Mathematics of Operations Research

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Multivariate Monotone Inclusions in Saddle Form

Minh N. Bùi, Patrick L. Combettes

MATHEMATICS OF OPERATIONS RESEARCH
Articles in Advance, pp. 1–28
ISSN 0364-765X (print), ISSN 1526-5471 (online)
http://pubsonline.informs.org/journal/moor

Multivariate Monotone Inclusions in Saddle Form

Minh N. Bùi,[a] Patrick L. Combettes[a]

[a] Department of Mathematics, North Carolina State University, Raleigh, North Carolina 27695

Contact: mnbui@ncsu.edu, https://orcid.org/0000-0003-2910-667X (MNB); plc@math.ncsu.edu, https://orcid.org/0000-0001-9730-5932 (PLC)

Received: February 22, 2020
Revised: October 29, 2020; February 14, 2021
Accepted: March 2, 2021
Published Online in Articles in Advance: December 21, 2021

MSC2010 Subject Classification: Primary: 47H05, 49M27; secondary: 65K05, 90C25
OR/MS Subject Classification. Primary: Programming: infinite dimensional, nonlinear: algorithms; secondary: nonlinear: convex, applications

https://doi.org/10.1287/moor.2021.1161

Abstract. We propose a novel approach to monotone operator splitting based on the notion of a saddle operator. Under investigation is a highly structured multivariate monotone inclusion problem involving a mix of set-valued, cocoercive, and Lipschitzian monotone operators, as well as various monotonicity-preserving operations among them. This model encompasses most formulations found in the literature. A limitation of existing primal-dual algorithms is that they operate in a product space that is too small to achieve full splitting of our problem in the sense that each operator is used individually. To circumvent this difficulty, we recast the problem as that of finding a zero of a saddle operator that acts on a bigger space. This leads to an algorithm of unprecedented flexibility, which achieves full splitting, exploits the specific attributes of each operator, is asynchronous, and requires to activate only blocks of operators at each iteration, as opposed to activating all of them. The latter feature is of critical importance in large-scale problems. The weak convergence of the main algorithm is established, as well as the strong convergence of a variant. Various applications are discussed, and instantiations of the proposed framework in the context of variational inequalities and minimization problems are presented.

Funding: This work was supported by the National Science Foundation [Grant CCF-1715671].

Keywords: monotone inclusion • monotone operator • saddle form • operator splitting • block-iterative algorithm • asynchronous algorithm • strong convergence

1. Introduction

In 1979, several methods appeared to solve the basic problem of finding a zero of the sum of two maximally monotone operators in a real Hilbert space (Lions and Mercier [37], Mercier [38], Passty [43]). Over the past 40 years, increasingly complex inclusion problems and solution techniques have been considered (Boţ and Hendrich [10], Briceño-Arias and Combettes [14], Briceño-Arias and Davis [17], Bùi and Combettes [19], Combettes [23], Combettes and Eckstein [25], Eckstein [29], Johnstone and Eckstein [34], Tseng [53]) to address concrete problems in fields as diverse as game theory (Attouch et al. [2], Briceño-Arias and Combettes [15], Yi and Pavel [56]), evolution inclusions (Attouch et al. [3]), traffic equilibrium (Attouch et al. [3], Fukushima [31]), domain decomposition (Attouch et al. [4]), machine learning (Bach et al. [6], Briceño-Arias et al. [12]), image recovery (Banert et al. [7], Boţ and Hendrich [11], Briceño-Arias et al. [16], Hintermüller and Stadler [33]), mean field games (Briceño-Arias et al. [18]), convex programming (Combettes [24], Li and Yuan [36]), statistics (Combettes and Müller [26], Yan and Bien [55]), neural networks (Combettes and Pesquet [27]), signal processing (Combettes and Wajs [28]), partial differential equations (Ghoussoub [32]), tensor completion (Mizoguchi and Yamada [39]), and optimal transport (Papadakis et al. [42]). In our view, two challenging issues in the field of monotone operator splitting algorithms are the following:

• A number of independent monotone inclusion models coexist with various assumptions on the operators and different types of operation among these operators. At the same time, as will be seen in Section 4, they are not sufficiently general to cover important applications.

• Most algorithms do not allow asynchrony and impose that all the operators be activated at each iteration. They can therefore not handle efficiently modern large-scale problems. The only methods that are asynchronous and block-iterative are limited to specific scenarios (Combettes and Eckstein [25], Eckstein [29], Johnstone and Eckstein [34]), and they do not cover inclusion models such as that of Combettes [23].

In an attempt to bring together and extend the application scope of the wide variety of unrelated models that coexist in the literature, we propose the following multivariate formulation that involves a mix of set-valued, cocoercive, and Lipschitzian monotone operators, as well as various monotonicity-preserving operations among them.

Problem 1. Let $(\mathcal{H}_i)_{i\in I}$ and $(\mathcal{G}_k)_{k\in K}$ be finite families of real Hilbert spaces with Hilbert direct sums $\boldsymbol{\mathcal{H}} = \bigoplus_{i\in I}\mathcal{H}_i$ and $\boldsymbol{\mathcal{G}} = \bigoplus_{k\in K}\mathcal{G}_k$. Denote by $\boldsymbol{x} = (x_i)_{i\in I}$ a generic element in $\boldsymbol{\mathcal{H}}$. For every $i \in I$ and every $k \in K$, let $s_i^* \in \mathcal{H}_i$, let $r_k \in \mathcal{G}_k$, and suppose that the following are satisfied:

[a] $A_i\colon \mathcal{H}_i \to 2^{\mathcal{H}_i}$ is maximally monotone, $C_i\colon \mathcal{H}_i \to \mathcal{H}_i$ is cocoercive with constant $\alpha_i^c \in]0, +\infty[$, $Q_i\colon \mathcal{H}_i \to \mathcal{H}_i$ is monotone and Lipschitzian with constant $\alpha_i^\ell \in [0, +\infty[$, and $R_i\colon \boldsymbol{\mathcal{H}} \to \mathcal{H}_i$.

[b] $B_k^m\colon \mathcal{G}_k \to 2^{\mathcal{G}_k}$ is maximally monotone, $B_k^c\colon \mathcal{G}_k \to \mathcal{G}_k$ is cocoercive with constant $\beta_k^c \in]0, +\infty[$, and $B_k^\ell\colon \mathcal{G}_k \to \mathcal{G}_k$ is monotone and Lipschitzian with constant $\beta_k^\ell \in [0, +\infty[$.

[c] $D_k^m\colon \mathcal{G}_k \to 2^{\mathcal{G}_k}$ is maximally monotone, $D_k^c\colon \mathcal{G}_k \to \mathcal{G}_k$ is cocoercive with constant $\delta_k^c \in]0, +\infty[$, and $D_k^\ell\colon \mathcal{G}_k \to \mathcal{G}_k$ is monotone and Lipschitzian with constant $\delta_k^\ell \in [0, +\infty[$.

[d] $L_{ki}\colon \mathcal{H}_i \to \mathcal{G}_k$ is linear and bounded.

In addition, it is assumed that

[e] $\boldsymbol{R}\colon \boldsymbol{\mathcal{H}} \to \boldsymbol{\mathcal{H}}\colon \boldsymbol{x} \mapsto (R_i\boldsymbol{x})_{i\in I}$ is monotone and Lipschitzian with constant $\chi \in [0, +\infty[$.

The objective is to solve the primal problem

$$\text{find}\ \ \overline{\boldsymbol{x}} \in \boldsymbol{\mathcal{H}}\ \ \text{such that}\ \ (\forall i \in I)\ \ s_i^* \in A_i\overline{x}_i + C_i\overline{x}_i + Q_i\overline{x}_i + R_i\overline{\boldsymbol{x}} + \sum_{k\in K} L_{ki}^*\Bigg(\big((B_k^m + B_k^c + B_k^\ell)\,\square\,(D_k^m + D_k^c + D_k^\ell)\big)\Bigg(\sum_{j\in I} L_{kj}\overline{x}_j - r_k\Bigg)\Bigg) \tag{1}$$

and the associated dual problem

$$\text{find}\ \ \overline{\boldsymbol{v}}^* \in \boldsymbol{\mathcal{G}}\ \text{such that}\ (\exists\, \boldsymbol{x} \in \boldsymbol{\mathcal{H}})(\forall i \in I)(\forall k \in K)\ \begin{cases} s_i^* - \displaystyle\sum_{j\in K} L_{ji}^*\overline{v}_j^* \in A_ix_i + C_ix_i + Q_ix_i + R_i\boldsymbol{x}\\ \overline{v}_k^* \in \big((B_k^m + B_k^c + B_k^\ell)\,\square\,(D_k^m + D_k^c + D_k^\ell)\big)\Bigg(\displaystyle\sum_{j\in I} L_{kj}x_j - r_k\Bigg). \end{cases} \tag{2}$$

Our highly structured model involves three basic monotonicity preserving operations, namely addition, composition with linear operators, and parallel sum. It extends the state-of-the-art model of Combettes [23], where the simpler form

$$(\forall i \in I)\quad s_i^* \in A_i\overline{x}_i + Q_i\overline{x}_i + \sum_{k\in K} L_{ki}^*\Bigg((B_k^m\,\square\,D_k^m)\Bigg(\sum_{j\in I} L_{kj}\overline{x}_j - r_k\Bigg)\Bigg) \tag{3}$$

of the system in (1) was investigated (see Attouch et al. [3] and Combettes and Eckstein [25] for special cases). In an increasing number of applications, the sets I and K can be sizable. To handle such large-scale problems, it is critical to implement block-iterative solution algorithms, in which only subgroups of the operators involved in the problem need to be activated at each iteration. In addition, it is desirable that the algorithm be asynchronous in the sense that, at any iteration, it has the ability to incorporate the result of calculations initiated at earlier iterations. Such methods have been proposed for special cases of Problem 1: first in Combettes and Eckstein [25] for the system

$$\text{find}\ \ \overline{\boldsymbol{x}} \in \boldsymbol{\mathcal{H}}\ \ \text{such that}\ \ (\forall i \in I)\ \ s_i^* \in A_i\overline{x}_i + \sum_{k\in K} L_{ki}^*\Bigg(B_k^m\Bigg(\sum_{j\in I} L_{kj}\overline{x}_j - r_k\Bigg)\Bigg), \tag{4}$$

and then in Eckstein [29] for the inclusion (we omit the subscript 1):

$$\text{find}\ \ \overline{x} \in \mathcal{H}\ \ \text{such that}\ \ 0 \in \sum_{k\in K} L_k^*(B_k^m(L_k\overline{x})), \tag{5}$$

and more recently in Johnstone and Eckstein [34] for the inclusion

$$\text{find}\ \ \overline{x} \in \mathcal{H}\ \ \text{such that}\ \ 0 \in A\overline{x} + Q\overline{x} + \sum_{k\in K} L_k^*((B_k^m + B_k^\ell)(L_k\overline{x})). \tag{6}$$

It is clear that the formulations (4) and (6) are not interdependent. Furthermore, as we shall see in Section 4, many applications of interest are not covered by either of them. From both a theoretical and a practical viewpoint, it is therefore important to unify and extend these approaches. To achieve this goal, we propose to design an algorithm for solving the general Problem 1 which possesses simultaneously the following features:

① It has the ability to process all the operators individually and exploit their specific attributes, for example, set-valuedness, cocoercivity, Lipschitz continuity, and linearity.

② It is block-iterative in the sense that it does not need to activate all the operators at each iteration, but only a subgroup of them.

③ It is asynchronous.

④ Each set-valued monotone operator is scaled by its own, iteration-dependent, parameter.

⑤ It does not require any knowledge of the norms of the linear operators involved in the model.

Let us observe that the method of Combettes and Eckstein [25] has features ①–⑤, but it is restricted to (4). Likewise, the method of Johnstone and Eckstein [34] has features ①–⑤, but it is restricted to (6).

Solving the intricate Problem 1 with the requirement ① does not seem possible with existing tools. The presence of requirements ②–⑤ further complicates this task. In particular, the Kuhn–Tucker approach initiated in Briceño-Arias and Combettes [14]—and further developed in Alotaibi et al. [1], Boţ and Hendrich [10], Combettes [23], Combettes and Eckstein [25], and Johnstone and Eckstein [34, 35]—relies on finding a zero of an operator acting on the primal-dual space $\boldsymbol{\mathcal{H}}\oplus\boldsymbol{\mathcal{G}}$. However, in the context of Problem 1, this primal-dual space is too small to achieve full splitting in the sense that each operator is used individually. To circumvent this difficulty, we propose a novel splitting strategy that consists of recasting the problem as that of finding a zero of a saddle operator acting on the bigger space $\boldsymbol{\mathcal{H}}\oplus\boldsymbol{\mathcal{G}}\oplus\boldsymbol{\mathcal{G}}\oplus\boldsymbol{\mathcal{G}}$. This is done in Section 2, where we define the saddle form of Problem 1, study its properties, and propose outer approximation principles to solve it. In Section 3, the main asynchronous block-iterative algorithm is presented, and we establish its weak convergence under mild conditions on the frequency at which the operators are selected. We also present a strongly convergent variant. The specializations to variational inequalities and multivariate minimization are discussed in Section 4, along with several applications. The appendix contains auxiliary results.

Notation. The notation used in this paper is standard and follows Bauschke and Combettes [9], to which one can refer for background and complements on monotone operators and convex analysis. Let $\mathcal{K}$ be a real Hilbert space. The symbols $\langle\cdot\mid\cdot\rangle$ and $\|\cdot\|$ denote the scalar product of $\mathcal{K}$ and the associated norm, respectively. The expressions $x_n\rightharpoonup x$ and $x_n\to x$ denote, respectively, the weak and the strong convergence of a sequence $(x_n)_{n\in\mathbb{N}}$ to x in $\mathcal{K}$, and $2^{\mathcal{K}}$ denotes the family of all subsets of $\mathcal{K}$. Let $A:\mathcal{K}\to 2^{\mathcal{K}}$. The graph of A is $\operatorname{gra}A=\{(x,x^*)\in\mathcal{K}\times\mathcal{K}\mid x^*\in Ax\}$, the set of zeros of A is $\operatorname{zer}A=\{x\in\mathcal{K}\mid 0\in Ax\}$, the inverse of A is $A^{-1}:\mathcal{K}\to 2^{\mathcal{K}}:x^*\mapsto\{x\in\mathcal{K}\mid x^*\in Ax\}$, and the resolvent of A is $J_A=(\mathrm{Id}+A)^{-1}$, where Id is the identity operator on $\mathcal{K}$. Furthermore, A is monotone if

$$(\forall(x,x^*)\in\operatorname{gra}A)(\forall(y,y^*)\in\operatorname{gra}A)\quad\langle x-y\mid x^*-y^*\rangle\geqslant 0,\tag{7}$$

and it is maximally monotone if, for every $(x,x^*)\in\mathcal{K}\times\mathcal{K}$,

$$(x,x^*)\in\operatorname{gra}A\quad\Leftrightarrow\quad(\forall(y,y^*)\in\operatorname{gra}A)\ \langle x-y\mid x^*-y^*\rangle\geqslant 0.\tag{8}$$

If A is maximally monotone, then J_A is a single-valued operator defined on $\mathcal{K}$. The parallel sum of $B:\mathcal{K}\to 2^{\mathcal{K}}$ and $D:\mathcal{K}\to 2^{\mathcal{K}}$ is $B\,\Box\,D=(B^{-1}+D^{-1})^{-1}$. An operator $C:\mathcal{K}\to\mathcal{K}$ is cocoercive with constant $\alpha\in\,]0,+\infty[$ if $(\forall x\in\mathcal{K})(\forall y\in\mathcal{K})\ \langle x-y\mid Cx-Cy\rangle\geqslant\alpha\|Cx-Cy\|^2$. We denote by $\Gamma_0(\mathcal{K})$ the class of lower semicontinuous convex functions $f:\mathcal{K}\to\,]-\infty,+\infty]$ such that $\operatorname{dom}f=\{x\in\mathcal{K}\mid f(x)<+\infty\}\neq\varnothing$. Let $f\in\Gamma_0(\mathcal{K})$. The conjugate of f is the function $\Gamma_0(\mathcal{K})\ni f^*:x^*\mapsto\sup_{x\in\mathcal{K}}(\langle x\mid x^*\rangle-f(x))$ and the subdifferential of f is the maximally monotone operator $\partial f:\mathcal{K}\to 2^{\mathcal{K}}:x\mapsto\{x^*\in\mathcal{K}\mid(\forall y\in\mathcal{K})\langle y-x\mid x^*\rangle+f(x)\leqslant f(y)\}$. In addition, $\operatorname{epi}f$ is the epigraph of f. For every $x\in\mathcal{K}$, the unique minimizer of $f+(1/2)\|\cdot-x\|^2$ is denoted by $\operatorname{prox}_f x$. We have $\operatorname{prox}_f=J_{\partial f}$. Given $h\in\Gamma_0(\mathcal{K})$, the infimal convolution of f and h is $f\,\Box\,h:\mathcal{K}\to[-\infty,+\infty]:x\mapsto\inf_{y\in\mathcal{K}}(f(y)+h(x-y))$; the infimal convolution $f\,\Box\,h$ is exact if the infimum is achieved everywhere, in which case we write $f\,\boxdot\,h$. Now let $(\mathcal{K}_i)_{i\in I}$ be a finite family of real Hilbert spaces and, for every $i\in I$, let $f_i:\mathcal{K}_i\to]-\infty,+\infty]$. Then

$$\bigoplus_{i\in I}f_i:\boldsymbol{\mathcal{K}}=\bigoplus_{i\in I}\mathcal{K}_i\to\,]-\infty,+\infty]:\boldsymbol{x}\mapsto\sum_{i\in I}f_i(x_i).\tag{9}$$

The partial derivative of a differentiable function $\Theta:\boldsymbol{\mathcal{K}}\to\mathbb{R}$ relative to $\mathcal{K}_i$ is denoted by $\nabla_i\Theta$. Finally, let C be a nonempty convex subset of $\mathcal{K}$. A point $x\in C$ belongs to the strong relative interior of C, in symbols $x\in\operatorname{sri}C$, if $\bigcup_{\lambda\in]0,+\infty[}\lambda(C-x)$ is a closed vector subspace of $\mathcal{K}$. If C is closed, the projection operator onto it is denoted by proj_C and the normal cone operator of C is the maximally monotone operator

$$N_C:\mathcal{K}\to 2^{\mathcal{K}}:x\mapsto\begin{cases}\{x^*\in\mathcal{K}\mid\sup\langle C-x\mid x^*\rangle\leqslant 0\}, & \text{if } x\in C;\\ \varnothing, & \text{otherwise.}\end{cases}\tag{10}$$

2. The Saddle Form of Problem 1

A classical Lagrangian setting for convex minimization is the following. Given real Hilbert spaces $\mathcal{H}$ and $\mathcal{G}$, $f \in \Gamma_0(\mathcal{H})$, $g \in \Gamma_0(\mathcal{G})$, and a bounded linear operator $L: \mathcal{H} \to \mathcal{G}$, consider the primal problem

$$\underset{x \in \mathcal{H}}{\text{minimize}}\; f(x) + g(Lx) \tag{11}$$

together with its Fenchel–Rockafellar dual (Rockafellar [47])

$$\underset{v^* \in \mathcal{G}}{\text{minimize}}\; f^*(-L^*v^*) + g^*(v^*). \tag{12}$$

The primal-dual pair (11)–(12) can be analyzed through the lens of Rockafellar's saddle formalism (Rockafellar [49, 50]) as follows. Set $h: \mathcal{H} \oplus \mathcal{G} \to \left]-\infty, +\infty\right]: (x, y) \mapsto f(x) + g(y)$ and $U: \mathcal{H} \oplus \mathcal{G} \to \mathcal{G}: (x, y) \mapsto Lx - y$, and note that $U^*: \mathcal{G} \to \mathcal{H} \oplus \mathcal{G}: v^* \mapsto (L^*v^*, -v^*)$. Then, upon defining $\mathcal{K} = \mathcal{H} \oplus \mathcal{G}$ and introducing the variable $z = (x, y) \in \mathcal{K}$, (11) is equivalent to

$$\underset{z \in \mathcal{K},\, Uz = 0}{\text{minimize}}\; h(z) \tag{13}$$

and (12) to

$$\underset{v^* \in \mathcal{G}}{\text{minimize}}\; h^*(-U^*v^*). \tag{14}$$

The Lagrangian associated with (13) is (see Rockafellar [51, example 4′] or Bauschke and Combettes [9, proposition 19.21])

$$\begin{aligned} \mathcal{L}: \mathcal{K} \oplus \mathcal{G} &\to \left]-\infty, +\infty\right] \\ (z, v^*) &\mapsto \begin{cases} h(z) + \langle Uz \mid v^* \rangle, & \text{if } z \in \operatorname{dom} h; \\ +\infty, & \text{otherwise}, \end{cases} \end{aligned} \tag{15}$$

and the associated saddle operator (Rockafellar [49, 50]) is the maximally monotone operator

$$\mathcal{S}: \mathcal{K} \oplus \mathcal{G} \to 2^{\mathcal{K} \oplus \mathcal{G}}: (z, v^*) \mapsto \partial\mathcal{L}(\cdot, v^*)(z) \times \partial(-\mathcal{L}(z, \cdot))(v^*) = (\partial h(z) + U^*v^*) \times \{-Uz\}. \tag{16}$$

As shown in Rockafellar [49], a zero $(\overline{z}, \overline{v}^*)$ of $\mathcal{S}$ is a saddle point of $\mathcal{L}$, and it has the property that $\overline{z}$ solves (13) and $\overline{v}^*$ solves (14). Thus, going back to the original Fenchel–Rockafellar pair (11)–(12), we learn that, if $(\overline{x}, \overline{y}, \overline{v}^*)$ is a zero of the saddle operator

$$\mathcal{S}: \mathcal{H} \oplus \mathcal{G} \oplus \mathcal{G} \to 2^{\mathcal{H} \oplus \mathcal{G} \oplus \mathcal{G}}: (x, y, v^*) \mapsto (\partial f(x) + L^*v^*) \times (\partial g(y) - v^*) \times \{-Lx + y\}, \tag{17}$$

then $\overline{x}$ solves (11) and $\overline{v}^*$ solves (12). As shown in Combettes [24, section 4.5], a suitable splitting of $\mathcal{S}$ leads to an implementable algorithm to solve (11)–(12).

A generalization of Fenchel–Rockafellar duality to monotone inclusions was proposed in Pennanen [44] and Robinson [46] and further extended in Combettes [23]. Given maximally monotone operators $A: \mathcal{H} \to 2^{\mathcal{H}}$ and $B: \mathcal{G} \to 2^{\mathcal{G}}$, and a bounded linear operator $L: \mathcal{H} \to \mathcal{G}$, the primal problem

$$\text{find } \overline{x} \in \mathcal{H} \text{ such that } 0 \in A\overline{x} + L^*(B(L\overline{x})) \tag{18}$$

is paired with the dual problem

$$\text{find } \overline{v}^* \in \mathcal{G} \text{ such that } 0 \in -L(A^{-1}(-L^*\overline{v}^*)) + B^{-1}\overline{v}^*. \tag{19}$$

Following the same pattern as that described above, let us consider the *saddle operator*

$$\mathcal{S}: \mathcal{H} \oplus \mathcal{G} \oplus \mathcal{G} \to 2^{\mathcal{H} \oplus \mathcal{G} \oplus \mathcal{G}}: (x, y, v^*) \mapsto (Ax + L^*v^*) \times (By - v^*) \times \{-Lx + y\}. \tag{20}$$

It is readily shown that, if $(\overline{x}, \overline{y}, \overline{v}^*)$ is a zero of $\mathcal{S}$, then $\overline{x}$ solves (18) and $\overline{v}^*$ solves (19). We call the problem of finding a zero of $\mathcal{S}$ the *saddle form* of (18)–(19). We now introduce a saddle operator for the general Problem 1.

Definition 1. In the setting of Problem 1, let $\mathbf{X} = \boldsymbol{\mathcal{H}} \oplus \boldsymbol{\mathcal{G}} \oplus \boldsymbol{\mathcal{G}} \oplus \boldsymbol{\mathcal{G}}$. The *saddle operator* associated with Problem 1 is

$$\begin{aligned} &\boldsymbol{\mathcal{S}}: \mathbf{X} \to 2^{\mathbf{X}}: (\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}, \boldsymbol{v}^*) \mapsto \\ &\quad \Bigg(\bigtimes_{i \in I} \Big(-s_i^* + A_i x_i + C_i x_i + Q_i x_i + R_i \boldsymbol{x} + \sum_{k \in K} L_{ki}^* v_k^* \Big), \bigtimes_{k \in K} (B_k^{m} y_k + B_k^{c} y_k + B_k^{\ell} y_k - v_k^*), \\ &\quad \bigtimes_{k \in K} (D_k^{m} z_k + D_k^{c} z_k + D_k^{\ell} z_k - v_k^*), \bigtimes_{k \in K} \Big\{ r_k + y_k + z_k - \sum_{i \in I} L_{ki} x_i \Big\} \Bigg), \end{aligned} \tag{21}$$

and the *saddle form* of Problem 1 is to

$$\text{find}\quad \overline{\mathbf{x}} \in \mathbf{X} \quad \text{such that}\quad \mathbf{0} \in \mathcal{S}\overline{\mathbf{x}}. \tag{22}$$

Next, we establish some properties of the saddle operator as well as connections with Problem 1.

