Fixed Point Theory and Applications
a SpringerOpen Journal

RESEARCH

Open Access

CrossMark

Mathematical programming for the sum of two convex functions with applications to lasso problem, split feasibility problems, and image deblurring problem

Chih Sheng Chuang[1], Zenn-Tsun Yu[2] and Lai-Jiu Lin[3*]

*Correspondence:
maljlin@cc.ncue.edu.tw
[3]Department of Mathematics,
National Changhua University of
Education, Changhua, 50058,
Taiwan
Full list of author information is
available at the end of the article

Abstract

In this paper, two iteration processes are used to find the solutions of the mathematical programming for the sum of two convex functions. In infinite Hilbert space, we establish two strong convergence theorems as regards this problem. As applications of our results, we give strong convergence theorems as regards the split feasibility problem with modified CQ method, strong convergence theorem as regards the lasso problem, and strong convergence theorems for the mathematical programming with a modified proximal point algorithm and a modified gradient-projection method in the infinite dimensional Hilbert space. We also apply our result on the lasso problem to the image deblurring problem. Some numerical examples are given to demonstrate our results. The main result of this paper entails a unified study of many types of optimization problems. Our algorithms to solve these problems are different from any results in the literature. Some results of this paper are original and some results of this paper improve, extend, and unify comparable results in existence in the literature.

MSC: 90C33; 90C34; 90C59

Keywords: lasso problem; mathematical programming for the sum of two functions; split feasibility problem; gradient-projection algorithm; proximal point algorithm

1 Introduction

Let $\mathbb{R}$ be the set of real numbers, H and H_1 be (real) Hilbert spaces with inner product $\langle \cdot, \cdot \rangle$ and norm $\|\cdot\|$. Let C and Q be nonempty, closed, convex subsets of H and H_1, respectively. Let $\Gamma_0(H)$ be the space of all proper lower semicontinuous convex functions from H to $(-\infty, \infty]$. In this paper, we consider the following minimization problem:

$$\arg\min_{x \in H} h_1(x) + g_1(x), \tag{1.1}$$

where $h_1, g_1 \in \Gamma_0(H)$, and $g_1 : H \to \mathbb{R}$ is a Fréchet differential function.

Let A be a $m \times n$ real matrix, $x \in \mathbb{R}^n$, $b \in R^m$, $\gamma \geq 0$ be a regularization parameter and $t \geq 0$. Tibshirani [1] studied the following minimization problem:

$$\min_x \frac{1}{2}\|Ax - b\|_2^2 \quad \text{subject to } \|x\|_1 \leq t, \tag{1.2}$$

where $\|x\|_1 = \sum_{i=1}^n |x_i|$ for $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$, and $\|y\|_2^2 = \sum_{i=1}^m (y_i)^2$ for $y = (y_1, y_2, \ldots, y_m) \in \mathbb{R}^m$. Problem (1.2) is called lasso, an abbreviation of 'the least absolute shrinkage and select operator'. This is a special case of problem (1.1). Besides, we know that problem (1.2) is equivalent to the following problem:

$$\min_x \frac{1}{2} \|Ax - b\|_2^2 + \gamma \|x\|_1. \tag{1.3}$$

Therefore, problems (1.2) and (1.3) are special cases of problem (1.1).

Due to the involvement of the ℓ_1 norm, which promotes the sparsity phenomena of many real world problems arising from image/signal processing, statistical regression, machining learning and so on, the lasso receives much attention (see Combettes and Wajs [2], Xu [3], and Wang and Xu [4]).

Let $f : H \to (-\infty, \infty]$ be proper and let C be a nonempty subset of $\mathrm{dom}(f)$. Then f is said to be uniformly convex on C if

$$f\big(\alpha x + (1-\alpha)y + \alpha(1-\alpha)p(\|x-y\|)\big) \le \alpha f(x) + (1-\alpha)f(y)$$

for all $x, y \in C$, and for all $\alpha \in (0,1)$, where $p : [0, \infty) \to [0, \infty]$ is an increasing function, p vanishes only at 0.

Let $g \in \Gamma_0(H)$ and $\lambda \in (0, \infty)$. The proximal operator of g is defined by

$$\mathrm{prox}_g x = \arg\min_{v \in H} \left(g(v) + \frac{1}{2} \|v - x\|^2 \right), \quad x \in H.$$

The proximal operator of $g \in \Gamma_0(H)$ of order $\lambda \in (0, \infty)$ is defined as the proximal operator of λg, that is,

$$\mathrm{prox}_{\lambda g} x = \arg\min_{v \in H} \left(g(v) + \frac{1}{\lambda} \|v - x\|^2 \right), \quad x \in H.$$

The following results are important results on the solution of the problem (1.1).

Theorem 1.1 (Douglas-Rachford-algorithm) [5] *Let f and g be functions in $\Gamma_0(H)$ such that $(\partial f + \partial g)^{-1} 0 \ne \emptyset$. Let $\{\lambda_n\}_{n \in \mathbb{N}}$ be a sequence in $[0, 2]$ such that $\sum_{n \in \mathbb{N}} \lambda_n (2 - \lambda_n) = +\infty$. Let $\gamma \in (0, \infty)$, and $x_0 \in H$. Set*

$$\begin{cases} y_n = \mathrm{prox}_{\gamma g} x_n, \\ z_n = \mathrm{prox}_{\gamma f}(2y_n - x_n), \\ x_{n+1} = x_n + \lambda_n(z_n - x_n), \quad n \in \mathbb{N}. \end{cases}$$

Then there exists $x \in H$ such that the following hold:

(i) $\mathrm{prox}_{\gamma g} x \in \arg\min_{x \in H}(f + g)(x)$;

(ii) $\{y_n - z_n\}_{n \in \mathbb{N}}$ *converges strongly to* 0;

(iii) $\{x_n\}_{n \in \mathbb{N}}$ *converges weakly to* x;

(iv) $\{y_n\}_{n \in \mathbb{N}}$ *and* $\{z_n\}_{n \in \mathbb{N}}$ *converge weakly to* $\mathrm{prox}_{\gamma g} x$;

(v) *suppose that one of the following holds*:

(a) f *is uniformly convex on every nonempty subset of* $\mathrm{dom}\,\partial f$;

(b) g *is uniformly convex on every nonempty bounded subset of* $\mathrm{dom}\,\partial g$.

Then $\{y_n\}_{n\in\mathbb{N}}$ and $\{z_n\}_{n\in\mathbb{N}}$ converge strongly to $\mathrm{prox}_{\gamma g}\, x$, which is the unique minimizer of $f + g$.

Theorem 1.2 (Forward-backward algorithm) [6] *Let $f \in \Gamma_0(H)$, let $g : H \to \mathbb{R}$ be convex and differentiable with a $\frac{1}{\beta}$-Lipschitz continuous gradient for some $\beta \in (0,\infty)$, let $\gamma \in (0, 2\beta)$, and set $\delta = \min\{1, \frac{\beta}{\gamma}\} + \frac{1}{2}$. Furthermore, let $\{\lambda_n\}_{n\in\mathbb{N}}$ be a sequence in $(0, 2\delta]$ such that $\sum_{n\in\mathbb{N}} \lambda_n(\delta - \lambda_n) = +\infty$. Suppose that $\arg\min(f + g) \neq \emptyset$ and let*

$$\begin{cases} y_n = x_n - \gamma \nabla g(x_n), \\ x_{n+1} = x_n + \lambda_n(\mathrm{prox}_{\gamma f}\, y_n - x_n), \quad n \in \mathbb{N}. \end{cases}$$

Then the following hold:

   (i) *$(x_n)_{n\in\mathbb{N}}$ converges weakly to a point in $\arg\min_{x\in H}(f + g)(x)$;*

   (ii) *suppose that $\inf_{n\in\mathbb{N}} \lambda_n \in (0, \infty)$ and one of the following hold*:

      (a) *f is uniformly convex on every nonempty bounded subset of $\mathrm{dom}\, \partial f$;*

      (b) *g is uniformly convex on every nonempty bounded subset of H.*

Then $\{x_n\}_{n\in\mathbb{N}}$ converges strongly to the unique minimizer of $f + g$

Theorem 1.3 (Tseng's algorithm) [7] *Let D be a nonempty subset of H, $f \in \Gamma_0(H)$ be such that $\mathrm{dom}\, \partial f \subset D$, and let $g \in \Gamma_0(H)$ be such that Gâteaux differentiable on D. Suppose that C is a nonempty, closed, convex subset of D such that $C \cap \arg\min_{x\in H}(f + g)(x) \neq \emptyset$, and that ∇g is $\frac{1}{\beta}$-Lipschitz continuous relative to $C \cup \mathrm{dom}\, \partial f$ for some $\beta \in (0,\infty)$. Let $x_0 \in C$ and $\gamma \in (0, \beta)$, and set*

$$\begin{cases} y_n = x_n - \nabla g(x_n), \\ z_n = \mathrm{prox}_{\gamma f}\, y_n, \\ r_n = z_n - \gamma \nabla g(z_n), \\ x_{n+1} = P_C(x_n - y_n + r_n), \quad n \in \mathbb{N}. \end{cases}$$

Then

   (a) *$\{x_n\}_{n\in\mathbb{N}}$ and $\{z_n\}_{n\in\mathbb{N}}$ converges weakly to a point in $C \cap \arg\min_{x\in H}(f + g)(x)$;*

   (b) *suppose that f or g is uniformly convex on every nonempty subset of $\mathrm{dom}\, \partial f$.*

Then $\{x_n\}_{n\in\mathbb{N}}$ and $\{z_n\}_{n\in\mathbb{N}}$ converges strongly in $C \cap \arg\min_{x\in H}(f + g)(x)$.

Combettes and Wajs [2] used the proximal gradient method to generalize a sequence $\{x_n\}$ by the algorithm: $x_0 \in H$ is chosen arbitrarily, and

$$x_{n+1} = \mathrm{prox}_{\lambda_n g}(I - \lambda_n \nabla f)x_n.$$

We observed that Combettes and Wajs [2] showed that $\{x_n\}$ converges weakly to a solution of the minimization problem (1.1) under suitable conditions. In 2014, Xu [3] gave an iteration process and proved weak convergence theorems to the solution for the problem (1.1). Next, Wang and Xu [4] studied problem (1.1) by the following two types of iteration processes:

   (a) $x_{n+1} := \mathrm{prox}_{\lambda_n g}((1 - \gamma_n)x_n - \delta_n \nabla f(x_n))$ for all $n \in \mathbb{N}$;

   (b) $x_{n+1} := (I - \gamma_n \mathrm{prox}_{\lambda_n g}(I - \delta_n \nabla f))x_n$ for all $n \in \mathbb{N}$.

For these two iteration processes, Wang and Xu [4] proved that they converge strongly to a solution of the problem (1.1) under suitable conditions.

