Solving monotone inclusions involving parallel sums of linearly composed maximally monotone operators

Radu Ioan Boţ [1] Christopher Hendrich [2]

April 28, 2016

Abstract. The aim of this article is to present two different primal-dual methods for solving structured monotone inclusions involving parallel sums of compositions of maximally monotone operators with linear bounded operators. By employing some elaborated splitting techniques, all of the operators occurring in the problem formulation are processed individually via forward or backward steps. The treatment of parallel sums of linearly composed maximally monotone operators is motivated by applications in imaging which involve first- and second-order total variation functionals, to which a special attention is given.

Keywords. convex optimization, Fenchel duality, infimal convolution, monotone inclusion, parallel sum, primal-dual algorithm

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

1 Introduction

In applied mathematics, a wide variety of convex optimization problems such as single- or multifacility location problems, support vector machine problems for classification and regression, problems in clustering and portfolio optimization as well as signal and image processing problems, all of them potentially possessing nonsmooth terms in their objectives, can be reduced to the solving of inclusion problems involving mixtures of monotone set-valued operators. Therefore, the solving of monotone inclusion problems involving maximally monotone operators continues to be one of the most attractive branches of research (see [1, 3, 5, 6, 9, 12, 13, 15–24, 26–29]).

1.1 Motivation

The problem formulation we consider in this article is inspired by a real-world application in image denoising, where first- and second-order total variation functionals are linked via infimal convolutions in order to reduce staircasing effects in the reconstructed images.

Let $b \in \mathbb{R}^n$ be the observed and vectorized noisy image of size $M \times N$ (with $n = MN$ for greyscale and $n = 3MN$ for colored images). For $k \in \mathbb{N}$ and $\omega = (\omega_1, \ldots, \omega_k) \in \mathbb{R}^k_{++}$

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

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

we consider on $\mathbb{R}^{k\times n}$ the following norm defined for $y = (y_1, \ldots, y_k)^T \in \mathbb{R}^{k\times n}$ as

$$\|y\|_{1,\omega} = \left\| (\omega_1 y_1^2 + \ldots + \omega_k y_k^2)^{\frac{1}{2}} \right\|_1,$$

where addition, multiplication and square root of vectors are understood to be componentwise. Further, we consider the forward difference matrix

$$D_k := \begin{bmatrix} -1 & 1 & 0 & 0 & \cdots & 0 \\ 0 & -1 & 1 & 0 & \cdots & 0 \\ \vdots & & \ddots & \ddots & & \vdots \\ 0 & \cdots & 0 & -1 & 1 & 0 \\ 0 & \cdots & 0 & 0 & -1 & 1 \\ 0 & \cdots & 0 & 0 & 0 & 0 \end{bmatrix} \in \mathbb{R}^{k\times k},$$

which models the discrete first-order derivative. Note that $-D_k^T D_k$ is then an approximation of the second-order derivative. We denote by $A \otimes B$ the Kronecker product of the matrices A and B and define

$$D_x = I_N \otimes D_M,\ D_y = D_N \otimes I_M \text{ and } \mathcal{D}_1 = \begin{bmatrix} D_x \\ D_y \end{bmatrix}, \tag{1.1}$$

where D_x and D_y represent the vertical and horizontal difference operators, respectively, and I_N and I_M are the identity matrices of sizes N and M, respectively. Further, we define the discrete second-order derivatives matrices

$$D_{xx} = I_N \otimes (-D_M^T D_M),\ D_{yy} = (-D_N^T D_N) \otimes I_M,\ \mathcal{D}_2 = \begin{bmatrix} D_{xx} \\ D_{yy} \end{bmatrix} \tag{1.2}$$

and

$$L_1 = \begin{bmatrix} -D_x^T & 0 \\ 0 & -D_y^T \end{bmatrix}$$

and notice that $\mathcal{D}_2 = L_1 \mathcal{D}_1$. This approach was initially proposed in [14] and further investigated in [27]. We refer the reader to [27] for other discrete second-order derivatives involving also mixed partial derivatives (in horizontal-vertical direction and vice versa).

The reconstructed image is obtained by solving one of the following convex optimization problems (see [27, Example 2.2 and Example 3.1])

$$(\ell_2^2\text{-IC/P}) \quad \inf_{x\in\mathbb{R}^n} \left\{ \frac{1}{2}\|x-b\|^2 + \Big((\alpha_1 \|\cdot\|_{1,\omega_1} \circ \mathcal{D}_1) \,\square\, (\alpha_2 \|\cdot\|_{1,\omega_2} \circ \mathcal{D}_2) \Big)(x) \right\} \tag{1.3}$$

and

$$(\ell_2^2\text{-MIC/P}) \quad \inf_{x\in\mathbb{R}^n} \left\{ \frac{1}{2}\|x-b\|^2 + \Big((\alpha_1 \|\cdot\|_{1,\omega_1}) \,\square\, (\alpha_2 \|\cdot\|_{1,\omega_2} \circ L_1) \Big)(\mathcal{D}_1 x) \right\}, \tag{1.4}$$

respectively, where $\alpha_1, \alpha_2 \in \mathbb{R}_{++}$ are the regularization parameters and the regularizers correspond to anistropic total variation functionals.

This is the reason why we are going to treat the following more general primal-dual pair of complexly structured convex optimization problems.

Problem 1.1. Let $\mathcal{H}$ be a real Hilbert space, $z \in \mathcal{H}$ and $f,\ h \in \Gamma(\mathcal{H})$ such that h is differentiable with μ-Lipschitzian gradient for $\mu \in \mathbb{R}_{++}$. Furthermore, for every $i = 1, \ldots, m$, let $\mathcal{G}_i$, $\mathcal{X}_i$, $\mathcal{Y}_i$ be real Hilbert spaces, $r_i \in \mathcal{G}_i$, let $g_i \in \Gamma(\mathcal{X}_i)$ and $l_i \in \Gamma(\mathcal{Y}_i)$

and consider the nonzero linear bounded operators $L_i : \mathcal{H} \to \mathcal{G}_i$, $K_i : \mathcal{G}_i \to \mathcal{X}_i$ and $M_i : \mathcal{G}_i \to \mathcal{Y}_i$. We want to solve the primal optimization problem

$$\inf_{x\in\mathcal{H}} \left\{ f(x) + \sum_{i=1}^{m} \Big((g_i \circ K_i) \,\Box\, (l_i \circ M_i) \Big)(L_i x - r_i) + h(x) - \langle x, z\rangle \right\} \tag{1.5}$$

together with its conjugate dual problem

$$\sup_{\substack{(\boldsymbol{p},\boldsymbol{q})\in\boldsymbol{\mathcal{X}}\oplus\boldsymbol{\mathcal{Y}},\\ K_i^* p_i = M_i^* q_i,\ i=1,\ldots,m}} \left\{ -(f^* \,\Box\, h^*)\left(z - \sum_{i=1}^{m} L_i^* K_i^* p_i \right) - \sum_{i=1}^{m} \Big[g_i^*(p_i) + l_i^*(q_i) + \langle p_i, K_i r_i\rangle \Big] \right\}. \tag{1.6}$$

By $\mathbb{R}_{++}$ we denote the set of strictly positive real numbers and by $\mathbb{R}_+ := \mathbb{R}_{++} \cup \{0\}$. For a function $f : \mathcal{H} \to \overline{\mathbb{R}} := \mathbb{R} \cup \{\pm\infty\}$, where $\mathcal{H}$ is a real Hilbert space, we denote by $\operatorname{dom} f := \{x \in \mathcal{H} : f(x) < +\infty\}$ its *effective domain* and call f *proper*, if $\operatorname{dom} f \neq \varnothing$ and $f(x) > -\infty$ for all $x \in \mathcal{H}$. We let

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

The *conjugate function* of f is $f^* : \mathcal{H} \to \overline{\mathbb{R}}$, $f^*(p) = \sup\{\langle p, x\rangle - f(x) : x \in \mathcal{H}\}$ for all $p \in \mathcal{H}$, and, if $f \in \Gamma(\mathcal{H})$, then $f^* \in \Gamma(\mathcal{H})$, as well. For a linear bounded operator $L : \mathcal{H} \to \mathcal{G}$, where $\mathcal{G}$ is another real Hilbert space, the operator $L^* : \mathcal{G} \to \mathcal{H}$ defined via $\langle Lx, y\rangle = \langle x, L^* y\rangle$ for all $x \in \mathcal{H}$ and all $y \in \mathcal{G}$ denotes its *adjoint*.

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

In order to solve the primal-dual pair of optimization problems (1.5)-(1.6), we will actually solve the corresponding system of optimality conditions (see (3.2)), which is nothing else than a system of monotone inclusions with a complex and intricate structure. This motivates the investigation of the following primal-dual pair of monotone inclusion problems.

Problem 1.2. Let $\mathcal{H}$ be a real Hilbert space, $z \in \mathcal{H}$, $A : \mathcal{H} \to 2^{\mathcal{H}}$ a maximally monotone operator and $C : \mathcal{H} \to \mathcal{H}$ a monotone μ^{-1}-cocoercive operator for $\mu \in \mathbb{R}_{++}$. Furthermore, for every $i = 1, \ldots, m$, let $\mathcal{G}_i$, $\mathcal{X}_i$, $\mathcal{Y}_i$ be real Hilbert spaces, $r_i \in \mathcal{G}_i$, $B_i : \mathcal{X}_i \to 2^{\mathcal{X}_i}$ and $D_i : \mathcal{Y}_i \to 2^{\mathcal{Y}_i}$ be maximally monotone operators and consider the nonzero linear bounded operators $L_i : \mathcal{H} \to \mathcal{G}_i$, $K_i : \mathcal{G}_i \to \mathcal{X}_i$ and $M_i : \mathcal{G}_i \to \mathcal{Y}_i$. We want to solve the primal inclusion

$$\text{find } \overline{x} \in \mathcal{H} \text{ such that } z \in A\overline{x} + \sum_{i=1}^{m} L_i^* \Big((K_i^* \circ B_i \circ K_i) \,\Box\, (M_i^* \circ D_i \circ M_i) \Big)(L_i \overline{x} - r_i) + C\overline{x} \tag{1.7}$$

together with its dual inclusion

$$\text{find } \begin{cases} \overline{p}_i \in \mathcal{X}_i, i = 1, ..., m, \\ \overline{q}_i \in \mathcal{Y}_i, i = 1, ..., m, \\ \overline{y}_i \in \mathcal{G}_i, i = 1, ..., m, \end{cases} \text{such that } \exists x \in \mathcal{H} : \begin{cases} z - \sum_{i=1}^{m} L_i^* K_i^* \overline{p}_i \in Ax + Cx, \\ K_i(L_i x - \overline{y}_i - r_i) \in B_i^{-1} \overline{p}_i, i = 1, ..., m, \\ M_i \overline{y}_i \in D_i^{-1} \overline{q}_i, i = 1, ..., m, \\ K_i^* \overline{p}_i = M_i^* \overline{q}_i, i = 1, ..., m. \end{cases} \tag{1.8}$$

We propose in this paper two iterative methods of forward-backward and forward-backward-forward type, respectively, for solving this primal-dual pair of monotone inclusion problems and investigate their asymptotic behavior. The two methods share the common feature to reduce the solving of the primal-dual pair of monotone inclusions to the determination of the set of the zeros of the sum of two maximally monotone operators defined on an suitable product space endowed with a topology that is synchronized with the problem. Depending on the nature of the two operators we employ the forward-backward algorithm (see [2]) or Tseng's forward-backward-forward algorithm (see [28]) and obtain easily implementable iterative schemes. These have the property that each of the operators arising in the formulation of the monotone inclusion problem (1.7) is evaluated separately. More precisely, the set-valued operators are evaluated via their resolvents, called backward steps, while the single-valued ones are accessed via explicit forward steps. A forward-backward-forward algorithm for solving the primal-dual pair of monotone inclusions (1.7) - (1.8), in the particular situation when L_i is the identity operator and $r_i = 0$ for any $i = 1, ..., m$, has been recently investigated in [3]. However, since it makes a forward step less, the forward-backward method is naturally more attractive from the perspective numerical implementations. This phenomenon is supported by our experimental results reported in Section 4.

After the appearance of the proximal point algorithm for approximating the set of zeros of a maximal monotone operator defined on a Hilbert space (see [26]), the attention of the community was drawn to iterative schemes for determining the zeros of the sum of two maximally monotone operators, due to the role played by these schemes in the minimization of the sum of two convex functions. To the most classical methods of this type belongs the Douglas-Rachford splitting algorithm (see [22]), which has the property that at each iteration the operators are processed separately via their resolvents. Of equal importance are methods designed to determine the zeros of the sum of a single-valued monotone operator and a maximally monotone operator, like the forward-backward [2] and Tseng's forward-backward-forward [28] algorithms, which evaluate the single-valued operator via a forward step and the set-valued one via its resolvent.

In the last years, motivated by different applications, the complexity of the monotone inclusion problems increased, by including sums of maximally monotone operators composed with linear bounded operators (see [13, 15]), (single-valued) Lipschitzian or cocoercive monotone operators and parallel sums of maximally monotone operators (see [3, 5, 12, 19–21, 29]). For some of these iterative schemes, under strong monotonicity assumptions accelerated versions have been provided (see [6, 9, 15]).

The article is organized as follows. In the remaining of this section we introduce notations and preliminary results in convex analysis and monotone operator theory. In Section 2 we formulate the two algorithms and study their convergence behavior. In Section 3 we employ the outcomes of the previous one to the simultaneously solving of convex minimization problems and their conjugate dual problems. Numerical experiments in the context of image denoising problems with first- and second-order total variation functionals are made in Section 4.

1.2 Notation and preliminaries

We are considering the real Hilbert space $\mathcal{H}$ endowed 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. Having the sequences $(x_n)_{n\geq 0}$ and $(y_n)_{n\geq 0}$ in $\mathcal{H}$, we mind errors in the implementation of the algorithm by using the following notation taken from [3]

$$(x_n \approx y_n \ \forall n \geq 0) \Leftrightarrow \sum_{n\geq 0} \|x_n - y_n\| < +\infty. \tag{1.9}$$

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

Let $\mu > 0$ be arbitrary. A single-valued operator $M : \mathcal{H} \to \mathcal{H}$ is said to be *μ-cocoercive* if $\langle x-y, Mx-My\rangle \geq \mu\|Mx-My\|^2$ for all $(x,y) \in \mathcal{H} \times \mathcal{H}$. Moreover, M is *μ-Lipschitzian* if $\|Mx - My\| \leq \mu\|x-y\|$ for all $(x,y) \in \mathcal{H} \times \mathcal{H}$. A linear bounded operator $M : \mathcal{H} \to \mathcal{H}$ is said to be *self-adjoint*, if $M = M^*$ and *skew*, if $M^* = -M$.

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

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

respectively. If M_1 and M_2 are monotone, then $M_1 + M_2$ and $M_1 \,\Box\, M_2$ are monotone, too. However, if M_1 and M_2 are maximally monotone, this property is in general not true neither for $M_1 + M_2$ nor for $M_1 \,\Box\, M_2$, unless some qualification conditions are fulfilled (see [2, 4, 30]).

