Méthodes d'éclatement basées sur les distances de Bregman pour les inclusions monotones composites et l'optimisation

Van Quang Nguyen

THÈSE DE DOCTORAT DE
L'UNIVERSITÉ PIERRE ET MARIE CURIE – PARIS VI

Spécialité :

MATHÉMATIQUES APPLIQUÉES

Présentée par :

Quang Văn Nguyễn

pour l'obtention du grade de

DOCTEUR DE L'UNIVERSITÉ PIERRE ET MARIE CURIE – PARIS VI

Sujet de la thèse :

Méthodes d'Éclatement Basées sur les Distances de Bregman pour les Inclusions Monotones Composites et l'Optimisation

Soutenue le 17 juillet 2015 devant le jury composé de :

Jean-Baptiste CAILLAU	*examinateur*
Patrick L. COMBETTES	*directeur de thèse*
Hélène FRANKOWSKA	*présidente*
Stéphane GAUBERT	*examinateur*
Bruno NAZARET	*examinateur*
Jean-Christophe PESQUET	*rapporteur*
Simeon REICH	*rapporteur*
Michel THÉRA	*examinateur*

Laboratoire Jacques-Louis Lions – UMR 7598

Remerciements

En tout premier lieu, je souhaite remercier profondément Monsieur Patrick Louis Combettes, mon directeur de thèse, d'avoir accepté de superviser mes premiers pas dans la recherche. Je lui suis très reconnaissant de son très grand soutien, de sa très grande disponibilité, et de son attention constante. Je tiens à lui exprimer ma gratitude pour sa patience en comprenant mon français et aussi mon anglais, en m'expliquant des questions mathématiques, et en corrigeant soigneusement toute ma thèse. Sa connaissance mathématique et son style de travail vont toujours m'influencer.

J'adresse mes remerciements à Monsieur Jean-Christophe Pesquet et à Monsieur Simeon Reich pour avoir accepté d'être rapporteurs de ma thèse. Monsieur Jean-Christophe Pesquet m'a par ailleurs apporté une aide précieuse sur certains aspects numériques, et j'aimerais lui exprimer ma gratitude. Je tiens également à exprimer ma grande reconnaissance aux membres du jury pour leur participation.

Cette thèse a été financée par le Ministère de l'Éducation et de la Formation du Vietnam et par l'Université Pierre et Marie Curie. J'ai eu de la chance de travailler dans l'environnement scientifique très fertile du Laboratoire Jacques-Louis Lions et du groupe de travail GdR MOA.

J'aimerais remercier mes professeurs des mathématiques : Monsieur Nguyen Thanh Anh pour son encouragement et Monsieur Do Duc Thai pour m'avoir présenté à mon directeur de thèse.

Je tiens à remercier mes amis : Bang, Kien, Thibault, et Viet pour leurs discussions mathématiques ; Hieu et Thi pour leur aide en écrivant les codes en Octave ; Cyril et Magali pour leurs corrections orthographiques ; et Cuong et Diep pour leur aide en arrangeant mes séjours en France.

Cette thèse est dédiée à ma famille : mes grands–parents, mes parents, et mes frères.

Paris, le 17 juillet 2015

Table des matières

Résumé

Méthodes d'Éclatement Basées sur les Distances de Bregman pour les Inclusions Monotones Composites et l'Optimisation

Le but de cette thèse est d'élaborer des méthodes d'éclatement basées sur les distances de Bregman pour la résolution d'inclusions monotones composites dans les espaces de Banach réels réflexifs. Ces résultats nous permettent d'étendre de nombreuses techniques, jusqu'alors limitées aux espaces hilbertiens. De plus, même dans le cadre restreint d'espaces euclidiens, ils donnent lieu à de nouvelles méthodes de décomposition qui peuvent s'avérer plus avantageuses numériquement que les méthodes classiques basées sur la distance euclidienne. Des applications numériques en traitement de l'image sont proposées.

Mots-clés : distance de Bregman, dualité, espace de Banach, Haugazeau, inclusion monotone, meilleure approximation, algorithme d'éclatement, algorithme à métrique variable, opérateur monotone, restauration d'images.

Tóm tắt

Các phương pháp chia tách toán tử dựa trên khoảng cách Bregman với các bao hàm thức đơn điệu chứa hợp thành và bài toán tối ưu

Mục đích của luận án là nghiên cứu một số phương pháp chia tách toán tử sử dụng khoảng cách Bregman để giải quyết các bao hàm thức đơn điệu trong các không gian Banach thực phản xạ. Các kết quả của luận án cho phép mở rộng nhiều kỹ thuật mà cho đến hiện tại mới chỉ giới hạn trong các không gian Hilbert. Hơn nữa, ngay cả khi hạn chế trong các không gian Euclid, những kết quả này cung cấp các phương pháp chia tách mới với nhiều ưu điểm về mặt tính toán khi so sánh với các phương pháp cổ điển dựa trên khoảng cách Euclid. Chúng tôi cũng đề xuất các ứng dụng thực tế trong bài toán xử lý ảnh.

Từ khóa : bao hàm thức đơn điệu, thuật toán biến thiên metric, đối ngẫu, Haugazeau, khoảng cách Bregman, khôi phục ảnh, không gian Banach, thuật toán chia tách, toán tử đơn điệu, xấp xỉ tốt nhất.

Abstract

Splitting Methods Based on Bregman Distances for Composite Monotone Inclusions and Optimization

The goal of this thesis is to design splitting methods based on Bregman distances for solving composite monotone inclusions in reflexive real Banach spaces. These results allow us to extend many techniques that were so far limited to Hilbert spaces. Furthermore, even when restricted to Euclidean spaces, they provide new splitting methods that may be more avantageous numerically than the classical methods based on the Euclidean distance. Numerical applications in image processing are proposed.

Key words : Banach space, best approximation, Bregman distance, duality, Haugazeau, image restoration, monotone inclusion, monotone operator, splitting algorithm, variable metric algorithm.

Notations et Glossaire

Les notations suivantes seront utilisées dans toute la thèse. De plus, nous rappelons certaines définitions de base en analyse convexe.

Notations générales

- $\mathcal{X}$, $\mathcal{Y}$: Espaces de Banach réels.
- $\mathcal{X}^*$, $\mathcal{Y}^*$: Duaux topologiques de $\mathcal{X}$ et $\mathcal{Y}$.
- $\langle \cdot, \cdot \rangle$: Crochet de dualité entre un espace de Banach et son dual topologique.
- $\| \cdot \|$: Norme d'un espace de Banach.
- $B(x; \rho)$: Boule de centre $x \in \mathcal{X}$ et de rayon $\rho \in \,]0, +\infty[$.
- $\mathcal{H}$: Espace hilbertien réel.
- Id : Opérateur identité.
- $2^{\mathcal{X}}$: Ensemble des parties de $\mathcal{X}$.
- $\Gamma_0(\mathcal{X})$: Ensemble des fonctions de $\mathcal{X}$ dans $]-\infty, +\infty]$, convexes, propres, et semicontinues inférieurement.
- $\mathcal{B}\left(\mathcal{X}, \mathcal{Y}\right)$: Espace des opérateurs linéaires bornés de $\mathcal{X}$ dans $\mathcal{Y}$.
- $\mathcal{B}\left(\mathcal{X}\right)$: $\mathcal{B}\left(\mathcal{X}, \mathcal{X}\right)$.
- L^* : Adjoint de l'opérateur $L \in \mathcal{B}\left(\mathcal{X}, \mathcal{Y}\right)$.
- $\to$: Convergence forte.
- $\rightharpoonup$: Convergence faible.
- $\overline{\lim}\, \alpha_n$: Limite supérieure de la suite $(\alpha_n)_{n\in\mathbb{N}}$ de $[-\infty, +\infty]$.
- $\underline{\lim}\, \alpha_n$: Limite inférieure de la suite $(\alpha_n)_{n\in\mathbb{N}}$ de $[-\infty, +\infty]$.

Soit C un sous-ensemble de $\mathcal{X}$.

- Fonction indicatrice de C :

$$\iota_C \colon x \mapsto \begin{cases} 0, & \text{si } \; x \in C; \\ +\infty, & \text{si } \; x \notin C. \end{cases}$$

- Opérateur cône normal à C :

$$N_C \colon x \mapsto \begin{cases} \big\{x^* \in \mathcal{X}^* \bigm| \sup \langle C - x, x^* \rangle \leqslant 0 \big\}, & \text{si } \; x \in C; \\ \varnothing, & \text{sinon.} \end{cases}$$

- $\operatorname{int} C$: Intérieur de C.
- $\overline{C}$: Adhérence de C.
- $\operatorname{bdry} C$: Frontière de C.
- $\operatorname{cone} C = \bigcup_{\lambda>0} \lambda C$.
- $\operatorname{sri} C = \big\{x \in C \;\big|\; \operatorname{cone}(C-x) = \overline{\operatorname{span}}(C-x)\big\}$: Intérieur relatif fort de C.

Notations et définitions relatives à un opérateur multivoque $M\colon \mathcal{X} \to 2^{\mathcal{Y}}$

- $\operatorname{dom} M = \big\{x \in \mathcal{X} \;\big|\; Mx \neq \varnothing\big\}$: Domaine de M.
- $\operatorname{gra} M = \big\{(x,y) \in \mathcal{X} \times \mathcal{Y} \;\big|\; y \in Mx\big\}$: Graphe de M.
- $M^{-1}\colon \mathcal{Y} \to 2^{\mathcal{X}}\colon y \mapsto \big\{x \in \mathcal{X} \;\big|\; y \in Mx\big\}$: Inverse de M.
- $\operatorname{zer} M = \big\{x \in \mathcal{X} \;\big|\; 0 \in Mx\big\}$: Zéros de M.
- $\operatorname{ran} M = \big\{y \in \mathcal{Y} \;\big|\; (\exists x \in \mathcal{X})\;\; y \in Mx\big\}$: Image de M.
- $\operatorname{Fix} M = \big\{x \in \mathcal{X} \;\big|\; x \in Mx\big\}$ (où $\mathcal{Y} = \mathcal{X}$) : Points fixes de M.

Définitions relatives à un opérateur multivoque $M\colon \mathcal{X} \to 2^{\mathcal{X}^*}$

- M est monotone :

$$(\forall (x_1, x_1^*) \in \operatorname{gra} M)(\forall (x_2, x_2^*) \in \operatorname{gra} M) \quad \langle x_1 - x_2, x_1^* - x_2^*\rangle \geqslant 0.$$

- M est maximalement monotone : M est monotone et il n'existe pas d'opérateur monotone de $\mathcal{X}$ dans $2^{\mathcal{X}^*}$ dont le graphe contient strictement $\operatorname{gra} M$.
- M est uniformément monotone en $x_1 \in \operatorname{dom} M$ s'il existe une fonction strictement croissante $\phi\colon\; [0,+\infty[\to [0,+\infty]$ qui s'annule seulement en 0 telle que

$$(\forall x_1^* \in Mx_1)(\forall (x_2, x_2^*) \in \operatorname{gra} M) \quad \langle x_1 - x_2, x_1^* - x_2^*\rangle \geqslant \phi(\|x_1 - x_2\|).$$

Notations relatives à une fonction $f \in \Gamma_0(\mathcal{X})$

- Domaine de f : $\operatorname{dom} f = \big\{x \in \mathcal{X} \;\big|\; f(x) < +\infty\big\}$.
- Ensemble des minimiseurs de f : $\operatorname{Argmin} f$.
- Conjuguée de f : $f^*\colon \mathcal{X}^* \to \left]-\infty,+\infty\right]\colon x^* \mapsto \sup\limits_{x \in \mathcal{X}} \big(\langle x, x^*\rangle - f(x)\big)$.
- Sous-différentiel de f :

$$\partial f\colon \mathcal{X} \to 2^{\mathcal{X}^*}\colon x \mapsto \big\{x^* \in \mathcal{X}^* \;\big|\; (\forall y \in \mathcal{X})\; \langle y - x, x^*\rangle + f(x) \leqslant f(y)\big\}.$$

Si $f \in \Gamma_0(\mathcal{X})$ est Gâteaux différentiable sur $\operatorname{int} \operatorname{dom} f \neq \varnothing$:

- Distance de Bregman associée à f :

$$D^f \colon \mathcal{X} \times \mathcal{X} \to [0, +\infty] : \\ (x, y) \mapsto \begin{cases} f(x) - f(y) - \big\langle x - y, \nabla f(y) \big\rangle, & \text{si } y \in \operatorname{int} \operatorname{dom} f; \\ +\infty, & \text{sinon.} \end{cases}$$

- D^f-résolvante de $M \colon \mathcal{X} \to 2^{\mathcal{X}^*}$:

$$R^f_M = (\nabla f + M)^{-1} \circ \nabla f \colon \mathcal{X} \to 2^{\mathcal{X}}.$$

- f-résolvante de $M \colon \mathcal{X} \to 2^{\mathcal{X}^*}$:

$$J^f_M = (\nabla f + M)^{-1} \colon \mathcal{X}^* \to 2^{\mathcal{X}}.$$

- Opérateur D^f-proximal de $\varphi \in \Gamma_0(\mathcal{X})$:

$$\operatorname{prox}^f_\varphi \colon \mathcal{X} \to 2^{\mathcal{X}} : y \mapsto \underset{x \in \mathcal{X}}{\operatorname{Argmin}} \; \big(\varphi(x) + D^f(x, y)\big).$$

- Opérateur f-proximal de $\varphi \in \Gamma_0(\mathcal{X})$:

$$\operatorname{Prox}^f_\varphi \colon \mathcal{X}^* \to 2^{\mathcal{X}} : x^* \mapsto \underset{x \in \mathcal{X}}{\operatorname{Argmin}} \; \Big(\varphi(x) + \big(f(x) - \langle x, x^* \rangle + f^*(x^*)\big)\Big).$$

- D^f-projecteur sur un sous-ensemble convexe C de $\mathcal{X}$.

$$P^f_C \colon \mathcal{X} \to 2^{\mathcal{X}} : y \mapsto \underset{x \in C}{\operatorname{Argmin}} \; D^f(x, y).$$

- D^f-distance à un sous-ensemble convexe $C \subset \mathcal{X}$:

$$D^f_C \colon \mathcal{X} \to [0, +\infty] : y \mapsto \inf_{x \in C} D^f(x, y).$$

Quelques fonctions dans $\Gamma_0(\mathbb{R})$

- Entropie de Boltzmann-Shannon

$$f \colon \mathbb{R} \to \left]-\infty, +\infty\right] : \xi \mapsto \begin{cases} \xi \ln \xi - \xi, & \text{si } \xi > 0; \\ 0, & \text{si } \xi = 0; \\ +\infty, & \text{sinon.} \end{cases}$$

- Entropie de Burg

$$f\colon \mathbb{R} \to \left]-\infty, +\infty\right] : \xi \mapsto \begin{cases} -\ln \xi, & \text{si } \ \xi > 0; \\ +\infty, & \text{sinon.} \end{cases}$$

- Entropie de Fermi-Dirac

$$f\colon \mathbb{R} \to \left]-\infty, +\infty\right] : \xi \mapsto \begin{cases} \xi \ln \xi + (1-\xi)\ln(1-\xi), & \text{si } 0 < \xi < 1; \\ 0, & \text{si } \xi \in \{0,1\}; \\ +\infty, & \text{sinon.} \end{cases}$$

- Fonction de type Hellinger

$$f\colon \mathbb{R} \to \left]-\infty, +\infty\right] : \xi \mapsto \begin{cases} -\sqrt{1-\xi^2}, & \text{si } \xi \in [-1,1]; \\ +\infty, & \text{sinon.} \end{cases}$$

Chapitre 1

Introduction

1.1 Inclusions monotones

La théorie des opérateurs monotones a émergé au début des années 1960 comme une branche bien structurée de l'analyse non linéaire [4, 15, 21, 45, 53, 54, 65] et elle reste très active [12, 15, 16, 35, 57, 60, 66]. L'une des raisons principales du succès de la théorie est qu'un nombre considérable de problèmes dans divers domaines (par exemple, l'optimisation, l'économie, les inégalités variationnelles, les équations aux dérivées partielles, la mécanique, le traitement d'images et du signal, le transport optimal, l'apprentissage automatique, la théorie des jeux, et la théorie du trafic) peut être réduit à la résolution du problème d'inclusion monotone suivant.

Problème 1.1 Soit $\mathcal{X}$ un espace de Banach réel réflexif, soit $\mathcal{X}^*$ le dual topologique de $\mathcal{X}$, et soit $M\colon \mathcal{X} \to 2^{\mathcal{X}^*}$ maximalement monotone. Le problème est de

$$\text{trouver } \; x \in \mathcal{X} \; \text{ tel que } \; 0 \in Mx. \tag{1.1}$$

L'ensemble des solution de (1.1) est noté zer M.

Une spécialisation du Problème 1.1 est le problème de minimisation convexe suivant.

Problème 1.2 Soit $\mathcal{X}$ un espace de Banach réel réflexif, soit $\mathcal{X}^*$ le dual topologique de $\mathcal{X}$, et soit $\varphi \in \Gamma_0(\mathcal{X})$. Le problème est de

$$\underset{x \in \mathcal{X}}{\text{minimiser}} \; \varphi(x). \tag{1.2}$$

Un objectif de la théorie des opérateurs monotones et de l'optimisation est de concevoir des méthodes itératives qui engendrent des suites convergeant vers une solution du Problème 1.1 ou du Problème 1.2.

1.1.1 Méthode proximale

Dans le cadre d'espaces hilbertiens, une méthode classique pour résoudre le Problème 1.1 est l'algorithme du point proximal.

Théorème 1.3 [23] et [12, Theorem 23.41] *Dans le Problème 1.1, supposons que $\mathcal{X}$ soit un espace hilbertien réel et que* $\operatorname{zer} M \neq \varnothing$. *Soit $x_0 \in \mathcal{X}$ et soit $(\gamma_n)_{n\in\mathbb{N}} \in \left]0,+\infty\right[^{\mathbb{N}}$ telle que $\sum_{n\in\mathbb{N}} \gamma_n^2 = +\infty$. Alors la suite $(x_n)_{n\in\mathbb{N}}$ engendrée par l'algorithme*

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = (\operatorname{Id} + \gamma_n M)^{-1}(x_n) \tag{1.3}$$

converge faiblement vers un point dans $\operatorname{zer} M$.

Une variante du Théorème 1.3 pour résoudre le Problème 1.2 est la suivante.

Théorème 1.4 [23] et [12, Theorem 27.1] *Dans le Problème 1.2, supposons que $\mathcal{X}$ soit un espace hilbertien réel et que* $\operatorname{Argmin} \varphi \neq \varnothing$. *Soit*

$$\operatorname{prox}_\varphi \colon \mathcal{X} \to \mathcal{X} \colon x \mapsto \operatorname{argmin}\big(\varphi + \|\cdot - x\|^2/2\big), \tag{1.4}$$

soit $x_0 \in \mathcal{X}$, et soit $(\gamma_n)_{n\in\mathbb{N}} \in \left]0,+\infty\right[^{\mathbb{N}}$ telle que $\sum_{n\in\mathbb{N}} \gamma_n = +\infty$. Alors la suite $(x_n)_{n\in\mathbb{N}}$ engendrée par l'algorithme

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = \operatorname{prox}_{\gamma_n\varphi} x_n \tag{1.5}$$

converge faiblement vers un point $\overline{x} \in$ $\operatorname{Argmin} \varphi$.

Au cœur de l'analyse de convergence des algorithmes (1.3) et (1.5) se trouvent le Théorème de Minty [54] qui dit que $\operatorname{Id} + M$ est surjective, et la monotonie fejérienne de $(x_n)_{n\in\mathbb{N}}$ sous la forme suivante [33, 40]

$$(\forall x \in \operatorname{zer} M)(\forall n \in \mathbb{N}) \quad \|x - x_{n+1}\| \leqslant \|x - x_n\|. \tag{1.6}$$

La notion de monotonie fejérienne et ses variantes jouent un rôle primordial dans l'analyse de la convergence de nombreux algorithmes de point fixe et d'optimisation dans les espaces hibertiens [8, 12, 32, 33, 40, 58]. Un développement récent dans ce domaine est l'extension de la notion de suite monotone (quasi)-fejérienne au cas où la métrique sous-jacente est autorisée à varier [37].

En cherchant une variante similaire des algorithmes (1.3) et (1.5) dans des espaces de Banach, on essaie de trouver une généralisation de l'opérateur $(\operatorname{Id} + M)^{-1}$ en même temps qu'une notion convenable de monotonie similaire à (1.6). Dans le cas d'un espace hilbertien réel identifié à son dual topologique, l'application de dualité $\Delta = \partial\big(\|\cdot\|^2/2\big)$ se réduit à l'opérateur identité. Dans un espace de Banach réel réflexif $\mathcal{X}$ tel que $\mathcal{X}^*$ est strictement convexe, Δ est univoque [22, I.1, Remarque 2] et le Théorème de Brézis-Browder [15, Theorem 3] affirme que $\Delta + M$ est surjective. En utilisant ce résultat, il est démontré dans [46] que s'il existe $y \in \mathcal{X}$ tel que $0 \in \operatorname{int} My$ alors la suite $(x_n)_{n\in\mathbb{N}}$ engendrée par l'algorithme

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = \big((\Delta + \gamma_n M)^{-1} \circ \Delta\big)(x_n) \tag{1.7}$$

converge vers un point dans zer M en un nombre fini d'itérations. Notons qu'en raison de la convexité stricte de $\mathcal{X}^*$, la fonction $\|\cdot\|^2/2$ est Gâteaux différentiable et son gradient est Δ. Il en résulte que Δ pourrait être remplacé par le gradient d'une fonction convenable. Dans cette direction, [10] unifie et propose une généralisation de (1.3) dans les espaces de Banach réels réflexifs. Ces travaux utilisent les définitions ci-après qui se basent sur la notion de distance de Bregman, apparue initialement dans [20].

Définition 1.5 [10] Soit $\mathcal{X}$ un espace de Banach réel réflexif, soit $\mathcal{X}^*$ le dual topologique de $\mathcal{X}$, et soit $f \in \Gamma_0(\mathcal{X})$ une fonction Gâteaux différentiable sur $\operatorname{int}\operatorname{dom} f \neq \varnothing$. La distance de Bregman associée à f est

$$\begin{aligned} D^f \colon \mathcal{X} \times \mathcal{X} &\to [0,+\infty] \\ (x,y) &\mapsto \begin{cases} f(x) - f(y) - \langle x-y, \nabla f(y)\rangle, & \text{si } \ y \in \operatorname{int}\operatorname{dom} f; \\ +\infty, & \text{sinon.} \end{cases} \end{aligned} \tag{1.8}$$

La D^f-résolvante de $M \colon \mathcal{X} \to 2^{\mathcal{X}^*}$ est

$$R_M^f = (\nabla f + M)^{-1} \circ \nabla f \colon \mathcal{X} \to 2^{\mathcal{X}}, \tag{1.9}$$

l'opérateur D^f-proximal de $\varphi \in \Gamma_0(\mathcal{X})$ est

$$\operatorname{prox}_\varphi^f \colon \mathcal{X} \to 2^{\mathcal{X}} \colon y \mapsto \underset{x\in\mathcal{X}}{\operatorname{Argmin}} \ \big(\varphi(x) + D^f(x,y)\big), \tag{1.10}$$

et le D^f-projecteur sur $C \subset \mathcal{X}$ est

$$P_C^f \colon \mathcal{X} \to 2^{\mathcal{X}} \colon y \mapsto \underset{x\in C}{\operatorname{Argmin}} \ D^f(x,y). \tag{1.11}$$

Les distances de Bregman ne sont pas des distances au sens usuel, car elles ne sont pas symétriques. Leurs propriétés sont toutefois suffisamment proches de celles d'une «vraie» distance pour qu'on puisse construire ds nouvelles méthodes adaptées aux géométries non euclidiennes (voir [10, 11, 13, 29] et leurs bibliographies). Du point de vue des applications, elles apparaissent dans de nombreuses disciplines telles que la théorie de l'information, l'apprentissage, les statistiques, le traitement de l'image et du signal [7, 14, 28, 30, 31, 39, 63, 47].

La notion de monotonie au sens de Bregman a été introduite dans [10].

Définition 1.6 [10, Definition 1.2] Soit $\mathcal{X}$ un espace de Banach réel réflexif, soit $\mathcal{X}^*$ le dual topologique de $\mathcal{X}$, soit $f \in \Gamma_0(\mathcal{X})$ une fonction Gâteaux différentiable sur $\operatorname{int}\operatorname{dom} f \neq \varnothing$, et soit $C \subset \mathcal{X}$ tel que $C \cap \operatorname{dom} f \neq \varnothing$. Une suite $(x_n)_{n\in\mathbb{N}}$ dans $\mathcal{X}$ est monotone au sens de Bregman par rapport à C si elle satisfait les conditions suivantes :

(i) $(x_n)_{n\in\mathbb{N}}$ est dans $\operatorname{int}\operatorname{dom} f$.

(ii) $(\forall x \in C \cap \operatorname{dom} f)(\forall n \in \mathbb{N}) \quad D^f(x, x_{n+1}) \leqslant D^f(x, x_n)$.

Considérons le Problème 1.1 et supposons que $\operatorname{zer} M \neq \varnothing$. En utilisant le cadre de suites monotones au sens de Bregman, il est montré dans [10, Theorem 5.18] que sous des hypothèses convenables sur f et M, étant donnée une suite $(\gamma_n)_{n\in\mathbb{N}}$ dans $]0, +\infty[$ telle que $\inf_{n\in\mathbb{N}} \gamma_n > 0$, la suite engendrée par l'algorithme

$$x_0 \in \operatorname{int}\operatorname{dom} f \quad \text{et} \quad (\forall n \in \mathbb{N}) \quad x_{n+1} = R^f_{\gamma_n M} x_n \tag{1.12}$$

converge faiblement vers un point dans $\operatorname{zer} M$.

Les auteurs dans [11] proposent un algorithme du type Haugazeau/Bregman pour résoudre le Problème 1.1. Pour décrire ce résultat, soit $f \in \Gamma_0(\mathcal{X})$ Gâteaux différentiable sur $\operatorname{int}\operatorname{dom} f \neq \varnothing$. On définit, pour $(x, y, z) \in (\operatorname{int}\operatorname{dom} f)^3$,

$$H^f(x, y) = \big\{u \in \mathcal{X} \;\big|\; \langle u - y, \nabla f(x) - \nabla f(y)\rangle \leqslant 0\big\}, \tag{1.13}$$

et on désigne par $Q^f(x, y, z)$ le minimiseur de $f - \nabla f(x)$ sur $H^f(x, y) \cap H^f(y, z)$, i.e., la projection au sens de Bregman (1.11) de x sur $H^f(x, y) \cap H^f(y, z)$. Il est démontré dans [11] qu'étant donnés $x_0 \in \operatorname{int}\operatorname{dom} f$ et une suite $(\gamma_n)_{n\in\mathbb{N}}$ dans $]0, +\infty[$ telle que $\inf_{n\in\mathbb{N}} \gamma_n > 0$, sous des hypothèses convenables sur f, la suite engendrée par l'algorithme

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = Q^f\Big(x_0, x_n, R^f_{\gamma_n M} x_n\Big) \tag{1.14}$$

converge fortement vers $P^f_{\operatorname{zer} M} x_0$.

1.1.2 Méthodes d'éclatement d'opérateurs monotones

Le principal inconvénient de l'algorithme (1.12) est qu'en général les opérateurs $(R^f_{\gamma_n M})_{n\in\mathbb{N}}$ sont difficile, voire impossible, à évaluer. Par exemple lorsque $M = A + L^* B L$ où $A \colon \mathcal{X} \to 2^{\mathcal{X}^*}$ et $B \colon \mathcal{Y} \to 2^{\mathcal{Y}^*}$ sont monotones et $L \colon \mathcal{X} \to \mathcal{Y}$ est linéaire, une bonne méthode itérative devrait exploiter la structure de M et utiliser ces opérateurs séparément. Une telle méthode s'appelle une méthode d'éclatement d'opérateurs. Le problème générique est le suivant.

Problème 1.7 Soient $\mathcal{X}$ et $\mathcal{Y}$ des espaces de Banach réels réflexifs, soient $\mathcal{X}^*$ et $\mathcal{Y}^*$ les duaux topologiques de $\mathcal{X}$ et $\mathcal{Y}$, respectivement, soient $A \colon \mathcal{X} \to 2^{\mathcal{X}^*}$ et $B \colon \mathcal{Y} \to 2^{\mathcal{Y}^*}$ maximalement monotones, et soit $L \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$. Considérons l'inclusion primale

$$\text{trouver } x \in \mathcal{X} \text{ tel que } 0 \in Ax + L^* B(Lx), \tag{1.15}$$

et l'inclusion duale

$$\text{trouver } y \in \mathcal{Y}^* \text{ tel que } 0 \in -LA^{-1}(-L^* y^*) + B^{-1} y^*. \tag{1.16}$$

Les ensembles des solutions du Problème (1.15) et du Problème (1.16) sont notés $\mathscr{P}$ et $\mathscr{D}$, respectivement.

De nombreuses méthodes d'éclatement existent dans le cadre des espaces hilbertiens où l'on compte deux types de méthodes :

- Les méthodes classiques pour résoudre le problème primal sans composition $0 \in Ax + Bx$: algorithme explicite-implicite [55], algorithme de Douglas-Rachford [50], algorithme de Peaceman-Rachford [50], algorithme doublement implicite [49, 55], le méthode de l'inverse partiel de Spingarn [61] ; voir aussi [12].
- Les méthodes dites «primales-duales» pour résoudre le Problème 1.7 : [1, 2, 3, 17, 18, 24, 34, 36, 62].

Dans la suite, nous rappelons des résultats récents concernant des méthodes d'éclatement dans des espaces hilbertiens qui vont servir de base à certains travaux de la thèse. Dans [38], les auteurs construisent une extension de l'algorithme explicite-implicite classique où la métrique est autorisée à varier à chaque itération.

Théorème 1.8 [38, Theorem 4.1] *Dans le Problème 1.7, supposons que $\mathcal{X}$ soit un espace hilbertien, que $\mathcal{Y} = \mathcal{X}$, que $L = \mathrm{Id}$, que $Z = \operatorname{zer}(A+B) \neq \varnothing$, et que $B^{-1} - \beta\,\mathrm{Id}$ soit monotone pour un certain $\beta \in \left]0,+\infty\right[$. Soient $(\eta_n)_{n\in\mathbb{N}} \in \ell_+^1(\mathbb{N})$ et $(U_n)_{n\in\mathbb{N}}$ une suite d'opérateurs auto-adjoints dans $\mathcal{B}(\mathcal{X})$ tels que $\mu = \sup_{n\in\mathbb{N}} \|U_n\| < +\infty$ et*

$$(\forall n \in \mathbb{N})(\forall x \in \mathcal{X}) \quad \langle x, U_n x\rangle \geqslant \alpha\|x\|^2 \quad \textit{et} \quad (1+\eta_n)\langle x, U_{n+1}x\rangle \geqslant \langle x, U_n x\rangle. \tag{1.17}$$

Soient $\varepsilon \in \left]0, \min\{1, 2\beta/(\mu+1)\}\right]$, $(\lambda_n)_{n\in\mathbb{N}}$ une suite dans $[\varepsilon, 1]$, $(\gamma_n)_{n\in\mathbb{N}}$ une suite dans $[\varepsilon, (2\beta-\varepsilon)/\mu]$, $x_0 \in \mathcal{X}$, et $(a_n)_{n\in\mathbb{N}}$ et $(b_n)_{n\in\mathbb{N}}$ des suites absolument sommables dans $\mathcal{X}$. Posons

$$(\forall n \in \mathbb{N}) \quad \left\lfloor \begin{array}{l} y_n = x_n - \gamma_n U_n(Bx_n + b_n) \\ x_{n+1} = x_n + \lambda_n\big((\mathrm{Id} + \gamma_n U_n A)^{-1} y_n + a_n - x_n\big). \end{array} \right. \tag{1.18}$$

Alors, il existe $\overline{x} \in Z$ tel que les propriétés suivantes soient satisfaites :

(i) $x_n \rightharpoonup \overline{x}$.

