Journal of Nonlinear and Convex Analysis
Volume 15, Number 1, 2014, 137–159

Yokohama Publishers
ISSN 1880-5221 ONLINE JOURNAL

AN ALGORITHM FOR SPLITTING PARALLEL SUMS OF LINEARLY COMPOSED MONOTONE OPERATORS, WITH APPLICATIONS TO SIGNAL RECOVERY

STEPHEN R. BECKER AND PATRICK L. COMBETTES*

Abstract. We present a new primal-dual splitting algorithm for structured monotone inclusions in Hilbert spaces and analyze its asymptotic behavior. A novelty of our framework, which is motivated by image recovery applications, is to consider inclusions that combine a variety of monotonicity-preserving operations such as sums, linear compositions, parallel sums, and a new notion of parallel composition. The special case of minimization problems is studied in detail, and applications to signal recovery are discussed. An image restoration example is provided to illustrate the numerical implementation of the algorithm.

1. Introduction

Let A and B be set-valued monotone operators acting on a real Hilbert space $\mathcal{H}$. The first operator splitting algorithms were developed in the late 1970s to solve inclusion problems of the form

$$\text{find } \overline{x} \in \mathcal{H} \text{ such that } 0 \in A\overline{x} + B\overline{x}, \tag{1.1}$$

by using separate applications of the operators A and B at each iteration; see [9, 15, 17, 18, 23, 24] and the references therein. Because of increasingly complex problem formulations, more sophisticated splitting algorithms have recently arisen. Thus, the splitting method proposed in [6] can solve problems of the type

$$\text{find } \overline{x} \in \mathcal{H} \text{ such that } 0 \in A\overline{x} + \sum_{k=1}^{r} \big(L_k^* \circ B_k \circ L_k\big)\overline{x}, \tag{1.2}$$

where B_k is a monotone operator acting on a real Hilbert space $\mathcal{G}_k$ and L_k is a bounded linear operator from $\mathcal{H}$ to $\mathcal{G}_k$. This model was further refined in [12] by considering inclusions of the form

$$\text{find } \overline{x} \in \mathcal{H} \text{ such that } 0 \in A\overline{x} + \sum_{k=1}^{r} \big(L_k^* \circ (B_k \,\Box\, D_k) \circ L_k\big)\overline{x} + C\overline{x}, \tag{1.3}$$

where D_k is a monotone operator acting on $\mathcal{G}_k$ and such that D_k^{-1} is Lipschitzian,

$$B_k \,\Box\, D_k = \big(B_k^{-1} + D_k^{-1}\big)^{-1} \tag{1.4}$$

2010 *Mathematics Subject Classification.* Primary 47H05; Secondary 65K05, 90C25, 94A08.
Key words and phrases. Duality, image recovery, infimal convolution, monotone operator, parallel composition, parallel sum, proximity operator, splitting algorithm.
*Contact author.
S. R. Becker is funded by the Fondation Sciences Mathématiques de Paris. The research of P. L. Combettes is supported in part by the European Union under the 7th Framework Programme "FP7-PEOPLE-2010-ITN", grant agreement number 264735-SADCO..

is the parallel sum of B_k and D_k, and $C\colon \mathcal{H} \to \mathcal{H}$ is a Lipschitzian monotone operator. More recent developments concerning splitting methods for models featuring parallel sums can be found in [4, 5, 13, 25]. In the present paper, motivated by variational problems arising in image recovery, we consider a new type of inclusions that involve both parallel sum and "parallel composition" operations in the sense we introduce below.

Definition 1.1. Let $\mathcal{H}$ and $\mathcal{G}$ be real Hilbert spaces, let $A\colon \mathcal{H} \to 2^{\mathcal{H}}$, and let $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$. The *parallel composition* of A by L is

$$L \rhd A = \big(L \circ A^{-1} \circ L^*\big)^{-1}. \tag{1.5}$$

The primal-dual inclusion problem under consideration will be the following (our notation is standard, see Section 2.1 for details).

Problem 1.2. Let $\mathcal{H}$ be a real Hilbert space, let r be a strictly positive integer, let $z \in \mathcal{H}$, let $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ be maximally monotone, and let $C\colon \mathcal{H} \to \mathcal{H}$ be monotone and μ-Lipschitzian for some $\mu \in [0, +\infty[$. For every integer $k \in \{1, \ldots, r\}$, let $\mathcal{G}_k$ and $\mathcal{K}_k$ be real Hilbert spaces, let $B_k\colon \mathcal{G}_k \to 2^{\mathcal{G}_k}$ and $D_k\colon \mathcal{K}_k \to 2^{\mathcal{K}_k}$ be maximally monotone, and let $L_k \in \mathcal{B}(\mathcal{H}, \mathcal{G}_k)$ and $M_k \in \mathcal{B}(\mathcal{H}, \mathcal{K}_k)$. It is assumed that

$$\beta = \mu + \sqrt{\sum_{k=1}^{r} \|L_k\|^2 + \max_{1 \leqslant k \leqslant r} \big(\|L_k\|^2 + \|M_k\|^2\big)} > 0 \tag{1.6}$$

and that the inclusion

$$\text{find } \overline{x} \in \mathcal{H} \text{ such that } z \in A\overline{x} + \sum_{k=1}^{r} \big((L_k^* \circ B_k \circ L_k) \,\Box\, (M_k^* \circ D_k \circ M_k)\big)\overline{x} + C\overline{x} \tag{1.7}$$

possesses at least one solution. Solve (1.7) together with the dual problem

$$\text{find } \overline{v_1} \in \mathcal{G}_1, \ \ldots, \ \overline{v_r} \in \mathcal{G}_r \text{ such that } (\forall k \in \{1, \ldots, r\}) \tag{1.8}$$
$$0 \in -L_k\Bigg((A + C)^{-1}\bigg(z - \sum_{l=1}^{r} L_l^* \overline{v_l}\bigg)\Bigg) + B_k^{-1}\overline{v_k} + L_k\Big(\big(M_k^* \rhd D_k^{-1}\big)(L_k^* \overline{v_k})\Big).$$

The paper is organized as follows. In Section 2 we define our notation and provide preliminary results. In particular, we establish some basic properties of the parallel composition operation introduced in Definition 1.1 and discuss an algorithm recently proposed in [10] that will serve as a basis for our splitting method. In Section 3, our algorithm is presented and weak and strong convergence results are established. Section 4 is devoted to the application of this algorithm to convex minimization problems. Finally, in Section 5, we propose applications of the results of Section 4 to a concrete problem in image recovery, along with numerical results.

2. Notation and preliminary results

2.1. **Notation and definitions.** The following notation will be used throughout. $\mathcal{H}$, $\mathcal{G}$, and $\mathcal{K}$ are real Hilbert spaces. We denote the scalar product of a Hilbert space by $\langle \cdot \mid \cdot \rangle$ and the associated norm by $\|\cdot\|$. The symbols $\rightharpoonup$ and $\to$ denote, respectively, weak and strong convergence. $\mathcal{B}(\mathcal{H}, \mathcal{G})$ is the space of bounded linear operators from $\mathcal{H}$ to $\mathcal{G}$. The Hilbert direct sum of $\mathcal{H}$ and $\mathcal{G}$ is denoted by $\mathcal{H}\oplus\mathcal{G}$. Given two sequences $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ in $\mathcal{H}$, it will be convenient to use the notation

$$\big[(\forall n\in\mathbb{N})\;\; x_n\approx y_n\big]\quad\Leftrightarrow\quad\sum_{n\in\mathbb{N}}\|x_n-y_n\|<+\infty \tag{2.1}$$

to model the tolerance to errors in the implementation of the algorithms.

The power set of $\mathcal{H}$ is denoted by $2^{\mathcal{H}}$. Let $A\colon\mathcal{H}\to 2^{\mathcal{H}}$ be a set-valued operator. We denote by $\operatorname{ran}A=\big\{u\in\mathcal{H}\;\big|\;(\exists\, x\in\mathcal{H})\; u\in Ax\big\}$ the range of A, by $\operatorname{dom}A=\big\{x\in\mathcal{H}\;\big|\;Ax\neq\varnothing\big\}$ the domain of A, by $\operatorname{gra}A=\big\{(x,u)\in\mathcal{H}\times\mathcal{H}\;\big|\;u\in Ax\big\}$ the graph of A, and by A^{-1} the inverse of A, i.e., the set-valued operator with graph $\big\{(u,x)\in\mathcal{H}\times\mathcal{H}\;\big|\;u\in Ax\big\}$. The resolvent of A is $J_A=(\mathrm{Id}+A)^{-1}$. Moreover, A is monotone if

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

and maximally monotone if there exists no monotone operator $B\colon\mathcal{H}\to 2^{\mathcal{H}}$ such that $\operatorname{gra}A\subset\operatorname{gra}B\neq\operatorname{gra}A$. In this case, J_A is a single-valued, nonexpansive operator defined everywhere in $\mathcal{H}$. We say that A is uniformly monotone at $x\in\operatorname{dom}A$ if there exists an increasing function ϕ: $[0,+\infty[\;\to[0,+\infty]$ that vanishes only at 0 such that

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

We denote by $\Gamma_0(\mathcal{H})$ the class of lower semicontinuous convex functions $f\colon\mathcal{H}\to\,]-\infty,+\infty]$ such that $\operatorname{dom}f=\big\{x\in\mathcal{H}\;\big|\;f(x)<+\infty\big\}\neq\varnothing$. Let $f\in\Gamma_0(\mathcal{H})$. The conjugate of f is the function $f^*\in\Gamma_0(\mathcal{H})$ defined by $f^*\colon u\mapsto\sup_{x\in\mathcal{H}}(\langle x\mid u\rangle-f(x))$. For every $x\in\mathcal{H}$, $f+\|x-\cdot\|^2/2$ possesses a unique minimizer, which is denoted by $\operatorname{prox}_f x$. The operator prox_f can also be defined as a resolvent, namely

$$\operatorname{prox}_f=(\mathrm{Id}+\partial f)^{-1}=J_{\partial f}, \tag{2.4}$$

where $\partial f\colon\mathcal{H}\to 2^{\mathcal{H}}\colon x\mapsto\big\{u\in\mathcal{H}\;\big|\;(\forall y\in\mathcal{H})\;\langle y-x\mid u\rangle+f(x)\leqslant f(y)\big\}$ is the subdifferential of f, which is maximally monotone. We say that f uniformly convex at $x\in\operatorname{dom}f$ if there exists an increasing function ϕ: $[0,+\infty[\;\to[0,+\infty]$ that vanishes only at 0 such that

$$(\forall y\in\operatorname{dom}f)(\forall\alpha\in\,]0,1[)\quad f(\alpha x+(1-\alpha)y)+\alpha(1-\alpha)\phi(\|x-y\|)\leqslant\alpha f(x)+(1-\alpha)f(y). \tag{2.5}$$

The infimal convolution of two functions f_1 and f_2 from $\mathcal{H}$ to $]-\infty,+\infty]$ is

$$f_1\,\Box\, f_2\colon\mathcal{H}\to[-\infty,+\infty]:x\mapsto\inf_{y\in\mathcal{H}}\big(f_1(x-y)+f_2(y)\big), \tag{2.6}$$

and the infimal postcomposition of $f\colon\mathcal{H}\to[-\infty,+\infty]$ by $L\colon\mathcal{H}\to\mathcal{G}$ is

$$L\rhd f\colon\mathcal{G}\to[-\infty,+\infty]:y\mapsto\inf f\big(L^{-1}\{y\}\big)=\inf_{\substack{x\in\mathcal{H}\\ Lx=y}}f(x). \tag{2.7}$$

Let C be a convex subset of $\mathcal{H}$. The indicator function of C is denoted by ι_C, and the strong relative interior of C, i.e., the set of points $x\in C$ such that the cone generated by $-x+C$ is a closed vector subspace of $\mathcal{H}$, by $\operatorname{sri}C$.

