# BILEVEL OPTIMIZATION AS A REGULARIZATION APPROACH TO PSEUDOMONOTONE EQUILIBRIUM PROBLEMS

**Bui V. Dinh**[1], **Pham G. Hung**[2], **and Le D. Muu**[3]

[1]*Le Quy Don Technical University, Hanoi, Vietnam*
[2]*Nha Trang University, Nha Trang, Vietnam*
[3]*Institute of Mathematics, VAST, Vietnam*

**Abstract**

We investigate some properties of an inexact proximal point method for pseudomonotone equilibrium problems in a real Hilbert space. Unlike monotone case, in pseudomonotone case, the regularized subproblems may not be strongly monotone, even not pseudomonotone. However, every proximal trajectory weakly converges to the same limit, We use these properties to extend a viscosity-proximal point algorithm developed in [29] to pseudomonotone equilibrium problems. Then we propose a hybrid extragradient-cutting plane algorithm for approximating the limit point by solving a bilevel strongly convex optimization problem. Finally, we show that by using this bilevel convex optimization, the proximal point method can be used for handling ill-possed pseudomonotone equilibrium problems.

**Keywords**: Pseudomonotone equilibrium problem, inexact proximal point, bilevel optimization, hybrid extragradient-cutting algorithm, regularization.
**Mathematics Subject Classifications**: 49J40, 90C33, 47H17.

# 1 INTRODUCTION

Throughout this article we assume that $\mathcal{H}$ is a real Hilbert space whose inner product and the associated norm are denoted by $\langle .,. \rangle$ and $\|.\|$ respectively. We say that a sequence $\{x^k\} \subset \mathcal{H}$ weakly converges to a vector $x \in \mathcal{H}$, and write $x^k \rightharpoonup x$, if $\{x^k\}$ converges to $x$ in the weak topology. Suppose that $C \subseteq \mathcal{H}$, is a nonempty closed convex set, and that $f : \mathcal{H} \times \mathcal{H} \to \mathbb{R} \cup \{+\infty\}$ satisfying $f(x,x) = 0$ for every $x \in C$, As usual, we call such a function $f$ a *equilibrium bifunction* on $C$.

In this paper we are concerned with the following equilibrium problem, shortly EP, which is also often called the Ky Fan inequality,

$$\text{Find } x \in C : f(x,y) \geq 0 \ \forall y \in C, \qquad (EP)$$

Address correspondence to Le Dung Muu, Institute of Mathematics, VAST, Vietnam; Email: ldmuu@math.ac.vn

EP is very wide in the sense that it contains some important problems such as optimization, variational inequalities, Kakutani fixed point, saddle point and Nash equilibriumm models as special cases (see e.g. [5, 21]). It is well known that EPs, in general, are ill-posed in the sense [30] that they are not uniquely solvable and the solutions do not depend continuously on the data.

The proximal point method, PPM for short, is a fundamental regularization technique for handling ill-posed problems. PPM first introduced by Martinet in [?] for monotone variational inequality and further extended by Rockafellar to maximal operator inclusions. Recently, the PPM is applied by Moudafi [18] to monotone equilibrium problems. The PPM applying to EPs consists of constructing iteratively regularized equilibrium subproblems of the form

$$\text{Find } x \in C : f_k(x,y) := f(x,y) + c_k g_k(x,y) \geq 0 \ \forall y \in C, \qquad (REP)$$

where $c_k > 0$ and $g_k$ are regularization parameter and regularization bifunction respectively. Usually, $g_k(x,y) = \langle x - x^{k-1}, y - x \rangle$, where $x^{k-1}$ is the iterate obtained at iteration $k-1$. In the case $f$ is monotone and $g_k$ is strongly monotone, thanks to the fact that the sum of a monotone and a strongly monotone bifunctions is strongly monotone, under certain assumptions on continuity of $f$, Problem $(REP)$ uniquely solvable and its solution tends to a solution of the original problem whenever $c_k$ is bounded below from zero, However, when $f$ is a generalized monotone bifunction, for instance, pseudomonotone, the sum of $f$ and a strongly monotone, in general, does not inherit any monotonicity property, and therefore the regularized subproblems, in general, are not uniquely solvable, even the solution sets are not convex. This fact might seem that one cannot directly apply PPM for handling ill-posed pseudomonotone EPs as in the case of monotonicity. In our recent paper [10], we have studied the Tikhonov regularization method for pseudomonotone EPs. There we have shown that, under certain mild assumptions, every Tikhonov trajectory has the same limit point which is the unique solution of the EP defined by the regularization bifunction and the solution set of the original problem. For the proximal point method the latter fact does not hold because, unlike the Tikhonov method, in the proximal point method, the regularization bifunction is updated at each iteration and the convergence, in general, is not strong.

The purpose this paper is three folds. First, we study an inexact proximal point method for pseudomonotone equilibrium problems in real Hilbert spaces. In this case, the regularized subproblems may not be strongly monotone, even not pseudomonotone. However, they are uniquely solvable at infinity, in the sense that every proximal trajectory has the same weakly limit, In order to make the convergence strong, and to show that the limit point is just the unique solution of the bilevel optimization whose objective function is the norm and the feasible domain is the solution set of the original problem (EP), we use the obtained result to extend a viscosity-proximal point algorithm developed in [29] to pseudomonotone EPs. Next, motivated by the fact that the subproblems in this algorithm cannot be solved by existing algorithms for EPs, we propose a hybrid extragradient-cutting plane algorithm using an Armijo linesearch for solving the resulting bilevel optimization problem. Finally, by using the bilevel optimization approach, we show that PPM can be used for handling ill-possed pseudomonotone equilibrium problems.

The organization of this paper is as follows. In the next section we investigate sone properties of an inexact proximal point for pseudomonotone EPs. Then we use these properties to prove the strong convergence of a viscosity-proximal point algorithm for pseudomonotone EPs. The third section is devoted to description of a hybrid extragradient-cutting plane algorithm for solving bilevel optimization problem and to prove its strong convergence. In the last section we discuss some stability issues by using the bilevel optimization.

## 2 AN INEXACT PPM FOR PSEUDOMONOTONE EPS

First we recall the following well-known definitions on monotonicity (see e.g. [4, 21]).

**Definition 2.1.** The bifunction $f : C \times C \to \mathbb{R} \cup \{+\infty\}$ is said to be

a) *strongly monotone* on $C$ with modulus $\gamma > 0$ if

$$f(x, y) + f(y, x) \le -\gamma \|x - y\|^2, \ \forall x, y \in C;$$

b) *monotone* on $C$ if

$$f(x, y) + f(y, x) \le 0, \ \forall x, y \in C;$$

c) *pseudomonotone* on $C$ if

$$f(x, y) \ge 0 \Rightarrow f(y, x) \le 0, \ \forall x, y \in C.$$

The following implications are obvious from the definition

$$a) \Rightarrow b) \Rightarrow c).$$

In this section we occasionally make use of the following blanket assumptions:

($A_1$) $f(., y)$ is weakly upper semicontinuous on $\mathcal{H}$ (shortly w.u.s.c.) for each $y \in C$;

($A_2$) $f(x, .)$ is weakly lower semicontinuous (shortly w.l.s.c.), convex on $\mathcal{H}$ and subdifferentiable on dom$f(x, .)$ for each $x \in C$;

($A_3$) There exist a closed ball $B \subset \mathcal{H}$ and a vector $y^0 \in B \cap K$ such that

$$f(x, y^0) < 0, \ \forall x \in C \backslash B.$$

Assumption ($A_3$) is often called the coercivity property. Note that if Assumption ($A_2$) is satisfied, then by convexity of $f(x, .)$, the lower level set

$$\{y \in C : f(x, y) \le \alpha\}$$

is weakly closed and convex for every $\alpha$. Since, for a convex set, the weak closedness is equivalent to the closedness, the weakly lower semicontinuity of $f(x, .)$ is equivalent to lower semicontinuity of it. The following well-known propositions will be used in the next section.

