Iterative Methods of Solving Stochastic Convex Feasibility Problems and Applications

Dan Butnariu
Department of Mathematics and Computer Science
University of Haifa, 31905 Haifa, Israel

Alfredo N. Iusem
Instituto de Matemática Pura e Aplicada
Estrada Dona Castorina 110, Jardim Botânico
Rio de Janeiro, R.J., CEP 22460-320, Brazil

Regina S. Burachik
Departamento de Matemática
Pontifícia Universidade Católica de Rio de Janeiro
Rua Marqués de São Vincente 225
Rio de Janeiro, RJ, CEP 22453-030, Brazil

Revised Version: September 1998

Abstract

The stochastic convex feasibility problem (SCFP) is the problem of finding almost common points of measurable families of closed convex subsets in reflexive and separable Banach spaces. In this paper we prove convergence criteria for two iterative algorithms devised to solve SCFPs. To do that, we first analyze the concepts of Bregman projection and Bregman function with emphasis on the properties of their local moduli of convexity. The areas of applicability of the algorithms we present include optimization problems, linear operator equations, inverse problems, etc., which can be represented as SCFPs and solved as such. Examples showing how these algorithms can be implemented are also given.

Keywords: Stochastic convex feasibility problem, Bregman projection, Bregman function, Modulus of convexity, Local moduli of convexity, Very convex function, Totally convex function, Regularly convex function, Bochner integral.

1 Introduction

1.1 In this work we study several iterative methods of solving the *stochastic convex feasibility problem* (*SCFP*) which, according to Butnariu and Flam [13], can be stated as follows: *Given a complete probability space* $(\Omega, \mathcal{A}, \mu)$, *a reflexive separable Banach space* B *and a measurable closed convex valued point-to-set mapping* $Q : \Omega \to B$, *find an almost common point of the sets* Q_ω, $\omega \in \Omega$, *provided that such a point exists.* Recall that the point-to-set mapping Q is called *measurable* if for each closed set $F \subseteq B$, we have

$$\{\omega \in \Omega : Q_\omega \cap F{=}\emptyset\} \in \mathcal{A} \tag{1}$$

An *almost common point* of the sets Q_ω, $\omega \in \Omega$, is a point $x^* \in B$ such that

$$\mu\left(\{\omega \in \Omega : x^* \in Q_\omega\}\right) = 1, \tag{2}$$

that is x^* belongs to almost all sets Q_ω. Note that, according to (1), the set

$$Q^{-1}(\{x^*\}) := \{\omega \in \Omega : x^* \in Q_\omega\}$$

is measurable and, thus, formula (2) makes sense. In the sequel we denote by $\sqcap_\mu Q_\omega$ the collection of almost common points of the sets Q_ω, $\omega \in \Omega$, i.e.,

$$\sqcap_\mu Q_\omega := \bigcup_{\{\Omega_0 \in \mathcal{A}; \mu(\Omega_0)=1\}} \bigcap_{\omega \in \Omega_0} Q_\omega.$$

1.2 The SCFP appears in applied mathematics in various forms. It was pointed out in [13] that the problem of determining the inverse of the x-ray transform in computed tomography as well as the best approximation problem in $\mathcal{L}^1([a,b])$ are, essentially, SCFPs in specific Hilbert spaces and can be solved as such. Also, Butnariu [10] has noted that the optimization problem of finding a minimizer for the convex differentiable function $f : \mathbb{R}^n \to \mathbb{R}$ over the nonempty compact convex set $X \subset \mathbb{R}^n$ reduces to a SCFP in $\mathbb{R}^n$. Kammerer and Nashed [25] made clear that Fredholm integral equations of the first kind in $\mathcal{L}^2([a,b])$ can be treated and solved as particular SCFPs. This idea was extrapolated by Butnariu [10] to a larger class of linear operator equations $Tx = b$ in the space $\mathcal{L}^2([a,b])$. These raise the question of how solutions of SCFPs *in Hilbert spaces* can be computed. This question was dealt with in [13], [10] and [12] where it was shown that a class of algorithms which are based on iterative averaging of metric or entropic projections of the current iterate onto the sets Q_ω are well-defined and converge weakly, and sometimes even strongly, to solutions of the given SCFP. What these algorithms have in common is that they proceed from an iterative step to another by averaging Bregman projections (see Section 2.1 below) onto the sets Q_ω, $\omega \in \Omega$, with respect to the square of the norm in the case of the algorithms involving metric projections and with respect to the negentropy (see Section 3.3) in the case of the algorithms involving entropic projections. This fact naturally led to the question of *whether it is possible to guarantee convergence of algorithms based on averaging Bregman projections with respect to functions other than the square of the norm and the negentropy.* By enlarging the pool of functions of this kind one improves the chances of finding computationally efficient algorithms for reaching solutions of SCFPs. The particular Bregman projection based algorithms approached in [13], [10] and [12] were proved to converge under the assumption that the space B in which the SCFP is given is Hilbertian. However, optimization problems and linear operator equations occurring in some non-Hilbertian Banach spaces can be also represented as SCFP (see Section 6 below). This makes us to push the previous question even further: *Can we find algorithms based on computing Bregman projections which converge weakly or strongly when applied to solving SCFPs in non-Hilbertian Banach spaces?*

1.3 The aim of this work is to show that the answers to both questions asked above are affirmative. Precisely, we define and analyze the convergence of two algorithms we call Algorithm A and Algorithm B which are based on computing Bregman projections onto closed convex sets in the Banach space B. Algorithm A (see Section 4.3) is a procedure oriented towards generating a sequence of points $\{x^k\}_{k\in\mathbb{N}}$ in B in such a way that the measure μ of the set of those $\omega \in \Omega$ such that $x^k \notin Q_\omega$ converges to zero as $k \to \infty$. Theorem

4.3 gives sufficient conditions for the sequences generated via Algorithm A to converge weakly or strongly to solutions of specific SCFPs. Algorithm B is a generalization of the algorithms based on averaging the Bregman projections onto the sets Q_ω, $\omega \in \Omega$, discussed above. It is oriented to producing sequences $\{x^k\}_{k\in\mathbb{N}}$ such that the averaged distance of x^k to the sets Q_ω converges to zero. Theorem 5.7 guarantees applicability of Algorithm B to a class of SCFPs which is significantly larger than that for which Algorithm A can be used. However, the conditions under which Theorem 5.7 is proved restrict the pool of functions f with respect to which the Bregman projections involved in the algorithm can be computed while ensuring convergence of the procedure. An advantage of the Algorithm B over the Algorithm A is that the Algorithm B may converge even if the given SCFP is inconsistent (i.e., the sets Q_ω have no almost common points). In such situations, the weak limits of the sequences generated by Algorithm B are averages of their own Bregman projections onto the sets Q_ω with respect to the chosen function f and, in some circumstances, they can be interpreted as natural replacements of the inexistent solution of the given SCFP. Also for this reason, the Algorithm B can be employed as an optimization procedure by taking into account that the sequences it generates approximate minimizers of the function Υ_f defined in Section 5.5. For instance, solving the problem of transport-efficient emplacement of a distribution center ([15, Section 5]) for a commodity (like electricity) whose transportation cost is proportional with the squared distance between the distribution center and the regions Q_ω, $\omega \in \Omega$, to which the commodity should be transported amounts to minimizing the function Υ_f associated to $f(x) = \|x\|^2$ in a space $\mathbb{R}^n$ and Algorithm B can be used for approximating a solution of it.

1.4 The conditions under which our convergence results for Algorithms A and B are proved are discussed in Sections 6.1 and 6.2 where we point out ways of relaxing some of the requirements of Theorems 4.3 and 5.7. It should be noted that even in their unrelaxed forms the conditions under which our convergence proofs are done are sufficient for allowing resolution of significant problems of applied mathematics along the lines suggested by the examples given in the Sections 6.5 and 6.6. The applications discussed in Section 6 refer to cases in which the set Ω involved in the given SCFP is infinite. When Ω is finite, $\mathcal{A} = 2^\Omega$ and $\mu(A) = \sum_{\omega\in A} \mu_\omega$ for some positive real numbers μ_ω, ($\omega \in \Omega$), then the SCFP reduces to the so called *convex feasibility problem* which is of interest in many practical applications (see Censor and Zenios [17]). In this setting, the Algorithm A is a new addition to the list of iterative methods of solving convex feasibility problems (see [4] and [20] for surveys) which starts with Johannes von Neumann's [28] successive projection algorithm. Placed in the context of the convex feasibility problems, the Algorithm B falls in the class of the "simultaneous" procedures of approximating common points of convex sets which originates in Cimmino's work [18]. By contrast to most of the existing methods, the Algorithms A and B work not only in finite dimensional Banach spaces or in Hilbert spaces but also in reflexive Banach spaces whenever suitable Bregman functions can be found. Even when applied to convex feasibility problems in spaces where other algorithms of similar nature are available, the Algorithms A and B present the advantage that their computational procedures can be eventually simplified by fitting the Bregman functions to the geometry of the sets Q_ω (see Section 1.3).

Solving SCFPs can sometimes be done by algorithms which are completely different from those presented in this work. For instance, one can interpret the SCFP as a problem of minimizing the average of the indicator functions I_{Q_ω} and, then, to apply the Progressive Hedging Algorithm of Rockefellar and Wets [31]. In practice, this amounts to solving a sequence of nonlinear programming problems. In specific situations as those discussed in the Sections 6.5 and 6.6 below, solving the SCFP by our algorithms presents the relative advantage of reducing the computational procedure to iterative application of closed formulae (see Section 6.3).

1.5 This work is organized as follows. In Sections 2 and 3 we present the notions and results concerning Bregman projections and Bregman functions which form the base of our convergence analysis of the Algorithms A and B. We mostly look at Bregman projections and Bregman functions through the properties of their local moduli of convexity because these are the tools we later use for establishing connections between the weak and strong topologies of the Banach space B and the initial topologies on B of specific Bregman functions. These connections are apparent in the proof of Lemma 2.5 and its consequences and make our convergence results possible. Section 3 focuses on the local moduli of convexity of particular Bregman functions which help us show that various sets of conditions involved in our convergence proofs are consistent. It should be noted that the Bregman functions discussed here are among the most likely to be used in practical applications. In

Sections 4 and 5 we present the convergence theorems discussed above and their proofs. Section 6 deals with the applicability of the convergence results. There we show alternative conditions under which the Algorithms A and B still converge and we give examples pointing out ways of implementation. In spite of the fact that in our examples we approach particular problems of applied mathematics, it is quite clear that many other problems of the same kind can be solved in the same manner.

2 Bregman projections

2.1 In this section we present a series of notions and results which are used in our analysis of algorithms for solving SCFPs. To this end, let B be a reflexive and separable Banach space and $f : B \to (-\infty, +\infty]$ a lower semicontinuous function whose domain $\mathcal{D} = Dom(f)$ has nonempty interior $\mathcal{D}^\circ$. We assume that f is convex on $\mathcal{D}$ and Gâteaux differentiable on $\mathcal{D}^\circ$ and denote $f'(x, d)$ the Gâteaux derivative of f at the point $x \in \mathcal{D}^\circ$ in the direction d. Recall (see, for instance, [29, Proposition 3.3 and Corollary 1.7]) that, in these circumstances, the function f is continuous on $\mathcal{D}^\circ$ and, therefore, for each $x \in \mathcal{D}^\circ$, the linear functional $f'(x, \cdot) : B \to \mathbb{R}$ is continuous too. By consequence, for each $x \in \mathcal{D}^\circ$, the function $D_f(\cdot, x) : \mathcal{D} \to \mathbb{R}_+$ given by

$$D_f(y, x) = f(y) - f(x) - f'(x, y - x), \tag{3}$$

is continuous on $\mathcal{D}^\circ$ and convex on $\mathcal{D}$. Note that, if f is strictly convex on $\mathcal{D}$, then so is $D_f(\cdot, x)$. Therefore, if f is strictly convex on $\mathcal{D}$ and if $K \subseteq B$ is a closed convex set such that $K \cap \mathcal{D}{=}\emptyset$, then there exists at most one point $\Pi_K^f(x) \in K \cap \mathcal{D}$ such that

$$D_f(\Pi_K^f(x), x) = \inf \left\{ D_f(y, x) : y \in K \cap \mathcal{D} \right\}. \tag{4}$$

This point (if any) was called *Bregman projection of $x \in \mathcal{D}^\circ$ onto K* by Censor and Lent [16] (see also [6]). A question of essential interest in the sequel is whether, and under which conditions concerning the function f, the Bregman projection $\Pi_K^f(x)$ exists and is unique, for each $x \in \mathcal{D}^\circ$.

2.2 The following three conditions are sufficient for ensuring the existence and uniqueness of the Bregman projection $\Pi_K^f(x)$ for each $x \in D^\circ$:

$(B1)$ *The domain $\mathcal{D}$ of f is closed with nonempty interior $\mathcal{D}^\circ$and f is Gâteaux differentiable on $\mathcal{D}^\circ$;*

$(B2)$ *The function f is strictly convex and continuous on $\mathcal{D}$;*

$(B3)$ *For each $x \in \mathcal{D}^\circ$ and for any real number α, the level set $L^f_\alpha(x)$ defined by*

$$L^f_\alpha(x) = \{y \in \mathcal{D} : D_f(y,x) \le \alpha\} \tag{5}$$

is bounded.

This is shown by the next result.

Lemma. *If f satisfies the conditions $(B1)$, $(B2)$ and $(B3)$, then, for every closed convex set $K \subseteq B$ such that $K \cap \mathcal{D}{=}\emptyset$, the Bregman projection $\Pi^f_K(x)$ of each $x \in \mathcal{D}^\circ$ onto the set K exists and is unique.*

Proof. Let z be any point in $K\cap\mathcal{D}$ and denote $r = D_f(z,x)$. The set $U = K\cap\mathcal{D}\cap L^f_r(x)$ is nonempty, closed, convex and bounded and, thus, it is weakly compact. The function $D_f(\cdot,x)$ is convex and continuous and, therefore, weakly lower semicontinuous on the weakly compact set U. Consequently, $D_f(\cdot,x)$ achieves its minimal value at some point in U. This point obviously satisfies (4) and it is the unique point with this property because $D_f(\cdot,x)$ is strictly convex.

2.3 In order to identify classes of functions satisfying the conditions $(B1)$-$(B3)$ we use the next elementary result:

Lemma. *If $h : B \to (-\infty,+\infty]$ is a convex function such that $C := Int(Dom(h)) = \emptyset$, then, for any $z \in C$, the right-hand side derivative in each direction $d \in B$,*

$$h'(z,d) = \lim_{\gamma\searrow 0} \frac{h(z+\gamma\cdot d) - h(z)}{\gamma}$$

exists and the following statements are equivalent:

(i) *The function h is strictly convex on C;*

(ii) *If $x,y \in C$, and $x = y$, then $h'(x,x-y) > h'(y,x-y)$.*

Proof. The existence of $h'(x,d)$ follows, for instance, from [29, Lemma 1.2]. The equivalence of (i) and (ii) results from the fact that the function $\gamma \to \frac{h[y+\gamma\cdot(x-y)]-h(y)}{\gamma}$ is (strictly) increasing on $[0,1]$ when h is (strictly) convex.

