Dong et al. *Journal of Inequalities and Applications* (2019) 2019:77
https://doi.org/10.1186/s13660-019-2030-x

Journal of Inequalities and Applications
a SpringerOpen Journal

RESEARCH Open Access

Some algorithms for classes of split feasibility problems involving paramonotone equilibria and convex optimization

Q.L. Dong[1], X.H. Li[1], D. Kitkuan[2], Y.J. Cho[3,4] and P. Kumam[2,5*]

*Correspondence:
poom.kum@kmutt.ac.th
[2]KMUTTFixed Point Research Laboratory, Department of Mathematics, Faculty of Science, King Mongkut's University of Technology Thonburi (KMUTT), Bangkok, Thailand
[5]Department of Medical Research, China Medical University Hospital, China Medical University, Taichung, Taiwan
Full list of author information is available at the end of the article

Abstract

In this paper, we first introduce a new algorithm which involves projecting each iteration to solve a split feasibility problem with paramonotone equilibria and using unconstrained convex optimization. The strong convergence of the proposed algorithm is presented. Second, we also revisit this split feasibility problem and replace the unconstrained convex optimization by a constrained convex optimization. We introduce some algorithms for two different types of objective function of the constrained convex optimization and prove some strong convergence results of the proposed algorithms. Third, we apply our algorithms for finding an equilibrium point with minimal environmental cost for a model in electricity production. Finally, we give some numerical results to illustrate the effectiveness and advantages of the proposed algorithms.

MSC: 47H05; 47H07; 47H10; 54H25

Keywords: Split feasibility problem; Equilibria; Constrained convex optimization; Practical model

1 Introduction and the problem statement

Let H_1 and H_2 be two real Hilbert spaces with inner product $\langle\cdot,\cdot\rangle$ and reduced norm $\|\cdot\|$, C and Q be nonempty closed convex subsets in H_1 and H_2, respectively.

In [7], Censor and Elving first introduced the *split feasibility problem* (shortly, SFP) in Euclidean space, which is formulated as follows:

$$\text{Find } x^* \in C \text{ such that } Ax^* \in Q,$$

where $A : H_1 \to H_2$ is a bounded linear operator. The SFP can be a model for many inverse problems where constraints are imposed on the solutions in the domain of the linear operator as well as in its range. It has a variety of specific applications in real world such as medical care, image reconstruction and signal processing (see [5, 14–17, 21, 38, 39] for more details).

Let $f : C \times C \to \mathbb{R}$ be a bifunction such that $f(x,x) = 0$ for all $x \in C$. The *equilibrium problem* (shortly, EP)

$$\text{Find } x^* \in C \text{ such that } f\left(x^*, y\right) \geq 0 \text{ for all } y \in C$$

was firstly introduced by Fan [19] and further studied by Blum and Oettli [2]. The solution set of the EP is denoted by Sol(EP). The EP is a generalization of many mathematical models, including variational inequality, fixed point, optimization, complementarity problems (see, for instance, [2, 4, 13, 18, 20, 22, 23, 25, 31, 33–36]).

Recently, Yen et al. [37] investigated the following *split feasibility problem* involving paramonotone equilibria and convex optimization (shortly, SEO):

Problem 1.1 Find $x^* \in C$ such that $f(x^*, y) \geq 0$ for all $y \in C$ and $g(Ax^*) \leq g(z)$ for all $z \in H_2$,

where g is a proper lower semi-continuous convex function on H_2. Also, they introduced the following algorithm to solve Problem 1.1:

Algorithm 1.1 For any $x^k \in C$, take $\eta^k \in \partial_2^{\epsilon_k} f(x^k, x^k)$ and define

$$\alpha_k = \frac{\beta_k}{\gamma_k},$$

where $\gamma_k = \max\{\delta_k, \|\eta^k\|\}$. Compute

$$y^k = P_C\big(x^k - \alpha_k \eta^k\big).$$

Take

$$\mu_k := \begin{cases} 0, & \text{if } \nabla h(y^k) = 0, \\ \rho_k \frac{h(y^k)}{\|\nabla h(y^k)\|^2}, & \text{if } \nabla h(y^k) \neq 0, \end{cases}$$

and compute

$$z^k = P_C\big(y^k - \mu_k A^*(I - \operatorname{prox}_{\lambda g})\big(Ay^k\big)\big).$$

Let

$$x^{k+1} = a_k x^k + (1 - a_k) z^k.$$

In Algorithm 1.1, $\operatorname{prox}_{\lambda g}$ denotes proximal mapping of the convex function g with $\lambda > 0$, and the parameters $\{a_k\}$, $\{\delta_k\}$, $\{\beta_k\}$, $\{\epsilon_k\}$ and $\{\rho_k\}$ are taken as in Algorithm 3.1 (see below Sect. 3).

Note that Algorithm 1.1 involves two exact projections onto the feasible set C, which limits the applicability of the method, especially when such projections are hard to compute. It is well known that only in a few specific instances the projection onto a convex set has an explicit formula. When the feasible set C is a general closed convex set, we must solve a nontrivial quadratic problem in order to compute the projection onto C.

In this paper, by expanding the domain of function f, we introduce a new algorithm which just involves a projection onto C. Also, we revisit Problem 1.1 and replace the unconstrained convex optimization by a constrained convex optimization. Further, we introduce two iterative algorithms to solve the new model and prove some strong convergence results of the proposed algorithms.

The paper is organized as follows: Sect. 2 deals with some definitions and lemmas for the main results in this paper. In Sect. 3, we introduce a new algorithm, which involves a projection in each iteration. In Sect. 4, we introduce two algorithms and study their convergence. In Sect. 5, we provide a practical model for an electricity market and some computational results for the model.

2 Preliminaries

The following definitions and lemmas are useful for the validity and convergence of the algorithms.

Definition 2.1 Let H be a Hilbert space, $T : H \to H$ be a mapping and let $K \subseteq H$.

(i) T is said to be *nonexpansive* if

$$\|Tx - Ty\| \le \|x - y\|$$

for all $x, y \in H$.

(ii) T is said to be *firmly nonexpansive* if

$$\|Tx - Ty\| \le \langle x - y, Tx - Ty\rangle$$

for all $x, y \in H$, or

$$0 \le \bigl\langle Tx - Ty, (I - T)x - (I - T)y\bigr\rangle$$

for all $x, y \in H$.

(iii) T is said to be *Lipschitz continuous* with Lipschitz constant L if

$$\|Tx - Ty\| \le L\|x - y\|$$

for all $x, y \in H$.

(iv) T is said to be α*-averaged* if

$$T = (1 - \alpha)I + \alpha S,$$

where $\alpha \in (0, 1)$ and $S : H \to H$ is a nonexpansive mapping.

Lemma 2.1 ([1, Proposition 4.4]) *Let H be a Hilbert space and $T : H \to H$ be a mapping. Then the following are equivalent*:

(i) *T is firmly nonexpansive*;

(ii) *$I - T$ is firmly nonexpansive*.

Lemma 2.2 ([3, 9]) *The composition of finitely many averaged mappings is averaged. In particular, if T_1 is α_1-averaged and T_2 is α_2-averaged, where $\alpha_1, \alpha_2 \in (0, 1)$, then the composition T_1T_2 is α-averaged, where $\alpha = \alpha_1 + \alpha_2 - \alpha_1\alpha_2$.*

It is easy to show that firmly nonexpansive mappings are $\frac{1}{2}$-averaged, and averaged mappings are nonexpansive.

For a mapping $T : H \to H$, $\mathrm{Fix}(T)$ denotes the set of fixed points of T, i.e.,

$$\mathrm{Fix}(T) := \{x \in H : Tx = x\}.$$

It is well known that every nonexpansive operator $T : H \to H$ satisfies the following inequality:

$$\bigl\langle (x - Tx) - (y - Ty), Ty - Tx \bigr\rangle \le \frac{1}{2} \bigl\| (x - Tx) - (y - Ty) \bigr\|^2$$

for all $x, y \in H$ and so

$$\langle x - Tx, y - Tx \rangle \le \frac{1}{2} \|x - Tx\|^2$$

for all $x \in H$ and $y \in \mathrm{Fix}(T)$ (see, for example, [11, Theorem 3], [12, Theorem 1]).

Let H be a real Hilbert space and K be a nonempty convex closed subset of H. For each point $x \in H$, there exists a unique nearest point in K, denoted by $P_K(x)$, such that

$$\bigl\| x - P_K(x) \bigr\| \le \|x - y\|$$

for all $y \in K$. The mapping $P_K : H \to K$ is called *the metric projection* of H onto K. It is well known that P_K is a nonexpansive mapping of H onto K and even a firmly nonexpansive mapping. So, P_K is also $\frac{1}{2}$-averaged, which is captured in the following lemma:

Lemma 2.3 *For any $x, y \in H$ and $z \in K$, the following hold*:
(i) $\|P_K(x) - P_K(y)\|^2 \le \|x - y\|$;
(ii) $\|P_K(x) - z\|^2 \le \|x - z\|^2 - \|P_K(x) - x\|^2$.

