A Douglas-Rachford type primal-dual method for solving inclusions with mixtures of composite and parallel-sum type monotone operators

Radu Ioan Boţ *   Christopher Hendrich †

October 2, 2013

Abstract. In this paper we propose two different primal-dual splitting algorithms for solving inclusions involving mixtures of composite and parallel-sum type monotone operators which rely on an inexact Douglas-Rachford splitting method, however applied in different underlying Hilbert spaces. Most importantly, the algorithms allow to process the bounded linear operators and the set-valued operators occurring in the formulation of the monotone inclusion problem separately at each iteration, the latter being individually accessed via their resolvents. The performance of the primal-dual algorithms is emphasized via some numerical experiments on location and image denoising problems.

Keywords. Douglas-Rachford splitting, monotone inclusion, Fenchel duality, convex optimization

AMS subject classification. 90C25, 90C46, 47A52

1 Introduction and preliminaries

In applied mathematics, a wide range of convex optimization problems such as single- or multifacility location problems, support vector machine problems for classification and regression, portfolio optimization problems as well as signal and image processing problems, all of them likely possessing nondifferentiable convex objectives, can be reduced to the solving of inclusions involving mixtures of monotone set-valued operators.

In this article we propose two different primal-dual iterative error-tolerant methods for solving inclusions with mixtures of composite and parallel-sum type monotone operators. Both algorithms rely on the inexact Douglas-Rachford algorithm (cf. [10, 11]), but still differ clearly from each other. An important feature of the two approaches and, simultaneously, an advantage over many existing methods is their capability of processing the set-valued operators separately via their resolvents, while the bounded linear operators are accessed via explicit forward steps on their own or on their adjoints. The resolvents of the maximally monotone operators are not always available in closed form expressions, fact which motivates the inexact versions of the algorithms, where implementation errors in the shape of summable sequences are allowed.

*University of Vienna, Faculty of Mathematics, Oskar-Morgenstern-Platz 1, A-1090 Vienna, Austria, e-mail: radu.bot@univie.ac.at. Research partially supported by DFG (German Research Foundation), project BO 2516/4-1.

†Chemnitz University of Technology, Department of Mathematics, D-09107 Chemnitz, Germany, e-mail: christopher.hendrich@mathematik.tu-chemnitz.de. Research supported by a Graduate Fellowship of the Free State Saxony, Germany.

The methods in this article are also perfectly parallelizable since the majority of their steps can be executed independently. Furthermore, when applied to subdifferential operators of proper, convex and lower semicontinuous functions, the solving of the monotone inclusion problems is, under appropriate qualification conditions (cf. [3, 5]), equivalent with finding optimal solutions to primal-dual pairs of convex optimization problems. The considered formulation also captures various types of primal convex optimization problems and corresponding conjugate duals appearing in wide ranges of applications. The resolvents of subdifferentials of proper, convex and lower semicontinuous functions are the proximal point mappings of these and are known to assume closed form expressions in many cases of interest.

Recent research (see [4, 6–8, 12, 19]) has shown that structured problems dealing with monotone inclusions can be efficiently solved via primal-dual splitting approaches. In [12], the problem involving sums of set-valued, composed, Lipschitzian and parallel-sum type monotone operators was decomposed and solved via an inexact Tseng algorithm having foward-backward-forward characteristics in a product Hilbert space. On the other hand, in [19], instead of Lipschitzian operators, the author has assumed cocoercive operators and solved the resulting problem with an inexact forward-backward algorithm. Thus, our methods can be seen as a continuation of these ideas, this time by making use of the inexact Douglas-Rachford method. Another primal-dual method relying on the same fundamental splitting algorithm is considered in [13] in the context of solving minimization problems having as objective the sum of two proper, convex and lower semicontinuous functions, one of them being composed with a bounded linear operator.

Due to the nature of Douglas-Rachford splitting, we will neither assume Lipschitz continuity nor cocoercivity for any of the operators present in the formulation of the monotone inclusion problem. The resulting drawback of not having operators which can be processed explicitly via forward steps is compensated by the advantage of allowing general maximal monotone operators in the parallel-sums, fact which relaxes the working hypotheses in [12, 19].

The article is organized as follows. In the remaining of this section we introduce the framework we work within and some necessary notations. The splitting algorithms and corresponding weak and strong convergence statements are subject of Section 2 while Section 3 is concerned with the application of the two methods to convex minimization problems. Finally, in Section 4 we make some numerical experiments and evaluate the obtained results.

We are considering the real Hilbert spaces $\mathcal{H}$ and $\mathcal{G}_i$ endowed with the *inner product* $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ and $\langle \cdot, \cdot \rangle_{\mathcal{G}_i}$ and associated *norm* $\|\cdot\|_{\mathcal{H}} = \sqrt{\langle \cdot, \cdot \rangle_{\mathcal{H}}}$ and $\|\cdot\|_{\mathcal{G}_i} = \sqrt{\langle \cdot, \cdot \rangle_{\mathcal{G}_i}}$, $i = 1, \ldots, m$, respectively. The symbols $\rightharpoonup$ and $\rightarrow$ denote weak and strong convergence, respectively, $\mathbb{R}_{++}$ denotes the set of strictly positive real numbers and $\mathbb{R}_+ = \mathbb{R}_{++} \cup \{0\}$. By $B(0, r)$ we denote the closed ball with center 0 and radius $r \in \mathbb{R}_{++}$. For a function $f : \mathcal{H} \to \overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$ we denote by $\operatorname{dom} f := \{x \in \mathcal{H} : f(x) < +\infty\}$ its *effective domain* and call f *proper* if $\operatorname{dom} f \neq \varnothing$ and $f(x) > -\infty$ for all $x \in \mathcal{H}$. Let be

$$\Gamma(\mathcal{H}) := \{f : \mathcal{H} \to \overline{\mathbb{R}} : f \text{ is proper, convex and lower semicontinuous}\}.$$

The *conjugate function* of f is $f^* : \mathcal{H} \to \overline{\mathbb{R}}$, $f^*(p) = \sup\{\langle p, x \rangle - f(x) : x \in \mathcal{H}\}$ for all $p \in \mathcal{H}$ and, if $f \in \Gamma(\mathcal{H})$, then $f^* \in \Gamma(\mathcal{H})$, as well. The *(convex) subdifferential* of $f : \mathcal{H} \to \overline{\mathbb{R}}$ at $x \in \mathcal{H}$ is the set $\partial f(x) = \{p \in \mathcal{H} : f(y) - f(x) \geq \langle p, y - x \rangle \ \forall y \in \mathcal{H}\}$, if $f(x) \in \mathbb{R}$, and is taken to be the empty set, otherwise. For a linear continuous operator

$L_i : \mathcal{H} \to \mathcal{G}_i$, the operator $L_i^* : \mathcal{G}_i \to \mathcal{H}$, defined via $\langle L_i x, y \rangle = \langle x, L_i^* y \rangle$ for all $x \in \mathcal{H}$ and all $y \in \mathcal{G}_i$, denotes its *adjoint*, for $i \in \{1, \dots, m\}$.

Having two functions $f, g : \mathcal{H} \to \overline{\mathbb{R}}$, their *infimal convolution* is defined by $f \,\Box\, g : \mathcal{H} \to \overline{\mathbb{R}}$, $(f \,\Box\, g)(x) = \inf_{y \in \mathcal{H}} \{f(y) + g(x - y)\}$ for all $x \in \mathcal{H}$, being a convex function when f and g are convex.

Let $M : \mathcal{H} \to 2^{\mathcal{H}}$ be a set-valued operator. We denote by $\operatorname{zer} M = \{x \in \mathcal{H} : 0 \in Mx\}$ its set of *zeros*, by $\operatorname{fix} M = \{x \in \mathcal{H} : x \in Mx\}$ its set of *fixed points*, by $\operatorname{gra} M = \{(x, u) \in \mathcal{H} \times \mathcal{H} : u \in Mx\}$ its *graph* and by $\operatorname{ran} M = \{u \in \mathcal{H} : \exists x \in \mathcal{H},\ u \in Mx\}$ its *range*. The *inverse* of M is $M^{-1} : \mathcal{H} \to 2^{\mathcal{H}}$, $u \mapsto \{x \in \mathcal{H} : u \in Mx\}$. We say that the operator M is *monotone* if $\langle x - y, u - v \rangle \geq 0$ for all $(x, u), (y, v) \in \operatorname{gra} M$ and it is said to be *maximally monotone* if there exists no monotone operator $M' : \mathcal{H} \to 2^{\mathcal{H}}$ such that $\operatorname{gra} M'$ properly contains $\operatorname{gra} M$. The operator M is said to be *uniformly monotone* with modulus $\phi_M : \mathbb{R}_+ \to [0, +\infty]$ if ϕ_M is increasing, vanishes only at 0, and $\langle x - y, u - v \rangle \geq \phi_M (\|x - y\|)$ for all $(x, u), (y, v) \in \operatorname{gra} M$.

The *resolvent* and the *reflected resolvent* of an operator $M : \mathcal{H} \to 2^{\mathcal{H}}$ are

$$J_M = (\operatorname{Id} + M)^{-1} \text{ and } R_M = 2J_M - \operatorname{Id},$$

respectively, the operator Id denoting the identity on the underlying Hilbert space. When M is maximally monotone, its resolvent (and, consequently, its reflected resolvent) is a single-valued operator and, by [1, Proposition 23.18], we have for $\gamma \in \mathbb{R}_{++}$

$$\operatorname{Id} = J_{\gamma M} + \gamma J_{\gamma^{-1} M^{-1}} \circ \gamma^{-1} \operatorname{Id}. \tag{1.1}$$

Moreover, for $f \in \Gamma(\mathcal{H})$ and $\gamma \in \mathbb{R}_{++}$ the subdifferential $\partial(\gamma f)$ is maximally monotone (cf. [20, Theorem 3.2.8]) and it holds $J_{\gamma \partial f} = (\operatorname{Id} + \gamma \partial f)^{-1} = \operatorname{Prox}_{\gamma f}$. Here, $\operatorname{Prox}_{\gamma f}(x)$ denotes the *proximal point* of γf at $x \in \mathcal{H}$ representing the unique optimal solution of the optimization problem

$$\inf_{y \in \mathcal{H}} \left\{ \gamma f(y) + \frac{1}{2} \|y - x\|^2 \right\}. \tag{1.2}$$

In this particular situation (1.1) becomes *Moreau's decomposition formula*

$$\operatorname{Id} = \operatorname{Prox}_{\gamma f} + \gamma \operatorname{Prox}_{\gamma^{-1} f^*} \circ \gamma^{-1} \operatorname{Id}. \tag{1.3}$$

When $\Omega \subseteq \mathcal{H}$ is a nonempty, convex and closed set, the function $\delta_\Omega : \mathcal{H} \to \overline{\mathbb{R}}$, defined by $\delta_\Omega(x) = 0$ for $x \in \Omega$ and $\delta_\Omega(x) = +\infty$, otherwise, denotes the *indicator function* of the set Ω. For each $\gamma > 0$ the proximal point of $\gamma \delta_\Omega$ at $x \in \mathcal{H}$ is nothing else than

$$\operatorname{Prox}_{\gamma \delta_\Omega}(x) = \operatorname{Prox}_{\delta_\Omega}(x) = \mathcal{P}_\Omega(x) = \operatorname*{arg\,min}_{y \in \Omega} \frac{1}{2} \|y - x\|^2,$$

which is the projection of x on Ω.

The *sum* and the *parallel sum* of two set-valued operators $M_1, M_2 : \mathcal{H} \to 2^{\mathcal{H}}$ are defined as $M_1 + M_2 : \mathcal{H} \to 2^{\mathcal{H}}, (M_1 + M_2)(x) = M_1(x) + M_2(x)\ \forall x \in \mathcal{H}$ and

$$M_1 \,\Box\, M_2 : \mathcal{H} \to 2^{\mathcal{H}}, M_1 \,\Box\, M_2 = \left(M_1^{-1} + M_2^{-1} \right)^{-1},$$

respectively. If M_1 and M_2 are monotone, then $M_1 + M_2$ and $M_1 \,\Box\, M_2$ are monotone, too. However, if M_1 and M_2 are maximally monotone, this property is in general neither for $M_1 + M_2$ nor for $M_1 \,\Box\, M_2$ true (see [3]).

2 Algorithms and convergence results

Within this section we provide two algorithms together with weak and strong convergence results for the following primal-dual pair of monotone inclusion problems.

Problem 2.1. Let $A : \mathcal{H} \to 2^{\mathcal{H}}$ be a maximally monotone operator and $z \in \mathcal{H}$. Furthermore, for every $i \in \{1, \dots, m\}$, let $r_i \in \mathcal{G}_i$, $B_i : \mathcal{G}_i \to 2^{\mathcal{G}_i}$ and $D_i : \mathcal{G}_i \to 2^{\mathcal{G}_i}$ be maximally monotone operators and $L_i : \mathcal{H} \to \mathcal{G}_i$ a nonzero bounded linear operator. The problem is to solve the primal inclusion

$$\text{find } \overline{x} \in \mathcal{H} \text{ such that } z \in A\overline{x} + \sum_{i=1}^{m} L_i^*(B_i \,\square\, D_i)(L_i\overline{x} - r_i) \tag{2.1}$$

together with the dual inclusion

$$\text{find } \overline{v}_1 \in \mathcal{G}_1, \dots, \overline{v}_m \in \mathcal{G}_m \text{ such that } (\exists x \in \mathcal{H}) \begin{cases} z - \sum_{i=1}^m L_i^*\overline{v}_i \in Ax \\ \overline{v}_i \in (B_i \,\square\, D_i)(L_i x - r_i),\ i = 1, \dots, m. \end{cases} \tag{2.2}$$

We say that $(\overline{x}, \overline{v}_1, \dots, \overline{v}_m) \in \mathcal{H} \times \mathcal{G}_1 \ldots \times \mathcal{G}_m$ is a primal-dual solution to Problem 2.1, if

$$z - \sum_{i=1}^{m} L_i^*\overline{v}_i \in A\overline{x} \text{ and } \overline{v}_i \in (B_i \,\square\, D_i)(L_i\overline{x} - r_i),\ \ i = 1, \dots, m. \tag{2.3}$$

If $(\overline{x}, \overline{v}_1, \dots, \overline{v}_m) \in \mathcal{H} \times \mathcal{G}_1 \ldots \times \mathcal{G}_m$ is a primal-dual solution to Problem 2.1, then $\overline{x}$ is a solution to (2.1) and $(\overline{v}_1, \dots, \overline{v}_m)$ is a solution to (2.2). Notice also that

$$\overline{x} \text{ solves (2.1)} \Leftrightarrow z - \sum_{i=1}^{m} L_i^*(B_i \,\square\, D_i)(L_i\overline{x} - r_i) \in A\overline{x} \Leftrightarrow$$

$$\exists\, \overline{v}_1 \in \mathcal{G}_1, \dots, \overline{v}_m \in \mathcal{G}_m \text{ such that } \begin{cases} z - \sum_{i=1}^m L_i^*\overline{v}_i \in A\overline{x}, \\ \overline{v}_i \in (B_i \,\square\, D_i)(L_i\overline{x} - r_i),\ \ i = 1, \dots, m. \end{cases}$$

Thus, if $\overline{x}$ is a solution to (2.1), then there exists $(\overline{v}_1, \dots, \overline{v}_m) \in \mathcal{G}_1 \times \ldots \mathcal{G}_m$ such that $(\overline{x}, \overline{v}_1, \dots, \overline{v}_m)$ is a primal-dual solution to Problem 2.1 and if $(\overline{v}_1, \dots, \overline{v}_m) \in \mathcal{G}_1 \times \ldots \mathcal{G}_m$ is a solution to (2.2), then there exists $\overline{x} \in \mathcal{H}$ such that $(\overline{x}, \overline{v}_1, \dots, \overline{v}_m)$ is a primal-dual solution to Problem 2.1.

