

Convergence analysis for a primal-dual monotone + skew splitting algorithm with applications to total variation minimization

Radu Ioan Boţ^{*} Christopher Hendrich[†]

November 7, 2012

Abstract. In this paper we investigate the convergence behavior of a primal-dual splitting method for solving monotone inclusions involving mixtures of composite, Lipschitzian and parallel sum type operators proposed by Combettes and Pesquet in [7]. Firstly, in the particular case of convex minimization problems, we derive convergence rates for the sequence of objective function values by making use of conjugate duality techniques. Secondly, we propose for the general monotone inclusion problem two new schemes which accelerate the sequences of primal and/or dual iterates, provided strong monotonicity assumptions for some of the involved operators are fulfilled. Finally, we apply the theoretical achievements in the context of different types of image restoration problems solved via total variation regularization.

Keywords. splitting method, Fenchel duality, convergence statements, image processing

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

1 Introduction and preliminaries

The last few years have shown a rising interest in solving structured nondifferentiable convex optimization problems within the framework of the theory of conjugate functions. Applications in fields like signal and image processing, location theory and supervised machine learning motivate these efforts.

In this article we investigate and improve the convergence behavior of the primal-dual monotone + skew splitting method for solving monotone inclusions which was proposed by Combettes and Pesquet in [7], itself being an extension of the algorithmic scheme from [4] obtained by allowing also Lipschitzian monotone operators and parallel sums in the problem formulation. In the mentioned works, by means of a product space approach, the problem is reduced to the one of finding the zeros of the sum of a Lipschitzian monotone operator with a maximally monotone operator. The latter is solved

^{*}Faculty of Mathematics, Chemnitz University of Technology, D-09107 Chemnitz, Germany, e-mail: radu.bot@mathematik.tu-chemnitz.de. Research partially supported by DFG (German Research Foundation), project BO 2516/4-1.

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

by using an error-tolerant version of Tseng's algorithm which has forward-backward-forward characteristics and allows to access the monotone Lipschitzian operators via explicit forward steps, while set-valued maximally monotone operators are processed via their resolvents. A notable advantage of this method is given by both its highly parallelizable character, most of its steps could be executed independently, and by the fact that allows to process maximal monotone operators and linear bounded operators separately, whenever they occur in the form of precompositions in the problem formulation.

Before coming to the description of the problem formulation and of the algorithm from [7], we introduce some preliminary notions and results which are needed throughout the paper.

We are considering the real Hilbert spaces \mathcal{H} and \mathcal{G}_i , $i = 1, \dots, m$, endowed with the *inner product* $\langle \cdot, \cdot \rangle$ and associated *norm* $\|\cdot\| = \sqrt{\langle \cdot, \cdot \rangle}$, for which we use the same notation, respectively, as there is no risk of confusion. The symbols \rightharpoonup and \rightarrow denote weak and strong convergence, respectively. By \mathbb{R}_{++} we denote the set of strictly positive real numbers, while the *indicator function* of a set $C \subseteq \mathcal{H}$ is $\delta_C : \mathcal{H} \rightarrow \overline{\mathbb{R}} := \mathbb{R} \cup \{\pm\infty\}$, defined by $\delta_C(x) = 0$ for $x \in C$ and $\delta_C(x) = +\infty$, otherwise. For a function $f : \mathcal{H} \rightarrow \overline{\mathbb{R}}$ we denote by $\text{dom } f := \{x \in \mathcal{H} : f(x) < +\infty\}$ its *effective domain* and call f *proper* if $\text{dom } f \neq \emptyset$ and $f(x) > -\infty$ for all $x \in \mathcal{H}$. Let be

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

The *conjugate function* of f is $f^* : \mathcal{H} \rightarrow \overline{\mathbb{R}}$, $f^*(p) = \sup \{\langle p, x \rangle - f(x) : x \in \mathcal{H}\}$ for all $p \in \mathcal{H}$ and, if $f \in \Gamma(\mathcal{H})$, then $f^* \in \Gamma(\mathcal{H})$, as well. The (*convex*) *subdifferential* of $f : \mathcal{H} \rightarrow \overline{\mathbb{R}}$ at $x \in \mathcal{H}$ is the set $\partial f(x) = \{p \in \mathcal{H} : f(y) - f(x) \geq \langle p, y - x \rangle \ \forall y \in \mathcal{H}\}$, if $f(x) \in \mathbb{R}$, and is taken to be the empty set, otherwise. For a linear continuous operator $L_i : \mathcal{H} \rightarrow \mathcal{G}_i$, the operator $L_i^* : \mathcal{G}_i \rightarrow \mathcal{H}$, defined via $\langle L_i x, y \rangle = \langle x, L_i^* y \rangle$ for all $x \in \mathcal{H}$ and all $y \in \mathcal{G}_i$, denotes its *adjoint operator*, for $i \in \{1, \dots, m\}$.

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

Let $M : \mathcal{H} \rightarrow 2^{\mathcal{H}}$ be a set-valued operator. We denote by $\text{gra } M = \{(x, u) \in \mathcal{H} \times \mathcal{H} : u \in Mx\}$ its *graph* and by $\text{ran } M = \{u \in \mathcal{H} : \exists x \in \mathcal{H}, u \in Mx\}$ its *range*. The *inverse operator* of M is defined as $M^{-1} : \mathcal{H} \rightarrow 2^{\mathcal{H}}$, $M^{-1}(u) = \{x \in \mathcal{H} : u \in Mx\}$. The operator M is called *monotone* if $\langle x - y, u - v \rangle \geq 0$ for all $(x, u), (y, v) \in \text{gra } M$ and it is called *maximally monotone* if there exists no monotone operator $M' : \mathcal{H} \rightarrow 2^{\mathcal{H}}$ such that $\text{gra } M'$ properly contains $\text{gra } M$. The operator M is called ρ -strongly monotone, for $\rho \in \mathbb{R}_{++}$, if $M - \rho \text{Id}$ is monotone, i. e. $\langle x - y, u - v \rangle \geq \rho \|x - y\|^2$ for all $(x, u), (y, v) \in \text{gra } M$, where Id denotes the identity on \mathcal{H} . The operator $M : \mathcal{H} \rightarrow \mathcal{H}$ is called ν -Lipschitzian for $\nu \in \mathbb{R}_{++}$ if it is single-valued and it fulfills $\|Mx - My\| \leq \nu \|x - y\|$ for all $x, y \in \mathcal{H}$.

The resolvent of a set-valued operator $M : \mathcal{H} \rightarrow 2^{\mathcal{H}}$ is $J_M : \mathcal{H} \rightarrow 2^{\mathcal{H}}$, $J_M = (\text{Id} + M)^{-1}$. When M is maximally monotone, the resolvent is a single-valued, 1-Lipschitzian and maximal monotone operator. Moreover, when $f \in \Gamma(\mathcal{H})$ and $\gamma \in \mathbb{R}_{++}$, $\partial(\gamma f)$ is maximally monotone (cf. [9, Theorem 3.2.8]) and it holds $J_{\gamma \partial f} = (\text{Id} + \gamma \partial f)^{-1} = \text{Prox}_{\gamma f}$. Here, $\text{Prox}_{\gamma f}(x)$ denotes the *proximal point* of parameter γ of

f at $x \in \mathcal{H}$ and it represents the unique optimal solution of the optimization problem

$$\inf_{y \in \mathcal{H}} \left\{ f(y) + \frac{1}{2\gamma} \|y - x\|^2 \right\}. \quad (1.1)$$

For a nonempty, convex and closed set $C \subseteq \mathcal{H}$ and $\gamma \in \mathbb{R}_{++}$ we have $\text{Prox}_{\gamma\delta_C} = \mathcal{P}_C$, where $\mathcal{P}_C : \mathcal{H} \rightarrow C$, $\mathcal{P}_C(x) = \arg \min_{z \in C} \|x - z\|$, denotes the *projection operator* on C .

Finally, the *parallel sum* of two set-valued operators $M_1, M_2 : \mathcal{H} \rightarrow 2^{\mathcal{H}}$ is defined as

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

We can formulate now the monotone inclusion problem which we investigate in this paper (see [7]).

Problem 1.1. Consider the real Hilbert spaces \mathcal{H} and $\mathcal{G}_i, i = 1, \dots, m$, $A : \mathcal{H} \rightarrow 2^{\mathcal{H}}$ a maximally monotone operator and $C : \mathcal{H} \rightarrow \mathcal{H}$ a monotone and μ -Lipschitzian operator for some $\mu \in \mathbb{R}_{++}$. Furthermore, let $z \in \mathcal{H}$ and for every $i \in \{1, \dots, m\}$, let $r_i \in \mathcal{G}_i$, let $B_i : \mathcal{G}_i \rightarrow 2^{\mathcal{G}_i}$ be maximally monotone operators, let $D_i : \mathcal{G}_i \rightarrow 2^{\mathcal{G}_i}$ be monotone operators such that D_i^{-1} is ν_i -Lipschitzian for some $\nu_i \in \mathbb{R}_{++}$, and let $L_i : \mathcal{H} \rightarrow \mathcal{G}_i$ be a nonzero linear continuous operator. The problem is to solve the primal inclusion

$$\text{find } \bar{x} \in \mathcal{H} \text{ such that } z \in A\bar{x} + \sum_{i=1}^m L_i^* ((B_i \square D_i)(L_i \bar{x} - r_i)) + C\bar{x}, \quad (1.2)$$

together with the dual inclusion

$$\text{find } \bar{v}_1 \in \mathcal{G}_1, \dots, \bar{v}_m \in \mathcal{G}_m \text{ such that } (\exists x \in \mathcal{H}) \begin{cases} z - \sum_{i=1}^m L_i^* \bar{v}_i \in Ax + Cx \\ \bar{v}_i \in (B_i \square D_i)(L_i x - r_i), i = 1, \dots, m. \end{cases} \quad (1.3)$$

Throughout this paper we denote by $\mathcal{G} := \mathcal{G}_1 \times \dots \times \mathcal{G}_m$ the Hilbert space equipped with the inner product

$$\langle (p_1, \dots, p_m), (q_1, \dots, q_m) \rangle = \sum_{i=1}^m \langle p_i, q_i \rangle \quad \forall (p_1, \dots, p_m) \quad \forall (q_1, \dots, q_m) \in \mathcal{G}$$

and the associated norm $\|(p_1, \dots, p_m)\| = \sqrt{\sum_{i=1}^m \|p_i\|^2}$ for all $(p_1, \dots, p_m) \in \mathcal{G}$. We introduce also the nonzero linear continuous operator $\mathbf{L} : \mathcal{H} \rightarrow \mathcal{G}$, $\mathbf{L}x = (L_1 x, \dots, L_m x)$, its adjoint being $\mathbf{L}^* : \mathcal{G} \rightarrow \mathcal{H}$, $\mathbf{L}^* v = \sum_{i=1}^m L_i^* v_i$.

We say that $(\bar{x}, \bar{v}_1, \dots, \bar{v}_m) \in \mathcal{H} \times \mathcal{G}$ is a primal-dual solution to Problem 1.1, if

$$z - \sum_{i=1}^m L_i^* \bar{v}_i \in A\bar{x} + C\bar{x} \text{ and } \bar{v}_i \in (B_i \square D_i)(L_i \bar{x} - r_i), i = 1, \dots, m. \quad (1.4)$$

If $(\bar{x}, \bar{v}_1, \dots, \bar{v}_m) \in \mathcal{H} \times \mathcal{G}$ is a primal-dual solution to Problem 1.1, then \bar{x} is a solution to (1.2) and $(\bar{v}_1, \dots, \bar{v}_m)$ is a solution to (1.3). Notice also that

$$\begin{aligned} \bar{x} \text{ solves (1.2)} &\Leftrightarrow z - \sum_{i=1}^m L_i^* (B_i \square D_i)(L_i \bar{x} - r_i) \in A\bar{x} + C\bar{x} \Leftrightarrow \\ &\exists \bar{v}_1 \in \mathcal{G}_1, \dots, \bar{v}_m \in \mathcal{G}_m \text{ such that } \begin{cases} z - \sum_{i=1}^m L_i^* \bar{v}_i \in A\bar{x} + C\bar{x}, \\ \bar{v}_i \in (B_i \square D_i)(L_i \bar{x} - r_i), i = 1, \dots, m. \end{cases} \end{aligned}$$

Thus, if \bar{x} is a solution to (1.2), then there exists $(\bar{v}_1, \dots, \bar{v}_m) \in \mathcal{G}$ such that $(\bar{x}, \bar{v}_1, \dots, \bar{v}_m)$ is a primal-dual solution to Problem 1.1 and if $(\bar{v}_1, \dots, \bar{v}_m)$ is a solution to (1.3), then there exists $\bar{x} \in \mathcal{H}$ such that $(\bar{x}, \bar{v}_1, \dots, \bar{v}_m)$ is a primal-dual solution to Problem 1.1.

The next result provides the error-free variant of the primal-dual algorithm in [7] and the corresponding convergence statements, as given in [7, Theorem 3.1].

