# Pseudomonotone$_*$ maps and the cutting plane property

Nicolas Hadjisavvas[1] and Siegfried Schaible[2]

**Abstract.** Pseudomonotone$_*$ maps are a generalization of paramonotone maps which is very closely related to the cutting plane property in variational inequality problems (VIP). In this paper, we first generalize the so-called minimum principle sufficiency and the maximum principle sufficiency for VIP with multivalued maps. Then we show that pseudomonotonicity$_*$ of the map implies the "maximum principle sufficiency" and, in fact, is equivalent to it in a sense. We then present two applications of pseudomonotone$_*$ maps. First we show that pseudomonotone$_*$ maps can be used instead of the much more restricted class of pseudomonotone$_+$ maps in a cutting plane method. Finally, an application to a proximal point method is given.

**Keywords:** Variational inequality; Pseudomonotone$_*$ map; Cutting plane method; Minimum principle sufficiency; Maximum principle sufficiency.

**Mathematics Subject Classification 2000:** 49J40; 47H04; 65K05.

[1]Corresponding author. Department of Product and Systems Design Engineering, University of the Aegean, 84100 Hermoupolis, Syros, Greece. Email: nhad@aegean.gr.

[2]A.G. Anderson Graduate School of Management, University of California, Riverside, California, CA 92512, U.S.A. Email: siegfried.schaible@ucr.edu.

# 1 Introduction

In recent years, several algorithms for the solution of variational inequalities were based on paramonotone maps. The introduction of these maps goes back to 1976 [2], but a more systematic study was initiated in [3] and continued in [14] and other papers thereafter. Paramonotone maps form a class which is more restricted than monotone but larger than strictly monotone maps, and contains the subdifferentials of proper, convex, lsc functions. They posses a property, the so-called cutting plane property, that makes them suitable in several interior point algorithms.

Paramonotone maps were subsequently generalized to pseudomonotone$_*$ maps. This was done in [4] for the single valued case and in [13] for the multivalued case. Pseudomonotone$_*$ maps form a class that lies between pseudomonotone and strictly pseudomonotone maps and contains the subdifferentials of locally Lipschitz pseudoconvex functions. These maps also have the cutting plane property, and in a sense are exactly the set of maps that always have this property [13].

The aim of the present paper is to show that pseudomonotone$_*$ maps may indeed succesfully replace other, more restricted, classes of maps in several applications, and exhibit their relation to the so-called "maximum priciple sufficiency" [17]. In Section 2 we recall the maximum principle sufficiency and minimum principle sufficiency as they were defined for variational inequality problems with single valued maps, and generalize them to the case of multivalued maps. Further, we show that pseudomonotone$_*$ maps are exactly the class of maps for which the maximum principle sufficiency always holds.

In Section 3 we present an application of pseudomonotone$_*$ maps to a cutting plane algorithm proposed in [9]. We obtain a result similar to the one in [9] but in a more general setting (with a map which is multivalued pseudomonotone$_*$ rather than single valued pseudomonotone$_+$), and with a weaker continuity assumption. In Section 4 we show how pseudomonotone$_*$ maps can be used instead of paramonotone maps, in a proximal point algorithm.

We first fix the notation and recall some preliminary results. In the following, $X$ will be a Banach space and $X^*$ its topological dual. For a set $K \subseteq X$ we denote by $N_K(x)$ the normal cone to $K$ at $x$. For $A \subseteq X^*$, we set $\mathbb{R}_{++}A = \{tx^* : t > 0, x^* \in A\}$. Given a multivalued map $T : X \to 2^{X^*}$, $D(T)$ will denote its domain and $Z_T$ its set of zeros, i.e., the set $Z_T = \{x \in X : 0 \in T(x)\}$. An element $x \in K$ is called a solution of the Stampacchia variational inequality problem VIP if

$$\exists x^* \in T(x) : \langle x^*, y - x\rangle \geq 0 \text{ for all } y \in K. \tag{VIP}$$

It is called a solution of the Minty variational inequality problem MVIP if

$$\forall y \in K, \forall y^* \in T(y),\ \langle y^*, y - x\rangle \geq 0. \tag{MVIP}$$

We denote by $S(T, K)$ (resp., $S_M(T, K)$) the set of solutions of the VIP (resp., MVIP).

There are two quite different notions of pseudomonotonicity we will employ. One is mainly related to continuity and is due to Brezis and Browder: we will follow [18] and call it t-pseudomonotonicity (t- stands for topological).

**Definition 1** *A multivalued map $T : X \to 2^{X^*}$ with closed convex domain $D(T)$ is called t-pseudomonotone on $D(T)$ if for every $\{x_k\} \subseteq D(T)$ converging weakly to some $x \in D(T)$ and every choice $x_k^* \in T(x_k)$ such that*

$$\limsup_{k\to+\infty} \langle x_k^*, x_k - x\rangle \le 0$$

*the following holds: For every $y \in D(T)$ there exists $x^* \in T(x)$ (depending in general on $y$) such that*

$$\langle x^*, x - y\rangle \le \liminf_{k\to+\infty} \langle x_k^*, x_k - y\rangle .$$

The other notion was introduced by Karamardian, and we will call it simply pseudomonotonicity, as it is indeed a generalization of monotonicity.

