



**HAL**  
open science

# A proximal variable smoothing for solving composite inclusion problems

Abdellatif Moudafi

► **To cite this version:**

Abdellatif Moudafi. A proximal variable smoothing for solving composite inclusion problems. Pure and Applied Functional Analysis, 2025, 10, pp.967-977. <hal-04992179>

**HAL Id: hal-04992179**

**<https://hal.science/hal-04992179v1>**

Submitted on 14 Mar 2025

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.



HAL Authorization

# A proximal variable smoothing for solving composite inclusion problems

Abdellatif MOUDAFI <sup>1</sup>

## Abstract

We propose a proximal variable smoothing algorithm for inclusion problem with a sum involving monotone composite operators. The proposed algorithm is designed as a forward-backward splitting algorithm with two steps: (i) a forward step with the gradient of a smoothed surrogate operator, designed with the Yosida approximate; (ii) the backward step with a resolvent operator of the remaining operator. For the proposed algorithm, we present a convergence analysis and develop three particular interesting cases, namely minimization, saddle point and equilibrium problems. Some concluding remarks are also provided.

**AMS Subject Classification.** Primary, 49J53, 65K10; Secondary, 49M37, 90C25.

**Keywords and Phrases.** Composite inclusion, Yosida approximate, Moreau envelope, variational composition, forward-backward algorithm, minimization, saddle point, equilibria.

## 1 Introduction

To begin with, let us recall the following concepts which are of common use in the context of convex and nonlinear analysis, see for example Rockafellar and Wets [21]. Throughout,  $\mathcal{G}$  is a real Hilbert space,  $\langle \cdot, \cdot \rangle$  denotes the associated scalar product and  $\| \cdot \|$  stands for the corresponding norm. An operator with domain  $dom(A)$  and range  $ran(A)$  is said to be monotone if

$$\langle u - v, x - y \rangle \geq 0, \text{ whenever } u \in A(x), v \in A(y).$$

It is said to be maximal monotone if, in addition, its graph,

$$gph A := \{(x, y) \in \mathcal{G} \times \mathcal{G} : y \in A(x)\},$$

is not properly contained in the graph of any other monotone operator. It is well-known that for each  $x \in \mathcal{G}$  and  $\mu > 0$  there is a unique  $z \in \mathcal{G}$  such that  $x \in (I + \mu A)z$ . The single-valued operator  $J_\mu^A := (I + \mu A)^{-1}$  is called the resolvent of  $A$  of parameter  $\mu$ . It is a nonexpansive mapping which is everywhere defined and is related to its Yosida approximate, namely

$$A_\mu(x) := \frac{x - J_\mu^A(x)}{\mu} \text{ by the relation } A_\mu(x) \in (J_\mu^A(x)).$$

---

<sup>1</sup>Aix-Marseille Université, Laboratoire d'Informatique et Systèmes (LIS UMR 7020 CNRS / AMU /UTLN) Marseille, France; e-mail: abdellatif.moudafi@univ-amu.fr

The latter is Lipschitz continuous with constant  $\frac{1}{\mu}$ .

Recall also that the inverse  $A^{-1}$  of  $A$  is the operator defined by  $x \in A^{-1}(y)$  if, and only if,  $y \in A(x)$ .

Finally, let  $T : \mathcal{H} \rightarrow \mathcal{G}$  be a bounded linear operator with adjoint  $T^*$ ,  $A : \mathcal{G} \rightarrow \mathcal{G}$  a maximal monotone operator,  $\mathcal{H}$  being another Hilbert space. It is easily checked that the composite operator  $T^*AT$  is monotone. This kind of operator appears, for example, in partial differential equations in divergence form. Without further conditions, however, it may fail to be maximal monotone, see for sufficient conditions [19].

Recall also that the idea of graphical convergence is to identify the operators with their graphs and to consider the Painlevé-Kuratowski convergence on them, see for example [1]. We will denote by  $g - \liminf_{\mu \rightarrow 0} C_\mu$  the mapping whose graph is the set of points  $(x, y)$  such that for every sequence  $\mu_k \rightarrow 0$  there exists a sequence  $(x_k, y_k) \rightarrow (x, y)$  with  $y_k \in C_{\mu_k}(x_k)$ . Similarly  $g - \limsup_{\mu \rightarrow 0} C_\mu = 0$  the mapping whose graph is the set of points  $(x, y)$  such that there exist a sequence  $(\mu_k) \rightarrow 0$  and  $(x_k, y_k) \rightarrow (x, y)$  with  $y_k \in C_{\mu_k}(x_k)$ . If

$$g - \liminf_{\mu \rightarrow 0} C_\mu = g - \limsup_{\mu \rightarrow 0} C_\mu,$$

one says that the family  $(C_\mu)$  graph-converges to the common limit which is denoted by  $g - \lim_{\mu \rightarrow 0} C_\mu$ .

