Journal of Convex Analysis
Volume 15 (2008), No. 3, 485–506

# Alternating Proximal Algorithms for Weakly Coupled Convex Minimization Problems. Applications to Dynamical Games and PDE's*

**H. Attouch**
*Institut de Mathématiques et de Modélisation de Montpellier, UMR CNRS 5149, CC 51, Université Montpellier II, Place Eugène Bataillon, 34095 Montpellier, France*
*attouch@math.univ-montp2.fr*

**J. Bolte**
*Equipe Combinatoire et Optimisation, Université Paris 6*
*bolte@math.jussieu.fr*

**P. Redont**
*Institut de Mathématiques et de Modélisation de Montpellier, UMR CNRS 5149, CC 51, Université Montpellier II, Place Eugène Bataillon, 34095 Montpellier, France*
*redont@math.univ-montp2.fr*

**A. Soubeyran**
*GREQAM UMR CNRS 6579, Université de la Méditerranée, 13290 Les Milles, France*
*antoine.soubeyran@univmed.fr*

*Dedicated to the memory of Thomas Lachand-Robert.*

Received: November 12, 2007

We introduce and study alternating minimization algorithms of the following type

$$\begin{gathered}(x_0, y_0) \in \mathcal{X} \times \mathcal{Y},\ \alpha, \mu, \nu > 0 \text{ given},\\ (x_k, y_k) \to (x_{k+1}, y_k) \to (x_{k+1}, y_{k+1}) \text{ as follows}\\ \begin{cases} x_{k+1} = \operatorname{argmin}\{f(\xi) + \frac{\mu}{2}Q(\xi, y_k) + \frac{\alpha}{2}\|\xi - x_k\|^2 :\ \xi \in \mathcal{X}\}\\ y_{k+1} = \operatorname{argmin}\{g(\eta) + \frac{\mu}{2}Q(x_{k+1}, \eta) + \frac{\nu}{2}\|\eta - y_k\|^2 :\ \eta \in \mathcal{Y}\}\end{cases}\end{gathered}$$

where $\mathcal{X}$ and $\mathcal{Y}$ are real Hilbert spaces, $f : \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$, $g : \mathcal{Y} \to \mathbb{R} \cup \{+\infty\}$ are closed convex proper functions, $Q : (x, y) \in \mathcal{X} \times \mathcal{Y} \to \mathbb{R}^+$ is a nonnegative quadratic form (hence convex, but possibly nondefinite) which couples the variables $x$ and $y$. A particular important situation is the "weak coupling" $Q(x, y) = \|Ax - By\|^2$ where $A \in L(\mathcal{X}, \mathcal{Z})$, $B \in L(\mathcal{Y}, \mathcal{Z})$ are continuous linear operators acting respectively from $\mathcal{X}$ and $\mathcal{Y}$ into a third Hilbert space $\mathcal{Z}$.

The "cost-to-move" terms $\|\xi - x\|^2$ and $\|\eta - y\|^2$ induce dissipative effects which are similar to friction in mechanics, anchoring and inertia in decision sciences. As a result, for each initial data $(x_0, y_0)$, the proximal-like algorithm generates a sequence $(x_k, y_k)$ which weakly converges to a minimum point of the convex function $L(x, y) = f(x) + g(y) + \frac{\mu}{2}Q(x, y)$. The cost-to-move terms, which vanish asymptotically, have a crucial role in the convergence of the algorithm. A direct alternating minimization of the function $L$ could fail to produce a convergent sequence in the weak coupling case.

Applications are given in game theory, variational problems and PDE's. These results are then extended to an arbitrary number of decision variables and to monotone inclusions.

*The first three authors acknowledge the support of the French ANR under grant ANR-05-BLAN-0248-01.

ISSN 0944-6532 / $ 2.50 

*Keywords:* Convex optimization, alternating minimization, splitting methods, proximal algorithm, weak coupling, quadratic coupling, costs to change, anchoring effect, dynamical games, best response, domain decomposition for PDE's, monotone inclusions

*1991 Mathematics Subject Classification:* 65K05, 65K10, 46N10, 49J40, 49M27, 90B50, 90C25

## 1. Introduction

### 1.1. Problem statement

The aim of this paper is to apply the alternating minimization algorithmic framework (with costs-to-move), recently introduced by Attouch, Redont and Soubeyran [4], to structured convex optimization problems of the following type

$$\min\left\{f(x)+g(y)+\frac{\mu}{2}Q(x,y):\ x\in\mathcal{X},\ y\in\mathcal{Y}\right\}, \tag{1}$$

where

- $\mathcal{X},\mathcal{Y}$ are real Hilbert spaces (possibly infinite dimensional);
- $f:\mathcal{X}\to\mathbb{R}\cup\{+\infty\}$, $g:\mathcal{Y}\to\mathbb{R}\cup\{+\infty\}$ are closed convex proper functions acting respectively on the spaces $\mathcal{X}$ and $\mathcal{Y}$;
- $Q:\mathcal{X}\times\mathcal{Y}\to\mathbb{R}^+$ is a continuous nonnegative quadratic form (hence convex, but possibly nondefinite) which couples the two variables $x\in\mathcal{X}$ and $y\in\mathcal{Y}$;
- $\mu$ is a positive parameter.

This is illustrated by the following model example: let $A\in L(\mathcal{X},\mathcal{Z})$ and $B\in L(\mathcal{Y},\mathcal{Z})$ be linear continuous operators acting respectively from $\mathcal{X}$ to $\mathcal{Z}$ and from $\mathcal{Y}$ to $\mathcal{Z}$. Take $Q(x,y)=\|Ax-By\|^2_{\mathcal{Z}}$ then (1) becomes

$$\min\left\{f(x)+g(y)+\frac{\mu}{2}\|Ax-By\|^2_{\mathcal{Z}}\ :\ x\in\mathcal{X},\ y\in\mathcal{Y}\right\}. \tag{2}$$

This type of variational problem occurs in various domains such as decision sciences and game theory [4], partial differential equations and mechanics [2], optimal control and approximation theory [13], image processing and signal theory [7, 11]. The original feature of the above problem (2) comes from the coupling term

$$Q(x,y)=\|Ax-By\|^2_{\mathcal{Z}}$$

that we shall call weak coupling, and which allows asymmetric, indirect or partial relations between the variables. This terminology is intended to make contrast with the classical variational setting, namely

$$\min\left\{f(x)+g(y)+\frac{\mu}{2}\|x-y\|^2_{\mathcal{H}}:\ x\in\mathcal{H},\ y\in\mathcal{H}\right\} \tag{3}$$

where the variables $x$ and $y$ lie in the same space $\mathcal{H}$ and the coupling, called "strong coupling", $Q(x,y)=\|x-y\|^2_{\mathcal{H}}$, involves the "whole" variables $x$ and $y$ in a symmetric way.

The variational formulation (2) covers many situations which do not enter into the formulation (3):

- In decision sciences and team games, one can consider players who only interact via some components, or some functions, of their decision variables.
- In the domain of elliptic partial differential equations, decomposition methods lead to transmission conditions through the common interface $\Gamma$ separating a domain $\Omega$ into two adjacent domains $\Omega_1$ and $\Omega_2$. A natural functional framework is

$$\begin{cases} \mathcal{X} = H^1(\Omega_1),\ \mathcal{Y} = H^1(\Omega_2),\ \mathcal{Z} = L^2(\Gamma) \\ A : \mathcal{X} = H^1(\Omega_1) \to \mathcal{Z} = L^2(\Gamma) \text{ is the trace operator} \\ B : \mathcal{Y} = H^1(\Omega_2) \to \mathcal{Z} = L^2(\Gamma) \text{ is the trace operator} \end{cases}$$

  Given some $h \in L^2(\Omega)$, problem (3) takes the following form

$$\min \left\{ \frac{1}{2}\int_{\Omega_1} |\nabla v_1|^2 + \frac{1}{2}\int_{\Omega_2} |\nabla v_2|^2 + \frac{\mu}{2}\int_\Gamma [v]^2 - \int_\Omega hv : \right.$$
$$\left. v_1 \in H^1(\Omega_1),\ v_2 \in H^1(\Omega_2),\ v = 0 \text{ on } \partial\Omega \right\}$$

  where $v = v_1$ on $\Omega_1$, $v = v_2$ on $\Omega_2$ and $[v] =$ jump of $v$ through the interface $\Gamma$.
- In optimal control theory, $x$ and $y$ respectively stand for the state and control variables, the coupling is a penalized function of the state equation $Ax = By$. See J.-L. Lions [13] for a systematic utilization of this device in the control of singular distributed systems.

Let us now come to the central subject of this paper which is the study of the convergence of alternating minimization algorithms with costs-to-move for such weakly coupled variational systems. The mathematical analysis will be carried out by considering the abstract general framework (1). The subject is of importance:

- First for numerical reasons: it is natural to exploit the "separable" structure of the minimization problem (1) and thus to study the convergence of the alternating minimization algorithm.
- Secondly for the purpose of modelling dynamical decision processes: for example, alternating minimization algorithms model "best response dynamics" in the case of Nash potential games (see [4]).

These questions have been intensively studied in the case of the "strong" coupling. Let us recall the classical result (1980) due to Acker and Prestel [1]: let $\mathcal{H}$ be a real Hilbert space and $f : \mathcal{H} \to \mathbb{R} \cup \{+\infty\}$, $g : \mathcal{H} \to \mathbb{R} \cup \{+\infty\}$ two closed convex proper functions. Consider the sequence $(x_k, y_k)_{k\in\mathbb{N}}$ generated by the alternating minimization algorithm

$$\begin{cases} x_{k+1} = \operatorname{argmin}\{f(\xi) + \frac{1}{2}\|\xi - y_k\|^2_{\mathcal{H}} :\ \xi \in \mathcal{H}\} \\ y_{k+1} = \operatorname{argmin}\{g(\eta) + \frac{1}{2}\|x_{k+1} - \eta\|^2_{\mathcal{H}} :\ \eta \in \mathcal{H}\} \end{cases}$$

Then, the sequence $(x_k, y_k)_{k\in\mathbb{N}}$ weakly converges to a solution of the joint minimization problem on $\mathcal{H} \times \mathcal{H}$

$$\min \left\{ f(x) + g(y) + \frac{1}{2}\|x - y\|^2_{\mathcal{H}} :\ (x, y) \in \mathcal{H} \times \mathcal{H} \right\}$$

if we assume that the minimum point set is nonempty. Note that, and this justifies the terminology, the preceding algorithm can be viewed as the alternating minimization of the bivariate function

$$L : (x, y) \in \mathcal{H} \times \mathcal{H} \mapsto L(x, y) = f(x) + g(y) + \frac{1}{2}\|x - y\|^2_{\mathcal{H}} \ \in \mathbb{R} \cup \{+\infty\}$$

and provides a minimizing sequence for $L$.

