A splitting algorithm for dual monotone inclusions involving cocoercive operators*

Bằng Công Vũ

UPMC Université Paris 06
Laboratoire Jacques-Louis Lions – UMR CNRS 7598
75005 Paris, France

vu@ann.jussieu.fr

Abstract

We consider the problem of solving dual monotone inclusions involving sums of composite parallel-sum type operators. A feature of this work is to exploit explicitly the cocoercivity of some of the operators appearing in the model. Several splitting algorithms recently proposed in the literature are recovered as special cases.

Keywords: cocoercivity, forward-backward algorithm, composite operator, duality, monotone inclusion, monotone operator, operator splitting, primal-dual algorithm

Mathematics Subject Classifications (2010) 47H05, 49M29, 49M27, 90C25

1 Introduction

Monotone operator splitting methods have found many applications in applied mathematics, e.g., evolution inclusions [2], partial differential equations [1, 20, 23], mechanics [21], variational inequalities [6, 19], Nash equilibria [8], and various optimization problems [7, 9, 10, 14, 15, 17, 25, 29]. In such formulations, cocoercivity often plays a central role; see for instance [2, 6, 11, 13, 19, 20, 21, 23, 28, 29, 30]. Recall that an operator $C\colon \mathcal{H} \to \mathcal{H}$ is cocoercive with constant $\beta \in \left]0, +\infty\right[$ if its inverse is β-strongly monotone, that is,

$$(\forall x \in \mathcal{H})(\forall y \in \mathcal{H}) \quad \langle x - y \mid Cx - Cy\rangle \geq \beta\|Cx - Cy\|^2. \tag{1.1}$$

*This work was supported by the Agence Nationale de la Recherche under grant ANR-08-BLAN-0294-02 and the Vietnam National Foundation for Science and Technology Development.

In this paper, we revisit a general primal-dual splitting framework proposed in [16] in the presence Lipschitzian operators in the context of cocoercive operators. This will lead to a new type of splitting technique and provide a unifying framework for some algorithms recently proposed in the literature. The problem under investigation is the following, where the parallel sum operation is denoted by $\square$ (see (2.4)).

Problem 1.1 Let $\mathcal{H}$ be a real Hilbert space, let $z \in \mathcal{H}$, let m be a strictly positive integer, let $(\omega_i)_{1\leq i\leq m}$ be real numbers in $]0,1]$ such that $\sum_{i=1}^m \omega_i = 1$, let $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ be maximally monotone, and let $C\colon \mathcal{H} \to \mathcal{H}$ be μ-cocoercive for some $\mu \in \left]0,+\infty\right[$. For every $i \in \{1,\ldots,m\}$, let $\mathcal{G}_i$ be a real Hilbert space, let $r_i \in \mathcal{G}_i$, let $B_i\colon \mathcal{G}_i \to 2^{\mathcal{G}_i}$ be maximally monotone, let $D_i\colon \mathcal{G}_i \to 2^{\mathcal{G}_i}$ be maximally monotone and ν_i-strongly monotone for some $\nu_i \in \left]0,+\infty\right[$, and suppose that $L_i\colon \mathcal{H} \to \mathcal{G}_i$ is a nonzero bounded linear operator. The problem is to solve the primal inclusion

$$\text{find } \overline{x} \in \mathcal{H} \text{ such that } z \in A\overline{x} + \sum_{i=1}^m \omega_i L_i^*\big((B_i \,\square\, D_i)(L_i\overline{x} - r_i)\big) + C\overline{x}, \tag{1.2}$$

together with the dual inclusion

$$\text{find } \overline{v}_1 \in \mathcal{G}_1,\ldots,\overline{v}_m \in \mathcal{G}_m \text{ such that } (\exists\, x \in \mathcal{H}) \begin{cases} z - \sum_{i=1}^m \omega_i L_i^*\overline{v}_i \in Ax + Cx \\ \big(\forall i \in \{1,\ldots,m\}\big)\ \overline{v}_i \in (B_i \,\square\, D_i)(L_i x - r_i). \end{cases} \tag{1.3}$$

We denote by $\mathcal{P}$ and $\mathcal{D}$ the sets of solutions to (1.2) and (1.3), respectively.

In the case when $(D_i^{-1})_{1\leq i\leq m}$ and C are general monotone Lipschitzian operators, Problem 1.1 was investigated in [16]. Here are a couple of special cases of Problem 1.1.

Example 1.2 In Problem 1.1, set $z = 0$ and

$$\big(\forall i \in \{1,\ldots,m\}\big) \quad B_i\colon v \mapsto \{0\} \quad \text{and} \quad D_i\colon v \mapsto \begin{cases} \mathcal{G}_i & \text{if } v = 0, \\ 0 & \text{if } v \neq 0. \end{cases} \tag{1.4}$$

The primal inclusion (1.2) reduces to

$$\text{find } \overline{x} \in \mathcal{H} \text{ such that } 0 \in A\overline{x} + C\overline{x}. \tag{1.5}$$

This problem is studied in [2, 11, 13, 17, 23, 28, 29].

Example 1.3 Suppose that in Problem 1.1 the operators $(D_i)_{1\leq i\leq m}$ are as in (1.4), and that

$$A\colon x \mapsto \{0\} \quad \text{and} \quad C\colon x \mapsto 0. \tag{1.6}$$

Then we obtain the primal-dual pair

$$\text{find } \overline{x} \in \mathcal{H} \text{ such that } z \in \sum_{i=1}^m \omega_i L_i^*\big(B_i(L_i\overline{x} - r_i)\big), \tag{1.7}$$

and

$$\text{find } \overline{v}_1 \in \mathcal{G}_1,\ldots,\overline{v}_m \in \mathcal{G}_m \text{ such that } \begin{cases} \sum_{i=1}^m \omega_i L_i^*\overline{v}_i = z, \\ (\exists\, x \in \mathcal{H})\big(\forall i \in \{1,\ldots,m\}\big)\ \overline{v}_i \in B_i(L_i x - r_i). \end{cases} \tag{1.8}$$

This framework is considered in [7], where further special cases will be found. In particular, it contains the classical Fenchel-Rockafellar [27] and Mosco [24] duality settings, as well as that of [3].

The paper is organized as follows. Section 2 is devoted to notation and background. In Section 3, we present our algorithm, prove its convergence, and compare it to existing work. Applications to minimization problems are provided in Section 4, where further connections with the state-of-the-art are made.

2 Notation and background

We recall some notation and background from convex analysis and monotone operator theory (see [6] for a detailed account).

Throughout, $\mathcal{H}$, $\mathcal{G}$, and $(\mathcal{G}_i)_{1\leq i\leq m}$ are real Hilbert spaces. The scalars product and the associated norms of both $\mathcal{H}$ and $\mathcal{G}$ are denoted respectively by $\langle\cdot\mid\cdot\rangle$ and $\|\cdot\|$. For every $i\in\{1,\ldots,m\}$, the scalar product and associated norm of $\mathcal{G}_i$ are denoted respectively by $\langle\cdot\mid\cdot\rangle_{\mathcal{G}_i}$ and $\|\cdot\|_{\mathcal{G}_i}$. We denote by $\mathcal{B}(\mathcal{H},\mathcal{G})$ the space of all bounded linear operators from $\mathcal{H}$ to $\mathcal{G}$. The symbols $\rightharpoonup$ and $\to$ denote respectively weak and strong convergence. Let $A\colon\mathcal{H}\to 2^{\mathcal{H}}$ be a set-valued operator. The domain and the graph of A are respectively defined by $\operatorname{dom}A=\big\{x\in\mathcal{H}\mid Ax\neq\varnothing\big\}$ and $\operatorname{gra}A=\big\{(x,u)\in\mathcal{H}\times\mathcal{H}\mid u\in Ax\big\}$. We denote by $\operatorname{zer}A=\big\{x\in\mathcal{H}\mid 0\in Ax\big\}$ the set of zeros of A, and by $\operatorname{ran}A=\big\{u\in\mathcal{H}\mid(\exists\,x\in\mathcal{H})\;u\in Ax\big\}$ the range of A. The inverse of A is $A^{-1}\colon\mathcal{H}\mapsto 2^{\mathcal{H}}\colon u\mapsto\big\{x\in\mathcal{H}\mid u\in Ax\big\}$. The resolvent of A is

