



# A Family of Projection Gradient Methods for Solving the Multiple-Sets Split Feasibility Problem

Jinhua Wang<sup>1</sup> · Yaohua Hu<sup>2</sup> · Carisa Kwok Wai Yu<sup>3</sup> · Xiaojun Zhuang<sup>4</sup>

Received: 20 May 2018 / Accepted: 21 June 2019 / Published online: 9 July 2019  
© Springer Science+Business Media, LLC, part of Springer Nature 2019

## Abstract

In the present paper, we explore a family of projection gradient methods for solving the multiple-sets split feasibility problem, which include the cyclic/simultaneous iteration methods introduced in Wen et al. (J Optim Theory Appl 166:844–860, 2015) as special cases. For the general case, where the involved sets are given by level sets of convex functions, the calculation of the projection onto the level sets is complicated in general, and thus, the resulting projection gradient method cannot be implemented easily. To avoid this difficulty, we introduce a family of relaxed projection gradient methods, in which the projections onto the approximated halfspaces are adopted in place of the ones onto the level sets. They cover the relaxed cyclic/simultaneous iteration methods introduced in Wen et al. (J Optim Theory Appl 166:844–860, 2015) as special cases. Global weak convergence theorems are established for these methods. In particular, as direct applications of the established theorems, our results fill some gaps and deal with the imperfections that appeared in Wen et al. (J Optim Theory Appl 166:844–860, 2015) and hence improve and extend the corresponding results therein.

**Keywords** Multiple-sets split feasibility problem · Projection gradient methods · Global weak convergence

---

Communicated by Julian P. Revalski.

---

✉ Yaohua Hu  
mayhhu@szu.edu.cn

<sup>1</sup> Department of Mathematics, Hangzhou Normal University, Hangzhou 311121, People's Republic of China

<sup>2</sup> Shenzhen Key Laboratory of Advanced Machine Learning and Applications, College of Mathematics and Statistics, Shenzhen University, Shenzhen 518060, People's Republic of China

<sup>3</sup> Department of Mathematics and Statistics, The Hang Seng University of Hong Kong, Shatin, Hong Kong

<sup>4</sup> Department of Mathematics, Zhejiang University of Technology, Hangzhou 310032, People's Republic of China

**Mathematics Subject Classification** 90C25 · 90C30 · 47J25

## 1 Introduction

The split feasibility problem (SFP) is to find a point such that itself and its image under a linear transformation fall within two closed convex sets in the space and the image space, respectively. The SFP was introduced by Censor and Elfving in [1] to resolve the phase retrieval problem and provided a unified framework for vast application problems in various areas such as signal processing and image reconstruction [2,3], intensity modulated radiation therapy [4] and systems biology [5,6]. Various numerical algorithms have been developed to solve the SFP; see [2,5,7–9] and references therein. One of the most popular and practical algorithms is the CQ algorithm, which was proposed by Byrne [2,7]. One can further refer to [10,11] and references therein for the study of nonconvex feasibility problems.

The multiple-sets split feasibility problem (MSSFP) is a generalization of the SFP with the involved sets being the intersection of a series of closed convex sets. The MSSFP was introduced by Censor et al. [12] to resolve the intensity modulated radiation therapy treatment planning [4]. Many numerical algorithms have been developed to solve the MSSFP; see [12–17] and references therein. More precisely, Censor et al. in [12] proposed a projection gradient method, where the step size depends on the Lipschitz constant of the gradient operator. To avoid the inconvenience of calculating the Lipschitz constant, Zhao and Yang in [15] (also see [14]) introduced a self-adaptive projection method by adopting the Armijo-like search rule; Zhao and Yang in [16] further suggested a new self-adaptive method with an implementable dynamic step-size rule. Moreover, considering the MSSFP in Hilbert spaces, Xu in [13] proposed a fixed-point method with the step size relying on the Lipschitz constant of the gradient operator. Adopting the self-adaptive step-size rule, Wen et al. in [17] introduced the cyclic/simultaneous iteration methods for the MSSFP and proposed the relaxed cyclic/simultaneous iteration methods for the general MSSFP where the involved sets are given by the level sets of a series of convex functions. Global weak convergence results of these theorems were established therein. However, there are some gaps in the proofs of [17, Theorems 3.1 and 4.1] (as explained in Remarks 3.2 and 4.2), and some imperfections in [17, Theorems 3.2 and 4.2] (as explained in Remarks 3.3, 4.3 and Examples 3.1, 3.2).

In this paper, we consider a family of projection gradient methods for solving the MSSFP in Hilbert spaces, which include the cyclic/simultaneous iteration methods introduced in [17] as special cases. For the general case where the involved sets are given by level sets of a series of convex functions, the projections onto the level sets are not easily implemented in general. To avoid this difficulty, we introduce a family of relaxed projection gradient methods, in which the projections onto the approximated halfspaces are adopted in place of the ones onto the level sets; they cover the relaxed cyclic/simultaneous iteration methods introduced in [17]. Global weak convergence theorems of these methods will be established in the present paper. In particular, when applied to the cyclic/simultaneous iteration methods, our results fill the gaps

and correct the imperfections that appeared in [17] mentioned above and hence extend and improve the corresponding results in [17]. This is also a motivation of this paper.

The rest of this paper is organized as follows. In Sect. 2, we introduce some basic notations and preliminary results that will be used in the remaining sections. In Sect. 3, we propose a family of projection gradient methods for solving the MSSFP and establish their global weak convergence. In Sect. 4, we investigate a family of relaxed projection gradient methods and explore their global weak convergence. Conclusions will be drawn in the last section.