For a detailed account of the above concepts, see [2].

2.2. **Parallel composition.** In this section we explore some basic properties of the parallel composition operation introduced in Definition 1.1 which are of interest in their own right. First, we justify the terminology via the following connection with the parallel sum.

Lemma 2.1. *Let* $A\colon \mathcal{H} \to 2^{\mathcal{H}}$, *let* $B\colon \mathcal{H} \to 2^{\mathcal{H}}$, *and let* $L\colon \mathcal{H} \oplus \mathcal{H} \to \mathcal{H}\colon (x,y) \mapsto x+y$. *Then* $L \triangleright (A \times B) = A \,\Box\, B$.

Proof. Since $L^*\colon \mathcal{H} \to \mathcal{H} \oplus \mathcal{H}\colon x \mapsto (x,x)$, the announced identity is an immediate consequence of (1.4) and (1.5). $\square$

Lemma 2.2. *Let* $A\colon \mathcal{H} \to 2^{\mathcal{H}}$, *let* $B\colon \mathcal{G} \to 2^{\mathcal{G}}$, *and let* $L \in \mathcal{B}(\mathcal{H},\mathcal{G})$. *Then the following hold.*

(i) $((L \triangleright A) \,\Box\, B)^{-1} = L \circ A^{-1} \circ L^* + B^{-1}$.

(ii) *Suppose that* A *and* B *are monotone. Then* $(L \triangleright A) \,\Box\, B$ *is monotone.*

(iii) *Suppose that* A *and* B *are maximally monotone and that the cone generated by* $L^*(\operatorname{ran} B) - \operatorname{ran} A$ *is a closed vector subspace. Then* $(L \triangleright A) \,\Box\, B$ *is maximally monotone.*

(iv) *Suppose that* A *is maximally monotone and that the cone generated by* $\operatorname{ran} L^* + \operatorname{ran} A$ *is a closed vector subspace. Then* $L \triangleright A$ *is maximally monotone.*

Proof. (i): This follows easily from (1.4) and (1.5).

(ii): By (i), $((L \triangleright A) \,\Box\, B)^{-1} = L \circ A^{-1} \circ L^* + B^{-1}$. Since A^{-1} and B^{-1} are monotone and monotonicity is preserved under inversion and this type of transformation [2, Proposition 20.10], the assertion is proved.

(iii): The operators A^{-1} and B^{-1} are maximally monotone [2, Proposition 20.22] and $L^*(\operatorname{ran} B) - \operatorname{ran} A = L^*(\operatorname{dom} B^{-1}) - \operatorname{dom} A^{-1}$. Hence, $L \circ A^{-1} \circ L^* + B^{-1}$ is maximally monotone [3, Section 24] and so is its inverse which, in view of (i), is $(L \triangleright A) \,\Box\, B$. This result can also be derived from [22, Theorems 16 and 21].

(iv): Set $B = \{0\}^{-1}$ in (iii). $\square$

Lemma 2.3. *Let* $A\colon \mathcal{H} \to 2^{\mathcal{H}}$, *let* $B\colon \mathcal{H} \to 2^{\mathcal{H}}$, *and let* $L \in \mathcal{B}(\mathcal{H},\mathcal{G})$. *Then* $L \triangleright (A \,\Box\, B) = (L \triangleright A) \,\Box\, (L \triangleright B)$.

Proof. It follows from (1.4) and (1.5) that

$$\begin{aligned} L \triangleright (A \,\Box\, B) &= \big(L \circ (A \,\Box\, B)^{-1} \circ L^*\big)^{-1} \\ &= \big(L \circ (A^{-1} + B^{-1}) \circ L^*\big)^{-1} \\ &= \big(L \circ A^{-1} \circ L^* + L \circ B^{-1} \circ L^*\big)^{-1} \\ &= \big((L \triangleright A)^{-1} + (L \triangleright B)^{-1}\big)^{-1} \\ &= (L \triangleright A) \,\Box\, (L \triangleright B), \end{aligned} \tag{2.8}$$

which proves the announced identity. $\square$

Lemma 2.4. *Let* $A\colon \mathcal{H} \to 2^{\mathcal{H}}$, *let* $L \in \mathcal{B}(\mathcal{H},\mathcal{G})$, *and let* $M \in \mathcal{B}(\mathcal{G},\mathcal{K})$. *Then* $M \triangleright (L \triangleright A) = (M \circ L) \triangleright A$.

Proof. Indeed, $M \triangleright (L \triangleright A) = (M \circ (L \triangleright A)^{-1} \circ M^*)^{-1} = (M \circ L \circ A^{-1} \circ L^* \circ M^*)^{-1} = (M \circ L) \triangleright A$. $\square$

Finally, in the next lemma we draw connections with the infimal convolution and postcomposition operations of (2.6) and (2.7).

Lemma 2.5. *Let $f \in \Gamma_0(\mathcal{H})$, let $g \in \Gamma_0(\mathcal{G})$, and let $L \in \mathcal{B}(\mathcal{H},\mathcal{G})$ be such that $0 \in \operatorname{sri}(L^*(\operatorname{dom} g^*) - \operatorname{dom} f^*)$. Then the following hold.*

(i) $(L \rhd f) \,\Box\, g \in \Gamma_0(\mathcal{G})$.

(ii) $\partial\big((L \rhd f) \,\Box\, g\big) = (L \rhd \partial f) \,\Box\, \partial g$.

Proof. (i): Since $0 \in L^*(\operatorname{dom} g^*) - \operatorname{dom} f^*$ and, by the Fenchel-Moreau theorem [2, Theorem 13.32], $f^* \in \Gamma_0(\mathcal{H})$ and $g^* \in \Gamma_0(\mathcal{G})$, we have $f^* \circ L^* + g^* \in \Gamma_0(\mathcal{G})$. Hence $(f^* \circ L^* + g^*)^* \in \Gamma_0(\mathcal{G})$. However, in view of [2, Theorem 15.27(i)], the assumptions also imply that $(f^* \circ L^* + g^*)^* = (L \rhd f) \,\Box\, g$.

(ii): Let y and v be in $\mathcal{G}$. Then (i), [2, Corollary 16.24, Proposition 13.21(i)&(iv), and Theorem 16.37(i)] enable us to write

$$
\begin{aligned}
v \in \partial\big((L \rhd f) \,\Box\, g\big)(y) \quad &\Leftrightarrow \quad y \in \Big(\partial\big((L \rhd f) \,\Box\, g\big)\Big)^{-1}(v) \\
&\Leftrightarrow \quad y \in \partial\big((L \rhd f) \,\Box\, g\big)^*(v) \\
&\Leftrightarrow \quad y \in \partial(f^* \circ L^* + g^*)(v) \\
&\Leftrightarrow \quad y \in (L \circ (\partial f^*) \circ L^* + \partial g^*)(v) \\
&\Leftrightarrow \quad y \in \big(L \circ (\partial f)^{-1} \circ L^* + (\partial g)^{-1}\big)(v) \\
&\Leftrightarrow \quad v \in \big((L \rhd \partial f) \,\Box\, \partial g\big)y, \qquad (2.9)
\end{aligned}
$$

which establishes the announced identity. ☐

Corollary 2.6. *Let $f \in \Gamma_0(\mathcal{H})$ and let $L \in \mathcal{B}(\mathcal{H},\mathcal{G})$ be such that $0 \in \operatorname{sri}(\operatorname{ran} L^* - \operatorname{dom} f^*)$. Then the following hold.*

(i) $L \rhd f \in \Gamma_0(\mathcal{G})$.

(ii) $\partial(L \rhd f) = L \rhd \partial f$.

Proof. Set $g = \iota_{\{0\}}$ in Lemma 2.5. ☐

2.3. **An inclusion problem.** Our main result in Section 3 will hinge on rewriting Problem 1.2 as an instance of the following formulation.

Problem 2.7. Let m and K be strictly positive integers, let $(\mathcal{H}_i)_{1\leqslant i\leqslant m}$ and $(\mathcal{G}_k)_{1\leqslant k\leqslant K}$ be real Hilbert spaces, and let $(\mu_i)_{1\leqslant i\leqslant m} \in [0,+\infty[^m$. For every $i \in \{1,\ldots,m\}$ and $k \in \{1,\ldots,K\}$, let $C_i \colon \mathcal{H}_i \to \mathcal{H}_i$ be monotone and μ_i-Lipschitzian, let $A_i \colon \mathcal{H}_i \to 2^{\mathcal{H}_i}$ and $B_k \colon \mathcal{G}_k \to 2^{\mathcal{G}_k}$ be maximally monotone, let $z_i \in \mathcal{H}_i$, and let $L_{ki} \in \mathcal{B}(\mathcal{H}_i,\mathcal{G}_k)$. It is assumed that

(2.10)
$$\beta = \sqrt{\lambda} + \max_{1\leqslant i\leqslant m} \mu_i > 0, \quad \text{where} \quad \lambda \in \left[\sup_{\sum_{i=1}^m \|x_i\|^2 \leqslant 1} \sum_{k=1}^K \left\| \sum_{i=1}^m L_{ki} x_i \right\|^2, +\infty\right[,$$

and that the system of coupled inclusions

(2.11) find $\overline{x_1} \in \mathcal{H}_1, \ldots, \overline{x_m} \in \mathcal{H}_m$ such that

$$\begin{cases} z_1 & \in A_1\overline{x_1} + \displaystyle\sum_{k=1}^{K} L_{k1}^*\Big(B_k\Big(\sum_{i=1}^{m} L_{ki}\overline{x_i}\Big)\Big) + C_1\overline{x_1} \\ \ \vdots & \\ z_m & \in A_m\overline{x_m} + \displaystyle\sum_{k=1}^{K} L_{km}^*\Big(B_k\Big(\sum_{i=1}^{m} L_{ki}\overline{x_i}\Big)\Big) + C_m\overline{x_m} \end{cases}$$

possesses at least one solution. Solve (2.11) together with the dual problem

(2.12) find $\overline{v_1} \in \mathcal{G}_1, \ldots, \overline{v_K} \in \mathcal{G}_K$ such that

$$\begin{cases} 0 \in -\displaystyle\sum_{i=1}^{m} L_{1i}\big(A_i + C_i\big)^{-1}\Big(z_i - \sum_{k=1}^{K} L_{ki}^*\overline{v_k}\Big) + B_1^{-1}\overline{v_1} \\ \ \vdots \\ 0 \in -\displaystyle\sum_{i=1}^{m} L_{Ki}\big(A_i + C_i\big)^{-1}\Big(z_i - \sum_{k=1}^{K} L_{ki}^*\overline{v_k}\Big) + B_K^{-1}\overline{v_K}. \end{cases}$$

The following result is a special case of [10, Theorem 2.4(iii)]. We use the notation (2.1) to model the possibility of inexactly evaluating the operators involved.

