Commun. Optim. Theory 1 (2012), 1-18

MATHRES
Communications in Optimization Theory
Available online at http://cot.mathres.org

AN APPROXIMATE BUNDLE METHOD FOR SOLVING VARIATIONAL INEQUALITIES

JIE SHEN[1,*], LI-PING PANG[2]

[1]School of Mathematics, Liaoning Normal University, Dalian 116029, China

[2]School of Mathematical Sciences, Dalian University of Technology, Dalian 116024, China

Abstract. We present an approximate algorithm for solving generalized variational inequality problem (P) by combining the bundle technique with the auxiliary problem method. For subproblems we construct a new lower approximation to the involved function φ, as in the bundle method for nonsmooth optimization, which only requires the approximate function values and approximate subgradients instead of the exact ones. This makes the subproblem easier to solve and more tractable. Besides that, a new stopping criterion is given to determine whether the current approximation is good enough, it not only takes into account the contribution of the operator F in problem (P), but also utilizes the inexact information of the involved function φ in subproblem. Finally, we study the convergence of the proposed algorithm for the case when the operator F is paramonotone, possibly multivalued and the stepsizes are chosen going to zero.

Keywords: generalized variational inequality; auxiliary problem principle; nonsmooth optimization; bundle method; approximate function value and subgradient.

2000 AMS Subject Classification: 49J40; 90C25; 90C90

1. Introduction

*Corresponding author

Received August 15, 2012

Supported by the National Nature Science Foundation of China (Grant 11171138, 11171049, 11071029) and Higher School Research Project of Educational Department of Liaoning Province (Grant L2010235).

Let F be a monotone multivalued operator defined on a real Hilbert space H with inner product $\langle\cdot\rangle$ and the associated norm $\|\cdot\|$, let C be a nonempty closed convex subset of H, and let $\varphi : H \to \bar{R} = R \cup \{+\infty\}$ be a l.s.c. proper convex function.

Consider the generalized variational inequality problem:

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

In this paper, we assume that $C \subseteq \text{int}(\text{dom}\varphi)$ and there exists at least one solution to problem (P).

In 1988, Cohen [1] developed an algorithm framework for solving problem (P) based on the so-called auxiliary problem principle. More precisely, let Ω be a strongly monotone and Lipschitz continuous auxiliary operator on H, and $\{\mu_k\}_{k\in N}$ be a sequence of positive real numbers. The problem considered at iteration k is the following:

$$
\begin{cases} \text{choose } r(x^k) \in F(x^k) \text{ and } x^{k+1} \in C \text{ such that} \\ \langle r(x^k) + \mu_k^{-1}[\Omega(x^{k+1}) - \Omega(x^k)], x - x^{k+1} \rangle + \varphi(x) - \varphi(x^{k+1}) \geq 0, \ \forall x \in C. \end{cases}
$$

In this paper, we choose Ω as the gradient of some continuously differentiable and strongly convex function h with Lipschitz continuous gradients. In that case, the subproblem can be equivalently written in the following minimization form:

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

When φ is a nonsmooth convex function, subproblem (AP^k) may be hard to solve. By employing bundle ideas, several authors proposed approximating φ by a sequence of more tractable convex functions, see [2, 3, 4, 5, 6, 7]. The strategy is to approximate the function φ, at the proximal iteration k, by a piecewise linear convex function, built step by step, and to move to the next iterate only when the approximation is suitable (good enough). This gives the following algorithm, see [8].

Bundle Algorithm for Solving Problem (P):

Let an initial point x^0 be given, together with a tolerance $m \in (0,1)$ and a positive number sequence $\{\mu_k\}_{k\in N}$. Compute $r(x^0) \in F(x^0)$. Set $y^0 = x^0, k = 0, i = 1$.

Step 1: Choose a piecewise linear convex function $\theta_a^i \leq \varphi$ and solve problem

$$(P_i^k) \qquad \min_{x\in C}\{\theta_a^i(x) + \langle r(x^k), x - x^k\rangle + \mu_k^{-1}[h(x) - h(x^k) - \langle \nabla h(x^k), x - x^k\rangle]\}$$

to obtain the unique solution $y^i \in C$.

Step 2: If the trial point y^i is good enough, i.e., if

$$\varphi(x^k) - \varphi(y^i) \geq m[\varphi(x^k) - \theta_a^i(y^i)] + (1-m)\langle r(x^k), y^i - x^k\rangle, \tag{1.1}$$

then set $x^{k+1} = y^i$, compute $r(x^{k+1}) \in F(x^{k+1})$ and let $k = k+1$.

Step 3: Let $i = i+1$ and go to Step 1.

The trial point $y^i \in C$ is obtained by solving (P_i^k), namely, subproblem (AP^k) with the function φ replaced by the approximation $\theta_a^i \leq \varphi$. This approximation is good enough if the stopping criterion (1.1) is satisfied. (1.1) is a new criterion when compared with the one in [2], it takes into account the contribution of the operator F. When the stopping criterion (1.1) holds, the outer iterate x^k is updated and we say that a serious step is made; otherwise, x^k is kept fixed for the next inner iteration, which will be performed with an improvement of the approximation θ_a^i. This step is called null-step. In this paper, we still employ the above framework to solve problem (P), but in our case, the function φ in subproblem (AP^k) will be replaced by some new approximation θ_a^i by utilizing the approximate values of φ and its approximate subgradients since in some cases computing the exact function value is not so easy. The assumptions for using approximate subgradients and approximate values of the function are realistic in many applications, for instance, the Lagrangian relaxation problem: if f is a max-type function of the form

$$f(y) = \sup\{F_z(y) \mid z \in Z\}, \tag{$*$}$$

where each $F_z(y)$ is convex and Z is an infinite set, then it may be impossible to calculate $f(y)$ since f itself is defined by a minimization problem involving another function F. However, we may still consider two cases. In the first case, for each positive $\varepsilon > 0$ one can find an ε-maximizer of (*), i.e., an element $z_y \in Z$ satisfying $F_{z_y}(y) \geq f(y) - \varepsilon$; in the second case, this may be possible only for some fixed (any possibly unknown) $\varepsilon < \infty$. In both cases we may set $\bar{f}_y = F_{z_y}(y) \geq f(y) - \varepsilon$. A special case of (*) arises from Lagrangian relaxation [9], where the problem $\min\{f(y) \,|\, y \in S\}$ with $S = R^n_+$ is the Lagrangian dual of the primal problem

