A BUNDLE METHOD TO SOLVE MULTIVALUED VARIATIONAL INEQUALITIES *

GENEVIÈVE SALMON[†], JEAN-JACQUES STRODIOT[‡], AND VAN HIEN NGUYEN[§]

Abstract. In this paper we present a bundle method for solving a generalized variational inequality problem. This problem consists in finding a zero of the sum of two multivalued operators defined on a real Hilbert space. The first one is monotone and the second one is the subdifferential of a lower semicontinuous proper convex function. The method is based on the auxiliary problem principle due to Cohen and the strategy is to approximate, in the subproblems, the nonsmooth convex function by a sequence of convex piecewise linear functions as in the bundle method in nonsmooth optimization. This makes these subproblems more tractable. Moreover to ensure the existence of subgradients at each iteration, we also introduce a barrier function in the subproblems. This function prevents the iterates to go outside the interior of the feasible domain. First we explain how to build, step by step, a suitable piecewise linear approximation and we give conditions to ensure the boundedness of the sequence generated by the algorithm. Then we study the properties that a gap function must satisfy to obtain that each weak limit point of this sequence is a solution of the problem. In particular, we give existence theorems of such a gap function when the first multivalued operator is paramonotone, weakly closed and Lipschitz continuous on bounded subsets of its domain and when it is the subdifferential of a convex function. When it is strongly monotone, we obtain that the sequence generated by the algorithm strongly converges to the unique solution of the problem.

Key words. generalized variational inequality, multivalued mapping, auxiliary problem principle, bundle method, gap functions, paramonotone operator

AMS subject classifications. 65K05, 90C25

1. Introduction. Let F be a monotone multivalued operator defined on a real Hilbert space H with inner product $\langle\cdot,\cdot\rangle$, let C be a nonempty closed convex subset of H and let $\varphi : H \to I\!R \cup \{+\infty\}$ be a lower semicontinuous proper convex function. We consider the following general variational inequality problem:

$$(P)\begin{cases} \text{find } x^* \in C \text{ and } r(x^*) \in F(x^*) \text{ such that, for all } x \in C, \\ \langle r(x^*), x - x^*\rangle + \varphi(x) - \varphi(x^*) \geq 0. \end{cases}$$

This problem can also be expressed under the form: Find x^* such that $0 \in F(x^*) + \partial(\varphi + \psi_C)(x^*)$ where ψ_C denotes the indicator function associated with C (i.e., $\psi_C(x) = 0$ if $x \in C$ and $+\infty$ otherwise) and $\partial(\varphi + \psi_C)(x^*)$, the subdifferential of the convex function $\varphi + \psi_C$ at x^*. Here we assume that $C \subseteq \text{dom}\varphi \subseteq \text{dom}F$ and that C is equal to the closure of its interior. Moreover, we suppose that there exists at least one solution to this problem. Existence results for problem (P) can be found, for example, in [3, 11] and [13]. When F is the subdifferential of a finite-valued convex continuous function f defined on H, problem (P) reduces to the nondifferentiable convex optimization problem:

$$(OP) \quad \min_{x \in C} \{f(x) + \varphi(x)\}.$$

* This work was supported in part by a grant from the Belgian Fonds National de la Recherche Scientifique (FNRS: B8/5-CB/MF-4.515 and B8/5-CB/SP-9.579).

[†]Assistant, Unité d'Optimisation, Département de Mathématique, Facultés Universitaires Notre Dame de la Paix, Namur, Belgium(genevieve.salmon@fundp.ac.be).

[‡]Professor, Unité d'Optimisation, Département de Mathématique, Facultés Universitaires Notre Dame de la Paix, Namur, Belgium(jean-jacques.strodiot@fundp.ac.be).

[§]Professor, Unité d'Optimisation, Département de Mathématique, Facultés Universitaires Notre Dame de la Paix, Namur, Belgium(vhnguyen@fundp.ac.be).

An important research work is devoted to the solution of problem (P) specially when $\varphi = 0$ (see, for example, [10, 13, 14, 15, 17, 18, 19, 25]). However, variational inequalities with a multivalued mapping F and a function $\varphi \neq 0$ are encountered in many applications. In particular, it is the case in mechanical problems (see e.g., [24]) and in equilibrium problems (see e.g., [7, 23]). So it is worth studying implementable methods for solving such problems. It is the purpose of this paper.

When $\varphi = 0$, problem (P) can be expressed under the form: Find x^* such that $0 \in F(x^*) + \partial\psi_C(x^*)$. Many methods have been proposed for solving this particular problem. When F is maximal monotone, the most famous method is the proximal method (see e.g., [20, 28]) which consists in finding a zero of the operator $F + \partial\psi_C$ by using the scheme:

$$x^{k+1} = [I + \mu_k(F + \partial\psi_C)]^{-1}(x^k), \tag{1.1}$$

where $\{\mu_k\}_{k\in\mathbb{N}}$ is a sequence of positive real numbers. More recently, this method has been generalized to avoid solving constrained subproblems at each iteration. The linear term $x^{k+1} - x^k$ in the equality (1.1) written under the form $x^k - x^{k+1} \in \mu_k(F + \partial\psi_C)(x^{k+1})$ has been replaced by some nonlinear functional $r(x^{k+1}, x^k)$ based on entropic proximal terms arising from appropriately formulated Bregman functions [5, 12] or entropic φ–divergence [32].

Splitting methods have also been studied for solving problem (P) when $\varphi = 0$. Here the multivalued operators F and $\partial\psi_C$ play separate roles. The simplest splitting method is the forward–backward scheme, see e.g., [33], whose iteration is given by

$$x^{k+1} \in [I + \mu_k\partial\psi_C]^{-1}[I - \mu_k F](x^k), \tag{1.2}$$

where $\{\mu_k\}_{k\in\mathbb{N}}$ is a sequence of positive real numbers. First, one element $r(x^k)$ is computed in $F(x^k)$ and then the vector $x^k - \mu_k r(x^k)$ is projected onto the closed convex set C.

When $\varphi \neq 0$, Cohen developed in [8] a general algorithmic framework for solving problem (P), based on the so–called auxiliary problem principle. The corresponding method is a generalization of the forward–backward method. More precisely, let $h : H \to \mathbb{R}$ be a continuously differentiable strongly convex function of modulus $\beta > 0$ on an open set containing C, and let $\{\mu_k\}_{k\in\mathbb{N}}$ be a sequence of positive real numbers. The problem considered at iteration k is the following:

$$x^{k+1} \in [\nabla h + \mu_k\partial(\varphi + \psi_C)]^{-1}[\nabla h - \mu_k F](x^k) \tag{1.3}$$

i.e.,

$$\begin{cases} \text{choose } r(x^k) \in F(x^k) \text{ and find } x^{k+1} \in C \text{ such that, for all } x \in C, \\ \langle r(x^k) + \mu_k^{-1}(\nabla h(x^{k+1}) - \nabla h(x^k)), x - x^{k+1}\rangle + \varphi(x) - \varphi(x^{k+1}) \geq 0. \end{cases}$$

This problem can also be equivalently written under the following minimization form:

$$(AP^k)\begin{cases} x^{k+1} \in \operatorname{argmin}_{x\in H}\{\varphi(x) + \psi_C(x) + \mu_k^{-1}[\ h(x) - h(x^k) - \langle z^k, x - x^k\rangle\ ]\ \}, \\ \text{with } z^k = \nabla h(x^k) - \mu_k r(x^k) \text{ and } r(x^k) \in F(x^k). \end{cases}$$

Since h is strongly convex on C, this problem has one and only one solution. The strong convergence of the sequence $\{x^k\}_{k\in\mathbb{N}}$ generated by this algorithm, has been proven by Cohen under the assumption that F is strongly monotone. More recently, Zhu ([34]) proved a weak convergence result when the operator F is paramonotone on C i.e., F is monotone on C and for all $x, y \in C$, and $r(x) \in F(x), r(y) \in F(y)$,

$$\langle r(x) - r(y), x - y \rangle = 0 \implies r(y) \in F(x) \text{ and } r(x) \in F(y). \tag{1.4}$$

This notion has been introduced by Bruck ([4]) and further studied in [16].

When φ is a nonsmooth convex function, subproblems (AP^k) may be very hard to solve and several authors proposed to approximate the function φ by a sequence $\{\psi^k\}_{k\in\mathbb{N}}$ composed of more tractable convex functions. When F is a single–valued operator, it is proven in [21, 29] that the resulting algorithm is convergent provided that the sequence $\{\psi^k\}_{k\in\mathbb{N}}$ converges to φ in the sense of Mosco [22]. Recently in the multivalued case, Salmon et al. ([30, 31]) suggested, when $C = H$, to approximate φ by an interior approximation $\psi^k \geq \varphi$ in such a way that the sequence $\{\psi^k\}_{k\in\mathbb{N}}$ converges sufficiently fast to φ in the sense that, for each solution x^* of problem (P), there exists a sequence $\{w^k\}_{k\in\mathbb{N}}$ in H such that

$$\sum_{k=1}^{+\infty} \|w^k - x^*\| < +\infty \quad \text{and} \quad \sum_{k=1}^{+\infty} \mid \psi^k(w^k) - \varphi(x^*) \mid < +\infty. \tag{1.5}$$

In [30], they proved the strong convergence when F is strongly monotone, and in [31] a weak convergence result when F is paramonotone.
An example of such an approximating sequence $\{\psi^k\}_{k\in\mathbb{N}}$ is the sequence of logarithmic barrier functions associated with a closed convex subset C described by finitely many convex inequalities: $g_i(x) \leq 0, i = 1, \ldots, m$. In this case, $\varphi = \psi_C$ and ψ^k is defined by

$$\psi^k(x) = -\nu_k^{-1} \sum_{i=1}^{m} \log\left(\min\left(\frac{1}{2}, -g_i(x)\right)\right), \quad x \in \operatorname{int} C. \tag{1.6}$$

Another example is the inverse barrier function defined by

$$\psi^k(x) = -\nu_k^{-1} \sum_{i=1}^{m} \frac{1}{g_i(x)}, \quad x \in \operatorname{int} C. \tag{1.7}$$

In these two examples, $\nu_k > 0$ is a barrier parameter and $\nu_k \to +\infty$. For the sequence $\{\psi^k\}_{k\in\mathbb{N}}$ of logarithmic barrier functions defined by (1.6), it is easy to see that there exists a sequence $\{w^k\}_{k\in\mathbb{N}}$ in int C such that (1.5) holds provided that, for all k, the barrier parameter ν_k be greater than k^γ with $\gamma > 1$. The same conclusion holds for the inverse barrier function (1.7).