Theorem 1.1. *For Problem 1.1 suppose that*

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

Let $x_0 \in \mathcal{H}$ and $(v_{1,0}, \dots, v_{m,0}) \in \mathcal{G}$, set

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

choose $\varepsilon \in (0, \frac{1}{\beta+1})$ and $(\gamma_n)_{n \geq 0}$ a sequence in $[\varepsilon, \frac{1-\varepsilon}{\beta}]$ and set

$$(\forall n \geq 0) \begin{cases} p_{1,n} = J_{\gamma_n A} (x_n - \gamma_n (C x_n + \sum_{i=1}^m L_i^* v_{i,n} - z)) \\ \text{For } i = 1, \dots, m \\ \quad \left[\begin{array}{l} p_{2,i,n} = J_{\gamma_n B_i^{-1}} (v_{i,n} + \gamma_n (L_i x_n - D_i^{-1} v_{i,n} - r_i)) \\ v_{i,n+1} = \gamma_n L_i (p_{1,n} - x_n) + \gamma_n (D_i^{-1} v_{i,n} - D_i^{-1} p_{2,i,n}) + p_{2,i,n} \\ x_{n+1} = \gamma_n \sum_{i=1}^m L_i^* (v_{i,n} - p_{2,i,n}) + \gamma_n (C x_n - C p_{1,n}) + p_{1,n}. \end{array} \right. \end{cases} \quad (1.5)$$

Then the following statements are true:

- (i) $\sum_{n \in \mathbb{N}} \|x_n - p_{1,n}\|^2 < +\infty$ and $\forall i \in \{1, \dots, m\} \sum_{n \in \mathbb{N}} \|v_{i,n} - p_{2,i,n}\|^2 < +\infty$.
- (ii) There exists a primal-dual solution $(\bar{x}, \bar{v}_1, \dots, \bar{v}_m) \in \mathcal{H} \times \mathcal{G}$ to Problem 1.1 such that the following hold:

- (a) $x_n \rightharpoonup \bar{x}$ and $p_{1,n} \rightharpoonup \bar{x}$.
- (b) $(\forall i \in \{1, \dots, m\}) v_{i,n} \rightharpoonup \bar{v}_i$ and $p_{2,i,n} \rightharpoonup \bar{v}_i$.

In this paper we consider first Problem 1.1 in its particular formulation as a primal-dual pair of convex minimization problems, approach which relies on the fact that the subdifferential of a proper, convex and lower semicontinuous function is maximally monotone, and show that the convergence rate of the sequence of objective function values on the iterates generated by (1.5) is of $\mathcal{O}(\frac{1}{n})$, where $n \in \mathbb{N}$ is the number of passed iterations. Further, in Section 3, we provide for the general monotone inclusion problem, as given in Problem 1.1, two new acceleration schemes which generate under strong monotonicity assumptions sequences of primal and/or dual iterates converge with improved convergence properties. The feasibility of the proposed methods is explicitly shown in Section 4 by means of numerical experiments in the context of solving image denoising, image deblurring and image inpainting problems via total variation regularization.

One of the iterative schemes to which we compare our algorithms is a primal-dual splitting method for solving highly structured monotone inclusions, as well, and it was provided by Vũ in [8]. Here, instead of monotone Lipschitzian operators, cocoercive operators were used and, consequently, instead of Tseng's splitting, the forward-backward splitting method has been used. The primal-dual method due to Chambolle and Pock described in [6, Algorithm 1] is a particular instance of Vũ's algorithm.

2 Convex minimization problems

The aim of this section is to provide a rate of convergence for the sequence of the values of the objective function at the iterates generated by the algorithm (1.5) when solving a convex minimization problem and its conjugate dual. The primal-dual pair under investigation is described in the following.

Problem 2.1. Consider the real Hilbert spaces \mathcal{H} and $\mathcal{G}_i, i = 1, \dots, m$, $f \in \Gamma(\mathcal{H})$ and $h : \mathcal{H} \rightarrow \mathbb{R}$ a convex and differentiable function with μ -Lipschitzian gradient for some $\mu \in \mathbb{R}_{++}$. Furthermore, let $z \in \mathcal{H}$ and for every $i \in \{1, \dots, m\}$, let $r_i \in \mathcal{G}_i$, $g_i, l_i \in \Gamma(\mathcal{G}_i)$ such that l_i is ν_i^{-1} -strongly convex for some $\nu_i \in \mathbb{R}_{++}$, and let $L_i : \mathcal{H} \rightarrow \mathcal{G}_i$ be a nonzero linear continuous operator. We consider the convex minimization problem

$$(P) \quad \inf_{x \in \mathcal{H}} \left\{ f(x) + \sum_{i=1}^m (g_i \square l_i)(L_i x - r_i) + h(x) - \langle x, z \rangle \right\} \quad (2.1)$$

and its dual problem

$$(D) \quad \sup_{(v_1, \dots, v_m) \in \mathcal{G}_1 \times \dots \times \mathcal{G}_m} \left\{ - (f^* \square h^*) \left(z - \sum_{i=1}^m L_i^* v_i \right) - \sum_{i=1}^m (g_i^*(v_i) + l_i^*(v_i) + \langle v_i, r_i \rangle) \right\}. \quad (2.2)$$

In order to investigate the primal-dual pair (2.1)-(2.2) in the context of Problem 2.1, one has to take

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

Then A and $B_i, i = 1, \dots, m$ are maximal monotone, C is monotone, by [1, Proposition 17.10], and $D_i^{-1} = \nabla l_i^*$ is monotone and ν_i -Lipschitz continuous for $i = 1, \dots, m$, according to [1, Proposition 17.10, Theorem 18.15 and Corollary 16.24]. One can easily see that (see, for instance, [7, Theorem 4.2]) whenever $(\bar{x}, \bar{v}_1, \dots, \bar{v}_m) \in \mathcal{H} \times \mathcal{G}$ is a primal-dual solution to Problem 1.1, with the above choice of the involved operators, \bar{x} is an optimal solution to (P), $(\bar{v}_1, \dots, \bar{v}_m)$ is an optimal solution to (D) and for (P)-(D) strong duality holds, thus the optimal objective values of the two problems coincide.

The primal-dual pair in Problem 2.1 captures various different types of optimization problems. One such particular instance is formulated as follows and we refer for more examples to [7].

Example 2.1. In Problem 2.1 take $z = 0$, let $l_i : \mathcal{G}_i \rightarrow \overline{\mathbb{R}}$, $l_i = \delta_{\{0\}}$ and $r_i = 0$ for $i = 1, \dots, m$, and set $h : \mathcal{H} \rightarrow \mathbb{R}$, $h(x) = 0$ for all $x \in \mathcal{H}$. Then (2.1) reduces to

$$(P) \quad \inf_{x \in \mathcal{H}} \left\{ f(x) + \sum_{i=1}^m g_i(L_i x) \right\},$$

while the dual problem (2.2) becomes

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

In order to simplify the upcoming formulations and calculations we introduce the following more compact notations. With respect to Problem 2.1, let $F : \mathcal{H} \rightarrow \overline{\mathbb{R}}$, $F(x) = f(x) + h(x) - \langle x, z \rangle$. Then $\text{dom } F = \text{dom } f$ and its conjugate $F^* : \mathcal{H} \rightarrow \overline{\mathbb{R}}$ is given by $F^*(p) = (f + h)^*(z + p) = (f^* \square h^*)(z + p)$, since $\text{dom } h = \mathcal{H}$. Further, we set

$$\mathbf{v} = (v_1, \dots, v_m), \quad \bar{\mathbf{v}} = (\bar{v}_1, \dots, \bar{v}_m), \quad \mathbf{p}_{2,n} = (p_{2,1,n}, \dots, p_{2,m,n}), \quad \mathbf{r} = (r_1, \dots, r_m).$$

We define the function $G : \mathcal{G} \rightarrow \overline{\mathbb{R}}$, $G(\mathbf{y}) = \sum_{i=1}^m (g_i \square l_i)(y_i)$ and observe that its conjugate $G^* : \mathcal{G} \rightarrow \overline{\mathbb{R}}$ is given by $G^*(\mathbf{v}) = \sum_{i=1}^m (g_i \square l_i)^*(v_i) = \sum_{i=1}^m (g_i^* + l_i^*)(v_i)$. Notice that, as $l_i^*, i = 1, \dots, m$, has full domain (cf. [1, Theorem 18.15]), we get

$$\text{dom } G^* = (\text{dom } g_1^* \cap \text{dom } l_1^*) \times \dots \times (\text{dom } g_m^* \cap \text{dom } l_m^*) = \text{dom } g_1^* \times \dots \times \text{dom } g_m^*, \quad (2.3)$$

The primal and the dual optimization problems given in Problem 2.1 can be equivalently represented as

$$(P) \quad \inf_{x \in \mathcal{H}} \{F(x) + G(\mathbf{L}x - \mathbf{r})\},$$

and, respectively,

$$(D) \quad \sup_{\mathbf{v} \in \mathcal{G}} \{-F^*(-\mathbf{L}^*\mathbf{v}) - G^*(\mathbf{v}) - \langle \mathbf{v}, \mathbf{r} \rangle\}.$$

Then $\bar{x} \in \mathcal{H}$ solves (P), $\bar{\mathbf{v}} \in \mathcal{G}$ solves (D) and for (P)-(D) strong duality holds if and only if (cf. [2, 3])

$$-\mathbf{L}^*\bar{\mathbf{v}} \in \partial F(\bar{x}) \text{ and } \mathbf{L}\bar{x} - \mathbf{r} \in \partial G^*(\bar{\mathbf{v}}). \quad (2.4)$$

Let us mention also that for $\bar{x} \in \mathcal{H}$ and $\bar{\mathbf{v}} \in \mathcal{G}$ fulfilling (2.4) it holds

$$[\langle \mathbf{L}x - \mathbf{r}, \bar{\mathbf{v}} \rangle + F(x) - G^*(\bar{\mathbf{v}})] - [\langle \mathbf{L}\bar{x} - \mathbf{r}, \mathbf{v} \rangle + F(\bar{x}) - G^*(\mathbf{v})] \geq 0 \quad \forall x \in \mathcal{H} \quad \forall \mathbf{v} \in \mathcal{G}.$$

For given sets $B_1 \subseteq \mathcal{H}$ and $B_2 \subseteq \mathcal{G}$ we introduce the so-called *primal-dual gap function*

$$\begin{aligned} \mathcal{G}_{B_1 \times B_2}(x, \mathbf{v}) &= \sup_{\tilde{\mathbf{v}} \in B_2} \{\langle \mathbf{L}x - \mathbf{r}, \tilde{\mathbf{v}} \rangle + F(x) - G^*(\tilde{\mathbf{v}})\} \\ &\quad - \inf_{\tilde{x} \in B_1} \{\langle \mathbf{L}\tilde{x} - \mathbf{r}, \mathbf{v} \rangle + F(\tilde{x}) - G^*(\mathbf{v})\}. \end{aligned} \quad (2.5)$$

We consider the following algorithm for solving (P)-(D), which differs from the one given in Theorem 1.1 by the fact that we are asking the sequence $(\gamma_n)_{n \geq 0} \subseteq \mathbb{R}_{++}$ to be nondecreasing.

Algorithm 2.1. Let $x_0 \in \mathcal{H}$ and $(v_{1,0}, \dots, v_{m,0}) \in \mathcal{G}$, set

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

choose $\varepsilon \in \left(0, \frac{1}{\beta+1}\right)$ and $(\gamma_n)_{n \geq 0}$ a nondecreasing sequence in $\left[\varepsilon, \frac{1-\varepsilon}{\beta}\right]$ and set

$$(\forall n \geq 0) \begin{cases} p_{1,n} = \text{Prox}_{\gamma_n f}(x_n - \gamma_n(\nabla h(x_n) + \sum_{i=1}^m L_i^* v_{i,n} - z)) \\ \text{For } i = 1, \dots, m \\ \quad \begin{cases} p_{2,i,n} = \text{Prox}_{\gamma_n g_i^*}(v_{i,n} + \gamma_n(L_i x_n - \nabla l_i^*(v_{i,n}) - r_i)) \\ v_{i,n+1} = \gamma_n L_i(p_{1,n} - x_n) + \gamma_n(\nabla l_i^*(v_{i,n}) - \nabla l_i^*(p_{2,i,n})) + p_{2,i,n} \\ x_{n+1} = \gamma_n \sum_{i=1}^m L_i^*(v_{i,n} - p_{2,i,n}) + \gamma_n(\nabla h(x_n) - \nabla h(p_{1,n})) + p_{1,n}. \end{cases} \end{cases} \quad (2.6)$$

Theorem 2.1. For Problem 2.1 suppose that

$$z \in \text{ran} \left(\partial f + \sum_{i=1}^m L_i^* (\partial g_i \square \partial l_i) (L_i \cdot -r_i) + \nabla h \right).$$

Then there exists an optimal solution $\bar{x} \in \mathcal{H}$ to (P) and an optimal solution $(\bar{v}_1, \dots, \bar{v}_m) \in \mathcal{G}$ to (D), such that the following holds for the sequences generated by Algorithm 2.1:

- (a) $z - \sum_{i=1}^m L_i^* \bar{v}_i \in \partial f(\bar{x}) + \nabla h(\bar{x})$ and $L_i \bar{x} - r_i \in \partial g_i^*(\bar{v}_i) + \nabla l_i^*(\bar{v}_i) \forall i \in \{1, \dots, m\}$.
- (b) $x_n \rightarrow \bar{x}$, $p_{1,n} \rightarrow \bar{x}$ and $v_{i,n} \rightarrow \bar{v}_i$, $p_{2,i,n} \rightarrow \bar{v}_i \forall i \in \{1, \dots, m\}$.
- (c) For $n \geq 0$ it holds

$$\frac{\|x_n - \bar{x}\|^2}{2\gamma_n} + \sum_{i=1}^m \frac{\|v_{i,n} - \bar{v}_i\|^2}{2\gamma_n} \leq \frac{\|x_0 - \bar{x}\|^2}{2\gamma_0} + \sum_{i=1}^m \frac{\|v_{i,0} - \bar{v}_i\|^2}{2\gamma_0}.$$

- (d) If $B_1 \subseteq \mathcal{H}$ and $B_2 \subseteq \mathcal{G}$ are bounded, then for $x^N := \frac{1}{N} \sum_{n=0}^{N-1} p_{1,n}$ and $v_i^N := \frac{1}{N} \sum_{n=0}^{N-1} p_{2,i,n}$, $i = 1, \dots, m$, the primal-dual gap has the upper bound

$$\mathcal{G}_{B_1 \times B_2}(x^N, v_1^N, \dots, v_m^N) \leq \frac{C(B_1, B_2)}{N}, \quad (2.7)$$

where

$$C(B_1, B_2) = \sup_{(x, v_1, \dots, v_m) \in B_1 \times B_2} \left\{ \frac{\|x_0 - x\|^2}{2\gamma_0} + \sum_{i=1}^m \frac{\|v_{i,0} - v_i\|^2}{2\gamma_0} \right\}.$$

- (e) The sequence $(x^N, v_1^N, \dots, v_m^N)$ converges weakly to $(\bar{x}, \bar{v}_1, \dots, \bar{v}_m)$.