2.4 Now, let $f : B \to (-\infty, +\infty]$ be a convex function which satisfies the condition $(B1)$. The *lower local modulus of convexity of f at $x \in \mathcal{D}^\circ$* is the function[1] $\nu_f^-(x, \cdot) : [0, +\infty) \to [0, +\infty]$ defined by

$$\nu_f^-(x,t) = \inf\{D_f(y,x) : y \in \mathcal{D}, \|y-x\| = t\}. \tag{6}$$

The function f is called *totally convex* if, for each $x \in \mathcal{D}^\circ$, its lower local modulus of convexity at x, $\nu_f^-(x,\cdot)$, is positive on $(0, +\infty)$.

The *upper local modulus of convexity of f at $x \in \mathcal{D}^\circ$*, $\nu_f^+(x,\cdot) : [0, +\infty) \to [0, +\infty]$, defined by

$$\nu_f^+(x,t) = \inf\{D_f(y,x) : y \in \mathcal{D}^\circ, \|y-x\| = t\}, \tag{7}$$

was studied in [12]. Recall (cf. [12]) that f is called *very convex* if, for each $x \in \mathcal{D}^\circ$, $\nu_f^+(x,t) > 0$, for all $t \in (0, +\infty)$. Obviously, $\nu_f^+(x,t) \geq \nu_f^-(x,t)$ and, therefore, if f is totally convex, then it is also very convex. The proof of Proposition 2.4 in [12] can be reproduced without essential modifications in order to show that, similarly to the upper local modulus of convexity, the lower local modulus of convexity has the property

$$c \geq 1 \quad and \quad t \geq 0 \Rightarrow \nu_f^-(x, c \cdot t) \geq c \cdot \nu_f^-(x,t). \tag{8}$$

By consequence, f is totally convex if and only if $\nu_f^-(x,\cdot)$ is strictly increasing on $[0, +\infty)$. Observe that, according to (8), we have $\nu_f^-(x,t) \geq t \cdot \nu_f^-(x,1)$, for any $t \geq 1$. Hence, if the function f is totally convex, then

$$\lim_{t\to\infty} \nu_f^-(x,t) = +\infty. \tag{9}$$

Note that, if $\mathcal{D} = B$, then the upper and the lower modulus of convexity coincide.

2.5 Lemma 2.2 shows that the conditions $(B1)$-$(B3)$ are sufficient for ensuring well-definiteness of the Bregman projection operator $x \to \Pi_K^f(x) : \mathcal{D}^\circ \to K$ for every closed, convex subset K of B such that $K \cap \mathcal{D} = \emptyset$. Verification of these conditions, and especially of $(B3)$, in practical situations may be, sometimes, difficult. The following result shows that, if B is of finite dimension, then functions f satisfying the conditions $(B1)$ and $(B2)$ also satisfy

[1] We make the usual convention that the infimum of an empty set of real numbers is $+\infty$.

the condition $(B3)$. If the space B has infinite dimension, then totally convex functions and continuous very convex functions which satisfy $(B1)$ satisfy $(B3)$ too.

Lemma. *Suppose that the convex function $f : B \to (-\infty, +\infty]$ satisfies the condition $(B1)$. Then the following implications hold:*

(i) If f is very convex, then f is strictly convex on $\mathcal{D}^\circ$.

(ii) If f is very convex and continuous on $\mathcal{D}$, then f satisfies $(B3)$.

(iii) If f is totally convex, then f satisfies $(B3)$.

Moreover, if the Banach space B has finite dimension and, in addition to $(B1)$, the function f also satisfies the condition $(B2)$, then f is totally convex, it satisfies the condition $(B3)$ and, for each $x \in B$, the lower local modulus of convexity, $\nu_f^-(x, \cdot)$, is continuous from the left on $(0, +\infty)$.

Proof. (i) Suppose, by contradiction, that f is very convex but not strictly convex on $\mathcal{D}^\circ$. Then, according to Lemma 2.3, for some $x, y \in \mathcal{D}^\circ$ such that $x{=}y$

$$f'(x, x-y) \leq f'(y, x-y).$$

Since f is convex, we also have

$$f'(x, x-y) \geq f(x) - f(y) \geq f'(y, x-y).$$

Hence,

$$f'(x, x-y) = f'(y, x-y). \tag{10}$$

Define the function $f_0 : \mathcal{D} \to \mathbb{R}$ by

$$f_0(z) = f(z) + f'(x, x-z).$$

Clearly, f_0 is convex on $\mathcal{D}$ and, for any $z \in \mathcal{D}^\circ$,

$$f'(z, \cdot) = f'(z, \cdot) - f'(x, \cdot).$$

This implies $f'(x, \cdot) = 0$, that is x is a global minimizer of f_0. Note that (10) implies

$$f(x) - f(y) \leq f'(x, x-y) = f'(y, x-y) \leq f(x) - f(y),$$

that is

$$f(x) - f(y) = f'(x, x-y) = f'(y, x-y).$$

Therefore,

$$f_0(x) = f(x) = f(y) + f'(x, x - y) = f_0(y).$$

Hence, y is also a global minimizer of f_0. Consequently,

$$0 = f'(y, \cdot) = f'(y, \cdot) - f'(x, \cdot),$$

showing that $f'(y, \cdot) = f'(x, \cdot)$. According to (3) this means

$$0 \leq D_f(x, y) = f(x) - f(y) - f'(y, x - y)$$

$$= f(x) - f(y) - f'(x, x - y) = -D_f(y, x) \leq 0.$$

In other words, we have $D_f(y, x) = 0$, which implies $\nu_f^+(x, \|y - x\|) = 0$. Since f is very convex, this is not possible unless $x = y$ that is, we reached a contradiction.

(ii) Suppose that the function f is very convex and continuous on $\mathcal{D}$ and that, for some $x \in B$ and for some real number $\alpha \geq 0$, the level set $L_\alpha^f(x)$ is unbounded. Then there exists an unbounded sequence $\left\{u^k\right\}_{k \in \mathbb{N}} \subset L_\alpha^f(x)$ and the function $D_f(\cdot, x)$ is continuous on $\mathcal{D}$. Consequently, for each nonnegative integer k, there exists a point $w^k \in \mathcal{D}^\circ$ such that $\left\|u^k - w^k\right\| < \frac{1}{k}$ and

$$D_f(w^k, x) \leq D_f(u^k, x) + \frac{1}{k} \leq \alpha + \frac{1}{k}.$$

These and [12, Proposition 2.4] which shows that $\nu_f^+(x, c \cdot t) \geq c\nu_f^+(x, t)$ when $c \geq 1$ and $t \geq 0$, imply that the sequence $\left\{w^k\right\}_{k \in \mathbb{N}}$ is unbounded and, for each k for which $\left\|w^k - x\right\| \geq 1$ we have

$$\nu_f^+(x, 1) \cdot \left\|w^k - x\right\| \leq \nu_f^+(x, \left\|w^k - x\right\|) \leq D_f(w^k, x) \leq \alpha + \frac{1}{k},$$

that is, α is not finite.

(iii) If the result does not hold, then, exactly as in (ii), there exists an unbounded sequence $\left\{u^k\right\}_{k \in \mathbb{N}}$ in $L_\alpha^f(x)$ and, for each nonnegative integer k, we have

$$\nu_f^-(x, \left\|u^k - x\right\|) \leq D_f(u^k, x) \leq \alpha.$$

Since, according to (8), for all positive integers k such that $\left\|u^k - x\right\| \geq 1$, we also have

$$\nu_f^-(x, \left\|u^k - x\right\|) \geq \left\|u^k - x\right\| \cdot \nu_f^-(x, 1),$$

where $\nu_f^-(x,1) > 0$, it results that α is not finite and this is a contradiction.

Now, suppose that B is of finite dimension and f is continuous and strictly convex on the closed set $\mathcal{D}$. Fix $x \in \mathcal{D}^\circ$ and $t \in (0,+\infty)$. Note that the set

$$\{y \in \mathcal{D} : \|y - x\| = t\}$$

is compact in B because it is bounded and closed. Also, $D_f(\cdot, x)$ is continuous on $\mathcal{D}$. Consequently, there exists a point $y \in \mathcal{D}$ with $\|y - x\| = t$ and such that

$$\nu_f^-(x,t) = D_f(y,x).$$

Since f is strictly convex on $\mathcal{D}$, we have

$$f'(x, y-x) \leq \frac{f[y + \gamma(x-y)] - f(y)}{\gamma} < f(y) - f(x),$$

for all $\gamma \in (0,1)$. This implies that $D_f(y,x) > 0$. Hence,

$$\nu_f^-(x,t) = D_f(y,x) > 0$$

and this shows that f is totally convex.

Now, suppose that $t_* \in (0,+\infty)$ and that the sequence$\{t_k\}_{k\in\mathbb{N}} \subseteq (0,+\infty)$ converges nondecreasingly to t_*. For each nonnegative integer k, let $y^k \in \mathcal{D}$ be a point such that $\left\|y^k - x\right\| = t_k$ and $\nu_f^-(x,t_k) = D_f(y^k, x)$. The sequence $\left\{y^k\right\}_{k\in\mathbb{N}}$ is bounded because $\{t_k\}_{k\in\mathbb{N}}$ converges. Hence, there exists a convergent subsequence $\left\{y^{k_p}\right\}_{p\in\mathbb{N}}$ of $\left\{y^k\right\}_{k\in\mathbb{N}}$. We denote $y^* = \lim_{p\to\infty} y^{k_p}$ and we have that $y^* \in \mathcal{D}$ and $\|y^* - x\| = t_*$. Since, for any $k \in \mathbb{N}$,

$$\nu_f^-(x,t_*) \geq \nu_f^-(x,t_k),$$

and the sequence $\left\{\nu_f^-(x,t_k)\right\}_{k\in\mathbb{N}}$ is nondecreasing, it results that

$$\nu_f^-(x,t_*) \geq \lim_{k\to\infty} \nu_f^-(x,t_k) = \lim_{p\to\infty} \nu_f^-(x,t_{k_p}) =$$

$$\lim_{p\to\infty} D_f(y^{k_p}, x) = D_f(y^*, x) \geq \nu_f^-(x,t_*).$$

Consequently, $\nu_f^-(x,t_*) = \lim_{k\to\infty} \nu_f^-(x,t_k)$, i.e., $\nu_f^-(x,\cdot)$ is continuous from the left at t_*.

2.6 Lemma 2.5 may leave the feeling that, whenever f satisfies the conditions $(B1)$ and $(B2)$ it is also totally convex and, thus, it automatically satisfies $(B3)$. This indeed happens when B is of finite dimension. The following example shows that, if B is of infinite dimension, this is not generally the case, i.e., on spaces of infinite dimension functions which satisfy $(B1)$ and $(B2)$ are not necessarily very convex and may not satisfy $(B3)$. Let $B = \ell^2$ be the Hilbert space of all square-summable sequences of real numbers and define $f : B \to \mathbb{R}$ by

$$f(\{t_k\}_{k\in\mathbb{N}}) = \sum_{k=1}^{\infty} \frac{t_k^2}{k}.$$

This function is strictly convex and Gâteaux differentiable on B and has

$$f'(\{t_k\}_{k\in\mathbb{N}}, \cdot) = \left\{2 \cdot \frac{t_k}{k}\right\}_{k\in\mathbb{N}_+}.$$

However, if $e^n = \{\delta_k^n\}_{k\in\mathbb{N}}$, where $\delta_k^n = 0$ if k=n and $\delta_k^n = 1$, otherwise, then

$$D_f(e^n, 0) = \frac{1}{n},$$

for any positive integer n. This and the fact that $\|e^n - 0\| = 1$, for any integer $n \geq 1$, imply that $\nu_f^+(0,1) = 0$, that is, the function f is not very convex. Moreover, the unbounded sequence $y^n := \{y_k^n\}_{k\in\mathbb{N}}$, $n \in \mathbb{N}$, which is defined by $y_k^n = \frac{1}{k}$, if n=k, and $y_n^n = \sqrt{n}$, belongs to $L_\alpha^f(0)$ for $\alpha = 1 + \sum_{k=1}^{\infty} \frac{1}{k^3}$. Hence, the function f does not satisfy $(B3)$.

3 Bregman functions

3.1 In this work we essentially deal with *Bregman functions*, that is functions $f : B \to (-\infty, +\infty]$ which, in addition to the conditions $(B1)$, $(B2)$ and $(B3)$, also satisfy the following requirements:

$(B4)$ *For each $y \in \mathcal{D}$ and for any real number $\alpha \geq 0$, the level set $R_\alpha^f(y)$ defined by*

$$R_\alpha^f(y) = \{x \in \mathcal{D}^\circ : D_f(y,x) \leq \alpha\}, \tag{11}$$

is bounded;

$(B5)$ *If the sequences* $\{x^k\}_{k\in\mathbb{N}} \subset \mathcal{D}^\circ$ *and* $\{y^k\}_{k\in\mathbb{N}} \subset \mathcal{D}$ *are such that* $\{x^k\}_{k\in\mathbb{N}}$ *is bounded and*

$$\lim_{k\to\infty} D_f(y^k, x^k) = 0,$$

then $\lim_{k\to\infty} \left\|x^k - y^k\right\| = 0.$

Bregman functions, first introduced in [6], were studied by Censor and Lent [16] and De Pierro and Iusem [22] and were placed in a more general context by Bauschke and Borwein [5]. The interest in Bregman functions is mostly due to the fact that they are good replacements for the squared norm in various algorithms whose convergence in specific settings is essentially conditioned by the properties $\|\cdot\|^2$. Imposing conditions on the function $\|\cdot\|^2$ means imposing conditions on the geometry of the space B in which the problem is placed and this implicitly leads to restricting the area of applicability of the algorithms. The algorithms we discuss in this work are applicable in Banach spaces under the only condition that an appropriate Bregman function $f : B \to (-\infty, +\infty]$ can be found in order to ensure convergence. Finding appropriate Bregman functions on specific Banach spaces is sometimes possible even if the space B itself has a geometry which makes the function $\|\cdot\|^2$ unfit for our purpose (e.g., $\|\cdot\|^2$ may be nondifferentiable, but a Bregman function on B may still exist).

3.2 In order to identify a class of Bregman functions which will be further used for illustrating the behavior of Algorithms A and B, recall (cf. [32]) that a proper convex function $f : B \to (-\infty, +\infty]$ is called *uniformly convex* if the function $\delta_f : [0, +\infty) \to [0, +\infty)$ defined by

$$\delta_f(t) = \inf \tau \in (0,1), x, y \in Dom(f), \|x-y\| = t \left\{ \frac{\tau \cdot f(x) + (1-\tau)\cdot f(y) - f[\tau\cdot x + (1-\tau)\cdot y]}{\tau\cdot(1-\tau)} \right\} \tag{12}$$

is positive on $(0, +\infty)$. The function δ_f is called *the modulus of convexity of* f. With that in mind we prove the following result:

Proposition. *If the function* $f : B \to (-\infty, +\infty]$ *is uniformly convex, satisfies the condition* $(B1)$ *and is continuous on* $\mathcal{D}$*, then* f *is a Bregman function. In particular, if* $p \geq 2$*, if* $B = L^p([a,b])$ *or* $B = \ell^p$*, and if* $f(x) = \|x\|^p$*, then* f *is an uniformly convex Bregman function having the modulus* *of convexity*

$$\delta_f(t) = \left(\frac{1}{2}\right)^{p-1} \cdot t^p. \tag{13}$$

Proof: Note that uniformly convex functions are strictly convex on their domain (cf. (12)) and, thus, f satisfies $(B2)$. It follows from [32, Remark 2] that $\delta_f(t) \leq \nu_f^-(x,t)$, for all $x \in \mathcal{D}^\circ$ and $t \geq 0$. Hence, if f is uniformly convex, then it is totally convex and, according to Lemma 2.5, f satisfies $(B3)$. It was proven in [32, Lemma 1] that, if f is uniformly convex, then

$$c \geq 1 \quad and \quad t \geq 0 \Rightarrow \delta_f(c \cdot t) \geq c^2 \cdot \delta_f(t). \tag{14}$$

Consequently, if we assume that there exists an unbounded sequence $\left\{x^k\right\}_{k\in\mathbb{N}} \subset R_\alpha^f(x)$, then, for each k for which $\left\|x^k - x\right\| \geq 1$, we have

$$\delta_f(1) \cdot \left\|x^k - x\right\|^2 \leq \delta_f(\left\|x^k - x\right\|) \leq D_f(x, x^k) \leq \alpha,$$

and this is a contradiction. Thus, f satisfies $(B4)$. Now, suppose that $\left\{y^k\right\}_{k\in\mathbb{N}} \subset \mathcal{D}$, $\left\{z^k\right\}_{k\in\mathbb{N}} \subset \mathcal{D}^\circ$ and that $\left\{z^k\right\}_{k\in\mathbb{N}}$ is bounded. Applying (14) we deduce that, for any nonnegative integer k,

$$\delta_f(1) \cdot \left\|y^k - x^k\right\|^2 \leq D_f(y^k, z^k).$$

Therefore, if $\lim_{k\to\infty} D_f(y^k, z^k) = 0$, then $\lim_{k\to\infty} \left\|y^k - x^k\right\| = 0$, that is the function f satisfies $(B5)$.

