ALTERNATING PROXIMAL ALGORITHMS FOR CONSTRAINED VARIATIONAL INEQUALITIES. APPLICATION TO DOMAIN DECOMPOSITION FOR PDE'S

H. ATTOUCH, A. CABOT, P. FRANKEL, AND J. PEYPOUQUET

ABSTRACT. Let $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ be real Hilbert spaces, let $f : \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$, $g : \mathcal{Y} \to \mathbb{R} \cup \{+\infty\}$ be closed convex functions and let $A : \mathcal{X} \to \mathcal{Z}$, $B : \mathcal{Y} \to \mathcal{Z}$ be linear continuous operators. Let us consider the constrained minimization problem

$$(\mathcal{P}) \qquad \min\{f(x) + g(y) : \quad Ax = By\}.$$

Given a sequence (γ_n) which tends toward 0 as $n \to +\infty$, we study the following alternating proximal algorithm

$$(\mathcal{A}) \quad \begin{cases} x_{n+1} = \operatorname{argmin}\left\{\gamma_{n+1} f(\zeta) + \frac{1}{2}\|A\zeta - By_n\|^2_{\mathcal{Z}} + \frac{\alpha}{2}\|\zeta - x_n\|^2_{\mathcal{X}};\ \zeta \in \mathcal{X}\right\} \\ y_{n+1} = \operatorname{argmin}\left\{\gamma_{n+1} g(\eta) + \frac{1}{2}\|Ax_{n+1} - B\eta\|^2_{\mathcal{Z}} + \frac{\nu}{2}\|\eta - y_n\|^2_{\mathcal{Y}};\ \eta \in \mathcal{Y}\right\}, \end{cases}$$

where α and ν are positive parameters. It is shown that if the sequence (γ_n) tends *moderately slowly* toward 0, then the iterates of $(\mathcal{A})$ weakly converge toward a solution of $(\mathcal{P})$. The study is extended to the setting of maximal monotone operators, for which a general ergodic convergence result is obtained. Applications are given in the area of domain decomposition for PDE's.

1. INTRODUCTION

Let $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ be real Hilbert spaces respectively endowed with the scalar products $\langle .,. \rangle_{\mathcal{X}}$, $\langle .,. \rangle_{\mathcal{Y}}$ and $\langle .,. \rangle_{\mathcal{Z}}$ and the corresponding norms. Let $f : \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$, $g : \mathcal{Y} \to \mathbb{R} \cup \{+\infty\}$ be closed convex proper functions and let $A : \mathcal{X} \to \mathcal{Z}$, $B : \mathcal{Y} \to \mathcal{Z}$ be linear continuous operators. In this study, our aim is to solve convex structured minimization problems of the form

$$(\mathcal{P}) \qquad \min\{f(x) + g(y) : \quad Ax = By\}.$$

In order to find a point that minimizes the map $(x, y) \mapsto \Phi(x, y) = f(x) + g(y)$ on the subspace $\{(x, y) \in \mathcal{X} \times \mathcal{Y}, \quad Ax = By\}$ we propose the following alternating algorithm:

$$(\mathcal{A}) \quad \begin{cases} x_{n+1} = \operatorname{argmin}\left\{\gamma_{n+1} f(\zeta) + \frac{1}{2}\|A\zeta - By_n\|^2_{\mathcal{Z}} + \frac{\alpha}{2}\|\zeta - x_n\|^2_{\mathcal{X}};\ \zeta \in \mathcal{X}\right\} \\ y_{n+1} = \operatorname{argmin}\left\{\gamma_{n+1} g(\eta) + \frac{1}{2}\|Ax_{n+1} - B\eta\|^2_{\mathcal{Z}} + \frac{\nu}{2}\|\eta - y_n\|^2_{\mathcal{Y}};\ \eta \in \mathcal{Y}\right\}, \end{cases}$$

Date: February 13, 2010.

1991 *Mathematics Subject Classification.* 65K05, 65K10, 49J40, 90C25.

Key words and phrases. Convex minimization, alternating minimization, proximal algorithm, variational inequalities, monotone inclusions, domain decomposition for PDE's.

With the support of the French ANR grant ANR-08-BLAN-0294-03. J. Peypouquet was partly supported by FONDECYT grant 11090023 and Basal Proyect, CMM, Universidad de Chile.

where α, ν are positive real numbers and (γ_n) is a positive sequence that tends[1] toward 0 as $n \to +\infty$. Due to the structured character of the objective function $\Phi(x,y) = f(x) + g(y)$, alternating algorithms imply a reduction on the size of the subproblems to be solved at each iteration. Our particular choice of $(\mathcal{A})$ is based on the following ideas:

a) Alternating algorithms with costs-to-move. Consider the convex function $\Phi_\gamma : \mathcal{X} \times \mathcal{Y} \to \mathbb{R} \cup \{+\infty\}$ defined by

$$\Phi_\gamma(x,y) = f(x) + g(y) + \frac{1}{2\gamma}\|Ax - By\|^2_{\mathcal{Z}},$$

where γ is a positive real parameter. The minimization of the function Φ_γ is studied in [12], where the authors introduce the alternating algorithm with costs-to-move

$$\begin{cases} x_{n+1} = \operatorname{argmin}\{f(\zeta) + \frac{1}{2\gamma}\|A\zeta - By_n\|^2_{\mathcal{Z}} + \frac{\alpha}{2}\|\zeta - x_n\|^2_{\mathcal{X}}; \quad \zeta \in \mathcal{X}\} \\ y_{n+1} = \operatorname{argmin}\{g(\eta) + \frac{1}{2\gamma}\|Ax_{n+1} - B\eta\|^2_{\mathcal{Z}} + \frac{\nu}{2}\|\eta - y_n\|^2_{\mathcal{Y}}; \quad \eta \in \mathcal{Y}\}, \end{cases}$$

α and ν being positive coefficients. If $\operatorname{argmin}\Phi_\gamma \neq \emptyset$, it is shown in [6] that the sequence (x_n, y_n) converges weakly toward a minimum of Φ_γ. The framework of [6, 12] extends the one of [1, 17] from the strong coupled problem to the weak coupled problem with costs-to-change. More precisely, $Q(x,y) = \|x - y\|^2_{\mathcal{Z}}$ is a strong coupling function with $\mathcal{X} = \mathcal{Y} = \mathcal{Z}$ and $A = B = \mathcal{I}$ while $Q(x,y) = \|Ax - By\|^2_{\mathcal{Z}}$ is now a weak coupling function which allows for asymmetric and partial relations between the variables x and y. The interest of the weak coupling term is to cover many situations, ranging from decomposition methods for PDE's to applications in game theory. In decision sciences, the term $Q(x,y) = \|Ax - By\|^2_{\mathcal{Z}}$ allows to consider agents who interplay, only via some components of their decision variables. For further details, the interested reader is referred to [6].

Observing that problem $(\mathcal{P})$ corresponds formally to the minimization of the function Φ_γ with $\gamma \to 0$, it is natural to consider a vanishing sequence (γ_n) in algorithm $(\mathcal{A})$.

b) Prox-penalization methods. Setting $\Psi(x,y) = \frac{1}{2}\|Ax - By\|^2_{\mathcal{Z}}$ and $\mathbf{x} = (x,y) \in \mathcal{X} \times \mathcal{Y} = \boldsymbol{\mathcal{X}}$, we can rewrite problem $(\mathcal{P})$ as

$$\min\{\ \Phi(\mathbf{x}) : \ \mathbf{x} \in \operatorname{argmin}\Psi\ \}.$$

This situation is studied in [11, 21], where the authors use a diagonal proximal point algorithm combined with a penalization scheme. This kind of technique can be traced back to the pioneering work [14]. The algorithm of [11, 21] applied to our setting reads as

$$(\mathcal{A}') \qquad \mathbf{x}_{n+1} \in \operatorname{argmin}\left\{\gamma_n\, \Phi(\mathbf{x}) + \Psi(\mathbf{x}) + \frac{1}{2}\|\mathbf{x} - \mathbf{x}_n\|^2_{\boldsymbol{\mathcal{X}}}\right\}.$$

[1] In another direction, algorithm $(\mathcal{A})$ has been recently studied in [22] in the case of a sequence (γ_n) increasing toward $+\infty$ as $n \to +\infty$.

Under suitable conditions on the sequence (γ_n), it is shown in [11, 21] that the iterates of algorithm $(\mathcal{A}')$ converge weakly to a solution of $(\mathcal{P})$.

These ideas lead us to the formulation of algorithm $(\mathcal{A})$, which has the following distinctive marks. First, it uses the structured character of the objective function to reduce the size of the subproblem solved at each iteration. Second, it combines proximal iterations with a penalization scheme in a simple way, meaning that no new nonlinearities are introduced by the latter, unlike most penalization procedures available in the literature. Consider, for instance, the functions θ described in [23] (see also the references therein).

The main result of the paper asserts that if the solution set is nonempty and if (γ_n) tends *moderately slowly* toward 0, then the iterates of $(\mathcal{A})$ weakly converge toward a solution of $(\mathcal{P})$. When the space $R(A) + R(B)$ is closed in $\mathcal{Z}$, the above condition on (γ_n) is satisfied if the sequence $\left(\frac{1}{\gamma_{n+1}} - \frac{1}{\gamma_n}\right)$ is bounded from above and if $(\gamma_n) \in l^2$.

We apply our abstract results to the framework of splitting methods for PDE's. For that purpose, we consider a domain $\Omega \subset \mathbb{R}^N$ that can be decomposed into two non overlapping subdomains Ω_1, Ω_2 with a common interface Γ. The functional spaces are $\mathcal{X} = H^1(\Omega_1)$, $\mathcal{Y} = H^1(\Omega_2)$ and $\mathcal{Z} = L^2(\Gamma)$, the operators $A : \mathcal{X} \to \mathcal{Z}$ and $B : \mathcal{Y} \to \mathcal{Z}$ being respectively the trace operators on Γ. The term $Au - Bv$ corresponds to the jump of the map $w = \begin{cases} u \text{ on } \Omega_1 \\ v \text{ on } \Omega_2 \end{cases}$ through the interface Γ. It is shown that algorithm $(\mathcal{A})$ allows to solve some given boundary value problem on Ω by solving separately mixed Dirichlet-Neumann problems on Ω_1 and Ω_2.

Finally observe that by writing down the optimality conditions satisfied by the iterates of algorithm $(\mathcal{A})$, we obtain

$$\left\{ \begin{array}{lllllll} 0 & \in & \gamma_{n+1}\,\partial f(x_{n+1}) & + & A^*(Ax_{n+1} - By_n) & + & \alpha(x_{n+1} - x_n) \\ 0 & \in & \gamma_{n+1}\,\partial g(y_{n+1}) & - & B^*(Ax_{n+1} - By_{n+1}) & + & \nu(y_{n+1} - y_n). \end{array} \right.$$

This suggests to extend the previous study to the framework of maximal monotone operators, by replacing respectively the subdifferential operators ∂f and ∂g with two maximal monotone operators M and N. Indeed, in this more general setting we are able to prove the convergence of the sequence of weighted averages.

The paper is organized as follows. Section 2 is devoted to fix the general setting and notations that are used throughout the paper. In section 3, we prove a general result of weak ergodic convergence for the iterates of $(\mathcal{A})$ in a maximal monotone setting. The key conditions are the closedness of the space $R(A) + R(B)$ and the assumption $(\gamma_n) \in l^2 \setminus l^1$. The subdifferential case is analyzed in section 4, where we establish a result of weak convergence toward a solution of $(\mathcal{P})$. Section 5 presents further convergence results for the strongly coupled problem without cost-to-move. Finally, the applications to domain decomposition for PDE's are illustrated in section 6.

2. General setting and notations

We recall that $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ are real Hilbert spaces respectively endowed with the scalar products $\langle ., . \rangle_{\mathcal{X}}$, $\langle ., . \rangle_{\mathcal{Y}}$, $\langle ., . \rangle_{\mathcal{Z}}$ and the corresponding norms. Let $M : \mathcal{X} \rightrightarrows \mathcal{X}$, $N : \mathcal{Y} \rightrightarrows \mathcal{Y}$ be maximal monotone operators such that $\operatorname{dom} M \neq \emptyset$, $\operatorname{dom} N \neq \emptyset$. Let $A : \mathcal{X} \to \mathcal{Z}$, $B : \mathcal{Y} \to \mathcal{Z}$ be linear continuous operators with adjoints $A^* : \mathcal{Z} \to \mathcal{X}$ and $B^* : \mathcal{Z} \to \mathcal{Y}$. Let (γ_n) be a positive sequence such that $\lim\limits_{n\to+\infty} \gamma_n = 0$. Given positive coefficients $\alpha, \nu > 0$ and initial data $(x_0, y_0) \in \mathcal{X} \times \mathcal{Y}$, let us consider the alternating proximal algorithm defined implicitly by

$$(\mathcal{A}) \qquad \left\{ \begin{array}{lllllll} 0 & \in & \gamma_{n+1}\, Mx_{n+1} & + & A^*(Ax_{n+1} - By_n) & + & \alpha(x_{n+1} - x_n) \\ 0 & \in & \gamma_{n+1}\, Ny_{n+1} & - & B^*(Ax_{n+1} - By_{n+1}) & + & \nu(y_{n+1} - y_n). \end{array} \right.$$

Observe that the linear continuous operator A^*A is maximal monotone, hence the operator $\gamma_{n+1}\, M + A^*A$ is also maximal monotone, see for example [18]. Therefore the iterate x_{n+1} is uniquely defined by Minty's theorem. The same holds true for the iterate y_{n+1}.

Remark 2.1 (Strong coupling without cost-to-move). Assume that $\mathcal{X} = \mathcal{Y} = \mathcal{Z}$ and that $A = B = \mathcal{I}$, which corresponds to a situation of strong coupling. In this case, algorithm $(\mathcal{A})$ is well-defined even if $\alpha = \nu = 0$. We denote by $(\mathcal{A}_0)$ the corresponding algorithm

$$(\mathcal{A}_0) \qquad \left\{ \begin{array}{lll} 0 & \in & \gamma_{n+1}\, Mx_{n+1} + x_{n+1} - y_n \\ 0 & \in & \gamma_{n+1}\, Ny_{n+1} + y_{n+1} - x_{n+1}, \end{array} \right.$$

that can be equivalently rewritten as

$$\left\{ \begin{array}{l} x_{n+1} = (I + \gamma_{n+1}\, M)^{-1}\, y_n \\ y_{n+1} = (I + \gamma_{n+1}\, N)^{-1}\, x_{n+1}. \end{array} \right.$$

It ensues that the sequences (x_n) and (y_n) satisfy the following recurrence formulae

$$x_{n+1} = (I + \gamma_{n+1}\, M)^{-1}(I + \gamma_n\, N)^{-1}\, x_n, \quad y_{n+1} = (I + \gamma_{n+1}\, N)^{-1}(I + \gamma_{n+1}\, M)^{-1}\, y_n.$$

This scheme consisting of a double backward step has been previously studied by Passty [30]. Algorithm $(\mathcal{A})$ can be viewed as an extension of iteration $(\mathcal{A}_0)$, so that our present paper appears as a continuation of the seminal work [30].

