RESEARCH   **Open Access**

CrossMark

A new criterion for an inexact parallel splitting augmented Lagrangian method

Xipeng Kou[*], Shengjie Li and Xueyong Wang

[*]Correspondence:
kouxipeng@126.com
College of Mathematics and
Statistics, Chongqing University,
Chongqing, 401331, China

Abstract

In this paper, we study the computational method for solving the variational inequality problem with the separable structure and linear constraints. We propose a new relaxed inexact criterion and a prediction-correction approach in the inexact splitting parallel augmented Lagrangian methods, which make it easier to solve the resulting subproblems. Under a mild condition, we prove the global convergence and establish a worst-case convergence rate for the new inexact algorithm. Some numerical experiments show the effectiveness and feasibility of the new inexact method.

Keywords: variational inequalities; proximal point method; decomposition splitting method; projection and contraction methods

1 Introduction

In this paper, we consider the following variational inequality problem (VI) with the separable structure:

$$\left(v - v^*\right)^T Q\left(v^*\right) \geq 0, \quad \forall v \in \Omega, \tag{1}$$

with

$$v = \begin{pmatrix} x \\ y \end{pmatrix}, \qquad Q(v) = \begin{pmatrix} f(x) \\ g(y) \end{pmatrix}, \quad \text{and} \tag{2}$$

$$\Omega := \left\{ v = (x, y) \mid Ax + By = b, x \in \mathcal{X}, y \in \mathcal{Y} \right\},$$

where $\mathcal{X} \subseteq \mathbb{R}^{n_1}$ and $\mathcal{Y} \subseteq \mathbb{R}^{n_2}$ are nonempty, closed, and convex sets; $A \in \mathbb{R}^{m \times n_1}$ and $B \in \mathbb{R}^{m \times n_2}$ are given matrices; $f : \mathcal{X} \to \mathbb{R}^{n_1}$ and $g : \mathcal{Y} \to \mathbb{R}^{n_2}$ are given monotone mappings; $b \in \mathbb{R}^m$ is a given vector and $n_1 + n_2 = n$.

The variational inequality problems with separable structures and linear constraints (1)-(2) have wide applications in some fields; see [1–7]. For solving the VIP, Glowinski and Marrocco [8] first proposed a Douglas-Rachford alternating direction method of multiplies (ADMM), which can decompose the original problems into subproblems with a smaller scale. The ADMM and its variants have been shown to be good efficient methods for many problems. However, the ADMM may fail since it is very difficult to solve the subproblems exactly in many practical applications. So, some strategies have been proposed

Springer

to overcome this drawback of the methods mentioned above, such as adding a proximal regularization term and transforming the nonlinear equation. For more details, one can refer to [9, 10].

In [11], He proposed a parallel splitting augmented Lagrangian method (PSALM) for solving the VI (1)-(2). The iterative process of the PSALM can be described to solve the following subproblems:

$$\widetilde{x} \in \mathcal{X}, \quad \langle x - \widetilde{x}, f(\widetilde{x}) - A^T[\lambda^k - H(A\widetilde{x} + By^k - b)] \rangle \geq 0, \quad \forall x \in \mathcal{X}, \tag{3}$$

$$\widetilde{y} \in \mathcal{Y}, \quad \langle y - \widetilde{y}, g(\widetilde{y}) - B^T[\lambda^k - H(Ax^k + B\widetilde{y} - b)] \rangle \geq 0, \quad \forall y \in \mathcal{Y}, \tag{4}$$

where $\lambda^k \in \mathbb{R}^m$ is the Lagrange multiplier associated with the constraint in (2) and $H \in \mathbb{R}^{m \times m}$ is a positive definite matrix which plays the role of the penalty parameter for the violation of the linear constraint in (2). The PSALM differs from other splitting methods, since the subproblems (3) and (4) can be computed in parallel. The advantage of parallel computation is easy implementation and attractive efficiency when the problems have a large scale. However, it is also very difficult to solve the subproblems (3) and (4) unless $f(x)$ and $g(y)$ have very particular structures. In [12], an inexact splitting parallel augmented Lagrangian method (IPSALM) was proposed to solve the subproblems (3) and (4) approximately so that their solutions satisfy a certain inexact criterion and are closed form ones. The prediction step is generated, for $\nu > 0$, by

$$\hat{x}^k = P_{\mathcal{X}} \left\{ x^k - \frac{1}{r_k}[f(x^k) - A^T(\lambda^k - Ax^k + By^k - b)] \right\}, \tag{5}$$

$$\hat{y}^k = P_{\mathcal{Y}} \left\{ y^k - \frac{1}{s_k}[g(y^k) - B^T(\lambda^k - Ax^k + By^k - b)] \right\}, \tag{6}$$

where r_k, s_k are chosen to satisfy the following conditions:

$$\|\xi_x^k + A^T HA(x^k - \hat{x}^k)\| \leq \nu r_k \|x^k - \hat{x}^k\| \quad \text{with } \xi_x^k := f(x^k) - f(\hat{x}^k), \tag{7}$$

$$\|\xi_y^k + B^T HB(y^k - \hat{y}^k)\| \leq \nu s_k \|y^k - \hat{y}^k\| \quad \text{with } \xi_y^k := g(y^k) - g(\hat{y}^k). \tag{8}$$

In [13], Zhang *et al.* proposed another inexact criterion for generating the prediction step, that is,

$$\langle x^k - \hat{x}^k, \xi_x^k \rangle + \|A(x^k - \hat{x}^k)\|_H^2 \leq \nu r_k \|x^k - \hat{x}^k\|^2, \tag{9}$$

$$\langle y^k - \hat{y}^k, \xi_y^k \rangle + \|B(y^k - \hat{y}^k)\|_H^2 \leq \nu s_k \|y^k - \hat{y}^k\|^2. \tag{10}$$

These inexact methods have the common feature that the subproblems or relevant problems are solved approximately at each iteration. Therefore, the effectiveness of the inexact methods depends greatly on the involved inexact criteria used to solve the subproblem.

Motivated and inspired by the inexact criteria in [12, 13], in this paper, we present a new inexact criterion (see (17) and (19)) to solve the subproblems under a very relaxed restriction. The new criterion improves the upper bound of $\langle x^k - \hat{x}^k, \xi_x^k \rangle + \|A(x^k - \hat{x}^k)\|_H^2$ and $\langle y^k - \hat{y}^k, \xi_y^k \rangle + \|B(y^k - \hat{y}^k)\|_H^2$. Thus, it reduces the computational load of the method considerably. Simultaneously, we also propose a prediction-correction approach in our algorithm analogous to [12–14]. Numerical applications to the multiple-sets split feasibility

problem (MSFP) and traffic equilibrium demonstrate that the proposed algorithm with the new criterion is very effective and feasible.

The rest of the paper is organized as follows. In Section 2, we summarize some concepts and properties, which are useful for further convergence analysis. In Sections 3 and 4, we present the inexact PSALM algorithm with the new criteria, and its global convergence and worst-case convergence rate, respectively. Preliminary numerical results of solving the multiple-sets split feasibility problem and the traffic equilibrium problem are presented in Section 5. Finally, we present a summary for our paper in Section 6.

2 Preliminaries

In this section, we summarize some basic properties and concepts, which will be used in the coming convergence analysis. Let G be a positive definite matrix. The G matrix norm of $v \in \mathbb{R}^n$ is defined by $\|v\|_G := \sqrt{v^T G v}$. In particular, $\|v\| := \sqrt{v^T v}$ is the Euclidean norm of $v \in \mathbb{R}^n$. $\langle \cdot, \cdot \rangle$ denotes the inner product in Euclidean spaces.

The following results are the well-known properties of the projection operator which will be used in the following analysis.

Lemma 2.1 *Let $\Omega \subset \mathbb{R}^n$ be nonempty, closed, and convex set, and let $P_\Omega[\cdot]$ be a projection operator onto the set Ω under the Euclidean norm. Then, for any $u, v \in \mathbb{R}^n$ and $w \in \Omega$, we have*

(1) $\langle u - P_\Omega[u], P_\Omega[u] - w \rangle \geq 0$;

(2) $\|P_\Omega[u] - P_\Omega[v]\|^2 \leq \langle u - v, P_\Omega[u] - P_\Omega[v] \rangle$;

(3) $\|P_\Omega[u] - w\|^2 \leq \|u - w\|^2 - \|u - P_\Omega[u]\|^2$.