**Definition 2** *A multivalued map $T : X \to 2^{X^*}$ is called pseudomonotone if for every $x, y \in X$ and every $x^* \in T(x)$, $y^* \in T(y)$, the following implication holds:*

$$\langle x^*, y - x\rangle \ge 0 \Rightarrow \langle y^*, y - x\rangle \ge 0.$$

We recall some more necessary definitions and facts. Details can be found in [11, 12]. Two pseudomonotone maps $T_1$, $T_2$ are called equivalent ($T_1 \sim T_2$) if they have the same domain, the same set of zeros, and for every $x \in X\backslash Z_{T_1}$, $\mathbb{R}_{++}T_1(x) = \mathbb{R}_{++}T_2(x)$. This is an equivalence relation. The equivalence class of a pseudomonotone map $T$ has a maximum element $\hat{T}$ with respect to graph inclusion, given by

$$\hat{T}(x) = \begin{cases} N_{L_{T,x}}(x), & \text{if } x \in Z_T \\ \mathbb{R}_{++}T(x), & \text{if } x \notin Z_T. \end{cases}$$

Here, $L_{T,x}$ is the set

$$L_{T,x} = \{y \in D(T) : \exists y^* \in T(y) : \langle y^*, y - x\rangle = 0\}$$

and $N_{L_{T,x}}(x)$ is the normal cone to $L_{T,x}$ at $x$:

$$N_{L_{T,x}}(x) = \{x^* \in X^* : \langle x^*, y - x\rangle \le 0, \forall y \in L_{T,x}\} .$$

**Definition 3** *A multivalued map $T : X \to 2^{X^*}$ is called:*

*(i) paramonotone [2, 3], if it is monotone and for every $x, y \in X$ and $x^* \in T(x)$, $y^* \in T(y)$, $\langle x^* - y^*, x - y\rangle = 0$ implies that $x^* \in T(y)$ and $y^* \in T(x)$*

*(ii) pseudomonotone$_*$ [13], if it is pseudomonotone and for every $x, y \in X$ and $x^* \in T(x)$, $y^* \in T(y)$, $\langle x^*, y - x\rangle = \langle y^*, y - x\rangle = 0$ implies that $x^* \in \hat{T}(y)$ and $y^* \in \hat{T}(x)$.*

The class of pseudomonotone$_*$ maps is significantly larger than the class of paramonotone maps. As an example, the Clarke subdifferential of a locally Lipschitz pseudoconvex function is pseudomonotone$_*$; also, if $T$ is pseudomonotone$_*$ and $S \sim T$, then $S$ is pseudomonotone$_*$ [13].

Paramonotone maps have the following "cutting plane property" (CPP). This property runs as follows:

$$\left.\begin{array}{c} x \in S(T,K) \\ z \in K \\ \langle z^*, x - z\rangle \geq 0 \text{ for some } z^* \in T(z) \end{array}\right\} \Rightarrow z \in S(T,K). \tag{CPP}$$

Also pseudomonotone$_*$ maps have the CPP. What is more interesting, these maps are in a sense characterized by the CPP: if a pseudomonotone map $T$ has convex, w$^*$-compact values and has the CPP on every convex, compact subset of $K$, then $T$ is pseudomonotone$_*$ on $\operatorname{int} K$ (cf Theorem 4.1 in [13]).

In Section 3 we will use the following very weak notion of continuity.

**Definition 4** *A multivalued map $T$ is called upper sign-continuous at $x \in D(T)$ if for every $v \in X$ the following implication holds:*

$$\forall t \in (0,1), \inf_{x^* \in T(x+tv)} \langle x^*, v\rangle \geq 0 \Rightarrow \sup_{x^* \in T(x)} \langle x^*, v\rangle \geq 0.$$

For instance, any upper hemicontinuous map (i.e., a map whose restriction on any line segment in $D(T)$ is upper semicontinuous with respect to the weak$^*$-topology on $X^*$) is upper sign-continuous. Any positive function in $\mathbb{R}$ is upper sign-continuous.

# 2 Maximum principle sufficiency and CPP

Let $f : K \to \mathbb{R}$ be a Gâteaux differentiale convex function and $K^* = \operatorname{argmin}_{x \in K} f(x)$. The so-called minimum principle for convex programs asserts that for every $x \in K^*$, the elements of $K^*$ minimize $\langle \nabla f(x), \cdot\rangle$ over $K$, i.e.,

$$z \in K^* \Rightarrow \forall y \in K,\ \langle \nabla f(x), y - z\rangle \geq 0. \tag{1}$$

If the reverse implication is also true for every $x \in K^*$, then one says that the minimum principle sufficiency (MinPS) holds (see for instance [7]). This has been generalized subsequently to variational inequalities for a single-valued monotone [8] or pseudomonotone [15] map $T$ with the help of the primal gap function $g : K \to \mathbb{R} \cup \{+\infty\}$ defined by

$$g(x) := \sup_{y \in K} \langle T(x), x - y\rangle\,. \tag{2}$$

It is known that $g(x) \geq 0$ for all $x \in K$, and $g(x) = 0$ if and only if $x \in S(T,K)$. For every $x \in X$ set

$$\begin{aligned} \Gamma(x) &= \operatorname*{argsup}_{y \in K} \langle T(x), x - y\rangle = \operatorname*{arginf}_{y \in K} \langle T(x), y\rangle \\ &= \{z \in K : \langle T(x), y - z\rangle \geq 0, \forall y \in K\}\,. \end{aligned}$$

Then pseudomonotonicity of $T$ implies that for each $x \in S(T,K)$ one has $S(T,K) \subseteq \Gamma(x)$ (minimum principle). If for each $x \in S(T,K)$ one has $S(T,K) = \Gamma(x)$ then one says that the minimum principle sufficiency holds.