(ii) $\sum_{n\in\mathbb{N}} \|Bx_n - B\overline{x}\|^2 < +\infty$.

(iii) *Supposons que l'une des conditions suivantes soit satisfaite.*

(a) $\underline{\lim}\, d_Z(x_n) = 0$.

(b) *En tout point dans Z, A ou B est demirégulier (i.e., en tout point $x \in Z$, si, pour toute suite $((x_n, x_n^*))_{n\in\mathbb{N}}$ dans $\operatorname{gra} A$ (respectivement $\operatorname{gra} B$) et pour tout $x^* \in Ax$ (respectivement $x^* \in Bx$) tels que $x_n \rightharpoonup x$ et $x_n^* \to x^*$, alors $x_n \to x$).*

(c) *$\operatorname{int} Z \neq \varnothing$ et il existe $(\nu_n)_{n\in\mathbb{N}} \in \ell_+^1(\mathbb{N})$ tel que $(\forall n \in \mathbb{N})$ $(\forall x \in \mathcal{X})$ $(1+\nu_n)\langle x, U_n x\rangle \geqslant \langle x, U_{n+1}x\rangle$.*

Alors $x_n \to \overline{x}$.

Dans l'approche primale-duale [2, 3, 24], on considère l'inclusion monotone composite et son inclusion duale simultanément. On leur associe l'ensemble de Kuhn-Tucker, et on construit des méthodes pour trouver un point dans cet ensemble. Nous rappelons les deux résultats suivants dans le cadre d'espaces hilbertiens. Le premier concerne les approximations fejériennes successives de l'ensemble de Kuhn-Tucker et le deuxième concerne la meilleure approximation d'un point initial $(x_0, y_0^*) \in \mathcal{X} \times \mathcal{Y}$ de $\boldsymbol{Z}$, i.e., la projection $P_{\boldsymbol{Z}}(x_0, y_0^*)$ de (x_0, y_0^*) sur $\boldsymbol{Z}$.

Proposition 1.9 [2, Proposition 3.5] *Dans le Problème 1.7, supposons que $\mathcal{X}$ et $\mathcal{Y}$ soient hilbertiens et désignons par*

$$\boldsymbol{Z} = \big\{(x, y^*) \in \mathcal{X} \times \mathcal{Y} \;\big|\; -L^*y^* \in Ax \quad \textit{et} \quad Lx \in B^{-1}y^*\big\} \tag{1.19}$$

l'ensemble de Kuhn-Tucker associé. Soient $(x_0, y_0^*) \in \mathcal{X} \times \mathcal{Y}$ et $\varepsilon \in \left]0, 1\right[$. On itère

$$\begin{array}{l}
\textit{pour } n = 0, 1, \ldots \\
\left\lfloor
\begin{array}{l}
(\gamma_n, \mu_n) \in [\varepsilon, 1/\varepsilon]^2 \\
a_n = (\mathrm{Id} + \gamma_n A)^{-1}(x_n - \gamma_n L^* y_n^*) \\
l_n = L x_n \\
b_n = (\mathrm{Id} + \mu_n B)^{-1}(l_n + \mu_n y_n^*) \\
s_n^* = \gamma_n^{-1}(x_n - a_n) + \mu_n^{-1} L^*(l_n - b_n) \\
t_n = b_n - L a_n \\
\tau_n = \|s_n^*\|^2 + \|t_n\|^2 \\
\textit{si } \tau_n = 0 \\
\left\lfloor
\begin{array}{l}
\overline{x} = a_n \\
\overline{y}^* = y_n^* + \mu_n^{-1}(l_n - b_n) \\
\textit{terminer}.
\end{array}
\right. \\
\lambda_n \in [\varepsilon, 2 - \varepsilon] \\
\theta_n = \lambda_n \big(\gamma_n^{-1}\|x_n - a_n\|^2 + \mu_n^{-1}\|l_n - b_n\|^2\big)/\tau_n \\
x_{n+1} = x_n - \theta_n s_n^* \\
y_{n+1}^* = y_n^* - \theta_n t_n.
\end{array}
\right.
\end{array} \tag{1.20}$$

Alors, soit (1.20) *termine à une solution $(\overline{x}, \overline{y}^*) \in \boldsymbol{Z}$ ou il engendre des suites infinies $(x_n)_{n\in\mathbb{N}}$ et $(y_n^*)_{n\in\mathbb{N}}$ et les propriétés suivantes sont satisfaites :*

(i) $\sum_{n\in\mathbb{N}} \|s_n^*\|^2 < +\infty$ *et* $\sum_{n\in\mathbb{N}} \|t_n\|^2 < +\infty$.

(ii) $\sum_{n\in\mathbb{N}} \|x_{n+1} - x_n\|^2 < +\infty$ *et* $\sum_{n\in\mathbb{N}} \|y_{n+1}^* - y_n^*\|^2 < +\infty$.

(iii) $\sum_{n\in\mathbb{N}} \|x_n - a_n\|^2 < +\infty$ *et* $\sum_{n\in\mathbb{N}} \|Lx_n - b_n\|^2 < +\infty$.

(iv) *$(x_n)_{n\in\mathbb{N}}$ converge faiblement vers un point $\overline{x} \in \mathscr{P}$, $(y_n^*)_{n\in\mathbb{N}}$ converge faiblement vers un point $\overline{y}^* \in \mathscr{D}$, et $(\overline{x}, \overline{y}^*) \in \boldsymbol{Z}$.*

Proposition 1.10 [3, Proposition 3.5] *Dans le Problème 1.7, supposons que $\mathcal{X}$ et $\mathcal{Y}$ soient hilbertiens et désignons par*

$$
\boldsymbol{Z} = \big\{(x, y^*) \in \mathcal{X} \times \mathcal{Y} \;\big|\; -L^*y^* \in Ax \quad \textit{et} \quad Lx \in B^{-1}y^*\big\} \tag{1.21}
$$

l'ensemble de Kuhn-Tucker associé. Soient $(x_0, y_0^*) \in \mathcal{X} \times \mathcal{Y}$ et $\varepsilon \in \left]0,1\right[$. On itère

$$
\begin{array}{l}
\textit{pour } n = 0, 1, \ldots\\
\left\lfloor
\begin{array}{l}
(\gamma_n, \mu_n) \in [\varepsilon, 1/\varepsilon]^2\\
a_n = (\mathrm{Id} + \gamma_n A)^{-1}(x_n - \gamma_n L^* y_n^*)\\
l_n = L x_n\\
b_n = (\mathrm{Id} + \mu_n B)^{-1}(l_n + \mu_n y_n^*)\\
s_n^* = \gamma_n^{-1}(x_n - a_n) + \mu_n^{-1} L^*(l_n - b_n)\\
t_n = b_n - L a_n\\
\tau_n = \|s_n^*\|^2 + \|t_n\|^2\\
\textit{si } \tau_n = 0\\
\left\lfloor\ \theta_n = 0\right.\\
\textit{si } \tau_n > 0\\
\left\lfloor
\begin{array}{l}
\lambda_n \in [\varepsilon, 1]\\
\theta_n = \lambda_n\big(\gamma_n^{-1}\|x_n - a_n\|^2 + \mu_n^{-1}\|l_n - b_n\|^2\big)/\tau_n
\end{array}\right.\\
x_{n+1/2} = x_n - \theta_n s_n^*\\
y_{n+1/2}^* = y_n^* - \theta_n t_n\\
\chi_n = \big\langle x_0 - x_n, x_n - x_{n+1/2}\big\rangle + \big\langle y_0^* - y_n^*, y_n^* - y_{n+1/2}^*\big\rangle\\
\mu_n = \|x_0 - x_n\|^2 + \|y_0^* - y_n^*\|^2\\
\nu_n = \|x_n - x_{n+1/2}\|^2 + \|y_n^* - y_{n+1/2}^*\|^2\\
\rho_n = \mu_n \nu_n - \chi_n^2\\
\textit{si } \rho_n = 0 \textit{ et } \chi_n \geqslant 0\\
\left\lfloor
\begin{array}{l}
x_{n+1} = x_{n+1/2}\\
y_{n+1}^* = y_{n+1/2}^*
\end{array}\right.\\
\textit{si } \rho_n > 0 \textit{ et } \chi_n \nu_n \geqslant \rho_n\\
\left\lfloor
\begin{array}{l}
x_{n+1} = x_0 + (1 + \chi_n/\nu_n)(x_{n+1/2} - x_n)\\
y_{n+1}^* = y_0^* + (1 + \chi_n/\nu_n)(y_{n+1/2}^* - y_n^*)
\end{array}\right.\\
\textit{si } \rho_n > 0 \textit{ et } \chi_n \nu_n < \rho_n\\
\left\lfloor
\begin{array}{l}
x_{n+1} = x_n + (\nu_n/\rho_n)\big(\chi_n(x_0 - x_n) + \mu_n(x_{n+1/2} - x_n)\big)\\
y_{n+1}^* = y_n^* + (\nu_n/\rho_n)\big(\chi_n(y_0^* - y_n^*) + \mu_n(y_{n+1/2}^* - y_n^*)\big).
\end{array}\right.
\end{array}\right.
\end{array}
\tag{1.22}
$$

Alors (1.22) engendre des suites infinies $(x_n)_{n\in\mathbb{N}}$ et $(y_n^*)_{n\in\mathbb{N}}$ et les propriétés suivantes sont satisfaites :

(i) $\sum_{n\in\mathbb{N}} \|x_{n+1} - x_n\|^2 < +\infty$ *et* $\sum_{n\in\mathbb{N}} \|y_{n+1}^* - y_n^*\|^2 < +\infty$.

(ii) $\sum_{n\in\mathbb{N}} \|s_n^*\|^2 < +\infty$ *et* $\sum_{n\in\mathbb{N}} \|t_n\|^2 < +\infty$.

(iii) $\sum_{n\in\mathbb{N}} \|x_n - a_n\|^2 < +\infty$ *et* $\sum_{n\in\mathbb{N}} \|Lx_n - b_n\|^2 < +\infty$.

(iv) $x_n \to \overline{x}$ *et* $y_n^* \to \overline{y}^*$*, où* $(\overline{x}, \overline{y}^*)$ *est la meilleure approximation de* (x_0, y_0^*) *dans* $\boldsymbol{Z}$.

1.2 Objectifs et motivations

Dans cette thèse, nous proposons des méthodes d'éclatement pour résoudre l'inclusion monotone composite primale-duale ci-après. Dans le cadre hilbertien, des algorithmes pour résoudre ce type de problème ont été introduits dans [24] et divers raffinements et extensions ont depuis été proposés [1, 2, 17, 36, 38]. Cependant, une inclusion monotone du type (1.1) est importante non seulement dans le cadre d'espaces hilbertiens mais aussi dans le cadre d'espaces de Banach réels. On trouve par exemple ce problème dans la théorie des équations aux dérivées partielles non-linéaires [25, 48, 26, 59, 66], dans le calcul des variations [27], en automatique [5], et en optimisation [16, 64] dans des espaces de Banach non-hilbertiens tels que les espaces de Lebesgue $L^p(\mathbb{R}^m)$, les espaces de Sobolev $W^p(\mathbb{R}^m)$, les espaces de Besov $B^s_{p,q}(\mathbb{R}^m)$ [66]. La théorie des opérateurs monotones non-linéaires dans des espaces de Banach réels demeure très active [6, 15, 16, 41, 60, 66].

Problème 1.11 Soient $\mathcal{X}$ et $\mathcal{Y}$ des espaces de Banach réels réflexifs tels que $\mathcal{X} \neq \{0\}$ et $\mathcal{Y} \neq \{0\}$, soient $\mathcal{X}^*$ et $\mathcal{Y}^*$ les duaux topologiques de $\mathcal{X}$ et $\mathcal{Y}$, respectivement, soient $A\colon \mathcal{X} \to 2^{\mathcal{X}^*}$ et $B\colon \mathcal{Y} \to 2^{\mathcal{Y}^*}$ maximalement monotones, et soit $L \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$. On considère l'inclusion primale

$$\text{trouver } x \in \mathcal{X} \text{ tel que } 0 \in Ax + L^* BLx \tag{1.23}$$

et l'inclusion duale

$$\text{trouver } y^* \in \mathcal{Y}^* \text{ tel que } 0 \in -LA^{-1}(-L^* y^*) + B^{-1} y^*. \tag{1.24}$$

On désigne par $\mathscr{P}$ l'ensemble des solutions de (1.23), par $\mathscr{D}$ l'ensemble des solutions de (1.24), et par

$$\boldsymbol{Z} = \big\{(x, y^*) \in \mathcal{X} \times \mathcal{Y}^* \;\big|\; -L^* y^* \in Ax \quad \text{et} \quad Lx \in B^{-1} y^*\big\} \tag{1.25}$$

l'ensemble de Kuhn-Tucker associé à (1.23)–(1.24).

Dans cette thèse, nous étudions aussi une synthèse des notions de monotonie fejérienne avec métriques variables [37] et de monotonie par rapport à une distance de Bregman au sens de la Définition 1.5. Ensuite, nous proposons un algorithme d'éclatement pour résoudre le problème suivant qui consiste à minimiser la somme composite pour deux fonctions convexes et un opérateur linéaire, où l'une des fonctions est différentiable.

Problème 1.12 Soient $\mathcal{X}$ et $\mathcal{Y}$ des espaces de Banach réels réflexifs, soit $\varphi \in \Gamma_0(\mathcal{X})$, soit $\psi \in \Gamma_0(\mathcal{Y})$ Gâteaux différentiable sur $\operatorname{int}\operatorname{dom}\psi \neq \varnothing$, et soit $L \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$. Le problème est de

$$\underset{x \in \mathcal{X}}{\text{minimiser}} \;\; \varphi(x) + \psi(Lx). \tag{1.26}$$

La convergence de l'algorithme proposé est démontrée en utilisant l'analyse d'une suite monotone par rapport à des distances de Bregman variables. Nous déduisons, dans le cadre d'espaces euclidiens, des algorithmes nouveaux et les appliquons en traitement de l'image avec les modèles proposés dans [28, 43, 51]. Plus généralement, notre cadre permet de résoudre le problème multivarié suivant.

Problème 1.13 Soient m et p deux entiers strictement positifs, soient $(\mathcal{X}_i)_{1\leqslant i\leqslant m}$ des espaces de Banach réels réflexifs, et soient $(\mathcal{Y}_k)_{1\leqslant k\leqslant p}$ des espaces de Banach réels. Pour tous $i \in \{1,\ldots,m\}$ et $k \in \{1,\ldots,p\}$, soient $\varphi_i \in \Gamma_0(\mathcal{X}_i)$ et $\psi_k \in \Gamma_0(\mathcal{Y}_k)$, et soit $L_{ki} \in \mathcal{B}(\mathcal{X}_i, \mathcal{Y}_k)$ non nul. Le problème est de

$$\underset{x_1\in\mathcal{X}_1,\ldots,x_m\in\mathcal{X}_m}{\text{minimiser}} \quad \sum_{i=1}^{m} \varphi_i(x_i) + \sum_{k=1}^{p} \psi_k\left(\sum_{i=1}^{m} L_{ki}x_i\right). \tag{1.27}$$

Du point de vue du développement de méthodes itératives, la principale difficulté à laquelle on est confronté dans les espaces de Banach non-hilbertiens est que des méthodes aussi élémentaires que (1.3) n'ont pas de sens puisque M est à valeurs dans l'espace dual $\mathcal{X}^*$ et qu'il est donc impossible d'effectuer le calcul $\mathrm{Id} + M$ ni dans $\mathcal{X}$ ni dans $\mathcal{X}^*$. De plus, l'opérateur proximal défini par (1.4) est difficile voire impossible à manipuler à cause de la non-linéarité de l'application de dualité. En revanche, les distances de Bregman donnent lieu à des outils biens plus flexibles et effectivement applicables dans ces espaces [10]. Une motivation pour nos travaux est d'utiliser ce type de distances pour construire des algorithmes d'éclatement dans les espaces de Banach réels réflexifs. Cette étude nous permet non seulement de résoudre des inclusions monotones dans ces espaces mais aussi de construire un cadre qui unifie les algorithmes existants dans des espaces hilbertiens réels. En particulier, dans le cadre d'espaces euclidiens, ces résultats nous donnent des algorithmes nouveaux.

Dans un espaces de Banach réel non-hilbertien, peu d'algorithmes sont disponibles pour résoudre les Problèmes 1.12 et 1.11, et il n'existe aucun algorithme pour résoudre le Problème 1.13. Un algorithme est proposé dans [19] pour résoudre un cas particulier du Problème 1.12 dans lequel $L = \mathrm{Id}$. Cet algorithme impose que $\nabla\psi$ soit $(p-1)$-höldérien sur les sous-ensembles bornés de $\mathcal{X}$ pour un certain $p \in \,]1,2]$ et il se décrit brièvement ainsi : étant donné x_n, l'itérée suivante est déterminée en résolvant le problème

$$\underset{x\in\mathcal{X}}{\text{minimiser}} \quad \frac{\|x-x_n\|^p}{p} + \tau_n\big(\langle x, \nabla\psi(x_n)\rangle + \varphi(x)\big), \tag{1.28}$$

où $(\tau_n)_{n\in\mathbb{N}}$ est une suite de paramètres convenables. Les résultats obtenus sont la propriété descente de $\big((\varphi+\psi)(x_n)\big)_{n\in\mathbb{N}}$. La convergence de $(x_n)_{n\in\mathbb{N}}$ n'est pas assurée, sauf quand $\varphi+\psi$ admet un unique minimiseur. L'utilisation du terme $\|\cdot\|^p/p$ dans (1.28) d'une part, rend ce problème difficile à résoudre, et d'autre part, implique que cet algorithme n'est pas une vraie technique d'éclatement. Une de nos stratégies est d'utiliser les distances de Bregman au lieu du terme $\|\cdot\|^p/p$ et nous proposons des algorithmes d'éclatement plus avantageux numériquement et démontrons leur convergence.

Dans le cas particulier où $L = \mathrm{Id}$ dans (1.23), un algorithme d'éclatement est proposé dans [44]. Cet algorithme trouve un point dans

$$S_e(A,B) = \big\{(x,y^*) \in \mathcal{X} \times \mathcal{X}^* \;\big|\; y^* \in Ax \text{ et } -y^* \in Bx\big\} \tag{1.29}$$

par la méthode qui se décrit brièvement ainsi :

— (x_n, y_n^*) est donné.

— Deux paramètres γ_n et μ_n sont choisis dans $]0,+\infty[$.

— Des points $(a_n, a_n^*) \in \operatorname{gra} A$ et $(b_n, b_n^*) \in \operatorname{gra} B$ sont obtenus en utilisant les opérateurs $(\nabla f + \gamma_n A)^{-1}$ et $(\nabla f + \mu_n B)^{-1}$.

— Un demi-espace affine $\boldsymbol{H}_n \subset \mathcal{X} \times \mathcal{X}^*$ séparant (x_n, y_n^*) de $S_e(A,B)$, est construit en utilisant (a_n, a_n^*) et (b_n, b_n^*).

— L'itérée suivante est déterminée par la résolution du problème

$$\underset{(x,y^*)\in \boldsymbol{H}_n}{\text{minimiser}} \; D^f(x,x_n) + D^g(y,y_n^*), \tag{1.30}$$

où $g\colon \mathcal{X}^* \to \,]-\infty,+\infty]$ est une fonction convenablement choisie.

Les auteurs montrent que $\big((x_n,y_n^*)\big)_{n\in\mathbb{N}}$ converge faiblement vers un point $(\overline{x},\overline{y}^*)$ dans $S_e(A,B)$ et donc, $\overline{x}$ est une solution de (1.23). Notons que $S_e(A,B)$ est en fait une spécialisation de $\boldsymbol{Z}$ défini dans (1.25) et donc, un point dans cet ensemble donne une solution primale-duale. Cette observation n'est pas faite dans [44] mais elle est exploitée dans [2, 3]. De plus, les fonctions f et g utilisées dans cet algorithme sont excessivement restrictives, par exemple, elles doivent être supercoercives, uniformément convexes avec des gradients faiblement séquentiellement continus. Beaucoup de fonctions utiles qui apparaissent en traitement de l'image et du signal comme les entropies ne vérifient pas ces conditions. En utilisant les distances de Bregman, combinées avec l'approche hilbertienne de [2, 3], nous visons à construire des méthodes pour résoudre le Problème 1.11 et établir leur convergence sous des conditions très générales ; en particulier, nous pourrons utiliser une classe de fonctions qui contient les principales notions d'entropie.

Finalement, une autre motivation pour nos travaux vient du point vue des applications et des aspects numériques pour des problèmes posés dans des espaces euclidiens. Dans ces espaces, les techniques basées sur la distance euclidienne ne sont pas toujours facilement applicables. Par exemple, afin d'utiliser l'algorithme explicite-implicite pour trouver un minimiseur de la somme de deux fonctions convexes, il faut qu'une fonction soit différentiable avec un gradient globalement lipschitzien et cette condition n'est pas vérifiée dans certains cas, par exemple, en traitement de l'image avec des modèles basés sur des entropies [14, 28, 43, 51]. De plus, pour certaines classes de fonctions, le calcul de l'opérateur proximal par rapport à une distance de Bregman défini par (1.10) s'avère plus avantageux numériquement que sa variante classique définie par (1.4).

1.3 Organisation

Au Chapitre 2, nous proposons un algorithme pour résoudre le Problème 1.13 en trouvant dans $\boldsymbol{Z}$ la meilleure approximation pour une distance de Bregman d'un point de référence.

Au Chapitre 3, nous étudions des variantes de l'algorithme du Chapitre 2 en utilisant les distances de Bregman en présence d'erreurs.

Au Chapitre 4, nous présentons un algorithme pour résoudre le Problème 1.13 en trouvant un point dans $\boldsymbol{Z}$ par les approximations bregmaniennes successives.

Au Chapitre 5, nous introduisons la notion de suites quasi-monotones au sens de Bregman et nous analysons leur comportement asymptotique. Nous appliquons ces résultats pour analyser le comportement asymptotique d'un algorithme du point proximal basé sur des distances de Bregman variables et d'un algorithme pour résoudre les problèmes d'admisibilité convexe dans des espaces de Banach réels réflexifs.

Au Chapitre 6, nous proposons une version de l'algorithme explicite-implicite dans des espaces de Banach réels réflexifs pour résoudre le Problème 1.11 et présentons des applications du modèle en restauration d'images.

Au Chapitre 7, nous étudions les notions de contractivité basées sur les distances de Bregman dans des espaces de Banach. Nous unifions et étudions une notion de cocoercivité en utilisant les distances de Bregman. Un algorithme d'approximations successives pour trouver un point fixe commun à une famille de contractions est aussi considéré. Nous présentons également une version particulière de l'algorithme explicite-implicite pour résoudre (1.23) lorsque B est le gradient d'une fonction convexe différentiable.

Au Chapitre 8, nous présentons nos conclusions et des questions ouvertes.

1.4 Contributions principales

- Conception des premières méthodes pour résoudre les inclusions monotones primales-duales composites dans des espaces de Banach réels réflexifs.
- Conception et étude d'un algorithme pour trouver la meilleure approximation pour une distance de Bregman d'un point de référence dans l'ensemble Kuhn-Tucker associé à des inclusions monotones composites primales-duales dans des espaces de Banach réels réflexifs.
- Introduction de la notion de suites quasi-monotones au sens de Bregman qui unifie la notion de suites quasi-fejériennes à métrique variable et celle de suites monotones au sens de Bregman. Les résultats obtenus sont les outils fondamentaux pour démontrer la convergence de processus itératifs dans des espaces de Banach réels réflexifs.

- Conception et étude asymptotique de l'algorithme du point proximal basé sur les distances de Bregman variables qui, d'une part, unifie les algorithmes du point proximal existants et d'autre part, donne des algorithmes nouveaux dans des espaces euclidiens.
- Conception et étude asymptotique de l'algorithme explicite-implicite pour résoudre les problèmes de minimisation convexe. Il constitue une nouvelle méthode d'éclatement dans des espaces de Banach réels. D'autre part, il donne lieu à de nouveaux algorithmes lorsqu'on l'appliquera au cadre d'espaces euclidiens.
- Analyse de diverses notions de contractivité et de cocoercivité basées sur les distances de Bregman qui nous permettent de proposer des algorithmes d'éclatement dans des espaces de Banach réels réflexifs.
- Unification des méthodes itératives pour trouver un point fixe commun à une famille de contractions dans des espaces de Banach réels réflexifs.

1.5 Bibliographie

[1] M. A. Alghamdi, A. Alotaibi, P. L. Combettes, and N. Shahzad, A primal-dual method of partial inverses for composite inclusions, *Optim. Lett.,* vol. 8, pp. 2271–2284, 2014.

[2] A. Alotaibi, P. L. Combettes, and N. Shahzad, Solving coupled composite monotone inclusions by successive Fejér approximations of their Kuhn-Tucker set, *SIAM J. Optim.,* vol. 24, pp. 2076–2095, 2014.

[3] A. Alotaibi, P. L. Combettes, and N. Shahzad, Best approximation from the Kuhn-Tucker set of composite monotone inclusions, *Numer. Funct. Anal. Optim.,* to appear.

[4] J.-P. Aubin and H. Frankowska, *Set-Valued Analysis.* Birkhäuser, Boston, MA, 1990.

[5] V. Barbu, *Analysis and Control of Nonlinear Infinite Dimensional Systems.* Academic Press, Boston, MA, 1993.

[6] V. Barbu, *Nonlinear Differential Equations of Monotone Types in Banach Spaces.* Springer, New York, 2010.

[7] M. Basseville, Divergence measures for statistical data processing : An annotated bibliography, *Signal Process.,* vol. 93, pp. 621–633, 2013.

[8] H. H. Bauschke and J. M. Borwein, On projection algorithms for solving convex feasibility problems, *SIAM Rev.,* vol. 38, pp. 367–426, 1996.

[9] H. H. Bauschke, J. M. Borwein, and P. L. Combettes, Essential smoothness, essential strict convexity and Legendre functions in Banach spaces, *Commun. Contemp. Math.,* vol. 3, pp. 615–647, 2001.

[10] H. H. Bauschke, J. M. Borwein, and P. L. Combettes, Bregman monotone optimization algorithms, *SIAM J. Control Optim.,* vol. 42, pp. 596–636, 2003.

[11] H. H. Bauschke and P. L. Combettes, Construction of best Bregman approximations in reflexive Banach spaces, *Proc. Amer. Math. Soc.,* vol. 131, pp. 3757–3766, 2003.

[12] H. H. Bauschke and P. L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces.* Springer, New York, 2011.

[13] H. H. Bauschke, P. L. Combettes, and D. Noll, Joint minimization with alternating Bregman proximity operators, *Pac. J. Optim.,* vol. 2, pp. 401–424, 2006.

[14] M. Bertero, P. Boccacci, G. Desiderà, and G. Vicidomini, Image deblurring with Poisson data : from cells to galaxies, *Inverse Problems,* vol. 25, Article ID 123006, 26 pp., 2009.

[15] J. M. Borwein, Fifty years of maximal monotonicity, *Optim. Lett.,* vol. 4, pp. 473–490, 2010.

[16] J. M. Borwein and Q. J. Zhu, *Techniques of Variational Analysis.* Springer, New York, 2005.

[17] R. I. Boţ, E. R. Csetnek, and A. Heinrich, A primal-dual splitting algorithm for finding zeros of sums of maximal monotone operators, *SIAM J. Optim.,* vol. 23, pp. 2011–2036, 2013.

[18] R. I. Boţ and C. Hendrich, A Douglas-Rachford type primal-dual method for solving inclusions with mixtures of composite and parallel-sum type monotone operators, *SIAM J. Optim.,* vol. 23, pp. 2541–2565, 2013.

[19] K. Bredies, A forward-backward splitting algorithm for the minimization of non-smooth convex functionals in Banach space, *Inverse Problems,* vol 25, Article ID 015005, 20 pp., 2009.

[20] L. M. Bregman, The relaxation method of finding the common point of convex sets and its application to the solution of problems in convex programming, *USSR Comput. Math. Math. Phys.*, vol 7, pp. 200–217, 1967.

[21] H. Brézis, *Opérateurs Maximaux Monotones et Semi-groupes de Contractions dans les Espaces de Hilbert.* North-Holland Publishing Co., Amsterdam, 1973.

[22] H. Brézis, *Analyse Fonctionnelle.* Masson, Paris, 1983.

[23] H. Brézis and P. L. Lions, Produits infinis de résolvantes, *Israel J. Math.,* vol. 29, pp. 329–345, 1978.

[24] L. M. Briceño-Arias and P. L. Combettes, A monotone+skew splitting model for composite monotone inclusions in duality, *SIAM J. Optim.*, vol. 21, pp. 1230–1250, 2011.

[25] F. E. Browder, Nonlinear elliptic boundary value problems, *Bull. Amer. Math. Soc.,* vol. 69, pp. 862–874, 1963.

[26] F. E. Browder, Nonlinear elliptic boundary value problems II, *Trans. Amer. Math. Soc.,* vol. 117, pp. 530–550, 1965.

[27] F. E. Browder, On the unification of the calculus of variations and the theory of monotone nonlinear operators in Banach spaces, *Proc. Nat. Acad. Sci. U.S.A.,* vol. 56, pp. 419–425, 1966.

[28] C. L. Byrne, Iterative image reconstruction algorithms based on cross-entropy minimization, *IEEE Trans. Image Process.,* vol. 2, pp. 96–103, 1993.

[29] Y. Censor and S. Zenios, Proximal minimization algorithm with D-functions, *J. Optim. Theory Appl.,* vol. 73, pp. 451–464, 1992.

[30] V. Chandrasekaran, P. A. Parrilo, and A. S. Willsky, Latent variable graphical model selection via convex optimization, *Ann. Statist.,* vol. 40, pp. 1935–1967, 2012.

[31] M. Collins, R. E. Schapire, and Y. Singer, Logistic regression, AdaBoost and Bregman distances, *Mach. Learn.,* vol. 48, pp. 253–285, 2002.

[32] P. L. Combettes, Quasi-Fejérian analysis of some optimization algorithms, in : *Inherently Parallel Algorithms for Feasibility and Optimization,* (D. Butnariu, Y. Censor, and S. Reich, eds.), pp. 115–152. Elsevier, New York, 2001.

[33] P. L. Combettes, Fejér monotonicity in convex optimization, in : *Encyclopedia of Optimization,* 2nd ed. (C. A. Floudas and P. M. Pardalos, eds.), pp. 1016–1024. Springer, New York, 2009.

