J Optim Theory Appl (2017) 172:914–928
DOI 10.1007/s10957-016-1051-6

On the Iteration Complexity of Some Projection Methods for Monotone Linear Variational Inequalities

Caihua Chen[1] · **Xiaoling Fu**[2] · **Bingsheng He**[3,4] · **Xiaoming Yuan**[5]

Received: 9 February 2015 / Accepted: 26 December 2016 / Published online: 9 January 2017

Abstract Projection-type methods are important for solving monotone linear variational inequalities. In this paper, we analyze the iteration complexity of two projection methods and accordingly establish their worst-case sublinear convergence rates measured by the iteration complexity in both the ergodic and nonergodic senses. Our analysis does not require any error bound condition or the boundedness of the feasible set, and it is scalable to other methods of the same kind.

Keywords Linear variational inequality · Projection methods · Convergence rate · Iteration complexity

Communicated by Patrice Marcotte.

✉ Xiaoming Yuan
xmyuan@hkbu.edu.hk

Caihua Chen
chchen@nju.edu.cn

Xiaoling Fu
fuxlnju@hotmail.com

Bingsheng He
hebs@sustc.edu.cn; hebma@nju.edu.cn

1 International Center of Management Science and Engineering, School of Management and Engineering, Nanjing University, Nanjing, China

2 School of Economics and Management, Southeast University, Nanjing, China

3 Department of Mathematics, South University of Science and Technology of China, Shenzhen 518055, China

4 Department of Mathematics, Nanjing University, Nanjing 210093, China

5 Department of Mathematics, Hong Kong Baptist University, Kowloon Tong, Hong Kong, China

Mathematics Subject Classification 90C33 · 90C25 · 90C30

1 Introduction

The variational inequality (VI) problem was originally introduced by Hartman and Stampacchia in [1] for the study of partial differential equations; it has received much attention from many researchers in different areas such as optimal control, economics, game theory and transportation science. In the literature, projection-type methods have been well studied for variational inequalities; they are particularly efficient for some special scenarios where the involved projection can be easily computed. We refer to [2] for a survey on projection methods in the variational inequality context.

In this paper, we consider two projection methods for monotone linear variational inequalities (LVI). The first one was proposed by Solodov and Tseng in [3]. The efficiency of this method has been numerically verified in the literature, and its iterative sequence has been proved to be strictly contractive with respect to the solution set of LVI. Hence, its convergence follows from the standard analytic framework of contraction methods; see, e.g., Blum and Oettli in [4]. A special case of this method was proposed by He in [5], and its linear convergence rate was also proved for the linear complementarity problem therein. Another representative projection method was proposed by He in [6,7], whose step sizes are the same as those defined in [3], but it utilizes different search directions and requires the computation of two projections at each iteration.

Our main purpose is to analyze the convergence rates for both the mentioned projection-type methods under mild assumptions. Indeed, if certain error bound condition is assumed, the strict contraction property enables us to establish some convergence rates in asymptotical sense immediately. In general, however, it is not easy to verify error bound conditions even for LVI. We thus consider the possibility of deriving the convergence rates without any error bound conditions. More specifically, we want to derive the worst-case sublinear convergence rate measured by the iteration complexity. This kind of convergence rate analysis based on the iteration complexity traces back to [8], and it has received much attention from the literature. Furthermore, our convergence rate analysis does not need to assume the boundedness of the feasible set which is usually required by iteration complexity analysis of projection methods for nonlinear variational inequalities; see the work [9] for the extragradient method studied in [10]. Finally, we emphasize that we only focus on LVI in this paper and do not discuss the iteration complexity analysis of projection methods for nonlinear variational inequalities. We refer to, e.g., [9,11], for some insightful discussions in this regard.

The rest of this paper is organized as follows. In Sect. 2, we mathematically state the LVI model and the projection methods to be discussed. Then, in Sect. 3, we summarize some preliminaries that are useful for further analysis. We prove two lemmas in Sect. 4 that are crucial for establishing the main convergence rate results. In Sect. 5, we present the main convergence rate results. Finally, some conclusions are drawn in Sect. 6.

2 Model and Projection Methods

We first present the LVI model and two projection methods to be discussed. Let Ω be a closed convex subset of $\mathbb{R}^n$, $M \in \mathbb{R}^{n\times n}$ and $q \in \mathbb{R}^n$. The linear variational inequality problem, denoted by LVI(Ω, M, q), is to find a vector $u^* \in \Omega$ such that

$$\mathrm{LVI}(\Omega, M, q) \qquad \langle u - u^*, Mu^* + q\rangle \geq 0, \quad \forall u \in \Omega. \tag{1}$$

We consider the case where the matrix M is positive semi-definite (but not necessarily symmetric). Moreover, the solution set of (1), denoted by Ω^*, is assumed to be nonempty.

