Mathematical Modelling and Analysis
Volume 27, Issue 2, 242–262, 2022
https://doi.org/10.3846/mma.2022.14479

VILNIUS TECH
Vilnius Gediminas Technical University

Strong Convergence of Multi-Parameter Projection Methods for Variational Inequality Problems

Dang Van Hieu[a], **Le Dung Muu**[a,b] **and Pham Kim Quy**[c]

[a] *TIMAS - Thang Long University*
Hanoi, Vietnam
[b] *Institute of Mathematics - VAST*
Hanoi, Vietnam
[c] *Department of Mathematics, College of Air Force*
Nha Trang, Vietnam

E-mail(*corresp.*): dv.hieu83@gmail.com
E-mail: ldmuu@math.ac.vn
E-mail: pk.quy85@gmail.com

Received February 26, 2021; revised February 27, 2022; accepted February 27, 2022

Abstract. In this paper, we introduce a multi-parameter projection method for solving a variational inequality problem, and establish its strong convergence in a Hilbert space under appropriate conditions. The method involves two projection-steps with different variable stepsizes where one of them is computed explicitly on a specifically structural half-space. The proof of strong convergence of the method is based on the regularization solutions depending on parameters of the original problem. It turns out that the solution obtained by the method is the solution of a bilevel variational inequality problem whose constraint is the solution set of our considered problem. In order to support the obtained theoretical results, we perform some experiments on transportation equilibrium and optimal control problems, and also involve comparisons. Numerical results show the computational effectiveness and the fast convergence of the new method over some existing ones.

Keywords: variational inequality, monotonicity, Lipschitz continuity, iterative method, regularization.

AMS Subject Classification: 65Y05; 65K15.

1 Introduction

Let $\mathcal{H}$ be a real Hilbert space and Ω be a nonempty closed convex subset of $\mathcal{H}$. Let $\mathcal{A} : \mathcal{H} \to \mathcal{H}$ be an operator. The variational inequality problem (VIP) for $\mathcal{A}$ on Ω is stated as follows:

$$\text{Find } u^* \in \Omega \text{ such that } \langle \mathcal{A}u^*, u - u^* \rangle \geq 0, \ \forall u \in \Omega. \tag{VIP}$$

Throughout this paper, we write $VI(\mathcal{A}, \Omega)$ to stand for the solution set of problem VIP for an operator $\mathcal{A}$ on a feasible set Ω. The VIP is a central problem in nonlinear analysis and it plays a key role in optimization theory. The problem unifies many important concepts in applied mathematics such as systems of nonlinear operator equations, necessary optimality conditions, complementarity problems, obstacle problems, or network equilibrium problems [14,22,24,32]. This explains why the VIP becomes an attractive field for many authors who have devoted to studying not only the existence and the stability of solutions but also the effective approximation methods for solving this problem.

The metric projection is an important tool used to construct numerical methods for solving constraint optimization problems as well as VIPs. The metric projection mapping $P_\Omega : \mathcal{H} \to \Omega$ is defined by

$$P_\Omega(u) = \arg\min \{\|u - v\| : v \in \Omega\}, \ u \in \mathcal{H}.$$

Since Ω is nonempty, closed and convex in $\mathcal{H}$, the point $P_\Omega(u)$ exists uniquely for each $u \in \mathcal{H}$. The oldest method for solving the VIP is the gradient projection method which is given by

$$u_{n+1} = P_\Omega(u_n - \lambda \mathcal{A}u_n), \ u_0 \in \Omega, \ n \geq 0,$$

where λ is a suitable parameter. This method comes from minimizing a differentiable convex function on a nonempty closed convex set. Although the gradient projection method has a simple and elegant structure, its convergence requires such a strict assumption, for example, that the operator $\mathcal{A}$ is either strongly monotone or inverse strongly monotone. Without such an additional condition, the gradient projection method can diverge, for example, when $\mathcal{A}$ is the rotation mapping in the plane. In order to overcome this drawback, the method used more popularly is the extragradient method which was introduced early by Korpelevich [26] for solving saddle point problems in finite dimensional spaces. The extragradient method, which is applied to the class of monotone and L - Lipschitz continuous operators, is of the form

$$\begin{cases} v_n = P_\Omega(u_n - \lambda_n \mathcal{A}u_n), \\ u_{n+1} = P_\Omega(u_n - \lambda_n \mathcal{A}v_n), \end{cases} \tag{1.1}$$

where $\{\lambda_n\} \subset [a, b] \subset (0, \frac{1}{L})$. The extragradient method has been used to solve some large optimization problems in the field of information science such as in machine learning, optical network, and speech recognition [13,25]. Because of its importance, in recent years, this method has been modified and improved in various ways by some authors. One of those methods is the subgradient extragradient method introduced by Censor et al. in [5, 6, 7]. This method is described as follows:

$$\begin{cases} v_n = P_\Omega(u_n - \lambda_n \mathcal{A} u_n), \\ T_n = \{u \in \mathcal{H} : \langle u_n - \lambda_n \mathcal{A} u_n - v_n, u - v_n \rangle \leq 0\}, \\ u_{n+1} = P_{T_n}(u_n - \lambda_n \mathcal{A} v_n), \end{cases} \tag{1.2}$$

where $\{\lambda_n\} \subset [a, b] \subset (0, \frac{1}{L})$. An advantage of the subgradient extragradient method is that the second projection is found on a half-space which is inherently explicit. Then, it is more particularly interesting over the extragradient method in the case where the feasible set Ω has a complicated structure and finding a projection on it is expensive. Some other projection methods for solving problem VIP, which not are mentioned in details here, can be found, for examples, in [1, 19, 21, 30, 31, 33, 38, 40, 41].

A class of notable projection methods for solving monotone VIPs is the projection and contraction method [16, 17, 18, 36, 37]. Recently, this method has been developed in various different forms because of its numerically computational performance (see, e.g., [4,8,9,10,11,12]). The projection and contraction method [16, 37] for solving problem VIP can be formulated in the form,

$$\begin{cases} v_n = P_\Omega(u_n - \lambda_n \mathcal{A} u_n), \\ u_{n+1} = P_\Omega(u_n - \gamma \rho_n \lambda_n \mathcal{A} v_n), \end{cases} \tag{1.3}$$

where $\gamma \in (0, 2)$, $\{\lambda_n\} \subset [a, b] \subset (0, \frac{1}{L})$, and

$$\rho_n = \frac{\|u_n - v_n\|^2 - \lambda_n \langle u_n - v_n, \mathcal{A} u_n - \mathcal{A} v_n \rangle}{\|(u_n - v_n) - \lambda_n(\mathcal{A} u_n - \mathcal{A} v_n)\|^2}.$$

As the extragradient method [26] and the subgradient extragradient method [5, 6, 7], at each iteration, the projection and contraction method includes two computational steps. The first step is called the prediction step for finding v_n, and the second one is for the next iterate u_{n+1}. However, at each iteration, the projection and contraction method has used two different stepsizes, namely λ_n and $\gamma \rho_n \lambda_n$. This makes a difference with two aforementioned extragradient methods. It is known that the effectiveness of iterative methods in general depends strictly on the used stepsizes, and thus the strategy of stepsize choice is important in real problems. Some numerical experiments implemented in [4] demonstrate that the computational load of the projection and contraction method is about half of that of the extragradient methods. Besides, schemes (1.1), (1.2) and (1.3) in general provide the weak convergence. As be seen in [3], the strong convergence is more useful than the weak convergence, especially in infinite Hilbert spaces.

In this paper, we introduce a new iterative method of multi-parameter form for solving a monotone and Lipschitz VIP, and prove the strong convergence of the proposed method in a Hilbert space. Our method is included two steps of projections with different stepsizes. We first use a line-search procedure to find a suitable stepsize in the first step which aims to avoid its dependence on the Lipschitz constant of cost operator. This projection step is performed on the feasible set of the problem. Next, we modify the stepsize in the second step in the spirit of the projection and contraction method [16,37] and use the explicit projection on a specifically constructed half-space to find the next iterate. Together with these constructions, we also incorporate the regularization technique which aims to get the strong convergence of the obtained method. The convergence is proved under suitable conditions imposed on control parameters. The technique to obtain the strong convergence here is different to known ones in the literature, for examples, the Halpern method [27], the viscosity method [29], the (shrinking) hybrid projection method [5]. It turns out that the obtained solution from our method is the solution of a bilevel variational inequality problem whose constraint is the solution set of the considered VIP. Bilevel-like problems have received a lot of attention by some authors in recent years (see, e.g. [28, 34, 44]). In order to illustrate the computational effectiveness of the new method, we implement some numerical experiments in comparision with known methods.

The paper is organized as follows: we recall in Section 2 some definitions and fundamental results used in the paper. Next, we describe the new method in Section 3 and analyze its convergence. Finally, several numerical experiments are performed in Section 4 to show the numerical behavior of the method, and compare it with existing ones.

2 Preliminaries

Let $\mathcal{H}$ be a real Hilbert space with the inner product $\langle\cdot,\cdot\rangle$ and the induced norm $\|\cdot\|$. Let us begin with some concepts of monotonicity of an operator.

