ISSN 1066-369X, Russian Mathematics, 2017, Vol. 61, No. 10, pp. 44–53. © Allerton Press, Inc., 2017.
Original Russian Text © I.V. Konnov, Salahuddin, 2017, published in Izvestiya Vysshikh Uchebnykh Zavedenii. Matematika, 2017, No. 10, pp. 50–61.

Two-Level Iterative Method for Non-Stationary Mixed Variational Inequalities

I. V. Konnov[1*] and Salahuddin[2]**

[1]*Kazan Federal University*
ul. Kremlyovskaya 18, Kazan, 420008 Russia

[2]*Jazan University*
Jazan, K. S. A.

Received June 3, 2016

Abstract—We consider a mixed variational inequality problem involving a set-valued non-monotone mapping and a general convex function, where only approximation sequences are known instead of exact values of the cost mapping and function, and feasible set. We suggest to apply a two-level approach with inexact solutions of each particular problem with a descent method and partial penalization and evaluation of accuracy with the help of a gap function. Its convergence is attained without concordance of penalty, accuracy, and approximation parameters under coercivity type conditions.

DOI: 10.3103/S1066369X17100061

Keywords: *mixed variational inequality, non-stationarity, non-monotone mappings, potential mappings, approximate solutions, penalty method, gap function.*

PRELIMINARIES

Let D be a nonempty set in the real n-dimensional space $\mathbb{R}^n$, $h : D \to \mathbb{R}$ a convex function, and let $G : D \to \Pi(\mathbb{R}^n)$ be a point-to-set mapping. Here $\Pi(A)$ denotes the family of all nonempty subsets of a set A.

Then one can define the *mixed variational inequality problem* (MVI, for short), which is to find an element $x^* \in D$ such that

$$\exists g^* \in G(x^*),\ \langle g^*, y - x^*\rangle + h(y) - h(x^*) \geq 0 \quad \forall y \in D. \tag{1}$$

Suppose also that D is a set of the form

$$D = V \bigcap W, \tag{2}$$

V and W are convex and closed sets in the space $\mathbb{R}^n$. This partition of the feasible set is optional and usually means that V represents "simple" constraints whereas W corresponds to complex or "functional" ones and a suitable penalty function should be used for this set.

Problem (1) was first proposed in [1], [2] (with the single-valued mapping G) and further investigated by many authors (see, e.g., [3–7]). MVIs give a suitable format for various problems arising in Economics, Mathematical Physics, and Operations Research. Besides, the usual variational inequalities and convex optimization problems can be viewed as particular cases of MVI (1).

We observe that most existing solution methods for these problems require exact values of the cost mapping G, function h, and feasible set D. However, this is often impossible due to the calculation errors and lack of the necessary information. The same situation arises if we find it useful to replace the initial problem by a sequence of auxiliary ones with better properties, as in regularization and penalty methods. Within this approach, we can also replace general nonlinear functions with their simple (for example,

*E-mail: konn-igor@yandex.ru.
**E-mail: salahuddin12@mailcity.com.

piecewise-linear) approximations, and the set-valued mapping G with a sequence of its single-valued approximations, etc. In other words, we have to develop methods for *non-stationary* problems, where only sequences of approximations are known instead of the exact values.

There exist a number of methods for non-stationary optimization and variational inequality problems, but they are based essentially upon monotonicity assumptions and restrictive concordance rules for accuracy, approximation, penalty, and iteration parameters, which creates serious difficulties for their implementation (see, e.g., [8–12] and the references therein).

In [13–15] several penalty based methods for non-stationary optimization and variational inequality problems were suggested. They do not require special concordance of parameters and their convergence was established under coercivity conditions without any monotonicity assumptions. However, the question about existence of an implementable iterative method within this approach, which is applied to approximate problems within some evaluated accuracy and can find a solution of the initial limit problem was still open, because exact solutions of approximate problems utilized in the above papers can not be found explicitly. In this paper, we give the positive answer to this question for some problems of form (1)–(2) and present an implementable method, which involves an inexact solution of approximate problems and does not require special concordance of the parameters or monotonicity assumptions.

Namely, we intend to describe an iterative method for the case when we have only some sequences, i.e.,

i) sets $\{V_l\}$ approximate the set V,

ii) auxiliary penalty functions $P_l : V_l \to \mathbb{R}$ approximate some penalty function $P : V \to \mathbb{R}$ for the set W,

iii) single-valued gradient mappings $\{G_l\}$ approximate the mapping G,

iv) functions $h_l : V_l \to \mathbb{R}$ approximate the function h.

We consider MVI (1)–(2) as an unknown limit problem. The above approximation properties will be specialized. In addition, we only notice that the approximation condition iii) implies certain potentiality properties of the mapping G, but no monotonicity will be assumed.