It is well known (see, e.g., [12, p. 267]) that u^* is a solution point of (1) if and only if it satisfies the following projection equation

$$u^* = P_{[\Omega,G]}[u^* - G^{-1}(Mu^* + q)], \tag{2}$$

where $G \in \mathbb{R}^{n\times n}$ is a symmetric positive definite matrix, and $P_{[\Omega,G]}(\cdot)$ denotes the projection onto Ω with respect to the G-norm:

$$P_{[\Omega,G]}(v) = \mathrm{argmin}\{\|u - v\|_G \,|\, u \in \Omega\},$$

and $\|u\|_G = \sqrt{u^T G u}$ for any $u \in \mathbb{R}^n$. When $G = I$, we simply use the notation $P_\Omega(\cdot)$ for $P_{[\Omega,I]}(\cdot)$. Moreover, for given $u \in \mathbb{R}^n$, we denote

$$\tilde{u} = P_\Omega[u - (Mu + q)].$$

Hence, we have

$$\tilde{u}^k = P_\Omega[u^k - (Mu^k + q)]. \tag{3}$$

We further use the notation

$$e(u^k) := u^k - \tilde{u}^k. \tag{4}$$

It follows from (2) that u is a solution point of LVI(Ω, M, q) if and only if $u = \tilde{u}$. Then, naturally, the projection equation residual $\|e(u^k)\|^2$ can be used to measure the accuracy of an iterate u^k to a solution point of LVI(Ω, M, q).

Indeed, the projection equation characterization (2) for LVI(Ω, M, q) is the basis of many algorithms in the literature, including the projection-type methods under our discussion, see, e.g., [3,5,6] to just mention a few. Because of their easiness in implementation, modest demand on storage and relatively fast convergence, projection-type methods are particularly efficient for the special scenario where the set Ω in (1) is simple in sense of that the projection onto it can be easily computed. The first method to be discussed is the projection method proposed in [5] and later generalized in [3] to the G-norm (see Algorithm 2.1 therein). More specifically, its iterative scheme is

$$\text{(Algorithm-I)} \quad u^{k+1} = u^k - \gamma\alpha_k^* G^{-1}(I + M^T)(u^k - \tilde{u}^k), \tag{5}$$

where $\gamma \in]0, 2[$ is a relaxation factor and the step size α_k^* is determined by

$$\alpha_k^* = \frac{\|u^k - \tilde{u}^k\|^2}{\|G^{-1}(I + M^T)(u^k - \tilde{u}^k)\|_G^2}. \tag{6}$$

Obviously, for the step size α_k^* defined in (6), we have

$$\alpha_k^* \geq \frac{1}{\|(I + M)G^{-1}(I + M)^T\|_2} := \alpha_{\min}, \tag{7}$$

where $\|\cdot\|_2$ denotes the spectral norm of a matrix. Therefore, the step size sequence of Algorithm-I is bounded away from zero; this is indeed an important property for both theoretically ensuring the convergence and numerically resulting in fast convergence for Algorithm-I. More specifically, as proved in [3], the sequence $\{u^k\}$ generated by Algorithm-I satisfies the inequality

$$\|u^{k+1} - u^*\|_G^2 \leq \|u^k - u^*\|_G^2 - \gamma(2 - \gamma)\alpha_k^* \|u^k - \tilde{u}^k\|^2, \tag{8}$$

where u^* is an arbitrary solution point of LVI(Ω, M, q). Recall the fact that $\|u^k - \tilde{u}^k\|^2 = 0$ if and only if u^k is a solution point of LVI(Ω, M, q). Thus, together with the property (7), inequality (8) essentially means that the sequence $\{u^k\}$ is strictly contractive with respect to the solution set of LVI(Ω, M, q).

In addition to Algorithm-I in (5), we consider another projection method

$$\text{(Algorithm-II)} \quad u^{k+1} = P_{[\Omega, G]}\left\{u^k - \gamma\alpha_k^* G^{-1}\left[\left(Mu^k + q\right) + M^T\left(u^k - \tilde{u}^k\right)\right]\right\}, \tag{9}$$

where the step size α_k^* is also defined in (6). As for Algorithm-I, we will show later (see Corollary 4.1) that the sequence $\{u^k\}$ generated by Algorithm-II also satisfies the property (8) and thus its convergence is ensured. The special case of Algorithm-II with $\Omega = \mathbb{R}_+^n$ and $G = I$ can be found in [7] and mentioned in [6], and its convergence proof can be found in [13]. Also, Algorithm-II differs from Algorithm 2.3 in [3] in that its step size is determined by (6) and thus it is bounded away from zero, while the latter may tend to zero, see (2.14) on [3, p. 1821].

In this paper, we establish the worst-case $O(1/t)$ convergence rate measured by the iteration complexity for Algorithm-I and Algorithm-II, where t is the iteration counter. More specifically, we will show that for a given $\epsilon > 0$, by implementing either Algorithm-I or Algorithm-II, we need at most $O(1/\epsilon)$ iterations to find an approximated solution point of LVI(Ω, M, q) with an accuracy of ϵ. In our analysis, we measure the accuracy of an approximated solution point of LVI(Ω, M, q) in two ways. The first one is the restricted merit function of LVI(Ω, M, q):

$$\sup_{u \in \mathcal{D}(v)} \left\{\langle v - u, Mu + q\rangle\right\},$$

where $\mathcal{D}(v)$ is some subset of Ω. The second one is simply the residual of the projection equation $\|e(u)\|^2$ as mentioned before. More details of the definition of an ϵ-approximated solution point of LVI(Ω, M, q) can be found in Sect. 3.2.

3 Preliminaries

In this section we summarize some preliminaries useful for further analysis. Throughout, the following notation is used. We use u^* to denote a fixed but arbitrary point in the solution set Ω^* of LVI. A superscript such as in u^k refers to a specific vector and usually denotes an iteration index. For any real matrix M and vector v, we denote their transposes by M^T and v^T, respectively. The Euclidean norm is denoted by $\|\cdot\|$.

