# Asynchronous Block-Iterative Primal-Dual Decomposition Methods for Monotone Inclusions*

Patrick L. Combettes [1] and Jonathan Eckstein [2]

[1]Sorbonne Universités – UPMC Univ. Paris 06
UMR 7598, Laboratoire Jacques-Louis Lions
F-75005, Paris, France
plc@ljll.math.upmc.fr

[2]Department of Managemement Science and Information Systems and RUTCOR
Rutgers University
Piscataway, NJ 08854, USA
jeckstei@rci.rutgers.edu

**Abstract**

We propose new primal-dual decomposition algorithms for solving systems of inclusions involving sums of linearly composed maximally monotone operators. The principal innovation in these algorithms is that they are block-iterative in the sense that, at each iteration, only a subset of the monotone operators needs to be processed, as opposed to all operators as in established methods. Deterministic strategies are used to select the blocks of operators activated at each iteration. In addition, we allow for operator processing "lags", permitting asynchronous implementation. The decomposition phase of each iteration of our methods is to generate points in the graphs of the selected monotone operators, in order to construct a half-space containing the Kuhn-Tucker set associated with the system. The coordination phase of each iteration involves a projection onto this half-space. We present two related methods: the first method provides weakly convergent primal and dual sequences under general conditions, while the second is a variant in which strong convergence is guaranteed without additional assumptions. Neither algorithm requires prior knowledge of bounds on the linear operators involved or the inversion of linear operators. Our algorithmic framework unifies and significantly extends the approaches taken in earlier work on primal-dual projective splitting methods.

**Keywords.** asynchronous algorithm, block-iterative algorithm, duality, monotone inclusion, monotone operator, primal-dual algorithm, splitting algorithm

---

*Contact author: P. L. Combettes, plc@ljll.math.upmc.fr, phone: +33 1 4427 6319, fax: +33 1 4427 7200. The work of P. L. Combettes was supported by the CNRS MASTODONS project under grant 2013MesureHD and by the CNRS Imag'in project under grant 2015OPTIMISME.

## 1 Introduction

This paper considers systems of monotone inclusions of the following general form.

**Problem 1.1** Let $m$ and $p$ be strictly positive integers, set $I = \{1,\ldots,m\}$ and $K = \{1,\ldots,p\}$, and let $(\mathcal{H}_i)_{i\in I}$ and $(\mathcal{G}_k)_{k\in K}$ be real Hilbert spaces. For every $i \in I$ and $k \in K$, let $A_i\colon \mathcal{H}_i \to 2^{\mathcal{H}_i}$ and $B_k\colon \mathcal{G}_k \to 2^{\mathcal{G}_k}$ be maximally monotone, let $z_i^* \in \mathcal{H}_i$, let $r_k \in \mathcal{G}_k$, and let $L_{ki}\colon \mathcal{H}_i \to \mathcal{G}_k$ be linear and bounded. Consider the coupled inclusions problem

$$\text{find}\ \ (\overline{x}_i)_{i\in I} \in \bigoplus_{i\in I}\mathcal{H}_i\ \ \text{such that}\ \ (\forall i\in I)\quad z_i^* \in A_i\overline{x}_i + \sum_{k\in K} L_{ki}^*\bigg(B_k\bigg(\sum_{j\in I} L_{kj}\overline{x}_j - r_k\bigg)\bigg), \tag{1.1}$$

its dual problem

$$\text{find}\ \ (\overline{v}_k^*)_{k\in K} \in \bigoplus_{k\in K}\mathcal{G}_k\ \ \text{such that}$$
$$(\forall k\in K)\quad -r_k \in -\sum_{i\in I} L_{ki}\bigg(A_i^{-1}\bigg(z_i^* - \sum_{l\in K} L_{li}^*\overline{v}_l^*\bigg)\bigg) + B_k^{-1}\overline{v}_k^*, \tag{1.2}$$

and the associated Kuhn-Tucker set

$$\boldsymbol{Z} = \bigg\{\big((\overline{x}_i)_{i\in I}, (\overline{v}_k^*)_{k\in K}\big)\ \bigg|\ (\forall i\in I)\ \ \overline{x}_i \in \mathcal{H}_i\ \ \text{and}\ \ z_i^* - \sum_{k\in K} L_{ki}^*\overline{v}_k^* \in A_i\overline{x}_i,\ \ \text{and}$$
$$(\forall k\in K)\ \ \overline{v}_k^* \in \mathcal{G}_k\ \ \text{and}\ \ \sum_{i\in I} L_{ki}\overline{x}_i - r_k \in B_k^{-1}\overline{v}_k^*\bigg\}. \tag{1.3}$$

The problem is to find a point in $\boldsymbol{Z}$. The sets of solutions to (1.1) and (1.2) are denoted by $\mathscr{P}$ and $\mathscr{D}$, respectively.

As discussed in [1], Problem 1.1 models a wide range of problems arising game theory, image recovery, evolution equations, machine learning, signal processing, mechanics, the cognitive sciences, and domain decomposition methods in partial differential equations. In [15, Section 5], it was shown that an important special case of Problem 1.1 is the following optimization problem, in which the monotone operators $(A_i)_{i\in I}$ and $(B_k)_{k\in K}$ are taken to be subdifferentials.

**Problem 1.2** Let $m$ and $p$ be strictly positive integers, set $I = \{1,\ldots,m\}$ and $K = \{1,\ldots,p\}$, and let $(\mathcal{H}_i)_{i\in I}$ and $(\mathcal{G}_k)_{k\in K}$ be real Hilbert spaces. For every $i \in I$ and $k \in K$, let $f_i\colon \mathcal{H}_i \to \left]-\infty,+\infty\right]$ and $g_k\colon \mathcal{G}_k \to \left]-\infty,+\infty\right]$ be proper lower semicontinuous convex functions, let $z_i^* \in \mathcal{H}_i$, let $r_k \in \mathcal{G}_k$, and let $L_{ki}\colon \mathcal{H}_i \to \mathcal{G}_k$ be linear and bounded. Suppose that

$$(\forall i\in I)\quad z_i^* \in \operatorname{range}\bigg(\partial f_i + \sum_{k\in K} L_{ki}^*\circ\partial g_k\circ\bigg(\sum_{j\in I} L_{kj}\cdot - r_k\bigg)\bigg). \tag{1.4}$$

The problem is to solve the primal minimization problem

$$\underset{(x_i)_{i\in I}\in\bigoplus_{i\in I}\mathcal{H}_i}{\text{minimize}}\quad \sum_{i\in I}\big(f_i(x_i) - \langle x_i \mid z_i^*\rangle\big) + \sum_{k\in K} g_k\bigg(\sum_{i\in I} L_{ki}x_i - r_k\bigg) \tag{1.5}$$

along with its dual problem

$$\underset{(v_k^*)_{k\in K}\in\bigoplus_{k\in K}\mathcal{G}_k}{\text{minimize}} \quad \sum_{i\in I} f_i^*\Big(z_i^* - \sum_{k\in K} L_{ki}^* v_k^*\Big) + \sum_{k\in K}\big(g_k^*(v_k^*) + \langle v_k^* \mid r_k\rangle\big). \tag{1.6}$$

In recent years, several decomposition algorithms have been proposed to solve Problem 1.1 (or at least the primal problem (1.1)) under various hypotheses [1, 2, 3, 9, 10, 11, 15, 16, 17, 25]. In such algorithms, the monotone operators as well as the linear operators are evaluated individually. The methods we propose in the present paper for solving Problem 1.1 are based on those of [1, 2], which are themselves based on the projective primal-dual methods initiated in [19, 20] for finding a zero of the sum of monotone operators. The basic idea underlying this class of methods is to generate at each iteration points in the graphs of all the monotone operators in such a way as to construct a half-space containing the Kuhn-Tucker set $\boldsymbol{Z}$. The calculations of each of these points are resolvent computations involving a single monotone operator $A_i$ or $B_k$, which is what makes the methods splitting algorithms. The coordination step of the method is to project the current iterate onto the recently constructed half-space. The advantages of this approach are that it does not impose additional assumptions on the operators present in the formulation, it does not require knowledge of the norm of the linear operators $(L_{ik})_{i\in I,k\in K}$ or of combinations thereof, and it does not involve the inversion of linear operators.

The methods of [1, 2] must evaluate all $m+p$ resolvents of the operators $(A_i)_{i\in I}$ and $(B_k)_{k\in K}$ at every iteration, with only limited ability to pass information between these calculations. Essentially, the resolvents of all the operators $(A_i)_{i\in I}$ must be evaluated independently, and then similarly for all the operators $(B_k)_{k\in K}$. In this setting, the only information flow within each iteration is from the $(A_i)_{i\in I}$ calculations to the $(B_k)_{k\in K}$ calculations. This property results in an algorithm in which large blocks of calculations must be performed before any information is exchanged between subsystems. Although in principle conducive to parallel computing, this kind of structure can still lead to difficulties even in a parallel execution environment: it requires an essentially synchronous implementation, so if some small subset of the subsystems represented by the operators $(A_i)_{i\in I}$ or $(B_k)_{k\in K}$ are more computation-intensive than others, load balancing can become problematic: most processors may have to sit idle while the remaining few complete their tasks. This kind of structure is common to nearly all prior splitting schemes for more than two monotone operators, the only exception we are aware of being that of [20] for the special case

$$\text{find } \overline{x}\in\mathcal{H} \text{ such that } 0\in\sum_{k\in K} B_k\overline{x} \tag{1.7}$$

of (1.1). In that case, information can flow in fairly arbitrary ways between the $p$ resolvent calculations comprising each iteration, as described by a set of algorithm parameters that is quadratic in $p$; the selection of these parameters is subject to a specific eigenvalue condition. However, the algorithm is still fundamentally synchronous, and it is has never been clear how to select its many parameters.

This paper presents a different approach to constructing more flexible and potentially asynchronous decomposition methods for problems fitting the general structure represented by Problem 1.1. The key idea is that our algorithm has the ability to process an essentially arbitrary subset of the operators between successive coordination/projection operations. The only restriction is one adapted from block-iterative methods for convex feasibility problems [6, 13, 23]: for some possibly large positive integer $M$, each operator must be processed at least once over every span of $M$ consecutive iterations. To our knowledge, this is the first application of this kind of versatile deterministic control scheme to finding zeros of sums of operators. Such control schemes have been used in convex feasibility problems [6, 13]. This aspect of our algorithm gives it potential flexibility absent from other splitting schemes for monotone inclusions: first, it provides the ability to find an arbitrary balance between computational effort expended on the subsystems and that expended on coordination. For example, if the subsystems are relatively time-consuming to process, one could perform as few as a single subsystem evaluation between successive projection steps, with the projections immediately spreading the information from each subsystem evaluation to each successive one. The second aspect of the flexibility of our approach involves the balance of computational effort between subsystems: in prior decomposition methods for monotone inclusions, every operator must be processed exactly the same number of times, but the class of algorithms proposed here is much more flexible. If, for example, some operators are less time-consuming to process than others, one has the option of processing them more frequently. Such features can be very useful in applications such as those described in [4].

Our analysis allows each activation of an operator to use information originating from an earlier iteration than the one in which its results are incorporated into the computation. This feature makes it possible to implement the algorithm asynchronously: the points in the graphs of the monotone operators incorporated into the projection step during a given iteration may be the results of resolvent computations initiated during earlier iterations. Our analysis shows that our method still converges so long as there is a fixed (but arbitrary) upper bound on the number of iterations between initiation and incorporation of a resolvent calculation. The potentially asynchronous nature of our method is a significant asset in the design of efficient parallel implementations.

Prior work on projective splitting methods has used two different approaches to constructing affine half-spaces to separate the target set $\boldsymbol{Z}$ from the current iterate. The original approach in [19, 20] was developed for the inclusion problem (1.7). In this special case of (1.1), it was possible to efficiently confine the iterates to a specific subspace $\boldsymbol{\mathcal{K}}$ of the primal-dual space, which can be numerically advantageous. In the general setting of Problem 1.1, the analysis of [1, 2] used an alternative half-space construction in which the iterates are not confined to a subspace. A secondary contribution of this paper is to develop a unifying framework for constructing separators for $\boldsymbol{Z}$ in which both prior approaches appear as special cases.

We present two classes of algorithms based on many of the same underlying building blocks and which may be viewed as asynchronous block-iterative extensions of the algorithms of [1, 2]. The first class uses a straightforward half-space projection at each iteration and allows for conventional overrelaxation of the projection steps by factors upper bounded by $2$. This class exhibits weak convergence to an unspecified Kuhn-Tucker point. The second class is a variant that involves a more complicated projection operation and does not use overrelaxation, but induces strong convergence to the unique point in the Kuhn-Tucker set that best approximates a given reference point. Numerical experiments with these new algorithms are being conducted and we shall report on their results elsewhere.

When applied in suitable product spaces, the block-coordinate methods of [16, 25] can be used to derive block-iterative splitting algorithms methods for a certain class of problems. However, unlike the methods we propose here, the resulting algorithms have been proved to converge only under random operator selection strategies, and they require either joint cocoercivity assumptions on the operators $(B_k)_{k\in K}$ or the ability to block-decompose the projection onto the graph of certain linear operators.

**Notation.** Our notation is standard and follows [8], which contains the necessary background on monotone operators and convex analysis. The scalar product of a Hilbert space is denoted by $\langle\cdot\mid\cdot\rangle$

and the associated norm by $\|\cdot\|$. The projection operator onto a nonempty closed convex subset $C$ of $\mathcal{H}$ is denoted by $P_C$. The symbols $\rightharpoonup$ and $\to$ denote respectively weak and strong convergence, and $\mathrm{Id}$ denotes the identity operator. The Hilbert direct sum of two Hilbert spaces $\mathcal{H}$ and $\mathcal{G}$ is denoted by $\mathcal{H}\oplus\mathcal{G}$, and the power set of $\mathcal{H}$ by $2^{\mathcal{H}}$. Given $A\colon\mathcal{H}\to 2^{\mathcal{H}}$, $\operatorname{gra} A$ denotes the graph of $A$, $A^{-1}$ denotes the inverse of $A$, and $J_A=(\mathrm{Id}+A)^{-1}$ denotes the resolvent of $A$.

## 2 Analysis of a generic primal-dual composite inclusion problem

### 2.1 Problem statement

Our investigation will be simplified by the analysis of the following problem, which can be regarded as a reduction of Problem 1.1 to the case when $m=K=1$, $z_1^*=0$, and $r_1=0$.