$$\sup \;\; \psi_0(z) \quad \text{ s.t. } \psi_j(z) \geq 0, \;\; j = 1, 2, \cdots, n, \;\; z \in Z,$$

with $F_{z_y}(y) = \psi_0(z) + \langle y, \psi(z) \rangle$ for $\psi = (\psi_1, \psi_2, \cdots, \psi_n)$. Then, for each multiplier $y \geq 0$, we need only find $z_y \in Z$ such that $\bar{f}_y = F_{z_y}(y) \geq f(y) - \varepsilon$, see [10]. Besides that, the study of approximate subgradients of convex functions is deserved since in some cases a subgradient $\xi(x) \in \partial f(x)$ is expensive to compute. But if we know an already computed subgradient $\xi(\bar{x}) \in \partial f(\bar{x})$, where $\bar{x}$ is near to x, then we have $\xi(\bar{x}) \in \partial_\varepsilon f(x)$ because

$$\begin{aligned} f(x) + \xi(\bar{x})^T(z - x) &= f(\bar{x}) + \xi(\bar{x})^T(z - \bar{x}) + \varepsilon \\ &\leq f(z) + \varepsilon, \qquad\quad \forall\, z \in R^n, \end{aligned}$$

where $\varepsilon = f(x) - f(\bar{x}) - \xi(\bar{x})^T(x - \bar{x}) \geq 0$.

On the basis of the above observation, we attempt to explore the possibility of using the approximate values and approximate subgradients of φ instead of the exact values for solving problem (P). Throughout this paper, we make the following assumption: at each given point $x \in \text{dom}\varphi$ and $\varepsilon \geq 0$ we can find some $\tilde{\varphi}^\varepsilon_x \in R$ and $\tilde{\xi}^\varepsilon_x \in H$ such that

$$\varphi(x) \geq \tilde{\varphi}^\varepsilon_x \geq \varphi(x) - \varepsilon, \quad \varphi(\zeta) \geq \tilde{\varphi}^\varepsilon_x + \langle \tilde{\xi}^\varepsilon_x, \zeta - x \rangle, \qquad \forall\, \zeta \in H., \tag{1.2}$$

The condition (1.2) means that $\tilde{\xi}^\varepsilon_x \in \partial_\varepsilon \varphi(x)$. This assumption is realistic in many applications, see [11, 12]. For more details and papers involving the approximate function values and subgradients, we refer to [13, 14, 15] and the references therein.

Other contributions to the construction of bundle methods for variational inequalities have appeared in the literature. For instance, in [16, 17], a bundle method is presented for finding a zero of a maximal monotone operator T defined on H. The main difference between [17] and our method is that our method takes into account the special structure of $T = F + \partial(\varphi + \psi_C)$, ($\psi_C = 0$ if $x \in C$; otherwise, $\psi_C = +\infty$) by using bundle technique not on the operator T but directly on the function φ. The bundle method for solving variational inequalities presented in [2] used the classical stopping criterion for inner iteration and the exact function values and its subgradients. In our method, we not only employ the inexact information of the involved function, but also utilize a new stopping criterion for inner iteration which takes into account the contribution of the operator F.

This paper is organized as follows: In Section 2, by employing the bundle idea we present an approximate algorithm for problem (P) and prove that if the stopping test of the inner iteration is suppressed, then the current iterate x^k is an approximate solution to problem (P). In Section 3 the strong convergence of the proposed algorithm is given under rather mild conditions that the operator F is paramonotone and possibly multivalued and the stepsizes are chosen going to zero.

2. An approximate algorithm for problem (P)

Under the assumption (1.2), we approximate φ from below by a piecewise linear convex function $\theta_a^i(y)$ defined by

$$\theta_a^i(y) = \max_{0\le j\le i-1}\{\tilde{\varphi}_y^j + \langle \tilde{s}_y^j, y - y^j\rangle\}, \tag{2.1}$$

where $\tilde{\varphi}_y^j \in H$ and $\tilde{s}_y^j \in R$ satisfy

$$\varphi(y^j) \ge \tilde{\varphi}_y^j \ge \varphi(y^j) - \varepsilon_j, \quad \varphi(\zeta) \ge \tilde{\varphi}_y^j + \langle \tilde{s}_y^j, \zeta - y^j\rangle, \qquad \forall\, \zeta \in H, \tag{2.2}$$

for given $\varepsilon_j \ge 0$, $\varepsilon_{j+1} = \gamma\varepsilon_j$, $0 < \gamma < 1$, $y^j \in H, j = 0, 1, 2, \cdots, i-1$. We give a new stopping criterion

$$\tilde{\varphi}_x^k - \tilde{\varphi}_y^i \ge m(\tilde{\varphi}_x^k - \theta_a^i(y^i)) + (1-m)\langle r(x^k), y^i - x^k\rangle, \tag{2.3}$$

where $m \in (0,1)$, $\tilde{\varphi}_x^k$ satisfies $\varphi(x^k) \ge \tilde{\varphi}_x^k \ge \varphi(x^k) - \varepsilon^k, \varepsilon_0$ is given, $\varepsilon^k = \min\{\varepsilon_0, \varepsilon_1, \cdots, \varepsilon_{i-1}, \varepsilon^{k-1}\}$, $y^i \in C$ is the unique solution to the subproblem

$$(BAP^k) \qquad \min_{y\in C}\{\theta_a^i(y) + \langle r(x^k), y - x^k\rangle + \mu_k^{-1}[h(y) - h(x^k) - \langle \nabla h(x^k), y - x^k\rangle]\},$$

for given $x^k \in C$, $\{\mu_k\}_{k\in N}$ is a sequence of positive real numbers, $\tilde{\varphi}_y^i$ satisfies $\varphi(y^i) \ge \tilde{\varphi}_y^i \ge \varphi(y^i) - \varepsilon_i$. In our algorithm, (2.3) will replace (1.1) to determine whether this approximation is good enough and it is an approximation to the original one coming from [8].

Approximate Bundle Algorithm for Solving Problem (P):

Let an initial point x^0 be given, together with a tolerance $m \in (0,1)$ and a positive number sequence $\{\mu_k\}_{k\in N}$. Compute $r(x^0) \in F(x^0)$. Set $y^0 = x^0, k = 0, i = 1$.