Let $\boldsymbol{\mathcal{X}} = \mathcal{X} \times \mathcal{Y}$ and denote by $\boldsymbol{\mathcal{V}}$ the closed subspace $\{(x, y) \in \boldsymbol{\mathcal{X}},\ Ax = By\}$. The normal cone operator $N_{\boldsymbol{\mathcal{V}}}$ takes the constant value $N_{\boldsymbol{\mathcal{V}}} \equiv \boldsymbol{\mathcal{V}}^{\perp}$ on its domain $\boldsymbol{\mathcal{V}}$. Setting $\mathbf{x} = (x, y)$, define the monotone operators $\mathbf{M} : \boldsymbol{\mathcal{X}} \rightrightarrows \boldsymbol{\mathcal{X}}$ and $\mathbf{T} : \boldsymbol{\mathcal{X}} \rightrightarrows \boldsymbol{\mathcal{X}}$ respectively by

$$\mathbf{Mx} = (Mx, Ny)$$

and

$$\mathbf{Tx} = \mathbf{Mx} + N_{\boldsymbol{\mathcal{V}}}(\mathbf{x}) = \left\{ \begin{array}{ll} \mathbf{Mx} + \boldsymbol{\mathcal{V}}^{\perp} & \text{if} \quad \mathbf{x} \in \boldsymbol{\mathcal{V}} \\ \emptyset & \text{if} \quad \mathbf{x} \notin \boldsymbol{\mathcal{V}}. \end{array} \right.$$

We denote by $\boldsymbol{\mathcal{S}} = \mathbf{T}^{-1}0$ the null set of $\mathbf{T}$. It is also convenient to define the bounded linear operator

$$\begin{array}{rcccc} \mathbf{A} & : & \boldsymbol{\mathcal{X}} & \to & \mathcal{Z} \\ & & (x, y) & \mapsto & Ax - By, \end{array}$$

and the map

$$\begin{array}{rccc} \Psi & : & \boldsymbol{\mathcal{X}} & \to & \mathbb{R} \\ & & (x,y) & \mapsto & \frac{1}{2}\|Ax - By\|^2_{\mathcal{Z}}. \end{array}$$

Recall that the Fenchel conjugate $\Psi^* : \boldsymbol{\mathcal{X}} \to \mathbb{R} \cup \{+\infty\}$ of the map Ψ is defined by $\Psi^*(\mathbf{p}) = \sup_{\mathbf{x}\in\boldsymbol{\mathcal{X}}} \{\langle \mathbf{p}, \mathbf{x}\rangle_{\boldsymbol{\mathcal{X}}} - \Psi(\mathbf{x})\}$ for every $\mathbf{p} \in \boldsymbol{\mathcal{X}}$. The next theorem shows that $\operatorname{dom}\Psi^* = R(\mathbf{A}^*)$ and gives the expression of the function Ψ^* on its domain.

Proposition 2.2. *With the same notations as above, we have* $\operatorname{dom}\Psi^* = R(\mathbf{A}^*)$ *and* $\Psi^*(\mathbf{A}^*z) = \frac{1}{2}d^2_{\mathcal{Z}}\,(z,\ Ker(\mathbf{A}^*))$ *for every* $z \in \mathcal{Z}$.

Proof. Let us fix $\mathbf{p} \in \boldsymbol{\mathcal{X}}$. From the definition of Ψ and Ψ^*, we have $\Psi^*(\mathbf{p}) = \sup_{\mathbf{x}\in\boldsymbol{\mathcal{X}}} \left\{\langle \mathbf{p}, \mathbf{x}\rangle_{\boldsymbol{\mathcal{X}}} - \frac{1}{2}\|\mathbf{A}\mathbf{x}\|^2_{\mathcal{Z}}\right\}$. This maximization problem can be reformulated as

$$-\inf_{\mathbf{x}\in\boldsymbol{\mathcal{X}}} \{F(\mathbf{x}) + G(\mathbf{A}\mathbf{x})\}, \tag{1}$$

where $F : \boldsymbol{\mathcal{X}} \to \mathbb{R}$ and $G : \mathcal{Z} \to \mathbb{R}$ are respectively defined by $F(\mathbf{x}) = -\langle \mathbf{p}, \mathbf{x}\rangle_{\boldsymbol{\mathcal{X}}}$ and $G(z) = \frac{1}{2}\|z\|^2_{\mathcal{Z}}$ for every $\mathbf{x} \in \boldsymbol{\mathcal{X}}$, $z \in \mathcal{Z}$. Let us introduce the following minimization problem

$$\begin{aligned} \inf_{z^*\in\mathcal{Z}} \{F^*(-\mathbf{A}^*z^*) + G^*(z^*)\} &= \inf_{z^*\in\mathcal{Z}} \left\{\delta_{\{-\mathbf{p}\}}(-\mathbf{A}^*z^*) + \frac{1}{2}\|z^*\|^2_{\mathcal{Z}}\right\} \qquad (2)\\ &= \inf_{\substack{z^* \in \mathcal{Z} \\ \mathbf{A}^*z^* = \mathbf{p}}} \frac{1}{2}\|z^*\|^2_{\mathcal{Z}}. \qquad (3) \end{aligned}$$

Since the functions F and G are convex and continuous, problems (1)-(2) are dual each to other, see for example [24, Chap. III]. Observing that the Moreau-Rockafellar qualification condition is satisfied, we derive from [24, Theorem 4.1, p. 59] that the infimum values of problems (1)-(2) are simultaneously finite and in this case they coincide. Expression (3) shows that the infimum in (2) is finite if and only if $\mathbf{p} \in R(\mathbf{A}^*)$. Coming back to problem (1), we deduce that $\mathbf{p} \in \operatorname{dom}\Psi^*$ if and only if $\mathbf{p} \in R(\mathbf{A}^*)$. Now assume that $\mathbf{p} = \mathbf{A}^*z$ for some $z \in \mathcal{Z}$. Then we have

$$\inf_{\substack{z^* \in \mathcal{Z} \\ \mathbf{A}^*z^* = \mathbf{p}}} \frac{1}{2}\|z^*\|^2_{\mathcal{Z}} = \inf_{\substack{z^* \in \mathcal{Z} \\ z^* - z \in \operatorname{Ker}(\mathbf{A}^*)}} \frac{1}{2}\|z^*\|^2_{\mathcal{Z}} = \frac{1}{2}d^2_{\mathcal{Z}}\,(z,\ \operatorname{Ker}(\mathbf{A}^*)),$$

which ends the proof. □

3. Maximal monotone framework: ergodic convergence results

The notations and hypotheses are the same as in the previous section. Given any initial point $(x_0, y_0) \in \boldsymbol{\mathcal{X}}$, the iterates generated by algorithm $(\mathcal{A})$ are denoted by (x_n, y_n), $n \in \mathbb{N}$.

3.1. **Preliminary results.** Let us start with an estimation that is at the core of the convergence analysis. For $(x, y) \in \boldsymbol{\mathcal{X}}$ set

$$h_n(x,y) = \alpha\|x_n - x\|^2_{\mathcal{X}} + \nu\|y_n - y\|^2_{\mathcal{Y}} + \|By_n - By\|^2_{\mathcal{Z}}. \tag{4}$$

Then we have the following:

Lemma 3.1. *For every $(x,y) \in \boldsymbol{\mathcal{V}}$ and $(\zeta,\eta) \in \mathbf{T}(x,y)$, there exists $\mathbf{p} \in \boldsymbol{\mathcal{V}}^{\perp}$ such that*

$$h_{n+1}(x,y) - h_n(x,y) + 2\gamma_{n+1}\left[\langle \zeta, x_{n+1} - x\rangle_{\mathcal{X}} + \langle \eta, y_{n+1} - y\rangle_{\mathcal{Y}}\right]$$
$$+ \alpha\|x_{n+1} - x_n\|_{\mathcal{X}}^2 + \nu\|y_{n+1} - y_n\|_{\mathcal{Y}}^2 + \|Ax_{n+1} - By_n\|_{\mathcal{Z}}^2 \leq 2\,\gamma_{n+1}^2 \Psi^*(\mathbf{p}). \tag{5}$$

Proof. To simplify the notation we set $h_n = h_n(x,y)$. The definition of (x_{n+1}) gives

$$\frac{\alpha}{\gamma_{n+1}}(x_{n+1} - x_n) + \frac{1}{\gamma_{n+1}} A^*(Ax_{n+1} - By_n) \in -Mx_{n+1}.$$

On the other hand since $(\zeta,\eta) \in \mathbf{T}(x,y)$, there exists $\mathbf{p} = (p,q) \in \boldsymbol{\mathcal{V}}^{\perp}$ such that

$$\zeta \in Mx + p \quad \text{and} \quad \eta \in Ny + q.$$

In particular, we have $p - \zeta \in -Mx$, which by the monotonicity of M implies

$$\frac{\alpha}{\gamma_{n+1}}\langle x_{n+1} - x_n, x_{n+1} - x\rangle_{\mathcal{X}} + \frac{1}{\gamma_{n+1}}\langle A^*(Ax_{n+1} - By_n), x_{n+1} - x\rangle_{\mathcal{X}} \leq \langle p - \zeta, x_{n+1} - x\rangle_{\mathcal{X}}.$$

This is equivalent to

$$\begin{aligned}\alpha\|x_{n+1} - x\|_{\mathcal{X}}^2 + \alpha\|x_{n+1} - x_n\|_{\mathcal{X}}^2 \;\leq\;& \alpha\|x_n - x\|_{\mathcal{X}}^2 - 2\langle Ax_{n+1} - By_n, Ax_{n+1} - Ax\rangle_{\mathcal{Z}} \\ &+ 2\gamma_{n+1}\langle p, x_{n+1} - x\rangle_{\mathcal{X}} - 2\gamma_{n+1}\langle \zeta, x_{n+1} - x\rangle_{\mathcal{X}}.\end{aligned}$$

In a similar way we obtain

$$\begin{aligned}\nu\|y_{n+1} - y\|_{\mathcal{Y}}^2 + \nu\|y_{n+1} - y_n\|_{\mathcal{Y}}^2 \;\leq\;& \nu\|y_n - y\|_{\mathcal{Y}}^2 - 2\langle By_{n+1} - Ax_{n+1}, By_{n+1} - By\rangle_{\mathcal{Z}} \\ &+ 2\gamma_{n+1}\langle q, y_{n+1} - y\rangle_{\mathcal{Y}} - 2\gamma_{n+1}\langle \eta, y_{n+1} - y\rangle_{\mathcal{Y}}.\end{aligned}$$

Using the properties of the inner product and the fact that $Ax = By$, we let the reader check that

$$-2\langle Ax_{n+1} - By_n, Ax_{n+1} - Ax\rangle_{\mathcal{Z}} - 2\langle By_{n+1} - Ax_{n+1}, By_{n+1} - By\rangle_{\mathcal{Z}} =$$
$$\|By_n - By\|_{\mathcal{Z}}^2 - \|By_{n+1} - By\|_{\mathcal{Z}}^2 - \|Ax_{n+1} - By_n\|_{\mathcal{Z}}^2 - \|Ax_{n+1} - By_{n+1}\|_{\mathcal{Z}}^2.$$

Since $(x,y) \in \boldsymbol{\mathcal{V}}$ and $\mathbf{p} = (p,q) \in \boldsymbol{\mathcal{V}}^{\perp}$, we have

$$\langle p, x_{n+1} - x\rangle_{\mathcal{X}} + \langle q, y_{n+1} - y\rangle_{\mathcal{Y}} = \langle p, x_{n+1}\rangle_{\mathcal{X}} + \langle q, y_{n+1}\rangle_{\mathcal{Y}} = \langle \mathbf{p}, (x_{n+1}, y_{n+1})\rangle_{\boldsymbol{\mathcal{X}}}.$$

Gathering all this information and writing

$$\begin{aligned}c_n \;=\;& h_{n+1} - h_n + 2\gamma_{n+1}\left[\langle \zeta, x_{n+1} - x\rangle_{\mathcal{X}} + \langle \eta, y_{n+1} - y\rangle_{\mathcal{Y}}\right] \\ &+ \alpha\|x_{n+1} - x_n\|_{\mathcal{X}}^2 + \nu\|y_{n+1} - y_n\|_{\mathcal{Y}}^2 + \|Ax_{n+1} - By_n\|_{\mathcal{Z}}^2\end{aligned}$$

we deduce that

$$\begin{aligned}c_n \;\leq\;& 2\gamma_{n+1}\langle \mathbf{p}, (x_{n+1}, y_{n+1})\rangle_{\boldsymbol{\mathcal{X}}} - \|Ax_{n+1} - By_{n+1}\|_{\mathcal{Z}}^2 \\ =\;& 2\left[\langle \gamma_{n+1}\mathbf{p}, (x_{n+1}, y_{n+1})\rangle_{\boldsymbol{\mathcal{X}}} - \Psi(x_{n+1}, y_{n+1})\right].\end{aligned}$$

By definition of Ψ^*, the term between brackets is majorized by $\Psi^*(\gamma_{n+1}\mathbf{p})$. Since $\Psi^*(\gamma_{n+1}\mathbf{p}) = \gamma_{n+1}^2\Psi^*(\mathbf{p})$, inequality (5) immediately follows. □

In order to exploit inequality (5), we may assume that $\Psi^*(\mathbf{p}) < +\infty$ for every $\mathbf{p} \in \boldsymbol{\mathcal{V}}^{\perp}$. In view of Proposition 2.2, this amounts to saying that $\boldsymbol{\mathcal{V}}^{\perp} \subset \operatorname{dom}\Psi^* = R(\mathbf{A}^*)$. Since $\boldsymbol{\mathcal{V}}^{\perp} = \operatorname{Ker}(\mathbf{A})^{\perp} = \overline{R(\mathbf{A}^*)}$, this condition is equivalent to the closedness of the space $R(\mathbf{A}^*)$, which is in turn equivalent to the closedness of $R(\mathbf{A})$ in $\mathcal{Z}$. From now on, we assume in this section that

$$R(\mathbf{A}) = R(A) + R(B) \quad \text{is closed in } \mathcal{Z}.$$

By using Lemma 3.1, we now prove the boundedness of the sequence (x_n, y_n) along with the summability of the sequences $(\|x_{n+1} - x_n\|^2_{\mathcal{X}})$, $(\|y_{n+1} - y_n\|^2_{\mathcal{Y}})$ and $(\|Ax_n - By_n\|^2_{\mathcal{Z}})$.

Proposition 3.2. *Assume that the space $R(\mathbf{A})$ is closed in $\mathcal{Z}$ and that $(\gamma_n) \in l^2$. Suppose that the set $\boldsymbol{\mathcal{S}}$ is non empty and let $(\overline{x}, \overline{y}) \in \boldsymbol{\mathcal{S}}$. We have the following*

(i) $\lim\limits_{n\to+\infty} h_n(\overline{x}, \overline{y})$ *exists, hence the sequence (x_n, y_n) is bounded.*

(ii) *The sequences $(\|x_{n+1} - x_n\|^2_{\mathcal{X}})$, $(\|y_{n+1} - y_n\|^2_{\mathcal{Y}})$ and $(\|Ax_n - By_n\|^2_{\mathcal{Z}})$ are summable. In particular,*

$$\lim_{n\to+\infty} \|x_{n+1} - x_n\|_{\mathcal{X}} = \lim_{n\to+\infty} \|y_{n+1} - y_n\|_{\mathcal{Y}} = \lim_{n\to+\infty} \|Ax_n - By_n\|_{\mathcal{Z}} = 0 \tag{6}$$

and every weak cluster point of the sequence (x_n, y_n) lies in $\boldsymbol{\mathcal{V}}$.