We suggest to find for each l an inexact solution of the auxiliary penalized MVI: find $z^l \in V_l$ such that

$$\langle G_l(z^l), v - z^l\rangle + h_l(v) - h_l(z^l) + \tau_l[P_l(v) - P_l(z^l)] \geq 0 \ \ \forall y \in V_l, \tag{3}$$

where $\tau_l > 0$, with a descent method in a finite number of inner iterations. Clearly, any descent method for the above MVI will require either monotonicity or potentiality of the mapping G_l together with the convexity of the function h_l for convergence (see, e.g., [5, 6]). However, we do not require the monotonicity of G_l and of G in this work. We impose the potentiality condition on G_l, i.e., our MVIs (3), hence (1), represent necessary optimality conditions for non-convex optimization problems, but the joint monotonicity/convexity does not hold here. For this reason, the solution set of MVI (3) need not to be convex, but we suggest to utilize a gap function and show that it enables one to evaluate a desired accuracy even in the non-monotone case that yields the general convergence. In such a way we create a two-level convergent iterative method for the initial limit problem.

1. DESCENT SPLITTING METHOD FOR THE AUXILIARY PROBLEM

Recall some definitions. Let X be a nonempty subset of a finite dimensional space E. A function $f : X \to \mathbb{R}$ is said to be

a) convex on a convex set $K \subseteq X$ if for each pair of points $x, y \in K$ and for all $\alpha \in [0, 1]$, it holds that

$$f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha)f(y);$$

b) strongly convex with constant $\varkappa > 0$ on a convex set $K \subseteq X$, if for each pair of points $x, y \in K$ and for all $\alpha \in [0, 1]$ it holds that

$$f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha)f(y) - 0.5\varkappa\alpha(1-\alpha)\|x-y\|^2;$$

c) upper (lower) semicontinuous on a set $K \subseteq X$ if for each sequence $\{x^l\} \to \overline{x}$, $x^l \in K$, we have

$$\limsup_{l\to\infty} f(x^l) \leq f(\overline{x}) \ \ (\liminf_{l\to\infty} f(x^l) \geq f(\overline{x})).$$

d) coercive if $f(x) \to +\infty$ as $\|x\| \to \infty$.

We see that b)$\Longrightarrow$a) and b)$\Longrightarrow$d), but converse is not true in general. A sequence of sets $\{X_k\}$ is called Mosco convergent to a set X if and only if

i) for each sequence $\{x^k\} \to \overline{x}$, $x^k \in X_k$, we have $\overline{x} \in X$;

ii) for each point $\overline{x} \in X$ there exists a sequence $\{x^k\} \to \overline{x}$ with $x^k \in X_k$.

Let us first consider the problem of finding the minimal value of some goal function $\mu : \mathbb{R}^n \to \mathbb{R}$ on a feasible set $X \subseteq \mathbb{R}^n$. For brevity, we write this problem as

$$\min_{x\in X} \to \mu(x), \tag{4}$$

its solution set is denoted by X^* and the optimal value of the function by μ^*. Next, we further suppose that the set $X \subseteq \mathbb{R}^n$ is non-empty, convex, and closed,

$$\mu(x) = \mu_1(x) + \mu_2(x) \tag{5}$$

where $\mu_1 : \mathbb{R}^n \to \mathbb{R}$ is a smooth, but not necessary convex function, and $\mu_2 : \mathbb{R}^n \to \mathbb{R}$ is not necessary smooth, but rather simple and convex function. We obtain a non-convex optimization problem, whose solution with respect to μ_2 is not supposed to be difficult. For this reason, we take the following MVI: Find a point $x^* \in X$ such that

$$\langle \mu_1'(x^*), y - x^* \rangle + [\mu_2(y) - \mu_2(x^*)] \geq 0 \quad \forall y \in X; \tag{6}$$

its solution set is denoted by X^0 (cf. (3)).

We recall the known relations between problems (4)–(5) and (6) (see [4], proposition 2.2.2 and also [5], p. 7) with their full proof.

Proposition 1. *Each solution to problem* (4)–(5) *is a solution to MVI* (6). *The reverse assertion is true if μ_1 is convex.*

Proof. If x^* solves MVI (6) and μ_1 is convex, then, by convexity, we have

$$\mu_1(y) - \mu_1(x^*) + \mu_2(y) - \mu_2(x^*) \geq \langle \mu_1'(x^*), y - x^* \rangle + \mu_2(y) - \mu_2(x^*) \geq 0$$

for every $y \in X$, i.e., $x^* \in X^*$. Conversely, let x^* solve problem (4)–(5). If $x^* \notin X^0$, there is a point $x' \in X$ such that

$$\langle \mu_1'(x^*), x' - x^* \rangle + \mu_2(x') - \mu_2(x^*) = \delta < 0.$$

Take $\lambda > 0$ and set $x(\lambda) = \lambda x' + (1-\lambda)x^*$. Then $x(\lambda) \in X$ if $\lambda \in (0, 1)$. At the same time, we have

$$\begin{aligned}&\langle \mu_1'(x^*), x(\lambda) - x^* \rangle + \mu_2(x(\lambda)) - \mu_2(x^*)\\ &\qquad \leq \lambda\langle \mu_1'(x^*), x' - x^* \rangle + \lambda\mu_2(x') + (1-\lambda)\mu_2(x^*) - \mu_2(x^*)\\ &\qquad\qquad = \lambda\left\{\langle \mu_1'(x^*), x' - x^* \rangle + \mu_2(x') - \mu_2(x^*)\right\} = \lambda\delta < 0.\end{aligned}$$

