Mathematics and Computers in Simulation 223 (2024) 86–107

Contents lists available at ScienceDirect

Mathematics and Computers in Simulation

journal homepage: www.elsevier.com/locate/matcom

Original articles

Extended splitting methods for systems of three-operator monotone inclusions with continuous operators

Yunda Dong

School of Mathematics and Statistics, Zhengzhou University, Zhengzhou, 450001, PR China

ARTICLE INFO

Keywords:
Monotone inclusion
Characteristic operator
Uniform continuity
Splitting method
Weak convergence

ABSTRACT

In this article, we propose two new splitting methods for solving systems of three-operator monotone inclusions in real Hilbert spaces, where the first operator is continuous monotone, the second is maximal monotone and the third is maximal monotone and is linearly composed. These methods primarily involve evaluating the first operator and computing resolvents with respect to the other two operators. For one method corresponding to Lipschitz continuous operator, we give back-tracking techniques to determine step lengths. Moreover, we propose a dual-first version of this method. For the other method, which corresponds to a uniformly continuous operator, we develop innovative back-tracking techniques, incorporating additional conditions to determine step lengths. The weak convergence of either method is proven using characteristic operator techniques. Notably, either method fully decouples the third operator from its linear composition operator. Numerical results demonstrate the effectiveness of our proposed splitting methods, together with their special cases and variants, in solving test problems.

1. Introduction

Let $\mathcal{H}_i$, $i = 1, \dots, n$, and $\mathcal{G}$ be real infinite-dimensional Hilbert spaces, with usual inner product $\langle \cdot, \cdot \rangle$ and induced norm $\|x\| = \sqrt{\langle x, x \rangle}$ for each vector x. In this article, we focus on the following system of three-operator monotone inclusions

$$0 \in F_i(x_i) + A_i(x_i) + Q_i^* B \left(\sum_{i=1}^n Q_i x_i - q \right), \quad i = 1, \dots, n, \tag{1}$$

where F_i, $A_i : \mathcal{H}_i \rightrightarrows \mathcal{H}_i$ are maximal monotone operators, $B : \mathcal{G} \rightrightarrows \mathcal{G}$ is a maximal monotone operator, and each $Q_i : \mathcal{H}_i \to \mathcal{G}$ is nonzero bounded linear operator with its adjoint operator Q_i^*, and $q \in \mathcal{G}$ is a vector. This problem models a wide range of problems arising from definite linear systems, linear/quadratic programming, semi-definite programming, complementarity problems, variational inequality problems, optimal control, traffic equilibrium, image reconstruction, portfolio selection. For further details, please refer to [3,6,14,20,22,29,32].

Recently, an extended Douglas-Rachford splitting method for solving (1) was proposed by the author [8], building upon prior work by [9,13,25]. At each iteration, the method requires computations of the resolvent for each individual operator in the monotone inclusion. Under the weakest possible assumptions, the author proved its weak convergence using characteristic operator techniques. Interestingly, the proposed method includes as special cases the Douglas-Rachford splitting method for a two-operator monotone inclusion and an equivalent version of a splitting method for convex minimization [18].

Now consider an important special case of the problem above. If $n = 1$, F is continuous, and B vanishes, then it reduces to

$$0 \in F(x) + A(x). \tag{2}$$

E-mail address: ydong@zzu.edu.cn.

https://doi.org/10.1016/j.matcom.2024.03.024
Received 31 December 2022; Received in revised form 25 December 2023; Accepted 18 March 2024
Available online 4 April 2024

In this case, Tseng's splitting method is particularly well-suited. Choose $x^0 \in \mathrm{dom}F \cap \mathrm{dom}A$. Choose $\sigma > 0$, $\beta \in (0,1)$, $\alpha_{-1} > 0$. At kth iteration, find the smallest i_k in $\alpha = \alpha_{k-1}\beta^i$, $i = 0, 1, \ldots$, such that $x^k(\alpha) = (I + \alpha A)^{-1}(x^k - \alpha F(x^k))$ satisfies

$$\alpha \| F(x^k) - F(x^k(\alpha))\| \le (1-\rho)\|x^k - x^k(\alpha)\|, \quad \rho \in (0,1). \tag{3}$$

Then take $\alpha_k = \alpha_{k-1}\beta^{i_k}$, and compute $\bar{x}^k = x^k(\alpha_k)$. Finally, compute

$$x^{k+1} = \bar{x}^k - \alpha_k F(\bar{x}^k) + \alpha_k F(x^k).$$

Tseng [32, Theorem 3.4] proved the method's individual weak convergence when F is monotone and Lipschitz continuous, and F is monotone and uniformly continuous respectively. However, to overcome its small steplength phenomenon, the author proposed an alternative in his Ph.D. dissertation [4]. Choose $x^0 \in \mathrm{dom}F \cap \mathrm{dom}A$. Choose $t \in (0,1)$, $\rho \in (0,1)$ and $\alpha_{-1} > 0$. At kth iteration, find the smallest j_k in $\alpha = \alpha_{k-1}t^j$, $j = 0, 1, \ldots$, such that $x^k(\alpha) = (I + \alpha A)^{-1}(x^k - \alpha F(x^k))$ satisfies

$$\alpha \langle x^k - x^k(\alpha),\ F(x^k) - F(x^k(\alpha))\rangle \le (1-\rho)\|x^k - x^k(\alpha)\|^2. \tag{4}$$

Take $\alpha_k = \alpha_{k-1}t^{j_k}$, and compute $\bar{x}^k = x^k(\alpha_k)$. Compute in order

$$d^k = x^k - \bar{x}^k - \alpha_k(F(x^k) - F(\bar{x}^k)), \quad \gamma_k = \langle x^k - \bar{x}^k,\ d^k\rangle / \|d^k\|^2.$$

Then choose $\theta_k \in (0,2]$ and compute

$$x^{k+1} = x^k - \theta_k \gamma_k d^k. \tag{5}$$

Notice that the condition (4) itself is originally due to [17,30,31] in the setting of monotone variational inequalities and later extended in [4] to such monotone inclusions. Almost in the same time, Noor [27] independently proposed a conceptual method similar to (5) but without our self-adaptive choice of α_k, and he proved the method's convergence in the finite-dimensional space.

Next, we will consider the problem (1) in the case where $n = 1$, i.e.,

$$0 \in F(x) + A(x) + Q^* B(Qx - q). \tag{6}$$

In this scenario, the primal monotone inclusion can be transformed into the following primal–dual monotone inclusion

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} \in \begin{pmatrix} F & Q^* \\ -Q & 0 \end{pmatrix} \begin{pmatrix} x \\ u \end{pmatrix} + \begin{pmatrix} 0 \\ q \end{pmatrix} + \begin{pmatrix} A & \\ & B^{-1} \end{pmatrix} \begin{pmatrix} x \\ u \end{pmatrix}. \tag{7}$$

As an application of the method (5), a special case can be obtained when $\mathrm{dom}F = \mathrm{dom}A = \mathcal{H}$. Choose $x^0 \in \mathcal{H}$ and $u^0 \in \mathcal{G}$. Choose $t \in (0,1)$ and $\alpha_{-1} > 0$. At kth iteration, find the smallest j_k in $\alpha = \alpha_{k-1}t^j$, $j = 0, 1, \ldots$, such that

$$x^k(\alpha) = (I + \alpha A)^{-1}(x^k - \alpha(F(x^k) + Q^* u^k)) \tag{8}$$

$$u^k(\alpha) = (I + \alpha B^{-1})^{-1}(u^k + \alpha(Qx^k - q)) \tag{9}$$

satisfies in order

$$\alpha \langle x^k - x^k(\alpha), F(x^k) - F(x^k(\alpha))\rangle \le (1-\rho)(\|x^k - x^k(\alpha)\|^2 + \|u^k - u^k(\alpha)\|^2). \tag{10}$$

Take $\alpha_k = \alpha_{k-1}t^{j_k}$, and compute $\bar{x}^k = x^k(\alpha_k)$, $\bar{u}^k = u^k(\alpha_k)$. Take

$$\begin{aligned} d_x^k &= x^k - \bar{x}^k - \alpha_k(F(x^k) - F(\bar{x}^k)) - \alpha_k Q^*(u^k - \bar{u}^k), \\ d_u^k &= u^k - \bar{u}^k + \alpha_k Q(x^k - \bar{x}^k), \\ \gamma_k &= \frac{\langle x^k - \bar{x}^k, d_x^k\rangle + \langle u^k - \bar{u}^k, d_u^k\rangle}{\|d_x^k\|^2 + \|d_u^k\|^2}. \end{aligned}$$

Then choose $\theta_k \in (0,2]$ and compute

$$x^{k+1} = x^k - \theta_k \gamma_k d_x^k, \quad u^{k+1} = u^k - \theta_k \gamma_k d_u^k. \tag{11}$$

The problem (6) above reminds us of a splitting method recently proposed by Johnstone and Eckstein [23] (JE splitting for short), which can be used to solve it. Consider the following monotone inclusion of finding a $z \in \mathcal{H}$ such that

$$0 \in G_1^* T_1 G_1 z + G_2^* T_2 G_2 z + T_3 z,$$

where $T_1 : \mathcal{G}_1 \rightrightarrows \mathcal{G}_1$, $T_2 : \mathcal{G}_2 \rightrightarrows \mathcal{G}_2$ are maximal monotone, and $\mathcal{G}_1$, $\mathcal{G}_2$ are Hilbert spaces, and $T_3 : \mathcal{H} \to \mathcal{H}$ is monotone and Lipschitz continuous, $G_1 : \mathcal{H} \to \mathcal{G}_1$ and $G_2 : \mathcal{H} \to \mathcal{G}_2$ are nonzero bounded linear. Their novel splitting method is given as follows. Choose $z^1 \in \mathcal{H}$, $w_1^1 \in \mathcal{G}_1$, $w_2^1 \in \mathcal{G}_2$, $w_3^1 \in \mathcal{H}$. Choose $\rho_{3,0} > 0$, $\gamma > 0$, $\Delta > 0$. For $i = 1, 2$, choose $\rho_{i,k} > 0$ and compute

$$x_i^k = (I + \rho_{i,k} T_i)^{-1}(G_i z^k + \rho_{i,k} w_i^k), \quad y_i^k = \rho_{i,k}^{-1}(G_i z^k + \rho_{i,k} w_i^k - x_i^k).$$

Choose $\rho_{3,k}$ to be the smallest element of $\rho_{3,k-1}\{1, 0.5, 0.5^2, \ldots\}$ such that

$$\rho_{3,k} \Delta \|T_3 z^k - w_3^k\|^2 \le \left\langle T_3 z^k - w_3^k, T_3(z^k - \rho_{3,k}(T_3 z^k - w_3^k)) - w_3^k \right\rangle. \tag{12}$$

Compute in order

$$x_3^k = z^k - \rho_{3,k}(T_3 z^k - w_3^k), \quad y_3^k = T_3 x_3^k,$$
$$u_i^k = x_i^k - G_i x_3^k, \ i = 1,2, \quad v^k = G_1^* y_1^k + G_2^* y_2^k + y_3^k.$$

If $\pi_k = \|u_1^k\|^2 + \|u_2^k\|^2 + \gamma^{-1}\|v^k\|^2 = 0$, then stop. Otherwise, choose $\beta_k \in [0.01, 1.99]$ and compute

$$\alpha_k = \pi_k^{-1}\beta_k \max\left\{0, \langle z^k, v^k\rangle + \langle w_1^k, u_1^k\rangle + \langle w_2^k, u_2^k\rangle - \sum_{i=1}^3 \langle x_i^k, y_i^k\rangle\right\},$$
$$z^{k+1} = z^k - \gamma^{-1}\alpha_k v^k,$$
$$w_i^{k+1} = w_i^k - \alpha_k u_i^k, \ i = 1,2, \quad w_3^{k+1} = -G_1^* w_1^{k+1} - G_2^* w_2^{k+1}.$$

At first glance, the method (11) has several similarities to JE splitting: At each iteration, the main costs are operator's evaluations of (Lipschitz) continuous operator, computations of each resolvent of other two. However, by comparing their individual ways of choosing steplengths, we find that the corresponding condition (12) in JE splitting has a different property from the one (10). The most striking feature that impresses us deeply is that the former merely requires operator's evaluations of continuous operator, whereas the latter requires operator evaluations of continuous operator and two resolvent computations. Thus, the condition (12) in JE splitting can considerably reduce computations in determining the associated steplengths.

Motivated by these observations, in this article, we aim to develop new splitting methods, where each involved steplength condition is similar to (10), but, as in JE splitting, no longer requires the resolvent of B^{-1}. Assuming that each F_i is further continuous, the key ingredients of new methods can be stated as follows. Choose $x^0 \in \mathcal{H}$ and $u^0 \in \mathcal{G}$. Choose $t \in (0,1)$ and $\alpha_{i,-1} > 0$ $(i = 1, \dots, n)$. At kth iteration, for the current primal–dual iterate (x^k, u^k), find the smallest j_k in $\alpha = \alpha_{k-1}t^j$, $j = 0, 1, \dots$, such that

$$x_i^k(\alpha_i) = (I + \alpha_i A_i)^{-1}(x_i^k - \alpha_i(F_i(x_i^k) + Q_i^* u^k))$$

satisfies

$$\alpha_i\langle x_i^k - x_i^k(\alpha_i),\ F_i(x_i^k) - F_i(x_i^k(\alpha_i))\rangle \le (1-\rho)\|x_i^k - x_i^k(\alpha_i)\|^2 \tag{13}$$

for $i = 1, \dots, n$. Take $\alpha_{i,k} = \alpha_{i,k-1}t^{j_k}$, and compute $\bar{x}_i^k = x_i^k(\alpha_{i,k})$. Choose $\beta_k > 0$ and find $\bar{u}^k$ such that

$$\left(\beta_k I + B^{-1}\right)(\bar{u}^k) \ni \beta_k u^k + \sum_{i=1}^n Q_i \bar{x}_i^k - q. \tag{14}$$

Finally, based on the information above, we obtain the new primal–dual iterate (x^{k+1}, u^{k+1}) in some cheap but different ways; see Algorithms 2.1 and 2.2 below for more details.

Clearly, when we compare the condition (13) (in the $n = 1$ case) with the one (10), we have removed resolvent computations with respect to B^{-1}. This design is particularly useful when such resolvent is difficult to evaluate, as shown in the second test problem for numerical demonstrations. Additionally, our proposed method is ideally suited for the case, where the operator is skew-adjoint (in part). Notice that the intermediate dual point in (14) (in the $n = 1$ case) is obviously different and can take full advantage of the latest information, when compared to the one in (9). In this sense, our proposed splitting methods are desirable and are not direct applications of the existing method (5).

