An inertial forward-backward-forward primal-dual splitting algorithm for solving monotone inclusion problems

Radu Ioan Boţ [*]   Ernö Robert Csetnek [†]

May 8, 2015

Abstract. We introduce and investigate the convergence properties of an inertial forward-backward-forward splitting algorithm for approaching the set of zeros of the sum of a maximally monotone operator and a single-valued monotone and Lipschitzian operator. By making use of the product space approach, we expand it to the solving of inclusion problems involving mixtures of linearly composed and parallel-sum type monotone operators. We obtain in this way an inertial forward-backward-forward primal-dual splitting algorithm having as main characteristic the fact that in the iterative scheme all operators are accessed separately either via forward or via backward evaluations. We present also the variational case when one is interested in the solving of a primal-dual pair of convex optimization problems with complexly structured objectives, which we also illustrate by numerical experiments in image processing.

Key Words. maximally monotone operator, resolvent, subdifferential, convex optimization, inertial splitting algorithm, primal-dual algorithm

AMS subject classification. 47H05, 65K05, 90C25

1   Introduction and preliminaries

Due to its wide applicability in different branches of the applied mathematics, especially in connection with real-life problems, the problem of solving inclusion problems involving mixtures of monotone operators in Hilbert spaces continues to attract the interest of many researchers (see [6, 10, 10, 14–16, 20, 21, 34]).

In this paper we will focus on the class of so-called *inertial proximal methods*, the origins of which go back to [1, 3]. The idea behind the iterative scheme relies on the use of an implicit discretization of a differential system of second-order in time and it was employed for the first time in the context of finding the zeros of a maximally monotone operator in [3]. One of the main features of the inertial proximal algorithm is that the next iterate is defined by making use of the last two iterates. It also turns out that the method is a generalization of the classical proximal-point one (see [30]). Since its

introduction, one can notice an increasing interest in the class of inertial type algorithms, see [1, 3, 3, 17, 24–26]. Especially noticeable is that these ideas where also used in the context of determining the zeros of the sum of a maximally monotone operator and a (single-valued) cocoercive operator, giving rise to the so-called inertial forward-backward algorithm [26]. This is an extension of the classical *forward-backward algorithm* (see [6,20]) and assumes the evaluation of the set-valued operator via its resolvent, called backward step, while the single-valued operator is evaluated via a forward step.

The first major aim of this manuscript to introduce and investigate an inertial forward-backward-forward splitting algorithm for finding the zeros of the sum of a maximally monotone operator and a monotone and Lipschitzian operator. The proposed scheme represents an extension of Tseng's *forward-backward-forward-type algorithm*, (see [6, 16, 32, 33]), however, for the study of its convergence properties we will use some generalizations of the Fejér monotonicity techniques provided in [3]. An essential argument in the favor of forward-backward-forward splitting algorithms is given by the fact that they can be used when solving a larger class of monotone inclusion problems, since it is known that there exist monotone and Lipschitzian operators which are not cocoercive, in which case the forward-backward algorithms cannot be applied (see [10, 16, 21]). This is for instance the case when considering primal-dual splitting methods, as one can notice by consulting [10, 16, 21].

Primal-dual splitting algorithms are modern techniques designed to solve inclusion problems where some complex structures of monotone operators are involved, such as mixtures of linearly composed and parallel-sum type monotone operators. The key feature of these algorithms is that they are fully decomposable, in the sense that each of the operators are evaluated in the algorithm separately, either via forward or via backward steps. It is also noticeable that the primal-dual algorithms solve concomitantly a (primal) monotone inclusion problem and its dual monotone inclusion problem in the sense of Attouch-Théra [5]. We invite the reader to consult [10, 11, 14–16, 19, 21, 22, 34] for further considerations concerning this class of algorithms. The second major aim of this paper will be to formulate an inertial primal-dual splitting algorithm relying on the inertial forward-backward-forward one.

Primal-dual splitting algorithms with inertial and memory effects relying on the Douglas-Rachford method have been recently introduced in [28] and [12]. Their advantages over corresponding non-inertial versions has been illustrated by numerical experiments in imaging (see [28]) and in clustering and optimal location selection (see [12]). Inertial algorithms prove to outperform their non-inertial versions counterparts even in the context of solving nonsmooth nonconvex optimization problems via the forward-backward method as pointed out in [13, 27].

The structure of the paper is the following. The remainder of this section is dedicated to some elements of the theory of maximal monotone operators and to the recall of some convergence results. In the next section we formulate the inertial forward-backward-forward splitting algorithm for finding the zeros of the sum of a maximally monotone operator and a monotone and Lipschitzian operator and investigate its convergence. In Section 3 we use the product space approach in order to obtain the inertial primal-dual splitting algorithm designed for solving monotone inclusion problems involving mixtures of linearly composed and parallel-sum type monotone operators. Finally, we show how the proposed iterative schemes can be used in order to solve primal-dual pairs of convex optimization problems and illustrate this approach by numerical experiments in image

denoising and deblurring.

For the notions and results presented as follows we refer the reader to [6–8, 23, 31, 35]. Let $\mathbb{N} = \{0, 1, 2, ...\}$ be the set of nonnegative integers. Let $\mathcal{H}$ be a real Hilbert space with *inner product* $\langle \cdot, \cdot \rangle$ and associated *norm* $\|\cdot\| = \sqrt{\langle \cdot, \cdot \rangle}$. The symbols $\rightharpoonup$ and $\rightarrow$ denote weak and strong convergence, respectively. When $\mathcal{G}$ is another Hilbert space and $K : \mathcal{H} \rightarrow \mathcal{G}$ a linear continuous operator, then the *norm* of K is defined as $\|K\| = \sup\{\|Kx\| : x \in \mathcal{H}, \|x\| \leq 1\}$, while $K^* : \mathcal{G} \rightarrow \mathcal{H}$, defined by $\langle K^*y, x \rangle = \langle y, Kx \rangle$ for all $(x, y) \in \mathcal{H} \times \mathcal{G}$, denotes the *adjoint operator* of K.

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*, by $\operatorname{dom} A = \{x \in \mathcal{H} : Ax \neq \emptyset\}$ its *domain*, by $\operatorname{ran} A = \cup_{x \in \mathcal{H}} Ax$ its *range* and by $A^{-1} : \mathcal{H} \rightrightarrows \mathcal{H}$ its *inverse operator*, defined by $(u, x) \in \operatorname{Gr} A^{-1}$ if and only if $(x, u) \in \operatorname{Gr} A$. 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 $J_A = (\operatorname{Id}_{\mathcal{H}} + A)^{-1}$, where $\operatorname{Id}_{\mathcal{H}} : \mathcal{H} \rightarrow \mathcal{H}, \operatorname{Id}_{\mathcal{H}}(x) = x$ for all $x \in \mathcal{H}$, is the *identity operator* on $\mathcal{H}$. Moreover, if A is maximally monotone, then $J_A : \mathcal{H} \rightarrow \mathcal{H}$ is single-valued and maximally monotone (see [6, Proposition 23.7 and Corollary 23.10]). For an arbitrary $\gamma > 0$ we have (see [6, Proposition 23.2])

$$p \in J_{\gamma A} x \text{ if and only if } (p, \gamma^{-1}(x - p)) \in \operatorname{Gr} A$$

and (see [6, Proposition 23.18])

$$J_{\gamma A} + \gamma J_{\gamma^{-1} A^{-1}} \circ \gamma^{-1} \operatorname{Id}_{\mathcal{H}} = \operatorname{Id}_{\mathcal{H}}. \tag{1}$$

Further, let us mention some classes of operators that are used in the paper. We say that A is *demiregular* at $x \in \operatorname{dom} A$ if, for every sequence $(x_n, u_n)_{n \in \mathbb{N}} \in \operatorname{Gr} A$ and every $u \in Ax$ such that $x_n \rightharpoonup x$ and $u_n \rightarrow u$, we have $x_n \rightarrow x$. We refer the reader to [4, Proposition 2.4] and [16, Lemma 2.4] for conditions ensuring this property. The operator A is said to be *uniformly monotone* at $x \in \operatorname{dom} A$ if there exists an increasing function $\phi_A : [0, +\infty) \rightarrow [0, +\infty]$ that vanishes only at 0, and $\langle x - y, u - v \rangle \geq \phi_A(\|x - y\|)$ for every $u \in Ax$ and $(y, v) \in \operatorname{Gr} A$. If this inequality holds for all $(x, u), (y, v) \in \operatorname{Gr} A$, we say that A is uniformly monotone. If A is uniformly monotone at $x \in \operatorname{dom} A$, then it is demiregular at x.

Prominent representatives of the class of uniformly monotone operators are the strongly monotone operators. Let $\gamma > 0$ be arbitrary. We say that A is γ-*strongly monotone*, if $\langle x - y, u - v \rangle \geq \gamma \|x - y\|^2$ for all $(x, u), (y, v) \in \operatorname{Gr} A$. Further, a single-valued operator $A : \mathcal{H} \rightarrow \mathcal{H}$ is said to be γ-*cocoercive* if $\langle x - y, Ax - Ay \rangle \geq \gamma \|Ax - Ay\|^2$ for all $(x, y) \in \mathcal{H} \times \mathcal{H}$. Moreover, A is γ-*Lipschitzian* if $\|Ax - Ay\| \leq \gamma \|x - y\|$ for all $(x, y) \in \mathcal{H} \times \mathcal{H}$. A single-valued linear operator $A : \mathcal{H} \rightarrow \mathcal{H}$ is said to be *skew*, if $\langle x, Ax \rangle = 0$ for all $x \in \mathcal{H}$. Finally, the *parallel sum* of two operators $A, B : \mathcal{H} \rightrightarrows \mathcal{H}$ is defined by $A \square B : \mathcal{H} \rightrightarrows \mathcal{H}, A \square B = (A^{-1} + B^{-1})^{-1}$.

We close this section by presenting three convergence results which will be crucial for the proof of the main results in the next section.

