Comput Optim Appl (2012) 51:649–679
DOI 10.1007/s10589-010-9372-0

Proximal-like contraction methods for monotone variational inequalities in a unified framework I: Effective quadruplet and primary methods

Bingsheng He · Li-Zhi Liao · Xiang Wang

Received: 27 May 2009 / Published online: 18 November 2010

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 directions, i.e., d_1^k and d_2^k for each step. In this paper, for various APPAs we 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 directions (d_1^k and d_2^k) those studied proximal-like methods always utilize the unit step size namely $\alpha_k \equiv 1$.

Keywords Variational inequality · Monotone · Contraction methods

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.

B. He (✉) · X. Wang
Department of Mathematics, Nanjing University, Nanjing, 210093, China
e-mail: hebma@nju.edu.cn

B. He
National Key Laboratory for Novel Software Technology, Nanjing University, Nanjing, China

L.-Z. Liao
Department of Mathematics, Hong Kong Baptist University, Hong Kong, China

1 Introduction

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

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

$\text{VI}(\Omega, F)$ 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

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

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

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

The variational inequality $\text{VI}(\Omega, F)$ is monotone when the mapping F is monotone. For solving a monotone variational inequality, a classical method is the proximal point algorithm (abbreviated as PPA) [17, 18]. 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:

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

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

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

where u^* is any solution point of $\text{VI}(\Omega, F)$ (for a proof, see [10] or [18]). 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, 20].

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

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

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 $F(\tilde{u}^k) = F(v^k)$. In this paper, for constructing various methods for VI(Ω, F), we rewrite the basic equation of APPAs (1.7) 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 two directions namely $d_1(u^k, v^k, \tilde{u}^k)$ and $d_2(u^k, v^k, \tilde{u}^k)$, where G is a symmetric positive definite matrix. In the following, we call $d_1(u^k, v^k, \tilde{u}^k)$ and $d_2(u^k, v^k, \tilde{u}^k)$ the *geminate directions*.

Because (1.8) is derived from the basic equation of APPAs, the methods developed in this paper are called *proximal-like methods*. Indeed, the 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 and 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 some more efficient methods than the primary ones with only minor extra computational loads. In this paper, however, we only propose the unified framework and different explicit formulations of the equation (1.8) inducing many existing methods for various kinds of monotone VIs. The further studies on more efficient methods are put forth in the descendent paper "Proximal-like contraction methods for monotone variational inequalities in a unified framework II", see [13].

Throughout this paper we assume that the operator F is point-to-point, 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., $0 < \beta_L \le \inf_{k=0}^{\infty} \beta_k \le \sup_{k=0}^{\infty} \beta_k \le \beta_U < +\infty$. Note that under our assumptions the solution set Ω^* is closed and convex (see pp. 158 in [6]). The projection mapping is a tool for the analysis in this paper. It has some important properties as shown in Lemma 1.1, which can be found in textbooks, e.g., [2].

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

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

$$\|P_\Omega(u) - P_\Omega(u')\| \le \|u - u'\|, \quad \forall u, \ u' \in R^n, \tag{1.10}$$

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

The paper is organized as follows. In Sect. 2, we propose the unified framework which consists of an effective quadruplet and an accepting rule. Under the framework, a class of primary methods is defined, followed by detailed analysis on their convergence.

For the rest of the paper, guided by the unified framework, we first 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 in Sects. 3, 4, 5 and 6, respectively. Finally, we illustrate how to construct effective quadruplets for two existing APPAs, i.e., Solodov and Svaiter's APPA (see Algorithm 2 in [20]) and the proximal alternating directions method (see [11]), abiding by their own accepting rules, in Sects. 7 and 8, respectively.

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 $(d_1(u^k, v^k, \tilde{u}^k), d_2(u^k, v^k, \tilde{u}^k), \varphi(u^k, v^k, \tilde{u}^k), \phi(u^k, v^k, \tilde{u}^k))$ 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 that

$$\tilde{u}^k = P_\Omega\{\tilde{u}^k - [d_2(u^k, v^k, \tilde{u}^k) - G d_1(u^k, v^k, \tilde{u}^k)]\} \tag{2.1a}$$

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

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

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

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

   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

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

According to the above four conditions (2.1a)–(2.1d), APPAs can be derived, in which v^k is 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 usually on β (see (1.7)). 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), the condition (2.1a) implies that

$$\tilde{u} \in \Omega, \qquad (u' - \tilde{u})^T \{d_2(u, v, \tilde{u}) - G d_1(u, v, \tilde{u})\} \geq 0, \quad \forall u' \in \Omega. \tag{2.2}$$

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 $(d_1(u, v, \tilde{u}), d_2(u, v, \tilde{u}), \varphi(u, v, \tilde{u}), \phi(u, v, \tilde{u}))$ 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

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

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

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

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.

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 $\frac{1}{2}\|u - u^*\|_G^2$ 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 $\frac{1}{2}\|u-u^*\|_G^2$ and $\frac{1}{2}\|u-u^*\|^2$ when $u \in \Omega \setminus \Omega^*$, respectively. The assertions are similar to Lemmas 3.1 and 3.2 in [14]. For completeness, the proofs are provided.

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

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

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

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

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

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

Assertion (2.5) follows from the above inequality directly. □

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 geminate descent directions ($d_1(u,v,\tilde{u})$ and $d_2(u,v,\tilde{u})$) provided by the effective quadruplet have similar properties (see (2.5) and (2.7)). However, there are two important differences:

- The 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 only true for all $u \in \Omega$ in Lemma 2.2.

