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

Caihua Chen* Xiaoling Fu† Bingsheng He‡ Xiaoming Yuan §

January 13, 2015

Abstract. Projection type methods are among the most important methods for solving monotone linear variational inequalities. In this note, we analyze the iteration complexity for two projection methods and accordingly establish their worst-case $O(1/t)$ convergence rates measured by the iteration complexity in both the ergodic and nonergodic senses, where t is the iteration counter. 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.

1 Introduction

Let Ω be a closed convex subset of R^n, $M \in R^{n\times n}$ and $q \in R^n$. The linear variational inequality problem, denoted by $\mathrm{LVI}(\Omega, M, q)$, is to find a vector $u^* \in \Omega$ such that

$$\mathrm{LVI}(\Omega, M, q) \qquad (u - u^*)^T(Mu^* + q) \geq 0, \qquad \forall u \in \Omega. \tag{1.1}$$

We consider the case where the matrix M is positive semi-definite (but could be asymmetric). Moreover, the solution set of (1.1), denoted by Ω^*, is assumed to be nonempty.

It is well known (see e.g. [1], pp. 267) that u^* is a solution point of (1.1) if and only if it satisfies the following projection equation

$$u^* = P_{[\Omega,G]}[u^* - G^{-1}(Mu^* + q)], \tag{1.2}$$

*International Centre of Management Science and Engineering, School of Management and Engineering, Nanjing University, Nanjing, China. This author was supported in part by the Natural Science Foundation of Jiangsu Province under project No. BK20130550 and the NSFC Grant 11401300 and 71271112. Email: chchen@nju.edu.cn

†Institute of Systems Engineering, Southeast University, Nanjing, 210096, China. This author was supported by the NSFC grant 70901018 and 71390333. Email: fuxlnju@hotmail.com

‡International Centre of Management Science and Engineering, School of Management and Engineering, Nanjing University, Nanjing, China; and Department of Mathematics, Nanjing University, Nanjing, China. This author was supported by the NSFC Grant 11471156 and the MOEC fund 20110091110004. Email: hebma@nju.edu.cn

§Department of Mathematics, Hong Kong Baptist University, Hong Kong, China. This author was supported by a General Research Fund from Hong Kong Research Grants Council. Email: xmyuan@hkbu.edu.hk

where $G \in R^{n\times n}$ is a symmetric positive definite matrix, $P_{[\Omega,G]}(\cdot)$ 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 $u \in \Re^n$. When $G = I$, we simply use the notation $P_\Omega(\cdot)$ for $P_{[\Omega,I]}(\cdot)$. Moreover, for given $u \in R^n$, we denote $\tilde{u} = P_\Omega[u - (Mu + q)]$. Hence, we have

$$\tilde{u}^k = P_\Omega[u^k - (Mu^k + q)]. \tag{1.3}$$

We further use

$$e(u^k) := u^k - \tilde{u}^k. \tag{1.4}$$

It follows from (1.2) that u is a solution 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 characterization (1.2) 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. [6, 7, 14] 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 (1.1) is simple in the sense the projection onto it can be easily computed. We refer to [3] for a survey. In this paper, we consider Algorithm 2.1 in [14], which is a representative projection method for LVI(Ω, M, q) whose efficiency has been verified numerically in the paper. More specifically, its iterative scheme is

$$\text{(Algorithm-I)} \qquad u^{k+1} = u^k - \gamma\alpha_k^* G^{-1}(I + M^T)(u^k - \tilde{u}^k), \tag{1.5}$$

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

$$\alpha_k^* = \frac{\|u^k - \tilde{u}^k\|^2}{\|G^{-1}(I + M^T)(u^k - \tilde{u}^k)\|_G^2}. \tag{1.6}$$

Obviously, for the step size α_k^* defined in (1.6), we have

$$\alpha_k^* \geq \frac{1}{\|(I + M)G^{-1}(I + M)^T\|_2} := \alpha_{\min}, \tag{1.7}$$

