Chuang and Chen *Journal of Inequalities and Applications* (2019) 2019:148
https://doi.org/10.1186/s13660-019-2084-9

Journal of Inequalities and Applications
a SpringerOpen Journal

RESEARCH **Open Access**

Split proximal linearized algorithm and convergence theorems for the split DC program

Chih-Sheng Chuang[1] and Chi-Ming Chen[2*]

*Correspondence:
ming@mail.nd.nthu.edu.tw
[2]Institute of Computational and Modeling Science, National Tsing Hua University, Hsinchu, Taiwan
Full list of author information is available at the end of the article

Abstract

In this paper, we study the split DC program by using the split proximal linearized algorithm. Further, linear convergence theorem for the proposed algorithm is established under suitable conditions. As applications, we first study the DC program (DCP). Finally, we give numerical results for the proposed convergence results.

MSC: 49J50; 49J53; 49M30; 49M37; 90C26

Keywords: DC program; Proximal linearized algorithm; Strongly monotonicity

1 Introduction

First, we recall the minimization problem for convex functions:

$$\text{Find } \bar{x} \in \arg\min\{f(x) : x \in H\}, \tag{MP1}$$

where H is a real Hilbert space and $f : H \to (-\infty, \infty]$ is a proper, lower semicontinuous, and convex function. This is a classical convex minimization problem with many applications. To study this problem, Martinet [11] introduced the proximal point algorithm

$$x_{n+1} = \arg\min_{y \in H}\left\{f(y) + \frac{1}{2\beta_n}\|y - x_n\|^2\right\}, \quad n \in \mathbb{N}, \tag{PPA}$$

and showed that $\{x_n\}_{n\in\mathbb{N}}$ converges weakly to a minimizer of f under suitable conditions. This algorithm is useful, however, only for convex problems, because the idea for this algorithm is based on the monotonicity of subdifferential operators of convex functions. So, it is important to consider the relation between nonconvex problems and a proximal point algorithm.

The following is a well-known nonconvex problem, known as DC program:

$$\text{Find } \bar{x} \in \arg\min_{x \in \mathbb{R}^n}\{f(x) = g(x) - h(x)\}, \tag{DCP}$$

where $g, h : \mathbb{R}^n \to \mathbb{R}$ are proper lower semicontinuous and convex functions. Here, the function f is called a DC function, and functions g and h are called DC components of f.

In the DC program, the convention $(+\infty) - (+\infty) = +\infty$ has been adopted to avoid the ambiguity $(+\infty)-(+\infty)$ that does not present any interest. It is well known that a necessary condition for $x \in \operatorname{dom}(f) := \{x \in \mathbb{R}^n : f(x) < \infty\}$ to be a local minimizer of f is $\partial h(x) \subseteq \partial g(x)$, where $\partial g(x)$ and $\partial h(x)$ are the subdifferentials of g and h, respectively (see Definition 2.4). But this condition is hard to be reached. So, many researchers focus their attention on finding points such that $\partial h(x) \cap \partial g(x) \neq \emptyset$, where x is called a critical point of f [8].

It is worth mentioning the richness of the class of DC functions which is a subspace containing the class of lower-$\mathcal{C}^2$ functions. In particular, $\mathcal{DC}(\mathbb{R}^n)$ contains the space $\mathcal{C}^{1,1}$ of functions whose gradient is locally Lipschitz continuous. Further, $\mathcal{DC}(\mathbb{R}^n)$ is closed under the operations usually considered in optimization. For example, a linear combination, a finite supremum, or the product of two DC functions remain DC. It is also known that the set of DC functions defined on a compact convex set of $\mathbb{R}^n$ is dense in the set of continuous functions on this set.

We also observed that the interest in the theory of DC functions has much increased in the last years. Some interesting optimality conditions and duality theorems related to the DC program have been given (for example, see [6, 7, 14]). Some algorithms for the DC program are proposed to analyze and solve a variety of highly structured and practical problems (for example, see [13]).

In 2003, Sun, Sampaio, and Candido [16] gave the following algorithm to study problem (DCP).

Algorithm 1.1 (Proximal point algorithm for (DCP) [16]) Let $\{\beta_n\}_{n\in\mathbb{N}}$ be a sequence in $(0,\infty)$, and let $g, h : \mathbb{R}^k \to \mathbb{R}$ be proper lower semicontinuous and convex functions. Let $\{x_n\}_{n\in\mathbb{N}}$ be generated by

$$\begin{cases} x_1 \in H_1 \quad \text{is chosen arbitrarily,} \\ \text{compute } w_n \in \partial h(x_n) \text{ and set } y_n = x_n + \beta_n w_n, \\ x_{n+1} := (I + \beta_n \partial g)^{-1}(y_n), \quad n \in \mathbb{N}, \\ \text{stop criteria: } x_{n+1} = x_n. \end{cases}$$

In 2016, Souza, Oliveira, and Soubeyran [15] gave the following algorithm to study the DC program.

Algorithm 1.2 (Proximal linearized algorithm for (DCP) [15]) Let $\{\beta_n\}_{n\in\mathbb{N}}$ be a sequence in $(0,\infty)$, and let $g, h : \mathbb{R}^k \to \mathbb{R}$ be proper lower semicontinuous and convex functions. Let $\{x_n\}_{n\in\mathbb{N}}$ be generated by

$$\begin{cases} x_1 \in H_1 \quad \text{is chosen arbitrarily,} \\ \text{compute } w_n \in \partial h(x_n), \\ x_{n+1} := \arg\min_{u\in H_1} \{g(u) + \frac{1}{2\beta_n}\|u - x_n\|^2 - \langle w_n, u - x_n\rangle\}, \quad n \in \mathbb{N}, \\ \text{stop criteria: } x_{n+1} = x_n. \end{cases}$$

In fact, if h is differentiable, then it is reduced to the following:

$$\begin{cases} x_1 \in H_1 \quad \text{is chosen arbitrarily}, \\ x_{n+1} := \arg\min_{u\in H_1}\{g(u) + \frac{1}{2\beta_n}\|u - x_n\|^2 - \langle \nabla h(x_n), u - x_n\rangle\}, \quad n \in \mathbb{N}. \\ \text{stop criteria: } x_{n+1} = x_n. \end{cases}$$

Further, Souza, Oliveira, and Soubeyran [15] gave the following convergence theorem for problem (DCP).

Theorem 1.1 ([15, Theorem 3]) *Let $g, h : \mathbb{R}^k \to \mathbb{R} \cup \{+\infty\}$ be proper, lower semicontinuous, and convex functions, and $g - h$ be bounded from below. Suppose that g is ρ-strongly convex, h is differentiable, and $\nabla h(x)$ is L-Lipschitz continuous. Let $\{\beta_n\}_{n\in\mathbb{N}}$ be a bounded sequence with $\liminf_{n\to\infty} \beta_n > 0$. Let $\{x_n\}_{n\in\mathbb{N}}$ be generated by Algorithm 1.2. If $\rho > 2L$, then $\{x_n\}_{n\in\mathbb{N}}$ converges linearly to a critical point $\bar{x}$ of problem (DCP).*

In this paper, we want to study the split DC program:

$$\text{Find } \bar{x} \in H_1 \text{ such that } \bar{x} \in \arg\min_{x\in H_1} f_1(x) \text{ and } A\bar{x} \in \arg\min_{y\in H_2} f_2(y), \tag{SDCP}$$

where H_1 and H_2 are real Hilbert spaces, $A : H_1 \to H_2$ is a nonzero linear and bounded mapping with adjoint operator A^*, $g_1, h_1 : H_1 \to \mathbb{R}$ are proper lower semicontinuous and convex functions, and $g_2, h_2 : H_2 \to \mathbb{R}$ are proper lower semicontinuous and convex functions, and $f_1(x) = g_1(x) - h_1(x)$ for all $x \in H_1$, and $f_2(y) = g_2(y) - h_2(y)$ for all $y \in H_2$.