Now, assume that $f(x) = \|x\|^p$ while $B = L^p([a,b])$ or $B = \ell^p$ and $p \geq 2$. Then, according to [32, Theorem 4], the function f is uniformly convex and (13) holds. As noted in [2, Section 5], the function f is Gâteaux differentiable with

$$f'(x, \cdot) = p \cdot |x|^{p-2} \cdot x. \tag{15}$$

Hence, f is a Bregman function.

3.3 Bregman functions are not always uniformly convex. An example of a nonuniformly convex Bregman function defined on a finite dimensional Banach space is the *negentropy*, that is the function $f : \mathbb{R}^n \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} \sum_{i=1}^n x_i \cdot \ln x_i & if \quad x_1, ..., x_n \geq 0, \\ +\infty & otherwise, \end{cases} \tag{16}$$

with the usual convention that $0 \cdot \ln 0 = 0$. It was shown in [12] that the negentropy is not uniformly convex. However, it has the following property:

Proposition. *The negentropy is a totally convex Bregman function and its lower local modulus of convexity satisfies the following inequality*

$$\nu_f^-(x,t) \geq \min_{1\leq i\leq n} \left\{ x_i \cdot \left[\left(1+\frac{t}{x_i \cdot \sqrt{n}}\right) \cdot \ln\left(1+\frac{t}{x_i \cdot \sqrt{n}}\right) - \frac{t}{x_i . \sqrt{n}} \right] \right\}. \quad (17)$$

Proof. The fact that the negentropy is a Bregman function was noted in [16]. As observed in [12], the lower and the upper local modulus of convexity of the function $g : \mathbb{R} \to \mathbb{R}$ defined by

$$g(t) = \begin{cases} t \cdot \ln t & if \quad t \geq 0, \\ +\infty & otherwise, \end{cases} \quad (18)$$

coincides with

$$\nu_g(x,t) = x \cdot \left[\left(1+\frac{t}{x}\right) \cdot \ln\left(1+\frac{t}{x}\right) - \frac{t}{x} \right]. \quad (19)$$

According to (16), we have

$$D_f(y,x) = \sum_{i=1}^{n} D_g(y_i, x_i).$$

Therefore, if $t > 0$ and if x is a vector in the positive orthant $\mathbb{R}^n_{++}$, then there exists a vector $y^* \in \mathbb{R}^n_+$ such that $\|y^* - x\| = t$ and

$$\nu_f^-(x,t) = D_f(y^*, x) = \sum_{i=1}^{n} D_g(y_i^*, x_i) \geq \sum_{i=1}^{n} \nu_g(x_i, |y_i^* - x_i|).$$

Since for at least one index i_0 we have $\left|y_{i_0}^* - x_{i_0}\right| \geq t/\sqrt{n}$, we deduce

$$\nu_f^-(x,t) \geq \nu_g(x_{i_0}, \left|y_{i_0}^* - x_{i_0}\right|) \geq \nu_g(x_{i_0}, \frac{t}{\sqrt{n}}) \geq \min_{1\leq i\leq n} \nu_g(x_i, , \frac{t}{\sqrt{n}}).$$

This and (19) complete the proof.

3.4 The negentropy is a nonuniformly convex Bregman function on a space of finite dimension. If $p \in (1,2)$ and if $B = \ell^p$ or $B = \mathcal{L}^p([a,b])$, then the function $f(x) = \|x\|^p$ is not uniformly convex (cf. [32, Theorem 5]). However, we have the next result proved in [14]:

Proposition. *If $p \in (1,2)$, then, in each of the following cases*
(i) $B = \mathbb{R}^n$ provided with the norm $\|x\| = \left(\sum_{i=1}^n |x_i|^p\right)^{1/p}$;
(ii) $B = \ell^p$;
(iii) $B = \mathcal{L}^p([a,b])$;
the function $f : B \to (-\infty, +\infty]$ defined by $f(x) = \|x\|^p$ is a totally convex Bregman function and, for every $x \in B$, we have

$$\nu_f^-(x,t) = (\|x\| + t)^p - \|x\|^p - p \cdot \|x\|^{p-1} \cdot t. \tag{20}$$

3.5 Remark. Bregman functions were studied in [6] as natural generalizations of the square of the norm in a Hilbert space. In fact, if B is a Hilbert space and $h(x) = \|x\|^2$, then h is an uniformly convex function with $\delta_h(t) = t^2$ and, thus, h is a Bregman function. Clearly, in this case, we have $D_h(y,x) = \|x - y\|^2$. The underlying reason for the interest in Bregman functions in a more general context consists of the similarity between the properties of the function $D_f(y,x)$ associated to an arbitrary Bregman function f and the properties of the square of the distance $D_h(y,x)$. This similarity was exploited by Alber [1] who bases his convergence results for cyclic Bregman projection algorithms on the fact that, in uniformly convex and uniformly smooth Banach spaces, the function $h(x) = \|x\|^2$ is a Bregman function.

4 The algorithm A

4.1 In this section we consider a SCFP as stated in Section 1.1 in which the point-to-set mapping $Q : \Omega \to B$ is defined by

$$Q_\omega = \{x \in B : g(\omega, x) \leq 0\}, \tag{21}$$

where $g : \Omega \times B \to \mathbb{R}$ is a *convex Carathéodory function*, that is, for each $\omega \in \Omega$, $g(\omega, \cdot) : B \to \mathbb{R}$ is convex and continuous and, for each $x \in B$,

$g(\cdot, x) : \Omega \to \mathbb{R}$ is measurable. Note that, in this case, the measurability of the point-to-set mapping Q defined by (21) is guaranteed (cf. [3, Theorem 8.2.9]) and so are the convexity and closedness of the sets Q_ω.

All over this section we assume that the convex Carathéodory function g satisfies the following conditions:

(I) There exists a point $\bar{x} \in B$ such that $g(\cdot, \bar{x})$ is integrable;

(II) There exists a function $h_1 : \mathbb{R}_+^2 \to \mathbb{R}_+$ and an integrable function $h_2 : \Omega \to \mathbb{R}_+$ such that,

(a) For each $s \in \mathbb{R}_+$, the function $h_1(s, \cdot)$ is continuous on the interval $[0, s]$ and vanishes at zero;

(b) For almost all $\omega \in \Omega$ and for all $x, y \in B$ and $r \in \mathbb{R}_+$,

$$\|x\| \leq r \quad \textit{and} \quad \|y\| \leq r \Rightarrow |g(\omega, x) - g(\omega, y)| \leq h_1(r, \frac{1}{2} \cdot \|x - y\|) \cdot h_2(\omega). \tag{22}$$

4.2 In order to describe the Algorithm A we need some preliminary results which are summarized in the following:

Lemma. *If the convex Carathéodory function g satisfies the conditions (I) and (II), then*

(i) For each $x \in B$, the function $g(\cdot, x) : \Omega \to \mathbb{R}$ is integrable;

(ii) The function $\Psi : B \times B \to \mathbb{R}$ given by

$$\Psi(x, y) = \int_{\Omega(y)} g(\omega, x) d\mu(\omega), \tag{23}$$

where

$$\Omega(y) := \{\omega \in \Omega : g(\omega, y) > 0\}, \tag{24}$$

is well-defined and, for each $y \in B$, the function $\Psi(\cdot, y)$ is convex and continuous on B;

(iii) For every $y \in B$, the set

$$K(y) := \{x \in B : \Psi(x, y) \leq 0\} \tag{25}$$

is convex, closed and includes $\bigcap_\mu Q_\omega$, the collection of all almost common points of the sets Q_ω;

(iv) If $\bigcap_\mu Q_\omega{=}\emptyset$ and if $f : B \to \mathbb{R}$ is a Bregman function, then the operator $\mathbf{P}^f : B \to B$ given by

$$\mathbf{P}^f(x) = \Pi^f_{K(x)}(x), \tag{26}$$

is well-defined.

Proof. (i) Note that, for $r = \max(\|x\|, \|\bar{x}\|)$, we have

$$|g(\omega, x)| \leq |g(\omega, \bar{x})| + |g(\omega, \bar{x}) - g(\omega, x)|$$

$$\leq |g(\omega, \bar{x})| + h_1(r, \tfrac{1}{2} \cdot \|x - \bar{x}\|) \cdot h_2(\omega), \quad a.e.,$$

where the right hand side is integrable. Thus, $g(\cdot, x)$ is integrable.

(ii) According to (i), the function Ψ is well-defined. For any $y \in B$, the function $\Psi(\cdot, y)$ is convex because the functions $g(\omega, \cdot)$, $\omega \in \Omega$, are convex. If $\{x^k\}_{k\in\mathbb{N}}$ is a sequence in B and if $\lim_{k\to\infty} x^k = x^*$, then, for any upperbound r of $\{\|x^k\|\}_{k\in\mathbb{N}}$, we have

$$\left|\Psi(x^k, y) - \Psi(x^*, y)\right| \leq h_1(r, \frac{1}{2} \cdot \left\|x^k - x^*\right\|) \cdot \int_{\Omega(y)} h_2(\omega) d\mu(\omega),$$

showing that $\Psi(\cdot, y)$ is continuous because $\lim_{k\to\infty} h_1(r, \frac{1}{2} \cdot \left\|x^k - x^*\right\|) = 0.$

(iii) is an immediate consequence of (ii).

(iv) According to (iii), if $\bigcap_\mu Q_\omega{=}\emptyset$, then the closed convex set $K(y)$ is nonempty. Thus, Lemma 2.2 applies and guarantees the well-definiteness of the operator $\mathbf{P}^f$.

4.3 The Algorithm A consists of the following procedure:

Algorithm *A. Choose an initial point $x^0 \in \mathcal{D}^\circ$ and, for any integer $k \geq 0$, set*

$$x^{k+1} := \mathbf{P}^f(x^k). \tag{27}$$

We show next that the sequences recursively generated by the Algorithm A converge weakly, and sometimes even strongly, to points in $\bigcap_\mu Q_\omega$.

Theorem. *Suppose that $\bigcap_\mu Q_\omega{=}\emptyset$ and that $f : B \to \mathbb{R}$ is a Bregman function. Then, for each initial point $x^0 \in B$, the following statements hold:*

(i) The sequence $\{x^k\}_{k\in\mathbb{N}}$ *recursively defined by* (27) *exists, is bounded, and all its weak accumulation points are in the set* $\bigcap_\mu Q_\omega$.

(ii) If the Gâteaux derivative $f' : B \to B^*$ *is sequentially weakly-to-weak** *continuous, then the sequence* $\{x^k\}_{k\in\mathbb{N}}$ *converges weakly to a point* $x^* \in \bigcap_\mu Q_\omega$. *In this case, the sequence* $\{x^k\}_{k\in\mathbb{N}}$ *converges strongly to* $x^* \in \bigcap_\mu Q_\omega$ *if and only if*

$$\underline{\lim}_{\,k\to\infty} f'(x^k, x^* - x^k) \geq 0. \tag{28}$$

Proof. (i) Let $x^0 \in B$ be fixed. The existence of the sequence $\{x^k\}_{k\in\mathbb{N}}$ defined by (27) follows from Lemma 4.2 which shows that $\mathbf{P}^f$ is well-defined on B. Lemma 4.2 also implies that

$$\bigcap_\mu Q_\omega \subseteq \bigcap_{x\in B} K(x). \tag{29}$$

Let z be a point in $\bigcap_\mu Q_\omega$. Then, for each $k \in \mathbb{N}$, we have $z \in K(x^k)$ and, thus, application of Corollary 2.6 in [2] yields

$$D_f(x^{k+1}, x^k) + D_f(z, x^{k+1}) \leq D_f(z, x^k). \tag{30}$$

From this inequality we deduce that the sequence $\{D_f(z, x^k)\}_{k\in\mathbb{N}}$ is nonincreasing. By consequence, for each $k \in \mathbb{N}$, we have $D_f(z, x^k) \leq D_f(z, x^0)$, showing that the sequence $\{x^k\}_{k\in\mathbb{N}}$ is included in the level set $R^f_\alpha(z)$ with $\alpha = D_f(z, x^0)$. Therefore, $\{x^k\}_{k\in\mathbb{N}}$ is bounded by $(B4)$. Since the space B is reflexive, it follows that $\{x^k\}_{k\in\mathbb{N}}$ has weak accumulation points.