Step 1: Choose an approximate piecewise linear convex function $\theta_a^i \leq \varphi$ and solve problem

$$(P_i^k) \quad \min_{x\in C}\{\theta_a^i(x) + \langle r(x^k), x - x^k\rangle + \mu_k^{-1}[h(x) - h(x^k) - \langle \nabla h(x^k), x - x^k\rangle]\}$$

to obtain the unique solution $y^i \in C$.

Step 2: If the trial point y^i is good enough, i.e., if

$$\tilde{\varphi}_x^k - \tilde{\varphi}_y^i \geq m(\tilde{\varphi}_x^k - \theta_a^i(y^i)) + (1-m)\langle r(x^k), y^i - x^k\rangle,$$

then set $x^{k+1} = y^i$, compute $r(x^{k+1}) \in F(x^{k+1})$ and let $k = k+1$.

Step 3: Let $i = i+1$ and go to Step 1.

Note that the optimality of $y^i \in C$, we have

$$\gamma^i \equiv \mu_k^{-1}[\nabla h(x^k) - \nabla h(y^i)] - r(x^k) \in \partial(\theta_a^i + \psi_C)(y^i). \tag{2.4}$$

Then we define the aggregate affine function l^i by

$$l^i(y) = \theta_a^i(y^i) + \langle \gamma^i, y - y^i\rangle, \tag{2.5}$$

therefore

$$l^i(y) \leq \theta_a^i(y), \qquad \forall\, y \in C. \tag{2.6}$$

If we approximate φ by θ_a^i defined by (2.1), then the following results hold naturally:

$(AC1)$ $\quad \theta_a^i \leq \varphi$ on C;

$(AC2)$ $\quad l^i \leq \theta_a^{i+1}$ on C, $\forall\, i \in (i_k, i_{k+1})$;

$(AC3)$ $\quad \tilde{\varphi}_y^i + \langle \tilde{s}_y^i, \cdot - y^i\rangle \leq \theta_a^{i+1}(\cdot)$, $\forall\, i \in (i_k, i_{k+1})$;

$(AC4)$ $\quad \varepsilon^k + \tilde{\varphi}_x^k + \langle \tilde{s}_x^k, x^{k+1} - x^k\rangle \leq \theta_a^{i_{k+1}}(x^{k+1})$, with $\varepsilon^k \geq 0$ small enough,

where i_k denotes the inner iteration that has produced x^k $(i_0 = 0)$. In what follows, we need to consider the following functions:

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

$\tilde{\theta}_a^i(y) = \theta_a^i(y) + \langle r(x^k), y - x^k\rangle + \mu_k^{-1}[h(y) - h(x_k) - \langle \nabla h(x^k), y - x^k\rangle]$. By a simple calculation, we have $\tilde{l}^i(y) = \tilde{l}^i(y^i) + \mu_k^{-1}[h(y) - h(y^i) - \langle \nabla h(y^i), y - y^i\rangle], \tilde{\theta}_a^i(x^k) = \theta_a^i(x^k),\ \tilde{l}^i(y^i) = \tilde{\theta}_a^i(y^i),\ \tilde{l}^i \leq \tilde{\theta}_a^{i+1}$ on C.

Now we study the convergence of bundle algorithm.

Assumption A:

- Problem (P) admits at least one solution;
- F is a monotone operator defined on H;
- $\varphi \in \Gamma_0(H)$, the set of l.s.c. proper convex functions form H into $\bar{R} = R \cup \{+\infty\}$;
- C is a nonempty closed convex subset of H such that $C \subseteq \mathrm{int}(\mathrm{dom}\varphi)$;
- $\partial_\varepsilon \varphi$ is uniformly bounded with respect to $\varepsilon \geq 0$ on bounded subsets of C;
- $h : H \to R$ is continuously differentiable and strongly convex over C with modulus $\beta > 0$, and its gradient ∇h is Lipschitz continuous over C with modulus $\gamma > 0$;
- the sequence $\{\theta_a^i\}_{i \in N}$ satisfies conditions $(AC1) - (AC3)$.

Proposition 2.1. *Suppose Assumption A holds. If the stopping test is suppressed in approximate bundle algorithm after some outer iterate x^k has been reached, then the sequence $\{y^i\}_{i\in N}$ generated by approximate bundle algorithm is bounded and $\tilde{\varphi}_y^i \to \theta_a^i(y^i)$, $y^i \to z(x^k)$ as $i \to \infty$, where $z(x^k) = argmin_{x \in C}\{\varphi(x) + \langle r(x^k), x - x^k\rangle + \mu_k^{-1}[h(x) - h(x^k) - \langle \nabla h(x^k), x - x^k\rangle]\}$.*

Proof. Since i_k denotes the inner iteration that has produced x^k, and only null steps are made after reaching x^k, all the following inequalities have to be understood for $i > i_k$. We proceed in three steps to show $\tilde{\varphi}_y^i \to \theta_a^i(y^i)$.

1. For all i, we have

$$\begin{aligned}
\varphi(x^k) &\geq \theta_a^{i+1}(x^k) \\
&= \tilde{\theta}_a^{i+1}(x^k) \\
&\geq \tilde{\theta}_a^{i+1}(y^{i+1}) \\
&= \tilde{l}^{i+1}(y^{i+1}) \\
&\geq \tilde{l}^i(y^{i+1}) \\
&\geq \tilde{l}^i(y^i) + (2\mu_k)^{-1}\beta\|y^{i+1} - y^i\|^2 \\
&\geq \tilde{l}^i(y^i),\ \text{where } D_h(y, z) = h(y) - h(z) - \langle \nabla h(z), y - z\rangle.
\end{aligned}$$

From the above relations, we deduce that $\{\tilde{l}^i(y^i)\}$ is nondecreasing and bounded above by $\varphi(x^k)$, so it is convergent. Moreover, we have

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

and then $y^{i+1} \to y^i$.