Lemma 2.2 [1] *Let $\Omega \subset \mathbb{R}^n$ be a nonempty, closed, and convex set. Let $P_\Omega(\cdot)$ be the projection operator onto Ω under the Euclidean norm. Then u^* is a solution of VI (Ω, F) if and only if it satisfies*

$$u^* = P_\Omega\big[u^* - \beta F\big(u^*\big)\big], \quad \forall \beta > 0. \tag{11}$$

We recall the definition of monotone and strongly monotone mappings.

Definition 2.1 Let F be a mapping defined on the closed convex set $\Omega \subset \mathbb{R}^n$. Then

(a) F is called monotone on Ω if

$$\langle u - v, F(u) - F(v) \rangle \geq 0, \quad \forall u, v \in \Omega;$$

(b) F is called strongly monotone with the modulus $\mu > 0$ on Ω if

$$\langle u - v, F(u) - F(v) \rangle \geq \mu \|u - v\|^2, \quad \forall u, v \in \Omega;$$

(c) F is called Lipschitz continuous on Ω if there exists a constant $L > 0$ such that

$$\big\|F(u) - F(v)\big\| \leq L \|u - v\|, \quad \forall u, v \in \Omega.$$

In this paper, by attaching a Lagrange multiplier $\lambda \in \mathbb{R}^m$ to the linear constraint $Ax + By = b$, one obtains a compact form of the problem (1)-(2):

$$\langle w' - w, F(w) \rangle \geq 0, \quad \forall w' \in \mathcal{W}, \tag{12}$$

where

$$\mathcal{W} := \mathcal{X} \times \mathcal{Y} \times \mathcal{R}^m \quad \text{and} \quad F(w) := \begin{pmatrix} f(x) - A^T\lambda \\ g(y) - B^T\lambda \\ Ax + By - b \end{pmatrix}. \tag{13}$$

Note that the mapping F is monotone whenever f and g are monotone. In the sequel, the problem (12)-(13) will be denoted by MVI $(\mathcal{W}, F)$.

Remark 2.1 Because of attaching the Lagrange multiplier $\lambda \in \mathbb{R}^m$ to the linear constraints $Ax + By = b$, the VI (1)-(2) amounts to finding $(x, y, \lambda) \in \mathcal{X} \times \mathcal{Y} \times \mathbb{R}^m$ such that we have

$$\begin{cases} (x' - x)^T[f(x) - A^T\lambda] \geq 0, \\ (y' - y)^T[g(y) - B^T\lambda] \geq 0, \quad \forall (x', y') \in \mathcal{X} \times \mathcal{Y}. \\ Ax + By - b = 0, \end{cases}$$

So, the above formula is equivalent to the MVI $(\mathcal{W}, F)$ (12)-(13).

Throughout the paper, we make the following assumptions:
(A1) It has a closed form solution to compute the projection onto the convex sets $\mathcal{X}$ and $\mathcal{Y}$ under Euclidean norm.
(A2) The mappings $f(x)$ and $g(y)$ are Lipschitz continuous on $\mathcal{X}$ and $\mathcal{Y}$, respectively. However, the Lipschitz constants are not necessarily known.
(A3) The solution set $\mathcal{W}^*$ of the MVI $(\mathcal{W}, F)$ is nonempty.

3 The inexact PSALM with new inexact criterion for MVI

In this section, we propose the inexact method for solving MVI $(\mathcal{W}, F)$. To simplify our coming analysis, we denote some matrices

$$R_k = r_k I_{n_1}, \qquad S_k = s_k I_{n_2}, \quad \text{and} \quad G_k = \begin{pmatrix} R_k & & \\ & S_k & \\ -A & -B & H^{-1} \end{pmatrix}, \tag{14}$$

where $r_k > 0$ and $s_k > 0$.

Now, we mention our algorithm.

The inexact PSALM with new inexact criterion for MVI

Step 0 Given $\nu \in (0, 1)$, $\mu > 1$, $\gamma \in (0, 2)$, $\|A^THA\|/\nu \geq r_0 > 0$, $\|B^THB\|/\nu \geq s_0 > 0$, $\varepsilon > 0$. Let $H \in \mathbb{R}^{m \times m}$ be positive definite, $w^0 = (x^0, y^0, \lambda^0) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times \mathbb{R}^m$, and $k = 0$.

Step 1 Prediction step: For a given $w^k = (x^k, y^k, z^k)$, Generate the trial iterate $\hat{x}^k$, $\hat{y}^k$ via S1.2 and S1.3, simultaneously.

S1.1
$$\hat{\lambda}^k = \lambda^k - H(Ax^k + By^k - b). \tag{15}$$

S1.2 Find the smallest nonnegative integer i_k such that $r_k = \mu^{i_k} r_0$ and

$$\hat{x}^k = P_{\mathcal{X}}\left[x^k - \frac{1}{r_k}\left[f(x^k) - A^T \hat{\lambda}^k\right]\right] \tag{16}$$

which satisfies

$$\langle x^k - \hat{x}^k, \xi_x^k \rangle + \left\|A(x^k - \hat{x}^k)\right\|_H^2$$
$$\leq \nu\left(r_k \left\|x^k - \hat{x}^k\right\|^2 + \left\|Ax^k - A\hat{x}^k - \frac{1}{2}H^{-1}(\lambda^k - \hat{\lambda}^k)\right\|_H^2\right). \tag{17}$$

S1.3 Find the smallest nonnegative integer j_k such that $s_k = \mu^{j_k} s_0$ and

$$\hat{y}^k = P_{\mathcal{Y}}\left[y^k - \frac{1}{s_k}\left[g(y^k) - B^T \hat{\lambda}^k\right]\right] \tag{18}$$

which satisfies

$$\langle y^k - \hat{y}^k, \xi_y^k \rangle + \left\|B(y^k - \hat{y}^k)\right\|_H^2$$
$$\leq \nu\left(s_k \left\|y^k - \hat{y}^k\right\|^2 + \left\|By^k - B\hat{y}^k - \frac{1}{2}H^{-1}(\lambda^k - \hat{\lambda}^k)\right\|_H^2\right). \tag{19}$$

Step 2 Convergence verification: if $\|w^k - \hat{w}^k\| \leq \epsilon$, then stop. $\hat{w}^k = (\hat{x}^k, \hat{y}^k, \hat{\lambda}^k)$ is an acceptable approximate solution.

Step 3 Correction step: generate the new iterate w^{k+1} via

Form I:
$$w_{\mathbf{I}}^{k+1} = w^k - \gamma \alpha_k^* d(w^k, \hat{w}^k, \xi^k), \tag{20}$$

where

$$d(w^k, \hat{w}^k, \xi^k) = G_k(w^k - \hat{w}^k) - \xi^k \quad \text{with } \xi^k = \begin{pmatrix} \xi_x^k \\ \xi_y^k \\ 0 \end{pmatrix} \tag{21}$$

or

Form II:
$$w_{\mathbf{II}}^{k+1} = P_{\mathcal{W}}\left[w^k - \gamma \alpha_k^* F(\hat{w}^k)\right]. \tag{22}$$

Here, the set $\mathcal{W}$, the mapping $F(w)$ and G_k are defined in (13) and (14), respectively, the step size α_k is determined by

$$\alpha_k^* = \frac{\phi(w^k, \hat{w}^k, \xi^k)}{\|d(w^k, \hat{w}^k, \xi^k)\|^2} \tag{23}$$

and

$$\phi(w^k, \hat{w}^k, \xi^k) := \langle w^k - \hat{w}^k, d(w^k, \hat{w}^k, \xi^k) \rangle. \tag{24}$$

Set $k := k + 1$ and go to Step 1.