Summing up the inequalities in (30) for all $k \in \mathbb{N}$ we deduce that

$$\sum_{k=0}^{\infty} D_f(x^{k+1}, x^k) < \infty,$$

and, thus,

$$\lim_{k\to\infty} D_f(x^{k+1}, x^k) = 0. \tag{31}$$

Since f satisfies $(B5)$ because it is a Bregman function and since $\{x^k\}_{k\in\mathbb{N}}$ is bounded, it results that

$$\lim_{k\to\infty} \left\| x^{k+1} - x^k \right\| = 0, \tag{32}$$

Denote $\Omega_k := \Omega(x^k)$. According to (24),

$$\Psi(x^k, x^k) = \int_{\Omega_k} g(\omega, x^k) d\mu(\omega) \geq 0. \tag{33}$$

From (26) it results that $x^{k+1} \in K(x^k)$ and, according to (25), this implies

$$\Psi(x^{k+1}, x^k) = \int_{\Omega_k} g(\omega, x^{k+1}) d\mu(\omega) \leq 0.$$

Hence, if $s > 0$ is an upperbound of the bounded sequence $\left\{\left\|x^k\right\|\right\}_{k\in\mathbb{N}}$, then

$$0 \leq \int_{\Omega_k} g(\omega, x^k) d\mu(\omega) \leq \int_{\Omega_k} g(\omega, x^k) d\mu(\omega) - \int_{\Omega_k} g(\omega, x^{k+1}) d\mu(\omega)$$

$$= \left| \int_{\Omega_k} [g(\omega, x^k) - g(\omega, x^{k+1})] d\mu(\omega) \right| \leq \int_{\Omega_k} \left| g(\omega, x^k) - g(\omega, x^{k+1}) \right| d\mu(\omega)$$

$$\leq h_1(s, \frac{1}{2} \cdot \left\|x^k - x^{k+1}\right\|) \cdot \int_{\Omega_k} h_2(\omega) d\mu(\omega)$$

$$\leq h_1(s, \frac{1}{2} \cdot \left\|x^k - x^{k+1}\right\|) \cdot \int_{\Omega} h_2(\omega) d\mu(\omega),$$

where the fourth inequality holds because of condition (II). Letting here $k \to \infty$ and taking into account the condition (II) again, we obtain

$$\lim_{k\to\infty} \int_{\Omega_k} g(\omega, x^k) d\mu(\omega) = 0. \tag{34}$$

Now, we complete the proof of (i) by showing that, if x^* is a weak accumulation point of $\left\{x^k\right\}_{k\in\mathbb{N}}$, then $x^* \in \bigcap_\mu Q_\omega$. In order to prove this claim, let

$$\Omega_* = \{\omega \in \Omega : g(\omega, x^*) > 0\}.$$

We aim at showing that $\mu(\Omega_*) = 0$. To this end, for any nonnegative integer k and for any real number $\varepsilon > 0$, denote

$$\Gamma^k(\varepsilon) = \left\{\omega \in \Omega : g(\omega, x^k) > \varepsilon\right\}.$$

Since $\Gamma^k(\varepsilon) \subseteq \Omega_k$, and since, for $\omega \in \Omega_k$, $g(\omega, x^k) > 0$, we have

$$0 \leq \mu(\Gamma^k(\varepsilon)) = \frac{1}{\varepsilon} \cdot \int_{\Gamma^k(\varepsilon)} \varepsilon d\mu(\omega)$$

$$\leq \frac{1}{\varepsilon} \cdot \int_{\Gamma^k(\varepsilon)} g(\omega, x^k) d\mu(\omega) \leq \frac{1}{\varepsilon} \cdot \int_{\Omega_k} g(\omega, x^k) d\mu(\omega).$$

Letting here $k \to \infty$ and taking into account (34), we deduce that, for any $\varepsilon > 0$,

$$\lim_{k\to\infty} \mu(\Gamma^k(\varepsilon)) = 0. \tag{35}$$

Let $\left\{x^{k_p}\right\}_{p\in\mathbb{N}}$ be a subsequence of $\left\{x^k\right\}_{k\in\mathbb{N}}$ which converges weakly to x^*. Let $\Delta^p(\varepsilon) = \bigcap_{j\geq p} \Gamma^{k_j}(\varepsilon)$. For any $j \geq p$, we have

$$0 \leq \mu(\Delta^p(\varepsilon)) \leq \mu(\Gamma^{k_j}(\varepsilon)).$$

This combined with (35) imply

$$0 \leq \mu(\Delta^p(\varepsilon)) \leq \lim_{j\to\infty} \mu(\Gamma^{k_j}(\varepsilon)) = 0,$$

that is $\mu(\Delta^p(\varepsilon)) = 0$ for any $p \in \mathbb{N}$ and $\varepsilon > 0$.

Denote

$$\Gamma^*(\varepsilon) = \{\omega \in \Omega : g(\omega, x^*) > \varepsilon\}.$$

For any $\omega \in \Omega$, the function $g(\omega, \cdot)$ is weakly lower semicontinuous because it is convex and continuous. Thus, if $\omega \in \Gamma^*(\varepsilon)$, then

$$\varepsilon < g(\omega, x^*) \leq \underline{\lim}_{\, p\to\infty} g(\omega, x^{k_p}).$$

Therefore, if $\omega \in \Gamma^*(\varepsilon)$, then there exists a positive integer $r = r(\omega)$ such that, for any integer $p \geq r$, we have $\frac{\varepsilon}{2} < g(\omega, x^{k_p})$, that is

$$\omega \in \bigcap_{p\geq r} \Gamma^{k_p}(\frac{\varepsilon}{2}) = \Delta^r(\frac{\varepsilon}{2}).$$

Hence, we have

$$\Gamma^*(\varepsilon) \subseteq \bigcup_{p=0}^{\infty} \Delta^p(\frac{\varepsilon}{2}),$$

where the set on the right-hand side has null measure μ as being a countable union of measurable sets whose measure μ is zero. Consequently, for any $\varepsilon > 0$, we have $\mu(\Gamma^*(\varepsilon)) = 0$. Observe that

$$\Omega_* \subseteq \bigcup_{n=1}^{\infty} \Gamma^*(\frac{1}{n}).$$

This shows that $\mu(\Omega_*) = 0$ because the set on the right-hand side is a countable union of sets of null measure μ and, thus, (i) is proven.

(ii) Assume that the Gâteaux derivative $f' : B \to B^*$ is sequentially weakly-to-weak* continuous and suppose, by contradiction, that the sequence $\left\{x^k\right\}_{k\in\mathbb{N}}$ has two different weak accumulation points x' and x''. According to (i), x' and x'' belong to $\bigcap_\mu Q_\omega$. Therefore, by applying (30), we deduce that the sequences $\left\{D_f(x', x^k)\right\}_{k\in\mathbb{N}}$ and $\left\{D_f(x'', x^k)\right\}_{k\in\mathbb{N}}$ are nonincreasing and bounded, hence convergent. Denote

$$a = \lim_{k\to\infty} \left[D_f(x', x^k) - D_f(x'', x^k)\right].$$

Clearly,

$$a = f(x') - f(x'') + \lim_{k\to\infty} f'(x^k, x'' - x').$$

If $\left\{x^{k_p}\right\}_{p\in\mathbb{N}}$ and $\left\{x^{k_q}\right\}_{q\in\mathbb{N}}$ are subsequences of $\left\{x^k\right\}_{k\in\mathbb{N}}$ which converge weakly to x' and x'', respectively, then

$$a = f(x') - f(x'') + \lim_{p\to\infty} f'(x^{k_p}, x'' - x') = D_f(x', x'') \geq 0.$$

because the Gâteaux derivative $f' : B \to B^*$ is sequentially weakly continuous. On the other hand,

$$a = f(x') - f(x'') + \lim_{q\to\infty} f'(x^{k_p}, x'' - x') = -D_f(x'', x') \leq 0.$$

Hence, $a = 0$, that is $D_f(x', x'') = 0$. This implies $\|x' - x''\| = 0$, because f is a Bregman function. Thus, we obtain $x' = x''$, and this is a contradiction. By consequence, $\left\{x^k\right\}_{k\in\mathbb{N}}$ has an unique weak accumulation point, i.e., $\left\{x^k\right\}_{k\in\mathbb{N}}$ converges weakly to some point x^* which necessarily belongs to $\bigcap_\mu Q_\omega$.

Now, assume that, in addition to the sequentially weak-to-weak* continuity of f, the inequality (28) is also satisfied. Observe that, for any nonnegative integers m, k,

$$\sum_{i=0}^{m} D_f(x^{k+i+1}, x^{k+i}) = f(x^{k+m+1}) - f(x^k) - \sum_{i=0}^{m} f'(x^{k+i}, x^{k+i+1} - x^{k+i})$$

$$= D_f(x^{k+m+1}, x^k) + f'(x^k, x^{k+m+1} - x^k) + f'(x^{k+m+1}, x^k - x^{k+m+1})$$

$$- \left[\sum_{i=0}^{m} f'(x^{k+i}, x^{k+i+1} - x^{k+i}) + f'(x^{k+m+1}, x^k - x^{k+m+1})\right].$$

The sum between square brackets is nonpositive because it is a $(m+1)$-cyclic sum associated to the convex function f (see [29, p. 26]). Thus, letting $m \to \infty$, we obtain

$$\sum_{i=0}^{\infty} D_f(x^{k+i+1}, x^{k+i}) \geq$$

$$\underline{\lim}_{\ m\to\infty} \left[D_f(x^{k+m+1}, x^k) + f'(x^k, x^{k+m+1} - x^k) + f'(x^{k+m+1}, x^k - x^{k+m+1})\right]$$

Since $D_f(\cdot, x^k)$ is convex, it is weakly lower semicontinuous and, therefore, the last inequality implies

$$\sum_{i=0}^{\infty} D_f(x^{k+i+1}, x^{k+i}) \geq \underline{\lim}_{\ m\to\infty} D_f(x^{k+m+1}, x^k) +$$

$$\underline{\lim}_{\ m\to\infty} f'(x^k, x^{k+m+1} - x^k) + \underline{\lim}_{\ m\to\infty} f'(x^{k+m+1}, x^k - x^{k+m+1}) \geq$$

$$D_f(x^*, x^k) + f'(x^k, x^* - x^k) + \underline{\lim}_{\ m\to\infty} f'(x^{k+m+1}, x^k - x^{k+m+1}).$$

Consequently, the weak sequential continuity of f' and (28) yield

$$\sum_{i=0}^{\infty} D_f(x^{k+i+1}, x^{k+i}) \geq D_f(x^*, x^k) + f'(x^k, x^* - x^k) +$$

$$\underline{\lim}_{\ m\to\infty} f'(x^{k+m+1}, x^k - x^*) + \underline{\lim}_{\ m\to\infty} f'(x^{k+m+1}, x^* - x^{k+m+1}) \geq$$

$$D_f(x^*, x^k) + f'(x^k, x^* - x^k) + f'(x^*, x^k - x^*).$$

The left-hand side of this inequality converges to zero as $k \to \infty$ because it is the reminder of a convergent series. The sequence $\left\{D_f(x^*, x^k)\right\}_{k\in\mathbb{N}}$ converges nonincreasingly because $x^* \in \bigcap_\mu Q_\omega$. Therefore, letting $k \to \infty$ in the last inequality, we obtain

$$0 \geq \overline{\lim}_{k\to\infty} \left[D_f(x^*, x^k) + f'(x^k, x^* - x^k) + f'(x^*, x^k - x^*) \right] =$$

$$\lim_{k\to\infty} \left[D_f(x^*, x^k) + f'(x^*, x^k - x^*) \right] + \overline{\lim}_{k\to\infty} f'(x^k, x^* - x^k)$$

$$= \lim_{k\to\infty} D_f(x^*, x^k) + \overline{\lim}_{k\to\infty} f'(x^k, x^* - x^k).$$

Applying (28) again we deduce

$$0 \geq -\underline{\lim}_{\ k\to\infty} f'(x^k, x^* - x^k) \geq$$

$$-\overline{\lim}_{k\to\infty} f'(x^k, x^* - x^k) \geq \lim_{k\to\infty} D_f(x^*, x^k) \geq 0.$$

Consequently,

$$\lim_{k\to\infty} D_f(x^*, x^k) = 0.$$

Since the sequence $\left\{x^k\right\}_{k\in\mathbb{N}}$ is bounded and since f is a Bregman function, this implies $\lim_{k\to\infty} \left\|x^* - x^k\right\| = 0$, showing that, when (28) holds, the sequence $\left\{x^k\right\}_{k\in\mathbb{N}}$ converges strongly to x^*.

Now, assume that $\left\{x^k\right\}_{k\in\mathbb{N}}$ converges strongly to x^*. For each $k \in \mathbb{N}$, we have

$$\left|f'(x^k, x^k - x^*)\right| \leq \left\|f'(x^k, \cdot)\right\|_* \cdot \left\|x^k - x^*\right\|.$$

The function $f' : B \to B^*$ is norm-to-weak* continuous (see [29, Proposition 2.8]) and, therefore, the sequence $\left\{\left\|f'(x^k, \cdot)\right\|\right\}_{k\in\mathbb{N}}$ is bounded. This shows that (28) holds with equality. The proof is complete.

4.4 A typical situations where Theorem 4.3 applies is that in which $B = \ell^p$ and $f(x) = \|x\|^p$, with $p \in (1, +\infty)$. In this case, the Gâteaux derivative f' is given by (15) and, therefore, it is sequentially weakly-to-weak* continuous (cf. [8, Proposition 8.2]). According to Proposition 3.4 combined with Theorem 4.3 we have the following result which is of use in Section 6.6 below:

Corollary. *Suppose that $B = \ell^p$ with $p \in (1, +\infty)$ and that $f(x) = \|x\|^p$. If $g : \Omega \times B \to \mathbb{R}$ is a convex Carathéodory function satisfying the conditions (I) and (II), then any sequence $\left\{x^k\right\}_{k\in\mathbb{N}}$ generated in B by the Algorithm A converges weakly to a point $x^* \in B$ satisfying $g(\omega, x^*) \leq 0$, a.e., provided that such a point exists.*

5 The Algorithm B

5.1 In Section 3 we have shown that solutions of SCFPs emerging from stochastic systems of convex inequalities can be weakly approximated by successive Bregman projections of the current iterate onto a set defined by averaging the restrictions the current iterate violates. However, the convergence of Algorithm A is proven under restrictive technical conditions on the given SCFP. This fact leads to the question of whether we can find alternative methods of solving SCFPs which will eventually work under different or less demanding requirements on the given problem. In what follows we study a method of solving SCFPs which we call Algorithm B. The basic idea of Algorithm B consists of recursively computing the "average" of the Bregman projections of the current iterate onto the sets Q_ω, $\omega \in \Omega$. When one analyses this algorithm one should observe that its convergence can be ensured without special requirements (like conditions (I) and (II) above) concerning the nature of the sets Q_ω. Also, by contrast to Algorithm A, the sequences generated by Algorithm B may converge weakly even in cases when the given SCFP is inconsistent. However, there is a price to be paid for these advantages of Algorithm B upon Algorithm A: In order to guarantee convergence of Algorithm B the function f with respect to which the Bregman projections we average are computed is subjected to conditions which sensibly restrict the pool from which such functions can be chosen.

5.2 For making these conditions precise, we call a convex function $f : B \to (-\infty, +\infty]$ *regularly convex on the subset* $K \subseteq \mathcal{D}^\circ$ if there exists a function $\eta : K \times [0, +\infty) \to [0, +\infty)$ such that, for each $x \in K$,

(i) The function $\eta(x, \cdot)$ is continuous, has $\eta(x, 0) = 0$, $\eta(x, 1) > 0$ and, for any $t > 0$,

$$c \geq 1 \Rightarrow \eta(c \cdot t) \geq c \cdot \eta(t); \tag{36}$$

(ii) If $y \in \mathcal{D}^\circ$, then

$$D_f(y, x) \geq \eta(x, \|y - x\|). \tag{37}$$

There are significant functions which are regularly convex. Some of them are described by the next result.

Lemma. *Let $f : B \to (-\infty, +\infty]$ be a convex function which satisfies the condition $(B1)$. Then the following statements hold:*

(a) *If the function f is very convex, then f is regularly convex on the set of the points $x \in \mathcal{D}^\circ$ such that $\nu_f^+(x,\cdot)$ is continuous from the left. In particular, if B is of finite dimension, then any differentiable, strictly convex real valued function on B is regularly convex on $\mathcal{D}^\circ$.*

(b) *If the function f is totally convex, then f is regularly convex on the set of the points $x \in \mathcal{D}^\circ$ such that $\nu_f^-(x,\cdot)$ is continuous. In particular, if $B = \mathcal{L}^p([a,b])$ or $B = \ell^p$ and $f(x) = \|x\|^p$ with $p \in (1,2)$, then f is regularly convex on B.*

(c) *If the function f is uniformly convex and has continuous modulus of convexity δ_f, then f is regularly convex on $\mathcal{D}^\circ$. In particular, if $B = \mathcal{L}^p([a,b])$ or $B = \ell^p$ with $p \geq 2$ and if $f(x) = \|x\|^p$, then f is regularly convex on B.*

(d) *The negentropy is regularly convex on $\mathbb{R}^n_{++}$.*

Proof. (a) Take $\eta = \nu_f^+$. Application of [12, Proposition 2.4] shows that η satisfies (i) and (ii) above. This and Lemma 2.5 complete the proof in this case.