When $F = 0$ and $C = H$, problem (P) reduces to the problem of minimizing the nondifferentiable convex function φ on H. This problem can be solved by the so–called bundle method, introduced in the eighties by Lemaréchal, see e.g., [9]. In this method, the effective domain of φ is supposed to be the whole space H and the strategy is to approximate, at iteration k, the function φ, step by step, by a piecewise linear convex function φ^k, and to move to the next iterate only when the approximation is suitable. The resulting step is called a *serious* step. As proven in [9], this method can be seen as a practical implementation of the classical proximal method in convex optimization. Our purpose in this paper is to use the bundle strategy for solving problem (P). However, there immediately appears a difficulty when we follow this strategy for solving approximately problem (AP^k). Indeed, a way to build a piecewise linear convex approximation $\varphi^k \leq \varphi$, is to generate points $y^1, \ldots, y^p$ in C, and to consider the function

$$\varphi^k(y) = \max_{1\leq i\leq p}\{\varphi(y^i) + \langle s(y^i), y - y^i\rangle\}, \quad y \in C, \tag{1.8}$$

where $s(y^i)$ denotes one subgradient of φ at y^i for $i = 1, \ldots, p$. Usually the points $y^1, \ldots, y^p$ are the trial points built from x^k. Doing that, we suppose that $s(y^i)$ exists for $i = 1, \ldots, p$. But we know that the subdifferential of a convex function is nonempty in the interior of its domain and may be empty on the boundary of its domain ([26]). Here, the latter case may occur because $C \subseteq \mathrm{dom}\varphi$ and the trial points are in C. So, in our method, to prevent the iterates to go to the boundary of C, we introduce a barrier function ψ^k in the objective function of problem (AP^k). Then this problem becomes an unconstrained problem.

The convergence of the resulting algorithm is presented in three steps. First, we prove that the sequence $\{x^k\}_{k\in\mathbb{N}}$ generated by the algorithm is bounded, then, that each weak limit point of this sequence is a solution of problem (P), and finally, that the sequence weakly and strongly converges to such a solution. To prove the boundedness of the sequence $\{x^k\}_{k\in\mathbb{N}}$, we have to impose, in particular, that the sequence $\{\psi^k\}_{k\in\mathbb{N}}$ satisfies property (1.5) with φ replaced by ψ_C. To obtain the property that each weak limit point of $\{x^k\}_{k\in\mathbb{N}}$ is a solution of (P), we use the concept of gap function associated with problem (P) and we observe that this property holds when there exists a gap function l that is weakly continuous on int C, and is such that $l(x^k) \to 0$. When l is a general Lipschitz continuous function (non necessarily a gap function), Cohen and Zhu [6] gave conditions on a sequence $\{x^k\}_{k\in\mathbb{N}}$ to obtain that $l(x^k) \to 0$. So, in a first part, we present three existence theorems of a gap function associated to (P) that is weakly continuous and Lipschitz continuous on the bounded sets of int C. The first one when F is paramonotone, weakly closed on C and Lipschitz continuous on bounded subsets of int C, the second one when F is the subdifferential of a convex continuous function and is bounded on bounded subsets of int C and finally the third one when F is strongly monotone on C and bounded on bounded subsets of int C. Then, in a second part, we prove that the conditions given by Cohen and Zhu on the sequence $\{x^k\}_{k\in\mathbb{N}}$ are satisfied by the sequence generated by our algorithm, provided that the sequence $\{\psi^k\}_{k\in\mathbb{N}}$ of barrier functions is uniformly Lipschitz continuous on the set $\{x^k \mid k \in \mathbb{N}\}$. This condition is satisfied by the logarithmic and inverse barrier functions. These results allow us to give a very general convergence theorem, not only for the weak limit points of $\{x^k\}_{k\in\mathbb{N}}$ but also for the weak convergence (when, in addition, ∇h is weakly continuous) and the strong convergence (when F is strongly monotone) of the sequence $\{x^k\}_{k\in\mathbb{N}}$. When $\varphi = 0$, we find again the results of [31] and when, in addition, ψ_C is not approximated by a barrier function, the results due to Zhu in [34].

The paper is organized as follows. In §2, first we define an approximation criterion for φ and we show how to use the bundle strategy to get a piecewise linear convex function φ^k that satisfies this criterion. Then we explain why the resulting algorithm is well defined. In §3, we present our convergence theorems concerning the boundedness of the sequence $\{x^k\}_{k\in\mathbb{N}}$ generated by the algorithm and the weak and strong convergence of this sequence to a solution of problem (P). Throughout this paper, we denote by $\Gamma_0(H)$ the set of lower semicontinuous proper convex functions from H into $\mathbb{R} \cup \{+\infty\}$. Any other undefined term or usage should be taken as in the books [3] and [26].

2. Approximate auxiliary problem. At iteration k, functions φ and ψ_C are approximated in the auxiliary subproblem (AP^k) by functions φ^k and ψ^k respectively. Then the approximate auxiliary problem can be expressed as: find x^{k+1} the solution of the unconstrained problem

$$(P^k) \quad \min\{\varphi^k(x) + \psi^k(x) + \mu_k^{-1}[\ h(x) - h(x^k) - \langle z^k, x - x^k\rangle\ ]\ \},$$

where $z^k = \nabla h(x^k) - \mu_k r(x^k)$ with $r(x^k) \in F(x^k)$. The functions φ^k are piecewise linear convex functions such that $\varphi^k \leq \varphi$ while the functions ψ^k are barrier functions associated with C, i.e., $\psi^k \in \Gamma_0(H), \psi_C \leq \psi^k, \mathrm{dom}\psi^k = \mathrm{int}\ C$ and $y^i \to y \in \mathrm{bdry}C$ implies that $\psi^k(y^i) \to +\infty$. Notice that int C is nonempty because C is the closure of its interior. Since h is strongly convex on C and ψ^k is a barrier function, it is easy to prove that there exists one and only one solution for problem (P^k) and that this solution belongs to int C. So x^{k+1} is well defined. On the sequence $\{\psi^k\}_{k\in\mathbb{N}}$, we impose the following condition: for each solution x^* of problem (P), there exists a sequence $\{w^k\}_{k\in\mathbb{N}}$ in int C such that (1.5) holds with $\varphi = \psi_C$, i.e.,

$$\sum_{k=1}^{+\infty} \|w^k - x^*\| < +\infty \quad \text{and} \quad \sum_{k=1}^{+\infty} \psi^k(w^k) < +\infty. \tag{2.1}$$

In order to obtain the approximate function φ^k, we use a bundle strategy, i.e., we build, successively, piecewise linear convex functions $\theta^1, \ldots, \theta^i, \ldots$ and we set $\varphi^k = \theta^i$ when the solution y^i of the unconstrained problem

$$(P_i^k) \quad \min\{\theta^i(x) + \psi^k(x) + \mu_k^{-1}[\ h(x) - h(x^k) - \langle z^k, x - x^k\rangle\ ]\ \},$$

is such that $\varphi(y^i) - \theta^i(y^i) \leq \Delta_k$ where $\Delta_k > 0$ is some tolerance. More precisely, one iteration of our algorithm is given by the following process:

Bundle algorithm

Let $x^k \in$ int C and $\mu_k, \Delta_k > 0$ be given. Compute $r(x^k) \in F(x^k)$. Set $z^k = \nabla h(x^k) - \mu_k r(x^k)$, $y^0 = x^k$ and $i = 1$.

Step 1. Choose a piecewise linear function $\theta^i \leq \varphi$ and solve problem (P_i^k) to obtain $y^i \in$ int C.

Step 2. If $\varphi(y^i) - \theta^i(y^i) \leq \Delta_k$, STOP and set $\varphi^k = \theta^i$ and $x^{k+1} = y^i$. Otherwise increase i by 1 and go to Step 1.

In order to prove that the STOP occurs after finitely many iterations, we have to impose conditions on the functions $\theta^i, i = 1, 2, \ldots$ Before presenting these conditions, first we observe that, by optimality of $y^i \in$ int C, we have

$$\gamma^i \equiv \mu_k^{-1}[z^k - \nabla h(y^i)] \in \partial[\theta^i + \psi^k](y^i). \tag{2.2}$$

Then we define the aggregate affine function l^i by

$$l^i(y) = \theta^i(y^i) + \psi^k(y^i) + \langle \gamma^i, y - y^i\rangle, \quad y \in \mathrm{int}\ C. \tag{2.3}$$

We have $l^i(y^i) = \theta^i(y^i) + \psi^k(y^i)$ and, using (2.2) and (2.3),

$$l^i(y) \leq \theta^i(y) + \psi^k(y), \quad \text{for all } y \in \mathrm{int}\ C. \tag{2.4}$$

Now we require the following conditions on the functions θ^i for $i = 1, 2, \ldots$:

(C1) $\theta^i \leq \varphi$,
(C2) $l^i \leq \theta^{i+1} + \psi^k$,
(C3) $\varphi(y^i) + \langle s(y^i), \cdot - y^i \rangle \leq \theta^{i+1}$,
(C4) $\varphi(y^0) + \langle s(y^0), \cdot - y^0 \rangle \leq \theta^i$,

where $s(y^i)$ denotes a subgradient of φ at y^i. Here we suppose that, at each point of int C, one subgradient of φ is available.