2. Let $y \in C$ be fixed, then we have

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

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

3. Using $(AC1), (AC3)$, note that $\tilde{s}_y^i \in \partial_{\varepsilon_i} f(y^i), \tilde{s}_y^{i+1} \in \partial_{\varepsilon_{i+1}} f(y^{i+1})$ and $\varepsilon_i \to 0$, we have

$$\begin{aligned}\langle \tilde{s}_y^i, y^{i+1} - y^i\rangle &\leq \theta_a^{i+1}(y^{i+1}) - \tilde{\varphi}_y^i \\ &\leq \tilde{\varphi}_y^{i+1} + \varepsilon_{i+1} - \tilde{\varphi}_y^i \\ &\leq \varepsilon_{i+1} - \tilde{\varphi}_y^i + \tilde{\varphi}_y^i + \varepsilon_i + \langle \tilde{s}_y^{i+1}, y^{i+1} - y^i\rangle \\ &= \varepsilon_{i+1} + \varepsilon_i + \langle \tilde{s}_y^{i+1}, y^{i+1} - y^i\rangle.\end{aligned}$$

Since we assume $\partial_\varepsilon\varphi$ is uniformly bounded with respect to ε on bounded sequence $\{y^i\}_{i\in N}$ and $\varepsilon_i \to 0$, as $\|y^{i+1} - y^i\|^2 \to 0$, hence

$$[\theta_a^{i+1}(y^{i+1}) - \tilde{\varphi}_y^i] \to 0, [\tilde{\varphi}_y^{i+1} - \tilde{\varphi}_y^i] \to 0, [\tilde{\varphi}_y^i - \varphi(y^i)] \to 0.$$

Thus,

$$\tilde{\varphi}_y^{i+1} - \theta_a^{i+1}(y^{i+1}) = \tilde{\varphi}_y^{i+1} - \tilde{\varphi}_y^i + \tilde{\varphi}_y^i - \theta_a^{i+1}(y^{i+1}) \to 0,$$

this establishes that $\tilde{\varphi}_y^i \to \theta_a^i(y^i)$.

Second, we show $y^i \to z(x^k)$. Since for all $y \in C$, for a given $\varepsilon \geq 0$,

$$\begin{aligned}\tilde{\varphi}_y^\varepsilon + \varepsilon &\geq \theta_a^i(y^i) + \langle r^i, y - y^i\rangle \\ &= \theta_a^i(y^i) + \mu_k^{-1}\langle \nabla h(x^k) - \nabla h(y^i), y - y^i\rangle - \langle r(x^k), y - y^i\rangle.\end{aligned} \tag{2.7}$$

where $\varphi(y) \geq \tilde{\varphi}_y^{\varepsilon} \geq \varphi(y) - \varepsilon$. Because $\{y^i\}_{i \in N}$ is bounded in C, we can extract a subsequence that weakly converges in C. Without loss of generality, suppose that $y^i \overset{w}{\to} \bar{y} \in C$. If we take $y = \bar{y}$, $\varepsilon = \bar{\varepsilon} = 0$ in (2.7) then

$$\mu_k[\tilde{\varphi}_{\bar{y}}^0 - \theta_a^i(y^i)] \geq -\mu_k \langle r(x^k), \bar{y} - y^i \rangle + \langle \nabla h(x^k) - \nabla h(y^i), \bar{y} - y^i \rangle + \beta \|\bar{y} - y^i\|^2. \quad (2.8)$$

Note that $y^i \overset{w}{\to} \bar{y}$, $\tilde{\varphi}_y^i \to \theta_a^i(y^i)$ and φ is weakly l.s.c., we directly have $\tilde{\varphi}_{\bar{y}}^0 \leq \overline{\lim}_{i\to\infty} [\tilde{\varphi}_y^i + \varepsilon_i] = \overline{\lim}_{i\to\infty} \tilde{\varphi}_y^i$ and $\overline{\lim}_{i\to\infty}[\tilde{\varphi}_{\bar{y}}^0 - \theta_a^i(y^i)] \leq 0$. Then passing to the superior limit in (2.8), we obtain

$$\overline{\lim}_{i\to\infty} \|y^i - \bar{y}\|^2 = 0, \quad (2.9)$$

and thus $y^i \to \bar{y}$. Now, from (2.7), we have for all $y \in C$, for any given $\varepsilon \geq 0$,

$$\tilde{\varphi}_y^{\varepsilon} + \varepsilon \geq [\theta_a^i(y^i) - \tilde{\varphi}_y^i] + [\tilde{\varphi}_y^i - \tilde{\varphi}_{\bar{y}}^0] + \tilde{\varphi}_{\bar{y}}^0 + \mu_k^{-1} \langle \nabla h(x^k) - \nabla h(y^i), y - y^i \rangle - \langle r(x^k), y - y^i \rangle.$$

By $\tilde{\varphi}_y^i \to \theta_a^i(y^i)$, $y^i \to \bar{y}$, $\varepsilon_i \to 0$, φ and ∇h are continuous on C, if we take the limit in the last inequality, we have that for all $y \in C$,

$$\begin{aligned} \tilde{\varphi}_y^{\varepsilon} + \varepsilon &\geq \tilde{\varphi}_{\bar{y}}^0 + \mu_k^{-1} \langle \nabla h(x^k) - \nabla h(\bar{y}), y - \bar{y} \rangle - \langle r(x^k), y - \bar{y} \rangle \\ &= \varphi(\bar{y}) + \mu_k^{-1} \langle \nabla h(x^k) - \nabla h(\bar{y}), y - \bar{y} \rangle - \langle r(x^k), y - \bar{y} \rangle. \end{aligned}$$

Therefore, $\mu_k^{-1}(\nabla h(x^k) - \nabla h(\bar{y})) - r(x^k) \in \partial(\varphi + \psi_C)(\bar{y})$, $\bar{y} = z(x^k)$, this completes the proof.

Definition 2.1. *If, for given $\varepsilon \geq 0$,*

$$\|z(x^k) - x^k\|^2 \leq \varepsilon,$$

then x^k is called an ε-optimal solution to problem (P), where $z(x^k) = argmin_{x \in C}\{\varphi(x) + \langle r(x^k), x - x^k \rangle + \mu_k^{-1}[h(x) - h(x^k) - \langle \nabla h(x^k), x - x^k \rangle]\}$.