Example 2.1. In Problem 2.1, set $m = 1$, $z = 0$ and $r_1 = 0$, let $\mathcal{G}_1 = \mathcal{G}$, $B_1 = B$, $L_1 = L$, $D_1 : \mathcal{G} \to 2^{\mathcal{G}}$, $D_1(0) = \mathcal{G}$ and $D_1(v) = \varnothing\ \forall v \in \mathcal{G} \setminus \{0\}$, and $A : \mathcal{H} \to 2^{\mathcal{H}}$ and $B : \mathcal{G} \to 2^{\mathcal{G}}$ be the convex subdifferentials of the functions $f \in \Gamma(\mathcal{H})$ and $g \in \Gamma(\mathcal{G})$, respectively. Then, under appropriate qualification conditions (see [3, 5]), to solve the primal inclusion problem (2.1) is equivalent to solve the optimization problem

$$\inf_{x \in \mathcal{H}} \{f(x) + g(Lx)\},$$

while to solve the dual inclusion problem (2.2) is nothing else than to solve its Fenchel-dual problem

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

For more primal-dual pairs of convex optimization problems which are particular instances of (2.1)-(2.2) we refer to [12, 19].

2.1 A first primal-dual algorithm

The first iterative scheme we propose in this paper has the particularity that it accesses the resolvents of A, B_i^{-1} and D_i^{-1}, $i = 1, \ldots, m$, and processes each operator L_i and its adjoint L_i^*, $i = 1, \ldots, m$ two times.

Algorithm 2.1.
Let $x_0 \in \mathcal{H}$, $(v_{1,0}, \ldots, v_{m,0}) \in \mathcal{G}_1 \times \ldots \times \mathcal{G}_m$ and τ and σ_i, $i = 1, \ldots, m$, be strictly positive real numbers such that

$$\tau \sum_{i=1}^{m} \sigma_i \|L_i\|^2 < 4.$$

Furthermore, let $(\lambda_n)_{n\geq 0}$ be a sequence in $(0, 2)$, $(a_n)_{n\geq 0}$ a sequence in $\mathcal{H}$, $(b_{i,n})_{n\geq 0}$ and $(d_{i,n})_{n\geq 0}$ sequences in $\mathcal{G}_i$ for all $i = 1, \ldots, m$ and set

$$(\forall n \geq 0) \quad \left\lfloor \begin{array}{l} p_{1,n} = J_{\tau A}\left(x_n - \frac{\tau}{2}\sum_{i=1}^m L_i^* v_{i,n} + \tau z\right) + a_n \\ w_{1,n} = 2p_{1,n} - x_n \\ \text{For } i = 1, \ldots, m \\ \quad \left\lfloor \begin{array}{l} p_{2,i,n} = J_{\sigma_i B_i^{-1}}\left(v_{i,n} + \frac{\sigma_i}{2} L_i w_{1,n} - \sigma_i r_i\right) + b_{i,n} \\ w_{2,i,n} = 2p_{2,i,n} - v_{i,n} \end{array} \right. \\ z_{1,n} = w_{1,n} - \frac{\tau}{2}\sum_{i=1}^m L_i^* w_{2,i,n} \\ x_{n+1} = x_n + \lambda_n(z_{1,n} - p_{1,n}) \\ \text{For } i = 1, \ldots, m \\ \quad \left\lfloor \begin{array}{l} z_{2,i,n} = J_{\sigma_i D_i^{-1}}\left(w_{2,i,n} + \frac{\sigma_i}{2} L_i(2z_{1,n} - w_{1,n})\right) + d_{i,n} \\ v_{i,n+1} = v_{i,n} + \lambda_n(z_{2,i,n} - p_{2,i,n}). \end{array} \right. \end{array} \right. \tag{2.4}$$

Theorem 2.1. *For Problem 2.1 assume that*

$$z \in \operatorname{ran}\left(A + \sum_{i=1}^{m} L_i^*(B_i \,\square\, D_i)(L_i \cdot - r_i)\right) \tag{2.5}$$

and consider the sequences generated by Algorithm 2.1.

(i) If

$$\sum_{n=0}^{+\infty} \lambda_n \|a_n\|_{\mathcal{H}} < +\infty, \quad \sum_{n=0}^{+\infty} \lambda_n(\|d_{i,n}\|_{\mathcal{G}_i} + \|b_{i,n}\|_{\mathcal{G}_i}) < +\infty, \;\; i = 1, \ldots, m,$$

and $\sum_{n=0}^{+\infty} \lambda_n(2 - \lambda_n) = +\infty$, *then*

(a) $(x_n, v_{1,n}, \ldots, v_{m,n})_{n\geq 0}$ *converges weakly to an element* $(\overline{x}, \overline{v}_1, \ldots, \overline{v}_m) \in \mathcal{H} \times \mathcal{G}_1 \times \ldots \times \mathcal{G}_m$ *such that, when setting*

$$\overline{p}_1 = J_{\tau A}\left(\overline{x} - \frac{\tau}{2}\sum_{i=1}^{m} L_i^* \overline{v}_i + \tau z\right),$$
$$\text{and } \overline{p}_{2,i} = J_{\sigma_i B_i^{-1}}\left(\overline{v}_i + \frac{\sigma_i}{2} L_i(2\overline{p}_1 - \overline{x}) - \sigma_i r_i\right), \;\; i = 1, \ldots, m,$$

the element $(\overline{p}_1, \overline{p}_{2,1}, \ldots, \overline{p}_{2,m})$ *is a primal-dual solution to Problem 2.1.*

(b) $\lambda_n(z_{1,n} - p_{1,n}) \to 0$ $(n \to +\infty)$ *and* $\lambda_n(z_{2,i,n} - p_{2,i,n}) \to 0$ $(n \to +\infty)$ *for* $i = 1, \ldots, m$.

(c) whenever $\mathcal{H}$ and $\mathcal{G}_i$, $i = 1, \ldots, m$, are finite-dimensional Hilbert spaces, $a_n \to 0$ $(n \to +\infty)$ and $b_{i,n} \to 0$ $(n \to +\infty)$ for $i = 1, \ldots, m$, then $(p_{1,n}, p_{2,1,n}, \ldots, p_{2,m,n})_{n \geq 0}$ converges to a primal-dual solution of Problem 2.1.

(ii) If

$$\sum_{n=0}^{+\infty} \|a_n\|_{\mathcal{H}} < +\infty, \quad \sum_{n=0}^{+\infty} (\|d_{i,n}\|_{\mathcal{G}_i} + \|b_{i,n}\|_{\mathcal{G}_i}) < +\infty, \ i = 1, \ldots, m, \quad \inf_{n \geq 0} \lambda_n > 0$$

$$\text{and } A \text{ and } B_i^{-1}, \ i = 1, \ldots, m, \text{ are uniformly monotone,}$$

then $(p_{1,n}, p_{2,1,n}, \ldots, p_{2,m,n})_{n \geq 0}$ converges strongly to the unique primal-dual solution of Problem 2.1.

Proof. Consider the Hilbert space $\boldsymbol{\mathcal{G}} = \mathcal{G}_1 \times \ldots \times \mathcal{G}_m$ endowed with inner product and associated norm defined, for $\boldsymbol{v} = (v_1, \ldots, v_m)$, $\boldsymbol{q} = (q_1, \ldots, q_m) \in \boldsymbol{\mathcal{G}}$, as

$$\langle \boldsymbol{v}, \boldsymbol{q} \rangle_{\boldsymbol{\mathcal{G}}} = \sum_{i=1}^{m} \langle v_i, q_i \rangle_{\mathcal{G}_i} \text{ and } \|\boldsymbol{v}\|_{\boldsymbol{\mathcal{G}}} = \sqrt{\sum_{i=1}^{m} \|v_i\|_{\mathcal{G}_i}^2}, \tag{2.6}$$

respectively. Furthermore, let $\boldsymbol{\mathcal{K}} = \mathcal{H} \times \boldsymbol{\mathcal{G}}$ be the Hilbert space endowed with inner product and associated norm defined, for $(x, \boldsymbol{v}), (y, \boldsymbol{q}) \in \boldsymbol{\mathcal{K}}$, as

$$\langle (x, \boldsymbol{v}), (y, \boldsymbol{q}) \rangle_{\boldsymbol{\mathcal{K}}} = \langle x, y \rangle_{\mathcal{H}} + \langle \boldsymbol{v}, \boldsymbol{q} \rangle_{\boldsymbol{\mathcal{G}}} \text{ and } \|(x, \boldsymbol{v})\|_{\boldsymbol{\mathcal{K}}} = \sqrt{\|x\|_{\mathcal{H}}^2 + \|\boldsymbol{v}\|_{\boldsymbol{\mathcal{G}}}^2}, \tag{2.7}$$

respectively. Consider the set-valued operator

$$\boldsymbol{M} : \boldsymbol{\mathcal{K}} \to 2^{\boldsymbol{\mathcal{K}}}, \quad (x, v_1, \ldots, v_m) \mapsto (-z + Ax, r_1 + B_1^{-1} v_1, \ldots, r_m + B_m^{-1} v_m),$$

which is maximally monotone, since A and B_i, $i = 1, \ldots, m$, are maximally monotone (cf. [1, Proposition 20.22 and Proposition 20.23]) and the bounded linear operator

$$\boldsymbol{S} : \boldsymbol{\mathcal{K}} \to \boldsymbol{\mathcal{K}}, \quad (x, v_1, \ldots, v_m) \mapsto \left(\sum_{i=1}^{m} L_i^* v_i, -L_1 x, \ldots, -L_m x \right),$$

which proves to be skew (i. e. $\boldsymbol{S}^* = -\boldsymbol{S}$) and hence maximally monotone (cf. [1, Example 20.30]). Further, consider the set-valued operator

$$\boldsymbol{Q} : \boldsymbol{\mathcal{K}} \to 2^{\boldsymbol{\mathcal{K}}}, \quad (x, v_1, \ldots, v_m) \mapsto \left(0, D_1^{-1} v_1, \ldots, D_m^{-1} v_m \right),$$

which is maximally monotone, as well, since D_i is maximally monotone for $i = 1, \ldots, m$. Therefore, since $\operatorname{dom} \boldsymbol{S} = \boldsymbol{\mathcal{K}}$, both $\frac{1}{2}\boldsymbol{S} + \boldsymbol{Q}$ and $\frac{1}{2}\boldsymbol{S} + \boldsymbol{M}$ are maximally monotone (cf. [1, Corollary 24.4(i)]). On the other hand, according to [12, Eq. (3.12)], it holds (2.5) $\Leftrightarrow$ $\operatorname{zer}(\boldsymbol{M} + \boldsymbol{S} + \boldsymbol{Q}) \neq \varnothing$, while [12, Eq. (3.21) and (3.22)] yield

$$\begin{aligned} &(x, v_1, \ldots, v_m) \in \operatorname{zer}(\boldsymbol{M} + \boldsymbol{S} + \boldsymbol{Q}) \\ \Rightarrow &(x, v_1, \ldots, v_m) \text{ is a primal-dual solution to Problem 2.1.} \end{aligned} \tag{2.8}$$

Finally, we introduce the bounded linear operator

$$\boldsymbol{V} : \boldsymbol{\mathcal{K}} \to \boldsymbol{\mathcal{K}}, \quad (x, v_1, \ldots, v_m) \quad \mapsto \left(\frac{x}{\tau} - \frac{1}{2} \sum_{i=1}^{m} L_i^* v_i, \frac{v_1}{\sigma_1} - \frac{1}{2} L_1 x, \ldots, \frac{v_m}{\sigma_m} - \frac{1}{2} L_m x \right).$$

It is a simple calculation to prove that $\boldsymbol{V}$ is self-adjoint, i. e. $\boldsymbol{V}^* = \boldsymbol{V}$. Furthermore, the operator $\boldsymbol{V}$ is ρ-strongly positive for

$$\rho = \left(1 - \frac{1}{2}\sqrt{\tau \sum_{i=1}^m \sigma_i \|L_i\|^2}\right) \min\left\{\frac{1}{\tau}, \frac{1}{\sigma_1}, \ldots, \frac{1}{\sigma_m}\right\},$$

which is a positive real number due to the assumption

$$\tau \sum_{i=1}^m \sigma_i \|L_i\|^2 < 4 \tag{2.9}$$

made in Algorithm 2.1. Indeed, using that $2ab \leq \alpha a^2 + \frac{b^2}{\alpha}$ for any a, $b \in \mathbb{R}$ and any $\alpha \in \mathbb{R}_{++}$, it yields for each $i = 1, \ldots, m$

$$2\|L_i\|\|x\|_{\mathcal{H}}\|v_i\|_{\mathcal{G}_i} \leq \frac{\sigma_i \|L_i\|^2}{\sqrt{\tau \sum_{i=1}^m \sigma_i \|L_i\|^2}} \|x\|_{\mathcal{H}}^2 + \frac{\sqrt{\tau \sum_{i=1}^m \sigma_i \|L_i\|^2}}{\sigma_i} \|v_i\|_{\mathcal{G}_i}^2 \tag{2.10}$$

and, consequently, for each $\boldsymbol{x} = (x, v_1, \ldots, v_m) \in \boldsymbol{\mathcal{K}}$, it follows that

$$\begin{aligned}
\langle \boldsymbol{x}, \boldsymbol{V}\boldsymbol{x} \rangle_{\boldsymbol{\mathcal{K}}} &= \frac{\|x\|_{\mathcal{H}}^2}{\tau} + \sum_{i=1}^m \frac{\|v_i\|_{\mathcal{G}_i}^2}{\sigma_i} - \sum_{i=1}^m \langle L_i x, v_i \rangle_{\mathcal{G}_i} \\
&\geq \frac{\|x\|_{\mathcal{H}}^2}{\tau} + \sum_{i=1}^m \frac{\|v_i\|_{\mathcal{G}_i}^2}{\sigma_i} - \sum_{i=1}^m \|L_i\|\|x\|_{\mathcal{H}}\|v_i\|_{\mathcal{G}_i} \\
&\overset{(2.10)}{\geq} \left(1 - \frac{1}{2}\sqrt{\tau \sum_{i=1}^m \sigma_i \|L_i\|^2}\right) \left(\frac{\|x\|_{\mathcal{H}}^2}{\tau} + \sum_{i=1}^m \frac{\|v_i\|_{\mathcal{G}_i}^2}{\sigma_i}\right) \\
&\geq \left(1 - \frac{1}{2}\sqrt{\tau \sum_{i=1}^m \sigma_i \|L_i\|^2}\right) \min\left\{\frac{1}{\tau}, \frac{1}{\sigma_1}, \ldots, \frac{1}{\sigma_m}\right\} \|\boldsymbol{x}\|_{\boldsymbol{\mathcal{K}}}^2 \\
&= \rho \|\boldsymbol{x}\|_{\boldsymbol{\mathcal{K}}}^2.
\end{aligned} \tag{2.11}$$

Since $\boldsymbol{V}$ is ρ-strongly positive, we have $\operatorname{cl}(\operatorname{ran} \boldsymbol{V}) = \operatorname{ran} \boldsymbol{V}$ (cf. [1, Fact 2.19]), $\operatorname{zer} \boldsymbol{V} = \{0\}$ and, as $(\operatorname{ran} \boldsymbol{V})^\perp = \operatorname{zer} \boldsymbol{V}^* = \operatorname{zer} \boldsymbol{V} = \{0\}$ (see, for instance, [1, Fact 2.18]), it holds $\operatorname{ran} \boldsymbol{V} = \boldsymbol{\mathcal{K}}$. Consequently, $\boldsymbol{V}^{-1}$ exists and $\|\boldsymbol{V}^{-1}\| \leq \frac{1}{\rho}$.