**Proposition 2.1.** (See [4, Propositions 3.1, 3.2]).

a) *If $f$ satisfies Assumptions $(A_1), (A_2)$ and is strongly monotone on $C$, then $EP(K, f)$ has a unique solution.*

b) *If $f$ satisfies Assumptions $(A_1), (A_2)$ and is pseudomonotone on $C$, then the solution set of (EP) is weakly closed, convex.*

c) *If $f$ satisfies Assumptions $(A_1), (A_2)$ and $(A_3)$, then the solution set of (EP) is nonempty.*

**Proposition 2.2.** *Suppose that $f$ satisfies Assumptions $(A_1)$ and $(A_2)$. Consider the following statements*

a) *There exists a vector $y^0 \in C$ such that the set*

$$L(y^0, f) := \{x \in C : f(x, y^0) \geq 0\}$$

*is bounded.*

b) *There exist a closed ball $B \subset \mathcal{H}$ and a vector $y^0 \in C \cap B$ such that*

$$f(x, y^0) < 0, \forall x \in C \setminus B.$$

c) *The solution set $S(C, f)$ of (EP) is nonempty and weakly compact.*

*It holds that a) $\Rightarrow$ b) $\Rightarrow$ c). In addition, if $f$ is pseudomonotone on $C$, then $S(C, f)$ is convex and the set*

$$L_{>}(y^0, f) := \{x \in C : f(x, y^0) > 0\}$$

*is empty for any $y^0 \in S(C, f)$.*

*Proof.* a) $\Rightarrow$ b): By the assumption a), we can take $B$ as a closed ball containing $L(y^0, f)$. Then it is obvious

$$\{x \in C \setminus B : f(x, y^0) \geq 0\} = \emptyset.$$

Hence b) holds.

b) $\Rightarrow$ c): By Proposition 2.1.c) we have $S(C, f) \neq \emptyset$. Since $C$ is weakly closed and $f(., y)$ is weakly upper semicontinuous on $C$, the solution set $S(C, f)$ is weakly closed. Moreover, from b) and the definition of $L(y^0, f)$ follows

$$S(C, f) \subseteq L(y^0, f) \subseteq C \cap B.$$

Thus $S(C, f)$ is weakly compact.

To see the last assertion, let $y^0 \in S(C, f)$. Then $f(y^0, x) \geq 0$ for every $x \in C$. By pseudomonotonicity, it follows that $f(x, y^0) \leq 0$ for every $x \in C$. Hence $L_{>}(y^0, f) = \emptyset$. The convexity of $S(C, f)$ follows from Proposition 2.1.c). $\square$

The first assertion in the next lemma was proved by Noor [24] (see also [11]) for exact PPM.

**Lemma 2.1.** *Suppose that $f$ is pseudomonotone on $C$. Then for any $\varepsilon > 0$, $\delta \geq 0$ , $\bar{x} \in S(C, f), x(\varepsilon) \in S_\delta(C, f_\varepsilon)$ and $x^g \in C$, it holds that*

a) $\|x^g - x(\varepsilon)\|^2 + \|x(\varepsilon) - \bar{x}\|^2 \leq \|x^g - \bar{x}\|^2 + 2\frac{\delta}{\varepsilon}$.

b) $S_\delta(C, f_\varepsilon) \subset \overline{B}\left(0, \|\frac{\bar{x}+x^g}{2}\| + \sqrt{\|\frac{\bar{x}-x^g}{2}\|^2 + \frac{\delta}{\varepsilon}}\right) \cap C$.

c) $\|x(\varepsilon) - x^g\| \leq \|\frac{\bar{x}-x^g}{2}\| + \sqrt{\|\frac{\bar{x}-x^g}{2}\|^2 + \frac{\delta}{\varepsilon}}$.

*where $\overline{B}(x, r)$ stands for the closed ball around $x$ with radius $r$.*

*Proof.* Since $\bar{x} \in S(C, f)$, by the pseudomonotonicity of $f$, we have

$$f(\bar{x}, y) \geq 0 \Rightarrow f(y, \bar{x}) \leq 0, \ \forall y \in C. \tag{1}$$

As $x(\varepsilon) \in S_\delta(C, f_\varepsilon)$, it holds that

$$f(x(\varepsilon), y) + \varepsilon\langle x(\varepsilon) - x^g, y - x(\varepsilon)\rangle \geq -\delta, \ \forall y \in C. \tag{2}$$

Substituting $y = x(\varepsilon)$ into the second inequality in (1) and $y = \bar{x}$ in (2) we obtain

$$f(x(\varepsilon), \bar{x}) \leq 0 \ \text{ and } \ f(x(\varepsilon), \bar{x}) + \varepsilon\langle x(\varepsilon) - x^g, \bar{x} - x(\varepsilon)\rangle \geq -\delta.$$

From which we deduce that

$$\frac{1}{2}\left[\|x^g - \bar{x}\|^2 - \|x^g - x(\varepsilon)\|^2 - \|x(\varepsilon) - \bar{x}\|^2\right] = \langle x(\varepsilon) - x^g, \bar{x} - x(\varepsilon)\rangle \geq -\frac{\delta}{\varepsilon}.$$

Hence, a) holds true. On the other hand

$$\|x(\varepsilon) - x^g\|^2 + \|[x(\varepsilon) - x^g] - [\bar{x} - x^g]\|^2 \leq \|\bar{x} - x^g\|^2 + 2\frac{\delta}{\varepsilon},$$

which implies

$$\|x(\varepsilon) - x^g\|^2 - \langle x(\varepsilon) - x^g, \bar{x} - x^g\rangle \leq \frac{\delta}{\varepsilon}.$$

Thus

$$\begin{aligned}\left\|x(\varepsilon) - \frac{\bar{x} + x^g}{2}\right\|^2 &= \left\|x(\varepsilon) - x^g - \frac{\bar{x} - x^g}{2}\right\|^2 \\ &= \|x(\varepsilon) - x^g\|^2 - \langle x(\varepsilon) - x^g, \bar{x} - x^g\rangle + \left\|\frac{\bar{x} - x^g}{2}\right\|^2 \\ &\leq \left\|\frac{\bar{x} - x^g}{2}\right\|^2 + \frac{\delta}{\varepsilon}\end{aligned}$$

which proves b) and c). □

**Lemma 2.2.** *Suppose that $f$ is pseudomonotone on $C$ and satisfies Assumptions $(A_1)$ and $(A_2)$. If the solution set $S(C, f)$ is nonempty, then for any $\varepsilon > 0$, $\delta \geq 0$, the $\delta-$ solution set $S_\delta(C, f_\varepsilon)$ is nonempty and weakly compact.*

*Proof.* According to Proposition 2.2, it is sufficient to find a vector $y^0 \in C$ such that the set

$$L_\delta(y^0, f_\varepsilon) := \{x \in C : f_\varepsilon(x, y^0) = f(x, y^0) + \varepsilon\langle x - x^g, y^0 - x\rangle \geq -\delta\}$$

is bounded. Let $y^0 \in S(C, f)$ and $x \in L_\delta(y^0, f_\varepsilon)$. By definition of $L_\delta(y^0, f_\varepsilon)$, we have

$$f_\varepsilon(x, y^0) = f(x, y^0) + \varepsilon\langle x - x^g, y^0 - x\rangle \geq -\delta.$$

Using the inequality a) in Lemma 2.1 with $x(\varepsilon) = x, \bar{x} = y^0$ we obtain

$$\|x^g - x\|^2 + \|x - y^0\|^2 \leq \|x^g - y^0\|^2 + 2\frac{\delta}{\varepsilon}$$

which implies

$$\|x - x^g\| \leq \sqrt{\|y^0 - x^g\|^2 + 2\frac{\delta}{\varepsilon}}$$

Thus

$$\|x\| \leq \|x^g\| + \sqrt{\|y^0 - x^g\|^2 + 2\frac{\delta}{\varepsilon}}, \ \forall x \in L_\delta(y^0, f_\varepsilon).$$

That means the set $L_\delta(y^0, f_\varepsilon)$ is bounded. □

Now we investigate the behaviour of iterates defined by an inexact proximal point algorithm for Problem (EP) where $f$ is a pseudomonotone bifunction. Starting from a given point $x^0 \in C$, at each iteration $k = 1, 2, ...,$ we consider the subproblem $EP(C, f_k)$ given as

$$\begin{cases} \text{Find } x^k \in C \text{ such that} \\ f_k(x^k, y) := f(x^k, y) + c_k\langle x^k - x^{k-1}, y - x^k\rangle \geq -\delta_k, \ \forall y \in C \end{cases} \tag{3}$$

where the regularization parameter $c_k > 0$ and the tolerance $\delta_k \geq 0$ are given. As usual, we call a solution of (3) a $\delta_k-$ solution to $EP(C, f_k)$ and we denote the set of all $\delta_k$-solutions by $S_{\delta_k}(C, f_k)$. We call a sequence $\{x^k\}$ with $x^k \in S_{\delta_k}(C, f_k)$ a *proximal trajectory*. The following theorem says that, for pseudomonotone EPs, although the regularized subproblems are not uniquely solvable, every proximal trajectory has the same limit.