Definition 2.2 According to the effective quadruplet, a method is called primary (or elementary) method when the new iterate u^{new} is generated by one of the following equations. The first kind updating form is

$$u^{\text{new}} = u - d_1(u,v,\tilde{u}), \quad \text{for all positive definite } G \text{ in (2.1)}. \tag{2.9a}$$

And the other two kinds are,

$$u^{\text{new}} = P_\Omega[u - d_1(u,v,\tilde{u})], \quad \text{when } G \text{ in (2.1) is the identity matrix}, \tag{2.9b}$$

and

$$u^{\text{new}} = P_\Omega[u - d_2(u,v,\tilde{u})], \quad \text{when } G \text{ in (2.1) is the identity matrix}. \tag{2.9c}$$

Note that, in the *primary methods*, we use the different search directions offered by the framework. The new iterates are updated with the unit step length 1 (with or without 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 method is generated by one of 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), we get

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

The last inequality is followed from (2.1c) 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 a certain accepting rule. Assume that the quadruplet $(d_1(u^k,v^k,\tilde{u}^k), d_2(u^k,v^k,\tilde{u}^k), \varphi(u^k,v^k,\tilde{u}^k), \phi(u^k,v^k,\tilde{u}^k))$ is effective and the sequence $\{u^k\}$ is generated by one of 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

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

where G is a positive definite matrix or the identity matrix. It follows from (2.14) that

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

and $\{u^k\}$ is bounded. Together with the condition (2.1d), it holds that

$$\lim_{k\to\infty}\|u^k-\tilde{u}^k\|=0. \tag{2.15}$$

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

$$\tilde{u}^{k_j}\in\Omega,\qquad (u'-\tilde{u}^{k_j})^T\left\{d_2(u^{k_j},v^{k_j},\tilde{u}^{k_j})-Gd_1(u^{k_j},v^{k_j},\tilde{u}^{k_j})\right\}\ge 0,\quad \forall u'\in\Omega.$$

Table 1 The mapping $F(u)$ in VIs of the different types

Type 1	$Hu+q$	$H^T = H$	Linear	&	Symmetric
Type 2	$Mu+q$	$M^T \neq M$	Linear	&	Asymmetric
Type 3	$\nabla f(u)$	$\nabla^2 f(u)$ is symmetric	Nonlinear	&	Symmetric
Type 4	$F(u)$	$\nabla F(u)$ is asymmetric	Nonlinear	&	Asymmetric

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 condition (2.13) that

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

The above variational inequality indicates that u^∞ is a solution of VI(Ω, F). By using (2.15) 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

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

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.

We divide the variational inequality (1.1) in four types according to the mapping $F(u)$ in VI(Ω, F) as shown in Table 1.

From Sect. 3 to Sect. 6, we will investigate the applications of the unified framework for VIs of the four different types with $G = I$. We construct the effective quadruplet in the following procedure:

1. First, according to the basic equation of APPAs (1.7), we find the geminate directions $d_1(u^k, v^k, \tilde{u}^k)$ and $d_2(u^k, v^k, \tilde{u}^k)$ which satisfy condition (2.1a).
2. After getting this pair of directions, we try to find an $\omega(u^k, v^k, \tilde{u}^k) \in R^n$ such that

$$(\tilde{u} - u^*)^T d_2(u, v, \tilde{u}) \ge \omega(u, v, \tilde{u}) \tag{2.16}$$

   and then let

$$\varphi(u, v, \tilde{u}) = \omega(u, v, \tilde{u}) + (u - \tilde{u})^T d_1(u, v, \tilde{u}). \tag{2.17}$$

   In this way condition (2.1b) is satisfied straightforwardly.
3. Finally, we find the accepting rule to guarantee conditions (2.1c) and (2.1d) in the unified framework.

3 Application to symmetric monotone linear VIs

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

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

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

$$\min\left\{\frac{1}{2}u^T Hu + q^T u \mid u \in \Omega\right\}. \tag{3.2}$$

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 $(d_1(u, v, \tilde{u}), d_2(u, v, \tilde{u}), \varphi(u, v, \tilde{u}), \phi(u, v, \tilde{u}))$ 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

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

and it can be rewritten as

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

Note that the left-hand-side of above inequality is $(\tilde{u} - u^*)^T d_2(u, v, \tilde{u})$. 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[1]

$$(\tilde{u} - u^*)^T d_2(u, v, \tilde{u}) \geq -\frac{1}{2}(v - \tilde{u})^T \beta H(v - \tilde{u}), \quad \forall u^* \in \Omega^*. \tag{3.6}$$

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

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

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

[1] The more tight inequality $x^T Hy \geq -\frac{1}{4}\|x - y\|_H^2$ may lead $(\tilde{u} - u^*)^T \beta(Hv + q) \geq -\frac{1}{4}\beta\|v - \tilde{u}\|_H^2$.

and defining

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

we get

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

thus the condition (2.1c) is satisfied.