Lemma 1 *(see [1–3]) Let $(\varphi_n)_{n \in \mathbb{N}}, (\delta_n)_{n \in \mathbb{N}}$ and $(\alpha_n)_{n \in \mathbb{N}}$ be sequences in $[0, +\infty)$ such that $\varphi_{n+1} \leq \varphi_n + \alpha_n(\varphi_n - \varphi_{n-1}) + \delta_n$ for all $n \geq 1$, $\sum_{n \in \mathbb{N}} \delta_n < +\infty$ and there exists a real number α with $0 \leq \alpha_n \leq \alpha < 1$ for all $n \in \mathbb{N}$. Then the following hold:*

(i) $\sum_{n\geq 1}[\varphi_n - \varphi_{n-1}]_+ < +\infty$, where $[t]_+ = \max\{t, 0\}$;

(ii) there exists $\varphi^* \in [0, +\infty)$ such that $\lim_{n\to+\infty} \varphi_n = \varphi^*$.

An easy consequence of Lemma 1 is the following result.

Lemma 2 Let $(\varphi_n)_{n\in\mathbb{N}}, (\delta_n)_{n\in\mathbb{N}}, (\alpha_n)_{n\in\mathbb{N}}$ and $(\beta_n)_{n\in\mathbb{N}}$ be sequences in $[0, +\infty)$ such that $\varphi_{n+1} \leq -\beta_n + \varphi_n + \alpha_n(\varphi_n - \varphi_{n-1}) + \delta_n$ for all $n \geq 1$, $\sum_{n\in\mathbb{N}} \delta_n < +\infty$ and there exists a real number α with $0 \leq \alpha_n \leq \alpha < 1$ for all $n \in \mathbb{N}$. Then the following hold:

(i) $\sum_{n\geq 1}[\varphi_n - \varphi_{n-1}]_+ < +\infty$, where $[t]_+ = \max\{t, 0\}$;

(ii) there exists $\varphi^* \in [0, +\infty)$ such that $\lim_{n\to+\infty} \varphi_n = \varphi^*$;

(iii) $\sum_{n\in\mathbb{N}} \beta_n < +\infty$.

Finally, we recall a well known result on weak convergence in Hilbert spaces.

Lemma 3 (Opial) Let C be a nonempty set of $\mathcal{H}$ and $(x_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{H}$ such that the following two conditions hold:

(a) for every $x \in C$, $\lim_{n\to+\infty} \|x_n - x\|$ exists;

(b) every sequential weak cluster point of $(x_n)_{n\in\mathbb{N}}$ is in C;

Then $(x_n)_{n\in\mathbb{N}}$ converges weakly to a point in C.

2 An inertial forward-backward-forward splitting algorithm

This section is dedicated to the formulation of an inertial forward-backward-forward splitting algorithm which approaches the set of zeros of the sum of two maximally monotone operators, one of them being single-valued and Lipschitzian, and to the investigation of its convergence properties.

Theorem 4 Let $A : \mathcal{H} \rightrightarrows \mathcal{H}$ be a maximally monotone operator and $B : \mathcal{H} \to \mathcal{H}$ a monotone and β-Lipschitzian operator for some $\beta > 0$. Suppose that $\mathrm{zer}(A + B) \neq \emptyset$ and consider the following iterative scheme:

$$(\forall n \geq 1) \quad \begin{cases} p_n = J_{\lambda_n A}[x_n - \lambda_n B x_n + \alpha_{1,n}(x_n - x_{n-1})] \\ x_{n+1} = p_n + \lambda_n(B x_n - B p_n) + \alpha_{2,n}(x_n - x_{n-1}), \end{cases}$$

where x_0 and x_1 are arbitrarily chosen in $\mathcal{H}$. Consider $\lambda, \sigma > 0$ and $\alpha_1, \alpha_2 \geq 0$ such that

$$12\alpha_2^2 + 9(\alpha_1 + \alpha_2) + 4\sigma < 1 \text{ and } \lambda \leq \lambda_n \leq \frac{1}{\beta}\sqrt{\frac{1 - 12\alpha_2^2 - 9(\alpha_1 + \alpha_2) - 4\sigma}{12\alpha_2^2 + 8(\alpha_1 + \alpha_2) + 4\sigma + 2}} \; \forall n \geq 1 \quad (2)$$

and for $i = 1, 2$ the nondecreasing sequences $(\alpha_{i,n})_{n\geq 1}$ with $\alpha_{i,1} = 0$ and $0 \leq \alpha_{i,n} \leq \alpha_i$ for all $n \geq 1$. Then there exists $\overline{x} \in \mathrm{zer}(A + B)$ such that the following statements are true:

(a) $\sum_{n\in\mathbb{N}} \|x_{n+1} - x_n\|^2 < +\infty$ and $\sum_{n\geq 1} \|x_n - p_n\|^2 < +\infty$;

(b) $x_n \rightharpoonup \overline{x}$ and $p_n \rightharpoonup \overline{x}$ as $n \to +\infty$;

4

(c) Suppose that one of the following conditions is satisfied:

(i) $A + B$ is demiregular at $\bar{x}$;

(ii) A or B is uniformly monotone at $\bar{x}$.

Then $x_n \to \bar{x}$ and $p_n \to \bar{x}$ as $n \to +\infty$.

Proof. Let z be a fixed element in $\mathrm{zer}(A + B)$, that is $-Bz \in Az$, and $n \geq 1$. From the definition of the resolvent we deduce

$$\frac{1}{\lambda_n}(x_n - p_n) - Bx_n + \frac{\alpha_{1,n}}{\lambda_n}(x_n - x_{n-1}) \in Ap_n.$$

Further, taking into account the relation between p_n and x_{n+1} in the algorithm, we obtain

$$\frac{1}{\lambda_n}(x_n - x_{n+1}) - Bp_n + \frac{\alpha_{1,n} + \alpha_{2,n}}{\lambda_n}(x_n - x_{n-1}) \in Ap_n. \tag{3}$$

The monotonicity of A delivers the inequality

$$0 \leq \left\langle \frac{1}{\lambda_n}(x_n - x_{n+1}) - Bp_n + \frac{\alpha_{1,n} + \alpha_{2,n}}{\lambda_n}(x_n - x_{n-1}) + Bz, p_n - z \right\rangle,$$

hence

$$0 \leq \frac{1}{\lambda_n}\langle x_n - x_{n+1}, p_n - z \rangle + \langle Bz - Bp_n, p_n - z \rangle + \frac{\alpha_{1,n} + \alpha_{2,n}}{\lambda_n}\langle x_n - x_{n-1}, p_n - z \rangle. \tag{4}$$

Since B is monotone, we have $\langle Bz - Bp_n, p_n - z \rangle \leq 0$. Moreover,

$$\langle x_n - x_{n+1}, p_n - z \rangle = \langle x_n - x_{n+1}, p_n - x_{n+1} \rangle + \langle x_n - x_{n+1}, x_{n+1} - z \rangle =$$

$$\frac{\|x_n - x_{n+1}\|^2}{2} + \frac{\|p_n - x_{n+1}\|^2}{2} - \frac{\|x_n - p_n\|^2}{2} + \frac{\|x_n - z\|^2}{2} - \frac{\|x_n - x_{n+1}\|^2}{2} - \frac{\|x_{n+1} - z\|^2}{2}.$$

In a similar way we obtain

$$\langle x_n - x_{n-1}, p_n - z \rangle = \langle x_n - x_{n-1}, x_n - z \rangle + \langle x_n - x_{n-1}, p_n - x_n \rangle =$$

$$\frac{\|x_n - x_{n-1}\|^2}{2} + \frac{\|x_n - z\|^2}{2} - \frac{\|x_{n-1} - z\|^2}{2} + \frac{\|p_n - x_{n-1}\|^2}{2} - \frac{\|x_n - x_{n-1}\|^2}{2} - \frac{\|x_n - p_n\|^2}{2}.$$

Further we have, by using that B is β-Lipschitzian,

$$\|x_{n+1} - p_n\|^2 \leq 2\lambda_n^2\beta^2\|x_n - p_n\|^2 + 2\alpha_{2,n}^2\|x_n - x_{n-1}\|^2$$

and

$$\|p_n - x_{n-1}\|^2 \leq 2\|x_n - p_n\|^2 + 2\|x_n - x_{n-1}\|^2.$$

The above estimates together with (4) imply

$$0 \leq \left(\frac{1}{2\lambda_n} + \frac{\alpha_{1,n} + \alpha_{2,n}}{2\lambda_n} \right)\|x_n - z\|^2 - \frac{1}{2\lambda_n}\|x_{n+1} - z\|^2 - \frac{\alpha_{1,n} + \alpha_{2,n}}{2\lambda_n}\|x_{n-1} - z\|^2 +$$

$$\left(\lambda_n\beta^2 - \frac{1}{2\lambda_n} + \frac{\alpha_{1,n} + \alpha_{2,n}}{\lambda_n} - \frac{\alpha_{1,n} + \alpha_{2,n}}{2\lambda_n} \right)\|x_n - p_n\|^2 +$$

$$\left(\frac{\alpha_{2,n}^2}{\lambda_n} + \frac{\alpha_{1,n} + \alpha_{2,n}}{\lambda_n} \right)\|x_n - x_{n-1}\|^2,$$

5

from which we further obtain, after multiplying with $2\lambda_n$,

$$\|x_{n+1} - z\|^2 - (1 + \alpha_{1,n} + \alpha_{2,n})\|x_n - z\|^2 + (\alpha_{1,n} + \alpha_{2,n})\|x_{n-1} - z\|^2 \leq$$
$$-(1 - \alpha_{1,n} - \alpha_{2,n} - 2\lambda_n^2\beta^2)\|x_n - p_n\|^2 + 2(\alpha_{2,n}^2 + \alpha_{1,n} + \alpha_{2,n})\|x_n - x_{n-1}\|^2. \quad (5)$$