3.1 Some Inequalities

We first recall several inequalities which will be frequently used in the upcoming analysis. First, since

$$P_{[\Omega,G]}(v) = \operatorname{argmin}\left\{\frac{1}{2}\|u - v\|_G^2 \mid u \in \Omega\right\},$$

we have

$$\langle v - P_{[\Omega,G]}(v), G(u - P_{[\Omega,G]}(v)\rangle \le 0, \quad \forall\, v \in \mathbb{R}^n, \forall\, u \in \Omega. \tag{10}$$

Let u^* be any fixed solution point of LVI(Ω, M, q). Since $\tilde{u}^k \in \Omega$, it follows from (1) that

$$\langle \tilde{u}^k - u^*, Mu^* + q\rangle \ge 0, \quad \forall\, u^* \in \Omega^*.$$

Setting $v = u^k - \left(Mu^k + q\right)$, $G = I$ and $u = u^*$ in (10), because of the notation $\tilde{u}^k$, we have

$$\langle \tilde{u}^k - u^*, [u^k - (Mu^k + q)] - \tilde{u}^k\rangle \ge 0, \quad \forall\, u^* \in \Omega^*.$$

Adding the last two inequalities together, we obtain

$$\langle \tilde{u}^k - u^*, (u^k - \tilde{u}^k) - M(u^k - u^*)\rangle \ge 0, \quad \forall\, u^* \in \Omega^*,$$

and consequently

$$\langle u^k - u^*, (I + M^T)(u^k - \tilde{u}^k)\rangle \ge \|u^k - \tilde{u}^k\|^2, \quad \forall\, u^* \in \Omega^*. \tag{11}$$

3.2 An ϵ-Approximated Solution Point of LVI(Ω, M, q)

To estimate the worst-case convergence rates measured by the iteration complexity for Algorithm-I or Algorithm-II, we need to precisely define an ϵ-approximated solution of LVI(Ω, M, q). We will consider the following two definitions, which are based on the variational inequality characterization and projection equation residual, respectively.

First, according to (2.3.2) in [2, p. 159], we know that Ω^* is convex and it can be characterized as follows:

$$\Omega^* = \bigcap_{u \in \Omega} \big\{ v \in \Omega : \langle u - v, Mu + q \rangle \geq 0 \big\}.$$

Therefore, motivated by [14], we call $v \in \Omega$ an ϵ-approximated solution point of LVI(Ω, M, q) in sense of the variational inequality characterization if it satisfies

$$v \in \Omega \quad \text{and} \quad \inf_{u \in \mathcal{D}(v)} \big\{ \langle u - v, Mu + q \rangle \big\} \geq -\epsilon,$$

where

$$\mathcal{D}(v) = \{ u \in \Omega \mid \|u - v\|_G \leq 1 \}.$$

Later, we will show that for given $\epsilon > 0$, after at most $O(1/\epsilon)$ iterations, both Algorithm-I and Algorithm-II can find v such that

$$v \in \Omega \quad \text{and} \quad \sup_{u \in \mathcal{D}(v)} \big\{ \langle v - u, Mu + q \rangle \big\} \leq \epsilon. \tag{12}$$

For the other definition, as mentioned, with $e(u)$ defined in (4), $\|e(u)\|^2$ is a measure of the distance between the iterate u and the solution set Ω^*. We thus call v an ϵ-approximated solution point of LVI(Ω, M, q) in sense of the projection equation residual if $\|e(v)\|^2 \leq \epsilon$.

4 Two Lemmas

We consider Algorithm-I and Algorithm-II simultaneously because of the similarity in their convergence rate analysis and they can be presented in a unified framework. In this section, we show that the sequences generated by both Algorithm-I and Algorithm-II satisfy one common inequality, which is indeed the key for estimating their convergence rates measured by the iteration complexity. For notation simplicity, we define

$$q_k(\gamma) = \gamma(2 - \gamma)\alpha_k^* \|u^k - \tilde{u}^k\|^2, \tag{13}$$

where α_k^* is given by (6). Moreover, let us use the notation

$$D = M + M^T,$$

where M is the matrix in (1).

In the following, we show that the sequence $\{u^k\}$ generated by either Algorithm-I or Algorithm-II satisfies the inequality

$$\gamma\alpha_k^*\langle u-\tilde{u}^k, Mu+q\rangle \geq \frac{1}{2}\big(\|u-u^{k+1}\|_G^2-\|u-u^k\|_G^2\big)+\frac{1}{2}q_k(\gamma),\ \ \forall\, u\in\Omega, \tag{14}$$

where $q_k(\gamma)$ is defined in (13). We present this conclusion in two lemmas.

Lemma 4.1 *For given $u^k\in\mathbb{R}^n$, let $\tilde{u}^k$ be defined by* (3) *and the new iterate u^{k+1} be generated by Algorithm-I in* (5)*. Then, the assertion* (14) is satisfied.