Minimum principle sufficiency is a very strong assumption, which is true in the special cases of quadratic programs and linear monotone complementarity problems with a nondegenerate solution [7]. It was studied in connection with the concept of weak sharp solutions and with the so-called pseudomonotone$_+$ maps, which are a rather restricted class of maps [15, 17].

In [17] another principle was introduced. Define the dual gap function

$$G(x) := \sup_{y \in K} \langle T(y), x - y \rangle$$

and

$$\begin{aligned} \Lambda(x) &= \operatorname*{argsup}_{y \in K} \langle T(y), x - y \rangle \\ &= \{ z \in K : \langle T(z), x - z \rangle \geq \langle T(y), x - y \rangle, \forall y \in K \} . \end{aligned}$$

One can easily show for a pseudomonotone map $T$ that for each $x \in S(T,K)$, $S(T,K) \subseteq \Lambda(x)$ holds. If for each $x \in S(T,K)$ one has $S(T,K) = \Lambda(x)$, then one says that the maximum principle sufficiency (MaxPS) holds.

On what follows, we will generalize these principles to the multivalued case and show the relation between them and the connection to pseudomonotone$_*$ maps.

Consider a multivalued map $T : K \to 2^{X^*} \backslash \{0\}$ with convex, weak$^*$-compact values. Then one can define the primal gap function $g : K \to \mathbb{R}$ by any of the following expressions

$$g(x) = \min_{x^* \in T(x)} \sup_{y \in K} \langle x^*, x - y \rangle = \sup_{y \in K} \min_{x^* \in T(x)} \langle x^*, x - y \rangle \tag{3}$$

where the two expressions on the right are equal by the Sion minimax theorem.

Likewise, one defines the dual gap function $G : K \to \mathbb{R}$ by

$$G(x) = \sup_{y \in K} \max_{y^* \in T(y)} \langle y^*, x - y \rangle .$$

The following properties are obvious: $g(x) \geq 0$ for all $x \in K$, and $g(x) = 0$ if and only if $x \in S(T,K)$. Likewise, $G(x) \geq 0$ for all $x \in K$, and $G(x) = 0$ if and only if $x \in S_M(T,K)$.

We can generalize the so-called minimum principle sufficiency (MinPS) and the maximum principle sufficiency (MaxPS) to the multivalued case as follows. Define

$$\Gamma(x) = \operatorname*{argsup}_{y \in K} \min_{x^* \in T(x)} \langle x^*, x - y \rangle =$$
$$\left\{ z \in K : \min_{x^* \in T(x)} \langle x^*, x - z \rangle \geq \min_{x^* \in T(x)} \langle x^*, x - y \rangle , \ \forall y \in K \right\}$$

and

$$\Lambda(x) = \operatorname*{argsup}_{y\in K} \max_{y^*\in T(y)} \langle y^*, x-y\rangle$$
$$= \left\{ z \in K : \max_{z^*\in T(z)} \langle z^*, x-z\rangle \geq \max_{y^*\in T(y)} \langle y^*, x-y\rangle, \forall y \in K \right\}.$$

**Proposition 5** *(i) For every $x \in S(T,K)$, $S_M(T,K) \subseteq \Gamma(x)$ holds.*

*(ii) For every $x \in S_M(T,K)$, $S(T,K) \subseteq \Lambda(x)$ holds.*

*Thus if, in particular, $T$ is pseudomonotone, then $x \in S(T,K)$ implies $S(T,K) \subseteq \Gamma(x)$ and $S(T,K) \subseteq \Lambda(x)$.*

**Proof.** Let $x \in S(T,K)$. Then $g(x) = 0$ hence

$$\min_{x^*\in T(x)} \langle x^*, x-y\rangle \leq 0,\ \forall y \in K. \tag{4}$$

If $z \in S_M(T,K)$, then $\langle y^*, y-z\rangle \geq 0$ for every $(y,y^*) \in \operatorname{gr} T$. This implies that $\min_{x^*\in T(x)} \langle x^*, x-z\rangle \geq 0$. Combining with (4) we obtain

$$\min_{x^*\in T(x)} \langle x^*, x-z\rangle = 0 \geq \min_{x^*\in T(x)} \langle x^*, x-y\rangle,\ \forall y \in K,$$

i.e, $z \in \Gamma(x)$. Thus $S_M(T,K) \subseteq \Gamma(x)$.