Remark 3.1 The prediction step of the proposed algorithm differs from both that of the method in [12] and that of the method in [13] in that we adopt a new criterion. Since $f(x)$ and $g(y)$ are Lipschitz continuous with constant L_f and L_g, respectively, we have $\|\xi_x^k\| \leq L_f\|x^k - \hat{x}^k\|$ and $\|\xi_y^k\| \leq L_g\|y^k - \hat{y}^k\|$. Using the Cauchy-Schwarz inequality, we have

$$\langle x^k - \hat{x}^k, \xi_x^k \rangle + \|A(x^k - \hat{x}^k)\|_H^2 \leq (L_f + \|A^T H A\|)\|x^k - \hat{x}^k\|^2.$$

Thus, the inequality (17) holds as long as

$$r_k \geq \frac{L_f + \|A^T H A\|}{\nu}.$$

Analogously, the inequality (19) holds when s_k satisfies

$$s_k \geq \frac{L_g + \|B^T H B\|}{\nu}.$$

Thus, in the implementation of the our algorithm, we choose, respectively, the values of r_k and s_k to satisfy the following conditions:

$$r_0 \leq r_k \leq r_{\max} := \frac{L_f + \|A^T H A\|}{\nu^2} \quad \text{and} \quad s_0 \leq s_k \leq s_{\max} =: \frac{L_g + \|B^T H B\|}{\nu^2}. \tag{25}$$

Remark 3.2 We note that $\hat{x}^k$ and $\hat{y}^k$ obtained by (16) and (18) are actually solutions of the following VIs, respectively:

$$\langle x - \hat{x}^k, f(x^k) - A^T \hat{\lambda}^k + R_k(\hat{x}^k - x^k) \rangle \geq 0, \quad \forall x \in \mathcal{X}, \tag{26}$$

$$\langle y - \hat{y}^k, g(y^k) - B^T \hat{\lambda}^k + S_k(\hat{y}^k - y^k) \rangle \geq 0, \quad \forall y \in \mathcal{Y}. \tag{27}$$

Combing (15)-(19) and (26)-(27), we have

$$\langle w - \hat{w}^k, F(\hat{w}^k) - d(w^k, \hat{w}^k, \xi^k) \rangle \geq 0, \quad \forall w \in \mathcal{W}. \tag{28}$$

4 Convergence
4.1 The global convergence
In this section, we establish the convergence of our algorithm.

Lemma 4.1 *Let the sequences $\hat{w}^k = (\hat{x}^k, \hat{y}^k, \hat{\lambda}^k)$ be generated by the prediction step from the given point $w^k = (x^k, y^k, \lambda^k)$. If $\|\hat{w}^k - w^k\| = 0$, then $\hat{w}^k$ is a solution of MVI $(\mathcal{W}, F)$.*

Proof Since $\|\hat{w}^k - w^k\| = 0$, it means that $\hat{x}^k = x^k$, $\hat{y}^k = y^k$, and $\hat{\lambda}^k = \lambda^k$. So, from (15) and (26)-(27), we obtain $A\hat{x}^k + B\hat{y}^k - b = 0$ and

$$\langle x' - \hat{x}^k, f(\hat{x}^k) - A^T \hat{\lambda}^k \rangle \geq 0, \quad \forall x' \in \mathcal{X},$$

$$\langle y' - \hat{y}^k, g(\hat{y}^k) - B^T \hat{\lambda}^k \rangle \geq 0, \quad \forall y' \in \mathcal{Y},$$

from which it follows that (12)-(13) hold for $\hat{w}^k$. The proof is completed. □

From Lemma 4.1, we terminate the algorithm if $\|w^k - \hat{w}^k\| \leq \epsilon$ holds for some k.

Lemma 4.2 *Let $\phi(w^k, \hat{w}^k)$ be the metrical function defined in* (24). *Then, for any $k \geq 1$, we have*

$$\phi(w^k, \hat{w}^k, \xi^k) \geq \frac{\delta(1-\nu)}{2}\big(\|w^k - \hat{w}^k\|^2 + \|A\hat{x}^k + B\hat{y}^k - b\|^2\big), \tag{29}$$

where $\delta := \min\{r_0, s_0, \lambda_m(H^{-1}), \lambda_m(H)\}$, and $\lambda_m(H^{-1})$ and $\lambda_m(H)$ denote the smallest eigenvalue of H^{-1} and H, respectively.

Proof By the definition of $\phi(w^k, \hat{w}^k, \xi^k)$, we get

$$\begin{aligned}
\phi(w^k, \hat{w}^k, \xi^k) &= \langle w^k - \hat{w}^k, d(w^k, \hat{w}^k, \xi^k)\rangle \\
&= \|x^k - \hat{x}^k\|_{R_k}^2 - \langle Ax^k - A\hat{x}^k, \xi_x^k\rangle + \|y^k - \hat{y}^k\|_{S_k}^2 - \langle By^k - B\hat{y}^k, \xi_y^k\rangle \\
&\quad - \langle Ax^k - A\hat{x}^k, \lambda^k - \hat{\lambda}^k\rangle - \langle By^k - B\hat{y}^k, \lambda^k - \hat{\lambda}^k\rangle + \|\lambda^k - \hat{\lambda}^k\|_{H^{-1}}^2 \\
&= \|x^k - \hat{x}^k\|_{R_k}^2 + \left\|Ax^k - A\hat{x}^k - \frac{1}{2}H^{-1}(\lambda^k - \hat{\lambda}^k)\right\|_H^2 \\
&\quad - \big(\langle Ax^k - A\hat{x}^k, \xi_x^k\rangle + \|Ax^k - A\hat{x}^k\|_H^2\big) \\
&\quad + \|y^k - \hat{y}^k\|_{S_k}^2 + \left\|By^k - B\hat{y}^k - \frac{1}{2}H^{-1}(\lambda^k - \hat{\lambda}^k)\right\|_H^2 \\
&\quad - \big(\langle By^k - B\hat{y}^k, \xi_y^k\rangle + \|By^k - B\hat{y}^k\|_H^2\big) + \frac{1}{2}\|\lambda^k - \hat{\lambda}^k\|_{H^{-1}}^2.
\end{aligned}$$

Using the inexact criteria (16)-(19), $2(\|a\|_H^2 + \|b\|_H^2) \geq \|a + b\|_H^2$, and the definition of δ, we have

$$\begin{aligned}
\phi(w^k, \hat{w}^k, \xi^k) &\geq (1-\nu)\left(\|x^k - \hat{x}^k\|_{R_k}^2 + \left\|Ax^k - A\hat{x}^k - \frac{1}{2}H^{-1}(\lambda^k - \hat{\lambda}^k)\right\|_H^2\right) \\
&\quad + (1-\nu)\left(\|y^k - \hat{y}^k\|_{S_k}^2 + \left\|By^k - B\hat{y}^k - \frac{1}{2}H^{-1}(\lambda^k - \hat{\lambda}^k)\right\|_H^2\right) \\
&\quad + \frac{1}{2}\|\lambda^k - \hat{\lambda}^k\|_{H^{-1}}^2 \\
&\geq (1-\nu)\big(\|x^k - \hat{x}^k\|_{R_k}^2 + \|y^k - \hat{y}^k\|_{S_k}^2\big) + \frac{1}{2}\|\lambda^k - \hat{\lambda}^k\|_{H^{-1}}^2 \\
&\quad + \frac{1-\nu}{2}\|Ax^k - A\hat{x}^k + By^k - B\hat{y}^k - H^{-1}(\lambda^k - \hat{\lambda}^k)\|_H^2 \\
&\geq \frac{1-\nu}{2}\big(\|x^k - \hat{x}^k\|_{R_k}^2 + \|y^k - \hat{y}^k\|_{S_k}^2 + \|\lambda^k - \hat{\lambda}^k\|_{H^{-1}}^2 + \|A\hat{x}^k + B\hat{y}^k - b\|_H^2\big) \\
&\geq \frac{\delta(1-\nu)}{2}\big(\|\hat{w}^k - w^k\|^2 + \|A\hat{x}^k + B\hat{y}^k - b\|^2\big).
\end{aligned}$$

The proof is completed. □

Lemma 4.3 *Let α_k^* be defined in* (23). *Then, for any $k \geq 1$, we get $\alpha_k^* \geq c > 0$.*

