Inducing strong convergence into the asymptotic behaviour of proximal splitting algorithms in Hilbert spaces

Radu Ioan Boţ [*]   Ernö Robert Csetnek [†]   Dennis Meier [‡]

November 18, 2017

Abstract. Proximal splitting algorithms for monotone inclusions (and convex optimization problems) in Hilbert spaces share the common feature to guarantee for the generated sequences in general weak convergence to a solution. In order to achieve strong convergence, one usually needs to impose more restrictive properties for the involved operators, like strong monotonicity (respectively, strong convexity for optimization problems). In this paper, we propose a modified Krasnosel'skiĭ–Mann algorithm in connection with the determination of a fixed point of a nonexpansive mapping and show strong convergence of the iteratively generated sequence to the minimal norm solution of the problem. Relying on this, we derive a forward-backward and a Douglas-Rachford algorithm, both endowed with Tikhonov regularization terms, which generate iterates that strongly converge to the minimal norm solution of the set of zeros of the sum of two maximally monotone operators. Furthermore, we formulate strong convergent primal-dual algorithms of forward-backward and Douglas-Rachford-type for highly structured monotone inclusion problems involving parallel-sums and compositions with linear operators. The resulting iterative schemes are particularized to the solving of convex minimization problems. The theoretical results are illustrated by numerical experiments on the split feasibility problem in infinite dimensional spaces.

Key Words. fixed points of nonexpansive mappings, Tikhonov regularization, splitting methods, forward-backward algorithm, Douglas-Rachford algorithm, primal-dual algorithm

AMS subject classification. 47J25, 47H09, 47H05, 90C25

1 Introduction and preliminaries

Let $\mathcal{H}$ be a real Hilbert space endowed with inner product $\langle \cdot, \cdot \rangle$ and associated norm $\|\cdot\| = \sqrt{\langle \cdot, \cdot \rangle}$. Let $T : \mathcal{H} \to \mathcal{H}$ be a nonexpansive mapping, that is $\|Tx - Ty\| \leq \|x - y\|$ for all $x, y \in \mathcal{H}$. One of the most popular iterative methods for finding a fixed point of the operator T is the Krasnosel'skiĭ–Mann algorithm

$$x_{n+1} = x_n + \lambda_n (Tx_n - x_n) \ \forall n \geq 0, \tag{1}$$

where $x_0 \in \mathcal{H}$ is arbitrary and $(\lambda_n)_{n \geq 0}$ is a sequence of nonnegative real numbers. Provided Fix $T = \{x \in \mathcal{H} : Tx = x\} \neq \emptyset$, one can show under mild conditions imposed on $(\lambda_n)_{n \geq 0}$, that the sequence $(x_n)_{n \geq 0}$ converges weakly to an element in Fix T (see for instance [6]).

The applications and the impact of this fundamental result go beyond the usual fixed point theory, representing in fact the starting point for the derivation of algorithms and related convergence statements in connection with the solving of monotone inclusions. In this context, we mention the classical *forward-backward algorithm* for determining a zero of the sum of a set-valued maximally monotone operator and a single-valued and cocoercive one and the *Douglas-Rachford algorithm* for determining a zero of the sum of two set-valued maximally monotone operators. The paradigms behind these classical methods can be transferred to the solving of convex optimization problems, too (see [6]).

The iterative algorithms mentioned above share the common property that the generated sequences converge weakly to a solution of the problem under investigation. However, for applications where infinite dimensional functional spaces are involved, weak convergence is not satisfactory. In order to achieve strong convergence, one usually needs to impose more restrictive properties for the involved operators, like strong monotonicity when considering monotone inclusions and strong convexity when solving optimization problems. Since there evidently are applications for which these stronger properties are not fulfilled, the interest of the applied mathematics community in developing algorithms which generate iterates that strongly convergence is justified.

We mention in this sense the *Halpern algorithm* and its numerous variants designed for finding a fixed point of a nonexpansive mapping (see for instance [13]). In the context of solving monotone inclusions, we mention the *proximal-Tikhonov algorithm*

$$x_{n+1} = \big(\operatorname{Id} + \lambda_n(A + \mu_n \operatorname{Id}) \big)^{-1}(x_n) \ \forall n \geq 0,$$

where $A : \mathcal{H} \rightrightarrows \mathcal{H}$ is a maximally monotone operator, Id is the identity operator on $\mathcal{H}$ and $(\lambda_n)_{n \geq 0}$ and $(\mu_n)_{n \geq 0}$ are sequences of nonnegative real numbers. Under mild conditions imposed on $(\lambda_n)_{n \geq 0}$ and $(\mu_n)_{n \geq 0}$, one can prove strong convergence of $(x_n)_{n \geq 0}$ to the minimal norm solution of the set of zeros of A (see [19, 27]). It is important to emphasize the Tikhonov regularization terms $(\mu_n \operatorname{Id})_{n \geq 0}$ in the above scheme, which actually enforces the strong convergence property. In the absence of the regularization term, the above numerical scheme becomes the classical *proximal algorithm* for determining a zero of the operator A, for which in general only weak convergence can be proved (see [23]). For more theoretical results concerning Tikhonov regularization and more motivational arguments for using such techniques, especially for optimization problems, we refer the reader to Attouch's paper [2]. For other techniques and tools in order to achieve strong convergence we mention also the works of Haugazeau [18] and [5].

In this article, we will first introduce and investigate a modified Krasnosel'skiĭ–Mann algorithm with relaxation parameters, having the outstanding property that it generates a sequence of iterates which converges strongly to the minimal norm solution of the fixed points set of a nonexpansive mapping. In contrast to [13, 20, 27] (see also the references therein), where the techniques and tools used have their roots in fixed point theory results for contractions, our convergence statements follow more directly. Relying on this, we derive a forward-backward and a Douglas-Rachford algorithm, both endowed with Tikhonov regularization terms, which generate iterates that strongly converge to the minimal norm solution of the set of zeros of the sum of two maximally monotone operators. The resulting iterative schemes are particularized to the minimization of the sum of two convex functions.

Furthermore, we deal with complexly structured monotone inclusions where parallel-sums and compositions with linear operators are involved. By making use of modern primal-dual

techniques (see [10,12,15,17,28] and also [8,9]), we derive strongly convergent numerical schemes of forward-backward and Douglas-Rachford type, both involving Tikhonov regularization terms and having the remarkable property that all the operators are evaluated separately. Moreover, the designed algorithms solve both the structured monotone inclusion problem and its dual monotone inclusion problem in the sense of Attouch-Théra (see [4]). When particularized to convex optimization problems, this means the concomitantly solving of a primal problem and its Fenchel dual one. For other types of primal-dual algorithms with strong convergence properties and their applications we refer the reader to [1, 3]. Finally, in the last section we carry out numerical experiments on the split feasibility problem in infinite dimensional Hilbert spaces which illustrate the potential of the algorithm endowed with Tikhonov regularization terms.

In the remaining of this section, we recall some results which will play a decisive role in the convergence analysis of the proposed algorithms. The following result is related to the convergence of a sequence satisfying a sharp quasi-Fejér monotonicity property and follows as a direct consequence of [27, Lemma 2.5].

Lemma 1 *Let $(a_n)_{n\geq 0}$ be a sequence of non-negative real numbers satisfying the inequality*

$$a_{n+1} \leq (1 - \theta_n)a_n + \theta_n b_n + \varepsilon_n \ \forall n \geq 0,$$

where
   (i) $0 \leq \theta_n \leq 1$ for all $n \geq 0$ and $\sum_{n\geq 0} \theta_n = +\infty$;
   (ii) $\limsup_{n\to+\infty} b_n \leq 0$;
   (iii) $\varepsilon_n \geq 0$ for all $n \geq 0$ and $\sum_{n\geq 0} \varepsilon_n < +\infty$.
   Then the sequence $(a_n)_{n\geq 0}$ converges to 0.

We close this section with a result that is a consequence of the demiclosedness principle (see [6, Corollary 4.18]) and it will be used in the proof of Theorem 3, which is the main result of this paper.

Lemma 2 *Let $T : \mathcal{H} \to \mathcal{H}$ be a nonexpansive operator and let $(x_n)_{n\geq 0}$ be a sequence in $\mathcal{H}$ and $x \in \mathcal{H}$ be such that $w - \lim_{n\to+\infty} x_n = x$ and $(Tx_n - x_n)_{n\geq 0}$ converges strongly to 0 as $n \to +\infty$. Then $x \in \operatorname{Fix} T$.*

2 A strongly convergent Krasnosel'skiĭ–Mann algorithm

In order to induce strong convergence into the asymptotic behaviour of the Krasnosel'skiĭ–Mann algorithm for determining a fixed point of a nonexpansive mapping $T : \mathcal{H} \to \mathcal{H}$, we propose the following modified version of it:

$$x_{n+1} = \beta_n x_n + \lambda_n \big(T(\beta_n x_n) - \beta_n x_n\big) \ \forall n \geq 0, \tag{2}$$

where $x_0 \in \mathcal{H}$ is the starting point and $(\lambda_n)_{n\geq 0}$ and $(\beta_n)_{n\geq 0}$ suitably chosen sequences of positive numbers. In the proof of the theorem below, we denote by $\operatorname{proj}_C : \mathcal{H} \to C, \operatorname{proj}_C(x) = \operatorname*{argmin}_{c\in C} \|x - c\|$, the projection operator onto the nonempty closed convex set $C \subseteq \mathcal{H}$. We notice that for a nonexpansive mapping $T : \mathcal{H} \to \mathcal{H}$, its set of fixed points $\operatorname{Fix} T$ is closed and convex (see [6, Corollary 4.15]).

Theorem 3 *Let $(\lambda_n)_{n\geq 0}$ and $(\beta_n)_{n\geq 0}$ be real sequences satisfying the conditions:*
   (i) $0 < \beta_n \leq 1$ for any $n \geq 0$, $\lim_{n\to+\infty} \beta_n = 1$, $\sum_{n\geq 0}(1-\beta_n) = +\infty$ and $\sum_{n\geq 1} |\beta_n - \beta_{n-1}| < +\infty$;

(ii) $0 < \lambda_n \le 1$ *for any* $n \ge 0$, $\liminf_{n \to +\infty} \lambda_n > 0$ *and* $\sum_{n \ge 1} |\lambda_n - \lambda_{n-1}| < +\infty$.

Consider the iterative scheme (2) with and arbitrary starting point $x_0 \in \mathcal{H}$ *and a nonexpansive mapping* $T : \mathcal{H} \to \mathcal{H}$ *fulfilling* $\mathrm{Fix}\, T \ne \emptyset$. *Then* $(x_n)_{n \ge 0}$ *converges strongly to* $\mathrm{proj}_{\mathrm{Fix}\, T}(0)$.

Proof. For the beginning, we prove that $(x_n)_{n \ge 0}$ is bounded. Let $x \in \mathrm{Fix}\, T$. Due to the nonexpansiveness of T, we have for any $n \ge 0$:

$$
\begin{aligned}
\|x_{n+1} - x\| &= \|(1 - \lambda_n)(\beta_n x_n - x) + \lambda_n\big(T(\beta_n x_n) - Tx\big)\| \\
&\le (1 - \lambda_n)\|\beta_n x_n - x\| + \lambda_n \|T(\beta_n x_n) - Tx\| \\
&\le \|\beta_n x_n - x\| \\
&= \|\beta_n(x_n - x) + (\beta_n - 1)x\| \\
&\le \beta_n \|x_n - x\| + (1 - \beta_n)\|x\|.
\end{aligned}
$$

A simple induction leads to the inequality

$$
\|x_n - x\| \le \max\{\|x_0 - x\|, \|x\|\} \ \forall n \ge 0,
$$

hence $(x_n)_{n \ge 0}$ is bounded.

We claim that

$$
\|x_{n+1} - x_n\| \to 0 \text{ as } n \to +\infty. \tag{3}
$$

Indeed, by taking into account that T is nonexpansive and that $(x_n)_{n \ge 0}$ is bounded, we obtain for any $n \ge 1$ the following estimates:

$$
\begin{aligned}
\|x_{n+1} - x_n\| &= \|(1 - \lambda_n)\beta_n x_n - (1 - \lambda_{n-1})\beta_{n-1} x_{n-1} + \lambda_n T(\beta_n x_n) - \lambda_{n-1} T(\beta_{n-1} x_{n-1})\| \\
&\le \|(1 - \lambda_n)(\beta_n x_n - \beta_{n-1} x_{n-1}) + (\lambda_{n-1} - \lambda_n)\beta_{n-1} x_{n-1}\| \\
&\quad + \|\lambda_n\big(T(\beta_n x_n) - T(\beta_{n-1} x_{n-1})\big) + (\lambda_n - \lambda_{n-1})T(\beta_{n-1} x_{n-1})\| \\
&\le \|\beta_n x_n - \beta_{n-1} x_{n-1}\| + |\lambda_n - \lambda_{n-1}| C_1,
\end{aligned}
$$

where $C_1 > 0$.