**Theorem 2.1.** *Suppose that $f$ is pseudomonotone on $C$, satisfies Assumptions $(A_1), (A_2)$, and that Problem (EP) admits a solution. Let $\{c_k\}$ and $\{\delta_k\}$ be two sequences of positive numbers such that $c_k \leq c < +\infty$ for every $k$, and $\sum_{k=1}^\infty \frac{\delta_k}{c_k} < +\infty$. Then*

a) *For every $k \in \mathbb{N}$, the solution set $S_{\delta_k}(C, f_k)$ is nonempty, closed, uniformly bounded, and it holds that*

$$\|x^{k-1} - x^k\|^2 + \|x^k - \bar{x}\|^2 \leq \|x^{k-1} - \bar{x}\|^2 + 2\frac{\delta_k}{c_k}, \tag{4}$$

*where $\bar{x} \in S(C, f)$ and $x^k \in S_{\delta_k}(C, f_{\varepsilon_k})$.*

b) *The sequence* $\{x^k\}$*, where* $x^k$ *is arbitrarily chosen in* $S_{\delta_k}(C, f_{\varepsilon_k})$*, weakly converges to a solution of (EP). Moreover, if* $\{x^k\}$ *has a strongly cluster point, then the whole sequence strongly converges to a solution of the original problem (EP).*

*Proof.* a) Using Lemma 2.2 with $x^g = x^{k-1} \in C$ and $\varepsilon = c_k > 0$, we see that, for every $k = 1, 2, ...,$ the solution set of Problem $EP(C, f_k)$ is nonempty, closed, uniformly bounded. To prove the inequality (4), just applying a) in Lemma 2.1 with

$$\varepsilon = c_k, x^g = x^{k-1}, x(\varepsilon) = x^k, \delta = \delta_k.$$

b) Fix any point $\bar{x}$ in the solution set of Problem (EP). Let $x^k \in S_{\delta_k}(C, f_k)$ with $k \geq 1$. From (4), one has

$$\|x^k - \bar{x}\|^2 \leq \|x^{k-1} - \bar{x}\|^2 + 2\frac{\delta_k}{c_k}. \tag{5}$$

Since $\sum_{k=1}^{+\infty} \frac{\delta_k}{c_k} < +\infty$, we have

$$\lim_{k\to\infty} \|x^k - \bar{x}\| = \mu < \infty. \tag{6}$$

Using again the inequality (4), we can write

$$\|x^k - x^{k-1}\|^2 \leq \|x^{k-1} - \bar{x}\|^2 - \|x^k - \bar{x}\|^2 + 2\frac{\delta_k}{c_k}.$$

Then, by (6) and $\frac{\delta_k}{c_k} \to 0$ as $k \to \infty$, we have

$$\lim_{k\to\infty} \|x^k - x^{k-1}\| = 0 \tag{7}$$

Now let

$$M := 2\sum_{j=1}^{\infty} \frac{\delta_j}{c_j} < +\infty$$

Then, from (5), it follows that

$$\begin{aligned}
\|x^k - \bar{x}\|^2 \leq \|x^g - \bar{x}\|^2 + 2\sum_{j=1}^{k} \frac{\delta_j}{c_j} \leq \|x^g - \bar{x}\|^2 + M \ \forall k \\
&\Rightarrow \|x^k - \bar{x}\| \leq \sqrt{\|x^g - \bar{x}\|^2 + M} \ \forall k \\
&\Rightarrow \|x^k\| \leq \|\bar{x}\| + \sqrt{\|x^g - \bar{x}\|^2 + M} \ \forall k \\
&\Rightarrow x^k \in S_{\delta_k}(K, f_k) \subset \overline{B}\left(0, \|\bar{x}\| + \sqrt{\|x^g - \bar{x}\|^2 + M}\right) \cap C \ \forall k.
\end{aligned}$$

So $\{x^k\}$ is bounded, and therefore there exists a subsequence $\{x^{k_j}\} \subseteq \{x^k\}$ such that

$$x^{k_j} \rightharpoonup x^* \in \overline{B}\left(0, \|\bar{x}\| + \sqrt{\|x^g - \bar{x}\|^2 + M}\right) \cap C.$$

Since $x^{k_j}$ is a $\delta_{k_j}-$ solution of $EP(C, f_{k_j})$ for every $k_j$, we have

$$f_{k_j}(x^{k_j}, y) = f(x^{k_j}, y) + c_{k_j}\langle x^{k_j} - x^{k_j-1}, y - x^{k_j}\rangle \geq -\delta_{k_j}, \ \forall y \in C. \tag{8}$$

Taking account of (7), the weakly upper semicontinuity of $f$ and the conditions $0 < c_{k_j} < c < +\infty$, $\delta_{k_j} \to 0^+$, we obtain from (8), in the limit, that

$$0 \leq \overline{\lim}_{k_j\to\infty} f_{k_j}(x^{k_j}, y) \leq \overline{\lim}_{k_j\to\infty} f(x^{k_j}, y) \leq f(x^*, y), \ \forall y \in C$$

which shows that $x^* \in S(C, f)$. Now, by using the same argument as in [27], we can show that $x^*$ is the uniquely weakly cluster point of $\{x^k\}$. In fact, suppose that $x_1^*$ and $x_2^*$ are two distinct weakly cluster points of $\{x^k\}$. Then $x_1^*, x_2^* \in S(C, f)$, as just we have seen. Then one can apply (6) with $x_i^*$ $(i = 1, 2)$ playing the role of $\bar{x}$ to obtain

$$\lim_{k\to\infty} \|x^k - x_i^*\| = \mu_i, \ i = 1, 2. \tag{9}$$

Clearly,

$$2\langle x^k - x_1^*, x_1^* - x_2^*\rangle = \|x^k - x_2^*\|^2 - \|x^k - x_1^*\|^2 - \|x_1^* - x_2^*\|^2. \tag{10}$$

As $x_1^*$ is a weakly cluster point of $\{x^k\}$, from (9) and (10) it follows that

$$0 = 2 \lim_{k\to\infty} \langle x^k - x_1^*, x_1^* - x_2^*\rangle = \mu_2^2 - \mu_1^2 - \|x_1^* - x_2^*\|^2.$$

Thus

$$\mu_2^2 - \mu_1^2 = \|x_1^* - x_2^*\|^2 > 0.$$

Changing the roles of $x_1^*$ and $x_2^*$ to each other, we also have $\mu_1^2 - \mu_2^2 > 0$. This contradiction asserts the uniqueness of $x^*$.