Definition 1. An operator $\mathcal{A}:\mathcal{H}\to\mathcal{H}$ is called:

(i) *monotone*, if

$$\langle \mathcal{A}u-\mathcal{A}v, u-v\rangle \geq 0,\ \forall u,\ v\in\mathcal{H};$$

(ii) γ - *strongly monotone*, if there exists a number $\gamma>0$ such that

$$\langle \mathcal{A}u-\mathcal{A}v, u-v\rangle \geq \gamma\|u-v\|^2,\ \forall u,\ v\in\mathcal{H};$$

(iii) L - *Lipschitz continuous*, if there exists a number $L>0$ such that

$$\|\mathcal{A}u-\mathcal{A}v\| \leq L\|u-v\|,\ \forall u,\ v\in\mathcal{H}.$$

In any Hilbert space $\mathcal{H}$, we have that, for all $u,\ v\in\mathcal{H}$,

$$\|u+v\|^2=\|u\|^2+2\langle u,v\rangle+\|v\|^2.$$

Let Ω be a nonempty closed convex subset in $\mathcal{H}$. The normal cone to Ω at a point $u\in\Omega$, given by

$$\mathcal{N}_\Omega(u)=\{w\in\mathcal{H}:\langle w,v-u\rangle\leq 0,\ \forall v\in\Omega\}.$$

The metric projection has the following characteristic properties (see, for example, [15, Proposition 3.5]).

Lemma 1. (i) $\langle P_\Omega(u) - P_\Omega(v), u - v\rangle \geq \|P_\Omega(u) - P_\Omega(v)\|^2$, $\forall u, v \in \mathcal{H}$;
(ii) $\|u - P_\Omega(v)\|^2 + \|P_\Omega(v) - v\|^2 \leq \|u - v\|^2, \forall u \in \Omega, v \in \mathcal{H}$;
(iii) $w = P_\Omega(u) \Leftrightarrow \langle u - w, v - w\rangle \leq 0, \quad \forall u \in \mathcal{H}, v \in \Omega$.

Given $x \in \mathcal{H}$ and $0 \neq v \in \mathcal{H}$, let $T = \{z \in \mathcal{H} : \langle v, z - x\rangle \leq 0\}$. Then, for all $u \in \mathcal{H}$, the projection $P_T(u)$ on the half-space T, is explicitly computed by

$$P_T(u) = u - \max\left\{0, \frac{\langle v, u - x\rangle}{\|v\|^2}\right\} v. \tag{2.1}$$

We have the following result regarding the regularization solutions of problem VIP (see, for example, [2, Lemma 6.6.1, page 353] or [20, Lemmas 3.1 and 3.2] for more general case).

Lemma 2. *Let Ω be a nonempty closed convex subset of $\mathcal{H}$ and $\mathcal{A} : \mathcal{H} \to \mathcal{H}$ be a monotone and Lipschitz continuous operator such that $VI(\mathcal{A}, \Omega)$ is nonempty. Suppose that $\mathcal{F} : \mathcal{H} \to \mathcal{H}$ is a strongly monotone and Lipschitz continuous mapping. For each $\alpha > 0$, set $\mathcal{A}_\alpha^\mathcal{F} = \mathcal{A} + \alpha\mathcal{F}$, and let $u_\alpha \in VI(\mathcal{A}_\alpha^\mathcal{F}, \Omega)$. Then, the followings hold:*
(i) *the sequence $\{u_\alpha\}$ is bounded;*
(ii) *there exists a number $M > 0$ such that $\|u_{\alpha_1} - u_{\alpha_2}\| \leq \frac{|\alpha_2 - \alpha_1|}{\alpha_1} M$ for all $\alpha_1 > 0$, $\alpha_2 > 0$;*
(iii) $\lim_{\alpha \to 0^+} u_\alpha = u^\dagger$, *where $u^\dagger \in VI(\mathcal{F}, VI(\mathcal{A}, \Omega))$.*

We need the following technical lemma to establish the convergence of our method.

Lemma 3. [43, Lemma 2.5] *Let $\{\Psi_n\}$ be a sequence of nonnegative real numbers. Suppose that*

$$\Psi_{n+1} \leq (1 - p_n)\Psi_n + q_n$$

for all $n \geq 0$, where the sequences $\{p_n\}$ in $(0, 1)$ and $\{q_n\}$ in $\Re$ satisfy the conditions: $\lim_{n\to\infty} p_n = 0$, $\sum_{n=1}^\infty p_n = \infty$ and $\limsup_{n\to\infty} \frac{q_n}{p_n} \leq 0$. Then $\lim_{n\to\infty} \Psi_n = 0$.

3 Multi-parameter projection methods

Let $\mathcal{F} : \mathcal{H} \to \mathcal{H}$ be a γ-strongly monotone and k-Lipschitz continuous operator. Consider our VIP where $\mathcal{A}$ is monotone and Lipschitz continuous. In addition, we take a sequence $\{\alpha_n\} \subset (0, +\infty)$ such that

$$\text{(C1)} : \lim_{n\to\infty} \alpha_n = 0; \quad \text{(C2)} : \sum_{n=1}^\infty \alpha_n = +\infty; \quad \text{(C3)} : \lim_{n\to\infty} (\alpha_n - \alpha_{n-1})\alpha_n^{-2} = 0.$$

The sequence $\alpha_n = \frac{1}{n^p}$ with $0 < p < 1$ satisfies conditions (C1)–(C3). For the sake of simplicity for the description of the method, we adopt the following conventions: $\frac{0}{0} = +\infty$ and $\frac{1}{0} = +\infty$. For each $n \in \mathbb{N}$, set $\mathcal{A}_{\alpha_n}^\mathcal{F} = \mathcal{A} + \alpha_n\mathcal{F}$. In order to solve problem VIP, we introduce the following iterative method.

Algorithm 1. [Multi-Parameter Projection Method - MPPM]

Initialization: Take $u_0 \in \mathcal{H}$, $\beta > 0$, $r \in (0,2)$, $\sigma > 0$, $l \in (0,1)$, $\mu \in (0,1)$.
Iterative Steps:

1. Compute
$$v_n = P_\Omega(u_n - \lambda_n \mathcal{A}^{\mathcal{F}}_{\alpha_n} u_n),$$
where λ_n is the largest number $\lambda \in \{\sigma, \sigma l, \sigma l^2, \dots\}$ such that
$$\lambda_n \|\mathcal{A}u_n - \mathcal{A}v_n\| \le \mu \|u_n - v_n\|. \tag{3.1}$$
2. Set $T_n = \{w \in \mathcal{H} : \langle u_n - \lambda_n \mathcal{A}^{\mathcal{F}}_{\alpha_n} u_n - v_n, w - v_n \rangle \le 0\}$. Compute
$$u_{n+1} = P_{T_n}(u_n - r\beta_n \lambda_n \mathcal{A}^{\mathcal{F}}_{\alpha_n} v_n - r\beta_n\lambda_n\alpha_n(\mathcal{F}u_n - \mathcal{F}v_n)),$$
where $d(u_n, v_n) = u_n - v_n - \lambda_n [\mathcal{A}u_n - \mathcal{A}v_n]$ and
$$\beta_n = \min\left\{\beta, \frac{\langle u_n - v_n, d(u_n, v_n)\rangle}{\|d(u_n, v_n)\|^2}\right\}.$$
3. Increase n by 1 and go back to Step 1.

Remark that the feasible set $\Omega \subset T_n$ for all $n \in \mathbb{N}$. This follows immediately from the definitions of v_n, T_n and Lemma 1(iii). The projection u_{n+1} is performed on a half-space T_n and is inherently explicit (see formula (2.1) in Section 2). We take a number $\beta > 0$ to aim that the sequence $\{\beta_n\}$ is bounded from up. In fact, this sequence is bounded from below by a positive number. When the Lipschitz constant of $\mathcal{A}$ is known, we can take $\{\lambda_n\} \subset [a, b] \subset \left(0, \frac{1}{L}\right)$, because in this case, stepsize rule (3.1) is satisfied immediately at the first checked step (with $0 < \sigma \le \frac{\mu}{L}$). Otherwise, when L is unknown, stepsize rule (3.1) is well-defined. Precisely, we have the following lemma.

Lemma 4. (i) *Stepsize rule (3.1) is well-defined and there exists a number* $\underline{\sigma}$ *such that*
$$0 < \underline{\sigma} \le \lambda_n \le \sigma, \ \forall n \in \mathbb{N}. \tag{3.2}$$
(ii) *There exists a number* $\underline{\beta}$ *such that*
$$0 < \underline{\beta} \le \beta_n \le \beta, \ \forall n \in \mathbb{N}.$$

Proof. (i) For each $n \in \mathbb{N}$, set
$$v_n^m = P_\Omega(u_n - \sigma l^m \mathcal{A}^{\mathcal{F}}_{\alpha_n} u_n), \ m \in \mathbb{N}.$$
Let L be the Lipschitz constant of $\mathcal{A}$. We have $\|\mathcal{A}v_n^m - \mathcal{A}u_n\| \le L\|v_n^m - u_n\|$, or
$$\frac{\mu}{L}\|\mathcal{A}v_n^m - \mathcal{A}u_n\| \le \mu\|v_n^m - u_n\|.$$
Therefore, if choose $m \in \mathbb{N}$ such that $\sigma l^m \le \frac{\mu}{L}$, then inequality (3.1) holds. Thus, this rule is well-defined. Now, we prove inequality (3.2). Indeed, from the definition, it is obvious that $\lambda_n \le \sigma$ for all $n \in \mathbb{N}$.

If $\lambda_n = \sigma$, inequality (3.2) holds for $\underline{\sigma} = \sigma$. Otherwise, if $\lambda_n < \sigma$, then λ_n/l violates inequality (3.2), i.e.,

$$\frac{\lambda_n}{l}\|\mathcal{A}u_n - \mathcal{A}v_n^{m-1}\| > \mu\|u_n - v_n^{m-1}\|.$$

Hence, since $\mathcal{A}$ is L-Lipschitz continuous, we obtain immediately that $\frac{\lambda_n}{l}L > \mu$ or $\lambda_n > \frac{\mu l}{L}$. Now, if take $\underline{\sigma} = \min\left\{\sigma, \frac{\mu l}{L}\right\} > 0$ then inequality (3.2) holds for all $n \in \mathbb{N}$.

