On the proximity operator of the sum of two closed and convex functions *

SAMIR ADLY†, LOÏC BOURDIN‡, AND FABIEN CAUBET§

Abstract. The main result of this paper provides an explicit decomposition of the proximity operator of the sum of two closed and convex functions. For this purpose, we introduce a new operator, called *f-proximity operator*, generalizing the classical notion. After providing some properties and characterizations, we discuss the relations between the *f-proximity operator* and the classical Douglas-Rachford operator. In particular we provide a one-loop algorithm allowing to compute numerically this new operator, and thus the proximity operator of the sum of two closed and convex functions. Finally we illustrate the usefulness of our main result in the context of sensitivity analysis of linear variational inequalities of second kind in a Hilbert space.

Key words. convex analysis, proximity operator, Douglas-Rachford operator, Forward-Backward operator, sensitivity analysis, variational inequality.

AMS subject classifications. 46N10, 47N10, 49J40, 49Q12

1. Introduction, notations and basics.

1.1. Introduction.

The *proximity operator* of a proper, closed, convex and extended-real-valued function was first introduced by J.-J. Moreau in 1965 in [18] and can be viewed as an extension of the projection operator on a closed and convex subset of a Hilbert space. This wonderful tool plays an important role, from both theoretical and numerical points of view, in convex optimization problems (see, e.g., [5, 16, 20, 22]), inverse problems (see, e.g., [4, 6]), signal processing (see, e.g., [8]), etc. We also refer to [7, 12] and references therein. For the rest of this introduction, we use standard notations of convex analysis. For the reader who is not acquainted with convex analysis, we refer to Section 1.2 for notations and basics.

Motivations from a sensitivity analysis. The present paper was initially motivated by the sensitivity analysis, with respect to a nonnegative parameter $t \geq 0$, of some parameterized linear variational inequalities of second kind in a Hilbert space H, with a corresponding functional denoted by $h \in \Gamma_0(\mathrm{H})$. In that framework, the solution $u(t) \in \mathrm{H}$ (that depends on the parameter t) can be expressed in terms of the proximity operator of h denoted by prox_h. As a consequence, the differentiability of $u(\cdot)$ at $t = 0$ is strongly related to the regularity of prox_h. If h is a smooth functional, one can easily compute (from the classical implicit function theorem for instance) the differential of prox_h, and then the sensitivity analysis can be achieved. In that smooth case, note that the variational inequality can actually be reduced to an equality. On the other hand, if $h = \iota_{\mathrm{K}}$ is the indicator function of a nonempty closed and convex subset K of H, then $\mathrm{prox}_h = \mathrm{proj}_{\mathrm{K}}$ is the classical projection operator on K. In that case, a result of F. Mignot (see [17, Theorem 2.1 p.145], see also [13, Theorem 2 p.620]) provides an asymptotic development of $\mathrm{prox}_h = \mathrm{proj}_{\mathrm{K}}$ and permits to obtain a differentiability result on $u(\cdot)$ at $t = 0$.

In a parallel work (in progress) of the authors on some shape optimization problems

*Submitted to the editors on March 31, 2017.

†Institut de recherche XLIM. UMR CNRS 7252. Université de Limoges, France (samir.adly@unilim.fr)

‡Institut de recherche XLIM. UMR CNRS 7252. Université de Limoges, France (loic.bourdin@unilim.fr),

§Institut de Mathématiques de Toulouse. UMR CNRS 5219. Université de Toulouse, France (fabien.caubet@math.univ-toulouse.fr)

with unilateral contact and friction, the considered variational inequality involves the sum of two functions, that is, $h = f + g$ where $f = \iota_{\mathrm{K}}$ is the indicator function of a nonempty closed and convex set of constraints K, and $g \in \Gamma_0(\mathrm{H})$ is a smooth functional (derived from the regularization of the friction functional in view of a numerical treatment). Despite the regularity of g, note that the variational inequality here cannot be reduced to an equality due to the presence of the constraint set K. In that framework, in order to get an asymptotic development of $\mathrm{prox}_h = \mathrm{prox}_{f+g}$, a first and natural strategy would be to develop a splitting method, looking for an explicit expression of prox_{f+g} depending only on the knowledge of prox_f and prox_g. Unfortunately, this question still remains an open challenge in the literature. Let us mention that Y.-L. Yu provides in [24] some necessary and/or sufficient conditions on general functions f, $g \in \Gamma_0(\mathrm{H})$ under which $\mathrm{prox}_{f+g} = \mathrm{prox}_f \circ \mathrm{prox}_g$. Unfortunately, these conditions are very restrictive and are not satisfied in most of cases.

Before coming to the main topic of this paper, we recall that a wide literature is already concerned with the sensitivity analysis of parameterized (linear and nonlinear) variational inequalities. We refer for instance to [3, 13, 19, 23] and references therein. The results in there are considered in very general frameworks. We precise that our original objective was to look for a simple and compact formula for the derivative $u'(0)$ in the very particular case described above, that is, in the context of a linear variational inequality and with $h = f + g$ where f is an indicator function and g is a smooth functional. For this purpose, we were led to consider the proximity operator of the sum of two proper closed and convex functions, to introduce a new operator and finally to prove the results presented in this paper.

Introduction of the f-proximity operator and main result. Let us consider general functions f, $g \in \Gamma_0(\mathrm{H})$. Section 2 is devoted to the introduction (see Definition 2.1) of a new operator denoted by prox^f_g, called *f-proximity operator* of g and defined by

$$\mathrm{prox}^f_g := \left(\mathrm{I} + \partial g \circ \mathrm{prox}_f\right)^{-1}.$$

We prove that its domain satisfies $\mathrm{D}(\mathrm{prox}^f_g) = \mathrm{H}$ if and only if $\partial(f + g) = \partial f + \partial g$ (see Proposition 2.4), and that prox^f_g can be seen as a generalization of prox_g in the sense that, if f is constant for instance, then $\mathrm{prox}^f_g = \mathrm{prox}_g$. More general sufficient (and necessary) conditions under which $\mathrm{prox}^f_g = \mathrm{prox}_g$ are provided in Propositions 2.11 and 2.14. Note that $\mathrm{prox}^f_g : \mathrm{H} \rightrightarrows \mathrm{H}$ is *a priori* a set-valued operator. We provide in Proposition 2.17 some sufficient conditions under which prox^f_g is single-valued. Some examples illustrate all the previous results throughout the section (see Examples 2.2, 2.3, 2.6 and 2.16).

Finally, if the additivity condition $\partial(f + g) = \partial f + \partial g$ is satisfied, the main result of the present paper (see Theorem 2.7) provides the equality

$$\mathrm{prox}_{f+g} = \mathrm{prox}_f \circ \mathrm{prox}^f_g.$$

Theorem 2.7 allows to prove in a simple and concise way almost all other results of this paper, making it central in our work.

Relations with the classical Douglas-Rachford operator and algorithms. Recall that the proximity operator prox_{f+g} is strongly related to the minimization problem

$$\mathrm{argmin}\ f + g,$$

since the solutions are exactly the fixed points of prox_{f+g}. In the sequel, we will assume that the above problem admits at least one solution. In most of cases, prox_{f+g} cannot be easily computable, even if prox_f and prox_g are known. As a consequence, to the best of our knowledge, no *proximal algorithm* $x_{n+1} = \mathrm{prox}_{f+g}(x_n)$, using only the knowledge of prox_f and prox_g, has been provided in the literature.

The classical *Douglas-Rachford operator*, introduced in [9] and denoted by $\mathcal{T}_{f,g}$ (see Section 3 for details), provides an algorithm $x_{n+1} = \mathcal{T}_{f,g}(x_n)$ that is weakly convergent to some $x^* \in \mathrm{H}$ satisfying

$$\mathrm{prox}_f(x^*) \in \mathrm{argmin}\ f + g.$$

Even if the *Douglas-Rachford algorithm* is not a proximal algorithm in general, it is a very powerful tool since it is a one-loop algorithm, allowing to solve the above minimization problem, that only requires the knowledge of prox_f and prox_g. We refer to [10, 15] and [2, Section 27.2 p.400] for more details.

Section 3 deals with the relations between the Douglas-Rachford operator $\mathcal{T}_{f,g}$ and the f-proximity operator prox^f_g introduced in this paper. Precisely, for all $x \in \mathrm{H}$, we prove in Proposition 3.2 that $\mathrm{prox}^f_g(x)$ coincides with the set of fixed points of $\overline{\mathcal{T}}_{f,g}(x,\cdot)$, where $\overline{\mathcal{T}}_{f,g}(x,\cdot)$ denotes a x-dependent generalization of the classical Douglas-Rachford operator $\mathcal{T}_{f,g}$. We refer to Section 3 for the precise definition of $\overline{\mathcal{T}}_{f,g}(x,\cdot)$ that only depends on the knowledge of prox_f and prox_g. In particular, if $x \in \mathrm{D}(\mathrm{prox}^f_g)$, we prove in Theorem 3.3 that the fixed-point algorithm $y_{k+1} = \overline{\mathcal{T}}_{f,g}(x, y_k)$, denoted by $(\mathcal{A}_1)$, weakly converges to some $y^* \in \mathrm{prox}^f_g(x)$. Moreover, if the additivity condition $\partial(f + g) = \partial f + \partial g$ is satisfied, we get from Theorem 2.7 that $\mathrm{prox}_f(y^*) = \mathrm{prox}_{f+g}(x)$. In that situation, we conclude that Algorithm $(\mathcal{A}_1)$ is a one-loop algorithm, that depends only on the knowledge of prox_f and prox_g, allowing to compute numerically $\mathrm{prox}_{f+g}(x)$.