## 2 Notations and Preliminary Results

The notations used in this paper are standard (cf. [18]). We consider a real Hilbert space  $\mathbb{H}$ , associated with the inner product  $\langle \cdot, \cdot \rangle$ , and the induced norm  $\| \cdot \|$ . As usual,  $\mathbb{N}^*$  denotes the set of all positive integers, and  $\mathbb{N} := \mathbb{N}^* \cup \{0\}$ . For  $\{x_n\} \subseteq \mathbb{H}$ ,  $x_n \rightarrow x$  denotes that  $\{x_n\}$  converges weakly to  $x$ . For  $C \subseteq \mathbb{H}$ , the distance function of  $C$  and the projection operator onto  $C$  are denoted by  $d_C : \mathbb{H} \rightarrow \mathbb{R}$  and  $\mathbb{P}_C : \mathbb{H} \rightarrow \mathbb{H}$ , respectively, namely

$$d_C(x) := \inf\{\|x - c\| : c \in C\} \quad \text{for each } x \in \mathbb{H},$$

and

$$\mathbb{P}_C(x) := \{\bar{x} \in C : \|\bar{x} - x\| = d_C(x)\} \quad \text{for each } x \in \mathbb{H}.$$

Some useful properties about  $\mathbb{P}_C(\cdot)$  are presented in the following lemma; see [18].

**Lemma 2.1** *Let  $C \subseteq \mathbb{H}$  be nonempty, closed and convex. Then, the following assertions hold for any  $x, y \in \mathbb{H}$  and  $z \in C$ :*

- (i)  $\langle x - \mathbb{P}_C(x), z - \mathbb{P}_C(x) \rangle \leq 0$ ;
- (ii)  $\|\mathbb{P}_C(x) - \mathbb{P}_C(y)\|^2 \leq \langle \mathbb{P}_C(x) - \mathbb{P}_C(y), x - y \rangle$ ;
- (iii)  $\|\mathbb{P}_C(x) - z\|^2 \leq \|x - z\|^2 - \|\mathbb{P}_C(x) - x\|^2$ .

Recall that a mapping  $T : \mathbb{H} \rightarrow \mathbb{H}$  is said to be nonexpansive, if

$$\|Tx - Ty\| \leq \|x - y\| \quad \text{for each } x, y \in \mathbb{H}.$$

Clearly, it follows from Lemma 2.1(ii) that  $\mathbb{P}_C$  is nonexpansive. The following demi-closedness principle of a nonexpansive mapping is well known in Hilbert spaces; see [18, Theorem 4.17]. Let  $\mathbb{I}$  be the identity operator in  $\mathbb{H}$ , and let  $\text{Fix}(T)$  denote the set of all fixed points of  $T$ , namely

$$\text{Fix}(T) := \{x \in \mathbb{H} : T(x) = x\}.$$

**Lemma 2.2** *Let  $C \subseteq \mathbb{H}$  be nonempty, closed and convex, and let  $T : C \rightarrow C$  be a nonexpansive mapping with  $\text{Fix}(T) \neq \emptyset$ . If  $\{x_n\} \subseteq C$  converges weakly to  $x$  and  $\{(\mathbb{I} - T)(x_n)\}$  converges strongly to  $y$ , then  $(\mathbb{I} - T)(x) = y$ . In particular, if  $y = 0$ , then  $x = T(x)$ .*

The following lemma (cf. [18, Theorem 5.5]) will be useful in convergence analysis of projection gradient methods.

**Lemma 2.3** *Let  $C \subseteq \mathbb{H}$  be nonempty, closed and convex. Let  $\{x_n\} \subseteq \mathbb{H}$  satisfy that  $\lim_{n \rightarrow \infty} \|x_n - x\|$  exists for each  $x \in C$  and that each weak cluster point of  $\{x_n\}$  belongs to  $C$ . Then,  $\{x_n\}$  converges weakly to a point in  $C$ .*

For a convex function  $f : \mathbb{H} \rightarrow \mathbb{R}$ , the subdifferential of  $f$  at  $x$  is denoted by  $\partial f(x)$  and defined by

$$\partial f(x) := \{v \in \mathbb{H} : f(x) + \langle v, y - x \rangle \leq f(y) \text{ for each } y \in \mathbb{H}\}.$$

$f$  is said to be subdifferentiable at  $x \in \mathbb{H}$  if  $\partial f(x) \neq \emptyset$ . It is shown in [18, Proposition 16.3] that the convex function is subdifferentiable everywhere.

Let  $\mathbb{H}_1$  and  $\mathbb{H}_2$  be two real Hilbert spaces, and let  $A : \mathbb{H}_1 \rightarrow \mathbb{H}_2$  be a bounded linear operator. Let  $t, r \in \mathbb{N}^*$ , and consider two families of closed convex sets in  $\mathbb{H}_1$  and  $\mathbb{H}_2$ :  $\{C_i : i = 1, \dots, t\}$  and  $\{Q_j : j = 1, \dots, r\}$ , respectively. In the present paper, we consider the following multiple-sets split feasibility problem (MSSFP): Find a point  $x^* \in \mathbb{H}_1$  such that

$$x^* \in \bigcap_{i=1}^t C_i \quad \text{and} \quad Ax^* \in \bigcap_{j=1}^r Q_j. \tag{1}$$

In particular, if  $t = r = 1$ , then the MSSFP is reduced to the classical split feasibility problem (SFP). Let  $S$  denote the solution set of problem (1). Throughout the paper, we always assume that the MSSFP is consistent, that is,  $S \neq \emptyset$ .

Let  $\beta_j > 0$  for  $j = 1, \dots, r$ . We recall that the proximity function associated with (1), which was introduced in Xu [13]:

$$p(x) := \frac{1}{2} \sum_{j=1}^r \beta_j \|Ax - \mathbb{P}_{Q_j}(Ax)\|^2 \quad \text{for each } x \in \mathbb{H}_1. \tag{2}$$

The following lemma describes the property of  $p(\cdot)$ ; see [13, p. 2028].

**Lemma 2.4** *The function  $p(\cdot)$  is differentiable with its gradient given by*

$$\nabla p(\cdot) = \sum_{j=1}^r \beta_j A^*(\mathbb{I} - \mathbb{P}_{Q_j})(A\cdot), \tag{3}$$

and the Lipschitz constant of  $\nabla p(\cdot)$  is  $L := \|A\|^2 \sum_{j=1}^r \beta_j$ .

We end this section by the following lemma, which will be useful in the sequel.