Let I be the identity function of H, and let $f_1 : C \times C \to H$ be a bifunction. Let $g_1 : H \to (-\infty, \infty]$ be a proper convex Fréchet differentiable function with Fréchet derivative ∇g_1 on $\mathrm{int}(\mathrm{dom}(g_1))$, $C \subset \mathrm{int}(\mathrm{dom}(g_1))$, and $h_1 : H \to (-\infty, \infty]$ be a proper convex lower semicontinuous function. Let P_C be the metric projection of H into C. Throughout this paper, we use these notations unless specified otherwise.

Motivated by the results of the above problems, in this paper, we introduce the following iterations to study problem (1.1).

Iteration (I) Let $x_1 \in C$ be chosen arbitrarily, and

$$\begin{cases} y_n = J_\lambda^{A f_1} P_C x_n, \\ x_{n+1} = \alpha_n x_n + (1 - \alpha_n)(\beta_n \theta_n + (I - \beta_n V) y_n), \quad n \in \mathbb{N}, \end{cases}$$

where $f_1(x, y) = \langle y - x, \nabla g_1(x) \rangle + h_1(y) - h_1(x)$.

Iteration (II) Let $x_1 \in C$ be chosen arbitrarily, and

$$\begin{cases} y_n = J_r^{\partial h_1}(I - r \nabla g_1) x_n, \\ x_{n+1} = \alpha_n x_n + (1 - \alpha_n)(\beta_n \theta_n + (I - \beta_n V) y_n), \quad n \in \mathbb{N}. \end{cases}$$

Then we establish two strong convergence theorems without the uniformly convex assumption on the functions we consider. Our results improve Combettes and Wajs [2], Xu [3], Douglas-Rachford [5], Theorem 1.2, and Tseng [7]. Our results is also different from Wang and Xu [4].

We also apply our results to study the following problems.

(AP1) Split feasibility problem:

$$\text{Find } \bar{x} \in H \text{ such that } \bar{x} \in C \text{ and } A\bar{x} \in Q, \qquad \text{(SFP)}$$

where $A : H \to H_1$ is a linear and bounded operator.

In 1994, the split feasibility problem (SFP) in finite dimensional Hilbert spaces was first introduced by Censor and Elfving [8] for modeling inverse problems which arise from phase retrievals and in medical image reconstruction. Since then, the split feasibility problem (SFP) has received much attention due to its applications in signal processing, image reconstruction, with particular progress in intensity-modulated radiation therapy, approximation theory, control theory, biomedical engineering, communications, and geophysics. For examples, one can refer to [8–13] and related literature.

In 2002, Byrne [9] first introduced the so-called CQ algorithm which generates a sequence $\{x_n\}$ by the following recursive procedure:

$$x_{n+1} = P_C\big(x_n - \rho_n A^*(I - P_Q) A x_n\big),$$

where the stepsize ρ_n is chosen in the interval $(0, 2/\|A\|^2)$, and P_C and P_Q are the metric projections onto $C \subseteq \mathbb{R}^n$ and $Q \subseteq \mathbb{R}^m$, respectively. Byrne [9] used the CQ iteration

method to study the split feasibility problem in finite dimensional spaces, but in the infinite dimensional Hilbert space, a strong convergence theorem may not be true for the split feasibility problem by the CQ algorithm [14]. Hence, some modified CQ algorithms are introduced.

In this paper, we give a new iteration to study the (SFP), we also give our modified CQ method to study (SFP), we study two strong convergence theorem to the solution of this problem. We establish a strong convergence of this problem. Our results are different from Theorem 5.1 of Xu [14] and improve other results of Xu [14].

(AP2) Lasso problem:

$$\arg\min_{x\in\mathbb{R}^n}\frac{1}{2}\|Ax-b\|_2^2+\gamma\|x\|_1.$$

We give two iterations to study the lasso problem, and we establish two strong convergence theorems of the lasso problem.

(AP3) Mathematical programming for convex function:

$$\arg\min_{x\in H}h_1(x),\text{ where }h_1\in\Gamma_0(H).$$

Rockafellar [15] used a proximal point algorithm and proved a weak convergence theorem to the solution of this problem. We establish a modified proximal point algorithm and prove a strongly convergence theorem to study this problem. Our result improves the results given by Rockafellar [15].

(AP4) Mathematical programming for convex Fréchet differentiable function:

$$\text{Find }\arg\min_{x\in H}g_1(x),\text{ where }g_1:H\to\mathbb{R}\text{ is a Fréchet differentiable function.}$$

Recently, Xu [16] studied weak convergence and strong convergence theorem of this problem with various types of relaxed gradient-projection iterations. He also used the viscosity nature of the gradient-projection method and the regularized method to study strong convergence theorems of this problem.

A special case of one of our iteration is modified gradient-projection algorithm. We use this modified gradient-projection algorithm to establish a strong convergence theorem of this problem (AP4), and our results improve recent results given by Xu in [16].

In this paper, we apply a recent result of Yu and Lin [17] to find the solution of the mathematical programming for two convex functions, then we apply our results on mathematical programming for two convex functions to study the above problems. We establish a strongly convergent theorem for these problems and apply our result on the lasso problem to the image deblurring problem. Some numerical examples are given to demonstrate our results. The main result of this paper gives a unified study of many types of optimization problems. Our algorithms to solve these problems are different from any results in the literature. Some results of this paper are original and some results of this paper improve, extend, and unify comparable results existence in the literature.

2 Preliminaries

Throughout this paper, we denote the strong convergence of $\{x_n\}$ to $x\in H$ by $x_n\to x$. Let $T:C\to H$ be mapping, and let $\text{Fix}(T):=\{x\in C:Tx=x\}$ denote the set of fixed points of T. Thus:

(i) T is called nonexpansive if $\|Tx - Ty\| \leq \|x - y\|$ for all $x, y \in C$.

(ii) T is strongly monotone if there exists $\bar{\gamma} > 0$ such that $\langle x - y, Tx - Ty \rangle \geq \bar{\gamma}\|x - y\|^2$ for all $x, y \in C$.

(iii) T is Lipschitz continuous if there exists $L > 0$ such that $\|Tx - Ty\| \leq L\|x - y\|$ for all $x, y \in C$.

(iv) Let $\alpha > 0$. Then T is α-inverse-strongly monotone if $\langle x - y, Tx - Ty \rangle \geq \alpha\|Tx - Ty\|^2$ for all $x, y \in C$. We denote T is α-ism if T is α-inverse-strongly monotone.

(v) T is firmly nonexpansive if

$$\|Tx - Ty\|^2 \leq \|x - y\|^2 - \left\|(I - T)x - (I - T)y\right\|^2 \quad \text{for every } x, y \in C,$$

that is,

$$\|Tx - Ty\|^2 \leq \langle x - y, Tx - Ty \rangle \quad \text{for every } x, y \in C.$$

Let $B : H \multimap H$ be a multivalued mapping. The effective domain of B is denoted by $D(B)$, that is, $D(B) = \{x \in H : Bx \neq \emptyset\}$. Thus:

(i) B is a monotone operator on H if $\langle x - y, u - v \rangle \geq 0$ for all $x, y \in D(B)$, $u \in Bx$ and $v \in By$.

(ii) B is a maximal monotone operator on H if B is a monotone operator on H and its graph is not properly contained in the graph of any other monotone operator on H.

Lemma 2.1 [18] *Let $G_1 : H \multimap H$ be maximal monotone. Let $J_r^{G_1}$ be the resolvent of G defined by $J_r^{G_1} = (I + rG_1)^{-1}$ for each $r > 0$. Then the following hold*:

(i) *For each $r > 0$, $J_r^{G_1}$ is single-valued and firmly nonexpansive.*

(ii) *$\mathcal{D}(J_r^{G_1}) = H_1$ and $\text{Fix}(J_r^{G_1}) = \{x \in \mathcal{D}(G_1) : 0 \in G_1 x\}$.*

Let C be a nonempty, closed, convex subset of a real Hilbert space H. Let $g : C \times C \to \mathbb{R}$ be a function. The Ky Fan inequality problem [19] is to find $z \in C$ such that

$$g(z, y) \geq 0 \quad \text{for each } y \in C. \tag{EP}$$

The solution set of Ky Fan inequality problem (KF) is denoted by $\text{KF}(C, g)$.

For solving the Ky Fan inequalities problem, let us assume that the bifunction $g : C \times C \to \mathbb{R}$ satisfies the following conditions:

(A1) $g(x, x) = 0$ for each $x \in C$;

(A2) g is monotone, *i.e.*, $g(x, y) + g(y, x) \leq 0$ for any $x, y \in C$;

(A3) for each $x, y, z \in C$, $\limsup_{t \downarrow 0} g(tz + (1 - t)x, y) \leq g(x, y)$;

(A4) for each $x \in C$, the scalar function $y \to g(x, y)$ is convex and lower semicontinuous.

We have the following result from Blum and Oettli [20].

Lemma 2.2 [20] *Let $g : C \times C \to \mathbb{R}$ be a bifunction which satisfies conditions* (A1)-(A4). *Then for each $r > 0$ and each $x \in H$, there exists $z \in C$ such that*

$$g(z, y) + \frac{1}{r}\langle y - z, z - x \rangle \geq 0$$

for all $y \in C$.

In 2005, Combettes and Hirstoaga [21] established the following important properties of resolvent operator.

Lemma 2.3 [21] *Let* $g : C \times C \to \mathbb{R}$ *be a function satisfying conditions* (A1)-(A4). *For* $r > 0$, *define* $T_r^g : H \to C$ *by*

$$T_r^g x = \left\{ z \in C : g(z, y) + \frac{1}{r} \langle y - z, z - x \rangle \geq 0, \forall y \in C \right\}$$

for all $x \in H$. *Then the following hold*:
　(i)　T_r^g *is single-valued*;
　(ii)　T_r^g *is firmly nonexpansive, that is,* $\|T_r^g x - T_r^g y\|^2 \leq \langle x - y, T_r^g x - T_r^g y \rangle$ *for all* $x, y \in H$;
　(iii)　$\{x \in H : T_r^g x = x\} = \{x \in C : g(x, y) \geq 0, \forall y \in C\}$;
　(iv)　$\{x \in C : g(x, y) \geq 0, \forall y \in C\}$ *is a closed and convex subset of* C.

We call such T_r^g the resolvent of g for $r > 0$.

Takahashi *et al.* [22] gave the following lemma.