Proof By the analysis in Remark 3.1, we have

$$\|\xi_x^k\| \leq L_f\|x^k - \hat{x}^k\| \quad \text{and} \quad \|\xi_y^k\| \leq L_g\|y^k - \hat{y}^k\|.$$

Moreover, it follows from (25) that $\forall k \geq 0$

$$
\begin{aligned}
\left\| G_k\left(w^k - \hat{w}^k\right) \right\| &= r_k \left\| x^k - \hat{x}^k \right\| + s_k \left\| y^k - \hat{y}^k \right\| \\
&\quad + \left\| Ax^k - A\hat{x}^k + By^k - B\hat{y}^k - H^{-1}\left(\lambda^k - \hat{\lambda}^k\right) \right\| \\
&\leq r_{\max} \left\| x^k - \hat{x}^k \right\| + s_{\max} \left\| y^k - \hat{y}^k \right\| + \left\| A\hat{x}^k + B\hat{y}^k - b \right\|.
\end{aligned}
$$

Thus, there exists a positive constant C such that

$$
\begin{aligned}
\left\| d\left(w^k, \hat{w}^k\right) \right\| &= \left\| G_k\left(w^k - \hat{w}^k\right) - \xi^k \right\| \\
&\leq \left\| G_k\left(w^k, \hat{w}^k\right) \right\| + \left\| \xi^k \right\| \\
&\leq C\left(\left\| x^k - \hat{x}^k \right\| + \left\| y^k - \hat{y}^k \right\| + \left\| A\hat{x}^k + B\hat{y}^k - b \right\| \right) \\
&\leq C\left(\left\| w^k - \hat{w}^k \right\| + \left\| A\hat{x}^k + B\hat{y}^k - b \right\| \right),
\end{aligned}
\tag{30}
$$

where $C := \max\{r_{\max}, s_{\max}, L_f, L_g, 1\}$. So, according to (29) and (30) and the Cauchy-Schwarz inequality, we obtain

$$
\alpha_k^* = \frac{\phi(w^k, \hat{w}^k, \xi^k)}{\| d(w^k, \hat{w}^k, \xi^k) \|^2} \geq c := \frac{\delta(1 - \nu)}{4C^2} > 0. \qquad \square
$$

Before proving the convergence of our algorithm, we explain why we choose the inexact criterion and the step size α_k given in (23). Let $w^* = (x^*, y^*, \lambda^*)$ be a solution of MVI (12)-(13). In order to find the proper step size α_k for correct step, by $w_I^{k+1}(\alpha)$ and $w_{II}^{k+1}(\alpha)$ we denote the correction form **I** and **II** with the undetermined step size, respectively, *i.e.*

$$
w_I^{k+1}(\alpha) := w^k - \alpha d\left(w^k, \hat{w}^k, \xi^k\right)
\tag{31}
$$

and

$$
w_{II}^{k+1}(\alpha) := P_{\mathcal{W}}\left[w^k - \alpha F\left(\hat{w}^k\right) \right].
\tag{32}
$$

Moreover, we measure the improvement obtained by the correction step as follows:

$$
\psi(\alpha) = \left\| w^k - w^* \right\|^2 - \left\| w^{k+1}(\alpha) - w^* \right\|^2.
\tag{33}
$$

Naturally, we choose the α_k, which maximizes the function $\psi(\alpha)$, as the step size of the correct step.

Theorem 4.1 *Let $w^{k+1}(\alpha)$ be the correction step* (31) *or* (32) *with an undetermined step size, and $\psi(\alpha)$ be defined in* (33). *Then*

$$
\psi(\alpha) \geq 2\alpha \phi\left(w^k, \hat{w}^k, \xi^k\right) - \alpha^2 \left\| d\left(w^k, \hat{w}^k, \xi^k\right) \right\|^2,
\tag{34}
$$

where $d(w^k, \hat{w}^k, \xi^k)$ and $\phi(w^k, \hat{w}^k, \xi^k)$ are defined in (21) *and* (24), *respectively.*

Proof We break our proof into two cases for the correct form **I** and **II**, respectively.

(**I**) First, we prove the assertion for the first correction form **I**. Since

$$
\begin{aligned}
\psi(\alpha) &= \left\| w^k - w^* \right\|^2 - \left\| w^{k+1}(\alpha) - w^* \right\|^2 \\
&= \left\| w^k - w^* \right\|^2 - \left\| w^k - \alpha d\left(w^k, \hat{w}^k, \xi^k\right) - w^* \right\|^2 \\
&= 2\alpha \left\langle w^k - w^*, d\left(w^k, \hat{w}^k, \xi^k\right) \right\rangle - \alpha^2 \left\| d\left(w^k, \hat{w}^k, \xi^k\right) \right\|^2,
\end{aligned}
$$

we only need to prove the following inequality:

$$
\left\langle w^k - w^*, d\left(w^k, \hat{w}^k, \xi^k\right) \right\rangle \geq \phi\left(w^k, \hat{w}^k, \xi^k\right) = \left\langle w^k - \hat{w}^k, d\left(w^k, \hat{w}^k, \xi^k\right) \right\rangle.
$$

In fact, by using Lemma 2.2, we can change the inequality (28) into an equality:

$$
\hat{w}^k = P_{\mathcal{W}}\left[\hat{w}^k - \left[F(\hat{w}^k) - d\left(w^k, \hat{w}^k, \xi^k\right) \right] \right].
$$

Setting $u = \hat{w}^k - [F(\hat{w}^k) - d(w^k, \hat{w}^k, \xi^k)]$ and $w = w^*$ in Lemma 2.1(1), we get

$$
\left\langle \hat{w}^k - w^*, \hat{w}^k - \left[F(\hat{w}^k) - d\left(w^k, \hat{w}^k, \xi^k\right) \right] - \hat{w}^k \right\rangle \geq 0,
$$

that is,

$$
\left\langle \hat{w}^k - w^*, d\left(w^k, \hat{w}^k, \xi^k\right) \right\rangle \geq \left\langle \hat{w}^k - w^*, F(\hat{w}^k) \right\rangle. \tag{35}
$$

Note that $w^* \in \mathcal{W}^*$ and the mapping $F(w)$ is monotone on $\mathcal{W}$. So, by (35) and the definition of $\phi(w^k, \hat{w}^k, \xi^k)$, we obtain

$$
\begin{aligned}
\left\langle w^k - w^*, d\left(w^k, \hat{w}^k, \xi^k\right) \right\rangle &\geq \left\langle w^k - \hat{w}^k, d\left(w^k, \hat{w}^k, \xi^k\right) \right\rangle + \left\langle \hat{w}^k - w^*, F(\hat{w}^k) \right\rangle \\
&\geq \phi\left(w^k, \hat{w}^k, \xi^k\right),
\end{aligned}
$$

which indicates that (34) holds for the correction form **I**.

(**II**) Now, we prove the assertion for the second correction form **II**. From (28), we obtain

$$
\left\langle w - \hat{w}^k, F(\hat{w}^k) \right\rangle \geq \left\langle w - \hat{w}^k, d\left(w^k, \hat{w}^k, \xi^k\right) \right\rangle.
$$

Then it follows from $w^* \in \mathcal{W}$, the correction form **II** and Lemma 2.1(3) that

$$
\left\| w^{k+1}(\alpha) - w^* \right\|^2 \leq \left\| w^k - \alpha F(\hat{w}^k) - w^* \right\|^2 - \left\| w^k - \alpha F(\hat{w}^k) - w^{k+1}(\alpha) \right\|^2.
$$

