Abstract

Kuhn-Tucker points play a fundamental role in the analysis and the numerical solution of monotone inclusion problems, providing in particular both primal and dual solutions. We propose a class of strongly convergent algorithms for constructing the best approximation to a reference point from the set of Kuhn-Tucker points of a general Hilbertian composite monotone inclusion problem. Applications to systems of coupled monotone inclusions are presented. Our framework does not impose additional assumptions on the operators present in the formulation, and it does not require knowledge of the norm of the linear operators involved in the compositions or the inversion of linear operators.

Keywords best approximation, duality, Haugazeau, monotone operator, primal-dual algorithm, splitting algorithm, strong convergence

Mathematics Subject Classifications (2010) Primary 47H05, 41A50; Secondary 65K05, 41A65, 90C25.

*Contact author: P. L. Combettes, plc@ljll.math.upmc.fr, phone: +33 1 4427 6319, fax: +33 1 4427 7200.
1 Introduction

Let $H$ and $G$ be real Hilbert spaces, let $L: H \to G$ be a bounded linear operator, and let $f: H \to [-\infty, +\infty]$ and $g: G \to [-\infty, +\infty]$ be proper lower semicontinuous convex functions. Classical Fenchel-Rockafellar duality [27] concerns the interplay between the optimization problem

$$\min_{x \in H} f(x) + g(Lx)$$

(1.1)

and its dual

$$\min_{v^* \in G} f^*(-L^*v^*) + g^*(v^*).$$

(1.2)

An essential ingredient in the analysis of such dual problems is the associated Kuhn-Tucker set [28]

$$Z = \{(x, v^*) \in H \oplus G \mid -L^*v^* \in \partial f(x) \text{ and } Lx \in \partial g^*(v^*)\},$$

(1.3)

which involves the maximally monotone subdifferential operators $\partial f$ and $\partial g^*$. A fruitful generalization of (1.1)–(1.2) is obtained by pairing the inclusion $0 \in Ax + L^*BLx$ on $H$ with the dual inclusion $0 \in -LA^{-1}(-L^*v^*) + B^{-1}v^*$ on $G$, where $A$ and $B$ are maximally monotone operators acting on $H$ and $G$, respectively. Such operator duality has been studied in [18, 24, 25, 26] and the first splitting algorithm for solving such composite inclusions was proposed in [11]. The strategy adopted in that paper was to use a standard 2-operator splitting method to construct a point in the Kuhn-Tucker set $Z = \{(x, v^*) \in H \oplus G \mid -L^*v^* \in \partial f(x) \text{ and } Lx \in \partial g^*(v^*)\}$ and hence obtain a primal-dual solution (see also [9, 14, 16, 17, 30] for variants of this approach). In [2] we investigated a different strategy based on an idea first proposed in [19] for solving the inclusion $0 \in Ax + Bx$. In this framework, at each iteration, one uses points in the graphs of $A$ and $B$ to construct a closed affine half-space of $H \oplus G$ containing $Z$; the primal-dual update is then obtained as the projection of the current iterate onto it. The resulting Fejér-monotone algorithm provides only weak convergence to an unspecified Kuhn-Tucker point. In the present paper we propose a strongly convergent modification of these methods for solving the following best approximation problem.

**Problem 1.1** Let $H$ and $G$ be real Hilbert spaces, and set $K = H \oplus G$. Let $A: H \to 2^H$ and $B: G \to 2^G$ be maximally monotone operators, and let $L: H \to G$ be a bounded linear operator. Let $(x_0, v_0^*) \in K$, assume that the inclusion problem

$$\text{find } x \in H \text{ such that } 0 \in Ax + L^*BLx$$

(1.4)

has at least one solution, and consider the dual problem

$$\text{find } v^* \in G \text{ such that } 0 \in -LA^{-1}(-L^*v^*) + B^{-1}v^*.$$ 

(1.5)

The problem is to find the best approximation $(x, v^*)$ to $(x_0, v_0^*)$ from the associated Kuhn-Tucker set

$$Z = \{(x, v^*) \in K \mid -L^*v^* \in Ax \text{ and } Lx \in B^{-1}v^*\}.$$ 

(1.6)
The principle of our algorithm goes back to the work of Yves Haugazeau [21] for finding the projection of a point onto the intersection of closed convex sets by means of projections onto the individual sets. Haugazeau’s method was generalized in several directions and applied to a variety of problems in nonlinear analysis and optimization in [13]. In [6], it was formulated as an abstract convergence principle for turning a class of weakly convergent methods into strongly convergent ones (see also [22] for recent related work). In the area of monotone inclusions, Haugazeau-like methods were used in [29] for solving $x \in A^{-1}0$ and in [6] for solving $x \in \bigcap_{i=1}^{m} A_i^{-1}0$. They were also used in splitting method for solving $0 \in Ax + Bx$ as a modification of the forward-backward splitting algorithm in [15] and [7, Corollary 29.5], and as a modification of the Douglas-Rachford algorithm in [8] and [31].

The paper is organized as follows. Section 2 is devoted to a version of an abstract Haugazeau principle. The algorithms for solving Problem 1.1 are presented in Section 3, where their strong convergence is established. In Section 4, we present an extension to systems of coupled monotone inclusions and consider applications to the relaxation of inconsistent common zero problems and to structured multivariate convex minimization problems.

**Notation.** Our notation is standard and follows [7], where the necessary background on monotone operators and convex analysis is available. The scalar product of a Hilbert space is denoted by $\langle \cdot | \cdot \rangle$ and the associated norm by $\| \cdot \|$. We denote respectively by $\rightarrow$ and $\rightharpoonup$ weak and strong convergence, and by $\text{Id}$ the identity operator. Let $\mathcal{H}$ and $\mathcal{G}$ be real Hilbert space. The Hilbert direct sum of $\mathcal{H}$ and $\mathcal{G}$ is denoted by $\mathcal{H} \oplus \mathcal{G}$, and the power set of $\mathcal{H}$ by $2^\mathcal{H}$. Now let $A: \mathcal{H} \rightarrow 2^\mathcal{H}$. Then ran $A$ is the range $A$, gra $A$ the graph of $A$, $A^{-1}$ the inverse of $A$, and $J_A = (\text{Id} + A)^{-1}$ the resolvent of $A$. The projection operator onto a nonempty closed convex subset $C$ of $\mathcal{H}$ is denoted by $P_C$ and $\Gamma_0(\mathcal{H})$ is the class of proper lower semicontinuous convex functions from $\mathcal{H}$ to $]-\infty, +\infty]$. Let $f \in \Gamma_0(\mathcal{H})$. The conjugate of $f$ is $f^*: u^* \mapsto \sup_{x \in \mathcal{H}} \langle x | u^* \rangle - f(x)$ and the subdifferential of $f$ is $\partial f: \mathcal{H} \rightarrow 2^\mathcal{H}: x \mapsto \{ u^* \in \mathcal{H} \mid (\forall y \in \mathcal{H}) \langle y - x | u^* \rangle + f(x) \leq f(y) \}$.

## 2 An abstract Haugazeau algorithm

In [21, Théorème 3-2] Haugazeau proposed an ingenious method for projecting a point onto the intersection of closed convex sets in a Hilbert space using the projections onto the individual sets. Abstract versions of his method for projecting onto a closed convex set in a real Hilbert space were devised in [13] and [6]. In this section, we present a formulation of this abstract principle which is better suited for our purposes.