Now, let us state the following key facts:

**Fact 1:** ([18]-Lemma 2.2.2) Let  $(a_k)$ ,  $(b_k)$  and  $(c_k)$  be three sequences of nonnegative numbers satisfying  $a_{k+1} \leq (1 + b_k)a_k + c_k$ . If  $\sum_{k=0}^{+\infty} a_k < +\infty$  and  $\sum_{k=0}^{+\infty} b_k < +\infty$ , then  $(a_k)$  is convergent.

**Fact 2:** [3, 16] Let  $S_k, S, T_k, T$  be maximal monotone operators such that  $g - \lim_{k \rightarrow +\infty} S_k = S$  and  $g - \lim_{k \rightarrow +\infty} T_k = T$ . If  $0 \in \text{int}(\text{dom}S - \text{dom}T)$ , then  $g - \lim_{k \rightarrow +\infty} (S_k + T_k) = S + T$ .

The idea of replacing  $T^*AT$  by  $T^*A_\mu T$  and taking the limit as  $\mu \rightarrow 0$ , has been already used in the proof of [21]-Theorem 12.43, where a sufficient constraint qualification condition was found in order to assure that the family  $T^*A_\mu T$  graph-converges to  $T^*AT$  (which in this case guarantees the maximality of  $T^*AT$ ). We've used it to present in [15] an algorithmic approach for the following split composite monotone problem of finding  $\bar{x}$  satisfying

$$0 \in T^*AT(\bar{x}) + B(\bar{x}).$$

This approach was very recently taken up in optimization to design prox-gradient methods that generated subsequences converging to a stationary point in the sense of zero of a general subdifferential [9, 8].

In [15], we proposed a smoothing forward-backward method by replacing  $T^*AT$  by  $C^\mu := T^*A_\mu T$ . Observe that since  $A_\mu$  is a monotone continuous mapping, the same is then true of  $C^\mu = T^*A_\mu T$ ; which guarantees the maximality. If one now takes the limit of  $C^\mu = T^*A_\mu T$  as  $\mu \rightarrow 0$ ; in the sense of graphical convergence, it turns out that one obtains a mapping that is more likely to be maximal monotone than the pointwise composition  $T^*AT$ : This limit mapping

$$C := (T^*AT)_v = g - \liminf_{\mu \rightarrow 0} T^*A_\mu T, \quad (1.0)$$

is the so-called variational composition of  $A$  and  $T$ , and we have that

$$(T^*AT)_v = g - \lim_{\mu \rightarrow 0} T^*A_\mu T,$$

whenever  $(T^*AT)_v$  is maximal monotone which is the interesting case.

**Fact 3:** [17]-Corollary 3.1. The variational composition  $(T^*AT)_v$  coincides with  $T^*AT$  when the latter is maximal monotone.

Motivated by the recent paper of Kume and Yamada [8], we are interested in the following problem of finding  $x^* \in \mathcal{H}$  such that

$$0 \in (T^*AT)_v(x^*) + B(x^*), \quad (1.1)$$

where  $A$  and  $B$  are two maximal monotone operators on two given Euclidean spaces, i.e., finite dimensional real Hilbert spaces  $\mathcal{H}, \mathcal{G}$ , respectively with  $(T^*AT)_v$  maximal monotone.

Observe that for a continuous linear operator  $T$  and a proper lower semicontinuous convex function  $f$  with  $\text{dom}(f \circ T) \neq \emptyset$ , see [17], we have

$$(T^*\partial f T)_v = \partial(f \circ T). \quad (1.2)$$

Hence, when  $A = \partial f$  and  $B = \partial g$ , problem (1.1) reduces to the problem studied in [8], namely the problem of finding  $x^* \in \mathcal{H}$  such that

$$0 \in \partial(f \circ T)(x^*) + \partial g(x^*), \quad (1.3)$$

where  $f$  and  $g$  are two proper lower semicontinuous convex functions.

A class of saddle point problems can be also written in our general model. More precisely, by considering  $A = \partial_1 K \times \partial_2(-K)$  and  $B = \partial_1 K \times \partial_2(-K)$ , problem (1.1) reduces to the problem of finding  $(x^*, y^*) \in \mathcal{H}$  such that

$$(0, 0) \in \partial(K \circ T)(x^*, y^*) + \partial L(x^*, y^*), \quad (1.4)$$

where  $K$  and  $L$  are two proper closed convex-concave functions defined on two Hilbert spaces  $\mathcal{H} = \mathcal{H}_1 \times \mathcal{H}_2$  and  $\mathcal{G} = \mathcal{G}_1 \times \mathcal{G}_2$ , respectively and  $\partial L = (\partial_1 L, -\partial_2(-L))$ .