Proof Setting $v=u^k-(Mu^k+q)$ in (10), and using $\tilde{u}^k=P_\Omega[u^k-(Mu^k+q)]$, we have

$$\langle u-\tilde{u}^k, (Mu^k+q)-(u^k-\tilde{u}^k)\rangle\geq 0,\quad \forall\, u\in\Omega.$$

This inequality can be rewritten as

$$\langle u-\tilde{u}^k, (Mu+q)-M(u-\tilde{u}^k)+(M+M^T)(u^k-\tilde{u}^k)-(I+M^T)(u^k-\tilde{u}^k)\rangle\geq 0.$$

Therefore, using the notation $M+M^T=D$ and the Cauchy–Schwarz inequality, we have

$$\begin{aligned}&\langle u-\tilde{u}^k, Mu+q\rangle\\ &\quad\geq \langle u-\tilde{u}^k, M(u-\tilde{u}^k)-(M+M^T)(u^k-\tilde{u}^k)+(I+M^T)(u^k-\tilde{u}^k)\rangle\\ &\quad=\langle u-\tilde{u}^k, (I+M^T)(u^k-\tilde{u}^k)\rangle+\frac{1}{2}\|u-\tilde{u}^k\|_D^2-\langle u-\tilde{u}^k, D(u^k-\tilde{u}^k)\rangle\\ &\quad\geq \langle u-\tilde{u}^k, (I+M^T)(u^k-\tilde{u}^k)\rangle-\frac{1}{2}\|u^k-\tilde{u}^k\|_D^2.\end{aligned}$$

Moreover, it follows from (5) that

$$\gamma\alpha_k^*(I+M^T)(u^k-\tilde{u}^k)=G(u^k-u^{k+1}).$$

Thus, we obtain

$$\gamma\alpha_k^*\langle u-\tilde{u}^k, Mu+q\rangle\geq\langle u-\tilde{u}^k, G(u^k-u^{k+1})\rangle-\frac{\gamma\alpha_k^*}{2}\|u^k-\tilde{u}^k\|_D^2. \tag{15}$$

For the crossed term in the right-hand side of (15): $\langle u-\tilde{u}^k, G(u^k-u^{k+1})\rangle$, it follows from the identity

$$\langle a-b, G(c-d)\rangle=\frac{1}{2}\left(\|a-d\|_G^2-\|a-c\|_G^2\right)+\frac{1}{2}\left(\|c-b\|_G^2-\|d-b\|_G^2\right)$$

that

$$
\begin{aligned}
&\langle u-\tilde{u}^k, G(u^k-u^{k+1})\rangle \\
&\quad= \frac{1}{2}\left(\|u-u^{k+1}\|_G^2-\|u-u^k\|_G^2\right)+\frac{1}{2}\left(\|u^k-\tilde{u}^k\|_G^2-\|u^{k+1}-\tilde{u}^k\|_G^2\right). \quad (16)
\end{aligned}
$$

Now, we treat the second part of the right-hand side of (16). Using (5), we get

$$
\begin{aligned}
&\|u^k-\tilde{u}^k\|_G^2-\|u^{k+1}-\tilde{u}^k\|_G^2 \\
&\quad= \|u^k-\tilde{u}^k\|_G^2-\|(u^k-\tilde{u}^k)-\gamma\alpha_k^* G^{-1}(I+M^T)(u^k-\tilde{u}^k)\|_G^2 \\
&\quad= 2\gamma\alpha_k^*\langle u^k-\tilde{u}^k, (I+M^T)(u^k-\tilde{u}^k)\rangle-\left(\gamma\alpha_k^*\right)^2\|G^{-1}(I+M^T)(u^k-\tilde{u}^k)\|_G^2 \\
&\quad= 2\gamma\alpha_k^*\|u^k-\tilde{u}^k\|^2+\gamma\alpha_k^*\|u^k-\tilde{u}^k\|_D^2 \\
&\qquad -(\gamma\alpha_k^*)2\|G^{-1}(I+M^T)(u^k-\tilde{u}^k)\|_G^2. \quad (17)
\end{aligned}
$$

Recall the definition of α_k^* in (6). It follows from (17) that

$$
\left(\gamma\alpha_k^*\right)^2\|G^{-1}(I+M^T)(u^k-\tilde{u}^k)\|_G^2=\gamma^2\alpha_k^*\|u^k-\tilde{u}^k\|^2,
$$

and consequently,

$$
\|u^k-\tilde{u}^k\|_G^2-\|u^{k+1}-\tilde{u}^k\|_G^2=\gamma(2-\gamma)\alpha_k^*\|u^k-\tilde{u}^k\|^2+\gamma\alpha_k^*\|u^k-\tilde{u}^k\|_D^2.
$$

Substituting it into the right-hand side of (16) and using the definition of $q_k(\gamma)$, we obtain

$$
\begin{aligned}
&\langle u-\tilde{u}^k, G(u^k-u^{k+1})\rangle \\
&\quad= \frac{1}{2}\left(\|u-u^{k+1}\|_G^2-\|u-u^k\|_G^2\right)+\frac{1}{2}q_k(\gamma)+\frac{\gamma\alpha_k^*}{2}\|u^k-\tilde{u}^k\|_D^2. \quad (18)
\end{aligned}
$$

Adding (15) and (18) together, we get the assertion (14) and the theorem is proved. □

Then, we prove the assertion (14) for Algorithm-II in the following lemma.

Lemma 4.2 *For given $u^k\in\mathbb{R}^n$, let $\tilde{u}^k$ be defined by* (3) *and the new iterate u^{k+1} be generated by Algorithm-II* (9). *Then, the assertion* (14) is satisfied.

