A proximal multiplier method for separable convex minimization

O. Sarmiento C.*, E. A. Papa Quiroz† and P. R. Oliveira‡

Systems Engineering and Computer Science, COPPE, Federal University of Rio de Janeiro, postal address: 68511, CEP: 21941-972, Rio de Janeiro, Brazil.

(*July 2014*)

In this paper, we propose an inexact proximal multiplier method using proximal distances for solving convex minimization problems with a separable structure. The proposed method unified the work of Chen and Teboulle (PCPM method), Kyono and Fukushima (NPCPMM) and Auslender and Teboulle (EPDM) and extends the convergence properties for the class of $\varphi-$divergence distances. We prove, under standard assumptions, that the iterations generated by the method are well defined and the sequence converges to an optimal solution of the problem.

Keywords: Proximal multiplier methods; separable convex problems; proximal distances; convex functions.

AMS Subject Classification:

1. Introduction

In this paper, we are interesting in solving the following separable convex optimization problem:

$$(CP) \quad \min\{f(x) + g(z) \,:\, Ax + Bz = b,\, x \in \bar{\mathtt{C}},\, z \in \bar{\mathtt{K}}\},$$

where $\mathtt{C} \subset I\!R^n$ and $\mathtt{K} \subset I\!R^p$ are nonempty open convex sets, $\bar{\mathtt{C}}$ and $\bar{\mathtt{K}}$ denote the closure (in the euclidean topology) of $\mathtt{C}$ and $\mathtt{K}$ respectively, $f : I\!R^n \to (-\infty, +\infty]$ and $g : I\!R^p \to (-\infty, +\infty]$ are closed proper convex functions and $A \in I\!R^{m\times n}$, $B \in I\!R^{m\times p}$, $b \in I\!R^m$.

The Fenchel dual problem, see [22], for (CP) is defined as:

$$(D) \quad \max\{-(f + \delta_{\bar{\mathtt{C}}})^*(-A^T y) - (g + \delta_{\bar{\mathtt{K}}})^*(-B^T y) - \langle y, b\rangle : y \in I\!R^m\}$$

where $\delta_{\mathtt{X}}$ denotes the indicator function of a subset X, $(f + \delta_{\bar{\mathtt{C}}})^*$, $(g + \delta_{\bar{\mathtt{K}}})^*$ are the conjugate functions of $f+\delta_{\bar{\mathtt{C}}}$ and $g+\delta_{\bar{\mathtt{K}}}$, respectively, and $\langle \cdot, \cdot\rangle$ denotes the canonical inner product .

The Lagrangian for (CP) is defined by $L : I\!R^n \times I\!R^p \times I\!R^m \to (-\infty, +\infty]$,

$$L(x, z, y) = (f + \delta_{\bar{\mathtt{C}}})(x) + (g + \delta_{\bar{\mathtt{K}}})(z) + \langle y, Ax + Bz - b\rangle \,.$$

*Email: osarmiento@cos.ufrj.br

†Corresponding author. Email: erikpapa@gmail.com

‡Email: poliveir@cos.ufrj.br

where y is the Lagrangian multiplier associated with the constraint $Ax + Bz = b$.

Let $(x^*, z^*) \in \mathbb{R}^n \times \mathbb{R}^p$ and $y^* \in \mathbb{R}^m$. In order that (x^*, z^*) be an optimal solution of (CP) and y^* be an optimal Lagrangian multiplier, it is necessary and sufficient that (x^*, z^*, y^*) be a saddle point of the Lagrangian L of (CP) (see [22], Theorem 28.3); i.e., for all $x \in dom f \cap \bar{\mathtt{C}}$, $z \in dom\, g \cap \bar{\mathtt{K}}$ and $y \in \mathbb{R}^m$, we have

$$L(x^*, z^*, y) \leq L(x^*, z^*, y^*) \leq L(x, z, y^*).$$

In recent decades a great interest has emerged in studying the separable structure of the problem (CP), this model has been found in various optimization problems, for example in telecommunications, see Mahey et. al. [20], management electricity, see Lenoir [17], and in computer science to solve matrix completion problems, see example 2 of Goldfarb et. al. [14].

Classical methods for solving (CP) based on the proximal point algorithm (PPA) are the method of multipliers, [6], which is obtained by applying the PPA to the dual problem (D), and the proximal method of multipliers [23], which is obtained by applying the PPA to the primal-dual problem. However, these algorithms do not take full advantage of the separable structure of the original problem (CP).

Various decomposition methods that exploit the special structure of the problem have been proposed. Some examples of such methods are: alternating directions method of multipliers [7, 11, 12, 18], partial inverse of Spingarn method [10, 19, 20, 24, 25], predictor corrector proximal multipliers methods of Chen and Teboulle and its extensions [3, 9, 16].

For the particular case when $\mathtt{C} = \mathbb{R}^n$, $\mathtt{K} = \mathbb{R}^m$, $B = -I$ and $b = 0$ in (CP), the exact version of Chen and Teboulle [9], called Predictor-Corrector Proximal Multiplier (PCPM) method, given an initial point (x^0, z^0, y^0), generates a sequence of points $\{(x^k, z^k, y^k)\}$ with the following iterative scheme:

$$p^{k+1} = y^k + \lambda_k(Ax^k - z^k) \tag{1}$$

$$x^{k+1} = \arg\min \left\{ f(x) + \langle p^{k+1}, Ax \rangle + (1/2\lambda_k)||x - x^k||^2 \right\} \tag{2}$$

$$z^{k+1} = \arg\min \left\{ g(z) - \langle p^{k+1}, z \rangle + (1/2\lambda_k)||z - z^k||^2 \right\} \tag{3}$$

$$y^{k+1} = y^k + \lambda_k(Ax^{k+1} - z^{k+1}) \tag{4}$$

where λ_k is a positive parameter. Assuming that there exists an optimal primal-dual solution (x^*, z^*, y^*), Chen and Teboulle [9] proved that the sequence $\{(x^k, z^k, y^k)\}$ globally converges to (x^*, z^*, y^*), that is, converges to a primal and dual solution respectively. Furthermore, they proved a linear rate of convergence of the (PCPM) method.

Kyono and Fukushima [16], studied the same particular problem, the exact version of their algorithm, called Nonlinear Predictor-Corrector Proximal Multiplier Method (NPCPMM), substitute (2) and (3) by the following iterations:

$$x^{k+1} = \arg\min \left\{ f(x) + \langle p^{k+1}, Ax \rangle + (1/\lambda_k) D_h(x, x^k) \right\}$$

$$z^{k+1} = \arg\min \left\{ g(z) - \langle p^{k+1}, z \rangle + (1/\lambda_k) D_h(z, z^k) \right\}$$

where D_h is a Bregman distance. An advantage of this method over the (PCPM) algorithm is that, by a suitable choice of the Bregman distance, each subproblem becomes in an unconstrained problem. Under simmilar assumptions that Chen and Teboulle, the authors have been proved the global convergence of the iterations to a primal and dual solution respectively.

Independently to the work of Kyono and Fukushima [16], Auslender and Teboulle [3], studied convex optimization and variational inequality problems with a separable structure. For the optimization case, they considered the problem (CP) with $\bar{\mathtt{C}} = \mathbb{R}^n_+$ and $\bar{\mathtt{K}} = \mathbb{R}^m$. The authors, motivated by the earlier work of Auslender, Teboulle and Ben-Tiba [2], substituted the iteration (2) by:

$$x^{k+1} = \arg\min \left\{ f(x) + \langle p^{k+1}, Ax\rangle + (1/\lambda_k) d(x, x^k) \right\}$$

where

$$d(u,v) = \begin{cases} \sum_{i=1}^{n} \frac{\sigma}{2}(u_i - v_i)^2 + \mu(v_i^2 \log(\frac{v_i}{u_i}) + u_i v_i - v_i^2), & \text{if } u \in \mathbb{R}^n_{++} \\ +\infty, & \text{otherwise} \end{cases}$$

Under the same above assumption they proved the global convergence of the iterations to a primal and dual solution respectively.

In this paper, we consider the general problem (CP) and propose the following inexact iteration: Given $(x^k, z^k, y^k) \in \mathtt{C} \times \mathtt{K} \times \mathbb{R}^m$, find $(x^{k+1}, \upsilon^{k+1}) \in \mathtt{C} \times \mathbb{R}^n$ and $(z^{k+1}, \xi^{k+1}) \in \mathtt{K} \times \mathbb{R}^p$ such that

$$\begin{aligned}
p^{k+1} &= y^k + \lambda_k(Ax^k + Bz^k - b) \\
\upsilon^{k+1} &\in \partial_{a_k} f^k(x^{k+1}), \qquad \upsilon^{k+1} + \lambda_k^{-1}\nabla_1 d(x^{k+1}, x^k) = 0, \\
\xi^{k+1} &\in \partial_{b_k} g^k(z^{k+1}), \qquad \xi^{k+1} + \lambda_k^{-1}\nabla_1 d'(z^{k+1}, z^k) = 0. \\
y^{k+1} &= y^k + \lambda_k(Ax^{k+1} + Bz^{k+_1} - b)
\end{aligned}$$

where the functions $f^k : \mathbb{R}^n \to (-\infty, +\infty]$ and $g^k : \mathbb{R}^p \to (-\infty, +\infty]$ are defined by $f^k(x) = f(x) + \langle p^{k+1}, Ax\rangle$ and $g^k(z) = g(z) + \langle p^{k+1}, Bz\rangle$, respectively, with p^{k+1} as above, $\partial_{a_k} f^k$ and $\partial_{b_k} g^k$ are the a_k- and b_k- Fenchel subdifferential of f^k and g^k respectively, and d, d' are regularized proximal distances, see sections 3 and 4.