Proposition 1. *Consider the setting of Problem 1 and Definition 1. Let $\mathscr{P}$ be the set of solutions to (1), let $\mathscr{D}$ be the set of solutions to (2), and let*

$$\mathsf{Z} = \Bigg\{(\overline{x},\overline{v}^*) \in \mathcal{H}\oplus\mathcal{G} \;\Bigg|\; (\forall i\in I)(\forall k\in K)\quad s_i^* - \sum_{j\in K} L_{ji}^*\overline{v}_j^* \in A_i\overline{x}_i + C_i\overline{x}_i + Q_i\overline{x}_i + R_i\overline{x}\ \text{ and }$$
$$\sum_{j\in I} L_{kj}\overline{x}_j - r_k \in (B_k^{m} + B_k^{c} + B_k^{\ell})^{-1}\overline{v}_k^* + (D_k^{m} + D_k^{c} + D_k^{\ell})^{-1}\overline{v}_k^*\Bigg\} \tag{23}$$

be the associated Kuhn–Tucker *set. Then the following hold:*

(i) $\mathcal{S}$ *is maximally monotone.*

(ii) zer $\mathcal{S}$ *is closed and convex.*

(iii) *Suppose that* $\overline{\mathbf{x}} = (\overline{x},\overline{y},\overline{z},\overline{v}^*) \in \operatorname{zer}\mathcal{S}$. *Then* $(\overline{x},\overline{v}^*) \in \mathsf{Z} \subset \mathscr{P}\times\mathscr{D}$.

(iv) $\mathscr{D} \neq \emptyset \iff \operatorname{zer}\mathcal{S} \neq \emptyset \iff \mathsf{Z} \neq \emptyset \Rightarrow \mathscr{P} \neq \emptyset$.

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

[a] I *is a singleton.*

[b] *For every* $k\in K$, $(B_k^{m} + B_k^{c} + B_k^{\ell}) \,\square\, (D_k^{m} + D_k^{c} + D_k^{\ell})$ *is at most single-valued.*

[c] *For every* $k\in K$, $(D_k^{m} + D_k^{c} + D_k^{\ell})^{-1}$ *is strictly monotone.*

[d] $I\subset K$, *the operators* $((B_k^{m} + B_k^{c} + B_k^{\ell})\,\square\,(D_k^{m} + D_k^{c} + D_k^{\ell}))_{k\in K\setminus I}$ *are at most single-valued, and* $(\forall i\in I)(\forall k\in I)\ k\neq i \Rightarrow L_{ki} = 0$.

Then $\mathscr{P}\neq\emptyset \Rightarrow \mathsf{Z}\neq\emptyset$.

Proof. Define

$$\begin{cases} \boldsymbol{A}\colon \mathcal{H}\to 2^{\mathcal{H}} : \boldsymbol{x}\mapsto \boldsymbol{R}\boldsymbol{x} + \bigtimes_{i\in I}(A_ix_i + C_ix_i + Q_ix_i)\\ \boldsymbol{B}\colon \mathcal{G}\to 2^{\mathcal{G}} : \boldsymbol{y}\mapsto \bigtimes_{k\in K}(B_k^{m}y_k + B_k^{c}y_k + B_k^{\ell}y_k)\\ \boldsymbol{D}\colon \mathcal{G}\to 2^{\mathcal{G}} : \boldsymbol{z}\mapsto \bigtimes_{k\in K}(D_k^{m}z_k + D_k^{c}z_k + D_k^{\ell}z_k)\\ \boldsymbol{L}\colon \mathcal{H}\to\mathcal{G} : \boldsymbol{x}\mapsto \left(\sum_{i\in I} L_{ki}x_i\right)_{k\in K}\\ \boldsymbol{s}^* = (s_i^*)_{i\in I}\ \text{ and }\ \boldsymbol{r} = (r_k)_{k\in K}. \end{cases} \tag{24}$$

Then the adjoint of $\boldsymbol{L}$ is

$$\boldsymbol{L}^*\colon \mathcal{G}\to\mathcal{H} : \boldsymbol{v}^*\mapsto \left(\sum_{k\in K} L_{ki}^*v_k^*\right)_{i\in I}. \tag{25}$$

Hence, in view of (21) and (24),

$$\mathcal{S}\colon \mathbf{X}\to 2^{\mathbf{X}} : (\boldsymbol{x},\boldsymbol{y},\boldsymbol{z},\boldsymbol{v}^*)\mapsto (-\boldsymbol{s}^* + \boldsymbol{A}\boldsymbol{x} + \boldsymbol{L}^*\boldsymbol{v}^*)\times(\boldsymbol{B}\boldsymbol{y} - \boldsymbol{v}^*)\times(\boldsymbol{D}\boldsymbol{z} - \boldsymbol{v}^*)\times\{\boldsymbol{r} - \boldsymbol{L}\boldsymbol{x} + \boldsymbol{y} + \boldsymbol{z}\}. \tag{26}$$

(i): Let us introduce the operators

$$\begin{cases} \mathbf{P}\colon \mathbf{X}\to 2^{\mathbf{X}} : (\boldsymbol{x},\boldsymbol{y},\boldsymbol{z},\boldsymbol{v}^*)\mapsto (-\boldsymbol{s}^* + \boldsymbol{A}\boldsymbol{x})\times\boldsymbol{B}\boldsymbol{y}\times\boldsymbol{D}\boldsymbol{z}\times\{\boldsymbol{r}\}\\ \mathbf{W}\colon \mathbf{X}\to\mathbf{X} : (\boldsymbol{x},\boldsymbol{y},\boldsymbol{z},\boldsymbol{v}^*)\mapsto (\boldsymbol{L}^*\boldsymbol{v}^*, -\boldsymbol{v}^*, -\boldsymbol{v}^*, -\boldsymbol{L}\boldsymbol{x} + \boldsymbol{y} + \boldsymbol{z}). \end{cases} \tag{27}$$

Using Problem 1[a]–[c], we derive from Bauschke and Combettes [9, example 20.31 and corollaries 20.28 and 25.5(i)] that, for every $i\in I$ and every $k\in K$, the operators $A_i + C_i + Q_i$, $B_k^{m} + B_k^{c} + B_k^{\ell}$, and $D_k^{m} + D_k^{c} + D_k^{\ell}$ are maximally monotone. At the same time, Problem 1[e] and Bauschke and Combettes [9, corollary 20.28] entail that $\boldsymbol{R}$ is maximally monotone. Therefore, it results from (24), Bauschke and Combettes [9, proposition 20.23 and corollary 25.5(i)], and (27) that $\mathbf{P}$ is maximally monotone. However, since Problem 1[d] and (27) imply that $\mathbf{W}$ is linear and bounded with $\mathbf{W}^* = -\mathbf{W}$, Bauschke and Combettes [9, example 20.35] asserts that $\mathbf{W}$ is maximally monotone. Hence, in view of Bauschke and Combettes [9, corollary 25.5(i)], we infer from (26)–(27) that $\mathcal{S} = \mathbf{P} + \mathbf{W}$ is maximally monotone.

(ii): This follows from (i) and Bauschke and Combettes [9, proposition 23.39].

(iii): Using (24) and (25), we deduce from (23) that

$$\mathsf{Z} = \big\{(\boldsymbol{x},\boldsymbol{v}^*) \in \boldsymbol{\mathcal{H}}\oplus\boldsymbol{\mathcal{G}} \;\big|\; \boldsymbol{s}^* - \boldsymbol{L}^*\boldsymbol{v}^* \in \boldsymbol{A}\boldsymbol{x} \;\text{ and }\; \boldsymbol{L}\boldsymbol{x} - \boldsymbol{r} \in \boldsymbol{B}^{-1}\boldsymbol{v}^* + \boldsymbol{D}^{-1}\boldsymbol{v}^*\big\} \tag{28}$$

and from (2) that

$$\mathscr{D} = \big\{\boldsymbol{v}^* \in \boldsymbol{\mathcal{G}} \;\big|\; -\boldsymbol{r} \in -\boldsymbol{L}(\boldsymbol{A}^{-1}(\boldsymbol{s}^* - \boldsymbol{L}^*\boldsymbol{v}^*)) + \boldsymbol{B}^{-1}\boldsymbol{v}^* + \boldsymbol{D}^{-1}\boldsymbol{v}^*\big\}. \tag{29}$$

Suppose that $(\boldsymbol{x},\boldsymbol{v}^*) \in \mathsf{Z}$. Then it follows from (28) that $\boldsymbol{x} \in \boldsymbol{A}^{-1}(\boldsymbol{s}^* - \boldsymbol{L}^*\boldsymbol{v}^*)$ and, in turn, that $-\boldsymbol{r} \in -\boldsymbol{L}\boldsymbol{x} + \boldsymbol{B}^{-1}\boldsymbol{v}^* + \boldsymbol{D}^{-1}\boldsymbol{v}^* \subset -\boldsymbol{L}(\boldsymbol{A}^{-1}(\boldsymbol{s}^* - \boldsymbol{L}^*\boldsymbol{v}^*)) + \boldsymbol{B}^{-1}\boldsymbol{v}^* + \boldsymbol{D}^{-1}\boldsymbol{v}^*$. Thus $\boldsymbol{v}^* \in \mathscr{D}$ by (29). In addition, (23) implies that

$$(\forall k \in K)\quad v_k^* \in ((B_k^{m} + B_k^{c} + B_k^{\ell}) \,\square\, (D_k^{m} + D_k^{c} + D_k^{\ell}))\Bigg(\sum_{j\in I} L_{kj}x_j - r_k\Bigg) \tag{30}$$

and, therefore, that

$$\begin{aligned}(\forall i \in I)\quad s_i^* &\in A_ix_i + C_ix_i + Q_ix_i + R_i\boldsymbol{x} + \sum_{k\in K} L_{ki}^*v_k^*\\ &\subset A_ix_i + C_ix_i + Q_ix_i + R_i\boldsymbol{x}\\ &\quad + \sum_{k\in K} L_{ki}^*\Bigg(((B_k^{m} + B_k^{c} + B_k^{\ell}) \,\square\, (D_k^{m} + D_k^{c} + D_k^{\ell}))\Bigg(\sum_{j\in I} L_{kj}x_j - r_k\Bigg)\Bigg).\end{aligned} \tag{31}$$

Hence, $\boldsymbol{x} \in \mathscr{P}$. To summarize, we have shown that $\mathsf{Z} \subset \mathscr{P} \times \mathscr{D}$. It remains to show that $(\overline{\boldsymbol{x}},\overline{\boldsymbol{v}}^*) \in \mathsf{Z}$. Since $\boldsymbol{0} \in \boldsymbol{\mathcal{S}}\overline{\mathbf{x}}$, we deduce from (26) that $\boldsymbol{s}^* - \boldsymbol{L}^*\overline{\boldsymbol{v}}^* \in \boldsymbol{A}\overline{\boldsymbol{x}}$, $\boldsymbol{L}\overline{\boldsymbol{x}} - \boldsymbol{r} = \overline{\boldsymbol{y}} + \overline{\boldsymbol{z}}$, $\boldsymbol{0} \in \boldsymbol{B}\overline{\boldsymbol{y}} - \overline{\boldsymbol{v}}^*$, and $\boldsymbol{0} \in \boldsymbol{D}\overline{\boldsymbol{z}} - \overline{\boldsymbol{v}}^*$. Therefore, $\boldsymbol{L}\overline{\boldsymbol{x}} - \boldsymbol{r} \in \boldsymbol{B}^{-1}\overline{\boldsymbol{v}}^* + \boldsymbol{D}^{-1}\overline{\boldsymbol{v}}^*$ and (28) thus yields $(\overline{\boldsymbol{x}},\overline{\boldsymbol{v}}^*) \in \mathsf{Z}$.

(iv): The implication $\operatorname{zer}\boldsymbol{\mathcal{S}} \neq \emptyset \Rightarrow \mathscr{P} \neq \emptyset$ follows from (iii). Next, we derive from (29) and (28) that

$$\begin{aligned}\mathscr{D} \neq \emptyset &\Leftrightarrow (\exists \overline{\boldsymbol{v}}^* \in \boldsymbol{\mathcal{G}})\quad -\boldsymbol{r} \in -\boldsymbol{L}(\boldsymbol{A}^{-1}(\boldsymbol{s}^* - \boldsymbol{L}^*\overline{\boldsymbol{v}}^*)) + \boldsymbol{B}^{-1}\overline{\boldsymbol{v}}^* + \boldsymbol{D}^{-1}\overline{\boldsymbol{v}}^*\\ &\Leftrightarrow (\exists (\overline{\boldsymbol{v}}^*,\overline{\boldsymbol{x}}) \in \boldsymbol{\mathcal{G}}\oplus\boldsymbol{\mathcal{H}})\quad -\boldsymbol{r} \in -\boldsymbol{L}\overline{\boldsymbol{x}} + \boldsymbol{B}^{-1}\overline{\boldsymbol{v}}^* + \boldsymbol{D}^{-1}\overline{\boldsymbol{v}}^* \;\text{ and }\; \overline{\boldsymbol{x}} \in \boldsymbol{A}^{-1}(\boldsymbol{s}^* - \boldsymbol{L}^*\overline{\boldsymbol{v}}^*)\\ &\Leftrightarrow (\exists (\overline{\boldsymbol{x}},\overline{\boldsymbol{v}}^*) \in \boldsymbol{\mathcal{H}}\oplus\boldsymbol{\mathcal{G}})\quad \boldsymbol{s}^* - \boldsymbol{L}^*\overline{\boldsymbol{v}}^* \in \boldsymbol{A}\overline{\boldsymbol{x}} \;\text{ and }\; \boldsymbol{L}\overline{\boldsymbol{x}} - \boldsymbol{r} \in \boldsymbol{B}^{-1}\overline{\boldsymbol{v}}^* + \boldsymbol{D}^{-1}\overline{\boldsymbol{v}}^*\\ &\Leftrightarrow \mathsf{Z} \neq \emptyset.\end{aligned} \tag{32}$$

However, (iii) asserts that $\operatorname{zer}\boldsymbol{\mathcal{S}} \neq \emptyset \Rightarrow \mathsf{Z} \neq \emptyset$. Therefore, it remains to show that $\mathsf{Z} \neq \emptyset \Rightarrow \operatorname{zer}\boldsymbol{\mathcal{S}} \neq \emptyset$. Towards this end, suppose that $(\overline{\boldsymbol{x}},\overline{\boldsymbol{v}}^*) \in \mathsf{Z}$. Then, by (28), $\boldsymbol{s}^* - \boldsymbol{L}^*\overline{\boldsymbol{v}}^* \in \boldsymbol{A}\overline{\boldsymbol{x}}$ and $\boldsymbol{L}\overline{\boldsymbol{x}} - \boldsymbol{r} \in \boldsymbol{B}^{-1}\overline{\boldsymbol{v}}^* + \boldsymbol{D}^{-1}\overline{\boldsymbol{v}}^*$. Hence, $\boldsymbol{0} \in -\boldsymbol{s}^* + \boldsymbol{A}\overline{\boldsymbol{x}} + \boldsymbol{L}^*\overline{\boldsymbol{v}}^*$, and there exists $(\overline{\boldsymbol{y}},\overline{\boldsymbol{z}}) \in \boldsymbol{\mathcal{G}}\oplus\boldsymbol{\mathcal{G}}$ such that $\overline{\boldsymbol{y}} \in \boldsymbol{B}^{-1}\overline{\boldsymbol{v}}^*$, $\overline{\boldsymbol{z}} \in \boldsymbol{D}^{-1}\overline{\boldsymbol{v}}^*$, and $\boldsymbol{L}\overline{\boldsymbol{x}} - \boldsymbol{r} = \overline{\boldsymbol{y}} + \overline{\boldsymbol{z}}$. We thus deduce that $\boldsymbol{0} \in \boldsymbol{B}\overline{\boldsymbol{y}} - \overline{\boldsymbol{v}}^*$, $\boldsymbol{0} \in \boldsymbol{D}\overline{\boldsymbol{z}} - \overline{\boldsymbol{v}}^*$, and $\boldsymbol{r} - \boldsymbol{L}\overline{\boldsymbol{x}} + \overline{\boldsymbol{y}} + \overline{\boldsymbol{z}} = \boldsymbol{0}$. Consequently, (26) implies that $(\overline{\boldsymbol{x}},\overline{\boldsymbol{y}},\overline{\boldsymbol{z}},\overline{\boldsymbol{v}}^*) \in \operatorname{zer}\boldsymbol{\mathcal{S}}$.

(v): In view of (iv), it suffices to establish that $\mathscr{P} \neq \emptyset \Rightarrow \mathscr{D} \neq \emptyset$. Suppose that $\overline{\boldsymbol{x}} \in \mathscr{P}$.

[a]: Suppose that $I = \{1\}$. We then infer from (1) that there exists $\overline{\boldsymbol{v}}^* \in \boldsymbol{\mathcal{G}}$ such that

$$\begin{cases} s_1^* \in A_1\overline{x}_1 + C_1\overline{x}_1 + Q_1\overline{x}_1 + R_1\overline{\boldsymbol{x}} + \sum_{k\in K} L_{k1}^*\overline{v}_k^*\\ (\forall k \in K)\; \overline{v}_k^* \in ((B_k^{m} + B_k^{c} + B_k^{\ell}) \,\square\, (D_k^{m} + D_k^{c} + D_k^{\ell}))(L_{k1}\overline{x}_1 - r_k).\end{cases} \tag{33}$$

Therefore, by (2), $\overline{\boldsymbol{v}}^* \in \mathscr{D}$.

[b]: Set $(\forall k \in K)\; \overline{v}_k^* = ((B_k^{m} + B_k^{c} + B_k^{\ell}) \,\square\, (D_k^{m} + D_k^{c} + D_k^{\ell}))(\sum_{j\in I} L_{kj}\overline{x}_j - r_k)$. Then $\overline{\boldsymbol{v}}^*$ solves (2).

[c] $\Rightarrow$ [b]: See Combettes [23, section 4].

[d]: Let $i \in I$. It results from our assumption that

$$\begin{aligned}s_i^* \in{}& A_i\overline{x}_i + C_i\overline{x}_i + Q_i\overline{x}_i + R_i\overline{\boldsymbol{x}} + L_{ii}^*(((B_i^{m} + B_i^{c} + B_i^{\ell}) \,\square\, (D_i^{m} + D_i^{c} + D_i^{\ell}))(L_{ii}\overline{x}_i - r_i))\\ &+ \sum_{k\in K\setminus I} L_{ki}^*\Bigg(((B_k^{m} + B_k^{c} + B_k^{\ell}) \,\square\, (D_k^{m} + D_k^{c} + D_k^{\ell}))\Bigg(\sum_{j\in I} L_{kj}\overline{x}_j - r_k\Bigg)\Bigg).\end{aligned} \tag{34}$$

Thus, there exists $\overline{v}_i^* \in \mathcal{G}_i$ such that $\overline{v}_i^* \in ((B_i^{m} + B_i^{c} + B_i^{\ell}) \,\square\, (D_i^{m} + D_i^{c} + D_i^{\ell}))(L_{ii}\overline{x}_i - r_i)$ and that

$$s_i^* \in A_i\overline{x}_i + C_i\overline{x}_i + Q_i\overline{x}_i + R_i\overline{\boldsymbol{x}} + L_{ii}^*\overline{v}_i^* + \sum_{k\in K\setminus I} L_{ki}^*\Bigg(((B_k^{m} + B_k^{c} + B_k^{\ell}) \,\square\, (D_k^{m} + D_k^{c} + D_k^{\ell}))\Bigg(\sum_{j\in I} L_{kj}\overline{x}_j - r_k\Bigg)\Bigg). \tag{35}$$

As a result, upon setting

$$(\forall k \in K\backslash I) \quad \overline{v}_k^* = ((B_k^m + B_k^c + B_k^\ell) \,\square\, (D_k^m + D_k^c + D_k^\ell))\left(\sum_{j\in I} L_{kj}\overline{x}_j - r_k\right), \tag{36}$$

we conclude that $\overline{\boldsymbol{v}}^* \in \mathscr{D}$. □

Remark 1. Some noteworthy observations about Proposition 1 are the following.

(i) The Kuhn–Tucker set (23) extends to Problem 1 the corresponding notion introduced for some special cases in Alotaibi et al. [1], Briceño-Arias and Combettes [14], and Combettes and Eckstein [25].

(ii) In connection with Proposition 1(v), we note that the implication $\mathscr{P} \neq \emptyset \Rightarrow \mathsf{Z} \neq \emptyset$ is implicitly used in Combettes and Eckstein [25, theorems 13 and 15], where one requires $\mathsf{Z} \neq \emptyset$ but merely assumes $\mathscr{P} \neq \emptyset$. However, this implication is not true in general (a similar oversight is found in Alotaibi et al. [1], Pesquet and Repetti [45], and Rosasco et al. [52]). Indeed, consider as a special case of (1), the problem of solving the system

$$\begin{cases} 0 \in B_1(x_1 + x_2) + B_2(x_1 - x_2) \\ 0 \in B_1(x_1 + x_2) - B_2(x_1 - x_2) \end{cases} \tag{37}$$

in the Euclidean plane $\mathbb{R}^2$. Then, by choosing $B_1 = \{0\}^{-1}$ and $B_2 = 1$, we obtain $\mathscr{P} = \{(x_1, -x_1) \mid x_1 \in \mathbb{R}\}$, whereas $\mathsf{Z} = \emptyset$.

(iii) As stated in Proposition 1(iii), any Kuhn–Tucker point is a solution to (1)–(2). In the simpler setting considered in Combettes and Eckstein [25], a splitting algorithm was devised for finding such a point. However, in the more general context of Problem 1, there does not seem to exist a path from the Kuhn–Tucker formalism in $\mathcal{H} \oplus \mathcal{G}$ to an algorithm that is fully split in the sense of ①. This motivates our approach, which seeks a zero of the saddle operator $\mathcal{S}$ defined on the bigger space $\mathbf{X}$ and, thereby, offers more flexibility.

(iv) Special cases of Problem 1 can be found in Alotaibi et al. [1], Combettes and Eckstein [25], and Johnstone and Eckstein [34, 35], where they were solved by algorithms that proceed by outer approximation of the Kuhn–Tucker set in $\mathcal{H} \oplus \mathcal{G}$. In those special cases, Algorithm 1 below does not reduce to those of Alotaibi et al. [1], Combettes and Eckstein [25], and Johnstone and Eckstein [34, 35] since it operates by outer approximation of the set of zeros of the saddle operator $\mathcal{S}$ in the bigger space $\mathbf{X}$.

The following operators will induce a decomposition of the saddle operator that will lead to a splitting algorithm which complies with our requirements ①–⑤.

Definition 2. In the setting of Definition 1, set

$$\begin{aligned} \mathbf{M} : \mathbf{X} \to 2^{\mathbf{X}} : (\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}, \boldsymbol{v}^*) \mapsto \Bigg(& \bigtimes_{i\in I}\left(-s_i^* + A_i x_i + Q_i x_i + R_i \boldsymbol{x} + \sum_{k\in K} L_{ki}^* v_k^*\right), \bigtimes_{k\in K}(B_k^m y_k + B_k^\ell y_k - v_k^*), \\ & \bigtimes_{k\in K}(D_k^m z_k + D_k^\ell z_k - v_k^*), \bigtimes_{k\in K}\left\{r_k + y_k + z_k - \sum_{i\in I} L_{ki} x_i\right\} \Bigg) \end{aligned} \tag{38}$$

and

$$\mathbf{C} : \mathbf{X} \to \mathbf{X} : (\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}, \boldsymbol{v}^*) \mapsto \Big((C_i x_i)_{i\in I}, (B_k^c y_k)_{k\in K}, (D_k^c z_k)_{k\in K}, \mathbf{0}\Big). \tag{39}$$

Proposition 2. *In the setting of Problem 1 and of Definitions 1 and 2, the following hold:*

(i) $\mathcal{S} = \mathbf{M} + \mathbf{C}$.