Proof It follows from Cauchy–Schwarz inequality that

$$
\begin{aligned}
&\langle u-\tilde{u}^k, Mu+q\rangle-\langle u-\tilde{u}^k, (Mu^k+q)+M^T(u^k-\tilde{u}^k)\rangle \\
&\quad= \langle u-\tilde{u}^k, M(u-u^k)-M^T(u^k-\tilde{u}^k)\rangle \\
&\quad= \langle u-\tilde{u}^k, M(u-\tilde{u}^k)-(M+M^T)(u^k-\tilde{u}^k)\rangle \\
&\quad= \frac{1}{2}\|u-\tilde{u}^k\|_D^2-\langle u-\tilde{u}^k, D(u^k-\tilde{u}^k)\rangle \\
&\quad\geq -\frac{1}{2}\|u^k-\tilde{u}^k\|_D^2.
\end{aligned}
$$

Consequently, we obtain

$$\begin{aligned}&\gamma\alpha_k^*\langle u-\tilde{u}^k, Mu+q\rangle\\&\geq \left\langle u-\tilde{u}^k, \gamma\alpha_k^*[(Mu^k+q)+M^T(u^k-\tilde{u}^k)]\right\rangle - \frac{\gamma\alpha_k^*}{2}\|u^k-\tilde{u}^k\|_D^2.\end{aligned} \tag{19}$$

Now we investigate the first term in the right-hand side of (19) and divide it into the following two terms, namely

$$\left\langle u^{k+1}-\tilde{u}^k, \gamma\alpha_k^*[(Mu^k+q)+M^T(u^k-\tilde{u}^k)]\right\rangle \tag{20a}$$

and

$$\left\langle u-u^{k+1}, \gamma\alpha_k^*[(Mu^k+q)+M^T(u^k-\tilde{u}^k)]\right\rangle. \tag{20b}$$

First, we deal with term (20a). Let us set $v=u^k-(Mu^k+q)$ in (10). Since $\tilde{u}^k = P_\Omega[u^k-(Mu^k+q)]$ and $u^{k+1}\in\Omega$, it follows that

$$\langle u^{k+1}-\tilde{u}^k, Mu^k+q\rangle \geq \langle u^{k+1}-\tilde{u}^k, u^k-\tilde{u}^k\rangle.$$

Adding the term $\langle u^{k+1}-\tilde{u}^k, M^T(u^k-\tilde{u}^k)\rangle$ to both sides in the above inequality, we obtain

$$\langle u-\tilde{u}^k, (Mu^k+q)+M^T(u^k-\tilde{u}^k)\rangle \geq \langle u-\tilde{u}^k, (I+M^T)(u^k-\tilde{u}^k)\rangle,$$

and it follows that

$$\begin{aligned}&\left\langle u^{k+1}-\tilde{u}^k, \gamma\alpha_k^*[(Mu^k+q)+M^T(u^k-\tilde{u}^k)]\right\rangle\\&\geq \gamma\alpha_k^*\langle u^{k+1}-\tilde{u}^k, (I+M^T)(u^k-\tilde{u}^k)\rangle\\&=\gamma\alpha_k^*\langle u^k-\tilde{u}^k, (I+M^T)(u^k-\tilde{u}^k)\rangle - \gamma\alpha_k^*\langle u^k-u^{k+1}, (I+M^T)(u^k-\tilde{u}^k)\rangle\\&\geq \gamma\alpha_k^*\|u^k-\tilde{u}^k\|^2 + \frac{\gamma\alpha_k^*}{2}\|u^k-\tilde{u}^k\|_D^2\\&\quad -\gamma\alpha_k^*\left\langle u^k-u^{k+1}, \left(I+M^T\right)\left(u^k-\tilde{u}^k\right)\right\rangle.\end{aligned} \tag{21}$$

For the crossed term of the right-hand side in (21), using Cauchy–Schwarz inequality and (6), we get

$$\begin{aligned}&-\gamma\alpha_k^*\langle u^k-u^{k+1}, (I+M^T)(u^k-\tilde{u}^k)\rangle\\&=-\left\langle u^k-u^{k+1}, G\left[\gamma\alpha_k^* G^{-1}(I+M^T)(u^k-\tilde{u}^k)\right]\right\rangle\\&\geq -\frac{1}{2}\|u^k-u^{k+1}\|_G^2 - \frac{1}{2}\gamma^2\left(\alpha_k^*\right)^2\|G^{-1}(I+M^T)(u^k-\tilde{u}^k)\|_G^2\\&=-\frac{1}{2}\|u^k-u^{k+1}\|_G^2 - \frac{1}{2}\gamma^2\alpha_k^*\|u^k-\tilde{u}^k\|^2.\end{aligned}$$

Substituting it into the right-hand side of (21) and using the notation $q_k(\gamma)$, we obtain

$$\begin{aligned}&\left\langle u^{k+1}-\tilde{u}^k, \gamma\alpha_k^*[(Mu^k+q)+M^T(u^k-\tilde{u}^k)]\right\rangle\\&\quad\geq \frac{1}{2}q_k(\gamma)+\frac{\gamma\alpha_k^*}{2}\|u^k-\tilde{u}^k\|_D^2-\frac{1}{2}\|u^k-u^{k+1}\|_G^2. \end{aligned}\tag{22}$$

Now, we turn to treat term (20b). The update form of Algorithm-II (9) means that u^{k+1} is the projection of the vector $\left(u^k-\gamma\alpha_k^*G^{-1}[(Mu^k+q)+M^T(u^k-\tilde{u}^k)]\right)$ onto Ω. Thus, it follows from (10) that

$$\left\langle (u^k-\gamma\alpha_k^*G^{-1}[(Mu^k+q)+M^T(u^k-\tilde{u}^k)])-u^{k+1}, G(u-u^{k+1})\right\rangle\leq 0,\quad \forall u\in\Omega,$$

and consequently

$$\left\langle u-u^{k+1}, \gamma\alpha_k^*[(Mu^k+q)+M^T(u^k-\tilde{u}^k)]\right\rangle\geq\langle u-u^{k+1}, G(u^k-u^{k+1})\rangle,\quad\forall u\in\Omega.$$