Lemma 2.4 [22] *Let* $g : C \times C \to \mathbb{R}$ *be a bifunction satisfying the conditions* (A1)-(A4). *Define* A_g *as follows*:

$$A_g x = \begin{cases} \{z \in H : g(x, y) \geq \langle y - x, z \rangle, \forall y \in C\} & \text{if } x \in C, \\ \emptyset & \text{if } x \notin C. \end{cases} \tag{L4.2}$$

Then $\mathrm{KF}(C, g) = A_g^{-1} 0$ *and* A_g *is a maximal monotone operator with the domain of* $A_g \subset C$. *Furthermore, for any* $x \in H$ *and* $r > 0$, *the resolvent* T_r^g *of* g *coincides with the resolvent of* A_g, *i.e.,* $T_r^g x = (I + r A_g)^{-1} x$.

Let $f : H \to (-\infty, \infty]$ be proper. The subdifferential ∂f of f is the set valued operator, defined $\partial f = \{u \in H : f(y) \geq f(x) + \langle y - x, u \rangle$ for all $y \in H\}$.

Let $x \in H$. Then f is subdifferentiable at x if $\partial f(x) \neq \emptyset$. Then the elements of $\partial f(x)$ are called the subgradient of f at x.

The directional derivative of f at x in the direction y is

$$f'(x, y) = \lim_{\alpha \downarrow 0} \frac{f(x + \alpha y) - f(x)}{\alpha},$$

provided that the limit exists in $[-\infty, \infty]$.

Let $x \in \mathrm{dom} f$ and suppose that $f'(x, y)$ is linear and continuous, then f is said to be Gâteaux differentiable at x. By the Riesz representation theorem, there exists a unique vector $\nabla f(x) \in H$ such that $f'(x, y) = \langle y, \nabla f(x) \rangle$ for all $y \in H$.

Let $x \in H$, let $\mu(x)$ denote the family of all neighborhood of x, let H_1 be a Hilbert space, let $C \in \mu(x)$ and let $f : C \to (-\infty, \infty]$. Then f is said to be Fréchet differentiable at x if there exists an operator $\nabla f(x) \in B(H, \mathbb{R})$, called the Fréchet derivative of f at x, such that

$$\lim_{0 \neq \|y\| \to 0} \frac{|f(x + y) - f(x) - \langle y, \nabla f(x) \rangle|}{\|y\|} = 0.$$

Further, if f is Frêchet differentiable at x, then f is Gâteaux differentiable at x.

Let C be a nonempty, closed, convex subset of H. The indicator function ι_C defined by

$$\iota_C x = \begin{cases} 0, & x \in C, \\ \emptyset, & x \notin C, \end{cases}$$

is a proper lower semicontinuous convex function and its subdifferential $\partial \iota_C$ defined by

$$\partial \iota_C x = \left\{ z \in H : \langle y - x, z \rangle \leq \iota_C(y) - \iota_C(x), \forall y \in H \right\}$$

is a maximal monotone operator (see Lemma 2.8). Furthermore, we also define the normal cone $N_C u$ of C at u as follows:

$$N_C u = \left\{ z \in H : \langle z, v - u \rangle \leq 0, \forall v \in C \right\}.$$

We can define the resolvent $J_\lambda^{\partial i_C}$ of ∂i_C for $\lambda > 0$, *i.e.*,

$$J_\lambda^{\partial i_C} x = (I + \lambda \partial i_C)^{-1} x$$

for all $x \in H$. Since

$$\begin{aligned} \partial i_C x &= \left\{ z \in H : i_C x + \langle z, y - x \rangle \leq i_C y, \forall y \in H \right\} \\ &= \left\{ z \in H : \langle z, y - x \rangle \leq 0, \forall y \in C \right\} \\ &= N_C x \end{aligned}$$

for all $x \in C$, we have

$$\begin{aligned} u = J_\lambda^{\partial i_C} x \quad &\Leftrightarrow \quad x \in u + \lambda \partial i_C u \\ &\Leftrightarrow \quad x - u \in \lambda N_C u \\ &\Leftrightarrow \quad \langle x - u, y - u \rangle \leq 0, \quad \forall y \in C \\ &\Leftrightarrow \quad u = P_C x. \end{aligned}$$

For details see [22].

Lemma 2.5 [4] *Let $g \in \Gamma_0(H)$ and $\lambda \in (0, \infty)$. Thus*:

(i) *If C is a nonempty, closed, convex subset of H and $g = i_C$ is the indicator function of C, then the proximal operator $\mathrm{prox}_{\lambda g} = P_C$ for all $\lambda \in (0, \infty)$, where P_C is the metric projection operator from H to C.*

(ii) *$\mathrm{prox}_{\lambda g}$ is firmly nonexpansive.*

(iii) *$\mathrm{prox}_{\lambda g} = (I + \lambda \partial g)^{-1} = J_\lambda^{\partial g}$.*

Lemma 2.6 [3] *Let $f, g \in \Gamma_0(H)$. Let $x^* \in H$ and $\lambda \in (0, \infty)$. Assume that f is finite valued and Fréchet differentiable function on H with Fréchet derivative ∇f. Then x^* is a solution to the problem $\arg \min_{x \in H} f(x) + g(x)$ if and only if $x^* = \mathrm{prox}_{\lambda g}(I - \lambda \nabla f) x^*$.*

Lemma 2.7 [23] *Let $C \subset H$ be nonempty, closed, convex subset, let $A : H \to H$, and let $f : H \to \mathbb{R}$ be convex and Fréchet differentiable. Let A be the Fréchet derivative of f. Then $\mathrm{VI}(C, A) = \arg \min_{x \in C} f(x)$.*

Lemma 2.8 [6] *Let $f \in \Gamma_0(H)$, then ∂f is maximum monotone.*

Lemma 2.9 [6] *Let f and g be functions in $\Gamma_0(H)$ such that one of the following holds*:

(i) $\operatorname{dom} f \cap \operatorname{int}(\operatorname{dom})g \neq \emptyset$;

(ii) $\operatorname{dom} g = H$.

Then $\partial(f + g) = \partial f + \partial g$.

A mapping $T_\alpha : H \to H$ is said to be averaged if $T_\alpha = (1 - \alpha)I + \alpha T$, where $\alpha \in (0, 1)$ and $T : H \to H$ is nonexpansive. In this case, we say that T_α is α-averaged. Clearly, a firmly nonexpansive mapping is $\frac{1}{2}$-averaged.

Lemma 2.10 [24] *Let $T : H \to H$ be a mapping. Then the following hold*:

(i) *T is nonexpansive if and only if the complement $(I - T)$ is $1/2$-ism*;

(ii) *if S is υ-ism, then for $\gamma > 0$, γS is υ/γ-ism*;

(iii) *S is averaged if and only if the complement $I - S$ is υ-ism for some $\upsilon > 1/2$*;

(iv) *if S and T are both averaged, then the product (composite) ST is averaged*;

(v) *if the mappings $\{T_i\}_{i=1}^n$ are averaged and have a common fixed point, then $\bigcap_{i=1}^n \operatorname{Fix}(T_i) = \operatorname{Fix}(T_1 \cdots T_n)$.*

Lemma 2.11 [6] *Let $f : H \to (-\infty, \infty]$ be proper and convex. Suppose that f is Gâteaux differentiable at x. Then $\partial f(x) = \{\nabla f(x)\}$.*

3 Common solution of variational inequality problem, fixed point, and Ky Fan inequalities problem

For each $i = 1, 2$, let $F_i : C \to H$ be a κ_i-inverse-strongly monotone mapping of C into H with $\kappa_i > 0$. For each $i = 1, 2$, let G_i be a maximal monotone mapping on H such that the domain of G_i is included in C and define the set $G_i^{-1}0$ as $G_i^{-1}0 = \{x \in H : 0 \in G_i x\}$. Let $J_\lambda^{G_1} = (I + \lambda G_1)^{-1}$ and $J_r^{G_2} = (I + rG_2)^{-1}$ for each $n \in \mathbb{N}$, $\lambda > 0$ and $r > 0$. Let $\{\theta_n\} \subset H$ be a sequence. Let V be a $\bar{\gamma}$-strongly monotone and L-Lipschitz continuous operator with $\bar{\gamma} > 0$ and $L > 0$. Let $T : C \to H$ be a nonexpansive mapping. Throughout this paper, we use these notations and assumptions unless specified otherwise.

In this paper, we say conditions (D) hold if the following conditions are satisfied:

(i) $0 < \liminf_{n \to \infty} \alpha_n \leq \limsup_{n \to \infty} \alpha_n < 1$;

(ii) $\lim_{n \to \infty} \beta_n = 0$, and $\sum_{n=1}^\infty \beta_n = \infty$;

(iii) $\lim_{n \to \infty} \theta_n = 0$.

The following strong convergence theorem is needed in this paper.

Theorem 3.1 [17] *Suppose that $\Omega_1 = \operatorname{Fix}(T) \cap \operatorname{Fix}(J_\lambda^{G_1}(I - \lambda F_1)) \cap \operatorname{Fix}(J_r^{G_2}(I - rF_2)) \neq \emptyset$. Take $\mu \in \mathbb{R}$ such that $0 < \mu < \frac{2\bar{\gamma}}{L^2}$. A sequence $\{x_n\} \subset H$ is defined as follows: $x_1 \in C$ is chosen arbitrarily,*

$$
\begin{cases}
y_n = J_\lambda^{G_1}(I - \lambda F_1)J_r^{G_2}(I - rF_2)x_n, \\
s_n = T y_n, \\
x_{n+1} = \alpha_n x_n + (1 - \alpha_n)(\beta_n \theta_n + (I - \beta_n V)s_n)
\end{cases}
$$

for each $n \in \mathbb{N}$, $\{\lambda, r\} \subset (0, \infty)$, $\{\alpha_n\} \subset (0, 1)$, and $\{\beta_n\} \subset (0, 1)$. Assume that conditions (D) hold and $0 < \lambda < 2\kappa_1$ and $0 < r < 2\kappa_2$. Then

$$
\lim_{n \to \infty} x_n = \bar{x},
$$

where

$$\bar{x} = P_{\mathrm{Fix}(T) \cap \mathrm{Fix}(J_\lambda^{G_1}(I-\lambda F_1)) \cap \mathrm{Fix}(J_r^{G_2}(I-rF_2))}(\bar{x} - V\bar{x}).$$

This point $\bar{x}$ is also the unique solution to the hierarchical variational inequality:

$$\langle V\bar{x}, q - \bar{x} \rangle \geq 0, \quad \forall q \in \mathrm{Fix}(T) \cap \mathrm{Fix}\big(J_\lambda^{G_1}(I - \lambda F_1)\big) \cap \mathrm{Fix}\big(J_r^{G_2}(I - rF_2)\big).$$

For each $i = 1, 2$, let $f_i : C \times C \to \mathbb{R}$ be a bifunction satisfying conditions (A1)-(A4). An iteration is used to find common solutions of a variational inequality problem, Ky Fan inequalities problems, and a fixed point set of a mapping:

$$\begin{cases} \text{Find } \bar{x} \in H \text{ such that } \bar{x} \in \mathrm{Fix}(T) \cap \mathrm{KF}(C, f_1) \cap \mathrm{KF}(C, f_2) \text{ and} \\ \langle V\bar{x}, q - \bar{x} \rangle \geq 0, \quad \forall q \in \mathrm{Fix}(T) \cap \mathrm{KF}(C, f_1) \cap \mathrm{KF}(C, f_2). \end{cases}$$