This method is an extension of the (PCPM) and (NPCPMM) methods and includes the class of $\varphi-$divergence distances, see Subsection 3.3 of [5], what to our knowledge, not yet been studied in this context. Observe also that the (EPDM) is a particular case of the our method when in our algorithm we consider exact iterations and we use the regularized log-quadratic distance, see Section 2 of [3]. Thus, the main result of this paper is show that the global convergence of the proposed method is still valid when we use regularized proximal distances.

The paper is organized as follows: In Section 2 we give some results in convex analysis. In Section 3 we present the class of proximal distances that we will use along the paper. In Section 4 we introduce the Proximal Multiplier Algorithm with Proximal Distances (PMAPD) and under some appropriate assumptions we prove the global convergence of the iterations to the solution of the (CP). Finally we present our conclusions and give some directions for future research.

2. Notations and some results in convex analysis

Throughout the paper $\mathbb{R}^n$ is the Euclidean space endowed with the canonical

inner product $\langle \cdot, \cdot \rangle$ and the norm of x given by $||x|| := \langle x, x\rangle^{1/2}$. For a matrix $M \in \mathbb{R}^{m\times n}$ we define $\|M\| := \max_{||x||\leq 1} \|Mx\|$. Given an extended real valued function $f : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ we denote its domain by $dom f := \{x \in \mathbb{R}^n : f(x) < +\infty\}$ and its epigraph $epi\, f := \{(x, \beta) \in \mathbb{R}^n \times \mathbb{R} : f(x) \leq \beta\}$. f is said to be proper, if $dom f \neq \emptyset$ and for all $x \in dom f$, we have $f(x) > -\infty$.
Also denote by $ri(X)$ the relative interior set of $X \subset \mathbb{R}^n$ and $\partial_\epsilon f$ is the $\epsilon-$subdiferential of f defined by

$$\partial_\epsilon f(u) = \{p \in \mathbb{R}^n : f(v) \geq f(u) + \langle p, v - u\rangle - \epsilon, \forall v \in dom\, f\}.$$

Finally, f is a lower semicontinuous function if for each $x \in \mathbb{R}^n$ we have that all $\{x^l\}$ such that $\lim_{l\to+\infty} x^l = x$ implies that $f(x) \leq \liminf_{l\to+\infty} f(x^l)$. It is easy to prove that the lower semicontinuity of f is equivalent to the closedness of the lower level set $L_f(\alpha) = \{x \in \mathbb{R}^n : f(x) \leq \alpha\}$, for each $\alpha \in \mathbb{R}$.

Definition 1 Let C be a nonempty convex set, the recession cone of C is definided by

$$0^+(C) = \{z \in \mathbb{R}^n \,:\, w + tz \in C,\ \forall w \in C,\, t \geq 0\},$$

furthermore, we define the the recession function $f0^+ : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ such that $epi(f0^+) = 0^+(epi\, f)$

Definition 2 Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ be a proper function, the conjugate of f, denoted by f^*, is defined by

$$f^*(x^*) = \sup_x [\langle x, x^*\rangle - f(x)].$$

PROPOSITION 2.1 ([22], *Corollary* 13.3.4, (*c*)) *Let $f : \mathbb{R}^n \to \mathbb{R}\cup\{\pm\infty\}$ be a closed proper convex function. Let x^* be a fixed vector and let $g(x) = f(x) - \langle x, x^*\rangle$. Then, $x^* \in int(dom\, f^*)$ if and only if $(g0^+)(y) > 0$ for every $y \neq 0$.*

PROPOSITION 2.2 ([22], *Theorem* 27.1, (*d*)) *Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ be a closed proper convex function. The minimum set of f is a non-empty and bounded set if and only if $0 \in int(dom\, f^*)$.*

PROPOSITION 2.3 ([22], *Theorem* 9.3) *Let $f_1, \ldots, f_m$ be proper convex functions on $\mathbb{R}^n$. If every f_i is closed and $f_1+\cdots+f_m$ is not identically $+\infty$, then $f_1+\cdots+f_m$ is a closed proper convex function and*

$$(f_1 + \cdots + f_m)0^+ = f_1 0^+ + \cdots + f_m 0^+.$$

PROPOSITION 2.4 ([22], *Theorem* 8.5) *Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ be a proper convex function. The recession function $f0^+$ of f is then a positively homogeneous proper convex function. For every vector y, one has*

$$(f0^+)(y) = \sup\{f(x + y) - f(x) \,:\, x \in dom f\}.$$

If f is closed, $f0^+$ is closed too, and for any $x \in dom f$, $f0^+$ is given by the formula:

$$(f0^+)(y) = \sup_{\lambda>0} \frac{f(x + \lambda y) - f(x)}{\lambda} = \lim_{\lambda\to+\infty} \frac{f(x + \lambda y) - f(x)}{\lambda}.$$

PROPOSITION 2.5 ([22], *Theorem* 23.8) *Let $f_1, \ldots, f_m$ be proper convex functions on $I\!R^n$ and let $f = f_1 + \cdots + f_m$. If the convex sets $ri(dom f_i)$, $i = 1, 2, \ldots, m$ have a point in common, then*

$$\partial f(x) = \partial f_1(x) + \cdots + \partial f_m(x), \ \ \forall x.$$

3. Proximal Distances

In this section, we present a variant of the definition of the proximal distance and induced proximal distance, introduced by Auslender and Teboulle [5].

Definition 3 A function $d : I\!R^n \times I\!R^n \to I\!R_+ \cup \{+\infty\}$ is called proximal distance with respect to an open nonempty convex set $\mathtt{C} \subset I\!R^n$ if for each $y \in \mathtt{C}$ it satisfies the following properties:

(i) $d(\cdot, y)$ is proper, lower semicontinuous, convex and continuously differentiable on $\mathtt{C}$;
(ii) $dom\ d(\cdot, y) \subset \bar{\mathtt{C}}$ and $dom\, \partial_1 d(\cdot, y) = \mathtt{C}$, where $\partial_1 d(\cdot, y)$ denotes the classical subgradient map of the function $d(\cdot, y)$ with respect to the first variable;
(iii) $d(\cdot, y)$ is coercive on $I\!R^n$ (i.e., $\lim_{||u|| \to \infty} d(u, y) = +\infty$).
(iv) $d(y, y) = 0$.

We denote by $\mathcal{D}(\mathtt{C})$ the family of functions satisfying this definition.

Property (i) is needed to preserve convexity of $d(\cdot, y)$, property (ii) will force the iteration of the proximal method to stay in $\mathtt{C}$, and property (iii) will be used to guarantee the existence of the proximal iterations. For each $y \in \mathtt{C}$, $\nabla_1 d(\cdot, y)$ denote the gradient map of the function $d(\cdot, y)$ with respect to the first variable. Note that by definition $d(\cdot, \cdot) \geq 0$ and from (iv) the global minimum of $d(\cdot, y)$ is obtained at y, which shows that $\nabla_1 d(y, y) = 0$.

Definition 4 Given $d \in \mathcal{D}(\mathtt{C})$, a function $H : I\!R^n \times I\!R^n \to I\!R_+ \cup \{+\infty\}$ is called the induced proximal distance to d if there exists $\gamma \in (0, 1]$ with H a finite valued on $\mathtt{C} \times \mathtt{C}$ and such that for each $a, b \in \mathtt{C}$, we have

(Ii) $H(a, a) = 0$.
(Iii) $\langle c - b, \nabla_1 d(b, a) \rangle \leq H(c, a) - H(c, b) - \gamma H(b, a), \quad \forall\, c \in \mathtt{C}$.

We write $(d, H) \in \mathcal{F}(\mathtt{C})$ to the proximal and induced proximal distance that satisfies the premises of Definition 4.
We also denote $(d, H) \in \mathcal{F}(\bar{\mathtt{C}})$ if there exists H such that:

(Iiii) H is finite valued on $\bar{\mathtt{C}} \times \mathtt{C}$ satisfying (Ii) and (Iii), for each $c \in \bar{\mathtt{C}}$.
(Iiv) For each $c \in \bar{\mathtt{C}}$, $H(c, \cdot)$ has level bounded sets on $\mathtt{C}$.

Finally, we write $(d, H) \in \mathcal{F}_+(\bar{\mathtt{C}})$ if

(Iv) $(d, H) \in \mathcal{F}(\bar{\mathtt{C}})$.
(Ivi) $\forall\, y \in \bar{\mathtt{C}}$ and $\forall\, \{y^k\} \subset \mathtt{C}$ bounded with $\lim_{k \to +\infty} H(y, y^k) = 0$, we have $\lim_{k \to +\infty} y^k = y$.
(Ivii) $\forall\, y \in \bar{\mathtt{C}}$, and $\forall\, \{y^k\} \subset \mathtt{C}$ such that $\lim\limits_{k \to +\infty} y^k = y$, we obtain $\lim_{k \to +\infty} H(y, y^k) = 0$.

The following additional conditions on H will be useful to prove the convergence of our algorithm.

Given $(d, H) \in \mathcal{F}_+(\bar{\mathtt{C}})$, H satisfies the following condition:

(Iviii) For all $c \in \bar{\mathtt{C}}$, and for all $\{y^k\} \subset \mathtt{C}$ such that $\lim\limits_{k\to+\infty} y^k = y$, then $\lim_{k\to+\infty} H(c, y^k) = H(c, y)$.

