HAL open science

Solution of Mismatched Monotone+Lipschitz Inclusion Problems

Emilie Chouzenoux, Jean-Christophe Pesquet, Fernando Roldán

To cite this version:

Emilie Chouzenoux, Jean-Christophe Pesquet, Fernando Roldán. Solution of Mismatched Monotone+Lipschitz Inclusion Problems. SIAM Journal on Optimization, In press, 10.48550/arXiv.2310.06402 . hal-04237154

HAL Id: hal-04237154

https://inria.hal.science/hal-04237154v1

Submitted on 11 Oct 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License

arXiv:2310.06402v1 [math.OC] 10 Oct 2023

SOLUTION OF MISMATCHED MONOTONE+LIPSCHITZ INCLUSION PROBLEMS

EMILIE CHOUZENOUX†, JEAN-CHRISTOPHE PESQUET†, AND FERNANDO ROLDÁN†

Abstract. In this article, we study the convergence of algorithms for solving monotone inclusions in the presence of adjoint mismatch. The adjoint mismatch arises when the adjoint of a linear operator is replaced by an approximation, due to computational or physical issues. This occurs in inverse problems, particularly in computed tomography. In real Hilbert spaces, monotone inclusion problems involving a maximally ρ-monotone operator, a cocoercive operator, and a Lipschitzian operator can be solved by the *Forward-Backward-Half-Forward*, the *Forward-Douglas-Rachford-Forward*, and the *Forward-Half-Reflected-Backward* methods. We investigate the case of a mismatched Lipschitzian operator. We propose variants of the three aforementioned methods to cope with the mismatch, and establish conditions under which the weak convergence to a solution is guaranteed for these variants. The proposed algorithms hence enable each iteration to be implemented with a possibly iteration-dependent approximation to the mismatch operator, thus allowing this operator to be modified at each iteration. Finally, we present numerical experiments on a computed tomography example in material science, showing the applicability of our theoretical findings.

Keywords. *Splitting algorithms, convergence analysis, fixed point theory, convex optimization, adjoint mismatch*

2020 *Mathematics Subject Classification.* 47H05, 47H10, 65K05, 90C25.

1. Introduction

A rich literature exists on monotone inclusion problems formulated on a Hilbert space $\mathcal{H}$ and their deep relations with optimization, game theory, and data science (see [3, 12, 14] and the references therein). In particular, splitting approaches have turned out to play a crucial role for solving complex formulations combining monotone and linear operators. A typical monotone inclusion problem involving the sum of several operators is the following one:

Problem 1.1. *Let $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ be a maximally ρ-monotone operator for some $\rho \in \mathbb{R}$, let $C\colon \mathcal{H} \to \mathcal{H}$ be a β-cocoercive operator for some $\beta \in \left]0,+\infty\right[$, let $B : \mathcal{G} \to \mathcal{G}$ be a monotone and ζ-Lipschitzian operator for some $\zeta \in \left]0,+\infty\right[$, let $L : \mathcal{H} \to \mathcal{G}$ be a linear bounded operator, let $c \in \mathcal{G}$, and let $\alpha \in [0,+\infty[$. We want to*

$$\text{find} \quad x \in \mathcal{H} \quad \text{such that} \quad 0 \in Ax + Cx + \alpha L^*(Lx - c) + L^*BLx, \tag{1.1}$$

under the assumption that the set of solutions is nonempty.

A particular case of this problem is the following optimization one:

Problem 1.2. *Let $f\colon \mathcal{H} \to \left]-\infty,+\infty\right[$ be a proper lower-semicontinuous ρ-strongly (resp. $(-\rho)$-weakly) convex function for some $\rho \in [0,+\infty[$ (resp $\rho < 0$), let $g\colon \mathcal{H} \to \mathbb{R}$ be a differentiable convex function with a $1/\beta$-Lischitzian gradient for some $\beta \in \left]0,+\infty\right[$, let $h\colon \mathcal{H} \to \mathbb{R}$ be a*

†Université Paris-Saclay, CentraleSupélec, CVN, Inria, Gif-sur-Yvette 91190, France.
E-mail addresses: emilie.chouzenoux@inria.fr, jean-christophe.pesquet@centralesupelec.fr, fernando.roldan-contreras@centralesupelec.fr.

differentiable convex function with a ζ-Lipschitzian gradient for some $\zeta \in]0,+\infty[$, let $L : \mathcal{H} \to \mathcal{G}$ be a linear bounded operator, let $c \in \mathcal{G}$, and let $\alpha \in [0,+\infty[$. Let

$$F\colon x \mapsto f(x) + g(x) + \alpha\frac{\|Lx-c\|^2}{2} + h(Lx). \tag{1.2}$$

We want to

$$\underset{x\in\mathcal{H}}{\text{minimize}}\ \ F(x) \tag{1.3}$$

under the assumption that the set of solutions is nonempty.

Let $\partial_{\mathrm{F}} f$ denotes the Fréchet subdifferential of f. The equivalence with Problem 1.1 is obtained by setting $A = \partial_{\mathrm{F}} f$, $C = \nabla g$, $B = \nabla h$, provided that every local minimizer of F is a global minimizer. The latter condition is satisfied when F is convex (which obviously arises when $\rho \geq 0$).
Another example is the following Nash equilibrium problem [1] involving 2 players:

Problem 1.3. *Let $\mathcal{H}_1$, $\mathcal{H}_2$, $\mathcal{G}_1$, and $\mathcal{G}_2$ be real Hilbert spaces. For every $i \in \{1,2\}$, let $f_i\colon \mathcal{H}_i \to]-\infty,+\infty]$ be a proper lower-semicontinuous ρ_i-strongly (resp. $(-\rho_i)$-weakly) convex function for some $\rho_i \in [0,+\infty[$ (resp $\rho_i < 0$) and let $g_i\colon \mathcal{H}_i \to \mathbb{R}$ be a differentiable convex function with a $1/\beta_i$-Lischitzian gradient for some $\beta_i \in]0,+\infty[$. Let R be a bounded linear operator from $\mathcal{G}_2$ to $\mathcal{G}_1$ and, for every $i \in \{1,2\}$, let L_i be a linear bounded operators from $\mathcal{H}_i$ to $\mathcal{G}_i$. Let $\alpha \in]0,+\infty[$ and, for every $i \in \{1,2\}$, let Q_i be a self-adjoint linear operator from $\mathcal{G}_i$ to $\mathcal{G}_i$ such that $Q_i - \alpha\mathrm{Id}_{\mathcal{G}_i}$ is positive. Let*

$$F_1\colon (x_1,x_2) \mapsto f_1(x_1) + g_1(x_1) + \left\langle L_1x_1 \mid \frac{1}{2}Q_1L_1x_1 + RL_2x_2 \right\rangle \tag{1.4}$$

$$F_2\colon (x_1,x_2) \mapsto f_2(x_2) + g_2(x_2) + \left\langle L_2x_2 \mid \frac{1}{2}Q_2L_2x_2 - R^*L_1x_1 \right\rangle. \tag{1.5}$$

We want to find $\overline{x}_1 \in \mathcal{H}_1$ and $\overline{x}_2 \in \mathcal{H}_2$ such that

$$\overline{x}_1 = \underset{x_1\in\mathcal{H}_1}{\arg\min}\, F_1(x_1,\overline{x}_2) \tag{1.6}$$

$$\overline{x}_2 = \underset{x_2\in\mathcal{H}_2}{\arg\min}\, F_2(\overline{x}_1,x_2), \tag{1.7}$$

under the assumption that such a pair $(\overline{x}_1,\overline{x}_2)$ exists.

Assume that, for every solution $(\overline{x}_1,\overline{x}_2)$ to Problem 1.3, every local minimizer of $x_1 \mapsto F_1(x_1,\overline{x}_2)$ (resp. $x_2 \mapsto F_2(\overline{x}_1,x_2)$) is a global minimizer. For example, this condition is satisfied if, for every $i \in \{1,2\}$, $x_i \mapsto f_i(x_i) + g_i(x_i) + \frac{1}{2}\langle L_ix_i \mid Q_iL_ix_i\rangle$ is convex. Then, the above game theory problem is an instance of Problem 1.1, where $\mathcal{H} = \mathcal{H}_1 \times \mathcal{H}_2$, $\mathcal{G} = \mathcal{G}_1 \times \mathcal{G}_2$, $\rho = \min\{\rho_1,\rho_2\}$, $A = A_1 \times A_2$, $(\forall i \in \{1,2\})$ $A_i = \partial_{\mathrm{F}} f_i$, $C\colon (x_1,x_2) \mapsto (\nabla g_1(x_1), \nabla g_2(x_2))$, $B\colon (y_1,y_2) \mapsto (Q_1y_1 + Sy_2 - \alpha y_1, -S^*y_1 + Q_2y_2 - \alpha y_2)$, $L\colon (x_1,x_2) \mapsto (L_1x_1, L_2x_2)$, and $c = 0$.

In this article, we will be interested in the following relaxation of Problem 1.1:

Problem 1.4. *Let $A : \mathcal{H} \to 2^{\mathcal{H}}$ be a maximally ρ-monotone operator for some $\rho \in \mathbb{R}$, let $C : \mathcal{H} \to \mathcal{H}$ be a β-cocoercive operator for some $\beta \in]0,+\infty[$, let $B : \mathcal{G} \to \mathcal{G}$ be a monotone and ζ-Lipschitzian operator for some $\zeta \in]0,+\infty[$, let $L : \mathcal{H} \to \mathcal{G}$ and $K : \mathcal{G} \to \mathcal{H}$ be linear bounded operators, let $c \in \mathcal{G}$, and let $\alpha \in [0,+\infty[$. We want to*

$$\textit{find}\quad x \in \mathcal{H} \quad \textit{such that}\quad 0 \in Ax + Cx + \alpha K(Lx-c) + KBLx, \tag{1.8}$$

under the assumption that the set of solutions is nonempty.

This formulation arises when L^* in Problem 1.1 is replaced by some approximation K, introducing a so-called adjoint mismatch. Such a mismatch is typically encountered in variational approaches for solving inverse problems, where L models a degradation process and its adjoint often needs to be approximated due to computational or physical issues. Adjoint mismatch problems have been the topic of a number of recent works where simpler scenarios than Problem 1.4 have been considered. The importance of adjoint mismatch in computer tomography has been early recognized in [45]. Then, various methods for solving mismatched forms of Problem 1.2 have been investigated in the literature. The analysis of the quadratic case ($f = g = h = 0$) in [19, 16] is grounded on algebraic tools. In the context of the randomized Kaczmarz method, affine admissibility problems, i.e., $f = g = 0$, $\alpha = 0$, and h is the indicator function of a singleton, have been addressed in [26]. The case when $\rho \geq 0$, $g = \frac{1}{2}\|\cdot\|^2$, and $h = 0$ is investigated in [10] by focusing on the proximal gradient algorithm. As an extension of [10], a new preconditioning strategy for the proximal gradient algorithm is proposed in [36]. The case when $\rho > 0$, $g = 0$, $\alpha = 0$, and the conjuguate of h is strongly convex is analyzed in [27] by using Chambolle-Pock algorithm with fixed and varying step sizes. A similar scenario where $\rho \geq 0$, $h = 0$, and $g = \ell \circ M$, where ℓ is a convex function and M is a bounded linear operator, has been studied in [8] by considering the Condat-Vũ [15, 39], Loris-Verhoeven [29], and Combettes-Pesquet [13] primal-dual methods. Note that the convergence proofs in [8, 10, 36] rely on cocoercivity properties of the underlying operators, while this paper puts emphasis on weaker Lipschitz properties.

Since the operator $L^*(\alpha \mathrm{Id}_{\mathcal{G}} + B)L$ is monotone and Lipschitzian, the methods proposed in [6, 30, 35] can be used to solve Problem 1.1. In particular, the authors in [6] proposed a method called *forward-backward-half-forward* (FBHF), which generalizes the *forward-backward* (FB) splitting [21, 25, 33] and the *forward-backward-forward* (FBF), also called *Tseng's splitting* [38]. FBHF involves two activations of the Lipschitzian operator, one activation of C, and one application of the resolvent of A (up to some scale factor), at each iteration. In turn, the authors in [30] proposed the *forward-half-reflected-backward* (FHRB). By storing the previous iterate, this method requires one activation of the Lipschitzian operator, one activation of C, and one computation of the resolvent of A, at each iteration. FHRB reduces to FB when the Lipschitzian operator is absent. On the other hand, the Forward-Douglas–Rachford-Forward (FDRF) splitting proposed in [35] involves two activations of the Lipschitzian operator and one computation of the resolvent of A and of the resolvent of B, at each iteration. FDRF reduces to the Douglas–Rachford splitting [18, 25] when $C = 0$ and it reduces to FBF when the Lipschitzian operator is absent. When dealing with Problem 1.4, the monotonicity of the operator $K(\alpha \mathrm{Id}_G + B)L$ is not guaranteed, and the existing convergence guarantees for the previously mentioned methods collapse.

In our work, we revisit FBHF, FHRB, and FDRF by proposing variants allowing to tackle Problem 1.4 and by studying conditions guaranteeing their convergence in this context. Additionally, our analysis will be carried out in the case when, at each iteration, K is not necessarily available, but an approximation K_n of it is available. Our results are therefore of potential interest in scenarios where K_n corresponds to a learned operator, for example in neural network architectures based on the unrolling of optimization algorithms [5, 7, 37]. In our analysis, we will also provide evaluations of the error incurred by the adjoint mismatch.

The outline of the paper is as follows. In section 2, we briefly introduce the necessary notation and mathematical background. Section 3 provides preliminary results concerning Problem 1.4. We also establish two lemmas which will be useful to prove convergence results for the considered algorithms. Sections 4, 5, and 6 are dedicated to the convergence analaysis of splitting methods based on FBHF, FDRF, and FHRB, respectively, for solving Problem 1.4.

In section 7, we present a numerical comparisons of the three algorithms in an image recovery problem arising in computer tomography. Some concluding remarks are drawn in section 8.

2. Notation and Background

Throughout this paper $\mathcal{H}$ and $\mathcal{G}$ are real Hilbert spaces with scalar product $\langle \cdot \mid \cdot \rangle$ and associated norm $\|\cdot\|$. The symbols $\rightharpoonup$ and $\to$ denote the weak and strong convergence, respectively. The identity operator on $\mathcal{H}$ is denoted by $\mathrm{Id}_{\mathcal{H}}$. We denote the set of bounded linear operator from $\mathcal{H}$ to $\mathcal{G}$ by $\mathcal{B}(\mathcal{H},\mathcal{G})$. Given a linear operator $M \in \mathcal{B}(\mathcal{H},\mathcal{G})$ we denote its adjoint by $M^* \in \mathcal{B}(\mathcal{G},\mathcal{H})$. Let $D \subset \mathcal{H}$ be non-empty set and let $T\colon D \to \mathcal{H}$. The set of fixed points of T is $\operatorname{Fix} T = \big\{x \in D \;\big|\; x = Tx\big\}$. Let $\beta \in \left]0,+\infty\right[$. The operator T is $\beta-$cocoercive if

$$(\forall x \in D)(\forall y \in D) \quad \langle x - y \mid Tx - Ty\rangle \geq \beta\|Tx - Ty\|^2 \tag{2.1}$$

and it is $\beta-$Lipschitzian if

$$(\forall x \in D)(\forall y \in D) \quad \|Tx - Ty\| \leq \beta\|x - y\|. \tag{2.2}$$

When the above inequality holds, the smallest constant $\beta \in [0,+\infty[$ allowing it to be satisfied is called the Lipschitz constant of T and denoted by $\operatorname{Lip} T$. Let $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ be a set-valued operator. The domain, range, zeros, and graph of A are $\operatorname{dom} A = \big\{x \in \mathcal{H} \;\big|\; Ax \neq \varnothing\big\}$, $\operatorname{ran} A = \big\{u \in \mathcal{H} \;\big|\; (\exists x \in \mathcal{H})\; u \in Ax\big\}$, $\operatorname{zer} A = \big\{x \in \mathcal{H} \;\big|\; 0 \in Ax\big\}$, and $\operatorname{gra} A = \big\{(x,u) \in \mathcal{H}\times\mathcal{H} \;\big|\; u \in Ax\big\}$, respectively. Moreover, the inverse of A is given by $A^{-1}\colon u \mapsto \big\{x \in \mathcal{H} \;\big|\; u \in Ax\big\}$. Let $\rho \in \mathbb{R}$, the operator A is ρ-monotone if, for every $(x,u) \in \operatorname{gra} A$ and $(y,v) \in \operatorname{gra} A$ we have

$$\langle x - y \mid u - v\rangle \geq \rho\|x - y\|^2. \tag{2.3}$$

Additionally, A is maximally ρ-monotone if it is ρ-monotone and its graph is maximal in the sense of inclusions among the graphs of ρ-monotone operators. In the case when $\rho = 0$, A is (maximally) monotone, and when $\rho > 0$ A is strongly (maximally) monotone. The resolvent of a maximally ρ-monotone operator A is defined by $J_A := (\mathrm{Id} + A)^{-1}$ and, if $\rho > -1$, J_A is single valued and $(1+\rho)$-cocoercive [4, Table 1]. Note that, if A is ρ-monotone, then, for every $\gamma \in \left]0,+\infty\right[$, γA is $\gamma\rho$-monotone.

We denote by $\Gamma_0(\mathcal{H})$ the class of proper lower semicontinuous convex functions $f\colon \mathcal{H} \to \left]-\infty,+\infty\right]$. Let $f \in \Gamma_0(\mathcal{H})$. The Fenchel conjugate of f is defined by $f^*\colon u \mapsto \sup_{x\in\mathcal{H}}(\langle x \mid u\rangle - f(x))$ and we have $f^* \in \Gamma_0(\mathcal{H})$. The Moreau subdifferential of f is the maximally monotone operator $\partial f\colon x \mapsto \big\{u \in \mathcal{H} \;\big|\; (\forall y \in \mathcal{H})\;\; f(x) + \langle y - x \mid u\rangle \leq f(y)\big\}$, we have that $(\partial f)^{-1} = \partial f^*$ and that $\operatorname{zer}\partial f$ is the set of minimizers of f, which is denoted by $\arg\min_{x\in\mathcal{H}} f$.

For further properties of monotone operators, nonexpansive mappings, and convex analysis, the reader is referred to [3].

3. Preliminary results

By simple calculation, we can show that, for every $M \in \mathcal{B}(\mathcal{G},\mathcal{H})$, the operator $M(\alpha\mathrm{Id}_{\mathcal{G}}+B)L$ is Lipschitzian. Indeed,

$$(\forall (x,y) \in \mathcal{H}^2) \quad \|M(\alpha\mathrm{Id}_{\mathcal{G}} + B)Lx - M(\alpha\mathrm{Id}_{\mathcal{G}} + B)Ly\| \leq (\alpha\|M\circ L\| + \zeta\|M\|\|L\|)\|x - y\|.$$

This applies, in particular, to MBL. Henceforth, we introduce the following notation.