(ii) By the definition, we see that $\beta_n \le \beta$ for all $n \in \mathbb{N}$. Also, from the definition of $d(u_n, v_n)$ and rule (3.1), we have

$$\begin{aligned}\langle u_n - v_n, d(u_n, v_n)\rangle &= \langle u_n - v_n, u_n - v_n - \lambda_n [\mathcal{A}u_n - \mathcal{A}v_n]\rangle \\ &= \|u_n - v_n\| - \lambda_n \langle u_n - v_n, \mathcal{A}u_n - \mathcal{A}v_n\rangle \\ &\ge \|u_n - v_n\| - \lambda_n\|u_n - v_n\|\,\|\mathcal{A}u_n - \mathcal{A}v_n\| \\ &\ge \|u_n - v_n\| - \mu\|u_n - v_n\|^2 = (1-\mu)\|u_n - v_n\|^2. \end{aligned} \tag{3.3}$$

On the other hand, from the triangle inequality and rule (3.1), one gets

$$\begin{aligned}\|d(u_n, v_n)\| &= \|u_n - v_n - \lambda_n [\mathcal{A}u_n - \mathcal{A}v_n]\| \\ &\le \|u_n - v_n\| + \lambda_n\|\mathcal{A}u_n - \mathcal{A}v_n\| \\ &\le \|u_n - v_n\| + \mu\|u_n - v_n\| = (1+\mu)\|u_n - v_n\|. \end{aligned} \tag{3.4}$$

Observe that if $v_n = u_n$, then $\beta_n = \beta$. Otherwise, if $v_n \ne u_n$, from relations (3.3) and (3.4), we derive

$$\beta_n = \min\left\{\beta, \frac{\langle u_n - v_n, d(u_n, v_n)\rangle}{\|d(u_n, v_n)\|^2}\right\} \ge \min\left\{\beta, \frac{1-\mu}{(1+\mu)^2}\right\}.$$

Set $\underline{\beta} = \min\left\{\beta, \frac{1-\mu}{(1+\mu)^2}\right\}$. Then $\beta_n \ge \underline{\beta} > 0$ for all $n \in \mathbb{N}$. This finishes the proof. □

Lemma 5. *For all $n \in \mathbb{N}$ and $\rho_n \in (0,1)$, we have*

$$\|u_{\alpha_{n+1}} - u\|^2 \le \frac{1}{1-\rho_n}\|u_{\alpha_n} - u\|^2 + \frac{M^2}{\rho_n}\frac{|\alpha_{n+1} - \alpha_n|^2}{\alpha_n^2}, \ \forall u \in \mathcal{H},$$

where $u_{\alpha_n} \in VI(\mathcal{A}^{\mathcal{F}}_{\alpha_n}, \Omega)$ and $u_{\alpha_{n+1}} \in VI(\mathcal{A}^{\mathcal{F}}_{\alpha_{n+1}}, \Omega)$.

Proof. By Lemma 2 (ii), we obtain

$$\|u_{\alpha_{n+1}} - u_{\alpha_n}\| \le \frac{|\alpha_{n+1} - \alpha_n|}{\alpha_n}M. \tag{3.5}$$

We have

$$\begin{aligned}\|u - u_{\alpha_n}\|^2 &= \|u - u_{\alpha_{n+1}}\|^2 + \|u_{\alpha_{n+1}} - u_{\alpha_n}\|^2 + 2\left\langle u - u_{\alpha_{n+1}}, u_{\alpha_{n+1}} - u_{\alpha_n}\right\rangle \\ &\ge \|u - u_{\alpha_{n+1}}\|^2 + \|u_{\alpha_{n+1}} - u_{\alpha_n}\|^2 - 2\|u - u_{\alpha_{n+1}}\|\,\|u_{\alpha_{n+1}} - u_{\alpha_n}\| \\ &\ge \|u - u_{\alpha_{n+1}}\|^2 + \|u_{\alpha_{n+1}} - u_{\alpha_n}\|^2 - \rho_n\|u - u_{\alpha_{n+1}}\|^2 - \frac{1}{\rho_n}\|u_{\alpha_{n+1}} - u_{\alpha_n}\|^2 \\ &= (1-\rho_n)\|u - u_{\alpha_{n+1}}\|^2 - \frac{1-\rho_n}{\rho_n}\|u_{\alpha_{n+1}} - u_{\alpha_n}\|^2, \end{aligned}$$

which, by relation (3.5), implies that

$$\|u-u_{\alpha_n}\|^2 \geq (1-\rho_n)\|u-u_{\alpha_{n+1}}\|^2 - \frac{M^2(1-\rho_n)}{\rho_n}\frac{|\alpha_{n+1}-\alpha_n|^2}{\alpha_n^2}.$$

Thus, since $\rho_n \in (0,1)$, we get

$$\|u-u_{\alpha_{n+1}}\|^2 \leq \frac{1}{1-\rho_n}\|u-u_{\alpha_n}\|^2 + \frac{M^2}{\rho_n}\frac{|\alpha_{n+1}-\alpha_n|^2}{\alpha_n^2},\ \forall \rho_n \in (0,1),\ \forall u \in \mathcal{H}.$$

The proof is completed. □

Lemma 6. *For all $n \in \mathbb{N}$ and some $\rho \in \left(0, \frac{\gamma}{k}\right)$, we have*

$$\|u_{n+1}-u_{\alpha_n}\|^2 \quad\leq\quad (1-\Xi\alpha_n)\|u_n-u_{\alpha_n}\|^2 - \Gamma_n\|d(u_n,v_n)\|^2,$$

where $\Xi = (\gamma-k\rho)r\underline{\beta}\underline{\sigma}$ and $\Gamma_n = \underline{\beta}^2 r(2-r) - \frac{kr\beta\sigma}{\rho(1-\mu)^2}\alpha_n$.

Proof. Set $t_n = r\beta_n\lambda_n\mathcal{A}^{\mathcal{F}}_{\alpha_n}v_n - r\beta_n\lambda_n\alpha_n(\mathcal{F}u_n - \mathcal{F}v_n)$. Then, u_{n+1} can be rewritten as $u_{n+1} = P_{T_n}(u_n - t_n)$. Therefore, it follows from Lemma 1 (iii) and the fact $u_{\alpha_n} \in VI(\mathcal{A}^{\mathcal{F}}_{\alpha_n}, \Omega) \subset \Omega \subset T_n$ that

$$\begin{aligned}\|u_{n+1}-u_{\alpha_n}\|^2 &\leq \|u_n-t_n-u_{\alpha_n}\|^2 - \|u_n-t_n-u_{n+1}\|^2\\ &= \|u_n-u_{\alpha_n}\|^2 - \|u_n-u_{n+1}\|^2 - 2\left\langle t_n, u_{n+1}-u_{\alpha_n}\right\rangle.\end{aligned}\tag{3.6}$$

We have from the definition of t_n that

$$\begin{aligned}\left\langle t_n, u_{n+1}-u_{\alpha_n}\right\rangle &= r\beta_n\lambda_n\left\langle \mathcal{A}^{\mathcal{F}}_{\alpha_n}v_n, u_{n+1}-u_{\alpha_n}\right\rangle\\ &\quad + r\beta_n\lambda_n\alpha_n\left\langle \mathcal{F}u_n-\mathcal{F}v_n, u_{n+1}-u_{\alpha_n}\right\rangle = r\beta_n\lambda_n\\ &\quad \times\left\langle \mathcal{A}v_n+\alpha_n\mathcal{F}v_n, u_{n+1}-u_{\alpha_n}\right\rangle + r\beta_n\lambda_n\alpha_n\left\langle \mathcal{F}u_n-\mathcal{F}v_n, u_{n+1}-u_{\alpha_n}\right\rangle\\ &= r\beta_n\lambda_n\left\langle \mathcal{A}v_n, u_{n+1}-u_{\alpha_n}\right\rangle + r\beta_n\lambda_n\alpha_n\left\langle \mathcal{F}u_n, u_{n+1}-u_{\alpha_n}\right\rangle.\end{aligned}\tag{3.7}$$

Substituting relation (3.7) into relation (3.6), we derive

$$\begin{aligned}\|u_{n+1}-u_{\alpha_n}\|^2 \quad\leq\quad &\|u_n-u_{\alpha_n}\|^2 - \|u_n-u_{n+1}\|^2 - 2r\beta_n\lambda_n\left\langle \mathcal{A}v_n, u_{n+1}-u_{\alpha_n}\right\rangle\\ &- 2r\beta_n\lambda_n\alpha_n\left\langle \mathcal{F}u_n, u_{n+1}-u_{\alpha_n}\right\rangle.\end{aligned}\tag{3.8}$$

By the monotonicity of $\mathcal{A}$ and $u_{\alpha_n} \in VI(\mathcal{A}^{\mathcal{F}}_{\alpha_n}, \Omega)$, we have

$$\left\langle \mathcal{A}v_n - \mathcal{A}u_{\alpha_n}, v_n-u_{\alpha_n}\right\rangle \quad\geq\quad 0, \tag{3.9}$$

$$\left\langle \mathcal{A}u_{\alpha_n} + \alpha_n\mathcal{F}u_{\alpha_n}, v_n-u_{\alpha_n}\right\rangle \quad\geq\quad 0. \tag{3.10}$$

Summing up two inequalities, we get $\left\langle \mathcal{A}v_n + \alpha_n\mathcal{F}u_{\alpha_n}, v_n-u_{\alpha_n}\right\rangle \geq 0$, or

$$\left\langle \mathcal{A}v_n, v_n-u_{\alpha_n}\right\rangle \geq \alpha_n\left\langle \mathcal{F}u_{\alpha_n}, u_{\alpha_n}-v_n\right\rangle.$$