**Lemma 2.5** *Let  $x \in \mathbb{H}_1$ ,  $z \in S$ , and  $\lambda > 0$ . Then,*

$$\|x - \lambda \nabla p(x) - z\|^2 \leq \|x - z\|^2 + \lambda^2 \|\nabla p(x)\|^2 - 4\lambda p(x). \quad (4)$$

**Proof** Note that

$$\|x - \lambda \nabla p(x) - z\|^2 = \|x - z\|^2 + \lambda^2 \|\nabla p(x)\|^2 - 2\lambda \langle \nabla p(x), x - z \rangle. \quad (5)$$

One has from (3) that

$$\langle \nabla p(x), x - z \rangle = \sum_{j=1}^r \beta_j \langle (\mathbb{I} - \mathbb{P}_{Q_j})(Ax), Ax - Az \rangle. \quad (6)$$

For  $j = 1, \dots, r$ , it follows from Lemma 2.1(i) that

$$\begin{aligned} & \langle (\mathbb{I} - \mathbb{P}_{Q_j})Ax, Ax - Az \rangle \\ &= \|Ax - \mathbb{P}_{Q_j}(Ax)\|^2 + \langle Ax - \mathbb{P}_{Q_j}(Ax), \mathbb{P}_{Q_j}(Ax) - Az \rangle \\ &\geq \|Ax - \mathbb{P}_{Q_j}(Ax)\|^2. \end{aligned}$$

Combining this with (6) yields that

$$\langle \nabla p(x), x - z \rangle \geq \sum_{j=1}^r \beta_j \|Ax - \mathbb{P}_{Q_j}(Ax)\|^2 = 2p(x).$$

This, together with (5), implies (4). The proof is complete.  $\square$

### 3 Projection Gradient Methods for MSSFP

In this section, for the case where the projections onto  $C_i$  and  $Q_j$  are easily implemented (e.g., they are halfspaces or box constraints), we propose a family of projection gradient methods for solving the MSSFP (1) and investigate its global weak convergence. Write  $I := \{1, \dots, t\}$  and recall  $p(\cdot)$  defined by (2). Throughout the paper, we adopt the convention that  $\frac{0}{0} := 0$ . The projection gradient method for solving the MSSFP is formally stated as follows.

#### Algorithm 3.1

Step 0 Choose  $x_0 \in \mathbb{H}_1$ ,  $\delta > 0$ ,  $0 < \underline{\rho} < \bar{\rho} < 4$ , and set  $n := 0$ .

Step 1 Determine the step size  $\lambda_n$  by

$$\lambda_n := \frac{\rho_n p(x_n)}{\|\nabla p(x_n)\|^2}, \text{ where } \underline{\rho} \leq \rho_n \leq \bar{\rho}. \quad (7)$$

Step 2 Take the weights  $\{\omega_i^n \in \mathbb{R}_+ : i \in I\}$  such that

$$\sum_{i=1}^t \omega_i^n = 1 \quad \text{and} \quad \inf_{i \in I_n} \omega_i^n > \delta, \quad \text{where } I_n := \{i \in I : \omega_i^n > 0\}. \quad (8)$$

Step 3 Set

$$x_{n+1} := \sum_{i=1}^t \omega_i^n \mathbb{P}_{C_i}(x_n - \lambda_n \nabla p(x_n)).$$

Step 4 Set  $n := n + 1$  and go to Step 1.

**Remark 3.1** Algorithm 3.1 includes the cyclic/simultaneous iteration methods introduced in [17] as special cases. In particular, we choose the weights as

$$\omega_i^n := \begin{cases} 1, & \text{if } i = (n \bmod t) + 1, \\ 0, & \text{otherwise.} \end{cases} \quad (9)$$

Then,  $\{x_n\}$  can be regarded as a sequence generated by the cyclic iteration method. Let  $\{w_i\}$  be defined in [17, Theorem 3.2], and set

$$\omega_i^n := w_i \quad \text{for each } i = 1, \dots, t. \quad (10)$$

Then,  $\{x_n\}$  is identical to a sequence generated by the simultaneous iteration method.

The notion of set control was introduced in [19,20] for feasibility problems, which is useful in convergence analysis of projection algorithms; see [8,21,22]. Here, we introduce a practical set control scheme for Algorithm 3.1, which will be useful in its convergence analysis.

**Definition 3.1** Let  $\{x_n\}$  be a sequence generated by Algorithm 3.1, and let  $q \in \mathbb{N}^*$ . We say that  $\{x_n\}$  satisfies the  $q$ -intermittent set control, if

$$I_n \cup \dots \cup I_{n+q-1} = I \quad \text{for each } n \in \mathbb{N}.$$

The main result of this section is presented in the following theorem, which establishes the global weak convergence property of the sequence generated by Algorithm 3.1. Recall that  $S$  is the solution set of (1).

**Theorem 3.1** *Let  $\{x_n\}$  be a sequence generated by Algorithm 3.1. Suppose that there exists  $q \in \mathbb{N}^*$  such that  $\{x_n\}$  satisfies the  $q$ -intermittent set control. Then,  $\{x_n\}$  converges weakly to a solution of (1).*

**Proof** Let  $z \in S$  and  $u_n := x_n - \lambda_n \nabla p(x_n)$ . In view of Algorithm 3.1, we obtain by the convexity of  $\|\cdot\|^2$  and Lemma 2.1(iii) that

$$\begin{aligned} \|x_{n+1} - z\|^2 &= \left\| \sum_{i=1}^t \omega_i^n (\mathbb{P}_{C_i}(u_n) - z) \right\|^2 \\ &\leq \sum_{i=1}^t \omega_i^n \|\mathbb{P}_{C_i}(u_n) - z\|^2 \\ &\leq \sum_{i=1}^t \omega_i^n (\|u_n - z\|^2 - \|\mathbb{P}_{C_i}(u_n) - u_n\|^2) \\ &\leq \|u_n - z\|^2 - \|x_{n+1} - u_n\|^2. \end{aligned} \quad (11)$$