Some characterizations of the metric projection P_K are given by the two properties in the following lemma:

Lemma 2.4 *Let $x \in H$ and $z \in K$. Then $z = P_K(x)$ if and only if $P_K(x) \in K$ and*

$$\bigl\langle x - P_K(x), P_K(x) - y \bigr\rangle \ge 0$$

for all $x \in H$ and $y \in K$.

Lemma 2.5 *Let C be a nonempty closed convex subset in a Hilbert space H and $P_C(x)$ be the metric projection of x onto C. Then we have*
(i) $\langle x - y, P_C(x) - P_C(y) \rangle \ge \|P_C(x) - P_C(y)\|^2$ *for all $x, y \in C$*;
(ii) $\|z^k - y^k\| \le \beta_k$.

Lemma 2.6 *Let $\{v^k\}$ and $\{\delta_k\}$ be the nonnegative sequences of real numbers satisfying $v^{k+1} \le v^k + \delta_k$ with $\sum_{k=1}^{\infty} \delta_k < +\infty$. Then the sequence $\{v^k\}$ is convergent.*

Lemma 2.7 *Let H be a real Hilbert space, $\{a_k\}$ be a sequence of real numbers such that $0 < a < a_k < b < 1$ for all $k \ge 1$ and $\{v^k\}$, $\{w^k\}$ be the sequences in H such that*

$$\limsup_{k \to +\infty} \bigl\| v^k \bigr\| \le c, \qquad \limsup_{k \to +\infty} \bigl\| w^k \bigr\| \le c$$

and, for some $c > 0$,

$$\limsup_{k\to+\infty}\left\|a_k v^k + \left(1-a^k\right)w^k\right\| = c.$$

Then $\lim_{k\to+\infty}\|v^k - w^k\| = 0$.

Definition 2.2 ([28]) The *normal cone* of K at $v \in K$, denote by N_K, is defined as follows:

$$N_K(v) := \left\{d \in H : \langle d, y - v\rangle \leq 0, \forall y \in K\right\}.$$

Definition 2.3 ([1, Definition 16.1]) The *subdifferential set* of a convex function c at a point x is defined as follows:

$$\partial c(x) := \left\{\xi \in H : c(y) \geq c(x) + \langle \xi, y - x\rangle, \forall y \in H\right\}.$$

Define by ι_K the *indicator function* of the set K, i.e.,

$$\iota_K(x) = \begin{cases} 0 & x \in K, \\ +\infty & \text{otherwise.}\end{cases}$$

It is well known that $\partial\iota_K(x) = N_K(x)$ and $(I + \lambda N_K)^{-1} = P_K$ for any $\lambda > 0$.

Let $f : H \times H \to \mathbb{R}$ be a bifunction. We need the following assumptions on $f(x,y)$ for our algorithms and convergence:

(A1) For each $x \in C$, $f(x,x) = 0$ and $f(x,\cdot)$ is lower semi-continuous and convex on C;

(A2) $\partial_2^\epsilon f(x,x)$ is nonempty for any $\epsilon > 0$ and $x \in C$ and is bounded on any bounded subset of C, where $\partial_2^\epsilon f(x,x)$ denotes ϵ-subdifferential of the convex function $f(x,\cdot)$ at x, that is,

$$\partial_2^\epsilon f(x,x) := \left\{\eta \in H_1 : \langle \eta, y - x\rangle + f(x,x) \leq f(x,y) + \epsilon, \forall y \in C\right\}; \tag{1}$$

(A3) f is pseudo-monotone on C with respect to every solution of the EP, that is, $f(x,x^*) \leq 0$ for any $x \in C$, $x^* \in \mathrm{Sol}(EP)$ and f satisfies the following condition, which is called the *para-monotonicity property*:

$$x^* \in \mathrm{Sol}(EP), \qquad y \in C, \qquad f\left(x^*, y\right) = f\left(y, x^*\right) = 0 \quad \Longrightarrow \quad y \in \mathrm{Sol}(EP);$$

(A4) For all $x \in K$, $f(\cdot, x)$ is weakly upper semi-continuous on C.

3 A new algorithm for Problem 1.1 and its convergence analysis

In this section we give a new algorithm for Problem 1.1 and analyze its convergence.

Recall that the proximal mapping of the convex function g with $\lambda > 0$, denoted by $\mathrm{prox}_{\lambda g}$, is defined as the unique solution of the strongly convex programming problem:

$$\mathrm{prox}_{\lambda g}(u) = \operatorname*{argmin}_{v\in H_2}\left\{g(v) + \frac{1}{2\lambda}\|v - u\|^2\right\}. \tag{P(u)}$$

The proximal mapping has some good properties, namely, it is firmly nonexpansive and $\mathrm{prox}_{\lambda g} = P_Q$ when $g = \delta_Q$ (see, e.g., [26]).

For any $\lambda > 0$, we set

$$h(x) := \frac{1}{2}\left\|(I - \operatorname{prox}_{\lambda g})Ax\right\|^2.$$

By using the necessary and sufficient optimality condition for convex programming, we can see that $h(x) = 0$ if and only if Ax solves $P(u)$ with $u = Ax$. Note that, even though g may not be differentiable, h is always differentiable and $\nabla h(x) = A^*(I - \operatorname{prox}_{\lambda g})Ax$ (see, for example, [28]).

Algorithm 3.1 Take positive parameters δ, ξ and the real sequences $\{a_k\}$, $\{\delta_k\}$, $\{\beta_k\}$, $\{\epsilon_k\}$, $\{\rho_k\}$ satisfying the following conditions: for each $k \in \mathbb{N}$,

$$\begin{aligned}
&0 < a \le a_k \le b < 1, \qquad 0 < \xi \le \rho_k \le 4 - \xi, \\
&\delta_k > \delta > 0, \qquad \beta_k > 0, \qquad \epsilon_k \ge 0, \\
&\lim_{k\to+\infty} a_k = \frac{1}{2}, \\
&\sum_{k=1}^{\infty} \frac{\beta_k}{\delta_k} = +\infty, \qquad \sum_{k=1}^{\infty} \beta_k^2 = +\infty, \qquad \sum_{k=1}^{\infty} \frac{\beta_k \epsilon_k}{\delta_k} < +\infty.
\end{aligned}$$

Step 1. Choose $x^1 \in C$ and let $k := 1$.
Step k. Have $x^k \in C$ and take

$$\mu_k := \begin{cases} 0 & \text{if } \nabla h(x^k) = 0, \\ \rho_k \frac{h(x^k)}{\|\nabla h(x^k)\|^2} & \text{if } \nabla h(x^k) \neq 0, \end{cases}$$

then compute

$$y^k = x^k - \mu_k A^*(I - \operatorname{prox}_{\lambda g})Ax^k.$$

Take $\eta_k \in \partial_2^{\epsilon_k} f(y^k, y^k)$ and define

$$\alpha_k = \frac{\beta_k}{\gamma_k},$$

where $\gamma_k = \max\{\delta_k, \|\eta_k\|\}$. Compute

$$z^k = P_C\bigl(y^k - \alpha_k \eta_k\bigr). \tag{2}$$

Let

$$x^{k+1} = a_k x^k + (1 - a_k) z^k.$$

Remark 3.1 It is obvious that Algorithm 3.1 involves only one projection onto C per each iteration. Note that the domain of function f is $H \times H$.

Lemma 3.1 ([24]) *Let S be the set of solutions of Problem 1.1 and $y \in S$. If $\nabla h(x^k) \neq 0$, then*

$$\|y^k - y\|^2 \leq \|x^k - y\|^2 - \rho_k(1-\rho_k)\frac{h^2(x^k)}{\|\nabla h(x^k)\|^2}.$$

Lemma 3.2 ([29]) *For each $k \geq 1$, the following inequalities hold*:

(i) $\alpha_k\|\eta_k\| \leq \beta_k$;
(ii) $\|z^k - y^k\| \leq \beta_k$.

Lemma 3.3 *Let $y \in S$. Then, for each $k \geq 1$ such that $\nabla h(x^k) \neq 0$, we have*

$$\begin{aligned}\|x^{k+1} - y\|^2 \leq{}& \|x^k - y\|^2 - (1-a_k)\rho_k(4-\rho_k)\frac{h^2(x^k)}{\|\nabla h(x^k)\|^2}\\ &+ 2(1-a_k)\alpha_k f(y^k, y) + A_k,\end{aligned}$$

and, for each $k \geq 1$ such that $\nabla h(x^k) = 0$, we have

$$\|x^{k+1} - y\|^2 \leq \|x^k - y\|^2 + 2(1-a_k)\alpha_k f(y^k, y) + A_k,$$

where $A_k = 2(1-a_k)(\alpha_k\epsilon_k + \beta_k^2)$.