A class of equilibrium problems can be included in our model as well. This can be reached via the associated monotone operators together with their related resolvent and Yosida mappings.

To solve Problem (1.1), we propose a forward-backward splitting-type algorithm designed with the Yosida approximate  $A_\mu$  of  $A$ . More precisely, the proposed algorithm consists of two steps: (i) the forward step with the gradient of a smoothed substitute operator  $T^*A_{\mu_k}T$  of  $T^*AT$  and (ii) the backward step with  $J_{\gamma_k}^B$  with a stepsize  $\gamma_k > 0$ . We present also an asymptotic convergence analysis of the proposed algorithm in Theorem 2.1. A particular attention is then given to the minimization, saddle point and equilibrium problems. Some concluding remarks are also stated, in particular the fact that the proposed convergence analysis may relies on a smooth approximation of a measure  $\mathcal{M}_{\gamma}^{C,B}$  as that considered in [9] and [8].

## 2 Proximal variable smoothing

In order to solve problem (1.1), we propose the following proximal variable smoothing algorithm which generates, from an arbitrary point  $x_0$ , a sequence  $(x_k)_{k \in \mathbb{N}}$  by the following rule

$$x_{k+1} = J_{\gamma_k}^B(x_k - \gamma_k C_k(x_k)), \quad (2.1)$$

where  $\gamma_k > 0$  for all  $k$  and  $C_k = T^* A_{\mu_k} T$ .

**Theorem 2.1** *Assume that problem (1.1) is consistent, i.e., its solution set  $\Gamma$  is nonempty. Choose arbitrarily  $x_0$  and let  $(\gamma_k)$  square summable verifying  $\gamma_k \leq \frac{\mu_k}{\|T\|^2}$  with  $\lim_{k \rightarrow \infty} \mu_k = 0$ .*

*Suppose that  $(x_k)_{k \in \mathbb{N}}$  is generated by algorithm (2.1) and such that  $\liminf_{k \rightarrow +\infty} \frac{x_k - x_{k+1}}{\gamma_k} = 0$ .*

*Then, the sequence  $(x_k)_{k \in \mathbb{N}}$  converges to point  $x^* \in \Gamma$ .*

*Proof:* Indeed, we successively have

$$\begin{aligned} \|x_{k+1} - x^*\|^2 &= \|J_{\gamma_k}^B(x_k - \gamma_k T^* A_{\mu_k} T x_k) - J_{\gamma_k}^B(x^* - \gamma_k T^* A T x^*)\|^2 \\ &\leq \|x_k - x^*\|^2 - 2\gamma_k \langle T^* A_{\mu_k} T x_k - T^* A T x^*, x_k - x^* \rangle + \gamma_k^2 \|T^* A_{\mu_k} T x_k - T^* A T x^*\|^2 \\ &= \|x_k - x^*\|^2 - 2\gamma_k \langle A_{\mu_k} T x_k - A T x^*, T x_k - T x^* \rangle + \gamma_k^2 \|T^* A_{\mu_k} T x_k - T^* A T x^*\|^2 \\ &\leq \|x_k - x^*\|^2 - 2\gamma_k \langle A_{\mu_k} T x_k - A T x^*, J_{\gamma_k}^B T x_k - T x^* \rangle \\ &\quad - 2\gamma_k \langle A_{\mu_k} T x_k - A T x^*, T x_k - J_{\gamma_k}^B T x^* \rangle + \gamma_k^2 \|A\|^2 \|A_{\mu_k} T x_k - A T x^*\|^2 \\ &\leq \|x_k - x^*\|^2 - 2\gamma_k \mu_k \langle A_{\mu_k} T x_k - A T x^*, A_{\gamma_k} T x_k \rangle + \gamma_k^2 \|A\|^2 \|A_{\mu_k} T x_k - A T x^*\|^2 \\ &= \|x_k - x^*\|^2 + \gamma_k^2 \|A\|^2 (-2 \langle A_{\mu_k} T x_k - A T x^*, A_{\gamma_k}^T x_k \rangle + \|A_{\mu_k} T x_k - A T x^*\|^2) \\ &\leq \|x_k - x^*\|^2 + \gamma_k^2 \|A\|^2 \|T x^*\|^2, \end{aligned}$$

where  $A(Tx^*)$  stands for any element in the set  $A(Tx^*)$ . By Fact 1, we have that

$$\lim_{k \rightarrow +\infty} \|x_k - x^*\|^2 < +\infty.$$

The latter ensures that the sequence is bounded and thus the existence of a subsequence converging to a cluster point  $\bar{x}$ .