Now, suppose that the subsequence $\{x^{k_j}\} \subseteq \{x^k\}$ strongly converges to some $x^* \in \mathcal{H}$. Then $x^* \in S(C, f)$. Applying (5) to $\bar{x} = x^*$, we obtain

$$\|x^k - x^*\|^2 \leq \|x^{k-1} - x^*\|^2 + 2\frac{\delta_k}{c_k}, \ \forall k \in \mathbb{N}. \tag{11}$$

For any $\gamma > 0$, as $\lim_{k_j\to\infty} \|x^{k_j} - x^*\| = 0$ and $\sum_{k=1}^{\infty} \frac{\delta_k}{c_k} < +\infty$, one can find some $l \in \mathbb{N}$ such that

$$\|x^{k_l} - x^*\| \leq \frac{\gamma}{\sqrt{2}} \quad \text{and} \quad \sum_{i=k_l+1}^{\infty} \frac{\delta_i}{c_i} \leq \frac{\gamma^2}{4}.$$

Hence, for $k > k_l + 1$, from (11), it holds that

$$\begin{aligned}
\|x^k - x^*\|^2 &\leq \|x^{k-1} - x^*\|^2 + 2\frac{\delta_k}{c_k} \\
&\leq \|x^{k-2} - x^*\|^2 + 2\left(\frac{\delta_k}{c_k} + \frac{\delta_{k-1}}{c_{k-1}}\right) \\
&\leq ... \\
&\leq \|x^{k_l} - x^*\|^2 + 2\left(\frac{\delta_k}{c_k} + \frac{\delta_{k-1}}{c_{k-1}} + ... + \frac{\delta_{k_l+1}}{c_{k_l+1}}\right) \\
&\leq \frac{\gamma^2}{2} + \frac{\gamma^2}{2} = \gamma^2.
\end{aligned}$$

Thus

$$\|x^k - x^*\| \leq \gamma. \ \forall k > k_l + 1.$$

Since $\gamma > 0$ is arbitrary, we can conclude that $\lim_{k\to\infty} \|x^k - x^*\| = 0$ as required. □

**Theorem 2.2.** *Suppose that $C$ is a nonempty closed convex subset of $\mathbb{R}^n$, that $f$ is pseudomonotone on $C$, $f(.,y)$ is upper semicontinuous for each $y \in C$, $f(x,.)$ is lower semicontinuous and convex for each $x \in C$, and that the problem (EP) admits a solution. Let $\{c_k\}$ and $\{\delta_k\}$ be two sequences of positive numbers such that $c_k < c < +\infty$, and $\sum_{k=1}^{\infty} \frac{\delta_k}{c_k} < +\infty$. Then*

a) *For any $k$, the $\delta_k-$solution set of Problem $EP(C, f_k)$ is nonempty and compact.*

b) *The sequence $\{x^k\}$, with $x^k$ being any $\delta_k-$solution of Problem $EP(C, f_k)$, strongly converges to some solution of (EP).*

*Proof.* Applying Theorem 2.1 and the property that, any bounded sequence in the space $\mathbb{R}^n$ must have a strongly convergent subsequence, we obtain the desired result. □

# 3 A HYBRID PROX-CUTTING ALGORITHM FOR PSEUDOMONOTONE EPS

The results obtained in the previous section show that every proximal trajectory has the same weakly limit point. However finding this limit point is a difficult task, since the convergence is not strong and the above results do not locate the limit point. In this section we use an iterative hybrid proximal point-cutting algorithm to force strong convergence and to locate the limit point for pseudomonotone EPs. This algorithm, which relies on the exact PPM, has been introduced in [29] for finding a solution of a monotone equilibrium problem which is also a fixed point of a nonexpansive mapping. This algorithm is modified for our problem as follows.

Suppose that $\delta_n \geq 0$ for every $n$, and that we are given a guessed solution $x^g \in C$. Starting from $x^1 := x^g$ the algorithm iteratively constructs two sequences $\{x^n\}$, $\{u^n\}$ satisfying

$$\begin{cases} f_n(u^n, y) := f(u^n, y) + c_n\langle y - u^n, u^n - x^n\rangle \ \geq -\delta_n \ \forall y \in C, \\ x^{n+1} := P_{B_n}(x^g) \end{cases} \tag{12}$$

with $B_n := C_n \cap D_n$, where $C_n$ and $D_n$ are the half spaces defined as

$$C_n := \{z \in \mathcal{H} : \ \|u^n - z\|^2 \leq \|x^n - z\|^2 + 2\frac{\delta_n}{c_n}\}, \tag{13}$$

$$D_n := \{z \in \mathcal{H} : \ \langle x^n - z, x^g - x^n\rangle \geq 0\} \tag{14}$$

The following convergence theorem whose proof follows some techniques used in the proof of Theorem 3.2 in [29]. However we emphasize that some parts of the proof in [29] is based upon the fact that when $f$ is monotone on $C$, the mapping $T_r : \mathcal{H} \to C$ defined by

$$T_r(x) := \{z \in C : f(z, y) + r\langle y - z, z - x\rangle \geq 0 \ \forall y \in C\}$$

with $r > 0$, is single valued and firmly nonexpansive, i.e.,

$$\|T_r(x) - T_r(y)\|^2 \leq \langle T_r(x) - T_r(y), x - y\rangle.$$

However when $f$ is pseudomonotone, $T_r(x)$ is not singleton, even not convex, and $T_r$ is not expansive. So the techniques in the proof of Theorem 3.2 in [29] that use these properties of $T_r$ can not be apply to our setting.

In the proof of the theorem below, we need the following lemma.

**Lemma 3.1.** *Suppose that $S \subset \mathcal{H}$ is a closed, convex set, $z^g \in \mathcal{H}$ and the sequence $\{z^n\} \subset \mathcal{H}$ satisfies the conditions*

$$\|z^n - z^g\| \leq \|s - z^g\| \ \forall n, \ \forall s \in S. \tag{15}$$

*and every weakly cluster point belongs to $S$. Then $\{z^n\}$ strongly converges to $P_S(z^g)$.*

*Proof.* By (15), the sequence $\{z^n\}$ is bounded. Let $z^*$ be any weakly cluster point of $\{z^n\}$. We may assume that $z^n \rightharpoonup z^* \in S$. Then, by l.s.c of the convex function $\|z^g - .\|$, from the condition (15) applying to $s = P_S(x^g)$, it follows that

$$\|z^g - P_S(x^g)\| \leq \|z^g - z^*\| \leq \underline{\lim}\|z^g - z^n\|$$
$$\leq \overline{\lim}\|z^g - z^n\| \leq \|z^g - P_S(z^g)\| \leq \|z^g - z^*\|$$

which implies $\lim \|z^g - z^n\| = \|z^g - z^*\| = \|z^g - P_S(z^g)\|$. Since $\mathcal{H}$ is a Hilbert space, by Opial's Theorem, we have $z^* = P_S(z^g)$ and $z^n \to P_S(z^g)$. □

**Theorem 3.1.** *Suppose that $f$ is pseudomonotone on $C$, that (EP) admits a solution and that $\infty > c' > c_n > c > 0$ for every $n$, $\sum_{n=1}^{\infty} \delta_n < \infty$. Then under Assumptions $(A_1)$, $(A_2)$, both sequences $\{u^n\}$, $\{x^n\}$ defined by (12) strongly convergence to the projection of $x^g$ onto the solution set $S(C,f)$ of Problem (EP).*

*Proof.* The theorem is proved through several claims.

*Claim 1:* $S(C,f) \subseteq B_n$ for every $n$ and

$$\|x^n - x^g\| \to \tau < +\infty \ \text{ as } n \to +\infty.$$

Indeed, apply Lemma 2.1 to (12) with $x^g := x^n$, $x^{k+1} := u^n$ and $\epsilon := c_k$, for every solution $x^*$ of Problem (EP), one has

$$\|x^n - u^n\|^2 + \|u^n - x^*\|^2 \leq \|x^n - x^*\|^2 + 2\frac{\delta_n}{c_n} \tag{16}$$

from which it follows that

$$\|u^n - x^*\|^2 \leq \|x^n - x^*\|^2 + 2\frac{\delta_n}{c_n}.$$

Hence, $S(C,f) \subseteq C_n$ for every $n$. To see that $S(C,f) \subseteq D_n$, we observe that $D_1 \equiv \mathcal{H}$. By induction, suppose that $S(C,f) \subseteq D_k$. Then, since $x^{k+1} = P_{D_k}(x^g)$, by induction, we can see easily that $S(C,f) \subseteq D_{k+1}$.