Notation 3.1. *In the context of Problem 1.4, for every $M \in \mathcal{B}(\mathcal{G},\mathcal{H})$, define $D_M\colon x \mapsto \alpha M(Lx - c) + MBLx$, $\kappa_M = \operatorname{Lip}(M(\alpha\mathrm{Id}_{\mathcal{G}} + B)L)$, and $\widetilde{\zeta}_M = \operatorname{Lip}(MBL)$. Let $\lambda_{\min} \in \mathbb{R}$ be defined by*

$$\lambda_{\min} = \inf\big\{\langle x \mid KLx\rangle \;\big|\; x \in \mathcal{H},\; \|x\| = 1\big\}. \tag{3.1}$$

In order to guarantee the convergence of methods for solving Problem 1.4, we introduce the following assumptions:

Assumption 3.2. *In the context of Problem 1.4 suppose that*

(i) $D_K \neq 0$,
(ii) $\hat{\rho} = \rho + \alpha\lambda_{\min} - \widetilde{\zeta}_{L^*-K} \geq 0$,
(iii) $(K_n)_{n\in\mathbb{N}}$ *is a sequence of* $\mathcal{B}(\mathcal{G},\mathcal{H})$ *such that, for every* $n \in \mathbb{N}$, $\|K_n - K\| \leq \omega_n$, *where* $\{\omega_n\}_{n\in\mathbb{N}} \subset [0,+\infty[$ *and* $\sum_{n\in\mathbb{N}} \omega_n < +\infty$.

Remark 3.3. *In the case when* $\alpha > 0$, $B = 0$, *and* A *is maximally monotone* ($\rho = 0$)*, Assumption 3.2 reduces to the monotonicity of* KL*, that is* $\lambda_{\min} \geq 0$*, which is a necessary condition for* KL *to be cocoercive* [10, Lemma 3.3]*, thus, for ensuring the convergence of cocoercive linear mismatch methods proposed in* [8, 10]*. In general, monotone linear operators are not necessarily cocoercive, for instance, consider the operator* $M\colon \mathbb{R}^2 \to \mathbb{R}^2\colon (x,y) \mapsto (-y,x)$.

Proposition 3.4. *In the context of Problem 1.4 and Assumption 3.2, the following assertions hold:*

(i) $A + D_K$ *is maximally monotone.*
(ii) $A + C + D_K$ *is maximally monotone.*
(iii) *Suppose that* $\hat{\rho} > 0$*. Then* $A + D_K$ *is* $\hat{\rho}$*-strongly monotone and* $\operatorname{zer}(A + C + D_K)$ *is a singleton.*

Proof.

(i) In view of the ρ-monotonicity of A, the definition of $\lambda_{\min}$ in (3.1), the Lipschitzianity of $(L^* - K)BL$, the monotonicity of L^*BL ([3, Proposition 20.10]), and Assumption 3.2(ii), we have, for every $\big((x,u),(y,v)\big) \in (\operatorname{gra} A)^2$,

$$\begin{aligned}
&\langle x - y \mid u + D_K x - (v + D_K y)\rangle \\
&= \langle x - y \mid u - v\rangle + \alpha\langle x - y \mid KLx - KLy\rangle \\
&\quad + \langle x - y \mid L^*BLx - L^*BLy\rangle + \langle x - y \mid (L^* - K)(BLx - BLy)\rangle \\
&\geq \langle x - y \mid u - v\rangle + \alpha\lambda_{\min}\|x - y\|^2 - \widetilde{\zeta}_{L^*-K}\|x - y\|^2 \\
&\geq \hat{\rho}\|x - y\|^2 \geq 0,
\end{aligned} \tag{3.2}$$

which shows the monotonicity of $A + D_K$. Now, by [4, Lemma 2.8] $A - \rho\mathrm{Id}_{\mathcal{H}}$ is maximally monotone, by Assumption 3.2(ii) and [3, Example 20.34] $\hat{\rho}\mathrm{Id}_{\mathcal{H}}$ is maximally monotone, by (3.1) and [3, Example 20.34] $\alpha(KL - \lambda_{\min}\mathrm{Id}_{\mathcal{H}})$ is maximally monotone, by [20, Lemma 2.12] $(K - L^*)BL + \widetilde{\zeta}_{L^*-K}\mathrm{Id}_{\mathcal{H}}$ is $(1/(2\widetilde{\zeta}_{L^*-K}))$-cocoercive with full domain, and, by [3, Corollary 25.6] and the full domain of B, L^*BL is maximally monotone. Since

$$\begin{aligned}
A + D_K =& (A - \rho\mathrm{Id}_{\mathcal{H}}) + \hat{\rho}\mathrm{Id}_{\mathcal{H}} + (\alpha KL - \alpha\lambda_{\min}\mathrm{Id}_{\mathcal{H}}) \\
&+ ((K - L^*)BL + \widetilde{\zeta}_{L^*-K}\mathrm{Id}_{\mathcal{H}}) + L^*BL - \alpha Kc,
\end{aligned} \tag{3.3}$$

the maximality of $A + D_K$ follows from [3, Corollary 25.5].

(ii) By (i), $A + D_K$ is maximally monotone. Since C is cocoercive with full domain, the operator $A + D_K + C$ is maximally monotone according to [3, Corollary 25.5].

(iii) The strong monotonicity of $A + D_K$ follows directly from (3.2). In view of the cocoercivity of C and [3, Corollary 23.37], we conclude that $\operatorname{zer}(A + C + D_K)$ is a singleton.

□

Remark 3.5.
Proposition 3.4(i) remains valid if Assumption 3.2(ii) is replaced by

$$\rho + \alpha\lambda_{\min} - \widetilde{\zeta}_K \geq 0. \tag{3.4}$$

Indeed, for every $\big((x,u),(y,v)\big) \in (\operatorname{gra} A)^2$,

$$\langle x-y \mid u + D_K x - (v + D_K y)\rangle \geq (\rho + \alpha\lambda_{\min} - \widetilde{\zeta}_K)\|x-y\|^2.$$

which shows the monotonicity of $A + D_K$. *The maximal monotonicity is deduced in the same way as at the end of the proof of Proposition 3.4(i). However, since* K *is a surrogate for operator* L^*, *it is expected that* $\widetilde{\zeta}_K \geq \widetilde{\zeta}_{L^*-K}$.

The following proposition provides an estimate of the distance between a solution to Problem 1.1 and a solution to Problem 1.4.

Proposition 3.6. *In the context of Problem 1.4, assume that* $\rho + \alpha\lambda_{\min} > 0$. *Then, there exists a unique solution* z^* *to Problem 1.1. Furthermore, every solution* z *to Problem 1.4 is such that*

$$\|z - z^*\| \leq \frac{1}{\rho + \alpha\lambda_{\min}} \|L^* - K\| \, \|\alpha(Lz - c) + BLz\|. \tag{3.5}$$

Proof. Since $\rho + \alpha\lambda_{\min} > 0$, it follows from Proposition 3.4(iii) when $K = L^*$ and the cocoercivity of C that $A + C + D_{L^*}$ is $(\rho + \alpha\lambda_{\min})$-strongly monotone and $\operatorname{zer}(A + C + D_{L^*})$ is a singleton $\{z^*\}$. Let $z \in \operatorname{zer}(A + C + D_K)$. Then

$$z = J_{A+C+D_{L^*}}(z + D_{L^*-K} z) \tag{3.6}$$

and

$$z^* = J_{A+C+D_{L^*}}(z^*). \tag{3.7}$$

Since $J_{A+C+D_{L^*}}$ is Lipschitzian with constant $1/(1 + \rho + \alpha\lambda_{\min})$ [3, Proposition 23.13], we deduce that

$$\begin{aligned}\|z - z^*\| &\leq \frac{1}{1 + \rho + \alpha\lambda_{\min}} \|z + \alpha(L^* - K)(Lz - c) + (L^* - K)BLz - z^*\| \\ &\leq \frac{1}{1 + \rho + \alpha\lambda_{\min}} \big(\|z - z^*\| + \|L^* - K\| \, \|\alpha(Lz - c) + BLz\|\big). \end{aligned} \tag{3.8}$$

The result follows from the last inequality. □

The following lemmas will play a prominent role to prove convergence properties of two of our proposed methods for solving Problem 1.4.

Lemma 3.7. *Let* $I \subset \,]0, +\infty[$ *and let* S *be a nonempty subset of* $\mathcal{H}$. *Suppose that, for every* $\gamma \in I$, $Q^\gamma : \mathcal{H} \to \mathcal{H}$ *is such that there exists a function* $\phi^\gamma \colon \mathcal{H} \to [0, +\infty[$ *satisfying, for every* $z \in \mathcal{H}$ *and* $z^* \in S$,

$$\|Q^\gamma z - z^*\|^2 \leq \|z - z^*\|^2 - \phi^\gamma(z). \tag{3.9}$$

For every $z^* \in S$, *let* $\{\varpi_n(z^*)\}_{n\in\mathbb{N}} \subset [0, +\infty[$ *and* $\{\eta_n(z^*)\}_{n\in\mathbb{N}} \subset [0, +\infty[$ *be such that* $\sum_{n\in\mathbb{N}} \varpi_n(z^*) < +\infty$ *and* $\sum_{n\in\mathbb{N}} \eta_n(z^*) < +\infty$. *For every* $n \in \mathbb{N}$ *and* $\gamma \in I$, *let* $Q_n^\gamma : \mathcal{H} \to \mathcal{H}$ *be such that*

$$(\forall z \in \mathcal{H}) \quad \|Q_n^\gamma z - Q^\gamma z\| \leq \varpi_n(z^*)\|z - z^*\| + \eta_n(z^*). \tag{3.10}$$

Let $\{\gamma_n\}_{n\in\mathbb{N}} \subset I$, *let* $z_0 \in \mathcal{H}$, *and define the sequence* $(z_n)_{n\in\mathbb{N}}$ *recursively by*

$$(\forall n \in \mathbb{N}) \quad z_{n+1} = Q_n^{\gamma_n} z_n. \tag{3.11}$$

Then, the following assertions hold:

(i) $(\|z_n - z^*\|)_{n\in\mathbb{N}}$ *is convergent.*
(ii) $\sum_{n\in\mathbb{N}} \|z_{n+1} - Q^{\gamma_n} z_n\| < +\infty$.
(iii) $\sum_{n\in\mathbb{N}} \phi^{\gamma_n}(z_n) < +\infty$.
(iv) *Suppose that every weak sequential cluster point of* $(z_n)_{n\in\mathbb{N}}$ *belongs to* S. *Then* $(z_n)_{n\in\mathbb{N}}$ *converges weakly to a point in* S.

Proof. Let $z^* \in S$.

(i) By (3.9) applied to $z = z_n$ and $\gamma = \gamma_n$ we obtain

$$\|Q^{\gamma_n} z_n - z^*\|^2 \leq \|z_n - z^*\|^2 - \phi^{\gamma_n}(z_n). \tag{3.12}$$

In particular,

$$\|Q^{\gamma_n} z_n - z^*\| \leq \|z_n - z^*\|. \tag{3.13}$$

Additionally, it follows from (3.10) that

$$\|Q_n^{\gamma_n} z_n - Q^{\gamma_n} z_n\| \leq \varpi_n(z^*)\|z_n - z^*\| + \eta_n(z^*). \tag{3.14}$$

Thus,

$$\|z_{n+1} - z^*\| \leq \|Q^{\gamma_n} z_n - z^*\| + \|Q_n^{\gamma_n} z_n - Q^{\gamma_n} z_n\| \tag{3.15}$$

$$\leq (1 + \varpi_n(z^*))\|z_n - z^*\| + \eta_n(z^*). \tag{3.16}$$

Therefore, from [3, Lemma 5.31], we conclude that $(\|z_n - z^*\|)_{n\in\mathbb{N}}$ is convergent.

(ii) We deduce from (i) that $\delta = \sup_{n\in\mathbb{N}} \|z_n - z^*\| < +\infty$. Since $(\varpi_n(z^*))_{n\in\mathbb{N}}$ and $(\eta_n(z^*))_{n\in\mathbb{N}}$ are summable sequences, we conclude from (3.14) that $\sum_{n\in\mathbb{N}} \|Q_n^{\gamma_n} z_n - Q^{\gamma_n} z_n\| < +\infty$.

(iii) By using Cauchy-Schwarz inequality, it follows from (3.12) that

$$\begin{aligned}
&\|z_{n+1} - z^*\|^2 \\
&= \|Q^{\gamma_n} z_n - z^*\|^2 + 2\langle Q^{\gamma_n} z_n - z^* \mid Q_n^{\gamma_n} z_n - Q^{\gamma_n} z_n\rangle + \|Q_n^{\gamma_n} z_n - Q^{\gamma_n} z_n\|^2 \\
&\leq \|z_n - z^*\|^2 - \phi^{\gamma_n}(z_n) + 2\|Q^{\gamma_n} z_n - z^*\|\|Q_n^{\gamma_n} z_n - Q^{\gamma_n} z_n\| + \|Q_n^{\gamma_n} z_n - Q^{\gamma_n} z_n\|^2 \\
&\leq \|z_n - z^*\|^2 - \phi^{\gamma_n}(z_n) + 2\delta\|Q_n^{\gamma_n} z_n - Q^{\gamma_n} z_n\| + \|Q_n^{\gamma_n} z_n - Q^{\gamma_n} z_n\|^2. \qquad (3.17)
\end{aligned}$$

In addition, according to (ii),

$$\sum_{n\in\mathbb{N}} 2\delta\|Q_n^{\gamma_n} z_n - Q^{\gamma_n} z_n\| + \|Q_n^{\gamma_n} z_n - Q^{\gamma_n} z_n\|^2 < +\infty. \tag{3.18}$$

Then, by invoking again [3, Lemma 5.31], we conclude that $\sum_{n\in\mathbb{N}} \phi^{\gamma_n}(z_n) < +\infty$.

(iv) Eq. (3.17) shows that $(z_n)_{n\in\mathbb{N}}$ is a quasi-Fejèr sequence with respect to S. The weak convergence of $(z_n)_{n\in\mathbb{N}}$ thus follows [3, Theorem 5.33(iv)].

□

Lemma 3.8. *Let $(\vartheta, \overline{\eta}) \in]0,1[^2$, let $\eta_0 \in [0,+\infty[$, let $\{\varpi_n\}_{n\in\mathbb{N}} \subset [0,+\infty[$ be such that $\lim_{n\to+\infty} \varpi_n = 0$, and let $\{a_n\}_{n\in\mathbb{N}} \subset [0,+\infty[$ be such that*

$$a_{n+1} \leq (\vartheta + \varpi_n)a_n + \eta_0\overline{\eta}^n. \tag{3.19}$$

Then, $(a_n)_{n\in\mathbb{N}}$ converges linearly to 0.

Proof. Since $(\varpi_n)_{n\in\mathbb{N}} \subset [0,+\infty[$ converges to zero and $\vartheta < 1$, there exist $n_0 \in \mathbb{N}$ and $\overline{\vartheta} \in]\vartheta, 1[$ such that, for every $n \geq n_0$,

$$a_{n+1} \leq \overline{\vartheta} a_n + \eta_0\overline{\eta}^n. \tag{3.20}$$

We deduce that, for every $n > n_0$,

$$a_n \leq \overline{\vartheta}^{n-n_0} a_{n_0} + \eta_0 \sum_{m=n_0}^{n-1} \overline{\eta}^m \overline{\vartheta}^{n-m-1}. \tag{3.21}$$

Without loss of generality, it can be assumed that $\overline{\vartheta} \neq \overline{\eta}$. We have then, for every $n > n_0$,

$$\begin{aligned} a_n &\leq \overline{\vartheta}^{n-n_0} a_{n_0} + \eta_0 \overline{\eta}^{n_0} \frac{\overline{\vartheta}^{n-n_0} - \overline{\eta}^{n-n_0}}{\overline{\vartheta} - \overline{\eta}} \\ &\leq \left(a_{n_0} + 2 \frac{\eta_0 \overline{\eta}^{n_0}}{|\overline{\vartheta} - \overline{\eta}|} \right) \max\{\overline{\vartheta}, \overline{\eta}\}^{n-n_0}, \end{aligned} \tag{3.22}$$

which shows the linear convergence of $(a_n)_{n\in\mathbb{N}}$ to 0. □

Lemma 3.9. *Let $I \subset]0, +\infty[$, let $z^* \in \mathcal{H}$, and let $(\vartheta, \overline{\eta}) \in]0, 1[^2$. Suppose that, for every $\gamma \in I$, $Q^\gamma \colon \mathcal{H} \to \mathcal{H}$ is such that, for every $z \in \mathcal{H}$,*

$$\|Q^\gamma z - z^*\|^2 \leq \vartheta \|z - z^*\|^2. \tag{3.23}$$

Let $\{\varpi_n(z^*)\}_{n\in\mathbb{N}} \subset [0, +\infty[$ be such that $\lim_{n\to+\infty} \varpi_n(z^*) = 0$ and let $\eta_0(z^*) \in [0, +\infty[$. For every $n \in \mathbb{N}$ and $\gamma \in I$, let $Q_n^\gamma : \mathcal{H} \to \mathcal{H}$ be such that (3.10) holds where $\eta_n(z^*) = \eta_0(z^*)\overline{\eta}^n$. Let $\{\gamma_n\}_{n\in\mathbb{N}} \subset I$ and let $z_0 \in \mathcal{H}$. Then the sequence $(z_n)_{n\in\mathbb{N}}$ defined by (3.11) converges linearly to z^*.

Proof. It follows from (3.15) that, for every $n \in \mathbb{N}$,

$$\|z_{n+1} - z^*\| \leq (\vartheta + \varpi_n(z^*))\|z_n - z^*\| + \eta_n(z^*). \tag{3.24}$$

The result then follows from Lemma 3.8. □

4. Forward-Backward-Half Forward Splitting

In this section, we will consider the following variant of the FBHF algorithm.