Some examples of proximal distances that satisfied (Ii)-(Iviii) are the following:

(a) **Bregman distances**. Let $S \subseteq I\!R^n$ be a nonempty open convex set, and let $\bar{S}$ be its closure. Let $h : \bar{S} \to I\!R \cup \{+\infty\}$ be a proper, lower semicontinuous and convex function with $dom\,\nabla h = S$, strictly convex and continuous on $dom\,h$, and continuously differentiable on S. Define

$$\begin{aligned} H(x,y) := D_h(x,y) &:= h(x) - [h(y) + \langle \nabla h(y), x - y\rangle] \;\; \forall\, x \in \bar{S}, \forall\, y \in S \\ &= +\infty \quad \text{otherwise.} \end{aligned}$$

The function D_h enjoys a remarkable three point identity (see [8], Lemma 3.1),

$$H(c,a) = H(c,b) + H(b,a) + \langle c - b, \nabla_1 H(b,a)\rangle \;\; \forall\, a, b \in S, \forall\, c \in dom\,h.$$

The function h is called a Bregman function with zone S, if it satisfies the following conditions:
(B_1) $dom\,h = \bar{S}$;
(B_2) (i) $\forall\, x \in \bar{S}$, $D_h(x, \cdot)$ is level bounded on $int(dom\,h)$;
(ii) $\forall\, y \in S$, $D_h(\cdot, y)$ is level bounded;
(B_3) $\forall\, y \in dom\,h$ and $\forall\,\{y^k\} \subset int(dom\,h)$ with $\lim_{k\to+\infty} y^k = y$, one has $\lim_{k\to+\infty} D_h(y, y^k) = 0$;
(B_4) If $\{y^k\}$ is a bounded sequence in $int(dom\,h)$ and $y \in dom\,h$ such that $\lim_{k\to+\infty} D_h(y, y^k) = 0$, then $\lim_{k\to+\infty} y^k = y$.

Note that (B_4) is a direct consequence of the first three properties, a fact proved by Kiwiel in ([15], Lemma 2.16).

Recall that a Bregman function h with zone S is said to be essentially smooth (see [22]) if it satisfies the following condition:

(B_5) If $\{y^k\} \subset S$ is convergent to a point on the boundary of S and $u \in S$, then $\lim_{k\to+\infty} D_h(u, y^k) = +\infty$.

Now, we show that this condition implies (Iviii) for $H = D_h$. Indeed, it is clear that, for $\mathtt{C} = S$ defining

$$d(x,y) := H(x,y) := D_h(x,y),$$

we obtain that $(d, H) \in \mathcal{F}_+(\bar{\mathtt{C}})$. Now, we will verify that when h is a essentially smooth Bregman function, H satisfies the condition (Iviii). Indeed, let $c \in \bar{\mathtt{C}}$ and a sequence $\{y^k\} \subset \mathtt{C}$ such that $\lim_{k\to+\infty} y^k = y$. Suppose that y is a point of the boundary of $\mathtt{C}$, then from condition (Ivii), we obtain

$$\lim_{k\to+\infty} D_h(y, y^k) = 0. \tag{5}$$

On the other hand, defining $u = y^k + (1-\theta)(y - y^k)$ with $0 < \theta < 1$, clearly,

$u \in \mathtt{C}$. Since ∇h is monotone,

$$\langle \nabla h(u), u - y^k \rangle \geq \langle \nabla h(y^k), u - y^k \rangle.$$

Taking into account that $u - y^k = (1-\theta)(y - y^k)$, the latter relation yields

$$\langle \nabla h(u), y - y^k \rangle \geq \langle \nabla h(y^k), y - y^k \rangle.$$

Then

$$\begin{aligned} D_h(y,u) + D_h(u,y^k) &= h(y) - h(u) - \langle \nabla h(u), y - u \rangle \\ &\quad + h(u) - h(y^k) - \langle \nabla h(y^k), u - y^k \rangle \\ &= h(y) - h(y^k) - \theta \langle \nabla h(u), y - y^k \rangle - (1-\theta)\langle \nabla h(y^k), y - y^k \rangle \\ &\leq h(y) - h(y^k) - \theta \langle \nabla h(y^k), y - y^k \rangle - (1-\theta)\langle \nabla h(y^k), y - y^k \rangle \\ &= h(y) - h(y^k) - \langle \nabla h(y^k), y - y^k \rangle \\ &= D_h(y, y^k). \end{aligned}$$

Thus,

$$D_h(y,u) + D_h(u,y^k) \leq D_h(y,y^k).$$

By taking $k \to +\infty$, and using the fact that h is essentially smooth, see condition B_5, and equality (5), we obtain $+\infty \leq 0$, which is a contradiction and therefore $y \in \mathtt{C}$.
Finally, by the continuity of h and ∇h, the condition (Iviii) is satisfied.

(b) **φ-divergence proximal distances**. Let $\varphi : I\!R \to I\!R \cup \{+\infty\}$ be an lower semicontinuous, convex, proper function such that $dom\, \varphi \subset I\!R_+$ and $dom \partial\varphi = I\!R_{++}$. We suppose in addition that φ is C^2, strictly convex, and nonnegative on $I\!R_{++}$ with $\varphi(1) = \varphi'(1) = 0$. We denote by Φ the class of such kernels and by Φ_1 the subclass of these kernels satisfying

$$\varphi''(1)\left(1 - \frac{1}{t}\right) \leq \varphi'(t) \leq \varphi''(1) log\, t \quad \forall\, t > 0.$$

The other subclass of Φ of interest is denoted by Φ_2, where the inequality above is replaced by

$$\varphi''(1)\left(1 - \frac{1}{t}\right) \leq \varphi'(t) \leq \varphi''(1)(t-1) \quad \forall\, t > 0.$$

Examples of functions in Φ_1, Φ_2 are (see, e.g., [2, 26])

$$\begin{aligned} \varphi_1(t) &= t\, log\, t - t + 1, \quad dom\, \varphi = [0, +\infty), \\ \varphi_2(t) &= -\, log\, t + t - 1, \quad dom\, \varphi = (0, +\infty), \\ \varphi_3(t) &= 2(\sqrt{t} - 1)^2, \quad dom\, \varphi = [0, +\infty), \end{aligned}$$

Corresponding to the classe Φ_1, we define a φ-divergence proximal distance by

$$d_\varphi(x,y) = \sum_{i=1}^n y_i \varphi\left(\frac{x_i}{y_i}\right).$$

For any $\varphi \in \Phi$, since argmin$\{\varphi(t)\ :\ t \in I\!R\} = \{1\}$, φ is coercive and thus it follows that $d_\varphi \in \mathcal{D}(\mathtt{C})$, with $\mathtt{C} = I\!R_{++}$.
Let $\varphi \in \Phi_1$ and define the regularized φ-divergence proximal distance (class Φ_1) by

$$\bar{d}_\varphi(x,y) := \sum_{i=1}^n y_i \varphi\left(\frac{x_i}{y_i}\right) + \frac{\sigma}{2}\|x-y\|^2.$$

where σ is a positive constant. Note also that $\bar{d}_\varphi \in \mathcal{D}(\mathtt{C})$. Taking $\varphi(t) = t - log\, t - 1$, we obtain

$$\bar{d}_\varphi(x,y) := \sum_{i=1}^n x_i - y_i - y_i log\frac{x_i}{y_i} + (\sigma/2)\|x-y\|^2.$$

Taking

$$\bar{H}(x,y) = \sum_{j=1}^n x_j log\,\frac{x_j}{y_j} + y_j - x_j + \frac{\sigma}{2}\|x-y\|^2$$

and one has

$$\langle c-b, \nabla_1 \bar{d}_\varphi(b,a)\rangle \le \bar{H}(c,a) - \bar{H}(c,b) - \gamma \bar{H}(b,a),$$

see Auslender and Teboulle [4] (Subsection 2.3, Case a2), and therefore $(\bar{d}_\varphi, \bar{H}) \in \mathcal{F}_+(I\!R^n_+)$. We note also that, definying

$$h(x) := \sum_{j=1}^n x_j log\, x_j + (\sigma/2)\|x\|^2$$

and adopting the convention $0\, log\, 0 = 0$, we have $D_h(x,y) = \bar{H}$ and thus $\bar{H}$ satisfies the condition (B_5). Therefore, as h is an essentially smooth Bregman function, we obtain that $\bar{H}$ satisfies the condition (Iviii).

(c) **Second order homogeneous proximal distances.** Let $p \in \Phi_2$, $\varphi(t) = \mu p(t) + \frac{\nu}{2}(t-1)^2$ with $\nu \ge \mu p''(1) > 0$, and let the associated proximal distance be defined by

$$d_\varphi(x,y) = \sum_{j=1}^n y_j^2 \varphi\left(\frac{x_j}{y_j}\right).$$

It is clear that, $d_\varphi \in \mathcal{D}(I\!R^n_{++})$ and from the key inequality (see formula (3.21) [2], Lemma 3.4) for any $p \in \Phi_2$ one has

$$\langle c-b, \nabla_1 d_\varphi(b,a)\rangle \le \bar{\eta}(\|c-a\|^2 - \|c-b\|^2 - \gamma\|b-a\|^2) \quad \forall\, a,b \in I\!R^n_{++},\ \forall\, c \in I\!R^n_+$$

where $\bar{\eta} = 2^{-1}(\nu + \mu p''(1))$. Therefore with

$$H(x,y) = \bar{\eta}\,\|x-y\|^2\,, \quad \gamma = \frac{\nu - \mu p''(1)}{\nu + \mu p''(1)}$$

it follows that $(d_\varphi, H) \in \mathcal{F}_+(I\!R^n_+)$, also clearly the condition (Iviii) is satisfied. The main result of the method will be when $(d,H) \in \mathcal{F}_+(\bar{\mathtt{C}})$ and the condition (Iviii) is satisfied.