(ii) $\mathbf{M}$ *is maximally monotone.*

(iii) *Set* $\alpha = \min\{\alpha_i^c, \beta_k^c, \delta_k^c\}_{i\in I, k\in K}$. *Then the following hold:*

(a) $\mathbf{C}$ *is* α*-cocoercive.*

(b) *Let* $(\mathbf{p}, \mathbf{p}^*) \in \operatorname{gra} \mathbf{M}$ *and* $\mathbf{q} \in \mathbf{X}$. *Then* $\operatorname{zer} \mathcal{S} \subset \left\{\mathbf{x} \in \mathbf{X} \mid \langle \mathbf{x} - \mathbf{p} \mid \mathbf{p}^* + \mathbf{C}\mathbf{q}\rangle \leqslant (4\alpha)^{-1}\|\mathbf{p} - \mathbf{q}\|^2\right\}$.

Proof.

(i): Clear from (21), (38), and (39).

(ii): This is a special case of Proposition 1(i), where, for every $i \in I$ and every $k \in K$, $C_i = 0$ and $B_k^c = D_k^c = 0$.

(iii)(a): Take $\mathbf{x} = (x, y, z, v^*)$ and $\mathbf{y} = (a, b, c, w^*)$ in $\mathbf{X}$. By (39) and Problem 1[a]–[c],

$$\begin{aligned}\langle \mathbf{x} - \mathbf{y} \mid \mathbf{C}\mathbf{x} - \mathbf{C}\mathbf{y}\rangle &= \sum_{i\in I}\langle x_i - a_i \mid C_i x_i - C_i a_i\rangle + \sum_{k\in K}\big(\langle y_k - b_k \mid B_k^c y_k - B_k^c b_k\rangle + \langle z_k - c_k \mid D_k^c z_k - D_k^c c_k\rangle\big)\\ &\geqslant \sum_{i\in I}\alpha_i^c\|C_i x_i - C_i a_i\|^2 + \sum_{k\in K}\big(\beta_k^c\|B_k^c y_k - B_k^c b_k\|^2 + \delta_k^c\|D_k^c z_k - D_k^c c_k\|^2\big)\\ &\geqslant \alpha\sum_{i\in I}\|C_i x_i - C_i a_i\|^2 + \alpha\sum_{k\in K}\big(\|B_k^c y_k - B_k^c b_k\|^2 + \|D_k^c z_k - D_k^c c_k\|^2\big)\\ &= \alpha\|\mathbf{C}\mathbf{x} - \mathbf{C}\mathbf{y}\|^2. \end{aligned} \tag{40}$$

(iii)(b): Suppose that $\mathbf{z} \in \operatorname{zer}\mathcal{S}$. We deduce from (i) that $-\mathbf{C}\mathbf{z} \in \mathbf{M}\mathbf{z}$ and from our assumption that $\mathbf{p}^* \in \mathbf{M}\mathbf{p}$. Hence, (ii) implies that $\langle \mathbf{z} - \mathbf{p} \mid \mathbf{p}^* + \mathbf{C}\mathbf{z}\rangle \leqslant 0$. Thus, we infer from (iii)(a) and the Cauchy–Schwarz inequality that

$$\begin{aligned}\langle \mathbf{z} - \mathbf{p} \mid \mathbf{p}^* + \mathbf{C}\mathbf{q}\rangle &= \langle \mathbf{z} - \mathbf{p} \mid \mathbf{p}^* + \mathbf{C}\mathbf{z}\rangle - \langle \mathbf{z} - \mathbf{q} \mid \mathbf{C}\mathbf{z} - \mathbf{C}\mathbf{q}\rangle + \langle \mathbf{p} - \mathbf{q} \mid \mathbf{C}\mathbf{z} - \mathbf{C}\mathbf{q}\rangle\\ &\leqslant -\alpha\|\mathbf{C}\mathbf{z} - \mathbf{C}\mathbf{q}\|^2 + \|\mathbf{p} - \mathbf{q}\|\,\|\mathbf{C}\mathbf{z} - \mathbf{C}\mathbf{q}\|\\ &= (4\alpha)^{-1}\|\mathbf{p} - \mathbf{q}\|^2 - \Big|(2\sqrt{\alpha})^{-1}\|\mathbf{p} - \mathbf{q}\| - \sqrt{\alpha}\|\mathbf{C}\mathbf{z} - \mathbf{C}\mathbf{q}\|\Big|^2\\ &\leqslant (4\alpha)^{-1}\|\mathbf{p} - \mathbf{q}\|^2, \end{aligned} \tag{41}$$

which establishes the claim. □

Next, we solve the saddle form (22) of Problem 1 via successive projections onto the outer approximations constructed in Proposition 2(iii)(b).

Proposition 3. *Consider the setting of Problem 1 and of Definitions 1 and 2, and suppose that* $\operatorname{zer}\mathcal{S} \neq \varnothing$. *Set* $\alpha = \min\{\alpha_i^c, \beta_k^c, \delta_k^c\}_{i\in I, k\in K}$, *let* $\mathbf{x}_0 \in \mathbf{X}$, *let* $\varepsilon \in\,]0,1[$, *and iterate*

$$\begin{array}{l} \text{for } n = 0,1,\ldots\\ \left\lfloor \begin{array}{l} (\mathbf{p}_n, \mathbf{p}_n^*) \in \operatorname{gra}\mathbf{M};\ \mathbf{q}_n \in \mathbf{X};\\ \mathbf{t}_n^* = \mathbf{p}_n^* + \mathbf{C}\mathbf{q}_n;\\ \Delta_n = \langle \mathbf{x}_n - \mathbf{p}_n \mid \mathbf{t}_n^*\rangle - (4\alpha)^{-1}\|\mathbf{p}_n - \mathbf{q}_n\|^2;\\ \text{if } \Delta_n > 0\\ \left\lfloor \begin{array}{l} \lambda_n \in [\varepsilon, 2-\varepsilon];\\ \mathbf{x}_{n+1} = \mathbf{x}_n - (\lambda_n\Delta_n/\|\mathbf{t}_n^*\|^2)\mathbf{t}_n^*; \end{array}\right.\\ \text{else}\\ \left\lfloor \begin{array}{l} \mathbf{x}_{n+1} = \mathbf{x}_n. \end{array}\right. \end{array}\right. \end{array} \tag{42}$$

Then the following hold:

(i) $(\forall \mathbf{z} \in \operatorname{zer}\mathcal{S})(\forall n \in \mathbb{N})\ \|\mathbf{x}_{n+1} - \mathbf{z}\| \leqslant \|\mathbf{x}_n - \mathbf{z}\|$.

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

(iii) *Suppose that* $(\mathbf{t}_n^*)_{n\in\mathbb{N}}$ *is bounded. Then* $\overline{\lim}\,\Delta_n \leqslant 0$.

(iv) *Suppose that* $\mathbf{x}_n - \mathbf{p}_n \rightharpoonup \mathbf{0}$, $\mathbf{p}_n - \mathbf{q}_n \to \mathbf{0}$, *and* $\mathbf{t}_n^* \to \mathbf{0}$. *Then* $(\mathbf{x}_n)_{n\in\mathbb{N}}$ *converges weakly to a point in* $\operatorname{zer}\mathcal{S}$.

Proof. (i) and (ii): Proposition 1(ii) and our assumption ensure that $\operatorname{zer}\mathcal{S}$ is a nonempty closed convex subset of $\mathbf{X}$. Now, for every $n \in \mathbb{N}$, set $\eta_n = (4\alpha)^{-1}\|\mathbf{p}_n - \mathbf{q}_n\|^2 + \langle \mathbf{p}_n \mid \mathbf{t}_n^*\rangle$ and $\mathbf{H}_n = \{\mathbf{x} \in \mathbf{X} \mid \langle \mathbf{x} \mid \mathbf{t}_n^*\rangle \leqslant \eta_n\}$. On the one hand, according to Proposition 2(iii)(b), $(\forall n \in \mathbb{N})\ \operatorname{zer}\mathcal{S} \subset \mathbf{H}_n$. On the other hand, (42) gives $(\forall n \in \mathbb{N})\ \Delta_n = \langle \mathbf{x}_n \mid \mathbf{t}_n^*\rangle - \eta_n$. Altogether, (42) is an instantiation of (A.3). The claims thus follow from Lemma A.4(i) and (ii).

(iii): Set $\mu = \sup_{n\in\mathbb{N}}\|\mathbf{t}_n^*\|$. For every $n \in \mathbb{N}$, if $\Delta_n > 0$, then (42) yields $\Delta_n = \lambda_n^{-1}\|\mathbf{t}_n^*\|\,\|\mathbf{x}_{n+1} - \mathbf{x}_n\| \leqslant \varepsilon^{-1}\mu\|\mathbf{x}_{n+1} - \mathbf{x}_n\|$; otherwise, $\Delta_n \leqslant 0 = \varepsilon^{-1}\mu\|\mathbf{x}_{n+1} - \mathbf{x}_n\|$. We therefore invoke (ii) to get $\overline{\lim}\,\Delta_n \leqslant \lim \varepsilon^{-1}\mu\|\mathbf{x}_{n+1} - \mathbf{x}_n\| = 0$.

(iv): Let $\mathbf{x} \in \mathbf{X}$, let $(k_n)_{n\in\mathbb{N}}$ be a strictly increasing sequence in $\mathbb{N}$, and suppose that $\mathbf{x}_{k_n} \rightharpoonup \mathbf{x}$. Then $\mathbf{p}_{k_n} = (\mathbf{p}_{k_n} - \mathbf{x}_{k_n}) + \mathbf{x}_{k_n} \rightharpoonup \mathbf{x}$. In addition, (42) and Proposition 2(i) imply that $(\mathbf{p}_{k_n}, \mathbf{p}_{k_n}^* + \mathbf{C}\mathbf{p}_{k_n})_{n\in\mathbb{N}}$ lies in $\operatorname{gra}(\mathbf{M} + \mathbf{C}) = \operatorname{gra}\mathcal{S}$. We also note that, since $\mathbf{C}$ is $(1/\alpha)$-Lipschitzian by Proposition 2(iii)(a), (42) yields $\|\mathbf{p}_n^* + \mathbf{C}\mathbf{p}_n\| = \|\mathbf{t}_n^* - \mathbf{C}\mathbf{q}_n + \mathbf{C}\mathbf{p}_n\| \leqslant \|\mathbf{t}_n^*\| + \|\mathbf{C}\mathbf{p}_n - \mathbf{C}\mathbf{q}_n\| \leqslant \|\mathbf{t}_n^*\| + \|\mathbf{p}_n - \mathbf{q}_n\|/\alpha \to 0$. Altogether, since $\mathcal{S}$ is maximally monotone by Proposition 1(i), Bauschke and Combettes [9, proposition 20.38(ii)] yields $\mathbf{x} \in \operatorname{zer}\mathcal{S}$. In turn, Lemma A.4(iii) guarantees that $(\mathbf{x}_n)_{n\in\mathbb{N}}$ converges weakly to a point in $\operatorname{zer}\mathcal{S}$. □

The next outer approximation scheme is a variant of the previous one that guarantees strong convergence to a specific zero of the saddle operator.

Proposition 4. *Consider the setting of Problem 1 and of Definitions 1 and 2, and suppose that* zer $\mathcal{S} \neq \emptyset$. *Define*

$$\Xi :]0,+\infty[\times]0,+\infty[\times\mathbb{R}\times\mathbb{R}\to\mathbb{R}^2$$
$$(\Delta,\tau,\varsigma,\chi)\mapsto\begin{cases}(1,\Delta/\tau), & \text{if } \rho=0;\\ (0,(\Delta+\chi)/\tau), & \text{if } \rho\neq 0 \text{ and } \chi\Delta\geqslant\rho; \quad \text{where } \rho=\tau\varsigma-\chi^2,\\ (1-\chi\Delta/\rho,\varsigma\Delta/\rho), & \text{if } \rho\neq 0 \text{ and } \chi\Delta<\rho,\end{cases} \tag{43}$$

set $\alpha=\min\{\alpha_i^c,\beta_k^c,\delta_k^c\}_{i\in I,k\in K}$, *and let* $\mathbf{x}_0\in\mathbf{X}$. *Iterate*

$$\begin{array}{l} \text{for } n=0,1,\ldots\\ \left\lfloor\begin{array}{l} (\mathbf{p}_n,\mathbf{p}_n^*)\in\text{gra }\mathbf{M};\ \mathbf{q}_n\in\mathbf{X};\\ \mathbf{t}_n^*=\mathbf{p}_n^*+\mathbf{C}\mathbf{q}_n;\\ \Delta_n=\langle\mathbf{x}_n-\mathbf{p}_n\,|\,\mathbf{t}_n^*\rangle-(4\alpha)^{-1}\|\mathbf{p}_n-\mathbf{q}_n\|^2;\\ \text{if } \Delta_n>0\\ \left\lfloor\begin{array}{l} \tau_n=\|\mathbf{t}_n^*\|^2;\ \varsigma_n=\|\mathbf{x}_0-\mathbf{x}_n\|^2;\ \chi_n=\langle\mathbf{x}_0-\mathbf{x}_n\,|\,\mathbf{t}_n^*\rangle;\\ (\kappa_n,\lambda_n)=\Xi(\Delta_n,\tau_n,\varsigma_n,\chi_n);\\ \mathbf{x}_{n+1}=(1-\kappa_n)\mathbf{x}_0+\kappa_n\mathbf{x}_n-\lambda_n\mathbf{t}_n^*;\end{array}\right.\\ \text{else}\\ \left\lfloor\ \mathbf{x}_{n+1}=\mathbf{x}_n.\right.\end{array}\right.\end{array} \tag{44}$$

Then the following hold:

(i) $(\forall n\in\mathbb{N})\ \|\mathbf{x}_n-\mathbf{x}_0\|\leqslant\|\mathbf{x}_{n+1}-\mathbf{x}_0\|\leqslant\|\text{proj}_{\text{zer }\mathcal{S}}\mathbf{x}_0-\mathbf{x}_0\|$.

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

(iii) *Suppose that* $(\mathbf{t}_n^*)_{n\in\mathbb{N}}$ *is bounded. Then* $\overline{\lim}\,\Delta_n\leqslant 0$.

(iv) *Suppose that* $\mathbf{x}_n-\mathbf{p}_n\rightharpoonup\mathbf{0}$, $\mathbf{p}_n-\mathbf{q}_n\to\mathbf{0}$, *and* $\mathbf{t}_n^*\to\mathbf{0}$. *Then* $\mathbf{x}_n\to\text{proj}_{\text{zer }\mathcal{S}}\mathbf{x}_0$.

Proof. Set $(\forall n\in\mathbb{N})\ \eta_n=(4\alpha)^{-1}\|\mathbf{p}_n-\mathbf{q}_n\|^2+\langle\mathbf{p}_n\,|\,\mathbf{t}_n^*\rangle$ and $\mathbf{H}_n=\{\mathbf{x}\in\mathbf{X}\mid\langle\mathbf{x}\,|\,\mathbf{t}_n^*\rangle\leqslant\eta_n\}$. As seen in the proof of Proposition 3, zer $\mathcal{S}$ is a nonempty closed convex subset of $\mathbf{X}$ and, for every $n\in\mathbb{N}$, zer $\mathcal{S}\subset\mathbf{H}_n$ and $\Delta_n=\langle\mathbf{x}_n\,|\,\mathbf{t}_n^*\rangle-\eta_n$. This and (43) make (44) an instance of (A.4).

(i) and (ii): Apply Lemma A.5(i) and (ii).

(iii): Set $\mu=\sup_{n\in\mathbb{N}}\|\mathbf{t}_n^*\|$. Take $n\in\mathbb{N}$. Suppose that $\Delta_n>0$. Then, by construction of $\mathbf{H}_n$, $\text{proj}_{\mathbf{H}_n}\mathbf{x}_n=\mathbf{x}_n-(\Delta_n/\|\mathbf{t}_n^*\|^2)\mathbf{t}_n^*$. This implies that $\Delta_n=\|\mathbf{t}_n^*\|\,\|\text{proj}_{\mathbf{H}_n}\mathbf{x}_n-\mathbf{x}_n\|\leqslant\mu\|\text{proj}_{\mathbf{H}_n}\mathbf{x}_n-\mathbf{x}_n\|$. Next, suppose that $\Delta_n\leqslant 0$. Then $\mathbf{x}_n\in\mathbf{H}_n$ and therefore $\Delta_n\leqslant 0=\mu\|\text{proj}_{\mathbf{H}_n}\mathbf{x}_n-\mathbf{x}_n\|$. Altogether, $(\forall n\in\mathbb{N})\ \Delta_n\leqslant\mu\|\text{proj}_{\mathbf{H}_n}\mathbf{x}_n-\mathbf{x}_n\|$. Consequently, Lemma A.5(ii) yields $\overline{\lim}\,\Delta_n\leqslant 0$.

(iv): Follow the same procedure as in the proof of Proposition 3(iv), invoking Lemma A.5(iii) instead of Lemma A.4(iii). □

3. Asynchronous Block-Iterative Outer Approximation Methods

We exploit the saddle form of Problem 1 described in Definition 1 to obtain splitting algorithms with features ①–⑤. Let us comment on the impact of requirements ①–④.

① For every $i\in I$ and every $k\in K$, each single-valued operator C_i, Q_i, R_i, B_k^c, B_k^ℓ, D_k^c, D_k^ℓ, and L_{ki} must be activated individually via a forward step, whereas each of the set-valued operators A_i, B_k^m, and D_k^m must be activated individually via a backward resolvent step.

② At iteration n, only operators indexed by subgroups $I_n\subset I$ and $K_n\subset K$ of indices need to be involved in the sense that the results of their evaluations are incorporated. This considerably reduces the computational load compared to standard methods, which require the use of all the operators at every iteration. Assumption 2 below regulates the frequency at which the indices should be chosen over time.

③ When an operator is involved at iteration n, its evaluation can be made at a point based on data available at an earlier iteration. This makes it possible to initiate a computation at a given iteration and incorporate its result at a later time. Assumption 3 below controls the lag allowed in the process of using past data.

④ Assumption 1 below describes the range allowed for the various scaling parameters in terms of the cocoercivity and Lipschitz constants of the operators.

Assumption 1. *In the setting of Problem 1, set* $\alpha=\min\{\alpha_i^c,\beta_k^c,\delta_k^c\}_{i\in I,k\in K}$, *let* $\sigma\in]0,+\infty[$ *and* $\varepsilon\in]0,1[$ *be such that*

$$\sigma>1/(4\alpha)\quad\text{and}\quad 1/\varepsilon>\max\{\alpha_i^\ell+\chi+\sigma,\beta_k^\ell+\sigma,\delta_k^\ell+\sigma\}_{i\in I,k\in K},\tag{45}$$

and suppose that the following are satisfied:

[a] *For every* $i\in I$ *and every* $n\in\mathbb{N}$, $\gamma_{i,n}\in[\varepsilon,1/(\alpha_i^\ell+\chi+\sigma)]$.

[b] *For every* $k\in K$ *and every* $n\in\mathbb{N}$, $\mu_{k,n}\in[\varepsilon,1/(\beta_k^\ell+\sigma)]$, $\nu_{k,n}\in[\varepsilon,1/(\delta_k^\ell+\sigma)]$, *and* $\sigma_{k,n}\in[\varepsilon,1/\varepsilon]$.

[c] *For every* $i\in I$, $x_{i,0}\in\mathcal{H}_i$; *for every* $k\in K$, $\{y_{k,0},z_{k,0},v_{k,0}^*\}\subset\mathcal{G}_k$.

Assumption 2. *I and K are finite sets,* $P\in\mathbb{N}$, $(I_n)_{n\in\mathbb{N}}$ *are nonempty subsets of I, and* $(K_n)_{n\in\mathbb{N}}$ *are nonempty subsets of K such that*

$$I_0=I,\quad K_0=K,\quad\text{and}\quad(\forall n\in\mathbb{N})\ \bigcup_{j=n}^{n+P}I_j=I\ \text{ and }\ \bigcup_{j=n}^{n+P}K_j=K.\tag{46}$$

Assumption 3. *I and K are finite sets,* $T\in\mathbb{N}$, *and, for every* $i\in I$ *and every* $k\in K$, $(\pi_i(n))_{n\in\mathbb{N}}$ *and* $(\omega_k(n))_{n\in\mathbb{N}}$ *are sequences in* $\mathbb{N}$ *such that* $(\forall n\in\mathbb{N})$ $n-T\leqslant\pi_i(n)\leqslant n$ *and* $n-T\leqslant\omega_k(n)\leqslant n$.

Our first algorithm is patterned after the abstract geometric outer approximation principle described in Proposition 3. As before, bold letters denote product space elements, for example, $\boldsymbol{x}_n=(x_{i,n})_{i\in I}\in\boldsymbol{\mathcal{H}}$.

Algorithm 1. Consider the setting of Problem 1 and suppose that Assumptions 1–3 are in force. Let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence in $[\varepsilon,2-\varepsilon]$ and iterate

$$
\begin{array}{l}
\text{for } n=0,1,\ldots\\
\left\lfloor
\begin{array}{l}
\text{for every } i\in I_n\\
\left\lfloor
\begin{array}{l}
l_{i,n}^*=Q_i x_{i,\pi_i(n)}+R_i\boldsymbol{x}_{\pi_i(n)}+\sum_{k\in K}L_{ki}^*v_{k,\pi_i(n)}^*;\\
a_{i,n}=J_{\gamma_{i,\pi_i(n)}A_i}\big(x_{i,\pi_i(n)}+\gamma_{i,\pi_i(n)}(s_i^*-l_{i,n}^*-C_ix_{i,\pi_i(n)})\big);\\
a_{i,n}^*=\gamma_{i,\pi_i(n)}^{-1}(x_{i,\pi_i(n)}-a_{i,n})-l_{i,n}^*+Q_ia_{i,n};\\
\xi_{i,n}=\|a_{i,n}-x_{i,\pi_i(n)}\|^2;
\end{array}
\right.\\
\text{for every } i\in I\smallsetminus I_n\\
\left\lfloor a_{i,n}=a_{i,n-1};\ a_{i,n}^*=a_{i,n-1}^*;\ \xi_{i,n}=\xi_{i,n-1};\right.\\
\text{for every } k\in K_n\\
\left\lfloor
\begin{array}{l}
u_{k,n}^*=v_{k,\omega_k(n)}^*-B_k^\ell y_{k,\omega_k(n)};\\
w_{k,n}^*=v_{k,\omega_k(n)}^*-D_k^\ell z_{k,\omega_k(n)};\\
b_{k,n}=J_{\mu_{k,\omega_k(n)}B_k^m}\big(y_{k,\omega_k(n)}+\mu_{k,\omega_k(n)}(u_{k,n}^*-B_k^cy_{k,\omega_k(n)})\big);\\
d_{k,n}=J_{\nu_{k,\omega_k(n)}D_k^m}\big(z_{k,\omega_k(n)}+\nu_{k,\omega_k(n)}(w_{k,n}^*-D_k^cz_{k,\omega_k(n)})\big);\\
e_{k,n}^*=\sigma_{k,\omega_k(n)}\left(\sum_{i\in I}L_{ki}x_{i,\omega_k(n)}-y_{k,\omega_k(n)}-z_{k,\omega_k(n)}-r_k\right)+v_{k,\omega_k(n)}^*;\\
q_{k,n}^*=\mu_{k,\omega_k(n)}^{-1}(y_{k,\omega_k(n)}-b_{k,n})+u_{k,n}^*+B_k^\ell b_{k,n}-e_{k,n}^*;\\
t_{k,n}^*=\nu_{k,\omega_k(n)}^{-1}(z_{k,\omega_k(n)}-d_{k,n})+w_{k,n}^*+D_k^\ell d_{k,n}-e_{k,n}^*;\\
\eta_{k,n}=\|b_{k,n}-y_{k,\omega_k(n)}\|^2+\|d_{k,n}-z_{k,\omega_k(n)}\|^2;\\
e_{k,n}=r_k+b_{k,n}+d_{k,n}-\sum_{i\in I}L_{ki}a_{i,n};
\end{array}
\right.\\
\text{for every } k\in K\smallsetminus K_n\\
\left\lfloor
\begin{array}{l}
b_{k,n}=b_{k,n-1};\ d_{k,n}=d_{k,n-1};\ e_{k,n}^*=e_{k,n-1}^*;\ q_{k,n}^*=q_{k,n-1}^*;\ t_{k,n}^*=t_{k,n-1}^*;\\
\eta_{k,n}=\eta_{k,n-1};\ e_{k,n}=r_k+b_{k,n}+d_{k,n}-\sum_{i\in I}L_{ki}a_{i,n};
\end{array}
\right.
\end{array}
\right.
\end{array}
\tag{47}
$$

$$
\begin{array}{l}
\text{for every } i \in I \\
\quad p^*_{i,n} = a^*_{i,n} + R_i a_n + \sum_{k\in K} L^*_{ki} e^*_{k,n}; \\
\Delta_n = -(4\alpha)^{-1}\left(\sum_{i\in I}\xi_{i,n} + \sum_{k\in K}\eta_{k,n}\right) + \sum_{i\in I}\langle x_{i,n} - a_{i,n} \,|\, p^*_{i,n}\rangle \\
\qquad + \sum_{k\in K}\big(\langle y_{k,n} - b_{k,n} \,|\, q^*_{k,n}\rangle + \langle z_{k,n} - d_{k,n} \,|\, t^*_{k,n}\rangle + \langle e_{k,n} \,|\, v^*_{k,n} - e^*_{k,n}\rangle\big); \\
\textit{if } \Delta_n > 0 \\
\quad \theta_n = \lambda_n\Delta_n \Big/ \left(\sum_{i\in I}\|p^*_{i,n}\|^2 + \sum_{k\in K}\left(\|q^*_{k,n}\|^2 + \|t^*_{k,n}\|^2 + \|e_{k,n}\|^2\right)\right); \\
\quad \text{for every } i \in I \\
\quad\quad x_{i,n+1} = x_{i,n} - \theta_n p^*_{i,n}; \\
\quad \text{for every } k \in K \\
\quad\quad y_{k,n+1} = y_{k,n} - \theta_n q^*_{k,n};\ z_{k,n+1} = z_{k,n} - \theta_n t^*_{k,n};\ v^*_{k,n+1} = v^*_{k,n} - \theta_n e_{k,n}; \\
\text{else} \\
\quad \text{for every } i \in I \\
\quad\quad x_{i,n+1} = x_{i,n}; \\
\quad \text{for every } k \in K \\
\quad\quad y_{k,n+1} = y_{k,n};\ z_{k,n+1} = z_{k,n};\ v^*_{k,n+1} = v^*_{k,n}.
\end{array}
$$

The convergence properties of Algorithm 1 are laid out in the following theorem.