Using again $x^{n+1} = P_{B_n}(x^g)$, and the fact that $S(C,f) \subseteq B_n$, for any $x^* \in S(C,f)$, one has

$$\|x^{n+1} - x^g\| \leq \|x^* - x^g\| \ \forall n, \tag{17}$$

which implies that the sequence $\{x^n\}$ is bounded. Then, since $\frac{\delta_k}{c_k} \to 0$, by (16), the sequence $\{u^n\}$ is bounded too. Moreover, since $x^{n+1} \in D_n$, by definition of $D_n$, we have $x^n = P_{D_n}(x^g)$. Thus,

$$\|x^n - x^g\| \leq \|x^{n+1} - x^g\| \ \forall n$$

which together with boundedness of $\{x^n\}$ implies that

$$\|x^n - x^g\| \to \tau < +\infty \ \text{ as } n \to +\infty.$$

*Claim 2:* The sequence is asymptotical regular, i.e., $\|x^{n+1} - x^n\| \to 0$ as $n \to +\infty$, and $\|x^n - u^n\| \to 0$. Indeed, since $x^n \in D_n$ and $x^{n+1} \in D_n$, by convexity of $D_n$, one has $\frac{x^{n+1}+x^n}{2} \in D_n$. Then from $x^n = P_{D_n}(x^g)$ by using the strong convexity of the function $\|x^g - .\|^2$ one can write

$$\begin{aligned}
\|x^g - x^n\|^2 &\leq \|x^g - \frac{x^{n+1} + x^n}{2}\|^2 \\
&= \|\frac{x^g - x^{n+1}}{2} + \frac{x^g - x^n}{2}\|^2 \\
&= \frac{1}{2}\|x^g - x^{n+1}\|^2 + \frac{1}{2}\|x^g - x^n\|^2 - \frac{1}{4}\|x^{n+1} - x^n\|^2
\end{aligned} \tag{18}$$

which implies that

$$\frac{1}{2}\|x^{n+1} - x^n\|^2 \leq \|x^g - x^{n+1}\|^2 - \|x^g - x^n\|^2.$$

Remember that $\lim \|x^n - x^g\|$ does exist, we obtain $\|x^{n+1} - x^n\| \to 0$ as $n \to \infty$. Moreover, according to the algorithm $x^{n+1} \in B_n \subset C_n$, by definition of $C_n$, one has

$$\|u^n - x^{n+1}\|^2 \leq \|x^n - x^{n+1}\|^2 + 2\frac{\delta_n}{c_n}.$$

Thus, since $\frac{\delta_n}{c_n} \to 0$ and $\|x^{n+1} - x^n\| \to 0$ as $n \to \infty$, we have $\|u^n - x^{n+1}\| \to 0$. On the other hand

$$\|u^n - x^n\| \leq \|u^n - x^{n+1}\| + \|x^{n+1} - x^n\|.$$

Hence $\|u^n - x^n\| \to 0$ as $n \to \infty$.

*Claim 3:* Any weakly cluster point of $\{x^n\}$ and $\{u^n\}$ is a solution of Problem (EP). Indeed, let $\overline{u}$ be any weakly cluster point of $\{u^n\}$. By taking subsequences, if necessary, we may assume that $u^n \rightharpoonup \overline{u}$ and $c_n \to \overline{c} > c > 0$ as $n \to \infty$. From the definition of $u^n$, it follows that

$$f(u^n, x) \geq c_n\langle u^n - x, u^n - x^n\rangle - \delta_n \ \forall x \in C.$$

Since $\|x^n - u^n\| \to 0$, we have that $x^n \rightharpoonup \overline{u}$. Moreover, letting $n \to \infty$ in the above inequality we obtain $\underline{\lim} f(u^n, x) \geq 0$ for every $x \in C$. Then, by weak upper semicontinuity of $f(.,x)$, we get $f(\overline{u},x) \geq 0$ for every $x \in C$, which implies that $\overline{u}$ solves (EP). Note that, by $\|u^n - x^n\| \to 0$, we can conclude that the sets of all weakly limit points of $\{x^n\}$ and $\{u^n\}$ are coincided.

*Claim 4:* Both sequences $\{x^n\}$ and $\{u^n\}$ strongly converge to the projection of the guessed solution $x^g$ onto the solution set $S(C,f)$ of Problem (EP). To see this claim, we observe, by (17), that

$$\|x^{n+1} - x^g\| \leq \|x^* - x^g\| \ \forall n, \ \forall x^* \in S(C,f).$$

Thus we can apply Lemma 3.1 to the sequence $\{x^n\}$ with $S := S(C,f)$, $z^g = x^g$ to obtain $x^n \to P_S(x^g)$. Then, since $\|u^n - x^n\| \to 0$, we have $z^n \to P_S(x^g)$ as well.

□

# 4 A BILEVEL OPTIMIZATION APPROACH

As we have mentioned, for monotone EPs, thanks to the strongly monotone of the regularized subproblems, the algorithm described in the previous section can be implemented by the available methods. For pseudomonotone EPs, however the regularized subproblems, in general, are not strongly monotone, even not pseudomonotone, thus the available methods that use any monotonicity property can not be applied. For this case, from the fact that the limit point is just the projection of the guessed solution $x^g$ onto the solution set of Problem (EP), it suggests that the limit point can be obtained by solving the bilevel optimization problem

$$\{\min \|x - x^g\|^2 : \ x \in S(C,f)\}. \qquad (BO)$$

It is well-known that, when $f$ is pseudomotone on $C$, the solution set $S(C,f)$ of Problem (EP) is a convex set. Thus (BO) is the problem of minimizing the norm over a convex set, which is independent of the regularization parameters $\{c_k\}$ in the PPM. The main difficulty here is that the feasible set $S(C,f)$ is not given in an explicit form as in a standard optimization problem. A penalty function method was proposed in [19] for the bilevel problem (BO) when the lower equilibrium problem (EP) is monotone. In [7] a gap function algorithm was proposed for a certain class of pseudomonotone EPs. Foundations of bilevel programming can be found in the monograph [6]. In this section, combining the extragradient method first introduced by [15] with the cutting plane technique used in the previous section, we propose an algorithm for solving Problem (BO) when the lower equilibrium problem (EP) is pseudomonotone. Note that, unlike the monotone case, in this case, the penalized equilibrium subproblems in [19], are not monotone, even not pseudomonotone.

In what follows we suppose that the solution set $S(C,f)$ of Problem (EP) is nonempty and that $f$ is weakly continuous, pseudomonotone on $C$

Following the auxiliary problem principle (see e.g. [16]) let us define a bifunction $L : \mathcal{H} \times \mathcal{H} \to I\!\!R$ satisfying

(B1) $L(x,x) = 0, \ \exists \beta > 0 : \ L(x,y) \geq \frac{\beta}{2}\|x-y\|^2, \forall x, y \in C;$

(B2) $L$ is weakly continuous, $L(x,.)$ is differentiable, strongly convex on $\mathcal{H}$ for every $x \in C$ and $\nabla_2 L(x,x) = 0$ for every $x \in \mathcal{H}$.

An example for such a bifunction is the Bregman distance (see e.g. [8])

$$L(x,y) := g(y) - g(x) - \langle \nabla g(x), y - x\rangle$$

with $g$ being any differentiable, strongly convex function on $\mathcal{H}$ with modulus $\beta > 0$, particularly, $g(x) = \frac{1}{2}\|x\|^2$.

The following lemma is well known from the auxiliary problem principle for EPs.

**Lemma 4.1.** *([16]) Suppose that $f$ satisfies (A1), (A2) and $L$ satisfies (B1), (B2). Then, for every $\rho > 0$, the following statements are equivalent:*

*a) $x^*$ is a solution to (EP);*

*b) $x^* \in C : f(x^*, y) + \frac{1}{\rho}L(x^*, y) \geq 0,\ \forall y \in C$;*

*c) $x^* = argmin\{f(x^*, y) + \frac{1}{\rho}L(x^*, y) : y \in C\}$.*

Now we describe a hybrid extragradient-cutting plane algorithm for solving bilevel problem (BO).

**Algorithm 1**. Choose $\rho > 0$ and $\eta \in (0,1)$. Starting from $x^1 := x^g \in C$ ($x^g$ plays the role of a guessed solution). If $x^1 \in S(C,f)$, terminate: $x^1$ is a solution of Problem (BO). Otherwise, perform iteration $k$ with $k = 1$.

*Iteration* $k(k = 1, 2, ...)$ Having $x^k$ do the following steps:

*Step 1.* Solve the strongly convex program

$$\min\{f(x^k, y) + \frac{1}{\rho}L(x^k, y) : y \in C\} \qquad CP(x^k)$$

to obtain its unique solution $y^k$.

If $y^k = x^k$, take $u^k := x^k$ and go to Step 3. Otherwise, go to Step 2.