The algorithmic scheme (2.4) is equivalent to

$$(\forall n \geq 0) \quad \left\lfloor \begin{array}{l}
\frac{x_n - p_{1,n}}{\tau} - \frac{1}{2}\sum_{i=1}^m L_i^* v_{i,n} \in A(p_{1,n} - a_n) - z - \frac{a_n}{\tau} \\
w_{1,n} = 2p_{1,n} - x_n \\
\text{For } i = 1, \ldots, m \\
\left\lfloor \begin{array}{l}
\frac{v_{i,n} - p_{2,i,n}}{\sigma_i} - \frac{1}{2}L_i(x_n - p_{1,n}) \in -\frac{1}{2}L_i p_{1,n} + B_i^{-1}(p_{2,i,n} - b_{i,n}) + r_i - \frac{b_{i,n}}{\sigma_i} \\
w_{2,i,n} = 2p_{2,i,n} - v_{i,n}
\end{array} \right. \\
\frac{w_{1,n} - z_{1,n}}{\tau} - \frac{1}{2}\sum_{i=1}^m L_i^* w_{2,i,n} = 0 \\
x_{n+1} = x_n + \lambda_n(z_{1,n} - p_{1,n}) \\
\text{For } i = 1, \ldots, m \\
\left\lfloor \begin{array}{l}
\frac{w_{2,i,n} - z_{2,i,n}}{\sigma_i} - \frac{1}{2}L_i(w_{1,n} - z_{1,n}) \in -\frac{1}{2}L_i z_{1,n} + D_i^{-1}(z_{2,i,n} - d_{i,n}) - \frac{d_{i,n}}{\sigma_i} \\
v_{i,n+1} = v_{i,n} + \lambda_n(z_{2,i,n} - p_{2,i,n}).
\end{array} \right.
\end{array} \right. \tag{2.12}$$

We introduce for every $n \geq 0$ the following notations:

$$\begin{cases} \boldsymbol{x}_n = (x_n, v_{1,n}, \dots, v_{m,n}) \\ \boldsymbol{y}_n = (p_{1,n}, p_{2,1,n}, \dots, p_{2,m,n}) \\ \boldsymbol{w}_n = (w_{1,n}, w_{2,1,n}, \dots, w_{2,m,n}) \\ \boldsymbol{z}_n = (z_{1,n}, z_{2,1,n}, \dots, z_{2,m,n}) \end{cases} \text{ and } \begin{cases} \boldsymbol{d}_n = (0, d_{1,n}, \dots, d_{m,n}) \\ \boldsymbol{d}_n^\sigma = (0, \frac{d_{1,n}}{\sigma_1}, \dots, \frac{d_{m,n}}{\sigma_m}) \\ \boldsymbol{b}_n = (a_n, b_{1,n}, \dots, b_{m,n}) \\ \boldsymbol{b}_n^\sigma = (\frac{a_n}{\tau}, \frac{b_{1,n}}{\sigma_1}, \dots, \frac{b_{m,n}}{\sigma_m}) \end{cases}. \tag{2.13}$$

The scheme (2.12) can equivalently be written in the form

$$(\forall n \geq 0) \left\lfloor \begin{array}{l} \boldsymbol{V}(\boldsymbol{x}_n - \boldsymbol{y}_n) \in \left(\frac{1}{2}\boldsymbol{S} + \boldsymbol{M}\right)(\boldsymbol{y}_n - \boldsymbol{b}_n) + \frac{1}{2}\boldsymbol{S}\boldsymbol{b}_n - \boldsymbol{b}_n^\sigma \\ \boldsymbol{w}_n = 2\boldsymbol{y}_n - \boldsymbol{x}_n \\ \boldsymbol{V}(\boldsymbol{w}_n - \boldsymbol{z}_n) \in \left(\frac{1}{2}\boldsymbol{S} + \boldsymbol{Q}\right)(\boldsymbol{z}_n - \boldsymbol{d}_n) + \frac{1}{2}\boldsymbol{S}\boldsymbol{d}_n - \boldsymbol{d}_n^\sigma \\ \boldsymbol{x}_{n+1} = \boldsymbol{x}_n + \lambda_n(\boldsymbol{z}_n - \boldsymbol{y}_n). \end{array} \right. \tag{2.14}$$

We set for every $n \geq 0$

$$\begin{aligned} \boldsymbol{e}_n^b &= \boldsymbol{V}^{-1}\left(\left(\frac{1}{2}\boldsymbol{S} + \boldsymbol{V}\right)\boldsymbol{b}_n - \boldsymbol{b}_n^\sigma\right) \\ \boldsymbol{e}_n^d &= \boldsymbol{V}^{-1}\left(\left(\frac{1}{2}\boldsymbol{S} + \boldsymbol{V}\right)\boldsymbol{d}_n - \boldsymbol{d}_n^\sigma\right). \end{aligned} \tag{2.15}$$

Next we introduce the Hilbert space $\boldsymbol{\mathcal{K}}_{\boldsymbol{V}}$ with inner product and norm respectively defined, for $\boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{\mathcal{K}}$, as

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle_{\boldsymbol{\mathcal{K}}_{\boldsymbol{V}}} = \langle \boldsymbol{x}, \boldsymbol{V}\boldsymbol{y} \rangle_{\boldsymbol{\mathcal{K}}} \text{ and } \|\boldsymbol{x}\|_{\boldsymbol{\mathcal{K}}_{\boldsymbol{V}}} = \sqrt{\langle \boldsymbol{x}, \boldsymbol{V}\boldsymbol{x} \rangle_{\boldsymbol{\mathcal{K}}}}, \tag{2.16}$$

respectively. Since $\frac{1}{2}\boldsymbol{S} + \boldsymbol{M}$ and $\frac{1}{2}\boldsymbol{S} + \boldsymbol{Q}$ are maximally monotone on $\boldsymbol{\mathcal{K}}$, the operators

$$\boldsymbol{B} := \boldsymbol{V}^{-1}\left(\frac{1}{2}\boldsymbol{S} + \boldsymbol{M}\right) \text{ and } \boldsymbol{A} := \boldsymbol{V}^{-1}\left(\frac{1}{2}\boldsymbol{S} + \boldsymbol{Q}\right) \tag{2.17}$$

are maximally monotone on $\boldsymbol{\mathcal{K}}_{\boldsymbol{V}}$. Moreover, since $\boldsymbol{V}$ is self-adjoint and ρ-strongly positive, one can easily see that weak and strong convergence in $\boldsymbol{\mathcal{K}}_{\boldsymbol{V}}$ are equivalent with weak and strong convergence in $\boldsymbol{\mathcal{K}}$, respectively.

Now, taking into account (2.14), for every $n \geq 0$, we have

$$\begin{aligned} & \boldsymbol{V}(\boldsymbol{x}_n - \boldsymbol{y}_n) \in \left(\frac{1}{2}\boldsymbol{S} + \boldsymbol{M}\right)(\boldsymbol{y}_n - \boldsymbol{b}_n) + \frac{1}{2}\boldsymbol{S}\boldsymbol{b}_n - \boldsymbol{b}_n^\sigma \\ \Leftrightarrow\ & \boldsymbol{V}\boldsymbol{x}_n \in \left(\boldsymbol{V} + \frac{1}{2}\boldsymbol{S} + \boldsymbol{M}\right)(\boldsymbol{y}_n - \boldsymbol{b}_n) + \left(\frac{1}{2}\boldsymbol{S} + \boldsymbol{V}\right)\boldsymbol{b}_n - \boldsymbol{b}_n^\sigma \\ \Leftrightarrow\ & \boldsymbol{x}_n \in \left(\mathrm{Id} + \boldsymbol{V}^{-1}\left(\frac{1}{2}\boldsymbol{S} + \boldsymbol{M}\right)\right)(\boldsymbol{y}_n - \boldsymbol{b}_n) + \boldsymbol{V}^{-1}\left(\left(\frac{1}{2}\boldsymbol{S} + \boldsymbol{V}\right)\boldsymbol{b}_n - \boldsymbol{b}_n^\sigma\right) \\ \Leftrightarrow\ & \boldsymbol{y}_n = \left(\mathrm{Id} + \boldsymbol{V}^{-1}\left(\frac{1}{2}\boldsymbol{S} + \boldsymbol{M}\right)\right)^{-1}\left(\boldsymbol{x}_n - \boldsymbol{e}_n^b\right) + \boldsymbol{b}_n \\ \Leftrightarrow\ & \boldsymbol{y}_n = (\mathrm{Id} + \boldsymbol{B})^{-1}\left(\boldsymbol{x}_n - \boldsymbol{e}_n^b\right) + \boldsymbol{b}_n \end{aligned} \tag{2.18}$$

and

$$\begin{aligned} & \boldsymbol{V}(\boldsymbol{w}_n - \boldsymbol{z}_n) \in \left(\frac{1}{2}\boldsymbol{S} + \boldsymbol{Q}\right)(\boldsymbol{z}_n - \boldsymbol{d}_n) + \frac{1}{2}\boldsymbol{S}\boldsymbol{d}_n - \boldsymbol{d}_n^\sigma \\ \Leftrightarrow\ & \boldsymbol{z}_n = \left(\mathrm{Id} + \boldsymbol{V}^{-1}\left(\frac{1}{2}\boldsymbol{S} + \boldsymbol{Q}\right)\right)^{-1}\left(\boldsymbol{w}_n - \boldsymbol{e}_n^d\right) + \boldsymbol{d}_n \\ \Leftrightarrow\ & \boldsymbol{z}_n = (\mathrm{Id} + \boldsymbol{A})^{-1}\left(\boldsymbol{w}_n - \boldsymbol{e}_n^d\right) + \boldsymbol{d}_n. \end{aligned} \tag{2.19}$$

Thus, the iterative rules in (2.14) become

$$(\forall n \geq 0) \quad \left\lfloor \begin{array}{l} \boldsymbol{y}_n = J_{\boldsymbol{B}}\left(\boldsymbol{x}_n - \boldsymbol{e}_n^b\right) + \boldsymbol{b}_n \\ \boldsymbol{z}_n = J_{\boldsymbol{A}}\left(2\boldsymbol{y}_n - \boldsymbol{x}_n - \boldsymbol{e}_n^d\right) + \boldsymbol{d}_n \\ \boldsymbol{x}_{n+1} = \boldsymbol{x}_n + \lambda_n(\boldsymbol{z}_n - \boldsymbol{y}_n) \end{array} \right. . \tag{2.20}$$

In addition, we have

$$\operatorname{zer}(\boldsymbol{A}+\boldsymbol{B}) = \operatorname{zer}\left(\boldsymbol{V}^{-1}(\boldsymbol{M}+\boldsymbol{S}+\boldsymbol{Q})\right) = \operatorname{zer}(\boldsymbol{M}+\boldsymbol{S}+\boldsymbol{Q}).$$

By defining for every $n \geq 0$

$$\boldsymbol{\beta}_n = J_{\boldsymbol{B}}\left(\boldsymbol{x}_n - \boldsymbol{e}_n^b\right) - J_{\boldsymbol{B}}(\boldsymbol{x}_n) + \boldsymbol{b}_n \text{ and } \boldsymbol{\alpha}_n = J_{\boldsymbol{A}}\left(2\boldsymbol{y}_n - \boldsymbol{x}_n - \boldsymbol{e}_n^d\right) - J_{\boldsymbol{A}}(2\boldsymbol{y}_n - \boldsymbol{x}_n) + \boldsymbol{d}_n,$$

the iterative scheme (2.20) becomes

$$(\forall n \geq 0) \quad \left\lfloor \begin{array}{l} \boldsymbol{y}_n = J_{\boldsymbol{B}}(\boldsymbol{x}_n) + \boldsymbol{\beta}_n \\ \boldsymbol{z}_n = J_{\boldsymbol{A}}(2\boldsymbol{y}_n - \boldsymbol{x}_n) + \boldsymbol{\alpha}_n \\ \boldsymbol{x}_{n+1} = \boldsymbol{x}_n + \lambda_n(\boldsymbol{z}_n - \boldsymbol{y}_n) \end{array} \right. . \tag{2.21}$$

Thus, it has the structure of an error-tolerant Douglas-Rachford algorithm (see [11]).

(i) The assumptions made on the error sequences yield

$$\sum_{n=0}^{+\infty} \lambda_n \|\boldsymbol{d}_n\|_{\boldsymbol{\mathcal{K}}} < +\infty, \ \sum_{n=0}^{+\infty} \lambda_n \|\boldsymbol{d}_n^\sigma\|_{\boldsymbol{\mathcal{K}}} < +\infty, \sum_{n=0}^{+\infty} \lambda_n \|\boldsymbol{b}_n\|_{\boldsymbol{\mathcal{K}}} < +\infty, \ \sum_{n=0}^{+\infty} \lambda_n \|\boldsymbol{b}_n^\sigma\|_{\boldsymbol{\mathcal{K}}} < +\infty \tag{2.22}$$

and, by the boundedness of $\boldsymbol{V}^{-1}$, $\boldsymbol{S}$ and $\boldsymbol{V}$, it follows

$$\sum_{n=0}^{+\infty} \lambda_n \|\boldsymbol{e}_n^b\|_{\boldsymbol{\mathcal{K}}} < +\infty \text{ and } \sum_{n=0}^{+\infty} \lambda_n \|\boldsymbol{e}_n^d\|_{\boldsymbol{\mathcal{K}}} < +\infty. \tag{2.23}$$

Further, by making use of the nonexpansiveness of the resolvents, the error sequences satisfy

$$\begin{aligned} \sum_{n=0}^{+\infty} \lambda_n \left[\|\boldsymbol{\alpha}_n\|_{\boldsymbol{\mathcal{K}}} + \|\boldsymbol{\beta}_n\|_{\boldsymbol{\mathcal{K}}}\right] \leq & \sum_{n=0}^{+\infty} \lambda_n \Big[\|J_{\boldsymbol{A}}\left(2\boldsymbol{y}_n - \boldsymbol{x}_n - \boldsymbol{e}_n^d\right) - J_{\boldsymbol{A}}(2\boldsymbol{y}_n - \boldsymbol{x}_n)\|_{\boldsymbol{\mathcal{K}}} + \|\boldsymbol{d}_n\|_{\boldsymbol{\mathcal{K}}} \\ & \qquad + \|J_{\boldsymbol{B}}\left(\boldsymbol{x}_n - \boldsymbol{e}_n^b\right) - J_{\boldsymbol{B}}(\boldsymbol{x}_n)\|_{\boldsymbol{\mathcal{K}}} + \|\boldsymbol{b}_n\|_{\boldsymbol{\mathcal{K}}}\Big] \\ \leq & \sum_{n=0}^{+\infty} \lambda_n \left[\|\boldsymbol{e}_n^d\|_{\boldsymbol{\mathcal{K}}} + \|\boldsymbol{d}_n\|_{\boldsymbol{\mathcal{K}}} + \|\boldsymbol{e}_n^b\|_{\boldsymbol{\mathcal{K}}} + \|\boldsymbol{b}_n\|_{\boldsymbol{\mathcal{K}}}\right] < +\infty. \end{aligned}$$

By the linearity and boundedness of $\boldsymbol{V}$ it follows that

$$\sum_{n=0}^{+\infty} \lambda_n \left[\|\boldsymbol{\alpha}_n\|_{\boldsymbol{\mathcal{K}}_{\boldsymbol{V}}} + \|\boldsymbol{\beta}_n\|_{\boldsymbol{\mathcal{K}}_{\boldsymbol{V}}}\right] < +\infty.$$

(i)(a) According to [11, Theorem 2.1(i)(a)] the sequence $(\boldsymbol{x}_n)_{n\geq 0}$ converges weakly in $\boldsymbol{\mathcal{K}}_{\boldsymbol{V}}$ and, consequently, in $\boldsymbol{\mathcal{K}}$ to an element $\overline{\boldsymbol{x}} \in \operatorname{fix}(R_{\boldsymbol{A}} R_{\boldsymbol{B}})$ with $J_{\boldsymbol{B}}\overline{\boldsymbol{x}} \in \operatorname{zer}(\boldsymbol{A}+\boldsymbol{B})$. The claim follows by identifying $J_{\boldsymbol{B}}\overline{\boldsymbol{x}}$ and by noting (2.8).