Proof. Theorem 4.2 in [7] guarantees the existence of an optimal solution $\bar{x} \in \mathcal{H}$ to (2.1) and of an optimal solution $(\bar{v}_1, \dots, \bar{v}_m) \in \mathcal{G}$ to (2.2) such that strong duality holds, $x_n \rightarrow \bar{x}$, $p_{1,n} \rightarrow \bar{x}$, as well as $v_{i,n} \rightarrow \bar{v}_i$ and $p_{2,i,n} \rightarrow \bar{v}_i$ for $i = 1, \dots, m$, when n converges to $+\infty$. Hence (a) and (b) are true. Thus, the solutions \bar{x} and $\bar{v} = (\bar{v}_1, \dots, \bar{v}_m)$ fulfill (2.4).

Regarding the sequences $(p_{1,n})_{n \geq 0}$ and $(p_{2,i,n})_{n \geq 0}$, $i = 1, \dots, m$, generated in Algorithm 2.1 we have for every $n \geq 0$

$$\begin{aligned} p_{1,n} &= (\text{Id} + \gamma_n \partial f)^{-1} (x_n - \gamma_n (\nabla h(x_n) + L^* v_n - z)) \\ &\Leftrightarrow \frac{x_n - p_{1,n}}{\gamma_n} - \nabla h(x_n) - L^* v_n + z \in \partial f(p_{1,n}) \end{aligned}$$

and, for $i = 1, \dots, m$,

$$\begin{aligned} p_{2,i,n} &= (\text{Id} + \gamma_n \partial g_i^*)^{-1} (v_{i,n} + \gamma_n (L_i x_n - \nabla l_i^*(v_{i,n}) - r_i)) \\ &\Leftrightarrow \frac{v_{i,n} - p_{2,i,n}}{\gamma_n} + L_i x_n - \nabla l_i^*(v_{i,n}) - r_i \in \partial g_i^*(p_{2,i,n}). \end{aligned}$$

In other words, it holds for every $n \geq 0$

$$f(x) \geq f(p_{1,n}) + \left\langle \frac{x_n - p_{1,n}}{\gamma_n} - \nabla h(x_n) - L^* v_n + z, x - p_{1,n} \right\rangle \quad \forall x \in \mathcal{H} \quad (2.8)$$

and, for $i = 1, \dots, m$,

$$g_i^*(v_i) \geq g_i^*(p_{2,i,n}) + \left\langle \frac{v_{i,n} - p_{2,i,n}}{\gamma_n} + L_i x_n - \nabla l_i^*(v_{i,n}) - r_i, v_i - p_{2,i,n} \right\rangle \quad \forall v_i \in \mathcal{G}_i. \quad (2.9)$$

In addition to that, using that h and l_i^* , $i = 1, \dots, m$, are convex and differentiable, it holds for every $n \geq 0$

$$h(x) \geq h(p_{1,n}) + \langle \nabla h(p_{1,n}), x - p_{1,n} \rangle \quad \forall x \in \mathcal{H} \quad (2.10)$$

and, for $i = 1, \dots, m$,

$$l_i^*(v_i) \geq l_i^*(p_{2,i,n}) + \langle \nabla l_i^*(p_{2,i,n}), v_i - p_{2,i,n} \rangle \quad \forall v_i \in \mathcal{G}_i. \quad (2.11)$$

Consider arbitrary $x \in \mathcal{H}$ and $\mathbf{v} = (v_1, \dots, v_m) \in \mathcal{G}$. Since

$$\begin{aligned} \left\langle \frac{x_n - p_{1,n}}{\gamma_n}, x - p_{1,n} \right\rangle &= \frac{\|x_n - p_{1,n}\|^2}{2\gamma_n} + \frac{\|x - p_{1,n}\|^2}{2\gamma_n} - \frac{\|x_n - x\|^2}{2\gamma_n} \\ \left\langle \frac{v_{i,n} - p_{2,i,n}}{\gamma_n}, v_i - p_{2,i,n} \right\rangle &= \frac{\|v_{i,n} - p_{2,i,n}\|^2}{2\gamma_n} + \frac{\|v_i - p_{2,i,n}\|^2}{2\gamma_n} - \frac{\|v_{i,n} - v_i\|^2}{2\gamma_n}, \quad i = 1, \dots, m, \end{aligned}$$

we obtain for every $n \geq 0$, by using the more compact notation of the elements in \mathcal{G} and by summing up the inequalities (2.8)–(2.11),

$$\begin{aligned} \frac{\|x_n - x\|^2}{2\gamma_n} + \frac{\|\mathbf{v}_n - \mathbf{v}\|^2}{2\gamma_n} &\geq \frac{\|x_n - p_{1,n}\|^2}{2\gamma_n} + \frac{\|x - p_{1,n}\|^2}{2\gamma_n} + \frac{\|\mathbf{v}_n - \mathbf{p}_{2,n}\|^2}{2\gamma_n} + \frac{\|\mathbf{v} - \mathbf{p}_{2,n}\|^2}{2\gamma_n} \\ &+ \sum_{i=1}^m \langle L_i x_n + \nabla l_i^*(p_{2,i,n}) - \nabla l_i^*(v_{i,n}) - r_i, v_i - p_{2,i,n} \rangle - \sum_{i=1}^m (g_i^* + l_i^*)(v_i) + (f + h)(p_{1,n}) \\ &+ \langle \nabla h(p_{1,n}) - \nabla h(x_n) - L^* v_n + z, x - p_{1,n} \rangle - \left[\sum_{i=1}^m -(g_i^* + l_i^*)(p_{2,i,n}) + (f + h)(x) \right]. \end{aligned}$$

Further, using again the update rules in Algorithm 2.1 and the equations

$$\left\langle \frac{p_{1,n} - x_{n+1}}{\gamma_n}, x - p_{1,n} \right\rangle = \frac{\|x_{n+1} - x\|^2}{2\gamma_n} - \frac{\|x_{n+1} - p_{1,n}\|^2}{2\gamma_n} - \frac{\|x - p_{1,n}\|^2}{2\gamma_n}$$

and, for $i = 1, \dots, m$,

$$\left\langle \frac{p_{2,i,n} - v_{i,n+1}}{\gamma_n}, v_i - p_{2,i,n} \right\rangle = \frac{\|v_{i,n+1} - v_i\|^2}{2\gamma_n} - \frac{\|v_{i,n+1} - p_{2,i,n}\|^2}{2\gamma_n} - \frac{\|v_i - p_{2,i,n}\|^2}{2\gamma_n},$$

we obtain for every $n \geq 0$

$$\begin{aligned} \frac{\|x_n - x\|^2}{2\gamma_n} + \frac{\|\mathbf{v}_n - \mathbf{v}\|^2}{2\gamma_n} &\geq \frac{\|x_{n+1} - x\|^2}{2\gamma_n} + \frac{\|\mathbf{v}_{n+1} - \mathbf{v}\|^2}{2\gamma_n} + \frac{\|x_n - p_{1,n}\|^2}{2\gamma_n} + \frac{\|\mathbf{v}_n - \mathbf{p}_{2,n}\|^2}{2\gamma_n} \\ &\quad - \frac{\|x_{n+1} - p_{1,n}\|^2}{2\gamma_n} - \frac{\|\mathbf{v}_{n+1} - \mathbf{p}_{2,n}\|^2}{2\gamma_n} + [\langle \mathbf{L}p_{1,n} - \mathbf{r}, \mathbf{v} \rangle - G^*(\mathbf{v}) + F(p_{1,n})] \\ &\quad - [\langle \mathbf{L}x - \mathbf{r}, \mathbf{p}_{2,n} \rangle - G^*(\mathbf{p}_{2,n}) + F(x)]. \end{aligned} \quad (2.12)$$

Further, we equip the Hilbert space $\mathcal{H} = \mathcal{H} \times \mathcal{G}$ with the inner product

$$\langle (y, \mathbf{p}), (z, \mathbf{q}) \rangle = \langle y, z \rangle + \langle \mathbf{p}, \mathbf{q} \rangle \quad \forall (y, \mathbf{p}), (z, \mathbf{q}) \in \mathcal{H} \times \mathcal{G} \quad (2.13)$$

and the associated norm $\|(y, \mathbf{p})\| = \sqrt{\|y\|^2 + \|\mathbf{p}\|^2}$ for every $(y, \mathbf{p}) \in \mathcal{H} \times \mathcal{G}$. For every $n \geq 0$ it holds

$$\frac{\|x_{n+1} - p_{1,n}\|^2}{2\gamma_n} + \frac{\|\mathbf{v}_{n+1} - \mathbf{p}_{2,n}\|^2}{2\gamma_n} = \frac{\|(x_{n+1}, \mathbf{v}_{n+1}) - (p_{1,n}, \mathbf{p}_{2,n})\|^2}{2\gamma_n}$$

and, consequently, by making use of the Lipschitz continuity of ∇h and ∇l_i^* , $i = 1, \dots, m$, it shows that

$$\begin{aligned} &\|(x_{n+1}, \mathbf{v}_{n+1}) - (p_{1,n}, \mathbf{p}_{2,n})\| \\ &= \gamma_n \|(\mathbf{L}^*(\mathbf{v}_n - \mathbf{p}_{2,n}), L_1(p_{1,n} - x_n), \dots, L_m(p_{1,n} - x_n)) \\ &\quad + (\nabla h(x_n) - \nabla h(p_{1,n}), \nabla l_1^*(v_{1,n}) - \nabla l_1^*(p_{2,1,n}), \dots, \nabla l_m^*(v_{m,n}) - \nabla l_m^*(p_{2,m,n}))\| \\ &\leq \gamma_n \|(\mathbf{L}^*(\mathbf{v}_n - \mathbf{p}_{2,n}), L_1(p_{1,n} - x_n), \dots, L_m(p_{1,n} - x_n))\| \\ &\quad + \gamma_n \|(\nabla h(x_n) - \nabla h(p_{1,n}), \nabla l_1^*(v_{1,n}) - \nabla l_1^*(p_{2,1,n}), \dots, \nabla l_m^*(v_{m,n}) - \nabla l_m^*(p_{2,m,n}))\| \\ &= \gamma_n \sqrt{\left\| \sum_{i=1}^m L_i^*(v_{i,n} - p_{2,i,n}) \right\|^2 + \sum_{i=1}^m \|L_i(p_{1,n} - x_n)\|^2} \\ &\quad + \gamma_n \sqrt{\|\nabla h(x_n) - \nabla h(p_{1,n})\|^2 + \sum_{i=1}^m \|\nabla l_i^*(v_{i,n}) - \nabla l_i^*(p_{2,i,n})\|^2} \\ &\leq \gamma_n \sqrt{\left(\sum_{i=1}^m \|L_i\|^2 \right) \sum_{i=1}^m \|v_{i,n} - p_{2,i,n}\|^2 + \left(\sum_{i=1}^m \|L_i\|^2 \right) \|p_{1,n} - x_n\|^2} \\ &\quad + \gamma_n \sqrt{\mu^2 \|x_n - p_{1,n}\|^2 + \sum_{i=1}^m \nu_i^2 \|v_{i,n} - p_{2,i,n}\|^2} \end{aligned}$$

$$\leq \gamma_n \left(\sqrt{\sum_{i=1}^m \|L_i\|^2 + \max\{\mu, \nu_1, \dots, \nu_m\}} \right) \|(x_n, \mathbf{v}_n) - (p_{1,n}, \mathbf{p}_{2,n})\|. \quad (2.14)$$

Hence, by taking into consideration the way in which $(\gamma_n)_{n \geq 0}$ is chosen, we have for every $n \geq 0$

$$\begin{aligned} & \frac{1}{2\gamma_n} \left[\|x_n - p_{1,n}\|^2 + \|\mathbf{v}_n - \mathbf{p}_{2,n}\|^2 - \|x_{n+1} - p_{1,n}\|^2 - \|\mathbf{v}_{n+1} - \mathbf{p}_{2,n}\|^2 \right] \\ & \geq \frac{1}{2\gamma_n} \left(1 - \gamma_n^2 \left(\sqrt{\sum_{i=1}^m \|L_i\|^2 + \max\{\mu, \nu_1, \dots, \nu_m\}} \right)^2 \right) \|(x, \mathbf{v}_n) - (p_{1,n}, \mathbf{p}_{2,n})\|^2 \geq 0. \end{aligned}$$

and, consequently, (2.12) reduces to

$$\begin{aligned} \frac{\|x_n - x\|^2}{2\gamma_n} + \frac{\|\mathbf{v}_n - \mathbf{v}\|^2}{2\gamma_n} & \geq \frac{\gamma_{n+1}}{\gamma_n} \frac{\|x_{n+1} - x\|^2}{2\gamma_{n+1}} + [\langle \mathbf{L}p_{1,n} - \mathbf{r}, \mathbf{v} \rangle - G^*(\mathbf{v}) + F(p_{1,n})] \\ & \quad + \frac{\gamma_{n+1}}{\gamma_n} \frac{\|\mathbf{v}_{n+1} - \mathbf{v}\|^2}{2\gamma_{n+1}} - \left[\langle \mathbf{L}x - \mathbf{r}, \mathbf{p}_{2,n} \rangle - G^*(\mathbf{p}_{2,n}) + F(x) \right]. \end{aligned}$$