It follows that

$$\mu_1(x(\lambda)) - \mu_1(x^*) + \mu_2(x(\lambda)) - \mu_2(x^*)$$
$$= \langle \mu_1'(x^*), x(\lambda) - x^* \rangle + o(\lambda) + \mu_2(x(\lambda)) - \mu_2(x^*) \le \lambda\delta + o(\lambda) < 0$$

for some $\lambda \in (0,1)$ small enough, a contradiction. □

Due to the above properties of the functions μ_1 and μ_2, we can solve MVI (6) with a splitting type descent method (see [16] and also [5, 17]).

Fix $\alpha > 0$. For each point $x \in X$ we can define $y(x) \in X$ such that

$$\langle \mu_1'(x) + \alpha^{-1}(y(x) - x), y - y(x) \rangle + [\mu_2(y) - \mu_2(y(x))] \ge 0 \quad \forall y \in X. \tag{7}$$

Clearly, this MVI determines the forward-backward splitting iteration with fixed stepsize. Besides, MVI (7) gives a necessary and sufficient optimality condition for the optimization problem:

$$\min_{y \in X} \to \Phi(x,y), \text{ where } \Phi(x,y) = \langle \mu_1'(x), y \rangle + 0.5\alpha^{-1}\|x - y\|^2 + \mu_2(y). \tag{8}$$

Under the above assumptions $\Phi(x,\cdot)$ is strongly convex, hence problem (8)(or (7)) has the unique solution $y(x)$, thus defining the single-valued mapping $x \mapsto y(x)$, which will be used for calculation of the descent direction. We recall its basic properties (see, e.g,. [17], [5], Chap. 2, and [7], lemma 9.5).

Lemma 1. a) $x = y(x) \Longleftrightarrow x \in X^0$,

b) *The mapping* $x \mapsto y(x)$ *is continuous on* X,

c) *For any point* $x \in X$ *it holds that*

$$\mu'(x; y(x) - x) \le -\alpha^{-1}\|y(x) - x\|^2.$$

Here and below, $\mu'(x;d)$ denotes the directional derivative of μ at a point x in d.

The descent splitting algorithm, which is based on these properties and utilizes an inexact Armijo type linesearch procedure (DSI for short), is described as follows.

Algorithm (DSI). Choose a point $x^0 \in X$ and numbers $\alpha > 0$, $\beta \in (0,1)$, $\gamma \in (0,1)$.

At the kth iteration, $k = 0, 1, \ldots$, we have a point $x^k \in X$, compute $y(x^k)$ and set $d^k = y(x^k) - x^k$. If $d^k = 0$, stop. Otherwise, we find m as the smallest non-negative integer such that

$$\mu(x^k + \gamma^m d^k) \le \mu(x^k) - \beta\gamma^m\|d^k\|^2,$$

set $\lambda_k = \gamma^m$, $x^{k+1} = x^k + \lambda_k d^k$ and go to the next iteration.

We give the known convergence properties of this algorithm (see, e.g., [17], [5], theorem 5.11, and [7], theorem 9.16).

Proposition 2. *Let a sequence* $\{x^k\}$ *be generated by Algorithm* (DSI). *Then*

i) *if the algorithm terminates,* $x^k \in X^0$*,*

ii) *the linesearch procedure is always finite,*

iii) *if the set*

$$X(x^0) = \{y \in X \mid \mu(y) \le \mu(x^0)\}$$

is bounded, then

$$\lim_{k\to\infty} \mu(x^k) = \widetilde{\mu} \text{ and } \lim_{k\to\infty} \|y(x^k) - x^k\| = 0; \tag{9}$$

$\{x^k\}$ *has limit points and all these limit points are solutions to MVI* (6)*. If, additionally,* μ_1 *is convex, then* $\widetilde{\mu} = \mu^*$*, and all the limit points of* $\{x^k\}$ *are solutions of problem (4)–(5).*

We are interested in finding an inexact solution of (4)–(5) in a finite number of iterations. In order to estimate a desired accuracy in the non-convex/non-monotone MVI we apply the gap function approach with respect to MVI (6) (see, e.g., [5, 7] and references therein).

Let us consider the function

$$\varphi_\alpha(x) = \max_{y\in X}\{\langle \mu_1'(x), x-y\rangle + \mu_2(x) - \mu_2(y) - 0.5\alpha^{-1}\|x-y\|^2\}$$
$$= \langle \mu_1'(x), x-y(x)\rangle + \mu_2(x) - \mu_2(y(x)) - 0.5\alpha^{-1}\|x-y(x)\|^2. \quad (10)$$

Calculation of $\varphi_\alpha(x)$ is equivalent to the usual forward-backward splitting iteration. We recall its basic properties (see, e.g., [5], Chap. 2).