Theorem 2.8. *Consider the setting of Problem 2.7. Let $x_{1,0} \in \mathcal{H}_1, \ldots, x_{m,0} \in \mathcal{H}_m$, $v_{1,0} \in \mathcal{G}_1, \ldots, v_{K,0} \in \mathcal{G}_K$, let $\varepsilon \in \left]0, 1/(\beta+1)\right[$, let $(\gamma_n)_{n\in\mathbb{N}}$ be a sequence in $[\varepsilon, (1-\varepsilon)/\beta]$, and set*

$$\begin{array}{l} \text{for } n = 0, 1, \ldots \\ \left\lfloor \begin{array}{l} \text{for } i = 1, \ldots, m \\ \left\lfloor \begin{array}{l} s_{1,i,n} \approx x_{i,n} - \gamma_n\big(C_i x_{i,n} + \sum_{k=1}^{K} L_{ki}^* v_{k,n}\big) \\ p_{1,i,n} \approx J_{\gamma_n A_i}(s_{1,i,n} + \gamma_n z_i) \end{array} \right. \\ \text{for } k = 1, \ldots, K \\ \left\lfloor \begin{array}{l} s_{2,k,n} \approx v_{k,n} + \gamma_n \sum_{i=1}^{m} L_{ki} x_{i,n} \\ p_{2,k,n} \approx s_{2,k,n} - \gamma_n J_{\gamma_n^{-1} B_k}(\gamma_n^{-1} s_{2,k,n}) \\ q_{2,k,n} \approx p_{2,k,n} + \gamma_n \sum_{i=1}^{m} L_{ki} p_{1,i,n} \\ v_{k,n+1} = v_{k,n} - s_{2,k,n} + q_{2,k,n} \end{array} \right. \\ \text{for } i = 1, \ldots, m \\ \left\lfloor \begin{array}{l} q_{1,i,n} \approx p_{1,i,n} - \gamma_n\big(C_i p_{1,i,n} + \sum_{k=1}^{K} L_{ki}^* p_{2,k,n}\big) \\ x_{i,n+1} = x_{i,n} - s_{1,i,n} + q_{1,i,n}. \end{array} \right. \end{array} \right. \end{array} \tag{2.13}$$

Then there exist a solution $(\overline{x_1}, \ldots, \overline{x_m})$ to (2.11) and a solution $(\overline{v_1}, \ldots, \overline{v_K})$ to (2.12) such that the following hold.

(i) $(\forall i \in \{1, \ldots, m\})$ $z_i - \sum_{k=1}^{K} L_{ki}^*\overline{v_k} \in A_i\overline{x_i} + C_i\overline{x_i}$.

(ii) $(\forall k \in \{1, \ldots, K\})$ $\sum_{i=1}^{m} L_{ki}\overline{x_i} \in B_k^{-1}\overline{v_k}$.

(iii) $(\forall i \in \{1, \ldots, m\})$ $x_{i,n} \rightharpoonup \overline{x_i}$.

(iv) $(\forall k \in \{1, \ldots, K\})$ $v_{k,n} \rightharpoonup \overline{v_k}$.

(v) *Suppose that A_1 or C_1 is uniformly monotone at $\overline{x_1}$. Then $x_{1,n} \to \overline{x_1}$.*

(vi) *Suppose that, for some $k \in \{1, \ldots, K\}$, B_k^{-1} is uniformly monotone at $\overline{v_k}$. Then $v_{k,n} \to \overline{v_k}$.*

3. Main algorithm

We start with the following facts.

Proposition 3.1. *Let $\mathcal{H}$ be a real Hilbert space, let r be a strictly positive integer, let $z \in \mathcal{H}$, let $A\colon \mathcal{H} \to 2^{\mathcal{H}}$, and let $C\colon \mathcal{H} \to \mathcal{H}$. For every integer $k \in \{1,\ldots,r\}$, let $\mathcal{G}_k$ and $\mathcal{K}_k$ be real Hilbert spaces, let $B_k\colon \mathcal{G}_k \to 2^{\mathcal{G}_k}$ and $D_k\colon \mathcal{K}_k \to 2^{\mathcal{K}_k}$, and let $L_k \in \mathcal{B}(\mathcal{H},\mathcal{G}_k)$ and $M_k \in \mathcal{B}(\mathcal{H},\mathcal{K}_k)$. Set*

$$\boldsymbol{\mathcal{H}} = \bigoplus_{k=1}^{r} \mathcal{H}, \quad \boldsymbol{\mathcal{G}} = \bigoplus_{k=1}^{r} \mathcal{G}_k, \quad \boldsymbol{\mathcal{K}} = \bigoplus_{k=1}^{r} \mathcal{K}_k, \tag{3.1}$$

and

$$\begin{cases} \boldsymbol{A}\colon \boldsymbol{\mathcal{H}} \to 2^{\boldsymbol{\mathcal{H}}}\colon (x,y_1,\ldots,y_r) \mapsto (Ax + Cx - z) \times \{0\} \times \cdots \times \{0\} \\ \boldsymbol{B}\colon \boldsymbol{\mathcal{G}} \oplus \boldsymbol{\mathcal{K}} \to 2^{\boldsymbol{\mathcal{G}} \oplus \boldsymbol{\mathcal{K}}}\colon (s_1,\ldots,s_r,t_1,\ldots,t_r) \mapsto B_1 s_1 \times \cdots \times B_r s_r \times D_1 t_1 \times \cdots \times D_r t_r \\ \boldsymbol{L}\colon \boldsymbol{\mathcal{H}} \to \boldsymbol{\mathcal{G}} \oplus \boldsymbol{\mathcal{K}}\colon (x,y_1,\ldots,y_r) \mapsto (L_1 x - L_1 y_1,\ldots,L_r x - L_r y_r, M_1 y_1,\ldots,M_r y_r). \end{cases} \tag{3.2}$$

Furthermore, suppose that

$$\big(\exists\, \overline{\boldsymbol{x}} = (\overline{x},\overline{y_1},\ldots,\overline{y_r}) \in \boldsymbol{\mathcal{H}}\big) \quad \boldsymbol{0} \in \boldsymbol{A}\overline{\boldsymbol{x}} + \boldsymbol{L}^*\big(\boldsymbol{B}(\boldsymbol{L}\overline{\boldsymbol{x}})\big). \tag{3.3}$$

Then the following hold for some $(\overline{v_1},\ldots,\overline{v_r}) \in \boldsymbol{\mathcal{G}}$ and $(\overline{w_1},\ldots,\overline{w_r}) \in \boldsymbol{\mathcal{K}}$.

(i) *$z - \sum_{k=1}^r L_k^* \overline{v_k} \in A\overline{x} + C\overline{x}$.*
(ii) *$(\forall k \in \{1,\ldots,r\})$ $L_k^*\overline{v_k} = M_k^*\overline{w_k}$, $L_k\overline{x} - L_k\overline{y}_k \in B_k^{-1}\overline{v_k}$, and $M_k\overline{y}_k \in D_k^{-1}\overline{w_k}$.*
(iii) *$\overline{x}$ solves (1.7).*
(iv) *$(\overline{v_1},\ldots,\overline{v_r})$ solves (1.8).*

Proof. (i) and (ii): It follows from (3.3) that there exists $\overline{\boldsymbol{v}} = (\overline{v_1},\ldots,\overline{v_r},\overline{w_1},\ldots,\overline{w_r}) \in \boldsymbol{\mathcal{G}} \oplus \boldsymbol{\mathcal{K}}$ such that $-\boldsymbol{L}^*\overline{\boldsymbol{v}} \in \boldsymbol{A}\overline{\boldsymbol{x}}$ and $\overline{\boldsymbol{v}} \in \boldsymbol{B}(\boldsymbol{L}\overline{\boldsymbol{x}})$, i.e.,

$$-\boldsymbol{L}^*\overline{\boldsymbol{v}} \in \boldsymbol{A}\overline{\boldsymbol{x}} \quad \text{and} \quad \boldsymbol{L}\overline{\boldsymbol{x}} \in \boldsymbol{B}^{-1}\overline{\boldsymbol{v}}. \tag{3.4}$$

Since

$$\boldsymbol{L}^*\colon \boldsymbol{\mathcal{G}} \oplus \boldsymbol{\mathcal{K}} \to \boldsymbol{\mathcal{H}}\colon (v_1,\ldots,v_r,w_1,\ldots,w_r) \mapsto \left(\sum_{k=1}^r L_k^* v_k, M_1^* w_1 - L_1^* v_1, \ldots, M_r^* w_r - L_r^* v_r\right), \tag{3.5}$$

it follows from (3.2) that (3.4) can be rewritten as

$$z - \sum_{k=1}^r L_k^*\overline{v_k} \in A\overline{x} + C\overline{x} \quad \text{and} \quad (\forall k \in \{1,\ldots,r\}) \quad \begin{cases} L_k^*\overline{v_k} = M_k^*\overline{w_k} \\ L_k\overline{x} - L_k\overline{y}_k \in B_k^{-1}\overline{v_k} \\ M_k\overline{y}_k \in D_k^{-1}\overline{w_k}. \end{cases} \tag{3.6}$$

(iii): For every $k \in \{1,\ldots,r\}$,

$$\text{(ii)} \Rightarrow \begin{cases} L_k^*\overline{v_k} = M_k^*\overline{w_k} \\ \overline{v_k} \in B_k(L_k\overline{x} - L_k\overline{y}_k) \\ \overline{w_k} \in D_k(M_k\overline{y}_k) \end{cases} \tag{3.7}$$

$$
\begin{aligned}
&\Rightarrow \begin{cases} L_k^*\overline{v_k} = M_k^*\overline{w_k} \\ L_k^*\overline{v_k} \in L_k^*\big(B_k(L_k\overline{x} - L_k\overline{y}_k)\big) \\ M_k^*\overline{w_k} \in M_k^*\big(D_k(M_k\overline{y}_k)\big) \end{cases} \\
&\Leftrightarrow \begin{cases} L_k^*\overline{v_k} = M_k^*\overline{w_k} \\ \overline{x} - \overline{y}_k \in (L_k^* \circ B_k \circ L_k)^{-1}(L_k^*\overline{v_k}) \\ \overline{y}_k \in (M_k^* \circ D_k \circ M_k)^{-1}(M_k^*\overline{w_k}) \end{cases} \tag{3.8} \\
&\Rightarrow \overline{x} \in (L_k^* \circ B_k \circ L_k)^{-1}(L_k^*\overline{v_k}) + (M_k^* \circ D_k \circ M_k)^{-1}(L_k^*\overline{v_k}) \\
&\Leftrightarrow L_k^*\overline{v_k} \in \big((L_k^* \circ B_k \circ L_k) \,\square\, (M_k^* \circ D_k \circ M_k)\big)(\overline{x}). \tag{3.9}
\end{aligned}
$$

Hence,

$$
\sum_{k=1}^{r} L_k^*\overline{v_k} \in \sum_{k=1}^{r} \big((L_k^* \circ B_k \circ L_k) \,\square\, (M_k^* \circ D_k \circ M_k)\big)(\overline{x}). \tag{3.10}
$$

Adding this inclusion to that of (i) shows that $\overline{x}$ solves (1.7).

(iv): It follows from (i) that

$$
(\forall k \in \{1, \ldots, r\}) \quad -L_k\overline{x} \in -L_k\left((A+C)^{-1}\left(z - \sum_{l=1}^{r} L_l^*\overline{v_l}\right)\right). \tag{3.11}
$$

On the other hand, (ii) yields

$$
(\forall k \in \{1, \ldots, r\}) \quad L_k\overline{x} - L_k\overline{y_k} \in B_k^{-1}\overline{v_k}, \tag{3.12}
$$

while (3.8) yields

$$
\begin{aligned}
(\forall k \in \{1, \ldots, r\}) \quad L_k\overline{y_k} &\in L_k\big((M_k^* \circ D_k \circ M_k)^{-1}(M_k^*\overline{w_k})\big) \\
&= L_k\big((M_k^* \circ D_k \circ M_k)^{-1}(L_k^*\overline{v_k})\big) \\
&= L_k\big((M_k^* \triangleright D_k^{-1})(L_k^*\overline{v_k})\big). \tag{3.13}
\end{aligned}
$$

Upon adding (3.11), (3.12), and (3.13), we obtain

$$
\begin{aligned}
(\forall k \in \{1, \ldots, r\}) \quad 0 \in -L_k&\left((A+C)^{-1}\left(z - \sum_{l=1}^{r} L_l^*\overline{v_l}\right)\right) \\
&+ B_k^{-1}\overline{v_k} + L_k\big((M_k^* \triangleright D_k^{-1})(L_k^*\overline{v_k})\big), \tag{3.14}
\end{aligned}
$$

which proves that $(\overline{v_1}, \ldots, \overline{v_r})$ solves (1.8). $\square$

We are now in a position to present our main result.