**Problem 2.1** Let $\mathcal{H}$ and $\mathcal{G}$ be real Hilbert spaces. Let $A\colon\mathcal{H}\to 2^{\mathcal{H}}$ and $B\colon\mathcal{G}\to 2^{\mathcal{G}}$ be maximally monotone operators, and let $L\colon\mathcal{H}\to\mathcal{G}$ be a bounded linear operator. Consider the inclusion problem

$$\text{find } \overline{x}\in\mathcal{H} \text{ such that } 0\in A\overline{x}+L^*BL\overline{x}, \tag{2.1}$$

its dual problem

$$\text{find } \overline{v}^*\in\mathcal{G} \text{ such that } 0\in -LA^{-1}(-L^*\overline{v}^*)+B^{-1}\overline{v}^*, \tag{2.2}$$

and the associated Kuhn-Tucker set

$$\boldsymbol{Z}=\big\{(x,v^*)\in\mathcal{H}\oplus\mathcal{G}\;\big|\;-L^*v^*\in Ax \text{ and } Lx\in B^{-1}v^*\big\}. \tag{2.3}$$

The problem is to find a point in $\boldsymbol{Z}$. The sets of solutions to (2.1) and (2.2) are denoted by $\mathscr{P}$ and $\mathscr{D}$, respectively.

**Proposition 2.2** *Consider the setting of Problem 2.1 and let $\boldsymbol{\mathcal{K}}$ be a closed vector subspace of $\mathcal{H}\oplus\mathcal{G}$ such that $\boldsymbol{Z}\subset\boldsymbol{\mathcal{K}}$. Then the following hold:*

(i) *$\boldsymbol{Z}$ is a closed convex subset of $\mathscr{P}\times\mathscr{D}$.*

(ii) $\mathscr{P}\neq\varnothing\Leftrightarrow\boldsymbol{Z}\neq\varnothing\Leftrightarrow\mathscr{D}\neq\varnothing$.

(iii) *For every $\mathsf{a}=(a,a^*)\in\operatorname{gra}A$ and $\mathsf{b}=(b,b^*)\in\operatorname{gra}B$, set $\boldsymbol{s}^*_{\mathsf{a},\mathsf{b}}=(a^*+L^*b^*,b-La)$, $\boldsymbol{t}^*_{\mathsf{a},\mathsf{b}}=P_{\boldsymbol{\mathcal{K}}}\boldsymbol{s}^*_{\mathsf{a},\mathsf{b}}$, $\eta_{\mathsf{a},\mathsf{b}}=\langle a\mid a^*\rangle+\langle b\mid b^*\rangle$, and*

$$\boldsymbol{H}_{\mathsf{a},\mathsf{b}}=\big\{\boldsymbol{x}\in\boldsymbol{\mathcal{K}}\;\big|\;\langle\boldsymbol{x}\mid\boldsymbol{t}^*_{\mathsf{a},\mathsf{b}}\rangle\leqslant\eta_{\mathsf{a},\mathsf{b}}\big\}. \tag{2.4}$$

*Then the following hold:*

(a) *Let $\mathsf{a}\in\operatorname{gra}A$ and $\mathsf{b}\in\operatorname{gra}B$. Then $\boldsymbol{H}_{\mathsf{a},\mathsf{b}}=\boldsymbol{\mathcal{K}}\;\Leftarrow\;\boldsymbol{s}^*_{\mathsf{a},\mathsf{b}}=\boldsymbol{0}\;\Rightarrow\;\big[(a,b^*)\in\boldsymbol{Z}$ and $\eta_{\mathsf{a},\mathsf{b}}=0\big]$.*

(b) $\boldsymbol{Z}=\bigcap_{\mathsf{a}\in\operatorname{gra}A}\bigcap_{\mathsf{b}\in\operatorname{gra}B}\boldsymbol{H}_{\mathsf{a},\mathsf{b}}$.

(iv) *Let $(a_n,a_n^*)_{n\in\mathbb{N}}$ be a sequence in $\operatorname{gra}A$, let $(b_n,b_n^*)_{n\in\mathbb{N}}$ be a sequence in $\operatorname{gra}B$, let $x\in\mathcal{H}$, and let $v^*\in\mathcal{G}$. Suppose that $a_n\rightharpoonup x$, $b_n^*\rightharpoonup v^*$, $a_n^*+L^*b_n^*\to 0$, and $La_n-b_n\to 0$. Then $(x,v^*)\in\boldsymbol{Z}$.*

*Proof*. (i): [12, Proposition 2.8(i)].

(ii): [12, Proposition 2.8(iii)-(v)]; see also [24].

(iii): For every $\mathsf{a} = (a, a^*) \in \operatorname{gra} A$ and $\mathsf{b} = (b, b^*) \in \operatorname{gra} B$, set

$$\boldsymbol{G}_{\mathsf{a},\mathsf{b}} = \big\{\boldsymbol{x} \in \mathcal{H} \oplus \mathcal{G} \;\big|\; \langle \boldsymbol{x} \mid \boldsymbol{s}^*_{\mathsf{a},\mathsf{b}}\rangle \leqslant \eta_{\mathsf{a},\mathsf{b}}\big\}, \tag{2.5}$$

and observe that

$$\begin{aligned}\boldsymbol{H}_{\mathsf{a},\mathsf{b}} &= \big\{\boldsymbol{x} \in \boldsymbol{\mathcal{K}} \;\big|\; \langle \boldsymbol{x} \mid \boldsymbol{t}^*_{\mathsf{a},\mathsf{b}}\rangle \leqslant \eta_{\mathsf{a},\mathsf{b}}\big\} \\ &= \big\{\boldsymbol{x} \in \boldsymbol{\mathcal{K}} \;\big|\; \langle \boldsymbol{x} \mid P_{\boldsymbol{\mathcal{K}}}\boldsymbol{s}^*_{\mathsf{a},\mathsf{b}}\rangle + \langle \boldsymbol{x} \mid P_{\boldsymbol{\mathcal{K}}^\perp}\boldsymbol{s}^*_{\mathsf{a},\mathsf{b}}\rangle \leqslant \eta_{\mathsf{a},\mathsf{b}}\big\} \\ &= \big\{\boldsymbol{x} \in \boldsymbol{\mathcal{K}} \;\big|\; \langle \boldsymbol{x} \mid \boldsymbol{s}^*_{\mathsf{a},\mathsf{b}}\rangle \leqslant \eta_{\mathsf{a},\mathsf{b}}\big\} \\ &= \boldsymbol{\mathcal{K}} \cap \boldsymbol{G}_{\mathsf{a},\mathsf{b}}.\end{aligned} \tag{2.6}$$

(iii)(a): By [1, Proposition 2.2(i)], $\boldsymbol{G}_{\mathsf{a},\mathsf{b}} = \mathcal{H} \oplus \mathcal{G} \;\Leftrightarrow\; \boldsymbol{s}^*_{\mathsf{a},\mathsf{b}} = \boldsymbol{0} \;\Rightarrow\; (a, b^*) \in \boldsymbol{Z}$ and $\eta_{\mathsf{a},\mathsf{b}} = 0$. The claim therefore follows from (2.6).

(iii)(b): By [1, Proposition 2.2(iii)] $\boldsymbol{Z} = \bigcap_{\mathsf{a}\in\operatorname{gra} A}\bigcap_{\mathsf{b}\in\operatorname{gra} B}\boldsymbol{G}_{\mathsf{a},\mathsf{b}}$. Hence, (2.6) yields $\boldsymbol{Z} = \boldsymbol{\mathcal{K}} \cap \boldsymbol{Z} = \bigcap_{\mathsf{a}\in\operatorname{gra} A}\bigcap_{\mathsf{b}\in\operatorname{gra} B}\boldsymbol{H}_{\mathsf{a},\mathsf{b}}$.

(iv): [1, Proposition 2.4]. ☐

**Remark 2.3** As will be seen in Remark 3.5, the subspace $\boldsymbol{\mathcal{K}}$ in Proposition 2.2 adds flexibility to the implementation of our proposed algorithms when certain structures are present in the problem formulation.

**Proposition 2.4** *Problem 1.1 is a special case of Problem 2.1.*

*Proof*. Let us set

$$\begin{cases}\mathcal{H} = \bigoplus_{i\in I}\mathcal{H}_i\\ \mathcal{G} = \bigoplus_{k\in K}\mathcal{G}_k\\ L\colon \mathcal{H} \to \mathcal{G}\colon (x_i)_{i\in I} \mapsto \big(\sum_{i\in I} L_{ki}x_i\big)_{k\in K}\\ A\colon \mathcal{H} \to 2^{\mathcal{H}}\colon (x_i)_{i\in I} \mapsto \times_{i\in I}(-z_i^* + A_ix_i)\\ B\colon \mathcal{G} \to 2^{\mathcal{G}}\colon (y_k)_{k\in K} \mapsto \times_{k\in K}B_k(y_k - r_k).\end{cases} \tag{2.7}$$

Then

$$L^*\colon \mathcal{G} \to \mathcal{H}\colon (y_k)_{k\in K} \mapsto \left(\sum_{k\in K} L^*_{ki}y_k\right)_{i\in I}. \tag{2.8}$$

With these settings, (2.1), (2.2), and (2.3) are respectively equivalent to (1.1), (1.2), and (1.3). ☐

## 2.2 A Fejér monotone algorithm

We first recall some basic results concerning Fejér monotone sequences.

**Proposition 2.5** [14] *Let $\boldsymbol{\mathcal{K}}$ be a real Hilbert space, let $\boldsymbol{C}$ be a nonempty closed convex subset of $\boldsymbol{\mathcal{K}}$, and let $\boldsymbol{x}_0 \in \boldsymbol{\mathcal{K}}$. Suppose that*

$$
\begin{array}{l}
\text{for } n = 0, 1, \ldots \\
\left\lfloor
\begin{array}{l}
\boldsymbol{t}_n^* \in \boldsymbol{\mathcal{K}} \text{ and } \eta_n \in \mathbb{R} \text{ are such that } \boldsymbol{C} \subset \boldsymbol{H}_n = \big\{\boldsymbol{x} \in \boldsymbol{\mathcal{K}} \;\big|\; \langle \boldsymbol{x} \mid \boldsymbol{t}_n^* \rangle \leqslant \eta_n \big\} \\
\lambda_n \in \left]0, 2\right[ \\
\boldsymbol{x}_{n+1} = \boldsymbol{x}_n + \lambda_n (P_{\boldsymbol{H}_n} \boldsymbol{x}_n - \boldsymbol{x}_n).
\end{array}
\right.
\end{array}
\tag{2.9}
$$

*Then the following hold:*

(i) $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ *is Fejér monotone with respect to* $\boldsymbol{C}$: $(\forall \boldsymbol{z} \in \boldsymbol{C})(\forall n \in \mathbb{N})\ \|\boldsymbol{x}_{n+1} - \boldsymbol{z}\| \leqslant \|\boldsymbol{x}_n - \boldsymbol{z}\|$.

(ii) $\sum_{n\in\mathbb{N}} \lambda_n (2 - \lambda_n) \|P_{\boldsymbol{H}_n} \boldsymbol{x}_n - \boldsymbol{x}_n\|^2 < +\infty$.

(iii) *Suppose that, for every $\boldsymbol{x} \in \boldsymbol{\mathcal{K}}$ and every strictly increasing sequence $(q_n)_{n\in\mathbb{N}}$ in $\mathbb{N}$, $\boldsymbol{x}_{q_n} \rightharpoonup \boldsymbol{x} \Rightarrow \boldsymbol{x} \in \boldsymbol{C}$. Then $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ converges weakly to a point in $\boldsymbol{C}$.*

**Algorithm 2.6** Consider the setting of Problem 2.1 and let $\boldsymbol{\mathcal{K}}$ be a closed vector subspace of $\mathcal{H} \oplus \mathcal{G}$ such that $\boldsymbol{Z} \subset \boldsymbol{\mathcal{K}}$. Let $\varepsilon \in \left]0, 1\right[$, let $(x_0, v_0^*) \in \boldsymbol{\mathcal{K}}$, and let $(\lambda_n)_{n\in\mathbb{N}} \in [\varepsilon, 2 - \varepsilon]^{\mathbb{N}}$. Iterate

$$
\begin{array}{l}
\text{for } n = 0, 1, \ldots \\
\left\lfloor
\begin{array}{l}
(a_n, a_n^*) \in \operatorname{gra} A \\
(b_n, b_n^*) \in \operatorname{gra} B \\
(t_n^*, t_n) = P_{\boldsymbol{\mathcal{K}}}(a_n^* + L^* b_n^*, b_n - L a_n) \\
\tau_n = \|t_n^*\|^2 + \|t_n\|^2 \\
\text{if } \tau_n > 0 \\
\left\lfloor
\theta_n = \dfrac{\lambda_n}{\tau_n} \max\Big\{0, \langle x_n \mid t_n^* \rangle + \langle t_n \mid v_n^* \rangle - \langle a_n \mid a_n^* \rangle - \langle b_n \mid b_n^* \rangle \Big\}
\right. \\
\text{else } \theta_n = 0 \\
x_{n+1} = x_n - \theta_n t_n^* \\
v_{n+1}^* = v_n^* - \theta_n t_n.
\end{array}
\right.
\end{array}
\tag{2.10}
$$

**Proposition 2.7** *Consider the setting of Problem 2.1 and Algorithm 2.6, and suppose that $\mathscr{P} \neq \varnothing$. Then the following hold:*

(i) *$(x_n, v_n^*)_{n\in\mathbb{N}}$ is a sequence in $\boldsymbol{\mathcal{K}}$ which is Fejér monotone with respect to $\boldsymbol{Z}$.*

(ii) $\sum_{n\in\mathbb{N}} \|x_{n+1} - x_n\|^2 < +\infty$ *and* $\sum_{n\in\mathbb{N}} \|v_{n+1}^* - v_n^*\|^2 < +\infty$.