Lemma 2. a) $\varphi_\alpha(x) \geq 0$ *for each* $x \in X$,

b) $x = y(x) \iff \varphi_\alpha(x) = 0 \iff x \in X^0$,

c) *the function* φ_α *is lower semi-continuous on* X.

In addition to Lemmas 1 and 2, we give the basic estimates.

Proposition 3. *The following inequalities hold true*:

i)

$$\varphi_\alpha(x) \geq (2\alpha)^{-1}\|x-y(x)\|^2 \quad \forall x \in X; \quad (11)$$

ii)

$$\langle \mu_1'(x) + \alpha^{-1}(y(x)-x), y-x\rangle + [\mu_2(y) - \mu_2(x)] \geq -\varphi_\alpha(x) + (2\alpha)^{-1}\|x-y(x)\|^2 \quad \forall y \in X. \quad (12)$$

Proof. By definition, from (7) with $y = x$ we have

$$\langle \mu_1'(x) + \alpha^{-1}(y(x)-x), x-y(x)\rangle + [\mu_2(x) - \mu_2(y(x))] \geq 0,$$

which yields (11). It also follows from (7) that

$$0 \leq \langle \mu_1'(x) + \alpha^{-1}(y(x)-x), y-x\rangle + [\mu_2(y) - \mu_2(x)]$$
$$+ \langle \mu_1'(x), x-y(x)\rangle + [\mu_2(x) - \mu_2(y(x))] - \alpha^{-1}\|x-y(x)\|^2$$
$$= \langle \mu_1'(x) + \alpha^{-1}(y(x)-x), y-x\rangle + [\mu_2(y) - \mu_2(x)] + \varphi_\alpha(x) - (2\alpha)^{-1}\|x-y(x)\|^2,$$

i.e., (12) holds true. □

Combining this property with Proposition 2, we obtain the inexact approximation estimate for Algorithm (DSI).

Theorem 1. *Suppose that the set* $X(x^0)$ *is bounded, a sequence* $\{x^k\}$ *is generated by Algorithm* (DSI). *Then,*

$$\lim_{k\to\infty} \varphi_\alpha(x^k) = 0, \quad (*)$$

for any number $\varepsilon > 0$, *there exists a number* $k = k(\varepsilon)$ *such that* $\varphi_\alpha(x^k) \leq \varepsilon$ *and*

$$\langle \mu_1'(x^k) + \alpha^{-1}(y(x^k)-x^k), y-x^k\rangle + [\mu_2(y) - \mu_2(x^k)] \geq -\varepsilon \quad \forall y \in X. \quad (13)$$

Proof. From Proposition 2 we now have that the sequence $\{x^k\}$ is bounded and all its limit points are solutions to MVI (6). Next, due to (10) we have

$$\varphi_\alpha(x) = \langle \mu_1'(x), x-y(x)\rangle + \mu_2(x) - \mu_2(y(x)) - 0.5\alpha^{-1}\|x-y(x)\|^2$$
$$\leq \langle \mu_1'(x) + g(x), x-y(x)\rangle - 0.5\alpha^{-1}\|x-y(x)\|^2,$$

where $g(x)$ is a subgradient of μ_2 at x. Now from (9) it follows $(*)$. Therefore, the number $k = k(\varepsilon)$ exists and (13) follows from (12). □

2. TWO-LEVEL METHOD AND ITS CONVERGENCE

We now intend to describe a general iterative method for MVI (1)–(2). We need several additional assumptions for its substantiation.

First we introduce the approximation assumptions.

(A1) There exists a sequence of nonempty convex closed sets $\{V_l\}$ which is Mosco convergent to the set V;

(A2) There exists a sequence of continuous mappings $G_l : V_l \to \mathbb{R}^n$, which are the gradients of functions $f_l : V_l \to \mathbb{R}$, $l = 1, 2, \dots$, such that the relations $\{y^l\} \to \overline{y}$ and $y^l \in V_l$ imply $\{G_l(y^l)\} \to \overline{g} \in G(\overline{y})$;

(A3) There exists a sequence of convex subdifferentiable functions $h_l : V_l \to \mathbb{R}$ such that the relations $\{y^l\} \to \overline{y}$ and $y^l \in V_l$ imply $\{h_l(y^l)\} \to h(\overline{y})$.

Moreover, the fixed set W is also approximated with a sequence of sets $\{W_l\}$, which will be defined implicitly. So, instead of MVI (1)–(2) we have in fact a sequence of MVIs: Find a point $\overline{z}^l \in D_l = V_l \bigcap W_l$ such that

$$\langle G_l(\overline{z}^l), y - \overline{z}^l \rangle + h_l(y) - h_l(\overline{z}^l) \geq 0 \quad \forall y \in D_l. \tag{14}$$

The perturbed set D_l can be empty for some l, although the limit feasible set D is usually supposed to be non-empty. Moreover, it seems more suitable to deal with an MVI having only "simple" constraints. For these reasons, we apply the penalty approach, i.e., the set W will be approximated via a sequence of auxiliary penalty functions. Let $P : \mathbb{R}^n \to \mathbb{R}$ be a general penalty function for W, i.e.,

$$P(w) \begin{cases} = 0, & \text{if } w \in W, \\ > 0, & \text{if } w \notin W. \end{cases}$$

We utilize its approximation sequence.

