x), x - z_0\rangle \ge \frac{(1-\delta)}{\beta}\|x - z_0\|^2.$$

Then, it follows that $x = z_0 = P_C(x - \beta T(x))$, *i.e.*, $x \in S_*$, a contradiction. ■

# 6 Conceptual algorithm with Linesearch B

In this section, we study the **Conceptual Algorithm B** in which **Linesearch B** is used to obtain the stepsizes. From now on, we assume that (**A1**) and (**A2**) hold. The next conceptual algorithm is related with **Algorithm 1.1** over Strategy (**b**) when non-null normal vectors are using in the steps (3a)-(3c).

**Conceptual Algorithm B** Given $\sigma > 0$, $\delta \in (0,1)$, and $M > 0$.

**Step 0 (Initialization):** Take $x^0 \in C$ and set $k \leftarrow 0$.

**Step 1 (Stopping Test 1):** If $x^k = P_C(x^k - T(x^k))$, *i.e.*, $x^k \in S_*$, then stop. Otherwise,

**Step 2 (Linesearch B):** Take $u^k \in \mathcal{N}_C(x^k)$ with $\|u^k\| \le M$ and set

$$(\alpha_k, z^k, v^k) = \textbf{Linesearch B } (x^k, u^k, \sigma, \delta, M),$$

*i.e.*, $(\alpha_k, z^k, v^k)$ satisfy

$$\begin{cases} v^k \in \mathcal{N}_C(z^k) \;\text{ with }\; \|v^k\| \le M \;; \quad \alpha_k \le \sigma; \\ z^k = P_C(x^k - \alpha_k(T(x^k) + \alpha_k u^k)); \\ \alpha_k \|T(z^k) - T(x^k) + \alpha_k(v^k - u^k)\| \le \delta \|z^k - x^k\|. \end{cases} \tag{22}$$

**Step 3 (Projection Step):** Set

$$\overline{v}^k := \alpha_k v^k \tag{23a}$$

$$\text{and} \quad x^{k+1} := \mathcal{F}(x^k). \tag{23b}$$

**Step 4 (Stopping Test 2):** If $x^{k+1} = x^k$ then stop. Otherwise, set $k \leftarrow k+1$ and go to **Step 1**.

We consider three variants of this algorithm. Their main difference lies in the computation (23b):

$$\mathcal{F}_{\mathrm{B.1}}(x^k) = P_C\big(P_{H(z^k,\overline{v}^k)}(x^k)\big); \qquad (\textbf{Variant B.1}) \tag{24}$$

$$\mathcal{F}_{\mathrm{B.2}}(x^k) = P_{C\cap H(z^k,\overline{v}^k)}(x^k); \qquad (\textbf{Variant B.2}) \tag{25}$$

$$\mathcal{F}_{\mathrm{B.3}}(x^k) = P_{C\cap H(z^k,\overline{v}^k)\cap W(x^k)}(x^0), \qquad (\textbf{Variant B.3}) \tag{26}$$

where

$$H(z,v) := \big\{ y \in \mathbb{R}^n : \langle T(z) + v, y - z\rangle \le 0 \big\}, \tag{27a}$$

$$\text{and} \qquad W(x) := \big\{ y \in \mathbb{R}^n : \langle y - x, x^0 - x\rangle \le 0 \big\}. \tag{27b}$$

These halfspaces have been widely used in the literature, *e.g.*, [6, 8, 34] and the references therein. Our goal is analyze the convergence of these variants. First, we start by showing that this conceptual algorithm is well-defined.

**Proposition 6.1** *Assume that* (23b) *is well-defined whenever* $x^k$ *is available. Then,* **Conceptual Algorithm B** *is also well-defined.*

*Proof.* If **Step 1** is not satisfied, then **Step 2** is guaranteed by Lemma 5.1. Thus, the entire algorithm is well-defined. ■

Next, we present a useful proposition for establishing the well-definition of the projection step (23b) over each variant.

**Proposition 6.2** *$x^k \in S_*$ if and only if $x^k \in H(z^k, \overline{v}^k)$, where $z^k$ and $\overline{v}^k$ are given respectively by* (22) *and* (23a)*.*

*Proof.* Suppose that $x^k \notin S_*$. Define $\bar{u}^k = \alpha_k u^k \in \mathcal{N}_C(x^k)$ and $w^k = x^k - \alpha_k(T(x^k) + \bar{u}^k)$. Then,

$$\begin{aligned} \alpha_k\langle T(z^k) + \overline{v}^k, x^k - z^k\rangle &= \alpha_k\langle T(z^k) - T(x^k) + \overline{v}^k - \bar{u}^k, x^k - z^k\rangle + \alpha_k\langle T(x^k) + \bar{u}^k, x^k - z^k\rangle \\ &= \alpha_k\langle T(z^k) - T(x^k) + \overline{v}^k - \bar{u}^k, x^k - z^k\rangle + \langle x^k - w^k, x^k - z^k\rangle \\ &\geq -\alpha_k\|T(z^k) - T(x^k) + \overline{v}^k - \bar{u}^k\| \cdot \|x^k - z^k\| + \|x^k - z^k\|^2 \\ &\geq -\delta\|x^k - z^k\|^2 + \|x^k - z^k\|^2 = (1-\delta)\|x^k - z^k\|^2 > 0, \end{aligned} \tag{28}$$

where we have used **Linesearch B** and Fact 2.7(iii) in the second inequality. Thus, $x^k \notin H(z^k, \overline{v}^k)$. Conversely, if $x^k \in S_*$ using Lemma 2.17, $x^k \in H(z^k, \overline{v}^k)$. ■

Now, we note a useful algebraic property on the sequence generated by **Conceptual Algorithm B**, which is a direct consequence of **Linesearch B**. Let $(x^k)_{k\in\mathbb{N}}$, $(z^k)_{k\in\mathbb{N}}$ and $(\alpha_k)_{k\in\mathbb{N}}$ be sequences generated by **Conceptual Algorithm B**, using (28), we get

$$\forall k \in \mathbb{N}: \quad \langle T(z^k) + \overline{v}^k, x^k - z^k\rangle \geq \frac{(1-\delta)}{\alpha_k}\delta\|x^k - z^k\|^2. \tag{29}$$

## 6.1 Convergence analysis of Variant B.1

In this section, all results are for **Variant B.1**, which is summarized below.

**Variant B.1** $x^{k+1} = \mathcal{F}_{\mathrm{B.1}}(x^k) = P_C(P_{H(z^k,\overline{v}^k)}(x^k)) = P_C\Big(x^k - \dfrac{\langle T(z^k) + \overline{v}^k, x^k - z^k\rangle}{\|T(z^k) + \overline{v}^k\|^2}\big(T(z^k) + \overline{v}^k\big)\Big)$
where

$$H(z^k, \overline{v}^k) = \big\{y \in \mathbb{R}^n : \langle T(z^k) + \overline{v}^k, y - z^k\rangle \leq 0\big\},$$

with $z^k$ and $\overline{v}^k$ are respectively given by (22) and (23a).

**Proposition 6.3** *If $x^{k+1} = x^k$, if and only if $x^k \in S_*$ and* **Variant B.1** *stops.*

*Proof.* If $x^{k+1} = P_C\big(P_{H(z^k,\overline{v}^k)}(x^k)\big) = x^k$, then Fact 2.7(ii) implies

$$\langle P_{H(z^k,\overline{v}^k)}(x^k) - x^k, z - x^k\rangle = \langle P_{H(z^k,\overline{v}^k)}(x^k) - x^{k+1}, z - x^{k+1}\rangle \leq 0, \tag{30}$$

for all $z \in C$. Using again Fact 2.7(ii),

$$\langle P_{H(z^k,\overline{v}^k)}(x^k) - x^k, P_{H(z^k,\overline{v}^k)}(x^k) - z\rangle \leq 0, \tag{31}$$

for all $z \in H(z^k, \overline{v}^k)$. Note that $C \cap H(z^k, \overline{v}^k) \neq \varnothing$, because $z^k$ belongs to it. So, for any $z \in C \cap H(z^k, \overline{v}^k)$, adding up (30) and (31) yields $\|x^k - P_{H(z^k,\overline{v}^k)}(x^k)\|^2 = 0$. Hence, $x^k = P_{H(z^k,\overline{v}^k)}(x^k)$, *i.e.*, $x^k \in H(z^k, \overline{v}^k)$. Thus, we have $x^k \in S_*$ by Proposition 6.2. Conversely, if $x^k \in S_*$, Proposition 6.2 implies $x^k \in H(z^k, \bar{v}^k)$ and together with (24), we get $x^k = x^{k+1}$. ■

As consequence of Proposition 6.3, we can assume that **Variant B.1** does not stop. Note that by Lemma 2.17, $H(z^k, \overline{v}^k)$ is nonempty for all $k$. So, the projection step (24) is well-defined. Thus, **Variant B.1** generates an infinite sequence $(x^k)_{k\in\mathbb{N}}$ such that $x^k \notin S_*$ for all $k \in \mathbb{N}$.

**Proposition 6.4** *The following hold:*

(i) *The sequence* $(x^k)_{k\in\mathbb{N}}$ *is Fejér convergent to* $S_*$.

(ii) *The sequence* $(x^k)_{k\in\mathbb{N}}$ *is bounded.*

(iii) $\lim_{k\to\infty}\langle T(z^k)+\overline{v}^k, x^k - z^k\rangle = 0$.

*Proof.* (i): Take $x_* \in S_*$. Note that, by definition $(z^k, \overline{v}^k) \in \mathrm{Gph}(\mathcal{N}_C)$. Using (24), Fact 2.7(i) and Lemma 2.17, we have

$$\begin{aligned}\|x^{k+1}-x_*\|^2 &= \|P_C(P_{H(z^k,\overline{v}^k)}(x^k)) - P_C(P_{H(z^k,\overline{v}^k)}(x_*))\|^2 \le \|P_{H(z^k,\overline{v}^k)}(x^k) - P_{H(z^k,\overline{v}^k)}(x_*)\|^2 \\ &\le \|x^k - x_*\|^2 - \|P_{H(z^k,\overline{v}^k)}(x^k) - x^k\|^2. \end{aligned} \tag{32}$$

So, $\|x^{k+1}-x_*\| \le \|x^k - x_*\|$.

(ii): Follows from (i) and Fact 2.12(i).