4. An inexact proximal multiplier method with proximal distances

4.1. *The Problem*

We are interested in solving the following convex minimization problem with separable structure:

$$(CP) \quad \min\{f(x) + g(z) \,:\, Ax + Bz = b,\, x \in \bar{\mathtt{C}},\, z \in \bar{\mathtt{K}}\},$$

where $\mathtt{C} \subset I\!R^n$ and $\mathtt{K} \subset I\!R^p$ are nonempty open convex sets, $\bar{\mathtt{C}}$ and $\bar{\mathtt{K}}$ denote the closure (in Euclidean topology) of $\mathtt{C}$ and $\mathtt{K}$ respectively, $f : I\!R^n \to (-\infty, +\infty]$ and $g : I\!R^p \to (-\infty, +\infty]$ are closed proper convex functions and $A \in I\!R^{m\times n}$, $B \in I\!R^{m\times p}$, $b \in I\!R^m$. Note that many convex programs can be formulated in the generic form (CP), see e.g. [13, 22].

The Lagrangian for (CP) is defined by $L : I\!R^n \times I\!R^p \times I\!R^m \to (-\infty, +\infty]$,

$$L(x,z,y) = (f + \delta_{\bar{\mathtt{C}}})(x) + (g + \delta_{\bar{\mathtt{K}}})(z) + \langle y, Ax + Bz - b\rangle\,.$$

where $\delta_{\mathtt{X}}$ denotes the indicator function of a subset X, $\langle\cdot,\cdot\rangle$ denotes the canonical inner product and y is the Lagrangian multiplier associated with the constraint $Ax + Bz = b$.

Let $(x^*, z^*) \in I\!R^n \times I\!R^p$ and $y^* \in I\!R^m$. In order that (x^*, z^*) be an optimal solution of (CP) and y^* be an optimal Lagrangian multiplier, it is necessary and sufficient that (x^*, z^*, y^*) be a saddle point of the Lagrangian L of (CP) (see [22], Theorem 28.3); i.e., for all $x \in dom f \cap \bar{\mathtt{C}}$, $z \in dom\, g \cap \bar{\mathtt{K}}$ and $y \in I\!R^m$, we have

$$L(x^*, z^*, y) \le L(x^*, z^*, y^*) \le L(x, z, y^*). \tag{6}$$

4.2. *The (PMAPD) Algorithm*

In the proposed algorithm we use the class of proximal distances $(d_0, H_0) \in \mathcal{F}_+(\bar{\mathtt{C}})$, $(d_0', H_0') \in \mathcal{F}_+(\bar{\mathtt{K}})$, satisfying the condition (Iviii) and given $\mu > 0, \mu' > 0$ we define the following functions:

$$d(x,y) = d_0(x,y) + (\mu/2)\,\|x-y\|^2\,, \tag{7}$$

$$H(x,y) = H_0(x,y) + (\mu/2)\,\|x-y\|^2\,, \tag{8}$$

$$d'(x,y) = d_0'(x,y) + (\mu'/2)\,\|x-y\|^2\,, \tag{9}$$

$$H'(x,y) = H_0'(x,y) + (\mu'/2)\,\|x-y\|^2\,. \tag{10}$$

It is easy to check that $(d,H) \in \mathcal{F}_+(\bar{\mathtt{C}})$ and $(d',H') \in \mathcal{F}_+(\bar{\mathtt{K}})$ (for the same value of γ and γ' respectively) and both satisfy the condition (Iviii).

The algorithm, which will be called Proximal Multiplier Algorithm with Proximal Distances (PMAPD) is as follows:

(PMAPD) Algorithm

Step 0. Choose two pairs $(d_0,H_0) \in \mathcal{F}_+(\bar{\mathtt{C}})$, $(d'_0,H'_0) \in \mathcal{F}_+(\bar{\mathtt{K}})$ satisfying the condition (Iviii) and define (d,H), (d',H') given by (7)-(8) and (9)-(10) respectively. Take three sequences $a_k \geq 0$, $b_k \geq 0$ and $\lambda_k > 0$ and choose an arbitrary starting point $(x^0, z^0, y^0) \in \mathtt{C} \times \mathtt{K} \times I\!R^m$.

Step 1. For $k = 0, 1, 2, \ldots$, calculate $p^{k+1} \in I\!R^m$ by

$$p^{k+1} = y^k + \lambda_k(Ax^k + Bz^k - b). \tag{11}$$

Step 2. Find $(x^{k+1}, \upsilon^{k+1}) \in \mathtt{C} \times I\!R^n$ and $(z^{k+1}, \xi^{k+1}) \in \mathtt{K} \times I\!R^p$ such that

$$\upsilon^{k+1} \in \partial_{a_k} f^k(x^{k+1}), \qquad \upsilon^{k+1} + \lambda_k^{-1}\nabla_1 d(x^{k+1}, x^k) = 0, \tag{12}$$

$$\xi^{k+1} \in \partial_{b_k} g^k(z^{k+1}), \qquad \xi^{k+1} + \lambda_k^{-1}\nabla_1 d'(z^{k+1}, z^k) = 0. \tag{13}$$

where the functions $f^k : I\!R^n \to (-\infty, +\infty]$ and $g^k : I\!R^p \to (-\infty, +\infty]$ are defined by $f^k(x) = f(x) + \langle p^{k+1}, Ax\rangle$ and $g^k(z) = g(z) + \langle p^{k+1}, Bz\rangle$, respectively.

Step 3. Compute

$$y^{k+1} = y^k + \lambda_k(Ax^{k+1} + Bz^{k+1} - b). \tag{14}$$

Stop criterium: If $x^{k+1} = x^k$, $z^{k+1} = z^k$ and $y^{k+1} = y^k$ then stop. Otherwise to do $k := k + 1$, and go to Step 1.

Remark 1 If the stopping criterion is satisfied, then the (PMAPD) find an $(a_k + b_k)-$solution of (CP). Of fact, if $x^{k+1} = x^k$, $z^{k+1} = z^k$ and $y^{k+1} = y^k$, from (12) and (13) we have

$$0 \in \partial_{a_k} f^k(x^k) \;\text{ and }\; 0 \in \partial_{b_k} g^k(z^k) \tag{15}$$

and from (11) and (14), we have

$$Ax^k + Bz^k = b \quad\text{and}\quad p^{k+1} = y^k, \tag{16}$$

from (15), we obtain

$$\begin{aligned} f^k(x) &\geq f^k(x^k) - a_k \quad \forall\, x \in dom f, \\ g^k(z) &\geq g^k(z^k) - b_k, \quad \forall\, z \in dom\, g, \end{aligned}$$

adding and rearranging terms, using (16), we get

$$f(x^k) + g(z^k) - (a_k + b_k) \leq f(x) + g(z) + \langle y^k, Ax + Bz - b\rangle,$$

then for all $x \in dom f \cap \bar{\mathtt{C}}$ and $z \in dom\, g \cap \bar{\mathtt{K}}$ such that $Ax + Bz = b$, we obtain

$$f(x^k) + g(z^k) - (a_k + b_k) \leq f(x) + g(z).$$

Therefore

$$f(x^k) + g(z^k) \leq \inf\{f(x) + g(z) \ : Ax + Bz = b, x \in \bar{\mathtt{C}}, z \in \bar{\mathtt{K}}\} + (a_k + b_k).$$

Thus, we obtain the result.

Remark 2 As we are interested in the asymptotic convergence of the method, we assume that $(x^{k+1}, z^{k+1}, y^{k+1}) \neq (x^k, z^k, y^k)$, for all k.

4.3. *Global convergence*

In this subsection, under appropriate assumptions, we prove that the iterations given by the algorithm are well defined and we establish the global convergence of the (PMAPD).

Throughout the paper we assume the following assumptions:

(H1) The problem (CP) has an optimal solution (x^*, z^*) and a corresponding Lagrange multiplier y^*.
(H2) There exist $x \in ri(dom\, d(\cdot, v)) \cap ri(dom\, f)$ and $z \in ri(dom\ d'(\cdot, v')) \cap ri(dom\, g)$ such that $Ax + Bz = b$.
(H3) The sequences $\{a_k\}$ and $\{b_k\}$ are nonnegative and

$$\sum_{k=0}^{\infty}(a_k + b_k) < \infty.$$

(H4) Given the parameters $\mu > 0$, $\mu' > 0$, defined in (7) and (9) respectively, the sequence $\{\lambda_k\}$ satisfies

$$\eta < \lambda_k < \bar{c} - \eta \tag{17}$$

where $\eta \in (0, \bar{c}/2)$ with $\bar{c} := \min\{\frac{\sqrt{\gamma\mu}}{2\|A\|}, \frac{\sqrt{\gamma'\mu'}}{2\|B\|}\}$ and γ, γ' are positive constants related to d and d', respectively, in Definition 4, (Iii).

Remark 3 Assumption (H1) implies that the Lagrangian function L has a saddle point (x^*, z^*, y^*). Assumption (H2) guaranteed the existence of an optimal dual Lagrange multiplier y^* associated with the constraint $Ax + Bz = b$, and associate with Assumption (H1) ensure that the iterations given by the (PMAPD) are well defined. The above assumptions are not very strong and these has been used in previous researches, see [3, 9, 16]. Assumption (H3) is a natural condition when we work with inexact iterations. Assumption (H4) will be used to ensures the convergence of the method. Observe that the interval $(\eta, \bar{c} - \eta)$ depends of μ and μ' which are arbitrary.