Further, we derive for any $n \ge 1$:

$$
\begin{aligned}
\|x_{n+1} - x_n\| &\le \|\beta_n(x_n - x_{n-1}) + (\beta_n - \beta_{n-1})x_{n-1}\| + |\lambda_n - \lambda_{n-1}| C_1 \\
&\le \beta_n \|x_n - x_{n-1}\| + |\beta_n - \beta_{n-1}| C_2 + |\lambda_n - \lambda_{n-1}| C_1,
\end{aligned}
$$

where $C_2 > 0$. Statement (3) is a consequence of Lemma 1, for $a_n := \|x_n - x_{n-1}\|, b_n := 0, \varepsilon_n := |\beta_n - \beta_{n-1}| C_2 + |\lambda_n - \lambda_{n-1}| C_1$ and $\theta_n := 1 - \beta_n, n \ge 1$.

In the following we prove that

$$
\|x_n - Tx_n\| \to 0 \text{ as } n \to +\infty. \tag{4}
$$

For any $n \ge 0$ we have the following inequalities:

$$
\begin{aligned}
\|x_n - Tx_n\| &\le \|x_{n+1} - x_n\| + \|x_{n+1} - Tx_n\| \\
&= \|x_{n+1} - x_n\| + \|(1 - \lambda_n)(\beta_n x_n - Tx_n) + \lambda_n\big(T(\beta_n x_n) - Tx_n\big)\| \\
&\le \|x_{n+1} - x_n\| + (1 - \lambda_n)\|\beta_n x_n - Tx_n\| + \lambda_n \|\beta_n x_n - x_n\| \\
&\le \|x_{n+1} - x_n\| + (1 - \lambda_n)\|\beta_n x_n - \beta_n Tx_n\| \\
&\quad + (1 - \lambda_n)\|\beta_n Tx_n - Tx_n\| + \lambda_n(1 - \beta_n)\|x_n\| \\
&\le \|x_{n+1} - x_n\| + (1 - \lambda_n)\|x_n - Tx_n\| \\
&\quad + (1 - \lambda_n)(1 - \beta_n)\|Tx_n\| + \lambda_n(1 - \beta_n)\|x_n\|.
\end{aligned}
$$

From here we deduce that for any $n \geq 0$:

$$\lambda_n \|x_n - Tx_n\| \leq \|x_{n+1} - x_n\| + (1 - \lambda_n)(1 - \beta_n)\|Tx_n\| + \lambda_n(1 - \beta_n)\|x_n\|.$$

Taking into account that $(x_n)_{n \geq 0}$ is bounded, (3) and the properties of the sequences involved, we derive from the last inequality that (4) holds.

In what follows we show that $(x_n)_{n \geq 0}$ actually converges strongly to $\text{proj}_{\text{Fix}\,T}(0) := \overline{x}$. Since T is nonexpansive, we have for any $n \geq 0$:

$$\begin{aligned}
\|x_{n+1} - \overline{x}\| &= \|(1 - \lambda_n)(\beta_n x_n - \overline{x}) + \lambda_n\big(T(\beta_n x_n) - T\overline{x}\big)\| \\
&\leq (1 - \lambda_n)\|\beta_n x_n - \overline{x}\| + \lambda_n\|T(\beta_n x_n) - T\overline{x}\| \\
&\leq \|\beta_n x_n - \overline{x}\|.
\end{aligned}$$

Hence,

$$\begin{aligned}
\|x_{n+1} - \overline{x}\|^2 &\leq \|\beta_n x_n - \overline{x}\|^2 \\
&= \|\beta_n(x_n - \overline{x}) + (\beta_n - 1)\overline{x}\|^2 \\
&= \beta_n^2 \|x_n - \overline{x}\|^2 + 2\beta_n(1 - \beta_n)\langle -\overline{x}, x_n - \overline{x}\rangle + (1 - \beta_n)^2\|\overline{x}\|^2 \\
&\leq \beta_n \|x_n - \overline{x}\|^2 + (1 - \beta_n)\big(2\beta_n\langle -\overline{x}, x_n - \overline{x}\rangle + (1 - \beta_n)\|\overline{x}\|^2\big) \; \forall n \geq 0. \quad (5)
\end{aligned}$$

Next we show that

$$\limsup_{n \to +\infty}\langle -\overline{x}, x_n - \overline{x}\rangle \leq 0. \quad (6)$$

Assuming the contrary, there would exist a positive real number l and a subsequence $(x_{k_j})_{j \geq 0}$ such that

$$\langle -\overline{x}, x_{k_j} - \overline{x}\rangle \geq l > 0 \; \forall j \geq 0.$$

Due to the boundedness of the sequence $(x_n)_{n \geq 0}$, we can assume without losing the generality that $(x_{k_j})_{j \geq 0}$ weakly converges to an element $y \in \mathcal{H}$. According to Lemma 2, by taking into consideration (4), it follows that $y \in \text{Fix}\,T$. From this and the variational characterization of the projection we easily derive

$$\lim_{j \to +\infty}\langle -\overline{x}, x_{k_j} - \overline{x}\rangle = \langle -\overline{x}, y - \overline{x}\rangle \leq 0,$$

which leads to a contradiction. This shows that (6) holds. Thus

$$\limsup_{n \to +\infty}\big(2\beta_n\langle -\overline{x}, x_n - \overline{x}\rangle + (1 - \beta_n)\|\overline{x}\|^2\big) \leq 0.$$

A direct application of Lemma 1 to (5), for $a_n := \|x_n - \overline{x}\|^2, b_n := 2\beta_n\langle -\overline{x}, x_n - \overline{x}\rangle + (1 - \beta_n)\|\overline{x}\|^2, \varepsilon_n := 0$ and $\theta_n := 1 - \beta_n, n \geq 0$, delivers the desired conclusion. ∎

Remark 4 Condition (ii) in the previous theorem is satisfied by every monotonically increasing (and, in consequence, convergent) sequence $(\lambda_n)_{n \geq 0} \subseteq (0, 1]$ and also by every monotonically decreasing (and, in consequence, convergent) sequence $(\lambda_n)_{n \geq 0} \subseteq (0, 1]$ having as limit a positive number.

Condition (i) in the previous theorem is satisfied by every monotonically increasing sequence $(\beta_n)_{n \geq 0} \subseteq (0, 1]$ which fulfills $\lim_{n \to +\infty}\beta_n = 1$ and $\sum_{n \geq 0}(1 - \beta_n) = +\infty$, as it is for instance the sequence with $\beta_0 \in \left(0, \frac{1}{2}\right)$ and $\beta_n = 1 - \frac{1}{n+1}$ for any $n \geq 1$.

An immediate consequence of Theorem 3 is the following corollary, which proposes an iterative scheme that finds a minimal norm solution of the set of fixed points of an averaged operator. Let $\alpha \in (0,1)$ be fixed. We say that $R : \mathcal{H} \to \mathcal{H}$ is an α-averaged operator if there exists a nonexpansive operator $T : \mathcal{H} \to \mathcal{H}$ such that $R = (1-\alpha)\operatorname{Id} + \alpha T$. It is obvious that α-averaged operators are also nonexpansive. The $\frac{1}{2}$-averaged operators are nothing else than the firmly nonexpansive ones and form the most most important representatives of this class. For properties and insights into these families of operators we refer the reader to [6]. The following result will play in the next section a determinant role in the converge analysis of the forward-backward method endowed with Tikhonov regularization term.

Corollary 5 *Consider the iterative scheme*

$$x_{n+1} = \beta_n x_n + \lambda_n \big(R(\beta_n x_n) - \beta_n x_n \big) \ \forall n \geq 0, \tag{7}$$

with $x_0 \in \mathcal{H}$ as starting point, $R : \mathcal{H} \to \mathcal{H}$ an α-averaged operator, for $\alpha \in (0,1)$, such that $\operatorname{Fix} R \neq \emptyset$ and $(\lambda_n)_{n\geq 0}$ and $(\beta_n)_{n\geq 0}$ real sequences satisfying the conditions:

(i) $0 < \beta_n \leq 1$ for any $n \geq 0$, $\lim_{n\to+\infty} \beta_n = 1$, $\sum_{n\geq 0}(1-\beta_n) = +\infty$ and $\sum_{n\geq 1}|\beta_n - \beta_{n-1}| < +\infty$;

(ii) $0 < \lambda_n \leq 1/\alpha$ for any $n \geq 0$, $\liminf_{n\to+\infty} \lambda_n > 0$ and $\sum_{n\geq 1}|\lambda_n - \lambda_{n-1}| < +\infty$. Then $(x_n)_{n\geq 0}$ converges strongly to $\operatorname{proj}_{\operatorname{Fix} R}(0)$.

Proof. Since R is α-averaged, there exists a nonexpansive operator $T : \mathcal{H} \to \mathcal{H}$ such that $R = (1-\alpha)\operatorname{Id} + \alpha T$. The conclusion follows from Theorem 3, by taking into account that (7) is equivalent to

$$x_{n+1} = \beta_n x_n + \alpha \lambda_n \big(T(\beta_n x_n) - \beta_n x_n \big) \ \forall n \geq 0$$

and that $\operatorname{Fix} R = \operatorname{Fix} T$. ∎

3 A forward-backward algorithm with Tikhonov regularization term

This section is dedicated to the formulation and convergence analysis of a forward-backward algorithm with Tikhonov regularization terms, which generates a sequence of iterates that converges strongly to the minimal norm solution of the set of zeros of the sum of two maximally monotone operators, one of them being single-valued.

For readers' convenience, we recall some standard notions and results in monotone operator theory which will be used in the following (see also [6,7,26]). For an arbitrary set-valued operator $A : \mathcal{H} \rightrightarrows \mathcal{H}$ we denote by $\operatorname{Gr} A = \{(x,u) \in \mathcal{H} \times \mathcal{H} : u \in Ax\}$ its graph. Then $A^{-1} : \mathcal{H} \rightrightarrows \mathcal{H}$, which is the operator with $\operatorname{Gr} A^{-1} = \{(x,u) \in \mathcal{H} \times \mathcal{H} : x \in Au\}$, denotes the inverse operator of A. We use also the notation $\operatorname{zer} A = \{x \in \mathcal{H} : 0 \in Ax\}$ for the set of zeros of A. We say that A is monotone, if $\langle x - y, u - v \rangle \geq 0$ for all $(x,u), (y,v) \in \operatorname{Gr} A$. A monotone operator A is said to be maximally monotone, if there exists no proper monotone extension of the graph of A on $\mathcal{H} \times \mathcal{H}$. The resolvent of A, $J_A : \mathcal{H} \rightrightarrows \mathcal{H}$, is defined by

$$J_A = (\operatorname{Id} + A)^{-1},$$

where $\operatorname{Id} : \mathcal{H} \to \mathcal{H}, \operatorname{Id}(x) = x$ for all $x \in \mathcal{H}$, is the identity operator on $\mathcal{H}$. Moreover, if A is maximally monotone, then $J_A : \mathcal{H} \to \mathcal{H}$ is single-valued and maximally monotone (see [6, Proposition 23.7 and Corollary 23.10]).

Let $\gamma > 0$. We say that $B : \mathcal{H} \to \mathcal{H}$ is γ-cocoercive, if $\langle x - y, Bx - By \rangle \geq \gamma \|Bx - By\|^2$ for all $x, y \in \mathcal{H}$.

The following technical result (see [21, Theorem 3(b)] and [16, Proposition 2.4]) gives an expression for the the averaged parameter of the composition of two averaged operators. We refer also to [6, Proposition 4.32] for other results of this type.

Proposition 6 *Let $T_i : \mathcal{H} \to \mathcal{H}$ be α_i-averaged, where $\alpha_i \in (0, 1)$, $i = 1, 2$. Then the composition $T_1 \circ T_2$ is α-averaged, where*

$$\alpha = \frac{\alpha_1 + \alpha_2 - 2\alpha_1\alpha_2}{1 - \alpha_1\alpha_2} \in (0, 1).$$

Theorem 7 *Let $A : \mathcal{H} \rightrightarrows \mathcal{H}$ be a maximally monotone operator and $B : \mathcal{H} \to \mathcal{H}$ a β-cocoercive operator, for $\beta > 0$, such that $\mathrm{zer}(A + B) \neq \emptyset$. Let $\gamma \in (0, 2\beta]$. Consider the iterative scheme*

$$x_{n+1} = (1 - \lambda_n)\beta_n x_n + \lambda_n J_{\gamma A}\big(\beta_n x_n - \gamma B(\beta_n x_n)\big) \ \forall n \geq 0, \qquad (8)$$

with $x_0 \in \mathcal{H}$ as starting point and $(\lambda_n)_{n \geq 0}$ and $(\beta_n)_{n \geq 0}$ real sequences satisfying the conditions:
  (i) $0 < \beta_n \leq 1$ for any $n \geq 0$, $\lim_{n \to +\infty} \beta_n = 1$, $\sum_{n \geq 0}(1 - \beta_n) = +\infty$ and $\sum_{n \geq 1}|\beta_n - \beta_{n-1}| < +\infty$;
  (ii) $0 < \lambda_n \leq \frac{4\beta - \gamma}{2\beta}$ for any $n \geq 0$, $\liminf_{n \to +\infty} \lambda_n > 0$ and $\sum_{n \geq 1}|\lambda_n - \lambda_{n-1}| < +\infty$.
  Then $(x_n)_{n \geq 0}$ converges strongly to $\mathrm{proj}_{\mathrm{zer}(A+B)}(0)$.