(iii): Take $x_* \in S_*$. Since $P_{H(z^k,\overline{v}^k)}(x^k) = x^k - \dfrac{\langle T(z^k)+\overline{v}^k, x^k-z^k\rangle}{\|T(z^k)+\overline{v}^k\|^2}\big(T(z^k)+\overline{v}^k\big)$, and combining it with (32), yields

$$\begin{aligned}\|x^{k+1}-x_*\|^2 &\le \|x^k-x_*\|^2 - \left\| x^k - \frac{\langle T(z^k)+\overline{v}^k, x^k-z^k\rangle}{\|T(z^k)+\overline{v}^k\|^2}\big(T(z^k)+\overline{v}^k\big) - x^k\right\|^2 \\ &= \|x^k-x_*\|^2 - \frac{(\langle T(z^k)+\overline{v}^k, x^k-z^k\rangle)^2}{\|T(z^k)+\overline{v}^k\|^2}.\end{aligned}$$

It follows from the last inequality that

$$\frac{(\langle T(z^k)+\overline{v}^k, x^k-z^k\rangle)^2}{\|T(z^k)+\overline{v}^k\|^2} \le \|x^k-x_*\|^2 - \|x^{k+1}-x_*\|^2. \tag{33}$$

Since $T$ and the projection are continuous and $(x^k)_{k\in\mathbb{N}}$ is bounded, $(z^k)_{k\in\mathbb{N}}$ is bounded. The boundedness of $\big(\|T(z^k)+\overline{v}^k\|\big)_{k\in\mathbb{N}}$ follows from (22). Using Fact 2.12(ii), the right hand side of (33) goes to 0, when $k$ goes to $\infty$. Then, the result follows. ■

Next we establish our main convergence result on **Variant B.1**.

**Theorem 6.5** *The sequence* $(x^k)_{k\in\mathbb{N}}$ *converges to a point in* $S_*$.

*Proof.* We claim that there exists an accumulation point of $(x^k)_{k\in\mathbb{N}}$ belonging to $S_*$. The existence of the accumulation points of $(x^k)_{k\in\mathbb{N}}$ follows from Proposition 6.4(ii). Let $(x^{i_k})_{k\in\mathbb{N}}$ be a convergent subsequence of $(x^k)_{k\in\mathbb{N}}$ such that, $(u^{i_k})_{k\in\mathbb{N}}$, $(v^{i_k})_{k\in\mathbb{N}}$ and $(\alpha_{i_k})_{k\in\mathbb{N}}$ also converge, and set $\lim_{k\to\infty} x^{i_k} = \tilde{x}$, $\lim_{k\to\infty} u^{i_k} = \tilde{u}$, $\lim_{k\to\infty} v^{i_k} = \tilde{v}$ and $\lim_{k\to\infty} \alpha^{i_k} = \tilde{\alpha}$. Using Proposition 6.4(iii), (29), and taking the limit along the subsequence $(i_k)_{k\in\mathbb{N}}$, we have $0 = \lim_{k\to\infty}\langle T(z^{i_k}) + \overline{v}^{i_k}, x^{i_k} - z^{i_k}\rangle \ge \frac{(1-\delta)}{\tilde{\alpha}} \lim_{k\to\infty}\|z^{i_k} - x^{i_k}\|^2 \ge 0$. Then,

$$\lim_{k\to\infty}\|x^{i_k} - z^{i_k}\| = 0. \tag{34}$$

Now we have two cases:

**Case 1:** $\lim_{k\to\infty} \alpha_{i_k} = \tilde{\alpha} > 0$. We have from (22), the continuity of the projection, and (34) that $\tilde{x} = \lim_{k\to\infty} x^{i_k} = \lim_{k\to\infty} z^{i_k} = P_C\big(\tilde{x} - \tilde{\alpha}(T(\tilde{x}) + \tilde{\alpha}\tilde{u})\big)$. Then, $\tilde{x} = P_C\big(\tilde{x} - \tilde{\alpha}(T(\tilde{x}) + \tilde{\alpha}\tilde{u})\big)$ and as consequence of Proposition 2.14, $\tilde{x} \in S_*$.

**Case 2:** $\lim_{k\to\infty} \alpha_{i_k} = \tilde{\alpha} = 0$. Define $\tilde{\alpha}_k := \frac{\alpha_k}{\theta}$. Hence,

$$\lim_{k\to\infty} \tilde{\alpha}_{i_k} = \lim_{k\to\infty} \frac{\alpha_{i_k}}{\theta} = 0. \tag{35}$$

Since $\tilde{\alpha}_k$ does not satisfy Armijo-type condition in **Linesearch B**, we have

$$\|T\big(\tilde{z}^k\big) - T(x^k) + \tilde{\alpha}_k \tilde{v}^k - \tilde{\alpha}_k u^k\| > \frac{\delta\|\tilde{z}^k - x^k\|}{\tilde{\alpha}_k}, \tag{36}$$

where $\tilde{v}^k \in \mathcal{N}_C(\tilde{z}^k)$ and

$$\tilde{z}^k = P_C(x^k - \tilde{\alpha}_k(T(x^k) + \tilde{\alpha}_k u^k)). \tag{37}$$

The left hand side of (36) goes to 0 along the subsequence $(i_k)_{k\in\mathbb{N}}$ by the continuity of $T$ and $P_C$. So,

$$\liminf_{k\to 0} \frac{\|\tilde{z}^k - x^k\|}{\tilde{\alpha}_k} = 0. \tag{38}$$

By Corollary 2.8, with $x = x^k$, $\alpha = \tilde{\alpha}_k$ and $p = T(x^k) + \tilde{\alpha}_k u^k$, we have

$$\frac{x^k - \tilde{z}^k}{\tilde{\alpha}_k} \in T(x^k) + \tilde{\alpha}_k u^k + \mathcal{N}_C(\tilde{z}^k).$$

Taking the limits along the subsequence $(i_k)_{k\in\mathbb{N}}$ and using (35), (37), (38), the continuity of $T$ and the closedness of $\mathrm{Gph}(\mathcal{N}_C)$, we get that $0 \in T(\tilde{x}) + \mathcal{N}_C(\tilde{x})$, thus, $\tilde{x} \in S_*$. ■

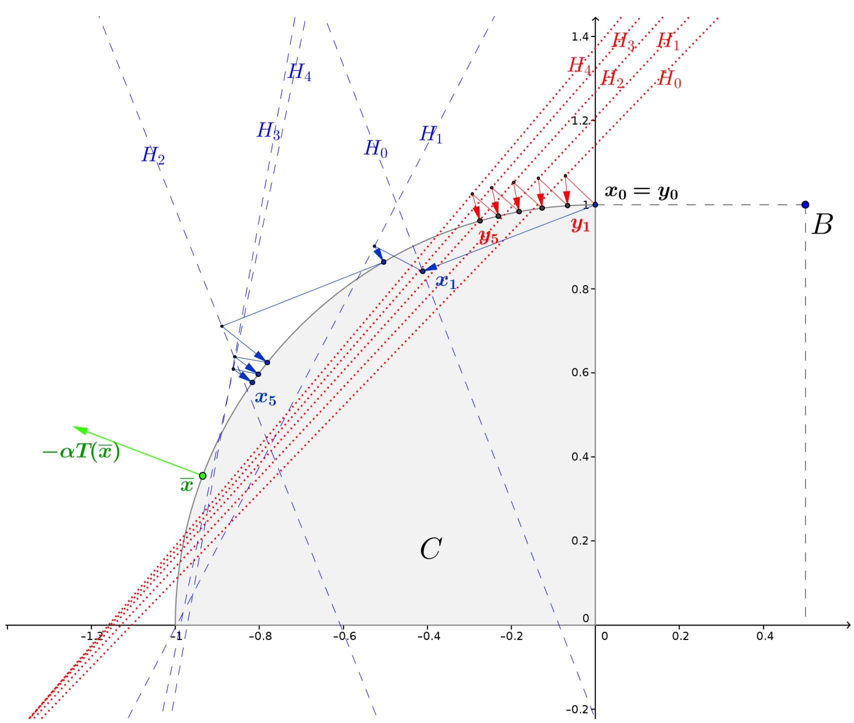


Figure 3: **Variant B.1** with and without normal vectors.

Figure 3 above examines the performance of **Variant B.1** for the variational inequality in Example 3.1 with and without normal vectors. It shows the first *five* elements of sequences $(y^k)_{k\in\mathbb{N}}$ (generated without normal vectors) and $(x^k)_{k\in\mathbb{N}}$ (generated with non-null normal vectors).

## 6.2 Convergence analysis on Variant B.2

In this section, all results are for **Variant B.2**, which is summarized below.

**Variant B.2** $x^{k+1} = \mathcal{F}_{\mathrm{B.2}}(x^k) = P_{C\cap H(z^k,\overline{v}^k)}(x^k)$ where

$$H(z^k,\overline{v}^k) = \big\{y \in \mathbb{R}^n : \langle T(z^k) + \overline{v}^k, y - z^k\rangle \le 0\big\},$$

$z^k$ and $\overline{v}^k$ are given by (22) and (23a), respectively.

**Proposition 6.6** *If $x^{k+1} = x^k$, if and only if $x^k \in S_*$ and* **Variant B.2** *stops.*

*Proof.* We have $x^{k+1} = x^k$ implies $x^k \in C \cap H(z^k,\overline{v}^k)$. Thus, $x^k \in S_*$ by Proposition 6.2. Conversely, if $x^k \in S_*$, then Proposition 6.2 implies $x^k \in H(z^k,\overline{v}^k)$. Then, the result follows from (25). ■

We study the case that **Variant B.2** does not stop, thus, it generates a sequence $(x^k)_{k\in\mathbb{N}}$.

**Proposition 6.7** *The sequence $(x^k)_{k\in\mathbb{N}}$ is Féjer convergent to $S_*$. Moreover, it is bounded and $\lim_{k\to\infty}\|x^{k+1} - x^k\| = 0$.*

*Proof.* Take $x_* \in S_*$. By Lemma 2.17, $x_* \in H(z^k,\overline{v}^k)$, for all $k$, and also $x_*$ belongs to $C$ implying that the projection step (25) is well-defined. Then, using Fact 2.7(i) for two points $x^k$, $x_*$ and the set $C \cap H(z^k,\overline{v}^k)$, we have

$$\|x^{k+1} - x_*\|^2 \le \|x^k - x_*\|^2 - \|x^{k+1} - x^k\|^2. \tag{39}$$

So, $(x^k)_{k\in\mathbb{N}}$ is Féjer convergent to $S_*$. Hence, by Fact 2.12(i) $(x^k)_{k\in\mathbb{N}}$ is bounded. Taking the limit in (39) and using Fact 2.12(ii), the result follows. ■

The next proposition shows a relation between the projection steps in **Variant B.1** and **Variant B.2**. This fact has a geometry interpretation: since the projection of **Variant B.2** is done onto a smaller set, it can improve the convergence of **Variant B.1**.