Theorem 1. *Consider the setting of Algorithm 1 and suppose that the dual problem (2) has a solution. Then the following hold:*

(i) *Let $i \in I$. Then $\sum_{n\in\mathbb{N}}\|x_{i,n+1} - x_{i,n}\|^2 < +\infty$.*

(ii) *Let $k \in K$. Then $\sum_{n\in\mathbb{N}}\|y_{k,n+1} - y_{k,n}\|^2 < +\infty$, $\sum_{n\in\mathbb{N}}\|z_{k,n+1} - z_{k,n}\|^2 < +\infty$, and $\sum_{n\in\mathbb{N}}\|v^*_{k,n+1} - v^*_{k,n}\|^2 < +\infty$.*

(iii) *Let $i \in I$ and $k \in K$. Then $x_{i,n} - a_{i,n} \to 0$, $y_{k,n} - b_{k,n} \to 0$, $z_{k,n} - d_{k,n} \to 0$, and $v^*_{k,n} - e^*_{k,n} \to 0$.*

(iv) *There exist a solution $\overline{\boldsymbol{x}}$ to (1) and a solution $\overline{\boldsymbol{v}}^*$ to (2) such that, for every $i \in I$ and every $k \in K$, $x_{i,n} \rightharpoonup \overline{x}_i$, $a_{i,n} \rightharpoonup \overline{x}_i$, and $v^*_{k,n} \rightharpoonup \overline{v}^*_k$. In addition, $(\overline{\boldsymbol{x}}, \overline{\boldsymbol{v}}^*)$ is a Kuhn–Tucker point of Problem 1 in the sense of (23).*

Proof. We use the notation of Definitions 1 and 2. We first observe that zer $\mathcal{S} \neq \varnothing$ by virtue of Proposition 1(iv). Next, let us verify that (47) is a special case of (42). For every $i \in I$, denote by $\overline{\vartheta}_i(n)$ the most recent iteration preceding an iteration n at which the results of the evaluations of the operators A_i, C_i, Q_i, and R_i were incorporated, and by $\vartheta_i(n)$ the iteration at which the corresponding calculations were initiated, that is,

$$\overline{\vartheta}_i(n) = \max\{j \in \mathbb{N} \mid j \leqslant n \ \text{ and } \ i \in I_j\} \quad \text{and} \quad \vartheta_i(n) = \pi_i(\overline{\vartheta}_i(n)). \tag{48}$$

Similarly, we define

$$(\forall k \in K)(\forall n \in \mathbb{N}) \quad \overline{\varrho}_k(n) = \max\{j \in \mathbb{N} \mid j \leqslant n \ \text{ and } \ k \in K_j\} \quad \text{and} \quad \varrho_k(n) = \omega_k(\overline{\varrho}_k(n)). \tag{49}$$

By virtue of (47),

$$(\forall i \in I)(\forall n \in \mathbb{N}) \quad a_{i,n} = a_{i,\overline{\vartheta}_i(n)}, \quad a^*_{i,n} = a^*_{i,\overline{\vartheta}_i(n)}, \quad \xi_{i,n} = \xi_{i,\overline{\vartheta}_i(n)}, \tag{50}$$

and likewise

$$(\forall k \in K)(\forall n \in \mathbb{N}) \quad \begin{cases} b_{k,n} = b_{k,\overline{\varrho}_k(n)}, \ d_{k,n} = d_{k,\overline{\varrho}_k(n)}, \ \eta_{k,n} = \eta_{k,\overline{\varrho}_k(n)} \\ e^*_{k,n} = e^*_{k,\overline{\varrho}_k(n)}, \ q^*_{k,n} = q^*_{k,\overline{\varrho}_k(n)}, \ t^*_{k,n} = t^*_{k,\overline{\varrho}_k(n)}. \end{cases} \tag{51}$$

To proceed further, set

$$(\forall n \in \mathbb{N}) \quad \begin{cases} \mathbf{x}_n = (\boldsymbol{x}_n, \boldsymbol{y}_n, \boldsymbol{z}_n, \boldsymbol{v}^*_n) \\ \mathbf{p}_n = (\boldsymbol{a}_n, \boldsymbol{b}_n, \boldsymbol{d}_n, \boldsymbol{e}^*_n) \\ \mathbf{p}^*_n = (\boldsymbol{p}^*_n - (C_i x_{i,\vartheta_i(n)})_{i\in I}, \boldsymbol{q}^*_n - (B^c_k y_{k,\varrho_k(n)})_{k\in K}, \boldsymbol{t}^*_n - (D^c_k z_{k,\varrho_k(n)})_{k\in K}, \boldsymbol{e}_n) \\ \mathbf{q}_n = ((x_{i,\vartheta_i(n)})_{i\in I}, (y_{k,\varrho_k(n)})_{k\in K}, (z_{k,\varrho_k(n)})_{k\in K}, (e^*_{k,n})_{k\in K}) \\ \mathbf{t}^*_n = (\boldsymbol{p}^*_n, \boldsymbol{q}^*_n, \boldsymbol{t}^*_n, \boldsymbol{e}_n). \end{cases} \tag{52}$$

For every $i \in I$ and every $n \in \mathbb{N}$, it follows from (50), (48), (47), and Bauschke and Combettes [9, proposition 23.2(ii)] that

$$\begin{aligned} a_{i,n}^* - C_i x_{i,\vartheta_i(n)} &= a_{i,\overline{\vartheta}_i(n)}^* - C_i x_{i,\pi_i(\overline{\vartheta}_i(n))} \\ &= \gamma_{i,\pi_i(\overline{\vartheta}_i(n))}^{-1}(x_{i,\pi_i(\overline{\vartheta}_i(n))} - a_{i,\overline{\vartheta}_i(n)}) - l_{i,\overline{\vartheta}_i(n)}^* - C_i x_{i,\pi_i(\overline{\vartheta}_i(n))} + Q_i a_{i,\overline{\vartheta}_i(n)} \\ &\in -s_i^* + A_i a_{i,\overline{\vartheta}_i(n)} + Q_i a_{i,\overline{\vartheta}_i(n)} \\ &= -s_i^* + A_i a_{i,n} + Q_i a_{i,n} \end{aligned} \tag{53}$$

and, therefore, that

$$\begin{aligned} p_{i,n}^* - C_i x_{i,\vartheta_i(n)} &= a_{i,n}^* - C_i x_{i,\vartheta_i(n)} + R_i \boldsymbol{a}_n + \sum_{k\in K} L_{ki}^* e_{k,n}^* \\ &\in -s_i^* + A_i a_{i,n} + Q_i a_{i,n} + R_i \boldsymbol{a}_n + \sum_{k\in K} L_{ki}^* e_{k,n}^*. \end{aligned} \tag{54}$$

Analogously, we invoke (51), (49), and (47) to obtain

$$(\forall k \in K)(\forall n \in \mathbb{N}) \quad q_{k,n}^* - B_k^c y_{k,\varrho_k(n)} \in B_k^m b_{k,n} + B_k^\ell b_{k,n} - e_{k,n}^* \tag{55}$$

and

$$(\forall k \in K)(\forall n \in \mathbb{N}) \quad t_{k,n}^* - D_k^c z_{k,\varrho_k(n)} \in D_k^m d_{k,n} + D_k^\ell d_{k,n} - e_{k,n}^*. \tag{56}$$

In addition, (47) states that

$$(\forall k \in K)(\forall n \in \mathbb{N}) \quad e_{k,n} = r_k + b_{k,n} + d_{k,n} - \sum_{i\in I} L_{ki} a_{i,n}. \tag{57}$$

Hence, using (52) and (38), we deduce that $(\mathbf{p}_n, \mathbf{p}_n^*)_{n\in\mathbb{N}}$ lies in gra $\mathbf{M}$. Next, it results from (52) and (39) that $(\forall n \in \mathbb{N})$ $\mathbf{t}_n^* = \mathbf{p}_n^* + \mathbf{C}\mathbf{q}_n$. Moreover, for every $n \in \mathbb{N}$, (47)–(52) entail that

$$\begin{aligned} \sum_{i\in I} \xi_{i,n} + \sum_{k\in K} \eta_{k,n} &= \sum_{i\in I} \xi_{i,\overline{\vartheta}_i(n)} + \sum_{k\in K} \eta_{k,\overline{\varrho}_k(n)} \\ &= \sum_{i\in I} \|a_{i,\overline{\vartheta}_i(n)} - x_{i,\pi_i(\overline{\vartheta}_i(n))}\|^2 + \sum_{k\in K} \Big(\|b_{k,\overline{\varrho}_k(n)} - y_{k,\omega_k(\overline{\varrho}_k(n))}\|^2 + \|d_{k,\overline{\varrho}_k(n)} - z_{k,\omega_k(\overline{\varrho}_k(n))}\|^2 \Big) \\ &= \sum_{i\in I} \|a_{i,n} - x_{i,\vartheta_i(n)}\|^2 + \sum_{k\in K} \Big(\|b_{k,n} - y_{k,\varrho_k(n)}\|^2 + \|d_{k,n} - z_{k,\varrho_k(n)}\|^2 \Big) \\ &= \|\mathbf{p}_n - \mathbf{q}_n\|^2 \end{aligned} \tag{58}$$

and, in turn, that

$$\Delta_n = \langle \mathbf{x}_n - \mathbf{p}_n \mid \mathbf{t}_n^* \rangle - (4\alpha)^{-1} \|\mathbf{p}_n - \mathbf{q}_n\|^2. \tag{59}$$

To sum up, (47) is an instantiation of (42). Therefore, Proposition 3(ii) asserts that

$$\sum_{n\in\mathbb{N}} \|\mathbf{x}_{n+1} - \mathbf{x}_n\|^2 < +\infty. \tag{60}$$

(i) and (ii): These follow from (60) and (52).

(iii) and (iv): Proposition 3(i) implies that $(\mathbf{x}_n)_{n\in\mathbb{N}}$ is bounded. It therefore results from (52) that

$$(\boldsymbol{x}_n)_{n\in\mathbb{N}}, \ (\boldsymbol{y}_n)_{n\in\mathbb{N}}, \ (\boldsymbol{z}_n)_{n\in\mathbb{N}}, \ \text{and} \ (\boldsymbol{v}_n^*)_{n\in\mathbb{N}} \ \text{are bounded.} \tag{61}$$

Hence, (51), (47), (49), and Assumption 1[b] ensure that

$$(\forall k \in K) \quad (e_{k,n}^*)_{n\in\mathbb{N}} = \left(\sigma_{k,\varrho_k(n)} \left(\sum_{i\in I} L_{ki} x_{i,\varrho_k(n)} - y_{k,\varrho_k(n)} - z_{k,\varrho_k(n)} - r_k \right) + v_{k,\varrho_k(n)}^* \right)_{n\in\mathbb{N}} \ \text{is bounded.} \tag{62}$$

Next, we deduce from (61) and Problem 1[e] that

$$(\forall i \in I) \quad (R_i \boldsymbol{x}_{\vartheta_i(n)})_{n\in\mathbb{N}} \ \text{is bounded.} \tag{63}$$

In turn, it follows from (47), (61), the fact that $(Q_i)_{i\in I}$ and $(C_i)_{i\in I}$ are Lipschitzian, and Assumption 1[a] that

$$(\forall i \in I) \quad \Big(x_{i,\vartheta_i(n)} + \gamma_{i,\vartheta_i(n)} \Big(s_i^* - l_{i,\overline{\vartheta}_i(n)}^* - C_i x_{i,\vartheta_i(n)} \Big) \Big)_{n\in\mathbb{N}} \ \text{is bounded.} \tag{64}$$

An inspection of (50), (47), (48), and Lemma A.1 reveals that

$$(\forall i \in I) \quad (a_{i,n})_{n\in\mathbb{N}} = \Big(J_{\gamma_{i,\vartheta_i(n)}A_i}\Big(x_{i,\vartheta_i(n)} + \gamma_{i,\vartheta_i(n)}\Big(s_i^* - l^*_{i,\overline{\vartheta}_i(n)} - C_i x_{i,\vartheta_i(n)}\Big)\Big)\Big)_{n\in\mathbb{N}} \;\; \text{is bounded.} \tag{65}$$

Hence, we infer from (50), (47), (61), and Assumption 1[a] that

$$(\forall i \in I) \quad (a^*_{i,n})_{n\in\mathbb{N}} \;\; \text{is bounded.} \tag{66}$$

Accordingly, by (47), (61), and Assumption 1[b],

$$(\forall k \in K) \quad \Big(y_{k,\varrho_k(n)} + \mu_{k,\varrho_k(n)}\Big(u^*_{k,\overline{\varrho}_k(n)} - B_k^c y_{k,\varrho_k(n)}\Big)\Big)_{n\in\mathbb{N}} \;\; \text{is bounded.} \tag{67}$$

Therefore, (51), (47), (49), and Lemma A.1 imply that

$$(\forall k \in K) \quad (b_{k,n})_{n\in\mathbb{N}} = \Big(J_{\mu_{k,\varrho_k(n)}B_k^m}\Big(y_{k,\varrho_k(n)} + \mu_{k,\varrho_k(n)}\Big(u^*_{k,\overline{\varrho}_k(n)} - B_k^c y_{k,\varrho_k(n)}\Big)\Big)\Big)_{n\in\mathbb{N}} \;\; \text{is bounded.} \tag{68}$$

Thus, (51), (47), (61), (62), and Assumption 1[b] yield

$$(\boldsymbol{q}^*_n)_{n\in\mathbb{N}} \;\; \text{is bounded.} \tag{69}$$

Likewise,

$$(\boldsymbol{d}_n)_{n\in\mathbb{N}} \;\; \text{and} \;\; (\boldsymbol{t}^*_n)_{n\in\mathbb{N}} \;\; \text{are bounded.} \tag{70}$$

We deduce from (57), (68), (70), and (65) that

$$(\boldsymbol{e}_n)_{n\in\mathbb{N}} \;\; \text{is bounded.} \tag{71}$$

On the other hand, (47), (66), (65), Problem 1[e], and (62) imply that

$$(\boldsymbol{p}^*_n)_{n\in\mathbb{N}} \;\; \text{is bounded.} \tag{72}$$

Hence, we infer from (52) and (69)–(71) that $(\mathbf{t}^*_n)_{n\in\mathbb{N}}$ is bounded. Consequently, (59) and Proposition 3(iii) yield

$$\overline{\lim}\,(\langle \mathbf{x}_n - \mathbf{p}_n \,|\, \mathbf{t}^*_n\rangle - (4\alpha)^{-1}\|\mathbf{p}_n - \mathbf{q}_n\|^2) = \overline{\lim}\,\Delta_n \leqslant 0. \tag{73}$$

Let $\boldsymbol{L}$ and $\mathbf{W}$ be as in (24) and (27). For every $n \in \mathbb{N}$, set

$$\begin{cases} (\forall i \in I)\;\; E_{i,n} = \gamma^{-1}_{i,\vartheta_i(n)}\mathrm{Id} - Q_i \\ (\forall k \in K)\;\; F_{k,n} = \mu^{-1}_{k,\varrho_k(n)}\mathrm{Id} - B_k^\ell, \;\; G_{k,n} = \nu^{-1}_{k,\varrho_k(n)}\mathrm{Id} - D_k^\ell \\ \mathbf{E}_n : \mathbf{X} \to \mathbf{X} : (\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}, \boldsymbol{v}^*) \mapsto ((E_{i,n}x_i)_{i\in I}, (F_{k,n}y_k)_{k\in K}, (G_{k,n}z_k)_{k\in K}, (\sigma^{-1}_{k,\varrho_k(n)}v^*_k)_{k\in K}) \end{cases} \tag{74}$$

and

$$\begin{cases} \widetilde{\mathbf{x}}_n = ((x_{i,\vartheta_i(n)})_{i\in I}, (y_{k,\varrho_k(n)})_{k\in K}, (z_{k,\varrho_k(n)})_{k\in K}, (v^*_{k,\varrho_k(n)})_{k\in K}) \\ \mathbf{v}^*_n = \mathbf{E}_n\mathbf{x}_n - \mathbf{E}_n\mathbf{p}_n, \; \mathbf{w}^*_n = \mathbf{W}\mathbf{p}_n - \mathbf{W}\mathbf{x}_n \\ \mathbf{r}^*_n = ((R_i\boldsymbol{a}_n - R_i\boldsymbol{x}_n)_{i\in I}, \mathbf{0}, \mathbf{0}, \mathbf{0}), \;\; \widetilde{\mathbf{r}}^*_n = ((R_i\boldsymbol{a}_n - R_i\boldsymbol{x}_{\vartheta_i(n)})_{i\in I}, \mathbf{0}, \mathbf{0}, \mathbf{0}) \\ \mathbf{l}^*_n = \left(\left(-\sum_{k\in K} L^*_{ki}v^*_{k,\vartheta_i(n)}\right)_{i\in I}, \left(v^*_{k,\varrho_k(n)}\right)_{k\in K}, \left(v^*_{k,\varrho_k(n)}\right)_{k\in K}, \left(\sum_{i\in I} L_{ki}x_{i,\varrho_k(n)} - y_{k,\varrho_k(n)} - z_{k,\varrho_k(n)}\right)_{k\in K}\right). \end{cases} \tag{75}$$

In view of Problem 1[a]–[c] and Assumption 1[a] and [b], we deduce from Lemma A.2 that

$$(\forall n \in \mathbb{N}) \quad \begin{cases} \text{the operators } \;(E_{i,n})_{i\in I}\; \text{ are } (\chi+\sigma)\text{-strongly monotone} \\ \text{the operators } \;(F_{k,n})_{k\in K}\; \text{ and } \;(G_{k,n})_{k\in K}\; \text{ are } \sigma\text{-strongly monotone,} \end{cases} \tag{76}$$

and from (74) that there exists $\kappa \in \left]0, +\infty\right[$ such that

$$\text{the operators } \;(\mathbf{E}_n)_{n\in\mathbb{N}}\; \text{ are } \kappa\text{-Lipschitzian.} \tag{77}$$

It results from (50), (47), (48), and (74) that

$$\begin{aligned} (\forall i \in I)(\forall n \in \mathbb{N}) \quad a^*_{i,n} &= a^*_{i,\overline{\vartheta}_i(n)} \\ &= \Big(\gamma^{-1}_{i,\pi_i(\overline{\vartheta}_i(n))}x_{i,\pi_i(\overline{\vartheta}_i(n))} - Q_i x_{i,\pi_i(\overline{\vartheta}_i(n))}\Big) - \Big(\gamma^{-1}_{i,\pi_i(\overline{\vartheta}_i(n))}a_{i,\overline{\vartheta}_i(n)} - Q_i a_{i,\overline{\vartheta}_i(n)}\Big) - R_i\boldsymbol{x}_{\pi_i(\overline{\vartheta}_i(n))} - \sum_{k\in K} L^*_{ki}v^*_{k,\pi_i(\overline{\vartheta}_i(n))} \\ &= E_{i,n}x_{i,\vartheta_i(n)} - E_{i,n}a_{i,n} - R_i\boldsymbol{x}_{\vartheta_i(n)} - \sum_{k\in K} L^*_{ki}v^*_{k,\vartheta_i(n)} \end{aligned} \tag{78}$$

and, therefore, that

$$(\forall i \in I)(\forall n \in \mathbb{N}) \quad p_{i,n}^* = a_{i,n}^* + R_i \boldsymbol{a}_n + \sum_{k\in K} L_{ki}^* e_{k,n}^* \\ = E_{i,n} x_{i,\vartheta_i(n)} - E_{i,n} a_{i,n} + R_i \boldsymbol{a}_n - R_i \boldsymbol{x}_{\vartheta_i(n)} - \sum_{k\in K} L_{ki}^* v_{k,\vartheta_i(n)}^* + \sum_{k\in K} L_{ki}^* e_{k,n}^*. \tag{79}$$

At the same time, (51), (47), (49), and (74) entail that

$$(\forall k \in K)(\forall n \in \mathbb{N}) \quad q_{k,n}^* = q_{k,\overline{\varrho}_k(n)}^* \\ = \Big(\mu_{k,\omega_k(\overline{\varrho}_k(n))}^{-1} y_{k,\omega_k(\overline{\varrho}_k(n))} - B_k^\ell y_{k,\omega_k(\overline{\varrho}_k(n))}\Big) - \Big(\mu_{k,\omega_k(\overline{\varrho}_k(n))}^{-1} b_{k,\overline{\varrho}_k(n)} - B_k^\ell b_{k,\overline{\varrho}_k(n)}\Big) + v_{k,\omega_k(\overline{\varrho}_k(n))}^* - e_{k,\overline{\varrho}_k(n)}^* \\ = F_{k,n} y_{k,\varrho_k(n)} - F_{k,n} b_{k,n} + v_{k,\varrho_k(n)}^* - e_{k,n}^* \tag{80}$$

and that

$$(\forall k \in K)(\forall n \in \mathbb{N}) \quad t_{k,n}^* = G_{k,n} z_{k,\varrho_k(n)} - G_{k,n} d_{k,n} + v_{k,\varrho_k(n)}^* - e_{k,n}^*. \tag{81}$$

Furthermore, we derive from (51), (47), and (49) that

$$(\forall k \in K)(\forall n \in \mathbb{N}) \quad r_k = \sigma_{k,\varrho_k(n)}^{-1} v_{k,\varrho_k(n)}^* - \sigma_{k,\varrho_k(n)}^{-1} e_{k,n}^* - y_{k,\varrho_k(n)} - z_{k,\varrho_k(n)} + \sum_{i\in I} L_{ki} x_{i,\varrho_k(n)} \tag{82}$$

and, in turn, from (57) that

$$(\forall k \in K)(\forall n \in \mathbb{N}) \quad e_{k,n} = \sigma_{k,\varrho_k(n)}^{-1} v_{k,\varrho_k(n)}^* - \sigma_{k,\varrho_k(n)}^{-1} e_{k,n}^* - y_{k,\varrho_k(n)} - z_{k,\varrho_k(n)} \\ + \sum_{i\in I} L_{ki} x_{i,\varrho_k(n)} + b_{k,n} + d_{k,n} - \sum_{i\in I} L_{ki} a_{i,n}. \tag{83}$$

Altogether, it follows from (52), (79)–(81), (83), (74), (75), (27), and (25) that

$$(\forall n \in \mathbb{N}) \quad \mathbf{t}_n^* = \mathbf{E}_n \widetilde{\mathbf{x}}_n - \mathbf{E}_n \mathbf{p}_n + \widetilde{\mathbf{r}}_n^* + \mathbf{l}_n^* + \mathbf{W}\mathbf{p}_n. \tag{84}$$