Acker and Prestel's theorem provides a natural extension of von Neumann's alternating projection theorem [14] for two closed convex nonempty sets of the Hilbert space $\mathcal{H}$ (take $f = \delta_{C_1}$, $g = \delta_{C_2}$).

### 1.2. Description of the alternating algorithm for weak coupling.

We now return to the description of the alternating algorithm with costs-to-change in the case of a general quadratic coupling (1). As we already stressed, one has to modify the classical alternating algorithm, which fails to converge in this general situation. Following Attouch, Redont and Soubeyran [4] we introduce costs-to-move and consider the following proximal-like alternating algorithm:

$$\begin{cases} x_{k+1} = \operatorname{argmin}\{f(\xi) + \frac{\mu}{2}Q(\xi, y_k) + \frac{\alpha}{2}\|\xi - x_k\|^2_{\mathcal{X}} : \ \xi \in \mathcal{X}\} \\ y_{k+1} = \operatorname{argmin}\{g(\eta) + \frac{\mu}{2}Q(x_{k+1}, \eta) + \frac{\nu}{2}\|\eta - y_k\|^2_{\mathcal{Y}} : \ \eta \in \mathcal{Y}\} \end{cases} \tag{4}$$

where $\alpha$ and $\nu$ are given fixed positive parameters.

In Section 2 it is shown that the sequence $(x_k, y_k)_{k\in\mathbb{N}}$ weakly converges to a minimum point of the convex function $L(x, y) = f(x) + g(y) + \frac{\mu}{2}Q(x, y)$.

Note that the cost-to-move terms $\|\xi - x\|^2_{\mathcal{X}}$ and $\|\eta - y\|^2_{\mathcal{Y}}$ induce dissipative effects which are similar to friction in mechanics, anchoring and inertia in decision sciences. They asymptotically vanish, but have a crucial role in the convergence of the algorithm.

In Section 3 we obtain a natural extension of this result to the case of $n$ variables. In that case, we restrict our analysis to the case of a sum of pairwise quadratic coupling functions.

### 1.3. Link with decision sciences and game theory: Inertial Nash equilibration processes.

Dynamical games provide both an important field of applications and a natural interpretation for the above algorithms. Let us briefly introduce some of the main ideas. The general context is that of noncooperative dynamical game theory. Consider two interrelated players 1 and 2 departing from strict individualism to take each other's decision into account via a coupling function. Their (per unit of time) static payoffs are made of two components: an individual payoff $f(\xi)$ and $g(\eta)$ for players 1 and 2 coming from their respective current actions $\xi \in \mathcal{X}$, $\eta \in \mathcal{Y}$ (decisions, strategies, performances...) and a common nonnegative payoff $Q(\xi, \eta)$ coming from their joint actions. In our presentation, individual payoffs $f(\xi)$ and $g(\eta)$ are cost functions (unsatisfied needs to be minimized) and $Q(\xi, \eta)$ is a joint cost function. Coefficients $\beta > 0$ and $\mu > 0$ represent how much each player contributes to the joint cost.

The static loss functions of players 1 and 2 are

$$F : (\xi, \eta) \in \mathcal{X} \times \mathcal{Y} \to F(\xi, \eta) = f(\xi) + \beta Q(\xi, \eta)$$
$$G : (\xi, \eta) \in \mathcal{X} \times \mathcal{Y} \to G(\xi, \eta) = g(\eta) + \mu Q(\xi, \eta).$$

The above formulation has been given in terms of costs in order to fit better with the literature concerning algorithms (minimization problems). Decision and Game theories use utility or benefit (profit) functions (maximization problems). The link between these two realms of literature is easy to do when one considers utility or benefit functions which are bounded from above (finite resource assumption). Let $\varphi(\xi)$, $\psi(\eta)$ and $R(\xi, \eta)$ be the respective individual and joint payoffs of the players, and let $\overline{\varphi}$, $\overline{\psi}$, $\overline{R}$ be their respective upper bounds. One reduces to our situation by setting $f(\xi) = \overline{\varphi} - \varphi(\xi) \geq 0$, $g(\eta) = \overline{\psi} - \psi(\eta) \geq 0$, $Q(\xi, \eta) = \overline{R} - R(\xi, \eta) \geq 0$.

The static gain functions of players 1 and 2 are $J_1(\xi, \eta) = \varphi(\xi) + \beta R(\xi, \eta)$ and $J_2(\xi, \eta) = \psi(\eta) + \mu R(\xi, \eta)$. Coefficients $\beta > 0$ and $\mu > 0$ represent how much each player benefits from the joint payoff $R(\xi, \eta)$. The model deals with the important case of a team whose members benefit (suffer) both from positive and negative externalities (interaction game).

The coupling term $R(\xi, \eta) \geq 0$ defines the more or less conflictual characteristic of the game, i.e. the degree of conflict between players which defines the nature of their interdependence. To better understand what a degree of conflict means, consider the one dimensional case where $\xi \in \mathcal{X} = \mathbb{R}$ and $\eta \in \mathcal{Y} = \mathbb{R}$ are effort levels, where the coupling term is the quadratic cost $Q(\xi, \eta) = (a\xi - b\eta)^2$, $a \in \mathbb{R}, b \in \mathbb{R}$ and where the joint payoff is the revenue $R(\xi, \eta) = \overline{R} - (a\xi - b\eta)^2 \geq 0$.

From $\partial R/\partial \xi = 2a(b\eta - a\xi)$, we see that, for a given $\eta$, the joint payoff $R(\xi, \eta)$ increases with $\xi$ up to the maximum $\xi^*(\eta) = (b/a)\eta$, and then decreases. If the individual payoff $\varphi(\xi)$ of player 1 increases with $\xi$, player 1 will have the incentive to continue to increase $\xi$ a bit higher than $\xi^*(\eta)$. This will hurt player 2 because, for a given $\eta$, the joint payoff will decrease as soon as $\xi$ is higher than $\xi^*(\eta)$. Similarly for player 2. If $b$ is much lower than $a$, $b/a$ is small, and, for a given $\eta$ the argument $\xi^*(\eta) = (b/a)\eta$ which maximizes the joint payoff is low. This will push player 1 to decrease very much the joint payoff, preferring to choose a $\xi$ high enough to increase his individual payoff $\varphi(\xi)$. Notice that, with respect to the joint payoff, effort levels are strategic substitutes or complements depending on the sign of $ab$: for example $ab < 0 \Longrightarrow \partial \xi^*(\eta)/\partial \eta = b/a < 0$.

The separable aspect of the payoff of each player between an individual and a joint payoff is central in this paper. It is shared by a lot of famous noncooperative games.

Let us now introduce some important dynamical aspects and make the connection with the algorithm (1). We add to this classical (per unit of time) static normal form game, some costs-to-change (or to move) for each player. Player 1 must pay the cost $h(x, \xi)$ to move from action $x \in \mathcal{X}$ to a new action $\xi \in \mathcal{X}$ and player 2 must pay the cost $k(y, \eta)$ to move from action $y \in \mathcal{Y}$ to a new action $\eta \in \mathcal{Y}$.

Let us briefly explain the role and the importance of these terms. The general idea is that, in real life, changing, improving the gain, the quality of actions has a cost. "Costs-to-change" cover various physical, physiological, psychological and cognitive aspects. They reflect the bounded rationality and behavioural features of decision processes in real life (see Kahneman [12], Camerer & Loewenstein [10] , Simon [16] for the concept of deliberation costs and Attouch & Soubeyran [5, 6] for the precise concept of costs to change).

Here, these costs mainly describe an anchoring effect. Agents have a (local) vision of their environment which depends on their current actions. Each action is anchored to the preceding one, which means that the perception the agents have of the quality of further actions depends on the current ones. In economics, management, one may think of actions as routines, ways of doing, with costs to change reflecting the difficulty to quit a routine or entering another one, or to change quickly (reactivity costs). In our situation, let us suppose that the current action of the two players is $(x, y) \in \mathcal{X} \times \mathcal{Y}$. Suppose now that only player 1 can choose a new action (i.e. player 1 chooses a new action $\xi$ while player 2 "stays" at $y$), then the inertial payoff of player 1 is

$$F(\xi, y) + \alpha h(x, \xi) = f(\xi) + \beta Q(y, \xi) + \alpha h(x, \xi).$$

The second member of this expression is the sum of three costs: a cost to be far from the objective (frustration), a cost to be far from (or close to!) the action of the other agent (coupling), and a cost to be far from the preceding action (anchoring or inertial effect). The coefficient $\alpha$ before the cost $h(x, \xi)$ usually reflects some dynamical cognitive features of player 1 (speed, reactivity, learning ability...).

Symmetrically, suppose now that only player 2 can choose a new action (i.e., player 2 chooses a new action $\eta$ while player 1 stays at $x$), then the inertial payoff of player 2 is

$$G(x, \eta) + \nu k(y, \eta) = g(\eta) + \mu Q(x, \eta) + \nu k(y, \eta)$$

with the $\nu$ coefficient reflecting some dynamical cognitive features of player 2. The timing of the game follows an asynchronous dynamic where players move in alternation. It is convenient, both for modelling and mathematical purpose, to describe it as a discrete dynamical system on the product space $\mathcal{X} \times \mathcal{Y}$.

The dynamic of actions works as follows:

$$(x_0, y_0) \longrightarrow (x_1, y_0) \longrightarrow (x_1, y_1) \longrightarrow \ldots (x_k, y_k) \longrightarrow (x_{k+1}, y_k) \longrightarrow (x_{k+1}, y_{k+1}) \longrightarrow \ldots$$

At stage $k = 0$, players 1 and 2 do the actions $x_0$ and $y_0$. At stage $k = 1$ player 1 can choose to do a new action $\xi = x_1$, while player 2 stays in $y_0$ (he can only repeat his previous action $y_0$). At stage $k = 2$ player 1 stays in $x_1$. He can only repeat his previous action $\xi = x_1$, while player 2 can choose to do a new action $\eta = y_1$ and so on.

When specializing the costs-to-change to be quadratic functions, namely $h(x, \xi) = \|x - \xi\|_{\mathcal{X}}^2$ and $k(y, \eta) = \|y - \eta\|_{\mathcal{Y}}^2$ we precisely obtain the dynamic described by algorithm (4).