Proof By the definition of x^{k+1}, we have

$$\begin{aligned}\|x^{k+1} - y\|^2 &= \|a_k x^k + (1-a_k)z^k - y\|^2\\ &\leq a_k\|x^k - y\|^2 + (1-a_k)\|z^k - y\|^2.\end{aligned} \tag{3}$$

Moreover, we have

$$\begin{aligned}\|z^k - y\|^2 &= \|y - y^k + y^k - z^k\|^2\\ &= \|y^k - y\|^2 - \|y^k - z^k\|^2 + 2\langle y^k - z^k, y - z^k\rangle\\ &\leq \|y^k - y\|^2 + 2\langle y^k - z^k, y - z^k\rangle.\end{aligned}$$

Since it follows from Lemma 2.4 and (2) that

$$\langle z^k - y^k + \alpha_k\eta_k, x - z^k\rangle \geq 0$$

for all $x \in C$, by taking $x = y$, we obtain

$$\langle z^k - y^k + \alpha_k\eta_k, y - z^k\rangle \geq 0 \quad\Longleftrightarrow\quad \langle \alpha_k\eta_k, y - z^k\rangle \geq \langle y^k - z^k, y - z^k\rangle$$

and hence

$$\begin{aligned}\|z^k - y\|^2 &\leq \|y^k - y\|^2 + 2\langle \alpha_k\eta_k, y - z^k\rangle\\ &= \|y^k - y\|^2 + 2\langle \alpha_k\eta_k, y - y^k\rangle + 2\langle \alpha_k\eta_k, y^k - z^k\rangle.\end{aligned} \tag{4}$$

It follows from $\eta_k \in \partial_2^{\epsilon_k} f(y^k, y^k)$ that

$$\begin{aligned} f(y^k, y) - f(y^k, y^k) \\ \geq \langle \eta_k, y - y^k \rangle - \epsilon_k \quad \Longleftrightarrow \quad f(y^k, y) + \epsilon_k \geq \langle \eta_k, y - y^k \rangle. \end{aligned} \tag{5}$$

On the other hand, from Lemma 3.2(ii), it follows that

$$\langle \alpha_k \eta_k, y^k - z^k \rangle \leq \alpha_k \|\eta_k\| \|y^k - z^k\| \leq \beta_k^2.$$

From (4), (5) and $\alpha_k > 0$, it follows that

$$\|z^k - y\|^2 \leq \|y^k - y\|^2 + 2\alpha_k f(y^k, y) + 2\alpha_k \epsilon_k + 2\beta_k^2. \tag{6}$$

Now, we consider two cases:

Case 1. If $\nabla h(x^k) \neq 0$, then, thanks to Lemma 3.1, we have

$$\|y^k - y\|^2 = \|x^k - y\|^2 - \rho_k(4 - \rho_k) \frac{h^2(x^k)}{\|\nabla h(x^k)\|^2}.$$

Combining this inequality with (3) and (6), we obtain

$$\begin{aligned} \|x^{k+1} - y\|^2 \leq \|x^k - y\|^2 + 2(1 - a_k)\alpha_k f(y^k, y) \\ - (1 - a_k)\rho_k(4 - \rho_k) \frac{h^2(x_k)}{\|\nabla h(x_k)\|^2} + A_k, \end{aligned}$$

where $A_k = 2(1 - a_k)(\alpha_k \epsilon_k + \beta_k^2)$.

Case 2. If $\nabla h(y_k) = 0$, then, by the definition of y^k, we can write $y^k = x^k$. Now, by the same argument as in Case 1, we have

$$\|z^k - y\|^2 \leq \|y^k - y\|^2 + 2\alpha_k f(y^k, y) + 2\alpha_k \epsilon_k + 2\beta_k^2.$$

Then we have

$$\|x^{k+1} - y\|^2 \leq \|x^k - y\|^2 + 2(1 - a_k)\alpha_k f(y^k, y) + A_k,$$

where $A_k = 2(1 - a_k)(\alpha_k \epsilon_k + \beta_k^2)$. This completes the proof. □

Theorem 3.1 *Suppose that Problem* 1.1 *admits a solution. Then, under Assumptions* (A1)–(A4), *the sequence* $\{x^k\}$ *generated by Algorithm* 3.1 *strongly converges to a solution of Problem* 1.1.

Proof *Claim* 1. The sequence $\{\|x^k - y\|^2\}$ is convergent for all $y \in S$. Indeed, let $y \in S$. Since $y \in \mathrm{Sol}(EP)$ and f is pseudomonotone on C with respect to every solution of (EP), we have

$$f(y^k, y) \leq 0.$$

If $\nabla h(x^k) \neq 0$, then, since

$$\rho_k(4-\rho_k)\frac{h^2(x^k)}{\|\nabla h(x_k)\|^2} \geq 0,$$

it follows from Lemma 3.3 that

$$\|x^{k+1}-y\|^2 \leq \|x^k-y\|^2 + A_k,$$

where $A_k = 2(1-a_k)(\alpha_k\epsilon_k + \beta_k^2)$. Since $\alpha_k = \frac{\beta_k}{\gamma_k}$ with $\gamma_k = \max\{\delta_k, \|\eta_k\|\}$,

$$\sum_{k=1}^{+\infty}\alpha_k\epsilon_k = \sum_{k=1}^{+\infty}\frac{\beta_k}{\gamma_k}\epsilon_k \leq \sum_{k=1}^{+\infty}\frac{\beta_k}{\delta_k}\epsilon_k < +\infty.$$

Note that $\sum_{k=1}^{+\infty}\beta_k^2 < +\infty$ and $0 < a < a_k < b < 1$ and so we have

$$\sum_{k=1}^{+\infty}A_k < 2(1-a)\sum_{k=1}^{+\infty}(\alpha_k\epsilon_k + \beta_k^2) < +\infty.$$

Now, using Lemma 2.6, we see that $\{\|x^k-y\|^2\}$ is convergent for all $y \in S$. Hence the sequence $\{x^k\}$ is bounded. Then, by Lemma 3.2, we can see that $\{y^k\}$ is bounded too.

Claim 2. $\limsup_{k\to\infty} f(y^k, y) = 0$ for all $y \in S$. By Lemma 3.3, for each $k \geq 1$, we have

$$-2(1-a_k)\alpha_k f(y^k, y) \leq \|x^k-y\|^2 - \|x^{k+1}-y\|^2 + A_k.$$

Summing up both sides in the above inequality, we obtain

$$\sum_{k=1}^{\infty} -2(1-a_k)\alpha_k f(y^k, y) < +\infty.$$

On the other hand, using Assumption (A2) and the fact that $\{x^k\}$ is bounded, we see that $\{\|\eta_k\|\}$ is bounded. Then there exists $L > \delta$ such that $\|\eta_k\| \leq L$ for each $k \geq 1$. Therefore, we have

$$\frac{\gamma_k}{\delta_k} = \max\left\{1, \frac{\|\eta_k\|}{\delta + k}\right\} \leq \frac{L}{\delta}$$

and hence

$$\alpha_k = \frac{\beta_k}{\gamma_k} \geq \frac{\delta}{L}\frac{\beta_k}{\delta_k}.$$

Since y is a solution, by pseudomonotonicity of f, we have $-f(y^k, y) \geq 0$, which together with $0 < a < a_k < b < 1$ implies

$$\sum_{k=1}^{\infty}(1-b)\frac{\beta_k}{\delta_k}\left[-f(y^k, y)\right] < +\infty.$$

But, from $\sum_{k=1}^{\infty} \frac{\beta_k}{\delta_k} = +\infty$, it follows that

$$\limsup_{k\to+\infty} f\left(y^k, y\right) = 0$$

for all $y \in S$.

Claim 3. For any $y \in S$, suppose that $\{y^{k_j}\}$ is the subsequence of $\{y^k\}$ such that

$$\limsup_{k\to+\infty} f\left(y^k, y\right) = \lim_{j\to+\infty} f\left(y^{k_j}, y\right) \tag{7}$$

and y^* is a weakly cluster point of $\{y^{k_j}\}$. Then y^* belongs to $\mathrm{Sol}(EP)$.