Thus

$$\begin{aligned}\left\langle \mathcal{A}v_n, u_{n+1}-u_{\alpha_n}\right\rangle &= \left\langle \mathcal{A}v_n, u_{n+1}-v_n\right\rangle + \left\langle \mathcal{A}v_n, v_n-u_{\alpha_n}\right\rangle\\ &\geq \left\langle \mathcal{A}v_n, u_{n+1}-v_n\right\rangle + \alpha_n\left\langle \mathcal{F}u_{\alpha_n}, u_{\alpha_n}-v_n\right\rangle.\end{aligned}\tag{3.11}$$

On the other hand, since $u_{n+1} \in T_n$, we derive that

$$\left\langle u_n - \lambda_n \mathcal{A}^{\mathcal{F}}_{\alpha_n} u_n - v_n, u_{n+1} - v_n \right\rangle \leq 0,$$

or, equivalently,

$$\langle u_n - \lambda_n \mathcal{A} u_n - \alpha_n \lambda_n \mathcal{F} u_n - v_n, u_{n+1} - v_n \rangle \leq 0.$$

Hence,

$$\begin{aligned}\langle u_n - v_n - \lambda_n \left(\mathcal{A} u_n - \mathcal{A} v_n\right), u_{n+1} - v_n \rangle \leq & \lambda_n \langle \mathcal{A} v_n, u_{n+1} - v_n \rangle \\ & + \alpha_n \lambda_n \langle \mathcal{F} u_n, u_{n+1} - v_n \rangle .\end{aligned}$$

Thus, by dividing both sides of the last inequality by $\lambda_n > 0$ and using the fact $d(u_n, v_n) = u_n - v_n - \lambda_n \left(\mathcal{A} u_n - \mathcal{A} v_n\right)$, one gets

$$\langle \mathcal{A} v_n, u_{n+1} - v_n \rangle \geq \frac{1}{\lambda_n} \langle d(u_n, v_n), u_{n+1} - v_n \rangle + \alpha_n \langle \mathcal{F} u_n, v_n - u_{n+1} \rangle . \quad (3.12)$$

Summing up inequalities (3.11) and (3.12), we come to the following one,

$$\begin{aligned}\langle \mathcal{A} v_n, u_{n+1} - u_{\alpha_n} \rangle \geq & \frac{1}{\lambda_n} \langle d(u_n, v_n), u_{n+1} - v_n \rangle \alpha_n \langle \mathcal{F} u_n, v_n - u_{n+1} \rangle \\ & + \alpha_n \langle \mathcal{F} u_{\alpha_n}, u_{\alpha_n} - v_n \rangle . \qquad (3.13)\end{aligned}$$

By relations (3.8) and (3.13), we get

$$\begin{aligned}\|u_{n+1} - u_{\alpha_n}\|^2 \leq & \|u_n - u_{\alpha_n}\|^2 - \|u_n - u_{n+1}\|^2 \\ & - 2r\beta_n \langle d(u_n, v_n), u_{n+1} - v_n \rangle\rangle - 2r\beta_n \lambda_n \alpha_n \langle \mathcal{F} u_n, v_n - u_{n+1} \rangle \\ & - 2r\beta_n \lambda_n \alpha_n \langle \mathcal{F} u_{\alpha_n}, u_{\alpha_n} - v_n \rangle - 2r\beta_n \lambda_n \alpha_n \langle \mathcal{F} u_n, u_{n+1} - u_{\alpha_n} \rangle \\ = & \|u_n - u_{\alpha_n}\|^2 - \left[\|u_n - u_{n+1}\|^2 + 2r\beta_n \langle d(u_n, v_n), u_{n+1} - v_n \rangle\rangle\right] \\ & - 2r\beta_n \lambda_n \alpha_n \langle \mathcal{F} u_n - \mathcal{F} u_{\alpha_n}, v_n - u_{\alpha_n} \rangle . \qquad (3.14)\end{aligned}$$

Moreover, we have

$$\begin{aligned}& - 2r\beta_n \langle d(u_n, v_n), u_{n+1} - v_n \rangle\rangle = -2r\beta_n \langle d(u_n, v_n), u_{n+1} - u_n \rangle \\ & - 2r\beta_n \langle d(u_n, v_n), u_n - v_n \rangle = \|u_{n+1} - u_n\|^2 + r^2 \beta_n^2 \|d(u_n, v_n)\|^2 \\ & - \|r\beta_n d(u_n, v_n) + u_{n+1} - u_n\|^2 - 2r\beta_n \langle d(u_n, v_n), u_n - v_n \rangle . \qquad (3.15)\end{aligned}$$

By the definition of β_n, we obtain $\langle d(u_n, v_n), u_n - v_n \rangle \geq \beta_n \|d(u_n, v_n)\|^2$. Hence, from (3.15), we come to the inequality,

$$\begin{aligned}& - 2r\beta_n \langle d(u_n, v_n), u_{n+1} - v_n \rangle\rangle \leq \|u_{n+1} - u_n\|^2 + r^2 \beta_n^2 \|d(u_n, v_n)\|^2 \\ & - \|r\beta_n d(u_n, v_n) + u_{n+1} - u_n\|^2 - 2r\beta_n^2 \|d(u_n, v_n)\|^2 \\ & = \|u_{n+1} - u_n\|^2 - \beta_n^2 r(2 - r) \|d(u_n, v_n)\|^2 - \|r\beta_n d(u_n, v_n) + u_{n+1} - u_n\|^2 \\ & \leq \|u_{n+1} - u_n\|^2 - \beta_n^2 r(2 - r) \|d(u_n, v_n)\|^2.\end{aligned}$$

Thus,

$$\|u_{n+1}-u_n\|^2 + 2r\beta_n \langle d(u_n,v_n), u_{n+1}-v_n\rangle \geq \beta_n^2 r(2-r)\|d(u_n,v_n)\|^2.$$

This together with (3.14) implies that

$$\begin{aligned}\|u_{n+1}-u_{\alpha_n}\|^2 \leq &\|u_n-u_{\alpha_n}\|^2 - \beta_n^2 r(2-r)\|d(u_n,v_n)\|^2 \\ &- 2r\beta_n\lambda_n\alpha_n \langle \mathcal{F}u_n - \mathcal{F}u_{\alpha_n}, v_n-u_{\alpha_n}\rangle .\end{aligned} \tag{3.16}$$

Now, we estimate the last term in the right-hand side of (3.16). Using the γ-strongly monotonicity and the L-Lipschitz continuity of $\mathcal{F}$, we obtain

$$\begin{aligned}2\langle \mathcal{F}u_n-\mathcal{F}u_{\alpha_n}, v_n-u_{\alpha_n}\rangle =&2\langle \mathcal{F}u_n-\mathcal{F}u_{\alpha_n}, v_n-u_n\rangle + 2\langle \mathcal{F}u_n-\mathcal{F}u_{\alpha_n}, u_n-u_{\alpha_n}\rangle \\ &\geq -2k\|u_n-u_{\alpha_n}\|\,\|v_n-u_n\| + \gamma\|u_n-u_{\alpha_n}\|^2 \\ &\geq -k\rho\|u_n-u_{\alpha_n}\|^2 - \frac{k}{\rho}\|v_n-u_n\|^2 + \gamma\|u_n-u_{\alpha_n}\|^2 \\ &\geq (\gamma-k\rho)\|u_n-u_{\alpha_n}\|^2 - \frac{k}{\rho}\|v_n-u_n\|^2.\end{aligned} \tag{3.17}$$

By relations (3.16) and (3.17), we get

$$\begin{aligned}\|u_{n+1}-u_{\alpha_n}\|^2 \leq &(1-(\gamma-k\rho)r\beta_n\lambda_n\alpha_n)\|u_n-u_{\alpha_n}\|^2 \\ &- \beta_n^2 r(2-r)\|d(u_n,v_n)\|^2 + \frac{kr\beta_n\lambda_n\alpha_n}{\rho}\|v_n-u_n\|^2.\end{aligned} \tag{3.18}$$

From relation (3.3), we have $(1-\mu)\|u_n-v_n\|^2 \leq \langle u_n-v_n, d(u_n,v_n)\rangle$. Thus,

$$(1-\mu)\|u_n-v_n\|^2 \leq \langle u_n-v_n, d(u_n,v_n)\rangle \leq \|u_n-v_n\|\,\|d(u_n,v_n)\|.$$

This implies that $\|u_n-v_n\| \leq \frac{1}{1-\mu}\|d(u_n,v_n)\|$. Hence, by relation (3.18), we obtain

$$\begin{aligned}\|u_{n+1}-u_{\alpha_n}\|^2\leq&(1-(\gamma-k\rho)r\beta_n\lambda_n\alpha_n)\|u_n-u_{\alpha_n}\|^2-\beta_n^2 r(2-r)\|d(u_n,v_n)\|^2 \\ &+ \frac{kr\beta_n\lambda_n\alpha_n}{\rho(1-\mu)^2}\|d(u_n,v_n)\|^2 = (1-(\gamma-k\rho)r\beta_n\lambda_n\alpha_n)\|u_n-u_{\alpha_n}\|^2 \\ &- \left[\beta_n^2 r(2-r) - \frac{kr\beta_n\lambda_n\alpha_n}{\rho(1-\mu)^2}\right]\|d(u_n,v_n)\|^2.\end{aligned} \tag{3.19}$$

Recall from Lemma 4 that $0<\underline{\sigma}\leq\lambda_n\leq\sigma$ and $0<\underline{\beta}\leq\beta_n\leq\beta$ for all $n\in\mathbb{N}$. Thus, by relation (3.19), we get

$$\begin{aligned}\|u_{n+1}-u_{\alpha_n}\|^2 \leq &(1-(\gamma-k\rho)r\underline{\beta}\underline{\sigma}\alpha_n)\|u_n-u_{\alpha_n}\|^2 - \Big[\underline{\beta}^2 r(2-r) \\ &- \frac{kr\beta\sigma\alpha_n}{\rho(1-\mu)^2}\Big]\|d(u_n,v_n)\|^2 = (1-\Xi\alpha_n)\|u_n-u_{\alpha_n}\|^2 - \mathit{\Gamma}_n\|d(u_n,v_n)\|^2.\end{aligned}$$

This completes the proof. □

Finally, we prove the main theorem.

