Proximal-like contraction methods for monotone variational inequalities in a unified framework[1]

Bingsheng He[†2] **Li-Zhi Liao**[‡] **Xiang Wang**[†]

[†]Department of Mathematics, Nanjing University, Nanjing, 210093, China

[‡]Department of Mathematics, Hong Kong Baptist University, Hong Kong, China

Abstract. Approximate proximal point algorithms (abbreviated as APPAs) are classical approaches for convex optimization problems and monotone variational inequalities. To solve the subproblems of these algorithms, the projection method takes the iteration in form of $u^{k+1} = P_\Omega[u^k - \alpha_k d^k]$. Interestingly, many of them can be paired such that $\tilde{u}^k = P_\Omega[u^k - \beta_k F(v^k)] = P_\Omega[\tilde{u}^k - (d_2^k - G d_1^k)]$, where $\inf\{\beta_k\} > 0$ and G is a symmetric positive definite matrix. In other words, this projection equation offers a pair of geminate directions d_1^k and d_2^k for each step. In this paper, for various APPAs we first present a unified framework involving the above equations. Unified characterization is investigated for the contraction and convergence properties under the framework. This shows some essential views behind various outlooks. To study and pair various APPAs for different types of variational inequalities, we thus construct the above equations in different expressions according to the framework. Based on our constructed frameworks, it is interesting to see that, by choosing one of the geminate directions those studied proximal-like methods always utilize the unit step size namely $\alpha_k \equiv 1$. With the same effective quadruplet and the accepting rule, we then present a more efficient class of methods (called extended or general contraction methods), in which only minor extra even negligible costs are needed for a different step size in each iteration. A set of matrix approximation examples as well as six other groups of numerical experiments are constructed to compare the performance between the primary and extended (general) methods. In general, our numerical experiments show the performance of the extended (general) methods are much more promising than that of the primary ones.

Keywords. Variational inequality, monotone, contraction methods.

1 Introduction

Let Ω be a nonempty closed convex subset of R^n and F be a continuous mapping from R^n into itself. Variational inequality problem is to determine a vector $u^* \in \Omega$ such that

$$\text{VI}(\Omega, F) \qquad (u - u^*)^T F(u^*) \ge 0, \qquad \forall u \in \Omega. \tag{1.1}$$

$\text{VI}(\Omega, F)$ problems include nonlinear complementarity problems (when $\Omega = R^n_+$) and systems of nonlinear equations (when $\Omega = R^n$), and thus have many important applications. Notice that $\text{VI}(\Omega, F)$ is invariant when we multiply F by some positive scalar $\beta > 0$. For any $\beta > 0$, it is well known ([1], p. 267) that

$$u^* \text{ is a solution of VI}(\Omega, F) \quad \Longleftrightarrow \quad u^* = P_\Omega[u^* - \beta F(u^*)], \tag{1.2}$$

[1] This work was supported by the National Natural Science Foundation of China (Grant No. 10971095), the Natural Science Foundation of Jiangsu Province (BK2008255), the Cultivation Fund of the Key Scientific and Technical Innovation Project, Ministry of Education of China (708044), and the Research Grant Council of Hong Kong

[2]Email: hebma@nju.edu.cn.

where $P_\Omega(\cdot)$ denotes the projection onto Ω with respect to the Euclidean norm, *i.e.*,

$$P_\Omega(v) = \text{argmin}\{\|u - v\| \mid u \in \Omega\}.$$

Since Ω is convex and closed, the projection onto Ω is unique. We say the mapping F is monotone with respect to Ω if

$$(u - v)^T(F(u) - F(v)) \ge 0, \quad \forall\, u, v \in \Omega.$$

Variational inequality VI(Ω, F) is monotone when the mapping F is monotone. For solving monotone variational inequality, a classical method is the proximal point algorithm (abbreviated as PPA) [26, 27]. For given $u^k \in \Omega$ and $\beta_k > 0$, the new iterate u^{k+1} of the exact PPA is the solution of the following variational inequality:

$$\text{(PPA)} \qquad u \in \Omega, \qquad (u' - u)^T F_k(u) \ge 0, \qquad \forall\, u' \in \Omega, \tag{1.3a}$$

where

$$F_k(u) = (u - u^k) + \beta_k F(u). \tag{1.3b}$$

According to (1.2), solving problem (1.3) is equivalent to finding a solution of the following projection equation

$$u = P_\Omega[u^k - \beta_k F(u)]. \tag{1.4}$$

The sequence $\{u^k\}$ generated by the exact PPA satisfies

$$\|u^{k+1} - u^*\|^2 \le \|u^k - u^*\|^2 - \|u^k - u^{k+1}\|^2, \tag{1.5}$$

where u^* is any solution point of VI(Ω, F) (for a proof, see [20]). Since the new iterate u^{k+1} is closer to the solution set than u^k, we say that the sequence $\{u^k\}$ is Fejér monotone with respect to the solution set. By using the terminology in [2], such methods are called *contraction methods* in this paper.

The ideal form (1.4) of the method is often impractical since in many cases solving problem (1.3) exactly is either impossible or expensive. Extensive developments on approximate proximal point algorithms (abbreviated as APPAs) are followed [4, 5, 29]. Let v^k be an approximate solution of the PPA's subproblem (1.3) accepted by a certain condition. In the sense of (1.4), we have

$$v^k \approx P_\Omega[u^k - \beta_k F(v^k)]. \tag{1.6}$$

Define the right-hand-side of (1.6) by $\tilde{u}^k$, we call

$$(\textbf{Basic equation of APPAs}) \qquad \tilde{u}^k = P_\Omega[u^k - \beta_k F(v^k)] \tag{1.7}$$

the basic equation of APPAs. It is clear that $\tilde{u}^k$ is the exact solution of the subproblem (1.3) if $\tilde{u}^k = v^k$ or $F(\tilde{u}^k) = F(v^k)$. Note that the basic equation of APPAs (1.7) can be always rewritten as

$$\tilde{u}^k = P_\Omega\{\tilde{u}^k - [d_2(u^k, v^k, \tilde{u}^k) - G d_1(u^k, v^k, \tilde{u}^k)]\} \tag{1.8}$$

with the geminate directions of $d_1(u^k, v^k, \tilde{u}^k)$ and $d_2(u^k, v^k, \tilde{u}^k)$, where G is a symmetric positive semi-definite matrix. Because (1.8) is derived from the basic equation of APPAs, the methods developed in this paper are all called *proximal-like methods.* Indeed, reformulation (1.8) plays the key role for constructing the unified framework in this paper. As readers can see in this paper, many existing projection contraction methods can be grouped as *primary methods* (which take $d_1(u^k, v^k, \tilde{u}^k)$ or $d_2(u^k, v^k, \tilde{u}^k)$ as the search direction and adopt the unit step size) under the unified framework. Moreover, according to the unified framework, we can construct more efficient methods (called *general/extended methods*) than the primary ones with only minor extra computational load for a different step size in each step.

Throughout this paper we assume that the operator F is monotone and continuous, the solution set of VI(Ω, F), denoted by Ω^*, is nonempty, and the sequence $\{\beta_k\}$ in (1.7) is bounded, *i.e.*,

$$0 < \beta_L \leq \inf_{k=0}^{\infty} \beta_k \leq \sup_{k=0}^{\infty} \beta_k \leq \beta_U < +\infty.$$

Note that under our assumptions the solution set Ω^* is closed and convex (see pp. 158 in [6]). The projection mapping is a tool in the analysis of this paper. We list its main properties in Lemmas 1.1 and 1.2. The proof of Lemma 1.1 can be found in textbooks, e. g., [2]. For a simple proof of Lemma 1.2, the readers can consult [34].

Lemma 1.1 *Let $\Omega \subset R^n$ be a closed convex set, then we have*

$$(w' - P_\Omega(w'))^T (w - P_\Omega(w')) \leq 0, \quad \forall\ w' \in R^n, \forall\ w \in \Omega. \tag{1.9}$$

Consequently, it follows that

$$\|P_\Omega(w) - P_\Omega(w')\| \leq \|w - w'\|, \qquad \forall\ w,\ w' \in R^n \tag{1.10}$$

and

$$\|w - P_\Omega(w')\|^2 \leq \|w - w'\|^2 - \|w' - P_\Omega(w')\|^2, \quad \forall\ w' \in R^n, \forall\ w \in \Omega. \tag{1.11}$$

Lemma 1.2 *Let $\Omega \subset R^n$ be a closed convex set, w and $d \neq 0$ be any given vectors in R^n. Then $\|P_\Omega(w - td) - w\|$ is a non-decreasing function of t for $t \geq 0$.*

The paper is organized as follows. In Section 2, we propose the unified framework which consists of an effective quadruplet and an accepting rule. Under the framework, the class of primary methods is defined, followed by detailed analysis on their convergence. For the rest of the paper, there are two parts, Part I (Sections 3–8) and Part II (Sections 9–11).

In Part I, guided by the unified framework, Sections 3, 4, 5 and 6 find the effective quadruplets and their related accepting rules for symmetric linear VIs, asymmetric linear VIs, symmetric nonlinear VIs (which equal to differentiable convex optimization problems) and nonlinear VIs, respectively. In Section 7, we present the general contraction methods, in which each convex combination of the geminate directions can be used as the search direction with selected step length. Furthermore, we provide comparisons on the efficiency among numerous directions and step lengths. Better and more reasonable ones are also addressed theoretically. In Section 8, the results in Section 7 are demonstrated by the numerical experiments.

In Part II, we first illustrate how to construct effective quadruplets for two existing APPAs, i.e., Solodov and Svaiter's APPA and the proximal alternating directions method, abide by their own accepting rules, in Sections 9 and 10, respectively. Then in Section 11, we propose a simple extended version of the primary method according to the quadruplet; and compare the proximal alternating directions method and the extended one though various numerical experiments, which strongly demonstrate the efficiency of the extended method.

Finally, some concluding remarks are drawn in Section 12.

2 The unified framework and the primary methods

2.1 The unified framework

Derived from the basic equation (1.7), the unified framework consists of an accepting rule and a related effective quadruplet described as follows.

Definition 2.1 (Accepting rule and effective quadruplet) *For the triplet* $(u^k, v^k, \tilde{u}^k)$ *in the basic equation* (1.7) *and a designed constraint condition, say* $(u^k, v^k, \tilde{u}^k) \in \mathcal{A}(u^k, v^k, \tilde{u}^k)$*, a quadruplet* $\big(d_1(u^k, v^k, \tilde{u}^k),\ d_2(u^k, v^k, \tilde{u}^k),\ \varphi(u^k, v^k, \tilde{u}^k),\ \phi(u^k, v^k, \tilde{u}^k)\big)$ *is called an effective quadruplet for contraction methods if the following conditions are satisfied:*
1. for the geminate directions $d_1(u^k, v^k, \tilde{u}^k)$, $d_2(u^k, v^k, \tilde{u}^k) \in R^n$*, it holds*

$$\tilde{u}^k = P_\Omega\{\tilde{u}^k - [d_2(u^k, v^k, \tilde{u}^k) - Gd_1(u^k, v^k, \tilde{u}^k)]\} \tag{2.1a}$$

where G *is symmetric positive definite;*
2. there is a continuous function $\varphi(u^k, v^k, \tilde{u}^k)$ *such that, for any* $u^* \in \Omega^*$,

$$(\tilde{u}^k - u^*)^T d_2(u^k, v^k, \tilde{u}^k) \geq \varphi(u^k, v^k, \tilde{u}^k) - (u^k - \tilde{u}^k)^T G d_1(u^k, v^k, \tilde{u}^k); \tag{2.1b}$$

3. for the given condition $(u^k, v^k, \tilde{u}^k) \in \mathcal{A}(u^k, v^k, \tilde{u}^k)$,

$$\varphi(u^k, v^k, \tilde{u}^k) \geq \frac{1}{2}\{\|d_1(u^k, v^k, \tilde{u}^k)\|_G^2 + \phi(u^k, v^k, \tilde{u}^k)\}, \tag{2.1c}$$

where $\phi(u^k, v^k, \tilde{u}^k)$ *is a non-negative continuous function;*
4. there is a positive constant $\kappa > 0$ *such that*

$$\kappa\|u^k - \tilde{u}^k\|^2 \leq \phi(u^k, v^k, \tilde{u}^k). \tag{2.1d}$$

According to the above four conditions (2.1a)-(2.1d)*, APPAs can be derived as implies* v^k *being an approximate solution of the subproblem* (1.3) *in the sense of* (1.7)*. Thus, we call the condition* $(u^k, v^k, \tilde{u}^k) \in \mathcal{A}(u^k, v^k, \tilde{u}^k)$ *the accepting rule in the individual APPAs.* □

Remark 2.1 *The condition* (2.1a) *gives two directions* d_1 *and* d_2 *for the projection contraction methods, whilst conditions* (2.1b)-(2.1d) *guarantee the convergence. The reader can see the details in the following two sub-sections.*

Remark 2.2 *Strongly speaking, the effective quadruplet are also depends on* β *usually. We omit the* β *in their expressions for convenience. The parameter* β *is adjusted mainly for satisfying the accepting rule, as can be seen in the following analyses.*

For the convenience of analysis, in what follows we ignore the index k. Namely, instead of β_k, u^k, v^k and $\tilde{u}^k$, we write β, u, v and $\tilde{u}$. From (1.2), condition (2.1a) implies that

$$\tilde{u} \in \Omega, \quad (u' - \tilde{u})^T\{d_2(u, v, \tilde{u}) - Gd_1(u, v, \tilde{u})\} \geq 0, \quad \forall\, u' \in \Omega. \tag{2.2}$$

Remark 2.3 *The exact PPA is a special case of APPAs* (1.7) *in which* $\tilde{u} = v$*, and thus its accepting rule is* $\|v - \tilde{u}\| = 0$*. Indeed, there is an effective quadruplet* $\big(d_1(u, v, \tilde{u}),\ d_2(u, v, \tilde{u}),\ \varphi(u, v, \tilde{u}),\ \phi(u, v, \tilde{u})\big)$ *which satisfies conditions* (2.1) *with* $G = I$*. According to* (1.2)*, we have*

$$\tilde{u} = P_\Omega\{\tilde{u} - [\beta F(\tilde{u}) - (u - \tilde{u})]\}.$$

The above expression is a form of (2.1a) *in which*

$$d_1(u, v, \tilde{u}) = u - \tilde{u} \qquad \text{and} \qquad d_2(u, v, \tilde{u}) = \beta F(\tilde{u}). \tag{2.3}$$

Since $\tilde{u} \in \Omega$, using the monotonicity of F, we have

$$(\tilde{u} - u^*)^T d_2(u, v, \tilde{u}) = (\tilde{u} - u^*)^T \beta F(\tilde{u}) \geq (\tilde{u} - u^*)^T \beta F(u^*), \quad \forall\, u^* \in \Omega^*.$$

Because $u^* \in \Omega^*$, we have

$$(\tilde{u} - u^*)^T \beta F(u^*) \geq 0 \qquad \textit{and thus} \qquad (\tilde{u} - u^*)^T d_2(u, v, \tilde{u}) \geq 0.$$

By setting

$$\varphi(u, v, \tilde{u}) = \phi(u, v, \tilde{u}) = \|u - \tilde{u}\|^2, \tag{2.4}$$

it is easy to check that conditions (2.1b)*,* (2.1c) *and* (2.1d) *(with $\kappa = 1$) are satisfied.*

In practice, we use the following procedure to find an effective quadruplet and the related accepting rule in the unified framework.

The Unified Framework
—— Procedure in finding an effective quadruplet and an accepting rule

1. Rewrite the basic equation of APPAs (1.7) to
   $$\tilde{u}^k = P_\Omega\{\tilde{u}^k - [d_2(u^k, v^k, \tilde{u}^k) - Gd_1(u^k, v^k, \tilde{u}^k)]\}.$$
2. With $d_1(u^k, v^k, \tilde{u}^k)$ and $d_2(u^k, v^k, \tilde{u}^k)$, define a function $\varphi(u^k, v^k, \tilde{u}^k)$ which satisfies
   $$(\tilde{u}^k - u^*)^T d_2(u^k, v^k, \tilde{u}^k) \geq \varphi(u^k, v^k, \tilde{u}^k) - (u^k - \tilde{u}^k)^T G d_1(u^k, v^k, \tilde{u}^k), \quad \forall u^* \in \Omega.$$
3. Find an accepting rule and a non-negative function $\phi(u^k, v^k, \tilde{u}^k)$ which satisfies
   $$2\varphi(u^k, v^k, \tilde{u}^k) \geq \|d_1(u^k, v^k, \tilde{u}^k)\|_G^2 + \phi(u^k, v^k, \tilde{u}^k).$$
4. Verify that there is a positive constant $\kappa > 0$ such that
   $$\kappa \|u^k - \tilde{u}^k\|^2 \leq \phi(u^k, v^k, \tilde{u}^k).$$

2.2 The descent directions and the primary methods

For any $u^* \in \Omega^*$, $G(u - u^*)$ is the gradient of the unknown distance function $\frac{1}{2}\|u - u^*\|_G^2$ at the point u. A direction d is called the descent direction of $\|u - u^*\|_G^2$ if and only if $(G(u - u^*))^T d < 0$. The following lemmas reveal that both $-d_1(u, v, \tilde{u})$ and $-d_2(u, v, \tilde{u})$ in the effective quadruplet are descent directions of the unknown distance function $\|u - u^*\|_G^2$ and $\|u - u^*\|^2$ when $u \in \Omega \setminus \Omega^*$, respectively. The assertions are similar to Lemmas 3.1 and 3.2 in [21]. For completeness, the proofs are provided.

Lemma 2.1 *If conditions* (2.1a) *and* (2.1b) *are satisfied, then*

$$(u - u^*)^T G d_1(u, v, \tilde{u}) \geq \varphi(u, v, \tilde{u}), \quad \forall u^* \in \Omega^*. \tag{2.5}$$

Proof. Since $u^* \in \Omega$, it follows from (2.2) that

$$(\tilde{u} - u^*)^T G d_1(u, v, \tilde{u}) \geq (\tilde{u} - u^*)^T d_2(u, v, \tilde{u}), \quad \forall u^* \in \Omega^*.$$

Substituting the right-hand-side of the above inequality by (2.1b), we obtain

$$(\tilde{u} - u^*)^T G d_1(u, v, \tilde{u}) \geq \varphi(u, v, \tilde{u}) - (u - \tilde{u})^T G d_1(u, v, \tilde{u}), \quad \forall u^* \in \Omega^*. \tag{2.6}$$

Assertion (2.5) follows from the above inequality directly. $\square$

Lemma 2.2 *If conditions* (2.1a) *and* (2.1b) *are satisfied, then*

$$(u-u^*)^T d_2(u,v,\tilde{u}) \ge \varphi(u,v,\tilde{u}), \quad \forall u \in \Omega, u^* \in \Omega^*. \tag{2.7}$$

Proof. Adding (2.2) and (2.1b), we obtain

$$(u'-u^*)^T d_2(u,v,\tilde{u}) \ge \varphi(u,v,\tilde{u}) + (u'-u)^T G d_1(u,v,\tilde{u}), \quad \forall u' \in \Omega, u^* \in \Omega^*. \tag{2.8}$$

Assertion (2.7) follows from the above inequality directly by setting $u' = u$. □

From Lemmas 2.1 and 2.2, the effective quadruplet provides geminate descent directions for contraction methods with very similar properties. However, there are two important differences:

- Condition (2.1c) implies that $\|d_1(u,v,\tilde{u})\| \to 0$ as $\varphi(u,v,\tilde{u}) \to 0$, while it does not ensure the same property for $\|d_2(u,v,\tilde{u})\|$.
- The assertion holds for all $u \in R^n$ in Lemma 2.1, while it is true for all $u \in \Omega$ in Lemma 2.2.