## 2. The case of two variables.

### 2.1. The variational case.

We will make the following assumptions

$$(\mathcal{H}) \begin{cases} \text{-} \ \mathcal{X}, \mathcal{Y} \text{ two real Hilbert spaces;} \\ \text{-} \ f : x \in \mathcal{X} \mapsto f(x) \in \mathbb{R} \cup \{+\infty\} \text{ and } g : y \in \mathcal{Y} \mapsto g(y) \in \mathbb{R} \cup \{+\infty\} \text{ two} \\ \quad \text{convex, lower semicontinuous, proper functionals;} \\ \text{-} \ Q : (x, y) \in \mathcal{X} \times \mathcal{Y} \mapsto Q(x, y) \in \mathbb{R} \text{ a nonnegative continuous quadratic form;} \\ \text{-} \ \text{the functional} \\ \qquad L : (x, y) \in \mathcal{X} \times \mathcal{Y} \to L(x, y) = f(x) + g(y) + \frac{1}{2} Q(x, y) \in \mathbb{R} \cup \{+\infty\} \\ \quad \text{has at least one minimum point.} \end{cases}$$

The spaces $\mathcal{X}$ and $\mathcal{Y}$ may be infinite dimensional, and for simplicity $\|u\|$ equally denotes the norm of an element $u \in \mathcal{X}$ or of an element $u \in \mathcal{Y}$; likewise, $\langle u, v\rangle$ denotes the inner product of $u$ and $v$ equally in $\mathcal{X} \times \mathcal{X}$ or in $\mathcal{Y} \times \mathcal{Y}$.

The quadratic form $Q$ is convex, since it is nonnegative, but possibly nondefinite (e.g. $\mathcal{X} = \mathcal{Y}$ and $Q(x, y) = \|x - y\|^2$).

Let $q : (\mathcal{X}\times\mathcal{Y})^2 \to \mathbb{R}$ denote the bilinear form associated with $Q$, i.e., for all $((x, y), (\xi, \eta)) \in (\mathcal{X} \times \mathcal{Y})^2$

$$q((x, y), (\xi, \eta)) = \frac{1}{2}\{Q(x + \xi, y + \eta) - Q(x, y) - Q(\xi, \eta)\}. \tag{5}$$

If we set $\eta = 0$ in equation (5), it provides us with an expression for the gradient $\nabla_x Q(x, y)$ at point $(x, y)$ with respect to the $x$ variable

$$\langle \nabla_x Q(x, y), \xi\rangle = 2q((x, y), (\xi, 0)).$$

Hence the map $(x, y) \in \mathcal{X} \times \mathcal{Y} \to \nabla_x Q(x, y) \in \mathcal{X} \times \mathcal{Y}$ is linear continuous; and further, in view of (5), we also have

$$\langle \nabla_x Q(x, y), \xi\rangle = Q(x, y) + Q(\xi, 0) - Q(x - \xi, y). \tag{6}$$

We are interested in finding the minimum of the functional $L(x, y) = f(x)+g(y)+\frac{1}{2}Q(x, y)$ where the function $\frac{1}{2}Q$ acts as a coupling between spaces $\mathcal{X}$ and $\mathcal{Y}$. According to our assumptions, $L$ has at least one minimum point, say $(\overline{x}, \overline{y})$. Due to the special form of $L$, its minimum points are also characterized as Nash equilibria of the pair of functions $(\xi, \eta) \to f(\xi) + \frac{1}{2}Q(\xi, \eta)$ and $(\xi, \eta) \to g(\eta) + \frac{1}{2}Q(\xi, \eta)$

$$\begin{cases} \overline{x} \in \operatorname{argmin}\{f(\xi) + \frac{1}{2}Q(\xi, \overline{y}) : \ \xi \in \mathcal{X}\} \\ \overline{y} \in \operatorname{argmin}\{g(\eta) + \frac{1}{2}Q(\overline{x}, \eta) : \ \eta \in \mathcal{Y}\}. \end{cases}$$

Writing the optimality condition for $(\overline{x}, \overline{y})$ to be a minimum point of $L$ we obtain

$$\begin{cases} \partial f(\overline{x}) + \frac{1}{2}\nabla_x Q(\overline{x}, \overline{y}) \ni 0 & (7) \\ \partial g(\overline{y}) + \frac{1}{2}\nabla_y Q(\overline{x}, \overline{y}) \ni 0. & (8) \end{cases}$$

To find minimum points of $L$ we propose the following alternate algorithm

$$(\mathcal{A}) \quad \begin{cases} x_{k+1} = \operatorname{argmin}\{f(\xi) + \frac{1}{2}Q(\xi, y_k) + \frac{1}{2}\|\xi - x_k\|^2 : \ \xi \in \mathcal{X}\} \\ y_{k+1} = \operatorname{argmin}\{g(\eta) + \frac{1}{2}Q(x_{k+1}, \eta) + \frac{1}{2}\|\eta - y_k\|^2 : \ \eta \in \mathcal{Y}\}. \end{cases}$$

The terms $\frac{1}{2}\|\xi - x_k\|^2$ and $\frac{1}{2}\|\eta - y_k\|^2$ are anchoring terms forcing $x_{k+1}$ and $y_{k+1}$ not to move too far away from $x_k$ and $y_k$. Observe also that these anchoring terms introduce strict convexity, and so $x_{k+1}$ and $y_{k+1}$ are uniquely defined; hence the equality sign above instead of the *member of* sign.

Points $x_{k+1}$ and $y_{k+1}$ are also characterized by the optimality conditions

$$\begin{cases} \partial f(x_{k+1}) + \frac{1}{2}\nabla_x Q(x_{k+1}, y_k) + (x_{k+1} - x_k) \ni 0 & (9) \\ \partial g(y_{k+1}) + \frac{1}{2}\nabla_y Q(x_{k+1}, y_{k+1}) + (y_{k+1} - y_k) \ni 0. & (10) \end{cases}$$

**Theorem 2.1.** *Under assumptions $(\mathcal{H})$, the sequence $(x_k, y_k)$ generated by the alternate algorithm $(\mathcal{A})$ is a minimizing sequence for $L$ converging weakly in $\mathcal{X} \times \mathcal{Y}$ to a minimum point $(x_\infty, y_\infty)$ of $L$. Moreover, $f(x_k) \to f(x_\infty)$, $g(y_k) \to g(y_\infty)$, $Q(x_k, y_k) \to Q(x_\infty, y_\infty)$, $\|x_{k+1} - x_k\| \to 0$ and $\|y_{k+1} - y_k\| \to 0$ as $k \to +\infty$.*

**Proof.** By the monotonicity of the subdifferential operator $\partial f$ at points $\overline{x}$ and $x_{k+1}$ we deduce from (7) and (9)

$$\langle (x_{k+1} - x_k) + \frac{1}{2}\nabla_x Q(x_{k+1}, y_k) - \frac{1}{2}\nabla_x Q(\overline{x}, \overline{y}), x_{k+1} - \overline{x}\rangle \leq 0. \tag{11}$$

On the one hand we have

$$\langle x_{k+1} - x_k, x_{k+1} - \overline{x}\rangle = \frac{1}{2}\|x_{k+1} - \overline{x}\|^2 - \frac{1}{2}\|x_k - \overline{x}\|^2 + \frac{1}{2}\|x_{k+1} - x_k\|^2. \tag{12}$$

On the other hand, in view of the linearity of the gradient map and of (6), we have

$$\begin{aligned} & \langle \nabla_x Q(x_{k+1}, y_k) - \nabla_x Q(\overline{x}, \overline{y}), x_{k+1} - \overline{x}\rangle \\ = \;& \langle \nabla_x Q(x_{k+1} - \overline{x}, y_k - \overline{y}), x_{k+1} - \overline{x}\rangle \\ = \;& Q(x_{k+1} - \overline{x}, y_k - \overline{y}) + Q(x_{k+1} - \overline{x}, 0) - Q(0, y_k - \overline{y}). \end{aligned} \tag{13}$$

Collecting (11, 12, 13) we obtain

$$\begin{aligned} & \|x_{k+1} - \overline{x}\|^2 - \|x_k - \overline{x}\|^2 + \|x_{k+1} - x_k\|^2 + Q(x_{k+1} - \overline{x}, y_k - \overline{y}) \\ & \quad + Q(x_{k+1} - \overline{x}, 0) - Q(0, y_k - \overline{y}) \leq 0. \end{aligned} \tag{14}$$

Similarly, expressing the monotonicity of the subgradient operator $\partial g$ at points $\overline{y}$ and $y_{k+1}$, we can deduce from (8) and (10)

$$\begin{aligned} & \|y_{k+1} - \overline{y}\|^2 - \|y_k - \overline{y}\|^2 + \|y_{k+1} - y_k\|^2 + Q(x_{k+1} - \overline{x}, y_{k+1} - \overline{y}) \\ & \quad + Q(0, y_{k+1} - \overline{y}) - Q(x_{k+1} - \overline{x}, 0) \leq 0. \end{aligned} \tag{15}$$

Adding inequalities (14) and (15) we obtain

$$\begin{aligned} & \|x_{k+1} - \overline{x}\|^2 + \|y_{k+1} - \overline{y}\|^2 + Q(0, y_{k+1} - \overline{y}) - \|x_k - \overline{x}\|^2 - \|y_k - \overline{y}\|^2 \\ & \quad - Q(0, y_k - \overline{y}) + \|x_{k+1} - x_k\|^2 + \|y_{k+1} - y_k\|^2 + Q(x_{k+1} - \overline{x}, y_k - \overline{y}) \\ & \quad + Q(x_{k+1} - \overline{x}, y_{k+1} - \overline{y}) \leq 0. \end{aligned}$$

And summing for $k$ from 0 to some $K$ we get further

$$\begin{aligned} & \sum_{k=0}^{K} \{\|x_{k+1} - x_k\|^2 + \|y_{k+1} - y_k\|^2 + Q(x_{k+1} - \overline{x}, y_k - \overline{y}) \\ & \quad + Q(x_{k+1} - \overline{x}, y_{k+1} - \overline{y})\} + \|x_{K+1} - \overline{x}\|^2 + \|y_{K+1} - \overline{y}\|^2 + Q(0, y_{K+1} - \overline{y}) \\ \leq \;& \|x_0 - \overline{x}\|^2 + \|y_0 - \overline{y}\|^2 + Q(0, y_0 - \overline{y}). \end{aligned}$$

Whence we draw the following consequences:

**a.** the sequence $(x_k, y_k)$ is bounded in $\mathcal{X} \times \mathcal{Y}$;

**b.** the sequence $k \to \|x_k - \overline{x}\|^2 + \|y_k - \overline{y}\|^2 + Q(0, y_k - \overline{y})$ is nonincreasing;
**c.** the quantities $\|x_{k+1} - x_k\|$, $\|y_{k+1} - y_k\|$ vanish as $k$ goes to $+\infty$.