As shown below, if each F_i is Lipschitz continuous, then we are able to use (13) to prove weak convergence of Algorithm 2.1. Impressively, if each F_i is merely assumed to be uniformly continuous but not necessarily Lipschitz continuous, we have to resort to additional conditions

$$\|x_i^k - x_i^k(\alpha) - \alpha(F_i(x_i^k) - F_i(x_i^k(\alpha)))\|^2 \le \tau\langle x_i^k - x_i^k(\alpha),\ x_i^k - x_i^k(\alpha) - \alpha(F_i(x_i^k) - F_i(x_i^k(\alpha)))\rangle$$

for $i = 1, \dots, n$, where $\tau > 1$ is any given positive number, to prove weak convergence of the corresponding Algorithm 2.2. To the best of our knowledge, this is a new idea of combining this group of conditions with (13) (further require $\alpha \leftarrow \alpha_i$) to determine steplengths in some algorithm. In particular, this new type of conditions still allows for larger steplengths than the condition (3); see Lemma 4.1 below for further explanations.

The rest of this article is organized as follows. In Section 2, we fully state our proposed splitting methods in Hilbert spaces in the setting of the monotone inclusions (1) above. In Section 3 and Section 4, under the weakest possible conditions, with the help of characteristic operator techniques [8,11,12,35], by assuming Lipschitz continuity and uniform continuity of the first operator, we prove individual weak convergence of the generated primal sequence of the iterates, respectively. In Section 5, we present the dual-first version of Algorithm 2.1. In Section 6 and Section 7, we discuss some other aspects of the proposed methods. In Section 8, we conducted numerical experiments to confirm the effectiveness of our proposed splitting methods, together with their special cases and variants, in solving our two test problems, when compared with other state-of-the-art algorithms. In Section 9, we conclude this article with some remarks. In Appendix A, we provide some useful concepts and preliminary results.

Throughout this article, we will agree that the notation $(x, a) \in A$ and the notation $x \in \mathrm{dom}A \subseteq \mathcal{H}$, $a \in A(x)$ have the same meaning, where $A : \mathcal{H} \rightrightarrows \mathcal{H}$ is maximal monotone and $\mathrm{dom}A = \{x \in \mathcal{H} : A(x) \neq \emptyset\}$ is its effective domain. We will denote by δ_C the indicator function (cf. [8]) of a nonempty set C.

2. Methods

In this section, we describe our proposed splitting method for systems of monotone inclusions (1) in details.
The method's design is based on the following

Assumption 2.1. For the system of monotone inclusions (1), we assume that: (i) For $i = 1, \dots, n$, each F_i is continuous. (ii) There exists an $l \in \{0, 1, \dots, n-1\}$ such that each of $F_1, \dots, F_l$ is skew-adjoint, linear operator. (iii) There exist $x_1^* \in \mathcal{H}_1$, ..., $x_n^* \in \mathcal{H}_n$, $u^* \in \mathcal{G}$ such that they solve

$$0 \in F_i(x_i) + A_i(x_i) + Q_i^* u, \quad i = 1, \dots, n, \tag{15}$$

$$0 \in -\sum_{i=1}^{n} Q_i x_i + q + B^{-1}(u). \tag{16}$$

Furthermore, $\emptyset \neq \mathrm{dom} F_i$, $\emptyset \neq \mathrm{dom} A_i$ for $i = 1, \dots, n$, $\emptyset \neq \mathrm{dom} B$. (iv) Every $\mathrm{dom} F_i \cap \mathrm{dom} A_i$ is closed for $i = 1, \dots, n$.

First of all, we would like to follow [8, Sect. 3] to explain Assumption 2.1 a bit. For example, we consider

$$\min \ \bar{f}(x) + f(x) + g(Qx - q),$$

where $\bar{f}, f : \mathcal{R}^n \to \mathcal{R}$, $g : \mathcal{R}^m \to \mathcal{R}$ are closed, proper convex functions and $\bar{f}$ is further continuously differentiable, Q is an $m \times n$ matrix, with its transpose Q^T, and $q \in \mathcal{R}^m$. If there exists an x such that

$$x \in \mathrm{ri}\ \mathrm{dom} \bar{f} \cap \mathrm{ri}\ \mathrm{dom} f, \quad Qx - q \in \mathrm{ri}\ \mathrm{dom} g, \tag{17}$$

then its optimality condition is

$$0 \in \nabla \bar{f}(x) + \partial f(x) + Q^T \partial g(Qx - q),$$

where ri stands for the relative interior, $\nabla \bar{f}$ and ∂f stand for the gradient and subdifferential of $\bar{f}$ and f respectively; see [28] for more details. If g is taken to be the indicator function $\delta_{\{0\}}$, then (17) reduces to

$$x \in \mathrm{ri}\ \mathrm{dom} \nabla \bar{f} \cap \mathrm{ri}\ \mathrm{dom} \partial f, \quad Qx - q = 0 \tag{18}$$

because the set ri dom$g$ becomes $\{0\}$, and we further have

$$0 \in \nabla \bar{f}(x) + \partial f(x) + Q^T \partial \delta_{\{0\}}(Qx - q).$$

Of course, we may replace (18) by

$$x \in \mathrm{int}\ \mathrm{dom} \nabla \bar{f} \cap \mathrm{int}\ \mathrm{dom} \partial f, \quad Qx - q = 0,$$

where int stands for the interior. This is stronger but more convenient.

For the associated steplengths of our proposed splitting algorithms below, at kth iteration, we adopt the following Armijo-like way. For $i = l+1, \dots, n$, choose $t \in (0, 1)$. We set

$$\alpha_i = \alpha_{i,k-1} t^j, \begin{cases} j = -1, 0, 1, \dots, & \text{if } F_i \text{ is strongly monotone,} \\ j = 0, 1, \dots, & \text{otherwise.} \end{cases} \tag{19}$$

2.1. F_i is Lipschitz continuous

In the case of F_i being Lipschitz continuous, we give

Algorithm 2.1. Our proposed splitting algorithm in Lipschitz continuity case

Step 0. For $i = 1, \dots, n$, choose $x_i^0 \in \mathcal{H}_i$, $u^0 \in \mathcal{G}$. Choose $\alpha_{i,-1} > 0$, $\rho \in (0, 1)$. Set $k := 0$.
Step 1. For $i = 1, \dots, l$, choose appropriate $\alpha_{i,k} > 0$. For $i = l+1, \dots, n$, find the smallest j_k in (19) such that

$$x_i^k(\alpha_i) = (I + \alpha_i A_i)^{-1}(x_i^k - \alpha_i(F_i(x_i^k) + Q_i^* u^k))$$

satisfies

$$\alpha_i \langle x_i^k - x_i^k(\alpha_i), F_i(x_i^k) - F_i(x_i^k(\alpha_i)) \rangle \leq (1 - \rho) \| x_i^k - x_i^k(\alpha_i) \|^2. \tag{20}$$

Take $\alpha_{i,k} = \alpha_{i,k-1} t^{j_k}$, and compute

$$\bar{x}_i^k = x_i^k(\alpha_{i,k}). \tag{21}$$

Choose $\beta_k > 0$ via (29) below and find $\bar{u}^k$ such that

$$\left(\beta_k I + B^{-1}\right)(\bar{u}^k) \ni \beta_k u^k + \sum_{i=1}^{n} Q_i \bar{x}_i^k - q. \tag{22}$$

If $\bar{x}_i^k = x_i^k$, $i = 1, \dots, n$, and $\bar{u}^k = u^k$, then stop. Otherwise go to Step 2.

Step 2. For $i = 1, \dots, n$, compute

$$d_{x_i}^k = \alpha_{i,k}^{-1}(x_i^k - \bar{x}_i^k) - (F_i(x_i^k) - F_i(\bar{x}_i^k)) - Q_i^*(u^k - \bar{u}^k), \tag{23}$$

$$d_u^k = \beta_k(u^k - \bar{u}^k), \tag{24}$$

$$\gamma_k = \frac{\sum_{i=1}^n \langle x_i^k - \bar{x}_i^k, d_{x_i}^k \rangle + \langle u^k - \bar{u}^k, d_u^k \rangle}{\sum_{i=1}^n \|d_{x_i}^k\|^2 + \|d_u^k\|^2}. \tag{25}$$

Then choose $\theta_k \in (0, 2]$ and the new iterates are given by

$$x_i^{k+1} = x_i^k - \theta_k \gamma_k d_{x_i}^k, \; i = 1, \dots, n, \tag{26}$$

$$u^{k+1} = u^k - \theta_k \gamma_k d_u^k. \tag{27}$$

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

Notice that, in the case of F_i being strongly monotone, it follows from (20) that

$$\alpha_i \mu_{F_i} \|x_i^k - x_i^k(\alpha_i)\|^2 \le (1 - \rho)\|x_i^k - x_i^k(\alpha_i)\|^2 \quad \Rightarrow \quad \alpha_i \le (1 - \rho)/\mu_{F_i}. \tag{28}$$

Thus, the sequence $\{\alpha_{i,k}\}$ must be uniformly bounded above as desired.

As to β_k, we set $\beta_k = \sum_{i=1}^n \beta_{i,k}$ and choose

$$\begin{aligned} \beta_{i,k} &\ge \frac{2\varepsilon - \varepsilon^2 + \alpha_{i,k}^2 \|Q_i\|^2}{2(2 - \varepsilon)\alpha_{i,k}}, \quad 0 < \varepsilon < 2, \quad i = 1, \dots, l, \\ \beta_{i,k} &\ge \frac{2\rho\varepsilon - \varepsilon^2 + \alpha_{i,k}^2 \|Q_i\|^2}{2(2\rho - \varepsilon)\alpha_{i,k}}, \quad 0 < \varepsilon < 2\rho, \quad \text{otherwise}, \end{aligned} \tag{29}$$

for $i = 1, \dots, n$. These two relations are obtained from

$$\begin{aligned} &1 + \alpha_{i,k}\beta_{i,k} - \sqrt{(1 - \alpha_{i,k}\beta_{i,k})^2 + \alpha_{i,k}^2 \|Q_i\|^2} \ge \varepsilon, \\ &\rho + \alpha_{i,k}\beta_{i,k} - \sqrt{(\rho - \alpha_{i,k}\beta_{i,k})^2 + \alpha_{i,k}^2 \|Q_i\|^2} \ge \varepsilon, \end{aligned} \tag{30}$$

respectively.

2.2. F_i is uniformly continuous

In the case of F_i being uniformly continuous, we give

Algorithm 2.2. Our proposed splitting algorithm in uniform continuity case

Step 0. Let $\Omega_i = \mathrm{dom} F_i \cap \mathrm{dom} A_i$ for $i = 1, \dots, n$. Choose $x_i^0 \in \Omega_i$ and $u^0 \in \mathcal{G}$. Choose $t \in (0, 1)$ and $\alpha_{-1} > 0$. Set $k := 0$.

Step 1. Find the smallest j_k in (19) such that

$$x_i^k(\alpha) = (I + \alpha A_i)^{-1}(x_i^k - \alpha(F_i(x_i^k) + Q_i^* u^k))$$

satisfies $x_i^k(\alpha) \in \Omega_i$ and (38)–(39) below. Take $\alpha_k = \alpha_{k-1} t^{j_k}$, and compute

$$\bar{x}_i^k = x_i^k(\alpha_k), \quad i = 1, \dots, n. \tag{31}$$

Choose β_k via (40) below and find $\bar{u}^k$ such that

$$\left(\beta_k I + B^{-1}\right)(\bar{u}^k) \ni \beta_k u^k + \sum_{i=1}^n Q_i \bar{x}_i^k - q. \tag{32}$$

If $\bar{x}_i^k = x_i^k$, $i = 1, \dots, n$, and $\bar{u}^k = u^k$, then stop. Otherwise go to Step 2.

Step 2. For $i = 1, \dots, n$, take

$$d_{x_i}^k = x_i^k - \bar{x}_i^k - \alpha_k(F_i(x_i^k) - F_i(\bar{x}_i^k)) - \alpha_k Q_i^*(u^k - \bar{u}^k), \tag{33}$$

$$d_u^k = \alpha_k \beta_k(u^k - \bar{u}^k), \tag{34}$$

$$\gamma_k = \frac{\sum_{i=1}^n \langle x_i^k - \bar{x}_i^k, d_{x_i}^k \rangle + \langle u^k - \bar{u}^k, d_u^k \rangle}{\sum_{i=1}^n \|d_{x_i}^k\|^2 + \|d_u^k\|^2}. \tag{35}$$

Then choose $\theta_k \in (0, 2]$ and the new iterates are given by

$$x_i^{k+1} = P_{\Omega_i}[x_i^k - \theta_k \gamma_k d_{x_i}^k], \quad i = 1, \dots, n, \tag{36}$$

$$u^{k+1} = u^k - \theta_k \gamma_k d_u^k. \tag{37}$$

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

First of all, we discuss how to choose $\alpha_k > 0$ in Step 1 of Algorithm 2.2. We give the following conditions

$$\alpha\langle x_i^k - x_i^k(\alpha),\, F_i(x_i^k) - F_i(x_i^k(\alpha))\rangle \le (1-\rho)\|x_i^k - x_i^k(\alpha)\|^2, \tag{38}$$

$$\|x_i^k - x_i^k(\alpha) - \alpha(F_i(x_i^k) - F_i(x_i^k(\alpha)))\|^2 \le \tau\langle x_i^k - x_i^k(\alpha),\, x_i^k - x_i^k(\alpha) - \alpha(F_i(x_i^k) - F_i(x_i^k(\alpha)))\rangle \tag{39}$$

for $i = 1, \ldots, n$, where $\tau > 1$ is any given positive number. Importantly, this is still a new idea of making use of this group of conditions to determine steplengths, although its root is in [5]. By the way, the inequality (38) always holds whenever such F_i is skew-adjoint.

As to β_k, we set $\beta_k = \sum_{i=1}^n \beta_{i,k}$ and choose

$$\begin{aligned} \beta_{i,k} &\ge \frac{2\varepsilon - \varepsilon^2 + \alpha_k^2\|Q_i\|^2}{2(2-\varepsilon)\alpha_k}, \quad 0<\varepsilon<2, \quad i=1,\ldots,l, \\ \beta_{i,k} &\ge \frac{2\rho\varepsilon - \varepsilon^2 + \alpha_k^2\|Q_i\|^2}{2(2\rho-\varepsilon)\alpha_k}, \quad 0<\varepsilon<2\rho, \quad \text{otherwise,} \end{aligned} \tag{40}$$

for $i = 1, \ldots, n$. These two relations are obtained from

$$\begin{aligned} 1 + \alpha_k\beta_{i,k} - \sqrt{(1-\alpha_k\beta_{i,k})^2 + \alpha_k^2\|Q_i\|^2} &\ge \varepsilon, \\ \rho + \alpha_k\beta_{i,k} - \sqrt{(\rho-\alpha_k\beta_{i,k})^2 + \alpha_k^2\|Q_i\|^2} &\ge \varepsilon, \end{aligned} \tag{41}$$

respectively. Please note that α_k here is not $\alpha_{i,k}$ in (30) there. We will discuss a practical way of choosing β_k for both Algorithms 2.1 and 2.2 in Section 7 below.

Interestingly, if B vanishes and $n = 1$, then Algorithm 2.2 coincides well with the method of [4, Algorithm 4.2.4].