This, together with (7) and (4), implies that

$$\begin{aligned} \|x_{n+1} - z\|^2 &\leq \|x_n - z\|^2 + \lambda_n^2 \|\nabla p(x_n)\|^2 - 4\lambda_n p(x_n) \\ &= \|x_n - z\|^2 - \rho_n(4 - \rho_n) \frac{p^2(x_n)}{\|\nabla p(x_n)\|^2}. \end{aligned} \quad (12)$$

Since  $0 < \rho_n < 4$ , it follows that  $\|x_{n+1} - z\| \leq \|x_n - z\|$ , which shows that  $\{x_n\}$  is a bounded sequence and  $\lim_{n \rightarrow \infty} \|x_n - z\|$  exists. Let  $x^*$  be a weak cluster point of  $\{x_n\}$ . That is, there exists a subsequence  $\{x_{n_l}\}$  such that  $x_{n_l} \rightharpoonup x^*$ . Below, we will show that  $x^* \in S$ . Granting this, one has from Lemma 2.3 that  $\{x_n\}$  converges weakly to a point in  $S$ ; hence the conclusion follows.

To complete the proof, it suffices to show that  $x^* \in S$ . By (12) and the fact that  $0 < \underline{\rho} \leq \rho_n \leq \bar{\rho} < 4$ , we obtain

$$\underline{\rho}(4 - \bar{\rho}) \frac{p^2(x_n)}{\|\nabla p(x_n)\|^2} \leq \|x_n - z\|^2 - \|x_{n+1} - z\|^2.$$

Letting  $n \rightarrow \infty$ , one has

$$\lim_{n \rightarrow \infty} \frac{p^2(x_n)}{\|\nabla p(x_n)\|^2} = 0. \quad (13)$$

Since  $\nabla p$  is Lipschitz continuous (cf. Lemma 2.4) and  $\{x_n\}$  is bounded,  $\{\|\nabla p(x_n)\|\}$  is bounded. Hence, we conclude from (13) that  $\lim_{n \rightarrow \infty} p(x_n) = 0$  and so by (2) that

$$\lim_{n \rightarrow \infty} \|Ax_n - \mathbb{P}_{Q_j}(Ax_n)\| = 0 \quad \text{for } j = 1, \dots, r.$$

Since  $x_{n_l} \rightharpoonup x^*$  and  $d_{Q_j}(A \cdot)$  is lower semicontinuous (thanks to its convexity), it follows that

$$d_{Q_j}(Ax^*) \leq \liminf_{l \rightarrow \infty} d_{Q_j}(Ax_{n_l}) = 0,$$

which implies  $Ax^* \in Q_j$  for  $j = 1, \dots, r$ ; consequently,  $Ax^* \in \bigcap_{j=1}^r Q_j$ .

Finally, it remains to show that  $x^* \in \bigcap_{i=1}^t C_i$ . Note that  $x_{n_l} \rightarrow x^*$  and it follows from (7) and (13) that

$$\lim_{n \rightarrow \infty} \lambda_n \|\nabla p(x_n)\| = 0. \tag{14}$$

This, together with the definition of  $u_n$ , implies that  $u_{n_l} \rightarrow x^*$  and

$$\lim_{n \rightarrow \infty} \|u_n - z\|^2 = \lim_{n \rightarrow \infty} \|x_n - z\|^2. \tag{15}$$

Observing from Lemma 2.1(iii) and the convexity of  $\|\cdot\|^2$ , we obtain

$$\begin{aligned} \sum_{i=1}^t \omega_i^n \|\mathbb{P}_{C_i}(u_n) - u_n\|^2 &\leq \|u_n - z\|^2 - \sum_{i=1}^t \omega_i^n \|\mathbb{P}_{C_i}(u_n) - z\|^2 \\ &\leq \|u_n - z\|^2 - \|x_{n+1} - z\|^2. \end{aligned}$$

Combining this with (15) yields

$$\lim_{n \rightarrow \infty} \sum_{i=1}^t \omega_i^n \|\mathbb{P}_{C_i}(u_n) - u_n\|^2 = 0. \tag{16}$$

Fix  $i \in I$  and  $n_l \in \mathbb{N}$ . Since  $\{x_n\}$  satisfies the  $q$ -intermittent set control, one has

$$I_{n_l} \cup \dots \cup I_{n_l+q-1} = I.$$

Hence, there exists  $q_i^l \in \{0, 1, \dots, q - 1\}$  such that  $i \in I_{n_l+q_i^l}$  and so

$$\omega_i^{n_l+q_i^l} > \delta > 0.$$

Let  $\tilde{n}_l := n_l + q_i^l$ . Note that, for each  $\tilde{n}_l$ ,

$$\delta \|\mathbb{P}_{C_i}(u_{\tilde{n}_l}) - u_{\tilde{n}_l}\| \leq \sum_{i=1}^t \omega_i^{\tilde{n}_l} \|\mathbb{P}_{C_i}(u_{\tilde{n}_l}) - u_{\tilde{n}_l}\|,$$

which, together with (16), implies

$$\lim_{l \rightarrow \infty} \|\mathbb{P}_{C_i}(u_{\tilde{n}_l}) - u_{\tilde{n}_l}\| = 0. \tag{17}$$

Below, we will show that  $u_{\tilde{n}_l} \rightarrow x^*$ . Granting this, it follows from (17) and Lemma 2.2 (with  $\mathbb{P}_{C_i}$  in place of  $T$ ) that  $x^* \in C_i$ . As  $i \in I$  is arbitrary, one has  $x^* \in \bigcap_{i=1}^t C_i$ , as desired.