Algorithm 4.1. *In the context of Problem 1.4, let $\{\gamma_n\}_{n\in\mathbb{N}} \subset]0, +\infty[$ be such that $(\forall n \in \mathbb{N})$ $\gamma_n \rho > -1$, and let $z_0 \in \mathcal{H}$. Consider the iteration*

$$(\forall n \in \mathbb{N}) \quad \left\lfloor \begin{aligned} u_n &= D_{K_n} z_n \\ y_n &= z_n - \gamma_n (C z_n + u_n) \\ x_n &= J_{\gamma_n A}(y_n) \\ z_{n+1} &= x_n + \gamma_n (u_n - D_{K_n} x_n). \end{aligned} \right. \tag{4.1}$$

Notation 4.2. *In the context of Problem 1.4, for every $\gamma \in]0, +\infty[$ such that $\gamma\rho > -1$, define the operators*

$$S^\gamma = J_{\gamma A}(\mathrm{Id}_{\mathcal{H}} - \gamma(C + D_K)), \quad T^\gamma = (\mathrm{Id}_{\mathcal{H}} - \gamma D_K) \circ S^\gamma + \gamma D_K \tag{4.2}$$

and, for every $n \in \mathbb{N}$,

$$S_n^\gamma = J_{\gamma A}(\mathrm{Id}_{\mathcal{H}} - \gamma(C + D_{K_n})), \quad T_n^\gamma = (\mathrm{Id}_{\mathcal{H}} - \gamma D_{K_n}) \circ S_n^\gamma + \gamma D_{K_n}. \tag{4.3}$$

Additionally, let $\chi \in]0, \min\{2\beta, 1/\kappa_K\}[$ be defined by

$$\chi = \begin{cases} \dfrac{4\beta}{1 + \sqrt{1 + 16\beta^2\kappa_K^2}} & \text{if } \rho \geq 0 \\ \min\left\{ \dfrac{4\beta}{1 + \sqrt{1 + 16\beta^2\kappa_K^2}}, -\dfrac{1}{\rho} \right\} & \text{otherwise.} \end{cases} \tag{4.4}$$

Proposition 4.3. *In the context of Problem 1.4 and Assumption 3.2, let $\gamma \in [\varepsilon, \chi - \varepsilon]$, for some $\varepsilon \in]0, \chi/2[$. Then, the following assertions hold:*

(i) $\mathrm{zer}(A + C + D_K) = \mathrm{Fix}\, T^\gamma$

(ii) *For every* $z \in \mathcal{H}$ *and every* $z^* \in \operatorname{Fix} T^\gamma$

$$\|T^\gamma z - z^*\|^2 \leq \|z - z^*\|^2 - \kappa_K^2 \varepsilon^2 \|z - S^\gamma z\|^2 - \frac{2\beta\varepsilon^2}{\chi}\|Cz - Cz^*\|^2. \tag{4.5}$$

(iii) *Suppose that* $\hat{\rho} = \rho + \alpha\lambda_{\min} - \widetilde{\zeta}_{L^*-K} > 0$. *Then* $\operatorname{Fix} T^\gamma$ *is a singleton* $\{z^*\}$ *and, for every* $z \in \mathcal{H}$,

$$\|T^\gamma z - z^*\| \leq \sqrt{1 - \varepsilon \min\{\kappa_K^2 \varepsilon/2, \hat{\rho}\}}\, \|z - z^*\|. \tag{4.6}$$

Proof.

(i) The property directly follows from the Lipschitzian property of D_K and [6, Proposition 2.1.1].

(ii) Note that, if $z^* \in \operatorname{zer}(A + C + D_K)$, then $-\gamma C z^* \in \gamma(A + D_K)z^*$. Additionally, by defining $y = z - \gamma(C + D_K)z$ and $x = S^\gamma z = J_{\gamma A} y$, we have $y - x + \gamma D_K x \in \gamma(A + D_K)x$. Therefore, the monotonicity of $A + D_K$ established in Proposition 3.4(i) yields

$$0 \leq \langle x - z^* \mid y - x + \gamma D_K x + \gamma C z^* \rangle \tag{4.7}$$

and we deduce that

$$\begin{aligned}\langle x - z^* \mid x - y - \gamma D_K x \rangle &= \langle x - z^* \mid \gamma C z^* \rangle - \langle x - z^* \mid y - x + \gamma D_K x + \gamma C z^* \rangle \\ &\leq \langle x - z^* \mid \gamma C z^* \rangle. \end{aligned} \tag{4.8}$$

By proceeding similarly to the proof of [6, Proposition 2.1.3],

$$\begin{aligned}&\|T^\gamma z - z^*\|^2 \\ &= \|x - z^* + \gamma(D_K z - D_K x)\|^2 \\ &\leq \|x - z^*\|^2 + 2\gamma\langle x - z^* \mid C z^* \rangle + 2\langle x - z^* \mid z - x - \gamma C z \rangle + \gamma^2 \|D_K z - D_K x\|^2 \\ &= \|z - z^*\|^2 - \|z - x\|^2 + 2\gamma\langle x - z^* \mid C z^* - C z \rangle + \gamma^2 \|D_K z - D_K x\|^2 \end{aligned} \tag{4.9}$$

By using the cocoercivity of C, for every $\eta \in]0, +\infty[$,

$$\begin{aligned}2\gamma\langle x - z^* \mid C z^* - C z \rangle &\leq 2\gamma\langle x - z \mid C z^* - C z \rangle - 2\gamma\beta\|C z^* - C z\|^2 \\ &\leq \eta\|x - z\|^2 + \gamma\left(\frac{\gamma}{\eta} - 2\beta\right)\|C z^* - C z\|^2. \end{aligned} \tag{4.10}$$

Combining (4.9), (4.10), and using the fact that D_K is κ_K-Lipschitz leads to

$$\begin{aligned}&\|T^\gamma z - z^*\|^2 \\ &\leq \|z - z^*\|^2 - (1 - \eta - \gamma^2\kappa_K^2)\|z - x\|^2 - \gamma\left(2\beta - \frac{\gamma}{\eta}\right)\|C z^* - C z\|^2. \end{aligned} \tag{4.11}$$

Let us choose $\eta < 1$ such that $\chi_0 = \frac{\sqrt{1-\eta}}{\kappa_K} = 2\beta\eta$ where $\chi_0 = 4\beta/(1 + \sqrt{1 + 16\beta^2\kappa_K^2})$. It follows from (4.11) that

$$\begin{aligned}&\|T^\gamma z - z^*\|^2 \\ &\leq \|z - z^*\|^2 - \kappa_K^2(\chi_0^2 - \gamma^2)\|z - x\|^2 - 2\beta\gamma\left(1 - \frac{\gamma}{\chi_0}\right)\|C z^* - C z\|^2. \end{aligned} \tag{4.12}$$

By observing that $\chi \leq \chi_0$ and taking into account the domain of variations of γ, (4.5) is deduced.

(iii) From (i) and Proposition 3.4(ii), we conclude that $\operatorname{Fix} T^\gamma$ is a singleton. The strong monotonicity of $A + D_K$ allows us to obtain the following inequality:

$$\gamma\hat{\rho}\|x - z^*\|^2 \leq \langle x - z^* \mid y - x + \gamma D_K x + \gamma C z^* \rangle. \tag{4.13}$$

Hence, by proceeding similarly to the proof of (ii), we obtain

$$\|T^\gamma z - z^*\|^2 \leq \|z - z^*\|^2 - \kappa_K^2\varepsilon^2\|z - S^\gamma z\|^2 - \frac{2\beta\varepsilon^2}{\chi}\|Cz - Cz^*\|^2 - 2\hat{\rho}\gamma\|x - z^*\|^2. \tag{4.14}$$

Therefore, since $\gamma \geq \varepsilon$,

$$\begin{aligned}\|T^\gamma z - z^*\|^2 \leq& \|z - z^*\|^2 - \kappa_K^2\varepsilon^2\|z - S^\gamma z\|^2 - 2\hat{\rho}\varepsilon\|S^\gamma z - z^*\|^2\\ &\leq \|z - z^*\|^2 - \min\{\kappa_K^2\varepsilon^2, 2\hat{\rho}\varepsilon\}(\|z - S^\gamma z\|^2 + \|S^\gamma z - z^*\|^2)\\ &\leq \|z - z^*\|^2 - \frac{\varepsilon}{2}\min\{\kappa_K^2\varepsilon, 2\hat{\rho}\}(\|z - z^*\|^2)\\ &= \left(1 - \varepsilon\min\{\kappa_K^2\varepsilon/2, \hat{\rho}\}\right)\|z - z^*\|^2.\end{aligned}$$

□

Proposition 4.4. *Consider the operators defined in (4.2) and (4.3). Then, there exists $(\theta_1, \theta_2, \theta_3, \theta_3', \theta_4, \theta_4', \theta_4'') \in \,]0, +\infty[^7$ such that, for every $(z, z^*) \in \mathcal{H}^2$, for every $\gamma \in]0, \chi[$, and for every $n \in \mathbb{N}$, the following inequalities hold:*

(i) $\|D_{K_n} z - D_K z\| \leq \omega_n(\theta_1\|z - z^*\| + \|(\alpha\mathrm{Id}_{\mathcal{G}} + B)Lz^*\|)$
(ii) $\|S_n^\gamma z - S^\gamma z\| \leq \frac{1}{1+\gamma\rho}\|D_{K_n} z - D_K z\|$
(iii) $\|S_n^\gamma z - S^\gamma z^*\| \leq \frac{1}{1+\gamma\rho}(\theta_2\|z - z^*\| + \|D_{K_n} z - D_K z\|)$
(iv) $\|D_{K_n} S_n^\gamma z - D_K S^\gamma z\| \leq \omega_n\left(\frac{\theta_3}{1+\gamma\rho}\|z - z^*\| + \|(\alpha\mathrm{Id}_{\mathcal{G}} + B)LS^\gamma z^*\| + \frac{\theta_3'}{1+\rho\gamma}\|(\alpha\mathrm{Id}_{\mathcal{G}} + B)Lz^*\|\right)$
(v) $\|T_n^\gamma z - T^\gamma z\| \leq \frac{\omega_n}{1+\gamma\rho}(\theta_4\|z - z^*\| + \theta_4'\|(\alpha\mathrm{Id}_{\mathcal{G}} + B)LS^\gamma z^*\| + \theta_4''\|(\alpha\mathrm{Id}_{\mathcal{G}} + B)Lz^*\|)$.

Proof. First note that, in view of Assumption 3.2, $\overline{\omega} = \sup_{n\in\mathbb{N}}\omega_n < +\infty$. Let $(z, z^*) \in \mathcal{H}^2$ and let $n \in \mathbb{N}$.

(i) It follows from Assumption 3.2(iii) that

$$\begin{aligned}\|D_{K_n} z - D_K z\| &= \|K_n(\alpha\mathrm{Id}_{\mathcal{G}} + B)Lz - K(\alpha\mathrm{Id}_{\mathcal{G}} + B)Lz\|\\ &\leq \|K_n - K\|\|(\alpha\mathrm{Id}_{\mathcal{G}} + B)Lz\|\\ &\leq \omega_n\|(\alpha\mathrm{Id}_{\mathcal{G}} + B)Lz\|\\ &\leq \omega_n(\|(\alpha\mathrm{Id}_{\mathcal{G}} + B)Lz - (\alpha\mathrm{Id}_{\mathcal{G}} + B)Lz^*\| + \|(\alpha\mathrm{Id}_{\mathcal{G}} + B)Lz^*\|)\\ &\leq \omega_n((\alpha + \zeta)\|L\|\|z - z^*\| + \|(\alpha\mathrm{Id}_{\mathcal{G}} + B)Lz^*\|).\end{aligned}$$

The result follows by setting

$$\theta_1 = (\alpha + \zeta)\|L\|. \tag{4.15}$$

(ii) It follows the $(1 + \gamma\rho)^{-1}$-Lipschitzianity of $J_{\gamma A}$ that

$$\begin{aligned}\|S_n^\gamma z - S^\gamma z\| &= \|J_{\gamma A}(\mathrm{Id}_{\mathcal{H}} - \gamma(C + D_{K_n}))z - J_{\gamma A}(\mathrm{Id}_{\mathcal{H}} - \gamma(C + D_K))z\|\\ &\leq \frac{1}{1 + \rho\gamma}\|(\mathrm{Id}_{\mathcal{H}} - \gamma(C + D_{K_n}))z - (\mathrm{Id}_{\mathcal{H}} - \gamma(C + D_K))z\|\\ &= \frac{1}{1 + \rho\gamma}\|D_{K_n} z - D_K z\|.\end{aligned}$$

(iii) Similarly, it follows from (ii) and the Lipschitzianity of $J_{\gamma A}$ that

$$\begin{aligned}\|S_n^\gamma z - S^\gamma z^*\| &\leq \|S^\gamma z - S^\gamma z^*\| + \|S_n^\gamma z - S^\gamma z\| \\ &\leq \frac{1}{1+\gamma\rho}(\|(\mathrm{Id}_{\mathcal{H}} - \gamma(C+D_K))z - (\mathrm{Id}_{\mathcal{H}} - \gamma(C+D_K))z^*\| + \|D_{K_n}z - D_K z\|) \\ &\leq \frac{1}{1+\gamma\rho}\big((1+\gamma(\beta^{-1}+\kappa_K))\|z-z^*\| + \|D_{K_n}z - D_K z\|\big).\end{aligned}$$

The conclusion follows by defining $\theta_2 = 1 + \chi(\beta^{-1} + \kappa_K)$.

(iv) It follows from (i), the Lipschitzian property of D_K, (iii), and (ii) that

$$\begin{aligned}&\|D_{K_n}S_n^\gamma z - D_K S^\gamma z\| \\ &= \|D_{K_n}S_n^\gamma z - D_K S_n^\gamma z + D_K S_n^\gamma z - D_K S^\gamma z\| \\ &\leq \|D_{K_n}S_n^\gamma z - D_K S_n^\gamma z\| + \|D_K S_n^\gamma z - D_K S^\gamma z\| \\ &\leq \omega_n(\theta_1\|S_n^\gamma z - S^\gamma z^*\| + \|(\alpha\mathrm{Id}_{\mathcal{G}} + B)LS^\gamma z^*\|) + \kappa_K\|S_n^\gamma z - S^\gamma z\| \\ &\leq \omega_n\Big(\frac{\theta_1}{1+\rho\gamma}(\theta_2\|z-z^*\| + \|D_{K_n}z - D_K z\|) + \|(\alpha\mathrm{Id}_{\mathcal{G}} + B)LS^\gamma z^*\|\Big) + \frac{\kappa_K}{1+\rho\gamma}\|D_{K_n}z - D_K z\| \\ &\leq \omega_n\left(\frac{\theta_1\theta_2}{1+\rho\gamma}\|z-z^*\| + \|(\alpha\mathrm{Id}_{\mathcal{G}} + B)LS^\gamma z^*\| + \frac{(\kappa_K + \theta_1\omega_n)}{1+\gamma\rho}(\theta_1\|z-z^*\| + \|(\alpha\mathrm{Id}_{\mathcal{G}} + B)Lz^*\|)\right) \\ &= \omega_n\left(\theta_1\frac{\kappa_K + \theta_1\omega_n + \theta_2}{1+\gamma\rho}\|z-z^*\| + \|(\alpha\mathrm{Id}_{\mathcal{G}} + B)LS^\gamma z^*\| + \frac{\kappa_K + \theta_1\omega_n}{1+\rho\gamma}\|(\alpha\mathrm{Id}_{\mathcal{G}} + B)Lz^*\|)\right).\end{aligned}$$

The result follows by defining $(\theta_3, \theta_3') = (\theta_1(\kappa_K + \theta_1\overline{\omega} + \theta_2), \kappa_K + \theta_1\overline{\omega})$.

(v) It follows from (i), (ii), and (iv) that

$$\begin{aligned}&\|T_n^\gamma z - T^\gamma z\| \\ &= \|(\mathrm{Id}_{\mathcal{H}} - \gamma D_{K_n})S_n^\gamma z + \gamma D_{K_n}z - (\mathrm{Id}_{\mathcal{H}} - \gamma D_K)S^\gamma z - \gamma D_K z\| \\ &\leq \|S_n^\gamma z - S^\gamma z\| + \gamma\|D_{K_n}S_n^\gamma z - D_K S^\gamma z\| + \gamma\|D_{K_n}z - D_K z\| \\ &\leq \left(\frac{1}{1+\gamma\rho} + \gamma\right)\|D_{K_n}z - D_K z\| \\ &\quad + \gamma\omega_n\left(\frac{\theta_3}{1+\gamma\rho}\|z-z^*\| + \|(\alpha\mathrm{Id}_{\mathcal{G}} + B)LS^\gamma z^*\| + \frac{\theta_3'}{1+\gamma\rho}\|(\alpha\mathrm{Id}_{\mathcal{G}} + B)Lz^*\|\right) \\ &\leq \omega_n\left(\left(\Big(\frac{1}{1+\gamma\rho} + \gamma\Big)\theta_1 + \frac{\gamma\theta_3}{1+\gamma\rho}\right)\|z-z^*\| + \gamma\|(\alpha\mathrm{Id}_{\mathcal{G}} + B)LS^\gamma z^*\|\right. \\ &\quad \left. + \Big(\frac{1+\gamma\theta_3'}{1+\gamma\rho} + \gamma\Big)\|(\alpha\mathrm{Id}_{\mathcal{G}} + B)Lz^*\|\right).\end{aligned}$$

We conclude by defining

$$\begin{aligned}\theta_4 &= ((1+\chi+\chi^2|\rho|)\theta_1 + \chi\theta) \\ \theta_4' &= \chi(1+\chi|\rho|) \\ \theta_4'' &= 1 + \chi(1+\theta_3') + \chi^2|\rho|.\end{aligned}$$

□

Theorem 4.5. *In the context of Problem 1.4 and Assumption 3.2, let $(\gamma_n)_{n\in\mathbb{N}}$ be a sequence in $[\varepsilon, \chi - \varepsilon]$, for some $\varepsilon \in]0, \chi/2[$, consider the sequence $(z_n)_{n\in\mathbb{N}}$ generated by Algorithm 4.1. Then the following hold.*

(i) *$(z_n)_{n\in\mathbb{N}}$ converges weakly to some solution to Problem 1.4.*

(ii) *If $\hat{\rho} > 0$ and there exists $\overline{\eta} \in [0,1[$ such that, for every $n \in \mathbb{N}$, $\omega_n = \omega_0\,\overline{\eta}^n$, then $(z_n)_{n\in\mathbb{N}}$ converges linearly to the unique solution to Problem 1.4.*

Proof. Let $z^* \in \operatorname{zer}(A + C + D_K)$ and, for every $\gamma \in [\varepsilon, \chi - \varepsilon]$, consider the operators S^γ, T^γ and $(S_n^\gamma)_{n\in\mathbb{N}}$, $(T_n^\gamma)_{n\in\mathbb{N}}$, defined in (4.2) and (4.3), respectively. Then, (4.1) can be reexpressed as

$$(\forall n \in \mathbb{N}) \quad x_n = S_n^{\gamma_n} z_n \text{ and } z_{n+1} = T_n^{\gamma_n} z_n. \tag{4.16}$$