The first three conditions are similar to the ones introduced in [9] in the framework of nonsmooth convex optimization. As in [9], they allow to prove that the STOP occurs in the bundle algorithm after finitely many iterations. Condition (C4) will be used in the next section to show the weak convergence of the sequence $\{x^k\}_{k\in\mathbb{N}}$ generated, step by step, by the bundle algorithm.

Let us now mention a few examples of functions θ^i satisfying Conditions (C1) to (C4). For the first function, we can take $\theta^1 = \varphi(y^0)+\langle s(y^0), \cdot - y^0\rangle$ and for $i = 1, 2, \ldots$, we can choose

$$\theta^{i+1} = \max_{0\leq j\leq i}\{\varphi(y^j) + \langle s(y^j), \cdot - y^j\rangle\}. \tag{2.5}$$

It is easy to see that (C1), (C3) and (C4) are satisfied. Since $\theta^i \leq \theta^{i+1}$, (C2) follows from (2.4). When ψ^k is differentiable on int C, other choices are possible, e.g.,

$$\theta^{i+1} = \max_{j\in\{0,i\}}\{\theta^i(y^i) + \langle \gamma^i - \nabla\psi^k(y^i), \cdot - y^i\rangle, \varphi(y^j) + \langle s(y^j), \cdot - y^j\rangle\}. \tag{2.6}$$

Indeed, (C3) and (C4) are obvious and (C1) is satisfied because $\gamma^i - \nabla\psi^k(y^i) \in \partial\theta^i(y^i)$ and $s(y^i) \in \partial\varphi(y^i)$. Finally, since

$$\theta^{i+1} \geq \theta^i(y^i) + \langle \gamma^i - \nabla\psi^k(y^i), \cdot - y^i\rangle = l^i - \psi^k(y^i) - \langle \nabla\psi^k(y^i), \cdot - y^i\rangle,$$

we have, using the subdifferential inequality, that

$$l^i \leq \theta^{i+1} + \psi^k(y^i) + \langle \nabla\psi^k(y^i), \cdot - y^i\rangle \leq \theta^{i+1} + \psi^k,$$

i.e., condition (C2).
In the sequel we will also need to consider the following functions:

$$\tilde{l}^i(y) = l^i(y) + \mu_k^{-1}\ [\ h(y) - h(x^k) - \langle z^k, y - x^k\rangle\ ],$$
$$\tilde{\theta}^i(y) = \theta^i(y) + \mu_k^{-1}\ [\ h(y) - h(x^k) - \langle z^k, y - x^k\rangle\ ].$$

Using (2.2) and (2.3), it is easy to see that, for all $y \in$ int C,

$$\tilde{l}^i(y) = \tilde{l}^i(y^i) + \mu_k^{-1}\ [\ h(y) - h(y^i) - \langle \nabla h(y^i), y - y^i\rangle\ ]. \tag{2.7}$$

Moreover, we have

$$\tilde{\theta}^i(x^k) = \theta^i(x^k) \quad \text{and} \quad \tilde{l}^i(y^i) = \tilde{\theta}^i(y^i) + \psi^k(y^i), \tag{2.8}$$

and, by condition (C2),

$$\tilde{l}^i \leq \tilde{\theta}^{i+1} + \psi^k. \tag{2.9}$$

PROPOSITION 2.1. *Suppose that $\partial\varphi$ is bounded on bounded subsets of int C. If the stopping test is suppressed in the bundle algorithm and if the sequence $\{\theta^i\}_{i\in\mathbb{N}_0}$ satisfies conditions (C1) to (C3), then $\varphi(y^i) - \theta^i(y^i) \to 0$.*

Proof. We proceed in three steps.

1. The sequence $\{\tilde{l}^i(y^i)\}_{i\in\mathbb{N}_0}$ is convergent and $y^{i+1} - y^i \to 0$.

   For all $i = 1, \dots$ we have

$$
\begin{aligned}
\varphi(x^k) + \psi^k(x^k) &\geq \theta^{i+1}(x^k) + \psi^k(x^k) && \text{(by } (C1)) \\
&= \tilde{\theta}^{i+1}(x^k) + \psi^k(x^k) && \text{(by (2.8))} \\
&\geq \tilde{\theta}^{i+1}(y^{i+1}) + \psi^k(y^{i+1}) && \text{(definition of } y^{i+1}) \\
&= \tilde{l}^{i+1}(y^{i+1}) && \text{(by (2.8))} \\
&\geq \tilde{l}^i(y^{i+1}) && \text{(by (2.9))} \\
&= \tilde{l}^i(y^i) + \mu_k^{-1}\ D_h(y^{i+1}, y^i) && \text{(by (2.7))},
\end{aligned}
$$

   where $D_h(y,z) = h(y) - h(z) - \langle \nabla h(z), y - z\rangle$.

   ¿From these relations, we deduce that the sequence $\{\tilde{l}^i(y^i)\}_{i\in\mathbb{N}_0}$ is nondecreasing and bounded above by $\varphi(x^k) + \psi^k(x^k)$. So it is convergent. Moreover, since h is strongly convex of modulus $\beta > 0$, we also obtain that

$$\tilde{l}^{i+1}(y^{i+1}) - \tilde{l}^i(y^i) \geq \mu_k^{-1}\ D_h(y^{i+1}, y^i) \geq (2\mu_k)^{-1}\beta\|y^{i+1} - y^i\|^2 \geq 0.$$

   But then $y^{i+1} - y^i \to 0$ (strongly) because the left hand side tends to zero.

2. The sequence $\{y^i\}_{i\in\mathbb{N}}$ is bounded.

   Let $y \in \text{int } C$ be fixed. Using successively (C1) and the definition of $\tilde{\theta}^{i+1}$, (2.9), (2.7) and the strong convexity of h, we have

$$
\begin{aligned}
\varphi(y) + \psi^k(y) + \mu_k^{-1}\ [\ h(y) - h(x^k) - \langle z^k, y - x^k\rangle\ ] &\geq \tilde{\theta}^{i+1}(y) + \psi^k(y) \\
\geq \tilde{l}^i(y^i) + \mu_k^{-1} D_h(y, y^i) &\geq \tilde{l}^i(y^i) + (2\mu_k)^{-1}\beta\|y - y^i\|^2.
\end{aligned}
$$

   Since the sequence $\{\tilde{l}^i(y^i)\}_{i\in\mathbb{N}_0}$ is convergent, the sequence $\{y - y^i\}_{i\in\mathbb{N}}$ must be bounded and thus also the sequence $\{y^i\}_{i\in\mathbb{N}}$.

3. $\varphi(y^{i+1}) - \theta^{i+1}(y^{i+1}) \to 0$.

   Using successively (C3), (C1) and the definition of the subgradient $s(y^{i+1})$, we obtain

$$\langle s(y^i), y^{i+1} - y^i\rangle \leq \theta^{i+1}(y^{i+1}) - \varphi(y^i) \leq \varphi(y^{i+1}) - \varphi(y^i) \leq \langle s(y^{i+1}), y^{i+1} - y^i\rangle.$$

   Since the subdifferential $\partial\varphi$ is bounded on the bounded sequence $\{y^i\}_{i\in\mathbb{N}}$, the sequence $\{s(y^i)\}_{i\in\mathbb{N}}$ is bounded and, as $\|y^{i+1} - y^i\| \to 0$, the opposite sides of the previous inequalities tend to zero. Hence

$$\theta^{i+1}(y^{i+1}) - \varphi(y^i) \to 0 \quad \text{and} \quad \varphi(y^{i+1}) - \varphi(y^i) \to 0,$$

   and $\quad \varphi(y^{i+1}) - \theta^{i+1}(y^{i+1}) = \varphi(y^{i+1}) - \varphi(y^i) + \varphi(y^i) - \theta^{i+1}(y^{i+1}) \to 0.$

This completes the proof. □

Since $\Delta_k > 0$, it follows from Proposition 2.1, that the STOP occurs after finitely many iterations in the bundle algorithm. So φ^k is well defined, x^{k+1} is the unique solution of problem (P^k), and

$$\varphi^k \leq \varphi \quad \text{and} \quad \varphi(x^{k+1}) - \varphi^k(x^{k+1}) \leq \Delta_k. \tag{2.10}$$

Finally, the sequence $\{x^k\}_{k\in\mathbb{N}}$ generated by applying, step by step, the bundle algorithm is well defined and its convergence can be studied. It is the purpose of the next section.

Remark. When $C \subseteq \operatorname{int}(\operatorname{dom}\varphi)$, the subdifferential $\partial\varphi(x)$ is nonempty on C and there is no need to suppose that int C is nonempty and to consider a barrier function ψ^k in the subproblems (P_i^k). In that case, $\psi^k(x)$ is replaced by $\psi_C(x)$ in problem (P_i^k). When C is given by linear inequalities and h is a strongly convex quadratic function as, for example, $h = 1/2\|\cdot\|^2$, observe that subproblems (P_i^k) become, in fact, convex quadratic programming problems.