THEOREM 4.1 *Let $d_0 \in \mathcal{D}(C)$ and $d'_0 \in \mathcal{D}(K)$. Suppose that assumptions (H1) and (H2) are satisfied, then for any $(x^k, z^k, y^k) \in C \times K \times I\!R^m$, $\lambda_k > 0$, exists $(x^{k+1}, z^{k+1}) \in C \times K$ satisfying (12) and (13).*

Proof. Let $F_k(x) = f(x) + \langle p^{k+1}, Ax\rangle + (1/\lambda_k)d(x, x^k)$, then from Assumption (H2) we have that $dom\, F_k = dom\, f \cap dom\, d(\cdot, x^k) \neq \emptyset$. Define $S_k = \arg\min_{x\in\bar{\mathtt{C}}}\{F_k(x)\} = \arg\min_{x\in I\!R^n}\{F_k(x) + \delta_{\bar{\mathtt{C}}}(x)\}$, where $\delta_{\bar{\mathtt{C}}}$ denote the indicator function, we will prove that S_k is nonempty. Indeed, because $f(\cdot)$, $\langle A^T p^{k+1}, \cdot\rangle$ and $(1/\lambda_k)d(\cdot, x^k)$ are closed proper convex functions, then F_k is closed proper and convex too and from propositions 2.1 and 2.2, it is suffices to show that $((F_k + \delta_{\bar{\mathtt{C}}})0^+)(y) > 0$, $\forall y \neq 0$.
Let $y \neq 0$, then by Proposition 2.3, we have that

$$((F_k+\delta_{\bar{\mathtt{C}}})0^+)(y) = (f0^+)(y)+(\langle A^T p^{k+1}, \cdot\rangle 0^+)(y)+(1/\lambda_k)(d(\cdot, x^k)0^+)(y)+(\delta_{\bar{\mathtt{C}}}0^+)(y). \tag{18}$$

It is easy to prove that

$$(f0^+)(y) > -\infty,\ \ (\delta_{\bar{\mathtt{C}}}0^+)(y) > -\infty,\ \ (\langle A^T p^{k+1}, \cdot\rangle 0^+)(y) = \langle A^T p^{k+1}, y\rangle.$$

so we prove that $(d(\cdot, x^k)0^+)(y) = +\infty$.
From Proposition 2.4, for $\bar{x} \in dom\, d_0(\cdot, x^k)$, we have that

$$(d(\cdot, x^k)0^+)(y) = \lim_{\lambda\to+\infty}(\frac{d_0(\bar{x} + \lambda y, x^k) - d_0(\bar{x}, x^k)}{\lambda} + \frac{\mu\lambda}{2}\|y\|^2) - \mu\left\langle \bar{x} - x^k, y\right\rangle,$$

and since $d_0(\cdot, x^k)$ convex function, we obtain that $(d(\cdot, x^k)0^+)(y) = +\infty$ and therefore from (18)

$$((F_k + \delta_{\bar{\mathtt{C}}})0^+)(y) = +\infty > 0,\ \ \forall\, y \neq 0.$$

Thus, there exists $\bar{x} := x^{k+1} \in I\!R^n$ such that

$$0 \in \partial\left(f(\cdot) + \langle A^T p^{k+1}, \cdot\rangle + (1/\lambda_k)d(\cdot, x^k) + \delta_{\bar{\mathtt{C}}}(\cdot)\right)(x^{k+1}).$$

Now by Assumption (H2) and Definition of $d(\cdot, x^k)$, there exists $x \in ri(dom\, f) \cap ri(dom\, d(\cdot, x^k)) \cap \mathtt{C}$, then from Proposition 2.5, we have that

$$0 \in \partial f(x^{k+1}) + A^T p^{k+1} + (1/\lambda_k)\partial_1 d(x^{k+1}, x^k) + \partial\delta_{\bar{\mathtt{C}}}(x^{k+1}).$$

From Definition 3, (ii), we have that $dom\,\partial_1 d(\cdot, x^k) = \mathtt{C}$, then $x^{k+1} \in \mathtt{C}$.
As $\partial f(x) \subseteq \partial_\epsilon f(x)$ for all $\epsilon \geq 0$, then there exists $x^{k+1} \in \mathtt{C}$ and $v^{k+1} \in I\!R^n$ satisfying (12).
For the existence of $z^{k+1} \in \mathtt{K}$ and $\xi^{k+1} \in I\!R^p$ satisfying (13) the analysis is similar.
Therefore the proof is completed. □

The subsequent convergence analysis follows a line of argument similar to that given in [9].

Lemma 4.2 ([*9*], *Lemma* 3.1) *Let $F : I\!R^m \to (-\infty, +\infty]$ a closed proper convex function, $\tau > 0$ and define:*

$$u^{k+1} = \arg\min_{u\in I\!R^m}\{F(u) + (1/(2\tau))\|u - u^k\|^2\}.$$

Then for any integer $k \geq 0$,

$$2\tau[F(u^{k+1}) - F(u)] \leq \|u^k - u\|^2 - \|u^{k+1} - u\|^2 - \|u^{k+1} - u^k\|^2,\ \forall\, u \in I\!R^m.$$

We will also use the following useful notation

$$\Delta_k(x,z) = H(x,x^k) - H(x,x^{k+1}) - \gamma H(x^{k+1},x^k) + H'(z,z^k) - H'(z,z^{k+1}) - \gamma' H'(z^{k+1},z^k) \tag{19}$$

LEMMA 4.3 *Let $(d_0, H_0) \in \mathcal{F}(\bar{C})$, $(d'_0, H'_0) \in \mathcal{F}(\bar{K})$. Suppose that the assumptions (H1) and (H2) are satisfied. Then, for all $x \in dom f \cap \bar{C}$, $z \in dom\, g \cap \bar{K}$, $y \in I\!R^m$ the following inequalities are satisfied:*

(*i*) $\lambda_k\{L(x^{k+1},z^{k+1},p^{k+1}) - L(x,z,p^{k+1})\} \le \Delta_k(x,z) + \lambda_k(a_k+b_k)$
(*ii*) $2\lambda_k\{L(x^k,z^k,y) - L(x^k,z^k,p^{k+1})\} \le \|y^k-y\|^2 - \|p^{k+1}-y\|^2 - \|p^{k+1}-y^k\|^2$
(*iii*) $2\lambda_k\{L(x^{k+1},z^{k+1},y) - L(x^{k+1},z^{k+1},y^{k+1})\} \le \|y^k-y\|^2 - \|y^{k+1}-y\|^2 - \|y^{k+1}-y^k\|^2.$

Proof. (*i*) Let $x \in dom f \cap \bar{C}$, $z \in dom\, g \cap \bar{K}$, $y \in I\!R^m$ from definition of L, it is clear that

$$L(x^{k+1},z^{k+1},p^{k+1}) - L(x,z,p^{k+1}) = (f^k(x^{k+1}) - f^k(x)) + (g^k(z^{k+1}) - g^k(z)). \tag{20}$$

Since $\upsilon^{k+1} \in \partial_{a_k} f^k(x^{k+1})$, then $f^k(x) \ge f^k(x^{k+1}) + \langle \upsilon^{k+1}, x - x^{k+1}\rangle - a_k$, then from (12) and from Definition 4, (Iii), we obtain

$$\begin{aligned} \lambda_k(f^k(x^{k+1}) - f^k(x)) &\le \langle \nabla_1 d(x^{k+1},x^k), x - x^{k+1}\rangle + \lambda_k a_k \\ &\le H(x,x^k) - H(x,x^{k+1}) - \gamma H(x^{k+1},x^k) + \lambda_k a_k. \end{aligned} \tag{21}$$

Analogously, since $\xi^{k+1} \in \partial_{b_k} g^k(z^{k+1})$, we have $g^k(z) \ge g^k(z^{k+1}) + \langle \xi^{k+1}, z - z^{k+1}\rangle - b_k$, then from (13) and from Definition 4, (Iii), we obtain

$$\begin{aligned} \lambda_k(g^k(z^{k+1}) - g^k(z)) &\le \nabla_1 d'(z^{k+1},z^k), z - z^{k+1}\rangle + \lambda_k b_k \\ &\le H'(z,z^k) - H'(z,z^{k+1}) - \gamma' H'(z^{k+1},z^k) + \lambda_k b_k, \end{aligned} \tag{22}$$

by adding (21) and (22), and from (20) and the notation in (19), we obtain

$$\lambda_k\{L(x^{k+1},z^{k+1},p^{k+1}) - L(x,z,p^{k+1})\} \le \Delta_k(x,z) + \lambda_k(a_k+b_k).$$

To prove (*ii*) and (*iii*) note that Steps 1 and 3 of (PMAPD) can be written equivalently as:

$$p^{k+1} = \arg\min_{y \in I\!R^m} \{-L(x^k,z^k,y) + (1/2\lambda_k)\|y-y^k\|^2\}$$

$$y^{k+1} = \arg\min_{y \in I\!R^m} \{-L(x^{k+1},z^{k+1},y) + (1/2\lambda_k)\|y-y^k\|^2\}$$

Then, using Lemma 4.2, twice with $\tau = \lambda_k$, $F(y) = -L(x^k,z^k,y)$ and $F(y) = -L(x^{k+1},z^{k+1},y)$ respectively, the first and second equations above respectively yield the desired inequalities (*ii*) and (*iii*). □

Remark 4 Throughout the rest of this paper, we denote $w^k = (x^k,z^k,y^k)$, $w^* = (x^*,z^*,y^*)$, $w = (x,z,y)$, $s = (l,q,r) \in I\!R^n \times I\!R^p \times I\!R^m$ and define the function $\hat{H} : X \times X \to I\!R_+ \cup \{+\infty\}$ by

$$\hat{H}(w,s) = \hat{H}((x,z,y),(l,q,r)) = H(x,l) + H'(z,q) + (1/2)\|y-r\|^2, \tag{23}$$

where $X = I\!R^n \times I\!R^p \times I\!R^m$, H and H' are defined in (8) and (10) respectively.