Next, in view of (60), (48), (49), and Assumptions 2 and 3, we learn from Lemma A.3 that

$$(\forall i \in I)(\forall k \in K) \quad \begin{cases} \boldsymbol{x}_{\vartheta_i(n)} - \boldsymbol{x}_n \to \mathbf{0}, \ \boldsymbol{x}_{\varrho_k(n)} - \boldsymbol{x}_n \to \mathbf{0}, \ \text{and} \ \boldsymbol{v}_{\vartheta_i(n)}^* - \boldsymbol{v}_n^* \to \mathbf{0} \\ \boldsymbol{y}_{\varrho_k(n)} - \boldsymbol{y}_n \to \mathbf{0}, \ \boldsymbol{z}_{\varrho_k(n)} - \boldsymbol{z}_n \to \mathbf{0}, \ \text{and} \ \boldsymbol{v}_{\varrho_k(n)}^* - \boldsymbol{v}_n^* \to \mathbf{0}. \end{cases} \tag{85}$$

Thus, (75), (27), (25), and (24) yield

$$\mathbf{l}_n^* + \mathbf{W}\mathbf{x}_n \to \mathbf{0}, \tag{86}$$

while Problem 1[e] gives

$$(\forall i \in I) \quad \|R_i \boldsymbol{x}_{\vartheta_i(n)} - R_i \boldsymbol{x}_n\| \leqslant \chi \|\boldsymbol{x}_{\vartheta_i(n)} - \boldsymbol{x}_n\| \to 0. \tag{87}$$

On the other hand, we infer from (77), (75), and (85) that

$$\|\mathbf{E}_n \widetilde{\mathbf{x}}_n - \mathbf{E}_n \mathbf{x}_n\| \leqslant \kappa \|\widetilde{\mathbf{x}}_n - \mathbf{x}_n\| \to 0. \tag{88}$$

Combining (84), (75), and (86)–(88), we obtain

$$\mathbf{t}_n^* - (\mathbf{v}_n^* + \mathbf{r}_n^* + \mathbf{w}_n^*) = \mathbf{l}_n^* + \mathbf{W}\mathbf{x}_n + \mathbf{E}_n \widetilde{\mathbf{x}}_n - \mathbf{E}_n \mathbf{x}_n + \widetilde{\mathbf{r}}_n^* - \mathbf{r}_n^* \to \mathbf{0}. \tag{89}$$

Now set

$$(\forall n \in \mathbb{N}) \quad \widetilde{\mathbf{q}}_n = (\boldsymbol{x}_n, \boldsymbol{y}_n, \boldsymbol{z}_n, \boldsymbol{e}_n^*). \tag{90}$$

Then $(\widetilde{\mathbf{q}}_n)_{n\in\mathbb{N}}$ is bounded by virtue of (61) and (62). On the one hand, (52), (62), (65), (68), and (70) imply that $(\mathbf{p}_n)_{n\in\mathbb{N}}$ is bounded. On the other hand, (52) and (85) give

$$\widetilde{\mathbf{q}}_n - \mathbf{q}_n \to \mathbf{0}. \tag{91}$$

Therefore, appealing to the Cauchy–Schwarz inequality, we obtain

$$|\langle \mathbf{p}_n - \widetilde{\mathbf{q}}_n \,|\, \widetilde{\mathbf{q}}_n - \mathbf{q}_n\rangle| \leqslant \Big(\sup_{m\in\mathbb{N}} \|\mathbf{p}_m\| + \sup_{m\in\mathbb{N}} \|\widetilde{\mathbf{q}}_m\|\Big)\|\widetilde{\mathbf{q}}_n - \mathbf{q}_n\| \to 0 \tag{92}$$

and, by (89),

$$|\langle \mathbf{x}_n - \mathbf{p}_n \,|\, \mathbf{t}_n^* - (\mathbf{v}_n^* + \mathbf{r}_n^* + \mathbf{w}_n^*)\rangle| \leqslant \Big(\sup_{m\in\mathbb{N}} \|\mathbf{x}_m\| + \sup_{m\in\mathbb{N}} \|\mathbf{p}_m\|\Big)\|\mathbf{t}_n^* - (\mathbf{v}_n^* + \mathbf{r}_n^* + \mathbf{w}_n^*)\| \to 0. \tag{93}$$

However, since $\mathbf{W}^* = -\mathbf{W}$ by (27), it results from (75) that $(\forall n \in \mathbb{N})$ $\langle \mathbf{x}_n - \mathbf{p}_n \,|\, \mathbf{w}_n^*\rangle = 0$. Thus, by (73) and (91)–(93),

$$\begin{aligned}
0 &\geqslant \overline{\lim}\,(\langle \mathbf{x}_n - \mathbf{p}_n \,|\, \mathbf{t}_n^*\rangle - (4\alpha)^{-1}\|\mathbf{p}_n - \mathbf{q}_n\|^2)\\
&= \overline{\lim}\,(\langle \mathbf{x}_n - \mathbf{p}_n \,|\, \mathbf{v}_n^* + \mathbf{r}_n^* + \mathbf{w}_n^*\rangle + \langle \mathbf{x}_n - \mathbf{p}_n \,|\, \mathbf{t}_n^* - (\mathbf{v}_n^* + \mathbf{r}_n^* + \mathbf{w}_n^*)\rangle - (4\alpha)^{-1}\|\mathbf{p}_n - \mathbf{q}_n\|^2)\\
&= \overline{\lim}\,(\langle \mathbf{x}_n - \mathbf{p}_n \,|\, \mathbf{v}_n^* + \mathbf{r}_n^*\rangle - (4\alpha)^{-1}(\|\mathbf{p}_n - \widetilde{\mathbf{q}}_n\|^2 + 2\langle \mathbf{p}_n - \widetilde{\mathbf{q}}_n \,|\, \widetilde{\mathbf{q}}_n - \mathbf{q}_n\rangle + \|\widetilde{\mathbf{q}}_n - \mathbf{q}_n\|^2))\\
&= \overline{\lim}\,(\langle \mathbf{x}_n - \mathbf{p}_n \,|\, \mathbf{v}_n^* + \mathbf{r}_n^*\rangle - (4\alpha)^{-1}\|\mathbf{p}_n - \widetilde{\mathbf{q}}_n\|^2).
\end{aligned} \tag{94}$$

On the other hand, we deduce from (75), (52), (74), (76), Assumption 1[b], the Cauchy–Schwarz inequality, Problem 1[e], and (90) that, for every $n \in \mathbb{N}$,

$$\begin{aligned}
&\langle \mathbf{x}_n - \mathbf{p}_n \,|\, \mathbf{v}_n^* + \mathbf{r}_n^*\rangle - (4\alpha)^{-1}\|\mathbf{p}_n - \widetilde{\mathbf{q}}_n\|^2\\
&= \langle \mathbf{x}_n - \mathbf{p}_n \,|\, \mathbf{E}_n\mathbf{x}_n - \mathbf{E}_n\mathbf{p}_n\rangle + \langle \mathbf{x}_n - \mathbf{p}_n \,|\, \mathbf{r}_n^*\rangle - (4\alpha)^{-1}\|\mathbf{p}_n - \widetilde{\mathbf{q}}_n\|^2\\
&= \sum_{i\in I} \langle x_{i,n} - a_{i,n} \,|\, E_{i,n}x_{i,n} - E_{i,n}a_{i,n}\rangle + \sum_{k\in K} \langle y_{k,n} - b_{k,n} \,|\, F_{k,n}y_{k,n} - F_{k,n}b_{k,n}\rangle\\
&\quad + \sum_{k\in K} \langle z_{k,n} - d_{k,n} \,|\, G_{k,n}z_{k,n} - G_{k,n}d_{k,n}\rangle + \sum_{k\in K} \sigma_{k,\varrho_k(n)}^{-1}\|v_{k,n}^* - e_{k,n}^*\|^2\\
&\quad + \langle \boldsymbol{x}_n - \boldsymbol{a}_n \,|\, \boldsymbol{R}\boldsymbol{a}_n - \boldsymbol{R}\boldsymbol{x}_n\rangle - (4\alpha)^{-1}\|\mathbf{p}_n - \widetilde{\mathbf{q}}_n\|^2\\
&\geqslant (\chi + \sigma)\|\boldsymbol{x}_n - \boldsymbol{a}_n\|^2 + \sigma\|\boldsymbol{y}_n - \boldsymbol{b}_n\|^2 + \sigma\|\boldsymbol{z}_n - \boldsymbol{d}_n\|^2\\
&\quad + \varepsilon\|\boldsymbol{v}_n^* - \boldsymbol{e}_n^*\|^2 - \|\boldsymbol{x}_n - \boldsymbol{a}_n\|\,\|\boldsymbol{R}\boldsymbol{a}_n - \boldsymbol{R}\boldsymbol{x}_n\| - (4\alpha)^{-1}\|\mathbf{p}_n - \widetilde{\mathbf{q}}_n\|^2\\
&\geqslant (\chi + \sigma)\|\boldsymbol{x}_n - \boldsymbol{a}_n\|^2 + \sigma\|\boldsymbol{y}_n - \boldsymbol{b}_n\|^2 + \sigma\|\boldsymbol{z}_n - \boldsymbol{d}_n\|^2\\
&\quad + \varepsilon\|\boldsymbol{v}_n^* - \boldsymbol{e}_n^*\|^2 - \chi\|\boldsymbol{x}_n - \boldsymbol{a}_n\|^2 - (4\alpha)^{-1}\|\mathbf{p}_n - \widetilde{\mathbf{q}}_n\|^2\\
&= (\sigma - (4\alpha)^{-1})(\|\boldsymbol{x}_n - \boldsymbol{a}_n\|^2 + \|\boldsymbol{y}_n - \boldsymbol{b}_n\|^2 + \|\boldsymbol{z}_n - \boldsymbol{d}_n\|^2) + \varepsilon\|\boldsymbol{v}_n^* - \boldsymbol{e}_n^*\|^2.
\end{aligned} \tag{95}$$

Hence, since $\sigma > 1/(4\alpha)$ by (45), taking the limit superior in (95) and invoking (94) yields

$$\boldsymbol{x}_n - \boldsymbol{a}_n \to \mathbf{0},\ \ \boldsymbol{y}_n - \boldsymbol{b}_n \to \mathbf{0},\ \ \boldsymbol{z}_n - \boldsymbol{d}_n \to \mathbf{0},\ \text{ and }\ \boldsymbol{v}_n^* - \boldsymbol{e}_n^* \to \mathbf{0}, \tag{96}$$

which establishes (iii). In turn, (52) and (77) force

$$\mathbf{x}_n - \mathbf{p}_n \to \mathbf{0} \quad\text{and}\quad \|\mathbf{E}_n\mathbf{x}_n - \mathbf{E}_n\mathbf{p}_n\| \leqslant \kappa\|\mathbf{x}_n - \mathbf{p}_n\| \to 0, \tag{97}$$

and (85) thus yields $\mathbf{p}_n - \mathbf{q}_n \to \mathbf{0}$. Furthermore, we infer from (75), (96), and Problem 1[e] that

$$\|\mathbf{r}_n^*\|^2 = \|\boldsymbol{R}\boldsymbol{a}_n - \boldsymbol{R}\boldsymbol{x}_n\|^2 \leqslant \chi^2\|\boldsymbol{a}_n - \boldsymbol{x}_n\|^2 \to 0. \tag{98}$$

Altogether, it follows from (75), (89), (97), and (98) that

$$\mathbf{t}_n^* = (\mathbf{t}_n^* - (\mathbf{v}_n^* + \mathbf{r}_n^* + \mathbf{w}_n^*)) + (\mathbf{E}_n\mathbf{x}_n - \mathbf{E}_n\mathbf{p}_n) + \mathbf{W}(\mathbf{p}_n - \mathbf{x}_n) + \mathbf{r}_n^* \to \mathbf{0}. \tag{99}$$

Hence, Proposition 3(iv) guarantees that there exists $\overline{\mathbf{x}} = (\overline{\boldsymbol{x}}, \overline{\boldsymbol{y}}, \overline{\boldsymbol{z}}, \overline{\boldsymbol{v}}^*) \in \operatorname{zer} \boldsymbol{\mathcal{S}}$ such that $\mathbf{x}_n \rightharpoonup \overline{\mathbf{x}}$. This and (96) imply that, for every $i \in I$ and every $k \in K$, $x_{i,n} \rightharpoonup \overline{x}_i$, $a_{i,n} \rightharpoonup \overline{x}_i$, and $v_{k,n}^* \rightharpoonup \overline{v}_k^*$. Finally, Proposition 1(iii) asserts that $(\overline{\boldsymbol{x}}, \overline{\boldsymbol{v}}^*)$ lies in the set of Kuhn–Tucker points (23), that $\overline{\boldsymbol{x}}$ solves (1), and that $\overline{\boldsymbol{v}}^*$ solves (2). □

Some infinite-dimensional applications require strong convergence of the iterates; see, for example, Attouch et al. [3, 4]. This will be guaranteed by the following variant of Algorithm 1, which hinges on the principle outlined in Proposition 4.

Algorithm 2. Consider the setting of Problem 1, define Ξ as in (43), and suppose that Assumptions 1–3 are in force. Iterate

$$
\begin{array}{l}
\text{for } n = 0, 1, \ldots \\
\quad \text{for every } i \in I_n \\
\quad\quad l_{i,n}^* = Q_i x_{i,\pi_i(n)} + R_i \boldsymbol{x}_{\pi_i(n)} + \sum_{k\in K} L_{ki}^* v_{k,\pi_i(n)}^*; \\
\quad\quad a_{i,n} = J_{\gamma_{i,\pi_i(n)} A_i}\big(x_{i,\pi_i(n)} + \gamma_{i,\pi_i(n)}(s_i^* - l_{i,n}^* - C_i x_{i,\pi_i(n)})\big); \\
\quad\quad a_{i,n}^* = \gamma_{i,\pi_i(n)}^{-1}(x_{i,\pi_i(n)} - a_{i,n}) - l_{i,n}^* + Q_i a_{i,n}; \\
\quad\quad \xi_{i,n} = \|a_{i,n} - x_{i,\pi_i(n)}\|^2; \\
\quad \text{for every } i \in I\backslash I_n \\
\quad\quad a_{i,n} = a_{i,n-1};\ a_{i,n}^* = a_{i,n-1}^*;\ \xi_{i,n} = \xi_{i,n-1}; \\
\quad \text{for every } k \in K_n \\
\quad\quad u_{k,n}^* = v_{k,\omega_k(n)}^* - B_k^{\ell} y_{k,\omega_k(n)}; \\
\quad\quad w_{k,n}^* = v_{k,\omega_k(n)}^* - D_k^{\ell} z_{k,\omega_k(n)}; \\
\quad\quad b_{k,n} = J_{\mu_{k,\omega_k(n)} B_k^{m}}\big(y_{k,\omega_k(n)} + \mu_{k,\omega_k(n)}(u_{k,n}^* - B_k^{c} y_{k,\omega_k(n)})\big); \\
\quad\quad d_{k,n} = J_{\nu_{k,\omega_k(n)} D_k^{m}}\big(z_{k,\omega_k(n)} + \nu_{k,\omega_k(n)}(w_{k,n}^* - D_k^{c} z_{k,\omega_k(n)})\big); \\
\quad\quad e_{k,n}^* = \sigma_{k,\omega_k(n)}\Big(\sum_{i\in I} L_{ki} x_{i,\omega_k(n)} - y_{k,\omega_k(n)} - z_{k,\omega_k(n)} - r_k\Big) + v_{k,\omega_k(n)}^*; \\
\quad\quad q_{k,n}^* = \mu_{k,\omega_k(n)}^{-1}(y_{k,\omega_k(n)} - b_{k,n}) + u_{k,n}^* + B_k^{\ell} b_{k,n} - e_{k,n}^*; \\
\quad\quad t_{k,n}^* = \nu_{k,\omega_k(n)}^{-1}(z_{k,\omega_k(n)} - d_{k,n}) + w_{k,n}^* + D_k^{\ell} d_{k,n} - e_{k,n}^*; \\
\quad\quad \eta_{k,n} = \|b_{k,n} - y_{k,\omega_k(n)}\|^2 + \|d_{k,n} - z_{k,\omega_k(n)}\|^2; \\
\quad\quad e_{k,n} = r_k + b_{k,n} + d_{k,n} - \sum_{i\in I} L_{ki} a_{i,n}; \\
\quad \text{for every } k \in K\backslash K_n \\
\quad\quad b_{k,n} = b_{k,n-1};\ d_{k,n} = d_{k,n-1};\ e_{k,n}^* = e_{k,n-1}^*;\ q_{k,n}^* = q_{k,n-1}^*;\ t_{k,n}^* = t_{k,n-1}^*; \\
\quad\quad \eta_{k,n} = \eta_{k,n-1};\ e_{k,n} = r_k + b_{k,n} + d_{k,n} - \sum_{i\in I} L_{ki} a_{i,n}; \\
\quad \text{for every } i \in I \\
\quad\quad p_{i,n}^* = a_{i,n}^* + R_i \boldsymbol{a}_n + \sum_{k\in K} L_{ki}^* e_{k,n}^*; \\
\quad \Delta_n = -(4\alpha)^{-1}\Big(\sum_{i\in I}\xi_{i,n} + \sum_{k\in K}\eta_{k,n}\Big) + \sum_{i\in I}\langle x_{i,n} - a_{i,n} \mid p_{i,n}^*\rangle \\
\quad\quad\quad + \sum_{k\in K}\big(\langle y_{k,n} - b_{k,n} \mid q_{k,n}^*\rangle + \langle z_{k,n} - d_{k,n} \mid t_{k,n}^*\rangle + \langle e_{k,n} \mid v_{k,n}^* - e_{k,n}^*\rangle\big); \\
\quad \text{if } \Delta_n > 0 \\
\quad\quad \tau_n = \sum_{i\in I}\|p_{i,n}^*\|^2 + \sum_{k\in K}\big(\|q_{k,n}^*\|^2 + \|t_{k,n}^*\|^2 + \|e_{k,n}\|^2\big); \\
\quad\quad \varsigma_n = \sum_{i\in I}\|x_{i,0} - x_{i,n}\|^2 \\
\quad\quad\quad + \sum_{k\in K}\big(\|y_{k,0} - y_{k,n}\|^2 + \|z_{k,0} - z_{k,n}\|^2 + \|v_{k,0}^* - v_{k,n}^*\|^2\big); \\
\quad\quad \chi_n = \sum_{i\in I}\langle x_{i,0} - x_{i,n} \mid p_{i,n}^*\rangle \\
\quad\quad\quad + \sum_{k\in K}\big(\langle y_{k,0} - y_{k,n} \mid q_{k,n}^*\rangle + \langle z_{k,0} - z_{k,n} \mid t_{k,n}^*\rangle + \langle e_{k,n} \mid v_{k,0}^* - v_{k,n}^*\rangle\big); \\
\quad\quad (\kappa_n, \lambda_n) = \Xi(\Delta_n, \tau_n, \varsigma_n, \chi_n); \\
\quad\quad \text{for every } i \in I \\
\quad\quad\quad x_{i,n+1} = (1-\kappa_n)x_{i,0} + \kappa_n x_{i,n} - \lambda_n p_{i,n}^*; \\
\quad\quad \text{for every } k \in K \\
\quad\quad\quad y_{k,n+1} = (1-\kappa_n)y_{k,0} + \kappa_n y_{k,n} - \lambda_n q_{k,n}^*; \\
\quad\quad\quad z_{k,n+1} = (1-\kappa_n)z_{k,0} + \kappa_n z_{k,n} - \lambda_n t_{k,n}^*; \\
\quad\quad\quad v_{k,n+1}^* = (1-\kappa_n)v_{k,0}^* + \kappa_n v_{k,n}^* - \lambda_n e_{k,n}; \\
\quad \text{else} \\
\quad\quad \text{for every } i \in I \\
\quad\quad\quad x_{i,n+1} = x_{i,n}; \\
\quad\quad \text{for every } k \in K \\
\quad\quad\quad y_{k,n+1} = y_{k,n};\ z_{k,n+1} = z_{k,n};\ v_{k,n+1}^* = v_{k,n}^*.
\end{array}
\tag{100}
$$

Theorem 2. *Consider the setting of Algorithm 2 and suppose that the dual problem (2) has a solution. Then the following hold:*

(i) *Let $i \in I$. Then $\sum_{n\in\mathbb{N}}\|x_{i,n+1}-x_{i,n}\|^2 < +\infty$.*

(ii) *Let $k \in K$. Then $\sum_{n\in\mathbb{N}}\|y_{k,n+1}-y_{k,n}\|^2 < +\infty$, $\sum_{n\in\mathbb{N}}\|z_{k,n+1}-z_{k,n}\|^2 < +\infty$, and $\sum_{n\in\mathbb{N}}\|v^*_{k,n+1}-v^*_{k,n}\|^2 < +\infty$.*

(iii) *Let $i \in I$ and $k \in K$. Then $x_{i,n}-a_{i,n} \to 0$, $y_{k,n}-b_{k,n} \to 0$, $z_{k,n}-d_{k,n} \to 0$, and $v^*_{k,n}-e^*_{k,n} \to 0$.*

(iv) *There exist a solution $\overline{\boldsymbol{x}}$ to (1) and a solution $\overline{\boldsymbol{v}}^*$ to (2) such that, for every $i \in I$ and every $k \in K$, $x_{i,n} \to \overline{x}_i$, $a_{i,n} \to \overline{x}_i$, and $v^*_{k,n} \to \overline{v}^*_k$. In addition, $(\overline{\boldsymbol{x}},\overline{\boldsymbol{v}}^*)$ is a Kuhn–Tucker point of Problem 1 in the sense of (23).*

Proof. Proceed as in the proof of Theorem 1 and use Proposition 4 instead of Proposition 3. □

4. Applications

In nonlinear analysis and optimization, problems with multiple variables occur in areas such as game theory (Attouch et al. [2], Briceño-Arias and Combettes [15], Yi and Pavel [56]), evolution inclusions (Attouch et al. [3]), traffic equilibrium (Attouch et al. [3], Fukushima [31]), domain decomposition (Attouch et al. [4]), machine learning (Bach et al. [6], Briceño-Arias et al. [12]), image recovery (Briceño-Arias and Combettes [13], Briceño-Arias et al. [16]), infimal-convolution regularization (Combettes [23]), statistics (Combettes and Müller [26], Yan and Bien [55]), neural networks (Combettes and Pesquet [27]), and variational inequalities (Fukushima [31]). The numerical methods used in these papers are limited to special cases of Problem 1, they do not perform block iterations, and they operate in synchronous mode. The methods presented in Theorems 1 and 2 provide a unified treatment of these problems and extensions, within a considerably more flexible algorithmic framework. In this section, we illustrate this in the context of variational inequalities and multivariate minimization. Below we present only the applications of Theorem 1, as similar applications of Theorem 2 follow using similar arguments.

4.1. Application to Variational Inequalities

The standard variational inequality problem associated with a closed convex subset D of a real Hilbert space $\mathcal{G}$ and a maximally monotone operator $B:\mathcal{G}\to\mathcal{G}$ is to

$$\text{find } \overline{y} \in D \text{ such that } (\forall y \in D)\ \langle \overline{y}-y \mid B\overline{y}\rangle \leqslant 0. \tag{101}$$

Classical methods require the ability to project onto D and specific assumptions on B such as cocoercivity, Lipschitz continuity, or the ability to compute the resolvent (Bauschke and Combettes [9], Facchinei and Pang [30], Tseng [53]). Let us consider a refined version of (101) in which B and D are decomposed into basic components and for which these classical methods are not applicable.

Problem 2. Let I be a nonempty finite set and let $(\mathcal{H}_i)_{i\in I}$ and $\mathcal{G}$ be real Hilbert spaces. For every $i \in I$, let E_i and F_i be closed convex subsets of $\mathcal{H}_i$ such that $E_i \cap F_i \neq \emptyset$ and let $L_i:\mathcal{H}_i\to\mathcal{G}$ be linear and bounded. In addition, let $B^m:\mathcal{G}\to 2^{\mathcal{G}}$ be at most single-valued and maximally monotone, let $B^c:\mathcal{G}\to\mathcal{G}$ be cocoercive with constant $\beta^c \in\,]0,+\infty[$, and let $B^\ell:\mathcal{G}\to\mathcal{G}$ be Lipschitzian with constant $\beta^\ell \in [0,+\infty[$. The objective is to

$$\text{find } \overline{y} \in \sum_{i\in I} L_i(E_i\cap F_i) \text{ such that } \left(\forall y \in \sum_{i\in I} L_i(E_i\cap F_i)\right)\ \langle \overline{y}-y \mid B^m\overline{y}+B^c\overline{y}+B^\ell\overline{y}\rangle \leqslant 0. \tag{102}$$

To motivate our analysis, let us consider an illustration of (102).