Now note that

$$(T^* A_{\mu_{k_\nu}} T) \text{ graph-converges to } (T^* A T)_v,$$

by (1.0) and

$$(T^* A_{\mu_{k_\nu}} T + B) \text{ graph-converges to } (T^* A T)_v + B,$$

thanks to Fact 2. Note also that algorithm (2.1) can be rewritten, for a subsequence, as

$$\frac{x_{k_\nu} - x_{k_\nu+1}}{\gamma_{k_\nu}} + T^* A_{\mu_{k_\nu}} T(x_{k_\nu+1}) - T^* A_{\mu_{k_\nu}} T(x_{k_\nu}) \in (T^* A_{\mu_{k_\nu}} T + B)(x_{k_\nu+1}). \quad (2.2)$$

By passing to the limit in the last inclusion, since  $\lim_{\nu \rightarrow +\infty} \frac{x_{k_\nu} - x_{k_\nu+1}}{\gamma_{k_\nu}} = 0$  (even if a new subsequence has to be extracted) and taking into account the fact that

$$\|T^* A_{\mu_{k_\nu}} T x_{k_\nu+1} - T^* A_{\mu_{k_\nu}} T x_{k_\nu}\| \leq \frac{\|A\|^2}{\mu_{k_\nu}} \|x_{k_\nu+1} - x_{k_\nu}\| \leq \frac{\|x_{k_\nu+1} - x_{k_\nu}\|}{\gamma_{k_\nu}},$$

we finally obtain

$$0 \in (T^*AT)_v(\bar{x}) + B(\bar{x}).$$

Hence any cluster point  $\bar{x}$  of  $(x_k)_{k \in \mathbb{N}}$  belongs to  $\Gamma$  and the uniqueness of the cluster point is now classical and derives from the celebrate Opial's Lemma. This completes the proof.  $\square$

**Corollary 2.2** *When  $T^*AT$  is maximal monotone, we recover a result proposed in [15].*

### 3 Minimization problems

When  $A = \partial f$  and  $B = \partial g$ , problem (1.1) reduces to the problem studied in [8], namely the problem of finding  $x^* \in \mathcal{H}$  such that

$$0 \in \partial(f \circ T)(x^*) + \partial g(x^*), \quad (3.1)$$

where  $f$  and  $g$  are two proper lower semicontinuous convex functions defined on two Hilbert spaces  $\mathcal{H}$  and  $\mathcal{G}$ , respectively.

It is well known (see for example [21]) that  $f_\mu$  is a convex  $C^1$ -function with  $(\partial f)_\mu = \nabla f_\mu$ , replacing  $f$  by its Moreau envelope  $f_\mu$  and using subdifferential calculus rules, we obtain

$$\text{find } x^* \in \mathcal{H} \text{ such that } 0 \in \partial(f_\mu \circ T + g)(x^*),$$

which is nothing else than the optimality condition of the following minimization problem

$$\min(f_\mu \circ T + g), \quad (3.2)$$

where  $f_\mu(x) = \min_{u \in \mathcal{G}} \{h(u) + \frac{1}{2\mu} \|u - x\|^2\}$  stands for the Moreau envelope of the function  $f$  of parameter  $\mu$  and  $\partial g(x) = \{u \in \mathcal{H} : g(y) \geq g(x) + \langle u, y - x \rangle, \forall y \in \mathcal{H}\}$  means the convex subdifferential of  $g$  at  $x$ .

Remembering that  $J_\gamma^{\partial g}(y) = \text{prox}_{\gamma g} := \text{argmin}_{u \in \mathcal{G}} \{h(u) + \frac{1}{2\mu} \|u - x\|^2\}$ , in this setting algorithm (2.1) reduces to that of generating a sequence  $(x_k)_{k \in \mathbb{N}}$  by the following proximal variable smoothing rule

$$x_{k+1} = \text{prox}_{\gamma_k g}(I - \gamma_k T^* \nabla f_{\mu_k} T)(x_k). \quad (3.3)$$

A direct application of Theorem 2.1 leads to the following convergence result.

**Proposition 3.1** *Assume that Problem (3.1) is consistent. Choose arbitrarily  $x_0$  and let  $(\gamma_k)$  square summable verifying  $\gamma_k \leq \frac{\mu_k}{\|T\|^2}$  with  $\lim_{k \rightarrow +\infty} \mu_k = 0$ . Suppose that  $(x_k)_{k \in \mathbb{N}}$  is generated by the prox-gradient algorithm (3.2) and such that  $\liminf_{k \rightarrow +\infty} \mathcal{M}_{\gamma_k}^{C_k, B}(x_k) = 0$ . Then, the sequence  $(x_k)_{k \in \mathbb{N}}$  converges to point  $x^*$  that solves Problem (3.1).*