(i)(b) According to [11, Theorem 2.1(i)(b)] it follows that $(R_{\boldsymbol{A}} R_{\boldsymbol{B}} \boldsymbol{x}_n - \boldsymbol{x}_n) \to 0$ $(n \to +\infty)$. From (2.21) it follows that for every $n \geq 0$

$$\lambda_n(\boldsymbol{z}_n - \boldsymbol{y}_n) = \frac{\lambda_n}{2}\left(R_{\boldsymbol{A}}(R_{\boldsymbol{B}}(\boldsymbol{x}_n) + 2\boldsymbol{\beta}_n) - \boldsymbol{x}_n + 2\boldsymbol{\alpha}_n\right),$$

thus, by taking into consideration the nonexpansiveness of the reflected resolvent and the boundedness of $(\lambda_n)_{n\geq 0}$, it yields

$$\begin{aligned}\|\lambda_n(\boldsymbol{z}_n - \boldsymbol{y}_n)\|_{\boldsymbol{\mathcal{K}_V}} \leq & \frac{\lambda_n}{2}\|R_{\boldsymbol{A}} R_{\boldsymbol{B}} \boldsymbol{x}_n - \boldsymbol{x}_n\|_{\boldsymbol{\mathcal{K}_V}} \\ & + \frac{\lambda_n}{2}\|R_{\boldsymbol{A}}(R_{\boldsymbol{B}}\boldsymbol{x}_n + 2\boldsymbol{\beta}_n) - R_{\boldsymbol{A}}(R_{\boldsymbol{B}}\boldsymbol{x}_n) + 2\boldsymbol{\alpha}_n\|_{\boldsymbol{\mathcal{K}_V}} \\ \leq & \|R_{\boldsymbol{A}} R_{\boldsymbol{B}} \boldsymbol{x}_n - \boldsymbol{x}_n\|_{\boldsymbol{\mathcal{K}_V}} + \lambda_n\left[\|\boldsymbol{\alpha}_n\|_{\boldsymbol{\mathcal{K}_V}} + \|\boldsymbol{\beta}_n\|_{\boldsymbol{\mathcal{K}_V}}\right].\end{aligned}$$

The claim follows by taking into account that $\lambda_n\left[\|\boldsymbol{\alpha}_n\|_{\boldsymbol{\mathcal{K}_V}} + \|\boldsymbol{\beta}_n\|_{\boldsymbol{\mathcal{K}_V}}\right] \to 0$ $(n \to +\infty)$.

(i)(c) As shown in (a), we have that $\boldsymbol{x}_n \to \overline{\boldsymbol{x}} \in \operatorname{fix}(R_{\boldsymbol{A}} R_{\boldsymbol{B}})$ $(n \to +\infty)$ with $J_{\boldsymbol{B}}\overline{\boldsymbol{x}} \in \operatorname{zer}(\boldsymbol{A} + \boldsymbol{B}) = \operatorname{zer}(\boldsymbol{M} + \boldsymbol{S} + \boldsymbol{Q})$. Moreover, by the assumptions and (2.13) we have $\boldsymbol{b}_n \to 0$ $(n \to +\infty)$, hence by (2.15) it holds that $\boldsymbol{e}_n^b \to 0$ $(n \to +\infty)$ and therefore $\boldsymbol{\beta}_n \to 0$ $(n \to +\infty)$. In conclusion, by the continuity of $J_{\boldsymbol{B}}$ and (2.21), we have

$$\boldsymbol{y}_n = J_{\boldsymbol{B}}(\boldsymbol{x}_n) + \boldsymbol{\beta}_n \to J_{\boldsymbol{B}}\overline{\boldsymbol{x}} \in \operatorname{zer}(\boldsymbol{M} + \boldsymbol{S} + \boldsymbol{Q}) \ (n \to +\infty).$$

(ii) The assumptions made on the error sequences yield

$$\sum_{n=0}^{+\infty}\|\boldsymbol{d}_n\|_{\boldsymbol{\mathcal{K}}} < +\infty, \ \sum_{n=0}^{+\infty}\|\boldsymbol{d}_n^{\sigma}\|_{\boldsymbol{\mathcal{K}}} < +\infty, \sum_{n=0}^{+\infty}\|\boldsymbol{b}_n\|_{\boldsymbol{\mathcal{K}}} < +\infty, \ \sum_{n=0}^{+\infty}\|\boldsymbol{b}_n^{\sigma}\|_{\boldsymbol{\mathcal{K}}} < +\infty,$$

thus,

$$\sum_{n=0}^{+\infty}\|\boldsymbol{e}_n^b\|_{\boldsymbol{\mathcal{K}}} < +\infty \text{ and } \sum_{n=0}^{+\infty}\|\boldsymbol{e}_n^d\|_{\boldsymbol{\mathcal{K}}} < +\infty.$$

This implies that

$$\sum_{n=0}^{+\infty}\left[\|\boldsymbol{\alpha}_n\|_{\boldsymbol{\mathcal{K}}} + \|\boldsymbol{\beta}_n\|_{\boldsymbol{\mathcal{K}}}\right] < +\infty$$

which, due to the linearity and boundedness of $\boldsymbol{V}$, yields

$$\sum_{n=0}^{+\infty}\left[\|\boldsymbol{\alpha}_n\|_{\boldsymbol{\mathcal{K}_V}} + \|\boldsymbol{\beta}_n\|_{\boldsymbol{\mathcal{K}_V}}\right] < +\infty.$$

Since A and B_i^{-1}, $i = 1,\ldots,m$, are uniformly monotone, there exist increasing functions $\phi_A : \mathbb{R}_+ \to [0,+\infty]$ and $\phi_{B_i^{-1}} : \mathbb{R}_+ \to [0,+\infty]$, $i = 1,\ldots,m$, vanishing only at 0, such that

$$\begin{aligned}&\langle x - y, u - z\rangle \geq \phi_A\left(\|x - y\|_{\mathcal{H}}\right) \ \forall\,(x,u),(y,z) \in \operatorname{gra} A \\ &\langle v - w, p - q\rangle \geq \phi_{B_i^{-1}}\left(\|v - w\|_{\mathcal{G}_i}\right) \ \forall\,(v,p),(w,q) \in \operatorname{gra} B_i^{-1} \ \forall i = 1,\ldots,m.\end{aligned} \tag{2.24}$$

The function $\phi_{\boldsymbol{M}} : \mathbb{R}_+ \to [0, +\infty]$,

$$\phi_{\boldsymbol{M}}(c) = \inf \left\{ \phi_A(a) + \sum_{i=1}^m \phi_{B_i^{-1}}(b_i) : \sqrt{a^2 + \sum_{i=1}^m b_i^2} = c \right\}, \tag{2.25}$$

is increasing and vanishes only at 0 and it fulfills for each $(\boldsymbol{x}, \boldsymbol{u}), (\boldsymbol{y}, \boldsymbol{z}) \in \operatorname{gra} \boldsymbol{M}$

$$\langle \boldsymbol{x} - \boldsymbol{y}, \boldsymbol{u} - \boldsymbol{z} \rangle_{\boldsymbol{\mathcal{K}}} \geq \phi_{\boldsymbol{M}} \left(\|\boldsymbol{x} - \boldsymbol{y}\|_{\boldsymbol{\mathcal{K}}} \right). \tag{2.26}$$

Thus, $\boldsymbol{M}$ is uniformly monotone on $\boldsymbol{\mathcal{K}}$.

The function $\phi_{\boldsymbol{B}} : \mathbb{R}_+ \to [0, +\infty]$, $\phi_{\boldsymbol{B}}(t) = \phi_{\boldsymbol{M}} \left(\frac{1}{\sqrt{\|\boldsymbol{V}\|}} t \right)$, is increasing and vanishes only at 0. Let be $(\boldsymbol{x}, \boldsymbol{u}), (\boldsymbol{y}, \boldsymbol{z}) \in \operatorname{gra} \boldsymbol{B}$. Then there exist $\boldsymbol{v} \in \boldsymbol{M}\boldsymbol{x}$ and $\boldsymbol{w} \in \boldsymbol{M}\boldsymbol{y}$ fulfilling $\boldsymbol{V}\boldsymbol{u} = \frac{1}{2}\boldsymbol{S}\boldsymbol{x} + \boldsymbol{v}$ and $\boldsymbol{V}\boldsymbol{z} = \frac{1}{2}\boldsymbol{S}\boldsymbol{y} + \boldsymbol{w}$ and it holds

$$\begin{aligned}
\langle \boldsymbol{x} - \boldsymbol{y}, \boldsymbol{u} - \boldsymbol{z} \rangle_{\boldsymbol{\mathcal{K}}_{\boldsymbol{V}}} &= \langle \boldsymbol{x} - \boldsymbol{y}, \boldsymbol{V}\boldsymbol{u} - \boldsymbol{V}\boldsymbol{z} \rangle_{\boldsymbol{\mathcal{K}}} \\
&= \left\langle \boldsymbol{x} - \boldsymbol{y}, \left(\frac{1}{2}\boldsymbol{S}\boldsymbol{x} + \boldsymbol{v} \right) - \left(\frac{1}{2}\boldsymbol{S}\boldsymbol{y} + \boldsymbol{w} \right) \right\rangle_{\boldsymbol{\mathcal{K}}} \\
&\overset{(2.26)}{\geq} \phi_{\boldsymbol{M}} \left(\|\boldsymbol{x} - \boldsymbol{y}\|_{\boldsymbol{\mathcal{K}}} \right) \\
&\geq \phi_{\boldsymbol{M}} \left(\frac{1}{\sqrt{\|\boldsymbol{V}\|}} \|\boldsymbol{x} - \boldsymbol{y}\|_{\boldsymbol{\mathcal{K}}_{\boldsymbol{V}}} \right) \\
&= \phi_{\boldsymbol{B}} \left(\|\boldsymbol{x} - \boldsymbol{y}\|_{\boldsymbol{\mathcal{K}}_{\boldsymbol{V}}} \right).
\end{aligned} \tag{2.27}$$

Consequently, $\boldsymbol{B}$ is uniformly monotone on $\boldsymbol{\mathcal{K}}_{\boldsymbol{V}}$ and, according to [11, Theorem 2.1(ii)(b)], $(J_{\boldsymbol{B}}\boldsymbol{x}_n)_{n \geq 0}$ converges strongly to the unique element $\overline{\boldsymbol{y}} \in \operatorname{zer}(\boldsymbol{A} + \boldsymbol{B}) = \operatorname{zer}(\boldsymbol{M} + \boldsymbol{S} + \boldsymbol{Q})$. In the light of (2.21) and using that $\boldsymbol{\beta}_n \to 0$ $(n \to +\infty)$, it follows that $\boldsymbol{y}_n \to \overline{\boldsymbol{y}}$ $(n \to +\infty)$. ☐

Remark 2.1. Some remarks concerning Algorithm 2.1 and Theorem 2.1 are in order.

(i) Algorithm 2.1 is a fully decomposable iterative method, as each of the operators occurring in Problem 2.1 is processed individually. Moreover, a considerable number of steps in (2.4) can be executed in parallel.

(ii) The proof of Theorem 2.1, which states the convergence of Algorithm 2.1, relies on the reformulation of the iterative scheme as an inexact Douglas-Rachford method in a specific real Hilbert space. For the use of a similar technique in the context of a forward-backward-type method we refer to [19].

(iii) We would like to notice that the assumption $\sum_{n=0}^{+\infty} \lambda_n \|a_n\|_{\mathcal{H}} < +\infty$ does not necessarily imply either that $(\|a_n\|_{\mathcal{H}})_{n \geq 0}$ is summable or that $(a_n)_{n \geq 0}$ (weakly or strongly) converges to 0 as $n \to +\infty$. We refer to [11, Remark 2.2(iii)] for further considerations on the conditions imposed on the error sequences in Theorem 2.1.

Remark 2.2. In the following we emphasize the relations between the proposed algorithm and other existent primal-dual iterative schemes.

(i) Other iterative methods for solving the primal-dual monotone inclusion pair introduced in Problem 2.1 were given in [12] and [19] for D_i^{-1}, $i = 1, \ldots, m$ monotone Lipschitzian and cocoercive operators, respectively. Different to the approach proposed in this subsection, there, the operators D_i^{-1}, $i = 1, \ldots, m$, are processed within some forward steps.

(ii) When for every $i = 1,\dots,m$ one takes $D_i(0) = \mathcal{G}_i$ and $D_i(v) = \varnothing \ \forall v \in \mathcal{G}_i \setminus \{0\}$, the algorithms proposed in [12, Theorem 3.1] (see, also, [8, Theorem 3.1] for the case $m = 1$) and [19, Theorem 3.1] applied to Problem 2.1 differ from Algorithm 2.1.

(iii) When solving the particular case of a primal-dual pair of convex optimization problems discussed in Example 2.1 one can make use of the iterative schemes provided in [13, Algorithm 3.1] and [9, Algorithm 1]. Let us notice that particularizing Algorithm 2.1 to this framework gives rise to a numerical scheme different to the ones in the mentioned literature.

2.2 A second primal-dual algorithm

In Algorithm 2.1 each operator L_i and its adjoint L_i^*, $i = 1,\dots,m$ are processed two times. However, for large-scale optimization problems these matrix-vector multiplications may be expensive compared with the computation of the resolvents of the operators A, B_i^{-1} and D_i^{-1}, $i = 1,\dots,m$.

The second primal-dual algorithm we propose for solving the monotone inclusions in Problem 2.1 has the particularity that it evaluates each operator L_i and its adjoint L_i^*, $i = 1,\dots,m$, only once.

Algorithm 2.2.
Let $x_0 \in \mathcal{H}$, $(y_{1,0},\dots,y_{m,0}) \in \mathcal{G}_1 \times \dots \times \mathcal{G}_m$, $(v_{1,0},\dots,v_{m,0}) \in \mathcal{G}_1 \times \dots \times \mathcal{G}_m$, and τ and σ_i, $i = 1,\dots,m$, be strictly positive real numbers such that

$$\tau \sum_{i=1}^m \sigma_i \|L_i\|^2 < \frac{1}{4}.$$

Furthermore, let $\gamma_i \le 2\sigma_i^{-1}\tau \sum_{i=1}^m \sigma_i \|L_i\|^2$, $i = 1,\dots,m$, let $(\lambda_n)_{n\ge 0}$ be a sequence in $(0,2)$, $(a_n)_{n\ge0}$ a sequence in $\mathcal{H}$, $(b_{i,n})_{n\ge0}$ and $(d_{i,n})_{n\ge0}$ sequences in $\mathcal{G}_i$ for all $i = 1,\dots,m$ and set

$$(\forall n \ge 0) \quad \left\lfloor \begin{array}{l} p_{1,n} = J_{\tau A}\left(x_n - \tau\left(\sum_{i=1}^m L_i^* v_{i,n} - z\right)\right) + a_n \\ x_{n+1} = x_n + \lambda_n(p_{1,n} - x_n) \\ \text{For } i = 1,\dots,m \\ \left\lfloor \begin{array}{l} p_{2,i,n} = J_{\gamma_i D_i}\left(y_{i,n} + \gamma_i v_{i,n}\right) + d_{i,n} \\ y_{i,n+1} = y_{i,n} + \lambda_n(p_{2,i,n} - y_{i,n}) \\ p_{3,i,n} = J_{\sigma_i B_i^{-1}}\left(v_{i,n} + \sigma_i\left(L_i(2p_{1,n} - x_n) - (2p_{2,i,n} - y_{i,n}) - r_i\right)\right) + b_{i,n} \\ v_{i,n+1} = v_{i,n} + \lambda_n(p_{3,i,n} - v_{i,n}). \end{array} \right. \end{array} \right. \tag{2.28}$$

Theorem 2.2. *In Problem 2.1 suppose that*

$$z \in \operatorname{ran}\left(A + \sum_{i=1}^m L_i^*(B_i \,\square\, D_i)(L_i \cdot - r_i)\right). \tag{2.29}$$

and consider the sequences generated by Algorithm 2.2.