[34] P. L. Combettes, Systems of structured monotone inclusions : Duality, algorithms, and applications, *SIAM J. Optim.,* vol. 23, pp. 2420–2447, 2013.

[35] P. L. Combettes and J.-C. Pesquet, Proximal splitting methods in signal processing, in *Fixed-Point Algorithms for Inverse Problems in Science and Engineering,* (H. H. Bauschke et al., eds), pp. 185–212. Springer, New York, 2011.

[36] P. L. Combettes and J.-C. Pesquet, Primal-dual splitting algorithm for solving inclusions with mixtures of composite, Lipschitzian, and parallel-sum type monotone operators, *Set-Valued Var. Anal.*, vol. 20, pp. 307–330, 2012.

[37] P. L. Combettes and B. C. Vũ, Variable metric quasi-Fejér monotonicity, *Nonlinear Anal.,* vol. 78, pp. 17–31, 2013.

[38] P. L. Combettes and B. C. Vũ, Variable metric forward-backward splitting with applications to monotone inclusions in duality, *Optimization,* vol. 63, pp. 1289–1318, 2014.

[39] I. S. Dhillon and J. A. Tropp, Matrix nearness problems with Bregman divergences, *SIAM J. Matrix Anal. Appl.,* vol. 29, pp. 1120–1146, 2007.

[40] Yu. M. Ermol'ev, On convergence of random quasi-Fejér sequences, *Cybernetics,* vol. 7, pp. 655–656, 1971.

[41] N. Ghoussoub, *Self-dual Partial Differential Systems and Their Variational Principles.* Springer, New York, 2009.

[42] Y. Haugazeau, *Sur les Inéquations Variationnelles et la Minimisation de Fonctionnelles Convexes.* Thèse, Université de Paris, Paris, France, 1968.

[43] L. Henri, R. Murel, and A. Claude, Penalized maximum likelihood image restoration with positivity constraints : multiplicative algorithms, *Inverse Problems,* vol. 18, pp. 1397–1419, 2002.

[44] A. N. Iusem and B. F. Svaiter, Splitting methods for finding zeroes of the sum of maximal monotone operators in Banach spaces, *J. Nonlinear Convex Anal.,* vol. 15, pp. 379–397, 2014.

[45] R. I. Kačurovskiĭ, On monotone operators and convex functionals, *Uspekhi Mat. Nauk*, vol. 15, pp. 213–215, 1960.

[46] G. Kassay, The proximal points algorithm for reflexive Banach spaces, *Studia Univ. Babeş-Bolyai Math.,* vol. 30, pp.9–17, 1985.

[47] J. Kivinen and M. K. Warmuth, Relative loss bounds for multidimensional regression problems, *Mach. Learn.,* vol. 45, pp. 301–329, 2001.

[48] J. Leray and J. L. Lions, Quelques résulatats de Višik sur les problèmes elliptiques nonlinéaires par les méthodes de Minty-Browder, *Bull. Soc. Math. France,* vol. 93, pp. 97–107, 1965.

[49] P. L. Lions, Une méthode itérative de résolution d'une inéquation variationnelle, *Israel J. Math.,* vol. 31, pp. 204–208, 1978.

[50] P. L. Lions and B. Mercier, Splitting algorithms for the sum of two nonlinear operators, *SIAM J. Numer. Anal.,* vol. 16, pp. 964–979, 1979.

[51] J. Markham and J. A. Conchello, Fast maximal-likelihood image-restoration algorithms for three dimensional fluorescence microscopy, *J. Opt. Soc. Am. A,* vol. 18, pp. 1062–1071, 2001.

[52] B. Mercier, *Topics in Finite Element Solution of Elliptic Problems* (Lectures on Mathematics, no. 63). Tata Institute of Fundamental Research, Bombay, 1979.

[53] G. J. Minty, On the maximal domain of a "monotone" function, *Michigan Math. J.*, vol. 8, pp. 135–137, 1961.

[54] G. J. Minty, Monotone (nonlinear) operators in Hilbert space, *Duke Math. J.*, vol. 29, pp. 341–346, 1962.

[55] G. B. Passty, Ergodic convergence to a zero of the sum of monotone operators in Hilbert space, *J. Math. Anal. Appl.,* vol. 72, pp. 382–390, 1979.

[56] T. Pennanen, Dualization of generalized equations of maximal monotone type, *SIAM J. Optim.,* vol. 10, pp. 809–835, 2000.

[57] J. Peypouquet and S. Sorin, Evolution equations for maximal monotone operators : asymptotic analysis in continuous and discrete time, *J. Convex Anal.,* vol. 17, pp. 1113–1163, 2010.

[58] E. Raik, Fejér type methods in Hilbert space, *Eesti NSV Tead. Akad. Toimetised Füüs.-Mat.,* vol. 16, pp. 286–293, 1967.

[59] R. E. Showalter, *Monotone Operators in Banach Space and Nonlinear Partial Differential Equations,* American Mathematical Society, Providence, RI, 1997.

[60] S. Simons, *From Hahn-Banach to Monotonicity,* Lecture Notes in Math. 1693, Springer-Verlag, New York, 2008.

[61] J. E. Spingarn, Partial inverse of a monotone operator, *Appl. Math. Optim.*, vol. 10, pp.247–265, 1983.

[62] B. C. Vũ, A splitting algorithm for dual monotone inclusions involving cocoercive operators, *Adv. Comput. Math.*, vol. 38, pp. 667–681, 2013.

[63] J. Yun, D. Sun, and K. C. Toh, A proximal point algorithm for log-determinant optimization with group Lasso regularization, *SIAM J. Optim.,* vol. 23, pp. 857–893, 2013.

[64] C. Zălinescu, *Convex Analysis in General Vector Spaces.* World Scientific Publishing Co., Inc., River Edge, NJ, 2002.

[65] E. H. Zarantonello, *Solving Functional Equations by Contractive Averaging*, Tech. Report 160, Math. Res. Center U.S. Army, Madison, University of Wisconsin, June 1960.

[66] E. Zeidler, *Nonlinear Functional Analysis and Its Applications,* vols. I–V. Springer-Verlag, New York, 1985–1993.

Chapitre 2

Algorithme de meilleure approximation des points de Kuhn-Tucker d'inclusions monotones

Nous proposons la première méthode d'éclatement pour des inclusions monotones composites primales-duales en dehors des espaces hilbertiens. La méthode primale-duale proposée construit successivement la meilleure approximation au sens de Bregman d'un point de référence dans l'ensemble de Kuhn-Tucker associé à une inclusion monotone composite. La convergence forte est établie dans des espaces de Banach réels réflexifs sans exiger de restriction supplémentaire sur les opérateurs monotones ou la connaissance de la norme de l'opérateur linéaire impliqué dans le modèle. Les opérateurs monotones sont activés via des résolvantes basées sur les distances de Bregman. La méthode est nouvelle même dans des espaces euclidiens, où elle fournit une alternative aux méthodes proximales basées sur les distances euclidiennes.

2.1 Description et résultats principaux

Nos résultats sont basés sur les définitions suivantes.

Définition 2.1 [7] Soit $\mathcal{X}$ un espace de Banach réel réflexif. Une fonction $f \in \Gamma_0(\mathcal{X})$ est *de Legendre* si elle est *essentiellement lisse* au sens où ∂f est à la fois localement borné et univoque sur son domaine, et *essentiellement strictement convexe* au sens où ∂f^* est localement borné sur son domaine et f est strictement convexe sur les sous-ensembles convexes de $\operatorname{dom} \partial f$. De plus, f est Gâteaux différentiable sur $\operatorname{int} \operatorname{dom} f \neq \varnothing$

et la distance de Bregman associée à f est

$$D^f\colon \mathcal{X}\times\mathcal{X} \to [0,+\infty]$$
$$(x,y) \mapsto \begin{cases} f(x)-f(y)-\big\langle x-y, \nabla f(y)\big\rangle, & \text{si } \ y\in \operatorname{int}\operatorname{dom} f;\\ +\infty, & \text{sinon.}\end{cases} \tag{2.1}$$

Soit C un sous-ensemble convexe fermé de $\mathcal{X}$ tel que $C\cap\operatorname{int}\operatorname{dom} f \neq \varnothing$. Le D^f*-projecteur*

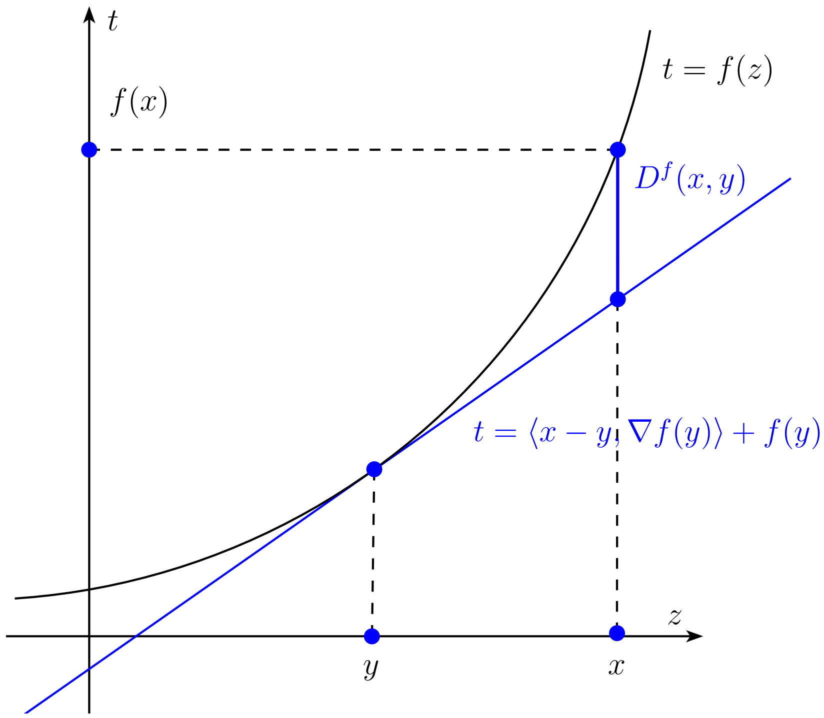

FIGURE 2.1 – Distance de Bregman.

sur C est

$$P^f_C\colon \operatorname{int}\operatorname{dom} f \to C\cap \operatorname{int}\operatorname{dom} f$$
$$y \mapsto \underset{x\in C}{\operatorname{argmin}}\, D^f(x,y). \tag{2.2}$$

Posons

$$(\forall (x,y)\in(\operatorname{int}\operatorname{dom} f)^2)\quad H^f(x,y)=\big\{z\in\mathcal{X} \;\big|\; \big\langle z-y, \nabla f(x)-\nabla f(y)\big\rangle \leqslant 0\big\} \tag{2.3}$$

et soit $(x,y,z)\in(\operatorname{int}\operatorname{dom} f)^3$ tel que $H^f(x,y)\cap H^f(y,z)\cap\operatorname{int}\operatorname{dom} f\neq\varnothing$. Alors

$$Q^f(x,y,z) = P^f_{H^f(x,y)\cap H^f(y,z)}x. \tag{2.4}$$

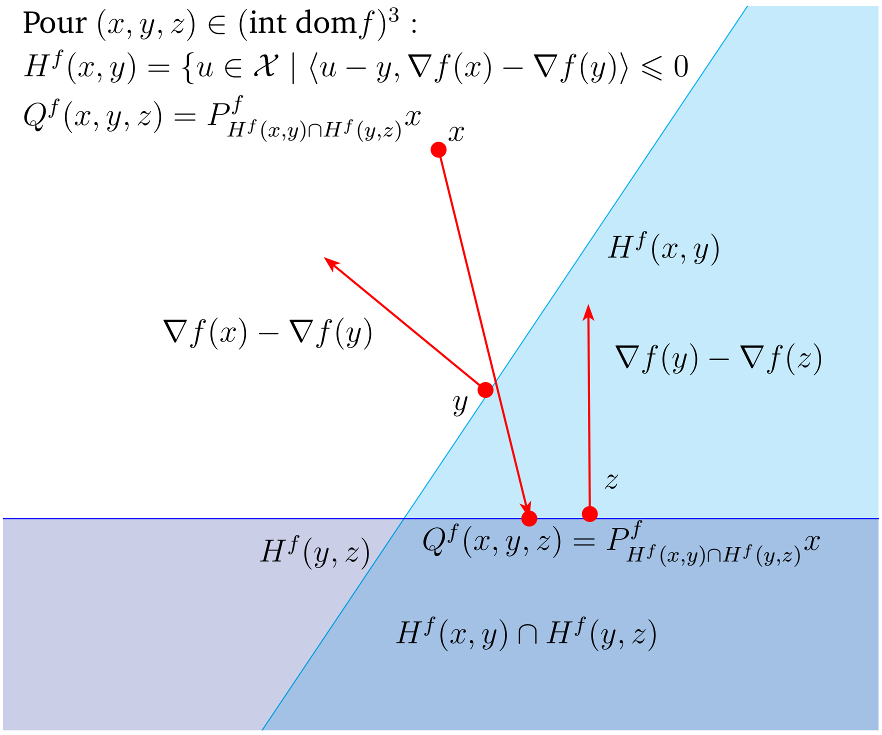

FIGURE 2.2 – H^f et Q^f.

Dans ce chapitre, nous considérons le problème suivant.

Problème 2.2 Soient $\mathcal{X}$ et $\mathcal{Y}$ des espaces de Banach réels réflexifs tels que $\mathcal{X} \neq \{0\}$ et $\mathcal{Y} \neq \{0\}$, soit $\boldsymbol{\mathcal{X}}$ l'espace vectoriel $\mathcal{X} \times \mathcal{Y}^*$ muni de la norme $(x,y^*) \mapsto \sqrt{\|x\|^2 + \|y^*\|^2}$, et soit $\boldsymbol{\mathcal{X}}^*$ le dual topologique de $\boldsymbol{\mathcal{X}}$, i.e., $\mathcal{X}^* \times \mathcal{Y}$ muni de la norme $(x^*,y) \mapsto \sqrt{\|x^*\|^2 + \|y\|^2}$. Soient $A\colon \mathcal{X} \to 2^{\mathcal{X}^*}$ et $B\colon \mathcal{Y} \to 2^{\mathcal{Y}^*}$ maximalement monotones et soit $L \in \mathcal{B}(\mathcal{X},\mathcal{Y})$. Considérons l'inclusion primale

$$\text{trouver } x \in \mathcal{X} \text{ tel que } 0 \in Ax + L^*BLx, \tag{2.5}$$

l'inclusion duale

$$\text{trouver } y^* \in \mathcal{Y}^* \text{ tel que } 0 \in -LA^{-1}(-L^*y^*) + B^{-1}y^*, \tag{2.6}$$

et l'ensemble de Kuhn-Tucker associé

$$\boldsymbol{Z} = \big\{(x,y^*) \in \boldsymbol{\mathcal{X}} \;\big|\; -L^*y^* \in Ax \quad \text{et} \quad Lx \in B^{-1}y^*\big\}. \tag{2.7}$$

Soient $f \in \Gamma_0(\mathcal{X})$ et $g \in \Gamma_0(\mathcal{Y})$ des fonctions de Legendre, soit $x_0 \in \text{int dom} f$, et soit $y_0^* \in \text{int dom}\, g^*$. Posons

$$\boldsymbol{f}\colon \boldsymbol{\mathcal{X}} \to]-\infty,+\infty] : (x,y^*) \mapsto f(x) + g^*(y^*), \tag{2.8}$$

et supposons que $\boldsymbol{Z} \cap \operatorname{int} \operatorname{dom} \boldsymbol{f} \neq \varnothing$. Le problème est de trouver la meilleure approximation au sens de Bregman $(\overline{x}, \overline{y}^*) = P_{\boldsymbol{Z}}^{\boldsymbol{f}}(x_0, y_0^*)$ de (x_0, y_0^*) dans $\boldsymbol{Z}$.

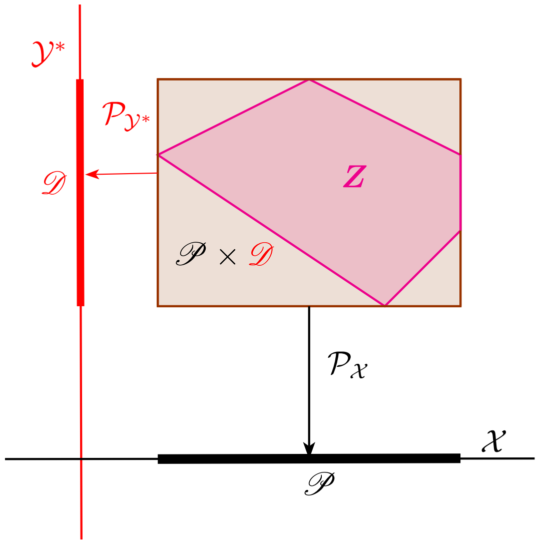

FIGURE 2.3 – Ensemble de Kuhn-Tucker.

Nous utilisons la condition suivante dans la suite.

Condition 2.3 Soit $\mathcal{X}$ un espace de Banach réel réflexif et soit $h \in \Gamma_0(\mathcal{X})$ Gâteaux différentiable sur $\operatorname{int} \operatorname{dom} h \neq \varnothing$. Pour toute suite $(x_n)_{n\in\mathbb{N}}$ dans $\operatorname{int} \operatorname{dom} h$ et toute suite bornée $(y_n)_{n\in\mathbb{N}}$ dans $\operatorname{int} \operatorname{dom} h$,

$$D^h(x_n, y_n) \to 0 \quad \Rightarrow \quad x_n - y_n \to 0. \tag{2.9}$$

Pour énoncer nos résultats, nous rappelons la notion de coercivité d'un opérateur multivoque.

Définition 2.4 Soit $\mathcal{X}$ un espace de Banach réel réflexif tel que $\mathcal{X} \neq \{0\}$, soit $\mathcal{X}^*$ le dual topologique de $\mathcal{X}$, et soit $M \colon \mathcal{X} \to 2^{\mathcal{X}^*}$. Alors M est *coercif* si

$$(\exists\, z \in \operatorname{dom} M) \quad \lim_{\|x\|\to+\infty} \inf \frac{\langle x - z, Mx\rangle}{\|x\|} = +\infty. \tag{2.10}$$

Le résultat principal dans ce chapitre est le théorème suivant.

Théorème 2.5 *Considérons le Problème 2.2. Soient $h \in \Gamma_0(\mathcal{X})$ et $j \in \Gamma_0(\mathcal{Y})$ des fonctions de Legendre telles que* $\operatorname{int} \operatorname{dom} f \subset \operatorname{int} \operatorname{dom} h$, $L(\operatorname{int} \operatorname{dom} f) \subset \operatorname{int} \operatorname{dom} j$, *et qu'il existe ε et*

δ *dans* $]0,+\infty[$ *tels que* $\nabla h+\varepsilon A$ *et* $\nabla j+\delta B$ *sont coercifs. Soit* $\sigma\in[\max\{\varepsilon,\delta\},+\infty[$ *et on itère*

$$
\begin{array}{l}
\textit{pour } n=0,1,\ldots\\
\left\lfloor
\begin{array}{l}
(\gamma_n,\mu_n)\in[\varepsilon,\sigma]\times[\delta,\sigma]\\
a_n=(\nabla h+\gamma_nA)^{-1}\big(\nabla h(x_n)-\gamma_nL^*y_n^*\big)\\
a_n^*=\gamma_n^{-1}\big(\nabla h(x_n)-\nabla h(a_n)\big)-L^*y_n^*\\
b_n=(\nabla j+\mu_nB)^{-1}\big(\nabla j(Lx_n)+\mu_ny_n^*\big)\\
b_n^*=\mu_n^{-1}\big(\nabla j(Lx_n)-\nabla j(b_n)\big)+y_n^*\\
\boldsymbol{H}_n=\big\{(x,y^*)\in\mathcal{X}\times\mathcal{Y}^*\;\big|\;\langle x,a_n^*+L^*b_n^*\rangle+\langle b_n-La_n,y^*\rangle\leqslant\langle a_n,a_n^*\rangle+\langle b_n,b_n^*\rangle\big\}\\
(x_{n+1/2},y_{n+1/2}^*)=P^{\boldsymbol{f}}_{\boldsymbol{H}_n}(x_n,y_n^*)\\
(x_{n+1},y_{n+1}^*)=Q^{\boldsymbol{f}}\big((x_0,y_0^*),(x_n,y_n^*),(x_{n+1/2},y_{n+1/2}^*)\big).
\end{array}
\right.
\end{array}
\tag{2.11}
$$

Alors les propriétés suivantes sont satisfaites.

(i) *Soit* $n\in\mathbb{N}$. *Les propriétés suivantes sont équivalentes :*

(a) $(x_n,y_n^*)=(\overline{x},\overline{y}^*)$.

(b) $(x_n,y_n^*)\in\boldsymbol{Z}$.

(c) $(x_n,y_n^*)\in\boldsymbol{H}_n$.

(d) $x_n=a_n$ *et* $y_n^*=b_n^*$.

(e) $La_n=b_n$ *et* $a_n^*=-L^*b_n^*$.

(f) $\boldsymbol{H}_n=\boldsymbol{\mathcal{X}}$.

(g) $(x_{n+1/2},y_{n+1/2}^*)=(x_n,y_n^*)$.

(h) $(x_{n+1},y_{n+1}^*)=(x_n,y_n^*)$.

(ii) $\sum_{n\in\mathbb{N}}D^f(x_{n+1},x_n)<+\infty$ *et* $\sum_{n\in\mathbb{N}}D^{g^*}(y_{n+1}^*,y_n^*)<+\infty$.

(iii) $\sum_{n\in\mathbb{N}}D^f(x_{n+1/2},x_n)<+\infty$ *et* $\sum_{n\in\mathbb{N}}D^{g^*}(y_{n+1/2}^*,y_n^*)<+\infty$.

(iv) *Supposons que* f, g^*, h, *et* j *vérifient la Condition 2.3, et que* ∇h *et* ∇j *soient uniformément continus sur les sous-ensembles bornés de* $\operatorname{int}\operatorname{dom}h$ *et* $\operatorname{int}\operatorname{dom}j$, *respectivement. Alors* $x_n\to\overline{x}$ *et* $y_n^*\to\overline{y}^*$.

Nous considérons à présent une spécialisation du Problème 2.2 au problème de minimisation multivarié.

Problème 2.6 Soient m et p des nombres entiers strictement positifs, soient $(\mathcal{X}_i)_{1\leqslant i\leqslant m}$ et $(\mathcal{Y}_k)_{1\leqslant k\leqslant p}$ des espaces de Banach réel réflexifs, et soit $\boldsymbol{\mathcal{X}}$ l'espace vectoriel $\big(\times_{i=1}^m\mathcal{X}_i\big)\times\big(\times_{k=1}^p\mathcal{Y}_k^*\big)$ muni de la norme

$$(x,y^*)=\big((x_i)_{1\leqslant i\leqslant m},(y_k^*)_{1\leqslant k\leqslant p}\big)\mapsto\sqrt{\sum_{i=1}^m\|x_i\|^2+\sum_{k=1}^p\|y_k^*\|^2}. \tag{2.12}$$

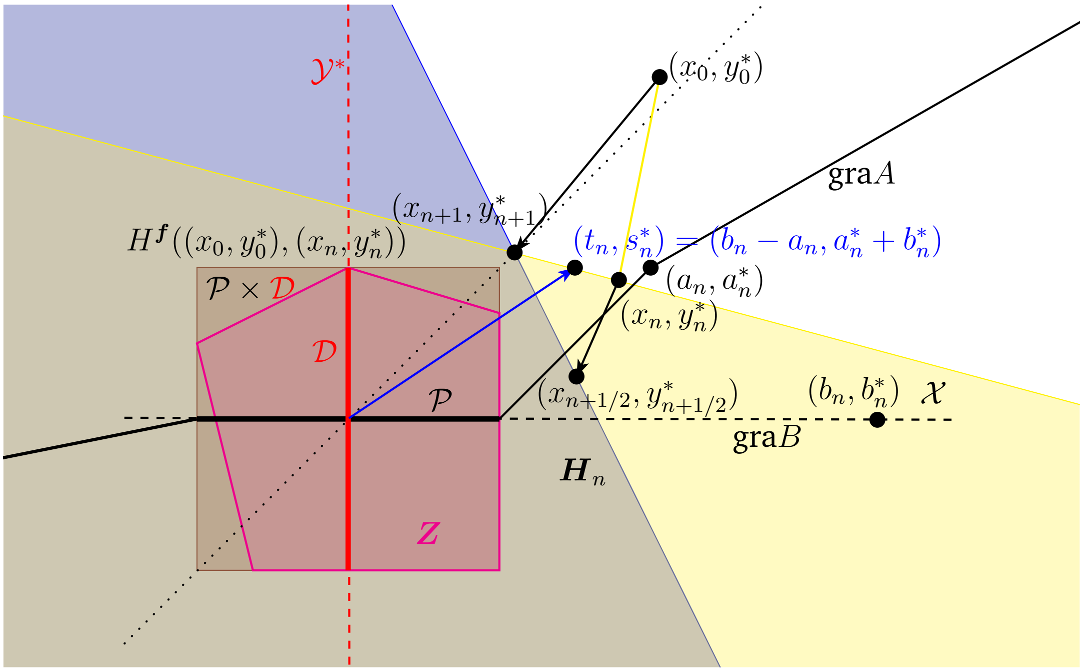

FIGURE 2.4 – Algorithme de meilleure approximation.

Pour tout $i \in \{1, \ldots, m\}$ et pour tout $k \in \{1, \ldots, p\}$, soit $\varphi_i \in \Gamma_0(\mathcal{X}_i)$, soit $\psi_k \in \Gamma_0(\mathcal{Y}_k)$, et soit $L_{ki} \in \mathcal{B}(\mathcal{X}_i, \mathcal{Y}_k)$. Considérons le problème primal

$$\underset{x_1 \in \mathcal{X}_1, \ldots, x_m \in \mathcal{X}_m}{\text{minimiser}} \quad \sum_{i=1}^m \varphi_i(x_i) + \sum_{k=1}^p \psi_k\bigg(\sum_{i=1}^m L_{ki} x_i\bigg), \tag{2.13}$$

le problème dual

$$\underset{y_1^* \in \mathcal{Y}_1^*, \ldots, y_p^* \in \mathcal{Y}_p^*}{\text{minimiser}} \quad \sum_{i=1}^m \varphi_i^*\bigg(-\sum_{k=1}^p L_{ki}^* y_k^*\bigg) + \sum_{k=1}^p \psi_k^*(y_k^*), \tag{2.14}$$

et l'ensemble de Kuhn-Tucker associé

$$\boldsymbol{Z} = \bigg\{(x, y^*) \in \boldsymbol{\mathcal{X}} \;\bigg|\; (\forall i \in \{1, \ldots, m\}) \quad -\sum_{k=1}^p L_{ki}^* y_k^* \in \partial\varphi_i(x_i) \;\text{ et } \quad (\forall k \in \{1, \ldots, p\}) \quad \sum_{i=1}^m L_{ki} x_i \in \partial\psi_k^*(y_k^*)\bigg\}. \tag{2.15}$$

Pour tout $i \in \{1, \ldots, m\}$, soit $f_i \in \Gamma_0(\mathcal{X}_i)$ une fonction de Legendre et soit $x_{i,0} \in \operatorname{int}\operatorname{dom} f_i$. Pour tout $k \in \{1, \ldots, p\}$, soit $g_k \in \Gamma_0(\mathcal{Y}_k)$ une fonction de Legendre et soit

$y^*_{k,0} \in \operatorname{int}\operatorname{dom} g^*_k$. Posons $x_0 = (x_{i,0})_{1\leqslant i\leqslant m}$, $y^*_0 = (y^*_{k,0})_{1\leqslant k\leqslant p}$, et

$$\boldsymbol{f}\colon \boldsymbol{\mathcal{X}} \to \left]-\infty,+\infty\right] : (x,y^*) \mapsto \sum_{i=1}^{m} f_i(x_i) + \sum_{k=1}^{p} g^*_k(y^*_k), \tag{2.16}$$

et supposons que $\boldsymbol{Z} \cap \operatorname{int}\operatorname{dom}\boldsymbol{f} \neq \varnothing$. Le but est de trouver la meilleure approximation au sens de Bregman $\big((\overline{x}_i)_{1\leqslant i\leqslant m}, (\overline{y}^*_k)_{1\leqslant k\leqslant p}\big) = P^{\boldsymbol{f}}_{\boldsymbol{Z}}(x_0,y^*_0)$ de (x_0,y^*_0) dans $\boldsymbol{Z}$.

Proposition 2.7 *Considérons le Problème 2.6. Pour tout $i \in \{1,\ldots,m\}$, soit $h_i \in \Gamma_0(\mathcal{X}_i)$ une fonction de Legendre telle que $\operatorname{int}\operatorname{dom} f_i \subset \operatorname{int}\operatorname{dom} h_i$ et $h_i + \varepsilon_i\varphi_i$ est supercoercive pour un certain $\varepsilon_i \in \left]0,+\infty\right[$. Pour tout $k \in \{1,\ldots,p\}$, soit $j_k \in \Gamma_0(\mathcal{Y}_k)$ une fonction de Legendre telle que $\sum_{i=1}^m L_{ki}(\operatorname{int}\operatorname{dom} f_i) \subset \operatorname{int}\operatorname{dom} j_k$ et $j_k + \delta_k\psi_k$ est supercoercive pour un certain $\delta_k \in \left]0,+\infty\right[$. Posons $\varepsilon = \max_{1\leqslant i\leqslant m}\varepsilon_i$ et $\delta = \max_{1\leqslant k\leqslant p}\delta_k$. Soit $\sigma \in \left[\max\{\varepsilon,\delta\},+\infty\right[$, et on itère*

$$\begin{array}{l}
\textit{pour } n = 0,1,\ldots\\
\left\lfloor\begin{array}{l}
(\gamma_n,\mu_n) \in [\varepsilon,\sigma]\times[\delta,\sigma]\\
\textit{pour } i = 1,\ldots,m\\
\left\lfloor\begin{array}{l}
a_{i,n} = (\nabla h_i + \gamma_n\partial\varphi_i)^{-1}\Big(\nabla h_i(x_{i,n}) - \gamma_n\sum_{k=1}^p L^*_{ki}y^*_{k,n}\Big)\\
a^*_{i,n} = \gamma_n^{-1}\big(\nabla h_i(x_{i,n}) - \nabla h_i(a_{i,n})\big) - \sum_{k=1}^p L^*_{ki}y^*_{k,n}
\end{array}\right.\\
\textit{pour } k = 1,\ldots,p\\
\left\lfloor\begin{array}{l}
b_{k,n} = (\nabla j_k + \mu_n\partial\psi_k)^{-1}\big(\nabla j_k\big(\sum_{i=1}^m L_{ki}x_{i,n}\big) + \mu_n y^*_{k,n}\big)\\
b^*_{k,n} = \mu_n^{-1}\big(\nabla j_k\big(\sum_{i=1}^m L_{ki}x_{i,n}\big) - \nabla j_k(b_{k,n})\big) + y^*_{k,n}\\
t_{k,n} = b_{k,n} - \sum_{i=1}^m L_{ki}a_{i,n}
\end{array}\right.\\
\textit{pour } i = 1,\ldots,m\\
\left\lfloor\begin{array}{l}
s^*_{i,n} = a^*_{i,n} + \sum_{k=1}^p L^*_{ki}b^*_{k,n}
\end{array}\right.\\
\eta_n = \sum_{i=1}^m \big\langle a_{i,n}, a^*_{i,n}\big\rangle + \sum_{k=1}^p \big\langle b_{k,n}, b^*_{k,n}\big\rangle\\
\boldsymbol{H}_n = \left\{(x,y^*)\in\boldsymbol{\mathcal{X}} \;\middle|\; \sum_{i=1}^m \big\langle x_i, s^*_{i,n}\big\rangle + \sum_{k=1}^p \langle t_{k,n}, y^*_k\rangle \leqslant \eta_n\right\}\\
\big(x_{n+1/2}, y^*_{n+1/2}\big) = P^{\boldsymbol{f}}_{\boldsymbol{H}_n}(x_n,y^*_n)\\
\big(x_{n+1}, y^*_{n+1}\big) = Q^f\big((x_0,y^*_0),(x_n,y^*_n),(x_{n+1/2},y^*_{n+1/2})\big),
\end{array}\right.
\end{array} \tag{2.17}$$

où nous utilisons les notions $(\forall n\in\mathbb{N})$ $x_n = (x_{i,n})_{1\leqslant i\leqslant m}$ et $y^*_n = (y^*_{k,n})_{1\leqslant k\leqslant p}$. Supposons que les conditions suivantes soient satisfaites.