Now let $x \in S_M(T,K)$. Then $G(x) = 0$ hence

$$\max_{y^*\in T(y)} \langle y^*, x-y\rangle \leq 0\ \forall y \in K.$$

If $z \in S(T,K)$ then for some $z_0^* \in T(z)$, $\langle z_0^*, y-z\rangle \geq 0$ holds for all $y \in K$; It follows that

$$\max_{z^*\in T(z)} \langle z^*, x-z\rangle \geq \langle z_0^*, x-z\rangle \geq 0 \geq \max_{y^*\in T(y)} \langle y^*, x-y\rangle\ \forall y \in K,$$

i.e., $z \in \Lambda(x)$. Thus $S(T,K) \subseteq \Lambda(x)$.

The last part of the proposition follows since for a pseudomonotone map $T$, $S(T,K) \subseteq S_M(T,K)$. ■

Generalizing [7] to the multivalued VIP, we call the property in Proposition 5(i) "minimum principle". Further, generalizing [7, 17] we say that VIP has the minimum principle sufficiency (MinPS) if $\Gamma(x) = S(T,K)$ for all $x \in S(T,K)$. Finally, generalizing [17], we will say that VIP has the maximum principle sufficiency (MaxPS) if $\Lambda(x) = S(T,K)$ for all $x \in S(T,K)$. The following proposition relates MinPS to MaxPS.

**Proposition 6** *Let $T$ be pseudomonotone. Then MinPS implies MaxPS.*

**Proof.** Let $x \in S(T,K)$. Then $x \in S_M(T,K)$, hence $G(x) = 0$. For each $z \in \Lambda(x)$, by definition of argsup we deduce that $\max_{z^*\in T(z)} \langle z^*, x-z\rangle = 0$ thus there exists $z_0^* \in T(z)$ such that $\langle z_0^*, x-z\rangle = 0$. Using pseudomonotonicity we deduce that $\langle x^*, x-z\rangle \geq 0$ for all $x^* \in T(x)$. Hence $\min_{x^*\in T(x)} \langle x^*, x-z\rangle \geq 0$.

Using $g(x) = 0$ and the definition of $\Gamma(x)$ we infer that $z \in \Gamma(x)$. By MinPS, $\Gamma(x) = S(T, K)$. Hence $z \in S(T, K)$, which shows that $\Lambda(x) \subseteq S(T, K)$. Since the reverse inclusion always holds, MaxPS follows. ■

When VIP stems from a convex program, i.e., $T = \nabla f$ where f is convex and Gâteaux differentiable (or more generally $T = \partial f$ where $f$ is convex and lsc) it is not hard to show that, in contrast to MinPS, MaxPS always holds. This is true in much more general situations. First we compare MaxPS with CPP:

**Proposition 7** *Let $T$ be pseudomonotone. VIP has the MaxPS if and only if CPP holds.*

**Proof.** Assume that CPP holds. Fix $x \in S(T, K)$. As in the proof of the previous proposition we get that for every $z \in \Lambda(x)$ there exists $z_0^* \in T(z)$ such that $\langle z_0^*, x - z\rangle = 0$. By CPP we infer that $z \in S(T, K)$, thus $\Lambda(x) \subseteq S(T, K)$.

Conversely, assume that MaxPS holds. Take $x \in S(T, K)$ and $z \in K$ such that $\langle z^*, x - z\rangle \geq 0$ for some $z^* \in T(z)$. Then $x \in S_M(T, K)$ implies

$$\langle z^*, x - z\rangle \geq 0 \geq \langle y^*, x - y\rangle \,,\ \forall y \in K,\ \forall y^* \in T(y).$$

Thus $z \in \Lambda(x)$. By the MaxPS, $z \in S(T, K)$ thus CPP holds. ■

An obvious consequence of the above proposition and the properties of pseudomonotone$_*$ maps that we mentioned in the introduction is:

**Corollary 8** *If $T$ is pseudomonotone$_*$ then MaxPS holds. Conversely, if $T$ is pseudomonotone with convex, w$^*$-compact values and MaxPS holds on every convex, compact subset of $K$, then $T$ is pseudomonotone$_*$ on* $\operatorname{int} K$.

In particular, if we have a program defined by a convex (resp., pseudoconvex) differentiable function $f$, then $\nabla f$ is a paramonotone (resp. pseudomonotone$_*$) map. Since paramonotone maps are pseudomonotone$_*$, in both cases MaxPS holds. By the same argument, in a program defined by a locally Lipschitz pseudoconvex function, MaxPS holds.

# 3 A cutting plane algorithm

Because of the close relation between pseudomonotone$_*$ maps and the cutting plane property, these maps are naturally suited for cutting plane algorithms for solving VIP. We illustrate this point by an application to a cutting plane algorithm for solving a multivalued VIP, where the cutting planes pass through approximate analytic centers. This algorithm was proposed in [9] for a VIP with a single valued, Lipschitz, pseudomonotone$_+$ map. In our case we will generalize to a multivalued, upper sign-continuous, pseudomonotone$_*$ map.

Let $K$ be a bounded full-dimensional polyhedron in $\mathbb{R}^n$:

$$K = \{x \in \mathbb{R}^n : Ax \leq b\}$$

where $A$ is a $m \times n$ matrix and $b \in \mathbb{R}^m$. Let $T : \mathbb{R}^n \to 2^{\mathbb{R}^n}$ be a multivalued map. We first recall the following existence result.