As a consequence, a *proximal-like algorithm* $x_{n+1} = \mathrm{prox}_{f+g}(x_n)$, denoted by $(\mathcal{A}_2)$, using only the knowledge of prox_f and prox_g, can be derived in the above framework (see Remark 3.7). We refer to Definition 3.5 for the precise meaning of *proximal-like algorithm*.

The aim of the present theoretical paper is not to discuss numerical experiments and comparisons between numerical algorithms (this should be the topic of future works). However, it should be noted that, in contrary to the classical Douglas-Rachford algorithm, a proximal-like algorithm is a two-loops algorithm. As a consequence, it should not be expected from Algorithm $(\mathcal{A}_2)$ better performances than the Douglas-Rachford algorithm for solving the minimization problem $\mathrm{argmin}\ f + g$.

Some other applications and forthcoming works. Section 4 can be seen as a conclusion of the paper. Its aim is to provide a glimpse of some other applications of our main result (Theorem 2.7) and to raise open questions for forthcoming works. This section is splitted into two parts.

In Section 4.1 we consider the framework where $f,\ g \in \Gamma_0(\mathrm{H})$ with g differentiable on H. In that framework, we prove from Theorem 2.7 that prox_{f+g} is related to the classical *Forward-Backward operator* (see [7, Section 10.3 p.191] for details) denoted by $\mathcal{F}_{f,g}$. Precisely, for all $x \in \mathrm{H}$, we prove in Proposition 4.1 that $\mathrm{prox}_{f+g}(x)$ coincides with the set of fixed points of $\overline{\mathcal{F}}_{f,g}(x,\cdot)$, where $\overline{\mathcal{F}}_{f,g}(x,\cdot)$ denotes a x-dependent generalization of the classical Forward-Backward operator $\mathcal{F}_{f,g}$. We refer to Section 4.1

for the precise definition of $\overline{\mathcal{F}}_{f,g}(x,\cdot)$ that only depends on the knowledge of prox_f and ∇g. From this point, one can develop a similar strategy as in Section 3. Precisely, for all $x \in \mathrm{H}$, one can consider the one-loop algorithm $y_{k+1} = \overline{\mathcal{F}}_{f,g}(x, y_k)$, denoted by $(\mathcal{A}_3)$, in order to compute numerically $\mathrm{prox}_{f+g}(x)$, with the only knowledge of prox_f and ∇g. Moreover, one can also deduce a two-loops algorithm denoted by $(\mathcal{A}_4)$ as a potential proximal-like algorithm $x_{n+1} = \mathrm{prox}_{f+g}(x_n)$, using only the knowledge of prox_f and ∇g. Convergence proofs (under some assumptions on f and g) and numerical experiments of Algorithms $(\mathcal{A}_3)$ and $(\mathcal{A}_4)$ should be the topic of future works.

In Section 4.2 we return back to our initial motivation, namely the sensitivity analysis, with respect to a nonnegative parameter $t \geq 0$, of some parameterized linear variational inequalities of second kind. Precisely, under some assumptions (see Proposition 4.3 for details), we derive from Theorem 2.7 that if

$$u(t) = \mathrm{prox}_{f+g}(r(t)),$$

where $f = \iota_{\mathrm{K}}$ (where K is a nonempty closed convex set) and where $g \in \Gamma_0(\mathrm{H})$ is a smooth enough functional, then

$$u'(0) = \mathrm{prox}_{\varphi_f + \varphi_g}(r'(0)),$$

where $\varphi_f := \iota_C$ (where C is a nonempty closed convex subset of H related to K) and where $\varphi_g(x) := \frac{1}{2}\langle \mathrm{D}^2 g(u(0))(x), x\rangle$ for all $x \in \mathrm{H}$. We refer to Proposition 4.3 for details. It should be noted that the assumptions of Proposition 4.3 are quite restrictive, raising open questions about their relaxations (see Remark 4.5). This also should be the subject of a forthcoming work.

1.2. Notations and basics.

In this section we introduce some notations available throughout the paper and we recall some basics of convex analysis. We refer to standard books like [2, 14, 21] and references therein.

Let H be a real Hilbert space and let $\langle \cdot, \cdot \rangle$ (resp. $\| \cdot \|$) be the corresponding scalar product (resp. norm). For all subset S of H, we denote respectively by $\mathrm{int}(S)$ and $\mathrm{cl}(S)$ its interior and its closure. In the sequel we denote by $\mathrm{I} : \mathrm{H} \to \mathrm{H}$ the identity application and by $L_x : \mathrm{H} \to \mathrm{H}$ the affine operator defined by

$$L_x(y) := x - y,$$

for all x, $y \in \mathrm{H}$.

For a set-valued map $A : \mathrm{H} \rightrightarrows \mathrm{H}$, the *domain* of A is given by

$$\mathrm{D}(A) := \{x \in \mathrm{H} \mid A(x) \neq \emptyset\}.$$

We denote by $A^{-1} : \mathrm{H} \rightrightarrows \mathrm{H}$ the set-valued map defined by

$$A^{-1}(y) := \{x \in \mathrm{H} \mid y \in A(x)\},$$

for all $y \in \mathrm{H}$. Note that $y \in A(x)$ if and only if $x \in A^{-1}(y)$, for all x, $y \in \mathrm{H}$. The *range* of A is given by

$$\mathrm{R}(A) := \{y \in \mathrm{H} \mid A^{-1}(y) \neq \emptyset\} = \mathrm{D}(A^{-1}).$$

We denote by $\mathrm{Fix}(A)$ the set of all fixed points of A, that is, the set given by

$$\mathrm{Fix}(A) := \{x \in \mathrm{H} \mid x \in A(x)\}.$$

Finally, if $A(x)$ is a singleton for all $x \in \mathrm{D}(A)$, we say that A is *single-valued*.

For all extended-real-valued functions $g : \mathrm{H} \to \mathbb{R} \cup \{+\infty\}$, the *domain* of g is given by

$$\mathrm{dom}(g) := \{x \in \mathrm{H} \mid g(x) < +\infty\}.$$

We say that g is *proper* if $\mathrm{dom}(g) \neq \emptyset$, and that g is *closed* (or *lower semi-continuous*) if its epigraph is a closed subset of $\mathrm{H} \times \mathbb{R}$.

Let $g : \mathrm{H} \to \mathbb{R} \cup \{+\infty\}$ be a proper extended-real-valued function. We denote by $g^* : \mathrm{H} \to \mathbb{R} \cup \{+\infty\}$ the *conjugate* of g defined by

$$g^*(y) := \sup_{z \in \mathrm{H}} \{\langle y, z \rangle - g(z)\},$$

for all $y \in \mathrm{H}$. Recall that g^* is closed and convex.

We denote by $\Gamma_0(\mathrm{H})$ the set of all extended-real-valued functions $g : \mathrm{H} \to \mathbb{R} \cup \{+\infty\}$ that are proper closed and convex. If $g \in \Gamma_0(\mathrm{H})$, recall that $g^* \in \Gamma_0(\mathrm{H})$. The *Fenchel-Moreau equality* is given by $g^{**} = g$. For all $g \in \Gamma_0(\mathrm{H})$, we denote by $\partial g : \mathrm{H} \rightrightarrows \mathrm{H}$ the *Fenchel-Moreau subdifferential* of g defined by

$$\partial g(x) := \{y \in \mathrm{H} \mid \langle y, z - x \rangle \leq g(z) - g(x),\ \forall z \in \mathrm{H}\},$$

for all $x \in \mathrm{H}$. It is easy to check that ∂g is a monotone operator and that, for all $x \in \mathrm{H}$, $0 \in \partial g(x)$ if and only if $x \in \operatorname{argmin} g$. Moreover, for all x, $y \in \mathrm{H}$, it holds that $y \in \partial g(x)$ if and only if $x \in \partial g^*(y)$. Recall that, if g is differentiable on H, then $\partial g(x) = \{\nabla g(x)\}$ for all $x \in \mathrm{H}$.