(i) If

$$\sum_{n=0}^{+\infty} \lambda_n \|a_n\|_{\mathcal{H}} < +\infty, \quad \sum_{n=0}^{+\infty} \lambda_n(\|d_{i,n}\|_{\mathcal{G}_i} + \|b_{i,n}\|_{\mathcal{G}_i}) < +\infty, \ i = 1,\dots,m,$$

and $\sum_{n=0}^{+\infty} \lambda_n(2-\lambda_n) = +\infty$, *then*

(a) $(x_n, y_{1,n}, \ldots, y_{m,n}, v_{1,n}, \ldots, v_{m,n})_{n\geq 0}$ *converges weakly to an element* $(\overline{x}, \overline{y}_1, \ldots, \overline{y}_m, \overline{v}_1, \ldots, \overline{v}_m) \in \mathcal{H} \times \mathcal{G}_1 \times \ldots \times \mathcal{G}_m \times \mathcal{G}_1 \times \ldots \times \mathcal{G}_m$ *such that* $(\overline{x}, \overline{v}_1, \ldots, \overline{v}_m)$ *is a primal-dual solution to Problem 2.1.*

(b) $\lambda_n(p_{1,n} - x_n) \to 0$ $(n \to +\infty)$, $\lambda_n(p_{2,i,n} - y_{i,n}) \to 0$ $(n \to +\infty)$ *and* $\lambda_n(p_{3,i,n} - v_{i,n}) \to 0$ $(n \to +\infty)$ *for* $i = 1, \ldots, m$.

(c) *whenever* $\mathcal{H}$ *and* $\mathcal{G}_i$, $i = 1, \ldots, m$, *are finite-dimensional Hilbert spaces,* $(x_n, v_{1,n}, \ldots, v_{m,n})_{n\geq 0}$ *converges to a primal-dual solution of Problem 2.1.*

(ii) *If*

$$\sum_{n=0}^{+\infty} \|a_n\|_{\mathcal{H}} < +\infty, \quad \sum_{n=0}^{+\infty} (\|d_{i,n}\|_{\mathcal{G}_i} + \|b_{i,n}\|_{\mathcal{G}_i}) < +\infty, \ i = 1, \ldots, m, \quad \inf_{n\geq 0} \lambda_n > 0$$

and A, B_i^{-1} *and* D_i, $i = 1, \ldots, m$, *are uniformly monotone,*

then $(p_{1,n}, p_{3,1,n}, \ldots, p_{3,m,n})_{n\geq 0}$ *converges strongly to the unique primal-dual solution of Problem 2.1.*

Proof. We let $\boldsymbol{\mathcal{G}} = \mathcal{G}_1 \times \ldots \times \mathcal{G}_m$ be the real Hilbert space endowed with the inner product and associated norm defined in (2.6) and consider

$$\boldsymbol{\mathcal{K}} = \mathcal{H} \times \boldsymbol{\mathcal{G}} \times \boldsymbol{\mathcal{G}},$$

the real Hilbert space endowed with inner product and associated norm defined for $\boldsymbol{x} = (x, \boldsymbol{y}, \boldsymbol{v}), \boldsymbol{u} = (u, \boldsymbol{q}, \boldsymbol{p}) \in \boldsymbol{\mathcal{K}}$ as

$$\langle \boldsymbol{x}, \boldsymbol{u} \rangle_{\boldsymbol{\mathcal{K}}} = \langle x, u \rangle_{\mathcal{H}} + \langle \boldsymbol{y}, \boldsymbol{q} \rangle_{\boldsymbol{\mathcal{G}}} + \langle \boldsymbol{v}, \boldsymbol{p} \rangle_{\boldsymbol{\mathcal{G}}} \text{ and } \|\boldsymbol{x}\|_{\boldsymbol{\mathcal{K}}} = \sqrt{\|x\|_{\mathcal{H}}^2 + \|\boldsymbol{y}\|_{\boldsymbol{\mathcal{G}}}^2 + \|\boldsymbol{v}\|_{\boldsymbol{\mathcal{G}}}^2}, \tag{2.30}$$

respectively. In what follows we set

$$\boldsymbol{y} = (y_1, \ldots, y_m), \quad \boldsymbol{v} = (v_1, \ldots, v_m), \quad \overline{\boldsymbol{y}} = (\overline{y}_1, \ldots, \overline{y}_m), \quad \overline{\boldsymbol{v}} = (\overline{v}_1, \ldots, \overline{v}_m).$$

Consider the set-valued operator

$$\boldsymbol{M} : \boldsymbol{\mathcal{K}} \to 2^{\boldsymbol{\mathcal{K}}}, \ (x, \boldsymbol{y}, \boldsymbol{v}) \ \mapsto (-z + Ax, D_1 y_1, \ldots, D_m y_m, r_1 + B_1^{-1} v_1, \ldots, r_m + B_m^{-1} v_m),$$

which is maximally monotone, since A, B_i and D_i, $i = 1, \ldots, m$, are maximally monotone (cf. [1, Proposition 20.22 and Proposition 20.23]) and the bounded linear operator

$$\boldsymbol{S} : \boldsymbol{\mathcal{K}} \to \boldsymbol{\mathcal{K}}, \quad (x, \boldsymbol{y}, \boldsymbol{v}) \mapsto \left(\sum_{i=1}^m L_i^* v_i, -v_1, \ldots, -v_m, -L_1 x + y_1, \ldots, -L_m x + y_m \right),$$

which proves to be skew (i. e. $\boldsymbol{S}^* = -\boldsymbol{S}$) and hence maximally monotone (cf. [1, Example 20.30]). Since $\operatorname{dom} \boldsymbol{S} = \boldsymbol{\mathcal{K}}$, the sum $\boldsymbol{M} + \boldsymbol{S}$ is maximally monotone, as well (cf. [1, Corollary 24.4(i)]). Further, we have

$$\begin{aligned}
(2.29) &\Leftrightarrow (\exists\, x \in \mathcal{H})\ z \in Ax + \sum_{i=1}^m L_i^* \left(B_i \,\Box\, D_i\right)(L_i x - r_i) \\
&\Leftrightarrow (\exists\, (x, \boldsymbol{v}) \in \mathcal{H} \times \boldsymbol{\mathcal{G}}) \left\{ \begin{array}{l} z \in Ax + \sum_{i=1}^m L_i^* v_i \\ v_i \in \left(B_i \,\Box\, D_i\right)(L_i x - r_i),\ i = 1, \ldots, m \end{array} \right. \\
&\Leftrightarrow (\exists\, (x, \boldsymbol{v}) \in \mathcal{H} \times \boldsymbol{\mathcal{G}}) \left\{ \begin{array}{l} z \in Ax + \sum_{i=1}^m L_i^* v_i \\ L_i x - r_i \in B_i^{-1} v_i + D_i^{-1} v_i,\ i = 1, \ldots, m \end{array} \right. \\
&\Leftrightarrow (\exists\, (x, \boldsymbol{y}, \boldsymbol{v}) \in \boldsymbol{\mathcal{K}}) \left\{ \begin{array}{l} 0 \in -z + Ax + \sum_{i=1}^m L_i^* v_i \\ 0 \in D_i y_i - v_i,\ i = 1, \ldots, m \\ 0 \in r_i + B_i^{-1} v_i - L_i x + y_i,\ i = 1, \ldots, m \end{array} \right. \\
&\Leftrightarrow (\exists\, (x, \boldsymbol{y}, \boldsymbol{v}) \in \boldsymbol{\mathcal{K}})\ (0, \ldots, 0) \in (\boldsymbol{M} + \boldsymbol{S})(x, \boldsymbol{y}, \boldsymbol{v}) \\
&\Leftrightarrow \operatorname{zer}(\boldsymbol{M} + \boldsymbol{S}) \neq \varnothing. \qquad (2.31)
\end{aligned}$$

From the above calculations it follows that

$$\begin{aligned}
(x, \boldsymbol{y}, \boldsymbol{v}) \in \operatorname{zer}(\boldsymbol{M} + \boldsymbol{S}) &\Rightarrow \left\{ \begin{array}{l} z - \sum_{i=1}^m L_i^* v_i \in Ax \\ v_i \in \left(B_i \,\Box\, D_i\right)(L_i x - r_i),\ i = 1, \ldots, m \end{array} \right. \\
&\Leftrightarrow (x, v_1, \ldots, v_m) \text{ is a primal-dual solution to Problem 2.1.} \qquad (2.32)
\end{aligned}$$

Finally, we introduce the bounded linear operator

$$\begin{aligned}
&\boldsymbol{V} : \boldsymbol{\mathcal{K}} \to \boldsymbol{\mathcal{K}} \\
&(x, \boldsymbol{y}, \boldsymbol{v}) \mapsto \left(\frac{x}{\tau} - \sum_{i=1}^m L_i^* v_i, \frac{y_1}{\gamma_1} + v_1, \ldots, \frac{y_m}{\gamma_m} + v_m, \frac{v_1}{\sigma_1} - L_1 x + y_1, \ldots, \frac{v_m}{\sigma_m} - L_m x + y_m \right),
\end{aligned}$$

which is self-adjoint, i. e. $\boldsymbol{V}^* = \boldsymbol{V}$. Furthermore, the operator $\boldsymbol{V}$ is ρ-strongly positive for

$$\rho = \left(1 - 2\sqrt{\tau \sum_{i=1}^m \sigma_i \|L_i\|^2}\right) \min\left\{ \frac{1}{\tau}, \frac{1}{\gamma_1}, \ldots, \frac{1}{\gamma_m}, \frac{1}{\sigma_1}, \ldots, \frac{1}{\sigma_m} \right\},$$

which is a positive real number due to the assumption

$$\tau \sum_{i=1}^m \sigma_i \|L_i\|^2 < \frac{1}{4} \qquad (2.33)$$

made in Algorithm 2.2. Indeed, for $\gamma_i \leq 2\sigma_i^{-1} \tau \sum_{i=1}^m \sigma_i \|L_i\|^2$ it yields for each $i = 1, \ldots, m$,

$$\begin{aligned}
2 \left\langle L_i x - y_i, v_i \right\rangle_{\mathcal{G}_i} &\leq 2\|L_i\|\|x\|_{\mathcal{H}}\|v_i\|_{\mathcal{G}_i} + 2\|y_i\|_{\mathcal{G}_i}\|v_i\|_{\mathcal{G}_i} \\
&\leq \frac{\sigma_i \|L_i\|^2 \|x\|_{\mathcal{H}}^2}{\sqrt{\tau \sum_{i=1}^m \sigma_i \|L_i\|^2}} + 2\sqrt{\tau \sum_{i=1}^m \sigma_i \|L_i\|^2} \frac{\|y_i\|_{\mathcal{G}_i}^2}{\gamma_i} + 2\sqrt{\tau \sum_{i=1}^m \sigma_i \|L_i\|^2} \frac{\|v_i\|_{\mathcal{G}_i}^2}{\sigma_i}
\end{aligned}$$

and, consequently, for each $\boldsymbol{x} = (x, \boldsymbol{y}, \boldsymbol{v}) \in \boldsymbol{\mathcal{K}}$, it follows that

$$\begin{aligned}
\langle \boldsymbol{x}, \boldsymbol{V}\boldsymbol{x} \rangle_{\boldsymbol{\mathcal{K}}} &= \frac{\|x\|_{\mathcal{H}}^2}{\tau} + \sum_{i=1}^m \left[\frac{\|y_i\|_{\mathcal{G}_i}^2}{\gamma_i} + \frac{\|v_i\|_{\mathcal{G}_i}^2}{\sigma_i} \right] - 2\sum_{i=1}^m \langle L_i x - y_i, v_i \rangle_{\mathcal{G}_i} \\
&\geq \left(1 - 2\sqrt{\tau \sum_{i=1}^m \sigma_i \|L_i\|^2} \right) \min\left\{ \frac{1}{\tau}, \frac{1}{\gamma_1}, \ldots, \frac{1}{\gamma_m}, \frac{1}{\sigma_1}, \ldots, \frac{1}{\sigma_m} \right\} \|\boldsymbol{x}\|_{\boldsymbol{\mathcal{K}}}^2 \\
&= \rho \|\boldsymbol{x}\|_{\boldsymbol{\mathcal{K}}}^2.
\end{aligned} \tag{2.34}$$

The algorithmic scheme (2.28) is equivalent to

$$(\forall n \geq 0) \quad \left\lfloor \begin{array}{l}
\frac{x_n - p_{1,n}}{\tau} - \sum_{i=1}^m L_i^* v_{i,n} \in -z + A(p_{1,n} - a_n) - \frac{a_n}{\tau} \\
x_{n+1} = x_n + \lambda_n (p_{1,n} - x_n) \\
\text{For } i = 1, \ldots, m \\
\left\lfloor \begin{array}{l}
\frac{y_{i,n} - p_{2,i,n}}{\gamma_i} + v_{i,n} \in D_i(p_{2,i,n} - d_{i,n}) - \frac{d_{i,n}}{\gamma_i} \\
y_{i,n+1} = y_{i,n} + \lambda_n (p_{2,i,n} - y_{i,n}) \\
\frac{v_{i,n} - p_{3,i,n}}{\sigma_i} - L_i(x_n - p_{1,n}) + y_{i,n} - p_{2,i,n} \\
\qquad\qquad \in r_i + B_i^{-1}(p_{3,i,n} - b_{i,n}) - L_i p_{1,n} + p_{2,i,n} - \frac{b_{i,n}}{\sigma_i} \\
v_{i,n+1} = v_{i,n} + \lambda_n (p_{3,i,n} - v_{i,n}).
\end{array} \right.
\end{array} \right. \tag{2.35}$$

We introduce for every $n \geq 0$ the following notations:

$$\left\{ \begin{array}{l}
\boldsymbol{x}_n = (x_n, y_{1,n}, \ldots, y_{m,n}, v_{1,n}, \ldots, v_{m,n}) \\
\boldsymbol{p}_n = (p_{1,n}, p_{2,1,n}, \ldots, p_{2,m,n}, p_{3,1,n}, \ldots, p_{3,m,n}) \\
\boldsymbol{a}_n = (a_n, d_{1,n}, \ldots, d_{m,n}, b_{1,n}, \ldots, b_{m,n}) \\
\boldsymbol{a}_n^\tau = (\frac{a_n}{\tau}, \frac{d_{1,n}}{\gamma_1}, \ldots, \frac{d_{m,n}}{\gamma_m}, \frac{b_{1,n}}{\sigma_1}, \ldots, \frac{b_{m,n}}{\sigma_m}).
\end{array} \right. \tag{2.36}$$

Hence, the scheme (2.35) can equivalently be written in the form

$$(\forall n \geq 0) \quad \left\lfloor \begin{array}{l}
\boldsymbol{V}(\boldsymbol{x}_n - \boldsymbol{p}_n) \in (\boldsymbol{S} + \boldsymbol{M})(\boldsymbol{p}_n - \boldsymbol{a}_n) + \boldsymbol{S}\boldsymbol{a}_n - \boldsymbol{a}_n^\tau \\
\boldsymbol{x}_{n+1} = \boldsymbol{x}_n + \lambda_n (\boldsymbol{p}_n - \boldsymbol{x}_n).
\end{array} \right. \tag{2.37}$$