To proceed, we obtain by (11) and (15) that

$$\lim_{n \rightarrow \infty} \|x_{n+1} - u_n\|^2 = 0. \tag{18}$$



Note by the definition of  $u_n$  that

$$\|x_{n+1} - u_n\|^2 = \|x_{n+1} - x_n\|^2 + \lambda_n^2 \|\nabla p(x_n)\|^2 + 2 \langle x_{n+1} - x_n, \lambda_n \nabla p(x_n) \rangle.$$

Letting  $n \rightarrow \infty$ , we get from (18) and (14) that  $\lim_{n \rightarrow \infty} \|x_{n+1} - x_n\| = 0$ . Then, it follows from the definition of  $u_n$  and (14) that

$$\lim_{n \rightarrow \infty} \|u_{n+1} - u_n\| = 0. \quad (19)$$

Let  $v \in \mathbb{H}_1$ . For each  $n_l \in \mathbb{N}$ , one has

$$\langle u_{\tilde{n}_l} - x^*, v \rangle = \langle u_{\tilde{n}_l} - u_{\tilde{n}_l-1}, v \rangle + \cdots + \langle u_{n_l} - x^*, v \rangle.$$

Since  $u_{n_l} \rightarrow x^*$  and  $\tilde{n}_l - n_l \leq q$ , it follows from (19) that

$$\lim_{l \rightarrow \infty} \langle u_{\tilde{n}_l} - x^*, v \rangle = 0,$$

and thus,  $u_{\tilde{n}_l} \rightarrow x^*$  because  $v \in \mathbb{H}_1$  is arbitrary. The proof is complete.  $\square$

By Remark 3.1 and Theorem 3.1, the convergence results of the iteration projection methods proposed in [17] can be directly obtained as follows.

**Corollary 3.1** *Let  $\{x_n\}$  be a sequence generated by Algorithm 3.1 with  $\{\omega_i^n\}$  given by (9). Then,  $\{x_n\}$  converges weakly to a solution of (1).*

**Proof** One can verify by (9) that  $I_n \cup \cdots \cup I_{n+t-1} = I$  for each  $n \in \mathbb{N}$ ; consequently,  $\{x_n\}$  satisfies the  $t$ -intermittent set control. Hence, Theorem 3.1 is applicable and the conclusion follows.  $\square$

**Corollary 3.2** *Let  $\{w_i\}_{i=1}^t \subset ]0, 1[$  be such that  $\sum_{i=1}^t w_i = 1$ . Let  $\{x_n\}$  be a sequence generated by Algorithm 3.1 with  $\{\omega_i^n\}$  given by (10). Then,  $\{x_n\}$  converges weakly to a solution of (1).*

**Proof** By the assumption, it is easy to see by (10) that  $I_n = I$  for each  $n \in \mathbb{N}$ , namely,  $\{x_n\}$  satisfies the 1-intermittent set control. Thus, Theorem 3.1 is applicable and the conclusion follows.  $\square$

**Remark 3.2** As mentioned in Remark 3.1, a sequence generated by Algorithm 3.1 with  $\{\omega_i^n\}$  given by (9) is identical to the one by the cyclic iteration method, whose weak convergence result was shown in [17, Theorem 3.1]. However, there is some gap in the proof of [17, Theorem 3.1]. More precisely, the authors claimed that “Notice that the pool  $\{1, \dots, t\}$  is finite, then for any  $i \in \{1, \dots, t\}$ , we can choose a subsequence  $\{n_{k_l}\} \subset \{n_k\}$  such that  $(n_{k_l} \bmod t) = i$ ,” whereas this is not true if the subsequence is  $\{n_k : n_k := tk + i_0\}$  for some fixed  $i_0$  or some other relevant cases. Corollary 3.1 fills the gap, and Theorem 3.1 extends the result of [17, Theorem 3.1].

**Remark 3.3** As mentioned in Remark 3.1, a sequence generated by Algorithm 3.1 with  $\{\omega_i^n\}$  given by (10) is identical to the one by the simultaneous iteration method, which was discussed in [17, Theorem 3.2]. In particular, the condition that  $\{w_i\}_{i=1}^t \subset [0, 1]$  is assumed in [17, Theorem 3.2]. As shown in Examples 3.1 and 3.2 below, [17, Theorem 3.2] may fail to (weakly) converge to a solution of (1) if  $\{w_i\}_{i=1}^t \subset [0, 1]$  is assumed. Our Corollary 3.2 and Theorem 3.1 correct and extend [17, Theorem 3.2], respectively.

We end this section by providing two counter examples of [17, Theorem 3.2], in which a sequence generated by the simultaneous iteration method with  $\{w_i\}_{i=1}^t \subset [0, 1]$  fails to (weakly) converge to a solution of (1). It should be remarked that  $\nabla p(x_n) = 0$  and  $\nabla p(x_n) \neq 0$  for each  $n \in \mathbb{N}$  in Examples 3.1 and 3.2, respectively.

**Example 3.1** Consider problem (1) with  $t := 3, r := 1, A := \mathbb{I}$  (the identity matrix in  $\mathbb{R}^2$ ) and

$$C_1 := [0, 4] \times [0, 4], \quad C_2 := [2, 6] \times [2, 6], \quad C_3 := [1, 3] \times [2, 6],$$

$$Q := [0, 4] \times [2, 6].$$

Clearly, the solution set of this problem is  $S = [2, 3] \times [2, 4]$ . Recall in [17, Theorem 3.2] that the simultaneous iteration method is presented as

$$x_{n+1} := \sum_{i=1}^3 \omega_i \mathbb{P}_{C_i}(x_n - \lambda_n \nabla p(x_n)) \quad \text{for each } n \geq 0, \tag{20}$$

where  $\omega_i \in [0, 1]$  such that  $\omega_1 + \omega_2 + \omega_3 = 1$ , and  $\lambda_n$  is given by (7). Select parameters  $\beta = 1, \omega_1 = \omega_2 = \frac{1}{2}, \omega_3 = 0$  and  $\rho_n \equiv 2$ , and take an initial point  $x_0 = (4, 6)^\top$ . Then, by (2) and (7), one has  $\nabla p(x_n) = 0$  and  $\lambda_n = 0$  for each  $n \in \mathbb{N}$ , and thus,  $\{x_n\}$  is well generated and

$$x_n = \left( 4, 4 + \frac{1}{2^{n-1}} \right)^\top \quad \text{for each } n \in \mathbb{N}.$$

Consequently,  $\{x_n\}$  converges to  $(4, 4)^\top \notin S$ . Hence, [17, Theorem 3.2] fails in this example.