Example 1. Let I be a nonempty finite set and let $(\mathcal{Z}_i)_{i\in I}$ and $\mathcal{K}$ be real Hilbert spaces. For every $i \in I$, let $S_i \subset \mathcal{Z}_i$ be closed and convex, and let $M_i:\mathcal{Z}_i\to\mathcal{K}$ be linear and bounded. In addition, let $f\in\Gamma_0(\mathcal{K})$ be Gâteaux differentiable on $\operatorname{dom}\partial f$, let $\varphi:\mathcal{K}\to\mathbb{R}$ be convex and differentiable with a Lipschitzian gradient, let $\mathcal{V}$ be a real Hilbert space, let $g\in\Gamma_0(\mathcal{V})$ be such that g^* is Gâteaux differentiable on $\operatorname{dom}\partial g^*$, let D be a closed convex subset of $\mathcal{V}$ such that

$$0 \in \operatorname{sri}(D - \operatorname{dom} g^*), \tag{103}$$

let $h\in\Gamma_0(\mathcal{V})$ be strongly convex, and let $L:\mathcal{K}\to\mathcal{V}$ be linear and bounded. By Bauschke and Combettes [9, theorem 18.15], h^* is differentiable on $\mathcal{V}$ and ∇h^* is cocoercive. The objective is to solve the Kuhn–Tucker problem

$$\text{find } (\overline{x},\overline{v}^*) \in \mathcal{K}\oplus\mathcal{V} \text{ such that}$$

$$\begin{bmatrix}0\\0\end{bmatrix} \in \underbrace{\begin{bmatrix}\nabla f & 0\\ 0 & \nabla g^*\end{bmatrix}}_{\text{monotone}}\begin{bmatrix}\overline{x}\\ \overline{v}^*\end{bmatrix} + \underbrace{\begin{bmatrix}\nabla\varphi & 0\\ 0 & \nabla h^*\end{bmatrix}}_{\text{cocoercive}}\begin{bmatrix}\overline{x}\\ \overline{v}^*\end{bmatrix} + \underbrace{\begin{bmatrix}0 & L^*\\ -L & 0\end{bmatrix}}_{\text{Lipschitzian}}\begin{bmatrix}\overline{x}\\ \overline{v}^*\end{bmatrix} + \underbrace{\begin{bmatrix}N_C & 0\\ 0 & N_D\end{bmatrix}}_{\text{normal cone}}\begin{bmatrix}\overline{x}\\ \overline{v}^*\end{bmatrix}, \tag{104}$$

where it is assumed that

$$C = \sum_{i \in I} M_i(S_i) \text{ is closed and } 0 \in \operatorname{sri}(C - \operatorname{dom} f). \tag{105}$$

Since $\operatorname{dom} h^* = \mathcal{V}$, we deduce from (103) and Bauschke and Combettes [9, proposition 15.7(i)] that $g \,\square\, h \,\square\, \sigma_D \in \Gamma_0(\mathcal{V})$. It follows from standard convex calculus (Bauschke and Combettes [9]) that a solution $(\overline{x}, \overline{v}^*)$ to (104) provides a solution $\overline{x}$ to

$$\underset{x \in C}{\text{minimize}}\; f(x) + (g \,\square\, h \,\square\, \sigma_D)(Lx) + \varphi(x), \tag{106}$$

as well as a solution $\overline{v}^*$ to the associated Fenchel–Rockafellar dual

$$\underset{v^* \in D}{\text{minimize}}\; \big((f + \varphi)^* \,\square\, \sigma_C\big)(-L^* v^*) + g^*(v^*) + h^*(v^*). \tag{107}$$

To see that (104)–(105) is a special case of Problem 2, set $\mathcal{G} = \mathcal{K} \oplus \mathcal{V}$ and

$$(\forall i \in I) \quad L_i : \mathcal{H}_i = \mathcal{Z}_i \oplus \mathcal{V} \to \mathcal{G} : (z_i, v^*) \mapsto (M_i z_i, v^*/\operatorname{card} I), \quad E_i = S_i \times D, \quad \text{and} \quad F_i = \mathcal{Z}_i \times \mathcal{V}. \tag{108}$$

Note that

$$C \times D = \sum_{i \in I} L_i(E_i \cap F_i). \tag{109}$$

Furthermore, in view of Bauschke and Combettes [9, proposition 17.31(i)], let us define

$$\begin{cases} B^{m} : \mathcal{G} \to 2^{\mathcal{G}} : (x, v^*) \mapsto \partial(f \oplus g^*)(x, v^*) = \begin{cases} (\nabla f(x), \nabla g^*(v^*)), & \text{if } (x, v^*) \in \operatorname{dom} \partial f \times \operatorname{dom} \partial g^*; \\ \varnothing, & \text{otherwise} \end{cases} \\ B^{c} : \mathcal{G} \to \mathcal{G} : (x, v^*) \mapsto (\nabla \varphi(x), \nabla h^*(v^*)) \\ B^{\ell} : \mathcal{G} \to \mathcal{G} : (x, v^*) \mapsto (L^* v^*, -Lx). \end{cases} \tag{110}$$

Then B^{m} is maximally monotone (Bauschke and Combettes [9, theorem 20.25]), B^{c} is cocoercive (Bauschke and Combettes [9, corollary 18.17]), and B^{ℓ} is a skew bounded linear operator and hence monotone and Lipschitzian (Bauschke and Combettes [9, example 20.35]). In turn, combining (108) and (110), we conclude that (104) can be written as

$$\text{find } (\overline{x}, \overline{v}^*) \in \mathcal{K} \oplus \mathcal{V} \text{ such that } (0,0) \in B^{m}(\overline{x}, \overline{v}^*) + B^{c}(\overline{x}, \overline{v}^*) + B^{\ell}(\overline{x}, \overline{v}^*) + N_{C \times D}(\overline{x}, \overline{v}^*), \tag{111}$$

which, in the light of (109), fits the format of (102). Special cases of (106) involving minimization over Minkowski sum of sets are found in areas such as signal and image processing (Aujol and Chambolle [5], Combettes and Wajs [28], Ono et al. [41]), location and network problems (Nam et al. [40]), and robotics and computational mechanics (Wang et al. [54]).

We are going to reformulate Problem 2 as a realization of Problem 1 and solve it via a block-iterative method derived from Algorithm 1. In addition, our approach employs the individual projection operators onto the sets $(E_i)_{i \in I}$ and $(F_i)_{i \in I}$ and the resolvents of the operator B^m. We are not aware of any method which features such flexibility. For instance, consider the special case discussed in Fukushima [31, section 4], where $\mathcal{G} = \mathbb{R}^N$, $B^c = B^\ell = 0$, $T : \mathbb{R}^N \to \mathbb{R}^M$ is a linear operator, and, for every $i \in I$, $\mathcal{H}_i = \mathbb{R}^N$, $L_i = \mathrm{Id}$, $E_i = T^{-1}(\{d_i\})$ for some $d_i \in \mathbb{R}^M$, and $F_i = [0, +\infty[^N$. There, the evaluations of all the projectors $(\operatorname{proj}_{E_i \cap F_i})_{i \in I}$ are required at every iteration. Note that there are no closed-form expressions for $(\operatorname{proj}_{E_i \cap F_i})_{i \in I}$ in general.

Corollary 1. *Consider the setting of Problem 2. Let $\sigma \in]1/(4\beta^c), +\infty[$, $\varepsilon \in]0, \min\{1, 1/(\beta^\ell + \sigma)\}[$, and $K = I \cup \{\overline{k}\}$, where $\overline{k} \notin I$. Suppose that Assumption 2 is in force, together with the following:*

[a] *For every $i \in I$ and every $n \in \mathbb{N}$, $\{\gamma_{i,n}, \mu_{i,n}, \nu_{i,n}\} \subset [\varepsilon, 1/\sigma]$ and $\sigma_{i,n} \in [\varepsilon, 1/\varepsilon]$.*

[b] *For every $n \in \mathbb{N}$, $\lambda_n \in [\varepsilon, 2-\varepsilon]$, $\mu_{\overline{k},n} \in [\varepsilon, 1/(\beta^\ell + \sigma)]$, $\nu_{\overline{k},n} \in [\varepsilon, 1/\sigma]$, and $\sigma_{\overline{k},n} \in [\varepsilon, 1/\varepsilon]$.*

[c] *For every $i \in I$, $\{x_{i,0}, y_{i,0}, z_{i,0}, v^*_{i,0}\} \subset \mathcal{H}_i$; $\{y_{\overline{k},0}, z_{\overline{k},0}, v^*_{\overline{k},0}\} \subset \mathcal{G}$.*

Iterate

$$
\begin{array}{l}
\text{for } n = 0, 1, \ldots \\
\quad \text{for every } i \in I_n \\
\quad\quad l^*_{i,n} = v^*_{i,n} + L^*_i v^*_{\overline{k},n}; \\
\quad\quad a_{i,n} = \operatorname{proj}_{E_i}(x_{i,n} - \gamma_{i,n} l^*_{i,n}); \\
\quad\quad a^*_{i,n} = \gamma^{-1}_{i,n}(x_{i,n} - a_{i,n}) - l^*_{i,n}; \\
\quad\quad \xi_{i,n} = \|a_{i,n} - x_{i,n}\|^2; \\
\quad \text{for every } i \in I \backslash I_n \\
\quad\quad a_{i,n} = a_{i,n-1};\ a^*_{i,n} = a^*_{i,n-1};\ \ \xi_{i,n} = \xi_{i,n-1}; \\
\quad \text{for every } k \in K_n \\
\quad\quad \text{if } k \in I \\
\quad\quad\quad b_{k,n} = \operatorname{proj}_{F_k}(y_{k,n} + \mu_{k,n} v^*_{k,n}); \\
\quad\quad\quad e^*_{k,n} = \sigma_{k,n}(x_{k,n} - y_{k,n} - z_{k,n}) + v^*_{k,n}; \\
\quad\quad\quad q^*_{k,n} = \mu^{-1}_{k,n}(y_{k,n} - b_{k,n}) + v^*_{k,n} - e^*_{k,n}; \\
\quad\quad\quad e_{k,n} = b_{k,n} - a_{k,n}; \\
\quad\quad \text{if } k = \overline{k} \\
\quad\quad\quad u^*_{k,n} = v^*_{k,n} - B^{\ell} y_{k,n}; \\
\quad\quad\quad b_{k,n} = J_{\mu_{k,n} B^{\mathfrak{m}}}(y_{k,n} + \mu_{k,n}(u^*_{k,n} - B^{\mathfrak{c}} y_{k,n})); \\
\quad\quad\quad e^*_{k,n} = \sigma_{k,n}\left(\sum_{i\in I} L_i x_{i,n} - y_{k,n} - z_{k,n}\right) + v^*_{k,n}; \\
\quad\quad\quad q^*_{k,n} = \mu^{-1}_{k,n}(y_{k,n} - b_{k,n}) + u^*_{k,n} + B^{\ell} b_{k,n} - e^*_{k,n}; \\
\quad\quad\quad e_{k,n} = b_{k,n} - \sum_{i\in I} L_i a_{i,n}; \\
\quad\quad t^*_{k,n} = \nu^{-1}_{k,n} z_{k,n} + v^*_{k,n} - e^*_{k,n}; \\
\quad\quad \eta_{k,n} = \|b_{k,n} - y_{k,n}\|^2 + \|z_{k,n}\|^2; \\
\quad \text{for every } k \in K \backslash K_n \\
\quad\quad b_{k,n} = b_{k,n-1};\ e^*_{k,n} = e^*_{k,n-1};\ \ q^*_{k,n} = q^*_{k,n-1};\ \ t^*_{k,n} = t^*_{k,n-1};\ \ \eta_{k,n} = \eta_{k,n-1}; \\
\quad\quad \text{if } k \in I \\
\quad\quad\quad e_{k,n} = b_{k,n} - a_{k,n}; \\
\quad\quad \text{if } k = \overline{k} \\
\quad\quad\quad e_{k,n} = b_{k,n} - \sum_{i\in I} L_i a_{i,n}; \\
\quad \text{for every } i \in I \\
\quad\quad p^*_{i,n} = a^*_{i,n} + e^*_{i,n} + L^*_i e^*_{\overline{k},n}; \\
\quad \Delta_n = -(4\beta^{\mathfrak{c}})^{-1}\left(\sum_{i\in I} \xi_{i,n} + \sum_{k\in K} \eta_{k,n}\right) + \sum_{i\in I} \langle x_{i,n} - a_{i,n} \mid p^*_{i,n}\rangle \\
\quad\quad\quad + \sum_{k\in K}\big(\langle y_{k,n} - b_{k,n} \mid q^*_{k,n}\rangle + \langle z_{k,n} \mid t^*_{k,n}\rangle + \langle e_{k,n} \mid v^*_{k,n} - e^*_{k,n}\rangle\big); \\
\quad \text{if } \Delta_n > 0 \\
\quad\quad \theta_n = \lambda_n \Delta_n \Big/ \left(\sum_{i\in I} \|p^*_{i,n}\|^2 + \sum_{k\in K}\left(\|q^*_{k,n}\|^2 + \|t^*_{k,n}\|^2 + \|e_{k,n}\|^2\right)\right); \\
\quad\quad \text{for every } i \in I \\
\quad\quad\quad x_{i,n+1} = x_{i,n} - \theta_n p^*_{i,n}; \\
\quad\quad \text{for every } k \in K \\
\quad\quad\quad y_{k,n+1} = y_{k,n} - \theta_n q^*_{k,n};\ z_{k,n+1} = z_{k,n} - \theta_n t^*_{k,n};\ \ v^*_{k,n+1} = v^*_{k,n} - \theta_n e_{k,n}; \\
\quad \text{else} \\
\quad\quad \text{for every } i \in I \\
\quad\quad\quad x_{i,n+1} = x_{i,n}; \\
\quad\quad \text{for every } k \in K \\
\quad\quad\quad y_{k,n+1} = y_{k,n};\ z_{k,n+1} = z_{k,n};\ \ v^*_{k,n+1} = v^*_{k,n}.
\end{array}
\tag{112}
$$

Furthermore, suppose that (102) *has a solution and that*

$$(\forall i \in I) \quad N_{E_i \cap F_i} = N_{E_i} + N_{F_i}. \tag{113}$$

Then there exists $(\overline{x}_i)_{i\in I} \in \bigoplus_{i\in I} \mathcal{H}_i$ *such that* $\sum_{i\in I} L_i \overline{x}_i$ *solves* (102) *and, for every* $i \in I$, $x_{i,n} \rightharpoonup \overline{x}_i$ *and* $a_{i,n} \rightharpoonup \overline{x}_i$.

Proof. Set $\boldsymbol{\mathcal{H}} = \bigoplus_{i\in I} \mathcal{H}_i$. Let us consider the problem

$$\text{find } \overline{\boldsymbol{x}} \in \boldsymbol{\mathcal{H}} \text{ such that } (\forall i \in I)\ 0 \in N_{E_i}\overline{x}_i + N_{F_i}\overline{x}_i + L_i^*(B^m + B^c + B^\ell)\Bigg(\sum_{j\in I} L_j \overline{x}_j\Bigg) \tag{114}$$

together with the associated dual problem

$$\text{find } (\overline{\boldsymbol{x}}^*, \overline{v}^*) \in \boldsymbol{\mathcal{H}} \oplus \mathcal{G} \text{ such that } (\exists \boldsymbol{x} \in \boldsymbol{\mathcal{H}}) \begin{cases} (\forall i \in I) \ -\overline{x}_i^* - L_i^*\overline{v}^* \in N_{E_i}x_i \text{ and } \overline{x}_i^* \in N_{F_i}x_i \\ \overline{v}^* = (B^m + B^c + B^\ell)\Bigg(\sum_{j\in I} L_j x_j\Bigg). \end{cases} \tag{115}$$

Denote by $\mathscr{P}$ and $\mathscr{D}$ the sets of solutions to (114) and (115), respectively. We observe that the primal-dual problem (114)–(115) is a special case of Problem 1 with

$$(\forall i \in I) \quad A_i = N_{E_i}, \quad C_i = Q_i = 0, \quad R_i = 0, \quad \text{and} \quad s_i^* = 0, \tag{116}$$

and

$$(\forall k \in K) \quad \begin{cases} \mathcal{G}_k = \mathcal{H}_k,\ B_k^m = N_{F_k},\ B_k^c = B_k^\ell = 0 \text{ if } k \in I; \\ \mathcal{G}_{\overline{k}} = \mathcal{G},\ B_{\overline{k}}^m = B^m,\ B_{\overline{k}}^c = B^c,\ B_{\overline{k}}^\ell = B^\ell \\ D_k^m = \{0\}^{-1},\ D_k^c = D_k^\ell = 0,\ r_k = 0 \\ (\forall j \in I)\ L_{kj} = \begin{cases} \mathrm{Id}, & \text{if } k = j; \\ 0, & \text{if } k \in I \text{ and } k \neq j; \\ L_j, & \text{if } k = \overline{k}. \end{cases} \end{cases} \tag{117}$$

Furthermore, we have

$$\begin{cases} (\forall i \in I)(\forall n \in \mathbb{N})\ J_{\gamma_{i,n}A_i} = \mathrm{proj}_{E_i} \\ (\forall k \in K)(\forall n \in \mathbb{N})\ J_{\nu_{k,n}D_k^m} = 0 \text{ and } J_{\mu_{k,n}B_k^m} = \begin{cases} \mathrm{proj}_{F_k}, & \text{if } k \in I; \\ J_{\mu_{k,n}B^m}, & \text{if } k = \overline{k}. \end{cases} \end{cases} \tag{118}$$

Therefore, (112) is a realization of Algorithm 1 in the context of (114)–(115). Now define $\boldsymbol{D} = \times_{i\in I}(E_i \cap F_i)$ and $\boldsymbol{L}: \boldsymbol{\mathcal{H}} \to \mathcal{G} : \boldsymbol{x} \mapsto \sum_{i\in I} L_i x_i$. Then $\boldsymbol{L}^*: \mathcal{G} \to \boldsymbol{\mathcal{H}} : y^* \mapsto (L_i^* y^*)_{i\in I}$. Hence, by (102), Bauschke and Combettes [9, proposition 16.9], and (113),

$$\begin{aligned}
(\forall \overline{y} \in \mathcal{G}) \quad \overline{y} \text{ solves (102)} &\Leftrightarrow (\exists\, \overline{\boldsymbol{x}} \in \boldsymbol{D}) \begin{cases} \overline{y} = \boldsymbol{L}\overline{\boldsymbol{x}} \\ (\forall \boldsymbol{x} \in \boldsymbol{D})\ \langle \boldsymbol{L}\overline{\boldsymbol{x}} - \boldsymbol{L}\boldsymbol{x} \mid (B^m + B^c + B^\ell)(\boldsymbol{L}\overline{\boldsymbol{x}})\rangle \leqslant 0 \end{cases} \\
&\Leftrightarrow (\exists\, \overline{\boldsymbol{x}} \in \boldsymbol{D}) \begin{cases} \overline{y} = \boldsymbol{L}\overline{\boldsymbol{x}} \\ (\forall \boldsymbol{x} \in \boldsymbol{D})\ \langle \overline{\boldsymbol{x}} - \boldsymbol{x} \mid \boldsymbol{L}^*((B^m + B^c + B^\ell)(\boldsymbol{L}\overline{\boldsymbol{x}}))\rangle \leqslant 0 \end{cases} \\
&\Leftrightarrow (\exists\, \overline{\boldsymbol{x}} \in \boldsymbol{\mathcal{H}}) \begin{cases} \overline{y} = \boldsymbol{L}\overline{\boldsymbol{x}} \\ \boldsymbol{0} \in N_{\boldsymbol{D}}\overline{\boldsymbol{x}} + \boldsymbol{L}^*((B^m + B^c + B^\ell)(\boldsymbol{L}\overline{\boldsymbol{x}})) \end{cases} \\
&\Leftrightarrow (\exists\, \overline{\boldsymbol{x}} \in \boldsymbol{\mathcal{H}}) \begin{cases} \overline{y} = \boldsymbol{L}\overline{\boldsymbol{x}} \\ (\forall i \in I)\ 0 \in N_{E_i \cap F_i}\overline{x}_i + L_i^*(B^m + B^c + B^\ell)\Bigg(\sum_{j\in I} L_j \overline{x}_j\Bigg) \end{cases} \\
&\Leftrightarrow (\exists\, \overline{\boldsymbol{x}} \in \boldsymbol{\mathcal{H}}) \begin{cases} \overline{y} = \boldsymbol{L}\overline{\boldsymbol{x}} \\ (\forall i \in I)\ 0 \in N_{E_i}\overline{x}_i + N_{F_i}\overline{x}_i + L_i^*(B^m + B^c + B^\ell)\Bigg(\sum_{j\in I} L_j \overline{x}_j\Bigg) \end{cases} \\
&\Leftrightarrow (\exists\, \overline{\boldsymbol{x}} \in \mathscr{P})\ \overline{y} = \boldsymbol{L}\overline{\boldsymbol{x}}.
\end{aligned} \tag{119}$$

In turn, $\mathscr{P} \neq \varnothing$ since (102) has a solution. Therefore, in view of (117), Proposition 1(v)[d] yields $\mathscr{D} \neq \varnothing$. As a result, Theorem 1(iv) asserts that there exists $(\overline{x}_i)_{i\in I} \in \mathscr{P}$ such that, for every $i \in I$, $x_{i,n} \rightharpoonup \overline{x}_i$ and $a_{i,n} \rightharpoonup \overline{x}_i$. Finally, using (119), we conclude that $\sum_{i\in I} L_i\overline{x}_i$ solves (102). □

Remark 2. Theorem 1 allows us to tackle other types of variational inequalities. For instance, let $(\mathcal{H}_i)_{i\in I}$ be a finite family of real Hilbert spaces and set $\boldsymbol{\mathcal{H}} = \bigoplus_{i\in I}\mathcal{H}_i$. For every $i \in I$, let $\varphi_i \in \Gamma_0(\mathcal{H}_i)$ and let $R_i : \boldsymbol{\mathcal{H}} \to \mathcal{H}_i$ be such that Problem 1[e] holds. The objective is to

$$\text{find } \ \overline{\boldsymbol{x}} \in \boldsymbol{\mathcal{H}} \ \text{ such that } \ (\forall i \in I) \ \ 0 \in \partial\varphi_i(\overline{x}_i) + R_i\overline{\boldsymbol{x}}. \tag{120}$$

This simple instantiation of Problem 1 shows up in neural networks (Combettes and Pesquet [27]) and in game theory (Attouch et al. [2], Briceño-Arias and Combettes [15]). Thanks to Theorem 1, it can be solved using an asynchronous block-iterative strategy, which is not possible with current splitting techniques such as those of Combettes and Eckstein [25] and Johnstone and Eckstein [34].

4.2. Application to Multivariate Minimization

We consider a composite multivariate minimization problem involving various types of convex functions and combinations between them.

Problem 3. Let $(\mathcal{H}_i)_{i\in I}$ and $(\mathcal{G}_k)_{k\in K}$ be finite families of real Hilbert spaces, and set $\boldsymbol{\mathcal{H}} = \bigoplus_{i\in I}\mathcal{H}_i$ and $\boldsymbol{\mathcal{G}} = \bigoplus_{k\in K}\mathcal{G}_k$. For every $i \in I$ and every $k \in K$, let $f_i \in \Gamma_0(\mathcal{H}_i)$, let $\alpha_i \in \left]0,+\infty\right[$, let $\varphi_i : \mathcal{H}_i \to \mathbb{R}$ be convex and differentiable with a $(1/\alpha_i)$-Lipschitzian gradient, let $g_k \in \Gamma_0(\mathcal{G}_k)$, let $h_k \in \Gamma_0(\mathcal{G}_k)$, let $\beta_k \in \left]0,+\infty\right[$, let $\psi_k : \mathcal{G}_k \to \mathbb{R}$ be convex and differentiable with a $(1/\beta_k)$-Lipschitzian gradient, and suppose that $L_{ki} : \mathcal{H}_i \to \mathcal{G}_k$ is linear and bounded. In addition, let $\chi \in \left[0,+\infty\right[$ and let $\Theta : \boldsymbol{\mathcal{H}} \to \mathbb{R}$ be convex and differentiable with a χ-Lipschitzian gradient. The objective is to

$$\underset{\boldsymbol{x}\in\boldsymbol{\mathcal{H}}}{\text{minimize}}\ \Theta(\boldsymbol{x}) + \sum_{i\in I}\big(f_i(x_i) + \varphi_i(x_i)\big) + \sum_{k\in K}\big((g_k+\psi_k)\,\square\, h_k\big)\Bigg(\sum_{j\in I} L_{kj}x_j\Bigg). \tag{121}$$

Special cases of Problem 3 are found in various contexts (Briceño-Arias and Combettes [13], Briceño-Arias et al. [16], Combettes [23], Combettes and Eckstein [25], Hintermüller and Stadler [33], Johnstone and Eckstein [34]). Formulation (121) brings together these disparate problems, and the following algorithm makes it possible to solve them in an asynchronous block-iterative fashion in full generality.