Considering again the Hilbert space $\boldsymbol{\mathcal{K}}_{\boldsymbol{V}}$ with inner product and norm respectively defined as in (2.16), since $\boldsymbol{V}$ is self-adjoint and ρ-strongly positive, weak and strong convergence in $\boldsymbol{\mathcal{K}}_{\boldsymbol{V}}$ are equivalent with weak and strong convergence in $\boldsymbol{\mathcal{K}}$, respectively. Moreover, $\boldsymbol{A} = \boldsymbol{V}^{-1}(\boldsymbol{S} + \boldsymbol{M})$ is maximally monotone on $\boldsymbol{\mathcal{K}}_{\boldsymbol{V}}$. Thus, by denoting $\boldsymbol{e}_n = \boldsymbol{V}^{-1}((\boldsymbol{S} + \boldsymbol{V})\boldsymbol{a}_n - \boldsymbol{a}_n^\tau)$ for every $n \geq 0$ the iterative scheme (2.37) becomes

$$(\forall n \geq 0) \quad \left\lfloor \begin{array}{l}
\boldsymbol{p}_n = J_{\boldsymbol{A}}(\boldsymbol{x}_n - \boldsymbol{e}_n) + \boldsymbol{a}_n \\
\boldsymbol{x}_{n+1} = \boldsymbol{x}_n + \lambda_n (\boldsymbol{p}_n - \boldsymbol{x}_n).
\end{array} \right. \tag{2.38}$$

Furthermore, introducing the maximal monotone operator $\boldsymbol{B} : \boldsymbol{\mathcal{K}} \to 2^{\boldsymbol{\mathcal{K}}}$, $\boldsymbol{x} \mapsto \{0\}$, and defining for every $n \geq 0$

$$\boldsymbol{\alpha}_n = J_{\boldsymbol{A}}(\boldsymbol{x}_n - \boldsymbol{e}_n) - J_{\boldsymbol{A}}(\boldsymbol{x}_n) + \boldsymbol{a}_n,$$

the iterative scheme (2.38) becomes (notice that $J_{\boldsymbol{B}} = \mathrm{Id}$)

$$(\forall n \geq 0) \quad \left\lfloor \begin{array}{l}
\boldsymbol{y}_n = J_{\boldsymbol{B}}(\boldsymbol{x}_n) \\
\boldsymbol{p}_n = J_{\boldsymbol{A}}(2\boldsymbol{y}_n - \boldsymbol{x}_n) + \boldsymbol{\alpha}_n \\
\boldsymbol{x}_{n+1} = \boldsymbol{x}_n + \lambda_n (\boldsymbol{p}_n - \boldsymbol{y}_n)
\end{array} \right. , \tag{2.39}$$

thus, it has the structure of the error-tolerant Douglas-Rachford algorithm from [11]. Obviously, $\operatorname{zer}(\boldsymbol{A}+\boldsymbol{B}) = \operatorname{zer}(\boldsymbol{M}+\boldsymbol{S})$.

(i) The assumptions made on the error sequences yield

$$\sum_{n=0}^{+\infty} \lambda_n \|\boldsymbol{a}_n\|_{\boldsymbol{\mathcal{K}}} < +\infty \text{ and } \sum_{n=0}^{+\infty} \lambda_n \|\boldsymbol{e}_n\|_{\boldsymbol{\mathcal{K}}} < +\infty.$$

Thus, by the nonexpansiveness of the resolvent of $\boldsymbol{A}$,

$$\sum_{n=0}^{+\infty} \lambda_n \|\boldsymbol{\alpha}_n\|_{\boldsymbol{\mathcal{K}}} < +\infty$$

and, consequently, by the linearity and boundedness of $\boldsymbol{V}$,

$$\sum_{n=0}^{+\infty} \lambda_n \|\boldsymbol{\alpha}_n\|_{\boldsymbol{\mathcal{K}}_{\boldsymbol{V}}} < +\infty.$$

(i)(a) Follows directly from [11, Theorem 2.1(i)(a)] by using that $J_{\boldsymbol{B}} = \operatorname{Id}$ and relation (2.32).

(i)(b) Follows in analogy to the proof of Theorem 2.1(i)(b).

(i)(c) Follows from Theorem 2.2(i)(a).

(ii) The iterative scheme (2.38) can be also formulated as

$$(\forall n \geq 0) \left\lfloor \begin{array}{l} \boldsymbol{p}_n = J_{\boldsymbol{A}}(\boldsymbol{x}_n) + \boldsymbol{\alpha}_n \\ \boldsymbol{y}_n = J_{\boldsymbol{B}}(2\boldsymbol{p}_n - \boldsymbol{x}_n) \\ \boldsymbol{x}_{n+1} = \boldsymbol{x}_n + \lambda_n(\boldsymbol{y}_n - \boldsymbol{p}_n) \end{array} \right. , \tag{2.40}$$

with the error sequence fulfilling

$$\sum_{n=0}^{+\infty} \|\boldsymbol{\alpha}_n\|_{\boldsymbol{\mathcal{K}}_{\boldsymbol{V}}} < +\infty.$$

The statement follows from [11, Theorem 2.1(ii)(b)] by taking into consideration the uniform monotonicity of $\boldsymbol{A}$ and relation (2.32). ☐

Remark 2.3. When for every $i = 1, \ldots, m$ one takes $D_i(0) = \mathcal{G}_i$ and $D_i(v) = \varnothing \ \forall v \in \mathcal{G}_i \setminus \{0\}$, and $(d_{i,n})_{n\geq 0}$ as a sequence of zeros, one can show that the assertions made in Theorem 2.2 hold true for step length parameters satisfying

$$\tau \sum_{i=1}^{m} \sigma_i \|L_i\|^2 < 1,$$

when choosing $(y_{1,0}, \ldots, y_{m,0}) = (0, \ldots, 0)$ in Algorithm 2.2, since the sequences $(y_{1,n}, \ldots, y_{m,n})_{n\geq 0}$ and $(v_{1,n}, \ldots, v_{m,n})_{n\geq 0}$ vanish in this particular situation.

Remark 2.4. In the following we emphasize the relations between Algorithm 2.2 and other existent primal-dual iterative schemes.

(i) When for every $i = 1, \ldots, m$ one takes $D_i(0) = \mathcal{G}_i$ and $D_i(v) = \varnothing \ \forall v \in \mathcal{G}_i \setminus \{0\}$, and $(d_{i,n})_{n\geq 0}$ as a sequence of zeros, Algorithm 2.2 with $(y_{1,0}, \ldots, y_{m,0}) = (0, \ldots, 0)$ as initial choice provides an iterative scheme which is identical to the one in [19, Eq. (3.3)], but differs from the one in [12, Theorem 3.1] (see, also, [8, Theorem 3.1] for the case $m = 1$) when the latter are applied to Problem 2.1.

(ii) When solving the particular case of a primal-dual pair of convex optimization problems discussed in Example 2.1 and when considering as initial choice $y_{1,0} = 0$, Algorithm 2.2 gives rise to an iterative scheme which is equivalent to [13, Algorithm 3.1]. In addition, under the assumption of exact implementations, the method in Algorithm 2.2 equals the one in [9, Algorithm 1], our choice of $(\lambda_n)_{n\geq 0}$ to be variable in the interval $(0, 2)$, however, relaxes the assumption in [9] that $(\lambda_n)_{n\geq 0}$ is a constant sequence in $(0, 1]$.

3 Application to convex minimization problems

In this section we particularize the two iterative schemes introduced and investigated in this paper in the context of solving a primal-dual pair of convex optimization problems. To this end we consider the following problem.

Problem 3.1. Let $\mathcal{H}$ be a real Hilbert space and let $f \in \Gamma(\mathcal{H})$, $z \in \mathcal{H}$. For every $i \in \{1, \dots, m\}$, suppose that $\mathcal{G}_i$ is a real Hilbert space, let g_i, $l_i \in \Gamma(\mathcal{G}_i)$, $r_i \in \mathcal{G}_i$ and let $L_i : \mathcal{H} \to \mathcal{G}_i$ be a nonzero bounded linear operator. Consider the convex optimization problem

$$(P) \quad \inf_{x\in\mathcal{H}} \left\{ f(x) + \sum_{i=1}^{m} (g_i \,\Box\, l_i)(L_i x - r_i) - \langle x, z\rangle \right\} \tag{3.1}$$

and its conjugate dual problem

$$(D) \quad \sup_{(v_1,\dots,v_m)\in\mathcal{G}_1\times\ldots\times\mathcal{G}_m} \left\{ -f^*\left(z - \sum_{i=1}^{m} L_i^* v_i \right) - \sum_{i=1}^{m} \left(g_i^*(v_i) + l_i^*(v_i) + \langle v_i, r_i\rangle \right) \right\}. \tag{3.2}$$

Considering the maximal monotone operators

$$A = \partial f, \;\; B_i = \partial g_i \text{ and } D_i = \partial l_i, \;\; i = 1, \dots, m,$$

the monotone inclusion problem (2.1) reads

$$\text{find } \overline{x} \in \mathcal{H} \text{ such that } z \in \partial f(\overline{x}) + \sum_{i=1}^{m} L_i^* (\partial g_i \,\Box\, \partial l_i)(L_i \overline{x} - r_i), \tag{3.3}$$

while the dual inclusion problem (2.2) reads

$$\text{find } \overline{v}_1 \in \mathcal{G}_1, \dots, \overline{v}_m \in \mathcal{G}_m \text{ such that } (\exists x \in \mathcal{H}) \begin{cases} z - \sum_{i=1}^{m} L_i^* \overline{v}_i \in \partial f(x) \\ \overline{v}_i \in (\partial g_i \,\Box\, \partial l_i)(L_i x - r_i), \; i = 1, \dots, m. \end{cases} \tag{3.4}$$

If $(\overline{x}, \overline{v}_1, \dots, \overline{v}_m) \in \mathcal{H} \times \mathcal{G}_1 \ldots \times \mathcal{G}_m$ is a primal-dual solution to (3.3)-(3.4), namely,

$$z - \sum_{i=1}^{m} L_i^* \overline{v}_i \in \partial f(\overline{x}) \text{ and } \overline{v}_i \in (\partial g_i \,\Box\, \partial l_i)(L_i \overline{x} - r_i), \; i = 1, \dots, m, \tag{3.5}$$

then $\overline{x}$ is an optimal solution to (P), $(\overline{v}_1, \dots, \overline{v}_m)$ is an optimal solution to (D) and the optimal objective values of the two problems, which we denote by $v(P)$ and $v(D)$, respectively, coincide (thus, strong duality holds).

Combining this statement with Algorithm 2.1 and Theorem 2.1 give rise to the following iterative scheme and corresponding convergence results for the primal-dual pair of optimization problems $(P) - (D)$. We also use that the subdifferential of a uniformly convex function is uniformly monotone (cf. [1, Example 22.3(iii)]).

Algorithm 3.1.
Let $x_0 \in \mathcal{H}$, $(v_{1,0}, \dots, v_{m,0}) \in \mathcal{G}_1 \times \dots \times \mathcal{G}_m$ and τ and σ_i, $i = 1, \dots, m$, be strictly positive real numbers such that

$$\tau \sum_{i=1}^{m} \sigma_i \|L_i\|^2 < 4.$$

Furthermore, let $(\lambda_n)_{n\geq 0}$ be a sequence in $(0, 2)$, $(a_n)_{n\geq 0}$ a sequence in $\mathcal{H}$, $(b_{i,n})_{n\geq 0}$ and $(d_{i,n})_{n\geq 0}$ sequences in $\mathcal{G}_i$ for all $i = 1, \dots, m$ and set

$$(\forall n \geq 0) \quad \left\lfloor \begin{array}{l} p_{1,n} = \operatorname{Prox}_{\tau f}\left(x_n - \frac{\tau}{2}\sum_{i=1}^m L_i^* v_{i,n} + \tau z\right) + a_n \\ w_{1,n} = 2p_{1,n} - x_n \\ \text{For } i = 1, \dots, m \\ \quad \left\lfloor \begin{array}{l} p_{2,i,n} = \operatorname{Prox}_{\sigma_i g_i^*}\left(v_{i,n} + \frac{\sigma_i}{2} L_i w_{1,n} - \sigma_i r_i\right) + b_{i,n} \\ w_{2,i,n} = 2p_{2,i,n} - v_{i,n} \end{array} \right. \\ z_{1,n} = w_{1,n} - \frac{\tau}{2}\sum_{i=1}^m L_i^* w_{2,i,n} \\ x_{n+1} = x_n + \lambda_n (z_{1,n} - p_{1,n}) \\ \text{For } i = 1, \dots, m \\ \quad \left\lfloor \begin{array}{l} z_{2,i,n} = \operatorname{Prox}_{\sigma_i l_i^*}\left(w_{2,i,n} + \frac{\sigma_i}{2} L_i (2z_{1,n} - w_{1,n})\right) + d_{i,n} \\ v_{i,n+1} = v_{i,n} + \lambda_n (z_{2,i,n} - p_{2,i,n}). \end{array} \right. \end{array} \right. \tag{3.6}$$

Theorem 3.1. *For Problem 3.1 suppose that*

$$z \in \operatorname{ran}\left(\partial f + \sum_{i=1}^{m} L_i^* (\partial g_i \,\square\, \partial l_i)(L_i \cdot - r_i)\right), \tag{3.7}$$

and consider the sequences generated by Algorithm 3.1.

(i) If

$$\sum_{n=0}^{+\infty} \lambda_n \|a_n\|_{\mathcal{H}} < +\infty, \quad \sum_{n=0}^{+\infty} \lambda_n (\|d_{i,n}\|_{\mathcal{G}_i} + \|b_{i,n}\|_{\mathcal{G}_i}) < +\infty, i = 1, \dots, m,$$

and $\sum_{n=0}^{+\infty} \lambda_n (2 - \lambda_n) = +\infty$, *then*

(i) $(x_n, v_{1,n}, \dots, v_{m,n})_{n\geq 0}$ *converges weakly to an element* $(\overline{x}, \overline{v}_1, \dots, \overline{v}_m) \in \mathcal{H} \times \mathcal{G}_1 \times \dots \times \mathcal{G}_m$ *such that, when setting*

$$\overline{p}_1 = \operatorname{Prox}_{\tau f}\left(\overline{x} - \frac{\tau}{2}\sum_{i=1}^{m} L_i^* \overline{v}_i + \tau z\right),$$

$$\text{and } \overline{p}_{2,i} = \operatorname{Prox}_{\sigma_i g_i^*}\left(\overline{v}_i + \frac{\sigma_i}{2} L_i (2\overline{p}_1 - \overline{x}) - \sigma_i r_i\right) \; i = 1, \dots, m,$$

$\overline{p}_1$ *is an optimal solution to* (P), $(\overline{p}_{2,1}, \dots, \overline{p}_{2,m})$ *is an optimal solution to* (D) *and* $v(P) = v(D)$.

(ii) $\lambda_n (z_{1,n} - p_{1,n}) \to 0$ $(n \to +\infty)$ *and* $\lambda_n (z_{2,i,n} - p_{2,i,n}) \to 0$ $(n \to +\infty)$ *for* $i = 1, \dots, m$.

(iii) whenever $\mathcal{H}$ *and* $\mathcal{G}_i, i = 1, \dots, m$, *are finite-dimensional Hilbert spaces,* $a_n \to 0$ $(n \to +\infty)$ *and* $b_{i,n} \to 0$ $(n \to +\infty)$ *for* $i = 1, \dots, m$, *then* $(p_{1,n})_{n\geq 0}$ *converges to an optimal solution to* (P) *and* $(p_{2,1,n}, \dots, p_{2,m,n})_{n\geq 0}$ *converges to an optimal solution to* (D).