(iii) *Suppose that the sequences $(a_n)_{n\in\mathbb{N}}$, $(b_n)_{n\in\mathbb{N}}$, $(a_n^*)_{n\in\mathbb{N}}$, and $(b_n^*)_{n\in\mathbb{N}}$ are bounded. Then*

$$
\varlimsup \big( \langle x_n - a_n \mid a_n^* + L^* v_n^* \rangle + \langle L x_n - b_n \mid b_n^* - v_n^* \rangle \big) \leqslant 0. \tag{2.11}
$$

(iv) *Suppose that, for every $(x, v^*) \in \boldsymbol{\mathcal{K}}$ and for every strictly increasing sequence $(q_n)_{n\in\mathbb{N}}$ in $\mathbb{N}$,*

$$
\big[\, x_{q_n} \rightharpoonup x \ \text{ and } \ v_{q_n}^* \rightharpoonup v^* \,\big] \quad \Rightarrow \quad (x, v^*) \in \boldsymbol{Z}. \tag{2.12}
$$

*Then $(x_n)_{n\in\mathbb{N}}$ converges weakly to a point $\overline{x} \in \mathscr{P}$, $(v_n^*)_{n\in\mathbb{N}}$ converges weakly to a point $\overline{v}^* \in \mathscr{D}$, and $(\overline{x}, \overline{v}^*) \in \boldsymbol{Z}$.*

*Proof*. Parts (i) and (ii) of Proposition 2.2 assert that $\boldsymbol{Z}$ is a nonempty, closed, and convex subset of $\boldsymbol{\mathcal{K}}$. Now set

$$(\forall n \in \mathbb{N}) \quad \boldsymbol{x}_n = (x_n, v_n^*), \quad \boldsymbol{s}_n^* = (s_n^*, s_n) = (a_n^* + L^* b_n^*, b_n - La_n), \quad \boldsymbol{t}_n^* = (t_n^*, t_n),$$
$$\eta_n = \langle a_n \mid a_n^* \rangle + \langle b_n \mid b_n^* \rangle, \quad \text{and} \quad \boldsymbol{H}_n = \big\{ \boldsymbol{x} \in \boldsymbol{\mathcal{K}} \;\big|\; \langle \boldsymbol{x} \mid \boldsymbol{t}_n^* \rangle \leqslant \eta_n \big\}. \quad (2.13)$$

Then it follows from (2.10) and Proposition 2.2(iii)(b) that $(\forall n \in \mathbb{N})$ $\boldsymbol{Z} \subset \boldsymbol{H}_n$. Set $(\forall n \in \mathbb{N})$ $\Delta_n = \sqrt{\tau_n}\theta_n/\lambda_n$. Using [8, Example 28.16(iii)], we get

$$(\forall n \in \mathbb{N}) \quad P_{\boldsymbol{H}_n}\boldsymbol{x}_n = \begin{cases} \boldsymbol{x}_n + \dfrac{\eta_n - \langle \boldsymbol{x}_n \mid \boldsymbol{t}_n^* \rangle}{\|\boldsymbol{t}_n^*\|^2}\boldsymbol{t}_n^*, & \text{if } \boldsymbol{t}_n^* \neq \boldsymbol{0} \text{ and } \langle \boldsymbol{x}_n \mid \boldsymbol{t}_n^* \rangle > \eta_n; \\ \boldsymbol{x}_n, & \text{otherwise.} \end{cases} \quad (2.14)$$

Hence,

$$(\forall n \in \mathbb{N}) \quad \Delta_n = \|P_{\boldsymbol{H}_n}\boldsymbol{x}_n - \boldsymbol{x}_n\| \quad \text{and} \quad \boldsymbol{x}_{n+1} = \boldsymbol{x}_n + \lambda_n(P_{\boldsymbol{H}_n}\boldsymbol{x}_n - \boldsymbol{x}_n). \quad (2.15)$$

Therefore, we derive from Proposition 2.5(ii) that

$$\sum_{n \in \mathbb{N}} \Delta_n^2 < +\infty. \quad (2.16)$$

(i): This follows from (2.15) and Proposition 2.5(i).

(ii): We derive from (2.10) that

$$(\forall n \in \mathbb{N}) \quad \|x_{n+1} - x_n\|^2 + \|v_{n+1}^* - v_n^*\|^2 = \theta_n^2\tau_n = \lambda_n^2\Delta_n^2 \leqslant 4\Delta_n^2. \quad (2.17)$$

Hence, the claim follows from (2.16).

(iii): Since $\|P_{\boldsymbol{\mathcal{K}}}\| \leqslant 1$, (2.10) and (2.13) yield

$$\begin{aligned}(\forall n \in \mathbb{N}) \quad \tau_n &= \|\boldsymbol{t}_n^*\|^2 \\ &\leqslant \|\boldsymbol{s}_n^*\|^2 \\ &= \|a_n^* + L^*b_n^*\|^2 + \|La_n - b_n\|^2 \\ &\leqslant 2\big(\|a_n^*\|^2 + \|L\|^2\,\|b_n^*\|^2 + \|L\|^2\|a_n\|^2 + \|b_n\|^2\big). \end{aligned} \quad (2.18)$$

Hence, $(\tau_n)_{n\in\mathbb{N}}$ is bounded. Therefore, since (2.16) implies that $\Delta_n \to 0$ and since $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ lies in $\boldsymbol{\mathcal{K}}$, we obtain

$$\begin{aligned}(\forall n \in \mathbb{N}) \quad & \max\big\{0, (\langle x_n \mid s_n^* \rangle + \langle s_n \mid v_n^* \rangle - \langle a_n \mid a_n^* \rangle - \langle b_n \mid b_n^* \rangle)\big\} \\ &= \max\big\{0, (\langle \boldsymbol{x}_n \mid \boldsymbol{s}_n^* \rangle - \langle a_n \mid a_n^* \rangle - \langle b_n \mid b_n^* \rangle)\big\} \\ &= \max\big\{0, (\langle \boldsymbol{x}_n \mid P_{\boldsymbol{\mathcal{K}}}\boldsymbol{s}_n^* \rangle - \langle a_n \mid a_n^* \rangle - \langle b_n \mid b_n^* \rangle)\big\} \\ &= \max\big\{0, (\langle x_n \mid t_n^* \rangle + \langle t_n \mid v_n^* \rangle - \langle a_n \mid a_n^* \rangle - \langle b_n \mid b_n^* \rangle)\big\} \\ &= \sqrt{\tau_n}\Delta_n \\ &\to 0. \end{aligned} \quad (2.19)$$

Consequently,

$$\varlimsup \big(\langle x_n - a_n \mid a_n^* + L^*v_n^* \rangle + \langle Lx_n - b_n \mid b_n^* - v_n^* \rangle\big)$$
$$= \varlimsup \big(\langle x_n \mid s_n^* \rangle + \langle s_n \mid v_n^* \rangle - \langle a_n \mid a_n^* \rangle - \langle b_n \mid b_n^* \rangle\big) \leqslant 0. \quad (2.20)$$

(iv): This follows from (2.15) and Proposition 2.5(iii). □

### 2.3 An Haugazeau-like algorithm

Algorithm 2.6 produces sequences that converge weakly to some undetermined point in $\boldsymbol{Z}$. We now describe an algorithm that provides strong convergence to the point in $\boldsymbol{Z}$ closest to some reference point $(x_0, v_0^*) \in \mathcal{H} \oplus \mathcal{G}$. This approach relies on a geometric construction going back to [21] and was used in the context of Problem 1.1 in [2].

Let $(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}) \in \boldsymbol{\mathcal{K}}^3$ be an ordered triplet from a real Hilbert space $\boldsymbol{\mathcal{K}}$. We define

$$H(\boldsymbol{x}, \boldsymbol{y}) = \big\{ \boldsymbol{h} \in \boldsymbol{\mathcal{K}} \;\big|\; \langle \boldsymbol{h} - \boldsymbol{y} \mid \boldsymbol{x} - \boldsymbol{y} \rangle \leqslant 0 \big\} \tag{2.21}$$

and, if the set $H(\boldsymbol{x}, \boldsymbol{y}) \cap H(\boldsymbol{y}, \boldsymbol{z})$ is nonempty, we denote by $Q(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})$ the projection of $\boldsymbol{x}$ onto it. The principle of the algorithm to project a point $\boldsymbol{x}_0 \in \boldsymbol{\mathcal{K}}$ onto a nonempty closed convex set $\boldsymbol{C} \subset \boldsymbol{\mathcal{K}}$ is to use at iteration $n$ the current iterate $\boldsymbol{x}_n$ to construct an outer approximation to $\boldsymbol{C}$ of the form $H(\boldsymbol{x}_0, \boldsymbol{x}_n) \cap H(\boldsymbol{x}_n, \boldsymbol{x}_{n+1/2})$; the update is then computed as the projection of $\boldsymbol{x}_0$ onto this intersection, i.e., $\boldsymbol{x}_{n+1} = Q(\boldsymbol{x}_0, \boldsymbol{x}_n, \boldsymbol{x}_{n+1/2})$. As the following lemma from [21] shows, this last computation is straightforward; an alternative derivation may be found in [8, Corollary 28.21].

**Lemma 2.8** ([21, Théorème 3-1]) *Let $\boldsymbol{\mathcal{K}}$ be a real Hilbert space, let $(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}) \in \boldsymbol{\mathcal{K}}^3$, and set $\boldsymbol{R} = H(\boldsymbol{x}, \boldsymbol{y}) \cap H(\boldsymbol{y}, \boldsymbol{z})$. Further, set $\chi = \langle \boldsymbol{x} - \boldsymbol{y} \mid \boldsymbol{y} - \boldsymbol{z} \rangle$, $\mu = \|\boldsymbol{x} - \boldsymbol{y}\|^2$, $\nu = \|\boldsymbol{y} - \boldsymbol{z}\|^2$, and $\rho = \mu\nu - \chi^2$. Then exactly one of the following holds:*

(i) *$\rho = 0$ and $\chi < 0$, in which case $\boldsymbol{R} = \varnothing$.*

(ii) *$[\,\rho = 0$ and $\chi \geqslant 0\,]$ or $\rho > 0$, in which case $\boldsymbol{R} \neq \varnothing$ and*

$$Q(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}) = \begin{cases} \boldsymbol{z}, & \text{if } \rho = 0 \text{ and } \chi \geqslant 0; \\ \boldsymbol{x} + (1 + \chi/\nu)(\boldsymbol{z} - \boldsymbol{y}), & \text{if } \rho > 0 \text{ and } \chi\nu \geqslant \rho; \\ \boldsymbol{y} + (\nu/\rho)\big(\chi(\boldsymbol{x} - \boldsymbol{y}) + \mu(\boldsymbol{z} - \boldsymbol{y})\big), & \text{if } \rho > 0 \text{ and } \chi\nu < \rho. \end{cases} \tag{2.22}$$

**Proposition 2.9** ([2, Proposition 2.1]) *Let $\boldsymbol{\mathcal{K}}$ be a real Hilbert space, let $\boldsymbol{C}$ be a nonempty closed convex subset of $\boldsymbol{\mathcal{K}}$, and let $\boldsymbol{x}_0 \in \boldsymbol{\mathcal{K}}$. Iterate*

$$\begin{array}{l} \text{for } n = 0, 1, \ldots \\ \left\lfloor \begin{array}{l} \text{take } \boldsymbol{x}_{n+1/2} \in \boldsymbol{\mathcal{K}} \text{ such that } \boldsymbol{C} \subset H(\boldsymbol{x}_n, \boldsymbol{x}_{n+1/2}) \\ \boldsymbol{x}_{n+1} = Q\big(\boldsymbol{x}_0, \boldsymbol{x}_n, \boldsymbol{x}_{n+1/2}\big). \end{array} \right. \end{array} \tag{2.23}$$

*Then the sequence $(\boldsymbol{x}_n)_{n \in \mathbb{N}}$ is well defined and the following hold:*

(i) $(\forall n \in \mathbb{N})\ \|\boldsymbol{x}_n - \boldsymbol{x}_0\| \leqslant \|\boldsymbol{x}_{n+1} - \boldsymbol{x}_0\| \leqslant \|P_{\boldsymbol{C}}\boldsymbol{x}_0 - \boldsymbol{x}_0\|$.

(ii) $(\forall n \in \mathbb{N})\ \boldsymbol{C} \subset H(\boldsymbol{x}_0, \boldsymbol{x}_n) \cap H(\boldsymbol{x}_n, \boldsymbol{x}_{n+1/2})$.

(iii) $\sum_{n \in \mathbb{N}} \|\boldsymbol{x}_{n+1} - \boldsymbol{x}_n\|^2 < +\infty$.

(iv) $\sum_{n \in \mathbb{N}} \|\boldsymbol{x}_{n+1/2} - \boldsymbol{x}_n\|^2 < +\infty$.

(v) *Suppose that, for every $\boldsymbol{x} \in \boldsymbol{\mathcal{K}}$ and every strictly increasing sequence $(q_n)_{n \in \mathbb{N}}$ in $\mathbb{N}$, $\boldsymbol{x}_{q_n} \rightharpoonup \boldsymbol{x} \Rightarrow \boldsymbol{x} \in \boldsymbol{C}$. Then $\boldsymbol{x}_n \to P_{\boldsymbol{C}}\boldsymbol{x}_0$.*

**Algorithm 2.10** Consider the setting of Problem 2.1 and let $\boldsymbol{\mathcal{K}}$ be a closed vector subspace of $\mathcal{H} \oplus \mathcal{G}$ such that $\boldsymbol{Z} \subset \boldsymbol{\mathcal{K}}$. Let $\varepsilon \in \,]0,1[$, let $(x_0, v_0^*) \in \boldsymbol{\mathcal{K}}$, and let $(\lambda_n)_{n\in\mathbb{N}} \in [\varepsilon, 1]^{\mathbb{N}}$. Iterate

$$
\begin{array}{l}
\text{for } n = 0,1,\ldots\\
\left\lfloor
\begin{array}{l}
(a_n, a_n^*) \in \operatorname{gra} A\\
(b_n, b_n^*) \in \operatorname{gra} B\\
(t_n^*, t_n) = P_{\boldsymbol{\mathcal{K}}}(a_n^* + L^*b_n^*, b_n - La_n)\\
\tau_n = \|t_n^*\|^2 + \|t_n\|^2\\
\text{if } \tau_n > 0\\
\left\lfloor
\theta_n = \dfrac{\lambda_n}{\tau_n} \max\Big\{0, \langle x_n \mid t_n^*\rangle + \langle t_n \mid v_n^*\rangle - \langle a_n \mid a_n^*\rangle - \langle b_n \mid b_n^*\rangle\Big\}
\right.\\
\text{else } \theta_n = 0\\
x_{n+1/2} = x_n - \theta_n t_n^*\\
v_{n+1/2}^* = v_n^* - \theta_n t_n\\
(x_{n+1}, v_{n+1}^*) = Q\big((x_0, v_0^*), (x_n, v_n^*), (x_{n+1/2}, v_{n+1/2}^*)\big).
\end{array}
\right.
\end{array}
\tag{2.24}
$$

