



**HAL**  
open science

# Convergence Results for Primal-Dual Algorithms in the Presence of Adjoint Mismatch

Emilie Chouzenoux, Andrés Contreras, Jean-Christophe Pesquet, M. Savanier

► **To cite this version:**

Emilie Chouzenoux, Andrés Contreras, Jean-Christophe Pesquet, M. Savanier. Convergence Results for Primal-Dual Algorithms in the Presence of Adjoint Mismatch. *SIAM Journal on Imaging Sciences*, 2023, 16 (1), pp.1-34. 10.1137/22M1490223 . hal-04210972v2

**HAL Id: hal-04210972**

**<https://inria.hal.science/hal-04210972v2>**

Submitted on 19 Sep 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

# 1 Convergence Results For Primal-Dual Algorithms in the Presence of Adjoint Mismatch\*

2 Emilie Chouzenoux<sup>†</sup>, Andrés Contreras<sup>†‡</sup>, Jean-Christophe Pesquet<sup>†</sup>, and Marion Savanier<sup>†§</sup>

3  
4 **Abstract.** Most optimization problems arising in imaging science involve high-dimensional linear operators and their  
5 adjoints. In the implementations of these operators, changes may be introduced for various practical consid-  
6 erations (e.g., memory limitation, computational cost, convergence speed), leading to an *adjoint mismatch*.  
7 This occurs for the X-ray tomographic inverse problems found in Computed Tomography (CT), where a  
8 surrogate operator often replaces the adjoint of the measurement operator (called projector). The resulting  
9 adjoint mismatch can jeopardize the convergence properties of iterative schemes used for image recovery.  
10 In this paper, we study the theoretical behavior of a panel of primal-dual proximal algorithms, which rely  
11 on forward-backward-(forward) splitting schemes when an adjoint mismatch occurs. We analyze these al-  
12 gorithms by focusing on the resolution of possibly non-smooth convex penalized minimization problems in  
13 an infinite-dimensional setting. Using tools from fixed point theory, we show that they can solve mono-  
14 tone inclusions beyond minimization problems. Such findings indicate these algorithms can be seen as a  
15 generalization of classical primal-dual formulations. The applicability of our findings is also demonstrated  
16 through two numerical experiments in the context of CT image reconstruction.

17  
18 **Key words.** Adjoint mismatch, convex optimization, primal-dual algorithms, convergence analysis, fixed point theory,  
19 CT reconstruction, forward-backward splitting, Tseng’s splitting.

20 **AMS subject classifications.** 65K10, 62H35, 94A08, 90C25, 47H05

21 **1. Introduction.** Many imaging science applications [2] can be formulated as inverse problems  
22 where a high-dimensional variable has to be estimated from noisy indirect observations. Solving  
23 an inverse problem can often be recast as the problem of minimizing a cost function expressed  
24 as a sum of convex terms composed with linear operators [41]. Especially when these functions  
25 are non-smooth, first-order proximal splitting algorithms appear as appealing minimization strate-  
26 gies [19, 11]. In these algorithms, each cost function term is handled separately, either through its  
27 gradient or through its proximity operator. At the core of the convergence analysis of most proximal  
28 splitting methods is the interpretation of the system of first-order optimality conditions associated  
29 with an optimization problem [21]. For example, the forward-backward (also called proximal gradi-  
30 ent [23] in the variational case) algorithm [24] solves an inclusion based on a sum of two monotone  
31 operators. Owing to its solid convergence guarantees, this scheme is widely used in inverse prob-  
32 lems involving sparsity priors [27, 15]. However, when the cost function involves a non-smooth term  
33 composed with a linear operator, the proximal step in the forward-backward algorithm may not  
34 have a closed form, so it must be computed using an inner loop at the expense of practical efficiency.  
35 A way to avoid performing these sub-iterations is to rely on primal-dual proximal splitting algo-  
36 rithms [36, 9, 22, 44, 1, 31]. These algorithmic schemes are grounded on splitting strategies such  
37 as the forward-backward, the Douglas-Rachford [5, 33, 6], or Tseng’s forward-backward-forward  
38 [53] algorithms for finding a zero of a sum of maximally monotone operators. For instance, the

---

\*Submitted to the editors September 19, 2023.

**Funding:** This work was supported by the European Research Council Starting Grant MAJORIS ERC-2019-STG-850925, the ANRT CIFRE Convention 2018/1587, and the ANR Research and Teaching Chair in Artificial Intelligence BRIDGEABLE.

<sup>†</sup>Université Paris-Saclay, CentraleSupélec, Centre pour la Vision Numérique, Inria, Gif-sur-Yvette, France. ([emilie.chouzenoux@centralesupelec.fr](mailto:emilie.chouzenoux@centralesupelec.fr), [jean-christophe.pesquet@centralesupelec.fr](mailto:jean-christophe.pesquet@centralesupelec.fr)).

<sup>‡</sup>Corresponding author, Universidad Católica de Temuco, Departamento de Ciencias Matemáticas y Físicas ([c5.contreras@gmail.com](mailto:c5.contreras@gmail.com), [acontreras@uct.cl](mailto:acontreras@uct.cl)).

<sup>§</sup>GE Healthcare, Buc, France. ([marion.savanier@gmail.com](mailto:marion.savanier@gmail.com)).

39 Chambolle-Pock algorithm [14] is one of the most popular methods. More flexibility can be added  
40 to the considered formulation by considering the gradient of Lipschitz differentiable terms in the  
41 objective function. The Combettes-Pesquet algorithm [22] based on Tseng’s forward-backward-  
42 forward splitting was the first algorithm offering this possibility. Then, the Condat-Vũ algorithm  
43 [56, 25] was proposed, which constitutes an extension of the Chambolle-Pock one. The Loris-  
44 Verhoeven algorithm [40] is another useful algorithm relying on the forward-backward splitting. It  
45 offers more freedom in choosing the algorithm parameters but deals with a more restricted class  
46 of optimization problems. Primal-dual algorithms have numerous applications in the resolution  
47 of signal/image processing problems [52, 45, 17, 7]. These schemes increase the dimension of the  
48 original minimization variables by introducing auxiliary dual variables, yielding the resolution of a  
49 system of two decoupled monotone inclusions in the primal-dual product space.

50 In real-life imaging applications, an *adjoint mismatch* frequently occurs. This term corresponds to  
51 the situation when, in the iterative process, the adjoints of the linear operators in the cost function  
52 are replaced by surrogate operators. An adjoint mismatch typically occurs in Computed Tomogra-  
53 phy (CT) image reconstruction. CT involves the resolution of a linear inverse problem, where the  
54 forward operator is a discrete Radon transform that relates an image to a set of projections measured  
55 at several angles during the rotation of an X-ray imaging system [35]. During the reconstruction,  
56 the adjoint of the projector, called backprojector, is often approximated to produce an estimate of  
57 the image more rapidly (by improving the conditioning [59] or reducing computational complexity  
58 per iteration [32]) or to improve the quality of the reconstruction [58]. An adjoint mismatch has also  
59 been advocated in SPECT (Single Photon Emission Computed Tomography) imaging [47, 59] and  
60 in microscopy imaging [46]. With such mathematical approximations, the convergence guarantees  
61 of classical minimization algorithms no longer hold. Hence there is a need to design and adapt  
62 existing results and schemes to these situations.

63 The effect of an adjoint mismatch has first been studied for CT in the context of simultaneous or  
64 row-action algebraic algorithms [35] based on projections onto convex sets (POCS). In that case,  
65 the convergence can be analyzed with tools from linear algebra. Some works gave convergence  
66 conditions [30, 59, 28, 38] and focused on fixing the convergence of these schemes [28]. In contrast,  
67 others proposed to use more general optimization schemes that can directly deal with non-symmetric  
68 normal fixed-point equations [42, 51]. The resulting algorithms are semi-convergent; the number  
69 of iterations acts as a regularization parameter. The literature on the influence of an adjoint mis-  
70 match on more complex proximal splitting schemes has been expanding recently. The convergence  
71 of the proximal gradient algorithm (PGA) has been studied in [18] when an adjoint mismatch is  
72 present on the forward operator. In the context of microscopy imaging, where the forward operator  
73 satisfies a specific orthogonality condition, [46] gave conditions for the Douglas-Rachford/ADMM  
74 iterations to converge in the Multi-Agent Consensus Equilibrium (MACE) framework [12]. They  
75 also highlighted that using an adjoint mismatch in a quadratic data fidelity term was equivalent to  
76 using the proximity operator (or agent) of the quadratic term with a different prior model for the  
77 image that depends on the mismatched adjoint. A similar observation was made in [50], where the  
78 authors showed that considering an adjoint mismatch in PGA could be seen as a problem of un-  
79 matched preconditioning where the metric used in the gradient step differs from the one used in the  
80 proximity step. Up to our knowledge, the first proposal to analyze a primal-dual proximal splitting  
81 method under an adjoint mismatch is [39] which studies a mismatched form of the Chambolle-Pock  
82 algorithm with fixed step sizes.

83 This analysis is conducted under strong convexity assumptions. The authors give conditions  
84 on the strong convexity modulus of the involved functions and derive formulas for the step sizes to  
85 recover a similar convergence rate to that of the matched scheme. However, they do not investigate  
86 the existence and uniqueness of the fixed point of the mismatched algorithm.

87 The contribution of this paper is to analyze the behavior of some of the most popular primal-dual  
 88 proximal splitting algorithms in the presence of adjoint mismatch by relying on fixed-point theory  
 89 [21]. Our study leverages some of our previous results in [18]. Our analysis applies to a broad  
 90 class of penalized least-squares problems. More specifically, we provide sufficient conditions for the  
 91 existence and uniqueness of fixed point of the modified schemes, and we also provide conditions  
 92 to ensure the convergence of the schemes to such a fixed point. Our methodology allows us to  
 93 develop extensions of existing algorithms that are robust to adjoint mismatches. Their resulting  
 94 fixed points satisfy general notions of equilibrium (similar to the one arising in neural network  
 95 architectures/paradigms, plug-and-play methods [21, 43, 34]) instead of being the solution to a  
 96 minimization problem. In this aspect, our work is in line with the recent efforts to design [50, 49]  
 97 and learn [3, 10, 43, 12] more expressive variants of well-known optimization schemes while keeping  
 98 convergence guarantees.

99

100 The paper is organized as follows. In section 2, we present our notation and useful definitions un-  
 101 derlying the class of minimization and monotone inclusion problems under the scope of our study.  
 102 We also present the penalized least-squares cost function we aim to minimize, which involves two  
 103 linear operators. The first linear operator appears in a quadratic term, while the second arises in a  
 104 non-necessarily smooth convex term. Sections 3, 4, and 5 gather our methodological contributions,  
 105 namely the study of the properties of three main classes of primal-dual splitting algorithms in the  
 106 presence of adjoint mismatch. Specifically, section 3 focuses on the Condat-Vũ algorithm [56, 25],  
 107 and its mismatched version. We consider the case when an adjoint mismatch appears on the linear  
 108 operator in the quadratic term and the second linear operator is perturbed. Note that this differs  
 109 from the mismatch considered in [39], which appears on the second linear operator. This analysis  
 110 is then extended to the projected form of the Condat-Vũ algorithm proposed by Briceño-Arias and  
 111 López [8]. In section 4, we perform an analysis of the Loris-Verhoeven algorithm [40] in the case of  
 112 an adjoint mismatch on the linear operator involved in the quadratic term. In section 5, we analyze  
 113 the Combettes-Pesquet algorithm [22] when the adjoints of both linear operators are modified. We  
 114 derive convergence results for all the above algorithms and characterize the resulting fixed points.  
 115 Finally, in section 6, we illustrate our theoretical findings on examples involving two different inverse  
 116 problems in CT reconstruction, with two types of regularization and noise modeling.

## 117 2. Preliminaries.

118 **2.1. Notation and background.** Let us introduce some basic notation and definitions on convex  
 119 analysis, proximity operators, and monotone operators. Most of our notation follow from the  
 120 reference book [4].

121 We consider optimization problems over real Hilbert spaces. We denote by  $\mathcal{H}$ ,  $\mathcal{G}$ ,  $\mathcal{L}$  some real  
 122 Hilbert spaces, and  $\mathcal{B}(\mathcal{H}, \mathcal{G})$  the set of bounded and linear operators from  $\mathcal{H}$  to  $\mathcal{G}$ . The norm of  $\mathcal{H}$   
 123 is denoted by  $\|\cdot\|_{\mathcal{H}}$  and the associated inner product  $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ . The operator norm of  $D \in \mathcal{B}(\mathcal{H}, \mathcal{G})$   
 124 is denoted by  $\|D\|_{\mathcal{H}, \mathcal{L}} = \sup \{\|Dx\|_{\mathcal{L}} \mid x \in \mathcal{H}, \|x\|_{\mathcal{H}} = 1\}$ . Given  $C \subset \mathcal{H}$ ,  $\text{sri } C$  denotes its strong  
 125 relative interior.

126  $2^{\mathcal{H}}$  denotes the power set of  $\mathcal{H}$ . Given a set-valued operator  $\mathcal{M} : \mathcal{H} \rightarrow 2^{\mathcal{H}}$ , we define  $\text{dom } \mathcal{M} =$   
 127  $\{x \in \mathcal{H} \mid \mathcal{M}x \neq \emptyset\}$  as its domain, and  $\text{ran } \mathcal{M}$  and  $\text{gra } \mathcal{M}$  as its range and graph, respectively.

128  $\mathcal{M}^{-1} : \mathcal{H} \rightarrow 2^{\mathcal{H}}$  denotes the inverse operator of  $\mathcal{M}$ , with domain  $\text{ran } \mathcal{M}$  and range  $\text{dom } \mathcal{M}$ , and  
 129  $\mathcal{M}^{-1}(0) = \{x \in \mathcal{H} \mid 0 \in \mathcal{M}x\}$  the set of zeros of  $\mathcal{M}$ .

130  $\mathcal{M}$  is said to be monotone when, for every  $(x, y) \in \text{gra } \mathcal{M}$  and  $(x', y') \in \text{gra } \mathcal{M}$ ,  $\langle y' - y, x' - x \rangle_{\mathcal{H}} \geq 0$ .

131  $\mathcal{M}$  is said to be maximal when  $\text{gra } \mathcal{M}$  is maximal in the sense of the inclusion in  $\mathcal{H} \times \mathcal{H}$  among  
 132 the graphs of monotone operators. Moreover,  $\mathcal{M}$  is strongly monotone with parameter  $\rho \geq 0$  if for  
 133 every  $(x, y) \in \text{gra } \mathcal{M}$  and  $(x', y') \in \text{gra } \mathcal{M}$ ,  $\langle y' - y, x' - x \rangle_{\mathcal{H}} \geq \rho \|x' - x\|_{\mathcal{H}}^2$ .

134 The resolvent of a maximally monotone operator  $\mathcal{M}$  scaled with a parameter  $\gamma > 0$  is the mapping

135  $J_{\gamma\mathcal{M}} : \mathcal{H} \rightarrow \mathcal{H} : x \mapsto J_{\gamma\mathcal{M}}(x) = (\text{Id}_{\mathcal{H}} + \gamma\mathcal{M})^{-1}x$ , where  $\text{Id}_{\mathcal{H}}$  refers to the identity operator in  $\mathcal{H}$ .  
 136 An operator  $\mathcal{C} : \mathcal{H} \rightarrow \mathcal{H}$  is cocoercive if there exists  $\Theta > 0$  such that for every  $(x, x') \in \mathcal{H}^2$ ,  
 137  $\langle \mathcal{C}x' - \mathcal{C}x, x' - x \rangle_{\mathcal{H}} \geq \Theta \|\mathcal{C}x' - \mathcal{C}x\|_{\mathcal{H}}^2$ .  
 138 The Fenchel-Legendre conjugate of a function  $f : \mathcal{H} \rightarrow ]-\infty, +\infty]$ , is expressed as  $f^* : \mathcal{H} \rightarrow$   
 139  $[-\infty, +\infty] : u \mapsto \sup_{x \in \mathcal{H}} (\langle x, u \rangle_{\mathcal{H}} - f(x))$ . The domain of  $f$  is  $\text{dom } f = \{x \in \mathcal{H} \mid f(x) < +\infty\}$  and this  
 140 function is said to be proper when  $\text{dom } f \neq \emptyset$ .  
 141 The class of convex, lower semicontinuous, and proper function defined on  $\mathcal{H}$  is denoted by  $\Gamma_0(\mathcal{H})$   
 142 and the Moreau subdifferential of any function  $f$  in  $\Gamma_0(\mathcal{H})$  is denoted by  $\partial f$ , while the proximity  
 143 operator of  $f$  is given by  $\text{prox}_f : \mathcal{H} \rightarrow \mathcal{H} : x \mapsto \text{argmin}_{v \in \mathcal{H}} \left( f(v) + \frac{1}{2} \|v - x\|^2 \right)$ .  
 144 From [4, Corollary 16.30] we have  $(\partial f)^{-1} = \partial f^*$  and, if  $f$  is strongly convex with modulus  $\sigma \geq 0$ ,  
 145 then  $\partial f$  is  $\sigma$ -strongly monotone [4, Example 22.4 (iv)].  
 146 Finally, given  $\alpha \in ]0, 1[$ , an operator  $T : \mathcal{H} \rightarrow \mathcal{H}$  is  $\alpha$ -averaged if, for every  $(x, y) \in \mathcal{H}^2$ , there exists  
 147 a nonexpansive operator  $\mathcal{R} : \mathcal{H} \rightarrow \mathcal{H}$  such that  $T = (1 - \alpha)\text{Id}_{\mathcal{H}} + \alpha\mathcal{R}$ .

148 **2.2. Problem statement.** Let us now state the generic minimization problem considered in this  
 149 paper. Let  $H \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ ,  $D \in \mathcal{B}(\mathcal{H}, \mathcal{L})$ ,  $f \in \Gamma_0(\mathcal{H})$ , and  $g \in \Gamma_0(\mathcal{L})$ . Given  $y \in \mathcal{G}$ , we are interested  
 150 in solving the primal optimization problem

$$151 \quad (2.1) \quad \underset{x \in \mathcal{H}}{\text{minimize}} \quad \frac{1}{2} \|y - Hx\|_{\mathcal{G}}^2 + f(x) + g(Dx).$$

152 By using first-order optimality conditions [4, Theorems 16.47, 27.1], Problem (2.1) can be reformu-  
 153 lated as finding a solution to the variational inclusion problem

$$154 \quad (2.2) \quad 0 \in \partial f(x) + D^* \partial g(Dx) + H^*(Hx - y).$$

155 A solution to (2.2) is a solution to (2.1), but the converse may not be true. From now on, the  
 156 solution set of (2.2) is supposed to be nonempty, and then so is the solution set of (2.1).

157 Under mild qualification constraints, the solutions to (2.1) and (2.2) are the same. One of such  
 158 qualification conditions is

$$159 \quad (2.3) \quad 0 \in \text{sri}(D(\text{dom } f) - \text{dom } g).$$

160 We refer the reader to [4, Proposition 27.5, Corollary 27.6] for other examples of qualification  
 161 constraints.

162 The dual problem associated with (2.1) reads

$$163 \quad (2.4) \quad \underset{u \in \mathcal{L}}{\text{minimize}} \quad (f + h)^*(-D^*u) + g^*(u),$$

164 where  $h : \mathcal{H} \rightarrow \mathbb{R} : x \mapsto \frac{1}{2} \|y - Hx\|_{\mathcal{G}}^2$ .

165 According to [4, Theorem 19.1],  $x$  is a solution to (2.1) and  $u$  is a solution to the dual problem  
 166 associated with (2.1) if  $z = (x, u) \in \mathcal{Z} = \mathcal{H} \times \mathcal{L}$  is a solution to

$$167 \quad (2.5) \quad \begin{pmatrix} 0 \\ 0 \end{pmatrix} \in \underbrace{\begin{pmatrix} \partial f(x) + D^*u \\ -Dx + \partial g^*(u) \end{pmatrix}}_{\mathcal{A}z} + \underbrace{\begin{pmatrix} H^*Hx - H^*y \\ 0 \end{pmatrix}}_{\mathcal{B}z}.$$

168 The operator  $\mathcal{A} : \mathcal{Z} \rightarrow 2^{\mathcal{Z}} : (x, u) \mapsto (\partial f(x) + D^*u, -Dx + \partial g^*(u))$  is maximally monotone [4,  
 169 Proposition 26.32 (iii)]. Moreover, the operator  $\mathcal{B} : \mathcal{Z} \rightarrow \mathcal{Z} : (x, u) \mapsto (H^*Hx - H^*y, 0)$  is  $\theta$ -  
 170 cocoercive, with  $\theta = 1/\|H\|_{\mathcal{H}, \mathcal{G}}^2$ . Based on this structure, the forward-backward splitting [24] can

171 be applied to solve (2.5). Combined with preconditioning [26], it yields a simple form of iteration  
 172 in terms of primal/dual updates.