Let $A : \mathrm{H} \to \mathrm{H}$ be a single-valued operator defined everywhere on H, and let $g \in \Gamma_0(\mathrm{H})$. We denote by $\mathrm{VI}(A, g)$ the variational inequality which consists of finding $y \in \mathrm{H}$ such that

$$-A(y) \in \partial g(y),$$

or equivalently,

$$\langle A(y), z - y \rangle + g(z) - g(y) \geq 0,$$

for all $z \in \mathrm{H}$. Then we denote by $\mathrm{Sol}_{\mathrm{VI}}(A, g)$ the set of solutions of $\mathrm{VI}(A, g)$. Recall that if A is Lipschitzian and strongly monotone, then $\mathrm{VI}(A, g)$ admits a unique solution, *i.e.* $\mathrm{Sol}_{\mathrm{VI}}(A, g)$ is a singleton.

Let $g \in \Gamma_0(\mathrm{H})$. The classical *proximity operator* of g is defined by

$$\mathrm{prox}_g := (\mathrm{I} + \partial g)^{-1}.$$

Recall that prox_g is a single-valued operator defined everywhere on H. Moreover, it can be characterized as follows:

$$\mathrm{prox}_g(x) = \operatorname{argmin} \left(g + \frac{1}{2} \| \cdot - x \|^2 \right) = \mathrm{Sol}_{\mathrm{VI}}(-L_x, g),$$

for all $x \in \mathrm{H}$. It is also well-known that

$$\mathrm{Fix}(\mathrm{prox}_g) = \operatorname{argmin} g.$$

The classical *Moreau's envelope* $\mathrm{M}_g : \mathrm{H} \to \mathbb{R}$ of g is defined by

$$\mathrm{M}_g(x) := \min \Big(g + \frac{1}{2}\| \cdot - x\|^2 \Big),$$

for all $x \in \mathrm{H}$. Recall that M_g is convex and differentiable on H with $\nabla \mathrm{M}_g = \mathrm{prox}_{g^*}$. Let us also recall the classical Moreau's decompositions

$$\mathrm{prox}_g + \mathrm{prox}_{g^*} = \mathrm{I} \qquad \text{and} \qquad \mathrm{M}_g + \mathrm{M}_{g^*} = \frac{1}{2}\| \cdot \|^2.$$

Finally, it is well-known that if $g = \iota_{\mathrm{K}}$ is the *indicator function* of a nonempty closed and convex subset K of H, that is, $\iota_{\mathrm{K}}(x) = 0$ if $x \in \mathrm{K}$ and $\iota_{\mathrm{K}}(x) = +\infty$ if not, then $\mathrm{prox}_g = \mathrm{proj}_{\mathrm{K}}$, where $\mathrm{proj}_{\mathrm{K}}$ denotes the classical projection operator on K.

2. The f-proximity operator.

2.1. Main result. Let $f,\ g \in \Gamma_0(\mathrm{H})$. In this section we introduce (see Definition 2.1) a new operator denoted by prox^f_g, generalizing the classical proximity operator prox_g. Under the additivity condition $\partial(f+g) = \partial f + \partial g$, we prove in Theorem 2.7 that prox_{f+g} can be written as the composition of prox_f with prox^f_g.

Definition 2.1 (f-proximity operator). *Let $f,\ g \in \Gamma_0(\mathrm{H})$. The f-proximity operator of g is the set-valued map $\mathrm{prox}^f_g : \mathrm{H} \rightrightarrows \mathrm{H}$ defined by*

$$\mathrm{prox}^f_g := (\mathrm{I} + \partial g \circ \mathrm{prox}_f)^{-1}.$$

Note that prox^f_g can be seen as a generalization of prox_g since $\mathrm{prox}^c_g = \mathrm{prox}_g$ for all constant $c \in \mathbb{R}$.

Example 2.2. Let us assume that $\mathrm{H} = \mathbb{R}$. We consider $f = \iota_{[-1,1]}$ and $g(x) = |x|$ for all $x \in \mathbb{R}$. In that case we obtain that $\partial g \circ \mathrm{prox}_f = \partial g$ and thus $\mathrm{prox}^f_g = \mathrm{prox}_g$.

Example 2.2 provides a simple situation where $\mathrm{prox}^f_g = \mathrm{prox}_g$ while f is not constant. We provide in Propositions 2.11 and 2.14 some general sufficient (and necessary) conditions under which $\mathrm{prox}^f_g = \mathrm{prox}_g$.

Example 2.3. Let us assume that $\mathrm{H} = \mathbb{R}$. We consider $f = \iota_{\{0\}}$ and $g(x) = |x|$ for all $x \in \mathbb{R}$. In that case we obtain that $\partial g \circ \mathrm{prox}_f(x) = [-1,1]$ for all $x \in \mathbb{R}$. As a consequence $\mathrm{prox}^f_g(x) = [x-1, x+1]$ for all $x \in \mathbb{R}$. See Figure 1 for graphic representations of prox_g and prox^f_g in that case.

Example 2.3 provides a simple illustration where prox^f_g is not single-valued. In particular it follows that prox^f_g cannot be written as a proximity operator prox_φ with $\varphi \in \Gamma_0(\mathrm{H})$. We provide in Proposition 2.17 some sufficient conditions under which prox^f_g is single-valued. Moreover, Example 2.3 provides a simple situation where $\partial g \circ \mathrm{prox}_f$ is not a monotone operator. As a consequence, it may be possible that $\mathrm{D}(\mathrm{prox}^f_g) \subsetneq \mathrm{H}$. In the next proposition, a necessary and sufficient condition under which $\mathrm{D}(\mathrm{prox}^f_g) = \mathrm{H}$ is derived.

Proposition 2.4. *Let $f,\ g \in \Gamma_0(\mathrm{H})$. It holds that $\mathrm{D}(\mathrm{prox}^f_g) = \mathrm{H}$ if and only if the additivity condition*

$$\partial(f+g) = \partial f + \partial g, \tag{1}$$

is satisfied.

Proof. We first assume that $\partial(f+g) = \partial f + \partial g$. Let $x \in \mathrm{H}$. Defining $w = \mathrm{prox}_{f+g}(x) \in \mathrm{H}$, we obtain that $x \in w + \partial(f+g)(w) = w + \partial f(w) + \partial g(w)$. Thus, there exist $w_f \in \partial f(w)$ and $w_g \in \partial g(w)$ such that $x = w + w_f + w_g$. We define $y = w + w_f \in w + \partial f(w)$. In particular we have $w = \mathrm{prox}_f(y)$. Moreover we obtain $x = y + w_g \in y + \partial g(w) = y + \partial g(\mathrm{prox}_f(y))$. We conclude that $y \in \mathrm{prox}^f_g(x)$.

Without any additional assumption and directly from the definition of the subdifferential, one can easily see that the inclusion $\partial f(w) + \partial g(w) \subset \partial(f+g)(w)$ is always satisfied for every $w \in \mathrm{H}$. Now let us assume that $\mathrm{D}(\mathrm{prox}^f_g) = \mathrm{H}$. Let $w \in \mathrm{H}$ and let $z \in \partial(f+g)(w)$. We consider $x = w + z \in w + \partial(f+g)(w)$. In particular it holds that $w = \mathrm{prox}_{f+g}(x)$. Since $\mathrm{D}(\mathrm{prox}^f_g) = \mathrm{H}$, there exists $y \in \mathrm{prox}^f_g(x)$ and thus it holds that $x \in y + \partial g(\mathrm{prox}_f(y))$. Moreover, since $y \in \mathrm{prox}_f(y) + \partial f(\mathrm{prox}_f(y))$, we get that $x \in \mathrm{prox}_f(y) + \partial f(\mathrm{prox}_f(y)) + \partial g(\mathrm{prox}_f(y)) \subset \mathrm{prox}_f(y) + \partial(f+g)(\mathrm{prox}_f(y))$. Thus it holds that $\mathrm{prox}_f(y) = \mathrm{prox}_{f+g}(x) = w$. Moreover, since $x \in \mathrm{prox}_f(y) + \partial f(\mathrm{prox}_f(y)) + \partial g(\mathrm{prox}_f(y))$, we obtain that $x \in w + \partial f(w) + \partial g(w)$. We have proved that $z = x - w \in \partial f(w) + \partial g(w)$. This concludes the proof. □

In most of the present paper, we will assume that Condition (1) is satisfied. It is not our aim here to discuss the weakest qualification condition ensuring Condition (1). A wide literature already deals with this topic (see, e.g., [1, 11, 20]). However, we recall in the following remark the classical sufficient condition of Moreau-Rockafellar under which Condition (1) holds true (see, e.g., [2, Corollary 16.38 p.234]), and we provide a simple example where Condition (1) does not holds and $\mathrm{D}(\mathrm{prox}^f_g) \subsetneq \mathrm{H}$.

Remark 2.5 (Moreau-Rockafellar theorem). Let $f, g \in \Gamma_0(\mathrm{H})$ such that $\mathrm{dom}(f) \cap \mathrm{int}(\mathrm{dom}(g)) \neq \emptyset$. Then $\partial(f+g) = \partial f + \partial g$.