Obviously, a nice feature of Algorithms 2.1 and 2.2 is that either decouples B from its linear composition operator. This feature is similar to that of [8, Algorithm 3.1], whose iterative formulae can be stated as follows. At kth iteration, for $x_i^k \in \mathcal{H}_i$, $a_i^k \in A_i(x_i^k)$, $i = 1, \ldots, n$, $x_{n+1}^k \in \mathcal{G}$, $u^k \in \mathcal{G}$. Compute

$$\begin{aligned} &\bar{u}^k = u^k - (x_{n+1}^k - \textstyle\sum_{i=1}^n Q_i x_i^k + q)/\beta, \\ &(\alpha_i I + F_i)(\bar{x}_i^k) = \alpha_i x_i^k - a_i^k - Q_i^*\bar{u}^k, \\ &(\alpha_{n+1} I + B)(\bar{x}_{n+1}^k) \ni \alpha_{n+1}x_{n+1}^k + \bar{u}^k. \end{aligned} \tag{42}$$

Then, choose $\theta \in (0, 2)$, calculate

$$\begin{aligned} \phi_k &:= \textstyle\sum_{i=1}^{n+1} \alpha_i \|x_i^k - \bar{x}_i^k\|^2 + \langle \bar{x}_{n+1}^k - \sum_{i=1}^n Q_i\bar{x}_i^k + q,\ u^k - \bar{u}^k\rangle, \\ \varphi_k &:= \textstyle\sum_{i=1}^{n+1} \|x_i^k - \bar{x}_i^k\|^2 + \|\bar{x}_{n+1}^k - \sum_{i=1}^n Q_i\bar{x}_i^k + q\|^2, \\ \gamma_k &:= \theta\phi_k/\varphi_k. \end{aligned}$$

Finally, for $i = 1, \ldots, n$, compute in order

$$\begin{aligned} &(\alpha_i I + A_i)(x_i^{k+1}) \ni \alpha_i x_i^k + a_i^k - \gamma_k(x_i^k - \bar{x}_i^k), \\ &\alpha_{n+1}x_{n+1}^{k+1} = \alpha_{n+1}x_{n+1}^k - \gamma_k(x_{n+1}^k - \bar{x}_{n+1}^k), \\ &u^{k+1} = u^k - \gamma_k\,(\bar{x}_{n+1}^k - \textstyle\sum_{i=1}^n Q_i\bar{x}_i^k + q), \end{aligned}$$

to get the new iterates. As proved in [8], if $\alpha_i > 0$, $i = 1, \ldots, n+1$, and

$$\beta > \textstyle\sum_{i=1}^n \|Q_i\|^2/(4\alpha_i) + 1/(4\alpha_{n+1}),$$

then the sequence $\{(x_1^k, \ldots, x_n^k)\}$ generated by the method of [8, Algorithm 3.1] weakly converges to an element of the solution set of (1).

However, Algorithms 2.1 and 2.2 differ widely from this method above because neither requires computing the resolvent of each F_i as in (42) at each iteration, but instead, they require appropriate steplength selection. This can be beneficial for numerical performance, especially when the resolvent of each F_i is difficult or impossible to compute exactly. Additionally, for the method of [8, Algorithm 3.1], it remains in force even F_i in (42) is replaced by (possibly multi-valued) maximal monotone operator $\bar{A}_i$.

In essence, Algorithms 2.1 and 2.2 are extensions of the method of [4, Algorithm 4.2.4], while the method of [8, Algorithm 3.1] is an extended DR splitting method.

3. Weak convergence in Lipschitz continuity case

In this section, we analyze convergence properties of the primal sequence and the dual sequence generated by Algorithm 2.1. Under the weakest possible assumptions, we prove the former's weak convergence to a solution of the problem (1) above.

Theorem 3.1. *Let $\{x_i^k\}(i = 1, \ldots, n)$, $\{u^k\}$ be the sequences generated by Algorithm 2.1. If Assumption 2.1 holds and F_i is Lipschitz continuous for $i = 1, \ldots, n$, each $\alpha_{i,k}$ is well-defined and can be determined within finite trials, and each sequence $\{\alpha_{i,k}\}$ has positive lower and upper bounds.*

Proof. *The cases of* $i = 1, \ldots, l$. Obvious.

The cases of $i = l+1, \ldots, n$. Since F_i is κ_i-Lipschitz continuous and monotone, we have

$$0 \le \alpha_i \langle x_i^k - x_i^k(\alpha_i),\, F_i(x_i^k) - F_i(x_i^k(\alpha_i))\rangle \le \alpha_i\,\kappa_i\,\|x_i^k - x_i^k(\alpha_i)\|^2.$$

If α_i takes the form of $\alpha_{i,k-1}t^j$ via (20), then the following relation

$$\alpha_i\,\kappa_i = \alpha_{i,k-1}t^j\,\kappa_i \le 1-\rho$$

will hold for some sufficiently large j. This indicates that $\alpha_{i,k}$ is well-defined and can be determined within finite trials.

In view of (19), the chosen step length falls into two cases. One is that $\alpha_{i,k}$ takes $\alpha_{i,k-1}t^{j_k}$ for $j_k = -1$ or $j_k = 0$. In such case, it is not difficult (but somewhat complicated) to get its positive lower bound. The other is that $\alpha_{i,k}$ takes $\alpha_{i,k-1}t^{j_k}$ for some $j_k \ge 1$. In such case, $\alpha_{i,k}t^{-1}$ does not satisfy (20), i.e.,

$$\alpha_{i,k}t^{-1}\langle x_i^k - x_i^k(\alpha_{i,k}t^{-1}),\, F_i(x_i^k) - F_i(x_i^k(\alpha_{i,k}t^{-1}))\rangle > (1-\rho)\|x_i^k - x_i^k(\alpha_{i,k}t^{-1})\|^2.$$

Combining this with the Cauchy–Schwarz inequality and F_i's Lipschitz continuity yields

$$\alpha_{i,k} > \underline{\alpha}_i = (1-\rho)t/\kappa_i,$$

where κ_i is F_i's Lipschitz constant, i.e., $\{\alpha_{i,k}\}$ has a positive lower bound. Of course, such sequence also has an upper bound from (19) and (28). □

Denote by

$$w = \begin{pmatrix} x_1 \\ \vdots \\ x_n \\ u \end{pmatrix}, \qquad d = \begin{pmatrix} d_{x_1} \\ \vdots \\ d_{x_n} \\ d_u \end{pmatrix}. \tag{43}$$

Theorem 3.2. *In the setting of Theorem 3.1, there exists some positive number* $\hat{\gamma}$ *such that*

$$\|w^{k+1} - w^*\|^2 \le \|w^k - w^*\|^2 - \theta_k(2-\theta_k)\hat{\gamma}\,\|w^k - \bar{w}^k\|^2,$$

where w *is defined in (43). Furthermore, if* $\sum_k \theta_k(2-\theta_k) = +\infty$*, the involved primal sequence is weakly convergent.*

Proof. For $i = 1, \ldots, n$, it follows from (21) that

$$A_i(\bar{x}_i^k) \ni \alpha_{i,k}^{-1}(x_i^k - \bar{x}_i^k) - F_i(x_i^k) - Q_i^* u^k, \tag{44}$$

which, together with (15)

$$A_i(x_i^*) \ni -F_i(x_i^*) - Q_i^* u^*,$$

and monotonicity of each A_i , implies

$$0 \le \langle \bar{x}_i^k - x_i^*,\, \alpha_{i,k}^{-1}(x_i^k - \bar{x}_i^k) - (F_i(x_i^k) - F_i(x_i^*)) - Q_i^*(u^k - u^*)\rangle,\ i = 1, \ldots, n.$$

Adding these relations to the following (due to each F_i's monotonicity)

$$0 \le \langle \bar{x}_i^k - x_i^*,\, F_i(\bar{x}_i^k) - F_i(x_i^*)\rangle,\ i = 1, \ldots, n$$

yields

$$0 \le \langle \bar{x}_i^k - x_i^*,\, \alpha_{i,k}^{-1}(x_i^k - \bar{x}_i^k) - (F_i(x_i^k) - F_i(\bar{x}_i^k)) - Q_i^*(u^k - u^*)\rangle,$$

i.e., for $i = 1, \ldots, n$, we have

$$0 \le \langle \bar{x}_i^k - x_i^*,\, \alpha_{i,k}^{-1}(x_i^k - \bar{x}_i^k) - (F_i(x_i^k) - F_i(\bar{x}_i^k))\rangle - \langle \bar{x}_i^k - x_i^*,\, Q_i^*(u^k - u^*)\rangle. \tag{45}$$

On the other hand, it follows from (22) that

$$B^{-1}(\bar{u}^k) \ni \beta_k(u^k - \bar{u}^k) + \sum\nolimits_{i=1}^n Q_i \bar{x}_i^k - q, \tag{46}$$

which, together with

$$B^{-1}(u^*) \ni \sum\nolimits_{i=1}^n Q_i x_i^* - q$$

and monotonicity of B^{-1}, implies

$$\begin{aligned} 0 &\le \langle \bar{u}^k - u^*,\, \beta_k(u^k - \bar{u}^k) + \sum\nolimits_{i=1}^n Q_i(\bar{x}_i^k - x_i^*)\rangle \\ &= \langle \bar{u}^k - u^*,\, \beta_k(u^k - \bar{u}^k)\rangle + \langle \bar{u}^k - u^*,\, \sum\nolimits_{i=1}^n Q_i(\bar{x}_i^k - x_i^*)\rangle \end{aligned}$$

$$= \langle \bar{u}^k - u^*, \beta_k(u^k - \bar{u}^k)\rangle + \sum_{i=1}^n \langle Q_i^*(\bar{u}^k - u^*), \bar{x}_i^k - x_i^*\rangle.$$

Thus, by this relation and (45), we can get

$$\begin{aligned} 0 &\le \sum_{i=1}^n \langle \bar{x}_i^k - x_i^*, \alpha_{i,k}^{-1}(x_i^k - \bar{x}_i^k) - (F_i(x_i^k) - F_i(\bar{x}_i^k))\rangle - \sum_{i=1}^n \langle \bar{x}_i^k - x_i^*, Q_i^*(u^k - \bar{u}^k)\rangle + \langle \bar{u}^k - u^*, \beta_k(u^k - \bar{u}^k)\rangle \\ &= \sum_{i=1}^n \langle \bar{x}_i^k - x_i^*, \alpha_{i,k}^{-1}(x_i^k - \bar{x}_i^k) - (F_i(x_i^k) - F_i(\bar{x}_i^k)) - Q_i^*(u^k - \bar{u}^k)\rangle + \langle \bar{u}^k - u^*, \beta_k(u^k - \bar{u}^k)\rangle. \end{aligned}$$

By using $\bar{x} - x^* = x - x^* - (x - \bar{x})$ and $\beta_k := \sum_{i=1}^n \beta_{i,k}$, we can further get

$$\begin{aligned} &\sum_{i=1}^n \langle x_i^k - x_i^*, \alpha_{i,k}^{-1}(x_i^k - \bar{x}_i^k) - (F_i(x_i^k) - F_i(\bar{x}_i^k)) - Q_i^*(u^k - \bar{u}^k)\rangle + \langle u^k - u^*, \beta_k(u^k - \bar{u}^k)\rangle \\ &\ge \sum_{i=1}^n \langle x_i^k - \bar{x}_i^k, \alpha_{i,k}^{-1}(x_i^k - \bar{x}_i^k) - (F_i(x_i^k) - F_i(\bar{x}_i^k)) - Q_i^*(u^k - \bar{u}^k)\rangle + \langle u^k - \bar{u}^k, \beta_k(u^k - \bar{u}^k)\rangle \\ &= \sum_{i=1}^n \alpha_{i,k}^{-1}\left(\|x_i^k - \bar{x}_i^k\|^2 - \alpha_{i,k}\langle x_i^k - \bar{x}_i^k, F_i(x_i^k) - F_i(\bar{x}_i^k)\rangle\right) \\ &\quad - \sum_{i=1}^n \langle Q_i(x_i^k - \bar{x}_i^k), u^k - \bar{u}^k\rangle + \beta_k\|u^k - \bar{u}^k\|^2. \end{aligned} \tag{47}$$

Since we have assumed that, for $i = 1, \dots, l$, F_i is further skew-adjoint, we have

$$\langle x_i^k - \bar{x}_i^k, F_i(x_i^k) - F_i(\bar{x}_i^k)\rangle = 0.$$

As to $i = l+1, \dots, n$, we adopt (20) to get

$$\|x_i^k - \bar{x}_i^k\|^2 - \alpha_{i,k}\langle x_i^k - \bar{x}_i^k, F_i(x_i^k) - F_i(\bar{x}_i^k)\rangle \ge \rho\|x_i^k - \bar{x}_i^k\|^2.$$

Therefore, (47) can be divided into two parts. One is that

$$\sum_{i=1}^l \left(\alpha_{i,k}^{-1}\|x_i^k - \bar{x}_i^k\|^2 - \langle Q_i(x_i^k - \bar{x}_i^k), u^k - \bar{u}^k\rangle + \beta_{i,k}\|u^k - \bar{u}^k\|^2\right). \tag{48}$$

The other is that

$$\sum_{i=l+1}^n \left(\alpha_{i,k}^{-1}\rho\|x_i^k - \bar{x}_i^k\|^2 - \langle Q_i(x_i^k - \bar{x}_i^k), u^k - \bar{u}^k\rangle + \beta_{i,k}\|u^k - \bar{u}^k\|^2\right). \tag{49}$$

Then, combining the sum of these two parts with (47), Lemma A5 and the conditions (29) and Theorem 3.1 yields

$$\langle w^k - w^*, d^k\rangle \ge \langle w^k - \bar{w}^k, d^k\rangle > 0.$$

So, it follows from this relation, (26) and (27) that

$$\begin{aligned} &\|w^{k+1} - w^*\|^2 \\ &= \|w^k - w^* - \theta_k\gamma_k d^k\|^2 \\ &= \|w^k - w^*\|^2 - 2\theta_k\gamma_k\langle w^k - w^*, d^k\rangle + \theta_k^2\gamma_k^2\|d^k\|^2 \\ &\le \|w^k - w^*\|^2 - 2\theta_k\gamma_k\langle w^k - \bar{w}^k, d^k\rangle + \theta_k^2\gamma_k^2\|d^k\|^2. \end{aligned}$$

Combining this with (25), i.e.,

$$\gamma_k = \langle w^k - \bar{w}^k, d^k\rangle/\|d^k\|^2$$

yields

$$\begin{aligned} &\|w^{k+1} - w^*\|^2 \\ &\le \|w^k - w^*\|^2 - 2\theta_k\gamma_k\langle w^k - \bar{w}^k, d^k\rangle + \theta_k^2\gamma_k^2\|d^k\|^2 \\ &= \|w^k - w^*\|^2 - \theta_k(2 - \theta_k)\left(\langle w^k - \bar{w}^k, d^k\rangle/\|d^k\|^2\right)\langle w^k - \bar{w}^k, d^k\rangle \\ &= \|w^k - w^*\|^2 - \theta_k(2 - \theta_k)\gamma_k\langle w^k - \bar{w}^k, d^k\rangle. \end{aligned} \tag{50}$$