Clearly, (SDCP) is a generalization of problem (DCP). Indeed, if $H_1 = H_2 = \mathbb{R}^n$, $A : \mathbb{R}^n \to \mathbb{R}^n$ is the identity mapping, $g_1 = g_2$, and $h_1 = h_2$, then problem (SDCP) is reduced to problem (DCP).

If $h_1(x) = 0$ and $h_2(y) = 0$ for all $x \in H_1$ and $y \in H_2$, then (SDCP) is reduced to the split minimization problems (SMP) for convex functions:

$$\text{Find } \bar{x} \in H_1 \text{ such that } g_1(\bar{x}) = \min_{u\in H_1} g_1(u) \text{ and } g_2(A\bar{x}) = \min_{v\in H_2} g_2(v), \tag{SMP}$$

where H_1 and H_2 are real Hilbert spaces, $A : H_1 \to H_2$ is a linear and bounded mapping with adjoint A^*, $g_1 : H_1 \to \mathbb{R}$ and $g_2 : H_2 \to \mathbb{R}$ are proper, lower semicontinuous, and convex functions. For example, one can see [4] and the related references.

If $H_1 = H_2 = H$ and $A : H \to H$ is the identity mapping, then problem (SMP) is reduced to the common minimization problem for convex functions:

$$\text{Find } \bar{x} \in H \text{ such that } g_1(\bar{x}) = \min_{u\in H} g_1(u) \text{ and } g_2(\bar{x}) = \min_{v\in H} g_2(v), \tag{CMP}$$

where H is a real Hilbert space, $g_1, g_2 : H \to \mathbb{R}$ are proper, lower semicontinuous, and convex functions. Further, if the solution set of problem (CMP) is nonempty, then problem (CMP) is equivalent to the following problem:

$$\text{Find } \bar{x} \in H \text{ such that } g_1(\bar{x}) + g_2(\bar{x}) = \min_{u\in H} g_1(u) + g_2(u), \tag{MP2}$$

where H is a real Hilbert space, $g_1, g_2 : H \to \mathbb{R}$ are proper, lower semicontinuous, and convex functions. This problem is well known and it has many important applications, including multiresolution sparse regularization, Fourier regularization, hard-constrained inconsistent feasibility, and alternating projection signal synthesis problems. For example, one can refer to [5, 9] and the related references.

On the other hand, Moudafi [12] introduced the split variational inclusion problem, which is a generalization of problem (SMP):

$$\text{Find } \bar{x} \in H_1 \text{ such that } 0_{H_1} \in B_1(\bar{x}) \text{ and } 0_{H_2} \in B_2(A\bar{x}), \tag{SVIP}$$

where H_1 and H_2 are real Hilbert spaces, $B_1 : H_1 \multimap H_1$ and $B_2 : H_2 \multimap H_2$ are set-valued maximal monotone mappings, $A : H_1 \to H_2$ is a linear and bounded operator, and A^* is the adjoint of A. Here, 0_{H_1} and 0_{H_2} are zero elements of real Hilbert spaces H_1 and H_2, respectively. To study problem (SVIP), Byrne et al. [3] gave the following algorithm and related convergence theorem in infinite dimensional Hilbert space.

Theorem 1.2 ([3]) *Let H_1 and H_2 be real Hilbert spaces, $A : H_1 \to H_2$ be a nonzero linear and bounded operator, and A^* denote the adjoint operator of A. Let $B_1 : H_1 \multimap H_1$ and $B_2 : H_2 \multimap H_2$ be set-valued maximal monotone mappings, $\beta > 0$, and $\gamma \in (0, \frac{2}{\|A\|^2})$. Let Ω be the solution set of (SVIP), and suppose that $\Omega \neq \emptyset$. Let $\{x_n\}_{n\in\mathbb{N}}$ be defined by*

$$x_{n+1} := J_\beta^{B_1}\bigl[x_n - \gamma A^*\bigl(I - J_\beta^{B_2}\bigr)Ax_n\bigr], \quad n \in \mathbb{N}.$$

Then $\{x_n\}$ converges weakly to an element $\bar{x} \in \Omega$.

If $B_1 = \partial g_1$ and $B_2 = \partial g_2$ (the subdifferential of g_i, $i = 1, 2$), then the algorithm given by Theorem 1.2 is reduced to the following algorithm:

$$\begin{cases} y_n = \arg\min_{z\in H_2}\{g(z) + \frac{1}{2\beta_n}\|z - Ax_n\|^2\}, \\ z_n = x_n - \gamma A^*(Ax_n - y_n), \\ x_{n+1} = \arg\min_{y\in H_1}\{g(y) + \frac{1}{2\beta_n}\|y - z_n\|^2\}, \quad n \in \mathbb{N}. \end{cases}$$

In this paper, motivated by the above works on DC programs and related problems, we want to study problem (SDCP) by using the split proximal linearized algorithm:

$$\begin{cases} x_1 \in H_1 \quad \text{is chosen arbitrarily}, \\ y_n := \arg\min_{v\in H_2}\{g_2(v) + \frac{1}{2\beta_n}\|v - Ax_n\|^2 - \langle\nabla h_2(Ax_n), v - Ax_n\rangle\}, \\ z_n := x_n - r_n A^*(Ax_n - y_n), \\ x_{n+1} := \arg\min_{u\in H_1}\{g_1(u) + \frac{1}{2\beta_n}\|u - z_n\|^2 - \langle\nabla h_1(z_n), u - z_n\rangle\}, \quad n \in \mathbb{N}, \end{cases}$$

where H_1 and H_2 are real Hilbert spaces, $A : H_1 \to H_2$ is a linear and bounded mapping with adjoint A^*, $g_1, h_1 : H_1 \to \mathbb{R}$ are proper lower semicontinuous and convex functions, and $g_2, h_2 : H_2 \to \mathbb{R}$ are proper lower semicontinuous and convex functions, g_1 and g_2 are strongly convex, h_1 and h_2 are Fréchet differentiable, ∇h_1 and ∇h_2 are L-Lipschitz continuous, and $f_1(x) = g_1(x) - h_1(x)$ for all $x \in H_1$, and $f_2(y) = g_2(y) - h_2(y)$ for all $y \in H_2$. Further, linear convergence theorems for the proposed algorithms are established under suitable conditions. Finally, we give numerical results for the proposed convergence theorems.

2 Preliminaries

Let H be a (real) Hilbert space with the inner product $\langle\cdot,\cdot\rangle$ and the norm $\|\cdot\|$. We denote the strong and weak convergence of $\{x_n\}_{n\in\mathbb{N}}$ to $x\in H$ by $x_n\to x$ and $x_n\rightharpoonup x$, respectively. For each $x,y,u,v\in H$ and $\lambda\in\mathbb{R}$, we have

$$\|x+y\|^2=\|x\|^2+2\langle x,y\rangle+\|y\|^2, \tag{2.1}$$

$$\big\|\lambda x+(1-\lambda)y\big\|^2=\lambda\|x\|^2+(1-\lambda)\|y\|^2-\lambda(1-\lambda)\|x-y\|^2, \tag{2.2}$$

$$2\langle x-y,u-v\rangle=\|x-v\|^2+\|y-u\|^2-\|x-u\|^2-\|y-v\|^2. \tag{2.3}$$

Definition 2.1 Let H be a real Hilbert space, $B:H\to H$ be a mapping, and $\rho>0$. Thus,

(i) B is monotone if $\langle x-y,Bx-By\rangle\geq 0$ for all $x,y\in H$.
(ii) B is ρ-strongly monotone if $\langle x-y,Bx-By\rangle\geq\rho\|x-y\|^2$ for all $x,y\in H$.