(B1) There exists a sequence of convex, subdifferentiable, and non-negative functions $P_l : V_l \to \mathbb{R}$;

(B2) if $v^l \in V_l$, $\{v^l\} \to \overline{w}$, and $\liminf\limits_{l\to\infty} P_l(v^l) = 0$, then $P(\overline{w}) = 0$;

(B3) for each point $\overline{w} \in D$ there exist a sequence $\{v^l\} \to \overline{w}$ with $v^l \in V_l$ and a number m such that $P_l(v^l) = 0$ if $l \geq m$.

Clearly, conditions **(B2)** and **(B3)** give a kind of the Mosco convergence of the functions $\{P_l\}$ to P. However, P_l is treated as a penalty function for some set W_l, then **(B2)** and **(B3)** give a kind of the Mosco convergence of the sequence $\{W_l\}$ to W.

For each $l = 1, 2, \dots$, instead of (14) we now consider the problem of finding a point $z^l \in V_l$ such that

$$\langle G_l(z^l), v - z^l \rangle + h_l(v) - h_l(z^l) + \tau_l[P_l(v) - P_l(z^l)] \geq 0 \quad \forall v \in V_l, \tag{15}$$

where $\tau_l > 0$ is a penalty parameter (cf. (3) and (6)).

Remark 1. Condition **(A2)** means that the limit set-valued mapping G at any point is approximated by a sequence of gradients $\{G_l\}$. It follows that G also possesses some properties of a generalized gradient set. For instance, if G is the Clarke subdifferential of a locally Lipschitz function f, it can be always approximated by a sequence of gradients within condition **(A2)** [18, 19]. This approach seems suitable from the computational point of view. Moreover, in the case where a sequence of non-smooth functions $\{\widetilde{f}_l\}$, which converges to the function f, is known, we can simply replace each $\widetilde{f}_l$ with its smooth approximation f_l.

The point z^l is an exact solution of the penalized MVI (15). We describe the two-level implementable combined penalty and descent method (PD for short), which utilizes approximate solutions. Let us define the gap function

$$\varphi_\alpha^l(z) = \max_{v \in V_l}\{\langle G_l(z), z - v\rangle + h_l(z) - h_l(v) + \tau_l[P_l(z) - P_l(v)] - 0.5\alpha^{-1}\|z - v\|^2\}$$
$$= \langle G_l(z), z - y^l(z)\rangle + h_l(z) - h_l(y^l(z)) + \tau_l[P_l(z) - P_l(y^l(z))] - 0.5\alpha^{-1}\|z - y^l(z)\|^2$$

(cf. (10)). Due to Theorem 1 and the absence of any concordance rules, we can take its values for error evaluations. Denote by $\pi_X(x)$ the projection of a point x onto a set X.

Method (PD). Choose a point $z^0 = \tilde{z}^0 \in V_0$ and positive sequences $\{\varepsilon_l\}$, $\{\tau_l\}$. Fix $\alpha > 0$.

At the lth stage, $l = 1, 2, \ldots$, we have a point $z^{l-1} \in V_{l-1}$ and a number ε_l. Set $\tilde{z}^{l-1} = \pi_{V_l}(z^{l-1})$ and apply Algorithm (DSI) with the starting point $x^0 = \tilde{z}^{l-1}$,

$$\mu_1(x) = f_l(x), \;\; \mu_2(x) = h_l(x) + \tau_l P_l(x),$$

obtain a point $\tilde{x} = x^k$ such that $\varphi_\alpha^l(\tilde{x}) \le \varepsilon_l$, and set $z^l = \tilde{x}$.

Remark 2. We can clearly remove the projection of the point z^{l-1} onto V_l above in case of the inner approximation of V, i.e., when $V_l \subseteq V_{l+1}$. This condition then should be substituted in **(A1)**.

Since the feasible set may be unbounded, we introduce certain coercivity conditions.

(C1) For each fixed $l = 1, 2, \ldots$, the function $f_l(x) + h_l(x)$ is coercive on the set V_l, i.e., $\{f_l(w^k) + h_l(w^k)\} \to +\infty$ if $\{w^k\} \subset V_l$, $\|w^k\| \to \infty$ as $k \to \infty$.