$$J_A=(\mathrm{Id}+A)^{-1},\tag{2.1}$$

where Id denotes the identity operator on $\mathcal{H}$. Moreover, A is monotone if

$$(\forall(x,y)\in\mathcal{H}\times\mathcal{H})\;(\forall(u,v)\in Ax\times Ay)\quad\langle x-y\mid u-v\rangle\geq 0,\tag{2.2}$$

and maximally monotone if it is monotone and there exists no monotone operator $B\colon\mathcal{H}\to 2^{\mathcal{H}}$ such that $\operatorname{gra}B$ properly contains $\operatorname{gra}A$. We say that A is uniformly monotone at $x\in\operatorname{dom}A$ if there exists an increasing function $\phi\colon\ [0,+\infty[\,\to[0,+\infty]$ vanishing only at 0 such that

$$\big(\forall u\in Ax\big)\big(\forall(y,v)\in\operatorname{gra}A\big)\quad\langle x-y\mid u-v\rangle\geq\phi(\|x-y\|).\tag{2.3}$$

If $A-\alpha\,\mathrm{Id}$ is monotone for some $\alpha\in\,]0,+\infty[$, then A is said to be α-strongly monotone. The parallel sum of two set-valued operators A and B from $\mathcal{H}$ to $2^{\mathcal{H}}$ is

$$A\,\Box\,B=(A^{-1}+B^{-1})^{-1}.\tag{2.4}$$

The class of all lower semicontinuous convex functions $f\colon\mathcal{H}\to\,]-\infty,+\infty]$ such that $\operatorname{dom}f=\big\{x\in\mathcal{H}\mid f(x)<+\infty\big\}\neq\varnothing$ is denoted by $\Gamma_0(\mathcal{H})$. Now, let $f\in\Gamma_0(\mathcal{H})$. The conjugate of f is the function $f^*\in\Gamma_0(\mathcal{H})$ defined by $f^*\colon u\mapsto\sup_{x\in\mathcal{H}}(\langle x\mid u\rangle-f(x))$, and the subdifferential of $f\in\Gamma_0(\mathcal{H})$ is the maximally monotone operator

$$\partial f\colon\mathcal{H}\to 2^{\mathcal{H}}\colon x\mapsto\big\{u\in\mathcal{H}\mid(\forall y\in\mathcal{H})\quad\langle y-x\mid u\rangle+f(x)\leq f(y)\big\}\tag{2.5}$$

with inverse given by

$$(\partial f)^{-1} = \partial f^*. \tag{2.6}$$

Moreover, the proximity operator of f is

$$\operatorname{prox}_f \colon \mathcal{H} \to \mathcal{H} \colon x \mapsto \underset{y\in\mathcal{H}}{\operatorname{argmin}}\, f(y) + \frac{1}{2}\|x-y\|^2. \tag{2.7}$$

We have

$$J_{\partial f} = \operatorname{prox}_f. \tag{2.8}$$

The infimal convolution of two functions f and g from $\mathcal{H}$ to $]-\infty,+\infty]$ is

$$f \,\square\, g \colon \mathcal{H} \to \,]-\infty,+\infty] : x \mapsto \inf_{y\in\mathcal{H}} (f(x) + g(x-y)). \tag{2.9}$$

Finally, let S be a convex subset of $\mathcal{H}$. The relative interior of S, i.e., the set of points $x \in S$ such that the cone generated by $x + S$ is a vector subspace of $\mathcal{H}$, is denoted by $\operatorname{ri} S$.

3 Algorithm and convergence

Our main result is the following theorem, in which we introduce our splitting algorithm and prove its convergence.

Theorem 3.1 *In Problem 1.1, suppose that*

$$z \in \operatorname{ran}\Bigg(A + \sum_{i=1}^m \omega_i L_i^*\big((B_i \,\square\, D_i)(L_i \cdot - r_i)\big) + C\Bigg). \tag{3.1}$$

Let τ and $(\sigma_i)_{1\leq i\leq m}$ be strictly positive numbers such that

$$2\rho \min\{\mu, \nu_1, \ldots, \nu_m\} > 1, \textit{ where } \rho = \min\Big\{\tau^{-1}, \sigma_1^{-1}, \ldots, \sigma_m^{-1}\Big\}\Bigg(1 - \sqrt{\tau \sum_{i=1}^m \sigma_i\omega_i\|L_i\|^2}\Bigg). \tag{3.2}$$

Let $\varepsilon \in \,]0,1[$, let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence in $[\varepsilon,1]$, let $x_0 \in \mathcal{H}$, let $(a_{1,n})_{n\in\mathbb{N}}$ and $(a_{2,n})_{n\in\mathbb{N}}$ be absolutely summable sequences in $\mathcal{H}$. For every $i \in \{1,\ldots,m\}$, let $v_{i,0} \in \mathcal{G}_i$, and let $(b_{i,n})_{n\in\mathbb{N}}$ and $(c_{i,n})_{n\in\mathbb{N}}$ be absolutely summable sequences in $\mathcal{G}_i$. Let $(x_n)_{n\in\mathbb{N}}$ and $(v_{1,n},\ldots,v_{m,n})_{n\in\mathbb{N}}$ be sequences generated by the following routine

$$(\forall n\in\mathbb{N}) \quad \left\lfloor \begin{array}{l} p_n = J_{\tau A}\Big(x_n - \tau\Big(\sum_{i=1}^m \omega_i L_i^* v_{i,n} + Cx_n + a_{1,n} - z\Big)\Big) + a_{2,n} \\ y_n = 2p_n - x_n \\ x_{n+1} = x_n + \lambda_n(p_n - x_n) \\ \text{for } i = 1,\ldots,m \\ \left\lfloor \begin{array}{l} q_{i,n} = J_{\sigma_i B_i^{-1}}\Big(v_{i,n} + \sigma_i\Big(L_i y_n - D_i^{-1} v_{i,n} - c_{i,n} - r_i\Big)\Big) + b_{i,n} \\ v_{i,n+1} = v_{i,n} + \lambda_n(q_{i,n} - v_{i,n}). \end{array}\right. \end{array}\right. \tag{3.3}$$

Then the following hold for some $\overline{x} \in \mathcal{P}$ and $(\overline{v}_1,\ldots,\overline{v}_m) \in \mathcal{D}$.