### 4 Saddle point problems

When  $A = \partial_1 K \times \partial_2(-K)$  and  $B = \partial_1 K \times \partial_2(-K)$ , problem (1.1) reduces to a saddle point problem, namely the problem of finding  $(x^*, y^*) \in \mathcal{H}$  such that

$$(0, 0) \in \partial(K \circ T)(x^*, y^*) + \partial L(x^*, y^*), \quad (4.1)$$

where  $K$  and  $L$  are two proper closed convex-concave functions defined on two Hilbert spaces  $\mathcal{H} = \mathcal{H}_1 \times \mathcal{H}_2$  and  $\mathcal{G} = \mathcal{G}_1 \times \mathcal{G}_2$ , respectively and  $\partial K = (\partial_1 K, -\partial_2(-K))$ . It was proved (see for example [2]) that the Moreau-Yosida approximate  $K_{\mu,\mu}$  of  $K$  of parameter  $\mu$  is a proper closed convex-concave  $C^1$ -function with  $A_\mu = (\nabla_1 K_{\mu,\mu} \times \nabla_2(-K_{\mu,\mu}))$ . Remember that

$$K_{\mu,\mu}(x, y) = \inf_u \sup_v \left( K(u, v) + \frac{1}{2\mu} \|u - x\|^2 - + \frac{1}{2\mu} \|v - y\|^2 \right).$$

Now, replacing  $K$  by its Moreau-Yosida envelope  $K_{\mu,\mu}$  in (4.1) and using subdifferential calculus rules, see [11]-Theorem 4, we get

$$\text{find } (x^*, y^*) \in \mathcal{H} \text{ such that } (0, 0) \in \partial(K_{\mu,\mu} \circ T + L)(x^*, y^*),$$

with is nothing else than the optimality condition of the following saddle point problem

$$\min_x \max_y (K_{\mu,\mu} \circ T + L)(x, y), \quad (4.2)$$

in other terms: find  $(x_\mu^*, y_\mu^*) \in \mathcal{H}$  such that  $\forall (x, y) \in \mathcal{H}$ , we have

$$K_{\mu,\mu}T(x_\mu^*, y) + L(x_\mu^*, y) \leq K_{\mu,\mu}T(x_\mu^*, y_\mu^*) + L(x_\mu^*, y_\mu^*) \leq K_{\mu,\mu}T(x, y_\mu^*) + L(x, y_\mu^*). \quad (4.3)$$

Taking in to account the fact that

$$J_\gamma^A(x, y) = \text{prox}_{\gamma A}(x, y) = \arg \min_u \max_v \left( K(u, v) + \frac{1}{2\mu} \|u - x\|^2 - + \frac{1}{2\mu} \|v - y\|^2 \right),$$

algorithm (2.1) reduces to that of generating a sequence  $(x_k, y_k)_k$  by the following rule

$$(x_{k+1}, y_{k+1}) = \text{prox}_{\gamma_k L} \left( I - \gamma_k T^* (\nabla_1 K_{\mu_k, \mu_k} \times \nabla_2(-K_{\mu_k, \mu_k})) T \right) (x_k, y_k). \quad (4.4)$$

Theorem 2.1 leads to the convergence result below.

**Proposition 4.1** *Assume that problem (4.1) is consistent. Choose arbitrarily  $(x_0, y_0)$  and let  $(\gamma_k)$  square summable verifying  $\gamma_k \leq \frac{\mu_k}{\|T\|^2}$  with  $\lim_{k \rightarrow +\infty} \mu_k = 0$ . Suppose that  $(x_k, y_k)_{k \in \mathbb{N}}$  is generated by the prox-gradient algorithm (4.4) and such that  $\liminf_{k \rightarrow +\infty} \mathcal{M}_{\gamma_k}^{C, B}(x_k) = 0$ . Then,  $(x_k, y_k)_{k \in \mathbb{N}}$  converges to point  $(x^*, y^*)$  that solves Problem (4.1).*

## 5 Equilibrium problems

When  $A = A^F$  and  $B = A^G$ , problem (1.1) reduces to the equilibrium problem of finding  $x^* \in C \subset \mathcal{H}$  such that

$$0 \in A^{F \circ T}(x^*) + A^G(x^*), \quad (5.1)$$

where  $F, G$  are two bifunctions defined on  $D \times D, C \times C$  with  $C$  and  $D$  two nonempty closed convex sets of two Hilbert spaces  $\mathcal{H}$  and  $\mathcal{G}$ , respectively.  $F \circ T$  being the composition of  $F$  by  $T$  with respect to the first variable and  $F, G$  verifying the usual conditions for a given bifunction  $L$ :

- (A1)  $L(x, x) = 0$  for all  $x, y \in C$ ;
- (A2)  $L$  is monotone  $L(x, y) + L(y, x) \leq 0$  for all  $x, y \in C$ ;
- (A3)  $\limsup_{t \downarrow 0} L(tz + (1-t)x, y) \leq F(x, y)$  for any  $x, y, z \in C$ ;
- (A4) for each  $x \in C, y \mapsto L(x, y)$  is convex and lower-semicontinuous.