Consequently, we have

$$
\begin{aligned}
\psi(\alpha) &= \left\| w^k - w^* \right\|^2 - \left\| w^{k+1}(\alpha) - w^* \right\|^2 \\
&\geq \left\| w^k - w^* \right\|^2 + \left\| w^k - \alpha F(\hat{w}^k) - w^{k+1}(\alpha) \right\|^2 - \left\| w^k - \alpha F(\hat{w}^k) - w^* \right\|^2 \\
&= \left\| w^k - w^{k+1}(\alpha) \right\|^2 + 2\alpha \left\langle w^{k+1}(\alpha) - w^*, F(\hat{w}^k) \right\rangle \\
&\geq \left\| w^k - w^{k+1}(\alpha) \right\|^2 + 2\alpha \left\langle w^{k+1}(\alpha) - \hat{w}^k, F(\hat{w}^k) \right\rangle
\end{aligned}
$$

$$\geq \left\| w^k - w^{k+1}(\alpha) \right\|^2 + 2\alpha \left\langle w^{k+1}(\alpha) - \hat{w}^k, d\left(w^k, \hat{w}^k, \xi^k\right) \right\rangle$$

$$= \left\| w^k - \alpha d\left(w^k, \hat{w}^k, \xi^k\right) - w^{k+1}(\alpha) \right\|^2 + 2\alpha \phi\left(w^k, \hat{w}^k, \xi^k\right) - \alpha^2 \left\| d\left(w^k, \hat{w}^k, \xi^k\right) \right\|^2,$$

which implies that (34) holds for the correction form **II**. The proof is completed. □

From Theorem 4.1, we may use the value of the maximum point α_k^* of the lower bound function $h(\alpha) = 2\alpha\phi(w^k, \hat{w}^k, \xi^k) - \alpha^2 \|d(w^k, \hat{w}^k, \xi^k)\|^2$ as the approximate solution of α. Since $h(\alpha)$ is a quadratic function, it reaches its maximum at

$$\alpha_k^* = \frac{\phi(w^k, \hat{w}^k, \xi^k)}{\|d(w^k, \hat{w}^k, \xi^k)\|^2}.$$

Furthermore, we introduce a relaxation factor γ to the step size, and then set $\alpha_k = \gamma \alpha_k^*$ in the correction step (20) or (22). We obtain

$$\psi(\alpha_k) = \psi(\gamma\alpha_k^*) = 2\gamma\alpha_k^*\phi(w^k, \hat{w}^k, \xi^k) - \gamma^2(\alpha_k^*)^2 \left\| d(w^k, \hat{w}^k, \xi^k) \right\|^2$$

$$= \gamma(2-\gamma)\alpha_k^*\phi(w^k, \hat{w}^k, \xi^k).$$

From Lemma 4.2, we have $\phi(w^k, \hat{w}^k, \xi^k) > 0$ if $w^k \neq \hat{w}^k$. Naturally, the step size $\alpha_k > 0$. So, the relaxation factor must satisfy $\gamma \in (0, 2)$ at each iteration. Based on the above analysis, we obtain the following corollary of Theorem 4.1.

Corollary 4.1 *Let $w^* \in \mathcal{W}^*$ and the sequence $\{w^k\}$ be generated by the proposed algorithm. Then the following inequality holds*:

$$\left\| w^{k+1} - w^* \right\|^2 \leq \left\| w^k - w^* \right\|^2 - \gamma(2-\gamma)\alpha_k^*\phi\left(w^k, \hat{w}^k, \xi^k\right)$$

$$\leq \left\| w^k - w^* \right\|^2 - \frac{c\delta\gamma(2-\gamma)(1-\nu)}{2} \left\| w^k - \hat{w}^k \right\|^2.$$

From Corollary 4.1, we note that $\{w^k\}$ is Fejér monotone. Using a similar proof procedure to [15] and [12], we can easily derive the following result.

Theorem 4.2 *The sequence $\{w^k\}$ as generated by the proposed algorithm converges to a solution of MVI $(\mathcal{W}, F)$.*

4.2 Convergence rate
Now, we show the worst-case $O(1/t)$ convergence rate for the proposed algorithm.

Lemma 4.4 *For the given $w^k \in \mathcal{W}$, let $\hat{w}^k$ be generated by the proposed algorithm and the new iterate is updated by correction form **I** or **II** with $\gamma > 0$. Then*

$$\left(w - \hat{w}^k\right)^T \gamma\alpha_k^* F\left(\hat{w}^k\right) + \frac{1}{2}\left(\left\| w - w^k \right\|^2 - \left\| w - w^{k+1} \right\|^2\right)$$

$$\geq \frac{1}{2}\gamma(2-\gamma)\phi\left(w^k, \hat{w}^k, \xi^k\right), \quad \forall w \in \mathcal{W}, \tag{36}$$

where $\phi(w^k, \hat{w}^k, \xi^k)$ is defined in (24).

Proof We divide our proof into two parts for correction form **I** and **II**, respectively.

(**I**) Due to (28)

$$\left(w - \hat{w}^k\right)^T F\left(\hat{w}^k\right) \geq \left(w - \hat{w}^k\right)^T d\left(w^k, \hat{w}^k, \xi^k\right), \quad \forall w \in \mathcal{W}, \tag{37}$$

and the correction form **I**

$$w^k - w^{k+1} = \gamma \alpha_k^* d\left(w^k, \hat{w}^k, \xi^k\right).$$

Using the identity

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

we have

$$\begin{aligned}
\left(w - \hat{w}^k\right)^T \gamma \alpha_k^* F\left(\hat{w}^k\right) &\geq \left(w - \hat{w}^k\right)^T \left(w^k - w^{k+1}\right) \\
&= \frac{1}{2}\left(\left\|w - w^{k+1}\right\|^2 - \left\|w - w^k\right\|^2\right) \\
&\quad + \frac{1}{2}\left(\left\|w^k - \hat{w}^k\right\|^2 - \left\|w^{k+1} - \hat{w}^k\right\|^2\right).
\end{aligned} \tag{38}$$

By the correction form **I** (20) and the definition of $\phi(w^k, \hat{w}^k, \xi^k)$, we obtain

$$\begin{aligned}
&\left\|w^k - \hat{w}^k\right\|^2 - \left\|w^{k+1} - \hat{w}^k\right\|^2 \\
&= \left\|w^k - \hat{w}^k\right\|^2 - \left\|w^k - \hat{w}^k - \gamma \alpha_k^* d\left(w^k, \hat{w}^k, \xi^k\right)\right\|^2 \\
&= 2\gamma \alpha_k^* \left(w^k - \hat{w}^k\right)^T d\left(w^k, \hat{w}^k, \xi^k\right) - \gamma^2 \left(\alpha_k^*\right)^2 \left\|d\left(w^k, \hat{w}^k, \xi^k\right)\right\|^2 \\
&= \gamma(2 - \gamma)\left(\alpha_k^*\right)^2 \left\|d\left(w^k, \hat{w}^k, \xi^k\right)\right\|^2 \\
&= \gamma(2 - \gamma)\phi\left(w^k, \hat{w}^k, \xi^k\right).
\end{aligned} \tag{39}$$

It follows from (38) and (39) that (36) holds for the correction form **I**.

(**II**) For the correction form **II**, we divide firstly $(w - \hat{w}^k)^T \gamma \alpha_k^* F(\hat{w}^k)$ in the terms

$$\left(w^{k+1} - \hat{w}^k\right)^T \gamma \alpha_k^* F\left(\hat{w}^k\right) \quad \text{and} \quad \left(w - w^{k+1}\right)^T \gamma \alpha_k^* F\left(\hat{w}^k\right). \tag{40}$$

First, we deal with the term $(w^{k+1} - \hat{w}^k)^T \gamma \alpha_k^* F(\hat{w}^k)$. Since the new iterate $w^{k+1} \in \mathcal{W}$, substituting $w = w^{k+1}$ into (37), we have

$$\begin{aligned}
\left(w^{k+1} - \hat{w}^k\right)^T \gamma \alpha_k^* F\left(\hat{w}^k\right) &\geq \gamma \alpha_k^* \left(w^{k+1} - \hat{w}^k\right)^T d\left(w^k, \hat{w}^k, \xi^k\right) \\
&= \gamma \alpha_k^* \left(w^k - \hat{w}^k\right)^T d\left(w^k, \hat{w}^k, \xi^k\right) \\
&\quad - \gamma \alpha_k^* \left(w^k - w^{k+1}\right)^T d\left(w^k, \hat{w}^k, \xi^k\right).
\end{aligned} \tag{41}$$