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 [14], the sequence $\{u^k\}$ generated by Algorithm-I satisfies the inequality

$$\|u^{k+1} - u^*\|_G^2 \leq \|u^k - u^*\|_G^2 - \gamma(2 - \gamma)\alpha_k^*\|u^k - \tilde{u}^k\|^2, \tag{1.8}$$

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 (1.7), the inequality (1.8) essentially means that the sequence $\{u^k\}$ is strictly contractive with respect to the solution set of LVI(Ω, M, q). Hence, the convergence of Algorithm-I follows from the standard analytic framework of contraction methods, as shown in [2]. A special case of Algorithm-I with $G = I$ and $\gamma = 1$ was proposed in [6] and its linear convergence was proved for the case where $\Omega = R_+^n$.

In addition to Algorithm-I in (1.5), we consider another projection method

$$\text{(Algorithm-II)} \qquad u^{k+1} = P_{[\Omega,G]}\{u^k - \gamma\alpha_k^* G^{-1}[(Mu^k + q) + M^T(u^k - \tilde{u}^k)]\}, \tag{1.9}$$

where the step size length α_k^* is also defined in (1.6). As Algorithm-I, we will show later (see Corollary 3.3) that the sequence $\{u^k\}$ generated by Algorithm-II also satisfies the property (1.8) and thus its convergence is ensured. The special case of Algorithm-II with $\Omega = R^n_+$ and $G = I$ can be found in [5] and mentioned in [7], and its convergence proof can be found in [8]. Also, Algorithm-II differs from Algorithm 2.3 in [14] in that its step size is determined by (1.6) and thus it is bounded away from zero, while the latter's may tend to zero, see (2.14) on pp. 1821 in [14]. Note that Algorithm-I and Algorithm-II utilize the same strategy (1.6) to determine their step sizes but along different search directions; and they require computing one and two projections at each iteration, respectively.

Our main purpose is analyzing the convergence rates for Algorithm-I and Algorithm-II under mild assumptions on (1.1). Indeed, if certain error bound condition is assumed, the strict contraction property (1.8) enables us to establish the asymptotical convergence rates for Algorithm-I and Algorithm-II immediately. More specifically, if we assume that there exists positive constant μ and δ (depending on M, q, Ω only) such that

$$d(u, \Omega^*) \le \mu\|u - \tilde{u}\|, \ \forall\, u \text{ with } \|u - \tilde{u}\| \le \delta,$$

where $d(\cdot, \Omega^*)$ denotes the 2-norm distance to Ω^*, then the linear convergence of Algorithm-I and Algorithm-II is just an obvious conclusion of (1.8). In general, however, it is not easy to verify error bound conditions even for LVI(Ω, M, q). We thus consider the possibility of deriving the convergence rates for Algorithm-I and Algorithm-II without any error bound conditions. Indeed, we want to derive 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. This kind of iteration-complexity-based analysis for convergence rate traces back to [11] and it has received much attention from the literature. 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. An ϵ-approximated solution point of LVI(Ω, M, q) will be defined precisely. Furthermore, it is noteworthy that our analysis for the convergence rates of Algorithm-I and Algorithm-II does not require the boundedness of the feasible set Ω in (1.1), which is usually required for the iteration complexity analysis of projection methods for nonlinear variational inequalities such as the work [10] for the extragradient method in [9]. Finally, we would mention that we only focus on the LVI(Ω, M, q) in this note and do not discuss the iteration complexity analysis of projection methods for nonlinear variational inequalities. We refer to, e.g., [10, 12], for some insightful discussions in this regard.

Throughout the paper, the following notational conventions are used. We use u^* to denote a fixed but arbitrary point in the solution set Ω^* of LVI(Ω, M, q). 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 will be denoted by $\|\cdot\|$.

2 Preliminaries

In this section we summarize some preliminaries which are useful for our analysis.