It remains to prove that the sequence $\{\gamma_k\}$ has a positive lower bound. In fact, it follows from (25) that

$$\gamma_k = \frac{\sum_{i=1}^n \langle x_i^k - \bar{x}_i^k, d_{x_i}^k\rangle + \langle u^k - \bar{u}^k, d_u^k\rangle}{\sum_{i=1}^n \|d_{x_i}^k\|^2 + \|d_u^k\|^2}.$$

In view of (23) and (24), the denominator is bounded above by

$$\left(\sum_{i=1}^n (\underline{\alpha}_i^{-2} + \kappa_i^2 + \|Q_i^*\|^2) + \beta_k^2\right)\left(\sum_{i=1}^n \|x_i^k - \bar{x}_i^k\|^2 + \|u^k - \bar{u}^k\|^2\right)$$

provided that F_i is κ_i-Lipschitz continuous. As to the numerator, by Lemma A5, (29) and (30), we know that (48) is bounded below by

$$\frac{1}{2}\sum_{i=1}^l \alpha_{i,k}^{-1}\varepsilon\left(\|x_i^k - \bar{x}_i^k\|^2 + \|u^k - \bar{u}^k\|^2\right) \ge \frac{1}{2}\varepsilon\bar{\alpha}_i^{-1}\left(\sum_{i=1}^l \|x_i^k - \bar{x}_i^k\|^2 + l\|u^k - \bar{u}^k\|^2\right),$$

where $\bar{\alpha}_i$ $(i = 1, \dots, l)$ stands for an upper bound of the sequence $\{\alpha_{i,k}\}$. Similarly, (49) is further bounded below by

$$\frac{1}{2}\sum_{i=l+1}^n \alpha_{i,k}^{-1}\varepsilon\left(\|x_i^k - \bar{x}_i^k\|^2 + \|u^k - \bar{u}^k\|^2\right) \ge \frac{1}{2}\varepsilon\bar{\alpha}_i^{-1}\left(\sum_{i=l+1}^n \|x_i^k - \bar{x}_i^k\|^2 + (n-l)\|u^k - \bar{u}^k\|^2\right),$$

where $\bar{\alpha}_i$ $(i = l+1,\ldots,n)$ stands for an upper bound of the sequence $\{\alpha_{i,k}\}$ and its existence follows from Theorem 3.1. So, the numerator must be bounded below by

$$\frac{\epsilon}{2\max\{\bar{\alpha}_i\}}\left(\sum_{i=1}^n \|x_i^k-\bar{x}_i^k\|^2+\|u^k-\bar{u}^k\|^2\right), \tag{51}$$

where $\max\{\bar{\alpha}_i\} = \max\{\bar{\alpha}_i : i = 1,\ldots,n\}$. Thus, it can be easily seen that the sequence $\{\gamma_k\}$ has a positive lower bound indeed.

Since the numerator in γ_k is equal to $\langle w^k - \bar{w}^k, d^k\rangle$ and is bounded below by (51), we have

$$\langle w^k-\bar{w}^k, d^k\rangle \geq \frac{\epsilon}{2\max\{\bar{\alpha}_i\}}\left(\sum_{i=1}^n \|x_i^k-\bar{x}_i^k\|^2+\|u^k-\bar{u}^k\|^2\right) = \frac{\epsilon}{2\max\{\bar{\alpha}_i\}}\|w^k-\bar{w}^k\|^2.$$

Thus, it follows from (50) that there exists some $\hat{\gamma} > 0$ such that

$$\|w^{k+1}-w^*\|^2 \leq \|w^k-w^*\|^2 - \theta_k(2-\theta_k)\hat{\gamma}\|w^k-\bar{w}^k\|^2.$$

So, we can conclude that: (a) The limit of the sequence $\{\|w^k - w^*\|\}$ exists and $\{w^k\}$ is bounded in norm; (b) The sequence $\{\theta_k(2-\theta_k)\|w^k-\bar{w}^k\|^2\}$ converges to zero. Combining this with $\sum_k \theta_k(2-\theta_k) = +\infty$ yields that there must exist some subsequence of $\{\|w^k-\bar{w}^k\|\}$ converging to zero. For notational simplicity, without loss of generality, we may assume that the sequence $\{\|w^k-\bar{w}^k\|\}$ itself converges to zero. Since each F_i is Lipschitz continuous, it can be easily seen that

$$(i)\ \ x_i^k-\bar{x}_i^k \to 0,\ i=1,\ldots,n,\ \ u^k-\bar{u}^k\to 0; \tag{52}$$

$$(ii)\ \ \{x_i^k\},\ \{F_i(x_i^k)\},\ i=1,\ldots,n,\ \ \{u^k\}\ \text{ are bounded in norm.} \tag{53}$$

Next, we will make use of Lemma A3 to prove the remaining part. To this end, we follow the definition of the set T to get

$$T(\bar{x}^k, F(x^k), \bar{u}^k) = \begin{pmatrix} A_1(\bar{x}_1^k)+F_1(x_1^k)+Q_1^*\bar{u}^k \\ \vdots \\ A_n(\bar{x}_n^k)+F_n(x_n^k)+Q_n^*\bar{u}^k \\ x_1^k-\bar{x}_1^k \\ \vdots \\ x_n^k-\bar{x}_n^k \\ B^{-1}(\bar{u}^k)-\sum_{i=1}^n Q_i\bar{x}_i^k+q \end{pmatrix}.$$

Combining this with (44) and (46) yields

$$T(\bar{x}^k, F(x^k), \bar{u}^k) \ni \begin{pmatrix} \alpha_{1,k}^{-1}(x_1^k-\bar{x}_1^k)-Q_1^*(u^k-\bar{u}^k) \\ \vdots \\ \alpha_{n,k}^{-1}(x_n^k-\bar{x}_n^k)-Q_n^*(u^k-\bar{u}^k) \\ x_1^k-\bar{x}_1^k \\ \vdots \\ x_n^k-\bar{x}_n^k \\ \beta_k(u^k-\bar{u}^k) \end{pmatrix}. \tag{54}$$

To invoke Lemma A3, we first consider all the terms on the right-hand side of (54). In fact, it follows from boundedness of $\alpha_{i,k}^{-1}$ $(i = 1,\ldots,n)$, β_k, Q_i^* $(i = 1,\ldots,n)$ and (52) that each term strongly converges to zero. On the other hand, according to (53), there exists one weak cluster point such that

$$x^{k_j} \rightharpoonup x^\infty,\quad F(x^{k_j}) \rightharpoonup F(x^\infty),\quad u^{k_j} \rightharpoonup u^\infty,$$

where the notation $\rightharpoonup$ stands for weak convergence. Combining this with (52) yields

$$\bar{x}^{k_j} \rightharpoonup x^\infty,\ \ F(x^{k_j}) \rightharpoonup F(x^\infty),\ \ \bar{u}^{k_j} \rightharpoonup u^\infty.$$

So, if we denote $F(x)$ by z, by Lemma A3, we can conclude that the cluster point $(x^\infty, z^\infty, u^\infty)$ satisfies $0 \in T(x, z, u)$, establishing the desired result. Furthermore, the primal cluster point x^∞ solves problem (1). To prove the uniqueness of the weak cluster point, we adopt a standard approach detailed in [9,29]. □

Remark 3.1. In analyzing the weak convergence of the primal sequence generated by Algorithm 2.1 in real Hilbert spaces, we leverage Lemmas A1 and A3. Our approach is more self-contained and less convoluted, and can be considered an enhancement of those presented in [7,12]. This is due to the modular proof that results from the introduction of the characteristic operator in Lemma A1, as discussed in subsequent works such as [11,35]. The basic idea originates from [12] and the 2017 draft of [8], independently.

Finally, we note that Algorithm 2.1 is related to the proximal point algorithm [26,29]. Specifically, if we set $n = 1$, $\theta_k \equiv 1$, and F and B to be zero, the main iterative formula in Algorithm 2.1 reduces to $(I + \alpha_k A)(x^{k+1}) \ni x^k$, $k = 0, 1, \ldots$, which happens to be the proximal point algorithm. Thus, as shown in [16], the sequence generated by Algorithm 2.1 may not converge strongly in general.

4. Weak convergence in uniform continuity case

In this section, we analyze convergence properties of the primal sequence and the dual sequence generated by Algorithm 2.2. Under the weakest possible assumptions, we prove the former's weak convergence to a solution of the problem (1) above.

First of all, we make the following three assumptions, which are standard in the literature [19,32].

Assumption 4.1. (i) For any $x_i^k \in \mathrm{dom} A_i$ and $y_i^k \in \mathrm{dom}\, A_i$, $i = 1, \dots, n$, if the sequences $\{x_i^k\}$ and $\{y_i^k\}$ converge weakly, respectively, and $\|x_i^k - y_i^k\| \to 0$, then $\|F_i(x_i^k) - F_i(y_i^k)\| \to 0$. (ii) x_i^k is always in $\mathrm{dom} A_i$ for $i = 1, \dots, n$.

Be aware that the first item is slightly different from the one in [32] because we add the assumption on weak convergence of the sequence $\{x_i^k\}$ and is the same as the one in [19]. It is certainly implied by Lipschitz continuity assumption. As to the second item, it holds provided that $\mathrm{dom} A_i = \mathcal{H}_i$.

Lemma 4.1. *Assume that $A : \mathcal{H} \rightrightarrows \mathcal{H}$ is maximal monotone, and $F : \mathcal{H} \to \mathcal{H}$ is uniformly continuous and monotone. Assume that $x \in \mathrm{dom} A$ and (x, u) is not a solution of $0 \in F(x) + A(x) + Q^* u$. Denote by*

$$x(\alpha) = (I + \alpha A)^{-1}(x - \alpha(F(x) + Q^* u)).$$

For any given $\hat{\alpha} > 0$ and $t \in (0,1)$ and $\tau > 1$, it needs finite trials to find out the smallest j such that the corresponding $\alpha = \hat{\alpha} t^j$, $j = -1, 0, 1, \dots$, satisfies

$$\|x - x(\alpha) - \alpha(F(x) - F(x(\alpha)))\|^2 \le \tau \langle x - x(\alpha),\ x - x(\alpha) - \alpha(F(x) - F(x(\alpha)))\rangle. \tag{55}$$

Proof. Assume that the assertion is not valid. Then for every $\alpha = \hat{\alpha} t^j$, $j = -1, 0, 1, \dots$, we always have

$$\|x - x(\alpha) - \alpha(F(x) - F(x(\alpha)))\|^2 > \tau \langle x - x(\alpha),\ x - x(\alpha) - \alpha(F(x) - F(x(\alpha)))\rangle.$$

So, we further have

$$\begin{aligned} &(\tau - 1)\frac{\|x - x(\alpha)\|^2}{\alpha^2} \\ <& (\tau - 2)\langle \frac{x - x(\alpha)}{\alpha}, F(x) - F(x(\alpha))\rangle + \|F(x) - F(x(\alpha))\|^2 \\ \le& |\tau - 2| \frac{\|x - x(\alpha)\|}{\alpha} \|F(x) - F(x(\alpha))\| + \|F(x) - F(x(\alpha))\|^2. \end{aligned} \tag{56}$$

Since $\alpha \to 0$, it follows from (A.2) and uniform continuity of F that $F(x) - F(x(\alpha)) \to 0$. Meanwhile, (A.3) tells us that $\alpha^{-1}\|x - x(\alpha)\|$ must be bounded. Consequently, either term on the right-hand side of (56) tends to zero, so does the term on the left-hand side. This is to say that $\alpha^{-1}\|x - x(\alpha)\| \to 0$ as $\alpha \to 0$. On the other hand, we have assumed that x is not a solution of $0 \in F(x) + A(x) + Q^* u$. Then, in view of (A.5), $\alpha^{-1}\|x - x(\alpha)\|$ has a positive lower bound. So, this is a contradiction. □

Lemma 4.2. *Let $\rho \in (0,1)$ is a prescribed real number. In the setting of Lemma 4.1, any positive number satisfying*

$$\alpha\|F(x) - F(x(\alpha))\| \le (1 - \rho)\|x - x(\alpha)\|$$

must satisfy (55) and

$$\alpha\langle x - x(\alpha),\ F(x) - F(x(\alpha))\rangle \le (1 - \rho)\|x - x(\alpha)\|^2.$$

Proof. Elementary. □

Theorem 4.1. *If Assumption 4.1 holds, then α_k in Algorithm 2.2 is well-defined and can be determined within finite trials, and the resulting sequence $\{\alpha_k\}$ has an upper bound $\bar{\alpha}$.*

Proof. In view of (19), we have $\alpha = \alpha_{k-1} t^j$. Assume that for all j the conditions (38) always fail to hold

$$\alpha\langle x_i^k - x_i^k(\alpha),\ F_i(x_i^k) - F_i(x_i^k(\alpha))\rangle > (1 - \rho)\|x_i^k - x_i^k(\alpha)\|^2$$

for $i = l + 1, \dots, n$. Thus, we can get

$$\|F_i(x_i^k) - F_i(x_i^k(\alpha))\| \ge \langle \frac{x_i^k - x_i^k(\alpha)}{\|x_i^k - x_i^k(\alpha)\|},\ F_i(x_i^k) - F_i(x_i^k(\alpha))\rangle > (1 - \rho)\,\alpha^{-1}\|x_i^k - x_i^k(\alpha)\|.$$

It follows from Lemma A2 that

$$\liminf_{\alpha \to 0}\ \alpha^{-1}\|x_i^k - x_i^k(\alpha)\| = \min\{\|w\|\ : w \in F_i(x_i^k) + A_i(x_i^k) + Q_i^* u^k\}.$$

Since x_i^k is not a solution to the problem, the closed convex set $F_i(x_i^k) + A_i(x_i^k) + Q_i^* u^k$ will not include the origin. So, we can get

$$\liminf_{\alpha \to 0}\ \alpha^{-1}\|x_i^k - x_i^k(\alpha)\| > 0.$$

Combining this with uniform continuity of F_i yields a contraction. This implies that, after finite trials, there exists the smallest $\tilde{j}_k$ such that $\alpha_{k-1}t^{\tilde{j}_k}$ satisfies (38).

Next, we consider

$$\alpha = \alpha_{k-1}t^j, \quad j \in \tilde{j}_k + \{0, 1, \dots\}.$$

Lemma 4.1 tells us that, after finite trials, there exists the smallest j_k such that $j_k \in \tilde{j}_k + \{0, 1, \dots\}$ and $\alpha_{k-1}t^{j_k}$ satisfies (39).

In conclusion, such steplength α_k is well-defined and can be determined with finite trials.