(ii) If

$$\sum_{n=0}^{+\infty} \|a_n\|_{\mathcal{H}} < +\infty, \quad \sum_{n=0}^{+\infty} (\|d_{i,n}\|_{\mathcal{G}_i} + \|b_{i,n}\|_{\mathcal{G}_i}) < +\infty,\ i = 1,\ldots,m, \quad \inf_{n\geq 0} \lambda_n > 0$$

and f and g_i^*, $i = 1,\ldots,m$, are uniformly convex,

then $(p_{1,n})_{n\geq 0}$ converges strongly to an optimal solution to (P), $(p_{2,1,n},\ldots,p_{2,m,n})_{n\geq 0}$ converges strongly to an optimal solution to (D) and $v(P) = v(D)$.

Algorithm 2.2 and Theorem 2.2 give rise to the following iterative scheme and corresponding convergence results for the primal-dual pair of optimization problems $(P)-(D)$.

Algorithm 3.2.
Let $x_0 \in \mathcal{H}$, $(y_{1,0},\ldots,y_{m,0}) \in \mathcal{G}_1 \times \ldots \times \mathcal{G}_m$, $(v_{1,0},\ldots,v_{m,0}) \in \mathcal{G}_1 \times \ldots \times \mathcal{G}_m$, and τ and σ_i, $i = 1,\ldots,m$, be strictly positive real numbers such that

$$\tau \sum_{i=1}^{m} \sigma_i \|L_i\|^2 < \frac{1}{4}.$$

Furthermore, let $\gamma_i \leq 2\sigma_i^{-1}\tau \sum_{i=1}^m \sigma_i\|L_i\|^2$, $i = 1,\ldots,m$, $(\lambda_n)_{n\geq 0}$ be a sequence in $(0,2)$, $(a_n)_{n\geq 0}$ a sequence in $\mathcal{H}$, $(b_{i,n})_{n\geq 0}$ and $(d_{i,n})_{n\geq 0}$ sequences in $\mathcal{G}_i$ for all $i = 1,\ldots,m$ and set

$$(\forall n \geq 0) \quad \left\lfloor \begin{array}{l} p_{1,n} = \operatorname{Prox}_{\tau f}\left(x_n - \tau\left(\sum_{i=1}^m L_i^* v_{i,n} - z\right)\right) + a_n \\ x_{n+1} = x_n + \lambda_n(p_{1,n} - x_n) \\ \text{For } i = 1,\ldots,m \\ \left\lfloor \begin{array}{l} p_{2,i,n} = \operatorname{Prox}_{\gamma_i l_i}\left(y_{i,n} + \gamma_i v_{i,n}\right) + d_{i,n} \\ y_{i,n+1} = y_{i,n} + \lambda_n(p_{2,i,n} - y_{i,n}) \\ p_{3,i,n} = \operatorname{Prox}_{\sigma_i g_i^*}\left(v_{i,n} + \sigma_i\left(L_i(2p_{1,n} - x_n) - (2p_{2,i,n} - y_{i,n}) - r_i\right)\right) + b_{i,n} \\ v_{i,n+1} = v_{i,n} + \lambda_n(p_{3,i,n} - v_{i,n}). \end{array} \right. \end{array} \right. \tag{3.8}$$

Theorem 3.2. *In Problem 3.1 suppose that*

$$z \in \operatorname{ran}\left(\partial f + \sum_{i=1}^m L_i^*(\partial g_i \,\square\, \partial l_i)(L_i \cdot - r_i)\right), \tag{3.9}$$

and consider the sequences generated by Algorithm 3.2.

(i) If

$$\sum_{n=0}^{+\infty} \lambda_n\|a_n\|_{\mathcal{H}} < +\infty, \quad \sum_{n=0}^{+\infty} \lambda_n(\|d_{i,n}\|_{\mathcal{G}_i} + \|b_{i,n}\|_{\mathcal{G}_i}) < +\infty,\ i = 1,\ldots,m,$$

and $\sum_{n=0}^{+\infty} \lambda_n(2-\lambda_n) = +\infty$, then

(a) $(x_n, y_{1,n},\ldots,y_{m,n}, v_{1,n},\ldots,v_{m,n})_{n\geq 0}$ converges weakly to an element $(\overline{x}, \overline{y}_1, \ldots, \overline{y}_m, \overline{v}_1, \ldots, \overline{v}_m) \in \mathcal{H} \times \mathcal{G}_1 \times \ldots \times \mathcal{G}_m \times \mathcal{G}_1 \times \ldots \times \mathcal{G}_m$ such that $\overline{x}$ is an optimal solution to (P), $(\overline{v}_1, \ldots, \overline{v}_m)$ is an optimal solution to (D) and $v(P) = v(D)$.

(b) $\lambda_n(p_{1,n}-x_n) \to 0$ $(n \to +\infty)$, $\lambda_n(p_{2,i,n}-y_{i,n}) \to 0$ $(n \to +\infty)$ *and* $\lambda_n(p_{3,i,n} - v_{i,n}) \to 0$ $(n \to +\infty)$ *for* $i = 1,\dots,m$.

(c) *whenever* $\mathcal{H}$ *and* $\mathcal{G}_i$, $i = 1,\dots,m$, *are finite-dimensional Hilbert spaces,* $(x_n)_{n\geq 0}$ *converges to an optimal solution to* (P) *and* $(v_{1,n},\dots,v_{m,n})_{n\geq 0}$ *converges to an optimal solution to* (D).

(ii) *If*

$$\sum_{n=0}^{+\infty} \|a_n\|_{\mathcal{H}} < +\infty, \quad \sum_{n=0}^{+\infty} (\|d_{i,n}\|_{\mathcal{G}_i} + \|b_{i,n}\|_{\mathcal{G}_i}) < +\infty, \ i = 1,\dots,m, \quad \inf_{n\geq 0} \lambda_n > 0$$

and f, l_i *and* g_i^*, $i = 1,\dots,m$, *are uniformly convex,*

then, $(p_{1,n})_{n\geq 0}$ *converges strongly to the unique optimal solution to* (P), $(p_{3,1,n},\dots,p_{3,m,n})_{n\geq 0}$ *converges strongly to the unique optimal solution of* (D) *and* $v(P) = v(D)$.

Remark 3.1. According to Remark 2.3, when $l_i : \mathcal{G}_i \to \overline{\mathbb{R}}$, $l_i = \delta_{\{0\}}$, and $(d_{i,n})_{n\geq 0}$ is chosen as a sequence of zeros for every $i = 1,\dots,m$, the assertions made in Theorem 3.2 hold true for step length parameters satisfying

$$\tau \sum_{i=1}^{m} \sigma_i \|L_i\|^2 < 1$$

when taking in Algorithm 3.2 as initial choice $(y_{1,0},\dots,y_{m,0}) = (0,\dots,0)$. In this case the sequences $(y_{1,n},\dots,y_{m,n})_{n\geq 0}$ and $(p_{2,1,n},\dots,p_{2,m,n})_{n\geq 0}$ vanish and (3.8) reduces to

$$(\forall n \geq 0) \quad \left\lfloor \begin{array}{l} p_{1,n} = \operatorname{Prox}_{\tau f}\left(x_n - \tau\left(\sum_{i=1}^m L_i^* v_{i,n} - z\right)\right) + a_n \\ x_{n+1} = x_n + \lambda_n(p_{1,n} - x_n) \\ \text{For } i = 1,\dots,m \\ \left\lfloor \begin{array}{l} p_{3,i,n} = \operatorname{Prox}_{\sigma_i g_i^*}\left(v_{i,n} + \sigma_i\left(L_i(2p_{1,n} - x_n) - r_i\right)\right) + b_{i,n} \\ v_{i,n+1} = v_{i,n} + \lambda_n(p_{3,i,n} - v_{i,n}). \end{array} \right. \end{array} \right. \tag{3.10}$$

Remark 3.2. Condition (3.7) in Theorem 3.1 (respectively, condition (3.9) in Theorem 3.2) is fulfilled, if the primal optimization problem (3.1) has an optimal solution,

$$0 \in \operatorname{sqri}\left(\operatorname{dom} g_i^* - \operatorname{dom} l_i^*\right), \ i = 1,\dots,m, \tag{3.11}$$

and (see, also, [12, Proposition 4.3])

$$(r_1,\dots,r_m) \in \operatorname{sqri} E, \tag{3.12}$$

where

$$E := \left\{(L_1 x - y_1,\dots,L_m x - y_m) : x \in \operatorname{dom} f \text{ and } y_i \in \operatorname{dom} g_i + \operatorname{dom} l_i, \ i = 1,\dots,m\right\}.$$

Here, for a nonempty convex set $\Omega \subseteq \mathcal{H}$, we denote by

$$\operatorname{sqri}\Omega = \left\{x \in \Omega : \bigcup_{\lambda\geq 0} \lambda(\Omega - x) \text{ is a closed linear subspace}\right\}$$

its *strong quasi-relative interior*. According to [1, Proposition 15.7], condition (3.11) guarantees that $g_i \,\Box\, l_i \in \Gamma(\mathcal{G}_i)$, $i = 1,\dots,m$.

If one of the following two conditions

(i) for any $i = 1, \dots, m$ one of the functions g_i and l_i is real-valued;

(ii) $\mathcal{H}$ and $\mathcal{G}_i$, $i = 1, \dots, m$, are finite-dimensional and there exists $x \in \operatorname{ri} \operatorname{dom} f$ such that $L_i x - y_i \in \operatorname{ri} \operatorname{dom} g_i + \operatorname{ri} \operatorname{dom} l_i$, $i = 1, \dots, m$;

is fulfilled, then condition (3.12) is obviously true. For (ii) one has to take into account that in finite-dimensional spaces the strong quasi-relative interior of a convex set is nothing else than its relative interior and to use the properties of the latter.

4 Numerical experiments

In this section we emphasize the performance of the algorithms introduced in this article in the context of two numerical experiments on location and image denoising problems.

4.1 The generalized Heron problem

We start by considering the *generalized Heron problem* which has been recently investigated in [14, 15] and where for its solving subgradient-type methods have been used.

While the *classical Heron problem* concerns the finding of a point $\overline{u}$ on a given straight line in the plane such that the sum of distances from $\overline{u}$ to given points u^1, u^2 is minimal, the problem that we address here aims to find a point in a closed convex set $\Omega \subseteq \mathbb{R}^n$ which minimizes the sum of the distances to given convex closed sets $\Omega_i \subseteq \mathbb{R}^n$, $i = 1, \dots, m$.

The distance from a point $x \in \mathbb{R}^n$ to a nonempty set $\Omega \subseteq \mathbb{R}^n$ is given by

$$d(x; \Omega) = (\| \cdot \| \,\Box\, \delta_\Omega)(x) = \inf_{z \in \Omega} \|x - z\|.$$

Thus the *generalized Heron problem* reads

$$\inf_{x \in \Omega} \sum_{i=1}^{m} d(x; \Omega_i), \tag{4.1}$$

where the sets $\Omega \subseteq \mathbb{R}^n$ and $\Omega_i \subseteq \mathbb{R}^n$, $i = 1, \dots, m$, are nonempty, closed and convex. We observe that (4.1) perfectly fits into the framework considered in Problem 3.1 when setting

$$f = \delta_\Omega, \text{ and } g_i = \| \cdot \|, \ l_i = \delta_{\Omega_i} \text{ for all } i = 1, \dots, m. \tag{4.2}$$

However, note that (4.1) cannot be solved via the primal-dual methods in [12] and [19] since they require the presence of at least one strongly convex function (cf. Baillon-Haddad Theorem, [1, Corollary 18.16]) in each of the infimal convolutions $\| \cdot \| \,\Box\, \delta_{\Omega_i}$, $i = 1, \dots, m$, a fact which is obviously not the case. Notice that

$$g_i^* : \mathbb{R}^n \to \overline{\mathbb{R}}, \ g_i^*(p) = \sup_{x \in \mathbb{R}^n} \{\langle p, x \rangle - \|x\|\} = \delta_{B(0,1)}(p), \ i = 1, \dots, m,$$

thus the proximal points of f, g_i^* and l_i^*, $i = 1, \dots, m$, can be calculated via projections, in case of the latter via Moreau's decomposition formula.

In the following we test our algorithms on some examples taken from [14, 15].

Example 4.1 (Example 5.5 in [15])**.** Consider problem (4.1) with the constraint set Ω being the closed ball centered at $(5, 0)$ having radius 2 and the sets Ω_i, $i = 1, \dots, 8$, being pairwise disjoint squares in *right position* in $\mathbb{R}^2$ (i. e. the edges are parallel to the x- and y-axes, respectively), with centers $(-2,4)$, $(-1,-8)$, $(0,0)$, $(0,6)$, $(5,-6)$, $(8,-8)$, $(8,9)$ and $(9,-5)$ and side length 1, respectively (see Figure 4.1).

When solving this problem with Algorithm 3.1 and Algorithm 3.2 (in the sequel called DR1 and DR2, respectively) and the choices made in (4.2), the following formulae for the proximal points involved in their formulations are necessary for $x, p \in \mathbb{R}^2$ and $\tau, \sigma_i \in \mathbb{R}_{++}$, $i = 1, \ldots, 8$:

$$\mathrm{Prox}_{\tau f}(x) = (5,0) + \underset{y \in B(0,2)}{\arg\min} \frac{1}{2}\|y - (x - (5,0))\|^2 = (5,0) + \mathcal{P}_{B(0,2)}(x - (5,0))$$

$$\mathrm{Prox}_{\sigma_i g_i^*}(p) = \underset{z \in B(0,1)}{\arg\min} \frac{1}{2}\|z - p\|^2 = \mathcal{P}_{B(0,1)}(p)$$

$$\mathrm{Prox}_{\sigma_i l_i^*}(p) \overset{(1.3)}{=} p - \sigma_i \,\mathrm{Prox}_{\sigma_i^{-1} l_i}\left(\frac{p}{\sigma_i}\right) = p - \sigma_i \underset{z \in \Omega_i}{\arg\min} \frac{1}{2}\left\|z - \frac{p}{\sigma_i}\right\|^2 = p - \sigma_i \mathcal{P}_{\Omega_i}\left(\frac{p}{\sigma_i}\right).$$

Figure 4.1 gives an insight into the performance of the proposed primal-dual methods when compared with the subgradient algorithm used in [15]. After a few milliseconds both splitting algorithms reach machine precision with respect to the *root-mean-square error* where the following parameters were used:

- DR1: $(\forall i \in \{1,\ldots,8\})$ $\sigma_i = 0.15$, $\tau = 2/(\sum_{j=1}^{8} \sigma_j)$, $\lambda_n = 1.5$, $x_0 = (5,2)$, $v_{i,0} = 0$,
- DR2: $(\forall i \in \{1,\ldots,8\})$ $\sigma_i = 0.1$, $\tau = 0.24/(\sum_{j=1}^{8} \sigma_j)$, $\lambda_n = 1.8$, $x_0 = (5,2)$, $v_{i,0} = 0$,
- Subgradient (cf. [15, Theorem 4.1]) $x_0 = (5,2)$, $\alpha_n = \frac{1}{n}$.