Theorem 2.1. *Consider the approximate bundle algorithm for problem (P). Suppose that Assumption A holds. If some iterate x^k is reached and, from then on, k remains fixed, i.e., only null steps are made, then x^k is an ε-optimal solution to problem (P) in the sense of definition 2.1, where $\varepsilon = \frac{2\mu_k}{\beta}(\varepsilon^k + \frac{\varepsilon^{z(x^k)}}{1-m})$.*

Proof. For all $i > i_k$,

$$\tilde{\varphi}_x^k - \tilde{\varphi}_y^i < m[\tilde{\varphi}_x^k - \theta_a^i(y^i)] + (1-m)\langle r(x^k), y^i - x^k\rangle.$$

We pass to the limit on i in this inequality, and if we note that the fact $\tilde{\varphi}_y^i \to \theta_a^i(y^i), y^i \to z(x^k)$, and φ is continuous on C, we obtain

$$\tilde{\varphi}_x^k - \tilde{\varphi}_{z(x^k)} - \varepsilon^{z(x^k)} \le \tilde{\varphi}_x^k - \varphi(z(x^k)) \le m[\tilde{\varphi}_x^k - \tilde{\varphi}_{z(x^k)}] + (1-m)\langle r(x^k), z(x^k) - x^k\rangle, \quad (2.10)$$

this means that

$$\tilde{\varphi}_x^k \le \tilde{\varphi}_{z(x^k)} + \langle r(x^k), z(x^k) - x^k\rangle + \frac{\varepsilon^{z(x^k)}}{1-m}.$$

On the other hand, by definition of $z(x^k)$, we have for all $y \in C$,

$$\begin{aligned}\tilde{\varphi}_{z(x^k)} &+ \langle r(x^k), z(x^k) - x^k\rangle + \mu_k^{-1}[h(z(x^k)) - h(x^k) - \langle \nabla h(x^k), z(x^k) - x^k\rangle] \\ &\le \varphi(y) + \langle r(x^k), y - x^k\rangle + \mu_k^{-1}[h(y) - h(x^k) - \langle \nabla h(x^k), y - x^k\rangle]\end{aligned}$$

If we take $y = x^k$ in this inequality, we deduce that

$$\begin{aligned}&\tilde{\varphi}_{z(x^k)} + \langle r(x^k), z(x^k) - x^k\rangle \\ &\qquad \le \tilde{\varphi}_x^k + \varepsilon^k + \mu_k^{-1}[h(x^k) - h(z(x^k)) + \langle \nabla h(x^k), z(x^k) - x^k\rangle].\end{aligned}$$

Since h is strongly convex with modulus $\beta > 0$, we have

$$h(x^k) - h(z(x^k)) + \langle \nabla h(x^k), z(x^k) - x^k\rangle \le -\frac{\beta}{2}\|z(x^k) - x^k\|^2.$$

Hence,

$$\tilde{\varphi}_{z(x^k)} + \langle r(x^k), z(x^k) - x^k\rangle \le \tilde{\varphi}_x^k + \varepsilon^k - \frac{\beta}{2\mu_k}\|z(x^k) - x^k\|^2. \quad (2.11)$$

Combining (2.10) and (2.11), we have

$$\tilde{\varphi}_x^k \le \tilde{\varphi}_x^k + \varepsilon^k + \frac{\varepsilon^{z(x^k)}}{1-m} + (-\frac{\beta}{2\mu_k}\|z(x^k) - x^k\|^2),$$

i.e.,

$$\|z(x^k) - x^k\|^2 \le \frac{2\mu_k}{\beta}(\varepsilon^k + \frac{\varepsilon^{z(x^k)}}{1-m}).$$

This completes the proof.

3. Convergence analysis

In this section, we suppose that the approximate bundle algorithm generates an infinite sequence $\{x^k\}_{k\in N}$. The operator F is multivalued and $\{\mu_k\}_{k\in N}$ is chosen to be

$$\begin{cases} \mu_k = \frac{\lambda_k}{\eta_k} \text{ with } \{\lambda_k\}_{k\in N} \text{ a sequence of positive numbers;} \\ \eta_k = \begin{cases} \max\{1, \|r(x^0)\|\}, & \text{if } k = 0, \\ \max\{\eta_{k-1}, \|r(x^k)\|\}, & \text{if } k \geq 1. \end{cases} \end{cases} \tag{3.1}$$

Consider the sequence $\{\Gamma^k(x^*, \cdot)\}_{k\in N}$ of Lyapunov functions defined on C by

$$\Gamma^k(x^*, x) = h(x^*)-h(x)-\langle\nabla h(x), x^*-x\rangle+\lambda_k(m\eta_k)^{-1}[\langle r(x^*), x-x^*\rangle+\varphi(x)-\varphi(x^*)], \tag{3.2}$$

where $x^* \in C$ denotes a solution to problem (P) and $r(x^*) \in F(x^*)$ with $\varepsilon^* = 0$. Since h is strongly convex with modulus $\beta > 0$, we have, for all $x \in C$,

$$\Gamma^k(x^*, x) \geq \frac{\beta}{2}\|x - x^*\|^2. \tag{3.3}$$

Lemma 3.1. *Suppose that Assumption A holds and that $\{\lambda_k\}_{k\in N}$ is a nonincreasing sequence of positive numbers. Then we have, for all $k \in N$,*

$$\begin{aligned} \Gamma^{k+1}(x^*, x^{k+1}) - \Gamma^k(x^*, x^k) \leq & -c\|x^{k+1} - x^k\|^2 + \lambda_k^2 u \\ & +\frac{\lambda_k}{\eta_k}[\langle r(x^k), x^* - x^k\rangle + \varphi(x^*) - \tilde{\varphi}_x^k + \frac{\varepsilon^k}{m}]. \end{aligned} \tag{3.4}$$

Proof. First observe the optimality conditions satisfied by $x^{k+1} \in C$ are

$$\begin{aligned} &\langle\eta_k^{-1} r(x^k) + \lambda_k^{-1}(\nabla h(x^{k+1}) - \nabla h(x^k)), x - x^{k+1}\rangle \\ &\quad +\eta_k^{-1}(\theta_a^{i_{k+1}}(x) - \theta_a^{i_{k+1}}(x^{k+1})) \geq 0, \qquad \forall\ x \in C. \end{aligned} \tag{3.5}$$