173 On the other hand, (2.5) is also equivalent to the following splitting

$$174 \quad (2.6) \quad \begin{pmatrix} 0 \\ 0 \end{pmatrix} \in \underbrace{\begin{pmatrix} \partial f(x) \\ \partial g^*(u) \end{pmatrix}}_{\mathcal{M}z} + \underbrace{\begin{pmatrix} H^*Hx - H^*y + D^*u \\ -Dx \end{pmatrix}}_{\mathbf{Q}z},$$

175 where  $\mathcal{M}$  is a maximally monotone operator (see [4, Theorem 26.32 (iii)]) and  $\mathbf{Q}$  is a monotone  
 176 operator which is Lipschitz continuous with constant  $\vartheta \leq \|H\|_{\mathcal{H}}^2 + \|D\|_{\mathcal{H},\mathcal{L}}$  (see [22]). We can use  
 177 Tseng's approach [4, Theorem 26.17], [9, Theorem 2.5 (ii)] to build primal-dual schemes delivering  
 178 a solution to splitting (2.6).

179 Both (2.5) and (2.6) splitting strategies will be at the basis of the primal-dual algorithms under  
 180 study. The Condat-Vũ (CV) algorithm (section 3) and the Loris-Verhoeven (LV) algorithm (sec-  
 181 tion 4) will be derived from (2.5) while the Combettes-Pesquet (CP) algorithm (section 5) will rely  
 182 on (2.6).

### 183 3. The mismatched Condat-Vũ algorithm.

184 **3.1. CV algorithm.** In this section, we use the forward-backward splitting to solve (2.6). The  
 185 resulting iteration by the implicit and preconditioned discretization of (2.5) reads: for every  $n \in \mathbb{N}$ ,

$$186 \quad (3.1) \quad 0 \in \mathcal{A}z_{n+\frac{1}{2}} + \mathcal{B}z_n + P(z_{n+\frac{1}{2}} - z_n),$$

187 where  $z_n = (x_n, u_n) \in \mathcal{Z}$ , and  $z_{n+\frac{1}{2}} = (x_{n+\frac{1}{2}}, u_{n+\frac{1}{2}}) \in \mathcal{Z}$ . The preconditioning metric  $P$  is defined  
 188 as

$$189 \quad (3.2) \quad (\forall z = (x, u) \in \mathcal{Z}) \quad Pz = \begin{pmatrix} \frac{1}{\tau}x - D^*u \\ -Dx + \frac{1}{\sigma}u \end{pmatrix},$$

190 where  $\tau$  and  $\sigma$  are two positive real parameters and initialization  $x_0 \in \mathcal{H}$  and  $u_0 \in \mathcal{L}$ .

191 Let  $\{\Theta_n\}_{n \in \mathbb{N}}$  be a sequence of relaxation parameters. The corresponding primal-dual forward-  
 192 backward iteration [4], often called the Condat-Vũ (CV) iteration in the literature [56, 25], reads

193 **CV iterations for solving (2.5):**

$$194 \quad (3.3) \quad \text{for } n = 0, 1, \dots \quad \begin{cases} x_{n+\frac{1}{2}} = \text{prox}_{\tau f}(x_n - \tau(H^*(Hx_n - y) + D^*u_n)) \\ u_{n+\frac{1}{2}} = \text{prox}_{\sigma g^*}(u_n + \sigma D(2x_{n+\frac{1}{2}} - x_n)) \\ x_{n+1} = x_n + \Theta_n(x_{n+\frac{1}{2}} - x_n) \\ u_{n+1} = u_n + \Theta_n(u_{n+\frac{1}{2}} - u_n). \end{cases}$$

196 The convergence of (3.3) to a solution pair  $(x, u)$  to (2.5) is guaranteed by [25, Theorem 3.1] for step  
 197 sizes  $\tau > 0$ ,  $\sigma > 0$  such that  $\tau \left( \sigma \|D\|_{\mathcal{H},\mathcal{L}}^2 + \frac{1}{2\theta} \right) < 1$ , and relaxation parameters  $\{\Theta_n\}_{n \in \mathbb{N}} \subset [0, \delta]$

198 such that  $\sum_{n \in \mathbb{N}} \Theta_n (\delta - \Theta_n) = +\infty$ , with  $\delta = 2 - \frac{1}{2\theta} \left( \frac{1}{\tau} - \sigma \|D\|_{\mathcal{H},\mathcal{L}}^2 \right)$  and  $\theta$  a cocoercivity constant  
 199 of operator  $\mathcal{B}$ .

200 *Remark 3.1.* The condition on the step sizes  $(\tau, \sigma)$  implies that  $\tau\sigma\|D\|_{\mathcal{H},\mathcal{L}}^2 < 1$ , which allows us  
 201 to conclude that operator  $P$  present in (3.1) is strongly positive (bounded from below) and therefore  
 202 invertible. Hereafter, we will consider the Hilbert space  $\mathcal{Z}_P$  obtained by equipping  $\mathcal{Z}$  with the inner  
 203 product  $\langle \cdot, \cdot \rangle_P: (z, z') \mapsto \langle z, z' \rangle_P = \langle z, Pz' \rangle_{\mathcal{Z}}$ . Hence, operators  $P^{-1}\mathcal{A}$  and  $P^{-1}\mathcal{B}$  are maximally  
 204 monotone [4, Proposition 20.24] and cocoercive [25, Theorem 3.1] in  $\mathcal{Z}_P$ , respectively.

205 **3.2. Convergence properties of the mismatched CV algorithm.** In this section, we will be  
 206 interested in studying the impact of replacing operator  $H^*$  in the  $n$ -th iteration of Algorithm (3.3)  
 207 by a surrogate  $K_n \in \mathcal{B}(\mathcal{G}, \mathcal{H})$ . This yields the following modified algorithm:

208 **Mismatched CV iterations:**

$$209 \quad (3.4) \quad \text{for } n = 0, 1, \dots \quad \begin{cases} x_{n+\frac{1}{2}} = \text{prox}_{\tau f}(x_n - \tau(K_n(Hx_n - y) + D^*u_n)) \\ u_{n+\frac{1}{2}} = \text{prox}_{\sigma g^*}(u_n + \sigma D(2x_{n+\frac{1}{2}} - x_n)) \\ x_{n+1} = x_n + \Theta_n(x_{n+\frac{1}{2}} - x_n) \\ u_{n+1} = u_n + \Theta_n(u_{n+\frac{1}{2}} - u_n), \end{cases}$$

211 The sequence of surrogates  $(K_n)_{n \in \mathbb{N}}$  is related to a constant linear operator  $\bar{K}$  through the following  
 212 assumption, which is similar to [18, Assumption 3.1],

213 **Assumption 3.2.** *There exist  $\bar{K} \in \mathcal{B}(\mathcal{G}, \mathcal{H})$  and  $\{\omega_n\}_{n \in \mathbb{N}} \subset [0, +\infty[$  such that  $\sum_{n \in \mathbb{N}} \omega_n < +\infty$*   
 214 *and*

$$215 \quad (3.5) \quad \bar{K}H \neq 0$$

$$216 \quad (3.6) \quad (\forall n \in \mathbb{N}) \quad \|K_n - \bar{K}\|_{\mathcal{G}, \mathcal{H}} \leq \omega_n.$$

218 To perform the convergence analysis of (3.4), we introduce notation involved in the characterization  
 219 of the spectra of the linear operators in the cost function.

220 **Notation 3.3.** *For every  $n \in \mathbb{N}$ , we define the following quantities:*

- 221 (i)  $L = \bar{K}H$
- 222 (ii)  $\lambda_{\min} = \inf \{ \langle x, Lx \rangle_{\mathcal{H}} \mid x \in \mathcal{H}, \|x\|_{\mathcal{H}} = 1 \}$
- 223 (iii)  $\tilde{L} = \Pi L$ ,  $\tilde{L}_n = \Pi K_n H$ ,  $\tilde{K} = \Pi K$ , and  $\tilde{K}_n = \Pi K_n$ , with  $\Pi: \mathcal{Z} \rightarrow \mathcal{Z}: (x, u) \mapsto (x, 0)$ .
- 224 (iv)  $T_n: \mathcal{Z} \rightarrow \mathcal{Z}: z \mapsto J_{P^{-1}\mathcal{A}} \left( z - P^{-1} \left( \tilde{L}_n z - \tilde{K}_n y \right) \right)$  and
- 225  $\bar{T}: \mathcal{Z} \rightarrow \mathcal{Z}: z \mapsto J_{P^{-1}\mathcal{A}} \left( z - P^{-1} \left( \tilde{L} z - \tilde{K} y \right) \right).$

226 Under this notation, Algorithm (3.4) can be rewritten more concisely in terms of a simple update  
 227 rule on the pair  $z_n = (x_n, u_n)$ :

$$228 \quad (3.7) \quad (\forall n \in \mathbb{N}) \quad z_{n+1} = z_n + \Theta_n(T_n(z_n) - z_n).$$

229 Algorithm (3.7) can be viewed as a mismatched form of a forward-backward algorithm for finding a  
 230 zero of a sum of maximally monotone operators in  $\mathcal{Z}_P$ . Similarly to [18], we rely on the cocoercivity  
 231 properties of operator  $\tilde{L}$  to study the convergence of (3.7).

232 **Proposition 3.4.** *Assume that  $(\tau, \sigma) \in ]0, +\infty[^2$  are such that  $\tau\sigma\|D\|_{\mathcal{H}, \mathcal{L}}^2 < 1$ . We have the fol-*  
 233 *lowing properties.*

- 234 (i)  $P^{-1}\tilde{L}$  is cocoercive in  $\mathcal{Z}_P$  if and only if  $L$  is cocoercive.
- 235 (ii) Suppose that  $\text{ran}(L + L^*)$  is closed. Then  $P^{-1}\tilde{L}$  is cocoercive in  $\mathcal{Z}_P$  with constant  $\tilde{\eta} > 0$  if  
 236 and only if  $\lambda_{\min} \geq 0$ ,  $\text{Ker}(L + L^*) = \text{Ker} L$ , and

$$237 \quad (3.8) \quad \tilde{\eta} \leq \tilde{\eta}_{\max} = \frac{2}{\|P^{-1/2}\Pi M\|_{\mathcal{H}, \mathcal{Z}}^2},$$

238 where

$$239 \quad (3.9) \quad M = (\text{Id}_{\mathcal{H}} + (L - L^*)(L + L^*)^\#)(L + L^*)^{1/2}$$

240 and  $(L + L^*)^\#$  denotes the pseudo-inverse of  $L + L^*$ . In addition,

$$241 \quad (3.10) \quad \tilde{\eta}_{\max} \geq \tau^{-1}(1 - \tau\sigma\|D\|_{\mathcal{H},\mathcal{L}}^2)\eta_{\max}$$

242 where  $\eta_{\max} = 2/\|M\|_{\mathcal{H},\mathcal{H}}^2$  is the largest cocoercivity constant of  $L$ .

243 *Proof.*

244 (i)  $P^{-1}\tilde{L}$  is cocoercive in  $\mathcal{Z}_P$  if and only if there exists  $\tilde{\eta} > 0$  such that, for every  $z \in \mathcal{Z}$ ,

$$245 \quad (\forall z \in \mathcal{Z}) \quad \langle z, P^{-1}\tilde{L}z \rangle_{\mathcal{Z}_P} \geq \tilde{\eta}\|P^{-1}\tilde{L}z\|_{\mathcal{Z}_P}^2$$

$$246 \quad (3.11) \quad \Leftrightarrow \quad (\forall z' \in \mathcal{Z}) \quad \langle z', P^{-1/2}\tilde{L}P^{-1/2}z' \rangle_{\mathcal{Z}} \geq \tilde{\eta}\|P^{-1/2}\tilde{L}P^{-1/2}z'\|_{\mathcal{Z}}^2.$$

248 Therefore  $P^{-1}\tilde{L}$  is cocoercive in  $\mathcal{Z}_P$  if and only if  $P^{-1/2}\tilde{L}P^{-1/2}$  is cocoercive in  $\mathcal{Z}$ . In turn,  
 249 it follows from [4, Proposition 4.12] that, if  $\tilde{L}$  is cocoercive, then  $P^{-1/2}\tilde{L}P^{-1/2}$  is cocoercive.  
 250 Conversely, if  $P^{-1/2}\tilde{L}P^{-1/2}$  is cocoercive, then  $P^{1/2}(P^{-1/2}\tilde{L}P^{-1/2})P^{1/2} = \tilde{L}$  is cocoercive,  
 251 that is

$$252 \quad (\forall z \in \mathcal{Z}) \quad \langle z, \tilde{L}z \rangle_{\mathcal{Z}} \geq \eta\|\tilde{L}z\|_{\mathcal{Z}}^2$$

$$253 \quad (3.12) \quad \Leftrightarrow \quad (\forall x \in \mathcal{H}) \quad \langle x, Lx \rangle_{\mathcal{H}} \geq \eta\|Lx\|_{\mathcal{H}}^2,$$

255 for some  $\eta > 0$ .

256 (ii) Let  $\tilde{L}_P^*$  denote the adjoint of  $P^{-1}\tilde{L}$  in  $\mathcal{Z}_P$  and let

$$257 \quad (3.13) \quad \tilde{\lambda}_{\min} = \inf \left\{ \langle z, P^{-1}\tilde{L}z \rangle_{\mathcal{Z}_P} \mid z \in \mathcal{Z}_P, \|z\|_{\mathcal{Z}_P} = 1 \right\}.$$

258 According to [18, Proposition 3.4(ii)], provided that  $\text{ran}(P^{-1}\tilde{L} + \tilde{L}_P^*)$  is closed,  $P^{-1}\tilde{L}$  is  
 259 cocoercive in  $\mathcal{Z}_P$  with constant  $\tilde{\eta}$  if and only if  $\tilde{\lambda}_{\min} \geq 0$ ,  $\text{Ker}(P^{-1}\tilde{L} + \tilde{L}_P^*) = \text{Ker}(P^{-1}\tilde{L})$ ,  
 260 and

$$261 \quad (3.14) \quad \tilde{\eta} \leq \frac{2}{\|(\text{Id}_{\mathcal{Z}} + (P^{-1}\tilde{L} - \tilde{L}_P^*)(P^{-1}\tilde{L} + \tilde{L}_P^*)^\#)(P^{-1}\tilde{L} + \tilde{L}_P^*)^{1/2}\|_{\mathcal{Z}_P, \mathcal{Z}_P}^2}.$$

262 Since  $\tilde{L}_P^* = P^{-1}\tilde{L}^*$ ,  $\text{Ker}(P^{-1}\tilde{L} + \tilde{L}_P^*) = \text{Ker}(\tilde{L} + \tilde{L}^*) = \text{Ker}(L + L^*) \times \mathcal{L}$  and, similarly,  
 263  $\text{Ker}(P^{-1}\tilde{L}) = \text{Ker} L \times \mathcal{L}$ . Therefore  $\text{Ker}(P^{-1}\tilde{L} + \tilde{L}_P^*) = \text{Ker}(P^{-1}\tilde{L})$  if and only if  $\text{Ker}(L + L^*) = \text{Ker} L$ .  
 264 Similarly,  $\text{ran}(P^{-1}\tilde{L} + \tilde{L}_P^*) = P^{-1}\text{ran}(\tilde{L} + \tilde{L}^*) = P^{-1}(\text{ran}(L + L^*) \times \{0\})$ .  
 265 Thus,  $\text{ran}(P^{-1}\tilde{L} + \tilde{L}_P^*)$  is closed if and only if  $\text{ran}(L + L^*)$  is closed.

266 For every  $z = (x, u) \in \mathcal{Z}$ ,

$$267 \quad \langle z, P^{-1}\tilde{L}z \rangle_{\mathcal{Z}_P} \geq \tilde{\lambda}_{\min}\|z\|_{\mathcal{Z}_P}^2$$

$$268 \quad \Leftrightarrow \quad \langle z, \tilde{L}z \rangle_{\mathcal{Z}} \geq \tilde{\lambda}_{\min}\|z\|_{\mathcal{Z}_P}^2$$

$$269 \quad (3.15) \quad \Leftrightarrow \quad \langle x, Lx \rangle_{\mathcal{H}} \geq \tilde{\lambda}_{\min}\|(x, u)\|_{\mathcal{Z}_P}^2.$$

271 So,  $\tilde{\lambda}_{\min} \geq 0$  if and only if, for every  $x \in \mathcal{H}$ ,  $\langle x, Lx \rangle_{\mathcal{H}} \geq 0$ , that is  $\lambda_{\min} \geq 0$ . In addition,  
 272 when condition  $\lambda_{\min} \geq 0$  is met,  $\tilde{\lambda}_{\min} = 0$ .

273 Since  $\|\cdot\|_{\mathcal{Z}_P, \mathcal{Z}_P} = \|P^{1/2} \cdot P^{-1/2}\|_{\mathcal{Z}, \mathcal{Z}}$ , we have

$$274 \quad \|(\text{Id}_{\mathcal{Z}} + (P^{-1}\tilde{L} - \tilde{L}_P^*)(P^{-1}\tilde{L} + \tilde{L}_P^*)^\#)(P^{-1}\tilde{L} + \tilde{L}_P^*)^{1/2}\|_{\mathcal{Z}_P, \mathcal{Z}_P}^2$$

$$275 \quad = \|(\text{Id}_{\mathcal{Z}} + (P^{-1}\tilde{L} - \tilde{L}_P^*)(P^{-1}\tilde{L} + \tilde{L}_P^*)^\#)(P^{-1}\tilde{L} + \tilde{L}_P^*)(\text{Id}_{\mathcal{Z}} + (P^{-1}\tilde{L} + \tilde{L}_P^*)^\#)(\tilde{L}_P^* - P^{-1}\tilde{L})\|_{\mathcal{Z}_P, \mathcal{Z}_P}$$

$$276 \quad = \|P^{1/2}(\text{Id}_{\mathcal{Z}} + P^{-1}(\tilde{L} - \tilde{L}^*)(\tilde{L} + \tilde{L}^*)^\#)P^{-1}(\tilde{L} + \tilde{L}^*)(\text{Id}_{\mathcal{Z}} + (\tilde{L} + \tilde{L}^*)^\#)PP^{-1}(\tilde{L}^* - \tilde{L})P^{-1/2}\|_{\mathcal{Z}, \mathcal{Z}}$$

$$277 \quad = \|P^{-1/2}(\text{Id}_{\mathcal{Z}} + (\tilde{L} - \tilde{L}^*)(\tilde{L} + \tilde{L}^*)^\#)(\tilde{L} + \tilde{L}^*)(\text{Id}_{\mathcal{Z}} + (\tilde{L} + \tilde{L}^*)^\#)(\tilde{L}^* - \tilde{L})P^{-1/2}\|_{\mathcal{Z}, \mathcal{Z}}$$

$$278 \quad = \|P^{-1/2}(\text{Id}_{\mathcal{Z}} + (\tilde{L} - \tilde{L}^*)(\tilde{L} + \tilde{L}^*)^\#)(\tilde{L} + \tilde{L}^*)^{1/2}\|_{\mathcal{Z}, \mathcal{Z}}^2. \quad \blacksquare$$



280 By using the specific form of  $\tilde{L}$ , we deduce that

$$281 \quad (3.16) \quad \|(\text{Id}_{\mathcal{Z}} + (P^{-1}\tilde{L} - \tilde{L}_P^*)(P^{-1}\tilde{L} + \tilde{L}_P^*)^\#)(P^{-1}\tilde{L} + \tilde{L}_P^*)^{1/2}\|_{\mathcal{Z}_P, \mathcal{Z}_P}^2 = \|P^{-1/2}\Pi M\|_{\mathcal{H}, \mathcal{Z}}^2.$$

283 Altogether with (3.14), this yields (3.8).

284 In addition,

$$285 \quad (3.17) \quad \|P^{-1/2}\Pi M\|_{\mathcal{H}, \mathcal{Z}} \leq \|P^{-1/2}\Pi\|_{\mathcal{Z}, \mathcal{Z}} \|M\|_{\mathcal{H}, \mathcal{H}} = \|P^{-1/2}\Pi\|_{\mathcal{Z}, \mathcal{Z}} \sqrt{\frac{2}{\eta_{\max}}},$$

286 where

$$287 \quad \|P^{-1/2}\Pi\|_{\mathcal{Z}, \mathcal{Z}} = \|\Pi^* P^{-1/2}\|_{\mathcal{Z}, \mathcal{Z}} = \|\Pi P^{-1/2}\|_{\mathcal{Z}, \mathcal{Z}} \\ 288 \quad = \sup_{z' \in \mathcal{Z} \setminus \{0\}} \frac{\|\Pi P^{-1/2} z'\|_{\mathcal{Z}}}{\|z'\|_{\mathcal{Z}}} \\ 289 \quad (3.18) \quad = \sup_{z=(x,u) \in \mathcal{Z} \setminus \{0\}} \frac{\|x\|_{\mathcal{H}}}{\sqrt{\langle z, Pz \rangle_{\mathcal{Z}}}}.$$

291 For every  $z = (x, u) \in \mathcal{Z}$ ,

$$292 \quad \langle z, Pz \rangle_{\mathcal{Z}} = \frac{1}{\tau} \|x\|_{\mathcal{H}}^2 - 2\langle Dx, u \rangle_{\mathcal{L}} + \frac{1}{\sigma} \|u\|_{\mathcal{L}}^2 \\ 293 \quad = \frac{1}{\tau} \|x\|_{\mathcal{H}}^2 + \frac{1}{\sigma} \|u - \sigma Dx\|_{\mathcal{L}}^2 - \sigma \|Dx\|_{\mathcal{L}}^2 \\ 294 \quad (3.19) \quad \geq \frac{1}{\tau} (1 - \tau\sigma \|D\|_{\mathcal{H}, \mathcal{L}}^2) \|x\|_{\mathcal{H}}^2.$$

296 We deduce from (3.18) and (3.19) that

$$297 \quad (3.20) \quad \|P^{-1/2}\Pi\|_{\mathcal{Z}, \mathcal{Z}}^2 \leq \frac{\tau}{1 - \tau\sigma \|D\|_{\mathcal{H}, \mathcal{L}}^2},$$

298 which, combined with (3.17), yields (3.10). ■

299 The following results provides a characterization of the fixed point set of  $\bar{T}$ .

300 **Proposition 3.5.** *Let  $(\tilde{x}, \tilde{u}) \in \mathcal{Z}$ . Then  $(\tilde{x}, \tilde{u}) \in \text{Fix}(\bar{T})$  if and only if  $(\tilde{x}, \tilde{u})$  belongs to*

$$301 \quad (3.21) \quad \bar{\mathcal{F}} = \{(x, u) \in \mathcal{Z} \mid \bar{K}y \in \partial f(x) + Lx + D^*u, u \in \partial g(Dx)\},$$

302 which is nonempty if  $L + \partial f + D^* \circ \partial g \circ D$  is surjective.