Definition 2.2 Let H be a real Hilbert space and $B:H\multimap H$ be a set-valued mapping with domain $\mathcal{D}(B):=\{x\in H:B(x)\neq\emptyset\}$. Thus,

(i) B is called monotone if $\langle u-v,x-y\rangle\geq 0$ for any $u\in B(x)$ and $v\in B(y)$.
(ii) B is maximal monotone if its graph $\{(x,y):x\in\mathcal{D}(B),y\in B(x)\}$ is not properly contained in the graph of any other monotone mapping.
(iii) B is ρ-strongly monotone if $\langle x-y,u-v\rangle\geq\rho\|x-y\|^2$ for all $x,y\in H$ and all $u\in B(x)$, and $v\in B(y)$.

Definition 2.3 Let H be a real Hilbert space, and $f:H\to\mathbb{R}$ be a function. Thus,

(i) f is proper if $\operatorname{dom}(f):=\{x\in H:f(x)<\infty\}\neq\emptyset$.
(ii) f is lower semicontinuous if $\{x\in H:f(x)\leq r\}$ is closed for each $r\in\mathbb{R}$.
(iii) f is convex if $f(tx+(1-t)y)\leq tf(x)+(1-t)f(y)$ for every $x,y\in H$ and $t\in[0,1]$.
(iv) f is ρ-strongly convex ($\rho>0$) if

$$f\big(tx+(1-t)y\big)+\frac{\rho}{2}\cdot t(1-t)\|x-y\|^2\leq tf(x)+(1-t)f(y)$$

for all $x,y\in H$ and $t\in(0,1)$.
(v) f is Gâteaux differentiable at $x\in H$ if there is $\nabla f(x)\in H$ such that

$$\lim_{t\to 0}\frac{f(x+ty)-f(x)}{t}=\big\langle y,\nabla f(x)\big\rangle$$

for each $y\in H$.
(vi) f is Fréchet differentiable at x if there is $\nabla f(x)$ such that

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

Remark 2.1 Let H be a real Hilbert space. Then $f(x):=\|x\|^2$ is a 2-strongly convex function. Besides, we know $g:H\to\mathbb{R}$ is ρ-strongly convex if and only if $g-\frac{\rho}{2}\|\cdot\|^2$ is convex [1, Proposition 10.6].

Example 2.1 Let $g(x):=\frac{1}{2}\langle Qx,x\rangle-\langle x,b\rangle$, where $Q\in\mathbb{R}^{n\times n}$ is a real symmetric positive definite matrix and $b\in\mathbb{R}^n$. Then g is a strongly convex function.

Definition 2.4 Let $f : H \to (-\infty, \infty]$ be a proper lower semicontinuous and convex function. Then the subdifferential ∂f of f is defined by

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

for each $x \in H$.

Lemma 2.1 *Let $f : H \to (-\infty, \infty]$ be a proper lower semicontinuous and convex function. Then the following are satisfied*:

(i) *∂f is a set-valued maximal monotone mapping.*

(ii) *f is Gâteaux differentiable at $x \in \operatorname{int}(\mathcal{D})$ if and only if $\partial f(x)$ consists of a single element. That is, $\partial f(x) = \{\nabla f(x)\}$* [2, Proposition 1.1.10].

(iii) *Suppose that f is Fréchet differentiable. Then f is convex if and only if ∇f is a monotone mapping.*

Lemma 2.2 ([1, Example 22.3(iv)]) *Let $\rho > 0$, H be a real Hilbert space and $f : H \to \mathbb{R}$ be a proper, lower-semicontinuous, and convex function. If f is ρ-strongly convex, then ∂f is ρ-strongly monotone.*

Lemma 2.3 ([1, Proposition 16.26]) *Let H be a real Hilbert space and $f : H \to (\infty, \infty]$ be a proper, lower semicontinuous, and convex function. If $\{u_n\}_{n\in\mathbb{N}}$ and $\{x_n\}_{n\in\mathbb{N}}$ are sequences in H with $u_n \in \partial f(x_n)$ for all $n \in \mathbb{N}$, and $x_n \rightharpoonup x$ and $u_n \to u$, then $u \in \partial f(x)$.*

Lemma 2.4 ([17, p. 114]) *Let $\{a_n\}_{n\in\mathbb{N}}$ and $\{b_n\}_{n\in\mathbb{N}}$ be sequences of nonnegative real numbers. If $\sum_{n=1}^{\infty} a_n = \infty$ and $\sum_{n=1}^{\infty} a_n b_n < \infty$, then $\liminf_{n\to\infty} b_n = 0$.*

Lemma 2.5 ([10]) *Let H be a real Hilbert space, $B : H \multimap H$ be a set-valued maximal monotone mapping, $\beta > 0$, and J_β^B be defined by $J_\beta^B(x) := (I + \beta B)^{-1}(x)$ for each $x \in H$. Then J_β^B is a single-valued mapping.*

3 Split proximal linearized algorithm

Throughout this section, we use the following notations and assumptions. Let $\rho \geq L > 0$. Let H_1 and H_2 be finite dimensional real Hilbert spaces, $A : H_1 \to H_2$ be a nonzero linear and bounded mapping, A^* be the adjoint of A, $g_1, h_1 : H_1 \to \mathbb{R}$ be proper lower semicontinuous and convex functions, $g_2, h_2 : H_2 \to \mathbb{R}$ be proper lower semicontinuous and convex functions, $f_1(x) = g_1(x) - h_1(x)$ for all $x \in H_1$, and $f_2(y) = g_2(y) - h_2(y)$ for all $y \in H_2$. Further, we assume that f_1 and f_2 are bounded from below, h_1 and h_2 are Fréchet differentiable, ∇h_1 and ∇h_2 are L-Lipschitz continuous, g_1 and g_2 are ρ-strongly convex. Let $\beta \in (0, \infty)$, and let $\{\beta_n\}_{n\in\mathbb{N}}$ be a sequence in $[a, b] \subseteq (0, \infty)$. Let $r \in (0, \frac{1}{\|A\|^2})$ and $\{r_n\}_{n\in\mathbb{N}}$ be a sequence in $(0, \frac{1}{\|A\|^2})$. Let Ω_{SDCP} be defined by

$$\Omega_{\mathrm{SDCP}} := \{x \in H_1 : \nabla h_1(x) \in \partial g_1(x), \nabla h_2(Ax) \in \partial g_2(Ax)\},$$

and assume that $\Omega_{\mathrm{SDCP}} \neq \emptyset$.

Proposition 3.1 *If $\rho > L$ and $\Omega_{\mathrm{SDCP}} \neq \emptyset$, then the set Ω_{SDCP} is a singleton.*

Proof If $x, y \in \Omega_{\mathrm{SDCP}}$, then

$$\begin{aligned}&\nabla h_1(x) \in \partial g_1(x), \qquad \nabla h_1(y) \in \partial g_1(y),\\&\nabla h_2(Ax) \in \partial g_2(Ax), \qquad \nabla h_2(Ay) \in \partial g_2(Ay).\end{aligned}$$

Since g_1 is ρ-strongly convex, we know ∂g_1 is ρ-strongly monotone. Thus,

$$\rho\|x-y\|^2 \le \langle x-y, \nabla h_1(x) - \nabla h_1(y)\rangle \le \|x-y\| \cdot \big\|\nabla h_1(x) - \nabla h_1(y)\big\|.$$

Since ∇h_1 is L-Lipschitz continuous, we have

$$\rho\|x-y\|^2 \le \|x-y\| \cdot \big\|\nabla h_1(x) - \nabla h_1(y)\big\| \le L\|x-y\|^2.$$

Since $\rho > L$, we have $x = y$. The proof is completed. □

In this section, we study the split DC program by the following algorithm.