Without loss of generality, we can assume that $\{y^{k_j}\}$ weakly converges to y^* as $j \to \infty$. Since $f(\cdot, y)$ is upper semi-continuous, by Claim 2, we have

$$f\left(y^*, y\right) \geq \limsup_{j\to+\infty} f\left(y^{k_j}, y\right) = 0.$$

Since $y \in S$ and f is pseudomonotone, we have $f(y^*, y) \leq 0$ and so $f(y^*, y) = 0$. Again, by pseudomonotonicity of f, $f(y, y^*) \leq 0$ and hence $f(y^*, y) = f(y, y^*) = 0$. Then, by paramonotonicity (Assumption (A3)), we can conclude that y^* is also a solution of (EP).

Claim 4. Every weakly cluster point $\bar{x}$ of the sequence $\{x^k\}$ satisfies $\bar{x} \in K$ and $A\bar{x} \in \operatorname{argmin} g$. Let $\bar{x}$ be a weakly cluster point of $\{x^k\}$ and $\{x^{k_j}\}$ be a subsequence of $\{x^k\}$ weakly converging to $\bar{x}$. Then $\bar{x} \in K$. From Lemma 3.3, if $\nabla h(x^k) \neq 0$, then we have

$$(1 - a_k)\rho_k(4 - \rho_k)\frac{h^2(x^k)}{\|\nabla h(x^k)\|^2} \leq \left\|x^k - z\right\|^2 - \left\|x^{k+1} - z\right\|^2 + A_k.$$

If $\nabla h(x^k) = 0$, then we have

$$0 \leq \left\|x^k - z\right\|^2 - \left\|x^{k+1} - z\right\|^2 + A_k.$$

Let $N_1 := \{k : \nabla h(x^k) \neq 0\}$. Summing up, we can write

$$\sum_{k\in N_1} (1 - a_k)\rho_k(4 - \rho_k)\frac{h^2(x^k)}{\|\nabla h(x^k)\|^2} \leq \left\|x^0 - z\right\|^2 + \sum_{k=1}^{\infty} A_k < +\infty.$$

Combining this fact with the assumption $\xi \leq \rho_k \leq 4 - \xi$ (for some $\xi > 0$) and $0 < a < a_k < b < 1$, we can conclude that

$$\sum_{k\in N_1} \frac{h^2(x^k)}{\|\nabla h(x^k)\|^2} < +\infty.$$

Moreover, since ∇h is Lipschitz continuous with constant $\|A\|^2$, we see that $\|\nabla h(x^k)\|^2$ is bounded. So, $h(x^k) \to 0$ as $k \in N_1$ and $k \to \infty$. Note that $h(x^k) = 0$ for $k \notin N_1$. Consequently, we have

$$\lim_{k\to+\infty} h\left(x^k\right) = 0. \tag{8}$$

By the lower semi-continuity of h,

$$0 \leq h(\bar{x}) \leq \liminf_{j\to+\infty} h\bigl(x^{k_j}\bigr) = \lim_{k\to+\infty} h\bigl(x^k\bigr) = 0,$$

which implies that $A\bar{x}$ is a fixed point of the proximal mapping of g. Thus $A\bar{x}$ is a minimizer of g. From (8) and the fact that $\|\nabla h(x^k)\|^2$ is bounded, it follows that

$$\lim_{k\to+\infty} \mu_k = 0,$$

which yields

$$\lim_{k\to+\infty} \bigl\|y^k - x^k\bigr\| = \lim_{k\to+\infty} \mu_k \bigl\|A^*(I - \operatorname{prox}_{\lambda g})\bigl(Ax^k\bigr)\bigr\| = 0.$$

Thus $\{y^{k_j}\}$ weakly converges to $\bar{x}$.

Claim 5. $\lim_{k\to+\infty} x^k = \lim_{k\to+\infty} y^k = \lim_{k\to+\infty} P(x^k) = x^*$, where x^* is a weakly cluster point of the sequence satisfying (7). From Claims 3 and 4, we can deduce that x^* belongs to S. By Claim 1, we can assume that

$$\lim_{k\to+\infty} \bigl\|x^k - x^*\bigr\| = c < +\infty.$$

By Lemma 3.2, we have

$$\begin{aligned}
\bigl\|z^k - x^*\bigr\| &\leq \bigl\|y^k - x^*\bigr\| + \bigl\|z^k - y^k\bigr\| \\
&\leq \bigl\|x^k - x^*\bigr\| + \bigl\|x^k - y^k\bigr\| + \beta_k,
\end{aligned}$$

which implies that

$$\limsup_{k\to+\infty} \bigl\|z^k - x^*\bigr\| \leq \limsup_{k\to+\infty} \bigl(\bigl\|x^k - x^*\bigr\| + \bigl\|x^k - y^k\bigr\| + \beta_k\bigr) = c.$$

On the other hand, we have

$$\lim_{k\to+\infty} \bigl\|a_k\bigl(x^k - x^*\bigr) + (1 - a_k)\bigl(z^k - x^*\bigr)\bigr\| = \lim_{k\to+\infty} \bigl\|x^{k+1} - x^*\bigr\| = c.$$

By applying Lemma 2.7 with $v^k := x^k - x^*$, $w^k := z^k - x^*$, we obtain

$$\lim_{k\to+\infty} \bigl\|z^k - x^k\bigr\| = 0.$$

Employing arguments, similar to those used in the proof of Theorem 1 in [37], we have

$$\lim_{k\to+\infty} x^k = x^*.$$

This completes the proof. □

4 Algorithms and convergence analysis

In [37], Yen et al. presented an application of Problem 1.1 to a model of electricity production, in which z denotes the quantity of the materials and $g(z)$ is the total environmental fee that companies have to pay for environmental pollution while using materials z for production. So, from $x \in C$, it follows that $z = Ax \in \{z : z = Ax, x \in C\}$.

However, in actual production, since the resources are limited, there are usually stricter constraints on the quantity of the materials such as $z \in Q$, where Q is a nonempty closed convex set of H_2. Therefore, it is necessary to replace the unconstrained convex optimization problem $\min_{x\in H_2} g(x)$ with the *constrained convex optimization* as follows:

$$\min_{x\in Q} g(x), \tag{9}$$

whose solution is denoted by $\mathrm{Sol}(Q,g)$. By using (9), Problem 1.1 becomes the following problem:

Problem 4.1 Find $x^* \in C$ such that $f(x^*, y) \geq 0$ for all $y \in C$ such that $Ax^* \in Q$ and $g(Ax^*) \leq g(z)$ for all $z \in Q$,

whose solution is denoted by

$$\Gamma := \Gamma(C, Q, f, g, A) := \bigl\{z \in \mathrm{Sol}(EP) : Az \in \mathrm{Sol}(Q,g)\bigr\}.$$

Throughout this paper, we assume $\Gamma \neq \emptyset$.

In this section, we discuss two cases that the function g is differentiable or non-differentiable. The corresponding algorithms and their convergence are provided next.

4.1 The case when *g* is differentiable

We need to make the following assumption on the mapping g:

(B) g is L-Lipschitz differentiable with $L > 0$, i.e.,

$$\bigl\|\nabla g(x) - \nabla g(y)\bigr\| \leq L\|x - y\|$$

for all $x, y \in H_2$.

It is easy to verify that the constrained convex optimization problem (9) is equivalent to the following *variational inequality problem*:

$$\bigl\langle \nabla g\bigl(z^*\bigr), z - z^*\bigr\rangle \geq 0 \quad \text{for all } z \in Q, \tag{10}$$

and the variational inequality problem (10) is equivalent to the following *fixed point problem*:

$$z^* = P_Q(I - \nu\nabla g)z^*, \tag{11}$$

where $\nu > 0$. So, the constrained optimization problem (9) and the fixed point problem (11) are equivalent. From the optimality condition of (9), we can also deduce the equivalence of problems (9) and (11) (see [24]). Next we construct an iterative algorithm based on this equivalence and Algorithm 3.1.

Firstly, define two functions

$$h(x) := \left\| \left(I - P_Q(I - \nu\nabla g)\right)Ax \right\|^2 \tag{12}$$

and

$$l(x) := \left\| A^*\left(I - P_Q(I - \nu\nabla g)\right)Ax \right\|^2,$$

where $\nu \in (0, \frac{2}{L})$.

Algorithm 4.1 Take the real sequences $\{a_k\}$, $\{\delta_k\}$, $\{\beta_k\}$, $\{\epsilon_k\}$ and $\{\rho_k\}$ as in Algorithm 3.1.

Step 1. Choose $x^1 \in C$ and let $k := 1$.