(b) Take $\eta = \nu_f^-$ and apply (8) and Proposition 3.4.

(c) Take $\eta(x,t) = \delta_f(t)$ and apply (14) and Proposition 3.2.

(d) This follows from Proposition 3.3 by observing that the right hand side of (17), which is the required function η, is the minimum of a family of functions which satisfy (8) because they are local moduli of convexity associated to the function g defined by (18).

5.3 As noted above, the Algorithm B is fit to approach SCFPs in their most general form. Precisely, we consider the SCFP as given in Section 1.1. We chose the function $f : B \to (-\infty, +\infty]$ to be a Bregman function which is Fréchet differentiable on the interior of its domain and satisfies the following conditions of compatibility with the given SCFP:

($C1$) *The interior $\mathcal{D}^\circ$ of the set $\mathcal{D} = Dom(f)$ covers the closed convexification of the range of the point-to-set mapping Q, that is,*

$$S := Cl\left[conv \bigcup_{\omega\in\Omega} Q_\omega\right] \subset \mathcal{D}^\circ \tag{38}$$

($C2$) *If the set Ω is not finite, then either the set $\bigcup_{\omega\in\Omega} Q_\omega$ is bounded or $\bigcup_{\omega\in\Omega} Q_\omega$ is unbounded but, then, the function f is regularly convex on $\mathcal{D}^\circ$.*

$(C3)$ *There exists a measurable selector* $\phi : \Omega \to B$ *of the point-to-set mapping* Q *such that, for any* $x \in \mathcal{D}^\circ$, *the function* $\omega \to D_f(\phi(\omega), x) : \Omega \to [0, +\infty)$ *is integrable.*

$(C4)$ *The restriction of* D_f *to the convex subset* $S \times S$ *of the product Banach space* $B \times B$ *is convex.*

If the given SCFP is consistent and if the condition $(C1)$ is satisfied, then $(C3)$ holds because, in this case, if $z \in \bigcap_\mu Q_\omega$, then the constant function $\phi(\omega) = z$ is a measurable selector of Q such that $D_f(\phi(\cdot), x)$ is integrable for any $x \in \mathcal{D}^\circ$.

A very demanding requirement is the condition $(C4)$. This condition is essentially involved in our proof of convergence for the Algorithm B and considerably restricts the pool of Bregman functions f which can be used in order to solve a priori given SCFPs. In spite of that, there are several significant Bregman functions which fulfil the conditions $(C1)$-$(C4)$ and, thus, can be used to solving SCFPs via the Algorithm B. A typical example of Bregman function which satisfies the conditions $(C1)$, $(C2)$ and $(C4)$ no matter how the SCFP is defined, is $f(x) = \|x\|^2$ when B is a Hilbert space (see Remark 3.5 and Lemma 5.2(c)). In this particular case, the condition $(C3)$ amounts to the requirement that the point-to-set mapping Q should have a square-integrably bounded selector. For SCFPs in $B = \mathbb{R}^n$ such that $S \subset \mathbb{R}^n_{++}$ the conditions $(C1)$, $(C2)$ and $(C4)$ hold with respect to the negentropy (see Proposition 3.3 and Lemma 5.2(d)). In this case, for satisfying $(C3)$ it is sufficient for the mapping Q to have an integrable selector. If $p \in (1, +\infty)$, if $B = \mathbb{R}^n$ provided with the ℓ_p-norm, or $B = \ell^p$, or $B = \mathcal{L}^p([a,b])$, and if $f(x) = \|x\|^p$, then the conditions $(C1)$ and $(C2)$ are automatically satisfied for any SCFP. In this case, the condition $(C3)$ amounts to the requirement that Q has a selector which is $p-$integrable. If $p \in (1,2)$, then the condition $(C4)$ is satisfied whenever S is a subset of the positive orthant (cf. Section 6.2 below).

5.4 According to Lemma 2.2, the fact that f is a Bregman function ensures that, for each $\omega \in \Omega$, the Bregman projection $\Pi^f_\omega(x) := \Pi^f_{Q_\omega}(x)$ exists and is unique whenever $x \in \mathcal{D}^\circ$. Hence, if $x \in \mathcal{D}^\circ$, then the function $\Pi^f_\bullet(x) : \omega \to \Pi^f_\omega(x) : \Omega \to B$ is well-defined. Note that, according to Theorem 8.2.11 in [3], the function $\Pi^f_\bullet(x)$ is measurable. A question essentially related with the well-definiteness of the Algorithm B is whether, or under what conditions, this function is (Bochner) integrable for every $x \in \mathcal{D}^\circ$. The next result shows that, in our setting, the function $\Pi^f_\bullet(x)$ in integrable.

Lemma. *Suppose that the conditions* $(C1)$, $(C2)$ *and* $(C3)$ *are satisfied. Then, for each* $x \in \mathcal{D}^\circ$, *the function* $\Pi^f_\bullet(x) : \Omega \to B$, defined by $\Pi^f_\omega(x) = \Pi^f_{Q_\omega}(x)$ *is* (*Bochner*) *integrable*[2] *and its integral*

$$\mathbf{\Pi}^f(x) := \int_\Omega \Pi^f_\omega(x) d\mu(\omega) \tag{39}$$

belongs to the set S.

Proof. The proof is made in several stages.

(I) We show that, whenever $\mathbf{\Pi}^f(x)$ exists, it belongs to S. It was observed above that the function $\Pi^f_\bullet(x)$ is well-defined and measurable. Note that, whenever it exists, the (Bochner) integral $\mathbf{\Pi}^f(x)$ is the limit of a sequence of points in B in which each term is of the form $\sum_{j=1}^p \mu(A_j) \cdot \Pi^f_{\omega_j}(x)$, where $A_1, ..., A_p$ is a measurable partition of Ω and $\omega_j \in A_j$, for $j = 1, ..., p$. This shows that $\mathbf{\Pi}^f(x)$ is the limit of a sequence of convex combinations of points in $\bigcup_{\omega\in\Omega} Q_\omega$, i.e., whenever it exists, $\mathbf{\Pi}^f(x) \in S$.

(II) Suppose that $\bigcup_{\omega\in\Omega} Q_\omega$ is bounded and that M is an upperbound of it. Then, M is an upperbound of S too. Since the function $\omega \to \Pi^f_\omega(x)$ is measurable, so is the function $\omega \to \left\|\Pi^f_\omega(x)\right\| : \Omega \to \mathbb{R}_+$. Therefore, the integral $\int_\Omega \left\|\Pi^f_\omega(x)\right\| d\mu(\omega)$ exists. Clearly, since, for each $\omega \in \Omega$, the Bregman projection $\Pi^f_\omega(x)$ belongs to S, we have

$$\int_\Omega \left\|\Pi^f_\omega(x)\right\| d\mu(\omega) \leq M,$$

showing that the function $\omega \to \left\|\Pi^f_\omega(x)\right\|$ is integrable and, hence, the function $\omega \to \Pi^f_\omega(x)$ is also integrable.

(III) Assume that $\bigcup_{\omega\in\Omega} Q_\omega$ is unbounded. Then, according to $(C2)$, the Bregman function f is regularly convex. Let $\eta : \mathcal{D}^\circ \times [0, +\infty) \to [0, +\infty)$ be the function associated to f which satisfies the conditions (i) and (ii) in Section 5.2. We prove the following

[2] In what follows integrals and integrability for functions from Ω to the Banach space B are in the sense of Bochner (see [27]).

Claim. For any $x \in \mathcal{D}^\circ$, the function $\eta(x,\cdot)$ is strictly increasing on $[0,+\infty)$ and has

$$\lim_{t\to\infty} \eta(x,t) = +\infty. \tag{40}$$

In order to prove that $\eta(x,\cdot)$ is strictly increasing, let $s,t \in (0,+\infty)$ such that $s > t$. Then, according to (36), we get

$$\eta(x,s) = \eta(x, \frac{s}{t}\cdot t) \geq \frac{s}{t}\cdot \eta(x,t) > \eta(x,t).$$

For proving (40), consider a sequence $\{t_k\}_{k\in\mathbb{N}}$ in $[0,+\infty)$ which converges nondecreasingly to $+\infty$. Then, there exists a positive integer k_0 such that, for any $k \geq k_0$, we have $t_k \geq 1$. Hence, for $k \geq k_0$, we have $\eta(x,t_k) \geq t_k \cdot \eta(x,1)$, where $\eta(x,1) > 0$. This shows that $\lim_{k\to\infty} \eta(x,t_k) = +\infty$ and the Claim is proved.

Fix $x \in \mathcal{D}^\circ$. From the condition (i) in Section 5.2 and the proved Claim it results that $\eta(x,\cdot)$ is continuous and strictly increasing on $[0,+\infty)$. Hence, $\eta(x,\cdot)$ is invertible. Let $\eta^{-1}(x,\cdot) : [0,+\infty) \to [0,+\infty)$ be its inverse. Clearly, the function $\eta^{-1}(x,\cdot)$ is continuous and strictly increasing. It also has the property that,

$$c \geq 1 \Rightarrow \eta^{-1}(x, c\cdot t) \leq c \cdot \eta^{-1}(x,t). \tag{41}$$

Indeed, by (36), for $c \geq 1$ and $s \geq 0$ we have

$$\eta^{-1}(x, \eta(x, c\cdot s)) \geq \eta^{-1}(x, c\cdot \eta(x,s)).$$

This implies

$$c\cdot s \geq \eta^{-1}(x, c\cdot \eta(x,s)).$$

Letting here $s = \eta^{-1}(x,t)$ we obtain (41).

Let $\phi : \Omega \to B$ be the selector of the mapping Q whose existence follows from the condition $(C3)$. According to (37) we have

$$\eta(x, \left\|\Pi_\omega^f(x) - x\right\|) \leq D_f(\Pi_\omega^f(x), x) \leq D_f(\phi(\omega), x), \quad a.e.$$

Hence,

$$\left\|\Pi_\omega^f(x) - x\right\| \leq \eta^{-1}(x, D_f(\Pi_\omega^f(x), x)) \leq \eta^{-1}(x, D_f(\phi(\omega), x)), \quad a.e.$$

The functions $\omega \to \left\| \Pi_\omega^f(x) - x \right\|$ and $\omega \to \eta^{-1}(x, D_f(\phi(\omega), x))$ are nonnegative and measurable. Thus, the following integrals exist and

$$\int_\Omega \left\| \Pi_\omega^f(x) - x \right\| d\mu(\omega) \leq \int_\Omega \eta^{-1}(x, D_f(\phi(\omega), x)) d\mu(\omega). \tag{42}$$

Since

$$\int_\Omega \left\| \Pi_\omega^f(x) \right\| d\mu(\omega) \leq \|x\| + \int_\Omega \left\| \Pi_\omega^f(x) - x \right\| d\mu(\omega), \tag{43}$$

if we show that

$$I := \int_\Omega \eta^{-1}(x, D_f(\phi(\omega), x)) d\mu(\omega) < +\infty, \tag{44}$$

then the existence of $\mathbf{\Pi}^f(x)$ will result from (42). In order to prove (44), recall that the function f is locally Lipschitz on $\mathcal{D}^\circ$ because it is convex and continuous (see $(B1)$ and $(B2)$) on this open set. Consequently, there exist two positive real numbers $r = r_f(x)$ and $L = L_f(x)$ such that, for each $y \in \mathcal{D}^\circ$ with $\|y - x\| < r$, we have

$$|f(x) - f(y)| \leq L \cdot \|x - y\|. \tag{45}$$

Denote

$$\Gamma = \{\omega \in \Omega : \|x - \phi(\omega)\| < r\}$$

and $\Lambda = \Omega \backslash \Gamma$. We have

$$I = \int_\Gamma \eta^{-1}(x, D_f(\phi(\omega), x)) d\mu(\omega) + \int_\Lambda \eta^{-1}(x, D_f(\phi(\omega), x)) d\mu(\omega). \tag{46}$$

Observe that, if $\omega \in \Gamma$, then

$$D_f(\phi(\omega), x) \leq |f(\phi(\omega)) - f(x)| + |f'(x, \phi(\omega) - x)| \leq$$

$$2 \cdot L \cdot \|x - \phi(\omega)\| \leq 2 \cdot L \cdot r,$$

because of (45) and of [19, Proposition 2.1.1] which shows that the norm of $f'(x, \cdot)$ in B^* satisfies $\|f'(x, \cdot)\|_* \leq L$. Hence, if $\omega \in \Gamma$, then

$$\eta^{-1}(x, D_f(\phi(\omega), x)) \leq \eta^{-1}(x, 2 \cdot L \cdot r),$$

showing that the first integral on the right hand of the inequality (46) is finite. For each $\omega \in \Lambda$, denote

$$c(\omega) = \frac{D_f(\phi(\omega), x)}{\eta(x, \|x - \phi(\omega)\|)}. \tag{47}$$

If $\omega \in \Lambda$, then $c(\omega) \geq 1$ because of (37). Therefore,

$$\begin{aligned} &\textstyle\int_\Lambda \eta^{-1}\left(x, D_f(\phi(\omega), x)\right) d\mu(\omega) = \int_\Lambda \eta^{-1}\left(x, c(\omega) \cdot \eta(x, \|x - \phi(\omega)\|)\right) d\mu(\omega) \\ &\leq \textstyle\int_\Lambda c(\omega) \cdot \eta^{-1}\left(x, \eta(x, \|x - \phi(\omega)\|)\right) d\mu(\omega) = \int_\Lambda c(\omega) \cdot \|x - \phi(\omega)\| \, d\mu(\omega), \end{aligned} \tag{48}$$

where the inequality follows from (41). Applying (36) we deduce that, if $\omega \in \Lambda$, then

$$\eta(x, \|x - \phi(\omega)\|) = \eta(x, \frac{\|x - \phi(\omega)\|}{r} \cdot r) \geq \frac{\|x - \phi(\omega)\|}{r} \cdot \eta(x, r),$$

and this, combined with (36) and with (47), gives

$$\int_\Lambda \eta^{-1}\left(x, D_f(\phi(\omega), x)\right) d\mu(\omega) \leq \int_\Lambda \frac{D_f(\phi(\omega), x)}{\frac{\|x - \phi(\omega)\|}{r} \cdot \eta(x, r)} \cdot \|x - \phi(\omega)\| \, d\mu(\omega)$$

$$= \frac{r}{\eta(x, r)} \int_\Lambda D_f(\phi(\omega), x) d\mu(\omega) < +\infty,$$

which shows that the second integral in (46) is finite. Hence, I is finite, that is, the function $\Pi^f_\bullet(x)$ is integrable.

5.5 Lemma 5.4 ensures that, if the conditions ($C1$), ($C2$) and ($C3$) hold, then the operator $\mathbf{\Pi}^f : \mathcal{D}^\circ \to \mathcal{D}^\circ$ given by (39) is well-defined. Therefore, the sequences recursively defined by the following procedure exist no matter how their initial points are chosen in $\mathcal{D}^\circ$.