Let $\mathcal{H}$ be a real Hilbert space. Given an ordered triplet $(x, y, z) \in \mathcal{H}^3$, we define

$$H(x, y) = \{ h \in \mathcal{H} \mid \langle h - y | x - y \rangle \leq 0 \}. \tag{2.1}$$

Moreover, if $R = H(x, y) \cap H(y, z) \neq \emptyset$, we denote by $Q(x, y, z)$ the projection of $x$ onto $R$. The principle of the algorithm to project a point $x_0 \in \mathcal{H}$ onto a nonempty closed convex set $C \subset \mathcal{H}$ is to use at iteration $n$ the current iterate $x_n$ to construct an outer approximation to $C$ of the form $H(x_n, x_{n+1}) \cap H(x_n, x_{n+1/2})$; the update is then computed as the projection of $x_0$ onto it, i.e., $x_{n+1} = Q(x_0, x_n, x_{n+1/2})$. 

3
**Proposition 2.1** Let $C$ be a nonempty closed convex subset of $\mathcal{H}$ and let $x_0 \in \mathcal{H}$. Iterate

\[
\text{for } n = 0, 1, \ldots \\
\begin{align*}
take & \ x_{n+1/2} \in \mathcal{H} \text{ such that } C \subset H(x_n, x_{n+1/2}) \\
& x_{n+1} = Q(x_0, x_n, x_{n+1/2}).
\end{align*} 
\tag{2.2}
\]

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

(i) $(\forall n \in \mathbb{N}) \ C \subset H(x_0, x_n) \cap H(x_n, x_{n+1/2})$.

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

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

(iv) Suppose that, for every $x \in \mathcal{H}$ and every strictly increasing sequence $(k_n)_{n \in \mathbb{N}}$ in $\mathbb{N}$, $x_{k_n} \rightarrow x$ implies $x \in C$. Then $x_n \rightarrow P_Cx_0$.

**Proof.** The proof is similar to those found in [6, Section 3] and [13, Section 3]. First, recall that the projector onto a nonempty closed convex subset $D$ of $\mathcal{H}$ is characterized by [7, Theorem 3.14]

$$(\forall x \in \mathcal{H}) \quad P_Dx \in D \quad \text{and} \quad D \subset H(x, P_Dx).$$

(i): Let $n \in \mathbb{N}$ be such that $x_n$ exists. Since by construction $C \subset H(x_n, x_{n+1/2})$, it is enough to show that $C \subset H(x_0, x_n)$. This inclusion is trivially true for $n = 0$ since $H(x_0, x_0) = \mathcal{H}$. Furthermore, it follows from (2.3) and (2.2) that

\[
C \subset H(x_0, x_n) \quad \Rightarrow \quad C \subset H(x_0, x_n) \cap H(x_n, x_{n+1/2}) \\
\Rightarrow \quad C \subset H(x_0, Q(x_0, x_n, x_{n+1/2})) \\
\Leftrightarrow \quad C \subset H(x_0, x_{n+1}),
\tag{2.4}
\]

which establishes the assertion by induction. This also shows that $H(x_0, x_n) \cap H(x_n, x_{n+1/2})$ is a nonempty closed convex set and therefore that the projection $x_{n+1}$ of $x_0$ onto it is well defined.

(ii): Let $n \in \mathbb{N}$. By construction, $x_{n+1} = Q(x_0, x_n, x_{n+1/2}) \in H(x_0, x_n) \cap H(x_n, x_{n+1/2})$. Consequently, since $x_n$ is the projection of $x_0$ onto $H(x_0, x_n)$ and $x_{n+1} \in H(x_0, x_n)$, we have $\|x_0 - x_n\| \leq \|x_0 - x_{n+1}\|$. On the other hand, since $P_Cx_0 \in C \subset H(x_0, x_n)$, we have $\|x_0 - x_n\| \leq \|x_0 - P_Cx_0\|$. It follows that $(\|x_0 - x_k\|)_{k \in \mathbb{N}}$ converges and that

\[
\lim \|x_0 - x_k\| \leq \|x_0 - P_Cx_0\|. 
\tag{2.5}
\]

On the other hand, since $x_{n+1} \in H(x_0, x_n)$, we have

\[
\|x_{n+1} - x_n\|^2 \leq \|x_{n+1} - x_n\|^2 + 2 \langle x_{n+1} - x_n, x_n - x_0 \rangle = \|x_0 - x_{n+1}\|^2 - \|x_0 - x_n\|^2. 
\tag{2.6}
\]

Hence, $\sum_{k=1}^{n} \|x_{k+1} - x_k\|^2 \leq \|x_0 - x_{n+1}\|^2 \leq \|x_0 - P_Cx_0\|^2$ and, in turn, $\sum_{k \in \mathbb{N}} \|x_{k+1} - x_k\|^2 < +\infty$. 

\[4\]
(iii): For every \( n \in \mathbb{N} \), we derive from the inclusion \( x_{n+1} \in H(x_n, x_{n+1/2}) \) that

\[
\|x_{n+1/2} - x_n\|^2 \leq \|x_{n+1} - x_{n+1/2}\|^2 + \|x_n - x_{n+1/2}\|^2 \\
\leq \|x_{n+1} - x_{n+1/2}\|^2 + 2\langle x_{n+1} - x_{n+1/2}, x_{n+1/2} - x_n \rangle + \|x_n - x_{n+1/2}\|^2 \\
= \|x_{n+1} - x_n\|^2. \tag{2.7}
\]

Hence, it follows from (ii) that \( \sum_{n \in \mathbb{N}} \|x_{n+1/2} - x_n\|^2 < +\infty \).

(iv): Let us note that (2.5) implies that \((x_n)_{n \in \mathbb{N}}\) is bounded. Now, let \( x \) be a weak sequential cluster point of \((x_n)_{n \in \mathbb{N}}\), say \( x_k \rightharpoonup x \). Then, by weak lower semicontinuity of \( \| \cdot \| \) [7, Lemma 2.35] and (2.5) \( |x_0 - x| \leq \lim_{k \to \infty} |x_0 - x_k| \leq \|x_0 - P_C x_0\| = \inf_{y \in C} \|x_0 - y\| \). Hence, since \( x \in C \), \( x = P_C x_0 \) is the only weak sequential cluster point of the sequence \((x_n)_{n \in \mathbb{N}}\) and it follows from [7, Lemma 2.38] that \( x_n \rightharpoonup P_C x_0 \). In turn (2.5) yields \( \|x_0 - P_C x_0\| \leq \lim_{k \to \infty} \|x_0 - x_k\| = \lim \|x_0 - x_n\| \leq \|x_0 - P_C x_0\| \). Thus, \( x_0 - x_n \rightharpoonup x_0 - P_C x_0 \) and \( x_0 - x_n \rightharpoonup x_0 - P_C x_0 \). We therefore derive from [7, Lemma 2.41(i)] that \( x_0 - x_n \rightharpoonup x_0 - P_C x_0 \), i.e., \( x_n \rightharpoonup P_C x_0 \). \( \square \)

**Remark 2.2** Suppose that, for some \( n \in \mathbb{N} \), \( x_n \in C \) in (2.2). Then \( \|x_0 - P_C x_0\| \leq \|x_0 - x_n\| \) and, since we always have \( \|x_0 - x_n\| \leq \|x_0 - P_C x_0\| \), we conclude that \( x_n = P_C x_0 \) and that the iterations can be stopped.

Algorithm (2.2) can easily be implemented thanks to the following lemma.

**Lemma 2.3** Let \((x, y, z) \in \mathcal{H}^3 \) and set \( R = H(x, y) \cap H(y, z) \). Moreover, set \( \chi = \langle x - y \mid y - z \rangle \), \( \mu = \|x - y\|^2 \), \( \nu = \|y - z\|^2 \), and \( \rho = \mu \nu - \chi^2 \). Then exactly one of the following holds:

(i) \( \rho = 0 \) and \( \chi < 0 \), in which case \( R = \emptyset \).