Certainly, in the cases of each F_i $(i = l+1, \dots, n)$ being strongly monotone, it follows from (38) that

$$\alpha\mu_{F_i}\|x_i^k - x_i^k(\alpha)\|^2 \le (1-\rho)\|x_i^k - x_i^k(\alpha)\|^2 \quad\Rightarrow\quad \alpha \le (1-\rho)/\mu_{F_i}.$$

Thus, the sequence $\{\alpha_k\}$ must be uniformly bounded above as desired. □

Theorem 4.2. *Let $\{\gamma_k\}$ be the sequence generated by Algorithm 2.2. If the sequence $\{\beta_k\}$ is bounded above, then there exists some $\hat{\gamma}$ such that*

$$\gamma_k \ge \hat{\gamma} > 0, \quad k = 0, 1, 2, \dots.$$

Proof. In this proof, we make use of $\sum$ to stand for $\sum_{i=1}^n$. Consider

$$\gamma_k = \frac{\sum\langle x_i^k - \bar{x}_i^k, d_{x_i}^k\rangle + \langle u^k - \bar{u}^k, d_u^k\rangle}{\sum\|d_{x_i}^k\|^2 + \|d_u^k\|^2}.$$

It follows from (39) and monotonicity of F_i that

$$\|x_i^k - \bar{x}_i^k - \alpha_k(F_i(x_i^k) - F_i(\bar{x}_i^k))\|^2 \le \tau\langle x_i^k - \bar{x}_i^k, x_i^k - \bar{x}_i^k - \alpha_k(F_i(x_i^k) - F_i(\bar{x}_i^k))\rangle \le \tau\|x_i^k - \bar{x}_i^k\|^2.$$

Thus, we get

$$\begin{aligned}
\|d_{x_i}^k\|^2 &= \|x_i^k - \bar{x}_i^k - \alpha_k(F_i(x_i^k) - F_i(\bar{x}_i^k)) - \alpha_k Q_i^*(u^k - \bar{u}^k)\|^2 \\
&\le (1 + \alpha_k^2\|Q_i\|^2)(\|x_i^k - \bar{x}_i^k - \alpha_k(F_i(x_i^k) - F_i(\bar{x}_i^k))\|^2 + \|u^k - \bar{u}^k\|^2) \\
&\le (1 + \alpha_k^2\|Q_i\|^2)(\tau\|x_i^k - \bar{x}_i^k\|^2 + \|u^k - \bar{u}^k\|^2), \\
\|d_u^k\|^2 &= \alpha_k^2\beta_k^2\|u^k - \bar{u}^k\|^2.
\end{aligned}$$

So, the denominator $\sum\|d_{x_i}^k\|^2 + \|d_u^k\|^2$ can be bounded above by

$$\begin{aligned}
&\max_{1\le i\le n}\{1 + \alpha_k^2\|Q_i\|^2\}(\tau\sum\|x_i^k - \bar{x}_i^k\|^2 + n\|u^k - \bar{u}^k\|^2) \\
&\quad + \alpha_k^2\beta_k^2\|u^k - \bar{u}^k\|^2 \\
\le&(\max_{1\le i\le n}\{1 + \alpha_k^2\|Q_i\|^2\} + \alpha_k^2\beta_k^2)(\tau\sum\|x_i^k - \bar{x}_i^k\|^2 + (n+1)\|u^k - \bar{u}^k\|^2) \\
\le&(\max_{1\le i\le n}\{1 + \alpha_k^2\|Q_i\|^2\} + \alpha_k^2\beta_k^2)(\tau + n + 1)(\sum\|x_i^k - \bar{x}_i^k\|^2 + \|u^k - \bar{u}^k\|^2).
\end{aligned}$$

As to the numerator,

$$\begin{aligned}
&\sum\left(\|x_i^k - \bar{x}_i^k\|^2 - \alpha_k\langle x_i^k - \bar{x}_i^k,\, F_i(x_i^k) - F_i(\bar{x}_i^k)\rangle\right) \\
&- \sum\alpha_k\langle Q_i(x_i^k - \bar{x}_i^k),\, u^k - \bar{u}^k\rangle + \sum\alpha_k\beta_{i,k}\|u^k - \bar{u}^k\|^2.
\end{aligned} \tag{57}$$

Since we have assumed that, for $i = 1, \dots, l$, F_i is further skew-adjoint, we have

$$\langle x_i^k - \bar{x}_i^k,\, F_i(x_i^k) - F_i(\bar{x}_i^k)\rangle = 0.$$

As to $i = l+1, \dots, n$, we adopted (38) to get

$$\|x_i^k - \bar{x}_i^k\|^2 - \alpha_k\langle x_i^k - \bar{x}_i^k,\, F_i(x_i^k) - F_i(\bar{x}_i^k)\rangle \ge \rho\|x_i^k - \bar{x}_i^k\|^2.$$

Therefore, (57) can be bounded below by the sum of two parts. One is that

$$\sum_{i=1}^l\left(\|x_i^k - \bar{x}_i^k\|^2 - \alpha_k\langle Q_i(x_i^k - \bar{x}_i^k),\, u^k - \bar{u}^k\rangle + \alpha_k\beta_{i,k}\|u^k - \bar{u}^k\|^2\right). \tag{58}$$

The other is that

$$\sum_{i=l+1}^n\left(\rho\|x_i^k - \bar{x}_i^k\|^2 - \alpha_k\langle Q_i(x_i^k - \bar{x}_i^k),\, u^k - \bar{u}^k\rangle + \alpha_k\beta_{i,k}\|u^k - \bar{u}^k\|^2\right).$$

Meanwhile, in view of Lemma A5, (40) and (41), they are bounded below by

$$\frac{1}{2}\varepsilon\left(\sum_{i=1}^l\|x_i^k - \bar{x}_i^k\|^2 + \|u^k - \bar{u}^k\|^2\right)$$

and

$$\frac{1}{2}\varepsilon\left(\sum_{i=l+1}^{n}\|x_i^k-\bar{x}_i^k\|^2+\|u^k-\bar{u}^k\|^2\right),$$

respectively. So, the numerator must be further bounded below by

$$\frac{1}{2}\varepsilon\left(\sum\|x_i^k-\bar{x}_i^k\|^2+\|u^k-\bar{u}^k\|^2\right). \tag{59}$$

Thus, it can be easily seen that the sequence $\{\gamma_k\}$ has a positive lower bound indeed, i.e.,

$$\gamma_k \geq \frac{\varepsilon}{2}\frac{1}{\max_{1\leq i\leq n}\{1+\alpha_k^2\|Q_i\|^2\}+\alpha_k^2\beta_k^2}\frac{1}{\tau+n+1}.$$

Denote by

$$\hat{\gamma}=\frac{\varepsilon}{2}\frac{1}{\max_{1\leq i\leq n}\{1+\bar{\alpha}^2\|Q_i\|^2\}+\bar{\alpha}^2\bar{\beta}^2}\frac{1}{\tau+n+1},$$

where $\bar{\beta}$ is an upper bound of the sequence $\{\beta_k\}$, which is well-defined due to boundedness of $\{\alpha_k\}$. Then the desired result follows. □

Denote by

$$w=\begin{pmatrix}x_1\\ \vdots\\ x_n\\ u\end{pmatrix},\qquad d=\begin{pmatrix}d_{x_1}\\ \vdots\\ d_{x_n}\\ d_u\end{pmatrix}. \tag{60}$$

Theorem 4.3. *If Assumption 4.1 holds and*

$$\sum_{k=0}^{+\infty}\theta_k(2-\theta_k)=+\infty, \tag{61}$$

then the involved primal sequence generated by Algorithm 2.2 is weakly convergent.

Proof. For simplicity, we will use $\sum$ to stand for $\sum_{i=1}^{n}$. For $i=1,\dots,n$, it follows from (31) that

$$A_i(\bar{x}_i^k)\ni\alpha_k^{-1}(x_i^k-\bar{x}_i^k)-F_i(x_i^k)-Q_i^*u^k, \tag{62}$$

which, together with (15), i.e., $A_i(x_i^*)\ni -F_i(x_i^*)-Q_i^*u^*$ and μ_{A_i}-monotonicity of each A_i and Lemma A4, implies

$$\begin{aligned}&\langle x_i^k-x_i^*,\,\alpha_k^{-1}(x_i^k-\bar{x}_i^k)-(F_i(x_i^k)-F_i(\bar{x}_i^k))\rangle-\langle Q_i(\bar{x}_i^k-x_i^*),u^k-u^*\rangle\\ \geq&\langle x_i^k-\bar{x}_i^k,\,\alpha_k^{-1}(x_i^k-\bar{x}_i^k)-(F_i(x_i^k)-F_i(\bar{x}_i^k))\rangle+(\mu_{F_i}+\mu_{A_i})\|\bar{x}_i^k-x_i^*\|^2.\end{aligned} \tag{63}$$

On the other hand, it follows from (32) that

$$B^{-1}(\bar{u}^k)\ni\beta_k(u^k-\bar{u}^k)+\sum Q_i\bar{x}_i^k-q, \tag{64}$$

which, together with $B^{-1}(u^*)\ni\sum Q_ix_i^*-q$ and monotonicity of B^{-1}, implies

$$0\leq\langle\bar{u}^k-u^*,\,\beta_k(u^k-\bar{u}^k)+\sum Q_i(\bar{x}_i^k-x_i^*)\rangle=\langle\bar{u}^k-u^*,\,\beta_k(u^k-\bar{u}^k)\rangle+\langle\sum Q_i(\bar{x}_i^k-x_i^*),\bar{u}^k-u^*\rangle.$$

Therefore

$$\langle u^k-u^*,\,\beta_k(u^k-\bar{u}^k)\rangle+\langle\sum Q_i(\bar{x}_i^k-x_i^*),\,\bar{u}^k-u^*\rangle\geq\beta_k\|u^k-\bar{u}^k\|^2. \tag{65}$$

Consider

$$\begin{aligned}&-\langle\sum Q_i(\bar{x}_i^k-x_i^*),\,u^k-u^*\rangle+\langle\sum Q_i(\bar{x}_i^k-x_i^*),\,\bar{u}^k-u^*\rangle\\ =&-\sum\langle\bar{x}_i^k-x_i^*,\,Q_i^*(u^k-\bar{u}^k)\rangle\\ =&-\sum\langle x_i^k-x_i^*,\,Q_i^*(u^k-\bar{u}^k)\rangle+\sum\langle x_i^k-\bar{x}_i^k,\,Q_i^*(u^k-\bar{u}^k)\rangle.\end{aligned}$$

Thus, by adding (65) to (63), we can get

$$\begin{aligned}&\sum\langle x_i^k-x_i^*,\,\alpha_k^{-1}(x_i^k-\bar{x}_i^k)-(F_i(x_i^k)-F_i(\bar{x}_i^k))-Q_i^*(u^k-\bar{u}^k)\rangle\\ &+\langle u^k-u^*,\,\beta_k(u^k-\bar{u}^k)\rangle\\ \geq&\sum\langle x_i^k-\bar{x}_i^k,\,\alpha_k^{-1}(x_i^k-\bar{x}_i^k)-(F_i(x_i^k)-F_i(\bar{x}_i^k))-Q_i^*(u^k-\bar{u}^k)\rangle\\ &+\beta_k\|u^k-\bar{u}^k\|^2+\sum(\mu_{F_i}+\mu_{A_i})\|\bar{x}_i^k-x_i^*\|^2\\ =&\sum\alpha_k^{-1}\left(\|x_i^k-\bar{x}_i^k\|^2-\alpha_k\langle x_i^k-\bar{x}_i^k,\,F_i(x_i^k)-F_i(\bar{x}_i^k)\rangle\right)\\ &-\sum\langle Q_i(x_i^k-\bar{x}_i^k),\,u^k-\bar{u}^k\rangle+\beta_k\|u^k-\bar{u}^k\|^2+\sum(\mu_{F_i}+\mu_{A_i})\|\bar{x}_i^k-x_i^*\|^2,\end{aligned}$$

which, together with $\alpha_k > 0$ and $\beta_k = \sum \beta_{i,k}$, yields

$$\begin{aligned}&\sum \langle x_i^k - x_i^*, x_i^k - \bar{x}_i^k - \alpha_k(F_i(x_i^k) - F_i(\bar{x}_i^k)) - \alpha_k Q_i^*(u^k - \bar{u}^k)\rangle \\ &+ \langle u^k - u^*, \alpha_k \beta_k (u^k - \bar{u}^k)\rangle \\ \geq &\sum \left(\|x_i^k - \bar{x}_i^k\|^2 - \alpha_k \langle x_i^k - \bar{x}_i^k, F_i(x_i^k) - F_i(\bar{x}_i^k)\rangle\right) \\ &- \sum \alpha_k \langle Q_i(x_i^k - \bar{x}_i^k), u^k - \bar{u}^k\rangle + \sum \alpha_k \beta_{i,k} \|u^k - \bar{u}^k\|^2 + \sum \alpha_k(\mu_{F_i} + \mu_{A_i})\|\bar{x}_i^k - x_i^*\|^2.\end{aligned}$$

Thus, in view of discussions from (57) to (59),

$$\begin{aligned}&\sum \left(\|x_i^k - \bar{x}_i^k\|^2 - \alpha_k \langle x_i^k - \bar{x}_i^k, F_i(x_i^k) - F_i(\bar{x}_i^k)\rangle\right) - \sum \alpha_k \langle Q_i(x_i^k - \bar{x}_i^k), u^k - \bar{u}^k\rangle + \sum \alpha_k \beta_{i,k}\|u^k - \bar{u}^k\|^2 \\ \geq &\frac{1}{2}\varepsilon \left(\sum \|x_i^k - \bar{x}_i^k\|^2 + \|u^k - \bar{u}^k\|^2\right).\end{aligned}$$

These relations show that

$$\langle w^k - w^*, d^k\rangle \geq \langle w^k - \bar{w}^k, d^k\rangle + \sum \alpha_k(\mu_{F_i} + \mu_{A_i})\|\bar{x}_i^k - x_i^*\|^2 \geq \frac{1}{2}\varepsilon \|w^k - \bar{w}^k\|^2 + \sum \alpha_k(\mu_{F_i} + \mu_{A_i})\|\bar{x}_i^k - x_i^*\|^2.$$