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

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* $(d_1(u, v, \tilde{u}), d_2(u, v, \tilde{u}), \varphi(u, v, \tilde{u}), \phi(u, v, \tilde{u}))$ *given by*

$$\text{(The quadruplet)} \qquad \begin{cases} d_1(u, v, \tilde{u}) = u - \tilde{u}, \\ d_2(u, v, \tilde{u}) = \beta(Hv + q), \\ \varphi(u, v, \tilde{u}) = \|u - \tilde{u}\|^2 - \frac{1}{2}(v - \tilde{u})^T \beta H(v - \tilde{u}), \\ \phi(u, v, \tilde{u}) = (1-\mu)\|u - \tilde{u}\|^2 \end{cases} \tag{3.10}$$

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

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

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 the 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

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

Note that

$$(u^k - \tilde{u}^k)^T \beta_k H(u^k - \tilde{u}^k) \le \beta_k \lambda_{\max} \|u^k - \tilde{u}^k\|^2,$$

where $\lambda_{\max}$ is the maximal eigenvalue of the symmetric positive semi-definite matrix H. Therefore, to satisfy (3.12b), β_k can always be selected as long as $\beta_k \le \mu/\lambda_{\max}$ with a certain positive lower bound (e.g. $0.1\mu/\lambda_{\max}$). 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

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

A small β will guarantee that the accepting rule (3.12b) is 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). $\beta_0 = 1$, $\mu = 0.9$.

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

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

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

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

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

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

where

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

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 (see (3.12a))

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

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

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

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

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

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)

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

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 $(d_1(u, v, \tilde{u}), d_2(u, v, \tilde{u}), \varphi(u, v, \tilde{u}), \phi(u, v, \tilde{u}))$ 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

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

and consequently

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

Adding the identity

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

to the both sides of (4.6), and using the definition of $d_2(u,v,\tilde{u})$ and $(v-u^*)^T\beta M^T \times (v-u^*) \geq 0$, we obtain

$$(\tilde{u}-u^*)^T d_2(u,v,\tilde{u}) \geq -(v-\tilde{u})^T\beta M^T(v-\tilde{u}). \tag{4.7}$$

By defining

$$\varphi(u,v,\tilde{u}) = (u-\tilde{u})^T d_1(u,v,\tilde{u}) - (v-\tilde{u})^T\beta M^T(v-\tilde{u}), \tag{4.8}$$

the inequality (4.7) 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 $(d_1(u,v,\tilde{u}), d_2(u,v,\tilde{u}), \varphi(u,v,\tilde{u}))$ 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 (4.8), by a straightforward manipulation, we get

$$\begin{aligned}
&2\varphi(u,v,\tilde{u}) - \|d_1(u,v,\tilde{u})\|^2 \\
&= 2(u-\tilde{u})^T d_1(u,v,\tilde{u}) - \|d_1(u,v,\tilde{u})\|^2 - 2(v-\tilde{u})^T\beta M^T(v-\tilde{u}) \\
&= d_1(u,v,\tilde{u})^T\big(2(u-\tilde{u}) - d_1(u,v,\tilde{u})\big) - 2(v-\tilde{u})^T\beta M^T(v-\tilde{u}) \\
&= \big((u-\tilde{u}) + \beta M^T(v-\tilde{u})\big)^T\big((u-\tilde{u}) - \beta M^T(v-\tilde{u})\big) - 2(v-\tilde{u})^T\beta M^T(v-\tilde{u}) \\
&= \|u-\tilde{u}\|^2 - \{2(v-\tilde{u})^T\beta M^T(v-\tilde{u}) + \|\beta M^T(v-\tilde{u})\|^2\}.
\end{aligned} \tag{4.9}$$

Let the approximate solution v be accepted when the accepting rule

$$\text{(Accepting rule)} \qquad 2(v-\tilde{u})^T\beta M^T(v-\tilde{u}) + \|\beta M^T(v-\tilde{u})\|^2 \leq \mu\|u-\tilde{u}\|^2, \quad \mu \in (0,1) \tag{4.10}$$

is satisfied. Then, we denote

$$\phi(u,v,\tilde{u}) := (1-\mu)\|u-\tilde{u}\|^2. \tag{4.11}$$

It follows from (4.9), (4.10) and (4.11) that

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

and thus condition (2.1c) is satisfied.

Condition (2.1d): From (4.11) condition (2.1d) holds with $\kappa = (1-\mu)$.
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* (4.10) *be satisfied. Then the quadruplet* $(d_1(u, v, \tilde{u}), d_2(u, v, \tilde{u}), \varphi(u, v, \tilde{u}), \phi(u, v, \tilde{u}))$ *given by*

$$\text{(The quadruplet)} \qquad \begin{cases} d_1(u, v, \tilde{u}) = (u - \tilde{u}) + \beta M^T (v - \tilde{u}), \\ d_2(u, v, \tilde{u}) = \beta(Mv + q) + \beta M^T (v - \tilde{u}), \\ \varphi(u, v, \tilde{u}) = (u - \tilde{u})^T d_1(u, v, \tilde{u}) - (v - \tilde{u})^T \beta M^T (v - \tilde{u}), \\ \phi(u, v, \tilde{u}) = (1 - \mu)\|u - \tilde{u}\|^2 \end{cases} \tag{4.12}$$

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

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

and thus

$$\lim_{k\to\infty} \|u^k - \tilde{u}^k\| = 0. \tag{4.13}$$

Since M is positive semi-definite, it follows from the accepting rule (4.10) that

$$\|\beta M^T (v^k - \tilde{u}^k)\| \le \mu\|u^k - \tilde{u}^k\|^2. \tag{4.14}$$

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 (4.13) and (4.14) that

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

and the 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 follows from the accepting rule (4.10)

$$\beta(v^k - \tilde{u}^k)^T (M + M^T)(v^k - \tilde{u}^k) = 2\beta(v^k - \tilde{u}^k)^T M^T (v^k - \tilde{u}^k) \le \mu\|u^k - \tilde{u}^k\|^2. \tag{4.15}$$

From the symmetry and positive semi-definiteness of $(M + M^T)$, (4.13) and (4.15), we obtain

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

and thus the 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

$$2(u^k - \tilde{u}^k)\beta_k M^T(u^k - \tilde{u}^k) + \|\beta_k M^T(u^k - \tilde{u}^k)\|^2 \le \mu\|u^k - \tilde{u}^k\|^2, \quad \mu \in (0, 1) \tag{4.16}$$

will be satisfied when the parameter β_k is small enough.