*Step 2* (Armijo linesearch) Find $m_k$ as the smallest nonnegative integer number $m$ satisfying

$$z^{k,m} := (1 - \eta^m)x^k + \eta^m y^k, \tag{19}$$

$$f(z^{k,m}, y^k) + \frac{1}{\rho}L(x^k, y^k) \leq 0. \tag{20}$$

Set $\eta_k = \eta^{m_k}$, $z^k := z^{k,m_k}$ and compute

$$\sigma_k = \frac{-\eta_k f(z^k, y^k)}{(1-\eta_k)\|g^k\|^2},\ u^k := P_C(x^k - \sigma_k g^k), \tag{21}$$

where $g^k \in \partial_2 f(z^k, z^k)$, the subgradient of the convex function $f(z^k, .)$ at $z^k$.

*Step 3.* Having $x^k$ and $u^k$ construct two halfspaces

$$C_k := \{y \in \mathcal{H} : \|u^k - y\|^2 \leq \|x^k - y\|^2\};$$

$$D_k := \{y \in \mathcal{H} : \langle x^g - x^k, y - x^k\rangle \leq 0\}.$$

*Step 4.* Set $B_k := C_k \cap D_k$ and compute $x^{k+1} := P_{B_k}(x^g)$.

If $x^{k+1} \in S(C,f)$, terminate: $x^{k+1}$ solves the bilevel problem (BO). Otherwise, increase $k$ by one and go to iteration $k$.

**Remark 4.1.** (i) The linesearch in Step 2 is well defined. Indeed, otherwise, for all nonnegative integer numbers $m$ one has

$$f(z^{k,m}, y^k) + \frac{1}{\rho}L(x^k, y^k) > 0. \tag{22}$$

Thus letting $m \to \infty$, by weakly upper semicontinuity of $f(., y^k)$, we have

$$f(x^k, y^k) + \frac{1}{\rho}L(x^k, y^k) \geq 0,$$

which, together with

$$f(x^k, x^k) + \frac{1}{\rho}L(x^k, x^k) = 0,$$

implies that $x^k$ is the solution of the strongly convex program $CP(x^k)$. Thus $x^k = y^k$ which contradicts to the fact that the linesearch is performed only when $y^k \neq x^k$.

Note that $m_k > 0$. Indeed, if $m_k = 0$, then, by the Armijo rule, we have $z^k = y^k$, and therefore

$$\frac{1}{\rho}L(x^k, y^k) = f(z^k, y^k) + \frac{1}{\rho}L(x^k, y^k) \leq 0,$$

which, together with nonnegativity of $L$, implies $L(x^k, y^k) = 0$. Since

$$L(x^k, y^k) \geq \frac{\beta}{2}\|x^k - y^k\|^2,$$

one has $x^k = y^k$.

(ii) $g^k \neq 0$ and the step size $\sigma_k$ defined by (21) is positive whenever $x^k \neq y^k$. Indeed, if $g^k = 0$, then, since $g^k \in \partial_2 f(z^k, z^k)$, we have

$$f(z^k, x) \geq \langle g^k, x - z^k\rangle + f(z^k, z^k) = 0 \ \forall x \in C$$

which implies that $z^k$ solves (EP). Then by (20), $L(x^k, y^k) \leq 0$. But, from Assumption (B1), it follows that $L(x^k, y^k) \geq \frac{\beta}{2}\|x^k - y^k\|^2$. Hence $x^k = y^k$.

**Lemma 4.2.** *Under the assumptions of Lemma 4.1, it holds that*

$$\|u^k - x^*\|^2 \leq \|x^k - x^*\|^2 - \sigma_k^2\|g^k\|^2, \ \forall x^* \in S(C, f), \ \forall k. \tag{23}$$

**Proof.** For simplicity of notation, we write $v^k$ for $x^k - \sigma_k g^k$. Since $u^k = P_C(v^k)$, by nonexpansiveness of the projection, we have

$$\begin{aligned} \|u^k - x^*\|^2 &= \|P_C(v^k) - P_C(x^*)\|^2 \leq \|v^k - x^*\|^2 \\ &= \|x^k - x^* - \sigma_k g^k\|^2 \\ &= \|x^k - x^*\|^2 + \sigma_k^2\|g^k\|^2 - 2\sigma_k\langle g^k, x^k - x^*\rangle. \end{aligned} \tag{24}$$

Note that $g^k \in \partial_2 f(z^k, z^k)$ we can write

$$\begin{aligned} \langle g^k, x^k - x^*\rangle &= \langle g^k, x^k - z^k + z^k - x^*\rangle \\ &= \langle g^k, x^k - z^k\rangle + \langle g^k, z^k - x^*\rangle \\ &\geq \langle g^k, x^k - z^k\rangle - f(z^k, x^*). \end{aligned} \tag{25}$$

Since $x^* \in S(C, f)$, we have $f(x^*, z^k) \geq 0$, which, by pseudomonotonicity of $f$, implies $-f(z^k, x^*) \geq 0$. Thus from (25), it follows that

$$\langle g^k, x^k - x^* \rangle \geq \langle g^k, x^k - z^k \rangle. \tag{26}$$

Remember that $x^k - z^k = \frac{\eta_k}{1-\eta_k}(z^k - y^k)$, we can write

$$\langle g^k, x^k - z^k \rangle = \frac{\eta_k}{1 - \eta_k} \langle g^k, z^k - y^k \rangle = \sigma_k \|g^k\|^2. \tag{27}$$

The last equality comes from the definition of $\sigma_k$ by (21) in the algorithm. Combining (24), (26) and (27) yields (23). □

The following theorem shows validity and convergence of the algorithm.

**Theorem 4.1.** *Suppose that the bifunction $f$ is weakly continuous, that $f(x, .)$ is convex, subdifferentiable on $C$ for any fixed $x \in C$, and that Problem (EP) admits a solution. Then both sequences $\{x^k\}$, $\{u^k\}$ converge to the unique solution of the bilevel problem (BO).*

**Proof.** As we have remarked, the linesearch used in the algorithm is well defined. To see validity of the algorithm, by the same argument as in the proof of Theorem 3.1, we can show that $S(C, f) \subseteq B_k$ for every $k$.

From definition of $D_k$, one has $x^k = P_{D_k}(x^g)$.

Note that $x^{k+1} \in D_k$, we can write

$$\|x^k - x^g\| \leq \|x^{k+1} - x^g\| \; \forall k.$$

Moreover, since $x^k = P_{D_k}(x^g)$ and $S(C, f) \subset D_k$ for every $k$, we have

$$\|x^k - x^g\| \leq \|x^* - x^g\| \; \forall x^* \in S(C, f), \; \forall k, \tag{28}$$

Hence $\{x^k\}$ is bounded. From the boundedness of $\{x^k\}$ and $\|x^k - x^g\| \leq \|x^{k+1} - x^g\|$ for every $k$, it follows that the $\lim_k \|x^k - x^g\|$ exists and is finite. Again by the same argument as in the proof of Theorem 3.1 we can see that i.e. $\|x^{k+1} - x^k\| \to 0$ as $k \to \infty$. On the other hand, since $x^{k+1} \in B_k \subseteq C_k$, by definition of $C_k$, we have

$$\|u^k - x^{k+1}\| \leq \|x^{k+1} - x^k\|.$$

Thus

$$\|u^k - x^k\| \leq \|u^k - x^{k+1}\| + \|x^{k+1} - x^k\| \leq 2\|x^{k+1} - x^k\|,$$

which together with $\|x^{k+1} - x^k\| \to 0$ implies $\|u^k - x^k\| \to 0$ as $k \to \infty$.

Next, we show that any weakly cluster point of the sequence $\{x^k\}$ is a solution to Problem (EP). Indeed, let $\overline{x}$ be any weakly cluster point of $\{x^k\}$. For simplicity of notation, without loss of generality we may assume that $x^k \rightharpoonup \overline{x}$. We consider two distinct cases:

*Case 1:* The linesearch is performed only for finitely many $k$. In this case, according to the algorithm, $u^k = x^k$ for infinitely many $k$. Thus $y^k = x^k$ is a solution to (EP) for every $k$, except a finitely many $k$. Hence, in this case the assertion is obvious.

*Case 2:* The linesearch is performed for infinitely many $k$. Then, by taking a subsequence, if necessary, we may assume that the linesearch is performed for every $k$.