Proof. It is immediate that the iterative scheme (8) can be written in the form

$$x_{n+1} = \beta_n x_n + \lambda_n\big(T(\beta_n x_n) - \beta_n x_n\big) \ \forall n \geq 0,$$

where $T = J_{\gamma A} \circ (\mathrm{Id} - \gamma B)$.

We consider two cases. The first one is when $\gamma \in (0, 2\beta)$.

According to [6, Corollary 23.8 and Remark 4.24(iii)], $J_{\gamma A}$ is $\frac{1}{2}$-cocoercive. Moreover, by [6, Proposition 4.33], $\mathrm{Id} - \gamma B$ is $\frac{\gamma}{2\beta}$-averaged. Combining this with Proposition 6, we derive that T is $\frac{2\beta}{4\beta - \gamma}$-averaged. The statement follows now from Corollary 5, by noticing that $\mathrm{Fix}\, T = \mathrm{zer}(A + B)$ (see [6, Proposition 25.1(iv)]).

The second case is when $\gamma = 2\beta$. The cocoercivity of B implies that $\mathrm{Id} - \gamma B$ is nonexpansive, hence the operator $T = J_{\gamma A} \circ (\mathrm{Id} - \gamma B)$ is nonexpansive, too, the conclusion follows in this situation from Theorem 3. ∎

Remark 8 The choice $\lambda_n = 1$ for any $n \geq 0$ in the previous theorem leads to the iterative scheme

$$x_{n+1} = J_{\gamma A}\big(\beta_n x_n - \gamma B(\beta_n x_n)\big) \ \forall n \geq 0,$$

which further becomes in case $B = 0$

$$x_{n+1} = J_{\gamma A}\big(\beta_n x_n\big) \ \forall n \geq 0.$$

This last relation can be equivalently written as

$$x_n \in \frac{1}{\beta_n} x_{n+1} + \frac{\gamma}{\beta_n} A x_{n+1} = \left(\mathrm{Id} + \varepsilon_n \,\mathrm{Id} + \frac{\gamma}{\beta_n} A\right)(x_{n+1}),$$

where $\varepsilon_n \,\mathrm{Id}$ (with $\varepsilon_n := \frac{1}{\beta_n} - 1 > 0$ and $\lim_{n \to +\infty} \varepsilon_n = 0$) represents the Tikhonov regularization term, which enforces the strong convergence of the sequence $(x_n)_{n \geq 0}$ to the minimal norm solution. For other types of Tikhonov-like methods for monotone inclusion problems we refer the reader to [19, 24, 27] and the references therein.

In the remaining of this section we turn our attention to the solving of optimization problems of the form

$$\min_{x \in \mathcal{H}} \{f(x) + g(x)\}, \tag{9}$$

where $f : \mathcal{H} \to \mathbb{R} \cup \{+\infty\}$ is a proper, convex and lower semicontinuous function and $g : \mathcal{H} \to \mathbb{R}$ is a convex and Fréchet differentiable function with $\frac{1}{\beta}$-Lipschitz continuous gradient, for $\beta > 0$.

For a proper, convex and lower semicontinuous function $f : \mathcal{H} \to \mathbb{R} \cup \{+\infty\}$, its (convex) subdifferential at $x \in \mathcal{H}$ is defined as

$$\partial f(x) = \{u \in \mathcal{H} : f(y) \geq f(x) + \langle u, y - x \rangle \ \forall y \in \mathcal{H}\},$$

for $x \in \mathcal{H}$ with $f(x) = +\infty$ and as $\partial f(x) = \emptyset$, otherwise. When seen as a set-valued mapping, the convex subdifferential is a maximally monotone operator (see [22]) and its resolvent is given by $J_{\partial f} = \mathrm{prox}_f$ (see [6]), where $\mathrm{prox}_f : \mathcal{H} \to \mathcal{H}$,

$$\mathrm{prox}_f(x) = \underset{y \in \mathcal{H}}{\mathrm{argmin}} \left\{ f(y) + \frac{1}{2} \|y - x\|^2 \right\}, \tag{10}$$

denotes the proximal operator of f.

Corollary 9 *Let $f : \mathcal{H} \to \mathbb{R} \cup \{+\infty\}$ be a proper, convex and lower semicontinuous function and $g : \mathcal{H} \to \mathbb{R}$ a convex and Fréchet differentiable function with $\frac{1}{\beta}$-Lipschitz continuous gradient, for $\beta > 0$, such that $\mathrm{argmin}_{x \in \mathcal{H}}\{f(x) + g(x)\} \neq \emptyset$. Let $\gamma \in (0, 2\beta]$. Consider the iterative scheme*

$$x_{n+1} = (1 - \lambda_n)\beta_n x_n + \lambda_n \, \mathrm{prox}_{\gamma f} \left(\beta_n x_n - \gamma \nabla g(\beta_n x_n) \right) \ \forall n \geq 0, \tag{11}$$

with $x_0 \in \mathcal{H}$ as starting point and $(\lambda_n)_{n \geq 0}$ and $(\beta_n)_{n \geq 0}$ real sequences satisfying the conditions:
  (i) $0 < \beta_n \leq 1$ for any $n \geq 0$, $\lim_{n \to +\infty} \beta_n = 1$, $\sum_{n \geq 0}(1 - \beta_n) = +\infty$ and $\sum_{n \geq 1} |\beta_n - \beta_{n-1}| < +\infty$;
  (ii) $0 < \lambda_n \leq \frac{4\beta - \gamma}{2\beta}$ for any $n \geq 0$, $\liminf_{n \to +\infty} \lambda_n > 0$ and $\sum_{n \geq 1} |\lambda_n - \lambda_{n-1}| < +\infty$.
Then $(x_n)_{n \geq 0}$ converges strongly to the minimal norm solution of (9).

Proof. The statement is a direct consequence of Theorem 7, by choosing $A := \partial f$ and $B := \nabla g$ and by taking into account that

$$\mathrm{zer}(\partial f + \nabla g) = \underset{x \in \mathcal{H}}{\mathrm{argmin}} \{f(x) + g(x)\}$$

and the fact that ∇g is β-cocoercive due to the Baillon-Haddad Theorem (see [6, Corollary 18.16]). ∎

4 A Douglas-Rachford algorithm with Tikhonov regularization term

In this section we derive from the Krasnosel'skiĭ–Mann algorithm formulated in Section 2 an iterative scheme of Douglas-Rachford-type, which generates sequences that strongly converge to a zero of of the sum of two set-valued maximally monotone operators.

In what follows, we denote by $R_A = 2J_A - \mathrm{Id}$ the reflected resolvent of a maximally monotone operator $A : \mathcal{H} \rightrightarrows \mathcal{H}$.

Theorem 10 *Let $A, B : \mathcal{H} \rightrightarrows \mathcal{H}$ be two maximally monotone operators such that $\mathrm{zer}(A+B) \neq \emptyset$ and $\gamma > 0$. Consider the following iterative scheme:*

$$
(\forall n \geq 0) \quad \begin{cases} y_n = J_{\gamma B}(\beta_n x_n) \\ z_n = J_{\gamma A}(2y_n - \beta_n x_n) \\ x_{n+1} = \beta_n x_n + \lambda_n(z_n - y_n) \end{cases}
$$

with $x_0 \in \mathcal{H}$ as starting point and $(\lambda_n)_{n \geq 0}$ and $(\beta_n)_{n \geq 0}$ real sequences satisfying the conditions:

(i) $0 < \beta_n \leq 1$ for any $n \geq 0$, $\lim_{n \to +\infty} \beta_n = 1$, $\sum_{n \geq 0}(1 - \beta_n) = +\infty$ and $\sum_{n \geq 1}|\beta_n - \beta_{n-1}| < +\infty$;

(ii) $0 < \lambda_n \leq 2$ for any $n \geq 0$, $\liminf_{n \to +\infty} \lambda_n > 0$ and $\sum_{n \geq 1}|\lambda_n - \lambda_{n-1}| < +\infty$.

Then the following statements are true:

(a) $(x_n)_{n \geq 0}$ converges strongly to $\overline{x} := \mathrm{proj}_{\mathrm{Fix}\, R_{\gamma A} R_{\gamma B}}(0)$ as $n \to +\infty$;

(b) $(y_n)_{n \geq 0}$ and $(z_n)_{n \geq 0}$ converge strongly to $J_{\gamma B}(\overline{x}) \in \mathrm{zer}(A+B)$ as $n \to +\infty$.

Proof. Taking into account the iteration rules and the definition of the reflected resolvent, the iterative scheme in the enunciation of the theorem can be equivalently written as

$$
\begin{aligned}
x_{n+1} &= \beta_n x_n + \lambda_n \left[J_{\gamma A} \circ (2J_{\gamma B} - \mathrm{Id})(\beta_n x_n) - J_{\gamma B}(\beta_n x_n) \right] \\
&= \beta_n x_n + \lambda_n \left[\left(\frac{\mathrm{Id} + R_{\gamma A}}{2} \circ R_{\gamma B} \right)(\beta_n x_n) - \frac{\mathrm{Id} + R_{\gamma B}}{2}(\beta_n x_n) \right] \\
&= \beta_n x_n + \frac{\lambda_n}{2}\left(T(\beta_n x_n) - \beta_n x_n \right) \ \forall n \geq 0,
\end{aligned} \tag{12}
$$

where $T := R_{\gamma A} \circ R_{\gamma B} : \mathcal{H} \to \mathcal{H}$ is a nonexpansive operator (see [6, Corollary 23.10(ii)]). From [6, Proposition 25.1(ii)] we have $\mathrm{zer}(A+B) = J_{\gamma B}(\mathrm{Fix}\, T)$, hence $\mathrm{Fix}\, T \neq \emptyset$. By applying Theorem 3, we obtain that $(x_n)_{n \geq 0}$ converges strongly to $\overline{x} := \mathrm{proj}_{\mathrm{Fix}\, T}(0)$ as $n \to +\infty$, hence (a) holds.

Further, by taking into account the definition of the sequence $(y_n)_{n \geq 0}$ and the continuity of the resolvent operator, we obtain that $(y_n)_{n \geq 0}$ converges strongly to $J_{\gamma B}\overline{x} \in \mathrm{zer}(A+B)$ as $n \to +\infty$. Finally, by taking the limit in the recursive formula of the sequence $(x_n)_{n \geq 0}$, we obtain that $z_n - y_n$ converges strongly to 0 as $n \to +\infty$, thus (b) holds, too. ∎

Remark 11 The classical Douglas-Rachford method, which reads

$$
(\forall k \geq 0) \quad \begin{cases} y_n = J_{\gamma B}(x_n) \\ z_n = J_{\gamma A}(2y_n - x_n) \\ x_{n+1} = x_n + \lambda_n(z_n - y_n), \end{cases}
$$

produces sequences for which in general only weak convergence to a zero of the $A + B$ can be proved (see for example [6, Theorem 25.6]). In order to ensure strong convergence, one usually needs to impose restrictve conditions on the monotone operators involved, like uniform monotonicity (which is a generalization of strong monotonicity). This is not the case for the iterative scheme stated in Theorem 10, where we are able to guarantee strong convergence in the very general situation of maximally monotone operators.

Further, we look at optimization problems of the form

$$
\min_{x \in \mathcal{H}} \{ f(x) + g(x) \}, \tag{13}
$$

where $f, g : \mathcal{H} \to \mathbb{R} \cup \{+\infty\}$ are proper, convex and lower semicontinuous functions. We denote by

$$\operatorname{dom} f = \{x \in \mathcal{H} : f(x) < +\infty\}$$

the effective domain of the function f.

In order to proceed, we need the following notion. For $S \subseteq \mathcal{H}$ a convex set, we denote by

$$\operatorname{sqri} S := \{x \in S : \cup_{\lambda > 0} \lambda(S - x) \text{ is a closed linear subspace of } \mathcal{H}\}$$

its strong quasi-relative interior. Notice that we always have $\operatorname{int} S \subseteq \operatorname{sqri} S$ (in general this inclusion may be strict). If $\mathcal{H}$ is finite-dimensional, then $\operatorname{sqri} S$ coincides with $\operatorname{ri} S$, the relative interior of S, which is the interior of S with respect to its affine hull. The notion of strong quasi-relative interior belongs to the class of generalized interiority notions and plays an important role in the formulation of regularity conditions, which are needed in convex optimization in order to guarantee duality results and also subdifferential sum formulas. The one considered in the next result is the so-called Attouch-Brézis regularity condition. We refer to [6, 7, 26, 29] for more interiority notions and their impact on the duality theory.

Corollary 12 *Let $f, g : \mathcal{H} \to \mathbb{R} \cup \{+\infty\}$ be proper, convex and lower semicontinuous functions such that $\operatorname{argmin}_{x \in \mathcal{H}} \{f(x) + g(x)\} \neq \emptyset$ and $0 \in \operatorname{sqri}(\operatorname{dom} f - \operatorname{dom} g)$ and $\gamma > 0$. Consider the following iterative scheme:*

$$(\forall n \geq 0) \quad \left\{ \begin{array}{l} y_n = \operatorname{prox}_{\gamma g}(\beta_n x_n) \\ z_n = \operatorname{prox}_{\gamma f}(2y_n - \beta_n x_n) \\ x_{n+1} = \beta_n x_n + \lambda_n(z_n - y_n) \end{array} \right.$$

with $x_0 \in \mathcal{H}$ as starting point and $(\lambda_n)_{n \geq 0}$ and $(\beta_n)_{n \geq 0}$ real sequences satisfying the conditions:

(i) $0 < \beta_n \leq 1$ for any $n \geq 0$, $\lim_{n \to +\infty} \beta_n = 1$, $\sum_{n \geq 0}(1 - \beta_n) = +\infty$ and $\sum_{n \geq 1} |\beta_n - \beta_{n-1}| < +\infty$;

(ii) $0 < \lambda_n \leq 2$ for any $n \geq 0$, $\liminf_{n \to +\infty} \lambda_n > 0$ and $\sum_{n \geq 1} |\lambda_n - \lambda_{n-1}| < +\infty$.