So, according to this relation, (36) and (37), we get

$$\begin{aligned}&\|w^{k+1} - w^*\|^2 \\ \leq &\|w^k - w^* - \theta_k \gamma_k d^k\|^2 \\ = &\|w^k - w^*\|^2 - 2\theta_k \gamma_k \langle w^k - w^*, d^k\rangle + \theta_k^2 \gamma_k^2 \|d^k\|^2 \\ \leq &\|w^k - w^*\|^2 - 2\theta_k \gamma_k \langle w^k - \bar{w}^k, d^k\rangle + \theta_k^2 \gamma_k^2 \|d^k\|^2 - 2\theta_k \gamma_k \sum \alpha_k(\mu_{F_i} + \mu_{A_i})\|\bar{x}_i^k - x_i^*\|^2,\end{aligned}$$

where the first inequality follows from non-expansiveness of projection operator. Combining this with (35), i.e., $\gamma_k = \langle w^k - \bar{w}^k, d^k\rangle / \|d^k\|^2$ yields

$$\begin{aligned}&\|w^{k+1} - w^*\|^2 \\ \leq &\|w^k - w^*\|^2 - \theta_k(2 - \theta_k)\left(\langle w^k - \bar{w}^k, d^k\rangle / \|d^k\|^2\right)\langle w^k - \bar{w}^k, d^k\rangle - 2\theta_k \gamma_k \sum \alpha_k(\mu_{F_i} + \mu_{A_i})\|\bar{x}_i^k - x_i^*\|^2 \\ = &\|w^k - w^*\|^2 - \theta_k(2 - \theta_k)\gamma_k \langle w^k - \bar{w}^k, d^k\rangle - 2\theta_k \gamma_k \sum \alpha_k(\mu_{F_i} + \mu_{A_i})\|\bar{x}_i^k - x_i^*\|^2 \\ \leq &\|w^k - w^*\|^2 - \theta_k(2 - \theta_k)\gamma_k \frac{1}{2}\varepsilon \|w^k - \bar{w}^k\|^2 - 2\theta_k \gamma_k \sum \alpha_k(\mu_{F_i} + \mu_{A_i})\|\bar{x}_i^k - x_i^*\|^2.\end{aligned}$$

So, it follows from Theorem 4.2 that

$$\|w^{k+1} - w^*\|^2 \leq \|w^k - w^*\|^2 - \theta_k(2 - \theta_k)\hat{\gamma}\frac{1}{2}\varepsilon \|w^k - \bar{w}^k\|^2 - 2\theta_k \hat{\gamma} \sum \alpha_k(\mu_{F_i} + \mu_{A_i})\|\bar{x}_i^k - x_i^*\|^2. \tag{66}$$

We can conclude that: (a) The limit of the sequence $\{\|w^k - w^*\|\}$ exists and $\{w^k\}$ is bounded in norm; (b) The condition (61) implies that $\{\|w^k - \bar{w}^k\|\}$ has some subsequence, say $\{\|w^{k_j} - \bar{w}^{k_j}\|\}$, which converges to zero. It can be easily seen that

$$(i)\ \ x_i^{k_j} - \bar{x}_i^{k_j} \to 0,\ i = 1, \ldots, n,\ \ u^{k_j} - \bar{u}^{k_j} \to 0; \tag{67}$$

$$(ii)\ \ \{x_i^{k_j}\},\ i = 1, \ldots, n,\ \ \{u^{k_j}\}\ \text{ are bounded in norm,} \tag{68}$$

It follows from (68) that $\{(x_i^{k_j}, u^{k_j})\}$ has some subsequence that converges weakly. In addition, α_{k_j} is bounded as well. Without loss of generality, we assume that

$$x^{k_j} \rightharpoonup x^\infty, \quad u^{k_j} \rightharpoonup u^\infty, \quad \alpha_{k_j} \to \alpha_\infty.$$

Meanwhile, in view of (67), we also have

$$\bar{x}^{k_j} \rightharpoonup x^\infty, \quad \bar{u}^{k_j} \rightharpoonup u^\infty.$$

Next, we will make use of Lemma A3 to prove the remaining part. To this end, we follow the definition of the set T to get

$$T(\bar{x}^k, F(x^k), \bar{u}^k) \ni \begin{pmatrix} A_1(\bar{x}_1^k) + F_1(x_1^k) + Q_1^*\bar{u}^k \\ \vdots \\ A_n(\bar{x}_n^k) + F_n(x_n^k) + Q_n^*\bar{u}^k \\ x_1^k - \bar{x}_1^k \\ \vdots \\ x_n^k - \bar{x}_n^k \\ B^{-1}(\bar{u}^k) - \sum_{i=1}^n Q_i \bar{x}_i^k + q \end{pmatrix}.$$

The inclusion relation above is also true for k_j,

$$T(\bar{x}^{k_j}, F(x^{k_j}), \bar{u}^{k_j}) \ni \begin{pmatrix} A_1(\bar{x}_1^{k_j}) + F_1(x_1^{k_j}) + Q_1^* \bar{u}^{k_j} \\ \vdots \\ A_n(\bar{x}_n^{k_j}) + F_n(x_n^{k_j}) + Q_n^* \bar{u}^{k_j} \\ x_1^{k_j} - \bar{x}_1^{k_j} \\ \vdots \\ x_n^{k_j} - \bar{x}_n^{k_j} \\ B^{-1}(\bar{u}^{k_j}) - \sum_{i=1}^n Q_i \bar{x}_i^{k_j} + q \end{pmatrix}.$$

Combining this with (62) and (64) yields

$$T(\bar{x}^{k_j}, F(x^{k_j}), \bar{u}^{k_j}) \ni \begin{pmatrix} \alpha_{k_j}^{-1}(x_1^{k_j} - \bar{x}_1^{k_j}) - Q_1^*(u^{k_j} - \bar{u}^{k_j}) \\ \vdots \\ \alpha_{k_j}^{-1}(x_n^{k_j} - \bar{x}_n^{k_j}) - Q_n^*(u^{k_j} - \bar{u}^{k_j}) \\ x_1^{k_j} - \bar{x}_1^{k_j} \\ \vdots \\ x_n^{k_j} - \bar{x}_n^{k_j} \\ \beta_k(u^{k_j} - \bar{u}^{k_j}) \end{pmatrix}. \tag{69}$$

To invoke Lemma A3, we first consider all the terms on the right-hand side of (69).

Case 1. $\alpha_\infty > 0$. In this case, we certainly have

$$\alpha_{k_j} > 0,\ \alpha_\infty > 0,\ x_i^{k_j} - \bar{x}_i^{k_j} \to 0 \quad \Rightarrow \quad \alpha_{k_j}^{-1}(x_i^{k_j} - \bar{x}_i^{k_j}) \to 0.$$

Case 2. $\alpha_\infty = 0$. Note that our choice of α_{k_j} implies that (38) and (39) fail to hold simultaneously for $\bar{\alpha}_{k_j} = t^{-1}\alpha_{k_j}$, if k_j is large enough. Thus, we get either

$$(1-\rho)\,\|x_i^{k_j} - x_i^{k_j}(\bar{\alpha}_{k_j})\|^2 < \bar{\alpha}_{k_j}\langle x_i^{k_j} - x_i^{k_j}(\bar{\alpha}_{k_j}),\ F_i(x_i^{k_j}) - F_i(x_i^{k_j}(\bar{\alpha}_{k_j}))\rangle$$

or

$$\|x_i^{k_j} - x_i^{k_j}(\bar{\alpha}_{k_j}) - \bar{\alpha}_{k_j}(F(x_i^{k_j}) - F(x_i^{k_j}(\bar{\alpha}_{k_j})))\|^2 > \tau\,\langle x_i^{k_j} - x_i^{k_j}(\bar{\alpha}_{k_j}),\ x_i^{k_j} - x_i^{k_j}(\bar{\alpha}_{k_j}) - \bar{\alpha}_{k_j}(F(x_i^{k_j}) - F(x_i^{k_j}(\bar{\alpha}_{k_j})))\rangle.$$

Denote by $\varphi(\alpha) = \alpha^{-1}\|x_i^{k_j} - x_i^{k_j}(\alpha)\|$. By making use of Cauchy–Schwarz inequality, we can further get

$$(1-\rho)\,\varphi^2(\bar{\alpha}_{k_j}) < \varphi(\bar{\alpha}_{k_j})\|F_i(x_i^{k_j}) - F_i(x_i^{k_j}(\bar{\alpha}_{k_j}))\|$$

and

$$|\tau - 2|\varphi(\bar{\alpha}_{k_j})\|F(x_i^{k_j}) - F(x_i^{k_j}(\bar{\alpha}_{k_j}))\| + \|F(x_i^{k_j}) - F(x_i^{k_j}(\bar{\alpha}_{k_j}))\|^2 > (\tau - 1)\varphi^2(\bar{\alpha}_{k_j}),$$

respectively. Note that the discussion of the latter's inequality is similar to that of (56). In either subcase, we confirm that $\varphi(\bar{\alpha}_{k_j}) \to 0$ as $\alpha_{k_j} \to 0$, since $\varphi(\bar{\alpha}_{k_j})$ is bounded and $F_i(x_i^{k_j}) - F_i(x_i^{k_j}(\bar{\alpha}_{k_j})) \to 0$ (the latter is implied by uniform continuity of F_i and $x_i^{k_j} - x_i^{k_j}(\bar{\alpha}_{k_j}) \to 0$). On the other hand, it follows from the properties (A.4) that $\varphi(\alpha_{k_j}) \le t^{-1}\varphi(\bar{\alpha}_{k_j})$. Therefore, we have

$$\varphi(\alpha_{k_j}) \to 0 \quad \Rightarrow \quad \alpha_{k_j}^{-1}(x_i^{k_j} - \bar{x}_i^{k_j}) \to 0, \quad \text{as } \alpha_{k_j} \to 0.$$

Since β_k, Q_i^* $(i = 1, \ldots, n)$ and (67) are bounded, each term on the right-hand side of (69) strongly converges to zero. Moreover, by the assumption and continuity of F, we have

$$\bar{x}^{k_j} \rightharpoonup x^\infty,\quad F(x^{k_j}) \rightharpoonup F(x^\infty),\quad \bar{u}^{k_j} \rightharpoonup u^\infty,$$

where the notation $\rightharpoonup$ stands for weak convergence. Denoting $F(x)$ by z, it follows from Lemma A3 that this cluster point $(x^\infty, z^\infty, u^\infty)$ solves $0 \in T(x, z, u)$ as desired and the primal cluster point x^∞ solves the problem (1) as well. The proof of uniqueness of weak cluster point is standard, see [9,29] for more details. □

5. The dual-first version of Algorithm 2.1

As we know, Algorithm 2.1 first computes the primal intermediate point $\bar{x}^k$ then the dual intermediate point $\bar{u}^k$ at each iteration. In this section, we describe its dual-first version. This means that it first computes the dual intermediate point then the primal intermediate point at each iteration. For convenience, the dual-first version of Algorithm 2.1 is called Algorithm 5.1 from now on.

Algorithm 5.1. The dual-first version of Algorithm 2.1

Step 0. For $i = 1, \dots, n$, choose $x_i^0 \in \mathcal{H}_i$, $u^0 \in \mathcal{G}$, $t \in (0,1)$, $\rho \in (0,1)$ and $\alpha_{i,-1} > 0$. Set $k := 0$.
Step 1. Choose $\beta_k > 0$ and find $\bar{u}^k$ such that

$$\left(\beta_k I + B^{-1}\right)(\bar{u}^k) \ni \beta_k u^k + \sum_{i=1}^n Q_i x_i^k - q.$$

Step 2. For $i = 1, \dots, l$, choose appropriate $\alpha_{i,k} > 0$. For $i = l+1, \dots, n$, find the smallest j_k in (19) such that

$$x_i^k(\alpha_i) = (I + \alpha_i A_i)^{-1}(x_i^k - \alpha_i(F_i(x_i^k) + Q_i^* \bar{u}^k))$$

satisfies

$$\alpha_i \langle x_i^k - x_i^k(\alpha_i),\ F_i(x_i^k) - F_i(x_i^k(\alpha_i))\rangle \le (1-\rho)\|x_i^k - x_i^k(\alpha_i)\|^2. \tag{70}$$

Take $\alpha_{i,k} = \alpha_{i,k-1} t^{j_k}$, and compute $\bar{x}_i^k = x_i^k(\alpha_{i,k})$. If $\bar{x}_i^k = x_i^k$, $i = 1, \dots, n$, and $\bar{u}^k = u^k$, then stop. Otherwise go to Step 3.
Step 3. Compute

$$d_{x_i}^k = \alpha_{i,k}^{-1}(x_i^k - \bar{x}_i^k) - (F_i(x_i^k) - F_i(\bar{x}_i^k)), \quad i = 1, \dots, n,$$
$$d_u^k = \beta_k(u^k - \bar{u}^k) + \sum_{i=1}^n Q_i(x_i^k - \bar{x}_i^k),$$
$$\gamma_k = \frac{\sum_{i=1}^n \langle x_i^k - \bar{x}_i^k, d_{x_i}^k\rangle + \langle u^k - \bar{u}^k, d_u^k\rangle}{\sum_{i=1}^n \|d_{x_i}^k\|^2 + \|d_u^k\|^2}.$$

Choose $\theta_k \in (0,2]$. Compute

$$x_i^{k+1} = x_i^k - \theta_k \gamma_k d_{x_i}^k,\ i = 1, \dots, n, \quad u^{k+1} = u^k - \theta_k \gamma_k d_u^k.$$

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

As to the β_k, formally we choose it via the same way as (29).

Theorem 5.1. *If Assumption 2.1 holds and $\sum_k \theta_k(2-\theta_k) = +\infty$, then Algorithm 5.1 is weakly convergent.*

Notice that the dual-first version of Algorithm 2.2 can be similarly derived and the associated weak convergence can be similarly analyzed, thus we omit it here.

Be aware that, generally speaking, Algorithm 5.1 may have the difficulty in implementing (70). This is because that $x_i^k(\alpha_i)$ relies on $\bar{u}^k$ whereas the latter is related to β_k. So, it seems somewhat impractical to choose appropriate β_k.

However, Algorithm 5.1 is well-suited for finding an $x \in \mathcal{H}$ such that

$$0 \in F(x) + A(x) + Q^* B(Qx - q),$$

where F is further assumed to be bounded linear and skew-adjoint. The resulting algorithm is the following

Algorithm 5.2. A special case of Algorithm 5.1

Step 0. Choose $x^0 \in \mathcal{H}$, $u^0 \in \mathcal{G}$. Choose $\alpha > 0$ and β via (73). Set $k := 0$.
Step 1. Find $\bar{u}^k$ such that

$$\left(\beta I + B^{-1}\right)(\bar{u}^k) \ni \beta u^k + Qx^k - q.$$

Step 2. Find $\bar{x}^k$ such that

$$(I + \alpha A)(\bar{x}^k) \ni x^k - \alpha(F(x^k) + Q^* \bar{u}^k).$$

If $\bar{x}^k = x^k$ and $\bar{u}^k = u^k$, then stop. Otherwise go to Step 3.
Step 3. Compute

$$d_x^k = \alpha^{-1}(x^k - \bar{x}^k) - (F(x^k) - F(\bar{x}^k)),$$
$$d_u^k = \beta(u^k - \bar{u}^k) + Q(x^k - \bar{x}^k),$$
$$\gamma_k = \frac{\langle x^k - \bar{x}^k, d_x^k\rangle + \langle u^k - \bar{u}^k, d_u^k\rangle}{\|d_x^k\|^2 + \|d_u^k\|^2}.$$

Choose $\theta_k \in (0,2]$. Compute

$$x^{k+1} = x^k - \theta_k \gamma_k d_x^k, \quad u^{k+1} = u^k - \theta_k \gamma_k d_u^k.$$

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

6. Applications to convex minimization

In this section, we briefly discuss how to apply Algorithms 2.1, 2.2 and 5.1 to solving convex minimization.