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

$$2(u^k - \tilde{u}^k)\beta_k M^T(u^k - \tilde{u}^k) + \|\beta_k M^T(u^k - \tilde{u}^k)\|^2 \le \mu_0\|u^k - \tilde{u}^k\|^2,$$

where μ_0 is relative smaller than μ. In order to find a suitable parameter β_k, the following procedure is recommended, in which we set $\mu_0 = 0.3$ based on our numerical experiments.

Procedure 4.1 Finding β_k to satisfy (4.16). $\beta_0 = 1$, $\mu = 0.9$.

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

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

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

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

(Accepting rule (4.16))

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

5 Application to symmetric monotone nonlinear VIs

In VI(Ω, F), when $F(u)$ is the gradient of a certain function, say $f(u)$, the Jacobian 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

$$\min\left\{\frac{1}{2}f(u) \mid u \in \Omega\right\}. \tag{5.1}$$

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 $(d_1(u, v, \tilde{u}), d_2(u, v, \tilde{u}), \varphi(u, v, \tilde{u}), \phi(u, v, \tilde{u}))$ 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 the 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

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

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

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

Note that the left-hand-side of (5.6) 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 $(d_1(u, v, \tilde{u}), d_2(u, v, \tilde{u}), \varphi(u, v, \tilde{u}))$ satisfies the condition (2.1b).

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

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

and defining

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

we get

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

thus the condition (2.1c) is satisfied.

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

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 a 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* $(d_1(u, v, \tilde{u}), d_2(u, v, \tilde{u}), \varphi(u, v, \tilde{u}), \phi(u, v, \tilde{u}))$ *given by*

$$\text{(The quadruplet)} \qquad \begin{cases} d_1(u, v, \tilde{u}) = u - \tilde{u}, \\ d_2(u, v, \tilde{u}) = \beta F(v), \\ \varphi(u, v, \tilde{u}) = \|u - \tilde{u}\|^2 - (v - \tilde{u})^T \beta(F(v) - F(\tilde{u})), \\ \phi(u, v, \tilde{u}) = (1 - \mu)\|u - \tilde{u}\|^2 \end{cases} \tag{5.10}$$

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

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

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 the 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 the 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

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

In other words, if F is Lipschitz continuous, by setting a 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

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

A small β will guarantee that the accepting rule (5.12b) is 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). $\beta_0 = 1, \mu = 0.9$.

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

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

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

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

6 Application to nonlinear monotone VIs

We consider the nonlinear monotone variational inequality (1.1). The mapping F in this section is nonlinear and ∇F (if it exists) is asymmetric. For the triplet $(u, v, \tilde{u})$ in (1.7), under the unified framework, we will find an effective quadruplet $(d_1(u, v, \tilde{u}), d_2(u, v, \tilde{u}), \varphi(u, v, \tilde{u}), \phi(u, v, \tilde{u}))$ 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 $(\tilde{u} - u^*)^T F(u^*) \geq 0$. Using (6.3) and the monotonicity of F we have

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

By letting

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

the triplet $(d_1(u, v, \tilde{u}), d_2(u, v, \tilde{u}), \varphi(u, v, \tilde{u}))$ 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

$$\text{(Accepting rule)} \qquad \beta\|F(v) - F(\tilde{u})\| \leq \mu\|u - \tilde{u}\|, \quad \mu \in (0, 1) \tag{6.7}$$

is satisfied. Define

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

Therefore,

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

and thus the condition (2.1c) is satisfied.

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

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* $(d_1(u, v, \tilde{u}), d_2(u, v, \tilde{u}), \varphi(u, v, \tilde{u}), \phi(u, v, \tilde{u}))$ *given by*

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

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

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

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

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

There are many primary methods in the literature which can be generated by the quadruplet $(d_1(u, v, \tilde{u}), d_2(u, v, \tilde{u}), \varphi(u, v, \tilde{u}), \phi(u, v, \tilde{u}))$ in (6.9). To mention a few, see [9, 16, 19, 21].

- In fact, the updating form of the prediction-correction methods in [9] 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 [19] is designed mainly for the point-to-set variational inequality which includes the point-to-point variational inequalities considered in this paper. When F is a point-to-point mapping, the hybrid projection-proximal point algorithm [19] 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 [21] generates the new iterate via

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

  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 [15, 16] produces the new iterate by

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

  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 results of these primary methods are consequences 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* $\text{VI}(\Omega, F)$.

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

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

and thus

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

From the accepting rule (6.7) and (6.16), it follows that

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

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 $\mathrm{VI}(\Omega, F)$. □

For the same reason of achieving faster convergence as mentioned in the previous three 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). $\beta_0 = 1$, $\mu = 0.9$.

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

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

UNTIL: $\beta_k\|F(u^k)-F(\tilde{u}^k)\| \le \mu\|u^k-\tilde{u}^k\|$. (Accepting rule (6.10b))

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

Indeed, for the different types of monotone VIs in Table 1, we have constructed their respective effective quadruplets and accepting rules. The key idea is to find the geminate directions $d_1(u^k,v^k,\tilde{u}^k)$ and $d_2(u^k,v^k,\tilde{u}^k)$ which satisfy the condition (2.1a) and to show that

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

Table 2 lists the function $\omega(u,v,\tilde{u})$ in Sects. 3–6 for the different types of VIs.