Theorem 1. *Suppose that $\mathcal{A} : \mathcal{H} \to \mathcal{H}$ is a monotone and Lipschitz continuous operator such that the solution set $VI(\mathcal{A}, \Omega)$ of problem* (VIP) *is nonempty. Let $\mathcal{F} : \mathcal{H} \to \mathcal{H}$ is a strongly monotone and Lipschitz continuous operator. Then, the sequence $\{u_n\}$ generated by Algorithm 1 converges strongly to a solution $u^\dagger$ of problem* (VIP)*, where $u^\dagger \in VI(\mathcal{F}, VI(\mathcal{A}, \Omega))$.*

Proof. Take $\rho_n = 0.5\Xi\alpha_n$. Since $\alpha_n \to 0$, there exists a number $n_0 \in \mathbb{N}$ such that $\rho_n \in (0,1)$ for all $n \geq n_0$. Using Lemma 5 for $\rho_n = 0.5\Xi\alpha_n$ and $u = u_{n+1}$ with $n \geq n_0$, we get

$$\|u_{n+1}-u_{\alpha_{n+1}}\|^2 \leq \frac{|u_{n+1}-u_{\alpha_n}\|^2}{1-0.5\Xi\alpha_n} + \frac{M^2}{0.5\Xi}\frac{|\alpha_{n+1}-\alpha_n|^2}{\alpha_n^3}, \quad \forall n \geq n_0. \tag{3.20}$$

Combining relation (3.20) and Lemma 6, we derive for all $n \geq n_0$ that

$$\begin{aligned}\|u_{n+1}-u_{\alpha_{n+1}}\|^2 \leq & \frac{1-\Xi\alpha_n}{1-0.5\Xi\alpha_n}\|u_n-u_{\alpha_n}\|^2 + \frac{M^2}{0.5\Xi}\frac{|\alpha_{n+1}-\alpha_n|^2}{\alpha_n^3} \\ & - \frac{\Gamma_n}{1-0.5\Xi\alpha_n}\|d(u_n,v_n)\|^2.\end{aligned}$$

Set $\Psi_n = \|u_n - u_{\alpha_n}\|^2$, $p_n = \frac{0.5\Xi\alpha_n}{1-0.5\Xi\alpha_n}$ and

$$q_n = \frac{M^2}{0.5\Xi}\frac{|\alpha_{n+1}-\alpha_n|^2}{\alpha_n^3} - \frac{\Gamma_n}{1-0.5\Xi\alpha_n}\|d(u_n,v_n)\|^2.$$

Thus,

$$\Psi_{n+1} \leq (1-p_n)\Psi_n + q_n, \ \forall n \geq n_0. \tag{3.21}$$

Note that from conditions (C1)–(C2), we have $p_n \to 0$ and $\sum_{n=n_0}^{\infty} p_n = +\infty$. Moreover, we have

$$\lim_{n\to\infty}\frac{q_n}{p_n} = \lim_{n\to\infty}\Big(\frac{M^2(1-0.5\Xi\alpha_n)}{(0.5\Xi)^2}\frac{|\alpha_{n+1}-\alpha_n|^2}{\alpha_n^4} - \frac{\Gamma_n}{0.5\Xi\alpha_n}\|d(u_n,v_n)\|^2\Big). \tag{3.22}$$

Since $\Gamma_n = \underline{\beta}^2 r(2-r) - \frac{kr\beta\sigma}{\rho(1-\mu)^2}\alpha_n$ and $\alpha_n \to 0$, we obtain that $\lim_{n\to\infty}\Gamma_n = \underline{\beta}^2 r(2-r) > 0$. Thus, by relation (3.22) and condition (C3), we get

$$\lim_{n\to\infty}\frac{q_n}{p_n} \leq \lim_{n\to\infty}\frac{M^2(1-0.5\Xi\alpha_n)}{(0.5\Xi)^2}\frac{|\alpha_{n+1}-\alpha_n|^2}{\alpha_n^4} = 0.$$

Consequently, Lemma 3 and relation (3.21) ensure that $\Psi_n = \|u_n - u_{\alpha_n}\|^2 \to 0$ as $n \to \infty$. Since $\alpha_n \to 0$ and Lemma 2 (iii), we find that $u_{\alpha_n} \to u^\dagger$ as $n \to \infty$. Thus, $u_n \to u^\dagger$ as $n \to \infty$. This completes the proof. □

In the case when $\mathcal{F} = I - u^g$, where u^g is a suggested point in $\mathcal{H}$. Then, from Lemma 1 (iii), we have

$$\begin{aligned} VI(\mathcal{F}, VI(\mathcal{A},\Omega)) &= \left\{u^\dagger \in VI(\mathcal{A},\Omega) : \left\langle \mathcal{F}u^\dagger, u^* - u^\dagger\right\rangle \geq 0, \ \forall u^* \in VI(\mathcal{A},\Omega)\right\} \\ &= \left\{u^\dagger \in VI(\mathcal{A},\Omega) : \left\langle u^\dagger - u^g, u^* - u^\dagger\right\rangle \geq 0, \ \forall u^* \in VI(\mathcal{A},\Omega)\right\} \\ &= \left\{u^\dagger \in VI(\mathcal{A},\Omega) : \left\langle u^g - u^\dagger, u^* - u^\dagger\right\rangle \leq 0, \ \forall u^* \in VI(\mathcal{A},\Omega)\right\} \\ &= \left\{P_{VI(\mathcal{A},\Omega)}(u^g)\right\}. \end{aligned}$$

Thus, the following corollary follows directly from Theorem 1.

Corollary 1. The sequence $\{u_n\}$ generated by Algorithm 1 with $\mathcal{F} = I - u^g$ converges strongly to a solution $u^\dagger$ of problem (VIP), where $u^\dagger = P_{VI(\mathcal{A},\Omega)}(u^g)$.

Remark 1. As the suggestion of a reviewer, we can replace the Armijo linesearch procedure at Step 1 of Algorithm 1 by a self-adaptive stepsize rule. More precisely, take $\lambda_0 > 0$, $\mu \in (0,1)$ and a sequence $\{p_n\} \subset [0, +\infty)$ is summable, i.e., $\sum_{n=1}^{\infty} p_n < +\infty$. At the n^{th}-step, compute $v_n = P_\Omega(u_n - \lambda_n \mathcal{A}^{\mathcal{F}}_{\alpha_n} u_n)$, after that, update λ_n for the next step by

$$\lambda_{n+1} = \min\left\{\lambda_n + p_n, \frac{\mu\|v_n - u_n\|}{\|\mathcal{A}u_n - \mathcal{A}v_n\|}\right\}. \tag{3.23}$$

It is not difficult to show that the sequence $\{\lambda_n\}$ generated by rule (3.23) is bounded. Actually, we have $\lambda_n > 0$ for all $n \geq 0$, $\lim_{n\to\infty} \lambda_n = \lambda > 0$ and

$$0 < \min\left\{\lambda_0, \frac{\mu}{L}\right\} = \underline{\sigma} \leq \lambda_n \leq \overline{\sigma} = \lambda_0 + \sum_{n=1}^{\infty} p_n < +\infty. \tag{3.24}$$

A proof for these claims can be found, for example, in [17]. By relation (3.23), we obtain

$$\|\mathcal{A}u_n - \mathcal{A}v_n\| \leq \frac{\mu\|v_n - u_n\|}{\lambda_{n+1}}, \ \forall n \geq 0. \tag{3.25}$$

Using relation (3.25) and repeating the proof of Lemma 4 (ii), we obtain

$$\min\left\{\beta, \left(1 - \mu\frac{\lambda_n}{\lambda_{n+1}}\right)\Big/\left(1 + \mu\frac{\lambda_n}{\lambda_{n+1}}\right)^2\right\} \leq \beta_n \leq \beta. \tag{3.26}$$

Since $\lambda_n \to \lambda > 0$ as $n \to \infty$, we derive

$$\lim_{n\to\infty} \min\left\{\beta, \left(1-\mu\frac{\lambda_n}{\lambda_{n+1}}\right)\Big/\left(1+\mu\frac{\lambda_n}{\lambda_{n+1}}\right)^2\right\} = \min\left\{\beta, \frac{1-\mu}{(1+\mu)^2}\right\} > 0. \tag{3.27}$$

Take a number $\underline{\beta}$ such that $0 < \underline{\beta} < \min\left\{\beta, \frac{1-\mu}{(1+\mu)^2}\right\}$. From relations (3.26) and (3.27), there exists $\overline{n}_0 \geq 1$ such that

$$0 < \underline{\beta} \leq \beta_n \leq \beta, \ \forall n \geq \overline{n}_0. \tag{3.28}$$

Employing relations (3.24) and (3.28), we also obtain the inequalities in Lemma 5 and Lemma 6. Thus, the conclusion of Theorem 1 still remains valuable in this case.

4 Numerical illustrations

This section is devoted to testing the computational effectiveness of Algorithm 1 (shortly, MPPM, with two stepsize rules: (3.1) and (3.23)) over some existing methods including the Halpern Subgradient Extragradient Method (HSEGM) [27], the Viscosity Subgradient Extragradient Method (VSEGM) [39, Algorithm 3.1], the Viscosity Tseng's Extragradient Method (VTEGM) [39, Algorithm 3.2], the Hybrid Extragradient Viscosity Method (HEGVM) [29].