3. Convergence of the algorithm. As usual, we proceed in three steps to prove the convergence of the algorithm. First we study the boundedness of the sequence $\{x^k\}_{k\in I\!N}$, then its weak convergence, and finally its strong convergence. In this paper, the sequence $\{\mu_k\}_{k\in I\!N}$ will be chosen under the following form:

$$\begin{cases} \mu_k = \lambda_k/\eta_k, \forall k \in I\!N, \text{ with } \{\lambda_k\}_{k\in I\!N} \text{ a sequence of positive numbers,} \\ \text{and } \eta_k = \begin{cases} \max\{1, \|r(x^0)\|\}, \text{ if } k = 0; \\ \max\{\eta_{k-1}, \|r(x^k)\|\}, \text{ if } k \geq 1. \end{cases} \end{cases}$$

The introduction of the sequence $\{\eta_k\}_{k\in I\!N}$ allows us to prove that the sequence $\{x^k\}_{k\in I\!N}$ is bounded without any additional assumption on the mapping F. Moreover, as it is classically assumed in the multivalued case (see for example [8]), the positive sequence $\{\lambda_k\}_{k\in I\!N}$ will be such that $\sum_{k=0}^{+\infty} \lambda_k^2 < +\infty$, and $\sum_{k=0}^{+\infty} \lambda_k = +\infty$. This rule is also considered in the literature for nonsmooth minimization problems, see, e.g., [1].

In the convergence proofs, we consider the sequence $\{\Gamma^k(x^*,\cdot)\}_{k\in I\!N}$ of Lyapunov functions defined on C by

$$\begin{aligned} \Gamma^k(x^*, x) &= h(x^*) - h(x) - \langle \nabla h(x), x^* - x\rangle \\ &+ (\lambda_k/\eta_k)[\langle r(x^*), x - x^*\rangle + \varphi(x) - \varphi(x^*)], \end{aligned} \tag{3.1}$$

where x^* denotes a solution of problem (P) and $r(x^*)$ is the element in $F(x^*)$ such that $\langle r(x^*), x - x^*\rangle + \varphi(x) - \varphi(x^*) \geq 0$, for all x in C. Since h is strongly convex with modulus $\beta > 0$, we have immediately that, for all $x \in C$,

$$\Gamma^k(x^*, x) \geq (\beta/2)\|x - x^*\|^2. \tag{3.2}$$

The next lemma gives an upper bound on $\Gamma^{k+1}(x^*, x^{k+1}) - \Gamma^k(x^*, x^k)$ which will be often used in the sequel.

LEMMA 3.1. *Suppose that the following conditions are satisfied:*

(*a*) ∇h *is a Lipschitz continuous mapping with Lipschitz constant* Λ *over* C;
(*b*) $\varphi \in \Gamma_0(H)$, *and* $\partial\varphi$ *is bounded on bounded subsets of int* C;
(*c*) $\{\lambda_k\}_{k\in I\!N}$ *is a nonincreasing sequence of positive numbers*;
(*d*) $\{\varphi^k\}_{k\in I\!N}, \varphi \in \Gamma_0(H)$ *are such that (2.10) is satisfied*;
(*e*) $\{\psi^k\}_{k\in I\!N}$ *is a sequence of barrier functions associated with* C *and there exists a sequence* $\{w^k\}_{k\in I\!N}$ *in int* C *such that (2.1) holds.*

Then, if $\{x^k\}_{k\in I\!N}$ *denotes the sequence generated by solving subproblems* (P^k), *we have for all* $k \in I\!N$,

$$\begin{aligned} \Gamma^{k+1}(x^*, x^{k+1}) - \Gamma^k(x^*, x^k) \leq -c\|x^{k+1} - x^k\|^2 + T^k + \lambda_k^2 u + \lambda_0\Delta_k \\ + (\lambda_k/\eta_k)[\langle r(x^k), x^* - x^k\rangle + \varphi(x^*) - \varphi(x^k)], \end{aligned} \tag{3.3}$$

with $c, u > 0$*,* $T^k \geq 0$*, and* $\sum_{k=0}^{+\infty} T^k < +\infty$*. When* H *is a finite dimensional space, the assumption "*$\partial\varphi$ *is bounded on bounded subsets of int* C*" is always true.*

Proof. First observe that the optimality conditions satisfied by x^{k+1} are

$$\begin{aligned} &\langle \eta_k^{-1} r(x^k) + \lambda_k^{-1}(\nabla h(x^{k+1}) - \nabla h(x^k)), x - x^{k+1}\rangle \\ &+\eta_k^{-1}(\varphi^k(x) - \varphi^k(x^{k+1}) + \psi^k(x) - \psi^k(x^{k+1})) \geq 0, \quad \text{for all } x \in \text{int } C, \end{aligned} \tag{3.4}$$

where $r(x^k) \in F(x^k)$. Using the definition of the Lyapunov function and noticing that $\lambda_{k+1} \leq \lambda_k$, and $\eta_{k+1} \geq \eta_k$ for all $k \in I\!N$, we can write

$$\Gamma^{k+1}(x^*, x^{k+1}) - \Gamma^k(x^*, x^k) \leq \Gamma^k(x^*, x^{k+1}) - \Gamma^k(x^*, x^k) = s_1 + s_2 + s_3, \tag{3.5}$$

with
$$\begin{aligned} s_1 &= h(x^k) - h(x^{k+1}) + \langle \nabla h(x^k), x^{k+1} - x^k\rangle, \\ s_2 &= \langle \nabla h(x^k) - \nabla h(x^{k+1}), x^* - x^{k+1}\rangle, \\ s_3 &= (\lambda_k/\eta_k)[\langle r(x^*), x^{k+1} - x^k\rangle + \varphi(x^{k+1}) - \varphi(x^k)]. \end{aligned}$$

For s_1, we derive easily from the strong convexity of h that

$$s_1 \leq -(\beta/2)\|x^{k+1} - x^k\|^2. \tag{3.6}$$

Now, using the sequence $\{w^k\}_{k\in I\!N}$ given in assumption (e), we can write s_2 as the sum of the two following terms:

$$\begin{aligned} s_{21} &= \langle \nabla h(x^k) - \nabla h(x^{k+1}), x^* - w^k\rangle, \\ s_{22} &= \langle \nabla h(x^k) - \nabla h(x^{k+1}), w^k - x^{k+1}\rangle. \end{aligned}$$

From the Lipschitz continuity of ∇h, we deduce that

$$\begin{aligned} s_{21} &\leq \Lambda\|x^{k+1} - x^k\| \, \|x^* - w^k\| \\ &\leq (\tau/2)\|x^{k+1} - x^k\|^2 + (\Lambda^2/(2\tau))\|x^* - w^k\|^2, \end{aligned} \tag{3.7}$$

where the second inequality holds for any $\tau > 0$.
Using (3.4) with $x = w^k$, and noticing that $\psi^k(x^{k+1}) \geq 0$, we obtain

$$\begin{aligned} s_{22} &\leq (\lambda_k/\eta_k)[\langle r(x^k), w^k - x^{k+1}\rangle + \varphi^k(w^k) - \varphi^k(x^{k+1}) + \psi^k(w^k) - \psi^k(x^{k+1})] \\ &= (\lambda_k/\eta_k)[\langle r(x^k), w^k - x^*\rangle \\ &\quad + \langle r(x^k), x^* - x^k\rangle + \varphi(x^*) - \varphi(x^k) \\ &\quad + \langle r(x^k), x^k - x^{k+1}\rangle \\ &\quad + \varphi^k(w^k) - \varphi(x^*) + \varphi(x^k) - \varphi^k(x^{k+1}) + \psi^k(w^k)]. \end{aligned} \tag{3.8}$$

From the definition of the sequence $\{\eta_k\}_{k\in I\!N}$ and the fact that $\lambda_k \leq \lambda_0$ for all k, we have successively

$$(\lambda_k/\eta_k)\langle r(x^k), w^k - x^*\rangle \leq \lambda_0\|w^k - x^*\|, \tag{3.9}$$

$$\begin{aligned} (\lambda_k/\eta_k)\langle r(x^k), x^k - x^{k+1}\rangle &\leq \lambda_k\|x^{k+1} - x^k\| \\ &\leq \lambda_k^2/(2\gamma) + (\gamma/2)\|x^{k+1} - x^k\|^2, \end{aligned} \tag{3.10}$$

where the last inequality holds for any $\gamma > 0$. Moreover, since $\varphi^k \leq \varphi$ for all k, we have

$$\varphi^k(w^k) - \varphi(x^*) \leq \varphi(w^k) - \varphi(x^*) \leq \langle e^k, w^k - x^*\rangle \leq \|e^k\|\|w^k - x^*\|,$$

where e^k is any subgradient of φ at w^k (it exists because $w^k \in \operatorname{int} C$). Then, since $\partial\varphi$ is bounded on the bounded sequence $\{w^k\}_{k\in\mathbb{N}}$, there exists $d > 0$ such that for all k,

$$\varphi^k(w^k) - \varphi(x^*) \le d\|w^k - x^*\|. \tag{3.11}$$

Finally for s_3, we obtain, using (2.10),

$$\begin{aligned} s_3 &\le \lambda_k\|r(x^*)\|\;\|x^{k+1} - x^k\| + (\lambda_k/\eta_k)[\varphi^k(x^{k+1}) + \Delta_k - \varphi(x^k)] \\ &\le (\lambda_k^2/(2\mu))\|r(x^*)\|^2 + (\mu/2)\|x^{k+1} - x^k\|^2 \\ &\quad +(\lambda_k/\eta_k)[\varphi^k(x^{k+1}) - \varphi(x^k)] + \lambda_0\Delta_k \end{aligned} \tag{3.12}$$

with μ any positive number.
Gathering inequalities (3.5)–(3.12) and rearranging the terms, we obtain that inequality (3.3) holds with

$$\begin{array}{lcl} c &=& (1/2)(\beta - \tau - \gamma - \mu), \\ T^k &=& \lambda_0(1+d)\|w^k - x^*\| + \lambda_0\psi^k(w^k) + (\Lambda^2/(2\tau))\|w^k - x^*\|^2, \\ u &=& (1/(2\gamma)) + (1/(2\mu))\|r(x^*)\|^2, \\ \tau, & \gamma, & \mu > 0 \ \text{ such that } \tau + \gamma + \mu < \beta. \end{array}$$

Since the sequence $\{w^k\}_{k\in\mathbb{N}}$ has been chosen such that (2.1) holds, we have that $\sum_{k=0}^{+\infty} T^k < +\infty$. Finally when H is a finite dimensional space, $\partial\varphi$ is always bounded on bounded subsets of int C (see, e.g., [26]). □

The next theorem gives conditions to ensure that the sequence $\{x^k\}_{k\in\mathbb{N}}$ be bounded.