(i) In view of Proposition 4.3(ii), Proposition 4.4(v), and Lemma 3.7 applied to $I = [\varepsilon, \chi - \varepsilon]$, $S = \operatorname{zer}(A + C + D_K)$, $Q^\gamma = T^\gamma$, $\phi^\gamma : z \mapsto \kappa_K^2 \varepsilon^2 \|z - S^\gamma z\|^2$, and

$$(\forall n \in \mathbb{N}) \quad \begin{cases} Q_n^\gamma = T_n^\gamma \\ \varpi_n(z^*) = \omega_n \upsilon \theta_4 \\ \eta_n(z^*) = \omega_n \upsilon(\theta_4' \|(\alpha \mathrm{Id}_{\mathcal{G}} + B) L S^\gamma z^*\| + \theta_4'' \|(\alpha \mathrm{Id}_{\mathcal{G}} + B) L z^*\|) \end{cases} \tag{4.17}$$

with

$$\upsilon = \begin{cases} 1 & \text{if } \rho \geq 0 \\ \frac{1}{1+\rho(\chi-\varepsilon)} & \text{if } \rho < 0, \end{cases} \tag{4.18}$$

$(\|z_n - z^*\|)_{n\in\mathbb{N}}$ is convergent, $\sum_{n\in\mathbb{N}} \|T_n^{\gamma_n} z_n - T^{\gamma_n} z_n\| < +\infty$, and $\sum_{n\in\mathbb{N}} \|z_n - S^{\gamma_n} z_n\|^2 < +\infty$. Moreover, by (4.16) and Proposition 4.4(i)&(ii) we obtain

$$\begin{aligned} (\forall n \in \mathbb{N}) \quad \|z_n - x_n\| &= \|z_n - S^{\gamma_n} z_n + S^{\gamma_n} z_n - S_n^{\gamma_n} z_n\| \\ &\leq \|z_n - S^{\gamma_n} z_n\| + \omega_n(\theta_1 \|z_n - z^*\| + \|(\alpha \mathrm{Id}_{\mathcal{G}} + B) L z^*\|) \\ &\leq \|z_n - S^{\gamma_n} z_n\| + \omega_n(\theta_1 \delta_z + \|(\alpha \mathrm{Id}_{\mathcal{G}} + B) L z^*\|), \end{aligned}$$

where

$$\delta_z = \sup_{n\in\mathbb{N}} \|z_n - z^*\| < +\infty. \tag{4.19}$$

Therefore

$$z_n - x_n \to 0. \tag{4.20}$$

Furthermore, by Proposition 4.4(i) and the Lipschitzianity of D_K, we have

$$\begin{aligned} \|D_{K_n} z_n - D_K x_n\| &\leq \|D_{K_n} z_n - D_K z_n\| + \|D_K z_n - D_K x_n\| \\ &\leq \omega_n(\theta_1 \|z_n - z^*\| + \|(\alpha \mathrm{Id}_{\mathcal{G}} + B) L z^*\|) + \kappa_K \|z_n - x_n\|, \end{aligned}$$

hence

$$D_{K_n} z_n - D_K x_n \to 0. \tag{4.21}$$

Now, let $\overline{z}$ be a weak cluster point of $(z_n)_{n\in\mathbb{N}}$ and let $(z_{k_n})_{n\in\mathbb{N}}$ be a subsequence such that $z_{k_n} \rightharpoonup \overline{z}$. It follows from (4.20) that $z_{k_n} - x_{k_n} \to 0$ and that $x_{k_n} \rightharpoonup \overline{z}$ and from (4.21) that $D_{K_{k_n}} z_{k_n} - D_K x_{k_n} \to 0$. Moreover, the cocoercivity of C yields $C z_{k_n} - C x_{k_n} \to 0$. In addition, for every $n \in \mathbb{N}$,

$$\begin{aligned} & x_{k_n} = S_{k_n}^{\gamma_{k_n}} z_{k_n} \\ \Leftrightarrow \quad & \frac{z_{k_n} - x_{k_n}}{\gamma_{k_n}} - (C + D_{K_{k_n}}) z_{k_n} \in A x_{k_n} \\ \Leftrightarrow \quad & \frac{z_{k_n} - x_{k_n}}{\gamma_{k_n}} - (C z_{k_n} - C x_{k_n}) - (D_{K_{k_n}} z_{k_n} - D_K x_{k_n}) \in (A + C + D_K) x_{k_n}. \end{aligned} \tag{4.22}$$

Since $\{\gamma_n\}_{n\in\mathbb{N}} \subset [\varepsilon, \chi - \varepsilon]$, the left-hand side converges strongly to 0 as $n \to +\infty$. By the weak-strong closure of the maximally monotone operator $A + C + D_K$ (see Proposition 3.4(iii) & [3, Proposition 20.38]), we conclude that $\overline{z} \in \operatorname{zer}(A + C + D_K)$. Finally, the weak convergence of $(z_n)_{n\in\mathbb{N}}$ to an element in $\operatorname{zer}(A + C + D_K)$, follows from Lemma 3.7(iv).

(ii) The result follows from Proposition 4.3(iii) and Lemma 3.9 with $I = [\varepsilon, \chi - \varepsilon]$, $S = \operatorname{zer}(A + C + D_K)$, $Q^\gamma = T^\gamma$, and

$$\begin{aligned}
&\vartheta = \sqrt{1 - \varepsilon \min\{\kappa_K^2 \varepsilon/2, \hat{\rho}\}}\\
&(\forall n \in \mathbb{N}) \quad Q_n^\gamma = T_n^\gamma\\
&(\forall n \in \mathbb{N}) \quad \varpi_n(z^*) = \omega_n \upsilon \theta_4\\
&\eta_0(z^*) = \omega_0 \upsilon(\theta_4' \|(\alpha \mathrm{Id}_{\mathcal{G}} + B) L S^\gamma z^*\| + \theta_4'' \|(\alpha \mathrm{Id}_{\mathcal{G}} + B) L z^*\|).
\end{aligned}$$

□

5. Forward-Douglas–Rachford-Forward Splitting

We will now turn our attention to the following algorithm.

Algorithm 5.1. *In the context of Problem 1.4, let $\gamma \in \,]0, +\infty[$ be such that $\gamma\rho > -1$, let $z_0 \in \mathcal{H}$, and consider the iteration*

$$(\forall n \in \mathbb{N}) \quad \left\lfloor \begin{aligned}
&x_n = J_{\gamma C} z_n\\
&w_n = D_{K_n} x_n\\
&y_n = J_{\gamma A}(2x_n - z_n - \gamma w_n)\\
&z_{n+1} = z_n + y_n - x_n - \gamma(D_{K_n} y_n - w_n).
\end{aligned} \right. \tag{5.1}$$

Notation 5.2. *In the context of Problem 1.4, for every $\gamma \in [0, +\infty[$ such that $\gamma\rho > -1$, define the operators*

$$R^\gamma = J_{\gamma A}(2J_{\gamma C} - \mathrm{Id}_{\mathcal{H}} - \gamma D_K \circ J_{\gamma C}), \quad V^\gamma = (\mathrm{Id}_{\mathcal{H}} - \gamma D_K) \circ R^\gamma + \mathrm{Id}_{\mathcal{H}} - (\mathrm{Id}_{\mathcal{H}} - \gamma D_K) \circ J_{\gamma C} \tag{5.2}$$

and, for every $n \in \mathbb{N}$,

$$R_n^\gamma = J_{\gamma A}(2J_{\gamma C} - \mathrm{Id}_{\mathcal{H}} - \gamma D_{K_n} \circ J_{\gamma C}), \quad V_n^\gamma = (\mathrm{Id}_{\mathcal{H}} - \gamma D_{K_n}) \circ R_n^\gamma + \mathrm{Id}_{\mathcal{H}} - (\mathrm{Id}_{\mathcal{H}} - \gamma D_{K_n}) \circ J_{\gamma C}. \tag{5.3}$$

Additionally, define the set

$$\Gamma = \left\{ \gamma \in \,]0, +\infty[\;\middle|\; \kappa_K^2 \gamma^2 \left(1 + \frac{\gamma}{2\beta}\right) < 1 \text{ and } \rho\gamma > -1 \right\}. \tag{5.4}$$

Note that $\Gamma \neq \varnothing$ since the involved conditions are always satisfied for γ small enough.

Proposition 5.3. *In the context of Problem 1.4 and Assumption 3.2, let $\gamma \in \Gamma$, let $\varepsilon_2 \in \,]0, +\infty[$ be such that*

$$\varepsilon_2 < \frac{1 - \kappa_K^2 \gamma^2 \left(1 + \frac{\gamma}{2\beta}\right)}{1 - \kappa_K^2 \gamma^2}, \tag{5.5}$$

and set $\varepsilon_1 = 1 - \kappa_K^2 \gamma^2 (1 + \gamma/(2\beta(1 - \varepsilon_2)))$. Then, the following assertions hold:

(i) $\operatorname{zer}(A + C + D_K) = J_{\gamma C}(\operatorname{Fix} V^\gamma)$.

(ii) *For every $z \in \mathcal{H}$ and every $z^* \in \operatorname{Fix} V^\gamma$*

$$\|V^\gamma z - z^*\|^2 \leq \|z - z^*\|^2 - \varepsilon_1 \|J_{\gamma C} z - R^\gamma z\|^2 - \frac{2\beta\varepsilon_2}{\gamma} \|J_{\gamma C} z - z + z^* - J_{\gamma C} z^*\|^2. \tag{5.6}$$

(iii) *Suppose that $\hat{\rho} > 0$. Then, for every $z \in \mathcal{H}$ and every $z^* \in \operatorname{Fix} T^\gamma$, we have*

$$\|V^\gamma z - z^*\| \leq \sqrt{1 - \frac{1}{3} \min\left\{\frac{2\beta\varepsilon_2}{\gamma}, \varepsilon_1, 2\gamma\hat{\rho}\right\}} \|z - z^*\|. \tag{5.7}$$

Proof.

(i) See [35, Lemma 4.1].

(ii) Let $z \in \mathcal{H}$ and set $x = J_{\gamma C} z$, $y = R^{\gamma} z$. Then, $2x - z - \gamma D_K x - y + \gamma D_K y \in \gamma(A + D_K)y$. Let $z^* \in \operatorname{Fix} V^{\gamma}$ and set $(x^*, u) = (J_{\gamma C} z^*, x - z + z^* - x^*)$. Since $x^* \in \operatorname{zer}(A + C + D_K)$, $x^* - z^* \in \gamma(A + D_K)x^*$. From the monotonicity of $A + D_K$ established in Proposition 3.4(i), we deduce that

$$\begin{aligned} 0 &\leq \langle y - x^* \mid 2x - z - \gamma D_K x - y + \gamma D_K y - x^* + z^* \rangle \\ &= \langle y - x^* \mid x - \gamma D_K x - y + \gamma D_K y \rangle + \langle y - x^* \mid u \rangle. \end{aligned}$$

Hence

$$\begin{aligned} 2\gamma \langle y - x^* \mid D_K x - D_K y \rangle &\leq 2\langle y - x^* \mid x - y \rangle + 2\langle y - x^* \mid u \rangle \\ &= \|x - x^*\|^2 - \|y - x^*\|^2 - \|x - y\|^2 + 2\langle y - x^* \mid u \rangle. \end{aligned} \tag{5.8}$$

We have then

$$\begin{aligned} &\|V^{\gamma} z - z^*\|^2 \\ &= \|(\mathrm{Id}_{\mathcal{H}} - \gamma D_K)y + z - (\mathrm{Id}_{\mathcal{H}} - \gamma D_K)x - z^*\|^2 \\ &= \|y - x^* + \gamma(D_K x - D_K y) - u\|^2 \\ &\leq \|x - x^*\|^2 - \|x - y\|^2 + \gamma^2 \|D_K x - D_K y\|^2 - 2\gamma \langle D_K x - D_K y \mid u \rangle + \|u\|^2. \end{aligned} \tag{5.9}$$

As C is β-cocoercive, it follows from [35, Lemma 3.2] that

$$\|x - x^*\|^2 = \|J_{\gamma C} z - J_{\gamma C} z^*\|^2 \leq \|z - z^*\|^2 - \left(1 + \frac{2\beta}{\gamma}\right) \|u\|^2. \tag{5.10}$$

We deduce from this inequality and (5.9) that

$$\begin{aligned} &\|V^{\gamma} z - z^*\|^2 \\ &\leq \|z - z^*\|^2 - \|x - y\|^2 + \gamma^2 \|D_K x - D_K y\|^2 - 2\gamma \langle D_K x - D_K y \mid u \rangle - \frac{2\beta}{\gamma} \|u\|^2 \\ &\leq \|z - z^*\|^2 - \|x - y\|^2 + \gamma^2 \left(1 + \frac{\gamma}{2\beta(1 - \varepsilon_2)}\right) \|D_K x - D_K y\|^2 - \frac{2\beta\varepsilon_2}{\gamma} \|u\|^2. \end{aligned} \tag{5.11}$$

By using the fact that D_K is κ_K-Lispchitzian, we get

$$\begin{aligned} &\|V^{\gamma} z - z^*\|^2 \\ &\leq \|z - z^*\|^2 - \left(1 - \kappa_K^2 \gamma^2 \Big(1 + \frac{\gamma}{2\beta(1 - \varepsilon_2)}\Big)\right) \|x - y\|^2 - \frac{2\beta\varepsilon_2}{\gamma} \|u\|^2, \end{aligned} \tag{5.12}$$

which yields (5.6). Condition (5.5) can be satisfied since $\gamma \in \Gamma$ and it guarantees that $\varepsilon_1 > 0$.

(iii) By Proposition 3.4(ii), $A + D_K$ is strongly monotone. Hence, similarly to (ii), we can show that

$$\begin{aligned} &2\gamma\hat{\rho}\|y - x^*\| + 2\gamma \langle y - x^* \mid Dx - Dy \rangle \\ &\qquad \leq \|x - x^*\|^2 - \|y - x^*\|^2 - \|x - y\|^2 + 2\langle y - x^* \mid u \rangle. \end{aligned} \tag{5.13}$$

and

$$\begin{aligned}
&\|V^\gamma z - z^*\|^2 \\
&\leq \|z - z^*\|^2 - \varepsilon_1\|x - y\|^2 - \frac{2\beta\varepsilon_2}{\gamma}\|x - z + z^* - x^*\|^2 - 2\gamma\hat{\rho}\|y - x^*\|^2 \\
&\leq \|z - z^*\|^2 - \min\left\{\frac{2\beta\varepsilon_2}{\gamma}, \varepsilon_1, 2\gamma\hat{\rho}\right\}(\|x - y\|^2 + \|y - x^*\|^2 + \|x - z + z^* - x^*\|^2) \\
&\leq \|z - z^*\|^2 - \frac{1}{3}\min\left\{\frac{2\beta\varepsilon_2}{\gamma}, \varepsilon_1, 2\gamma\hat{\rho}\right\}\|z - z^*\|^2. \qquad (5.14)
\end{aligned}$$

□

Proposition 5.4. *Consider the operators defined by (5.2) and (5.3). Let $\gamma \in \Gamma$. Then, there exists $(\lambda_1, \lambda_2, \lambda_3, \lambda_3', \lambda_4, \lambda_4') \in]0, +\infty[^6$ such that, for every $(z, z^*) \in \mathcal{H}^2$, for every $\gamma \in \Gamma$, and for every $n \in \mathbb{N}$, the following inequalities hold:*

(i) $\|R_n^\gamma z - R^\gamma z\| \leq \omega_n \left(\lambda_1\|z - z^*\| + \frac{\gamma}{1+\rho\gamma}\|(\alpha\mathrm{Id}_{\mathcal{G}} + B)LJ_{\gamma C}z^*\|\right)$

(ii) $\|R_n^\gamma z - R^\gamma z^*\| \leq \lambda_2\|z - z^*\| + \frac{\omega_n\gamma}{1+\rho\gamma}\|(\alpha\mathrm{Id}_{\mathcal{G}} + B)LJ_{\gamma C}z^*\|$

(iii) $\|D_{K_n}R_n^\gamma z - D_K R^\gamma z\| \leq \omega_n(\lambda_3\|z - z^*\| + \lambda_3'\|(\alpha\mathrm{Id}_{\mathcal{G}} + B)LJ_{\gamma C}z^*\| + \|(\alpha\mathrm{Id}_{\mathcal{G}} + B)LR^\gamma z^*\|)$

(iv) $\|V_n^\gamma z - V^\gamma z\| \leq \omega_n(\lambda_4\|z - z^*\| + \lambda_4'\|(\alpha\mathrm{Id}_{\mathcal{G}} + B)LJ_{\gamma C}z^*\| + \gamma\|(\alpha\mathrm{Id}_{\mathcal{G}} + B)LR^\gamma z^*\|)$.

Proof. Recall that, in view of Assumption 3.2, $\overline{\omega} = \sup_{n\in\mathbb{N}}\omega_n < +\infty$. Let $(z, z^*) \in \mathcal{H}^2$.

(i) It follows from the $(1+\gamma\rho)^{-1}$-Lipschitzianity of $J_{\gamma A}$, the nonexpansiveness of $J_{\gamma C}$, and Proposition 4.4(i) that

$$\begin{aligned}
\|R_n^\gamma z - R^\gamma z\| &= \|J_{\gamma A}(2J_{\gamma C} - \mathrm{Id}_{\mathcal{H}} - \gamma D_{K_n} \circ J_{\gamma C})z - J_{\gamma A}(2J_{\gamma C} - \mathrm{Id}_{\mathcal{H}} - \gamma D_K \circ J_{\gamma C})z\| \\
&\leq \frac{\gamma}{1+\rho\gamma}\|D_{K_n}J_{\gamma C}z - D_K J_{\gamma C}z\| \\
&\leq \frac{\gamma\omega_n}{1+\rho\gamma}(\theta_1\|J_{\gamma C}z - J_{\gamma C}z^*\| + \|(\alpha\mathrm{Id}_{\mathcal{G}} + B)LJ_{\gamma C}z^*\|) \\
&\leq \frac{\gamma\omega_n}{1+\rho\gamma}(\theta_1\|z - z^*\| + \|(\alpha\mathrm{Id}_{\mathcal{G}} + B)LJ_{\gamma C}z^*\|). \qquad (5.15)
\end{aligned}$$

The result follows by setting $\lambda_1 = \gamma\theta_1/(1+\rho\gamma)$, θ_1 being given by (4.15).