303 (i) If  $\lambda_{\min} \geq 0$ , then  $\bar{\mathcal{F}}$  is closed and convex.

304 (ii) Let  $\bar{\mathcal{F}}_1 = \{x \in \mathcal{H} \mid (\exists u \in \mathcal{L}) (x, u) \in \bar{\mathcal{F}}\}$ .

305  $\bar{\mathcal{F}}_1$  has at most one element if one of the following conditions holds:

306 (a)  $L + \partial f + D^* \circ \partial g \circ D$  is strictly monotone.

307 (b)  $\lambda_{\min} \geq 0$  and  $g \circ D + f$  is strictly convex.

308  $\bar{\mathcal{F}}_1$  is a singleton if (2.3) is satisfied and one of the following conditions holds:

309 (c)  $\lambda_{\min} \geq 0$  and  $L + \partial f + D^* \circ \partial g \circ D$  is strongly monotone.

310 (d)  $\lambda_{\min} > 0$ .

311 (e)  $\lambda_{\min} \geq 0$ , and  $f$  is strongly convex or [ $g$  is strongly convex and  $D^*D$  is strongly positive].

312 (f)  $L$  is cocoercive, and  $f$  is strongly convex or [ $g$  is strongly convex and  $D^*D$  is strongly positive].

*Proof.*

$$315 \quad \tilde{z} = (\tilde{x}, \tilde{u}) \in \text{Fix}(\bar{T}) \Leftrightarrow \tilde{z} = \bar{T}\tilde{z} \Leftrightarrow 0 \in P^{-1}(\tilde{L}\tilde{z} - \tilde{K}y) + P^{-1}\mathcal{A}\tilde{z} \Leftrightarrow \bar{K}y \in \partial f(\tilde{x}) + L\tilde{x} + D^*\tilde{u}, \quad D\tilde{x} \in \partial g^*(\tilde{u}),$$

316 that is  $\tilde{z} \in \bar{\mathcal{F}}$ . In addition,

$$317 \quad (3.22) \quad (\exists \tilde{u} \in \mathcal{L}) \quad \begin{cases} \bar{K}y \in \partial f(\tilde{x}) + L\tilde{x} + D^*\tilde{u} \\ D\tilde{x} \in \partial g^*(\tilde{u}) \end{cases} \Leftrightarrow \bar{K}y \in \partial f(\tilde{x}) + L\tilde{x} + D^*\partial g(D\tilde{x}),$$

318 and the latter condition is satisfied if  $\partial f + L + D^* \circ \partial g \circ D$  is surjective.

319 (i) If  $\lambda_{\min} \geq 0$ , then, for every  $z = (x, u) \in \mathcal{Z}$ ,  $\langle \tilde{L}z, z \rangle_{\mathcal{Z}} = \langle Lx, x \rangle_{\mathcal{H}} \geq 0$  and, since  $\tilde{L}$  is  
320 continuous,  $\tilde{L}$  is maximally monotone on  $\mathcal{Z}$ . Operator  $\mathcal{A}$  being maximally monotone, by [4,  
321 Proposition 23.39], we conclude that  $\bar{\mathcal{F}} = \text{zer}(\tilde{L} - \tilde{K}y + \mathcal{A})$  is closed and convex.

322 (ii) According to (3.22),  $\bar{\mathcal{F}}_1 = \text{zer}(L - \bar{K}y + \partial f + D^* \circ \partial g \circ D)$ .

323 (a) Follows from [4, Proposition 23.35].

324 (b) From standard subdifferential calculus rules,  $\bar{\mathcal{F}}_1 \subset \text{zer}(L - Ky + \partial(f + g \circ D))$ .  
325  $\lambda_{\min} \geq 0 \Leftrightarrow L + L^*$  positive  $\Leftrightarrow L$  is monotone. In addition,  $f + g \circ D$  is strictly  
326 convex if and only if  $\partial(f + g \circ D)$  is strictly monotone. Thus, under the stated  
327 condition,  $L - Ky + \partial(f + g \circ D)$  is strictly monotone, and  $\bar{\mathcal{F}}_1$  has at most one  
328 element.

329 (c) Since (2.3) holds, it follows from [4, Theorem 16.47] that  $\partial f + D^* \circ \partial g \circ D = \partial(f + g \circ D)$   
330 is maximally monotone. Since  $L$  is maximally monotone and has a full domain,  
331  $L + \partial f + D^* \circ \partial g \circ D$  is thus maximally monotone. Then we deduce the result from  
332 [4, Proposition 23.37].

333 (d) If  $\lambda_{\min} > 0$ ,  $L$  is strongly monotone and (c) is satisfied.

334 (e) If  $g$  is  $\rho_G$ -strongly convex with  $\rho_G > 0$ , there exists  $h \in \Gamma_0(\mathcal{L})$  such that  $g =$   
335  $h + \rho_G \|\cdot\|_{\mathcal{L}}^2/2$  and  $D^* \circ \partial g \circ D = D^* \circ \partial h \circ D + \rho_G D^* D$  is strongly monotone as  
336  $D^* D$  is strongly positive. If  $f$  is strongly convex, then  $\partial f$  is strongly monotone. The  
337 result then follows from (c).

338 (f) Since  $L$  is cocoercive, from [18, Lemma 3.3(i)],  $\lambda_{\min} \geq 0$  and the result follows  
339 from (e). ■

340 We will now turn our attention to the convergence of the mismatched CV algorithm by estab-  
341 lishing a first result regarding the averagedness properties of operator  $\bar{T}$ .

342 **Lemma 3.6.** *Let  $\tilde{\eta} \in ]\frac{1}{2}, +\infty[$ ,  $\bar{\alpha} = \frac{1}{2 - \frac{1}{2\tilde{\eta}}}$ , and  $\bar{W} = \text{Id}_{\mathcal{Z}_P} - P^{-1}\tilde{L}$ . Then, if  $P^{-1}\tilde{L}$  is  $\tilde{\eta}$ -cocoercive  
343 in  $\mathcal{Z}_P$ , then  $\bar{T}$  is  $\bar{\alpha}$ -averaged in  $\mathcal{Z}_P$ .*

344 *Proof.* If  $P^{-1}\tilde{L}$  is  $\tilde{\eta}$ -cocoercive in  $\mathcal{Z}_P$  and  $\tilde{\eta} > 1/2$ , according to [4, Proposition 4.39],  $\bar{W}$  is  
345  $\frac{1}{2\tilde{\eta}}$ -averaged in  $\mathcal{Z}_P$ . Similarly to the proof of [18, Lemma 3.14], we deduce that

$$346 \quad (3.23) \quad (\forall z \in \mathcal{Z}_P) \quad \|\bar{W}z - 2(1 - \bar{\alpha})z\|_{\mathcal{Z}_P} + \|\bar{W}z\|_{\mathcal{Z}_P} \leq 2\bar{\alpha}\|z\|_{\mathcal{Z}_P}.$$

347 Since  $P^{-1}\mathcal{A}$  is maximally monotone on  $\mathcal{Z}_P$ , then  $J_{P^{-1}\mathcal{A}}$  is firmly nonexpansive [4, Corollary 23.9].  
348 Finally, it follows from (3.23) and [20, Theorem 3.8], that  $\bar{T} = J_{P^{-1}\mathcal{A}}(\bar{W} \cdot + P^{-1}\tilde{K}y)$  is  $\bar{\alpha}$ -averaged. ■

349 The following theorem provides conditions under which iteration (3.4) converges to a fixed point of  
350  $\bar{T}$ .

351 **Theorem 3.7.** *Let  $(\tau, \sigma) \in ]0, +\infty[^2$  be such that  $\tau\sigma\|D\|_{\mathcal{H}, \mathcal{L}}^2 < 1$ . Assume that  $\tilde{\eta} \in ]1/2, +\infty[$  is a  
352 cocoercivity constant of  $P^{-1}\tilde{L}$  in  $\mathcal{Z}_P$ . For  $\delta = 2 - 1/(2\tilde{\eta})$ , let  $\{\Theta_n\}_{n \in \mathbb{N}} \subset [0, \delta]$  be a sequence such*

353 that  $\sum_{n \in \mathbb{N}} \Theta_n (\delta - \Theta_n) = +\infty$ , and suppose that  $\overline{\mathcal{F}} \neq \emptyset$ . Then the sequence  $((x_n, u_n))_{n \in \mathbb{N}}$  given by

354 (3.4) converges weakly to some point in  $\overline{\mathcal{F}}$ .

355 *Proof.* Let  $z_0 = r_0 \in \mathcal{Z}$ , let  $(z_n)_{n \geq 1}$  be given by (3.7), and let  $(r_n)_{n \geq 1}$  be defined as

$$356 \quad (3.24) \quad (\forall n \in \mathbb{N}) \quad r_{n+1} = r_n + \Theta_n (\overline{T}(r_n) - r_n).$$

357 By applying Lemma 3.6, operator  $\overline{T}$  is  $\overline{\alpha}$ -averaged in  $\mathcal{Z}_P$  with  $\overline{\alpha} = 1/(2 - \frac{1}{2\eta})$ . Thus, the se-  
358 quence  $(r_n)_{n \in \mathbb{N}}$  converges weakly to some  $\bar{r} \in \text{Fix}(\overline{T})$  [4, Proposition 5.16], which implies that  
359  $\varrho = \sup_{n \in \mathbb{N}} \|r_n\|_{\mathcal{Z}_P} < +\infty$ .

360 For every  $n \in \mathbb{N}$ , let us bound  $z_{n+1} - r_{n+1}$  as follows

$$361 \quad \|z_{n+1} - r_{n+1}\|_{\mathcal{Z}_P} = \|z_n - r_n + \Theta_n (r_n - z_n) + \Theta_n (T_n(z_n) - \overline{T}(r_n))\|_{\mathcal{Z}_P}$$

$$362 \quad (3.25) \quad \leq \|z_n - r_n + \Theta_n (\overline{T}(z_n) - \overline{T}(r_n) - z_n + r_n)\|_{\mathcal{Z}_P} + \Theta_n \|T_n(z_n) - \overline{T}(z_n)\|_{\mathcal{Z}_P}.$$

363  $\overline{T}$  being  $\overline{\alpha}$ -averaged in  $\mathcal{Z}_P$ , there exist a nonexpansive operator  $\mathbf{W}: \mathcal{Z}_P \rightarrow \mathcal{Z}_P$  such that  $\overline{T} =$   
364  $(1 - \overline{\alpha})\text{Id} + \overline{\alpha}\mathbf{W}$ . Since  $\{\Theta_n\}_{n \in \mathbb{N}} \subset [0, 1/\overline{\alpha}]$ , we have

$$365 \quad \|z_n - r_n + \Theta_n (\overline{T}(z_n) - \overline{T}(r_n) - z_n + r_n)\|_{\mathcal{Z}_P} = \|(1 - \overline{\alpha}\Theta_n)(z_n - r_n) + \overline{\alpha}\Theta_n (\mathbf{W}(z_n) - \mathbf{W}(r_n))\|_{\mathcal{Z}_P}$$

$$366 \quad \leq (1 - \overline{\alpha}\Theta_n) \|z_n - r_n\|_{\mathcal{Z}_P} + \overline{\alpha}\Theta_n \|\mathbf{W}(z_n) - \mathbf{W}(r_n)\|_{\mathcal{Z}_P}$$

$$367 \quad (3.26) \quad \leq \|z_n - r_n\|_{\mathcal{Z}_P}.$$

368 Since  $J_{P^{-1}\mathcal{A}}$  is firmly nonexpansive in  $\mathcal{Z}_P$ , for every  $z \in \mathcal{Z}$ ,

$$369 \quad \|T_n(z) - \overline{T}(z)\|_{\mathcal{Z}_P} \leq \|P^{-1}(\tilde{L}_n z - \tilde{K}_n y - \tilde{L}z + \tilde{K}y)\|_{\mathcal{Z}_P}$$

$$370 \quad \leq \|P^{-1}(\tilde{L}_n - \tilde{L})z\|_{\mathcal{Z}_P} + \|P^{-1}(\tilde{K}_n - \tilde{K})y\|_{\mathcal{Z}_P}$$

$$371 \quad \leq \|P^{-1}(\tilde{L}_n - \tilde{L})\|_{\mathcal{Z}_P, \mathcal{Z}_P} \|z\|_{\mathcal{Z}_P} + \|P^{-1}(\tilde{K}_n - \tilde{K})\|_{\mathcal{Z}_P, \mathcal{Z}_P} \|y\|_{\mathcal{Z}_P}$$

$$372 \quad = \|P^{-1/2}(\tilde{L}_n - \tilde{L})P^{-1/2}\|_{\mathcal{Z}, \mathcal{Z}} \|z\|_{\mathcal{Z}_P} + \|P^{-1/2}(\tilde{K}_n - \tilde{K})P^{-1/2}\|_{\mathcal{Z}, \mathcal{Z}} \|y\|_{\mathcal{Z}_P}$$

$$373 \quad \leq \|P^{-1}\|_{\mathcal{Z}, \mathcal{Z}} (\|\tilde{L}_n - \tilde{L}\|_{\mathcal{Z}, \mathcal{Z}} \|z\|_{\mathcal{Z}_P} + \|\tilde{K}_n - \tilde{K}\|_{\mathcal{Z}, \mathcal{Z}} \|y\|_{\mathcal{Z}_P})$$

$$374 \quad = \|P^{-1}\|_{\mathcal{Z}, \mathcal{Z}} (\|L_n - L\|_{\mathcal{H}, \mathcal{H}} \|z\|_{\mathcal{Z}_P} + \|K_n - K\|_{\mathcal{G}, \mathcal{H}} \|y\|_{\mathcal{Z}_P})$$

$$375 \quad = \|P^{-1}\|_{\mathcal{Z}, \mathcal{Z}} (\|(K_n - K)H\|_{\mathcal{H}, \mathcal{H}} \|z\|_{\mathcal{Z}_P} + \|K_n - K\|_{\mathcal{G}, \mathcal{H}} \|y\|_{\mathcal{Z}_P})$$

$$376 \quad \leq \|P^{-1}\|_{\mathcal{Z}, \mathcal{Z}} (\|H\|_{\mathcal{H}, \mathcal{G}} \|z\|_{\mathcal{Z}_P} + \|y\|_{\mathcal{Z}_P}) \|K_n - K\|_{\mathcal{G}, \mathcal{H}}$$

$$377 \quad (3.27) \quad \leq \|P^{-1}\|_{\mathcal{Z}, \mathcal{Z}} (\|H\|_{\mathcal{H}, \mathcal{G}} \|z\|_{\mathcal{Z}_P} + \|y\|_{\mathcal{Z}_P}) \omega_n,$$

379 where the last inequality follows from Assumption 3.2. Altogether (3.25), (3.26), and (3.27) yield,  
380 for every  $n \in \mathbb{N}$ ,

$$381 \quad \|z_{n+1} - r_{n+1}\|_{\mathcal{Z}_P} \leq \|z_n - r_n\|_{\mathcal{Z}_P} + \Theta_n \|P^{-1}\|_{\mathcal{Z}, \mathcal{Z}} (\|H\|_{\mathcal{H}, \mathcal{G}} \|z_n\|_{\mathcal{Z}_P} + \|y\|_{\mathcal{Z}_P}) \omega_n$$

$$382 \quad \leq \|z_n - r_n\|_{\mathcal{Z}_P} + \Theta_n \|P^{-1}\|_{\mathcal{Z}, \mathcal{Z}} (\|H\|_{\mathcal{H}, \mathcal{G}} (\|z_n - r_n\|_{\mathcal{Z}_P} + \|r_n\|_{\mathcal{Z}_P}) + \|y\|_{\mathcal{Z}_P}) \omega_n$$

$$383 \quad (3.28) \quad \leq (1 + \mu_n) \|z_n - r_n\|_{\mathcal{Z}_P} + \nu_n$$

384 with

$$385 \quad (3.29) \quad \mu_n = \delta \|P^{-1}\|_{\mathcal{Z}, \mathcal{Z}} \|H\|_{\mathcal{H}, \mathcal{G}} \omega_n$$

$$386 \quad (3.30) \quad \nu_n = \delta \|P^{-1}\|_{\mathcal{Z}, \mathcal{Z}} (\varrho \|H\|_{\mathcal{H}, \mathcal{G}} + \|y\|_{\mathcal{Z}_P}) \omega_n.$$

388 Since  $(\mu_n) \in \ell_+^1$  and  $(\nu_n) \in \ell_+^1$ , according to [4, Lemma 5.31],  $\|z_n - r_n\|_{\mathcal{Z}_P} < +\infty$ . Consequently,  
 389  $\delta' = \sup_{n \in \mathbb{N}} \|z_n\|_{\mathcal{Z}_P} < \varrho + \sup_{n \in \mathbb{N}} \|z_n - r_n\|_{\mathcal{Z}_P} < +\infty$ .  
 390 Let us now define

$$391 \quad (3.31) \quad (\forall n \in \mathbb{N}) \quad e_n = \frac{T_n(z_n) - \bar{T}(z_n)}{\bar{\alpha}},$$

392 and it follows from (3.27) that

$$393 \quad (3.32) \quad \sum_{n \in \mathbb{N}} \|e_n\|_{\mathcal{Z}_P} \leq \frac{\|P^{-1}\|_{\mathcal{Z}, \mathcal{Z}}}{\bar{\alpha}} (\|H\|_{\mathcal{H}, \mathcal{G}} \delta' + \|y\|_{\mathcal{Z}_P}) \sum_{n \in \mathbb{N}} \omega_n < +\infty.$$

394 From (3.31), Algorithm (3.7) can be re-expressed as

$$395 \quad (3.33) \quad (\forall n \in \mathbb{N}) \quad z_{n+1} = z_n + \Theta'_n (\mathbf{W}(z_n) + e_n - z_n) \quad \text{with} \quad \Theta'_n = \bar{\alpha} \Theta_n \in ]0, 1[.$$

396 Therefore, the weak convergence of  $(z_n)_{n \in \mathbb{N}}$  to some  $\tilde{z} \in \text{Fix}(\mathbf{W}) = \text{Fix}(\bar{T}) = \bar{\mathcal{F}}$  follows from [4,  
 397 Theorem 5.5]. ■

398 We deduce the following more restrictive convergence result which is an extension of standard  
 399 convergence results for CV algorithm.

400 **Corollary 3.8.** *Assume that  $\text{ran}(L + L^*)$  is closed,  $\lambda_{\min} \geq 0$ , and  $\text{Ker}(L + L^*) = \text{Ker} L$ . Let  
 401  $(\tau, \sigma) \in ]0, +\infty[^2$  be such that  $\tau^{-1} - \sigma \|D\|_{\mathcal{H}, \mathcal{L}}^2 > \|M\|_{\mathcal{H}, \mathcal{H}}^2/4$ , where  $M$  is given by (3.9). For*

$$402 \quad (3.34) \quad \delta = 2 - \frac{1}{4} \left( \frac{1}{\tau} - \sigma \|D\|_{\mathcal{H}, \mathcal{L}}^2 \right)^{-1} \|M\|_{\mathcal{H}, \mathcal{H}}^2,$$

403 let  $\{\Theta_n\}_{n \in \mathbb{N}} \subset [0, \delta]$  be a sequence such that  $\sum_{n \in \mathbb{N}} \Theta_n (\delta - \Theta_n) = +\infty$ , and suppose that  $\bar{\mathcal{F}} \neq \emptyset$ . Then

404 the sequence  $((x_n, u_n))_{n \in \mathbb{N}}$  given by (3.4) converges weakly to some point in  $\bar{\mathcal{F}}$ .

405 *Proof.* Note that, if  $\tau^{-1} - \sigma \|D\|_{\mathcal{H}, \mathcal{L}}^2 > \|M\|_{\mathcal{H}, \mathcal{H}}^2/4$ , then  $\tau\sigma \|D\|_{\mathcal{H}, \mathcal{L}}^2 < 1$ . In addition, it follows  
 406 from Proposition 3.4 that, when  $\text{ran}(L + L^*)$  is closed,  $\lambda_{\min} \geq 0$ , and  $\text{Ker}(L + L^*) = \text{Ker} L$ ,  $P^{-1}\tilde{L}$   
 407 is  $\tilde{\eta}$ -cocoercive in  $\mathcal{Z}_P$  with

$$408 \quad (3.35) \quad \tilde{\eta} = \tau^{-1} (1 - \tau\sigma \|D\|_{\mathcal{H}, \mathcal{L}}^2) \eta_{\max}$$

409 and  $\eta_{\max} = 2/\|M\|_{\mathcal{H}, \mathcal{H}}^2$ . The result then follows by applying Theorem 3.7. ■

410 We now provide an estimate of the distance between a Kuhn-Tucker pair  $(\hat{x}, \hat{u})$  of the original  
 411 optimization problem and a fixed point  $(\tilde{x}, \tilde{u})$  of  $\bar{T}$ .