Example 2.6. Let us assume that $\mathrm{H} = \mathbb{R}$. We consider $f = \iota_{\mathbb{R}^-}$ and $g(x) = \iota_{\mathbb{R}^+}(x) - \sqrt{x}$ for all $x \in \mathbb{R}$. In that case, one can easily check that $\partial f(0) + \partial g(0) = \emptyset \subsetneq \mathbb{R} = \partial(f+g)(0)$ and $\mathrm{D}(\mathrm{prox}^f_g) = \emptyset \subsetneq \mathrm{H}$.

We are now in position to state and prove the main result of the present paper.

THEOREM 2.7. *Let $f, g \in \Gamma_0(\mathrm{H})$ such that $\partial(f+g) = \partial f + \partial g$. It holds that*

$$\mathrm{prox}_{f+g} = \mathrm{prox}_f \circ \mathrm{prox}^f_g.$$

In other words, for every $x \in \mathrm{H}$, we have $\mathrm{prox}_{f+g}(x) = \mathrm{prox}_f(z)$ for all $z \in \mathrm{prox}^f_g(x)$.

Proof. Let $x \in \mathrm{H}$ and let $y \in \mathrm{prox}^f_g(x)$ constructed as in the first part of the proof of Proposition 2.4. In particular it holds that $\mathrm{prox}_f(y) = \mathrm{prox}_{f+g}(x)$. Let $z \in \mathrm{prox}^f_g(x)$. We know that $x - y \in \partial g(\mathrm{prox}_f(y))$ and $x - z \in \partial g(\mathrm{prox}_f(z))$. Since ∂g is a monotone operator, we obtain that

$$\langle (x-y) - (x-z), \mathrm{prox}_f(y) - \mathrm{prox}_f(z) \rangle \geq 0.$$

From the cocoercivity (see [2, Definition 4.4 p.60]) of the proximity operator, we obtain that

$$0 \geq \langle y - z, \mathrm{prox}_f(y) - \mathrm{prox}_f(z) \rangle \geq \| \mathrm{prox}_f(y) - \mathrm{prox}_f(z) \|^2 \geq 0.$$

We deduce that $\mathrm{prox}_f(z) = \mathrm{prox}_f(y) = \mathrm{prox}_{f+g}(x)$. The proof is complete. □

Remark 2.8. Let $f, g \in \Gamma_0(\mathrm{H})$ such that $\partial(f+g) = \partial f + \partial g$ and let $x \in \mathrm{H}$. Theorem 2.7 states that, even if $\mathrm{prox}^f_g(x)$ is not a singleton, all elements of $\mathrm{prox}^f_g(x)$

has the same value through the proximity operator prox_f, and this value is equal to $\mathrm{prox}_{f+g}(x)$.

Remark 2.9. Note that the additivity condition $\partial(f+g) = \partial f + \partial g$ is not only sufficient, but also necessary for the validity of the equality $\mathrm{prox}_{f+g} = \mathrm{prox}_f \circ \mathrm{prox}_g^f$. Indeed, from Proposition 2.4, if $\partial f + \partial g \subsetneq \partial(f+g)$, then there exists $x \in \mathrm{H}$ such that $\mathrm{prox}_g^f(x) = \emptyset$ and thus $\mathrm{prox}_{f+g}(x) \neq \mathrm{prox}_f \circ \mathrm{prox}_g^f(x)$.

Remark 2.10. Let f, $g \in \Gamma_0(\mathrm{H})$ such that $\partial(f+g) = \partial f + \partial g$. From Theorem 2.7, we deduce that $\mathrm{R}(\mathrm{prox}_{f+g}) \subset \mathrm{R}(\mathrm{prox}_f) \cap \mathrm{R}(\mathrm{prox}_g)$. If the additivity condition $\partial(f+g) = \partial f + \partial g$ is not satisfied, this remark does not hold true anymore. Indeed, with the framework of Example 2.6, we have $\mathrm{R}(\mathrm{prox}_{f+g}) = \{0\}$ while $0 \notin \mathrm{R}(\mathrm{prox}_g)$.

2.2. Additional results. Let f, $g \in \Gamma_0(\mathrm{H})$. We know that prox_g^f is a generalization of prox_g in the sense that $\mathrm{prox}_g^f = \mathrm{prox}_g$ if f is constant for instance. In the next proposition, our aim is to provide more general sufficient (and necessary) conditions under which $\mathrm{prox}_g^f = \mathrm{prox}_g$. We will base our discussion on the following conditions:

$$\forall x \in \mathrm{H}, \quad \partial g(x) \subset \partial g(\mathrm{prox}_f(x)), \tag{2}$$

$$\forall x \in \mathrm{H}, \quad \partial g(\mathrm{prox}_f(x)) \subset \partial g(x). \tag{3}$$

Note that Condition (2) has been introduced by Y.-L. Yu in [24] as a sufficient condition under which $\mathrm{prox}_{f+g} = \mathrm{prox}_f \circ \mathrm{prox}_g$.

PROPOSITION 2.11. *Let* f, $g \in \Gamma_0(\mathrm{H})$.

(i) *If Condition (2) is satisfied, then* $\mathrm{prox}_g(x) \in \mathrm{prox}_g^f(x)$ *for all* $x \in \mathrm{H}$.
(ii) *If Conditions (1) and (3) are satisfied, then* $\mathrm{prox}_g^f(x) = \mathrm{prox}_g(x)$ *for all* $x \in \mathrm{H}$.

In both cases, it holds that $\mathrm{prox}_{f+g} = \mathrm{prox}_f \circ \mathrm{prox}_g$.

Proof. Let $x \in \mathrm{H}$. If Condition (2) is satisfied, considering $y = \mathrm{prox}_g(x)$, we get that $x \in y + \partial g(y) \subset y + \partial g(\mathrm{prox}_f(y))$ and thus $y \in \mathrm{prox}_g^f(x)$. In particular, it holds that $\mathrm{D}(\mathrm{prox}_g^f) = \mathrm{H}$ and thus $\partial(f+g) = \partial f + \partial g$ from Proposition 2.4. On the other hand, if Conditions (1) and (3) are satisfied, then $\mathrm{D}(\mathrm{prox}_g^f) = \mathrm{H}$ from Proposition 2.4. Considering $y \in \mathrm{prox}_g^f(x)$, we get that $x \in y + \partial g(\mathrm{prox}_f(y)) \subset y + \partial g(y)$ and thus $y = \mathrm{prox}_g(x)$. The last assertion of Proposition 2.11 directly follows from Theorem 2.7. ∎

Remark 2.12. From Proposition 2.11, we deduce that Condition (2) implies Condition (1).

In the first item of Proposition 2.11 and if prox_g^f is set-valued, we are in the situation where prox_g is a selection of prox_g^f. Proposition 2.14 specifies this selection in the case where $\partial(f+g) = \partial f + \partial g$.

LEMMA 2.13. *Let* f, $g \in \Gamma_0(\mathrm{H})$. *Then* $\mathrm{prox}_g^f(x)$ *is a nonempty closed and convex subset of* H *for all* $x \in \mathrm{D}(\mathrm{prox}_g^f)$.

Proof. The proof of Lemma 2.13 is provided after the proof of Proposition 3.2 (required). ∎

PROPOSITION 2.14. *Let* f, $g \in \Gamma_0(\mathrm{H})$ *such that* $\partial(f+g) = \partial f + \partial g$ *and let* $x \in \mathrm{H}$. *If* $\mathrm{prox}_g(x) \in \mathrm{prox}_g^f(x)$, *then*

$$\mathrm{prox}_g(x) = \mathrm{proj}_{\mathrm{prox}_g^f(x)}(\mathrm{prox}_{f+g}(x)).$$

Proof. If $\mathrm{prox}_g(x) \in \mathrm{prox}_g^f(x)$, then $x \in \mathrm{D}(\mathrm{prox}_g^f)$ and thus $\mathrm{prox}_g^f(x)$ is a nonempty closed and convex subset of H from Lemma 2.13. Let $z \in \mathrm{prox}_g^f(x)$. In particular we have $\mathrm{prox}_f(z) = \mathrm{prox}_{f+g}(x)$ from Theorem 2.7. Using the fact that $x - \mathrm{prox}_g(x) \in \partial g(\mathrm{prox}_g(x))$ and $x - z \in \partial g(\mathrm{prox}_f(z)) = \partial g(\mathrm{prox}_{f+g}(x))$ together with the monotonicity of ∂g, we obtain that

$$\begin{aligned}\langle \mathrm{prox}_{f+g}(x) - \mathrm{prox}_g(x), z - \mathrm{prox}_g(x)\rangle \\ = \langle \mathrm{prox}_{f+g}(x) - \mathrm{prox}_g(x), (x - \mathrm{prox}_g(x)) - (x - z)\rangle \leq 0.\end{aligned}$$

Since $\mathrm{prox}_g(x) \in \mathrm{prox}_g^f(x)$, we conclude the proof from the classical characterization of $\mathrm{proj}_{\mathrm{prox}_g^f(x)}$. ∎

Remark 2.15. Let $f = \iota_{\{\omega\}}$ with $\omega \in \mathrm{H}$ and let $g \in \Gamma_0(\mathrm{H})$ such that $\omega \in \mathrm{int}(\mathrm{dom}(g))$. Hence the additivity condition $\partial(f+g) = \partial f + \partial g$ is satisfied from Remark 2.5. From Remark 2.10 and since $\mathrm{prox}_f = \mathrm{proj}_{\{\omega\}}$, we easily deduce that $\mathrm{R}(\mathrm{prox}_{f+g}) = \{\omega\}$. Let $x \in \mathrm{H}$ such that $\mathrm{prox}_g(x) \in \mathrm{prox}_g^f(x)$. From Proposition 2.14 we get that

$$\mathrm{prox}_g(x) = \mathrm{proj}_{\mathrm{prox}_g^f(x)}(\omega).$$

If moreover $\omega = 0$, we deduce that $\mathrm{prox}_g(x)$ is the particular selection that corresponds to the element of minimal norm in $\mathrm{prox}_g^f(x)$ (also known as the *lazy selection*). The following example is in this sense.