The *resolvent* of an operator $M : \mathcal{H} \to 2^{\mathcal{H}}$ is

$$J_M = (\mathrm{Id} + M)^{-1},$$

where the operator Id denotes the *identity* on $\mathcal{H}$. When M is maximally monotone, its resolvent is a single-valued firmly nonexpansive operator and, by [2, Proposition 23.18], we have for $\gamma \in \mathbb{R}_{++}$

$$\mathrm{Id} = J_{\gamma M} + \gamma J_{\gamma^{-1}M^{-1}} \circ \gamma^{-1}\mathrm{Id}. \tag{1.10}$$

Moreover, for $f \in \Gamma(\mathcal{H})$ and $\gamma \in \mathbb{R}_{++}$, the subdifferential $\partial(\gamma f)$ is maximally monotone (cf. [25]) and it holds $J_{\gamma\partial f} = (\mathrm{Id} + \gamma\partial f)^{-1} = \mathrm{Prox}_{\gamma f}$. Recall that the *(convex) subdifferential* of $f : \mathcal{H} \to \overline{\mathbb{R}}$ at $x \in \mathcal{H}$ is the set $\partial f(x) = \{p \in \mathcal{H} : f(y) - f(x) \geq \langle p, y-x\rangle \ \ \forall y \in \mathcal{H}\}$, if $f(x) \in \mathbb{R}$, and is taken to be the empty set, otherwise. Furthermore, $\mathrm{Prox}_{\gamma f}(x)$ denotes the *proximal point* of γf at $x \in \mathcal{H}$, representing the unique optimal solution of the optimization problem

$$\inf_{y\in\mathcal{H}} \left\{\gamma f(y) + \frac{1}{2}\|y-x\|^2\right\}. \tag{1.11}$$

In this particular situation, relation (1.10) becomes *Moreau's decomposition formula*

$$\mathrm{Id} = \mathrm{Prox}_{\gamma f} + \gamma\,\mathrm{Prox}_{\gamma^{-1}f^*} \circ \gamma^{-1}\mathrm{Id}. \tag{1.12}$$

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

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

which is the *projection* of x on Ω.

Finally, when for $i = 1, \dots, m$ the real Hilbert spaces $\mathcal{H}_i$ are endowed with inner product $\langle \cdot, \cdot \rangle_{\mathcal{H}_i}$ and associated norm $\|\cdot\|_{\mathcal{H}_i} = \sqrt{\langle \cdot, \cdot \rangle_{\mathcal{H}_i}}$, we denote by

$$\boldsymbol{\mathcal{H}} = \mathcal{H}_1 \oplus \ldots \oplus \mathcal{H}_m$$

their direct sum. For $\boldsymbol{v} = (v_1, \dots, v_m)$, $\boldsymbol{q} = (q_1, \dots, q_m) \in \boldsymbol{\mathcal{H}}$, this real Hilbert space is endowed with inner product and associated norm defined via

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

2 The primal-dual iterative schemes

Within this section we provide two different algorithms for solving the primal-dual inclusions introduced in Problem 1.2 and discuss their asymptotic convergence. In Subsection 2.2, however, the assumptions imposed on the monotone operator $C : \mathcal{H} \to \mathcal{H}$ are weakened by assuming that C is only μ-Lipschitz continuous for some $\mu \in \mathbb{R}_{++}$.

In the following, we let

$$\boldsymbol{\mathcal{X}} = \mathcal{X}_1 \oplus \ldots \oplus \mathcal{X}_m, \ \boldsymbol{\mathcal{Y}} = \mathcal{Y}_1 \oplus \ldots \oplus \mathcal{Y}_m, \ \boldsymbol{\mathcal{G}} = \mathcal{G}_1 \oplus \ldots \oplus \mathcal{G}_m$$

and

$$\boldsymbol{p} = (p_1, \dots, p_m) \in \boldsymbol{\mathcal{X}}, \ \boldsymbol{q} = (q_1, \dots, q_m) \in \boldsymbol{\mathcal{Y}}, \ \boldsymbol{y} = (y_1, \dots, y_m) \in \boldsymbol{\mathcal{G}}.$$

We say that $(\overline{x}, \overline{\boldsymbol{p}}, \overline{\boldsymbol{q}}, \overline{\boldsymbol{y}}) \in \mathcal{H} \oplus \boldsymbol{\mathcal{X}} \oplus \boldsymbol{\mathcal{Y}} \oplus \boldsymbol{\mathcal{G}}$ is a primal-dual solution to Problem 1.2, if

$$\begin{aligned} & z - \sum_{i=1}^{m} L_i^* K_i^* \overline{p}_i \in A\overline{x} + C\overline{x} \text{ and} \\ & K_i(L_i\overline{x} - \overline{y}_i - r_i) \in B_i^{-1}\overline{p}_i, \ M_i\overline{y}_i \in D_i^{-1}\overline{q}_i, \ K_i^*\overline{p}_i = M_i^*\overline{q}_i, \ i = 1, \dots, m. \end{aligned} \tag{2.1}$$

If $(\overline{x}, \overline{\boldsymbol{p}}, \overline{\boldsymbol{q}}, \overline{\boldsymbol{y}}) \in \mathcal{H} \oplus \boldsymbol{\mathcal{X}} \oplus \boldsymbol{\mathcal{Y}} \oplus \boldsymbol{\mathcal{G}}$ is a primal-dual solution to Problem 1.2, then $\overline{x}$ is a solution to (1.7) and $(\overline{\boldsymbol{p}}, \overline{\boldsymbol{q}}, \overline{\boldsymbol{y}})$ is a solution to (1.8). Notice also that

$$\begin{aligned} & \overline{x} \text{ solves (1.7)} \Leftrightarrow z \in A\overline{x} + \sum_{i=1}^{m} L_i^* \Big((K_i^* \circ B_i \circ K_i) \,\square\, (M_i^* \circ D_i \circ M_i) \Big)(L_i\overline{x} - r_i) + C\overline{x} \\ & \Leftrightarrow \exists\, \overline{\boldsymbol{v}} \in \boldsymbol{\mathcal{G}} \text{ such that} \begin{cases} z - \sum_{i=1}^{m} L_i^* \overline{v}_i \in A\overline{x} + C\overline{x}, \\ L_i\overline{x} - r_i \in \big(K_i^* \circ B_i \circ K_i\big)^{-1}(\overline{v}_i) + \big(M_i^* \circ D_i \circ M_i\big)^{-1}(\overline{v}_i), \\ i = 1, \dots, m \end{cases} \\ & \Leftrightarrow \exists\, (\overline{\boldsymbol{v}}, \overline{\boldsymbol{y}}) \in \boldsymbol{\mathcal{G}} \oplus \boldsymbol{\mathcal{G}} \text{ such that} \begin{cases} z - \sum_{i=1}^{m} L_i^* \overline{v}_i \in A\overline{x} + C\overline{x}, \\ \overline{v}_i \in \big(K_i^* \circ B_i \circ K_i\big)(L_i\overline{x} - \overline{y}_i - r_i), \ i = 1, \dots, m, \\ \overline{v}_i \in \big(M_i^* \circ D_i \circ M_i\big)(\overline{y}_i), \ i = 1, \dots, m \end{cases} \\ & \Leftrightarrow \exists\, (\overline{\boldsymbol{p}}, \overline{\boldsymbol{q}}, \overline{\boldsymbol{y}}) \in \boldsymbol{\mathcal{X}} \oplus \boldsymbol{\mathcal{Y}} \oplus \boldsymbol{\mathcal{G}} \text{ such that} \begin{cases} z - \sum_{i=1}^{m} L_i^* K_i^* \overline{p}_i \in A\overline{x} + C\overline{x}, \\ \overline{p}_i \in \big(B_i \circ K_i\big)(L_i\overline{x} - \overline{y}_i - r_i), \ i = 1, \dots, m, \\ \overline{q}_i \in \big(D_i \circ M_i\big)(\overline{y}_i), \ i = 1, \dots, m, \\ K_i^*\overline{p}_i = M_i^*\overline{q}_i, \ i = 1, \dots, m \end{cases} \end{aligned}$$

$$\Leftrightarrow \exists\, (\overline{\boldsymbol{p}}, \overline{\boldsymbol{q}}, \overline{\boldsymbol{y}}) \in \boldsymbol{\mathcal{X}} \oplus \boldsymbol{\mathcal{Y}} \oplus \boldsymbol{\mathcal{G}} \text{ such that } \begin{cases} z - \sum_{i=1}^m L_i^* K_i^* \overline{p}_i \in A\overline{x} + C\overline{x}, \\ K_i(L_i\overline{x} - \overline{y}_i - r_i) \in B_i^{-1}\overline{p}_i,\ i = 1, \ldots, m, \\ M_i\overline{y}_i \in D_i^{-1}\overline{q}_i,\ \ i = 1, \ldots, m, \\ K_i^*\overline{p}_i = M_i^*\overline{q}_i,\ \ i = 1, \ldots, m \end{cases} \tag{2.2}$$

Thus, if $\overline{x}$ is a solution to (1.7), then there exists $(\overline{\boldsymbol{p}}, \overline{\boldsymbol{q}}, \overline{\boldsymbol{y}}) \in \boldsymbol{\mathcal{X}} \oplus \boldsymbol{\mathcal{Y}} \oplus \boldsymbol{\mathcal{G}}$ such that $(\overline{x}, \overline{\boldsymbol{p}}, \overline{\boldsymbol{q}}, \overline{\boldsymbol{y}})$ is a primal-dual solution to Problem 1.2 and if $(\overline{\boldsymbol{p}}, \overline{\boldsymbol{q}}, \overline{\boldsymbol{y}})$ is a solution to (1.8), then there exists $\overline{x} \in \mathcal{H}$ such that $(\overline{x}, \overline{\boldsymbol{p}}, \overline{\boldsymbol{q}}, \overline{\boldsymbol{y}})$ is a primal-dual solution to Problem 1.2.

Remark 2.1. The notations (1.9) have been introduced in order to allow errors in the implementation of the algorithm, without affecting the readability of the paper in the sequel. This is reasonable since errors preserve their summability under addition, scalar multiplication and linear bounded mappings.

2.1 An algorithm of forward-backward type

In this subsection we propose a forward-backward type algorithm for solving Problem 1.2 and prove its convergence by showing that it can be reduced to an error-tolerant forward-backward iterative scheme.

Algorithm 2.1.
Let $x_0 \in \mathcal{H}$, and for any $i = 1, \ldots, m$, let $p_{i,0} \in \mathcal{X}_i$, $q_{i,0} \in \mathcal{Y}_i$ and $z_{i,0},\ y_{i,0},\ v_{i,0} \in \mathcal{G}_i$. For any $i = 1, \ldots, m$, let τ, $\theta_{1,i}$, $\theta_{2,i}$, $\gamma_{1,i}$, $\gamma_{2,i}$ and σ_i be strictly positive real numbers such that

$$2\mu^{-1}\left(1 - \overline{\alpha}\right) \min_{i=1,\ldots,m} \left\{ \frac{1}{\tau}, \frac{1}{\theta_{1,i}}, \frac{1}{\theta_{2,i}}, \frac{1}{\gamma_{1,i}}, \frac{1}{\gamma_{2,i}}, \frac{1}{\sigma_i} \right\} > 1, \tag{2.3}$$

for

$$\overline{\alpha} = \max\left\{ \sqrt{\tau \sum_{i=1}^m \sigma_i \|L_i\|^2},\ \max_{j=1,\ldots,m} \left\{ \sqrt{\theta_{1,j}\gamma_{1,j}\|K_j\|^2}, \sqrt{\theta_{2,j}\gamma_{2,j}\|M_j\|^2} \right\} \right\}.$$

Furthermore, let $\varepsilon \in (0,1)$, $(\lambda_n)_{n\geq 0}$ be a sequence in $[\varepsilon, 1]$ and set

$$(\forall n \geq 0) \quad \left\lfloor \begin{array}{l} \widetilde{x}_n \approx J_{\tau A}\left(x_n - \tau\left(Cx_n + \sum_{i=1}^m L_i^* v_{i,n} - z\right)\right) \\ \text{For } i = 1, \ldots, m \\ \left\lfloor \begin{array}{l} \widetilde{p}_{i,n} \approx J_{\theta_{1,i}B_i^{-1}}\left(p_{i,n} + \theta_{1,i}K_i z_{i,n}\right) \\ \widetilde{q}_{i,n} \approx J_{\theta_{2,i}D_i^{-1}}\left(q_{i,n} + \theta_{2,i}M_i y_{i,n}\right) \\ u_{1,i,n} \approx z_{i,n} + \gamma_{1,i}\left(K_i^*\left(p_{i,n} - 2\widetilde{p}_{i,n}\right) + v_{i,n} + \sigma_i\left(L_i(2\widetilde{x}_n - x_n) - r_i\right)\right) \\ u_{2,i,n} \approx y_{i,n} + \gamma_{2,i}\left(M_i^*\left(q_{i,n} - 2\widetilde{q}_{i,n}\right) + v_{i,n} + \sigma_i\left(L_i(2\widetilde{x}_n - x_n) - r_i\right)\right) \\ \widetilde{z}_{i,n} \approx \frac{1+\sigma_i\gamma_{2,i}}{1+\sigma_i(\gamma_{1,i}+\gamma_{2,i})}\left(u_{1,i,n} - \frac{\sigma_i\gamma_{1,i}}{1+\sigma_i\gamma_{2,i}}u_{2,i,n}\right) \\ \widetilde{y}_{i,n} \approx \frac{1}{1+\sigma_i\gamma_{2,i}}\left(u_{2,i,n} - \sigma_i\gamma_{2,i}\widetilde{z}_{i,n}\right) \\ \widetilde{v}_{i,n} \approx v_{i,n} + \sigma_i\left(L_i(2\widetilde{x}_n - x_n) - r_i - \widetilde{z}_{i,n} - \widetilde{y}_{i,n}\right) \end{array} \right. \\ x_{n+1} = x_n + \lambda_n(\widetilde{x}_n - x_n) \\ \text{For } i = 1, \ldots, m \\ \left\lfloor \begin{array}{l} p_{i,n+1} = p_{i,n} + \lambda_n(\widetilde{p}_{i,n} - p_{i,n}) \\ q_{i,n+1} = q_{i,n} + \lambda_n(\widetilde{q}_{i,n} - q_{i,n}) \\ z_{i,n+1} = z_{i,n} + \lambda_n(\widetilde{z}_{i,n} - z_{i,n}) \\ y_{i,n+1} = y_{i,n} + \lambda_n(\widetilde{y}_{i,n} - y_{i,n}) \\ v_{i,n+1} = v_{i,n} + \lambda_n(\widetilde{v}_{i,n} - v_{i,n}). \end{array} \right. \end{array} \right. \tag{2.4}$$

Theorem 2.1. *For Problem 1.2, suppose that*

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

and consider the sequences generated by Algorithm 2.1. Then there exists a primal-dual solution $(\overline{x},\overline{\boldsymbol{p}},\overline{\boldsymbol{q}},\overline{\boldsymbol{y}})$ to Problem 1.2 such that

(i) $x_n \rightharpoonup \overline{x}$, $p_{i,n} \rightharpoonup \overline{p}_i$, $q_{i,n} \rightharpoonup \overline{q}_i$ *and* $y_{i,n} \rightharpoonup \overline{y}_i$ *for any* $i=1,\dots,m$ *as* $n\to+\infty$.

(ii) if C is uniformly monotone at $\overline{x}$, then $x_n \to \overline{x}$ as $n\to+\infty$.