Now, rewriting equalities (9,10) in the following way

$$\begin{cases} -(x_{k+1} - x_k) + \frac{1}{2}\nabla_x Q(0, y_{k+1} - y_k) \in \partial f(x_{k+1}) + \frac{1}{2}\nabla_x Q(x_{k+1}, y_{k+1}) \\ \qquad\qquad\qquad\qquad -(y_{k+1} - y_k) \in \partial g(y_{k+1}) + \frac{1}{2}\nabla_y Q(x_{k+1}, y_{k+1}) \end{cases}$$

shows that $(u_{k+1}, v_{k+1}) = \big(-(x_{k+1} - x_k) + \frac{1}{2}\nabla_x Q(0, y_{k+1} - y_k), -(y_{k+1} - y_k)\big)$ is a subgradient of $L$ at point $(x_{k+1}, y_{k+1})$ that converges in norm to 0 in $\mathcal{X} \times \mathcal{Y}$ according to **c**. Writing the subgradient inequality at that point yields, for all $(\xi, \eta) \in \mathcal{X} \times \mathcal{Y}$

$$L(\xi, \eta) \geq L(x_{k+1}, y_{k+1}) + \langle u_{k+1}, \xi - x_{k+1}\rangle + \langle v_{k+1}, \eta - y_{k+1}\rangle.$$

Let $(x_{k'+1}, y_{k'+1})$ be a subsequence that converges weakly to some limit $(x_\infty, y_\infty)$. In view of points **a** and **c** above and of the continuity of the gradient, we have

$$L(\xi, \eta) \geq \liminf_{k' \to \infty} L(x_{k'+1}, y_{k'+1}) \geq L(x_\infty, y_\infty). \tag{16}$$

Hence $(x_\infty, y_\infty)$ is a minimum point of $L$.

Now, $N(u, v) = (\|u\|^2 + \|v\|^2 + Q(0, v))^{1/2}$ is a norm on $\mathcal{X} \times \mathcal{Y}$ derived from the inner product $((u_1, v_1), (u_2, v_2)) \in (\mathcal{X} \times \mathcal{Y})^2 \to \langle u_1, u_2\rangle + \langle v_1, v_2\rangle + q(0, v_1; 0, v_2)$; since $Q$ is continuous, norm $N$ is equivalent to the canonical norm. Moreover $N((x_k, y_k) - (\overline{x}, \overline{y}))$ does have a limit (point **b**). Opial's lemma [15] then shows that $(x_k, y_k)$ converges weakly to some limit, still denoted $(x_\infty, y_\infty)$, in argmin $L$.

As the subdifferential of $L$ at point $(x_k, y_k)$ contains a sequence $(u_k, v_k)$ that converges strongly to 0, inequality (16) shows that the bounded sequence $(x_k, y_k)$ does minimize $L$. Hence $f(x_k) + g(y_k) + \frac{1}{2}Q(x_k, y_k) \to f(x_\infty) + g(y_\infty) + \frac{1}{2}Q(x_\infty, y_\infty)$. Noticing that $f(x_\infty) \leq \liminf f(x_k)$, $g(y_\infty) \leq \liminf g(y_k)$ and $Q(x_\infty, y_\infty) \leq \liminf Q(x_k, y_k)$, we easily infer $f(x_k) \to f(x_\infty)$, $g(y_k) \to g(y_\infty)$ and $Q(x_k, y_k) \to Q(x_\infty, y_\infty)$.

## 2.2. A typical example.

A particular important situation is the weak coupling $Q(x, y) = \|Ax - By\|^2$ where $A \in L(\mathcal{X}, \mathcal{Z})$, $B \in L(\mathcal{Y}, \mathcal{Z})$ are linear continuous operators acting respectively from $\mathcal{X}$ and $\mathcal{Y}$ into a third Hilbert space $\mathcal{Z}$. Algorithm $(\mathcal{A})$ then reads

$$\begin{cases} \partial f(x_{k+1}) + A^*(Ax_{k+1} - By_k) + (x_{k+1} - x_k) \ni 0 \\ \partial g(y_{k+1}) + B^*(By_{k+1} - Ax_{k+1}) + (y_{k+1} - y_k) \ni 0 \end{cases}$$

where $A^* : \mathcal{Z} \to \mathcal{X}$ and $B^* : \mathcal{Z} \to \mathcal{Y}$ are the adjoint operators of $A$ and $B$. Then the conclusions of Theorem 2.1 hold. This situation is illustrated in Sections 4 and 5.

Still more particularly, if $f = \delta_C$ and $g = \delta_D$ are the indicator functions of nonvoid closed convex sets $C \subseteq \mathcal{X}$ and $D \subseteq \mathcal{Y}$, then the minimum points $(\overline{x}, \overline{y})$ of function $L(x, y) = \delta_C(x) + \delta_D(y) + \|Ax - By\|^2$ satisfy $\mathrm{dist}_{\mathcal{Z}}(\Gamma, \Delta) = \inf\{\|x - y\|,\ x \in \Gamma,\ y \in \Delta\} = \|A\overline{x} - B\overline{y}\|$, where $\Gamma = AC$ and $\Delta = BD$ are convex subsets of $\mathcal{Z}$ that *need not be closed.*

Thus, if there do exist points realizing the distance $\mathrm{dist}_{\mathcal{Z}}(\Gamma, \Delta)$, then algorithm $(\mathcal{A})$ yields limit points $(x_\infty, y_\infty)$ such that $(Ax_\infty, By_\infty)$ realize the distance. In particular, if $\Gamma \cap \Delta \neq \emptyset$, the algorithm finds a point $Ax_\infty = By_\infty$ in the intersection.

### 2.3. The case of maximal monotone operators.

Subdifferentials of closed convex proper functions are particular cases of maximal monotone operators (see [9]). Actually the preceding analysis applies with only minor changes to maximal monotone operators. Our basic assumptions are the following

$$(\mathcal{H}')\quad\begin{cases}\text{- } \mathcal{X},\ \mathcal{Y} \text{ two real Hilbert spaces;}\\ \text{- } A : x \in \mathcal{X} \to A(x) \subseteq \mathcal{X} \text{ and } B : y \in \mathcal{Y} \to B(y) \subseteq \mathcal{Y} \text{ two maximal monotone operators;}\\ \text{- } Q : (x,y) \in \mathcal{X}\times\mathcal{Y} \mapsto Q(x,y) \in \mathbb{R} \text{ a nonnegative continuous quadratic form;}\\ \text{- the maximal monotone operator}\\ \qquad T = (A,B) + \dfrac{1}{2}\nabla Q : \mathcal{X}\times\mathcal{Y} \rightrightarrows \mathcal{X}\times\mathcal{Y}\\ \quad\text{has at least one zero point.}\end{cases}$$

Let $(\overline{x},\overline{y}) \in \mathcal{X}\times\mathcal{Y}$ be a zero point of operator $T$, i.e.

$$\begin{cases} A(\overline{x}) + \frac{1}{2}\nabla_x Q(\overline{x},\overline{y}) \ni 0 \\ B(\overline{y}) + \frac{1}{2}\nabla_y Q(\overline{x},\overline{y}) \ni 0. \end{cases}$$

To find zero points of $T$ we propose the following alternate algorithm that computes the sequence $(x_k, y_k)$ iteratively

$$(\mathcal{A}')\quad \begin{cases} A(x_{k+1}) + \frac{1}{2}\nabla_x Q(x_{k+1}, y_k) + (x_{k+1} - x_k) \ni 0 & (17)\\ B(y_{k+1}) + \frac{1}{2}\nabla_y Q(x_{k+1}, y_{k+1}) + (y_{k+1} - y_k) \ni 0. & (18)\end{cases}$$

Equation (17) uniquely defines $x_{k+1}$ in function of $x_k$ and $y_k$, because the operator $\xi \in \mathcal{X} \to A(\xi) + \frac{1}{2}\nabla Q(\xi, y_k) \subseteq \mathcal{X}$ is maximal monotone, and hence $A + \frac{1}{2}\nabla Q(., y_k) + I$ is invertible (see [9]). Similarly, $y_{k+1}$ is well defined by (18).

Obviously, if $A$ and $B$ are subdifferentials of closed convex proper functions, then assumptions $(\mathcal{H})$ and $(\mathcal{H}')$ and algorithms $(\mathcal{A})$ and $(\mathcal{A}')$ coincide.

**Theorem 2.2.** *Under assumptions $(\mathcal{H}')$, the sequence $(x_k, y_k)$ generated by the alternate algorithm $(\mathcal{A}')$ converges weakly in $\mathcal{X}\times\mathcal{Y}$ to a zero point $(x_\infty, y_\infty)$ of the maximal monotone operator $T = (A,B) + \frac{1}{2}\nabla Q$. Moreover $\|x_{k+1} - x_k\| \to 0$ and $\|y_{k+1} - y_k\| \to 0$.*

**Proof.** The proof runs along the same line as in the variational case except in the end where we resort to the lower semicontinuity of $L$ to conclude that any limit point of $(x_k, y_k)$ is a minimum point of $L$.

Let $(x_{k'+1}, y_{k'+1})$ be a subsequence that converges weakly to $(x_\infty, y_\infty)$, and let us rewrite equations (17, 18) as follows

$$\begin{cases} -(x_{k'+1} - x_{k'}) + \frac{1}{2}\nabla_x Q(0, y_{k'+1} - y_{k'}) \in A(x_{k'+1}) + \frac{1}{2}\nabla_x Q(x_{k'+1}, y_{k'+1}) \\ -(y_{k'+1} - y_{k'}) \in B(y_{k'+1}) + \frac{1}{2}\nabla_y Q(x_{k'+1}, y_{k'+1}) \end{cases}$$

equivalently

$$\left(-(x_{k'+1} - x_{k'}) + \frac{1}{2}\nabla_x Q(0, y_{k'+1} - y_{k'}), -(y_{k'+1} - y_{k'})\right) \in T(x_{k'+1}, y_{k'+1}).$$

In view of $\|x_{k'+1} - x_{k'}\| \to 0$, $\|y_{k'+1} - y_{k'}\| \to 0$, of the boundedness of the sequence $(x_k, y_k)$ and of the continuity of the gradient $\nabla_x Q$, the left hand side above converges strongly to zero, while in the right hand side the argument $(x_{k'+1}, y_{k'+1})$ converges weakly to $(x_\infty, y_\infty)$. Owing to the weak-strong closedness of maximal monotone operators, the limit $(x_\infty, y_\infty)$ is a zero point of $T$.

The proof is completed by invoking Opial's lemma like in the variational case.

## 3. The case of n variables.

The analysis for two variables in Section 2 extends to more than two variables provided some restriction is made on the quadratic form $Q$, namely the coupling between the variables is pairwise.

### 3.1. The variational case.

We will make the following assumptions

$$(\mathcal{H}_n)\begin{cases} \text{- } (\mathcal{X}_i)_{i\in\{1,\dots,n\}}\ n \text{ real Hilbert spaces;} \\ \text{- for each } i \in \{1,\dots,n\},\ f_i : x \in \mathcal{X}_i \mapsto f_i(x) \in \mathbb{R}\cup\{+\infty\} \text{ is a convex, lower} \\ \quad \text{semicontinuous, proper functional;} \\ \text{- for each } i,j \text{ with } 1 \le i < j \le n,\ Q_{ij} : (x_i,x_j) \in \mathcal{X}_i \times \mathcal{X}_j \mapsto Q_{ij}(x_i,x_j) \in \mathbb{R} \\ \quad \text{is a nonnegative continuous quadratic form; and define } Q : (x_1,\dots,x_n) \in \\ \quad \prod_{i=1}^n \mathcal{X}_i \mapsto Q(x_1,\dots,x_n) = \sum_{1\le i<j\le n} Q_{ij}(x_i,x_j) \in \mathbb{R}; \\ \text{- the function} \\ \qquad L : (x_1,\dots,x_n) \in \prod_{i=1}^n \mathcal{X}_i \\ \qquad \mapsto\ L(x_1,\dots,x_n) = \sum_{i=1}^n f_i(x_i) + \frac{1}{2}Q(x_1,\dots,x_n) \in \mathbb{R}\cup\{+\infty\} \\ \quad \text{has at least one minimum point.} \end{cases}$$

For simplicity $\|u\|$ will denote the norm of an element $u \in \mathcal{X}_i$ for any $i$; similarly, $\langle u, v\rangle$ will denote the dot product of $u$ and $v$ in $\mathcal{X}_i$ for any $i$.