2.1 Some Inequalities

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

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

we have

$$(v - P_{[\Omega,G]}(v))^T G(u - P_{[\Omega,G]}(v)) \le 0, \quad \forall\, v \in R^n, \forall\, u \in \Omega. \tag{2.1}$$

Let u^* be any fixed solution point. Since $\tilde{u}^k \in \Omega$, it follows from (1.1) that

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

Set $v = u^k - (Mu^k + q)$, $G = I$ and $u = u^*$ in (2.1), because of the notation $\tilde{u}^k$, we have

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

Adding the last two inequalities, we obtain

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

and consequently

$$(u^k - u^*)^T (I + M^T)(u^k - \tilde{u}^k) \ge \|u^k - \tilde{u}^k\|^2, \quad \forall\, u^* \in \Omega^*. \tag{2.2}$$

2.2 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 clearly 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) on Page 159 in [3], we know that Ω^* is convex and it can be characterized by

$$\Omega^* = \bigcap_{u \in \Omega} \big\{ v \in \Omega : (u - v)^T (Mu + q) \ge 0 \big\}.$$

Therefore, motivated by [13], 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

$$v \in \Omega \quad \text{and} \quad \inf_{u \in \mathcal{D}(v)} \big\{ (u - v)^T (Mu + q) \big\} \ge -\epsilon,$$

where

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

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

$$v \in \Omega \quad \text{and} \quad \sup_{u \in \mathcal{D}(v)} \big\{ (v - u)^T (Mu + q) \big\} \le \epsilon. \tag{2.3}$$

The other definition comes from the mentioned fact that $\|e(u)\|^2$ defined in (1.4) serves as 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 $\|e(v)\|^2 \le \epsilon$.

3 Two Lemmas

One reason we consider Algorithm-I and Algorithm-II simultaneously is that their iteration-complexity-based convergence rate analysis share a great degree of similarity 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 iteration complexities. For notation simplicity, we define

$$q_k(\gamma) = \gamma(2-\gamma)\alpha_k^*\|u^k - \tilde{u}^k\|^2, \tag{3.1}$$

where α_k^* is given by (1.6). Moreover, let us use the notation

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

where M is the matrix in (1.1).

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

$$\gamma\alpha_k^*(u - \tilde{u}^k)^T(Mu+q) \geq \frac{1}{2}\big(\|u - u^{k+1}\|_G^2 - \|u - u^k\|_G^2\big) + \frac{1}{2}q_k(\gamma), \;\; \forall\, u \in \Omega, \tag{3.2}$$

where $q_k(\gamma)$ is defined in (3.1). We will present this conclusion in two lemmas.

Lemma 3.1. *For given $u^k \in R^n$, let $\tilde{u}^k$ be defined by* (1.3) *and the new iterate u^{k+1} be generated by Algorithm-I* (1.5)*. Then, the assertion (3.2) is satisfied.*

Proof. Set $v = u^k - (Mu^k + q)$ in (2.1) and we know $\tilde{u}^k = P_\Omega[u^k - (Mu^k+q)]$. Thus, we have

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

This inequality can be rewritten as

$$(u - \tilde{u}^k)^T\{(Mu+q) - M(u - \tilde{u}^k) + (M + M^T)(u^k - \tilde{u}^k) - (I + M^T)(u^k - \tilde{u}^k)\} \geq 0, \; \forall\, u \in \Omega.$$

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

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

Moreover, it follows from (1.5) that

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

Thus, we obtain

$$\gamma\alpha_k^*(u - \tilde{u}^k)^T(Mu+q) \geq (u - \tilde{u}^k)^T G(u^k - u^{k+1}) - \frac{\gamma\alpha_k^*}{2}\|u^k - \tilde{u}^k\|_D^2. \tag{3.3}$$

For the crossed term in the right-hand-side of (3.3): $(u - \tilde{u}^k)^T G(u^k - u^{k+1})$, it follows from the identity

$$(a-b)^T G(c-d) = \frac{1}{2}\big(\|a-d\|_G^2 - \|a-c\|_G^2\big) + \frac{1}{2}\big(\|c-b\|_G^2 - \|d-b\|_G^2\big)$$

that

$$(u-\tilde{u}^k)^T G(u^k-u^{k+1}) = \frac{1}{2}\big(\|u-u^{k+1}\|_G^2 - \|u-u^k\|_G^2\big) + \frac{1}{2}\big(\|u^k-\tilde{u}^k\|_G^2 - \|u^{k+1}-\tilde{u}^k\|_G^2\big). \quad (3.4)$$