Let $N \geq 1$ be an arbitrary natural number. Summing the above inequality up from $n = 0$ to $N - 1$ and using the fact that $(\gamma_n)_{n \geq 0}$ is nondecreasing, it follows that

$$\begin{aligned} \frac{\|x_0 - x\|^2}{2\gamma_0} + \frac{\|\mathbf{v}_0 - \mathbf{v}\|^2}{2\gamma_0} & \geq \frac{\|x_N - x\|^2}{2\gamma_N} + \sum_{n=0}^{N-1} [\langle \mathbf{L}p_{1,n} - \mathbf{r}, \mathbf{v} \rangle - G^*(\mathbf{v}) + F(p_{1,n})] \\ & \quad + \frac{\|\mathbf{v}_N - \mathbf{v}\|^2}{2\gamma_N} - \sum_{n=0}^{N-1} \left[\langle \mathbf{L}x - \mathbf{r}, \mathbf{p}_{2,n} \rangle - G^*(\mathbf{p}_{2,n}) + F(x) \right]. \end{aligned} \quad (2.15)$$

Replacing $x = \bar{x}$ and $\mathbf{v} = \bar{\mathbf{v}}$ in the above estimate, since they fulfill (2.4), we obtain

$$\sum_{n=0}^{N-1} [\langle \mathbf{L}p_{1,n} - \mathbf{r}, \bar{\mathbf{v}} \rangle - G^*(\bar{\mathbf{v}}) + F(p_{1,n})] - \sum_{n=0}^{N-1} \left[\langle \mathbf{L}\bar{x} - \mathbf{r}, \mathbf{p}_{2,n} \rangle - G^*(\mathbf{p}_{2,n}) + F(\bar{x}) \right] \geq 0.$$

Consequently,

$$\frac{\|x_0 - \bar{x}\|^2}{2\gamma_0} + \frac{\|\mathbf{v}_0 - \bar{\mathbf{v}}\|^2}{2\gamma_0} \geq \frac{\|x_N - \bar{x}\|^2}{2\gamma_N} + \frac{\|\mathbf{v}_N - \bar{\mathbf{v}}\|^2}{2\gamma_N}$$

and statement (c) follows. On the other hand, dividing (2.15) by N , using the convexity of F and G^* , and denoting $x^N := \frac{1}{N} \sum_{n=0}^{N-1} p_{1,n}$ and $\mathbf{v}_i^N := \frac{1}{N} \sum_{n=0}^{N-1} p_{2,i,n}$, $i = 1, \dots, m$, we obtain

$$\begin{aligned} \frac{1}{N} \left(\frac{\|x_0 - x\|^2}{2\gamma_0} + \frac{\|\mathbf{v}_0 - \mathbf{v}\|^2}{2\gamma_0} \right) & \geq \left[\langle \mathbf{L}x^N - \mathbf{r}, \mathbf{v} \rangle - G^*(\mathbf{v}) + F(x^N) \right] \\ & \quad - \left[\langle \mathbf{L}x - \mathbf{r}, \mathbf{v}^N \rangle - G^*(\mathbf{v}^N) + F(x) \right], \end{aligned}$$

which shows (2.7) when passing to the supremum over $x \in B_1$ and $\mathbf{v} \in B_2$. In this way statement (d) is verified. The weak convergence of (x^N, \mathbf{v}^N) to $(\bar{x}, \bar{\mathbf{v}})$ when N converges to $+\infty$ is an easy consequence of the Stolz–Cesàro Theorem, fact which shows (e). \square

Remark 2.1. In the situation when the functions g_i are Lipschitz continuous on \mathcal{G}_i , $i = 1, \dots, m$, inequality (2.7) provides for the sequence of the values of the objective of (P) taken at $(x^N)_{N \geq 1}$ a convergence rate of $\mathcal{O}(\frac{1}{N})$, namely, it holds

$$F(x^N) + G(\mathbf{L}x^N - \mathbf{r}) - F(\bar{x}) - G(\mathbf{L}\bar{x} - \mathbf{r}) \leq \frac{C(B_1, B_2)}{N} \quad \forall N \geq 1. \quad (2.16)$$

Indeed, due to statement (b) of the previous theorem, the sequence $(p_{1,n})_{n \geq 0} \subseteq \mathcal{H}$ is bounded and one can take $B_1 \subset \mathcal{H}$ being a bounded, convex and closed set containing this sequence. Obviously, $\bar{x} \in B_1$. On the other hand, we take $B_2 = \text{dom } g_1^* \times \dots \times \text{dom } g_m^*$, which is in this situation a bounded set. Then it holds, using the Fenchel-Moreau Theorem and the Young-Fenchel inequality, that

$$\begin{aligned} \mathcal{G}_{B_1 \times B_2}(x^N, \mathbf{v}^N) &= F(x^N) + G(\mathbf{L}x^N - \mathbf{r}) + G^*(\mathbf{v}^N) - \inf_{\tilde{x} \in B_1} \left\{ \langle \mathbf{L}\tilde{x} - \mathbf{r}, \mathbf{v}^N \rangle + F(\tilde{x}) \right\} \\ &\geq F(x^N) + G(\mathbf{L}x^N - \mathbf{r}) + G^*(\mathbf{v}^N) - \langle \mathbf{L}\bar{x} - \mathbf{r}, \mathbf{v}^N \rangle - F(\bar{x}) \\ &\geq F(x^N) + G(\mathbf{L}x^N - \mathbf{r}) - F(\bar{x}) - G(\mathbf{L}\bar{x} - \mathbf{r}). \end{aligned}$$

Hence, (2.16) follows by statement (d) in Theorem 2.1.

In a similar way, one can show that, whenever f is Lipschitz continuous, (2.7) provides for the sequence of the values of the objective of (D) taken at $(v^N)_{N \geq 1}$ a convergence rate of $\mathcal{O}(\frac{1}{N})$.

Remark 2.2. If \mathcal{G}_i , $i = 1, \dots, m$, are finite-dimensional real Hilbert spaces, then (2.16) is true, even under the weaker assumption that the convex functions g_i , $i = 1, \dots, m$, have full domain, without necessarily being Lipschitz continuous. The set $B_1 \subset \mathcal{H}$ can be chosen as in Remark 2.1, but this time we take $B_2 = \times_{i=1}^m \bigcup_{n \geq 0} \partial g_i(L_i p_{1,n}) \subset \mathcal{G}$, by noticing also that the functions g_i , $i = 1, \dots, m$, are everywhere subdifferentiable.

The set B_2 is bounded, as for every $i = 1, \dots, m$ the set $\bigcup_{n \geq 0} \partial g_i(L_i p_{1,n})$ is bounded. Let be $i \in \{1, \dots, m\}$ fixed. Indeed, as $p_{1,n} \rightarrow \bar{x}$, it follows that $L_i p_{1,n} \rightarrow L_i \bar{x}$ for $i = 1, \dots, m$. Using the fact that the subdifferential of g_i is a locally bounded operator at $L_i \bar{x}$, the boundedness of $\bigcup_{n \geq 0} \partial g_i(L_i p_{1,n})$ follows automatically.

For this choice of the sets B_1 and B_2 , by using the same arguments as in the previous remark, it follows that (2.16) is true.

3 Zeros of sums of monotone operators

In this section we turn our attention to the primal-dual monotone inclusion problems formulated in Problem 1.1 with the aim to provide accelerations of the iterative method described in Theorem 1.1 under the additional strong monotonicity assumptions.

3.1 The case when $A + C$ is strongly monotone

We focus first on the case when $A + C$ is ρ -strongly monotone for some $\rho \in \mathbb{R}_{++}$ and investigate the impact of this assumption on the convergence rate of the sequence of primal iterates. The condition $A + C$ is ρ -strongly monotone is fulfilled when either $A : \mathcal{H} \rightarrow 2^{\mathcal{H}}$ or $C : \mathcal{H} \rightarrow \mathcal{H}$ is ρ -strongly monotone. In case that A is ρ_1 -monotone and C is ρ_2 -monotone, the sum $A + C$ is ρ -monotone with $\rho = \rho_1 + \rho_2$.

Remark 3.1. The situation when $B_i^{-1} + D_i^{-1}$ is τ_i -strongly monotone with $\tau_i \in \mathbb{R}_{++}$ for $i = 1, \dots, m$, which improves the convergence rate of the sequence of dual iterates, can be handled with appropriate modifications.

Due to technical reasons we assume in the following that the operators D_i^{-1} in Problem 1.1 are zero for $i = 1, \dots, m$, thus, $D_i(0) = \mathcal{G}_i$ and $D_i(x) = \emptyset$ for $x \neq 0$, for $i = 1, \dots, m$. In Remark 3.2 we show how the results given in this particular context can be employed when treating the primal-dual pair of monotone inclusions (1.2)-(1.3). Consequently, the problem we deal with in this subsection is as follows.

Problem 3.1. Consider the real Hilbert spaces \mathcal{H} and $\mathcal{G}_i, i = 1, \dots, m$, $A : \mathcal{H} \rightarrow 2^{\mathcal{H}}$ a maximally monotone operator and $C : \mathcal{H} \rightarrow \mathcal{H}$ a monotone and μ -Lipschitzian operator for some $\mu \in \mathbb{R}_{++}$. Furthermore, let $z \in \mathcal{H}$ and for every $i \in \{1, \dots, m\}$, let $r_i \in \mathcal{G}_i$, let $B_i : \mathcal{G}_i \rightarrow 2^{\mathcal{G}_i}$ be maximally monotone operators and let $L_i : \mathcal{H} \rightarrow \mathcal{G}_i$ be a nonzero linear continuous operator. The problem is to solve the primal inclusion

$$\text{find } \bar{x} \in \mathcal{H} \text{ such that } z \in A\bar{x} + \sum_{i=1}^m L_i^* B_i(L_i \bar{x} - r_i) + C\bar{x}, \quad (3.1)$$

together with the dual inclusion

$$\text{find } \bar{v}_1 \in \mathcal{G}_1, \dots, \bar{v}_m \in \mathcal{G}_m \text{ such that } (\exists x \in \mathcal{H}) \begin{cases} z - \sum_{i=1}^m L_i^* \bar{v}_i \in Ax + Cx \\ \bar{v}_i \in B_i(L_i x - r_i), i = 1, \dots, m. \end{cases} \quad (3.2)$$

The subsequent algorithm represents an accelerated version of the one given in Theorem 1.1 and relies on the fruitful idea of using a second sequence of variable step length parameters $(\sigma_n)_{n \geq 0} \subseteq \mathbb{R}_{++}$, which, together with the sequence of parameters $(\gamma_n)_{n \geq 0} \subseteq \mathbb{R}_{++}$, play an important role in the convergence analysis.

Algorithm 3.1. Let $x_0 \in \mathcal{H}$, $(v_{1,0}, \dots, v_{m,0}) \in \mathcal{G}$,

$$\gamma_0 \in \left(0, \min \left\{ 1, \frac{\sqrt{1+4\rho}}{2(1+2\rho)\mu} \right\} \right) \text{ and set } \sigma_0 = \frac{1}{2\gamma_0(1+2\rho)\sum_{i=1}^m \|L_i\|^2}.$$

Consider the following updates:

$$(\forall n \geq 0) \begin{cases} p_{1,n} = J_{\gamma_n A}(x_n - \gamma_n(Cx_n + \sum_{i=1}^m L_i^* v_{i,n} - z)) \\ \text{For } i = 1, \dots, m \\ \quad p_{2,i,n} = J_{\sigma_n B_i^{-1}}(v_{i,n} + \sigma_n(L_i x_n - r_i)) \\ \quad v_{i,n+1} = \sigma_n L_i(p_{1,n} - x_n) + p_{2,i,n} \\ x_{n+1} = \gamma_n \sum_{i=1}^m L_i^*(v_{i,n} - p_{2,i,n}) + \gamma_n(Cx_n - Cp_{1,n}) + p_{1,n} \\ \theta_n = 1/\sqrt{1+2\rho\gamma_n(1-\gamma_n)}, \gamma_{n+1} = \theta_n \gamma_n, \sigma_{n+1} = \sigma_n/\theta_n. \end{cases} \quad (3.3)$$

Theorem 3.1. In Problem 3.1 suppose that $A+C$ is ρ -strongly monotone with $\rho \in \mathbb{R}_{++}$ and let $(\bar{x}, \bar{v}_1, \dots, \bar{v}_m) \in \mathcal{H} \times \mathcal{G}$ be a primal-dual solution to Problem 3.1. Then for every $n \geq 0$ it holds

$$\|x_n - \bar{x}\|^2 + \gamma_n \sum_{i=1}^m \frac{\|v_{i,n} - \bar{v}_i\|^2}{\sigma_n} \leq \gamma_n^2 \left(\frac{\|x_0 - \bar{x}\|^2}{\gamma_0^2} + \sum_{i=1}^m \frac{\|v_{i,0} - \bar{v}_i\|^2}{\gamma_0 \sigma_0} \right), \quad (3.4)$$

where $\gamma_n, \sigma_n \in \mathbb{R}_{++}$, $x_n \in \mathcal{H}$ and $(v_{1,n}, \dots, v_{m,n}) \in \mathcal{G}$ are the iterates generated by Algorithm 3.1.