Example 2.16. Let us consider the framework of Example 2.3. In that case, Conditions (1) and (2) are satisfied. We deduce from Proposition 2.11 that $\mathrm{prox}_g(x) \in \mathrm{prox}_g^f(x)$ for all $x \in \mathbb{R}$. From Remark 2.15, we conclude that $\mathrm{prox}_g(x)$ is exactly the element of minimal norm in $\mathrm{prox}_g^f(x)$ for all $x \in \mathbb{R}$. This result is clearly illustrated by the graphs of prox_g and prox_g^f provided in Figure 1.

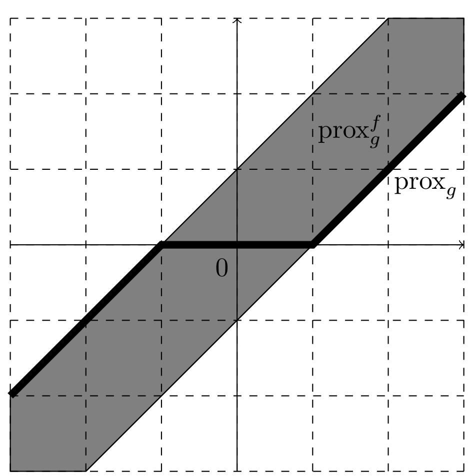

FIG. 1. *Examples 2.3 and 2.16, graph of* prox_g *in bold line, and graph of* prox_g^f *in gray.*

Let $f, g \in \Gamma_0(\mathrm{H})$ such that $\partial(f+g) = \partial f + \partial g$. From Theorem 2.7, one can easily see that, if prox_f is injective, then prox_g^f is single-valued. Since the injection of prox_f is too restrictive, other sufficient conditions under which prox_g^f is single-valued are provided from Theorem 2.7 in the next proposition.

PROPOSITION 2.17. *Let $f,\ g \in \Gamma_0(\mathrm{H})$ such that $\partial(f+g) = \partial f + \partial g$. If either ∂f or ∂g is single-valued, then prox_g^f is single-valued.*

Proof. Let $x \in \mathrm{H}$ and let $z_1,\ z_2 \in \mathrm{prox}_g^f(x)$. From Theorem 2.7, it holds that $\mathrm{prox}_f(z_1) = \mathrm{prox}_f(z_2) = \mathrm{prox}_{f+g}(x)$. If the operator ∂f is single-valued, we obtain that $z_1 = \mathrm{prox}_{f+g}(x) + \partial f(\mathrm{prox}_{f+g}(x)) = z_2$. If the operator ∂g is single-valued, we get $x - z_1 = \partial g(\mathrm{prox}_f(z_1)) = \partial g(\mathrm{prox}_f(z_2)) = x - z_2$ and thus $z_1 = z_2$. □

3. Relations with the Douglas-Rachford operator.

Let $f,\ g \in \Gamma_0(\mathrm{H})$. The classical *Douglas-Rachford operator* $\mathcal{T}_{f,g} : \mathrm{H} \to \mathrm{H}$ associated to f and g is usually defined by

$$\mathcal{T}_{f,g}(y) := y - \mathrm{prox}_f(y) + \mathrm{prox}_g(2\mathrm{prox}_f(y) - y),$$

for all $y \in \mathrm{H}$. We refer to [9, 10, 15] and to [2, Section 27.2 p.400] for more details.

One aim of this section is to study the relations between the f-proximity operator prox_g^f introduced in this paper and the Douglas-Rachford operator $\mathcal{T}_{f,g}$. For this purpose, we introduce an extension $\overline{\mathcal{T}}_{f,g} : \mathrm{H}\times\mathrm{H} \to \mathrm{H}$ of the classical Douglas-Rachford operator defined by

$$\overline{\mathcal{T}}_{f,g}(x,y) := y - \mathrm{prox}_f(y) + \mathrm{prox}_g(x + \mathrm{prox}_f(y) - y),$$

for all $x,\ y \in \mathrm{H}$.

Note that $\mathcal{T}_{f,g}(y) = \overline{\mathcal{T}}_{f,g}(\mathrm{prox}_f(y), y)$ for all $y \in \mathrm{H}$, and that the definition of $\overline{\mathcal{T}}_{f,g}$ only depends on the knowledge of prox_f and prox_g.

3.1. Several characterizations of prox_g^f.

Let $f,\ g \in \Gamma_0(\mathrm{H})$. In this section, our aim is to derive several characterizations of prox_g^f in terms of solutions of variational inequalities, of minimization problems and of fixed point problems (see Proposition 3.2).

LEMMA 3.1. *Let $f,\ g \in \Gamma_0(\mathrm{H})$. It holds that*

$$\overline{\mathcal{T}}_{f,g}(x,\cdot) = \mathrm{prox}_{g^* \circ L_x} \circ \mathrm{prox}_{f^*},$$

for all $x \in \mathrm{H}$.

Proof. Let $x \in \mathrm{H}$. Lemma 3.1 directly follows from the equality $\mathrm{prox}_{g^*\circ L_x} = L_x \circ \mathrm{prox}_{g^*} \circ L_x$ (see [2, Proposition 23.29 p.342]) and from Moreau's decompositions. □

PROPOSITION 3.2. *Let $f,\ g \in \Gamma_0(\mathrm{H})$. It holds that*

$$\mathrm{prox}_g^f(x) = \mathrm{Sol}_{\mathrm{VI}}(\mathrm{prox}_f, g^* \circ L_x) = \mathrm{argmin}\,(\mathrm{M}_{f^*} + g^* \circ L_x) = \mathrm{Fix}(\overline{\mathcal{T}}_{f,g}(x,\cdot)),$$

for all $x \in \mathrm{H}$.

Proof. In this proof we will use standard properties of convex analysis recalled in Section 1.2. Let $x \in \mathrm{H}$. One can easily prove that $\partial(g^* \circ L_x) = -\partial g^* \circ L_x$. For all $y \in \mathrm{H}$, it holds that

$$\begin{aligned} y \in \mathrm{prox}_g^f(x) &\Longleftrightarrow x - y \in \partial g(\mathrm{prox}_f(y)) \\ &\Longleftrightarrow \mathrm{prox}_f(y) \in \partial g^*(x-y) \\ &\Longleftrightarrow -\mathrm{prox}_f(y) \in \partial(g^* \circ L_x)(y). \end{aligned}$$

Moreover, since $\mathrm{dom}(\mathrm{M}_{f^*}) = \mathrm{H}$ and from Remark 2.5, we have

$$\begin{aligned} -\mathrm{prox}_f(y) \in \partial(g^* \circ L_x)(y) &\Longleftrightarrow 0 \in \nabla \mathrm{M}_{f^*}(y) + \partial(g^* \circ L_x)(y) \\ &\Longleftrightarrow 0 \in \partial(\mathrm{M}_{f^*} + g^* \circ L_x)(y). \end{aligned}$$

Finally,

$$-\mathrm{prox}_f(y) \in \partial(g^* \circ L_x)(y) \Longleftrightarrow \mathrm{prox}_{f^*}(y) \in y + \partial(g^* \circ L_x)(y)$$
$$\Longleftrightarrow y = \mathrm{prox}_{g^* \circ L_x} \circ \mathrm{prox}_{f^*}(y).$$

This concludes the proof from Lemma 3.1. ∎

Proof of Lemma 2.13. Let $x \in \mathrm{D}(\mathrm{prox}_g^f)$. In particular $\mathrm{prox}_g^f(x)$ is not empty. From Proposition 3.2, we have

$$\mathrm{prox}_g^f(x) = \mathrm{argmin}\,(\mathrm{M}_{f^*} + g^* \circ L_x).$$

Since $\mathrm{M}_{f^*} + g^* \circ L_x \in \Gamma_0(\mathrm{H})$, one can easily deduce that $\mathrm{prox}_g^f(x)$ is closed and convex. ∎

3.2. A one-loop algorithm in order to compute prox_g^f numerically. Let f, $g \in \Gamma_0(\mathrm{H})$. In this section, our aim is to derive a one-loop algorithm, that depends only on the knowledge of prox_f and prox_g, allowing to compute numerically an element of $\mathrm{prox}_g^f(x)$ for all $x \in \mathrm{D}(\mathrm{prox}_g^f)$. We refer to Algorithm $(\mathcal{A}_1)$ in Theorem 3.3.