**Proposition 6.8** *Let $(x^k)_{k\in\mathbb{N}}$ be the sequence generated by* **Variant B.2**. *Then:*

(i) $x^{k+1} = P_{C\cap H(z^k,\overline{v}^k)}(P_{H(z^k,\overline{v}^k)}(x^k))$.

(ii) $\lim_{k\to\infty}\langle A(z^k) + \overline{v}^k, x^k - z^k\rangle = 0$.

*Proof.* (i): Since $x^k \in C$ but $x^k \notin H(z^k,\overline{v}^k)$ and $C \cap H_k \neq \varnothing$, the result follows from Lemma 2.9.

(ii): Take $x_* \in S_*$. Notice that $x^{k+1} = P_{C\cap H(z^k,\overline{v}^k)}(x^k)$ and that projections onto convex sets are firmly-nonexpansive (see Fact 2.7(i)), we have

$$\|x^{k+1} - x_*\|^2 = \|x^k - x_*\|^2 - \|x^{k+1} - x^k\|^2 \le \|x^k - x_*\|^2 - \|P_{H(z^k,\overline{v}^k)}(x^k) - x^k\|^2.$$

The rest of the proof is analogous to Proposition 6.4(iii). ■

Finally we present the convergence result for **Variant B.2**.

**Proposition 6.9** *The sequence* $(x^k)_{k\in\mathbb{N}}$ *converges to a point in* $S_*$.

*Proof.* Similar to the proof of Theorem 6.5. ■

Next, we examine the performance of **Variant B.2** for the variational inequality in Example 3.1 with and without normal vectors.

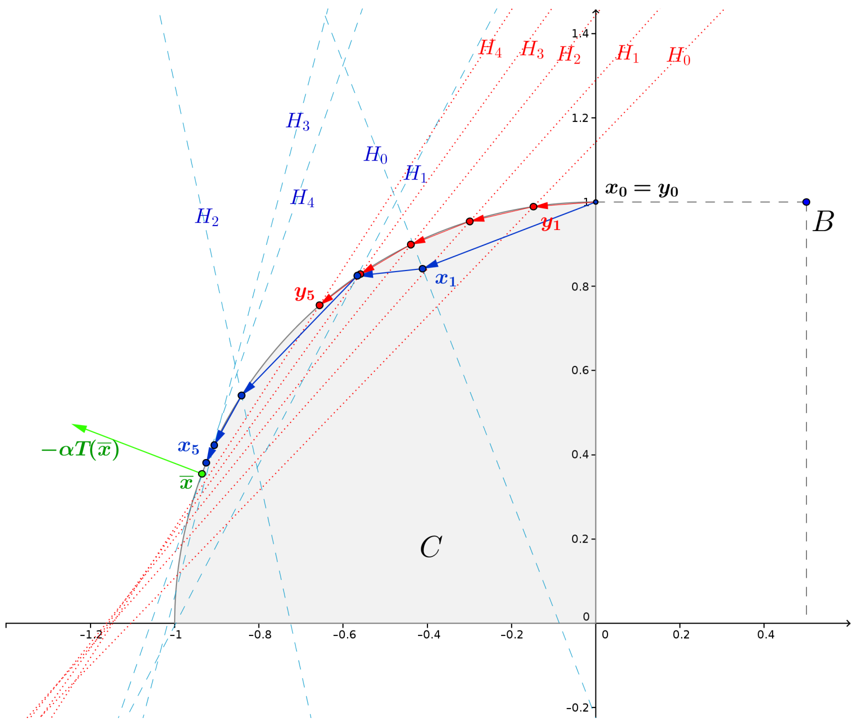


Figure 4: **Variant B.2** with and without normal vectors.

Figure 4 shows the first *five* elements of sequences $(y^k)_{k\in\mathbb{N}}$ (generated without normal vectors) and $(x^k)_{k\in\mathbb{N}}$ (generated with non-null normal vectors).

### 6.3 Convergence analysis on Variant B.3

In this section, all results are for **Variant B.3**, which is summarized below.

**Variant B.3** $x^{k+1} = \mathcal{F}_{\mathrm{B.3}}(x^k) = P_{C\cap H(z^k,\bar{v}^k)\cap W(x^k)}(x^0)$ where

$$W(x^k) = \{y \in \mathbb{R}^n : \langle y - x^k, x^0 - x^k\rangle \le 0\},$$
$$H(z^k,\bar{v}^k) = \{y \in \mathbb{R}^n : \langle T(z^k) + \bar{v}^k, y - z^k\rangle \le 0\},$$

$z^k$ and $\bar{v}^k$ are defined by (22) and (23a), respectively.

**Proposition 6.10** *If $x^{k+1} = x^k$, then $x^k \in S_*$ and* **Variant B.3** *stops.*

*Proof.* We have $x^{k+1} = P_{C\cap H(z^k,\overline{v}^k)\cap W(x^k)}(x^0) = x^k$. So, $x^k \in C \cap H(z^k,\overline{v}^k) \cap W(x^k) \subseteq H(z^k,\overline{v}^k)$. Finally, $x^k \in S_*$ by Proposition 6.2. ■

We now consider the case **Variant B.3** does not stop. Observe that $W(x^k)$ and $H(z^k,\overline{v}^k)$ are closed halfspaces, for each $k$. Therefore, $C \cap H(z^k,\overline{v}^k) \cap W(x^k)$ is a closed convex set. So, if the set $C \cap H(z^k,\overline{v}^k) \cap W(x^k)$ is nonempty, then the next iterate, $x^{k+1}$, is well-defined. The following lemma guarantees the non-emptiness of this set.

**Lemma 6.11** *For all $k \in \mathbb{N}$, we have $S_* \subseteq C \cap H(z^k,\overline{v}^k) \cap W(x^k)$.*

*Proof.* We proceed by induction. By definition, $S_* \neq \varnothing$ and $S_* \subseteq C$. By Lemma 2.17, $S_* \subseteq H(z^k,\overline{v}^k)$, for all $k$. For $k = 0$, as $W(x^0) = \mathbb{R}^n$, $S_* \subseteq H(z^0,\overline{v}^0) \cap W(x^0)$. Assume that $S_* \subseteq H(z^k,\overline{v}^k) \cap W(x^k)$. Then, $x^{k+1} = P_{C\cap H(z^k,\overline{v}^k)\cap W(x^k)}(x^0)$ is well-defined. By Fact 2.7(ii), we obtain $\langle x_* - x^{k+1}\,,\, x^0 - x^{k+1}\rangle \leq 0$, for all $x_* \in S_*$. This implies $x_* \in W(x^{k+1})$. Hence, $S_* \subseteq H(z^{k+1},\overline{v}^{k+1}) \cap W(x^{k+1})$. Then, the statement follows by induction. ■

As previous variants we establish the conversely part of Proposition 6.10, which is a direct consequence of Lemma 6.11.

**Corollary 6.12** *If $x^k \in S_*$, then $x^{k+1} = x^k$, and* **Variant B.3** *stops.*

**Corollary 6.13** **Variant B.3** *is well-defined.*

*Proof.* Lemma 6.11 shows that $C \cap H(z^k,\overline{v}^k) \cap W(x^k)$ is nonempty for all $k \in \mathbb{N}$. So, the projection step (26) is well-defined. Thus, **Variant B.3** is well-defined by using Proposition 6.1. ■

Before proving the convergence of the sequence $(x^k)_{k\in\mathbb{N}}$, we study its boundedness. The next lemma shows that the sequence remains in a ball determined by the initial point.

**Lemma 6.14** *Let $\overline{x} = P_{S_*}(x^0)$ and $\rho = \operatorname{dist}(x^0, S_*)$. Then $(x^k)_{k\in\mathbb{N}} \subset B\left[\frac{1}{2}(x^0 + \overline{x}), \frac{1}{2}\rho\right] \cap C$. In particular, $(x^k)_{k\in\mathbb{N}}$ is bounded.*

*Proof.* $S_* \subseteq H(z^k,\overline{v}^k) \cap W(x^k)$ follows from Lemma 6.11. Using Lemma 2.10, with $S = S_*$ and $x = x^k$, we have $x^k \in B\left[\frac{1}{2}(x^0 + \overline{x}), \frac{1}{2}\rho\right]$, for all $k \in \mathbb{N}$. Finally, notice that $(x^k)_{k\in\mathbb{N}} \subset C$. ■

Now, we focus on the properties of the accumulation points.