412 **Proposition 3.9.** *Assume that (2.3) holds and  $L$  is a cocoercive operator.*

413 *If  $f + g \circ D$  is strongly convex with modulus  $\rho > 0$ , then there exists a unique vector  $\tilde{x}$  in  $\bar{\mathcal{F}}_1$  and a  
 414 unique solution  $\hat{x}$  to the primal problem (2.1), and we have*

$$415 \quad (3.36) \quad \|\tilde{x} - \hat{x}\|_{\mathcal{H}} \leq \frac{1}{\rho} \|(\bar{K} - H^*) (H\hat{x} - y)\|_{\mathcal{H}}.$$

416 *In addition, if  $g$  is  $\beta$ -Lipschitz differentiable with  $\beta \in [0, +\infty[$ , there exists a unique  $(\tilde{x}, \tilde{u}) \in \bar{\mathcal{F}}$  and  
 417 a unique solution  $\hat{u}$  to the dual problem, and we have*

$$418 \quad (3.37) \quad \|\tilde{u} - \hat{u}\|_{\mathcal{L}} \leq \frac{\beta}{\rho} \|D\|_{\mathcal{H}, \mathcal{L}} \|(\bar{K} - H^*) (H\hat{x} - y)\|_{\mathcal{H}}.$$

419 *Proof.* Since (2.3) is satisfied,  $\partial f + D^* \circ \partial g \circ D = \partial(f + g \circ D)$  [4, Theorem 16.47 (i)]. As we  
 420 have assumed that  $f + g \circ D$  is  $\rho$ -strongly convex and  $L$  is cocoercive,  $L + \partial f + D^* \circ \partial g \circ D$  is  
 421 strongly monotone. It follows from Proposition 3.5(c) that there exists a single element  $\tilde{x}$  in  $\overline{\mathcal{F}}_1$   
 422 which is such that

$$423 \quad (3.38) \quad \overline{K}y \in L\tilde{x} + \partial f(\tilde{x}) + D^* \partial g(D\tilde{x}).$$

424 For any  $\gamma > 0$ , (3.38) is equivalent to

$$425 \quad (3.39) \quad \tilde{x} = \text{prox}_{\gamma(f+g \circ D)}(\tilde{x} - \gamma \overline{K}(H\tilde{x} - y)).$$

426 For similar reasons, there exists a unique solution  $\hat{x}$  to the primal problem, which satisfies the fixed  
 427 point equation

$$428 \quad (3.40) \quad \hat{x} = \text{prox}_{\gamma(f+g \circ D)}(\hat{x} - \gamma H^*(H\hat{x} - y)).$$

429 Because of the  $\rho$ -strong convexity of  $f + g \circ D$ ,  $\text{prox}_{\gamma(f+g \circ D)}$  is  $(1 + \gamma\rho)^{-1}$ -Lipschitzian [4, Proposition  
 430 23.13]. The error bound in (3.36) is thus derived by the same arguments as in the proof of [18,  
 431 Theorem 3.11].

432 In addition, if  $g$  is Gâteaux differentiable, there exists a unique  $\tilde{u} \in \mathcal{L}$  such that  $(\tilde{x}, \tilde{u}) \in \overline{\mathcal{F}}$ , which  
 433 is given by

$$434 \quad (3.41) \quad \tilde{u} = \nabla g(D\tilde{x}),$$

435 where  $\nabla g$  is the gradient of  $g$ . Similarly, there exists a unique solution  $\hat{u}$  to the dual problem, given  
 436 by

$$437 \quad (3.42) \quad \hat{u} = \nabla g(D\hat{x}).$$

438 By using the fact that  $\nabla g$  is  $\beta$ -Lipschitzian, we deduce that

$$439 \quad (3.43) \quad \begin{aligned} \|\tilde{u} - \hat{u}\|_{\mathcal{L}} &\leq \beta \|D(\tilde{x} - \hat{x})\|_{\mathcal{L}} \\ &\leq \beta \|D\|_{\mathcal{H}, \mathcal{L}} \|\tilde{x} - \hat{x}\|_{\mathcal{H}}. \end{aligned}$$

442 The upper error bound in (3.37) then follows from (3.36). ■

443 **3.3. Convergence properties of the mismatched CV algorithm with modified  $D$ .** We now  
 444 provide a brief analysis of the convergence of Condat-Vũ algorithm when  $D$  is replaced by an  
 445 operator  $V \in \mathcal{B}(\mathcal{H}, \mathcal{L})$  such that  $\|V - D\|_{\mathcal{H}, \mathcal{L}} \leq \epsilon$  for some  $\epsilon > 0$ . Let

$$446 \quad (1) \quad \tilde{\mathcal{A}} : \mathcal{Z} \rightarrow 2^{\mathcal{Z}} : z = (x, u) \mapsto (\partial f(x) + V^*u, -Vx + \partial g^*(u))$$

$$447 \quad (2) \quad \tilde{P} : \mathcal{Z} \rightarrow \mathcal{Z} : z = (x, u) \mapsto \left(\frac{x}{\tau} - V^*u, -Vx + \frac{u}{\sigma}\right)$$

448 with  $(\tau, \sigma) \in ]0, +\infty[^2$ .

449 Here again  $\tilde{\mathcal{A}}$  is maximally monotone on  $\mathcal{Z}$  and  $\tilde{P}$  is self adjoint and strongly positive for all  
 450  $(\tau, \sigma) \in ]0, +\infty[^2$  satisfying  $\tau\sigma(\|D\|_{\mathcal{H}, \mathcal{L}} + \epsilon)^2 < 1$ , hence  $\tilde{P}^{-1}\tilde{\mathcal{A}}$  is maximally monotone in  $\mathcal{Z}_{\mathcal{P}}$  [4,  
 451 Proposition 20.24]. Let the operators  $L, \tilde{L}, \tilde{K}, \tilde{K}_n$ , and  $\tilde{L}_n$  be defined by Notation 3.3 and suppose  
 452 that  $(K_n, \overline{K}) \in (\mathcal{B}(\mathcal{G}, \mathcal{H}))^2$  satisfy Assumption 3.2.

453 For every  $n \in \mathbb{N}$ , we define the operators

$$454 \quad (3.44) \quad \tilde{T} = J_{\tilde{P}^{-1}\tilde{\mathcal{A}}} \circ \left(\text{Id}_{\mathcal{Z}} - \tilde{P}^{-1}(\tilde{L} - \tilde{K})\right), \quad \tilde{T}_n = J_{\tilde{P}^{-1}\tilde{\mathcal{A}}} \circ \left(\text{Id}_{\mathcal{Z}} - \tilde{P}^{-1}(\tilde{L}_n - \tilde{K}_n)\right).$$

455 Then the mismatched form of CV algorithm can be rewritten as

$$456 \quad (3.45) \quad (\forall n \in \mathbb{N}) \quad z_{n+1} = z_n + \Theta_n \left(\tilde{T}_n(z_n) - z_n\right),$$

457 with  $z_0 \in \mathcal{Z}$ . Under conditions similar to those in Theorem 3.7, the weak convergence of sequence  
 458  $(z_n)_{n \in \mathbb{N}}$  to some  $\tilde{z} = (\tilde{x}, \tilde{u}) \in \text{Fix}(\tilde{T})$  can be readily established.

459 The following result provides bounds on the errors incurred in this additional change.

460 **Proposition 3.10.** *Assume that  $L$  is a cocoercive operator,  $f$  is strongly convex with modulus*  
 461  *$\rho > 0$ , and  $g$  is Lipschitz-differentiable with constant  $\beta > 0$ . Then, there exists a unique fixed point*  
 462  *$(\tilde{x}, \tilde{u})$  of  $\tilde{T}$  and unique solution  $(\hat{x}, \hat{u})$  to the primal-dual problem (2.5). In addition,*

$$463 \quad (3.46) \quad \|\tilde{x} - \hat{x}\|_{\mathcal{H}} \leq \frac{1}{\rho} \left\| (\bar{K} - H^*)(H\hat{x} - y) + V^*\nabla g(V\hat{x}) - D^*\nabla g(D\hat{x}) \right\|_{\mathcal{H}}$$

$$464 \quad (3.47) \quad \|\tilde{u} - \hat{u}\|_{\mathcal{L}} \leq \beta(\|V\|_{\mathcal{H},\mathcal{L}}\|\tilde{x} - \hat{x}\|_{\mathcal{H}} + \|(V - D)\hat{x}\|_{\mathcal{L}}).$$

466 *Proof.* Since  $\text{dom } g = \mathcal{L}$ , (2.3) holds. The existence and uniqueness of  $(\tilde{x}, \tilde{u})$  and  $(\hat{x}, \hat{u})$  follows  
 467 from arguments similar to those in the proof of Proposition 3.9.

468 In addition, it follows from (3.38) that, for every  $\gamma > 0$ ,

$$469 \quad (3.48) \quad \tilde{x} = \text{prox}_{\gamma f}(\tilde{x} - \gamma(\bar{K}(H\tilde{x} - y) + V^*\nabla g(V\tilde{x})))$$

470 Similarly,

$$471 \quad (3.49) \quad \hat{x} = \text{prox}_{\gamma f}(\hat{x} - \gamma(H^*(H\hat{x} - y) + D^*\nabla g(D\hat{x}))).$$

472 Since  $f$  is  $\rho$ -strongly convex,  $\text{prox}_{\gamma f}$  is Lipschitz-continuous with constant  $(1 + \gamma\rho)^{-1}$  [4, Proposition  
 473 23.13]. Thus, using the triangle inequality and the Lipschitz property of  $\text{prox}_{\gamma f}$  yields

$$474 \quad \|\tilde{x} - \hat{x}\|_{\mathcal{H}} \leq \frac{1}{1 + \gamma\rho} \left\| \tilde{x} - \hat{x} - \gamma(\bar{K}(H\tilde{x} - y) + V^*\nabla g(V\tilde{x}) - H^*(H\hat{x} - y) - D^*\nabla g(D\hat{x})) \right\|_{\mathcal{H}}$$

$$475 \quad \leq \frac{1}{1 + \gamma\rho} \left( \|(\text{Id}_{\mathcal{H}} - \gamma L)(\tilde{x} - \hat{x}) - \gamma(V^*\nabla g(V\tilde{x}) - V^*\nabla g(V\hat{x}))\|_{\mathcal{H}} \right.$$

$$476 \quad (3.50) \quad \left. + \gamma \left\| (\bar{K} - H^*)(H\hat{x} - y) + V^*\nabla g(V\hat{x}) - D^*\nabla g(D\hat{x}) \right\|_{\mathcal{H}} \right).$$

478 Since  $L$  is cocoercive with constant  $\eta_{\max}$  and  $\nabla g$  is  $\beta$ -Lipschitzian, hence cocoercive with constant  
 479  $1/\beta$ ,  $L + V^* \circ \nabla g \circ V$  is cocoercive with constant  $\zeta = ((\eta_{\max})^{-1} + \beta\|V\|_{\mathcal{H},\mathcal{L}}^2)^{-1}$  [4, Proposition 4.12].  
 480 Consequently, by choosing  $\gamma \in [0, 2\zeta]$ ,  $\text{Id}_{\mathcal{H}} - \gamma(L + V^* \circ \nabla g \circ V)$  is nonexpansive [4, Proposition  
 481 4.39]. We then deduce from (3.50) that

$$482 \quad (3.51) \quad \|\tilde{x} - \hat{x}\|_{\mathcal{H}} \leq \frac{1}{1 + \gamma\rho} \left( \|\tilde{x} - \hat{x}\|_{\mathcal{H}} + \gamma \left\| (\bar{K} - H^*)(H\hat{x} - y) + V^*\nabla g(V\hat{x}) - D^*\nabla g(D\hat{x}) \right\|_{\mathcal{H}} \right),$$

483 which results in (3.46).

484 According to the first-order characterization of  $\text{Fix}(\tilde{T})$ ,

$$485 \quad (3.52) \quad \tilde{u} = \nabla g(V\tilde{x}).$$

486 Using this equation and (3.42) allows us to derive the following inequality:

$$487 \quad \|\tilde{u} - \hat{u}\|_{\mathcal{L}} = \|\nabla g(V\tilde{x}) - \nabla g(D\hat{x})\|_{\mathcal{L}}$$

$$488 \quad \leq \beta\|V\tilde{x} - D\hat{x}\|_{\mathcal{L}}$$

$$489 \quad (3.53) \quad \leq \beta(\|V(\tilde{x} - \hat{x})\|_{\mathcal{L}} + \|(V - D)\hat{x}\|_{\mathcal{L}}),$$

491 so yielding (3.47). ■

492 **3.4. Remarks on the mismatched projected primal-dual splitting algorithm.** When an addi-  
 493 tional constraint is added to our initial primal problem (2.1), the optimization problem becomes

$$494 \quad (3.54) \quad \text{Find } \hat{x} \in C \cap \underset{x \in \mathcal{H}}{\text{Argmin}} \frac{1}{2} \|Hx - y\|_{\mathcal{G}}^2 + f(x) + g(Dx),$$

495 where  $C$  is a closed and convex nonempty subset of  $\mathcal{H}$ , and  $H$ ,  $D$ ,  $f$  and  $g$  are defined as previously.  
 496 The dual problem reads

$$497 \quad (3.55) \quad \text{Find } \hat{u} \in E \cap \underset{u \in \mathcal{L}}{\text{Argmin}} (f + h)^*(-D^*u) + g^*(u),$$

498 where  $E$  is a closed vector subspace of  $\mathcal{L}$  such that  $\text{ran } D \subset E$ . Such a problem can be solved  
 499 by the projected form of CV algorithm proposed by Briceño-Arias and López in [8]. We will be  
 500 interested in the mismatched form of this algorithm with a fixed operator  $\bar{K}$ : Given  $(x_0, u_0) \in \mathcal{Z}$   
 501 and  $(\tau, \sigma) \in ]0, +\infty[^2$ ,

502 **Projected mismatched primal-dual algorithm:**

$$503 \quad (3.56) \quad \text{for } n = 0, 1, \dots \quad \begin{cases} p_n = \text{prox}_{\tau f}(x_n - \tau(\bar{K}(Hx_n - y) + D^*u_n)) \\ x_{n+1} = \text{proj}_C(p_n) \\ \bar{x}_n = x_{n+1} + p_n - x_n \\ q_n = \text{prox}_{\sigma g^*}(u_n + \sigma D\bar{x}_n) \\ u_{n+1} = \text{proj}_E(q_n) \end{cases} .$$

505 In particular, if  $\bar{K} = H^*$ ,  $C = \mathcal{H}$ , and  $E = \mathcal{L}$ , we recover (3.3) in the case when, for every  $n \in \mathbb{N}$ ,  
 506  $\Theta_n = 1$ .

507 We then have the following convergence result where  $L$  and  $\bar{\mathcal{F}}$  are defined as previously.

508 **Proposition 3.11.** *Assume that  $\text{ran}(L + L^*)$  is closed,  $\lambda_{\min} \geq 0$ , and  $\text{Ker}(L + L^*) = \text{Ker } L$ . Let  
 509  $(\tau, \sigma) \in ]0, +\infty[^2$  be such that  $\tau^{-1} - \sigma \|D\|_{\mathcal{H}, \mathcal{L}}^2 > \|M\|_{\mathcal{H}, \mathcal{H}}^2/4$ , where  $M$  is given by (3.9). Suppose  
 510 that  $\bar{\mathcal{F}} \cap (C \times E) \neq \emptyset$ . Then the sequence  $((x_n, u_n))_{n \in \mathbb{N}}$  generated by (3.56) converges weakly to  
 511 some point in  $\bar{\mathcal{F}} \cap (C \times E)$ . In addition  $(p_n - x_n)_{n \in \mathbb{N}}$  and  $(q_n - u_n)_{n \in \mathbb{N}}$  converge strongly to 0.*

512 *Proof.* On the one hand, under the assumptions made on  $L$ , we have seen that  $x \mapsto \bar{K}(Hx - y)$   
 513 is  $\eta_{\max}$ -cocoercive with  $\eta_{\max} = 2/\|M\|_{\mathcal{H}, \mathcal{H}}^2$ . On the other hand, [8, Theorem 3.2 (ii)] guarantees the  
 514 weak convergence of  $(x_n, u_n)_{n \in \mathbb{N}}$  under the condition

$$515 \quad (3.57) \quad \|D\|_{\mathcal{H}, \mathcal{L}}^2 < \frac{1}{\sigma} \left( \frac{1}{\tau} - \frac{1}{2\eta_{\max}} \right).$$

516 The strong convergence properties of  $(p_n - x_n)_{n \in \mathbb{N}}$  and  $(q_n - u_n)_{n \in \mathbb{N}}$  are also stated in the proof of  
 517 [8, Theorem 3.2]. ■

518 Note that the error related to the mismatch can still be quantified by Proposition 3.9.

519 **4. The mismatched Loris-Verhoeven algorithm.** In this section, we consider a specific instance  
 520 of our original template (2.1) where  $f = \frac{\kappa}{2} \|\cdot\|_{\mathcal{H}}^2$  with  $\kappa \in ]0, +\infty[$ . This problem can be solved by  
 521 the primal-dual algorithm proposed by Loris and Verhoeven [40], which also relies on the forward-  
 522 backward splitting. This algorithm, which will be designated as LV algorithm, also appears under  
 523 the name of *Primal-Dual Fixed-Point algorithm based on the Proximity Operator* (PDFP2O) [16]  
 524 and *Proximal Alternating Predictor-Corrector* (PAPC) algorithm [29].

525 The LV iterations can still be described through the implicit inclusion (3.1) where  $\mathcal{A}$ ,  $\mathcal{B}$ , and  $P$  are

526 now given by

$$527 \quad (4.1) \quad (\forall z = (x, u) \in \mathcal{Z}) \quad \mathcal{A}z = \begin{pmatrix} D^*u \\ -Dx + \partial g^*(u) \end{pmatrix}$$

$$528 \quad (4.2) \quad \mathcal{B}z = \begin{pmatrix} (H^*H + \kappa \text{Id}_{\mathcal{H}})x - H^*y \\ 0 \end{pmatrix}$$

$$529 \quad (4.3) \quad \mathcal{P}z = \begin{pmatrix} \frac{1}{\tau}x \\ (\frac{1}{\sigma} \text{Id}_{\mathcal{L}} - \tau DD^*)u \end{pmatrix}$$

531 with  $(\tau, \sigma) \in ]0, +\infty[^2$ . The iterations of LV algorithm are then given by

532 **LV iterations:**

$$533 \quad (4.4) \quad \text{for } n = 0, 1, \dots \quad \begin{cases} t_n = H^*(Hx_n - y) + \kappa x_n \\ u_{n+\frac{1}{2}} = \text{prox}_{\sigma g^*} \left( u_n + \sigma D \left( x_n - \tau(t_n + D^*u_n) \right) \right) \\ x_{n+1} = x_n - \Theta_n \tau \left( t_n + D^*u_{n+\frac{1}{2}} \right) \\ u_{n+1} = u_n + \Theta_n \left( u_{n+\frac{1}{2}} - u_n \right) \end{cases}$$

535 where  $\{\Theta_n\}_{n \in \mathbb{N}}$  is a sequence of relaxation parameters and  $(x_0, u_0) \in \mathcal{Z}$ .

536 When, at iteration  $n \in \mathbb{N}$ ,  $H^*$  is replaced by an operator  $K_n \in \mathcal{B}(\mathcal{H}, \mathcal{G})$  satisfying [Assumption 3.2](#),  
537 the mismatched form of LV algorithm reads

538 **Mismatched LV iterations:**

$$539 \quad (4.5) \quad \text{for } n = 0, 1, \dots \quad \begin{cases} t_n = K_n(Hx_n - y) + \kappa x_n \\ u_{n+\frac{1}{2}} = \text{prox}_{\sigma g^*} \left( u_n + \sigma D \left( x_n - \tau(t_n + D^*u_n) \right) \right) \\ x_{n+1} = x_n - \Theta_n \tau \left( t_n + D^*u_{n+\frac{1}{2}} \right) \\ u_{n+1} = u_n + \Theta_n \left( u_{n+\frac{1}{2}} - u_n \right) \end{cases}$$

541 This iteration can be reexpressed as [\(3.7\)](#) where [Notation 3.3\(ii\)-\(iv\)](#) holds, but  $L$  is now defined as

$$542 \quad (4.6) \quad L = \overline{K}H + \kappa \text{Id}_{\mathcal{H}}.$$

543 It follows that all the results in [subsection 3.2](#) can be extended to the mismatched LV algorithm.

544 [Proposition 4.1](#) is a straightforward adaptation of [Proposition 3.5](#) to characterize the fixed point  
545 set of the nonlinear mapping  $\overline{T}$  (see [Notation 3.3\(iv\)](#)).