Proof. (i) Taking $(\zeta, \eta) = (0, 0)$ in inequality (5) and setting $h_n = h_n(\overline{x}, \overline{y})$, we obtain

$$h_{n+1} - h_n + \alpha\|x_{n+1} - x_n\|^2_{\mathcal{X}} + \nu\|y_{n+1} - y_n\|^2_{\mathcal{Y}} + \|Ax_{n+1} - By_n\|^2_{\mathcal{Z}} \leq 2\gamma^2_{n+1}\, \Psi^*(\mathbf{p}). \tag{7}$$

In particular, $h_{n+1} - h_n \leq 2\gamma^2_{n+1}\, \Psi^*(\mathbf{p})$. Since $(\gamma_n) \in l^2$ and $\Psi^*(\mathbf{p}) < +\infty$, the following lemma shows that (h_n) converges, which in turn implies that the sequence (x_n, y_n) is bounded.

Lemma 3.3. *Let (a_n) and (ε_n) be two real sequences. Assume that (a_n) is minorized, that $(\varepsilon_n) \in l^1$ and that $a_{n+1} \leq a_n + \varepsilon_n$ for every $n \in \mathbb{N}$. Then (a_n) converges.*

Proof of Lemma 3.3. Define the sequence (w_n) by $w_n = a_n - \sum_{k=0}^{n-1} \varepsilon_k$. The sequence (w_n) is bounded from below and nonincreasing, hence convergent. It follows that $\lim_{n\to+\infty} a_n = \sum_{k=0}^{+\infty} \varepsilon_k + \lim_{n\to+\infty} w_n$. □

(ii) Let us sum up inequality (7) from $n = 0$ to $+\infty$. Recalling that $(\gamma_n) \in l^2$, that $\Psi^*(\mathbf{p}) < +\infty$ and that $h_n \geq 0$, we immediately deduce the summability of the sequences $(\|x_{n+1} - x_n\|^2_{\mathcal{X}})$, $(\|y_{n+1} - y_n\|^2_{\mathcal{Y}})$ and $(\|Ax_{n+1} - By_n\|^2_{\mathcal{Z}})$. Since $\|Ax_n - By_n\|^2_{\mathcal{Z}} \leq 2\|Ax_{n+1} - By_n\|^2_{\mathcal{Z}} + 2\|Ax_{n+1} - Ax_n\|^2_{\mathcal{Z}}$, the sequence $(\|Ax_n - By_n\|^2_{\mathcal{Z}})$ is also summable. For the last part use the fact that $\lim\limits_{n\to+\infty} \|Ax_n - By_n\|^2_{\mathcal{Z}} = 0$ and the weak lower-semicontinuity of the function $(x, y) \mapsto \|Ax - By\|^2_{\mathcal{Z}}$. □

Remark 3.4. Proposition 3.2 still holds if one assumes only that $\overline{R(\mathbf{A}^*)} \cap R(\mathbf{M}) \subset R(\mathbf{A}^*)$, a condition that is weaker than the closedness of $R(\mathbf{A})$. The reason is that one uses Lemma 3.1 for $(x, y) = (\overline{x}, \overline{y}) \in \boldsymbol{\mathcal{S}}$.

3.2. **Ergodic convergence.** From now on we assume that $(\gamma_n) \in l^2 \setminus l^1$. Condition $(\gamma_n) \notin l^1$ is standard and common to most proximal-type algorithms. Roughly speaking it states that the corresponding time $\sigma_n = \sum_{k=1}^{n} \gamma_k$ goes to $+\infty$. It also means that the sequence (γ_n) does not tend to 0 too fast as $n \to +\infty$ while condition $(\gamma_n) \in l^2 \setminus l^1$ expresses that the sequence (γ_n) tends *moderately* slowly toward 0. This kind of assumption appears in several works related to proximal algorithms involving maximal monotone operators with alternating features. See for example the seminal work [30]. Let us define the averages

$$\widetilde{x}_n = \frac{1}{\sigma_n} \sum_{k=1}^{n} \gamma_k x_k \quad \text{and} \quad \widetilde{y}_n = \frac{1}{\sigma_n} \sum_{k=1}^{n} \gamma_k y_k, \tag{8}$$

and prove that the sequence $(\widetilde{x}_n, \widetilde{y}_n)$ converges weakly to a point in $\boldsymbol{\mathcal{S}}$.

Theorem 3.5. *Assume that the space $R(\mathbf{A})$ is closed in $\mathcal{Z}$ and that $(\gamma_n) \in l^2 \setminus l^1$. Assume moreover that the operator $\mathbf{T}$ is maximal monotone with $\boldsymbol{\mathcal{S}} = \mathbf{T}^{-1}0 \neq \emptyset$. Then the sequence $(\widetilde{x}_n, \widetilde{y}_n)$ of averages converges weakly as $n \to +\infty$ to a point in $\boldsymbol{\mathcal{S}}$.*

Proof. Let us first prove that every weak cluster point of the sequence $(\widetilde{x}_n, \widetilde{y}_n)$ is in $\boldsymbol{\mathcal{S}}$. Fix $(\zeta, \eta) \in \mathbf{T}(x, y)$. By summing up inequality (5) of Lemma 3.1 for $k = 0, \dots, n-1$, we obtain

$$\langle \zeta, \widetilde{x}_n - x \rangle_{\mathcal{X}} + \langle \eta, \widetilde{y}_n - y \rangle_{\mathcal{Y}} \leq \frac{1}{\sigma_n} \left[h_0(x, y) + 2\,\Psi^*(\mathbf{p}) \sum_{k=1}^{n} \gamma_k^2 \right].$$

Let $(\widetilde{x}_\infty, \widetilde{y}_\infty)$ be a weak cluster point of $(\widetilde{x}_n, \widetilde{y}_n)$ as $n \to +\infty$ and let n tend to $+\infty$ in the above inequality. By using the fact that $\Psi^*(\mathbf{p}) < +\infty$ and that $(\gamma_n) \in l^2 \setminus l^1$, we deduce that

$$\langle \zeta, \widetilde{x}_\infty - x \rangle_{\mathcal{X}} + \langle \eta, \widetilde{y}_\infty - y \rangle_{\mathcal{Y}} \leq 0. \tag{9}$$

Since this holds whenever $(\zeta, \eta) \in \mathbf{T}(x, y)$ we conclude $(\widetilde{x}_\infty, \widetilde{y}_\infty) \in \boldsymbol{\mathcal{S}}$ by maximality of the monotone operator $\mathbf{T}$.

Now observe that the sequence $(\widetilde{x}_n, \widetilde{y}_n)$ is bounded by Proposition 3.2. In order to establish the weak convergence of the sequence $(\widetilde{x}_n, \widetilde{y}_n)$ it suffices to prove that it has at most one weak cluster point[2]. Indeed, let (x, y) and (x', y') be two such points, which must belong to $\boldsymbol{\mathcal{S}}$. Define the quantity $Q(u, v) = \alpha \|u\|_{\mathcal{X}}^2 + \nu \|v\|_{\mathcal{Y}}^2 + \|Bv\|_{\mathcal{Z}}^2$ for every $(u, v) \in \boldsymbol{\mathcal{X}}$. From Proposition 3.2 (i), the limits

$$\ell(x, y) = \lim_{n \to +\infty} Q(x_n - x, y_n - y) \qquad \text{and} \qquad \ell(x', y') = \lim_{n \to +\infty} Q(x_n - x', y_n - y')$$

exist. Observe that

$$\begin{aligned} Q(x_n - x, y_n - y) \;\; &= \;\; Q(x_n - x', y_n - y') + Q(x - x', y - y') \\ &\quad + 2\,\alpha \langle x_n - x', x' - x \rangle_{\mathcal{X}} + 2\,\nu \langle y_n - y', y' - y \rangle_{\mathcal{Y}} \\ &\quad + 2\,\langle B(y_n - y'), B(y' - y) \rangle_{\mathcal{Z}}. \end{aligned} \tag{10}$$

Taking the average and letting $(\widetilde{x}_{n_k}, \widetilde{y}_{n_k}) \rightharpoonup (x', y')$ as $k \to +\infty$ we obtain

$$\ell(x, y) = \ell(x', y') + Q(x - x', y - y').$$

In a similar fashion we deduce that

$$\ell(x', y') = \ell(x, y) + Q(x - x', y - y')$$

and hence $Q(x - x', y - y') = 0$ which implies $(x, y) = (x', y')$. $\square$

3.3. **Links with Passty theorem.** Assume that $\mathcal{X} = \mathcal{Y} = \mathcal{Z}$ and that $A = B = \mathcal{I}$, along with $\alpha = \nu = 0$. This induces a situation of strong coupling without cost-to-move. The corresponding algorithm is denoted by $(\mathcal{A}_0)$, see Remark 2.1. Since $R(\mathbf{A}) = \mathcal{X}$, the closedness of $R(\mathbf{A})$ is automatically satisfied. It is immediate that $0 \in \mathbf{T}(x, y)$ if and only if $x = y$ and $Mx + Nx \ni 0$. Therefore we have

$$\boldsymbol{\mathcal{S}} = \mathbf{T}^{-1}0 = \left\{ (x, x) \in \mathcal{X}^2, \quad x \in (M + N)^{-1}0 \right\}.$$

Assume that the operator $M + N$ is maximal monotone with $(M + N)^{-1}0 \neq \emptyset$. Let (γ_n) be a positive sequence such that $(\gamma_n) \in l^2 \setminus l^1$. By arguing as in the proof of Proposition 3.2, we obtain that

[2]This idea, inspired by the Opial lemma [29] (see Lemma 4.9 below), can also be found in [30].

(a) $\lim_{n\to+\infty} \|y_n - \overline{y}\|^2_{\mathcal{X}}$ exists for every $\overline{y} \in (M+N)^{-1}0$.

(b) The sequence $\left(\|x_{n+1} - y_n\|^2_{\mathcal{X}}\right)$ is summable, hence $\lim_{n\to+\infty} \|x_{n+1} - y_n\|_{\mathcal{X}} = 0$.

Take $\zeta = 0$ and $x = y$ in the proof of Theorem 3.5. Observe that $(0,\eta) \in T(y,y)$ holds if and only if $\eta \in (M+N)y$. Hence formula (9) implies that $\langle \eta, \widetilde{y}_\infty - y\rangle_{\mathcal{Y}} \leq 0$ for every $\eta \in (M+N)y$. We deduce that $\widetilde{y}_\infty \in (M+N)^{-1}0$ by maximality of the monotone operator $M+N$. This proves that every weak cluster point of the sequence $(\widetilde{y}_n)$ lies in $(M+N)^{-1}0$. Then we prove that the sequence $(\widetilde{y}_n)$ has at most one weak cluster point. It suffices to adapt the proof of Theorem 3.5 by invoking point (a) above and by using the quantity $Q(v) = \|v\|^2_{\mathcal{X}}$ (instead of $Q(u,v)$). We obtain that the sequence $(\widetilde{y}_n)$ weakly converges toward some $\widetilde{y}_\infty \in (M+N)^{-1}0$. By using point (b) above, we infer that the sequence $(\widetilde{x}_n)$ weakly converges toward $\widetilde{x}_\infty = \widetilde{y}_\infty$. As a conclusion, we recover the following result

Theorem (Passty [30]) *Assume that the operator $M+N$ is maximal monotone with $(M+N)^{-1}0 \neq \emptyset$. Let (γ_n) be a positive sequence such that $(\gamma_n) \in l^2 \setminus l^1$ and let (x_n, y_n) be any sequence generated by algorithm $(\mathcal{A}_0)$. Then there exists $\widetilde{x}_\infty \in (M+N)^{-1}0$ such that both sequences of averages $(\widetilde{x}_n)$ and $(\widetilde{y}_n)$ converge weakly toward $\widetilde{x}_\infty$.*

3.4. **Strong monotonicity.** Under strong monotonicity assumptions, we are able to prove the strong convergence of the sequence (x_n, y_n) itself (not only in average). Let us recall that the operator M is said to be strongly monotone with parameter a if, for every x_1, $x_2 \in \operatorname{dom} M$ and every $\xi_1 \in Mx_1$, $\xi_2 \in Mx_2$, we have

$$\langle \xi_2 - \xi_1, x_2 - x_1\rangle_{\mathcal{X}} \geq a\,\|x_2 - x_1\|^2_{\mathcal{X}}.$$

Assuming in the same way that the operator N is strongly monotone, we obtain that the operators $\mathbf{M}$ and $\mathbf{T} = \mathbf{M} + N_{\mathcal{V}}$ are strongly monotone. Hence if the set $\boldsymbol{\mathcal{S}} = \mathbf{T}^{-1}0$ is nonempty it must be reduced to a single point, say $\boldsymbol{\mathcal{S}} = \{(\overline{x}, \overline{y})\}$.

Proposition 3.6. *Assume that the space $R(\mathbf{A})$ is closed in $\mathcal{Z}$ and that $(\gamma_n) \in l^2 \setminus l^1$. If the operators M and N are strongly monotone and if $\boldsymbol{\mathcal{S}} \neq \emptyset$ then the sequence (x_n, y_n) converges strongly to the unique $(\overline{x}, \overline{y}) \in \boldsymbol{\mathcal{S}}$.*

Proof. Let us suppose that the operators M and N are strongly monotone, respectively with parameters $a, b > 0$. We let the reader check that this assumption leads to a stronger form of inequality (5), which in turn implies

$$h_{n+1}(\overline{x},\overline{y}) - h_n(\overline{x},\overline{y}) + 2a\gamma_{n+1}\|x_{n+1} - \overline{x}\|^2_{\mathcal{X}} + 2b\gamma_{n+1}\|y_{n+1} - \overline{y}\|^2_{\mathcal{Y}} \leq 2\gamma^2_{n+1}\,\Psi^*(\mathbf{p}).$$

Since $(\gamma_n) \in l^2$ and $\Psi^*(\mathbf{p}) < +\infty$, and recalling that $h_n(\overline{x},\overline{y}) \geq 0$, the summation of the above inequality implies

$$\sum_{n=1}^{+\infty} \gamma_n \left[\|x_n - \overline{x}\|^2_{\mathcal{X}} + \|y_n - \overline{y}\|^2_{\mathcal{Y}}\right] < +\infty$$

and hence

$$\sum_{n=1}^{+\infty} \gamma_n h_n(\overline{x},\overline{y}) \leq \alpha \sum_{n=1}^{+\infty} \gamma_n \|x_n - \overline{x}\|^2_{\mathcal{X}} + (\nu + \|B\|^2) \sum_{n=1}^{+\infty} \gamma_n \|y_n - \overline{y}\|^2_{\mathcal{Y}} < +\infty.$$

Since $(\gamma_n) \notin l^1$ and since $\lim_{n\to+\infty} h_n(\overline{x},\overline{y})$ exists, this limit must be equal to 0 and we deduce that $\lim_{n\to+\infty} (x_n, y_n) = (\overline{x}, \overline{y})$. □

Remark 3.7. Observe that the maximality of the operator $\mathbf{T}$ does not come into play in the previous proof. Notice also that if the operator $\mathbf{T}$ is both maximal and strongly monotone, then condition $\boldsymbol{\mathcal{S}} \neq \emptyset$ is automatically satisfied, see for example [18, Cor. 2.4] or [34, Prop. 12. 54].