Due to the special form of $L$, a minimum point $(\overline{x}_1, \dots, \overline{x}_n)$ is also characterized by the $n$ relations

$$\overline{x}_i \in \operatorname{argmin}\left\{ f_i(\xi) + \frac{1}{2}Q(\overline{x}_1,\dots,\overline{x}_{i-1},\xi,\overline{x}_{i+1},\dots,\overline{x}_n) :\ \xi \in \mathcal{X}_i \right\},\ i \in \{1,..,n\}.$$

Writing the optimality condition for $(\overline{x}_1, \dots, \overline{x}_n)$ to be a minimum point of $L$ we obtain

$$\partial f_i(\overline{x}_i) + \frac{1}{2}\nabla_{x_i} Q(\overline{x}_1,\dots,\overline{x}_n) \ni 0,\ i \in \{1,..,n\}. \tag{19}$$

To find minimum points of $L$ we propose the following alternate algorithm for successively computing the components $(x_{i,k+1})_{i\in\{1,..,n\}}$ at the $(k+1)$-th step

$$(\mathcal{A}_n)\begin{cases} x_{1,k+1} = \operatorname{argmin}\{f_1(\xi)+\frac{1}{2}Q(\xi,x_{2,k},\dots,x_{n,k})+\frac{1}{2}\|\xi-x_{1,k}\|^2:\ \xi\in\mathcal{X}_1\}\\ \vdots\\ x_{i,k+1} = \operatorname{argmin}\{f_i(\xi)+\frac{1}{2}Q(x_{1,k+1},\dots,x_{i-1,k+1},\xi,x_{i+1,k},\dots,x_{n,k})\\ \qquad\qquad +\frac{1}{2}\|\xi-x_{i,k}\|^2:\ \xi\in\mathcal{X}_i\}\\ \vdots\\ x_{n,k+1} = \operatorname{argmin}\{f_n(\xi)+\frac{1}{2}Q(x_{1,k+1},\dots,x_{n-1,k+1},\xi)+\frac{1}{2}\|\xi-x_{n,k}\|^2:\ \xi\in\mathcal{X}_n\}\end{cases}$$

The terms $\|\xi-x_{i,k}\|^2$ are anchoring terms forcing $x_{i,k+1}$ not to move too far away from $x_{i,k}$. Observe also that these anchoring terms introduce strict convexity, and so $x_{i,k+1}$ is uniquely defined; hence the equality sign above instead of the *element of* sign.

The components $x_{i,k+1}$ are also characterized by the optimality conditions

$$\begin{aligned}\partial f_i(x_{i,k+1})+\frac{1}{2}\nabla_{x_i}Q(x_{1,k+1},\dots,x_{i-1,k+1},x_{i,k+1},x_{i+1,k},\dots,x_{n,k})\\ +(x_{i,k+1}-x_{i,k})\ni 0,\ i\in\{1,..,n\}\end{aligned}\tag{20}$$

**Theorem 3.1.** *Under assumptions* $(\mathcal{H}_n)$*, the sequence* $k\to(x_{i,k})_{i\in\{1,..,n\}}$ *generated by the alternate algorithm* $(\mathcal{A}_n)$ *is a minimizing sequence for* $L$ *converging weakly in* $\prod_{i=1}^n\mathcal{X}_i$ *to a minimum point* $(x_{i,\infty})_{i\in\{1,..,n\}}$ *of* $L$*. Moreover* $f_i(x_{i,k})\to f_i(x_{i,\infty})$ *and* $Q(x_{1,k},\dots,x_{n,k})\to Q(x_{1,\infty},\dots,x_{n,\infty})$ *as* $k\to+\infty$*.*

**Proof.** By the monotonicity of the subdifferential operator $\partial f_i$ at points $\overline{x}_i$ and $x_{i,k+1}$ we deduce from (19) and (20)

$$\begin{aligned}\langle(x_{i,k+1}-x_{i,k})+\frac{1}{2}\nabla_{x_i}Q(x_{1,k+1},\dots,x_{i-1,k+1},x_{i,k+1},x_{i+1,k},\dots,x_{n,k})\\ -\frac{1}{2}\nabla_{x_i}Q(\overline{x}_1,\dots,\overline{x}_n),x_{i,k+1}-\overline{x}_i\rangle\le 0.\end{aligned}\tag{21}$$

On the one hand we have

$$\langle x_{i,k+1}-x_{i,k},x_{i,k+1}-\overline{x}_i\rangle=\frac{1}{2}\|x_{i,k+1}-\overline{x}_i\|^2-\frac{1}{2}\|x_{i,k}-\overline{x}_i\|^2+\frac{1}{2}\|x_{i,k+1}-x_{i,k}\|^2.\tag{22}$$

On the other hand, set for $i\in\{1,\dots,n\}$

$$\begin{aligned}S_i=&\langle\nabla_{x_i}Q(x_{1,k+1},\dots,x_{i-1,k+1},x_{i,k+1},x_{i+1,k},\dots,x_{n,k})\\ &-\nabla_{x_i}Q(\overline{x}_1,\dots,\overline{x}_n),x_{i,k+1}-\overline{x}_i\rangle.\end{aligned}\tag{23}$$

In view of the linearity of the gradient $\nabla_{x_i}Q$ and of the expression of $Q$ as the sum of the quadratic forms $Q_{ij}$, we have

$$\begin{aligned}S_i\ &=\ \langle\nabla_{x_i}Q(x_{1,k+1}-\overline{x}_1,\dots,x_{i-1,k+1}-\overline{x}_{i-1},\\ &\qquad x_{i,k+1}-\overline{x}_i,x_{i+1,k}-\overline{x}_{i+1},\dots,x_{n,k}-\overline{x}_n),x_{i,k+1}-\overline{x}_i\rangle\\ &=\ \sum_{1\le j<i}\langle\nabla_{x_i}Q_{ji}(x_{j,k+1}-\overline{x}_j,x_{i,k+1}-\overline{x}_i),x_{i,k+1}-\overline{x}_i\rangle\\ &\qquad+\sum_{i<j\le n}\langle\nabla_{x_i}Q_{ij}(x_{i,k+1}-\overline{x}_i,x_{j,k}-\overline{x}_j),x_{i,k+1}-\overline{x}_i\rangle.\end{aligned}$$

Invoking (6) we have further

$$S_i = \sum_{1\le j<i} \{Q_{ji}(x_{j,k+1} - \overline{x}_j, x_{i,k+1} - \overline{x}_i) + Q_{ji}(0, x_{i,k+1} - \overline{x}_i) - Q_{ji}(x_{j,k+1} - \overline{x}_j, 0)\} + \sum_{i<j\le n} \{Q_{ij}(x_{i,k+1} - \overline{x}_i, x_{j,k} - \overline{x}_j) + Q_{ij}(x_{i,k+1} - \overline{x}_i, 0) - Q_{ij}(0, x_{j,k} - \overline{x}_j)\}. \tag{24}$$

Now we want to compute $\sum_{i=1}^n S_i$. The sum is given by the right hand member of (24) where $i$ is allowed to vary from 1 to $n$. But then, permuting $i$ and $j$ in the first sum $\sum_{1\le j<i\le n}$ is valid; so we have

$$\sum_{i=1}^n S_i = \sum_{1\le i<j\le n} \{Q_{ij}(x_{i,k+1} - \overline{x}_i, x_{j,k+1} - \overline{x}_j) + Q_{ij}(0, x_{j,k+1} - \overline{x}_j) - Q_{ij}(x_{i,k+1} - \overline{x}_i, 0)\} + \sum_{1\le i<j\le n} \{Q_{ij}(x_{i,k+1} - \overline{x}_i, x_{j,k} - \overline{x}_j) + Q_{ij}(x_{i,k+1} - \overline{x}_i, 0) - Q_{ij}(0, x_{j,k} - \overline{x}_j)\}.$$

Rearranging the terms we obtain

$$\begin{aligned}\sum_{i=1}^n S_i &= \sum_{1\le i<j\le n} \{Q_{ij}(x_{i,k+1} - \overline{x}_i, x_{j,k+1} - \overline{x}_j) + Q_{ij}(x_{i,k+1} - \overline{x}_i, x_{j,k} - \overline{x}_j)\} \\ &\quad + \sum_{1\le i<j\le n} Q_{ij}(0, x_{j,k+1} - \overline{x}_j) - \sum_{1\le i<j\le n} Q_{ij}(0, x_{j,k} - \overline{x}_j) \\ &= Q(x_{1,k+1} - \overline{x}_1, \dots, x_{n,k+1} - \overline{x}_n) + \sum_{1\le i<j\le n} Q_{ij}(x_{i,k+1} - \overline{x}_i, x_{j,k} - \overline{x}_j) \\ &\quad + \sum_{1\le i<j\le n} Q_{ij}(0, x_{j,k+1} - \overline{x}_j) - \sum_{1\le i<j\le n} Q_{ij}(0, x_{j,k} - \overline{x}_j).\end{aligned} \tag{25}$$

Finally, collecting (21, 22, 23, 25) we obtain

$$\begin{aligned}&\sum_{i=1}^n \|x_{i,k+1} - x_{i,k}\|^2 + Q(x_{1,k+1} - \overline{x}_1, \dots, x_{n,k+1} - \overline{x}_n) \\ &\quad + \sum_{1\le i<j\le n} Q_{ij}(x_{i,k+1} - \overline{x}_i, x_{j,k} - \overline{x}_j) + \sum_{i=1}^n \|x_{i,k+1} - \overline{x}_i\|^2 \\ &\quad + \sum_{1\le i<j\le n} Q_{ij}(0, x_{j,k+1} - \overline{x}_j) - \sum_{i=1}^n \|x_{i,k} - \overline{x}_i\|^2 - \sum_{1\le i<j\le n} Q_{ij}(0, x_{j,k} - \overline{x}_j) \le 0\end{aligned}$$

Thence we draw the following consequences:

- the sequence $k \to (x_{1,k}, \dots, x_{n,k})$ is bounded in $\prod_{i=1}^n \mathcal{X}_i$;
- the sequence $k \to \sum_{i=1}^n \|x_{i,k} - \overline{x}_i\|^2 + \sum_{1\le i<j\le n} Q_{ij}(0, x_{j,k} - \overline{x}_j)$ is nonincreasing;
- the quantities $\|x_{i,k+1} - x_{i,k}\|$ vanish as $k$ goes to $+\infty$.