Definition 2.2 *According to the effective quadruplet, a method is called primary method when the new iterate* u^{new} *is generated by one of the following equations. The first kind updating form is,* $\forall G$,

$$u^{\text{new}} = u - d_1(u,v,\tilde{u}). \tag{2.9a}$$

And the other two kinds are,

$$u^{\text{new}} = P_\Omega[u - d_1(u,v,\tilde{u})], \quad G = I, \tag{2.9b}$$

$$u^{\text{new}} = P_\Omega[u - d_2(u,v,\tilde{u})], \quad G = I. \tag{2.9c}$$

Note that the first kind primary methods (2.9a) need no additional projection.

Remark 2.4 *In the exact PPA, the new iterate* u^{k+1} *is given by* $\tilde{u}^k$. *Note that*

$$\tilde{u}^k = u^k - (u^k - \tilde{u}^k) = P_\Omega[u^k - \beta_k F(\tilde{u}^k)].$$

With the directions $d_1(u,v,\tilde{u})$ *and* $d_2(u,v,\tilde{u})$ *defined in Remark 2.3, the new iterate of the exact PPA methods is generated by the primary methods.*

For the primary methods, the following results are straightforward consequences of Lemmas 2.1 and 2.2.

Proposition 2.1 *Let conditions* (2.1a)-(2.1c) *be satisfied and the new iterate be generated by* (2.9a) *or* (2.9b). *Then we have*

$$\|u^{\text{new}} - u^*\|_G^2 \le \|u-u^*\|_G^2 - \phi(u,v,\tilde{u}), \quad \forall u^* \in \Omega^*. \tag{2.10}$$

Proof: It follows from (1.10), (2.9a) and (2.9b) that

$$\|u^{\text{new}} - u^*\|_G^2 \le \|u - d_1(u,v,\tilde{u}) - u^*\|_G^2.$$

Consequently, applying (2.5) and (2.1c), we have

$$\begin{aligned} \|u^{\text{new}} - u^*\|_G^2 &\le \|u - d_1(u,v,\tilde{u}) - u^*\|_G^2 \\ &= \|u-u^*\|_G^2 - 2(u-u^*)^T G d_1(u,v,\tilde{u}) + \|d_1(u,v,\tilde{u})\|_G^2 \\ &\le \|u-u^*\|_G^2 - \phi(u,v,\tilde{u}). \end{aligned}$$

The assertion is proved. □

Proposition 2.2 *For $G = I$, let conditions* (2.1a)-(2.1c) *be satisfied and the new iterate be generated by* (2.9c)*. Then we have*

$$\|u^{\text{new}} - u^*\|^2 \le \|u - u^*\|^2 - \phi(u, v, \tilde{u}), \quad \forall u^* \in \Omega^*. \tag{2.11}$$

Proof: Since $u^* \in \Omega$, it follows from (2.9c) and (1.11) that

$$\begin{aligned} \|u^{\text{new}} - u^*\|^2 &\le \|u - d_2(u, v, \tilde{u}) - u^*\|^2 - \|u - d_2(u, v, \tilde{u}) - u^{\text{new}}\|^2 \\ &= \|u - u^*\|^2 - 2(u^{\text{new}} - u^*)^T d_2(u, v, \tilde{u}) - \|u - u^{\text{new}}\|^2. \end{aligned} \tag{2.12}$$

Since $u^{\text{new}} \in \Omega$, it follows from (2.8) that

$$(u^{\text{new}} - u^*)^T d_2(u, v, \tilde{u}) \ \ge \ \varphi(u, v, \tilde{u}) + (u^{\text{new}} - u)^T d_1(u, v, \tilde{u}).$$

Substituting it in the right-hand-side of (2.12) and using (2.1c), we get

$$\begin{aligned} \|u^{\text{new}} - u^*\|^2 &\le \|u - u^*\|^2 - 2\varphi(u, v, \tilde{u}) - 2(u^{\text{new}} - u)^T d_1(u, v, \tilde{u}) - \|u - u^{\text{new}}\|^2 \\ &= \|u - u^*\|^2 - 2\varphi(u, v, \tilde{u}) + \|d_1(u, v, \tilde{u})\|^2 - \|u - u^{\text{new}} - d_1(u, v, \tilde{u})\|^2 \\ &\le \|u - u^*\|^2 - \phi(u, v, \tilde{u}). \end{aligned}$$

and the assertion is proved. □

2.3 Convergence of the primary methods

For the convergence of the primary methods, we need the following **additional conditions**: The geminate directions $d_1(u, v, \tilde{u})$ and $d_2(u, v, \tilde{u})$ in the effective quadruplet satisfy

$$\lim_{k\to\infty} d_1(u^k, v^k, \tilde{u}^k) = 0 \tag{2.13a}$$

and

$$\lim_{k\to\infty} \{d_2(u^k, v^k, \tilde{u}^k) - \beta_k F(\tilde{u}^k)\} = 0. \tag{2.13b}$$

Theorem 2.1 *For given $u^k \in \Omega$ and $\beta_k \ge \beta_L > 0$, let $v^k \in \Omega$ be an approximate solution of* (1.3) *with certain accepting rule. Assume that the quadruplet $\big(d_1(u^k, v^k, \tilde{u}^k)$, $d_2(u^k, v^k, \tilde{u}^k)$, $\varphi(u^k, v^k, \tilde{u}^k)$, $\phi(u^k, v^k, \tilde{u}^k)\big)$ is effective and the sequence $\{u^k\}$ is generated by the primary methods. If the additional conditions* (2.13) *are satisfied, then $\{u^k\}$ converges to some u^∞ which is a solution point of VI(Ω, F).*

Proof. According to Propositions 2.1 and 2.2, for the sequence generated by the primary methods, we have

$$\|u^{k+1} - u^*\|_Q^2 \le \|u^k - u^*\|_Q^2 - \phi(u^k, v^k, \tilde{u}^k), \quad \forall\, u^* \in \Omega^*, \tag{2.14}$$

where $Q = G$ or I. It follows that

$$\lim_{k\to\infty} \phi(u^k, v^k, \tilde{u}^k) = 0$$

and $\{u^k\}$ is bounded. Together with condition (2.1d), $\{\tilde{u}^k\}$ is also bounded. Let u^∞ be a cluster point of $\{\tilde{u}^k\}$ and $\{\tilde{u}^{k_j}\}$ is a subsequence which converges to u^∞. Condition (2.1a) means that

$$\tilde{u}^{k_j} \in \Omega, \quad (u' - \tilde{u}^{k_j})^T \big\{d_2(u^{k_j}, v^{k_j}, \tilde{u}^{k_j}) - G d_1(u^{k_j}, v^{k_j}, \tilde{u}^{k_j})\big\} \ge 0, \qquad \forall\, u' \in \Omega.$$

Since $0 < \inf_{k=0}^{\infty} \beta_k \le \sup_{k=0}^{\infty} \beta_k < +\infty$, it follows from the continuity of F and the additional conditions that

$$u^\infty \in \Omega, \quad (u' - u^\infty)^T F(u^\infty) \ge 0, \qquad \forall\, u' \in \Omega.$$

The above variational inequality indicates that u^∞ is a solution of VI(Ω, F). Because $\lim_{k\to\infty} \|u^k - \tilde{u}^k\|_M = 0$ and $\lim_{j\to\infty} \tilde{u}^{k_j} = u^\infty$, the subsequence $\{u^{k_j}\}$ converges to u^∞. Due to (2.14), we have

$$\|u^{k+1} - u^\infty\|_Q \le \|u^k - u^\infty\|_Q$$

and $\{u^k\}$ converges to u^∞. □

Remark 2.5 *For the exact PPA which can be viewed as a primary method with the quadruplet defined in Remark 2.3, it is easy to check that the additional conditions are satisfied.*

3 Application to symmetric monotone linear VIs

From the basic equation of APPAs, this section derives the unified framework for symmetric linear variational inequality

$$u \in \Omega, \quad (u' - u)^T (Hu + q) \geq 0, \quad \forall\, u' \in \Omega, \tag{3.1}$$

where $H \in R^{n \times n}$ is symmetric positive semi-definite and $q \in R^n$. Symmetric linear variational inequality is equivalent to the quadratic programming

$$\min\{\frac{1}{2}u^T Hu + q^T u \,|\, u \in \Omega\}. \tag{3.2}$$

Since $F(u) = Hu + q$, the basic equation of form (1.7) is

$$\tilde{u} = P_\Omega[u - \beta(Hv + q)]. \tag{3.3}$$

Under the unified framework, we will find an effective quadruplet $\big(d_1(u, v, \tilde{u}),\, d_2(u, v, \tilde{u}),\, \varphi(u, v, \tilde{u}),\, \phi(u, v, \tilde{u})\big)$ and its related accepting rule which satisfy conditions (2.1) with $G = I$.

Condition (2.1a): The basic equation (3.3) can be rewritten as

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

By setting

$$d_1(u, v, \tilde{u}) = u - \tilde{u} \tag{3.4}$$

and

$$d_2(u, v, \tilde{u}) = \beta(Hv + q), \tag{3.5}$$

the geminate directions $(d_1(u, v, \tilde{u}), d_2(u, v, \tilde{u}))$ satisfy condition (2.1a).

Condition (2.1b): Since $\tilde{u} \in \Omega$ and H is positive semi-definite, we have

$$(\tilde{u} - u^*)^T \beta(Hu^* + q) \geq 0, \qquad \forall\, u^* \in \Omega^*$$

and it can be rewritten as

$$(\tilde{u} - u^*)^T \beta(Hv + q) \geq (\tilde{u} - u^*)^T \beta H(v - u^*), \qquad \forall\, u^* \in \Omega^*.$$

Because H is symmetric, using $x^T Hy \geq -\frac{1}{2}(x - y)^T H(x - y)$ to the right-hand-side of the above inequality, we get

$$(\tilde{u} - u^*)^T \beta(Hv + q) \geq -\frac{1}{2}(v - \tilde{u})^T \beta H(v - \tilde{u}), \qquad \forall\, u^* \in \Omega^*. \tag{3.6}$$

Note that the left-hand-side of above inequality is $(\tilde{u} - u^*)^T d_2(u, v, \tilde{u})$, by defining

$$\varphi(u, v, \tilde{u}) = \|u - \tilde{u}\|^2 - \frac{1}{2}(v - \tilde{u})^T \beta H(v - \tilde{u}), \tag{3.7}$$

the triplet $\big(d_1(u, v, \tilde{u}), d_2(u, v, \tilde{u}), \varphi(u, v, \tilde{u})\big)$ satisfies condition (2.1b).

Condition (2.1c): Since $d_1(u, v, \tilde{u}) = u - \tilde{u}$, using the following accepting rule

$$(\textbf{Accpting rule}) \qquad (v - \tilde{u})^T \beta H(v - \tilde{u}) \leq \nu \|u - \tilde{u}\|^2, \quad \nu \in (0, 1) \tag{3.8}$$

and defining

$$\phi(u, v, \tilde{u}) = (1-\nu)\|u - \tilde{u}\|^2, \tag{3.9}$$

we get

$$2\varphi(u, v, \tilde{u}) \geq \|d_1(u, v, \tilde{u})\|^2 + \phi(u, v, \tilde{u}),$$

thus condition (2.1c) is satisfied.

Condition (2.1d): This follows from (3.9) directly with $\kappa = (1-\nu)$.

Based on the above verification, we have proved the following theorem.

Theorem 3.1 *For solving problem* (3.1)*, let the triplet* $(u, v, \tilde{u})$ *be defined in* (3.3) *and the accepting rule* (3.8) *be satisfied. Then the quadruplet* $\big(d_1(u, v, \tilde{u}),\ d_2(u, v, \tilde{u}),\ \varphi(u, v, \tilde{u}),\ \phi(u, v, \tilde{u})\big)$ *given by*

$$(\textbf{The quadruplet}) \qquad \left\{ \begin{array}{rcl} d_1(u, v, \tilde{u}) & = & u - \tilde{u}, \\ d_2(u, v, \tilde{u}) & = & \beta(Hv + q), \\ \varphi(u, v, \tilde{u}) & = & \|u - \tilde{u}\|^2 - \frac{1}{2}(v - \tilde{u})^T \beta H (v - \tilde{u}), \\ \phi(u, v, \tilde{u}) & = & (1-\nu)\|u - \tilde{u}\|^2 \end{array} \right. \tag{3.10}$$

is an effective quadruplet which fulfills conditions (2.1) *with* $G = I$.

According to Theorem 2.1, for the convergence of the primary methods, we need only to verify the additional conditions (2.13).

Theorem 3.2 *Let the conditions in Theorem 3.1 be satisfied. Then the sequence* $\{u^k\}$ *generated by the primary methods converges to some* u^∞ *which is a solution point of problem* (3.1).

Proof. Since $\phi(u^k, v^k, \tilde{u}^k) = (1-\nu)\|u^k - \tilde{u}^k\|^2$, it follows from Propositions 2.1 and 2.2 that

$$\|u^{k+1} - u^*\|^2 \leq \|u^k - u^*\|^2 - (1-\nu)\|u^k - \tilde{u}^k\|^2$$

and thus

$$\lim_{k\to\infty} \|u^k - \tilde{u}^k\| = 0. \tag{3.11}$$

Because $d_1(u^k, v^k, \tilde{u}^k) = (u^k - \tilde{u}^k)$, it follows that

$$\lim_{k\to\infty} d_1(u^k, v^k, \tilde{u}^k) = 0$$

and condition (2.13a) holds. Note that

$$d_2(u^k, v^k, \tilde{u}^k) - \beta F(\tilde{u}^k) = d_2(u^k, v^k, \tilde{u}^k) - \beta(H\tilde{u}^k + q) = \beta H(v^k - \tilde{u}^k).$$

From the symmetry and positive semi-definiteness of H, the accepting rule (3.8) and (3.11), we obtain

$$\lim_{k\to\infty} H(v^k - \tilde{u}^k) = 0$$

and thus condition (2.13b) is satisfied. The proof is complete. □

Remark 3.1 *Instead of* (3.3) *and* (3.8), *in the k-th iteration, we can set*

$$\tilde{u}^k = P_\Omega[u^k - \beta_k(Hu^k + q)] \tag{3.12a}$$

by choosing a suitable β_k, *such that*

$$(u^k - \tilde{u}^k)^T \beta_k H(u^k - \tilde{u}^k) \le \nu \|u^k - \tilde{u}^k\|^2, \quad \nu \in (0,1). \tag{3.12b}$$

In other words, if the parameter β_k *is small enough, the accepting rule will be satisfied even if* $v^k = u^k$ *is taken as the approximate solution in the k-th iteration. In this case, the geminate directions are given by*

$$d_1(u^k, v^k, \tilde{u}^k) = u^k - \tilde{u}^k \qquad \textit{and} \qquad d_2(u^k, v^k, \tilde{u}^k) = \beta_k H(u^k + q). \tag{3.12c}$$

A small β will guarantee the accepting rule (3.12b) to be satisfied. However, a too small positive parameter β will lead to slow convergence. It should be noticed that, in practical computation, the increase of parameter β is necessary when

$$(u^k - \tilde{u}^k)^T \beta_k H(u^k - \tilde{u}^k) \ll \|u^k - \tilde{u}^k\|^2.$$

Hence, we suggest to use the following procedure for finding a suitable parameter β_k.

Procedure 3.1 Finding β_k to satisfy (3.12b). $\beta_0 = 1$, $\nu = 0.9$.

REPEAT: $\tilde{u}^k = P_\Omega[u^k - \beta_k(Hu^k + q)]$.

If $r_k := \dfrac{(u^k - \tilde{u}^k)^T \beta_k H(u^k - \tilde{u}^k)}{\|u^k - \tilde{u}^k\|^2} \ge \nu$, $\quad \beta_k := \beta_k * 0.7 * \min\{1, \frac{1}{r_k}\}$.

UNTIL: $(u^k - \tilde{u}^k)^T \beta_k H(u^k - \tilde{u}^k) \le \nu \|u^k - \tilde{u}^k\|^2$. (Accepting rule (3.12b))

ADJUST: $\beta_{k+1} := \begin{cases} \beta_k * \nu * 0.9 / r_k & \text{if } r_k \le 0.3, \\ \beta_k & \text{otherwise.} \end{cases}$

Remark 3.2 *When* $\Omega = R^n$, *problem* (3.2) *is reduced to an unconstrained convex quadratic optimization problem. In this case, the recursion of the steepest descent method is*

$$u^{k+1} = u^k - \alpha_k^{\mathrm{SD}}(Hu^k + q), \tag{3.13}$$

where

$$\alpha_k^{\mathrm{SD}} = \frac{\|Hu^k + q\|^2}{(Hu^k + q)^T H(Hu^k + q)} \tag{3.14}$$

is the step size in the steepest descent method. Most recently, from numerous tests we are surprised that the numerical performance is improved significantly via scaling the step-sizes α_k^{SD} *simply by a multiplier in [0.3, 0.9]. The experiments coincide with the observations of Dai and Yuan [3].*

For unconstrained convex quadratic programming, the recursion of the method with selected parameter β_k *in this section is*

$$u^{k+1} = u^k - \beta_k(Hu^k + q). \tag{3.15}$$

Since β_k *satisfies condition* (3.12b). *By a simple manipulation,*

$$\beta_k \le \frac{\nu \|Hu^k + q\|^2}{(Hu^k + q)^T H(Hu^k + q)}. \tag{3.16}$$

The restriction $0.3 \le r_k \le 0.9$ *in Procedure 3.1 leads to*

$$\beta_k \in [0.3\alpha_k^{\mathrm{SD}}, 0.9\alpha_k^{\mathrm{SD}}].$$

Therefore, the methods can be viewed as the extension of the steepest descent methods with shortened step size to constrained convex quadratic optimization.

4 Application to asymmetric monotone linear VIs

From the basic equation of APPAs, this section turns to the linear variational inequality (without symmetry)

$$u \in \Omega, \quad (u' - u)^T (Mu + q) \geq 0, \quad \forall\, u' \in \Omega, \tag{4.1}$$

where $M \in R^{n\times n}$ is positive semi-definite (but not necessarily symmetric) and $q \in R^n$. Since $F(u) = Mu + q$, the basic equation of form (1.7) is

$$\tilde{u} = P_\Omega[u - \beta(Mv + q)]. \tag{4.2}$$

For the triplet $(u, v, \tilde{u})$ in (4.2), under the unified framework, we will find an effective quadruplet $\big(d_1(u, v, \tilde{u}),\ d_2(u, v, \tilde{u}),\ \varphi(u, v, \tilde{u}),\ \phi(u, v, \tilde{u})\big)$ and its related accepting rule which satisfy conditions (2.1) with $G = I$.

Condition (2.1a): Equation (4.2) can be written as

$$\tilde{u} = P_\Omega\{\tilde{u} - \{[\beta(Mv + q) + \beta M^T(v - \tilde{u})] - [(u - \tilde{u}) + \beta M^T(v - \tilde{u})]\}\}. \tag{4.3}$$

By defining

$$d_1(u, v, \tilde{u}) = (u - \tilde{u}) + \beta M^T(v - \tilde{u}) \tag{4.4}$$

and

$$d_2(u, v, \tilde{u}) = \beta(Mv + q) + \beta M^T(v - \tilde{u}), \tag{4.5}$$

the geminate directions ($d_1(u, v, \tilde{u})$ and $d_2(u, v, \tilde{u})$) satisfy condition (2.1a).