**Example 3.2** Consider problem (1) with  $t = r = 2, A := \mathbb{I}$  and

$$C_1 := \{(x_1, x_2) \in \mathbb{R}^2 : -1 \leq x_1 \leq 1, x_1 = x_2\},$$

$$C_2 := \{(x_1, x_2) \in \mathbb{R}^2 : -2 \leq x_1 \leq 0, -2 \leq x_2 \leq 0, -2 \leq x_1 + x_2 \leq -1\},$$

$$Q_1 := [-1, 0] \times [-1, 1], \quad Q_2 := [-1, 1] \times [-1, 0].$$

Then, the solution set of (1) is  $S = \{(x_1, x_2) \in \mathbb{R}^2 : -1 \leq x_1 \leq -\frac{1}{2}, x_1 = x_2\}$ . Select parameters  $\beta_1 = \beta_2 = \frac{1}{2}, \omega_1 = 1, \omega_2 = 0$  and  $\rho_n \equiv 1$ , and take an initial point

$x_0 = (1, 1)^\top$ . Then, by (2) and (20), one has  $\nabla p(x_n) \neq 0$  and  $\lambda_n$  is well defined for each  $n \in \mathbb{N}$ , and thus,  $\{x_n\}$  is well generated and

$$x_n = (2^{-n}, 2^{-n})^\top \quad \text{for each } n \in \mathbb{N}.$$

Consequently,  $\{x_n\}$  converges to the origin, which is not a solution of (1). Hence, [17, Theorem 3.2] fails in this example.

### 4 Relaxed Projection Gradient Methods for MSSFP

In this section, we consider a general case of the MSSFP, where  $\{C_i\}'_{i=1}$  and  $\{Q_j\}'_{j=1}$  in (1) are given by level sets of convex functions. Throughout this section, we assume that each  $c_i : \mathbb{H}_1 \rightarrow \mathbb{R}$  and  $q_j : \mathbb{H}_2 \rightarrow \mathbb{R}$  are convex functions and the sets  $C_i$  and  $Q_j$  are given, respectively, by

$$C_i := \{x \in \mathbb{H}_1 : c_i(x) \leq 0\} \quad \text{and} \quad Q_j := \{y \in \mathbb{H}_2 : q_j(y) \leq 0\}. \quad (21)$$

We assume that  $\partial c_i$  and  $\partial q_j$  are bounded operators (i.e., bounded on any bounded set) for each  $i = 1, \dots, t$  and  $j = 1, \dots, r$ .

In this situation, the projections onto  $C_i$  and  $Q_j$  are not easily implemented in general. To avoid this difficulty, we introduce a family of relaxed projection gradient methods, in which the projections onto the approximated halfspaces are adopted in place of the projections onto  $C_i$  and  $Q_j$  (as in Algorithm 3.1). In particular, letting  $n \in \mathbb{N}$ ,  $i \in \{1, \dots, t\}$  and  $j \in \{1, \dots, r\}$ , we use  $C_i^n$  and  $Q_j^n$  to denote the approximated halfspaces of  $C_i$  and  $Q_j$  at  $x_n$ , i.e.,

$$C_i^n := \{x \in \mathbb{H}_1 : c_i(x_n) + \langle \xi_i^n, x - x_n \rangle \leq 0\}, \quad \text{where } \xi_i^n \in \partial c_i(x_n),$$

and

$$Q_j^n := \{y \in \mathbb{H}_2 : q_j(Ax_n) + \langle \eta_j^n, y - Ax_n \rangle \leq 0\}, \quad \text{where } \eta_j^n \in \partial q_j(Ax_n),$$

respectively. By the definition of subgradient, it follows that  $C_i \subseteq C_i^n$  and  $Q_j \subseteq Q_j^n$ . As in [17], letting  $\beta_j > 0$  for  $j = 1, \dots, r$ , we define the following (relaxed) proximity function

$$p_n(x) := \frac{1}{2} \sum_{j=1}^r \beta_j \|Ax - \mathbb{P}_{Q_j^n}(Ax)\|^2 \quad \text{for each } x \in \mathbb{H}_1. \quad (22)$$

It follows from Lemma 2.4 that  $p_n(\cdot)$  is differentiable with its gradient given by

$$\nabla p_n(x) = \sum_{j=1}^r \beta_j A^*(\mathbb{I} - \mathbb{P}_{Q_j^n})(Ax).$$

Recall that  $I := \{1, \dots, t\}$ . The relaxed projection gradient method for solving the MSSFP (1) with  $\{C_i\}$  and  $\{Q_j\}$  given by (21) is formulated as follows.

**Algorithm 4.1**

Step 0 Choose  $x_0 \in \mathbb{H}_1, \delta > 0, 0 < \underline{\rho} < \bar{\rho} < 4$ , and set  $n := 0$ .

Step 1 Choose the step size  $\lambda_n$  by

$$\lambda_n := \frac{\rho_n p_n(x_n)}{\|\nabla p_n(x_n)\|^2}, \text{ where } \underline{\rho} \leq \rho_n \leq \bar{\rho}.$$

Step 2 Take the weights  $\{\omega_i^n \in \mathbb{R}_+ : i \in I\}$  satisfying (8).

Step 3 Set

$$x_{n+1} := \sum_{i=1}^I \omega_i^n \mathbb{P}_{C_i^n}(x_n - \lambda_n \nabla p_n(x_n)).$$

Step 4 Set  $n := n + 1$  and go to Step 1.

**Remark 4.1** Algorithm 4.1 includes the relaxed cyclic/simultaneous iteration methods introduced in [17] as special cases where the weights  $\{\omega_i^n : i \in I\}$  are chosen by (9) and (10), respectively.

The main result of this section is stated in the following theorem, in which we prove the global weak convergence of the sequence generated by Algorithm 4.1. Recall that  $S$  is the solution of (1).