Proof. Taking into account the definitions of the resolvents occurring in Algorithm 3.1 we obtain

$$\text{and } \begin{aligned} \frac{x_n - p_{1,n}}{\gamma_n} - Cx_n - \sum_{i=1}^m L_i^* v_{i,n} + z &\in Ap_{1,n} \\ \frac{v_{i,n} - p_{2,i,n}}{\sigma_n} + L_i x_n - r_i &\in B_i^{-1} p_{2,i,n}, i = 1, \dots, m, \end{aligned}$$

which, in the light of the updating rules in (3.3), furnishes for every $n \geq 0$

$$\text{and } \begin{aligned} \frac{x_n - x_{n+1}}{\gamma_n} - \sum_{i=1}^m L_i^* p_{2,i,n} + z &\in (A + C)p_{1,n} \\ \frac{v_{i,n} - v_{i,n+1}}{\sigma_n} + L_i p_{1,n} - r_i &\in B_i^{-1} p_{2,i,n}, i = 1, \dots, m. \end{aligned} \quad (3.5)$$

The primal-dual solution $(\bar{x}, \bar{v}_1, \dots, \bar{v}_m) \in \mathcal{H} \times \mathcal{G}$ to Problem 3.1 fulfills (see (1.4), where D_i^{-1} are taken to be zero for $i = 1, \dots, m$)

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

Since the sum $A + C$ is ρ -strongly monotone, we have for every $n \geq 0$

$$\left\langle p_{1,n} - \bar{x}, \frac{x_n - x_{n+1}}{\gamma_n} - \sum_{i=1}^m L_i^* p_{2,i,n} + z - \left(z - \sum_{i=1}^m L_i^* \bar{v}_i \right) \right\rangle \geq \rho \|p_{1,n} - \bar{x}\|^2 \quad (3.6)$$

while, due to the monotonicity of $B_i^{-1} : \mathcal{G}_i \rightarrow 2^{\mathcal{G}_i}$, we obtain for every $n \geq 0$

$$\left\langle p_{2,i,n} - \bar{v}_i, \frac{v_{i,n} - v_{i,n+1}}{\sigma_n} + L_i p_{1,n} - r_i - (L_i \bar{x} - r_i) \right\rangle \geq 0, i = 1, \dots, m. \quad (3.7)$$

Further, we set

$$\bar{\mathbf{v}} = (\bar{v}_1, \dots, \bar{v}_m), \quad \mathbf{v}_n = (v_{1,n}, \dots, v_{m,n}), \quad \mathbf{p}_{2,n} = (p_{2,1,n}, \dots, p_{2,m,n}).$$

Summing up the inequalities (3.6) and (3.7), it follows that

$$\begin{aligned} &\left\langle p_{1,n} - \bar{x}, \frac{x_n - x_{n+1}}{\gamma_n} \right\rangle + \left\langle \mathbf{p}_{2,n} - \bar{\mathbf{v}}, \frac{\mathbf{v}_n - \mathbf{v}_{n+1}}{\sigma_n} \right\rangle + \left\langle p_{1,n} - \bar{x}, \mathbf{L}^*(\bar{\mathbf{v}} - \mathbf{p}_{2,n}) \right\rangle \\ &+ \left\langle \mathbf{p}_{2,n} - \bar{\mathbf{v}}, \mathbf{L}(p_{1,n} - \bar{x}) \right\rangle \geq \rho \|p_{1,n} - \bar{x}\|^2. \end{aligned} \quad (3.8)$$

and, from here,

$$\left\langle p_{1,n} - \bar{x}, \frac{x_n - x_{n+1}}{\gamma_n} \right\rangle + \left\langle \mathbf{p}_{2,n} - \bar{\mathbf{v}}, \frac{\mathbf{v}_n - \mathbf{v}_{n+1}}{\sigma_n} \right\rangle \geq \rho \|p_{1,n} - \bar{x}\|^2 \quad \forall n \geq 0. \quad (3.9)$$

In the light of the equations

$$\begin{aligned} \left\langle p_{1,n} - \bar{x}, \frac{x_n - x_{n+1}}{\gamma_n} \right\rangle &= \left\langle p_{1,n} - x_{n+1}, \frac{x_n - x_{n+1}}{\gamma_n} \right\rangle + \left\langle x_{n+1} - \bar{x}, \frac{x_n - x_{n+1}}{\gamma_n} \right\rangle \\ &= \frac{\|x_{n+1} - p_{1,n}\|^2}{2\gamma_n} - \frac{\|x_n - p_{1,n}\|^2}{2\gamma_n} + \frac{\|x_n - \bar{x}\|^2}{2\gamma_n} - \frac{\|x_{n+1} - \bar{x}\|^2}{2\gamma_n}, \end{aligned}$$

and

$$\begin{aligned} \left\langle \mathbf{p}_{2,n} - \bar{\mathbf{v}}, \frac{\mathbf{v}_n - \mathbf{v}_{n+1}}{\sigma_n} \right\rangle &= \left\langle \mathbf{p}_{2,n} - \mathbf{v}_{n+1}, \frac{\mathbf{v}_n - \mathbf{v}_{n+1}}{\sigma_n} \right\rangle + \left\langle \mathbf{v}_{n+1} - \bar{\mathbf{v}}, \frac{\mathbf{v}_n - \mathbf{v}_{n+1}}{\sigma_n} \right\rangle \\ &= \frac{\|\mathbf{v}_{n+1} - \mathbf{p}_{2,n}\|^2}{2\sigma_n} - \frac{\|\mathbf{v}_n - \mathbf{p}_{2,n}\|^2}{2\sigma_n} + \frac{\|\mathbf{v}_n - \bar{\mathbf{v}}\|^2}{2\sigma_n} - \frac{\|\mathbf{v}_{n+1} - \bar{\mathbf{v}}\|^2}{2\sigma_n} \end{aligned}$$

inequality (3.9) reads for every $n \geq 0$

$$\begin{aligned} \frac{\|x_n - \bar{x}\|^2}{2\gamma_n} + \frac{\|\mathbf{v}_n - \bar{\mathbf{v}}\|^2}{2\sigma_n} &\geq \rho \|p_{1,n} - \bar{x}\|^2 + \frac{\|x_{n+1} - \bar{x}\|^2}{2\gamma_n} + \frac{\|\mathbf{v}_{n+1} - \bar{\mathbf{v}}\|^2}{2\sigma_n} + \frac{\|x_n - p_{1,n}\|^2}{2\gamma_n} \\ &\quad + \frac{\|\mathbf{v}_n - \mathbf{p}_{2,n}\|^2}{2\sigma_n} - \frac{\|x_{n+1} - p_{1,n}\|^2}{2\gamma_n} - \frac{\|\mathbf{v}_{n+1} - \mathbf{p}_{2,n}\|^2}{2\sigma_n}. \end{aligned} \quad (3.10)$$

Using that $2ab \leq \alpha a^2 + \frac{b^2}{\alpha}$ for all $a, b \in \mathbb{R}$, $\alpha \in \mathbb{R}_{++}$, we obtain for $\alpha := \gamma_n$,

$$\begin{aligned} \rho \|p_{1,n} - \bar{x}\|^2 &\geq \rho \left(\|x_{n+1} - \bar{x}\|^2 - 2\|x_{n+1} - \bar{x}\| \|x_{n+1} - p_{1,n}\| + \|x_{n+1} - p_{1,n}\|^2 \right) \\ &\geq \frac{2\rho\gamma_n(1-\gamma_n)}{2\gamma_n} \|x_{n+1} - \bar{x}\|^2 - \frac{2\rho(1-\gamma_n)}{2\gamma_n} \|x_{n+1} - p_{1,n}\|^2, \end{aligned}$$

which, in combination with (3.10), yields for every $n \geq 0$

$$\begin{aligned} \frac{\|x_n - \bar{x}\|^2}{2\gamma_n} + \frac{\|\mathbf{v}_n - \bar{\mathbf{v}}\|^2}{2\sigma_n} &\geq \frac{(1 + 2\rho\gamma_n(1-\gamma_n))\|x_{n+1} - \bar{x}\|^2}{2\gamma_n} + \frac{\|\mathbf{v}_{n+1} - \bar{\mathbf{v}}\|^2}{2\sigma_n} \\ &\quad + \frac{\|x_n - p_{1,n}\|^2}{2\gamma_n} + \frac{\|\mathbf{v}_n - \mathbf{p}_{2,n}\|^2}{2\sigma_n} - \frac{(1 + 2\rho(1-\gamma_n))\|x_{n+1} - p_{1,n}\|^2}{2\gamma_n} - \frac{\|\mathbf{v}_{n+1} - \mathbf{p}_{2,n}\|^2}{2\sigma_n}. \end{aligned} \quad (3.11)$$

Investigating the last two terms in the right-hand side of the above estimate it shows for every $n \geq 0$ that

$$\begin{aligned} &- \frac{(1 + 2\rho(1-\gamma_n))\|x_{n+1} - p_{1,n}\|^2}{2\gamma_n} \\ &\geq - \frac{(1 + 2\rho)\gamma_n}{2} \left\| \sum_{i=1}^m L_i^*(v_{i,n} - p_{2,i,n}) + (Cx_n - Cp_{1,n}) \right\|^2 \\ &\geq - \frac{2(1 + 2\rho)\gamma_n}{2} \left(\left(\sum_{i=1}^m \|L_i\|^2 \right) \|\mathbf{v}_n - \mathbf{p}_{2,n}\|^2 + \mu^2 \|x_n - p_{1,n}\|^2 \right), \end{aligned}$$

and

$$- \frac{\|\mathbf{v}_{n+1} - \mathbf{p}_{2,n}\|^2}{2\sigma_n} = - \frac{\sigma_n}{2} \left(\sum_{i=1}^m \|L_i(p_{1,n} - x_n)\|^2 \right) \geq - \frac{\sigma_n}{2} \left(\sum_{i=1}^m \|L_i\|^2 \right) \|p_{1,n} - x_n\|^2.$$

Hence, for every $n \geq 0$ it holds

$$\begin{aligned}
& \frac{\|x_n - p_{1,n}\|^2}{2\gamma_n} + \frac{\|\mathbf{v}_n - \mathbf{p}_{2,n}\|^2}{2\sigma_n} - \frac{(1 + 2\rho(1 - \gamma_n))\|x_{n+1} - p_{1,n}\|^2}{2\gamma_n} - \frac{\|\mathbf{v}_{n+1} - \mathbf{p}_{2,n}\|^2}{2\sigma_n} \\
& \geq \frac{(1 - \gamma_n\sigma_n \sum_{i=1}^m \|L_i\|^2 - 2(1 + 2\rho)\gamma_n^2\mu^2)}{2\gamma_n} \|p_{1,n} - x_n\|^2 \\
& \quad + \frac{(1 - 2\gamma_n\sigma_n(1 + 2\rho) \sum_{i=1}^m \|L_i\|^2)}{2\sigma_n} \|\mathbf{v}_n - \mathbf{p}_{2,n}\|^2 \\
& \geq 0.
\end{aligned}$$

The nonnegativity of the expression in the above relation follows because of the sequence $(\gamma_n)_{n \geq 0}$ is nonincreasing, $\gamma_n\sigma_n = \gamma_0\sigma_0$ for every $n \geq 0$ and

$$\gamma_0 \in \left(0, \min \left\{1, \frac{\sqrt{1 + 4\rho}}{2(1 + 2\rho)\mu}\right\}\right) \text{ and } \sigma_0 = \frac{1}{2\gamma_0(1 + 2\rho) \sum_{i=1}^m \|L_i\|^2}.$$

Consequently, inequality (3.11) becomes

$$\frac{\|x_n - \bar{x}\|^2}{2\gamma_n} + \frac{\|\mathbf{v}_n - \bar{\mathbf{v}}\|^2}{2\sigma_n} \geq \frac{(1 + 2\rho\gamma_n(1 - \gamma_n))\|x_{n+1} - \bar{x}\|^2}{2\gamma_n} + \frac{\|\mathbf{v}_{n+1} - \bar{\mathbf{v}}\|^2}{2\sigma_n} \quad \forall n \geq 0. \tag{3.12}$$

Notice that we have $\gamma_{n+1} < \gamma_n$, $\sigma_{n+1} > \sigma_n$ and $\gamma_{n+1}\sigma_{n+1} = \gamma_n\sigma_n$ for every $n \geq 0$. Dividing (3.12) by γ_n and making use of

$$\theta_n = \frac{1}{\sqrt{1 + 2\rho\gamma_n(1 - \gamma_n)}}, \quad \gamma_{n+1} = \theta_n\gamma_n, \quad \sigma_{n+1} = \frac{\sigma_n}{\theta_n},$$

we obtain

$$\frac{\|x_n - \bar{x}\|^2}{2\gamma_n^2} + \frac{\|\mathbf{v}_n - \bar{\mathbf{v}}\|^2}{2\gamma_n\sigma_n} \geq \frac{\|x_{n+1} - \bar{x}\|^2}{2\gamma_{n+1}^2} + \frac{\|\mathbf{v}_{n+1} - \bar{\mathbf{v}}\|^2}{2\gamma_{n+1}\sigma_{n+1}} \quad \forall n \geq 0.$$

Let be $N \geq 1$. Summing this inequalities from $n = 0$ to $N - 1$, we finally get

$$\frac{\|x_0 - \bar{x}\|^2}{2\gamma_0^2} + \frac{\|\mathbf{v}_0 - \bar{\mathbf{v}}\|^2}{2\gamma_0\sigma_0} \geq \frac{\|x_N - \bar{x}\|^2}{2\gamma_N^2} + \frac{\|\mathbf{v}_N - \bar{\mathbf{v}}\|^2}{2\gamma_N\sigma_N}. \tag{3.13}$$

In conclusion,

$$\frac{\|x_n - \bar{x}\|^2}{2} + \gamma_n \frac{\|\mathbf{v}_n - \bar{\mathbf{v}}\|^2}{2\sigma_n} \leq \gamma_n^2 \left(\frac{\|x_0 - \bar{x}\|^2}{2\gamma_0^2} + \frac{\|\mathbf{v}_0 - \bar{\mathbf{v}}\|^2}{2\gamma_0\sigma_0} \right) \quad \forall n \geq 0, \tag{3.14}$$

which completes the proof. \square

Next we show that $\rho\gamma_n$ converges like $\frac{1}{n}$ as $n \rightarrow +\infty$.