All the programs are written in Matlab 7.0 and performed on a PC Desktop Intel(R) Core(TM) i5-3210M CPU @ 2.50GHz, RAM 2.00 GB.

Example 1 [Transportation Equilibrium Problems]. Consider a traffic network consisting of N nodes connected by oriented edges [25, Chapter 6]. Suppose that D is the set of edges of the network and W is the set of oriented pairs of the nodes. Each element $w \in W$ is of the form $w = (a, b)$ where a is the original node and b is the destination one. Let P_w be the set of all paths from node a to node b, and $\mathbb{P}$ be the set of all paths in the network, i.e., $\mathbb{P} = \cup_{w \in W} P_w$. Let us denote by u_p, for each $p \in \mathbb{P}$, the path flow for the path p. Each pair $w \in W$ is associated with a positive number d_w which gives the flow demand from a to b. The feasible set of flows Ω can be defined as follows:

$$\Omega = \Big\{ u \in \mathbb{R}^{|\mathbb{P}|} : \sum_{p \in P_w} u_p = d_w,\ \forall w \in W;\ u_p \geq 0,\ \forall p \in P_w \Big\}.$$

We can write $\Omega = \Pi_{w \in W} \Omega_w$, where

$$\Omega_w = \Big\{ u \in \mathbb{R}^{|P_w|} : \sum_{p \in P_w} u_p = d_w;\ u_p \geq 0,\ \forall p \in P_w \Big\}.$$

If, the flow vector u is known, we can define the value of edge flow f_d for each edge $d \in D$ by

$$f_d = \sum_{p \in \mathbb{P}} \xi_{pd} u_p = \sum_{p \in P_w} \sum_{w \in W} \xi_{pd} u_p.$$

Here

$$\xi_{pd} = \begin{cases} 1 & \text{if } d \text{ belongs to path } p, \\ 0 & \text{otherwise.} \end{cases}$$

When all the values of edge flows are known, we can define the value of costs (expenses) for each edge $d \in D$ as follows $t_d = C_d(f_d)$, which in general depends on flows for other edges and uses some mapping C_d that is defined in the space of flows. Then we can find the value of costs for each path p as follows:

$$\mathcal{A}_p(u) = \sum_{d \in D} \xi_{pd} t_d.$$

A feasible flow vector $u^* \in \Omega$ is called the equilibrium vector if, it satisfies the following conditions:

$$\forall q \in P_w,\ u_q^* > 0 \Rightarrow \mathcal{A}_q(u^*) = \min_{p \in P_w} \mathcal{A}_p(u^*),\ \forall w \in W. \tag{4.1}$$

This means, when the traffic network is at equilibrium, among all paths of P_w, the path with traffic has the lowest cost. The problem of finding u^* satisfying (4.1) is equivalent to our VIP (see, [25, Theorem 6.1]):

$$\text{Find } u^* \in \Omega \text{ such that } \langle \mathcal{A}u^*, u - u^* \rangle \geq 0,\ \forall u \in \Omega.$$

Here $\mathcal{A}u$ is a vector in $\mathbb{R}^{|\mathbb{P}|}$ with components $\mathcal{A}_p(u)$.

For experiment, we consider a traffic network with five nodes and eight edges di $(i = 1, 2, \ldots, 8)$ as in Figure 1, and the cost function C_d is given by

$$C_d(f) = \begin{cases} \alpha_d f + \beta_d & \text{if} \quad 0 \le f \le \kappa_n, \\ \gamma_d f + \alpha_d \kappa_d + \beta_d - \gamma_d \kappa_d & \text{if} \quad f > \kappa_d. \end{cases} \tag{4.2}$$

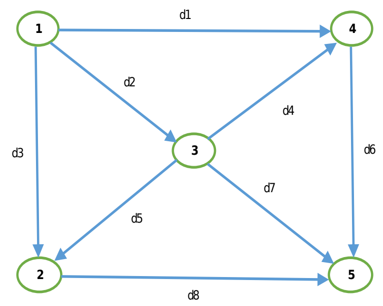

Figure 1. Traffic network with 5 nodes.

Here α_d, β_d, γ_d, κ_d are given in Table 1.

Table 1. Parameters of the cost functions in the transport equilibrium problem.

d	α_d	β_d	γ_d	κ_d
d1	1	100	10	100
d2	1.1	120	11	120
d3	0.9	80	9	80
d4	0.1	150	8	150
d5	0.1	70	11	70
d6	0.7	140	12	210
d7	1.2	150	13	150
d8	0.6	160	14	250

We can explain formula (4.2) as follows:

a) When the traffic f increases, but is still within the allowable limit of edge (see, $f \le \kappa_d$), the costs increases with a small rate α_d. We can see that β_d is the minimal cost on edge d.

b) Otherwise, when the traffic f exceeds the allowable limit κ_d, the costs increases with the huge rate γ_d (traffic jam).

Suppose that $W = \{(1, 5)\}$ and $d_{(1,5)} = 1000$. The set $P_{(1,5)}$ includes five elements: $p_1 = d_1 \to d6$, $p_2 = d3 \to d8$, $p_3 = d2 \to d7$, $p_4 = d2 \to d5 \to d8$, and $p_5 = d2 \to d4 \to d6$.

Since the cost function C_d increases, $\mathcal{A}$ is monotone. We use the aforementioned methods to compute numerically. The parameters for MPPM with stepsize rule (3.1) are $\beta = 1$, $r = 1$, $\sigma = 1$, $l = \mu = 0.5$, denoted by MPPM1, and with rule (3.23) are $\lambda_0 = 1$, $\mu = 0.5$, $p_n = (n+1)^{-1.1}$, denoted by MPPM2.

We choose $\tau_0 = 1$, $\mu = 0.5$ for VSEGM and VTEGM. Since the information on the Lipschitz constant of $\mathcal{A}$ is unknown, we also use a linesearch rule to find stepsizes for HSEGM and HEGVM. The sequence $\{\alpha_n\}$ is $\alpha_n = \frac{1}{(n+1)^p}$, with $p = 0.9$ or $p = 0.5$, for all the methods. The operator $\mathcal{F}$ is of the form $\mathcal{F}u = Qu+q$, where q is a random vector and Q is a random positive definite and symmetric matrix. The starting point u_0 is chosen in Ω. The results are shown in Figures 2 and 3. In view of these figures, we see that the method MPPM (with both the two stepsize rules) works the best when the time elapses. The number of iterations of MPPM which needs to compute is smaller than that of other methods because at each iteration it requires some tested-computations.

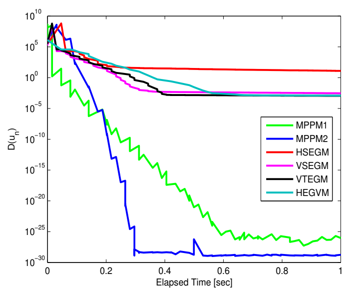

Figure 2. Methods for Example 1 with $\alpha_n = (n+1)^{-0.9}$. Number of iterations is 81, 94, 159, 254, 341, 390, respectively.

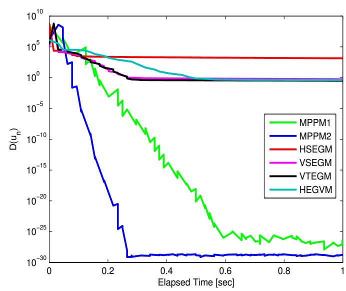

Figure 3. Methods for Example 1 with $\alpha_n = (n+1)^{-0.5}$. Number of iterations is 82, 95, 181, 263, 348, 387, respectively.

Example 2 [Bilevel Optimization Problem]. In this example, we consider a bilevel optimization problem (where we use the terminology of inner and outer levels) [34, 44]. The outer level is given by the following constraint minimization problem

$$\min \omega(u) \text{ s.t. } u \in \underset{\mathbb{R}^m}{\operatorname{Argmin}} \varphi \subset \mathbb{R}^m, \tag{4.3}$$

where ω is a strongly convex and differentiable function while $\underset{\mathbb{R}^m}{\operatorname{Argmin}} \varphi$ is the, assumed nonempty, set of minimizers of the inner level problem for a convex function φ on $\mathbb{R}^m$ of the form $\varphi(u) = h(u) + g(u)$. For experiment, we take $w(u) = \frac{1}{2}u^T Qu$ and $\varphi(u) = \frac{1}{2}\|Tu - y\|^2 + \delta_\Omega$, where $Q \in \mathbb{R}^{m\times m}$ is a certain positive definite matrix, $T \in \mathbb{R}^{m\times l}$, $y \in \mathbb{R}^l$, δ_Ω is the indicator function over the nonnegative orthant $\Omega = \{u \in \mathbb{R}^m : u \geq 0\}$ with $m = 100$, $l = 10$. Problem (4.3) is equivalent to our problem with $\mathcal{A} = T^*(T - b)$ and $\mathcal{F} = \nabla w$. The data is generated randomly. The results for this example are shown in Figures 4–5.