Now, we treat the second part of the right-hand-side of (3.4). Using (1.5), we get

$$\begin{aligned} & \|u^k-\tilde{u}^k\|_G^2 - \|u^{k+1}-\tilde{u}^k\|_G^2 \\ = \;& \|u^k-\tilde{u}^k\|_G^2 - \|(u^k-\tilde{u}^k) - \gamma\alpha_k^* G^{-1}(I+M^T)(u^k-\tilde{u}^k)\|_G^2 \\ = \;& 2\gamma\alpha_k^*(u^k-\tilde{u}^k)^T(I+M^T)(u^k-\tilde{u}^k) - (\gamma\alpha_k^*)^2\|G^{-1}(I+M^T)(u^k-\tilde{u}^k)\|_G^2 \\ = \;& 2\gamma\alpha_k^*\|u^k-\tilde{u}^k\|^2 + \gamma\alpha_k^*\|u^k-\tilde{u}^k\|_D^2 - (\gamma\alpha_k^*)^2\|G^{-1}(I+M^T)(u^k-\tilde{u}^k)\|_G^2. \end{aligned} \quad (3.5)$$

Recall (1.6). Thus, it follows from (3.5) that

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

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 (3.4) and using the definition of $q_k(\gamma)$, we obtain

$$(u-\tilde{u}^k)^T G(u^k-u^{k+1}) = \frac{1}{2}\big(\|u-u^{k+1}\|_G^2 - \|u-u^k\|_G^2\big) + \frac{1}{2}q_k(\gamma) + \frac{\gamma\alpha_k^*}{2}\|u^k-\tilde{u}^k\|_D^2. \quad (3.6)$$

Adding (3.3) and (3.6) together, we get the assertion (3.2) and the theorem is proved. □

Then, we prove the assertion (3.2) for Algorithm-II in the following lemma.

Lemma 3.2. *For given $u^k \in R^n$, let $\tilde{u}^k$ be defined by* (1.3) *and the new iterate u^{k+1} be generated by Algorithm-II* (1.9)*. Then, the assertion (3.2) is satisfied.*

Proof. It follows from Cauchy-Schwarz Inequality that

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

Consequently, we obtain

$$\gamma\alpha_k^*(u-\tilde{u}^k)^T(Mu+q) \geq (u-\tilde{u}^k)^T\gamma\alpha_k^*[(Mu^k+q) + M^T(u^k-\tilde{u}^k)] - \frac{\gamma\alpha_k^*}{2}\|u^k-\tilde{u}^k\|_D^2. \quad (3.7)$$

Now we investigate the first term in the right-hand-side of (3.7) and divide it into the following two terms, namely

$$(u^{k+1}-\tilde{u}^k)^T\gamma\alpha_k^*[(Mu^k+q) + M^T(u^k-\tilde{u}^k)] \quad (3.8a)$$

and

$$(u-u^{k+1})^T\gamma\alpha_k^*[(Mu^k+q) + M^T(u^k-\tilde{u}^k)]. \quad (3.8b)$$

First, we deal with the term (3.8a). Set $v = u^k - (Mu^k + q)$ in (2.1). Since $\tilde{u}^k = P_\Omega[u^k - (Mu^k + q)]$ and $u^{k+1} \in \Omega$, it follows that

$$(u^{k+1} - \tilde{u}^k)^T (Mu^k + q) \geq (u^{k+1} - \tilde{u}^k)^T (u^k - \tilde{u}^k).$$