(ii) Using the nonexpansiveness of $J_{\gamma A}$, the Lipschitzianity of D_K, and the nonexpansiveness of $2J_{\gamma C} - \mathrm{Id}_H$ [3, Corollary 23.11], we deduce from (i) that

$$\begin{aligned}
&\|R_n^\gamma z - R^\gamma z^*\| \\
&\leq \|R^\gamma z - R^\gamma z^*\| + \|R_n^\gamma z - R^\gamma z\| \qquad (5.16) \\
&= \frac{1}{1+\gamma\rho}\|(2J_{\gamma C} - \mathrm{Id}_{\mathcal{H}} - \gamma D_K J_{\gamma C})z - (2J_{\gamma C} - \mathrm{Id}_{\mathcal{H}} - \gamma D_K J_{\gamma C})z^*\| + \|R_n^\gamma z - R^\gamma z\| \\
&\leq \left(\frac{1+\gamma\kappa_K}{1+\gamma\rho} + \omega_n\lambda_1\right)\|z - z^*\| + \frac{\omega_n\gamma}{1+\gamma\rho}\|(\alpha\mathrm{Id}_{\mathcal{G}} + B)LJ_{\gamma C}z^*\|. \qquad (5.17)
\end{aligned}$$

We conclude by defining

$$\lambda_2 = \overline{\omega}\lambda_1 + \frac{\gamma\kappa_K + 1}{1+\gamma\rho}. \qquad (5.18)$$

(iii) It follows from (i), (ii), and Proposition 4.4(i) that

$$\begin{aligned}
&\|D_{K_n}R_n^\gamma z - D_K R^\gamma z\| \\
&\leq \|D_{K_n}R_n^\gamma z - D_K R_n^\gamma z\| + \|D_K R_n^\gamma z - D_K R^\gamma z\| \\
&\leq \omega_n(\theta_1\|R_n^\gamma z - R^\gamma z^*\| + \|(\alpha\mathrm{Id}_{\mathcal{G}} + B)LR^\gamma z^*\|) + \kappa_K\|R_n^\gamma z - R^\gamma z\| \\
&\leq \omega_n\Big((\theta_1\lambda_2 + \kappa_K\lambda_1)\|z - z^*\| + \frac{\gamma(\omega_n\theta_1 + \kappa_K)}{1+\gamma\rho}\|(\alpha\mathrm{Id}_{\mathcal{G}} + B)LJ_{\gamma C}z^*\| \\
&\qquad\qquad + \|(\alpha\mathrm{Id}_{\mathcal{G}} + B)LR^\gamma z^*\|\Big). \qquad (5.19)
\end{aligned}$$

The result is obtained by defining $\lambda_3 = (\theta_1\lambda_2 + \kappa_K\lambda_1)$ and $\lambda_3' = \gamma(\overline{\omega}\lambda_1 + \kappa_K)/(1+\gamma\rho)$.

(iv) It follows from (i), (iii), Proposition 4.4(i), and the nonexpansiveness of $J_{\gamma C}$ that

$$\begin{aligned}
\|V_n^\gamma z - V^\gamma z\| &\leq \|R_n^\gamma z - R^\gamma z\| + \gamma\|D_{K_n}R_n^\gamma z - D_K R^\gamma z\| + \gamma\|D_{K_n}J_{\gamma C}z - D_K J_{\gamma C}z\| \\
&\leq \omega_n\Bigg((\lambda_1 + \gamma(\lambda_3 + \theta_1))\|z - z^*\| + \gamma\Big(\lambda_3' + 1 + \frac{1}{1+\gamma\rho}\Big)(\alpha\mathrm{Id}_{\mathcal{G}} + B)LJ_{\gamma C}z^*\| \\
&\qquad + \gamma\|(\alpha\mathrm{Id}_{\mathcal{G}} + B)LR^\gamma z^*\|\Bigg). \qquad (5.20)
\end{aligned}$$

This yields the sought inequality by defining $\lambda_4 = \lambda_1 + \gamma(\lambda_3 + \theta_1)$ and

$$\lambda_4' = \lambda_3' + 1 + \frac{1}{1+\gamma\rho}. \qquad (5.21)$$

□

Theorem 5.5. *In the context of Problem 1.4 and Assumption 3.2, let $\gamma \in \Gamma$, and consider the sequences $(z_n)_{n\in\mathbb{N}}$ and $(x_n)_{n\in\mathbb{N}}$ generated by Algorithm 5.1. Then the following hold.*

(i) *$(z_n)_{n\in\mathbb{N}}$ converges weakly to some $\overline{z} \in \operatorname{Fix} V^\gamma$ and $(x_n)_{n\in\mathbb{N}}$ converges weakly to $J_{\gamma C}\overline{z} \in \operatorname{zer}(A + C + D_K)$.*

(ii) *If $\hat{\rho} > 0$ and there exists $\overline{\eta} \in [0, 1[$ such that, for every $n \in \mathbb{N}$, $\omega_n = \omega_0\,\overline{\eta}^n$, then $(z_n)_{n\in\mathbb{N}}$ converges linearly to $\overline{z} \in \operatorname{Fix} V^\gamma$ and $(x_n)_{n\in\mathbb{N}}$ converges linearly to $J_{\gamma C}\overline{z}$, which is the unique solution to Problem 1.4.*

Proof. Let $\gamma \in \Gamma$. Consider the operators R^γ, V^γ and $(R_n^\gamma)_{n\in\mathbb{N}}$, $(V_n^\gamma)_{n\in\mathbb{N}}$ defined in (5.2) and (5.3), respectively. Let $x^* \in \operatorname{zer}(A + C + D_K)$. According to Proposition 5.3(i), there exists $z^* \in \operatorname{Fix} V^\gamma$ such that $x^* = J_{\gamma C}z^*$. Note that (5.1) is equivalent to

$$(\forall n \in \mathbb{N}) \quad y_n = R_n^\gamma z_n \text{ and } z_{n+1} = V_n^\gamma z_n. \qquad (5.22)$$

(i) In view of Proposition 5.3(ii) and Proposition 5.4(iv), Lemma 3.7 can be applied to $I = \{\gamma\}$, $S = \operatorname{Fix} V^\gamma$, $Q^\gamma = V^\gamma$, $\phi^\gamma : z \mapsto \varepsilon_1\|J_{\gamma C}z - R^\gamma z\|^2 + \frac{2\beta\varepsilon_2}{\gamma}\|J_{\gamma C}z - z + z^* - J_{\gamma C}z^*\|^2$, and

$$(\forall n \in \mathbb{N}) \quad \begin{cases} Q_n^\gamma = V_n^\gamma \\ \varpi_n(z^*) = \omega_n\lambda_4 \\ \eta_n(z^*) = \omega_n(\lambda_4'\|(\alpha\mathrm{Id}_{\mathcal{G}} + B)LJ_{\gamma C}z^*\| + \gamma\|(\alpha\mathrm{Id}_{\mathcal{G}} + B)LR^\gamma z^*\|). \end{cases} \qquad (5.23)$$

This allows us to deduce that $(\|z_n - z^*\|)_{n\in\mathbb{N}}$ is convergent, $\sum_{n\in\mathbb{N}}\|V_n^\gamma z_n - V^\gamma z_n\| < +\infty$, $\sum_{n\in\mathbb{N}}\|x_n - R^\gamma z_n\| < +\infty$, and $\sum_{n\in\mathbb{N}}\|x_n - z_n - x^* + z^*\| < +\infty$. Moreover,

according to (5.22) and Proposition 5.4(i),

$$\begin{aligned}\|x_n - y_n\| &= \|x_n - R^\gamma z_n + R^\gamma z_n - R_n^\gamma z_n\| \\ &\leq \|x_n - R^\gamma z_n\| + \omega_n \left(\lambda_1 \|z_n - z^*\| + \frac{\gamma}{1+\gamma\rho}\|(\alpha \mathrm{Id}_{\mathcal{G}} + B)LJ_{\gamma C} z^*\|\right) \\ &\leq \|x_n - R^\gamma z_n\| + \omega_n \left(\lambda_1 \delta_z + \frac{\gamma}{1+\gamma\rho}\|(\alpha \mathrm{Id}_{\mathcal{G}} + B)LJ_{\gamma C} z^*\|\right),\end{aligned}$$

where δ_z is given by (4.19). Therefore,

$$y_n - x_n \to 0 \tag{5.24}$$

and, it follows from the cocoercivity of C and the Lipschitzian property of D that

$$Cy_n - Cx_n \to 0 \text{ and } D_K y_n - D_K x_n \to 0. \tag{5.25}$$

Since $z_n - x_n = \gamma C x_n$, we deduce that

$$\frac{z_n - x_n}{\gamma} - Cy_n \to 0. \tag{5.26}$$

Furthermore, by Proposition 4.4(i) and the nonexpansiveness of $J_{\gamma C}$ we have

$$\begin{aligned}\|w_n - D_K y_n\| &\leq \|D_{K_n} x_n - D_K x_n\| + \|D_K y_n - D_K x_n\| \\ &\leq \omega_n(\theta_1 \|z_n - z^*\| + \|(\alpha \mathrm{Id}_{\mathcal{G}} + B)LJ_{\gamma C} z^*\|) + \kappa_K \|y_n - x_n\|.\end{aligned}$$

Thus

$$w_n - D_K y_n \to 0. \tag{5.27}$$

Now, let $\overline{z}$ be a weak cluster point of $(z_n)_{n\in\mathbb{N}}$ and let $(z_{k_n})_{n\in\mathbb{N}}$ be a subsequence such that $z_{k_n} \rightharpoonup \overline{z}$. Since $x_{k_n} - z_{k_n} \to x^* - z^*$,

$$x_{k_n} \rightharpoonup \overline{x} = \overline{z} + x^* - z^*. \tag{5.28}$$

According to (5.24), $x_{k_n} - y_{k_n} \to 0$, hence that $y_{k_n} \rightharpoonup \overline{x}$. It follows from (5.25), (5.26), and (5.27) that $Dx_{k_n} - Dy_{k_n} \to 0$, $(x_{k_n} - z_{k_n})/\gamma - Cy_{k_n} \to 0$, and $D_K y_{k_n} - w_{k_n} \to 0$. Furthermore, from (5.1),

$$\begin{aligned}&\frac{2x_{k_n} - z_{k_n} - y_{k_n}}{\gamma} - w_{k_n} \in Ay_{k_n} \\ \Leftrightarrow\ &\frac{x_{k_n} - y_{k_n}}{\gamma} - \left(\frac{z_{k_n} - x_{k_n}}{\gamma} - Cy_{k_n}\right) - (D_K x_{k_n} - D_K y_{k_n}) + (D_K x_{k_n} - w_{k_n}) \\ &\qquad \in (A + C + D_K)y_{k_n}.\end{aligned} \tag{5.29}$$

Altogether, by the weak-strong closure of the maximally monotone operator $A+C+D_K$ (see Proposition 3.4(iii) & [3, Proposition 20.38]), we conclude that $\overline{x} \in \mathrm{zer}(A + C + D_K)$. We can thus choose $x^* = \overline{x}$ and (5.28) yields $\overline{z} = z^* = J_{\gamma C}\overline{x} \in \mathrm{Fix}\, V^\gamma$. The weak convergence of $(z_n)_{n\in\mathbb{N}}$ to $\overline{z}$ follows from Lemma 3.7(iv). Finally $x_n - z_n \to \overline{x} - \overline{z}$ $\Rightarrow x_n \rightharpoonup \overline{x}$.

(ii) The linear convergence of $(z_n)_{n\in\mathbb{N}}$ to z^* follows from Proposition 5.3(ii) and Lemma 3.9 with $I = \{\gamma\}$, $S = \mathrm{Fix}\, V^\gamma$, $Q^\gamma = V^\gamma$, and

$$\vartheta = \sqrt{1 - \frac{1}{3}\min\left\{\frac{2\beta\varepsilon_2}{\gamma}, \varepsilon_1, 2\gamma\hat{\rho}\right\}} \tag{5.30}$$

$$(\forall n \in \mathbb{N}) \quad Q_n^\gamma = V_n^\gamma \tag{5.31}$$

$$(\forall n \in \mathbb{N}) \quad \varpi_n(z^*) = \omega_n \lambda_4 \tag{5.32}$$

$$\eta_0(z^*) = \omega_0(\lambda_4' \|(\alpha \mathrm{Id}_{\mathcal{G}} + B)LJ_{\gamma C} z^*\| + \gamma\|(\alpha \mathrm{Id}_{\mathcal{G}} + B)LR^\gamma z^*\|). \tag{5.33}$$

Since $J_{\gamma C}$ is nonexpansive,

$$(\forall n \in \mathbb{N}) \quad \|x_n - x^*\| \leq \|z_n - z^*\|.$$

This shows that $(x_n)_{n\in\mathbb{N}}$ converges linearly to x^*, which is the unique solution to Problem 1.4.

□

6. Forward-Half-Reflected-Backward Splitting

Finally, in this section, we will study the following algorithm.

Algorithm 6.1. *In the context of Problem 1.4, let $\gamma \in]0,+\infty[$ be such that $\gamma\rho > -1$, let $(z_{-1}, z_0) \in \mathcal{H}^2$, set $y_{-1} = D_{K_0} z_{-1}$, and consider the iteration*

$$(\forall n \in \mathbb{N}) \quad \left\lfloor \begin{array}{l} y_n = D_{K_{n+1}} z_n \\ x_n = z_n - \gamma(2y_n - y_{n-1} + Cz_n) \\ z_{n+1} = J_{\gamma A} x_n. \end{array} \right. \tag{6.1}$$

Notation 6.2. *In the context of Problem 1.4, for every $\gamma \in]0,+\infty[$ such that $\gamma\rho > -1$, and for every $n \in \mathbb{N}$, define the operator*

$$M_n^\gamma \colon \mathcal{H}^2 \to \mathcal{H} \colon (z,w) \mapsto M_n^\gamma(z,w) = J_{\gamma A}(z - \gamma(2D_{K_{n+1}} z - D_{K_n} w + Cz)). \tag{6.2}$$

Additionally, let

$$\chi' = \begin{cases} \dfrac{2\beta}{4\beta\kappa_K + 1} & \text{if } \rho \geq 0, \\ \min\left\{\dfrac{2\beta}{4\beta\kappa_K + 1}, -\dfrac{1}{\rho}\right\} & \text{if } \rho < 0. \end{cases} \tag{6.3}$$

Proposition 6.3. *In the context of Problem 1.4 and Assumption 3.2, let $\gamma \in]0, \chi'[$. Then, the following hold.*

(i) *There exists a sequence $(\epsilon_n)_{n\in\mathbb{N}}$ of Lipschitz constants of $(\gamma D_{K_n})_{n\in\mathbb{N}}$ such that*

$$(\exists N_0 \in \mathbb{N})(\forall n \geq N_0) \quad \epsilon_n \in [\delta, 1-\delta] \tag{6.4}$$

with $\delta \in]0, 1/2[$.

(ii) *There exists $\{\Upsilon_n\}_{n\in\mathbb{N}} \subset [0,+\infty[$ such that $\sum_{n\in\mathbb{N}} \Upsilon_n < +\infty$ and, for every $n \geq N_0$, for every $(z,w) \in \mathcal{H}^2$, and for every $z^* \in \operatorname{zer}(A + C + D_K)$,*

$$\begin{aligned} &2\gamma|\langle D_{K_{n+1}} z - D_{K_n} z + D_{K_{n+1}} z_+ - D_K z_+ \mid z_+ - z^* \rangle| \\ &\quad \leq \Upsilon_n(\|\xi_n^\gamma(z,w) - z^*\|^2 + \epsilon_n(1-\epsilon_n)\|z - w\|^2 + \|(\alpha \mathrm{Id}_{\mathcal{G}} + B)Lz^*\|^2), \end{aligned} \tag{6.5}$$

where $z_+ = M_n^\gamma(z,w)$ and $\xi_n^\gamma(z,w) = z + \gamma(D_{K_n} w - D_{K_n} z)$.

(iii) *If there exists $\overline{\eta} \in [0,1[$ such that, for every $n \in \mathbb{N}$, $\omega_n = \omega_0 \overline{\eta}^n$, then the sequence $(\Upsilon_n)_{n\in\mathbb{N}}$ in (ii) can be chosen such that, for every $n \in \mathbb{N}$, $\Upsilon_n = \Upsilon_0 \overline{\eta}^n$.*

Proof.

(i) Let $(x,z) \in \mathcal{H}^2$ and let $n \in \mathbb{N}$. Then

$$\begin{aligned} \|D_{K_n} z - D_{K_n} x\| &= \|K_n(\alpha \mathrm{Id}_{\mathcal{G}} + B)Lz - K_n(\alpha \mathrm{Id} + B)Lx\| && (6.6) \\ &\leq \|(K_n - K)((\alpha \mathrm{Id} + B)Lz - (\alpha \mathrm{Id} + B)Lx)\| + \|D_K z - D_K x\| \\ &\leq (\omega_n(\alpha + \zeta)\|L\| + \kappa_K)\|z - x\|. && (6.7) \end{aligned}$$

This shows that

$$\epsilon_n = \gamma(\omega_n(\alpha + \zeta)\|L\| + \kappa_K) \tag{6.8}$$

is a Lipschitz constant of γD_{K_n}. We conclude by noticing that

$$\gamma(\omega_n(\alpha+\beta^{-1})\|L\|+\kappa_K)\to\gamma\kappa_K<1/2.$$

(ii) Let $n\in\mathbb{N}$. It follows from Cauchy-Schwarz inequality that

$$\begin{aligned}&2\gamma|\langle D_{K_{n+1}}z-D_{K_n}z+D_{K_{n+1}}z_+-D_Kz_+\mid z_+-z^*\rangle|\\&\leq 2\gamma(\|D_{K_{n+1}}z-D_{K_n}z+D_{K_{n+1}}z_+-D_Kz_+\|)\|z_+-z^*\|\\&\leq 2\gamma(\|D_{K_{n+1}}z-D_{K_n}z\|+\|D_{K_{n+1}}z_+-D_Kz_+\|)\|z_+-z^*\|.\end{aligned}\tag{6.9}$$

Now, we proceed by bounding each of the terms $\|z_+-z^*\|$, $\|D_nz-D_{n-1}z\|$, and $\|D_nz_+-Dz_+\|$. First, since $z^*\in\operatorname{zer}(A+C+D_K)$, we have $z^*=J_{\gamma A}(z^*-\gamma Cz^*-\gamma D_Kz^*)$ and, by the $(1+\gamma\rho)$-Lipschitzianity of $J_{\gamma A}$,

$$\begin{aligned}&\|z_+-z^*\|\\&=\|J_{\gamma A}(\xi_n^\gamma(z,w)-2\gamma D_{K_{n+1}}z+\gamma D_{K_n}z-\gamma Cz)-J_{\gamma A}(z^*-\gamma Cz^*-\gamma D_Kz^*)\|\\&\leq\frac{1}{1+\gamma\rho}\|\xi_n^\gamma(z,w)-2\gamma D_{K_{n+1}}z+\gamma D_{K_n}z-\gamma Cz-z^*+\gamma Cz^*+\gamma D_Kz^*\|\\&\leq\frac{1}{1+\gamma\rho}(\|\xi_n^\gamma(z,w)-z^*\|+\gamma\|D_{K_{n+1}}z-D_Kz^*\|+\gamma\|D_{K_{n+1}}z-D_{K_n}z\|+\gamma\|Cz-Cz^*\|)\\&\leq\frac{1}{1+\gamma\rho}(\|\xi_n^\gamma(z,w)-z^*\|+2\gamma\|D_{K_{n+1}}z-D_Kz^*\|+\gamma\|D_{K_n}z-D_Kz^*\|+\gamma\|Cz-Cz^*\|).\end{aligned}\tag{6.10}$$