Then the following statements are true:

(a) $(x_n)_{n \geq 0}$ converges strongly to $\overline{x} := \operatorname{proj}_{\operatorname{Fix} T}(0)$ as $n \to +\infty$, where $T = (2 \operatorname{prox}_{\gamma f} - \operatorname{Id}) \circ (2 \operatorname{prox}_{\gamma g} - \operatorname{Id})$;

(b) $(y_n)_{n \geq 0}$ and $(z_n)_{n \geq 0}$ converge strongly to $\operatorname{prox}_{\gamma g}(\overline{x}) \in \operatorname{argmin}_{x \in \mathcal{H}} \{f(x) + g(x)\}$ as $n \to +\infty$.

Proof. The results is a direct consequence of Theorem 10 for $A = \partial f$ and $B = \partial g$ and by noticing that the regularity condition $0 \in \operatorname{sqri}(\operatorname{dom} f - \operatorname{dom} g)$ ensures the relation (see [6, Proposition 7.2])

$$\operatorname{zer}(\partial f + \partial g) = \operatorname*{argmin}_{x \in \mathcal{H}} \{f(x) + g(x)\}.$$

∎

5 Strongly convergent primal-dual algorithms

The aim of this section is to induce strong convergence in the nowadays so popular *primal-dual algorithms* designed for solving highly structured monotone inclusions involving parallel sums and compositions with linear operators.

5.1 A primal-dual algorithm of forward-backward-type with Tikhonov regularization terms

In this subsection, the following monotone inclusion problem will be in the focus of our investigations.

Problem 13 Let $A : \mathcal{H} \rightrightarrows \mathcal{H}$ be a maximally monotone operator and $C : \mathcal{H} \to \mathcal{H}$ a μ-cocoercive operator, for some $\mu > 0$. Let m be a strictly positive integer and for any $i = 1, ..., m$, let $\mathcal{G}_i$ be a real Hilbert space, $B_i, D_i : \mathcal{G}_i \rightrightarrows \mathcal{G}_i$ be maximally monotone operators such that D_i are ν_i-strongly monotone, for some $\nu_i > 0$, and $L_i : \mathcal{H} \to \mathcal{G}_i$ be a nonzero linear continuous operator. The problem is to solve the primal inclusion

$$\text{find } \overline{x} \in \mathcal{H} \text{ such that } 0 \in A\overline{x} + \sum_{i=1}^{m} L_i^*(B_i \square D_i)(L_i\overline{x}) + C\overline{x}, \tag{14}$$

together with the dual inclusion of Attouch-Théra type (see [4, 15, 28])

$$\text{find } \overline{v}_1 \in \mathcal{G}_1, ..., \overline{v}_m \in \mathcal{G}_m \text{ such that } \exists x \in \mathcal{H} : \begin{cases} -\sum_{i=1}^{m} L_i^*\overline{v}_i \in Ax + Cx \\ \overline{v}_i \in (B_i \square D_i)(L_i x), \ i = 1, ..., m. \end{cases} \tag{15}$$

Some of the notations used above are to be specified. 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 the adjoint of the linear continuous operator $L_i : \mathcal{H} \to \mathcal{G}_i$, for $i = 1, \ldots, m$. We say that $D_i : \mathcal{G}_i \rightrightarrows \mathcal{G}_i$ is ν_i-strongly monotone, for some $\nu_i > 0$, if $\langle x - y, u - v \rangle \geq \nu_i \|x - y\|^2$ for all $(x, u), (y, v) \in \text{Gr } D_i, i = 1, ..., m$. The parallel sum of the set-valued operators $B_i, D_i : \mathcal{G}_i \rightrightarrows \mathcal{G}_i$ is defined as $B_i \square D_i : \mathcal{G}_i \rightrightarrows \mathcal{G}_i, B_i \square D_i = \left(B_i^{-1} + D_i^{-1}\right)^{-1}$, for $i = 1, ..., m$.

We say that $(\overline{x}, \overline{v}_1, ..., \overline{v}_m) \in \mathcal{H} \times \mathcal{G}_1 \times ... \times \mathcal{G}_m$ is a primal-dual solution to Problem 13, if

$$-\sum_{i=1}^{m} L_i^*\overline{v}_i \in A\overline{x} + C\overline{x} \text{ and } \overline{v}_i \in (B_i \square D_i)(L_i\overline{x}), \ i = 1, ..., m. \tag{16}$$

It is easy to see that, if $(\overline{x}, \overline{v}_1, ..., \overline{v}_m) \in \mathcal{H} \times \mathcal{G}_1 \times ... \times \mathcal{G}_m$ is a primal-dual solution to Problem 13, then $\overline{x}$ is a solution to (14) and $(\overline{v}_1, ..., \overline{v}_m) \in \mathcal{G}_1 \times ... \times \mathcal{G}_m$ is a solution to (15). Moreover, if $\overline{x} \in \mathcal{H}$ is a solution to (14), then there exists $(\overline{v}_1, ..., \overline{v}_m) \in \mathcal{G}_1 \times ... \times \mathcal{G}_m$ such that $(\overline{x}, \overline{v}_1, ..., \overline{v}_m)$ is a primal-dual solution to Problem 13 and, if $(\overline{v}_1, ..., \overline{v}_m) \in \mathcal{G}_1 \times ... \times \mathcal{G}_m$ is a solution to (15), then there exists $\overline{x} \in \mathcal{H}$ such that $(\overline{x}, \overline{v}_1, ..., \overline{v}_m)$ is a primal-dual solution to Problem 13.

Theorem 14 *In Problem 13, suppose that*

$$0 \in \text{ran} \left(A + \sum_{i=1}^{m} L_i^* \circ (B_i \square D_i) \circ L_i + C \right). \tag{17}$$

Let τ and σ_i, $i = 1, ..., m$, be strictly positive numbers such that

$$2 \cdot \min\{\tau^{-1}, \sigma_1^{-1}, ..., \sigma_m^{-1}\} \cdot \min\{\mu, \nu_1, ..., \nu_m\} \left(1 - \sqrt{\tau \sum_{i=1}^{m} \sigma_i \|L_i\|^2} \right) \geq 1.$$

Let be the starting point $(x_0, v_{1,0}, ..., v_{m,0}) \in \mathcal{H} \times \mathcal{G}_1 \times ... \times \mathcal{G}_m$ and set:

$$(\forall n \geq 0) \left| \begin{array}{l} p_n = J_{\tau A}\big[\beta_n x_n - \tau\big(\beta_n \sum_{i=1}^m L_i^* v_{i,n} + C(\beta_n x_n)\big)\big] \\ x_{n+1} = \beta_n x_n + \lambda_n(p_n - \beta_n x_n) \\ \text{For } i = 1,...,m \\ \left| \begin{array}{l} q_{i,n} = J_{\sigma_i B_i^{-1}}\big[\beta_n v_{i,n} + \sigma_i\big(L_i(2p_n - \beta_n x_n) - D_i^{-1}(\beta_n v_{i,n})\big)\big] \\ v_{i,n+1} = \beta_n v_{i,n} + \lambda_n(q_{i,n} - \beta_n v_{i,n}) \end{array} \right. \end{array} \right. \tag{18}$$

where $(\lambda_n)_{n\geq0}$ and $(\beta_n)_{n\geq0}$ are real sequences satisfying the conditions:

(i) $0 < \beta_n \leq 1$ for any $n \geq 0$, $\lim_{n\to+\infty} \beta_n = 1$, $\sum_{n\geq0}(1-\beta_n) = +\infty$ and $\sum_{n\geq1}|\beta_n - \beta_{n-1}| < +\infty$;

(ii) $0 < \lambda_n \leq \frac{4\beta\rho-1}{2\beta\rho}$ for any $n \geq 0$, $\liminf_{n\to+\infty} \lambda_n > 0$ and $\sum_{n\geq1}|\lambda_n - \lambda_{n-1}| < +\infty$, for

$$\beta = \min\{\mu, \nu_1, ..., \nu_m\}$$

and

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

Then there exists a primal-dual solution $(\overline{x}, \overline{v}_1, ..., \overline{v}_m)$ to Problem 13 such that the sequence of primal-dual iterates $(x_n, v_{1,n}, ..., v_{m,n})$ converges strongly to $(\overline{x}, \overline{v}_1, ..., \overline{v}_m)$ as $n \to +\infty$.

Remark 15 (i) Since $D_i : \mathcal{G}_i \rightrightarrows \mathcal{G}_i$ is ν_i-strongly monotone, we have that $D_i^{-1} : \mathcal{G}_i \to \mathcal{G}_i$ is ν_i-cocoercive, for $i = 1, ..., m$.

(ii) The resolvent of the inverse operator of a maximally monotone operator $M : \mathcal{H} \rightrightarrows \mathcal{H}$ can be computed as follows (see [6]):

$$\text{Id} = J_{\gamma M} + \gamma J_{\gamma^{-1}M^{-1}} \circ \gamma^{-1}\,\text{Id}. \tag{19}$$

Proof. The idea is to apply Theorem 7 in an appropriate product space under the use of appropriate renorming techniques (see [28]). We consider the Hilbert space $\mathcal{K} = \mathcal{H} \times \mathcal{G}_1 \times ... \times \mathcal{G}_m$ endowed with inner product and associated norm defined, for $(x, v_1,..., v_m), (y, q_1,..., q_m) \in \mathcal{K}$, via

$$\langle (x, v_1,..., v_m), (y, q_1,..., q_m)\rangle_{\mathcal{K}} = \langle x, y\rangle_{\mathcal{H}} + \sum_{i=1}^m \langle v_i, q_i\rangle_{\mathcal{G}_i}$$

$$\text{and } \|(x, v_1,..., v_m)\|_{\mathcal{K}} = \sqrt{\|x\|_{\mathcal{H}}^2 + \sum_{i=1}^m \|v_i\|_{\mathcal{G}_i}^2}, \tag{20}$$

respectively. Furthermore, we consider the set-valued operator

$$\boldsymbol{M} : \mathcal{K} \rightrightarrows \mathcal{K}, \quad (x, v_1,..., v_m) \mapsto (Ax, B_1^{-1}v_1,..., B_m^{-1}v_m),$$

which is maximally monotone, since A and B_i, $i = 1,..., m$, are maximally monotone (see [6, Proposition 20.22 and Proposition 20.23]), and the linear continuous operator

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

which is skew-symmetric (i.e. $\boldsymbol{S}^* = -\boldsymbol{S}$) and hence maximally monotone (see [6, Example 20.30]). We also consider the single-valued operator

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

which is once again maximally monotone, since C and D_i are maximally monotone for $i = 1,..., m$. Therefore, since $\text{dom}\,\boldsymbol{S} = \mathcal{K}$, $\boldsymbol{M} + \boldsymbol{S}$ is maximally monotone (see [6, Corollary 24.4(i)]). According to [28, page 672]

$$\boldsymbol{Q} \text{ is } \beta - \text{cocoercive}.$$

Further, one can easily verify that (17) is equivalent to $\text{zer}\,(\boldsymbol{M} + \boldsymbol{S} + \boldsymbol{Q}) \neq \varnothing$ and (see also [15, page 317])

$$\begin{aligned} &(x, v_1,..., v_m) \in \text{zer}\,(\boldsymbol{M} + \boldsymbol{S} + \boldsymbol{Q}) \\ \Leftrightarrow &(x, v_1,..., v_m) \text{ is a primal-dual solution to Problem 13.} \end{aligned} \tag{21}$$

We also introduce the linear continuous operator

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

which is self-adjoint and ρ-strongly positive (see [28]), namely, the following inequality holds

$$\langle \boldsymbol{x}, \boldsymbol{V}\boldsymbol{x}\rangle_{\mathcal{K}} \geq \rho\|\boldsymbol{x}\|_{\mathcal{K}}^2 \ \forall \boldsymbol{x} \in \mathcal{K}.$$

Therefore, its inverse operator $\boldsymbol{V}^{-1}$ exists and it fulfills $\|\boldsymbol{V}^{-1}\| \leq \frac{1}{\rho}$.