Step k. Have $x^k \in C$ and take

$$\mu_k := \begin{cases} 0 & \text{if } l(x^k) = 0, \\ \rho_k \frac{h(x^k)}{l(x^k)} & \text{if } l(x^k) \neq 0, \end{cases}$$

then compute

$$y^k = x^k - \mu_k A^*\left(I - P_Q(I - \nu\nabla g)\right)\left(Ax^k\right). \tag{13}$$

Take $\eta_k \in \partial_2^{\epsilon_k} f(y^k, y^k)$ and define

$$\alpha_k = \frac{\beta_k}{\gamma_k},$$

where $\gamma_k = \max\{\delta_k, \|\eta_k\|\}$. Compute

$$z^k = P_C\left(y^k - \alpha_k \eta_k\right).$$

Let

$$x^{k+1} = a_k x^k + (1 - a_k) z^k.$$

Now, we need the following lemmas to prove the convergence of Algorithm 4.1:

Lemma 4.1 ([8, Lemma 6.2]) *Assume that a mapping* $g : H_2 \to H_2$ *satisfies Assumption* (B) *and* $\nu \in (0, \frac{2}{L})$. *Let* $y \in \Gamma$. *If* $\|l(x^k)\| \neq 0$, *then it follows that*

$$\left\|y^k - y\right\|^2 \le \left\|x^k - y\right\|^2 - \rho_k(1 - \rho_k)\frac{h^2(x^k)}{l(x^k)}.$$

Proof Let $T = P_Q(I - \nu\nabla g)$. Since $y \in \Gamma$, it follows from (11) that Ay is a fixed point of T. From the proof of [32, Theorem 4.1], it follows that T is $\frac{2+\nu L}{4}$-averaged and so it is nonexpansive. By (13) and Lemma 2.3(i), we have

$$\left\|y^k - y\right\|^2 \le \left\|x^k - y\right\|^2 + \mu_k^2 \left\|A^*(I - T)\left(Ax^k\right)\right\|^2$$

$$- 2\mu_k\langle x^k - y, A^*(I - T)(Ax^k)\rangle. \tag{14}$$

By the nonexpansivity of T and (2), we have

$$\begin{aligned}
&\langle x^k - y, A^*(I - T)(Ax^k)\rangle \\
&\quad = \langle A(x^k - y), (I - T)(Ax^k)\rangle \\
&\quad = \langle A(x^k - y) - (I - T)(Ax^k) + (I - T)(Ax^k), (I - T)(Ax^k)\rangle \\
&\quad = \langle T(Ax^k) - Ay, Ax^k - T(Ax^k)\rangle + \|(I - T)(Ax^k)\|^2 \\
&\quad \geq \frac{1}{2}\|(I - T)(Ax^k)\|^2.
\end{aligned} \tag{15}$$

Combining (14) and (15) and using the definitions of $h(x)$ and $l(x)$, we obtain

$$\begin{aligned}
\|y^k - y\|^2 &\leq \|x^k - y\|^2 + \mu_k^2 l(x^k) - \mu_k h(x^k) \\
&= \|x^k - y\|^2 - \rho_k(1 - \rho_k)\frac{h^2(x^k)}{l(x^k)}.
\end{aligned}$$

This completes the proof. □

Remark 4.1 From (15), it follows that $l(x) = 0$ implies $h(x) = 0$.

Using Lemma 4.1 and following the lines of the proof of Lemma 3.3, we have the following:

Lemma 4.2 *Let $y \in \Gamma$. Then, for each $k \geq 1$ such that $l(x^k) \neq 0$, we have*

$$\begin{aligned}
\|x^{k+1} - y\|^2 \leq{}& \|x^k - y\|^2 - \rho_k(1 - \rho_k)\frac{h^2(x^k)}{l(x^k)} \\
&+ 2(1 - a_k)\alpha_k f(y^k, y) + A_k
\end{aligned}$$

and, for each $k \geq 1$ such that $l(x^k) = 0$, we have

$$\|x^{k+1} - y\|^2 \leq \|x^k - y\|^2 + 2(1 - a_k)\alpha_k f(y^k, y) + A_k,$$

where $A_k = (1 - a_k)(\alpha_k\epsilon_k + \beta_k^2)$.

Next we establish the convergence of Algorithm 4.1.

Theorem 4.1 *Under Assumptions* (A1)–(A4) *and* (B), *the sequence $\{x^k\}$ generated by Algorithm 4.1 strongly converges to a solution of Problem 4.1.*

The proof of Theorem 4.1 is similar with that of Theorem 3.1, so here we omit it.

The only thing to note about the proof of Theorem 4.1 is that from $h(\bar{x}) = 0$ it follows that $A\bar{x}$ is a fixed point of $P_Q(I - \nu\nabla g)$. Thus $A\bar{x}$ is a solution of (9).

4.2 The case when *g* is non-differentiable

Let $g : Q \to \mathbb{R} \cup \{+\infty\}$ be a proper convex lower semi-continuous function. Denote by

$$g_\lambda(z) := \min_{u\in Q}\left\{g(u) + \frac{1}{2\lambda}\|u - y\|^2\right\} \tag{16}$$

the *Moreau–Yosida approximate* of the function g with the parameter λ. It is easy to see that the solution of (16) converges to that of $\min_{x\in Q} g(x)$ as $\lambda \to \infty$.

For the mapping defined in (16), we have following result.

Lemma 4.3 *The constrained optimization problem*

$$\min_{x\in Q} g_\lambda(x) \tag{17}$$

is equivalent to the fixed point formulation

$$x^* = P_Q\bigl(x^* - \nu(I - \operatorname{prox}_{\lambda g})x^*\bigr), \tag{18}$$

where $\nu \in (0, +\infty)$.

Proof It is well known that problem (17) is equivalent to the following problem:

$$\min_{x\in H_2}\bigl\{\iota_Q(x) + g_\lambda(x)\bigr\}. \tag{19}$$

Note that the differentiability of the Yosida-approximate g_λ (see, for instance, [28]) secures the additivity of the subdifferentials, and so we can write

$$\partial\bigl(\iota_Q(x) + g_\lambda(x)\bigr) = \partial\iota_Q(x) + \frac{I - \operatorname{prox}_{\lambda g}}{\lambda}(x).$$

The optimality condition of (19) can be then written as follows:

$$0 \in \lambda\partial\iota_Q(x) + (I - \operatorname{prox}_{\lambda g})(x), \tag{20}$$

where the subdifferential of ι_C at x is $N_C(x)$. The inclusion (20) in turn yields (18). This completes the proof. □

Set

$$h(x) = \bigl\|\bigl(I - P_Q\bigl(I - \nu(I - \operatorname{prox}_{\lambda g})\bigr)\bigr)Ax\bigr\|^2$$

and

$$l(x) = \bigl\|A^*\bigl(I - P_Q\bigl(I - \nu(I - \operatorname{prox}_{\lambda g})\bigr)\bigr)Ax\bigr\|^2.$$

Similar to Algorithm 3.1, using Lemma 4.3, we introduce the following algorithm:

Algorithm 4.2 Take the real sequences $\{a_k\}$, $\{\delta_k\}$, $\{\beta_k\}$, $\{\epsilon_k\}$ and $\{\rho_k\}$ as in Algorithm 3.1. Take a positive parameter ν.

Step 1. Choose $x^1 \in C$ and let $k := 1$.

Step k. Have $x^k \in C$ and take

$$\mu_k := \begin{cases} 0 & \text{if } l(x^k) = 0, \\ \rho_k \frac{h(x^k)}{l(x^k)} & \text{if } l(x^k) \neq 0, \end{cases}$$

then compute

$$y^k = x^k - \mu_k \bigl(A^*\bigl(I - P_Q\bigl(I - \nu(I - \operatorname{prox}_{\lambda g})\bigr)\bigr)\bigl(Ax^k\bigr)\bigr). \tag{21}$$

Take $\eta_k \in \partial_2^{\epsilon_k} f(y^k, y^k)$ and define

$$\alpha_k = \frac{\beta_k}{\gamma_k},$$

where $\gamma_k = \max\{\delta_k, \|\eta_k\|\}$. Compute

$$z^k = P_C\bigl(y^k - \alpha_k \eta_k\bigr).$$

Let

$$x^{k+1} = a_k x^k + (1 - a_k) z^k.$$

Remark 4.2 Let $\nu = 1$ in Algorithm 4.2, then formula (21) becomes

$$y^k = x^k - \mu_k \bigl(A^*(I - P_Q \circ \operatorname{prox}_{\lambda g})\bigl(Ax^k\bigr)\bigr),$$

which yields

$$y^k = x^k - \mu_k A^*(I - \operatorname{prox}_{\lambda g})\bigl(Ax^k\bigr),$$

when $Q = H_2$. So, Algorithm 3.1 is a special case of Algorithm 4.2.

We need the following lemmas for the proof of the convergence of Algorithm 4.2.