**Theorem 4.1** *Let  $\{x_n\}$  be a sequence generated by Algorithm 4.1. Suppose that there exists  $q \in \mathbb{N}^*$  such that  $\{x_n\}$  satisfies the  $q$ -intermittent set control. Then,  $\{x_n\}$  converges weakly to a solution of (1).*

**Proof** Let  $z \in S$  and  $u_n := x_n - \lambda_n \nabla p_n(x_n)$ . Following the same lines as in the beginning of the proof of Theorem 3.1, we can deduce that  $\lim_{n \rightarrow \infty} \|x_n - z\|$  exists and

$$\lim_{n \rightarrow \infty} p_n(x_n) = 0 \quad \text{and} \quad \lim_{n \rightarrow \infty} \|u_n - x_n\| = 0. \tag{23}$$

Let  $x^*$  be a weak cluster point of  $\{x_n\}$ , namely, there exists a subsequence  $\{x_{n_k}\}$  such that  $x_{n_k} \rightharpoonup x^*$ . Below, we will show that  $x^* \in S$ . Granting this, one has from Lemma 2.3 that  $\{x_n\}$  converges weakly to a point in  $S$ ; hence the conclusion follows.

To accomplish the proof, it suffices to show that  $x^* \in S$ . Note by (22) and the first equality of (23) that

$$\lim_{n \rightarrow \infty} \|Ax_n - \mathbb{P}_{Q_j^n}(Ax_n)\| = 0 \quad \text{for } j = 1, \dots, r. \tag{24}$$

By assumption that  $\partial q_j$  is bounded on bounded set and noting that  $\{x_n\}$  is bounded, there exists  $\eta > 0$  such that  $\|\eta_j^n\| \leq \eta$  for each  $n \in \mathbb{N}$  and  $j = 1, \dots, r$ . Hence, one has by the definition of  $Q_j^n$  that

$$q_j(Ax_n) \leq \langle \eta_j^n, Ax_n - \mathbb{P}_{Q_j^n}(Ax_n) \rangle \leq \eta \|Ax_n - \mathbb{P}_{Q_j^n}(Ax_n)\|. \tag{25}$$

By the convexity of  $q_j$ , it is weakly lower semicontinuous, and then, we obtain from (24) and (25) that  $q_j(Ax^*) \leq \liminf_{k \rightarrow \infty} q_j(Ax_{n_k}) \leq 0$ , which says  $Ax^* \in Q_j$  for  $j = 1, \dots, r$ ; consequently,  $Ax^* \in \bigcap_{j=1}^r Q_j$ .

Finally, it remains to show that  $x^* \in \bigcap_{i=1}^t C_i$ . Fix  $i \in I$ . Following the same arguments as in the proof of Theorem 3.1 [cf. (18)], we can show that there exists a subsequence  $\{\tilde{n}_k\}$  such that

$$\lim_{k \rightarrow \infty} \|\mathbb{P}_{C_i^{\tilde{n}_k}}(u_{\tilde{n}_k}) - u_{\tilde{n}_k}\| = 0. \tag{26}$$

and  $x_{\tilde{n}_k} \rightharpoonup x^*$ . Note that there exists  $\zeta > 0$  such that  $\|\xi_i^{\tilde{n}_k}\| \leq \zeta$  for each  $k \in \mathbb{N}$  (by the assumption that  $\partial c_i$  is bounded on bounded set); consequently, one has by the definition of  $C_j^n$  that

$$\begin{aligned} c_i(x_{\tilde{n}_k}) &\leq \langle \xi_i^{\tilde{n}_k}, x_{\tilde{n}_k} - \mathbb{P}_{C_i^{\tilde{n}_k}}(u_{\tilde{n}_k}) \rangle \\ &\leq \zeta (\|x_{\tilde{n}_k} - u_{\tilde{n}_k}\| + \|u_{\tilde{n}_k} - \mathbb{P}_{C_i^{\tilde{n}_k}}(u_{\tilde{n}_k})\|). \end{aligned}$$

By the lower semicontinuity of  $c_i(\cdot)$  (thanks to its convexity) and  $x_{\tilde{n}_k} \rightharpoonup x^*$ , this, together with (26) and the second equality of (23), yields that

$$c_i(x^*) \leq \liminf_{k \rightarrow \infty} c_i(x_{\tilde{n}_k}) \leq 0;$$

consequently,  $x^* \in C_i$ . As  $i \in I$  is arbitrary, it follows that  $x^* \in \bigcap_{i=1}^t C_i$ , as desired. The proof is complete. □

By applying Remark 4.1 and Theorem 4.1, the convergence results of the relaxed cyclic/simultaneous iteration methods proposed in [17] are directly obtained as follows.

**Corollary 4.1** *Let  $\{x_n\}$  be a sequence generated by Algorithm 4.1 with  $\{\omega_i^n\}$  given by (9). Then,  $\{x_n\}$  converges weakly to a solution of (1).*

**Corollary 4.2** *Let  $\{w_i\}_{i=1}^t \subset ]0, 1[$  such that  $\sum_{i=1}^t w_i = 1$ . Let  $\{x_n\}$  be a sequence generated by Algorithm 4.1 with  $\{\omega_i^n\}$  given by (10). Then,  $\{x_n\}$  converges weakly to a solution of (1).*

**Remark 4.2** As mentioned in Remark 4.1, a sequence generated by Algorithm 4.1 with  $\{\omega_i^n\}$  given by (9) is identical to the one by the relaxed cyclic iteration method, whose weak convergence result was shown in [17, Theorem 4.1]. However, there is some gap in the proof of [17, Theorem 4.1]. More precisely, the authors claimed that “Since the pool of convex sets  $\{C_i\}_{i=1}^t$  is finite. For any  $i \in \{1, \dots, t\}$ , we can choose a subsequence  $\{n_{k_i}\} \subset \{n_k\}$  such that  $(n_{k_i} \bmod t) = i$ ,” whereas this is not true if the subsequence is  $\{n_k : n_k := tk + i_0\}$  for some fixed  $i_0$  or some other relevant cases. Our Corollary 4.1 fills the gap, and Theorem 4.1 extends the result in [17, Theorem 4.1].