By using the bounds given for the sequences $(\lambda_n)_{n\geq 1}$, $(\alpha_{1,n})_{n\geq 1}$ and $(\alpha_{2,n})_{n\geq 1}$ one can easily show by taking into account (2) that

$$2\lambda_n^2\beta^2 < 1 - \alpha_1 - \alpha_2 \leq 1 - \alpha_{1,n} - \alpha_{2,n},$$

thus

$$1 - \alpha_{1,n} - \alpha_{2,n} - 2\lambda_n^2\beta^2 > 0.$$

Taking into account that

$$\|x_{n+1} - x_n\|^2 = \|p_n - x_n + \lambda_n(Bx_n - Bp_n) + \alpha_{2,n}(x_n - x_{n-1})\|^2$$
$$\leq 2(1 + \lambda_n\beta)^2\|x_n - p_n\|^2 + 2\alpha_{2,n}^2\|x_n - x_{n-1}\|^2,$$

we obtain from (5)

$$\|x_{n+1} - z\|^2 - (1 + \alpha_{1,n} + \alpha_{2,n})\|x_n - z\|^2 + (\alpha_{1,n} + \alpha_{2,n})\|x_{n-1} - z\|^2 \leq$$
$$-\frac{1 - \alpha_{1,n} - \alpha_{2,n} - 2\lambda_n^2\beta^2}{2(1 + \lambda_n\beta)^2}\|x_{n+1} - x_n\|^2 + \gamma_n\|x_n - x_{n-1}\|^2, \quad (6)$$

where

$$\gamma_n := 2(\alpha_{2,n}^2 + \alpha_{1,n} + \alpha_{2,n}) + \frac{\alpha_{2,n}^2(1 - \alpha_{1,n} - \alpha_{2,n} - 2\lambda_n^2\beta^2)}{(1 + \lambda_n\beta)^2} > 0.$$

(a) For the proof of this statement we are going to use some techniques from [3]. We define the sequences $\varphi_n := \|x_n - z\|^2$ for all $n \in \mathbb{N}$ and $\mu_n := \varphi_n - (\alpha_{1,n} + \alpha_{2,n})\varphi_{n-1} + \gamma_n\|x_n - x_{n-1}\|^2$ for all $n \geq 1$. Using the monotonicity of $(\alpha_{i,n})_{n\geq 1}$, $i = 1, 2$, and the fact that $\varphi_n \geq 0$ for all $n \in \mathbb{N}$ we get

$$\mu_{n+1} - \mu_n \leq$$
$$\varphi_{n+1} - (1 + \alpha_{1,n} + \alpha_{2,n})\varphi_n + (\alpha_{1,n} + \alpha_{2,n})\varphi_{n-1} + \gamma_{n+1}\|x_{n+1} - x_n\|^2 - \gamma_n\|x_n - x_{n-1}\|^2,$$

which gives by (6)

$$\mu_{n+1} - \mu_n \leq -\left(\frac{1 - \alpha_{1,n} - \alpha_{2,n} - 2\lambda_n^2\beta^2}{2(1 + \lambda_n\beta)^2} - \gamma_{n+1}\right)\|x_{n+1} - x_n\|^2 \ \forall n \geq 1. \quad (7)$$

We claim that

$$\frac{1 - \alpha_{1,n} - \alpha_{2,n} - 2\lambda_n^2\beta^2}{2(1 + \lambda_n\beta)^2} - \gamma_{n+1} \geq \sigma \ \forall n \geq 1. \quad (8)$$

Indeed, this follows by taking into account that for all $n \geq 1$

$$\alpha_{1,n} + \alpha_{2,n} + 2(\lambda_n\beta)^2 + 2(1 + \lambda_n\beta)^2(\gamma_{n+1} + \sigma) \leq$$
$$\alpha_1 + \alpha_2 + 2(\lambda_n\beta)^2 + 2(1 + \lambda_n\beta)^2(3\alpha_2^2 + 2(\alpha_1 + \alpha_2) + \sigma) \leq$$
$$\alpha_1 + \alpha_2 + 2(\lambda_n\beta)^2 + 4(1 + (\lambda_n\beta)^2)(3\alpha_2^2 + 2(\alpha_1 + \alpha_2) + \sigma) \leq 1.$$

In the above estimates we used the upper bounds for $(\alpha_{i,n})_{n\geq 1}$, $i = 1, 2$, that

$$\gamma_{n+1} \leq 2(\alpha_2^2 + \alpha_1 + \alpha_2) + \alpha_2^2 \ \forall n \in \mathbb{N}$$

and the assumptions in (2).

We obtain from (7) and (8) that

$$\mu_{n+1} - \mu_n \leq -\sigma \|x_{n+1} - x_n\|^2 \ \forall n \geq 1. \tag{9}$$

The sequence $(\mu_n)_{n\geq 1}$ is nonincreasing and the bounds for $(\alpha_{i,n})_{n\geq 1}$, $i = 1, 2$, deliver

$$-(\alpha_1 + \alpha_2)\varphi_{n-1} \leq \varphi_n - (\alpha_1 + \alpha_2)\varphi_{n-1} \leq \mu_n \leq \mu_1 \ \forall n \geq 1. \tag{10}$$

We obtain

$$\varphi_n \leq (\alpha_1 + \alpha_2)^n \varphi_0 + \mu_1 \sum_{k=0}^{n-1} (\alpha_1 + \alpha_2)^k \leq (\alpha_1 + \alpha_2)^n \varphi_0 + \frac{\mu_1}{1 - \alpha_1 - \alpha_2} \ \forall n \geq 1,$$

where we notice that $\mu_1 = \varphi_1 \geq 0$ (due to the relation $\alpha_{1,1} = \alpha_{2,1} = 0$). Combining (9) and (10) we get for all $n \geq 1$

$$\sigma \sum_{k=1}^{n} \|x_{k+1} - x_k\|^2 \leq \mu_1 - \mu_{n+1} \leq \mu_1 + (\alpha_1 + \alpha_2)\varphi_n \leq (\alpha_1 + \alpha_2)^{n+1} \varphi_0 + \frac{\mu_1}{1 - \alpha_1 - \alpha_2},$$

which shows that $\sum_{n \in \mathbb{N}} \|x_{n+1} - x_n\|^2 < +\infty$.

Combining this relation with (5) and Lemma 2 it yields

$$\sum_{n \geq 1} \left(1 - \alpha_{1,n} - \alpha_{2,n} - 2\lambda_n^2 \beta^2\right) \|x_n - p_n\|^2 < +\infty.$$

Moreover, from (8) we have $1 - \alpha_{1,n} - \alpha_{2,n} - 2\lambda_n^2 \beta^2 \geq 2\sigma(1 + \lambda\beta)^2$ for all $n \geq 1$ and obtain, consequently, $\sum_{n \geq 1} \|x_n - p_n\|^2 < +\infty$.

(b) We are going to use Lemma 3. We proved above that for an arbitrary $z \in \operatorname{zer}(A+B)$ the inequality (5) is true. By part (a) and Lemma 2 it follows that $\lim_{n \to +\infty} \|x_n - z\|$ exists. On the other hand, let x be a sequential weak cluster point of $(x_n)_{n \in \mathbb{N}}$, that is, it has a subsequence $(x_{n_k})_{k \in \mathbb{N}}$ fulfilling $x_{n_k} \rightharpoonup x$ as $k \to +\infty$. Since $x_n - p_n \to 0$ as $n \to +\infty$, we get $p_{n_k} \rightharpoonup x$ as $k \to +\infty$. Since $A + B$ is maximally monotone (see [6, Corollary 20.25 and Corollary 24.4]), its graph is sequentially closed in the weak-strong topology of $\mathcal{H} \times \mathcal{H}$ (see [6, Proposition 20.33(ii)]). As $(\lambda_n)_{n \geq 1}$ and $(\alpha_{i,n})_{n \geq 1}$, $i = 1, 2$, are bounded, we derive from (3) and part (a) that $0 \in (A+B)x$, hence $x \in \operatorname{zer}(A + B)$. By Lemma 3 there exists $\overline{x} \in \operatorname{zer}(A+B)$ such that $x_n \rightharpoonup \overline{x}$ as $n \to +\infty$. In view of (a) we have $p_n \rightharpoonup \overline{x}$ as $n \to +\infty$.

(c) Since (ii) implies that $A + B$ is uniformly monotone at $\overline{x}$, hence demiregular at $\overline{x}$, it is sufficient to prove the statement under condition (i). Since $p_n \rightharpoonup \overline{x}$ and $\frac{1}{\lambda_n}(x_n - x_{n+1}) + \frac{\alpha_{1,n} + \alpha_{2,n}}{\lambda_n}(x_n - x_{n-1}) \to 0$ as $n \to +\infty$, the result follows easily from (3) and the definition of demiregular operators. ∎

Remark 5 Let us mention that the conclusion of the theorem holds also in case one assumes that the sequence $(\alpha_{1,n} + \alpha_{2,n})_{n \geq 1}$ is nondecreasing. Moreover, the condition $\alpha_{1,1} = \alpha_{2,1} = 0$ was imposed in order to ensure $\mu_1 \geq 0$, which is needed in the proof. An alternative is to require that $x_0 = x_1$, in which case the assumption $\alpha_{1,1} = \alpha_{2,1} = 0$ is not anymore necessary.

Remark 6 Assuming that $\alpha_2 = 0$, which enforces $\alpha_{2,n} = 0$ for all $n \geq 1$, the conclusions of Theorem 4 remains valid if one takes as upper bound for $(\lambda_n)_{n \geq 1}$ the expression $\frac{1}{\beta} \sqrt{\frac{1 - 5\alpha_1 - 2\sigma}{4\alpha_1 + 2\sigma + 1}}$. This is due to the fact in this situation one can use in its proof the improved inequalities $\|x_{n+1} - p_n\|^2 \leq \lambda_n^2 \beta^2 \|x_n - p_n\|^2$ and $\|x_{n+1} - x_n\|^2 \leq (1 + \lambda_n \beta)^2 \|x_n - p_n\|^2$ for all $n \geq 1$. On the other hand, let us also notice that the algorithmic scheme obtained in this way and its convergence properties can be seen as generalizations of the corresponding statements given for the error-free case of the classical forward-backward-forward algorithm proposed by Tseng in [33] (see also [16, Theorem 2.5]). Indeed, if we further set $\alpha_1 = 0$, having as consequence that $\alpha_{1,n} = 0$ for all $n \geq 1$, we obtain nothing else than the iterative scheme from [16,33]. Notice that for $\varepsilon \in (0, 1/(\beta+1))$, one can chose $\lambda := \varepsilon$ and $\sigma := \frac{1 - (1-\varepsilon)^2}{2(1 + (1-\varepsilon)^2)}$. In this case the sequence $(\lambda_n)_{n \geq 1}$ must fulfill the inequalities $\varepsilon \leq \lambda_n \leq \frac{1}{\beta} \sqrt{\frac{1 - 2\sigma}{2\sigma + 1}} = \frac{1-\varepsilon}{\beta}$ for all $n \geq 1$, which is exactly the situation considered in [16].