4. The subdifferential case: weak convergence results

4.1. **Preliminaries.** Let $f : \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$, $g : \mathcal{Y} \to \mathbb{R} \cup \{+\infty\}$ be closed convex proper functions. Define the maximal monotone operators M and N respectively by $M = \partial f$ and $N = \partial g$. The operator $\mathbf{M}$ coincides with the subdifferential of the function Φ defined by $\Phi(x, y) = f(x) + g(y)$ for every $(x, y) \in \boldsymbol{\mathcal{X}}$. Observe that the monotone operator $\mathbf{T} = \partial\Phi + N_{\boldsymbol{\mathcal{V}}} = \partial\Phi + \partial\delta_{\boldsymbol{\mathcal{V}}}$ is maximal if, and only if,

$$\partial\Phi + \partial\delta_{\boldsymbol{\mathcal{V}}} = \partial\left(\Phi + \delta_{\boldsymbol{\mathcal{V}}}\right).$$

Maximality is guaranteed if one assumes some qualification condition such as the Moreau-Rockafellar one [28, 33] or the Attouch-Brézis one [7]. In order to cover various applications to PDE's (see paragraph 6.2), we assume the following Attouch-Brézis qualification condition

$$(QC) \qquad \bigcup_{\lambda>0} \lambda(\operatorname{dom} f \times \operatorname{dom} g - \boldsymbol{\mathcal{V}}) \quad \text{is a closed subspace of } \mathcal{X} \times \mathcal{Y}.$$

Under (QC) the following claim shows that the set $\boldsymbol{\mathcal{S}} = \mathbf{T}^{-1}0$ can be interpreted as the set of minima of a suitable function.

Claim 4.1. *We have*

$$\boldsymbol{\mathcal{S}} \subset \operatorname{argmin}_{\boldsymbol{\mathcal{V}}}\Phi = \operatorname{argmin}\{f(x) + g(y) : Ax = By\}.$$

If condition (QC) is satisfied, the above inclusion holds true as an equality.

Proof. First recall that the inclusion $\partial\Phi + \partial\delta_{\boldsymbol{\mathcal{V}}} \subset \partial\left(\Phi + \delta_{\boldsymbol{\mathcal{V}}}\right)$ is always satisfied. It ensues immediately that

$$\boldsymbol{\mathcal{S}} = \mathbf{T}^{-1}0 = \left[\partial\Phi + \partial\delta_{\boldsymbol{\mathcal{V}}}\right]^{-1} 0 \subset \left[\partial\left(\Phi + \delta_{\boldsymbol{\mathcal{V}}}\right)\right]^{-1} 0 = \operatorname{argmin}_{\boldsymbol{\mathcal{V}}}\Phi.$$

If condition (QC) is satisfied, the set $\bigcup_{\lambda>0} \lambda(\operatorname{dom}\Phi - \operatorname{dom}\delta_{\boldsymbol{\mathcal{V}}})$ is a closed subspace of $\boldsymbol{\mathcal{X}}$. This classically implies that $\partial\Phi + \partial\delta_{\boldsymbol{\mathcal{V}}} = \partial(\Phi + \delta_{\boldsymbol{\mathcal{V}}})$ and the conclusion follows. $\square$

Recall that the iterate (x_{n+1}, y_{n+1}) of algorithm $(\mathcal{A})$ is implicitly defined by

$$\left\{\begin{array}{lll} 0 \in & \gamma_{n+1}\,\partial f(x_{n+1}) + A^*(Ax_{n+1} - By_n) & +\alpha(x_{n+1} - x_n) \\ 0 \in & \gamma_{n+1}\,\partial g(y_{n+1}) - B^*(Ax_{n+1} - By_{n+1}) & +\nu(y_{n+1} - y_n). \end{array}\right. \tag{11}$$

These are the optimality conditions associated to the following minimization problems

$$(\mathcal{A}) \quad \left\{\begin{array}{l} x_{n+1} = \operatorname{argmin}\{\gamma_{n+1} f(\zeta) + \dfrac{1}{2}\|A\zeta - By_n\|^2_{\mathcal{Z}} + \dfrac{\alpha}{2}\|\zeta - x_n\|^2_{\mathcal{X}}; \quad \zeta \in \mathcal{X}\} \\ y_{n+1} = \operatorname{argmin}\{\gamma_{n+1} g(\eta) + \dfrac{1}{2}\|Ax_{n+1} - B\eta\|^2_{\mathcal{Z}} + \dfrac{\nu}{2}\|\eta - y_n\|^2_{\mathcal{Y}}; \quad \eta \in \mathcal{Y}\}. \end{array}\right.$$

For each $(x, y) \in \boldsymbol{\mathcal{X}}$, define

$$h_n(x, y) = \alpha\|x_n - x\|^2_{\mathcal{X}} + \nu\|y_n - y\|^2_{\mathcal{Y}} + \|By_n - By\|^2_{\mathcal{Z}} \tag{12}$$

as in section 3. Define also the sequence (φ_n) by

$$\varphi_n = f(x_n) + g(y_n) + \frac{1}{2\gamma_n}\|Ax_n - By_n\|^2_{\mathcal{Z}}. \tag{13}$$

Lemma 4.2. *With the above notations and hypotheses, we have the following*[3]

(i) *For every* $(x,y) \in \operatorname{argmin}_{\mathcal{V}}\Phi$ *and for every* $n \geq 0$,

$$h_{n+1}(x,y) - h_n(x,y) + 2\gamma_{n+1}\left(f(x_{n+1}) + g(y_{n+1}) - \min_{\mathcal{V}}\Phi\right) + \|Ax_{n+1} - By_{n+1}\|^2_{\mathcal{Z}}$$
$$+\|Ax_{n+1} - By_n\|^2_{\mathcal{Z}} + \alpha\|x_{n+1} - x_n\|^2_{\mathcal{X}} + \nu\|y_{n+1} - y_n\|^2_{\mathcal{Y}} \leq 0. \tag{14}$$

(ii) *For every* $n \geq 0$,

$$\varphi_{n+1} - \varphi_n \leq \frac{1}{2}\left(\frac{1}{\gamma_{n+1}} - \frac{1}{\gamma_n}\right)\|Ax_n - By_n\|^2_{\mathcal{Z}}. \tag{15}$$

Proof. In view of the optimality conditions (11), for all $(x,y) \in \mathcal{X}\times\mathcal{Y}$ we can write the subdifferential inequalities

$$\gamma_{n+1}\left(f(x) - f(x_{n+1})\right) \geq -\langle Ax_{n+1} - By_n, Ax - Ax_{n+1}\rangle_{\mathcal{Z}} - \alpha\langle x_{n+1} - x_n, x - x_{n+1}\rangle_{\mathcal{X}} \tag{16}$$

and

$$\gamma_{n+1}\left(g(y) - g(y_{n+1})\right) \geq \langle Ax_{n+1} - By_{n+1}, By - By_{n+1}\rangle_{\mathcal{Z}} - \nu\langle y_{n+1} - y_n, y - y_{n+1}\rangle_{\mathcal{Y}}. \tag{17}$$

Using the properties of the inner product, the reader can check that

$$\begin{aligned}
\|By - By_n\|^2_{\mathcal{Z}} - \|By - By_{n+1}\|^2_{\mathcal{Z}} \quad = \quad & \|By_{n+1} - Ax_{n+1}\|^2_{\mathcal{Z}} - \|By - Ax\|^2_{\mathcal{Z}} \\
& + \|By - Ax - (By_n - Ax_{n+1})\|^2_{\mathcal{Z}} \\
& + 2\langle By - By_{n+1}, By_{n+1} - Ax_{n+1}\rangle_{\mathcal{Z}} \\
& + 2\langle By_n - Ax_{n+1}, Ax_{n+1} - Ax\rangle_{\mathcal{Z}}.
\end{aligned}$$

Combining (16) and (17) we deduce that

$$\begin{aligned}
&\|By - By_n\|^2_{\mathcal{Z}} - \|By - By_{n+1}\|^2_{\mathcal{Z}} \\
&\quad \geq \quad \|By_{n+1} - Ax_{n+1}\|^2_{\mathcal{Z}} - \|By - Ax\|^2_{\mathcal{Z}} + \|By - Ax - (By_n - Ax_{n+1})\|^2_{\mathcal{Z}} \\
&\qquad\quad + 2\gamma_{n+1}\left[f(x_{n+1}) - f(x) + g(y_{n+1}) - g(y)\right] \\
&\qquad\quad + 2\alpha\langle x_{n+1} - x_n, x_{n+1} - x\rangle_{\mathcal{X}} + 2\nu\langle y_{n+1} - y_n, y_{n+1} - y\rangle_{\mathcal{Y}}, \\
&\quad = \quad \|Ax_{n+1} - By_{n+1}\|^2_{\mathcal{Z}} - \|Ax - By\|^2_{\mathcal{Z}} + \|By - Ax - (By_n - Ax_{n+1})\|^2_{\mathcal{Z}} \\
&\qquad\quad + 2\gamma_{n+1}(f(x_{n+1}) + g(y_{n+1}) - f(x) - g(y)) \\
&\qquad\quad + \alpha(\|x_{n+1} - x_n\|^2_{\mathcal{X}} + \|x_{n+1} - x\|^2_{\mathcal{X}} - \|x_n - x\|^2_{\mathcal{X}}) \\
&\qquad\quad + \nu(\|y_{n+1} - y_n\|^2_{\mathcal{Y}} + \|y_{n+1} - y\|^2_{\mathcal{Y}} - \|y_n - y\|^2_{\mathcal{Y}}).
\end{aligned}$$

We infer that for all $(x,y) \in \mathcal{X}\times\mathcal{Y}$,

$$\begin{aligned}
h_n(x,y) - h_{n+1}(x,y) \quad \geq \quad & 2\gamma_{n+1}(f(x_{n+1}) + g(y_{n+1}) - f(x) - g(y)) - \|Ax - By\|^2_{\mathcal{Z}} \\
& + \|Ax_{n+1} - By_{n+1}\|^2_{\mathcal{Z}} + \|By - Ax - (By_n - Ax_{n+1})\|^2_{\mathcal{Z}} \\
& + \alpha\|x_{n+1} - x_n\|^2_{\mathcal{X}} + \nu\|y_{n+1} - y_n\|^2_{\mathcal{Y}}.
\end{aligned} \tag{18}$$

---

[3] Inequalities (5) and (14) are closely related, even if they rely on different techniques (monotonicity in the first case and subdifferential inequalities in the second one).

Now let $(x,y) \in \operatorname{argmin}_{\boldsymbol{\mathcal{V}}}\Phi$. Then $Ax = By$ and $f(x) + g(y) = \min_{\boldsymbol{\mathcal{V}}} \Phi$ so that inequality (18) becomes (14). On the other hand, by using inequality (18) with $x = x_n$ and $y = y_n$, we infer that

$$2\gamma_{n+1}(f(x_{n+1}) + g(y_{n+1}) - f(x_n) - g(y_n)) + \|Ax_{n+1} - By_{n+1}\|^2_{\mathcal{Z}} \le \|Ax_n - By_n\|^2_{\mathcal{Z}}. \tag{19}$$

We finally divide by $2\gamma_{n+1}$ and rearrange the terms to obtain (15). □

4.2. **Weak convergence.** Assuming that $\operatorname{argmin}_{\boldsymbol{\mathcal{V}}}\Phi \neq \emptyset$, let us set

$$\begin{aligned} \omega_n &= \inf_{(x,y)\in\mathcal{X}\times\mathcal{Y}} \left\{ \frac{1}{2}\|Ax - By\|^2_{\mathcal{Z}} + \gamma_n \left(f(x) + g(y) - \min_{\boldsymbol{\mathcal{V}}} \Phi \right) \right\} \\ &= \inf_{\mathbf{x}\in\boldsymbol{\mathcal{X}}} \left\{ \Psi(\mathbf{x}) + \gamma_n \left(\Phi(\mathbf{x}) - \min_{\boldsymbol{\mathcal{V}}} \Phi \right) \right\}. \end{aligned} \tag{20}$$

Denote by (ω_n^-) the negative part of (ω_n). In the sequel, we will assume the key condition

$$(\omega_n^-) \in l^1.$$

This kind of hypothesis was introduced by the second author in [21].

Proposition 4.3. *Assuming that* $\operatorname{argmin}_{\boldsymbol{\mathcal{V}}}\Phi \neq \emptyset$, *consider the following assertions:*

(i) $(\gamma_n) \in l^2$, *the space* $R(\mathbf{A})$ *is closed in* $\mathcal{Z}$ *and condition* (QC) *is satisfied.*
(ii) $(\gamma_n) \in l^2$ *and there exists* $\overline{\mathbf{x}} \in \boldsymbol{\mathcal{V}}$ *and* $\mathbf{p} \in R(\mathbf{A}^*)$ *such that* $-\mathbf{p} \in \partial\Phi(\overline{\mathbf{x}})$.
(iii) $(\omega_n^-) \in l^1$.

Then we have the implications (i) $\Longrightarrow$ (ii) $\Longrightarrow$ (iii).

Proof. (i) $\Longrightarrow$ (ii) Let $\overline{\mathbf{x}} \in \operatorname{argmin}_{\boldsymbol{\mathcal{V}}}\Phi$. Since condition (QC) is satisfied, we deduce from Claim 4.1 that $\overline{\mathbf{x}} \in \boldsymbol{\mathcal{S}} = \left[\partial\Phi + \boldsymbol{\mathcal{V}}^\perp\right]^{-1}0$. Hence there exists $\mathbf{p} \in \boldsymbol{\mathcal{V}}^\perp$ such that $-\mathbf{p} \in \partial\Phi(\overline{\mathbf{x}})$. The closedness of $R(\mathbf{A})$ implies the closedness of $R(\mathbf{A}^*)$, hence we have $\boldsymbol{\mathcal{V}}^\perp = \operatorname{Ker}(\mathbf{A})^\perp = R(\mathbf{A}^*)$ and finally $\mathbf{p} \in R(\mathbf{A}^*)$.
(ii) $\Longrightarrow$ (iii) The subdifferential inequality gives for every $\mathbf{x} \in \boldsymbol{\mathcal{X}}$,

$$\Phi(\mathbf{x}) - \Phi(\overline{\mathbf{x}}) \ge \langle -\mathbf{p}, \mathbf{x} - \overline{\mathbf{x}}\rangle_{\boldsymbol{\mathcal{X}}} = \langle -\mathbf{p}, \mathbf{x}\rangle_{\boldsymbol{\mathcal{X}}},$$

where the last equality is a consequence of $\mathbf{p} \in R(\mathbf{A}^*) \subset \boldsymbol{\mathcal{V}}^\perp$ and $\overline{\mathbf{x}} \in \boldsymbol{\mathcal{S}} \subset \boldsymbol{\mathcal{V}}$. Since $\Phi(\overline{\mathbf{x}}) = \min_{\boldsymbol{\mathcal{V}}} \Phi$, we deduce that

$$\Psi(\mathbf{x}) + \gamma_n \left(\Phi(\mathbf{x}) - \min_{\boldsymbol{\mathcal{V}}} \Phi \right) \ge \Psi(\mathbf{x}) - \gamma_n \langle \mathbf{p}, \mathbf{x}\rangle_{\boldsymbol{\mathcal{X}}}.$$

Taking the infimum over $\mathbf{x} \in \boldsymbol{\mathcal{X}}$, we find

$$\omega_n \ge -\sup_{\mathbf{x}\in\boldsymbol{\mathcal{X}}} \{\gamma_n \langle \mathbf{p}, \mathbf{x}\rangle_{\boldsymbol{\mathcal{X}}} - \Psi(\mathbf{x})\} = -\Psi^*(\gamma_n \mathbf{p}) = -\gamma_n^2 \Psi^*(\mathbf{p}).$$

It ensues that $\omega_n^- \le \gamma_n^2 \Psi^*(\mathbf{p})$. Since $\mathbf{p} \in R(\mathbf{A}^*) = \operatorname{dom}\Psi^*$ (see Proposition 2.2), the conclusion follows from the summability of (γ_n^2). □

Notice that in infinite dimensional spaces, conditions (ii) or (iii) can be satisfied even if the space $R(\mathbf{A})$ is not closed. An example will be provided in the last section.