Lemma 4.4 *Let* $\nu \in (0, 1]$. *Then operator* $P_Q(I - \nu(I - \operatorname{prox}_{\lambda g}))$ *is nonexpansive.*

Proof By the fact that $\operatorname{prox}_{\lambda g}$ is firmly nonexpansive and Lemma 2.1, $I - \operatorname{prox}_{\lambda g}$ and $\nu(I - \operatorname{prox}_{\lambda g})$ are also firmly nonexpansive. So, using Lemma 2.1 again, $I - \nu(I - \operatorname{prox}_{\lambda g})$ is firmly nonexpansive. Thus, from Lemma 2.2, it follows that $P_Q(I - \nu(I - \operatorname{prox}_{\lambda g}))$ is $\frac{3}{4}$-averaged and hence nonexpansive. This completes the proof. □

Using Lemma 4.4 and following the proof of Theorem 4.1, we obtain the convergence result of Algorithm 4.2.

Theorem 4.2 *Let* $\nu \in (0,1]$. *Then, under Assumptions* (A1)–(A4), *the sequence* $\{x^k\}$ *generated by Algorithm* 4.2 *strongly converges to a solution of Problem* 4.1.

The proof of Theorem 4.2 is similar to that of Theorem 3.1, so here we omit it.

One thing to note about proof of Theorem 4.2 is that from $h(\bar{x}) = 0$ it follows that $A\bar{x}$ is a fixed point of $P_Q(I - \nu(I - \mathrm{prox}_{\lambda g}))$. Thus, by Lemma 4.3, $A\bar{x}$ is a solution of (17).

5 Numerical examples

In this section, we provide two numerical examples to compare different algorithms. All programs are written in Matlab version 7.0 and performed on a desktop PC with Intel(R) Core(TM) i5-4200U CPU @ 2.30 GHz, RAM 4.00 GB.

Example 5.1 First, we consider an equilibrium-optimization model which was investigated by Yen et al. [37]. This model can be regarded as an extension of a Nash–Cournot oligopolistic equilibrium model in electricity markets. The latter model has been investigated in some research papers (see, for example, [10, 27]).

In this equilibrium model, it is assumed that there are n companies. Let x denote the vector whose entry x_i stands for the power generated by company i. Following Contreras et al. [10], we suppose that the price $p_i(s)$ is a decreasing affine function of s with $s = \sum_{i=1}^n x_i$, that is,

$$p_i(s) = \alpha - \beta_i s.$$

Then the profit made by company i is given by

$$f_i(x) = p_i(s)x_i - c_i(x_i),$$

where $c_i(x_i)$ is the cost for generating x_i by the company i.

Suppose that C_i is the strategy set of company i, that is, condition $x_i \in C_i$ must be satisfied for each i. Then the strategy set of the model is $C := C_1 \times C_2 \times \cdots \times C_n$.

Actually, each company seeks to maximize its profit by choosing the corresponding production level under the presumption that the production of the other companies are parametric input. A commonly used approach to this model is based upon the famous Nash equilibrium concept.

Now, we recall that a point $x^* \in C = C_1 \times C_2 \times \cdots \times C_n$ is an *equilibrium point* of the model if

$$f_i(x^*) \geq f_i(x^*[x_i])$$

for all $x_i \in C_i$ and $i = 1, 2, \ldots, n$, where $x^*[x_i]$ stands for the vector obtained from x^* by replacing x_i^* with x_i. By taking

$$f(x,y) := \Psi(x,y) - \Psi(x,x)$$

with

$$\Psi(x,y) := -\sum_{i=1}^n f_i(x[y_i]), \tag{22}$$

the problem of finding a Nash equilibrium point of the model can be formulated as follows:

$$\text{Find } x^* \in C \text{ such that } f\left(x^*, x\right) \geq 0 \text{ for all } x \in C. \tag{EP}$$

In [37], Yen et al. extended this equilibrium model by additionally assuming that the companies use some materials to produce electricity.

Let $a_{l,i}$ denote the quantity of material l ($l = 1, \ldots, m$) for producing one unit of electricity by company i ($i = 1, \ldots, n$). Let A be the matrix whose entries are $a_{l,i}$. Then entry l of the vector Ax is the quantity of material l for producing x. Using materials for production may cause environmental pollution, for which companies have to pay a fee. Suppose that $g(Ax)$ is the total environmental fee for producing x.

The task now is to find a production x^* such that it is a Nash equilibrium point with a minimum environmental fee, while the quantity of the materials satisfies constraint Q. This problem can be formulated as the *split feasibility problem* of the following form:

$$\begin{aligned} &\text{Find } x^* \in C \text{ such that } f\left(x^*, x\right) \geq 0 \text{ for all } x \in C \\ &\quad \text{and } g\left(Ax^*\right) \leq g(Ax) \text{ for all } Ax \in Q. \end{aligned} \tag{SEP}$$

Suppose that, for every i, cost c_i for production and environmental fee g are increasing convex functions. The convexity assumption here means that both the cost and fee for producing a unit of product increase as the quantity of the product gets larger.

Under this convexity assumption, it is not hard to see (see also Quoc et al. [27]) that problem (EP) with f given by (22) can be formulated as follows:

$$\text{Find } x^* \in C \text{ such that } f\left(x^*, x\right) := \left\langle \tilde{B}_1 x^* - \bar{a}, x - x^* \right\rangle + \varphi(x) - \varphi\left(x^*\right) \geq 0 \text{ for all } x \in C, \tag{23}$$

where

$$\begin{aligned} &\bar{a} := (\alpha, \alpha, \ldots, \alpha)^T, \\ &B_1 := \begin{pmatrix} b_1 & 0 & 0 & \ldots & 0 \\ 0 & b_2 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & b_n \end{pmatrix}, \qquad \tilde{B}_1 := \begin{pmatrix} 0 & b_1 & b_1 & \ldots & b_1 \\ b_2 & 0 & b_2 & \ldots & b_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ b_n & b_n & b_n & \ldots & 0 \end{pmatrix}, \\ &\varphi(x) := x^T B_1 x + \sum_{i=1}^{n} c_i(x_i). \end{aligned} \tag{24}$$

Note that, when c_i is differentiable and convex for each i, problem (23) is equivalent to the following *variational inequality problem*:

$$\text{Find } x^* \in C \text{ such that } \left\langle \tilde{B}_1 x^* - \bar{a} + \nabla\varphi\left(x^*\right), x - x^* \right\rangle \geq 0 \text{ for all } x \in C.$$

We tested the proposed algorithm with the cost function given by

$$c_i(x_i) = \frac{1}{2} p_i x_i^2 + q_i x_i, \tag{25}$$

where $p_i \geq 0$. In [37], Yen et al. showed that function $f(x,y)$ defined by (23), (24) and (25) satisfies Assumptions (A1), (A2) and (A4).

In [37], the author denoted by $g(z)$ the total environmental fee. It is unreasonable. Firstly, the total environmental fee should be included in the cost, that is, it is a part of $c_i(x_i)$. Secondly, it is supposed that the companies behave as players in an oligopolistic market, but at the same time they are subordinated to the centralized planning decision in order to minimize the total environmental fee for the whole system. That is, the model is not concordant with the real system behavior.

It may be reasonable to denote by $g(z)$ the restriction for the emission of contaminants. To protect the environment, governments generally adopt policies to restrict emissions of contaminants.

Assume that the production of electricity brings p contaminants and governments require that the quantity of contaminants brought by the production of one unit of electricity is in a given region. We use a set $K \subset \mathbb{R}^p$ to denote this region.

Let $b_{k,l}$ denote the quantity of the contaminant k ($k = 1,\dots,p$) for consuming one unit of material l ($l = 1,\dots,m$). Let B be the matrix whose entries are $b_{k,l}$. Then entry k of vector Bz is the quantity of contaminant k for consuming one unit of material z_l ($l = 1,\dots,m$). So, the quantity of contaminant k ($k = 1,\dots,p$) for producing one unit of electricity is entry k of BAx, and BAx should be in the set K, i.e., $BAx \in K$. We get $Bz \in K$ when letting $z = Ax$.

Therefore we define function $g(z)$ as follows:

$$g(z) = \frac{1}{2}\left\| Bz - P_K(Bz)\right\|^2, \tag{26}$$

which is differentiable and $\nabla g(z) = B^T(I - P_K)(Bz)$ (see, e.g., [6]).