**Proposition 6.15** *All accumulation points of $(x^k)_{k\in\mathbb{N}}$ belong to $S_*$.*

*Proof.* Notice that $W(x^k)$ is a halfspace with normal $x^0 - x^k$, we have $x^k = P_{W(x^k)}(x^0)$. Moreover, $x^{k+1} \in W(x^k)$. Thus, by the firm non-expansiveness of $P_{W(x^k)}$ (see Fact 2.7(i)), we have $\|x^{k+1} - x^k\|^2 \leq \|x^{k+1} - x^0\|^2 - \|x^k - x^0\|^2$. So, the sequence $(\|x^k - x^0\|)_{k\in\mathbb{N}}$ is monotone and nondecreasing. In addition, $(\|x^k - x^0\|)_{k\in\mathbb{N}}$ is bounded by Lemma 6.14. Thus, $(x^k)_{k\in\mathbb{N}}$ converges. So,

$$\lim_{k\to\infty} \|x^{k+1} - x^k\| = 0. \tag{40}$$

Since $x^{k+1} \in H(z^k,\overline{v}^k)$, we get $\langle T(z^k) + \overline{v}^k, x^{k+1} - z^k\rangle \leq 0$, where $\overline{v}^k$ and $z^k$ are defined by (23a) and (22), respectively. Combining the above inequality with (29), we get $0 \geq \langle T(z^k) + \overline{v}^k, x^{k+1} -$

$x^k\rangle + \langle T(z^k) + \overline{v}^k, x^k - z^k\rangle \geq -\|T(z^k) + \overline{v}^k\| \cdot \|x^{k+1} - x^k\| + \frac{1-\delta}{\alpha_k}\|x^k - z^k\|^2$. After simple algebra and using (22),

$$\|x^k - z^k\|^2 \leq \frac{\sigma}{1-\delta}\|T(z^k) + \overline{v}^k\| \cdot \|x^{k+1} - x^k\|. \tag{41}$$

Choosing a subsequence $(i_k)_{k\in\mathbb{N}}$ such that, the subsequences $(\alpha_{i_k})_{k\in\mathbb{N}}$, $(x^{i_k})_{k\in\mathbb{N}}$ and $(\overline{v}^{i_k})_{k\in\mathbb{N}}$ converge to $\tilde{\alpha}$, $\tilde{x}$ and $\tilde{v}$, respectively. This is possible by the boundedness of $(\overline{v}^k)_{k\in\mathbb{N}}$ and $(x^k)_{k\in\mathbb{N}}$ . Taking the limits in (41) and using (40), we obtain $\lim_{k\to\infty}\|x^{i_k} - z^{i_k}\|^2 = 0$. and as consequence $\tilde{x} = \lim_{k\to\infty} z^{i_k}$. Now we consider two cases:

**Case 1:** $\lim_{k\to\infty}\alpha_{i_k} = \tilde{\alpha} > 0$. By (22) and the continuity of the projection, $\tilde{x} = \lim_{k\to\infty} z^{i_k} = P_C\big(\tilde{x} - \tilde{\alpha}(T(\tilde{x}) + \tilde{\alpha}\tilde{u})\big)$ and hence by Proposition 2.14, $\tilde{x} \in S_*$.

**Case 2:** $\lim_{k\to\infty}\alpha_{i_k} = \tilde{\alpha} = 0$. Then, $\lim_{k\to\infty}\frac{\alpha_{i_k}}{\theta} = 0$. The rest part of this case is analogous to the proof of Theorem 6.5. ■

Finally, we are ready to prove the convergence of the sequence $(x^k)_{k\in\mathbb{N}}$ generated by **Variant B.3**, to the solution closest to $x^0$.

**Theorem 6.16** *Define $\overline{x} = P_{S_*}(x^0)$. Then, $(x^k)_{k\in\mathbb{N}}$ converges to $\overline{x}$.*

*Proof.* First note that from Lemma 2.18 $\bar{x}$ is well-defined. It follows from Lemma 6.14 that $(x^k)_{k\in\mathbb{N}} \subset B\left[\frac{1}{2}(x^0+\overline{x}), \frac{1}{2}\rho\right] \cap C$, so it is bounded. Let $(x^{i_k})_{k\in\mathbb{N}}$ be a convergent subsequence of $(x^k)_{k\in\mathbb{N}}$, and let $\hat{x}$ be its limit. Thus, $\hat{x} \in B\left[\frac{1}{2}(x^0+\overline{x}), \frac{1}{2}\rho\right] \cap C$. Furthermore, by Proposition 6.15, $\hat{x} \in S_*$. So, $\hat{x} \in S_* \cap B\left[\frac{1}{2}(x^0+\overline{x}), \frac{1}{2}\rho\right] = \{\overline{x}\}$, implying $\hat{x} = \overline{x}$, *i.e.*, $\overline{x}$ is the unique limit point of $(x^k)_{k\in\mathbb{N}}$. Hence, $(x^k)_{k\in\mathbb{N}}$ converges to $\overline{x} \in S_*$. ■

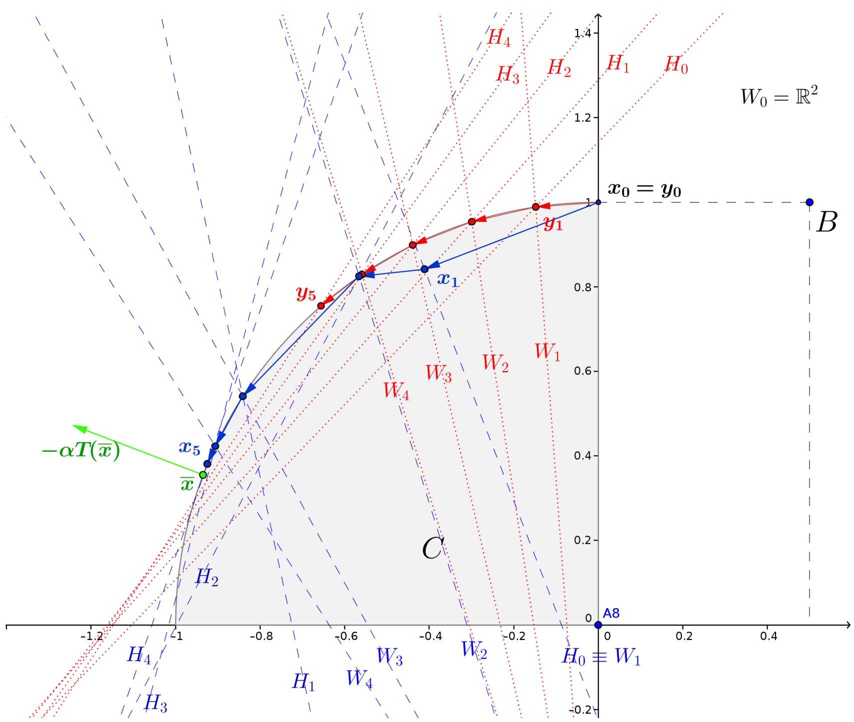


Figure 5: **Variant B.3** with and without normal vectors.

Figure 5 above shows the performance of the first *five* elements of the sequence generated by **Variant B.3** in Example 3.1 with and without normal vectors. Recall that $(y^k)_{k\in\mathbb{N}}$ (is the sequence generated without normal vectors) and $(x^k)_{k\in\mathbb{N}}$ (generated with non-null normal vectors).

## 7 Conceptual algorithm with Linesearch F

We continue presenting our second conceptual algorithm for solving (1) using **Linesearch F**.

**Conceptual Algorithm F** Given $(\beta_k)_{k\in\mathbb{N}} \subset [\check{\beta}, \hat{\beta}]$ $0 < \check{\beta} \le \hat{\beta} < +\infty$, $\delta \in (0,1)$, and $M > 0$.

**Step 1 (Initialization):** Take $x^0 \in C$ and set $k \leftarrow 0$.

**Step 1 (Stopping Test 1):** If $x^k = P_C(x^k - T(x^k))$, then stop. Otherwise,

**Step 2 (Linesearch F):** Take $u^k \in \mathcal{N}_C(x^k)$ with $\|u^k\| \le M$ and set

$$(\alpha_k, z^k, \bar{v}^k) = \textbf{Linesearch F } (x^k, u^k, \beta_k, \delta, M), \tag{42}$$

*i.e.*, $(\alpha_k, z^k, \bar{v}^k)$ satisfy

$$\begin{cases} \bar{v}^k \in \mathcal{N}_C(\alpha_k z^k + (1-\alpha_k)x^k) \text{ with } \|\bar{v}^k\| \le M \ ; \quad \alpha_k \le 1 \ ; \\ z^k = P_C(x^k - \beta_k(T(x^k) + \alpha_k u^k)); \\ \langle T(\alpha_k z^k + (1-\alpha_k)x^k) + \bar{v}^k, x^k - z^k\rangle \ge \delta \langle T(x^k) + \alpha_k u^k, x^k - z^k\rangle. \end{cases} \tag{43}$$

**Step 3 (Projection Step):** Set

$$\bar{x}^k := \alpha_k z^k + (1-\alpha_k)x^k, \tag{44a}$$

$$\text{and} \quad x^{k+1} := \mathcal{F}(x^k). \tag{44b}$$

**Step 4 (Stopping Test 2):** If $x^{k+1} = x^k$, then stop. Otherwise, set $k \leftarrow k+1$ and go to **Step 1**.

Also, we consider three different projection steps (called Variants **F.1**, **F.2** and **F.3**) which are analogous to **Conceptual Algorithm B**. These variants of **Conceptual Algorithm F** have their main difference in the projection step given in (44b).

$$\mathcal{F}_{\mathrm{F}.1}(x^k) = P_C\big(P_{H(\bar{x}^k,\bar{v}^k)}(x^k)\big); \qquad (\textbf{Variant F.1}) \tag{45}$$

$$\mathcal{F}_{\mathrm{F}.2}(x^k) = P_{C\cap H(\bar{x}^k,\bar{v}^k)}(x^k); \qquad (\textbf{Variant F.2}) \tag{46}$$

$$\mathcal{F}_{\mathrm{F}.3}(x^k) = P_{C\cap H(\bar{x}^k,\bar{v}^k)\cap W(x^k)}(x^0), \qquad (\textbf{Variant F.3}) \tag{47}$$

where $H(x,v)$ and $W(x)$ are defined by (27). Now, we analyze some general properties of **Conceptual Algorithm F**.

**Proposition 7.1** *Assuming that* (44b) *is well-defined whenever* $x^k$ *is available. Then,* **Conceptual Algorithm F** *is well-defined.*

*Proof.* If **Step 1** is not satisfied, then **Step 2** is guaranteed by Lemma 5.2. Thus, the entire algorithm is well-defined. ∎

**Proposition 7.2** *$x^k \in H(\overline{x}^k, \overline{v}^k)$ for $\overline{x}^k$ and $\overline{v}^k$ as in* (44a) *and* (42) *respectively, if and only if, $x^k \in S_*$.*

*Proof.* Since $x^k \in H(\overline{x}^k, \overline{v}^k)$, $\langle T(\overline{x}^k)+\overline{v}^k, x^k-\overline{x}^k\rangle \leq 0$. Using the definition of $\alpha_k$ in **Algorithm F**, (43) and (44a), we have

$$\begin{aligned}0 &\geq \langle T(\overline{x}^k)+\overline{v}^k, x^k-\overline{x}^k\rangle = \alpha_k\langle T(\overline{x}^k)+\overline{v}^k, x^k-z^k\rangle \geq \alpha_k\delta\big\langle T(x^k)+\alpha_k u^k, x^k-z^k\big\rangle \\ &\geq \frac{\alpha_k}{\beta_k}\delta\|x^k-z^k\|^2 \geq \frac{\alpha_k}{\hat{\beta}}\delta\|x^k-z^k\|^2, \end{aligned} \tag{48}$$

implying that $x^k = z^k$. Then, from (43) and Proposition 2.14, $x^k \in S_*$ by Proposition 2.14. Conversely, if $x^k \in S_*$, then $x^k \in H(\overline{x}^k, \overline{v}^k)$ by Lemma 2.17. ■

Next, we note a useful algebraic property on the sequence generated by **Conceptual Algorithm F**, which is a direct consequence of the linesearch in **Algorithm F**. Let $(x^k)_{k\in\mathbb{N}}$ and $(\alpha_k)_{k\in\mathbb{N}}$ be sequences generated by **Conceptual Algorithm F**, using (48), we get

$$\forall k \in \mathbb{N}: \quad \langle T(\overline{x}^k)+\overline{v}^k, x^k-\overline{x}^k\rangle \geq \frac{\alpha_k}{\hat{\beta}}\delta\|x^k-z^k\|^2. \tag{49}$$

## 7.1 Convergence analysis on Variant F.1

In this subsection all results are for **Variant F.1**, which is summarized below.