Remark 7 In case $Bx = 0$ for all $x \in \mathcal{H}$ the proposed iterative scheme becomes

$$x_{n+1} = J_{\lambda_n A}[x_n + \alpha_{1,n}(x_n - x_{n-1})] + \alpha_{2,n}(x_n - x_{n-1}) \ \forall n \geq 1,$$

and is to the best of our knowledge new and can be regarded as an extension of the classical proximal-point algorithm (see [30]) in the context of solving the monotone inclusion problem $0 \in Ax$. If, additionally, $\alpha_2 = 0$, which enforces as already noticed $\alpha_{2,n} = 0$ for all $n \geq 1$, we get the algorithm

$$x_{n+1} = J_{\lambda_n A}[x_n + \alpha_{1,n}(x_n - x_{n-1})],$$

the convergence of which has been investigated in [3].

3 Solving monotone inclusion problems involving mixtures of linearly composed and parallel-sum type operators

In this section we employ the inertial forward-backward-forward splitting algorithm proposed above to the concomitantly solving of a primal monotone inclusion problem involving mixtures of linearly composed and parallel-sum type operators and its Attouch-Théra-type dual problem. We consider the following setting.

Problem 8 Let $\mathcal{H}$ be a real Hilbert space, $z \in \mathcal{H}$, $A : \mathcal{H} \rightrightarrows \mathcal{H}$ a maximally monotone operator and $C : \mathcal{H} \to \mathcal{H}$ a monotone and μ-Lipschitzian operator for $\mu > 0$. Let m be a strictly positive integer and, for any $i \in \{1,...,m\}$, let $\mathcal{G}_i$ be a real Hilbert space, $r_i \in \mathcal{G}_i$, let $B_i : \mathcal{G}_i \rightrightarrows \mathcal{G}_i$ be a maximally monotone operator, let $D_i : \mathcal{G}_i \rightrightarrows \mathcal{G}_i$ be monotone such that D_i^{-1} is ν_i-Lipschitzian for $\nu_i > 0$ and let $L_i : \mathcal{H} \to \mathcal{G}_i$ be a nonzero linear continuous operator. The problem is to solve the primal inclusion

$$\text{find } \overline{x} \in \mathcal{H} \text{ such that } z \in A\overline{x} + \sum_{i=1}^{m} L_i^* \big((B_i \square D_i)(L_i \overline{x} - r_i) \big) + C\overline{x}, \tag{11}$$

together with the dual inclusion

$$\text{find } \overline{v}_1 \in \mathcal{G}_1,..., \overline{v}_m \in \mathcal{G}_m \text{ such that } \exists x \in \mathcal{H} : \begin{cases} z - \sum_{i=1}^{m} L_i^* \overline{v}_i \in Ax + Cx \\ \overline{v}_i \in (B_i \square D_i)(L_i x - r_i), \ i = 1,...,m. \end{cases} \tag{12}$$

8

We say that $(\bar{x}, \bar{v}_1,..., \bar{v}_m) \in \mathcal{H} \times \mathcal{G}_1 \times ... \times \mathcal{G}_m$ is a primal-dual solution to Problem 8, if

$$z - \sum_{i=1}^{m} L_i^* \bar{v}_i \in A\bar{x} + C\bar{x} \text{ and } \bar{v}_i \in (B_i \square D_i)(L_i\bar{x} - r_i), \ i = 1,..., m. \tag{13}$$

If $\bar{x} \in \mathcal{H}$ is a solution to (11), then there exists $(\bar{v}_1,..., \bar{v}_m) \in \mathcal{G}_1 \times ... \times \mathcal{G}_m$ such that $(\bar{x}, \bar{v}_1,..., \bar{v}_m)$ is a primal-dual solution to Problem 8 and, if $(\bar{v}_1,..., \bar{v}_m) \in \mathcal{G}_1 \times ... \times \mathcal{G}_m$ is a solution to (12), then there exists $\bar{x} \in \mathcal{H}$ such that $(\bar{x}, \bar{v}_1,..., \bar{v}_m)$ is a primal-dual solution to Problem 8. Moreover, if $(\bar{x}, \bar{v}_1,..., \bar{v}_m) \in \mathcal{H} \times \mathcal{G}_1 \times ... \times \mathcal{G}_m$ is a primal-dual solution to Problem 8, then $\bar{x}$ is a solution to (11) and $(\bar{v}_1,..., \bar{v}_m) \in \mathcal{G}_1 \times ... \times \mathcal{G}_m$ is a solution to (12).

Problem 8 covers a large class of monotone inclusion problems and we refer the reader to consult [21] for several interesting particular instances of it. The main result of this section follows.

Theorem 9 *In Problem 8 suppose that*

$$z \in \mathrm{ran}\left(A + \sum_{i=1}^{m} L_i^*\big((B_i \square D_i)(L_i \cdot -r_i)\big) + C\right). \tag{14}$$

Chose $x_0, x_1 \in \mathcal{H}$ and $v_{i,0}, v_{i,1} \in \mathcal{G}_i$, $i = 1, ..., m$, and set

$$(\forall n \geq 1) \begin{cases} p_{1,n} = & J_{\lambda_n A}[x_n - \lambda_n(Cx_n + \sum_{i=1}^{m} L_i^* v_{i,n} - z) + \alpha_{1,n}(x_n - x_{n-1})] \\ p_{2,i,n} = & J_{\lambda_n B_i^{-1}}[v_{i,n} + \lambda_n(L_i x_n - D_i^{-1} v_{i,n} - r_i) + \alpha_{1,n}(v_{i,n} - v_{i,n-1})], \\ & i = 1, ..., m \\ v_{i,n+1} = & \lambda_n L_i(p_{1,n} - x_n) + \lambda_n(D_i^{-1} v_{i,n} - D_i^{-1} p_{2,i,n}) + p_{2,i,n} \\ & + \alpha_{2,n}(v_{i,n} - v_{i,n-1}), i = 1, ..., m \\ x_{n+1} = & \lambda_n \sum_{i=1}^{m} L_i^*(v_{i,n} - p_{2,i,n}) + \lambda_n(Cx_n - Cp_{1,n}) + p_{1,n} \\ & + \alpha_{2,n}(x_n - x_{n-1}). \end{cases}$$

Consider $\lambda, \sigma > 0$ and $\alpha_1, \alpha_2 \geq 0$ such that

$$12\alpha_2^2 + 9(\alpha_1 + \alpha_2) + 4\sigma < 1 \text{ and } \lambda \leq \lambda_n \leq \frac{1}{\beta}\sqrt{\frac{1 - 12\alpha_2^2 - 9(\alpha_1 + \alpha_2) - 4\sigma}{12\alpha_2^2 + 8(\alpha_1 + \alpha_2) + 4\sigma + 2}} \ \forall n \geq 1,$$

where

$$\beta = \max\{\mu, \nu_1, ..., \nu_m\} + \sqrt{\sum_{i=1}^{m} \|L_i\|^2},$$

and for $i = 1, 2$ the nondecreasing sequences $(\alpha_{i,n})_{n \geq 1}$ with $\alpha_{i,1} = 0$ and $0 \leq \alpha_{i,n} \leq \alpha_i$ for all $n \geq 1$. Then the following statements are true:

(a) $\sum_{n \in \mathbb{N}} \|x_{n+1} - x_n\|^2 < +\infty$, $\sum_{n \geq 1} \|x_n - p_{1,n}\|^2 < +\infty$ and, for $i = 1, ..., m$, $\sum_{n \in \mathbb{N}} \|v_{i,n+1} - v_{i,n}\|^2 < +\infty$ and $\sum_{n \geq 1} \|v_{i,n} - p_{2,i,n}\|^2 < +\infty$;

(b) There exists $(\bar{x}, \bar{v}_1,..., \bar{v}_m) \in \mathcal{H} \times \mathcal{G}_1 \times ... \times \mathcal{G}_m$ a primal-dual solution to Problem 8 such that the following hold:

(i) $x_n \rightharpoonup \overline{x}$, $p_{1,n} \rightharpoonup \overline{x}$ and, for $i = 1, ..., m$, $v_{i,n} \rightharpoonup \overline{v}_i$ and $p_{2,i,n} \rightharpoonup \overline{v}_i$ as $n \to +\infty$;

(ii) If $A + C$ is uniformly monotone at $\overline{x}$, then $x_n \to \overline{x}$ and $p_{1,n} \to \overline{x}$ as $n \to +\infty$.

(iii) If $B_i^{-1} + D_i^{-1}$ is uniformly monotone at $\overline{v}_i$ for some $i \in \{1, ..., m\}$, then $v_{i,n} \to \overline{v}_i$ and $p_{2,i,n} \to \overline{v}_i$ as $n \to +\infty$.

Proof. We will apply Theorem 4 in an appropriate product space and will make use to this end of a construction similar to the one considered in [21]. We endow the product space $\mathcal{K} = \mathcal{H} \times \mathcal{G}_1 \times ... \times \mathcal{G}_m$ with the inner product and the associated norm defined for all $(x, v_1, ..., v_m), (y, w_1, ..., w_m) \in \mathcal{K}$ as

$$\langle (x, v_1, ..., v_m), (y, w_1, ..., w_m) \rangle_{\mathcal{K}} = \langle x, y \rangle_{\mathcal{H}} + \sum_{i=1}^m \langle v_i, w_i \rangle_{\mathcal{G}_i}$$

and

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

respectively.