Proposition 4.4 *Let* $(d_0, H_0) \in \mathcal{F}(\bar{C})$, $(d_0', H_0') \in \mathcal{F}(\bar{K})$. *Suppose that Assumptions* $(H1), (H2)$ *and* $(H4)$ *are satisfied. Let* $\{w^k\}$ *and* $\{p^k\}$ *be sequences generated by the* $(PMAPD)$; *let* (x^*, z^*) *be an optimal solution of* (CP), *and let* y^* *be a corresponding Lagrange multiplier, then we have, for each* k,

$$\begin{aligned}\hat{H}(w^*, w^{k+1}) \le \hat{H}(w^*, w^k) &- [\gamma H_0(x^{k+1}, x^k) + (1/2)(\gamma\mu - 4\lambda_k^2\|A\|^2)\|x^{k+1} - x^k\|^2] + \\ &-(\gamma' H_0'(z^{k+1}, z^k) + (1/2)(\gamma'\mu' - 4\lambda_k^2\|B\|^2)\|z^{k+1} - z^k\|^2) + \\ &-(1/2)(\|p^{k+1} - y^{k+1}\|^2 + \|p^{k+1} - y^k\|^2) + \lambda_k(a_k + b_k),\end{aligned} \tag{24}$$

where $w^* = (x^*, z^*, y^*)$. *In particular, we have, for all* k,

$$\hat{H}(w^*, w^{k+1}) \le \hat{H}(w^*, w^k) + \lambda_k(a_k + b_k). \tag{25}$$

Proof. Taking $(x, z) = (x^*, z^*)$ in Lemma 4.3, (i), we obtain

$$\begin{aligned}\lambda_k(L(x^{k+1}, z^{k+1}, p^{k+1}) - L(x^*, z^*, p^{k+1})) \le\ & H(x^*, x^k) - H(x^*, x^{k+1}) + H'(z^*, z^k) + \\ & -H'(z^*, z^{k+1}) - (\gamma H(x^{k+1}, x^k) + \\ & +\gamma' H'(z^{k+1}, z^k)) + \lambda_k(a_k + b_k).\end{aligned} \tag{26}$$

Since (x^*, z^*, y^*) is a saddle point of L, we obtain from (6)

$$\lambda_k(L(x^*, z^*, p^{k+1}) - L(x^{k+1}, z^{k+1}, y^*)) \le 0. \tag{27}$$

By adding (26) and (27) and rearranging terms, we get

$$\begin{aligned}H(x^*, x^{k+1}) + H'(z^*, z^{k+1}) \le\ & H(x^*, x^k) + H'(z^*, z^k) - \left[\gamma H(x^{k+1}, x^k) + \gamma' H'(z^{k+1}, z^k)\right] + \\ & -\lambda_k \left\langle p^{k+1} - y^*, Ax^{k+1} + Bz^{k+1} - b\right\rangle + \lambda_k(a_k + b_k).\end{aligned} \tag{28}$$

Now, using Lemma 4.3, (ii), with $y := y^{k+1}$ and $y := y^*$ for (iii), we obtain, respectively,

$$2\lambda_k(L(x^k, z^k, y^{k+1}) - L(x^k, z^k, p^{k+1})) \le \|y^k - y^{k+1}\|^2 - \|p^{k+1} - y^{k+1}\|^2 - \|p^{k+1} - y^k\|^2 \tag{29}$$

$$2\lambda_k(L(x^{k+1}, z^{k+1}, y^*) - L(x^{k+1}, z^{k+1}, y^{k+1})) \le \|y^k - y^*\|^2 - \|y^{k+1} - y^*\|^2 - \|y^{k+1} - y^k\|^2 \tag{30}$$

adding (29) and (30) and rearranging terms, we get

$$\begin{aligned}(1/2)\left\|y^{k+1} - y^*\right\|^2 \le\ & (1/2)\left\|y^k - y^*\right\|^2 - (1/2)\left[\left\|p^{k+1} - y^{k+1}\right\|^2 + \left\|p^{k+1} - y^k\right\|^2\right] + \\ & -\lambda_k(\left\langle y^* - y^{k+1}, Ax^{k+1} + Bz^{k+1} - b\right\rangle + \\ & + \left\langle y^{k+1} - p^{k+1}, Ax^k + Bz^k - b\right\rangle).\end{aligned} \tag{31}$$

Then, by adding (28) and (31) and from (23), we obtain

$$\begin{aligned}\hat{H}(w^*, w^{k+1}) \le\ & \hat{H}(w^*, w^k) - (\gamma H(x^{k+1}, x^k) + \gamma' H'(z^{k+1}, z^k)) + \\ & - (1/2)\left(\|p^{k+1} - y^{k+1}\|^2 + \|p^{k+1} - y^k\|^2\right) + \lambda_k(a_k + b_k) + \\ & + \lambda_k \left\langle y^{k+1} - p^{k+1}, A(x^{k+1} - x^k) + B(z^{k+1} - z^k)\right\rangle \\ =\ & \hat{H}(w^*, w^k) - (\gamma H(x^{k+1}, x^k) + \gamma' H'(z^{k+1}, z^k)) + \\ & - (1/2)\left(\|p^{k+1} - y^{k+1}\|^2 + \|p^{k+1} - y^k\|^2\right) + \\ & + \lambda_k^2 \left\|A(x^{k+1} - x^k) + B(z^{k+1} - z^k)\right\|^2 + \lambda_k(a_k + b_k),\end{aligned} \tag{32}$$

where the last equality follows from (11) and (14).
Using, in (32), the inequality

$$\left\|A(x^{k+1}-x^k)+B(z^{k+1}-z^k)\right\|^2 \le 2\left(\|A\|^2\,\|x^{k+1}-x^k\|^2+\|B\|^2\,\|z^{k+1}-z^k\|^2\right),$$

and also,

$$\begin{aligned} H(x^{k+1},x^k) &= H_0(x^{k+1},x^k)+(\mu/2)\|x^{k+1}-x^k\|^2,\\ H'(z^{k+1},z^k) &= H_0'(z^{k+1},z^k)+(\mu'/2)\|z^{k+1}-z^k\|^2, \end{aligned}$$

we obtain

$$\begin{aligned} \hat{H}(w^*,w^{k+1}) \le \hat{H}(w^*,w^k) &- [\gamma H_0(x^{k+1},x^k)+(1/2)(\gamma\mu-4\lambda_k^2\,\|A\|^2)\|x^{k+1}-x^k\|^2]+\\ &-(\gamma' H_0'(z^{k+1},z^k)+(1/2)(\gamma'\mu'-4\lambda_k^2\,\|B\|^2)\|z^{k+1}-z^k\|^2)+\\ &-(1/2)\left(\|p^{k+1}-y^{k+1}\|^2+\|p^{k+1}-y^k\|^2\right)+\lambda_k(a_k+b_k). \end{aligned}$$

Then, the inequality (25) follows immediately from (24) and Assumption (H4). □

LEMMA 4.5 ([*21*], *Lemma* 2) *Let* $\{\alpha_k\}$ *and* $\{\beta_k\}$ *be nonnegative sequences such that* $\alpha_{k+1}\le\alpha_k+\beta_k$, *for all* k *and* $\sum_{k=0}^{\infty}\beta_k<+\infty$. *Then* $\{\alpha_k\}$ *is convergent.*

PROPOSITION 4.6 *Let* $(d_0,H_0)\in\mathcal{F}_+(\bar{C})$, $(d_0',H_0')\in\mathcal{F}_+(\bar{K})$ *satisfying the condition (Iviii). Suppose that assumptions (H1),(H2), (H3) and (H4) are satisfied. Let* $\{w^k\}$ *be a sequence generated by the* $(PMAPD)$ *algorithm, then* $\{w^k\}$ *is bounded and every accumulation point of* $\{w^k\}$ *is a saddle point of the Lagrangian L.*

Proof. First, note that Assumption (H3) together with Assumption (H4) ensures that

$$\sum_{k=0}^{\infty}\lambda_k(a_k+b_k)<+\infty. \tag{33}$$

Then, from (25), we have

$$w^k\in L_{\hat{H}}(w^*,\bar{\alpha}):=\{w\,:\,\hat{H}(w^*,w)\le\bar{\alpha}\},\ \text{ for all } k,$$

where $\bar{\alpha}=\hat{H}(w^*,w^0)+\sum_{k=0}^{\infty}\lambda_k(a_k+b_k)$.
This implies that $\{w^k\}$ is bounded because of condition (Iiv) in Definition 4 for H and H'.
Now, let (x^*,z^*) be an optimal solution of (CP), let y^* be a corresponding Lagrange multiplier, and let $w^*=(x^*,z^*,y^*)$.
Then, from (25), (33) and Lemma 4.5, we obtain that $\{\hat{H}(w^*,w^k)\}$ is convergent, i.e., there exists $\beta\ge 0$ such that

$$\lim_{k\to+\infty}\hat{H}(w^*,w^k)=\beta, \tag{34}$$

Taking limit in the inequality of Lemma (24), from Assumption (H3) and rearranging terms, we obtain

$$\limsup_{k\to+\infty}(\gamma H_0(x^{k+1},x^k)+\frac{1}{2}(\gamma\mu-4\lambda_k^2\,\|A\|^2)\|x^{k+1}-x^k\|^2+\frac{1}{2}(\gamma'\mu'-4\lambda_k^2\,\|B\|^2)\|z^{k+1}-z^k\|^2+ \\ +\gamma' H_0'(z^{k+1},z^k)+\tfrac{1}{2}(\|p^{k+1}-y^{k+1}\|^2+\|p^{k+1}-y^k\|^2))\leq 0. \tag{35}$$

It then follows from (34) and assumption (H4) that

$$\begin{aligned} H_0(x^{k+1},x^k)\to 0,\ \|x^{k+1}-x^k\|\to 0 \\ H_0'(z^{k+1},z^k)\to 0,\ \|z^{k+1}-z^k\|\to 0 \\ \|p^{k+1}-y^{k+1}\|\to 0,\ \|p^{k+1}-y^k\|\to 0. \end{aligned} \tag{36}$$

Now, let $\bar{w}=(\bar{x},\bar{z},\bar{y})$ be a cluster point of $\{w^k\}$, and let $\{w^k\}_{k\in K}$ a subsequence converging to $\bar{w}$. Using Lemma 4.3, (i), for all $x\in dom f\cap\bar{\mathtt{C}}$ and $z\in dom\, g\cap\bar{\mathtt{K}}$, we have

$$\lambda_k\{L(x^{k+1},z^{k+1},p^{k+1})-L(x,z,p^{k+1})\}\leq\Delta_k(x,z)+\lambda_k(a_k+b_k). \tag{37}$$

Taking the limit over the appropriate subsequences on both sides of (46), using (36) and the Assumptions (H3), (H4), and condition (Iviii) for H and H', we obtain

$$L(\bar{x},\bar{z},\bar{y})-L(x,z,\bar{y})\leq 0, \tag{38}$$

since λ_k is bounded below by $\eta>0$.