Moreover, if the additivity condition $\partial(f+g) = \partial f + \partial g$ is satisfied, it follows from Theorem 2.7 that Algorithm $(\mathcal{A}_1)$ is a one-loop algorithm allowing to compute numerically $\mathrm{prox}_{f+g}(x)$ for all $x \in \mathrm{H}$ with the only knowledge of prox_f and prox_g.

THEOREM 3.3. *Let f, $g \in \Gamma_0(\mathrm{H})$ and let $x \in \mathrm{D}(\mathrm{prox}_g^f)$ be fixed. Then, Algorithm $(\mathcal{A}_1)$ given by*

$$(\mathcal{A}_1) \qquad \begin{cases} y_0 \in \mathrm{H}, \\ y_{k+1} = \overline{\mathcal{T}}_{f,g}(x, y_k), \end{cases}$$

weakly converges to an element $y^* \in \mathrm{prox}_g^f(x)$. Moreover, if the additivity condition $\partial(f+g) = \partial f + \partial g$ is satisfied, it holds that $\mathrm{prox}_f(y^*) = \mathrm{prox}_{f+g}(x)$.

Proof. From Lemma 3.1, $\overline{\mathcal{T}}_{f,g}(x,\cdot)$ coincides with the composition of two firmly non-expansive operators, and thus of two non-expansive and $\frac{1}{2}$-averaged operators (see [2, Remark 4.24(iii) p.68]). Since $x \in \mathrm{D}(\mathrm{prox}_g^f)$, it follows from Proposition 3.2 and Lemma 3.1 that $\mathrm{Fix}(\mathrm{prox}_{g^* \circ L_x} \circ \mathrm{prox}_{f^*}) \neq \emptyset$. We conclude from [2, Theorem 5.22 p.82] that Algorithm $(\mathcal{A}_1)$ weakly converges to a fixed point y^* of $\overline{\mathcal{T}}_{f,g}(x,\cdot)$. From Proposition 3.2, it holds that $y^* \in \mathrm{prox}_g^f(x)$. Finally, if the additivity condition $\partial(f+g) = \partial f + \partial g$ is satisfied, we conclude that $\mathrm{prox}_f(y^*) = \mathrm{prox}_{f+g}(x)$ from Theorem 2.7. ∎

Remark 3.4. Let f, $g \in \Gamma_0(\mathrm{H})$ and let $x \in \mathrm{D}(\mathrm{prox}_g^f)$. Algorithm $(\mathcal{A}_1)$ consists in a fixed-point algorithm from the characterization given in Proposition 3.2 by

$$\mathrm{prox}_g^f(x) = \mathrm{Fix}(\overline{\mathcal{T}}_{f,g}(x,\cdot)).$$

Actually, one can easily see that Algorithm $(\mathcal{A}_1)$ also coincides with the well-known *Forward-Backward algorithm* (see [7, Section 10.3 p.191] for details) from the characterization given in Proposition 3.2 by

$$\mathrm{prox}_g^f(x) = \mathrm{argmin}\,(\mathrm{M}_{f^*} + g^* \circ L_x).$$

Indeed, we recall that M_{f^*} is differentiable with $\nabla \mathrm{M}_{f^*} = \mathrm{prox}_f$. We also refer to Section 4.1 for a brief discussion about the Forward-Backward algorithm.

Let $f,\, g \in \Gamma_0(\mathrm{H})$. As mentioned in the introduction, no proximal algorithm $x_{n+1} = \mathrm{prox}_{f+g}(x_n)$, using only the knowledge of prox_f and prox_g, has been provided in the literature. This remains a very interesting open challenge in the literature. However, we will introduce now the notion of *proximal-like algorithm* (see Definition 3.5) and we will provide in Remark 3.7 such a proximal-like algorithm $x_{n+1} = \mathrm{prox}_{f+g}(x_n)$ requiring only the knowledge of prox_f and prox_g.

DEFINITION 3.5 (Proximal-like algorithm). *Let $g \in \Gamma_0(\mathrm{H})$. An algorithm is said to be a proximal-like algorithm $x_{n+1} = \mathrm{prox}_g(x_n)$ if it can be written as*

$$\begin{cases} x_0 \in \mathrm{H}, \\ x_{n+1} = P_1(y_n^*), \\ \qquad \qquad \textit{where } y_n^* \textit{ is given by solving a weakly} \\ \qquad \qquad \textit{convergent auxiliary subalgorithm} \\ \qquad \qquad \begin{cases} y_{n,0} \in \mathrm{H}, \\ y_{n,k+1} = P_2(x_n, y_{n,k}), \end{cases} \end{cases}$$

where $P_1 : \mathrm{H} \to \mathrm{H}$ and $P_2 : \mathrm{H} \times \mathrm{H} \to \mathrm{H}$ are two given operators satisfying

$$P_1(\mathrm{Fix}(P_2(x,\cdot))) = \mathrm{prox}_g(x),$$

for all $x \in \mathrm{H}$.

Remark 3.6. In contrary to the classical proximal, Douglas-Rachford and Forward-Backward algorithms, it should be noted that a proximal-like algorithm is a two-loops algorithm.

Remark 3.7. Let $f,\, g \in \Gamma_0(\mathrm{H})$ such that $\partial(f+g) = \partial f + \partial g$. From Theorem 2.7, Proposition 3.2 and Theorem 3.3, Algorithm $(\mathcal{A}_2)$ given by

$$(\mathcal{A}_2) \qquad \begin{cases} x_0 \in \mathrm{H}, \\ x_{n+1} = \mathrm{prox}_f(y_n^*), \\ \qquad \qquad \text{where } y_n^* \text{ is given by solving the weakly} \\ \qquad \qquad \text{convergent auxiliary subalgorithm} \\ \qquad \qquad \begin{cases} y_{n,0} \in \mathrm{H}, \\ y_{n,k+1} = \overline{\mathcal{T}}_{f,g}(x_n, y_{n,k}), \end{cases} \end{cases}$$

is a *proximal-like algorithm* $x_{n+1} = \mathrm{prox}_{f+g}(x_n)$ that only requires the knowledge of prox_f and prox_g.

Remark 3.8. As mentioned in the introduction, the aim of the present theoretical paper is not to discuss numerical experiments and comparisons between numerical algorithms (this should be the topic of future works). However, in contrary to the classical Douglas-Rachford algorithm, it should be noted that Algorithm $(\mathcal{A}_2)$ is a two-loops algorithm. As a consequence, it should not be expected from Algorithm $(\mathcal{A}_2)$ better performances than the Douglas-Rachford algorithm for solving the minimization problem $\mathrm{argmin}\, f+g$.

3.3. An additional result on the Douglas-Rachford operator. Let $f,\, g \in \Gamma_0(\mathrm{H})$. It is well-known in the literature (and it can be easily proved) that

$$\mathrm{prox}_f(\mathrm{Fix}(\mathcal{T}_{f,g})) \subset \mathrm{argmin}\ f+g.$$

Our aim in this section is to prove, with the help of Theorem 2.7, that the opposite inclusion holds true under the additivity condition $\partial(f+g) = \partial f + \partial g$. To the best of our knowledge, this result is new in the literature.

LEMMA 3.9. *Let f, $g \in \Gamma_0(\mathrm{H})$. It holds that*

$$\mathrm{Fix}(\mathcal{T}_{f,g}) = \mathrm{Fix}(\mathrm{prox}_g^f \circ \mathrm{prox}_f).$$

Proof. Let $z \in \mathrm{H}$. It holds from Proposition 3.2 that

$$\begin{aligned} z \in \mathrm{Fix}(\mathcal{T}_{f,g}) &\iff z = \mathcal{T}_{f,g}(z) = \overline{\mathcal{T}}_{f,g}(\mathrm{prox}_f(z), z) \\ &\iff z \in \mathrm{Fix}(\overline{\mathcal{T}}_{f,g}(\mathrm{prox}_f(z), \cdot)) = \mathrm{prox}_g^f(\mathrm{prox}_f(z)) \\ &\iff z \in \mathrm{Fix}(\mathrm{prox}_g^f \circ \mathrm{prox}_f). \end{aligned}$$

The proof is complete. ∎

PROPOSITION 3.10. *Let f, $g \in \Gamma_0(\mathrm{H})$ such that $\partial(f+g) = \partial f + \partial g$. It holds that*

$$\mathrm{argmin}\ f + g = \mathrm{prox}_f(\mathrm{Fix}(\mathcal{T}_{f,g})).$$

Proof. Let $y \in \mathrm{Fix}(\mathcal{T}_{f,g})$. Then $y \in \mathrm{Fix}(\mathrm{prox}_g^f \circ \mathrm{prox}_f)$ from Lemma 3.9. Thus $y \in \mathrm{prox}_g^f \circ \mathrm{prox}_f(y)$. From Theorem 2.7, we get that $\mathrm{prox}_f(y) = \mathrm{prox}_{f+g}(\mathrm{prox}_f(y))$ and thus $\mathrm{prox}_f(y) \in \mathrm{argmin}\ f + g$.