(C2) There exist a number $\sigma > 0$ and a point $\overline{v} \in D$ such that for any sequences $\{u^l\}$, $\{v^l\}$, and $\{d^l\}$, satisfying the conditions:

$$u^l \in V_l, v^l \in V_l, \;\; \{v^l\} \to \overline{v}, \{\|u^l\|\} \to +\infty, \;\; \{d^l\} \to \mathbf{0};$$

it holds that

$$\liminf_{l \to \infty} \left\{\langle G_l(u^l) + d^l, v^l - u^l\rangle + [h_l(v^l) - h_l(u^l)]\right\} \le -\sigma. \tag{16}$$

Clearly, **(C1)** presents a rather mild coercivity condition for each function $f_l(x) + h_l(x)$ and is destined for providing existence of solutions of each particular problem (14) or (15). Obviously, **(C1)** holds if V_l is bounded. At the same time, **(C2)** gives a similar coercivity condition for the whole sequence of these problems approximating the limit MVI (1)–(2). It also holds if the sequence $\{V_l\}$ is uniformly bounded. Let us turn to the unbounded case. Then we can consider the following coercivity condition for the limit problem:

(C2′). There exist a point $\overline{v} \in D$ such that for any sequence $\{u^l\}$ with $\{\|u^l\|\} \to +\infty$ it holds that

$$\|\overline{v} - u^l\|^{-1}\left\{\langle G(u^l), \overline{v} - u^l\rangle + [h(\overline{v}) - h(u^l)]\right\} \to -\infty, \;\text{ as }\; l \to \infty.$$

This condition is rather usual for providing both existence of solutions and convergence of penalty methods solving problem (1) (see, e.g., [20, 21]). However, it implies that

$$\|\overline{v} - u^l\|^{-1}\left\{\langle G(u^l) + d^l, \overline{v} - u^l\rangle + [h(\overline{v}) - h(u^l)]\right\} \to -\infty \;\text{ as }\; l \to \infty$$

for each sequence $\{d^l\}$ such that $\{\|d^l\|\} \to 0$ as $l \to \infty$. Obviously, the condition

$$\liminf_{l \to \infty} \left\{\|\overline{v} - u^l\|^{-1}\{\langle G(u^l) + d^l, \overline{v} - u^l\rangle + [h(\overline{v}) - h(u^l)]\}\right\} \le -\sigma \tag{17}$$

for some $\sigma > 0$ is weaker essentially. Replacing G and h in (17) with their approximations G_l and h_l, respectively, we obtain (16). We therefore conclude that conditions **(C1)** and **(C2)** are not restrictive.

We now establish the main convergence result.

Theorem 2. *Suppose that assumptions* **(A1)–(A3)**, **(B1)– (B3)**, *and* **(C1)–(C2)** *are fulfilled, the parameters* $\{\varepsilon_l\}$ *and* $\{\tau_l\}$ *satisfy*

$$\{\varepsilon_l\} \searrow 0, \ \ \{\tau_l\} \nearrow +\infty.$$

Then

i) *problem (15) has a solution for any* $\tau_l > 0$,

ii) *the number of iterations at each stage of Method* (PD) *is finite*,

iii) *the sequence* $\{z^l\}$ *generated by Method* (PD) *has limit points and all these limit points are solutions of MVI* (1)–(2).

Proof. We first observe that **(C1)** implies that each problem MVI (15) has a solution since the cost function

$$\mu(x) = f_l(x) + h_l(x) + \tau_l P_l(x)$$

becomes coercive, hence the set

$$V_l(x^0) = \{y \in V_l \mid \mu(y) \le \mu(x^0)\}$$

is bounded. It follows that the optimization problem

$$\min_{x\in V_l} \to \mu(x)$$

has a solution and so is MVI (15) due to Proposition 1. Hence, assertion i) is true. From Theorem 1 we now have that assertion ii) is also true.

By ii), the sequence $\{z^l\}$ is well-defined and (13) implies

$$\langle G_l(z^l) + \alpha^{-1}(y^l(z^l) - z^l), y - z^l\rangle + [h_l(y) - h_l(z^l)] + \tau_l[P_l(y) - P_l(z^l)] \ge -\varepsilon_l \quad \forall y \in V_l. \tag{18}$$

Besides, (11) gives

$$\varepsilon_l \ge \varphi^l_\alpha(z^l) \ge (2\alpha)^{-1}\|z^l - y^l(z^l)\|^2. \tag{19}$$

We now proceed to show that $\{z^l\}$ is bounded. Conversely, suppose that $\{\|z^l\|\} \to +\infty$. By definition, $z^l \in V_l$, besides, by **(B3)** and **(C2)** there exists a sequence $\{v^l\} \to \overline{v}$ such that $v^l \in V_l$ and $P_l(v^l) = 0$ for l large enough. Applying (18), we have

$$\begin{aligned} 0 &\le \langle g^l + d^l, v^l - z^l\rangle + [h_l(v^l) - h_l(z^l)] + \tau_l[P_l(v^l) - P_l(z^l)] + \varepsilon_l \\ &= \langle g^l + d^l, v^l - z^l\rangle + [h_l(v^l) - h_l(z^l)] - \tau_l P_l(z^l) + \varepsilon_l \\ &\le \langle g^l + d^l, v^l - z^l\rangle + [h_l(v^l) - h_l(z^l)] + \varepsilon_l. \end{aligned}$$