Proof. We introduce the real Hilbert space $\boldsymbol{\mathcal{K}}=\mathcal{H}\oplus\boldsymbol{\mathcal{X}}\oplus\boldsymbol{\mathcal{Y}}\oplus\boldsymbol{\mathcal{G}}\oplus\boldsymbol{\mathcal{G}}\oplus\boldsymbol{\mathcal{G}}$ and let

$$\begin{cases}\boldsymbol{p}=(p_1,\dots,p_m)\\ \boldsymbol{q}=(q_1,\dots,q_m)\\ \boldsymbol{y}=(y_1,\dots,y_m)\end{cases} \text{and} \begin{cases}\boldsymbol{z}=(z_1,\dots,z_m)\\ \boldsymbol{v}=(v_1,\dots,v_m)\\ \boldsymbol{r}=(r_1,\dots,r_m)\end{cases}. \tag{2.6}$$

We introduce the maximally monotone operators

$$\boldsymbol{B}:\boldsymbol{\mathcal{X}}\to 2^{\boldsymbol{\mathcal{X}}},\ \boldsymbol{p}\mapsto B_1p_1\times\ldots\times B_mp_m \text{ and } \boldsymbol{D}:\boldsymbol{\mathcal{Y}}\to 2^{\boldsymbol{\mathcal{Y}}},\ \boldsymbol{q}\mapsto D_1q_1\times\ldots\times D_mq_m.$$

Further, consider the set-valued operator

$$\boldsymbol{M}:\boldsymbol{\mathcal{K}}\to 2^{\boldsymbol{\mathcal{K}}},\ (x,\boldsymbol{p},\boldsymbol{q},\boldsymbol{z},\boldsymbol{y},\boldsymbol{v})\mapsto(-z+Ax)\times\boldsymbol{B}^{-1}\boldsymbol{p}\times\boldsymbol{D}^{-1}\boldsymbol{q}\times(-\boldsymbol{v},-\boldsymbol{v},\boldsymbol{r}+\boldsymbol{z}+\boldsymbol{y}),$$

which is maximally monotone, since A, $\boldsymbol{B}$ and $\boldsymbol{D}$ are maximally monotone (cf. [2, Proposition 20.22 and Proposition 20.23]) and the linear bounded operator

$$(x,\boldsymbol{p},\boldsymbol{q},\boldsymbol{y},\boldsymbol{z},\boldsymbol{v})\mapsto(0,\boldsymbol{0},\boldsymbol{0},-\boldsymbol{v},-\boldsymbol{v},\boldsymbol{z}+\boldsymbol{y})$$

is skew and hence maximally monotone (cf. [2, Example 20.30]). Therefore, $\boldsymbol{M}$ can be written as the sum of two maximally monotone operators, one of them having full domain, fact which leads to the maximality of $\boldsymbol{M}$ (see, for instance, [2, Corollary 24.4(i)]). Furthermore, consider the linear bounded operators

$$\widetilde{K}:\boldsymbol{\mathcal{G}}\to\boldsymbol{\mathcal{X}},\ \boldsymbol{z}\mapsto(K_1z_1,\dots,K_mz_m),\ \widetilde{M}:\boldsymbol{\mathcal{G}}\to\boldsymbol{\mathcal{Y}},\ \boldsymbol{y}\mapsto(M_1y_1,\dots,M_my_m),$$

and

$$\begin{aligned}&\boldsymbol{S}:\boldsymbol{\mathcal{K}}\to\boldsymbol{\mathcal{K}},\\ &(x,\boldsymbol{p},\boldsymbol{q},\boldsymbol{z},\boldsymbol{y},\boldsymbol{v})\mapsto\left(\sum_{i=1}^m L_i^*v_i,-\widetilde{K}\boldsymbol{z},-\widetilde{M}\boldsymbol{y},\widetilde{K}^*\boldsymbol{p},\widetilde{M}^*\boldsymbol{q},-L_1x,\dots,-L_mx\right).\end{aligned}$$

The operator $\boldsymbol{S}$ is skew as well and therefore maximally monotone. As $\operatorname{dom}\boldsymbol{S}=\boldsymbol{\mathcal{K}}$, the sum $\boldsymbol{M}+\boldsymbol{S}$ is maximally monotone (see [2, Corollary 24.4(i)]).

Finally, we introduce the monotone operator

$$\boldsymbol{Q}:\boldsymbol{\mathcal{K}}\to\boldsymbol{\mathcal{K}},\ (x,\boldsymbol{p},\boldsymbol{q},\boldsymbol{z},\boldsymbol{y},\boldsymbol{v})\mapsto(Cx,\boldsymbol{0},\boldsymbol{0},\boldsymbol{0},\boldsymbol{0},\boldsymbol{0})$$

which is, obviously, μ^{-1}-cocoercive. By making use of (2.2), we observe that

$$(2.5)\ \Leftrightarrow\exists\,(x,\boldsymbol{p},\boldsymbol{q},\boldsymbol{y})\in\mathcal{H}\oplus\boldsymbol{\mathcal{X}}\oplus\boldsymbol{\mathcal{Y}}\oplus\boldsymbol{\mathcal{G}}:\begin{cases}z-\sum_{i=1}^m L_i^*K_i^*p_i\in Ax+Cx,\\ K_i(L_ix-y_i-r_i)\in B_i^{-1}p_i,\ i=1,\dots,m,\\ M_iy_i\in D_i^{-1}q_i,\ i=1,\dots,m,\\ K_i^*p_i=M_i^*q_i,\ i=1,\dots,m.\end{cases}$$

$$\Leftrightarrow \begin{array}{l} \exists\,(x,\boldsymbol{p},\boldsymbol{q}) \in \mathcal{H} \oplus \boldsymbol{\mathcal{X}} \oplus \boldsymbol{\mathcal{Y}} \\ \exists\,(\boldsymbol{z},\boldsymbol{y},\boldsymbol{v}) \in \boldsymbol{\mathcal{G}} \oplus \boldsymbol{\mathcal{G}} \oplus \boldsymbol{\mathcal{G}} \end{array} : \begin{cases} 0 \in -z + Ax + \sum_{i=1}^m L_i^* v_i + Cx, \\ 0 \in -K_i z_i + B_i^{-1} p_i, \ i = 1,\dots,m, \\ 0 \in -M_i y_i + D_i^{-1} q_i, \ i = 1,\dots,m, \\ 0 = K_i^* p_i - v_i, \ i = 1,\dots,m, \\ 0 = M_i^* q_i - v_i, \ i = 1,\dots,m, \\ 0 = r_i + z_i + y_i - L_i x, \ i = 1,\dots,m \end{cases}$$

$$\Leftrightarrow \exists\,(x,\boldsymbol{p},\boldsymbol{q},\boldsymbol{z},\boldsymbol{y},\boldsymbol{v}) \in \operatorname{zer}(\boldsymbol{M}+\boldsymbol{S}+\boldsymbol{Q}).$$

From here it follows that

$$\begin{aligned} &(\overline{x},\overline{\boldsymbol{p}},\overline{\boldsymbol{q}},\overline{\boldsymbol{z}},\overline{\boldsymbol{y}},\overline{\boldsymbol{v}}) \in \operatorname{zer}(\boldsymbol{M}+\boldsymbol{S}+\boldsymbol{Q}) \\ &\Rightarrow \begin{cases} z - \sum_{i=1}^m L_i^* K_i^* \overline{p}_i \in A\overline{x} + C\overline{x}, \\ K_i(L_i\overline{x} - \overline{y}_i - r_i) \in B_i^{-1}\overline{p}_i, \ i = 1,\dots,m, \\ M_i \overline{y}_i \in D_i^{-1} \overline{q}_i, \ i = 1,\dots,m, \\ K_i^* \overline{p}_i = M_i^* \overline{q}_i, \ i = 1,\dots,m. \end{cases} \\ &\Leftrightarrow (\overline{x},\overline{\boldsymbol{p}},\overline{\boldsymbol{q}},\overline{\boldsymbol{y}}) \text{ is a primal-dual solution to Problem 1.2.} \end{aligned} \tag{2.7}$$

Further, for positive real values $\tau, \theta_{1,i}, \theta_{2,i}, \gamma_{1,i}, \gamma_{2,i}, \sigma_i \in \mathbb{R}_{++}$, $i = 1,\dots,m$, we introduce the notations

$$\begin{cases} \frac{\boldsymbol{p}}{\theta_1} = \left(\frac{p_1}{\theta_{1,1}},\dots,\frac{p_m}{\theta_{1,m}}\right) \\ \frac{\boldsymbol{q}}{\theta_2} = \left(\frac{q_1}{\theta_{2,1}},\dots,\frac{q_m}{\theta_{2,m}}\right) \end{cases}, \quad \begin{cases} \frac{\boldsymbol{z}}{\gamma_1} = \left(\frac{z_1}{\gamma_{1,1}},\dots,\frac{z_m}{\gamma_{1,m}}\right) \\ \frac{\boldsymbol{y}}{\gamma_2} = \left(\frac{y_1}{\gamma_{2,1}},\dots,\frac{y_m}{\gamma_{2,m}}\right) \end{cases}, \quad \left\{ \frac{\boldsymbol{v}}{\sigma} = \left(\frac{v_1}{\sigma_1},\dots,\frac{v_m}{\sigma_m}\right), \right.$$

and define the linear bounded operator

$$\begin{aligned} \boldsymbol{V} : \boldsymbol{\mathcal{K}} \to \boldsymbol{\mathcal{K}}, \ (x,\boldsymbol{p},\boldsymbol{q},\boldsymbol{z},\boldsymbol{y},\boldsymbol{v}) \mapsto &\left(\frac{x}{\tau}, \frac{\boldsymbol{p}}{\theta_1}, \frac{\boldsymbol{q}}{\theta_2}, \frac{\boldsymbol{z}}{\gamma_1}, \frac{\boldsymbol{y}}{\gamma_2}, \frac{\boldsymbol{v}}{\sigma}\right) + \\ &\left(-\sum_{i=1}^m L_i^* v_i, \widetilde{K}\boldsymbol{z}, \widetilde{M}\boldsymbol{y}, \widetilde{K}^*\boldsymbol{p}, \widetilde{M}^*\boldsymbol{q}, -L_1 x, \dots, -L_m x\right). \end{aligned}$$

It is a simple calculation to prove that $\boldsymbol{V}$ is self-adjoint. Furthermore, the operator $\boldsymbol{V}$ is ρ-strongly positive with

$$\rho = (1-\overline{\alpha}) \min_{i=1,\dots,m} \left\{ \frac{1}{\tau}, \frac{1}{\theta_{1,i}}, \frac{1}{\theta_{2,i}}, \frac{1}{\gamma_{1,i}}, \frac{1}{\gamma_{2,i}}, \frac{1}{\sigma_i} \right\} > 0,$$

for

$$\overline{\alpha} = \max \left\{ \sqrt{\tau \sum_{i=1}^m \sigma_i \|L_i\|^2}, \max_{j=1,\dots,m} \left\{ \sqrt{\theta_{1,j}\gamma_{1,j}\|K_j\|^2}, \sqrt{\theta_{2,j}\gamma_{2,j}\|M_j\|^2} \right\} \right\}.$$

The fact that ρ is a positive real number follows by the assumptions made in Algorithm 2.1. Indeed, using that $2ab \le \alpha a^2 + \frac{b^2}{\alpha}$ for every $a, b \in \mathbb{R}$ and every $\alpha \in \mathbb{R}_{++}$, it yields for any $i = 1,\dots,m$

$$\begin{aligned} 2\|L_i\|\|x\|_{\mathcal{H}}\|v_i\|_{\mathcal{G}_i} &\le \frac{\sigma_i\|L_i\|^2}{\sqrt{\tau\sum_{i=1}^m \sigma_i\|L_i\|^2}}\|x\|_{\mathcal{H}}^2 + \frac{\sqrt{\tau\sum_{i=1}^m \sigma_i\|L_i\|^2}}{\sigma_i}\|v_i\|_{\mathcal{G}_i}^2, \\ 2\|K_i\|\|p_i\|_{\mathcal{X}_i}\|z_i\|_{\mathcal{G}_i} &\le \frac{\gamma_{1,i}\|K_i\|}{\sqrt{\theta_{1,i}\gamma_{1,i}}}\|p_i\|_{\mathcal{X}_i}^2 + \frac{\sqrt{\theta_{1,i}\gamma_{1,i}\|K_i\|^2}}{\gamma_{1,i}}\|z_i\|_{\mathcal{G}_i}^2, \\ 2\|M_i\|\|q_i\|_{\mathcal{Y}_i}\|y_i\|_{\mathcal{G}_i} &\le \frac{\gamma_{2,i}\|M_i\|}{\sqrt{\theta_{2,i}\gamma_{2,i}}}\|q_i\|_{\mathcal{Y}_i}^2 + \frac{\sqrt{\theta_{2,i}\gamma_{2,i}\|M_i\|^2}}{\gamma_{2,i}}\|y_i\|_{\mathcal{G}_i}^2. \end{aligned} \tag{2.8}$$

Consequently, for each $\boldsymbol{x} = (x, \boldsymbol{p}, \boldsymbol{q}, \boldsymbol{z}, \boldsymbol{y}, \boldsymbol{v}) \in \boldsymbol{\mathcal{K}}$, using the Cauchy–Schwarz inequality and (2.8), it follows that

$$\begin{aligned}
\langle \boldsymbol{x}, \boldsymbol{V}\boldsymbol{x}\rangle_{\boldsymbol{\mathcal{K}}} &= \frac{\|x\|^2_{\mathcal{H}}}{\tau} + \sum_{i=1}^{m}\left[\frac{\|p_i\|^2_{\mathcal{X}_i}}{\theta_{1,i}} + \frac{\|q_i\|^2_{\mathcal{Y}_i}}{\theta_{2,i}} + \frac{\|z_i\|^2_{\mathcal{G}_i}}{\gamma_{1,i}} + \frac{\|y_i\|^2_{\mathcal{G}_i}}{\gamma_{2,i}} + \frac{\|v_i\|^2_{\mathcal{G}_i}}{\sigma_i}\right] \\
&\quad - 2\sum_{i=1}^{m}\langle L_i x, v_i\rangle_{\mathcal{G}_i} + 2\sum_{i=1}^{m}\langle p_i, K_i z_i\rangle_{\mathcal{X}_i} + 2\sum_{i=1}^{m}\langle q_i, M_i y_i\rangle_{\mathcal{Y}_i} \\
&\geq (1-\overline{\alpha})\min_{i=1,\dots,m}\left\{\frac{1}{\tau}, \frac{1}{\theta_{1,i}}, \frac{1}{\theta_{2,i}}, \frac{1}{\gamma_{1,i}}, \frac{1}{\gamma_{2,i}}, \frac{1}{\sigma_i}\right\}\|\boldsymbol{x}\|^2_{\boldsymbol{\mathcal{K}}} \\
&= \rho\|\boldsymbol{x}\|^2_{\boldsymbol{\mathcal{K}}}.
\end{aligned} \tag{2.9}$$

Since $\boldsymbol{V}$ is maximally monotone (cf. [2, Example 20.29]) and ρ-strongly positive, it is strongly monotone and therefore, by [2, Proposition 22.8], it holds that $\boldsymbol{V}$ is surjective. Consequently, $\boldsymbol{V}^{-1}$ exists and $\|\boldsymbol{V}^{-1}\| \leq \frac{1}{\rho}$.