Algorithm 3.1 (Split proximal linearized algorithm)

$$\begin{cases} x_1 \in H_1 \quad \text{is chosen arbitrarily,}\\ y_n := \arg\min_{v\in H_2}\{g_2(v) + \frac{1}{2\beta_n}\|v - Ax_n\|^2 - \langle\nabla h_2(Ax_n), v - Ax_n\rangle\},\\ z_n := x_n - r_n A^*(Ax_n - y_n),\\ x_{n+1} := \arg\min_{u\in H_1}\{g_1(u) + \frac{1}{2\beta_n}\|u - z_n\|^2 - \langle\nabla h_1(z_n), u - z_n\rangle\}, \quad n \in \mathbb{N}.\end{cases}$$

Theorem 3.1 *Let* $\{r_n\}_{n\in\mathbb{N}}$ *be a sequence in* $(0, \frac{1}{\|A\|^2})$ *with* $0 < \liminf_{n\to\infty} r_n \le \limsup_{n\to\infty} r_n < \frac{1}{\|A\|^2}$. *Let* $\{x_n\}_{n\in\mathbb{N}}$ *be generated by Algorithm* 3.1. *Then* $\{x_n\}_{n\in\mathbb{N}}$ *converges to an element* $\bar{x} \in \Omega_{\mathrm{SDCP}}$.

Proof Take any $w \in \Omega_{\mathrm{SDCP}}$ and $n \in \mathbb{N}$, and let w and n be fixed. First, from the second line of Algorithm 3.1, we get

$$0 \in \partial g_2(y_n) + \frac{1}{\beta_n}(y_n - Ax_n) - \nabla h_2(Ax_n). \tag{3.1}$$

By (3.1), there exists $u_n \in \partial g_2(y_n)$ such that

$$\nabla h_2(Ax_n) = u_n + \frac{1}{\beta_n}(y_n - Ax_n). \tag{3.2}$$

Since $w \in \Omega_{\mathrm{SDCP}}$, we know that $\nabla h_2(Aw) \in \partial g_2(Aw)$. By Lemma 2.2, ∂g_2 is ρ-strongly monotone, and then

$$0 \le \langle y_n - Aw, u_n - \nabla h_2(Aw)\rangle - \rho\|y_n - Aw\|^2. \tag{3.3}$$

By (3.2) and (3.3),

$$0 \le \left\langle y_n - Aw, \nabla h_2(Ax_n) + \frac{1}{\beta_n}(Ax_n - y_n) - \nabla h_2(Aw)\right\rangle - \rho\|y_n - Aw\|^2. \tag{3.4}$$

Hence, by (3.4), we have

$$\begin{aligned}
0 &\le 2\beta_n\bigl\langle y_n - Aw, \nabla h_2(Ax_n) - \nabla h_2(Aw)\bigr\rangle + 2\langle y_n - Aw, Ax_n - y_n\rangle \\
&\quad - 2\beta_n\rho\|y_n - Aw\|^2 \\
&\le 2\beta_n L\|y_n - Aw\| \cdot \|Ax_n - Aw\| - 2\beta_n\rho\|y_n - Aw\|^2 \\
&\quad + \|Ax_n - Aw\|^2 - \|y_n - Ax_n\|^2 - \|y_n - Aw\|^2 \\
&\le \beta_n L\|y_n - Aw\|^2 + \beta_n L\|Ax_n - Aw\|^2 - 2\beta_n\rho\|y_n - Aw\|^2 \\
&\quad + \|Ax_n - Aw\|^2 - \|y_n - Ax_n\|^2 - \|y_n - Aw\|^2.
\end{aligned} \tag{3.5}$$

By (3.5), we obtain

$$\begin{aligned}
\|y_n - Aw\|^2 &\le \frac{\beta_n L + 1}{1 + 2\beta_n\rho - \beta_n L}\|Ax_n - Aw\|^2 - \frac{\|y_n - Ax_n\|^2}{1 + 2\beta_n\rho - \beta_n L} \\
&\le \|Ax_n - Aw\|^2 - \frac{\|y_n - Ax_n\|^2}{1 + 2\beta_n\rho - \beta_n L}.
\end{aligned} \tag{3.6}$$

In the same way, one obtains

$$\|x_{n+1} - w\|^2 \le \|z_n - w\|^2 - \frac{1}{1 + 2\beta_n\rho - \beta_n L}\|x_{n+1} - z_n\|^2 \le \|z_n - w\|^2. \tag{3.7}$$

Next, we have

$$\begin{aligned}
2\|z_n - w\|^2 &= 2\bigl\langle z_n - w, x_n - r_n A^*(Ax_n - y_n) - w\bigr\rangle \\
&= 2\langle z_n - w, x_n - w\rangle - 2r_n\bigl\langle z_n - w, A^*(Ax_n - y_n)\bigr\rangle \\
&= 2\langle z_n - w, x_n - w\rangle - 2r_n\langle Az_n - Aw, Ax_n - y_n\rangle \\
&= \|z_n - w\|^2 + \|x_n - w\|^2 - \|x_n - z_n\|^2 - r_n\|Az_n - y_n\|^2 \\
&\quad - r_n\|Ax_n - Aw\|^2 + r_n\|Az_n - Ax_n\|^2 + r_n\|y_n - Aw\|^2.
\end{aligned} \tag{3.8}$$

By (3.6), (3.7), and (3.8),

$$\begin{aligned}
&\|x_{n+1} - w\|^2 \\
&\quad \le \|z_n - w\|^2 \\
&\quad = \|x_n - w\|^2 - \|x_n - z_n\|^2 - r_n\|Az_n - y_n\|^2 \\
&\qquad - r_n\|Ax_n - Aw\|^2 + r_n\|Az_n - Ax_n\|^2 + r_n\|y_n - Aw\|^2 \\
&\quad \le \|x_n - w\|^2 - \|x_n - z_n\|^2 - r_n\|Az_n - y_n\|^2 - r_n\|Ax_n - Aw\|^2 \\
&\qquad + r_n\|A\|^2 \cdot \|z_n - x_n\|^2 + r_n \cdot \frac{\beta_n L + 1}{1 + 2\beta_n\rho - \beta_n L}\|Ax_n - Aw\|^2 \\
&\quad = \|x_n - w\|^2 - \bigl(1 - r_n\|A\|^2\bigr)\|x_n - z_n\|^2 - r_n\|Az_n - y_n\|^2 \\
&\qquad - r_n\left(1 - \frac{\beta_n L + 1}{1 + 2\beta_n\rho - \beta_n L}\right)\|Ax_n - Aw\|^2
\end{aligned}$$

$$= \|x_n - w\|^2 - \bigl(1 - r_n\|A\|^2\bigr)\|x_n - z_n\|^2 - r_n\|Az_n - y_n\|^2$$
$$- r_n\left(\frac{2\beta_n(\rho - L)}{1 + 2\beta_n\rho - \beta_n L}\right)\|Ax_n - Aw\|^2$$
$$\le \|x_n - w\|^2. \tag{3.9}$$