We introduce the operators $\boldsymbol{M} : \mathcal{K} \rightrightarrows \mathcal{K}$,

$$\boldsymbol{M}(x, v_1, ..., v_m) = (-z + Ax) \times (r_1 + B_1^{-1} v_1) \times .... \times (r_m + B_m^{-1} v_m)$$

and $\boldsymbol{Q} : \mathcal{K} \to \mathcal{K}$,

$$\boldsymbol{Q}(x, v_1, ..., v_m) = \left(Cx + \sum_{i=1}^m L_i^* v_i, -L_1 x + D_1^{-1} v_1, ..., -L_m x + D_m^{-1} v_m \right)$$

and show that Theorem 4 can be applied for the operators $\boldsymbol{M}$ and $\boldsymbol{Q}$ in the product space $\mathcal{K}$. Let us start by noticing that

$$(14) \Leftrightarrow \mathrm{zer}(\boldsymbol{M} + \boldsymbol{Q}) \neq \emptyset$$

and

$(x, v_1, ..., v_m) \in \mathrm{zer}(\boldsymbol{M} + \boldsymbol{Q}) \Leftrightarrow (x, v_1, ..., v_m)$ is a primal-dual solution of Problem 8.

(15)

Further, since A and B_i, $i = 1, ..., m$ are maximally monotone, $\boldsymbol{M}$ is maximally monotone, too (see [6, Props. 20.22, 20.23]). On the other hand, $\boldsymbol{Q}$ is a monotone and β-Lipschitzian (see, for instance, the proof of [21, Theorem 3.1]).

For every $(x, v_1, ..., v_m) \in \mathcal{K}$ and every $\lambda > 0$ we have (see [6, Proposition 23.16])

$$J_{\lambda \boldsymbol{M}}(x, v_1, ..., v_m) = (J_{\lambda A}(x + \lambda z), J_{\lambda B_1^{-1}}(v_1 - \lambda r_1), ..., J_{\lambda B_m^{-1}}(v_m - \lambda r_m)).$$

Set

$$\boldsymbol{x}_n = (x_n, v_{1,n}, ..., v_{m,n}) \ \forall n \in \mathbb{N} \text{ and } \boldsymbol{p}_n = (p_{1,n}, p_{2,1,n}, ..., p_{2,m,n}) \ \forall n \geq 1.$$

In the light of the above considerations it follows that the iterative scheme in the statement of Theorem 9 can be equivalently written as

$$\forall n \geq 1 \ \begin{cases} \boldsymbol{p}_n = J_{\lambda_n \boldsymbol{M}}[\boldsymbol{x}_n - \lambda_n \boldsymbol{Q} \boldsymbol{x}_n + \alpha_{1,n}(\boldsymbol{x}_n - \boldsymbol{x}_{n-1})] \\ \boldsymbol{x}_{n+1} = \boldsymbol{p}_n + \lambda_n(\boldsymbol{Q} \boldsymbol{x}_n - \boldsymbol{Q} \boldsymbol{p}_n) + \alpha_{2,n}(\boldsymbol{x}_n - \boldsymbol{x}_{n-1}), \end{cases}$$

which is nothing else than the algorithm stated in Theorem 4 formulated for the operators $\boldsymbol{M}$ and $\boldsymbol{Q}$.

(a) Is a direct consequence of Theorem 4(a).

(b)(i) Is a direct consequence of Theorem 4(b) and (15).

(b)(ii) Let $n \geq 1$ be fixed. From the definition of the resolvent we get

$$\frac{1}{\lambda_n}(x_n - p_{1,n}) - Cx_n - \sum_{i=1}^{m} L_i^* v_{i,n} + z + \frac{\alpha_{1,n}}{\lambda_n}(x_n - x_{n-1}) \in Ap_{1,n}.$$

The update rule for x_n yields

$$\frac{1}{\lambda_n}(p_{1,n} - x_{n+1}) + Cx_n + \sum_{i=1}^{m} L_i^*(v_{i,n} - p_{2,i,n}) + \frac{\alpha_{2,n}}{\lambda_n}(x_n - x_{n-1}) = Cp_{1,n},$$

hence,

$$\frac{1}{\lambda_n}(x_n - x_{n+1}) - \sum_{i=1}^{m} L_i^* p_{2,i,n} + z + \frac{\alpha_{1,n} + \alpha_{2,n}}{\lambda_n}(x_n - x_{n-1}) \in (A + C)p_{1,n}.$$

Further, since $z - \sum_{i=1}^{m} L_i^* \bar{v}_i \in (A + C)\bar{x}$ and $A + C$ is uniformly monotone at $\bar{x}$, there exists an increasing function $\phi_{A,C} : [0, +\infty) \to [0, +\infty]$ that vanishes only at 0, such that

$$\left\langle p_{1,n} - \bar{x}, \frac{1}{\lambda_n}(x_n - x_{n+1}) - \sum_{i=1}^{m} L_i^* p_{2,i,n} + z + \frac{\alpha_{1,n} + \alpha_{2,n}}{\lambda_n}(x_n - x_{n-1}) - \left(z - \sum_{i=1}^{m} L_i^* \bar{v}_i \right) \right\rangle$$

$$\geq \phi_{A,C}(\|p_{1,n} - \bar{x}\|),$$

thus

$$\frac{1}{\lambda_n}\langle p_{1,n} - \bar{x}, x_n - x_{n+1}\rangle + \left\langle p_{1,n} - \bar{x}, \sum_{i=1}^{m} L_i^*(\bar{v}_i - p_{2,i,n}) \right\rangle$$

$$+ \frac{\alpha_{1,n} + \alpha_{2,n}}{\lambda_n}\langle p_{1,n} - \bar{x}, x_n - x_{n-1}\rangle \geq \phi_{A,C}(\|p_{1,n} - \bar{x}\|). \tag{16}$$

In a similar way, for $i = 1, ..., m$, the definition of $p_{2,i,n}$ yields

$$\frac{1}{\lambda_n}(v_{i,n} - p_{2,i,n}) + L_i x_n - D_i^{-1} v_{i,n} - r_i + \frac{\alpha_{1,n}}{\lambda_n}(v_{i,n} - v_{i,n-1}) \in B_i^{-1} p_{2,i,n}$$

and from

$$\frac{1}{\lambda_n}(p_{2,i,n} - v_{i,n+1}) + L_i p_{1,n} - L_i x_n + D_i^{-1} v_{i,n} + \frac{\alpha_{2,n}}{\lambda_n}(v_{i,n} - v_{i,n-1}) = D_i^{-1} p_{2,i,n}$$

we further obtain

$$\frac{1}{\lambda_n}(v_{i,n} - v_{i,n+1}) + L_i p_{1,n} - r_i + \frac{\alpha_{1,n} + \alpha_{2,n}}{\lambda_n}(v_{i,n} - v_{i,n-1}) \in (B_i^{-1} + D_i^{-1})p_{2,i,n}.$$

Moreover, since $L_i \bar{x} - r_i \in (B_i^{-1} + D_i^{-1})\bar{v}_i$, the monotonicity of $B_i^{-1} + D_i^{-1}, i = 1, ..., m$, yields the inequality

$$\left\langle \frac{1}{\lambda_n}(v_{i,n} - v_{i,n+1}) + L_i p_{1,n} - r_i + \frac{\alpha_{1,n} + \alpha_{2,n}}{\lambda_n}(v_{i,n} - v_{i,n-1}) - (L_i \bar{x} - r_i), p_{2,i,n} - \bar{v}_i \right\rangle \geq 0,$$

hence

$$\frac{1}{\lambda_n}\sum_{i=1}^{m}\langle v_{i,n} - v_{i,n+1}, p_{2,i,n} - \overline{v}_i\rangle + \left\langle p_{1,n} - \overline{x}, \sum_{i=1}^{m} L_i^*(p_{2,i,n} - \overline{v}_i)\right\rangle$$

$$+\frac{\alpha_{1,n} + \alpha_{2,n}}{\lambda_n}\sum_{i=1}^{m}\langle v_{i,n} - v_{i,n-1}, p_{2,i,n} - \overline{v}_i\rangle \geq 0. \tag{17}$$

Summing up the inequalities (16) and (17) we obtain for all $n \geq 1$

$$\frac{1}{\lambda_n}\langle p_{1,n} - \overline{x}, x_n - x_{n+1}\rangle + \frac{\alpha_{1,n} + \alpha_{2,n}}{\lambda_n}\langle p_{1,n} - \overline{x}, x_n - x_{n-1}\rangle$$

$$+\frac{1}{\lambda_n}\sum_{i=1}^{m}\langle v_{i,n} - v_{i,n+1}, p_{2,i,n} - \overline{v}_i\rangle + \frac{\alpha_{1,n} + \alpha_{2,n}}{\lambda_n}\sum_{i=1}^{m}\langle v_{i,n} - v_{i,n-1}, p_{2,i,n} - \overline{v}_i\rangle \tag{18}$$

$$\geq \phi_{A,C}(\|p_{1,n} - \overline{x}\|).$$

It then follows from (a), (b)(i) and the boundedness of the sequences $(\alpha_{i,n})_{n\geq1}$, $i = 1, 2$ and $(\lambda_n)_{n\geq1}$ that $\lim_{n\to+\infty}\phi_{A,C}(\|p_{1,n} - \overline{x}\|) = 0$, thus $p_{1,n} \to \overline{x}$ as $n \to +\infty$. From (a) we get that $x_n \to \overline{x}$ as $n \to +\infty$.