Table 2 The function $\omega(u,v,\tilde{u})$ in Sects. 3–6 for different types of VIs

Section 3	$\omega(u,v,\tilde{u}) = -1/2(v-\tilde{u})^T\beta H(v-\tilde{u})$	pointed in (3.6)
Section 4	$\omega(u,v,\tilde{u}) = -(v-\tilde{u})^T\beta M^T(v-\tilde{u})$	pointed in (4.7)
Section 5	$\omega(u,v,\tilde{u}) = -(v-\tilde{u})^T\beta(F(v)-F(\tilde{u}))$	pointed in (5.6)
Section 6	$\omega(u,v,\tilde{u}) = 0$	pointed in (6.4)

7 Effective quadruplet for Solodov-Svaiter's APPA

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). As in Sect. 6, the mapping F in this section is nonlinear and ∇F (if it exists) is asymmetric. In the APPA for (1.1) proposed by Solodov and Svaiter, see Algorithm 2 in [20], the triplet $(u,v,\tilde{u})$ in (1.7) is accepted when

$$\text{(Accepting rule)} \qquad \Delta(u,v,\tilde{u}) \le \mu\|u-v\|^2, \quad \mu\in(0,1), \tag{7.1a}$$

is satisfied, where

$$\Delta(u,v,\tilde{u}) = 2(v-\tilde{u})^T\{(v-u)+\beta F(v)\} - \|v-\tilde{u}\|^2. \tag{7.1b}$$

Under the accepting rule (7.1), the point $\tilde{u}^k$ was accepted as the new iterate u^{k+1} by Solodov and Svaiter (see [20], pp. 385). For the triplet $(u,v,\tilde{u})$ in (1.7) satisfying the accepting rule (7.1), we define the quadruplet $(d_1(u,v,\tilde{u}), d_2(u,v,\tilde{u}), \varphi(u,v,\tilde{u}), \phi(u,v,\tilde{u}))$ by setting

$$\text{(The quadruplet)} \qquad \begin{cases} d_1(u,v,\tilde{u}) = u-\tilde{u}, \\ d_2(u,v,\tilde{u}) = \beta F(v), \\ \varphi(u,v,\tilde{u}) = \|u-\tilde{u}\|^2 - (v-\tilde{u})^T\beta F(v), \\ \phi(u,v,\tilde{u}) = (1-\mu)\|u-v\|^2. \end{cases} \tag{7.2}$$

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

$$\tilde{u} = P_{\Omega}\{\tilde{u} - [\beta F(v) - (u-\tilde{u})]\}. \tag{7.3}$$

Using (7.2), the geminate directions ($d_1(u,v,\tilde{u})$ and $d_2(u,v,\tilde{u})$) satisfy the condition (2.1a).

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

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

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

$$(\tilde{u}-u^*)^T\beta F(v) \ge (\tilde{u}-v)^T\beta F(v). \tag{7.4}$$

Note that the left-hand-side of (7.4) is

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

while the right-hand-side of (7.4) is

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

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

Condition (2.1c): It follows from (7.1) and the definition of $\phi(u,v,\tilde{u})$ in (7.2) that

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

and consequently

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

An equivalent form of the above inequality is

$$2\|u-\tilde{u}\|^2-2(v-\tilde{u})^T\beta F(v)\ge\|u-\tilde{u}\|^2+\phi(u,v,\tilde{u}). \tag{7.5}$$

Note that the left-hand-side of (7.5) is $2\varphi(u,v,\tilde{u})$ (see (7.2)), while the right-hand-side of (7.5) is $\|d_1(u,v,\tilde{u})\|^2+\phi(u,v,\tilde{u})$. Therefore, the 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

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

and thus

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

From (7.1b) and the above inequality we obtain

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

Together with the accepting rule (7.1), we get

$$\|v-\tilde{u}\|\le\sqrt{\mu}\|u-v\|$$

and thus

$$\|u-\tilde{u}\|\le\|u-v\|+\|v-\tilde{u}\|\le(1+\sqrt{\mu})\|u-v\|. \tag{7.6}$$

Finally, from (7.6) and the definition of $\phi(u,v,\tilde{u})$ in (7.2), we obtain

$$\|u-\tilde{u}\|^2\le\left(\frac{1+\sqrt{\mu}}{1-\sqrt{\mu}}\right)\phi(u,v,\tilde{u})$$

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

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

Theorem 7.1 *For solving problem* (1.1), *let the triplet* $(u, v, \tilde{u})$ *be defined in* (1.7) *and the accepting rule* (7.1) *be satisfied. Then the quadruplet* $(d_1(u, v, \tilde{u}), d_2(u, v, \tilde{u}), \varphi(u, v, \tilde{u}), \phi(u, v, \tilde{u}))$ *defined in* (7.2) *is an effective quadruplet which fulfills conditions* (2.1).

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

Theorem 7.2 *Let the conditions in Theorem* 7.1 *be satisfied. Then the sequence* $\{u^k\}$ *generated by the primary methods converges to some* u^∞ *which is a solution point of* $\mathrm{VI}(\Omega, F)$.