By (3.9), $\lim_{n\to\infty}\|x_n - w\|$ exists and $\{x_n\}_{n\in\mathbb{N}}$ is a bounded sequence. Further, $\{Ax_n\}_{n\in\mathbb{N}}$, $\{y_n\}_{n\in\mathbb{N}}$, $\{z_n\}_{n\in\mathbb{N}}$ are bounded sequences. By (3.9) again, we know that

$$\lim_{n\to\infty}\|x_n - w\| = \lim_{n\to\infty}\|z_n - w\|, \tag{3.10}$$

and

$$\lim_{n\to\infty}\frac{\|x_{n+1} - z_n\|^2}{1 + 2\beta_n\rho - \beta_n L} = \lim_{n\to\infty} r_n\|Az_n - y_n\|^2 = \lim_{n\to\infty}\bigl(1 - r_n\|A\|^2\bigr)\|x_n - z_n\|^2 = 0. \tag{3.11}$$

It follows from $\{\beta_n\}_{n\in\mathbb{N}} \subseteq (a,b)$, $0 < \liminf_{n\to\infty} r_n \le \limsup_{n\to\infty} r_n < \frac{1}{\|A\|^2}$, and (3.11) that

$$\lim_{n\to\infty}\|x_{n+1} - z_n\| = \lim_{n\to\infty}\|Az_n - y_n\| = \lim_{n\to\infty}\|x_n - z_n\| = 0. \tag{3.12}$$

Since $\{x_n\}_{n\in\mathbb{N}}$ is bounded, there exists a subsequence $\{x_{n_k}\}_{k\in\mathbb{N}}$ of $\{x_n\}_{n\in\mathbb{N}}$ such that $x_{n_k} \to \bar{x} \in H_1$. Clearly, $Ax_{n_k} \to A\bar{x}$, $z_{n_k} \to \bar{x}$, $Az_{n_k} \to A\bar{x}$, $y_{n_k} \to A\bar{x}$, and $x_{n_k+1} \to \bar{x}$. Further, by (3.2), we obtain

$$\nabla h_2(Ax_{n_k}) \in \partial g_2(y_{n_k}) + \frac{1}{\beta_{n_k}}(y_{n_k} - Ax_{n_k}). \tag{3.13}$$

By (3.12), (3.13), Lemma 2.3, and $\{\beta_n\}_{n\in\mathbb{N}} \subseteq (a,b)$, we determine that

$$\nabla h_2(A\bar{x}) \in \partial g_2(A\bar{x}). \tag{3.14}$$

Similarly, we have

$$\nabla h_1(\bar{x}) \in \partial g_1(\bar{x}). \tag{3.15}$$

By (3.14) and (3.15), we know that $\bar{x} \in \Omega_{\mathrm{SDCP}}$. Further, $\lim_{n\to\infty}\|x_n - \bar{x}\| = \lim_{k\to\infty}\|x_{n_k} - \bar{x}\| = 0$. Therefore, the proof is completed. □

Remark 3.1

(i) In Algorithm 3.1, if $y_n = Ax_n$ and $x_{n+1} = z_n$, then $x_n = z_n$, and this implies that $\nabla h_1(x_n) \in \partial g_1(x_n)$ and $\nabla h_2(Ax_n) \in \partial g_2(Ax_n)$. Thus, $x_n \in \Omega_{\mathrm{SDCP}}$.

(ii) In Algorithm 3.1, if $x_{n+1} \neq z_n$, then $f_1(x_{n+1}) < f_1(z_n)$. Indeed, it follows from $\partial h_1(z_n) = \{\nabla h_1(z_n)\}$ and the definition of x_{n+1} that

$$g_1(x_{n+1}) - h_1(x_{n+1}) + \frac{1}{2\beta_n}\|x_{n+1} - z_n\|^2 \le g_1(z_n) - h_1(z_n).$$

So, if $x_{n+1} \neq z_n$, then $f_1(x_{n+1}) < f_1(z_n)$.

(iii) In Algorithm 3.1, if $y_n \neq Ax_n$, then $f_2(y_n) < f_2(Ax_n)$. Indeed, it follows from $\partial h_2(Ax_n) = \{\nabla h_2(Ax_n)\}$ and the definition of y_n that

$$g_2(y_n) - h_2(y_n) + \frac{1}{2\beta_n}\|y_n - Ax_n\|^2 \leq g_2(Ax_n) - h_2(Ax_n).$$

So, if $y_n \neq Ax_n$, then $f_2(y_n) < f_2(Ax_n)$.

(iv) If $\rho > L$, then it follows from (3.7) that (3.9) can be rewritten as

$$\|x_{n+1} - w\|^2 \leq k_n\|z_n - w\|^2 \leq k_n\|x_n - w\|^2,$$

where

$$k_n := \frac{1 + \beta_n L}{1 + 2\beta_n\rho - \beta_n L} \in (0,1).$$

Hence, $\{x_n\}_{n\in\mathbb{N}}$ converges linearly to $\bar{x}$, where $\Omega_{\mathrm{SDCP}} = \{\bar{x}\}$.

Remark 3.2 From the proof of Theorem 3.1, we know that

$$\nabla h_2(Ax_n) + \frac{1}{\beta_n}(Ax_n - y_n) \in \partial g_2(y_n), \tag{3.16}$$

and this implies that

$$Ax_n + \beta_n\nabla h_2(Ax_n) \in y_n + \beta_n\partial g_2(y_n) = (I_{H_2} + \beta_n\partial g_2)(y_n), \tag{3.17}$$

where I_{H_2} is the identity mapping on H_2. Since g_2 is proper, lower semicontinuous, and convex, we know that ∂g_2 is maximal monotone. So, by Lemma 2.5, we determine that

$$y_n = (I_{H_2} + \beta_n\partial g_2)^{-1}\big(Ax_n + \beta_n\nabla h_2(Ax_n)\big). \tag{3.18}$$

Similarly, we have

$$x_{n+1} = (I_{H_1} + \beta_n\partial g_1)^{-1}\big(z_n + \beta_n\nabla h_1(z_n)\big), \tag{3.19}$$

where I_{H_1} is the identity mapping on H_1. Therefore, Algorithm 3.1 can be rewritten as the following algorithm:

$$\begin{cases} y_n := (I_{H_2} + \beta_n\partial g_2)^{-1}(Ax_n + \beta_n\nabla h_2(Ax_n)), \\ z_n := x_n - r_n A^*(Ax_n - y_n), \\ x_{n+1} := (I_{H_1} + \beta_n\partial g_1)^{-1}(z_n + \beta_n\nabla h_1(z_n)), \quad n \in \mathbb{N}. \end{cases} \tag{Algorithm 3.2}$$

In fact, we observe that the idea of Algorithm 3.2 is the same as the proposed algorithm by Sun, Sampaio, and Candido [16, Algorithm 2.3]. Hence, a modified algorithm and related convergence theorems could be presented by using the idea of [16, Algorithm 5.3].