In consideration of (1.9), the algorithmic scheme (2.4) can equivalently be written in the form

$$(\forall n \geq 0) \quad \left\lfloor \begin{array}{l}
\frac{x_n - \widetilde{x}_n}{\tau} - \sum_{i=1}^{m} L_i^*(v_{i,n} - \widetilde{v}_{i,n}) - Cx_n \\
\qquad\qquad \in -z + A(\widetilde{x}_n - a_n) + \sum_{i=1}^{m} L_i^*\widetilde{v}_{i,n} - \frac{a_n}{\tau} \\
\text{For } i = 1, \dots, m \\
\left\lfloor \begin{array}{l}
\frac{p_{i,n} - \widetilde{p}_{i,n}}{\theta_{1,i}} + K_i(z_{i,n} - \widetilde{z}_{i,n}) \in B_i^{-1}(\widetilde{p}_{i,n} - b_{i,n}) - K_i\widetilde{z}_{i,n} - \frac{b_{i,n}}{\theta_{1,i}} \\
\frac{q_{i,n} - \widetilde{q}_{i,n}}{\theta_{2,i}} + M_i(y_{i,n} - \widetilde{y}_{i,n}) \in D_i^{-1}(\widetilde{q}_{i,n} - d_{i,n}) - M_i\widetilde{y}_{i,n} - \frac{d_{i,n}}{\theta_{2,i}} \\
\frac{z_{i,n} - \widetilde{z}_{i,n}}{\gamma_{1,i}} + K_i^*(p_{i,n} - \widetilde{p}_{i,n}) = -\widetilde{v}_{i,n} + K_i^*\widetilde{p}_{i,n} - e_{1,i,n} \\
\frac{y_{i,n} - \widetilde{y}_{i,n}}{\gamma_{2,i}} + M_i^*(q_{i,n} - \widetilde{q}_{i,n}) = -\widetilde{v}_{i,n} + M_i^*\widetilde{q}_{i,n} - e_{2,i,n} \\
\frac{v_{i,n} - \widetilde{v}_{i,n}}{\sigma_i} - L_i(x_n - \widetilde{x}_n) = r_i + \widetilde{z}_{i,n} + \widetilde{y}_{i,n} - L_i\widetilde{x}_n - e_{3,i,n}
\end{array} \right. \\
\boldsymbol{x}_{n+1} = \boldsymbol{x}_n + \lambda_n(\widetilde{\boldsymbol{x}}_n - \boldsymbol{x}_n),
\end{array} \right. \tag{2.10}$$

where

$$\left\{ \begin{array}{l}
\boldsymbol{p}_n = (p_{1,n}, \dots p_{m,n}) \in \boldsymbol{\mathcal{X}} \\
\boldsymbol{q}_n = (q_{1,n}, \dots, q_{m,n}) \in \boldsymbol{\mathcal{Y}} \\
\boldsymbol{z}_n = (z_{1,n}, \dots, z_{m,n}) \in \boldsymbol{\mathcal{G}} \\
\boldsymbol{y}_n = (y_{1,n}, \dots, y_{m,n}) \in \boldsymbol{\mathcal{G}} \\
\boldsymbol{v}_n = (v_{1,n}, \dots, v_{m,n}) \in \boldsymbol{\mathcal{G}}
\end{array} \right. \qquad
\left\{ \begin{array}{l}
\widetilde{\boldsymbol{p}}_n = (\widetilde{p}_{1,n}, \dots \widetilde{p}_{m,n}) \in \boldsymbol{\mathcal{X}} \\
\widetilde{\boldsymbol{q}}_n = (\widetilde{q}_{1,n}, \dots, \widetilde{q}_{m,n}) \in \boldsymbol{\mathcal{Y}} \\
\widetilde{\boldsymbol{z}}_n = (\widetilde{z}_{1,n}, \dots, \widetilde{z}_{m,n}) \in \boldsymbol{\mathcal{G}} \\
\widetilde{\boldsymbol{y}}_n = (\widetilde{y}_{1,n}, \dots, \widetilde{y}_{m,n}) \in \boldsymbol{\mathcal{G}} \\
\widetilde{\boldsymbol{v}}_n = (\widetilde{v}_{1,n}, \dots, \widetilde{v}_{m,n}) \in \boldsymbol{\mathcal{G}}
\end{array} \right.$$

and

$$\left\{ \begin{array}{l}
\boldsymbol{x}_n = (x_n, \boldsymbol{p}_n, \boldsymbol{q}_n, \boldsymbol{z}_n, \boldsymbol{y}_n, \boldsymbol{v}_n) \in \boldsymbol{\mathcal{K}} \\
\widetilde{\boldsymbol{x}}_n = (\widetilde{x}_n, \widetilde{\boldsymbol{p}}_n, \widetilde{\boldsymbol{q}}_n, \widetilde{\boldsymbol{z}}_n, \widetilde{\boldsymbol{y}}_n, \widetilde{\boldsymbol{v}}_n) \in \boldsymbol{\mathcal{K}}.
\end{array} \right.$$

Also, for any $n \geq 0$, we consider sequences defined by

$$\left\{ \begin{array}{l}
a_n \in \mathcal{H} \\
\boldsymbol{b}_n = (b_{1,n}, \dots b_{m,n}) \in \boldsymbol{\mathcal{X}} \\
\boldsymbol{d}_n = (d_{1,n}, \dots, d_{m,n}) \in \boldsymbol{\mathcal{Y}}
\end{array} \right. \quad \text{and} \quad
\left\{ \begin{array}{l}
\boldsymbol{e}_{1,n} = (e_{1,1,n}, \dots, e_{1,m,n}) \in \boldsymbol{\mathcal{G}} \\
\boldsymbol{e}_{2,n} = (e_{2,1,n}, \dots, e_{2,m,n}) \in \boldsymbol{\mathcal{G}}, \\
\boldsymbol{e}_{3,n} = (e_{3,1,n}, \dots, e_{3,m,n}) \in \boldsymbol{\mathcal{G}}
\end{array} \right. \tag{2.11}$$

that are summable in the corresponding norm. Further, by denoting for any $n \geq 0$

$$\left\{ \begin{array}{l}
\boldsymbol{e}_n = (a_n, \boldsymbol{b}_n, \boldsymbol{d}_n, \boldsymbol{0}, \boldsymbol{0}, \boldsymbol{0}) \in \boldsymbol{\mathcal{K}} \\
\boldsymbol{e}_n^{\tau} = \left(\frac{a_n}{\tau}, \frac{\boldsymbol{b}_n}{\theta_1}, \frac{\boldsymbol{d}_n}{\theta_2}, \boldsymbol{e}_{1,n}, \boldsymbol{e}_{2,n}, \boldsymbol{e}_{3,n}\right) \in \boldsymbol{\mathcal{K}},
\end{array} \right.$$

which are also terms of summable sequences in the corresponding norm, it yields that the scheme in (2.10) is equivalent to

$$(\forall n \geq 0) \left\lfloor \begin{array}{l} \boldsymbol{V}(\boldsymbol{x}_n - \widetilde{\boldsymbol{x}}_{\boldsymbol{n}}) - \boldsymbol{Q}\boldsymbol{x}_n \in (\boldsymbol{M}+\boldsymbol{S})(\widetilde{\boldsymbol{x}}_{\boldsymbol{n}} - \boldsymbol{e}_n) + \boldsymbol{S}\boldsymbol{e}_n - \boldsymbol{e}_n^\tau \\ \boldsymbol{x}_{n+1} = \boldsymbol{x}_n + \lambda_n (\widetilde{\boldsymbol{x}}_{\boldsymbol{n}} - \boldsymbol{x}_n). \end{array} \right. \tag{2.12}$$

We now introduce the notations

$$\boldsymbol{A}_{\boldsymbol{\mathcal{K}}} := \boldsymbol{V}^{-1}(\boldsymbol{M}+\boldsymbol{S}) \text{ and } \boldsymbol{B}_{\boldsymbol{\mathcal{K}}} := \boldsymbol{V}^{-1}\boldsymbol{Q} \tag{2.13}$$

and the summable sequence with terms $\boldsymbol{e}_n^{\boldsymbol{V}} = \boldsymbol{V}^{-1}((\boldsymbol{V}+\boldsymbol{S})\boldsymbol{e}_n - \boldsymbol{e}_n^\tau)$ for any $n \geq 0$. Then, for any $n \geq 0$, we have

$$\begin{aligned} & \boldsymbol{V}(\boldsymbol{x}_n - \widetilde{\boldsymbol{x}}_{\boldsymbol{n}}) - \boldsymbol{Q}\boldsymbol{x}_n \in (\boldsymbol{M}+\boldsymbol{S})(\widetilde{\boldsymbol{x}}_{\boldsymbol{n}} - \boldsymbol{e}_n) + \boldsymbol{S}\boldsymbol{e}_n - \boldsymbol{e}_n^\tau \\ \Leftrightarrow\ & \boldsymbol{V}\boldsymbol{x}_n - \boldsymbol{Q}\boldsymbol{x}_n \in (\boldsymbol{V}+\boldsymbol{M}+\boldsymbol{S})(\widetilde{\boldsymbol{x}}_{\boldsymbol{n}} - \boldsymbol{e}_n) + (\boldsymbol{V}+\boldsymbol{S})\boldsymbol{e}_n - \boldsymbol{e}_n^\tau \\ \Leftrightarrow\ & \boldsymbol{x}_n - \boldsymbol{V}^{-1}\boldsymbol{Q}\boldsymbol{x}_n \in \left(\mathrm{Id} + \boldsymbol{V}^{-1}(\boldsymbol{M}+\boldsymbol{S})\right)(\widetilde{\boldsymbol{x}}_{\boldsymbol{n}} - \boldsymbol{e}_n) + \boldsymbol{V}^{-1}((\boldsymbol{V}+\boldsymbol{S})\boldsymbol{e}_n - \boldsymbol{e}_n^\tau) \\ \Leftrightarrow\ & \widetilde{\boldsymbol{x}}_{\boldsymbol{n}} = \left(\mathrm{Id} + \boldsymbol{V}^{-1}(\boldsymbol{M}+\boldsymbol{S})\right)^{-1}\left(\boldsymbol{x}_n - \boldsymbol{V}^{-1}\boldsymbol{Q}\boldsymbol{x}_n - \boldsymbol{e}_n^{\boldsymbol{V}}\right) + \boldsymbol{e}_n \\ \Leftrightarrow\ & \widetilde{\boldsymbol{x}}_{\boldsymbol{n}} = (\mathrm{Id} + \boldsymbol{A}_{\boldsymbol{\mathcal{K}}})^{-1}\left(\boldsymbol{x}_n - \boldsymbol{B}_{\boldsymbol{\mathcal{K}}}\boldsymbol{x}_n - \boldsymbol{e}_n^{\boldsymbol{V}}\right) + \boldsymbol{e}_n. \end{aligned} \tag{2.14}$$

Taking into account that the resolvent is Lipschitz continuous, the sequence having as terms

$$\boldsymbol{e}_n^{\boldsymbol{A}_{\boldsymbol{\mathcal{K}}}} = J_{\boldsymbol{A}_{\boldsymbol{\mathcal{K}}}}\left(\boldsymbol{x}_n - \boldsymbol{B}_{\boldsymbol{\mathcal{K}}}\boldsymbol{x}_n - \boldsymbol{e}_n^{\boldsymbol{V}}\right) - J_{\boldsymbol{A}_{\boldsymbol{\mathcal{K}}}}(\boldsymbol{x}_n - \boldsymbol{B}_{\boldsymbol{\mathcal{K}}}\boldsymbol{x}_n) + \boldsymbol{e}_n \ \forall n \geq 0$$

is summable and we have

$$\widetilde{\boldsymbol{x}}_{\boldsymbol{n}} = J_{\boldsymbol{A}_{\boldsymbol{\mathcal{K}}}}(\boldsymbol{x}_n - \boldsymbol{B}_{\boldsymbol{\mathcal{K}}}\boldsymbol{x}_n) + \boldsymbol{e}_n^{\boldsymbol{A}_{\boldsymbol{\mathcal{K}}}} \ \forall n \geq 0.$$

Thus, the iterative scheme in (2.12) becomes

$$(\forall n \geq 0) \left\lfloor \begin{array}{l} \widetilde{\boldsymbol{x}}_{\boldsymbol{n}} \approx J_{\boldsymbol{A}_{\boldsymbol{\mathcal{K}}}}(\boldsymbol{x}_n - \boldsymbol{B}_{\boldsymbol{\mathcal{K}}}\boldsymbol{x}_n) \\ \boldsymbol{x}_{n+1} = \boldsymbol{x}_n + \lambda_n(\widetilde{\boldsymbol{x}}_{\boldsymbol{n}} - \boldsymbol{x}_n), \end{array} \right. \tag{2.15}$$

which shows that the algorithm we propose in this subsection has the structure of a forward-backward method.

In addition, let us observe that

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

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

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

Since $\boldsymbol{M}+\boldsymbol{S}$ and $\boldsymbol{Q}$ are maximally monotone on $\boldsymbol{\mathcal{K}}$, the operators $\boldsymbol{A}_{\boldsymbol{\mathcal{K}}}$ and $\boldsymbol{B}_{\boldsymbol{\mathcal{K}}}$ are maximally monotone on $\boldsymbol{\mathcal{K}}_{\boldsymbol{V}}$. Moreover, since $\boldsymbol{V}$ is self-adjoint and ρ-strongly positive, one can easily see that weak and strong convergence in $\boldsymbol{\mathcal{K}}_{\boldsymbol{V}}$ are equivalent with weak and strong convergence in $\boldsymbol{\mathcal{K}}$, respectively. By making use of $\|V^{-1}\| \leq \frac{1}{\rho}$, one can show that

$\boldsymbol{B_\mathcal{K}}$ is $(\mu^{-1}\rho)$-cocoercive on $\boldsymbol{\mathcal{K}_V}$. Indeed, we get for $\boldsymbol{x}$, $\boldsymbol{y} \in \boldsymbol{\mathcal{K}_V}$ that (see, also, [29, Eq. (3.35)])

$$\begin{aligned}
\langle \boldsymbol{x}-\boldsymbol{y}, \boldsymbol{B_\mathcal{K}}\boldsymbol{x}-\boldsymbol{B_\mathcal{K}}\boldsymbol{y}\rangle_{\boldsymbol{\mathcal{K}_V}} &= \langle \boldsymbol{x}-\boldsymbol{y}, \boldsymbol{Q}\boldsymbol{x}-\boldsymbol{Q}\boldsymbol{y}\rangle_{\boldsymbol{\mathcal{K}}} \\
&\geq \mu^{-1}\|\boldsymbol{Q}\boldsymbol{x}-\boldsymbol{Q}\boldsymbol{y}\|^2_{\boldsymbol{\mathcal{K}}} \\
&\geq \mu^{-1}\|\boldsymbol{V}^{-1}\|^{-1}\|\boldsymbol{V}^{-1}\boldsymbol{Q}\boldsymbol{x}-\boldsymbol{V}^{-1}\boldsymbol{Q}\boldsymbol{y}\|_{\boldsymbol{\mathcal{K}}}\|\boldsymbol{Q}\boldsymbol{x}-\boldsymbol{Q}\boldsymbol{y}\|_{\boldsymbol{\mathcal{K}}} \\
&\geq \mu^{-1}\|\boldsymbol{V}^{-1}\|^{-1}\langle \boldsymbol{B_\mathcal{K}}\boldsymbol{x}-\boldsymbol{B_\mathcal{K}}\boldsymbol{y}, \boldsymbol{Q}\boldsymbol{x}-\boldsymbol{Q}\boldsymbol{y}\rangle_{\boldsymbol{\mathcal{K}}} \\
&= \mu^{-1}\|\boldsymbol{V}^{-1}\|^{-1}\|\boldsymbol{B_\mathcal{K}}\boldsymbol{x}-\boldsymbol{B_\mathcal{K}}\boldsymbol{y}\|^2_{\boldsymbol{\mathcal{K}_V}} \\
&\geq \mu^{-1}\rho\|\boldsymbol{B_\mathcal{K}}\boldsymbol{x}-\boldsymbol{B_\mathcal{K}}\boldsymbol{y}\|^2_{\boldsymbol{\mathcal{K}_V}}.
\end{aligned} \tag{2.17}$$

As our assumption imposes that $2\mu^{-1}\rho > 1$, we can use the statements given in [17, Corollary 6.5] in the context of an error tolerant forward-backward algorithm in order to establish the desired convergence results.

(i) By Corollary 6.5 in [17], the sequence $(\boldsymbol{x}_n)_{n\geq 0}$ converges weakly in $\boldsymbol{\mathcal{K}_V}$ (and therefore in $\boldsymbol{\mathcal{K}}$) to some $\overline{\boldsymbol{x}} = (\overline{x}, \overline{\boldsymbol{p}}, \overline{\boldsymbol{q}}, \overline{\boldsymbol{z}}, \overline{\boldsymbol{y}}, \overline{\boldsymbol{v}}) \in \operatorname{zer}(\boldsymbol{A_\mathcal{K}} + \boldsymbol{B_\mathcal{K}}) = \operatorname{zer}(\boldsymbol{M}+\boldsymbol{S}+\boldsymbol{Q})$. By (2.7), it thus follows that $(\overline{x}, \overline{\boldsymbol{p}}, \overline{\boldsymbol{q}}, \overline{\boldsymbol{y}})$ is a primal-dual solution with respect to Problem 1.2.