(ii) \( \rho = 0 \) and \( \chi \geq 0 \) or \( \rho > 0 \), in which case \( R \neq \emptyset \) and

\[
Q(x, y, z) = \begin{cases} 
  z, & \text{if } \rho = 0 \text{ and } \chi \geq 0; \\
  x + (1 + \chi/\nu)(y - z), & \text{if } \rho > 0 \text{ and } \chi \nu \geq \rho; \\
  y + (\nu/\rho)(\chi(x - y) + \mu(z - y)), & \text{if } \rho > 0 \text{ and } \chi \nu < \rho. 
\end{cases} \tag{2.8}
\]

**Proof.** See [21, Théorème 3-1] for the original proof and [7, Corollary 28.21] for an alternate derivation. \( \square \)

### 3 Main result

In this section, we devise a strongly convergent algorithm for solving Problem 1.1 by coupling Proposition 2.1 with the construction of [2] to determine the half-spaces \((H(x_n, x_{n+1/2}))_{n \in \mathbb{N}}\). First, we need a couple of facts.

**Proposition 3.1** [11, Proposition 2.8] In the setting of Problem 1.1, \( Z \) is a nonempty closed convex set and, if \((x, \nu^*) \in Z \), then \( x \) solves (1.4) and \( \nu^* \) solves (1.5).
Proposition 3.2 [2, Proposition 2.4] In the setting of Problem 1.1, let \((a_n, a_n^*)_{n \in \mathbb{N}}\) be a sequence in gra \(A\), let \((b_n, b_n^*)_{n \in \mathbb{N}}\) be a sequence in gra \(B\), and let \((x, v^*) \in \mathcal{K}\). Suppose that \(a_n \rightharpoonup x\), \(b_n^* \rightharpoonup v^*\), \(a_n^* + L^*b_n^* \to 0\), and \(La_n - b_n \to 0\). Then \(\langle a_n \mid a_n^* \rangle + \langle b_n \mid b_n^* \rangle \to 0\) and \((x, v^*) \in \mathbb{Z}\).

The next result features our general algorithm for solving Problem 1.1.

Theorem 3.3 Consider the setting of Problem 1.1. Let \(\varepsilon \in ]0, 1]\), let \(\alpha \in ]0, +\infty]\), and set, for every \((x, v^*) \in \mathcal{K}\),

\[ G_\alpha(x, v^*) = \left\{ (a, b, a^*, b^*) \in \mathcal{K} \times \mathcal{K} \mid (a, a^*) \in \text{gra } A, \ (b, b^*) \in \text{gra } B, \ 
\langle x-a \mid a^* + L^*v^* \rangle + \langle Lx - b \mid b^* - v^* \rangle \geq \alpha (\|a^* + L^*b^*\|^2 + \|La - b\|^2) \right\}. \] (3.1)

Iterate

\[
\text{for } n = 0, 1, \ldots \\
\begin{align*} 
(a_n, b_n, a_n^*, b_n^*) & \in G_\alpha(x_n, v_n^*) \\
n_n & = a_n^* + L^*b_n^* \\
t_n & = b_n - La_n \\
n_n & = \|n_n\|^2 + \|t_n\|^2 \\
if \& n_n = 0 \\
if \& n_n > 0 \\
\lambda_n & \in [\varepsilon, 1] \\
\theta_n & = \lambda_n \left( \langle x_n \mid n_n \rangle + \langle t_n \mid v_n^* \rangle - \langle a_n \mid a_n^* \rangle - \langle b_n \mid b_n^* \rangle \right) / n_n \\
x_{n+1/2} & = x_n - \theta_n n_n \\
n_n & = v_n^* - \theta_n t_n \\
\chi_n & = \langle x_0 - x_n \mid x_n - x_{n+1/2} \rangle + \langle v_n^* - v_n^* \mid v_n^* - v_{n+1/2}^* \rangle \\
\mu_n & = \|x_0 - x_n\|^2 + \|v_0^* - v_n^*\|^2 \\
v_n & = \|x_n - x_{n+1/2}\|^2 + \|v_n^* - v_{n+1/2}^*\|^2 \\
\rho_n & = \mu_n v_n - \chi_n^2 \\
if \& \rho_n = 0 \text{ and } \chi_n \geq 0 \\
x_{n+1} & = x_{n+1/2} \\
v_{n+1} & = v_{n+1/2} \\
if \& \rho_n > 0 \text{ and } \chi_n v_n \geq \rho_n \\
x_{n+1} & = x_0 + (1 + \chi_n / v_n)(x_{n+1/2} - x_n) \\
v_{n+1} & = v_0^* + (1 + \chi_n / v_n)(v_{n+1/2}^* - v_n^*) \\
if \& \rho_n > 0 \text{ and } \chi_n v_n < \rho_n \\
x_{n+1} & = x_n + \left( v_n / \rho_n \right) \left( \chi_n (x_0 - x_n) + \mu_n (x_{n+1/2} - x_n) \right) \\
v_{n+1} & = v_n^* + \left( v_n / \rho_n \right) \left( \chi_n (v_n^* - v_n^*) + \mu_n (v_{n+1/2}^* - v_n^*) \right). 
\end{align*}
\] (3.2)

Then (3.2) generates infinite sequences \((x_n)_{n \in \mathbb{N}}\) and \((v_n^*)_{n \in \mathbb{N}}\), and the following hold:

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

(ii) \(\sum_{n \in \mathbb{N}} \|n_n\|^2 < +\infty\) and \(\sum_{n \in \mathbb{N}} \|t_n\|^2 < +\infty\).
(iii) Suppose that \( x_n - a_n \to 0 \) and \( v^*_n - b^*_n \to 0 \). Then \( x_n \to x \) and \( v^*_n \to x^* \).

Proof. We are going to show that the claims follow from Proposition 2.1 applied in \( K \) to the set \( Z \) of (1.6), which is nonempty, closed, and convex by Proposition 3.1. First, let us set

\[
(\forall n \in \mathbb{N}) \quad x_n = (x_n, v^*_n) \quad \text{and} \quad x_{n+1/2} = (x_{n+1/2}, v^*_{n+1/2}).
\]

We deduce from (3.2) that

\[
(\forall (x, v^*) \in K (\forall n \in \mathbb{N}) ) \quad \langle x | s^*_n \rangle + \langle t_n \ | v^* \rangle - \langle a_n | a^*_n \rangle - \langle b_n \ | b^*_n \rangle = \langle x | a^*_n + L^* b^*_n \rangle + \langle b_n - L a_n \ | v^* \rangle - \langle a_n \ | a^*_n \rangle - \langle b_n \ | b^*_n \rangle = \langle x - a_n \ | a^*_n + L^* v^* \rangle + \langle L x - b_n \ | b^*_n - v^* \rangle.
\]

Next, let us show that

\[
(\forall n \in \mathbb{N}) \quad Z \subset H(x_n, x_{n+1/2}).
\]

To this end, let \( z = (x, v^*) \in Z \) and let \( n \in \mathbb{N} \). We must show that \( \langle z - x_{n+1/2} | x_n - x_{n+1/2} \rangle \leq 0 \). If \( \tau_n = 0 \), then \( x_{n+1/2} = x_n \) and the inequality is trivially satisfied. Now suppose that \( \tau_n > 0 \). Then (3.4) and (3.1) yield

\[
\theta_n = \lambda_n \frac{\langle x_n - a_n | a^*_n + L^* v^*_n \rangle + \langle L x_n - b_n \ | b^*_n - v^*_n \rangle}{\tau_n} \\
\geq \varepsilon \alpha \\
> 0.
\]