Theorem 3.2 *For each $i = 1, 2$, let $f_i : C \times C \to \mathbb{R}$ be a bifunction satisfying conditions* (A1)-(A4), *and let A_{f_i} be defined as* (L4.2) *in Lemma* 2.4. *Suppose that $\Omega_2 := \mathrm{Fix}(T) \cap \mathrm{KF}(C, f_1) \cap \mathrm{KF}(C, f_2) \neq \emptyset$. Take $\mu \in \mathbb{R}$ such that $0 < \mu < \frac{2\bar{\gamma}}{L^2}$. A sequence $\{x_n\} \subset H$ is defined as follows: $x_1 \in C$ chosen arbitrarily, and*

$$\begin{cases} y_n = J_\lambda^{A_{f_1}} J_r^{A_{f_2}} x_n, \\ s_n = T y_n, \\ x_{n+1} = \alpha_n x_n + (1 - \alpha_n)(\beta_n \theta_n + (I - \beta_n V) s_n) \end{cases}$$

for each $n \in \mathbb{N}$, $\{\lambda, r\} \subset (0, \infty)$, and $\{\alpha_n, \beta_n\} \subset (0, 1)$. Assume that conditions (D) *hold. Then $\lim_{n\to\infty} x_n = \bar{x}$, where $\bar{x} = P_{\Omega_2}(\bar{x} - V\bar{x})$. This point $\bar{x} \in \Omega_2$ is also the unique solution to the hierarchical variational inequality*:

$$\langle V\bar{x}, q - \bar{x} \rangle \geq 0, \quad \forall q \in \Omega_2.$$

Proof For each $i = 1, 2$, by Lemma 2.4, we know that A_{f_i} is a maximal monotone operator with the domain of $A_{f_i} \subset C$ and $\mathrm{KF}(C, f_i) = A_{f_i}^{-1} 0$. For each $i = 1, 2$, let $F_i = 0$, and $G_i = A_{f_i}$ in Theorem 3.1. By Lemma 2.1(ii), we have, for each $i = 1, 2$,

$$\mathrm{Fix}\big(J_\lambda^{G_i}(I - \lambda F_i)\big) = G_i^{-1} 0 = A_{f_i}^{-1} 0 = \mathrm{KF}(C, f_i).$$

This implies that $\Omega_1 = \Omega_2$. By Theorem 3.1, $\lim_{n\to\infty} x_n = \bar{x}$, where $\bar{x} = P_{\Omega_2}(\bar{x} - V\bar{x})$. This point $\bar{x} \in \Omega_2$ is also the unique solution to the hierarchical variational inequality:

$$\langle V\bar{x}, q - \bar{x} \rangle \geq 0, \quad \forall q \in \Omega_2.$$

Thus,

$$\bar{x} \in \mathrm{Fix}(T) \cap \mathrm{KF}(C, f_1) \cap \mathrm{KF}(C, f_2)$$

and

$$\langle V\bar{x}, q - \bar{x} \rangle \geq 0, \quad \forall q \in \mathrm{Fix}(T) \cap \mathrm{KF}(C, f_1) \cap \mathrm{KF}(C, f_2).$$

Therefore, the proof is completed. □

As a simple consequence of Theorem 3.2, we study the common solution of the Ky Fan inequalities problems.

Theorem 3.3 *Let* $f_1 : C \times C \to \mathbb{R}$ *be a bifunction satisfying conditions* (A1)-(A4) *and let* A_{f_1} *be defined as* (L4.2) *in Lemma* 2.4. *Suppose that* $\Omega_3 := \mathrm{KF}(C, f_1) \neq \emptyset$. *Take* $\mu \in \mathbb{R}$ *such that* $0 < \mu < \frac{2\bar{\gamma}}{L^2}$. *A sequence* $\{x_n\} \subset H$ *is defined as follows*: $x_1 \in C$ *is chosen arbitrarily, and*

$$
\begin{cases}
y_n = J_\lambda^{A_{f_1}} P_C x_n, \\
x_{n+1} = \alpha_n x_n + (1 - \alpha_n)(\beta_n \theta_n + (I - \beta_n V) y_n)
\end{cases}
$$

for each $n \in \mathbb{N}$, $\{\lambda, r\} \subset (0, \infty)$, *and* $\{\alpha_n, \beta_n\} \subset (0, 1)$. *Assume that conditions* (D) *hold. Then* $\lim_{n \to \infty} x_n = \bar{x}$, *where* $\bar{x} = P_{\Omega_3}(\bar{x} - V\bar{x})$. *This point* $\bar{x} \in \mathrm{KF}(C, f_1)$ *is also the unique solution to the hierarchical variational inequality*:

$$
\langle V\bar{x}, q - \bar{x} \rangle \geq 0, \quad \forall q \in \Omega_3.
$$

Proof Let $I|_C$ and i_C be the restriction of the identity function on C and the indicate function on C respectively and let $T = I|_C$, $f_2 = i_C$ in Theorem 3.2, then Theorem 3.3 follows from Theorem 3.2. □

Theorem 3.4 *Let* $\Omega_4 := \mathrm{Fix}(T) \neq \emptyset$. *Take* $\mu \in \mathbb{R}$ *such that* $0 < \mu < \frac{2\bar{\gamma}}{L^2}$. *A sequence* $\{x_n\} \subset H_1$ *is defined as follows*: $x_1 \in C$ *is chosen arbitrarily, and*

$$
\begin{cases}
s_n = T P_C x_n, \\
x_{n+1} = \alpha_n x_n + (1 - \alpha_n)(\beta_n \theta_n + (1 - \beta_n) s_n)
\end{cases}
$$

for each $n \in \mathbb{N}$, *and* $\{\alpha_n, \beta_n\} \subset (0, 1)$. *Assume that conditions* (D) *hold. Then* $\lim_{n \to \infty} x_n = \bar{x}$, *where* $\bar{x} = P_{\Omega_4}(\bar{x} - V\bar{x})$. *This point* $\bar{x} \in \mathrm{Fix}(T)$ *is also the unique solution to the hierarchical variational inequality*:

$$
\langle V\bar{x}, q - \bar{x} \rangle \geq 0, \quad \forall q \in \Omega_4.
$$

Proof For each $i = 1, 2$, let $f_i = i_C$, $A_{f_i} = \partial i_C$ in Theorem 3.2, where i_C is the indicator function of C. Then $\mathrm{KF}(C, f_i) = C$ and $J_r^{A_{f_i}} = P_C$. Therefore, Theorem 3.4 follows immediately from Theorem 3.2. □

4 Mathematical programming for the sum of two convex functions

In the following theorem, an iteration is used to find the solution of the optimization problem for the sum of two convex functions:

$$
\arg \min_{y \in C} (g_2 + h_2)(y).
$$

Theorem 4.1 *Let* $g_1 : H \to (-\infty, \infty)$ *be a convex Fréchet differentiable function with Fréchet derivative* ∇g_1 *on* H, *and* $h_1 : H \to (-\infty, \infty]$ *be a proper convex lower semicontinuous function. Let* $f_1(x, y) = \langle y - x, \nabla g_1 x \rangle + h_1(y) - h_1(x)$ *for all* $x, y \in H$ *and let* A_{f_1} *be*

defined as (L4.2) *in Lemma* 2.4. *Take* $\mu \in \mathbb{R}$ *such that* $0 < \mu < \frac{2\bar{\gamma}}{L^2}$. *Suppose that* $\Pi_1 := H - \text{VI}(C, \nabla g_1, h_1) \neq \emptyset$, *where*

$$H - \text{VI}(C, \nabla g_1, h_1) = \left\{ x \in C : \langle y - x, \nabla g_1 x \rangle + h_1(y) - h_1(x) \geq 0 \text{ for all } y \in C \right\}.$$

A sequence $\{x_n\} \subset H$ *is defined as follows*: $x_1 \in C$ *is chosen arbitrarily, and*

$$\begin{cases} y_n = J_\lambda^{A_{f_1}} P_C x_n, \\ x_{n+1} = \alpha_n x_n + (1 - \alpha_n)(\beta_n \theta_n + (I - \beta_n V) y_n) \end{cases}$$

for each $n \in \mathbb{N}$, $\lambda \in (0, \infty)$, *and* $\{\alpha_n, \beta_n\} \subset (0, 1)$. *Assume that conditions* (D) *hold. Then* $\lim_{n \to \infty} x_n = \bar{x}$, *where* $\bar{x} = P_{\Pi_1}(\bar{x} - V\bar{x})$. *Further,* $\bar{x}$ *is the unique solution to the hierarchical variational inequality*:

$$\langle V\bar{x}, q - \bar{x} \rangle \geq 0, \quad \forall q \in \Pi_1.$$

Proof Since ∇g_1 is the Fréchet derivative of the convex function g_1, it follows from Corollary 17.33 of [6] that ∇g_1 is continuous on C. By Proposition 17.10 of [6], ∇g_1 is monotone on C. Hence, ∇g_1 is bounded on any line segment of C. By Proposition 17.2 of [6],

$$\langle y - x, \nabla g_1(x) \rangle \leq g_1(y) - g_1(x) \quad \text{for all } x, y \in C. \tag{4.1}$$

Since $h_1 : C \to \mathbb{R}$ is a proper convex lower semicontinuous function, it is easy to see that for each $x, y, z \in C$,

$$\limsup_{t \downarrow 0} f_1\big(tz + (1 - t)x, y\big)$$

$$\leq \limsup_{t \downarrow 0} \langle y - (tz + (1 - t)x), \nabla g_1(tz + (1 - t)x) \rangle + \limsup_{t \downarrow 0} \big(h_1(y) - h_1(tz + (1 - t)x)\big)$$

$$= \limsup_{t \downarrow 0} \big(\langle y - tz - (1 - t)x - (y - x), \nabla g_1(tz + (1 - t)x) \rangle + \langle (y - x), \nabla g_1(tz + (1 - t)x) \rangle\big)$$

$$\quad + \limsup_{t \downarrow 0} \big(h_1(y) - h_1(tz + (1 - t)x)\big)$$

$$\leq \langle y - x, \nabla g_1 x \rangle + h_1(y) - h_1(x)$$

$$= f_1(x, y).$$

This shows that condition (A3) is satisfied. It is easy to see that f_1 also satisfies conditions (A1), (A2), and (A4). We see $\text{KF}(C, f_1) = H - \text{VI}(C, \nabla g_1, h_1)$ and $\Pi_1 = \Omega_3 \neq \emptyset$. By Theorem 3.3, $\lim_{n \to \infty} x_n = \bar{x}$, where $\bar{x} = P_{\Omega_3}(\bar{x} - V\bar{x})$, $\bar{x} \in \text{KF}(C, f_1)$. This point $\bar{x} \in \Omega_3$ is also the unique solution to the hierarchical variational inequality:

$$\langle V\bar{x}, q - \bar{x} \rangle \geq 0, \quad \forall q \in \Omega_3.$$

By $\Omega_3 = \Pi_1$ and $\bar{x} \in \text{KF}(C, f_1)$, we have

$$\langle V\bar{x}, q - \bar{x} \rangle \geq 0, \quad \forall q \in \Pi_1$$

and

$$\langle y - \bar{x}, \nabla g_1(\bar{x})\rangle + h_1(y) - h_1(\bar{x}) \geq 0 \quad \text{for all } y \in C. \tag{4.2}$$