Additionally, since $\gamma D_{K_{n+1}}$ (resp γD_{K_n}) is Lipschitzian with modulus ϵ_{n+1} (resp. ϵ_n),

$$\begin{aligned}&\gamma\|D_{K_{n+1}}z-D_Kz^*\|\\&\leq\gamma(\|D_{K_{n+1}}z-D_{K_{n+1}}z^*\|+\|D_{K_{n+1}}z^*-D_Kz^*\|)\\&\leq\epsilon_{n+1}\|z-z^*\|+\gamma\omega_{n+1}\|(\alpha\mathrm{Id}_{\mathcal{G}}+B)Lz^*\|\\&\leq\epsilon_{n+1}(\|\xi_n^\gamma(z,w)-z^*\|+\gamma\|D_{K_n}w-D_{K_n}z\|)+\gamma\omega_{n+1}\|(\alpha\mathrm{Id}_{\mathcal{G}}+B)Lz^*\|\\&\leq\epsilon_{n+1}\|\xi_n^\gamma(z,w)-z^*\|+\gamma\epsilon_{n+1}\epsilon_n\|z-w\|+\gamma\omega_{n+1}\|(\alpha\mathrm{Id}_{\mathcal{G}}+B)Lz^*\|.\end{aligned}\tag{6.11}$$

Similarly,

$$\gamma\|D_{K_n}z-D_Kz^*\|\leq\epsilon_n\|\xi_n^\gamma(z,w)-z^*\|+\gamma\epsilon_n^2\|z-w\|+\gamma\omega_n\|(\alpha\mathrm{Id}_{\mathcal{G}}+B)Lz^*\|,\tag{6.12}$$

and the β-cocoercivity of C yields

$$\begin{aligned}\|Cz-Cz^*\|&\leq\frac{1}{\beta}\|z-z^*\|\\&\leq\frac{1}{\beta}(\|\xi_n^\gamma(z,w)-z^*\|+\gamma\|D_{K_n}z-D_{K_n}w\|)\\&\leq\frac{1}{\beta}(\|\xi_n^\gamma(z,w)-z^*\|+\epsilon_n\|z-w\|).\end{aligned}\tag{6.13}$$

Altogether (6.10)-(6.13) lead to

$$\begin{aligned}\|z_+-z^*\|\leq\frac{1}{1+\gamma\rho}\Bigg(&\left(1+\frac{\gamma}{\beta}+2\epsilon_{n+1}+\epsilon_n\right)\|\xi_n^\gamma(z,w)-z^*\|+\gamma\epsilon_n\left(\frac{1}{\beta}+2\epsilon_{n+1}+\epsilon_n\right)\|z-w\|\\&+\gamma(2\omega_{n+1}+\omega_n)\|(\alpha\mathrm{Id}_{\mathcal{G}}+B)Lz^*\|\Bigg).\end{aligned}\tag{6.14}$$

Now, using Proposition 4.4(i) yields

$$\begin{aligned}
&\|D_{K_{n+1}}z - D_{K_n}z\| \\
&\leq \|D_{K_n}z - D_K z\| + \|D_{K_{n+1}}z - D_K z\| \\
&\leq (\omega_{n+1}+\omega_n)(\theta_1\|z-z^*\| + \|(\alpha\mathrm{Id}_{\mathcal{G}}+B)Lz^*\|) \qquad (6.15)\\
&\leq (\omega_{n+1}+\omega_n)(\theta_1\|\xi_n^\gamma(z,w)-z^*\| + \theta_1\gamma\|D_{K_n}w - D_{K_n}z\| + \|(\alpha\mathrm{Id}_{\mathcal{G}}+B)Lz^*\|) \\
&\leq (\omega_{n+1}+\omega_n)(\theta_1\|\xi_n^\gamma(z,w)-z^*\| + \theta_1\epsilon_n\|z-w\| + \|(\alpha\mathrm{Id}_{\mathcal{G}}+B)Lz^*\|). \qquad (6.16)
\end{aligned}$$

By invoking again Proposition 4.4(i), we have

$$\|D_{K_{n+1}}z_+ - D_K z_+\| \leq \omega_{n+1}(\theta_1\|z_+-z^*\| + \|(\alpha\mathrm{Id}_{\mathcal{G}}+B)Lz^*\|). \qquad (6.17)$$

Hence, since $(\epsilon_n)_{n\in\mathbb{N}}$ is bounded and $(\omega_n)_{n\in\mathbb{N}}$ converges to 0, we deduce from (6.14) and (6.16) that there exists $\mu_1 \in \,]0,+\infty[$ such that

$$\begin{aligned}
&\|D_{K_{n+1}}z - D_{K_n}z\|\|z_+-z^*\| \\
&\qquad \leq \mu_1(\omega_{n+1}+\omega_n)(\|\xi_n^\gamma(z,w)-z^*\|^2 + \epsilon_n^2\|z-w\|^2 + \|(\alpha\mathrm{Id}_{\mathcal{G}}+B)Lz^*\|^2). \qquad (6.18)
\end{aligned}$$

In addition, we deduce from (6.14) and (6.17) that there exists $\mu_2 \in \,]0,+\infty[$ such that

$$\begin{aligned}
&\|D_{K_{n+1}}z_+ - D_K z_+\|\|z_+-z^*\| \\
&\leq \omega_{n+1}(\theta_1\|z_+-z^*\| + \|(\alpha\mathrm{Id}_{\mathcal{G}}+B)Lz^*\|)\|z_+-z^*\| \\
&\leq \frac{\omega_{n+1}}{2}\left((2\theta_1+1)\|z_+-z^*\|^2 + \|(\alpha\mathrm{Id}_{\mathcal{G}}+B)Lz^*\|^2\right) \\
&\leq \mu_2\omega_{n+1}(\|\xi_n^\gamma(z,w)-z^*\|^2 + \epsilon_n^2\|z-w\|^2 + \|(\alpha\mathrm{Id}_{\mathcal{G}}+B)Lz^*\|^2). \qquad (6.19)
\end{aligned}$$

The following inequality is thus obtained by combining (6.9), (6.18), and (6.19):

$$\begin{aligned}
&2\gamma|\langle D_{K_{n+1}}z - D_{K_n}z + D_{K_{n+1}}z_+ - D_K z_+ \mid z_+ - z^*\rangle| \\
&\quad \leq 2\gamma(\mu_1(\omega_{n+1}+\omega_n) + \mu_2\omega_{n+1})(\|\xi_n^\gamma(z,w)-z^*\|^2 + \epsilon_n^2\|z-w\|^2 + \|(\alpha\mathrm{Id}_{\mathcal{G}}+B)Lz^*\|^2).
\end{aligned}$$

Finally, according to (i), there exists $N_0 \in \mathbb{N}$ such that, for every $n \geq N_0$, $\epsilon_n/(1-\epsilon_n) \in [\delta/(1-\delta),(1-\delta)/\delta]$. Since $\delta < 1/2$, we deduce that, for every $n \geq N_0$,

$$\begin{aligned}
&2\gamma\langle D_{K_{n+1}}z - D_{K_n}z + D_{K_{n+1}}z_+ - D_K z_+ \mid z_+ - z^*\rangle| \\
&\leq \frac{2\gamma(1-\delta)}{\delta}(\mu_1(\omega_{n+1}+\omega_n) + \mu_2\omega_{n+1})(\|\xi_n^\gamma(z,w)-z^*\|^2 \\
&\qquad\qquad + \epsilon_n(1-\epsilon_n)\|z-w\|^2 + \|(\alpha\mathrm{Id}_{\mathcal{G}}+B)Lz^*\|^2). \qquad (6.20)
\end{aligned}$$

The result follows by setting $\Upsilon_n = 2\gamma(1-\delta)(\mu_1(\omega_{n+1}+\omega_n) + \mu_2\omega_{n+1})/\delta$.

(iii) Note that, with the previous choice of $(\Upsilon_n)_{n\in\mathbb{N}}$, for every $n \in \mathbb{N}$,

$$\Upsilon_n = \frac{2\gamma}{\delta}(1-\delta)(\mu_1(\overline{\eta}+1) + \mu_2\overline{\eta})\omega_0\overline{\eta}^n = \Upsilon_0\overline{\eta}^n. \qquad (6.21)$$

□

Theorem 6.4. *In the context of Problem 1.4 and Assumption 3.2, let $\gamma \in]0,\chi'[$, and consider the sequence $(z_n)_{n\in\mathbb{N}}$ generated by Algorithm 6.1. Then the following hold.*

(i) *$(z_n)_{n\in\mathbb{N}}$ converges weakly to some solution to Problem 1.4.*

(ii) *If $\hat{\rho} > 0$ and there exists $\overline{\eta} \in [0,1[$ such that, for every $n \in \mathbb{N}$, $\omega_n = \omega_0\,\overline{\eta}^n$, then $(z_n)_{n\in\mathbb{N}}$ converges linearly to the unique solution to Problem 1.4.*

Proof. Let $z^* \in \operatorname{zer}(A + C + D_K)$, and consider the operators $(M_n^\gamma)_{n\in\mathbb{N}}$ defined by (6.2). Note that (6.1) can be reexpressed as

$$(\forall n \in \mathbb{N}) \quad z_{n+1} = M_n^\gamma(z_n, z_{n-1}). \tag{6.22}$$

For every $n \in \mathbb{N}$, set $u_n = \gamma(y_{n-1} - D_{K_n} z_n)$ and $\xi_n = z_n + u_n$.

(i) It follows from (6.1) that

$$\begin{aligned} & z_n - z_{n+1} - \gamma(2y_n - y_{n-1} + Cz_n) \in \gamma A z_{n+1} \\ \Leftrightarrow\ & \xi_n - \xi_{n+1} - \gamma(Cz_n + D_{K_{n+1}} z_n - D_{K_n} z_n + D_{K_{n+1}} z_{n+1} - D_K z_{n+1}) \in \gamma(A + D_K) z_{n+1}. \end{aligned} \tag{6.23}$$

Note that $-\gamma C z^* \in \gamma(A + D_K) z^*$. From the monotonicity of $A + D_K$, we derive the following inequality:

$$\begin{aligned} 0 &\le \big\langle \xi_n - \xi_{n+1} - \gamma(Cz_n + D_{K_{n+1}} z_n - D_{K_n} z_n + D_{K_{n+1}} z_{n+1} - D_K z_{n+1} - Cz^*) \mid z_{n+1} - z^* \big\rangle \\ &\le \langle \xi_n - \xi_{n+1} \mid z_{n+1} - z^* \rangle + \frac{\gamma}{4\beta} \|z_{n+1} - z_n\|^2 \\ &\qquad - \gamma \big\langle D_{K_{n+1}} z_n - D_{K_n} z_n + D_{K_{n+1}} z_{n+1} - D_K z_{n+1} \mid z_{n+1} - z^* \big\rangle, \end{aligned} \tag{6.24}$$

where we have used the thee point inequality for cocoercive operators [32, Equation (3)]

$$\langle Cz_n - Cz^* \mid z_{n+1} - z^* \rangle \ge -\frac{1}{4\beta} \|z_{n+1} - z_n\|^2. \tag{6.25}$$

Additionally, we have

$$\begin{aligned} 2\langle \xi_n - \xi_{n+1} \mid z_{n+1} - z^* \rangle &= \|\xi_n - z^*\|^2 - \|\xi_{n+1} - z^*\|^2 - \|\xi_n - z_{n+1}\|^2 + \|\xi_{n+1} - z_{n+1}\|^2 \\ &= \|\xi_n - z^*\|^2 - \|\xi_{n+1} - z^*\|^2 - \|u_n - (z_{n+1} - z_n)\|^2 + \|u_{n+1}\|^2, \end{aligned} \tag{6.26}$$

and we deduce from Proposition 6.3(i) that

$$\|u_n\| = \gamma \|D_{K_n} z_n - \gamma D_{K_n} z_{n-1}\| \le \epsilon_n \|z_n - z_{n-1}\|, \tag{6.27}$$

and

$$\begin{aligned} \|u_n - (z_{n+1} - z_n)\|^2 &= \|u_n\|^2 - 2\langle u_n \mid z_{n+1} - z_n \rangle + \|z_{n+1} - z_n\|^2 \\ &\ge \|u_n\|^2 - \frac{1}{\epsilon_n} \|u_n\|^2 - \epsilon_n \|z_{n+1} - z_n\|^2 + \|z_{n+1} - z_n\|^2 \\ &\ge -\left(\frac{1}{\epsilon_n} - 1\right) \|u_n\|^2 + (1 - \epsilon_n) \|z_{n+1} - z_n\|^2 \\ &\ge (1 - \epsilon_n)(\|z_{n+1} - z_n\|^2 - \epsilon_n \|z_n - z_{n-1}\|^2). \end{aligned} \tag{6.28}$$

Define, for every $n \in \mathbb{N}$,

$$\mu_n = 1 - \frac{\gamma}{2\beta} - \epsilon_n - \epsilon_{n+1} \tag{6.29}$$

$$a_n = \|\xi_n - z^*\|^2 + \epsilon_n (1 - \epsilon_n) \|z_n - z_{n-1}\|^2. \tag{6.30}$$

Altogether (6.24) and (6.26)-(6.28) yield

$$\begin{aligned}0 &\leq \|\xi_n - z^*\|^2 - \|\xi_{n+1} - z^*\|^2 - \|u_n - (z_{n+1} - z_n)\|^2 + \|u_{n+1}\|^2 + \frac{\gamma}{2\beta}\|z_{n+1} - z_n\|^2 \\ &\qquad - 2\gamma\big\langle D_{K_{n+1}} z_n - D_{K_n} z_n + D_{K_{n+1}} z_{n+1} - D_K z_{n+1}) \mid z_{n+1} - z^*\big\rangle \\ &\leq a_n - \|\xi_{n+1} - z^*\|^2 - \big(\mu_n + \epsilon_{n+1}(1-\epsilon_{n+1})\big)\|z_{n+1} - z_n\|^2 \\ &\qquad - 2\gamma\big\langle D_{K_{n+1}} z_n - D_{K_n} z_n + D_{K_{n+1}} z_{n+1} - D_K z_{n+1}) \mid z_{n+1} - z^*\big\rangle.\end{aligned} \tag{6.31}$$

Therefore, by Proposition 6.3(ii) (with $z = z_n$ and $w = z_{n-1}$), there exists $N_0 \in \mathbb{N}$ such that, for every $n \geq N_0$

$$\begin{aligned}a_{n+1} &\leq a_n - \mu_n\|z_{n+1} - z_n\|^2 \\ &\qquad - 2\gamma\big\langle D_{K_{n+1}} z_n - D_{K_n} z_n + D_{K_{n+1}} z_{n+1} - D_K z_{n+1}) \mid z_{n+1} - z^*\big\rangle \\ &\leq (1 + \Upsilon_n)a_n + \Upsilon_n\|(\alpha \mathrm{Id}_{\mathcal{G}} + B)Lz^*\|^2 - \mu_n\|z_{n+1} - z_n\|^2.\end{aligned} \tag{6.32}$$

According to (6.8),

$$\mu_n = 1 - \frac{\gamma}{2\beta} - 2\gamma\kappa_K - \gamma(\alpha + \zeta)\|L\|(\omega_n + \omega_{n+1}). \tag{6.33}$$

Since $\gamma < 2\beta/(4\beta\kappa + 1)$ and $\omega_n + \omega_{n+1} \to 0$, there exists $N_1 \geq N_0$ such that $\inf_{n \geq N_1} \mu_n > 0$. Then, by [3, Lemma 5.31], $(a_n)_{n\in\mathbb{N}}$ is convergent and $\sum_{n\in\mathbb{N}} \|z_{n+1} - z_n\|^2 < +\infty$. Since $z_{n+1} - z_n \to 0$, we deduce from (6.27) that $u_n \to 0$. From the convergence $(a_n)_{n\in\mathbb{N}}$ we deduce the convergence of $(\|\xi_n - z^*\|)_{n\in\mathbb{N}}$, which, from the definition of $(\xi_n)_{n\in\mathbb{N}}$, allows us to conclude that $(\|z_n - z^*\|)_{n\in\mathbb{N}}$ is convergent.

Additionally, it follows from (6.23) that

$$\begin{aligned}w_n = z_n - z_{n+1} - u_{n+1} + u_n - \gamma(Cz_n - Cz_{n+1} + D_{K_{n+1}} z_n - D_{K_n} z_n + D_{K_{n+1}} z_{n+1} - D_K z_{n+1}) \\ \in \gamma(A + D_K + C)z_{n+1}.\end{aligned} \tag{6.34}$$

Let $(z_{k_n})_{n\in\mathbb{N}}$ be a weakly convergent subsequence of $(z_n)_{n\geq N_1}$. Then

$$\begin{aligned}\|w_{k_n}\| &\leq (1 + \gamma\beta^{-1})\|z_{k_n} - z_{k_n+1}\| + \|u_{k_n+1}\| + \|u_{k_n}\| \\ &\qquad + \gamma(\|D_{K_{k_n+1}} z_{k_n} - D_{K_{k_n}} z_{k_n}\| + \|D_{K_{k_n+1}} z_{k_n+1} - D_K z_{k_n+1}\|).\end{aligned} \tag{6.35}$$

The sequence $(z_{k_n})_{n\in\mathbb{N}}$ is bounded and, since $z_{k_n} - z_{k_n+1} \to 0$, $(z_{k_n+1})_{n\in\mathbb{N}}$ is also bounded. According to Proposition 4.4(i),

$$\|D_{K_{k_n+1}} z_{k_n+1} - D_K z_{k_n+1}\| \leq \omega_{k_n+1}(\theta_1\|z_{k_n+1} - z^*\| + \|(\alpha \mathrm{Id}_{\mathcal{G}} + B)Lz^*\|), \tag{6.36}$$

and thus $D_{K_{k_n+1}} z_{k_n+1} - D_K z_{k_n+1} \to 0$. In addition, as shown in (6.15),

$$\|D_{K_{k_n+1}} z_{k_n} - D_{K_{k_n}} z_{k_n}\| \leq (\omega_{k_n+1} + \omega_{k_n})(\theta_1\|z_{k_n} - z^*\| + \|(\alpha \mathrm{Id}_{\mathcal{G}} + B)Lz^*\|). \tag{6.37}$$

We deduce from the previous two inequalities that $D_{K_{k_n+1}} z_{k_n+1} - D_K z_{k_n+1} \to 0$ and $D_{K_{k_n+1}} z_{k_n} - D_{K_{k_n}} z_{k_n} \to 0$, hence $w_{k_n} \to 0$. By the weak-strong closure of the maximally monotone operator $A + C + D_K$ (Proposition 3.4(ii) & [3, Proposition 20.38]), we conclude that every weak cluster point of $(z_n)_{n\in\mathbb{N}}$ belongs to $\mathrm{zer}(A + C + D_K)$. The result follows from Opial's lemma ([3, Lemma 2.47]).