(i) *$x_n \rightharpoonup \overline{x}$ and $(v_{1,n},\ldots,v_{m,n}) \rightharpoonup (\overline{v}_1,\ldots,\overline{v}_m)$.*

(ii) *Suppose that C is uniformly monotone at $\overline{x}$. Then $x_n \to \overline{x}$.*

(iii) *Suppose that D_j^{-1} is uniformly monotone at $\overline{v}_j$ for some $j \in \{1,\ldots,m\}$. Then $v_{j,n} \to \overline{v}_j$.*

Proof. We define $\boldsymbol{\mathcal{G}}$ as the real Hilbert space obtained by endowing the Cartesian product $\mathcal{G}_1 \times \ldots \times \mathcal{G}_m$ with the scalar product and the associated norm respectively defined by

$$\langle \cdot \mid \cdot \rangle_{\boldsymbol{\mathcal{G}}} : (\boldsymbol{v},\boldsymbol{w}) \mapsto \sum_{i=1}^m \omega_i \langle v_i \mid w_i \rangle_{\mathcal{G}_i} \quad \text{and} \quad \|\cdot\|_{\boldsymbol{\mathcal{G}}} \colon \boldsymbol{v} \mapsto \sqrt{\sum_{i=1}^m \omega_i \|v_i\|_{\mathcal{G}_i}^2}, \tag{3.4}$$

where $\boldsymbol{v} = (v_1,\ldots,v_m)$ and $\boldsymbol{w} = (w_1,\ldots,w_m)$ denote generic elements in $\boldsymbol{\mathcal{G}}$. Next, we let $\boldsymbol{\mathcal{K}}$ be the Hilbert direct sum

$$\boldsymbol{\mathcal{K}} = \mathcal{H} \oplus \boldsymbol{\mathcal{G}}. \tag{3.5}$$

Thus, the scalar product and the norm of $\boldsymbol{\mathcal{K}}$ are respectively defined by

$$\langle \cdot \mid \cdot \rangle_{\boldsymbol{\mathcal{K}}} : \big((x,\boldsymbol{v}),(y,\boldsymbol{w})\big) \mapsto \langle x \mid y \rangle + \langle \boldsymbol{v} \mid \boldsymbol{w} \rangle_{\boldsymbol{\mathcal{G}}} \quad \text{and} \quad \|\cdot\|_{\boldsymbol{\mathcal{K}}} \colon (x,\boldsymbol{v}) \mapsto \sqrt{\|x\|^2 + \|\boldsymbol{v}\|_{\boldsymbol{\mathcal{G}}}^2}. \tag{3.6}$$

Let us set

$$\begin{aligned} \boldsymbol{M} \colon \boldsymbol{\mathcal{K}} &\to 2^{\boldsymbol{\mathcal{K}}} \\ (x,v_1,\ldots,v_m) &\mapsto \big(-z + Ax\big) \times \big(r_1 + B_1^{-1}v_1\big) \times \ldots \times \big(r_m + B_m^{-1}v_m\big). \end{aligned} \tag{3.7}$$

Since the operators A and $(B_i)_{1\leq i\leq m}$ are maximally monotone, $\boldsymbol{M}$ is maximally monotone [6, Propositions 20.22 and 20.23]. We also introduce

$$\boldsymbol{S} \colon \boldsymbol{\mathcal{K}} \to \boldsymbol{\mathcal{K}} \tag{3.8}$$

$$(x,v_1,\ldots,v_m) \mapsto \Bigg(\sum_{i=1}^m \omega_i L_i^* v_i, -L_1 x, \ldots, -L_m x\Bigg). \tag{3.9}$$

Note that $\boldsymbol{S}$ is linear, bounded, and skew (i.e, $\boldsymbol{S}^* = -\boldsymbol{S}$). Hence, $\boldsymbol{S}$ is maximally monotone [6, Example 20.30]. Moreover, since $\operatorname{dom} \boldsymbol{S} = \boldsymbol{\mathcal{K}}$, $\boldsymbol{M} + \boldsymbol{S}$ is maximally monotone [6, Corollary 24.24(i)]. Since, for every $i \in \{1,\ldots,m\}$, D_i is ν_i-strongly monotone, D_i^{-1} is ν_i-cocoercive. Let us prove that

$$\begin{aligned} \boldsymbol{Q} \colon \boldsymbol{\mathcal{K}} &\to \boldsymbol{\mathcal{K}} \\ (x,v_1,\ldots,v_m) &\mapsto \big(Cx, D_1^{-1}v_1, \ldots, D_m^{-1}v_m\big) \end{aligned} \tag{3.10}$$

is β-cocoercive with

$$\beta = \min\{\mu, \nu_1, \ldots, \nu_m\}. \tag{3.11}$$

For every $(x,v_1,\dots,v_m)$ and every $(y,w_1,\dots,w_m)$ in $\boldsymbol{\mathcal{K}}$, we have

$$\begin{aligned}
&\langle (x,v_1,\dots,v_m)-(y,w_1,\dots,w_m) \mid \boldsymbol{Q}(x,v_1,\dots,v_m)-\boldsymbol{Q}(y,w_1,\dots,w_m)\rangle_{\boldsymbol{\mathcal{K}}}\\
&= \langle x-y \mid Cx-Cy\rangle + \sum_{i=1}^m \omega_i \left\langle v_i-w_i \mid D_i^{-1}v_i - D_i^{-1}w_i\right\rangle_{\mathcal{G}_i}\\
&\geq \mu\|Cx-Cy\|^2 + \sum_{i=1}^m \nu_i\omega_i\|D_i^{-1}v_i-D_i^{-1}w_i\|^2_{\mathcal{G}_i}\\
&\geq \beta\bigg(\|Cx-Cy\|^2 + \sum_{i=1}^m \omega_i\|D_i^{-1}v_i-D_i^{-1}w_i\|^2_{\mathcal{G}_i}\bigg)\\
&= \beta\|\boldsymbol{Q}(x,v_1,\dots,v_m)-\boldsymbol{Q}(y,w_1,\dots,w_m)\|^2_{\boldsymbol{\mathcal{K}}}. \qquad (3.12)
\end{aligned}$$

Therefore, by (1.1), $\boldsymbol{Q}$ is β-cocoercive. It is shown in [16, Eq. (3.12)] that under the condition (3.1), $\operatorname{zer}(\boldsymbol{M}+\boldsymbol{S}+\boldsymbol{Q})\neq\varnothing$. Moreover, [16, Eq. (3.21)] and [16, Eq. (3.22)] yield

$$(\overline{x},\overline{\boldsymbol{v}})\in\operatorname{zer}(\boldsymbol{M}+\boldsymbol{S}+\boldsymbol{Q}) \Rightarrow \overline{x}\in\mathcal{P} \quad\text{and}\quad \overline{\boldsymbol{v}}\in\mathcal{D}. \qquad (3.13)$$

Now, define

$$\begin{aligned}
&\boldsymbol{V}\colon \boldsymbol{\mathcal{K}}\to\boldsymbol{\mathcal{K}}\\
&(x,v_1,\dots,v_m)\mapsto \bigg(\tau^{-1}x-\sum_{i=1}^m\omega_iL_i^*v_i, \sigma_1^{-1}v_1-L_1x,\dots,\sigma_m^{-1}v_m-L_mx\bigg). \qquad (3.14)
\end{aligned}$$

Then $\boldsymbol{V}$ is self-adjoint. Let us check that $\boldsymbol{V}$ is ρ-strongly positive. To this end, define

$$\boldsymbol{T}\colon \mathcal{H}\to\boldsymbol{\mathcal{G}}\colon x\mapsto \Big(\sqrt{\sigma_1}L_1x,\dots,\sqrt{\sigma_m}L_mx\Big). \qquad (3.15)$$

Then,

$$(\forall x\in\mathcal{H})\quad \|\boldsymbol{T}x\|^2_{\boldsymbol{\mathcal{G}}} = \sum_{i=1}^m\omega_i\sigma_i\|L_ix\|^2_{\mathcal{G}_i} \leq \|x\|^2\sum_{i=1}^m\omega_i\sigma_i\|L_i\|^2, \qquad (3.16)$$

which implies that

$$\|\boldsymbol{T}\|^2\leq\sum_{i=1}^m\omega_i\sigma_i\|L_i\|^2. \qquad (3.17)$$

Now set

$$\delta=\left(\sqrt{\tau\sum_{i=1}^m\sigma_i\omega_i\|L_i\|^2}\right)^{-1}-1. \qquad (3.18)$$

Then, it follows from (3.2) that $\delta>0$. Moreover, (3.17) and (3.18) yield

$$\tau\|\boldsymbol{T}\|^2(1+\delta)\leq\tau(1+\delta)\sum_{i=1}^m\omega_i\sigma_i\|L_i\|^2=(1+\delta)^{-1}. \qquad (3.19)$$