By (4.1) and (4.2), we have

$$g_1(y) + h_1(y) - g_1(\bar{x}) - h_1(\bar{x}) \geq \langle y - x, \nabla g_1(x)\rangle + h_1(y) - h_1(\bar{x}) \geq 0$$

for all $y \in C$. Then $\bar{x} \in \arg\min_{y \in C}(g_1 + h_1)(y)$. $\qquad\square$

Example 4.1 Let $h_1(x) = x^2$, $g_1(x) = x^2 + 2x + 1$, $C = [-1,1]$, $H = \mathbb{R}$, $\lambda = 1$, $V = I$, $\alpha_n = \frac{1}{2}$ for all $n \in \mathbb{N}$, $\beta_n = \frac{1}{1,000n}$, $C = [-1,1]$, $\theta_n = 0$, $f_1(x,y) = \langle y - x, \nabla g_1(x)\rangle + h_1(y) - h_1(x)$. Then $f_1(x,y) = (y-x)(2x+y+x+2)$, this implies that $f_1(-\frac{1}{2},y) = (y+\frac{1}{2})^2 \geq 0$ for all $y \in [-1,1]$, and $-\frac{1}{2} \in H - \text{VI}([-1,1], \nabla g_1, h_1) \neq \emptyset$.

We also see $A_{f_1}(-1) = (-\infty, -2]$, $A_{f_1}(1) = [6, \infty)$, and $A_{f_1}(x) = 4x + 2$ if $x \in (-1,1)$. Let $y_n = J_1^{A_{f_1}} P_C x_n$, $x_{n+1} = \frac{1}{2}x_n + \frac{1}{2}(1 - \frac{1}{1,000n})y_n$.

It is easy to see that $P_C x_n = 5y_n + 2$ and $y_n = \frac{1}{5}(P_C x_n - 2)$.

Hence $x_{n+1} = \frac{1}{2}x_n + \frac{1}{10}(1 - \frac{1}{1,000n})(P_C x_n - 2)$. It is easy to see all the conditions of Theorem 4.1 are satisfied.

Let $x_1 = 0$, then $x_2 = -0.1998$, $x_3 = -0.31977001$, $x_4 = -0.39177967$, $x_5 = -0.435008, \ldots$, we see $\lim_{n \to \infty} x_n = \bar{x} = -\frac{1}{2} \in \arg\min_{x \in [-1,1]} g_1 x + h_1 x$.

Next, an iteration is used to find the solution of the following optimization problem for the convex differentiable function:

$$\arg\min_{y \in C} g_1(y).$$

Corollary 4.1 *Let $g_1 : H \to \mathbb{R}$ be a convex Fréchet differentiable function with Fréchet derivative ∇g_1. Let $f_1(x,y) = \langle y - x, \nabla g_1 x\rangle$ for all $x, y \in H$ and let A_{f_1} be defined as (L4.2) in Lemma 2.4. Suppose that $\Pi_{1,1} := \arg\min_{y \in C} g_1(y) \neq \emptyset$. Take $\mu \in \mathbb{R}$ such that $0 < \mu < \frac{2\bar{\gamma}}{L^2}$. A sequence $\{x_n\} \subset H_1$ is defined as follows: $x_1 \in C$ is chosen arbitrarily, and*

$$\begin{cases} y_n = J_\lambda^{A_{f_1}} P_C x_n, \\ x_{n+1} = \alpha_n x_n + (1 - \alpha_n)(\beta_n \theta_n + (I - \beta_n V)y_n) \end{cases}$$

for each $n \in \mathbb{N}$, $\lambda \in (0, \infty)$ and $\{\alpha_n, \beta_n\} \subset (0,1)$. Assume that conditions (D) hold. Then $\lim_{n \to \infty} x_n = \bar{x}$, where $\bar{x} = P_{\Pi_{1,1}}(\bar{x} - V\bar{x})$. Further, $\bar{x}$ is also the unique solution to the hierarchical variational inequality:

$$\langle V\bar{x}, q - \bar{x}\rangle \geq 0, \quad \forall q \in \Pi_{1,1}.$$

Proof By Lemma 2.7, we know that $\text{VI}(C, \nabla g_1) = \arg\min_{y \in C} g_1(y)$. Therefore, Corollary 4.1 follows immediately from Theorem 4.1 by letting $h_1 = 0$. $\qquad\square$

Next, another iteration is used to find the solution of the following optimization problem for a convex function:

$$\arg\min_{z \in C} h_1(z).$$

Corollary 4.2 *Let $f_1(x, y) = h_1(y) - h_1(x)$, for all $x, y \in C$, and let A_{f_1} be defined as* (L4.2) *in Lemma 2.4. Suppose that $\Pi_{1,2} := \arg\min_{y \in C} h_1(y) \neq \emptyset$. Take $\mu \in \mathbb{R}$ such that $0 < \mu < \frac{2\bar{\gamma}}{L^2}$. A sequence $\{x_n\} \subset H_1$ is defined as follows: $x_1 \in C$ is chosen arbitrarily, and*

$$\begin{cases} y_n = J_\lambda^{A_{f_1}} P_C x_n, \\ x_{n+1} = \alpha_n x_n + (1 - \alpha_n)(\beta_n \theta_n + (I - \beta_n V) y_n) \end{cases}$$

for each $n \in \mathbb{N}$, $\lambda \in (0, \infty)$, and $\{\alpha_n, \beta_n\} \subset (0, 1)$. Assume that conditions (D) *hold. Then $\lim_{n \to \infty} x_n = \bar{x}$, where $\bar{x} = P_{\Pi_{1,2}}(\bar{x} - V\bar{x})$. Further, $\bar{x}$ is also the unique solution to the hierarchical variational inequality*:

$$\langle V\bar{x}, q - \bar{x} \rangle \geq 0, \quad \forall q \in \Pi_{1,2}.$$

Proof Put $g_1 = 0$ in Theorem 4.1. Then Corollary 4.2 follows from Theorem 4.1. □

In the following theorem, an iteration is used to find the solution of the following optimization problem for the sum of two convex functions:

$$\arg\min_{y \in H} g_1(y) + h_1(y).$$

Theorem 4.2 *Let $g_1 : H \to (-\infty, \infty)$ be a convex Fréchet differentiable function with Fréchet derivative ∇g_1 on H, and $h_1 : H \to (-\infty, \infty]$ be a proper convex lower semicontinuous function. Suppose that ∇g_1 is Lipschitz with Lipschitz constant $\frac{1}{L_1}$ and $\Pi_2 := \arg\min_{x \in H}(g_1 + h_1)(x) \neq \emptyset$. Take $\mu \in \mathbb{R}$ such that $0 < \mu < \frac{2\bar{\gamma}}{L^2}$. A sequence $\{x_n\} \subset H_1$ is defined as follows: $x_1 \in H$ is chosen arbitrarily, and*

$$\begin{cases} y_n = \text{prox}_{rh_1}(I - r\nabla g_1)x_n = J_r^{\partial h_1}(I - r\nabla g_1)x_n, \\ x_{n+1} = \alpha_n x_n + (1 - \alpha_n)(\beta_n \theta_n + (I - \beta_n V) y_n) \end{cases}$$

for each $n \in \mathbb{N}$, $r \in (0, \frac{2}{L_1})$, and $\{\alpha_n, \beta_n\} \subset (0, 1)$. Assume that conditions (D) *hold. Then $\lim_{n \to \infty} x_n = \bar{x}$, where $\bar{x} = P_{\Pi_2}(\bar{x} - V\bar{x})$. Further, $\bar{x}$ is also the unique solution to the hierarchical variational inequality*:

$$\langle V\bar{x}, q - \bar{x} \rangle \geq 0, \quad \forall q \in \Pi_2.$$

We give two different proofs for this theorem.

Proof I Put $C = H$, $G_1 = \partial h_1$, $F_1 = \nabla g_1$, $T = I|_C$, $F_2 = 0$, $G_2 = \partial i_H$ in Theorem 3.1, where $I|_C$ is the restriction of I on C and i_C is the indicate function of C. Since ∇g_1 is Lipschitz continuous with Lipschitz constant L_1, it follows from Corollary 10 of [25] that ∇g_1 is $\frac{1}{L_1}$-strongly-inverse-monotone. Since h_1 is a proper convex lower semicontinuous function, it follows from Lemma 2.8 that ∂h_1 is a set valued maximum monotone mapping. By Lemma 2.11, $\partial g_1 = \{\nabla g_1\}$. It follows from $\text{dom}(f) \cap \text{int}(\text{dom}(g)) \neq \emptyset$, $\text{dom}(g) = H$, and Lemma 2.9 that

$$\partial(h_1 + g_1)(x) = \partial h_1(x) + \partial g_1(x) = \partial h_1(x) + \nabla g_1(x).$$

Hence,

$$
\begin{aligned}
\Pi_2 = \underset{y \in H}{\arg\min}(g_1 + h_1)(y) &= \left\{ x \in H : x \in (\partial h_1 + \nabla g_1)^{-1} 0 \right\} \\
&= \mathrm{Fix}(I|_C) \cap \mathrm{Fix}\big(J_1^{\partial h_1}(I - \nabla g_1)\big) \cap \mathrm{Fix}(P_C) \\
&= \mathrm{Fix}(T) \cap \mathrm{Fix}\big(J_1^{\partial h_1}(I - \nabla g_1)\big) \cap \mathrm{Fix}\big(J_1^{\partial i_H}\big) \\
&= \Omega_1.
\end{aligned}
$$