Algorithm 3. Consider the setting of Problem 3 and suppose that Assumptions 2 and 3 are in force. Set $\alpha = \min\{\alpha_i,\beta_k\}_{i\in I,k\in K}$, let $\sigma \in \left]1/(4\alpha),+\infty\right[$, and let $\varepsilon \in \left]0,\min\{1,1/(\chi+\sigma)\}\right[$. For every $i \in I$, every $k \in K$, and every $n \in \mathbb{N}$, let $\gamma_{i,n} \in [\varepsilon, 1/(\chi+\sigma)]$, let $\{\mu_{k,n},\nu_{k,n}\} \subset [\varepsilon,1/\sigma]$, let $\sigma_{k,n} \in [\varepsilon,1/\varepsilon]$, and let $\lambda_n \in [\varepsilon, 2-\varepsilon]$. In addition, let $\boldsymbol{x}_0 \in \boldsymbol{\mathcal{H}}$ and $\{\boldsymbol{y}_0, \boldsymbol{z}_0, \boldsymbol{v}_0^*\} \subset \boldsymbol{\mathcal{G}}$. Iterate

$$
\begin{array}{l}
\text{for } n = 0,1,\ldots\\
\left\lfloor
\begin{array}{l}
\text{for every } i \in I_n\\
\left\lfloor
\begin{array}{l}
l_{i,n}^* = \nabla_i\Theta(\boldsymbol{x}_{\pi_i(n)}) + \sum_{k\in K} L_{ki}^* v_{k,\pi_i(n)}^*;\\
a_{i,n} = \text{prox}_{\gamma_{i,\pi_i(n)} f_i}\Big(x_{i,\pi_i(n)} - \gamma_{i,\pi_i(n)}\big(l_{i,n}^* + \nabla\varphi_i(x_{i,\pi_i(n)})\big)\Big);\\
a_{i,n}^* = \gamma_{i,\pi_i(n)}^{-1}(x_{i,\pi_i(n)} - a_{i,n}) - l_{i,n}^*;\\
\xi_{i,n} = \|a_{i,n} - x_{i,\pi_i(n)}\|^2;
\end{array}
\right.\\
\text{for every } i \in I\smallsetminus I_n\\
\left\lfloor
a_{i,n} = a_{i,n-1};\ a_{i,n}^* = a_{i,n-1}^*;\ \xi_{i,n} = \xi_{i,n-1};
\right.\\
\text{for every } k \in K_n\\
\left\lfloor
\begin{array}{l}
b_{k,n} = \text{prox}_{\mu_{k,\omega_k(n)} g_k}\Big(y_{k,\omega_k(n)} + \mu_{k,\omega_k(n)}\big(v_{k,\omega_k(n)}^* - \nabla\psi_k(y_{k,\omega_k(n)})\big)\Big);\\
d_{k,n} = \text{prox}_{\nu_{k,\omega_k(n)} h_k}\big(z_{k,\omega_k(n)} + \nu_{k,\omega_k(n)} v_{k,\omega_k(n)}^*\big);\\
e_{k,n}^* = \sigma_{k,\omega_k(n)}\Bigg(\sum_{i\in I} L_{ki}x_{i,\omega_k(n)} - y_{k,\omega_k(n)} - z_{k,\omega_k(n)}\Bigg) + v_{k,\omega_k(n)}^*;\\
q_{k,n}^* = \mu_{k,\omega_k(n)}^{-1}(y_{k,\omega_k(n)} - b_{k,n}) + v_{k,\omega_k(n)}^* - e_{k,n}^*;\\
t_{k,n}^* = \nu_{k,\omega_k(n)}^{-1}(z_{k,\omega_k(n)} - d_{k,n}) + v_{k,\omega_k(n)}^* - e_{k,n}^*;\\
\eta_{k,n} = \|b_{k,n} - y_{k,\omega_k(n)}\|^2 + \|d_{k,n} - z_{k,\omega_k(n)}\|^2;\\
e_{k,n} = b_{k,n} + d_{k,n} - \sum_{i\in I} L_{ki}a_{i,n};
\end{array}
\right.
\end{array}
\right.
\end{array}
\tag{122}
$$

$$
\begin{array}{l}
\text{for every } k \in K\setminus K_n \\
\left\lfloor \begin{array}{l} b_{k,n} = b_{k,n-1};\ d_{k,n} = d_{k,n-1};\ e^*_{k,n} = e^*_{k,n-1};\ q^*_{k,n} = q^*_{k,n-1};\ t^*_{k,n} = t^*_{k,n-1}; \\ \eta_{k,n} = \eta_{k,n-1};\ e_{k,n} = b_{k,n} + d_{k,n} - \sum_{i\in I} L_{ki} a_{i,n}; \end{array}\right. \\
\text{for every } i \in I \\
\left\lfloor p^*_{i,n} = a^*_{i,n} + \nabla_i \Theta(\boldsymbol{a}_n) + \sum_{k\in K} L^*_{ki} e^*_{k,n};\right. \\
\Delta_n = -(4\alpha)^{-1}\left(\sum_{i\in I}\xi_{i,n} + \sum_{k\in K}\eta_{k,n}\right) + \sum_{i\in I}\langle x_{i,n} - a_{i,n} \mid p^*_{i,n}\rangle \\
\qquad + \sum_{k\in K}\Big(\langle y_{k,n} - b_{k,n} \mid q^*_{k,n}\rangle + \langle z_{k,n} - d_{k,n} \mid t^*_{k,n}\rangle + \langle e_{k,n} \mid v^*_{k,n} - e^*_{k,n}\rangle\Big); \\
\text{if } \Delta_n > 0 \\
\left\lfloor \begin{array}{l} \theta_n = \lambda_n\Delta_n \Big/ \left(\sum_{i\in I}\|p^*_{i,n}\|^2 + \sum_{k\in K}\Big(\|q^*_{k,n}\|^2 + \|t^*_{k,n}\|^2 + \|e_{k,n}\|^2\Big)\right); \\
\text{for every } i \in I \\
\lfloor x_{i,n+1} = x_{i,n} - \theta_n p^*_{i,n}; \\
\text{for every } k \in K \\
\lfloor y_{k,n+1} = y_{k,n} - \theta_n q^*_{k,n};\ z_{k,n+1} = z_{k,n} - \theta_n t^*_{k,n};\ v^*_{k,n+1} = v^*_{k,n} - \theta_n e_{k,n}; \end{array}\right. \\
\text{else} \\
\left\lfloor \begin{array}{l} \text{for every } i \in I \\
\lfloor x_{i,n+1} = x_{i,n}; \\
\text{for every } k \in K \\
\lfloor y_{k,n+1} = y_{k,n};\ z_{k,n+1} = z_{k,n};\ v^*_{k,n+1} = v^*_{k,n}. \end{array}\right.
\end{array}
$$

Corollary 2. *Consider the setting of Algorithm 3. Suppose that*

$$(\forall k \in K) \quad \text{epi}\,(g_k + \psi_k) + \text{epi}\, h_k \text{ is closed} \tag{123}$$

and that Problem 3 admits a Kuhn–Tucker point, that is, there exist $\widetilde{\boldsymbol{x}} \in \boldsymbol{\mathcal{H}}$ and $\widetilde{\boldsymbol{v}}^* \in \boldsymbol{\mathcal{G}}$ such that

$$(\forall i \in I)(\forall k \in K) \quad \begin{cases} -\sum_{j\in K} L^*_{ji}\widetilde{v}^*_j \in \partial f_i(\widetilde{x}_i) + \nabla\varphi_i(\widetilde{x}_i) + \nabla_i\Theta(\widetilde{\boldsymbol{x}}) \\ \sum_{j\in I} L_{kj}\widetilde{x}_j \in \partial(g^*_k \,\square\, \psi^*_k)(\widetilde{v}^*_k) + \partial h^*_k(\widetilde{v}^*_k). \end{cases} \tag{124}$$

Then there exists a solution $\overline{\boldsymbol{x}}$ to (121) such that, for every $i \in I$, $x_{i,n} \rightharpoonup \overline{x}_i$ and $a_{i,n} \rightharpoonup \overline{x}_i$.

Proof. Set

$$\begin{cases} (\forall i \in I) \;\; A_i = \partial f_i, \quad C_i = \nabla\varphi_i, \quad \text{and} \quad R_i = \nabla_i\,\Theta \\ (\forall k \in K) \;\; B^{m}_k = \partial g_k, \quad B^{c}_k = \nabla\psi_k, \quad \text{and} \quad D^{m}_k = \partial h_k. \end{cases} \tag{125}$$

First, Bauschke and Combettes [9, theorem 20.25] asserts that the operators $(A_i)_{i\in I}$, $(B^{m}_k)_{k\in K}$, and $(D^{m}_k)_{k\in K}$ are maximally monotone. Second, it follows from Bauschke and Combettes [9, corollary 18.17] that, for every $i \in I$, C_i is α_i-cocoercive and, for every $k \in K$, B^{c}_k is β_k-cocoercive. Third, in view of (125) and Bauschke and Combettes [9, proposition 17.7], $\boldsymbol{R} = \nabla\Theta$ is monotone and χ-Lipschitzian. Now consider the problem

$$\text{find } \overline{\boldsymbol{x}} \in \boldsymbol{\mathcal{H}} \text{ such that } (\forall i \in I)\;\; 0 \in A_i\overline{x}_i + C_i\overline{x}_i + R_i\overline{\boldsymbol{x}} + \sum_{k\in K} L^*_{ki}\left(\left((B^{m}_k + B^{c}_k) \,\square\, D^{m}_k\right)\left(\sum_{j\in I} L_{kj}\overline{x}_j\right)\right) \tag{126}$$

together with its dual

$$\text{find } \overline{\boldsymbol{v}}^* \in \boldsymbol{\mathcal{G}} \text{ such that } (\exists \boldsymbol{x} \in \boldsymbol{\mathcal{H}})(\forall i \in I)(\forall k \in K) \quad \begin{cases} -\sum_{j\in K} L^*_{ji}\overline{v}^*_j \in A_i x_i + C_i x_i + R_i\boldsymbol{x} \\ \overline{v}^*_k \in \left((B^{m}_k + B^{c}_k) \,\square\, D^{m}_k\right)\left(\sum_{j\in I} L_{kj}x_j\right). \end{cases} \tag{127}$$

Denote by $\mathscr{P}$ and $\mathscr{D}$ the sets of solutions to (126) and (127), respectively. We observe that, by (125) and Bauschke and Combettes [9, example 23.3], Algorithm 3 is an application of Algorithm 1 to the primal-dual problem (126)–(127). Furthermore, it results from (124) and Proposition 1(iv) that $\mathscr{D} \neq \varnothing$. According to Theorem 1(iv), there exist $\overline{\boldsymbol{x}} \in \mathscr{P}$ and $\overline{\boldsymbol{v}}^* \in \mathscr{D}$ such that, for every $i \in I$ and every $k \in K$,

$$x_{i,n} \rightharpoonup \overline{x}_i, \quad a_{i,n} \rightharpoonup \overline{x}_i, \quad \text{and} \quad \begin{cases} -\sum_{j\in K} L_{ji}^* \overline{v}_j^* \in A_i\overline{x}_i + C_i\overline{x}_i + R_i\overline{\boldsymbol{x}} \\ \overline{v}_k^* \in ((B_k^{m} + B_k^{c}) \,\square\, D_k^{m})\left(\sum_{j\in I} L_{kj}\overline{x}_j\right). \end{cases} \tag{128}$$

It remains to show that $\overline{\boldsymbol{x}}$ solves (121). Define

$$\begin{cases} \boldsymbol{f} = \bigoplus_{i\in I} f_i, \ \boldsymbol{\varphi} = \bigoplus_{i\in I} \varphi_i, \ \boldsymbol{g} = \bigoplus_{k\in K} g_k, \ \boldsymbol{h} = \bigoplus_{k\in K} h_k, \ \text{and} \ \boldsymbol{\psi} = \bigoplus_{k\in K} \psi_k \\ \boldsymbol{L} : \boldsymbol{\mathcal{H}} \to \boldsymbol{\mathcal{G}} : \boldsymbol{x} \mapsto \left(\sum_{i\in I} L_{ki}x_i\right)_{k\in K}. \end{cases} \tag{129}$$

We deduce from Bauschke and Combettes [9, theorem 15.3] that $(\forall k \in K)\ (g_k + \psi_k)^* = g_k^* \boxdot \psi_k^*$. In turn, (124) implies that

$$(\forall k \in K) \quad \varnothing \neq \operatorname{dom}(g_k^* \boxdot \psi_k^*) \cap \operatorname{dom} h_k^* = \operatorname{dom}(g_k + \psi_k)^* \cap \operatorname{dom} h_k^*. \tag{130}$$

On the other hand, since the sets $(\operatorname{epi}(g_k + \psi_k) + \operatorname{epi} h_k)_{k\in K}$ are convex, it follows from (123) and Bauschke and Combettes [9, theorem 3.34] that they are weakly closed. Therefore, Burachik and Jeyakumar [20, theorem 1] and the Fenchel–Moreau theorem (Bauschke and Combettes [9, theorem 13.37]) imply that

$$(\forall k \in K) \quad ((g_k + \psi_k)^* + h_k^*)^* = (g_k + \psi_k)^{**} \boxdot h_k^{**} = (g_k + \psi_k) \boxdot h_k. \tag{131}$$

Hence, we derive from (125), Bauschke and Combettes [9, corollaries 16.48(iii) and 16.30], (131), and Bauschke and Combettes [9, proposition 16.42] that

$$\begin{aligned} (\forall k \in K) \quad (B_k^{m} + B_k^{c}) \,\square\, D_k^{m} &= (\partial g_k + \nabla \psi_k) \,\square\, (\partial h_k) \\ &= ((\partial(g_k + \psi_k))^{-1} + (\partial h_k)^{-1})^{-1} \\ &= (\partial(g_k + \psi_k)^* + \partial h_k^*)^{-1} \\ &= (\partial((g_k + \psi_k)^* + h_k^*))^{-1} \\ &= \partial((g_k + \psi_k)^* + h_k^*)^* \\ &= \partial((g_k + \psi_k) \boxdot h_k). \end{aligned} \tag{132}$$

Since it results from (129) and (131) that

$$(\boldsymbol{g} + \boldsymbol{\psi}) \,\square\, \boldsymbol{h} = (\boldsymbol{g} + \boldsymbol{\psi}) \boxdot \boldsymbol{h} = \bigoplus_{k\in K} ((g_k + \psi_k) \boxdot h_k), \tag{133}$$

we deduce from Bauschke and Combettes [9, proposition 16.9] and (132) that

$$\partial((\boldsymbol{g} + \boldsymbol{\psi}) \boxdot \boldsymbol{h}) = \bigtimes_{k\in K} \partial((g_k + \psi_k) \boxdot h_k) = \bigtimes_{k\in K} ((B_k^{m} + B_k^{c}) \,\square\, D_k^{m}). \tag{134}$$

It thus follows from (128) and (129) that $\overline{\boldsymbol{v}}^* \in \partial((\boldsymbol{g} + \boldsymbol{\psi}) \boxdot \boldsymbol{h})(\boldsymbol{L}\overline{\boldsymbol{x}})$. On the other hand, since $\boldsymbol{L}^* : \boldsymbol{\mathcal{G}} \to \boldsymbol{\mathcal{H}} : \boldsymbol{v}^* \mapsto (\sum_{k\in K} L_{ki}^* v_k^*)_{i\in I}$, we infer from (128), (125), (129), and Bauschke and Combettes [9, proposition 16.9] that $-\boldsymbol{L}^*\overline{\boldsymbol{v}}^* \in (C_i\overline{x}_i)_{i\in I} + \boldsymbol{R}\overline{\boldsymbol{x}} + \bigtimes_{i\in I} A_i\overline{x}_i = \nabla\boldsymbol{\varphi}(\overline{\boldsymbol{x}}) + \nabla\Theta(\overline{\boldsymbol{x}}) + \partial\boldsymbol{f}(\overline{\boldsymbol{x}})$. Hence, we invoke Bauschke and Combettes [9, proposition 16.6(ii)] to obtain

$$\begin{aligned} \boldsymbol{0} &\in \partial\boldsymbol{f}(\overline{\boldsymbol{x}}) + \nabla\boldsymbol{\varphi}(\overline{\boldsymbol{x}}) + \nabla\Theta(\overline{\boldsymbol{x}}) + \boldsymbol{L}^*\overline{\boldsymbol{v}}^* \\ &\subset \partial\boldsymbol{f}(\overline{\boldsymbol{x}}) + \nabla\boldsymbol{\varphi}(\overline{\boldsymbol{x}}) + \nabla\Theta(\overline{\boldsymbol{x}}) + \boldsymbol{L}^*(\partial((\boldsymbol{g} + \boldsymbol{\psi}) \boxdot \boldsymbol{h})(\boldsymbol{L}\overline{\boldsymbol{x}})) \\ &\subset \partial(\boldsymbol{f} + \boldsymbol{\varphi} + \Theta + ((\boldsymbol{g} + \boldsymbol{\psi}) \boxdot \boldsymbol{h}) \circ \boldsymbol{L})(\overline{\boldsymbol{x}}). \end{aligned} \tag{135}$$

However, thanks to (129) and (133), (121) is equivalent to

$$\underset{\boldsymbol{x}\in\boldsymbol{\mathcal{H}}}{\text{minimize}}\ \boldsymbol{f}(\boldsymbol{x}) + \boldsymbol{\varphi}(\boldsymbol{x}) + \Theta(\boldsymbol{x}) + ((\boldsymbol{g} + \boldsymbol{\psi}) \boxdot \boldsymbol{h})(\boldsymbol{L}\boldsymbol{x}). \tag{136}$$

Consequently, in view of Fermat's rule (Bauschke and Combettes [9, theorem 16.3]), (135) implies that $\overline{\boldsymbol{x}}$ solves (121). □

Remark 3. In Briceño-Arias et al. [16], multicomponent image recovery problems were approached by applying the forward-backward and the Douglas–Rachford algorithms in a product space. Using Corollary 2, we can now solve these problems with asynchronous block-iterative algorithms and more sophisticated formulations. For instance, the standard total variation loss used in Briceño-Arias et al. [16] can be replaced by the pth order Huber total variation penalty of Hintermüller and Stadler [33], which turns out to involve an infimal convolution.

To conclude, we provide some scenarios in which condition (123) is satisfied.

Proposition 5. *Consider the setting of Problem 3. Suppose that there exist $\widetilde{\boldsymbol{x}} \in \boldsymbol{\mathcal{H}}$ and $\widetilde{\boldsymbol{v}}^* \in \boldsymbol{\mathcal{G}}$ such that*

$$(\forall i \in I)(\forall k \in K) \quad \begin{cases} -\sum_{j\in K} L_{ji}^*\widetilde{v}_j^* \in \partial f_i(\widetilde{x}_i) + \nabla\varphi_i(\widetilde{x}_i) + \nabla_i\Theta(\widetilde{\boldsymbol{x}}) \\ \sum_{j\in I} L_{kj}\widetilde{x}_j \in \partial(g_k^* \,\square\, \psi_k^*)(\widetilde{v}_k^*) + \partial h_k^*(\widetilde{v}_k^*) \end{cases} \tag{137}$$

and that, for every $k \in K$, one of the following is satisfied:

[a] $0 \in \operatorname{sri}(\operatorname{dom} g_k^* + \operatorname{dom} \psi_k^* - \operatorname{dom} h_k^*)$.

[b] *$\mathcal{G}_k$ is finite-dimensional, h_k is polyhedral, and* $\operatorname{dom} h_k^* \cap \operatorname{ri}\operatorname{dom}(g_k + \psi_k)^* \neq \varnothing$.

[c] *$\mathcal{G}_k$ is finite-dimensional, g_k and h_k are polyhedral, and* $\psi_k = 0$.

Then, for every $k \in K$, $\operatorname{epi}(g_k + \psi_k) + \operatorname{epi} h_k$ *is closed.*

Proof. Let $k \in K$. Since $\operatorname{dom}\psi_k = \mathcal{G}_k$, Bauschke and Combettes [9, theorem 15.3] yields

$$(g_k + \psi_k)^* = g_k^* \,\square\, \psi_k^*. \tag{138}$$

Therefore, (137) implies that

$$\varnothing \neq \operatorname{dom}(g_k^* \,\square\, \psi_k^*) \cap \operatorname{dom} h_k^* = \operatorname{dom}(g_k + \psi_k)^* \cap \operatorname{dom} h_k^*. \tag{139}$$

In view of (139), Burachik and Jeyakumar [20, theorem 1], and Bauschke and Combettes [9, theorem 3.34], it suffices to show that $((g_k + \psi_k)^* + h_k^*)^* = (g_k + \psi_k)^{**} \,\square\, h_k^{**}$.

[a]: We deduce from Bauschke and Combettes [9, proposition 12.6(ii)] and (138) that $0 \in \operatorname{sri}(\operatorname{dom}(g_k^* \,\square\, \psi_k^*) - \operatorname{dom} h_k^*) = \operatorname{sri}(\operatorname{dom}(g_k + \psi_k)^* - \operatorname{dom} h_k^*)$. In turn, Bauschke and Combettes [9, theorem 15.3] gives $((g_k + \psi_k)^* + h_k^*)^* = (g_k + \psi_k)^{**} \,\square\, h_k^{**}$.

[b]: Since Rockafellar [48, theorem 19.2] asserts that h_k^* is polyhedral, we infer from Rockafellar [48, theorem 20.1] that $((g_k + \psi_k)^* + h_k^*)^* = (g_k + \psi_k)^{**} \,\square\, h_k^{**}$.

[c]: Since g_k^* and h_k^* are polyhedral by Rockafellar [48, theorem 19.2], it follows from (139) and Rockafellar [48, theorem 20.1] that $(g_k^* + h_k^*)^* = g_k^{**} \,\square\, h_k^{**}$. □

Appendix

In this section, $\mathcal{K}$ is a real Hilbert space.