**Theorem 9** *Assume that $T$ has nonempty, convex, compact values on $K$. Assume further that $T$ is either upper semicontinuous, or it is upper sign-continuous and pseudomonotone. Then $S(T,K) \neq \emptyset$.*

In the upper semicontinuous case, this result is well-known. In the pseudomonotone case, one may consult [18].

The algorithm for finding an element of $S(T,K)$ relies on the following idea. Assume that $T$ is pseudomonotone$_*$. If an element $x_0$ is not a solution of VIP, then by the CPP we know that all solutions are in the intersection

$$K_1 := K \cap \{x \in \mathbb{R}^n : \langle x_0^*, x - x_0 \rangle < 0\}$$

where $x_0^*$ is any element of $T(x_0^*)$. Thus, if $x_0$ is "somewhere in the middle" of $K$, we will be sure that $S(T,K)$ is bound to be in $K_1$ which is a smaller set. By continuing in this way, one can hope to produce a decreasing sequence of sets and, at the limit, arrive to a solution of VIP. The role of points "somewhere in the middle" will be played by approximate analytic centers. For the convenience of the reader, we briefly recall the definition from [9]. We write $K$ as

$$K = \{x \in \mathbb{R}^n : Ax + s = b, s \in \mathbb{R}^m_+\}.$$

A point $\bar{x} \in K$ is the analytic center if the vector $\bar{s} = b - A\bar{x} \in \mathbb{R}^m_+$ maximizes the strictly concave potential $\varphi(s) = \sum_{j=1}^{m} \ln(s_j)$. The analytic center is the unique solution of the Karush-Kuhn-Tucker system

$$\begin{aligned} A^T\bar{y} &= 0 \\ A\bar{x} + \bar{s} &= b \\ \bar{Y}\bar{s} &= e \end{aligned}$$

where $\bar{y} \in \mathbb{R}^m$ is the corresponding dual vector, $\bar{Y}$ is the $m \times m$ diagonal matrix whose diagonal elements are the coordinates of $\bar{y}$, and $e$ is the vector $(1, 1, \dots 1) \in \mathbb{R}^m$. Given $\eta \in (0,1)$, an approximate analytic center is a vector $x \in C$ such that for some $s \in \mathbb{R}^m_+$ and $y \in \mathbb{R}^m$ one has

$$\begin{aligned} A^T y &= 0 \\ Ax + s &= b \\ \|Ys - e\| &\leq \eta \end{aligned}$$

where $Y$ is the diagonal matrix whose diagonal elements are the coordinates of $y$.

We will use the same $\eta \in (0,1)$ at all iterations of the algorithm below. The algorithm runs as follows.

**1. Initialization.**
$k = 0$, $A_k = A$, $b_k = b$.

**2. Computation of an approximate analytic center.**

Set $C_k = \{x \in \mathbb{R}^n : A_k x \le b_k\}$. Find an approximate analytic center $x_k$ of $C_k$.

**3. Stopping rule.**
Compute $g(x_k)$. If $g(x_k) = 0$, then stop. Otherwise go to Step 4.

**4. Generation of a cutting plane.**
Choose $x_k^* \in T(x_k)$ and set

$A_{k+1} = \begin{pmatrix} A_k \\ x_k^* \end{pmatrix}$, $b_{k+1} = \begin{pmatrix} b_k \\ \langle x_k^*, x_k \rangle \end{pmatrix}$.

Set $k := k+1$ and go to Step 2.

We will use the following property P of the algorithm, taken from [9]:

$$\begin{array}{c}\text{Given any open ball } B(x,\varepsilon) \text{ lying in } K, \text{ there exists} \\ \text{an iteration index } k \text{ such that } B(x,\varepsilon) \nsubseteq C_k.\end{array} \tag{P}$$

Note that property (P) holds for any method of generating a cutting plane at step 4; in particular it does not depend on the properties of $T$.

We first examine the case where $T$ is upper semicontinuous.

**Theorem 10** *Let the set $K = \{x \in \mathbb{R}^n : Ax \le b\}$ be bounded and such that $\operatorname{int} K \ne \emptyset$. Assume that $T : K \to 2^{\mathbb{R}^n}$ is a pseudomonotone$_*$, upper semicontinuous map with nonempty, convex, compact values. Then the algorithm either stops with a solution of VIP or defines an infinite sequence which has a limit point that is a solution of VIP.*

**Proof.** Assume that the algorithm does not stop with a solution of VIP. Then it defines a bounded sequence $\{x_k\}$ such that $x_k \notin S(T,K)$ for all $k \in \mathbb{N}$.