Recall also that

$$A^L(x) = \{v \in C; L(x, y) \geq \langle v, y - x \rangle \forall y \in C\} \text{ if } x \in C, A^F(x) = \emptyset \text{ if } x \notin C,$$

is the associated maximal monotone operator. Its related resolvent mapping,  $J_\mu^L$ , appeared first in [6] which properties as well as those of its approximate Yosida,  $A_\mu^F := \frac{I - J_\mu^L}{\mu}$  were given first in [12, 14]. Furthermore, we can easily check that we have the following nice formula

$$A_\mu^F = A^{F_\mu},$$

where  $F_\mu$  stands for the Moreau envelope of  $F$  with respect to the first variable.

By replacing  $F$  by its Moreau envelope  $F_\mu$  and using subdifferential calculus rules, we obtain

$$\text{find } x^* \in C \text{ such that } 0 \in A^{F_\mu \circ T + G}(x^*),$$

which is nothing else than the optimality condition of the following equilibrium problem

$$F_\mu(T(x^*), y) + G(x^*, y) \geq 0 \quad \forall y \in C, \quad (5.2)$$

where  $F_\mu(x, y) = \inf_{u \in D} \{F(u, y) + \frac{1}{2\mu} \|u - x\|^2\}$  stands for the Moreau envelope of the function  $F$  of parameter  $\mu$  with respect to the first variable.

In this setting algorithm (2.1) reduces to that of generating a sequence  $(x_k)_{k \in \mathbb{N}}$  by the following proximal variable smoothing rule

$$x_{k+1} = J_{\gamma_k}^G(I - \gamma_k T^* A^{F_{\mu_k}} T)(x_k) = J_{\gamma_k}^G(I - \gamma_k T^* A_{\mu_k}^F T)(x_k). \quad (5.3)$$

Remembering that  $J_\gamma^G(x)$  is characterized by

$$G(J_\gamma^G(x), y) + \gamma^{-1} \langle J_\gamma^G(x) - x, y - J_\gamma^G(x) \rangle \geq 0,$$

setting  $z_k = x_k - \gamma_k T^* A_{\mu_k}^F T(x_k)$  together with (5.3) lead to  $x_{k+1}$  is the unique solution of

$$G(x_{k+1}, y) + \gamma_k^{-1} \langle x_{k+1} - z_k, y - x_{k+1} \rangle \geq 0 \quad \forall y \in C.$$

A direct application of Theorem 2.1 leads to the following convergence result.

**Proposition 5.1** *Assume that Problem (5.1) is consistent. Choose arbitrarily  $x_0 \in C$  and let  $(\gamma_k)$  square summable verifying  $\gamma_k \leq \frac{\mu_k}{\|T\|^2}$  with  $\lim_{k \rightarrow +\infty} \mu_k = 0$ . Suppose that  $(x_k)_{k \in \mathbb{N}}$  is generated by the prox-gradient algorithm (5.3) and such that  $\liminf_{k \rightarrow +\infty} \mathcal{M}_{\gamma_k}^{C, B}(x_k) = 0$ . Then, the sequence  $(x_k)_{k \in \mathbb{N}}$  converges to point  $x^*$  that solves Problem (5.1).*

## 6 Concluding remarks

In what follows, we provide some comments and potential extensions.

- A connection between the variational composition and the variational sum of two maximal monotone operators  $A$  and  $B$ , namely  $g - \liminf_{\lambda, \mu \rightarrow 0; \lambda\mu \neq 0} A_\lambda + B_\mu$  was highlighted in [17]. More precisely, in  $\mathcal{H} \times \mathcal{H}$ , by setting

$$Tx = (x, x) \text{ and } \mathbf{A}(x_1, x_2) = A(x_1) \times B(x_2),$$

we obtain

$$(T^* \mathbf{A} T)_v = g - \liminf_{\mu \rightarrow 0} A_\mu + B_\mu = A +_v B,$$

whenever the latter is maximal monotone (which is the interesting case).