Condition (2.1b): Since $\tilde{u} \in \Omega$, we have

$$(\tilde{u} - u^*)^T\beta(Mu^* + q) \geq 0, \qquad \forall\, u^* \in \Omega^*$$

and consequently

$$\begin{aligned}(\tilde{u} - u^*)^T\beta(Mv + q) &\geq (\tilde{u} - u^*)^T\beta M(v - u^*)\\ &= (v - u^*)^T\beta M^T(\tilde{u} - u^*). \quad \forall\, u^* \in \Omega^*.\end{aligned} \tag{4.6}$$

On the other hand, we have the identity

$$(\tilde{u} - u^*)^T\beta M^T(v - \tilde{u}) = (v - u^*)^T\beta M^T(v - \tilde{u}) - (v - \tilde{u})^T\beta M^T(v - \tilde{u}). \tag{4.7}$$

Adding (4.6) and (4.7) and using $(v - u^*)^T\beta M^T(v - u^*) \geq 0$, we obtain

$$(\tilde{u} - u^*)^T\{\beta(Mv + q) + \beta M^T(v - \tilde{u})\} \geq -(v - \tilde{u})^T\beta M^T(v - \tilde{u}). \tag{4.8}$$

Note that the left-hand-side of above inequality is $(\tilde{u} - u^*)^T d_2(u, v, \tilde{u})$. By defining

$$\varphi(u, v, \tilde{u}) = (u - \tilde{u})^T d_1(u, v, \tilde{u}) - (v - \tilde{u})^T\beta M^T(v - \tilde{u}), \tag{4.9}$$

inequality (4.8) becomes

$$(\tilde{u} - u^*)^T d_2(u, v, \tilde{u}) \geq \ \varphi(u, v, \tilde{u}) - (u - \tilde{u})^T d_1(u, v, \tilde{u})$$

and thus the triplet $\big(d_1(u, v, \tilde{u}), d_2(u, v, \tilde{u}), \varphi(u, v, \tilde{u})\big)$ satisfies condition (2.1b).

Condition (2.1c): For $d_1(u, v, \tilde{u})$ defined in (4.4) and $\varphi(u, v, \tilde{u})$ defined in (4.9), by a straightforward manipulation, we get

$$\begin{aligned}
&2\varphi(u, v, \tilde{u}) - \|d_1(u, v, \tilde{u})\|^2 \\
&= 2(u - \tilde{u})^T d_1(u, v, \tilde{u}) - \|d_1(u, v, \tilde{u})\|^2 - 2(v - \tilde{u})^T \beta M^T (v - \tilde{u}) \\
&= d_1(u, v, \tilde{u})^T \big(2(u - \tilde{u}) - d_1(u, v, \tilde{u})\big) - 2(v - \tilde{u})^T \beta M^T (v - \tilde{u}) \\
&= \big((u - \tilde{u}) + \beta M^T (v - \tilde{u})\big)^T \big((u - \tilde{u}) - \beta M^T (v - \tilde{u})\big) - 2(v - \tilde{u})^T \beta M^T (v - \tilde{u}) \\
&= \|u - \tilde{u}\|^2 - \{2(v - \tilde{u})^T \beta M^T (v - \tilde{u}) + \|\beta M^T (v - \tilde{u})\|^2\}. \qquad (4.10)
\end{aligned}$$

Let the approximate solution v be accepted when the accepting rule

$$(\textbf{Accepting rule}) \qquad 2(v-\tilde{u})^T \beta M^T (v-\tilde{u}) + \|\beta M^T (v-\tilde{u})\|^2 \le \nu \|u-\tilde{u}\|^2, \quad \nu \in (0,1) \qquad (4.11)$$

is satisfied and let

$$\phi(u, v, \tilde{u}) := (1 - \nu)\|u - \tilde{u}\|^2. \qquad (4.12)$$

It follows from (4.10), (4.11) and (4.12) that

$$2\varphi(u, v, \tilde{u}) \ge \|d_1(u, v, \tilde{u})\|^2 + \phi(u, v, \tilde{u})$$

and thus condition (2.1c) is satisfied.

Condition (2.1d): From (4.12) condition (2.1d) holds with $\kappa = (1 - \nu)$.

Based on the above verification, we have proved the following theorem.

Theorem 4.1 *For solving problem* (4.1)*, let the triplet* $(u, v, \tilde{u})$ *be defined in* (4.2) *and the accepting rule* (4.11) *be satisfied. Then the quadruplet* $\big(d_1(u, v, \tilde{u}),\ d_2(u, v, \tilde{u}),\ \varphi(u, v, \tilde{u}),\ \phi(u, v, \tilde{u})\big)$ *given by*

$$(\textbf{The quadruplet}) \qquad \left\{ \begin{array}{rcl} d_1(u, v, \tilde{u}) &=& (u - \tilde{u}) + \beta M^T (v - \tilde{u}), \\ d_2(u, v, \tilde{u}) &=& \beta(Mv + q) + \beta M^T (v - \tilde{u}), \\ \varphi(u, v, \tilde{u}) &=& (u - \tilde{u})^T d_1(u, v, \tilde{u}) - (v - \tilde{u})^T \beta M^T (v - \tilde{u}), \\ \phi(u, v, \tilde{u}) &=& (1 - \nu)\|u - \tilde{u}\|^2 \end{array} \right. \qquad (4.13)$$

is an effective quadruplet which fulfills conditions (2.1) *with* $G = I$.

According to Theorem 2.1, for the convergence of the primary methods, we need only to verify the additional conditions (2.13).

Theorem 4.2 *Let the conditions in Theorem 4.1 be satisfied. Then the sequence* $\{u^k\}$ *generated by the primary methods converges to some* u^∞ *which is a solution point of problem* (4.1).

Proof. We need only to verify the additional conditions (2.13). Since $\phi(u^k, v^k, \tilde{u}^k) = (1 - \nu)\|u^k - \tilde{u}^k\|^2$, it follows from Propositions 2.1 and 2.2 that

$$\|u^{k+1} - u^*\|^2 \le \|u^k - u^*\|^2 - (1 - \nu)\|u^k - \tilde{u}^k\|^2$$

and thus

$$\lim_{k\to\infty} \|u^k - \tilde{u}^k\| = 0. \qquad (4.14)$$

Since M is positive semi-definite, it follows from the accepting rule (4.11) that

$$\|\beta M^T (v^k - \tilde{u}^k)\| \le \nu \|u^k - \tilde{u}^k\|^2 \qquad (4.15)$$

Because $d_1(u^k, v^k, \tilde{u}^k) = (u^k - \tilde{u}^k) + \beta M^T(v^k - \tilde{u}^k)$, it follows from (4.14) and (4.15) that

$$\lim_{k\to+\infty} d_1(u^k, v^k, \tilde{u}^k) = 0$$

and condition (2.13a) holds. Note that

$$d_2(u^k, v^k, \tilde{u}^k) - \beta(M\tilde{u}^k + q) = \beta(M + M^T)(v^k - \tilde{u}^k).$$

It follow from the accepting rule (4.11)

$$\beta(v^k - \tilde{u}^k)^T(M + M^T)(v^k - \tilde{u}^k) \le \nu\|u^k - \tilde{u}^k\|^2. \tag{4.16}$$

From the symmetry and positive semi-definiteness of $(M + M^T)$, (4.14) and (4.16), we obtain

$$\lim_{k\to\infty}(M + M^T)(v^k - \tilde{u}^k) = 0$$

and thus condition (2.13b) is satisfied. □

Remark 4.1 *Similar to Remark 3.2, we can take $v^k = u^k$ as the approximate solution in the k-th iteration. The accepting rule*

$$2(u^k - \tilde{u}^k)\beta_k M^T(u^k - \tilde{u}^k) + \|\beta_k M^T(u^k - \tilde{u}^k)\|^2 \le \nu\|u^k - \tilde{u}^k\|^2, \quad \nu \in (0, 1). \tag{4.17}$$

will be satisfied when the parameter β_k is small enough.

Reconsidering that a too small positive parameter β will lead to slow convergence, similarly to Procedure 3.1, we will increase the parameter β when

$$2(u^k - \tilde{u}^k)\beta_k M^T(u^k - \tilde{u}^k) + \|\beta_k M^T(u^k - \tilde{u}^k)\|^2 \le \nu'\|u^k - \tilde{u}^k\|^2,$$

where ν' is relative smaller than ν. In order to find a suitable parameter β_k, the following procedure is recommended, in which we set $\nu' = 0.3$ based on our numerical experiments.

Procedure 4.1 Finding β_k to satisfy (4.17). $\beta_0 = 1$, $\nu = 0.9$.

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

If $$r_k := \frac{2(u^k - \tilde{u}^k)\beta_k M^T(u^k - \tilde{u}^k) + \|\beta_k M^T(u^k - \tilde{u}^k)\|^2}{\|u^k - \tilde{u}^k\|^2} \ge \nu,$$

$\beta_k := \beta_k * 0.7 * \min\{1, \frac{1}{r_k}\}$.

UNTIL: $2(u^k - \tilde{u}^k)\beta_k M^T(u^k - \tilde{u}^k) + \|\beta_k M^T(u^k - \tilde{u}^k)\|^2 \le \nu\|u^k - \tilde{u}^k\|^2$.

(Accepting rule (4.17))

ADJUST: $$\beta_{k+1} := \begin{cases} \beta_k * \nu * 0.9/r_k & \text{if } r_k \le 0.3, \\ \beta_k & \text{otherwise.} \end{cases}$$

5 Application to symmetric monotone nonlinear VIs

In VI(Ω, F), when $F(u)$ is the gradient of certain function, say $f(u)$, the Jacobian matrix of $F(u)$ (if it exists) is symmetric. In this sense, we call the related VI(Ω, F) as the symmetric VIs. In other words, the symmetric monotone nonlinear variational inequality is equivalent to the differentiable convex optimization

$$\min\{\frac{1}{2}f(u) \,|\, u \in \Omega\}. \tag{5.1}$$

Since $F(u) = \nabla f(u)$, a basic property of the differentiable convex function is

$$f(v) \geq f(u) + (v-u)^T F(u), \quad \forall u, v \in R^n. \tag{5.2}$$

The basic equation of form (1.7) is

$$\tilde{u} = P_\Omega[u - \beta F(v)]. \tag{5.3}$$

Under the unified framework, we will find an effective quadruplet $\big(d_1(u,v,\tilde{u}),\, d_2(u,v,\tilde{u}),\, \varphi(u,v,\tilde{u}),$ $\phi(u,v,\tilde{u})\big)$ and its related accepting rule which satisfy conditions (2.1) with $G = I$.

Condition (2.1a): The basic equation (5.3) can be rewritten as

$$\tilde{u} = P_\Omega\{\tilde{u} - [\beta F(v) - (u - \tilde{u})].$$

By setting

$$d_1(u,v,\tilde{u}) = u - \tilde{u} \tag{5.4}$$

and

$$d_2(u,v,\tilde{u}) = \beta F(v), \tag{5.5}$$

the geminate directions $(d_1(u,v,\tilde{u}), d_2(u,v,\tilde{u}))$ satisfy condition (2.1a).

Condition (2.1b): Using the basic property of the differentiable convex function (5.2), we have $f(u^*) \geq f(v) + (u^* - v)^T F(v)$ and thus

$$(v - u^*)F(v) \geq f(v) - f(u^*).$$

Since $\tilde{u} \in \Omega$ and u^* is a solution of the convex optimization problem (5.1), thus $f(\tilde{u}) \geq f(u^*)$ and consequently

$$(v - u^*)F(v) \geq f(v) - f(\tilde{u}) \geq (v - \tilde{u})^T F(\tilde{u}), \qquad \forall\, u^* \in \Omega^*.$$

Again, the last inequality follows from (5.2). Adding $(\tilde{u} - v)^T F(v)$ to the both sides of the above inequality and multiplying the positive factor β, we get

$$(\tilde{u} - u^*)^T \beta F(v) \geq -(v - \tilde{u})^T \beta (F(v) - F(\tilde{u})), \qquad \forall\, u^* \in \Omega^*. \tag{5.6}$$

Note that the left-hand-side of above inequality is $(\tilde{u} - u^*)^T d_2(u,v,\tilde{u})$, by defining

$$\varphi(u,v,\tilde{u}) = \|u - \tilde{u}\|^2 - (v - \tilde{u})^T \beta (F(v) - F(\tilde{u})), \tag{5.7}$$

the triplet $\big(d_1(u,v,\tilde{u}), d_2(u,v,\tilde{u}), \varphi(u,v,\tilde{u})\big)$ satisfies condition (2.1b).

Condition (2.1c): Since $d_1(u,v,\tilde{u}) = u - \tilde{u}$, using the following accepting rule

$$(\textbf{Accpting rule}) \qquad (v - \tilde{u})^T 2\beta (F(v) - F(\tilde{u})) \leq \nu \|u - \tilde{u}\|^2, \quad \nu \in (0,1) \tag{5.8}$$

and defining

$$\phi(u,v,\tilde{u}) = (1 - \nu)\|u - \tilde{u}\|^2, \tag{5.9}$$

we get

$$2\varphi(u,v,\tilde{u}) \geq \|d_1(u,v,\tilde{u})\|^2 + \phi(u,v,\tilde{u}),$$

thus condition (2.1c) is satisfied.

Condition (2.1d): This follows from (5.9) directly with $\kappa = (1 - \nu)$.

Based on the above verification, we have proved the following theorem.

Theorem 5.1 *For solving problem* (1.1) *whose mapping* F *is the gradient of certain convex function, let the triplet* $(u, v, \tilde{u})$ *be defined in* (5.3) *and the accepting rule* (5.8) *be satisfied. Then the quadruplet* $\big(d_1(u,v,\tilde{u}),\ d_2(u,v,\tilde{u}),\ \varphi(u,v,\tilde{u}),\ \phi(u,v,\tilde{u})\big)$ *given by*

$$\text{(\textbf{The quadruplet})} \qquad \left\{ \begin{array}{rcl} d_1(u,v,\tilde{u}) &=& u-\tilde{u}, \\ d_2(u,v,\tilde{u}) &=& \beta F(v), \\ \varphi(u,v,\tilde{u}) &=& \|u-\tilde{u}\|^2 - (v-\tilde{u})^T\beta(F(v)-F(\tilde{u})), \\ \phi(u,v,\tilde{u}) &=& (1-\nu)\|u-\tilde{u}\|^2 \end{array} \right. \tag{5.10}$$

is an effective quadruplet which fulfills conditions (2.1) *with* $G = I$.

According to Theorem 2.1, for the convergence of the primary methods, we need only to verify the additional conditions (2.13).

Theorem 5.2 *Let the conditions in Theorem 5.1 be satisfied. Then the sequence* $\{u^k\}$ *generated by the primary methods converges to some* u^∞ *which is a solution point of problem* (1.1) *when* F *is the gradient of certain convex function.*

Proof. Since $\phi(u^k, v^k, \tilde{u}^k) = (1-\nu)\|u^k - \tilde{u}^k\|^2$, it follows from Propositions 2.1 and 2.2 that

$$\|u^{k+1} - u^*\|^2 \le \|u^k - u^*\|^2 - (1-\nu)\|u^k - \tilde{u}^k\|^2$$

and thus

$$\lim_{k\to\infty} \|u^k - \tilde{u}^k\| = 0. \tag{5.11}$$

Because $d_1(u^k, v^k, \tilde{u}^k) = (u^k - \tilde{u}^k)$, it follows that

$$\lim_{k\to+\infty} d_1(u^k, v^k, \tilde{u}^k) = 0$$

and condition (2.13a) holds. Note that

$$d_2(u^k, v^k, \tilde{u}^k) - \beta F(\tilde{u}^k) = \beta(F(v^k) - F(\tilde{u}^k)).$$

From the accepting rule (5.8) and (5.11), we obtain

$$\lim_{k\to\infty} \beta(F(v^k) - F(\tilde{u}^k)) = 0$$

and thus condition (2.13b) is satisfied. The proof is complete. □

Remark 5.1 *Instead of* (5.3) *and* (5.8), *in the* k*-th iteration, we can set*

$$\tilde{u}^k = P_\Omega[u^k - \beta_k F(u^k)] \tag{5.12a}$$

by choosing a suitable β_k, *such that*

$$(u^k - \tilde{u}^k)^T 2\beta_k (F(u^k) - F(\tilde{u}^k)) \le \nu\|u^k - \tilde{u}^k\|^2, \quad \nu \in (0,1). \tag{5.12b}$$

In other words, if F *is Lipschitz continuous, by setting sa suitable small* β_k, *the accepting rule will be satisfied even if* $v^k = u^k$ *is taken as the approximate solution in the* k*-th iteration. In this case, the geminate directions are given by*

$$d_1(u^k, v^k, \tilde{u}^k) = u^k - \tilde{u}^k \qquad \text{and} \qquad d_2(u^k, v^k, \tilde{u}^k) = \beta_k F(u^k). \tag{5.12c}$$

A small β will guarantee the accepting rule (5.12b) to be satisfied. However, a too small positive parameter β will lead to slow convergence. It should be noticed that, in practical computation, the increase of parameter β is necessary when

$$(u^k - \tilde{u}^k)^T 2\beta_k (F(u^k) - F(\tilde{u}^k)) \ll \|u^k - \tilde{u}^k\|^2.$$

Hence, we suggest to use the following procedure for finding a suitable parameter β_k.

Procedure 5.1 Finding β_k to satisfy (5.12b). $\beta_0 = 1$, $\nu = 0.9$.

REPEAT: $\tilde{u}^k = P_\Omega[u^k - \beta_k F(u^k)]$.

If $r_k := \dfrac{(u^k - \tilde{u}^k)^T 2\beta_k (F(u^k) - F(\tilde{u}^k))}{\|u^k - \tilde{u}^k\|^2} \ge \nu$, $\beta_k := \beta_k * 0.7 * \min\{1, \frac{1}{r_k}\}$.

UNTIL: $(u^k - \tilde{u}^k)^T 2\beta_k (F(u^k - F(\tilde{u}^k))\| \le \nu \|u^k - \tilde{u}^k\|^2$. (Accepting rule (5.12b))

ADJUST: $\beta_{k+1} := \begin{cases} \beta_k * \nu * 0.9/r_k & \text{if } r_k \le 0.3, \\ \beta_k & \text{otherwise.} \end{cases}$

6 Application to nonlinear monotone VIs

We consider the nonlinear monotone variational inequality (1.1). For the triplet $(u, v, \tilde{u})$ in (1.7), under the unified framework, we will find an effective quadruplet $\big(d_1(u, v, \tilde{u}),\ d_2(u, v, \tilde{u}),\ \varphi(u, v, \tilde{u}),\ \phi(u, v, \tilde{u})\big)$ and its related accepting rule which satisfy conditions (2.1).

Condition (2.1a): The basic equation (1.7) can be written as

$$\tilde{u} = P_\Omega\{\tilde{u} - \{\beta F(\tilde{u}) - [(u - \tilde{u}) - \beta(F(v) - F(\tilde{u}))]\}\}. \tag{6.1}$$

By setting

$$d_1(u, v, \tilde{u}) = u - \tilde{u} - \beta(F(v) - F(\tilde{u})) \tag{6.2}$$

and

$$d_2(u, v, \tilde{u}) := \beta F(\tilde{u}), \tag{6.3}$$

the geminate directions ($d_1(u, v, \tilde{u})$ and $d_2(u, v, \tilde{u})$) satisfy condition (2.1a).