Example 3 [Optimal Control Problems (see, [42])]. Let $H = L^2[0, T]$ be the space of square-integrable functions on the interval $[0, T]$ with the inner product $\langle x, y\rangle = \int_0^T x(t)y(t)dt$ and the induced norm $\|x\| = \sqrt{\langle x, x\rangle}$. Let $k \in \mathbb{N}^*$ be a natural number. Consider the product Hilbert space $\mathcal{H} = H^k = H\times H\times\ldots\times H$. Let Ω be a k-dimensional box of piecewise continuous functions, given by

$$\Omega = \{u \in \mathcal{H} : u_i(t) \in [\underline{u}, \bar{u}],\ \forall t \in [0, T],\ \forall i = 1, 2, \ldots, k\}.$$

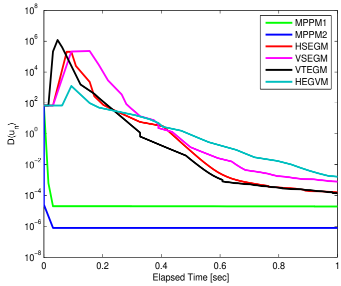

Figure 4. Methods for Example 2 with $\alpha_n = (n+1)^{-0.9}$. Number of iterations is 55, 67, 104, 143, 199, 223, respectively.

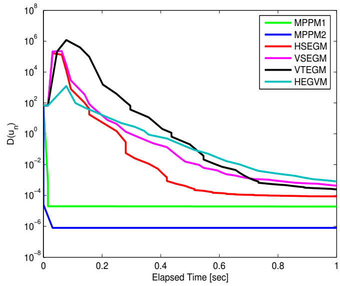

Figure 5. Methods for Example 2 with $\alpha_n = (n+1)^{-0.5}$. Number of iterations is 61, 70, 118, 164, 212, 239, respectively.

For each $u(t) \in \Omega$, let us denote by $x(t) \in H^m$ the state trajectory vector with continuous components and piecewise continuous derivatives satisfying the following system of genaral differential equations,

$$\dot{x}(t) = \mathbb{A}(t)x(t) + \mathbb{B}(t)u(t),\ x(0) = x_0,\ t \in [0, T],$$

where $\mathbb{A}(t) \in \mathbb{R}^{m\times m}$ and $\mathbb{B}(t) \in \mathbb{R}^{m\times k}$ are the matrices of continuous functions in $[0, T]$. In this example, we consider the following optimal control problem

$$\min \{f(u) : u \in \Omega\},$$

where the terminal objective function f is of the form $f(u) = \varphi(x(T))$ and φ is a differentiable convex function defined on the attainability set.

By the Pontryagin maximum principle, each pair (x^*, u^*), together with a corresponding absolutely continuous function $p^* : [0, T] \to \mathbb{R}^m$, satisfies the following system of equations:

$$\dot{x^*}(t) = \mathbb{A}(t)x^*(t) + \mathbb{B}(t)u^*(t),\ x^*(0) = x_0, \tag{4.4}$$

$$\dot{p^*}(t) = -\mathbb{A}^\intercal(t)p^*(t),\ p^*(T) = \nabla\varphi(x(T)), \tag{4.5}$$

$$0 \in \mathbb{B}^\intercal(t)p^*(t) + \mathcal{N}_\Omega(u^*(t)), \tag{4.6}$$

where $\mathcal{N}_\Omega(u^*(t))$ is the normal cone to Ω at $u^*(t)$. Let $\mathcal{A}u(t) = \mathbb{B}^\intercal(t)p(t)$, then $\mathcal{A}u(t)$ is the gradient of f (see, [23]). From the definition of the normal cone, relation (4.6) can be rewritten as our VIP:

$$\text{Find } u^* \in \Omega \text{ such that } \langle \mathcal{A}u^*, u - u^*\rangle \geq 0,\ \forall u \in \Omega.$$

In order to compute, we divide the interval $[0, T]$ by the points $t_i = ih$ $(i = 0, 1, \ldots, K)$, and discretize continuous functions by these points, where K is a chosen natural number and $h = T/K$ is the mesh size. We indentity a discretized-control ${}^K u = (u_0, u_1, \ldots, u_K)$ by its piecewise continuous extension,

$${}^K u(t) = u_i,\ \forall t \in [t_i, t_{i+1}),\ i = 0, 1, \ldots, K-1.$$

Moreover, we also identity a discretized-state ${}^{K}x = (x_0, x_1, \dots, x_K)$ with its piece-wise linear interpolation:

$$ {}^{K}x(t) = x_i + \frac{t - t_i}{h}(x_{i+1} - x_i),\ \forall t \in [t_i, t_{i+1}),\ i = 0, 1, \dots, K-1. $$

Similarly, with the co-state variable ${}^{K}p = (p_0, p_1, \dots, p_K)$ and

$$ {}^{K}\mathcal{A}u = (\mathbb{B}^{\intercal}(t_0)p_0, \mathbb{B}^{\intercal}(t_1)p_1, \dots, \mathbb{B}^{\intercal}(t_K)p_K). $$

We use the Euler method to discretize the system of ODEs (4.4)–(4.5). This means that at each iteration we need to solve the system of linear equations,

$$ \begin{aligned} &x_{i+1} = x_i + h\left[\mathbb{A}(t_i)x_i + \mathbb{B}(t_i)u_i\right], \quad x(0) = x_0, \\ &p_i = p_{i+1} + h\mathbb{A}^{\intercal}(t_i)p_{i+1}, \quad p_K = \nabla\varphi(x_K). \end{aligned} $$

Now, we consider the following optimal control problem (see, Example 1.2 [35])

$$ \begin{aligned} \text{minimize} \quad & x_1(1) \\ \text{subject to} \quad & \dot{x}_j(t) = s_j x_{j+1}(t) + u(t),\ s_j = -2(m-j+1),\ j = 1, 2, \cdots, m, \\ & \dot{x}_{m+1}(t) = u(t),\ t \in [0, 1], \\ & x(0) = 0, \quad u(t) \in [-1, 1], \end{aligned} $$

where m is a natural number. Observe that the matrix-valued functions $\mathbb{A}(t)$ and $\mathbb{B}(t)$ are given by

$$ \mathbb{A}(t) = \begin{pmatrix} 0 & s_1 & 0 & \dots & 0 & 0 \\ 0 & 0 & s_2 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \ddots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 & s_m \\ 0 & 0 & 0 & \dots & 0 & 0 \end{pmatrix}_{m+1,m+1}, \quad \mathbb{B}(t) = \begin{pmatrix} 1 \\ 1 \\ 1 \\ \vdots \\ 1 \\ 1 \end{pmatrix}_{m+1,1}. $$

We take the mesh size $h = T/K$ with $K = 256$. The starting point u_0 is generated randomly in Ω and $\mathcal{F}u = 0.5u - u_0$. The methods are performed for the sequence of control parameters $\alpha_n = (n+1)^{-0.9}$. Other parameters are chosen as in Example 1. The numerical results are illustrated in Figures 6–9 for $m \in \{1, 2, 3, 4\}$. These results also demonstrate that the method MPPM has competitive advantages over others.

5 Conclusions

In this paper, we have proposed a new method of multi-parameter form for solving a monotone and Lipschitz variational inequality problem in a Hilbert space. The method is constructed around the projection method incorporated with regularization terms. The strong convergence of the method has been established under appropriate conditions imposed on control parameters. The obtained solution from the method is a solution of a bilevel variational inequality problem. The computational efficiency of the method over some existing ones is illustrated by numerical experiments on transportation equilibrium, bilevel optimization, and optimal control problems.

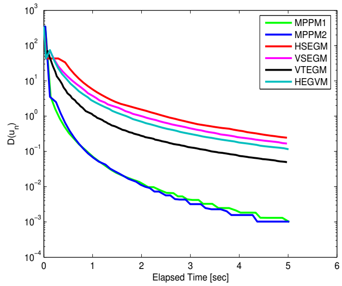

Figure 6. Methods for Example 3 with $m = 1$. Number of iterations is 48, 59, 48, 58, 107, 57, respectively.

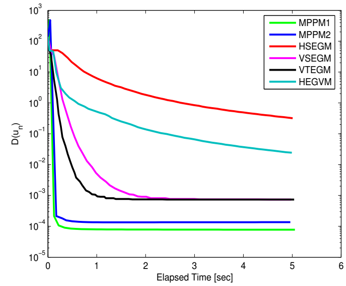

Figure 7. Methods for Example 3 with $m = 2$. Number of iterations is 49, 57, 56, 52, 94, 54, respectively.

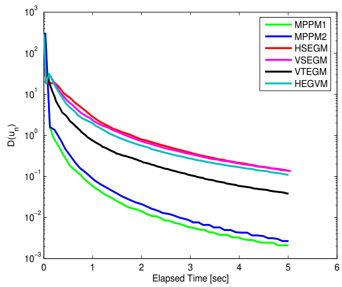

Figure 8. Methods for Example 3 with $m = 3$. Number of iterations is 55, 58, 56, 56, 108, 55, respectively.

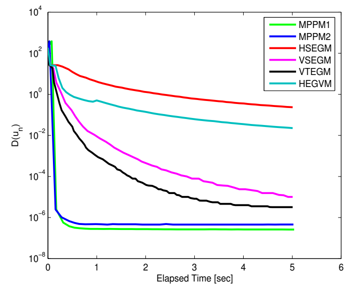

Figure 9. Methods for Example 3 with $m = 4$. Number of iterations is 43, 52, 56, 50, 89, 56, respectively.

Acknowledgements

The authors would like to thank the Associate Editor and the anonymous referee for their valuable comments and suggestions which helped us very much in improving the original version of this paper. The research of the first two authors is funded by Vietnam National Foundation for Science and Technology Development (NAFOSTED) under Grant Number 101.01-2020.06.

References

[1] M. Abbas and H. Iqbal. Two inertial extragradient viscosity algorithms for solving variational inequality and fixed point problems. *Journal of Nonlinear and Variational Analysis*, **4**(3):377–398, 2020. https://doi.org/10.23952/jnva.4.2020.3.04.

[2] Ya.I. Alber and I. Ryazantseva. *Nonlinear Ill-posed Problems of Monotone Type*. Springer, Dordrecht, 2006. https://doi.org/10.1007/1-4020-4396-1.