We can thus handle the example given in [10]. Indeed, let us consider in  $\mathbb{R}^3$  the problem of finding the closest point to the origin in the intersection of a two-dimensional subspace  $L$  and a cylinder  $C$  lying on that subspace, namely

$$\min_{x \in L \cap C} \frac{1}{2} \|x\|^2. \quad (6.1)$$

If we want to achieve a decomposition of the constraints, we can split them in the following way: find  $x^*$  verifying

$$0 \in A(x^*) + B(x^*),$$

where  $A = I + \partial i_L$  and  $B = \partial i_C$ . Then, we get that  $\text{dom}(A + B) = L \cap C$ . In our setting, with  $B = 0$ ,  $Tx = (x, x)$  and  $\mathbf{A}(x_1, x_2) = A(x_1) \times B(x_2)$ , algorithm (2.1) reduces to the following parallel projection scheme

$$x_{k+1} = \left(1 - \frac{2\gamma}{\mu_k}\right)x_k + \frac{\gamma}{\mu_k} \text{Proj}_L\left(\frac{x_k}{1 + \mu_k}\right) + \frac{\gamma}{\mu_k} \text{Proj}_C(x_k). \quad (6.2)$$

Thanks to our convergence result, the generated sequence converges to a solution of (5.1), as was the case in [10] by means of the alternating projection algorithm

$$x_{k+1} = \text{Proj}_C\left(\text{Proj}_L\left(\frac{x_k}{1 + \mu}\right)\right), \quad (6.3)$$

and by making  $\mu$  tend towards 0.

In [17]-Section 6, an example was given of a problem for which the constraint qualification fails, but where the variational composition can be computed.

- The result can be extended to  $\eta$ -hypomonotone operators  $A$  with  $\mu \in (0, \eta^{-1})$ . Indeed, its Yosida approximate of index  $\mu$  is Lipschitz continuous with constant  $\max(\frac{1}{\mu}, \frac{\eta}{1 - \mu\eta})$  and when  $A = \partial f$ , the function  $f$  is weakly convex which is the case in [7, 9, 8].
- The convergence results still hold true in infinite Hilbert spaces by replacing the convergence by weak convergence.

- In order to measure at a given  $\bar{x}$  with  $\gamma > 0$  an achievement level of (1.1), we may consider

$$\mathcal{M}_\gamma^{C,B}(\bar{x}) = \inf\{\gamma^{-1}\|\bar{x} - J_\gamma^B(\bar{x} - \gamma v)\|, v \in C(\bar{x})\}, \quad (6.4)$$

with  $C := (T^*AT)_v$ , this measure was employed in [9] and [8] for finding a stationary point in optimization setting.

$\mathcal{M}_\gamma^{C,B}$  works as a stationarity measure at  $\bar{x}$  for Problem (1.1) because

$$\forall \gamma > 0 \quad 0 \in (T^*AT)_v(\bar{x}) + B(\bar{x}) \Leftrightarrow \mathcal{M}_\gamma^{C,B}(\bar{x}) = 0. \quad (6.5)$$

Thus, finding a solution  $x^* \in \Gamma$  is equivalent to finding  $x^*$  such that  $\mathcal{M}_\gamma^{C,B}(x^*) = 0$ . From a viewpoint of approximating a solution  $x^*$  with an iterative update of estimates  $(x_k)_{k \in \mathbb{N}}$  of  $x^*$ , a measure is desired to have a lower semicontinuous property which is the case of  $\mathcal{M}_\gamma^{C,B}$  defined in (5.4). The latter can be approximated with a smoothed substitute operator  $T^*A_\mu T$  of  $T^*AT$ .

Observe that for arbitrarily given  $\bar{x}$  and  $(x_k)_{k \in \mathbb{N}}$  such that  $\lim_{n \rightarrow +\infty} x_k = \bar{x}$ , we have

$$\forall \gamma > 0 \quad \liminf_{k \rightarrow +\infty} \mathcal{M}_\gamma^{C,B}(x_k) \geq \mathcal{M}_\gamma^{C,B}(\bar{x}).$$

Furthermore, Let  $\mu_k$  such that  $\lim_{k \rightarrow +\infty} \mu_k = 0$  and  $C_k := T^*A_{\mu_k}T$ . Then,

$$\forall \gamma > 0 \quad \liminf_{n \rightarrow +\infty} \mathcal{M}_\gamma^{C_k,B}(x_k) \geq \mathcal{M}_\gamma^{C,B}(\bar{x}),$$

because the graph convergence of the sequence  $(C_k)_k$  to  $C$  is equivalent to the pointwise convergence of the corresponding resolvent operators, see [1].