Similarly, by taking the limit, over the appropriate subsequences, on both sides of the inequality in Lemma 4.3, (ii), we have that, for all $y\in I\!R^m$,

$$L(\bar{x},\bar{z},y)-L(\bar{x},\bar{z},\bar{y})\leq 0. \tag{39}$$

From (38) and (39), we obtain that, for all $x\in dom f\cap\bar{\mathtt{C}}$, $z\in dom\, g\cap\bar{\mathtt{K}}$, $y\in I\!R^m$,

$$L(\bar{x},\bar{z},y)\leq L(\bar{x},\bar{z},\bar{y})\leq L(x,z,\bar{y}).$$

Therefore, $\bar{w}$ is a saddle point of L. □

From the above propositions, we obtain the following convergence theorem for the (PMAPD).

Theorem 4.7 *Let $(d_0,H_0)\in\mathcal{F}_+(\bar{C})$, $(d_0',H_0')\in\mathcal{F}_+(\bar{K})$ satisfying the condition (Iviii). Suppose that the Assumptions $(H1),(H2),(H3)$ and $(H4)$ are satisfied and let $\{(x^k,z^k,y^k)\}$ be the sequence generated by the (PMAPD), then the sequence $\{(x^k,z^k,y^k)\}$ globally convergences to (x^*,z^*,y^*), with (x^*,z^*) optimal for (CP) and y^* be a corresponding Lagrange multiplier.*

Proof. From Proposition 4.6, $\{w^k\}=\{(x^k,z^k,y^k)\}$ is bounded and any cluster point $\bar{w}$ is a saddle point of L.
Let $\{w^k\}_{k\in K}$ be a subsequence which converges to $\bar{w}$, then $(\bar{x},\bar{z})$ is an optimal solution of (CP) and $\bar{y}$ be a corresponding Lagrange multiplier (see [22], Theorem 28.3). Then, it is suffices to prove that $\{w^k\}$ converges to $\bar{w}$.
Indeed, since

$$\lim_{k\to+\infty,\,k\in K}x^k=\bar{x},\qquad\lim_{k\to+\infty,\,k\in K}z^k=\bar{z},\qquad\lim_{k\to+\infty,\,k\in K}y^k=\bar{y},$$

from Definition 4, (Ivii), we have that

$$\lim_{k\to+\infty,\, k\in K} H(\bar{x}, x^k) = 0, \quad \lim_{k\to+\infty,\, k\in K} H'(\bar{z}, z^k) = 0, \quad \lim_{k\to+\infty,\, k\in K} (1/2)\|y^k - \bar{y}\|^2 = 0$$

then, taking $w = \bar{w}$ and $s = w^k$ in (23), we have

$$\lim_{k\to+\infty,\, k\in K} \hat{H}(\bar{w}, w^k) = 0, \tag{40}$$

From (25) substituting w^* by $\bar{w}$ we obtain

$$0 \leq \hat{H}(\bar{w}, w^{k+1}) \leq \hat{H}(\bar{w}, w^k) + \lambda_k(a_k + b_k), \text{ for all } k,$$

Then, from Assumptions (H3), (H4) and Lemma 4.5 we have that $\hat{H}(\bar{w}, w^k)$ converges and as there exists a subsequence which converges to zero, (see (40)), then all the sequence converges, that is,

$$\lim_{k\to+\infty} \hat{H}(\bar{w}, w^k) = 0.$$

Finally, from (23) and Definition 4, (Ivi), for H and H', we obtain that

$$\lim_{k\to+\infty} w^k = \bar{w}.$$

Thus, we obtain the result. □

4.4. *Convergence for a general case*

In the previous section we analyzed the convergence of the algorithm (PMAPD) when $(d_0, H_0) \in \mathcal{F}_+(\bar{\mathtt{C}})$, $(d_0', H_0') \in \mathcal{F}_+(\bar{\mathtt{K}})$, in particular when the proximal distances satisfy the condition (Iii) of Definition 4, i.e.,

$$\langle c - b, \nabla_1 d(b, a)\rangle \leq H(c, a) - H(c, b) - \gamma H(b, a), \text{ with } \gamma \in (0, 1],$$

$$\langle c - b, \nabla_1 d'(b, a)\rangle \leq H'(c, a) - H'(c, b) - \gamma' H'(b, a), \text{ with } \gamma' \in (0, 1].$$

The above conditions are satisfied by the class of Bregman distances, second order homogeneous distances and for some φ-divergences distances. However, it is well known that all φ-divergences distance satisfy the above inequalities when $\gamma = 0$ and $\gamma' = 0$ that is,

$$(\text{Iii})' : \ \langle c - b, \nabla_1 d(b, a)\rangle \leq H(c, a) - H(c, b),$$

see Teboulle [26], Lemma 4.1, (ii).

In this subsection we analyze the convergence of the algorithm (PMAPD) substituting in Definition 4 the condition (Iii) by (Iii)$'$. We also replace the assumption

(H4) by the following:

$$(\mathrm{H4})' : \quad \begin{array}{c} \kappa < \lambda_k < \bar{\lambda} \\ \sum_{k=0}^{\infty} \|x^{k+1} - x^k\|^2 < +\infty \\ \sum_{k=0}^{\infty} \|z^{k+1} - z^k\|^2 < +\infty \end{array}$$

where κ and $\bar{\lambda}$ are positive constants.

Observe that, under the new assumptions, Theorem 4.1, Lemma 4.2, Lemma 4.3 (with $\Delta_k(x,z) = H(x,x^k) - H(x,x^{k+1}) + H'(z,z^k) - H'(z,z^{k+1})$) continue to be true but Proposition 4.4 becomes

PROPOSITION 4.8 *Let $(d_0, H_0) \in \mathcal{F}(\bar{C})$, $(d'_0, H'_0) \in \mathcal{F}(\bar{K})$. Suppose that Assumptions $(H1), (H2)$ are satisfied. Let $\{w^k\}$ and $\{p^k\}$ be sequences generated by the $(PMAPD)$; let (x^*, z^*) be an optimal solution of (CP), and let y^* be a corresponding Lagrange multiplier, then we have, for each k,*

$$\hat{H}(w^*, w^{k+1}) \leq \hat{H}(w^*, w^k) - (1/2)(\|p^{k+1} - y^{k+1}\|^2 + \|p^{k+1} - y^k\|^2) + \\ + 2\lambda_k^2(\|A\|^2\|x^{k+1} - x^k\|^2 + \|B\|^2\|z^{k+1} - z^k\|^2) + \lambda_k(a_k + b_k),$$

where $w^* = (x^*, z^*, y^*)$. In particular, we have, for all k,

$$\hat{H}(w^*, w^{k+1}) \leq \hat{H}(w^*, w^k) + 2\lambda_k^2(\|A\|^2\|x^{k+1} - x^k\|^2 + \|B\|^2\|z^{k+1} - z^k\|^2) + \lambda_k(a_k + b_k). \quad (41)$$

Proof. Proceeding analogously as the proof of Proposition 4.4, using Lemma 4.3 with $\Delta_k(x,z) = H(x,x^k) - H(x,x^{k+1}) + H'(z,z^k) - H'(z,z^{k+1})$, we obtain the result. □

Then, we obtain the following result

PROPOSITION 4.9 *Let $(d_0, H_0) \in \mathcal{F}(\bar{C})$, $(d'_0, H'_0) \in \mathcal{F}(\bar{K})$ (substituting (Iii) by (Iii)$'$) satisfying the condition (Iviii). Suppose that the Assumptions $(H1), (H2), (H3)$ and $(H4)'$ are satisfied. Let $\{w^k\}$ be a sequence generated by the $(PMAPD)$ algorithm, then $\{w^k\}$ is bounded and every accumulation point of $\{w^k\}$ is a saddle point of the Lagrangian L.*

Proof. The proof is analogous as the proof of Proposition 4.6. First, note that Assumption (H3) together with Assumption (H4)$'$ ensures that

$$\sum_{k=0}^{\infty} 2\lambda_k^2(\|A\|^2\|x^{k+1} - x^k\|^2 + \|B\|^2\|z^{k+1} - z^k\|^2) + \lambda_k(a_k + b_k) < +\infty. \quad (42)$$

Then, from (41), we have

$$w^k \in L_{\hat{H}}(w^*, \bar{\alpha}) := \{w \,:\, \hat{H}(w^*, w) \leq \bar{\alpha}\}, \text{ for all } k,$$

where

$$\bar{\alpha} = \hat{H}(w^*, w^0) + \sum_{k=0}^{+\infty} 2\lambda_k^2(\|A\|^2\|x^{k+1} - x^k\|^2 + \|B\|^2\|z^{k+1} - z^k\|^2) + \lambda_k(a_k + b_k).$$