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

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

and it follows that

$$\begin{aligned}
&(u^{k+1} - \tilde{u}^k)^T \gamma\alpha_k^*[(Mu^k + q) + M^T(u^k - \tilde{u}^k)] \\
&\geq \gamma\alpha_k^*(u^{k+1} - \tilde{u}^k)^T (I + M^T)(u^k - \tilde{u}^k) \\
&= \gamma\alpha_k^*(u^k - \tilde{u}^k)^T (I + M^T)(u^k - \tilde{u}^k) - \gamma\alpha_k^*(u^k - u^{k+1})^T (I + M^T)(u^k - \tilde{u}^k) \\
&\geq \gamma\alpha_k^*\|u^k - \tilde{u}^k\|^2 + \frac{\gamma\alpha_k^*}{2}\|u^k - \tilde{u}^k\|_D^2 - \gamma\alpha_k^*(u^k - u^{k+1})^T (I + M^T)(u^k - \tilde{u}^k). \qquad (3.9)
\end{aligned}$$

For the crossed term of the right-hand-side in (3.9), using Cauchy-Schwarz Inequality and (1.6), we get

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

Substituting it into the right-hand-side of (3.9) and using the notation $q_k(\gamma)$, we obtain

$$(u^{k+1} - \tilde{u}^k)^T \gamma\alpha_k^*[(Mu^k + q) + M^T(u^k - \tilde{u}^k)] \geq \frac{1}{2}q_k(\gamma) + \frac{\gamma\alpha_k^*}{2}\|u^k - \tilde{u}^k\|_D^2 - \frac{1}{2}\|u^k - u^{k+1}\|_G^2. \quad (3.10)$$

Now, we turn to treat the term (3.8b). The update form of Algorithm-II (1.9) means that u^{k+1} is the projection of $\big(u^k - \gamma\alpha_k^* G^{-1}[(Mu^k + q) + M^T(u^k - \tilde{u}^k)]\big)$ on Ω. Thus, it follows from (2.1) that

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

and consequently

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

Using the identity

$$a^T G b = \frac{1}{2}\{\|a\|_G^2 - \|a - b\|_G^2 + \|b\|_G^2\}$$

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

$$\big(u - u^{k+1}\big)^T \gamma\alpha_k^*[(Mu^k + q) + M^T(u^k - \tilde{u}^k)] \geq \frac{1}{2}\big(\|u - u^{k+1}\|_G^2 - \|u - u^k\|_G^2\big) + \frac{1}{2}\|u^k - u^{k+1}\|_G^2. \quad (3.11)$$

Adding (3.10) and (3.11) together, we get

$$(u - \tilde{u}^k)^T \gamma\alpha_k^*[(Mu^k + q) + M^T(u^k - \tilde{u}^k)] \geq \frac{1}{2}\big(\|u - u^{k+1}\|_G^2 - \|u - u^k\|_G^2\big) + \frac{1}{2}q_k(\gamma) + \frac{\gamma\alpha_k^*}{2}\|u^k - \tilde{u}^k\|_D^2.$$

Finally, substituting it into (3.7), the proof is complete. □

Based on Lemmas 3.1 and 3.2, 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 3.3. *The sequence $\{u^k\}$ generated by Algorithm-I or Algorithm-II is strictly contractive with respect to the solution set Ω^* of LVI(Ω, M, q).*

Proof. In Lemmas 3.1 and 3.2, we have proved that the sequence $\{u^k\}$ generated by either Algorithm-I or Algorithm-II satisfies the inequality (3.2). Setting $u = u^*$ in (3.2) where $u^* \in \Omega^*$ is an arbitrary solution point of LVI(Ω, M, q), we get

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

Because $(\tilde{u}^k - u^*)^T(Mu^* + q) \geq 0$, it follows from the last inequality and (3.1) 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 Algorithm-I or Algorithm-II is strictly contractive with respect to the solution set Ω^*. The proof is complete. □

4 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.

4.1 Iteration Complexity in the Ergodic Sense

We first derive 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 (2.3).