THEOREM 3.2. *Let x^* be a solution of (P). Assume that all assumptions of Lemma 3.1 hold. If $\sum_{k=0}^{+\infty}\lambda_k^2 < +\infty$ and $\sum_{k=0}^{+\infty}\Delta_k < +\infty$, then, provided that $x^0 \in \operatorname{int} C$, the sequence $\{\Gamma^k(x^*, x^k)\}_{k\in\mathbb{N}}$ is convergent, the sequence $\{x^k\}_{k\in\mathbb{N}}$ is bounded, $\sum_{k=0}^{+\infty}\|x^{k+1} - x^k\|^2 < +\infty$ and*

$$\sum_{k=0}^{+\infty}(\lambda_k/\eta_k)[\langle r(x^k), x^k - x^*\rangle + \varphi(x^k) - \varphi(x^*)] < +\infty. \tag{3.13}$$

Proof. Since $r(x^k) \in F(x^k)$ for all k and F is monotone, we have that

$$(\lambda_k/\eta_k)[\langle r(x^k), x^* - x^k\rangle + \varphi(x^*) - \varphi(x^k)] \le 0.$$

So, we derive from (3.3) that

$$\Gamma^{k+1}(x^*, x^{k+1}) - \Gamma^k(x^*, x^k) \le T^k + u\lambda_k^2 + \lambda_0\Delta_k. \tag{3.14}$$

Since the series $\sum_{k=0}^{+\infty} T^k$, $\sum_{k=0}^{+\infty}\lambda_k^2$, $\sum_{k=0}^{+\infty}\Delta_k$ are convergent, it follows that the sequence $\{\Gamma^k(x^*, x^k)\}_{k\in\mathbb{N}}$ is a convergent sequence in H. Using inequality (3.2), we conclude that the sequence $\{x^k\}_{k\in\mathbb{N}}$ is bounded. Then, rearranging the terms of inequality (3.3) as follows

$$\begin{aligned} &c\|x^{k+1} - x^k\|^2 + (\lambda_k/\eta_k)[\,\langle r(x^k), x^k - x^*\rangle + \varphi(x^k) - \varphi(x^*)\,] \\ &\qquad \le \Gamma^k(x^*, x^k) - \Gamma^{k+1}(x^*, x^{k+1}) + T^k + \lambda_k^2 u + \lambda_0\Delta_k, \end{aligned} \tag{3.15}$$

we obtain, using the convergence of the sequence $\{\Gamma^k(x^*, x^k)\}_{k\in \mathbb{N}}$ and of the series given in the statement of the theorem, that $\sum_{k=0}^{+\infty} \|x^{k+1} - x^k\|^2 < +\infty$ and (3.13) holds. $\square$

To prove that any weak limit point of the sequence $\{x^k\}_{k\in \mathbb{N}}$ is a solution of problem (P), we will use the concept of gap function (see e.g., [2]). We recall that a function $l : C \longrightarrow \mathbb{R} \cup \{+\infty\}$ is a gap function with respect to problem (P) if

$$\text{for all } x \in C, l(x) \geq 0 \text{ and } l(\bar{x}) = 0 \text{ if and only if } \bar{x} \text{ is a solution of } (P).$$

We will say that l is weakly lower semicontinuous (l.s.c.) on int C if $x^k \rightharpoonup \bar{x}$, $x^k \in$ int C implies that $\liminf_{k\to+\infty} l(x^k) \geq l(\bar{x})$. In our context, the usefulness of the gap functions appears in the next proposition.

PROPOSITION 3.3. *Let l be a gap function with respect to (P). If l is a weakly l.s.c. function on int C and if $l(x^k) \to 0$, then any weak limit point of the sequence $\{x^k\}_{k\in \mathbb{N}}$ generated by the algorithm is a solution of (P).*

Proof. First, notice that the sequence $\{x^k\}_{k\in \mathbb{N}}$ is contained in int C. Then, let $\bar{x}$ be a weak limit point of this sequence. We have $x^{k_i} \rightharpoonup \bar{x}$ and, by assumption, that

$$0 = \lim_{k\to+\infty} l(x^k) = \liminf_{i\to+\infty} l(x^{k_i}) \geq l(\bar{x}) \geq 0,$$

i.e., $l(\bar{x}) = 0$ and $\bar{x}$ is a solution of (P). $\square$

To prove that $l(x^k) \to 0$, we will use the following lemma due to Cohen and Zhu ([6], Lemma 4).

LEMMA 3.4. *If l is a Lipschitz continuous function on $\{x^k | \, k \in \mathbb{N}\}$ and if $\{\lambda_k\}$ is a sequence of positive numbers such that*
(a) $\sum \lambda_k = +\infty$;
(b) $\sum \lambda_k l(x^k) < +\infty$;
(c) $\exists\, \delta > 0$ such that $\forall k \in \mathbb{N}$, $\|x^{k+1} - x^k\| \leq \delta\lambda_k$,
then $l(x^k) \to 0$.

First we give three existence results of gap functions weakly l.s.c. on int C and Lipschitz continuous on bounded subsets of int C. Then we prove that assumptions (b) and (c) of Lemma 3.4 are satisfied for our algorithm. However, before giving these results, we need to recall some definitions and properties concerning multivalued operators. A multivalued operator F is said to be Lipschitz continuous on a subset B of C if

$$\exists\, L > 0 \text{ such that } \forall x, y \in B \quad e(F(x), F(y)) \leq L \, \|x - y\|,$$

where $e(F(x), F(y)) = \sup_{r\in F(x)} \inf_{s\in F(y)} \|r - s\|$. The next lemma will be used in the sequel.

LEMMA 3.5. *Let B be a bounded subset of C. If F is Lipschitz continuous on B and if there exists $\bar{y} \in B$ such that $F(\bar{y})$ is bounded, then F is bounded on B, i.e., there exists $\alpha > 0$ such that $\|r(x)\| \leq \alpha$ for all $x \in B$ and $r(x) \in F(x)$.*

Proof. Let $\epsilon > 0$. Then, by assumption, $e(F(x), F(\bar{y})) \leq L\|x - \bar{y}\|$ for all $x \in B$, i.e.,

$$\forall x \in B, \forall r(x) \in F(x) \quad \exists\, r(\bar{y}) \in F(\bar{y}) \text{ such that } \|r(x) - r(\bar{y})\| \leq L\|x - \bar{y}\| + \epsilon.$$

Since B and $F(\bar{y})$ are bounded, there exist $\alpha_1 > 0$ and $\alpha_2 > 0$ such that $\|x\| \leq \alpha_1$ for all $x \in B$ and $\|r(\bar{y})\| \leq \alpha_2$ for all $r(\bar{y}) \in F(\bar{y})$. Then, for all $x \in B$ and $r(x) \in F(x)$, we have successively

$$\begin{aligned} \|r(x)\| &\leq \| r(x) - r(\bar{y}) \| + \| r(\bar{y}) \| \\ &\leq L\, [\, \|x\| + \| \bar{y} \| \,] + \epsilon + \alpha_2 \\ &\leq L\, [\, \alpha_1 + \| \bar{y} \| \,] + \epsilon + \alpha_2, \end{aligned}$$

i.e., what we have to prove. □

A multivalued operator F is said to be weakly closed on C if

$$z^k \rightharpoonup \bar{z},\ z^k \in C \text{ and } r^k \rightharpoonup \bar{r}, r^k \in F(z^k) \implies \bar{r} \in F(\bar{z}).$$

In particular, when F is weakly closed on C, then $F(z)$ is a weakly closed subset of H for each $z \in C$. Finally let us mention the following result due to Iusem [16]: If F is paramonotone and if x^* is a solution of (P), then $\bar{x}$ is a solution of (P) if and only if

$$\bar{x} \in C \text{ and } \exists\, \bar{r} \in F(\bar{x}) \text{ such that } \langle \bar{r}, x^* - \bar{x} \rangle + \varphi(x^*) - \varphi(\bar{x}) \geq 0. \tag{3.16}$$

PROPOSITION 3.6. *Let x^* denote any solution of problem (P).*
(a) If F is paramonotone on C and $F(x)$ is a bounded and weakly closed subset of H for all $x \in C$, then $l(x) = \inf_{r(x) \in F(x)} \langle r(x), x - x^* \rangle + \varphi(x) - \varphi(x^*)$ is a gap function.
(b) If, in addition, F and φ are Lipschitz continuous on bounded subsets of int C, then l is Lipschitz continuous on bounded subsets of int C.
(c) If, in addition, F is weakly closed on C, then l is weakly l.s.c. on int C.

Proof. (a) Since F is monotone and x^* is a solution of (P), for each $x \in C$ and $r(x) \in F(x)$, we have

$$\begin{aligned} \langle r(x), x - x^* \rangle + \varphi(x) - \varphi(x^*) &= \langle r(x) - r(x^*), x - x^* \rangle \\ &\quad + \langle r(x^*), x - x^* \rangle + \varphi(x) - \varphi(x^*) \geq 0. \end{aligned}$$

So, using the definition of l, we obtain that $l(x) \geq 0$. Now if $\bar{x}$ is a solution of (P), then we have immediately that

$$l(\bar{x}) \leq \langle r(\bar{x}), \bar{x} - x^* \rangle + \varphi(\bar{x}) - \varphi(x^*) \leq 0 \leq l(\bar{x}).$$

So, $l(\bar{x}) = 0$. Conversely, suppose that $l(\bar{x}) = 0$. Then, by definition of the infimum, there exists a sequence $\{r_k\}_{k \in \mathbb{N}}$ contained in $F(\bar{x})$ such that, for all $k \geq 1$,

$$0 \leq \langle r_k, \bar{x} - x^* \rangle + \varphi(\bar{x}) - \varphi(x^*) < 1/k.$$

Since the subset $F(\bar{x})$ is bounded and weakly closed, there exists a subsequence of $\{r_k\}_{k \in \mathbb{N}}$ that weakly converges to some $r \in F(\bar{x})$. Then $0 \leq \langle r, \bar{x} - x^* \rangle + \varphi(\bar{x}) - \varphi(x^*) \leq 0$, and by (3.16), $\bar{x}$ is a solution of (P) because F is paramonotone.