Therefore, we get the conclusion of Theorem 4.2 from Theorem 3.1. □

Proof II Let $C = H$, $T = \mathrm{prox}_{rh_1}(I - r\nabla g_1)$ in Theorem 3.4. Since ∇g_1 is Lipschitz constant L_1, it follows from Corollary 10 of [25] that ∇g_1 is $\frac{1}{L_1}$-inverse- strongly-monotone. By Lemma 2.10, $r\nabla g_1$ is $\frac{1}{rL_1}$-ism and $(I - r\nabla g_1)$ is averaged. Since ∂h_1 is maximum monotone, it follows from Lemma 2.5, $\mathrm{prox}_{rh_1} = J_r^{\partial h_1}$ is firmly nonexpansive. Hence prox_{rh_1} is $\frac{1}{2}$-averaged. Then by Lemma 2.10, T is averaged and nonexpansive. We have

$$
\begin{aligned}
\mathrm{Fix}(T) &= \mathrm{Fix}\big(J_r^{\partial h_1}(I - r\nabla g_1)P_H\big) \\
&= \mathrm{Fix}\big(\mathrm{prox}_{rh_1}(I - r\nabla g_1)\big) \cap \mathrm{Fix}(P_H) \\
&= \underset{x \in H}{\arg\min}\, g_1(x) + h_1(x) \cap H \\
&= \underset{x \in C}{\arg\min}\, g_1(x) + h_1(x).
\end{aligned}
$$

Hence, $\Pi_2 = \Omega_4$, and we get the conclusion of Theorem 4.2 from Theorem 3.4. □

Remark 4.1

(a) The iterations in Theorems 4.1 and 4.2 are different.
(b) Theorem 1.1, Theorem 1.2, Theorem 1.3, and the results given by Combettes and Wajs [2], and Xu [3] are weak convergence theorems for the problem:

$$
\underset{x \in H}{\arg\min}\, g_1(x) + h_1(x).
$$

Further, Theorem 1.1, Theorem 1.2, and Theorem 1.3 gave strong convergence theorems of this problem under the uniform convex assumption on h_1 or g_1. Therefore, Theorem 4.2 is different from these results. Besides, Theorem 4.2 is also different from the result given by Wang and Xu [4] and related algorithms in the literature.

Example 4.2 Let $h_1(x) = x^2$, $g_1(x) = x^2 + 2x + 1$, $H = \mathbb{R}$, $\alpha = 1$, $V = I$, $\alpha_n = \frac{1}{2}$ for all $n \in \mathbb{N}$, $\beta_n = \frac{1}{1,000n}$, $C = [-1,1]$, $\theta_n = 0$. Then $\partial h_1(x) = \{2x\}$, $\nabla g_1(x) = 2x + 2$, $\theta_n = 0$ for all $n \in \mathbb{N}$. We see $-\frac{1}{2} \in \arg\min_{x \in \mathbb{R}} g_1(x) + h_1(x) \neq \emptyset$, let

$$
\begin{cases}
y_n = J_1^{\partial h_1}(I - \nabla h_1 g_1)x_n = (I + \partial h_1)^{-1}(I - \nabla g_1)x_n, \\
x_{n+1} = \frac{1}{2}x_n + \frac{1}{2}(1 - \frac{1}{1,000n})y_n
\end{cases}
$$

for each $n \in \mathbb{N}$.

We see all conditions of Theorem 4.2 are satisfied.

We obtain $(I - \nabla g_1)x_n = (I + \partial h_1)y_n = y_n + 2y_n = 3y_n = x_n - 2x_n - 2$.
From this we obtain $y_n = \frac{-(x_n+2)}{3}$ and

$$
\begin{aligned}
x_{n+1} &= \frac{1}{2}x_n + \frac{1}{2}\left(1 - \frac{1}{1,000n}\right)y_n \\
&= \frac{1}{2}x_n - \frac{1}{2}\left(1 - \frac{1}{1,000n}\right)\frac{(x_n + 2)}{3} \\
&= \frac{1}{2}x_n - \left(1 - \frac{1}{1,000n}\right)\frac{(x_n + 2)}{6}.
\end{aligned}
$$

Let $x_1 = 1$, then $x_2 = 0.0005$, $x_3 = -0.33075$, $x_4 = -0.4434906$, $x_5 = -04810987$, $x_6 = -0.493649$, $x_7 = -0.4978412$, $x_8 = -0.4992446$, $x_9 = -0.4997169$, $x_{10} = -0.49987785$, $x_{11} = -0.49993425$, $x_{12} = -4999553$, $x_{13} = -0.4999642$, ... From the relation

$$
x_{n+1} = \frac{1}{2}x_n - \left(1 - \frac{1}{1,000n}\right)\frac{-(x_n + 2)}{6}
$$

and $x_1 = 1$, it is easy to see that the sequence $\{x_n\}_{n\in\mathbb{N}}$ is nonincreasing for some $n \geq m$ and bounded. Hence $\lim_{n\to\infty} x_n$ exists. Let $\lim_{n\to\infty} x_n = \bar{x}$. From the relation

$$
x_{n+1} = \frac{1}{2}x_n - \left(1 - \frac{1}{1,000n}\right)\frac{(x_n + 2)}{6},
$$

we see that $\bar{x} = \frac{1}{2}\bar{x} - \frac{\bar{x}+2}{6}$. Therefore $\bar{x} = -\frac{1}{2} \in \arg\min_{x\in H} g_1(x) + h_1(x)$.

In the following corollary, an iteration is used to find the solution of the following optimization problem:

Find $\bar{x} \in \arg\min_{y\in H} h_1(y)$.

Corollary 4.3 *Let $h_1 : H \to (-\infty, \infty]$ be a proper convex lower semicontinuous function. Take $\mu \in \mathbb{R}$ such that $0 < \mu < \frac{2\bar{\gamma}}{L^2}$. Suppose that $\Pi_{2,1} := \arg\min_{y\in H} h_1(y) \neq \emptyset$. A sequence $\{x_n\} \subset H$ is defined as follows: $x_1 \in H$ is chosen arbitrarily, and*

$$
\begin{cases}
y_n = J_\alpha^{\partial h_1} x_n, \\
x_{n+1} = \alpha_n x_n + (1 - \alpha_n)(\beta_n \theta_n + (I - \beta_n V)y_n)
\end{cases}
$$

for each $n \in \mathbb{N}$, $\alpha \in (0, \infty)$ and $\{\alpha_n, \beta_n\} \subset (0, 1)$. Assume that conditions (D) *hold. Then $\lim_{n\to\infty} x_n = \bar{x}$, where $\bar{x} = P_{\Pi_{2,1}}(\bar{x} - V\bar{x})$. Further, $\bar{x}$ is also the unique solution to the hierarchical variational inequality:*

$$
\langle V\bar{x}, q - \bar{x} \rangle \geq 0, \quad \forall q \in \Pi_{2,1}.
$$

Remark 4.2 In 1976, Rockafellar [15] proved the following in the Hilbert space setting: If h_1 is a proper convex lower semicontinuous function on H, the solution set $\arg\min_{y\in H} h_1(y)$ is nonempty and $\liminf_{n\to\infty} \beta_n > 0$. Let

$$
x_{n+1} = \arg\min_{y\in H}\left\{h_1(y) + \frac{1}{2\beta_n}\|y - x_n\|^2\right\} = \text{prox}_{\beta_n h_1} x_n = J_{\beta_n}^{\partial h_1} x_n, \quad n \in \mathbb{N}, \tag{4.3}
$$

then $\{x_n\}$ converges weakly to a minimizer of h_1. We see that Corollary 4.3 gives a different iteration which converges strongly to the solution of the following problem: Find $\bar{x} \in \arg\min_{y \in H} h_1(y)$.

Next, a modified gradient-projection algorithm is used to find the solution of the following mathematical program:

$$\text{Find } \bar{x} \in \arg\min_{y \in C} g_1(y).$$

Theorem 4.3 *Let $g_1 : H \to (-\infty, \infty)$ be a convex Fréchet differentiable function with Fréchet derivative ∇g_1 on H. Take $\mu \in \mathbb{R}$ such that $0 < \mu < \frac{2\bar{\gamma}}{L^2}$. Suppose that ∇g_1 is Lipschitz continuous with Lipschitz constant L_1 and $\Pi_3 := \arg\min_{y \in C} g_1(y) \neq \emptyset$. A sequence $\{x_n\} \subset H$ is defined as follows: $x_1 \in C$ is chosen arbitrarily, and*

$$\begin{cases} y_n = P_C(I - \alpha \nabla g_1)x_n, \\ x_{n+1} = \alpha_n x_n + (1 - \alpha_n)(\beta_n \theta_n + (I - \beta_n V)y_n) \end{cases}$$

for each $n \in \mathbb{N}$, $\alpha \in (0, \frac{2}{L_1})$, and $\{\alpha_n, \beta_n\} \subset (0,1)$. Assume that conditions (D) *hold. Then $\lim_{n \to \infty} x_n = \bar{x}$, where $\bar{x} = P_{\Pi_3}(\bar{x} - V\bar{x})$. Further, $\bar{x}$ is the unique solution to the hierarchical variational inequality:*

$$\langle V\bar{x}, q - \bar{x} \rangle \geq 0, \quad \forall q \in \Pi_3.$$

Proof Let $h_1 = i_C$, where i_C denotes the indicator function of C. From Lemma 2.5, $\text{prox}_{\lambda h_1} = P_C$, and

$$\arg\min_{x \in C}\left(h_1(x) + g_1(x)\right) = \arg\min_{x \in C} g_1(x),$$

Theorem 4.3 follows immediately from Theorem 4.2. □

Remark 4.3 We know an iteration, defined by

$$x_{n+1} = P_C(I - \alpha_n \nabla g_1)x_n, \quad n \in \mathbb{N},$$

is called a gradient-projection algorithm, where ∇g_1 is Lipschitz continuous. In 2011, Xu [16] used the gradient-projection algorithm and the relaxed gradient-projection algorithm and studied the problem

$$\text{Find } \bar{x} \in \arg\min_{y \in C} g_1(y),$$

and gave weak convergence theorems. Xu also used the viscosity nature of the gradient-projection algorithms and regularized algorithm to study strong convergence theorems for this problem [16]. In Theorem 4.3, we establish a strong convergence theorem for this problem by a different modified gradient-projection algorithm and a different approach.

In the end of this section, an iteration is used to find the solution of the following optimization problem:

Find $\bar{x} \in \arg\min_{y \in H} g_1(y)$.