Remark 3.3 Under the assumptions in this section, consider the following:

$$\begin{cases} y := \arg\min_{v\in H_2}\{g_2(v) + \frac{1}{2\beta}\|v - Ax\|^2 - \langle \nabla h_2(Ax), v - Ax\rangle\}, \\ z := x - rA^*(Ax - y), \\ w := \arg\min_{u\in H_1}\{g_1(u) + \frac{1}{2\beta}\|u - z\|^2 - \langle \nabla h_1(z), u - z\rangle\}, \end{cases} \tag{3.20}$$

that is,

$$\begin{cases} y := (I_{H_2} + \beta\partial g_2)^{-1}(Ax + \beta\nabla h_2(Ax)), \\ z := x - rA^*(Ax - y), \\ w := (I_{H_1} + \beta\partial g_1)^{-1}(z + \beta\nabla h_1(z)), \end{cases} \tag{3.21}$$

we know that

$$Ax = y \quad \text{and} \quad z = w \quad \Leftrightarrow \quad x = z \in \Omega_{\mathrm{SDCP}}. \tag{3.22}$$

Proof For this equivalence relation, we only need to show that $x = z \in \Omega_{\mathrm{SDCP}}$ implies that $Ax = y$ and $z = w$. Indeed, since $x = z \in \Omega_{\mathrm{SDCP}}$, we know that $\nabla h_1(z) \in \partial g_1(z)$ and $\nabla h_2(Ax) \in \partial g_2(Ax)$. By Lemma 2.5,

$$\begin{cases} (I_{H_2} + \beta\partial g_2)^{-1}(Ax + \beta\nabla h_2(Ax)) = Ax, \\ (I_{H_1} + \beta\partial g_1)^{-1}(z + \beta\nabla h_1(z)) = z. \end{cases} \tag{3.23}$$

By (3.21) and (3.23), we know that $Ax = y$ and $z = w$. □

Remark 3.4 In Algorithm 3.1, if $\beta_n = \beta$ and $r_n = r$ for each $n \in \mathbb{N}$, and $x_{N+1} = x_N$ for some $N \in \mathbb{N}$, then $x_n = x_N$, $y_n = y_N$, and $z_n = z_N$ for each $n \in \mathbb{N}$ with $n \geq N$. By Theorem 3.1, we know that $\lim_{n\to\infty} x_n = x_N \in \Omega_{\mathrm{SDCP}}$. So, $x_{n+1} = x_n$ could be set as a stop criterion in Algorithm 3.1. Further, from (3.21), we have

$$\begin{aligned} x = w \quad &\Rightarrow \quad x \in \Omega_{\mathrm{SDCP}} \\ &\Rightarrow \quad x \in \Omega_{\mathrm{SDCP}} \quad \text{and} \quad y = Ax \\ &\Rightarrow \quad x = z \in \Omega_{\mathrm{SDCP}} \quad \text{and} \quad y = Ax \\ &\Rightarrow \quad x = z = w \in \Omega_{\mathrm{SDCP}} \quad \text{and} \quad y = Ax \\ &\Rightarrow \quad x = w. \end{aligned}$$

This equivalence relation is important for the split DC program.

By Remark 3.4, we give the following result.

Proposition 3.2 *Under the assumptions in this section, and*

$$\begin{cases} y := \arg\min_{v\in H_2}\{g_2(v) + \frac{1}{2\beta}\|v - Ax\|^2 - \langle \nabla h_2(Ax), v - Ax\rangle\}, \\ z := x - rA^*(Ax - y), \\ w := \arg\min_{u\in H_1}\{g_1(u) + \frac{1}{2\beta}\|u - z\|^2 - \langle \nabla h_1(z), u - z\rangle\}. \end{cases} \tag{3.24}$$

Then $x \in \Omega_{\mathrm{SDCP}}$ if and only if $x = w$.

Next, we give another convergence theorem for the split proximal linearized algorithm under different assumptions on $\{r_n\}_{n\in\mathbb{N}}$.

Theorem 3.2 *Let $\{r_n\}_{n\in\mathbb{N}}$ be a sequence in $(0, \frac{1}{\|A\|^2})$ with $\lim_{n\to\infty} r_n = 0$ and $\sum_{n=1}^{\infty} r_n = \infty$. Let $\{x_n\}_{n\in\mathbb{N}}$ be generated by Algorithm 3.1. Then $\{x_n\}_{n\in\mathbb{N}}$ converges to an element $\bar{x} \in \Omega_{\mathrm{SDCP}}$.*

Proof Take any $w \in \Omega_{\mathrm{SDCP}}$ and $n \in \mathbb{N}$, and let w and n be fixed. From the proof of Theorem 3.1, we have

$$\nabla h_2(Ax_n) - \frac{1}{\beta_n}(y_n - Ax_n) \in \partial g_2(Ax_n), \tag{3.25}$$

$$\nabla h_1(z_n) - \frac{1}{\beta_n}(x_{n+1} - z_n) \in \partial g_1(x_{n+1}), \tag{3.26}$$

and

$$\begin{aligned} \|x_{n+1} - w\|^2 &\le \|x_n - w\|^2 - \bigl(1 - r_n\|A\|^2\bigr)\|x_n - z_n\|^2 - r_n\|Az_n - y_n\|^2 \\ &\quad - r_n\left(\frac{2\beta_n(\rho - L)}{1 + 2\beta_n\rho - \beta_n L}\right)\|Ax_n - Aw\|^2 - \frac{1}{1 + 2\beta_n\rho - \beta_n L}\|x_{n+1} - z_n\|^2 \\ &\le \|x_n - w\|^2. \end{aligned} \tag{3.27}$$

Further, the following are satisfied:

$$\begin{cases} \lim_{n\to\infty}\|x_n - w\| \text{ exists}, \\ \{x_n\}_{n\in\mathbb{N}}, \{Ax_n\}_{n\in\mathbb{N}}, \{y_n\}_{n\in\mathbb{N}}, \{z_n\}_{n\in\mathbb{N}} \text{ are bounded sequences}, \\ \lim_{n\to\infty}\|x_{n+1} - z_n\| = 0, \\ \lim_{n\to\infty}(1 - r_n\|A\|^2)\|x_n - z_n\|^2 = 0. \end{cases}$$

Since $\lim_{n\to\infty} r_n = 0$, we know that

$$\lim_{n\to\infty}\|x_n - z_n\| = 0. \tag{3.28}$$

By (3.27), we have

$$\sum_{n=1}^{\infty} r_n\|Az_n - y_n\|^2 \le \sum_{n=1}^{\infty}\bigl(\|x_n - w\|^2 - \|x_{n+1} - w\|^2\bigr) < \infty. \tag{3.29}$$

By (3.29) and Lemma 2.4, we determine that

$$\liminf_{n\to\infty} \|Az_n - y_n\|^2 = 0. \tag{3.30}$$

Then there exist a subsequence $\{y_{n_k}\}_{k\in\mathbb{N}}$ of $\{y_n\}_{n\in\mathbb{N}}$, a subsequence $\{z_{n_k}\}_{k\in\mathbb{N}}$ of $\{z_n\}_{n\in\mathbb{N}}$, and $\bar{x} \in H_1$ such that $z_{n_k} \to \bar{x}$, $y_{n_k} \to A\bar{x}$, and

$$\liminf_{n\to\infty} \|Az_n - y_n\|^2 = \lim_{k\to\infty} \|Az_{n_k} - y_{n_k}\|^2 = 0. \tag{3.31}$$

Clearly, $x_{n_k} \to \bar{x}$, $x_{n_k+1} \to \bar{x}$, and $Ax_{n_k} \to A\bar{x}$. By (3.25) and (3.26), we know that $\bar{x} \in \Omega_{\mathrm{SDCP}}$. Thus, $\bar{x} = \hat{x}$. Since $\lim_{n\to\infty} \|x_n - \bar{x}\|$ exists, we know $\lim_{n\to\infty} \|x_n - \bar{x}\| = \lim_{k\to\infty} \|x_{n_k} - \bar{x}\| = 0$. Therefore, the proof is completed. □

4 Application to the DC program and numerical results

Let $\rho > L \geq 0$. Let H be a finite dimensional Hilbert space, $g, h : H \to \mathbb{R}$ be proper, lower semicontinuous, and convex functions. Besides, we also assume that h is Fréchet differentiable, ∇h is L-Lipschitz continuous, g is ρ-strongly convex. Let $\{\beta_n\}_{n\in\mathbb{N}}$ be a sequence in $(a, b) \subseteq (0, \infty)$. Let $\{r_n\}_{n\in\mathbb{N}}$ be a sequence in $(0, 1)$ with $0 < \liminf_{n\to\infty} r_n \leq \limsup_{n\to\infty} r_n < 1$.