(ii) From [17, Remark 3.4], it follows

$$\sum_{n\geq 0}\|\boldsymbol{B_\mathcal{K}}\boldsymbol{x}_n - \boldsymbol{B_\mathcal{K}}\overline{\boldsymbol{x}}\|^2_{\boldsymbol{\mathcal{K}_V}} < +\infty,$$

and therefore we have $\boldsymbol{B_\mathcal{K}}\boldsymbol{x}_n \to \boldsymbol{B_\mathcal{K}}\overline{\boldsymbol{x}}_n$ or, equivalently, $\boldsymbol{Q}\boldsymbol{x}_n \to \boldsymbol{Q}\overline{\boldsymbol{x}}$ as $n \to +\infty$. Considering the definition of $\boldsymbol{Q}$, one can see that this implies $Cx_n \to C\overline{x}$ as $n \to +\infty$. As C is uniformly monotone, there exists an increasing function $\phi_C : [0, +\infty) \to [0, +\infty]$ vanishing only at 0 such that

$$\phi_C(\|x_n - \overline{x}\|) \leq \langle x_n - \overline{x}, Cx_n - C\overline{x}\rangle \leq \|x_n - \overline{x}\|\|Cx_n - C\overline{x}\| \ \forall n \geq 0.$$

The boundedness of $(x_n - \overline{x})_{n\geq 0}$ and the convergence $Cx_n \to C\overline{x}$ further imply that $x_n \to \overline{x}$ as $n \to +\infty$. □

Remark 2.2. Suppose that $C : \mathcal{H} \to \mathcal{H}$, $x \mapsto \{0\}$, in Problem 1.2. Then condition (2.3) simplifies to

$$\max\left\{\tau\sum_{i=1}^m \sigma_i\|L_i\|^2, \max_{j=1,\dots,m}\left\{\theta_{1,j}\gamma_{1,j}\|K_j\|^2, \theta_{2,j}\gamma_{2,j}\|M_j\|^2\right\}\right\} < 1.$$

In this case, the scheme (2.15) reads

$$(\forall n \geq 0) \left\lfloor \ \boldsymbol{x}_{n+1} \approx \boldsymbol{x}_n + \lambda_n(J_{\boldsymbol{A_\mathcal{K}}}\boldsymbol{x}_n - \boldsymbol{x}_n), \right. \tag{2.18}$$

and it can be shown to convergence under the relaxed assumption that $(\lambda_n)_{n\geq 0} \subseteq [\varepsilon, 2-\varepsilon]$, for $\varepsilon \in (0,1)$ (see, for instance, [16, 17, 23]).

Remark 2.3. (i) When implementing Algorithm 2.1, the term $L_i(2\widetilde{x}_n - x_n)$ should be stored in a separate variable for any $i = 1, \dots, m$. Taking this into account, each linear bounded operator occurring in Problem 1.2 needs to be processed once via some forward evaluation and once via its adjoint.

(ii) The maximally monotone operators A, B_i and D_i, $i = 1, \dots, m$, in Problem 1.2 are accessed via their resolvents (so-called backward steps), also by taking into account the relation between the resolvent of a maximally monotone operator and its inverse given in (1.10).

(iii) The possibility of performing a forward step for the cocoercive monotone operator C is an important aspect, since forward steps are usually much easier to implement than resolvent steps (resp. evaluations of proximal operators). Due to the Baillon–Haddad theorem (cf. [2, Corollary 18.16]), each μ-Lipschitzian gradient with $\mu \in \mathbb{R}_{++}$ of a convex and Fréchet differentiable function $f : \mathcal{H} \to \mathbb{R}$ is μ^{-1}-cocoercive.

2.2 An algorithm of forward-backward-forward type

In this subsection we propose a forward-backward-forward type algorithm for solving Problem 1.2, with the modification that the operator $C : \mathcal{H} \to \mathcal{H}$ is assumed to be μ-Lipschitz continuous for some $\mu \in \mathbb{R}_{++}$, but not necessarily μ^{-1}-cocoercive.

Algorithm 2.2.
Let $x_0 \in \mathcal{H}$, and for any $i = 1, ..., m$, let $p_{i,0} \in \mathcal{X}_i$, $q_{i,0} \in \mathcal{Y}_i$, and $z_{i,0},\ y_{i,0},\ v_{i,0} \in \mathcal{G}_i$. Set

$$\beta = \mu + \sqrt{\max\left\{ \sum_{i=1}^{m} \|L_i\|^2, \max_{j=1,\dots,m} \left\{\|K_j\|^2, \|M_j\|^2\right\}\right\}}, \tag{2.19}$$

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

$$(\forall n \geq 0) \quad \left\lfloor \begin{array}{l} \widetilde{x}_n \approx J_{\gamma_n A}\left(x_n - \gamma_n \left(C x_n + \sum_{i=1}^m L_i^* v_{i,n} - z\right)\right) \\ \text{For } i = 1, \dots, m \\ \left\lfloor \begin{array}{l} \widetilde{p}_{i,n} \approx J_{\gamma_n B_i^{-1}}\left(p_{i,n} + \gamma_n K_i z_{i,n}\right) \\ \widetilde{q}_{i,n} \approx J_{\gamma_n D_i^{-1}}\left(q_{i,n} + \gamma_n M_i y_{i,n}\right) \\ u_{1,i,n} \approx z_{i,n} - \gamma_n\left(K_i^* p_{i,n} - v_{i,n} - \gamma_n\left(L_i x_n - r_i\right)\right) \\ u_{2,i,n} \approx y_{i,n} - \gamma_n\left(M_i^* q_{i,n} - v_{i,n} - \gamma_n\left(L_i x_n - r_i\right)\right) \\ \widetilde{z}_{i,n} \approx \frac{1+\gamma_n^2}{1+2\gamma_n^2}\left(u_{1,i,n} - \frac{\gamma_n^2}{1+\gamma_n^2} u_{2,i,n}\right) \\ \widetilde{y}_{i,n} \approx \frac{1}{1+\gamma_n^2}\left(u_{2,i,n} - \gamma_n^2 \widetilde{z}_{i,n}\right) \\ \widetilde{v}_{i,n} \approx v_{i,n} + \gamma_n\left(L_i x_n - r_i - \widetilde{z}_{i,n} - \widetilde{y}_{i,n}\right) \end{array}\right. \\ x_{n+1} \approx \widetilde{x}_n + \gamma_n (C x_n - C\widetilde{x}_n + \sum_{i=1}^m L_i^*(v_{i,n} - \widetilde{v}_{i,n})) \\ \text{For } i = 1, \dots, m \\ \left\lfloor \begin{array}{l} p_{i,n+1} \approx \widetilde{p}_{i,n} - \gamma_n (K_i(z_{i,n} - \widetilde{z}_{i,n})) \\ q_{i,n+1} \approx \widetilde{q}_{i,n} - \gamma_n (M_i(y_{i,n} - \widetilde{y}_{i,n})) \\ z_{i,n+1} \approx \widetilde{z}_{i,n} + \gamma_n (K_i^*(p_{i,n} - \widetilde{p}_{i,n})) \\ y_{i,n+1} \approx \widetilde{y}_{i,n} + \gamma_n (M_i^*(q_{i,n} - \widetilde{q}_{i,n})) \\ v_{i,n+1} \approx \widetilde{v}_{i,n} - \gamma_n (L_i(x_n - \widetilde{x}_n)). \end{array}\right. \end{array}\right. \tag{2.20}$$

Theorem 2.2. *In Problem 1.2, let $C : \mathcal{H} \to \mathcal{H}$ be μ-Lipschitz continuous for $\mu \in \mathbb{R}_{++}$, suppose that*

$$z \in \operatorname{ran}\left(A + \sum_{i=1}^{m} L_i^* \Big((K_i^* \circ B_i \circ K_i) \,\square\, (M_i^* \circ D_i \circ M_i) \Big)(L_i \cdot - r_i) + C\right), \tag{2.21}$$

and consider the sequences generated by Algorithm 2.2. Then there exists a primal-dual solution $(\overline{x}, \overline{\boldsymbol{p}}, \overline{\boldsymbol{q}}, \overline{\boldsymbol{y}})$ to Problem 1.2 such that

(i) $\sum_{n\geq 0} \|x_n - \widetilde{x}_n\|^2 < +\infty$ *and for any* $i = 1, ..., m$

$$\sum_{n\geq 0} \|p_{i,n} - \widetilde{p}_{i,n}\|^2 < +\infty, \quad \sum_{n\geq 0} \|q_{i,n} - \widetilde{q}_{i,n}\|^2 < +\infty \ \text{ and } \sum_{n\geq 0} \|y_{i,n} - \widetilde{y}_{i,n}\|^2 < +\infty.$$

(ii) $x_n \rightharpoonup \overline{x}$, $\widetilde{x}_n \rightharpoonup \overline{x}$, *and for any* $i = 1, ..., m$

$$\left\{ \begin{array}{l} p_{i,n} \rightharpoonup \overline{p}_{i,n} \\ \widetilde{p}_{i,n} \rightharpoonup \overline{p}_{i,n} \end{array} \right., \left\{ \begin{array}{l} q_{i,n} \rightharpoonup \overline{q}_{i,n} \\ \widetilde{q}_{i,n} \rightharpoonup \overline{q}_{i,n} \end{array} \right. \text{ and } \left\{ \begin{array}{l} y_{i,n} \rightharpoonup \overline{y}_{i,n} \\ \widetilde{y}_{i,n} \rightharpoonup \overline{y}_{i,n} \end{array} \right. .$$

Proof. As in the proof of Theorem 2.1, consider $\boldsymbol{\mathcal{K}} = \mathcal{H} \oplus \boldsymbol{\mathcal{X}} \oplus \boldsymbol{\mathcal{Y}} \oplus \boldsymbol{\mathcal{G}} \oplus \boldsymbol{\mathcal{G}} \oplus \boldsymbol{\mathcal{G}}$ along with the notations introduced in (2.6). Further, let the operators $\boldsymbol{M} : \boldsymbol{\mathcal{K}} \to 2^{\boldsymbol{\mathcal{K}}}$, $\boldsymbol{S} : \boldsymbol{\mathcal{K}} \to \boldsymbol{\mathcal{K}}$ and $\boldsymbol{Q} : \boldsymbol{\mathcal{K}} \to \boldsymbol{\mathcal{K}}$ be defined as in the proof of the same result. The operator $\boldsymbol{S} + \boldsymbol{Q}$ is monotone, Lipschitz continuous, hence maximally monotone (cf. [2, Corollary 20.25]), and it fulfills $\operatorname{dom}(\boldsymbol{S} + \boldsymbol{Q}) = \boldsymbol{\mathcal{K}}$. Therefore the sum $\boldsymbol{M} + \boldsymbol{S} + \boldsymbol{Q}$ is maximally monotone as well (see [2, Corollary 24.4(i)]).

In the following we derive the Lipschitz constant of $\boldsymbol{S} + \boldsymbol{Q}$. For arbitrary

$$\boldsymbol{x} = (x, \boldsymbol{p}, \boldsymbol{q}, \boldsymbol{z}, \boldsymbol{y}, \boldsymbol{v}) \text{ and } \widetilde{\boldsymbol{x}} = (\widetilde{x}, \widetilde{\boldsymbol{p}}, \widetilde{\boldsymbol{q}}, \widetilde{\boldsymbol{z}}, \widetilde{\boldsymbol{y}}, \widetilde{\boldsymbol{v}}) \in \boldsymbol{\mathcal{K}},$$

by using the Cauchy–Schwarz inequality, it yields

$$\begin{aligned}
&\|(\boldsymbol{S} + \boldsymbol{Q})\boldsymbol{x} - (\boldsymbol{S} + \boldsymbol{Q})\widetilde{\boldsymbol{x}}\| \leq \|\boldsymbol{Q}\boldsymbol{x} - \boldsymbol{Q}\widetilde{\boldsymbol{x}}\| + \|\boldsymbol{S}\boldsymbol{x} - \boldsymbol{S}\widetilde{\boldsymbol{x}}\| \\
&\leq \mu\|x - \widetilde{x}\| + \left\| \left(\sum_{i=1}^m L_i^*(v_i - \widetilde{v}_i), -\widetilde{K}(\boldsymbol{z} - \widetilde{\boldsymbol{z}}), -\widetilde{M}(\boldsymbol{y} - \widetilde{\boldsymbol{y}}), \widetilde{K}^*(\boldsymbol{p} - \widetilde{\boldsymbol{p}}), \right.\right. \\
&\qquad\qquad \left.\left. \widetilde{M}^*(\boldsymbol{q} - \widetilde{\boldsymbol{q}}), -L_1(x - \widetilde{x}), \ldots, -L_m(x - \widetilde{x}) \right) \right\| \\
&= \mu\|x - \widetilde{x}\| + \left(\left\| \sum_{i=1}^m L_i^*(v_i - \widetilde{v}_i) \right\|^2 + \sum_{i=1}^m \Big[\|K_i(z_i - \widetilde{z}_i)\|^2 + \|M_i(y_i - \widetilde{y}_i)\|^2 \right. \\
&\qquad\qquad \left. + \|K_i^*(p_i - \widetilde{p}_i)\|^2 + \|M_i^*(q_i - \widetilde{q}_i)\|^2 + \|L_i(x - \widetilde{x})\|^2 \Big] \right)^{\frac{1}{2}} \\
&\leq \mu\|x - \widetilde{x}\| + \left(\left(\sum_{i=1}^m \|L_i\|^2 \right) \left(\|x - \widetilde{x}\|^2 + \sum_{i=1}^m \|v_i - \widetilde{v}_i\|^2 \right) + \sum_{i=1}^m \Big[\|K_i\|^2 \|z_i - \widetilde{z}_i\|^2 \right. \\
&\qquad\qquad \left. + \|M_i\|^2 \|y_i - \widetilde{y}_i\|^2 + \|K_i\|^2 \|p_i - \widetilde{p}_i\|^2 + \|M_i\|^2 \|q_i - \widetilde{q}_i\|^2 \Big] \right)^{\frac{1}{2}} \\
&\leq \left(\mu + \sqrt{\max \left\{ \sum_{i=1}^m \|L_i\|^2, \max_{j=1,\ldots,m} \left\{ \|K_j\|^2, \|M_j\|^2 \right\} \right\}} \right) \|\boldsymbol{x} - \widetilde{\boldsymbol{x}}\|. \qquad (2.22)
\end{aligned}$$

In the following we use the sequences in (2.11) for modeling summable errors in the implementation. In addition we consider the summable sequences in $\boldsymbol{\mathcal{K}}$ with terms defined for any $n \geq 0$ as

$$\boldsymbol{e}_n = (a_n, \boldsymbol{b}_n, \boldsymbol{d}_n, \boldsymbol{0}, \boldsymbol{0}, \boldsymbol{0}) \text{ and } \widetilde{\boldsymbol{e}}_n = (\boldsymbol{0}, \boldsymbol{0}, \boldsymbol{0}, \boldsymbol{e}_{1,n}, \boldsymbol{e}_{2,n}, \boldsymbol{e}_{3,n}).$$

Note that (2.20) can equivalently be written as

$$(\forall n \geq 0) \quad \left\lfloor \begin{array}{l} x_n - \gamma_n (Cx_n + \sum_{i=1}^m L_i^* v_{i,n}) \in (\mathrm{Id} + \gamma_n(-z + A))(\widetilde{x}_n - a_n) \\ \text{For } i = 1, \dots, m \\ \left\lfloor \begin{array}{l} p_{i,n} + \gamma_n K_i z_{i,n} \in (\mathrm{Id} + \gamma_n B_i^{-1})(\widetilde{p}_{i,n} - b_{i,n}) \\ q_{i,n} + \gamma_n M_i y_{i,n} \in (\mathrm{Id} + \gamma_n D_i^{-1})(\widetilde{q}_{i,n} - d_{i,n}) \\ z_{i,n} - \gamma_n K_i^* p_{i,n} = \widetilde{z}_{i,n} - \gamma_n \widetilde{v}_{i,n} - e_{1,i,n} \\ y_{i,n} - \gamma_n M_i^* q_{i,n} = \widetilde{y}_{i,n} - \gamma_n \widetilde{v}_{i,n} - e_{2,i,n} \\ v_{i,n} + \gamma_n L_i x_n = \widetilde{v}_{i,n} + \gamma_n (r_i + \widetilde{z}_{i,n} + \widetilde{y}_{i,n}) - e_{3,i,n} \end{array} \right. \\ x_{n+1} \approx \widetilde{x}_n + \gamma_n (Cx_n - C\widetilde{x}_n + \sum_{i=1}^m L_i^* (v_{i,n} - \widetilde{v}_{i,n})) \\ \text{For } i = 1, \dots, m \\ \left\lfloor \begin{array}{l} p_{i,n+1} \approx \widetilde{p}_{i,n} - \gamma_n (K_i (z_{i,n} - \widetilde{z}_{i,n})) \\ q_{i,n+1} \approx \widetilde{q}_{i,n} - \gamma_n (M_i (y_{i,n} - \widetilde{y}_{i,n})) \\ z_{i,n+1} \approx \widetilde{z}_{i,n} + \gamma_n (K_i^* (p_{i,n} - \widetilde{p}_{i,n})) \\ y_{i,n+1} \approx \widetilde{y}_{i,n} + \gamma_n (M_i^* (q_{i,n} - \widetilde{q}_{i,n})) \\ v_{i,n+1} \approx \widetilde{v}_{i,n} - \gamma_n (L_i (x_n - \widetilde{x}_n)). \end{array} \right. \end{array} \right. \tag{2.23}$$