For every $\boldsymbol{x} = (x, v_1, \ldots, v_m)$ in $\boldsymbol{\mathcal{K}}$, by using (3.19), we obtain

$$
\begin{aligned}
\langle \boldsymbol{x} \mid \boldsymbol{V}\boldsymbol{x}\rangle_{\boldsymbol{\mathcal{K}}} &= \tau^{-1}\|x\|^2 + \sum_{i=1}^m \sigma_i^{-1}\omega_i\|v_i\|^2_{\mathcal{G}_i} - 2\sum_{i=1}^m \omega_i \left\langle L_i x \mid v_i\right\rangle_{\mathcal{G}_i} \\
&= \tau^{-1}\|x\|^2 + \sum_{i=1}^m \sigma_i^{-1}\omega_i\|v_i\|^2_{\mathcal{G}_i} - 2\sum_{i=1}^m \omega_i \left\langle \sqrt{\sigma_i} L_i x \mid \sqrt{\sigma_i}^{-1} v_i\right\rangle_{\mathcal{G}_i} \\
&= \tau^{-1}\|x\|^2 + \sum_{i=1}^m \sigma_i^{-1}\omega_i\|v_i\|^2_{\mathcal{G}_i} - 2\left\langle \boldsymbol{T}x \mid (\sqrt{\sigma_1}^{-1}v_1, \ldots, \sqrt{\sigma_m}^{-1}v_m)\right\rangle_{\boldsymbol{\mathcal{G}}} \\
&\geq \tau^{-1}\|x\|^2 + \sum_{i=1}^m \sigma_i^{-1}\omega_i\|v_i\|^2_{\mathcal{G}_i} - \left(\frac{\|\boldsymbol{T}x\|^2_{\boldsymbol{\mathcal{G}}}}{\tau(1+\delta)\|\boldsymbol{T}\|^2} + \tau(1+\delta)\|\boldsymbol{T}\|^2\sum_{i=1}^m \sigma_i^{-1}\omega_i\|v_i\|^2_{\mathcal{G}_i}\right) \\
&\geq \left(1-(1+\delta)^{-1}\right)\left(\tau^{-1}\|x\|^2 + \sum_{i=1}^m \sigma_i^{-1}\omega_i\|v_i\|^2_{\mathcal{G}_i}\right) \\
&\geq \left(1-(1+\delta)^{-1}\right)\min\{\tau^{-1}, \sigma_1^{-1}, \ldots, \sigma_m^{-1}\}\|\boldsymbol{x}\|^2_{\boldsymbol{\mathcal{K}}} \\
&= \rho\|\boldsymbol{x}\|^2_{\boldsymbol{\mathcal{K}}}.
\end{aligned} \tag{3.20}
$$

Therefore, $\boldsymbol{V}$ is ρ-strongly positive. Furthermore, it follows from (3.20) that

$$
\boldsymbol{V}^{-1} \text{ exists and } \|\boldsymbol{V}^{-1}\| \leq \rho^{-1}. \tag{3.21}
$$

(i): We first observe that (3.3) is equivalent to

$$
(\forall n \in \mathbb{N}) \quad
\left\lfloor
\begin{array}{l}
\tau^{-1}(x_n - p_n) - \sum_{i=1}^m \omega_i L_i^* v_{i,n} - Cx_n \in \\
\qquad\qquad -z + A(p_n - a_{2,n}) + a_{1,n} - \tau^{-1}a_{2,n} \\
x_{n+1} = x_n + \lambda_n(p_n - x_n) \\
\text{for } i = 1, \ldots, m \\
\left\lfloor
\begin{array}{l}
\sigma_i^{-1}(v_{i,n} - q_{i,n}) - L_i(x_n - p_n) - D_i^{-1}v_{i,n} \in \\
\qquad\qquad r_i + B_i^{-1}(q_{i,n} - b_{i,n}) - L_i p_n + c_{i,n} - \sigma_i^{-1}b_{i,n} \\
v_{i,n+1} = v_{i,n} + \lambda_n(q_{i,n} - v_{i,n}).
\end{array}
\right.
\end{array}
\right. \tag{3.22}
$$

Now set

$$
(\forall n \in \mathbb{N}) \quad
\begin{cases}
\boldsymbol{x}_n = (x_n, v_{1,n}, \ldots, v_{m,n}) \\
\boldsymbol{y}_n = (p_n, q_{1,n}, \ldots, q_{m,n}) \\
\boldsymbol{a}_n = (a_{2,n}, b_{1,n}, \ldots, b_{m,n}) \\
\boldsymbol{c}_n = (a_{1,n}, c_{1,n}, \ldots, c_{m,n}) \\
\boldsymbol{d}_n = (\tau^{-1}a_{2,n}, \sigma_1^{-1}b_{1,n}, \ldots, \sigma_m^{-1}b_{m,n}).
\end{cases} \tag{3.23}
$$

We have

$$
\sum_{n\in\mathbb{N}} \|\boldsymbol{a}_n\|_{\boldsymbol{\mathcal{K}}} < +\infty, \quad \sum_{n\in\mathbb{N}} \|\boldsymbol{c}_n\|_{\boldsymbol{\mathcal{K}}} < +\infty, \quad \text{and} \quad \sum_{n\in\mathbb{N}} \|\boldsymbol{d}_n\|_{\boldsymbol{\mathcal{K}}} < +\infty. \tag{3.24}
$$

Furthermore, (3.22) yields

$$
(\forall n \in \mathbb{N}) \quad
\left\lfloor
\begin{array}{l}
\boldsymbol{V}(\boldsymbol{x}_n - \boldsymbol{y}_n) - \boldsymbol{Q}\boldsymbol{x}_n \in (\boldsymbol{M}+\boldsymbol{S})(\boldsymbol{y}_n - \boldsymbol{a}_n) + \boldsymbol{S}\boldsymbol{a}_n + \boldsymbol{c}_n - \boldsymbol{d}_n \\
\boldsymbol{x}_{n+1} = \boldsymbol{x}_n + \lambda_n(\boldsymbol{y}_n - \boldsymbol{x}_n).
\end{array}
\right. \tag{3.25}
$$

Next, we set

$$(\forall n \in \mathbb{N}) \quad \boldsymbol{b}_n = \boldsymbol{V}^{-1}\big((\boldsymbol{S}+\boldsymbol{V})\boldsymbol{a}_n + \boldsymbol{c}_n - \boldsymbol{d}_n\big). \tag{3.26}$$

Then (3.24) implies that

$$\sum_{n\in\mathbb{N}} \|\boldsymbol{b}_n\|_{\boldsymbol{\mathcal{K}}} < +\infty. \tag{3.27}$$

Moreover, using (3.21) and (3.26), we have

$$\begin{aligned}
&(\forall n \in \mathbb{N}) \quad \boldsymbol{V}(\boldsymbol{x}_n - \boldsymbol{y}_n) - \boldsymbol{Q}\boldsymbol{x}_n \in (\boldsymbol{M}+\boldsymbol{S})(\boldsymbol{y}_n - \boldsymbol{a}_n) + \boldsymbol{S}\boldsymbol{a}_n + \boldsymbol{c}_n - \boldsymbol{d}_n\\
\Leftrightarrow\ &(\forall n \in \mathbb{N}) \quad (\boldsymbol{V}-\boldsymbol{Q})\boldsymbol{x}_n \in (\boldsymbol{M}+\boldsymbol{S}+\boldsymbol{V})(\boldsymbol{y}_n - \boldsymbol{a}_n) + (\boldsymbol{S}+\boldsymbol{V})\boldsymbol{a}_n + \boldsymbol{c}_n - \boldsymbol{d}_n\\
\Leftrightarrow\ &(\forall n \in \mathbb{N}) \quad \boldsymbol{y}_n = \big(\boldsymbol{M}+\boldsymbol{S}+\boldsymbol{V}\big)^{-1}\Big((\boldsymbol{V}-\boldsymbol{Q})\boldsymbol{x}_n - (\boldsymbol{S}+\boldsymbol{V})\boldsymbol{a}_n - \boldsymbol{c}_n + \boldsymbol{d}_n\Big) + \boldsymbol{a}_n\\
\Leftrightarrow\ &(\forall n \in \mathbb{N}) \quad \boldsymbol{y}_n = \Big(\mathbf{Id} + \boldsymbol{V}^{-1}(\boldsymbol{M}+\boldsymbol{S})\Big)^{-1}\Big((\mathbf{Id} - \boldsymbol{V}^{-1}\boldsymbol{Q})\boldsymbol{x}_n - \boldsymbol{b}_n\Big) + \boldsymbol{a}_n.
\end{aligned} \tag{3.28}$$