Let $x \in \mathrm{argmin}\ f + g$. Since $\mathrm{D}(\mathrm{prox}_g^f) = \mathrm{H}$ from Proposition 2.4, let us consider $y \in \mathrm{prox}_g^f(x)$. From Theorem 2.7, it holds that $x = \mathrm{prox}_{f+g}(x) = \mathrm{prox}_f(y)$. Let us prove that $y \in \mathrm{Fix}(\mathcal{T}_{f,g})$. Since $y \in \mathrm{prox}_g^f(x) = \mathrm{Fix}(\overline{\mathcal{T}}_{f,g}(x, \cdot))$, we get that $y = \overline{\mathcal{T}}_{f,g}(x, y) = \overline{\mathcal{T}}_{f,g}(\mathrm{prox}_f(y), y) = \mathcal{T}_{f,g}(y)$. The proof is complete. ∎

4. Some other applications and forthcoming works.

This section can be seen as a conclusion of the paper. Its aim is to provide a glimpse of some other applications of our main result (Theorem 2.7) and to raise open questions for forthcoming works. This section is splitted into two parts.

4.1. Relations with the classical Forward-Backward operator.

Let f, $g \in \Gamma_0(\mathrm{H})$ such that g is differentiable on H. In that situation, note that the additivity condition $\partial(f+g) = \partial f + \partial g$ is satisfied from Remark 2.5, and that Proposition 2.17 implies that prox_g^f is single-valued.

In that framework, the classical *Forward-Backward operator* $\mathcal{F}_{f,g} : \mathrm{H} \to \mathrm{H}$ associated to f and g is usually defined by

$$\mathcal{F}_{f,g}(y) := \mathrm{prox}_f(y - \nabla g(y)),$$

for all $y \in \mathrm{H}$. We refer to [7, Section 10.3 p.191] for more details. Let us introduce the extension $\overline{\mathcal{F}}_{f,g} : \mathrm{H} \times \mathrm{H} \to \mathrm{H}$ defined by

$$\overline{\mathcal{F}}_{f,g}(x, y) := \mathrm{prox}_f(x - \nabla g(y)),$$

for all x, $y \in \mathrm{H}$. In particular, it holds that $\mathcal{F}_{f,g}(y) = \overline{\mathcal{F}}_{f,g}(y, y)$ for all $y \in \mathrm{H}$. The following result follows from Theorem 2.7.

PROPOSITION 4.1. *Let f, $g \in \Gamma_0(\mathrm{H})$ such that g is differentiable on* H. *Then*

$$\mathrm{prox}_{f+g}(x) = \mathrm{Fix}(\overline{\mathcal{F}}_{f,g}(x, \cdot)),$$

for all $x \in \mathrm{H}$.

Proof. Let $x \in \mathrm{H}$. Firstly, let $z = \mathrm{prox}_{f+g}(x)$ and let $y = \mathrm{prox}_g^f(x)$. In particular, we have $x = y + \nabla g(\mathrm{prox}_f(y))$. From Theorem 2.7, we get that $z = \mathrm{prox}_f(y) = \mathrm{prox}_f(x - \nabla g(\mathrm{prox}_f(y))) = \mathrm{prox}_f(x - \nabla g(z)) = \overline{\mathcal{F}}_{f,g}(x,z)$. Conversely, let $z \in \mathrm{Fix}(\overline{\mathcal{F}}_{f,g}(x,\cdot))$, that is, $z = \mathrm{prox}_f(x - \nabla g(z))$. Considering $y = x - \nabla g(z)$, we have $z = \mathrm{prox}_f(y)$ and thus $x = y + \nabla g(\mathrm{prox}_f(y))$, that is, $y = \mathrm{prox}_g^f(x)$. Finally, from Theorem 2.7, we get that $z = \mathrm{prox}_f \circ \mathrm{prox}_g^f(x) = \mathrm{prox}_{f+g}(x)$. ∎

From Proposition 4.1, we retrieve the following classical result.

PROPOSITION 4.2. *Let $f,\ g \in \Gamma_0(\mathrm{H})$ such that g is differentiable on* H. *Then*

$$\mathrm{argmin}\ f + g = \mathrm{Fix}(\mathcal{F}_{f,g}).$$

Proof. Let $x \in \mathrm{H}$. It holds that

$$\begin{aligned} x \in \mathrm{argmin}\ f + g &\Longleftrightarrow x = \mathrm{prox}_{f+g}(x) \\ &\Longleftrightarrow x \in \mathrm{Fix}(\overline{\mathcal{F}}_{f,g}(x,\cdot)) \\ &\Longleftrightarrow x = \overline{\mathcal{F}}_{f,g}(x,x) = \mathcal{F}_{f,g}(x) \\ &\Longleftrightarrow x \in \mathrm{Fix}(\mathcal{F}_{f,g}). \end{aligned}$$

The proof is complete. ∎

Let $f,\ g \in \Gamma_0(\mathrm{H})$ such that g is differentiable on H. The classical *Forward-Backward algorithm* $x_{n+1} = \mathcal{F}_{f,g}(x_n)$ is a powerful tool since it provides a one-loop algorithm, only requiring the knowledge of prox_f and ∇g, that weakly converges (under some conditions on g, see [2, Section 27.3 p.405] for details) to a fixed point of $\mathcal{F}_{f,g}$, and thus to a minimizer of $f + g$.

From Proposition 4.1, and for all $x \in \mathrm{H}$, one can consider the one-loop algorithm (potentially weakly convergent) given by

$$(\mathcal{A}_3) \qquad \begin{cases} y_0 \in \mathrm{H}, \\ y_{k+1} = \overline{\mathcal{F}}_{f,g}(x, y_k), \end{cases}$$

in order to compute numerically $\mathrm{prox}_{f+g}(x)$, with the only knowledge of prox_f and ∇g. Finally, one can also consider the two-loops algorithm

$$(\mathcal{A}_4) \qquad \begin{cases} x_0 \in \mathrm{H}, \\ x_{n+1} = y_n^*, \\ \qquad \text{where } y_n^* \text{ is given by solving the auxiliary subalgorithm} \\ \qquad \begin{cases} y_{n,0} \in \mathrm{H}, \\ y_{n,k+1} = \overline{\mathcal{F}}_{f,g}(x_n, y_{n,k}), \end{cases} \end{cases}$$

as a potential proximal-like algorithm $x_{n+1} = \mathrm{prox}_{f+g}(x_n)$, using only the knowledge of prox_f and ∇g.

Convergence proofs (under some assumptions on f and g) and numerical experiments of Algorithms $(\mathcal{A}_3)$ and $(\mathcal{A}_4)$, and eventually comparisons with other known algorithms in the literature, should be the subject of future works.

4.2. Application to sensitivity analysis for variational inequalities. As a conclusion of the present paper, we return back to our initial motivation, namely the sensitivity analysis, with respect to a nonnegative parameter $t \geq 0$, of some parameterized linear variational inequalities of second kind in a real Hilbert space H. More precisely, for all $t \geq 0$, we consider the variational inequality which consists of finding $u(t) \in \mathrm{K}$ such that

$$\langle u(t), z - u(t)\rangle + g(z) - g(u(t)) \geq \langle r(t), z - u(t)\rangle,$$

for all $z \in \mathrm{K}$, where $\mathrm{K} \subset \mathrm{H}$ is a nonempty closed and convex set of constraints, and where $g \in \Gamma_0(\mathrm{H})$ and $r : \mathbb{R}^+ \to \mathrm{H}$ are assumed to be given. Note that the above problem admits a unique solution given by

$$u(t) = \mathrm{prox}_{f+g}(r(t)),$$

where $f = \iota_{\mathrm{K}}$ is the indicator function of K.

Our aim is to provide from Theorem 2.7 a simple and compact formula for the derivative $u'(0)$ under some assumptions (see Proposition 4.3 for details). Following the idea of F. Mignot in [17] (see also [13, Theorem 2 p.620]), we first introduce the following sets

$$O_v := \{w \in \mathrm{H} \mid \exists \lambda > 0,\ \mathrm{proj}_{\mathrm{K}}(v) + \lambda w \in \mathrm{K}\} \cap [v - \mathrm{proj}_{\mathrm{K}}(v)]^{\perp},$$

$$C_v := \mathrm{cl}\Big(\{w \in \mathrm{H} \mid \exists \lambda > 0,\ \mathrm{proj}_{K}(v) + \lambda w \in \mathrm{K}\} \Big) \cap [v - \mathrm{proj}_{\mathrm{K}}(v)]^{\perp},$$

for all $v \in \mathrm{H}$, where $\perp$ denotes the classical orthogonal of a set.