Proof Due to Theorem 7.1 the quadruplet $(d_1(u, v, \tilde{u}), d_2(u, v, \tilde{u}), \varphi(u, v, \tilde{u}), \phi(u, v, \tilde{u}))$ in (7.2) satisfies the condition (2.1). We need only to verify the additional conditions (2.13). It follows from Propositions 2.1 and 2.2 that

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

and thus

$$\lim_{k\to\infty} \|u^k - v^k\| = 0. \tag{7.7}$$

From (7.6) we know that

$$\|u^k - \tilde{u}^k\| \le \|u^k - v^k\| + \|v^k - \tilde{u}^k\| \le (1+\sqrt{\mu})\|u^k - v^k\|. \tag{7.8}$$

Because $d_1(u^k, v^k, \tilde{u}^k) = u^k - \tilde{u}^k$ (see the form in (7.2)), it follows from (7.7) and (7.8) that

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

From (7.6), (7.7) 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 $\mathrm{VI}(\Omega, F)$. □

For the accepting rule (7.1) in APPAs for nonlinear VIs, the method proposed by Solodov and Svaiter, see Algorithm 2 in [20], adopts $\tilde{u}^k$ as the new iterate. Using $d_1(u^k, v^k, \tilde{u}^k) = u^k - \tilde{u}^k$ defined in (7.2), it can be rewritten as

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

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

Remark 7.1 As long as the parameter β_k is small enough, the accepting rules in the last four sections (see (3.8), (4.10), (5.8) 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

$$\tilde{u}^k = P_\Omega[u^k - \beta_k F(u^k)]. \tag{7.9}$$

However, the accepting rule (7.1) cannot be satisfied by setting $v^k = u^k$ for any $\beta_k > 0$. For $u \in \Omega \setminus \Omega^*$ and $\tilde{u} = P_\Omega[u - \beta F(u)] \neq u$, 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 (7.1b) we obtain

$$\Delta(u, u, \tilde{u}) \ge \|u - \tilde{u}\|^2 > 0.$$

Thus the accepting rule $\Delta(u, u, \tilde{u}) \le 0$ (see (7.1a)) can never be satisfied.

8 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:

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

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 sets, $f : \mathcal{X} \to R^n$ 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

$$u \in \Omega, \qquad \begin{cases} (x' - x)^T \{f(x) - A^T \lambda\} \ge 0, \\ (y' - y)^T \{g(y) + \lambda\} \ge 0, \\ Ax - y = 0, \end{cases} \quad \forall u' \in \Omega \tag{8.1}$$

where

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

The problem (8.1) is referred as a structured variational inequality (SVI for short) [12]. The compact form is

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

where

$$F(u) = \begin{pmatrix} f(x) - A^T\lambda \\ g(y) + \lambda \\ Ax - y \end{pmatrix}. \tag{8.2}$$

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:

$$\tilde{u} \in \Omega, \qquad \begin{pmatrix} x' - \tilde{x} \\ y' - \tilde{y} \\ \lambda' - \tilde{\lambda} \end{pmatrix}^T \left\{ \begin{pmatrix} f(\tilde{x}) - A^T\tilde{\lambda} \\ g(\tilde{y}) + \tilde{\lambda} \\ A\tilde{x} - \tilde{y} \end{pmatrix} + \begin{pmatrix} r(\tilde{x} - x) \\ s(\tilde{y} - y) \\ \beta^{-1}(\tilde{\lambda} - \lambda) \end{pmatrix} \right\} \geq 0, \quad \forall u' \in \Omega, \tag{8.3}$$

where $r, s, \beta^{-1} > 0$ are called the proximal coefficients. By using the notation of $F(u)$ (see (8.2)) and $\Omega = \mathcal{X} \times \mathcal{Y} \times R^m$ (in particular, $(A\tilde{x} - \tilde{y}) + \beta^{-1}(\tilde{\lambda} - \lambda) = 0$), (8.3) can be rewritten as

$$\tilde{x} \in \mathcal{X}, \qquad (x' - \tilde{x})^T \big\{ f(\tilde{x}) - A^T[\lambda - \beta(A\tilde{x} - \tilde{y})] + r(\tilde{x} - x) \big\} \geq 0, \quad \forall x' \in \mathcal{X}, \tag{8.4a}$$

$$\tilde{y} \in \mathcal{Y}, \qquad (y' - \tilde{y})^T \big\{ g(\tilde{y}) + [\lambda - \beta(A\tilde{x} - \tilde{y})] + s(\tilde{y} - y) \big\} \geq 0, \qquad \forall y' \in \mathcal{Y}, \tag{8.4b}$$

$$\tilde{\lambda} = \lambda - \beta(A\tilde{x} - \tilde{y}). \tag{8.4c}$$

The shortcoming of (8.4) is that subproblems (8.4a) and (8.4b) need to be solved jointly. In order to overcome this disadvantage, the proximal alternating directions method [11] 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

$$\begin{aligned} &\tilde{x}^k \in \mathcal{X}, \\ &(x' - \tilde{x}^k)^T \big\{ f(\tilde{x}^k) - A^T[\lambda^k - \beta(A\tilde{x}^k - y^k)] + r(\tilde{x}^k - x^k) \big\} \geq 0, \\ &\quad \forall x' \in \mathcal{X}. \end{aligned} \tag{8.5a}$$

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

$$\tilde{y}^k \in \mathcal{Y}, \qquad (y' - \tilde{y}^k)^T \big\{ g(\tilde{y}^k) + [\lambda^k - \beta(A\tilde{x}^k - \tilde{y}^k)] + s(\tilde{y}^k - y^k) \big\} \geq 0, \quad \forall y' \in \mathcal{Y}. \tag{8.5b}$$

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

$$\tilde{\lambda}^k = \lambda^k - \beta(A\tilde{x}^k - \tilde{y}^k). \tag{8.5c}$$

The only difference between (8.4) and (8.5) is to substitute $\tilde{y}^k$ (in (8.4a)) by y^k (in (8.5a)). Since (8.4) is a PPA updating form, the proximal alternating directions method (8.5) can be viewed as an approximate proximal point algorithm for the structured variational inequality (8.1). Note that the alternating directions method in [7, 8] is an extreme case of (8.5) which does not have the proximal terms $r(\tilde{x}^k - x^k)$ and $s(\tilde{y}^k - y^k)$.