We derive from (3.25) that

$$\begin{aligned}
(\forall n \in \mathbb{N}) \quad \boldsymbol{x}_{n+1} &= \boldsymbol{x}_n + \lambda_n\Big(\big(\mathbf{Id} + \boldsymbol{V}^{-1}(\boldsymbol{M}+\boldsymbol{S})\big)^{-1}\big(\boldsymbol{x}_n - \boldsymbol{V}^{-1}\boldsymbol{Q}\boldsymbol{x}_n - \boldsymbol{b}_n\big) + \boldsymbol{a}_n - \boldsymbol{x}_n\Big)\\
&= \boldsymbol{x}_n + \lambda_n\Big(J_{\boldsymbol{A}}\big(\boldsymbol{x}_n - \boldsymbol{B}\boldsymbol{x}_n - \boldsymbol{b}_n\big) + \boldsymbol{a}_n - \boldsymbol{x}_n\Big),
\end{aligned} \tag{3.29}$$

where

$$\boldsymbol{A} = \boldsymbol{V}^{-1}(\boldsymbol{M}+\boldsymbol{S}) \quad \text{and} \quad \boldsymbol{B} = \boldsymbol{V}^{-1}\boldsymbol{Q}. \tag{3.30}$$

Algorithm (3.29) has the structure of the forward-backward splitting algorithm [13]. Hence, it is sufficient to check the convergence conditions of the forward-backward splitting algorithm [13, Corollary 6.5] to prove our claims. To this end, let us introduce the real Hilbert space $\boldsymbol{\mathcal{K}}_{\boldsymbol{V}}$ with scalar product and norm defined by

$$\big(\forall(\boldsymbol{x},\boldsymbol{y}) \in \boldsymbol{\mathcal{K}}\times\boldsymbol{\mathcal{K}}\big) \quad \langle \boldsymbol{x} \mid \boldsymbol{y}\rangle_{\boldsymbol{V}} = \langle \boldsymbol{x} \mid \boldsymbol{V}\boldsymbol{y}\rangle_{\boldsymbol{\mathcal{K}}} \quad \text{and} \quad \|\boldsymbol{x}\|_{\boldsymbol{V}} = \sqrt{\langle \boldsymbol{x} \mid \boldsymbol{V}\boldsymbol{x}\rangle_{\boldsymbol{\mathcal{K}}}}, \tag{3.31}$$

respectively. Since $\boldsymbol{V}$ is a bounded linear operator, it follows from (3.24) and (3.27) that

$$\sum_{n\in\mathbb{N}} \|\boldsymbol{a}_n\|_{\boldsymbol{V}} < +\infty \quad \text{and} \quad \sum_{n\in\mathbb{N}} \|\boldsymbol{b}_n\|_{\boldsymbol{V}} < +\infty. \tag{3.32}$$

Moreover, since $\boldsymbol{M}+\boldsymbol{S}$ is monotone on $\boldsymbol{\mathcal{K}}$, we have

$$\begin{aligned}
\big(\forall(\boldsymbol{x},\boldsymbol{y}) \in \boldsymbol{\mathcal{K}}\times\boldsymbol{\mathcal{K}}\big) \quad \langle \boldsymbol{x}-\boldsymbol{y} \mid \boldsymbol{A}\boldsymbol{x} - \boldsymbol{A}\boldsymbol{y}\rangle_{\boldsymbol{V}} &= \langle \boldsymbol{x}-\boldsymbol{y} \mid \boldsymbol{V}\boldsymbol{A}\boldsymbol{x} - \boldsymbol{V}\boldsymbol{A}\boldsymbol{y}\rangle_{\boldsymbol{\mathcal{K}}}\\
&= \langle \boldsymbol{x}-\boldsymbol{y} \mid (\boldsymbol{M}+\boldsymbol{S})\boldsymbol{x} - (\boldsymbol{M}+\boldsymbol{S})\boldsymbol{y}\rangle_{\boldsymbol{\mathcal{K}}} &(3.33)\\
&\geq 0. &(3.34)
\end{aligned}$$

Hence, $\boldsymbol{A}$ is monotone on $\boldsymbol{\mathcal{K}}_{\boldsymbol{V}}$. Likewise, $\boldsymbol{B}$ is monotone on $\boldsymbol{\mathcal{K}}_{\boldsymbol{V}}$. Since $\boldsymbol{V}$ is strongly positive, and since $\boldsymbol{M}+\boldsymbol{S}$ is maximally monotone on $\boldsymbol{\mathcal{K}}$, $\boldsymbol{A}$ is maximally monotone on $\boldsymbol{\mathcal{K}}_{\boldsymbol{V}}$. Next, let us show that $\boldsymbol{B}$ is $(\beta\rho)$-cocoercive on $\boldsymbol{\mathcal{K}}_{\boldsymbol{V}}$. Using (3.12), (3.20) and (3.21), we have

$$
\begin{aligned}
\big(\forall(\boldsymbol{x},\boldsymbol{y})\in\boldsymbol{\mathcal{K}}_{\boldsymbol{V}}\times\boldsymbol{\mathcal{K}}_{\boldsymbol{V}}\big)\quad \langle\boldsymbol{x}-\boldsymbol{y}\mid\boldsymbol{B}\boldsymbol{x}-\boldsymbol{B}\boldsymbol{y}\rangle_{\boldsymbol{V}} &= \langle\boldsymbol{x}-\boldsymbol{y}\mid\boldsymbol{V}\boldsymbol{B}\boldsymbol{x}-\boldsymbol{V}\boldsymbol{B}\boldsymbol{y}\rangle_{\boldsymbol{\mathcal{K}}}\\
&= \langle\boldsymbol{x}-\boldsymbol{y}\mid\boldsymbol{Q}\boldsymbol{x}-\boldsymbol{Q}\boldsymbol{y}\rangle_{\boldsymbol{\mathcal{K}}}\\
&\geq \beta\|\boldsymbol{Q}\boldsymbol{x}-\boldsymbol{Q}\boldsymbol{y}\|^2_{\boldsymbol{\mathcal{K}}}\\
&= \beta\|\boldsymbol{Q}\boldsymbol{x}-\boldsymbol{Q}\boldsymbol{y}\|_{\boldsymbol{\mathcal{K}}}\|\boldsymbol{Q}\boldsymbol{x}-\boldsymbol{Q}\boldsymbol{y}\|_{\boldsymbol{\mathcal{K}}}\\
&= \beta\|\boldsymbol{V}^{-1}\|^{-1}\|\boldsymbol{V}^{-1}\|\|\boldsymbol{Q}\boldsymbol{x}-\boldsymbol{Q}\boldsymbol{y}\|_{\boldsymbol{\mathcal{K}}}\|\boldsymbol{Q}\boldsymbol{x}-\boldsymbol{Q}\boldsymbol{y}\|_{\boldsymbol{\mathcal{K}}}\\
&\geq \beta\|\boldsymbol{V}^{-1}\|^{-1}\|\boldsymbol{V}^{-1}\boldsymbol{Q}\boldsymbol{x}-\boldsymbol{V}^{-1}\boldsymbol{Q}\boldsymbol{y}\|_{\boldsymbol{\mathcal{K}}}\|\boldsymbol{Q}\boldsymbol{x}-\boldsymbol{Q}\boldsymbol{y}\|_{\boldsymbol{\mathcal{K}}}\\
&\geq \beta\|\boldsymbol{V}^{-1}\|^{-1}\big\langle\boldsymbol{V}^{-1}\boldsymbol{Q}\boldsymbol{x}-\boldsymbol{V}^{-1}\boldsymbol{Q}\boldsymbol{y}\mid\boldsymbol{Q}\boldsymbol{x}-\boldsymbol{Q}\boldsymbol{y}\big\rangle_{\boldsymbol{\mathcal{K}}}\\
&= \beta\|\boldsymbol{V}^{-1}\|^{-1}\langle\boldsymbol{B}\boldsymbol{x}-\boldsymbol{B}\boldsymbol{y}\mid\boldsymbol{Q}\boldsymbol{x}-\boldsymbol{Q}\boldsymbol{y}\rangle_{\boldsymbol{\mathcal{K}}}\\
&= \beta\|\boldsymbol{V}^{-1}\|^{-1}\|\boldsymbol{B}\boldsymbol{x}-\boldsymbol{B}\boldsymbol{y}\|^2_{\boldsymbol{V}}\\
&\geq \beta\rho\|\boldsymbol{B}\boldsymbol{x}-\boldsymbol{B}\boldsymbol{y}\|^2_{\boldsymbol{V}}.
\end{aligned}
\tag{3.35}
$$