[3] H.H. Bauschke and P.L. Combettes. *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. Springer, New York, 2011.

[4] X. Cai, G. Gu and B. He. On the $o(1/t)$ convergence rate of the projection and contraction methods for variational inequalities with Lipschitz continuous monotone operators. *Computational Optimization and Applications*, **57**:339–363, 2014. https://doi.org/10.1007/s10589-013-9599-7.

[5] Y. Censor, A. Gibali and S. Reich. Strong convergence of subgradient extragradient methods for the variational inequality problem in Hilbert space. *Optimization Methods and Software*, **26**(4-5):827–845, 2011. https://doi.org/10.1080/10556788.2010.551536.

[6] Y. Censor, A. Gibali and S. Reich. The subgradient extragradient method for solving variational inequalities in Hilbert space. *Journal of Optimization Theory and Applications*, **148**:318–335, 2011. https://doi.org/10.1007/s10957-010-9757-3.

[7] Y. Censor, A. Gibali and S. Reich. Extensions of Korpelevich's extragradient method for the variational inequality problem in Euclidean space. *Optimization*, **61**(9):1119–1132, 2012. https://doi.org/10.1080/02331934.2010.539689.

[8] Q.L. Dong, Y.J. Cho and T.M. Rassias. The projection and contraction methods for finding common solutions to variational inequality problems. *Optim. Lett.*, **12**:1871–1896, 2018. https://doi.org/10.1007/s11590-017-1210-1.

[9] Q.L. Dong, Y.J. Cho, L.L. Zhong and T.M. Rassias. Inertial projection and contraction algorithms for variational inequalities. *Journal of Global Optimization*, **70**:687–704, 2018. https://doi.org/10.1007/s10898-017-0506-0.

[10] Q.L. Dong, A. Gibali, D. Jiang and S.H. Ke. Convergence of projection and contraction algorithms with outer perturbations and their applications to sparse signals recovery. *Journal of Fixed Point Theory and Applications*, **20**(16), 2018. https://doi.org/10.1007/s11784-018-0501-1.

[11] Q.L. Dong, D. Jiang and A. Gibali. A modified subgradient extragradient method for solving the variational inequality problem. *Numerical Algorithms*, **79**:927–940, 2018. https://doi.org/10.1007/s11075-017-0467-x.

[12] N.T. Vinh D.V. Thong and Y.J. Cho. New strong convergence theorem of the inertial projection and contraction method for variational inequality problems. *Numerical Algorithms*, **84**:285–305, 2020. https://doi.org/10.1007/s11075-019-00755-1.

[13] F. Facchinei and J.S. Pang. *Finite-Dimensional Variational Inequalities and Complementarity Problems.* Springer, Berlin, 2003. https://doi.org/10.1007/b97544.

[14] R. Glowinski. *Numerical Methods for Nonlinear Variational Problems*. Springer, New York, 1984. https://doi.org/10.1007/978-3-662-12613-4.

[15] K. Goebel and S. Reich. *Uniform Convexity, Hyperbolic Geometry, and Nonexpansive Mappings*. Marcel Dekker, New York and Basel, 1984.

[16] B.S. He. A class of projection and contraction methods for monotone variational inequalities. *Applied Mathematics and Optimization*, **35**:69–76, 1997. https://doi.org/10.1007/BF02683320.

[17] B.S. He and L.-Z. Liao. Improvements of some projection methods for monotone nonlinear variational inequalities. *Journal of Optimization Theory and Applications*, **112**:111–128, 2002. https://doi.org/10.1023/A:1013096613105.

[18] B.S. He, X.M. Yuan and J.J.Z. Zhang. Comparison of two kinds of prediction-correction methods for monotone variational inequalities. *Computational Optimization and Applications*, **27**:247–267, 2004. https://doi.org/10.1023/B:COAP.0000013058.17185.90.

[19] D.V. Hieu, Y.J. Cho, Y.B. Xiao and P. Kumam. Modified extragradient method for pseudomonotone variational inequalities in infinite dimensional Hilbert spaces. *Vietnam Journal of Mathematics*, **49**:1165–1183, 2021. https://doi.org/10.1007/s10013-020-00447-7.

[20] D.V. Hieu, L.D. Muu, P.K. Quy and H.N. Duong. Regularization extragradient methods for equilibrium programming in Hilbert spaces. *Optimization*, **0**(0):1–31, 2021. https://doi.org/10.1080/02331934.2021.1873988.

[21] D.V. Hieu, J.J. Strodiot and L.D. Muu. An explicit extragradient algorithm for solving variational inequalities. *Journal of Optimization Theory and Applications*, **185**:476–503, 2020. https://doi.org/10.1007/s10957-020-01661-6.

[22] I. Hlavacek, J. Haslinger, J. Necas and J. Lovicek. *Solution of Variational Inequalities in Mechanics.* Springer, New York, 1982. https://doi.org/10.1007/978-1-4612-1048-1.

[23] E.V. Khoroshilova. Extragradient-type method for optimal control problem with linear constraints and convex objective function. *Optimization Letters*, **7**:1193–1214, 2013. https://doi.org/10.1007/s11590-012-0496-2.

[24] D. Kinderlehrer and G. Stampacchia. *An Introduction to Variational Inequalities and Their Applications.* Academic Press, New York, 1980.

[25] I.V. Konnov. *Equilibrium Models and Variational Inequalities.* Amsterdam, Elsevier, 2007.

[26] G.M. Korpelevich. The extragradient method for finding saddle points and other problems. *Ekonomikai Matematicheskie Metody*, **12**:747–756, 1976.

[27] R. Kraikaew and S. Saejung. Strong convergence of the Halpern subgradient extragradient method for solving variational inequalities in Hilbert spaces. *Journal of Optimization Theory and Applications*, **163**:399–412, 2014. https://doi.org/10.1007/s10957-013-0494-2.

[28] L. Lampariello, C. Neumann, J. M. Ricci, S. Sagratella and O. Stein. An explicit Tikhonov algorithm for nested variational inequalities. *Computational Optimization and Applications*, **77**:335–350, 2020. https://doi.org/10.1007/s10589-020-00210-1.

[29] P.E. Maingé. A hybrid extragradient-viscosity method for monotone operators and fixed point problems. *SIAM Journal on Control and Optimization*, **47**(3):1499–1515, 2008. https://doi.org/10.1137/060675319.

[30] Y.V. Malitsky. Projected reflected gradient methods for monotone variational inequalities. *SIAM Journal on Optimization*, **25**(1):502–520, 2015. https://doi.org/10.1137/14097238X.

[31] Y.V. Malitsky. Golden ratio algorithms for variational inequalities. *Mathematical Programming*, **184**:383–410, 2020. https://doi.org/10.1007/s10107-019-01416-w.

[32] P.D. Panagiotopoulos. *Inequality Problems in Mechanics and Applications.* Boston, Birkhäuser, 1985. https://doi.org/10.1007/978-1-4612-5152-1.

[33] L.D. Popov. A modification of the Arrow-Hurwicz method for searching for saddle points. *Mathematical notes of the Academy of Sciences of the USSR*, **28**:845–848, 1980. https://doi.org/10.1007/BF01141092.

[34] S. Sabach and S. Shtern. A first-order method for solving convex bilevel optimization problems. *SIAM Journal on Optimization*, **27**(2):640–660, 2017. https://doi.org/10.1137/16M105592X.

[35] M. Seydenschwanz. Convergence results for the discrete regularization of linear-quadratic control problems with bang-bang solutions. *Computational Optimization and Applications*, **629**:731–760, 2015. https://doi.org/10.1007/s10589-015-9730-z.

[36] M.V. Solodov and P. Tseng. Modified projection-type methods for monotone variational inequalities. *SIAM Journal on Control and Optimization*, **34**(5):1814–1830, 1996. https://doi.org/10.1137/S0363012994268655.

[37] D.F. Sun. A class of iterative methods for solving nonlinear projection equations. *Journal of Optimization Theory and Applications*, **91**:123–140, 1996. https://doi.org/10.1007/BF02192286.

[38] B. Tan and S. Xu. Strong convergence of two inertial projection algorithms in Hilbert spaces. *Journal of Applied and Numerical Optimization*, **2**(2):171–186, 2020. https://doi.org/10.23952/jano.2.2020.2.04.

[39] D.V. Thong and D.V. Hieu. Some extragradient-viscosity algorithms for solving variational inequality problems and fixed point problems. *Numerical Algorithms*, **82**:761–789, 2019. https://doi.org/10.1007/s11075-018-0626-8.

[40] M. Tian and G. Xu. Inertial modified Tseng's extragradient algorithms for solving monotone variational inequalities and fixed point problems. *Journal of Nonlinear Functional Analysis*, **2020**:1–19, 2020. https://doi.org/10.23952/jnfa.2020.35.

[41] P. Tseng. A modified forward-backward splitting method for maximal monotone mappings. *SIAM Journal on Control and Optimization*, **38**(2):431–446, 2000. https://doi.org/10.1137/S0363012998338806.

[42] P.T. Vuong and Y. Shehu. Convergence of an extragradient-type method for variational inequality with applications to optimal control problems. *Numerical Algorithms*, **81**:269–291, 2019. https://doi.org/10.1007/s11075-018-0547-6.

[43] H.K. Xu. Iterative algorithms for nonlinear operators. *Journal of the London Mathematical Society*, **66**(1):240–256, 2002. https://doi.org/10.1112/S0024610702003332.

[44] P.T. Vuong Y. Shehu and A. Zemkoho. An inertial extrapolation method for convex simple bilevel optimization. *Optimization Methods and Software*, **36**(1):1–19, 2021. https://doi.org/10.1080/10556788.2019.1619729.