Using the definition of step size α, we get

$$\gamma \alpha_k^* \left(w^k - \hat{w}^k\right)^T d\left(w^k, \hat{w}^k, \xi^k\right) = \gamma \left(\alpha_k^*\right)^2 \left\|d\left(w^k, \hat{w}^k, \xi^k\right)\right\|^2.$$

Naturally, we have

$$-\gamma \alpha_k^* (w^k - w^{k+1})^T d(w^k, \hat{w}^k, \xi^k) = -\frac{1}{2} \left(\|w^k - w^{k+1}\|^2 + \gamma^2 (\alpha_k^*)^2 \|d(w^k, \hat{w}^k, \xi^k)\|^2 \right)$$
$$+ \frac{1}{2} \|w^k - w^{k+1} - \gamma \alpha_k^* d(w^k, \hat{w}^k, \xi^k)\|^2.$$

So, we get

$$(w^{k+1} - \hat{w}^k)^T \gamma \alpha_k^* F(\hat{w}^k) \geq \frac{1}{2} \gamma (2 - \gamma)(\alpha_k^*)^2 \|d(w^k, \hat{w}^k, \xi^k)\|^2 - \frac{1}{2} \|w^k - w^{k+1}\|^2$$
$$= \frac{1}{2} \gamma (2 - \gamma)\phi(w^k, \hat{w}^k, \xi^k) - \frac{1}{2} \|w^k - w^{k+1}\|^2. \tag{42}$$

Now, we turn to treat the second term $(w - w^{k+1})^T \gamma \alpha_k^* F(\hat{w}^k)$ in (40). Since w^{k+1} is corrected by the correction form **II** (22), w^{k+1} is the projection of $w^k - \gamma \alpha_k^* F(\hat{w}^k)$ on $\mathcal{W}$, it follows from Lemma 2.1(1) that

$$(w - w^{k+1})^T (w^{k+1} - w^k + \gamma \alpha_k^* F(\hat{w}^k)) \geq 0, \quad \forall w \in \mathcal{W},$$

and consequently, using the identity $a^T b = \frac{1}{2}(\|a\|^2 + \|b\|^2 - \|a - b\|^2)$, we obtain

$$(w - w^{k+1})^T \gamma \alpha_k^* F(\hat{w}^k) \geq (w - w^{k+1})^T (w^k - w^{k+1})$$
$$= \frac{1}{2} (\|w - w^{k+1}\|^2 - \|w - w^k\|^2) + \frac{1}{2} \|w^k - w^{k+1}\|^2. \tag{43}$$

Adding (42) and (43), we get (36). The proof is completed. □

Theorem 4.3 *For an integer $t > 0$, there is a $\hat{w}_t \in \mathcal{W}$, which is a convex combination of the prediction iterates $\hat{w}^0, \hat{w}^1, \ldots, \hat{w}^n$, satisfying*

$$(\hat{w}_t - w)^T F(w) \leq \frac{1}{2\gamma \Gamma_t} \|w^0 - w\|^2, \quad \forall w \in \mathcal{W},$$

where

$$\Gamma_t := \sum_{k=0}^{t} \alpha_k^* \quad and \quad \hat{w}_t := \frac{1}{\Gamma_t} \sum_{k=0}^{t} \alpha_k^* \hat{w}^k.$$

Proof It follows from (29) and (36) that

$$\langle w - \hat{w}^k, \alpha_k^* F(\hat{w}^k) \rangle + \frac{1}{2\gamma} \|w - w^k\|^2 \geq \frac{1}{2\gamma} \|w - w^{k+1}\|^2, \quad \forall w \in \mathcal{W}.$$

By the monotonicity of $F(w)$ and the above inequality, we have

$$\langle w - \hat{w}^k, \alpha_k^* F(w) \rangle + \frac{1}{2\gamma} \|w - w^k\|^2 \geq \frac{1}{2\gamma} \|w - w^{k+1}\|^2, \quad \forall w \in \mathcal{W}. \tag{44}$$

Summing the inequality (44) over $k = 0, 1, \ldots, t$, we get

$$\left[\left(\sum_{k=0}^{t} \alpha_k^* \right) w - \left(\sum_{k=0}^{t} \alpha_k^* \hat{w}^k \right) \right]^T F(w) + \frac{1}{2\gamma} \|w - w^0\|^2 \geq \frac{1}{2\gamma} \|w - w^{k+1}\|^2 \geq 0, \quad \forall w \in \mathcal{W}.$$

Since $\sum_{k=0}^{t} \alpha_k^* / \Gamma_t = 1$, $\hat{w}_t$ is a convex combination of $\hat{w}^0, \hat{w}^1, \dots, \hat{w}^t$ and thus $\hat{w}_t \in \mathcal{W}$. By the definitions of Γ_t and $\hat{w}_t$, we derive

$$(w - \hat{w}_t)^T F(w) + \frac{1}{2\gamma \Gamma_t} \| w - w^0 \|^2 \geq 0.$$

The assertion follows from the above inequality immediately.　　　　　　　　□

So, we have gotten a worst-case $O(1/t)$ convergence rate of the proposed algorithm in the ergodic sense.

5 Numerical results

In this section, we illustrate the effectiveness of our algorithm by comparing with some existing algorithms. We denote the proposed method in this paper as 'NEW', the alternating projection-based prediction-correction method in [15] as 'HLQ', the inexact parallel splitting augmented Lagrangian method in [12] as 'TY' and the simultaneous method in [13] as 'ZHY'. All codes are written in Matlab and run on i3-2100 CPU 3.1GHZ and 3GB memory.

5.1 Multiple-sets split feasibility problem

The multiple-sets split feasibility problem (MSFP) is to find a point in the intersection of a family of closed convex sets in one space such that its image under a certain operator is in the intersection of another family of a closed convex sets in image space. The MSFP plays a significant role in diversified areas, such as image restoration, signal processing, and medical care; see [16–18]. In the paper, we consider the constrained MSFP in the following form:

$$x^* \in \mathcal{X} \cap \left(\bigcap_{i=1}^{t_1} C_i \right) \quad \text{and} \quad Ax^* \in \mathcal{Y} \cap \left(\bigcap_{j=1}^{t_2} Q_j \right). \tag{45}$$

Censor *et al.* [17] proposed the proximity function to measure the aggregate distance to the involved sets C_i's and Q_j's as

$$p(x) := \frac{1}{2} \sum_{i=1}^{t_1} a_i \| x - P_{C_i}(x) \|^2 + \frac{1}{2} \sum_{j=1}^{t_2} b_j \| Ax - P_{Q_j}(Ax) \|^2, \tag{46}$$

where $a_i > 0$ ($i = 1, \dots, t_1$) and $b_j > 0$ ($j = 1, \dots, t_2$) are coefficients which can be defined as weights of important attached to sets. Note that the condition $\sum_{i=1}^{t_1} a_i + \sum_{j=1}^{t_2} b_j = 1$ is usually assumed in practice. With the proximity function (46), Censor *et al.* [17] proposed the optimization model

$$\min \{ p(x) \mid x \in \mathcal{X} \} \tag{47}$$

to approximate the constrained MSFP and used the following projection gradient method to solve this model:

$$x^{k+1} = P_{\mathcal{X}} \left[x^k - s\nabla P(x^k) \right],$$

where the gradient of $p(x)$ is as follows:

$$\nabla p(x) = \sum_{i=1}^{t_1} a_i \big(x - P_{C_i}(x)\big) + \sum_{j=1}^{t_2} b_i A^T \big(Ax - P_{Q_j}(Ax)\big),$$

where P_{C_i} and P_{Q_j} is projection mapping onto set C_i and Q_j, respectively. Zhang *et al.* changed (45) into the following optimization model with separable structure:

$$\min\big\{\theta_1(x) + \theta_2(y) \mid Ax - y = 0, x \in \mathcal{X}, y \in \mathcal{Y}\big\}, \tag{48}$$

where

$$\theta_1(x) =: \frac{1}{2} \sum_{i=1}^{t_1} a_i \big\|x - P_{C_i}(x)\big\|^2 \quad \text{and} \quad \theta_2(y) =: \frac{1}{2} \sum_{j=1}^{t_2} b_j \big\|y - P_{Q_j}(y)\big\|^2.$$