By Theorem 9, $S(T,K)$ is nonempty. Choose $z \in S(T,K)$ and let $\{z_i\}$, $i \in \mathbb{N}$ be any sequence in $\operatorname{int} K$ converging to $z$. For each $i \in \mathbb{N}$ choose $\varepsilon_i > 0$ such that $B(z_i, \varepsilon_i) \subseteq \operatorname{int} K$ and $\lim_{i\to+\infty} \varepsilon_i = 0$. By Property (P), for each $i \in \mathbb{N}$ we can choose inductively $k(i) \in \mathbb{N}$ such that $B(z_i,\varepsilon_i) \nsubseteq C_k$ for all $k \ge k(i)$ and $k(i) > k(i-1)$, $i \ge 2$. Choose $y_i \in B(z_i,\varepsilon_i) \backslash C_{k(i)}$. Since $y_i \notin C_{k(i)}$, one has

$$\forall i \in \mathbb{N},\ \left\langle x_{k(i)}^*, y_i \right\rangle > \left\langle x_{k(i)}^*, x_{k(i)} \right\rangle. \tag{5}$$

Furthermore, $y_i \in B(z_i,\varepsilon_i)$ together with $\lim_{i\to+\infty} z_i = z$ and $\lim_{i\to+\infty} \varepsilon_i = 0$ imply that $\lim_{i\to+\infty} y_i = z$. By compactness of $K$, the sequence $\{x_{k(i)}\}$ has a subsequence $\{x_{k(i')}\}$, $i' \in \mathbb{N}'$ with $\mathbb{N}' \subseteq \mathbb{N}$ which converges to some element $x_0 \in K$. Since $T$ is upper semicontinuous with compact values, the sequence $\{x_{k(i')}^*\}$ has a subsequence $\{x_{k(i'')}^*\}$, $i'' \in \mathbb{N}''$ with $\mathbb{N}'' \subseteq \mathbb{N}'$, converging to some $x_0^* \in T(x_0)$. By taking the limit in (5), we deduce that $\langle x_0^*, z - x_0\rangle \ge 0$. Using property CPP we deduce that $x_0 \in S(T,K)$.

This element $x_0$ cannot be equal to any element of the sequence $\{x_k\}$ since $x_k \notin S(T,K)$ for all $k \in \mathbb{N}$. Hence it is a limit point of $\{x_k\}$. ■

The case of an upper sign-continuous map can be reduced to the upper semicontinuous case, by using the properties of the equivalence relation.

**Corollary 11** *Let the set $K = \{x \in \mathbb{R}^n : Ax \leq b\}$ be bounded and such that $\operatorname{int} K \neq \emptyset$. Assume that $T : \mathbb{R}^n \to 2^{\mathbb{R}^n}$ is a pseudomonotone$_*$, upper sign-continuous map such that $K \subseteq \operatorname{int} D(T)$, with nonempty, convex, compact values on $\operatorname{int} D(T)$. Then the algorithm either stops with a solution of VIP or defines an infinite sequence which has a limit point that is a solution of VIP.*

**Proof.** We assume again that the algorithm does not stop, hence it generates and infinite sequence $\{x_k\}$ whose elements are not solutions of VIP. We will show that we can assume, without loss of generality, that $T$ is upper semicontinous on $K$. By restricting the domain of $T$ if necessary, we can assume that $D(T) = \operatorname{int} D(T)$. According to Theorem 3.7 of [11], there exists an upper semicontinuous pseudomonotone map $T'$ with nonempty, compact convex values which is equivalent to $T$. By Proposition 3.11 of [13], $T'$ is pseudomonotone$_*$. We show by induction that $T$ and $T'$ produce the same sets $C_k$ and the same sequence $\{x_k\}$. Assume that this statement is true for some $k \in \mathbb{N}$, i.e., at iteration $k$ both maps give rise to the same set $C_k$ and the same element $x_k$.

Since $x_k \notin S(T, K)$, it is clear that $x_k \notin Z_T = Z_{T'}$. By the definition of the equivalence relation, for any choice of $x_k^* \in T(x_k)$, there exists $\lambda_k > 0$ such that $\lambda_k x_k^* \in T'(x_k)$ and vice versa. It follows that the set $C_{k+1}$ is the same for the two maps $T$ and $T'$. Also, the element $x_{k+1}$ depends only on $C_{k+1}$ and not on the map $T$ or $T'$.

The corollary follows from Theorem 10. ∎

# 4 A proximal point algorithm for solving VIP

In the following, $X$ will be a reflexive Banach space. Given a multivalued map $T : X \to 2^{X^*}$ and a nonempty closed convex set $K \subseteq X$, Burachik and Scheimberg [1] (and many others) used Bregman functions in a version of the proximal point method, in order to solve the Stampacchia variational inequality.

Let $f : X \to \mathbb{R} \cup \{+\infty\}$ be a strictly convex, proper, lsc function with closed domain $\mathcal{D} = \operatorname{dom}(f)$ which is Gâteaux differentiable on the nonempty interior of $\mathcal{D}$. For this function, the *Bregman distance* $D_f(y, x)$ is defined on $\mathcal{D} \times \operatorname{int} \mathcal{D}$ by

$$D_f(y, x) = f(y) - f(x) - \langle \nabla f(x), y - x \rangle .$$

The following "three point property" is obvious:

$$D_f(x, z) + D_f(z, y) = D_f(x, y) + \langle \nabla f(z) - \nabla f(y), z - x \rangle . \tag{6}$$

Following [1] we will say that $f$ is a Bregman function if the following assumptions hold:

B$_1$. The right level sets of $D_f(x, \cdot)$:

$$R^f_\alpha(x) = \{y \in \operatorname{int} \mathcal{D} : D_f(x, y) \leq \alpha\}$$

are bounded, for all $\alpha \in \mathbb{R}$ and $x \in \mathcal{D}$.