(a) Problem with optimizer (b) Progress of the RMSE values

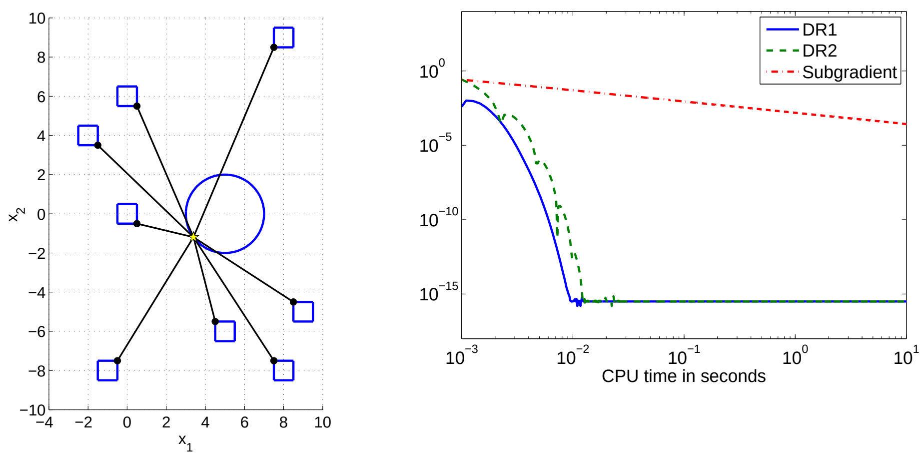


Figure 4.1: Example 4.1. Generalized Heron problem with squares and disc constraint set on the left-hand side, performance evaluation for the *root-mean-square error* (RMSE) on the right-hand side.

Example 4.2 (Example 4.3 in [14]). In this example we solve the generalized Heron problem (4.1) in $\mathbb{R}^3$, where the constraint set Ω is the closed ball centered at $(0,2,0)$ with radius 1 and Ω_i, $i = 1,\ldots,5$, are cubes in right position with center at $(0,-4,0)$, $(-4,2,-3)$, $(-3,-4,2)$, $(-5,4,4)$ and $(-1,8,1)$ and side length 2, respectively.

Figure 4.2 shows that also for this instance the primal-dual approaches outperform the subgradient method from [15]. In this example we used the following parameters:

- DR1: $(\forall i \in \{1, \ldots, 5\})$ $\sigma_i = 0.3$, $\tau = 2/(\sum_{j=1}^{5} \sigma_j)$, $\lambda_n = 1.5$, $x_0 = (5, 2)$, $v_{i,0} = 0$,
- DR2: $(\forall i \in \{1, \ldots, 5\})$ $\sigma_i = 0.2$, $\tau = 0.24/(\sum_{j=1}^{5} \sigma_j)$, $\lambda_n = 1.8$, $x_0 = (5, 2)$, $v_{i,0} = 0$,
- Subgradient (cf. [14, Theorem 4.1]) $x_0 = (0, 2, 0)$, $\alpha_n = \frac{1}{n}$.

(a) Problem with optimizer (b) Progress of the RMSE values

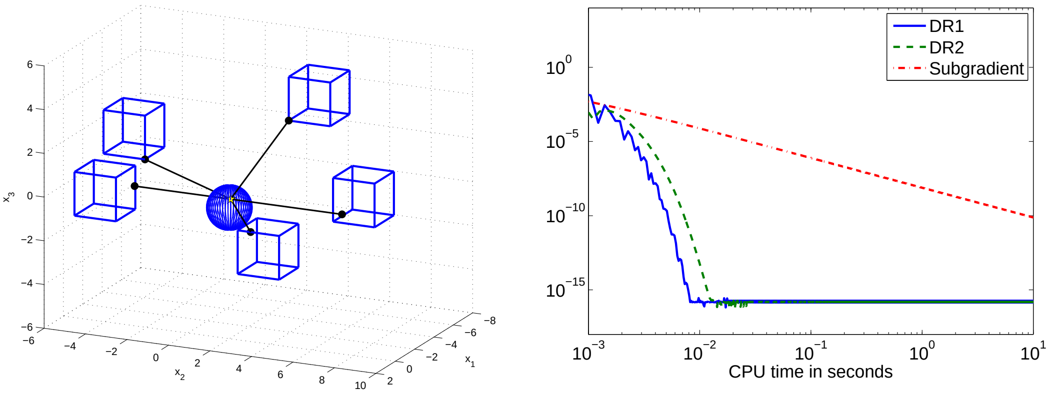


Figure 4.2: Example 4.2. Generalized Heron problem with cubes and ball constraint set on the left-hand side, performance evaluation for the RMSE on the right-hand side.

4.2 Image denoising

The second numerical experiment concerns the solving of a problem arising in image denoising. To this end, we consider images of size $M \times N$ as vectors $x \in \mathbb{R}^n$ for $n = MN$, where each pixel denoted by $x_{i,j}$, $1 \leq i \leq M$, $1 \leq j \leq N$, ranges in the closed interval from 0 to 1.

We are solving the regularized convex nondifferentiable problem

$$\inf_{x \in \mathbb{R}^n} \left\{ \frac{1}{2} \|x - b\|^2 + \lambda TV(x) \right\}, \tag{4.3}$$

where $b \in \mathbb{R}^n$ is our given noisy image. Here, $\lambda \in \mathbb{R}_{++}$ is a regularization parameter and $TV : \mathbb{R}^n \to \mathbb{R}$ is the discrete anisotropic total variation function.

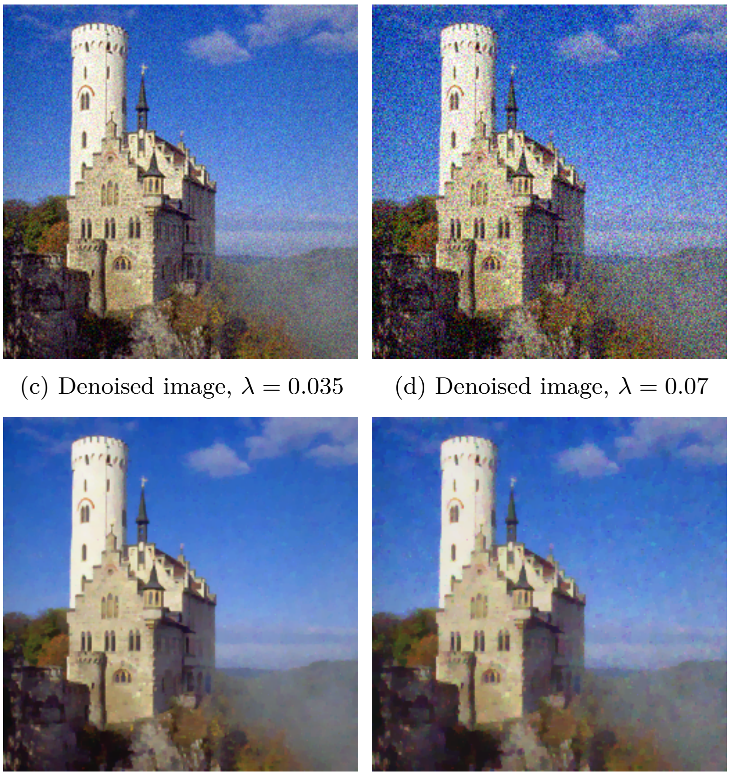
We let $\mathcal{Y} = \mathbb{R}^n \times \mathbb{R}^n$, then the operator $L : \mathbb{R}^n \to \mathcal{Y}$, $x_{i,j} \mapsto (L_1 x_{i,j}, L_2 x_{i,j})$,

$$L_1 x_{i,j} = \begin{cases} x_{i+1,j} - x_{i,j}, & \text{if } i < M \\ 0, & \text{if } i = M \end{cases} \text{ and } L_2 x_{i,j} = \begin{cases} x_{i,j+1} - x_{i,j}, & \text{if } j < N \\ 0, & \text{if } j = N \end{cases},$$

represents a discretization of the gradient in horizontal and vertical direction for each $x \in \mathbb{R}^n$ with $\|L\| \leq \sqrt{8}$. Furthermore, considering the discrete anisotropic total variation functional

$$\begin{aligned} TV(x) = & \sum_{i=1}^{M-1} \sum_{j=1}^{N-1} |x_{i+1,j} - x_{i,j}| + |x_{i,j+1} - x_{i,j}| \\ & + \sum_{i=1}^{M-1} |x_{i+1,N} - x_{i,N}| + \sum_{j=1}^{N-1} |x_{M,j+1} - x_{M,j}|, \end{aligned}$$

where reflexive boundary conditions are assumed, it holds that $TV(x) = \|Lx\|_1$.



(a) Noisy image, $\sigma = 0.06$ (b) Noisy image, $\sigma = 0.12$ (c) Denoised image, $\lambda = 0.035$ (d) Denoised image, $\lambda = 0.07$

Figure 4.3: The noisy image in (a) was obtained after adding white Gaussian noise with standard deviation $\sigma = 0.06$ to the original 256×256 lichtenstein test image, (c) shows the denoised image for $\lambda = 0.035$. Likewise, the noisy image when choosing $\sigma = 0.12$ and the denoised one for $\lambda = 0.07$ are shown in (b) and (d), respectively.

Consequently, the optimization problem (4.3) can be equivalently written as

$$\inf_{x \in \mathbb{R}^n} \{f(x) + g(Lx)\}, \tag{4.4}$$

where $f : \mathbb{R}^n \to \mathbb{R}$, $f(x) = \frac{1}{2}\|x - b\|^2$, and $g : \mathcal{Y} \to \mathbb{R}$, $g(v, w) = \lambda\|(v, w)\|_1$. Since for every $(p, q) \in \mathcal{Y}$ we have $g^*(p, q) = \delta_S(p, q)$ (see [3]), $S \subseteq \mathcal{Y}$ being the closed convex set

$$S = [-\lambda, \lambda]^n \times [-\lambda, \lambda]^n.$$

It is easy to see that for all x, p, $q \in \mathbb{R}^n$ and τ, $\sigma \in \mathbb{R}_{++}$, it holds

$$\operatorname{Prox}_{\tau f}(x) = \operatorname*{arg\,min}_{z \in \mathbb{R}^n} \left\{ \frac{\tau}{2}\|z - b\|^2 + \frac{1}{2}\|z - x\|^2 \right\} = \frac{1}{1 + \tau}(x + \tau b)$$

$$\operatorname{Prox}_{\sigma g^*}(p, q) = \operatorname*{arg\,min}_{(z_1, z_2) \in S} \frac{1}{2}\|(z_1, z_2) - (p, q)\|^2 = \mathcal{P}_S(p, q).$$

We solved the regularized image denoising problem with the two Douglas-Rachford type primal-dual methods DR1 (Algorithm 3.1) and DR2 (Algorithm 3.2), the forward-backward-forward type primal dual method (FBF, cf. [12]) and its acceleration (FBF Acc, cf. [6]), the primal-dual method (PD) and its accelerated version (PD Acc), both given in [9], the alternating minimization algorithm (AMA) from [18] together with its Nesterov-type acceleration (cf. [17]), as well as the Nesterov algorithm (cf. [16]) and the FISTA algorithm (cf. [2]), the latter operating on the dual problem. A comparison of the obtained results is shown in Table 4.1 while the test images can be found in Figure 4.3. It is noticeable that, especially when a low level of tolerance is assumed, the primal-dual algorithms proposed in this paper definitely outperform the other methods.

	$\sigma = 0.12,\ \lambda = 0.07$		$\sigma = 0.06,\ \lambda = 0.035$	
	$\varepsilon = 10^{-4}$	$\varepsilon = 10^{-6}$	$\varepsilon = 10^{-4}$	$\varepsilon = 10^{-6}$
DR1	1.40s (48)	3.35s (118)	1.31s (45)	2.93s (103)
DR2	1.24s (75)	2.82s (173)	1.12s (66)	2.57s (147)
FBF	8.89s (343)	58.96s (2271)	4.86s (187)	41.21s (1586)
FBF Acc	2.63s (101)	11.73s (451)	1.93s (73)	8.07s (308)
PD	5.26s (337)	35.53s (2226)	2.77s (183)	25.53s (1532)
PD Acc	1.42s (96)	7.26s (447)	1.20s (70)	5.44s (319)
AMA	7.29s (471)	46.76s (3031)	3.98s (254)	34.36s (2184)
AMA Acc	1.83s (89)	11.68s (561)	1.41s (63)	8.39s (383)
Nesterov	1.97s (102)	12.45s (595)	1.51s (72)	8.77s (415)
FISTA	1.71s (100)	10.92s (645)	1.14s (70)	7.41s (429)

Table 4.1: Performance evaluation for the images in Figure 4.3. The entries refer to the CPU times in seconds and the number of iterations, respectively, needed in order to attain a *root-mean-square error* (RMSE) for the iterates below the tolerance ε.

Acknowledgements. The authors are grateful to their colleague E.R. Csetnek and to the anonymous reviewers for remarks which improved the quality of the paper.

References

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

[2] A. Beck and M. Teboulle. A fast iterative shrinkage-tresholding algorithm for linear inverse problems. *SIAM J. Imaging Sci.* **2**(1), 183–202, 2009.

[3] R.I. Boţ. *Conjugate Duality in Convex Optimization*. Lecture Notes in Economics and Mathematical Systems, Vol. 637, Springer, Berlin, 2010.

[4] R.I. Boţ, E.R. Csetnek and A. Heinrich. A primal-dual splitting algorithm for finding zeros of sums of maximally monotone operators. *arXiv:1206.5953v1 [math.OC]*, 2012.

[5] R.I. Boţ, S.M. Grad and G. Wanka. *Duality in Vector Optimization*. Springer, Berlin, 2009.

[6] R.I. Boţ and C. Hendrich. Convergence analysis for a primal-dual monotone + skew splitting algorithm with applications to total variation minimization. *arXiv:1211.1706v1 [math.OC]*, 2012.

[7] R.I. Boţ and C. Hendrich. An algorithm for solving monotone inclusions involving parallel sums of linearly composed maximally monotone operators. *arXiv:1306.3191v1 [math.FA]*, 2013.

[8] L.M. Briceño-Arias and P.L. Combettes. A monotone + skew splitting model for composite monotone inclusions in duality. *SIAM J. Optim.* **21**(4), 1230–1250, 2011.

[9] A. Chambolle and T. Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. *J. Math. Imaging Vis.* **40**(1), 120–145, 2011.

[10] P.L. Combettes. Solving monotone inclusions via compositions of nonexpansive averaged operators. *Optimization*, **53**(5–6), 475–504, 2004.

[11] P.L. Combettes. Iterative construction of the resolvent of a sum of maximal monotone operators. *J. Convex Anal.* **16**(3–4), 727–748, 2009.

[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**(2), 307–330, 2012.

[13] L. Condat. A primal-dual splitting method for convex optimization involving Lipschitzian, proximable and linear composite terms. *J. Optim. Theory Appl.* **158**(2), 460–479, 2012.

[14] B.S. Mordukhovich, N.M. Nam and J. Salinas. Solving a generalized Heron problem by means of convex analysis. *Amer. Math. Monthly*, **119**(2), 87–99, 2012.

[15] B.S. Mordukhovich, N.M. Nam and J. Salinas. Applications of variational analysis to a generalized Heron problem. *Appl. Anal.* **91**(10), 1915–1942, 2012.

[16] Y. Nesterov. Smooth minimization of non-smooth functions. *Math. Program.* **103**(1), 127–152, 2005.

[17] Y. Nesterov. A method for unconstrained convex minimization problem with the rate of convergence $\mathcal{O}(1/k^2)$. *Doklady AN SSSR (translated as Soviet Math. Docl.)* **269**, 543–547, 1983.

[18] P. Tseng. Applications of a splitting algorithm to decomposition in convex programming and variational inequalities. *SIAM J. Optim.* **29**(1), 119–138, 1991.

[19] B.C. Vũ. A splitting algorithm for dual monotone inclusions involving cocoercive operators. *Adv. Comp. Math.* **38**(3), 667–681, 2013.

[20] C. Zălinescu. *Convex Analysis in General Vector Spaces.* World Scientific, Singapore, 2002.