On the other hand, it follows from (3.2) and (1.6) that \( a^*_n \in A a_n \) and \(-L^* v^* \in A x \). Hence, since \( A \) is monotone, \( \langle x - a_n | a^*_n + L^* v^* \rangle \leq 0 \). Similarly, since \( v^* \in B(L x) \) and \( b^*_n \in B b_n \), the monotonicity of \( B \) implies that \( \langle L x - b_n \ | b^*_n - v^* \rangle \leq 0 \). Consequently, we derive from (3.2), (3.4), and (3.1) that

\[
\langle z - x_{n+1/2} | x_n - x_{n+1/2} \rangle / \theta_n \\
= \langle z | x_n - x_{n+1/2} \rangle / \theta_n + \langle x_{n+1/2} | x_n - x_{n+1/2} \rangle / \theta_n \\
= \langle x | x_n - x_{n+1/2} \rangle / \theta_n + \langle v^* | v^* - v^*_{n+1/2} \rangle / \theta_n \\
+ \langle x_{n+1/2} | x_n - x_{n+1/2} \rangle / \theta_n + \langle v^*_{n+1/2} | v^*_{n+1/2} \rangle / \theta_n \\
= \langle x | s^*_n \rangle + \langle t_n \ | v^* \rangle - \langle x_n | s^*_n \rangle - \langle t_n \ | v^*_n \rangle + \theta_n (\|s^*_n\|_2^2 + \|t_n\|_2^2) \\
= \langle x | s^*_n \rangle + \langle t_n \ | v^* \rangle - \langle x_n | s^*_n \rangle - \langle t_n \ | v^*_n \rangle + \lambda_n \langle x_n | s^*_n \rangle + \langle t_n \ | v^*_n \rangle - \langle a_n | a^*_n \rangle - \langle b_n \ | b^*_n \rangle \\
= \langle x | s^*_n \rangle + \langle t_n \ | v^* \rangle - \langle a_n \ | a^*_n \rangle - \langle b_n \ | b^*_n \rangle \\
-(1 - \lambda_n) \langle x_n | s^*_n \rangle + \langle t_n \ | v^*_n \rangle - \langle a_n \ | a^*_n \rangle - \langle b_n \ | b^*_n \rangle \\
= \langle x - a_n | a^*_n + L^* v^* \rangle + \langle L x - b_n \ | b^*_n - v^* \rangle \\
-(1 - \lambda_n) \langle x_n - a_n | a^*_n + L^* v^*_n \rangle + \langle L x_n - b_n \ | b^*_n - v^*_n \rangle \\
\leq \langle x - a_n | a^*_n + L^* v^* \rangle + \langle L x - b_n \ | b^*_n - v^* \rangle - \alpha (1 - \lambda_n) (\|a^*_n + L^* b^*_n\|_2^2 + \|L a_n - b_n\|_2^2) \\
\leq \langle x - a_n | a^*_n + L^* v^* \rangle + \langle L x - b_n \ | b^*_n - v^* \rangle \\
\leq 0.
\]
This verifies (3.5). It therefore follows from (2.8) that (3.2) is an instance of (2.2).

(i): It follows from (3.3) and Proposition 2.1(ii) that \( \sum_{n \in \mathbb{N}} \| x_{n+1} - x_n \|^2 + \sum_{n \in \mathbb{N}} \| v_{n+1}^* \|^2 = \sum_{n \in \mathbb{N}} \| x_{n+1} - x_n \|^2 < +\infty \).

(ii): Let \( n \in \mathbb{N} \). We consider two cases.

- \( \tau_n = 0 \): Then (3.2) yields \( \| s_n^* \|^2 + \| t_n \|^2 = 0 = \| x_{n+1/2} - x_n \|^2 / (\alpha \varepsilon)^2 \).
- \( \tau_n > 0 \): Then it follows from (3.1) and (3.2) that
  \[
  \| s_n^* \|^2 + \| t_n \|^2 = \tau_n \leq \frac{1}{\alpha^2 \varepsilon^2} \left( \frac{x_n - a_n \langle a_n^* + L^* v_n^* \rangle + \langle L x_n - b_n \mid b_n^* - v_n^* \rangle}{\alpha^2 \tau_n} \right) \leq \frac{\lambda_n^2 \| x_n \|^2 + \| t_n \|^2 - \| a_n \| \| a_n^* \| - \langle b_n \mid b_n^* \rangle}{\alpha^2 \varepsilon^2 \tau_n} \leq \frac{\theta^2 \tau_n}{\alpha^2 \varepsilon^2} \leq \frac{\| x_{n+1/2} - x_n \|^2 + \| v_{n+1/2}^* - v_n^* \|^2}{\alpha^2 \varepsilon^2} \leq \frac{\| x_{n+1/2} - x_n \|^2}{\alpha^2 \varepsilon^2}.
  \]

Altogether, it follows from Proposition 2.1(iii) that \( \sum_{n \in \mathbb{N}} \| s_n^* \|^2 + \sum_{n \in \mathbb{N}} \| t_n \|^2 < +\infty \).

(iii): Take \( x \in \mathcal{H} \), \( v^* \in \mathcal{G} \), and a strictly increasing sequence \( (k_n)_{n \in \mathbb{N}} \) in \( \mathbb{N} \), such that \( x_{k_n} \to x \) and \( v_{k_n}^* \to v^* \). We derive from (ii) and (3.2) that \( a_n^* + L^* b_n^* \to 0 \) and \( L a_n - b_n \to 0 \). Hence, the assumptions yield

\[
(3.9)
\]

On the other hand, (3.1) also asserts that \( \forall n \in \mathbb{N} \) \( (a_n, a_n^*) \in \text{gra} A \) and \( (b_n, b_n^*) \in \text{gra} B \). Altogether, Proposition 3.2 implies that \( (x, v^*) \in Z \). In view of Proposition 2.1(iv), the proof is complete. \( \square \)

Remark 3.4 Here are a few observations pertaining to Theorem 3.3.

(i) These results appear to provide the first algorithmic framework for composite inclusions problems that does not require additional assumptions on the constituents of the problem to achieve strong convergence.

(ii) If the second half of (3.2) is by-passed, i.e., if we set \( x_{n+1} = x_{n+1/2} \) and \( v_{n+1}^* = v_{n+1/2}^* \), and if the relaxation parameter \( \lambda_n \) is chosen in the range \([\varepsilon, 2 - \varepsilon]\), one recovers the algorithm of [2, Corollary 3.3]. However, this algorithm provides only weak convergence.
to an unspecified Kuhn-Tucker point, whereas (3.2) guarantees strong convergence to the best Kuhn-Tucker approximation to \((x_0, v_0^*)\). This can be viewed as another manifestation of the weak-to-strong convergence principle investigated in [6] in a different setting (\(\Sigma\)-class operators).

The following proposition is an application of Theorem 3.3 which describes a concrete implementation of (3.2) with a specific rule for selecting \((a_n, b_n, a^*_n, b^*_n) \in G_\alpha(x_n, v^*_n)\).