**Remark 2.11** Using Lemma 2.8, the computation of the update $(x_{n+1}, v_{n+1}^*)$ in (2.24) can be explicitly broken into the following steps:

$$
\begin{array}{l}
\chi_n = \langle x_0 - x_n \mid x_n - x_{n+1/2}\rangle + \langle v_0^* - v_n^* \mid v_n^* - v_{n+1/2}^*\rangle\\
\mu_n = \|x_0 - x_n\|^2 + \|v_0^* - v_n^*\|^2\\
\nu_n = \|x_n - x_{n+1/2}\|^2 + \|v_n^* - v_{n+1/2}^*\|^2\\
\rho_n = \mu_n\nu_n - \chi_n^2\\
\text{if } \rho_n = 0 \text{ and } \chi_n \geqslant 0\\
\left\lfloor
\begin{array}{l}
x_{n+1} = x_{n+1/2}\\
v_{n+1}^* = v_{n+1/2}^*
\end{array}
\right.\\
\text{if } \rho_n > 0 \text{ and } \chi_n\nu_n \geqslant \rho_n\\
\left\lfloor
\begin{array}{l}
x_{n+1} = x_0 + (1 + \chi_n/\nu_n)(x_{n+1/2} - x_n)\\
v_{n+1}^* = v_0^* + (1 + \chi_n/\nu_n)(v_{n+1/2}^* - v_n^*)
\end{array}
\right.\\
\text{if } \rho_n > 0 \text{ and } \chi_n\nu_n < \rho_n\\
\left\lfloor
\begin{array}{l}
x_{n+1} = x_n + (\nu_n/\rho_n)\big(\chi_n(x_0 - x_n) + \mu_n(x_{n+1/2} - x_n)\big)\\
v_{n+1}^* = v_n^* + (\nu_n/\rho_n)\big(\chi_n(v_0^* - v_n^*) + \mu_n(v_{n+1/2}^* - v_n^*)\big).
\end{array}
\right.
\end{array}
\tag{2.25}
$$

**Proposition 2.12** *Consider the setting of Problem 2.1 and Algorithm 2.10. Suppose that $\mathscr{P} \neq \varnothing$ and set $(\overline{x}, \overline{v}^*) = P_{\boldsymbol{Z}}(x_0, v_0^*)$. Then the following hold:*

(i) *$(x_n)_{n\in\mathbb{N}}$ and $(v_n^*)_{n\in\mathbb{N}}$ are bounded.*

(ii) *$\sum_{n\in\mathbb{N}} \|x_{n+1} - x_n\|^2 < +\infty$ and $\sum_{n\in\mathbb{N}} \|v_{n+1}^* - v_n^*\|^2 < +\infty$.*

(iii) *$\sum_{n\in\mathbb{N}} \|x_{n+1/2} - x_n\|^2 < +\infty$ and $\sum_{n\in\mathbb{N}} \|v_{n+1/2}^* - v_n^*\|^2 < +\infty$.*

(iv) *Suppose that the sequences $(a_n)_{n\in\mathbb{N}}$, $(b_n)_{n\in\mathbb{N}}$, $(a_n^*)_{n\in\mathbb{N}}$, and $(b_n^*)_{n\in\mathbb{N}}$ are bounded. Then*

$$\varlimsup \big(\langle x_n - a_n \mid a_n^* + L^*v_n^*\rangle + \langle Lx_n - b_n \mid b_n^* - v_n^*\rangle\big) \leqslant 0. \tag{2.26}$$

(v) *Suppose that, for every $(x, v^*) \in \boldsymbol{\mathcal{K}}$ and every strictly increasing sequence $(q_n)_{n\in\mathbb{N}}$ in $\mathbb{N}$,*

$$\big[\, x_{q_n} \rightharpoonup x \;\text{ and }\; v_{q_n}^* \rightharpoonup v^* \,\big] \quad\Rightarrow\quad (x, v^*) \in \boldsymbol{Z}. \tag{2.27}$$

*Then $(x_n)_{n\in\mathbb{N}}$ converges strongly to $\overline{x} \in \mathscr{P}$ and $(v_n^*)_{n\in\mathbb{N}}$ converges strongly to $\overline{v}^* \in \mathscr{D}$.*

*Proof*. We first show that we recover the setting of Proposition 2.9 applied in $\boldsymbol{\mathcal{K}}$ to the set $\boldsymbol{Z}$ of (2.3), which is nonempty, closed, and convex by Proposition 2.2(i)–(ii). Set

$$(\forall n\in\mathbb{N})\quad \boldsymbol{x}_n=(x_n,v_n^*),\quad \boldsymbol{x}_{n+1/2}=(x_{n+1/2},v_{n+1/2}^*),\quad \boldsymbol{t}_n^*=(t_n^*,t_n),\quad\text{and}\quad \eta_n=\langle a_n\mid a_n^*\rangle+\langle b_n\mid b_n^*\rangle. \tag{2.28}$$

If, for some $n\in\mathbb{N}$, we have $\boldsymbol{x}_{n+1/2}=\boldsymbol{x}_n$, then trivially $\boldsymbol{Z}\subset H(\boldsymbol{x}_n,\boldsymbol{x}_{n+1/2})=\boldsymbol{\mathcal{K}}$; otherwise, (2.24) imposes that $\langle \boldsymbol{x}_n\mid \boldsymbol{t}_n^*\rangle>\eta_n$ and therefore that

$$\begin{aligned}\eta_n&\leqslant\langle \boldsymbol{x}_n\mid \boldsymbol{t}_n^*\rangle-\lambda_n\big(\langle \boldsymbol{x}_n\mid \boldsymbol{t}_n^*\rangle-\eta_n\big)\\&=\langle \boldsymbol{x}_n\mid \boldsymbol{t}_n^*\rangle-\theta_n\tau_n\\&=\langle \boldsymbol{x}_n-\theta_n\boldsymbol{t}_n^*\mid \boldsymbol{t}_n^*\rangle\\&=\langle \boldsymbol{x}_{n+1/2}\mid \boldsymbol{t}_n^*\rangle,\end{aligned} \tag{2.29}$$

from which we deduce using Proposition 2.2(iii) that

$$\begin{aligned}\boldsymbol{Z}&\subset\big\{\boldsymbol{x}\in\boldsymbol{\mathcal{K}}\;\big|\;\langle \boldsymbol{x}\mid \boldsymbol{t}_n^*\rangle\leqslant\eta_n\big\}\\&\subset\big\{\boldsymbol{x}\in\boldsymbol{\mathcal{K}}\;\big|\;\langle \boldsymbol{x}\mid \boldsymbol{t}_n^*\rangle\leqslant\langle \boldsymbol{x}_{n+1/2}\mid \boldsymbol{t}_n^*\rangle\big\}\\&=\big\{\boldsymbol{x}\in\boldsymbol{\mathcal{K}}\;\big|\;\langle \boldsymbol{x}\mid \boldsymbol{x}_n-\boldsymbol{x}_{n+1/2}\rangle\leqslant\langle \boldsymbol{x}_{n+1/2}\mid \boldsymbol{x}_n-\boldsymbol{x}_{n+1/2}\rangle\big\}\\&=H\big(\boldsymbol{x}_n,\boldsymbol{x}_{n+1/2}\big).\end{aligned} \tag{2.30}$$

Altogether, (2.24) is an instance of (2.23) with $\boldsymbol{C}=\boldsymbol{Z}$, and we can apply Proposition 2.9. In particular, Proposition 2.9(ii) asserts that $(x_n,v_n^*)_{n\in\mathbb{N}}$ is well defined. We can now establish the claims of the proposition as follows.

(i): This is a consequence of Proposition 2.9(i).

(ii): It follows from (2.28) and Proposition 2.9(iii) that $\sum_{n\in\mathbb{N}}\|x_{n+1}-x_n\|^2+\sum_{n\in\mathbb{N}}\|v_{n+1}^*-v_n^*\|^2=\sum_{n\in\mathbb{N}}\|\boldsymbol{x}_{n+1}-\boldsymbol{x}_n\|^2<+\infty$.

(iii): In view of (2.28) and Proposition 2.9(iv), $\sum_{n\in\mathbb{N}}\|x_{n+1/2}-x_n\|^2+\sum_{n\in\mathbb{N}}\|v_{n+1/2}^*-v_n^*\|^2=\sum_{n\in\mathbb{N}}\|\boldsymbol{x}_{n+1/2}-\boldsymbol{x}_n\|^2<+\infty$.

(iv): Set $(\forall n\in\mathbb{N})$ $\Delta_n=\sqrt{\tau_n}\theta_n/\lambda_n$. We derive from (2.24) and (iii) that

$$\sum_{n\in\mathbb{N}}\Delta_n^2=\sum_{n\in\mathbb{N}}\frac{\tau_n\theta_n^2}{\lambda_n^2}\leqslant\sum_{n\in\mathbb{N}}\frac{\tau_n\theta_n^2}{\varepsilon^2}=\sum_{n\in\mathbb{N}}\frac{\|\boldsymbol{x}_{n+1/2}-\boldsymbol{x}_n\|^2}{\varepsilon^2}<+\infty. \tag{2.31}$$

The claim is then obtained by arguing as in the proof of Proposition 2.7(iii).

(v): This follows directly from Proposition 2.9(v). □

**Remark 2.13** Proposition 2.12 guarantees strong convergence to the projection of the initial point $(x_0,v_0^*)$ onto the Kuhn-Tucker set under the same conditions that provide weak convergence to an unspecified Kuhn-Tucker point in Proposition 2.7. This phenomenon is akin to the weak-to-strong convergence principle investigated in a fixed-point setting in [7].

## 3 Solving Problem 1.1

### 3.1 Block iterations and asynchronicity

In existing monotone operator splitting methods, each operator in the inclusion problem must be used at each iteration $n$ in a resolvent calculation that must be based on information available at the current iteration. For instance, the methods of [1, 2] require points $(a_{i,n}, a^*_{i,n}) \in \operatorname{gra} A_i$ and $(b_{k,n}, b^*_{k,n}) \in \operatorname{gra} B_k$ for every $i \in I$ and every $k \in K$, and these points must be computed using the current values of the primal variables $(x_{i,n})_{i\in I}$ and of the dual variables $(v^*_{k,n})_{k\in K}$. The earlier work in [19, 20] in the context of (1.7) is similar. The two main novelties we present in this paper are to depart from this approach by allowing asynchronous block iterations. Specifically, we allow:

**Block iterations:** At iteration $n$, we require calculation of new points in the graphs of only some of the operators, say $(A_i)_{i\in I_n\subset I}$ and $(B_k)_{k\in K_n\subset K}$. The control sequences $(I_n)_{n\in\mathbb{N}}$ and $(K_n)_{n\in\mathbb{N}}$ dictate how frequently the various operators are used.

**Asynchronicity:** A new point $(a_{i,n}, a^*_{i,n}) \in \operatorname{gra} A_i$ being incorporated into the calculations at iteration $n$ may be based on data $x_{i,c_i(n)}$ and $(v^*_{k,c_i(n)})_{k\in K}$ available at some possibly earlier iteration $c_i(n) \leqslant n$. Therefore, the calculation of $(a_{i,n}, a^*_{i,n})$ could have been initiated at iteration $c_i(n)$, with its results becoming available only at iteration $n$. Likewise, for every $k \in K_n$, the computation of $(b_{k,n}, b^*_{k,n}) \in \operatorname{gra} B_k$ can be initiated at some iteration $d_k(n) \leqslant n$, based on $(x_{i,d_k(n)})_{i\in I}$ and $v^*_{k,d_k(n)}$.

To establish convergence, there needs to be some limits on the asynchronous asynchronicity lag of the algorithm and the spacing between successive calculations involving each operator, as described in the following assumption.

**Assumption 3.1**

(i) *$M$ is a strictly positive integer, $(I_n)_{n\in\mathbb{N}}$ is a sequence of nonempty subsets of $I$, and $(K_n)_{n\in\mathbb{N}}$ is a sequence of nonempty subsets of $K$ such that*

$$I_0 = I,\ K_0 = K, \quad \textit{and} \quad (\forall n \in \mathbb{N}) \quad \left( \bigcup_{j=n}^{n+M-1} I_j = I \quad \textit{and} \quad \bigcup_{j=n}^{n+M-1} K_j = K \right). \tag{3.1}$$

(ii) *$D$ is a positive integer and, for every $i \in I$ and every $k \in K$, $(c_i(n))_{n\in\mathbb{N}}$ and $(d_k(n))_{n\in\mathbb{N}}$ are sequences in $\mathbb{N}$ such that*

$$(\forall n \in \mathbb{N}) \quad \Big( (\forall i \in I)\ \ n - D \leqslant c_i(n) \leqslant n \quad \textit{and} \quad (\forall k \in K)\ \ n - D \leqslant d_k(n) \leqslant n \Big). \tag{3.2}$$

(iii) *$\varepsilon \in \left]0,1\right[$ and, for every $i \in I$ and every $k \in K$, $(\gamma_{i,n})_{n\in\mathbb{N}}$ and $(\mu_{k,n})_{n\in\mathbb{N}}$ are sequences in $[\varepsilon, 1/\varepsilon]$.*

At iteration $n$, our algorithms incorporates points in the graphs of the operators $(A_i)_{i\in I_n}$ and $(B_k)_{k\in K_n}$. Condition (3.1) ensures that over any span of $M$ consecutive iterations, each operator is incorporated into the algorithm at least once. The standard case corresponds to using all the operators at each iteration, *i.e.* $(\forall n \in \mathbb{N})$ $I_n = I$ and $K_n = K$. Toward the other extreme, it is possible to use just one of the operators from $(A_i)_{i\in I}$ and $(B_k)_{k\in K}$ at iteration $n$. For example, such a control regime could be achieved by setting $M = \max\{m,p\}$ and sweeping though the operators in a periodic manner. Condition (3.2) guarantees that the points in the graphs incorporated into the algorithm are based on information at most $D$ iterations out of date. If the algorithm is being implemented synchronously, then one can simply set $D = 0$, in which case $(\forall n \in \mathbb{N})(\forall i \in I)(\forall k \in K)$ $c_i(n) = n$ and $d_k(n) = n$. Finally, the positive scalars $(\gamma_{i,n})_{n\in\mathbb{N}}$ and $(\mu_{k,n})_{n\in\mathbb{N}}$ in (iii) are the proximal parameters used in the resolvent calculations. The assumption requires that they be bounded above and also away from 0.

The following result is the key asymptotic principle on which our two main theorems will rest. The key idea of our algorithm is to simply recycle an old point in the graph of each operator for which new information is not available.