(b)(iii) In this case one can show that instead of (18) one has for all $n \geq 1$

$$\frac{1}{\lambda_n}\langle p_{1,n} - \overline{x}, x_n - x_{n+1}\rangle + \frac{\alpha_{1,n} + \alpha_{2,n}}{\lambda_n}\langle p_{1,n} - \overline{x}, x_n - x_{n-1}\rangle$$

$$+\frac{1}{\lambda_n}\sum_{j=1}^{m}\langle v_{j,n} - v_{j,n+1}, p_{2,j,n} - \overline{v}_j\rangle + \frac{\alpha_{1,n} + \alpha_{2,n}}{\lambda_n}\sum_{j=1}^{m}\langle v_{j,n} - v_{j,n-1}, p_{2,j,n} - \overline{v}_j\rangle \tag{19}$$

$$\geq \phi_{B_i^{-1}, D_i^{-1}}(\|p_{2,i,n} - \overline{v}_i\|).$$

where $\phi_{B_i^{-1}, D_i^{-1}} : [0, +\infty) \to [0, +\infty]$ is an increasing function that vanishes only at 0. The same arguments as in (b)(ii) provide the desired conclusion. ∎

Remark 10 The case $\alpha_1 = \alpha_2 = 0$, which enforces $\alpha_{1,n} = \alpha_{2,n} = 0$ for all $n \geq 1$, shows that error-free case of the forward-backward-forward algorithm considered in [21, Theorem 3.1] is a particular case of the iterative scheme introduced in Theorem 9. We refer to Remark 6 for a discussion on how to choose the parameters λ and σ in order to get exactly the bounds from [21, Theorem 3.1].

4 Convex optimization problems

The aim of this section is to show how the inertial forward-backward-forward primal-dual algorithm can be implemented when solving a primal-dual pair of convex optimization problems.

For a function $f : \mathcal{H} \to \overline{\mathbb{R}}$, where $\overline{\mathbb{R}} := \mathbb{R} \cup \{\pm\infty\}$ is the extended real line, we denote by $\text{dom } f = \{x \in \mathcal{H} : f(x) < +\infty\}$ its *effective domain* and say that f is *proper* if $\text{dom } f \neq \emptyset$ and $f(x) \neq -\infty$ for all $x \in \mathcal{H}$. We denote by $\Gamma(\mathcal{H})$ the family of proper, convex and lower semi-continuous extended real-valued functions defined on $\mathcal{H}$. Let $f^* : \mathcal{H} \to \overline{\mathbb{R}}$, $f^*(u) = \sup_{x\in\mathcal{H}}\{\langle u, x\rangle - f(x)\}$ for all $u \in \mathcal{H}$, be the *conjugate function* of f. The *subdifferential* of f at $x \in \mathcal{H}$, with $f(x) \in \mathbb{R}$, is the set $\partial f(x) := \{v \in \mathcal{H} : f(y) \geq f(x) + \langle v, y - x\rangle \ \forall y \in \mathcal{H}\}$.

We take by convention $\partial f(x) := \emptyset$, if $f(x) \in \{\pm\infty\}$. Notice that if $f \in \Gamma(\mathcal{H})$, then ∂f is a maximally monotone operator (see [29]) and it holds $(\partial f)^{-1} = \partial f^*$. For two proper functions $f, g : \mathcal{H} \to \overline{\mathbb{R}}$, we consider their *infimal convolution*, which is the function $f \square g : \mathcal{H} \to \overline{\mathbb{R}}$, defined by $(f \square g)(x) = \inf_{y \in \mathcal{H}} \{f(y) + g(x - y)\}$, for all $x \in \mathcal{H}$.

Let $S \subseteq \mathcal{H}$ be a nonempty set. The *indicator function* of S, $\delta_S : \mathcal{H} \to \overline{\mathbb{R}}$, is the function which takes the value 0 on S and $+\infty$ otherwise. The subdifferential of the indicator function is the *normal cone* of S, that is $N_S(x) = \{u \in \mathcal{H} : \langle u, y - x \rangle \leq 0 \ \forall y \in S\}$, if $x \in S$ and $N_S(x) = \emptyset$ for $x \notin S$.

When $f \in \Gamma(\mathcal{H})$ and $\gamma > 0$, for every $x \in \mathcal{H}$ we denote by $\operatorname{prox}_{\gamma f}(x)$ the *proximal point* of parameter γ of f at x, which is the unique optimal solution of the optimization problem

$$\inf_{y \in \mathcal{H}} \left\{ f(y) + \frac{1}{2\gamma} \|y - x\|^2 \right\}. \tag{20}$$

Notice that $J_{\gamma \partial f} = (\operatorname{Id}_{\mathcal{H}} + \gamma \partial f)^{-1} = \operatorname{prox}_{\gamma f}$, thus $\operatorname{prox}_{\gamma f} : \mathcal{H} \to \mathcal{H}$ is a single-valued operator fulfilling the extended *Moreau's decomposition formula*

$$\operatorname{prox}_{\gamma f} + \gamma \operatorname{prox}_{(1/\gamma)f^*} \circ \gamma^{-1} \operatorname{Id}_{\mathcal{H}} = \operatorname{Id}_{\mathcal{H}}. \tag{21}$$

Let us also recall that a proper function $f : \mathcal{H} \to \overline{\mathbb{R}}$ is said to be *uniformly convex*, if there exists an increasing function $\phi : [0, +\infty) \to [0, +\infty]$ which vanishes only at 0 and such that $f(tx + (1-t)y) + t(1-t)\phi(\|x - y\|) \leq tf(x) + (1-t)f(y)$ for all $x, y \in \operatorname{dom} f$ and $t \in (0, 1)$. In case this inequality holds for $\phi = (\beta/2)(\cdot)^2$, where $\beta > 0$, then f is said to be β-*strongly convex*. Let us mention that this property implies β-strong monotonicity of ∂f (see [6, Example 22.3]) (more general, if f is uniformly convex, then ∂f is uniformly monotone, see [6, Example 22.3]).

Finally, we notice that for $f = \delta_S$, where $S \subseteq \mathcal{H}$ is a nonempty convex and closed set, it holds

$$J_{\gamma N_S} = J_{N_S} = J_{\partial \delta_S} = (\operatorname{Id}_{\mathcal{H}} + N_S)^{-1} = \operatorname{prox}_{\delta_S} = P_S, \tag{22}$$

where $P_S : \mathcal{H} \to C$ denotes the *projection operator* on S (see [6, Example 23.3 and Example 23.4]).

We investigate the applicability of the algorithm introduced in Section 3 in the context of the solving of the following primal-dual pair of convex optimization problems.

Problem 11 Let $\mathcal{H}$ be a real Hilbert space, $z \in \mathcal{H}$, $f \in \Gamma(\mathcal{H})$ and $h : \mathcal{H} \to \mathbb{R}$ a convex and differentiable function with a μ-Lipschitzian gradient for $\mu > 0$. Let m be a strictly positive integer and for any $i \in \{1,...,m\}$ let $\mathcal{G}_i$ be a real Hilbert space, $r_i \in \mathcal{G}_i$, $g_i, l_i \in \Gamma(\mathcal{G}_i)$ such that l_i is ν_i^{-1}-strongly convex for $\nu_i > 0$ and $L_i : \mathcal{H} \to \mathcal{G}_i$ a nonzero linear continuous operator. Consider the convex optimization problem

$$\inf_{x \in \mathcal{H}} \left\{ f(x) + \sum_{i=1}^{m} (g_i \square l_i)(L_i x - r_i) + h(x) - \langle x, z \rangle \right\} \tag{23}$$

and its *Fenchel-type dual* problem

$$\sup_{v_i \in \mathcal{G}_i, \, i=1,..,m} \left\{ -\left(f^* \square h^*\right) \left(z - \sum_{i=1}^{m} L_i^* v_i \right) - \sum_{i=1}^{m} \left(g_i^*(v_i) + l_i^*(v_i) + \langle v_i, r_i \rangle \right) \right\}. \tag{24}$$

13

Considering the maximal monotone operators

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

according to [6, Proposition 17.10, Theorem 18.15], $D_i^{-1} = \nabla l_i^*$ is a monotone and ν_i-Lipschitzian operator for $i = 1,...,m$. The monotone inclusion problem (11) reads

$$\text{find } \overline{x} \in \mathcal{H} \text{ such that } z \in \partial f(\overline{x}) + \sum_{i=1}^{m} L_i^*((\partial g_i \square \partial l_i)(L_i\overline{x} - r_i)) + \nabla h(\overline{x}), \qquad (25)$$

while the dual inclusion problem (12) reads

$$\text{find } \overline{v}_1 \in \mathcal{G}_1,..., \overline{v}_m \in \mathcal{G}_m \text{ such that } \exists x \in \mathcal{H} : \left\{ \begin{array}{l} z - \sum_{i=1}^{m} L_i^* \overline{v}_i \in \partial f(x) + \nabla h(x) \\ \overline{v}_i \in (\partial g_i \square \partial l_i)(L_i x - r_i), \ i = 1,...,m. \end{array} \right.$$
$$(26)$$

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 (25)-(26), namely,

$$z - \sum_{i=1}^{m} L_i^* \overline{v}_i \in \partial f(\overline{x}) + \nabla h(\overline{x}) \text{ and } \overline{v}_i \in (\partial g_i \square \partial l_i)(L_i\overline{x} - r_i), \ i = 1,...,m, \qquad (27)$$

then $\overline{x}$ is an optimal solution of the problem (23), $(\overline{v}_1,..., \overline{v}_m)$ is an optimal solution of (24) and the optimal objective values of the two problems coincide. Notice that (27) is nothing else than the system of optimality conditions for the primal-dual pair of convex optimization problems (23)-(24).

In case a regularity condition is fulfilled, the optimality conditions (27) are also necessary. More precisely, if the primal problem (23) has an optimal solution $\overline{x}$ and a suitable regularity condition is fulfilled, then there exists $(\overline{v}_1,..., \overline{v}_m)$ an optimal solution to (24) such that $(\overline{x}, \overline{v}_1,..., \overline{v}_m)$ satisfies the optimality conditions (27).

For the readers convenience, we discuss some regularity conditions which are suitable in this context. One of the weakest qualification conditions of interiority-type reads (see, for instance, [21, Proposition 4.3, Remark 4.4])

$$(r_1,..., r_m) \in \text{sqri} \left(\prod_{i=1}^{m} (\text{dom} \, g_i + \text{dom} \, l_i) - \{(L_1 x,..., L_m x) : x \in \text{dom} \, f\} \right). \qquad (28)$$

Here, for $\mathcal{H}$ a real Hilbert space and $S \subseteq \mathcal{H}$ a convex set, we denote by

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

its *strong quasi-relative interior*. Notice that we always have $\text{int} \, S \subseteq \text{sqri} \, S$ (in general this inclusion may be strict). If $\mathcal{H}$ is finite-dimensional, then $\text{sqri} \, S$ coincides with $\text{ri} \, S$, the relative interior of S, which is the interior of S with respect to its affine hull. The condition (28) is fulfilled, if: (i) for all $i = 1,...,m$, $\text{dom} \, g_i = \mathcal{G}_i$ or $\text{dom} \, h_i = \mathcal{G}_i$, or (ii) $\mathcal{H}$ and $\mathcal{G}_i$ are finite-dimensional spaces and there exists $x \in \text{ri} \, \text{dom} \, f$ such that $L_i x - r_i \in \text{ri} \, \text{dom} \, g_i + \text{ri} \, \text{dom} \, l_i, \ i = 1,...,m$ (see [21, Proposition 4.3]). For other regularity conditions we refer the reader to consult [6–9, 35].