According to the first order optimality condition, the constrained MSFP is equivalent to finding $w^* = (x^*, y^*, \lambda^*) \in \mathcal{W} := \mathcal{X} \times \mathcal{Y} \times \mathbb{R}^m$ such that, for all $\tilde{w} = (\tilde{x}, \tilde{y}, \tilde{\lambda}) \in \mathcal{W}$,

$$\begin{cases} \langle \tilde{x} - x^*, f(x^*) - A^T \lambda^* \rangle \geq 0, \\ \langle \tilde{y} - y^*, g(y^*) - \lambda^* \rangle \geq 0, \\ \langle \tilde{\lambda} - \lambda^*, Ax^* - y^* \rangle \geq 0, \end{cases} \tag{49}$$

where

$$f(x) := \nabla\theta_1(x) = \sum_{i=1}^{t_1} a_i \big(x - P_{C_i}(x)\big) \quad \text{and} \quad g(y) := \nabla\theta_2(x) = \sum_{j=1}^{t_2} b_j \big(y - P_{Q_j}(y)\big).$$

From [13], Lemma 3, $f(x)$ and $g(y)$ are Lipschitz continuous on $\mathcal{X}$ and $\mathcal{Y}$, the Lipschitz constant is $L_1 = \sum_{i=1}^{t_1} a_i$ and $L_2 = \sum_{j=1}^{t_2} b_j$, respectively.

Now, we consider the special MSFP problem (45) tested in [13] with the sets

$$C_i = \big\{x \in \mathbb{R}^n \mid \|x - D_i\| \leq R_i\big\}, \quad i = 1, \ldots, t_1,$$

$$Q_j = \{y \in \mathbb{R}^n \mid L_j \leq y \leq U_j\}, \quad j = 1, \ldots, t_2,$$

where $D_i \in \mathbb{R}^n$ is the center of the ball C_i and is randomly generated in $(0, 10)$; $R_i \in \mathbb{R}^n$ is the radius of the ball C_i and is randomly generated in $(40, 50)$; L_j and U_j are lower and upper bounds of the box set Q_j and are randomly generated in $(10, 30)$ and $(40, 80)$, respectively. The components of the linear operator $A \in \mathbb{R}^{n \times n}$ are generated randomly with eigenvalues in $(10, 20)$. The constraints $\mathcal{X}$ and $\mathcal{Y}$ in (45) are the sets $\mathbb{R}^n_+$ and $\mathbb{R}^n$, respectively.

We set also the same initial iterate $x^0 = \mathbf{0}$ and $y^0 = \lambda^0 = \mathbf{1}$ for HLQ, TY, ZHY, and NEW where $\mathbf{0}$ and $\mathbf{1}$ are vectors whose elements are all $\mathbf{0}$ and $\mathbf{1}$, respectively. The choice of the involved parameters for the tested methods is as follows: $\nu = 0.95$, $H = \beta I$ with $\beta = 0.0002$, $\gamma = 1.2$, $\mu = 1.8$, $r_0 = s_0 = 1$, and the stopping criterion $\|\hat{w}^k - w^k\| \leq \epsilon$ for HLQ, TY, ZHY, and NEW.

We report the numerical performance of various methods for the MSFP problem with different scenarios of t_1 and t_2 in Table 1. The data in Table 1 exhibits the effectiveness of the proposed method and its superiority to HLQ, TY, and ZHY when the dimensionality of the MSFP or the number of set components is large.

Table 1 Numerical performance for MSFP

	Method		n								
			20	**30**	**40**	**50**	**60**	**70**	**80**	**90**	**100**
$t_1 = 500$	HLQ	Iter.	10	22	13	28	18	31	26	23	30
$t_2 = 500$		Tim.	0.06	0.11	0.07	0.16	0.13	0.24	0.16	0.16	0.18
	TY	Iter.	30	31	28	26	27	28	27	26	27
		Tim.	0.15	0.17	0.17	0.24	0.22	0.24	0.21	0.20	0.18
	ZHY	Iter.	26	28	30	26	29	24	25	27	20
		Tim.	0.11	0.23	0.25	0.33	0.16	0.31	0.18	0.27	0.27
	NEW	Iter.	11	17	21	24	18	21	21	19	19
		Tim.	0.05	0.10	0.10	0.14	0.10	0.17	0.14	0.13	0.11
$t_1 = 1{,}000$	HLQ	Iter.	34	35	33	24	30	33	25	29	25
$t_2 = 1{,}000$		Tim.	0.33	0.41	0.36	0.25	0.32	0.40	0.33	0.33	0.43
	TY	Iter.	32	33	31	24	28	30	24	24	28
		Tim.	0.33	0.40	0.35	0.24	0.32	0.37	0.30	0.33	0.41
	ZHY	Iter.	30	33	31	22	25	16	22	22	31
		Tim.	0.34	0.42	0.35	0.27	0.36	0.27	0.31	0.33	0.44
	NEW	Iter.	26	10	22	15	20	12	18	22	21
		Tim.	0.22	0.11	0.21	0.15	0.21	0.14	0.19	0.27	0.25
$t_1 = 5{,}000$	HLQ	Iter.	29	32	33	27	28	27	28	29	25
$t_2 = 5{,}000$		Tim.	1.55	1.53	1.66	1.40	1.80	1.64	1.6	1.82	1.6
	TY	Iter.	27	28	31	28	27	28	27	28	21
		Tim.	0.97	0.96	0.99	0.89	1.03	1.09	1.01	1.02	1.04
	ZHY	Iter.	24	27	30	16	33	33	22	23	21
		Tim.	0.88	0.85	0.94	0.58	1.3	1.12	0.93	0.85	1.17
	NEW	Iter.	11	12	22	16	18	22	18	18	18
		Tim.	0.84	0.36	0.66	0.49	0.55	0.77	0.72	0.65	0.65
$t_1 = 10{,}000$	HLQ	Iter.	36	36	34	26	30	33	26	23	33
$t_2 = 10{,}000$		Tim.	3.30	3.50	3.20	2.60	2.23	3.24	3.28	2.83	3.10
	TY	Iter.	29	30	32	25	32	31	27	26	24
		Tim.	1.52	1.82	2.1	1.61	2.01	2.09	1.77	1.97	1.81
	ZHY	Iter.	23	30	33	25	32	30	26	25	22
		Tim.	1.35	1.90	2.20	1.55	2.21	2.11	1.86	1.94	1.66
	NEW	Iter.	28	29	19	23	23	14	21	23	22
		Tim.	1.50	1.70	1.30	1.34	1.38	1.02	1.34	1.48	1.53

5.2 Traffic equilibrium problems

In this subsection, we apply the proposed method to solve the traffic equilibrium problems with link capacity bounds, which have been well studied in the literature of transportation. Since both the travel cost and the travel disutility are functions of the path flow $\mathbf{x}$, the traffic network equilibrium problem with link capacity is to seek the path flow $\mathbf{x}^*$:

$$\langle \mathbf{x} - \mathbf{x}^*, f(\mathbf{x}^*) \rangle \geq 0, \quad \forall \mathbf{x} \in \Pi, \tag{50}$$

with

$$\Pi = \left\{ \mathbf{x} \in \mathbb{R}^n : A^T \mathbf{x} \leq b \right\}, \tag{51}$$

where $\mathbf{x} \in \mathbb{R}^n$ represents the traffic flow on paths, b is the vector indicating the capacities on the links, $A \in \mathbb{R}^{n \times m}$ is the path-link indicating matrix, and f is the vector indicating the traffic flows on the links. By introducing the slack variable $\mathbf{y} \geq 0$, the traffic equilibrium problem (50) is equivalent to

$$\langle \mathbf{x} - \mathbf{x}^*, f(\mathbf{x}^*) \rangle \geq 0, \quad \forall \mathbf{x} \in \Omega, \tag{52}$$