This implies that $\{w^k\}$ is bounded due to condition (Iiv) in Definition 4 for H and H'.
Now, let (x^*, z^*) be an optimal solution of (CP), let y^* be a corresponding Lagrange multiplier, and let $w^* = (x^*, z^*, y^*)$.
Then, from (41), (42) and Lemma 4.5, we obtain that $\{\hat{H}(w^*, w^k)\}$ is convergent, i.e., there exists $\beta \geq 0$ such that

$$\lim_{k\to+\infty} \hat{H}(w^*, w^k) = \beta. \tag{43}$$

Taking limsup in the first inequality of Proposition 4.8, we obtain

$$\begin{aligned}\limsup_{k\to+\infty} \frac{1}{2}(\|p^{k+1} - y^{k+1}\|^2 + \|p^{k+1} - y^k\|^2)) - 2\lambda_k^2(\|A\|^2\|x^{k+1} - x^k\|^2 + \|B\|^2\|z^{k+1} - z^k\|^2)\\ -\lambda_k(a_k + b_k) \leq 0,\end{aligned} \tag{44}$$

then from Assumptions (H3) and (H4)$'$, we obtain

$$\begin{aligned}\|x^{k+1} - x^k\| &\to 0\\ \|z^{k+1} - z^k\| &\to 0\\ \|p^{k+1} - y^{k+1}\| &\to 0\\ \|p^{k+1} - y^k\| &\to 0.\end{aligned} \tag{45}$$

Now, let $\bar{w} = (\bar{x}, \bar{z}, \bar{y})$ be a cluster point of $\{w^k\}$, and let $\{w^k\}_{k\in K}$ a subsequence converging to $\bar{w}$. Using Lemma 4.3, (i), for all $x \in dom f \cap \bar{\mathtt{C}}$ and $z \in dom\, g \cap \bar{\mathtt{K}}$, we have

$$\lambda_k\{L(x^{k+1}, z^{k+1}, p^{k+1}) - L(x, z, p^{k+1})\} \leq \Delta_k(x, z) + \lambda_k(a_k + b_k). \tag{46}$$

Taking the limit over the appropriate subsequences on both sides of (46), using (45) and the Assumptions (H3), (H4)$'$, and condition (Iviii) for H and H', we obtain

$$L(\bar{x}, \bar{z}, \bar{y}) - L(x, z, \bar{y}) \leq 0, \tag{47}$$

since λ_k is bounded below by $\kappa > 0$.

Similarly, by taking the limit, over the appropriate subsequences, on both sides of the inequality in Lemma 4.3, (ii), we have that, for all $y \in I\!R^m$,

$$L(\bar{x}, \bar{z}, y) - L(\bar{x}, \bar{z}, \bar{y}) \leq 0. \tag{48}$$

From (47) and (48), we obtain that, for all $x \in dom f \cap \bar{\mathtt{C}}$, $z \in dom\, g \cap \bar{\mathtt{K}}$, $y \in I\!R^m$,

$$L(\bar{x}, \bar{z}, y) \leq L(\bar{x}, \bar{z}, \bar{y}) \leq L(x, z, \bar{y}).$$

Therefore, $\bar{w}$ is a saddle point of L. □

Finally, proceeding analogously as the proof of Theorem 4.7, using the result of Proposition 4.9, we have the following main result:

THEOREM 4.10 *Let $(d_0, H_0) \in \mathcal{F}_+(\bar{C})$, $(d_0', H_0') \in \mathcal{F}_+(\bar{K})$ (substituting (Iii) by (Iii)$'$) satisfying the condition (Iviii). Suppose that the Assumptions*

$(H1), (H2), (H3)$ and $(H4)'$ are satisfied and let $\{(x^k, z^k, y^k)\}$ be the sequence generated by the (PMAPD), then the sequence $\{(x^k, z^k, y^k)\}$ globally convergences to (x^*, z^*, y^*), with (x^*, z^*) optimal for (CP) and y^* be a corresponding Lagrange multiplier.

Conclusions and future researches

In this article, we have proposed an inexact proximal multiplier method using regularized proximal distances called (PMAPD) for solving convex minimization problems with a separable structure. This method is an extension of the (PCPM) and (NPCPMM) methods given by Chen-Teboulle and Kyono-Fukushima respectively, and includes the class of regularized $\varphi-$divergence distances. Note also that the (EPDM) is a particular case of the our method when we consider exact iterations and we use the regularized log-quadratic distance. We have established, under standard assumptions, that the iterations generated by the method are well defined and the sequence converges to an optimal solution of the problem.

A future research may include to find a variant of the (PMAPD) method to obtain superlinear rate of convergence, as also, to perform computational experiments and comparison with other methods in the literature. The extension of the (PMAPD) to solve more general problems, such as, quasiconvex optimization and variational inequalities are desirable.

Acknowledgements

This research, for the first author, was supported by CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior). Part of the research of the second author was supported by the PosDoctoral Scolarchip CAPES-FAPERJ Edital PAPD-2011. The authors wish to thank at Systems Engineering and Computer Science Program, COPPE, Federal University of Rio de Janeiro for its backing.

References

[1] A. Auslender, M. Teboulle, S. Ben-Tiba, *A logarithmic-quadratic proximal method for variational inequalities,* Comput. Optim. Appl., 12 (1999), pp. 31-40.

[2] A. Auslender, M. Teboulle, S. Ben-Tiba, *Interior proximal and multiplier methods based on second order homogeneous functionals,* Math Oper. Research 24, 3, (1999), pp. 645-668.

[3] A. Auslender, M. Teboulle, *Entropic proximal decomposition methods for convex programs and variational inequalities,* Math. Program., Ser. A, vol. 91, (2001), pp. 33-47.

[4] A. Auslender, M. Teboulle, *Interior projection-like methods for monotone variational inequalities,* Math. Program., Ser. A, 104, (2005), pp. 39-68.

[5] A. Auslender, M. Teboulle, *Interior gradient and proximal methods for convex and conic optimization,* SIAM J. Optim., vol. 16, 3, (2006), pp. 697-725.

[6] D. P. Bertsekas, *Constrained Optimization and Lagrange Multiplier Methods,* Academic Press, (1982), New York, NY.

[7] D. P. Bertsekas and J. N. Tsitsiklis, *Parallel and distributed computation, Numerical Methods,* Prentice Hall, (1989), Englewood Cliffs, New Jersey.

[8] G. Chen and M. Teboulle, *Convergence analysis of a proximal-like minimization algorithm using Bregman functions,* SIAM J. Optim., 3 (1993), pp. 538-543.

[9] G. Chen and M. Teboulle, *A proximal-Based decomposition method for convex minimization problems,* Mathematical Programming, Serie A, 1, vol. 64, (1994), pp. 81-101.

[10] J. Chifflet, P. Mahey and V. Reiner, *Proximal decomposition for multicommodity flow problems with convex costs,* Telecommun. Syst. vol. 3, (1994), pp. 1-10.

[11] D. Gabay, *Chapter ix applications of the method of multipliers to variational inequalities,* Studies in mathematics and its applications, vol. 15, (1983), pp. 299-331.

[12] D. Gabay and B. Mercier, *A Dual algorithm for the solution of nonlinear variational problems via finite element approximations,* Computers and Mathematics with Applications, vol. 2, (1976), pp. 17-40.

[13] R. Glowinski and P. Le Tallec, *Augmented lagrangian and operator-splitting methods in nonlinear mechanics,* SIAM, Studies in Applied Mathematics SIAM, (1989), Philadelphia, PA.

[14] D. Goldfarb, S. Ma and K. Scheinberg, *Fast alternating linearization methods for minimizing the sum of two convex functions,* Math. Program. 141 (2013), no. 1-2, Ser. A, 349-382.

[15] K. C. Kiwiel, *Proximal minimization methods with generalized Bregman functions,* SIAM J. Control Optim., 35 (1997), pp. 1142-1168.

[16] M. Kyono and M. Fukushima, *Nonlinear proximal decomposition method for convex programming,* JOTA, vol. 106, N. 2, (2000), pp. 357-372.

[17] A. Lenoir, *Modèles et algorithmes pour la planification de production à moyen terme en environnement incertain. Application de méthodes de décomposition proximales,* Ph.D. Thesis, Université Blaise Pascal - Clermont-Ferrand II, 2008.

[18] P. L. Lions and B. Mercier, *Splitting algorithms for the sum of two nonlinear operators,* SIAM Journal on Numerical Analysis, vol. 16, (1979), pp. 964-979.

[19] P. Mahey, Oualibouch and D.T. Pham, *Proximal decomposition on the graph of a maximal monotone operator,* SIAM J. Optim. vol. 5, N 2, (1995), pp. 454-466.

[20] P. Mahey, A. Ouorou, L. LeBlanc and J. Chifflet, *A new proximal decomposition algorithm for routing in telecommunication networks,* Networks, vol. 31, (1997), pp. 227-238.

[21] B. T. Polyak, *Introduction to Optimization,* Optimization Software, Publications division, New York, NY, 1987.

[22] R. T. Rockafellar, *Convex Analysis,* Princeton University Press, Princeton New Jersey, 1970.

[23] R. T. Rockafellar, *Augmented Lagrangians and Applications of the Proximal Point Algorithm in Convex Programming,* Mathematics of Operations Research, vol. 1, (1976), pp. 97-116.

[24] J. E. Spingarn, *Partial Inverse of a Monotone Operator,* Applied Mathematics and Optimization 10, (1983), pp. 247-265.

[25] J. E. Spingarn, *Applications of the method of partial inverse to convex programming decomposition,* Math. Program. vol. 32, (1985), pp. 199-223.

[26] M. Teboulle, *Convergence of proximal-like algorithms,* SIAM J. Optim., 7 (1997), pp. 1069-1083.