Theorem 3.2. *Consider the setting of Problem 1.2. Let* $x_0 \in \mathcal{H}$, $y_{1,0} \in \mathcal{H}$, $\ldots$, $y_{r,0} \in \mathcal{H}$, $v_{1,0} \in \mathcal{G}_1$, $\ldots$, $v_{r,0} \in \mathcal{G}_r$, $w_{1,0} \in \mathcal{K}_1$, $\ldots$, $w_{r,0} \in \mathcal{K}_r$, *let* $\varepsilon \in \left]0, 1/(\beta+1)\right[$, *let* $(\gamma_n)_{n\in\mathbb{N}}$ *be a sequence in* $[\varepsilon,(1-\varepsilon)/\beta]$, *and set*

$$
\begin{array}{l}
\text{for } n=0,1,\ldots\\
\left\lfloor
\begin{array}{l}
s_{1,1,n}\approx x_n-\gamma_n(Cx_n+\sum_{k=1}^r L_k^* v_{k,n})\\
p_{1,1,n}\approx J_{\gamma_n A}(s_{1,1,n}+\gamma_n z)\\
\text{for } k=1,\ldots,r\\
\left\lfloor
\begin{array}{l}
p_{1,k+1,n}\approx y_{k,n}+\gamma_n(L_k^* v_{k,n}-M_k^* w_{k,n})\\
s_{2,k,n}\approx v_{k,n}+\gamma_n L_k(x_n-y_{k,n})\\
p_{2,k,n}\approx s_{2,k,n}-\gamma_n J_{\gamma_n^{-1}B_k}(\gamma_n^{-1}s_{2,k,n})\\
q_{2,k,n}\approx p_{2,k,n}+\gamma_n L_k(p_{1,1,n}-p_{1,k+1,n})\\
v_{k,n+1}=v_{k,n}-s_{2,k,n}+q_{2,k,n}\\
s_{2,k+r,n}\approx w_{k,n}+\gamma_n M_k y_{k,n}\\
p_{2,k+r,n}\approx s_{2,k+r,n}-\gamma_n\big(J_{\gamma_n^{-1}D_k}(\gamma_n^{-1}s_{2,k+r,n})\big)\\
q_{1,k+1,n}\approx p_{1,k+1,n}+\gamma_n(L_k^* p_{2,k,n}-M_k^* p_{2,k+r,n})\\
q_{2,k+r,n}\approx p_{2,k+r,n}+\gamma_n M_k p_{1,k+1,n}\\
w_{k,n+1}=w_{k,n}-s_{2,k+r,n}+q_{2,k+r,n}
\end{array}
\right.\\
q_{1,1,n}\approx p_{1,1,n}-\gamma_n(Cp_{1,1,n}+\sum_{k=1}^r L_k^* p_{2,k,n})\\
x_{n+1}=x_n-s_{1,1,n}+q_{1,1,n}\\
\text{for } k=1,\ldots,r\\
\left\lfloor\; y_{k,n+1}=y_{k,n}-p_{1,k+1,n}+q_{1,k+1,n}.\right.
\end{array}
\right.
\end{array}
\tag{3.15}
$$

Then the following hold for some solution $\overline{x}$ *to* (1.7) *and some solution* $(\overline{v_1},\ldots,\overline{v_r})$ *to* (1.8).

(i) $x_n \rightharpoonup \overline{x}$ *and* $(\forall k\in\{1,\ldots,r\})$ $v_{k,n}\rightharpoonup\overline{v_k}$.
(ii) *Suppose that* A *or* C *is uniformly monotone at* $\overline{x}$. *Then* $x_n\to\overline{x}$.
(iii) *Suppose that, for some* $k\in\{1,\ldots,r\}$, B_k^{-1} *is uniformly monotone at* $\overline{v_k}$. *Then* $v_{k,n}\to\overline{v_k}$.

Proof. We introduce the auxiliary problem

$$
\text{find } \overline{x}\in\mathcal{H},\ \overline{y_1}\in\mathcal{H},\ldots,\overline{y_r}\in\mathcal{H} \text{ such that }
\begin{cases}
z & \in A\overline{x}+\sum_{k=1}^r L_k^*\big(B_k(L_k\overline{x}-L_k\overline{y_k})\big)+C\overline{x}\\
0 & \in -L_1^*\big(B_1(L_1\overline{x}-L_1\overline{y_1})\big)+M_1^*\big(D_1(M_1\overline{y_1})\big)\\
 & \vdots\\
0 & \in -L_r^*\big(B_r(L_r\overline{x}-L_r\overline{y_r})\big)+M_r^*\big(D_r(M_r\overline{y_r})\big).
\end{cases}
\tag{3.16}
$$

Let $x\in\mathcal{H}$. Then

$$
x \text{ solves (1.7)} \Leftrightarrow z\in Ax+\sum_{k=1}^r\big((L_k^*\circ B_k\circ L_k)\,\Box\,(M_k^*\circ D_k\circ M_k)\big)x+Cx
$$

$$\Leftrightarrow \big(\exists\,(u_k)_{1\leqslant k\leqslant r} \in \mathcal{H}^r\big) \begin{cases} z & \in Ax + \sum_{k=1}^r u_k + Cx \\ u_1 & \in \big((L_1^* \circ B_1 \circ L_1) \,\square\, (M_1^* \circ D_1 \circ M_1)\big)x \\ & \vdots \\ u_r & \in \big((L_r^* \circ B_r \circ L_r) \,\square\, (M_r^* \circ D_r \circ M_r)\big)x \end{cases}$$

$$\Leftrightarrow \big(\exists\,(u_k)_{1\leqslant k\leqslant r} \in \mathcal{H}^r\big) \begin{cases} z & \in Ax + \sum_{k=1}^r u_k + Cx \\ x & \in (L_1^* \circ B_1 \circ L_1)^{-1}u_1 + (M_1^* \circ D_1 \circ M_1)^{-1}u_1 \\ & \vdots \\ x & \in (L_r^* \circ B_r \circ L_r)^{-1}u_r + (M_r^* \circ D_r \circ M_r)^{-1}u_r \end{cases}$$

$$\Leftrightarrow \big(\exists\,(u_k)_{1\leqslant k\leqslant r} \in \mathcal{H}^r\big)\big(\exists\,(y_k)_{1\leqslant k\leqslant r} \in \mathcal{H}^r\big)$$
$$\begin{cases} z & \in Ax + \sum_{k=1}^r u_k + Cx \\ x - y_1 & \in (L_1^* \circ B_1 \circ L_1)^{-1}u_1 \\ y_1 & \in (M_1^* \circ D_1 \circ M_1)^{-1}u_1 \\ & \vdots \\ x - y_r & \in (L_r^* \circ B_r \circ L_r)^{-1}u_r \\ y_r & \in (M_r^* \circ D_r \circ M_r)^{-1}u_r \end{cases}$$

$$\Leftrightarrow \big(\exists\,(u_k)_{1\leqslant k\leqslant r} \in \mathcal{H}^r\big)\big(\exists\,(y_k)_{1\leqslant k\leqslant r} \in \mathcal{H}^r\big)$$
$$\begin{cases} z & \in Ax + \sum_{k=1}^r u_k + Cx \\ u_1 & \in (L_1^* \circ B_1 \circ L_1)(x - y_1) \\ u_1 & \in (M_1^* \circ D_1 \circ M_1)y_1 \\ & \vdots \\ u_r & \in (L_r^* \circ B_r \circ L_r)(x - y_r) \\ u_r & \in (M_r^* \circ D_r \circ M_r)y_r \end{cases}$$

$$\Rightarrow \big(\exists\,(y_k)_{1\leqslant k\leqslant r} \in \mathcal{H}^r\big)$$
$$\begin{cases} z & \in Ax + \sum_{k=1}^r L_k^*\big(B_k(L_k x - L_k y_k)\big) + Cx \\ 0 & \in -L_1^*\big(B_1(L_1 x - L_1 y_1)\big) + M_1^*\big(D_1(M_1 y_1)\big) \\ & \vdots \\ 0 & \in -L_r^*\big(B_r(L_r x - L_r y_r)\big) + M_r^*\big(D_r(M_r y_r)\big). \end{cases} \tag{3.17}$$

Hence since, by assumption, (1.7) has at least one solution,

$$\text{(3.16) has at least one solution.} \tag{3.18}$$

Next, we set

$$\begin{cases} m = r+1 \\ K = 2r \\ \mathcal{H}_1 = \mathcal{H} \\ A_1 = A \\ C_1 = C \\ \mu_1 = \mu \\ z_1 = z \end{cases} \quad \text{and} \quad (\forall k \in \{1,\ldots,r\}) \quad \begin{cases} \mathcal{H}_{k+1} = \mathcal{H} \\ A_{k+1} = 0 \\ B_{k+r} = D_k \\ C_{k+1} = 0 \\ \mu_{k+1} = 0 \\ z_{k+1} = 0. \end{cases} \tag{3.19}$$

We also define

$$(\forall k \in \{1,\ldots,r\}) \quad \mathcal{G}_{k+r} = \mathcal{K}_k \quad \text{and} \tag{3.20}$$

$$(\forall i \in \{1,\ldots,m\}) \quad L_{ki} = \begin{cases} L_k, & \text{if } 1 \leqslant k \leqslant r \text{ and } i = 1; \\ -L_k, & \text{if } 1 \leqslant k \leqslant r \text{ and } i = k+1; \\ M_{k-r}, & \text{if } r+1 \leqslant k \leqslant 2r \text{ and } i = k-r+1; \\ 0, & \text{otherwise.} \end{cases}$$

We observe that in this setting

$$\text{(3.16) is a special case of (2.11).} \tag{3.21}$$

Moreover, if we set $\lambda = \sum_{k=1}^{r} \|L_k\|^2 + \max_{1\leqslant k\leqslant r}(\|L_k\|^2 + \|M_k\|^2)$, we deduce from the Cauchy-Schwarz inequality in $\mathbb{R}^2$ that, for every $(x_i)_{1\leqslant i\leqslant m} = (x, y_1, \ldots, y_r) \in \bigoplus_{i=1}^{m} \mathcal{H}$,

$$\begin{aligned}
\sum_{k=1}^{K} \left\| \sum_{i=1}^{m} L_{ki} x_i \right\|^2 &= \|(L_1 x - L_1 y_1, \ldots, L_r x - L_r y_r, M_1 y_1, \ldots, M_r y_r)\|^2 \\
&\leqslant \big(\|(L_1 x, \ldots, L_r x)\| + \|(L_1 y_1, \ldots, L_r y_r, M_1 y_1, \ldots, M_r y_r)\|\big)^2 \\
&= \left(\sqrt{\sum_{k=1}^{r} \|L_k x\|^2} + \sqrt{\sum_{k=1}^{r} \big(\|L_k y_k\|^2 + \|M_k y_k\|^2\big)} \right)^2 \\
&\leqslant \left(\sqrt{\sum_{k=1}^{r} \|L_k\|^2}\, \|x\| + \sqrt{\sum_{k=1}^{r} (\|L_k\|^2 + \|M_k\|^2)\, \|y_k\|^2} \right)^2 \\
&\leqslant \left(\sqrt{\sum_{k=1}^{r} \|L_k\|^2}\, \|x\| + \max_{1\leqslant k\leqslant r} \sqrt{\|L_k\|^2 + \|M_k\|^2}\, \|(y_1, \ldots, y_r)\| \right)^2 \\
&\leqslant \left(\sum_{k=1}^{r} \|L_k\|^2 + \max_{1\leqslant k\leqslant r} \big(\|L_k\|^2 + \|M_k\|^2\big) \right) \big(\|x\|^2 + \|(y_1, \ldots, y_r)\|^2\big) \\
&= \lambda \sum_{i=1}^{m} \|x_i\|^2.
\end{aligned} \tag{3.22}$$