(b) Let B be a bounded subset of int C and $\alpha_1 > 0$ be such that $\|x\| \leq \alpha_1$ for all $x \in B$. Since φ is Lipschitz continuous on B, it is sufficient to prove that there exists $L_1 > 0$ such that, for all $x, y \in B$,

$$\inf_{r \in F(x)} \langle r, x - x^* \rangle + \sup_{s \in F(y)} \langle s, x^* - y \rangle \leq L_1 \|x - y\|. \tag{3.17}$$

Let $x, y \in B, \epsilon > 0$ and $s \in F(y)$. Since $e(F(y), F(x)) \leq L\|x - y\|$, we have

$$\inf_{r \in F(x)} \|r - s\| \leq L\|x - y\|.$$

So, there exists $r \in F(x)$ such that $\|r - s\| \leq L\|x - y\| + \epsilon/(\alpha_1 + \|x^*\|)$. Then

$$\begin{aligned} \langle r, x - x^* \rangle + \langle s, x^* - y \rangle &= \langle r, x - y \rangle + \langle r - s, y - x^* \rangle \\ &\leq \|r\|\|x - y\| + \|r - s\|\|y - x^*\| \\ &\leq \|r\|\|x - y\| + L\|x - y\|(\alpha_1 + \|x^*\|) + \epsilon. \end{aligned} \tag{3.18}$$

Moreover, by Lemma 3.5, F is bounded on B and consequently, there exists $\alpha > 0$ such that $\|r\| \leq \alpha$ for all $x \in B$ and $r \in F(x)$. Then, from (3.18), we deduce that

$$\inf_{r \in F(x)} \langle r, x - x^* \rangle + \langle s, x^* - y \rangle \leq L_1 \|x - y\| + \epsilon,$$

where $L_1 = \alpha + L(\alpha_1 + \|x^*\|)$. Since this inequality is satisfied for all $s \in F(y)$ and $\epsilon > 0$, we obtain (3.17).

(c) Suppose that F is weakly closed on C. Since φ is weakly l.s.c. on C, we have only to prove that

$$l^1(x) \equiv \inf_{r(x) \in F(x)} \langle r(x), x - x^* \rangle$$

is weakly l.s.c. on int C. Let $x^k \rightharpoonup \bar{x}$ with $x^k \in$ int C, and let $\bar{l}^1$ be a limit point of the sequence $\{l^1(x^k)\}_{k \in I\!N}$. We have to prove that $\bar{l}^1 \geq l^1(\bar{x})$. Without loss of generality, we can suppose that $l^1(x^k) \to \bar{l}^1$. Let then $\epsilon > 0$. By definition of the infimum, for each k, there exists $r(x^k) \in F(x^k)$ such that

$$\langle r(x^k), x^k - x^* \rangle \leq l^1(x^k) + \epsilon. \tag{3.19}$$

Since the sequence $\{x^k\}_{k \in I\!N}$ is bounded and contained in int C, and since F is bounded on bounded subsets of int C, the sequence $\{r(x^k)\}_{k \in I\!N}$ is bounded and thus there exists a subsequence $\{r(x^{k'})\}_{k' \in K}$ weakly converging to some $\bar{r}$. Since F is weakly closed, it follows that $\bar{r} \in F(\bar{x})$. Now, F being monotone, we have that $\langle r(x^{k'}) - r(\bar{x}), x^{k'} - \bar{x} \rangle \geq 0$ and thus that

$$\langle r(x^{k'}), x^{k'} - x^* \rangle \geq \langle r(\bar{x}), x^{k'} - \bar{x} \rangle + \langle r(x^{k'}), \bar{x} - x^* \rangle. \tag{3.20}$$

Gathering (3.19) and (3.20), we obtain

$$l^1(x^{k'}) + \epsilon \geq \langle r(\bar{x}), x^{k'} - \bar{x} \rangle + \langle r(x^{k'}), \bar{x} - x^* \rangle. \tag{3.21}$$

Passing to the limit in (3.21) and noticing that $\langle \bar{r}, \bar{x} - x^* \rangle \geq l^1(\bar{x})$, we have that $\bar{l}^1 + \epsilon \geq l^1(\bar{x})$. Since ϵ is arbitrary, we have that $\bar{l}^1 \geq l^1(\bar{x})$ and consequently l is weakly l.s.c. on int C. □

PROPOSITION 3.7. *Let x^* denote any solution of problem (P). If $F = \partial f, f \in \Gamma_0(H)$ and $C \subseteq int(domf)$, then $l(x) = f(x) + \varphi(x) - f(x^*) - \varphi(x^*)$ is a gap function such that, for all $x \in C$ and $r(x) \in F(x)$,*

$$\langle r(x), x - x^* \rangle + \varphi(x) - \varphi(x^*) \geq l(x).$$

The function l is convex and weakly l.s.c. on C and if, in addition, f and φ are Lipschitz continuous on bounded subsets of int C, then l is also Lipschitz continuous on bounded subsets of int C.

Proof. For all $x \in C, r(x) \in F(x) = \partial f(x)$, we have $f(x^*) \geq f(x) + \langle r(x), x^* - x \rangle$. So, we obtain

$$\langle r(x), x - x^* \rangle + \varphi(x) - \varphi(x^*) \geq f(x) - f(x^*) + \varphi(x) - \varphi(x^*) = l(x).$$

The rest of the proof is obvious. □

PROPOSITION 3.8. *If F is strongly monotone of modulus $\alpha > 0$ on C, then $l(x) = \|x - x^*\|^2$ is a gap function such that, for all $x \in C$ and $r(x) \in F(x)$,*

$$\langle r(x), x - x^* \rangle + \varphi(x) - \varphi(x^*) \geq \alpha l(x).$$

(here x^* denotes the unique solution of (P)). Moreover l is strongly convex, weakly l.s.c. on H and Lipschitz continuous on bounded subsets of C.

Proof. Since x^* is the unique solution of problem (P), it is obvious that l is a gap function and that l is strongly convex and weakly l.s.c. on H. Moreover, for all $x \in C$, we have

$$\langle r(x), x - x^* \rangle + \varphi(x) - \varphi(x^*) = \langle r(x) - r(x^*), x - x^* \rangle + \langle r(x^*), x - x^* \rangle + \varphi(x) - \varphi(x^*).$$

Since F is strongly monotone of modulus α and x^* is the solution of (P), we obtain immediately that the right–hand side of the previous equality is greater than $\alpha l(x)$. Finally, let B be a bounded subset of C. Then there exists $\alpha_1 > 0$ such that $\|z\| \leq \alpha_1$ for all $z \in B$. So, for $x, y \in B$, we have successively

$$\begin{aligned} \|x - x^*\|^2 - \|y - x^*\|^2 &= \|x - y\|^2 + 2\langle x - y, y - x^* \rangle \\ &\leq \|x - y\| [\ \|x - y\| + 2\|y - x^*\|\ ] \\ &\leq \|x - y\| [\ 4\alpha_1 + 2\|x^*\|\ ] \end{aligned}$$

i.e., l is Lipschitz continuous on B. □

In order to get a most general convergence result, we put together, in the same assumption, the properties requested on the gap function. These properties are satisfied in the three situations described in Propositions 3.6, 3.7 and 3.8.

Assumption (I):
(i)$\exists\, \alpha > 0\ \exists\, l : C \to I\!R \cup \{+\infty\}$ such that

$$\forall x \in C,\ \forall r(x) \in F(x) \quad \langle r(x), x - x^* \rangle + \varphi(x) - \varphi(x^*) \geq \alpha\ l(x);$$

(ii) for all $x \in C$, $l(x) \geq 0$ and $l(\bar{x}) = 0 \Leftrightarrow \bar{x}$ is a solution of (P);
(iii) l is weakly l.s.c. on int C and Lipschitz continuous on bounded subsets of int C.

The purpose of the next proposition is to prove that conditions (b) and (c) of Lemma 3.4 are satisfied.

PROPOSITION 3.9. *(a) Assume that assumptions of Theorem 3.2 are satisfied as well as Assumption (I)(i) and (ii). If F is bounded on bounded subsets of int C, then* $\sum \lambda_k l(x^k) < +\infty$.
(b) If $\partial\varphi$ is bounded on bounded subsets of int C and if there exists $\delta_\psi > 0$ such that

$$\psi^k(x^k) - \psi^k(x^{k+1}) \leq \delta_\psi \|x^{k+1} - x^k\| \text{ for all } k \geq 1, \tag{3.22}$$

then there exists $\delta > 0$ such that, for all $k \geq 1$, $\|x^{k+1} - x^k\| \leq \delta\lambda_k$.
(c) When $C = \{x \mid g_i(x) \leq 0, i = 1, \ldots, m\}$ with $g_1, \ldots, g_m$ convex functions from $\mathbb{R}^n$ to $\mathbb{R}$, inequality (3.22) is satisfied by the logarithmic barrier functions (1.6) and by the inverse barrier functions (1.7) provided that the barrier parameters be large enough.