The algorithmic scheme (18) in the statement of the theorem can be written by using this notations as

$$(\forall n \geq 0) \quad \left| \begin{array}{l} \beta_n(\tau^{-1}x_n - \sum_{i=1}^{m} L_i^* v_{i,n}) - \tau^{-1}p_n + \sum_{i=1}^{m} L_i^* q_{i,n} - C(\beta_n x_n) \in Ap_n + \sum_{i=1}^{m} L_i^* q_{i,n} \\ x_{n+1} = \beta_n x_n + \lambda_n(p_n - \beta_n x_n) \\ \text{For } i = 1,..., m \\ \quad \left| \begin{array}{l} \beta_n(\sigma_i^{-1}v_{i,n} - L_i x_n) - \sigma_i^{-1}q_{i,n} + L_i p_n - D_i^{-1}(\beta_n v_{i,n}) \in B_i^{-1}(q_{i,n}) - L_i p_n \\ v_{i,n+1} = \beta_n v_{i,n} + \lambda_n(q_{i,n} - \beta_n v_{i,n}), \end{array} \right. \end{array} \right. \tag{22}$$

By introducing the sequences

$$\boldsymbol{x}_n = (x_n, v_{1,n},..., v_{m,n}), \text{ and } \boldsymbol{y}_n = (p_n, q_{1,n},..., q_{m,n}) \ \forall n \geq 0,$$

the scheme (22) can equivalently be written in the form

$$(\forall n \geq 0) \quad \left| \begin{array}{l} \beta_n \boldsymbol{V}(\boldsymbol{x}_n) - \boldsymbol{V}(\boldsymbol{y}_n) - \boldsymbol{Q}(\beta_n \boldsymbol{x}_n) \in (\boldsymbol{M} + \boldsymbol{S})\,(\boldsymbol{y}_n) \\ \boldsymbol{x}_{n+1} = \beta_n \boldsymbol{x}_n + \lambda_n\,(\boldsymbol{y}_n - \beta_n \boldsymbol{x}_n). \end{array} \right. \tag{23}$$

Furthermore, we have for any $n \geq 0$

$$\begin{aligned} &\beta_n \boldsymbol{V}(\boldsymbol{x}_n) - \boldsymbol{V}(\boldsymbol{y}_n) - \boldsymbol{Q}(\beta_n \boldsymbol{x}_n) \in (\boldsymbol{M} + \boldsymbol{S})\,(\boldsymbol{y}_n) \\ \Leftrightarrow &\left(\beta_n \boldsymbol{V} - \boldsymbol{Q} \circ (\beta_n\,\text{Id})\right)(\boldsymbol{x}_n) \in (\boldsymbol{M} + \boldsymbol{S} + \boldsymbol{V})(\boldsymbol{y}_n) \\ \Leftrightarrow &\boldsymbol{y}_n = (\boldsymbol{M} + \boldsymbol{S} + \boldsymbol{V})^{-1}\left(\beta_n \boldsymbol{V} - \boldsymbol{Q} \circ (\beta_n\,\text{Id})\right)(\boldsymbol{x}_n) \\ \Leftrightarrow &\boldsymbol{y}_n = \left(\text{Id} + \boldsymbol{V}^{-1}(\boldsymbol{M} + \boldsymbol{S})\right)^{-1}\left(\beta_n\,\text{Id} - \boldsymbol{V}^{-1} \circ \boldsymbol{Q} \circ (\beta_n\,\text{Id})\right)(\boldsymbol{x}_n) \\ \Leftrightarrow &\boldsymbol{y}_n = J_{\boldsymbol{A}}\left(\beta_n \boldsymbol{x}_n - \boldsymbol{B}(\beta_n \boldsymbol{x}_n)\right), \end{aligned}$$

where

$$\boldsymbol{A} := \boldsymbol{V}^{-1}(\boldsymbol{M} + \boldsymbol{S}) \text{ and } \boldsymbol{B} := \boldsymbol{V}^{-1}\boldsymbol{Q}. \tag{24}$$

Let $\mathcal{K}_{\boldsymbol{V}}$ be the Hilbert space with inner product and norm defined, for $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{K}$, by

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle_{\mathcal{K}_{\boldsymbol{V}}} = \langle \boldsymbol{x}, \boldsymbol{V}\boldsymbol{y} \rangle_{\mathcal{K}} \text{ and } \|\boldsymbol{x}\|_{\mathcal{K}_{\boldsymbol{V}}} = \sqrt{\langle \boldsymbol{x}, \boldsymbol{V}\boldsymbol{x} \rangle_{\mathcal{K}}}, \tag{25}$$

respectively. As the set-valued operators $\boldsymbol{M} + \boldsymbol{S}$ and $\boldsymbol{Q}$ are maximally monotone on $\mathcal{K}$, the operators $\boldsymbol{A}$ and $\boldsymbol{B}$ are maximally monotone on $\mathcal{K}_{\boldsymbol{V}}$ (see also [28]). Furthermore, $\boldsymbol{B}$ is $\beta\rho$-cocoercive on $\mathcal{K}_{\boldsymbol{V}}$. Moreover, since $\boldsymbol{V}$ is self-adjoint and ρ-strongly positive, weak and strong convergence in $\mathcal{K}_{\boldsymbol{V}}$ are equivalent with weak and strong convergence in $\mathcal{K}$, respectively.

Taking this into account, it follows that (23) becomes

$$(\forall n \geq 0) \ \boldsymbol{x}_{n+1} = \beta_n \boldsymbol{x}_n + \lambda_n \Big(J_{\boldsymbol{A}} \big(\beta_n \boldsymbol{x}_n - \boldsymbol{B}(\beta_n \boldsymbol{x}_n) \big) - \beta_n \boldsymbol{x}_n \Big),$$

which is the algorithm presented in Theorem 7 for determining the zeros of $\boldsymbol{A} + \boldsymbol{B}$ in case $\gamma = 1$. However, we have

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

According to Theorem 7, $\boldsymbol{x}_n$ converges strongly to $\text{proj}_{\text{zer}(\boldsymbol{A}+\boldsymbol{B})}(0, 0, ..., 0)$ in the space $\mathcal{K}_{\boldsymbol{V}}$ as $n \to +\infty$ and the conclusion follows from (21). $\blacksquare$

In the remaining of this subsection we investigate the convergence property of the algorithm (18) in the context of simultaneously solving complexly structured convex optimization problems and their Fenchel duals. The problem under investigation is the following one.

Problem 16 Let $f \in \Gamma(\mathcal{H})$ and $h : \mathcal{H} \to \mathbb{R}$ be a convex and differentiable function with a μ^{-1}-Lipschitz continuous gradient, for some $\mu > 0$. Let m be a strictly positive integer and for $i = 1, ..., m$, let $\mathcal{G}_i$ be a real Hilbert space, $g_i, l_i \in \Gamma(\mathcal{G}_i)$ such that l_i is ν_i-strongly convex, for some $\nu_i > 0$ and $L_i : \mathcal{H} \to \mathcal{G}_i$ a nonzero linear continuous operator. Consider the convex optimization problem

$$\inf_{x \in \mathcal{H}} \left\{ f(x) + \sum_{i=1}^{m} (g_i \square l_i)(L_i x) + h(x) \right\} \tag{26}$$

and its Fenchel-type dual problem

$$\sup_{v_i \in \mathcal{G}_i, i=1,..,m} \left\{ -(f^* \square h^*) \left(-\sum_{i=1}^{m} L_i^* v_i \right) - \sum_{i=1}^{m} \big(g_i^*(v_i) + l_i^*(v_i) \big) \right\}. \tag{27}$$

We denote by $\Gamma(\mathcal{H})$ the set of proper, convex and lower semicontinuous functions defined on $\mathcal{H}$ with values in the extended real line $\mathbb{R} \cup \{+\infty\}$. The conjugate of a function f is $f^* : \mathcal{H} \to \overline{\mathbb{R}}$, $f^*(p) = \sup \{ \langle p, x \rangle - f(x) : x \in \mathcal{H} \}$ for all $p \in \mathcal{H}$. Moreover, if $f \in \Gamma(\mathcal{H})$, then $f^* \in \Gamma(\mathcal{H})$, as well, and $(\partial f)^{-1} = \partial f^*$. Finally, having two proper functions $f, g : \mathcal{H} \to \overline{\mathbb{R}}$, their infimal convolution is defined by $f \square g : \mathcal{H} \to \overline{\mathbb{R}}$, $(f \square g)(x) = \inf_{y \in \mathcal{H}} \{ f(y) + g(x - y) \}$ for all $x \in \mathcal{H}$.

Corollary 17 *In Problem 16, suppose that*

$$0 \in \text{ran} \left(\partial f + \sum_{i=1}^{m} L_i^* \circ (\partial g_i \square \partial l_i) \circ L_i + \nabla h \right). \tag{28}$$

Let τ and σ_i, $i = 1, ..., m$, be strictly positive numbers such that

$$2 \cdot \min\{\tau^{-1}, \sigma_1^{-1}, ..., \sigma_m^{-1}\} \cdot \min\{\mu, \nu_1, ..., \nu_m\} \left(1 - \sqrt{\tau \sum_{i=1}^{m} \sigma_i \|L_i\|^2}\right) \geq 1.$$

Let be the starting point $(x_0, v_{1,0}, ..., v_{m,0}) \in \mathcal{H} \times \mathcal{G}_1 \times ... \times \mathcal{G}_m$ and set:

$$(\forall n \geq 0) \quad \left|\begin{array}{l} p_n = \text{prox}_{\tau f}\left[\beta_n x_n - \tau\left(\beta_n \sum_{i=1}^{m} L_i^* v_{i,n} + \nabla h(\beta_n x_n)\right)\right] \\ x_{n+1} = \beta_n x_n + \lambda_n(p_n - \beta_n x_n) \\ \text{For } i = 1, ..., m \\ \quad q_{i,n} = \text{prox}_{\sigma_i g_i^*}\left[\beta_n v_{i,n} + \sigma_i\left(L_i(2p_n - \beta_n x_n) - \nabla l_i^*(\beta_n v_{i,n})\right)\right] \\ \quad v_{i,n+1} = \beta_n v_{i,n} + \lambda_n(q_{i,n} - \beta_n v_{i,n}), \end{array}\right. \tag{29}$$

where $(\lambda_n)_{n \geq 0}$ and $(\beta_n)_{n \geq 0}$ are real sequences satisfying the conditions:

(i) $0 < \beta_n \leq 1$ for any $n \geq 0$, $\lim_{n \to +\infty} \beta_n = 1$, $\sum_{n \geq 0}(1 - \beta_n) = +\infty$ and $\sum_{n \geq 1} |\beta_n - \beta_{n-1}| < +\infty$;

(ii) $0 < \lambda_n \leq \frac{4\beta\rho - 1}{2\beta\rho}$ for any $n \geq 0$, $\liminf_{n \to +\infty} \lambda_n > 0$ and $\sum_{n \geq 1} |\lambda_n - \lambda_{n-1}| < +\infty$,

for

$$\beta := \min\{\mu, \nu_1, ..., \nu_m\}$$

and

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

Then there exists $(\bar{x}, \bar{v}_1, ..., \bar{v}_m) \in \mathcal{H} \times \mathcal{G}_1 \times ... \times \mathcal{G}_m$ such that $(x_n, v_{1,n}, ..., v_{m,n})$ converges strongly to $(\bar{x}, \bar{v}_1, ..., \bar{v}_m)$ as $n \to +\infty$ and $\bar{x}$ is an optimal solution of the problem (26), $(\bar{v}_1, ..., \bar{v}_m)$ is an optimal solution of (27) and the optimal objective values of the two optimization problems coincide.

Remark 18 The proximal-point operator of the conjugate function can be computed via the Moreau's decomposition formula

$$\text{prox}_{\gamma f} + \gamma \, \text{prox}_{(1/\gamma) f^*} \circ \gamma^{-1} \, \text{Id} = \text{Id}, \tag{30}$$

which is valid for $\gamma > 0$ and $f \in \Gamma(\mathcal{H})$ (see [6]).

Proof. Consider the maximal monotone operators

$$A = \partial f, C = \nabla h, B_i = \partial g_i \text{ and } D_i = \partial l_i, \ i = 1, ..., m.$$

The Baillon-Haddad Theorem (see [6, Corollary 18.16]) ensures that C is μ-cocoercive. Since l_i is ν_i-strongly convex, D_i is ν_i-strongly monotone, for $i = 1, ..., m$. According to [6, Proposition 17.10, Theorem 18.15], $D_i^{-1} = \nabla l_i^*$ is a monotone and ν_i^{-1}-Lipschitz continuous operator for $i = 1, ..., m$. The strong convexity of the functions l_i guarantees that $g_i \square l_i \in \Gamma(\mathcal{G}_i)$ (see [6, Corollary 11.16, Proposition 12.14]) and $\partial(g_i \square l_i) = \partial g_i \square \partial l_i, i = 1, ..., m$, (see [6, Proposition 24.27]).