Lemma A.1. *Let $A : \mathcal{K} \to 2^{\mathcal{K}}$ be maximally monotone, let $(x_n)_{n\in\mathbb{N}}$ be a bounded sequence in $\mathcal{K}$, and let $(\gamma_n)_{n\in\mathbb{N}}$ be a bounded sequence in $]0, +\infty[$. Then $(J_{\gamma_n A}x_n)_{n\in\mathbb{N}}$ is bounded.*

Proof. Fix $x \in \mathcal{K}$. Using the triangle inequality, the nonexpansiveness of $(J_{\gamma_n A})_{n\in\mathbb{N}}$, and Bauschke and Combettes [9, proposition 23.31(iii)], we obtain $(\forall n \in \mathbb{N}) \quad \|J_{\gamma_n A}x_n - J_A x\| \leqslant \|J_{\gamma_n A}x_n - J_{\gamma_n A}x\| + \|J_{\gamma_n A}x - J_A x\| \leqslant \|x_n - x\| + |1 - \gamma_n|\, \|J_A x - x\| \leqslant \|x\| + \sup_{m\in\mathbb{N}}\|x_m\| + (1 + \sup_{m\in\mathbb{N}}\gamma_m)\|J_A x - x\|$. □

Lemma A.2. *Let $\alpha \in [0, +\infty[$, let $A : \mathcal{K} \to \mathcal{K}$ be α-Lipschitzian, let $\sigma \in]0, +\infty[$, and let $\gamma \in]0, 1/(\alpha + \sigma)]$. Then $\gamma^{-1}\mathrm{Id} - A$ is σ-strongly monotone.*

Proof. By Cauchy–Schwarz,

$$\begin{aligned}(\forall x \in \mathcal{K})(\forall y \in \mathcal{K}) \quad & \langle x - y \mid (\gamma^{-1}\mathrm{Id} - A)x - (\gamma^{-1}\mathrm{Id} - A)y\rangle \\ &= \gamma^{-1}\|x - y\|^2 - \langle x - y \mid Ax - Ay\rangle \\ &\geqslant (\alpha + \sigma)\|x - y\|^2 - \|x - y\|\,\|Ax - Ay\| \\ &\geqslant (\alpha + \sigma)\|x - y\|^2 - \alpha\|x - y\|^2 \\ &= \sigma\|x - y\|^2,\end{aligned} \tag{A.1}$$

which proves the assertion. □

Lemma A.3. *Let I be a nonempty finite set, let $(I_n)_{n\in\mathbb{N}}$ be nonempty subsets of I, let $P\in\mathbb{N}$, and let $(x_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{K}$. Suppose that $\sum_{n\in\mathbb{N}}\|x_{n+1}-x_n\|^2<+\infty$, $I_0=I$, and $(\forall n\in\mathbb{N})\ \bigcup_{j=n}^{n+P}I_j=I$. Furthermore, let $T\in\mathbb{N}$, let $i\in I$, and let $(\pi_i(n))_{n\in\mathbb{N}}$ be a sequence in $\mathbb{N}$ such that $(\forall n\in\mathbb{N})\ n-T\leqslant\pi_i(n)\leqslant n$. For every $n\in\mathbb{N}$, set $\overline{\vartheta}_i(n)=\max\{j\in\mathbb{N}\mid j\leqslant n \text{ and } i\in I_j\}$ and $\vartheta_i(n)=\pi_i(\overline{\vartheta}_i(n))$. Then $x_{\vartheta_i(n)}-x_n\to 0$.*

Proof. For every integer $n\geqslant P$, since $i\in\bigcup_{j=n-P}^{n}I_j$, we have $n\leqslant\overline{\vartheta}_i(n)+P\leqslant\pi_i(\overline{\vartheta}_i(n))+P+T=\vartheta_i(n)+P+T$. Hence, $\vartheta_i(n)\to+\infty$, and therefore, $\sum_{j=\vartheta_i(n)}^{\vartheta_i(n)+P+T}\|x_{j+1}-x_j\|^2\to 0$. However, it results from our assumption that $(\forall n\in\mathbb{N})\ \vartheta_i(n)=\pi_i(\overline{\vartheta}_i(n))\leqslant\overline{\vartheta}_i(n)\leqslant n$. We thus deduce from the triangle and Cauchy–Schwarz inequalities that

$$\|x_n-x_{\vartheta_i(n)}\|^2\leqslant\left|\sum_{j=\vartheta_i(n)}^{\vartheta_i(n)+P+T}\|x_{j+1}-x_j\|\right|^2\leqslant(P+T+1)\sum_{j=\vartheta_i(n)}^{\vartheta_i(n)+P+T}\|x_{j+1}-x_j\|^2\to 0. \tag{A.2}$$

Consequently, $x_{\vartheta_i(n)}-x_n\to 0$. □

Lemma A.4 (Combettes [22]). *Let Z be a nonempty closed convex subset of $\mathcal{K}$, $x_0\in\mathcal{K}$, and $\varepsilon\in\,]0,1[$. Suppose that*

$$\begin{array}{l}
\text{for } n=0,1,\ldots\\
\left\lfloor\begin{array}{l}
t_n^*\in\mathcal{K} \text{ and } \eta_n\in\mathbb{R} \text{ satisfy } Z\subset H_n=\{x\in\mathcal{K}\mid\langle x\,|\,t_n^*\rangle\leqslant\eta_n\};\\
\Delta_n=\langle x_n\,|\,t_n^*\rangle-\eta_n;\\
\text{if } \Delta_n>0\\
\left\lfloor\begin{array}{l}
\lambda_n\in[\varepsilon,2-\varepsilon];\\
x_{n+1}=x_n-(\lambda_n\Delta_n/\|t_n^*\|^2)\,t_n^*;
\end{array}\right.\\
\text{else}\\
\left\lfloor x_{n+1}=x_n.\right.
\end{array}\right.
\end{array} \tag{A.3}$$

Then the following hold:

(i) $(\forall z\in Z)(\forall n\in\mathbb{N})\ \|x_{n+1}-z\|\leqslant\|x_n-z\|$.

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

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

We now revisit ideas found in Bauschke and Combettes [8] and Combettes [21] in a format that is more suited for our purposes.

Lemma A.5. *Let Z be a nonempty closed convex subset of $\mathcal{K}$ and let $x_0\in\mathcal{K}$. Suppose that*

$$\begin{array}{l}
\text{for } n=0,1,\ldots\\
\left\lfloor\begin{array}{l}
t_n^*\in\mathcal{K} \text{ and } \eta_n\in\mathbb{R} \text{ satisfy } Z\subset H_n=\{x\in\mathcal{K}\mid\langle x\,|\,t_n^*\rangle\leqslant\eta_n\};\\
\Delta_n=\langle x_n\,|\,t_n^*\rangle-\eta_n;\\
\text{if } \Delta_n>0\\
\left\lfloor\begin{array}{l}
\tau_n=\|t_n^*\|^2;\ \varsigma_n=\|x_0-x_n\|^2;\ \chi_n=\langle x_0-x_n\,|\,t_n^*\rangle;\ \rho_n=\tau_n\varsigma_n-\chi_n^2;\\
\text{if } \rho_n=0\\
\left\lfloor \kappa_n=1;\ \lambda_n=\Delta_n/\tau_n;\right.\\
\text{else}\\
\left\lfloor\begin{array}{l}
\text{if } \chi_n\Delta_n\geqslant\rho_n\\
\left\lfloor \kappa_n=0;\ \lambda_n=(\Delta_n+\chi_n)/\tau_n;\right.\\
\text{else}\\
\left\lfloor \kappa_n=1-\chi_n\Delta_n/\rho_n;\ \lambda_n=\varsigma_n\Delta_n/\rho_n;\right.
\end{array}\right.\\
x_{n+1}=(1-\kappa_n)x_0+\kappa_nx_n-\lambda_nt_n^*;
\end{array}\right.\\
\text{else}\\
\left\lfloor x_{n+1}=x_n.\right.
\end{array}\right.
\end{array} \tag{A.4}$$

Then the following hold:

(i) $(\forall n\in\mathbb{N})\ \|x_n-x_0\|\leqslant\|x_{n+1}-x_0\|\leqslant\|\mathrm{proj}_Zx_0-x_0\|$.

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

(iii) *Suppose that, for every $x\in\mathcal{K}$ and every strictly increasing sequence $(k_n)_{n\in\mathbb{N}}$ in $\mathbb{N}$, $x_{k_n}\rightharpoonup x\Rightarrow x\in Z$. Then $x_n\to\mathrm{proj}_Zx_0$.*

Proof. Define $(\forall n \in \mathbb{N})\ G_n = \{x \in \mathcal{K} \mid \langle x - x_n \mid x_0 - x_n \rangle \leqslant 0\}$. Then, by virtue of (A.4),

$$(\forall n \in \mathbb{N}) \quad x_n = \mathrm{proj}_{G_n} x_0 \quad \text{and} \quad \Big[\Delta_n > 0 \ \Rightarrow\ \mathrm{proj}_{H_n} x_n = x_n - (\Delta_n/\|t_n^*\|^2)\, t_n^*\Big]. \tag{A.5}$$

Let us establish that

$$(\forall n \in \mathbb{N}) \quad Z \subset H_n \cap G_n \quad \text{and} \quad x_{n+1} = \mathrm{proj}_{H_n \cap G_n} x_0. \tag{A.6}$$

Since $G_0 = \mathcal{K}$, (A.4) yields $Z \subset H_0 = H_0 \cap G_0$. Hence, we derive from (A.5) and (A.4) that $\Delta_0 > 0 \Rightarrow [\mathrm{proj}_{H_0} x_0 = x_0 - (\Delta_0/\tau_0)\, t_0^*$ and $\rho_0 = 0] \Rightarrow [\mathrm{proj}_{H_0} x_0 = x_0 - (\Delta_0/\tau_0)\, t_0^*,\ \kappa_0 = 1,$ and $\lambda_0 = \Delta_0/\tau_0] \Rightarrow x_1 = x_0 - (\Delta_0/\tau_0)\, t_0^* = \mathrm{proj}_{H_0}\ x_0 = \mathrm{proj}_{H_0 \cap G_0} x_0$. On the other hand, $\Delta_0 \leqslant 0 \Rightarrow x_1 = x_0 \in H_0 = H_0 \cap G_0 \Rightarrow x_1 = \mathrm{proj}_{H_0 \cap G_0} x_0$. Now assume that, for some integer $n \geqslant 1$, $Z \subset H_{n-1} \cap G_{n-1}$ and $x_n = \mathrm{proj}_{H_{n-1} \cap G_{n-1}} x_0$. Then, according to Bauschke and Combettes [9, theorem 3.16], $Z \subset H_{n-1} \cap G_{n-1} \subset \{x \in \mathcal{K} \mid \langle x - x_n \mid x_0 - x_n \rangle \leqslant 0\} = G_n$. In turn, (A.4) entails that $Z \subset H_n \cap G_n$. Next, it follows from (A.4), (A.5), and Bauschke and Combettes [9, proposition 29.5] that $\Delta_n \leqslant 0 \Rightarrow [x_{n+1} = x_n$ and $\mathrm{proj}_{G_n} x_0 = x_n \in H_n] \Rightarrow x_{n+1} = \mathrm{proj}_{G_n} x_0 = \mathrm{proj}_{H_n \cap G_n} x_0$. To complete the induction argument, it remains to verify that $\Delta_n > 0 \Rightarrow x_{n+1} = \mathrm{proj}_{H_n \cap G_n} x_0$. Assume that $\Delta_n > 0$ and set

$$y_n = \mathrm{proj}_{H_n} x_n, \quad \widetilde{\chi}_n = \langle x_0 - x_n \mid x_n - y_n \rangle, \quad \widetilde{\nu}_n = \|x_n - y_n\|^2, \quad \text{and} \quad \widetilde{\rho}_n = \varsigma_n \widetilde{\nu}_n - \widetilde{\chi}_n^2. \tag{A.7}$$

Since $\Delta_n > 0$, we have $H_n = \{x \in \mathcal{K} \mid \langle x - y_n \mid x_n - y_n \rangle \leqslant 0\}$ and $y_n = x_n - \theta_n t_n^*$, where $\theta_n = \Delta_n/\tau_n > 0$. In turn, we infer from (A.7) and (A.4) that

$$\widetilde{\chi}_n = \theta_n \chi_n, \quad \widetilde{\nu}_n = \theta_n^2 \tau_n = \theta_n \Delta_n, \quad \text{and} \quad \widetilde{\rho}_n = \theta_n^2 \rho_n. \tag{A.8}$$

Furthermore, (A.4) and the Cauchy–Schwarz inequality ensure that $\rho_n \geqslant 0$, which leads to two cases.

- $\rho_n = 0$: On the one hand, (A.4) asserts that $x_{n+1} = x_n - (\Delta_n/\tau_n)\, t_n^* = y_n$. On the other hand, (A.8) yields $\widetilde{\rho}_n = 0$ and, therefore, since $H_n \cap G_n \neq \emptyset$, Bauschke and Combettes [9, corollary 29.25(ii)] yields $\mathrm{proj}_{H_n \cap G_n} x_0 = y_n$. Altogether, $x_{n+1} = \mathrm{proj}_{H_n \cap G_n} x_0$.
- $\rho_n > 0$: By (A.8), $\widetilde{\rho}_n > 0$. First, suppose that $\chi_n \Delta_n \geqslant \rho_n$. It follows from (A.4) that $x_{n+1} = x_0 - ((\Delta_n + \chi_n)/\tau_n)\, t_n^*$ and from (A.8) that $\widetilde{\chi}_n \widetilde{\nu}_n = \theta_n^2 \chi_n \Delta_n \geqslant \theta_n^2 \rho_n = \widetilde{\rho}_n$. Thus, Bauschke and Combettes [9, corollary 29.25(ii)] and (A.8) imply that

$$\begin{aligned}
\mathrm{proj}_{H_n \cap G_n} x_0 &= x_0 + \left(1 + \frac{\widetilde{\chi}_n}{\widetilde{\nu}_n}\right)(y_n - x_n) \\
&= x_0 - \left(1 + \frac{\chi_n}{\theta_n \tau_n}\right)\theta_n t_n^* \\
&= x_0 - \frac{\theta_n \tau_n + \chi_n}{\tau_n}\, t_n^* \\
&= x_0 - \frac{\Delta_n + \chi_n}{\tau_n}\, t_n^* \\
&= x_{n+1}.
\end{aligned} \tag{A.9}$$

Now suppose that $\chi_n \Delta_n < \rho_n$. Then $\widetilde{\chi}_n \widetilde{\nu}_n < \widetilde{\rho}_n$, and hence, it results from Bauschke and Combettes [9, corollary 29.25(ii)], (A.8), and (A.4) that

$$\begin{aligned}
\mathrm{proj}_{H_n \cap G_n} x_0 &= x_n + \frac{\widetilde{\nu}_n}{\widetilde{\rho}_n}\big(\widetilde{\chi}_n (x_0 - x_n) + \varsigma_n (y_n - x_n)\big) \\
&= \frac{\widetilde{\chi}_n \widetilde{\nu}_n}{\widetilde{\rho}_n}\, x_0 + \left(1 - \frac{\widetilde{\chi}_n \widetilde{\nu}_n}{\widetilde{\rho}_n}\right) x_n + \frac{\widetilde{\nu}_n \varsigma_n}{\widetilde{\rho}_n}(y_n - x_n) \\
&= \frac{\chi_n \Delta_n}{\rho_n}\, x_0 + \left(1 - \frac{\chi_n \Delta_n}{\rho_n}\right) x_n - \frac{\tau_n \varsigma_n}{\rho_n} \frac{\Delta_n}{\tau_n}\, t_n^* \\
&= x_{n+1}.
\end{aligned} \tag{A.10}$$

(i): Let $n \in \mathbb{N}$. We derive from (A.6) that $\|x_{n+1} - x_0\| = \|\mathrm{proj}_{H_n \cap G_n} x_0 - x_0\| \leqslant \|\mathrm{proj}_Z x_0 - x_0\|$. On the other hand, since $x_{n+1} \in G_n$ by virtue of (A.6), we have

$$\begin{aligned}
\|x_n - x_0\|^2 + \|x_{n+1} - x_n\|^2 &\leqslant \|x_n - x_0\|^2 + \|x_{n+1} - x_n\|^2 + 2\langle x_{n+1} - x_n \mid x_n - x_0 \rangle \\
&= \|x_{n+1} - x_0\|^2.
\end{aligned} \tag{A.11}$$

(ii): Let $N \in \mathbb{N}$. In view of (A.11) and (i), $\sum_{n=0}^{N} \|x_{n+1} - x_n\|^2 \leqslant \sum_{n=0}^{N} (\|x_{n+1} - x_0\|^2 - \|x_n - x_0\|^2) = \|x_{N+1} - x_0\|^2 \leqslant \|\mathrm{proj}_Z x_0 - x_0\|^2$. Therefore, $\sum_{n\in\mathbb{N}} \|x_{n+1} - x_n\|^2 < +\infty$. However, for every $n \in \mathbb{N}$, since (A.6) asserts that $x_{n+1} \in H_n$, we have $\|\mathrm{proj}_{H_n} x_n - x_n\| \leqslant \|x_{n+1} - x_n\|$. Thus, $\sum_{n\in\mathbb{N}} \|\mathrm{proj}_{H_n} x_n - x_n\|^2 < +\infty$.

(iii): It results from (i) that $(x_n)_{n\in\mathbb{N}}$ is bounded. Now let $x \in \mathcal{K}$, let $(k_n)_{n\in\mathbb{N}}$ be a strictly increasing sequence in $\mathbb{N}$, and suppose that $x_{k_n} \rightharpoonup x$. Using Bauschke and Combettes [9, lemma 2.42] and (i), we deduce that $\|x - x_0\| \leqslant \underline{\lim}\|x_{k_n} - x_0\| \leqslant \|\mathrm{proj}_Z x_0 - x_0\|$. Thus, since it results from our assumption that $x \in Z$, we have $x = \mathrm{proj}_Z x_0$, which implies that $x_n \rightharpoonup \mathrm{proj}_Z x_0$ (Bauschke and Combettes [9, lemma 2.46]). In turn, since $\overline{\lim}\|x_n - x_0\| \leqslant \|\mathrm{proj}_Z x_0 - x_0\|$ by (i), Bauschke and Combettes [9, lemma 2.51(i)] forces $x_n \to \mathrm{proj}_Z x_0$. □

References

[1] Alotaibi A, Combettes PL, Shahzad N (2014) Solving coupled composite monotone inclusions by successive Fejér approximations of their Kuhn-Tucker set. *SIAM J. Optim.* 24:2076–2095.
[2] Attouch H, Bolte J, Redont P, Soubeyran A (2008) Alternating proximal algorithms for weakly coupled convex minimization problems. Applications to dynamical games and PDE's. *J. Convex Anal.* 15:485–506.
[3] Attouch H, Briceño-Arias LM, Combettes PL (2010) A parallel splitting method for coupled monotone inclusions. *SIAM J. Control Optim.* 48:3246–3270.
[4] Attouch H, Briceño-Arias LM, Combettes PL (2016) A strongly convergent primal-dual method for nonoverlapping domain decomposition. *Numer. Math.* 133:433–470.
[5] Aujol J-F, Chambolle A (2005) Dual norms and image decomposition models. *Int. J. Comput. Vision* 63:85–104.
[6] Bach F, Jenatton R, Mairal J, Obozinski G (2012) Optimization with sparsity-inducing penalties. *Found. Trends Machine Learn.* 4:1–106.
[7] Banert S, Ringh A, Adler J, Karlsson J, Öktem O (2020) Data-driven nonsmooth optimization. *SIAM J. Optim.* 30:102–131.
[8] Bauschke HH, Combettes PL (2001) A weak-to-strong convergence principle for Fejér-monotone methods in Hilbert spaces. *Math. Oper. Res.* 26:248–264.
[9] Bauschke HH, Combettes PL (2017) *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, 2nd ed. (Springer, New York).
[10] Boţ RI, Hendrich C (2013) A Douglas-Rachford type primal-dual method for solving inclusions with mixtures of composite and parallel-sum type monotone operators. *SIAM J. Optim.* 23:2541–2565.
[11] Boţ RI, Hendrich C (2014) Convergence analysis for a primal-dual monotone + skew splitting algorithm with applications to total variation minimization. *J. Math. Imaging Vision* 49:551–568.
[12] Briceño-Arias LM, Chierchia G, Chouzenoux E, Pesquet J-C (2019) A random block-coordinate Douglas–Rachford splitting method with low computational complexity for binary logistic regression. *Comput. Optim. Appl.* 72:707–726.
[13] Briceño-Arias LM, Combettes PL (2009) Convex variational formulation with smooth coupling for multicomponent signal decomposition and recovery. *Numer. Math. Theory Methods Appl.* 2:485–508.
[14] Briceño-Arias LM, Combettes PL (2011) A monotone+skew splitting model for composite monotone inclusions in duality. *SIAM J. Optim.* 21:1230–1250.
[15] Briceño-Arias LM, Combettes PL (2013) Monotone operator methods for Nash equilibria in non-potential games. Bailey D et al., eds. *Computational and Analytical Mathematics* (Springer, New York), 143–159.
[16] Briceño-Arias LM, Combettes PL, Pesquet J-C, Pustelnik N (2011) Proximal algorithms for multicomponent image recovery problems. *J. Math. Imaging Vision* 41:3–22.
[17] Briceño-Arias LM, Davis D (2018) Forward-backward-half forward algorithm for solving monotone inclusions. *SIAM J. Optim.* 28:2839–2871.
[18] Briceño-Arias LM, Kalise D, Silva FJ (2018) Proximal methods for stationary mean field games with local couplings. *SIAM J. Control Optim.* 56:801–836.
[19] Bùi MN, Combettes PL (2020) Warped proximal iterations for monotone inclusions. *J. Math. Anal. Appl.* 491:124315.
[20] Burachik RS, Jeyakumar V (2005) A new geometric condition for Fenchel's duality in infinite dimensional spaces. *Math. Program.* B104: 229–233.
[21] Combettes PL (2000) Strong convergence of block-iterative outer approximation methods for convex optimization. *SIAM J. Control Optim.* 38:538–565.
[22] Combettes PL (2001) Fejér-monotonicity in convex optimization. Floudas CA, Pardalos PM, eds. *Encyclopedia of Optimization*, vol. 2 (Springer-Verlag, New York), 106–114.
[23] Combettes PL (2013) Systems of structured monotone inclusions: Duality, algorithms, and applications. *SIAM J. Optim.* 23:2420–2447.
[24] Combettes PL (2018) Monotone operator theory in convex optimization. *Math. Program.* B170:177–206.
[25] Combettes PL, Eckstein J (2018) Asynchronous block-iterative primal-dual decomposition methods for monotone inclusions. *Math. Program.* B168:645–672.
[26] Combettes PL, Müller CL (2020) Perspective maximum likelihood-type estimation via proximal decomposition. *Electron. J. Stat.* 14:207–238.
[27] Combettes PL, Pesquet J-C (2020) Deep neural network structures solving variational inequalities. *Set-Valued Var. Anal.* 28:491–518.
[28] Combettes PL, Wajs VR (2005) Signal recovery by proximal forward-backward splitting. *Multiscale Model. Simul.* 4:1168–1200.
[29] Eckstein J (2017) A simplified form of block-iterative operator splitting, and an asynchronous algorithm resembling the multi-block alternating direction method of multipliers. *J. Optim. Theory Appl.* 173:155–182.
[30] Facchinei F, Pang J-S (2003) *Finite-Dimensional Variational Inequalities and Complementarity Problems* (Springer-Verlag, New York).
[31] Fukushima M (1996) The primal Douglas-Rachford splitting algorithm for a class of monotone mappings with applications to the traffic equilibrium problem. *Math. Program.* 72:1–15.
[32] Ghoussoub N (2009) *Self-Dual Partial Differential Systems and Their Variational Principles* (Springer-Verlag, New York).
[33] Hintermüller M, Stadler G (2006) An infeasible primal-dual algorithm for total bounded variation-based inf-convolution-type image restoration. *SIAM J. Sci. Comput.* 28:1–23.
[34] Johnstone PR, Eckstein J (2020) Projective splitting with forward steps. *Math. Program.*
[35] Johnstone PR, Eckstein J (2021) Single-forward-step projective splitting: Exploiting cocoercivity. *Comput. Optim. Appl.* 78:125–166.

[36] Li M, Yuan X (2015) A strictly contractive Peaceman-Rachford splitting method with logarithmic-quadratic proximal regularization for convex programming. *Math. Oper. Res.* 40:842–858.
[37] Lions PL, Mercier B (1979) Splitting algorithms for the sum of two nonlinear operators. *SIAM J. Numer. Anal.* 16:964–979.
[38] Mercier B (1979) *Topics in Finite Element Solution of Elliptic Problems*. Lectures on Mathematics, no. 63 (Tata Institute of Fundamental Research, Bombay, India).
[39] Mizoguchi T, Yamada I (2019) Hypercomplex tensor completion via convex optimization. *IEEE Trans. Signal Process.* 67:4078–4092.
[40] Nam NM, Nguyen TA, Rector RB, Sun J (2014) Nonsmooth algorithms and Nesterov's smoothing technique for generalized Fermat-Torricelli problems. *SIAM J. Optim.* 24:1815–1839.
[41] Ono S, Miyata T, Yamada I (2014) Cartoon-texture image decomposition using blockwise low-rank texture characterization. *IEEE Trans. Signal Process.* 23:1128–1142.
[42] Papadakis N, Peyré G, Oudet E (2014) Optimal transport with proximal splitting. *SIAM J. Imaging Sci.* 7:212–238.
[43] Passty GB (1979) Ergodic convergence to a zero of the sum of monotone operators in Hilbert space. *J. Math. Anal. Appl.* 72:383–390.
[44] Pennanen T (2000) Dualization of generalized equations of maximal monotone type. *SIAM J. Optim.* 10:809–835.
[45] Pesquet J-C, Repetti A (2015) A class of randomized primal-dual algorithms for distributed optimization. *J. Nonlinear Convex Anal.* 16: 2453–2490.
[46] Robinson SM (1999) Composition duality and maximal monotonicity. *Math. Program.* 85:1–13.
[47] Rockafellar RT (1967) Duality and stability in extremum problems involving convex functions. *Pacific J. Math.* 21:167–187.
[48] Rockafellar RT (1970) *Convex Analysis* (Princeton University Press, Princeton, NJ).
[49] Rockafellar RT (1970) Monotone operators associated with saddle-functions and minimax problems. Browder FE, ed. *Nonlinear Functional Analysis, Part 1* (AMS, Providence, RI), 241–250.
[50] Rockafellar RT (1971) Saddle points and convex analysis. Kuhn HW, Szegö GP, eds. *Differential Games and Related Topics* (North-Holland, New York), 109–127.
[51] Rockafellar RT (1974) *Conjugate Duality and Optimization* (SIAM, Philadelphia).
[52] Rosasco L, Villa S, Vũ BC (2016) A stochastic inertial forward-backward splitting algorithm for multivariate monotone inclusions. *Optimization* 65:1293–1314.
[53] Tseng P (2000) A modified forward-backward splitting method for maximal monotone mappings. *SIAM J. Control Optim.* 38:431–446.
[54] Wang X, Zhang J, Zhang W (2020) The distance between convex sets with Minkowski sum structure: Application to collision detection. *Comput. Optim. Appl.* 77:465–490.
[55] Yan X, Bien J (2021) Rare feature selection in high dimensions. *J. Amer. Statist. Assoc*. 116:887–900.
[56] Yi P, Pavel L (2019) An operator splitting approach for distributed generalized Nash equilibria computation. *Automatica* 102: 111–121.