**Proposition 3.2** *Consider the setting of Problem 1.1 and suppose that the following are satisfied:*

(a) *For every $i \in I$, $(x_{i,n})_{n\in\mathbb{N}}$ is a bounded sequence in $\mathcal{H}_i$ and, for every $k \in K$, $(v^*_{k,n})_{n\in\mathbb{N}}$ is a bounded sequence in $\mathcal{G}_k$.*

(b) *Assumption 3.1 is in force.*

(c) *For every $n \in \mathbb{N}$, set*

$$
\begin{array}{l}
\text{for every } i \in I_n\\
\left\lfloor
\begin{array}{l}
l^*_{i,n} = \sum_{k\in K} L^*_{ki} v^*_{k,c_i(n)}\\
(a_{i,n}, a^*_{i,n}) = \Big(J_{\gamma_{i,c_i(n)}A_i}\big(x_{i,c_i(n)} + \gamma_{i,c_i(n)}(z^*_i - l^*_{i,n})\big), \gamma^{-1}_{i,c_i(n)}(x_{i,c_i(n)} - a_{i,n}) - l^*_{i,n}\Big)
\end{array}
\right.\\
\text{for every } i \in I \smallsetminus I_n\\
\left\lfloor\; (a_{i,n}, a^*_{i,n}) = (a_{i,n-1}, a^*_{i,n-1})\right.\\
\text{for every } k \in K_n\\
\left\lfloor
\begin{array}{l}
l_{k,n} = \sum_{i\in I} L_{ki} x_{i,d_k(n)}\\
(b_{k,n}, b^*_{k,n}) = \Big(r_k + J_{\mu_{k,d_k(n)}B_k}\big(l_{k,n} + \mu_{k,d_k(n)} v^*_{k,d_k(n)} - r_k\big), v^*_{k,d_k(n)} + \mu^{-1}_{k,d_k(n)}(l_{k,n} - b_{k,n})\Big)
\end{array}
\right.\\
\text{for every } k \in K \smallsetminus K_n\\
\left\lfloor\; (b_{k,n}, b^*_{k,n}) = (b_{k,n-1}, b^*_{k,n-1}),\right.
\end{array}
\tag{3.3}
$$

*and define*

$$(\forall n \in \mathbb{N}) \quad a_n = (a_{i,n})_{i\in I}, \quad a^*_n = (a^*_{i,n})_{i\in I}, \quad b_n = (b_{k,n})_{k\in K}, \quad \textit{and} \quad b^*_n = (b^*_{k,n})_{k\in K}. \tag{3.4}$$

*Then the following hold:*

(i) *Define $A$ and $B$ as in (2.7). Then $(\forall n \in \mathbb{N})$ $(a_n, a^*_n) \in \operatorname{gra} A$ and $(b_n, b^*_n) \in \operatorname{gra} B$.*

(ii) *$(a_n)_{n\in\mathbb{N}}$, $(a^*_n)_{n\in\mathbb{N}}$, $(b_n)_{n\in\mathbb{N}}$, and $(b^*_n)_{n\in\mathbb{N}}$ are bounded.*

(iii) *Suppose that the following are satisfied:*

  (d) *$(\forall i \in I)$ $\sum_{n\in\mathbb{N}} \|x_{i,n+1} - x_{i,n}\|^2 < +\infty$ and $(\forall k \in K)$ $\sum_{n\in\mathbb{N}} \|v^*_{k,n+1} - v^*_{k,n}\|^2 < +\infty$.*

  (e) $\varlimsup \big(\sum_{i\in I} \langle x_{i,n} - a_{i,n} \mid a^*_{i,n} + \sum_{k\in K} L^*_{ki} v^*_{k,n}\rangle + \sum_{k\in K} \langle \sum_{i\in I} L_{ki} x_{i,n} - b_{k,n} \mid b^*_{k,n} - v^*_{k,n}\rangle\big) \leqslant 0.$

  (f) *$(q_n)_{n\in\mathbb{N}}$ is a strictly increasing sequence in $\mathbb{N}$, for every $i \in I$, $x_i \in \mathcal{H}_i$ and $x_{i,q_n} \rightharpoonup x_i$, and, for every $k \in K$, $v^*_k \in \mathcal{G}_k$ and $v^*_{k,q_n} \rightharpoonup v^*_k$.*

*Then* $((x_i)_{i\in I}, (v_k^*)_{k\in K}) \in \boldsymbol{Z}$.

*Proof*. Define $\mathcal{H}$, $\mathcal{G}$, and $L$ as in (2.7) and set

$$(\forall n \in \mathbb{N}) \quad x_n = (x_{i,n})_{i\in I} \quad \text{and} \quad v_n^* = (v_{k,n}^*)_{k\in K}. \tag{3.5}$$

(i): This follows from (3.3) and basic resolvent calculus rules [8, Propositions 23.15 and 23.16].

(ii): Let $i \in I$. We derive from hypothesis (a) and Assumption 3.1(iii) that the sequence $\big(x_{i,c_i(n)} - \gamma_{i,c_i(n)} \sum_{k\in K} L_{ki}^* v_{k,c_i(n)}^*\big)_{n\in\mathbb{N}}$ is bounded. Since the operators $(J_{\gamma_{i,c_i(n)}A_i})_{n\in\mathbb{N}}$ are nonexpansive [8, Corollary 23.8], it follows from (3.3) that $(a_{i,n})_{n\in\mathbb{N}}$ is bounded, and hence that $(a_{i,n}^*)_{n\in\mathbb{N}}$ is also bounded. Likewise, for every $k \in K$, $(\sum_{i\in I} L_{ki}x_{i,d_k(n)} + \mu_{k,d_k(n)}v_{k,d_k(n)}^*)_{n\in\mathbb{N}}$ is bounded and we deduce from (3.3) that $(b_{k,n})_{n\in\mathbb{N}}$ and $(b_{k,n}^*)_{n\in\mathbb{N}}$ are bounded. In view of (3.4), this establishes the claim.

(iii): For every every $i \in I$ and every $n \in \mathbb{N}$, define $\bar{\ell}_i(n)$ as the most recent iteration at which a new point in the graph of $A_i$ was incorporated into the algorithm, that is,

$$(\forall i \in I)(\forall n \in \mathbb{N}) \quad \bar{\ell}_i(n) = \max\big\{j \in S_i \;\big|\; j \leqslant n\big\}, \quad \text{where} \quad S_i = \big\{j \in \mathbb{N} \;\big|\; i \in I_j\big\}. \tag{3.6}$$

Note that (3.3) implies that

$$(\forall i \in I)(\forall n \in \mathbb{N}) \quad (a_{i,n}, a_{i,n}^*) = \big(a_{i,\bar{\ell}_i(n)}, a_{i,\bar{\ell}_i(n)}^*\big). \tag{3.7}$$

For every $i \in I$, (3.1) yields $\sup_{n\in\mathbb{N}}(n - \bar{\ell}_i(n)) \leqslant M$ and hence $\lim_{n\to+\infty} \bar{\ell}_i(n) = +\infty$. Next, we define

$$(\forall i \in I)(\forall n \in \mathbb{N}) \quad \ell_i(n) = c_i\big(\bar{\ell}_i(n)\big). \tag{3.8}$$

Thus, $\ell_i(n)$ is the iteration from which the computation of the most recent point in the graph of $A_i$ was initiated. It follows from (3.2) that

$$(\forall i \in I)(\forall n \in \mathbb{N}) \quad n - \ell_i(n) = n - \bar{\ell}_i(n) + \bar{\ell}_i(n) - \ell_i(n) \leqslant M + D. \tag{3.9}$$

Hence, $(\forall i \in I)$ $\lim_{n\to+\infty} \ell_i(n) = +\infty$. Since $\max_{i\in I} \sum_{j\in\mathbb{N}} \|x_{i,j+1} - x_{i,j}\|^2 < +\infty$ by (iii)(d), we deduce that

$$\begin{aligned}(\forall i \in I) \quad \|x_{i,n} - x_{i,\ell_i(n)}\|^2 &\leqslant \left(\sum_{j=\ell_i(n)}^{\ell_i(n)+M+D-1} \|x_{i,j+1} - x_{i,j}\|\right)^2 \\ &\leqslant (M + D) \sum_{j=\ell_i(n)}^{\ell_i(n)+M+D-1} \|x_{i,j+1} - x_{i,j}\|^2 \\ &\leqslant (M + D) \sum_{j=\ell_i(n)}^{+\infty} \|x_{i,j+1} - x_{i,j}\|^2 \\ &\to 0. \end{aligned} \tag{3.10}$$

Likewise, since (iii)(d) asserts that $\max_{k\in K} \sum_{j\in\mathbb{N}} \|v_{k,j+1}^* - v_{k,j}^*\|^2 < +\infty$, we have

$$(\forall i \in I)(\forall k \in K) \quad \|v_{k,n}^* - v_{k,\ell_i(n)}^*\|^2 \leqslant (M + D) \sum_{j=\ell_i(n)}^{+\infty} \|v_{k,j+1}^* - v_{k,j}^*\|^2 \to 0. \tag{3.11}$$

Next, let us set

$$(\forall i\in I)(\forall n\in\mathbb{N})\quad \begin{cases}\phi_{i,n}=\Big\langle x_{i,n}-a_{i,n}\;\Big|\; a^*_{i,n}+\sum_{k\in K}L^*_{ki}v^*_{k,n}\Big\rangle\\ \widetilde{\phi}_{i,n}=\Big\langle x_{i,\ell_i(n)}-a_{i,\bar{\ell}_i(n)}\;\Big|\; a^*_{i,\bar{\ell}_i(n)}+\sum_{k\in K}L^*_{ki}v^*_{k,\ell_i(n)}\Big\rangle.\end{cases}\tag{3.12}$$

Then it follows from (3.7), (a), (ii), (3.10), and (3.11) that

$$\begin{aligned}(\forall i\in I)\quad \phi_{i,n}-\widetilde{\phi}_{i,n}&=\Big\langle x_{i,n}-a_{i,n}\;\Big|\; a^*_{i,n}+\sum_{k\in K}L^*_{ki}v^*_{k,n}\Big\rangle\\ &\quad-\Big\langle x_{i,\ell_i(n)}-a_{i,n}\;\Big|\; a^*_{i,n}+\sum_{k\in K}L^*_{ki}v^*_{k,\ell_i(n)}\Big\rangle\\ &=\Big\langle x_{i,n}-a_{i,n}\;\Big|\;\sum_{k\in K}L^*_{ki}(v^*_{k,n}-v^*_{k,\ell_i(n)})\Big\rangle\\ &\quad+\Big\langle x_{i,n}-x_{i,\ell_i(n)}\;\Big|\; a^*_{i,n}+\sum_{k\in K}L^*_{ki}v^*_{k,\ell_i(n)}\Big\rangle\\ &\leqslant\bigg(\sum_{k\in K}\|L_{ki}\|\sup_{j\in\mathbb{N}}\big(\|x_{i,j}\|+\|a_{i,j}\|\big)\bigg)\|v^*_{k,n}-v^*_{k,\ell_i(n)}\|\\ &\quad+\bigg(\sup_{j\in\mathbb{N}}\|a^*_{i,j}\|+\sum_{k\in K}\|L_{ki}\|\sup_{j\in\mathbb{N}}\|v^*_{k,j}\|\bigg)\|x_{i,n}-x_{i,\ell_i(n)}\|\\ &\to 0.\end{aligned}\tag{3.13}$$

We also derive from (3.12), (3.7), and (3.3) that

$$\begin{aligned}(\forall i\in I)(\forall n\in\mathbb{N})\quad \widetilde{\phi}_{i,n}&=\Big\langle x_{i,\ell_i(n)}-a_{i,n}\;\Big|\; a^*_{i,n}+\sum_{k\in K}L^*_{ki}v^*_{k,\ell_i(n)}\Big\rangle\\ &=\gamma^{-1}_{i,\ell_i(n)}\|x_{i,\ell_i(n)}-a_{i,n}\|^2\\ &=\gamma_{i,\ell_i(n)}\Big\|a^*_{i,n}+\sum_{k\in K}L^*_{ki}v^*_{k,\ell_i(n)}\Big\|^2,\end{aligned}\tag{3.14}$$

which yields

$$(\forall i\in I)(\forall n\in\mathbb{N})\quad \widetilde{\phi}_{i,n}=\gamma^{-1}_{i,\ell_i(n)}\|x_{i,\ell_i(n)}-a_{i,n}\|^2\geqslant\varepsilon\|x_{i,\ell_i(n)}-a_{i,n}\|^2\tag{3.15}$$

and

$$(\forall i\in I)(\forall n\in\mathbb{N})\quad \widetilde{\phi}_{i,n}=\gamma_{i,\ell_i(n)}\Big\|a^*_{i,n}+\sum_{k\in K}L^*_{ki}v^*_{k,\ell_i(n)}\Big\|^2\geqslant\varepsilon\Big\|a^*_{i,n}+\sum_{k\in K}L^*_{ki}v^*_{k,\ell_i(n)}\Big\|^2.\tag{3.16}$$

It follows from (3.15) that

$$\begin{aligned}(\forall i\in I)(\forall n\in\mathbb{N})\quad \|x_{i,n}-a_{i,n}\|^2&\leqslant 2\big(\|x_{i,n}-x_{i,\ell_i(n)}\|^2+\|x_{i,\ell_i(n)}-a_{i,n}\|^2\big)\\ &\leqslant 2\big(\|x_{i,n}-x_{i,\ell_i(n)}\|^2+\varepsilon^{-1}(\widetilde{\phi}_{i,n}-\phi_{i,n})+\varepsilon^{-1}\phi_{i,n}\big)\end{aligned}\tag{3.17}$$

and from (3.16) that

$$
\begin{aligned}
(\forall i\in I)(\forall n\in\mathbb{N})\quad & \bigg\|a^*_{i,n}+\sum_{k\in K}L^*_{ki}v^*_{k,n}\bigg\|^2\\
&\leqslant \bigg\|a^*_{i,n}+\sum_{k\in K}L^*_{ki}v^*_{k,\ell_i(n)}-\sum_{k\in K}L^*_{ki}(v^*_{k,\ell_i(n)}-v^*_{k,n})\bigg\|^2\\
&\leqslant 2\Bigg(\bigg\|a^*_{i,n}+\sum_{k\in K}L^*_{ki}v^*_{k,\ell_i(n)}\bigg\|^2+\sum_{k\in K}\|L_{ki}\|^2\,\|v^*_{k,\ell_i(n)}-v^*_{k,n}\|^2\Bigg)\\
&\leqslant 2\Bigg(\varepsilon^{-1}(\widetilde{\phi}_{i,n}-\phi_{i,n})+\varepsilon^{-1}\phi_{i,n}+\sum_{k\in K}\|L_{ki}\|^2\,\|v^*_{k,\ell_i(n)}-v^*_{k,n}\|^2\Bigg).
\end{aligned}
\tag{3.18}
$$

We now perform a similar analysis for the operators $(B_k)_{k\in K}$. Much as in (3.6), for every $k\in K$ and every $n\in\mathbb{N}$, define $\bar{\vartheta}_k(n)$ as the most recent iteration at which a new point in the graph of $B_k$ was incorporated into the algorithm, that is,

$$
(\forall k\in K)(\forall n\in\mathbb{N})\quad \bar{\vartheta}_k(n)=\max\big\{j\in T_k \;\big|\; j\leqslant n\big\},\quad \text{where}\quad T_k=\big\{j\in\mathbb{N}\;\big|\;k\in K_j\big\},
\tag{3.19}
$$

and observe that

$$
(\forall k\in K)(\forall n\in\mathbb{N})\quad (b_{k,n},b^*_{k,n})=\big(b_{k,\bar{\vartheta}_k(n)},b^*_{k,\bar{\vartheta}_k(n)}\big).
\tag{3.20}
$$