**Variant F.1** $x^{k+1} = \mathcal{F}_{\text{F.1}}(x^k) = P_C(P_{H(\overline{x}^k,\overline{v}^k)}(x^k)) = P_C\Big(x^k - \dfrac{\langle T(\overline{x}^k)+\overline{v}^k, x^k-\overline{x}^k\rangle}{\|T(\overline{x}^k)+\overline{v}^k\|^2}(T(\overline{x}^k)+\overline{v}^k)\Big)$
where

$$H(\overline{x}^k, \overline{v}^k) = \big\{y \in \mathbb{R}^n : \langle T(z^k)+\overline{v}^k, y-z^k\rangle \leq 0\big\},$$

$\overline{x}^k$ and $\overline{v}^k$ as (44a) and (42), respectively.

**Proposition 7.3** *If $x^{k+1} = x^k$, if and only if $x^k \in S_*$, and* **Variant F.1** *stops.*

*Proof.* We have $x^{k+1} = P_C\big(P_{H(\overline{x}^k,\overline{v}^k)}(x^k)\big) = x^k$. Then, Fact 2.7(ii) implies

$$\langle P_{H(\overline{x}^k,\overline{v}^k)}(x^k) - x^k, z - x^k\rangle \leq 0, \tag{50}$$

for all $z \in C$. Again, using Fact 2.7(ii),

$$\langle P_{H(\overline{x}^k,\overline{v}^k)}(x^k) - x^k, P_{H(\overline{x}^k,\overline{v}^k)}(x^k) - z\rangle \leq 0, \tag{51}$$

for all $z \in H(\overline{x}^k, \overline{v}^k)$. Note that $C \cap H(\overline{x}^k, \overline{v}^k) \neq \varnothing$. So, for any $z \in C \cap H(\overline{x}^k, \overline{v}^k)$, adding up (50) and (51) yields $\|x^k - P_{H(\overline{x}^k,\overline{v}^k)}(x^k)\|^2 = 0$. Hence, $x^k = P_{H(\overline{x}^k,\overline{v}^k)}(x^k)$, *i.e.*, $x^k \in H(\overline{x}^k, \overline{v}^k)$. Finally, we have $x^k \in S_*$ by Proposition 7.2. Conversely, if $x^k \in S_*$, Proposition 7.2 implies $x^k \in H(\bar{x}^k, \bar{v}^k)$ and together with (45), we get $x^k = x^{k+1}$. ■

From now on, we assume that **Variant F.1** does not stop. Note that by Lemma 2.17, $H(\overline{x}^k, \overline{v}^k)$ is nonempty for all $k$. Then, the projection step (45) is well-defined. Thus, **Variant F.1** generates an infinite sequence $(x^k)_{k\in\mathbb{N}}$ such that $x^k \notin S_*$ for all $k \in \mathbb{N}$.

**Proposition 7.4** *The following hold:*

(i) *The sequence* $(x^k)_{k\in\mathbb{N}}$ *is Fejér convergent to* $S_*$.

(ii) *The sequence* $(x^k)_{k\in\mathbb{N}}$ *is bounded.*

(iii) $\lim_{k\to\infty}\langle A(\overline{x}^k)+\overline{v}^k, x^k-\overline{x}^k\rangle = 0.$

*Proof.* (i): Take $x_*\in S_*$. Note that, by definition $(\overline{x}^k,\overline{v}^k)\in \mathrm{Gph}(\mathcal{N}_C)$. Using (45), Fact 2.7(i) and Lemma 2.17, we have

$$\begin{aligned}\|x^{k+1}-x_*\|^2 &= \|P_C(P_{H(\overline{x}^k,\overline{v}^k)}(x^k)) - P_C(P_{H(\overline{x}^k,\overline{v}^k)}(x_*))\|^2\\ &\le \|P_{H(\overline{x}^k,\overline{v}^k)}(x^k) - P_{H(\overline{x}^k,\overline{v}^k)}(x_*)\|^2 \le \|x^k-x_*\|^2 - \|P_{H(\overline{x}^k,\overline{v}^k)}(x^k)-x^k\|^2. \end{aligned} \tag{52}$$

So, $\|x^{k+1}-x_*\|\le\|x^k-x_*\|$.

(ii): Follows immediately from the previous item and Fact 2.12(i).

(iii): Take $x_*\in S_*$. Using (44a) and $P_{H(\overline{x}^k,\overline{v}^k)}(x^k) = x^k - \dfrac{\langle T(\overline{x}^k)+\overline{v}^k, x^k-\overline{x}^k\rangle}{\|T(\overline{x}^k)+\overline{v}^k\|^2}\big(T(\overline{x}^k)+\overline{v}^k\big)$ combining with (52), yields

$$\begin{aligned}\|x^{k+1}-x_*\|^2 &\le \|x^k-x_*\|^2 - \left\| x^k - \frac{\langle T(\overline{x}^k)+\overline{v}^k, x^k-\overline{x}^k\rangle}{\|T(\overline{x}^k)+\overline{v}^k\|^2}\big(T(\overline{x}^k)+\overline{v}^k\big) - x^k\right\|^2\\ &= \|x^k-x_*\|^2 - \frac{(\langle T(\overline{x}^k)+\overline{v}^k, x^k-\overline{x}^k\rangle)^2}{\|T(\overline{x}^k)+\overline{v}^k\|^2}.\end{aligned}$$

It follows that $\dfrac{(\langle T(\overline{x}^k)+\overline{v}^k, x^k-\overline{x}^k\rangle)^2}{\|T(\overline{x}^k)+\overline{v}^k\|^2} \le \|x^k-x_*\|^2-\|x^{k+1}-x_*\|^2$. Using Fact 2.12(ii), the right side of the above inequality goes to 0, when $k$ goes to $\infty$ and since $T$ is continuous and $(x^k)_{k\in\mathbb{N}}$, $(z^k)_{k\in\mathbb{N}}$ and $(\overline{x}^k)_{k\in\mathbb{N}}$ are bounded. Implying the boundedness of $\big(\|T(\overline{x}^k)+\overline{v}^k\|\big)_{k\in\mathbb{N}}$ and the desired result. ■

Next, we establish our main convergence result on **Variant F.1**.

**Theorem 7.5** *The sequence* $(x^k)_{k\in\mathbb{N}}$ *converges to a point in* $S_*$.

*Proof.* We claim that there exists an accumulation point of $(x^k)_{k\in\mathbb{N}}$ belonging to $S_*$. The existence of the accumulation points follows from Proposition 7.4(ii). Let $(x^{i_k})_{k\in\mathbb{N}}$ be a convergent subsequence of $(x^k)_{k\in\mathbb{N}}$ such that, $(\overline{x}^{i_k})$, $(\overline{v}^{i_k})$, $(u^{i_k})$, $(\alpha_{i_k})_{k\in\mathbb{N}}$ and $(\beta_{i_k})_{k\in\mathbb{N}}$ also converge, and set $\lim_{k\to\infty}x^{i_k}=\tilde{x}$, $\lim_{k\to\infty}u^{i_k}=\tilde{u}$, $\lim_{k\to\infty}\alpha_{i_k}=\tilde{\alpha}$ and $\lim_{k\to\infty}\beta_{i_k}=\tilde{\beta}$. Using Proposition 7.4(iii) and by passing to the limit in (49), over the subsequence $(i_k)_{k\in\mathbb{N}}$, we have $0=\lim_{k\to\infty}\langle T(\overline{x}^{i_k})+\bar{u}^{i_k}, x^{i_k}-\overline{x}^{i_k}\rangle \ge \lim_{k\to\infty}\frac{\alpha_{i_k}}{\hat{\beta}}\delta\|x^{i_k}-z^{i_k}\|^2\ge 0$. Therefore,

$$\lim_{k\to\infty}\alpha_{i_k}\|x^{i_k}-z^{i_k}\| = 0. \tag{53}$$

Now we consider two cases.

**Case 1:** $\lim_{k\to\infty} \alpha_{i_k} = \tilde{\alpha} > 0$. In the view of (53), $\lim_{k\to\infty} \|x^{i_k} - z^{i_k}\| = 0$. Using the continuity of $T$, and the projection imply $\tilde{x} = \lim_{k\to\infty} x^{i_k} = \lim_{k\to\infty} z^{i_k} = P_C\big(\tilde{x} - \tilde{\beta}(T(\tilde{x}) + \tilde{\alpha}\tilde{u})\big)$. Then, $\tilde{x} = P_C\big(\tilde{x} - \tilde{\beta}(T(\tilde{x}) + \tilde{\alpha}\tilde{u})\big)$, and Proposition 2.14 implies that $\tilde{x} \in S_*$.