Note that $\lambda_{k+1} \leq \lambda_k$, $\eta_{k+1} \geq \eta_k$, we can write

$$\begin{aligned} \Gamma^{k+1}(x^*, x^{k+1}) - \Gamma^k(x^*, x^k) &\leq \Gamma^k(x^*, x^{k+1}) - \Gamma^k(x^*, x^k) \\ &\leq s_1 + s_2 + s_3 \end{aligned} \tag{3.6}$$

with

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

For s_1, we have

$$s_1 \leq -\frac{\beta}{2}\|x^{k+1} - x^k\|^2. \tag{3.7}$$

Using (3.5) with $x = x^*$, we obtain

$$\begin{aligned} s_2 &\leq \tfrac{\lambda_k}{\eta_k}[\langle r(x^k), x^* - x^{k+1}\rangle + \theta_a^{i_{k+1}}(x^*) - \theta_a^{i_{k+1}}(x^{k+1})] \\ &= \tfrac{\lambda_k}{\eta_k}[\langle r(x^k), x^* - x^k\rangle + \varphi(x^*) - \tilde{\varphi}_x^k + \langle r(x^k), x^k - x^{k+1}\rangle \\ &\qquad + \theta_a^{i_{k+1}}(x^*) - \varphi(x^*) + \tilde{\varphi}_x^k - \theta_a^{i_{k+1}}(x^{k+1})]. \end{aligned} \tag{3.8}$$

From the stopping test (2.3), we deduce that

$$\tilde{\varphi}_x^k - \theta_a^{i_{k+1}}(x^{k+1}) \leq \frac{1}{m}[\tilde{\varphi}_x^k - \tilde{\varphi}_x^{k+1}] - \frac{1-m}{m}\langle r(x^k), x^{k+1} - x^k\rangle. \tag{3.9}$$

Combining the fact $\theta_a^{i_{k+1}} \leq \varphi$ with (3.8) and (3.9) we derive that

$$\begin{aligned} s_2 + s_3 &\leq \tfrac{\lambda_k}{\eta_k}[\langle r(x^k), x^* - x^k\rangle + \varphi(x^*) - \tilde{\varphi}_x^k \\ &\qquad + \tfrac{1}{m}\langle r(x^k) - r(x^*), x^k - x^{k+1}\rangle + \varepsilon^k] \\ &\leq \tfrac{\lambda_k}{\eta_k}[\langle r(x^k), x^* - x^k\rangle + \varphi(x^*) - \tilde{\varphi}_x^k] \\ &\qquad + \tfrac{1}{m}[\tfrac{1}{2\tau}(\lambda_k^2/\eta_k^2)\|r(x^k - r(x^*))\|^2 + \tfrac{\tau}{2}\|x^k - x^{k+1}\|^2] + \tfrac{\lambda_k}{m\eta_k}\varepsilon^k, \end{aligned} \tag{3.10}$$

where τ is any positive constant. From the definition of $\{\eta_k\}$, we have

$$\frac{1}{\eta_k^2}\|r(x^k) - r(x^*)\|^2 \leq [1 + \|r(x^*)\|^2]^2. \tag{3.11}$$

Combining (3.6), (3.7), (3.10) and (3.11), we obtain

$$\begin{aligned} \Gamma^{k+1}(x^*, x^{k+1}) - \Gamma^k(x^*, x^k) &\leq -\tfrac{1}{2}(\beta - \tfrac{\tau}{m})\|x^{k+1} - x^k\|^2 + \tfrac{1}{2m\tau}[1 + \|r(x^*)\|^2]^2\lambda_k^2 \\ &\qquad + \tfrac{\lambda_k}{\eta_k}[\langle r(x^k), x^* - x^k\rangle + \varphi(x^*) - \tilde{\varphi}_x^k + \tfrac{\varepsilon^k}{m}]. \end{aligned} \tag{3.12}$$

If we choose τ such that $0 < \tau < \beta m$, then we obtain (3.4) with $c = \frac{1}{2}(\beta - \frac{\tau}{m}) > 0$, $u = \frac{1}{2m\tau}[1 + \|r(x^*)\|^2]^2 > 0$.

Theorem 3.1. *Assume that the conditions of Lemma 3.1 hold. If* $\sum_{k=0}^{+\infty}\lambda_k^2 < +\infty$ *and* $\sum_{k=0}^{+\infty}\lambda_k\varepsilon^k < +\infty$*, then* $\{\Gamma^k(x^*, x^k)\}_{k\in N}$ *is convergent,* $\{x^k\}_{k\in N}$ *is bounded,* $\sum_{k=0}^{+\infty}\|x^{k+1} - x^k\|^2 < +\infty$ *and*

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

Proof. Since x^* is a solution to problem (P), if we take $\varepsilon^* = 0$, we have, for all k,

$$\frac{\lambda_k}{\eta_k}[\langle r(x^k), x^* - x^k\rangle + \varphi(x^*) - \tilde{\varphi}_x^k - \varepsilon^k] \leq 0,$$

we derive from (3.4) that

$$\Gamma^{k+1}(x^*, x^{k+1}) - \Gamma^k(x^*, x^k) \leq \lambda_k^2 u + \frac{\lambda_k \varepsilon^k}{\eta_k}(1 + \frac{1}{m}).$$

Since $\sum_{k=0}^{+\infty} \lambda_k^2$ and $\sum_{k=0}^{+\infty} \lambda_k \varepsilon^k$ are convergent, it follows that $\{\Gamma^k(x^*, x^k))\}_{k \in N}$ is convergent. Using inequality (3.3) we have $\{x^k\}_{k\in N}$ is bounded. Rearranging the terms of (3.4) as

$$\begin{aligned} c\|x^{k+1} - x^k\|^2 + \tfrac{\lambda_k}{\eta_k}[\langle r(x^k), x^k - x^*\rangle + \tilde{\varphi}_x^k + \varepsilon^k - \varphi(x^*)] \\ \leq \Gamma^k(x^*, x^k) - \Gamma^{k+1}(x^*, x^{k+1}) + \lambda_k^2 u + (1 + \tfrac{1}{m})\tfrac{\lambda_k \varepsilon^k}{\eta_k}, \end{aligned}$$

we obtain, using $\eta_k \geq 1$ and the convergence of $\sum_{k=0}^{+\infty} \lambda_k^2$, $\sum_{k=0}^{+\infty} \lambda_k \varepsilon^k$ and $\{\Gamma^k(x^*, x^k))\}_{k\in N}$, that $\sum_{k=0}^{+\infty} \|x^{k+1} - x^k\|^2 < +\infty$ and (3.13) holds.