Finally, if  $\liminf_{k \rightarrow +\infty} \mathcal{M}_\gamma^{C_k,B}(x_k) = 0$  with  $\gamma > 0$ , then  $\bar{x}$  solves problem (1.1). Hence finding a solution  $x^*$  of Problem (1.1) is reduced to finding a convergent sequence  $(x_k)_{k \in \mathbb{N}}$  such that

$$\liminf_{k \rightarrow +\infty} \mathcal{M}_\gamma^{C_k,B}(x_k) = 0.$$

An open question: Can we remove the assumption  $\liminf_{k \rightarrow +\infty} \mathcal{M}_\gamma^{C_k,B}(x_k) = 0$  in Theorem 2.1 at least in the convex-concave case, namely in Proposition 4.1 and even in Theorem 2.1, since this is the case in optimization setting, see [8]. It was also shown that a subsequence can be chosen such that  $\liminf_{\nu \rightarrow +\infty} \mathcal{M}_\gamma^{f_{k_\nu}, g}(x_{k_\nu}) = 0$  and that every cluster point  $\bar{x}$  is a stationary point of the minimization problem  $\min(f \circ T + g)$ .

## References

- [1] H. Attouch, Variational convergence for functions and operators, Pitman Advanced Publishing Program, Applicable Mathematics Series, 1984.
- [2] Attouch, H., Aze, D., Wets, R.J.-B. (1986). On Continuity Properties of the Partial Legendre-Fenchel Transform: Convergence of Sequences of Augmented Lagrangian Functions, Moreau-Yosida Approximates and Subdifferential Operators. IIASA Collaborative Paper. IIASA, Laxenburg, Austria: CP-86-016.

- [3] H. Attouch, H. Riahi, M. Théra, Somme Ponctuelle D'opérateurs Maximaux Monotones, *Serdica Mathematical Journal* (1996) Volume: 22, Issue: 3, 267-292.
- [4] D. Azé, (1988). Rate of convergence for the saddle points of convex-concave functions. In: Hoffmann, KH., Zowe, J., Hiriart-Urruty, JB., Lemarechal, C. (eds) *Trends in Mathematical Optimization*. International Series of Numerical Mathematics/Internationale Schriftenreihe zur Numerischen Mathematik/Série internationale d'Analyse numérique, vol 84. Birkhäuser Basel.
- [5] H. H. Bauschke, P. L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, second edn. CMS Books in Mathematics. Springer, Cham, 2017.
- [6] E. Blum and W. Oettli, From optimization and variational inequalities to equilibrium problems, *Math. Stud.* **63** (1994), 123-145.
- [7] A. Bohm and S. J. Wright, Variable smoothing for weakly convex composite functions, *Journal of Optimization Theory and Applications*, vol. 188, no. 3, 628-649, 2021.
- [8] K. Kume, I. Yamada, A Proximal Variable Smoothing for Nonsmooth Minimization Involving Weakly Convex Composite with MIMO Application. ArXiv:2409.10934v2, 27 Sep 2024.
- [9] T. Liu, F. Xia, Proximal variable smoothing method for three-composite nonconvex nonsmooth minimization with a linear operator. *Numer. Algor.* **96**, 237-266 (2024).
- [10] P. Mahey, Pham Dinh Tao, Partial regularization of the sum of two maximal monotone operators, *M2AN - Modélisation mathématique et analyse numérique*, tome 27, no 3 (1993), 375-392.
- [11] L. McLinden, Dual Operations On Saddle Functions, *Transactions of the American Mathematical Society*, Volume 179, May 1973.
- [12] A. Moudafi, Mixed equilibrium problems: Sensitivity analysis and algorithmic aspect, *Computers & mathematics with applications Journal*, (2002), Vol. 44, 1099-1108.
- [13] A. Moudafi, M. Théra, Finding the zero for the sum of two maximal monotone operators, *Journal Optimization Theory Applications*, vol. 94, N2, (1997), 425-448.
- [14] A. Moudafi, M. Théra, Proximal and dynamical approaches to equilibrium problems, *Lecture Notes in Economics and Mathematical Systems*, 477, (1999), 187-201, Springer.
- [15] A. Moudafi, Y. Shehu, About a regularization of split monotone inclusions with composite operators, *PanAmerican Mathematical Journal* Volume 24(2014), Number 3, 82-90.
- [16] T. Pennanen, Dualization of generalized equations of maximal monotone type, *SIAM J. Optim.* **10** (2000), 809-835.
- [17] T. Pennanen, J. Revalski and M. Théra, Variational composition of a monotone operator and a linear mapping with applications to elliptic PDEs with singular coefficients. *Journal of Functional Analysis* **198** (2003), 84-105.
- [18] B. T. Polyak, *Introduction to Optimization*. Optimization Software Inc., New York, 1987.

- [19] S.M. Robinson, Composition duality and maximal monotonicity, *Math. Programing Ser. A* 85 (1999) 1-13.
- [20] R. T. Rockafellar, Monotone operators associated with saddle-functions and minimax problems, in: *Non- linear Functional Analysis, Part 1*, (F. E. Browder, ed.), 241-250. AMS, Providence, RI, 1970.
- [21] R. T. Rockafellar and R. J-B Wets, *Variational Analysis*, Springer-Verlag, 1998.