Hence, the monotone inclusion problem (14) reads

$$\text{find } \bar{x} \in \mathcal{H} \text{ such that } 0 \in \partial f(\bar{x}) + \sum_{i=1}^{m} L_i^*(\partial g_i \square \partial l_i)(L_i \bar{x}) + \nabla h(\bar{x}), \tag{31}$$

while the dual monotone inclusion problem (15) reads

$$\text{find } \overline{v}_1 \in \mathcal{G}_1,...,\overline{v}_m \in \mathcal{G}_m \text{ such that } \exists x \in \mathcal{H}: \begin{cases} -\sum_{i=1}^m L_i^* \overline{v}_i \in \partial f(x) + \nabla h(x) \\ \overline{v}_i \in (\partial g_i \square \partial l_i)(L_i x), \ i = 1,...,m. \end{cases} \quad (32)$$

One can see that if $(\overline{x}, \overline{v}_1,...,\overline{v}_m) \in \mathcal{H} \times \mathcal{G}_1 \times ... \times \mathcal{G}_m$ is a primal-dual solution to (31)-(32), namely,

$$-\sum_{i=1}^m L_i^* \overline{v}_i \in \partial f(\overline{x}) + \nabla h(\overline{x}) \text{ and } \overline{v}_i \in (\partial g_i \square \partial l_i)(L_i \overline{x}), \ i = 1,...,m, \quad (33)$$

then $\overline{x}$ is an optimal solution of the problem (26), $(\overline{v}_1,...,\overline{v}_m)$ is an optimal solution of (27) and the optimal objective values of the two problems coincide. Notice that (33) is nothing else than the system of optimality conditions for the primal-dual pair of convex optimization problems (26)-(27).

The conclusion follows now from Theorem 14. ∎

Remark 19 (i) The relation (28) in the above theorem is fulfilled if the primal problem (26) has an optimal solution $\overline{x} \in \mathcal{H}$ and a suitable regularity condition holds. Under these auspices there exists an optimal solution to (27) $(\overline{v}_1,...,\overline{v}_m) \in \mathcal{G}_1 \times ... \times \mathcal{G}_m$, such that $(\overline{x}, \overline{v}_1,...,\overline{v}_m)$ satisfies the optimality conditions (33) and, consequently, (28) holds.

(ii) Further, let us discuss some conditions ensuring the existence of a primal optimal solution. Suppose that the primal problem (26) is feasible, which means that its optimal objective value is not identical $+\infty$. The existence of optimal solutions for (26) is guaranteed if, for instance, $f + h$ is coercive (that is $\lim_{\|x\| \to \infty}(f + h)(x) = +\infty$) and for all $i = 1,...,m$, g_i is bounded from below. Indeed, under these circumstances, the objective function of (26) is coercive (use also [6, Corollary 11.16 and Proposition 12.14] to show that for all $i = 1,...,m$, $g_i \square l_i$ is bounded from below and $g_i \square l_i \in \Gamma(\mathcal{G}_i)$) and the statement follows via [6, Corollary 11.15]. On the other hand, if $f + h$ is strongly convex, then the objective function of (26) is strongly convex, too, thus (26) has a unique optimal solution (see [6, Corollary 11.16]).

(iii) We discuss at this point a suitable regularity condition as mentioned at item (i) above. Since $\text{dom}(g_i \square l_i) = \text{dom } g_i + \text{dom } l_i, \ i = 1,...,m$, one can use to this end the regularity condition of interiority-type (see also [15])

$$(0,...,0) \in \text{sqri} \left(\prod_{i=1}^m (\text{dom } g_i + \text{dom } l_i) - \{(L_1 x,..., L_m x) : x \in \text{dom } f\} \right). \quad (34)$$

This is fulfilled provided that one of the following conditions is verified (see [15, Proposition 4.3]):

(a) $\text{dom } g_i + \text{dom } l_i = \mathcal{G}_i, \ i = 1,...,m$;

(b) $\mathcal{H}$ and $\mathcal{G}_i$ are finite-dimensional and there exists $x \in \text{ri dom } f$ such that $L_i x - r_i \in \text{ri dom } g_i + \text{ri dom } l_i, \ i = 1,...,m$.

5.2 A primal-dual algorithm of Douglas-Rachford-type with Tikhonov regularization terms

The problem that we investigate in this section reads as follows.

Problem 20 Let $A : \mathcal{H} \rightrightarrows \mathcal{H}$ be a maximally monotone operator. Let m be a strictly positive integer and for any $i = 1, ..., m$, let $\mathcal{G}_i$ be a real Hilbert space, $B_i, D_i : \mathcal{G}_i \rightrightarrows \mathcal{G}_i$ be maximally monotone operators and $L_i : \mathcal{H} \to \mathcal{G}_i$ a nonzero linear continuous operator. The problem is to solve the primal inclusion

$$\text{find } \overline{x} \in \mathcal{H} \text{ such that } 0 \in A\overline{x} + \sum_{i=1}^{m} L_i^* (B_i \square D_i)(L_i \overline{x}) \tag{35}$$

together with the dual inclusion

$$\text{find } \overline{v}_1 \in \mathcal{G}_1, ..., \overline{v}_m \in \mathcal{G}_m \text{ such that } (\exists x \in \mathcal{H}) \begin{cases} -\sum_{i=1}^{m} L_i^* \overline{v}_i \in Ax \\ \overline{v}_i \in (B_i \square D_i)(L_i x), \ i = 1, ..., m. \end{cases} \tag{36}$$

Different to Problem 13, the operators $D_i, i = 1, ..., m$ are general maximally monotone operators, thus they will have to be addressed through their resolvents. This is why in this context a primal-dual algorithm relying on the Douglas-Rachford paradigm is more appropriate.

Theorem 21 *In Problem 20, suppose that*

$$0 \in \mathrm{ran}\left(A + \sum_{i=1}^{m} L_i^* \circ (B_i \square D_i) \circ L_i \right). \tag{37}$$

Let $\tau, \sigma_i > 0$, $i = 1, ..., m$, be strictly positive numbers such that

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

Let be the starting point $(x_0, v_{1,0}, ..., v_{m,0}) \in \mathcal{H} \times \mathcal{G}_1 ... \times \mathcal{G}_m$ and set:

$$(\forall n \geq 0) \quad \left|
\begin{aligned}
& p_{1,n} = J_{\tau A}\left(\beta_n x_n - \tfrac{\tau}{2}\beta_n \sum_{i=1}^{m} L_i^* v_{i,n} \right) \\
& w_{1,n} = 2p_{1,n} - \beta_n x_n \\
& \text{For } i = 1, ..., m \\
& \quad \left|
\begin{aligned}
& p_{2,i,n} = J_{\sigma_i B_i^{-1}}\left(\beta_n v_{i,n} + \tfrac{\sigma_i}{2} L_i w_{1,n} \right) \\
& w_{2,i,n} = 2p_{2,i,n} - \beta_n v_{i,n}
\end{aligned}\right. \\
& z_{1,n} = w_{1,n} - \tfrac{\tau}{2} \sum_{i=1}^{m} L_i^* w_{2,i,n} \\
& x_{n+1} = \beta_n x_n + \lambda_n(z_{1,n} - p_{1,n}) \\
& \text{For } i = 1, ..., m \\
& \quad \left|
\begin{aligned}
& z_{2,i,n} = J_{\sigma_i D_i^{-1}}\left(w_{2,i,n} + \tfrac{\sigma_i}{2} L_i(2z_{1,n} - w_{1,n}) \right) \\
& v_{i,n+1} = \beta_n v_{i,n} + \lambda_n(z_{2,i,n} - p_{2,i,n}),
\end{aligned}\right.
\end{aligned}\right. \tag{38}$$

where $(\lambda_n)_{n \geq 0}$ and $(\beta_n)_{n \geq 0}$ real sequences satisfying the conditions:

(i) $0 < \beta_n \leq 1$ for any $n \geq 0$, $\lim_{n \to +\infty} \beta_n = 1$, $\sum_{n \geq 0}(1 - \beta_n) = +\infty$ and $\sum_{n \geq 1} |\beta_n - \beta_{n-1}| < +\infty$;

(ii) $0 < \lambda_n \leq 2$ for any $n \geq 0$, $\liminf_{n \to +\infty} \lambda_n > 0$ and $\sum_{n \geq 1} |\lambda_n - \lambda_{n-1}| < +\infty$.

Then there exists an element $(\overline{x}, \overline{v}_1, ..., \overline{v}_m) \in \mathcal{H} \times \mathcal{G}_1 ... \times \mathcal{G}_m$ such that the following statements are true:

(a) by setting

$$\overline{p}_1 = J_{\tau A}\left(\overline{x} - \frac{\tau}{2} \sum_{i=1}^{m} L_i^* \overline{v}_i \right),$$

$$\overline{p}_{2,i} = J_{\sigma_i B_i^{-1}}\left(\overline{v}_i + \frac{\sigma_i}{2} L_i(2\overline{p}_1 - \overline{x}) \right), \ i = 1, ..., m,$$

the element $(\overline{p}_1, \overline{p}_{2,1}, ..., \overline{p}_{2,m}) \in \mathcal{H} \times \mathcal{G}_1 \times ... \times \mathcal{G}_m$ is a primal-dual solution to Problem 20;

(b) $(x_n, v_{1,n},..., v_{m,n})$ converges strongly to $(\overline{x}, \overline{v}_1,..., \overline{v}_m)$ as $n \to +\infty$;

(c) $(p_{1,n}, p_{2,1,n},..., p_{2,m,n})$ and $(z_{1,n}, z_{2,1,n},..., z_{2,m,n})$ converge strongly to $(\overline{p}_1, \overline{p}_{2,1},..., \overline{p}_{2,m})$ as $n \to +\infty$.

Proof. For the proof we use Theorem 10 (see also [9]) in the same setting as in the proof of Theorem 14, namely, by considering $\mathcal{K} = \mathcal{H} \times \mathcal{G}_1 \times ... \times \mathcal{G}_m$ endowed with inner product and associated norm defined in (20). Furthermore, we consider again the maximally monotone operator

$$\boldsymbol{M} : \mathcal{K} \rightrightarrows \mathcal{K}, \quad (x, v_1,..., v_m) \mapsto (Ax, B_1^{-1}v_1,..., B_m^{-1}v_m),$$

the linear continuous skew-symmetric operator

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

and the (this time not necessarily single-valued) maximally monotone operator

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

Since $\text{dom}\,\boldsymbol{S} = \mathcal{K}$, both $\frac{1}{2}\boldsymbol{S} + \boldsymbol{Q}$ and $\frac{1}{2}\boldsymbol{S} + \boldsymbol{M}$ are maximally monotone (see [6, Corollary 24.4(i)]). Furthermore,

$$(37) \Leftrightarrow \text{zer}\,(\boldsymbol{M} + \boldsymbol{S} + \boldsymbol{Q}) \neq \varnothing$$

and

$$
\begin{aligned}
&(x, v_1,..., v_m) \in \text{zer}\,(\boldsymbol{M} + \boldsymbol{S} + \boldsymbol{Q}) \\
&\Leftrightarrow (x, v_1,..., v_m) \text{ is a primal-dual solution to Problem 20.}
\end{aligned}
\tag{39}
$$

We introduce the linear continuous operator

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

which is self-adjoint and ρ-strongly positive (see also [9]), for

$$\rho := \left(1 - \frac{1}{2}\sqrt{\tau \sum_{i=1}^m \sigma_i \|L_i\|^2} \right) \min\left\{ \frac{1}{\tau}, \frac{1}{\sigma_1}, \ldots, \frac{1}{\sigma_m} \right\} > 0,$$

namely, the following inequality holds

$$\langle \boldsymbol{x}, \boldsymbol{V}\boldsymbol{x} \rangle_{\mathcal{K}} \geq \rho \|\boldsymbol{x}\|_{\mathcal{K}}^2 \ \forall \boldsymbol{x} \in \mathcal{K}.$$

Therefore, its inverse operator $\boldsymbol{V}^{-1}$ exists and it fulfills $\|\boldsymbol{V}^{-1}\| \leq \frac{1}{\rho}$.