Definition 3.1. *The function $l : C \to R \cup \{+\infty\}$ is called a gap function with respect to problem (P) if for all $x \in C$, $l(x) \geq 0$ and $l(\bar{x}) = 0$ iff $\bar{x}$ is a solution to problem (P).*

Proposition 3.1. [8] *Let l be a gap function with respect to problem (P). If l is weakly l.s.c. on C and $l(x^k) \to 0$, then any weak limit point of the sequence $\{x^k\}_{k\in N}$ generated by the algorithm is a solution to problem (P).*

Lemma 3.2. [18] *If l is a Lipschitz function on $\{x^k\}_{k\in N}$, and if $\{\lambda_k\}_{k\in N}$ is a sequence of positive numbers such that*

$$(a) \sum \lambda_k = +\infty, (b) \sum \lambda_k l(x^k) < +\infty, (c)\ \exists\ \delta > 0, \textit{such that}\ \forall\ k \in N, \|x^{k+1} - x^k\|^2 \leq \delta\lambda_k,$$

then $l(x^k) \to 0$.

Definition 3.2. *A multivalued operator F is said to be Lipschitz continuous on a subset B of C if $\exists\ \delta > 0$ such that for all $x, y \in B$,*

$$e(F(x), F(y)) \leq L\|x - y\|,$$

where $e(F(x), F(y)) = sup_{r\in F(x)} inf_{s\in F(y)} \|r - s\|$.

Lemma 3.3. [8] *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 $\|r(x)\| \leq \alpha, \forall\, x \in B$ and $r(x) \in F(x)$.*

Definition 3.3. [8] *A multivalued operator F is said to be weakly closed on C if*

$$z^k \overset{w}{\to} \bar{z}, z^k \in C, r^k \overset{w}{\to} \bar{r}, r^k \in F(z^k) \Longrightarrow \bar{r} \in F(\bar{z}).$$

Definition 3.4. [8] *A multivalued operator F is said to be paramonotone on C if F is monotone on C and, for all $x, y \in C$ and $r(x) \in F(x), r(y) \in F(y)$,*

$$\langle r(x) - r(y), x - y\rangle = 0 \Longrightarrow r(y) \in F(x), r(x) \in F(y).$$

Proposition 3.2. [19] *If F is paramonotone and if x^* is a solution to (P), then $\bar{x}$ is a solution to (P) iff*

$$\bar{x} \in C \text{ and } \bar{r} \in F(\bar{x}) \text{ such that } \langle \bar{r}, x^* - \bar{x}\rangle + \varphi(x^*) - \varphi(\bar{x}) \geq 0, \tag{3.14}$$

where we take $\bar{\varepsilon} = \varepsilon^* = 0$, the approximation errors at the solutions $\bar{x}$ and x^*.

Proposition 3.3. [8] *Let x^* denote any solution to (P).*

(a)If F is paramonotone on C, and F 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 C, then l is Lipschitz continuous on bounded subsets of C.

(c)If, in addition, F is weakly closed on C, then l is weakly l.s.c. on C.

Proposition 3.4. [8] *Let x^* denote any solution to (P). If $F = \partial f$, $f = \Gamma_0(H)$ 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 C, then l is also Lipschitz continuous on bounded subsets of C.

Proposition 3.5. [8] *If F is strongly monotone with modulus $\alpha > 0$ on C, then $l(x) = \|x - x^*\|^2$ is a gap function such that, for all $x \in C$ and $r(x) \in F(x)$,*

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

where x^* denotes the unique solution to (P) with $\varepsilon^* = 0$. Moreover, l is strongly convex, weakly l.s.c. on H, and Lipschitz continuous on bounded subsets of C.

We put together the properties requested on the gap function. These properties are satisfied in the three situations described in Propositions 3.3, 3.4, 3.5.

Assumption I:

(i) $\exists\, \alpha > 0, \exists\, l : C \to R \cup \{+\infty\}$ such that for all $x \in C, r(x) \in F(x)$,

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

(ii) for all $x \in C$, $l(x) \geq 0$ and $l(\bar{x}) = 0$ iff $\bar{x}$ is a solution to (P);

(iii) l is weakly l.s.c. on C and Lipschitz continuous on bounded subsets of C.

The purpose of the next proposition is to prove the conditions of (b) and (c) in Lemma 3.2 are satisfied.

Proposition 3.6. *(a) Assume that the conditions of Theorem 3.1 and Assumption I (i) and (II) are satisfied. If F is bounded on bounded subsets of C, then $\sum \lambda_k l(x^k) < +\infty$.*
(b) Assume that Assumption A holds and that the sequence $\{\theta_a^i\}_{k \in N}$ satisfies condition $(AC4)$. Then there exists $\delta > 0$ such that for all $k \geq 1$, $\|x^{k+1} - x^k\| \leq \delta \lambda_k$.

Proof. (a) Since $\{x^k\}_{k \in N}$ is bounded and F is bounded on bounded subsets of C, the sequences $\{r(x^k)\}$ and $\{\eta_k\}$ are bounded. Then, using successively Theorem 3.1, Assumption I (i)(ii) and $\sum_{k=1}^{+\infty} \lambda_k \varepsilon^k < +\infty$, we have

$$\alpha \sum_{k=1}^{+\infty} \lambda_k l(x^k) \leq \sum_{k=1}^{+\infty} \lambda_k [\langle r(x^k), x^k - x^* \rangle + \tilde{\varphi}_x^k - \varphi(x^*)] + \sum_{k=1}^{+\infty} \lambda_k \varepsilon^k < +\infty.$$

(b) From the optimality conditions (3.5) applied to $x = x^k$, we obtain

$$\begin{aligned}&\langle \nabla h(x^{k+1}) - \nabla h(x^k), x^{k+1} - x^k\rangle \\ &\qquad \leq (\lambda_k/\eta_k)[\langle r(x^k), x^k - x^{k+1}\rangle + \theta_a^{i_{k+1}}(x^k) - \theta_a^{i_{k+1}}(x^{k+1})].\end{aligned} \tag{3.15}$$

Since h is strongly convex and $\|r(x^k)\| \leq \eta_k$, we derive from (3.15) that

$$\beta\|x^{k+1} - x^k\|^2 \leq \lambda_k\|x^{k+1} - x^k\| + (\lambda_k/\eta_k)[\theta_a^{i_{k+1}}(x^k) - \theta_a^{i_{k+1}}(x^{k+1})]. \tag{3.16}$$