Hence, by (1.1), $\boldsymbol{B}$ is $(\beta\rho)$-cocoercive on $\boldsymbol{\mathcal{K}}_{\boldsymbol{V}}$. Moreover, it follows from our assumption that $2\beta\rho > 1$. Altogether, by [13, Corollary 6.5] the sequence $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ converges weakly in $\boldsymbol{\mathcal{K}}_{\boldsymbol{V}}$ to some $\overline{\boldsymbol{x}} = (\overline{x},\overline{v}_1,\ldots,\overline{v}_m)\in\operatorname{zer}(\boldsymbol{A}+\boldsymbol{B}) = \operatorname{zer}(\boldsymbol{M}+\boldsymbol{S}+\boldsymbol{Q})$. Since $\boldsymbol{V}$ is self-adjoint and $\boldsymbol{V}^{-1}$ exists, the weak convergence of the sequence $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ to $\overline{\boldsymbol{x}}$ in $\boldsymbol{\mathcal{K}}_{\boldsymbol{V}}$ is equivalent to the weak convergence of $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ to $\overline{\boldsymbol{x}}$ in $\boldsymbol{\mathcal{K}}$. Hence, $\boldsymbol{x}_n \rightharpoonup \overline{\boldsymbol{x}}\in\operatorname{zer}(\boldsymbol{M}+\boldsymbol{S}+\boldsymbol{Q})$. It follows from (3.13) that $\overline{x}\in\mathcal{P}$ and $(\overline{v}_1,\ldots,\overline{v}_m)\in\mathcal{D}$. This proves (i).

(ii)&(iii): It follows from [13, Remark 3.4] that

$$\sum_{n\in\mathbb{N}}\|\boldsymbol{B}\boldsymbol{x}_n-\boldsymbol{B}\overline{\boldsymbol{x}}\|^2_{\boldsymbol{V}} < +\infty. \tag{3.36}$$

On the other hand, from (3.20) and (3.36) yield $\boldsymbol{B}\boldsymbol{x}_n-\boldsymbol{B}\overline{\boldsymbol{x}} = \boldsymbol{V}^{-1}(\boldsymbol{Q}\boldsymbol{x}_n-\boldsymbol{Q}\overline{\boldsymbol{x}})\to 0$, which implies that $\boldsymbol{Q}\boldsymbol{x}_n-\boldsymbol{Q}\overline{\boldsymbol{x}}\to 0$. Hence,

$$Cx_n\to C\overline{x}\quad\text{and}\quad\big(\forall i\in\{1,\ldots,m\}\big)\quad D_i^{-1}v_{i,n}\to D_i^{-1}\overline{v}_i. \tag{3.37}$$

If C is uniformly monotone at $\overline{x}$, then there exists an increasing function $\phi_C\colon\ [0,+\infty[\,\to[0,+\infty]$ vanishing only at 0 such that

$$\phi_C(\|x_n-\overline{x}\|)\leq\langle x_n-\overline{x}\mid Cx_n-C\overline{x}\rangle\leq\|x_n-\overline{x}\|\,\|Cx_n-C\overline{x}\|. \tag{3.38}$$

Notice that $(x_n-\overline{x})_{n\in\mathbb{N}}$ is bounded. It follows from (3.37) and (3.38) that $x_n\to\overline{x}$. This proves (ii), and (iii) is proved in a similar fashion. ☐

Remark 3.2 Here are some remarks concerning the connections between our framework and existing work.

(i) The strategy used in the proof of Theorem 3.1(i) is to reformulate algorithm (3.3) as a forward-backward splitting algorithm in a real Hilbert space endowed with a suitable norm. This renorming technique was used in [22] for a minimization problem in finite-dimensional spaces. The same technique is also used in the primal-dual minimization problem of [18].

(ii) Consider the special case when $z = 0$, and $(B_i)_{1\leq i\leq m}$ and $(D_i)_{1\leq i\leq m}$ are as in (1.4). Then algorithm (3.3) reduces to

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = x_n + \lambda_n\Big(J_{\tau A}\big(x_n - \tau(Cx_n + a_{1,n})\big) + a_{2,n} - x_n\Big), \tag{3.39}$$

which is the standard forward-backward splitting algorithm [13, Algorithm 6.4] where the sequence $(\gamma_n)_{n\in\mathbb{N}}$ in [13, Eq. (6.3)] is constant.

(iii) The inclusions (1.7) and (1.8) in Example 1.3 can be solved by [7, Theorem 3.8]. However, the algorithm resulting from (3.3) in this special case is different from that of [7, Theorem 3.8].

(iv) In Problem 1.1, since C and $(D_i^{-1})_{1\leq i\leq m}$ are cocoercive, they are Lipschitzian. Hence, Problem 1.1 can be solved by the algorithm proposed in [16, Theorem 3.1], which has a different structure from the present algorithm.

(v) Consider the special case when $z = 0$ and $(\forall i \in \{1,\ldots,m\})$ $\mathcal{G}_i = \mathcal{H}, L_i = \mathrm{Id}, D_i^{-1} = 0, r_i = 0$. Then the primal inclusion (1.2) reduces to

$$\text{find } \overline{x} \in \mathcal{H} \text{ such that } 0 \in A\overline{x} + \sum_{i=1}^{m} \omega_i B_i\overline{x} + C\overline{x}. \tag{3.40}$$

This inclusion can be solved by the algorithm proposed in [26], which is not designed as a primal-dual scheme.

4 Application to minimization problems

We provide an application of the algorithm (3.3) to minimization problems, by revisiting [16, Problem 4.1].