Theorem 4.1. *Let the sequence $\{u^k\}$ be generated by Algorithm-I or Algorithm-II starting from u^0, and $\tilde{u}^k$ be given by* (1.3)*. For any integer $t > 0$, let*

$$\tilde{u}_t = \frac{1}{\Upsilon_t}\sum_{k=0}^{t}\alpha_k^*\tilde{u}^k \qquad \text{and} \qquad \Upsilon_t = \sum_{k=0}^{t}\alpha_k^*. \tag{4.1}$$

Then, it holds that

$$(\tilde{u}_t - u)^T(Mu + q) \leq \frac{\|u - u^0\|_G^2}{2\alpha_{\min}\gamma(t+1)}, \quad \forall u \in \Omega. \tag{4.2}$$

Proof. Note that although Lemmas 3.1 and 3.2 still hold for any $\gamma > 0$; and the strict contraction in Corollary 3.3 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, (3.1) and (3.2) that

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

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

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

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

$$(\tilde{u}_t - u)^T(Mu+q) \le \frac{\|u-u^0\|_G^2}{2\gamma\Upsilon_t}, \quad \forall u \in \Omega. \tag{4.3}$$

Indeed, $\tilde{u}_t \in \Omega$ because it is a convex combination of $\tilde{u}^0, \tilde{u}^1, \ldots, \tilde{u}^t$. Because $\alpha_k^* \ge \alpha_{\min}$ (see (1.7)), it follows from (4.1) that

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

Substituting it into (4.3), 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 4.2. *For any $\epsilon > 0$ and $u^* \in \Omega^*$, starting from u^0, the 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 variation inequality characterization (2.3), where*

$$d := 3 + 9\|u^0 - u^*\|_G^2 + \frac{6\|G\|_2\|u^0-u^*\|_G^2}{\gamma(2-\gamma)\alpha_{\min}}. \tag{4.4}$$

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

$$\begin{aligned}\|u-u^0\|_G^2 &\le 3\|u-\tilde{u}_t\|_G^2 + 3\|u^0-u^*\|_G^2 + 3\|\tilde{u}_t - u^*\|_G^2 \\ &\le 3 + 3\|u^0-u^*\|_G^2 + 3\max_{0\le k\le t}\|\tilde{u}^k - u^*\|_G^2 \\ &\le 3 + 3\|u^0-u^*\|_G^2 + 6\max_{0\le k\le t}\|u^k-u^*\|_G^2 + 6\max_{0\le k\le t}\|u^k-\tilde{u}^k\|_G^2. \end{aligned} \tag{4.5}$$

On the other hand, it follows from (1.8) that

$$\|u^k-u^*\|_G^2 \le \|u^0-u^*\|_G^2 \tag{4.6}$$

and

$$\|u^k-\tilde{u}^k\|^2 \le \frac{\|u^0-u^*\|_G^2}{\gamma(2-\gamma)\alpha_k^*} \le \frac{\|u^0-u^*\|_G^2}{\gamma(2-\gamma)\alpha_{\min}}. \tag{4.7}$$

Using the inequality

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

it follows from (4.5), (4.6) and (4.7) that

$$\|u-u_0\|_G^2 \le 3 + 9\|u^0-u^*\|_G^2 + \frac{6\|G\|_2\|u^0-u^*\|_G^2}{\gamma(2-\gamma)\alpha_{\min}} = d. \tag{4.8}$$

This, together with (4.2), completes the proof of the theorem. □

4.2 Iteration Complexity in a Nonergodic Sense

In this subsection, we derive worst-case $O(1/t)$ convergence rate 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 Section 2.2.

Theorem 4.3. *Let the sequence $\{u^k\}$ be generated by Algorithm-I or Algorithm-II starting from u^0, and $e(u^k)$ be defined in (1.4). For any integer $t > 0$, it holds that*

$$\min_{0 \le k \le t} \|e(u^k)\|^2 \le \frac{\|u^0 - u^*\|_G^2}{\alpha_{\min}\gamma(2-\gamma)(t+1)}. \tag{4.9}$$

Proof. Summarizing the inequality (1.8) over $k = 0, 1, \ldots, t$ and using the inequality $\alpha_k \ge \alpha_{\min}$, we derive that

$$\sum_{k=0}^{\infty} \|e(u^k)\|^2 \le \frac{\|u^0 - u^*\|_G^2}{\alpha_{\min}\gamma(2-\gamma)}. \tag{4.10}$$

This implies

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

which proves the assertion (4.9). The proof is complete. □

Based on Theorem 4.3, we can easily show the worst-case $O(1/t)$ convergence rates in a nonergodic sense for Algorithm I and Algorithm-II. We omit the proof.