Now since $\theta_a^{i_{k+1}}(x^k) \leq \tilde{\varphi}_x^k + \varepsilon^k$ and by condition $(AC4)$

$$\varepsilon^k + \tilde{\varphi}_x^k + \langle \tilde{s}_x^k, x^{k+1} - x^k\rangle \leq \theta_a^{i_{k+1}}(x^{k+1}),$$

we have

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

Hence, since $\partial_\varepsilon\varphi$ is uniformly bounded with respect to ε on bounded subsets of C, we have that there exists $\delta_\varphi > 0$ such that, for all k,

$$\theta_a^{i_{k+1}}(x^k) - \theta_a^{i_{k+1}}(x^{k+1}) \leq \delta_\varphi\|x^{k+1} - x^k\|. \tag{3.17}$$

Finally, from (3.16) (3.17) and since $\eta_k \geq 1$, we deduce that

$$\|x^{k+1} - x^k\| \leq \delta\lambda_k, \ \ \forall k, \ \text{with } \delta = \frac{1}{\beta}(1 + \delta_\varphi).$$

Now, let's state the main convergence result.

Theorem 3.2. *Suppose that the following conditions are satisfied:*

- *Assumption A and I hold.*
- *F is bounded on bounded subsets of C.*
- *$\{\lambda_k\}_{k\in N}$ is nonincreasing and $\sum \lambda_k = +\infty, \sum \lambda_k^2 < +\infty, \sum \lambda_k\varepsilon^k < +\infty$.*

Then the sequence $\{x^k\}_{k\in N}$ is bounded, $l(x^k) \to 0$, and any weak limit point of $\{x^k\}_{k\in N}$ is a solution to problem (P). If, in addition, ∇h is weakly continuous on C, then $\{x^k\}_{k\in N}$ weakly converges to a solution to problem (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 to problem (P).

Proof. The first part of the theorem follows from Lemma 3.1, Theorem 3.1, Proposition 3.6, Lemma 3.2 and Proposition 3.1. Suppose now that ∇h is weakly continuous on C and that the sequence $\{x^k\}_{k\in N}$ has two different weak limit points $\bar{x}^1$ and $\bar{x}^2$. If we take the approximation errors at solution points $\bar{\varepsilon}^1 = \bar{\varepsilon}^2 = 0$, the proof can be obtained by imitating Theorem 3.10 in [8].

References

[1] G. Cohen, Auxiliary problem principle extended to variational inequalities, Journal of Optimization Theory and Applications. 59 (1988), 325-333.

[2] T. T. Hue, J. J. Strodiot and V. H. Nguyen, Convergence of the approximate auxiliary problem method for solving generalized variational inequalities, Journal of Optimization Theory and Applications. 121 (2004), 119-145.

[3] B. Lemaire, Coupling optimization methods and variational convergence, in Trends in Mathematical Optimization, K. H. Hoffmann, J. B. Hiriari-Urruty, C. Lemaréchal, and J. Zowe, eds., Int. Ser. Numer. Math., Birkhauser-Verlag, Basel, 1988, 163-179.

[4] S. Makler-Scheimberg, V. H. Nguyen and J. J. Strodiot, Family of perturbation methods for variational inequalities, Journal of Optimization Theory and Applications. 89 (1996), 423-452.

[5] G. Salmon, V. H. Nguyen and J. J. Strodiot, Coupling the auxiliary porblem principle and the epiconvergence theory to solve general variational inequalities, Journal of Optimization Theory and Applications. 104 (2000), 629-657.

[6] G. Salmon, J. J. Strodiot and V. H. Nguyen, A perturbed auxiliary problem method for paramonotone multivalued mappings, in Advances in Convex Analysis and Global Optimization, N. Hadjisavvas and P. Pardalos, eds., Nonconvex Optim. Appl. 54, Kluwer, Dordrecht, 2001, 515-529.

[7] Y. Sonntag, Convergence au sens de Mosco: Theorie et applications a l'approximation des solutions d'inequations, Ph.D. thesis, Universite de Provence, 1982.

[8] G. Salmon, J. J. Strodiot and V. H. Nguyen, A bundle method for solving variational inequalities, SIAM J. OPTIM. 14 (3) (2004), 869-893.

[9] D. P. Bertsekas, Nonlinear programming, 2nd ed., Athena Scientific, Belmont, MA, 1999.

[10] K. C. Kiwiel, A proximal bundle method with approximate subgradient linearization, SIAM J. OPTIM. 16 (2006), 1007-1023.

[11] K. C. Kiwiel, Approximations in proximal bundle methods and decomposition of convex programs, Journal of Optimization Theory and Applications. 84 (1995), 529-548.

[12] M. V. Solodov, On approximation with finite precision in bundle methods for nonsmooth optimization, Journal of Optimization Theory and Applications. 119 (2003), 151-165.

[13] Jie Shen, Zun-Quan Xia and Li-Ping Pang, A proximal bundle method with inexact data for convex nondifferentiable minimization, Nonlinear Analysis A : theory, method and applications, 66 (2007), 2016-2027.

[14] Shen Jie and Pang Li-Ping, A bundle-type auxiliary problem method for generalized variational-like inequality, Computers and Mathematics with Applications, 55 (2008), 2993–2998.

[15] Shen Jie and Pang Li-Ping, An approximate bundle-type auxiliary problem method for generalized variational inequality, Mathematical and Computer Modeling, 48 (2008), 769-775.

[16] R. S. Burachik, C. A. Sagastizabal and B. F. Svaiter, Episilon-enlargement of maximal monotone operators with application to variational inequalities, in Reformulation-Nonsmooth, Piecewise S-mooth, Semismooth and Smoothing Methods, M. Fukushima and L. Qi, eds., Kluwer, Dordrecht, 1997, 25-43.

[17] R. S. Burachik, C. A. Sagastizabal and B. F. Svaiter, Bundle methods for maximal monotone operators, in Ill-posed Variational Problems and Regularization Techniques, R. Tichatschke and M. Thera, eds., Springer, Berlin, 1999, 49-64.

[18] 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, Vol.1, J. B. Cruz, ed., JAI Press, Greenwich, CT, 1984, 203-266.

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