Problem 4.1 Let $\mathcal{H}$ be a real Hilbert space, let $z \in \mathcal{H}$, let m be a strictly positive integer, let $(\omega_i)_{1\leq i\leq m}$ be real numbers in $]0,1]$ such that $\sum_{i=1}^m \omega_i = 1$, let $f \in \Gamma_0(\mathcal{H})$, and let $h\colon \mathcal{H} \to \mathbb{R}$ be convex and differentiable with a μ^{-1}-Lipschitzian gradient for some $\mu \in]0,+\infty[$. For every $i \in \{1,\ldots,m\}$, let $\mathcal{G}_i$ be a real Hilbert space, let $r_i \in \mathcal{G}_i$, let $g_i \in \Gamma_0(\mathcal{G}_i)$, let $\ell_i \in \Gamma_0(\mathcal{G}_i)$ be ν_i-strongly convex, for some $\nu_i \in]0,+\infty[$, and suppose that $L_i\colon \mathcal{H} \to \mathcal{G}_i$ is a nonzero bounded linear operator. Consider the primal problem

$$\underset{x\in\mathcal{H}}{\text{minimize}}\; f(x) + \sum_{i=1}^{m} \omega_i (g_i \,\Box\, \ell_i)(L_i x - r_i) + h(x) - \langle x \mid z\rangle\,, \tag{4.1}$$

and the dual problem

$$\underset{v_1\in\mathcal{G}_1,\ldots,v_m\in\mathcal{G}_m}{\text{minimize}}\; (f^* \,\Box\, h^*)\Big(z - \sum_{i=1}^{m} \omega_i L_i^* v_i\Big) + \sum_{i=1}^{m} \omega_i\big(g_i^*(v_i) + \ell_i^*(v_i) + \langle v_i \mid r_i\rangle_{\mathcal{G}_i}\big). \tag{4.2}$$

We denote by $\mathcal{P}_1$ and $\mathcal{D}_1$ the sets of solutions to (4.1) and (4.2), respectively.

Corollary 4.2 *In Problem 4.1, suppose that*

$$z \in \operatorname{ran}\left(\partial f + \sum_{i=1}^{m} \omega_i L_i^*\big((\partial g_i \,\square\, \partial \ell_i)(L_i \cdot - r_i)\big) + \nabla h\right). \tag{4.3}$$

Let τ and $(\sigma_i)_{1\leq i\leq m}$ be strictly positive numbers such that

$$2\rho \min\{\mu, \nu_1, \ldots, \nu_m\} > 1, \textit{where } \rho = \min\left\{\tau^{-1}, \sigma_1^{-1}, \ldots, \sigma_m^{-1}\right\}\left(1 - \sqrt{\tau \sum_{i=1}^{m} \sigma_i \omega_i \|L_i\|^2}\right). \tag{4.4}$$

Let $\varepsilon \in \left]0,1\right[$ and let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence in $[\varepsilon, 1]$, let $x_0 \in \mathcal{H}$, let $(a_{1,n})_{n\in\mathbb{N}}$ and $(a_{2,n})_{n\in\mathbb{N}}$ be absolutely summable sequences in $\mathcal{H}$. For every $i \in \{1,\ldots,m\}$, let $v_{i,0} \in \mathcal{G}_i$, and let $(b_{i,n})_{n\in\mathbb{N}}$ and $(c_{i,n})_{n\in\mathbb{N}}$ be absolutely summable sequences in $\mathcal{G}_i$. Let $(x_n)_{n\in\mathbb{N}}$ and $(v_{1,n},\ldots,v_{m,n})_{n\in\mathbb{N}}$ be sequences generated by the following routine

$$(\forall n \in \mathbb{N}) \quad \left\lfloor \begin{array}{l} p_n = \operatorname{prox}_{\tau f}\left(x_n - \tau\Big(\sum_{i=1}^{m} \omega_i L_i^* v_{i,n} + \nabla h(x_n) + a_{1,n} - z\Big)\right) + a_{2,n} \\ y_n = 2p_n - x_n \\ x_{n+1} = x_n + \lambda_n(p_n - x_n) \\ \text{for } i = 1, \ldots, m \\ \left\lfloor \begin{array}{l} q_{i,n} = \operatorname{prox}_{\sigma_i g_i^*}\left(v_{i,n} + \sigma_i\Big(L_i y_n - \nabla \ell_i^*(v_{i,n}) + c_{i,n} - r_i\Big)\right) + b_{i,n} \\ v_{i,n+1} = v_{i,n} + \lambda_n(q_{i,n} - v_{i,n}). \end{array}\right. \end{array}\right. \tag{4.5}$$

Then the following hold for some $\overline{x} \in \mathcal{P}_1$ and $(\overline{v}_1,\ldots,\overline{v}_m) \in \mathcal{D}_1$.

(i) $x_n \rightharpoonup \overline{x}$ *and* $(v_{1,n},\ldots,v_{m,n}) \rightharpoonup (\overline{v}_1,\ldots,\overline{v}_m)$.

(ii) *Suppose that h is uniformly convex at $\overline{x}$. Then $x_n \to \overline{x}$.*

(iii) *Suppose that ℓ_j^* is uniformly convex at $\overline{v}_j$ for some $j \in \{1,\ldots,m\}$. Then $v_{j,n} \to \overline{v}_j$.*

Proof. The connection between Problem 4.1 and Problem 1.1 is established in the proof of [16, Theorem 4.2]. Since ∇h is μ^{-1}-Lipschitz continuous, by the Baillon-Haddad Theorem [4, 5], it is μ-cocoercive. Moreover since, for every $i \in \{1,\ldots,m\}$, ℓ_i is ν_i-strongly convex, $\partial \ell_i$ is ν_i-strongly monotone. Hence, by applying Theorem 3.1(i) with $A = \partial f$, $J_{\tau A} = \operatorname{prox}_{\tau f}$, $C = \nabla h$ and for every $i \in \{1,\ldots,m\}$, $D_i^{-1} = \nabla \ell_i^*$, $B_i = \partial g_i$, $J_{\sigma_i B_i^{-1}} = \operatorname{prox}_{\sigma_i g_i^*}$, we obtain that the sequence $(x_n)_{n\in\mathbb{N}}$ converges weakly to some $\overline{x} \in \mathcal{H}$ such that

$$z \in \partial f(\overline{x}) + \sum_{i=1}^{m} \omega_i L_i^*\big((\partial g_i \,\square\, \partial \ell_i)(L_i \overline{x} - r_i)\big) + \nabla h(\overline{x}), \tag{4.6}$$

and the sequence $((v_{1,n},\ldots,v_{m,n}))_{n\in\mathbb{N}}$ converges weakly to some $(\overline{v}_1,\ldots,\overline{v}_m)$ such that

$$(\exists\, x \in \mathcal{H}) \quad \begin{cases} z - \sum_{i=1}^{m} \omega_i L_i^* \overline{v}_i \in \partial f(x) + \nabla h(x) \\ (\forall i \in \{1,\ldots,m\}) \quad \overline{v}_i \in (\partial g_i \,\square\, \partial \ell_i)(L_i x - r_i). \end{cases} \tag{4.7}$$

As shown in the proof of [16, Theorem 4.2], $\overline{x} \in \mathcal{P}_1$ and $(\overline{v}_1,\ldots,\overline{v}_m) \in \mathcal{D}_1$. This proves (i). Now, if h is uniformly convex at $\overline{x}$, then ∇h is uniformly monotone at $\overline{x}$. Hence, (ii) follows from Theorem 3.1(ii). Similarly, (iii) follows from Theorem 3.1(iii). ◻

Remark 4.3 Here are some observations on the above results.

(i) If a function $\varphi\colon \mathcal{H} \to \mathbb{R}$ is convex and differentiable function with a β^{-1}-Lipschitzian gradient, then $\nabla\varphi$ is β-cocoercive [4, 5]. Hence, in the context of convex minimization problems, the restriction of cocoercivity made in Problem 1.1 with respect to the problem considered in [16] disappears. Yet, the algorithm we obtain is quite different from that proposed in [16, Theorem 4.2].

(ii) Sufficient conditions which ensure that (4.3) is satisfied are provided in [16, Proposition 4.3]. For instance, if (4.1) has at least one solution, and if $\mathcal{H}$ and $(\mathcal{G}_i)_{1\leq i\leq m}$ are finite-dimensional, and there exists $x \in \operatorname{ri}\operatorname{dom} f$ such that

$$\big(\forall i \in \{1,\ldots,m\}\big) \quad L_i x - r_i \in \operatorname{ri}\operatorname{dom} g_i + \operatorname{ri}\operatorname{dom} \ell_i, \tag{4.8}$$

then (4.3) holds.