**Proposition 3.5** Consider the setting of Problem 1.1. Let \(\varepsilon \in ]0, 1[\) and iterate

\[
\begin{align*}
\gamma_n, \mu_n &\in [\varepsilon, 1/\varepsilon]^2 \\
 a_n &= J_{\gamma_n}A(x_n - \gamma_nL^*v^*_n) \\
l_n &= Lx_n \\
b_n &= J_{\mu_n}B(l_n + \mu_n v^*_n) \\
s_n = \gamma_n^{-1}(x_n - a_n) + \mu_n^{-1}L^*(l_n - b_n) \\
t_n &= b_n - L_a_n \\
\tau_n &= \|s_n\|^2 + \|t_n\|^2 \\
\text{if } \tau_n = 0 &\Rightarrow \theta_n = 0 \\
\text{if } \tau_n > 0 &\begin{cases} \\
\lambda_n \in [\varepsilon, 1] \\
\theta_n = \lambda_n(\gamma_n^{-1}\|x_n - a_n\|^2 + \mu_n^{-1}\|l_n - b_n\|^2)/\tau_n \\
x_{n+1/2} = x_n - \theta_ns_n \\
v_n^{*+1/2} = v_n^* - \theta_nt_n \\
\chi_n = (x_0 - x_n \mid x_n - x_{n+1/2}) + (v_0^* - v_n^* \mid v_n^* - v_{n+1/2}^*) \\
\mu_n = \|x_0 - x_n\|^2 + \|v_0^* - v_n^*\|^2 \\
\nu_n = \|x_n - x_{n+1/2}\|^2 + \|v_n^* - v_{n+1/2}^*\|^2 \\
\rho_n = \mu_n\nu_n - \chi_n^2 \\
\text{if } \rho_n = 0 \text{ and } \chi_n \geq 0 &\begin{cases} \\
x_{n+1} = x_{n+1/2} \\
v^*_n = v_{n+1/2} \\
\text{if } \rho_n > 0 \text{ and } \chi_n\nu_n \geq \rho_n &\begin{cases} \\
x_{n+1} = x_0 + (1 + \chi_n/\nu_n)(x_{n+1/2} - x_n) \\
v^*_n = v_0^* + (1 + \chi_n/\nu_n)(v_{n+1/2}^* - v_n^*) \\
\text{if } \rho_n > 0 \text{ and } \chi_n\nu_n < \rho_n &\begin{cases} \\
x_{n+1} = x_n + (\nu_n/\rho_n)(\chi_n(x_0 - x_n) + \mu_n(x_{n+1/2} - x_n)) \\
v^*_n = v_n^* + (\nu_n/\rho_n)(\chi_n(v_0^* - v_n^*) + \mu_n(v_{n+1/2}^* - v_n^*)). \\
\end{cases}
\end{cases}
\end{cases}
\end{align*}
\]

Then (3.10) generates infinite sequences \((x_n)_{n \in \mathbb{N}}\) and \((v^*_n)_{n \in \mathbb{N}}\), and the following hold:

- (i) \(\sum_{n \in \mathbb{N}} \|x_{n+1} - x_n\|^2 < +\infty\) and \(\sum_{n \in \mathbb{N}} \|v^*_{n+1} - v^*_n\|^2 < +\infty\).
- (ii) \(\sum_{n \in \mathbb{N}} \|s_n\|^2 < +\infty\) and \(\sum_{n \in \mathbb{N}} \|t_n\|^2 < +\infty\).
- (iii) \(\sum_{n \in \mathbb{N}} \|x_n - a_n\|^2 < +\infty\) and \(\sum_{n \in \mathbb{N}} \|Lx_n - b_n\|^2 < +\infty\).
(iv) $x_n \to \overline{x}$ and $v_n^* \to \overline{v}^*$.

Proof. Let us define

$$
\alpha = \frac{\varepsilon}{1 + \|L\|^2 + 2(1 - \varepsilon^2)\max\{1, \|L\|^2\}}
$$

and

$$(\forall n \in \mathbb{N}) \quad a_n^* = \gamma_n^{-1}(x_n - a_n) - L^* v_n^* \quad \text{and} \quad b_n^* = \mu_n^{-1}(L x_n - b_n) + v_n^*. \quad \text{(3.12)}$$

Then it is shown in [2, proof of Proposition 3.5] that

$$(\forall n \in \mathbb{N}) (a_n, b_n, a_n^*, b_n^*) \in G_{\alpha}(x_n, v_n^*) \quad \text{(3.13)}$$

and

$$(\forall n \in \mathbb{N}) \quad \|x_n - a_n\|^2 \leq 2\varepsilon^{-2}(\|s_n^*\|^2 + \varepsilon^{-2}\|L\| \|t_n\|^2). \quad \text{(3.14)}$$

We deduce from (3.12) and (3.13) that (3.10) is a special case of (3.2). Consequently, assertions (i) and (ii) follow from their counterparts in Theorem 3.3. To show (iii) it suffices to note that (3.14) and (ii) imply that

$$\sum_{n \in \mathbb{N}} \|x_n - a_n\|^2 < +\infty \quad \text{(3.15)}$$

and hence that $\sum_{n \in \mathbb{N}} \|L x_n - b_n\|^2 < +\infty$ since

$$\langle \forall n \in \mathbb{N} \rangle \quad \|L x_n - b_n\|^2 = \|L(x_n - a_n) + L a_n - b_n\|^2 \leq 2(\|L\|^2 \|x_n - a_n\|^2 + \|t_n\|^2). \quad \text{(3.16)}$$

In turn, (3.12) yields

$$\sum_{n \in \mathbb{N}} \|v_n^* - b_n^*\|^2 = \sum_{n \in \mathbb{N}} \mu_n^{-2} \|L x_n - b_n\|^2 \leq \varepsilon^{-2} \sum_{n \in \mathbb{N}} \|L x_n - b_n\|^2 < +\infty. \quad \text{(3.17)}$$

Altogether, (iv) follows from (3.15), (3.17), and Theorem 3.3(iii). \(\square\)

Remark 3.6 In (3.10), the identity $\tau_n = 0$ can be used as a stopping rule. Indeed, $\tau_n = 0 \iff (a_n^* + L^* b_n^*, b_n - L a_n) = (0, 0) \iff (-L^* b_n^*, L a_n) = (a_n^*, b_n) \in A a_n \times B^{-1} b_n \iff (a_n, b_n^*) \in \mathbb{Z}$. On the other hand, it follows from (3.14) and (3.16) that $\tau_n = 0 \Rightarrow (x_n, v_n^*) = (a_n, b_n^*)$. Altogether, Remark 2.2 yields $(x_n, v_n^*) = P_{\mathbb{Z}}(x_0, v_0^*) = (\overline{x}, \overline{v}^*)$.

Remark 3.7 An important feature of algorithm (3.10) which is inherited from that of [2, Proposition 3.5] is that it does not require the knowledge of $\|L\|$ or necessitate potentially hard to implement inversions of linear operators.
4 Application to systems of monotone inclusions

As discussed in \([2, 3, 5, 10, 12, 14, 20]\), various problems in applied mathematics can be modeled by systems of coupled monotone inclusions. In this section, we consider the following setting.