Proposition 3.2. *Let $\gamma_0 \in (0, 1)$ and consider the sequence $(\gamma_n)_{n \geq 0} \subseteq \mathbb{R}_{++}$, where*

$$\gamma_{n+1} = \frac{\gamma_n}{\sqrt{1 + 2\rho\gamma_n(1 - \gamma_n)}} \quad \forall n \geq 0. \tag{3.15}$$

Then $\lim_{n \rightarrow +\infty} n\rho\gamma_n = 1$.

Proof. Since the sequence $(\gamma_n)_{n \geq 0} \subseteq (0, 1)$ is bounded and decreasing, it converges towards some $l \in [0, 1)$ as $n \rightarrow +\infty$. We let $n \rightarrow +\infty$ in (3.15) and obtain

$$l^2(1 + 2\rho l(1 - l)) = l^2 \Leftrightarrow 2\rho l^3(1 - l) = 0,$$

which shows that $l = 0$, i. e. $\gamma_n \rightarrow 0$ ($n \rightarrow +\infty$). On the other hand, (3.15) implies that $\frac{\gamma_n}{\gamma_{n+1}} \rightarrow 1$ ($n \rightarrow +\infty$). As $(\frac{1}{\gamma_n})_{n \geq 0}$ is a strictly increasing and unbounded sequence, by applying the Stolz–Cesàro Theorem it shows that

$$\begin{aligned} \lim_{n \rightarrow +\infty} n\gamma_n &= \lim_{n \rightarrow +\infty} \frac{n}{\frac{1}{\gamma_n}} = \lim_{n \rightarrow +\infty} \frac{n+1-n}{\frac{1}{\gamma_{n+1}} - \frac{1}{\gamma_n}} = \lim_{n \rightarrow +\infty} \frac{\gamma_n \gamma_{n+1}}{\gamma_n - \gamma_{n+1}} \\ &= \lim_{n \rightarrow +\infty} \frac{\gamma_n \gamma_{n+1} (\gamma_n + \gamma_{n+1})}{\gamma_n^2 - \gamma_{n+1}^2} \stackrel{(3.15)}{=} \lim_{n \rightarrow +\infty} \frac{\gamma_n \gamma_{n+1} (\gamma_n + \gamma_{n+1})}{2\rho \gamma_{n+1}^2 \gamma_n (1 - \gamma_n)} \\ &= \lim_{n \rightarrow +\infty} \frac{\gamma_n + \gamma_{n+1}}{2\rho \gamma_{n+1} (1 - \gamma_n)} = \lim_{n \rightarrow +\infty} \frac{\frac{\gamma_n}{\gamma_{n+1}} + 1}{2\rho(1 - \gamma_n)} = \frac{2}{2\rho} = \frac{1}{\rho}, \end{aligned}$$

which completes the proof. \square

Hence, we have shown the following result.

Theorem 3.3. *In Problem 3.1 suppose that $A + C$ is ρ -strongly monotone and let $(\bar{x}, \bar{v}_1, \dots, \bar{v}_m) \in \mathcal{H} \times \mathcal{G}$ be a primal-dual solution to Problem 3.1. Then, for any $\varepsilon > 0$, there exists some $n_0 \in \mathbb{N}$ (depending on ε and $\rho\gamma_0$) such that for any $n \geq n_0$*

$$\|x_n - \bar{x}\|^2 \leq \frac{1 + \varepsilon}{n^2} \left(\frac{\|x_0 - \bar{x}\|^2}{\rho^2 \gamma_0^2} + \sum_{i=1}^m \frac{\|v_{i,0} - \bar{v}_i\|^2}{\rho^2 \gamma_0 \sigma_0} \right), \quad (3.16)$$

where $\gamma_n, \sigma_n \in \mathbb{R}_{++}$, $x_n \in \mathcal{H}$ and $(v_{1,n}, \dots, v_{m,n}) \in \mathcal{G}$ are the iterates generated by Algorithm 3.1.

Remark 3.2. In Algorithm 3.1 and Theorem 3.3 we assumed that $D_i^{-1} = 0$ for $i = 1, \dots, m$, however, similar statements can be also provided for Problem 1.1 under the additional assumption that the operators $D_i : \mathcal{G}_i \rightarrow 2^{\mathcal{G}_i}$ are such that D_i^{-1} is ν_i^{-1} -cocoercive with $\nu_i \in \mathbb{R}_{++}$ for $i = 1, \dots, m$. This assumption is in general stronger than assuming that D_i is monotone and D_i^{-1} is ν_i -Lipschitzian for $i = 1, \dots, m$. However, it guarantees that D_i is ν_i^{-1} -strongly monotone and maximally monotone for $i = 1, \dots, m$ (see [1, Example 20.28, Proposition 20.22 and Example 22.6]). We introduce the Hilbert space $\mathcal{H} = \mathcal{H} \times \mathcal{G}$, the element $z = (z, 0, \dots, 0) \in \mathcal{H}$ and the maximally monotone operator $\mathbf{A} : \mathcal{H} \rightarrow 2^{\mathcal{H}}$, $\mathbf{A}(x, y_1, \dots, y_m) = (Ax, D_1 y_1, \dots, D_m y_m)$ and the monotone and Lipschitzian operator $\mathbf{C} : \mathcal{H} \rightarrow \mathcal{H}$, $\mathbf{C}(x, y_1, \dots, y_m) = (Cx, 0, \dots, 0)$. Notice also that $\mathbf{A} + \mathbf{C}$ is strongly monotone. Furthermore, we introduce the element $\mathbf{r} = (r_1, \dots, r_m) \in \mathcal{G}$, the maximally monotone operator $\mathbf{B} : \mathcal{G} \rightarrow 2^{\mathcal{G}}$, $\mathbf{B}(y_1, \dots, y_m) = (B_1 y_1, \dots, B_m y_m)$, and the linear continuous operator $\mathbf{L} : \mathcal{H} \rightarrow \mathcal{G}$, $\mathbf{L}(x, y_1, \dots, y_m) = (L_1 x - y_1, \dots, L_m x - y_m)$, having as adjoint $\mathbf{L}^* : \mathcal{G} \rightarrow \mathcal{H}$, $\mathbf{L}^*(q_1, \dots, q_m) = (\sum_{i=1}^m L_i^* q_i, -q_1, \dots, -q_m)$. We consider the primal problem

$$\text{find } \bar{x} = (\bar{x}, \bar{p}_1 \dots \bar{p}_m) \in \mathcal{H} \text{ such that } z \in \mathbf{A}\bar{x} + \mathbf{L}^* \mathbf{B}(\mathbf{L}\bar{x} - \mathbf{r}) + \mathbf{C}\bar{x}, \quad (3.17)$$

together with the dual inclusion problem

$$\text{find } \bar{\mathbf{v}} \in \mathcal{G} \text{ such that } (\exists \mathbf{x} \in \mathcal{H}) \begin{cases} \mathbf{z} - \mathbf{L}^* \bar{\mathbf{v}} \in \mathbf{A} \mathbf{x} + \mathbf{C} \mathbf{x} \\ \bar{\mathbf{v}} \in \mathbf{B}(\mathbf{L} \mathbf{x} - \mathbf{r}) \end{cases}. \quad (3.18)$$

We notice that Algorithm 3.1 can be employed for solving this primal-dual pair of monotone inclusion problems and, by separately involving the resolvents of A, B_i and $D_i, i = 1, \dots, m$, as for $\gamma \in \mathbb{R}_{++}$

$$\begin{aligned} J_{\gamma \mathbf{A}}(x, y_1, \dots, y_m) &= (J_{\gamma \mathbf{A}} x, J_{\gamma D_1} y_1, \dots, J_{\gamma D_m} y_m) \quad \forall (x, y_1, \dots, y_m) \in \mathcal{H} \\ J_{\gamma \mathbf{B}}(q_1, \dots, q_m) &= (J_{\gamma B_1} q_1, \dots, J_{\gamma B_m} q_m) \quad \forall (q_1, \dots, q_m) \in \mathcal{G}. \end{aligned}$$

Having $(\bar{\mathbf{x}}, \bar{\mathbf{v}}) \in \mathcal{H} \times \mathcal{G}$ a primal-dual solution to the primal-dual pair of monotone inclusion problems (3.17)-(3.18), Algorithm 3.1 generates a sequence of primal iterates fulfilling (3.16) in \mathcal{H} . Moreover, $(\bar{\mathbf{x}}, \bar{\mathbf{v}})$ is a primal-dual solution to (3.17)-(3.18) if and only if

$$\begin{aligned} &\mathbf{z} - \mathbf{L}^* \bar{\mathbf{v}} \in \mathbf{A} \bar{\mathbf{x}} + \mathbf{C} \bar{\mathbf{x}} \text{ and } \bar{\mathbf{v}} \in \mathbf{B}(\mathbf{L} \bar{\mathbf{x}} - \mathbf{r}) \\ \Leftrightarrow &\mathbf{z} - \sum_{i=1}^m L_i^* \bar{v}_i \in \mathbf{A} \bar{\mathbf{x}} + \mathbf{C} \bar{\mathbf{x}} \text{ and } \bar{v}_i \in D_i \bar{p}_i, \bar{v}_i \in B_i(L_i \bar{\mathbf{x}} - \bar{p}_i - r_i), i = 1, \dots, m \\ \Leftrightarrow &\mathbf{z} - \sum_{i=1}^m L_i^* \bar{v}_i \in \mathbf{A} \bar{\mathbf{x}} + \mathbf{C} \bar{\mathbf{x}} \text{ and } \bar{v}_i \in D_i \bar{p}_i, L_i \bar{\mathbf{x}} - r_i \in B_i^{-1} \bar{v}_i + \bar{p}_i, i = 1, \dots, m. \end{aligned}$$

Thus, if $(\bar{\mathbf{x}}, \bar{\mathbf{v}})$ is a primal-dual solution to (3.17)-(3.18), then $(\bar{\mathbf{x}}, \bar{\mathbf{v}})$ is a primal-dual solution to Problem 1.1. Viceversa, if $(\bar{\mathbf{x}}, \bar{\mathbf{v}})$ is a primal-dual solution to Problem 1.1, then, choosing $\bar{p}_i \in D_i^{-1} \bar{v}_i, i = 1, \dots, m$, and $\bar{\mathbf{x}} = (\bar{\mathbf{x}}, \bar{p}_1 \dots \bar{p}_m)$, it yields that $(\bar{\mathbf{x}}, \bar{\mathbf{v}})$ is a primal-dual solution to (3.17)-(3.18). In conclusion, the first component of every primal iterate in \mathcal{H} generated by Algorithm 3.1 for finding the primal-dual solution $(\bar{\mathbf{x}}, \bar{\mathbf{v}})$ to (3.17)-(3.18) will furnish a sequence of iterates verifying (3.16) in \mathcal{H} for the primal-dual solution $(\bar{\mathbf{x}}, \bar{\mathbf{v}})$ to Problem 1.1.

3.2 The case when $A + C$ and $B_i^{-1} + D_i^{-1}, i = 1, \dots, m$, are strongly monotone

Within this subsection we consider the case when $A + C$ is ρ -strongly monotone with $\rho \in \mathbb{R}_{++}$ and $B_i^{-1} + D_i^{-1}$ is τ_i -strongly monotone with $\tau_i \in \mathbb{R}_{++}$ for $i = 1, \dots, m$, and provide an accelerated version of the algorithm in Theorem 1.1 which generates sequences of primal and dual iterates that converge to the primal-dual solution to Problem 1.1 with an improved rate of convergence.

Algorithm 3.2. Let $x_0 \in \mathcal{H}, (v_{1,0}, \dots, v_{m,0}) \in \mathcal{G}$, and $\gamma \in (0, 1)$ such that

$$\gamma \leq \frac{1}{\sqrt{1 + 2 \min \{\rho, \tau_1, \dots, \tau_m\} \left(\sqrt{\sum_{i=1}^m \|L_i\|^2} + \max \{\mu, \nu_1, \dots, \nu_m\} \right)}}.$$

Consider the following updates:

$$(\forall n \geq 0) \begin{cases} p_{1,n} = J_{\gamma A} (x_n - \gamma (Cx_n + \sum_{i=1}^m L_i^* v_{i,n} - z)) \\ \text{For } i = 1, \dots, m \\ \left[\begin{array}{l} p_{2,i,n} = J_{\gamma B_i^{-1}} (v_{i,n} + \gamma (L_i x_n - D_i^{-1} v_{i,n} - r_i)) \\ v_{i,n+1} = \gamma L_i (p_{1,n} - x_n) + \gamma (D_i^{-1} v_{i,n} - D_i^{-1} p_{2,i,n}) + p_{2,i,n} \end{array} \right. \\ x_{n+1} = \gamma \sum_{i=1}^m L_i^* (v_{i,n} - p_{2,i,n}) + \gamma (Cx_n - Cp_{1,n}) + p_{1,n}. \end{cases} \quad (3.19)$$

Theorem 3.4. *In Problem 1.1 suppose that $A + C$ is ρ -strongly monotone with $\rho \in \mathbb{R}_{++}$, $B_i^{-1} + D_i^{-1}$ is τ_i -strongly monotone with $\tau_i \in \mathbb{R}_{++}$ for $i = 1, \dots, m$, and let $(\bar{x}, \bar{v}_1, \dots, \bar{v}_m) \in \mathcal{H} \times \mathcal{G}$ be a primal-dual solution to Problem 1.1. Then for every $n \geq 0$ it holds*

$$\|x_n - \bar{x}\|^2 + \sum_{i=1}^m \|v_{i,n} - \bar{v}_i\|^2 \leq \left(\frac{1}{1 + 2\rho_{\min} \gamma (1 - \gamma)} \right)^n \left(\|x_0 - \bar{x}\|^2 + \sum_{i=1}^m \|v_{i,0} - \bar{v}_i\|^2 \right),$$

where $\rho_{\min} = \min \{\rho, \tau_1, \dots, \tau_m\}$ and $x_n \in \mathcal{H}$ and $(v_{1,n}, \dots, v_{m,n}) \in \mathcal{G}$ are the iterates generated by Algorithm 3.2.