$B_2$. If $\{x_k\} \subseteq \operatorname{int}\mathcal{D}$ and $\{y_k\} \subseteq \operatorname{int}\mathcal{D}$ converge weakly to $x$ and $\lim_{k\to+\infty} D_f(x_k, y_k) = 0$, then

$$\lim_{k\to+\infty} \left(D_f(x, x_k) - D_f(x, y_k)\right) = 0.$$

$B_3$. If $\{x_k\} \subseteq \mathcal{D}$ is bounded, $\{y_k\} \subseteq \operatorname{int}\mathcal{D}$ is such that w-$\lim_{k\to+\infty} y_k = y$ and $\lim_{k\to+\infty} D_f(x_k, y_k) = 0$, then w-$\lim_{k\to+\infty} x_k = y$.

For example, in $X = L^p$ or $l_p$ with $p \in (0,1)$, $f(x) = \|x\|_p^p$ is a Bregman function. A Bregman function in $\mathbb{R}^n$ whose domain is the positive orthant $\mathbb{R}^n_+$ is the Shannon function $h(x) = \sum\limits_{i=1}^{n} x_i \ln x_i$ with the convention $0 \ln 0 = 0$.

Let a Bregman function $f$ be given. Assume that $K \subseteq D(T) \cap \operatorname{int}\mathcal{D}$. Consider the following proximal point algorithm.

**1. Initialization.**
Choose $x_0 \in K$.

**2. Solution of a simpler VIP.**
Given $x_k$, define $x_{k+1}$ by the inclusion

$$0 \in (\mu_k T + N_K + \nabla f)\, x_{k+1} - \nabla f(x_k) \tag{7}$$

for some $\mu_k > 0$.

**3. Stoping rule.**
If $x_{k+1} = x_k$, then stop. Otherwise set $k = k+1$ and go to Step 2.

Note that (7) can be stated equivalently as follows: $x_{k+1}$ is a solution of the variational inequality: for some $x^*_{k+1} \in T(x_{k+1})$,

$$\left\langle \mu_k x^*_{k+1} + \nabla f(x_{k+1}) - \nabla f(x_k), y - x_{k+1} \right\rangle \geq 0, \quad \forall y \in K. \tag{8}$$

If the algorithm stops at $x_k$, then it is obvious that $x_k$ is a solution of VIP. Otherwise, the goal is to show that the infinite sequence provided by the algorithm (or a subsequence of it) converges to a solution of VIP. For this, we will use the following theorem of [1].

**Theorem 12** *Assume that the sequence generated by the algorithm is well defined and infinite. If $X^* \neq \emptyset$, then the following hold:*
*(i) The sequence $\{x_k\}$ is bounded,*
*(ii) $\sum_{k=0}^{+\infty} D_f(x_{k+1}, x_k) < +\infty$,*
*(iii) For each $\bar{x} \in S(T,K)$, $D_f(\bar{x}, x_k)$ converges.*

The following theorem generalizes Theorem 3.4.($A_1$) of [1].

**Theorem 13** *Assume that $S(T,K)$ is nonempty, that $\mu_k > \mu$ for some $\mu > 0$, and that the algorithm produces a well-defined infinite sequence. If $T$ is t-pseudomonotone and pseudomonotone$_*$, then every weak limit point of $\{x_k\}_{k\in\mathbb{N}}$ belongs to $S(T,K)$.*

**Proof.** We follow mainly the arguments of [1]. By Theorem 12, the sequence $\{x_k\}_{k\in\mathbb{N}}$ is bounded. Let $\left\{x_{k_j}\right\}_{j\in\mathbb{N}}$ be a subsequence of $\{x_k\}_{k\in\mathbb{N}}$, weakly converging to some point $x$. By Theorem 12(ii), $\lim_{k_j\to+\infty} D_f(x_{k_j+1}, x_{k_j}) = 0$. By condition $B_3$, w-$\lim x_{k_j+1} = x$.

Using $B_2$ we deduce that

$$\lim_j \left(D_f\left(x, x_{k_j+1}\right) - D_f\left(x, x_{k_j}\right)\right) = 0. \tag{9}$$

By using successively (8), the three-point property (6) and (9), we get

$$\begin{aligned}\liminf_j \mu_{k_j}\left\langle x^*_{k_j+1}, x - x_{k_j+1}\right\rangle &\geq \liminf_j \left\langle \nabla f(x_{k_j}) - \nabla f(x_{k_j+1}), x - x_{k_j+1}\right\rangle \\ &= \liminf_j \left(D_f\left(x, x_{k_j+1}\right) - D_f\left(x, x_{k_j}\right) + D_f\left(x_{k_j+1}, x_{k_j}\right)\right) = 0.\end{aligned}$$

Since $\mu_{k_j} > \mu$, we arrive at $\liminf_j \left\langle x^*_{k_j+1}, x - x_{k_j+1}\right\rangle \geq 0$. By t-pseudomonotonicity, given $\bar{x} \in S(T, K)$ there exists $x^* \in T(x)$ such that

$$\langle x^*, x - \bar{x}\rangle \leq \liminf_j \left\langle x^*_{k_j+1}, x_{k_j+1} - \bar{x}\right\rangle. \tag{10}$$