(i) *Pour tout $i \in \{1,\ldots,m\}$, f_i et h_i satisfont la Condition 2.3 et ∇h_i est uniformément continu sur les sous-ensembles bornés de $\operatorname{int}\operatorname{dom} h_i$.*

(ii) *Pour tout $k \in \{1,\ldots,p\}$, g^*_k et j_k satisfont la Condition 2.3 et ∇j_k est uniformément continu sur les sous-ensembles bornés de $\operatorname{int}\operatorname{dom} j_k$.*

Alors

$$(\forall i\in\{1,\ldots,m\})\quad x_{i,n}\to\overline{x}_i \quad \textit{et} \quad (\forall k\in\{1,\ldots,p\})\quad y^*_{k,n}\to\overline{y}^*_k. \tag{2.18}$$

Dans les espaces de dimensions finies, la convergence de l'algorithme 2.11 peut-être obtenue sous des hypothèses plus générales.

Proposition 2.8 *Dans le Problème 2.2, supposons que $\mathcal{X}$ et $\mathcal{Y}$ soient de dimensions finies. Soient $h \in \Gamma_0(\mathcal{X})$ et $j \in \Gamma_0(\mathcal{Y})$ des fonctions de Legendre telles que $\operatorname{int}\operatorname{dom} f \subset \operatorname{int}\operatorname{dom} h$, $L(\operatorname{int}\operatorname{dom} f) \subset \operatorname{int}\operatorname{dom} j$, et il existe ε et δ dans $]0,+\infty[$ tels que $\nabla h + \varepsilon A$ et $\nabla j + \delta B$ sont coercifs. Soit $\sigma \in [\max\{\varepsilon,\delta\},+\infty[$ et on exécute l'algorithme 2.11. Alors $(x_n, y_n^*) \to (\overline{x}, \overline{y}^*)$.*

2.2 Article en anglais

SOLVING COMPOSITE MONOTONE INCLUSIONS IN REFLEXIVE BANACH SPACES BY CONSTRUCTING BEST BREGMAN APPROXIMATIONS FROM THEIR KUHN-TUCKER SET [1]

2.2.1 Introduction

Let $\mathcal{X}$ be a reflexive real Banach space with norm $\|\cdot\|$ and let $\langle\cdot,\cdot\rangle$ be the duality pairing between $\mathcal{X}$ and its topological dual $\mathcal{X}^*$. A set-valued operator $M\colon \mathcal{X} \to 2^{\mathcal{X}^*}$ with graph $\operatorname{gra} M = \big\{(x,x^*) \in \mathcal{X}\times\mathcal{X}^* \,\big|\, x^* \in Mx\big\}$ is monotone if

$$(\forall (x_1,x_1^*) \in \operatorname{gra} M)(\forall (x_2,x_2^*) \in \operatorname{gra} M) \quad \langle x_1 - x_2, x_1^* - x_2^*\rangle \geqslant 0, \tag{2.19}$$

and maximally monotone if, furthermore, there exists no monotone operator from $\mathcal{X}$ to $2^{\mathcal{X}^*}$ the graph of which properly contains $\operatorname{gra} M$. Monotone operator theory emerged in the early 1960s as a well-structured branch of nonlinear analysis [23, 28, 29, 40], and its remains very active [9, 10, 34, 41]. One of the main reasons for the success of the theory is that a significant range of problems in areas such as optimization, economics, variational inequalities, partial differential equations, mechanics, signal and image processing, optimal transportation, machine learning, and traffic theory can be reduced to solving inclusions of the type

$$\text{find } x \in \mathcal{X} \quad \text{such that} \quad 0 \in Mx, \tag{2.20}$$

where $M\colon \mathcal{X} \to 2^{\mathcal{X}^*}$ is maximally monotone. Conceptually, this inclusion can be solved via the Bregman proximal point algorithm, special instances of which go back to [21, 24, 35]. To present its general form [7], we need the following definitions, which revolve around the notion of a Bregman distance pioneered in [13].

1. P. L. Combettes and V. Q. Nguyen, Solving composite monotone inclusions in reflexive Banach spaces by constructing best Bregman approximations from their Kuhn-Tucker set, *submitted*.

Definition 2.9 [6, 7] Let $\mathcal{X}$ be a reflexive real Banach space and let $f\colon \mathcal{X} \to \left]-\infty,+\infty\right]$ be a proper lower semicontinuous convex function, with conjugate $f^*\colon \mathcal{X}^* \to \left]-\infty,+\infty\right] : x^* \mapsto \sup_{x\in\mathcal{X}}(\langle x, x^*\rangle - f(x))$ and Moreau subdifferential [31]

$$\partial f\colon \mathcal{X} \to 2^{\mathcal{X}^*}\colon x \mapsto \big\{x^* \in \mathcal{X}^* \;\big|\; (\forall y \in \mathcal{X})\ \langle y - x, x^*\rangle + f(x) \leqslant f(y)\big\}. \tag{2.21}$$

Then f is a *Legendre function* if it is *essentially smooth* in the sense that ∂f is both locally bounded and single-valued on its domain, and *essentially strictly convex* in the sense that ∂f^* is locally bounded on its domain and f is strictly convex on every convex subset of $\operatorname{dom}\partial f$. Moreover, f is Gâteaux differentiable on $\operatorname{int}\operatorname{dom} f \neq \varnothing$ and the associated *Bregman distance* is

$$\begin{aligned} D^f\colon \mathcal{X}\times\mathcal{X} &\to [0,+\infty] \\ (x,y) &\mapsto \begin{cases} f(x) - f(y) - \langle x - y, \nabla f(y)\rangle\,, & \text{if } y \in \operatorname{int}\operatorname{dom} f;\\ +\infty, & \text{otherwise}. \end{cases} \end{aligned} \tag{2.22}$$

Let C be a closed convex subset of $\mathcal{X}$ such that $C\cap\operatorname{int}\operatorname{dom} f \neq \varnothing$. The *Bregman projector* onto C induced by f is

$$\begin{aligned} P_C^f\colon \operatorname{int}\operatorname{dom} f &\to C\cap\operatorname{int}\operatorname{dom} f \\ y &\mapsto \operatorname*{argmin}_{x\in C} D^f(x,y). \end{aligned} \tag{2.23}$$

It follows from [7, Theorem 5.18] that, under suitable assumptions on f and M, given a sequence $(\gamma_n)_{n\in\mathbb{N}}$ in $\left]0,+\infty\right[$ such that $\inf_{n\in\mathbb{N}}\gamma_n > 0$, the sequence defined by

$$x_0 \in \operatorname{int}\operatorname{dom} f \quad\text{and}\quad (\forall n\in\mathbb{N})\quad x_{n+1} = (\nabla f + \gamma_n M)^{-1}\circ\nabla f(x_n) \tag{2.24}$$

converges weakly to a solution to (2.20) (in the case when $\mathcal{X}$ is a Hilbert space and $f = \|\cdot\|^2/2$, $(\nabla f + \gamma_n M)^{-1}\circ\nabla f$ reduces to the standard resolvent $J_{\gamma_n M}$ and we obtain the classical result of [33, Theorem 1]). A strongly convergent variant of (2.24) was proposed in [8]. In applications, however, M is typically too complex for (2.24) to be implementable. For instance, given a real Banach space $\mathcal{Y}$, a typical composite model is $M = A + L^*BL$, where $A\colon \mathcal{X}\to 2^{\mathcal{X}^*}$ and $B\colon \mathcal{Y}\to 2^{\mathcal{Y}^*}$ are monotone, and $L\colon \mathcal{X}\to\mathcal{Y}$ is linear and bounded. In Hilbert spaces, if $\mathcal{X} = \mathcal{Y}$ and $L = \operatorname{Id}$, several well-established splitting methods are available to solve (2.20), i.e., to find a zero of $A + B$ using A and B separately at each iteration [9, 26, 27, 36]. Splitting methods for the more versatile composite model $M = A + L^*BL$ in Hilbert spaces were first proposed in [14] (see [1, 11, 12, 18, 19, 37] for subsequent developments). These methods provide in general only weak convergence to an unspecified solution and, in addition, they require knowledge of $\|L\|$ or potentially costly inversions of linear operators. The recent method primal-dual method of [3] circumvents these limitations and, in addition, converges to the best approximation to a reference point from the Kuhn-Tucker set relative to the underlying hilbertian distance. The objective of this paper is to extend it to reflexive

Banach spaces and to best approximation relative to general Bregman distances. Let us stress that the theory of splitting algorithms in Banach spaces is rather scarce as most hilbertian splitting methods cannot be naturally extended to that setting; in particular, to the best of our knowledge there exists at present no splitting algorithm for finding a zero of $M = A + L^*BL$ outside of Hilbert spaces. By contrast, the geometric primal-dual construction of [3], which consists in projecting a reference point onto successive simple outer approximations to the Kuhn-Tucker set of the inclusion $0 \in Ax + L^*BLx$, lends itself to such an extension. Our analysis will borrow tools on Legendre functions and Bregman-based algorithms from [6] and [7], as well as geometric constructs from [3] and [8]. The proposed results will provide not only the first splitting methods for composite inclusions outside of Hilbert spaces, but also new algorithms in Hilbert, and even Euclidean, spaces.

The problem under consideration is the following.

Problem 2.10 Let $\mathcal{X}$ and $\mathcal{Y}$ be reflexive real Banach spaces such that $\mathcal{X} \neq \{0\}$ and $\mathcal{Y} \neq \{0\}$, let $\boldsymbol{\mathcal{X}}$ be the standard product vector space $\mathcal{X} \times \mathcal{Y}^*$ equipped with the norm $(x, y^*) \mapsto \sqrt{\|x\|^2 + \|y^*\|^2}$, and let $\boldsymbol{\mathcal{X}}^*$ be its topological dual, that is, $\mathcal{X}^* \times \mathcal{Y}$ equipped with the norm $(x^*, y) \mapsto \sqrt{\|x^*\|^2 + \|y\|^2}$. Let $A\colon \mathcal{X} \to 2^{\mathcal{X}^*}$ and $B\colon \mathcal{Y} \to 2^{\mathcal{Y}^*}$ be maximally monotone, and let $L\colon \mathcal{X} \to \mathcal{Y}$ be linear and bounded. Consider the inclusion problem

$$\text{find} \quad x \in \mathcal{X} \quad \text{such that} \quad 0 \in Ax + L^*BLx, \tag{2.25}$$

the dual problem

$$\text{find} \quad y^* \in \mathcal{Y}^* \quad \text{such that} \quad 0 \in -LA^{-1}(-L^*y^*) + B^{-1}y^*, \tag{2.26}$$

and let

$$\boldsymbol{Z} = \big\{(x, y^*) \in \boldsymbol{\mathcal{X}} \;\big|\; -L^*y^* \in Ax \quad \text{and} \quad Lx \in B^{-1}y^*\big\} \tag{2.27}$$

be the associated Kuhn-Tucker set. Let $f\colon \mathcal{X} \to \left]-\infty, +\infty\right]$ and $g\colon \mathcal{Y} \to \left]-\infty, +\infty\right]$ be Legendre functions, set

$$\boldsymbol{f}\colon \boldsymbol{\mathcal{X}} \to \left]-\infty, +\infty\right] : (x, y^*) \mapsto f(x) + g^*(y^*), \tag{2.28}$$

let $x_0 \in \operatorname{int}\operatorname{dom} f$, let $y_0^* \in \operatorname{int}\operatorname{dom} g^*$, and suppose that $\boldsymbol{Z} \cap \operatorname{int}\operatorname{dom} \boldsymbol{f} \neq \varnothing$. The problem is to find the best Bregman approximation $(\overline{x}, \overline{y}^*) = P_{\boldsymbol{Z}}^{\boldsymbol{f}}(x_0, y_0^*)$ to (x_0, y_0^*) from $\boldsymbol{Z}$.

Notation. The symbols $\rightharpoonup$ and $\to$ denote respectively weak and strong convergence. The set of weak sequential cluster points of a sequence $(x_n)_{n\in\mathbb{N}}$ is denoted by $\mathfrak{W}(x_n)_{n\in\mathbb{N}}$. The closed ball of center $x \in \mathcal{X}$ and radius $\rho \in \left]0, +\infty\right[$ is denoted by $B(x; \rho)$. Let $M\colon \mathcal{X} \to 2^{\mathcal{X}^*}$ be a set-valued operator. The domain of M is $\operatorname{dom} M = \big\{x \in \mathcal{X} \;\big|\; Mx \neq \varnothing\big\}$, the range of M is $\operatorname{ran} M = \big\{x^* \in \mathcal{X}^* \;\big|\; (\exists\, x \in \mathcal{X})\, x^* \in Mx\big\}$, and the set of zeros of M is $\operatorname{zer} M = \big\{x \in \mathcal{X} \;\big|\; 0 \in Mx\big\}$. $\Gamma_0(\mathcal{X})$ is the class of all lower semicontinuous convex functions $f\colon \mathcal{X} \to \left]-\infty, +\infty\right]$ such that $\operatorname{dom} f = \big\{x \in \mathcal{X} \;\big|\; f(x) < +\infty\big\} \neq \varnothing$. Let $f\colon \mathcal{X} \to \left]-\infty, +\infty\right]$. Then f is coercive if $\lim_{\|x\|\to+\infty} f(x) = +\infty$ and supercoercive if $\lim_{\|x\|\to+\infty} f(x)/\|x\| = +\infty$.

2.2.2 Preliminary results

2.2.2.1 Properties of the Kuhn-Tucker set

The following proposition revisits and complements some results of [2] and [14] on the properties of the Kuhn-Tucker set in the more general setting of Problem 2.10.

Proposition 2.11 *Consider the setting of Problem 2.10. Then the following hold :*

(i) *Let $\mathscr{P}$ be the set of solutions to (2.25) and let $\mathscr{D}$ be the set of solutions to (2.26). Then the following hold :*

(a) *$\boldsymbol{Z}$ is a closed convex subset of $\mathscr{P} \times \mathscr{D}$.*

(b) *Set $\mathcal{Q}_{\mathcal{X}} \colon \boldsymbol{\mathcal{X}} \to \mathcal{X} \colon (x, y^*) \mapsto x$ and $\mathcal{Q}_{\mathcal{Y}^*} \colon \boldsymbol{\mathcal{X}} \to \mathcal{Y}^* \colon (x, y^*) \mapsto y^*$. Then $\mathscr{P} = \mathcal{Q}_{\mathcal{X}}(\boldsymbol{Z})$ and $\mathscr{D} = \mathcal{Q}_{\mathcal{Y}^*}(\boldsymbol{Z})$.*

(c) *$\mathscr{P} \neq \varnothing \Leftrightarrow \boldsymbol{Z} \neq \varnothing \Leftrightarrow \mathscr{D} \neq \varnothing$.*

(ii) *For every $\mathsf{a} = (a, a^*) \in \operatorname{gra} A$ and $\mathsf{b} = (b, b^*) \in \operatorname{gra} B$, set $\boldsymbol{s}^*_{\mathsf{a},\mathsf{b}} = (a^* + L^*b^*, b - La)$, $\eta_{\mathsf{a},\mathsf{b}} = \langle a, a^*\rangle + \langle b, b^*\rangle$, and $\boldsymbol{H}_{\mathsf{a},\mathsf{b}} = \big\{\boldsymbol{x} \in \boldsymbol{\mathcal{X}} \mid \langle \boldsymbol{x}, \boldsymbol{s}^*_{\mathsf{a},\mathsf{b}}\rangle \leqslant \eta_{\mathsf{a},\mathsf{b}}\big\}$. Then the following hold :*

(a) *$(\forall \mathsf{a} \in \operatorname{gra} A)(\forall \mathsf{b} \in \operatorname{gra} B)\ [\,\boldsymbol{s}^*_{\mathsf{a},\mathsf{b}} = \boldsymbol{0} \ \Leftrightarrow\ \boldsymbol{H}_{\mathsf{a},\mathsf{b}} = \boldsymbol{\mathcal{X}} \ \Rightarrow\ (a, b^*) \in \boldsymbol{Z}$ and $\eta_{\mathsf{a},\mathsf{b}} = 0\,]$.*

(b) *$\boldsymbol{Z} = \bigcap_{\mathsf{a} \in \operatorname{gra} A} \bigcap_{\mathsf{b} \in \operatorname{gra} B} \boldsymbol{H}_{\mathsf{a},\mathsf{b}}$.*

(iii) *Let $(a_n, a_n^*)_{n\in\mathbb{N}}$ be a sequence in $\operatorname{gra} A$, let $(b_n, b_n^*)_{n\in\mathbb{N}}$ be a sequence in $\operatorname{gra} B$, and let $(x, y^*) \in \boldsymbol{\mathcal{X}}$. Suppose that $a_n \rightharpoonup x$, $b_n^* \rightharpoonup y^*$, $a_n^* + L^*b_n^* \to 0$, and $La_n - b_n \to 0$. Then $(x, y^*) \in \boldsymbol{Z}$.*

Proof. Set $\boldsymbol{M} \colon \boldsymbol{\mathcal{X}} \to 2^{\boldsymbol{\mathcal{X}}^*} \colon (x, y^*) \mapsto Ax \times B^{-1}y^*$ and $\boldsymbol{S} \colon \boldsymbol{\mathcal{X}} \to \boldsymbol{\mathcal{X}}^* \colon (x, y^*) \mapsto (L^*y^*, -Lx)$. Since A and B^{-1} are maximally monotone, so is $\boldsymbol{M}$. On the other hand, $\boldsymbol{S}$ is linear, bounded, and positive since

$$(\forall (x, y^*) \in \boldsymbol{\mathcal{X}}) \quad \langle \boldsymbol{S}(x, y^*), (x, y^*)\rangle = \langle x, L^*y^*\rangle + \langle -Lx, y^*\rangle = 0. \tag{2.29}$$

Thus, it follows from [34, Section 17] that $\boldsymbol{S}$ is maximally monotone with $\operatorname{dom} \boldsymbol{S} = \boldsymbol{\mathcal{X}}$. In turn, we derive from [34, Theorem 24.1(a)] that

$$\boldsymbol{M} + \boldsymbol{S} \text{ is maximally monotone.} \tag{2.30}$$

(i)(a) : Let $(x, y^*) \in \boldsymbol{\mathcal{X}}$. Then

$$\begin{aligned} \boldsymbol{0} \in \boldsymbol{M}(x, y^*) + \boldsymbol{S}(x, y^*) \ &\Leftrightarrow\ 0 \in Ax + L^*y^* \quad \text{and} \quad 0 \in B^{-1}y^* - Lx \\ &\Leftrightarrow\ -L^*y^* \in Ax \quad \text{and} \quad Lx \in B^{-1}y^* \\ &\Leftrightarrow\ (x, y^*) \in \boldsymbol{Z}. \end{aligned} \tag{2.31}$$

Therefore, we derive from (2.30) and [15, Lemma 1.1(a)] that $\boldsymbol{Z} = \operatorname{zer}(\boldsymbol{M}+\boldsymbol{S}) = (\boldsymbol{M}+\boldsymbol{S})^{-1}(0)$ is closed and convex.

(i)(b) : Let $x \in \mathcal{X}$. Then $x \in \mathscr{P} \Leftrightarrow 0 \in Ax + L^*BLx \Leftrightarrow (\exists\, y^* \in \mathcal{Y}^*)\ [-L^*y^* \in Ax$ and $y^* \in BLx] \Leftrightarrow (\exists\, y^* \in \mathcal{Y}^*)\ (x,y^*) \in \boldsymbol{Z}$. Hence $\mathscr{P} \neq \varnothing \Leftrightarrow \boldsymbol{Z} \neq \varnothing$. Likewise, let $y^* \in \mathcal{Y}^*$. Then $y^* \in \mathscr{D} \Leftrightarrow 0 \in -LA^{-1}(-L^*y^*) + B^{-1}y^* \Leftrightarrow (\exists\, x \in \mathcal{X})\ [x \in A^{-1}(-L^*y^*)$ and $0 \in -Lx + B^{-1}y^*] \Leftrightarrow (\exists\, x \in \mathcal{X})\ [-L^*y^* \in Ax$ and $Lx \in B^{-1}y^*] \Leftrightarrow (\exists\, x \in \mathcal{X})\ (x,y^*) \in \boldsymbol{Z}$.

(i)(c) : Clear from (i)(b) (see also [32, Corollary 2.4]).

(ii)(a) : Let $\mathsf{a} \in \operatorname{gra} A$ and $\mathsf{b} \in \operatorname{gra} B$. Then $\boldsymbol{s}^*_{\mathsf{a},\mathsf{b}} = \boldsymbol{0} \Rightarrow [-L^*b^* = a^* \in Aa$ and $La = b \in B^{-1}b^*] \Rightarrow (a,b^*) \in \boldsymbol{Z}$. In addition,

$$\boldsymbol{s}^*_{\mathsf{a},\mathsf{b}} = \boldsymbol{0} \quad \Rightarrow \quad \eta_{\mathsf{a},\mathsf{b}} = \langle a, a^*\rangle + \langle b, b^*\rangle = \langle a, -L^*b^*\rangle + \langle La, b^*\rangle = -\langle La, b^*\rangle + \langle La, b^*\rangle = 0. \tag{2.32}$$

Thus $\boldsymbol{s}^*_{\mathsf{a},\mathsf{b}} = \boldsymbol{0} \Rightarrow \boldsymbol{H}_{\mathsf{a},\mathsf{b}} = \boldsymbol{\mathcal{X}}$. Conversely, $\boldsymbol{H}_{\mathsf{a},\mathsf{b}} = \boldsymbol{\mathcal{X}} \Rightarrow \boldsymbol{s}^*_{\mathsf{a},\mathsf{b}} = \boldsymbol{0} \Rightarrow \eta_{\mathsf{a},\mathsf{b}} = 0$.

(ii)(b) : First, suppose that $\boldsymbol{x} = (x,y^*) \in \bigcap_{\mathsf{a}\in\operatorname{gra} A}\bigcap_{\mathsf{b}\in\operatorname{gra} B} \boldsymbol{H}_{\mathsf{a},\mathsf{b}}$. Then

$$\begin{aligned}(\forall \mathsf{a} \in \operatorname{gra} A)(\forall \mathsf{b} \in \operatorname{gra} B) \quad & \langle (a,b^*) - (x,y^*), (a^*,b) - (-L^*y^*, Lx)\rangle \\ &= \langle (a-x, b^*-y^*), (a^*+L^*y^*, b-Lx)\rangle \\ &= \langle a-x, a^*+L^*y^*\rangle + \langle b-Lx, b^*-y^*\rangle \\ &= \eta_{\mathsf{a},\mathsf{b}} - \left\langle \boldsymbol{x}, \boldsymbol{s}^*_{\mathsf{a},\mathsf{b}}\right\rangle \\ &\geqslant 0. \end{aligned} \tag{2.33}$$

On the other hand, since

$$\big\{\big((a,b^*),(a^*,b)\big) \;\big|\; \mathsf{a} \in \operatorname{gra} A, \mathsf{b} \in \operatorname{gra} B\big\} = \operatorname{gra} \boldsymbol{M}, \tag{2.34}$$

it follows from (2.30) and (2.33) that $((x,y^*),(-L^*y^*,Lx)) \in \operatorname{gra}\boldsymbol{M}$, i.e., $\boldsymbol{x} \in \boldsymbol{Z}$. Thus

$$\bigcap_{\mathsf{a}\in\operatorname{gra} A}\ \bigcap_{\mathsf{b}\in\operatorname{gra} B} \boldsymbol{H}_{\mathsf{a},\mathsf{b}} \subset \boldsymbol{Z}. \tag{2.35}$$

Conversely, let $\mathsf{a} \in \operatorname{gra} A$, let $\mathsf{b} \in \operatorname{gra} B$, and let $(x,y^*) \in \boldsymbol{Z}$. Then $(x,-L^*y^*) \in \operatorname{gra} A$ and $(Lx,y^*) \in \operatorname{gra} B$. Since A and B are monotone, we obtain

$$\langle a-x, a^*+L^*y^*\rangle \geqslant 0 \quad \text{and} \quad \langle b-Lx, b^*-y^*\rangle \geqslant 0. \tag{2.36}$$

Adding these two inequalities yields

$$\langle x-a, a^*+L^*y^*\rangle + \langle Lx-b, b^*-y^*\rangle \leqslant 0 \tag{2.37}$$

and, therefore,

$$
\begin{aligned}
\langle \boldsymbol{x}, \boldsymbol{s}^*_{\mathsf{a},\mathsf{b}} \rangle &= \langle x, a^* + L^*b^* \rangle + \langle b - La, y^* \rangle \\
&= \langle x, a^* + L^*y^* \rangle + \langle Lx, b^* - y^* \rangle + \langle b - Lx, y^* \rangle + \langle x - a, L^*y^* \rangle \\
&= \langle x - a, a^* + L^*y^* \rangle + \langle a, a^* \rangle + \langle La, y^* \rangle \\
&\quad + \langle Lx - b, b^* - y^* \rangle + \langle b, b^* \rangle - \langle b, y^* \rangle + \langle b - Lx, y^* \rangle + \langle x - a, L^*y^* \rangle \\
&\leqslant \langle a, a^* \rangle + \langle b, b^* \rangle + \langle La - b, y^* \rangle + \langle b - Lx, y^* \rangle + \langle x - a, L^*y^* \rangle \\
&= \langle a, a^* \rangle + \langle b, b^* \rangle \\
&= \eta_{\mathsf{a},\mathsf{b}}.
\end{aligned} \tag{2.38}
$$

This implies that $(x, y^*) \in \boldsymbol{H}_{\mathsf{a},\mathsf{b}}$. Hence, $\boldsymbol{Z} \subset \boldsymbol{H}_{\mathsf{a},\mathsf{b}}$.

(iii) : Set $(\forall n \in \mathbb{N})$ $\boldsymbol{x}_n = (a_n, b^*_n)$ and $\boldsymbol{x}^*_n = (a^*_n + L^*b^*_n, b_n - La_n)$. Then $\boldsymbol{x}_n \rightharpoonup (x, y^*)$, $\boldsymbol{x}^*_n \to \mathbf{0}$, and $(\forall n \in \mathbb{N})$ $(\boldsymbol{x}_n, \boldsymbol{x}^*_n) \in \operatorname{gra}(\boldsymbol{M} + \boldsymbol{S})$. However, it follows from (2.30) that $\operatorname{gra}(\boldsymbol{M} + \boldsymbol{S})$ is sequentially closed in $\boldsymbol{\mathcal{X}}^{\text{weak}} \times \boldsymbol{\mathcal{X}}^{*\text{strong}}$ [15, Lemma 1.2]. Hence, $\mathbf{0} \in (\boldsymbol{M} + \boldsymbol{S})(x, y^*)$, i.e., by (2.31), $(x, y^*) \in \boldsymbol{Z}$. □

2.2.2.2 Best Bregman approximation algorithm

The approach we present goes back to Haugazeau's algorithm [22, Théorème 3-2] (see also [9, Theorem 29.3]) for projecting a point onto the intersection of closed convex sets in a Hilbert space using the projections onto the individual sets. The method was extended in [17] to minimize certain convex functions over the intersection of closed convex sets in Banach spaces. The adaptation to the problem of finding the best Bregman approximation from a closed convex set was investigated in [8].

Definition 2.12 [7, Definition 3.1] and [8, Section 3] Let $\boldsymbol{\mathcal{X}}$ be a reflexive real Banach space, let $\boldsymbol{f} \in \Gamma_0(\boldsymbol{\mathcal{X}})$ be a Legendre function, let $\boldsymbol{x}_0 \in \operatorname{int}\operatorname{dom}\boldsymbol{f}$, let $\boldsymbol{x} \in \operatorname{int}\operatorname{dom}\boldsymbol{f}$, and let $\boldsymbol{y} \in \operatorname{int}\operatorname{dom}\boldsymbol{f}$. Then

$$
\begin{aligned}
H^{\boldsymbol{f}}(\boldsymbol{x}, \boldsymbol{y}) &= \big\{ \boldsymbol{z} \in \boldsymbol{\mathcal{X}} \;\big|\; \langle \boldsymbol{z} - \boldsymbol{y}, \nabla \boldsymbol{f}(\boldsymbol{x}) - \nabla \boldsymbol{f}(\boldsymbol{y}) \rangle \leqslant 0 \big\} \\
&= \big\{ \boldsymbol{z} \in \boldsymbol{\mathcal{X}} \;\big|\; D^{\boldsymbol{f}}(\boldsymbol{z}, \boldsymbol{y}) + D^{\boldsymbol{f}}(\boldsymbol{y}, \boldsymbol{x}) \leqslant D^{\boldsymbol{f}}(\boldsymbol{z}, \boldsymbol{x}) \big\}
\end{aligned} \tag{2.39}
$$

is the closed affine half-space onto which $\boldsymbol{y}$ is the Bregman projection of $\boldsymbol{x}$ if $\boldsymbol{x} \neq \boldsymbol{y}$. Moreover, if $H^{\boldsymbol{f}}(\boldsymbol{x}_0, \boldsymbol{x}) \cap H^{\boldsymbol{f}}(\boldsymbol{x}, \boldsymbol{y}) \cap \operatorname{int}\operatorname{dom}\boldsymbol{f} \neq \varnothing$, then

$$
Q^{\boldsymbol{f}}(\boldsymbol{x}_0, \boldsymbol{x}, \boldsymbol{y}) = P^{\boldsymbol{f}}_{H^{\boldsymbol{f}}(\boldsymbol{x}_0, \boldsymbol{x}) \cap H^{\boldsymbol{f}}(\boldsymbol{x}, \boldsymbol{y})} \boldsymbol{x}_0. \tag{2.40}
$$

Lemma 2.13 [7, Lemma 3.2] *Let $\boldsymbol{\mathcal{X}}$ be a reflexive real Banach space, and let $\boldsymbol{C}_1$ and $\boldsymbol{C}_2$ be convex subsets of $\boldsymbol{\mathcal{X}}$ such that $\boldsymbol{C}_1$ is closed and $\boldsymbol{C}_1 \cap \operatorname{int}\boldsymbol{C}_2 \neq \varnothing$. Then $\overline{\boldsymbol{C}_1 \cap \operatorname{int}\boldsymbol{C}_2} = \boldsymbol{C}_1 \cap \overline{\boldsymbol{C}}_2$.*

Proposition 2.14 *Let $\boldsymbol{\mathcal{X}}$ be a reflexive real Banach space, let $\boldsymbol{f} \in \Gamma_0(\boldsymbol{\mathcal{X}})$ be a Legendre function, let $\boldsymbol{C}$ be a closed convex subset of $\overline{\operatorname{dom}}\,\boldsymbol{f}$ such that $\boldsymbol{C} \cap \operatorname{int}\operatorname{dom}\boldsymbol{f} \neq \varnothing$, let $\boldsymbol{x}_0 \in \operatorname{int}\operatorname{dom}\boldsymbol{f}$, and set $\overline{\boldsymbol{x}} = P^{\boldsymbol{f}}_{\boldsymbol{C}}\boldsymbol{x}_0$. At every iteration $n \in \mathbb{N}$, find $\boldsymbol{x}_{n+1/2} \in \operatorname{int}\operatorname{dom}\boldsymbol{f}$ such that $\boldsymbol{C} \subset H^{\boldsymbol{f}}(\boldsymbol{x}_n, \boldsymbol{x}_{n+1/2})$ and set*

$$\boldsymbol{x}_{n+1} = Q^{\boldsymbol{f}}\big(\boldsymbol{x}_0, \boldsymbol{x}_n, \boldsymbol{x}_{n+1/2}\big). \tag{2.41}$$

Then the following hold :

(i) $(\forall n \in \mathbb{N})\ \boldsymbol{C} \subset H^{\boldsymbol{f}}(\boldsymbol{x}_0, \boldsymbol{x}_n)$.