Algorithm B. *Choose the initial point $x^0 \in \mathcal{D}^\circ$ and, for each nonnegative integer k, let*

$$x^{k+1} = \int_\Omega \Pi^f_\omega(x^k) d\mu(\omega). \tag{49}$$

In what follows we show that, under the conditions ($C1$)-($C4$), the sequences generated by the Algorithm B converge weakly to fixed points of $\mathbf{\Pi}^f$ whenever the function $\Upsilon_f : \mathcal{D}^\circ \to [0, \infty]$ given by

$$\Upsilon_f(x) = \int_\Omega D_f(\Pi^f_\omega(x), x) d\mu(\omega), \tag{50}$$

has a global minimizer and $f' : B \to B^*$ is sequentially weakly-to-weak* continuous. Note that Υ_f is well-defined because the function $\omega \to D_f(\Pi^f_\omega(x), x) : \Omega \to [0, +\infty)$ is measurable (it is the composition of the continuous function $D_f(\cdot, x)$ with the measurable function $\Pi^f_\bullet(x)$). The function Υ_f is relevant in our context because, when the given SCFP is consistent, Υ_f vanishes at almost common points of the sets Q_ω, $\omega \in \Omega$, only. This shows that solving consistent SCFPs in our setting amounts to finding global minimizers of the function Υ_f, that is, points in the set *Argmin*(Υ_f). A fact first emphasized in [24] for the particular case of SCFPs in $B = \mathbb{R}^n$ and involving finite sets Ω while $f(x) = \|x\|^2$ is that, even if the given SCFP is not consistent, but *Argmin*(Υ_f)=$\emptyset$, the sequences generated by Algorithm B, whenever they converge weakly, have their weak limits in *Argmin*(Υ_f) and that the points in *Argmin*(Υ_f) are exactly the fixed points of the operator $\mathbf{\Pi}^f : \mathcal{D}^\circ \to \mathcal{D}^\circ$ defined by (39). This idea was further developed in [10] where finiteness of Ω was dropped and B was assumed to be a Hilbert space. The convergence analysis of Algorithm B we present below shows that this is a more general property of sequences generated by the Algorithm B.

5.6 The next result summarizes the properties of the function Υ_f which are of use in our study of Algorithm B and, mainly, the relationship between the set *Argmin*(Υ_f) of the global minimizers of Υ_f and the set *Fix*$(\mathbf{\Pi}^f)$ of the fixed points of the operator $\mathbf{\Pi}^f$.

Lemma. *If the conditions* $(C1)$-$(C4)$ *are satisfied, then*

(i) *The function* Υ_f *is everywhere finite and convex on* $\mathcal{D}^\circ$;

(ii) *For each* $x \in \mathcal{D}^\circ$,

$$\Upsilon_f(\mathbf{\Pi}^f(x)) + D_f(\mathbf{\Pi}^f(x), x) \leq \Upsilon_f(x); \tag{51}$$

(iii) *If* $z \in$*Fix*$(\mathbf{\Pi}^f)$ *and if* $x \in \mathcal{D}^\circ$, *then,*

$$D_f(z, \mathbf{\Pi}^f(x)) + \Upsilon_f(x) \leq D_f(z, x) + \Upsilon_f(z); \tag{52}$$

(iv) *If Argmin*(Υ_f)=$\emptyset$, *then Argmin*$(\Upsilon_f) =$ *Fix*$(\mathbf{\Pi}^f)$.

Proof. (i) Let $x \in \mathcal{D}^\circ$ and let ϕ be the measurable selector of the point-to-set mapping Q whose existence follows from the condition $(C3)$. Then, $D_f(\Pi_\omega^f(x), x) \leq D_f(\phi(\omega), x)$, a.e., and, thus,

$$\Upsilon_f(x) \leq \int_\Omega D_f(\phi(\omega), x) d\mu(\omega) < +\infty.$$

Now, in order to prove that Υ_f is convex, let $x, y \in \mathcal{D}^\circ$ and $\alpha \in [0,1]$. Then,

$$\Upsilon_f(\alpha \cdot x + (1-\alpha) \cdot y) =$$

$$\int_\Omega D_f(\Pi_\omega^f(\alpha \cdot x + (1-\alpha) \cdot y), \alpha \cdot x + (1-\alpha) \cdot y) d\mu(\omega) \leq$$

$$\int_\Omega D_f(\alpha \cdot \Pi_\omega^f(x) + (1-\alpha) \cdot \Pi_\omega^f(y), \alpha \cdot x + (1-\alpha) \cdot y) d\mu(\omega) \leq$$

$$\alpha \cdot \Upsilon_f(x) + (1-\alpha) \cdot \Upsilon_f(y),$$

where the first inequality follows from the definition of the Bregman projection $\Pi_\omega^f(\alpha \cdot x + (1-\alpha) \cdot y)$ and the second inequality results from $(C4)$.

(ii) Observe that, for any $\omega \in \Omega$,

$$D_f(\Pi_\omega^f(\mathbf{\Pi}^f(x)), \mathbf{\Pi}^f(x)) \leq D_f(\Pi_\omega^f(x), \mathbf{\Pi}^f(x)) =$$

$$D_f(\Pi_\omega^f(x), x) + D_f(x, \mathbf{\Pi}^f(x)) + \left[f'(x, \Pi_\omega^f(x) - x) - f'(\mathbf{\Pi}^f(x), \Pi_\omega^f(x) - x) \right].$$

Integrating this inequality and taking into account that, for each $u \in \mathcal{D}^\circ$, the function $f'(u, \cdot)$ is linear and continuous, we obtain

$$\Upsilon_f(\mathbf{\Pi}^f(x)) \leq \Upsilon_f(x) + D_f(x, \mathbf{\Pi}^f(x)) + f'(x, \mathbf{\Pi}^f(x) - x) - f'(\mathbf{\Pi}^f(x), \mathbf{\Pi}^f(x) - x),$$

where

$$f'(\mathbf{\Pi}^f(x), \mathbf{\Pi}^f(x) - x) - f'(x, \mathbf{\Pi}^f(x) - x) = D_f(\mathbf{\Pi}^f(x), x) + D_f(x, \mathbf{\Pi}^f(x))$$

and this proves (51).

(iii) Observe that

$$D_f(\Pi_\omega^f(x), x) + D_f(\Pi_\omega^f(z), \Pi_\omega^f(x)) - D_f(z, x) - D_f(\Pi_\omega^f(z), z) =$$

$$\left[f'(x, \Pi_\omega^f(z) - \Pi_\omega^f(x)) - f'(\Pi_\omega^f(x), \Pi_\omega^f(z) - \Pi_\omega^f(x)) \right] - \tag{53}$$

$$\left[f'(x, \Pi_\omega^f(z) - z) - f'(z, \Pi_\omega^f(z) - z) \right].$$

Recall (see [2, Proposition 2.5]) that the function $D_f(\cdot, x)$ is Gâteaux differentiable and its derivative at $\Pi_\omega^f(x)$ in the direction $\Pi_\omega^f(z) - \Pi_\omega^f(x)$ is exactly

$$f'(\Pi_\omega^f(x), \Pi_\omega^f(z) - \Pi_\omega^f(x)) - f'(x, \Pi_\omega^f(z) - \Pi_\omega^f(x)).$$

Since $\Pi_\omega^f(z) \in Q_\omega$ and $\Pi_\omega^f(x)$ is a minimizer of $D_f(\cdot, x)$ over Q_ω, this derivative is nonnegative.. Hence, the first term on the right hand side in (53) is nonpositive. Therefore,

$$D_f(\Pi_\omega^f(x), x) + D_f(\Pi_\omega^f(z), \Pi_\omega^f(x)) - D_f(z, x) - D_f(\Pi_\omega^f(z), z) \leq$$

$$f'(z, \Pi_\omega^f(z) - z) - f'(x, \Pi_\omega^f(z) - z).$$

Observe that the function D_f is continuous on $\mathcal{D}^\circ \times \mathcal{D}^\circ$ because f is Fréchet differentiable. Consequently, the function $\omega \to D_f(\Pi_\omega^f(x), \Pi_\omega^f(x))$ is measurable. Thus, integrating the inequality above, we obtain

$$\begin{gathered} \Upsilon_f(x) + \textstyle\int_\Omega D_f(\Pi_\omega^f(z), \Pi_\omega^f(x)) d\mu(\omega) - D_f(z, x) - \Upsilon_f(z) \leq \\ f'(z, \mathbf{\Pi}^f(z) - z) - f'(x, \mathbf{\Pi}^f(z) - z) = 0, \end{gathered} \tag{54}$$

where the last equality holds because $z \in Fix(\mathbf{\Pi}^f)$. Now, applying the condition $(C4)$ combined with Jensen's inequality, we get

$$\int_\Omega D_f(\Pi_\omega^f(z), \Pi_\omega^f(x)) d\mu(\omega) \geq D_f(\mathbf{\Pi}^f(z), \mathbf{\Pi}^f(x)) = D_f(z, x).$$

This and (54) implies (52).

(iv) Let z be a point in $Argmin(\Upsilon_f)$. Then, according to (51), we have

$$0 \leq D_f(\mathbf{\Pi}^f(z), z) \leq \Upsilon_f(z) - \Upsilon_f(\mathbf{\Pi}^f(z)) \leq 0,$$

that is, $D_f(\mathbf{\Pi}^f(z), z) = 0$. This cannot happen unless $z = \mathbf{\Pi}^f(z)$ because f is a Bregman function. Hence, $z \in Fix(\mathbf{\Pi}^f)$. Conversely, assume that $x \in Fix(\mathbf{\Pi}^f)$. Fix $z \in Argmin(\Upsilon_f)$. Then, according to (52), we deduce

$$D_f(z, x) + \Upsilon_f(x) - \Upsilon_f(z) = D_f(z, \mathbf{\Pi}^f(x)) + \Upsilon_f(x) - \Upsilon_f(z) \leq D_f(z, x).$$

Hence, $\Upsilon_f(x) \leq \Upsilon_f(z)$ and this shows that $x \in Argmin(\Upsilon_f)$.

5.7 The next result summarizes the convergence properties of the sequences generated by Algorithm B.

Theorem. *Let* $\{x^k\}_{k\in\mathbb{N}}$ *be a sequence generated by Algorithm B.*

(A) *If Argmin*(Υ_f)=$\emptyset$ *and if the conditions* $(C1)$, $(C2)$, $(C3)$ *and* $(C4)$ *are satisfied, then*

(i) *The sequence* $\{x^k\}_{k\in\mathbb{N}}$ *is bounded and it has weak accumulation points;*

(ii) *The weak accumulation points of* $\{x^k\}_{k\in\mathbb{N}}$ *belong to Argmin*(Υ_f) *and are fixed points of the operator* $\mathbf{\Pi}^f : \mathcal{D}^\circ \to \mathcal{D}^\circ$ *defined by* (39).

(B) *If Argmin*(Υ_f)=$\emptyset$*, if the conditions* $(C1)$, $(C2)$, $(C3)$ *and* $(C4)$ *hold and if the Gâteaux derivative,* $f' : B \to B^*$ *is sequentially weakly-to-weak** *continuous on* S*, then*

(iii) *The sequence* $\{x^k\}_{k\in\mathbb{N}}$ *converges weakly to a point* x^* *in Argmin*(Υ_f)*;*

(iv) *The sequence* $\{x^k\}_{k\in\mathbb{N}}$ *converges strongly if and only if* (28) *is satisfied.*

(C) *If the given SCFP is consistent* $(i.e., \bigsqcap_\mu Q_\omega = \emptyset)$ *and if the conditions* $(C1)$, $(C2)$ *and* $(C4)$ *are satisfied, then*

(v) *Fix*$(\mathbf{\Pi}^f)$= *Argmin*$(\Upsilon_f) = \bigsqcap_\mu Q_\omega$*;*

(vi) *The sequence* $\{x^k\}_{k\in\mathbb{N}}$ *has accumulation points and all its accumulation points are almost common points of the sets* Q_ω, $\omega \in \Omega$*;*

(vii) *If the Gâteaux derivative,* $f' : B \to B^*$ *is sequentially weakly-to-weak** *continuous on* S*, then the sequence* $\{x^k\}_{k\in\mathbb{N}}$ *converges weakly to a point* $x^* \in \bigsqcap_\mu Q_\omega$.

Proof. In order to prove (A) fix a point $z \in Argmin(\Upsilon_f)$. According to Lemma 5.6, we have $\mathbf{\Pi}^f(z) = z$ and, for each nonnegative integer k,

$$D_f(z, x^{k+1}) \le D_f(z, x^{k+1}) + \Upsilon_f(x^k) - \Upsilon_f(z) \le D_f(z, x^k). \tag{55}$$

This shows that the sequence $\{D_f(z, x^k)\}_{k\in\mathbb{N}}$ is nonincreasing and, hence, bounded by $\alpha = D_f(z, x^0)$. Therefore, $\{x^k\}_{k\in\mathbb{N}} \subset R^f_\alpha(z)$ and, thus, $\{x^k\}_{k\in\mathbb{N}}$ is bounded. Since B is reflexive, this implies that $\{x^k\}_{k\in\mathbb{N}}$ has at least one weak accumulation point and (i) is proven. Let x^* be a weak accumulation point of $\{x^k\}_{k\in\mathbb{N}}$. Note that, by (55), we have

$$0 \le \Upsilon_f(x^k) - \Upsilon_f(z) \le D_f(z, x^k) - D_f(z, x^{k+1}),$$

and taking the limit for $k \to \infty$, we get

$$\Upsilon_f(x^*) - \Upsilon_f(z) \leq \underline{\lim}_{\ k\to\infty} \left[\Upsilon_f(x^k) - \Upsilon_f(z)\right] \leq$$

$$\underline{\lim}_{\ k\to\infty} \left[D_f(z, x^k) - D_f(z, x^{k+1})\right] = 0,$$

where the first inequality holds because Υ_f is convex (see Lemma 5.6). This implies that $\Upsilon_f(x^*) \leq \Upsilon_f(z)$ and, thus, $x^* \in Argmin(\Upsilon_f) = Fix(\mathbf{\Pi}^f)$ and this proves (ii).

The proof of (B) reproduces without essential modifications the argument which proves Theorem 4.3(ii) by first observing that, whenever x^* is a weak accumulation point of $\left\{x^k\right\}_{k\in\mathbb{N}}$, the inequality (55) still holds while z is replaced by x^*. This implies that the sequence $\left\{D_f(x^*, x^k)\right\}_{k\in\mathbb{N}}$ is nonincreasing and bounded, hence, convergent.

For proving (C) note that, when the given SCFP is consistent, the condition $(C3)$ automatically holds because any constant function $\phi : \Omega \to B$ defined by $\phi(\omega) = z \in \bigcap_\mu Q_\omega$ is a measurable selector of the mapping Q. Also, if $z \in \bigcap_\mu Q_\omega$, then $\Upsilon_f(z) = 0$, that is $z \in Argmin(\Upsilon_f)$. Hence, $\bigcap_\mu Q_\omega \subseteq Argmin(\Upsilon_f)$. According to Lemma 5.6, this implies that $Argmin(\Upsilon_f) = Fix(\mathbf{\Pi}^f)$. Suppose that $y \in Argmin(\Upsilon_f)$. Then, $\Upsilon_f(y) \leq \Upsilon_f(z) = 0$, i.e., $\Upsilon_f(y) = 0$. Hence, for almost all $\omega \in \Omega$, $D_f(\Pi_\omega^f(y), y) = 0$. Since f is a Bregman function this implies that, for almost all $\omega \in \Omega$, we have $y = \Pi_\omega^f(y) \in Q_\omega$, that is $y \in \bigcap_\mu Q_\omega$ completing the proof of (v). Taking into account that the conditions $(C1)$-$(C4)$ are satisfied and that $Argmin(\Upsilon_f) = \emptyset$, we apply (B) in order to deduce that (vi) also holds. Also, (vii) results from (C) and the proof of the theorem is complete.