We omit the rest of the proof which runs as before in the proof of Theorem 2.1.

### 3.2. The case of maximal monotone operators.

This section is an easy extension of Sections 2.3 and 3.1. We only state the assumptions, the algorithm and the theorem without proof.

$(\mathcal{H}'_n)$
- $(\mathcal{X}_i)_{i\in\{1,\ldots,n\}}$ $n$ real Hilbert spaces;
- for each $i \in \{1,\ldots,n\}$, $A_i : x \in \mathcal{X}_i \to A_i(x) \subseteq \mathcal{X}_i$ is a maximal monotone operator;
- for each $i, j$ with $1 \le i < j \le n$, $Q_{ij} : (x_i, x_j) \in \mathcal{X}_i \times \mathcal{X}_j \mapsto Q_{ij}(x_i, x_j) \in \mathbb{R}$ is a nonnegative continuous quadratic form; and define $Q : (x_1,\ldots,x_n) \in \prod_{i=1}^n \mathcal{X}_i \mapsto Q(x_1,\ldots,x_n) = \sum_{1\le i<j\le n} Q_{ij}(x_i,x_j) \in \mathbb{R}$;
- the maximal monotone operator
$$T = (A_1,\ldots,A_n) + \frac{1}{2}\nabla Q : \prod_{i=1}^n \mathcal{X}_i \rightrightarrows \prod_{i=1}^n \mathcal{X}_i$$
has at least one zero point.

To find zero points of $T$ we propose the following alternate algorithm for successively computing the components $(x_{i,k+1})_{i\in\{1,..,n\}}$ at the $(k+1)$-th step

$$(\mathcal{A}'_n) \qquad A_i(x_{i,k+1}) + \frac{1}{2}\nabla_{x_i} Q(x_{1,k+1},\ldots,x_{i-1,k+1},x_{i,k+1},x_{i+1,k},\ldots,x_{n,k}) + (x_{i,k+1} - x_{i,k}) \ni 0,\ i \in \{1,..,n\}$$

**Theorem 3.2.** *Under assumptions $(\mathcal{H}'_n)$, the sequence $k \to (x_{i,k})_{i\in\{1,..,n\}}$ generated by the alternate algorithm $(\mathcal{A}'_n)$ converges weakly in $\prod_{i=1}^n \mathcal{X}_i$ to a zero point $(\overline{x}_i)_{i\in\{1,..,n\}}$ of the maximal monotone operator $T = (A_1,\ldots,A_n) + \frac{1}{2}\nabla Q$.*

## 4. Applications to decision sciences.

### 4.1. Dynamical team games.

Let us study some dynamical games with the help of the results of Sections 2 and 3. We will be mostly concerned with the description and the study of their equilibration processes. The general setting of this study has been introduced in Section 1.3.

Let us suppose that the cost to change for player 1 when passing from performance $x$ to $\xi$ is given by $h(x,\xi) = \frac{\alpha}{2}\|x-\xi\|^2_{\mathcal{X}}$. Similarly, the cost to change for player 2 when passing from performance $y$ to $\eta$ is given by $k(y,\eta) = \frac{\nu}{2}\|y-\eta\|^2_{\mathcal{Y}}$. These are **low local costs to change**. The quadratic character of these costs reflects the fact that small changes are nearly costless for the players. On the opposite, big changes are very costly. Besides these qualitative aspects, parameters $\alpha$ and $\nu$ allow a quantitative description of these phenomena.

Note that the metric structures on the decision spaces $\mathcal{X}$ and $\mathcal{Y}$ are not a priori given. Indeed, they result from the modelling of the physiological, psychological features of the players, and of the perception the players have of their environment. They reflect the difficulties faced by the players when they want to progress, to move, to change in their respective decision spaces. Among many other features, they take into account their risk aversion, their inertial features (difficulty to quit a routine, reactivity). The coupling function $Q : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$ is supposed to be of the following form

$$Q(x, y) = \|Ax - By\|^2_{\mathcal{Z}}$$

where $\mathcal{Z}$ is a third Hilbert space, $A \in L(\mathcal{X}, \mathcal{Z})$ and $B \in L(\mathcal{Y}, \mathcal{Z})$ are linear continuous operators acting respectively from $\mathcal{X}$ to $\mathcal{Z}$ and from $\mathcal{Y}$ to $\mathcal{Z}$. Thus, $Q$ is a continuous convex quadratic function. Let us list some properties resulting from this type of coupling:

1. The coupling function can be interpreted as a joint payoff, either a joint cost or a joint revenue (see Section 1.3).
2. The above formulation allows to consider the general situation where the decision spaces are distinct one from the other, each player having his own decision space.
3. As before, the quadratic character of the coupling coordination cost reflects the fact that small deviations of $Ax - By$ from zero (which is the perfect coupling) are costless.

As introduced in Section 1.3, we consider the following dynamic of actions where players 1 and 2 play alternatively:

$$(x_k, y_k) \longrightarrow (x_{k+1}, y_k) \longrightarrow (x_{k+1}, y_{k+1}) \quad k = 0, 1, ...$$

$$\begin{cases} x_{k+1} = \text{argmin}\{f(\xi) + \frac{\mu_1}{2}\|A\xi - By_k\|^2_{\mathcal{Z}} + \frac{\alpha}{2}\|\xi - x_k\|^2_{\mathcal{X}} : \ \xi \in \mathcal{X}\} \\ y_{k+1} = \text{argmin}\{g(\eta) + \frac{\mu_2}{2}\|Ax_{k+1} - B\eta\|^2_{\mathcal{Z}} + \frac{\nu}{2}\|\eta - y_k\|^2_{\mathcal{Y}} : \ \eta \in \mathcal{Y}\} \end{cases} \tag{26}$$

Coefficients $\mu_1$ and $\mu_2$, when there are distinct, allow to distinguish the respective importance of the collective aspects in the decisions of the players.

One can easily recover the situation studied in Section 2. Just notice that

$$\begin{cases} x_{k+1} = \text{argmin}\left\{\frac{1}{\mu_1}f(\xi) + \frac{1}{2}\|A\xi - By_k\|^2_{\mathcal{Z}} + \frac{\alpha}{2\mu_1}\|\xi - x_k\|^2_{\mathcal{X}} : \ \xi \in \mathcal{X}\right\} \\ y_{k+1} = \text{argmin}\left\{\frac{1}{\mu_2}g(\eta) + \frac{1}{2}\|Ax_{k+1} - B\eta\|^2_{\mathcal{Z}} + \frac{\nu}{2\mu_2}\|\eta - y_k\|^2_{\mathcal{Y}} : \ \eta \in \mathcal{Y}\right\}. \end{cases} \tag{27}$$

The equilibria are the solutions of the following convex minimization problem:

$$\min\left\{\frac{1}{\mu_1}f(\xi) + \frac{1}{\mu_2}g(\eta) + \frac{1}{2}\|A\xi - B\eta\|^2_{\mathcal{Z}} : \ \xi \in \mathcal{X}, \ \eta \in \mathcal{Y}.\right\} \tag{28}$$

Assuming that assumptions $(\mathcal{A})$ of Theorem 2.1 are satisfied, we conclude that the sequence $(x_k, y_k)$ generated by the alternating dynamical game (26) converges weakly in $\mathcal{X} \times \mathcal{Y}$ to a solution $(x_\infty, y_\infty)$ of (28). Equivalently, $(x_\infty, y_\infty)$ is a minimum point of the functional

$$L_{\mu_1,\mu_2}(\xi, \eta) = \mu_2 f(\xi) + \mu_1 g(\eta) + \frac{\mu_1\mu_2}{2}\|A\xi - B\eta\|^2_{\mathcal{Z}}. \tag{29}$$

**Comments**:

- The above dynamic is an "inertial Nash equilibration process" [4]. Its trajectories converge to the Nash equilibria of the potential game associated with the following payoff functions (respectively to the first and second player):

$$\begin{cases} F(\xi, \eta) = f(\xi) + \frac{\mu_1}{2}\|A\xi - B\eta\|^2_{\mathcal{Z}} \\ G(\xi, \eta) = g(\eta) + \frac{\mu_2}{2}\|A\xi - B\eta\|^2_{\mathcal{Z}} \end{cases} \tag{30}$$

Because of the partial separable structure of these payoffs, of the fact that the coupling function is the same in the two expressions above (up to a multiplicative factor) and of the convexity assumptions, one obtains that these Nash equilibria are the minimizers of the functional $L_{\mu_1,\mu_2}$.

- Another important feature of the above dynamic is its sequential character. The players make decisions in a predefined order after having observed the decisions of the players who preceded them. We have examined here the quite simple situation with two players who play alternatively. Let us notice that this kind of dynamic is decentralized and does not require a high degree of rationality from the agents (one can model myopic local behaviours as well as inertial features, risk aversion...).
  This makes contrast with "simultaneous games" where players play simultaneously without knowing the decisions of the other players (but they may know the past, i.e. what has been played before). In that case, an efficient coupling between the players usually requires the help of a co-ordinator, a coach who has a global vision of the collective interest.
- In the $n$ dimensional case, our model can be applied to examine bilateral nonsymmetric interaction games (then, they fail to be potential games, see [18]). There are $n$ players with static utility functions

  $$J_i(x_1, x_2, ..., x_n) = \Sigma_{j \neq i} w_{ij}(x_i, x_j) - h_i(x_i), x_i \in X_i, i \in \{1, 2, ..., n\}.$$

  We do not need symmetric interaction functions $w_{ij}(x_i, x_j) = w_{ji}(x_j, x_i)$. But our interaction functions are quadratic. Cournot games with a linear final demand function can be examined in this way.
- We can apply our model to pollution games where the static utility functions of the $n$ players are

  $$J_i(x_1, x_2, ..., x_n) = h_i(x_i) - D\left[\Sigma_{j=1}^n e_j x_j\right], i \in \{1, 2, ..., n\}.$$

  The individual payoff of player $i$ is his profit function $h_i(x_i)$, $x_i \in X_i = R^+$ is his effort level, his pollution emission level is $e_j x_j, e_j > 0$, and the total damage function is $D[\Sigma_{j=1}^n e_j x_j] = (\Sigma_{j=1}^n e_j x_j)^2$.

### 4.2. Solving dissimilar problems: a cognitive alternating algorithm.

We consider the following model situation: an agent with human cognitive features (for example a manager, a decision maker, a research worker...) has several problems to solve, or has to solve a problem involving several criteria. When these problems are difficult to solve or involve dissimilar features, it is impossible for the agent to treat them simultaneously. He has to accept to handle them **sequentially, alternatively**. So doing, he enters into a dynamical process. We will touch only some aspects of this complex dynamic. Our main purpose is to show its close connection with the inertial equilibration process for weakly coupled systems, which has been described in the previous section.