**Case 2:** $\lim_{k\to\infty} \alpha_{i_k} = \tilde{\alpha} = 0$. Define $\tilde{\alpha}_k = \frac{\alpha_{i_k}}{\theta}$. Then,

$$\lim_{k\to\infty} \tilde{\alpha}_{i_k} = 0. \tag{54}$$

Define $\tilde{y}^k := \tilde{\alpha}_k \tilde{z}^k + (1 - \tilde{\alpha}_k)x^k$, where $\tilde{z}^k = P_C(x^k - \beta_k(T(x^k) + \tilde{\alpha}_k u^k))$, as (43). Hence,

$$\lim_{k\to\infty} y^{i_k} = \tilde{x}. \tag{55}$$

From the definition $\alpha_k$ in **Algorithm F**, $\tilde{y}^k$ does not satisfy the condition, *i.e.*,

$$\Big\langle T(\tilde{y}^k) + \tilde{v}^k, x^k - \tilde{z}^k \Big\rangle < \delta \langle T(x^k) + \tilde{\alpha}_k u^k, x^k - \tilde{z}^k \rangle, \tag{56}$$

for $\tilde{v}^k \in \mathcal{N}_C(\tilde{y}^k)$ and all $k$. Taking a subsequence $(i_k)_{k\in\mathbb{N}}$ and relabeling if necessary, we assume that $(v^k_{\tilde{\alpha}_k})_{k\in\mathbb{N}}$ converges to $\tilde{v}$ over the subsequence $(i_k)_{k\in\mathbb{N}}$. By Fact 2.6, $\tilde{v}$ belongs to $\mathcal{N}_C(\tilde{x})$. Using (43) and (54), $\lim_{k\to\infty} \tilde{z}^{i_k} = \tilde{z} = P_C\big(\tilde{x} - \tilde{\beta}T(\tilde{x})\big)$. By passing to the limit in (56) over the subsequence $(i_k)_{k\in\mathbb{N}}$ and using (55), we get $\langle T(\tilde{x}) + \tilde{v}, \tilde{x} - \tilde{z} \rangle \le \delta \langle T(\tilde{x}), \tilde{x} - \tilde{z} \rangle$. Then,

$$\begin{aligned} 0 &\ge (1-\delta)\big\langle T(\tilde{x}), \tilde{x} - \tilde{z} \big\rangle + \big\langle \tilde{v}, \tilde{x} - \tilde{z} \big\rangle \ge (1-\delta)\big\langle T(\tilde{x}), \tilde{x} - \tilde{z} \big\rangle \\ &= \frac{(1-\delta)}{\tilde{\beta}} \langle \tilde{x} - (\tilde{x} - \tilde{\beta}T(\tilde{x})), \tilde{x} - \tilde{z} \rangle \ge \frac{(1-\delta)}{\tilde{\beta}} \|\tilde{x} - \tilde{z}\|^2 \ge \frac{(1-\delta)}{\hat{\beta}} \|\tilde{x} - \tilde{z}\|^2 \ge 0. \end{aligned}$$

This means $\tilde{x} = \tilde{z}$, which implies $\tilde{x} \in S_*$. ■

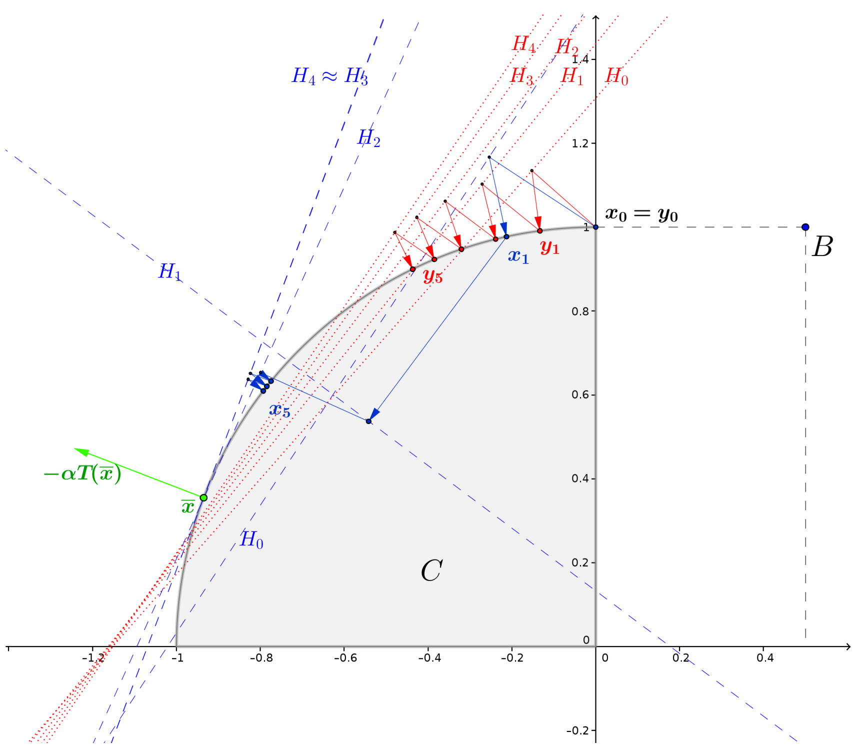


Figure 6: **Variant F.1** with and without normal vectors.

Figure 6 examines the performance of **Variant F.1** for the variational inequality in Example 3.1 with and without normal vectors. Figure 6 shows the first *five* elements of sequences $(y^k)_{k\in\mathbb{N}}$ (generated without normal vectors) and $(x^k)_{k\in\mathbb{N}}$ (generated with non-null normal vectors).

## 7.2 Convergence analysis on Variant F.2

In this section, all results are for **Variant F.2**, which is summarized below.

**Variant F.2** $x^{k+1} = \mathcal{F}_{\mathrm{F.2}}(x^k) = P_{C\cap H(\overline{x}^k,\overline{v}^k)}(x^k)$ where

$$H(\overline{x}^k,\overline{v}^k) = \left\{y \in \mathbb{R}^n : \langle T(z^k) + \overline{v}^k, y - z^k\rangle \leq 0\right\},$$

$\overline{x}^k$ and $\overline{v}^k$ given by (44a) and (42), respectively.

**Proposition 7.6** *If $x^{k+1} = x^k$, if and only if $x^k \in S_*$, and* **Variant F.2** *stops.*

*Proof.* We have $x^{k+1} = P_{C\cap H(\overline{x}^k,\overline{v}^k)}(x^k) = x^k$. So, $x^k \in C \cap H(\overline{x}^k,\overline{v}^k)$. Hence, $x^k \in S_*$ by Proposition 7.2. Conversely, if $x^k \in S_*$, Proposition 7.2 implies $x^k \in H(\bar{x}^k,\bar{v}^k)$ and together with (46), we get $x^k = x^{k+1}$. ■

From now on, we assume that **Variant F.2** does not stop.

**Proposition 7.7** *The sequence $(x^k)_{k\in\mathbb{N}}$ is Féjer convergent to $S_*$. Moreover, it is bounded and $\lim_{k\to\infty}\|x^{k+1} - x^k\| = 0$.*

*Proof.* Take $x_* \in S_*$. By Lemma 2.17, $x_* \in H(\overline{x}^k,\overline{v}^k)$, for all $k$ and also $x_*$ belongs to $C$, so, the projection step (46) is well-defined. Then, using Fact 2.7(i) for the projection operator $P_{H(\overline{x}^k,\overline{v}^k)}$, we obtain

$$\|x^{k+1} - x_*\|^2 \leq \|x^k - x_*\|^2 - \|x^{k+1} - x^k\|^2. \tag{57}$$

The above inequality implies that $(x^k)_{k\in\mathbb{N}}$ is Féjer convergent to $S_*$. Hence, by Fact 2.12(i)&(ii), $(x^k)_{k\in\mathbb{N}}$ is bounded and thus $(\|x^k - x_*\|)_{k\in\mathbb{N}}$ is a convergent sequence. By passing to the limit in (57) and using Fact 2.12(ii), we get $\lim_{k\to\infty}\|x^{k+1} - x^k\| = 0$. ■

The next proposition shows a relation between the projection steps in **Variant F.1** and **Variant F.2**. This fact has a geometry interpretation: since the projection of **Variant F.2** is done over a small set, it may improve the convergence behaviour of the sequence generated by **Variant F.1**.

**Proposition 7.8** *Let $(x^k)_{k\in\mathbb{N}}$ be the sequence generated by* **Variant F.2**. *Then,*

(i) $x^{k+1} = P_{C\cap H(\overline{x}^k,\overline{v}^k)}(P_{H(\overline{x}^k,\overline{v}^k)}(x^k))$.

(ii) $\lim_{k\to\infty}\langle A(\overline{x}^k) + \overline{v}^k, x^k - \overline{x}^k\rangle = 0$.

*Proof.* (i): Since $x^k \in C$ but $x^k \notin H(\overline{x}^k,\overline{v}^k)$ and $C \cap H_k \neq \varnothing$, by Lemma 2.9, we have the result.

(ii): Take $x_* \in S_*$. Notice that $x^{k+1} = P_{C\cap H(z^k,\overline{v}^k)}(x^k)$ and that projections onto convex sets are firmly-nonexpansive (see Fact 2.7(i)), we have

$$\|x^{k+1} - x_*\|^2 \leq \|x^k - x_*\|^2 - \|x^{k+1} - x^k\|^2 \leq \|x^k - x_*\|^2 - \|P_{H(\overline{x}^k,\overline{v}^k)}(x^k) - x^k\|^2.$$

The rest of the proof is analogous to Proposition 7.4(iii). ■

**Proposition 7.9** *The sequence $(x^k)_{k\in\mathbb{N}}$ converges to a point in $S_*$.*

*Proof.* Similar to the proof of Theorem 7.5. ■

Next, we examine the performance of **Variant F.2** for the variational inequality in Example 3.1 with and without normal vectors. Figure 7 shows the first *five* elements of sequences $(y^k)_{k\in\mathbb{N}}$ (generated without normal vectors) and $(x^k)_{k\in\mathbb{N}}$ (generated with non-null normal vectors).