Now, we recall the DC program:

$$\text{Find } \bar{x} \in \arg\min_{x\in H} \bigl\{ f(x) = g(x) - h(x) \bigr\}. \tag{DCP}$$

Let Ω_{DCP} be defined by

$$\Omega_{\mathrm{DCP}} := \bigl\{ x \in H : \nabla h(x) \in \partial g(x) \bigr\},$$

and assume that $\Omega_{\mathrm{DCP}} \neq \emptyset$. If $H_1 = H_2 = H$, $g_1 = g_2 = g$, and $h_1 = h_2 = h$, then we get the following algorithm and convergence theorem from Algorithm 3.1 and Theorem 3.1, respectively.

Algorithm 4.1

$$\begin{cases} x_1 \in H \quad \text{is chosen arbitrarily}, \\ y_n := \arg\min_{v\in H} \{ g(v) + \frac{1}{2\beta_n} \|v - x_n\|^2 - \langle \nabla h(x_n), v - x_n \rangle \}, \\ z_n := (1 - r_n)x_n + r_n y_n, \\ x_{n+1} := \arg\min_{u\in H} \{ g(u) + \frac{1}{2\beta_n} \|u - z_n\|^2 - \langle \nabla h(z_n), u - z_n \rangle \}, \quad n \in \mathbb{N}. \end{cases}$$

Theorem 4.1 *Let $\{x_n\}_{n\in\mathbb{N}}$ be generated by Algorithm 4.1. Then $\{x_n\}_{n\in\mathbb{N}}$ converges to an element $\bar{x} \in \Omega_{\mathrm{DCP}}$.*

In fact, we can get the following algorithm and convergence theorem by Algorithm 4.1 and Theorem 4.1, respectively.

Algorithm 4.2

$$\begin{cases} x_1 \in H \quad \text{is given,} \\ z_n := \arg\min_{u\in H}\{g(u) + \frac{1}{2\beta_n}\|u - x_n\|^2 - \langle \nabla h(x_n), u - x_n\rangle\}, \\ y_n := \arg\min_{v\in H}\{g(v) + \frac{1}{2\beta_n}\|v - z_n\|^2 - \langle \nabla h(z_n), v - z_n\rangle\}, \\ x_{n+1} := (1 - r_n)z_n + r_n y_n, \quad n \in \mathbb{N}. \end{cases}$$

Theorem 4.2 *Let $\{x_n\}_{n\in\mathbb{N}}$ be generated by Algorithm 4.2. Then $\{x_n\}_{n\in\mathbb{N}}$ converges to an element $\bar{x} \in \Omega_{\mathrm{DCP}}$.*

Example 4.1 Let $g, h : \mathbb{R}^3 \to \mathbb{R}$ be defined by $g(x_1, x_2, x_3) := 2x_1^2 + 2x_2^2 + 2x_3^2$ and $h(x_1, x_2, x_3) := 4x_1 + 8x_2 + 12x_3$ for all $(x_1, x_2, x_3) \in \mathbb{R}^3$. Then $\Omega_{\mathrm{DCP}} := \{x \in H : \nabla h(x) \in \partial g(x)\} = \{(1, 2, 3)\}$.

Example 4.2 Let $g_1, h_1 : \mathbb{R}^3 \to \mathbb{R}$ be defined by $g_1(x_1, x_2, x_3) := 2x_1^2 + 2x_2^2 + 2x_3^2$ and $h_1(x_1, x_2, x_3) := 4x_1 + 8x_2 + 12x_3$ for all $(x_1, x_2, x_3) \in \mathbb{R}^3$. Let $g_2, h_2 : \mathbb{R}^2 \to \mathbb{R}$ be defined by $g_2(y_1, y_2) := y_1^2 + y_2^2$ and $h_2(y_1, y_2) := 28y_1 + 64y_2$ for all $(y_1, y_2) \in \mathbb{R}^2$. Let $A : \mathbb{R}^3 \to \mathbb{R}^2$ be defined by $A(x_1, x_2, x_3) = (x_1 + 2x_2 + 3x_3, 4x_1 + 5x_2 + 6x_3)$ for all $(x_1, x_2, x_3) \in \mathbb{R}^3$. Here, $\|A\| \approx 0.10517$ and $\Omega_{\mathrm{SDCP}} = \{(1, 2, 3)\}$.

From Table 1, we see that Algorithm 4.2 reaches the required errors only need six iterations if $\beta_n = 500$ for all $n \in \mathbb{N}$, but Algorithm 4.2 reaches the required errors need 73 iterations if $\beta_n = 0.1$ for all $n \in \mathbb{N}$.

From Table 2, we see that Algorithm 4.2 reaches the required errors only need seven iterations if $\beta_n = 100$ for all $n \in \mathbb{N}$, but Algorithm 4.2 reaches the required errors need 72 iterations if $\beta_n = 0.1$ for all $n \in \mathbb{N}$.

From Table 3, we see that Algorithm 3.1 reaches the required errors only need seven iterations if $\beta_n = 700$ for all $n \in \mathbb{N}$, but Algorithm 3.1 reaches the required errors need 99 iterations if $\beta_n = 0.1$ for all $n \in \mathbb{N}$.

From Table 3 and Table 4, we see that Algorithm 3.1 reaches the required errors need 283 iterations if $\beta_n = 1$ and $r_n = 0.05$ for all $n \in \mathbb{N}$, but Algorithm 3.1 reaches the required errors need 39 iterations if $\beta_n = 1$ and $r_n = 0.09$ for all $n \in \mathbb{N}$. On the other hand, for other settings of β_n, we know the numerical results in Table 3 and Table 4 show that there are no significant differences in the setting of $\{r_n\}_{n\in\mathbb{N}}$.

Table 1 Numerical results for Example 4.1.

Algorithm 4.2	$x_1 = (88, 2000, 500)$, and $r_n = 0.5$ for all $n \in \mathbb{N}$	
$\varepsilon = 10^{-12}$	Iteration	Approximate solution
$\beta_n = 0.1$	73	(1.00000000000004, 2.00000000000091, 3.00000000000023)
$\beta_n = 1$	18	(1.00000000000002, 2.00000000000044, 3.00000000000011)
$\beta_n = 10$	10	(1.00000000000000, 2.00000000000002, 3.00000000000000)
$\beta_n = 20$	8	(1.00000000000003, 2.00000000000074, 3.00000000000018)
$\beta_n = 30$	8	(1.00000000000000, 2.00000000000004, 3.00000000000001)
$\beta_n = 40$	8	(1.00000000000000, 2.00000000000001, 3.00000000000000)
$\beta_n = 50$	7	(1.00000000000002, 2.00000000000049, 3.00000000000012)
$\beta_n = 100$	7	(1.00000000000000, 2.00000000000001, 3.00000000000030)
$\beta_n = 500$	6	(1, 2, 3)

Table 2 Numerical results for Example 4.1.