By a simple manipulation, (8.5) can be written as: $\tilde{u} = (\tilde{x}, \tilde{y}, \tilde{\lambda}) \in \Omega$,

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

$\forall u' \in \Omega$.

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

$$\tilde{u} \in \Omega, \qquad (u' - \tilde{u})^T \{ \big(F(\tilde{u}) + \eta(u, \tilde{u})\big) + G(\tilde{u} - u) \} \geq 0, \quad \forall u' \in \Omega, \tag{8.6}$$

where

$$\eta(u, \tilde{u}) = \beta \begin{pmatrix} -A^T(y - \tilde{y}) \\ (y - \tilde{y}) \\ 0 \end{pmatrix} \tag{8.7}$$

and

$$G = \begin{pmatrix} rI_n & & \\ & (\beta + s)I_m & \\ & & \beta^{-1}I_m \end{pmatrix}. \tag{8.8}$$

To force G to be positive definite, r, β and $(\beta + s)$ should be positive. For convenience, however, we set $r,\ s$ and $\beta > 0$.

According to (1.2), since $\tilde{u}$ is a solution of (8.6), we have

$$\tilde{u} = P_\Omega\{\tilde{u} - [F(\tilde{u}) + \eta(u, \tilde{u}) + G(\tilde{u} - u)]\}. \tag{8.9}$$

Although v is degenerated, we can still regard $(F(\tilde{u}) + \eta(u, \tilde{u}) + G(\tilde{u} - u)) + (u - \tilde{u})$ as $F(v)$. Then, we can 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): $\quad (u, \tilde{u})$ satisfies (8.6) with $r,\ s$ and $\beta > 0$.

And, we define the effective quadruplet

$$\text{(The quadruplet)} \qquad \begin{cases} d_1(u, v, \tilde{u}) = u - \tilde{u}, \\ d_2(u, v, \tilde{u}) = F(\tilde{u}) + \eta(u, \tilde{u}), \\ \varphi(u, v, \tilde{u}) = \|u - \tilde{u}\|_G^2 - (\lambda - \tilde{\lambda})^T(y - \tilde{y}), \\ \phi(u, v, \tilde{u}) = \|u - \tilde{u}\|_G^2 - 2(\lambda - \tilde{\lambda})^T(y - \tilde{y}), \end{cases} \tag{8.10}$$

which involves 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 (8.8).

Condition (2.1a): Note that u and $\tilde{u}$ in sub-problem (8.6) can be written as

$$\tilde{u} = P_{\Omega}\{\tilde{u} - [(F(\tilde{u}) + \eta(u, \tilde{u})) - G(u - \tilde{u})]\}. \tag{8.11}$$

By combining (8.10) with (8.11), the geminate directions ($d_1(u, v, \tilde{u})$ and $d_2(u, v, \tilde{u})$) satisfy the condition (2.1a).

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

$$(\tilde{u} - u^*)^T F(\tilde{u}) \geq 0. \tag{8.12}$$

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 (8.10) and the above two inequalities that

$$(\tilde{u} - u^*)^T d_2(u, v, \tilde{u}) \geq (y - \tilde{y})^T (\tilde{\lambda} - \lambda) = \varphi(u, v, \tilde{u}) - \|d_1(u, v, \tilde{u})\|_G. \tag{8.13}$$

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

Condition (2.1c): For $d_1(u, v, \tilde{u})$ and $\varphi(u, v, \tilde{u})$ defined in (8.10), we have

$$2\varphi(u, v, \tilde{u}) - \|d_1(u, v, \tilde{u})\|_G^2 = \|u - \tilde{u}\|_G^2 - 2(\lambda - \tilde{\lambda})^T (y - \tilde{y}) = \phi(u, v, \tilde{u}). \tag{8.14}$$

Thus the condition (2.1c) follows from (8.14) 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 (8.8) and the definition of $\phi(u, v, \tilde{u})$ (8.10) that

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

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

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

Theorem 8.1 *For solving problem* (8.1), *let the duplet* $(u, \tilde{u})$ *be defined in* (8.6). *Then the quadruplet* $(d_1(u, v, \tilde{u}), d_2(u, v, \tilde{u}), \varphi(u, v, \tilde{u}), \phi(u, v, \tilde{u}))$ *given by* (8.10) *is an effective quadruplet which fulfills conditions* (2.1) *with* G *defined in* (8.8).

Theorem 8.2 *Let the conditions in Theorem* 8.1 *be satisfied. Then the sequence* $\{u^k\}$ *generated by the primary methods converges to some* u^{∞} *which is a solution point of problem* (8.1).

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

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

and thus

$$\lim_{k\to\infty} \|u^k - \tilde{u}^k\| = 0. \tag{8.15}$$

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 the 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 (8.7) and (8.15) that

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

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

9 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 different types of monotone VIs in Sects. 3–6, 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. By setting $v = u$, the important formulae are grouped in Tables 3 and 4.