Proposition 4.3. *Let $v(t) := r(t) - \nabla g(u(t))$ for all $t \in \mathbb{R}$. If the following assertions are satisfied:*

(i) *r is differentiable at $t = 0$;*
(ii) *g is twice differentiable on* H*;*
(iii) *$O_{v(0)}$ is dense in $C_{v(0)}$;*
(iv) *u is differentiable at $t = 0$;*

then the derivative $u'(0)$ is given by

$$u'(0) = \mathrm{prox}_{\varphi_f + \varphi_g}(r'(0)),$$

where $\varphi_f := \iota_{C_{v(0)}}$ and $\varphi_g(x) := \frac{1}{2}\langle \mathrm{D}^2 g(u(0))(x), x\rangle$ for all $x \in \mathrm{H}$.

Proof. Note that v is differentiable at $t = 0$ with

$$v'(0) = r'(0) - \mathrm{D}^2 g(u(0))(u'(0)).$$

Note that prox_g^f is single-valued from Proposition 2.17 and Remark 2.5. From Theorem 2.7, one can easily obtain that

$$v(t) = \mathrm{prox}_g^f(r(t)), \qquad \text{and thus} \qquad u(t) = \mathrm{prox}_f \circ \mathrm{prox}_g^f(r(t)) = \mathrm{proj}_{\mathrm{K}}(v(t)),$$

for all $t \geq 0$. Since $O_{v(0)}$ is dense in $C_{v(0)}$, we use the asymptotic development of F. Mignot [17, Theorem 2.1 p.145] and we obtain that

$$u'(0) = \mathrm{proj}_{C_{v(0)}}(v'(0)).$$

We deduce that

$$v'(0) + \mathrm{D}^2 g(u(0)) \circ \mathrm{proj}_{C_{v(0)}}(v'(0)) = r'(0).$$

Since g is convex and since $C_{v(0)}$ is a nonempty closed convex subset of H, we deduce that φ_f, $\varphi_g \in \Gamma_0(\mathrm{H})$. Moreover $\partial(\varphi_f + \varphi_g) = \partial\varphi_f + \partial\varphi_g$ from Remark 2.5 and $\mathrm{prox}_{\varphi_g}^{\varphi_f}$

is single-valued from Proposition 2.17. It also should be noted that $\nabla\varphi_g = \mathrm{D}^2 g(u(0))$. As a consequence, we have obtained that

$$v'(0) + \nabla\varphi_g \circ \mathrm{prox}_{\varphi_f}(v'(0)) = r'(0),$$

that is, $v'(0) = \mathrm{prox}_{\varphi_g}^{\varphi_f}(r'(0))$. We conclude the proof from the equality $u'(0) = \mathrm{prox}_{\varphi_f}(v'(0))$ and from Theorem 2.7. ∎

Remark 4.4. Proposition 4.3 provides an expression of $u'(0)$ in terms of the proximity operator of a sum of two closed and convex functions. Hence, it could be numerically computed from Algorithm $(\mathcal{A}_1)$, requiring the knowledge of $\mathrm{proj}_{C_{v(0)}}$ and $\mathrm{prox}_{\varphi_g}$. Alternatively, if the convergence is proved, one can also consider Algorithm $(\mathcal{A}_3)$ requiring the knowledge of $\mathrm{proj}_{C_{v(0)}}$ and $\nabla\varphi_g = \mathrm{D}^2 g(u(0))$.

Remark 4.5. The relaxations in special frameworks of the assumptions of Proposition 4.3 should be the subject of future works. In particular, it would be relevant to provide sufficient conditions on K and g ensuring that u is differentiable at $t = 0$.

The application of Proposition 4.3 in the context of some shape optimization problems with unilateral contact and friction is the subject of a forthcoming research paper (work in progress).

REFERENCES

[1] H. Attouch and H. Brezis, *Duality for the sum of convex functions in general Banach spaces*, in Aspects of mathematics and its applications, vol. 34 of North-Holland Math. Library, North-Holland, Amsterdam, 1986, pp. 125–133, https://doi.org/10.1016/S0924-6509(09)70252-1.

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

[3] A. Bensoussan and J.-L. Lions, *Inéquations quasi-variationnelles dépendant d'un paramètre*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4), 4 (1977), pp. 231–255.

[4] C. Chaux, P. L. Combettes, J.-C. Pesquet, and V. R. Wajs, *A variational formulation for frame-based inverse problems*, Inverse Problems, 23 (2007), pp. 1495–1518, https://doi.org/10.1088/0266-5611/23/4/008.

[5] P. L. Combettes and J.-C. Pesquet, *Proximal thresholding algorithm for minimization over orthonormal bases*, SIAM J. Optim., 18 (2007), pp. 1351–1376, https://doi.org/10.1137/060669498.

[6] P. L. Combettes and J.-C. Pesquet, *A proximal decomposition method for solving convex variational inverse problems*, Inverse Problems, 24 (2008), pp. 065014, 27, https://doi.org/10.1088/0266-5611/24/6/065014.

[7] P. L. Combettes and J.-C. Pesquet, *Proximal splitting methods in signal processing*, in Fixed-point algorithms for inverse problems in science and engineering, vol. 49 of Springer Optim. Appl., Springer, New York, 2011, pp. 185–212, https://doi.org/10.1007/978-1-4419-9569-8_10.

[8] P. L. Combettes and V. R. Wajs, *Signal recovery by proximal forward-backward splitting*, Multiscale Model. Simul., 4 (2005), pp. 1168–1200, https://doi.org/10.1137/050626090.

[9] J. Douglas and H. H. Rachford, *On the numerical solution of heat conduction problems in two and three space variables*, Trans. Amer. Math. Soc., 82 (1956), pp. 421–439, https://doi.org/10.2307/1993056.

[10] J. Eckstein and D. P. Bertsekas, *On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators*, Math. Programming, 55 (1992), pp. 293–318, https://doi.org/10.1007/BF01581204.

[11] I. Ekeland and R. Temam, *Analyse convexe et problèmes variationnels*, Dunod; Gauthier-Villars, Paris-Brussels-Montreal, Que., 1974. Collection Études Mathématiques.

[12] M. S. Gockenbach, B. Jadamba, A. A. Khan, C. Tammer, and B. Winkler, *Proximal methods for the elastography inverse problem of tumor identification using an equation error approach*, in Advances in variational and hemivariational inequalities, vol. 33 of Adv. Mech. Math., Springer, Cham, 2015, pp. 173–197, https://doi.org/10.1007/978-3-319-14490-0_7.

[13] A. Haraux, *How to differentiate the projection on a convex set in Hilbert space. Some applications to variational inequalities*, J. Math. Soc. Japan, 29 (1977), pp. 615–631, https://doi.org/10.2969/jmsj/02940615.

[14] J.-B. Hiriart-Urruty and C. Lemaréchal, *Fundamentals of convex analysis*, Grundlehren Text Editions, Springer-Verlag, Berlin, 2001, https://doi.org/10.1007/978-3-642-56468-0.

[15] P.-L. Lions and B. Mercier, *Splitting algorithms for the sum of two nonlinear operators*, SIAM J. Numer. Anal., 16 (1979), pp. 964–979, https://doi.org/10.1137/0716071.

[16] B. Martinet, *Détermination approchée d'un point fixe d'une application pseudo-contractante. Cas de l'application prox*, C. R. Acad. Sci. Paris Sér. A-B, 274 (1972), pp. A163–A165.

[17] F. Mignot, *Contrôle dans les inéquations variationelles elliptiques*, J. Functional Analysis, 22 (1976), pp. 130–185.

[18] J.-J. Moreau, *Proximité et dualité dans un espace hilbertien*, Bull. Soc. Math. France, 93 (1965), pp. 273–299.

[19] Y. Qiu and T. L. Magnanti, *Sensitivity analysis for variational inequalities*, Math. Oper. Res., 17 (1992), pp. 61–76, https://doi.org/10.1287/moor.17.1.61.

[20] R. T. Rockafellar, *Convex analysis*, Princeton Mathematical Series, No. 28, Princeton University Press, Princeton, N.J., 1970.

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

[22] R. T. Rockafellar, *Monotone operators and the proximal point algorithm*, SIAM J. Control Optimization, 14 (1976), pp. 877–898, https://doi.org/10.1137/0314056.

[23] A. Shapiro, *Sensitivity analysis of parameterized variational inequalities*, Math. Oper. Res., 30 (2005), pp. 109–126, https://doi.org/10.1287/moor.1040.0115.

[24] Y.-L. Yu, *On decomposing the proximal map*, Advances in Neural Information Processing Systems, 26 (2013), pp. 91–99.