(ii) $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ *is a well-defined bounded sequence in* $\operatorname{int}\operatorname{dom}\boldsymbol{f}$.

(iii) *Suppose that, for some $n \in \mathbb{N}$, $\boldsymbol{x}_n \in \boldsymbol{C}$. Then $(\forall k \in \mathbb{N})\ \boldsymbol{x}_{n+k} = \overline{\boldsymbol{x}}$.*

(iv) $\sum_{n\in\mathbb{N}} D^{\boldsymbol{f}}(\boldsymbol{x}_{n+1}, \boldsymbol{x}_n) < +\infty$.

(v) $(\forall n \in \mathbb{N})\ D^{\boldsymbol{f}}(\boldsymbol{x}_{n+1/2}, \boldsymbol{x}_n) \leqslant D^{\boldsymbol{f}}(\boldsymbol{x}_{n+1}, \boldsymbol{x}_n)$.

(vi) $\sum_{n\in\mathbb{N}} D^{\boldsymbol{f}}(\boldsymbol{x}_{n+1/2}, \boldsymbol{x}_n) < +\infty$.

(vii) $[\,\boldsymbol{x}_n \rightharpoonup \overline{\boldsymbol{x}}$ *and* $\boldsymbol{f}(\boldsymbol{x}_n) \to \boldsymbol{f}(\overline{\boldsymbol{x}})\,] \Leftrightarrow D^{\boldsymbol{f}}(\boldsymbol{x}_n, \overline{\boldsymbol{x}}) \to 0 \Leftrightarrow \mathfrak{W}(\boldsymbol{x}_n)_{n\in\mathbb{N}} \subset \boldsymbol{C}$.

Proof. Item (i) is found in [8, Proof of Proposition 3.3]. The first equivalence in (vii) follows from [8, Propositions 2.2(ii)]. To establish the remaining assertions, set

$$(\forall n \in \mathbb{N}) \quad T_n = P^{\boldsymbol{f}}_{H^{\boldsymbol{f}}(\boldsymbol{x}_n, \boldsymbol{x}_{n+1/2})}. \tag{2.42}$$

Then, for every $n \in \mathbb{N}$, Definition 2.12 yields $T_n\boldsymbol{x}_n = \boldsymbol{x}_{n+1/2}$, [7, Proposition 3.32(ii)(b)] yields $\operatorname{Fix} T_n = H^{\boldsymbol{f}}(\boldsymbol{x}_n, \boldsymbol{x}_{n+1/2}) \cap \operatorname{int}\operatorname{dom}\boldsymbol{f}$, and we derive from Lemma 2.13 that

$$\boldsymbol{C} \subset H^{\boldsymbol{f}}(\boldsymbol{x}_n, \boldsymbol{x}_{n+1/2}) \cap \overline{\operatorname{dom}}\,\boldsymbol{f} = \overline{H^{\boldsymbol{f}}(\boldsymbol{x}_n, \boldsymbol{x}_{n+1/2}) \cap \operatorname{int}\operatorname{dom}\boldsymbol{f}} = \overline{\operatorname{Fix}}\,T_n. \tag{2.43}$$

On the other hand,

$$(\forall n \in \mathbb{N}) \quad \varnothing \neq \boldsymbol{C} \cap \operatorname{int}\operatorname{dom}\boldsymbol{f} \subset H^{\boldsymbol{f}}(\boldsymbol{x}_n, \boldsymbol{x}_{n+1/2}) \cap \operatorname{int}\operatorname{dom}\boldsymbol{f} = \operatorname{Fix} T_n \tag{2.44}$$

and, therefore, $\bigcap_{n\in\mathbb{N}} \operatorname{Fix} T_n \neq \varnothing$. Altogether, [8, Condition 3.2] is satisfied and it follows from [8, Propositions 3.3 and 3.4] and [8, Proof of Proposition 3.4(vii)] that the proof is complete. □

Remark 2.15 Proposition 2.14 addresses the problem of finding the Bregman projection of $\boldsymbol{x}_0$ onto $\boldsymbol{C}$ by replacing it with a sequence of much simpler problems, namely those of finding at each iteration the Bregman projection of $\boldsymbol{x}_0$ onto the intersection of two half-spaces. The latter amounts to minimizing the convex function $\boldsymbol{f} - \nabla\boldsymbol{f}(\boldsymbol{x}_0)$ subject to two linear inequality constraints.

2.2.2.3 Coercivity and boundedness of monotone operators

Definition 2.16 Let $\mathcal{X}$ be a reflexive real Banach space such that $\mathcal{X} \neq \{0\}$ and let $M\colon \mathcal{X} \to 2^{\mathcal{X}^*}$. Then M is *coercive* if

$$(\exists\, z \in \operatorname{dom} M) \qquad \lim_{\|x\|\to+\infty} \inf \frac{\langle x-z, Mx\rangle}{\|x\|} = +\infty, \tag{2.45}$$

and it is *bounded* if it maps bounded sets to bounded set.

Lemma 2.17 *Let $\mathcal{X}$ be a reflexive real Banach space such that $\mathcal{X} \neq \{0\}$ and let $M\colon \mathcal{X} \to 2^{\mathcal{X}^*}$. Suppose that one of the following holds :*

(i) $\operatorname{dom} M$ *is nonempty and bounded.*

(ii) M *is uniformly monotone at some point $z \in \operatorname{dom} M$ with a supercoercive modulus : there exists a strictly increasing function $\phi\colon\ [0,+\infty[\, \to [0,+\infty]$ that vanishes only at 0 such that $\lim_{t\to+\infty} \phi(t)/t = +\infty$ and*

$$(\forall (x,x^*) \in \operatorname{gra} M)(\forall z^* \in Mz) \quad \langle x-z, x^*-z^*\rangle \geqslant \phi(\|x-z\|). \tag{2.46}$$

(iii) $M = \partial\varphi$*, where φ is a supercoercive function in $\Gamma_0(\mathcal{X})$.*

Then M is coercive.

Proof. (i) : Let $x \in \mathcal{X}$ and let $z \in \operatorname{dom} M$. Then, if $\|x\|$ is sufficiently large, we have $Mx = \varnothing$ and therefore $\inf \langle x-z, Mx\rangle/\|x\| = +\infty$.

(ii) : We have

$$(\forall (x,x^*) \in \operatorname{gra} M)(\forall z^* \in Mz) \quad \langle x-z, x^*-z^*\rangle \geqslant \phi(\|x-z\|). \tag{2.47}$$

Hence, for every $x \in \operatorname{dom} M$ such that $\|x\| > \|z\|$, we have

$$\begin{aligned}(\forall x^* \in Mx)(\forall z^* \in Mz) \quad \frac{\langle x-z, x^*\rangle}{\|x\|} &\geqslant \frac{\phi(\|x-z\|) - \|x-z\|\,\|z^*\|}{\|x\|} \\ &= \frac{\|x-z\|}{\|x\|}\left(\frac{\phi(\|x-z\|)}{\|x-z\|} - \|z^*\|\right).\end{aligned} \tag{2.48}$$

Thus,

$$\lim_{\|x\|\to+\infty} \inf \frac{\langle x-z, Mx\rangle}{\|x\|} = +\infty. \tag{2.49}$$

(iii) : In view of (i), we suppose that $\operatorname{dom} M$ is unbounded. Let $z \in \operatorname{dom} M$. Then we derive from (2.21) that, for every $x \in \operatorname{dom} M \smallsetminus \{0\}$,

$$(\forall x^* \in Mx) \quad \frac{\varphi(x)-\varphi(z)}{\|x\|} \leqslant \frac{\langle x-z, x^*\rangle}{\|x\|}. \tag{2.50}$$

Hence, the supercoercivity of φ yields

$$\lim_{\|x\|\to+\infty} \inf \frac{\langle x-z, Mx\rangle}{\|x\|} = +\infty \tag{2.51}$$

and M is therefore coercive. □

Lemma 2.18 *Let $\mathcal{X}$ be a reflexive real Banach space such that $\mathcal{X} \neq \{0\}$, let $M_1\colon \mathcal{X} \to 2^{\mathcal{X}^*}$, and let $M_2\colon \mathcal{X} \to 2^{\mathcal{X}^*}$ be monotone. Suppose that there exists $z \in \operatorname{dom} M_1 \cap \operatorname{dom} M_2$ such that*

$$\lim_{\|x\|\to+\infty} \inf \frac{\langle x-z, M_1x\rangle}{\|x\|} = +\infty. \tag{2.52}$$

Then $M_1 + M_2$ is coercive.

Proof. Suppose that $x \in (\operatorname{dom} M_1 \cap \operatorname{dom} M_2) \smallsetminus \{0\}$, let $x^* \in (M_1 + M_2)x$, and let $z^* \in (M_1 + M_2)z$. Then there exist $x_1^* \in M_1x$, $x_2^* \in M_2x$, $z_1^* \in M_1z$, and $z_2^* \in M_2z$ such that $x^* = x_1^* + x_2^*$ and $z^* = z_1^* + z_2^*$. In turn, the monotonicity of M_2 yields

$$\begin{aligned}\frac{\langle x-z, x^*\rangle}{\|x\|} &= \frac{\langle x-z, x_1^*-z_1^*\rangle}{\|x\|} + \frac{\langle x-z, x_2^*-z_2^*\rangle}{\|x\|} + \frac{\langle x-z, z^*\rangle}{\|x\|}\\ &\geqslant \frac{\langle x-z, x_1^*\rangle}{\|x\|} + \frac{\langle x-z, z_2^*\rangle}{\|x\|}\\ &\geqslant \frac{\langle x-z, x_1^*\rangle - \|x-z\|\,\|z_2^*\|}{\|x\|}\end{aligned} \tag{2.53}$$

and (2.52) implies that $M_1 + M_2$ is coercive. □

Lemma 2.19 *Let $\mathcal{X}$ be a reflexive real Banach space such that $\mathcal{X} \neq \{0\}$, let $M_1\colon \mathcal{X} \to 2^{\mathcal{X}^*}$, let $M_2\colon \mathcal{X} \to 2^{\mathcal{X}^*}$ be monotone, let $(x_n^*)_{n\in\mathbb{N}}$ be a bounded sequence in $\mathcal{X}^*$, and let $(\gamma_n)_{n\in\mathbb{N}}$ be a bounded sequence in $]0,+\infty[$. Suppose that there exists $z \in \operatorname{dom} M_1 \cap \operatorname{dom} M_2$ such that*

$$\lim_{\|x\|\to+\infty} \inf \frac{\langle x-z, M_1x\rangle}{\|x\|} = +\infty, \tag{2.54}$$

and that

$$(\forall n \in \mathbb{N}) \quad x_n \in (M_1 + \gamma_n M_2)^{-1} x_n^*. \tag{2.55}$$

Then $(x_n)_{n\in\mathbb{N}}$ is bounded.

Proof. Set $\beta = \sup_{n\in\mathbb{N}} \|x_n^*\|$ and $\sigma = \sup_{n\in\mathbb{N}} \gamma_n$. It follows from (2.55) that, for every $n \in \mathbb{N}$, there exist $a_n^* \in M_1x_n$ and $b_n^* \in M_2x_n$ such that $x_n^* = a_n^* + \gamma_n b_n^*$. If $(x_n)_{n\in\mathbb{N}}$

is unbounded, there exists a strictly increasing sequence $(k_n)_{n\in\mathbb{N}}$ in $\mathbb{N}$ such that $0 < \|x_{k_n}\| \uparrow +\infty$. Therefore, (2.54) yields

$$\lim_{n\to+\infty} \frac{\left\langle x_{k_n} - z, a^*_{k_n}\right\rangle}{\|x_{k_n}\|} = +\infty. \tag{2.56}$$

Now let $z^* \in M_2 z$. By monotonicity of M_2, $(\forall n \in \mathbb{N})$ $\left\langle x_{k_n} - z, b^*_{k_n} - z^*\right\rangle \geqslant 0$. Hence, (2.56) implies that

$$\begin{aligned}
\beta\left(1 + \frac{\|z\|}{\|x_{k_0}\|}\right) &\geqslant \beta\left(1 + \frac{\|z\|}{\|x_{k_n}\|}\right)\\
&\geqslant \beta\frac{\|x_{k_n} - z\|}{\|x_{k_n}\|}\\
&\geqslant \frac{\left\langle x_{k_n} - z, x^*_{k_n}\right\rangle}{\|x_{k_n}\|}\\
&= \frac{\left\langle x_{k_n} - z, a^*_{k_n}\right\rangle}{\|x_{k_n}\|} + \gamma_{k_n}\frac{\left\langle x_{k_n} - z, b^*_{k_n} - z^*\right\rangle}{\|x_{k_n}\|} + \gamma_{k_n}\frac{\left\langle x_{k_n} - z, z^*\right\rangle}{\|x_{k_n}\|}\\
&\geqslant \frac{\left\langle x_{k_n} - z, a^*_{k_n}\right\rangle}{\|x_{k_n}\|} - \sigma\frac{\|x_{k_n} - z\|\,\|z^*\|}{\|x_{k_n}\|}\\
&\geqslant \frac{\left\langle x_{k_n} - z, a^*_{k_n}\right\rangle}{\|x_{k_n}\|} - \sigma\|z^*\|\left(1 + \frac{\|z\|}{\|x_{k_0}\|}\right)\\
&\to +\infty,
\end{aligned} \tag{2.57}$$

and we reach a contradiction. □

Corollary 2.20 *Let $\mathcal{X}$ be a reflexive real Banach space such that $\mathcal{X} \neq \{0\}$ and let $M\colon \mathcal{X} \to 2^{\mathcal{X}^*}$ be coercive. Then M^{-1} is bounded.*

Proof. Take $M_1 = M$ and $M_2 = 0$ in Lemma 2.19. □

Proposition 2.21 *Let $\mathcal{X}$ be a reflexive real Banach space such that $\mathcal{X} \neq \{0\}$, let $h \in \Gamma_0(\mathcal{X})$ be essentially smooth, and let $M\colon \mathcal{X} \to 2^{\mathcal{X}^*}$ be such that $\operatorname{dom} M \cap \operatorname{int}\operatorname{dom} h \neq \varnothing$. Suppose that one of the following holds :*

(i) $\operatorname{dom} M \cap \operatorname{int}\operatorname{dom} h$ *is bounded.*

(ii) *There exists $z \in \operatorname{dom} M \cap \operatorname{int}\operatorname{dom} h$ such that*

$$\lim_{\|x\|\to+\infty} \inf \frac{\langle x - z, Mx\rangle}{\|x\|} = +\infty. \tag{2.58}$$

(iii) *M is uniformly monotone at a point $z \in \operatorname{dom} M \cap \operatorname{int}\operatorname{dom} h$ with a supercoercive modulus.*

(iv) *M is monotone and h is supercoercive.*

(v) *M is monotone and h is uniformly convex at a point $z \in \operatorname{dom} M \cap \operatorname{int}\operatorname{dom} h$, i.e., there exists an increasing function $\phi\colon [0,+\infty[\to [0,+\infty]$ that vanishes only at 0 such that*

$$(\forall y \in \operatorname{dom} h)(\forall \alpha \in \,]0,1[)\;\; h(\alpha y+(1-\alpha)z)+\alpha(1-\alpha)\phi(\|y-z\|) \leqslant \alpha h(y)+(1-\alpha)h(z). \tag{2.59}$$

Then $\nabla h+M$ is coercive. If, in addition, M is maximally monotone, then $\operatorname{dom}(\nabla h+M)^{-1} = \mathcal{X}^*$.

Proof. We first observe that [34, Theorem 18.7] and [6, Theorem 5.6] imply that ∇h is maximally monotone and that $\operatorname{dom}\nabla h = \operatorname{int}\operatorname{dom} h$.

(i) : Lemma 2.17(i).

(ii) : It follows from Lemma 2.18 that $\nabla h + M$ is coercive.

(iii) : Since $\nabla h + M$ is uniformly monotone at z with a supercoercive modulus, the claim follows from Lemma 2.17(ii).

(iv) : Let $z \in \operatorname{dom} M \cap \operatorname{int}\operatorname{dom} h = \operatorname{dom} M \cap \operatorname{dom}\partial h$. Then we derive from (2.51) that

$$\lim_{\|x\|\to+\infty} \frac{\langle x-z, \nabla h(x)\rangle}{\|x\|} = +\infty. \tag{2.60}$$

Thus, ∇h satisfies (2.52) and it follows from Lemma 2.18 that $\nabla h + M$ is coercive.

(v) : It follows from [38, Definition 2.2 and Remark 2.8] that ∇h is uniformly monotone at z with a supercoercive modulus. Hence, $\nabla h + M$ is likewise and Lemma 2.17(ii) implies that $\nabla h + M$ is coercive. Alternatively, this is a special case of (iv).

Finally, suppose that M is maximally monotone. Then [34, Theorem 24.1(a)] asserts that $\nabla h + M$ is maximally monotone. Consequently, since $\nabla h + M$ is coercive, it follows from [41, Corollary II-B.32.35] that $\operatorname{dom}(\nabla h + M)^{-1} = \operatorname{ran}(\nabla h + M) = \mathcal{X}^*$. ☐

Lemma 2.22 *Let $\mathcal{X}$ and $\mathcal{Y}$ be real Banach spaces, let $D \subset \mathcal{X}$ be a nonempty open set, and let C be a nonempty bounded convex subset of D. Suppose that $T\colon D \to \mathcal{Y}$ is uniformly continuous on C in the sense that*

$$(\forall \varepsilon \in \,]0,+\infty[)(\exists\, \delta \in \,]0,+\infty[)(\forall x \in C)(\forall y \in C) \quad \|x-y\| \leqslant \delta \;\Rightarrow\; \|Tx-Ty\| \leqslant \varepsilon. \tag{2.61}$$

Then T is bounded on C.

Proof. In view of (2.61), there exists $\delta \in \,]0,+\infty[$ such that

$$(\forall x \in C)(\forall y \in C) \quad \|x-y\| \leqslant \delta \quad\Rightarrow\quad \|Tx-Ty\| \leqslant 1. \tag{2.62}$$

Now fix $z \in C$ and $\rho \in \left]0,+\infty\right[$ such that $C \subset \{x \in \mathcal{X} \mid \|x-z\| \leqslant \rho\}$, and take an integer $m \geqslant 1+\rho/\delta$. Let $x \in C$ and set

$$(\forall n \in \{0,\ldots,m\}) \quad x_n = x + \frac{n}{m}(z-x) \in C. \tag{2.63}$$

Then, for every $n \in \{0,\ldots,m-1\}$, $\|x_{n+1}-x_n\| = \|z-x\|/m \leqslant \rho/m \leqslant \delta$ and (2.62) yields $\|Tx_{n+1}-Tx_n\| \leqslant 1$. Hence, $\|Tz-Tx\| \leqslant \sum_{n=0}^{m-1}\|Tx_{n+1}-Tx_n\| \leqslant m$ and therefore $\|Tx\| < \|Tz\|+m$. We conclude that $\sup_{x\in C}\|Tx\| \leqslant \|Tz\|+m$. ☐

2.2.3 Best Bregman approximation algorithm

Proposition 2.11(i)(a) asserts that Problem 2.10 reduces to finding the Bregman projection of a reference point (x_0, y_0^*) onto the closed convex subset $\boldsymbol{C} = \boldsymbol{Z} \cap \operatorname{dom} \boldsymbol{f}$ of $\operatorname{dom} \boldsymbol{f}$. Our strategy is to employ Proposition 2.14 for this task. The following condition will be used subsequently (see [7, Examples 4.10, 5.11, and 5.13] for special cases).

Condition 2.23 [8, Condition 4.3(ii)] Let $\mathcal{X}$ be a reflexive real Banach space and let $h \in \Gamma_0(\mathcal{X})$ be Gâteaux differentiable on $\operatorname{int}\operatorname{dom} h \neq \varnothing$. For every sequence $(x_n)_{n\in\mathbb{N}}$ in $\operatorname{int}\operatorname{dom} h$ and every bounded sequence $(y_n)_{n\in\mathbb{N}}$ in $\operatorname{int}\operatorname{dom} h$,

$$D^h(x_n,y_n) \to 0 \quad \Rightarrow \quad x_n - y_n \to 0. \tag{2.64}$$

We now derive from Proposition 2.14 our best Bregman approximation algorithm to solve Problem 2.10.

Theorem 2.24 *Consider the setting of Problem 2.10. Let $h \in \Gamma_0(\mathcal{X})$ and $j \in \Gamma_0(\mathcal{Y})$ be Legendre functions such that $\operatorname{int}\operatorname{dom} f \subset \operatorname{int}\operatorname{dom} h$, $L(\operatorname{int}\operatorname{dom} f) \subset \operatorname{int}\operatorname{dom} j$, and there exist ε and δ in $\left]0,+\infty\right[$ such that $\nabla h + \varepsilon A$ and $\nabla j + \delta B$ are coercive. Let $\sigma \in \left[\max\{\varepsilon,\delta\},+\infty\right[$ and iterate*

$$\begin{array}{l}
\text{for } n = 0,1,\ldots \\
\left\lfloor \begin{array}{l}
(\gamma_n,\mu_n) \in [\varepsilon,\sigma]\times[\delta,\sigma] \\
a_n = (\nabla h + \gamma_n A)^{-1}\big(\nabla h(x_n) - \gamma_n L^* y_n^*\big) \\
a_n^* = \gamma_n^{-1}\big(\nabla h(x_n) - \nabla h(a_n)\big) - L^* y_n^* \\
b_n = (\nabla j + \mu_n B)^{-1}\big(\nabla j(Lx_n) + \mu_n y_n^*\big) \\
b_n^* = \mu_n^{-1}\big(\nabla j(Lx_n) - \nabla j(b_n)\big) + y_n^* \\
\boldsymbol{H}_n = \big\{(x,y^*) \in \boldsymbol{\mathcal{X}} \;\big|\; \langle x, a_n^* + L^* b_n^*\rangle + \langle b_n - La_n, y^*\rangle \leqslant \langle a_n, a_n^*\rangle + \langle b_n, b_n^*\rangle\big\} \\
(x_{n+1/2}, y_{n+1/2}^*) = P_{\boldsymbol{H}_n}^{\boldsymbol{f}}(x_n, y_n^*) \\
(x_{n+1}, y_{n+1}^*) = Q^{\boldsymbol{f}}\big((x_0,y_0^*),(x_n,y_n^*),(x_{n+1/2},y_{n+1/2}^*)\big).
\end{array}\right.
\end{array} \tag{2.65}$$

Then the following hold :

(i) *Let* $n \in \mathbb{N}$. *Then the following are equivalent :*

(a) $(x_n, y_n^*) = (\overline{x}, \overline{y}^*)$.

(b) $(x_n, y_n^*) \in \boldsymbol{Z}$.

(c) $(x_n, y_n^*) \in \boldsymbol{H}_n$.

(d) $x_n = a_n$ *and* $y_n^* = b_n^*$.

(e) $La_n = b_n$ *and* $a_n^* = -L^* b_n^*$.

(f) $\boldsymbol{H}_n = \boldsymbol{\mathcal{X}}$.

(g) $(x_{n+1/2}, y_{n+1/2}^*) = (x_n, y_n^*)$.

(h) $(x_{n+1}, y_{n+1}^*) = (x_n, y_n^*)$.

(ii) $\sum_{n\in\mathbb{N}} D^f(x_{n+1}, x_n) < +\infty$ *and* $\sum_{n\in\mathbb{N}} D^{g^*}(y_{n+1}^*, y_n^*) < +\infty$.

(iii) $\sum_{n\in\mathbb{N}} D^f(x_{n+1/2}, x_n) < +\infty$ *and* $\sum_{n\in\mathbb{N}} D^{g^*}(y_{n+1/2}^*, y_n^*) < +\infty$.

(iv) *Suppose that* f, g^*, h, *and* j *satisfy Condition 2.23, and that* ∇h *and* ∇j *are uniformly continuous on every bounded subset of* $\operatorname{int}\operatorname{dom} h$ *and* $\operatorname{int}\operatorname{dom} j$, *respectively. Then* $x_n \to \overline{x}$ *and* $y_n^* \to \overline{y}^*$.

Proof. We apply Proposition 2.14 to

$$\boldsymbol{C} = \boldsymbol{Z} \cap \overline{\operatorname{dom} \boldsymbol{f}}. \tag{2.66}$$

It follows from Proposition 2.11(i)(a) and our assumptions that $\boldsymbol{C}$ is a closed convex subset of $\overline{\operatorname{dom} \boldsymbol{f}}$ and that $\boldsymbol{C} \cap \operatorname{int}\operatorname{dom} \boldsymbol{f} \neq \varnothing$. Moreover, since f and g are Legendre functions, so are f^* and g^* [6, Corollary 5.5]. Therefore, it follows from [6, Theorem 5.6(iii)] that ∂f and ∂g^* are single-valued on $\operatorname{dom} \partial f = \operatorname{int}\operatorname{dom} f$ and $\operatorname{dom} \partial g^* = \operatorname{int}\operatorname{dom} g^*$, respectively. On the other hand, we derive from (2.28) that $\operatorname{dom} \boldsymbol{f} = \operatorname{dom} f \times \operatorname{dom} g^*$ and that

$$\partial \boldsymbol{f} \colon \boldsymbol{\mathcal{X}} \to 2^{\boldsymbol{\mathcal{X}}^*} \colon (x, y^*) \mapsto \partial f(x) \times \partial g^*(y^*). \tag{2.67}$$

Thus, $\partial \boldsymbol{f}$ is single-valued on

$$\operatorname{dom} \partial \boldsymbol{f} = \operatorname{dom} \partial f \times \operatorname{dom} \partial g^* = \operatorname{int}\operatorname{dom} f \times \operatorname{int}\operatorname{dom} g^* = \operatorname{int}(\operatorname{dom} f \times \operatorname{dom} g^*) = \operatorname{int}\operatorname{dom} \boldsymbol{f}. \tag{2.68}$$

Likewise, since

$$\boldsymbol{f}^* \colon \mathcal{X}^* \to \left]-\infty, +\infty\right] \colon (x^*, y) \mapsto f^*(x^*) + g(y), \tag{2.69}$$

we deduce that $\partial \boldsymbol{f}^*$ is single-valued on $\operatorname{dom} \partial \boldsymbol{f}^* = \operatorname{int}\operatorname{dom} \boldsymbol{f}^*$. Consequently, [6, Theorems 5.4 and 5.6] assert that

$$\boldsymbol{f} \text{ is a Legendre function.} \tag{2.70}$$

Now let $\gamma \in [\varepsilon, +\infty[$ and let $\mu \in [\delta, +\infty[$. Since h is strictly convex, ∇h is strictly monotone [39, Theorem 2.4.4(ii)] and $\nabla h + \gamma A$ is likewise. Let (x^*, x_1) and (x^*, x_2) be two elements in $\operatorname{gra}(\nabla h + \gamma A)^{-1}$ such that $x_1 \neq x_2$. Then (x_1, x^*) and (x_2, x^*) lie in $\operatorname{gra}(\nabla h + \gamma A)$ and the strict monotonicity of $\nabla h + \gamma A$ implies that

$$0 = \langle x_1 - x_2, x^* - x^* \rangle > 0, \tag{2.71}$$

which is impossible. Thus,

$$(\nabla h + \gamma A)^{-1} \quad \text{is at most single-valued.} \tag{2.72}$$

The same argument shows that

$$(\nabla j + \mu B)^{-1} \quad \text{is at most single-valued.} \tag{2.73}$$

On the other hand, by assumption, there exists $(x, y^*) \in \boldsymbol{Z} \cap \operatorname{int}\operatorname{dom} \boldsymbol{f}$. It follows from (2.27) that $x \in \operatorname{dom} A$ and $Lx \in \operatorname{dom} B$. Furthermore, (2.68) yields $x \in \operatorname{int}\operatorname{dom} f$. Therefore

$$\begin{cases} x \in \operatorname{dom} A \cap \operatorname{int}\operatorname{dom} f \subset \operatorname{dom} A \cap \operatorname{int}\operatorname{dom} h \\ Lx \in \operatorname{dom} B \cap L(\operatorname{int}\operatorname{dom} f) \subset \operatorname{dom} B \cap \operatorname{int}\operatorname{dom} j. \end{cases} \tag{2.74}$$

Thus, $\operatorname{dom} A \cap \operatorname{int}\operatorname{dom} h \neq \varnothing$ and $\operatorname{dom} B \cap \operatorname{int}\operatorname{dom} j \neq \varnothing$. It therefore follows from Lemma 2.18 that

$$\nabla h + \gamma A = (\nabla h + \varepsilon A) + (\gamma - \varepsilon) A \quad \text{and} \quad \nabla j + \mu B = (\nabla j + \delta B) + (\mu - \delta) B \ \text{ are coercive.} \tag{2.75}$$

Altogether, (2.72), (2.73), (2.75), and Proposition 2.21 assert that the operators

$$\begin{cases} (\nabla h + \gamma A)^{-1} \colon \mathcal{X}^* \to \operatorname{dom} A \cap \operatorname{int}\operatorname{dom} h \\ (\nabla j + \mu B)^{-1} \colon \mathcal{Y}^* \to \operatorname{dom} B \cap \operatorname{int}\operatorname{dom} j \end{cases} \tag{2.76}$$

are well defined and single-valued. Now set

$$(\forall n \in \mathbb{N}) \quad \boldsymbol{x}_n = (x_n, y_n^*) \quad \text{and} \quad \boldsymbol{x}_{n+1/2} = (x_{n+1/2}, y_{n+1/2}^*). \tag{2.77}$$

Since (2.65) yields

$$(\forall n \in \mathbb{N}) \quad (a_n, a_n^*) \in \operatorname{gra} A \quad \text{and} \quad (b_n, b_n^*) \in \operatorname{gra} B, \tag{2.78}$$

it follows from (2.66), Proposition 2.11(ii)(b), (2.65), and Definition 2.12 that

$$(\forall n \in \mathbb{N}) \quad \varnothing \neq \boldsymbol{C} \subset \boldsymbol{Z} \subset \boldsymbol{H}_n = H^{\boldsymbol{f}}(\boldsymbol{x}_n, \boldsymbol{x}_{n+1/2}). \tag{2.79}$$

Hence, appealing to (2.70) and (2.23), we see that

$$(\forall n \in \mathbb{N}) \quad P^{\boldsymbol{f}}_{\boldsymbol{H}_n} \colon \operatorname{int}\operatorname{dom} \boldsymbol{f} \to \boldsymbol{H}_n \cap \operatorname{int}\operatorname{dom} \boldsymbol{f} \tag{2.80}$$

and that

$$(\forall n \in \mathbb{N}) \quad \boldsymbol{x}_{n+1} = Q^{\boldsymbol{f}}\big(\boldsymbol{x}_0, \boldsymbol{x}_n, \boldsymbol{x}_{n+1/2}\big). \tag{2.81}$$

Thus, we derive from (2.77), (2.79), and Proposition 2.14(ii) that $(x_n)_{n\in\mathbb{N}}$ and $(y_n^*)_{n\in\mathbb{N}}$ are well-defined sequences in $\operatorname{int}\operatorname{dom} f$ and $\operatorname{int}\operatorname{dom} g^*$, respectively.