Table 3 The mapping F and the geminate directions $d_1(u, u, \tilde{u})$ and $d_2(u, u, \tilde{u})$

Sections	$F(u)$	$d_1(u, u, \tilde{u})$ and which is given in		$d_2(u, u, \tilde{u})$ and which is given in	
Section 3	$Hu + q$	$(u - \tilde{u})$	(3.4)	$\beta(Hu + q)$	(3.5)
Section 4	$Mu + q$	$(u - \tilde{u}) + \beta M^T(u - \tilde{u})$	(4.4)	$\beta(Mu + q) + \beta M^T(u - \tilde{u})$	(4.5)
Section 5	$\nabla f(u)$	$(u - \tilde{u})$	(5.4)	$\beta F(u)$	(5.5)
Section 6	$F(u)$	$(u - \tilde{u}) - \beta(F(u) - F(\tilde{u}))$	(6.2)	$\beta F(\tilde{u})$	(6.3)

Table 4 The functions $\omega(u, u, \tilde{u})$ and $\varphi(u, u, \tilde{u})$

Section 3	$-\frac{1}{2}(u - \tilde{u})^T \beta H(u - \tilde{u})$	(3.6)	$\|u - \tilde{u}\|^2 - \frac{1}{2}(u - \tilde{u})^T \beta H(u - \tilde{u})$	(3.7)
Section 4	$-(u - \tilde{u})^T \beta M^T(u - \tilde{u})$	(4.7)	$\|u - \tilde{u}\|^2$	(4.8)
Section 5	$-(u - \tilde{u})^T \beta(F(u) - F(\tilde{u}))$	(5.6)	$\|u - \tilde{u}\|^2 - (u - \tilde{u})^T \beta(F(u) - F(\tilde{u}))$	(5.7)
Section 6	0	(6.4)	$\|u - \tilde{u}\|^2 - (u - \tilde{u})^T \beta(F(u) - F(\tilde{u}))$	(6.5)

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.

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 primary methods. At the same time, many new efficient methods (which can be viewed as the extension of the primary ones) could be also constructed. In Part II of this paper [13], we will show that those more efficient corresponding methods (called general or extended methods) can be constructed with only minor extra costs under the same effective quadruplet and accepting rule for each method. A set of numerical experiments are tested. From the numerical results, the efficiencies of the general methods are significant and convincing.

References

1. Bertsekas, D.P., Tsitsiklis, J.N.: Parallel and Distributed Computation, Numerical Methods. Prentice-Hall, Englewood Cliffs (1989)
2. Blum, E., Oettli, W.: Mathematische Optimierung. Grundlagen und Verfahren. Ökonometrie und Unternehmensforschung. Springer, Berlin (1975)
3. Dai, Y.H., Yuan, Y.: Alternate minimization gradient method. IMA J. Numer. Anal. **23**, 377–393 (2003)
4. Eckstein, J.: Nonlinear proximal point algorithms using Bregman functions, with applications to convex programing. Math. Oper. Res. **18**, 202–226 (1993)
5. Eckstein, J.: Approximate iterations in Bregman-function-based proximal algorithms. Math. Program. **83**, 113–123 (1998)
6. Facchinei, F., Pang, J.S.: Finite-Dimensional Variational Inequalities and Complementarity Problems. Springer Series in Operations Research, vol. I. Springer, New York (2003)
7. Glowinski, R.: Numerical Methods for Nonlinear Variational Problems. Springer, New York (1984)
8. Glowinski, R., Le Tallec, P.: Augmented Lagrangian and Operator Splitting Methods in Nonlinear Mechanics. SIAM Studies in Applied Mathematics. Philadelphia, PA (1989)
9. He, B.S.: Some predict-correct projection methods for monotone variational inequalities. Report 95-68, Faculty of Technical Mathematics and Informatics, Delf University of Technology (1995)
10. He, B.S., Liao, L.-Z.: Improvements of some projection methods for monotone nonlinear variational inequalities. J. Optim. Theory Appl. **112**, 111–128 (2002)
11. He, B.S., Liao, L.-Z., Han, D.R., Yang, H.: A new inexact alternating directions method for monotone variational inequalities. Math. Program. **92**, 103–118 (2002)
12. He, B.S., Liao, L.-Z., Qian, M.J.: Alternating projection based prediction-correction methods for structured variational inequalities. J. Comput. Math. **24**, 693–710 (2006)
13. He, B.S., Liao, L.-Z., Wang, X.: Proximal-like contraction methods for monotone variational inequalities in a unified framework II: General methods and numerical experiments. To appear in COAP
14. He, B.S., Xu, M.H.: A general framework of contraction methods for monotone variational inequalities. Pac. J. Optim. **4**, 195–212 (2008)
15. Khobotov, E.N.: Modification of the extragradient method for solving variational inequalities and certain optimization problems. USSR Comput. Math. Phys. **27**, 120–127 (1987)
16. Korpelevich, G.M.: The extragradient method for finding saddle points and other problems. Ekon. Mat. Metody **12**, 747–756 (1976)
17. Rockafellar, R.T.: Augmented Lagrangians and applications of the proximal point algorithm in convex programming. Math. Oper. Res. **1**, 97–116 (1976)
18. Rockafellar, R.T.: Monotone operators and the proximal point algorithm. SIAM J. Control Optim. **14**, 877–898 (1976)
19. Solodov, M.V., Svaiter, B.F.: A hybrid projection-proximal point algorithm. J. Convex Anal. **6**, 59–70 (1999)
20. Solodov, M.V., Svaiter, B.F.: Error bounds for proximal point subproblems and associated inexact proximal point algorithms. Math. Program. **88**, 371–389 (2000)
21. Tseng, P.: A modified forward-backword splitting method for maximal monotone mappings. SIAM J. Control Optim. **38**, 431–446 (2000)