Let us now state the main result of the paper.

Theorem 4.4. *Let $f : \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ and $g : \mathcal{Y} \to \mathbb{R} \cup \{+\infty\}$ be closed convex proper functions. Let $A : \mathcal{X} \to \mathcal{Z}$ and $B : \mathcal{Y} \to \mathcal{Z}$ be linear continuous operators. Assume that the qualification condition (QC) holds and that* $\operatorname{argmin}\{f(x) + g(y) : Ax = By\} \neq \emptyset$. *Let (γ_n) be a positive sequence such that $\left(\frac{1}{\gamma_{n+1}} - \frac{1}{\gamma_n}\right)$ is majorized by some $M > 0$. Finally suppose that condition $(\omega_n^-) \in l^1$ holds, where the sequence (ω_n) is defined by (20). Then we have*

(i) *(x_n, y_n) converges weakly to a point $(x_\infty, y_\infty) \in$* $\operatorname{argmin}\{f(x) + g(y) : Ax = By\}$.
(ii) $\lim_{n\to+\infty} f(x_n) = f(x_\infty)$ *and* $\lim_{n\to+\infty} g(y_n) = g(y_\infty)$.

Proof. Let us start with several preliminary claims.

Claim 4.5. *For every $(x, y) \in \operatorname{argmin}_{\mathcal{V}} \Phi$,*

$$\lim_{n\to+\infty} \alpha\|x_n - x\|^2_{\mathcal{X}} + \nu\|y_n - y\|^2_{\mathcal{Y}} + \|By_n - By\|^2_{\mathcal{Z}} \quad \textit{exists in } \mathbb{R}.$$

Proof of Claim 4.5. Fix $(x, y) \in \operatorname{argmin}_{\mathcal{V}} \Phi$ and set $h_n = \alpha\|x_n - x\|^2_{\mathcal{X}} + \nu\|y_n - y\|^2_{\mathcal{Y}} + \|By_n - By\|^2_{\mathcal{Z}}$ as in (12). From inequality (14) we deduce that

$$h_{n+1} - h_n + 2\omega_{n+1} + \|Ax_{n+1} - By_n\|^2_{\mathcal{Z}} + \alpha\|x_{n+1} - x_n\|^2_{\mathcal{X}} + \nu\|y_{n+1} - y_n\|^2_{\mathcal{Y}} \leq 0.$$

This implies

$$h_{n+1} - h_n + \|Ax_{n+1} - By_n\|^2_{\mathcal{Z}} + \alpha\|x_{n+1} - x_n\|^2_{\mathcal{X}} + \nu\|y_{n+1} - y_n\|^2_{\mathcal{Y}} \leq 2\omega^-_{n+1}. \quad (21)$$

It ensues that $h_{n+1} - h_n \leq 2\omega^-_{n+1}$. Since $(\omega_n^-) \in l^1$, owing to Lemma 3.3, we conclude that $\lim_{n\to+\infty} h_n$ exists. ☐

Claim 4.6. *The sequence $(\|Ax_n - By_n\|^2_{\mathcal{Z}})$ is summable, and therefore $\lim_{n\to+\infty} \|Ax_n - By_n\|_{\mathcal{Z}} = 0$.*

Proof of Claim 4.6. Let us sum up inequalities (21) which are obtained for $n = 0$ to $+\infty$. Recalling that $(\omega_n^-) \in l^1$ and that $h_n \geq 0$, we immediately deduce the summability of the sequences $(\|x_{n+1} - x_n\|^2_{\mathcal{X}})$, $(\|y_{n+1} - y_n\|^2_{\mathcal{Y}})$ and $(\|Ax_{n+1} - By_n\|^2_{\mathcal{Z}})$. Since $\|Ax_n - By_n\|^2_{\mathcal{Z}} \leq 2\|Ax_{n+1} - By_n\|^2_{\mathcal{Z}} + 2\|Ax_{n+1} - Ax_n\|^2_{\mathcal{Z}}$, the sequence $(\|Ax_n - By_n\|^2_{\mathcal{Z}})$ is also summable. ☐

Claim 4.7. *Setting $\varphi_n = f(x_n) + g(y_n) + \frac{1}{2\gamma_n}\|Ax_n - By_n\|^2_{\mathcal{Z}}$ as in (13), we have $\lim_{n\to+\infty} \varphi_n = \min_{\mathcal{V}} \Phi$.*

Proof of Claim 4.7. Since $\left(\frac{1}{\gamma_{n+1}} - \frac{1}{\gamma_n}\right) \leq M$, we derive from inequality (15) that

$$\varphi_{n+1} - \varphi_n \leq \frac{M}{2}\|Ax_n - By_n\|^2_{\mathcal{Z}}. \quad (22)$$

From the previous claim the sequence $(\|Ax_n - By_n\|^2_{\mathcal{Z}})$ is summable. By applying Lemma 3.3 we deduce that the sequence (φ_n) converges. Let us now set

$$a_N = 2\sum_{n=0}^{N} \left\{\gamma_n \left(f(x_n) + g(y_n) - \min_{\mathcal{V}} \Phi\right) + \frac{1}{2}\|Ax_n - By_n\|^2_{\mathcal{Z}}\right\}.$$

From inequality (14), the sequence $(h_N + a_N)$ is nonincreasing. Moreover the assumption $(\omega_n^-) \in l^1$ allows us to assert that, for all $n \in \mathbb{N}$,

$$a_N \geq -2 \sum_{n=0}^{+\infty} \omega_n^- > -\infty.$$

Thus the sequence $(h_N + a_N)$ is bounded from below, hence convergent. As a consequence,

$$\lim_{N \to +\infty} a_N = \lim_{N \to +\infty} 2 \sum_{n=0}^{N} \gamma_n \left(\varphi_n - \min_{\boldsymbol{\mathcal{V}}} \Phi \right) \text{ exists in } \mathbb{R}. \tag{23}$$

Since $\frac{1}{\gamma_{n+1}} - \frac{1}{\gamma_n} \leq M$ for every $n \geq 0$, we deduce that $\gamma_n \geq \frac{1}{Mn+\frac{1}{\gamma_0}}$, hence $(\gamma_n) \notin l^1$. Recalling that $\lim_{n \to +\infty} \varphi_n$ exists in $\mathbb{R}$, we infer from (23) that $\lim_{n \to +\infty} \varphi_n = \min_{\boldsymbol{\mathcal{V}}} \Phi$. □

Claim 4.8. $\lim_{n \to +\infty} \Phi(x_n, y_n) = \min_{\boldsymbol{\mathcal{V}}} \Phi$.

Proof of Claim 4.8. Let $(x, y) \in \operatorname{argmin}_{\boldsymbol{\mathcal{V}}} \Phi$. Since condition (QC) holds, we deduce from Claim 4.1 that $(x, y) \in \boldsymbol{\mathcal{S}} = \mathbf{T}^{-1}0$. Hence there exists $(p, q) \in \boldsymbol{\mathcal{V}}^\perp$ such that $-(p, q) \in \partial\Phi(x, y)$. The convex subdifferential inequality then gives

$$\begin{aligned} \Phi(x_n, y_n) &\geq \Phi(x, y) + \langle -(p, q), (x_n, y_n) - (x, y) \rangle_{\mathcal{X} \times \mathcal{Y}} \\ &= \min_{\boldsymbol{\mathcal{V}}} \Phi - \langle (p, q), (x_n, y_n) \rangle_{\mathcal{X} \times \mathcal{Y}}. \end{aligned} \tag{24}$$

Let us prove that $\lim_{n \to +\infty} \langle (p, q), (x_n, y_n) \rangle_{\mathcal{X} \times \mathcal{Y}} = 0$. From Claim 4.5 the sequence (x_n, y_n) is bounded, hence it suffices to prove that 0 is the unique limit point of $\big(\langle (p, q), (x_n, y_n) \rangle_{\mathcal{X} \times \mathcal{Y}} \big)$. Let $\big(\langle (p, q), (x_{n_k}, y_{n_k}) \rangle_{\mathcal{X} \times \mathcal{Y}} \big)$ be a convergent subsequence. We can extract a subsequence of (x_{n_k}, y_{n_k}), still denoted by (x_{n_k}, y_{n_k}), which weakly converges toward $(\overline{x}, \overline{y})$. The weak lower semicontinuity of the function $(x, y) \mapsto \|Ax - By\|_{\mathcal{Z}}^2$ combined with Claim 4.6 implies that

$$\|A\overline{x} - B\overline{y}\|_{\mathcal{Z}}^2 \leq \liminf_{k \to +\infty} \|Ax_{n_k} - By_{n_k}\|_{\mathcal{Z}}^2 = \lim_{n \to +\infty} \|Ax_n - By_n\|_{\mathcal{Z}}^2 = 0,$$

hence $(\overline{x}, \overline{y}) \in \boldsymbol{\mathcal{V}}$. Recalling that $(p, q) \in \boldsymbol{\mathcal{V}}^\perp$, we infer that

$$\lim_{k \to +\infty} \langle (p, q), (x_{n_k}, y_{n_k}) \rangle_{\mathcal{X} \times \mathcal{Y}} = \langle (p, q), (\overline{x}, \overline{y}) \rangle_{\mathcal{X} \times \mathcal{Y}} = 0.$$

We immediately deduce that the whole sequence $\big(\langle (p, q), (x_n, y_n) \rangle_{\mathcal{X} \times \mathcal{Y}} \big)$ converges toward 0. Hence from (24) we obtain that $\liminf_{n \to +\infty} \Phi(x_n, y_n) \geq \min_{\boldsymbol{\mathcal{V}}} \Phi$. On the other hand, since $\Phi(x_n, y_n) \leq \varphi_n$, we have in view of Claim 4.7

$$\limsup_{n \to +\infty} \Phi(x_n, y_n) \leq \lim_{n \to +\infty} \varphi_n = \min_{\boldsymbol{\mathcal{V}}} \Phi.$$

We conclude that $\lim_{n \to +\infty} \Phi(x_n, y_n) = \min_{\boldsymbol{\mathcal{V}}} \Phi$. □

The proof of (i) relies on the Opial's lemma [29], that we recall for the sake of completeness.

Lemma 4.9 (Opial)**.** *Let $\mathcal{H}$ be a Hilbert space endowed with the norm N. Let (ξ_n) be a sequence of $\mathcal{H}$ such that there exists a nonempty set $\Xi \subset \mathcal{H}$ which verifies*

(a) *for all* $\xi \in \Xi$, $\lim_{n\to+\infty} N(\xi_n - \xi)$ *exists,*

(b) *if* $(\xi_{n_k}) \rightharpoonup \overline{\xi}$ *weakly in* $\mathcal{H}$ *as* $k \to +\infty$, *we have* $\overline{\xi} \in \Xi$.

Then the sequence (ξ_n) *weakly converges in* $\mathcal{H}$ *as* $n \to +\infty$ *toward a point of* Ξ.

Let us define the norm $N(u,v) = \left[\alpha\|u\|^2_{\mathcal{X}} + \nu\|v\|^2_{\mathcal{Y}} + \|Bv\|^2_{\mathcal{Z}}\right]^{1/2}$ on the space $\mathcal{X} \times \mathcal{Y}$. Since the linear operator B is continuous, the norm N is equivalent to the canonical norm on $\mathcal{X} \times \mathcal{Y}$. In view of Claim 4.5, the quantity $N(x_n - x, y_n - y)$ does have a limit as $n \to +\infty$ for every $(x,y) \in \operatorname{argmin}_{\boldsymbol{\mathcal{V}}}\Phi$, which shows point (a). Let (x_{n_k}, y_{n_k}) be a subsequence of (x_n, y_n) which weakly converges towards $(\overline{x}, \overline{y})$. The weak lower semicontinuity of the function $(x,y) \mapsto \|Ax - By\|^2_{\mathcal{Z}}$ combined with Claim 4.6 implies that

$$\|A\overline{x} - B\overline{y}\|^2_{\mathcal{Z}} \le \liminf_{k\to+\infty} \|Ax_{n_k} - By_{n_k}\|^2_{\mathcal{Z}} = \lim_{n\to+\infty} \|Ax_n - By_n\|^2_{\mathcal{Z}} = 0,$$

hence $(\overline{x}, \overline{y}) \in \boldsymbol{\mathcal{V}}$. In the same way, using Claim 4.8 and the weak lower semicontinuity of Φ, we infer that $(\overline{x}, \overline{y}) \in \operatorname{argmin}_{\boldsymbol{\mathcal{V}}}\Phi$. This shows point (b) of Opial's lemma and ends the proof of (i).

Let us now prove that $\lim_{n\to+\infty} f(x_n) = f(x_\infty)$. Using the weak lower semicontinuity of f, we have $f(x_\infty) \le \liminf_{n\to+\infty} f(x_n)$. On the other hand, we deduce from Claim 4.8 that

$$\begin{aligned}\limsup_{n\to+\infty} f(x_n) &= \limsup_{n\to+\infty}(f(x_n) + g(y_n) - g(y_n)) \\ &= f(x_\infty) + g(y_\infty) - \liminf_{n\to+\infty} g(y_n).\end{aligned}$$

By the weak lower semicontinuity of g, we have $g(y_\infty) \le \liminf_{n\to+\infty} g(y_n)$. We infer that $\limsup_{n\to+\infty} f(x_n) \le f(x_\infty)$, and finally $\lim_{n\to+\infty} f(x_n) = f(x_\infty)$. In the same way, we have $\lim_{n\to+\infty} g(y_n) = g(y_\infty)$, which ends the proof of (ii). □

5. FURTHER CONVERGENCE RESULTS FOR STRONGLY COUPLED PROBLEMS

In this section, we assume that $\mathcal{X} = \mathcal{Y} = \mathcal{Z}$ and that $A = B = \mathcal{I}$, along with $\alpha = \nu = 0$. Given closed convex functions $f, g : \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$, consider the following particular case[4] of algorithm $(\mathcal{A})$

$$(\mathcal{A}_0) \qquad \begin{cases} x_{n+1} = \operatorname{argmin}\left\{\gamma_{n+1} f(\zeta) + \frac{1}{2}\|\zeta - y_n\|^2_{\mathcal{X}};\ \zeta \in \mathcal{X}\right\} \\ y_{n+1} = \operatorname{argmin}\left\{\gamma_{n+1} g(\eta) + \frac{1}{2}\|x_{n+1} - \eta\|^2_{\mathcal{X}};\ \eta \in \mathcal{X}\right\}. \end{cases}$$

Using the same notations as in the previous sections, we have

$$\boldsymbol{\mathcal{V}} = \{(x,x); x \in \mathcal{X}\} \quad \text{and} \quad \operatorname{argmin}_{\boldsymbol{\mathcal{V}}}\Phi = \{(x,x); x \in \operatorname{argmin}(f+g)\}.$$

Let us first start with an example.

[4]See Remark 2.1, where algorithm $(\mathcal{A}_0)$ has been introduced in the framework of maximal monotone operators.