Next, we define

$$
(\forall k\in K)(\forall n\in\mathbb{N})\quad \vartheta_k(n)=d_k\big(\bar{\vartheta}_k(n)\big).
\tag{3.21}
$$

Then, we derive from (3.1) and (3.2) that

$$
(\forall k\in K)(\forall n\in\mathbb{N})\quad n-\vartheta_k(n)=n-\bar{\vartheta}_k(n)+\bar{\vartheta}_k(n)-\vartheta_k(n)\leqslant M+D
\tag{3.22}
$$

and therefore that $(\forall k\in K)\ \lim_{k\to+\infty}\vartheta_k(n)=+\infty$. Since $\max_{i\in I}\sum_{j\in\mathbb{N}}\|x_{i,j+1}-x_{i,j}\|^2<+\infty$ by (iii)(d), we then deduce that

$$
(\forall k\in K)\quad \|x_{i,n}-x_{i,\vartheta_k(n)}\|^2\leqslant (M+D)\sum_{j=\vartheta_k(n)}^{+\infty}\|x_{i,j+1}-x_{i,j}\|^2\to 0.
\tag{3.23}
$$

Similarly since, $\max_{k\in K}\sum_{j\in\mathbb{N}}\|v^*_{k,j+1}-v^*_{k,j}\|^2<+\infty$, we have

$$
(\forall k\in K)\quad \|v^*_{k,n}-v^*_{k,\vartheta_k(n)}\|^2\leqslant (M+D)\sum_{j=\vartheta_k(n)}^{+\infty}\|v^*_{k,j+1}-v^*_{k,j}\|^2\to 0.
\tag{3.24}
$$

Now, let us set

$$
(\forall k\in K)(\forall n\in\mathbb{N})\quad
\begin{cases}
\psi_{k,n}=\Big\langle \sum_{i\in I}L_{ki}x_{i,n}-b_{k,n}\;\Big|\;b^*_{k,n}-v^*_{k,n}\Big\rangle\\
\widetilde{\psi}_{k,n}=\Big\langle \sum_{i\in I}L_{ki}x_{i,\vartheta_k(n)}-b_{k,\bar{\vartheta}_k(n)}\;\Big|\;b^*_{k,\bar{\vartheta}_k(n)}-v^*_{k,\vartheta_k(n)}\Big\rangle.
\end{cases}
\tag{3.25}
$$

Then it follows from (3.20), (a), (ii), (3.23), and (3.24) that

$$
\begin{aligned}
(\forall k \in K) \quad \psi_{k,n} - \widetilde{\psi}_{k,n} &= \Big\langle \sum_{i\in I} L_{ki} x_{i,n} - b_{k,n} \;\Big|\; b^*_{k,n} - v^*_{k,n} \Big\rangle \\
&\quad - \Big\langle \sum_{i\in I} L_{ki} x_{i,\vartheta_k(n)} - b_{k,n} \;\Big|\; b^*_{k,n} - v^*_{k,\vartheta_k(n)} \Big\rangle \\
&= \Big\langle \sum_{i\in I} L_{ki}(x_{i,n} - x_{i,\vartheta_k(n)}) \;\Big|\; b^*_{k,n} - v^*_{k,n} \Big\rangle \\
&\quad + \Big\langle \sum_{i\in I} L_{ki} x_{i,\vartheta_k(n)} - b_{k,n} \;\Big|\; v^*_{k,\vartheta_k(n)} - v^*_{k,n} \Big\rangle \\
&\leqslant \bigg( \sum_{i\in I} \|L_{ki}\| \sup_{j\in\mathbb{N}} \big(\|b^*_{k,j}\| + \|v^*_{k,j}\|\big) \bigg) \|x_{i,n} - x_{i,\vartheta_k(n)}\| \\
&\quad + \sup_{j\in\mathbb{N}} \bigg( \sum_{i\in I} \|L_{ki}\| \, \|x_{i,\vartheta_k(j)}\| + \|b_{k,j}\| \bigg) \|v^*_{k,n} - v^*_{k,\vartheta_k(n)}\| \\
&\to 0.
\end{aligned} \tag{3.26}
$$

In addition, (3.20) and (3.3) yield

$$
\begin{aligned}
(\forall k \in K)(\forall n \in \mathbb{N}) \quad \widetilde{\psi}_{k,n} &= \Big\langle \sum_{i\in I} L_{ki} x_{i,\vartheta_k(n)} - b_{k,n} \;\Big|\; b^*_{k,n} - v^*_{k,\vartheta_k(n)} \Big\rangle \\
&= \mu^{-1}_{k,\vartheta_k(n)} \Big\| \sum_{i\in I} L_{ki} x_{i,\vartheta_k(n)} - b_{k,n} \Big\|^2 \\
&= \mu_{k,\vartheta_k(n)} \|b^*_{k,n} - v^*_{k,\vartheta_k(n)}\|^2.
\end{aligned} \tag{3.27}
$$

Consequently,

$$
(\forall k \in K)(\forall n \in \mathbb{N}) \quad \widetilde{\psi}_{k,n} = \frac{\Big\| \sum_{i\in I} L_{ki} x_{i,\vartheta_k(n)} - b_{k,\vartheta_k(n)} \Big\|^2}{\mu_{k,\vartheta_k(n)}} \geqslant \varepsilon \Big\| \sum_{i\in I} L_{ki} x_{i,\vartheta_k(n)} - b_{k,\vartheta_k(n)} \Big\|^2 \tag{3.28}
$$

and

$$
(\forall k \in K)(\forall n \in \mathbb{N}) \quad \widetilde{\psi}_{k,n} = \mu_{k,\vartheta_k(n)} \|b^*_{k,n} - v^*_{k,\vartheta_k(n)}\|^2 \geqslant \varepsilon \|b^*_{k,n} - v^*_{k,\vartheta_k(n)}\|^2. \tag{3.29}
$$

It follows from (3.28) that

$$
\begin{aligned}
(\forall k \in K)(\forall n \in \mathbb{N}) \quad & \Big\| \sum_{i\in I} L_{ki} x_{i,n} - b_{k,n} \Big\|^2 \\
&= \Big\| \sum_{i\in I} L_{ki}(x_{i,n} - x_{i,\vartheta_k(n)}) + \sum_{i\in I} L_{ki} x_{i,\vartheta_k(n)} - b_{k,n} \Big\|^2 \\
&\leqslant 2 \bigg( \sum_{i\in I} \|L_{ki}\|^2 \, \|x_{i,n} - x_{i,\vartheta_k(n)}\|^2 + \Big\| \sum_{i\in I} L_{ki} x_{i,\vartheta_k(n)} - b_{k,n} \Big\|^2 \bigg) \\
&\leqslant 2 \bigg( \sum_{i\in I} \|L_{ki}\|^2 \, \|x_{i,n} - x_{i,\vartheta_k(n)}\|^2 + \varepsilon^{-1}(\widetilde{\psi}_{k,n} - \psi_{k,n}) + \varepsilon^{-1}\psi_{k,n} \bigg),
\end{aligned} \tag{3.30}
$$

and from (3.29) that

$$(\forall k\in K)(\forall n\in\mathbb{N})\quad \|b^*_{k,n}-v^*_{k,n}\|^2\leqslant 2\big(\|b^*_{k,n}-v^*_{k,\vartheta_k(n)}\|^2+\|v^*_{k,n}-v^*_{k,\vartheta_k(n)}\|^2\big)$$
$$\leqslant 2\big(\varepsilon^{-1}(\widetilde{\psi}_{k,n}-\psi_{k,n})+\varepsilon^{-1}\psi_{k,n}+\|v^*_{k,n}-v^*_{k,\vartheta_k(n)}\|^2\big). \tag{3.31}$$

On the one hand, we derive from (3.5), (3.17), and (3.31) that

$$\begin{aligned}(\forall n\in\mathbb{N})\quad \|x_n-a_n\|^2+\|b^*_n-v^*_n\|^2&=\sum_{i\in I}\|x_{i,n}-a_{i,n}\|^2+\sum_{k\in K}\|b^*_{k,n}-v^*_{k,n}\|^2\\
&\leqslant 2\sum_{i\in I}\|x_{i,n}-x_{i,\ell_i(n)}\|^2+2\sum_{k\in K}\|v^*_{k,n}-v^*_{k,\vartheta_k(n)}\|^2\\
&\quad+2\varepsilon^{-1}\sum_{i\in I}(\widetilde{\phi}_{i,n}-\phi_{i,n})+2\varepsilon^{-1}\sum_{k\in K}(\widetilde{\psi}_{k,n}-\psi_{k,n})\\
&\quad+2\varepsilon^{-1}\Bigg(\sum_{i\in I}\phi_{i,n}+\sum_{k\in K}\psi_{k,n}\Bigg).\end{aligned}\tag{3.32}$$

On the other hand, we derive from (3.5), (3.18), and (3.30) that

$$\begin{aligned}(\forall n\in\mathbb{N})\quad &\|a^*_n+L^*v^*_n\|^2+\|Lx_n-b_n\|^2\\
&=\sum_{i\in I}\Bigg\|a^*_{i,n}+\sum_{k\in K}L^*_{ki}v^*_{k,n}\Bigg\|^2+\sum_{k\in K}\Bigg\|\sum_{i\in I}L_{ki}x_{i,n}-b_{k,n}\Bigg\|^2\\
&\leqslant 2\varepsilon^{-1}\sum_{i\in I}(\widetilde{\phi}_{i,n}-\phi_{i,n})+2\sum_{i\in I}\sum_{k\in K}\|L_{ki}\|^2\,\|v^*_{k,\ell_i(n)}-v^*_{k,n}\|^2\\
&\quad+2\sum_{k\in K}\sum_{i\in I}\|L_{ki}\|^2\,\|x_{i,n}-x_{i,\vartheta_k(n)}\|^2+2\varepsilon^{-1}\sum_{k\in K}(\widetilde{\psi}_{k,n}-\psi_{k,n})\\
&\quad+2\varepsilon^{-1}\Bigg(\sum_{i\in I}\phi_{i,n}+\sum_{k\in K}\psi_{k,n}\Bigg).\end{aligned}\tag{3.33}$$

We deduce from (3.12), (3.25), (3.4), (3.5), and (iii)(e) that

$$\varlimsup\Bigg(\sum_{i\in I}\phi_{i,n}+\sum_{k\in K}\psi_{k,n}\Bigg)=\varlimsup\big(\langle x_n-a_n\mid a^*_n+L^*v^*_n\rangle+\langle Lx_n-b_n\mid b^*_n-v^*_n\rangle\big)\leqslant 0. \tag{3.34}$$

Altogether, taking the limit superior in (3.32) and (3.33), and using (3.10), (3.24), (3.13), (3.26), (3.11), (3.23), and (3.34), we obtain

$$x_n-a_n\to 0,\quad a^*_n+L^*v^*_n\to 0,\quad Lx_n-b_n\to 0,\quad\text{and}\quad b^*_n-v^*_n\to 0. \tag{3.35}$$

Now set $x=(x_i)_{i\in I}$ and $v^*=(v^*_k)_{k\in K}$. Then (iii)(f) and (3.35) yield

$$a_{q_n}\rightharpoonup x,\quad b^*_{q_n}\rightharpoonup v^*,\quad a^*_{q_n}+L^*b^*_{q_n}\to 0,\quad\text{and}\quad La_{q_n}-b_{q_n}\to 0. \tag{3.36}$$

In turn, (i) and Proposition 2.2(iv) imply that $(x,v^*)\in\boldsymbol{Z}$. ☐

**Remark 3.3** In (3.3), the resolvents are assumed to be computed exactly to simplify the presentation. However, it is possible to allow for relative errors in these computations in the spirit of [20, Algorithm 3]. More precisely, we can replace the calculation

$$(a_{i,n},a^*_{i,n})=\Big(J_{\gamma_{i,c_i(n)}A_i}\big(x_{i,c_i(n)}+\gamma_{i,c_i(n)}(z^*_i-l^*_{i,n})\big),\gamma^{-1}_{i,c_i(n)}(x_{i,c_i(n)}-a_{i,n})-l^*_{i,n}\Big) \tag{3.37}$$

by any choice of $(a_{i,n}, a^*_{i,n}) \in \mathcal{H}_i^2$ such that

$$(a_{i,n}, z^*_i + a^*_{i,n}) \in \text{gra}\, A_i \quad \text{and} \quad a_{i,n} + \gamma_{i,c_i(n)} a^*_{i,n} = x_{i,c_i(n)} - \gamma_{i,c_i(n)} l^*_{i,n} + e_{i,n}, \tag{3.38}$$

where the error $e_{i,n}$ satisfies

$$\begin{cases} \|e_{i,n}\| \leqslant \beta \\ \langle e_{i,n} \mid a^*_{i,n} + l^*_{i,n} \rangle \leqslant \sigma \gamma_{i,c_i(n)} \|a^*_{i,n} + l^*_{i,n}\|^2 \\ \langle x_{i,c_i(n)} - a_{i,n} \mid e_{i,n} \rangle \geqslant -\sigma \|x_{i,c_i(n)} - a_{i,n}\|^2 \end{cases} \tag{3.39}$$

for some constants $\beta \in \left]0, +\infty\right[$ and $\sigma \in \left]0, 1\right[$ that are independent of $i$ and $n$. It follows from [8, Proposition 23.21] that (3.38) can also be written as

$$(a_{i,n}, a^*_{i,n}) = \Big( J_{\gamma_{i,c_i(n)} A_i}\big(x_{i,c_i(n)} + \gamma_{i,c_i(n)}(z^*_i - l^*_{i,n}) + e_{i,n}\big), \gamma^{-1}_{i,c_i(n)}(x_{i,c_i(n)} - a_{i,n} + e_{i,n}) - l^*_{i,n} \Big). \tag{3.40}$$

It may easily be seen that the calculations (3.37) satisfy (3.38) with $e_{i,n} = 0$, trivially fulfilling (3.39). In the setting of (3.39), (3.15) becomes

$$\begin{aligned} (\forall i \in I)(\forall n \in \mathbb{N}) \quad \widetilde{\phi}_{i,n} &= \gamma^{-1}_{i,\ell_i(n)} \big( \|x_{i,\ell_i(n)} - a_{i,n}\|^2 + \langle x_{i,\ell_i(n)} - a_{i,n} \mid e_{i,n} \rangle \big) \\ &\geqslant \varepsilon(1-\sigma) \|x_{i,\ell_i(n)} - a_{i,n}\|^2 \end{aligned} \tag{3.41}$$

and (3.16) becomes

$$\begin{aligned} (\forall i \in I)(\forall n \in \mathbb{N}) \quad \widetilde{\phi}_{i,n} &= \gamma_{i,\ell_i(n)} \Big\| a^*_{i,n} + \sum_{k \in K} L^*_{ki} v^*_{k,\ell_i(n)} \Big\|^2 - \Big\langle e_{i,n} \Big| a^*_{i,n} + \sum_{k \in K} L^*_{ki} v^*_{k,\ell_i(n)} \Big\rangle \\ &\geqslant \varepsilon(1-\sigma) \Big\| a^*_{i,n} + \sum_{k \in K} L^*_{ki} v^*_{k,\ell_i(n)} \Big\|^2. \end{aligned} \tag{3.42}$$