Proof. (a) Since the sequence $\{x^k\}_{k\in\mathbb{N}}$ is bounded and F is bounded on bounded subsets of int C, the sequence $\{r(x^k)\}_{k\in\mathbb{N}}$ is bounded and also the sequence $\{\eta_k\}_{k\in\mathbb{N}}$. Then, using successively Theorem 3.2 and Assumption (I)(i) and (ii), we have

$$\sum_{k=1}^{+\infty} \lambda_k[\langle r(x^k), x^k - x^* \rangle + \varphi(x^k) - \varphi(x^*)] < +\infty \text{ and } \sum_{k=1}^{+\infty} \lambda_k l(x^k) < +\infty.$$

(b) From the optimality conditions (3.4) applied to $x = x^k$, we obtain

$$\begin{aligned} \langle \nabla h(x^{k+1}) - \nabla h(x^k), x^{k+1} - x^k \rangle &\leq (\lambda_k/\eta_k)[\langle r(x^k), x^k - x^{k+1} \rangle \\ &\quad + \varphi^k(x^k) - \varphi^k(x^{k+1}) + \psi^k(x^k) - \psi^k(x^{k+1})]. \end{aligned} \tag{3.23}$$

Since h is strongly convex and $\|r(x^k)\| \leq \eta_k$, we derive from (3.23) that

$$\begin{aligned} \beta\|x^{k+1} - x^k\|^2 &\leq \lambda_k \|x^{k+1} - x^k\| \\ &\quad + (\lambda_k/\eta_k)[\varphi^k(x^k) - \varphi^k(x^{k+1}) + \psi^k(x^k) - \psi^k(x^{k+1})]. \end{aligned} \tag{3.24}$$

Now since $\varphi^k \leq \varphi$ and, by construction (see condition (C4)),

$$\varphi^k(x) \geq \varphi(x^k) + \langle s(x^k), x - x^k \rangle \text{ for all } x \in \text{int } C,$$

we have

$$\begin{aligned} \varphi^k(x^k) - \varphi^k(x^{k+1}) &\leq \varphi(x^k) - \varphi(x^k) - \langle s(x^k), x^{k+1} - x^k \rangle \\ &= \langle s(x^k), x^k - x^{k+1} \rangle \\ &\leq \|s(x^k)\| \, \|x^{k+1} - x^k\|. \end{aligned}$$

So, as $\partial\varphi$ is bounded on bounded subsets of int C, the sequence $\{\|s(x^k)\|\}_{k\in\mathbb{N}}$ is bounded and there exists $\delta_\varphi > 0$ such that, for all k,

$$\varphi^k(x^k) - \varphi^k(x^{k+1}) \leq \delta_\varphi \, \|x^{k+1} - x^k\|. \tag{3.25}$$

Finally, from (3.24), (3.25), (3.22), and since $\eta_k \geq 1$, we deduce that $\|x^{k+1} - x^k\| \leq \delta\lambda_k$ for all k, with $\delta = (1/\beta)[1 + \delta_\varphi + \delta_\psi]$.

(c) Let $\psi^k(x) = \nu_k^{-1} b(x)$ with

$$b(x) = -\sum_{i=1}^{m} \log(\min(\frac{1}{2}, -g_i(x))) \qquad \text{or} \qquad b(x) = -\sum_{i=1}^{m} \frac{1}{g_i(x)}.$$

Let $e^k \in \partial b(x^k)$ (it exists because b is convex and $x^k \in \text{int } C$). Then

$$\begin{array}{rcl} \psi^k(x^k) - \psi^k(x^{k+1}) & \leq & \nu_k^{-1}\langle e^k, x^k - x^{k+1}\rangle \\ & \leq & \nu_k^{-1}\|e^k\|\,\|x^k - x^{k+1}\|. \end{array}$$

So, if $\nu_k \geq \|e^k\|$, then $\psi^k(x^k) - \psi^k(x^{k+1}) \leq \|x^{k+1} - x^k\|$. □

Notice that the choice $\nu_k \geq \|e^k\|$ is possible because ψ^k is built once x^k is known.

We are now ready to state our main convergence result.

THEOREM 3.10. *Suppose that the following conditions are satisfied:*
(a) Assumption (I) holds;
(b) Assumptions of Lemma 3.1 hold;
(c) F is bounded on bounded subsets of int C;
(d) Inequality (3.22) holds;
(e) $\sum \lambda_k = +\infty$, $\sum \lambda_k^2 < +\infty$, $\sum \Delta_k < +\infty$.
Then the sequence $\{x_k\}_{k\in\mathbb{N}}$ is bounded, $l(x^k) \to 0$ and any weak limit point of $\{x_k\}_k$ (and there exists at least one such point) is a solution of problem (P).
If, in addition, ∇h is weakly continuous on C, then $x^k \rightharpoonup \bar{x}$ and $\bar{x}$ is a solution of (P). If, in addition, the gap function l is strongly convex on an open set containing C, then $x^k \to x^*$, the unique solution of (P). When H is a finite dimensional space, the (strong) convergence of the whole sequence toward a solution is true under the only assumptions (a)–(e).

Proof. The first part of the theorem follows immediately from Lemma 3.1, Theorem 3.2, Proposition 3.9, Lemma 3.4 and Proposition 3.3. Suppose now that ∇h is weakly continuous on C and that the sequence $\{x^k\}_{k\in\mathbb{N}}$ has two different weak limit points x^1 and x^2. Let $\{x^{m(k)}\}_{k\in\mathbb{N}}$ the subsequence of $\{x^k\}_{k\in\mathbb{N}}$ weakly converging to x^1 and $\{x^{n(k)}\}_{k\in\mathbb{N}}$ the subsequence weakly converging to x^2. By the first part of the theorem, x^1 and x^2 are solutions of problem (P). Then, by Theorem 3.2, the sequences of Lyapunov functions $\{\Gamma^k(x^1, x^k)\}_{k\in\mathbb{N}}$ and $\{\Gamma^k(x^2, x^k)\}_{k\in\mathbb{N}}$ are convergent in $\mathbb{R}$. We denote respectively by Γ_1 and Γ_2 their limits. By definition of the Lyapunov function, we have

$$\begin{array}{l} \Gamma^{n(k)}(x^1, x^{n(k)}) - \Gamma^{n(k)}(x^2, x^{n(k)}) \\ = h(x^1) - h(x^2) - \langle\nabla h(x^{n(k)}), x^1 - x^2\rangle \\ +(\lambda_{n(k)}/\eta_{n(k)})[\langle r(x^1), x^{n(k)} - x^1\rangle - \langle r(x^2), x^{n(k)} - x^2\rangle + \varphi(x^2) - \varphi(x^1)]. \end{array}$$

Since ∇h is weakly continuous on C, since $\eta_k \geq 1$ for all k and since, by (e), $\lambda_k \to 0$, we obtain, taking the limit on k in the last equality, that

$$\Gamma_1 - \Gamma_2 = h(x^1) - h(x^2) - \langle\nabla h(x^2), x^1 - x^2\rangle. \tag{3.26}$$

Since the role of x^1 and x^2 is symmetric, we also have that

$$\Gamma_1 - \Gamma_2 = h(x^1) - h(x^2) - \langle \nabla h(x^1), x^1 - x^2 \rangle. \tag{3.27}$$

Comparing (3.26) and (3.27), we obtain $\langle \nabla h(x^1) - \nabla h(x^2), x^1 - x^2 \rangle = 0$. Since ∇h is strongly monotone, this inequality implies that $x^1 = x^2$. So the sequence $\{x^k\}_{k \in \mathbb{N}}$ weakly converges to a solution of (P).
If the gap function l is strongly convex with constant $s > 0$ on an open convex set containing C, then x^* is the unique solution of problem (P), $\partial l(x^*)$ is nonempty, and for any $e^* \in \partial l(x^*)$,

$$l(x^k) - l(x^*) - \langle e^*, x^k - x^* \rangle \geq (s/2)\|x^k - x^*\|^2. \tag{3.28}$$

Since $l(x^k) \to 0$, $l(x^*) = 0$ and $x^k \rightharpoonup x^*$, we obtain, passing to the limit in (3.28) that $\|x^k - x^*\| \to 0$, i.e., $x^k \to x^*$ strongly.
Finally when H is a finite dimensional space, it follows from assumption (a) of Lemma 3.1 that ∇h is continuous in the strong topology and thus in the weak topology. Then the convergence of the whole sequence toward a solution is established under the only assumptions (a)–(e). This completes the proof. □

For example, if $h(x) = (1/2)x^T x$, for all $x \in H$, then ∇h is weakly continuous on H.

Using Propositions 3.6, 3.7 and 3.8 which give sufficient conditions to ensure that Assumption (I) is satisfied, we can particularize our main result (Theorem 3.10), to get two other convergence theorems. However, before presenting them, and for the sake of simplicity, we collect in a statement, several assumptions used previously, and we prove a preliminary lemma.

Assumption (II):
(i) Assumptions of Lemma 3.1 hold;
(ii) ∇h is weakly continuous on C;
(iii) Inequality (3.22) holds;
(iv) $\sum \lambda_k = +\infty$, $\sum \lambda_k^2 < +\infty$, $\sum \Delta_k < +\infty$.

Notice that when H is a finite dimensional space, (ii) is always true as also it is the case for the second assumption of Lemma 3.1: $\partial\varphi$ is bounded on bounded subsets of int C.