5.8 Remarks. (i) Theorem 5.7 generalizes and improves upon the main convergence result in [12] which shows that Algorithm B converges to an almost common point of the sets Q_ω, $\omega \in \Omega$, provided that $B = \mathbb{R}^n$, that the given SCFP is consistent and that the function $f : B \to \mathbb{R}$ is the negentropy. A question asked in [12, Section 1.5] is whether this particular convergence result can be expanded to functions other than the negentropy. As shown in Section 5.3, there are significant Bregman functions satisfying the conditions $(C1)$-$(C4)$ not only on $\mathbb{R}^n$ but also on some infinite dimensional Banach spaces too and, therefore, Theorem 5.7 is an answer to that question.

(ii) Theorem 5.7 generalizes in some respects a convergence result in [10] showing that, if B is a Hilbert space and if $f(x) = \|x\|^2$, then the sequences generated by the Algorithm B converge weakly to minimizers of Υ_f which are necessarily among the fixed points of the operator $\mathbf{\Pi}^f$. However, Theorem 2.2 in [10] says more: It shows that a sequence generated by Algorithm B in the given particular setting converges weakly if and only if *Fix(*$\mathbf{\Pi}^f$)=$\emptyset$*. This raises the question, whose answer we do not know, of whether, in general, the fact that a sequence generated by Algorithm B converges weakly implies that Argmin*(Υ_f)=$\emptyset$ *or, eventually, Fix*($\mathbf{\Pi}^f$)=$\emptyset$.

6 Comments and Applications

6.1 The aim of this section is to discuss the assumptions under which the main results presented in this work, namely Theorem 4.3 and Theorem 5.7, were proved and to emphasize situations in which these results are applicable. We start our discussion by observing that Theorems 4.3 and 5.7 ensure that the sequences generated by Algorithms A and B, respectively, accumulate weakly to solutions of the SCFPs they are supposed to solve. Convergence of these algorithms is guaranteed if the Bregman function f with respect to which the Bregman projections required by the algorithms are computed has the property that its Gâteaux derivative $f' : B \to B^*$ is sequentially weakly-to-weak* continuous. This condition is satisfied by any Bregman function when B is of finite dimension and, also, by significant Bregman functions in spaces of infinite dimension as, for instance, the function $f(x) = \|x\|^2$ when B is a Hilbert space and the functions $f(x) = \|x\|^p$ when $B = \ell^p$, where $p \in (1, +\infty)$. The following condition, suggested in [9], can be used instead of the sequentially weak-to-weak* continuity of f' in order to ensure weak convergence of the sequences generated by Algorithms A and B.

($D1$) *If* $\{z^k\}_{k\in\mathbb{N}}$ *and* $\{y^k\}_{k\in\mathbb{N}}$ *are sequences in S which converge weakly to z and y, respectively, and if z=y, then*

$$\underline{\lim}_{\ k\to\infty} \left| f'(z^k, z-y) - f'(y^k, z-y) \right| > 0. \tag{56}$$

It should be noted that, according to Lemma 2.3, a Bregman function f which has sequentially weakly-to-weak* continuous Gâteaux derivative f' necessarily satisfies ($D1$). However, *we do not know whether the converse implication is true, that is, we do not have examples of Bregman functions which satisfy ($D1$) without being sequentially weakly-to-weak* continuous.* The fact that condition $(D1)$ is sufficient for ensuring weak convergence of the algorithms A and B is shown by the following result:

Lemma. *Theorem* 4.3(ii) *and Theorem* 5.7(B) *still hold when one replaces the requirement that the Gâteaux derivative* $f' : \mathcal{D}^\circ \to B^*$ *is sequentially weakly-to-weak* continuous by condition* $(D1)$.

Proof. Let $\left\{x^k\right\}_{k\in\mathbb{N}}$ be a sequence generated either by the Algorithm A or by the Algorithm B under the circumstances of Theorem 4.3(ii) or of Theorem 5.7(B), respectively, in which the sequential weak-to-weak* continuity of f' is replaced by $(D1)$. Suppose, by contradiction, that the sequence $\left\{x^k\right\}_{k\in\mathbb{N}}$ does not converge weakly. Then, $\left\{x^k\right\}_{k\in\mathbb{N}}$ has two weakly convergent subsequences $\{x^{t_p}\}_{p\in\mathbb{N}}$ and $\{x^{s_p}\}_{p\in\mathbb{N}}$ included in S whose weak limits, x' and x'', respectively, are different. Since

$$|f'(x^{t_p}, x' - x'') - f'(x^{s_p}, x' - x'')| =$$

$$|(D_f(x', x^{t_p}) - D_f(x', x^{s_p})) - (D_f(x'', x^{t_p}) - D_f(x'', x^{s_p}))| \leq$$

$$|D_f(x', x^{t_p}) - D_f(x', x^{s_p})| + |D_f(x'', x^{t_p}) - D_f(x'', x^{s_p})|$$

and since, in the specific circumstances of each of the mentioned theorems, the sequences $\{D_f(x', x^k)\}_{k\in\mathbb{N}}$ and $\{D_f(x'', x^k)\}_{k\in\mathbb{N}}$ are convergent, by taking in the last inequality the limit for $p \to \infty$ one obtains

$$\underline{\lim}_{\ p\to\infty} \left| f'(x^{t_p}, x' - x'') - f'(x^{s_p}, x' - x'')\right| = 0$$

and this contradicts $(D1)$.

6.2 Theorem 5.7 ensures weak convergence of Algorithm B under the condition $(C2)$ which requires that, if Ω is not finite and if $\bigcup_{\omega\in\Omega} Q_\omega$ is unbounded, then the Bregman function f with respect to which the projections $\Pi_\omega^f(\cdot)$ are computed should be regularly convex. The fact is that we do not have examples of Bregman functions which are not regularly convex. This leads to the interesting open question of *whether any Bregman function is regularly convex.* As long as we do not have a positive answer to that question, we would like to replace the regular convexity requirement in Theorem 5.7 by a condition which is easier to verify in practical situations. To this end, note that very convex functions with continuous upper local moduli of convexity $\nu_f^+(x, \cdot)$ at all points $x \in \mathcal{D}^\circ$ are regularly convex. Proposition

2.4 in [12] shows that $\nu_f^+(x,\cdot)$ is necessarily continuous from the right on $[0,+\infty)$. Therefore, any very convex function whose upper local modulus of convexity $\nu_f^+(x,\cdot)$ is continuous from the left on $(0,+\infty)$ for all $x \in \mathcal{D}^\circ$ is regularly convex. Lemma 5.2 above enumerates significant examples of Bregman functions having $\nu_f^+(x,\cdot)$ continuous from the left on $(0,+\infty)$. We conjecture that the upper local modulus of convexity of any Bregman function is continuous from the left on $(0,+\infty)$. Obviously, if this conjecture is proven to be true, then the question asked above about the regular convexity of Bregman functions is affirmatively settled. Also, in this case, condition $(C2)$ involved in Theorem 5.7 can be replaced by the following, somewhat weaker, requirement:

$(D2)$ *If Ω is not finite, then either the set $\bigcup_{\omega\in\Omega} Q_\omega$ is bounded or the set $\bigcup_{\omega\in\Omega} Q_\omega$ is unbounded but, then, the function f is very convex.*

Even if the answer to the open question asked above is negative, one can replace in the proof of Theorem 5.7 the condition $(C2)$ by the condition $(D2)$ but at the expense of strenghtening the condition $(C3)$. Precisely, we have the following result:

Lemma. *Theorem* 5.7 *still holds if one replaces conditions $(C2)$ and $(C3)$ by conditions $(D2)$ and, respectively,*

$(D3)$ *There exists a measurable selector $\phi : \Omega \to B$ of the point-to-set mapping Q such that, for any $x \in \mathcal{D}^\circ$, the function $\omega \to D_f(\phi(\omega), x)$ is essentially bounded.*

Proof. Let $x \in \mathcal{D}^\circ$. Observe that the function $D_f(\cdot, x)$ is continuous. Therefore, the function $\omega \to D_f(\phi(\omega), x)$ is measurable whenever $\phi : \Omega \to B$ is a measurable selector of Q. Hence, $(D3)$ implies $(C3)$. Note that the only place in the proof of Theorem 5.7 where condition $(C2)$ is used is the proof of Lemma 5.4. Thus, it is enough to show that the conclusion of Lemma 5.4 remains true when one replaces conditions $(C2)$ and $(C3)$ by conditions $(D2)$ and $(D3)$, respectively. To this end, suppose that Ω is not finite and $\bigcup_{\omega\in\Omega} Q_\omega$ is unbounded. Let $M > 0$ be an essential upperbound of the function $D_f(\phi(\cdot), x)$. Then, we have

$$\nu_f^+(x, \left\|\Pi_\omega^f(x) - x\right\|) \le D_f(\Pi_\omega^f(x), x) \le D_f(\phi(\omega), x) \le M, \quad a.s.$$

The function $\nu_f^+(x,\cdot)$ is strictly increasing and has $\lim_{t\to\infty} \nu_f^+(x,t) = +\infty$. Hence, for some $t_0 > 0$, we have

$$\nu_f^+(x, \left\|\Pi_\omega^f(x) - x\right\|) \le M \le \nu_f^+(x, t_0), \quad a.s.$$

This shows that $\left\|\Pi_\omega^f(x) - x\right\| \leq t_0$, a.s., which implies that $\Pi_\bullet^f(x)$ is integrable because of (43).

6.3 As noted in Section 5.3, a demanding requirement involved in our convergence proof of Algorithm B (i.e., in the proof of Theorem 5.7) is the condition $(C4)$ which asks that the Bregman function f should be such that its associated functional D_f is convex on $S \times S$. It was already mentioned in Section 5.3 that this condition is automatically satisfied when B is a Hilbert space and $f(x) = \|x\|^2$ as well as in the case of $B = \mathbb{R}^n$ when $f(x)$ is the negentropy and condition $(C1)$ holds. Many other Bregman functions on $B = \mathbb{R}^n$ behave similarly. For instance, if $\alpha > 1$ and $f(x)$ is the negentropy, then the function $g : \mathbb{R}^n \to \mathbb{R}$ defined by

$$g(x) = c_0 \cdot f(x) + \sum_{i=1}^{n} c_i \cdot |x_i|^\alpha ,$$

where $c_0 \geq 0$ and $c_i > 0$ for $i = 1, ..., n$, is a Bregman function whose associated functional D_g is convex on $\mathbb{R}^n_{++} \times \mathbb{R}^n_{++}$.

In infinite dimensional Banach spaces the pool of functions which satisfy the condition $(C4)$ (and, depending on how the SCFP is given, the other conditions of Theorem 5.7) is not restricted to the square of the norm on Hilbert spaces. A significant example in this sense, which will be used below, is the function $f(x) = \|x\|^p$ when $B = \mathcal{L}^p([a,b])$ or $B = \ell^p$ with $p \in (1,2]$. Precisely, we have the following result:

Lemma. *If $p \in (1,2]$ and $B = \mathcal{L}^p([a,b])$ or $B = \ell^p$, then the functional D_f associated to $f(x) = \|x\|^p$ is convex on the set $B_+ \times B_+$, where B_+ denotes the subset of B consisting of all elements of B which are almost everywhere nonnegative.*

Proof. Note that for $p = 2$ the statement results from Remark 3.5. Assume that $p \in (1,2)$. Observe that, according to (15), we have

$$D_f(x,y) = \sum_{j=1}^{\infty} \left[|x_j|^p + (p-1) \cdot |y_j|^p - p \cdot |y_j|^{p-2} \cdot x_j \cdot y_j \right],$$

when $B = \ell^p$ and

$$D_f(x,y) = \int_a^b \left[|x|^p + (p-1) \cdot |y|^p - p \cdot |y|^{p-2} \cdot x \cdot y \right] dt,$$

when $B = \mathcal{L}^p([a,b])$. This and the continuity of D_f on $B \times B$ imply that if we prove that the function

$$\psi(s,t) = |s|^p + (p-1) \cdot |t|^p - p \cdot |t|^{p-2} \cdot s \cdot t$$

is convex on $\mathbb{R}_{++} \times \mathbb{R}_{++}$, then the convexity of D_f on $B_+ \times B_+$ results. Note that the Hessian matrix of the function ψ at the point $(s,t) \in \mathbb{R}_{++} \times \mathbb{R}_{++}$ is exactly $p \cdot (p-1) \cdot H(s,t)$, where

$$H(s,t) = \begin{bmatrix} s^{p-2} & -t^{p-2} \\ -t^{p-2} & t^{p-3} \cdot [(p-1) \cdot t + (2-p) \cdot s] \end{bmatrix}.$$

Hence, it suffices to prove that the matrix $H(s,t)$ is positive semidefinite. Since for $s,t > 0$ the elements of the diagonal of $H(s,t)$ are nonnegative, this amounts to showing that

$$Det(H(s,t)) := t^{p-3} \cdot \left[(p-1) \cdot s^{p-2} \cdot t + (2-p) \cdot s^{p-1} - t^{p-1}\right] \geq 0.$$

Let $\rho(s,t)$ be the expression between square brackets in the last equation. We show that $\rho(s,t) \geq 0$ when $s,t > 0$. To this end, note that $\rho(t,t) = 0$ and that, for each $t > 0$,

$$\lim_{s \to 0} \rho(s,t) = \lim_{s \to \infty} \rho(s,t) = +\infty.$$

This and the fact that $\rho(\cdot,t)$ is continuous on $(0,+\infty)$ imply that $\rho(\cdot,t)$ has at least one minimizer in $(0,+\infty)$. Clearly, $\frac{\partial \rho}{\partial s}(s,t)$ vanishes at $s = t$ only. Thus, the only minimizer of $\rho(\cdot,t)$ in $(0,+\infty)$ is $s = t$. Hence, $\rho(s,t) \geq \rho(t,t) = 0$, for all $s,t > 0$ and this completes our argument.

6.4 It was noted in Section 1.3 (see also [12]) that the computational efficiency of Bregman projection based algorithms for solving SCFPs can be improved when one fits the Bregman function f involved in the algorithm to the geometry of the sets Q_ω. The improvement results from reducing the computational effort of determining the Bregman projections required at each stage of the procedure. In some circumstances, determining Bregman projections on the sets Q_ω can be done by using closed formulae. This is typically the case when $B = \mathcal{L}^p([a,b])$ or $B = \ell^p$ and $f(x) = \|x\|^p$, where $p \in (1,+\infty)$ and the sets Q_ω are hyperplanes of the form

$$C = \{x \in B :< u, x >= v\} \tag{57}$$

or half-spaces of the form

$$C = \{x \in B :< u, x >\geq v\}, \tag{58}$$

with $u \in B^*\backslash\{0\}$ and $v \in \mathbb{R}$. A careful analysis of the proof of Proposition 5.5 in [2] shows[3] that the following result, which is used below, is true.