Therefore, (2.23) is nothing else than

$$(\forall n \geq 0) \quad \left\lfloor \begin{array}{l} \boldsymbol{x}_n - \gamma_n (\boldsymbol{S} + \boldsymbol{Q}) \boldsymbol{x}_n \in (\mathrm{Id} + \gamma_n \boldsymbol{M}) (\widetilde{\boldsymbol{x}}_n - \boldsymbol{e}_n) - \widetilde{\boldsymbol{e}}_n \\ \boldsymbol{x}_{n+1} \approx \widetilde{\boldsymbol{x}}_n + \gamma_n ((\boldsymbol{S} + \boldsymbol{Q}) \boldsymbol{x}_n - (\boldsymbol{S} + \boldsymbol{Q}) \boldsymbol{p}_n). \end{array} \right. \tag{2.24}$$

We now introduce the notations

$$\boldsymbol{A}_{\boldsymbol{\mathcal{K}}} := \boldsymbol{M} \text{ and } \boldsymbol{B}_{\boldsymbol{\mathcal{K}}} := \boldsymbol{S} + \boldsymbol{Q}. \tag{2.25}$$

Then (2.24) is

$$(\forall n \geq 0) \quad \left\lfloor \begin{array}{l} \widetilde{\boldsymbol{x}}_n = J_{\gamma_n \boldsymbol{A}_{\boldsymbol{\mathcal{K}}}} (\boldsymbol{x}_n - \gamma_n \boldsymbol{B}_{\boldsymbol{\mathcal{K}}} \boldsymbol{x}_n + \widetilde{\boldsymbol{e}}_n) + \boldsymbol{e}_n \\ \boldsymbol{x}_{n+1} \approx \widetilde{\boldsymbol{x}}_n + \gamma_n (\boldsymbol{B}_{\boldsymbol{\mathcal{K}}} \boldsymbol{x}_n - \boldsymbol{B}_{\boldsymbol{\mathcal{K}}} \widetilde{\boldsymbol{x}}_n). \end{array} \right. \tag{2.26}$$

We observe that for

$$\boldsymbol{e}_n^{\boldsymbol{\mathcal{K}}} := J_{\gamma_n \boldsymbol{A}_{\boldsymbol{\mathcal{K}}}} (\boldsymbol{x}_n - \gamma_n \boldsymbol{B}_{\boldsymbol{\mathcal{K}}} \boldsymbol{x}_n + \widetilde{\boldsymbol{e}}_n) - J_{\gamma_n \boldsymbol{A}_{\boldsymbol{\mathcal{K}}}} (\boldsymbol{x}_n - \gamma_n \boldsymbol{B}_{\boldsymbol{\mathcal{K}}} \boldsymbol{x}_n) + \boldsymbol{e}_n,$$

one has $\widetilde{\boldsymbol{x}}_n = J_{\gamma_n \boldsymbol{A}_{\boldsymbol{\mathcal{K}}}} (\boldsymbol{x}_n - \gamma_n \boldsymbol{B}_{\boldsymbol{\mathcal{K}}} \boldsymbol{x}_n) + \boldsymbol{e}_n^{\boldsymbol{\mathcal{K}}}$ for any $n \geq 0$ and it holds

$$\begin{aligned} \sum_{n \geq 0} \|\boldsymbol{e}_n^{\boldsymbol{\mathcal{K}}}\| &= \sum_{n \geq 0} \|J_{\gamma_n \boldsymbol{A}_{\boldsymbol{\mathcal{K}}}} (\boldsymbol{x}_n - \gamma_n \boldsymbol{B}_{\boldsymbol{\mathcal{K}}} \boldsymbol{x}_n + \widetilde{\boldsymbol{e}}_n) - J_{\gamma_n \boldsymbol{A}_{\boldsymbol{\mathcal{K}}}} (\boldsymbol{x}_n - \gamma_n \boldsymbol{B}_{\boldsymbol{\mathcal{K}}} \boldsymbol{x}_n) + \boldsymbol{e}_n\| \\ &\leq \sum_{n \geq 0} [\|J_{\gamma_n \boldsymbol{A}_{\boldsymbol{\mathcal{K}}}} (\boldsymbol{x}_n - \gamma_n \boldsymbol{B}_{\boldsymbol{\mathcal{K}}} \boldsymbol{x}_n + \widetilde{\boldsymbol{e}}_n) - J_{\gamma_n \boldsymbol{A}_{\boldsymbol{\mathcal{K}}}} (\boldsymbol{x}_n - \gamma_n \boldsymbol{B}_{\boldsymbol{\mathcal{K}}} \boldsymbol{x}_n)\| + \|\boldsymbol{e}_n\|] \\ &\leq \sum_{n \geq 0} [\|\widetilde{\boldsymbol{e}}_n\| + \|\boldsymbol{e}_n\|] < +\infty. \end{aligned}$$

Thus, (2.26) becomes

$$(\forall n \geq 0) \quad \left\lfloor \begin{array}{l} \widetilde{\boldsymbol{x}}_n \approx J_{\gamma_n \boldsymbol{A}_{\boldsymbol{\mathcal{K}}}} (\boldsymbol{x}_n - \gamma_n \boldsymbol{B}_{\boldsymbol{\mathcal{K}}} \boldsymbol{x}_n) \\ \boldsymbol{x}_{n+1} \approx \widetilde{\boldsymbol{x}}_n + \gamma_n (\boldsymbol{B}_{\boldsymbol{\mathcal{K}}} \boldsymbol{x}_n - \boldsymbol{B}_{\boldsymbol{\mathcal{K}}} \widetilde{\boldsymbol{x}}_n), \end{array} \right. \tag{2.27}$$

which is an error-tolerant forward-backward-forward method in $\boldsymbol{\mathcal{K}}$ whose convergence has been investigated in [13]. Note that the exact version of this algorithm was proposed by Tseng in [28].

(i) By [13, Theorem 2.5(i)] we have

$$\sum_{n\geq 0}\|\boldsymbol{x}_n-\widetilde{\boldsymbol{x}}_{\boldsymbol{n}}\|^2<+\infty,$$

which yields $\sum_{n\geq 0}\|x_n-\widetilde{x}_n\|^2<+\infty$ and for any $i=1,...,m$,

$$\sum_{n\geq 0}\|p_{i,n}-\widetilde{p}_{i,n}\|^2<+\infty,\ \sum_{n\geq 0}\|q_{i,n}-\widetilde{q}_{i,n}\|^2<+\infty \text{ and } \sum_{n\geq 0}\|y_{i,n}-\widetilde{y}_{i,n}\|^2<+\infty.$$

(ii) Let $\overline{\boldsymbol{x}}=(\overline{x},\overline{\boldsymbol{p}},\overline{\boldsymbol{q}},\overline{\boldsymbol{z}},\overline{\boldsymbol{y}},\overline{\boldsymbol{v}})\in\operatorname{zer}(\boldsymbol{M}+\boldsymbol{S}+\boldsymbol{Q})$. Using [13, Theorem 2.5(ii)], we obtain $\boldsymbol{x}_n\rightharpoonup\overline{\boldsymbol{x}}$ and $\widetilde{\boldsymbol{x}}_{\boldsymbol{n}}\rightharpoonup\overline{\boldsymbol{x}}$. In consideration of (2.7), it follows that $(\overline{x},\overline{\boldsymbol{p}},\overline{\boldsymbol{q}},\overline{\boldsymbol{v}})$ is a primal-dual solution to Problem 1.2, $x_n\rightharpoonup\overline{x}$, $\widetilde{x}_n\rightharpoonup\overline{x}$, and for $i=1,...,m$

$$\left\{\begin{array}{l}p_{i,n}\rightharpoonup\overline{p}_{i,n}\\ \widetilde{p}_{i,n}\rightharpoonup\overline{p}_{i,n}\end{array}\right.,\ \left\{\begin{array}{l}q_{i,n}\rightharpoonup\overline{q}_{i,n}\\ \widetilde{q}_{i,n}\rightharpoonup\overline{q}_{i,n}\end{array}\right.,\text{ and }\left\{\begin{array}{l}y_{i,n}\rightharpoonup\overline{y}_{i,n}\\ \widetilde{y}_{i,n}\rightharpoonup\overline{y}_{i,n}\end{array}\right..$$

□

Remark 2.4. (i) In contrast to Algorithm 2.1, the iterative scheme in Algorithm 2.2 requires twice the amount of forward steps and is therefore more time-intensive. On the other hand, many steps in Algorithm 2.2 can be processed in parallel.

(ii) A forward-backward-forward type algorithm for solving the primal-dual pair of monotone inclusions (1.7) - (1.8), in the particular situation when L_i is the identity operator and $r_i=0$ for any $i=1,...,m$, has been recently investigated in [3].

3 Application to convex minimization

In this section we employ the algorithms introduced in the previous one in the context of solving the primal-dual pairs of convex optimization problems introduced in Problem 1.1.

For every $x\in\mathcal{H}$ and $(\boldsymbol{p},\boldsymbol{q})\in\boldsymbol{\mathcal{X}}\oplus\boldsymbol{\mathcal{Y}}$ with $K_i^*p_i=M_i^*q_i$, $i=1,\dots,m$, by the Young-Fenchel inequality, it holds

$$f(x)+h(x)+(f^*\,\Box\,h^*)\left(z-\sum_{i=1}^m L_i^*K_i^*p_i\right)\geq\left\langle z-\sum_{i=1}^m L_i^*K_i^*p_i,x\right\rangle$$

and, for any $i=1,\dots,m$ and $y_i\in\mathcal{G}$,

$$g_i(K_i(L_ix-r_i-y_i))+g_i^*(p_i)\geq\langle p_i,K_i(L_ix-r_i-y_i)\rangle=\langle K_i^*p_i,L_ix-r_i-y_i\rangle$$

and

$$l_i(M_iy_i)+l_i^*(q_i)\geq\langle q_i,M_iy_i\rangle=\langle M_i^*q_i,y_i\rangle.$$

This yields

$$\begin{aligned}
&\inf_{x\in\mathcal{H}}\left\{f(x)+\sum_{i=1}^m\Big((g_i\circ K_i)\,\Box\,(l_i\circ M_i)\Big)(L_ix-r_i)+h(x)-\langle x,z\rangle\right\}\\
&=\inf_{(x,\boldsymbol{y})\in\mathcal{H}\oplus\boldsymbol{\mathcal{G}}}\left\{f(x)+\sum_{i=1}^m\Big(g_i(K_i(L_ix-r_i-y_i))+l_i(M_iy_i)\Big)+h(x)-\langle x,z\rangle\right\} \qquad (3.1)\\
&\geq\sup_{\substack{(\boldsymbol{p},\boldsymbol{q})\in\boldsymbol{\mathcal{X}}\oplus\boldsymbol{\mathcal{Y}},\\ K_i^*p_i=M_i^*q_i,\ i=1,\dots,m}}\left\{-(f^*\,\Box\,h^*)\left(z-\sum_{i=1}^m L_i^*K_i^*p_i\right)-\sum_{i=1}^m\Big[g_i^*(p_i)+l_i^*(q_i)+\langle p_i,K_ir_i\rangle\Big]\right\},
\end{aligned}$$

which means that for the primal-dual pair of optimization problems (1.5)-(1.6) weak duality is always given.

Considering $(\overline{x}, \overline{\boldsymbol{p}}, \overline{\boldsymbol{q}}, \overline{\boldsymbol{y}}) \in \mathcal{H} \oplus \boldsymbol{\mathcal{X}} \oplus \boldsymbol{\mathcal{Y}} \oplus \boldsymbol{\mathcal{G}}$ a solution of the primal-dual system of monotone inclusions

$$\begin{aligned}
&z - \sum_{i=1}^{m} L_i^* K_i^* \overline{p}_i \in \partial f(\overline{x}) + \nabla h(\overline{x}) \text{ and} \\
&K_i(L_i\overline{x} - \overline{y}_i - r_i) \in \partial g_i^*(\overline{p}_i),\ M_i\overline{y}_i \in \partial l_i^*(\overline{q}_i),\ K_i^*\overline{p}_i = M_i^*\overline{q}_i,\ i = 1, \ldots, m,
\end{aligned} \tag{3.2}$$

it follows that $\overline{x}$ is an optimal solution to (1.5) and that $(\overline{\boldsymbol{p}}, \overline{\boldsymbol{q}})$ is an optimal solution to (1.6). Indeed, as h is convex and everywhere differentiable, it holds

$$z - \sum_{i=1}^{m} L_i^* K_i^* \overline{p}_i \in \partial f(\overline{x}) + \nabla h(\overline{x}) \subseteq \partial(f + h)(\overline{x}),$$

thus

$$f(\overline{x}) + h(\overline{x}) + (f^* \,\square\, h^*) \left(z - \sum_{i=1}^{m} L_i^* K_i^* \overline{p}_i \right) = \left\langle z - \sum_{i=1}^{m} L_i^* K_i^* \overline{p}_i, \overline{x} \right\rangle.$$

On the other hand, since $g_i \in \Gamma(\mathcal{X}_i)$ and $l_i \in \Gamma(\mathcal{Y}_i)$, we have for any $i = 1, \ldots, m$

$$g_i(K_i(L_i\overline{x} - \overline{y}_i - r_i)) + g_i^*(\overline{p}_i) = \langle K_i^*\overline{p}_i, L_i\overline{x} - r_i - \overline{y}_i \rangle$$

and

$$l_i(M_i\overline{y}_i) + l_i^*(\overline{q}_i) = \langle M_i^*\overline{q}_i, \overline{y}_i \rangle.$$

By summing up these equations and using (3.2), it yields

$$\begin{aligned}
& f(\overline{x}) + \sum_{i=1}^{m} \Big((g_i \circ K_i) \,\square\, (l_i \circ M_i) \Big)(L_i\overline{x} - r_i) + h(\overline{x}) - \langle \overline{x}, z \rangle \\
\leq\ & f(\overline{x}) + \sum_{i=1}^{m} \Big(g_i(K_i(L_i\overline{x} - r_i - \overline{y}_i)) + l_i(M_i\overline{y}_i) \Big) + h(\overline{x}) - \langle \overline{x}, z \rangle \\
=\ & -(f^* \,\square\, h^*) \left(z - \sum_{i=1}^{m} L_i^* K_i^* \overline{p}_i \right) - \sum_{i=1}^{m} \Big[g_i^*(\overline{p}_i) + l_i^*(\overline{q}_i) + \langle \overline{p}_i, K_i r_i \rangle \Big],
\end{aligned}$$

which, together with (3.1), leads to the desired conclusion.