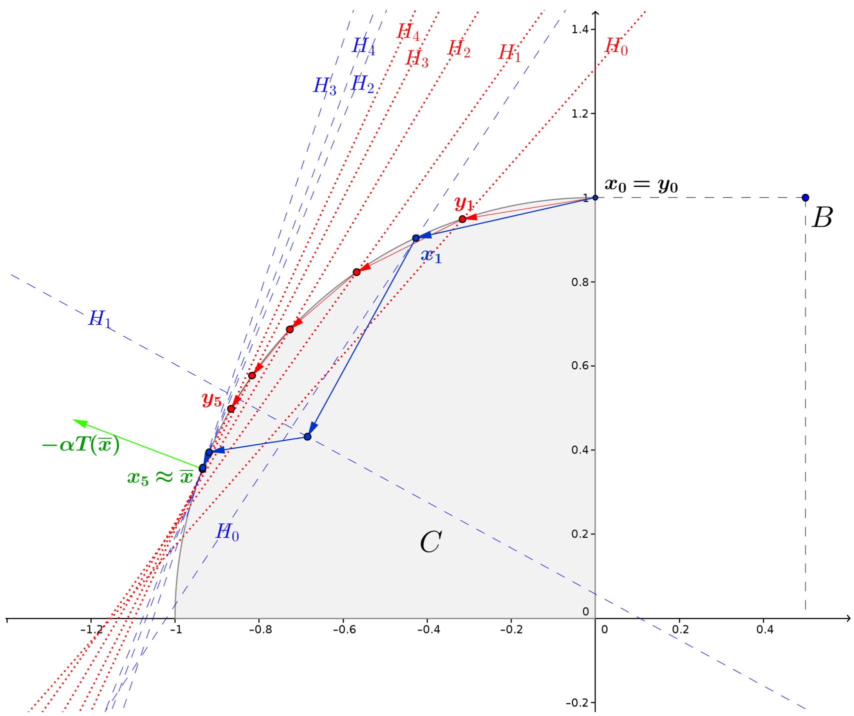


Figure 7: **Variant F.2** with and without normal vectors.

## 7.3 Convergence analysis on Variant F.3

In this section, all results are for **Variant F.3**, which is summarized below.

**Variant F.3** $x^{k+1} = \mathcal{F}_{\text{F.3}}(x^k) = P_{C\cap H(\overline{x}^k,\overline{v}^k)\cap W(x^k)}(x^0)$ where

$$W(x^k) = \{y \in \mathbb{R}^n : \langle y - x^k, x^0 - x^k\rangle \le 0\},$$
$$H(\overline{x}^k,\overline{v}^k) = \{y \in \mathbb{R}^n : \langle T(z^k) + \overline{v}^k, y - z^k\rangle \le 0\},$$

$\overline{x}^k$ and $\overline{v}^k$ given by (44a) and (42), respectively.

**Proposition 7.10** *If $x^{k+1} = x^k$, then $x^k \in S_*$ and* **Variant F.3** *stops.*

*Proof.* We have $x^{k+1} = P_{C\cap H(\overline{x}^k,\overline{v}^k)\cap W(x^k)}(x^0) = x^k$. So, $x^k \in C \cap H(\overline{x}^k, \overline{v}^k) \cap W(x^k) \subseteq H(\overline{x}^k, \overline{v}^k)$. Thus, $x^k \in S_*$ by Proposition 7.2. ■

From now on we assume that **Variant F.3** does not stop. Observe that, by the virtue of their definitions, $W(x^k)$ and $H(\overline{x}^k, \overline{v}^k)$ are convex and closed halfspaces, for each $k$. Therefore, $C \cap H(\overline{x}^k, \overline{v}^k) \cap W(x^k)$ is a closed convex set. So, if $C \cap H(\overline{x}^k, \overline{v}^k) \cap W(x^k)$ is nonempty, then the next iterate, $x^{k+1}$, is well-defined. The following lemma guarantees this fact and the proof is very similar to Lemma 6.11.

**Lemma 7.11** *For all $k \in \mathbb{N}$, we have $S_* \subset C \cap H(\overline{x}^k, \overline{v}^k) \cap W(x^k)$.*

*Proof.* We proceed by induction. By definition, $S_* \neq \varnothing$ and $S_* \subseteq C$. By Lemma 2.17, $S_* \subseteq H(\overline{x}^k, \overline{v}^k)$, for all $k$. For $k = 0$, as $W(x^0) = \mathbb{R}^n$, $S_* \subseteq H(\overline{x}^0, \overline{v}^0) \cap W(x^0)$. Assume that $S_* \subseteq H(\overline{x}^\ell, \overline{v}^\ell) \cap W(x^\ell)$, for $\ell \leq k$. Henceforth, $x^{k+1} = P_{C\cap H(\overline{x}^k,\overline{v}^k)\cap W(x^k)}(x^0)$ is well-defined. Then, by Fact 2.7(ii), we have $\langle x_* - x^{k+1}, x^0 - x^{k+1}\rangle \leq 0$, for all $x_* \in S_*$. This implies $x_* \in W(x^{k+1})$, and hence, $S_* \subseteq H(\overline{x}^{k+1}, \overline{v}^{k+1}) \cap W(x^{k+1})$. Then, the result follows by induction. ■

The above lemma shows that the set $C \cap H(\overline{x}^k, \overline{v}^k) \cap W(x^k)$ is nonempty and as consequence the projection step, given in (47), is well-defined. Before proving the convergence of the sequence, we study its boundedness. The next lemma shows that the sequence remains in a ball determined by the initial point.

As previous variants we establish the conversely part of Proposition 7.10, which is a direct consequence of Lemma 7.11.

**Corollary 7.12** *If $x^k \in S_*$, then $x^{k+1} = x^k$, and* **Variant F.3** *stops.*

**Lemma 7.13** *The sequence $(x^k)_{k\in\mathbb{N}}$ is bounded. Furthermore, $(x^k)_{k\in\mathbb{N}} \subset B\left[\frac{1}{2}(x^0 + \overline{x}), \frac{1}{2}\rho\right] \cap C$, where $\overline{x} = P_{S_*}(x^0)$ and $\rho = \operatorname{dist}(x^0, S_*)$.*

*Proof.* It follows from Lemma 7.11 that $S_* \subseteq H(\overline{x}^k, \overline{v}^k) \cap W(x^k)$, for all $k \in \mathbb{N}$. The proof now follows by repeating the proof of Lemma 6.14. ■

Finally we prove the convergence of the sequence generated by **Variant F.3** to the solution closest to $x^0$.

**Theorem 7.14** *Define $\overline{x} = P_{S_*}(x^0)$. Then, $(x^k)_{k\in\mathbb{N}}$ converges to $\overline{x}$.*

*Proof.* First we prove the optimality of the all accumulation points of $(x^k)_{k\in\mathbb{N}}$. Notice that $W(x^k)$ is a halfspace with normal $x^0 - x^k$, we have $x^k = P_{W(x^k)}(x^0)$. Moreover, $x^{k+1} \in W(x^k)$. So, by the firm nonexpansiveness of $P_{W(x^k)}$, we have $0 \leq \|x^{k+1} - x^k\|^2 \leq \|x^{k+1} - x^0\|^2 - \|x^k - x^0\|^2$, which implies that the sequence $(\|x^k - x^0\|)_{k\in\mathbb{N}}$ is monotone and nondecreasing. From Lemma 7.13, we have that $(\|x^k - x^0\|)$ is bounded, thus, convergent. So,

$$\lim_{k\to\infty} \|x^{k+1} - x^k\| = 0. \tag{58}$$

Since $x^{k+1} \in H(\overline{x}^k, \overline{v}^k)$, we get

$$\langle T(\overline{x}^k) + \overline{v}^k, x^{k+1} - \overline{x}^k\rangle \leq 0, \tag{59}$$

with $\overline{v}^k$ and $\overline{x}^k$ given by (42) and (44a). Substituting (44a) into (59), we have $\langle T(\overline{x}^k)+\overline{v}^k, x^{k+1} - x^k\rangle + \alpha_k\big\langle T(\overline{x}^k)+\overline{v}^k, x^k-z^k\big\rangle \le 0$. Combining the above inequality with (43), we get

$$\langle T(\overline{x}^k)+\overline{v}^k, x^{k+1}-x^k\rangle + \alpha_k\delta\langle T(x^k)+\alpha_k u^k, x^k-z^k\rangle \le 0. \tag{60}$$

Combining the above inequality with (43) and using Fact 2.7(iii), we can check that $\beta_k\langle T(x^k)+\alpha_k u^k, x^k-z^k\rangle \ge \|x^k-z^k\|^2$. Thus, after use the last inequality and the Cauchy-Schwartz inequality in (60), we get

$$\frac{\alpha_k}{\beta_k}\delta\|x^k-z^k\|^2 \le \|T(\overline{x}^k)+\overline{v}^k\|\cdot\|x^{k+1}-x^k\|. \tag{61}$$

Choosing a subsequence $(i_k)$ such that, the subsequences $(\alpha_{i_k})_{k\in\mathbb{N}}$, $(u^{i_k})_{k\in\mathbb{N}}$, $(\beta_{i_k})_{k\in\mathbb{N}}$, $(x^{i_k})_{k\in\mathbb{N}}$ and $(\overline{v}^{i_k})_{k\in\mathbb{N}}$ converge to $\tilde{\alpha}$, $\tilde{u}$, $\tilde{\beta}$, $\tilde{x}$ and $\tilde{v}$ respectively (this is possible by the boundedness of all of these sequences) and taking limit in (61) along the subsequence $(i_k)_{k\in\mathbb{N}}$, we get, from (58),

$$\lim_{k\to\infty}\alpha_{i_k}\|x^{i_k}-z^{i_k}\|^2 = 0. \tag{62}$$

Now we consider two cases,

**Case 1:** $\lim_{k\to\infty}\alpha_{i_k} = \tilde{\alpha} > 0$. By (62), $\lim_{k\to\infty}\|x^{i_k}-z^{i_k}\|^2 = 0$. By continuity of the projection, we have $\tilde{x} = P_C\big(\tilde{x} - \tilde{\beta}(T(\tilde{x})+\tilde{\alpha}\tilde{u})\big)$. So, $\tilde{x}\in S_*$ by Proposition 2.14.

**Case 2:** $\lim_{k\to\infty}\alpha_{i_k} = 0$. Then, $\lim_{k\to\infty}\frac{\alpha_{i_k}}{\theta} = 0$. The rest is similar to the proof of Theorem 7.5.

Thus, all accumulation points of $(x^k)_{k\in\mathbb{N}}$ are in $S_*$. The proof follows similar to Theorem 6.16. ■

Next, we examine the performance of **Variant F.3** for the variational inequality in Example 3.1 with and without normal vectors. Figure 8 shows the first *five* elements of sequences $(y^k)_{k\in\mathbb{N}}$ (generated without normal vectors) and $(x^k)_{k\in\mathbb{N}}$ (generated with non-null normal vectors).