Using successively: the fact that $\bar{x}$ is a solution of MVIP, (8) and the three-point property, we deduce

$$\begin{aligned}0 \leq \mu_{k_j}\left\langle x^*_{k_j+1}, x_{k_j+1} - \bar{x}\right\rangle &\leq \left\langle \nabla f(x_{k_j+1}) - \nabla f\left(x_{k_j}\right), \bar{x} - x_{k_j+1}\right\rangle \\ &= D_f\left(\bar{x}, x_{k_j}\right) - D_f\left(\bar{x}, x_{k_j+1}\right) - D_f\left(x_{k_j+1}, x_{k_j}\right).\end{aligned}$$

By theorem 12(iii) the sequence $\{D_f\left(\bar{x}, x_k\right)\}$ is converging. Hence the above relation implies that $\lim_j \left\langle x^*_{k_j+1}, x_{k_j+1} - \bar{x}\right\rangle = 0$. Combining with (10), we arrive at $\langle x^*, x - \bar{x}\rangle \leq 0$. Finally, using that $T$ is pseudomonotone$_*$, we infer from (CPP) that $x \in S(T, K)$. ■

If further $\nabla f$ is weakly continuous, then it can be shown by standard arguments (see for instance Theorem 3.5 in [1]) that the whole sequence $\{x_k\}$ weakly converges to a solution of VIP.

**Acknowledgement.** The authors should like to thank the referee for his/her very useful comments and suggestions. The research of the first author was supported by the grant no. 227-$\varepsilon$ of GSRT, Greece.

# References

[1] Burachik, S.R., Scheimberg, S.: A proximal point method for the variational inequality problem in Banach spaces. SIAM J. Optim. 39, 1633-1649 (2001).

[2] Bruck, R.E. Jr.: An iterative solution of a variational inequality for certain monotone operators in Hilbert space. Research announcement. Bull. Amer. Math. Soc. 81, 890-892 (1976); Bruck, R.E. Jr.: Corrigendum. Bull. Amer. Math. Soc. 82, 353 (1976).

[3] Censor, Y., Iusem, A.N., Zenios, S.A.: An interior point method with Bregman functions for the variational inequality problem with paramonotone operators. Math. Program. 81, 373-400 (1998).

[4] Crouzeix, J.P., Marcotte, P., Zhu, D.: Conditions ensuring the applicability of cutting-plane methods for solving variational inequalities. Math. Program. 88, 521-539 (2000).

[5] Denault, M., Goffin, J.L.: On a primal-dual analytic center cutting plane method for variational inequalities. Comput. Optim. Appl. 12, 127–155 (1999).

[6] Eckstein, J., Ferris, M.C.: Smooth methods of multipliers for complementarity problems. Math. Program. 86, 65-90 (1999).

[7] Ferris, M.C., Mangasarian, O.L.: Minimum principle sufficiency. Math. Program. 57, 1-14 (1992).

[8] Ferris, M.C., Pang, J.S.: Nondegenerate Solutions and Related Concepts in Affine Variational Inequalities. SIAM J. Control Optim. 34, 244-263 (1996).

[9] Goffin, J.L., Marcotte, P., Zhu, D.: An analytic center cutting plane method for pseudomonotone variational inequalities. Oper. Res. Lett. 20, 1-6 (1997).

[10] Goffin, J.L., Vial, J.P.: Interior point methods for nondifferentiable optimization. In: Kischka, P. et al (eds) Operations Research Proceedings 1997, pp. 35-49. Springer, Berlin Heidelberg New York (1998).

[11] Hadjisavvas, N.: Continuity and Maximality Properties of Pseudomonotone Operators, J. Convex Anal. 10, 465-475 (2003).

[12] Hadjisavvas, N.: Maximal pseudomonotone operators. In: Crespi, G.P., Guerraggio, A., Miglierina, E., Rocca, M. (eds) Proceedings of the Workshop 'Optimization', pp. 87-100. Datanova Editrice, Milano (2003).

[13] Hadjisavvas, N., Schaible, S.: On a generalization of paramonotone maps and its application to solving the Stampacchia variational inequality. Optimization 55, 593-604 (2006).

[14] Iusem, A.N.: On some properties of paramonotone operators. J. Convex Anal. 5, 269-278 (1998).

[15] Marcotte, P., Zhu, D.: Weak sharp solutions of variational inequalities. SIAM J. Optim., 9, 179-189 (1998).

[16] Solodov, M.V., Svaiter, B.F.: An inexact hybrid generalized proximal point algorithm and some new results on the theory of Bregman functions. Math. Oper. Res. 25, 214-230 (2000).

[17] Wu, Z., Wu, S.-Y.: Weak sharp solutions of variational inequalities in Hilbert spaces. SIAM J. Optim. 14, 1011-1027 (2004).

[18] Yao, J.C, Chadli, O.: Pseudomonotone complementarity problems and variational inequalities. In: Hadjisavvas, N., Komlosi, S., Schaible, S. (eds.) Handbook of generalized convexity and generalized monotonicity, pp. 501-558. Springer, Berlin Heidelberg New York (2005).