In the following, by extending the result in [3, Proposition 4.2] to our setting, we provide sufficient conditions which guarantee the validity of (2.5) when applied to convex minimization problems. To this end we mention that the *strong quasi-relative interior* of a nonempty convex set $\Omega \subseteq \mathcal{H}$ is defined as

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

Proposition 3.1. *Suppose that the primal problem (1.5) has an optimal solution, that*

$$0 \in \operatorname{sqri}\left(\operatorname{dom}(g_i \circ K_i)^* - \operatorname{dom}(l_i \circ M_i)^* \right),\ i = 1, \ldots, m \tag{3.3}$$

and

$$0 \in \operatorname{sqri} \boldsymbol{E}, \tag{3.4}$$

where

$$\boldsymbol{E}:=\left\{\bigtimes_{i=1}^{m}\Big\{K_i(L_i(\operatorname{dom} f)-r_i-y_i)-\operatorname{dom} g_i\Big\}\times\bigtimes_{i=1}^{m}\Big\{M_iy_i-\operatorname{dom} l_i\Big\}: y_i\in\mathcal{G}_i,\ i=1,...,m\right\}.$$

Then

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

Proof. Let $\overline{x}\in\mathcal{H}$ be an optimal solution to (1.5). Since (3.4) holds, we have that $(g_i\circ K_i)$, $(l_i\circ M_i)\in\Gamma(\mathcal{G}_i)$, $i=1,\dots,m$. Further, because of (3.3), [2, Proposition 15.7] guarantees for any $i=1,\dots,m$ the existence of $\overline{y}_i\in\mathcal{G}_i$ such that

$$((g_i\circ K_i)\,\square\,(l_i\circ M_i))(\overline{x})=(g_i\circ K_i)(\overline{x}-\overline{y}_i)+(l_i\circ M_i)(\overline{y}_i).$$

Hence, $(\overline{x},\overline{\boldsymbol{y}})=(\overline{x},\overline{y}_1,\dots,\overline{y}_m)$ is an optimal solution to the convex optimization problem

$$\inf_{(x,\boldsymbol{y})\in\mathcal{H}\oplus\boldsymbol{\mathcal{G}}}\left\{f(x)+h(x)-\langle x,z\rangle+\sum_{i=1}^{m}\Big[g_i(K_i(L_ix-r_i-y_i))+l_i(M_iy_i)\Big]\right\}\tag{3.5}$$

By denoting

$$\begin{aligned}
&\boldsymbol{f}:\mathcal{H}\oplus\boldsymbol{\mathcal{G}}\to\overline{\mathbb{R}},\ \boldsymbol{f}(x,\boldsymbol{y})=f(x)+h(x)-\langle x,z\rangle\\
&\boldsymbol{g}:\boldsymbol{\mathcal{X}}\oplus\boldsymbol{\mathcal{Y}}\to\overline{\mathbb{R}},\ \boldsymbol{g}(\boldsymbol{x},\boldsymbol{y})=\sum_{i=1}^{m}\Big[g_i(x_i-K_ir_i)+l_i(y_i)\Big]\\
&\boldsymbol{L}:\mathcal{H}\oplus\boldsymbol{\mathcal{G}}\to\boldsymbol{\mathcal{X}}\oplus\boldsymbol{\mathcal{Y}},\ (x,\boldsymbol{y})\mapsto\bigtimes_{i=1}^{m}\Big\{K_i(L_ix-y_i)\Big\}\times\bigtimes_{i=1}^{m}\Big\{M_iy_i\Big\},
\end{aligned}\tag{3.6}$$

problem (3.5) can be equivalently written as

$$\inf_{(x,\boldsymbol{y})\in\mathcal{H}\oplus\boldsymbol{\mathcal{G}}}\left\{\boldsymbol{f}(x,\boldsymbol{y})+\boldsymbol{g}(\boldsymbol{L}(x,\boldsymbol{y}))\right\}.\tag{3.7}$$

Thus,

$$0\in\partial(\boldsymbol{f}+\boldsymbol{g}\circ\boldsymbol{L})(\overline{x},\overline{\boldsymbol{y}}).$$

Since $\boldsymbol{E}=\boldsymbol{L}(\operatorname{dom}\boldsymbol{f})-\operatorname{dom}\boldsymbol{g}$ and (3.4) is fulfilled, it holds (see, for instance, [2, 4, 7])

$$0\in\partial(\boldsymbol{f}+\boldsymbol{g}\circ\boldsymbol{L})(\overline{x},\overline{\boldsymbol{y}})=\partial\boldsymbol{f}(\overline{x},\overline{\boldsymbol{y}})+(\boldsymbol{L}^*\circ\partial\boldsymbol{g}\circ\boldsymbol{L})(\overline{x},\overline{\boldsymbol{y}}),$$

where

$$\boldsymbol{L}^*:\boldsymbol{\mathcal{X}}\oplus\boldsymbol{\mathcal{Y}}\to\mathcal{H}\oplus\boldsymbol{\mathcal{G}},\ (\boldsymbol{p},\boldsymbol{q})\mapsto\left(\sum_{i=1}^{m}L_i^*K_i^*p_i,-K_1^*p_1+M_1^*q_1,\dots,-K_m^*p_m+M_m^*q_m\right).$$

We obtain

$$\begin{aligned}
&0\in\partial\boldsymbol{f}(\overline{x},\overline{\boldsymbol{y}})+(\boldsymbol{L}^*\circ\partial\boldsymbol{g}\circ\boldsymbol{L})(\overline{x},\overline{\boldsymbol{y}})\\
\Leftrightarrow&\left\{\begin{array}{l}0\in\partial f(\overline{x})+\nabla h(\overline{x})-z+\sum_{i=1}^{m}L_i^*(K_i^*\circ\partial g_i\circ K_i)(L_i\overline{x}-r_i-\overline{y}_i)\\ 0\in-(K_i^*\circ\partial g_i\circ K_i)(L_i\overline{x}-r_i-\overline{y}_i)+(M_i^*\circ\partial l_i\circ M_i)\overline{y}_i,\ i=1,\dots,m\end{array}\right.\\
\Leftrightarrow&\ \exists\boldsymbol{v}\in\boldsymbol{\mathcal{G}}:\left\{\begin{array}{l}0\in\partial f(\overline{x})+\nabla h(\overline{x})-z+\sum_{i=1}^{m}L_i^*v_i\\ v_i\in(K_i^*\circ\partial g_i\circ K_i)(L_i\overline{x}-r_i-\overline{y}_i),\ i=1,\dots,m\\ v_i\in(M_i^*\circ\partial l_i\circ M_i)\overline{y}_i,\ i=1,\dots,m\end{array}\right.
\end{aligned}$$

$$\Leftrightarrow \exists \boldsymbol{v} \in \boldsymbol{\mathcal{G}} : \begin{cases} 0 \in \partial f(\overline{x}) + \nabla h(\overline{x}) - z + \sum_{i=1}^m L_i^* v_i \\ L_i\overline{x} - r_i - \overline{y}_i \in \left(K_i^* \circ \partial g_i \circ K_i\right)^{-1} v_i, \ i = 1, \ldots, m \\ \overline{y}_i \in \left(M_i^* \circ \partial l_i \circ M_i\right)^{-1} v_i, \ i = 1, \ldots, m \end{cases}$$

$$\Leftrightarrow \exists \boldsymbol{v} \in \boldsymbol{\mathcal{G}} : \begin{cases} 0 \in \partial f(\overline{x}) + \nabla h(\overline{x}) - z + \sum_{i=1}^m L_i^* v_i \\ v_i \in \Big(\left(K_i^* \circ \partial g_i \circ K_i\right) \Box \left(M_i^* \circ \partial l_i \circ M_i\right)\Big)(L_i\overline{x} - r_i), \ i = 1, \ldots, m \end{cases}$$

$$\Leftrightarrow z \in \partial f(\overline{x}) + \sum_{i=1}^m L_i^* \Big(\left(K_i^* \circ \partial g_i \circ K_i\right) \Box \left(M_i^* \circ \partial l_i \circ M_i\right)\Big)(L_i\overline{x} - r_i) + \nabla h(\overline{x}),$$

which completes the proof. □

Remark 3.1. If one of the following two conditions

- f is real-valued and the operators L_i, K_i and M_i are surjective for any $i = 1, \ldots, m$;
- the functions g_i and l_i are real-valued for any $i = 1, \ldots, m$;

is fulfilled, then $\boldsymbol{E} = \boldsymbol{\mathcal{X}} \oplus \boldsymbol{\mathcal{Y}}$ and (3.4) is obviously true.
On the other hand, if $\mathcal{H}$, $\mathcal{G}_i, \mathcal{X}_i$ and $\mathcal{Y}_i$, $i = 1, \ldots, m$ are finite dimensional and

$$\text{for any } i = 1, \ldots, m, \text{ there exists } y_i \in \mathcal{G}_i : \begin{cases} K_i y_i \in K_i(L_i(\operatorname{ri}\operatorname{dom} f) - r_i) - \operatorname{ri}\operatorname{dom} g_i, \\ M_i y_i \in \operatorname{ri}\operatorname{dom} l_i \end{cases},$$

then (3.4) is also true. This follows by using that in finite dimensional spaces the strong quasi-relative interior of a convex set is nothing else than its relative interior and by taking into account the properties of the latter.

3.1 An algorithm of forward-backward type

When applied to (3.2), the iterative scheme introduced in (2.4) and the corresponding convergence statements read as follows.

Algorithm 3.1.
Let $x_0 \in \mathcal{H}$, and for any $i = 1, \ldots, m$, let $p_{i,0} \in \mathcal{X}_i$, $q_{i,0} \in \mathcal{Y}_i$ and $y_{i,0}, \ z_{i,0}, \ v_{i,0} \in \mathcal{G}_i$. For any $i = 1, \ldots, m$, let τ, $\theta_{1,i}$, $\theta_{2,i}$, $\gamma_{1,i}$, $\gamma_{2,i}$ and σ_i be strictly positive real numbers such that

$$2\mu^{-1}\left(1 - \overline{\alpha}\right) \min_{i=1,\ldots,m} \left\{\frac{1}{\tau}, \frac{1}{\theta_{1,i}}, \frac{1}{\theta_{2,i}}, \frac{1}{\gamma_{1,i}}, \frac{1}{\gamma_{2,i}}, \frac{1}{\sigma_i}\right\} > 1, \tag{3.8}$$

for

$$\overline{\alpha} = \max\left\{\sqrt{\tau \sum_{i=1}^m \sigma_i \|L_i\|^2}, \max_{j=1,\ldots,m}\left\{\sqrt{\theta_{1,j}\gamma_{1,j}\|K_j\|^2}, \sqrt{\theta_{2,j}\gamma_{2,j}\|M_j\|^2}\right\}\right\}.$$

Furthermore, let $\varepsilon \in (0,1)$, $(\lambda_n)_{n\geq 0}$ be a sequence in $[\varepsilon, 1]$ and set

$$
(\forall n \geq 0) \quad \left\lfloor \begin{array}{l}
\widetilde{x}_n \approx \operatorname{Prox}_{\tau f}\left(x_n - \tau\left(Cx_n + \sum_{i=1}^m L_i^* v_{i,n} - z\right)\right) \\
\text{For } i = 1, \ldots, m \\
\left\lfloor \begin{array}{l}
\widetilde{p}_{i,n} \approx \operatorname{Prox}_{\theta_{1,i} g_i^*}\left(p_{i,n} + \theta_{1,i} K_i z_{i,n}\right) \\
\widetilde{q}_{i,n} \approx \operatorname{Prox}_{\theta_{2,i} l_i^*}\left(q_{i,n} + \theta_{2,i} M_i y_{i,n}\right) \\
u_{1,i,n} \approx z_{i,n} + \gamma_{1,i}\left(K_i^*\left(p_{i,n} - 2\widetilde{p}_{i,n}\right) + v_{i,n} + \sigma_i\left(L_i(2\widetilde{x}_n - x_n) - r_i\right)\right) \\
u_{2,i,n} \approx y_{i,n} + \gamma_{2,i}\left(M_i^*\left(q_{i,n} - 2\widetilde{q}_{i,n}\right) + v_{i,n} + \sigma_i\left(L_i(2\widetilde{x}_n - x_n) - r_i\right)\right) \\
\widetilde{z}_{i,n} \approx \frac{1+\sigma_i\gamma_{2,i}}{1+\sigma_i(\gamma_{1,i}+\gamma_{2,i})}\left(u_{1,i,n} - \frac{\sigma_i\gamma_{1,i}}{1+\sigma_i\gamma_{2,i}} u_{2,i,n}\right) \\
\widetilde{y}_{i,n} \approx \frac{1}{1+\sigma_i\gamma_{2,i}}\left(u_{2,i,n} - \sigma_i\gamma_{2,i}\widetilde{z}_{i,n}\right) \\
\widetilde{v}_{i,n} \approx v_{i,n} + \sigma_i\left(L_i(2\widetilde{x}_n - x_n) - r_i - \widetilde{z}_{i,n} - \widetilde{y}_{i,n}\right)
\end{array}\right. \\
x_{n+1} = x_n + \lambda_n(\widetilde{x}_n - x_n) \\
\text{For } i = 1, \ldots, m \\
\left\lfloor \begin{array}{l}
p_{i,n+1} = p_{i,n} + \lambda_n(\widetilde{p}_{i,n} - p_{i,n}) \\
q_{i,n+1} = q_{i,n} + \lambda_n(\widetilde{q}_{i,n} - q_{i,n}) \\
z_{i,n+1} = z_{i,n} + \lambda_n(\widetilde{z}_{i,n} - z_{i,n}) \\
y_{i,n+1} = y_{i,n} + \lambda_n(\widetilde{y}_{i,n} - y_{i,n}) \\
v_{i,n+1} = v_{i,n} + \lambda_n(\widetilde{v}_{i,n} - v_{i,n}).
\end{array}\right.
\end{array}\right. \tag{3.9}
$$

Theorem 3.2. *For the optimization problem* (1.5)*, suppose that*

$$
z \in \operatorname{ran}\left(\partial f + \sum_{i=1}^m L_i^*\left((K_i^* \circ \partial g_i \circ K_i) \,\square\, (M_i^* \circ \partial l_i \circ M_i)\right)(L_i \cdot - r_i) + \nabla h\right) \tag{3.10}
$$

and consider the sequences generated by Algorithm 3.1. Then there exists an optimal solution $\overline{x}$ to (1.5) *and optimal solution $(\overline{p}, \overline{q})$ to* (1.6) *such that*

(i) $x_n \rightharpoonup \overline{x}$, $p_{i,n} \rightharpoonup \overline{p}_i$ *and* $q_{i,n} \rightharpoonup \overline{q}_i$ *for any* $i = 1, \ldots, m$ *as* $n \to +\infty$.

(ii) if h is uniformly convex at $\overline{x}$, then $x_n \to \overline{x}$ as $n \to +\infty$.

Proof. The results is a direct consequence of Theorem 2.1 when taking

$$
A = \partial f, \;\; C = \nabla h, \;\; \text{and} \;\; B_i = \partial g_i, \;\; D_i = \partial l_i, \;\; i = 1, \ldots, m. \tag{3.11}
$$

We also notice that, according to Theorem 20.40 in [2], the operators in (3.11) are maximally monotone, while, by [2, Corollary 16.24], we have $A^{-1} = \partial f^*$, $C^{-1} = \partial h^*$, $B_i^{-1} = \partial g_i^*$ and $D_i^{-1} = \partial l_i^*$ for $i = 1, \ldots, m$. Furthermore, by [2, Corollary 18.16], $C = \nabla h$ is μ^{-1}-cocoercive, while, if h is uniformly convex at $\overline{x} \in \mathcal{H}$, then $C = \nabla h$ is uniformly monotone at $\overline{x}$ (cf. [30, Section 3.4]). □

Remark 3.2. If $h \in \Gamma(\mathcal{H})$ such that $\nabla h(x) = 0$ for all $x \in \mathcal{H}$, then condition (3.8) simplifies to

$$
\max\left\{\tau \sum_{i=1}^m \sigma_i \|L_i\|^2, \max_{j \in \mathcal{I}}\left\{\theta_{1,j}\gamma_{1,j}\|K_j\|^2, \theta_{2,j}\gamma_{2,j}\|M_j\|^2\right\}\right\} < 1.
$$

In this situation Algorithm 3.1 converges under the relaxed assumption that $(\lambda_n)_{n\geq 0} \subseteq [\varepsilon, 2-\varepsilon]$ for $\varepsilon \in (0,1)$ (see also Remark 2.2).

3.2 An algorithm of forward-backward-forward type

On the other hand, when applied to (3.2), the iterative scheme introduced in (2.20) and the corresponding convergence statements read as follows.