Take the sequences $\{\beta_k\}$, $\{\epsilon_k\}$, $\{\delta_k\}$ of the parameters as follows:

$$\beta_k = \frac{4}{k+1}, \qquad \epsilon_k = 0, \qquad \delta_k = 3, \qquad \gamma_k = \max\left\{3, \|\eta_k\|\right\}$$

for each $k \geq 1$ and take $\nu = \frac{1.99}{\|B\|}$. The entries of matrix A were randomly generated in the interval $[0,5]$. In the bifunction $f(x,y)$ defined by (23), (24) and (25), the parameters $\alpha = 0.5$ and b_i, p_i and q_i for each $i = 1,\dots,n$ were generated randomly in the interval $(0,1]$, $[1,3]$, and $[1,3]$, respectively. In the function $g(z)$, we take $B \in \mathbb{R}^p \times \mathbb{R}^m$, and its elements are generated randomly in $(0,1)$.

Since function $g(z)$ is differentiable, we use Algorithm 4.1 to solve Problem 4.1 and compare it with Algorithms 1.1 and 3.1. In Algorithms 1.1 and 3.1 we substitute $\mathrm{prox}_{\lambda g}$ with $I - \nu\nabla g$ and do not consider the constraint set Q.

The computational results are shown in Figs. 1 and 2. The horizontal and vertical axes show iteration k, as well as $\mathrm{error1}(k) := \|x^k - x^{k-1}\|$ and $\mathrm{error2}(k) := \|Ax^k - P_Q(Ay^k)\|$, respectively. We solve the model with $m = 15$ and take $n = 10$ as the number of companies.

From Figs. 1 and 2, we have two conclusions as follows:

(a) The "error1" of Algorithm 4.1 is smaller than that of Algorithms 1.1 and 3.1 and the "error1" of Algorithm 3.1 is slightly smaller than that of Algorithm 1.1.

(b) The "error2" of Algorithm 4.1 decreases with the iteration number k, while the "error2" of Algorithms 1.1 and 3.1 increases with the iteration number k. The "error2" of Algorithm 4.1 is smaller than those of Algorithms 1.1 and 3.1.

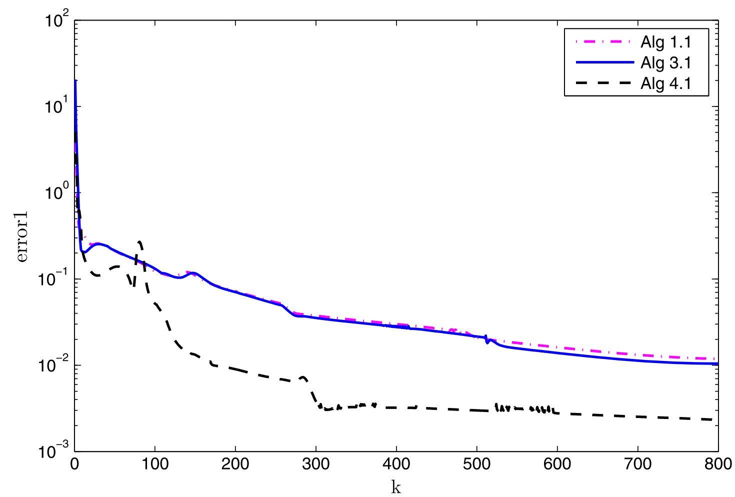

Figure 1 Comparison of Algorithms 1.1, 3.1 and 4.1

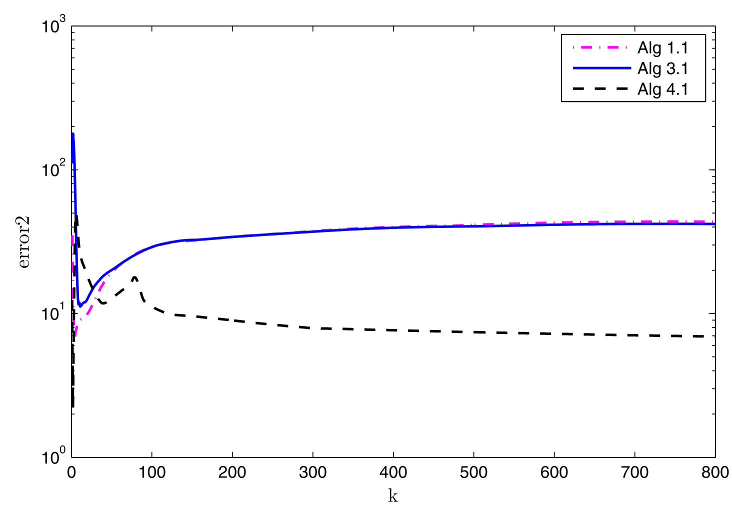

Figure 2 Comparison of Algorithms 1.1, 3.1 and 4.1

Next we give a numerical procedure in an infinite-dimensional space and compare Algorithm 4.1 with a numerical algorithm which is based on the Halpern modification of [8, Algorithm 6.1] as follows:

Algorithm 5.1

$$x^{k+1} = \tau_k x^1 + (1 - \tau_k) U\big(x^k + \gamma A^*(T - I)\big(Ax^k\big)\big),$$

where $T := P_Q(I - \lambda\nabla g)$, $U := P_C(I - \lambda f)$, and $\gamma \in (0, 1/L)$, L is the spectral radius of the operators A^*A, denoted by $\rho(A^*A)$. The parameter λ depends on the constants of the inverse strong monotonicity of ∇g and f.

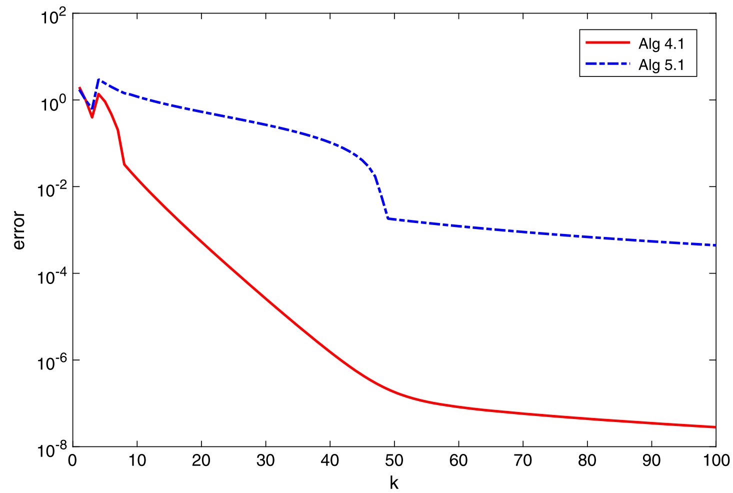

Figure 3 Comparison of Algorithms 4.1 and 5.1

According to the condition of the convergence of Halpern-type algorithm, we assume that $\lim_{k\to\infty} \tau_k = 0$ and $\sum_{k=1}^{\infty} \tau_k = \infty$.

Example 5.2 Suppose that $H = L^2([0,1])$ with norm $\|x\| := (\int_0^1 |x(t)|^2\,dt)^{\frac{1}{2}}$ and inner product $\langle x,y\rangle := \int_0^1 x(t)y(t)\,dt$, $x,y \in H$. Let $C := \{x \in H : \|x\| \leq 1\}$ be the unit ball, $Q := \{x \in H : \langle x(t), \sin(10x(t))\rangle \leq 1\}$. Define an operator $F : C \to H$ by

$$F(x)(t) = \int_0^1 \bigl(x(t) - B(t,s)p\bigl(x(s)\bigr)\bigr)\,ds + q(t)$$

for all $x \in C$ and $t \in [0,1]$, where

$$B(t,s) = \frac{2tse^{t+s}}{e\sqrt{e^2-1}}, \qquad p(x) = \cos x, \qquad q(t) = \frac{2te^t}{e\sqrt{e^2-1}}.$$

As shown in [30], F is monotone and L-Lipschitz-continuous with $L = 2$. Let $f(x(t), y(t)) = \langle Fx(t), y(t) - x(t)\rangle$, $g(x)(t) = \frac{1}{2}\|x(t)\|^2$ and $(Ax)(t) = 3x(t)$ for all $x \in H$.

Let $x^1(t) = 1$. Take the sequences $\{\alpha_k\}$, $\{\beta_k\}$, $\{\epsilon_k\}$, $\{\delta_k\}$ of the parameters as follows:

$$\alpha_k = \frac{1}{2}, \qquad \beta_k = \frac{4}{k+1}, \qquad \epsilon_k = 0, \qquad \delta_k = 3, \qquad \gamma_k = \max\bigl\{3, \|\eta_k\|\bigr\}$$

for each $k \geq 1$ and take $\nu = \frac{1.99}{L_g}$. We take $\lambda = 1$ according to the numerical tests since the constants of the inverse strong monotonicity of ∇g and f are unknown. Take $\tau_k = \frac{1}{k+1}$ and $\gamma = \frac{0.9}{\rho(A^*A)}$ for Algorithm 5.1. We use error $= \frac{1}{2}\|P_C(x^k) - x^k\|^2 + \frac{1}{2}\|P_Q(Ax^k) - Ax^k\|^2$ to measure the error of the kth iteration.