Corollary 4.4 *Let $g_1 : H \to (-\infty, \infty)$ be a convex Fréchet differentiable function with Fréchet derivative ∇g_1 on H. Take $\mu \in \mathbb{R}$ such that $0 < \mu < \frac{2\bar{\gamma}}{L^2}$. Suppose that ∇g_1 is Lipschitz continuous with Lipschitz constant L_1 and $\Pi_{3,1} := \arg\min_{y \in H} g_1(y) \neq \emptyset$. A sequence $\{x_n\} \subset H$ is defined as follows: $x_1 \in H$ is chosen arbitrarily, and*

$$\begin{cases} y_n = (I - \alpha \nabla g_1)x_n, \\ x_{n+1} = \alpha_n x_n + (1 - \alpha_n)(\beta_n \theta_n + (I - \beta_n V)y_n) \end{cases}$$

for each $n \in \mathbb{N}$, $\alpha \in (0, \frac{2}{L_1})$, and $\{\alpha_n, \beta_n\} \subset (0,1)$. Assume that conditions (D) *hold. Then $\lim_{n \to \infty} x_n = \bar{x}$, where $\bar{x} = P_{\Pi_{3,1}}(\bar{x} - V\bar{x})$. Further, $\bar{x}$ is also the unique solution to the hierarchical variational inequality:*

$$\langle V\bar{x}, q - \bar{x} \rangle \geq 0, \quad \forall q \in \Pi_{3,1}.$$

Proof Let $h_1 = i_H$. By Lemma 2.5, we know that $\text{prox}_{\lambda h_1} = J_\lambda^{\partial i_H} = P_H = I$, and

$$\arg\min_{x \in H}\big(h_1(x) + g_1(x)\big) = \arg\min_{x \in H} g_1(x).$$

Hence, Corollary 4.4 follows immediately from Theorem 4.3.　　　　　　　　□

5 Split feasibility problems and lasso problems

In the following theorem, a modified Byrne CQ iteration is used to find the solution of the following split feasibility problem: Find $\bar{x} \in C$ such that $A\bar{x} \in Q$.

Theorem 5.1 *Let $A : H \to H_1$ be a bounded linear operator, A^* be the adjoint of A. Take $\mu \in \mathbb{R}$ such that $0 < \mu < \frac{2\bar{\gamma}}{L^2}$. Suppose that $\Delta_1 := \{x \in C : Ax \in Q\} \neq \emptyset$. A sequence $\{x_n\} \subset H$ is defined as follows: $x_1 \in C$ is chosen arbitrarily, and*

$$\begin{cases} y_n = P_C(I - \alpha A^*(A - P_Q A))x_n, \\ x_{n+1} = \alpha_n x_n + (1 - \alpha_n)(\beta_n \theta_n + (I - \beta_n V)y_n) \end{cases}$$

for each $n \in \mathbb{N}$, $\alpha \in (0, \frac{2}{\|A\|^2})$, and $\{\alpha_n, \beta_n\} \subset (0,1)$. Assume that conditions (D) *hold. Then $\lim_{n \to \infty} x_n = \bar{x}$, where $\bar{x} = P_{\Delta_1}(\bar{x} - V\bar{x})$. Further, $\bar{x}$ is also the unique solution to the hierarchical variational inequality:*

$$\langle V\bar{x} - \theta, q - \bar{x} \rangle \geq 0, \quad \forall q \in \Delta_1.$$

Proof Let $g_1(x) = \frac{\|Ax - P_Q Ax\|_2^2}{2}$. It is easy to see that $g_1(x) = \min_{y \in Q} \frac{1}{2}\|Ax - y\|^2$ is a convex function. Then for any $v \in H_1$, we have

$$2g_1(x + v) - 2g_1(x) - 2\big\langle Av, (I - P_Q)Ax \big\rangle$$

$$= \big\langle (I - P_Q)A(x + v), (I - P_Q)A(x + v) \big\rangle - \big\langle (I - P_Q)Ax, (I - P_Q)Ax \big\rangle$$

$$- 2\langle Av, (I - P_Q)Ax \rangle$$

$$= \langle Ax + Av - P_QAx + \big(P_QAx - P_QA(x + v)\big), Ax + Av - P_QAx$$

$$+ \big(P_QAx - P_QA(x + v)\big) \rangle - \langle (I - P_Q)Ax, (I - P_Q)Ax \rangle - 2\langle Av, (I - P_Q)Ax \rangle$$

$$= \langle Av, Av \rangle + \langle P_QAx - P_QA(x + v), P_QAx - P_QA(x + v) \rangle$$

$$+ 2\langle (I - P_Q)Ax, P_QAx - P_QA(x + v) \rangle + 2\langle Av, P_QAx - P_QA(x + v) \rangle. \tag{5.1}$$

P_Q is a self-adjoint operator and $P_Q^2 = P_Q$, therefore, we have

$$\big| \langle (I - P_Q)Ax, P_QAx - P_QA(x + v) \rangle \big|$$

$$= \big| \langle (I - P_Q)Ax, P_QAx \rangle - \langle (I - P_Q)Ax, P_QA(x + v) \rangle \big|$$

$$= \langle P_Q(I - P_Q)Ax, Ax \rangle - \langle P_Q(I - P_Q)Ax, A(x + v) \rangle = 0. \tag{5.2}$$

Since $P_Q : H \to Q$ is a nonexpansive mapping, we have

$$\big| \langle P_QAx - P_QA(x + v), P_QAx - P_QA(x + v) \rangle \big|$$

$$= \big\| P_QAx - P_QA(x + v) \big\|_2^2$$

$$\leq \big\| Ax - (Ax + Av) \big\|_2^2$$

$$= \| Av \|_2^2$$

$$\leq \| A \|_2^2 \cdot \| v \|_2^2 \tag{5.3}$$

and

$$2 \big| \langle Av, P_QAx - P_QA(x + v) \rangle \big|$$

$$\leq 2 \big| \| Av \|_2 \big\| P_QAx - P_QA(x + v) \big\|_2 \big|$$

$$\leq 2 \| Av \|_2 \big\| Ax - A(x + v) \big\|_2$$

$$\leq 2 \| Av \|_2 \| Av \|_2$$

$$\leq 2 \| A \|_2^2 \cdot \| v \|_2^2. \tag{5.4}$$

We also see that

$$\langle Av, Av \rangle \leq \| A \|_2^2 \cdot \| v \|_2^2. \tag{5.5}$$

By (5.1), (5.2), (5.3), (5.4), and (5.5), we have

$$\lim_{v \to 0} \frac{|g_1(x + v) - g_1(x) - \langle v, A^*(I - P_Q)Ax \rangle|}{\| v \|_2}$$

$$= \lim_{v \to 0} \frac{|2g_1(x + v) - 2g_1(x) - 2\langle v, A^*(I - P_Q)Ax \rangle|}{\| 2v \|_2}$$

$$\leq \lim_{v \to 0} 2 \| A \|_2^2 \| v \|_2$$

$$= 0. \tag{5.6}$$

This shows that g_1 is Fréchet differentiable with Fréchet derivative $\nabla g_1 = A^*(A - P_Q A)$. Since A is a bounded operator and P_Q is a firmly nonexpansive mapping,

$$
\begin{aligned}
\left\| \nabla g_1(x) - \nabla g_1(y) \right\| &= \left\| A^*(A - P_Q A)x - A^*(A - P_Q A)y \right\| \\
&\leq \|A\| \left\| (I - P_Q)Ax - (I - P_Q)Ay \right\| \\
&\leq \|A\| \|Ax - Ay\| \\
&\leq \|A\|^2 \cdot \|x - y\|.
\end{aligned} \tag{5.7}
$$

This shows that ∇g_1 is a Lipschitz function with Lipschitz constant $\|A\|^2$. By the assumption $\Delta_1 \neq \emptyset$, we know that $\Delta_1 = \Pi_3$. Then we get the conclusion of Theorem 5.1 from Theorem 4.3. □

Remark 5.1 In 2010, Xu [14] used various algorithms to establish weak convergence theorems in infinite dimensional Hilbert spaces for the split feasibility problem (see Theorems 3.1, 3.3, 3.4, 4.1 and 5.7 of [14]). Also, Xu [14] established a strongly theorem for this problem in the infinite dimensional Hilbert space (see Theorem 5.5 of [14]). Now, Theorem 5.1 gives an algorithm to study the split feasibility problem which converges strongly to the split feasibility problem, and this result improves Byrne's CQ algorithm [9], and the results given by Xu [14].

By Theorem 5.1, we get the following result for split feasibility problem.

Corollary 5.1 *Let $A : H \to H_1$ be a bounded linear operator, A^* be the adjoint of A. Take $\mu \in \mathbb{R}$ such that $0 < \mu < 2$. Suppose that $\Delta_1 := \{x \in C : Ax \in Q\} \neq \emptyset$. A sequence $\{x_n\} \subset H$ is defined as follows: $x_1 \in C$ is chosen arbitrarily, and*

$$
\begin{cases}
y_n = P_C(I - \alpha A^*(A - P_Q A))x_n, \\
x_{n+1} = \alpha_n x_n + (1 - \alpha_n)((1 - \beta_n)y_n)
\end{cases}
$$

for each $n \in \mathbb{N}$, $\alpha \in (0, \frac{2}{\|A\|^2})$, and $\{\alpha_n, \beta_n\} \subset (0, 1)$. Assume that conditions (D) hold. Then $\lim_{n \to \infty} x_n = \bar{x}$, where $\bar{x} = P_{\Delta_1} 0$.

Proof Let $\theta_n = \theta = 0$ for all $n \in \mathbb{N}$ and $V = I$ in Theorem 5.1, then Corollary 5.1 follows from Theorem 5.1. □

Next, an iteration is used to find the solution of the following lasso problem:

$$
\text{Find } \bar{x} \in \underset{x \in \mathbb{R}^n}{\arg\min} \left(\frac{1}{2} \|Ax - b\|_2^2 + \gamma \|x\|_1 \right).
$$

Theorem 5.2 *Let A be a $m \times n$ real matrix, $x \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, and $\gamma \geq 0$ be a regularization parameter. Take $\mu \in \mathbb{R}$ such that $0 < \mu < \frac{2\bar{\gamma}}{L^2}$. Suppose that $\Delta_2 := \arg\min_{x \in \mathbb{R}^n} \frac{1}{2} \|Ax - b\|_2^2 + \gamma \|x\|_1 \neq \emptyset$. A sequence $\{x_n\} \subset H$ is defined as follows: $x_1 \in \mathbb{R}^n$ is chosen arbitrarily, and*

$$
\begin{cases}
y_n = J_\alpha^{\partial \gamma \|\cdot\|_1}(x_n - \alpha A^*(Ax_n - b)), \\
x_{n+1} = \alpha_n x_n + (1 - \alpha_n)(\beta_n \theta_n + (I - \beta_n V)y_n)
\end{cases}
$$

for each $n \in \mathbb{N}$, $\alpha \in (0, \frac{2}{\|A^2\|})$, and $\{\alpha_n, \beta_n\} \subset (0, 1)$. Assume that conditions (D) hold. Then $\lim_{n \to \infty} x_n = \bar{x}$, where $\bar{x} = P_{\Delta_2}(\bar{x} - V\bar{x})$. Further, $\bar{x}$ is also the unique solution to the hierarchical variational inequality:

$$\langle V\bar{x} - \theta, q - \bar{x} \rangle \geq 0, \quad \forall q \in \Delta_2.$$

Proof Let $g_1(x) = \frac{1}{2} \|Ax - b\|_2^2$, and $h_1(x) = \gamma \|x\|_1$. For each $v \in \mathbb{R}^n$,

$$\lim_{v \to 0} \frac{|g_1(x + v) - g_1(x) - \langle v, A^*(Ax - b) \rangle|}{\|v\|_2}$$

$$= \lim_{v \to 0} \frac{|\langle A(x + v) - b, A(x + v) - b \rangle - \langle Ax - b, Ax - b \rangle - 2\langle v, A^*(Ax - b) \rangle|}{2\|v\|_2}$$

$$= 0. \tag{5.8}$$

Then $\nabla g_1(x) = A^*(Ax - b)$, and h_1 is a proper convex and lower semicontinuous function on H. Hence,

$$\|\nabla g_1(x) - \nabla g_1(y)\| = \|A^*(Ax - b) - A^*(Ay - b)\|_2 \leq \|A\|^2 \|x - y\|$$

and ∇g_1 is a Lipschitz function with Lipschitz constant $\|A\|^2$. Therefore, Theorem 5.2 follows from Theorem 4.2. □

The following is a special case of Theorem 5.2.