Likewise, we can replace the calculation

$$(b_{k,n}, b^*_{k,n}) = \Big( r_k + J_{\mu_{k,d_k(n)} B_k}\big(l_{k,n} + \mu_{k,d_k(n)} v^*_{k,d_k(n)} - r_k\big), v^*_{k,d_k(n)} + \mu^{-1}_{k,d_k(n)}(l_{k,n} - b_{k,n}) \Big) \tag{3.43}$$

by any choice of $(b_{k,n}, b^*_{k,n}) \in \mathcal{G}_k^2$ such that

$$(b_{k,n} - r_k, b^*_{k,n}) \in \text{gra}\, B_k \quad \text{and} \quad b_{k,n} + \mu_{k,d_k(n)} b^*_{k,n} = l_{k,n} + \mu_{k,d_k(n)} v^*_{k,d_k(n)} + f_{k,n}, \tag{3.44}$$

where the error $f_{k,n}$ satisfies

$$\begin{cases} \|f_{k,n}\| \leqslant \delta \\ \langle l_{k,n} - b_{k,n} \mid f_{k,n} \rangle \geqslant -\zeta \|l_{k,n} - b_{k,n}\|^2 \\ \langle f_{k,n} \mid b^*_{k,n} - v^*_{k,d_k(n)} \rangle \leqslant \zeta \mu_{k,d_k(n)} \|b^*_{k,n} - v^*_{k,d_k(n)}\|^2 \end{cases} \tag{3.45}$$

for some constants $\delta \in \left]0, +\infty\right[$ and $\zeta \in \left]0, 1\right[$ that are independent of $k$ and $n$. Altogether, the effect of such approximate resolvent evaluations is to replace $\varepsilon^{-1}$ by $\varepsilon^{-1}(1-\sigma)^{-1}$ or $\varepsilon^{-1}(1-\zeta)^{-1}$ in (3.32)–(3.33), with the remainder of the proof of Proposition 3.2 remaining unchanged.

### 3.2 A weakly convergent algorithm for finding a Kuhn-Tucker point

We propose a Fejér monotone primal-dual algorithm based on the results of Section 2.2 to find a point in the Kuhn-Tucker set (1.3).

**Algorithm 3.4** Consider the setting of Problem 1.1, let $\boldsymbol{\mathcal{K}}$ be a closed vector subspace of $\bigoplus_{i\in I}\mathcal{H}_i \oplus \bigoplus_{k\in K}\mathcal{G}_k$ such that $\boldsymbol{Z}\subset\boldsymbol{\mathcal{K}}$, and suppose that Assumption 3.1 is in force. Let $(\lambda_n)_{n\in\mathbb{N}}\in[\varepsilon,2-\varepsilon]^{\mathbb{N}}$, let $((x_{i,0})_{i\in I},(v^*_{k,0})_{k\in K})\in\boldsymbol{\mathcal{K}}$, and iterate

$$
\begin{array}{l}
\text{for } n=0,1,\ldots\\
\left\lfloor
\begin{array}{l}
\text{for every } i\in I_n\\
\left\lfloor
\begin{array}{l}
l^*_{i,n}=\sum_{k\in K}L^*_{ki}v^*_{k,c_i(n)}\\
(a_{i,n},a^*_{i,n})=\Big(J_{\gamma_{i,c_i(n)}A_i}\big(x_{i,c_i(n)}+\gamma_{i,c_i(n)}(z^*_i-l^*_{i,n})\big),\gamma^{-1}_{i,c_i(n)}(x_{i,c_i(n)}-a_{i,n})-l^*_{i,n}\Big)
\end{array}
\right.\\
\text{for every } i\in I\smallsetminus I_n\\
\lfloor\ (a_{i,n},a^*_{i,n})=(a_{i,n-1},a^*_{i,n-1})\\
\text{for every } k\in K_n\\
\left\lfloor
\begin{array}{l}
l_{k,n}=\sum_{i\in I}L_{ki}x_{i,d_k(n)}\\
(b_{k,n},b^*_{k,n})=\Big(r_k+J_{\mu_{k,d_k(n)}B_k}\big(l_{k,n}+\mu_{k,d_k(n)}v^*_{k,d_k(n)}-r_k\big),v^*_{k,d_k(n)}+\mu^{-1}_{k,d_k(n)}(l_{k,n}-b_{k,n})\Big)
\end{array}
\right.\\
\text{for every } k\in K\smallsetminus K_n\\
\lfloor\ (b_{k,n},b^*_{k,n})=(b_{k,n-1},b^*_{k,n-1})\\
\big((t^*_{i,n})_{i\in I},(t_{k,n})_{k\in K}\big)=P_{\boldsymbol{\mathcal{K}}}\big((a^*_{i,n}+\sum_{k\in K}L^*_{ki}b^*_{k,n})_{i\in I},(b_{k,n}-\sum_{i\in I}L_{ki}a_{i,n})_{k\in K}\big)\\
\tau_n=\sum_{i\in I}\|t^*_{i,n}\|^2+\sum_{k\in K}\|t_{k,n}\|^2\\
\text{if } \tau_n>0\\
\left\lfloor\ \theta_n=\dfrac{\lambda_n}{\tau_n}\max\left\{0,\displaystyle\sum_{i\in I}\big(\langle x_{i,n}\mid t^*_{i,n}\rangle-\langle a_{i,n}\mid a^*_{i,n}\rangle\big)+\sum_{k\in K}\big(\langle t_{k,n}\mid v^*_{k,n}\rangle-\langle b_{k,n}\mid b^*_{k,n}\rangle\big)\right\}\right.\\
\text{else } \theta_n=0\\
\text{for every } i\in I\\
\lfloor\ x_{i,n+1}=x_{i,n}-\theta_n t^*_{i,n}\\
\text{for every } k\in K\\
\lfloor\ v^*_{k,n+1}=v^*_{k,n}-\theta_n t_{k,n}.
\end{array}
\right.
\end{array}
\tag{3.46}
$$

**Remark 3.5** When Problem 2.1 has no special structure, on can take $\boldsymbol{\mathcal{K}}=\bigoplus_{i\in I}\mathcal{H}_i\oplus\bigoplus_{k\in K}\mathcal{G}_k$ in Algorithm 3.4. In other instances, it may be advantageous computationally to use a suitable proper subspace $\boldsymbol{\mathcal{K}}$. For instance, if $I=\{1\}$, $z^*_1=0$, and $A_1\colon\mathcal{H}_1\to\mathcal{H}_1$ is linear, then (1.3) reduces to

$$
\boldsymbol{Z}=\bigg\{\big(x_1,(v^*_k)_{k\in K}\big)\;\bigg|\; x_1\in\mathcal{H}_1,\ A_1x_1+\sum_{k\in K}L^*_{k1}v^*_k=0,\ \text{ and }\ (\forall k\in K)\ \ v^*_k\in\mathcal{G}_k\ \text{ and }\ L_{k1}x_1-r_k\in B_k^{-1}v^*_k\bigg\},\tag{3.47}
$$

and we can use

$$
\boldsymbol{\mathcal{K}}=\bigg\{\big(x_1,(v^*_k)_{k\in K}\big)\in\mathcal{H}_1\oplus\bigoplus_{k\in K}\mathcal{G}_k\;\bigg|\;A_1x_1+\sum_{k\in K}L^*_{k1}v^*_k=0\bigg\}.\tag{3.48}
$$

In effect, this approach was adopted in [20] in the further special case in which $A_1 = 0$ and $(\forall k \in K)$ $\mathcal{G}_k = \mathcal{H}_1$, $L_{k1} = \mathrm{Id}$, and $r_k = 0$.

**Theorem 3.6** *Consider the setting of Problem 1.1 and Algorithm 3.4, suppose that $\mathscr{P} \neq \varnothing$, and let*

$$(\forall n \in \mathbb{N}) \quad x_n = (x_{i,n})_{i\in I} \quad \textit{and} \quad v_n^* = (v_{k,n}^*)_{k\in K}. \tag{3.49}$$

*Then $(x_n)_{n\in\mathbb{N}}$ converges weakly to a point $\overline{x} \in \mathscr{P}$, $(v_n^*)_{n\in\mathbb{N}}$ converges weakly to a point $\overline{v} \in \mathscr{D}$, and $(\overline{x}, \overline{v}^*) \in \boldsymbol{Z}$.*

*Proof*. Define $\mathcal{H}$, $\mathcal{G}$, $L$, $A$, and $B$ as in (2.7), and $(a_n)_{n\in\mathbb{N}}$, $(a_n^*)_{n\in\mathbb{N}}$, $(b_n)_{n\in\mathbb{N}}$, and $(b_n^*)_{n\in\mathbb{N}}$ as in (3.4). Further, define $(\forall n \in \mathbb{N})$ $t_n = (t_{k,n})_{k\in K}$ and $t_n^* = (t_{i,n}^*)_{i\in I}$. It follows from (3.46), (3.49), (2.8), and Proposition 3.2(i) that Algorithm 3.4 is a special case of Algorithm 2.6. Hence, upon invoking Proposition 2.4, we can apply the results of Proposition 2.7 in this setting. First, Proposition 2.7(i) implies that the boundedness assumption (a) in Proposition 3.2 is satisfied. Second, in view of (3.46), the sequence $(a_n, a_n^*)_{n\in\mathbb{N}}$ and $(b_n, b_n^*)_{n\in\mathbb{N}}$ are constructed according to assumption (c) in Proposition 3.2. We thus derive from Proposition 3.2(ii) that

$$(a_n)_{n\in\mathbb{N}}, \quad (a_n^*)_{n\in\mathbb{N}}, \quad (b_n)_{n\in\mathbb{N}}, \quad \text{and} \quad (b_n^*)_{n\in\mathbb{N}} \quad \text{are bounded.} \tag{3.50}$$

Furthermore, the summability assumption (iii)(d) in Proposition 3.2 is secured by Proposition 2.7(ii), while the limit superior assumption (iii)(e) in Proposition 3.2 holds by Proposition 2.7(iii). We therefore use Proposition 3.2(iii) to conclude by applying Proposition 2.7(iv). To this end, take $(x, v^*) \in \boldsymbol{\mathcal{K}}$ and a strictly increasing sequence $(q_n)_{n\in\mathbb{N}}$ in $\mathbb{N}$ such that $x_{q_n} \rightharpoonup x$ and $v_{q_n}^* \rightharpoonup v^*$. Then Proposition 3.2(iii) assert that $(x, v^*) \in \boldsymbol{Z}$. Thus, (2.12) is satisfied and the proof is complete. ☐

**Remark 3.7** Theorem 3.6 subsumes [1, Theorem 4.3], which required the following additional assumptions: the implementation is synchronous, *i.e.*,

$$(\forall n \in \mathbb{N})(\forall i \in I)(\forall k \in K) \quad c_i(n) = d_k(n) = n, \tag{3.51}$$

no proper subspace is used, *i.e.*,

$$\boldsymbol{\mathcal{K}} = \bigoplus_{i\in I} \mathcal{H}_i \oplus \bigoplus_{k\in K} \mathcal{G}_k, \tag{3.52}$$

the control is fully parallel, *i.e.*,

$$(\forall n \in \mathbb{N}) \quad I_n = I \quad \text{and} \quad K_n = K, \tag{3.53}$$

and common proximal parameters are used in the sense that

$$(\forall n \in \mathbb{N})(\forall i \in I)(\forall k \in K) \quad \gamma_{i,n} = \gamma_n \quad \text{and} \quad \mu_{k,n} = \mu_n. \tag{3.54}$$

Therefore, the proposed method also subsumes [18] and [19, Proposition 3] (see also [5]), which are special cases of [1, Theorem 4.3]; see [1, Examples 3.7 and 3.8] for details.

**Remark 3.8** Theorem 3.6 is closely related to [20, Proposition 4.2] (see also [5]), which considers the special case of Problem 1.1 in which $I = \{1\}$, $z_1^* = 0$, $A_1 = 0$, and $(\forall k \in K)$ $\mathcal{G}_k = \mathcal{H}_1$ and $L_{k1} = \mathrm{Id}$. If in this case one sets

$$\boldsymbol{\mathcal{K}} = \left\{ \big(x_1, (v_k^*)_{k\in K}\big) \in \mathcal{H}_1^{p+1} \;\middle|\; \sum_{k\in K} v_k^* = 0 \right\} \tag{3.55}$$

in our algorithm, we recover the special case of the method of [20, Section 4] in which the parameter $\alpha_{ij}^k$ of [20, Proposition 4.2] is $1$ if $i = j$, and $0$ otherwise. Other settings of $\alpha_{ij}^k$ in [20] produce algorithms that are not special cases of our scheme, but must process the resolvent of every operator at every iteration and remain fully synchronous as in (3.51) and (3.53).

**Remark 3.9** Recall that the resolvent of the subdifferential of a proper lower semicontinuous convex function $f\colon \mathcal{H} \to \left]-\infty,+\infty\right]$ is Moreau's proximity operator $(\mathrm{Id} + \partial f)^{-1} = \mathrm{prox}_f\colon x \mapsto \mathrm{argmin}_{y\in\mathcal{H}}(f(y) + \|x-y\|^2/2)$ [8, 22]. Now consider the setting of Problem 1.2 and execute Algorithm 3.4 with $(\forall i \in I)\ A_i = \partial f_i$ and $(\forall k \in K)\ B_k = \partial g_k$. Then, using the same arguments as in [15, Proposition 5.4], it follows from Theorem 3.6 that $(x_n)_{n\in\mathbb{N}}$ converges weakly to a solution to (1.5) and that $(v_n^*)_{n\in\mathbb{N}}$ converges weakly to a solution to (1.6).

**Remark 3.10** The framework of [20, Algorithm 3] for solving (1.7) allows for relative errors in the computation of the resolvents. Similar errors may be incorporated in Algorithm 3.4 by adopting the approximate evaluation scheme of Remark 3.3 to select points in the graphs of the monotone operators in (3.46). Since Proposition 3.2 remains valid with such approximate resolvent computations, so does Theorem 3.6.

### 3.3 A best approximation result

In this section we use the abstract Haugazeau-like algorithm of Section 2.3 to devise a strongly convergent asynchronous block-iterative method to construct the best approximation to a reference point from the Kuhn-Tucker set (1.3).