Table 2 Numerical performance for traffic equilibrium problems

Level	Method	No. of iter.			No. of f eval. s			CPU		
		HLQ	TY	NEW	HLQ	TY	NEW	HLQ	TY	NEW
$b = 30$	Ex 7.4	291	509	305	624	1,046	859	0.119	0.094	0.033
	Ex 7.5	234	664	239	492	1,358	550	0.093	0.188	0.030
$b = 40$	Ex 7.4	387	397	303	1,146	1,118	808	0.172	0.188	0.031
	Ex 7.5	294	512	278	653	1,048	607	0.141	0.188	0.031

Table 3 The optimal link flow and toll for Example 7.4 in [7] with $b = 40$

Link	Flow	Charge	Link	Flow	Charge	Link	Flow	Charge
1	0	0	11	1.1631	0			
2	2.6941	0	12	33.8285	0	21	1.1631	0
3	40	−27.078	13	25.8473	0	22	34.9916	0
4	14.1527	0	14	1.987	0	23	1.8702	0
5	12.1658	0	15	36.1428	0	24	3.85072	0
6	40	−134.009	16	6.9776	0	25	40	−118.4
7	33.0224	0	17	1.0895	0	26	30.9547	0
8	31.9328	0	18	31.9328	0	27	32.0442	0
9	0	0	19	0	0	28	0	0
10	0	0	20	0	0			

Table 4 The optimal link flow and toll for Example 7.5 in [7] with $b = 40$

Link	Flow	Charge	Link	Flow	Charge	Link	Flow	Charge
1	40	−34.1851	14	0	0			
2	40	0	15	0	0	27	0	0
3	40	−70.1851	16	0	0	28	0	0
4	0	0	17	0	0	29	32.4413	0
5	0	0	18	0	0	30	0	0
6	0	0	19	0	0	31	7.5587	0
7	0	0	20	40	−9.7608	32	40	−134.445
8	0	0	21	40	−12.9483	33	40	−51.6542
9	0	0	22	40	−124.363	34	32.4413	0
10	40	−13.3271	23	40	−51.6045	35	32.4413	0
11	0	0	24	0	0	36	40	−275.088
12	7.5587	0	25	0	0	37	0	0
13	40	−70.6569	26	0	0			

with

$$\Omega = \left\{ \mathbf{x} \in \mathbb{R}^n : A^T \mathbf{x} + \mathbf{y} = b, \mathbf{x} \geq 0, \mathbf{y} \geq 0 \right\}, \tag{53}$$

which is a special case of VIP with $g(\mathbf{y}) \equiv 0$, $B = I$, $\mathcal{X} = \mathbb{R}^n_+$, and $\mathcal{Y} = \mathbb{R}^m_+$. We test particularly Examples 7.4 and 7.5 in [7] to compare the new method with HLQ [15] and TY [12]. For Example 7.4, $n = 49$, $m = 28$, and $A \in \mathbb{R}^{49 \times 28}$. For Example 7.5, $n = 55$, $m = 37$, and $A \in \mathbb{R}^{55 \times 37}$. As recommended in [12], we choose the parameters in HLQ and TY as $\nu = 0.95$, $H = \beta I$ with $\beta = 1.1$, $\gamma = 1.85$, $\mu = 1.25$, $r_0 = 1$, $s_0 = 1.1$. In our algorithm, we choose the parameter $r_0 = s_0 = 1.25$. The different parameters were selected to make the related methods to achieve better results. Simultaneously, we choose the same initial iterative points $x^0 = \mathbf{1}$, $y^0 = \lambda^0 = \mathbf{0}$, and the stopping criterion $\|\hat{w}^k - w^k\| \leq \epsilon$ in all methods.

For Examples 7.4 and 7.5, the numerical results of HLQ, TM, and NEW are reported in Table 2. By No. of iter., No. of f eval. and CPU we denote the number of iterations, the number of function evaluations, and the CPU time in seconds, respectively.

From Table 2, we can see that the proposed method with the new criterion is comparable to HLQ and superior to TY when the capacity b = 30. For the capacity b = 40, Table 2 shows that our method is more effective than HLQ and TY in terms of the number of iterations, number of function evaluations, and the CPU time.

Moreover, we report the optimal link flow generated by the proposed method. As illustrated in [7], the absolute value of the Lagrange multiplier λ^* actually means the toll that should be charged on the links to avoid congestion. The numerical results of Examples 7.4 and 7.5 [7] with the capacity b = 40 are reported in Tables 3 and 4, respectively. We can see that no toll is charged on the links whose flows are lower than their capacities.

6 Conclusions

In this paper, we study an inexact criterion for solving the convex problems and variational inequalities with separable structures. Based on the prediction-correction approach, two correction forms were derived. Preliminary numerical results with MSFP and traffic equilibrium problems indicate that our method is efficient in practice. In addition, the reported numerical experiments for MSFP are confined to artificial randomly generated data. In the future, we shall test the real world split inversion problems with the inexact criterion.

References
 1. Bertsekas, DP, Tsitsiklis, JN: Parallel and Distributed Computation: Numerical Methods. Prentice Hall, Englewood Cliffs (1989)
 2. Chen, G, Teboulle, M: A proximal-based decomposition method for convex minimization problems. Math. Program. **64**, 81-101 (1994)
 3. Chen, YM, Hager, WW, Yashtini, M, Ye, XJ, Zhang, HC: Bregman operator splitting with variable stepsize for total variation image reconstruction. Comput. Optim. Appl. **54**, 317-342 (2013)
 4. Facchinei, F, Pang, JS: Finite-Dimensional Variational Inequalities and Complementarity Problems. Springer, New York (2003)
 5. Fukushima, M: Application of the alternating direction of multipliers to separable convex programming problems. Comput. Optim. Appl. **2**, 93-111 (1992)
 6. Glowinski, R, Le Tallec, P: Augmented Lagrangian and Operator Splitting Methods in Nonlinear Mechanics. SIAM Studies in Applied Mathematics. SIAM, Philadelphia (1989)
 7. Nagurney, A, Zhang, D: Projection Dynamical Systems and Variational Inequalities with Applications. Kluwer Academic, Boston (1996)
 8. Glowinski, R, Marrocco, A: Sur l'approximation par éléments finis d'ordre un et la résolution par pénalisation-dualité d'une classe de problèmes de Dirichlet nonlinéaires. Rev. Fr. Autom. Inform. Rech. Opér., Anal. Numér. **2**, 41-76 (1975)
 9. He, BS, Liao, LZ, Han, DR, Yang, H: A new inexact alternating directions method for monotone variational inequalities. Math. Program. **92**(1), 103-118 (2002)
10. Martinet, B: Régularisation d'inéquations variationnelles par approximations successives. Rev. Fr. Inform. Rech. Opér. **4**, 154-159 (1970)
11. He, BS: Parallel splitting augmented Lagrangian method for solving for monotone structured variational inequalities. Comput. Optim. Appl. **42**, 195-212 (2009)
12. Tao, M, Yuan, XM: An inexact parallel splitting augmented Lagrangian method for monotone variational inequalities with separable structures. Comput. Optim. Appl. **52**, 439-461 (2012)
13. Zhang, WX, Han, DR, Yuan, XM: An efficient simultaneous method for the constrained multiple-sets split feasibility problem. Comput. Optim. Appl. **52**, 825-843 (2012)
14. Cai, XJ, Gu, GY, He, BS: On the $O(1/t)$ convergence rate of the projection and contraction methods for variational inequalities with Lipschitz continuous monotone operators. Comput. Optim. Appl. **57**, 339-363 (2014)

15. He, BS, Liao, LZ, Qian, MJ: Alternating projection based prediction-correction methods for structured variational inequalities. J. Comput. Math. **24**, 693-710 (2006)

16. Byrne, C: A unified treatment of some iterative algorithms in signal processing and image reconstruction. Inverse Probl. **20**, 103-120 (2004)

17. Censor, Y, Elfving, T, Kopf, N, Bortfeld, T: The multiple-sets split feasibility problem and its appplications for inverse problems. Inverse Probl. **21**, 2071-2084 (2005)

18. Censor, Y, Bortfeld, T, Martin, B, Trofimov, A: A unified approach for inversion problems in intensity-modulated radiation therapy. Phys. Med. Biol. **51**, 2353-2365 (2006)