Example 5.1. Take $\mathcal{X} = \mathbb{R}$ and define the functions f, $g : \mathbb{R} \to \mathbb{R}$ respectively by $f(x) = \frac{1}{2}(x-1)^2$ and $g(y) = \frac{1}{2}(y+1)^2$. We then have $\operatorname{argmin}(f+g) = \{0\}$. By writing down the optimality conditions for algorithm $(\mathcal{A}_0)$, we immediately obtain the following recurrence formulae (see also Remark 2.1)

$$\begin{cases} \gamma_{n+1}(x_{n+1}-1) + x_{n+1} - y_n = 0 \\ \gamma_{n+1}(y_{n+1}+1) + y_{n+1} - x_{n+1} = 0. \end{cases}$$

We infer that

$$y_{n+1} = \frac{1}{(1+\gamma_{n+1})^2} y_n - \frac{\gamma_{n+1}^2}{(1+\gamma_{n+1})^2}.$$

Let us set $a_n = \frac{1}{(1+\gamma_{n+1})^2}$ and $b_n = \frac{\gamma_{n+1}^2}{(1+\gamma_{n+1})^2}$. We deduce from the above equality that $|y_{n+1}| \le a_n |y_n| + b_n$. To prove the convergence of the sequence (y_n), we use the following lemma borrowed from [31, Lemma 3, p. 45].

Lemma 5.2. *Let (a_n) and (b_n) be real sequences such that $0 \le a_n < 1$ and $b_n \ge 0$ for every $n \in \mathbb{N}$. Assume moreover that $(1-a_n) \notin l^1$ and that $\lim_{n\to+\infty} \frac{b_n}{1-a_n} = 0$. Let (u_n) be a real sequence such that $u_{n+1} \le a_n u_n + b_n$ for every $n \in \mathbb{N}$. Then we have $\limsup_{n\to+\infty} u_n \le 0$.*

It is easy to check that, if the sequence (γ_n) is not summable, then the sequence $(1-a_n)$ is not summable. Moreover we have $\lim_{n\to+\infty} \frac{b_n}{1-a_n} = \lim_{n\to+\infty} \frac{\gamma_{n+1}}{2+\gamma_{n+1}} = 0$. Thus the previous lemma implies that $\limsup_{n\to+\infty} |y_n| \le 0$, hence $\lim_{n\to+\infty} y_n = 0$. Finally we have proved that if $(\gamma_n) \notin l^1$ then $\lim_{n\to+\infty}(x_n, y_n) = (0,0)$.

It is worthwhile noticing that the assumption $(\gamma_n) \in l^2$ does not come into play in the above example. This is in fact a consequence of a general result that will be brought to light by Theorem 5.4 (i), see also Remark 5.5. Before stating Theorem 5.4, we need the following preliminary result.

Proposition 5.3. *Let f, $g : \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ be closed convex functions which are bounded from below and such that $\operatorname{dom} f \cap \operatorname{dom} g \neq \emptyset$. Let (γ_n) be a positive nonincreasing sequence such that $\lim_{n\to\infty} \gamma_n = 0$. Then any sequence (x_n, y_n) generated by $(\mathcal{A}_0)$ satisfies $\lim_{n\to+\infty} \|x_n - y_n\|_{\mathcal{X}} = 0$.*

Proof. Let us define the sequence (ψ_n) by

$$\psi_n = \gamma_n\big(f(x_n) + g(y_n)\big) + \frac{1}{2}\|x_n - y_n\|_{\mathcal{X}}^2. \tag{25}$$

We have

$$\psi_n \ge \gamma_n \inf \Phi + \frac{1}{2}\|x_n - y_n\|_{\mathcal{X}}^2, \tag{26}$$

hence the sequence $(\psi_n - \gamma_n \inf \Phi)$ is nonnegative. By using inequality (19) with $A = B = \mathcal{I}$, we deduce that, for every $n \in \mathbb{N}$,

$$\psi_{n+1} - \psi_n \le (\gamma_{n+1} - \gamma_n) \inf \Phi.$$

This shows that the sequence $(\psi_n - \gamma_n \inf \Phi)$ is nonincreasing, hence convergent. Since $\lim_{n\to+\infty} \gamma_n = 0$, the sequence (ψ_n) also converges. Let us apply inequality (18) with $A = B = \mathcal{I}$, $\alpha = \nu = 0$ and $x = y$; we find for all $x \in \operatorname{dom} f \cap \operatorname{dom} g \neq \emptyset$,

$$2\psi_{n+1} - 2\gamma_{n+1}(f(x) + g(x)) \le \|y_n - x\|_{\mathcal{X}}^2 - \|y_{n+1} - x\|_{\mathcal{X}}^2.$$

By summing the above inequalities for $n = 0, ..., N$, we obtain

$$2\sum_{n=0}^{N}\left[\psi_{n+1} - \gamma_{n+1}(f(x)+g(x))\right] \le \|y_0 - x\|^2_{\mathcal{X}}.$$

Since this is true for every $N \in \mathbb{N}$, we derive that

$$\liminf_{n\to+\infty}\left[\psi_{n+1} - \gamma_{n+1}(f(x)+g(x))\right] \le 0.$$

But both terms are convergent, so we have $\lim_{n\to+\infty}\psi_n \le 0$. From (26) we immediately deduce that $\lim\limits_{n\to+\infty}\|x_n - y_n\|^2_{\mathcal{X}} = 0$. $\square$

The approach that we now develop relies on topological ingredients that can already be found in [5, 9, 16, 21]. The result below shows that if $(\gamma_n) \notin l^1$ and $\lim\limits_{n\to+\infty}\gamma_n = 0$, the iterates x_n, y_n of algorithm $(\mathcal{A}_0)$ approach the set $\operatorname{argmin}(f+g)$ as $n \to +\infty$. Weak convergence is obtained under the extra assumption $(\gamma_n) \in l^2$. In the next statement, we denote by $d\big(\cdot, \operatorname{argmin}(f+g)\big)$ the distance function to the set $\operatorname{argmin}(f+g)$.

Theorem 5.4. *Let $f, g : \mathcal{X} \to \mathbb{R}\cup\{+\infty\}$ be closed convex functions which are bounded from below and such that $\operatorname{dom} f \cap \operatorname{dom} g \ne \emptyset$. Assume that either f or g is inf-compact[5]. Let (γ_n) be a positive nonincreasing sequence such that $(\gamma_n) \notin l^1$ and $\lim\limits_{n\to+\infty}\gamma_n = 0$. Finally, let (x_n, y_n) be a sequence generated by $(\mathcal{A}_0)$. Then*

(i) $\lim\limits_{n\to+\infty} d_{\mathcal{X}}\big(x_n, \operatorname{argmin}(f+g)\big) = \lim\limits_{n\to+\infty} d_{\mathcal{X}}\big(y_n, \operatorname{argmin}(f+g)\big) = 0.$

(ii) *If $(\gamma_n) \in l^2$, and if condition (QC) is satisfied[6], then the sequence (x_n, y_n) converges weakly to a point $(\overline{x}, \overline{x})$ with $\overline{x} \in \operatorname{argmin}(f+g)$.*

(iii) *If moreover the sequence $\left(\frac{1}{\gamma_{n+1}} - \frac{1}{\gamma_n}\right)$ is majorized by some $M > 0$, then the sequence (x_n, y_n) converges strongly in $\mathcal{X}$.*

Proof. Without loss of generality, we can assume that the function f is inf-compact. Since the function g is bounded from below, we derive that the function $f+g$ is inf-compact. On the other hand, the assumption $\operatorname{dom} f \cap \operatorname{dom} g \ne \emptyset$ ensures that $f+g$ is a proper function. It follows that $\operatorname{argmin}(f+g) \ne \emptyset$.
(i) In view of Proposition 5.3, it suffices to prove that $\lim\limits_{n\to+\infty} d_{\mathcal{X}}\big(y_n, \operatorname{argmin}(f+g)\big) = 0$. Set $A = B = \mathcal{I}$ and $\alpha = \nu = 0$ in inequality (14) to deduce that for every $y \in \operatorname{argmin}(f+g)$ and every $n \in \mathbb{N}$ we have

$$\|y_{n+1} - y\|^2_{\mathcal{X}} - \|y_n - y\|^2_{\mathcal{X}} + 2\gamma_{n+1}\left(\Phi(x_{n+1}, y_{n+1}) - \min_{\mathcal{V}}\Phi\right) + \|x_{n+1} - y_{n+1}\|^2_{\mathcal{X}} \le 0. \tag{27}$$

[5] Recall that a function is said to be inf-compact if its sublevel sets are relatively compact.

[6] In our present setting, it is easy to check that condition (QC) is satisfied if and only if

$$\bigcup_{\lambda>0}\lambda(\operatorname{dom} f - \operatorname{dom} g) \text{ is a closed subspace of } \mathcal{X}.$$

This is precisely the Attouch-Brézis condition, which ensures that $\partial f + \partial g = \partial(f+g)$ and hence $(\partial f + \partial g)^{-1}0 = \operatorname{argmin}(f+g)$.

Let P denote the projection operator onto the closed convex set $\operatorname{argmin}(f+g)$ and take $y = P(y_n)$. Setting $u_n = d^2_{\mathcal{X}}\big(y_n, \operatorname{argmin}(f+g)\big)$, we derive from (27) that

$$u_{n+1} - u_n + 2\gamma_{n+1}\left(\Phi(x_{n+1}, y_{n+1}) - \min_{\mathcal{V}} \Phi\right) \leq 0. \tag{28}$$

We now follow the same arguments as those used by the second author in [21, Theorem 3.1]. We distinguish two cases:

(a) There exists $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$, $\Phi(x_n, y_n) > \min_{\mathcal{V}} \Phi$.
(b) For all $n_0 \in \mathbb{N}$, there exists $n \geq n_0$ such that $\Phi(x_n, y_n) \leq \min_{\mathcal{V}} \Phi$.

Case (a). Assume there exists $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$, $\Phi(x_n, y_n) > \min_{\mathcal{V}} \Phi$. From inequality (28), we deduce that the sequence $(u_n)_{n \geq n_0}$ is nonincreasing and convergent. We must prove that $\lim_{n\to+\infty} u_n = 0$. Using again inequality (28), we can assert that the sequence $(\gamma_n(\Phi(x_n, y_n) - \min_{\mathcal{V}} \Phi))$ is summable. Moreover, since $(\gamma_n) \notin l^1$, we have $\liminf_{n\to+\infty} \Phi(x_n, y_n) = \min_{\mathcal{V}} \Phi$. Consider a subsequence of (x_n, y_n), still denoted by (x_n, y_n), such that $\lim\limits_{n\to+\infty} \Phi(x_n, y_n) = \min_{\mathcal{V}} \Phi$. Since the function g is bounded from below, the sequence $(f(x_n))$ is majorized. Using the inf-compactness of the map f, we obtain that the sequence (x_n) is relatively compact in $\mathcal{X}$. Thus there exist a subsequence (x_{n_k}) along with $\overline{x} \in \mathcal{X}$ such that $\lim\limits_{k\to+\infty} x_{n_k} = \overline{x}$ strongly in $\mathcal{X}$. In view of Proposition 5.3 we also have $\lim\limits_{k\to+\infty} y_{n_k} = \overline{x}$ strongly in $\mathcal{X}$. The closedness of the function Φ allows to assert that $\Phi(\overline{x}, \overline{x}) \leq \liminf_{k\to+\infty} \Phi(x_{n_k}, y_{n_k}) = \min_{\mathcal{V}} \Phi$. Hence $(\overline{x}, \overline{x}) \in \operatorname{argmin}_{\mathcal{V}} \Phi$, *i.e.* $\overline{x} \in \operatorname{argmin}(f+g)$. Thus

$$\lim_{k\to+\infty} u_{n_k} = \lim_{n\to+\infty} d^2_{\mathcal{X}}(y_{n_k}, \operatorname{argmin}(f+g)) = d^2_{\mathcal{X}}(\overline{x}, \operatorname{argmin}(f+g)) = 0.$$

Recalling that the sequence (u_n) is convergent, we conclude that $\lim\limits_{n\to+\infty} u_n = 0$.

Case (b). We assume that for all $n_0 \in \mathbb{N}$ there exists $n \geq n_0$ such that $\Phi(x_n, y_n) \leq \min_{\mathcal{V}} \Phi$. Let us define

$$\tau_N = \max\{n \in \mathbb{N}, n \leq N \text{ and } \Phi(x_n, y_n) \leq \min_{\mathcal{V}} \Phi\}.$$

The integer τ_N is well-defined for N large enough and $\lim_{N\to+\infty} \tau_N = \infty$. If $\tau_N < N$ inequality (28) implies $u_{n+1} \leq u_n$ whenever $\tau_N \leq n \leq N-1$. In particular,

$$u_N \leq u_{\tau_N}. \tag{29}$$

Notice that if $\tau_N = N$ this inequality is still true. Therefore it suffices to prove that $\lim\limits_{n\to+\infty} u_{\tau_n} = 0$. First observe that $\Phi(x_{\tau_n}, y_{\tau_n}) \leq \min_{\mathcal{V}} \Phi$ for all sufficiently large n by definition. We deduce, as before, that the sequence (x_{τ_n}) is relatively compact, hence bounded in $\mathcal{X}$. In view of Proposition 5.3, the sequence (y_{τ_n}) is also bounded in $\mathcal{X}$, whence the boundedness of the real sequence (u_{τ_n}). The proof will be complete if we verify that every convergent subsequence of (u_{τ_n}) must vanish. Indeed, assume that $\lim\limits_{k\to+\infty} u_{\tau_{n_k}}$ exists. We may assume, upon passing to a subsequence if necessary, that $\lim_{k\to+\infty} x_{\tau_{n_k}} = \lim_{k\to+\infty} y_{\tau_{n_k}} = \overline{x}$ for some $\overline{x} \in \mathcal{X}$. The closedness of Φ then gives

$$\Phi(\overline{x}, \overline{x}) \leq \liminf_{k\to\infty} \Phi(x_{\tau_{n_k}}, y_{\tau_{n_k}}) \leq \min_{\mathcal{V}} \Phi,$$

which implies $(\overline{x}, \overline{x}) \in \operatorname{argmin}_{\boldsymbol{\mathcal{V}}} \Phi$. As before, this implies $\lim\limits_{k \to +\infty} u_{\tau_{n_k}} = 0$ and we deduce that the whole sequence (u_{τ_n}) converges toward 0. Then inequality (29) shows that $\lim_{n \to +\infty} u_n = 0$.

(ii) Let us assume that $(\gamma_n) \in l^2$ and that condition (QC) is satisfied. Observe that $R(A) + R(B) = R(\mathcal{I}) = \mathcal{X}$, so the closedness of the space $R(A) + R(B)$ is fulfilled. By Proposition 4.3, the sequence (ω_n) defined by

$$\omega_n = \inf_{(x,y) \in \mathcal{X}^2} \left\{ \frac{1}{2} \|x - y\|^2_{\mathcal{X}} + \gamma_n \left(f(x) + g(y) - \min_{\boldsymbol{\mathcal{V}}} \Phi \right) \right\}$$

satisfies $(\omega_n^-) \in l^1$. Let $y \in \operatorname{argmin}(f+g)$. From inequality (27), we obtain

$$\|y_{n+1} - y\|^2_{\mathcal{X}} - \|y_n - y\|^2_{\mathcal{X}} \le 2\omega^-_{n+1}.$$

Since $(\omega_n^-) \in l^1$ this implies in view of Lemma 3.3 that

$$\forall y \in \operatorname{argmin}(f+g), \quad \lim_{n \to +\infty} \|y_n - y\|^2_{\mathcal{X}} \text{ exists.} \tag{30}$$

On the other hand, recalling that $\lim\limits_{n \to +\infty} d_{\mathcal{X}}(y_n, \operatorname{argmin}(f+g)) = 0$, every weak cluster point of the sequence (y_n) lies in $\operatorname{argmin}(f+g)$. We infer from Lemma 4.9 that the sequence (y_n) weakly converges toward some point in $\operatorname{argmin}(f+g)$. Finally Proposition 5.3 shows that the sequences (x_n) and (y_n) tend weakly toward the same limit.