Condition (2.1b): Since $\tilde{u} \in \Omega$, we have $(\tilde{u} - u^*)^T F(u^*) \ge 0$. Using (6.3) and the monotonicity of F we have

$$(\tilde{u} - u^*)^T d_2(u, v, \tilde{u}) = (\tilde{u} - u^*)^T \beta F(\tilde{u}) \ge 0. \tag{6.4}$$

By letting

$$\varphi(u, v, \tilde{u}) := (u - \tilde{u})^T d_1(u, v, \tilde{u}) \tag{6.5}$$

the triplet $\big(d_1(u, v, \tilde{u}), d_2(u, v, \tilde{u}), \varphi(u, v, \tilde{u})\big)$ satisfies condition (2.1b).

Condition (2.1c): For $d_1(u, v, \tilde{u})$ defined in (6.2) and $\varphi(u, v, \tilde{u})$ defined in (6.5), we have

$$2\varphi(u, v, \tilde{u}) - \|d_1(u, v, \tilde{u})\|^2 = \|u - \tilde{u}\|^2 - \beta^2 \|F(v) - F(\tilde{u})\|^2. \tag{6.6}$$

Let the approximate solution v be accepted when the following rule

$$(\textbf{Accepting rule}) \qquad \beta\|F(v) - F(\tilde{u})\| \le \nu \|u - \tilde{u}\|, \quad \nu \in (0, 1) \tag{6.7}$$

is satisfied. Define

$$\phi(u, v, \tilde{u}) = (1 - \nu^2)\|u - \tilde{u}\|^2. \tag{6.8}$$

Therefore,

$$2\varphi(u,v,\tilde{u}) \ge \|d_1(u,v,\tilde{u})\|^2 + \phi(u,v,\tilde{u})$$

and thus condition (2.1c) is satisfied.

Condition (2.1d): From (6.8) condition (2.1d) holds with $\kappa = (1-\nu^2)$.

Based on the above verification, we have proved the following theorem.

Theorem 6.1 *For solving problem* (1.1)*, let the triplet* $(u,v,\tilde{u})$ *be defined in* (1.7) *and the accepting rule* (6.7) *be satisfied. Then the quadruplet* $\big(d_1(u,v,\tilde{u}),\, d_2(u,v,\tilde{u}),\, \varphi(u,v,\tilde{u}),\, \phi(u,v,\tilde{u})\big)$ *given by*

$$(\textbf{The quadruplet}) \qquad \begin{cases} d_1(u,v,\tilde{u}) &=& u-\tilde{u}-\beta(F(v)-F(\tilde{u})), \\ d_2(u,v,\tilde{u}) &=& \beta F(\tilde{u}), \\ \varphi(u,v,\tilde{u}) &=& (u-\tilde{u})^T d_1(u,v,\tilde{u}), \\ \phi(u,v,\tilde{u}) &=& (1-\nu^2)\|u-\tilde{u}\|^2 \end{cases} \tag{6.9}$$

is an effective quadruplet which fulfills conditions (2.1) *with* $G = I$.

Remark 6.1 *If* F *is Lipschitz continuous, instead of* (1.7) *and* (6.7)*, in the* k*-th iteration, we can set*

$$\tilde{u}^k = P_\Omega[u^k - \beta_k F(u^k)] \tag{6.10a}$$

by choosing a suitable β_k*, such that*

$$\beta_k \|F(u^k) - F(\tilde{u}^k)\| \le \nu \|u^k - \tilde{u}^k\|, \quad \nu \in (0,1). \tag{6.10b}$$

That is, if the parameter β_k *is small enough, the accepting rule will be satisfied even if* $v^k = u^k$ *is taken as the approximate solution in the* k*-th iteration. In this case, the geminate directions are given by*

$$d_1(u^k,v^k,\tilde{u}^k) = u^k - \tilde{u}^k - \beta(F(u^k) - F(\tilde{u}^k)) \qquad \textit{and} \qquad d_2(u^k,v^k,\tilde{u}^k) = \beta F(\tilde{u}^k).$$

There are many primary methods in the literature which can be generated by the quadruplet $\big(d_1(u,v,\tilde{u}), d_2(u,v,\tilde{u}), \varphi(u,v,\tilde{u}), \phi(u,v,\tilde{u})\big)$ in (6.9. To mention a few, see [16, 25, 28, 33]. In fact, the updating form of the prediction correction methods in [16] can be written as

$$u^{k+1} = u^k - d_1(u^k, v^k, \tilde{u}^k). \tag{6.11}$$

The method proposed in [28] is designed mainly for the point-to-set variational inequality. Accordingly, it can be applied to point-to-point variational inequalities. When F is a point-to-point mapping, the hybrid projection-proximal point algorithm [28] can be interpreted as the primary method

$$u^{k+1} = P_\Omega[u^k - d_1(u^k, v^k, \tilde{u}^k)].$$

For $v = u$, the accepting rule (6.7) can be satisfied with a suitable β if F is Lipschitz continuous. By setting $J_{\beta A} = P_\Omega$, the forward-backward splitting method [33] generates the new iterate via

$$(\textbf{FB}) \qquad u_{\rm FB}^{\rm new} = P_\Omega[\tilde{u} + \beta(F(u) - F(\tilde{u}))] \tag{6.12}$$

and it can be viewed as the primary method

$$u^{k+1} = P_\Omega[u^k - d_1(u^k, v^k, \tilde{u}^k)]. \tag{6.13}$$

The extra-gradient method [24, 25] produces the new iterate by

$$\textbf{(EG)} \qquad u_{\rm EG}^{\rm new} = P_\Omega[u - \beta F(\tilde{u})] \tag{6.14}$$

and it can be viewed as the primary method

$$u^{k+1} = P_\Omega[u^k - d_2(u^k, v^k, \tilde{u}^k)]. \tag{6.15}$$

The convergence of these primary methods is a consequence of Theorem 2.1.

Theorem 6.2 *Let the conditions in Theorem 6.1 be satisfied. Then the sequence $\{u^k\}$ generated by the primary method* (6.13) *or* (6.15) *converges to some u^∞ which is a solution point of $VI(\Omega, F)$.*

Proof. Due to Theorem 6.1 the quadruplet $\big(d_1(u,v,\tilde{u}), d_2(u,v,\tilde{u}), \varphi(u,v,\tilde{u}), \phi(u,v,\tilde{u})\big)$ in this section satisfies condition (2.1). We only need to verify the additional conditions (2.13). It follows from Propositions 2.1 and 2.2 that

$$\|u^{k+1} - u^*\|^2 \le \|u^k - u^*\|^2 - (1-\nu^2)\|u^k - \tilde{u}^k\|^2$$

and thus

$$\lim_{k\to\infty} \|u^k - \tilde{u}^k\| = 0. \tag{6.16}$$

From the accepting rule (6.7) and (6.16) follows

$$\begin{aligned} \lim_{k\to\infty} \|d_1(u^k, v^k, \tilde{u}^k)\| &\le \lim_{k\to\infty} \{\|u^k - \tilde{u}^k\| + \beta_k\|F(v^k) - F(\tilde{u}^k)\|\} \\ &\le \lim_{k\to\infty} (1+\nu)\|u^k - \tilde{u}^k\| = 0. \end{aligned}$$

The last additional condition holds because $d_2(u^k, v^k, \tilde{u}^k) = \beta_k F(\tilde{u}^k)$ (see (6.3)). All the conditions in Theorem 2.1 are satisfied and thus $\{u^k\}$ converges to a solution point of VI(Ω, F). $\square$

For the same reason of achieving faster convergence as mentioned in the previous two sections, in practice, we increase the parameter β when

$$\beta_k\|F(u^k) - F(\tilde{u}^k)\| \ll \|u^k - \tilde{u}^k\|.$$

The following procedure is recommended to find a suitable parameter β_k.

Procedure 6.1 Finding β_k to satisfy (6.10b). $\beta_0 = 1$, $\nu = 0.9$.

REPEAT:	$\tilde{u}^k = P_\Omega[u^k - \beta_k F(u^k)]$.
	If $r_k := \dfrac{\beta_k\|F(u^k) - F(\tilde{u}^k)\|}{\|u^k - \tilde{u}^k\|} \ge \nu$, $\quad \beta_k := \beta_k * 0.7 * \min\{1, \frac{1}{r_k}\}$.
UNTIL:	$\beta_k\|F(u^k) - F(\tilde{u}^k)\| \le \nu\|u^k - \tilde{u}^k\|$. (Accepting rule (6.10b))
ADJUST:	$\beta_{k+1} := \begin{cases} \beta_k * \nu * 0.9/r_k & \text{if } r_k \le 0.3, \\ \beta_k & \text{otherwise.} \end{cases}$

7 The general contraction methods

In the case $G = I$, the primary methods (2.9b) and (2.9c) take the similar iterations. Besides the primary methods, this section considers the construction of the *general contraction methods*. In the general methods, instead of $d_1(u,v,\tilde{u})$ and/or $d_2(u,v,\tilde{u})$, we use their convex combination

$$d(u,v,\tilde{u},t) = (1-t)d_1(u,v,\tilde{u}) + t d_2(u,v,\tilde{u}) \quad t \in [0,1] \tag{7.1}$$

as the search direction. In the case $G = I$, it follows from Lemmas 2.1 and 2.2 that $-\boldsymbol{d}(u, v, \tilde{u}, t)$ is a descent direction of the unknown distance function $\|u - u^*\|^2$ for any $u \in \Omega \setminus \Omega^*$. Let the new iterate be given by[3]

$$u(\alpha, t) = P_\Omega[u - \alpha \boldsymbol{d}(u, v, \tilde{u}, t)]. \tag{7.2}$$

We discuss how to select a reasonable step length α and analyze which direction is better. For these purposes, we define

$$\theta(\alpha, t) = \|u - u^*\|^2 - \|u(\alpha, t) - u^*\|^2 \tag{7.3}$$

as the progress function in the k-th iteration. In order to achieve more progress in each iteration, the ideal thought is to maximize $\theta(\alpha, t)$. Unfortunately, because $\theta(\alpha, t)$ involves the unknown vector u^*, we cannot maximize it directly. The following theorem provides a lower bound for $\theta(\alpha, t)$, namely, $\vartheta(\alpha, t)$, which does not include the unknown solution u^*.

Theorem 7.1 *For any $u^* \in \Omega^*$, $t \in [0, 1]$ and $\alpha \geq 0$, we have*

$$\theta(\alpha, t) \geq \vartheta(\alpha, t), \tag{7.4}$$

where

$$\vartheta(\alpha, t) = q(\alpha) + \varpi(\alpha, t), \tag{7.5}$$

$$q(\alpha) = 2\alpha\varphi(u, v, \tilde{u}) - \alpha^2\|d_1(u, v, \tilde{u})\|^2 \tag{7.6}$$

and

$$\varpi(\alpha, t) = \|u(\alpha, t) - [u - \alpha d_1(u, v, \tilde{u})]\|^2. \tag{7.7}$$

Proof. First, since $u(\alpha, t) = P_\Omega[u - \alpha \boldsymbol{d}(u, v, \tilde{u}, t)]$ and $u^* \in \Omega$, it follows from (1.11) that

$$\|u(\alpha, t) - u^*\|^2 \leq \|u - \alpha \boldsymbol{d}(u, v, \tilde{u}, t) - u^*\|^2 - \|u - \alpha \boldsymbol{d}(u, v, \tilde{u}, t) - u(\alpha, t)\|^2. \tag{7.8}$$

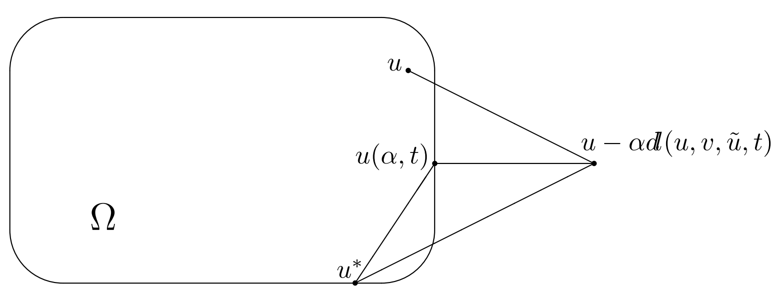

Fig. 1. Geometric interpretation of inequality (7.8) with $G = I$.

Consequently, using the definition of $\theta(\alpha, t)$, we get

$$\begin{aligned} \theta(\alpha, t) &\geq \|u - u^*\|^2 - \|u - \alpha \boldsymbol{d}(u, v, \tilde{u}, t) - u^*\|^2 + \|u - \alpha \boldsymbol{d}(u, v, \tilde{u}, t) - u(\alpha, t)\|^2 \\ &= \|u - u(\alpha, t)\|^2 + 2\alpha(u - u^*)^T \boldsymbol{d}(u, v, \tilde{u}, t) + 2\alpha(u(\alpha, t) - u)^T \boldsymbol{d}(u, v, \tilde{u}, t) \\ &= 2\alpha(u(\alpha, t) - u^*)^T \boldsymbol{d}(u, v, \tilde{u}, t) + \|u - u(\alpha, t)\|^2. \end{aligned} \tag{7.9}$$

Since $u(\alpha, t) \in \Omega$, and $\boldsymbol{d}(u, v, \tilde{u}, t)$ is a convex combination of $d_1(u, v, \tilde{u})$ and $d_2(u, v, \tilde{u})$, it follows from (2.6) and (2.8) that

$$(u(\alpha, t) - u^*)^T \boldsymbol{d}(u, v, \tilde{u}, t) \geq \varphi(u, v, \tilde{u}) + (u(\alpha, t) - u)^T d_1(u, v, \tilde{u}). \tag{7.10}$$

[3]Since $-S^{-1}\boldsymbol{d}(u, v, \tilde{u}, t)$ is a descent direction of $\|u - u^*\|_S^2$ for any given symmetric positive definite matrix S, the analysis for the general case $u(\alpha, t) = P_\Omega[u - \alpha S^{-1}\boldsymbol{d}(u, v, \tilde{u}, t)]$ is similar.

Substituting (7.10) into the right-hand-side of (7.9), we obtain

$$\begin{aligned}\theta(\alpha,t) &\geq 2\alpha\varphi(u,v,\tilde{u})+2\alpha(u(\alpha,t)-u)^T d_1(u,v,\tilde{u})+\|u-u(\alpha,t)\|^2\\ &= 2\alpha\varphi(u,v,\tilde{u})-\alpha^2\|d_1(u,v,\tilde{u})\|^2+\|u-u(\alpha,t)-\alpha d_1(u,v,\tilde{u})\|^2\\ &= q(\alpha)+\varpi(\alpha,t). \end{aligned}\tag{7.11}$$

The proof is complete. □

In general, $\vartheta(\alpha,t)$ is a tight lower bound of $\theta(\alpha,t)$ (for an example, see [23]). Note that $\vartheta(\alpha,t)$ involves two parts. The first part, $q(\alpha)$, offers us the rule to determine the step length and the second part, $\varpi(\alpha,t)$, tells us how to choose the search directions.

7.1 Selecting reasonable step lengths and the convergence

When $\Omega = R^n$ and $d\!\!l(u,v,\tilde{u},0) = d_1(u,v,\tilde{u})$, we have $\varpi(\alpha,0) = 0$ (see (7.2)) for any $\alpha \geq 0$. Therefore, in the process of determining the step length α, we ignore $\varpi(\alpha,t)$ and use the function $q(\alpha)$ only. Note that $q(\alpha)$ (independent of t) is a quadratic of α, it reaches its maximum at

$$\alpha^* = \frac{\varphi(u,v,\tilde{u})}{\|d_1(u,v,\tilde{u})\|^2}. \tag{7.12}$$

For $u^k \notin \Omega^*$, it follows from condition (2.1c) that

$$\alpha^* > \frac{1}{2}.$$

The step size α^* is only dependent on $\varphi(u,v,\tilde{u})$ and $d_1(u,v,\tilde{u})$, regardless of what convex combination factor t is selected. In other words, the methods use different search directions but the same step size.

Since some inequalities are used in the analysis, the 'optimal' step size (7.12) is usually conservative for contraction methods. It is encouraged to use a relaxation factor $\gamma \in [1,2)$ and the new iterate is updated by

$$u^{k+1} = P_\Omega[u^k - \gamma\alpha_k^* d\!\!l(u^k,v^k,\tilde{u}^k,t)]. \tag{7.13}$$

For updating form (7.13), using Theorem 7.1, by a simple manipulation we get

$$\begin{aligned}\|u^{k+1}-u^*\|^2 &\leq \|u^k-u^*\|^2 - q(\gamma\alpha_k^*)\\ &= \|u^k-u^*\|^2 - 2\gamma\alpha_k^*\varphi(u^k,v^k,\tilde{u}^k)+\gamma^2(\alpha_k^*)^2\|d_1(u^k,v^k,\tilde{u}^k)\|^2\\ &= \|u^k-u^*\|^2 - \gamma(2-\gamma)\alpha_k^*\varphi(u^k,v^k,\tilde{u}^k). \end{aligned}\tag{7.14}$$

Since $\alpha_k^* > \frac{1}{2}$, from (7.14) we obtain

$$\|u^{k+1}-u^*\|^2 \leq \|u^k-u^*\|^2 - \frac{\gamma(2-\gamma)}{2}\varphi(u^k,v^k,\tilde{u}^k). \tag{7.15}$$

As a conclusion, we get the following convergence results.

Theorem 7.2 *For given $u^k \in \Omega$ and $\beta_k \geq \beta_L > 0$, let $v^k \in \Omega$ be an approximate solution of* (1.3) *via certain accepting rule. Assume that the common conditions are satisfied and the sequence $\{u^k\}$ is generated by the general methods. In addition, if*

$$\lim_{k\to\infty}\{d_2(u^k,v^k,\tilde{u}^k) - \beta_k F(\tilde{u}^k)\} = 0, \tag{7.16}$$

then $\{u^k\}$ converges to some u^∞ which is a solution point of VI(Ω,F).

Proof. Based on the proof of Theorem 2.1, it is sufficient to verify

$$\lim_{k\to\infty} d_1(u^k, v^k, \tilde{u}^k) = 0. \tag{7.17}$$

From (7.14), it follows that

$$\lim_{k\to\infty} \varphi(u^k, v^k, \tilde{u}^k) = 0$$

and (7.17) is satisfied due to condition (2.1c). The proof is complete. □

Instead of choosing a fixed $\gamma \in [1, 2)$, we could adopt a dynamical relaxation factor $\gamma_k = 1/\alpha_k^*$ in (7.13),

$$u^{k+1} = P_\Omega[u^k - \mathbf{d}(u^k, v^k, \tilde{u}^k, t)]. \tag{7.18}$$

For this updating form, it follows from (7.14) and (7.12) that

$$\begin{aligned} \|u^{k+1} - u^*\|^2 &\le \|u^k - u^*\|^2 - \gamma(2-\gamma)\alpha_k^* \varphi(u^k, v^k, \tilde{u}^k) \\ &= \|u^k - u^*\|^2 - \big(2 - 1/\alpha_k^*\big)\varphi(u^k, v^k, \tilde{u}^k) \\ &= \|u^k - u^*\|^2 - \big(2\varphi(u^k, v^k, \tilde{u}^k) - \|d_1(u, v, \tilde{u})\|^2\big) \\ &\le \|u^k - u^*\|^2 - \phi(u^k, v^k, \tilde{u}^k). \end{aligned}$$

This is the same result as the ones in Propositions 2.1 and 2.2. Thus (7.18) can be viewed as a primary method.

7.2 Choosing better directions

This subsection answers which direction is better for getting more progress in the k-th iteration. Since the first part of $\vartheta(\alpha, t)$, namely $q(\alpha)$, is independent of t (see (7.6)), we need only to investigate the magnitude $\varpi(\alpha, t)$ for different $t \in [0, 1]$.