546 **Proposition 4.1.** *Let  $(\tilde{x}, \tilde{u}) \in \mathcal{Z}$ . Then  $(\tilde{x}, \tilde{u}) \in \text{Fix}(\overline{T})$  if and only if  $(\tilde{x}, \tilde{u})$  belongs to*

$$547 \quad (4.7) \quad \overline{\mathcal{F}} = \{(x, u) \in \mathcal{Z} \mid \overline{K}y \in Lx + D^*u, u \in \partial g(Dx)\},$$

548 *which is nonempty if  $L + D^* \circ \partial g \circ D$  is surjective.*

549 (i) *If  $\lambda_{\min} \geq 0$ , then  $\overline{\mathcal{F}}$  is closed and convex.*

550 (ii) *Let  $\overline{\mathcal{F}}_1 = \{x \in \mathcal{H} \mid (\exists u \in \mathcal{L}) (x, u) \in \overline{\mathcal{F}}\}$ .*

551  *$\overline{\mathcal{F}}_1$  has at most one element if one of the following conditions holds:*

552 (a)  *$L + D^* \circ \partial g \circ D$  is strictly monotone.*

553 (b)  *$\lambda_{\min} \geq 0$  and  $g \circ D$  is strictly convex.*

554  *$\overline{\mathcal{F}}_1$  is a singleton if [\(2.3\)](#) is satisfied and one of the following conditions holds:*

555 (c)  *$\lambda_{\min} \geq 0$  and  $L + D^* \circ \partial g \circ D$  is strongly monotone.*

556 (d)  *$\lambda_{\min} > 0$ .*



- 557 (e)  $\lambda_{\min} \geq 0$ ,  $g$  is strongly convex, and  $D^*D$  is strongly positive.  
 558 (f)  $L$  is cocoercive,  $g$  is strongly convex, and  $D^*D$  is strongly positive.

559 Similarly, we derive an equivalent of [Corollary 3.8](#) concerning the convergence of the mismatched  
 560 LV.

561 **Proposition 4.2.** *Assume that  $\text{ran}(L + L^*)$  is closed,  $\lambda_{\min} \geq 0$ , and  $\text{Ker}(L + L^*) = \text{Ker} L$ . Let  
 562  $(\tau, \sigma) \in ]0, +\infty[^2$  be such that  $\tau < 4/\|M\|_{\mathcal{H},\mathcal{H}}^2$  and  $\tau\sigma\|D\|_{\mathcal{H},\mathcal{L}}^2 < 1$ , where  $M$  is given by [\(3.9\)](#). For*

$$563 \quad (4.8) \quad \delta = 2 - \frac{\tau}{4}\|M\|_{\mathcal{H},\mathcal{H}}^2,$$

564 let  $\{\Theta_n\}_{n \in \mathbb{N}} \subset [0, \delta]$  be a sequence such that  $\sum_{n \in \mathbb{N}} \Theta_n (\delta - \Theta_n) = +\infty$ , and suppose that  $\overline{\mathcal{F}} \neq \emptyset$ . Then  
 565 the sequence  $((x_n, u_n))_{n \in \mathbb{N}}$  given by [\(4.5\)](#) converges weakly to some point in  $\overline{\mathcal{F}}$ .

566 *Proof.* The result from [Proposition 3.4](#) stating that, when  $\text{ran}(L + L^*)$  is closed,  $\lambda_{\min} \geq 0$ , and  
 567  $\text{Ker}(L + L^*) = \text{Ker} L$ ,  $P^{-1}\tilde{L}$  is cocoercive in  $\mathcal{Z}_P$  with constant  $\tilde{\eta}_{\max} = 2/\|P^{-1/2}\Pi M\|_{\mathcal{H},\mathcal{Z}}^2$  still holds  
 568 for LV algorithm. We also have

$$569 \quad (4.9) \quad \tilde{\eta}_{\max} \geq \frac{2}{\|P^{-1/2}\Pi\|_{\mathcal{H},\mathcal{Z}}^2\|M\|_{\mathcal{H},\mathcal{H}}^2} = \frac{2}{\tau\|M\|_{\mathcal{H},\mathcal{H}}^2} = \tilde{\eta},$$

570 which shows that  $P^{-1}\tilde{L}$  is  $\tilde{\eta}$ -cocoercive in  $\mathcal{Z}_P$ . The result then follows from [Theorem 3.7](#), which  
 571 remains valid in this context. ■

572 **Remark 4.3.** When there is no mismatch,  $M = \sqrt{2}(H^*H + \kappa \text{Id}_{\mathcal{H}})^{1/2}$  and we recover the condi-  
 573 tions derived in [\[40, Theorem 3.1\]](#), [\[26, Theorem 3.1\]](#) for the convergence of sequences  $(x_n)_{n \in \mathbb{N}}$  and  
 574  $(u_n)_{n \in \mathbb{N}}$  generated by [\(4.4\)](#).

575 Similarly to [Proposition 3.9](#), we can provide an estimate of the distance between a Kuhn-Tucker  
 576 pair  $(\hat{x}, \hat{u})$  of the original problem and a fixed point  $(\tilde{x}, \tilde{u})$  of  $\overline{T}$ .

577 **Proposition 4.4.** *Assume that  $0 \in \text{sri}(\text{ran}(D) - \text{dom}(g))$  and  $L$ , given in [\(4.6\)](#), is a cocoercive  
 578 operator.*

579 *Let  $\rho \in ]0, +\infty[$ . If  $\kappa \geq \rho$  or  $g \circ D$  is strongly convex with modulus  $\rho$ , then there exists a unique  
 580 vector  $\tilde{x}$  in  $\overline{\mathcal{F}}_1$ , defined in [Proposition 4.1\(ii\)](#), and a unique solution  $\hat{x}$  to the primal problem [\(2.1\)](#).  
 581 Moreover, inequality [\(3.36\)](#) holds.*

582 *In addition, if  $g$  is  $\beta$ -Lipschitz differentiable with  $\beta \in [0, +\infty[$ , there exists a unique  $(\tilde{x}, \tilde{u}) \in \overline{\mathcal{F}}$ ,  
 583 defined in [\(4.7\)](#), and a unique solution  $\hat{u}$  to the dual problem. Finally, inequality [\(3.37\)](#) is satisfied.*

## 584 **5. The mismatched Combettes - Pesquet algorithm.**

585 **5.1. CP algorithm.** In this section, we explore the benefits of the Combettes - Pesquet algorithm  
 586 (CP), which relies on Tseng's splitting [\(2.6\)](#) to solve [\(2.1\)](#)-[\(2.4\)](#). This algorithm was introduced in  
 587 [\[22, Theorem 4.2\]](#) and constitutes a generalization of the one in [\[9\]](#). It reads

588 **CP-iterations for [\(2.1\)](#) and [\(2.4\)](#):**

$$589 \quad (5.1) \quad \text{for } n = 0, 1, \dots \quad \left\{ \begin{array}{l} v_{1,n} = x_n - \gamma(H^*(Hx_n - y) + D^*u_n) \\ p_{1,n} = \text{prox}_{\gamma f}(v_{1,n}) \\ v_{2,n} = u_n + \gamma Dp_{1,n} \\ p_{2,n} = \text{prox}_{\gamma g^*}(v_{2,n}) \\ q_{2,n} = p_{2,n} + \gamma Dp_{1,n} \\ q_{1,n} = p_{1,n} - \gamma(H^*(Hp_{1,n} - y) + D^*p_{2,n}) \\ x_{n+1} = x_n - v_{1,n} + q_{1,n} \\ u_{n+1} = u_n - v_{2,n} + q_{2,n}, \end{array} \right.$$

590

591 where  $\gamma > 0$  and  $(x_0, u_0) \in \mathcal{Z}$ . By setting, for every  $n \in \mathbb{N}$ ,

$$592 \quad (5.2) \quad z_n = (x_n, u_n), \quad v_n = (v_{1,n}, v_{2,n}), \quad p_n = (p_{1,n}, p_{2,n}), \quad \text{and} \quad q_n = (q_{1,n}, q_{2,n}),$$

593 iterations (5.1) can be rewritten as

594 **Tseng iterations:**

$$595 \quad (5.3) \quad \text{for } n = 0, 1, \dots \quad \left\{ \begin{array}{l} v_n = z_n - \gamma \mathbf{Q}(z_n) \\ p_n = J_{\gamma \mathcal{M}}(v_n) \\ q_n = p_n - \gamma \mathbf{Q}(p_n) \\ z_{n+1} = z_n - v_n + q_n, \end{array} \right.$$

596 with

$$598 \quad (5.4) \quad \mathcal{M}: \mathcal{Z} \rightarrow 2^{\mathcal{Z}}: (x, u) \mapsto (\partial f(x), \partial g^*(u))$$

$$599 \quad (5.5) \quad \mathbf{Q}: \mathcal{Z} \rightarrow \mathcal{Z}: (x, u) \mapsto (H^*Hx + D^*u - H^*y, -Dx).$$

601 Since (5.3) is an instance of Tseng's algorithm, according to [9, Theorem 2.5 (ii)], if  $\text{zer}(\mathcal{M} +$   
 602  $\mathbf{Q}) \neq \emptyset$  and  $\gamma \in ]0, 1/\vartheta[$ , sequences  $(z_n)_{n \in \mathbb{N}} = ((x_n, u_n))_{n \in \mathbb{N}}$  and  $(p_n)_{n \in \mathbb{N}}$  converge weakly to some  
 603  $\hat{z} = (\hat{x}, \hat{u}) \in \text{zer}(\mathcal{M} + \mathbf{Q})$ . Therefore, sequences  $(x_n)_{n \in \mathbb{N}}$  (resp.  $(u_n)_{n \in \mathbb{N}}$ ) generated by (5.1) converge  
 604 weakly to  $\hat{x}$  (resp.  $\hat{u}$ ), which is a solution to (2.1) (resp. (2.4)).

605 **5.2. Convergence properties of the twice mismatched CP algorithm.** For algorithm (5.1), we  
 606 consider a mismatched form obtained by substituting a fixed operator  $\bar{K} \in \mathcal{B}(\mathcal{G}, \mathcal{H})$  for  $H^*$  as well  
 607 as an operator  $V^* \in \mathcal{B}(\mathcal{L}, \mathcal{H})$  for  $D^*$ . This leads to the following iterations:

608 **Mismatched CP-iterations:**

$$609 \quad (5.6) \quad \text{for } n = 0, 1, \dots \quad \left\{ \begin{array}{l} v_{1,n} = x_n - \gamma (\bar{K}(Hx_n - y) + V^*u_n) \\ p_{1,n} = \text{prox}_{\gamma f}(v_{1,n}) \\ v_{2,n} = u_n + \gamma(Dx_n + \varepsilon u_n) \\ p_{2,n} = \text{prox}_{\gamma g^*}(v_{2,n}) \\ q_{2,n} = p_{2,n} + \gamma(Dp_{1,n} + \varepsilon p_{2,n}) \\ q_{1,n} = p_{1,n} - \gamma (\bar{K}(Hp_{1,n} - y) + V^*p_{2,n}) \\ x_{n+1} = x_n - v_{1,n} + q_{1,n} \\ u_{n+1} = u_n - v_{2,n} + q_{2,n}, \end{array} \right.$$

611 where  $\varepsilon > 0$  is an additional parameter. In addition to the presence of mismatched adjoints, we see  
 612 that there are two algorithmic modifications in the updates rules of variables  $v_{2,n}$  and  $q_{2,n}$ .  
 613 By making the change of variables (5.2), algorithm (5.6) can be rewritten, in the product space  $\mathcal{Z}$ ,  
 614 as

615 **Tseng form of (5.6):**

$$616 \quad (5.7) \quad \text{for } n = 0, 1, \dots \quad \left\{ \begin{array}{l} v_n = z_n - \gamma \tilde{\mathbf{Q}}(z_n) \\ p_n = J_{\gamma \mathcal{M}}(v_n) \\ q_n = p_n - \gamma \tilde{\mathbf{Q}}(p_n) \\ z_{n+1} = z_n - v_n + q_n, \end{array} \right.$$

617 where

$$619 \quad (5.8) \quad \tilde{\mathbf{Q}}: \mathcal{Z} \rightarrow \mathcal{Z}: (x, u) \mapsto (Lx + V^*u - \bar{K}y, -Dx + \varepsilon u)$$

620 and

$$621 \quad (5.9) \quad L = \overline{K}H,$$

622 as in [section 3](#).

623 We first provide some preliminary results about operator  $\tilde{\mathbf{Q}}$ .

624 **Proposition 5.1.** *Let  $\lambda_{\min}$  defined in [Notation 3.3](#) and*

$$625 \quad (5.10) \quad \varepsilon \lambda_{\min} = \lambda_{\min} - \frac{1}{4\varepsilon} \|V - D\|_{\mathcal{H}, \mathcal{L}}^2,$$

$$626 \quad (5.11) \quad \varepsilon \vartheta_1 = \max\{\|L\|_{\mathcal{H}, \mathcal{H}}, \varepsilon\} + \max\{\|D\|_{\mathcal{H}, \mathcal{L}}, \|V\|_{\mathcal{H}, \mathcal{L}}\},$$

$$627 \quad (5.12) \quad \varepsilon \vartheta_2 = \sqrt{\|L\|_{\mathcal{H}, \mathcal{H}}^2 + \|V\|_{\mathcal{H}, \mathcal{L}}^2 + \|D\|_{\mathcal{H}, \mathcal{L}}^2 + \varepsilon^2},$$

629 and  $\varepsilon \vartheta = \min\{\varepsilon \vartheta_1, \varepsilon \vartheta_2\}$ . We have the following properties:

630 (i)  $\tilde{\mathbf{Q}}$  is Lipschitz continuous with constant  $\varepsilon \vartheta$ .

631 (ii) If  $\varepsilon \lambda_{\min} \geq 0$ , then  $\tilde{\mathbf{Q}}$  is monotone.

632 (iii) If  $\varepsilon \lambda_{\min} > 0$ , then  $\tilde{\mathbf{Q}}$  is strongly monotone and cocoercive.

633 *Proof.* Since  $\tilde{\mathbf{Q}}$  is an affine operator, its monotonicity, Lipschitz continuity, and cocoercivity  
634 properties are the same as those of the linear operator

$$635 \quad (5.13) \quad \overline{\mathbf{Q}} = \tilde{\mathbf{Q}} + (\overline{K}y, 0).$$

636 (i) On the one hand

$$637 \quad \|\overline{\mathbf{Q}}\|_{\mathcal{Z}, \mathcal{Z}} \leq \left\| \begin{bmatrix} L & 0 \\ 0 & \varepsilon \text{Id}_{\mathcal{L}} \end{bmatrix} \right\|_{\mathcal{Z}, \mathcal{Z}} + \left\| \begin{bmatrix} 0 & V^* \\ -D & 0 \end{bmatrix} \right\|_{\mathcal{Z}, \mathcal{Z}}$$

$$638 \quad (5.14) \quad \leq \max\{\|L\|_{\mathcal{H}, \mathcal{H}}, \varepsilon\} + \max\{\|D\|_{\mathcal{H}, \mathcal{L}}, \|V\|_{\mathcal{H}, \mathcal{L}}\} = \varepsilon \vartheta_1.$$

640 On the other hand, for every  $z = (x, u) \in \mathcal{Z}$ ,

$$641 \quad \|\overline{\mathbf{Q}}z\|_{\mathcal{Z}}^2 = \|Lx + V^*u\|_{\mathcal{H}}^2 + \|-Dx + \varepsilon u\|_{\mathcal{L}}^2$$

$$642 \quad \leq \|LL^* + V^*V\|_{\mathcal{H}, \mathcal{H}} \|z\|_{\mathcal{Z}}^2 + \|DD^* + \varepsilon^2 \text{Id}_{\mathcal{L}}\|_{\mathcal{L}, \mathcal{L}} \|z\|_{\mathcal{Z}}^2$$

$$643 \quad (5.15) \quad \leq (\|L\|_{\mathcal{H}, \mathcal{H}}^2 + \|V\|_{\mathcal{H}, \mathcal{L}}^2 + \|D\|_{\mathcal{H}, \mathcal{L}}^2 + \varepsilon^2) \|z\|_{\mathcal{Z}}^2,$$

645 which implies that

$$646 \quad (5.16) \quad \|\overline{\mathbf{Q}}\|_{\mathcal{Z}, \mathcal{Z}} \leq \varepsilon \vartheta_2.$$

647 In summary,  $\overline{\mathbf{Q}}$ , and thus  $\tilde{\mathbf{Q}}$ , are Lipschitz continuous with constant  $\|\overline{\mathbf{Q}}\|_{\mathcal{Z}, \mathcal{Z}} \leq \varepsilon \vartheta$ .

648 (ii) By using Cauchy-Schwarz inequality, for every  $z = (x, u) \in \mathcal{Z}$ ,

$$649 \quad \langle \overline{\mathbf{Q}}z, z \rangle_{\mathcal{Z}} = \langle Lx, x \rangle_{\mathcal{H}} + \varepsilon \|u\|_{\mathcal{L}}^2 + \langle u, (V - D)x \rangle_{\mathcal{L}}$$

$$650 \quad \geq \lambda_{\min} \|x\|_{\mathcal{H}}^2 + \varepsilon \|u\|_{\mathcal{L}}^2 - \|u\|_{\mathcal{L}} \|(V - D)x\|_{\mathcal{L}}$$

$$651 \quad \geq \lambda_{\min} \|x\|_{\mathcal{H}}^2 + \varepsilon \|u\|_{\mathcal{L}}^2 - \|V - D\|_{\mathcal{H}, \mathcal{L}} \|u\|_{\mathcal{L}} \|x\|_{\mathcal{H}}$$

$$652 \quad (5.17) \quad = [\|x\|_{\mathcal{H}} \quad \|u\|_{\mathcal{L}}] C \begin{bmatrix} \|x\|_{\mathcal{H}} \\ \|u\|_{\mathcal{L}} \end{bmatrix},$$

654 where

$$655 \quad (5.18) \quad C = \begin{bmatrix} \lambda_{\min} & -\frac{1}{2} \|V - D\|_{\mathcal{H}, \mathcal{L}} \\ -\frac{1}{2} \|V - D\|_{\mathcal{H}, \mathcal{L}} & \varepsilon \end{bmatrix}.$$

656  $C$  is positive semidefinite if and only if

$$657 \quad (5.19) \quad \begin{cases} \operatorname{tr}(C) = \lambda_{\min} + \varepsilon \geq 0 \\ \det(C) = \lambda_{\min}\varepsilon - \frac{1}{4}\|V - D\|_{\mathcal{H},\mathcal{L}}^2 \geq 0, \end{cases}$$

658 that is  ${}^\varepsilon\lambda_{\min} \geq 0$ . We deduce from (5.17) that, subject to this condition,  $\overline{\mathbf{Q}}$  and thus  $\tilde{\mathbf{Q}}$  are  
659 monotone.