(i) : We prove the following implications.

(i)(a)⇒(i)(b) : Clear.

(i)(b)⇒(i)(a) : Proposition 2.14(iii).

(i)(b)⇒(i)(c) : Clear by (2.79).

(i)(c)⇒(i)(d) : In view of (2.65),

$$\begin{aligned} 0 &\geqslant \langle x_n, a_n^* + L^* b_n^* \rangle + \langle b_n - L a_n, y_n^* \rangle - \langle a_n, a_n^* \rangle - \langle b_n, b_n^* \rangle \\ &= \langle x_n - a_n, a_n^* + L^* y_n^* \rangle + \langle L x_n - b_n, b_n^* - y_n^* \rangle \\ &= \gamma_n^{-1} \langle x_n - a_n, \nabla h(x_n) - \nabla h(a_n) \rangle + \mu_n^{-1} \langle L x_n - b_n, \nabla j(L x_n) - \nabla j(b_n) \rangle . \end{aligned} \tag{2.82}$$

Consequently, the strict monotonicity of ∇h and ∇j yields

$$x_n = a_n \quad \text{and} \quad L x_n = b_n. \tag{2.83}$$

Furthermore,

$$b_n^* = \mu_n^{-1}\big(\nabla j(L x_n) - \nabla j(b_n)\big) + y_n^* = \mu_n^{-1}\big(\nabla j(b_n) - \nabla j(b_n)\big) + y_n^* = y_n^*. \tag{2.84}$$

(i)(d)⇒(i)(e) : We derive from (2.65) that $a_n^* = \gamma_n^{-1}(\nabla h(x_n) - \nabla h(a_n)) - L^* y_n^* = -L^* y_n^* = -L^* b_n^*$. On the other hand, since

$$\begin{aligned} \langle L a_n - b_n, \nabla j(L a_n) - \nabla j(b_n) \rangle &= \langle L x_n - b_n, \nabla j(L x_n) - \nabla j(b_n) \rangle \\ &= \mu_n \langle L x_n - b_n, b_n^* - y_n^* \rangle \\ &= 0, \end{aligned} \tag{2.85}$$

the strict monotonicity of ∇j yields $L a_n = b_n$.

(i)(e)⇔(i)(f) : Proposition 2.11(ii)(a).

(i)(f)⇒(i)(g) : Indeed, $\boldsymbol{x}_{n+1/2} = P^{\boldsymbol{f}}_{\boldsymbol{H}_n} \boldsymbol{x}_n = \boldsymbol{x}_n$.

(i)(g)⇒(i)(h) : We have

$$\boldsymbol{x}_{n+1} = Q^{\boldsymbol{f}}(\boldsymbol{x}_0, \boldsymbol{x}_n, \boldsymbol{x}_{n+1/2}) = P^{\boldsymbol{f}}_{H^{\boldsymbol{f}}(\boldsymbol{x}_0, \boldsymbol{x}_n) \cap H^{\boldsymbol{f}}(\boldsymbol{x}_n, \boldsymbol{x}_{n+1/2})} \boldsymbol{x}_0 = P^{\boldsymbol{f}}_{H^{\boldsymbol{f}}(\boldsymbol{x}_0, \boldsymbol{x}_n)} \boldsymbol{x}_0 = \boldsymbol{x}_n. \tag{2.86}$$

(i)(h)⇒(i)(g) : By Proposition 2.14(v), $0 \leqslant D^{\boldsymbol{f}}(\boldsymbol{x}_{n+1/2}, \boldsymbol{x}_n) \leqslant D^{\boldsymbol{f}}(\boldsymbol{x}_{n+1}, \boldsymbol{x}_n) = 0$. Therefore $D^{\boldsymbol{f}}(\boldsymbol{x}_{n+1/2}, \boldsymbol{x}_n) = 0$ and we derive from [6, Lemma 7.3(vi)] that $\boldsymbol{x}_{n+1/2} = \boldsymbol{x}_n$.

(i)(g)⇒(i)(c) : Indeed, $\boldsymbol{x}_n = \boldsymbol{x}_{n+1/2} = P^{\boldsymbol{f}}_{\boldsymbol{H}_n}\boldsymbol{x}_n \in \boldsymbol{H}_n$.

(i)(d)⇒(i)(b) : We derive from (2.65) that

$$\nabla h(x_n) - \gamma_n L^* y_n^* \in \nabla h(a_n) + \gamma_n A a_n = \nabla h(x_n) + \gamma_n A x_n. \tag{2.87}$$

Hence $-L^* y_n^* \in Ax_n$. Likewise, as in (2.85), we first obtain $Lx_n = b_n$ and then

$$\nabla j(Lx_n) + \mu_n y_n^* \in \nabla j(b_n) + \mu_n B b_n = \nabla j(Lx_n) + \mu_n B(Lx_n). \tag{2.88}$$

Thus, $y_n^* \in B(Lx_n)$, i.e., $Lx_n \in B^{-1} y_n^*$. In view of (2.27), the implication is proved.

(ii) : Proposition 2.14(iv) yields

$$\sum_{n\in\mathbb{N}} D^f(x_{n+1}, x_n) + \sum_{n\in\mathbb{N}} D^{g^*}(y_{n+1}^*, y_n^*) = \sum_{n\in\mathbb{N}} D^{\boldsymbol{f}}(\boldsymbol{x}_{n+1}, \boldsymbol{x}_n) < +\infty. \tag{2.89}$$

(iii) : Proposition 2.14(vi) yields

$$\sum_{n\in\mathbb{N}} D^f(x_{n+1/2}, x_n) + \sum_{n\in\mathbb{N}} D^{g^*}(y_{n+1/2}^*, y_n^*) = \sum_{n\in\mathbb{N}} D^{\boldsymbol{f}}(\boldsymbol{x}_{n+1/2}, \boldsymbol{x}_n) < +\infty. \tag{2.90}$$

(iv) : Proposition 2.14(ii) implies that $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ is a bounded sequence in $\operatorname{int}\operatorname{dom}\boldsymbol{f}$. In turn,

$$(x_n)_{n\in\mathbb{N}} \in (\operatorname{int}\operatorname{dom} f)^{\mathbb{N}} \quad \text{and} \quad (y_n^*)_{n\in\mathbb{N}} \in (\operatorname{int}\operatorname{dom} g^*)^{\mathbb{N}} \quad \text{are bounded.} \tag{2.91}$$

On the other hand, by (2.65),

$$(\forall n \in \mathbb{N}) \quad (x_{n+1/2}, y_{n+1/2}^*) = \boldsymbol{x}_{n+1/2} = P^{\boldsymbol{f}}_{\boldsymbol{H}_n}\boldsymbol{x}_n \in \boldsymbol{H}_n \tag{2.92}$$

and

$$(\forall n \in \mathbb{N}) \quad \left\langle x_{n+1/2}, a_n^* + L^* b_n^* \right\rangle + \left\langle b_n - La_n, y_{n+1/2}^* \right\rangle = \langle a_n, a_n^* \rangle + \langle b_n, b_n^* \rangle . \tag{2.93}$$

Therefore,

$$\begin{aligned}(\forall n \in \mathbb{N}) \quad & \|x_n - x_{n+1/2}\| \, \|a_n^* + L^* b_n^*\| + \|b_n - La_n\| \, \|y_n^* - y_{n+1/2}^*\| \\ & \geqslant \left\langle x_n - x_{n+1/2}, a_n^* + L^* b_n^* \right\rangle + \left\langle b_n - La_n, y_n^* - y_{n+1/2}^* \right\rangle \\ & = \langle x_n, a_n^* + L^* b_n^* \rangle + \langle b_n - La_n, y_n^* \rangle - \langle a_n, a_n^* \rangle - \langle b_n, b_n^* \rangle \\ & = \langle x_n - a_n, a_n^* + L^* y_n^* \rangle + \langle Lx_n - b_n, b_n^* - y_n^* \rangle \\ & = \gamma_n^{-1} \langle x_n - a_n, \nabla h(x_n) - \nabla h(a_n) \rangle + \mu_n^{-1} \langle Lx_n - b_n, \nabla j(Lx_n) - \nabla j(b_n) \rangle \\ & \geqslant \sigma^{-1}\big(D^h(x_n, a_n) + D^h(a_n, x_n) + D^j(Lx_n, b_n) + D^j(b_n, Lx_n)\big) \\ & \geqslant \sigma^{-1}\big(D^h(x_n, a_n) + D^j(Lx_n, b_n)\big). \end{aligned} \tag{2.94}$$

However, since (iii) yields

$$D^f(x_{n+1/2}, x_n) \to 0 \quad \text{and} \quad D^{g^*}(y^*_{n+1/2}, y^*_n) \to 0 \tag{2.95}$$

and since f and g^* satisfy Condition 2.23, (2.64) yields

$$x_{n+1/2} - x_n \to 0 \quad \text{and} \quad y^*_{n+1/2} - y^*_n \to 0. \tag{2.96}$$

Since ∇h is uniformly continuous on every bounded subset of $\operatorname{int}\operatorname{dom} h$, Lemma 2.22 asserts that ∇h is bounded on every bounded subset of $\operatorname{int}\operatorname{dom} h$ and hence, since $\operatorname{int}\operatorname{dom} f \subset \operatorname{int}\operatorname{dom} h$ and L^* is bounded, it follows from (2.91) that $\big(\nabla h(x_n) - \gamma_n L^* y^*_n\big)_{n\in\mathbb{N}}$ is bounded. We therefore deduce from (2.76), (2.65), and Lemma 2.19 that

$$(a_n)_{n\in\mathbb{N}} \in (\operatorname{int}\operatorname{dom} h)^{\mathbb{N}} \;\text{ is bounded.} \tag{2.97}$$

Similarly, since ∇j is uniformly continuous on every bounded subset of $\operatorname{int}\operatorname{dom} j$ and $L(\operatorname{int}\operatorname{dom} f) \subset \operatorname{int}\operatorname{dom} j$, it follows from (2.91) and Lemma 2.22 that $\big(\nabla j(Lx_n) + \mu_n y^*_n\big)_{n\in\mathbb{N}}$ is bounded and hence (2.76), (2.65), and Lemma 2.19 yield

$$(b_n)_{n\in\mathbb{N}} \in (\operatorname{int}\operatorname{dom} j)^{\mathbb{N}} \;\text{ is bounded.} \tag{2.98}$$

Thus, $(\nabla h(x_n))_{n\in\mathbb{N}}$, $(\nabla h(a_n))_{n\in\mathbb{N}}$, $(\nabla j(Lx_n))_{n\in\mathbb{N}}$, and $(\nabla j(b_n))_{n\in\mathbb{N}}$ are bounded and we deduce from (2.65) that

$$(a^*_n)_{n\in\mathbb{N}} \quad \text{and} \quad (b^*_n)_{n\in\mathbb{N}} \quad \text{are bounded.} \tag{2.99}$$

We therefore derive from (2.94), (2.96), (2.97), and (2.98) that

$$D^h(x_n, a_n) \to 0 \quad \text{and} \quad D^j(Lx_n, b_n) \to 0. \tag{2.100}$$

Since h and j satisfy Condition 2.23, we get

$$x_n - a_n \to 0 \quad \text{and} \quad Lx_n - b_n \to 0. \tag{2.101}$$

Therefore, since ∇h is uniformly continuous on every bounded subset of $\operatorname{int}\operatorname{dom} h$ and ∇j is uniformly continuous on every bounded subset of $\operatorname{int}\operatorname{dom} j$,

$$\nabla h(x_n) - \nabla h(a_n) \to 0 \quad \text{and} \quad \nabla j(Lx_n) - \nabla j(b_n) \to 0. \tag{2.102}$$

Hence, using (2.65), we get

$$a^*_n + L^* y^*_n \to 0 \quad \text{and} \quad b^*_n - y^*_n \to 0. \tag{2.103}$$

Now, let $\boldsymbol{x} = (x, y^*) \in \mathfrak{W}(\boldsymbol{x}_n)_{n\in\mathbb{N}}$, say $\boldsymbol{x}_{k_n} \rightharpoonup \boldsymbol{x}$. Then $x_{k_n} \rightharpoonup x$ and $y^*_{k_n} \rightharpoonup y^*$, and we derive from (2.101) and (2.103) that

$$\begin{cases} a_{k_n} \rightharpoonup x \\ b^*_{k_n} \rightharpoonup y^* \end{cases} \quad \text{and} \quad \begin{cases} La_{k_n} - b_{k_n} \to 0 \\ a^*_{k_n} + L^* b^*_{k_n} \to 0. \end{cases} \tag{2.104}$$

It therefore follows from (2.78), Proposition 2.11(iii), and (2.91) that $\boldsymbol{x} \in \boldsymbol{Z} \cap \overline{\operatorname{dom} \boldsymbol{f}} = \boldsymbol{C}$. Hence, we derive from Proposition 2.14(vii) that

$$D^f(x_n, \overline{x}) + D^{g^*}(y_n^*, \overline{y}^*) = D^{\boldsymbol{f}}(\boldsymbol{x}_n, \overline{\boldsymbol{x}}) \to 0, \tag{2.105}$$

where $\overline{\boldsymbol{x}} = (\overline{x}, \overline{y}^*)$. Hence, $D^f(x_n, \overline{x}) \to 0$, $D^{g^*}(y_n^*, \overline{y}^*) \to 0$, and, since f and g^* satisfy Condition 2.23, we conclude that $x_n \to \overline{x}$ and $y_n^* \to \overline{y}^*$. □

Remark 2.25 We provide a couple of settings that satisfy the assumptions of Theorem 2.24.

(i) In Problem 2.10, suppose that $\mathcal{X}$ and $\mathcal{Y}$ are Hilbert spaces, that $f = \|\cdot\|^2/2$, and that $g = \|\cdot\|^2/2$. Furthermore, in Theorem 2.24, set $h = f$ and $j = g$, and note that, for any $\varepsilon \in \left]0, +\infty\right[$, $\nabla h + \varepsilon A = \mathrm{Id} + \varepsilon A$ and $\nabla j + \varepsilon B = \mathrm{Id} + \varepsilon B$ are strongly monotone and hence coercive by Lemma 2.17(ii). Then we recover the framework of [3], which has been applied to domain decomposition problems in [4].

(ii) Let $(\Omega_1, \mathcal{F}_1, \mu_1)$ and $(\Omega_2, \mathcal{F}_2, \mu_2)$ be measure spaces, let p and q be in $\left]1, +\infty\right[$, and set $p^* = p/(p-1)$ and $q^* = q/(q-1)$. In Problem 2.10, suppose that $\mathcal{X} = L^p(\Omega_1, \mathcal{F}_1, \mu_1)$, $\mathcal{Y} = L^q(\Omega_2, \mathcal{F}_2, \mu_2)$, $f = \|\cdot\|^p/p$, and $g = \|\cdot\|^q/q$. Then $\mathcal{X}^* = L^{p^*}(\Omega_1, \mathcal{F}_1, \mu_1)$, $\mathcal{Y}^* = L^{q^*}(\Omega_2, \mathcal{F}_2, \mu_2)$, and $g^* = \|\cdot\|^{q^*}/q^*$. Moreover, it follows from Clarkson's theorem [16, Theorem II.4.7] that $\mathcal{X}$, $\mathcal{X}^*$, $\mathcal{Y}$, and $\mathcal{Y}^*$ are uniformly convex and uniformly smooth. Hence, we derive from [6, Corollary 5.5 and Example 6.5] that f, g, and g^* are Legendre functions which are uniformly convex on every bounded set, and which therefore satisfy Condition 2.23 by virtue of [7, Example 4.10(i)]. Now set $h = f$ and $j = g$ in Theorem 2.24. We derive from [16, Theorem II.2.16(i)] that ∇h and ∇j are uniformly continuous on every bounded subset of $\mathcal{X}$ and $\mathcal{Y}$, respectively. In addition, h and j are supercoercive and therefore, for any $\varepsilon \in \left]0, +\infty\right[$, Proposition 2.21(iv) asserts that $\nabla h + \varepsilon A$ and $\nabla j + \varepsilon B$ are coercive. Finally, it follows from [16, Proposition II.4.9] that $\nabla h\colon x \mapsto |x|^{p-1}\mathrm{sign}(x)$ and $\nabla j\colon y \mapsto |y|^{q-1}\mathrm{sign}(y)$.

Next, we consider a specialization of Problem 2.10 to multivariate structured minimization.

Problem 2.26 Let m and p be strictly positive integers, let $(\mathcal{X}_i)_{1 \leqslant i \leqslant m}$ and $(\mathcal{Y}_k)_{1 \leqslant k \leqslant p}$ be reflexive real Banach spaces, and let $\boldsymbol{\mathcal{X}}$ be the standard vector product space $\left(\bigtimes_{i=1}^m \mathcal{X}_i\right) \times \left(\bigtimes_{k=1}^p \mathcal{Y}_k^*\right)$ equipped with the norm

$$(x, y^*) = \big((x_i)_{1 \leqslant i \leqslant m}, (y_k^*)_{1 \leqslant k \leqslant p}\big) \mapsto \sqrt{\sum_{i=1}^m \|x_i\|^2 + \sum_{k=1}^p \|y_k^*\|^2}. \tag{2.106}$$

For every $i \in \{1, \ldots, m\}$ and every $k \in \{1, \ldots, p\}$, let $\varphi_i \in \Gamma_0(\mathcal{X}_i)$, let $\psi_k \in \Gamma_0(\mathcal{Y}_k)$, and let $L_{ki} \colon \mathcal{X}_i \to \mathcal{Y}_k$ be linear and bounded. Consider the primal problem

$$\underset{x_1 \in \mathcal{X}_1, \ldots, x_m \in \mathcal{X}_m}{\text{minimize}} \quad \sum_{i=1}^{m} \varphi_i(x_i) + \sum_{k=1}^{p} \psi_k\bigg(\sum_{i=1}^{m} L_{ki} x_i\bigg), \tag{2.107}$$

the dual problem

$$\underset{y_1^* \in \mathcal{Y}_1^*, \ldots, y_p^* \in \mathcal{Y}_p^*}{\text{minimize}} \quad \sum_{i=1}^{m} \varphi_i^*\bigg(-\sum_{k=1}^{p} L_{ki}^* y_k^*\bigg) + \sum_{k=1}^{p} \psi_k^*(y_k^*), \tag{2.108}$$

and let

$$\begin{aligned}\boldsymbol{Z} = \Bigg\{ (x, y^*) \in \boldsymbol{\mathcal{X}} \;\Bigg|\; & (\forall i \in \{1, \ldots, m\}) \quad -\sum_{k=1}^{p} L_{ki}^* y_k^* \in \partial\varphi_i(x_i) \;\text{ and} \\ & (\forall k \in \{1, \ldots, p\}) \quad \sum_{i=1}^{m} L_{ki} x_i \in \partial\psi_k^*(y_k^*) \Bigg\}\end{aligned} \tag{2.109}$$

be the associated Kuhn-Tucker set. For every $i \in \{1, \ldots, m\}$, let $f_i \in \Gamma_0(\mathcal{X}_i)$ be a Legendre function and let $x_{i,0} \in \operatorname{int}\operatorname{dom} f_i$. For every $k \in \{1, \ldots, p\}$, let $g_k \in \Gamma_0(\mathcal{Y}_k)$ be a Legendre function and let $y_{k,0}^* \in \operatorname{int}\operatorname{dom} g_k^*$. Set $x_0 = (x_{i,0})_{1 \leqslant i \leqslant m}$, $y_0^* = (y_{k,0}^*)_{1 \leqslant k \leqslant p}$, and

$$\boldsymbol{f} \colon \boldsymbol{\mathcal{X}} \to \left]-\infty, +\infty\right] : (x, y^*) \mapsto \sum_{i=1}^{m} f_i(x_i) + \sum_{k=1}^{p} g_k^*(y_k^*), \tag{2.110}$$

and suppose that $\boldsymbol{Z} \cap \operatorname{int}\operatorname{dom} \boldsymbol{f} \neq \varnothing$. The objective is to find the best Bregman approximation $\big((\overline{x}_i)_{1 \leqslant i \leqslant m}, (\overline{y}_k^*)_{1 \leqslant k \leqslant p}\big) = P_{\boldsymbol{Z}}^{\boldsymbol{f}}(x_0, y_0^*)$ to (x_0, y_0^*) from $\boldsymbol{Z}$.

We derive from Theorem 2.24 the following convergence result for a splitting algorithm to solve Problem 2.26.

Proposition 2.27 *Consider the setting of Problem 2.26. For every $i \in \{1, \ldots, m\}$, let $h_i \in \Gamma_0(\mathcal{X}_i)$ be a Legendre function such that $\operatorname{int}\operatorname{dom} f_i \subset \operatorname{int}\operatorname{dom} h_i$ and $h_i + \varepsilon_i \varphi_i$ is supercoercive for some $\varepsilon_i \in \left]0, +\infty\right[$. For every $k \in \{1, \ldots, p\}$, let $j_k \in \Gamma_0(\mathcal{Y}_k)$ be a Legendre function such that $\sum_{i=1}^{m} L_{ki}(\operatorname{int}\operatorname{dom} f_i) \subset \operatorname{int}\operatorname{dom} j_k$ and $j_k + \delta_k \psi_k$ is supercoercive for some $\delta_k \in \left]0, +\infty\right[$. Set $\varepsilon = \max_{1 \leqslant i \leqslant m} \varepsilon_i$ and $\delta = \max_{1 \leqslant k \leqslant p} \delta_k$, let $\sigma \in \left[\max\{\varepsilon, \delta\}, +\infty\right[$, and* *iterate*

$$
\begin{array}{l}
\textit{for } n=0,1,\ldots\\
\left\lfloor
\begin{array}{l}
(\gamma_n,\mu_n)\in[\varepsilon,\sigma]\times[\delta,\sigma]\\
\textit{for } i=1,\ldots,m\\
\left\lfloor
\begin{array}{l}
a_{i,n}=(\nabla h_i+\gamma_n\partial\varphi_i)^{-1}\Big(\nabla h_i(x_{i,n})-\gamma_n\sum_{k=1}^{p}L_{ki}^*y_{k,n}^*\Big)\\
a_{i,n}^*=\gamma_n^{-1}\big(\nabla h_i(x_{i,n})-\nabla h_i(a_{i,n})\big)-\sum_{k=1}^{p}L_{ki}^*y_{k,n}^*
\end{array}
\right.\\
\textit{for } k=1,\ldots,p\\
\left\lfloor
\begin{array}{l}
b_{k,n}=(\nabla j_k+\mu_n\partial\psi_k)^{-1}\big(\nabla j_k\big(\sum_{i=1}^{m}L_{ki}x_{i,n}\big)+\mu_n y_{k,n}^*\big)\\
b_{k,n}^*=\mu_n^{-1}\big(\nabla j_k\big(\sum_{i=1}^{m}L_{ki}x_{i,n}\big)-\nabla j_k(b_{k,n})\big)+y_{k,n}^*\\
t_{k,n}=b_{k,n}-\sum_{i=1}^{m}L_{ki}a_{i,n}
\end{array}
\right.\\
\textit{for } i=1,\ldots,m\\
\left\lfloor
\begin{array}{l}
s_{i,n}^*=a_{i,n}^*+\sum_{k=1}^{p}L_{ki}^*b_{k,n}^*
\end{array}
\right.\\
\eta_n=\sum_{i=1}^{m}\big\langle a_{i,n},a_{i,n}^*\big\rangle+\sum_{k=1}^{p}\big\langle b_{k,n},b_{k,n}^*\big\rangle\\
\boldsymbol{H}_n=\Big\{(x,y^*)\in\boldsymbol{\mathcal{X}}\;\Big|\;\sum_{i=1}^{m}\big\langle x_i,s_{i,n}^*\big\rangle+\sum_{k=1}^{p}\langle t_{k,n},y_k^*\rangle\leqslant\eta_n\Big\}\\
\big(x_{n+1/2},y_{n+1/2}^*\big)=P_{\boldsymbol{H}_n}^{\boldsymbol{f}}(x_n,y_n^*)\\
\big(x_{n+1},y_{n+1}^*\big)=Q^f\big((x_0,y_0^*),(x_n,y_n^*),(x_{n+1/2},y_{n+1/2}^*)\big),
\end{array}
\right.
\end{array}
\tag{2.111}
$$

where we use the notation $(\forall n\in\mathbb{N})$ $x_n=(x_{i,n})_{1\leqslant i\leqslant m}$ *and* $y_n^*=(y_{k,n}^*)_{1\leqslant k\leqslant p}$. *Suppose that the following hold :*

(i) *For every* $i\in\{1,\ldots,m\}$, f_i *and* h_i *satisfy Condition 2.23 and* ∇h_i *is uniformly continuous on every bounded subset of* $\operatorname{int}\operatorname{dom}h_i$.

(ii) *For every* $k\in\{1,\ldots,p\}$, g_k^* *and* j_k *satisfy Condition 2.23 and* ∇j_k *is uniformly continuous on every bounded subset of* $\operatorname{int}\operatorname{dom}j_k$.

Then

$$
(\forall i\in\{1,\ldots,m\})\quad x_{i,n}\to\overline{x}_i\quad\textit{and}\quad(\forall k\in\{1,\ldots,p\})\quad y_{k,n}^*\to\overline{y}_k^*. \tag{2.112}
$$

Proof. Denote by $\mathcal{X}$ and $\mathcal{Y}$ the standard vector product spaces $\bigtimes_{i=1}^m\mathcal{X}_i$ and $\bigtimes_{k=1}^p\mathcal{Y}_k$ equipped with the norms $x=(x_i)_{1\leqslant i\leqslant m}\mapsto\sqrt{\sum_{i=1}^m\|x_i\|^2}$ and $y=(y_k)_{1\leqslant k\leqslant p}\mapsto\sqrt{\sum_{k=1}^p\|y_k\|^2}$, respectively. Then $\mathcal{X}^*$ is the vector product space $\bigtimes_{i=1}^m\mathcal{X}_i^*$ equipped with the norm $x^*\mapsto\sqrt{\sum_{i=1}^m\|x_i^*\|^2}$ and $\mathcal{Y}^*$ is the vector product space $\bigtimes_{k=1}^p\mathcal{Y}_k^*$ equipped with the norm $y^*\mapsto\sqrt{\sum_{k=1}^p\|y_k^*\|^2}$. Let us introduce the operators

$$
\begin{cases}
A\colon\mathcal{X}\to 2^{\mathcal{X}^*}\colon x\mapsto\bigtimes_{i=1}^m\partial\varphi_i(x_i)\\
B\colon\mathcal{Y}\to 2^{\mathcal{Y}^*}\colon y\mapsto\bigtimes_{k=1}^p\partial\psi_k(y_k)\\
L\colon\mathcal{X}\to\mathcal{Y}\colon x\mapsto\big(\sum_{i=1}^m L_{ki}x_i\big)_{1\leqslant k\leqslant p}
\end{cases}
\tag{2.113}
$$

and the functions

$$\begin{cases} f\colon \mathcal{X} \to \left]-\infty,+\infty\right] : x \mapsto \sum_{i=1}^m f_i(x_i) \\ h\colon \mathcal{X} \to \left]-\infty,+\infty\right] : x \mapsto \sum_{i=1}^m h_i(x_i) \\ \varphi\colon \mathcal{X} \to \left]-\infty,+\infty\right] : x \mapsto \sum_{i=1}^m \varphi_i(x_i) \\ g\colon \mathcal{Y} \to \left]-\infty,+\infty\right] : y \mapsto \sum_{k=1}^p g_k(y_k) \\ j\colon \mathcal{Y} \to \left]-\infty,+\infty\right] : y \mapsto \sum_{k=1}^p j_k(y_k). \end{cases} \tag{2.114}$$

Then it follows from [39, Theorem 3.1.11] that A and B are maximally monotone. In addition, the adjoint of L is $L^*\colon \mathcal{Y}^* \to \mathcal{X}^*\colon y^* \mapsto (\sum_{k=1}^p L_{ki}^* y_k^*)_{1\leqslant i\leqslant m}$, and, as in (2.70), f and g are Legendre functions. Thus, Problem 2.26 is a special case of Problem 2.10. Furthermore, h and j are Legendre functions,

$$\operatorname{int}\operatorname{dom} f = \bigtimes_{i=1}^m \operatorname{int}\operatorname{dom} f_i \subset \bigtimes_{i=1}^m \operatorname{int}\operatorname{dom} h_i = \operatorname{int}\operatorname{dom} h, \tag{2.115}$$

and

$$L\big(\operatorname{int}\operatorname{dom} f\big) = \bigtimes_{k=1}^p \sum_{i=1}^m L_{ki}(\operatorname{int}\operatorname{dom} f_i) \subset \bigtimes_{k=1}^p \operatorname{int}\operatorname{dom} j_k = \operatorname{int}\operatorname{dom} j. \tag{2.116}$$