Let us first describe some typical situations involving the resolution of several dissimilar problems. For a university staff member, it is a recurrent question to decide how to organize his activities between teaching, research and administrative duties. In the economic field, it is difficult for a firm (or an institution, or a state) to progress simultaneously in the production, competition, social and environmental aspects. Most often, the agent chooses to concentrate his activity on one of these tasks during a sufficiently long period, and then to switch over to another one and so on. At each step, he does his best to get his current problem or criterion progressing. Then he turns to the next problem or criterion which he tries to improve, and so on.

Passing from one activity to the other usually involves classical costs to move which account for various physical or physiological aspects (reactivity, inertia, speed of decision making, see [5]). When solving different problems alternatively, the difficulty for the agent comes mostly from the dissimilarity features of these different activities. He has to manage his brain, switch off the neural network corresponding to the first activity, and then, switch on the neural network corresponding to the next one. This leads to introduce *dissimilarity costs* (viewed as transition costs) between these different activities. They prevent him from switching too often from one activity to the other.

A closely related situation occurs when the agent has a problem to solve involving different, antagonistic criteria. The agent has to be careful not to destroy too much the quality of the preceding criteria, when improving the current one. This leads, in a similar way, to introduce costs to change which anchor each decision in the preceding ones and prevent them from being too different.

Thus, we are led to introduce the following mathematical model. As we will see, it is very similar to the previous one (despite the fact that we have now only one player).

1. A decision $x = (x, y)$ is a couple where $x \in \mathcal{X}$ is the component relative to the first problem while $y \in \mathcal{Y}$ is the component corresponding to the second one. Problems are seen as tasks to do and the state space $\mathcal{X}$ (respectively $\mathcal{Y}$) is the state space of performance levels $x$ (respectively $y$). For a given task, say task 1, $x \in \mathcal{X}$ is a vector of performance levels $x_i$, given a list of subtasks $i \in I$. For example if $\mathcal{X} = \mathbb{R}^n$, the first task, with degree of complexity $n$, is made of $n$ subtasks $I = \{1, 2, \dots, n\}$. Given $x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$, $x_i \in \mathbb{R}$ is the level of performance reached for the subtask $i$. In order to represent both the underlying state space of actions and the physical, psychological or cognitive efforts which are needed and help to reach such performances, it is convenient to equip these spaces with *metric* structures: actually, $\mathcal{X}$, the state space of performance levels for task 1, is a real Hilbert space, with $\|x\|_{\mathcal{X}}$ denoting the norm of an element $x \in \mathcal{X}$; the distance between two vectors of performance levels $x \in \mathcal{X}$ and $\xi \in \mathcal{X}$ is measured by $\|x - \xi\|$. Similarly for the second task and space $\mathcal{Y}$.
2. We suppose that the cost to change for the task 1, when passing from the performance level $x$ to $\xi$, is given by $h(x, \xi) = \frac{\alpha}{2}\|x - \xi\|^2_{\mathcal{X}}$. Similarly, the cost to change for the task 2, when passing from the performance level $y$ to $\eta$, is given by $k(y, \eta) = \frac{\nu}{2}\|y - \eta\|^2_{\mathcal{Y}}$. This corresponds to low local costs to change. These costs mainly describe an anchoring effect. This means that the present perception of the cost of the way $\xi$ of doing task 1 depends on the previous way $x$ of doing this task before. For one part, they are psychological costs of perception, the perception of the costs of doing something in a new way being influenced by the old way of doing. They are also costs to learn how to do better and better the task, given the way this task has been done before.
3. To each $x \in \mathcal{X}$ the agent attaches a value $f(x) \in \mathbb{R} \cup \{+\infty\}$ which measures the quality of the decision $x \in \mathcal{X}$ in relation to the first problem. Similarly, to each $y \in \mathcal{Y}$ the agent attaches a value $g(y) \in \mathbb{R} \cup \{+\infty\}$ which measures the quality of

the decision $y \in \mathcal{Y}$ in relation to the second problem. The value $+\infty$ takes account of possible constraints.

4. Let us complete the picture by introducing, as a basic ingredient, the switching cost $Q$ which couples the two resolution processes. Depending on the context, we use as synonyms switching costs, dissimilarity costs or coupling functions.
The switching cost is supposed to be a quadratic function of the following form

$$Q(x, y) = \|Ax - By\|^2_{\mathcal{Z}}$$

where $\mathcal{Z}$ is a third Hilbert space, $A \in L(\mathcal{X}, \mathcal{Z})$ and $B \in L(\mathcal{Y}, \mathcal{Z})$ are linear continuous operators acting respectively from $\mathcal{X}$ to $\mathcal{Z}$ and from $\mathcal{Y}$ to $\mathcal{Z}$.
We stress the fact that the **weak coupling** property, as described above, is essential in order to reflect the fact that, when comparing two vectors $x$ and $y$ attached to different activities, it has no meaning to compare them directly. By contrast, in order to evaluate the difficulty to pass from performance $x$ relative to the first problem to performance $y$ relative to the second one, the agent has to construct some "**cognitive representations**" $Ax$ and $By$ of them. For example $Ax$ is a vector whose components are made of some general psychological or cognitive features of the performance $x$, like its degree of difficulty, intensity (speed, energy), level of interest, social status... Similarly for $By$. Now it makes sense to compare $Ax$ and $By$. For example, passing from $x$ to $y$ may correspond to the agent spending more energy, going faster, for a more (or less) interesting activity... with a corresponding transition cost. We can say that $u = Ax$ and $v = By$ represent different "states of mind". The coupling function $Q(x, y)$ means that it is costly to change one's state of mind, to stop focusing on a goal and to redirect one's attention to a new goal.

5. Another interpretation of the coupling term $Q(x, y)$ as a nonnegative quadratic form comes from fatigue costs.
When an agent does a first complex task made of several subtasks $i \in \{1, 2, ..., n\}$ and makes the efforts $x_i \in \mathbb{R}^+$ to do each subtask $i$, his fatigue feeling is $\Phi(x) = \Sigma_i \lambda_i x_i, \lambda_i > 0$. When this agent starts doing a new complex task made of several subtasks and makes an effort $y_j$ to do each new subtask $j \in \{1, 2, ..., m\}$, his unit cost $c_j$ to do a subtask $j$ increases with his fatigue $c_j = c_j(x) = \mu_j \Phi(x)$, with $\mu_j > 0$. Then, the total cost of doing the new task, having done before the first task, is the quadratic term $c(x, y) = \Sigma_{j=1}^m c_j(x) y_j = \Sigma_{ij} \lambda_i \mu_j x_i y_j$.

6. We can now describe the alternate dynamical system

$$(x_k, y_k) \to (x_{k+1}, y_k) \to (x_{k+1}, y_{k+1}) \quad k = 0, 1, ...$$

and the costs to change which accompany each of its steps.
When passing from $(x_k, y_k)$ to $(x_{k+1}, y_k)$, the performance $x_{k+1}$ is chosen in an optimal way among performances $\xi \in \mathcal{X}$ as follows:

$$x_{k+1} = \operatorname{argmin}\left\{ f(\xi) + \frac{\mu_1}{2}\|A\xi - By_k\|^2_{\mathcal{Z}} + \frac{\alpha}{2}\|\xi - x_k\|^2_{\mathcal{X}} : \ \xi \in \mathcal{X} \right\}. \tag{31}$$

Similarly, when passing from $(x_{k+1}, y_k)$ to $(x_{k+1}, y_{k+1})$

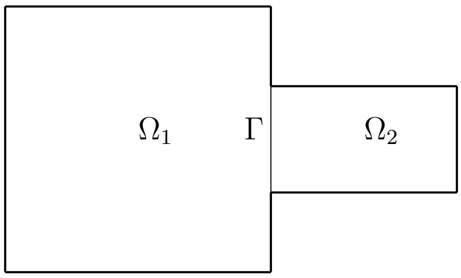
$$y_{k+1} = \operatorname{argmin}\left\{ g(\eta) + \frac{\mu_2}{2}\|Ax_{k+1} - B\eta\|^2_{\mathcal{Z}} + \frac{\nu}{2}\|\eta - y_k\|^2_{\mathcal{Y}} : \ \eta \in \mathcal{Y} \right\}. \tag{32}$$

That is exactly the same dynamic as in the previous section with a similar conclusion, that is the convergence of the sequence $(x_k, y_k)$ to a solution $(x_\infty, y_\infty)$ of the convex minimization problem (28) (when assumptions $(\mathcal{A})$ are met).

7. Note that a more realistic model should address the difficult questions:
   - How to choose the optimal, or nearly optimal time schedule for doing each task in alternation, like the length of time spent each period to do a task and then, the other one, when to stop one task, and start the second (this leads to consider parameters $\alpha$, $\mu$, $\nu$ as control variables).
   - In our situation, the switching cost satisfies a symmetry condition, i.e. the cost that the agent endures when passing from a performance $x$ related to the first problem to a performance $y$ related to the second one is the same as the cost that he endures when passing from $y$ to $x$. This is quite a restrictive assumption, a more realistic model would allow dissymmetry (one may think of using a relative entropy, but, in that case, one loses the quadratic property).
   - Enlarge the class of payoff functions $f$ and $g$ passing, for example, from convex to quasiconvex or subanalytic functions.

## 5. Application to splitting methods for partial differential equations.

Let us show how algorithm $(\mathcal{A})$ can be applied so as to obtain new PDE domain decomposition techniques (see [17] for a reference). Making here a systematic study of this important subject would exceed by far the scope of the paper. We just focus our attention on a model situation with the description of the algorithm together with its convergence properties. As a general rule for this kind of problem, the coupling terms involve the Sobolev trace operators on the common interfaces between the subdomains. Let us illustrate this in the following case, with $n = 2$ subdomains, and consider a decomposition of the domain $\Omega = \Omega_1 \cup \Omega_2 \cup \Gamma$ into two nonoverlapping adjacent subdomains $\Omega_1$ and $\Omega_2$ with a common interface $\Gamma$. This is illustrated in the following figure




Given some $h \in L^2(\Omega)$ and some parameter $\mu > 0$ let us consider the following variational problem on $\Omega$

$$\min\left\{\frac{1}{2}\int_{\Omega_1} \nabla v_1|^2 + \frac{1}{2}\int_{\Omega_2} |\nabla v_2|^2 + \frac{\mu}{2}\int_{\Gamma}[v]^2 - \int_{\Omega} hv \;:\; v_1 \in H^1(\Omega_1), v_2 \in H^1(\Omega_2), v = 0 \text{ on } \partial\Omega\right\} \tag{33}$$

where

$$\begin{cases} v = v_1 & \text{on } \Omega_1 \\ v = v_2 & \text{on } \Omega_2 \end{cases}$$

and

$$[v] = \text{jump of } v \text{ through the interface } \Gamma.$$

This kind of variational problem frequently occurs in various domains like mechanics (transmission through thin isolating layers, cracks, fissures), imaging (segmentation, image restoration), thermodynamics (phase transitions), more generally in the description of phenomena involving discontinuities on the interfaces between subdomains, see [2, 3, 7, 8, 11].