660 (iii) Assume now that  ${}^\varepsilon\lambda_{\min} > 0$ . Then  $C$  is positive definite and its smallest eigenvalue is

$$661 \quad (5.20) \quad {}^\varepsilon v = \frac{\lambda_{\min} + \varepsilon - \sqrt{(\lambda_{\min} + \varepsilon)^2 - 4{}^\varepsilon\lambda_{\min}\varepsilon}}{2} > 0.$$

662 It follows from (5.17) that  $\overline{\mathbf{Q}}$  is strongly monotone with constant  ${}^\varepsilon v$ .  
663 We have then, for every  $z \in \mathcal{Z}$ ,

$$664 \quad \langle \overline{\mathbf{Q}}z, z \rangle_{\mathcal{Z}} \geq {}^\varepsilon v \|z\|_{\mathcal{Z}}^2 \geq {}^\varepsilon v \frac{\|\overline{\mathbf{Q}}z\|_{\mathcal{Z}}^2}{\|\overline{\mathbf{Q}}\|_{\mathcal{Z},\mathcal{Z}}^2} \geq \frac{{}^\varepsilon v}{({}^\varepsilon\vartheta)^2} \|\overline{\mathbf{Q}}z\|_{\mathcal{Z}}^2.$$

665 This shows that  $\overline{\mathbf{Q}}$  (and thus  $\tilde{\mathbf{Q}}$ ) is cocoercive with constant

$$666 \quad (5.21) \quad {}^\varepsilon\eta = \frac{{}^\varepsilon v}{({}^\varepsilon\vartheta)^2}. \quad \blacksquare$$

667 Conditions of convergence for the mismatched CP algorithm are deduced from this result.

668 **Proposition 5.2.** *Let  ${}^\varepsilon\lambda_{\min}$  and  ${}^\varepsilon\vartheta$  be defined as in Proposition 5.1. Let  $\gamma \in ]0, ({}^\varepsilon\vartheta)^{-1}[$ . Assume  
669 that  $\operatorname{zer}(\mathcal{M} + \tilde{\mathbf{Q}}) \neq \emptyset$  and  ${}^\varepsilon\lambda_{\min} \geq 0$ . Then the sequences  $((x_n, u_n))_{n \in \mathbb{N}}$  and  $((p_{1,n}, p_{2,n}))_{n \in \mathbb{N}}$   
670 generated by Algorithm (5.6) converge weakly to  $(\tilde{x}, \tilde{u}) \in \operatorname{zer}(\mathcal{M} + \tilde{\mathbf{Q}})$ . In addition, if  ${}^\varepsilon\lambda_{\min} > 0$ ,  
671 then  $((x_n, u_n))_{n \in \mathbb{N}}$  and  $((p_{1,n}, p_{2,n}))_{n \in \mathbb{N}}$  converge strongly to the unique zero of  $\mathcal{M} + \tilde{\mathbf{Q}}$ .*

672 *Proof.* Under the considered assumptions,  $\tilde{\mathbf{Q}}$  is monotone and  ${}^\varepsilon\vartheta$ -Lipschitzian. Thus, the result  
673 follows from standard conditions for the convergence of Tseng's algorithm [4, Theorem 26.17] applied  
674 to (5.7). Then, if  ${}^\varepsilon\lambda_{\min} > 0$ ,  $\tilde{\mathbf{Q}}$  is strongly monotone and the strong convergence property follows  
675 from [4, Theorem 26.17 (iii)].  $\blacksquare$

676 We now characterize the set of limit points.

677 **Proposition 5.3.** *Let  ${}^\varepsilon g$  be the Moreau envelope of  $g$  of parameter  $\varepsilon > 0$  defined as*

$$678 \quad (5.22) \quad (\forall v \in \mathcal{L}) \quad {}^\varepsilon g(v) = \inf_{w \in \mathcal{L}} g(w) + \frac{1}{2\varepsilon} \|w - v\|_{\mathcal{L}}^2.$$

679 *Let  $(\tilde{x}, \tilde{u}) \in \mathcal{Z}$ . Then  $(\tilde{x}, \tilde{u}) \in \operatorname{zer}(\mathcal{M} + \tilde{\mathbf{Q}})$  if and only if  $(\tilde{x}, \tilde{u})$  belongs to*

$$680 \quad (5.23) \quad {}^\varepsilon\overline{\mathcal{F}} = \left\{ (x, u) \in \mathcal{Z} \mid \overline{K}y \in \partial f(x) + Lx + V^*u, u = \nabla {}^\varepsilon g(Dx) = \frac{Dx - \operatorname{prox}_{\varepsilon g}(Dx)}{\varepsilon} \right\},$$

681 *which is nonempty if  $L + \partial f + V^* \circ \nabla {}^\varepsilon g \circ D$  is surjective.*

682 (i) *If  ${}^\varepsilon\lambda_{\min}$  defined by (5.10) is nonnegative, then  ${}^\varepsilon\overline{\mathcal{F}}$  is closed and convex.*

683 (ii) *Let  ${}^\varepsilon\overline{\mathcal{F}}_1 = \{x \in \mathcal{H} \mid (x, \nabla {}^\varepsilon g(Dx)) \in {}^\varepsilon\overline{\mathcal{F}}\}$  and let*

$$684 \quad (5.24) \quad {}^\varepsilon\lambda_{1,\min} = \lambda_{\min} - \frac{1}{\varepsilon} \|V - D\|_{\mathcal{H},\mathcal{L}} \|D\|_{\mathcal{H},\mathcal{L}} \geq 0.$$

685  *${}^\varepsilon\overline{\mathcal{F}}_1$  has at most one element if one of the following conditions holds:*

- 686 (a)  $L + \partial f + V^* \circ \nabla^\varepsilon g \circ D$  is strictly monotone.  
 687 (b)  ${}^\varepsilon\lambda_{1,\min} \geq 0$  and  ${}^\varepsilon g \circ D + f$  is strictly convex.  
 688  $\overline{\mathcal{F}}_1$  is a singleton if one of the following conditions holds:  
 689 (c)  $L + \partial f + V^* \circ \nabla^\varepsilon g \circ D$  is strongly monotone.  
 690 (d)  ${}^\varepsilon\lambda_{1,\min} > 0$   
 691 (e)  ${}^\varepsilon\lambda_{1,\min} \geq 0$ , and  $f$  is strongly convex or  $[g^*$  is Lipschitz-differentiable and  $D^*D$  is  
 692 strongly positive].

693 *Proof.* The proof follows the same lines as in the proof of [Proposition 3.5](#). In the following, we  
 694 point out the main differences.

695 By using [\(3.9\)](#) and [\(5.8\)](#), we have

$$696 \quad (\tilde{x}, \tilde{u}) \in \text{zer}(\mathcal{M} + \tilde{\mathbf{Q}})$$

$$697 \quad (5.25) \quad \Leftrightarrow \begin{cases} 0 \in \partial f(\tilde{x}) + L\tilde{x} + V^*\tilde{u} - \overline{K}y \\ 0 \in \partial g^*(\tilde{u}) - D\tilde{x} + \varepsilon\tilde{u}. \end{cases}$$

698

699 We know that  $({}^\varepsilon g)^* = g^* + \frac{\varepsilon}{2}\|\cdot\|^2 \Rightarrow \partial({}^\varepsilon g)^*(\tilde{u}) = \partial g^*(\tilde{u}) + \varepsilon\tilde{u}$  [[4](#), Proposition 14.1] and  ${}^\varepsilon g$  is  
 700 differentiable with gradient  $\nabla^\varepsilon g = \varepsilon^{-1}(\text{Id}_{\mathcal{L}} - \text{prox}_{\varepsilon g})$  [[4](#), Proposition 12.30]. We thus deduce that

$$701 \quad (\tilde{x}, \tilde{u}) \in \text{zer}(\mathcal{M} + \tilde{\mathbf{Q}})$$

$$702 \quad \Leftrightarrow \begin{cases} \overline{K}y \in \partial f(\tilde{x}) + L\tilde{x} + V^*\tilde{u} \\ D\tilde{x} \in \partial({}^\varepsilon g)^*(\tilde{u}) \end{cases}$$

$$703 \quad (5.26) \quad \Leftrightarrow \begin{cases} \overline{K}y \in \partial f(\tilde{x}) + L\tilde{x} + V^*\tilde{u} \\ \tilde{u} = \nabla^\varepsilon g(D\tilde{x}). \end{cases}$$

704

705 This shows that  $(\tilde{x}, \tilde{u}) \in {}^\varepsilon\overline{\mathcal{F}}$ .

- 706 (i) According to [Proposition 5.1\(i\)-\(ii\)](#), if  ${}^\varepsilon\lambda_{\min} \geq 0$ ,  $\tilde{\mathbf{Q}}$  is monotone and continuous. Since  $\mathcal{M}$   
 707 is maximally monotone,  $\mathcal{M} + \tilde{\mathbf{Q}}$  is also maximally monotone and  $\text{zer}(\mathcal{M} + \tilde{\mathbf{Q}})$  is closed and  
 708 convex.  
 709 (ii) We can perform the decomposition

$$710 \quad (5.27) \quad L + \partial f + V^* \circ \nabla^\varepsilon g \circ D = \partial f + D^* \circ \nabla^\varepsilon g \circ D + L + (V - D)^* \circ \nabla^\varepsilon g \circ D.$$

711 Subdifferential  $\partial f + D^* \circ \nabla^\varepsilon g \circ D = \partial(f + {}^\varepsilon g \circ D)$  is maximally monotone. Using the  
 712 Cauchy-Schwarz inequality, for every  $(x, x') \in \mathcal{H}^2$ ,

$$713 \quad \langle L(x - x'), x - x' \rangle_{\mathcal{H}} + \langle (V - D)^* \nabla^\varepsilon g(Dx) - (V - D)^* \nabla^\varepsilon g(Dx'), x - x' \rangle_{\mathcal{H}}$$

$$714 \quad \geq \lambda_{\min} \|x - x'\|_{\mathcal{H}}^2 - \|(V - D)^*(\nabla^\varepsilon g(Dx) - \nabla^\varepsilon g(Dx'))\|_{\mathcal{H}} \|x - x'\|_{\mathcal{H}}$$

$$715 \quad \geq \lambda_{\min} \|x - x'\|_{\mathcal{H}}^2 - \|V - D\|_{\mathcal{H}, \mathcal{L}} \|\nabla^\varepsilon g(Dx) - \nabla^\varepsilon g(Dx')\|_{\mathcal{L}} \|x - x'\|_{\mathcal{H}}$$

$$716 \quad \geq \lambda_{\min} \|x - x'\|_{\mathcal{H}}^2 - \frac{1}{\varepsilon} \|V - D\|_{\mathcal{H}, \mathcal{L}} \|D(x - x')\|_{\mathcal{L}} \|x - x'\|_{\mathcal{H}}$$

$$717 \quad (5.28) \quad \geq {}^\varepsilon\lambda_{1,\min} \|x - x'\|_{\mathcal{H}}^2,$$

718

719 where we have used the  $\varepsilon^{-1}$ -Lipschitz continuity of  $\nabla^\varepsilon g$ . This shows that  $L + (V - D)^* \circ$   
 720  $\nabla^\varepsilon g \circ D$  is monotone, if  ${}^\varepsilon\lambda_{1,\min} \geq 0$ . In addition, it is strongly monotone (hence, strictly  
 721 monotone) if  ${}^\varepsilon\lambda_{1,\min} > 0$ . Since this operator is also continuous, it is maximally monotone  
 722 in both cases. The rest of the proof is similar to that of [Proposition 3.5](#), by noticing that  
 723  ${}^\varepsilon g$  is strongly convex  $\Leftrightarrow ({}^\varepsilon g)^*$  is Lipschitz-differentiable  $\Leftrightarrow g^*$  is Lipschitz-differentiable. ■

724 We now provide a bound on the distance between an optimal pair of solutions  $(\hat{x}, \hat{u})$  to problem  
725 (2.1)-(2.4) and  $(\tilde{x}, \tilde{u}) \in \varepsilon \overline{\mathcal{F}}$ .

726 **Proposition 5.4.** *Let  $\varepsilon \lambda_{\min}$  be defined by (5.10). Assume that  $\varepsilon \lambda_{\min} > 0$ ,  $f$  is strongly convex  
727 with modulus  $\rho > 0$ , and  $g$  is Lipschitz-differentiable with constant  $\beta > 0$ . Then, there exists a  
728 unique pair  $\tilde{z} = (\tilde{x}, \tilde{u}) \in \varepsilon \overline{\mathcal{F}}$  and a unique solution  $(\hat{x}, \hat{u})$  to the primal-dual problem. In addition,*

$$729 \quad (5.29) \quad \sqrt{\|\tilde{x} - \hat{x}\|_{\mathcal{H}}^2 + \|\tilde{u} - \hat{u}\|_{\mathcal{L}}^2} \leq \frac{1}{\mu} \left( \|(\overline{K} - H^*)(H\hat{x} - y)\|_{\mathcal{H}} + \sqrt{\|V - D\|_{\mathcal{H}, \mathcal{L}}^2 + \varepsilon^2 \|\hat{u}\|_{\mathcal{L}}} \right),$$

730 where  $\mu = \min\{\rho, 1/\beta\}$ .

731 *Proof.*  $f$  is  $\rho$ -strongly convex and  $g$  is  $\beta$ -Lipschitz differentiable (i.e.,  $g^*$  is  $\beta^{-1}$ -strongly convex),  
732  $\partial f$  and  $\partial g^*$  are strongly monotone with constants  $\rho$  and  $1/\beta$ , respectively.  $\mathcal{M}$  is thus strongly  
733 monotone with constant  $\mu$ . Since  $\mathbf{Q}$  is continuous and monotone,  $\mathcal{M} + \mathbf{Q}$  is maximally monotone  
734 and strongly monotone. The existence of a unique zero  $\hat{z}$  to  $\mathcal{M} + \mathbf{Q}$  is thus guaranteed by [4,  
735 Corollary 23.37]. Similarly, it follows from Proposition 5.1(i) and Proposition 5.1(iii) that  $\tilde{\mathbf{Q}}$  is  
736 continuous and strongly monotone. Hence  $\mathcal{M} + \tilde{\mathbf{Q}}$  is maximally monotone and strongly monotone  
737 and  $\text{zer}(\mathcal{M} + \tilde{\mathbf{Q}})$  is a singleton  $\{\tilde{z}\}$ .

738 For every  $\gamma > 0$ ,

$$739 \quad (5.30) \quad \hat{z} \in \text{zer}(\mathcal{M} + \mathbf{Q}) \Leftrightarrow \hat{z} = J_{\gamma, \mathcal{M}}(\hat{z} - \gamma \mathbf{Q} \hat{z})$$

$$740 \quad (5.31) \quad \tilde{z} \in \text{zer}(\mathcal{M} + \tilde{\mathbf{Q}}) \Leftrightarrow \tilde{z} = J_{\gamma, \mathcal{M}}(\tilde{z} - \gamma \tilde{\mathbf{Q}} \tilde{z})$$

742 Since  $\mathcal{M}$  is strongly monotone with constant  $\mu$ ,  $J_{\gamma, \mathcal{M}}$  is Lipschitz continuous with constant  $1/(1+\gamma\mu)$   
743 [4, Proposition 23.13]. From (5.30) and (5.31), we deduce that

$$744 \quad \|\tilde{z} - \hat{z}\|_{\mathcal{Z}} \leq \frac{1}{1+\gamma\mu} \|\tilde{z} - \hat{z} - \gamma(\tilde{\mathbf{Q}}\tilde{z} - \mathbf{Q}\hat{z})\|_{\mathcal{Z}} \\ 745 \quad \leq \frac{1}{1+\gamma\mu} \|(\text{Id}_{\mathcal{Z}} - \gamma\tilde{\mathbf{Q}})(\tilde{z} - \hat{z}) - \gamma(\tilde{\mathbf{Q}} - \mathbf{Q})\hat{z}\|_{\mathcal{Z}} \\ 746 \quad (5.32) \quad \leq \frac{1}{1+\gamma\mu} \left( \|\text{Id}_{\mathcal{Z}} - \gamma\tilde{\mathbf{Q}}\|_{\mathcal{Z}, \mathcal{Z}} \|\tilde{z} - \hat{z}\|_{\mathcal{Z}} + \gamma \|(\tilde{\mathbf{Q}} - \mathbf{Q})\hat{z}\|_{\mathcal{Z}} \right). \\ 747$$

748 According to Proposition 5.1(iii),  $\tilde{\mathbf{Q}}$  is cocoercive with constant  $\varepsilon \eta$  given by (5.21). Therefore, by  
749 assuming that  $\gamma \in ]0, \varepsilon \eta]$ , we have  $\|\text{Id}_{\mathcal{Z}} - \gamma\tilde{\mathbf{Q}}\|_{\mathcal{Z}, \mathcal{Z}} \leq 1$ . We deduce from (5.32) that

$$750 \quad \|\tilde{z} - \hat{z}\|_{\mathcal{Z}} \leq \frac{1}{\mu} \|(\tilde{\mathbf{Q}} - \mathbf{Q})\hat{z}\|_{\mathcal{Z}} \\ 751 \quad = \frac{1}{\mu} \|((\overline{K} - H^*)(H\hat{x} - y) + (V - D)^*\hat{u}, \varepsilon\hat{u})\|_{\mathcal{Z}} \\ 752 \quad \leq \frac{1}{\mu} \left( \|((\overline{K} - H^*)(H\hat{x} - y), 0)\|_{\mathcal{Z}} + \|((V - D)^*\hat{u}, \varepsilon\hat{u})\|_{\mathcal{Z}} \right) \\ 753 \quad (5.33) \quad \leq \frac{1}{\mu} \left( \|(\overline{K} - H^*)(H\hat{x} - y)\|_{\mathcal{H}} + \sqrt{\|V - D\|_{\mathcal{H}, \mathcal{L}}^2 + \varepsilon^2 \|\hat{u}\|_{\mathcal{L}}} \right). \quad \blacksquare \\ 754$$

755 **Remark 5.5.**

756 (i) In the absence of mismatch on  $D^*$  (i.e.,  $V^* = D^*$ ), one can choose  $\varepsilon = 0$  in (5.6) and (5.8).  
757  $\text{zer}(\mathcal{M} + \tilde{\mathbf{Q}})$  is then equal to the set  $\overline{\mathcal{F}}$  characterized in Proposition 3.5. If  $\overline{\mathcal{F}} \neq \emptyset$ ,  $\lambda_{\min} \geq$   
758  $0$ , and  $\gamma \in ]0, (\theta^0)^{-1}[$ , then the sequences  $((x_n, u_n))_{n \in \mathbb{N}}$  and  $((p_{1,n}, p_{2,n}))_{n \in \mathbb{N}}$  generated by  
759 Algorithm (5.6) converge weakly to  $(\tilde{x}, \tilde{u}) \in \overline{\mathcal{F}}$ . Proposition 3.9 applies to evaluate the error  
760 incurred by the mismatch.

761 (ii) When  $H = \overline{K} = 0$ , it follows from [39, Theorem 1.2] that, if  $f$  is strongly convex with  
 762 modulus  $\rho > 0$ ,  $(\tilde{x}, \tilde{u}) \in \varepsilon \overline{\mathcal{F}}$  and  $\hat{x}$  is the solution to the primal problem, then

$$763 \quad (5.34) \quad \|\tilde{x} - \hat{x}\|_{\mathcal{H}} \leq \frac{1}{\rho} \|V - D\|_{\mathcal{H}, \mathcal{L}} \|\tilde{u}\|_{\mathcal{L}}.$$

764 **6. Numerical experiments.** This section illustrates the applicability of our theoretical results to  
 765 the resolution of 2D image reconstruction problems arising in Computed Tomography (CT). All the  
 766 simulations presented in this section are performed using the ASTRA Toolbox [55, 54] implemented  
 767 in Matlab.

768 **6.1. Example 1: reconstruction from few CT views.** We aim at recovering an image  $\bar{x}$  with  $N$   
 769 pixels, reshaped as a vector in the Euclidean space  $\mathcal{H} = \mathbb{R}^N$ . A set of noisy tomographic projections  
 770  $p \in \mathcal{G} = \mathbb{R}^S$  of the original image  $\bar{x}$  is available, according to the following observation model:

$$771 \quad (6.1) \quad p = R\bar{x} + b$$

772 where  $R \in \mathbb{R}^{S \times N}$  is the forward operator called projector, and  $b$  is an additive i.i.d. zero-mean  
 773 Gaussian noise. The projector is a discrete implementation of the Radon transform, chosen here  
 774 as the line-length ray-driven projector [57]. An approximate adjoint of  $R$ , denoted by  $B \in \mathbb{R}^{N \times S}$ ,  
 775 is derived from an alternative discretization model. In our simulations,  $B$  is the pixel-driven back-  
 776 projector which is particularly suited for a GPU implementation compared to the adjoint of  $R$   
 777 [32].

778 For  $(u, v)$  i.i.d. uniformly sampled on  $([0, 1]^N)^2$ , the average over 20 realizations of the ratio  
 779  $\langle Ru \mid v \rangle / \langle u \mid Bv \rangle$  is 1.005.

780 Our goal is to retrieve an estimate of  $\bar{x}$  given  $p$ , the projector  $R$ , and its surrogate adjoint  $B$ .

781 **6.1.1. Data.** Let us specify the data and settings used for this experiment. In model (6.1),  $\bar{x}$  is  
 782 an axial slice of an abdomen of size 41 cm whose values range from 1000 to 3000, containing intense  
 783 inserts with pixel intensity ranges in [3500, 4200]. These values correspond to positive Hounsfield  
 784 Unit (HU) (air is 0 HU and water is 1000 HU). The source-to-object distance is 800 mm, and the  
 785 source-to-image distance is 1200 mm, which corresponds to a magnification factor of 1.5, typically  
 786 encountered in clinical scanners. The projector  $R$  describes a fan-beam geometry over  $180^\circ$  using 50  
 787 regularly spaced angular steps. The detector has a length of 40 cm. The bin grid is twice upsampled  
 788 with respect to the pixel grid: the detector has 250 bins of size 1.6 mm, so that  $S = 250 \times 250$ .  
 789 The image is reconstructed on a grid of  $N = 160 \times 160$  pixels, with size  $2 \times 1.6/1.5 = 2.13$  mm.  
 790 Data  $p$  is obtained after adding 1% relative Gaussian noise on  $R\bar{x}$ . The inverse problem (6.1) is  
 791 highly ill-posed because of the small detector field of view (FOV) and the limited angular density.  
 792 This problem corresponds to a common setup in CT when the detector is not large enough to  
 793 measure the projections of large body parts such as the abdomen, as it happens in image guidance  
 794 for interventional radiology. The set of pixels in the image whose projections belong to the detector  
 795 FOV defines the so-called FOV image. Because of truncation, the exterior of the FOV has to be  
 796 estimated for the reconstruction of the FOV to be accurate. The size of the reconstruction grid  
 797 is thus slightly larger than the support of the FOV. Figure 1 shows the original image  $\bar{x}$  with its  
 798 FOV and the observed data  $p$ . Due to the use of a short detector, the projections suffer from axial  
 799 truncation, as shown by non-zero values at the border of some vertical lines of the sinogram.