LEMMA 3.11. *Let $g \in \Gamma_0(H)$ and let B be a bounded subset of int(dom g). If ∂g is bounded on B, then g is Lipschitz continuous on B.*

Proof. Let $x, y \in B$. Since $B \subseteq$int(dom g), the subdifferentials $\partial g(x)$ and $\partial g(y)$ are nonempty. Let $s(x) \in \partial g(x)$ and $s(y) \in \partial g(y)$. Then

$$\begin{array}{lll} g(x) - g(y) & \leq \langle s(x), x - y \rangle & \leq \|s(x)\| \, \|x - y\| \\ g(y) - g(x) & \leq \langle s(y), y - x \rangle & \leq \|s(y)\| \, \|y - x\|. \end{array}$$

So $|g(x) - g(y)| \leq L\|x - y\|$ where $L = \sup\{\|s(z)\| \mid z \in B, s(z) \in \partial g(z)\}$. Since ∂g is bounded on B, this constant L is finite and thus g is Lipschitz continuous on B. □

THEOREM 3.12. *Suppose that Assumption (II) holds.*
(a) If F is paramonotone, weakly closed on C and Lipschitz continuous on bounded subsets of int C, and if $F(x)$ is a bounded subset of H for all $x \in C$, then the whole sequence $x^k \rightharpoonup \bar{x}$ where $\bar{x}$ is a solution of (P).
(b) If $F = \partial f$ with $f \in \Gamma_0(H)$ and $C \subseteq$ int(dom$f$), and if ∂f is bounded on bounded subsets of int C, then the whole sequence $x^k \rightharpoonup \bar{x}$ where $\bar{x}$ is a solution of (P).
When H is a finite dimensional space, the assumption on ∂f is always true.

Proof. By Theorem 3.10, it is sufficient to prove that Assumption (I) holds and that F is bounded on bounded subsets of int C.
(a) Since $\partial\varphi$ is bounded on bounded subsets of int C, it follows from Lemma 3.11 that φ is Lipschitz continuous on bounded subsets of int C. All the assumptions of Proposition 3.6 are then satisfied and thus Assumption (I) is satisfied. Finally, using Lemma 3.5, F is bounded on bounded subsets of int C.
(b) By Lemma 3.11, f and φ are Lipschitz continuous on bounded subsets of int C. So, using Proposition 3.7, Assumption (I) is satisfied. The conclusion follows because $F = \partial f$ is bounded on bounded subsets of int C. □

THEOREM 3.13. *Suppose Assumption (II) holds. If F is strongly monotone on C and bounded on bounded subsets of int C, then the whole sequence x^k strongly converges to x^*, the unique solution of (P).*

Proof. ¿From Proposition 3.8, we have that Assumption (I) is satisfied. Then the conclusion follows from Theorem 3.10 because F is bounded on bounded subsets of int C and the gap function $l(x) = \|x - x^*\|^2$ is strongly convex on H. □

Remark. When $C \subseteq \text{int}(\text{dom}\varphi)$, the subdifferential $\partial\varphi(x)$ is nonempty on C and there is no need to suppose that int C is nonempty and to introduce a barrier function in subproblems (P^k). However, all our convergence results remain true in that case, provided that each assumption made on int C be extended to C.

To conclude, all our convergence results suppose that the sequence $\{\varphi^k\}_{k\in\mathbb{N}}$ of functions approximating φ, satisfy inequalities (2.10) where $\{\Delta_k\}_{k\in\mathbb{N}}$ is a sequence of positive numbers such that $\sum \Delta_k < +\infty$. But practically, the choice of the sequence $\{\Delta_k\}_{k\in\mathbb{N}}$ remains an open issue. To be efficient, the tolerance Δ_k should be determined, not in advance, but once x^k has been found by taking into account the behavior and the progress of the iterates to the solution. Such a strategy has been proposed in nonsmooth optimization, for solving nondifferentiable convex minimization problems (see [9]). In our context, this question deserves more investigations.

REFERENCES

[1] Y.I. ALBER, A.N. IUSEM AND M.V. SOLODOV, *On the projected subgradient method for nonsmooth convex optimization in a Hilbert space*, Math. Program., 81 (1998), pp. 23–35.

[2] G. AUCHMUTY, *Variational principles for variational inequalities*, Numerical Functional Analysis and Optimization, 10 (1989), pp. 863–874.

[3] J.P. AUBIN AND I. EKELAND, *Applied Nonlinear Analysis*, Wiley, New York, New York, 1984.

[4] R.D. BRUCK, *An iterative solution of a variational inequality for certain monotone operators in Hilbert space*, Bulletin of the American Math. Soc., 81 (1975), pp. 890–892.

[5] R.S. BURACHIK AND A.N. IUSEM, *A generalized proximal point algorithm for the variational inequality problem in a Hilbert space*, SIAM J. Optim. 8 (1998), pp. 197–216.

[6] G. COHEN AND D.L. ZHU, *Decomposition coordination methods in large scale optimization problems: the nondifferentiable case and the use of augmented lagrangians* in: Advances in Large Scale Systems Theory and Applications, J.B. Cruz, ed., Vol. 1, JAI Press, Greenwich, Connecticut, USA, 1984, pp. 203–266.

[7] G. COHEN, *Nash equilibria: gradient and decomposition algorithms*, Large Scale Systems, 12 (1987), pp. 173–184.

[8] G. COHEN, *Auxiliary problem principle extended to variational inequalities*, J. Optim. Theory Appl., 59 (1988), pp. 325–333.

[9] R. CORREA AND C. LEMARÉCHAL, *Convergence of some algorithms for convex minimization*, Math. Program., 62 (1993), pp. 261–275.

[10] R.W. COTTLE, F. GIANNESSI AND J.L. LIONS, *Variational Inequalities and Complementarity Problems: Theory and Applications*, Wiley, New York, New York, 1980.

[11] J.P. CROUZEIX, *Pseudomonotone variational inequalitiy problems: existence of solutions* Math. Program., 78 (1997), pp. 305–314.

[12] J. ECKSTEIN, *Approximate iterations in Bregman–function–based proximal algorithms*, Math. Program., 83 (1998), pp. 113–124.

[13] I. EKELAND AND R. TEMAM, *Convex Analysis and Variational Inequalities*, North–Holland, Amsterdam, 1976.

[14] R. GLOWINSKI, J.L. LIONS AND R. TRÉMOLIÈRE, *Numerical Analysis of Variational Inequalities*, North–Holland, Amsterdam, 1981.

[15] P.T. HARKER AND J.S. PANG, *Finite–dimensional variational inequality and nonlinear complementarity problems. A survey of theory, algorithms and applications*, Math. Program., 48 (1990), pp. 161–220.

[16] A.N. IUSEM, *On some properties of paramonotone operators*, J. Convex Anal., 5 (1998), pp. 269–278.

[17] D. KINDERLEHRER AND G. STAMPACCHIA, *An Introduction to Variational Inequalities and their Application*, Academic Press, New York, New York, 1980.

[18] I. KONNOV, *Combined Relaxation Methods for Variational Inequalities*, Lecture Notes in Economics and Mathematical Systems, Vol. 495, Springer–Verlag, Berlin, 2001.

[19] P.L. LIONS AND B. MERCIER, *Splitting algorithms for the sum of two nonlinear operators*, SIAM J. Num. Anal., 16 (1979), pp. 964–979.

[20] B. MARTINET, *Régularisation d'inéquations variationnelles par approximations successives*, Revue Française d'Automatique de Informatique Recherche Opérationnelle, 4 (1970), pp. 154–159.

[21] S. MAKLER–SCHEIMBERG, V.H. NGUYEN AND J.J. STRODIOT, *Family of perturbation methods for variational inequalities*, J. Optim. Theory Appl., 89 (1996), pp. 423–452.

[22] U. MOSCO, *Convergence of convex sets and of solutions of variational inequalities*, Advances in Mathematics, 3 (1969), pp. 510–585.

[23] A. MOUDAFI AND M., NOOR, *New convergence results of iterative methods for set–valued mixed variational inequalities*, Mathematical Inequalities and Applications, 3 (2000), pp 295–303.

[24] P. PANAGIOTOPOULOS AND G. STAVROULAKIS, *New types of variational principles based on the notion of quasidifferentiability*, Acta Mechanics, 94 (1994), pp. 171–194.

[25] J.S. PANG, *Asymmetric variational inequality problems over product sets: applications and iterative methods*, Math. Program., 31 (1985), pp. 206–219.

[26] R.T. ROCKAFELLAR, *Convex Analysis*, Princeton University Press, Princeton, New Jersey, 1970.

[27] R.T. ROCKAFELLAR, *On the maximal monotonicity of subdifferential mapping*, Pacific Journal of Mathematics, 33 (1970), pp. 209–216.

[28] R.T. ROCKAFELLAR, *Monotone operators and the proximal point algorithm*, SIAM J. Control Optim., 14 (1976), pp. 877–898.

[29] G. SALMON, V.H. NGUYEN AND J.J. STRODIOT, *Coupling the auxiliary problem principle and the epiconvergence theory to solve general variational inequalities*, J. Optim. Theory Appl.,

104 (2000), pp. 629–657.

[30] G. Salmon, V.H. Nguyen and J.J. Strodiot, *A perturbed and inexact version of the auxiliary problem method for solving general variational inequalities with a multivalued operator*, in: Optimization, V.H. Nguyen, J.J. Strodiot and P. Tossings, eds., Lecture Notes in Economics and Mathematical Systems, Vol. 481, Springer–Verlag, Berlin, 2000, pp. 396–418.

[31] G. Salmon, J.J. Strodiot and V.H. Nguyen, *A pertubed auxiliary problem method for paramonotone multivalued mappings*, in: Advances in Convex Analysis and Global Optimization, N. Hadjisavvas and P. Pardalos, eds., Non Convex Optimization and Applications Series, Kluwer Academic Publishers, Dordrecht, 2001, to appear.

[32] M. Teboulle, *Convergence of proximal–like algorithms*, SIAM J. Optim., 7 (1997), pp. 1069–1083.

[33] P. Tseng, *Applications of a splitting algorithm to decomposition in convex programming and variational inequalities*, SIAM J. Control Optim., 29 (1991), pp. 119–138.

[34] D. Zhu, *The decomposition method for the variational inequalities with multivalued mapping*, Private communication, 1999.