The algorithmic scheme (38) is equivalent to

$$
(\forall n \geq 0) \quad
\left|
\begin{aligned}
&\frac{\beta_n x_n - p_{1,n}}{\tau} - \frac{1}{2}\sum_{i=1}^m L_i^*(\beta_n v_{i,n} - p_{2,i,n}) \in \frac{1}{2}\sum_{i=1}^m L_i^* p_{2,i,n} + A p_{1,n} \\
&w_{1,n} = 2p_{1,n} - \beta_n x_n \\
&\text{For } i = 1,..., m \\
&\quad\left|
\begin{aligned}
&\frac{\beta_n v_{i,n} - p_{2,i,n}}{\sigma_i} - \frac{1}{2}L_i(\beta_n x_n - p_{1,n}) \in -\frac{1}{2}L_i p_{1,n} + B_i^{-1} p_{2,i,n} \\
&w_{2,i,n} = 2p_{2,i,n} - \beta_n v_{i,n}
\end{aligned}
\right. \\
&\frac{w_{1,n} - z_{1,n}}{\tau} - \frac{1}{2}\sum_{i=1}^m L_i^* w_{2,i,n} = 0 \\
&x_{n+1} = \beta_n x_n + \lambda_n(z_{1,n} - p_{1,n}) \\
&\text{For } i = 1,..., m \\
&\quad\left|
\begin{aligned}
&\frac{w_{2,i,n} - z_{2,i,n}}{\sigma_i} - \frac{1}{2}L_i(w_{1,n} - z_{1,n}) \in -\frac{1}{2}L_i z_{1,n} + D_i^{-1} z_{2,i,n} \\
&v_{i,n+1} = \beta_n v_{i,n} + \lambda_n(z_{2,i,n} - p_{2,i,n}).
\end{aligned}
\right.
\end{aligned}
\right.
\tag{40}
$$

By considering for any $n \geq 0$ the notations

$$\boldsymbol{x}_n = (x_n, v_{1,n}, ..., v_{m,n}), \boldsymbol{y}_n = (p_{1,n}, p_{2,1,n}, ..., p_{2,m,n}) \text{ and } \boldsymbol{z}_n = (z_{1,n}, z_{2,1,n}, ..., z_{2,m,n}),$$

the iterative scheme (40) can be written as

$$(\forall n \geq 0) \left|
\begin{array}{l}
\boldsymbol{V}(\beta_n \boldsymbol{x}_n - \boldsymbol{y}_n) \in \left(\frac{1}{2}\boldsymbol{S} + \boldsymbol{M}\right) \boldsymbol{y}_n \\
\boldsymbol{V}(2\boldsymbol{y}_n - \beta_n \boldsymbol{x}_n - \boldsymbol{z}_n) \in \left(\frac{1}{2}\boldsymbol{S} + \boldsymbol{Q}\right) \boldsymbol{z}_n \\
\boldsymbol{x}_{n+1} = \beta_n \boldsymbol{x}_n + \lambda_n (\boldsymbol{z}_n - \boldsymbol{y}_n)
\end{array}
\right. \tag{41}$$

which is further equivalent to

$$(\forall n \geq 0) \left|
\begin{array}{l}
\boldsymbol{y}_n = \left(\mathrm{Id} + \boldsymbol{V}^{-1}(\frac{1}{2}\boldsymbol{S} + \boldsymbol{M})\right)^{-1} (\beta_n \boldsymbol{x}_n) \\
\boldsymbol{z}_n = \left(\mathrm{Id} + \boldsymbol{V}^{-1}(\frac{1}{2}\boldsymbol{S} + \boldsymbol{Q})\right)^{-1} (2\boldsymbol{y}_n - \beta_n \boldsymbol{x}_n) \\
\boldsymbol{x}_{n+1} = \beta_n \boldsymbol{x}_n + \lambda_n (\boldsymbol{z}_n - \boldsymbol{y}_n).
\end{array}
\right. \tag{42}$$

Let $\mathcal{K}_{\boldsymbol{V}}$ be the Hilbert space with inner product and norm defined, for $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{K}$, via

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle_{\mathcal{K}_{\boldsymbol{V}}} = \langle \boldsymbol{x}, \boldsymbol{V}\boldsymbol{y} \rangle_{\mathcal{K}} \text{ and } \|\boldsymbol{x}\|_{\mathcal{K}_{\boldsymbol{V}}} = \sqrt{\langle \boldsymbol{x}, \boldsymbol{V}\boldsymbol{x} \rangle_{\mathcal{K}}}, \tag{43}$$

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

$$\boldsymbol{B} := \boldsymbol{V}^{-1}\left(\frac{1}{2}\boldsymbol{S} + \boldsymbol{M}\right) \text{ and } \boldsymbol{A} := \boldsymbol{V}^{-1}\left(\frac{1}{2}\boldsymbol{S} + \boldsymbol{Q}\right) \tag{44}$$

are maximally monotone on $\mathcal{K}_{\boldsymbol{V}}$. Furthermore, since $\boldsymbol{V}$ is self-adjoint and ρ-strongly positive, strong convergence in $\mathcal{K}_{\boldsymbol{V}}$ is equivalent with strong convergence in $\mathcal{K}$.

Taking this into account, (42) becomes

$$(\forall n \geq 0) \left|
\begin{array}{l}
\boldsymbol{y}_n = J_{\boldsymbol{B}} (\beta_n \boldsymbol{x}_n) \\
\boldsymbol{z}_n = J_{\boldsymbol{A}} (2\boldsymbol{y}_n - \beta_n \boldsymbol{x}_n) \\
\boldsymbol{x}_{n+1} = \beta_n \boldsymbol{x}_n + \lambda_n (\boldsymbol{z}_n - \boldsymbol{y}_n),
\end{array}
\right. \tag{45}$$

which is the Douglas–Rachford algorithm formulated in Theorem 10 in case $\gamma = 1$ for determining the zeros of $\boldsymbol{A} + \boldsymbol{B}$. It is easy to see that

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

By Theorem 10 (a), there exists $\overline{\boldsymbol{x}} = (\overline{x}, \overline{v}_1, ..., \overline{v}_m) \in \mathrm{Fix}(R_{\boldsymbol{A}} R_{\boldsymbol{B}})$, such that $J_{\boldsymbol{B}} \overline{\boldsymbol{x}} \in \mathrm{zer}(\boldsymbol{A} + \boldsymbol{B}) = \mathrm{zer}(\boldsymbol{M} + \boldsymbol{S} + \boldsymbol{Q})$. The claim follows from Theorem 10, (39) and by writing $J_{\boldsymbol{B}} \overline{\boldsymbol{x}}$ in terms of the resolvents of the operators involved in the expression of $\boldsymbol{B}$. $\blacksquare$

We close this section by considering the variational case.

Problem 22 Let $f \in \Gamma(\mathcal{H})$, m be a strictly positive integer and for $i = 1, ..., m$, let $\mathcal{G}_i$ be a real Hilbert space, $g_i, l_i \in \Gamma(\mathcal{G}_i)$ and $L_i : \mathcal{H} \to \mathcal{G}_i$ a nonzero linear continuous operator. Consider the convex optimization problem

$$\inf_{x \in \mathcal{H}} \left\{ f(x) + \sum_{i=1}^{m} (g_i \square l_i)(L_i x) \right\} \tag{46}$$

and its conjugate dual problem

$$\sup_{(v_1, ..., v_m) \in \mathcal{G}_1 \times ... \times \mathcal{G}_m} \left\{ -f^*\left(-\sum_{i=1}^{m} L_i^* v_i\right) - \sum_{i=1}^{m} (g_i^*(v_i) + l_i^*(v_i)) \right\}. \tag{47}$$

Corollary 23 *In Problem 22, suppose that*

$$0 \in \mathrm{ran}\left(\partial f + \sum_{i=1}^{m} L_i^* \circ (\partial g_i \square \partial l_i) \circ L_i\right). \tag{48}$$

Let $\tau, \sigma_i > 0$, $i = 1,...,m$, be strictly positive numbers such that

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

Let be the starting point $(x_0, v_{1,0}, ..., v_{m,0}) \in \mathcal{H} \times \mathcal{G}_1 ... \times \mathcal{G}_m$ and set:

$$(\forall n \geq 0) \quad \left| \begin{array}{l} p_{1,n} = \mathrm{prox}_{\tau f}\left(\beta_n x_n - \frac{\tau}{2}\beta_n \sum_{i=1}^{m} L_i^* v_{i,n}\right) \\ w_{1,n} = 2p_{1,n} - \beta_n x_n \\ \text{For } i = 1,...,m \\ \quad \left| \begin{array}{l} p_{2,i,n} = \mathrm{prox}_{\sigma_i g_i^*}\left(\beta_n v_{i,n} + \frac{\sigma_i}{2}L_i w_{1,n}\right) \\ w_{2,i,n} = 2p_{2,i,n} - \beta_n v_{i,n} \end{array}\right. \\ z_{1,n} = w_{1,n} - \frac{\tau}{2}\sum_{i=1}^{m} L_i^* w_{2,i,n} \\ x_{n+1} = \beta_n x_n + \lambda_n(z_{1,n} - p_{1,n}) \\ \text{For } i = 1,...,m \\ \quad \left| \begin{array}{l} z_{2,i,n} = \mathrm{prox}_{\sigma_i l_i^*}\left(w_{2,i,n} + \frac{\sigma_i}{2}L_i(2z_{1,n} - w_{1,n})\right) \\ v_{i,n+1} = \beta_n v_{i,n} + \lambda_n(z_{2,i,n} - p_{2,i,n}), \end{array}\right. \end{array}\right. \tag{49}$$

where $(\lambda_n)_{n\geq 0}$ and $(\beta_n)_{n\geq 0}$ are real sequences satisfying the conditions:
    (i) $0 < \beta_n \leq 1$ for any $n \geq 0$, $\lim_{n\to+\infty} \beta_n = 1$, $\sum_{n\geq 0}(1-\beta_n) = +\infty$ and $\sum_{n\geq 1}|\beta_n - \beta_{n-1}| < +\infty$;
    (ii) $0 < \lambda_n \leq 2$ for any $n \geq 0$, $\liminf_{n\to+\infty} \lambda_n > 0$ and $\sum_{n\geq 1}|\lambda_n - \lambda_{n-1}| < +\infty$.
    Then there exists an element $(\bar{x}, \bar{v}_1, ..., \bar{v}_m) \in \mathcal{H} \times \mathcal{G}_1 ... \times \mathcal{G}_m$ such that the following statements are true:

  (a) by setting

$$\bar{p}_1 = \mathrm{prox}_{\tau f}\left(\bar{x} - \frac{\tau}{2}\sum_{i=1}^{m} L_i^* \bar{v}_i\right),$$

$$\bar{p}_{2,i} = \mathrm{prox}_{\sigma_i g_i^*}\left(\bar{v}_i + \frac{\sigma_i}{2}L_i(2\bar{p}_1 - \bar{x})\right), \quad i = 1,...,m,$$

  the element $(\bar{p}_1, \bar{p}_{2,1}, ..., \bar{p}_{2,m}) \in \mathcal{H} \times \mathcal{G}_1 \times ... \times \mathcal{G}_m$ is a primal-dual solution to Problem 20, namely,

$$-\sum_{i=1}^{m} L_i^* \bar{v}_i \in \partial f(\bar{x}) \text{ and } \bar{v}_i \in (\partial g_i \square \partial l_i)(L_i \bar{x}), \quad i = 1,...,m, \tag{50}$$

  hence $\bar{p}_1$ is an optimal solution to (46) and $(\bar{p}_{2,1}, ..., \bar{p}_{2,m})$ is an optimal solution to (47);

  (b) $(x_n, v_{1,n}, ..., v_{m,n})$ converges strongly to $(\bar{x}, \bar{v}_1, ..., \bar{v}_m)$ as $n \to +\infty$;

  (c) $(p_{1,n}, p_{2,1,n}, ..., p_{2,m,n})$ and $(z_{1,n}, z_{2,1,n}, ..., z_{2,m,n})$ converge strongly to $(\bar{p}_1, \bar{p}_{2,1}, ..., \bar{p}_{2,m})$ as $n \to +\infty$.

Remark 24 The hypothesis (48) in Theorem 23 is fulfilled, if the primal problem (46) has an optimal solution, the regularity condition (34) holds and

$$0 \in \mathrm{sqri}(\mathrm{dom}\, g_i^* - \mathrm{dom}\, l_i^*) \text{ for } i = 1, ..., m.$$

According to [6, Proposition 15.7], the latter also guarantees that $g_i \square l_i \in \Gamma(\mathcal{G}_i)$, $i = 1, ..., m$.

6 Numerical experiments: applications to the split feasibility problem

Let $\mathcal{H}$ and $\mathcal{G}$ be real Hilbert spaces and $L : \mathcal{H} \to \mathcal{G}$ a bounded linear operator. Let C and Q be nonempty, closed and convex subsets of $\mathcal{H}$ and $\mathcal{G}$, respectively. The *split feasibility problem* (SFP) searches a point $x \in \mathcal{H}$ with the property

$$x \in C \text{ and } Lx \in Q. \tag{51}$$

The (SFP) was originally introduced by Censor and Elfving [11] for solving inverse problems in the context of phase retrieval, medical image reconstruction and intensity modulated radiation therapy.

We show that the strongly convergent primal-dual algorithms which we have investigated in Section 5 are excellently suited to solve the (SFP), especially in the case when infinite dimensional Hilbert spaces are involved. For this purpose, we note that problem (51) can be written equivalently in the form

$$\min_{x \in \mathcal{H}} \left\{ \delta_C(x) + \delta_Q(Lx) \right\}, \tag{52}$$

where

$$\delta_S(x) := \left\{ \begin{array}{ll} 0, & \text{if } x \in S \\ +\infty, & \text{else} \end{array} \right.$$

denotes the indicator function of a subset S of a Hilbert space. Another alternative way to write problem (51) is as the minimization problem

$$\min_{x \in \mathcal{H}} \left\{ \frac{1}{2} d_C^2(x) + \delta_Q(Lx) \right\}, \tag{53}$$

where $d_C(x) = \inf_{u \in \mathcal{C}} \| x - u \|$ denotes the distance of the point $x \in \mathcal{H}$ to the set C. Both optimization problems are special instances of Problem 16, consequently, we will solve them by making use of the algorithm stated in Corollary 17.