800 **6.1.2. Regularization.** We provide an estimate of  $\bar{x}$  by solving the following penalized least  
 801 squares problem:

$$802 \quad (6.2) \quad \underset{x \in \mathbb{R}^N}{\text{minimize}} \quad \frac{1}{2} \|p - Rx\|^2 + f(x) + g(Dx) + \frac{\kappa}{2} \|x\|^2$$

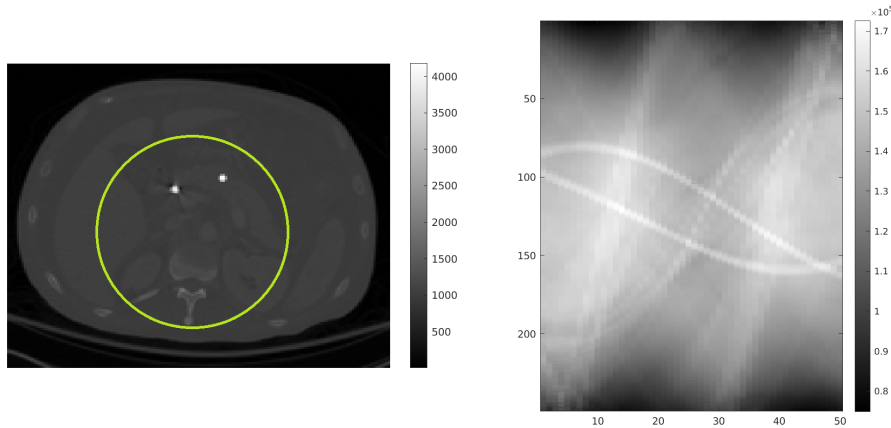


Figure 1: Original image  $\bar{x}$  with highlighted FOV (left) and observed projections  $p$  (right).

803 with  $\kappa \in [0, +\infty[$ . Functions  $f$  and  $g \circ D$  promote desired properties on the image. Here, we  
 804 promote sparsity of the image vertical and horizontal gradients [13]. We additionally constraint  
 805 the nonnegativity of the reconstructed pixel intensities. This leads us to set  $f = \iota_{[0, +\infty[^N}$  where  
 806  $\iota_C$  denotes the indicator function of a set  $C$ . Moreover, we set  $g = \xi \|\cdot\|_{1,2}$  with  $D = \nabla = \begin{bmatrix} \nabla^h \\ \nabla^v \end{bmatrix}$ ,  
 807 where  $\nabla^h \in \mathbb{R}^{N \times N}$ ,  $\nabla^v \in \mathbb{R}^{N \times N}$  are, respectively, the horizontal and vertical discrete gradient  
 808 operators (assuming zero-padding) and  $\|\cdot\|_{1,2}$  is the  $\ell_{1,2}$ -norm of  $\mathbb{R}^N$ , so that  $g \circ D$  is the discrete  
 809 total variation penalty weighted by  $\xi \in [0, +\infty[$  [48]. We set the regularization hyperparameter  $\xi$   
 810 to 800 through a grid search, minimizing the FOV image error.

811 **6.1.3. Condat-Vũ algorithm.** Problem (6.2) can be rewritten as (2.1) with  $H = \begin{bmatrix} R \\ \sqrt{\kappa} \text{Id}_{\mathbb{R}^N} \end{bmatrix}$ ,

812  $\kappa \in [0, +\infty[$ , and  $y = \begin{bmatrix} p \\ 0 \end{bmatrix}$ . The surrogate adjoint of  $H$  is  $\bar{K} = [B \quad \sqrt{\kappa} \text{Id}_{\mathbb{R}^N}]$ . For such a problem,

813 we can apply the CV approach presented in section 3.

814 We run Algorithms (3.3) and (3.4) (i.e., CV algorithm without/with an adjoint mismatch, respec-  
 815 tively) for  $\kappa \in \{\kappa_1, \kappa_2\} \subset ]0, +\infty[$  where  $L = \bar{K}H = BR + \kappa \text{Id}_{\mathbb{R}^N}$  is only cocoercive for  $\kappa = \kappa_2$ . In the  
 816 latter case, the condition given in Proposition 3.5(d) holds, which proves the existence of a unique  
 817 fixed point  $\tilde{x}$  of scheme (3.4) and its convergence is ensured according to Corollary 3.8. In contrast,  
 818 nothing can be said about the convergence of the scheme in the case involving  $\kappa_1$ . We set  $\kappa_1 = 1$   
 819 and  $\kappa_2$  to  $-\tilde{\lambda}_{\min} + 0.01$  where  $\tilde{\lambda}_{\min} = -28.24$  is the minimum spectral value of  $(BR + R^*B^*)/2$   
 820 estimated from the power iterative method. The cocoercivity constant  $\eta$  is computed using Propo-  
 821 sition 3.4(ii).

822 The convergence parameter  $\sigma$  is set to  $10^{-3}$ . For Algorithm (3.3), the step-size  $\tau$  is set to  $0.99/(8\sigma +$   
 823  $0.5/\theta)$  with  $\theta = 1/\|H\|_{\mathcal{H}, \mathcal{G}}^2$ . To illustrate the instabilities incurred by the use of the mismatched  
 824 adjoint  $\bar{K}$  when using  $\kappa_1$ , the same step-size value  $\tau$  is used as in the matched case. With  $\kappa_2$ , the  
 825 convergence of Algorithm (3.4) is ensured, as stated by Corollary 3.8 and Proposition 3.4, by setting  
 826  $\tau = 0.99/(8\sigma + \|M\|_{\mathcal{H}, \mathcal{H}}^2/4)$ , where  $M$  is defined in (3.9). Both algorithms are run until a maximum  
 827 number  $3 \times 10^4$  of iterations is reached. Initial iterates  $x_0$  and  $u_0$  are set to zero.

828 **6.1.4. Loris-Verhoeven algorithm.** Problem (6.2) can also be solved with the LV approach  
 829 presented in section 4. More precisely, the cost function can be rewritten as in (2.1) by setting



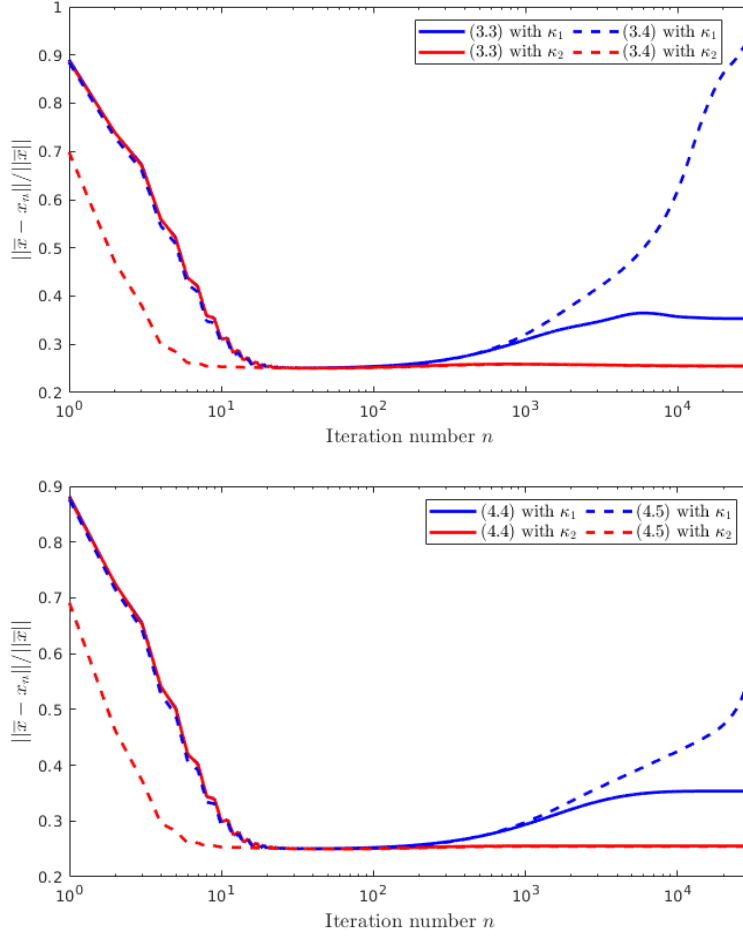


Figure 2: Evolution of the error  $(\|\bar{x} - x_n\|/\|\bar{x}\|)_n$  along iterations for Algorithms (3.3)-(3.4) (top) and Algorithms (4.4)-(4.5) (bottom), for two settings of parameter  $\kappa$ .

$$830 \quad H = R, \quad y = p, \quad D = \begin{bmatrix} \nabla \\ \text{Id}_{\mathbb{R}^N} \end{bmatrix}, \quad f = \frac{\kappa}{2} \|\cdot\|^2, \quad \text{and}$$

$$831 \quad (6.3) \quad (\forall (z_1, z_2) \in (\mathbb{R}^N)^2) \quad g \left( \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \right) = \|z_1\|_{1,2} + \iota_{[0,+\infty[^N}(z_2).$$

832 Similarly to the CV case,  $L = BR + \kappa \text{Id}_{\mathbb{R}^N}$ . Therefore, for the mismatched LV algorithm (4.5),  
 833 the existence and uniqueness of a fixed point  $\tilde{x}$  in  $\bar{\mathcal{F}}_1$ , defined in Proposition 4.1(ii)(c), is only  
 834 guaranteed when  $\kappa = \kappa_2$ , but not when  $\kappa = \kappa_1$ . The convergence parameter  $\sigma$  is here set as  
 835  $1.99/(9\tau)$ . For Algorithm (4.4) with both values of  $\kappa$  and Algorithm (4.5) with  $\kappa_1$ , step-size  $\tau$  is  
 836 set to  $1.99/(\|H\|_{\mathcal{H},\mathcal{G}}^2 + \kappa)$ . The convergence of Algorithm (4.5) with  $\kappa_2$  is ensured by setting  $\tau$  to  
 837  $3.99/\|M\|_{\mathcal{H},\mathcal{H}}^2$ , where  $M$  is the same as in the CV case, in accordance with Proposition 4.2.

838 **6.1.5. Results.** Figure 2 displays the normalized root mean square error (NMSE) defined as  
 839  $(\|\bar{x} - x_n\|/\|\bar{x}\|)_n$ , computed along the iterations when applying CV Algorithms (3.3)-(3.4) and LV  
 840 Algorithms (4.4)-(4.5). We recall that Algorithms (3.3) and (4.4) require the use of the exact adjoint  
 841 of  $H$ . The plots confirm that, with value  $\kappa_1$ , both CV and LV algorithms converge when this exact  
 842 adjoint  $H^*$  is used, but diverge when  $H^*$  is replaced by  $\bar{K}$ , as was expected from our theoretical

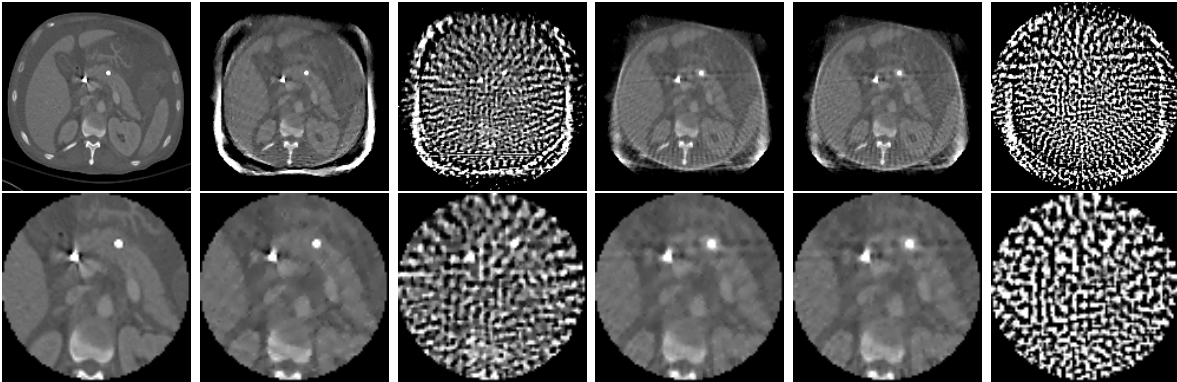


Figure 3: Reconstructed images (top) and zoomed FOVs (bottom). From left to right:  $\bar{x}$ , reconstructions obtained using (3.3) with  $\kappa_1$ , (3.4) with  $\kappa_1$ , (3.3) with  $\kappa_2$ , (3.4) with  $\kappa_2$ , (4.5) with  $\kappa_1$ .

843 analysis.

844 In the latter case, CV and LV algorithms show an initial convergence trend before diverging. We  
 845 notice that on this example, the mismatched CV (3.4) diverges faster than the mismatched LV (4.5).  
 846 When reaching the maximal number of iterations, the NMSE associated to (3.4) is 0.93 whereas  
 847 the NMSE associated to (4.5) is 0.57.

848 When using  $\kappa_2$ , all algorithms converge to close fixed points, as expected by our theory. The  
 849 corresponding NMSE values are 0.251 for (3.3), 0.242 for (3.4), 0.253 for (4.4) and 0.252 for (4.5).  
 850 Remarkably, our mismatched algorithms (3.4)-(4.5) with  $\kappa_2$  lead to lower reconstruction error in  
 851 the first iterations of the algorithms.

852 Reconstructed images and their FOVs are displayed in Figure 3 using the same windowing. Note  
 853 that the reconstructions obtained using (4.4) look the same as those obtained with (3.3). Likewise,  
 854 the same reconstruction is obtained with (3.4) and (4.5), when  $\kappa_2$  is used. For all the reconstructed  
 855 images, we also provide the NMSE and the maximum absolute error (MAE) computed in the FOV  
 856 image, defined as  $\max_{i \in \{1, \dots, N\}} |[\text{mask}_{\text{FOV}}(\bar{x} - x)]_i|$  where  $[\text{mask}_{\text{FOV}}(x_n)]_i = [x_n]_i$  if pixel  $i$  of  $x_n$  is in  
 857 the FOV, and  $[\text{mask}_{\text{FOV}}(x_n)]_i = 0$  otherwise. When parameter  $\kappa_2$  is used, the reconstructed image  
 858 obtained by CV/LV with the mismatched adjoint  $\bar{K}$  (MAE=461, NMSE=0.040) is very similar  
 859 to the image obtained without mismatch (MAE=495, NMSE=0.041). In contrast, combining the  
 860 setting  $\kappa_1$  with the mismatched adjoint yields reconstructions that are highly deteriorated by high-  
 861 frequency patterns leading to a higher error (MAE=2735, NMSE=0.637 for CV and MAE=1249,  
 862 NMSE=0.249 for LV) compared to the solution obtained when using the exact adjoint (MAE=215,  
 863 NMSE=0.026 for both CV and LV).

864 **6.2. Example 2: reconstruction from Poisson data.** In this second example, we focus on an  
 865 another acquisition scenario, namely we consider the case when the projection data  $p$  contain photon  
 866 count views which follow a Poisson distribution:

$$867 \quad (6.4) \quad p = \mathcal{P}(R\bar{x}),$$

868 that each component  $p_s$  of  $p \in \mathbb{R}^S$ ,  $s \in \{1, \dots, S\}$ , is drawn independently from a Poisson distribu-  
 869 tion with mean  $[R\bar{x}]_s$ . The goal is again to restore an estimate of  $\bar{x}$  given  $p$ ,  $R$ , and its mismatched  
 870 adjoint  $B$ .

871 **6.2.1. Data.** This second test problem uses a fan-beam geometry with 200 views. In model  
 872 (6.4),  $\bar{x}$  is an axial slice of an abdomen. Contrary to Example 1, it is now made only of an anatomical

873 background (see [Figure 4](#)). We set the source-to-object distance, the source-to-image distance, the  
 874 detector length, the number of bins, and the bin size as in Example 1. Hence,  $S = 250 \times 200$ .  
 875 Projections  $p$  are then simulated by using model (6.4). The image is reconstructed on a discrete  
 876 grid of  $N = 220 \times 220$  pixels, with size 2.13 mm. The mismatched backprojector  $B$  is derived from  
 877 the same discretization scheme as  $B$  in Example 1.

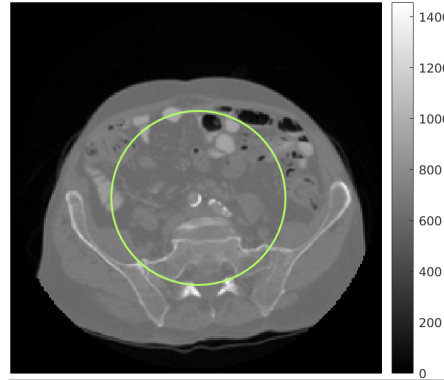


Figure 4: Phantom  $\bar{x}$  with highlighted FOV

878 **6.2.2. Regularization.** Due to the presence of Poisson noise, the data discrepancy term in the  
 879 cost function differs from the one in Example 1. Namely, we make use of the negative log-likelihood  
 880 of the image given the data [37], to define  $\ell \circ R$  with

$$881 \quad (6.5) \quad (\forall z = (z_s)_{1 \leq s \leq S} \in \mathbb{R}^S) \quad \ell(z) = \sum_{s=1}^S \mathcal{KL}(z_s, p_s)$$

882 and

$$883 \quad (6.6) \quad (\forall (u, v) \in \mathbb{R}^2) \quad \mathcal{KL}(u, v) = \begin{cases} -v \log u + u & \text{if } u > 0, v > 0 \\ u & \text{if } u \geq 0, v = 0 \\ +\infty & \text{otherwise.} \end{cases}$$

884 Furthermore, we introduce a nonnegativity constraint on the components of the solution and a  
 885 Tikhonov-based penalty. Altogether, the resulting minimization problem reads

$$886 \quad (6.7) \quad \underset{x \in \mathbb{R}^N}{\text{minimize}} \quad \chi \ell(Rx) + \frac{1}{2} \|Hx\|^2 + \iota_{[0, +\infty[^N}(x)$$

887 where  $\chi \in [0, +\infty[$  weights the Poisson data fidelity term and we define the linear operator  $H =$   
 888  $\begin{bmatrix} \Delta \\ \kappa \text{Id}_{\mathbb{R}^N} \end{bmatrix}$ . Moreover,  $\Delta \in \mathbb{R}^{N \times N}$  refers to the 2D discrete Laplacian operator (here, implemented in  
 889 the 2D Fourier domain), and  $\kappa \in [0, +\infty[$ .

890 Problem (6.7) can be rewritten in the form (2.1) with  $y = 0_{\mathbb{R}^{2N}}$ ,  $D = R$ ,  $g = \chi \ell$ , and  $f = \iota_{[0, +\infty[^N}$ .

891 **6.2.3. Combettes-Pesquet algorithm.** Problem (6.7) is solved with CP algorithm (5.1)-(5.6).  
 892 In our configuration, an adjoint mismatch only arises on  $D$  (i.e., the projector  $R$ ), and again denoting  
 893 by  $B$  the mismatched adjoint. We hence have  $V = B^*$  and  $\bar{K} = H^*$  in Algorithm (5.6).

894 To guarantee the convergence of Algorithm (5.6) to a unique fixed point, we must choose  $(\kappa, \varepsilon)$   
 895 to satisfy the conditions of [Proposition 5.1](#) and [Proposition 5.3](#). We first set  $\kappa$  and then choose  
 896  $4\varepsilon = \|V - D\|_{\mathcal{H}, \mathcal{L}}^2 / (\kappa^2 - 0.02)$  so that  $\varepsilon \lambda_{\min} = 0.02 > 0$  in (5.10). In particular, we consider the

897 setting  $(\kappa_2, \varepsilon_2) = (3.3, 6.6)$  for which convergence is guaranteed and the setting  $(\kappa_1, \varepsilon_1) = (0.6, 0)$   
 898 for which it is not guaranteed.  
 899 In Algorithm (5.1), parameter  $\gamma$  is set to  $0.99/(4 + \kappa^2 + \|D\|_{\mathcal{H}, \mathcal{L}})$ . When Algorithm (5.6) is run  
 900 with  $\kappa_1$ ,  $\gamma$  is set as in the matched case. When Algorithm (5.6) is run with  $\kappa_2$  and  $\varepsilon_2$ , we ensure its  
 901 convergence using Proposition 5.2 by setting  $\gamma = 0.99/\varepsilon\vartheta$  with  $\varepsilon\vartheta = \min(\varepsilon\vartheta_1, \varepsilon\vartheta_2) = \varepsilon\vartheta_1$  where  
 902  $\varepsilon\vartheta_1$  and  $\varepsilon\vartheta_2$  are defined respectively by (5.11) and (5.12). We set the data fidelity parameter  $\chi$  to  
 903 5000, and we perform 4000 iterations of CP algorithm. Similarly to Example 1, the initial iterates  
 904  $x_0$  and  $u_0$  are set to zero.

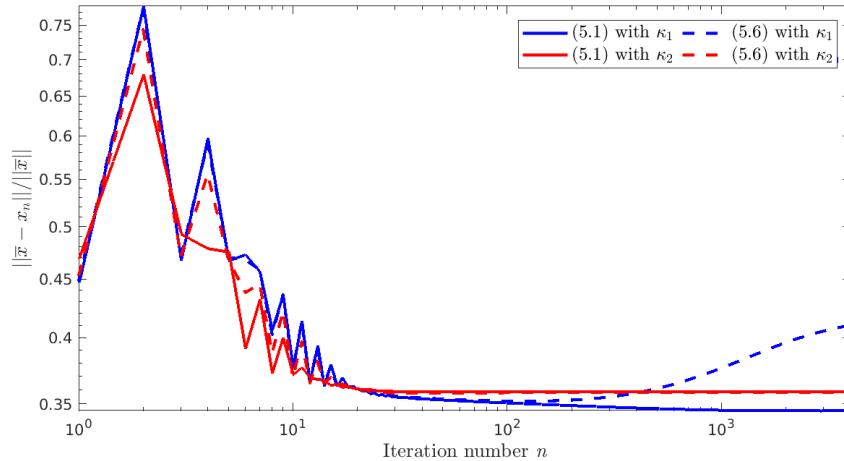


Figure 5: Evolution of the error  $\|\bar{x} - x_n\|/\|\bar{x}\|$  along iterations for CP Algorithms (5.1) and (5.6), for two settings of parameter  $\kappa$ .