Proof. Taking into account the definitions of the resolvents occurring in Algorithm 3.2 we obtain for every $n \geq 0$

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

and

$$\frac{v_{i,n} - v_{i,n+1}}{\gamma} + L_i p_{1,n} - r_i \in (B_i^{-1} + D_i^{-1})p_{2,i,n}, \quad i = 1, \dots, m.$$

The primal-dual solution $(\bar{x}, \bar{v}_1, \dots, \bar{v}_m) \in \mathcal{H} \times \mathcal{G}$ to Problem 1.1 fulfills (see (1.4))

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

By the strong monotonicity of $A + C$ and $B_i^{-1} + D_i^{-1}$, $i = 1, \dots, m$, we obtain for every $n \geq 0$

$$\left\langle p_{1,n} - \bar{x}, \frac{x_n - x_{n+1}}{\gamma} - \sum_{i=1}^m L_i^* p_{2,i,n} + z - \left(z - \sum_{i=1}^m L_i^* \bar{v}_i \right) \right\rangle \geq \rho \|p_{1,n} - \bar{x}\|^2 \quad (3.20)$$

and, respectively,

$$\left\langle p_{2,i,n} - \bar{v}_i, \frac{v_{i,n} - v_{i,n+1}}{\gamma} + L_i p_{1,n} - r_i - (L_i \bar{x} - r_i) \right\rangle \geq \tau_i \|p_{2,i,n} - \bar{v}_i\|^2, \quad i = 1, \dots, m. \quad (3.21)$$

Consider the Hilbert space $\mathcal{H} = \mathcal{H} \times \mathcal{G}$, equipped with the inner product defined in (2.13) and associated norm, and set

$$\bar{\mathbf{x}} = (\bar{x}, \bar{v}_1, \dots, \bar{v}_m), \quad \mathbf{x}_n = (x_n, v_{1,n}, \dots, v_{m,n}), \quad \mathbf{p}_n = (p_{1,n}, p_{2,1,n}, \dots, p_{2,m,n}).$$

Summing up the inequalities (3.20) and (3.21) and using

$$\left\langle \mathbf{p}_n - \bar{\mathbf{x}}, \frac{\mathbf{x}_n - \mathbf{x}_{n+1}}{\gamma} \right\rangle = \frac{\|\mathbf{x}_{n+1} - \mathbf{p}_n\|^2}{2\gamma} - \frac{\|\mathbf{x}_n - \mathbf{p}_n\|^2}{2\gamma} + \frac{\|\mathbf{x}_n - \bar{\mathbf{x}}\|^2}{2\gamma} - \frac{\|\mathbf{x}_{n+1} - \bar{\mathbf{x}}\|^2}{2\gamma},$$

we obtain for every $n \geq 0$

$$\frac{\|\mathbf{x}_n - \bar{\mathbf{x}}\|^2}{2\gamma} \geq \rho_{\min} \|\mathbf{p}_n - \bar{\mathbf{x}}\|^2 + \frac{\|\mathbf{x}_{n+1} - \bar{\mathbf{x}}\|^2}{2\gamma} + \frac{\|\mathbf{x}_n - \mathbf{p}_n\|^2}{2\gamma} - \frac{\|\mathbf{x}_{n+1} - \mathbf{p}_n\|^2}{2\gamma}. \quad (3.22)$$

Further, using the estimate $2ab \leq \gamma a^2 + \frac{b^2}{\gamma}$ for all $a, b \in \mathbb{R}$, we obtain

$$\begin{aligned} \rho_{\min} \|\mathbf{p}_n - \bar{\mathbf{x}}\|^2 &\geq \frac{2\rho_{\min}\gamma(1-\gamma)}{2\gamma} \|\mathbf{x}_{n+1} - \bar{\mathbf{x}}\|^2 - \frac{2\rho_{\min}(1-\gamma)}{2\gamma} \|\mathbf{x}_{n+1} - \mathbf{p}_n\|^2 \\ &\geq \frac{2\rho_{\min}\gamma(1-\gamma)}{2\gamma} \|\mathbf{x}_{n+1} - \bar{\mathbf{x}}\|^2 - \frac{2\rho_{\min}}{2\gamma} \|\mathbf{x}_{n+1} - \mathbf{p}_n\|^2 \quad \forall n \geq 0. \end{aligned}$$

Hence, (3.22) reduces to

$$\begin{aligned} \frac{\|\mathbf{x}_n - \bar{\mathbf{x}}\|^2}{2\gamma} &\geq \frac{(1 + 2\rho_{\min}\gamma(1-\gamma))\|\mathbf{x}_{n+1} - \bar{\mathbf{x}}\|^2}{2\gamma} \\ &\quad + \frac{\|\mathbf{x}_n - \mathbf{p}_n\|^2}{2\gamma} - \frac{(1 + 2\rho_{\min})\|\mathbf{x}_{n+1} - \mathbf{p}_n\|^2}{2\gamma} \quad \forall n \geq 0. \end{aligned}$$

Using the same arguments as in (2.14), it is easy to check that for every $n \geq 0$

$$\begin{aligned} &\frac{\|\mathbf{x}_n - \mathbf{p}_n\|^2}{2\gamma} - \frac{(1 + 2\rho_{\min})\|\mathbf{x}_{n+1} - \mathbf{p}_n\|^2}{2\gamma} \\ &\geq \left(1 - (1 + 2\rho_{\min})\gamma^2 \left(\sqrt{\sum_{i=1}^m \|L_i\|^2} + \max\{\mu, \nu_1, \dots, \nu_m\} \right)^2 \right) \frac{\|\mathbf{x}_n - \mathbf{p}_n\|^2}{2\gamma} \\ &\geq 0, \end{aligned}$$

whereby the nonnegativity of this term is ensured by the assumption that

$$\gamma \leq \frac{1}{\sqrt{1 + 2\rho_{\min}} \left(\sqrt{\sum_{i=1}^m \|L_i\|^2} + \max\{\mu, \nu_1, \dots, \nu_m\} \right)}.$$

Therefore, we obtain

$$\|\mathbf{x}_n - \bar{\mathbf{x}}\|^2 \geq (1 + 2\rho_{\min}\gamma(1-\gamma))\|\mathbf{x}_{n+1} - \bar{\mathbf{x}}\|^2 \quad \forall n \geq 0,$$

which leads to

$$\|\mathbf{x}_n - \bar{\mathbf{x}}\|^2 \leq \left(\frac{1}{1 + 2\rho_{\min}\gamma(1-\gamma)} \right)^n \|\mathbf{x}_0 - \bar{\mathbf{x}}\|^2 \quad \forall n \geq 0.$$

□

4 Numerical experiments in imaging

In this section we test the feasibility of Algorithm 2.1 and of its accelerated version Algorithm 3.1 in the context of different problem formulations occurring in imaging and compare their performances to the ones of two other popular primal-dual algorithms introduced in [6]. For all applications discussed in this section the images have been normalized, in order to make their pixels range in the closed interval from 0 (pure black) to 1 (pure white).

4.1 TV-based image denoising

Our first numerical experiment aims the solving of an image denoising problem via total variation regularization. More precisely, we deal with the convex optimization problem

$$\inf_{x \in \mathbb{R}^n} \left\{ \lambda TV(x) + \frac{1}{2} \|x - b\|^2 \right\}, \quad (4.1)$$

where $\lambda \in \mathbb{R}_{++}$ is the regularization parameter, $TV : \mathbb{R}^n \rightarrow \mathbb{R}$ is a discrete total variation functional and $b \in \mathbb{R}^n$ is the observed noisy image.

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 $i = 1, \dots, M$ and $j = 1, \dots, N$. Two popular choices for the discrete total variation functional are the *isotropic total variation* $TV_{\text{iso}} : \mathbb{R}^n \rightarrow \mathbb{R}$,

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

and the *anisotropic total variation* $TV_{\text{aniso}} : \mathbb{R}^n \rightarrow \mathbb{R}$,

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

where in both cases reflexive (Neumann) boundary conditions are assumed.

We denote $\mathcal{Y} = \mathbb{R}^n \times \mathbb{R}^n$ and define the linear operator $L : \mathbb{R}^n \rightarrow \mathcal{Y}$, $x_{i,j} \mapsto (L_1 x_{i,j}, L_2 x_{i,j})$, 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. One can easily check that $\|L\|^2 \leq 8$ and that its adjoint $L^* : \mathcal{Y} \rightarrow \mathbb{R}^n$ is as easy to implement as the operator itself (cf. [5]).

	$\sigma = 0.12, \lambda = 0.07$		$\sigma = 0.06, \lambda = 0.035$	
	$\varepsilon = 10^{-4}$	$\varepsilon = 10^{-6}$	$\varepsilon = 10^{-4}$	$\varepsilon = 10^{-6}$
ALG1	350 (7.03 s)	2989 (59.82 s)	184 (3.69 s)	1454 (29.07 s)
ALG2	101 (2.28 s)	442 (9.91 s)	72 (1.62 s)	298 (6.68 s)
PD1	342 (3.59 s)	3133 (32.68 s)	180 (1.91 s)	1427 (14.87 s)
PD2	96 (1.02 s)	442 (4.67 s)	69 (0.76 s)	319 (3.39 s)

Table 4.1: Performance evaluation for the images in Figure 4.1. The entries represent to the number of iterations and the CPU times in seconds, respectively, needed in order to attain a root mean squared error for the iterates below the tolerance ε .

Within this example we will focus on the anisotropic total variation function which is nothing else than the composition of the l_1 -norm in \mathcal{Y} with the linear operator L . Due to the full splitting characteristics of the iterative methods presented in this paper, we need only to compute the proximal point of the conjugate of the l_1 -norm, the latter being the indicator function of the dual unit ball. Thus, the calculation of the proximal point will result in the computation of a projection, which has an easy implementation. The more challenging isotropic total variation functional is employed in the forthcoming subsection in the context of an image deblurring problem.

Thus, problem (4.1) reads equivalently

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

where $h : \mathbb{R}^n \rightarrow \mathbb{R}$, $h(x) = \frac{1}{2}\|x - b\|^2$, is 1-strongly monotone and differentiable with 1-Lipschitzian gradient and $g : \mathcal{Y} \rightarrow \mathbb{R}$ is defined as $g(y_1, y_2) = \lambda\|(y_1, y_2)\|_1$. Then its conjugate $g^* : \mathcal{Y} \rightarrow \overline{\mathbb{R}}$ is nothing else than

$$g^*(p_1, p_2) = (\lambda\|\cdot\|_1)^*(p_1, p_2) = \lambda \left\| \begin{pmatrix} p_1 \\ p_2 \end{pmatrix} \right\|_1^* = \delta_S(p_1, p_2),$$

where $S = [-\lambda, \lambda]^n \times [-\lambda, \lambda]^n$. Taking $x_0 \in \mathcal{H}$, $v_0 \in \mathcal{Y}$,

$$\gamma_0 \in \left(0, \min \left\{ 1, \frac{\sqrt{1+4\rho}}{2(1+2\rho)\mu} \right\} \right) \text{ and } \sigma_0 = \frac{1}{2\gamma_0(1+2\rho)\sum_{i=1}^m \|L_i\|^2},$$

Algorithm 3.1 looks for this particular problem like

$$(\forall n \geq 0) \begin{cases} p_{1,n} = x_n - \gamma_n(x_n - b + L^*v_n) \\ p_{2,n} = \mathcal{P}_S(v_n + \sigma_n Lx_n) \\ v_{n+1} = \sigma_n L(p_{1,n} - x_n) + p_{2,n} \\ x_{n+1} = \gamma_n L^*(v_n - p_{2,n}) + \gamma_n(x_n - p_{1,n}) + p_{1,n} \\ \theta_n = 1/\sqrt{1+2\rho\gamma_n(1-\gamma_n)}, \gamma_{n+1} = \theta_n \gamma_n, \sigma_{n+1} = \sigma_n/\theta_n. \end{cases}$$

However, we solved the regularized image denoising problem with Algorithm 2.1, the primal-dual iterative scheme from [6] (see, also, [8]) and the accelerated version of the latter presented in [6, Theorem 2], as well, and refer the reader to Table 4.1 for a comparison of the obtained results:



Figure 4.1: $TV-l_2$ image denoising. The noisy image in (a) was obtained after adding white Gaussian noise with standard deviation $\sigma = 0.06$ to the original 256×256 lichtenstein test image, while the output of Algorithm 3.1, for $\lambda = 0.035$, after 100 iterations is shown in (c). Likewise, the noise image when choosing $\sigma = 0.12$ and the output of the same algorithm, for $\lambda = 0.07$, after 100 iterations are shown in (b) and (d), respectively.

- ALG1: Algorithm 2.1 with $\gamma = \frac{1-\tilde{\varepsilon}}{\sqrt{8}}$, small $\tilde{\varepsilon} > 0$ and by taking the last iterate instead of the averaged sequence.
- ALG2: Algorithm 3.1 with $\rho = 0.3$, $\mu = 1$ and $\gamma_0 = \frac{\sqrt{1+4\rho}}{2(1+2\rho)\mu}$.
- PD1: Algorithm 1 in [6] with $\tau = \frac{1}{\sqrt{8}}$, $\tau\sigma_0 = 1$ and by taking the last iterate instead of the averaged sequence.
- PD2: Algorithm 2 in [6] with $\rho = 0.3$, $\tau_0 = \frac{1}{\sqrt{8}}$, $\tau_0\sigma_0 = 1$.