**Remark 4.3** As mentioned in Remark 4.1, a sequence generated by Algorithm 4.1 with  $\{\omega_i^n\}$  given by (10) is identical to the one by the relaxed simultaneous iteration method, which was discussed in [17, Theorem 4.2]. The condition  $\{w_i\}_{i=1}^t \subset [0, 1]$  is assumed in [17, Theorem 4.2]; however, the convergence may fail under this assumption (see Example 3.1). Hence, our Corollary 4.2 and Theorem 4.1 correct and extend [17, Theorem 4.2], respectively.

## 5 Conclusions

The multiple-sets split feasibility problem (MSSFP) is a generalization of the split feasibility problem, which has attracted a great amount of attention in numerical algorithms and has been widely applied in various fields. In the present paper, a family of projection gradient methods for the MSSFP were proposed for the cases where the projections onto the involved sets are easily implemented, which include the cyclic/simultaneous iteration methods introduced in [17] as special cases. For the general case where the involved sets are given by level sets of convex functions, we proposed a family of relaxed projection gradient methods (with projections onto the approximated halfspaces in place of the ones onto the level sets), which cover the relaxed cyclic/simultaneous iteration method introduced in [17] as special cases. Global weak convergence theorems of these methods were established in the present paper. In particular, as their direct applications, our results fill some gaps and imperfections and improve/extend the corresponding results in [17].

**Acknowledgements** The authors are grateful to the anonymous referees and the editor for their valuable comments and suggestions that helped to improve the quality of this paper. Jinhua Wang was supported in part by the National Natural Science Foundation of China (Grant 11771397) and Zhejiang Provincial Natural Science Foundation of China (Grant LY17A010021). Yaohua Hu was supported in part by National Natural Science Foundation of China (11601343, 11601344, 11871347), Natural Science Foundation of Guangdong (2016A030310038), Natural Science Foundation of Shenzhen (JCYJ20170817100950436, JCYJ20170818091621856) and Interdisciplinary Innovation Team of Shenzhen University. Carisa Kwok Wai Yu is supported in part by the Research Grants Council of the Hong Kong Special Administrative Region, China (UGC/FDS14/P02/15 and UGC/FDS14/P02/17).

## References

1. Censor, Y., Elfving, T.: A multiprojection algorithm using Bregman projections in a product space. *Numer. Algorithms* **8**, 221–239 (1994)
2. Byrne, C.: A unified treatment of some iterative algorithms in signal processing and image reconstruction. *Inverse Probl.* **20**, 103–120 (2004)
3. He, H., Ling, C., Xu, H.K.: An implementable splitting algorithm for the  $\ell_1$ -norm regularized split feasibility problem. *J. Sci. Comput.* **67**, 1–18 (2015)
4. 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)
5. Wang, J.H., Hu, Y.H., Li, C., Yao, J.C.: Linear convergence of CQ algorithms and applications in gene regulatory network inference. *Inverse Probl.* **33**, 055017 (2017)
6. Hu, Y.H., Li, C., Meng, K.W., Qin, J., Yang, X.Q.: Group sparse optimization via  $\ell_{p,q}$  regularization. *J. Mach. Learn. Res.* **18**, 1–52 (2017)
7. Byrne, C.: Iterative oblique projection onto convex sets and the split feasibility problem. *Inverse Probl.* **18**, 441–453 (2002)

8. Bauschke, H.H., Borwein, J.M.: On projection algorithms for solving convex feasibility problems. *SIAM Rev.* **38**, 367–426 (1996)
9. Bauschke, H.H., Kruk, S.G.: Reflection-projection method for convex feasibility problems with an obtuse cone. *J. Optim. Theory Appl.* **120**, 503–531 (2004)
10. Hu, Y.H., Li, C., Yang, X.Q.: On convergence rates of linearized proximal algorithms for convex composite optimization with applications. *SIAM J. Optim.* **26**, 1207–1235 (2016)
11. Li, G.Y., Pong, T.K.: Douglas–Rachford splitting for nonconvex optimization with application to nonconvex feasibility problems. *Math. Program. Ser. A* **159**, 371–401 (2016)
12. Censor, Y., Elfving, T., Kopf, N., Bortfeld, T.: The multiple-sets split feasibility problem and its applications for inverse problems. *Inverse Probl.* **21**, 2071–2084 (2005)
13. Xu, H.K.: A variable Krasnosel’skiĭ–Mann algorithm and the multiple-set split feasibility problem. *Inverse Probl.* **22**, 2021–2034 (2006)
14. Zhang, W.X., Han, D.R., Li, Z.B.: A self-adaptive projection method for solving the multiple-sets split feasibility problem. *Inverse Probl.* **25**, 115001 (2009)
15. Zhao, J.L., Yang, Q.Z.: Self-adaptive projection methods for the multiple-sets split feasibility problem. *Inverse Probl.* **27**, 035009 (2011)
16. Zhao, J.L., Yang, Q.Z.: A simple projection method for solving the multiple-sets split feasibility problem. *Inverse Probl. Sci. Eng.* **21**, 537–546 (2013)
17. Wen, M., Peng, J.G., Tang, Y.C.: A cyclic and simultaneous iterative method for solving the multiple-sets split feasibility problem. *J. Optim. Theory Appl.* **166**, 844–860 (2015)
18. Bauschke, H.H., Combettes, P.L.: *Convex Analysis and Monotone Operator Theory in Hilbert Space*. Springer, London (2011)
19. Bregman, L.M.: The method of successive projection for finding a common point of convex sets. *Sov. Math. Dokl.* **6**, 688–692 (1965)
20. Gubin, L.G., Polyak, B.T., Raik, E.V.: The method of projections for finding the common point of convex sets. *USSR Comput. Math. Math. Phys.* **7**, 1–24 (1967)
21. Combettes, P.L.: Hilbertian convex feasibility problem: convergence of projection methods. *Appl. Math. Optim.* **35**, 311–330 (1997)
22. Zhao, X., Ng, K.F., Li, C., Yao, J.C.: Linear regularity and linear convergence of projection-based methods for solving convex feasibility problems. *Appl. Math. Optim.* **78**, 613–641 (2018)

**Publisher’s Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.