(ii) Proceeding similarly to (i) and using the definition of $(a_n)_{n\in\mathbb{N}}$ in (6.30), the $\hat{\rho}$-strongly monotonicity of $A + D_K$ allows us to refine (6.32). More precisely, there exist $N_0 \in \mathbb{N}$, such that, for every $n \geq N_0$,

$$(1 + \Upsilon_n)a_n + \Upsilon_n\|(\alpha \mathrm{Id}_{\mathcal{G}} + B)Lz^*\|^2 \geq \mu_n\|z_{n+1} - z_n\|^2 + \gamma\hat{\rho}\|z_{n+1} - z^*\|^2 + a_{n+1}. \tag{6.38}$$

Now, by Proposition 6.3(i), for every $n \geq N_0$, we have $\epsilon_n \in [\delta, 1-\delta]$ with $\delta \in]0, 1/2[$, and

$$\begin{aligned}
&(1+\Upsilon_n)a_n + \Upsilon_n \|(\alpha \mathrm{Id}_{\mathcal{G}} + B)Lz^*\|^2 \\
&\quad \geq \frac{\mu_n}{2}\|z_{n+1} - z_n\|^2 + \frac{\mu_n}{2}\|z_{n+1} - z_n\|^2 + \gamma\hat{\rho}\|z_{n+1} - z^*\|^2 + a_{n+1} \\
&\quad \geq \frac{\mu_n \epsilon_{n+1}(1-\epsilon_{n+1})}{2(1-\delta)^2}\|z_{n+1} - z_n\|^2 + \frac{\mu_n}{2\epsilon_{n+1}}\|u_{n+1}\|^2 + \gamma\hat{\rho}\|z_{n+1} - z^*\|^2 + a_{n+1} \\
&\quad \geq \frac{\mu_n \epsilon_{n+1}(1-\epsilon_{n+1})}{2(1-\delta)^2}\|z_{n+1} - z_n\|^2 + \frac{1}{2}\min\left\{\frac{\mu_n}{2\delta}, \gamma\hat{\rho}\right\}\|\xi_{n+1} - z^*\|^2 + a_{n+1} \\
&\quad \geq \left(\frac{1}{2}\min\left\{\frac{\mu_n}{(1-\delta)^2}, \frac{\mu_n}{2\delta}, \gamma\hat{\rho}\right\} + 1\right) a_{n+1}. \qquad (6.39)
\end{aligned}$$

According to Proposition 6.3(iii), we can choose $(\Upsilon_n)_{n\in\mathbb{N}}$ such that $(\forall n \in \mathbb{N})\ \Upsilon_n = \Upsilon_0 \overline{\eta}^n$. Now, set

$$\vartheta^{-1} = \frac{1}{2}\min\left\{\frac{\mu_n}{(1-\delta^2)}, \frac{\mu_n}{2\delta}, \gamma\hat{\rho}\right\} + 1 > 1 \qquad (6.40)$$

$$\eta_0 = \vartheta\Upsilon_0\|(\alpha \mathrm{Id}_{\mathcal{G}} + B)Lz^*\|^2 \qquad (6.41)$$

$$(\forall n \in \mathbb{N}) \quad \varpi_n = \vartheta\Upsilon_n. \qquad (6.42)$$

By Lemma 3.8, $(a_n)_{n\in\mathbb{N}}$ converges linearly to 0. We deduce the linear convergence of $(z_n)_{n\in\mathbb{N}}$ to z^* by noticing that, for every $n \geq N_0$,

$$\begin{aligned}
a_n &= \|\xi_n - z^*\|^2 + \epsilon_n(1-\epsilon_n)\|z_n - z_{n-1}\|^2 \\
&\geq \|\xi_n - z^*\|^2 + \epsilon_n^2\|z_n - z_{n-1}\|^2 \\
&\geq \|\xi_n - z^*\|^2 + \|\gamma D_{K_n} z_n - \gamma D_{K_n} z_{n-1}\|^2 \\
&\geq \frac{1}{2}\|z_n - z^*\|^2. \qquad (6.43)
\end{aligned}$$

□

7. Numerical Experiments

This section is devoted to illustrate our theoretical results, through numerical experiments on an image reconstruction problem arising in Computed Tomography (CT), in material science.

7.1. **Problem formulation and settings.** In CT [23], one aims at solving the inverse problem of retrieving an estimate of a sought image $\bar{x} \in \mathbb{R}^N$, with $N \geq 1$ pixels, from acquisitions

$$c = \mathcal{D}(L\bar{x}), \qquad (7.1)$$

where $L \in \mathbb{R}^{N\times M}$ is a forward linear operator acting as a discretized Radon projector, $\mathcal{D} : \mathbb{R}^M \to \mathbb{R}^M$ models some noise perturbing the acquisitions, and $c \in \mathbb{R}^M$ is the noisy tomographic projection. We focus on the challenging situation when the back-projector matrix $L^\top \colon \mathbb{R}^M \to \mathbb{R}^N$ is approximated by $K \colon \mathbb{R}^M \to \mathbb{R}^N$. This is a current situation in practical CT reconstruction, where operator L (and thus, its transpose) cannot be stored, for memory reasons. It is instead implemented as a function, which computes on-the-fly projection and back-projections operations, making use of fast operations involving advanced interpolation strategies [42]. The adjoint mismatch is thus inherent to this application [22, 46] and, except in special simplistic cases, cannot be avoided.

An efficient approach to retrieve an estimate $\bar{x}$ from c, L, and K, consists of minimizing a penalized cost function, in the form of Problem 1.2. However, as explained earlier, due to the

adjoint mismatch, the formulation in Problem 1.2 is not well suited, and we propose instead to solve the following mismatched monotone inclusion:

$$\text{find } x \in \mathbb{R}^N \text{ such that } 0 \in \partial_{\mathrm{F}} f(x) + \nabla g(x) + \alpha K(Lx - c) + K\nabla h(Lx), \tag{7.2}$$

with $f\colon \mathbb{R}^N \to \mathbb{R}$ and $g\colon \mathbb{R}^N \to \mathbb{R}$ playing the role of regularization terms favoring a priori properties on the estimated image, and $h \circ L$ the data fidelity term accounting for the noise model. The latter inclusion problem reads as a particular instance of Problem 1.4, by setting $A = \partial_{\mathrm{F}} f$, $B = \nabla h$, and $C = \nabla g$ under suitable assumptions on the involved functions. In particular, we will choose f and g so that $\hat{\rho} > 0$ and the inclusion in (7.2) has a unique solution (Proposition 3.4.(iii)).

Data fidelity term: We consider a general mixed multiplicative/additive noise model, as discussed for instance in [9]. The vector c is related to $\bar{x}$ through

$$c = z + e, \tag{7.3}$$

with $z|\bar{x} \sim \mathcal{P}(L\bar{x})$ (i.e., Poisson distribution with mean $L\bar{x}$) and $e \sim \mathcal{N}(0, \sigma^2 \mathrm{Id})$ (i.e., i.i.d. Gaussian distribution with zero-mean and variance σ^2). Such a noise model allows to both account for multiplicative noise typical from emission tomography scenarios, and additive noise coming from the sensors. As shown in [43, 31], a suitable choice for the data fidelity term in such case is the Generalized Anscombe function, which is a smoothed approximation of the neg-log-likelihood associated to a Gauss-Poisson noise model. Under the assumption that $c = (c_m)_{1 \leq m \leq M} \in [-\frac{3}{8} - \sigma^2, +\infty[^M$ (which can be satisfied by basic cropping), function h reads

$$(\forall y = (y_m)_{1 \leq m \leq M} \in \mathbb{R}^M) \quad h(y) = \sum_{m=1}^{M} \varphi(y_m; c_m), \tag{7.4}$$

where, for every $a \in \mathbb{R}$, and every $b \in [-\frac{3}{8} - \sigma^2, +\infty[$,

$$\varphi(a; b) = \begin{cases} 2\left(\sqrt{b + \frac{3}{8} + \sigma^2} - \sqrt{a + \frac{3}{8} + \sigma^2}\right)^2 & \text{if} \quad a \geq 0, \\ \varphi(0; b) + \dot{\varphi}(0; b)a + \frac{1}{2}\nu(b)a^2 & \text{otherwise,} \end{cases} \tag{7.5}$$

with

$$\left(\forall b \geq -\frac{3}{8} - \sigma^2\right) \quad \nu(b) = \left(\frac{3}{8} + \sigma^2\right)^{-\frac{3}{2}} \sqrt{\frac{3}{8} + b + \sigma^2}. \tag{7.6}$$

Basic calculus shows that, for every $b \geq -\frac{3}{8} - \sigma^2$, the derivative of $\varphi(\cdot; b)$ at $a \geq 0$ reads

$$\dot{\varphi}(a; b) = 2 - \frac{2\sqrt{8b + 8\sigma^2 + 3}}{\sqrt{8a + 8\sigma^2 + 3}}. \tag{7.7}$$

Under this definition, we can readily show that, for every $b \geq -\frac{3}{8} - \sigma^2$, $\dot{\varphi}(\cdot\,; b)$ is Lipschitzian on $\mathbb{R}$, with constant $\nu(b)$. Assuming that the observed data satisfies $c \in [-\frac{3}{8} - \sigma^2, +\infty[^M$, we deduce that h is ζ-Lipschitz differentiable on $\mathbb{R}^N$ with

$$\zeta = \max_{m \in \{1, \ldots, M\}} \nu(c_m). \tag{7.8}$$

Regularization terms: Function f imposes the range of the restored image and controls the image energy, and is defined as

$$(\forall x \in \mathbb{R}^N) \quad f(x) = \iota_{[0, x_{\max}]^N}(x) + \frac{\rho}{2}\|x\|^2 \tag{7.9}$$

with $\rho \in]0, +\infty[$. Function f is ρ-strongly convex on $\mathbb{R}^N$. Its proximity operator has the following closed form expression:

$$(\forall \gamma \in]0, +\infty[) \quad \operatorname{prox}_{\gamma f}(x) = \min\{\max\{(\gamma\rho + 1)^{-1}x, 0\}, x_{\max}\}. \tag{7.10}$$

Function g promotes sparsity of the image in a transformed domain defined by a linear operator $W \in \mathbb{R}^{N\times N}$:

$$(\forall x \in \mathbb{R}^N) \quad g(x) = (\Phi_\delta \circ W)(x). \tag{7.11}$$

Hereabove, Φ_δ is the Huber function defined, for $\delta > 0$, as

$$(\forall x = (x_i)_{1\leq i\leq N} \in \mathbb{R}^N) \quad \Phi_\delta(x) = \sum_{i=1}^{N} \phi_\delta(x_i), \tag{7.12}$$

with

$$(\forall \eta \in \mathbb{R}) \quad \phi_\delta(\eta) = \begin{cases} |\eta| - \frac{\delta}{2}, & \text{if } |\eta| > \delta, \\ \frac{\eta^2}{2\delta}, & \text{otherwise.} \end{cases} \tag{7.13}$$

Function Φ_δ can be viewed as a smoothed approximation of the ℓ_1 penalty, promoting the sparsity of its argument. Function g belongs to $\Gamma_0(\mathbb{R}^N)$. Moreover, the derivative of ϕ_δ reads

$$(\forall \eta \in \mathbb{R}) \quad \dot{\phi}_\delta(\eta) = \begin{cases} \frac{|\eta|}{\eta}, & \text{if } |\eta| > \delta, \\ \frac{\eta}{\delta}, & \text{otherwise,} \end{cases} \tag{7.14}$$

which shows that Φ_δ has $(1/\delta)$-Lipschitizian gradient. We set $W \in \mathbb{R}^{N\times N}$ as an orthonormal wavelet transform [34], that leads to efficient penalties in tomography [24, 28]. Then $\|W\| = 1$ and g also has $(1/\delta)$-Lipschitizian gradient. Additionally, by orthogonality of W, [3, Corollary 23.27] yields

$$(\forall \gamma \in \left]0, +\infty\right[) \quad \operatorname{prox}_{\gamma g} = W^\top \circ \operatorname{prox}_{\gamma\Phi_\delta} \circ W, \tag{7.15}$$

with $W^\top = W^{-1}$, and, by [2, Proposition 24.11],

$$(\forall \gamma \in \left]0, +\infty\right[)(\forall x = (x_i)_{1\leq i\leq N} \in \mathbb{R}^N) \quad \operatorname{prox}_{\gamma\Phi_\delta}(x) = \big(\operatorname{prox}_{\gamma\phi_\delta}(x_i)\big)_{1\leq i\leq N}, \tag{7.16}$$

with

$$(\forall \gamma \in \left]0, +\infty\right[)(\forall \eta \in \mathbb{R}) \quad \operatorname{prox}_{\gamma\phi_\delta}(\eta) = \begin{cases} \eta - \frac{\gamma|\eta|}{\eta}, & \text{if } |\eta| > \delta + \gamma, \\ \frac{\delta\eta}{\gamma+\delta}, & \text{if } |\eta| \leq \delta + \gamma. \end{cases} \tag{7.17}$$

Algorithms implementation: We are now ready to apply Algorithm 4.1 (MMFBHF), Algorithm 5.1 (MMFDRF), and Algorithm 6.1 (MMFHRB) to solve Problem 1.2 (MM stands for MisMatched). In the considered setting, the algorithms read as follows.

Algorithm 7.1 (MMFBHF)**.** *Let $\gamma > 0$, let $z_0 \in \mathbb{R}^N$, and consider the iteration*

$$(\forall n \in \mathbb{N}) \quad \left\lfloor \begin{aligned} u_n &= K(\alpha\mathrm{Id} + \nabla h)(Lz_n) - \alpha Kc \\ y_n &= z_n - \gamma(\nabla g(z_n) + u_n) \\ x_n &= \operatorname{prox}_{\gamma f}(y_n) \\ z_{n+1} &= x_n + \gamma(u_n - K(\alpha\mathrm{Id} + \nabla h)(Lx_n) + \alpha Kc). \end{aligned} \right. \tag{7.18}$$

Algorithm 7.2 (MMFDRF)**.** *Let $\gamma > 0$, let $z_0 \in \mathbb{R}^N$, and consider the iteration*

$$(\forall n \in \mathbb{N}) \quad \left\lfloor \begin{aligned} x_n &= \operatorname{prox}_{\gamma g} z_n \\ w_n &= K(\alpha\mathrm{Id} + \nabla h)(Lx_n) - \alpha Kc \\ y_n &= \operatorname{prox}_{\gamma f}(2x_n - z_n - \gamma w_n) \\ z_{n+1} &= z_n + y_n - x_n - \gamma(K(\alpha\mathrm{Id} + \nabla h)(Ly_n) - \alpha Kc - w_n). \end{aligned} \right. \tag{7.19}$$

Algorithm 7.3 (MMFHRB). *Let $\gamma > 0$, let $z_0, z_{-1} \in \mathbb{R}^N$, let $y_{-1} = K_0(\alpha\mathrm{Id} + \nabla h)(Lz_{-1})$, and consider the iteration*

$$(\forall n \in \mathbb{N}) \quad \begin{cases} y_n = K(\alpha\mathrm{Id} + \nabla h)(Lz_n) - \alpha Kc \\ x_n = z_n - \gamma(2y_n - y_{n-1} + \nabla g(z_n)) \\ z_{n+1} = \operatorname{prox}_{\gamma f}(x_n). \end{cases} \tag{7.20}$$

The projector L is given by the line length ray-driven projector [44] and implemented in MATLAB using the line fan-beam projector provided by the ASTRA toolbox [40, 41]. Moreover, a constant mismatch, i.e., for every $n \in \mathbb{N}$, $K_n = K$, is considered, where the mismatched backprojector K is the adjoint of the strip fan-beam projector from the ASTRA toolbox.

In order to set up the stepsize parameters guaranteeing the convergence of our algorithms, we need to evaluate $\lambda_{\min}$, defined in (3.1). To do so, we compute the eigenvalues of the operator $(KL + L^\top K^\top)/2$ by using the function *eigs* from MATLAB, yielding $\lambda_{\min} \approx -6.0082$. Note that, it would also be possible to estimate $\lambda_{\min}$ avoiding an explicit implementation of $L^\top$ and $K^\top$ by the strategy proposed in [17]. In order to guarantee that Assumption 3.2(ii) holds, we set $\rho = -\alpha\lambda_{\min} + \widetilde{\zeta}_{L^\top - K} + 10^{-3}$, where $\widetilde{\zeta}_{L^\top - K}$ is estimated as $\|L^\top - K\|\|L\|\zeta$. The spectral norms $\|L^\top - K\|$ and $\|L\|$ are computed using the power iterative method. We implement MMFBHF with constant step-size $\gamma = (3.99\beta)/(1 + \sqrt{1 + 16\beta^2\kappa_K^2})$, MMFDRF with $\gamma = (4.53\beta)/(1 + \sqrt{1 + 16\beta^2\kappa_K^2})$, and MMFHRB with $\gamma = 0.999\hat{\gamma}$, where $\hat{\gamma}$ is the largest solution to the equation $\kappa_K^2\hat{\gamma}^2(1 + \hat{\gamma}/(2\beta))) = 1$, computed numerically. These choices allow satisfying our technical assumptions, so that the convergence theorems hold.

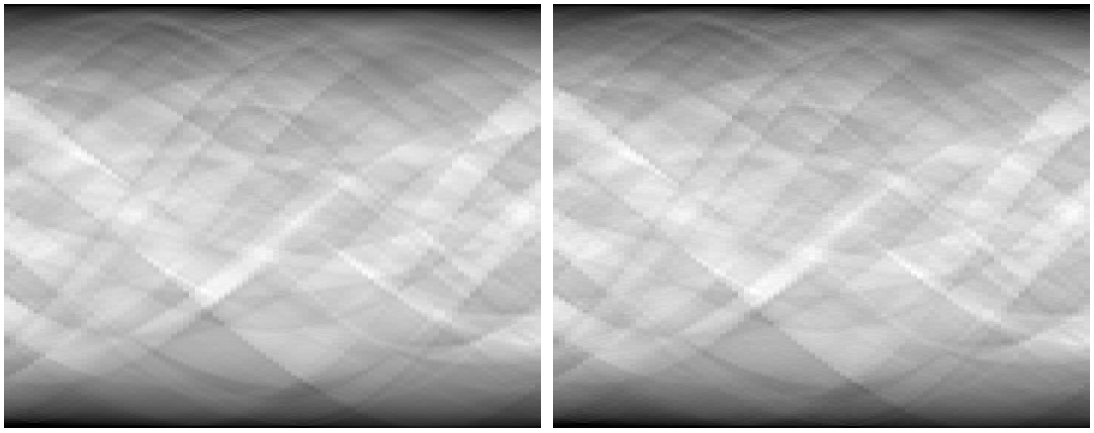
7.2. **Experimental results.** We now present our experimental results. In the observation model (7.1), the ground truth image $\overline{x}$ represents a part of a high resolution scan of a phase-separated barium borosilicate glass imaged at the ESRF synchrotron [11][1]. The image size is $N = 128 \times 128$ pixels. The projector L describes a fan-beam geometry over 180° using 90 regularly spaced angular steps. The source-to-object distance is 800 mm, and the source-to-image distance is 1200 mm. The bin grid is twice upsampled with respect to the pixel grid, the detector has 249 bins of size 1.6 mm, so that $M = 90 \times 249$. The pixel values of $\overline{x}$ outside a circle of diameter 128 pixels are set to 0, to guarantee that the object of interest lies within the field of view.