Using the identity

$$\langle a, Gb\rangle=\frac{1}{2}\left\{\|a\|_G^2-\|a-b\|_G^2+\|b\|_G^2\right\}$$

for the right-hand side of the last inequality, we obtain

$$\begin{aligned}&\left\langle u-u^{k+1}, \gamma\alpha_k^*[(Mu^k+q)+M^T(u^k-\tilde{u}^k)]\right\rangle\\&\quad\geq\frac{1}{2}\left(\|u-u^{k+1}\|_G^2-\|u-u^k\|_G^2\right)+\frac{1}{2}\|u^k-u^{k+1}\|_G^2.\end{aligned}\tag{23}$$

Adding (22) and (23) together, we get

$$\left\langle u-\tilde{u}^k, \gamma\alpha_k^*[(Mu^k+q)+M^T(u^k-\tilde{u}^k)]\right\rangle\tag{24}$$

$$\geq\frac{1}{2}\left(\|u-u^{k+1}\|_G^2-\|u-u^k\|_G^2\right)+\frac{1}{2}q_k(\gamma)+\frac{\gamma\alpha_k^*}{2}\|u^k-\tilde{u}^k\|_D^2.\tag{25}$$

Finally, substituting it into (19), the proof is complete. □

Based on Lemmas 4.1 and 4.2, the strict contraction property of the sequences generated by Algorithm-I and Algorithm-II can be easily derived. We summarize them in the following corollary.

Corollary 4.1 *The sequence $\{u^k\}$ generated by either Algorithm-I or Algorithm-II is strictly contractive with respect to the solution set Ω^* of LVI(Ω, M, q).*

Proof In Lemmas 4.1 and 4.2, we have proved that the sequence $\{u^k\}$ generated by either Algorithm-I or Algorithm-II satisfies inequality (14). Setting $u = u^*$ in (14) where $u^* \in \Omega^*$ is an arbitrary solution point of LVI(Ω, M, q), we get

$$\|u^k - u^*\|_G^2 - \|u^{k+1} - u^*\|_G^2 \geq 2\gamma\alpha_k^* \langle \tilde{u}^k - u^*, Mu^* + q\rangle + q_k(\gamma).$$

Because $\langle \tilde{u}^k - u^*, Mu^* + q\rangle \geq 0$, it follows from the last inequality and (13) that

$$\|u^{k+1} - u^*\|_G^2 \leq \|u^k - u^*\|_G^2 - \gamma(2-\gamma)\alpha_k^* \|u^k - \tilde{u}^k\|^2,$$

which means that the sequence $\{u^k\}$ generated by either Algorithm-I or Algorithm-II is strictly contractive with respect to the solution set Ω^*. The proof is complete. $\square$

5 Estimates on Iteration Complexity

In this section, we estimate the worst-case convergence rates measured by the iteration complexity for Algorithm-I and Algorithm-II. We discuss both the ergodic and nonergodic senses.

5.1 Iteration Complexity in the Ergodic Sense

We first derive the worst-case convergence rates measured by the iteration complexity in the ergodic sense. For this purpose, we need the definition of an ϵ-approximated solution point of LVI(Ω, M, q) in sense of the variational inequality characterization (12).

Theorem 5.1 *Let the sequence $\{u^k\}$ be generated by either Algorithm-I or Algorithm-II starting from u^0, and $\tilde{u}^k$ be given by (3). For any integer $t > 0$, let*

$$\tilde{u}_t = \frac{1}{\Upsilon_t}\sum_{k=0}^{t} \alpha_k^* \tilde{u}^k \quad \text{and} \quad \Upsilon_t = \sum_{k=0}^{t} \alpha_k^*. \tag{26}$$

Then, it holds that

$$\langle \tilde{u}_t - u, Mu + q\rangle \leq \frac{\|u - u^0\|_G^2}{2\alpha_{\min}\gamma(t+1)}, \quad \forall u \in \Omega. \tag{27}$$

Proof Note that Lemmas 4.1 and 4.2 still hold for any $\gamma > 0$; the strict contraction in Corollary 4.1 is guaranteed for $\gamma \in]0, 2[$. In this proof, we can slightly extend the restriction of γ to $\gamma \in]0, 2]$. Clearly, for this case, we still have $q_k(\gamma) \geq 0$. It follows from the positivity of M, (13) and (14) that

$$\langle u - \tilde{u}^k, \alpha_k^*(Mu + q)\rangle + \frac{1}{2\gamma}\|u - u^k\|_G^2 \geq \frac{1}{2\gamma}\|u - u^{k+1}\|_G^2, \quad \forall u \in \Omega.$$

Summarizing the above inequality over $k = 0, \dots, t$, we obtain

$$\left\langle \left(\sum_{k=0}^{t} \alpha_k^*\right) u - \sum_{k=0}^{t} \alpha_k^* \tilde{u}^k, Mu + q \right\rangle + \frac{1}{2\gamma} \|u - u^0\|_G^2 \geq 0, \quad \forall u \in \Omega.$$