We distinguish two possibilities:

(a) $\overline{\lim}_k \eta_k > 0$. From $x^k \rightharpoonup \overline{x}$ and $\|u^k - x^k\| \to 0$ follows $u^k \rightharpoonup \overline{x}$. Then applying (23) with some $x^* \in S(C, f)$ we see that $\sigma_k \; \|g^k\|^2 \to 0$. Then by definition of $\sigma_k$, we have $-\frac{\eta_k}{1-\eta_k}\langle g^k, y^k - z^k\rangle \to 0$. Since $\overline{\lim}_k \eta_k > 0$, by taking again a subsequence if necessary, we may assume that $\langle g^k, y^k - z^k\rangle \to 0$. On the other hand, using Assumption (B1) and the Armijo rule, we can write

$$0 \leq \frac{\beta}{2\rho}\|x^k - y^k\|^2 \leq \frac{1}{2\rho}L(x^k, y^k) \leq -\langle g^k, y^k - z^k\rangle \to 0.$$

Hence $\|x^k - y^k\| \to 0$, which, together with $x^k \rightharpoonup \overline{x}$, implies $y^k \rightharpoonup \overline{x}$. Note that $y^k$ being the solution of the problem

$$\min\Big\{f(x^k, y) + \frac{1}{\rho}L(x^k, y) : \; y \in C\Big\}, \qquad CP(x^k)$$

we can write

$$f(x^k, y) + \frac{1}{\rho}L(x^k, y) \geq f(x^k, y^k) + \frac{1}{\rho}L(x^k, y^k) \; \forall y \in C.$$

Letting $k$ to infinity, by the weak continuity of $f$ and $L$, we obtain

$$f(\bar{x}, y) + \frac{1}{\rho}L(\bar{x}, y) \geq f(\bar{x}, \bar{y}) + \frac{1}{\rho}L(\bar{x}, \bar{y}) \; \forall y \in C.$$

which means that $\bar{y}$ is a solution of Problem CP($\bar{x}$). Remember that $\|x^k - y^k\| \to 0$ and $x^k \rightharpoonup \bar{x}, y^k \rightharpoonup \bar{y}$, we can see that $\bar{x} = \bar{y} \in C$. Then, by Lemma 4.1, $\overline{x}$ is a solution of (EP).

(b) $\lim_k \eta_k = 0$. In this case the sequence $\{y^k\}$ also is bounded. Indeed, since $y^k$ is the solution of Problem $CP(x^k)$, whose objective function is weakly continuous, strongly convex and feasible set is constant, by Berge's maximum theorem (Proposition 23 in [2], see also [3]) the mapping $x^k \to s(x^k) := y^k$ is weakly continuous. Then from the boundedness of $\{x^k\}$, it follows that $\{y^k\}$ is bounded. Thus, we may assume, taking a subsequence if necessary, that $y^k \rightharpoonup \overline{y}$ for some $\overline{y}$. By the same argument as before we have

$$f(\overline{x}, \overline{y}) + \frac{1}{\rho}L(\overline{x}, \overline{y}) \leq f(\overline{x}, y) + \frac{1}{\rho}L(\overline{x}, y) \; \forall y \in C. \qquad (29)$$

On the other hand, as $m_k$ is the smallest natural number satisfying the Armijo linesearch rule, one has

$$f(z^{k,m_k-1}, y^k) + \frac{1}{\rho}L(x^k, y^k) > 0.$$

Note that $z^{k,m_k-1} \rightharpoonup \overline{x}$ as $k \to \infty$, from the last inequality, by the weak continuity of $f$ and $L$, we obtain in the limit that

$$f(\overline{x}, \overline{y}) + \frac{1}{\rho} L(\overline{x}, \overline{y}) \geq 0. \tag{30}$$

Substituting $y = \overline{x}$ into (29) we get

$$f(\overline{x}, \overline{y}) + \frac{1}{\rho} L(\overline{x}, \overline{y}) \leq 0,$$

which, together with (30), yields

$$f(\overline{x}, \overline{y}) + \frac{1}{\rho} L(\overline{x}, \overline{y}) = 0. \tag{31}$$

From (31) and

$$f(\overline{x}, \overline{x}) + \frac{1}{\rho} L(\overline{x}, \overline{x}) = 0,$$

it follows that both $\overline{x}$ and $\overline{y}$ are solutions of the strongly convex program

$$\min\{f(\overline{x}, y) + \frac{1}{\rho} L(\overline{x}, y) : y \in C\}.$$

Hence $\overline{x} = \overline{y}$ and therefore, by Lemma 4.1, $\overline{x}$ solves (EP). Moreover, from $\|u^k - x^k\| \to 0$, we can also conclude that every weakly cluster point of $\{u^k\}$ is a solution to (EP).

Finally, we show that $\{x^k\}$ converges to the unique solution of the bilevel problem (BO). To this end, we note that, by (28) and the fact that any weakly cluster point of $\{x^k\}$ belongs to the solution set $S(C, f)$, all conditions in Lemma 3.1 are satisfied for $S(C, f)$, $x^g$ and $\{x^k\}$. Thus $x^k \to P_S(x^g)$, which is the solution of (BO). Then, since $\|x^k - u^k\| \to 0$, we obtain $u^k \to P_S(x^g)$. □

## 5 STABILITY ISSUES

As we have mentioned, when $f$ is monotone, the bifunction $f_n$ defined in (12) is strongly monotone, and therefore the iterate $u^n$ can be computed by some existing methods (see e.g., [11, 13, 16, 17, 22, 23, 24, 26, 31] and the references therein). However, when $f$ is pseudomonotone, $f_n$ does not inherit any monotonicity property from $f$, and therefore computing a $\delta_n$-solution by (12) is a difficult task.

In order to overcome this difficulty we can use the algorithm described in the previous section to compute the limit point of the sequences $\{x^n\}$ and $\{u^n\}$ by solving the bilevel optimization problem (BO). Since the objective function $\|x - x^g\|^2$ is strongly convex and the constrained set $S(C, f)$ is convex, Problem (BO) is uniquely solvable. This facts suggests that PPM can be used for handling ill-posed pseudomonotone EPs. For this purpose, in the sequel, we suppose that $T$ is a Banach space and $C : T \to 2^{\mathcal{H}}$ is a operator form $T$ to

$\mathcal{H}$ such that $C(t)$ is nonempty, closed, convex for every $t \in T$, and we consider the parametric EPs of the form

$$\text{Find } x(t) \in C(t) : f(x(t), y) \geq 0 \ \forall y \in C(t). \tag{$EP_t$}$$

Let us denote the solution set of this problem by $S(t)$. Then the corresponding bilevel optimization problem takes the from

$$\min\{\|x - x^g\|^2 : x \in S(t)\}. \tag{$BO(t)$}$$

In addition to Assumptions $(A_1)$, $(A_2)$, suppose that $S(0) \neq \emptyset$ and that the mapping $C(.)$ is upper semicontinuous in a neighbourhood of 0. Then, by the well-known Berge maximum theorem [3], the unique solution $x(t)$ of Problem $(BO(t))$ is continuous at 0.

Below we give a particular case, where the solution set mapping $S(.)$ is upper semicontinuous. First we recall [3] that, a multivalued mapping $F$ from a Banach space $X$ to Banach space $Y$ is closed (resp. convex) if its graph is closed (resp. convex) in $X \times Y$.

Now we suppose that the feasible domain $C := F^{-1}(0)$ and we consider the parametric problem

$$\begin{cases} \text{Find } \ x(t) \in F^{-1}(t) \ \text{ such that} \\ f(x(t), x) \geq 0, \ \forall x \in F^{-1}(t). \end{cases} \tag{32}$$

Let $S(t)$ denote the solution set of this problem. We need the following lemma:

**Lemma 5.1.** ([20, Lemma 1]) *Suppose that $F$ is a multivalued mapping from $X$ into $Y$ satisfying*

a) *$F$ is convex and closed;*

b) *$F(X) = Y$;*

c) *$F^{-1}(0)$ is bounded.*

*Then for each bounded neighborhood $V_0$ of $0 \in Y$ there is a bounded closed set $B \subset X$ such that $F^{-1}(t) \subset B$ for all $t \in V_0$ and $F^{-1}$ is upper semicontinuous in $V_0$.*

The following proposition on the upper semicontinuity of the solution set mapping has been proved in [20] for monotone bifunction. Now we extend it to the problem (32) with $f$ being an equilibrium pseudomonotone bifunction on $X = \mathcal{H}$.