It falls into the mathematical formulation (2):

$$\min\left\{f_1(v_1) + f_2(v_2) + \frac{\mu}{2}\|A_1 v_1 - A_2 v_2\|^2_{\mathcal{Z}} : \ v_1 \in \mathcal{X}_1, \ v_2 \in \mathcal{X}_2\right\} \tag{34}$$

where

$\mathcal{X}_1 = \{v \in H^1(\Omega_1), v = 0 \text{ on } \partial\Omega \cap \partial\Omega_1\}, A_1 : H^1(\Omega_1) \to \mathcal{Z} = L^2(\Gamma)$ is the trace operator
$\mathcal{X}_2 = \{v \in H^1(\Omega_2), v = 0 \text{ on } \partial\Omega \cap \partial\Omega_2\}, A_2 : H^1(\Omega_2) \to \mathcal{Z} = L^2(\Gamma)$ is the trace operator

and

$$f_1(v_1) = \frac{1}{2}\int_{\Omega_1} |\nabla v_1|^2 - \int_{\Omega_1} h v_1$$
$$f_2(v_2) = \frac{1}{2}\int_{\Omega_2} |\nabla v_2|^2 - \int_{\Omega_2} h v_2.$$

Note that

$$A_1(v_1) - A_2(v_2) = [v] \text{ is the jump of } v \text{ through the interface } \Gamma.$$

Let us explicit algorithm $(\mathcal{A})$.

Let us denote by $u_k = (u_{1,k}, u_{2,k}) \in \mathcal{X}_1 \times \mathcal{X}_2$ the current point at step $k$ generated by the algorithm. We have

$$\begin{cases} u_{1,k+1} = \operatorname{argmin}\{f_1(v_1) + \frac{\mu}{2}\|A_1 v_1 - A_2 u_{2,k}\|^2_{\mathcal{Z}} + \frac{\alpha}{2}\|v_1 - u_{1,k}\|^2_{\mathcal{X}_1} : \ v_1 \in \mathcal{X}_1\} \\ u_{2,k+1} = \operatorname{argmin}\{f_2(v_2) + \frac{\mu}{2}\|A_1 u_{1,k+1} - A_2 v_2\|^2_{\mathcal{Z}} + \frac{\nu}{2}\|v_2 - u_{2,k}\|^2_{\mathcal{X}_2} : \ v_2 \in \mathcal{X}_2\} \end{cases} \tag{35}$$

where $\alpha$ and $\nu$ are given fixed positive parameters.

Let us write the optimality conditions (Euler equations) of the above variational problems. Let us denote by $Q : \mathcal{X}_1 \times \mathcal{X}_2 \to \mathcal{Z} = L^2(\Gamma)$

$$Q(v) = \frac{1}{2}\|A_1 v_1 - A_2 v_2\|^2_{L^2(\Gamma)} = \frac{1}{2}\int_\Gamma |A_1 v_1 - A_2 v_2|^2$$

the convex quadratic coupling function.

An elementary directional derivative computation yields

$$\lim_{t\to 0}\frac{1}{t}[Q(u + tv) - Q(u)] = \int_\Gamma (A_1 u_1 - A_2 u_2)(A_1 v_1 - A_2 v_2).$$

At this point, in order to go further, we need to specify the scalar product on $\mathcal{X}_1$ and $\mathcal{X}_2$

$$\langle u_i, v_i\rangle = \int_{\Omega_i} \nabla u_i . \nabla v_i, \ i \in \{1, 2\}$$

with the corresponding norm

$$\|v_i\|^2_{\mathcal{X}_i} = \int_{\Omega_i} |\nabla v_i|^2.$$

Let us recall that the elements of $\mathcal{X}_i$ are equal to zero on $\partial\Omega \cap \partial\Omega_i$. By Poincaré inequality, when $\Omega$ is a bounded set, this scalar product induces on $\mathcal{X}_i$ the usual topology of $H^1(\Omega_i)$.

With a similar directional compution as above we finally obtain the following weak variational formulation of algorithm $(\mathcal{A})$:

$$\forall v_1 \in \mathcal{X}_1 \quad \int_{\Omega_1} \nabla u_{1,k+1}.\nabla v_1 + \mu \int_{\Gamma} (A_1 u_{1,k+1} - A_2 u_{2,k}) A_1 v_1 + \alpha \int_{\Omega_1} (\nabla u_{1,k+1} - \nabla u_{1,k}) \nabla v_1 = \int_{\Omega_1} h v_1$$

$$\forall v_2 \in \mathcal{X}_2 \quad \int_{\Omega_2} \nabla u_{2,k+1}.\nabla v_2 + \mu \int_{\Gamma} (A_2 u_{2,k+1} - A_1 u_{1,k+1}) A_2 v_2 + \nu \int_{\Omega_2} (\nabla u_{2,k+1} - \nabla u_{2,k}) \nabla v_2 = \int_{\Omega_1} h v_2.$$

These are the variational weak formulations of the following Dirichlet-Neumann boundary value problems respectively on $\Omega_1$

$$\begin{cases} -(1-\alpha)\Delta u_{1,k+1} = h + \alpha \Delta u_{1,k} & \text{on } \Omega_1 \\ (1+\alpha)\dfrac{\partial u_{1,k+1}}{\partial \nu_1} + \mu u_{1,k+1} = \mu u_{2,k} + \alpha \dfrac{\partial u_{1,k}}{\partial \nu_1} & \text{on } \Gamma \\ u_{1,k+1} = 0 & \text{on } \partial\Omega_1 \cap \partial\Omega \end{cases}$$

and $\Omega_2$

$$\begin{cases} -(1-\nu)\Delta u_{2,k+1} = h + \nu \Delta u_{2,k} & \text{on } \Omega_2 \\ (1+\nu)\dfrac{\partial u_{2,k+1}}{\partial \nu_2} + \mu u_{2,k+1} = \mu u_{1,k+1} + \nu \dfrac{\partial u_{2,k}}{\partial \nu_2} & \text{on } \Gamma \\ u_{2,k+1} = 0 & \text{on } \partial\Omega_2 \cap \partial\Omega \end{cases}$$

We have adopted the classical notations, $\frac{\partial z_i}{\partial \nu_i}$ is the derivative of $z_i$ in the direction of $\nu_i$ which is the normal to $\Gamma$ oriented outwards of $\Omega_i$.

Theorem 2.1 now shows that the above algorithm converges. Indeed, we have both weak convergence of the sequences $u_{1,k}$, $u_{2,k}$ and convergence of the corresponding energy functionals. As a result, the sequence $(u_{1,k}, u_{2,k})$ converges strongly in $H^1(\Omega_1) \times H^1(\Omega_2)$ to a minimum point $(\overline{u}_1, \overline{u}_2)$ of problem (33).

Let us notice that the initial problem on $\Omega$ has been entirely decomposed into subproblems on $\Omega_1$ and $\Omega_2$. Let us stress that functionals $f_1$ and $f_2$ have just been assumed to be closed and convex, which allows to treat by this method a large class of nonlinear problems. The method does apply to an arbitrary decomposition of the domain $\Omega$ into subdomains $\Omega_1, \Omega_2, ..., \Omega_m$ (non overlapping like in the above example or possibly overlapping). There is still much to do in order to develop the algorithm in an operational way. One has to consider the limiting case $\mu = +\infty$ in order to treat PDE decomposition problems. We have only considered here a relaxed (penalized) version of the variational formulation.

Lagrangian and multipliers methods are certainly in the picture. Finally, in order to end up by solving finite dimensional problems, one has to combine the algorithm with a numerical method like finite elements, finite difference, spectral method.

## References

[1] F. Acker, M.-A. Prestel: Convergence d'un schéma de minimisation alternée, Ann. Fac. Sci. Toulouse, V. Ser., Math. 2 (1980) 1–9.

[2] H. Attouch: Variational Convergence for Functions and Operators, Applicable Mathematics Series, Pitman, London (1984).

[3] H. Attouch, G. Buttazzo, G. Michaille: Variational Analysis in Sobolev and BV Spaces. Applications to PDEs and Optimization, MPS/SIAM Series on Optimization 6, SIAM, Philadelphia (2006).

[4] H. Attouch, P. Redont, A. Soubeyran: A new class of alternating proximal minimization algorithms with costs-to-move, SIAM J. Optim. 18 (2007) 1061–1081.

[5] H. Attouch, A. Soubeyran: Inertia and reactivity in decision making as cognitive variational inequalities, J. Convex Analysis 13 (2006) 207–224.

[6] H. Attouch, A. Soubeyran: A worthwhile to move approach of satisficing with not too much sacrificing, J. Math. Psychol., submitted.

[7] H. H. Bauschke, P. L. Combettes, D. Noll: Joint minimization with alternating Bregman proximity operators, Pac. J. Optim. 2(3) (2006) 401–424.

[8] H. H. Bauschke, P. L. Combettes, S. Reich: The asymptotic behavior of the composition of two resolvents, Nonlinear Anal., Theory Methods Appl. 60A (2005) 283–301.

[9] H. Brézis: Opérateurs maximaux monotones dans les espaces de Hilbert et équations d'évolution, Lecture Notes 5, North-Holland, Amsterdam (1972).

[10] C. Camerer, G. Loewenstein: Behavioral economics: past, present, future, in: Advances in Behavioral Economics, C. Camerer, G. Loewenstein, M. Rabin (eds.), Princeton University Press, Princeton (2003) 3–51.

[11] P. L. Combettes: The foundations of set theoretic estimation, Proc. IEEE 81 (1993) 182–208.

[12] D. Kahneman: Maps of bounded rationality: psychology for behavioral economics, Amer. Econ. Rev. (2003) 1449–1475.

[13] J.-L. Lions: Contrôle des Systèmes Distribués Singuliers, Gauthier-Villars, Paris (1983).

[14] J. von Neumann: Functional Operators. Vol. II. The Geometry of Orthogonal Spaces, Annals of Mathematics Studies 22, Princeton University Press, Princeton (1950).

[15] Z. Opial: Weak convergence of the sequence of successive approximations for nonexpansive mappings, Bull. Amer. Math. Soc. 73 (1967) 591–597.

[16] H. Simon: A behavioral model of rational choice, Quart. J. Econ. 69 (1975) 99–118.

[17] P. Le Tallec: Domain decomposition methods in computational mechanics, Comput. Mech. Adv. 1(2) (1994) 121–220.

[18] T. Ui: A Shapley value representation of potential games, Games Econ. Behav. 31 (2000) 121–135.