Theorem 4.4. *For any $\epsilon > 0$ and $u^* \in \Omega^*$, starting from u^0, 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 $\|e(u^k)\|^2 \le \epsilon$.*

Indeed, the worst-case $O(1/t)$ convergence rate in a nonergodic sense established in Theorem 4.3 can be easily refined as an $o(1/t)$ order. We summarize it in the following corollary.

Corollary 4.5. *Let the sequence $\{u^k\}$ be generated by Algorithm-I or Algorithm-II; and $e(u^k)$ be defined in (1.4). For any integer $t > 0$, it holds that*

$$\min_{0 \le k \le t} \|e(u^k)\|^2 = o(1/t), \quad \text{as } t \to \infty \tag{4.11}$$

Proof. Notice that

$$\frac{t}{2} \min_{0 \le k \le t} \|e(u^k)\|^2 \le \sum_{i=\lfloor \frac{t}{2} \rfloor + 1}^{t} \|e(u^k)\|^2 \to 0 \tag{4.12}$$

as $t \to \infty$, where $\lfloor t/2 \rfloor$ denotes the greatest integer no greater than $t/2$ and the equation (4.12) holds due to (4.10) and the Cauchy principle. The proof is complete. □

5 Conclusions

We study the iteration complexity for 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 reply on the strict contraction property of the sequences generated by these two algorithms. Our analysis is conducted under mild assumptions, and the derived worst-case convergence rates are sublinear. We do not require any error bound conditions which are usually needed for deriving asymptotically linear convergence rates of projection type methods, or 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-like methods for nonlinear variational inequalities such as the extragradient methods in [9] and the modified forward-backward methods in [15]. We leave it as our future work.

References

[1] D. P. Bertsekas and J. N. Tsitsiklis, *Parallel and Distributed Computation, Numerical Methods*, Prentice-Hall, Englewood Cliffs, NJ, 1989.

[2] E. Blum and W. Oettli, *Mathematische Optimierung: Grundlagen und Verfahren. Ökonometrie und Unternehmensforschung.* Berlin-Heidelberg-New York, Springer-Verlag, 1975.

[3] F. Facchinei and J.S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems, Vol. I and II.* Springer Series in Operations Research. Springer Verlag, New York, 2003.

[4] B.S. He, A projection and contraction method for a class of linear complementarity problems and its application in convex quadratic programming, *Applied Math. Optim.*, **25**, pp. 247–262, 1992.

[5] B.S. He, 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.

[6] B.S. He, A new method for a class of linear variational inequalities, *Math. Program.*, **66**, pp. 137–144, 1994.

[7] B.S. He, Solving a class of linear projection equations, *Numer. Math.*, **68**, pp. 71–80, 1994.

[8] B.S. He, A modified projection and contraction method for a class of linear complementarity problems, *J. Comput. Math.*, **14**, pp. 54–63, 1996.

[9] G.M. Korpelevich, The extragradient method for finding saddle points and other problems, *Ekonomika i Matematicheskie Metody*, **12**, pp. 747-756, 1976.

[10] A. Nemirovski, 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**, pp. 229-251, 2005.

[11] Y. E. Nesterov, A method for solving the convex programming problem with convergence rate $O(1/k^2)$, Dokl. Akad. Nauk SSSR, 269 (1983), pp. 543-547.

[12] Y. E. Nesterov, Dual extrapolation and its applications to solving variational inequalities and related problems, *Math. Program.,* **109**, pp. 319-344, 2007.

[13] Y. E. Nesterov, Gradient methods for minimizing composite objective functions, *Math. Program.*, 140(1) (2013), pp. 125-161

[14] M.V. Solodov and P. Tseng, Modified projection-type methods for monotone variational inequalities, *SIAM J. Con. Optim.*, **34**, pp. 1814–1830, 1996.

[15] P. Tseng, A modified forward-backward splitting method for maximal monotone mappings, *SIAM J. Con. Optim.*, **38**, pp. 431-446, 2000.