(iii) Let us first prove that the sequence (φ_n) defined by formula (13) is bounded. By applying inequality (18) with $A = B = \mathcal{I}$, $\alpha = \nu = 0$ and $(x,y) = (x_n, y_n)$, we easily find

$$\varphi_{n+1} - \varphi_n + \frac{1}{2\,\gamma_{n+1}} \left(\|x_{n+1} - x_n\|^2_{\mathcal{X}} + \|y_{n+1} - y_n\|^2_{\mathcal{X}} \right) \le \frac{1}{2} \left(\frac{1}{\gamma_{n+1}} - \frac{1}{\gamma_n} \right) \|x_n - y_n\|^2_{\mathcal{X}}. \tag{31}$$

Observe that this inequality is slightly more precise than (15), where two terms were omitted. Since $\frac{1}{\gamma_{n+1}} - \frac{1}{\gamma_n} \le M$ by assumption and since $\|x_n - y_n\|^2_{\mathcal{X}} \le 2\,\|x_{n+1} - y_n\|^2_{\mathcal{X}} + 2\,\|x_{n+1} - x_n\|^2_{\mathcal{X}}$, inequality (31) implies

$$\varphi_{n+1} - \varphi_n + \frac{1}{2\,\gamma_{n+1}} \|x_{n+1} - x_n\|^2_{\mathcal{X}} \le M \left(\|x_{n+1} - y_n\|^2_{\mathcal{X}} + \|x_{n+1} - x_n\|^2_{\mathcal{X}} \right).$$

From the fact that $\lim_{n \to +\infty} \gamma_n = 0$, we immediately derive that for n large enough

$$\varphi_{n+1} - \varphi_n \le M\,\|x_{n+1} - y_n\|^2_{\mathcal{X}}. \tag{32}$$

Recall that the sequence (ω_n^-) is summable, see the proof of (ii). The summability of $(\|x_{n+1} - y_n\|^2_{\mathcal{X}})$ is then an immediate consequence of inequality (21), with $A = B = \mathcal{I}$ and $\alpha = \nu = 0$. In view of (32), we infer from Lemma 3.3 that the sequence (φ_n) is convergent, hence bounded. Since the function g is bounded from below, the sequence $(f(x_n))$ is majorized. The inf-compactness of f allows to deduce that the sequence (x_n) is relatively compact in $\mathcal{X}$. Hence there exists $\overline{x} \in \mathcal{X}$ along with a subsequence (x_{n_k}) such that $\lim_{k \to +\infty} x_{n_k} = \overline{x}$ strongly in $\mathcal{X}$. From Proposition 5.3, we also have $\lim_{k \to +\infty} y_{n_k} = \overline{x}$ strongly in $\mathcal{X}$. In view of (i), it is clear that $\overline{x} \in \operatorname{argmin}(f+g)$. Taking $y = \overline{x}$ in assertion (30), we deduce that $\lim_{n\to+\infty} \|y_n - \overline{x}\|_{\mathcal{X}} = 0$. Owing to Proposition 5.3, we conclude that $\lim_{n\to+\infty} x_n = \lim_{n\to+\infty} y_n = \overline{x}$ strongly in $\mathcal{X}$. □

Remark 5.5. Observe that if $\operatorname{argmin}(f+g) = \{\overline{\xi}\}$, Theorem 5.4 (i) shows that any sequence generated by $(\mathcal{A}_0)$ converges strongly to $(\overline{\xi}, \overline{\xi})$, even if $(\gamma_n) \notin l^2$.

Remark 5.6. No qualification condition is required in the proof of Theorem 5.4 (i), which is a distinctive mark with respect to the proof of Theorem 4.4 (see specially Claim 4.8).

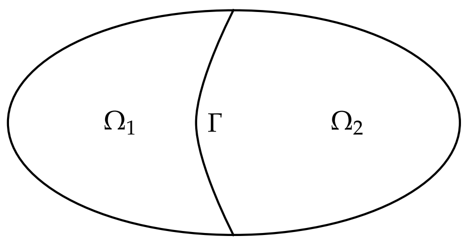
Remark 5.7. Recall from Remark 2.1 that the iterates of algorithm $(\mathcal{A}_0)$ satisfy the following equalities

$$\begin{aligned} x_{n+1} &= (I + \gamma_{n+1}\,\partial f)^{-1}(I + \gamma_n\,\partial g)^{-1}\,x_n, \\ y_{n+1} &= (I + \gamma_{n+1}\,\partial g)^{-1}(I + \gamma_{n+1}\,\partial f)^{-1}\,y_n. \end{aligned}$$

This corresponds to a double resolvent scheme studied by Passty in [30]. In this reference, weak ergodic convergence of such sequences is established for general maximal monotone operators such that the sum is itself maximal, provided that $(\gamma_n) \in l^2 \setminus l^1$. Under some inf-compactness assumption, Theorem 5.4 (ii) (resp. (iii)) shows that weak ergodic convergence is replaced by weak (resp. strong) convergence in the subdifferential framework. Hence our result is an improvement of Passty theorem when applied to subdifferential operators.

6. Application to domain decomposition for PDE's

Let us consider a bounded domain $\Omega \subset \mathbb{R}^N$ with $\mathcal{C}^2$ boundary. Assume that the set Ω is decomposed in two nonoverlapping Lipschitz subdomains Ω_1 and Ω_2 with a common interface Γ. This situation is illustrated in the next figure.



6.1. **Neumann problem.** Given a function $h \in L^2(\Omega)$, let us consider the following Neumann boundary value problem on Ω

$$\begin{cases} -\Delta w &= h \quad \text{on} \quad \Omega \\ \frac{\partial w}{\partial n} &= 0 \quad \text{on} \quad \partial\Omega, \end{cases}$$

where $\frac{\partial w}{\partial n} = \nabla w.\vec{n}$ and $\vec{n}$ is the unit outward normal to $\partial\Omega$. We assume that $\int_\Omega h = 0$, which is a necessary and sufficient condition for the existence of a solution. The weak solutions of the above Neumann problem satisfy the following minimization problem

$$\min\left\{\frac{1}{2}\int_\Omega |\nabla w|^2 - \int_\Omega hw; \quad w \in H^1(\Omega)\right\}, \tag{33}$$

see for example [8, 19, 26, 32]. Moreover, denoting by $\widehat{w}$ a particular solution, the solution set of (33) is of the form $\{\widehat{w} + C,\ C \in \mathbb{R}\}$. Assuming that Ω is of class $\mathcal{C}^2$, we know from the regularity theory of weak solutions that $\widehat{w} \in H^2(\Omega)$, see for instance [3, 4, 25]. Notice that, if $w \in H^1(\Omega)$ then the restrictions $u = w_{|\Omega_1}$ and $v = w_{|\Omega_2}$ belong respectively to $H^1(\Omega_1)$ and $H^1(\Omega_2)$ and moreover $u_{|\Gamma} = v_{|\Gamma}$. Conversely, if $u \in H^1(\Omega_1)$, $v \in H^1(\Omega_2)$ and if $u_{|\Gamma} = v_{|\Gamma}$, then the function w defined by $w = \begin{cases} u \text{ on } \Omega_1 \\ v \text{ on } \Omega_2 \end{cases}$ belongs to $H^1(\Omega)$. As a consequence, problem (33) can be reformulated as

$$(\mathcal{P}) \qquad \min\left\{f(u) + g(v); \quad (u,v) \in H^1(\Omega_1) \times H^1(\Omega_2) \text{ and } u_{|\Gamma} = v_{|\Gamma}\right\},$$

where

$$f(u) = \frac{1}{2}\int_{\Omega_1} |\nabla u|^2 - \int_{\Omega_1} hu \quad \text{and} \quad g(v) = \frac{1}{2}\int_{\Omega_2} |\nabla v|^2 - \int_{\Omega_2} hv. \tag{34}$$

Let us show how the algorithm $(\mathcal{A})$ can be applied so as to solve problem $(\mathcal{P})$. The set $\mathcal{X} = H^1(\Omega_1)$ is equipped with the scalar product $\langle u_1, u_2\rangle_{\mathcal{X}} = \int_{\Omega_1}(\nabla u_1.\nabla u_2 + u_1 u_2)$ and the corresponding norm. The same holds for $\mathcal{Y} = H^1(\Omega_2)$ by replacing Ω_1 with Ω_2. The set $\mathcal{Z} = L^2(\Gamma)$ is equipped with the scalar product $\langle z_1, z_2\rangle_{\mathcal{Z}} = \int_\Gamma z_1 z_2$ and the corresponding norm. The operators $A : \mathcal{X} \to \mathcal{Z}$ and $B : \mathcal{Y} \to \mathcal{Z}$ are respectively the trace operators on Γ, which are well-defined by the Lipschitz character of the boundaries of Ω_1 and Ω_2 (see [15, Theorem II.46] or [27, Theorem 2]). Algorithm $(\mathcal{A})$ runs as follows

$$\begin{cases} u_{n+1} = \operatorname{argmin}\{\gamma_{n+1} f(u) + \frac{1}{2}\|Au - Bv_n\|^2_{\mathcal{Z}} + \frac{\alpha}{2}\|u - u_n\|^2_{\mathcal{X}}; \quad u \in \mathcal{X}\} \\ v_{n+1} = \operatorname{argmin}\{\gamma_{n+1} g(v) + \frac{1}{2}\|Au_{n+1} - Bv\|^2_{\mathcal{Z}} + \frac{\nu}{2}\|v - v_n\|^2_{\mathcal{Y}}; \quad v \in \mathcal{Y}\}, \end{cases}$$

where α and ν are fixed positive parameters. An elementary directional derivative computation shows that the weak variational formulation of algorithm $(\mathcal{A})$ is given by

$$\begin{aligned} \gamma_{n+1}\int_{\Omega_1} \nabla u_{n+1}.\nabla u + \alpha\int_{\Omega_1}(\nabla u_{n+1} - \nabla u_n).\nabla u & \\ + \alpha\int_{\Omega_1}(u_{n+1} - u_n)u + \int_\Gamma (Au_{n+1} - Bv_n)Au &= \gamma_{n+1}\int_{\Omega_1} hu \end{aligned}$$

and

$$\begin{aligned} \gamma_{n+1}\int_{\Omega_2} \nabla v_{n+1}.\nabla v + \nu\int_{\Omega_2}(\nabla v_{n+1} - \nabla v_n).\nabla v & \\ + \nu\int_{\Omega_2}(v_{n+1} - v_n)v + \int_\Gamma (Bv_{n+1} - Au_{n+1})Bv &= \gamma_{n+1}\int_{\Omega_2} hv \end{aligned}$$

for all $u \in \mathcal{X}$ and $v \in \mathcal{Y}$. These are the variational weak formulations of the following mixed Dirichlet-Neumann boundary value problems respectively on Ω_1

$$\begin{cases} -(\gamma_{n+1}+\alpha)\Delta u_{n+1}+\alpha u_{n+1} &= \gamma_{n+1}h-\alpha\Delta u_n+\alpha u_n \quad \text{on} \quad \Omega_1 \\ (\gamma_{n+1}+\alpha)\frac{\partial u_{n+1}}{\partial n} &= \alpha\frac{\partial u_n}{\partial n} \quad \text{on} \quad \partial\Omega_1\cap\partial\Omega \\ (\gamma_{n+1}+\alpha)\frac{\partial u_{n+1}}{\partial n}+u_{n+1} &= \alpha\frac{\partial u_n}{\partial n}+v_n \quad \text{on} \quad \Gamma, \end{cases}$$

and Ω_2

$$\begin{cases} -(\gamma_{n+1}+\nu)\Delta v_{n+1}+\nu v_{n+1} &= \gamma_{n+1}h-\nu\Delta v_n+\nu v_n \quad \text{on} \quad \Omega_2 \\ (\gamma_{n+1}+\nu)\frac{\partial v_{n+1}}{\partial n} &= \nu\frac{\partial v_n}{\partial n} \quad \text{on} \quad \partial\Omega_2\cap\partial\Omega \\ (\gamma_{n+1}+\nu)\frac{\partial v_{n+1}}{\partial n}+v_{n+1} &= \nu\frac{\partial v_n}{\partial n}+u_{n+1} \quad \text{on} \quad \Gamma. \end{cases}$$

Let us now check the validity of the assumptions of Theorem 4.4. The qualification condition (QC) is automatically satisfied since $\operatorname{dom} f = \mathcal{X}$ and $\operatorname{dom} g = \mathcal{Y}$. In view of Proposition 4.3, assumption $(\omega_n^-) \in l^1$ is verified[7] if $(\gamma_n) \in l^2$ and if there exist $(\widehat{u}, \widehat{v}) \in \mathcal{X}\times\mathcal{Y}$ such that $\widehat{u}_{|\Gamma} = \widehat{v}_{|\Gamma}$ along with $z \in \mathcal{Z}$ satisfying

$$-A^*z \in \partial f(\widehat{u}) \quad \text{and} \quad B^*z \in \partial g(\widehat{v}). \tag{35}$$

Take $\widehat{u} = \widehat{w}_{|\Omega_1}$ and $\widehat{v} = \widehat{w}_{|\Omega_2}$ the restrictions of $\widehat{w}$ respectively to Ω_1 and Ω_2. Let us multiply the equality $-\Delta\widehat{u} = h$ by $u \in H^1(\Omega_1)$ and integrate on Ω_1. Using Green's formula and the fact that $\frac{\partial \widehat{u}}{\partial n} = 0$ on $\partial\Omega\cap\partial\Omega_1$, we obtain

$$\forall u \in H^1(\Omega_1), \quad \int_{\Omega_1} \nabla\widehat{u}.\nabla u - \int_\Gamma \frac{\partial\widehat{u}}{\partial n}u = \int_{\Omega_1} hu.$$

Hence we deduce that for every $u \in H^1(\Omega_1)$,

$$f(u) = \frac{1}{2}\int_{\Omega_1}|\nabla u|^2 - \int_{\Omega_1}\nabla\widehat{u}.\nabla u + \int_\Gamma \frac{\partial\widehat{u}}{\partial n}\,u$$

and therefore

$$\begin{aligned} f(u) - f(\widehat{u}) &= \frac{1}{2}\int_{\Omega_1}|\nabla u - \nabla\widehat{u}|^2 + \int_\Gamma \frac{\partial\widehat{u}}{\partial n}(u-\widehat{u}) \\ &\geq \int_\Gamma \frac{\partial\widehat{u}}{\partial n}(u-\widehat{u}) = \left\langle A^*\frac{\partial\widehat{u}}{\partial n}, u-\widehat{u}\right\rangle_{\mathcal{X}}. \end{aligned}$$

This shows that $A^*\frac{\partial\widehat{u}}{\partial n} \in \partial f(\widehat{u})$ and we find in the same way $B^*\frac{\partial\widehat{v}}{\partial n} \in \partial g(\widehat{v})$. Since $\frac{\partial\widehat{u}}{\partial n}_{|\Gamma} = -\frac{\partial\widehat{v}}{\partial n}_{|\Gamma}$, condition (35) is proved with $z = \frac{\partial\widehat{v}}{\partial n}$, which belongs to $L^2(\Gamma)$ since $\widehat{v} \in H^2(\Omega_2)$.