Next we observe that, for every $i \in \{1,\ldots,m\}$, since $h_i+\varepsilon\varphi_i$ is supercoercive, $(h_i+\varepsilon\varphi_i)^*$ is bounded above on every bounded subset of $\mathcal{X}_i^*$ [6, Theorem 3.3]. As a result, $(h+\varepsilon\varphi)^*\colon x^* \mapsto \sum_{i=1}^m (h_i+\varepsilon\varphi_i)^*(x_i^*)$ is bounded above on every bounded subset of $\mathcal{X}^*$, and it follows from [6, Theorem 3.3] that $h+\varepsilon\varphi$ is supercoercive. In turn since, as in (2.74), $\varnothing \neq \operatorname{dom} A \cap \operatorname{int}\operatorname{dom} f \subset \operatorname{dom}\varphi \cap \operatorname{int}\operatorname{dom} f$, we derive from [39, Theorem 2.8.3] and Lemma 2.17(iii) that

$$\nabla h + \varepsilon A = \nabla h + \varepsilon\partial\varphi = \partial(h+\varepsilon\varphi) \tag{2.117}$$

is coercive. We show in a similar fashion that $\nabla j + \delta B$ is coercive. Now set, for every $n \in \mathbb{N}$, $a_n = (a_{i,n})_{1\leqslant i\leqslant m}$, $a_n^* = (a_{i,n}^*)_{1\leqslant i\leqslant m}$, $b_n = (b_{k,n})_{1\leqslant k\leqslant p}$, and $b_n^* = (b_{k,n}^*)_{1\leqslant k\leqslant p}$. Then, for every $n \in \mathbb{N}$, we have

$$\begin{aligned} &(\forall i \in \{1,\ldots,m\}) \quad a_{i,n} = (\nabla h_i + \gamma_n\partial\varphi_i)^{-1}\Bigg(\nabla h_i(x_{i,n}) - \gamma_n \sum_{k=1}^p L_{ki}^* y_{k,n}^*\Bigg) \\ &\quad\Leftrightarrow (\forall i \in \{1,\ldots,m\}) \quad \nabla h_i(x_{i,n}) - \gamma_n \sum_{k=1}^p L_{ki}^* y_{k,n}^* \in \nabla h_i(a_{i,n}) + \gamma_n\partial\varphi_i(a_{i,n}) \\ &\quad\Leftrightarrow \nabla h(x_n) - \gamma_n L^* y_n^* \in \nabla h(a_n) + \gamma_n A a_n \\ &\quad\Leftrightarrow a_n = (\nabla h + \gamma_n A)^{-1}\big(\nabla h(x_n) - \gamma_n L^* y_n^*\big). \end{aligned} \tag{2.118}$$

Likewise,

$$(\forall n \in \mathbb{N}) \quad b_n = (\nabla j + \mu_n B)^{-1}\big(\nabla j(Lx_n) + \mu_n y_n^*\big). \tag{2.119}$$

Thus, (2.111) is a special case of (2.65). In addition, it follows from our assumptions and (2.114) that f, g^*, h, and j satisfy Condition 2.23, and that ∇h and ∇j are uniformly continuous on every bounded subset of $\operatorname{int}\operatorname{dom} h$ and $\operatorname{int}\operatorname{dom} j$, respectively. Altogether, the conclusions follow from Theorem 2.24(iv), with $\overline{x} = (\overline{x}_i)_{1\leqslant i\leqslant m}$ and $\overline{y}^* = (\overline{y}^*_k)_{1\leqslant k\leqslant p}$. ☐

Remark 2.28 In Problem 2.26, suppose that, for every $i \in \{1,\ldots,m\}$ and every $k \in \{1,\ldots,p\}$, φ_i and ψ_k are supercoercive Legendre functions satisfying Condition 2.23, that $\nabla\varphi_i$ and $\nabla\psi_k$ are uniformly continuous on bounded subset of $\operatorname{int}\operatorname{dom}\varphi_i$ and $\operatorname{int}\operatorname{dom}\psi_k$, respectively, and that $\sum_{i=1}^m L_{ki}(\operatorname{int}\operatorname{dom}\varphi_i) \subset \operatorname{int}\operatorname{dom}\psi_k$. Then, in Proposition 2.27, we can choose, for every $i \in \{1,\ldots,m\}$ and every $k \in \{1,\ldots,p\}$, $h_i = \varphi_i$ and $j_k = \psi_k$, and in (2.111), we obtain

$$a_{i,n} = \nabla h_i^*\left(\frac{\nabla h_i(x_{i,n}) - \gamma_n\sum_{k=1}^p L^*_{ki}y^*_{k,n}}{1+\gamma_n}\right) \tag{2.120}$$

and

$$b_{k,n} = \nabla j_k^*\left(\frac{\nabla j_k\left(\sum_{i=1}^m L_{ki}x_{i,n}\right) + \mu_n y^*_{k,n}}{1+\mu_n}\right). \tag{2.121}$$

For example, suppose that, for every $i \in \{1,\ldots,m\}$ and every $k \in \{1,\ldots,p\}$, $\mathcal{X}_i = \mathbb{R}$, $\mathcal{Y}_k = \mathbb{R}$, and $\varphi_i = h_i$ is the Hellinger-like function, i.e.,

$$\varphi_i\colon \mathbb{R} \to \left]-\infty,+\infty\right] : x_i \mapsto \begin{cases} -\sqrt{1-x_i^2}, & \text{if } x_i \in [-1,1]; \\ +\infty, & \text{otherwise.} \end{cases} \tag{2.122}$$

Then (2.120) becomes

$$a_{i,n} = \frac{x_{i,n} - \gamma_n\left(\sum_{k=1}^p L^*_{ki}y^*_{k,n}\right)\sqrt{1-x_{i,n}^2}}{\sqrt{(1+\gamma_n)^2(1-x_{i,n}^2) + \left(x_{i,n} - \gamma_n\left(\sum_{k=1}^p L^*_{ki}y^*_{k,n}\right)\sqrt{1-x_{i,n}^2}\right)^2}}. \tag{2.123}$$

Furthermore, as shown in the next section, in finite-dimensional spaces, we can remove Condition 2.23 and the assumption on the uniform continuity of $(\nabla\varphi_i)_{1\leqslant i\leqslant m}$ and $(\nabla\psi_k)_{1\leqslant k\leqslant p}$.

2.2.4 Finite-dimensional setting

In finite-dimensional spaces, the convergence of algorithm (2.65) can be obtained under more general assumptions. To establish the corresponding results, the following technical facts will be needed.

Lemma 2.29 *Let $\mathcal{X}$ be a finite-dimensional real Banach space and let $f \in \Gamma_0(\mathcal{X})$ be a Legendre function. Then the following hold :*

(i) *f and ∇f are continuous on* $\operatorname{int}\operatorname{dom} f$ [9, Corollaries 8.30(iii), 17.34, and 17.35].

(ii) $\nabla f\colon \operatorname{int}\operatorname{dom} f \to \operatorname{int}\operatorname{dom} f^*$ *is bijective with inverse* $\nabla f^*\colon \operatorname{int}\operatorname{dom} f^* \to \operatorname{int}\operatorname{dom} f$ [6, Theorem 5.10].

(iii) *Let $x \in \operatorname{int}\operatorname{dom} f$, let $y \in \overline{\operatorname{dom}}\, f$, and let $(y_n)_{n\in\mathbb{N}} \in (\operatorname{int}\operatorname{dom} f)^{\mathbb{N}}$. Suppose that $y_n \to y$ and that $(D^f(x, y_n))_{n\in\mathbb{N}}$ is bounded. Then $y \in \operatorname{int}\operatorname{dom} f$ and $D^f(y, y_n) \to 0$* [5, Theorem 3.8(ii)].

(iv) *Let $x \in \operatorname{int}\operatorname{dom} f$, let $y \in \operatorname{int}\operatorname{dom} f$, let $(x_n)_{n\in\mathbb{N}} \in (\operatorname{dom} f)^{\mathbb{N}}$, and let $(y_n)_{n\in\mathbb{N}} \in (\operatorname{int}\operatorname{dom} f)^{\mathbb{N}}$. Suppose that $D^f(x_n, y_n) \to 0$. Then $x = y$* [5, Theorem 3.9(iii)].

(v) *Let $y \in \operatorname{int}\operatorname{dom} f$. Then $D^f(\cdot, y)$ is coercive* [6, Lemma 7.3(v)].

(vi) *Let $\{x, y\} \subset \operatorname{int}\operatorname{dom} f$. Then $D^f(x, y) = D^{f^*}(\nabla f(y), \nabla f(x))$* [6, Lemma 7.3(vii)].

Proposition 2.30 *In Problem 2.10, suppose that $\mathcal{X}$ and $\mathcal{Y}$ are finite-dimensional. Let $h \in \Gamma_0(\mathcal{X})$ and $j \in \Gamma_0(\mathcal{Y})$ be Legendre functions such that $\operatorname{int}\operatorname{dom} f \subset \operatorname{int}\operatorname{dom} h$, $L(\operatorname{int}\operatorname{dom} f) \subset \operatorname{int}\operatorname{dom} j$, and there exist ε and δ in $]0, +\infty[$ such that $\nabla h + \varepsilon A$ and $\nabla j + \delta B$ are coercive. Let $\sigma \in [\max\{\varepsilon, \delta\}, +\infty[$ and execute algorithm (2.65). Then $(x_n, y_n^*) \to (\overline{x}, \overline{y}^*)$.*

Proof. Set $\boldsymbol{C} = \boldsymbol{Z} \cap \overline{\operatorname{dom}}\, \boldsymbol{f}$. We first observe that, as in (2.91),

$$(x_n)_{n\in\mathbb{N}} \in (\operatorname{int}\operatorname{dom} f)^{\mathbb{N}} \quad \text{and} \quad (y_n^*)_{n\in\mathbb{N}} \in (\operatorname{int}\operatorname{dom} g^*)^{\mathbb{N}} \quad \text{are bounded.} \tag{2.124}$$

In addition, we deduce from (2.77), (2.81), and Proposition 2.14(i) that $\overline{\boldsymbol{x}} = (\overline{x}, \overline{y}^*) \in \boldsymbol{C} \subset \bigcap_{n\in\mathbb{N}} H^{\boldsymbol{f}}(\boldsymbol{x}_0, \boldsymbol{x}_n)$, and hence from (2.39) that

$$(\forall n \in \mathbb{N}) \quad D^f(\overline{x}, x_n) + D^{g^*}(\overline{y}^*, y_n^*) = D^{\boldsymbol{f}}\big(\overline{\boldsymbol{x}}, \boldsymbol{x}_n\big) \leqslant D^{\boldsymbol{f}}\big(\overline{\boldsymbol{x}}, \boldsymbol{x}_0\big). \tag{2.125}$$

By virtue of Proposition 2.14(vii) and (2.124), it suffices to show that every cluster point of $(x_n, y_n^*)_{n\in\mathbb{N}}$ belongs to $\boldsymbol{Z}$. To this end, take $x \in \mathcal{X}$, $y^* \in \mathcal{Y}$, and a strictly increasing sequence $(k_n)_{n\in\mathbb{N}}$ in $\mathbb{N}$ such that $x_{k_n} \to x$ and $y_{k_n}^* \to y^*$. Then $Lx_{k_n} \to Lx$, $x \in \overline{\operatorname{dom}}\, f$, and $y^* \in \overline{\operatorname{dom}}\, g^*$. Since $\overline{x} \in \operatorname{int}\operatorname{dom} f$ and since (2.125) implies that $(D^f(\overline{x}, x_{k_n}))_{n\in\mathbb{N}}$ is bounded, it follows from Lemma 2.29(iii) that $x \in \operatorname{int}\operatorname{dom} f$. Analogously, $y^* \in \operatorname{int}\operatorname{dom} g^*$. In turn, Lemma 2.29(i) asserts that

$$\nabla f\big(x_{k_n}\big) \to \nabla f(x) \quad \text{and} \quad \nabla g^*\big(y_{k_n}^*\big) \to \nabla g^*(y^*). \tag{2.126}$$

Furthermore, since $\operatorname{int}\operatorname{dom} f \subset \operatorname{int}\operatorname{dom} h$ and $L(\operatorname{int}\operatorname{dom} f) \subset \operatorname{int}\operatorname{dom} j$, we obtain $x \in \operatorname{int}\operatorname{dom} h$ and $Lx \in \operatorname{int}\operatorname{dom} j$. Thus, there exists $\rho \in\,]0, +\infty[$ such that $B(x; \rho) \subset \operatorname{int}\operatorname{dom} h$ and $B(Lx; \rho) \subset \operatorname{int}\operatorname{dom} j$. We therefore assume without loss of generality that

$$(x_{k_n})_{n\in\mathbb{N}} \in B(x; \rho)^{\mathbb{N}} \quad \text{and} \quad (Lx_{k_n})_{n\in\mathbb{N}} \in B(Lx; \rho)^{\mathbb{N}}. \tag{2.127}$$

In view of Lemma 2.29(i), $h(B(x;\rho))$ and $\nabla h(B(x;\rho))$ are therefore compact, which implies that $(h(x_{k_n}))_{n\in\mathbb{N}}$ and $(\nabla h(x_{k_n}))_{n\in\mathbb{N}}$ are bounded. Hence $(D^h(\overline{x},x_{k_n}))_{n\in\mathbb{N}}$ is bounded and, moreover, it follows from (2.65), (2.124), Lemma 2.19, and (2.76) that $(a_{k_n})_{n\in\mathbb{N}}$ is a bounded sequence in $\operatorname{int}\operatorname{dom} h$. We show likewise that $(D^j(L\overline{x},Lx_{k_n}))_{n\in\mathbb{N}}$ and $(b_{k_n})_{n\in\mathbb{N}}$ are bounded. Next, since the convexity of h yields

$$\begin{aligned}(\forall n\in\mathbb{N})\quad D^h(\overline{x},a_{k_n}) &= h(\overline{x})-h(a_{k_n})-\langle\overline{x}-a_{k_n},\nabla h(a_{k_n})\rangle\\ &= h(\overline{x})-h(x_{k_n})-\langle\overline{x}-x_{k_n},\nabla h(x_{k_n})\rangle\\ &\quad+\big\langle\overline{x}-a_{k_n},\nabla h(x_{k_n})-\nabla h(a_{k_n})\big\rangle\\ &\quad-\big(h(a_{k_n})-h(x_{k_n})-\langle a_{k_n}-x_{k_n},\nabla h(x_{k_n})\rangle\big)\\ &\leqslant D^h(\overline{x},x_{k_n})+\big\langle\overline{x}-a_{k_n},\nabla h(x_{k_n})-\nabla h(a_{k_n})\big\rangle,\end{aligned}\tag{2.128}$$

we derive from (2.65) that

$$\begin{aligned}(\forall n\in\mathbb{N})\quad \sigma^{-1}D^h(\overline{x},a_{k_n}) &\leqslant \gamma_{k_n}^{-1}D^h(\overline{x},a_{k_n})\\ &\leqslant \varepsilon^{-1}D^h(\overline{x},x_{k_n})+\big\langle\overline{x}-a_{k_n},\gamma_{k_n}^{-1}\big(\nabla h(x_{k_n})-\nabla h(a_{k_n})\big)\big\rangle\\ &= \varepsilon^{-1}D^h(\overline{x},x_{k_n})+\big\langle\overline{x}-a_{k_n},a^*_{k_n}+L^*y^*_{k_n}\big\rangle\\ &= \varepsilon^{-1}D^h(\overline{x},x_{k_n})+\big\langle\overline{x}-a_{k_n},a^*_{k_n}+L^*\overline{y}^*\big\rangle\\ &\quad+\big\langle L\overline{x}-La_{k_n},y^*_{k_n}-\overline{y}^*\big\rangle.\end{aligned}\tag{2.129}$$

Similarly,

$$\begin{aligned}(\forall n\in\mathbb{N})\quad \sigma^{-1}D^j(L\overline{x},b_{k_n}) &\leqslant \delta^{-1}D^j(L\overline{x},Lx_{k_n})+\big\langle L\overline{x}-b_{k_n},b^*_{k_n}-y^*_{k_n}\big\rangle\\ &\leqslant \delta^{-1}D^j(L\overline{x},Lx_{k_n})+\big\langle L\overline{x}-b_{k_n},b^*_{k_n}-\overline{y}^*\big\rangle\\ &\quad+\big\langle L\overline{x}-b_{k_n},\overline{y}^*-y^*_{k_n}\big\rangle.\end{aligned}\tag{2.130}$$

Since (2.79) entails that

$$\overline{\boldsymbol{x}}\in\boldsymbol{C}\subset\bigcap_{n\in\mathbb{N}}\boldsymbol{H}_{k_n}=\bigcap_{n\in\mathbb{N}}\big\{(x,y^*)\in\boldsymbol{\mathcal{X}}\mid\big\langle x-a_{k_n},a^*_{k_n}+L^*y^*\big\rangle+\big\langle Lx-b_{k_n},b^*_{k_n}-y^*\big\rangle\leqslant 0\big\},\tag{2.131}$$

we deduce from (2.129) and (2.130) that

$$\begin{aligned}&(\forall n\in\mathbb{N})\quad \sigma^{-1}\big(D^h(\overline{x},a_{k_n})+D^j(L\overline{x},b_{k_n})\big)\\ &\quad\leqslant \varepsilon^{-1}D^h(\overline{x},x_{k_n})+\delta^{-1}D^j(L\overline{x},Lx_{k_n})+\big\langle L\overline{x}-La_{k_n},y^*_{k_n}-\overline{y}^*\big\rangle+\big\langle L\overline{x}-b_{k_n},\overline{y}^*-y^*_{k_n}\big\rangle\\ &\quad= \varepsilon^{-1}D^h(\overline{x},x_{k_n})+\delta^{-1}D^j(L\overline{x},Lx_{k_n})+\big\langle b_{k_n}-La_{k_n},y^*_{k_n}-\overline{y}^*\big\rangle\\ &\quad\leqslant \varepsilon^{-1}D^h(\overline{x},x_{k_n})+\delta^{-1}D^j(L\overline{x},Lx_{k_n})+(\|b_{k_n}\|+\|L\|\,\|a_{k_n}\|)(\|y^*_{k_n}\|+\|\overline{y}^*\|).\end{aligned}\tag{2.132}$$

Hence, the boundedness of $(a_{k_n})_{n\in\mathbb{N}}$, $(b_{k_n})_{n\in\mathbb{N}}$, $(y^*_{k_n})_{n\in\mathbb{N}}$, $(D^h(\overline{x},x_{k_n}))_{n\in\mathbb{N}}$, and $(D^j(L\overline{x},Lx_{k_n}))_{n\in\mathbb{N}}$ implies that of $(D^h(\overline{x},a_{k_n}))_{n\in\mathbb{N}}$ and $(D^j(L\overline{x},b_{k_n}))_{n\in\mathbb{N}}$. In turn, by Lemma 2.29(vi),

$$\Big(D^{h^*}(\nabla h(a_{k_n}),\nabla h(\overline{x}))\Big)_{n\in\mathbb{N}}\quad\text{and}\quad\Big(D^{j^*}(\nabla j(b_{k_n}),\nabla j(L\overline{x}))\Big)_{n\in\mathbb{N}}\quad\text{are bounded}\tag{2.133}$$

and, since Lemma 2.29(v) asserts that $D^{h^*}(\cdot, \nabla h(\overline{x}))$ and $D^{j^*}(\cdot, \nabla j(L\overline{x}))$ are coercive, it follows from (2.133) that $(\nabla h(a_{k_n}))_{n\in\mathbb{N}}$ and $(\nabla j(b_{k_n}))_{n\in\mathbb{N}}$ are bounded. Thus, since $(y^*_{k_n})_{n\in\mathbb{N}}$, $(\nabla h(x_{k_n}))_{n\in\mathbb{N}}$ and $(\nabla j(Lx_{k_n}))_{n\in\mathbb{N}}$ are bounded, we infer from (2.65) that

$$(a^*_{k_n})_{n\in\mathbb{N}} \quad \text{and} \quad (b^*_{k_n})_{n\in\mathbb{N}} \quad \text{are bounded.} \tag{2.134}$$

On the other hand, since (2.79) yields $\boldsymbol{\overline{x}} \in \boldsymbol{C} \subset \bigcap_{n\in\mathbb{N}} H^{\boldsymbol{f}}(\boldsymbol{x}_{k_n}, \boldsymbol{x}_{k_n+1/2})$, (2.39) and (2.125) imply that

$$(\forall n \in \mathbb{N}) \quad D^f(\overline{x}, x_{k_n+1/2}) + D^{g^*}(\overline{y}^*, y^*_{k_n+1/2}) = D^{\boldsymbol{f}}\big(\boldsymbol{\overline{x}}, \boldsymbol{x}_{k_n+1/2}\big) \leqslant D^{\boldsymbol{f}}\big(\boldsymbol{\overline{x}}, \boldsymbol{x}_{k_n}\big) \leqslant D^{\boldsymbol{f}}\big(\boldsymbol{\overline{x}}, \boldsymbol{x}_0\big). \tag{2.135}$$

Thus, Lemma 2.29(vi) yields

$$\begin{aligned}(\forall n \in \mathbb{N}) \quad & D^{f^*}\big(\nabla f(x_{k_n+1/2}), \nabla f(\overline{x})\big) + D^g\big(\nabla g^*(y^*_{k_n+1/2}), \nabla g^*(\overline{y}^*)\big) \\ &= D^f(\overline{x}, x_{k_n+1/2}) + D^{g^*}(\overline{y}^*, y^*_{k_n+1/2}) \\ &\leqslant D^{\boldsymbol{f}}\big(\boldsymbol{\overline{x}}, \boldsymbol{x}_0\big)\end{aligned} \tag{2.136}$$

and, since $D^{f^*}(\cdot, \nabla f(\overline{x}))$ and $D^g(\cdot, \nabla g^*(\overline{y}^*))$ are coercive by Lemma 2.29(v), it follows that

$$(\nabla f(x_{k_n+1/2}))_{n\in\mathbb{N}} \quad \text{and} \quad (\nabla g^*(y^*_{k_n+1/2}))_{n\in\mathbb{N}} \quad \text{are bounded.} \tag{2.137}$$

However, as in (2.95), $D^f(x_{k_n+1/2}, x_{k_n}) \to 0$ and $D^{g^*}(y^*_{k_n+1/2}, y^*_{k_n}) \to 0$, and it therefore follows from Lemma 2.29(vi) that

$$D^{f^*}\big(\nabla f(x_{k_n}), \nabla f(x_{k_n+1/2})\big) \to 0 \quad \text{and} \quad D^g\big(\nabla g^*(y^*_{k_n}), \nabla g^*(y^*_{k_n+1/2})\big) \to 0. \tag{2.138}$$

In view of Lemma 2.29(iv), we infer from (2.126), (2.137), and (2.138) that there exists a strictly increasing sequence $(p_{k_n})_{n\in\mathbb{N}}$ in $\mathbb{N}$ such that

$$\nabla f\big(x_{p_{k_n}+1/2}\big) \to \nabla f(x) \quad \text{and} \quad \nabla g^*\big(y^*_{p_{k_n}+1/2}\big) \to \nabla g^*(y^*). \tag{2.139}$$

Since, by Lemma 2.29(i)–(ii), $(\nabla f)^{-1} = \nabla f^*$ is continuous on $\operatorname{int}\operatorname{dom} f^*$ and $(\nabla g^*)^{-1} = \nabla g$ is continuous on $\operatorname{int}\operatorname{dom} g$, we obtain $x_{p_{k_n}+1/2} \to x$ and $y^*_{p_{k_n}+1/2} \to y^*$. Thus,

$$x_{p_{k_n}+1/2} - x_{p_{k_n}} \to 0 \quad \text{and} \quad y^*_{p_{k_n}+1/2} - y^*_{p_{k_n}} \to 0. \tag{2.140}$$

On the other hand, as in (2.94),

$$\begin{aligned}(\forall n \in \mathbb{N}) \quad & \|x_{p_{k_n}} - x_{p_{k_n}+1/2}\|\,\|a^*_{p_{k_n}} + L^*b^*_{p_{k_n}}\| + \|b_{p_{k_n}} - La_{p_{k_n}}\|\,\|y^*_{p_{k_n}} - y^*_{p_{k_n}+1/2}\| \\ &\geqslant \sigma^{-1}\big(D^h(x_{p_{k_n}}, a_{p_{k_n}}) + D^j(Lx_{p_{k_n}}, b_{p_{k_n}})\big),\end{aligned} \tag{2.141}$$

and hence, since $(a_{p_{k_n}})_{n\in\mathbb{N}}$ and $(b_{p_{k_n}})_{n\in\mathbb{N}}$ are bounded, we deduce from (2.134) and (2.140) that

$$\begin{cases} D^h\big(x_{p_{k_n}}, a_{p_{k_n}}\big) \to 0 \\ D^j\big(Lx_{p_{k_n}}, b_{p_{k_n}}\big) \to 0 \\ x_{p_{k_n}} \to x \\ Lx_{p_{k_n}} \to Lx \\ (a_{p_{k_n}})_{n\in\mathbb{N}} \text{ has a cluster point} \\ (b_{p_{k_n}})_{n\in\mathbb{N}} \text{ has a cluster point.} \end{cases} \tag{2.142}$$

Consequently, by dropping to a subsequence if necessary and invoking Lemma 2.29(iv), we get

$$a_{p_{k_n}} \to x \quad \text{and} \quad b_{p_{k_n}} \to Lx. \tag{2.143}$$

Hence, using the fact that $\nabla h(x_{p_{k_n}}) \to \nabla h(x)$ and $\nabla j(Lx_{p_{k_n}}) \to \nabla j(Lx)$, we derive that $\nabla h(x_{p_{k_n}}) - \nabla h(a_{p_{k_n}}) \to 0$ and $\nabla j(Lx_{p_{k_n}}) - \nabla j(b_{p_{k_n}}) \to 0$, which, in view of (2.65), yields

$$a^*_{p_{k_n}} + L^* y^*_{p_{k_n}} \to 0 \quad \text{and} \quad b^*_{p_{k_n}} - y^*_{p_{k_n}} \to 0. \tag{2.144}$$

Thus, since $y^*_{p_{k_n}} \to y^*$, it follows that $a^*_{p_{k_n}} \to -L^*y^*$ and $b^*_{p_{k_n}} \to y^*$. In summary,

$$\operatorname{gra} A \ni (a_{p_{k_n}}, a^*_{p_{k_n}}) \to (x, -L^*y^*) \quad \text{and} \quad \operatorname{gra} B \ni (b_{p_{k_n}}, b^*_{p_{k_n}}) \to (Lx, y^*). \tag{2.145}$$

Since $\operatorname{gra} A$ and $\operatorname{gra} B$ are closed [9, Proposition 20.33(iii)], we conclude that $(x, -L^*y^*) \in \operatorname{gra} A$ and $(Lx, y^*) \in \operatorname{gra} B$, and therefore that $(x, y^*) \in \boldsymbol{Z}$. □

Let us note that, even in Euclidean spaces, it may be easier to evaluate $(\nabla h + \gamma\partial\varphi)^{-1}$ than the usual proximity operator $\operatorname{prox}_{\gamma\varphi} = (\operatorname{Id} + \gamma\partial\varphi)^{-1}$ introduced by Moreau [30], which is based on $h = \|\cdot\|^2/2$. We provide illustrations of such instances in the standard Euclidean space $\mathbb{R}^m$.

Example 2.31 Let $\gamma \in \left]0, +\infty\right[$, let $\phi \in \Gamma_0(\mathbb{R})$ be such that $\operatorname{dom} \phi \cap \left]0, +\infty\right[\neq \varnothing$, and let ϑ be the Boltzmann-Shannon entropy, i.e.,

$$\vartheta\colon \xi \mapsto \begin{cases} \xi \ln \xi - \xi, & \text{if } \ \xi \in \left]0, +\infty\right[; \\ 0, & \text{if } \ \xi = 0; \\ +\infty, & \text{otherwise.} \end{cases} \tag{2.146}$$

Set $\varphi\colon (\xi_i)_{1\leqslant i\leqslant m} \mapsto \sum_{i=1}^m \phi(\xi_i)$ and $h\colon (\xi_i)_{1\leqslant i\leqslant m} \mapsto \sum_{i=1}^m \vartheta(\xi_i)$. Note that h is a supercoercive Legendre function [5, Sections 5 and 6] and hence Proposition 2.21(iv) asserts that $\nabla h + \gamma\partial\varphi$ is coercive and $\operatorname{dom}(\nabla h + \gamma\partial\varphi)^{-1} = \mathbb{R}^m$. Now let $(\xi_i)_{1\leqslant i\leqslant m} \in \mathbb{R}^m$, set $(\eta_i)_{1\leqslant i\leqslant m} = (\nabla h + \gamma\partial\varphi)^{-1}(\xi_i)_{1\leqslant i\leqslant m}$, let W be the Lambert function [20, 25], i.e., the inverse of $\xi \mapsto \xi e^{\xi}$ on $\left[0, +\infty\right[$, and let $i \in \{1, \ldots, m\}$. Then η_i can be computed as follows.

(i) Let $\omega \in \mathbb{R}$ and suppose that

$$\phi\colon \xi \mapsto \begin{cases} \xi \ln \xi - \omega\xi, & \text{if } \xi \in \,]0,+\infty[\,; \\ 0, & \text{if } \xi = 0; \\ +\infty, & \text{otherwise.} \end{cases} \tag{2.147}$$

Then $\eta_i = e^{(\xi_i+\omega-1)/(\gamma+1)}$.

(ii) Let $p \in [1,+\infty[$ and suppose that either $\phi = |\cdot|^p/p$ or

$$\phi\colon \xi \mapsto \begin{cases} \xi^p/p, & \text{if } \xi \in [0,+\infty[\,; \\ +\infty, & \text{otherwise.} \end{cases} \tag{2.148}$$

Then

$$\eta_i = \begin{cases} \left(\dfrac{W(\gamma(p-1)e^{(p-1)\xi_i})}{\gamma(p-1)}\right)^{\frac{1}{p-1}}, & \text{if } p \in \,]1,+\infty[\,; \\ e^{\xi_i-\gamma}, & \text{if } p = 1. \end{cases} \tag{2.149}$$

(iii) Let $p \in [1,+\infty[$ and suppose that

$$\phi\colon \xi \mapsto \begin{cases} \xi^{-p}/p, & \text{if } \xi \in \,]0,+\infty[\,; \\ +\infty, & \text{otherwise.} \end{cases} \tag{2.150}$$

Then

$$\eta_i = \left(\frac{W(\gamma(p+1)e^{-(p+1)\xi_i})}{\gamma(p+1)}\right)^{\frac{-1}{p+1}}. \tag{2.151}$$

(iv) Let $p \in \,]0,1[$ and suppose that

$$\phi\colon \xi \mapsto \begin{cases} -\xi^p/p, & \text{if } \xi \in [0,+\infty[\,; \\ +\infty, & \text{otherwise.} \end{cases} \tag{2.152}$$

Then

$$\eta_i = \left(\frac{W(\gamma(1-p)e^{(p-1)\xi_i})}{\gamma(1-p)}\right)^{\frac{1}{p-1}}. \tag{2.153}$$

Example 2.32 Let $\phi \in \Gamma_0(\mathbb{R})$ be such that $\operatorname{dom}\phi \cap]0,1[\neq \varnothing$ and let ϑ be the Fermi-Dirac entropy, i.e.,

$$\vartheta\colon \xi \mapsto \begin{cases} \xi \ln \xi - (1-\xi)\ln(1-\xi), & \text{if } \xi \in \,]0,1[\,; \\ 0 & \text{if } \xi \in \{0,1\}; \\ +\infty, & \text{otherwise.} \end{cases} \tag{2.154}$$

Set $\varphi\colon (\xi_i)_{1\leqslant i\leqslant m} \mapsto \sum_{i=1}^m \phi(\xi_i)$ and $h\colon (\xi_i)_{1\leqslant i\leqslant m} \mapsto \sum_{i=1}^m \vartheta(\xi_i)$. Note that h is a Legendre function [5, Sections 5 and 6] and that $\operatorname{int}\operatorname{dom} h = \,]0,1[^m$ is bounded. Therefore, Proposition 2.21(i) asserts that $\nabla h + \partial\varphi$ is coercive and that $\operatorname{dom}(\nabla h + \partial\varphi)^{-1} = \mathbb{R}^m$. Now let $(\xi_i)_{1\leqslant i\leqslant m} \in \mathbb{R}^m$, set $(\eta_i)_{1\leqslant i\leqslant m} = (\nabla h + \partial\varphi)^{-1}(\xi_i)_{1\leqslant i\leqslant m}$, and let $i \in \{1,\ldots,m\}$. Then η_i can be computed as follows.