**Problem 4.1** Let \(m\) and \(K\) be strictly positive integers, let \((\mathcal{H}_i)_{1 \leq i \leq m}\) and \((\mathcal{G}_k)_{1 \leq k \leq K}\) be real Hilbert spaces, and set \(\mathcal{K} = \mathcal{H}_1 \oplus \cdots \oplus \mathcal{H}_m \oplus \mathcal{G}_1 \oplus \cdots \oplus \mathcal{G}_K\). For every \(i \in \{1, \ldots, m\}\) and every \(k \in \{1, \ldots, K\}\), let \(A_i : \mathcal{H}_i \to 2^{\mathcal{H}_i}\) and \(B_k : \mathcal{G}_k \to 2^{\mathcal{G}_k}\) be maximally monotone, let \(z_i \in \mathcal{H}_i\), let \(r_k \in \mathcal{G}_k\), and let \(L_{ki} : \mathcal{H}_i \to \mathcal{G}_k\) be linear and bounded. Let \((x_0, v_0^\ast) = (x_{1,0}, \ldots, x_{m,0}, v_{1,0}^\ast, \ldots, v_{K,0}^\ast) \in \mathcal{K}\), assume that the coupled inclusions problem

\[
\begin{align*}
\text{find } & \; x_1, \ldots, x_m \in \mathcal{H}_m \text{ such that } \\
& (\forall i \in \{1, \ldots, m\}) \quad z_i \in A_i x_i + \sum_{k=1}^K L_{ki}^* \left( B_k \left( \sum_{j=1}^m L_{kj}^* x_j - r_k \right) \right) 
\end{align*}
\]  

has at least one solution, and consider the dual problem

\[
\begin{align*}
\text{find } & \; v_1^\ast, \ldots, v_K^\ast \in \mathcal{G}_K \text{ such that } \\
& (\forall k \in \{1, \ldots, K\}) \quad -r_k \in -\sum_{i=1}^m L_{ki} \left( A_i^{-1} \left( z_i - \sum_{l=1}^K L_{il}^* v_l^\ast \right) \right) + B_k^{-1} v_k^\ast.
\end{align*}
\]  

The problem is to find the best approximation \((x_1, \ldots, x_m, v_1^\ast, \ldots, v_K^\ast)\) to \((x_0, v_0^\ast)\) from the associated Kuhn-Tucker set

\[
Z = \left\{ (x_1, \ldots, x_m, v_1^\ast, \ldots, v_K^\ast) \in \mathcal{K} \mid \right. \\
(\forall i \in \{1, \ldots, m\}) \quad z_i - \sum_{k=1}^K L_{ki}^* v_k^\ast \in A_i x_i \text{ and } \\
(\forall k \in \{1, \ldots, K\}) \quad \sum_{i=1}^m L_{ki} x_i - r_k \in B_k^{-1} v_k^\ast \left\}.
\]  

The next result presents a strongly convergent method for solving Problem 4.1. Let us note that existing methods require stringent additional conditions on the operators to achieve strong convergence, produce only unspecified points in the Kuhn-Tucker set, and necessitate the knowledge of the norms of the linear operators present in the model \([3, 14]\) or costly – sometimes unimplementable – linear inversions \([1]\). These shortcomings are simultaneously circumvented in the proposed algorithm. In \([4]\), these features are exploited to construct an implementable algorithm to solve domain decomposition methods in partial differential equations.
Proposition 4.2 Consider the setting of Problem 4.1. Let $\varepsilon \in ]0,1[$ and iterate

For $n = 0, 1, \ldots$

$$\gamma_n, \mu_n \in [\varepsilon, 1/\varepsilon]^2$$

For $i = 1, \ldots, m$

$$a_{i,n} = J_{\gamma_n} A_i (x_{i,n} + \gamma_n (z_i - \sum_{k=1}^{K} L_{k_i}^* v_{k,n}^*))$$

For $k = 1, \ldots, K$

$$l_{k,n} = \sum_{i=1}^{n} L_{k_i} x_{i,n}$$

$$b_{k,n} = r_k + J_{\mu_n B_k} (l_{k,n} + \mu_n v_{k,n}^* - r_k)$$

$$t_{k,n} = b_{k,n} - \sum_{i=1}^{m} L_{k_i} a_{i,n}$$

For $i = 1, \ldots, m$

$$s_{i,n}^* = \gamma_n^{-1} (x_{i,n} - a_{i,n}) + \mu_n^{-1} \sum_{k=1}^{K} L_{k_i}^* (l_{k,n} - b_{k,n})$$

$$\tau_n = \sum_{i=1}^{m} \|s_{i,n}^*\|^2 + \sum_{k=1}^{K} \|l_{k,n}\|^2$$

If $\tau_n = 0$

$$\theta_n = 0$$

If $\tau_n > 0$

$$\lambda_n \in [\varepsilon, 1]$$

$$\theta_n = \lambda_n (\gamma_n^{-1} \sum_{i=1}^{m} \|x_{i,n} - a_{i,n}\|^2 + \mu_n^{-1} \sum_{k=1}^{K} \|l_{k,n} - b_{k,n}\|^2) / \tau_n$$

For $i = 1, \ldots, m$

$$x_{i,n+1/2} = x_{i,n} - \theta_n s_{i,n}^*$$

For $k = 1, \ldots, K$

$$v_{k,n+1/2}^* = v_{k,n}^* - \theta_n t_{k,n}$$

$$\chi_n = \sum_{i=1}^{m} \langle x_{i,n} - x_{i,n} | x_{i,n} - x_{i,n+1/2} \rangle + \sum_{k=1}^{K} (v_{k,0}^* - v_{k,n}^* | v_{k,n}^* - v_{k,n+1/2}^*)$$

$$\mu_n = \sum_{i=1}^{m} \|x_{i,n} - x_{i,n}\|^2 + \sum_{k=1}^{K} \|v_{k,0}^* - v_{k,n}^*\|^2$$

$$\nu_n = \sum_{i=1}^{m} \|x_{i,n} - x_{i,n+1/2}\|^2 + \sum_{k=1}^{K} \|v_{k,n}^* - v_{k,n+1/2}\|^2$$

$$\rho_n = \mu_n \nu_n - \chi_n^2$$

If $\rho_n = 0$ and $\chi_n \geq 0$

For $i = 1, \ldots, m$

$$x_{i,n+1} = x_{i,n+1/2}$$

For $k = 1, \ldots, K$

$$v_{k,n+1}^* = v_{k,n+1/2}^*$$

If $\rho_n > 0$ and $\chi_n \nu_n \geq \rho_n$

For $i = 1, \ldots, m$

$$x_{i,n+1} = x_{i,n} + (1 + \chi_n / \nu_n) (x_{i,n+1/2} - x_{i,n})$$

For $k = 1, \ldots, K$

$$v_{k,n+1}^* = v_{k,0}^* + (1 + \chi_n / \nu_n) (v_{k,n+1/2}^* - v_{k,n}^*)$$

If $\rho_n > 0$ and $\chi_n \nu_n < \rho_n$

For $i = 1, \ldots, m$

$$x_{i,n+1} = x_{i,n} + (\nu_n / \rho_n) (\chi_n (x_{i,0} - x_{i,n}) + \mu_n (x_{i,n+1/2} - x_{i,n}))$$

For $k = 1, \ldots, K$

$$v_{k,n+1}^* = v_{k,n}^* + (\nu_n / \rho_n) (\chi_n (v_{k,0}^* - v_{k,n}^*) + \mu_n (v_{k,n+1/2}^* - v_{k,n}^*))$$

Then (4.4) generates infinite sequences $(x_{1,n})_{n \in \mathbb{N}}, \ldots, (x_{m,n})_{n \in \mathbb{N}}, (v_{1,n}^*)_{n \in \mathbb{N}}, \ldots, (v_{K,n}^*)_{n \in \mathbb{N}}$, and the
following hold:

(i) Let \( i \in \{1, \ldots, m\} \). Then \( \sum_{n \in \mathbb{N}} \| s_{i,n}^* \|^2 < +\infty \), \( \sum_{n \in \mathbb{N}} \| x_{i,n+1} - x_{i,n} \|^2 < +\infty \), \( \sum_{n \in \mathbb{N}} \| x_{i,n} - a_{i,n} \|^2 < +\infty \), and \( x_{i,n} \to \varpi_i \).