The following statement is a particular instance of Theorem 9.

Theorem 12 *Suppose that the primal optimization problem* (23) *has an optimal solution and the regularity condition* (28) *is fulfilled. Chose $x_0, x_1 \in \mathcal{H}$ and $v_{i,0}, v_{i,1} \in \mathcal{G}_i$, $i = 1, ..., m$, and set*

$$(\forall n \geq 1) \begin{cases} p_{1,n} = \operatorname{prox}_{\lambda_n f}[x_n - \lambda_n(\nabla f(x_n) + \sum_{i=1}^m L_i^* v_{i,n} - z) + \alpha_{1,n}(x_n - x_{n-1})] \\ p_{2,i,n} = \operatorname{prox}_{\lambda_n g_i^*}[v_{i,n} + \lambda_n(L_i x_n - \nabla l_i^*(v_{i,n}) - r_i) + \alpha_{1,n}(v_{i,n} - v_{i,n-1})], \\ \qquad i = 1, ..., m \\ v_{i,n+1} = \lambda_n L_i(p_{1,n} - x_n) + \lambda_n(\nabla l_i^*(v_{i,n}) - \nabla l_i^*(p_{2,i,n})) + p_{2,i,n} \\ \qquad +\alpha_{2,n}(v_{i,n} - v_{i,n-1}), i = 1, ..., m \\ x_{n+1} = \lambda_n \sum_{i=1}^m L_i^*(v_{i,n} - p_{2,i,n}) + \lambda_n(\nabla h(x_n) - \nabla h(p_{1,n})) + p_{1,n} \\ \qquad +\alpha_{2,n}(x_n - x_{n-1}). \end{cases}$$

Consider $\lambda, \sigma > 0$ and $\alpha_1 \geq 0, \alpha_2 \geq 0$ such that

$$12\alpha_2^2 + 9(\alpha_1 + \alpha_2) + 4\sigma < 1 \text{ and } \lambda \leq \lambda_n \leq \frac{1}{\beta}\sqrt{\frac{1 - 12\alpha_2^2 - 9(\alpha_1 + \alpha_2) - 4\sigma}{12\alpha_2^2 + 8(\alpha_1 + \alpha_2) + 4\sigma + 2}} \; \forall n \geq 1,$$

where

$$\beta = \max\{\mu, \nu_1, ..., \nu_m\} + \sqrt{\sum_{i=1}^m \|L_i\|^2},$$

and for $i = 1, 2$ the nondecreasing sequences $(\alpha_{i,n})_{n\geq 1}$ with $\alpha_{i,1} = 0$ and $0 \leq \alpha_{i,n} \leq \alpha_i$ for all $n \geq 1$. Then the following statements are true:

(a) $\sum_{n\in\mathbb{N}} \|x_{n+1} - x_n\|^2 < +\infty$, $\sum_{n\geq 1} \|x_n - p_{1,n}\|^2 < +\infty$ *and, for $i = 1, ..., m$,*
$\sum_{n\in\mathbb{N}} \|v_{i,n+1} - v_{i,n}\|^2 < +\infty$ *and* $\sum_{n\geq 1} \|v_{i,n} - p_{2,i,n}\|^2 < +\infty$;

(b) *There exists $(\bar{x}, \bar{v}_1, ..., \bar{v}_m) \in \mathcal{H} \times \mathcal{G}_1 \times ... \times \mathcal{G}_m$ satisfying the optimality conditions* (27), *hence $\bar{x}$ is an optimal solution of the problem* (23), $(\bar{v}_1, ..., \bar{v}_m)$ *is an optimal solution of* (24) *and the optimal objective values of the two problems coincide, such that the following hold:*

    (i) $x_n \rightharpoonup \bar{x}$, $p_{1,n} \rightharpoonup \bar{x}$ *and, for $i = 1, ..., m$, $v_{i,n} \rightharpoonup \bar{v}_i$ and $p_{2,i,n} \rightharpoonup \bar{v}_i$ as $n \to +\infty$;*

    (ii) *If $f + h$ is uniformly convex, then $x_n \to \bar{x}$ and $p_{1,n} \to \bar{x}$ as $n \to +\infty$;*

    (iii) *If $g_i^* + l_i^*$ is uniformly convex for some $i \in \{1, ..., m\}$, then $v_{i,n} \to \bar{v}_i$ and $p_{2,i,n} \to \bar{v}_i$ as $n \to +\infty$.*

Remark 13 Suppose that the primal optimization problem (23) is feasible, which means that its optimal objective value is not identical $+\infty$. The existence of optimal solutions of (23) is guaranteed if for instance, $f + h + \langle \cdot, -z \rangle$ is coercive (that is $\lim_{\|x\|\to\infty}(f + h + \langle \cdot, -z \rangle)(x) = +\infty$) and for all $i = 1., , ., ., m$, g_i is bounded from below. Indeed, under these circumstances, the objective function of (23) is coercive (one can use [6, Corollary 11.16 and Proposition 12.14] to show that $g_i \square l_i$ is bounded from below and $g_i \square l_i \in \Gamma(\mathcal{G}_i)$ for $i = 1, ..., m$) and the statement follows via [6, Corollary 11.15]. On the other hand, when $f + h$ is strongly convex, then the objective function of (23) is strongly convex, too, thus (23) has a unique optimal solution (see [6, Corollary 11.16]).

Remark 14 Let us mention that for $i \in \{1, ..., m\}$ the function $g_i^* + l_i^*$ is uniformly convex, if $g_i^* + l_i^*$ is δ_i-strongly convex for $\delta_i > 0$. This is the case, for example, when g_i^* (or l_i^*) is δ_i-strongly convex or when g_i^* is α_i-strongly convex and l_i^* is β_i-strongly convex, where $\alpha_i, \beta_i > 0$ are such that $\alpha_i + \beta_i \geq \delta_i$. Let us also notice that, according to [6, Theorem 18.15], g_i^* is α_i-strongly convex if and only if g_i is Fréchet-differentiable and ∇g_i is α_i^{-1}-Lipschitzian.

5 Numerical experiments in image deblurring and denoising

The aim of this section is to illustrate the theoretical results obtained in the previous section in the context of treating a a problem occurring in imaging. For the application discussed in this section the images have been normalized, in order to make their pixels range in the closed interval from 0 to 1.

The numerical experiment that we consider addresses an ill-conditioned linear inverse problem which arises in image deblurring and denoising. For a given matrix $A \in \mathbb{R}^{n \times n}$ describing a blur operator and a given vector $b \in \mathbb{R}^n$ representing the blurred and noisy image, the task is to estimate the unknown original image $\overline{x} \in \mathbb{R}^n$ fulfilling

$$A\overline{x} = b.$$

To this end we solve the following regularized convex minimization problem

$$\inf_{x \in [0,1]^n} \{\|Ax - b\|_1 + \delta TV_{\text{iso}}(x)\}, \tag{29}$$

where $\delta > 0$ denotes a regularization parameter and $TV_{\text{iso}} : \mathbb{R}^n \to \mathbb{R}$ is the discrete isotropic total variation functional. In this context, $x \in \mathbb{R}^n$ represents the vectorized image $X \in \mathbb{R}^{M \times N}$, where $n = M \cdot N$ and $x_{i,j}$ denotes the normalized value of the pixel located in the i-th row and the j-th column, for $i = 1, \ldots, M$ and $j = 1, \ldots, N$.

The *isotropic total variation* $TV_{\text{iso}} : \mathbb{R}^n \to \mathbb{R}$ is defined by

$$TV_{\text{iso}}(x) = \sum_{i=1}^{M-1} \sum_{j=1}^{N-1} \sqrt{(x_{i+1,j} - x_{i,j})^2 + (x_{i,j+1} - x_{i,j})^2}$$
$$+ \sum_{i=1}^{M-1} |x_{i+1,N} - x_{i,N}| + \sum_{j=1}^{N-1} |x_{M,j+1} - x_{M,j}|.$$

We show first that the optimization problem (29) can be written in the framework of Problem 11.

We denote $\mathcal{Y} = \mathbb{R}^n \times \mathbb{R}^n$ and define the linear operator $L : \mathbb{R}^n \to \mathcal{Y}$, $x_{i,j} \mapsto (L_1 x_{i,j}, L_2 x_{i,j})$, where

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

The operator L represents a discretization of the gradient using reflexive (Neumann) boundary conditions and standard finite differences and fulfills $\|L\|^2 \leq 8$. For the formula for its adjoint operator $L^* : \mathcal{Y} \to \mathbb{R}^n$ we refer to [18].