The image intensity range lies in $[0, x_{\max}]$, with $x_{\max} = 900$. The Gaussian noise level is set to $\sigma = 200$. The input signal-to-noise-ratio (SNR) in decibels (dB), between the clean projection $L\bar{x}$ and c (both displayed on Figure 1), defined as

$$\mathrm{SNR}_{\mathrm{input}} = 20\log_{10}\left(\frac{\|L\bar{x}\|}{\|L\bar{x} - c\|}\right), \tag{7.21}$$

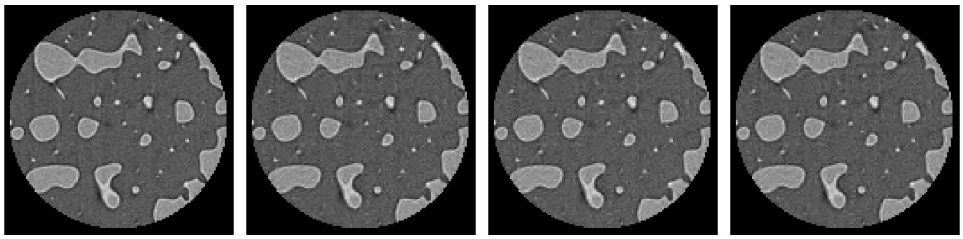
is here equal to 42.18 dB. Problem (1.3) is solved using an orthonormal Symmlet basis with 4 vanishing moments, and 2 resolution levels for W operator. The following penalty parameter values are chosen: $\lambda = 150$, $\delta = 5$, and $\alpha = 0.1$. The reconstructed images using MMFBHF, MMFDRF, and MMFHRB with 10^4 iterations are presented in Figure 2. We also present the results within a zoomed region-of-interest (ROI), with size 80×80 pixels and circular shape, in Figure 2 (bottom). We evaluate, for each algorithm, the quantitative error between the original image $\bar{x}$ and its recovered version $\hat{x}$, through the normalized mean squared error NMSE $= \|\bar{x} - \hat{x}\|^2/\|\bar{x}\|^2$, the mean absolute error MAE $= \|\bar{x} - \hat{x}\|_\infty$ and the SNR $= 10\log_{10}$ NMSE. Similar formula are used to determine SNR, MAE and NMSE scores inside the ROI. The obtained values are provided in the caption of Figure 2.

[1] https://www.esrf.fr/ - The dataset is a courtesy of David Bouttes.

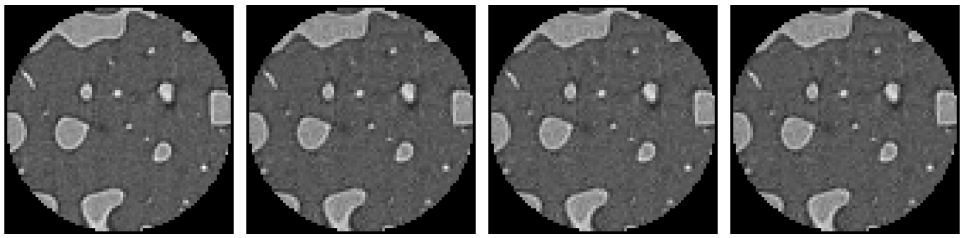


(A) Tomographic projection $L\overline{x}$ (B) Noisy projection c.

FIGURE 1. Clean projection and its noisy version, with $\mathrm{SNR}_{\mathrm{input}} = 42.18$ dB



(A) Original image $\overline{x}$ (B) MMFBHF result, SNR=23.29 dB, NMSE $= 4.71 \times 10^{-3}$, MAE $=$ 120.58. (C) MMFDRF result, SNR=23.20 dB, NMSE $= 4.72 \times 10^{-3}$, MAE $=$ 119.76. (D) MMFHRB result, SNR=23.28 dB, NMSE $= 4.72 \times 10^{-3}$, MAE $=$ 121.22.



(E) Zoomed original image $\overline{x}$ (F) Zoomed MMFBHF result, SNR=23.20 dB, NMSE $= 4.80 \times 10^{-3}$, MAE $=$ 120.58. (G) Zoomed MMFDRF result, SNR=23.11 dB, NMSE $= 4.81 \times 10^{-3}$, MAE $=$ 116.90. (H) Zoomed MMFHRB result, SNR=23.18 dB, NMSE $= 4.83 \times 10^{-3}$, MAE $=$ 121.22.

FIGURE 2. Original and reconstructed images (full view, and zoom) after 10^5 iterations of MMFBHF, MMFDRF, and MMFHRB algorithms, respectively.

In Figure 3, we display the evolution of the SNR, along iterations and times, for codes running in MATLAB R2023a, on a laptop with AMD Ryzen 5 3550Hz, Radeon Vega Mobile Gfx, and 32 Gb RAM. One can notice that MMFBHF and MMFDRF behave similarly, while MMFHRB is slightly behind, in terms of convergence speed. Still, all algorithms reach convergence in about 2000 iterations, and 200 seconds, confirming the validity of our theoretical results.

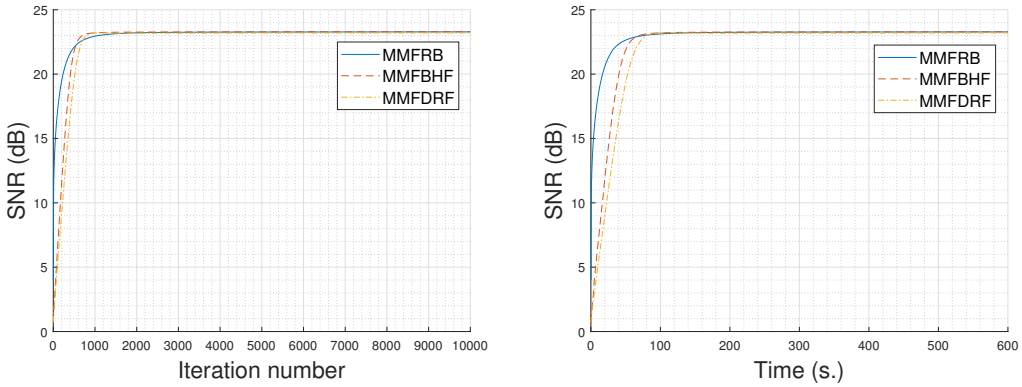


FIGURE 3. Evolution along iterations (left) and computational time (right) in seconds, of the SNR (in dB) between the true image and its reconstruction, for MMFBHF, MMFDRF, and MMFHRB algorithms.

8. Conclusion

In this paper, we introduced three iterative algorithms for numerically solving monotone inclusions involving the sum of a maximally ρ-monotone operator, a cocoercive operator, and a mismatched Lipschitzian operator. The proposed schemes can be viewed as extensions of the *Forward-Backward-Half-Forward*, the *Forward-Douglas-Rachford-Forward*, and the *Forward-Half-Reflected-Backward* splitting methods, that use an approximation to an adjoint operator at each iteration. We provided conditions under which the sequence generated by these variants weakly converges to a solution to the mismatched inclusion. We also showed that, under some strong monotonicity assumptions, a linear convergence rate is obtained for the three algorithms. The applicability of our study is illustrated by numerical experiments in the context of imaging of materials. When appplied to variational problems, the main advantage of our work with respect to [19, 16, 8, 10] is to allow to deal with mistmatches on more sophisticated functions than quadratic ones.

Acknowledgements

E.C. acknowledges support from the European Research Council Starting Grant MAJORIS ERC-2019-STG-850925.

References

[1] Briceño Arias, L.M., Combettes, P.L.: Monotone operator methods for Nash equilibria in non-potential games. In: Computational and analytical mathematics, *Springer Proc. Math. Stat.*, vol. 50, pp. 143–159. Springer, New York (2013). DOI 10.1007/978-1-4614-7621-4_9

[2] Bauschke, H.H., Combettes, P.L.: Convex analysis and monotone operator theory in Hilbert spaces. CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC. Springer, New York (2011). DOI 10.1007/978-1-4419-9467-7

[3] Bauschke, H.H., Combettes, P.L.: Convex analysis and monotone operator theory in Hilbert spaces, second edn. CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC. Springer, Cham (2017). DOI 10.1007/978-3-319-48311-5

[4] Bauschke, H.H., Moursi, W.M., Wang, X.: Generalized monotone operators and their averaged resolvents. Math. Program. **189**(1-2, Ser. B), 55–74 (2021). DOI 10.1007/s10107-020-01500-6

[5] Bertocchi, C., Chouzenoux, E., Corbineau, M.C., Pesquet, J.C., Prato, M.: Deep unfolding of a proximal interior point method for image restoration. Inverse Problems **36**(3), 034,005, 27 (2020). DOI 10.1088/1361-6420/ab460a

[6] Briceño-Arias, L.M., Davis, D.: Forward-backward-half forward algorithm for solving monotone inclusions. SIAM J. Optim. **28**(4), 2839–2871 (2018). DOI 10.1137/17M1120099

[7] Bubba, T.A., Galinier, M., Lassas, M., Prato, M., Ratti, L., Siltanen, S.: Deep neural networks for inverse problems with pseudodifferential operators: An application to limited-angle tomography. SIAM Journal on Imaging Sciences **14**(2), 470–505 (2021). DOI 10.1137/20M1343075

[8] Chouzenoux, E., Contreras, A., Pesquet, J.C., Savanier, M.: Convergence results for primal-dual algorithms in the presence of adjoint mismatch. SIAM Journal on Imaging Sciences **16**(1), 1–34 (2023)

[9] Chouzenoux, E., Jezierska, A., Pesquet, J.C., Talbot, H.: A convex approach for image restoration with exact Poisson-Gaussian likelihood. SIAM Journal on Imaging Sciences **8**(4), 2662–2682 (2015)

[10] Chouzenoux, E., Pesquet, J.C., Riddell, C., Savanier, M., Trousset, Y.: Convergence of proximal gradient algorithm in the presence of adjoint mismatch. Inverse Problems **37**(6), Paper No. 065,009, 29 (2021). DOI 10.1088/1361-6420/abd85c

[11] Chouzenoux, E., Zolyniak, F., Gouillart, E., Talbot, H.: A majorize-minimize memory gradient algorithm applied to X-ray tomography. In: Proceedings of the 20th IEEE International Conference on Image Processing (ICIP 2013), pp. 1011–1015. Melbourne, Australia (2013)

[12] Combettes, P.L.: Monotone operator theory in convex optimization. Math. Program. **170**(1), 177–206 (2018). DOI 10.1007/s10107-018-1303-3

[13] Combettes, P.L., Pesquet, J.C.: 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). DOI 10.1007/s11228-011-0191-y

[14] Combettes, P.L., Pesquet, J.C.: Fixed point strategies in data science. IEEE Transactions on Signal Processing **69**, 3878–3905 (2021). DOI 10.1109/TSP.2021.3069677

[15] Condat, L.: A primal-dual splitting method for convex optimization involving Lipschitzian, proximable and linear composite terms. J. Optim. Theory Appl. **158**(2), 460–479 (2013). DOI 10.1007/s10957-012-0245-9

[16] Dong, Y., Hansen, P.C., Hochstenbach, M.E., Brogaard Riis, N.A.: Fixing nonconvergence of algebraic iterative reconstruction with an unmatched backprojector. SIAM Journal on Scientific Computing **41**(3), A1822–A1839 (2019). DOI 10.1137/18M1206448

[17] Dong, Y., Hansen, P.C., Hochstenbach, M.E., Brogaard Riis, N.A.: Fixing nonconvergence of algebraic iterative reconstruction with an unmatched backprojector. SIAM Journal on Scientific Computing **41**(3), A1822–A1839 (2019). DOI 10.1137/18M1206448

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

[19] Elfving, T., Hansen, P.C.: Unmatched projector/backprojector pairs: Perturbation and convergence analysis. SIAM Journal on Scientific Computing **40**(1), A573–A591 (2018). DOI 10.1137/17M1133828

[20] Giselsson, P., Moursi, W.M.: On compositions of special cases of Lipschitz continuous operators. Fixed Point Theory Algorithms Sci. Eng. pp. Paper No. 25, 38 (2021). DOI 10.1186/s13663-021-00709-0

[21] Goldstein, A.A.: Convex programming in Hilbert space. Bulletin of the American Mathematical Society **70**(5), 709 – 710 (1964). DOI bams/1183526263. URL https://doi.org/

[22] Hansen, P.C., Hayami, K., Morikuni, K.: GMRES methods for tomographic reconstruction with an unmatched back projector. Tech. rep. (2022). URL https://arxiv.org/pdf/2110.01481.pdf

[23] Kak, A.C., Slaney, M.: Principles of Computerized Tomographic Imaging. Society of Industrial and Applied Mathematics (2001)

[24] Klann, E., Quinto, E.T., Ramlau, R.: Wavelet methods for a weighted sparsity penalty for region of interest tomography. Inverse Problems **31**(2), 025,001 (2015)

[25] Lions, P.L., Mercier, B.: Splitting algorithms for the sum of two nonlinear operators. SIAM J. Numer. Anal. **16**(6), 964–979 (1979). DOI 10.1137/0716071

[26] Lorenz, D.A., Rose, S., Schöpfer, F.: The randomized Kaczmarz method with mismatched adjoint. BIT **58**(4), 1079–1098 (2018). DOI 10.1007/s10543-018-0717-x

[27] Lorenz, D.A., Schneppe, F.: Chambolle-Pock's primal-dual method with mismatched adjoint. Appl. Math. Optim. **87**(2), Paper No. 22, 26 (2023). DOI 10.1007/s00245-022-09933-5

[28] Loris, I., Nolet, G., Daubechies, I., Dahlen, F.A.: Tomographic inversion using l1-norm regularization of wavelet coefficients. Geophysical Journal International **170**(1), 359–370 (2007)

[29] Loris, I., Verhoeven, C.: On a generalization of the iterative soft-thresholding algorithm for the case of non-separable penalty. Inverse Problems **27**(12), 125,007, 15 (2011). DOI 10.1088/0266-5611/27/12/125007

[30] Malitsky, Y., Tam, M.K.: A forward-backward splitting method for monotone inclusions without cocoercivity. SIAM J. Optim. **30**(2), 1451–1472 (2020). DOI 10.1137/18M1207260

[31] Marnissi, Y., Zheng, Y., Chouzenoux, E., Pesquet, J.C.: A variational bayesian approach for image restoration. application to image deblurring with Poisson-Gaussian noise. IEEE Transactions on Computational Imaging **3**(4), 722–737 (2017)

[32] Morin, M., Banert, S., Giselsson, P.: Nonlinear forward-backward splitting with momentum correction (2022)

[33] Passty, G.: Ergodic convergence to a zero of the sum of monotone operators in hilbert space. J. Math. Anal. Appl. **72**(2), 383–390 (1979)

[34] Pustelnik, N., Benazza-Benhayia, A., Zheng, Y., Pesquet, J.C.: Wavelet-based image deconvolution and reconstruction, pp. 1–34. Wiley Encyclopedia of Electrical and Electronics Engineering (2016)

[35] Ryu, E.K., Vũ, B.C.: Finding the forward-Douglas-Rachford-forward method. J. Optim. Theory Appl. **184**(3), 858–876 (2020). DOI 10.1007/s10957-019-01601-z

[36] Savanier, M., Chouzenoux, E., Pesquet, J.C., Riddell, C.: Unmatched preconditioning of the proximal gradient algorithm. IEEE Signal Processing Letters **29**, 1122–1126 (2022). DOI 10.1109/LSP.2022.3169088

[37] Savanier, M., Chouzenoux, E., Pesquet, J.C., Riddell, C.: Deep unfolding of the dbfb algorithm with application to roi ct imaging with limited angular density. IEEE Transactions on Computational Imaging **9**, 502–516 (2023). DOI 10.1109/TCI.2023.3279053

[38] Tseng, P.: A modified forward-backward splitting method for maximal monotone mappings. SIAM J. Control Optim. **38**(2), 431–446 (2000). DOI 10.1137/S0363012998338806

[39] Vũ, B.C.: A splitting algorithm for dual monotone inclusions involving cocoercive operators. Adv. Comput. Math. **38**(3), 667–681 (2013). DOI 10.1007/s10444-011-9254-8

[40] van Aarle, W., Palenstijn, W.J., Cant, J., Janssens, E., Bleichrodt, F., Dabravolski, A., Beenhouwer, J.D., Batenburg, K.J., Sijbers, J.: Fast and flexible x-ray tomography using the astra toolbox. Opt. Express **24**(22), 25,129–25,147 (2016). DOI 10.1364/OE.24.025129

[41] van Aarle, W., Palenstijn, W.J., De Beenhouwer, J., Altantzis, T., Bals, S., Batenburg, K.J., Sijbers, J.: The astra toolbox: A platform for advanced algorithm development in electron tomography. Ultramicroscopy **157**, 35–47 (2015). DOI https://doi.org/10.1016/j.ultramic.2015.05.002

[42] Xu, F., Mueller, K.: A comparative study of popular interpolation and integration methods for use in computed tomography. In: Proceedings of the 3rd IEEE International Symposium on Biomedical Imaging: Nano to Macro, 2006., pp. 1252–1255 (2006)

[43] Zanni, L., Benfenati, A., Bertero, M., Ruggiero, V.: Numerical methods for parameter estimation in poisson data inversion. Journal of Mathematical Imaging and Vision **52**(3), 397—-413 (2015)

[44] Zeng, G., Gullberg, G.: A ray-driven backprojector for backprojection filtering and filtered backprojection algorithms. In: 1993 IEEE Conference Record Nuclear Science Symposium and Medical Imaging Conference, pp. 1199–1201 (1993). DOI 10.1109/NSSMIC.1993.701833

[45] Zeng, G., Gullberg, G.: Unmatched projector/backprojector pairs in an iterative reconstruction algorithm. IEEE Transactions on Medical Imaging **19**(5), 548–555 (2000). DOI 10.1109/42.870265

[46] Zeng, G., Gullberg, G.: Unmatched projector/backprojector pairs in an iterative reconstruction algorithm. IEEE Transactions on Medical Imaging **19**(5) (2000)