**Algorithm 3.11** Consider the setting of Problem 1.1, let $\boldsymbol{\mathcal{K}}$ be a closed vector subspace of $\bigoplus_{i\in I}\mathcal{H}_i \oplus \bigoplus_{k\in K}\mathcal{G}_k$ such that $\boldsymbol{Z} \subset \boldsymbol{\mathcal{K}}$, and suppose that Assumption 3.1 is in force. Let $(\lambda_n)_{n\in\mathbb{N}} \in [\varepsilon, 1]^{\mathbb{N}}$, let $((x_{i,0})_{i\in I}, (v_{k,0}^*)_{k\in K}) \in \boldsymbol{\mathcal{K}}$, and iterate

$$
\begin{array}{l}
\text{for } n = 0, 1, \ldots \\
\quad \text{for every } i \in I_n \\
\quad\quad l_{i,n}^* = \sum_{k \in K} L_{ki}^* v_{k,c_i(n)}^* \\
\quad\quad (a_{i,n}, a_{i,n}^*) = \Big( J_{\gamma_{i,c_i(n)} A_i}\big(x_{i,c_i(n)} + \gamma_{i,c_i(n)}(z_i^* - l_{i,n}^*)\big), \gamma_{i,c_i(n)}^{-1}(x_{i,c_i(n)} - a_{i,n}) - l_{i,n}^* \Big) \\
\quad \text{for every } i \in I \smallsetminus I_n \\
\quad\quad (a_{i,n}, a_{i,n}^*) = (a_{i,n-1}, a_{i,n-1}^*) \\
\quad \text{for every } k \in K_n \\
\quad\quad l_{k,n} = \sum_{i \in I} L_{ki} x_{i,d_k(n)} \\
\quad\quad (b_{k,n}, b_{k,n}^*) = \Big( r_k + J_{\mu_{k,d_k(n)} B_k}\big(l_{k,n} + \mu_{k,d_k(n)} v_{k,d_k(n)}^* - r_k\big), v_{k,d_k(n)}^* + \mu_{k,d_k(n)}^{-1}(l_{k,n} - b_{k,n}) \Big) \\
\quad \text{for every } k \in K \smallsetminus K_n \\
\quad\quad (b_{k,n}, b_{k,n}^*) = (b_{k,n-1}, b_{k,n-1}^*) \\
\quad \big((t_{i,n}^*)_{i \in I}, (t_{k,n})_{k \in K}\big) = P_{\boldsymbol{\mathcal{K}}}\big((a_{i,n}^* + \sum_{k \in K} L_{ki}^* b_{k,n}^*)_{i \in I}, (b_{k,n} - \sum_{i \in I} L_{ki} a_{i,n})_{k \in K}\big) \\
\quad \tau_n = \sum_{i \in I} \|t_{i,n}^*\|^2 + \sum_{k \in K} \|t_{k,n}\|^2 \\
\quad \text{if } \tau_n > 0 \\
\quad\quad \theta_n = \dfrac{\lambda_n}{\tau_n} \max\left\{ 0, \sum_{i \in I} \big(\langle x_{i,n} \mid t_{i,n}^* \rangle - \langle a_{i,n} \mid a_{i,n}^* \rangle\big) + \sum_{k \in K} \big(\langle t_{k,n} \mid v_{k,n}^* \rangle - \langle b_{k,n} \mid b_{k,n}^* \rangle\big) \right\} \\
\quad \text{else } \theta_n = 0 \\
\quad \text{for every } i \in I \\
\quad\quad x_{i,n+1/2} = x_{i,n} - \theta_n t_{i,n}^* \\
\quad \text{for every } k \in K \\
\quad\quad v_{k,n+1/2}^* = v_{k,n}^* - \theta_n t_{k,n} \\
\quad \chi_n = \sum_{i \in I} \langle x_{i,0} - x_{i,n} \mid x_{i,n} - x_{i,n+1/2} \rangle + \sum_{k \in K} \langle v_{k,0}^* - v_{k,n}^* \mid v_{k,n}^* - v_{k,n+1/2}^* \rangle \\
\quad \mu_n = \sum_{i \in I} \|x_{i,0} - x_{i,n}\|^2 + \sum_{k \in K} \|v_{k,0}^* - v_{k,n}^*\|^2 \\
\quad \nu_n = \sum_{i \in I} \|x_{i,n} - x_{i,n+1/2}\|^2 + \sum_{k \in K} \|v_{k,n}^* - v_{k,n+1/2}^*\|^2 \\
\quad \rho_n = \mu_n \nu_n - \chi_n^2 \\
\quad \text{if } \rho_n = 0 \text{ and } \chi_n \geqslant 0 \\
\quad\quad \text{for every } i \in I \\
\quad\quad\quad x_{i,n+1} = x_{i,n+1/2} \\
\quad\quad \text{for every } k \in K \\
\quad\quad\quad v_{k,n+1}^* = v_{k,n+1/2}^* \\
\quad \text{if } \rho_n > 0 \text{ and } \chi_n \nu_n \geqslant \rho_n \\
\quad\quad \text{for every } i \in I \\
\quad\quad\quad x_{i,n+1} = x_{i,0} + (1 + \chi_n/\nu_n)(x_{i,n+1/2} - x_{i,n}) \\
\quad\quad \text{for every } k \in K \\
\quad\quad\quad v_{k,n+1}^* = v_{k,0}^* + (1 + \chi_n/\nu_n)(v_{k,n+1/2}^* - v_{k,n}^*) \\
\quad \text{if } \rho_n > 0 \text{ and } \chi_n \nu_n < \rho_n \\
\quad\quad \text{for every } i \in I \\
\quad\quad\quad x_{i,n+1} = x_{i,n} + (\nu_n/\rho_n)\big(\chi_n(x_{i,0} - x_{i,n}) + \mu_n(x_{i,n+1/2} - x_{i,n})\big) \\
\quad\quad \text{for every } k \in K \\
\quad\quad\quad v_{k,n+1}^* = v_{k,n}^* + (\nu_n/\rho_n)\big(\chi_n(v_{k,0}^* - v_{k,n}^*) + \mu_n(v_{k,n+1/2}^* - v_{k,n}^*)\big).
\end{array}
\tag{3.56}
$$

**Theorem 3.12** *Consider the setting of Problem 1.1 and Algorithm 3.11, and suppose that $\mathscr{P} \neq \varnothing$. Define*

$$(\forall n \in \mathbb{N}) \quad x_n = (x_{i,n})_{i \in I} \quad \text{and} \quad v_n^* = (v_{k,n}^*)_{k \in K} \tag{3.57}$$

*and set $(\overline{x}, \overline{v}^*) = P_{\boldsymbol{Z}}(x_0, v_0^*)$. Then $(x_n)_{n\in\mathbb{N}}$ converges strongly to $\overline{x} \in \mathscr{P}$ and $(v_n^*)_{n\in\mathbb{N}}$ converges strongly to $\overline{v}^* \in \mathscr{D}$.*

*Proof.* Define $\mathcal{H}$, $\mathcal{G}$, $L$, $A$, and $B$ as in (2.7), $(a_n)_{n\in\mathbb{N}}$, $(a_n^*)_{n\in\mathbb{N}}$, $(b_n)_{n\in\mathbb{N}}$, and $(b_n^*)_{n\in\mathbb{N}}$ as in (3.4), and set $(\forall n \in \mathbb{N})$ $t_n = (t_{k,n})_{k\in K}$ and $t_n^* = (t_{i,n}^*)_{i\in I}$. In view of (3.56), (3.57), (2.8), and Proposition 3.2(i), Algorithm 3.11 is an instance of Algorithm 2.10. Hence, upon invoking Proposition 2.4, we can apply the results of Proposition 2.12 in this setting. First, Proposition 2.12(i) implies that assumption (a) in Proposition 3.2 is satisfied. Second, in view of (3.56), assumption (c) in Proposition 3.2 is satisfied as well. Thus, Proposition 3.2(ii) asserts that the sequences $(a_n)_{n\in\mathbb{N}}$, $(a_n^*)_{n\in\mathbb{N}}$, $(b_n)_{n\in\mathbb{N}}$, and $(b_n^*)_{n\in\mathbb{N}}$ are bounded. Third, assumption (iii)(d) in Proposition 3.2 is secured by Proposition 2.12(ii). Finally, assumption (iii)(e) in Proposition 3.2 holds by Proposition 2.12(iv). We therefore use Proposition 3.2(iii) to conclude by invoking Proposition 2.12(v). Take $(x, v^*) \in \boldsymbol{\mathcal{K}}$ and a strictly increasing sequence $(q_n)_{n\in\mathbb{N}}$ in $\mathbb{N}$ such that $x_{q_n} \rightharpoonup x$ and $v_{q_n}^* \rightharpoonup v^*$. Then it follows from Proposition 3.2(iii) that $(x, v^*) \in \boldsymbol{Z}$, which completes the proof. ☐

**Remark 3.13** As in Remark 3.9, consider the setting of Problem 1.2 and execute Algorithm 3.11 with $(\forall i \in I)$ $A_i = \partial f_i$ and $(\forall k \in K)$ $B_k = \partial g_k$. Then Theorem 3.12 asserts that $(x_n)_{n\in\mathbb{N}}$ converges strongly to a solution $\overline{x}$ to (1.5) and that $(v_n^*)_{n\in\mathbb{N}}$ converges strongly to a solution $\overline{v}^*$ to (1.6) such that $(\overline{x}, \overline{v}^*)$ is the projection of $(x_0, v_0^*)$ onto the corresponding Kuhn-Tucker set (1.3).

**Remark 3.14** Theorem 3.12 improves upon [2, Proposition 4.2], which addresses the special case in which the algorithm is synchronous and the restrictions (3.51)–(3.54) are imposed. The latter was applied in the context of Remark 3.13 to domain decomposition methods in [4]; Theorem 3.12 provides a new range of ways to revisit such applications using asynchronous block-iterative calculations.

**Remark 3.15** By an argument similar to that of Remark 3.10, Theorem 3.12 remains valid if the resolvent computations in (3.56) are replaced by approximate evaluations meeting the conditions in Remark 3.3.

## References


[1] A. Alotaibi, P. L. Combettes, and N. Shahzad, Solving coupled composite monotone inclusions by successive Fejér approximations of their Kuhn-Tucker set, *SIAM J. Optim.*, vol. 24, pp. 2076–2095, 2014.

[2] A. Alotaibi, P. L. Combettes, and N. Shahzad, Best approximation from the Kuhn-Tucker set of composite monotone inclusions, *Numer. Funct. Anal. Optim.*, to appear. http://arxiv.org/abs/1401.8005

[3] 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.

[4] H. Attouch, L. M. Briceño-Arias, and P. L. Combettes, A strongly convergent primal-dual method for nonoverlapping domain decomposition, *Numer. Math.*, published online 2015-07-10.

[5] H. H. Bauschke, A note on the paper by Eckstein and Svaiter on "General projective splitting methods for sums of maximal monotone operators", *SIAM J. Control Optim.*, vol. 48, pp. 2513–2515, 2009.

[6] H. H. Bauschke and J. M. Borwein, On projection algorithms for solving convex feasibility problems, *SIAM Rev.,* vol. 38, pp. 367–426, 1996.

[7] H. H. Bauschke and P. L. Combettes, A weak-to-strong convergence principle for Fejér-monotone methods in Hilbert spaces, *Math. Oper. Res.,* vol. 26, pp. 248–264, 2001.

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

[9] R. I. Boţ, E. R. Csetnek, A. Heinrich, and C. Hendrich, On the convergence rate improvement of a primal-dual splitting algorithm for solving monotone inclusion problems, *Math. Programming,* vol. A150, pp. 251–279, 2015.

[10] R. I. Boţ, E. R. Csetnek, and E. Nagy, Solving systems of monotone inclusions via primal-dual splitting techniques, *Taiwanese J. Math.,* vol. 17, pp. 1983–2009, 2013.

[11] L. M. Briceño-Arias, Forward-partial inverse-forward splitting for solving monotone inclusions, *J. Optim. Theory Appl.,* vol. 166, pp. 391–413, 2015.

[12] L. M. Briceño-Arias and P. L. Combettes, A monotone+skew splitting model for composite monotone inclusions in duality, *SIAM J. Optim.,* vol. 21, pp. 1230–1250, 2011.

[13] P. L. Combettes, Hilbertian convex feasibility problem: Convergence of projection methods, *Appl. Math. Optim.,* vol. 35, pp. 311–330, 1997.

[14] P. L. Combettes, Fejér-monotonicity in convex optimization, in: *Encyclopedia of Optimization,* (C. A. Floudas and P. M. Pardalos, Eds.), vol. 2, Springer-Verlag, New York, 2001, pp. 106–114. (Also available in 2nd ed., pp. 1016–1024, 2009.)

[15] P. L. Combettes, Systems of structured monotone inclusions: Duality, algorithms, and applications, *SIAM J. Optim.,* vol. 23, pp. 2420–2447, 2013.

[16] P. L. Combettes and J.-C. Pesquet, Stochastic quasi-Fejér block-coordinate fixed point iterations with random sweeping, *SIAM J. Optim.,* vol. 25, pp. 1221–1248, 2015.

[17] D. Davis, Convergence rate analysis of primal-dual splitting schemes, *SIAM J. Optim.,* vol. 25, pp. 1912–1943, 2015.

[18] Y. Dong, An LS-free splitting method for composite mappings, *Appl. Math. Lett.,* vol. 18, pp. 843–848, 2005.

[19] J. Eckstein and B. F. Svaiter, A family of projective splitting methods for the sum of two maximal monotone operators, *Math. Programming,* vol. 111, pp. 173–199, 2008.

[20] J. Eckstein and B. F. Svaiter, General projective splitting methods for sums of maximal monotone operators, *SIAM J. Control Optim.,* vol. 48, pp. 787–811, 2009.

[21] Y. Haugazeau, *Sur les Inéquations Variationnelles et la Minimisation de Fonctionnelles Convexes.* Thèse, Université de Paris, Paris, France, 1968.

[22] J. J. Moreau, Fonctions convexes duales et points proximaux dans un espace hilbertien, *C. R. Acad. Sci. Paris Sér. A,* vol. 255, pp. 2897–2899, 1962.

[23] N. Ottavy, Strong convergence of projection-like methods in Hilbert spaces, *J. Optim. Theory Appl.*, vol. 56, pp. 433–461, 1988.

[24] T. Pennanen, Dualization of generalized equations of maximal monotone type, *SIAM J. Optim.,* vol. 10, pp. 809–835, 2000.

[25] J.-C. Pesquet and A. Repetti, A class of randomized primal-dual algorithms for distributed optimization, *J. Nonlinear Convex Anal.,* to appear. http://arxiv.org/abs/1406.6404