We conclude from Theorem 4.4 (i) and the preceding argument that if $\left(\frac{1}{\gamma_{n+1}} - \frac{1}{\gamma_n}\right)$ is bounded from above and if $(\gamma_n) \in l^2$, then any sequence (u_n, v_n) generated by

[7] Observe that we have $R(A) = R(B) = H^{1/2}(\Gamma)$. Hence the set $R(A)+R(B) = H^{1/2}(\Gamma)$ is dense in $\mathcal{Z} = L^2(\Gamma)$ and condition $(\omega_n^-) \in l^1$ cannot be verified by using assertion (i) of proposition 4.3. This remark may suggest to take $\mathcal{Z} = H^{1/2}(\Gamma)$ endowed with the corresponding norm. In this case, the closedness of the set $R(A)+R(B)$ is automatically ensured. However the practical implementation of algorithm $(\mathcal{A})$ will be more complicated due to the use of the $H^{1/2}(\Gamma)$ norm. The details are out of the scope of the paper.

$(\mathcal{A})$ weakly converges in $H^1(\Omega_1) \times H^1(\Omega_2)$ to a minimum point $(\widehat{u}+C, \widehat{v}+C)$, $(C \in \mathbb{R})$ of problem $(\mathcal{P})$. Without loss of generality, we can assume that $C = 0$. Since Ω_1 and Ω_2 are Lipschitz domains, the injections $H^1(\Omega_1) \hookrightarrow L^2(\Omega_1)$ and $H^1(\Omega_2) \hookrightarrow L^2(\Omega_2)$ are compact by the Rellich-Kondrachov Theorem (see [2, Theorem 6.2] or [15, Theorem II.55]). It ensues that the sequence (u_n, v_n) converges to $(\widehat{u}, \widehat{v})$ strongly in $L^2(\Omega_1) \times L^2(\Omega_2)$. Moreover, from Theorem 4.4 (ii), we have $\lim_{n\to+\infty} f(u_n) = f(\widehat{u})$ and $\lim_{n\to+\infty} g(v_n) = g(\widehat{v})$, hence $\lim_{n\to+\infty} \int_{\Omega_1} |\nabla u_n|^2 = \int_{\Omega_1} |\nabla \widehat{u}|^2$ and $\lim_{n\to+\infty} \int_{\Omega_2} |\nabla v_n|^2 = \int_{\Omega_2} |\nabla \widehat{v}|^2$. As a consequence, we have

$$\lim_{n\to+\infty} \|(u_n, v_n)\|_{H^1(\Omega_1)\times H^1(\Omega_2)} = \|(\widehat{u}, \widehat{v})\|_{H^1(\Omega_1)\times H^1(\Omega_2)}.$$

Since (u_n, v_n) weakly converges in $H^1(\Omega_1) \times H^1(\Omega_2)$ toward $(\widehat{u}, \widehat{v})$, the convergence is strong in $H^1(\Omega_1) \times H^1(\Omega_2)$. We have proved the following:

Theorem 6.1. *Let $\Omega \subset \mathbb{R}^N$ be a bounded domain which can be decomposed in two nonoverlapping Lipschitz subdomains Ω_1 and Ω_2 with a common interface Γ. We assume that the set Ω is of class $\mathcal{C}^2$. Let $h \in L^2(\Omega)$ be such that $\int_\Omega h = 0$ and define the functions $f : H^1(\Omega_1) \to \mathbb{R}$ and $g : H^1(\Omega_2) \to \mathbb{R}$ by formulas (34). Assume that (γ_n) is a positive sequence such that $(\gamma_n) \in l^2$ and the sequence $\left(\frac{1}{\gamma_{n+1}} - \frac{1}{\gamma_n}\right)$ is bounded from above. Then any sequence (u_n, v_n) generated by algorithm $(\mathcal{A})$ strongly converges in $H^1(\Omega_1) \times H^1(\Omega_2)$ and the limit $(\widehat{u}, \widehat{v})$ is such that the map $\widehat{w} = \begin{cases} \widehat{u} \text{ on } \Omega_1 \\ \widehat{v} \text{ on } \Omega_2 \end{cases}$ is a solution of the Neumann problem (33).*

Algorithm $(\mathcal{A})$ allows to solve the initial Neumann problem on Ω by solving separately mixed Dirichlet-Neumann problems on Ω_1 and Ω_2. A similar method is developped in [13], where the authors consider alternating minimization algorithms based on augmented Lagrangian approach.

6.2. **Problem with an obstacle.** As a model situation, let us consider the variational problem with an obstacle constraint

$$\min\left\{\frac{1}{2}\int_\Omega |\nabla w|^2 - \int_\Omega hw; \quad w \in H^1(\Omega),\ w \geq 0 \text{ on } \Omega\right\}. \tag{36}$$

It can be cast into our framework by taking

$$f(u) = \frac{1}{2}\int_{\Omega_1} |\nabla u|^2 - \int_{\Omega_1} hu + \delta_{C_1}(u) \quad \text{and} \quad g(v) = \frac{1}{2}\int_{\Omega_2} |\nabla v|^2 - \int_{\Omega_2} hv + \delta_{C_2}(v),$$

where δ_{C_1} is the indicator function of the convex set $C_1 = \{u \geq 0; \quad u \in H^1(\Omega_1)\}$ and δ_{C_2} is the indicator function of the convex set $C_2 = \{v \geq 0; \quad v \in H^1(\Omega_2)\}$. Problem (36) can be reformulated as

$$(\mathcal{P}) \qquad \min\left\{f(u) + g(v); \quad (u, v) \in H^1(\Omega_1) \times H^1(\Omega_2) \text{ and } u_{|\Gamma} = v_{|\Gamma}\right\}.$$

Let us show that Attouch-Brézis qualification condition (QC) is satisfied in this situation (by contrast with Moreau-Rockafellar condition which fails to be satisfied for $N \geq 2$). Indeed, we are going to verify that

$$\operatorname{dom} f \times \operatorname{dom} g - \boldsymbol{\mathcal{V}} = H^1(\Omega_1) \times H^1(\Omega_2).$$

To that end we introduce two trace lifting operators

$$\begin{aligned} r_1 : \quad & H^{1/2}(\Gamma) \to H^1(\Omega_1) \\ r_2 : \quad & H^{1/2}(\Gamma) \to H^1(\Omega_2) \end{aligned}$$

such that for every $z \in H^{1/2}(\Gamma), \quad z \geq 0 \Rightarrow r_i(z) \geq 0,\ i = 1,2$. Such operators can be easily obtained by taking any lifting operator and then taking its positive part. Precisely, we use that for any $u \in H^1(\Omega_1)$, $u^+ = \max\{u, 0\} \in H^1(\Omega_1)$, $u^- = \max\{0, -u\} \in H^1(\Omega_1)$ and $u = u^+ - u^-$. Similarly, for any $v \in H^1(\Omega_2)$, $v^+ \in H^1(\Omega_2)$, $v^- \in H^1(\Omega_2)$ and $v = v^+ - v^-$. For any $u \in H^1(\Omega_1)$ and $v \in H^1(\Omega_2)$, we denote respectively by $u_{|\Gamma}$ and $v_{|\Gamma}$ their Sobolev traces on Γ. Let us now perform the following decomposition: for any $(u, v) \in H^1(\Omega_1) \times H^1(\Omega_2)$

$$\begin{aligned} (u, v) &= (u^+ - u^-, v) \\ &= \Big(u^+, v + r_2\big((u^-)_{|\Gamma}\big)\Big) - \Big(u^-, r_2\big((u^-)_{|\Gamma}\big)\Big). \end{aligned} \tag{37}$$

Let us notice that $\Big(u^-, r_2\big((u^-)_{|\Gamma}\big)\Big)$ belongs to $\boldsymbol{\mathcal{V}}$ because u^- and $r_2\big((u^-)_{|\Gamma}\big)$ have the same trace on Γ. Let us perform once more this operation: set $\mathrm{v} = v + r_2\big((u^-)_{|\Gamma}\big)$ which belongs to $H^1(\Omega_2)$.

$$\begin{aligned} (u^+, \mathrm{v}) &= (u^+, \mathrm{v}^+ - \mathrm{v}^-) \\ &= \Big(u^+ + r_1\big((\mathrm{v}^-)_{|\Gamma}\big), \mathrm{v}^+\Big) - \Big(r_1\big((\mathrm{v}^-)_{|\Gamma}\big), \mathrm{v}^-\Big). \end{aligned} \tag{38}$$

Combining (37) and (38) we finally obtain

$$(u, v) = \Big(u^+ + r_1\big((\mathrm{v}^-)_{|\Gamma}\big), \mathrm{v}^+\Big) - \Big[\Big(r_1\big((\mathrm{v}^-)_{|\Gamma}\big), \mathrm{v}^-\Big) + \Big(u^-, r_2\big((u^-)_{|\Gamma}\big)\Big)\Big].$$

By construction of the trace lifting operator, and by $\mathrm{v}^- \geq 0$ we have $r_1\big((\mathrm{v}^-)_{|\Gamma}\big) \geq 0$. Thus, we have obtained a decomposition of (u, v) as a difference of an element of $H^1(\Omega_1)^+ \times H^1(\Omega_2)^+$ and an element of $H^1(\Omega)$. The decomposition algorithm can now be developped in a very similar way as in the unconstrained case.

REFERENCES

[1] F. Acker, M.A. Prestel, Convergence d'un schéma de minimisation alternée, *Annales de la Faculté des Sciences de Toulouse V, série Mathématiques*, 2, (1980), pp. 1-9.

[2] R. Adams, Sobolev spaces, Academic Press, New-York, 1975.

[3] S. Agmon, A. Douglis, L. Nirenberg, Estimates near the boundary for solutions of elliptic partial differential equations satisfying general boundary conditions I., *Comm. Pure Appl. Math.*, 12 (1959), pp. 623-727.

[4] S. Agmon, A. Douglis, L. Nirenberg, Estimates near the boundary for solutions of elliptic partial differential equations satisfying general boundary conditions II., *Comm. Pure Appl. Math.*, 17 (1964), pp. 35-92.

[5] F. Alvarez, R. Cominetti, Primal and dual convergence of a proximal point exponential penalty method for linear programming, *Math. Programming*, 93 (2002), pp. 87-96.

[6] H. Attouch, J. Bolte, P. Redont, A. Soubeyran, Alternating proximal algorithms for weakly coupled convex minimization problems. Applications to dynamical games and PDE's, *Journal of Convex Analysis*, 15 (2008), No 3, pp. 485-506.

[7] H. Attouch, H. Brézis, Duality for the sum of convex functions in general Banach spaces, *Aspects of Mathematics and Its Applications*, J. Barroso Ed., North-Holland, Amsterdam (1986), pp. 125-133.

[8] H. Attouch, G. Buttazzo, G. Michaille, Variational analysis in Sobolev and BV spaces. Applications to PDE's and optimization, *MPS/SIAM Series on Optimization*, 6. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2006.

[9] H. Attouch, M.-O. Czarnecki, Asymptotic control and stabilization of nonlinear oscillators with non isolated equilibria, *Journal of Differential Equations*, 179 (2002), pp. 278-310.

[10] H. Attouch, M.-O. Czarnecki, Asymptotic behavior of coupled dynamical systems with multiscale aspects, *Journal of Differential Equations*, 248 (2010), pp. 1315-1344.

[11] H. Attouch, M.-O. Czarnecki, J. Peypouquet, Prox-penalization and splitting methods for constrained variational problems, *paper under review*.

[12] H. Attouch, P. Redont, A. Soubeyran, A new class of alternating proximal minimization algorithms with costs-to-move, *SIAM Journal on Optimization*, 18 (2007), pp. 1061-1081.

[13] H. Attouch, M. Soueycatt, Augmented Lagrangian and proximal alternating direction methods of multipliers in Hilbert spaces. Applications to games , PDE's and control, *Pacific Journal of Optimization*, 5 (2009), No 1, pp. 17-37.

[14] A. Auslender, J.-P. Crouzeix, P. Fedit, Penalty-proximal methods in convex programming, *J. Optim. Theory Appl.*, 55 (1987), pp. 1-21.

[15] D. Azé, Eléments d'analyse convexe et variationnelle, Ellipses, Paris, 1997.

[16] J.-B. Baillon, R. Cominetti, A convergence result for non-autonomous subgradient evolution equations and its application to the steepest descent exponential penalty trajectory in linear programming, *J. Funct. Anal.*, 187 (2001), pp. 263-273.

[17] H.H. Bauschke, P.L. Combettes, S. Reich, The asymptotic behavior of the composition of two resolvents, *Nonlinear analysis*, 60 (2005), pp. 283-301.

[18] H. Brézis, Opérateurs maximaux monotones dans les espaces de Hilbert et équations d'évolution, Lecture Notes 5, North Holland, 1972.

[19] H. Brézis, Analyse fonctionnelle: théorie et applications, Dunod, Paris, 1999.

[20] H. Brézis, P. L. Lions, Produits infinis de résolvantes, *Israel J. Math.*, 29 (1978), pp. 329-345.

[21] A. Cabot, Proximal point algorithm controlled by a slowly vanishing term: Applications to hierarchical minimization, *SIAM Journal on Optimization*, 15 (2005), No 2, pp. 555-572.

[22] A. Cabot, P. Frankel, Alternating proximal algorithms with costs-to-move and asymptotically vanishing coupling, submitted to *Optimization*.

[23] R. Cominetti, M. Courdourier, Coupling general penalty schemes for convex programming with the steepest descent method and the proximal point algorithm, *SIAM J. Optim*, **13** (2002), 745-765.

[24] I. Ekeland, R. Temam, Convex Analysis and Variational Problems, SIAM Classics in Applied Mathematics, 28, 1999.

[25] D. Gilbarg, N. Trudinger, Elliptic partial differential equations of second order, Springer-Verlag, Berlin, 1977.

[26] J.-L. Lions, E. Magenes, Non-homogeneous boundary value problems and applications, Die Grundlehren der mathematischen Wissenschaften, Springer-Verlag, New York-Heidelberg, 1972.

[27] J. Marschall, The trace of Sobolev-Slobodeckij spaces on Lipschitz domains, *Manuscripta Math.*, **58** (1987), pp. 47–65.

[28] J.J. Moreau, Fonctionnelles convexes, Cours Collège de France 1967, new edition CNR, Facoltà di Ingegneria di Roma, 2003.

[29] Z. Opial, Weak convergence of the sequence of successive approximations for nonexpansive mappings, *Bull. Amer. Math. Soc.*, **73** (1967), pp. 591-597.

[30] G. Passty , Ergodic convergence to a zero of the sum of monotone operators in Hilbert space, *J. Math. Anal. Appl.*, **72** (1979), no. 2, pp. 383-390.

[31] B.T. Polyak, Introduction to optimization, Opimization Software, Inc. Publications Division, New york, 1987.

[32] P.-A. Raviart, J.-M. Thomas, Introduction à l'analyse numérique des équations aux dérivées partielles, Masson, Paris, 1993.

[33] R.T. Rockafellar, Convex Analysis, Princeton University Press, Princeton, 1970.

[34] R.T. Rockafellar, R. Wets, Variational analysis, Springer, Berlin, 1998.

DÉPARTEMENT DE MATHÉMATIQUES, UNIVERSITÉ MONTPELLIER II, CC 051, PLACE EUGÈNE BATAILLON, 34095 MONTPELLIER CEDEX 5, FRANCE

E-mail address: attouch@math.univ-montp2.fr, acabot@math.univ-montp2.fr, p.frankel30@orange.fr

DEPARTAMENTO DE MATEMÁTICA, UNIVERSIDAD TÉCNICA FEDERICO SANTA MARÍA, AVENIDA ESPAÑA 1680, VALPARAÍSO, CHILE

E-mail address: juan.peypouquet@usm.cl