A first application is to the separable convex minimization

$$\text{minimize} \quad \sum\nolimits_{i=1}^{n} (f_i + g_i)(x_i), \quad \text{subject to} \quad \sum\nolimits_{i=1}^{n} Q_i x_i - q \in \mathcal{X}, \tag{71}$$

where $f_i,\ g_i : \mathcal{H}_i \to \mathcal{R}$ are closed proper convex functions, f_i are further assumed to be continuously differentiable with the gradient ∇f_i, and $\mathcal{X}$ is a nonempty closed convex set, onto which it is easy to project, say the first orthant or a ball. This problem can be rewritten as

$$\text{minimize} \quad \sum\nolimits_{i=1}^{n} (f_i + g_i)(x_i) + \delta_{\mathcal{X}} \left(\sum\nolimits_{i=1}^{n} Q_i x_i - q \right).$$

Under suitable assumptions, its optimal conditions are the following system of monotone inclusions

$$0 \in \nabla f_i(x_i) + \partial g_i(x_i) + Q_i^* \partial \delta_{\mathcal{X}} \left(\sum\nolimits_{i=1}^{n} Q_i x_i - q \right), \ i = 1, \dots, n.$$

This corresponds to (1) with $F_i = \nabla f_i$, $A_i = \partial g_i$ and $B = \partial \delta_{\mathcal{X}}$ respectively.

A second application is to the following convex minimization

$$\text{minimize} \quad f(x) + g(x) - \sum\nolimits_{i=1}^{p} \ln x_i - \ln(1 - \langle e, x \rangle),$$

where $f : \mathcal{R}^p \to \mathcal{R}$ is a continuously differentiable convex function with the gradient ∇f, and $g : \mathcal{R}^p \to \mathcal{R}$ is a closed proper convex function, and e is a p-dimensional vector of all ones. Additionally, we assume that the set $\mathcal{S} = \{x \in \mathcal{R}^p : x \geq 0, \langle e, x \rangle \leq 1\}$ is included in both $\mathrm{dom} f$ and $\mathrm{dom} g$.

Under suitable assumptions, its optimal condition reads

$$0 \in \nabla f(x) + \partial g(x) + Q^T B (Qx - q).$$

This corresponds to (1) with $n = 1$, $F = \nabla f$, $A = \partial g$ and

$$B(z) = -\mathrm{diag} \left[\frac{1}{z_1}, \dots, \frac{1}{z_{p+1}} \right], \quad Q = \begin{pmatrix} I_p \\ -e^T \end{pmatrix}, \quad q = (0, \dots, 0, -1)^T,$$

respectively, where I_p is the identity matrix. Obviously, B is continuous and monotone, thus maximal monotone, in its effective domain $\{z \in \mathcal{R}^{p+1} : z_i > 0, i = 1, \dots, p+1\}$.

7. Implementation details

In this section, we discuss some key implementation details of the aforementioned splitting methods so as to assure and improve their individual easiness and efficiency in practice.

7.1. How to choose beta

Let

$$\beta_{i,k}(\varepsilon; t) = \frac{2t\varepsilon - \varepsilon^2 + \alpha_{i,k}^2 \|Q_i\|_1 \|Q_i\|_\infty}{2(2t - \varepsilon)\alpha_{i,k}} = \frac{\varepsilon}{2\alpha_{i,k}} + \frac{\alpha_{i,k} \|Q_i\|_1 \|Q_i\|_\infty}{2(2t - \varepsilon)},$$

where $t > 0$ is given. The function β here is increasing with respect to $\varepsilon \in (0, 2t)$.

For Algorithms 2.1 and 2.2, we choose β_k via

$$\begin{aligned} \beta_k &= \sum\nolimits_{i=1}^{n} \beta_{i,k} \\ &= \sum\nolimits_{i=1}^{l} \max\{\beta_{i,k}(\varepsilon; 1), \alpha_{i,k}\} + \sum\nolimits_{i=l+1}^{n} \max\{\beta_{i,k}(\varepsilon; \rho), \alpha_{i,k}\}. \end{aligned} \tag{72}$$

For Algorithm 5.2, we choose β via

$$\beta = \max\{\beta(\varepsilon; 1), \alpha\}, \tag{73}$$

where $\varepsilon > 0$ is sufficiently small, say $\varepsilon = 10^{-9}$, even smaller.

7.2. How to adapt Q

If $n = 1$, then the problem (71) reduces to

$$\text{minimize} \quad f(x) + g(x), \quad \text{subject to} \quad Qx = q.$$

In this case, we check if $\sqrt{\|Q\|_1 \|Q\|_\infty} < 2$ is satisfied. If not, we adapt

$$(Q, q) \leftarrow (Q, q) / \sqrt{\|Q\|_1 \|Q\|_\infty}.$$

7.3. How to solve subproblem

In these algorithms above, there is one type of sub-problem $(I+\alpha A)(x) \ni w$. Now we discuss how to solve it. (i) If A is further linear, then we may use Matlab solver via

$$x = (I+\alpha A) \setminus w.$$

(ii) If $A = \nabla f$ is the gradient of some continuously differentiable convex function f, then we may resort to quasi-Newton method with novel conditions using solely gradient to locate steplength; see [34, Sect. 5.3]. (iii) If $A = F$ is continuously differentiable, then we may use some Newton-type method to solve this sub-problem.

The other type of sub-problem is $\left(\beta I + B^{-1}\right)(u) \ni p$. We may solve it directly. Of course, if it is easier to evaluate the resolvent of B, we instead consider Moreau identity

$$\left(\beta I + B^{-1}\right)^{-1} \equiv \beta^{-1} I - \beta^{-1}(I+\beta B)^{-1}, \quad \forall \beta > 0.$$

Thus, the process of solving this sub-problem can be divided into

$$\tilde{p} = (I+\beta B)^{-1}(p), \quad u = \beta^{-1}(p - \tilde{p}).$$

7.4. How to deal with (39) in Algorithm 2.2

Upon reviewing the convergence proofs of Algorithm 2.2, we observed that (38) and (39) serve different purposes. The former ensures monotone decrease in $\|w^k - w^*\|$, while the latter's significance lies in providing a theoretically rigorous proof of convergence for Algorithm 2.2. Consequently, (39) may no longer be necessary in practice, as Algorithm 2.2 often performs well for the first several hundred iterations (assuming it performs well at all).

8. Rudimentary experiments

In this section, we confirmed the effectiveness of our proposed splitting algorithms. In our writing style, rather than striving for maximal test problems, we tried to make the basic ideas and techniques as clear as possible.

We performed all numerical experiments on a desktop computer equipped with a 3.00 GHz Intel(R) Core(TM) i5-7400 CPU and 8.00 GB of memory. The MATLAB R2020a platform was used as the implementation environment.

We compared our proposed splitting algorithms with other state-of-the-art splitting algorithms, selected for their similarities in features, applicability, and implementation effort.

Our first test problem is from [7], which is to find an $x \in \mathcal{R}^m$ such that

$$0 \in Dx - d + Q^* \partial \delta_C (Qx - q),$$

where

$$D = \text{tridiag}\,(-1-h, 4+2h, -1), \quad h = 1/(m+1),$$

and

$$Q = [eye(m); (-1/m) * ones(1,m)]; \quad q = [zeros(m,1); -1/m]$$

and $C \subseteq \mathcal{R}^{m+1}$ is the first orthant. To ensure that $e_1 = (1,0,\ldots,0)^T$ solves it, we set $d = De_1$ in our practical implements. Thus, the problem's unique solution is $x^* = e_1$. We chose

$$F = 0.5(D - D^T), \quad A(x) = 0.5(D + D^T)x - d, \quad B = \partial \delta_C$$

to match the problem (1), and we chose $x^0 = (0,\ldots,0)^T$ as the starting point, as done in [7].

We chose this particular problem because it features the constraint set (given by a general half-space and the first orthant) onto which it is easy to project individually, but whose intersection poses a more difficult projection problem. This property makes it relatively tricky to apply some splitting methods such as those [10,21] as they cannot fully split the problem and thus have to perform an extra and nontrivial projection per iteration.

In practical implementations, we set $m = 1000$. In this case, we via Matlab got

$$\frac{1}{4}\|Q\|_1\|Q\|_\infty \approx 0.2508, \quad \frac{1}{4}\|Q\|^2 \approx 0.2503.$$

VC splitting: A splitting method of Vu [33] and Condat [2], also described in [8, Algorithm 6]. We implemented it in the same way as [8].

ES1: An extended splitting method recently proposed in [7]. We implemented it in the same way as [7].

JE splitting: A splitting method of Johnstone and Eckstein, whose parameters, suggested in [23], were

$$\rho_{3,0} = 1, \quad \rho_{1,k} = \rho_{2,k} \equiv 1, \quad \Delta = 1, \quad \gamma = 10.$$

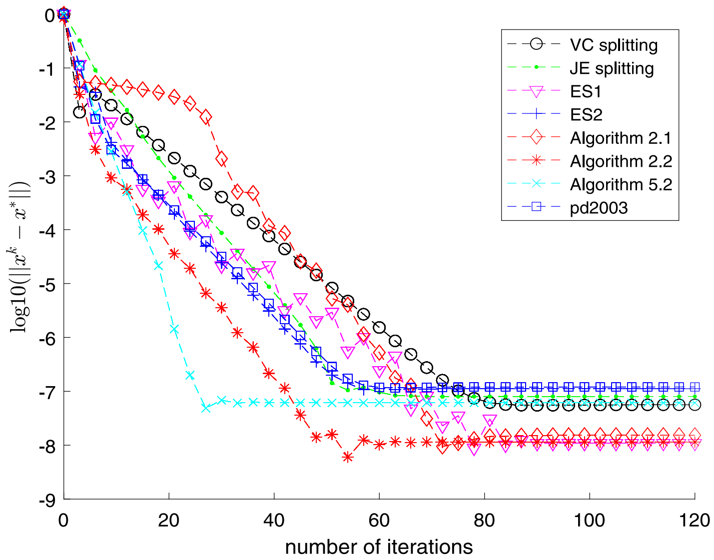

Fig. 1. Numerical results on the first test problem.

Moreover, we set $G_1 = Q$, $G_2 = eye(m)$, $z^1 = zeros(m, 1)$ and

$$w_1^1 = zeros(m+1, 1), \quad w_2^1 = zeros(m, 1), \quad w_3^1 = -G_1^T w_1^1 - G_2^T w_2^1.$$

ES2: An extended splitting method recently proposed in [8]. We implemented it in the same way as [8].

Notice that, for the following four algorithms, we chose

$$\alpha \in \{0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8\}$$

to be close to $1/p \approx 0.25$, with $p = \text{trace}(0.5(D + D^T))/m$.

Algorithm 2.1: We chose $\alpha = 0.6$ and $\theta_k \equiv \theta = 1.0$, and we chose β via (73), with $\epsilon = 10^{-9}$, i.e., $\beta = \alpha$.

Algorithm 2.2: We chose $\alpha = 0.5$, and we chose θ_k and β in the same as those in Algorithm 2.1.

Algorithm 5.2: We chose α, θ_k and β in the same as those in Algorithm 2.2.

pd2003: A primal–dual splitting method described in (11). We chose $\alpha = 0.5$ and $\theta_k \equiv \theta = 1.8$.

For the last four algorithms above, we set $x^0 = zeros(m, 1)$ and $u^0 = zeros(m + 1, 1)$.

Numerical results on the first test problem were given in Figs. 1–2. From Fig. 1, we can see convergence behaviors of these eight different algorithms in the first 120 iterations. From Fig. 2, we can see that ES1, Algorithms 2.1 and 2.2 (in the $n = 1$ case) achieved the accuracy 10^{-8}, respectively. Furthermore, Algorithms 2.2 and 5.2 were faster than all the others.

An interesting observation is that $\{\|x^k - x^*\|\}$, where x^k is generated by Algorithm 2.2, is not necessarily monotonically decreasing. This is because that what we have proved in (66) is merely for monotonicity of $\{\|w^k - w^*\|\}$.

Our second test problem is to solve the following monotone inclusion

$$0 \in F(x) + A(x) + Q^* B(Qx - q),$$

where $Q = \text{diag}(1, \dots, n)$, $q = (0, \dots, 0)^T$ and

$$F(x) = (\sqrt[3]{x_1}, x_2, \dots, x_n)^T, \; A = \text{tridiag}(n, 1, -n), \; B = \text{diag}(\arctan, \dots, \arctan).$$

Note that such F is uniformly continuous and x^*=0 is its unique solution. We took $x^0 = \frac{1}{n}(1, 1, \dots, 1)^T$ as the starting point and $\|x^k - x^*\| \le \epsilon$ as the stopping criterion.

Algorithm 2.2: We chose β_k via (72)

$$\beta_k = \max\{\beta_k(\epsilon; \rho), \alpha_k\}, \qquad \text{with} \quad \epsilon = 10^{-9},$$

and chose $t = 0.5$, $\rho = 0.1$ and $\theta_k \equiv \theta = 1.0$. By the way, if we chose $\epsilon = 10^{-40}$, we observed the same numerical results.

pd2003: A primal–dual splitting method described in (11). We chose $t = 0.5$, $\rho = 0.1$, $\theta_k \equiv \theta = 1.8$.

Numerical results on the second test problem were reported in Table 1, where the format "number of iteration/ CPU time (in seconds)" was used and "–" means failure of the desired accuracy within the first 200 iterations. For Algorithm 2.2, the condition (39) was no longer applied.

From Table 1, we can see that Algorithm 2.2 was by far faster than pd2003 because the former no longer entails time-consuming computations of B's resolvent.

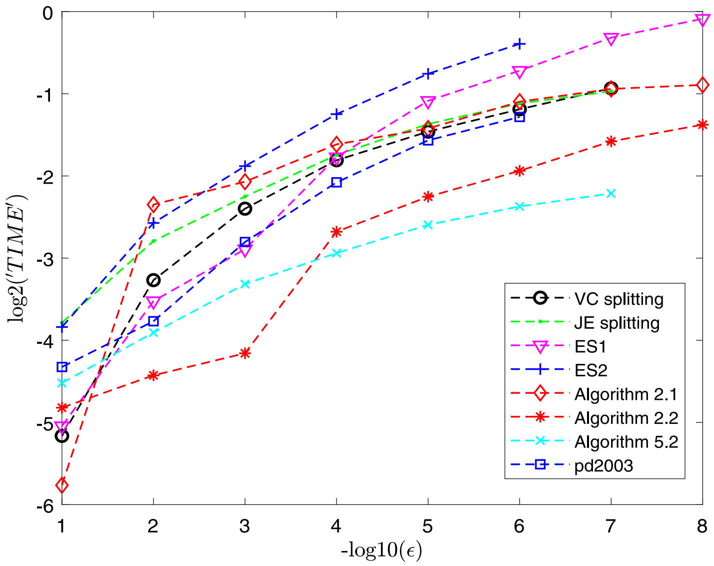

Fig. 2. Numerical results on the first test problem.

Table 1
Numerical results on the second test problem.