Lemma. *If $p \in (1, +\infty)$ and if C is the set given by (57) or the set given by (58), then, for any $x \in B$,*

$$\Pi_C^f(x) = \begin{cases} x & if \quad x \in C, \\ \left|s \cdot u + |x|^{p-2} \cdot x\right|^{q-2} \cdot \left(s \cdot u + |x|^{p-2} \cdot x\right) & if \quad x \notin C, \end{cases} \tag{59}$$

where $q = \frac{p}{p-1}$ and s represents the (necessarily unique) nonnegative solution of the equation

$$< u, \left|s \cdot u + |x|^{p-2} \cdot x\right|^{q-2} \cdot \left(s \cdot u + |x|^{p-2} \cdot x\right) > -v = 0. \tag{60}$$

6.5 The facts established in Sections 6.3 and 6.4 combined with Proposition 3.4 allow us to give the following example which illustrates how the Algorithm B can be applied to solving linear operator equations in $B = \mathcal{L}^p([a, b])$ with $p \in (1, 2]$. For making a long story shorter, we discuss a particular case of a Fredholm integral equation of the first kind, but one should note that the same idea can be generalized along the lines traced in [10, Section 4]. Precisely, we consider the following problem: *Find a function $x \in \mathcal{L}^p([a, b])$ such that*

$$\int_a^b K(\omega, s) \cdot x(s)ds = w(\omega) and x(\omega) \geq z(\omega), \quad a.s., \tag{61}$$

where $K : [a, b]^2 \to \mathbb{R}$ is continuous, $z \in \mathcal{L}^p([a, b])$ and $w \in \mathcal{L}^1([a, b])$. We assume that a solution of this problem exists and want to find it via Algorithm B. We aim at showing that finding a solution of problem (61) can be done by solving a consistent SCFP in the space $\mathcal{L}^p([a, b])$ and that Theorem 5.7 ensures that sequences $\{x^k\}_{k\in\mathbb{N}}$ generated by Algorithm B can be used in order to approximate solutions of it. To this end, observe that finding a

[3] Proposition 5.5 in [2] is stated under the assumption that $p > 2$, but it can be easily verified that its proof remains true when $p \in (1, 2]$.

solution $x \in \mathcal{L}^p([a,b])$ of (61) amounts to finding a solution $y \in \mathcal{L}^p([a,b])$ of the problem

$$\int_a^b K(\omega, s) \cdot y(s) d\mu(s) = v(\omega) and y(\omega) \geq 0, \quad a.s., \tag{62}$$

where

$$v(\omega) = (b-a)w + \int_a^b K(\omega, s) \cdot z(s) d\mu(s)$$

and $\mu = (b-a)^{-1}\lambda$ with λ being the Lebesgue measure on the $\sigma-$algebra $\mathcal{A}$ of Lebesgue measurable subsets of $\Omega = [a,b]$. This is true because, if y is a solution of (62), then $x = y + z$ is a solution of (61).

For each $\omega \in [a,b]$ denote

$$Q_\omega := \{y \in \mathcal{L}_+^p([a,b]) : T_\omega y = v(\omega)\}, \tag{63}$$

with

$$T_\omega y := \int_a^b K(\omega, s) \cdot y(s) d\mu(s). \tag{64}$$

The next result ensures that the SCFP associated to the point-to-set mapping Q given by (63) is correctly posed and that Algorithm B produces sequences which approximate solutions of it and, implicitly, solutions of the Fredholm integral equation (61).

Lemma. *If the problem* (61) *has solutions in* $\mathcal{L}^p([a,b])$ *with* $p \in (1,2]$, *then*

(i) *The sets* Q_ω *are convex, closed, included in* $\mathcal{L}_+^p([a,b])$ *and nonempty for almost all* $\omega \in [a,b]$;

(ii) *The point-to-set mapping* $\omega \to Q_\omega : [a,b] \to \mathcal{L}^p([a,b])$ *is measurable;*

(iii) *Any almost common point of the sets* Q_ω, $\omega \in [a,b]$, *is a solution of* (62) *and conversely;*

(iv) *If* $f(x) = \|x\|^p$, *then any sequence* $\{x^k\}_{k\in\mathbb{N}}$ *recursively defined in* $\mathcal{L}_+^p([a,b])$ *by*

$$x^{k+1} = \int_a^b \Pi_{Q_\omega}^f (x^k) d\mu(\omega), \tag{65}$$

with $x^0 \in \mathcal{L}_+^p([a,b])$ *exists, has weak accumulation points and all its weak accumulation points are solutions of* (62); *In this case, the sequence*

$$y^k = x^k + z \tag{66}$$

has weak accumulation points and each such point is a solution of (61);

(v) *If* $p=2$, *then the sequence* $\{y^k\}_{k\in\mathbb{N}}$ *defined by* (65) *and* (66) *converges weakly to a solution of* (61).

Proof. (i) Observe that if x is a solution of problem (61), then $x-z\in Q_\omega$ for almost all $\omega\in[a,b]$. Therefore, only closedness of the sets Q_ω has to be proved. To this end, let $\{y^k\}_{k\in\mathbb{N}}$ be a sequence in Q_ω and assume that $\{y^k\}_{k\in\mathbb{N}}$ converges strongly in $\mathcal{L}^p([a,b])$ to y^*. Then, there exists a subsequence $\{y^{p_k}\}_{k\in\mathbb{N}}$ such that $\lim_{k\to\infty} y^{p_k}(t)=y^*(t)$, a.s. (see, for instance, [7, Proposition 4.9]). This implies that $y^*(t)\geq z(t)$, a.s. Since the functional T_ω is continuous on B we also have $T_\omega y^*=\lim_{k\to\infty} T_\omega y^k=v(\omega)$, i.e., $y^*\in Q_\omega$.

(ii) According to (i), Proposition 8.2.9 in [3] applies to the point-to-set mapping $Q:\omega\to Q_\omega$ and it shows that this mapping is measurable.

(iii) results immediately from (i) and (ii).

(iv) We apply Theorem 5.7 to the SCFP associated to the point-to-set mapping Q defined by (63). To this end, observe that this SCFP is consistent (cf. (iii)) and that the conditions ($C1$)-($C4$) are satisfied when $f(x)=\|x\|^p$ because the set $S=Cl(conv\bigcup_{\omega\in\Omega}Q_\omega)$ is included in $\mathcal{L}^p_+([a,b])$ and D_f is convex on $\mathcal{L}^p_+([a,b])\times\mathcal{L}^p_+([a,b])$ as shown by Lemma 6.3. Thus, Theorem 5.7 ensures that Algorithm B and, hence, formula (65) which is a particular version of that algorithm, generates bounded sequences whose weak accumulation points are solutions of the given SCFP and, implicitly, solutions of (62).

(v) Let $p=2$. In this case the function f' is sequentially weakly-to-weak* continuous and, thus, Theorem 5.7 implies that the sequence $\{x^k\}_{k\in\mathbb{N}}$ is weakly convergent to an almost common point x^* of the sets Q_ω, $\omega\in[a,b]$. Consequently, the sequence $\{y^k\}_{k\in\mathbb{N}}$ converges weakly to x^*+z and this is a solution of (61).

6.6 The Algorithm A is not applicable to the SCFP described in Section 6.5 because this SCFP is not in the form described in Section 4.1. However, the Algorithm A can be applied to SCFPs which do not satisfy all conditions ($C1$)-($C4$). Such a SCFP is that emerging from the following *approximation problem* which is frequently encountered in electrical engineering: *Given the real number* $\varepsilon>0$, *an integrable function* $\varphi_0:[a,b]\to\mathbb{R}$, *and the sequence* $\{\varphi_k\}_{k\in\mathbb{N}_0}$ *of continuous nonnegative real functions defined on the interval* $[a,b]$ *such that, for each* $t\in[a,b]$, *the sequence* $\varphi(t)=(\varphi_1(t),\varphi_2(t),...)\in\ell^q$, *where* $q = p/(p-1)$ *with* $p \in (1, +\infty)$, *find* $x = (x_1, x_2, ...) \in \ell^p$ *such that, for any* $t \in [a, b]$,

$$\left| \varphi_0(t) - \sum_{j=1}^{\infty} x_j \cdot \varphi_j(t) \right| \leq \varepsilon, \tag{67}$$

provided that such an $x \in \ell^p$ *exists.*

In order to rewrite this problem as a SCFP in ℓ^p denote $\Omega = [a, 2 \cdot b - a]$ and provide it with the σ-algebra of its Lebesgue measurable subsets and with the probability measure $\mu = (2 \cdot b - 2 \cdot a)^{-1} \cdot \lambda$, where λ represents the Lebesgue measure. Define the function $g : \Omega \times \ell^p \to \mathbb{R}$ by

$$g(\omega, x) = \begin{cases} \varphi_0(\omega) - \sum_{j=1}^{\infty} x_j \cdot \varphi_j(\omega) - \varepsilon & if \quad \omega \in [a, b], \\ \\ \sum_{j=1}^{\infty} x_j \cdot \varphi_j(\omega - b + a) - \varphi_0(\omega - b + a) - \varepsilon & if \quad \omega \in (b, 2 \cdot b - a]. \end{cases}$$

This is a convex Carathéodory function which satisfies the requirements (I) and (II) in Section 4.1 because $g(\cdot, 0)$ is integrable and

$$|g(\omega, x) - g(\omega, y)| \leq \|\varphi(\omega)\| \cdot \|x - y\| .$$

Therefore, Theorem 4.3 (and, specifically, Corollary 4.4) applies to the SCFP with the sets Q_ω defined by (21) and taking $f(x) = \|x\|^p$. It ensures that the sequences generated by the Algorithm A converge weakly to an almost common point of the sets Q_ω, $\omega \in \Omega$. It can be easily seen that any such point is a solution of the approximation problem posed above. Also, formula (59) can be used in order to determine the points $\mathbf{P}^f(x^k)$ at each stage of the algorithm.

Acknowledgments

The work of Dan Butnariu was partially done during his visit to the Institute of Pure and Applied Mathematics (IMPA) of Rio de Janeiro, Brazil. Also, Dan Butnariu gratefully acknowledges the support of the Israel Science Foundation founded by the Israel Academy of Sciences and Humanities.

The authors are grateful to Yakov Alber, Benar Svaiter and to the reviewers for competent comments which led to improvements of earlier versions of this work.

References

[1] Alber, Ya., Metric and generalized projection operators in Banach spaces, in: M.E. Drakhlin and E. Litsyn (Eds.), "Functional Differential Equations", pp.1-21, *The Research Institute of the College of Iudea and Samaria Publications*, Kedumim-Ariel, Israel, 1993.

[2] Alber, Y. and Butnariu, D., Convergence of Bregman-projection methods for solving convex feasibility problems in reflexive Banach spaces, *J. Optim. Theory Appl.*, 92, 1997, 33-61.

[3] Aubin, J.-P. and Frankowska, H.: Set-Valued Analysis, *Birkhäuser,* Basel, 1990.

[4] Bauschke, H.H. and Borwein, J.M.: On projection algorithms for solving convex feasibility problems, *SIAM Rev.*, 38, 1996, 367-426.

[5] Bauschke, H.H. and Borwein, J.M., Legendre functions and the method of random Bregman projections, *J. Convex Anal.,* 4, 1997, 27-67.

[6] Bregman, L.M., The relaxation method for finding common points of convex sets and its application to the solution of convex programming, *USSR Comp. Math. and Math. Phys.*, 7, 1967, 200-217.

[7] Brezis, H., Analyse Fonctionelle: Théorie et Applications, *Masson*, Paris, 1983.

[8] Browder, F.E., Nonlinear Operators and Nonlinear Equations of Evolution in Banach Spaces, *Proceedings of Symposia in Pure Mathematics,* Vol. XVIII, Part 2, *American Mathematical Society*, Providence, Rhode Island, 1976.

[9] Burachik, R.S., Generalized proximal point methods for the variational inequality problem, Ph.D.Thesis, *Instituto de Matemática Pura e Aplicada, Rio de Janeiro*, Brazil, 1995.

[10] Butnariu, D., The expected-projection methods: Its behavior and applications to linear operator equations and convex optimization, *J. Applied Analysis*, Vol. 1, No. 1, 1995, 95-108.

[11] Butnariu, D. and Censor, Y., Strong convergence of almost simultaneous projection methods in Hilbert spaces, *J. Comput. Appl. Math.*, 53, 1994, 33-42.

[12] Butnariu, D., Censor, Y. and Reich, S., Iterative averaging of entropic projections for solving stochastic convex feasibility problems, *Comput. Optim. Appl.,* 8, 1997, 21-38.

[13] Butnariu, D. and Flam, S.D., Strong convergence of expected projection methods in Hilbert spaces, *Numer. Funct. Anal. Optim.*, 16, 1995, 601-636.

[14] Butnariu, D. and Iusem, A.N., Local moduli of convexity and their applications to finding almost common fixed points of mesurable families of operators, in: Y. Censor and S. Reich (Eds.): "Recent Developments in Optimization and Nonlinear Analysis", *Contemporary Mathematics*, Vol. 204, pp. 33-61, *American Mathematical Society*, Providence, Rhode Island, 1997.

[15] Butnariu, D. and Mehrez, A., Convergence criteria for generalized gradient methods of solving locally Lipschitz feasibility problems, *Comput. Optim. Appl.*, 1, 1992, 307-326.

[16] Censor, Y. and Lent, A., An iterative row action method for interval convex programming, *J. Optim. Theory Appl.*, 34(3), 1981, 321-353.

[17] Censor, Y. and Zenios, S., Parallel Optimization: Theory, Algorithms and Applications, *Oxford University Press*, 1997.

[18] Cimmino, G., Calcolo approsimato per le soluzioni di sistemi di ecuazioni lineari, *La Riserca Scientifica, Roma,* XVI, Anno IX, 2, 1938, 326-333.

[19] Clarke, F.H., Optimization and Nonsmooth Analysis, *John Wiley and Sons*, New York, 1983.

[20] Combettes, P.L., The Convex Feasibility Problem in Image Recovery, in: P. Hawkes (ed.), *Advances in Imaging and Electron Phiyisc, pp. 155-270, Vol. 95, Academic Press*, New York, 1995.

[21] De Pierro, A.R. and Iusem, A.N., A parallel projection method for finding a common point of a family of convex sets, *Pesquisa Operacional*, 5, 1985, 243-253.

[22] De Pierro, A.R. and Iusem, A.N., A relaxed version of Bregman's method for convex programming, *J. Optim. Theory Appl.*, 51, 1986, 421-440.

[23] Isratescu, V.I., Strict Convexity and Complex Strict Convexity, Lecture Notes in Pure and Applied Mathematics, vol. 89, *Marcel Dekker*, New York, 1984.

[24] Iusem, A.N. and De Pierro, A.R., Convergence results for an accelerated Cimmino algorithm, *Numer. Math.*, 49, 1986, 347-368.

[25] Kammerer, W.J. and Nashed, M.Z., Iterative methods for best approximate solutions of linear integral equations of the first and second kind, *J. Math. Anal. Appl.*, 40, 1972, 547-573.

[26] Kantorovich, L.V.K. and Akilov, G.P., Functional Analysis, *Pergamon Press*, Oxford, 1982.

[27] Mikusinski, J., The Bochner Integral, *Academic Press*, New York 1978.

[28] Neumann, J. von, Functional Operators - Vol. II, The Geometry of Orthogonal Spaces, in: *Annals of Mathematics Studies*, No. 22, *Princeton University Press*, 1950. [Reprint of mimeographed lecture notes distributed in 1933].

[29] Phelps, R.R., Convex Functions, Monotone Operators and Differentiability, 2-nd Edition, *Springer Verlag*, Berlin, 1993.

[30] Reich, S., A weak convergence theorem for the alternanting method with Bregman distances, in: A.G. Kartsatos (Ed.), "Theory and Applications of Nonlinear Operators of Accretive and Monotone Types", pp. 313-318, *Marcel Dekker*, New York, 1996.

[31] Rockafellar, R.T. and Wets, R.J.-B., Scenarios of policy aggregation in optimization under uncertainty, *Math. Oper. Res.*, 16, 1991, 119-147.

[32] Vladimirov, A.A., Nesterov, Y.E. and Chekanov, Y.N., Uniformly convex functionals (Russian), *Vestnik Moskovskaya Universiteta, Series Mathematika i Kybernetika*, 3, 1978, 12-23.