Thus,

$$(1.6) \text{ is a special case of } (2.10). \tag{3.23}$$

Now, let us define

$$(\forall n \in \mathbb{N}) \quad x_{1,n} = x_n \quad \text{and} \quad (\forall k \in \{1,\ldots,r\}) \quad \begin{cases} x_{k+1,n} = y_{k,n} \\ v_{k+r,n} = w_{k,n}. \end{cases} \tag{3.24}$$

Then it follows from (3.19) that

$$(3.15) \text{ is a special case of } (2.13). \tag{3.25}$$

Altogether, Theorem 2.8(i)–(iv) asserts that there exist a solution $\overline{\boldsymbol{x}} = (\overline{x_1},\ldots,\overline{x_m}) = (\overline{x},\overline{y_1},\ldots,\overline{y_r})$ to (2.11) and a solution $(\overline{v_1},\ldots,\overline{v_K}) = (\overline{v_1},\ldots,\overline{v_r},\overline{w_1},\ldots,\overline{w_r})$ to (2.12) which satisfy

$$x_n \rightharpoonup \overline{x} \quad \text{and} \quad (\forall k \in \{1,\ldots,r\}) \quad \begin{cases} v_{k,n} \rightharpoonup \overline{v_k} \\ w_{k,n} \rightharpoonup \overline{w_k}, \end{cases} \tag{3.26}$$

together with the inclusions

$$z - \sum_{k=1}^{r} L_k^* \overline{v_k} \in A\overline{x} + C\overline{x} \quad \text{and} \quad (\forall k \in \{1,\ldots,r\}) \quad \begin{cases} L_k^*\overline{v_k} = M_k^*\overline{w_k} \\ L_k\overline{x} - L_k\overline{y}_k \in B_k^{-1}\overline{v_k} \\ M_k\overline{y}_k \in D_k^{-1}\overline{w_k}. \end{cases} \tag{3.27}$$

Using the notation (3.2), we can rewrite (3.27) as

$$\mathbf{0} \in \boldsymbol{A}\overline{\boldsymbol{x}} + \boldsymbol{L}^*\big(\boldsymbol{B}(\boldsymbol{L}\overline{\boldsymbol{x}})\big). \tag{3.28}$$

In turn, it follows from Proposition 3.1(iii)–(iv) that

$$\overline{x} \text{ solves (1.7) and } (\overline{v_1},\ldots,\overline{v_r}) \text{ solves (1.8).} \tag{3.29}$$

This and (3.26) prove (i). Finally, (ii) and (iii) follow from (3.19) and Theorem 2.8(v)–(vi). ☐

Remark 3.3. In the spirit of the splitting methods of [10, 12], the algorithm described in (3.15) achieves full decomposition in that every operator is used individually at each iteration.

4. Application to convex minimization

In this section we consider a structured minimization problem of the following format.

Problem 4.1. Let $\mathcal{H}$ be a real Hilbert space, let r be a strictly positive integer, let $z \in \mathcal{H}$, let $f \in \Gamma_0(\mathcal{H})$, and let $\ell\colon \mathcal{H} \to \mathbb{R}$ be a differentiable convex function such that $\nabla\ell$ is μ-Lipschitzian for some $\mu \in [0,+\infty[$. For every integer $k \in \{1,\ldots,r\}$, let $\mathcal{G}_k$ and $\mathcal{K}_k$ be real Hilbert spaces, let $g_k \in \Gamma_0(\mathcal{G}_k)$ and $h_k \in \Gamma_0(\mathcal{K}_k)$, and let $L_k \in \mathcal{B}(\mathcal{H},\mathcal{G}_k)$ and $M_k \in \mathcal{B}(\mathcal{H},\mathcal{K}_k)$. It is assumed that

$$\beta = \mu + \sqrt{\sum_{k=1}^{r} \|L_k\|^2 + \max_{1\leqslant k\leqslant r}\big(\|L_k\|^2 + \|M_k\|^2\big)} > 0, \tag{4.1}$$

that

$$(\forall k \in \{1,\ldots,r\}) \quad 0 \in \operatorname{sri}\big(\operatorname{dom}(g_k \circ L_k)^* - M_k^*(\operatorname{dom} h_k^*)\big), \tag{4.2}$$

that

$$(\forall k \in \{1,\ldots,r\}) \quad 0 \in \operatorname{sri}\big(\operatorname{ran} M_k - \operatorname{dom} h_k\big), \tag{4.3}$$

and that

$$z \in \operatorname{ran}\left(\partial f + \sum_{k=1}^{r} \big((L_k^* \circ (\partial g_k) \circ L_k) \,\square\, (M_k^* \circ (\partial h_k) \circ M_k)\big) + \nabla \ell\right). \tag{4.4}$$

Solve the primal problem

$$\underset{x \in \mathcal{H}}{\text{minimize}} \;\; f(x) + \sum_{k=1}^{r} \big((g_k \circ L_k) \,\square\, (h_k \circ M_k)\big)(x) + \ell(x) - \langle x \mid z\rangle, \tag{4.5}$$

together with the dual problem

$$\underset{v_1 \in \mathcal{G}_1,\ldots,\, v_r \in \mathcal{G}_r}{\text{minimize}} \;\; (f^* \,\square\, \ell^*)\left(z - \sum_{k=1}^{r} L_k^* v_k\right) + \sum_{k=1}^{r} \big(g_k^*(v_k) + (M_k^* \rhd h_k^*)(L_k^* v_k)\big). \tag{4.6}$$

Special cases when (4.2) and (4.3) are satisfied can be derived from [2, Proposition 15.24]. The next proposition describes scenarios in which (4.4) holds.

Proposition 4.2. *Consider the same setting as in Problem 4.1 with the exception that assumption* (4.4) *is not made and is replaced by the assumptions that*

$$\begin{aligned}\boldsymbol{E} = \Big\{ \big(L_1(x-y_1)-s_1,\ldots,L_r(x-y_r)-s_r, M_1y_1-t_1,\ldots,M_ry_r-t_r\big) \;\Big|\; x \in \operatorname{dom} f, \\ y_1 \in \mathcal{H},\ldots,y_r \in \mathcal{H}, s_1 \in \operatorname{dom} g_1,\ldots,s_r \in \operatorname{dom} g_r, t_1 \in \operatorname{dom} h_1,\ldots,t_r \in \operatorname{dom} h_r\Big\} \neq \varnothing\end{aligned} \tag{4.7}$$

and that (4.5) *has a solution. Then* (4.4) *is satisfied in each of the following cases.*

(i) $\mathbf{0} \in \operatorname{sri} \boldsymbol{E}$.

(ii) $\boldsymbol{E}$ *is a closed vector subspace.*

(iii) f *is real-valued and, for every* $k \in \{1,\ldots,r\}$, *the operators* L_k *and* M_k *are surjective.*

(iv) *For every* $k \in \{1,\ldots,r\}$, g_k *and* h_k *are real-valued.*

(v) $\mathcal{H}$, $(\mathcal{G}_k)_{1\leqslant k\leqslant r}$, *and* $(\mathcal{K}_k)_{1\leqslant k\leqslant r}$ *are finite-dimensional, and*

$$(\exists\, x \in \operatorname{ri}\operatorname{dom} f)(\forall k \in \{1,\ldots,r\})(\exists\, y_k \in \mathcal{H}) \quad \begin{cases} L_k(x-y_k) \in \operatorname{ri}\operatorname{dom} g_k \\ M_k y_k \in \operatorname{ri}\operatorname{dom} h_k. \end{cases} \tag{4.8}$$

Proof. Let us define $\boldsymbol{\mathcal{H}}$, $\boldsymbol{\mathcal{G}}$, and $\boldsymbol{\mathcal{K}}$ as in (3.1), $\boldsymbol{L}$ as in (3.2), and let us set

$$\begin{cases} \boldsymbol{f}\colon \boldsymbol{\mathcal{H}} \to \left]-\infty,+\infty\right] : \boldsymbol{x} = (x,y_1,\ldots,y_r) \mapsto f(x) + \ell(x) - \langle x \mid z\rangle \\ \boldsymbol{g}\colon \boldsymbol{\mathcal{G}} \oplus \boldsymbol{\mathcal{K}} \to \left]-\infty,+\infty\right] : \boldsymbol{s} = (s_1,\ldots,s_r,t_1,\ldots,t_r) \mapsto \sum_{k=1}^{r} \big(g_k(s_k) + h_k(t_k)\big). \end{cases} \tag{4.9}$$

Then we can rewrite (4.7) as

$$\boldsymbol{E} = \boldsymbol{L}(\operatorname{dom} \boldsymbol{f}) - \operatorname{dom} \boldsymbol{g}. \tag{4.10}$$

(i): Since $\boldsymbol{E} \neq \varnothing$, the functions $(g_k \circ L_k)_{1 \leqslant k \leqslant r}$ and $(h_k \circ M_k)_{1 \leqslant k \leqslant r}$ are proper and therefore in $\Gamma_0(\mathcal{H})$. In turn, the Fenchel-Moreau theorem [2, Theorem 13.32] asserts that the functions $((g_k \circ L_k)^*)_{1 \leqslant k \leqslant r}$ and $((h_k \circ M_k)^*)_{1 \leqslant k \leqslant r}$ are in $\Gamma_0(\mathcal{H})$. On the other hand, since (4.3) and [2, Corollary 15.28(i)] imply that

$$(\forall k \in \{1, \ldots, r\}) \quad (h_k \circ M_k)^* = M_k^* \rhd h_k^*, \tag{4.11}$$

(4.2) and [2, Proposition 12.34(i)] yield

$$\begin{aligned}(\forall k \in \{1, \ldots, r\}) \quad 0 &\in \operatorname{sri}\big(\operatorname{dom}(g_k \circ L_k)^* - M_k^*(\operatorname{dom} h_k^*)\big)\\ &= \operatorname{sri}\big(\operatorname{dom}(g_k \circ L_k)^* - \operatorname{dom}(M_k^* \rhd h_k^*)\big)\\ &= \operatorname{sri}\big(\operatorname{dom}(g_k \circ L_k)^* - \operatorname{dom}(h_k \circ M_k)^*\big).\end{aligned} \tag{4.12}$$

Hence, we derive from [2, Proposition 15.7] that

$$\begin{aligned}&(\forall k \in \{1, \ldots, r\})(\forall x \in \mathcal{H})(\exists\, y_k \in \mathcal{H})\\ &\qquad \big((g_k \circ L_k) \,\square\, (h_k \circ M_k)\big)(x) = g_k(L_k x - L_k y_k) + h_k(M_k y_k),\end{aligned} \tag{4.13}$$

which allows us to rewrite (4.5) as a minimization problem on $\boldsymbol{\mathcal{H}}$, namely

$$\underset{x \in \mathcal{H}, y_1 \in \mathcal{H}, \ldots, y_r \in \mathcal{H}}{\operatorname{minimize}} \; f(x) + \ell(x) - \langle x \mid z \rangle + \sum_{k=1}^{r} \big(g_k(L_k x - L_k y_k) + h_k(M_k y_k)\big) \tag{4.14}$$

or, equivalently,

$$\underset{\boldsymbol{x} \in \boldsymbol{\mathcal{H}}}{\operatorname{minimize}} \; \boldsymbol{f}(\boldsymbol{x}) + \boldsymbol{g}(\boldsymbol{L}\boldsymbol{x}). \tag{4.15}$$