Theorem 7.3 *For any $\alpha \ge 0$ and $t \in [0, 1]$, $\varpi(\alpha, t)$ is a nondecreasing function of t. Especially, we have*

$$\varpi(\alpha, t) - \varpi(\alpha, 0) \ge \|u(\alpha, t) - u(\alpha, 0)\|^2, \quad \forall\ \alpha \ge 0. \tag{7.19}$$

Proof. Note that $\mathbf{d}(u, v, \tilde{u}, t)$ in (7.1) can be rewritten as

$$\mathbf{d}(u, v, \tilde{u}, t) = d_1(u, v, \tilde{u}) + t(d_2(u, v, \tilde{u}) - d_1(u, v, \tilde{u})), \tag{7.20}$$

and thus $u(\alpha, t)$ in (7.2) is

$$u(\alpha, t) = P_\Omega\{[u - \alpha d_1(u, v, \tilde{u})] - t\alpha[d_2(u, v, \tilde{u}) - d_1(u, v, \tilde{u})]\}.$$

By using the notation

$$\bar{u}(\alpha) = u - \alpha d_1(u, v, \tilde{u}), \tag{7.21}$$

we have (see (7.2))

$$u(\alpha, t) = P_\Omega\{\bar{u}(\alpha) - t\alpha[d_2(u, v, \tilde{u}) - d_1(u, v, \tilde{u})]\},$$

and (see (7.7))

$$\varpi(\alpha, t) = \|P_\Omega\{\bar{u}(\alpha) - t\alpha[d_2(u, v, \tilde{u}) - d_1(u, v, \tilde{u})]\} - \bar{u}(\alpha)\|^2.$$

It follows from Lemma 1.2 that $\varpi(\alpha, t)$ is a nondecreasing function of t for $t \ge 0$. According to (7.7) and (7.21)

$$\varpi(\alpha, t) - \varpi(\alpha, 0) = \|u(\alpha, t) - \bar{u}(\alpha)\|^2 - \|u(\alpha, 0) - \bar{u}(\alpha)\|^2. \tag{7.22}$$

Note that (see the notation (7.2), (7.20) and (7.21))

$$u(\alpha, 0) = P_\Omega[\bar{u}(\alpha)]. \tag{7.23}$$

Since $u(\alpha, t) \in \Omega$, by using (1.11) we obtain

$$\|u(\alpha, t) - u(\alpha, 0)\|^2 \le \|\bar{u}(\alpha) - u(\alpha, t)\|^2 - \|\bar{u}(\alpha) - u(\alpha, 0)\|^2. \tag{7.24}$$

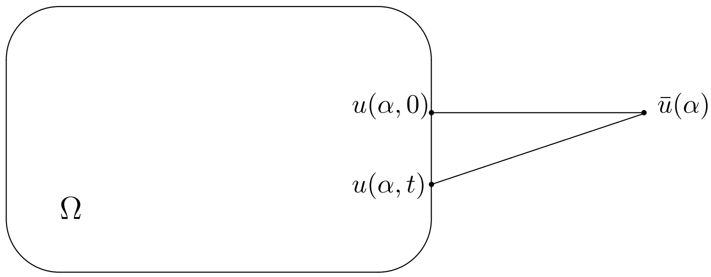

Fig. 2. Geometric interpretation of inequality (7.24).

The assertion of this theorem follows directly from (7.22) and (7.24). □

Theorem 7.3 indicates that in each iterative step, we may expect the contraction methods using direction $d_2(u, v, \tilde{u})$ will get more progress than the one using $d_1(u, v, \tilde{u})$. Under the common conditions, the new iterate should be updated by

$$u^{k+1} = P_\Omega[u^k - \gamma\alpha_k^* d_2(u^k, v^k, \tilde{u}^k)], \tag{7.25}$$

where $\gamma \in [1, 2)$ and α_k^* is defined in (7.12). Note that other relaxation factors than $\gamma \in [1.8, 1.95]$ could be used but experiments [20, 23] have shown that these values generate good results.

8 Numerical experiments of part I

This section tests the methods based on the effective quadruplets in Sections 3–6, namely, the primary methods and the related general methods described in Section 7. The quadruplets and the accepting rules will be satisfied by choosing a suitable parameter β_k. Thus the main computational load of the methods is the evaluation of the mapping F. All of the recursions are in the form

$$u^{k+1} = P_\Omega[u^k - \alpha_k d(u^k, v^k, \tilde{u}^k)].$$

We differentiate the test methods by using the following abbreviations:

- P and G denote the primary methods and general methods, respectively. In the primary methods, $\alpha_k \equiv 1$. In the general methods, $\alpha_k = \gamma\alpha_k^*$, where $\alpha_k^* = \varphi(u, v, \tilde{u})/\|d_1(u, v, \tilde{u})\|^2$ (see (7.12)) and $\gamma = 1.8$.

- The directions $d_1(u, v, \tilde{u})$ and $d_2(u, v, \tilde{u})$ are abbreviated as D1 and D2, respectively;

- SL, L, and NL are abbreviations for Symmetric Linear VI, Linear VI, and Nonlinear VI, respectively.

The abbreviations for different methods are listed in the following table.

The marks of the methods with different search directions and step sizes

Step size α	Quadruplet in Sec. 3		Quadruplet in Sec. 4		Quadruplet in Sec. 6	
	$d_1(u,v,\tilde{u})$	$d_2(u,v,\tilde{u})$	$d_1(u,v,\tilde{u})$	$d_2(u,v,\tilde{u})$	$d_1(u,v,\tilde{u})$	$d_2(u,v,\tilde{u})$
$\alpha_k \equiv 1$	SLD1-P	SLD2-P	LD1-P	LD2-P	NLD1-P	NLD2-P
$\alpha_k = \gamma\alpha_k^*$	SLD1-G	SLD2-G	LD1-G	LD2-G	NLD1-G	NLD2-G

Because

$$\{\text{Symmetric Nonlinear VIs}\} \subset \{\text{Nonlinear VIs}\}$$

and

$$\{\text{Symmetric Linear VIs}\} \subset \{\text{Linear VIs}\} \subset \{\text{Nonlinear VIs}\},$$

we test the Nonlinear VIs followed by the tests of the Symmetric Nonlinear VIs by using the symmetry and the Linear VIs (without symmetry) by using the linearity, respectively. Finally, the Symmetric Linear VIs are tested by using the linearity and symmetry.

We use *No. it* and *No. F* to denote the numbers of iterations and the evaluations of the mapping F, respectively. The size of the tested problems is from 100 to 1000. All codes are written in Matlab and run on a notebook computer. The iterations begin with $u^0 = 0$, $\beta = 1$ and stop as soon as

$$\frac{\|u^k - P_\Omega[u^k - F(u^k)]\|_\infty}{\|u^0 - P_\Omega[u^0 - F(u^0)]\|_\infty} \le 10^{-6}.$$

8.1 The numerical results for nonlinear variational inequalities

Test examples of nonlinear VIs. The mapping $F(u)$ in the tested nonlinear VIs is given by

$$F(u) = D(u) + Mu + q, \tag{8.1}$$

where $D(u) : R^n \to R^n$ is the nonlinear part, M is an $n \times n$ matrix, and $q \in R^n$ is a vector.

- In $D(u)$, the nonlinear part of $F(u)$, the components are

  $$D_j(u) = d_j * \arctan(a_j * u_j),$$

  where a and d are random vectors[4] whose elements are in $(0, 1)$.

- The matrix M in the linear part is given by $M = A^TA + B$. A is an $n \times n$ matrix whose entries are randomly generated in the interval $(-5, +5)$, and B is an $n \times n$ skew-symmetric random matrix whose entries[5] are in the interval $(-5, +5)$.

It is clear that the mapping composed in this way is monotone. We construct the following 6 sets of test examples by different combinations of Ω and q.

1. In the the first set of test examples, $\Omega = R^n_+$ is the non-negative orthant. The elements of vector q is generated from a uniform distribution in the interval $(-1000, 1000)$.

2. The second set of test examples is modified from the first set, the only difference is $\Omega = \{u \in R^n \,|\, 0 \le u \le b\}$. Each elements of b equals $u^*_{\max}$ by a positive factor less than 1, where $u^*_{\max}$ is the maximal element of the solution of the related problem in the first class.

[4] A similar type of (small) problems was tested in [32] where the components of the nonlinear mapping $D(u)$ are $D_j(u) = constant * \arctan(u_j)$.

[5] In the paper by Harker and Pang [11], the matrix $M = A^TA + B + D$, where A and B are the same matrices as what we use here, and D is a diagonal matrix with uniformly distributed random variable $d_{jj} \in (0.0, 0.3)$.

3. The 3-rd set[6] of test examples is similar to the first set. Instead of $q \in (-1000, 1000)$, the vector q is generated from a uniform distribution in the interval $(-1000, 0)$.

4. The 4-th set of test examples is modified from the third set, the only difference is $\Omega = \{u \in R^n \mid 0 \le u \le b\}$. Each elements of b equals $u^*_{\max}$ by a positive factor less than 1, where $u^*_{\max}$ is the maximal element of the solution of the related problem in the 3-rd class.

5. The 5-th set of test examples has a known solution u^* and $\Omega = R^n_+$. Let vector p be generated from a uniform distribution in the interval $(-10, 10)$ and

$$u^* = \max(p, 0). \tag{8.2}$$

For given $\eta > 0$, by setting

$$w = \eta * \max(-p, 0) \quad \text{and} \quad q = w - F(u^*),$$

we form a test problem

$$0 \le u \perp F(u) \ge 0$$

with a known solution u^* described in (8.2).

6. The 6-th set of test examples: The test problem has a known solution u^* and $\Omega = \{u \in R^n \mid 0 \le u \le b\}$ is a box. Let vector p be generated from a uniform distribution in the interval $(-5, 15)$, $b = 10$ and

$$u^* = p + \max(-p, 0) - \max(b - p, 0), \tag{8.3}$$

For given $\eta_1, \eta_2 > 0$, by setting

$$w = \max(-p, 0) * \eta_1 - \max(b - p, 0) * \eta_2 \qquad \text{and} \qquad q = w - F(u^*),$$

we form a test problem

$$u = P_\Omega[u - F(u)]$$

with a known solution u^* described in (8.3).

The tested methods and the numerical results. For nonlinear variational inequalities, we test the problems by using the accepting rule (6.10) which is fulfilled by Procedure 6.1. The quadruplet $\big(d_1(u, v, \tilde{u}),\, d_2(u, v, \tilde{u}),\, \varphi(u, v, \tilde{u}),\, \phi(u, v, \tilde{u})\big)$ is described in (6.9) with $v = u$. It is worth observing the effectiveness of different search directions and the different step-size rules. Thus, we compare the convergence behaviors of the following 4 methods:

NLD1-P, NLD2-P, NLD1-G and NLD2-G.

Each iteration of the test methods needs at least 2 times of evaluations of the mapping F. The test results for the 6 sets of nonlinear variational inequalities are given in Tables 8.1.1–8.1.6. Because u^* in the 5-th and 6-th sets of test examples is known, the difference $\|u^k - u^*\|$ is reported when the stopping criterium is satisfied.

Table 8.1.1. Numerical results for Nonlinear VIs of the 1-st set examples

Problem	Method NLD1-P		Method NLD2-P		Method NLD1-G		Method NLD2-G		max element
size n	No. It	No. F	No. It	No. F	No. It	No. F	No. It	No. F	of u^*
100	466	1011	391	815	236	490	200	434	5.1207
200	640	1361	568	1184	316	668	282	611	3.3988
500	697	1471	596	1244	343	728	310	672	1.4260
800	565	1203	483	999	278	589	259	546	0.8282
1000	601	1265	520	1086	297	631	267	580	0.5933

[6] In [11], the similar problems in the first set are called easy problems while the 3-rd problems are called hard problems.

Table 8.1.2. Numerical results for Nonlinear VIs of the 2-nd set examples

Problem size n	Method NLD1-P		Method NLD2-P		Method NLD1-G		Method NLD2-G		The vector b in $u \in [0, b]$
	No. It	No. F	No. It	No. F	No. It	No. F	No. It	No. F	
100	576	1243	473	986	285	591	243	527	4.0
200	676	1441	594	1237	329	695	296	641	3.0
500	730	1547	627	1302	361	764	325	704	1.0
800	713	1527	624	1296	358	755	350	727	0.6
1000	787	1663	677	1411	384	814	349	757	0.5

Table 8.1.3. Numerical results for Nonlinear VIs of the 3-rd set examples

Problem size n	Method NLD1-P		Method NLD2-P		Method NLD1-G		Method NLD2-G		max element of u^*
	No. It	No. F	No. It	No. F	No. It	No. F	No. It	No. F	
100	952	2021	884	1841	478	1018	448	969	14.437
200	1189	2246	1105	2296	594	1270	561	1214	9.0339
500	1453	3000	1402	2922	733	1571	711	1538	3.7623
800	1434	2952	1344	2802	730	1560	683	1478	2.5715
1000	1532	3159	1424	2968	772	1652	720	1557	2.4738

Table 8.1.4. Numerical results for Nonlinear VIs of the 4-th set examples

Problem size n	Method NLD1-P		Method NLD2-P		Method NLD1-G		Method NLD2-G		The vector b in $u \in [0, b]$
	No. It	No. F	No. It	No. F	No. It	No. F	No. It	No. F	
100	932	1961	837	1741	453	958	418	904	10.0
200	1346	2767	1246	2593	682	1457	635	1374	6.0
500	1723	3554	1626	3390	875	1873	830	1795	3.0
800	1759	3619	1639	3416	889	1898	830	1795	2.0
1000	1962	4046	1850	3855	994	2127	940	2032	2.0

Table 8.1.5. Numerical results for Nonlinear VIs of the 5-th set examples

Problem size n	Method NLD1-P		Method NLD2-P		Method NLD1-G		Method NLD2-G		The bounds of $\|u^k - u^*\|_\infty$
	No. It	No. F	No. It	No. F	No. It	No. F	No. It	No. F	
100	762	1550	670	1393	379	804	337	730	$2.0e-6$
200	964	1972	869	1812	478	1017	442	957	$8.0e-7$
500	969	1962	868	1809	487	1031	442	956	$3.0e-7$
800	1030	2106	928	1933	526	1120	470	1017	$1.6e-7$
1000	1026	2074	926	1928	518	1101	473	1023	$1.2e-7$

Table 8.1.6. Numerical results for Nonlinear VIs of the 6-th set examples

Problem size n	Method NLD1-P		Method NLD2-P		Method NLD1-G		Method NLD2-G		The bounds of $\|u^k - u^*\|_\infty$
	No. It	No. F	No. It	No. F	No. It	No. F	No. It	No. F	
100	1221	2507	1141	2374	618	1318	573	1238	$2.0e-7$
200	1157	2406	1052	2191	571	1216	533	1153	$7.0e-8$
500	1433	2908	1311	2729	723	1535	665	1437	$3.0e-8$
800	1312	2683	1212	2524	654	1391	613	1324	$1.8e-8$
1000	1133	2311	1034	2152	564	1198	521	1126	$1.2e-8$

The numerical results coincide with our theoretical results and analysis.

- In both of the primary methods and general methods, the methods with direction $d_2(u, v, \tilde{u})$ require fewer iterations than the corresponding methods with direction $d_1(u, v, \tilde{u})$. In particular,
$$\frac{\text{Computational load of NLD2-P}}{\text{Computational load of NLD1-P}}, \; \frac{\text{Computational load of NLD2-G}}{\text{Computational load of NLD1-G}} < 95\%.$$
- For the methods adopting the same direction, the general methods converge much faster than the primary methods.
$$\frac{\text{Computational load of NLD1-G}}{\text{Computational load of NLD1-P}}, \; \frac{\text{Computational load of NLD2-G}}{\text{Computational load of NLD2-P}} \approx 50\text{-}55\%.$$

Therefore, for nonlinear variational inequalities, we suggest to use method NLD2-G with $\gamma = 1.8$.

8.2 The numerical results for symmetric nonlinear variational inequalities

Test examples of Symmetric Nonlinear VIs. The test problems of symmetric nonlinear variational inequalities are formed by deleting the asymmetric part of the matrix M. in (8.1). In details,

$$F(u) = D(u) + Mu + q, \qquad M = A^T A.$$

The other data in the test problems are same as those described in Subsection 8.1.

$$F(u) = \nabla f(u). \tag{8.4}$$

The test problems are formed by $H = A^T A$. The other data in the test problems are same as those described in Subsection 8.3. In these test examples, the Jacobian matrix of $F(u)$ is symmetric and $F(u)$ can be viewed as the gradient of certain convex function.

The tested methods and the numerical results. We use the accepting rule (5.8) which is fulfilled by Procedure 5.1. The quadruplet $\big(d_1(u, v, \tilde{u}),\, d_2(u, v, \tilde{u}),\, \varphi(u, v, \tilde{u}),\, \phi(u, v, \tilde{u})\big)$ is described in (5.10) with $v = u$. Note that from (5.3), (5.4) and (5.5), it follows that SNLD1-P=SNLD2-P, thus we call the method as SNLD-P. Because

$$\{\text{Symmetric noninear VIs}\} \subset \{\text{Nonlinear VIs}\},$$

and the general contraction methods outperform the primary methods, we test the symmetric nonlinear problems with the method for nonlinear problem

NLD1-G and NLD2-G.

In addition, by using the symmetry, we use the method

SLD-P.

It is worth comparing the effectiveness of the following 3 methods:

NLD1-G, NLD2-G and SNLD-P.

Without the trial computation for finding the accepted parameter β_k, each iteration of SNLD-P needs only one evaluation of the mapping F. The test results for the 6 sets of symmetric linear variational inequalities are given in Tables 8.2.1–8.2.6. Also, in the 5-th and 6-th sets of test examples, because u^* is known, we also report the difference $\|u^k - u^*\|$ when the stopping criterion is satisfied.

Table 8.2.1. Numerical results for Symmetric NL-VIs of the 1-st set examples

Problem	Method NLD1-G		Method NLD2-G		Method SNLD-P		max element
size n	No. It	No. F	No. It	No. F	No. It	No. F	of u^*
100	235	485	205	445	113	146	5.1150
200	292	618	252	546	125	181	3.2853
500	345	732	308	667	142	202	1.4302
800	278	589	526	538	110	158	0.8389
1000	293	623	267	580	122	172	0.6027

Table 8.2.2. Numerical results for Symmetric NL-VIs of the 2-nd set examples

Problem	Method NLD1-G		Method NLD2-G		Method SNLD-P		The vector b
size n	No. It	No. F	No. It	No. F	No. It	No. F	in $u \in [0, b]$
100	299	630	265	575	156	187	4.0
200	307	649	274	593	140	194	3.0
500	351	713	319	691	181	246	1.0
800	356	751	340	710	167	222	0.6
1000	385	817	349	757	168	238	0.5

Table 8.2.3. Numerical results for Symmetric NL-VIs of the 3-rd set examples

Problem size n	Method NLD1-G		Method NLD2-G		Method SNLD-P		max element of u^*
	No. It	No. F	No. It	No. F	No. It	No. F	
100	617	1307	565	1222	249	336	17.2269
200	648	1386	609	1318	247	343	8.8365
500	738	1580	692	1497	318	427	3.8270
800	703	1502	657	1422	284	382	2.6666
1000	769	1666	725	1568	275	375	2.5822

Table 8.2.4. Numerical results for Symmetric NL-VIs of the 4-th set examples

Problem size n	Method NLD1-G		Method NLD2-G		Method SNLD-P		The vector b in $u \in [0, b]$
	No. It	No. F	No. It	No. F	No. It	No. F	
100	514	1079	462	998	210	299	12
200	779	1667	735	1590	294	406	6
500	929	1990	884	1912	362	486	3
800	869	1855	803	1737	296	408	2
1000	956	2046	910	1967	347	474	2

Table 8.2.5. Numerical results for Symmetric NL-VIs of the 5-th set examples

Problem size n	Method NLD1-G		Method NLD2-G		Method SNLD-P		The bounds of $\|u^k - u^*\|_\infty$
	No. It	No. F	No. It	No. F	No. It	No. F	
100	419	890	374	810	190	275	$1.4e-6$
200	507	1078	464	1004	208	295	$7.8e-7$
500	491	1039	446	965	213	302	$2.3e-7$
800	524	1116	473	1023	233	330	$1.0e-6$
1000	522	1109	474	1025	220	311	$1.3e-7$

Table 8.2.6. Numerical results for Symmetric NL-VIs of the 6-th set examples

Problem size n	Method NLD1-G		Method NLD2-G		Method SNLD-P		The bounds of $\|u^k - u^*\|_\infty$
	No. It	No. F	No. It	No. F	No. It	No. F	
100	794	1695	723	1562	349	480	$1.4e-7$
200	619	1318	576	1067	286	394	$5.6e-8$
500	734	1560	684	1478	361	496	$2.7e-8$
800	654	1392	611	1320	291	391	$1.6e-8$
1000	570	1210	528	1141	243	333	$1.1e-8$

Similarly as in Subsection 8.1, the numerical results coincide with our theoretical results and analysis.