Here and below, for brevity we set $g^l = G_l(z^l)$ and $d^l = \alpha^{-1}(y(z^l) - z^l)$. Take a subsequence $\{l_s\}$ such that

$$\lim_{s\to\infty}\{\langle g^{l_s} + d^{l_s}, v^{l_s} - z^{l_s}\rangle + [h_{l_s}(v^{l_s}) - h_{l_s}(z^{l_s})]\} = \liminf_{l\to\infty}\{\langle g^l + d^l, v^l - z^l\rangle + [h_l(v^l) - h_l(z^l)]\},$$

then, by **(C2)**, we have

$$0 \le \lim_{s\to\infty}\{\langle g^{l_s} + d^{l_s}, v^{l_s} - z^{l_s}\rangle + [h_{l_s}(v^{l_s}) - h_{l_s}(z^{l_s})]\} \le -\sigma < 0,$$

a contradiction. Therefore, the sequence $\{z^l\}$ is bounded and has limit points. Let $\overline{z}$ be an arbitrary limit point for $\{z^l\}$, i.e., $\overline{z} = \lim\limits_{s\to\infty} z^{l_s}$. Since $z^l \in V_l$, we have $\overline{z} \in V$ due to **(A1)**. From (18) it follows that

$$0 \le P_{l_s}(z^{l_s}) \le \tau^{-1}_{l_s}\{\langle g^{l_s} + d^{l_s}, v - z^{l_s}\rangle + [h_{l_s}(v) - h_{l_s}(z^{l_s})]\} + P_{l_s}(v) + \tau^{-1}_{l_s}\varepsilon_{l_s} \quad \forall v \in V_{l_s}.$$

Note that the sequence $\{g^{l_s}\}$ is bounded due to **(A2)**.

For any $w \in D$ there exists a sequence $\{v^l\} \to w$ with $v^l \in V_l$ and $P_l(v^l) = 0$ for l large enough due to **(B3)**. Taking $v = v^{l_s}$ above, we obtain

$$0 \leq \liminf_{s\to\infty} P_{l_s}(z^{l_s}) \leq \limsup_{s\to\infty} P_{l_s}(z^{l_s}) \leq \limsup_{s\to\infty} \tau_{l_s}^{-1}\big\{\langle g^{l_s} + d^{l_s}, v^{l_s} - z^{l_s}\rangle + [h_{l_s}(v^{l_s}) - h_{l_s}(z^{l_s})]\big\} = 0$$

on account of **(A2)**, **(A3)**, and (19), i.e., $\lim\limits_{s\to\infty} P_{l_s}(z^{l_s}) = 0$. Due to **(B2)**, this gives $\overline{z} \in W$, i.e., $\overline{z} \in D$.

Now, by **(B3)**, there exists a sequence $\{v^l\} \to \overline{z}$ with $v^l \in V_l$ and $P_l(v^l) = 0$ for l large enough. Again from (18) and **(A3)** we have

$$0 \leq \tau_{l_s} P_{l_s}(z^{l_s}) \leq \langle g^{l_s} + d^{l_s}, v^{l_s} - z^{l_s}\rangle + [h_{l_s}(v^{l_s}) - h_{l_s}(z^{l_s})] + \varepsilon_{l_s} \to 0$$

as $s \to \infty$, hence $\lim\limits_{s\to\infty} \big[\tau_{l_s} P_{l_s}(z^{l_s})\big] = 0$. Take an arbitrary point $w \in D$, then, again by **(B3)**, there exists a sequence $\{v^l\} \to w$ with $v^l \in V_l$ and $P_l(v^l) = 0$ for l large enough. Using again (18), we have

$$\langle g^{l_s} + d^{l_s}, v^{l_s} - z^{l_s}\rangle + [h_{l_s}(v^{l_s}) - h_{l_s}(z^{l_s})] - \tau_{l_s} P_{l_s}(z^{l_s}) + \varepsilon_{l_s} = \langle g^{l_s} + d^{l_s}, v^{l_s} - z^{l_s}\rangle + [h_{l_s}(v^{l_s}) - h_{l_s}(z^{l_s})] + \tau_{l_s}[P_{l_s}(v^{l_s}) - P_{l_s}(z^{l_s})] + \varepsilon_{l_s} \geq 0.$$

It now follows from **(A2)** that $\lim_{s\to\infty} g^{l_s} = \overline{g} \in G(\overline{z})$ and

$$\langle \overline{g}, w - \overline{z}\rangle + [h(w) - h(\overline{z})] = \lim_{s\to\infty} \big\{\langle g^{l_s}, v^{l_s} - z^{l_s}\rangle + [h_{l_s}(v^{l_s}) - h_{l_s}(z^{l_s})]\big\} = \lim_{s\to\infty} \big\{\langle g^{l_s} + d^{l_s}, v^{l_s} - z^{l_s}\rangle + [h_{l_s}(v^{l_s}) - h_{l_s}(z^{l_s})]\big\} \geq \lim_{s\to\infty} \Big[\tau_{l_s} P_{l_s}(z^{l_s})\Big] = 0,$$

therefore $\overline{z}$ solves MVI (1)–(2) and assertion iii) holds true. ☐

We observe that the above proof implies that MVI (1)–(2) has a solution.