(ii) Let \( k \in \{1, \ldots, K\} \). Then \( \sum_{n \in \mathbb{N}} \| t_{k,n} \|^2 < +\infty \), \( \sum_{n \in \mathbb{N}} \| v_{k,n+1}^* - v_{k,n}^* \|^2 < +\infty \), \( \sum_{n \in \mathbb{N}} \| \sum_{i=1}^m L_{k,i} x_{i,n} - b_{k,n} \|^2 < +\infty \), and \( v_{k,n}^* \to \varpi_k^* \).

\[
\begin{align*}
\text{Proof.} \quad & \text{Let } \mathcal{H} = \bigoplus_{i=1}^m \mathcal{H}_i \text{ and } \mathcal{G} = \bigoplus_{k=1}^K \mathcal{G}_k, \text{ and let us introduce the operators} \\
& \begin{cases} \\
A: \mathcal{H} \to 2^\mathcal{H}: (x_i)_{1 \leq i \leq m} \mapsto \sum_{i=1}^m (-z_i + A_i x_i) \\
B: \mathcal{G} \to 2^\mathcal{G}: (y_k)_{1 \leq k \leq K} \mapsto \sum_{k=1}^K B_k (y_k - r_k) \\
L: \mathcal{H} \to \mathcal{G}: (x_i)_{1 \leq i \leq m} \mapsto \left( \sum_{i=1}^m L_{k,i} x_i \right)_{1 \leq k \leq K}.
\end{cases}
\end{align*}
\]

Then \( L^* : \mathcal{G} \to \mathcal{H} : (y_k)_{1 \leq k \leq K} \mapsto \left( \sum_{k=1}^K L_{k,i}^* y_k \right)_{1 \leq i \leq m} \) and, in this setting, Problem 1.1 becomes Problem 4.1. Next, for every \( n \in \mathbb{N} \), let us introduce the variables \( a_n = (a_{i,n})_{1 \leq i \leq m} \), \( s_n = (s_{i,n})_{1 \leq i \leq m} \), \( x_n = (x_{i,n})_{1 \leq i \leq m} \), \( x_{n+1/2} = (x_{i,n+1/2})_{1 \leq i \leq m} \), \( b_n = (b_{k,n})_{1 \leq k \leq K} \), \( l_n = (l_{k,n})_{1 \leq k \leq K} \), \( t_n = (t_{k,n})_{1 \leq k \leq K} \), \( v_n = (v_{k,n})_{1 \leq k \leq K} \), and \( v_{n+1/2} = (v_{k,n+1/2})_{1 \leq k \leq K} \). Since [7, Propositions 23.15 and 23.16] assert that

\[
(\forall n \in \mathbb{N})(\forall(x_i)_{1 \leq i \leq m} \in \mathcal{H})(\forall(y_k)_{1 \leq k \leq K} \in \mathcal{G}) \quad J_{\gamma_n A}(x_i)_{1 \leq i \leq m} = (J_{\gamma_n A_i}(x_i + \gamma_n z_i))_{1 \leq i \leq m}
\]

and 

\[
(4.6)
\]

(3.10) reduces in the present scenario to (4.4). Thus, the results follow from Proposition 3.5. \( \square \)

**Example 4.3** Let \( A, (B_k)_{1 \leq k \leq K} \), and \( (S_k)_{1 \leq k \leq K} \) be maximally monotone operators acting on a real Hilbert space \( \mathcal{H} \). We revisit a problem discussed in [14, Section 4], namely the relaxation of the possibly inconsistent inclusion problem

\[
\text{find } \varpi \in \mathcal{H} \text{ such that } 0 \in A \varpi \cap \bigcap_{k=1}^K B_k \varpi
\]

\[
\begin{align*}
\text{to}
\end{align*}
\]

\[
\text{find } \varpi \in \mathcal{H} \text{ such that } 0 \in A \varpi + \sum_{k=1}^K (B_k \square S_k) \varpi, \quad \text{where } B_k \square S_k = (B_k^{-1} + S_k^{-1})^{-1}
\]

We assume that (4.8) has at least one solution and that, for every \( k \in \{1, \ldots, K\} \), \( S_k^{-1} \) is at most single-valued and strictly monotone, with \( S_k^{-1} 0 = \{0\} \). Hence, (4.8) is a relaxation of (4.7) in the sense that if the latter happens to have solutions, they coincide with those of the former [14, Proposition 4.2]. As shown in [14], this framework captures many relaxation schemes, and a point \( \varpi_1 \in \mathcal{H} \) solves (4.8) if and only if \( (\varpi_1, \varpi_2, \ldots, \varpi_m) \) solves (4.1), where \( m = K + 1, \mathcal{H}_1 = \mathcal{H}, A_1 = A, z_1 = 0, \) and, for every \( k \in \{1, \ldots, K\}, \)

\[
\begin{align*}
\begin{cases}
\mathcal{H}_{k+1} = \mathcal{H} \\
\mathcal{G}_k = \mathcal{H} \\
A_{k+1} = S_k \\
z_{k+1} = 0 \\
r_k = 0
\end{cases} \quad \text{and} \quad \begin{cases}
L_{k1} = \text{Id} \\
(\forall i \in \{2, \ldots, m\}) \quad L_{ki} = \begin{cases}
-\text{Id}, & \text{if } i = k + 1; \\
0, & \text{otherwise}.
\end{cases}
\end{cases}
\end{align*}
\]
Thus (4.4) can be reduced to

for $n = 0, 1, \ldots$

$$\begin{align*}
(\gamma_n, \mu_n) &\in [\varepsilon, 1/\varepsilon]^2 \\
a_{1,n} &= J_{\gamma_n, A}(x_1, n - \gamma_n \sum_{k=1}^n v^*_k, n)
\end{align*}$$

for $k = 1, \ldots, K$

$$\begin{align*}
 a_{k+1,n} &= J_{\gamma_n, S_k}(x_{k+1,n} + \gamma_n v^*_k, n) \\
 l_{k,n} &= x_1, n - x_{k+1,n} \\
b_{k,n} &= J_{\mu_n, B_k}(l_{k,n} + \mu_n v^*_k, n) \\
t_{k,n} &= b_{k,n} + a_{k+1,n} - a_{1,n} \\
s^*_{k+1,n} &= \gamma_n^{-1}(x_{k+1,n} - a_{k+1,n}) + \mu_n^{-1}(b_{k,n} - l_{k,n}) \\
 s^*_1,n &= \gamma_n^{-1}(x_{1,n} - a_{1,n}) + \mu_n^{-1}\sum_{k=1}^K (l_{k,n} - b_{k,n}) \\
\tau_n &= \sum_{k=1}^{K+1} ||s^*_k,n||^2 + \sum_{k=1}^K ||t_{k,n}||^2 \\
\theta_n &= 0 \\
\theta_n &= 0 \\
\lambda_n &\in [\varepsilon, 1] \\
\theta_n &= \lambda_n \left( \gamma_n^{-1} \sum_{k=1}^{K+1} ||x_{k,n} - a_{k,n}||^2 + \mu_n^{-1} \sum_{k=1}^K ||l_{k,n} - b_{k,n}||^2 \right) / \tau_n \\
x_{1,n+1/2} &= x_{1,n} - \theta_n s^*_1,n
\end{align*}$$

for $k = 1, \ldots, K$

$$\begin{align*}
x_{k+1,n+1/2} &= x_{k+1,n} - \theta_n s^*_k,n \\
v^*_{k,n+1/2} &= v^*_k, n - \theta_n t_{k,n}
\end{align*}$$