From the point of view of the number of iterations, one can notice similarities between both the primal-dual algorithms ALG1 and PD1 and the accelerated versions ALG2 and PD2. From this point of view they behave almost equal. When comparing the CPU times, it shows that the methods in this paper need almost twice amount of time. This is since ALG1 and ALG2 lead back to a forward-backward-forward splitting, whereas PD1 and PD2 rely on a forward-backward splitting scheme, meaning that ALG1 and ALG2 process the double amount of forward steps than PD1 and PD2. In this example the evaluation of forward steps (i.e. which constitute in matrix-vector multiplications involving the linear operators and their adjoints) is, compared with the calculation of projections when computing the resolvents, the most costly step.

4.2 TV -based image deblurring

The second numerical experiment that we consider concerns the solving of an extremely 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 $\bar{x} \in \mathbb{R}^n$ fulfilling

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

To this end we basically solve the following regularized convex nondifferentiable problem

$$\inf_{x \in \mathbb{R}^n} \left\{ \|Ax - b\|_1 + \lambda_1 TV_{\text{iso}}(x) + \lambda_2 \|x\|_1 + \delta_{[0,1]^n}(x) \right\}, \quad (4.2)$$

where $\lambda_1, \lambda_2 \in \mathbb{R}_{++}$ are regularization parameters and $TV_{\text{iso}} : \mathbb{R}^n \rightarrow \mathbb{R}$ is the discrete isotropic total variation function. Notice that none of the functions occurring in (4.2) is differentiable, while the regularization is done by a combination of two regularization functionals with different properties.

The blurring operator is constructed by making use of the Matlab routines `imfilter` and `fspecial` as follows:

```

1 H=fspecial('gaussian',9,4); % gaussian blur of size 9 times 9
2                               % and standard deviation 4
3 B=imfilter(X,H,'conv','symmetric'); % B=observed blurred image
4                               % X=original image

```

The function `fspecial` returns a rotationally symmetric Gaussian lowpass filter of size 9×9 with standard deviation 4, the entries of H being nonnegative and their sum adding up to 1. The function `imfilter` convolves the filter H with the image X and furnishes the blurred image B . The boundary option “symmetric” corresponds to reflexive boundary conditions. Thanks to the rotationally symmetric filter H , the linear operator A defined via the routine `imfilter` is symmetric, too. By making use of the real spectral decomposition of A , it shows that $\|A\|^2 = 1$.

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 $TV_{\text{iso}}(x) = \|Lx\|_{\times}$, where L is the linear operator defined in the previous section. The conjugate function $(\|\cdot\|_{\times})^* : \mathcal{Y} \rightarrow \overline{\mathbb{R}}$ of $\|\cdot\|_{\times}$ is for every $(p, q) \in \mathcal{Y}$ given by (see, for instance, [3])

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

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 (4.2) can be written in the form of

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

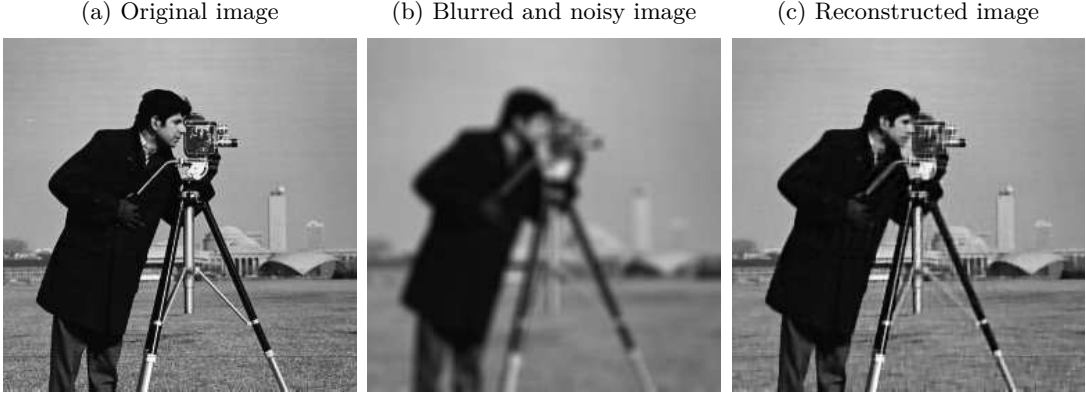


Figure 4.2: $TV-l_1-l_1$ image deblurring. Figure (a) shows the clean 256×256 cameraman test image, (b) shows the image obtained after multiplying it with a blur operator and adding white Gaussian noise and (c) shows the averaged sequence generated by Algorithm 2.1 after 400 iterations.

where $f : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$, $f(x) = \lambda_2 \|x\|_1 + \delta_{[0,1]^n}(x)$, $g_1 : \mathbb{R}^n \rightarrow \mathbb{R}$, $g_1(y) = \|y - b\|_1$ and $g_2 : \mathcal{Y} \rightarrow \mathbb{R}$, $g_2(y, z) = \lambda_1 \|(y, z)\|_\times$. For every $p \in \mathbb{R}^n$ it holds $g_1^*(p) = \delta_{[-1,1]^n}(p) + p^T b$ (see, for instance, [2]), while for any $(p, q) \in \mathcal{Y}$ we have $g_2^*(p, q) = \delta_S(p, q)$, with $S = \{(p, q) \in \mathcal{Y} : \|(p, q)\|_{\times*} \leq \lambda_1\}$. We solved this problem by Algorithm 2.1 and to this end we made use of the following formulae for the proximal points involved in the formulation of this iterative scheme:

$$\begin{aligned} \text{Prox}_{\gamma f}(x) &= \arg \min_{z \in [0,1]^n} \left\{ \gamma \lambda_2 \|z\|_1 + \frac{1}{2} \|z - x\|^2 \right\} = \mathcal{P}_{[0,1]^n}(x - \gamma \lambda_2 \mathbb{1}^n) \quad \forall x \in \mathbb{R}^n, \\ \text{Prox}_{\gamma g_1^*}(p) &= \arg \min_{z \in [-1,1]^n} \left\{ \gamma \langle z, b \rangle + \frac{1}{2} \|z - p\|^2 \right\} = \mathcal{P}_{[-1,1]^n}(p - \gamma b) \quad \forall p \in \mathbb{R}^n, \end{aligned}$$

and

$$\text{Prox}_{\gamma g_2^*}(p, q) = \arg \min_{(y,z) \in S} \frac{1}{2} \|(y, z) - (p, q)\|^2 = \mathcal{P}_S(p, q) \quad \forall (p, q) \in \mathcal{Y},$$

where $\gamma \in \mathbb{R}_{++}$, $\mathbb{1}^n$ is the vector in \mathbb{R}^n with all entries equal to 1 and the projection operator $\mathcal{P}_S : \mathcal{Y} \rightarrow S$ is defined as

$$(p_{i,j}, q_{i,j}) \mapsto \left(\frac{p_{i,j}}{\max \left\{ 1, \frac{\sqrt{p_{i,j}^2 + q_{i,j}^2}}{\lambda_1} \right\}}, \frac{q_{i,j}}{\max \left\{ 1, \frac{\sqrt{p_{i,j}^2 + q_{i,j}^2}}{\lambda_1} \right\}} \right).$$

Taking $x_0 \in \mathcal{H}$, $(v_{1,0}, v_{2,0}) \in \mathbb{R}^n \times \mathcal{Y}$, $\beta = \sqrt{1+8} = 3$, $\varepsilon \in \left(0, \frac{1}{\beta+1}\right)$ and $(\gamma_n)_{n \geq 0}$ a

nondecreasing sequence in $[\varepsilon, \frac{1-\varepsilon}{\beta}]$, Algorithm 2.1 looks for this particular problem like

$$(\forall n \geq 0) \quad \begin{cases} p_{1,n} = \mathcal{P}_{[0,1]^n} (x_n - \gamma_n (A^* v_{1,n} + L^* v_{2,n} + \lambda_2 \mathbb{1}^n)) \\ p_{2,1,n} = \mathcal{P}_{[-1,1]^n} (v_{1,n} + \gamma_n (Ax_n - b)) \\ p_{2,2,n} = \mathcal{P}_S (v_{2,n} + \gamma_n Lx_n) \\ v_{1,n+1} = \gamma_n A(p_{1,n} - x_n) + p_{2,1,n} \\ v_{2,n+1} = \gamma_n L(p_{1,n} - x_n) + p_{2,2,n} \\ x_{n+1} = \gamma_n (A^*(v_{1,n} - p_{2,1,n}) + L^*(v_{2,n} - p_{2,2,n})) + p_{1,n}. \end{cases}$$

Figure 4.2 shows the original cameraman test image, which is part of the image processing toolbox in Matlab, the image obtained after multiplying it with the blur operator and adding after that normally distributed white Gaussian noise with standard deviation 10^{-3} and the image reconstructed by Algorithm 2.1 when taking as regularization parameters $\lambda_1 = 3e-3$ and $\lambda_2 = 2e-5$.

4.3 TV-based image inpainting

In the last section of the paper we show how image inpainting problems, which aim for recovering lost information, can be efficiently solved via the primal-dual methods investigated in this work. To this end, we consider the following TV - l_1 model

$$\inf_{x \in \mathbb{R}^n} \left\{ \lambda TV_{\text{iso}}(x) + \|Kx - b\|_1 + \delta_{[0,1]^n}(x) \right\}, \quad (4.3)$$

where $\lambda \in \mathbb{R}_{++}$ is the regularization parameter and $TV_{\text{iso}} : \mathbb{R}^n \rightarrow \mathbb{R}$ is the isotropic total variation functional and $K \in \mathbb{R}^{n \times n}$ is the diagonal matrix, where for $i = 1, \dots, n$, $K_{i,i} = 0$, if the pixel i in the noisy image $b \in \mathbb{R}^n$ is lost (in our case pure black) and $K_{i,i} = 1$, otherwise. The induced linear operator $K : \mathbb{R}^n \rightarrow \mathbb{R}^n$ fulfills $\|K\| = 1$, while, in the light of the considerations made in the previous two subsections, we have that $TV_{\text{iso}}(x) = \|Lx\|_\infty$ for all $x \in \mathbb{R}^n$.

Thus, problem (4.3) can be formulated as

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

where $f : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$, $f(x) = \delta_{[0,1]^n}$, $g_1 : \mathcal{Y} \rightarrow \mathbb{R}$, $g_1(y_1, y_2) = \|(y_1, y_2)\|_\infty$ and $g_2 : \mathbb{R}^n \rightarrow \mathbb{R}$, $g_2(y) = \|y - b\|_1$. We solve it by Algorithm 2.1, the formulae for the proximal points involved in this iterative scheme been already given in Subsection 4.2. Figure 4.3 shows the original fruit image, the image obtained from it after setting to pure black 80% randomly chosen pixels and the image reconstructed by Algorithm 2.1 when taking as regularization parameter $\lambda = 0.05$.

References

- [1] H.H. Bauschke and P.L. Combettes. *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. CMS Books in Mathematics, Springer, 2011.
- [2] R.I. Boţ. *Conjugate Duality in Convex Optimization*. Lecture Notes in Economics and Mathematical Systems, Vol. 637, Springer-Verlag Berlin Heidelberg, 2010.

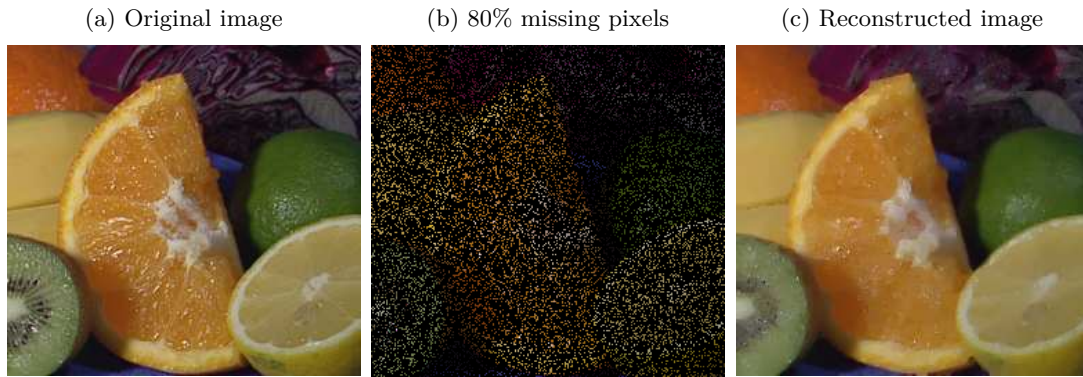


Figure 4.3: $TV-l_1$ image inpainting. Figure (a) shows the 240×256 clean fruits image, (b) shows the same image for which 80% randomly chosen pixels were set to pure black and (c) shows the nonaveraged iterate generated by Algorithm 2.1 after 200 iterations.

- [3] R.I. Boţ, S.M. Grad and G. Wanka. *Duality in Vector Optimization*. Springer-Verlag Berlin Heidelberg, 2009.
- [4] L.M. Briceño-Arias and P.L. Combettes. A monotone + skew splitting model for composite monotone inclusions in duality. *SIAM Journal on Optimization* 21(4):1230–1250, 2011.
- [5] A. Chambolle. An algorithm for total variation minimization and applications. *Journal of Mathematical Imaging and Vision* 20(1–2):89–97, 2004.
- [6] A. Chambolle and 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.
- [7] P.L. Combettes and J.-C. Pesquet. Primal-dual splitting algorithm for solving inclusions with mixtures of composite, Lipschitzian, and parallel-sum type monotone operators. *Set-Valued and Variational Analysis* 20(2):307–330, 2012.
- [8] B.C. Vũ. A splitting algorithm for dual monotone inclusions involving cocoercive operators. *Advances in Computational Mathematics*, 2011. <http://dx.doi.org/10.1007/s10444-011-9254-8>
- [9] C. Zălinescu. *Convex Analysis in General Vector Spaces*. World Scientific, 2002.