Algorithm 4.2	$x_1 = (123, 456, 789)$, and $r_n = 0.5$ for all $n \in \mathbb{N}$	
$\varepsilon = 10^{-12}$	Iteration	Approximate solution
$\beta_n = 0.1$	72	(1.00000000000009, 2.00000000000034, 3.00000000000059)
$\beta_n = 1$	18	(1.00000000000003, 2.00000000000010, 3.00000000000017)
$\beta_n = 10$	9	(1.00000000000007, 2.00000000000027, 3.00000000000047)
$\beta_n = 20$	8	(1.00000000000005, 2.00000000000017, 3.00000000000029)
$\beta_n = 30$	8	(1.00000000000000, 2.00000000000001, 3.00000000000002)
$\beta_n = 40$	7	(1.00000000000011, 2.00000000000042, 3.00000000000073)
$\beta_n = 50$	7	(1.00000000000003, 2.00000000000011, 3.00000000000019)
$\beta_n = 100$	7	(1, 2, 3)

Table 3 Numerical results for Example 4.2.

Algorithm 3.1	$x_1 = (123, 456, 789)$, and $r_n = 0.05$ for all $n \in \mathbb{N}$	
$\varepsilon = 10^{-12}$	Iteration	Approximate solution
$\beta_n = 0.1$	99	(0.99999999999927, 1.99999999999990, 3.00000000000052)
$\beta_n = 1$	39	(1.00000000000036, 2.00000000000048, 3.00000000000059)
$\beta_n = 10$	15	(1.00000000000018, 2.00000000000024, 3.00000000000030)
$\beta_n = 20$	12	(0.99999999999973, 1.99999999999964, 2.99999999999955)
$\beta_n = 30$	11	(1.00000000000013, 2.00000000000017, 3.00000000000021)
$\beta_n = 40$	10	(0.99999999999963, 1.99999999999952, 2.99999999999940)
$\beta_n = 50$	10	(0.99999999999995, 1.99999999999993, 2.99999999999992)
$\beta_n = 100$	9	(1.00000000000001, 2.00000000000002, 3.00000000000002)
$\beta_n = 700$	7	(1, 2, 3)

Table 4 Numerical results for Example 4.2.

Algorithm 3.1	$x_1 = (123, 456, 789)$, and $r_n = 0.09$ for all $n \in \mathbb{N}$	
$\varepsilon = 10^{-12}$	Iteration	Approximate solution
$\beta_n = 0.1$	98	(0.99999999999931, 1.99999999999990, 3.00000000000050)
$\beta_n = 1$	283	(1.00000000000038, 2.00000000000051, 3.00000000000063)
$\beta_n = 10$	21	(1.00000000000008, 2.00000000000010, 3.00000000000013)
$\beta_n = 20$	15	(1.00000000000042, 2.00000000000056, 3.00000000000069)
$\beta_n = 30$	14	(0.99999999999997, 1.99999999999996, 2.99999999999995)
$\beta_n = 40$	12	(0.99999999999959, 1.99999999999946, 2.99999999999933)
$\beta_n = 50$	12	(0.99999999999996, 1.99999999999995, 2.99999999999994)
$\beta_n = 100$	10	(0.99999999999994, 1.99999999999992, 2.99999999999990)
$\beta_n = 1300$	7	(1, 2, 3)

Table 5 Numerical results for Example 4.2.

Algorithm 3.1	$x_1 = (123, 456, 789)$, and $\beta_n = 1$ for all $n \in \mathbb{N}$	
$\varepsilon = 10^{-12}$	Iteration	Approximate solution
$r_n = 0.05$	39	(1.00000000000036, 2.00000000000048, 3.00000000000059)
$r_n = 0.09$	283	(1.00000000000038, 2.00000000000051, 3.00000000000063)
$r_n = 0.095$	611	(1.00000000000041, 2.00000000000054, 3.00000000000067)
$r_n = 0.099$	5129	(1.00000000000043, 2.00000000000056, 3.00000000000070)
$r_n = 0.0994$	18,434	(0.99999999999957, 1.99999999999943, 2.99999999999930)

However, in Table 5, if $\beta_n = 1$ for all $n \in \mathbb{N}$, then we know the numerical results have big differences in the setting of $\{r_n\}_{n\in\mathbb{N}}$.

Acknowledgements
The authors would like to thank referee(s) for many useful comments and suggestions for the improvement of the article. Prof. Chi-Ming Chen was supported by Grant No. MOST 107-2115-M-007-008 of the Ministry of Science and Technology of the Republic of China.

Funding
Not applicable.

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

Authors' contributions
The authors contributed equally and significantly in writing this paper. The authors read and approved the final manuscript.

Author details
[1]Department of Applied Mathematics, National Chiayi University, Chiayi, Taiwan. [2]Institute of Computational and Modeling Science, National Tsing Hua University, Hsinchu, Taiwan.

Publisher's Note

Received: 4 October 2018 Accepted: 29 April 2019 Published online: 22 May 2019

References

1. Bauschke, H.H., Combettes, P.L.: Convex Analysis and Monotone Operator Theory in Hilbert Spaces. Springer, Berlin (2011)
2. Butnariu, D., Iusem, A.N.: Totally Convex Functions for Fixed Points Computation and Infinite Dimensional Optimization. Kluwer Academic, Norwell (2000)
3. Byrne, C., Censor, Y., Gibali, A., Reich, S.: Weak and strong convergence of algorithms for the split common null point problem. J. Nonlinear Convex Anal. **13**, 759–775 (2011)
4. Chuang, C.S.: Strong convergence theorems for the split variational inclusion problem in Hilbert spaces. Fixed Point Theory Appl. **2013**, 350 (2013)
5. Combettes, P.L., Wajs, V.R.: Signal recovery by proximal forward-backward splitting. Multiscale Model. Simul. **4**, 1168–1200 (2005)
6. Fujikara, Y., Kuroiwa, D.: Lagrange duality in canonical DC programming. J. Math. Anal. Appl. **408**, 476–483 (2013)
7. Harada, R., Kuroiwa, D.: Lagrange-type in DC programming. J. Math. Anal. Appl. **418**, 415–424 (2014)
8. Hiriart-Urruty, J.B., Tuy, H.: Essays on Nonconvex Optimization. Mathematical Programming, vol. 41. North-Holland, Amsterdam (1988)
9. Levy, A.J.: A fast quadratic programming algorithm for positive signal restoration. IEEE Trans. Acoust. Speech Signal Process. **31**, 1337–1341 (1983)
10. Marino, G., Xu, H.K.: Convergence of generalized proximal point algorithm. Commun. Pure Appl. Anal. **3**, 791–808 (2004)
11. Martinet, B.: Régularisation d'inéquations variationnelles par approximations successives. Rev. Fr. Inform. Rech. Opér. **4**(Ser R–3), 154–158 (1970)
12. Moudafi, A.: Split monotone variational inclusions. J. Optim. Theory Appl. **150**, 275–283 (2011)
13. Pham, D.T., An, L.T.H., Akoa, F.: The DC (difference of convex functions) programming and DCA revisited with DC models of real world nonconvex optimization problems. Ann. Oper. Res. **133**, 23–46 (2005)
14. Saeki, Y., Kuroiwa, D.: Optimality conditions for DC programming problems with reverse convex constraints. Nonlinear Anal. **80**, 18–27 (2013)
15. Souza, J.C.O., Oliveira, R.P., Soubeyran, A.: Global convergence of a proximal linearized algorithm for difference of convex functions. Optim. Lett. **10**, 1529–1539 (2016)
16. Sun, W., Sampaio, R.J.B., Candido, M.A.B.: Proximal point algorithm for minimization of DC functions. J. Comput. Math. **21**, 451–462 (2003)
17. Takahashi, W.: Introduction to Nonlinear and Convex Analysis. Yokohoma Publishers, Yokohoma (2009)