- Again, the general method NLD2-G requires fewer iterations than NLD1-G.

$$\frac{\text{Computational load of NLD2-G}}{\text{Computational load of NLD1-G}} < 0.95.$$

- For symmetric nonlinear VIs, the method SNLD-P converges much faster than the method NLD2-G.

$$\frac{\text{Computational load of SNLD-P}}{\text{Computational load of NLD2-G}} \in [0.25, 0.35],$$

  it seems that we should use symmetry when the mapping F is the gradient of certain convex function.

8.3 The numerical results for asymmetric linear variational inequalities

Test examples of Linear VIs. In the linear variational inequalities (4.1), the mapping

$$F(u) = Mu + q. \tag{8.5}$$

The test problems are formed by deleting the nonlinear part $D(u)$ in (8.1). The other data in the test problems are same as those described in Subsection 8.1.

The tested methods and the numerical results. We use the accepting rule (4.17) which is fulfilled by Procedure 4.1. The quadruplet $\big(d_1(u,v,\tilde{u}),\, d_2(u,v,\tilde{u}),\, \varphi(u,v,\tilde{u}),\, \phi(u,v,\tilde{u})\big)$ is described in (4.13) with $v = u$. Because

$$\{\text{Linear VIs}\} \subset \{\text{Nonlinear VIs}\},$$

and the general contraction methods outperform the primary methods, we test the linear problems with nonlinear methods

$$\text{NLD1-G} \qquad \text{and} \qquad \text{NLD2-G}.$$

In addition, by using the linearity, we use the methods

$$\text{LD1-G} \qquad \text{and} \qquad \text{LD2-G}.$$

It is worth comparing the effectiveness of the methods:

$$\text{NLD1-G},\;\; \text{NLD2-G},\quad \text{LD1-G} \quad \text{and} \quad \text{LD2-G}.$$

Each iteration of the test methods needs at least 2 times of evaluations of the mapping F. The test results for the 6 sets of (asymmetric) linear variational inequalities are given in Tables 8.3.1–8.3.6. In the 5-th and 6-th sets of test examples, because u^* is known, the difference $\|u^k - u^*\|$ is reported when the stopping criterion is satisfied.

Table 8.3.1. Numerical results for Asymmetric LVIs of the 1-st set examples

Problem	Method NLD1-G		Method NLD2-G		Method LD1-G		Method LD2-G		max element
size n	No. It	No. F	No. It	No. F	No. It	No. F	No. It	No. F	of u^*
100	236	490	200	434	348	783	178	363	5.1242
200	315	666	282	611	520	1132	261	529	3.4029
500	342	726	310	672	596	1269	279	563	1.4266
800	278	589	259	546	417	964	228	463	0.8284
1000	296	629	267	580	470	1051	240	487	0.5933

Table 8.3.1. Numerical results for Asymmetric LVIs of the 2-nd set examples

Problem	Method NLD1-G		Method NLD2-G		Method LD1-G		Method LD2-G		The vector b
size n	No. It	No. F	No. It	No. F	No. It	No. F	No. It	No. F	in $u \in [0, b]$
100	286	593	244	529	429	959	216	439	4.0
200	329	695	297	643	538	1182	273	553	3.0
500	359	760	325	704	618	1323	291	587	1.0
800	358	755	351	729	505	1235	287	580	0.6
1000	384	814	349	757	648	1406	312	631	0.5

Table 8.3.3. Numerical results for Asymmetric LVIs of the 3-rd set examples

Problem	Method NLD1-G		Method NLD2-G		Method LD1-G		Method LD2-G		max element
size n	No. It	No. F	No. It	No. F	No. It	No. F	No. It	No. F	of u^*
100	479	1020	448	969	822	1792	432	874	14.449
200	595	1272	561	1214	927	2142	521	1047	9.0412
500	740	1586	711	1538	1293	2832	636	1301	3.7641
800	730	1560	684	1480	1256	2737	641	1296	2.5727
1000	771	1650	720	1557	1405	2963	665	1356	2.4747

Table 8.3.4. Numerical results for Asymmetric LVIs of the 4-th set examples

Problem	Method NLD1-G		Method NLD2-G		Method LD1-G		Method LD2-G		The vector b
size n	No. It	No. F	No. It	No. F	No. It	No. F	No. It	No. F	in $u \in [0, b]$
100	453	958	418	904	775	1680	415	840	10.0
200	684	1461	640	1385	1207	2593	589	1183	6.0
500	875	1873	830	1795	1553	3349	768	1555	3.0
800	894	1909	835	1806	1590	3397	781	1576	2.0
1000	995	2129	940	2032	1819	3852	875	1767	2.0

Table 8.3.5. Numerical results for Asymmetric LVIs of the 5-th set examples

Problem size n	Method NLD1-G		Method NLD2-G		Method LD1-G		Method LD2-G		The bounds of $\|u^k - u^*\|_\infty$
	No. It	No. F	No. It	No. F	No. It	No. F	No. It	No. F	
100	379	804	334	723	544	1268	320	646	$2.0e-6$
200	478	1017	442	957	796	1754	416	840	$8.0e-7$
500	487	1031	442	956	710	1667	417	840	$3.0e-7$
800	525	1118	470	1017	841	1859	442	905	$1.6e-7$
1000	518	1101	473	1023	880	1904	412	861	$1.2e-7$

Table 8.3.6. Numerical results for Asymmetric LVIs of the 6-th set examples

Problem size n	Method NLD1-G		Method NLD2-G		Method LD1-G		Method LD2-G		The bounds of $\|u^k - u^*\|_\infty$
	No. It	No. F	No. It	No. F	No. It	No. F	No. It	No. F	
100	617	1316	573	1238	1062	2301	553	1120	$2.0e-7$
200	572	1218	533	1153	983	2137	505	1022	$7.0e-8$
500	724	1537	666	1439	1226	2666	613	1251	$3.0e-8$
800	657	1398	616	1331	1029	2361	570	1161	$1.8e-8$
1000	565	1200	521	1126	947	2074	493	997	$1.2e-8$

Similarly as in Subsection 8.4, the numerical results coincide with our theoretical results and analysis.

- Using the general methods for nonlinear VIs in Section 6 to solve the linear VIs, the method with direction $d_2(u, v, \tilde{u})$ requires fewer iterations than the corresponding methods with direction $d_1(u, v, \tilde{u})$,
$$\frac{\text{Computational load of NLD2-G}}{\text{Computational load of NLD1-G}} \approx 90\%.$$
- For the general methods with either direction $d_1(u, v, \tilde{u})$ or direction $d_2(u, v, \tilde{u})$ for linear VIs in Section 4, the method with direction $d_2(u, v, \tilde{u})$ converges much faster than the corresponding method with direction $d_1(u, v, \tilde{u})$,
$$\frac{\text{Computational load of LD2-G}}{\text{Computational load of LD1-G}} \approx 50\%.$$
- For linear VIs, the general method with direction $d_2(u, v, \tilde{u})$ in Section 4 requires fewer iterations than the corresponding method in Section 6,
$$\frac{\text{Computational load of LD2-G}}{\text{Computational load of NLD2-G}} \approx 90\%.$$

The method LD2-G converges faster than all other tested methods, which implies that we should use the linearity when the variational inequality is linear.

8.4 The numerical results for symmetric linear variational inequalities

Test examples of Linear VIs. In the symmetric linear variational inequalities (3.1), the mapping

$$F(u) = Hu + q. \tag{8.6}$$

The test problems are formed by $H = A^T A$. The other data in the test problems are same as those described in Subsection 8.3.

The tested methods and the numerical results. We use the accepting rule (3.12) which is fulfilled by Procedure 3.1. The quadruplet $\big(d_1(u, v, \tilde{u}),\, d_2(u, v, \tilde{u}),\, \varphi(u, v, \tilde{u}),\, \phi(u, v, \tilde{u})\big)$ is described in (3.10) with $v = u$. Note that from (3.3), (3.4) and (3.5), it follows that SLD1-P=SLD2-P, thus we call the method as SLD-P. Because

$$\{\text{Symmetric Linear VIs}\} \subset \{\text{Linear VIs}\} \subset \{\text{Nonlinear VIs}\},$$

and the general contraction methods outperform the primary methods, we test the symmetric linear problems with the method for nonlinear problem

$$\text{NLD2-G} \qquad \text{and} \qquad \text{LD2-G}.$$

In addition, by using the symmetry, we use the method

$$\text{SLD-P}.$$

It is worth comparing the effectiveness of the following 3 methods:

$$\text{NLD2-G}, \ \text{LD2-G} \quad \text{and} \quad \text{SLD-P}.$$

Without the trial computation for finding the accepted parameter β_k, each iteration of SLD-P needs only one evaluation of the mapping F (here is $Hu + q$). The test results for the 6 sets of symmetric linear variational inequalities are given in Tables 8.4.1–8.4.6. Also, in the 5-th and 6-th sets of test examples, because u^* is known, we also report the difference $\|u^k - u^*\|$ when the stopping criterion is satisfied.

Table 8.4.1. Numerical results for Symmetric LVIs of the 1-st set examples

Problem	Method NLD2-G		Method LD2-G		Method SLD-P		max element
size n	No. It	No. F	No. It	No. F	No. It	No. F	of u^*
100	205	445	185	375	74	103	5.1184
200	256	555	232	471	83	116	3.2891
500	308	667	279	563	97	123	1.4308
800	256	538	226	459	139	160	0.8391
1000	267	580	210	497	134	160	0.6028

Table 8.4.2. Numerical results for Symmetric LVIs of the 2-nd set examples

Problem	Method NLD2-G		Method LD2-G		Method SLD-P		The vector b
size n	No. It	No. F	No. It	No. F	No. It	No. F	in $u \in [0, b]$
100	377	817	345	704	115	157	4.0
200	468	1013	436	890	164	223	3.0
500	602	1303	553	1112	153	204	1.0
800	510	1104	475	957	137	192	0.6
1000	585	1266	547	1103	184	240	0.5

Table 8.4.3. Numerical results for Symmetric LVIs of the 3-rd set examples

Problem	Method NLD2-G		Method LD2-G		Method SLD-P		max element
size n	No. It	No. F	No. It	No. F	No. It	No. F	of u^*
100	566	1224	531	1068	163	220	17.2432
200	609	1318	577	1159	182	251	8.8438
500	692	1497	638	1298	150	209	3.8289
800	658	1424	613	1244	289	349	2.6678
1000	725	1568	674	1362	161	219	2.5832

Table 8.4.4. Numerical results for Symmetric LVIs of the 4-th set examples

Problem	Method NLD2-G		Method LD2-G		Method SLD-P		The vector b
size n	No. It	No. F	No. It	No. F	No. It	No. F	in $u \in [0, b]$
100	461	999	426	867	146	198	12
200	839	1814	793	1600	166	225	6
500	830	1795	786	1577	172	236	3
800	859	1858	806	1628	230	292	2
1000	1008	2179	956	1928	267	352	2

Table 8.4.5. Numerical results for Symmetric LVIs of the 5-th set examples

Problem size n	Method NLD2-G		Method LD2-G		Method SLD-P		The bounds of $\|u^k - u^*\|_\infty$
	No. It	No. F	No. It	No. F	No. It	No. F	
100	370	802	352	715	151	190	$2.0e-6$
200	464	1004	438	881	170	219	$8.0e-7$
500	446	965	417	857	174	230	$3.0e-7$
800	477	1032	451	909	179	225	$1.6e-7$
1000	474	1025	430	883	150	197	$1.2e-7$

Table 8.4.6. Numerical results for Symmetric LVIs of the 6-th set examples

Problem size n	Method NLD2-G		Method LD2-G		Method SLD-P		The bounds of $\|u^k - u^*\|_\infty$
	No. It	No. F	No. It	No. F	No. It	No. F	
100	725	1566	704	1415	208	274	$2.0e-6$
200	577	1248	513	1067	189	253	$8.0e-7$
500	684	1478	634	1287	204	268	$3.0e-7$
800	612	1322	555	1145	216	270	$1.6e-7$
1000	528	1141	492	1003	155	210	$1.2e-7$

Similarly as in Subsection 8.1, the numerical results coincide with our theoretical results and analysis.

- For symmetric linear VIs, the general method LD2-G requires fewer iterations than NLD2-G.

  $$\frac{\text{Computational load of LD2-G}}{\text{Computational load of NLD2-G}} < 95\%,$$

  this means that we should use the linearity if the variational inequality is linear.

- For symmetric linear VIs, the method SLD-P converges much faster than the method LD2-G.

  $$\frac{\text{Computational load of SLD-P}}{\text{Computational load of LD2-G}} \approx \text{20-25}\%,$$

  it seems that we should use symmetry when the linear VIs are symmetric.

9 Effective quadruplet for Solodov-Svaiter's APPA

Now, we begin the new part of this paper. For some existing APPAs and their accepting rules, we will find the effective quadruplets and point out that the existing methods are primary methods of the form (2.9a). Using the unified framework, we will give a simple extended contraction method. The significance of the extension will be reported in Section 11.

The APPA proposed by Solodov and Svaiter, see Algorithm 2 in [29], the triplet $(u, v, \tilde{u})$ in (1.7) is accepted when

$$(\textbf{Accepting rule}) \qquad \Delta(u, v, \tilde{u}) \le \nu \|u - v\|^2, \quad \nu \in (0, 1), \tag{9.1a}$$

is satisfied, where

$$\Delta(u, v, \tilde{u}) = 2(v - \tilde{u})^T \{(v - u) + \beta F(v)\} - \|v - \tilde{u}\|^2. \tag{9.1b}$$

For fixed $\beta > 0$ and any small $\tau > 0$, if $u \in \Omega$ is not a solution point of (1.3), one can find a good approximate solution v in (1.6) such that

$$\|v - \tilde{u}\| \le \tau \, \|u - \tilde{u}\|$$

and thus, in principal, the accepting rule (9.1) is implementable. Under the accepting rule (9.1), the point $\tilde{u}^k$ was accepted as the new iterate u^{k+1} by Solodov and Svaiter (see [29], pp. 385). For the triplet $(u, v, \tilde{u})$ in (1.7) satisfying accepting rule (9.1), we define the quadruplet $\big(d_1(u, v, \tilde{u}), d_2(u, v, \tilde{u}), \varphi(u, v, \tilde{u}), \phi(u, v, \tilde{u})\big)$ by setting

$$\textbf{(The quadruplet)} \qquad \begin{cases} d_1(u, v, \tilde{u}) &= u - \tilde{u}, \\ d_2(u, v, \tilde{u}) &= \beta F(v), \\ \varphi(u, v, \tilde{u}) &= \|u - \tilde{u}\|^2 - (v - \tilde{u})^T \beta F(v), \\ \phi(u, v, \tilde{u}) &= (1 - \nu)\|u - v\|^2. \end{cases} \tag{9.2}$$

In the following we show that this quadruplet is effective which satisfies conditions (2.1) with $G = I$.

Condition (2.1a): The basic equation (1.7) can be written as

$$\tilde{u} = P_\Omega\{\tilde{u} - [\beta F(v) - (u - \tilde{u})]\}. \tag{9.3}$$

Using (9.2), the duplet $(d_1(u, v, \tilde{u}), d_2(u, v, \tilde{u}))$ satisfies condition (2.1a).

Condition (2.1b): Since $v \in \Omega$ and $u^* \in \Omega^*$, we have $(v - u^*)^T F(u^*) \geq 0$. From the monotonicity of F we have

$$(v - u^*)^T \beta F(v) \geq 0.$$

Adding $(\tilde{u} - v)^T \beta F(v)$ to both sides of the above inequality, we get

$$(\tilde{u} - u^*)^T \beta F(v) \geq (\tilde{u} - v)^T \beta F(v). \tag{9.4}$$

Note that the left-hand-side of (9.4) is

$$(\tilde{u} - u^*)^T d_2(u, v, \tilde{u})$$

while the right-hand-side of (9.4) is

$$\varphi(u, v, \tilde{u}) - (u - \tilde{u})^T d_1(u, v, \tilde{u}).$$

Thus the triplet $\big(d_1(u, v, \tilde{u}), d_2(u, v, \tilde{u}), \varphi(u, v, \tilde{u})\big)$ satisfies condition (2.1b).

Condition (2.1c): It follows from (9.1) and the definition of $\phi(u, v, \tilde{u})$ in (9.2) that

$$2(v - \tilde{u})^T\{(v - u) + \beta F(v)\} - \|v - \tilde{u}\|^2 + \phi(u, v, \tilde{u}) \leq \|u - v\|^2$$

and consequently

$$2(v - \tilde{u})^T \beta F(v) + \phi(u, v, \tilde{u}) \leq \|u - \tilde{u}\|^2.$$

An equivalent form of the above inequality is

$$2\|u - \tilde{u}\|^2 - 2(v - \tilde{u})^T \beta F(v) \geq \|u - \tilde{u}\|^2 + \phi(u, v, \tilde{u}). \tag{9.5}$$

Note that the left-hand-side of (9.5) is $2\varphi(u, v, \tilde{u})$ (see (9.2)), while the right-hand-side of (9.5) is $\|d_1(u, v, \tilde{u})\|^2 + \phi(u, v, \tilde{u})$. Therefore, condition (2.1c) is satisfied.

Condition (2.1d): Since $\tilde{u} = P_\Omega[u - \beta F(v)]$ and $v \in \Omega$, it follows from (1.9) that

$$\{[u - \beta F(v)] - \tilde{u}\}^T (v - \tilde{u}) \leq 0$$

and thus

$$(v - \tilde{u})^T\{(v - u) + \beta F(v)\} \geq \|v - \tilde{u}\|^2.$$

From (9.1b) and the above inequality we obtain

$$\Delta(u, v, \tilde{u}) \geq \|v - \tilde{u}\|^2.$$

Together with the accepting rule (9.1), we get

$$\|v - \tilde{u}\| \leq \sqrt{\nu}\|u - v\|$$

and thus

$$\|u - \tilde{u}\| \leq \|u - v\| + \|v - \tilde{u}\| \leq (1 + \sqrt{\nu})\|u - v\|. \tag{9.6}$$

Finally, from (9.6) and the definition of $\phi(u, v, \tilde{u})$ in (9.2), we obtain

$$\|u - \tilde{u}\|^2 \leq \Big(\frac{1 + \sqrt{\nu}}{1 - \sqrt{\nu}}\Big)\phi(u, v, \tilde{u})$$

and condition (2.1d) holds with $\kappa = (1 - \sqrt{\nu})/(1 + \sqrt{\nu})$.