For (y, z), $(p, q) \in \mathcal{Y}$, we introduce the inner product

$$\langle (y, z), (p, q) \rangle = \sum_{i=1}^{M} \sum_{j=1}^{N} y_{i,j} p_{i,j} + z_{i,j} q_{i,j}$$

and define $\|(y, z)\|_{\times} = \sum_{i=1}^{M} \sum_{j=1}^{N} \sqrt{y_{i,j}^2 + z_{i,j}^2}$. One can check that $\| \cdot \|_{\times}$ is a norm on $\mathcal{Y}$ and that for every $x \in \mathbb{R}^n$ it holds $TV_{\text{iso}}(x) = \|Lx\|_{\times}$. The conjugate function $(\| \cdot \|_{\times})^* : \mathcal{Y} \to \overline{\mathbb{R}}$ of $\| \cdot \|_{\times}$ is for every $(p, q) \in \mathcal{Y}$ given by

$$(\| \cdot \|_{\times})^*(p, q) = \begin{cases} 0, & \text{if } \|(p, q)\|_{\times*} \leq 1 \\ +\infty, & \text{otherwise} \end{cases},$$

where

$$\|(p, q)\|_{\times*} = \sup_{\|(y,z)\|_{\times} \leq 1} \langle (p, q), (y, z) \rangle = \max_{\substack{1 \leq i \leq M \\ 1 \leq j \leq N}} \sqrt{p_{i,j}^2 + q_{i,j}^2}.$$

Therefore, the optimization problem (29) can be written in the form of

$$\inf_{x \in \mathbb{R}^n} \{ f(x) + g_1(Ax) + g_2(Lx) \}, \tag{30}$$

where $f : \mathbb{R}^n \to \overline{\mathbb{R}}$, $f(x) = \delta_{[0,1]^n}(x)$, $g_1 : \mathbb{R}^n \to \mathbb{R}$, $g_1(y) = \|y - b\|_1$, $g_2 : \mathcal{Y} \to \mathbb{R}$, $g_2(y, z) = \lambda \|(y, z)\|_{\times}$ (notice that the functions l_i are taken to be $\delta_{\{0\}}$ for $i = 1, 2$ and $h \equiv 0$). For every $p \in \mathbb{R}^n$, it holds $g_1^*(p) = \delta_{[-1,1]^n}(p) + p^T b$, while for every $(p, q) \in \mathcal{Y}$, we have $g_2^*(p, q) = \delta_S(p, q)$, with $S = \{(p, q) \in \mathcal{Y} : \|(p, q)\|_{\times*} \leq \lambda\}$.

Figure 1: The figure shows the original 256×256 cameraman test image, the blurred and noisy image and the reconstructed image after 1500 iterations.

We solved problem (30) with both the inertial forward-backward-forward primal-dual algorithm stated in Theorem 12 and its non-inertial version from [21]. When implementing these we used the following formulas for the occurring proximal mappings

$$\text{prox}_{\gamma f}(x) = P_{[0,1]^n}(x) \ \forall x \in \mathbb{R}^n$$
$$\text{prox}_{\gamma g_1^*}(p) = P_{[-1,1]^n}(p - \gamma b) \ \forall p \in \mathbb{R}^n$$
$$\text{prox}_{\gamma g_2^*}(p, q) = P_S(p, q) \ \forall (p, q) \in \mathcal{Y},$$

where $\gamma > 0$ and the projection operator $P_S : \mathcal{Y} \to S$ is defined as (see [14])

$$(p_{i,j}, q_{i,j}) \mapsto \lambda \frac{(p_{i,j}, q_{i,j})}{\max \left\{ \lambda, \sqrt{p_{i,j}^2 + q_{i,j}^2} \right\}}, \ 1 \leq i \leq M, \ 1 \leq j \leq N.$$

For the numerical experiments we considered from the 256×256 cameraman test image and constructed a blurred and noisy image by making first use of a Gaussian blur operator of size 9×9 and standard deviation 4. Afterwards, we added a zero-mean white Gaussian noise with standard deviation 10^{-3}. We considered as regularization parameter $\delta = 0.001$, while the other parameter needed in the inertial forward-backward-forward primal-dual algorithm (IFBF) has been chosen as $\sigma = 0.01$, $\alpha_1 = \alpha_{1,n} = \alpha_2 = \alpha_{2,n} = 0.01$ for $n \geq 2$ and (noticing that $\beta = 3$)

$$\lambda_n = \frac{1}{3}\sqrt{\frac{1 - 12\alpha_2^2 - 9(\alpha_1 + \alpha_2) - 4\sigma}{12\alpha_2^2 + 8(\alpha_1 + \alpha_2) + 4\sigma + 2}} \ \forall n \geq 1.$$

Figure 1 shows the original cameraman test image, the blurred and noisy image and the image reconstructed by using the inertial forward-backward-forward primal-dual algorithm after 1500 iterations.

We compared the IFBF algorithm with its non-inertial version (FBF) from [21] in terms of the improvement in signal-to-noise ratio (ISNR), which is defined as

$$\text{ISNR}(n) = 10 \log_{10}\left(\frac{\|x - b\|^2}{\|x - x_n\|^2}\right),$$

where x, b and x_n denote the original, observed and estimated image at iteration n, respectively, and represents a measure for the quality of the reconstructed images. The evolution of the ISNR values of the two algorithms is shown in Figure 2 and it emphasizes that the standard version of the FBF algorithm is slightly outperformed by its inertial version after a certain number of iterations.
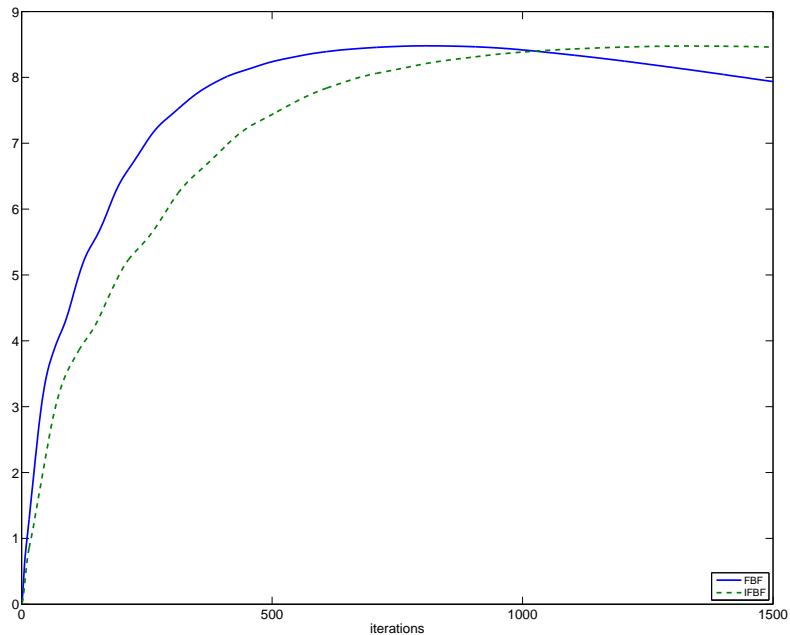

Figure 2: The figure shows the evolution of the ISNR values of the FBF and IFBF algorithms.

18

References

[1] F. Alvarez, *On the minimizing property of a second order dissipative system in Hilbert spaces*, SIAM Journal on Control and Optimization 38(4), 1102–1119, 2000

[2] F. Alvarez, *Weak convergence of a relaxed and inertial hybrid projection-proximal point algorithm for maximal monotone operators in Hilbert space*, SIAM Journal on Optimization 14(3), 773–782, 2004

[3] F. Alvarez, H. Attouch, *An inertial proximal method for maximal monotone operators via discretization of a nonlinear oscillator with damping*, Set-Valued Analysis 9, 3–11, 2001

[4] H. Attouch, L.M. Briceño-Arias, P.L. Combettes, *A parallel splitting method for coupled monotone inclusions*, SIAM Journal on Control and Optimization 48(5), 3246–3270, 2010

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

[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, *Conjugate Duality in Convex Optimization*, Lecture Notes in Economics and Mathematical Systems, Vol. 637, Springer, Berlin Heidelberg, 2010

[9] R.I. Boţ, E.R. Csetnek, *Regularity conditions via generalized interiority notions in convex optimization: new achievements and their relation to some classical statements*, Optimization 61(1), 35–65, 2012

[10] R.I. Boţ, E.R. Csetnek, A. Heinrich, *A primal-dual splitting algorithm for finding zeros of sums of maximally monotone operators*, SIAM Journal on Optimization, 23(4), 2011–2036, 2013

[11] 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

[12] R.I. Boţ, E.R. Csetnek, C. Hendrich, *Inertial Douglas-Rachford splitting for monotone inclusion problems*, Applied Mathematics and Computation 256, 472–487, 2015

[13] R.I. Boţ, E.R. Csetnek, S. László, *An inertial forward-backward algorithm for the minimization of the sum of two nonconvex functions*, arXiv:1410.0641

[14] R.I. Boţ, C. Hendrich, *Convergence analysis for a primal-dual monotone + skew splitting algorithm with applications to total variation minimization*, Journal of Mathematical Imaging and Vision 49(3), 551–568, 2014

[15] 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

[16] 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

[17] A. Cabot, P. Frankel, *Asymptotics for some proximal-like method involving inertia and memory aspects*, Set-Valued and Variational Analysis 19, 59–74, 2011

[18] A. Chambolle, *An algorithm for total variation minimization and applications*, Journal of Mathematical Imaging and Vision 20(1–2), 89–97, 2004

[19] 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

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

[21] 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

[22] 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

[23] I. Ekeland, R. Temam, *Convex Analysis and Variational Problems*, North-Holland Publishing Company, Amsterdam, 1976

[24] P.-E. Maingé, *Convergence theorems for inertial KM-type algorithms*, Journal of Computational and Applied Mathematics 219, 223–236, 2008

[25] P.-E. Maingé, A. Moudafi, *Convergence of new inertial proximal methods for dc programming*, SIAM Journal on Optimization 19(1), 397–413, 2008

[26] A. Moudafi, M. Oliny, *Convergence of a splitting inertial proximal method for monotone operators*, Journal of Computational and Applied Mathematics 155, 447–454, 2003

[27] P. Ochs, Y. Chen, T. Brox, T. Pock, *iPiano: Inertial proximal algorithm for nonconvex optimization*, SIAM Journal of Imaging Sciences 7(2), 1388–1419, 2014

[28] J.-C. Pesquet, N. Pustelnik, *A parallel inertial proximal optimization method*, Pacific Journal of Optimization 8(2), 273–305, 2012

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

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

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

[32] P. Tseng. *Applications of a splitting algorithm to decomposition in convex programming and variational inequalities*, SIAM Journal on Control and Optimization 29(1), 119–138, 1991

[33] P. Tseng, *A modified forward-backward splitting method for maximal monotone mappings*, SIAM Journal on Control and Optimization 38(2), 431–446, 2000

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

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