Then, using the notation of Υ_t and $\tilde{u}_t$ in the above inequality, we derive

$$\langle \tilde{u}_t - u, Mu + q \rangle \leq \frac{\|u - u^0\|_G^2}{2\gamma \Upsilon_t}, \quad \forall u \in \Omega. \tag{28}$$

Indeed, $\tilde{u}_t \in \Omega$ because it is a convex combination of the iterates $\tilde{u}^0, \tilde{u}^1, \dots, \tilde{u}^t$. Because of $\alpha_k^* \geq \alpha_{\min}$ (see (7)), it follows from (26) that

$$\Upsilon_t \geq (t+1)\alpha_{\min}.$$

Substituting it into (28), the proof is complete. □

The next theorem shows clearly the worst-case $O(1/t)$ convergence rate measured by the iteration complexity in the ergodic sense for Algorithm-I and Algorithm-II.

Theorem 5.2 *For any $\epsilon > 0$ and $u^* \in \Omega^*$, with an initial iterate u^0, either Algorithm-I or Algorithm-II requires no more iterations than $\left\lceil \frac{d}{2\alpha_{\min}\gamma\epsilon} \right\rceil$ to produce an ϵ-approximated solution point of LVI(Ω, M, q) in sense of the variational inequality characterization (12), where*

$$d := 3 + 9\|u^0 - u^*\|_G^2 + \frac{6\|G\|_2 \|u^0 - u^*\|_G^2}{\gamma(2-\gamma)\alpha_{\min}}. \tag{29}$$

Proof For $u \in \mathcal{D}(\tilde{u}_t)$, it follows from Cauchy–Schwarz inequality and the convexity of $\|\cdot\|_G^2$ that

$$\begin{aligned} \|u - u^0\|_G^2 &\leq 3\|u - \tilde{u}_t\|_G^2 + 3\|u^0 - u^*\|_G^2 + 3\|\tilde{u}_t - u^*\|_G^2 \\ &\leq 3 + 3\|u^0 - u^*\|_G^2 + 3 \max_{0 \leq k \leq t} \|\tilde{u}^k - u^*\|_G^2 \\ &\leq 3 + 3\|u^0 - u^*\|_G^2 + 6 \max_{0 \leq k \leq t} \|u^k - u^*\|_G^2 \\ &\quad + 6 \max_{0 \leq k \leq t} \|u^k - \tilde{u}^k\|_G^2. \end{aligned} \tag{30}$$

On the other hand, it follows from (8) that

$$\|u^k - u^*\|_G^2 \leq \|u^0 - u^*\|_G^2 \tag{31}$$

and

$$\|u^k - \tilde{u}^k\|^2 \leq \frac{\|u^0 - u^*\|_G^2}{\gamma(2-\gamma)\alpha_k^*} \leq \frac{\|u^0 - u^*\|_G^2}{\gamma(2-\gamma)\alpha_{\min}}. \tag{32}$$

Because of the inequality

$$\|u^k - \tilde{u}^k\|_G^2 \le \|G\|_2 \|u^k - \tilde{u}^k\|^2,$$

it follows from (30), (31) and (32) that

$$\|u - u^0\|_G^2 \le 3 + 9\|u^0 - u^*\|_G^2 + \frac{6\|G\|_2\|u^0 - u^*\|_G^2}{\gamma(2-\gamma)\alpha_{\min}} = d. \tag{33}$$

This, together with (27), completes the proof of the theorem. □

Remark 5.1 Recall that $v \in \Omega$ is called an ϵ-approximated solution point of LVI(Ω, M, q) in sense of the variational inequality characterization if it satisfies (12). Replacing $\mathcal{D}(v)$ with Ω in (12), we obtain the standard merit function

$$\sup_{u \in \Omega} \left\{ \langle v - u, Mu + q \rangle \right\},$$

which has been widely used in the literature, e.g., [9,10]. Moreover, if Ω is bounded, it follows from (28) that

$$\sup_{u \in \Omega} \left\{ \langle \tilde{u}_t - u, Mu + q \rangle \right\} \le \frac{d}{2\alpha_{\min}\gamma(t+1)},$$

where $d = \sup\{\|u - u^0\|_G^2 \mid u \in \Omega\}$. Then, for any $\epsilon > 0$, with an initial iterate u^0, either Algorithm-I or Algorithm-II requires no more iterations than $\left\lceil \frac{d}{2\alpha_{\min}\gamma\epsilon} \right\rceil$ to produce an ϵ-approximated solution in sense of

$$\sup_{u \in \Omega} \left\{ \langle \tilde{u}_t - u, Mu + q \rangle \right\} \le \epsilon.$$

5.2 Iteration Complexity in a Nonergodic Sense

In this subsection, we derive the worst-case $O(1/t)$ convergence rates measured by the iteration complexity in a nonergodic sense for Algorithm-I and Algorithm-II. For this purpose, we need the definition of an ϵ-approximated solution point of LVI(Ω, M, q) in sense of the projection equation residual characterization mentioned in Sect. 3.2.

The worst-case $O(1/t)$ convergence rates in a nonergodic sense for Algorithm-I and Algorithm-II are proved in the following theorem.