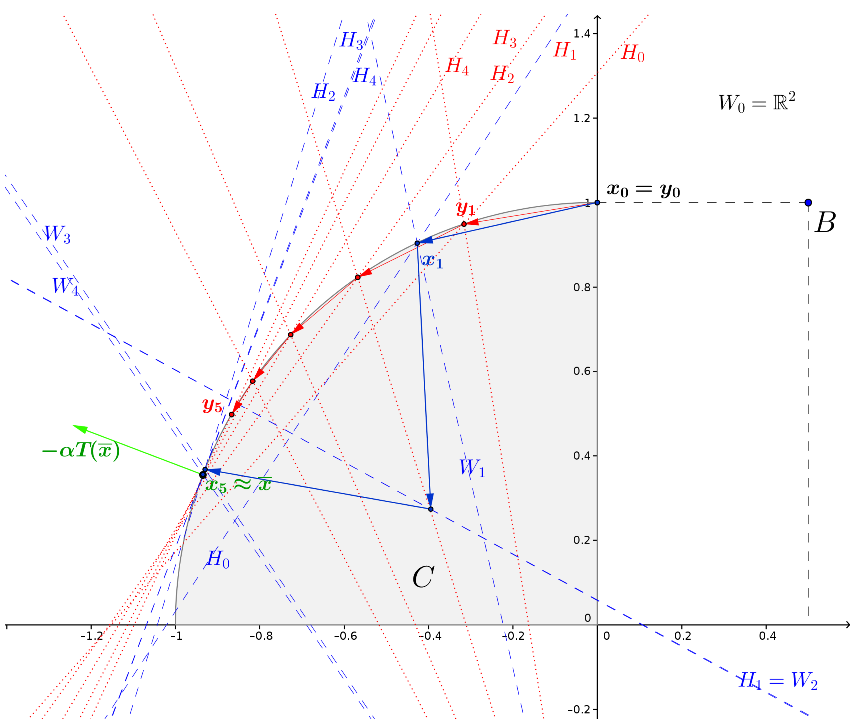


Figure 8: **Variant F.3** with and without normal vectors.

# 8 Final remarks

In this paper we have proposed two conceptual conditional extragradient algorithms generalizing classical extragradient algorithms for solving constrained variational inequality problems. The main idea here comes from the (sub)gradient algorithms where non-null normal vectors of the feasible set improve the convergence, avoiding zigzagging. The scheme proposed here mainly contains two parts:

(**1**) Two different linesearches are analyzed. The linesearches allow us to find suitable halfspaces containing the solution set of the problem using non-null normal vectors of the feasible set. It is well-known in the literature that such procedures are very effective in absence of Lipschitz continuity and they use more information available at each iteration, allowing long steplength.

(**2**) Many projection steps are performed, which yield different and interesting features extending several known projection algorithms. Furthermore, the convergence analysis of both conceptual algorithms was established assuming existence of solutions, continuity and a weaker condition than pseudomonotonicity on the operator, showing examples when the non-null vectors in the normal cone archive better performance.

We hope that this study will serve as a basis for future research on other more efficient variants, as well as including sophisticated linesearches permitting optimal choice for the vectors in the normal cone of the feasible set. Several of the ideas of this paper merit further investigation, some of which will be presented in future work. In particular we discuss in a separate paper variants of the projection algorithms proposed in [9] for solving nonsmooth variational inequalities. The difficulties of extending this previous result to point-to-set operators are non-trivial, the main obstacle yields in the impossibility to use linesearches or separating techniques as was strongly used in this paper. To our knowledge, variants of the linesearches for variational inequalities require smoothness of $T$, even for the nonlinear convex optimization problems ($T = \partial f$) is not possible make linesearch, because the negative subgradients are not always a descent direction. Actually, a few explicit methods have been proposed in the literature for solving nonsmooth monotone variational inequality problems, examples of such methods appear in [13, 20]. Future work will be addressed to further investigation on the modified Forward-Backward splitting iteration for inclusion problems [7, 8, 36], exploiting the additive structure of the main operator and adding dynamic choices of the stepsizes with conditional and deflected techniques [15, 30].

## Acknowledgments

This work was completed while JYBC was visiting the University of British Columbia. This author is very grateful for the warm hospitality of the Irving K. Barber School of Arts and Sciences at the University of British Columbia. RDM was supported by his scholarship for his doctoral studies, granted jointly by CAPES. HMP was partially supported by an NSERC accelerator grant of Heinz H. Bauschke (University of British Columbia, Kelowna, B.C., Canada, V1V 1V7).

## References

[1] Auslender, A., Teboulle, M.: Interior projection-like methods for monotone variational inequalities. *Mathematical Programming* **104** (2005) 39–68.

[2] Bauschke, H.H., Borwein, J.M.: On projection algorithms for solving convex feasibility problems. *SIAM Review* **38** (1996) 367–426.

[3] Bauschke, H.H., Combettes, Patrick L.: *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. Springer, New York (2011).

[4] Bello Cruz, J.Y., Iusem, A.N.: A strongly convergent direct method for monotone variational inequalities in Hilbert spaces. *Numerical Functional Analysis and Optimization* **30** (2009) 23–36.

[5] Bello Cruz, J.Y., Iusem, A.N.: Convergence of direct methods for paramonotone variational inequalities. *Computation Optimization and Applications* **46** (2010) 247–263.

[6] Bello Cruz, J.Y., Iusem, A.N.: A strongly convergent method for nonsmooth convex minimization in Hilbert spaces. *Numerical Functional Analysis and Optimization* **32** (2011) 1009–1018.

[7] Bello Cruz, J.Y., Díaz Millán, R.: A direct splitting method for nonsmooth variational inequalities. *Journal Optimization Theory and Application* **161** (2014) 728–737.

[8] Bello Cruz, J.Y., Díaz Millán, R.: A variant of Forward-Backward splitting method for the sum of two monotone operators with a new search strategy. *Optimization*, DOI: 10.1080/02331934.2014.883510 (2014).

[9] Bello Cruz, J.Y., Díaz Millán, R.: *A relaxed-projection splitting algorithm for variational inequalities in Hilbert spaces*. http://arxiv.org/pdf/1312.3921.pdf (2014).

[10] Bonnans, J.F., Shapiro, A.: *Perturbation Analysis of Optimization Problems*. Springer, New York (2000).

[11] Browder, F.E.: Convergence theorems for sequences of nonlinear operators in Banach spaces. *Mathematische Zeitschrift* **100** (1967) 201–225.

[12] Burachik, R.S., Iusem, A.N.: *Set-Valued Mappings and Enlargements of Monotone Operators*. Springer, Berlin (2008).

[13] Burachik, R.S., Lopes, J.O., Svaiter, B.F.: An outer approximation method for the variational inequality problem. *SIAM Journal on Control and Optimization* **43** (2005) 2071–2088.

[14] Censor, Y., Gibali, A., Reich, S.: The subgradient extragradient method for solving variational inequalities in Hilbert space. *Journal of Optimization Theory and Applications* **148** (2011) 318–335.

[15] d'Antonio, G., Frangioni, A.: Convergence analysis of deflected conditional approximate subgradient methods. *SIAM Journal on Optimization* **20** (2009) 357–386.

[16] Facchinei, F., Pang, J.S.: *Finite-dimensional Variational Inequalities and Complementarity Problems*. Springer, Berlin (2003).

[17] Ferris, M.C., Pang, J.S.: Engineering and economic applications of complementarity problems. *SIAM Review* **39** (1997) 669–713.

[18] Harker, P.T., Pang, J.S.: Finite dimensional variational inequalities and nonlinear complementarity problems: a survey of theory, algorithms and applications. *Mathematical Programming* **48** (1990) 161–220.

[19] Hartman, P., Stampacchia, G.: On some non-linear elliptic differential-functional equations. *Acta Mathematica* **115** (1966) 271–310.

[20] He, B.S.: A new method for a class of variational inequalities. *Mathematical Programming* **66** (1994) 137–144.

[21] Iusem, A.N.: An iterative algorithm for the variational inequality problem. *Computational and Applied Mathematics* **13** (1994) 103–114.

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

[23] Iusem, A.N., Svaiter, B.F., Teboulle, M.: Entropy-like proximal methods in convex programming. *Mathematics of Operations Research* **19** (1994) 790–814.

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

[25] Khobotov, E.N.: Modifications of the extragradient method for solving variational inequalities and certain optimization problems. *USSR Computational Mathematics and Mathematical Physics* **27** (1987) 120–127.

[26] Konnov, I.V.: *Combine Relaxation Methods for Variational Inequalities.* Lecture Notes in Economics and Mathematical Systems. **495** Springer-Velarg, Berlin (2001).

[27] Konnov, I.V.: A combined relaxation method for variational inequalities with nonlinear constraints. *Mathematical Programming.* **80** (1998) 239–252.

[28] Konnov, I. V.: A class of combined iterative methods for solving variational inequalities. *Journal of Optimization Theory and Applications.* **94** (1997) 677–693.

[29] Korpelevich, G.M.: The extragradient method for finding saddle points and other problems. *Ekonomika i Matematcheskie Metody* **12** (1976) 747–756.

[30] Larson, T., Patriksson, M., Stromberg, A-B.: Conditional subgradient optimization - Theory and application. *European Journal of Operational Research.* **88** (1996) 382–403.

[31] Rockafellar, R.T.: *Convex Analysis.* Princeton, New York (1970).

[32] Rockafellar, R.T., Wets, R.J-B.: *Variational Analysis.* Springer, Berlin (1998).

[33] Solodov, M.V., Svaiter, B.F.: A new projection method for monotone variational inequality problems. *SIAM Journal on Control and Optimization* **37** (1999) 765–776.

[34] Solodov, M.V., Svaiter, B.F.: Forcing strong convergence of proximal point iterations in a Hilbert space. *Mathematical Programming* **87** (2000) 189-202.

[35] Solodov, M.V., Tseng, P.: Modified projection-type methods for monotone variational inequalities. *SIAM Journal on Control and Optimization* **34** (1996) 1814–1830.

[36] Tseng, P.: A modified forward-backward splitting method for maximal monotone mappings. *SIAM on Journal Control Optimization* **38** (2000) 431–446.

[37] Zaraytonelo, E.H.: *Projections on Convex Sets in Hilbert Space and Spectral Theory.* in Contributions to Nonlinear Functional Analysis, E. Zarantonello, Academic Press, New York (1971) 237–424.