The optimization problem (52) can be stated in the framework of Problem 16 by taking $f = \delta_C$, $m = 1$, $g_1 = \delta_Q$, $l_1 = \delta_{\{0\}}$, $L_1 = L$ and $h = 0$. The iterative scheme stated in Corollary 17 reads:

$$(\forall n \geq 0) \quad \left| \begin{array}{l} p_n = P_C(\beta_n x_n - \tau \beta_n L^* v_n) \\ x_{n+1} = \beta_n x_n + \lambda_n(p_n - \beta_n x_n) \\ q_n = \beta_n v_n + \sigma L(2p_n - \beta_n x_n) - \sigma P_Q(\sigma^{-1}\beta_n v_n + L(2p_n - \beta_n x_n)) \\ v_{n+1} = \beta_n v_n + \lambda_n(q_n - \beta_n v_n). \end{array} \right. \tag{54}$$

The optimization problem (52) can be stated in the framework of Problem 16 by taking $f = 0$, $m = 1$, $g_1 = \delta_Q$, $l_1 = \delta_{\{0\}}$, $L_1 = L$ and $h = \frac{1}{2}d_C^2$. Noticing that $\nabla\left(\frac{1}{2}d_C^2\right) = \mathrm{Id} - P_C$, the iterative scheme stated in Corollary 17 reads:

$$(\forall n \geq 0) \quad \left| \begin{array}{l} p_n = \beta_n x_n - \tau(\beta_n L^* v_n + \beta_n x_n - P_C(\beta_n x_n)) \\ x_{n+1} = \beta_n x_n + \lambda_n(p_n - \beta_n x_n) \\ q_n = \beta_n v_n + \sigma L(2p_n - \beta_n x_n) - \sigma P_Q(\sigma^{-1}\beta_n v_n + L(2p_n - \beta_n x_n)) \\ v_{n+1} = \beta_n v_n + \lambda_n(q_n - \beta_n v_n). \end{array} \right. \tag{55}$$

For the numerical example we consider the following setup: let $\mathcal{H} = \mathcal{G} = L^2([0, 2\pi]) := \left\{ f : [0, 2\pi] \to \mathbb{R} : \int_0^{2\pi} |f(t)|^2 dt < +\infty \right\}$ equipped with the scalar product $\langle f, g \rangle := \int_0^{2\pi} f(t)g(t)dt$ and the associated norm $\|f\| := \left(\int_0^{2\pi} |f(t)|^2 dt \right)^{1/2}$ for all $f, g \in L^2([0, 2\pi])$. The sets

$$C := \left\{ x \in L^2([0, 2\pi]) : \int_0^{2\pi} x(t)dt \leq 1 \right\}$$

and

$$Q := \left\{ x \in L^2([0, 2\pi]) : \int_0^{2\pi} |x(t) - \sin(t)|^2 dt \leq 16 \right\}$$

are nonempty, closed and convex subsets of $L^2([0, 2\pi])$. Notice that $C = \{x \in L^2([0, 2\pi]) : \langle x, u \rangle \leq 1\}$ and $Q = \{x \in L^2([0, 2\pi]) : \|x - f\| \leq 4\}$, where $u : [0, 2\pi] \to \mathbb{R}, u(t) = 1$ for all $t \in [0, 2\pi]$, and $f : [0, 2\pi] \to \mathbb{R}, f(t) = \sin t$ for all $t \in [0, 2\pi]$. We define the linear continuous operator $L : L^2([0, 2\pi]) \to L^2([0, 2\pi])$ as

$$(Lx)(t) := \left(\int_0^{2\pi} x(s)ds \right) \cdot t.$$

Let $x, y \in L^2([0, 2\pi])$. Then

$$\langle Lx, y \rangle = \int_0^{2\pi} Lx(t)y(t)dt = \int_0^{2\pi} x(s) \int_0^{2\pi} ty(t)dt\ ds,$$

and therefore

$$(L^*y)(t) = \int_0^{2\pi} sy(s)ds.$$

Furthermore, by using the Cauchy-Schwarz inequality yields

$$\|L\|^2 = \sup_{\|x\|=1} \int_0^{2\pi} Tx(t)^2 dt = \sup_{\|x\|=1} \int_0^{2\pi} t^2 dt \left(\int_0^{2\pi} x(s)ds \right)^2$$

$$\leq \sup_{\|x\|=1} \frac{8}{3}\pi^3 \int_0^{2\pi} x(s)^2 ds \cdot 2\pi = \frac{16}{3}\pi^4.$$

The projection onto the set C can be computed as (see [6, Example 28.16])

$$P_C(x) = \begin{cases} \frac{1 - \int_0^{2\pi} x(t)dt}{2\pi} + x, & \text{if } \int_0^{2\pi} x(t)dt > 1 \\ x, & \text{else.} \end{cases}$$

On the other hand P_Q is given by [6, Example 28.10]

$$P_Q(x) = \begin{cases} \sin + \frac{4(x - \sin)}{(\int_0^{2\pi} |x(t) - \sin(t)|^2 dt)^{1/2}}, & \text{if } \int_0^{2\pi} |x(t) - \sin(t)|^2 dt > 16 \\ x, & \text{else.} \end{cases}$$

We implemented the algorithms (54) and (55) in MATLAB used symbolic computation for generating the sequences of iterates. One can easily notice that, in this particular setting, the split feasibility problem and, consequently, both addressed optimization problems are solvable.

		Number of iterations	
x_0	v_0	$\beta_n = 1$	$\beta_n = 1 - \frac{1}{n+1}$
$\frac{t^2}{10}$	$\frac{t^2}{10}$	13	1
$\frac{t^2}{10}$	$\frac{1}{2}e^t$	20	11
$\frac{t^2}{10}$	$e^t + \frac{t^2}{24}$	21	12
$\frac{1}{2}e^t$	$\frac{t^2}{10}$	> 150	11
$\frac{1}{2}e^t$	$\frac{1}{2}e^t$	20	12
$\frac{1}{2}e^t$	$e^t + \frac{t^2}{24}$	21	13
$e^t + \frac{t^2}{24}$	$\frac{t^2}{10}$	> 150	15
$e^t + \frac{t^2}{24}$	$\frac{1}{2}e^t$	20	13
$e^t + \frac{t^2}{24}$	$e^t + \frac{t^2}{24}$	21	13

Table 1: Comparison of the variants without and with Tikhonov regularization terms of the primal-dual algorithm (54), for different starting values and step sizes $\tau = 0.1$ and $\sigma = 0.01$.

		Number of iterations	
x_0	v_0	$\beta_n = 1$	$\beta_n = 1 - \frac{1}{n+1}$
$\frac{t^2}{10}$	$\frac{t^2}{10}$	24	1
$\frac{t^2}{10}$	$\frac{1}{2}e^t$	46	10
$\frac{t^2}{10}$	$e^t + \frac{t^2}{24}$	46	10
$\frac{1}{2}e^t$	$\frac{t^2}{10}$	30	6
$\frac{1}{2}e^t$	$\frac{1}{2}e^t$	24	11
$\frac{1}{2}e^t$	$e^t + \frac{t^2}{24}$	35	21
$e^t + \frac{t^2}{24}$	$\frac{t^2}{10}$	32	6
$e^t + \frac{t^2}{24}$	$\frac{1}{2}e^t$	36	12
$e^t + \frac{t^2}{24}$	$e^t + \frac{t^2}{24}$	24	11

Table 2: Comparison of the variants without and with Tikhonov regularization terms of the primal-dual algorithm (55), for different starting values and step sizes $\tau = 0.1$ and $\sigma = 0.01$.

The numerical experiments confirmed that the primal-dual algorithms involving Tikhonov regularization terms outperform the ones without Tikhonov regularization terms, which correspond to the case when $\beta_n = 1$ for every $n \geq 0$ and for which is known that they weakly convergence to a primal-dual solution of the corresponding KKT system of optimality conditions.

In the two tables above we present the numbers of iterations needed by the two algorithms to approach a solution of the split feasibility problem (SFP). We consider as stopping criterion

$$E(x_n) := \frac{1}{2}||P_C(x_n) - x_n||^2 + \frac{1}{2}||P_Q(Lx_n) - Lx_n||^2 \leq 10^{-3},$$

while taking as relaxation variables $\lambda_n = 0.4$ for every $n \geq 0$ and as Tikhonov regularization parameters $\beta_n := 1 - \frac{1}{n+1}$ for every $n \geq 0$, respectively, $\beta_n = 1$ for every $n \geq 0$ for the variants without Tikhonov regularization terms.

7 Further work

We point out some directions of research related to proximal methods with Tikhonov regularization terms, which merit to be addressed starting from the investigations made in this paper:

1. To consider in the numerical algorithms of type forward-backward and Douglas-Rachford proposed in this paper *dynamic step sizes*, which are known to increase the flexibility of the algorithms. This can be for instance done by formulating first a Krasnosel'skiĭ–Mann–type algorithm for determining an element in the intersection of the sets of fixed points of a family of nonexpansive operators $(T_k)_{k \geq 0}$. Suitable choices of the operators in this family, in the spirit of the investigations we made in the sections (3) and (4), can lead to iterative algorithms with dynamic step sizes (see [14]).

2. To employ in the iterative schemes discussed in this article *inertial and memory effects*, which are known to contribute to the acceleration of the convergence behaviour of the algorithms.

3. To translate the proposed numerical methods to a framework that goes *beyond the Hilbert space setting* considered in this article, in order to allow applications where functional spaces are involved. We refer the reader to [24, 25, 27] and the references therein for tools and techniques which allow to prove convergence statements in Banach spaces with some particular underlying geometric structures.

References

[1] A. Alotaibi, P.L. Combettes, N. Shahzad, *Best approximation from the Kuhn-Tucker set of composite monotone inclusions*, Numerical Functional Analysis and Optimization 36(12), 1513–1532, 2015

[2] H. Attouch, *Viscosity solutions of minimization problems*, SIAM Journal on Optimization 6(3), 769-806, 1996

[3] H. Attouch, L.M. Briceño-Arias, P.L. Combettes, *A strongly convergent primal-dual method for nonoverlapping domain decomposition*, Numerische Mathematik 133(3), 443–470, 2016

[4] H. Attouch, M. Théra, *A general duality principle for the sum of two operators*, Journal of Convex Analysis 3(1), 1–24, 1996

[5] H.H. Bauschke, P.L. Combettes, *A weak-to-strong convergence principle for Fejér-monotone methods in Hilbert spaces*, Mathematics of Operations Research 26(2), 248–264, 2001

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

[7] J.M. Borwein and J.D. Vanderwerff, *Convex Functions: Constructions, Characterizations and Counterexamples*, Cambridge University Press, Cambridge, 2010

[8] R.I. Boţ, E.R. Csetnek, A. Heinrich, C. Hendrich, *On the convergence rate improvement of a primal-dual splitting algorithm for solving monotone inclusion problems*, Mathematical Programming 150(2), 251–279, 2015

[9] R.I. Boţ, C. Hendrich, *A Douglas-Rachford type primal-dual method for solving inclusions with mixtures of composite and parallel-sum type monotone operators*, SIAM Journal on Optimization 23(4), 2541–2565, 2013

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

[11] Y. Censor, T. Elfving, *A multiprojection algorithm using Bregman projections in a product space*, Numerical Algorithms 8, 221–239, 1994

[12] A. Chambolle, T. Pock, *A first-order primal-dual algorithm for convex problems with applications to imaging*, Journal of Mathematical Imaging and Vision 40(1), 120–145, 2011

[13] C. Chidume, *Geometric properties of Banach spaces and nonlinear iterations*, Lecture Notes in Mathematics 1965, Springer-Verlag London, 2009

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

[15] P.L. Combettes, J.-C. Pesquet, *Primal-dual splitting algorithm for solving inclusions with mixtures of composite, Lipschitzian, and parallel-sum type monotone operators*, Set-Valued and Variational Analysis 20(2), 307–330, 2012

[16] P.L. Combettes, I. Yamada, *Compositions and convex combinations of averaged nonexpansive operators*, Journal of Mathematical Analysis and Applications 425(1), 55–70, 2015

[17] L. Condat, *A primal-dual splitting method for convex optimization involving Lipschitzian, proximable and linear, composite terms*, Journal of Optimization Theory and Applications 158(2), 460–479, 2013

[18] Y. Haugazeau, *Sur les inéquations variationnelles et la minimisation de fonctionnelles convexes*, Thése, Université de Paris, France, 1968

[19] N. Lehdili, A. Moudafi, *Combining the proximal algorithm and Tikhonov regularization*, Optimization 37(3), 239–252, 1996

[20] A. Moudafi, *Viscosity approximation methods for fixed-points problems*, Journal of Mathematical Analysis and Applications 241(1), 46–55, 2000

[21] N. Ogura, I. Yamada, *Non-strictly convex minimization over the fixed point set of an asymptotically shrinking nonexpansive mapping*, Numerical Functional Analysis and Optimization 23(1-2), 113–137, 2002

[22] R.T. Rockafellar, *On the maximal monotonicity of subdifferential mappings*, Pacific Journal of Mathematics 33(1), 209–216, 1970

[23] R.T. Rockafellar, *Monotone operators and the proximal point algorithm*, SIAM Journal on Control and Optimization 14(5), 877–898, 1976

[24] D.R. Sahu, Q.H. Ansari, J.C. Yao, *The prox-Tikhonov-like forward-backward method and applications*, Taiwanese Journal of Mathematics 19(2), 481–503, 2015

[25] D.R. Sahu, Y.C. Yao, *The prox-Tikhonov regularization method for the proximal point algorithm in Banach spaces*, Journal of Global Optimization 51(4), 641–655, 2011

[26] S. Simons, *From Hahn-Banach to Monotonicity*, Springer, Berlin, 2008

[27] H.-K. Xu, *Iterative algorithms for nonlinear operators*, Journal of the London Mathematical Society (2), 66(1), 240–256, 2002

[28] B.C. Vũ, *A splitting algorithm for dual monotone inclusions involving cocoercive operators*, Advances in Computational Mathematics 38(3), 667–681, 2013

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