Theorem 5.3 *For any $\epsilon > 0$ and $u^* \in \Omega^*$, with an initial iterate u^0, either Algorithm-I or Algorithm-II requires no more iterations than $\left\lceil \frac{\|u^0 - u^*\|_G^2}{\alpha_{\min}\gamma(2-\gamma)\epsilon} \right\rceil$ to obtain an ϵ-approximated solution point of LVI(Ω, M, q) in the sense of $\|e(u^k)\|^2 \le \epsilon$.*

Proof Summarizing inequality (8) over $k = 0, 1, \ldots, t$ and using the inequality $\alpha_k \geq \alpha_{\min}$, we derive that

$$\sum_{k=0}^{\infty} \|e(u^k)\|^2 \leq \frac{\|u^0 - u^*\|_G^2}{\alpha_{\min}\gamma(2-\gamma)}. \tag{34}$$

This implies

$$(t+1) \min_{0\leq k\leq t} \|e(u^k)\|^2 \leq \sum_{k=0}^{t} \|e(u^k)\|^2 \leq \frac{\|u^0 - u^*\|_G^2}{\alpha_{\min}\gamma(2-\gamma)},$$

which proves this theorem. □

Indeed, the worst-case $O(1/t)$ convergence rates in a nonergodic sense established in Theorem 5.3 can be easily refined as an $o(1/t)$ order. We summarize it in the following corollary.

Corollary 5.1 *Let the sequence $\{u^k\}$ be generated by either Algorithm-I or Algorithm-II; $e(u^k)$ be defined in (4). For any integer $t > 0$, it holds that*

$$\min_{0\leq k\leq t} \|e(u^k)\|^2 = o(1/t), \quad \text{as } t \to \infty \tag{35}$$

Proof Notice that

$$\frac{t}{2} \min_{0\leq k\leq t} \|e(u^k)\|^2 \leq \sum_{i=\lfloor \frac{t}{2} \rfloor + 1}^{t} \|e(u^k)\|^2 \to 0 \tag{36}$$

as $t \to \infty$, where equation (36) holds due to (34) and Cauchy principle. The proof is complete. □

6 Conclusions

We study the iteration complexity of two projection methods for monotone linear variational inequalities and derive their worst-case convergence rates measured by the iteration complexity in both the ergodic and nonergodic senses. The proofs critically depend on the strict contraction property of the sequences generated by these two projection methods. Our analysis is conducted under mild assumptions, and the derived worst-case convergence rates are sublinear. We do not need any error bound conditions which are usually required by projection methods for deriving asymptotically linear convergence rates, nor the boundedness restriction onto the feasible set which is usually required by estimating iteration complexity-based convergence rates for some algorithms to solve nonlinear variational inequalities. It is interesting to consider extending our analysis to projection-type methods for nonlinear variational inequalities such as the extragradient methods in [10] and the modified forward–backward methods in [15]. We leave this topic as a future work.

Acknowledgements The authors are grateful to the anonymous referees and the editor for their valuable comments which have helped them to improve the presentation of this paper. This research was supported in part by the Natural Science Foundation of Jiangsu Province under Project No. BK20130550, the NSFC Grant 11401300, 71271112, 71390333, 70901018, 11471156 and 91530115, and the General Research Fund from Hong Kong Research Grants Council: HKBU 12302514.

References

1. Hartman, P., Stampacchia, G.: On some nonlinear elliptic differential functional equations. Acta Math. **115**, 271–310 (1966)
2. Facchinei, F., Pang, J.S.: Finite-Dimensional Variational Inequalities and Complementarity Problems. Vol. I and II, Springer Series in Operations. Springer, New York (2003)
3. Solodov, M.V., Tseng, P.: Modified projection-type methods for monotone variational inequalities. SIAM J. Control Optim. **34**, 1814–1830 (1996)
4. Blum, E., Oettli, W.: Mathematische Optimierung: Grundlagen und Verfahren. Ökonometrie und Unternehmensforschung. Springer, Berlin (1975)
5. He, B.S.: A new method for a class of linear variational inequalities. Math. Program. **66**, 137–144 (1994)
6. He, B.S.: Solving a class of linear projection equations. Numer. Math. **68**, 71–80 (1994)
7. He, B.S. : A globally linearly convergent projection and contraction method for a class of linear complementarity problems. Schwerpunktprogramm der Deutschen Forschungsgemeinschaft Anwendungsbezogene Optimierung und Steuerung, Report No. 352 (1992)
8. Nesterov, Y.E.: A method for solving the convex programming problem with convergence rate $O(1/k^2)$. Dokl. Akad. Nauk SSSR **269**, 543–547 (1983)
9. Nemirovski, A.: Prox-method with rate of convergence $O(1/t)$ for variational inequality with Lipschitz continuous monotone operators and smooth convex-concave saddle point problems. SIAM J. Optim. **15**, 229–251 (2005)
10. Korpelevich, G.M.: The extragradient method for finding saddle points and other problems. Ekonomika i Matematicheskie Metody **12**, 747–756 (1976)
11. Nesterov, Y.E.: Dual extrapolation and its applications to solving variational inequalities and related problems. Math. Program. **109**, 319–344 (2007)
12. Bertsekas, D.P., Tsitsiklis, J.N.: Parallel and Distributed Computation, Numerical Methods. Prentice-Hall, Englewood Cliffs (1989)
13. He, B.S.: A modified projection and contraction method for a class of linear complementarity problems. J. Comput. Math. **14**, 54–63 (1996)
14. Nesterov, Y.E.: Gradient methods for minimizing composite objective functions. Math. Program. **140**(1), 125–161 (2013)
15. Tseng, P.: A modified forward-backward splitting method for maximal monotone mappings. SIAM J. Con. Optim. **38**, 431–446 (2000)