Numerical results are given in Fig. 3, which illustrate that Algorithm 4.1 behaves better than Algorithm 5.1.

6 Conclusions

We first introduce a new algorithm, which involves a projection of each iteration, and show its strong convergence. We also improve the model proposed in [37] by adding a constraint to the minimization problem of the total environmental fee. Two algorithms are introduced to approximate the solution and their strong convergence is analyzed.

Acknowledgements
We sincerely thank Prof. S. He for his helpful discussion and the reviewers for their valuable suggestions and useful comments that have led to the present improved version of the original manuscript.

Funding
The first author was supported by the scientific research project of Tianjin Municipal Education Commission (No. 2018KJ253). The fifth author was supported by the Theoretical and Computational Science (TaCS) Center under Computational and Applied Science for Smart Innovation Cluster (CLASSIC), Faculty of Science, KMUTT. The authors acknowledge the financial support provided by King Mongkut's University of Technology Thonburi through the "KMUTT 55th Anniversary Commemorative Fund". Furthermore, Poom Kumam was supported by he Thailand Research Fund (TRF) and the King Mongkut's University of Technology Thonburi (KMUTT) under the TRF Research Scholar Award (Grant No. RSA6080047).

Availability of data and materials
Data sharing not applicable to this article as no datasets were generated during the current study.

Competing interests
The authors declare that they have no competing interests.

Authors' contributions
All authors contributed equally to the writing of this paper. All authors read and approved the final manuscript.

Author details
[1]Tianjin Key Laboratory for Advanced Signal Processing, College of Science, Civil Aviation University of China, Tianjin, P.R. China. [2]KMUTTFixed Point Research Laboratory, Department of Mathematics, Faculty of Science, King Mongkut's University of Technology Thonburi (KMUTT), Bangkok, Thailand. [3]Department of Mathematics Education, Gyeongsang National University, Jinju, Korea. [4]School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu, P.R. China. [5]Department of Medical Research, China Medical University Hospital, China Medical University, Taichung, Taiwan.

Publisher's Note

Received: 17 December 2018 Accepted: 18 March 2019 Published online: 27 March 2019

References

1. Bauschke, H.H., Combettes, P.L.: Convex Analysis and Monotone Operator Theory in Hilbert Spaces. Springer, Berlin (2011)
2. Blum, E., Oettli, W.: From optimization and variational inequalities to equilibrium problems. Math. Stud. **63**, 123–145 (1994)
3. Byrne, C.: A unified treatment of some iterative algorithms in signal processing and image reconstruction. Inverse Probl. **20**, 103–120 (2004)
4. Ceng, L.C.: Approximation of common solutions of a split inclusion problem and a fixed-point problem. J. Appl. Numer. Optim. **1**, 1–12 (2019)
5. 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)
6. 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)
7. Censor, Y., Elving, T.: A multiprojections algorithm using Bregman projections in a product spaces. Numer. Algorithms **8**, 221–239 (1994)
8. Censor, Y., Gibali, A., Reich, S.: Algorithms for the split variational inequality problem. Numer. Algorithms **59**, 301–323 (2012)
9. Combettes, P.L.: Solving monotone inclusions via compositions of nonexpansive averaged operators. Optimization **53**, 475–504 (2004)
10. Contreras, J., Klusch, M., Krawczyk, J.B.: Numerical solution to Nash–Cournot equilibria in coupled constraint electricity markets. IEEE Trans. Power Syst. **19**, 195–206 (2004)
11. Crombez, G.: A geometrical look at iterative methods for operators with fixed points. Numer. Funct. Anal. Optim. **26**, 157–175 (2005)
12. Crombez, G.: A hierarchical presentation of operators with fixed points on Hilbert spaces. Numer. Funct. Anal. Optim. **27**, 259–277 (2006)
13. Dong, Q.L., Cho, Y.J., Zhong, L.L., Rassias, T.M.: Inertial projection and contraction algorithms for variational inequalities. J. Glob. Optim. **70**, 687–704 (2018)
14. Dong, Q.L., He, S., Zhao, J.: Solving the split equality problem without prior knowledge of operator norms. Optimization **64**, 1887–1906 (2015)
15. Dong, Q.L., Lu, Y.Y., Yang, J.: The extragradient algorithm with inertial effects for solving the variational inequality. Optimization **65**, 2217–2226 (2016)

16. Dong, Q.L., Tang, Y.C., Cho, Y.J., Rassias, T.M.: "Optimal" choice of the step length of the projection and contraction methods for solving the split feasibility problem. J. Glob. Optim. **71**, 341–360 (2018)
17. Dong, Q.L., Yao, Y., He, S.: Weak convergence theorems of the modified relaxed projection algorithms for the split feasibility problem in Hilbert spaces. Optim. Lett. **8**, 1031–1046 (2014)
18. Dong, Q.L., Yuan, H.B., Cho, Y.J., Rassias, T.M.: Modified inertial Mann algorithm and inertial CQ-algorithm for nonexpansive mappings. Optim. Lett. **12**, 87–102 (2018)
19. Fan, K.: Fixed point and minimax theorems in locally convex topological linear spaces. Proc. Natl. Acad. Sci. USA **38**, 121–126 (1952)
20. He, S.N., Tian, H.L.: Selective projection methods for solving a class of variational inequalities. Numer. Algorithms **80**, 617–634 (2019)
21. He, S.N., Tian, H.L., Xu, H.K.: The selective projection method for convex feasibility and split feasibility problems. J. Nonlinear Convex Anal. **19**, 1199–1215 (2018)
22. Konnov, I.V.: Combined Relaxation Methods for Variational Inequalities. Springer, Berlin (2000)
23. Konnov, I.V.: The method of pairwise variations with tolerances for linearly constrained optimization problems. J. Nonlinear Var. Anal. **1**, 25–41 (2017)
24. Moudafi, A., Thakur, B.S.: Solving proximal split feasibility problems without prior knowledge of operator norms. Optim. Lett. **8**, 2099–2110 (2014)
25. Muu, L.D., Oettli, W.: Convergence of an adaptive penalty scheme for finding constrained equilibria. Nonlinear Anal. **18**, 1159–1166 (1992)
26. Parikh, N., Boyd, S.: Proximal algorithms. Found. Trends Optim. **1**, 123–231 (2013)
27. Quoc, T.D., Muu, L.D.: Iterative methods for solving monotone equilibrium problems via dual gap functions. Comput. Optim. Appl. **51**, 709–728 (2012)
28. Rockafellar, T.R., Wets, R.: Variational Analysis. Springer, Berlin (1998)
29. Santos, P., Scheimberg, S.: An inexact subgradient algorithm for equilibrium problems. Comput. Appl. Math. **30**, 91–107 (2011)
30. Shehu, Y., Dong, Q.L., Jiang, D.: Single projection method for pseudo-monotone variational inequality in Hilbert spaces. Optimization **68**, 385–409 (2019)
31. Xiao, Y.B., Huang, N.J., Cho, Y.J.: A class of generalized evolution variational inequalities in Banach spaces. Appl. Math. Lett. **25**, 914–920 (2012)
32. Xu, H.K.: Averaged mappings and the gradient-projection algorithm. J. Optim. Theory Appl. **150**, 360–378 (2011)
33. Yao, Y., Leng, L., Postolache, M., Zheng, X.: Mann-type iteration method for solving the split common fixed point problem. J. Nonlinear Convex Anal. **18**, 875–882 (2017)
34. Yao, Y., Liou, Y.C., Yao, J.C.: Iterative algorithms for the split variational inequality and fixed point problems under nonlinear transformations. J. Nonlinear Sci. Appl. **10**, 843–854 (2017)
35. Yao, Y., Yao, J.C., Liou, Y.C., Postolache, M.: Iterative algorithms for split common fixed points of demicontractive operators without priori knowledge of operator norms. Carpath. J. Math. **34**, 459–466 (2018)
36. Yao, Y.H., Postolache, M., Liou, Y.C.: Strong convergence of a self-adaptive method for the split feasibility problem. Fixed Point Theory Appl. **2013**, Article ID 201 (2013)
37. Yen, L.H., Muu, L.D., Huyen, N.T.T.: An algorithm for a class of split feasibility problems: application to a model in electricity production. Math. Methods Oper. Res. **84**, 549–565 (2016)
38. Zhao, J.: Solving split equality fixed-point problem of quasi-nonexpansive mappings without prior knowledge of operators norms. Optimization **64**, 2619–2630 (2015)
39. Zhao, J., Zong, H.: Iterative algorithms for solving the split feasibility problem in Hilbert spaces. J. Fixed Point Theory Appl. **21**, 11 (2018)