Based on the above analysis, we have proved the following theorem.

Theorem 9.1 *For solving problem* (1.1)*, let the triplet* $(u, v, \tilde{u})$ *be defined in* (1.7) *and the accepting rule* (9.1) *be satisfied. Then the quadruplet* $\big(d_1(u, v, \tilde{u}),\ d_2(u, v, \tilde{u}),\ \varphi(u, v, \tilde{u}),\ \phi(u, v, \tilde{u})\big)$ *defined in* (9.2) *is an effective quadruplet which fulfills conditions* (2.1)*.*

The convergence of this primary method is a consequence of Theorem 2.1.

Theorem 9.2 *Let the conditions in Theorem 9.1 be satisfied. Then the sequence* $\{u^k\}$ *generated by the primary methods converges to some* u^∞ *which is a solution point of VI*(Ω, F)*.*

Proof. Due to Theorem 9.1 the quadruplet $\big(d_1(u, v, \tilde{u}), d_2(u, v, \tilde{u}), \varphi(u, v, \tilde{u}), \phi(u, v, \tilde{u})\big)$ in (9.2) satisfies condition (2.1). We need only to verify the additional conditions (2.13). It follows from Propositions 2.1 and 2.2 that

$$\|u^{k+1} - u^*\|^2 \leq \|u^k - u^*\|^2 - (1 - \nu)\|u^k - v^k\|^2$$

and thus

$$\lim_{k\to\infty} \|u^k - v^k\| = 0. \tag{9.7}$$

From (9.6) we know that

$$\|u^k - \tilde{u}^k\| \leq \|u^k - v^k\| + \|v^k - \tilde{u}^k\| \leq (1 + \sqrt{\nu})\|u^k - v^k\|. \tag{9.8}$$

Because $d_1(u^k, v^k, \tilde{u}^k) = u^k - \tilde{u}^k$ (see the form in (9.2)), it follows from (9.7) and (9.8) that

$$\lim_{k\to+\infty} d_1(u^k, v^k, \tilde{u}^k) = 0.$$

From (9.6), (9.7) and the continuity of F, we get

$$\lim_{k\to+\infty} \{d_2(u^k, v^k, \tilde{u}^k) - \beta_k F(\tilde{u}^k)\} = \beta_k \lim_{k\to+\infty} (F(v^k) - F(\tilde{u}^k)) = 0.$$

All the conditions in Theorem 2.1 hold and $\{u^k\}$ converges to a solution point of VI(Ω, F). □

For the accepting rule (9.1) in APPAs for nonlinear VIs, the method proposed by Solodov and Svaiter [29] adopts $\tilde{u}^k$ as the new iterate. Using $d_1(u^k, v^k, \tilde{u}^k)$ defined in (9.2), it can be rewritten as

$$u^{k+1} = u^k - d_1(u^k, v^k, \tilde{u}^k).$$

Thus, using the unified framework, Solodov-Svaiter's APPA is the primary method of form (2.9a).

Remark 9.1 *As long as the parameter β_k is small enough, the accepting rules in the last three sections (see* (3.8), (4.11), *and* (6.7)) *will be satisfied even if $v^k = u^k$ is taken as the approximate solution. In this way we get $\tilde{u}^k$ by*

$$\tilde{u}^k = P_\Omega[u^k - \beta_k F(u^k)]. \tag{9.9}$$

However, the accepting rule (9.1) *cannot be satisfied by setting $v^k = u^k$ for any $\beta_k > 0$. For $u \in \Omega$ and $\tilde{u} = P_\Omega[u - \beta F(u)]$, it follows from* (1.9) *that*

$$\{[u - \beta F(u)] - \tilde{u}\}^T (u - \tilde{u}) \le 0$$

and thus

$$(u - \tilde{u})^T \beta F(u) \ge \|u - \tilde{u}\|^2.$$

From the above inequality and (9.1b) *we obtain*

$$\Delta(u, u, \tilde{u}) \ge \|u - \tilde{u}\|^2.$$

Thus the accepting rule $\Delta(u, u, \tilde{u}) \le 0$ (see (9.1a)) *can never be satisfied.*

10 Effective quadruplet for proximal alternating directions methods

In this section, we define the effective quadruplet of the unified framework for proximal alternating directions methods. Consider the variational inequality problem:

$$(x^*, y^*) \in \mathcal{D}, \quad \left\{ \begin{array}{l} (x - x^*)^T f(x^*) \ge 0, \\ (y - y^*)^T g(y^*) \ge 0, \end{array} \right. \qquad \forall\, (x, y) \in \mathcal{D},$$

where

$$\mathcal{D} = \{(x, y) | x \in \mathcal{X},\ y \in \mathcal{Y},\ Ax - y = 0\},$$

$A \in R^{m \times n}$, $\mathcal{X} \subset R^n$ and $\mathcal{Y} \subset R^m$ are given nonempty closed convex subsets, $f : \mathcal{X} \to R^n$ and $g : \mathcal{Y} \to R^m$ are monotone operators. By attaching a Lagrange multiplier vector $\lambda \in R^m$ to the linear constraints $Ax - y = 0$, this problem can be explained equivalently as the following form: Find

$$u \in \Omega, \qquad \left\{ \begin{array}{l} (x' - x)^T \{f(x) - A^T \lambda\} \ge 0, \\ (y' - y)^T \{g(y) + \lambda\} \ge 0, \\ Ax - y = 0, \end{array} \right. \qquad \forall\, u' \in \Omega \tag{10.1}$$

where

$$\Omega = \mathcal{X} \times \mathcal{Y} \times R^m.$$

Problem (10.1) is referred as a structured variational inequality (SVI for short). The compact form is

$$u \in \Omega, \quad (u' - u)^T F(u) \ge 0, \quad \forall u' \in \Omega,$$

where

$$F(u) = \begin{pmatrix} f(x) - A^T \lambda \\ g(y) + \lambda \\ Ax - y \end{pmatrix}. \tag{10.2}$$

From the current point $u = (x, y, \lambda)$, the new iterate $\tilde{u}$ of the proximal point algorithm is the solution of the following variational inequality:

$$\tilde{u} \in \Omega, \qquad \begin{pmatrix} x' - \tilde{x} \\ y' - \tilde{y} \\ \lambda' - \tilde{\lambda} \end{pmatrix}^T \left\{ \begin{pmatrix} f(\tilde{x}) - A^T\tilde{\lambda} \\ g(\tilde{y}) + \tilde{\lambda} \\ A\tilde{x} - \tilde{y} \end{pmatrix} + \begin{pmatrix} r(\tilde{x} - x) \\ s(\tilde{y} - y) \\ \beta^{-1}(\tilde{\lambda} - \lambda) \end{pmatrix} \right\} \geq 0, \qquad \forall\, u' \in \Omega, \tag{10.3}$$

where $r, s, \beta^{-1} > 0$ are called the proximal coefficients. By using the notation of $F(u)$ (see (10.2)) and $\Omega = \mathcal{X} \times \mathcal{Y} \times R^m$ (in particular, $(A\tilde{x} - \tilde{y}) + \beta^{-1}(\tilde{\lambda} - \lambda) = 0$), (10.3) can be rewritten as

$$\tilde{x} \in \mathcal{X}, \quad (x' - \tilde{x})^T \big\{ f(\tilde{x}) - A^T[\lambda - \beta(A\tilde{x} - \tilde{y})] + r(\tilde{x} - x) \big\} \geq 0, \quad \forall\, x' \in \mathcal{X}, \tag{10.4a}$$

$$\tilde{y} \in \mathcal{Y}, \quad (y' - \tilde{y})^T \big\{ g(\tilde{y}) + [\lambda - \beta(A\tilde{x} - \tilde{y})] + s(\tilde{y} - y) \big\} \geq 0, \qquad \forall\, y' \in \mathcal{Y}, \tag{10.4b}$$

$$\tilde{\lambda} = \lambda - \beta(A\tilde{x} - \tilde{y}). \tag{10.4c}$$

The shortcoming of (10.4) is that subproblems (10.4a) and (10.4b) are coupled. In order to overcome this disadvantage, the proximal alternating directions method [19] generates $\tilde{u}^k = (\tilde{x}^k, \tilde{y}^k, \tilde{\lambda}^k) \in \Omega$ via the following procedure: First find an $\tilde{x}^k \in \mathcal{X}$ such that

$$\tilde{x}^k \in \mathcal{X}, \quad (x' - \tilde{x}^k)^T \big\{ f(\tilde{x}^k) - A^T[\lambda^k - \beta(A\tilde{x}^k - y^k)] + r(\tilde{x}^k - x^k) \big\} \geq 0, \quad \forall\ x' \in \mathcal{X}. \tag{10.5a}$$

Then find a $\tilde{y}^k \in \mathcal{Y}$ such that

$$\tilde{y}^k \in \mathcal{Y}, \quad (y' - \tilde{y}^k)^T \big\{ g(\tilde{y}^k) + [\lambda^k - \beta(A\tilde{x}^k - \tilde{y}^k)] + s(\tilde{y}^k - y^k) \big\} \geq 0, \qquad \forall y' \in \mathcal{Y}. \tag{10.5b}$$

Finally, update $\tilde{\lambda}^k$ via

$$\tilde{\lambda}^k = \lambda^k - \beta(A\tilde{x}^k - \tilde{y}^k). \tag{10.5c}$$

The only difference between (10.4) and (10.5) is to substitute $\tilde{y}^k$ (in (10.4a)) by y^k (in (10.5a)). Since (10.4) is a PPA updating form, the proximal alternating directions method (10.5) can be viewed as an approximate proximal point algorithm for the structured variational inequality (10.1). Note that the alternating directions method in [9, 10] is an extreme case of (10.5) which does not have the proximal terms $r(\tilde{x}^k - x^k)$ and $s(\tilde{y}^k - y^k)$.

By a simple manipulation, (10.5) can be written as: $\tilde{u} = (\tilde{x}, \tilde{y}, \tilde{\lambda}) \in \Omega$,

$$\begin{pmatrix} x' - \tilde{x} \\ y' - \tilde{y} \\ \lambda' - \tilde{\lambda} \end{pmatrix}^T \left\{ \begin{pmatrix} f(\tilde{x}) - A^T\tilde{\lambda} \\ g(\tilde{y}) + \tilde{\lambda} \\ A\tilde{x} - \tilde{y} \end{pmatrix} + \beta \begin{pmatrix} -A^T(y - \tilde{y}) \\ (y - \tilde{y}) \\ 0 \end{pmatrix} + \begin{pmatrix} r(\tilde{x} - x) \\ (\beta + s)(\tilde{y} - y) \\ (\tilde{\lambda} - \lambda) \end{pmatrix} \right\} \geq 0, \quad \forall u' \in \Omega.$$

By using the notation of $F(u)$, a compact form of the proximal alternating directions method is

$$\tilde{u} \in \Omega, \quad (u' - \tilde{u})^T \{ \big(F(\tilde{u}) + \eta(u, \tilde{u}) \big) - G(u - \tilde{u}) \} \geq 0, \quad \forall\, u' \in \Omega, \tag{10.6}$$

where

$$\eta(u, \tilde{u}) = \beta \begin{pmatrix} -A^T(y - \tilde{y}) \\ (y - \tilde{y}) \\ 0 \end{pmatrix} \tag{10.7}$$

and

$$G = \begin{pmatrix} rI_n & & \\ & (\beta + s)I_m & \\ & & \beta^{-1}I_m \end{pmatrix}. \tag{10.8}$$

To force G to be symmetric positive definite, r and β should be set positive. In order to construct the effective quadruplet, however, we set $r,\ s,\ \beta > 0$.

Although v is degenerated thus the equation (1.7) is not exist, we can still follow the unified framework to construct the accepting rule and the effective quadruplet for this proximal alternating directions method. The convergence analysis is also a descendant to Theorem 2.1 with verifying the two additional conditions (2.1a) and (2.1b) under our following proposed accepting rule and effective quadruplet.

(**Accepting rule**) : $(u, \tilde{u})$ satisfies (10.6) with $r,\ s,\ \beta > 0$.

And, we define the effective quadruplet

$$
(\textbf{The quadruplet}) \qquad \begin{cases} d_1(u,v,\tilde{u}) &= u - \tilde{u}, \\ d_2(u,v,\tilde{u}) &= F(\tilde{u}) + \eta(u,\tilde{u}), \\ \varphi(u,v,\tilde{u}) &= \|u-\tilde{u}\|_G^2 - (\lambda - \tilde{\lambda})^T(y-\tilde{y}), \\ \phi(u,v,\tilde{u}) &= \|u-\tilde{u}\|_G^2 - 2(\lambda-\tilde{\lambda})^T(y-\tilde{y}), \end{cases} \tag{10.9}
$$

which involve v in expressions to be coincident with Definition 2.1. In the following we show that the quadruplet is effective which satisfies conditions (2.1) with G defined in (10.8).

Condition (2.1a): Note that u and $\tilde{u}$ in sub-problem (10.6) can be written as

$$
\tilde{u} = P_\Omega\{\tilde{u} - [(F(\tilde{u}) + \eta(u,\tilde{u})) - G(u-\tilde{u})]\}. \tag{10.10}
$$

By combining (10.9) with (10.10), the geminate directions ($d_1(u,v,\tilde{u})$ and $d_2(u,v,\tilde{u})$) satisfy condition (2.1a).

Condition (2.1b): Using the monotonicity of F and $(\tilde{u}-u^*)^T F(u^*) \geq 0$ we have

$$
(\tilde{u}-u^*)^T F(\tilde{u}) \geq 0. \tag{10.11}
$$

Since $Ax^* - y^* = 0$ and $\beta(A\tilde{x} - \tilde{y}) = \lambda - \tilde{\lambda}$, we obtain

$$
\begin{aligned} (\tilde{u}-u^*)^T \eta(u,\tilde{u}) &= (y-\tilde{y})^T \beta(-A\tilde{x} + Ax^* + \tilde{y} - y^*) \\ &= (y-\tilde{y})^T(\tilde{\lambda} - \lambda). \end{aligned}
$$

Thus, it follows from $d_2(u,v,\tilde{u})$ in (10.9) and the above two inequalities that

$$
(\tilde{u}-u^*)^T d_2(u,v,\tilde{u}) \geq (y-\tilde{y})^T(\tilde{\lambda}-\lambda) = \varphi(u,v,\tilde{u}) - \|d_1(u,v,\tilde{u})\|_G. \tag{10.12}
$$

Thus, the triplet $\big(d_1(u,v,\tilde{u}), d_2(u,v,\tilde{u}), \varphi(u,v,\tilde{u})\big)$ satisfies condition (2.1b).

Condition (2.1c): For $d_1(u,v,\tilde{u})$ and $\varphi(u,v,\tilde{u})$ defined in (10.9), we have

$$
2\varphi(u,v,\tilde{u}) - \|d_1(u,v,\tilde{u})\|_G^2 = \|u-\tilde{u}\|_G^2 - 2(\lambda-\tilde{\lambda})^T(y-\tilde{y}) = \phi(u,v,\tilde{u}). \tag{10.13}
$$

Thus condition (2.1c) follows from (10.13) directly.

Condition (2.1d): Because $(\beta + s)\cdot\frac{1}{\beta} > 1$, there exists a constant $\varsigma > 0$ such that

$$
(\beta+s)\|y^k - \tilde{y}^k\|^2 + \frac{1}{\beta}\|\lambda-\tilde{\lambda}\|^2 - 2(\lambda-\tilde{\lambda})^T(y-\tilde{y}) \geq \varsigma\big(\|y-\tilde{y}\|^2 + \|\lambda-\tilde{\lambda}\|^2\big).
$$

It follows from (10.8) and the definition of $\phi(u,v,\tilde{u})$ (10.9) that

$$
\phi(u,v,\tilde{u}) \geq r\|x-\tilde{x}\|^2 + \varsigma\big(\|y-\tilde{y}\|^2 + \|\lambda-\tilde{\lambda}\|^2\big) \geq \min\{r,\varsigma\}\|u-\tilde{u}\|^2.
$$

Condition (2.1d) holds with $\kappa = \min\{r,\varsigma\}$.

Based on the above analysis, we have proved the following theorem.

Theorem 10.1 *For solving problem* (10.1)*, let the duplet* $(u, \tilde{u})$ *be defined in* (10.6)*. Then the quadruplet* $\big(d_1(u,v,\tilde{u}),\ d_2(u,v,\tilde{u}),\ \varphi(u,v,\tilde{u}),\ \phi(u,v,\tilde{u})\big)$ *given by* (10.9) *is an effective quadruplet which fulfills conditions* (2.1) *with* G *defined in* (10.8)*.*

Theorem 10.2 *Let the conditions in Theorem 10.1 be satisfied. Then the sequence* $\{u^k\}$ *generated by the primary methods converges to some* u^∞ *which is a solution point of Problem* (10.1)*.*

Proof: We need only to verify the additional conditions (2.13). Since

$$\phi(u^k, v^k, \tilde{u}^k) \geq \min\{r, \varsigma\}\|u^k - \tilde{u}^k\|^2,$$

it follows from Propositions 2.1 and 2.2 that

$$\|u^{k+1} - u^*\|_G^2 \leq \|u^k - u^*\|_G^2 - \min\{r, \varsigma\}\|u^k - \tilde{u}^k\|^2$$

and thus

$$\lim_{k\to\infty} \|u^k - \tilde{u}^k\| = 0. \tag{10.14}$$

Because $d_1(u^k, v^k, \tilde{u}^k) = (u^k - \tilde{u}^k)$, it follows that

$$\lim_{k\to+\infty} d_1(u^k, v^k, \tilde{u}^k) = 0$$

and condition (2.13a) holds. Note that

$$d_2(u^k, v^k, \tilde{u}^k) - F(\tilde{u}^k) = \eta(u^k, \tilde{u}^k).$$

It follows from (10.7) and (10.14) that

$$\lim_{k\to\infty} \{d_2(u^k, v^k, \tilde{u}^k) - F(\tilde{u}^k)\} = 0$$

and thus condition (2.13b) is satisfied. The proof is complete. □

11 Simple extended contraction method and its numerical results

The purpose of constructing the effective quadruplets for existing APPAs is to develop more efficient contraction methods. Under their own accepting rules, both Solodov-Svaiter's APPA and the proximal alternating directions method take $\tilde{u}^k$ as the new iterate. The updating form can be written as

$$u^{k+1} = u^k - d_1(u^k, v^k, \tilde{u}^k), \tag{11.1}$$

which belongs to the primary method of form (2.9a). In this section, we presents a simple extended contraction method and demonstrates its efficiency.

11.1 The extended contraction method

Under the unified framework, instead of (11.1), we take the new iterate by

(Extended contraction method) $$u^{k+1} = u^k - \gamma\alpha_k^* d_1(u^k, v^k, \tilde{u}^k), \tag{11.2}$$

where

$$\alpha_k^* = \frac{\varphi(u^k, v^k, \tilde{u}^k)}{\|d_1(u^k, v^k, \tilde{u}^k)\|_G^2} \qquad \text{and} \qquad \gamma \in (1, 2). \tag{11.3}$$

Compared with the primary method (11.1), the extended method needs only a few extra computational loads. The following theorem is essential for the convergence of the extended contraction method.