ϵ	Algorithm	$n = 5$	$n = 10$	$n = 50$	$n = 100$
10^{-2}	Algo2.2	11/2.655	8/1.953	38/40.23	84/176.2
	pd2003	59/13.01	10/6.637	33/75.65	60/250.1
10^{-3}	Algo2.2	28/3.104	27/5.908	141/146.2	145/305.6
	pd2003	–	–	–	93/394.5
10^{-4}	Algo2.2	183/18.87	173/36.51	–	–
	pd2003	–	–	–	–

9. Conclusions

In this article, we have proposed two new splitting methods for solving systems of three-operator monotone inclusions in real Hilbert spaces, where the third operator is linearly composed. These methods primarily involve evaluating the first operator and computing resolvents with respect to the other two operators. Importantly, they fully decouple the third operator from its linear composition operator. One of these methods is specifically designed for the case where the first operator is Lipschitz continuous. We have provided back-tracking techniques to determine appropriate step lengths and also propose a dual-first version of this method. For the other method, which corresponds to a uniformly continuous operator, we have developed innovative back-tracking techniques, incorporating additional conditions to determine step lengths. The weak convergence of either method is proven using characteristic operator techniques. We also have discussed implementation details to enhance the user-friendliness of these methods. To validate the efficiency of our proposed splitting methods, including their special cases and variations, we have conducted numerical experiments and compared their performance with other state-of-the-art methods. The results demonstrated the effectiveness of our proposed splitting methods in solving test problems. In future work, we plan to further investigate these splitting methods and propose their variable metric variants, incorporating relative errors as discussed in Ref. [5]. By exploring these extensions, we aim to enhance the capabilities and applicability of our proposed splitting methods for solving systems of three-operator monotone inclusions.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The author is greatly indebted to the Editor-in-Chief, Professor Laura Gardini, the handling editor and the referees for their encouraging words and insightful suggestions, which improve the quality of this article. Special thanks go to Yiyi Li for writing Matlab codes and Yuanyuan Zhang and Mengke Zhu for careful reading of the current version of this manuscript. This work was partially supported by the National Natural Science Foundation of China (12201075).

Appendix A

Let $A : \mathcal{H} \rightrightarrows \mathcal{H}$ be an operator. It is called monotone iff

$$\langle x - x', a - a' \rangle \geq 0, \quad \forall (x, a) \in A, \ \forall (x', a') \in A;$$

maximal monotone iff it is monotone and for given $\hat{x} \in \mathcal{H}$ and $\hat{a} \in \mathcal{H}$ the following implication relation holds

$$\langle x - \hat{x}, a - \hat{a} \rangle \geq 0, \quad \forall (x, a) \in A \quad \Rightarrow \quad (\hat{x}, \hat{a}) \in A.$$

Let $A : \mathcal{H} \rightrightarrows \mathcal{H}$ be an operator. It is called μ-monotone if there exists some $\mu \geq 0$ such that

$$\langle x - x', a - a' \rangle \geq \mu \|x - x'\|^2, \quad \forall (x, a) \in A, \ \forall (x', a') \in A.$$

If $\mu > 0$, then it is usually called μ-strongly monotone.

Denote by $x = (x_1^T, \dots, x_n^T)^T$ and

$$F = \begin{pmatrix} F_1 & & \\ & \ddots & \\ & & F_n \end{pmatrix}, \ A = \begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_n \end{pmatrix}, \ Q = [Q_1, \dots, Q_n].$$

Lemma A1 ([8]). *For the system of monotone inclusions (1), we introduce the dual variable $u \in \mathcal{G}$. Then*

$$T(x, z, u) = \begin{pmatrix} A & & \\ & F^{-1} & \\ & & B^{-1} \end{pmatrix} \begin{pmatrix} x \\ z \\ u \end{pmatrix} + \begin{pmatrix} 0 & I & Q^* \\ -I & 0 & 0 \\ -Q & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ z \\ u \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ q \end{pmatrix} \tag{A.1}$$

must be maximal monotone.

In this article, such T is called the characteristic operator or eigenoperator [8,11,35] with respect to the problem (1) above.

Lemma A2. *Let $F : \mathcal{H} \to \mathcal{H}$ and $A : \mathcal{H} \rightrightarrows \mathcal{H}$ be maximal monotone. Denote by*

$$x(\alpha) = (I + \alpha A)^{-1}(x - \alpha F(x)).$$

Assume that $x \in domF \cap domA$ and there exists $\bar{\alpha}$ such that

$$x(\alpha) \in domA, \quad \forall \alpha \in (0, \bar{\alpha}).$$

Then the following hold

$$(a) \ \ \alpha \to 0 \quad \Rightarrow \quad x - x(\alpha) \to 0, \tag{A.2}$$

$$(b) \ \ \alpha^{-1}\|x - x(\alpha)\| \leq \min\{\|w\| \ : w \in F(x) + A(x)\}, \tag{A.3}$$

$$(c) \ \ \liminf_{\alpha \to 0} \alpha^{-1}\|x - x(\alpha)\| = \min\{\|w\| \ : w \in F(x) + A(x)\}, \tag{A.4}$$

$$(d) \ \ 1 \leq \frac{\|x - x(\alpha)\|}{\|x - x(\alpha')\|} \leq \frac{\alpha}{\alpha'}, \quad \forall \alpha' \in (0, \alpha]. \tag{A.5}$$

Moreover, if $F + A$ is maximal monotone on $domF \cap domA$, then the minimum on the right-hand side of either (A.3) or (A.4) must be uniquely attainable.

Proof. We first prove (A.2) and (A.3). In view of the notation $x(\alpha)$, we have

$$\frac{x - x(\alpha)}{\alpha} - F(x) \in A(x(\alpha)), \quad \alpha > 0,$$

which, together with $w - F(x) \in A(x)$ and monotonicity of A, implies

$$0 \leq \langle x(\alpha) - x, \frac{x - x(\alpha)}{\alpha} - F(x) - w + F(x) \rangle = \langle x(\alpha) - x, \frac{x - x(\alpha)}{\alpha} - w \rangle.$$

By making use of the Cauchy–Schwarz inequality, we further get

$$\frac{1}{\alpha}\|x - x(\alpha)\|^2 \leq \langle w, x - x(\alpha) \rangle \leq \|w\|\|x - x(\alpha)\|.$$

So, we conclude that either $x - x(\alpha) = 0$ or $\|x - x(\alpha)\|/\alpha \leq \|w\|$.

Finally, we refer to [1,19] for proofs of (A.4) and (A.5) respectively. □

Below, we give a well-known result.

Lemma A3. *Consider any maximal monotone operator $T : \mathcal{H} \rightrightarrows \mathcal{H}$. Assume that the sequence $\{w^k\}$ in $\mathcal{H}$ converges weakly to w, and the sequence $\{s^k\}$ on $domT$ converges strongly to s. If $T(w^k) \ni s^k$ for all k, then the relation $T(w) \ni s$ must hold.*

For a short proof of Lemma A3, we refer to [32] and the references cited therein.

Lemma A4. *Assume that $F : \mathcal{H} \to \mathcal{H}$ is continuous and monotone and $A : \mathcal{H} \rightrightarrows \mathcal{H}$ is maximal monotone. If F and A are further μ_F-monotone and μ_A-monotone, respectively, and*

$$A(\bar{x}) \ni \alpha^{-1}(x - \bar{x}) - F(x) - Q^* u, \tag{A.6}$$

$$A(x^*) \ni -F(x^*) - Q^* u^*, \tag{A.7}$$

where $\alpha > 0$, then the following inequality holds

$$\begin{aligned} &\langle x - x^*, \alpha^{-1}(x - \bar{x}) - (F(x) - F(\bar{x}))\rangle - \langle Q(\bar{x} - x^*), u - u^*\rangle \\ &\geq \langle x - \bar{x}, \alpha^{-1}(x - \bar{x}) - (F(x) - F(\bar{x}))\rangle + (\mu_F + \mu_A)\|\bar{x} - x^*\|^2. \end{aligned}$$

Proof. In view of (A.6) and (A.7) and A's μ_A-monotonicity, we get

$$\langle \bar{x} - x^*, \alpha^{-1}(x - \bar{x}) - (F(x) - F(x^*)) - Q^*(u - u^*)\rangle \geq \mu_A\|\bar{x} - x^*\|^2.$$

Combining this with F's μ_F-monotonicity, i.e.,

$$\langle \bar{x} - x^*, F(\bar{x}) - F(x^*)\rangle \geq \mu_F\|\bar{x} - x^*\|^2$$

yields the desired result. □

Lemma A5. *Let $Q : \mathcal{H} \to \mathcal{G}$ be nonzero, bounded and linear operator, and let $\alpha > 0$, $t \in \mathcal{R}$. If $4\alpha > t^2\beta\|Q\|^2$, then the following*

$$\langle x, \alpha x\rangle + \langle u, \beta u\rangle - t\langle Qx, \beta u\rangle \geq \varphi(\alpha, \beta, tQ)\left(\|x\|^2 + \|u\|^2\right)$$

holds for all $x \in \mathcal{H}$ and all $u \in \mathcal{G}$, where

$$\varphi(\alpha, \beta, tQ) = \frac{1}{2}\left(\alpha + \beta - \sqrt{(\alpha - \beta)^2 + t^2\beta^2\|Q\|^2}\right).$$

To our best knowledge, Lemma A5 or its equivalent version was given in [15, Sect. 3] and [24, Lemma 5.1]. Very recently, such a nice result was used in [10] and generalized in the author's 2017 manuscript of [8].

References

[1] H. Brézis, Opérateurs Maximaux Monotones, North-Holland, Amsterdam, 1973.
[2] L. Condat, A primal–dual splitting method for convex optimization involving Lipschitzian, proximable and linear composite terms, J. Optim. Theory Appl. 158 (2) (2013) 460–479.
[3] S. Dafermos, Traffic equilibrium and variational inequalities, Transp. Sci. 14 (1980) 42–54.
[4] Y.D. Dong, Splitting Methods for Monotone Inclusions, (PhD dissertation), Nanjing University, 2003.
[5] Y.D. Dong, A variable metric proximal-descent algorithm for monotone operators, J. Appl. Math. Comput. 38 (2012) 505–521.
[6] Y.D. Dong, Douglas-Rachford splitting method for semi-definite programming, J. Appl. Math. Comput. 51 (2016) 569–591.
[7] Y.D. Dong, Weak convergence of an extended splitting method for monotone inclusions, J. Global Optim. 79 (1) (2021) 257–277.
[8] Y.D. Dong, A new splitting method for systems of monotone inclusions in Hilbert spaces, Math. Comput. Simulation 203 (2023) 518–537.
[9] Y.D. Dong, A. Fischer, A family of operator splitting methods revisited, Nonlinear Anal. 72 (2010) 4307–4315.
[10] Y.D. Dong, X.H. Yu, A new splitting method for monotone inclusions of three operators, Calcolo 56 (1) (2019) Art. 3.
[11] Y.D. Dong, X. Zhu, An inertial splitting method for monotone inclusions of three operators, Int. J. Math. Stat. Oper. Res. 2 (1) (2022) 43–60.
[12] J. Eckstein, A simplified form of block-iterative operator splitting and an asynchronous algorithm resembling the multi-block alternating direction method of multipliers, J. Optim. Theory Appl. 173 (1) (2017) 155–182.
[13] J. Eckstein, D.P. Bertsekas, On the Douglas-Rachford splitting method and the proximal algorithm for maximal monotone operators, Math. Program. 55 (1992) 293–318.
[14] J. Eckstein, M.C. Ferris, Operator-splitting methods for monotone affine variational inequalities, with a parallel application to optimal control, Inf. J. Comput. 10 (1998) 218–235.
[15] J. Eckstein, B.F. Svaiter, A family of projective splitting methods for the sum of two maximal monotone operators, Math. Program. 111 (2008) 173–199.
[16] O. Güler, On the convergence of the proximal point algorithm for convex minimization, SIAM J. Control Optim. 29 (1991) 403–419.
[17] B.S. He, A class of projection and contraction methods for monotone variational inequalities, Appl. Math. Optim. 35 (1997) 69–76.
[18] H.J. He, D.R. Han, A distributed douglas-rachford splitting method for multi-block convex minimization problems, Adv. Comput. Math. 42 (1) (2016) 27–53.
[19] Y.Y. Huang, Y.D. Dong, New properties of forward–backward splitting and a practical proximal-descent algorithm, Appl. Math. Comput. 237 (2014) 60–68.
[20] A. Irschara, C. Zach, M. Klopschitz, H. Bischof, Large-scale, dense city reconstruction from user-contributed photos, Comput. Vis. Image Und. 116 (2012) 2–15.
[21] O.S. Iyiola, C.D. Enyi, Y. Shehu, Reflected three-operator splitting method for monotone inclusion problem, Optim. Methods Softw. 37 (4) (2022) 1527–1565.
[22] P.R. Johnstone, J. Eckstein, Single-forward-step projective splitting: exploiting cocoercivity, Comput. Optim. Appl. 78 (2021) 125–166.
[23] P.R. Johnstone, J. Eckstein, Projective splitting with forward steps, Math. Program. 191 (2022) 631–670.
[24] P. Latafat, P. Patrinos, Asymmetric forward–backward-adjoint splitting for solving monotone inclusions involving three operators, Comput. Optim. Appl. 68 (1) (2017) 57–93.
[25] P.L. Lions, B. Mercier, Splitting algorithms for the sum of two nonlinear operators, SIAM J. Numer. Anal. 16 (1979) 964–979.
[26] B. Martinet, Regularisation d'inéquations variationelles par approximations successives, Rev. Fr. d'Informatique Recherche Opér. 4 (1970) 154–158.
[27] M.A. Noor, Mixed quasi variational inequalities, Appl. Math. Comput. 146 (2003) 553–578.
[28] T. Pennanen, Dualization of generalized equations of maximal monotone type, SIAM J. Optim. 10 (2000) 809–835.

[29] R.T. Rockafellar, Monotone operators and the proximal point algorithm, SIAM J. Control Optim. 14 (1976) 877–898.
[30] M.V. Solodov, P. Tseng, Modified projection-type methods for monotone variational inequalities, SIAM J. Control Optim. 34 (1996) 1814–1830.
[31] D.F. Sun, A class of iterative methods for solving nonlinear projection equations, J. Optim. Theory Appl. 91 (1) (1996) 123–140.
[32] P. Tseng, A modified forward–backward splitting method for maximal monotone mappings, SIAM J. Control Optim. 38 (2) (2000) 431–446.
[33] B.C. Vu, A splitting algorithm for dual monotone inclusions involving cocoercive operators, Adv. Comput. Math. 38 (3) (2013) 667–681.
[34] D. Wicht, M. Schneider, T. Böhlke, An efficient solution scheme for small-strain crystal-elasto-viscoplasticity in a dual framework, Comput. Methods Appl. Mech. Engrg. 358 (2020) 112611.
[35] X. Zhu, Inertial Splitting Methods for Monotone Inclusions of Three Operators (Thesis), Zhengzhou University, 2020.