3. CONCLUSIONS

We considered a mixed variational inequality problem involving a set-valued non-monotone potential mapping and a convex function, where only approximation sequences are known instead of exact values of the cost mapping and function, and feasible set. In particular, the cost mapping is approximated by a sequence of gradient mappings. We proposed to apply a two-level approach with inexact solutions of each particular problem with a descent splitting method and partial penalization. Its convergence is attained without concordance of penalty, accuracy, and approximation parameters under coercivity type conditions. We suggested to utilize a gap function for evaluation of accuracy approximation of particular penalized problems.

ACKNOWLEDGMENTS

The authors are grateful to an anonymous referee for his/her valuable comments.

In this work, the first author was supported by the RFBR grant, project No. 16-01-00109a and by grant No. 297689 from Academy of Finland and within the frames of State assignment of Ministry of Education and Science of Russian Federation, project No. 1.460.2016/1.4.

REFERENCES

1. Lescarret, C. "Cas d'Addition des Applications Monotones Maximales dan un Espace de Hilbert", in *Compt. Rend. Acad. Sci.* (Paris, 1965), **261**, pp. 1160–1163.
2. Browder, F. E. "On the Unification of the Calculus of Variations and the Theory of Monotone Nonlinear Operators in Banach Spaces", Proc. Nat. Acad. Sci. USA **56** (2), 419–425 (1966).
3. Duvaut, D. and Lions, J.-L. *Les Inéquations en Mechanique et Physique*, Dunod, Paris, 1972; Nauka, Moscow, 1980).

4. Panagiotopoulos, P. D. *Inequality Problems in Mechanics and Their Applications*, Birkhauser, Boston, 1985; Mir, Moscow, 1989).
5. Patriksson, M. *Nonlinear Programming and Variational Inequality Problems*: *A Unified Approach* (Kluwer Academic Publishers, Dordrecht, 1999).
6. Konnov, I. V. *Combined Relaxation Methods for Variational Inequalities* (Springer, Berlin, 2001).
7. Konnov, I. V. *Nonlinear Optimization and Variational Inequalities* (Kazan Univ. Press, Kazan, 2013) [in Russian].
8. Alart, P., Lemaire, B. "Penalization in Non-Classical Convex Programming via Variational Convergence", Math. Program. **51**, No. 1, 307–331 (1991).
9. Cominetti, R. "Coupling the Proximal Point Algorithm With Approximation Methods", J. Optim. Theory Appl. **95**, No. 3, 581–600 (1997).
10. Antipin, A. S. and Vasil'ev, F. P. "A Stabilization Method for Equilibrium Programming Problems With an Approximately Given Set", Comput. Math. Math. Phys. **39**, No. 11, 1707–1714 (1999).
11. Salmon, G., Nguyen, V. H., Strodiot, J. J. "Coupling the Auxiliary Problem Principle and Epiconvergence Theory for Solving General Variational Inequalities", J. Optim. Theory Appl. **104**, No. 3, 629–657 (2000).
12. Kaplan, A., Tichatschke, R. "A General View on Proximal Point Methods for Variational Inequalities in Hilbert Spaces", J. Nonlin. Conv. Anal. **2**, No. 3, 305–332 (2001).
13. Konnov, I. V. "Application of Penalty Methods to Non-Stationary Variational Inequalities", Nonlinear Analysis: Theory, Methods and Appl. **92**, No. 1, 177–182 (2013).
14. Konnov, I. V. "Application of the Penalty Method to Nonstationary Approximation of an Optimization Problem", Russian Mathematics **58**, No. 8, 49–55 (2014).
15. Konnov I.V. "An Inexact Penalty Method for Non-Stationary Generalized Variational Inequalities", Set-valued and variational anal. **23**, No. 2, 239–248 (2015).
16. Fukushima, M., Mine, H. "A Generalized Proximal Point Algorithm for Certain Non-Convex Minimization Problems", Intern. J. Syst. Sci. **12**, No. 8, 989–1000 (1981).
17. Patriksson, M. "Cost Approximation: A Unified Framework of Descent Algorithms for Nonlinear Programs", SIAM J. Optim. **8**, No. 2, 561–582 (1998).
18. Ermoliev, Y. M., Norkin, V. I., Wets, R. J. B. "The Minimization of Semicontinuous Functions: Mollifier Subgradient", SIAM J. Contr. Optim. **33**, No. 1, 149–167 (1995).
19. Czarnecki, M.-O., Rifford, L. "Approximation and Regularization of Lipschitz Functions: Convergence of the Gradients", Trans. Amer. Math. Soc. **358**, No. 10, 4467–4520 (2006).
20. Gwinner, J. "On the Penalty Method for Constrained Variational Inequalities", in *Optimization*: *Theory and Algorithms*, Ed. by J.-B. Hiriart-Urruty, W. Oettli, and J. Stoer (Marcel Dekker, New York, 1981), pp. 197–211.
21. Blum, E., Oettli, W. "From Optimization and Variational Inequalities to Equilibrium Problems", The Math. Student **63**, No. 1, 123–145 (1994).

Authorized translation