It follows from (4.10) that $\mathbf{0} \in \operatorname{sri}\big(\boldsymbol{L}(\operatorname{dom} \boldsymbol{f}) - \operatorname{dom} \boldsymbol{g}\big)$ and therefore from [2, Theorem 16.37(i)], that

$$\partial(\boldsymbol{f} + \boldsymbol{g} \circ \boldsymbol{L}) = \partial \boldsymbol{f} + \boldsymbol{L}^* \circ (\partial \boldsymbol{g}) \circ \boldsymbol{L}. \tag{4.16}$$

Since, by assumption, (4.5) has a solution, so does (4.15). By Fermat's rule [2, Theorem 16.2], this means that $\mathbf{0} \in \operatorname{ran} \partial(\boldsymbol{f} + \boldsymbol{g} \circ \boldsymbol{L})$. Thus (4.16) yields

$$\mathbf{0} \in \operatorname{ran}\big(\partial \boldsymbol{f} + \boldsymbol{L}^* \circ (\partial \boldsymbol{g}) \circ \boldsymbol{L}\big). \tag{4.17}$$

Let us introduce the operators

$$A = \partial f, \; C = \nabla \ell, \quad \text{and} \quad (\forall k \in \{1, \ldots, r\}) \quad \begin{cases} B_k = \partial g_k \\ D_k = \partial h_k. \end{cases} \tag{4.18}$$

We derive from [2, Proposition 17.10] that C is monotone and from [2, Theorem 20.40] that the operators A, $(B_k)_{1 \leqslant k \leqslant r}$, and $(D_k)_{1 \leqslant k \leqslant r}$ are maximally monotone. Next, let us define $\boldsymbol{A}$ and $\boldsymbol{B}$ as in (3.2). Then it follows from (4.17) and [2, Proposition 16.8] that (3.3) holds. In turn, Proposition 3.1(iii) asserts that (4.4) is satisfied.

(ii)⇒(i): This follows from [2, Proposition 6.19(i)].

(iii)⇒(i) and (iv)⇒(i): In both cases $\boldsymbol{E} = \boldsymbol{\mathcal{G}} \oplus \boldsymbol{\mathcal{K}}$.

(v)⇒(i): Since $\boldsymbol{\mathcal{H}}$, $\boldsymbol{\mathcal{G}}$, and $\boldsymbol{\mathcal{K}}$ are finite-dimensional, (4.10) and [2, Corollary 6.15] imply that

$$(4.8) \Leftrightarrow (\exists\, \boldsymbol{x} \in \operatorname{ri} \operatorname{dom} \boldsymbol{f}) \quad \boldsymbol{L}\boldsymbol{x} \in \operatorname{ri} \operatorname{dom} \boldsymbol{g}$$

$$
\begin{aligned}
&\Leftrightarrow \mathbf{0} \in \big(\boldsymbol{L}(\operatorname{ri}\operatorname{dom}\boldsymbol{f}) - \operatorname{ri}\operatorname{dom}\boldsymbol{g}\big)\\
&\Leftrightarrow \mathbf{0} \in \operatorname{ri}\big(\boldsymbol{L}(\operatorname{dom}\boldsymbol{f}) - \operatorname{dom}\boldsymbol{g}\big)\\
&\Leftrightarrow \mathbf{0} \in \operatorname{ri}\boldsymbol{E}\\
&\Leftrightarrow \mathbf{0} \in \operatorname{sri}\boldsymbol{E},
\end{aligned}
\tag{4.19}
$$

which completes the proof. □

Next, we propose our algorithm for solving Problem 4.1.

Theorem 4.3. *Consider the setting of Problem 4.1. Let* $x_0 \in \mathcal{H}$, $y_{1,0} \in \mathcal{H}$, $\ldots$, $y_{r,0} \in \mathcal{H}$, $v_{1,0} \in \mathcal{G}_1$, $\ldots$, $v_{r,0} \in \mathcal{G}_r$, $w_{1,0} \in \mathcal{K}_1$, $\ldots$, $w_{r,0} \in \mathcal{K}_r$, *let* $\varepsilon \in \left]0, 1/(\beta+1)\right[$, *let* $(\gamma_n)_{n\in\mathbb{N}}$ *be a sequence in* $[\varepsilon, (1-\varepsilon)/\beta]$, *and set*

$$
\begin{array}{l}
\text{for } n = 0,1,\ldots\\
\left\lfloor
\begin{array}{l}
s_{1,1,n} \approx x_n - \gamma_n\big(\nabla\ell(x_n) + \sum_{k=1}^r L_k^* v_{k,n}\big)\\
p_{1,1,n} \approx \operatorname{prox}_{\gamma_n f}(s_{1,1,n} + \gamma_n z)\\
\text{for } k = 1,\ldots,r\\
\left\lfloor
\begin{array}{l}
p_{1,k+1,n} \approx y_{k,n} + \gamma_n(L_k^* v_{k,n} - M_k^* w_{k,n})\\
s_{2,k,n} \approx v_{k,n} + \gamma_n L_k(x_n - y_{k,n})\\
p_{2,k,n} \approx s_{2,k,n} - \gamma_n \operatorname{prox}_{\gamma_n^{-1} g_k}(\gamma_n^{-1} s_{2,k,n})\\
q_{2,k,n} \approx p_{2,k,n} + \gamma_n L_k(p_{1,1,n} - p_{1,k+1,n})\\
v_{k,n+1} = v_{k,n} - s_{2,k,n} + q_{2,k,n}\\
s_{2,k+r,n} \approx w_{k,n} + \gamma_n M_k y_{k,n}\\
p_{2,k+r,n} \approx s_{2,k+r,n} - \gamma_n\big(\operatorname{prox}_{\gamma_n^{-1} h_k}(\gamma_n^{-1} s_{2,k+r,n})\big)\\
q_{1,k+1,n} \approx p_{1,k+1,n} + \gamma_n(L_k^* p_{2,k,n} - M_k^* p_{2,k+r,n})\\
q_{2,k+r,n} \approx p_{2,k+r,n} + \gamma_n M_k p_{1,k+1,n}\\
w_{k,n+1} = w_{k,n} - s_{2,k+r,n} + q_{2,k+r,n}
\end{array}
\right.\\
q_{1,1,n} \approx p_{1,1,n} - \gamma_n\big(\nabla\ell(p_{1,1,n}) + \sum_{k=1}^r L_k^* p_{2,k,n}\big)\\
x_{n+1} = x_n - s_{1,1,n} + q_{1,1,n}\\
\text{for } k = 1,\ldots,r\\
\left\lfloor\ y_{k,n+1} = y_{k,n} - p_{1,k+1,n} + q_{1,k+1,n}.\right.
\end{array}
\right.
\end{array}
\tag{4.20}
$$

Then the following hold for some solution $\overline{x}$ *to* (4.5) *and some solution* $(\overline{v_1}, \ldots, \overline{v_r})$ *to* (4.6).

(i) $x_n \rightharpoonup \overline{x}$ *and* $(\forall k \in \{1,\ldots,r\})$ $v_{k,n} \rightharpoonup \overline{v_k}$.
(ii) *Suppose that* f *or* ℓ *is uniformly convex at* $\overline{x}$. *Then* $x_n \to \overline{x}$.
(iii) *Suppose that, for some* $k \in \{1,\ldots,r\}$, g_k^* *is uniformly convex at* $\overline{v_k}$. *Then* $v_{k,n} \to \overline{v_k}$.

Proof. Set

$$
A = \partial f,\ C = \nabla\ell, \quad \text{and} \quad (\forall k \in \{1,\ldots,r\}) \quad \begin{cases} B_k = \partial g_k\\ D_k = \partial h_k. \end{cases}
\tag{4.21}
$$

We derive from [2, Proposition 17.10] that C is monotone. Furthermore, [2, Theorem 20.40 and Corollary 16.24] assert that the operators A, $(B_k)_{1\leqslant k\leqslant r}$, and $(D_k)_{1\leqslant k\leqslant r}$ are maximally monotone with inverses respectively given by ∂f^*, $(\partial g_k^*)_{1\leqslant k\leqslant r}$, and $(\partial h_k^*)_{1\leqslant k\leqslant r}$. Moreover, (4.4) implies that (1.7) has a solution. Now let x and $\boldsymbol{v} = (v_k)_{1\leqslant k\leqslant r}$ be, respectively, the solutions to (1.7) and (1.8) produced by Theorem 3.2. Since the uniform convexity of a function at a point implies the uniform monotonicity of its subdifferential at that point [27, Section 3.4] and since, in the setting of (4.21), (4.20) reduces to (3.15) thanks to (2.4), it is enough to show that x solves (4.5) and $\boldsymbol{v}$ solves (4.6). To this end, we first derive from (4.12) and [2, Propositions 16.5(ii) and 24.27] that

$$(\forall k\in\{1,\ldots,r\})\quad \big(L_k^*\circ(\partial g_k)\circ L_k\big)\,\Box\,\big(M_k^*\circ(\partial h_k)\circ M_k\big)\subset\partial(g_k\circ L_k)\,\Box\,\partial(h_k\circ M_k) = \partial\big((g_k\circ L_k)\,\Box\,(h_k\circ M_k)\big). \tag{4.22}$$

Hence, it follows from (4.21) and Fermat's rule [2, Theorem 16.2] that

$$\begin{aligned} x \text{ solves (1.7)} &\Rightarrow z\in\partial f(x)+\sum_{k=1}^r\big((L_k^*\circ(\partial g_k)\circ L_k)\,\Box\,(M_k^*\circ(\partial h_k)\circ M_k)\big)x+\nabla\ell(x)\\ &\Rightarrow z\in\partial f(x)+\sum_{k=1}^r\partial\Big((g_k\circ L_k)\,\Box\,(h_k\circ M_k)\Big)x+\partial\ell(x)\\ &\Rightarrow 0\in\partial\Big(f+\sum_{k=1}^r\big((g_k\circ L_k)\,\Box\,(h_k\circ M_k)\big)+\ell-\langle\cdot\mid z\rangle\Big)(x)\\ &\Rightarrow x \text{ solves (4.5)}. \end{aligned}\tag{4.23}$$

On the other hand, (4.3) and Corollary 2.6(ii) yield

$$(\forall k\in\{1,\ldots,r\})\quad M_k^*\rhd\partial h_k^* = \partial(M_k^*\rhd h_k^*), \tag{4.24}$$

while [2, Proposition 16.5(ii)] yields

$$(\forall k\in\{1,\ldots,r\})\quad \partial g_k^*+L_k\circ\big(\partial(M_k^*\rhd h_k^*)\big)\circ L_k^*\subset\partial\big(g_k^*+(M_k^*\rhd h_k^*)\circ L_k^*\big). \tag{4.25}$$

Now define $\boldsymbol{\mathcal{G}}$ as in (3.1) and

$$\begin{cases}\varphi\colon\mathcal{H}\to\,]-\infty,+\infty]:u\mapsto(f^*\,\Box\,\ell^*)(z+u)\\ \boldsymbol{\psi}\colon\boldsymbol{\mathcal{G}}\to\,]-\infty,+\infty]:\boldsymbol{v}\mapsto\sum_{k=1}^r\big(g_k^*(v_k)+(M_k^*\rhd h_k^*)(L_k^*v_k)\big)\\ \boldsymbol{M}\colon\boldsymbol{\mathcal{G}}\to\mathcal{H}\colon\boldsymbol{v}\mapsto-\sum_{k=1}^r L_k^*v_k.\end{cases}\tag{4.26}$$