905 **6.2.4. Results.** In Figure 5, we plot the relative error between the ground truth  $\bar{x}$  and the  
 906 estimate along the iterations. We observe the same behavior as in our previous example. Algorithm  
 907 (5.6) with  $(\kappa_1, \varepsilon_1)$  diverges quickly, while it converges to a fixed point with  $(\kappa_2, \varepsilon_2)$ . This fixed point  
 908 is indistinguishable from the minimizer of (6.7) with  $\kappa = \kappa_2$ , both in terms of NMSE/MAE and  
 909 visual inspection (see Figure 6).

910 **7. Conclusion.** We thoroughly analyzed a set of classical primal-dual proximal splitting algo-  
 911 rithms when the adjoints of the involved linear operators are modified. Our study strongly relies  
 912 on fixed-point theory tools. We established necessary conditions to ensure the convergence of these  
 913 unmatched schemes when applied to non-smooth convex penalized least-squares problems. We  
 914 showed that unmatched algorithms solve more general monotone inclusions than those derived from  
 915 the original minimization problem. We illustrated our results through two numerical examples of  
 916 CT image reconstruction when the adjoint of the projector is replaced by a surrogate backpro-  
 917 jector. A quadratic and a more sophisticated Poisson fidelity term have been considered in our  
 918 experiments. In both cases, we showed that convergence can still be guaranteed for an unmatched  
 919 projector-backprojector pair. It would be interesting to extend our analysis to algorithms for finding  
 920 a zero of a sum of more than two maximally monotone operators [21]. Handling more general form  
 921 of cost functions would also be of further interest in future research. Other interesting research  
 922 avenues would be the study of accelerated versions of mismatched primal-dual algorithms presented  
 923 in this paper and possibly the derivation of convergence rates.

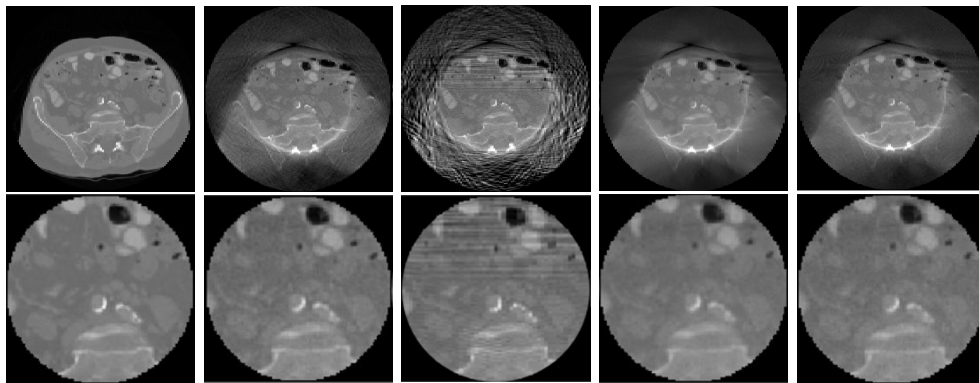


Figure 6: Reconstructed images (top) and zoomed FOVs (bottom). From left to right:  $\bar{x}$ , reconstructed images using (5.1) with  $\kappa_1$  (NMSE = 0.052, MAE = 201), (5.6) with  $\kappa_1$  (NMSE = 0.079, MAE = 401), (5.1) with  $\kappa_2$  (NMSE = 0.056, MAE = 232), (5.6) with  $\kappa_2$  (NMSE = 0.055, MAE = 206).

- 925 [1] A. ALOTAIBI, P. L. COMBETTES, AND N. SHAHZAD, *Solving coupled composite monotone inclusions by successive*  
 926 *Fejér approximations of their Kuhn–Tucker set*, SIAM Journal on Optimization, 24 (2014), pp. 2076–2095,  
 927 <https://doi.org/10.1137/130950616>.
- 928 [2] S. ARRIDGE, P. MAASS, O. ÖKTEM, AND C. SCHÖNLIEB, *Solving inverse problems using data-driven models*,  
 929 *Acta Numerica*, 28 (2019), pp. 1 – 174.
- 930 [3] S. BANERT, J. RUDZUSIKA, O. ÖKTEM, AND J. ADLER, *Accelerated forward-backward optimization using deep*  
 931 *learning*, 2021, <https://doi.org/10.48550/arxiv.2105.05210>.
- 932 [4] H. H. BAUSCHKE AND P. L. COMBETTES, *Convex Analysis And Monotone Operator Theory In Hilbert Spaces*,  
 933 Springer, New York, 2nd ed., 2017, <https://hal.sorbonne-universite.fr/hal-01517477>.
- 934 [5] S. BOYD, *Distributed optimization and statistical learning via the alternating direction method of multipliers*,  
 935 *Foundations and Trends in Machine Learning*, 3 (2010), pp. 1–122, <https://doi.org/10.1561/22000000016>.
- 936 [6] R. I. BOŦ, E. R. CSETNEK, AND A. HEINRICH, *A primal-dual splitting algorithm for finding zeros of sums of*  
 937 *maximal monotone operators*, SIAM Journal on Optimization, 23 (2013), pp. 2011–2036, [https://doi.org/](https://doi.org/10.1137/12088255)  
 938 [10.1137/12088255](https://doi.org/10.1137/12088255).
- 939 [7] R. I. BOŦ AND C. HENDRICH, *Convergence analysis for a primal-dual monotone + skew splitting algorithm*  
 940 *with applications to total variation minimization*, Journal of Mathematical Imaging and Vision, 49 (2014),  
 941 pp. 551–568, <https://doi.org/10.1007/s10851-013-0486-8>.
- 942 [8] L. BRICEÑO-ARIAS AND S. LÓPEZ, *A projected primal–dual method for solving constrained monotone inclusions*,  
 943 *Journal of Optimization Theory and Applications*, 180 (2019), <https://doi.org/10.1007/s10957-018-1430-2>.
- 944 [9] L. M. BRICEÑO-ARIAS AND P. L. COMBETTES, *A monotone+skew splitting model for composite monotone*  
 945 *inclusions in duality*, SIAM Journal on Optimization, 21 (2011), pp. 1230–1250, [https://doi.org/10.1137/](https://doi.org/10.1137/10081602X)  
 946 [10081602X](https://doi.org/10.1137/10081602X).
- 947 [10] T. A. BUBBA, M. GALINIER, M. LASSAS, M. PRATO, L. RATTI, AND S. SILTANEN, *Deep neural networks*  
 948 *for inverse problems with pseudodifferential operators: An application to limited-angle tomography*, SIAM  
 949 *Journal on Imaging Sciences*, 14 (2021), pp. 470–505, <https://doi.org/10.1137/20M1343075>.
- 950 [11] M. BURGER, A. SAWATZKY, AND G. STEIDL, *First Order Algorithms in Variational Image Processing*, Springer  
 951 *International Publishing*, Cham, 2016, pp. 345–407, [https://doi.org/10.1007/978-3-319-41589-5\\_10](https://doi.org/10.1007/978-3-319-41589-5_10).
- 952 [12] G. T. BUZZARD, S. H. CHAN, S. SREEHARI, AND C. A. BOUMAN, *Plug-and-play unplugged: Optimization-free*  
 953 *reconstruction using consensus equilibrium*, SIAM Journal on Imaging Sciences, 11 (2018), pp. 2001–2020,  
 954 <https://doi.org/10.1137/17M1122451>.
- 955 [13] E. J. CANDÉS, J. ROMBERG, AND T. TAO, *Robust uncertainty principles: exact signal reconstruction from highly*  
 956 *incomplete frequency information*, IEEE Transactions on Information Theory, 52 (2006), pp. 489–509.
- 957 [14] A. CHAMBOLE AND T. POCK, *A first-order primal-dual algorithm for convex problems with applications to*  
 958 *imaging*, Journal of Mathematical Imaging and Vision, 40 (2010), pp. 120–145.
- 959 [15] C. CHAUX, P. L. COMBETTES, J.-C. PESQUET, AND V. R. WAJS, *A variational formulation for frame-based*  
 960 *inverse problems*, Inverse Problems, 23 (2007), pp. 1495–1518, <https://doi.org/10.1088/0266-5611/23/4/008>.
- 961 [16] P. CHEN, J. C. HUANG, AND X. ZHANG, *A primal-dual fixed point algorithm for convex separable minimization*  
 962 *with applications to image restoration*, Inverse Problems, 29 (2013), p. 025011.

- 963 [17] E. CHOUZENOUX, A. JEZIERSKA, J. PESQUET, AND H. TALBOT, *A convex approach for image restoration*  
964 *with exact poisson–gaussian likelihood*, SIAM Journal on Imaging Sciences, 8 (2015), pp. 2662–2682, <https://doi.org/10.1137/15M1014395>.  
965
- 966 [18] E. CHOUZENOUX, J.-C. PESQUET, C. RIDDELL, M. SAVANIER, AND Y. TROUSSET, *Convergence of proximal*  
967 *gradient algorithm in the presence of adjoint mismatch*, Inverse Problems, 37 (2021), <http://iopscience.iop.org/article/10.1088/1361-6420/abd85c>.  
968
- 969 [19] P. COMBETTES AND J.-C. PESQUET, *Proximal Splitting Methods in Signal Processing*, Springer-Verlag, NY,  
970 2010, pp. 185–212.
- 971 [20] P. COMBETTES AND J.-C. PESQUET, *Deep neural network structures solving variational inequalities*, Set-Valued  
972 and Variational Analysis, 28 (2020), pp. 491–518, <https://doi.org/10.1007/s11228-019-00526-z>.
- 973 [21] P. L. COMBETTES AND J. PESQUET, *Fixed point strategies in data science*, IEEE Transactions on Signal Pro-  
974 cessing, 69 (2021), pp. 3878–3905.
- 975 [22] P. L. COMBETTES AND J.-C. PESQUET, *Primal-dual splitting algorithm for solving inclusions with mixtures of*  
976 *composite, lipschitzian, and parallel-sum type monotone operators*, Set-Valued and Variational Analysis, 20  
977 (2011), pp. 307–330, <https://doi.org/10.1007/s11228-011-0191-y>.
- 978 [23] P. L. COMBETTES AND J.-C. PESQUET, *Proximal Splitting Methods in Signal Processing*, in Fixed-Point Algo-  
979 rithms for Inverse Problems in Science and Engineering, Bauschke, H. Burachik, R. Combettes, P. Elser,  
980 V. Luke, D. Wolkowicz, and H. (Eds.), eds., 2011, pp. 185–212.
- 981 [24] P. L. COMBETTES AND V. WAJS, *Signal recovery by proximal forward-backward splitting*, Multiscale Modeling  
982 and Simulation: A SIAM Interdisciplinary Journal, 4 n°4 (2005), pp. 1164–1200.
- 983 [25] L. CONDAT, *A primal–dual splitting method for convex optimization involving lipschitzian, proximable and linear*  
984 *composite terms*, Journal of Optimization Theory and Applications, 158 (2013), pp. 460–479.
- 985 [26] L. CONDAT, D. KITAHARA, A. CONTRERAS, AND A. HIRABAYASHI, *Proximal splitting algorithms: A tour of*  
986 *recent advances, with new twists.*, arXiv: Optimization and Control, (2020).
- 987 [27] I. DAUBECHIES, M. DEFRISE, AND C. DE MOL, *An iterative thresholding algorithm for linear inverse problems*  
988 *with a sparsity constraint*, Communications on Pure and Applied Mathematics, 57 (2004), pp. 1413–1457,  
989 <https://doi.org/https://doi.org/10.1002/cpa.20042>.
- 990 [28] Y. DONG, P. C. HANSEN, M. E. HOCHSTENBACH, AND N. A. BROGAARD RIIS, *Fixing nonconvergence of*  
991 *algebraic iterative reconstruction with an unmatched backprojector*, SIAM Journal on Scientific Computing,  
992 41 (2019), pp. A1822–A1839.
- 993 [29] Y. DRORI, S. SABACH, AND M. TEBoulLE, *A simple algorithm for a class of nonsmooth convex-concave saddle-*  
994 *point problems*, Oper. Res. Lett., 43 (2015), pp. 209–214, <https://doi.org/10.1016/j.orl.2015.02.001>.
- 995 [30] T. ELFVING AND P. C. HANSEN, *Unmatched projector/backprojector pairs: Perturbation and convergence analy-*  
996 *sis*, SIAM Journal on Scientific Computing, 40 (2018), pp. A573–A591.
- 997 [31] E. ESSER, X. ZHANG, AND T. F. CHAN, *A general framework for a class of first order primal-dual algorithms*  
998 *for convex optimization in imaging science*, SIAM Journal on Imaging Sciences, 3 (2010), pp. 1015–1046,  
999 <https://doi.org/10.1137/09076934X>.
- 1000 [32] F. XU AND K. MUELLER, *A comparative study of popular interpolation and integration methods for use in*  
1001 *computed tomography*, in Proceedings of the 3rd IEEE International Symposium on Biomedical Imaging:  
1002 Nano to Macro, 2006., Apr 2006, pp. 1252–1255.
- 1003 [33] D. GABAY AND B. MERCIER, *A dual algorithm for the solution of nonlinear variational problems via finite*  
1004 *element approximation*, Computers & Mathematics with Applications, 2 (1976), pp. 17–40, [https://doi.org/10.1016/0898-1221\(76\)90003-1](https://doi.org/10.1016/0898-1221(76)90003-1).
- 1005
- 1006 [34] J. HERTRICH, S. NEUMAYER, AND G. STEIDL, *Convolutional proximal neural networks and Plug-and-Play al-*  
1007 *gorithms*, Linear Algebra and its Applications, 631 (2021), pp. 203–234, <https://doi.org/10.1016/j.laa.2021.09.004>.  
1008
- 1009 [35] A. C. KAK AND M. SLANEY, *Principles of Computerized Tomographic Imaging*, Society of Industrial and Applied  
1010 Mathematics, 2001.
- 1011 [36] N. KOMODAKIS AND J.-C. PESQUET, *Playing with duality: An overview of recent primal-dual approaches for*  
1012 *solving large-scale optimization problems*, IEEE Signal Processing Magazine, 32 (2015), pp. 31–54, <https://doi.org/10.1109/MSP.2014.2377273>.  
1013
- 1014 [37] S. KULLBACK AND R. A. LEIBLER, *On information and sufficiency*, Ann. Math. Statist., 22 (1951), pp. 79–86.
- 1015 [38] D. A. LORENZ, S. ROSE, AND F. SCHÖPFER, *The randomized Kaczmarz method with mismatched adjoint*, BIT  
1016 Numerical Mathematics, 58 (2018), pp. 1079–1098.
- 1017 [39] D. A. LORENZ AND F. SCHNEPPE, *Chambolle-pock’s primal-dual method with mismatched adjoint*, 2022, <https://doi.org/10.48550/arxiv.2201.04928>.  
1018
- 1019 [40] I. LORIS AND C. VERHOEVEN, *On a generalization of the iterative soft-thresholding algorithm for the case of*  
1020 *non-separable penalty*, Inverse Problems, 27 (2011), p. 125007.
- 1021 [41] J. NUYTS, B. D. MAN, J. A. FESSLER, W. ZBIJEWSKI, AND F. J. BEEKMAN, *Modelling the physics in the*  
1022 *iterative reconstruction for transmission computed tomography*, Physics in Medicine and Biology, 58 (2013),

- 1023 pp. R63–R96, <https://doi.org/10.1088/0031-9155/58/12/r63>.
- 1024 [42] K. M. PER CHRISTIAN HANSEN, KEN HAYAMI, *Gmres methods for tomographic reconstruction with an un-*  
1025 *matched back projector*, tech. report, 2022, <https://arxiv.org/pdf/2110.01481.pdf>.
- 1026 [43] J.-C. PESQUET, A. REPETTI, M. TERRIS, AND Y. WIAUX, *Learning maximally monotone operators for image re-*  
1027 *covery*, SIAM Journal on Imaging Sciences, 14 (2021), pp. 1206–1237, <https://doi.org/10.1137/20M1387961>.
- 1028 [44] C. H. R. I. BOT, *A douglas-rachford type primal-dual method for solving inclusions with mixtures of composite*  
1029 *and parallel-sum type monotone operators*, SIAM Journal on Optimization, 23 (2013), p. 2541–2565, <https://doi.org/10.1137/120901106>.
- 1030
- 1031 [45] Z. RAMZI, P. CIUCIU, AND J.-L. STARCK, *Benchmarking proximal methods acceleration enhancements for*  
1032 *cs-acquired mr image analysis reconstruction*, in SPARS 2019 - Signal Processing with Adaptive Sparse  
1033 Structured Representations Workshop, 2019, <https://arxiv.org/abs/1807.04005>.
- 1034 [46] E. J. REID, L. F. DRUMMY, C. A. BOUMAN, AND G. T. BUZZARD, *Multi-resolution data fusion for super*  
1035 *resolution imaging*, IEEE Transactions on Computational Imaging, 8 (2022), pp. 81–95, <https://doi.org/10.1109/TCI.2022.3140551>.
- 1036
- 1037 [47] C. RIDDELL, B. BENDRIEM, M. H. BOURGUIGNON, AND J. P. KERNEVEZ, *The approximate inverse and con-*  
1038 *jugate gradient: non-symmetrical algorithms for fast attenuation correction in SPECT*, Physics in Medicine  
1039 and Biology, 40 (1995), pp. 269–281.
- 1040 [48] L. RUDIN, S. OSHER, AND E. FATEMI, *Nonlinear total variation based noise removal algorithms*, Physica D:  
1041 Nonlinear Phenomena, 60 (1992), pp. 259–268.
- 1042 [49] H. SADEGHI, S. BANERT, AND P. GISELSSON, *Forward-backward splitting with deviations for monotone inclu-*  
1043 *sions*, 2021, <https://doi.org/10.48550/arxiv.2112.00776>.
- 1044 [50] M. SAVANIER, E. CHOUZENOUX, J.-C. PESQUET, AND C. RIDDELL, *Unmatched preconditioning of the prox-*  
1045 *imal gradient algorithm*, tech. report, 2022, [https://arxiv.org/abs/http://www.optimization-online.org/](https://arxiv.org/abs/http://www.optimization-online.org/DB_FILE/2021/12/8741.pdf)  
1046 [DB\\_FILE/2021/12/8741.pdf](http://www.optimization-online.org/DB_FILE/2021/12/8741.pdf).
- 1047 [51] E. Y. SIDKY, P. C. HANSEN, J. S. JØRGENSEN, AND X. PAN, *Iterative image reconstruction for CT with*  
1048 *unmatched projection matrices using the generalized minimal residual algorithm*, tech. report, 2022.
- 1049 [52] E. Y. SIDKY, J. H. JØRGENSEN, AND X. PAN, *Convex optimization problem prototyping for image reconstruction*  
1050 *in computed tomography with the Chambolle-Pock algorithm.*, Physics in medicine and biology, 57 10 (2012),  
1051 pp. 3065–91.
- 1052 [53] P. TSENG, *A modified forward-backward splitting method for maximal monotone mappings*, SIAM Journal on  
1053 Control and Optimization, 38 (2000), pp. 431–446, <https://doi.org/10.1137/S0363012998338806>.
- 1054 [54] W. VAN AARLE, W. J. PALENSTIJN, J. CANT, E. JANSSENS, F. BLEICHRODT, A. DABRAVOLSKI, J. D. BEEN-  
1055 HOUWER, K. J. BATENBURG, AND J. SIJBERS, *Fast and flexible x-ray tomography using the ASTRA toolbox*,  
1056 Opt. Express, 24 (2016), pp. 25129–25147, <https://doi.org/10.1364/OE.24.025129>.
- 1057 [55] W. VAN AARLE, W. J. PALENSTIJN, J. DE BEENHOUWER, T. ALTANTZIS, S. BALS, K. J. BATENBURG, AND  
1058 J. SIJBERS, *The ASTRA toolbox: A platform for advanced algorithm development in electron tomography*,  
1059 Ultramicroscopy, 157 (2015), pp. 35–47, <https://doi.org/10.1016/j.ultramic.2015.05.002>.
- 1060 [56] B. C. VU, *A splitting algorithm for dual monotone inclusions involving cocoercive operators*, Advances in Com-  
1061 putational Mathematics, 38 (2013), pp. 667–681, <https://doi.org/10.1007/s10444-011-9254-8>.
- 1062 [57] G. ZENG AND G. GULLBERG, *A ray-driven backprojector for backprojection filtering and filtered backprojection*  
1063 *algorithms*, in 1993 IEEE Conference Record Nuclear Science Symposium and Medical Imaging Conference,  
1064 1993, pp. 1199–1201, <https://doi.org/10.1109/NSSMIC.1993.701833>.
- 1065 [58] G. L. ZENG, *Counter examples for unmatched projector/backprojector in an iterative algorithm*, Chinese Journal  
1066 of Academic Radiology, (2019).
- 1067 [59] G. L. ZENG AND G. T. GULLBERG, *Unmatched projector/backprojector pairs in an iterative reconstruction*  
1068 *algorithm*, IEEE Transactions on Medical Imaging, (2000).