Theorem 11.1 *Assume that the quadruplet* $\big(d_1(u^k,v^k,\tilde u^k),\, d_2(u^k,v^k,\tilde u^k),\, \varphi(u^k,v^k,\tilde u^k),\, \phi(u^k,v^k,\tilde u^k)\big)$ *is effective and the sequence* $\{u^k\}$ *is generated by the extended method* (11.2)-(11.3)*. Then we have*

$$\|u^{k+1}-u^*\|_G^2 \le \|u^k-u^*\|_G^2 - \frac{\gamma(2-\gamma)}{2}\varphi(u^k,v^k,\tilde u^k), \quad \forall u^*\in\Omega^*. \tag{11.4}$$

Proof. Since $u^*\in\Omega$, it follows from (11.2)-(11.3) and (2.5) that

$$\begin{array}{rcl}\|u^{k+1}-u^*\|_G^2 &=& \|u^k-\gamma\alpha^* d_1(u,v,\tilde u)-u^*\|_G^2\\ &=& \|u-u^*\|_G^2 - 2\gamma\alpha^*(u-u^*)^T G d_1(u,v,\tilde u) + (\gamma\alpha^*)^2\|d_1(u,v,\tilde u)\|_G^2\\ &\le& \|u-u^*\|_G^2 - 2\gamma\alpha^*\varphi(u,v,\tilde u) + (\gamma\alpha^*)^2\|d_1(u,v,\tilde u)\|_G^2\\ &=& \|u-u^*\|_G^2 - \gamma(2-\gamma)\alpha^*\varphi(u,v,\tilde u).\end{array}$$

In addition, from (2.1c) we have $\alpha^* > \frac{1}{2}$. Thus, assertion (11.4) is proved. □

In [22], the updating form of Solodov-Svaiter's APPA is extended by

$$u^{k+1} = u^k - \alpha_k(u^k-\tilde u^k).$$

The extended method is just the forms of (11.2)-(11.3) related to the unified framework. Numerical experiments in [22] show that the extended method is much more efficient than the primary one. Thus, in this section, we only compare the efficiency of the proximal alternating directions method and its extended version.

11.2 Test problems and the equivalent structured variational inequality

The problems arise from finance and statistics and we form the test problems similarly as in [8]. Let H_L, H_U and C be given $n\times n$ symmetric matrices, $H_L \le H_U$ in element wise. The problem considered in this section is

$$\min\{\frac{1}{2}\|X-C\|_F^2 \,|\, X\in S_+^n\cap\mathcal{B}\}, \tag{11.5}$$

where $\|\cdot\|_F$ is the matrix Fröbenis norm, *i.e.*, $\|C\|_F = \Big(\sum_{i=1}^n\sum_{j=1}^n |C_{ij}|^2\Big)^{1/2}$,

$$S_+^n = \{H\in R^{n\times n} \,|\, H^T = H,\ \ H\succeq 0\}$$

is the semi-definite cone and

$$\mathcal{B} = \{H\in R^{n\times n} \,|\, H^T=H,\ H_L\le H\le H_U\}. \tag{11.6}$$

Note that the matrix Fröbenis norm is induced by the inner product

$$\langle A,B\rangle = \mathrm{Trace}(A^TB).$$

The data in the test problems.

- The entries of diagonal elements of C are randomly generated in the interval $(0,2)$, the entries of off-diagonal elements of C are randomly generated in the interval $(-1,1)$.
- $(H_U)_{jj} = (H_L)_{jj} = 1$, and $(H_U)_{ij} = -(H_L)_{ij} = 0.1$, $\forall\, i\ne j,\ i,j=1,2,\ldots,n$.

Note that problem (11.5) is equivalent to the following one:

$$\begin{array}{rl} \min & \frac{1}{2}\|X-C\|_F^2 + \frac{1}{2}\|Y-C\|_F^2 \\ \text{s.t} & X - Y = 0, \\ & X \in S_+^n,\ Y \in \mathcal{B}. \end{array} \tag{11.7}$$

The mathematical form of the equivalent structured variational inequality is

$$u = (X, Y, Z) \in \Omega, \quad \left\{ \begin{array}{l} \langle X' - X,\ (X - C) - Z\rangle \ge 0, \\ \langle Y' - Y,\ (Y - C) + Z\rangle \ge 0, \\ X - Y = 0, \end{array} \right. \quad \forall\, u' \in \Omega, \tag{11.8}$$

where

$$\Omega = S_+^n \times \mathcal{B} \times R^{n\times n}. \tag{11.9}$$

11.3 Implementation of the proximal alternating directions method

If we use proximal alternating directions method (10.5) to solve the structured variational inequality (11.8)-(11.9), for given triplet $u^k = (X^k, Y^k, Z^k)$, $\tilde{u}^k = (\tilde{X}^k, \tilde{Y}^k, \tilde{Z}^k)$ is obtained by

$$\tilde{X}^k \in S_+^n, \quad \langle X' - \tilde{X}^k,\ (\tilde{X}^k - C) - [Z^k - \beta(\tilde{X}^k - Y^k)] + r(\tilde{X}^k - X^k)\rangle \ge 0, \quad \forall\ X' \in S_+^n, \tag{11.10a}$$

$$\tilde{Y}^k \in \mathcal{B}, \quad \langle Y' - \tilde{Y}^k,\ (\tilde{Y}^k - C) + [Z^k - \beta(\tilde{X}^k - \tilde{Y}^k)] + s(\tilde{Y}^k - Y^k)\rangle \ge 0, \quad \forall\ Y' \in \mathcal{B}, \tag{11.10b}$$

and

$$\tilde{Z}^k = Z^k - \beta(\tilde{X}^k - \tilde{Y}^k). \tag{11.10c}$$

Using the principle of (1.2) and the special structures of (11.10a) and (11.10b), $\tilde{X}^k$ and $\tilde{Y}^k$ can be directly obtained by

$$\tilde{X}^k = P_{S_+^n}\{\frac{1}{1+\beta+r}(\beta Y^k + Z^k + C + rX^k)\} \tag{11.11}$$

and

$$\tilde{Y}^k = P_{\mathcal{B}}\{\frac{1}{1+\beta+s}(\beta \tilde{X}^k - Z^k + C + sY^k)\}, \tag{11.12}$$

respectively. For given symmetric matrix $A \in R^{n\times n}$, let its eigenvalue decomposition be

$$A = V\Lambda V^T, \tag{11.13}$$

where $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$. Then

$$P_{S_+^n}(A) = V\tilde{\Lambda}V^T, \tag{11.14}$$

where

$$\tilde{\Lambda} = \mathrm{diag}(\tilde{\lambda}_1, \ldots, \tilde{\lambda}_n), \qquad \tilde{\lambda}_i = \max\{0, \lambda_i\}.$$

Note that the computational loads of (11.13) and (11.14) are about $9n^3$ and n^3 flops, respectively [12]. The projection $P_{\mathcal{B}}(A)$ is easy to be carried out, namely, in element wise

$$P_{\mathcal{B}}(A) = \max\{H_L, \min\{A, H_U\}\}.$$

For this matrix optimization problem, using the notations in Section 10, we have

$$u = \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \quad \text{and} \quad G = \begin{pmatrix} rI & & \\ & (\beta+s)I & \\ & & \frac{1}{\beta}I \end{pmatrix}. \tag{11.15}$$

According to the effective quadruplet defined in Theorem 10.1, for this problem

$$d_1(u,v,\tilde{u}) = u - \tilde{u}, \tag{11.16}$$

and

$$\varphi(u,v,\tilde{u}) = r\|X-\tilde{X}\|_F^2 + (\beta+s)\|Y-\tilde{Y}\|_F^2 + \frac{1}{\beta}\|Z-\tilde{Z}\|_F^2 - \langle Y^k-\tilde{Y}^k,\, Z^k-\tilde{Z}^k\rangle. \tag{11.17}$$

In line of preliminary method (see Definition 2.2),

$$u^{k+1} = u^k - d_1(u^k, v^k, \tilde{u}^k).$$

Since $d_1(u,v,\tilde{u}) = u - \tilde{u}$, the proximal alternating directions method adopts $\tilde{u}^k$ as the new iterate, the iteration form can be written as

$$\begin{pmatrix} X^{k+1} \\ Y^{k+1} \\ Z^{k+1} \end{pmatrix} = \begin{pmatrix} X^k \\ Y^k \\ Z^k \end{pmatrix} - \begin{pmatrix} X^k-\tilde{X}^k \\ Y^k-\tilde{Y}^k \\ Z^k-\tilde{Z}^k \end{pmatrix} = \begin{pmatrix} \tilde{X}^k \\ \tilde{Y}^k \\ \tilde{Z}^k \end{pmatrix}. \tag{11.18}$$

The most time consuming operation is to compute $\tilde{X}^k$ in (11.11) which is about $10n^3$ flops.

11.4 Implementation of the extended contraction method

Let $(\tilde{X}^k, \tilde{Y}^k, \tilde{Z}^k)$ be generated by (11.10) from given (X^k, Y^k, Z^k). In line with the extended contraction method with updating form (11.2)-(11.3), the new iterate is generated by

$$\begin{pmatrix} X^{k+1} \\ Y^{k+1} \\ Z^{k+1} \end{pmatrix} = \begin{pmatrix} X^k \\ Y^k \\ Z^k \end{pmatrix} - \gamma\alpha_k^* \begin{pmatrix} X^k-\tilde{X}^k \\ Y^k-\tilde{Y}^k \\ Z^k-\tilde{Z}^k \end{pmatrix}. \tag{11.19}$$

According to (11.3) and (11.15)-(11.17), we have

$$\alpha_k^* = \frac{\|u^k-\tilde{u}^k\|_G^2 - \langle Y^k-\tilde{Y}^k,\, Z^k-\tilde{Z}^k\rangle}{\|u^k-\tilde{u}^k\|_G^2}$$

where

$$\|u^k-\tilde{u}^k\|_G^2 = r\|X^k-\tilde{X}^k\|_F^2 + (\beta+s)\|Y^k-\tilde{Y}^k\|_F^2 + \frac{1}{\beta}\|Z^k-\tilde{Z}^k\|_F^2.$$

Note that the computational load for α_k^* is $4n^2$. In comparison with the cost for getting $\tilde{X}^k$ in (11.11), the extra work in the extended contraction method is negligible.

Remark 11.1 *The procedures* (11.10) *and* (11.19) *are only provided for the comparison of the primary proximal alternating directions method and its extended one with the effective quadruplet proposed in Section 10. It should be noticed that, for solving the similar optimization problem* (11.5) *(with less than* 5% *inequality constraints on the off-diagonal elements) the most recent method in [8] reported the better numerical performance.*

11.5 Numerical results

We test the problem with the following data.

- The entries of diagonal elements of C are randomly generated in the interval $(0,2)$, the entries of off-diagonal elements of C are randomly generated in the interval $(-1,1)$.

- $(H_U)_{jj} = (H_L)_{jj} = 1$, and $(H_U)_{ij} = -(H_L)_{ij} = 0.1$, $\forall\, i \neq j$, $i, j = 1, 2, \ldots, n$.

We take $u^0 = (X^0, Y^0, Z^0) = (I_n, I_n, 0_n)$ as the initial point in the test. The iteration is stopped as soon as

$$\frac{\max(\text{abs}(u^k - \tilde{u}^k))}{\max(\text{abs}(u^0 - \tilde{u}^0))} \leq \varepsilon = 10^{-6}.$$

For $r = s = 1$, $\beta = 10$ and the relaxation factor $\gamma = 1.5$, Table 11.1 reports the iteration numbers and the CPU times of the various methods. Since the complexity of each iteration is $O(n^3)$ (about $10n^3$), the CPU time is proportional to the product of the iteration number by n^3. From Table 11.1 we find

$$\frac{\text{It No. of the extended contraction method}}{\text{It No. of the proximal alternating directions method}} \approx 0.60.$$

The improvement of the extended contraction method is significant.

Table 11.1. Numerical results for $r = s = 1$, $\beta = 10$ and $\gamma = 1.5$

$n \times n$ Matrix	Primary Method Updating form (11.18)		Extended Method Updating form (11.19)	
n	No. It	CPU Sec.	No. It	CPU Sec.
100	71	1.07	46	0.85
200	67	4.08	44	2.97
500	79	40.44	49	28.28
1000	91	367.46	56	250.58

For these test problems, Table 11.2 shows that the influence of the different proximal coefficients is insensitive. The iteration numbers of the extended contraction method by using different relaxation factors are reported in Table 11.3. Indeed, as the numerical experiences in [20, 23], a good experiential choice of the relaxation factor is $\gamma \in [1.2, 1.8]$.

Table 11.2. Iteration number for proximal coefficients $r = s = 1$ and different β

$n \times n$ Matrix	Primary Method Updating form (11.18)				Extended Method Updating form (11.19), $\gamma = 1.5$			
n	$\beta = 2$	$\beta = 10$	$\beta = 30$	$\beta = 50$	$\beta = 2$	$\beta = 10$	$\beta = 30$	$\beta = 50$
100	285	71	167	269	187	46	108	176
200	371	67	149	240	245	44	96	156
500	443	79	146	235	292	49	94	153
1000	499	91	149	240	329	56	95	156

Table 11.3. Iteration number for different relaxation factor γ $(r = s = 1,\ \beta = 10)$

Matrix$_{n \times n}$	Extended method with updating form (11.19)						
n	$\gamma = 0.8$	$\gamma = 1.0$	$\gamma = 1.2$	$\gamma = 1.5$	$\gamma = 1.8$	$\gamma = 1.9$	$\gamma = 2.0$
100	85	71	57	46	69	145	div
200	82	67	54	44	71	147	div
500	97	79	63	49	69	145	div
1000	111	91	72	56	68	140	div

12 Concluding remarks

In this paper, we introduce a unified framework for proximal-like contraction methods. The framework is based on an effective quadruplet along with an accepting rule. For symmetric linear VIs, linear VIs and nonlinear VIs, we have constructed their respective effective quadruplets and accepting rules. With these effective quadruplets and rules, various existing APPAs can be viewed as primary methods deduced by the framework. Other than the primary methods, we have also proposed the general contraction methods, in which we can use the convex combinations of the geminate directions in the quadruplet as the search directions with selected step lengths. Besides the theoretical comparisons on the efficiency of the different directions and step lengths, we present numerous numerical results which clearly confirm with our theoretical results and analysis. From the numerical results, the numbers of the iterations and function evaluations are reduced significantly for the general contraction methods. Our numerical experiments also indicate that, special properties such as symmetry, linearity, etc., should be considered in solving these problems.

For the two exiting popular methods, i.e., Solodov and Svaiter's APPA and the proximal alternating directions method, we present their respective effective quadruplets and accepting rules. Based on these two methods, we can construct more efficient corresponding methods (called extended methods) under our new framework but with only minor extra costs. As an application, we test a matrix approximation problem to compare the efficiency of the proximal alternating directions method and its extended method. From the numerical results, the efficiency of the extended method is significant and convincing.

For our final remark, under our new unified framework, many existing proximal-like methods can be generated directly and therefore viewed as a class of the primal methods, at the same time, many new efficient methods (which can be viewed as the extension of the primal ones) could be also constructed.

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. Springer-Verlag, Berlin-Heidelberg-New York (1975).

[3] Y.H. Dai and Y. Yuan. Alternate minimization gradient method. IMA J. Numer. Anal. 23, 377-393 (2003).

[4] J. Eckstein, Nonlinear proximal point algorithms using Bregman functions, with applications to convex programing. Math. Oper. Res. 18, 202-226 (1993).

[5] J. Eckstein, Approximate iterations in Bregman-function-based proximal algorithms. Math. Program. 83, 113-123 (1998).

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

[7] R. Fletcher, Practical Methods of Optimization, Second Edition. John Wiley & Sons, (1987).

[8] Y. Gao and D.F. Sun, Calibrating least squares covariance matrix problems with equality and inequality constraints, Technical Report, Department of Mathematics, National University of Singapore, 2008. http://www.math.nus.edu.sg/ matsundf/

[9] R. Glowinski, Numerical Methods for Nonlinear Variational Problems, Springer-Verlag, New York, Berlin, Heidelberg, Tokyo, 1984.

[10] R. Glowinski and P. Le Tallec, Augmented Lagrangian and Operator Splitting Methods in Nonlinear Mechanics, SIAM Studies in Applied Mathematics, Philadelphia, PA, 1989.

[11] P.T. Harker and J.S. Pang, A Damped-Newton Method for the Linear Complementarity Problem. Lectures in Applied Mathematics, 26, 265–284 (1990).

[12] G. H. Golub and C. F. van Loan, Matrix Computation, third edition, The Johns Hopkins University Press, 1996.

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

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

[15] B. S. He, Solving a class of linear projection equations. Numerische Mathematik. 68, 71–80 (1994).

[16] B.S. He, Some predict-correct projection methods for monotone variational inequalities. Report 95-68, Faculty of Technical Mathematics and Informatics, Delf University of Technology. (1995).

[17] B.S. He, A class of projection and contraction methods for monotone variational inequalities. Appl. Math. Optim. 35, 69–76 (1997).

[18] B.S. He, Inexact implicit methods for monotone general variational inequalities. Math. Program. 86, 199–217 (1999).

[19] B.S. He, L.-Z. Liao, D.R. Han and H. Yang A new inexact alternating directions method for monotone variational inequalities. Math Program. 92, 103–118, (2002).

[20] B.S He and L.-Z Liao, Improvements of some projection methods for monotone nonlinear variational inequalities. J. Optim. Theory Appl. 112, 111–128 (2002).

[21] B.S. He and M. H. Xu, A general framework of contraction methods for monotone variational inequalities. Pac. J. Optim. 4, 195–212 (2008).

[22] B.S. He, Z.H. Yang and X. M. Yuan, An Approximate proximal-extragradient type method for monotone variational inequalities, J. Mathematical Analysis and Applications **300** 362-374, 2004

[23] B.S. He, X.M. Yuan and J.J.Z. Zhang, Comparison of two kinds of prediction-correction methods for monotone variational inequalities. Comput. Optim. Appl. 27, 247–267 (2004).

[24] E. N. Khobotov, Modification of the extragradient method for solving variational inequalities and certain optimization problems. USSR. Comput. Math. Phys. 27, 120–127 (1987).

[25] G. M. Korpelevich, The Extragradient Method for Finding Saddle Points and Other Problems. Ekonomika i Matematchskie Metody. 12, 747-756 (1976).

[26] R.T. Rockafellar, Augmented Lagrangians and applications of the proximal point algorithm in convex programming. Math. Oper. Res. 1, 97–116, (1976).

[27] R.T. Rockafellar, Monotone operators and the proximal point algorithm. SIAM, J. Control Optim. 14, 877–898 (1976).

[28] M.V. Solodov and B.F. Svaiter, A hybrid projection-proximal point algorithm. J. Convex Anal. 6, 59–70 (1999).

[29] M.V. Solodov and B.F. Svaiter, Error bounds for proximal point subproblems and associated inexact proximal point algorithms. Math. Program. 88, 371–389, (2000).

[30] M.V. Solodov and B.F. Svaiter, A unified framework for some inexact proximal point algorithms. Numer. Funct. Anal. Optim. 22, 1013-1035, 2001.

[31] M.V. Solodov and P. Tseng, Modified projection-type methods for monotone variational inequalities. SIAM J. Control Optim. 34, 1814–1830 (1996).

[32] K. Taji, M. Fukushima and I. Ibaraki, A globally convergent Newton mmethod for solving strongly monotone variational inequalities. Math. Program. 58, 369-383 (1993).

[33] P. Tseng, A Modified forward-backword splitting method for maximal monotone mappings. SIAM. J. Control Optim. 38, 431–446 (2000).

[34] T. Zhu and Z. G. Yu, A simple proof for some important properties of the projection mapping. Math. Inequal. Appl. 7, 453–456 (2004).