Then

$$(\forall\boldsymbol{v}\in\boldsymbol{\mathcal{G}})\quad \varphi(\boldsymbol{M}\boldsymbol{v})+\boldsymbol{\psi}(\boldsymbol{v}) = (f^*\,\Box\,\ell^*)\Big(z-\sum_{k=1}^r L_k^*v_k\Big)+\sum_{k=1}^r\big(g_k^*(v_k)+(M_k^*\rhd h_k^*)(L_k^*v_k)\big). \tag{4.27}$$

Invoking successively (4.21), (4.24), (4.25), [2, Proposition 16.8], (4.26), (4.27), and Fermat's rule, we get

$$\boldsymbol{v}\text{ solves (1.8)}\Rightarrow(\forall k\in\{1,\ldots,r\})\;\;0\in -L_k\Big(\partial(f+\ell)^*\Big(z-\sum_{l=1}^r L_l^*v_l\Big)\Big)$$

$$
\begin{aligned}
&+ \partial g_k^*(v_k) + L_k\Big(\big(M_k^* \rhd \partial h_k^*\big)(L_k^* v_k)\Big) \\
&\Rightarrow (\forall k \in \{1,\ldots,r\})\ \ 0 \in -L_k\Bigg(\partial(f^* \,\square\, \ell^*)\Bigg(z - \sum_{l=1}^{r} L_l^* v_l\Bigg)\Bigg) \\
&+ \partial\big(g_k^* + (M_k^* \rhd h_k^*) \circ L_k^*\big)(v_k) \\
&\Rightarrow \mathbf{0} \in \big(\boldsymbol{M}^* \circ (\partial\varphi) \circ \boldsymbol{M}\big)(\boldsymbol{v}) + \partial\boldsymbol{\psi}(\boldsymbol{v}) \\
&\Rightarrow \mathbf{0} \in \partial\big(\varphi \circ \boldsymbol{M} + \boldsymbol{\psi}\big)(\boldsymbol{v}) \\
&\Rightarrow \boldsymbol{v} \text{ solves (4.6)},
\end{aligned}
\tag{4.28}
$$

which completes the proof. □

Theorem 4.3 enables us to solve a new class of structured minimization problems featuring both infimal convolutions and postcompositions. The special cases of this model which arise in the area of image recovery [7, 21] initially motivated our investigation. Such applications are considered in the next section.

5. Image restoration application

5.1. **Image restoration.** Proximal splitting methods were introduced in the field of image recovery in [14] for variational models of the form

$$
\underset{x\in\mathcal{H}}{\text{minimize}}\ \ f(x) + \ell(x), \tag{5.1}
$$

where f and ℓ are as in Problem 4.1 (see [11] for recent developments in this application area). In this section we show a full fledged implementation of the algorithm in Theorem 4.3 in the Euclidean setting ($\mathcal{H} = \mathbb{R}^N$) which goes much beyond (5.1). For this purpose, we consider the problem of image restoration from a blurred image [1]. Imaging devices, such as cameras, microscopes, and telescopes, distort the light field due to both optical imperfections and diffraction; another source of blur is relative movement of the scene and the device during the exposure, as happens when taking a photo in low-light without a tripod or when a telescope observes the stars with imperfect motion compensation. The effect is that the recorded image is the convolution of the true scene with a function known as the point-spread function. The resulting convolution operator T is called the blur operator.

The original N-pixel ($N = 512^2$) image shown in Fig. 1 (a) is degraded by a linear blurring operator T associated with a 21-pixel long point-spread function corresponding to motion blur, followed by addition of a noise component w. Images in their natural matrix form are converted to vectors $x \in \mathbb{R}^N$ by stacking columns together. We write the coefficients of x as $x = (\xi_i)_{1\leqslant i\leqslant N}$, but when we wish to make use of the 2-dimensional nature of the image (as a $\sqrt{N} \times \sqrt{N}$ image), we use the convention $\xi_{i,j} = \xi_{(j-1)\sqrt{N}+i}$ for every i and j in $\{1,\ldots,\sqrt{N}\}$, so that i and j refer to the row and column indices, respectively. The degraded image

$$
y = T\overline{x} + w \tag{5.2}
$$

is shown in Fig. 1 (b). The noise level is chosen to give y a signal-to-noise ratio of 45 dB relative to $T\overline{x}$. The variational formulation we propose to recover $\overline{x}$ is an instantiation of Problem 4.1 with $r = 2$, namely,

$$\underset{x\in C}{\text{minimize}} \ \big((\alpha\|\cdot\|_{1,2}\circ D^{(1)})\,\Box\,(\beta\|\cdot\|_{1,2}\circ D^{(2)})\big)(x) + \gamma\|Wx\|_1 + \frac{1}{2}\|Tx-y\|_2^2 \tag{5.3}$$

or, equivalently,

$$\begin{aligned}\underset{x\in\mathcal{H}}{\text{minimize}} \ \ & \underbrace{\iota_C}_{f}(x) + \Big(\underbrace{(\alpha\|\cdot\|_{1,2}\circ D^{(1)})}_{g_1\circ L_1}\,\Box\,\underbrace{(\beta\|\cdot\|_{1,2}\circ D^{(2)})}_{h_1\circ M_1}\Big)(x)\\ & + \big(\underbrace{\gamma\|W\cdot\|_1}_{g_2\circ L_2}\,\Box\,\underbrace{(\iota_{\{0\}}\circ \mathrm{Id})}_{h_2\circ M_2}\big)(x) + \underbrace{\frac{1}{2}\|T\cdot-y\|_2^2}_{\ell}(x).\end{aligned} \tag{5.4}$$

In this model, α, β, and γ are strictly positive constants, and C is a constraint set modeling the known amplitude bounds on pixel values; here $C = [0,1]^N$. To promote the piecewise smoothness of $\overline{x}$ we use an inf-convolution term mixing first- and second-order total variation potentials, in a fashion initially advocated in [7] and further explored in [21]. First-order total variation is commonly used in image processing, but suffers from staircase effects (see, e.g., [7]), which are reduced by using the inf-convolution model. The operators $D^{(1)}$ and $D^{(2)}$ are, respectively, first and second order discrete gradient operators that map $\mathbb{R}^N$ to $\mathbb{R}^{N\times M}$ for $M = 2$ and $M = 3$, respectively (see section 5.2 for details). The functions g_1 and h_1 are the usual mixed norms defined on $\mathbb{R}^{N\times M}$ as

$$\|\cdot\|_{1,2}\colon x \mapsto \sum_{i=1}^{N}\sqrt{\sum_{j=1}^{M}\xi_{i,j}^2}, \tag{5.5}$$

which is the sum of the norm of the rows of x. The potential

$$x \mapsto \|Wx\|_1, \tag{5.6}$$

where W is the analysis operator of a weighted 9/7 biorthogonal wavelet frame [8], is intended to promote sparsity of the wavelet coefficients of x. Since natural images are known to have approximately sparse wavelet representations, this term penalizes noise, which does not have a sparse wavelet representation. Such wavelet terms are standard in the literature, and are often used in conjunction with a first-order TV term [19]. Finally, data fidelity is promoted by the potential

$$\ell\colon x \mapsto \frac{1}{2}\|Tx-y\|^2. \tag{5.7}$$

Remark 5.1. Here are some comments on the implementation of the algorithm from Theorem 4.3 in the setting of (5.4).

(i) The proximity operator of $f = \iota_C$ is simply the projector onto a hypercube, which is straightforward.

(ii) By [2, Example 14.5], for every $x \in \mathcal{H}\smallsetminus\{0\}$,

$$\mathrm{prox}_{\|\cdot\|}x = \left(1 - \frac{1}{\|x\|}\right)x \tag{5.8}$$

and $\mathrm{prox}_{\|\cdot\|}0 = 0$. Since $\|x\|_{1,2}$ is separable in the rows of x, $\mathrm{prox}_{\|\cdot\|_{1,2}}x$ is computed by applying (5.8) to each row.

(iii) The gradient of ℓ is $\nabla\ell\colon x \mapsto T^\top(Tx-y)$, which is Lipschitz continuous with constant $\|T\|^2$.
(iv) The proximity operator of $\|\cdot\|_1$ is implemented by soft-thresholding of each component [11].
(v) No special assumption is required on the structure of W (e.g., the frame need not be tight or, in particular, an orthonormal basis). Without assumptions on W, there is no known closed-form proximity operator of $x \mapsto \gamma\|Wx\|_1$, which is why it is important to treat $\|\cdot\|_1$ and W separately.
(vi) We have used only one hard constraint set C, but it is clear that our framework can accommodate an arbitrary number of constraint sets, hence permitting one to inject easily *a priori* information in the restoration process. Each additional hard constraint of the type $L_k\overline{x} \in C_k$ can be handled by setting $g_k = \iota_{C_k}$, $h_k = \iota_{\{0\}}$, and $M_k = \mathrm{Id}$.

Remark 5.2. Remark 5.1 shows that the computation of proximity operators for each function involved in (5.4) is implementable. It is also possible to compute proximity operators for scaled versions of the above functions. Let $\rho \in \left]0,+\infty\right[$. Then given $\varphi \in \Gamma_0(\mathcal{H})$ and $\widetilde{\varphi}\colon x \mapsto \varphi(\rho x)$, [2, Corollary 23.24] implies that

$$(\forall x \in \mathcal{H}) \quad \mathrm{prox}_{\widetilde{\varphi}}\, x = \rho^{-1}\mathrm{prox}_{\rho^2\varphi}(\rho x). \tag{5.9}$$

This gives the possibility of writing $f(Lx)$ as $\tilde{f}(\tilde{L}x)$ for $\tilde{L} = \rho^{-1}L$. Our implementation will exploit this flexibility in order to rescale all L_k and M_k operators to have unit operator norm. Numerical evidence suggests that this improves convergence profiles since all dual variables $(v_k)_{1\leqslant k\leqslant r}$ and $(w_k)_{1\leqslant k\leqslant r}$ are approximately of the same scale.

5.2. **Total variation.** Total variation can be defined for mathematical objects such as measures and functions [28]. In a discrete setting, there are many possible definitions of total variation. We use the standard isotropic discretization,

$$\begin{aligned} \mathrm{tv}(x) &= \sum_{i=1}^{\sqrt{N}-1}\sum_{j=1}^{\sqrt{N}-1}\sqrt{(\xi_{i+1,j}-\xi_{i,j})^2+(\xi_{i,j+1}-\xi_{i,j})^2}, \\ x &= (\xi_k)_{1\leqslant k\leqslant N},\ \xi_{i,j} = \xi_{(j-1)\sqrt{N}+i}, \end{aligned} \tag{5.10}$$

originally advocated in [20]. There is no known closed form expression for the proximity operator of (5.10).

Infimal-convolution with a second-order total variation term was first suggested in [7]. We use the particular second-order total variation term corresponding to "$\mathcal{D}_{2,b}$" (with weights $b = (1,1/2,1)$) from [21]. We now show how to recover the relation $\mathrm{tv}(x) = \|D^{(1)}x\|_{1,2}$. Define the horizontal finite-difference operator by

$$\begin{aligned} &D_{\leftrightarrow}\colon \mathbb{R}^N \to \mathbb{R}^{\sqrt{N}\times\sqrt{N}}\colon x \mapsto z = (\zeta_{i,j})_{1\leqslant i,j\leqslant\sqrt{N}}, \\ &\zeta_{i,j} = \begin{cases} \xi_{i,j+1}-\xi_{i,j}, & \text{if } 1 \leqslant j < \sqrt{N}; \\ 0, & \text{if } j = \sqrt{N}, \end{cases} \end{aligned} \tag{5.11}$$

and the vertical operator $D_{\updownarrow}$ by $D_{\updownarrow}\colon\ x \mapsto (D_{\leftrightarrow}(x^\top))^\top$. Let $\operatorname{vec}(\cdot)$ be the mapping that re-orders a matrix by stacking the columns together, and define $D^{(1)}\colon x \mapsto (\operatorname{vec}(D_{\leftrightarrow}(x)), \operatorname{vec}(D_{\updownarrow}(x)))$. Then by comparing (5.5) with (5.10), we observe that $\operatorname{tv}(x) = \|D^{(1)}x\|_{1,2}$.