(iii) Consider the special case when $z = 0$ and, for every $i \in \{1,\ldots,m\}$, $r_i = 0, \sigma_i = \sigma \in \left]0,+\infty\right[$, and

$$\ell_i\colon v \mapsto \begin{cases} 0 & \text{if } v = 0,\\ +\infty & \text{otherwise.}\end{cases} \tag{4.9}$$

Then, (4.5) reduces to

$$(\forall n \in \mathbb{N}) \quad \left\lfloor \begin{array}{l} p_n = \operatorname{prox}_{\tau f}\Big(x_n - \tau\Big(\sum_{i=1}^m \omega_i L_i^* v_{i,n} + \nabla h(x_n) + a_{1,n}\Big)\Big) + a_{2,n}\\ y_n = 2p_n - x_n\\ x_{n+1} = x_n + \lambda_n(p_n - x_n)\\ \text{for } i = 1,\ldots,m\\ \left\lfloor \begin{array}{l} q_{i,n} = \operatorname{prox}_{\sigma g_i^*}\Big(v_{i,n} + \sigma\big(L_i y_n + c_{i,n}\big)\Big) + b_{i,n}\\ v_{i,n+1} = v_{i,n} + \lambda_n(q_{i,n} - v_{i,n}), \end{array}\right. \end{array}\right. \tag{4.10}$$

which is the method proposed in [18, Eq. (36)]. However, in this setting, the conditions (4.4) and (4.3) are different from the conditions [18, Eq. (38)] and [18, Eq. (39)], respectively. Moreover, the present paper provides the strong convergence conditions.

(iv) In finite-dimensional spaces, with exact implementation of the operators, and with the further restriction that $m = 1$, $h\colon x \mapsto 0$, ℓ_1 is as in (4.9), $r_1 = 0$, and $z = 0$, (4.5) remains convergent if $\lambda_n \equiv \lambda \in \left]0,2\right[$ under the same condition presented here [22, Remark 5.4]. If we further impose the restriction $\lambda_n \equiv 1$, then (4.5) reduces to the method proposed in [10, Algorithm 1]. An alternative primal-dual algorithm for this problem is proposed in [12].

Acknowledgement. I thank Professor Patrick L. Combettes for bringing this problem to my attention and for helpful discussions.

References

[1] H. Attouch, J. Bolte, P. Redont, and A. Soubeyran, Alternating proximal algorithms for weakly coupled convex minimization problems – Applications to dynamical games and PDE's, *J. Convex Anal.,* vol. 15, pp. 485–506, 2008.

[2] H. Attouch, L. M. Briceño-Arias, and P. L. Combettes, A parallel splitting method for coupled monotone inclusions, *SIAM J. Control Optim.*, vol. 48, pp. 3246–3270, 2010.

[3] H. Attouch and M. Théra, A general duality principle for the sum of two operators, *J. Convex Anal.,* vol. 3, pp. 1–24, 1996.

[4] J.-B. Baillon and G. Haddad, Quelques propriétés des opérateurs angle-bornés et n-cycliquement monotones, *Israel J. Math.,* vol. 26, pp. 137–150, 1977.

[5] H. H. Bauschke and P. L. Combettes, The Baillon-Haddad theorem revisited, *J. Convex Anal.*, vol. 17, pp. 781–787, 2010.

[6] H. H. Bauschke and P. L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, Springer, New York, 2011.

[7] L. M. Briceño-Arias and P. L. Combettes, A monotone+skew splitting model for composite monotone inclusions in duality, *SIAM J. Optim.*, 2011, to appear.

[8] L. M. Briceño-Arias and P. L. Combettes, Monotone operator methods for Nash equilibria in non-potential games. http://arxiv.org/abs/1106.0144

[9] J.-F. Cai, R. H. Chan, and Z. Shen, Simultaneous cartoon and texture inpainting, *Inverse Probl. Imaging,* vol. 4, pp. 379–395, 2010.

[10] A. Chambolle and T. Pock, A first order primal dual algorithm for convex problems with applications to imaging, *J. Math. Imaging Vision*, vol. 40, pp. 120–145, 2011.

[11] G. H-G. Chen and R. T. Rockafellar, Convergence rates in forward-backward splitting, *SIAM J. Optim.,* vol. 7, pp. 421–444, 1997.

[12] G. Chen and M. Teboulle, A proximal based decomposition method for minimization problems, *Math. Program.*, vol. 64, pp. 81–101, 1994.

[13] P. L. Combettes, Solving monotone inclusions via compositions of nonexpansive averaged operators, *Optimization*, vol. 53, pp. 475–504, 2004.

[14] P. L. Combettes, Dinh Dũng, and B. C. Vũ, Proximity for sums of composite functions, *J. Math. Anal. Appl.*, vol. 380, pp. 680–688, 2011.

[15] P. L. Combettes and J.-C. Pesquet, Proximal splitting methods in signal processing, in: *Fixed-Point Algorithms for Inverse Problems in Science and Engineering,* H. H. Bauschke *et al.* eds., Springer, New York, pp. 185–212, 2011.

[16] P. L. Combettes and J.-C. Pesquet, Primal-dual splitting algorithm for solving inclusions with mixtures of composite, Lipschitzian, and parallel-sum monotone operators, *Set-Valued Var. Anal.*, published on-line August 27, 2011.
http://www.springerlink.com/content/e80v7w7h68534335/.

[17] P. L. Combettes and V. R. Wajs, Signal recovery by proximal forward-backward splitting, *Multiscale Model. Simul.*, vol. 4, pp. 1168–1200, 2005.

[18] L. Condat, A generic first-order primal-dual method for convex optimization involving Lipschitzian, proximable and linear composite terms, 2011.
http://hal.archives-ouvertes.fr/hal-00609728/fr/.

[19] F. Facchinei and J.-S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems*, Springer-Verlag, New York, 2003.

[20] D. Gabay, Applications of the method of multipliers to variational inequalities, in: M. Fortin and R. Glowinski (eds.), *Augmented Lagrangian Methods: Applications to the Numerical Solution of Boundary Value Problems,* pp. 299–331. North-Holland, Amsterdam, 1983.

[21] R. Glowinski and P. Le Tallec, *Augmented Lagrangian and Operator-Splitting Methods in Nonlinear Mechanics*, SIAM, Philadelphia, 1989.

[22] B. He and X. Yuan, Convergence analysis of primal-dual algorithms for total variation image restoration, 2010. http://www.optimization-online.org/DB_HTML/2010/11/2790.html.

[23] B. Mercier, *Topics in Finite Element Solution of Elliptic Problems* (Lectures on Mathematics, no. 63). Tata Institute of Fundamental Research, Bombay, 1979.

[24] U. Mosco, Dual variational inequalities, *J. Math. Anal. Appl.,* vol. 40, pp. 202–206, 1972.

[25] J.-C. Pesquet and N. Pustelnik, A parallel inertial proximal optimization method, 2010. http://www.optimization-online.org/DB_HTML/2010/11/2825.html.

[26] H. Raguet, J. Fadili, and G. Peyré, Generalized forward-backward splitting, 2011.
http://arxiv.org/abs/1108.4404.

[27] R. T. Rockafellar, Duality and stability in extremum problems involving convex functions, *Pacific J. Math.,* vol. 21, pp. 167–187, 1967.

[28] P. Tseng, Further applications of a splitting algorithm to decomposition in variational inequalities and convex programming, *Math. Programming*, vol. 48, pp. 249–263, 1990.

[29] P. Tseng, Applications of a splitting algorithm to decomposition in convex programming and variational inequalities, *SIAM J. Control Optim.*, vol. 29, pp. 119–138, 1991.

[30] D. L. Zhu and P. Marcotte, Co-coercivity and its role in the convergence of iterative schemes for solving variational inequalities, *SIAM J. Optim.*, vol. 6, pp. 714–726, 1996.