Corollary 5.2 Let A be a $m \times n$ real matrix, $x \in \mathbb{R}^n$, $b \in R^m$, and $\gamma \geq 0$ be a regularization parameter. Take $\mu \in \mathbb{R}$ such that $0 < \mu < 2$. Suppose that $\Delta_2 := \arg\min_{x \in \mathbb{R}^n} \frac{1}{2} \|Ax - b\|_2^2 + \gamma \|x\|_1 \neq \emptyset$. A sequence $\{x_n\} \subset \mathbb{R}^n$ is defined as follows: $x_1 \in \mathbb{R}^n$ is chosen arbitrarily, and

$$\begin{cases} y_n = J_\alpha^{\partial \gamma \|\cdot\|_1}(x_n - \alpha A^*(Ax_n - b)), \\ x_{n+1} = \alpha_n x_n + (1 - \alpha_n)((1 - \beta_n)y_n) \end{cases}$$

for each $n \in \mathbb{N}$, $\alpha \in (0, \frac{2}{\|A\|^2})$, and $\{\alpha_n, \beta_n\} \subset (0, 1)$. Assume that conditions (D) hold. Then $\lim_{n \to \infty} x_n = \bar{x}$, where $\bar{x} = P_{\Delta_2}(0)$.

Proof Let $\theta_n = \theta = 0$ for all $n \in \mathbb{N}$ and $V = I$ in Theorem 5.2, then Corollary 5.2 follows from Theorem 5.2. □

Apply Corollary 4.1, an iteration is used to find the solution to the split feasibility problem: Find $\bar{x} \in C$, $A\bar{x} \in Q$.

Theorem 5.3 Let $A : H \to H_1$ be a bounded linear operator, A^* be the adjoint of A. Let A_{f_1} be defined as (L4.2) in Lemma 2.4. Take $\mu \in \mathbb{R}$ such that $0 < \mu < \frac{2\bar{\gamma}}{L^2}$. Suppose that $\Pi_{13} := \{x \in C : Ax \in Q\} \neq \emptyset$. Let $f_1(x, y) = \langle y - x, A^*(A - P_Q A)x \rangle$. A sequence $\{x_n\} \subset H$ is defined as follows: $x_1 \in C$ is chosen arbitrarily, and

$$\begin{cases} y_n = J_\lambda^{A_{f_1}} P_C x_n, \\ x_{n+1} = \alpha_n x_n + (1 - \alpha_n)(\beta_n \theta_n + (I - \beta_n V)y_n) \end{cases}$$

for each $n \in \mathbb{N}$, $\alpha \in (0, \frac{2}{\|A\|^2})$, and $\{\alpha_n, \beta_n\} \subset (0,1)$. Assume that conditions (D) *hold. Then* $\lim_{n \to \infty} x_n = \bar{x}$, *where* $\bar{x} = P_{\Pi_{13}}(\bar{x} - V\bar{x})$. *Further, $\bar{x}$ is also the unique solution to the hierarchical variational inequality*:

$$\langle V\bar{x} - \theta, q - \bar{x} \rangle \geq 0, \quad \forall q \in \Pi_{13}.$$

Proof Let $g_1(x) = \frac{\|Ax - P_Q Ax\|_2^2}{2}$, then shows that g_1 is Fréchet differentiable with Fréchet derivative $\nabla g_1 = A^*(A - P_Q A)$, and ∇g_1 a Lipschitz function with Lipschitz constant $\|A\|^2$. Applying Corollary 4.1 and following the same argument as Theorem 5.1, we can prove Theorem 5.3. □

6 Image deblurring problem

This section mainly focuses on the image deblurring problems, which has received a lot of attention in recent years. Until now, some researchers have proposed many novel algorithms for this problem based on different deblurring models; for examples, see [26]. Now, by Corollary 5.2, we can consider the image deblurring problem.

All pixels of the original images described in the examples were first scaled into the range between 0 and 1.

The image went through a Gaussian blur of size 9×9 and standard deviation 4 (applied by the MATLAB functions imfilter and fspecial) followed by an additive zero-mean white Gaussian noise with standard deviation 10^{-3}. The original and observed images are given in Figures 1-3.

Remark 6.1 In the literature, we may observe that there are many fast algorithms for the image deblurring problem. Here, we show that we can also consider this problem by Corollary 5.2.

Figure 1 The original image.

Figure 2 The blurred image.

Figure 3 The deblurred image.
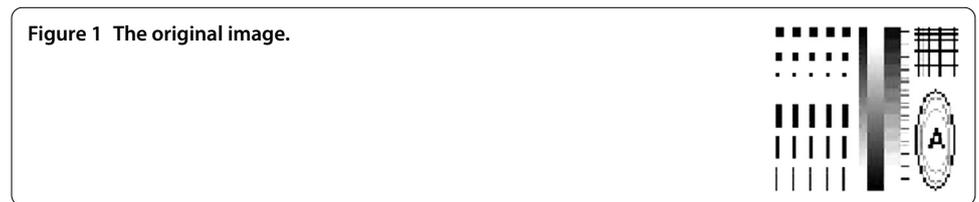
7 Conclusion and remarks

In this paper, we apply a recent fixed point theorem in [17] to study mathematical programming for the sum of two convex functions, mathematical programming of convex function, the split feasibility problem, and the lasso problem. We establish strong convergence theorems as regards these problems. The study of such problems will give many other applications in science, nonlinear analysis, and statistics.

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

Author details
[1]Department of Applied Mathematics, National Sun Yat Sen University, Kaohsiung, 700, Taiwan. [2]Department of Electronic Engineering, Nan Kai University of Technology, Nantou, 542, Taiwan. [3]Department of Mathematics, National Changhua University of Education, Changhua, 50058, Taiwan.

References
1. Tibshirani, R: Regression shrinkage and selection for the lasso. J. R. Stat. Soc., Ser. B **58**, 267-288 (1996)
2. Combettes, PL, Wajs, R: Signal recovery by proximal forward-backward splitting. Multiscale Model. Simul. **4**, 1168-1200 (2005)
3. Xu, HK: Properties and iterative methods for the lasso and its variants. Chin. Ann. Math., Ser. B **35**, 501-518 (2014)
4. Wang, Y, Xu, HK: Strong convergence for the proximal gradient methods. J. Nonlinear Convex Anal. **15**, 581-593 (2014)
5. Douglas, J, Rachford, HH: On the numerical solution of heat conduction in two and three space variable. Trans. Am. Math. Soc. **82**, 421-439 (1956)
6. Bauschken, HH, Combettes, PL: Convex Analysis and Monotone Operator Theory in Hilbert Space. Springer, Berlin (2011)
7. Tseng, P: Further applications of a splitting algorithm to decomposition in variational inequalities and convex programming. Math. Program., Ser. B **48**, 249-263 (1990)
8. Censor, Y, Elfving, T: A multiprojection algorithm using Bregman projection in a product space. Numer. Algorithms **8**, 221-239 (1994)
9. Byrne, C: Iterative oblique projection onto convex sets and the split feasibility problem. Inverse Probl. **18**, 441-453 (2002)
10. Byrne, C: A unified treatment of some iterative algorithms in signal processing and image reconstruction. Inverse Probl. **20**, 103-120 (2004)
11. 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 (2003)
12. López, G, Martín-Márquez, V, Xu, HK: Iterative algorithms for the multiple-sets split feasibility problem. In: Censor, Y, Jiang, M, Wang, G (eds.) Biomedical Mathematics: Promising Directions in Imaging, Therapy Planning and Inverse Problems, pp. 243-279. Medical Physics Publishing, Madison (2010)
13. Stark, H: Image Recovery: Theory and Applications. Academic Press, San Diego (1987)
14. Xu, HK: Iterative methods for the split feasibility problem in infinite-dimensional Hilbert spaces. Inverse Probl. **26**, 105018 (2010)
15. Rockafellar, TA: Monotone operators and the proximal point algorithm. SIAM J. Control Optim. **14**, 877-898 (1976)
16. Xu, HK: Averaged mappings and the gradient projection algorithm. J. Optim. Theory Appl. **150**, 360-378 (2011)
17. Yu, ZT, Lin, LJ: Hierarchical problems with applications to mathematical programming with multiple sets split feasibility constraints. Fixed Point Theory Appl. **2013**, 283 (2013)
18. Takahashi, W: Nonlinear Functional Analysis: Fixed Point Theory and Its Applications. Yokohama Publishers, Yokohama (2000)
19. Fan, K: A minimax inequalities and its applications. In: Shisha, O (ed.) Inequalities III, pp. 103-113. Academic Press, San Diego (1972)
20. Blum, E, Oettli, W: From optimization and variational inequalities to equilibrium problems. Math. Stud. **63**, 123-146 (1994)
21. Combettes, PL, Hirstoaga, SA: Equilibrium programming in Hilbert spaces. J. Nonlinear Convex Anal. **6**, 117-136 (2005)
22. Takahashi, S, Takahashi, W, Toyoda, M: Strong convergence theorems for maximal monotone operators with nonlinear mappings in Hilbert spaces. J. Optim. Theory Appl. **147**, 27-41 (2010)
23. Ekaland, I, Temam, R: Convex Analysis and Variational Problems. North-Holland, Amsterdam (1976)
24. Combettes, PL: Solving monotone inclusions via compositions of nonexpansive averaged operators. Optimization **53**(5-6), 475-504 (2004)
25. Baillon, JB, Haddad, G: Quelques propriétés des opérateurs angle-bornés et *n*-cycliquement monotones. Isr. J. Math. **26**, 137-150 (1977)
26. Chambolle, A: An algorithm for total variation minimization and applications. J. Math. Imaging Vis. **20**, 89-97 (2004)