The second-order total variation potential makes use of an additional set of first-order difference operators that have different boundary conditions, namely

$$\widetilde{D}_{\leftrightarrow} : \mathbb{R}^N \to \mathbb{R}^{\sqrt{N}\times\sqrt{N}} : x \mapsto z = (\zeta_{i,j})_{1\leqslant i,j\leqslant\sqrt{N}}, \tag{5.12}$$
$$\zeta_{i,j} = \begin{cases} \xi_{i,j} - \xi_{i,j-1}, & \text{if } 1 < j < \sqrt{N};\\ \xi_{i,j}, & \text{if } j = 1;\\ -\xi_{i,j-1}, & \text{if } j = \sqrt{N}, \end{cases}$$

and $\widetilde{D}_{\updownarrow}\colon x \mapsto (\widetilde{D}_{\leftrightarrow}(x^\top))^\top$. Then define

$$D^{(2)}\colon x \mapsto \left(\operatorname{vec}(\widetilde{D}_{\leftrightarrow}(D_{\leftrightarrow}x)), \frac{\operatorname{vec}(\widetilde{D}_{\leftrightarrow}(D_{\updownarrow}x)) + \operatorname{vec}(\widetilde{D}_{\updownarrow}(D_{\leftrightarrow}x))}{\sqrt{2}}, \operatorname{vec}(\widetilde{D}_{\updownarrow}(D_{\updownarrow}x))\right). \tag{5.13}$$

The second-order total variation potential is defined as $x \mapsto \|D^{(2)}x\|_{1,2}$.

5.3. **Constraint qualifications.** To apply the results of Theorem 4.3, we need to check that the constraint qualifications (4.2), (4.3), and (4.4) hold. Starting with (4.2), for each $k \in \{1,2\}$ we have

$$\begin{aligned} \operatorname{sri}\big(\operatorname{dom}(g_k \circ L_k)^* - M_k^*(\operatorname{dom} h_k^*)\big) &= \operatorname{sri}\big(\operatorname{dom}(L_k^* \triangleright g_k^*) - M_k^*(\operatorname{dom} h_k^*)\big)\\ &= \operatorname{sri}\big(L_k^*(\operatorname{dom} g_k^*) - M_k^*(\operatorname{dom} h_k^*)\big)\\ &= L_k^*\big(\operatorname{ri}\operatorname{dom} g_k^*) - M_k^*(\operatorname{ri}\operatorname{dom} h_k^*\big), \end{aligned} \tag{5.14}$$

where the first line follows from [2, Proposition 15.28] and the fact that g_k has full domain, the second line follows from [2, Proposition 12.34(i)], and the third line follows from [2, Corollary 6.15]. Since g_1, g_2, and h_1 are coercive, their conjugates all include 0 in the interior of their domain [2, Theorem 14.17]. Furthermore, the conjugate of $h_2 = \iota_{\{0\}}$ is $h_2^* = 0$ which has full domain. Thus,

$$(\forall k \in \{1,2\}) \quad 0 \in L_k^*(\operatorname{ri}\operatorname{dom} g_k^*) \quad \text{and} \quad 0 \in M_k^*(\operatorname{ri}\operatorname{dom} h_k^*). \tag{5.15}$$

Altogether, (5.14) is satisfied for each $k \in \{1,2\}$ and hence so is (4.2). The qualification (4.3) holds for $k = 1$ since $h_1 = \|\cdot\|_{1,2}$ has full domain. For $k = 2$, since $h_2 = \iota_{\{0\}}$, using [2, Corollary 6.15] and the linearity of M_2, we obtain

$$\operatorname{sri}(\operatorname{ran} M_2 - \operatorname{dom} h_2) = \operatorname{sri}(\operatorname{ran} M_2) = \operatorname{ri}(\operatorname{ran} M_2) = \operatorname{ran} M_2. \tag{5.16}$$

Thus, since $0 \in \operatorname{ran} M_2$, (4.3) is satisfied. On the other hand, since $\mathcal{H}$ is finite-dimensional, the constraint qualification (4.4) is implied by Proposition 4.2(v). Both g_1 and h_1 are norms and therefore have full domain, so (4.8) is satisfied for $k = 1$. For $k = 2$, g_2 is a norm and has full domain while $h_2 = \iota_{\{0\}}$, so $0 \in \operatorname{ri}\operatorname{dom} h_2$ and hence (4.8) holds for $k = 2$.

To apply Proposition 4.2, the primal problem must have a solution. Here existence of a solution follows from the compactness of C [2, Proposition 11.14(ii)].

(a) Original image (b) Degraded image (c) Solution to (5.4)

FIGURE 1. Original, blurred, and restored images.

TABLE 1. Quantitative measurements of performance.

Method	Peak signal-to-noise ratio	Structural similarity index
Blurred and noisy image	20.32 dB	0.545
Restoration	25.42 dB	0.803

5.4. **Numerical experiments.** Experiments are made on a quad-core 1.60 GHz Intel i7 laptop, with the algorithms and analysis implemented using the free software package GNU Octave [16]. The authors are grateful for the support of the Octave development community.

Note that in (4.20), the update for $s_{1,1,n}$ and for $p_{1,k+1,n}$ both involve $L_k^* v_{k,n}$, hence it is possible to prevent redundant computation by storing $L_k^* v_{k,n}$ as a temporary variable. Similarly, the updates for $q_{1,1,n}$ and $q_{1,k+1,n}$ both involve $L_k^* p_{2,k,n}$, which can also be stored as a temporary variable for savings. With this approach, each L_k and M_k is applied exactly twice per iteration, and each L_k^* and M_k^* is also applied exactly twice. The restored image is displayed in Fig. 1 (c). The algorithm uses all variables initialized to 0. The values of the parameters are as follows: $\alpha = \beta = \gamma = 10^{-2}$. Figures of merit relative to these experiments are provided in Table 1. Given a reference image x and an estimate $\overline{x} = (\overline{\xi}_i)_{1 \leqslant i \leqslant N}$, the peak signal-to-noise ratio (PSNR), a standard measure of image quality, is defined by

$$\mathrm{PSNR}_x(\overline{x}) = 10 \log_{10} \left(\frac{N \max_{1 \leqslant i \leqslant N} \xi_i^2}{\sum_{i=1}^N (\xi_i - \overline{\xi}_i)^2} \right) \tag{5.17}$$

and reported in units of decibels (dB). The structural similarity index attempts to quantify human visual response to images; details can be found in [26].

ACKNOWLEDGMENT.

The authors thank J.-C. Pesquet for the implementation of the wavelet transform.

REFERENCES

[1] H. C. Andrews and B. R. Hunt, *Digital Image Restoration*, Prentice-Hall, NJ, 1977.

[2] H. H. Bauschke and P. L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, Springer, New York, 2011.
[3] R. I. Boţ, *Conjugate Duality in Convex Optimization*, Lecture Notes in Economics and Mathematical Systems, vol. 637, Springer, New York, 2010.
[4] R. I. Boţ, E. R. Csetnek and E. Nagy, *Solving systems of monotone inclusions via primal-dual splitting techniques*, Taiwan. J. Math., to appear.
[5] R. I. Boţ and C. Hendrich, *Convex risk minimization via proximal splitting methods*, `http://arxiv.org/abs/1304.7694`.
[6] L. M. Briceño-Arias and P. L. Combettes, *A monotone+skew splitting model for composite monotone inclusions in duality*, SIAM J. Optim. **21** (2011), 1230–1250.
[7] A. Chambolle and P.-L. Lions, *Image recovery via total variation minimization and related problems*, Numer. Math. **76** (1997), 167–188.
[8] A. Cohen, I. Daubechies and J.-C. Feauveau, *Biorthogonal bases of compactly supported wavelets*, Comm. Pure Appl. Math. **45** (1992), 485–560.
[9] P. L. Combettes, *Solving monotone inclusions via compositions of nonexpansive averaged operators*, Optimization **53** (2004), 475–504.
[10] P. L. Combettes, *Systems of structured monotone inclusions: duality, algorithms, and applications*, SIAM J. Optim., to appear.
[11] P. L. Combettes and J.-C. Pesquet, *Proximal splitting methods in signal processing*, in Fixed-Point Algorithms for Inverse Problems in Science and Engineering, (H. H. Bauschke et al., eds), Springer, New York, 2011, pp. 185–212.
[12] P. L. Combettes and J.-C. Pesquet, *Primal-dual splitting algorithm for solving inclusions with mixtures of composite, Lipschitzian, and parallel-sum type monotone operators*, Set-Valued Var. Anal. **20** (2012), 307–330.
[13] P. L. Combettes and B. C. Vũ, *Variable metric forward-backward splitting with applications to monotone inclusions in duality*, Optimization, to appear.
[14] P. L. Combettes and V. R. Wajs, *Signal recovery by proximal forward-backward splitting*, Multiscale Model. Simul. **4** (2005), 1168–1200.
[15] J. Eckstein and D. P. Bertsekas, *On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators*, Math. Programming **55** (1992), 293–318.
[16] GNU Octave scientific package, `http://www.octave.org`.
[17] P. L. Lions and B. Mercier, *Splitting algorithms for the sum of two nonlinear operators*, SIAM J. Numer. Anal. **16** (1979), 964–979.
[18] B. Mercier, *Topics in Finite Element Solution of Elliptic Problems* (Lectures on Mathematics, no. 63). Tata Institute of Fundamental Research, Bombay, 1979.
[19] N. Pustelnik, C. Chaux and J.-C. Pesquet, *Parallel proximal algorithm for image restoration using hybrid regularization*, IEEE Trans. Image Process. **20** (2011), 2450–2462.
[20] L. I. Rudin, S. Osher and E. Fatemi, *Nonlinear total variation based noise removal algorithms*, Phys. D **60** (1992), 259–268.
[21] S. Setzer, G. Steidl and T. Teuber, *Infimal convolution regularizations with discrete ℓ_1-type functionals*, Commun. Math. Sci. **9** (2011), 797–827.
[22] S. Simons, *Quadrivariate existence theorems and strong representability*, Optimization **60** (2011), 875–891.
[23] P. Tseng, *Applications of a splitting algorithm to decomposition in convex programming and variational inequalities*, SIAM J. Control Optim. **29** (1991), 119–138.
[24] P. Tseng, *A modified forward-backward splitting method for maximal monotone mappings*, SIAM J. Control Optim. **38** (2000), 431–446.
[25] B. C. Vũ, *A splitting algorithm for dual monotone inclusions involving cocoercive operators*, Adv. Comput. Math. **38** (2013), 667–681.
[26] Z. Wang, A. C. Bovik, H. R. Sheikh and E. P. Simoncelli, *Image quality assessment: From error visibility to structural similarity*, IEEE Trans. Image Process. **13** (2004), 600–612.
[27] C. Zălinescu, *Convex Analysis in General Vector Spaces*, World Scientific, River Edge, NJ, 2002.
[28] W. P. Ziemer, *Weakly Differentiable Functions*, Springer, New York, 1989.

Manuscript received May 24, 2013
revised October 7, 2013

Stephen R. Becker
UPMC Université Paris 06, Laboratoire Jacques-Louis Lions – UMR CNRS 7598, 75005 Paris, France
E-mail address: becker@ljll.math.upmc.fr

Patrick L. Combettes
UPMC Université Paris 06, Laboratoire Jacques-Louis Lions – UMR CNRS 7598, 75005 Paris, France; and
Department of Mathematics, King Abdulaziz University KAU, Jeddah, Saudi Arabia
E-mail address: plc@math.jussieu.fr