$$\begin{align*}
\chi_n &= \sum_{k=1}^{K+1} \langle x_{k,0} - x_{k,n} | x_{k,n} - x_{k,n+1/2} \rangle + \sum_{k=1}^{K+1} \langle v^*_k, 0 - v^*_k, n | v^*_k, n - v^*_k, n+1/2 \rangle \\
\mu_n &= \sum_{k=1}^{K+1} ||x_{k,0} - x_{k,n}||^2 + \sum_{k=1}^{K+1} ||v^*_k, 0 - v^*_k, n||^2 \\
\nu_n &= \sum_{k=1}^{K+1} ||x_{k,n} - x_{k,n+1/2}||^2 + \sum_{k=1}^{K+1} ||v^*_k, n - v^*_k, n+1/2||^2 \\
\rho_n &= \mu_n \nu_n - \chi_n^2
\end{align*}$$

if $\rho_n = 0$ and $\chi_n \geq 0$

$$\begin{align*}
x_{1,n+1} &= x_{1,n+1/2} \\
\chi_{k,n+1} &= x_{k+1,n+1/2} \\
v^*_{k,n+1} &= v^*_k, n+1/2
\end{align*}$$

if $\rho_n > 0$ and $\chi_n \nu_n \geq \rho_n$

$$\begin{align*}
x_{1,n+1} &= x_{1,0} + (1 + \chi_n / \nu_n) (x_{1,n+1/2} - x_{1,n}) \\
x_{k+1,n+1} &= x_{k+1,0} + (1 + \chi_n / \nu_n) (x_{k+1,n+1/2} - x_{k+1,n}) \\
v^*_{k,n+1} &= v^*_k, 0 + (1 + \chi_n / \nu_n) (v^*_k, n+1/2 - v^*_k, n)
\end{align*}$$

if $\rho_n > 0$ and $\chi_n \nu_n < \rho_n$

$$\begin{align*}
x_{1,n+1} &= x_{1,n} + (\nu_n / \rho_n) (\chi_n (x_{1,0} - x_{1,n}) + \mu_n (x_{1,n+1/2} - x_{1,n})) \\
x_{k+1,n+1} &= x_{k+1,n} + (\nu_n / \rho_n) (\chi_n (x_{k+1,0} - x_{k+1,n}) + \mu_n (x_{k+1,n+1/2} - x_{k+1,n})) \\
v^*_{k,n+1} &= v^*_k, n + (\nu_n / \rho_n) (\chi_n (v^*_k, 0 - v^*_k, n) + \mu_n (v^*_k, n+1/2 - v^*_k, n))
\end{align*}$$

(4.10)

and it follows from Proposition 4.2 that $(x_{1,n})_{n \in \mathbb{N}}$ converges strongly to a solution $\pi_1$ to the relaxed problem (4.8). Let us note that the algorithm proposed in [14, Proposition 4.2] to solve
(4.8) requires that $A$ be uniformly monotone at $\mathfrak{p}$, to guarantee strong convergence, whereas this assumption is not needed here. In addition, the scaling parameters used in the resolvents of the monotone operators in [14, Proposition 4.2] must be identical at each iteration and bounded by a fixed constant: $(\forall n \in \mathbb{N}) \gamma_n = \mu_n \in [\varepsilon, (1 - \varepsilon)/\sqrt{K + 1}]$. By contrast, the parameters $\mu_n$ and $\gamma_n$ in (4.10) may differ and they can be arbitrarily large since $\varepsilon$ can be arbitrarily small, which could have some beneficial impact in terms of speed of convergence.

As a second illustration of Proposition 4.2, we consider the following multivariate minimization problem.

**Problem 4.4** Let $m$ and $K$ be strictly positive integers, let $(\mathcal{H}_i)_{1 \leq i \leq m}$ and $(\mathcal{G}_k)_{1 \leq k \leq K}$ be real Hilbert spaces, and set $\mathcal{K} = \mathcal{H}_1 \oplus \cdots \mathcal{H}_m \oplus \mathcal{G}_1 \oplus \cdots \oplus \mathcal{G}_K$. For every $i \in \{1, \ldots, m\}$ and every $k \in \{1, \ldots, K\}$, let $f_i \in \Gamma_0(\mathcal{H}_i)$ and $g_k \in \Gamma_0(\mathcal{G}_k)$, let $z_i \in \mathcal{H}_i$, let $r_k \in \mathcal{G}_k$, and let $L_{ki} : \mathcal{H}_i \to \mathcal{G}_k$ be linear and bounded. Let $(x_0, v_0^*) = (x_{1,0}, \ldots, x_{m,0}, v_{1,0}^*, \ldots, v_{K,0}^*) \in \mathcal{K}$ and assume that

\[
(\forall i \in \{1, \ldots, m\}) \quad z_i \in \text{ran} \left( \partial f_i + \sum_{k=1}^K L_{ki}^* \circ \partial g_k \circ \left( \sum_{j=1}^m L_{kj} \cdot r_k \right) \right). \tag{4.11}
\]

Consider the primal problem

\[
\text{minimize} \quad \sum_{i=1}^m (f_i(x_i) - \langle x_i \mid z_i \rangle) + \sum_{k=1}^K g_k \left( \sum_{i=1}^m L_{ki} x_i - r_k \right) \tag{4.12}
\]

and the dual problem

\[
\text{minimize} \quad \sum_{i=1}^m f_i^*(z_i - \sum_{k=1}^K L_{ki}^* v_k^*) + \sum_{k=1}^K \langle g_k^*(v_k^*) + (v_k^* \mid r_k) \rangle. \tag{4.13}
\]

The objective is to find the best approximation $(\bar{x}_1, \ldots, \bar{x}_m, \bar{v}_1^*, \ldots, \bar{v}_K^*)$ to $(x_0, v_0^*)$ from the associated Kuhn-Tucker set

\[
Z = \left\{ (x_1, \ldots, x_m, v_1^*, \ldots, v_K^*) \in \mathcal{K} \middle| (\forall i \in \{1, \ldots, m\}) \quad z_i - \sum_{k=1}^K L_{ki}^* v_k^* \in \partial f_i(x_i) \quad \text{and} \right. \\
\left. (\forall k \in \{1, \ldots, K\}) \quad \sum_{i=1}^m L_{ki} x_i - r_k \in \partial g_k^*(v_k^*) \right\}. \tag{4.14}
\]

The following corollary provides a strongly convergent method to solve Problem 4.4. Recall that the Moreau proximity operator [23] of a function $\varphi \in \Gamma_0(\mathcal{H})$ is $\text{prox}_\varphi = \partial_\varphi$, i.e., the operator which maps every point $x \in \mathcal{H}$ to the unique minimizer of the function $y \mapsto \varphi(y) + \|x - y\|^2/2$.

**Corollary 4.5** Consider the setting of Problem 4.4. Let $\varepsilon \in [0, 1]$ and execute (4.4), where $J_{\gamma_n A_i}$ is replaced by $\text{prox}_{\gamma_n f_i}$ and $J_{\mu_n B_k}$ is replaced by $\text{prox}_{\mu_n g_k}$. Then the following hold:

(i) $(\bar{x}_1, \ldots, \bar{x}_m)$ solves (4.12) and $(\bar{v}_1^*, \ldots, \bar{v}_m^*)$ solves (4.13).
(ii) For every $i \in \{1, \ldots, m\}$, $x_{i,n} \to x_i$.

(iii) For every $k \in \{1, \ldots, K\}$, $v^*_{k,n} \to v^*_k$.

**Proof.** Let us define $(\forall i \in \{1, \ldots, m\}) A_i = \partial f_i$ and $(\forall k \in \{1, \ldots, K\}) B_k = \partial g_k$. Then, as shown in the proof of [14, Proposition 5.4], (4.11) implies that Problem 4.1 assumes the form of Problem 4.4 and that Kuhn-Tucker points provide primal and dual solutions. Hence, applying Proposition 4.2 in this setting yields the claims. \(\square\)

**References**