Algorithm 3.2.
Let $x_0 \in \mathcal{H}$, and for any $i = 1, \ldots, m$, let $p_{i,0} \in \mathcal{X}_i$, $q_{i,0} \in \mathcal{Y}_i$, and $z_{i,0},\ y_{i,0},\ v_{i,0} \in \mathcal{G}_i$. Set

$$\beta = \mu + \sqrt{\max\left\{\sum_{i=1}^{m} \|L_i\|^2, \max_{j=1,\ldots,m}\{\|K_j\|^2, \|M_j\|^2\}\right\}}, \tag{3.12}$$

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

$$(\forall n \geq 0) \quad \left\lfloor \begin{array}{l} \widetilde{x}_n \approx \operatorname{Prox}_{\gamma_n f}\left(x_n - \gamma_n\left(Cx_n + \sum_{i=1}^m L_i^* v_{i,n} - z\right)\right) \\ \text{For } i = 1, \ldots, m \\ \left\lfloor \begin{array}{l} \widetilde{p}_{i,n} \approx \operatorname{Prox}_{\gamma_n g_i^*}\left(p_{i,n} + \gamma_n K_i z_{i,n}\right) \\ \widetilde{q}_{i,n} \approx \operatorname{Prox}_{\gamma_n l_i^*}\left(q_{i,n} + \gamma_n M_i y_{i,n}\right) \\ u_{1,i,n} \approx z_{i,n} - \gamma_n\left(K_i^* p_{i,n} - v_{i,n} - \gamma_n\left(L_i x_n - r_i\right)\right) \\ u_{2,i,n} \approx y_{i,n} - \gamma_n\left(M_i^* q_{i,n} - v_{i,n} - \gamma_n\left(L_i x_n - r_i\right)\right) \\ \widetilde{z}_{i,n} \approx \frac{1+\gamma_n^2}{1+2\gamma_n^2}\left(u_{1,i,n} - \frac{\gamma_n^2}{1+\gamma_n^2} u_{2,i,n}\right) \\ \widetilde{y}_{i,n} \approx \frac{1}{1+\gamma_n^2}\left(u_{2,i,n} - \gamma_n^2 \widetilde{z}_{i,n}\right) \\ \widetilde{v}_{i,n} \approx v_{i,n} + \gamma_n\left(L_i x_n - r_i - \widetilde{z}_{i,n} - \widetilde{y}_{i,n}\right) \end{array}\right. \\ x_{n+1} \approx \widetilde{x}_n + \gamma_n(Cx_n - C\widetilde{x}_n + \sum_{i=1}^m L_i^*(v_{i,n} - \widetilde{v}_{i,n})) \\ \text{For } i = 1, \ldots, m \\ \left\lfloor \begin{array}{l} p_{i,n+1} \approx \widetilde{p}_{i,n} - \gamma_n(K_i(z_{i,n} - \widetilde{z}_{i,n})) \\ q_{i,n+1} \approx \widetilde{q}_{i,n} - \gamma_n(M_i(y_{i,n} - \widetilde{y}_{i,n})) \\ z_{i,n+1} \approx \widetilde{z}_{i,n} + \gamma_n(K_i^*(p_{i,n} - \widetilde{p}_{i,n})) \\ y_{i,n+1} \approx \widetilde{y}_{i,n} + \gamma_n(M_i^*(q_{i,n} - \widetilde{q}_{i,n})) \\ v_{i,n+1} \approx \widetilde{v}_{i,n} - \gamma_n(L_i(x_n - \widetilde{x}_n)). \end{array}\right. \end{array}\right. \tag{3.13}$$

Theorem 3.3. *For the optimization problem* (1.5)*, suppose that*

$$z \in \operatorname{ran}\left(\partial f + \sum_{i=1}^{m} L_i^*\Big((K_i^* \circ \partial g_i \circ K_i) \,\Box\, (M_i^* \circ \partial l_i \circ M_i)\Big)(L_i \cdot - r_i) + \nabla h\right), \tag{3.14}$$

and consider the sequences generated by Algorithm 3.2. Then there exists an optimal solution $\overline{x}$ *to* (1.5) *and optimal solution* $(\overline{p}, \overline{q})$ *to* (1.6) *such that*

(i) $\sum_{n\geq 0} \|x_n - \widetilde{x}_n\|^2 < +\infty$ *and for any* $i = 1,\ldots,m$

$$\sum_{n\geq 0} \|p_{i,n} - \widetilde{p}_{i,n}\|^2 < +\infty \text{ and } \sum_{n\geq 0} \|q_{i,n} - \widetilde{q}_{i,n}\|^2 < +\infty.$$

(ii) $x_n \rightharpoonup \overline{x}$, $\widetilde{x}_n \rightharpoonup \overline{x}$ *and for any* $i = 1,\ldots,m$

$$\left\{ \begin{array}{l} p_{i,n} \rightharpoonup \overline{p}_{i,n} \\ \widetilde{p}_{i,n} \rightharpoonup \overline{p}_{i,n} \end{array} \right. \text{ and } \left\{ \begin{array}{l} q_{i,n} \rightharpoonup \overline{q}_{i,n} \\ \widetilde{q}_{i,n} \rightharpoonup \overline{q}_{i,n} \end{array} \right. .$$

Proof. The conclusions follow by using the statements in the proof of Theorem 3.2 and by applying Theorem 2.2. □

4 Numerical experiments

Within this section we show how the two algorithms we propose in this paper have performed when solving the image denoising problems (1.3) and (1.4) formulated in the introductory section, namely

$$(\ell_2^2\text{-IC/P}) \qquad \inf_{x\in\mathbb{R}^n} \left\{ \frac{1}{2}\|x-b\|^2 + \Big((\alpha_1\|\cdot\|_{1,\omega_1} \circ \mathcal{D}_1) \,\square\, (\alpha_2\|\cdot\|_{1,\omega_2} \circ \mathcal{D}_2) \Big)(x) \right\}$$

and

$$(\ell_2^2\text{-MIC/P}) \qquad \inf_{x\in\mathbb{R}^n} \left\{ \frac{1}{2}\|x-b\|^2 + \Big((\alpha_1\|\cdot\|_{1,\omega_1}) \,\square\, (\alpha_2\|\cdot\|_{1,\omega_2} \circ L_1) \Big)(\mathcal{D}_1 x) \right\},$$

respectively, also in comparison with other numerical schemes from the literature.

First, we notice that for both problems a condition of type (3.3) is fulfilled, thus the infimal convolutions are proper, convex and lower semicontinuous functions. Due to the fact that the objective functions of the two optimization problems are proper, strongly convex and lower semicontinuous, these have unique optimal solutions. Finally, in the light of Remark 3.1, a condition of type (3.4) holds. Thus, according to Proposition 3.1, the hypotheses of the theorems 3.2 and 3.3 are for both optimization problems (ℓ_2^2-IC/P) and (ℓ_2^2-MIC/P) fulfilled.

In order to compare our two primal-dual iterative schemes with algorithms relying on (augmented) Lagrangian and smoothing techniques, we formulated using the definition of the infimal convolution (1.3) and (1.4) as optimization problems with constraints of the form

$$\begin{aligned} (\ell_2^2\text{-IC/P}) \qquad & \inf_{x_1,x_2,z_1,z_2} \left\{ \frac{1}{2}\|x_1+x_2-b\|^2 + \alpha_1\|z_1\|_{1,\omega_1} + \alpha_2\|z_2\|_{1,\omega_2} \right\}, \\ & \text{subject to} \begin{pmatrix} \mathcal{D}_1 & 0 \\ 0 & \mathcal{D}_2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \end{aligned} \tag{4.1}$$

and

$$\begin{aligned} (\ell_2^2\text{-MIC/P}) \qquad & \inf_{x,y_1,y_2,z} \left\{ \frac{1}{2}\|x-b\|^2 + \alpha_1\|y_1\|_{1,\omega_1} + \alpha_2\|z\|_{1,\omega_2} \right\}, \\ & \text{subject to} \begin{pmatrix} \mathcal{D}_1 & -\mathrm{Id} \\ 0 & L_1 \end{pmatrix} \begin{pmatrix} x \\ y_2 \end{pmatrix} = \begin{pmatrix} y_1 \\ z \end{pmatrix} \end{aligned} \tag{4.2}$$

respectively.

We performed our numerical tests on the colored test image lichtenstein (see Figure 4.1) of size 256×256 making each color ranging in the closed interval from 0 to 1. By adding white Gaussian noise with standard deviation 0.08, we obtained the noisy image $b \in \mathbb{R}^n$. We took $\omega_1 = (1,1)$ and $\omega_2 = (1,1)$, the regularization parameters in (ℓ_2^2-IC/P) and (ℓ_2^2-MIC/P) were set to $\alpha_1 = 0.06$ and $\alpha_2 = 0.2$, while the tests were made on an Intel Core i7-3770 processor.

When measuring the quality of the restored images, we used the improvement in signal-to-noise ratio (ISNR), which is given by

$$\mathrm{ISNR}_k = 10\log_{10}\left(\frac{\|x-b\|^2}{\|x-x_k\|^2}\right),$$

(a) Original image (b) Noisy image (c) Reconstructed image

Figure 4.1: Figure (a) shows the clean 256×256 lichtenstein test image, (b) shows the image obtained after adding white Gaussian noise with standard deviation 0.08 and (c) shows the reconstructed image.

(a) ISNR values for (ℓ_2^2-IC/P) (b) ISNR values for (ℓ_2^2-MIC/P)

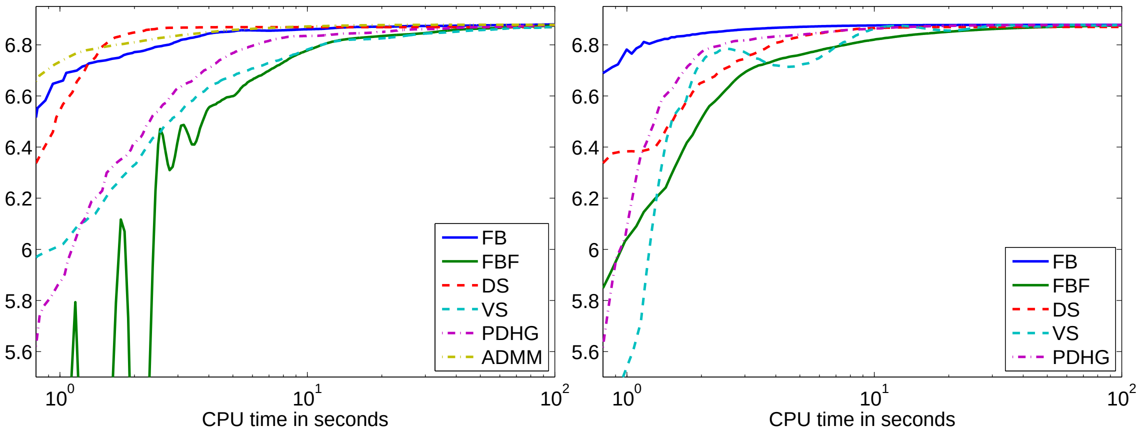

Figure 4.2: Figure (a) shows the evolution of the ISNR for the (ℓ_2^2-IC/P) problem w.r.t. the CPU times (in seconds) in log scale. Figure (b) shows the evolution of the ISNR for the (ℓ_2^2-MIC/P) problem w.r.t. the CPU times (in seconds) in log scale.

where x, b, and x_k are the original, the observed noisy and the reconstructed image at iteration $k \in \mathbb{N}$, respectively.

In Figure 4.2 we compare the performances of Algorithm 3.1 (FB) and Algorithm 3.2 (FBF) in the context of solving the optimization problems (1.3) and (1.4) to the ones of different optimization algorithms.

The double smoothing (DS) algorithm, as proposed in [11], is applied to the Fenchel dual problems of (4.1) and (4.2) by considering the acceleration strategies in [10]. One should notice that, since the smoothing parameters are constant, (DS) solves continuously differentiable approximations of (4.1) and (4.2) and does therefore not necessarily converge to the unique minimizers of (1.3) and (1.4). As a second smoothing algorithm, we considered the variable smoothing technique (VS) in [8], which successively reduces the smoothing parameter in each iteration and therefore solves the primal optimization problems as the iteration counter increases. We further considered the primal-dual hybrid gradient method (PDHG) as discussed in [27], which is nothing else than the primal-dual method in [15]. Finally, the alternating direction method of multipliers (ADMM) was applied to (4.1), as it was also done in [27]. Here, one makes use of the Moore-Penrose inverse of a special linear bounded operator which can be implemented in this setting efficiently, since $\mathcal{D}_1^T\mathcal{D}_1$ and $\mathcal{D}_2^T\mathcal{D}_2$ can be diagonalized by the discrete cosine transform. The problem which arises in (4.2), however, is far more difficult to be solved with this method (and was therefore not implemented), since the linear bounded operator assumed to be inverted has a more complicated structure. This reveals a typical drawback of ADMM given by the fact that this method does not provide a full splitting, like primal-dual or smoothing algorithms do.

As it follows from the comparisons shown in Figure 4.2, the FBF method suffers because of its additional forward step. However, many time-intensive steps in this algorithm could have been executed in parallel, which would lead to a significant decrease of the execution time. On the other hand, the FB method performs fast and stable in both examples, while optical differences in the reconstructions for (ℓ_2^2-IC/P) and (ℓ_2^2-MIC/P) are not observable.

Acknowledgements. The authors are thankful to two anonymous reviewers for comments and recommendations which improved the quality of the paper.

References

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

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

[3] S. Becker and P.L. Combettes. An algorithm for splitting parallel sums of linearly composed monotone operators, with applications to signal recovery. *J. Nonlinear Convex A.* **15**(1), 137–159, 2015.

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

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

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

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

[8] R.I. Boţ and C. Hendrich. A variable smoothing algorithm for solving convex optimization problems. *TOP* **23**(1), 124–150, 2015.

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

[10] R.I. Boţ and C. Hendrich. On the acceleration of the double smoothing technique for unconstrained convex optimization problems. *Optimization* **64**(2), 265–288, 2015.

[11] R.I. Boţ and C. Hendrich. A double smoothing technique for solving unconstrained nondifferentiable convex optimization problems. *Comput. Optim. Appl.* **54**(2), 239–262, 2013.

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

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

[14] A. Chambolle and P.-L. Lions. Image recovery via total variation minimization and related problems. *Numer. Math.* **76**(2), 167–188, 1997.

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

[16] P.L. Combettes. Quasi-Fejérian analysis of some optimization algorithms. In: D. Butnariu, Y. Censor and S. Reich (Eds.), *Inherently Parallel Algorithms in Feasibility and Optimization and their Applications*, Elsevier, New York, 115–152, 2001.

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

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

[19] P.L. Combettes. Systems of structured monotone inclusions: duality, algorithms, and applications. *SIAM J. Optim.* **23**(4), 2420–2447, 2013.

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

[21] L. Condat. A primal-dual splitting method for convex optimization involving Lipschitzian, proximable and linear composite terms. *J. Optim. Theory Appl.* **158**(2), 460–479, 2013.

[22] J. Douglas and H.H. Rachford. On the numerical solution of the heat conduction problem in two and three space variables. *Trans. of the Amer. Math. Soc.* **82**(2), 421–439, 1956.

[23] J. Eckstein and D.P. Bertsekas. On the Douglas–Rachford splitting method and the proximal point algorithm for maximal monotone operators. *Math. Program.* **55**(1–3), 293–318, 1992.

[24] S. Harizanov, J.-C. Pesquet and G. Steidl. Epigraphical projection for solving least squares anscombe transformed constrained optimization problems. In: *Scale Space and Variational Methods in Computer Vision*, Springer, Berlin Heidelberg, 125–136, 2013.

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

[26] R.T. Rockafellar. Monotone operators and the proximal point algorithm. *SIAM J. Control Optim.* **14**(5), 877–898, 1976.

[27] S. Setzer, G. Steidl and T. Teuber. Infimal convolution regularizations with discrete ℓ_1-type functionals. *Commun. Math. Sci.* **9**(3), 797–827, 2011.

[28] P. Tseng. A modified forward-backward splitting method for maximal monotone mappings. *SIAM J. Control Optim.* **38**(2), 431–446, 2000.

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

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