**Proposition 5.1.** *Suppose that $f$ is pseudomonotone on $X$. Then under Assumptions $(A_1), (A_2)$ and the conditions specified in Lemma 5.1, there exists a neighborhood $V_0$ of $0 \in Y$ such that problem* (32) *has a solution for every $t \in V_0$ and the mapping $S(.)$ is upper semicontinuous at* 0.

*Proof.* We outline the proof because it can be done a similar way as in the proof of Theorem 1 in [20] for monotone case. Since $F^{-1}$ is convex and closed, $F^{-1}(t)$ is convex and closed for every $t \in V_0$. Moreover, by Lemma 5.1, $F^{-1}(V_0)$ is contained in a bounded closed set. Then from Assumptions $(A_1)$ and $(A_2)$ it follows that the problem (32) has a solution for every $t \in V_0$ [?]. In addition, the pseudomonotonicity of $f$ implies that, $x(t)$ is a solution of (32) if and only if it is a solution of the dual problem (see e.g. [12, Proposition 2.1.15]), that is

$$x(t) \in F^{-1}(t): \ f(x(t), x) \geq 0, \ \forall t \in F^{-1}(t)$$

if and only if

$$x(t) \in F^{-1}(t): \ f(x, x(t)) \leq 0, \ \forall t \in F^{-1}(t).$$

Now take $h(t, x') := \max\{f(x, x') : x \in F^{-1}(t)\}$. Then by Lemma 5.1 and the well-known Berge maximum theorem, $h$ is lower semicontinuous in $X \times V_0$. As $S(t)$ is contained in a bounded set, to see the upper semicontinuity of the solution mapping $S(.)$, we need to show the closedness of its graph. Indeed, let $(t^0, x^0) \notin graphS$. Then

$$x^0 \notin F^{-1}(t^0) \text{ or } h(t^0, x^0) > 0 \text{ or both.}$$

Then, by the closedness of $F^{-1}(t^0)$ and lower semicontinuity of $h$, there exists a neighborhood $V \times U$ of $(t^0, x^0)$ such that

$$x \notin F^{-1}(t) \text{ or } h(t, x) > 0 \text{ or both,}$$

which implies that $(U \times V) \cap graphS = \emptyset$. □

To illustrate the result, let us consider an example, where $F(x) := M - G(x)$ with $G$ being a mapping from $X$ into $Y$ and $M$ a closed convex cone in $Y$. We suppose that

(i) $G$ is continuous and $-G(X) + M = Y$;

(ii) $G$ is $M-$ convex on $X$, i.e.

$$G(tx + (1-t)y) \in tG(x) + (1-t)G(y) + M, \ \forall x, y, \ \forall t \in [0, 1].$$

Then it is not hard to verify that $F^{-1}(t) = \{x \in X : G(x) + t \in M\}$ and that all assumptions imposed on $F$ are satisfied.

General conditions ensuring the upper semicontinuity of the solution mappings of parametric equilibrium problems can be found, for example, in [?] and the references therein.

**Conclusion.** We have considered an inexact proximal point algorithm for pseudomonotone EPs in real Hilbert spaces and have shown that, although the regularized subproblems are not uniquely solvable, any proximal trajectory weakly converges to the same limit. In order to make the convergence strong and to locate the limit point, we have extended a viscosity-proximal algorithm in [29] to pseudomonotne EPs. Then we have proposed a hybrid extragradint-cutting plane algorithm for solving the resulting strongly convex bilevel optimization problem, thereby to obtain the limit point. The obtained results allow possibilities to use bilevel convex optimization as a regularization tool for handling ill-possed pseudomonotone EPs.

## 6 ACKNOWLEDGMENTS

This research is funded by Vietnam National Foundation for Science and Technology Development (NAFOSTED) under grant number 101.02-2011.19.

# References

[1] L.Q. Anh, P.Q. Khanh (2008). Semicontinuity of the approximate solution sets of multivalued quasiequilibrium problems. *Numer. Funct. Anal. Optim.* 29:24-42.

[2] J.B. Aubin, I. Ekeland (1984). *Applied nonlinear analysis*. John Wiley and Sons

[3] C. Berge (1968). *Topological spaces*. MacMillan, New York.

[4] M. Bianchi, S. Schaible (1996). Generalized monotone bifunctions and equilibrium problems. *J. Optim. Theory Appl.* 90:31-43.

[5] E. Blum, W. Oettli (1994). From optimization and variational inequalities to equilibrium problems. *Math. Student*. 62:127-169.

[6] S. Dempe 2002. *Foundations of Bilevel Programming*. Kluwer Academic Press, Dordrecht.

[7] B.V. Dinh, L.D. Muu (2011). On penalty and gap function methods for bilevel equilibrium problems. *J. Appl. Math.* doi:10.1155/2011/646452.

[8] Y. Censor, A. Lent (1981). An iterative row-action method for interval convex programming. *J. Optim. Theory Appl.* 34:321-353.

[9] F. Facchinei, J.S. Pang (2003). *Finite-Dimensional Variational Inequalities and Complementarity Problems*. Springer, New York.

[10] P.G. Hung, L.D. Muu (2011). The Tikhonov regularization extended to equilibrium problems involving pseudomonotone bifunctions. *Nonlinear Anal. TMA*. 74:6121-6129.

[11] A. Iusem, W. Sosa (2010). On the proximal point method for equilibrium problems in Hilbert spaces. *Optimization*. 59:1259-1274.

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

[13] I.V. Konnov (2003). Application of the proximal point method nonmonotone equilibrium problems. *J. Optim. Theory Appl.* 119:317-333.

[14] I.V. Konnov (2009). Regularization methods for nonmonone equilibrium problems. *J. Nonlinear and Convex Anal.* 10:93-101.

[15] G.M. Korpelevich (1976). The extragradient method for finding saddle points and other problems. *Ekon. Math. Methody.* 12:747-756.

[16] G. Mastroeni (2003). On Auxiliary priciple for equilibrium problems. *J. Glob. Optim.* 27:411-426.

[17] G. Mastroeni (2003). Gap functions for equilibrium problems. *J. Glob. Optim.* 27:411-426.

[18] A. Moudafi (1999). Proximal point algorithm extended to equilibrium problems. *J. of Natural Geometry.* 15:91-100.

[19] A. Moudafi (2010). Proximal methods for a class of bilevel monotone equilibrium problems. *J. Glob. Optim.* 47:287-292.

[20] L.D. Muu (1984). Stability property of a class of variational inequality. *Optimization.* 15:347-351.

[21] L.D. Muu, W. Oettli (1992). Convergence of an adaptive penalty scheme for finding constrained equilibria. *Nonlinear Anal.* 18:1159-1166.

[22] L.D. Muu, T.D. Quoc (2009). Regularization algorithms for solving monotone Ky Fan inequalities with application to a Nash-Cournot equilibrium model. *J. Optim. Theory Appl.* 142:185-204.

[23] L.D. Muu, V.H. Nguyen and T.D. Quoc (2008). Extragradient algorithms extended to equilibrium problems. *Optimization.* 57:749-776.

[24] M.A. Noor (2004). Auxiliary principle technique for equilibrium problems. *J. of Optim. Theory and Appl.* 122: 371-386.

[25] M.A. Noor, K.I. Noor (2004). On equilibrium problems. *Appl. Math. E-Notes.* 4:125-132.

[26] T.D. Quoc and L. D. Muu . Iterative methods for solving monotone equilibrium problems. *Comput. Optim. Appl.* DOI 10.1007/s10589-010-5360-4.

[27] R.T. Rockafellar (1976). Monotone operators and the proximal point algorithm. *SIAM J. Control and Optim.* 5:877-898.

[28] M.V. Solodov and B.F. Svaiter (1999). A hybrid projection-proximal point algorithm. *J. of Convex Anal.* 6:59-70.

[29] A. Tada and W. Takahashi (2007). Weak and strong convergence theorem for nonexpansive mapping and equilibrium problem. *J. Optim. Theory and Appl.* 133:359-370.

[30] A.N. Tikhonov, V.Y. Arsenin (1977). *Solutions of Ill-Posed Problems.* John Wiley and Sons, New York.

[31] N.T.T. Van, V.H. Nguyen and J.J. Strodiot (2009). A bundle method for solving equilibrium problems. *Math. Prog.* 116:529-552.