(i) Let $\omega \in \mathbb{R}$ and suppose that

$$\phi\colon \xi \mapsto \begin{cases} \xi\ln\xi - \omega\xi, & \text{if } \ \xi \in \,]0,+\infty[\,;\\ 0, & \text{if } \ \xi = 0;\\ +\infty, & \text{otherwise.}\end{cases} \tag{2.155}$$

Then $\eta_i = -e^{\xi_i+\omega-1}/2 + \sqrt{e^{2(\xi_i+\omega-1)}/4 + e^{\xi_i+\omega-1}}$.

(ii) Suppose that

$$\phi\colon \xi \mapsto \begin{cases} (1-\xi)\ln(1-\xi) + \xi, & \text{if } \ \xi \in \,]-\infty,1[\,;\\ 1, & \text{if } \ \xi = 1;\\ +\infty, & \text{otherwise.}\end{cases} \tag{2.156}$$

Then $\eta_i = 1 + e^{-\xi_i}/2 - \sqrt{e^{-\xi_i} + e^{-2\xi_i}/4}$.

2.3 Bibliographie

[1] M. A. Alghamdi, A. Alotaibi, P. L. Combettes, and N. Shahzad, A primal-dual method of partial inverses for composite inclusions, *Optim. Lett.*, vol. 8, pp. 2271–2284, 2014.

[2] A. Alotaibi, P. L. Combettes, and N. Shahzad, Solving coupled composite monotone inclusions by successive Fejér approximations of their Kuhn-Tucker set, *SIAM J. Optim.*, vol. 24, pp. 2076–2095, 2014.

[3] A. Alotaibi, P. L. Combettes, and N. Shahzad, Best approximation from the Kuhn-Tucker set of composite monotone inclusions, *Numer. Funct. Anal. Optim.*, to appear.

[4] H. Attouch, L. M. Briceño-Arias, and P. L. Combettes, A strongly convergent primal-dual method for nonoverlapping domain decomposition, *Numer. Math.*, to appear.

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

[6] H. H. Bauschke, J. M. Borwein, and P. L. Combettes, Essential smoothness, essential strict convexity, and Legendre functions in Banach spaces, *Commun. Contemp. Math.*, vol. 3, pp. 615–647, 2001.

[7] H. H. Bauschke, J. M. Borwein, and P. L. Combettes, Bregman monotone optimization algorithms, *SIAM J. Control Optim.*, vol. 42, pp. 596–636, 2003.

[8] H. H. Bauschke and P. L. Combettes, Construction of best Bregman approximations in reflexive Banach spaces, *Proc. Amer. Math. Soc.*, vol. 131, pp. 3757–3766, 2003.

[9] H. H. Bauschke and P. L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces.* Springer, New York, 2011.

[10] J. M. Borwein, Fifty years of maximal monotonicity, *Optim. Lett.,* vol. 4, pp. 473–490, 2010.

[11] R. I. Boţ, E. R. Csetnek, and A. Heinrich, A primal-dual splitting algorithm for finding zeros of sums of maximal monotone operators, *SIAM J. Optim.,* vol. 23, pp. 2011–2036, 2013.

[12] R. I. Boţ and C. Hendrich, A Douglas-Rachford type primal-dual method for solving inclusions with mixtures of composite and parallel-sum type monotone operators, *SIAM J. Optim.,* vol. 23, pp. 2541–2565, 2013.

[13] L. M. Bregman, The relaxation method of finding the common point of convex sets and its application to the solution of problems in convex programming, *USSR Comput. Math. Math. Phys.*, vol 7, pp. 200–217, 1967.

[14] L. M. Briceño-Arias and P. L. Combettes, A monotone+skew splitting model for composite monotone inclusions in duality, *SIAM J. Optim.*, vol. 21, pp. 1230–1250, 2011.

[15] F. E. Browder, Multi-valued monotone nonlinear mappings and duality mappings in Banach spaces, *Trans. Amer. Math. Soc.,* vol. 118, pp. 338–351, 1965.

[16] I. Cioranescu, *Geometry of Banach spaces, Duality Mappings and Nonlinear Problems.* Kluwer, Dordrecht, 1990.

[17] P. L. Combettes, Strong convergence of block-iterative outer approximation methods for convex optimization, *SIAM J. Control Optim.*, vol. 38, pp. 538–565, 2000.

[18] P. L. Combettes, Systems of structured monotone inclusions : Duality, algorithms, and applications, *SIAM J. Optim.,* vol. 23, pp. 2420–2447, 2013.

[19] P. L. Combettes and B. C. Vũ, Variable metric forward-backward splitting with applications to monotone inclusions in duality, *Optimization,* vol. 63, pp. 1289–1318, 2014.

[20] R. M. Corless, G. H. Gonnet, D. E. G. Hare, D. J. Jeffrey, and D. E. Knuth, On the Lambert W function, *Adv. Comput. Math.,* vol. 5, pp. 329–359, 1996.

[21] J. Eckstein, Nonlinear proximal point algorithms using Bregman functions, with applications to convex programming, *Math. Oper. Res.*, vol. 18, pp. 202–226, 1993.

[22] Y. Haugazeau, *Sur les Inéquations Variationnelles et la Minimisation de Fonctionnelles Convexes.* Thèse, Université de Paris, Paris, France, 1968.

[23] R. I. Kačurovskiĭ, On monotone operators and convex functionals, *Uspekhi Mat. Nauk*, vol. 15, pp. 213–215, 1960.

[24] G. Kassay, The proximal points algorithm for reflexive Banach spaces, *Studia Univ. Babeş-Bolyai Math.*, vol. 30, pp. 9–17, 1985.

[25] J. H. Lamberti, Observationes variae in mathesin puram, *Acta Helvetica, Physico-Mathematico-Anatomico-Botanico-Medica,* vol. 3, pp. 128–168, 1758.

[26] P. L. Lions and B. Mercier, Splitting algorithms for the sum of two nonlinear operators, *SIAM J. Numer. Anal.*, vol. 16, pp. 964–979, 1979.

[27] B. Mercier, *Topics in Finite Element Solution of Elliptic Problems* (Lectures on Mathematics, no. 63). Tata Institute of Fundamental Research, Bombay, 1979.

[28] G. J. Minty, On the maximal domain of a “monotone” function, *Michigan Math. J.*, vol. 8, pp. 135–137, 1961.

[29] G. J. Minty, Monotone (nonlinear) operators in Hilbert space, *Duke Math. J.*, vol. 29, pp. 341–346, 1962.

[30] J. J. Moreau, Fonctions convexes duales et points proximaux dans un espace hilbertien, *C. R. Acad. Sci. Paris Sér. A Math.,* vol. 255, pp. 2897–2899, 1962.

[31] J. J. Moreau, Fonctionnelles sous-différentiables, *C. R. Acad. Sci. Paris Sér. A Math.,* vol. 257, pp. 4117–4119, 1963.

[32] T. Pennanen, Dualization of generalized equations of maximal monotone type, *SIAM J. Optim.,* vol. 10, pp. 809–835, 2000.

[33] R. T. Rockafellar, Monotone operators and the proximal point algorithm, *SIAM J. Control Optim.,* vol. 14, pp. 877–898, 1976.

[34] S. Simons, *From Hahn-Banach to Monotonicity,* Lecture Notes in Math. 1693, Springer-Verlag, New York, 2008.

[35] M. Teboulle, Entropic proximal mappings with applications to nonlinear programming, *Math. Oper. Res.*, vol. 17, pp. 670–690, 1992.

[36] P. Tseng, A modified forward-backward splitting method for maximal monotone mappings, *SIAM J. Control Optim.*, vol. 38, pp. 431–446, 2000.

[37] B. C. Vũ, A splitting algorithm for dual monotone inclusions involving cocoercive operators, *Adv. Comput. Math.*, vol. 38, pp. 667–681, 2013.

[38] C. Zălinescu, On uniformly convex functions, *J. Math. Anal. Appl.,* vol. 95, pp. 344–374, 1983.

[39] C. Zălinescu, *Convex Analysis in General Vector Spaces.* World Scientific Publishing, River Edge, NJ, 2002.

[40] E. H. Zarantonello, *Solving functional equations by contractive averaging,* Tech. Report 160, Math. Res. Center U.S. Army, Madison, University of Wisconsin, June 1960.

[41] E. Zeidler, *Nonlinear Functional Analysis and Its Applications,* vols. I–V. Springer-Verlag, New York, 1985–1993.

Chapitre 3

Algorithme d'Haugazeau-Bregman avec erreurs

Dans [4] (cf. Chapitre 2), les auteurs introduisent un algorithme de type Haugazeau-Bregman pour trouver la meilleure approximation au sens de Bregman d'un point donné dans un sous-ensemble convexe fermé d'un espace de Banach réel réflexif. Nous étudions dans ce chapitre le comportement de cette méthode en présence d'erreurs.

3.1 Notations et introduction

Nous rappelons les notations et les définitions qui seront utilisées dans ce chapitre.

Notation 3.1 La norme d'un espace de Banach est noté $\|\cdot\|$ et le crochet de dualité entre un espace de Banach et son dual topologique est noté $\langle\cdot,\cdot\rangle$. Les symboles $\rightharpoonup$ et $\to$ représentent la convergence faible et forte, respectivement. L'ensemble des points faiblement adhérents à une suite $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ est noté $\mathfrak{W}(\boldsymbol{x}_n)_{n\in\mathbb{N}}$.

Définition 3.2 [1, 2, 4] Soit $\boldsymbol{\mathcal{X}}$ un espace de Banach réel réflexif. Une fonction $\boldsymbol{f} \in \Gamma_0(\boldsymbol{\mathcal{X}})$ est *de Legendre* si elle est *essentiellement lisse* au sens où $\partial\boldsymbol{f}$ est à la fois localement borné et univoque sur son domaine, et *essentiellement strictement convexe* au sens où $\partial\boldsymbol{f}^*$ est localement borné sur son domaine et $\boldsymbol{f}$ est strictement convexe sur les sous-ensembles convexes de $\operatorname{dom}\partial\boldsymbol{f}$. De plus, $\boldsymbol{f}$ est Gâteaux différentiable sur $\operatorname{int}\operatorname{dom}\boldsymbol{f} \neq \varnothing$ et la distance de Bregman associée à $\boldsymbol{f}$ est

$$\begin{aligned} D^{\boldsymbol{f}} \colon \boldsymbol{\mathcal{X}} \times \boldsymbol{\mathcal{X}} &\to [0,+\infty] \\ (\boldsymbol{x},\boldsymbol{y}) &\mapsto \begin{cases} \boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{f}(\boldsymbol{y}) - \langle \boldsymbol{x}-\boldsymbol{y}, \nabla\boldsymbol{f}(\boldsymbol{y})\rangle, & \text{si } \boldsymbol{y} \in \operatorname{int}\operatorname{dom}\boldsymbol{f}; \\ +\infty, & \text{sinon.} \end{cases} \end{aligned} \tag{3.1}$$

Soit $\boldsymbol{C}$ un sous-ensemble convexe fermé de $\boldsymbol{\mathcal{X}}$ tel que $\boldsymbol{C} \cap \operatorname{int} \operatorname{dom} \boldsymbol{f} \neq \varnothing$. Alors

$$\begin{aligned} P_{\boldsymbol{C}}^{\boldsymbol{f}} \colon \operatorname{int} \operatorname{dom} \boldsymbol{f} &\to \boldsymbol{C} \cap \operatorname{int} \operatorname{dom} \boldsymbol{f} \\ \boldsymbol{y} &\mapsto \underset{\boldsymbol{x} \in \boldsymbol{C}}{\operatorname{argmin}}\, D^{\boldsymbol{f}}(\boldsymbol{x}, \boldsymbol{y}). \end{aligned} \tag{3.2}$$

De plus, posons

$$\begin{aligned} (\forall (\boldsymbol{x}, \boldsymbol{y}) \in (\operatorname{int} \operatorname{dom} \boldsymbol{f})^2) \quad H^{\boldsymbol{f}}(\boldsymbol{x}, \boldsymbol{y}) &= \big\{ \boldsymbol{z} \in \boldsymbol{\mathcal{X}} \bigm| \big\langle \boldsymbol{z} - \boldsymbol{y}, \nabla \boldsymbol{f}(\boldsymbol{x}) - \nabla \boldsymbol{f}(\boldsymbol{y}) \big\rangle \leqslant 0 \big\} \\ &= \big\{ \boldsymbol{z} \in \boldsymbol{\mathcal{X}} \bigm| D^{\boldsymbol{f}}(\boldsymbol{z}, \boldsymbol{y}) + D^{\boldsymbol{f}}(\boldsymbol{y}, \boldsymbol{x}) \leqslant D^{\boldsymbol{f}}(\boldsymbol{z}, \boldsymbol{x}) \big\}. \end{aligned} \tag{3.3}$$

Soit $(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}) \in (\operatorname{int} \operatorname{dom} \boldsymbol{f})^3$ tel que $H^{\boldsymbol{f}}(\boldsymbol{x}, \boldsymbol{y}) \cap H^{\boldsymbol{f}}(\boldsymbol{y}, \boldsymbol{z}) \cap \operatorname{int} \operatorname{dom} \boldsymbol{f} \neq \varnothing$. Alors

$$Q^{\boldsymbol{f}}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}) = P^{\boldsymbol{f}}_{H^{\boldsymbol{f}}(\boldsymbol{x}, \boldsymbol{y}) \cap H^{\boldsymbol{f}}(\boldsymbol{y}, \boldsymbol{z})} \boldsymbol{x}. \tag{3.4}$$

Nous rappelons le résultat suivant, qui lui-même comprend le cadre hilbertien de [3], ce dernier remontant aux travaux de [7] (cf. [5]).

Proposition 3.3 [4, Section 5.1] *Soit $\boldsymbol{\mathcal{X}}$ un espace de Banach réel réflexif, soit $\boldsymbol{C}$ un sous-ensemble convexe fermé de $\overline{\operatorname{dom}}\, \boldsymbol{f}$, soit $\boldsymbol{f} \in \Gamma_0(\boldsymbol{\mathcal{X}})$ une fonction de Legendre telle que $\boldsymbol{C} \cap \operatorname{int} \operatorname{dom} \boldsymbol{f} \neq \varnothing$, et soit $\boldsymbol{x}_0 \in \operatorname{int} \operatorname{dom} \boldsymbol{f}$. À chaque itération $n \in \mathbb{N}$, on prend $\boldsymbol{x}_{n+1/2} \in \boldsymbol{\mathcal{X}}$ tel que $\boldsymbol{C} \subset H^{\boldsymbol{f}}(\boldsymbol{x}_n, \boldsymbol{x}_{n+1/2})$, et on pose*

$$\boldsymbol{x}_{n+1} = Q^{\boldsymbol{f}}(\boldsymbol{x}_0, \boldsymbol{x}_n, \boldsymbol{x}_{n+1/2}). \tag{3.5}$$

Alors $[\, \boldsymbol{x}_n \rightharpoonup P^{\boldsymbol{f}}_{\boldsymbol{C}} \boldsymbol{x}_0$ et $\boldsymbol{f}(\boldsymbol{x}_n) \to \boldsymbol{f}(P^{\boldsymbol{f}}_{\boldsymbol{C}} \boldsymbol{x}_0) \,] \Leftrightarrow D^{\boldsymbol{f}}(\boldsymbol{x}_n, P^{\boldsymbol{f}}_{\boldsymbol{C}} \boldsymbol{x}_0) \to 0 \Leftrightarrow \mathfrak{W}(\boldsymbol{x}_n)_{n \in \mathbb{N}} \subset \boldsymbol{C}$.

Remarque 3.4 Dans la démonstration de convergence de l'algorithme (3.5), on utilise le fait que

$$\boldsymbol{C} \subset \bigcap_{n \in \mathbb{N}} H^{\boldsymbol{f}}(\boldsymbol{x}_0, \boldsymbol{x}_n). \tag{3.6}$$

Cette inclusion est une conséquence des inclusions

$$(\forall n \in \mathbb{N}) \quad \boldsymbol{C} \subset H^{\boldsymbol{f}}(\boldsymbol{x}_n, \boldsymbol{x}_{n+1/2}). \tag{3.7}$$

Lorsque $\boldsymbol{C} = \operatorname{zer} \boldsymbol{M}$ où $\boldsymbol{M} \colon \boldsymbol{\mathcal{X}} \to 2^{\boldsymbol{\mathcal{X}}^*}$ est un opérateur maximalement monotone, il est proposé de choisir dans [4]

$$(\forall n \in \mathbb{N}) \quad \boldsymbol{x}_{n+1/2} = (\nabla \boldsymbol{f} + \gamma_n \boldsymbol{M})^{-1}(\nabla \boldsymbol{f}(\boldsymbol{x}_n)), \tag{3.8}$$

où $(\gamma_n)_{n \in \mathbb{N}} \in \left]0, +\infty\right[^{\mathbb{N}}$ est telle que $\inf_{n \in \mathbb{N}} \gamma_n > 0$. Puisque les opérateurs $((\nabla \boldsymbol{f} + \gamma_n \boldsymbol{M})^{-1} \circ \nabla \boldsymbol{f})_{n \in \mathbb{N}}$ peuvent être difficiles à évaluer, il est utile d'introduire des tolérances dans (3.8) pour le calcul de la suite $(\boldsymbol{x}_{n+1/2})_{n \in \mathbb{N}}$. Dans ce cas, les inclusions (3.6) et (3.7) ne sont pas nécessairement satisfaites et nous nous proposons d'étudier le comportement asymptotique de la méthode dans ce contexte.

3.2 Algorithme d'Haugazeau-Bregman abstrait avec erreurs

Dans l'algorithme (3.5), on projette l'itérée courante sur l'intersection de deux demi-espaces affines fermés séparant cette itérée et $\boldsymbol{C}$. On peut remplacer l'un des demi-espaces par un sous-ensemble convexe fermé convenable.

Proposition 3.5 *Soit* $\boldsymbol{f} \in \Gamma_0(\boldsymbol{\mathcal{X}})$ *une fonction de Legendre, soit* $\boldsymbol{C}$ *un sous-ensemble convexe fermé de* $\overline{\operatorname{dom} \boldsymbol{f}}$ *tel que* $\boldsymbol{C} \cap \operatorname{int} \operatorname{dom} \boldsymbol{f} \neq \varnothing$*, soit* $\boldsymbol{x}_0 \in \operatorname{int} \operatorname{dom} \boldsymbol{f}$*, et soit* $\overline{\boldsymbol{x}} = P^{\boldsymbol{f}}_{\boldsymbol{C}} \boldsymbol{x}_0$*. Pour chaque* $n \in \mathbb{N}$*, soit* $\boldsymbol{C}_n \subset \boldsymbol{\mathcal{X}}$ *un sous-ensemble convexe fermé tel que* $\boldsymbol{C}_n \cap H^{\boldsymbol{f}}(\boldsymbol{x}_0, \boldsymbol{x}_n) \cap \operatorname{int} \operatorname{dom} \boldsymbol{f} \neq \varnothing$*, et soit*

$$\boldsymbol{x}_{n+1} = P^{\boldsymbol{f}}_{\boldsymbol{C}_n \cap H^{\boldsymbol{f}}(\boldsymbol{x}_0, \boldsymbol{x}_n)} \boldsymbol{x}_0. \tag{3.9}$$

Supposons qu'il existe $(\boldsymbol{a}_n)_{n\in\mathbb{N}} \in \boldsymbol{\mathcal{X}}^{\mathbb{N}}$ *telle que* $\boldsymbol{a}_n \to 0$ *et que*

$$(\forall n \in \mathbb{N}) \quad \overline{\boldsymbol{x}} + \boldsymbol{a}_n \in H^{\boldsymbol{f}}(\boldsymbol{x}_0, \boldsymbol{x}_n). \tag{3.10}$$

Alors les propriétés suivantes sont satisfaites :

(i) $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ *est une suite bornée dans* $\operatorname{int} \operatorname{dom} \boldsymbol{f}$.

(ii) $\sum_{n\in\mathbb{N}} D^{\boldsymbol{f}}(\boldsymbol{x}_{n+1}, \boldsymbol{x}_n) < +\infty$.

(iii) $[\, \boldsymbol{x}_n \rightharpoonup \overline{\boldsymbol{x}}$ *et* $\boldsymbol{f}(\boldsymbol{x}_n) \to \boldsymbol{f}(\overline{\boldsymbol{x}}) \,] \Leftrightarrow D^{\boldsymbol{f}}(\boldsymbol{x}_n, \overline{\boldsymbol{x}}) \to 0 \Leftrightarrow \mathfrak{W}(\boldsymbol{x}_n)_{n\in\mathbb{N}} \subset \boldsymbol{C}$.

Démonstration. Pour tout $n \in \mathbb{N}$, parce que $\boldsymbol{C}_n \cap H^{\boldsymbol{f}}(\boldsymbol{x}_0, \boldsymbol{x}_n)$ est un sous-ensemble convexe fermé et non vide de $\boldsymbol{\mathcal{X}}$ et $\boldsymbol{C}_n \cap H^{\boldsymbol{f}}(\boldsymbol{x}_0, \boldsymbol{x}_n) \cap \operatorname{int} \operatorname{dom} \boldsymbol{f} \neq \varnothing$, on déduit de (4.7) que

$$P^{\boldsymbol{f}}_{\boldsymbol{C}_n \cap H^{\boldsymbol{f}}(\boldsymbol{x}_0, \boldsymbol{x}_n)} \colon \operatorname{int} \operatorname{dom} \boldsymbol{f} \to \boldsymbol{C}_n \cap H^{\boldsymbol{f}}(\boldsymbol{x}_0, \boldsymbol{x}_n) \cap \operatorname{int} \operatorname{dom} \boldsymbol{f}. \tag{3.11}$$

En raisonnant par récurrence, on déduit que $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ est une suite bien définie dans $\operatorname{int} \operatorname{dom} \boldsymbol{f}$. Il résulte de (3.9) que

$$(\forall n \in \mathbb{N}) \quad \boldsymbol{x}_{n+1} \in H^{\boldsymbol{f}}(\boldsymbol{x}_0, \boldsymbol{x}_n), \tag{3.12}$$

et donc on déduit de (3.3) que

$$(\forall n \in \mathbb{N}) \quad D^{\boldsymbol{f}}(\boldsymbol{x}_{n+1}, \boldsymbol{x}_n) + D^{\boldsymbol{f}}(\boldsymbol{x}_n, \boldsymbol{x}_0) \leqslant D^{\boldsymbol{f}}(\boldsymbol{x}_{n+1}, \boldsymbol{x}_0), \tag{3.13}$$

et en particulier,

$$(\forall n \in \mathbb{N}) \quad D^{\boldsymbol{f}}(\boldsymbol{x}_n, \boldsymbol{x}_0) \leqslant D^{\boldsymbol{f}}(\boldsymbol{x}_{n+1}, \boldsymbol{x}_0). \tag{3.14}$$

Cela implique que $(D^{\boldsymbol{f}}(\boldsymbol{x}_n,\boldsymbol{x}_0))_{n\in\mathbb{N}}$ est croissante. Comme $\operatorname{int}\operatorname{dom}\boldsymbol{f}$ est ouvert et $\overline{\boldsymbol{x}}+\boldsymbol{a}_n\to\overline{\boldsymbol{x}}\in\operatorname{int}\operatorname{dom}\boldsymbol{f}$, on suppose sans perte de généralité que $(\overline{\boldsymbol{x}}+\boldsymbol{a}_n)_{n\in\mathbb{N}}\in(\operatorname{int}\operatorname{dom}\boldsymbol{f})^{\mathbb{N}}$. Par ailleurs, comme (3.10) implique que

$$(\forall n\in\mathbb{N})\quad \big\langle\overline{\boldsymbol{x}}+\boldsymbol{a}_n-\boldsymbol{x}_n,\nabla\boldsymbol{f}(\boldsymbol{x}_0)-\nabla\boldsymbol{f}(x_n)\big\rangle\leqslant 0, \tag{3.15}$$

on obtient

$$(\forall n\in\mathbb{N})\quad D^{\boldsymbol{f}}(\overline{\boldsymbol{x}}+\boldsymbol{a}_n,\boldsymbol{x}_n)+D^{\boldsymbol{f}}(\boldsymbol{x}_n,\boldsymbol{x}_0)\leqslant D^{\boldsymbol{f}}(\overline{\boldsymbol{x}}+\boldsymbol{a}_n,\boldsymbol{x}_0), \tag{3.16}$$

et donc

$$(\forall n\in\mathbb{N})\quad D^{\boldsymbol{f}}(\boldsymbol{x}_n,\boldsymbol{x}_0)\leqslant D^{\boldsymbol{f}}(\overline{\boldsymbol{x}}+\boldsymbol{a}_n,\boldsymbol{x}_0). \tag{3.17}$$

De plus, comme $\overline{\boldsymbol{x}}+\boldsymbol{a}_n\to\overline{\boldsymbol{x}}$, et comme $\boldsymbol{f}$ est continue sur $\operatorname{int}\operatorname{dom}\boldsymbol{f}$, on obtient

$$D^{\boldsymbol{f}}(\overline{\boldsymbol{x}}+\boldsymbol{a}_n,\boldsymbol{x}_0)\to D^{\boldsymbol{f}}(\overline{\boldsymbol{x}},\boldsymbol{x}_0). \tag{3.18}$$

Par suite, on déduit de (3.17) que $(D^{\boldsymbol{f}}(\boldsymbol{x}_n,\boldsymbol{x}_0))_{n\in\mathbb{N}}$ est bornée. Elle est donc convergente.

(i) : Puisque $D^{\boldsymbol{f}}(\cdot,\boldsymbol{x}_0)$ est coercive [1, Lemma 7.3(v)], on déduit que $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ est bornée.

(ii) : On déduit de (3.13) que

$$\sum_{n\in\mathbb{N}}D^{\boldsymbol{f}}(\boldsymbol{x}_{n+1},\boldsymbol{x}_n)\leqslant\sum_{n\in\mathbb{N}}\big(D^{\boldsymbol{f}}(\boldsymbol{x}_{n+1},\boldsymbol{x}_0)-D^{\boldsymbol{f}}(\boldsymbol{x}_n,\boldsymbol{x}_0)\big)\leqslant\sup_{n\in\mathbb{N}}D^{\boldsymbol{f}}(\boldsymbol{x}_n,\boldsymbol{x}_0)<+\infty. \tag{3.19}$$

(iii) : La première equivalence est établie dans [4, Proposition 2.2(ii)]. Pour la deuxième équivalence, soit $\boldsymbol{x}\in\mathfrak{W}(\boldsymbol{x}_n)_{n\in\mathbb{N}}$, i.e., $\boldsymbol{x}_{k_n}\rightharpoonup\boldsymbol{x}$. D'abord, supposons que $D^{\boldsymbol{f}}(\boldsymbol{x}_n,\overline{\boldsymbol{x}})\to 0$. Alors, il résulte de [1, Lemma 7.3(ii)] que

$$D^{\boldsymbol{f}}(\boldsymbol{x},\overline{\boldsymbol{x}})\leqslant\varliminf D^{\boldsymbol{f}}(\boldsymbol{x}_{k_n},\overline{\boldsymbol{x}})=0, \tag{3.20}$$

et donc $D^{\boldsymbol{f}}(\boldsymbol{x},\overline{\boldsymbol{x}})=0$. Par suite, $\boldsymbol{x}=\overline{\boldsymbol{x}}\in\boldsymbol{C}$. Maintenant, supposons que $\mathfrak{W}(\boldsymbol{x}_n)_{n\in\mathbb{N}}\subset\boldsymbol{C}$. On déduit de [1, Lemma 7.3(ii)], de (3.17), et de (3.18) que

$$D^{\boldsymbol{f}}(\boldsymbol{x},\boldsymbol{x}_0)\leqslant\varliminf D^{\boldsymbol{f}}(\boldsymbol{x}_{k_n},\boldsymbol{x}_0)=\lim D^{\boldsymbol{f}}(\boldsymbol{x}_n,\boldsymbol{x}_0)\leqslant\lim D^{\boldsymbol{f}}(\overline{\boldsymbol{x}}+\boldsymbol{a}_n,\boldsymbol{x}_0)=D^{\boldsymbol{f}}(\overline{\boldsymbol{x}},\boldsymbol{x}_0), \tag{3.21}$$

et ainsi $\boldsymbol{x}=\overline{\boldsymbol{x}}$. Cela montre que $\mathfrak{W}(\boldsymbol{x}_n)_{n\in\mathbb{N}}=\{\overline{\boldsymbol{x}}\}$, et donc, comme $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ est bornée, $\boldsymbol{x}_n\rightharpoonup\overline{\boldsymbol{x}}$. Par conséquent, comme (3.21) implique que

$$D^{\boldsymbol{f}}(\overline{\boldsymbol{x}},\boldsymbol{x}_0)\leqslant\lim D^{\boldsymbol{f}}(\boldsymbol{x}_n,\boldsymbol{x}_0)=D^{\boldsymbol{f}}(\overline{\boldsymbol{x}},\boldsymbol{x}_0), \tag{3.22}$$

on obtient $D^{\boldsymbol{f}}(\boldsymbol{x}_n,\boldsymbol{x}_0)\to D^{\boldsymbol{f}}(\overline{\boldsymbol{x}},\boldsymbol{x}_0)$, ce qui implique que

$$D^{\boldsymbol{f}}(\boldsymbol{x}_n,\overline{\boldsymbol{x}})=D^{\boldsymbol{f}}(\boldsymbol{x}_n,\boldsymbol{x}_0)-D^{\boldsymbol{f}}(\overline{\boldsymbol{x}},\boldsymbol{x}_0)-\big\langle\boldsymbol{x}_n-\overline{\boldsymbol{x}},\nabla\boldsymbol{f}(\boldsymbol{x}_0)\big\rangle\to 0. \tag{3.23}$$

□

Remarque 3.6 Considérons la Proposition 3.5. Dans l'algorithme (3.9), une petite erreur se produit à chaque itération $n \in \mathbb{N}$. On n'exige pas que $H^{\boldsymbol{f}}(\boldsymbol{x}_0, \boldsymbol{x}_n)$ contienne $\boldsymbol{C}$ mais qu'il contienne $\overline{\boldsymbol{x}} + \boldsymbol{a}_n$.

Remarque 3.7 Dans la Proposition 3.5, supposons que $\boldsymbol{C} = \bigcap_{i \in I} \operatorname{zer} \boldsymbol{A}_i$, où I est un ensemble totalement ordonné au plus dénombrable et $(\boldsymbol{A}_i)_{i \in I}$ sont des opérateurs maximalement monotones de $\boldsymbol{\mathcal{X}}$ dans $2^{\boldsymbol{\mathcal{X}}^*}$ tels que $\bigcap_{i \in I} \operatorname{zer} \boldsymbol{A}_i \neq \varnothing$.

(i) À chaque itération $n \in \mathbb{N}$, on prend $\mathrm{i}(n) \in I$, $\gamma_n \in \left]0, +\infty\right[$, et on pose

$$\boldsymbol{C}_n = H^{\boldsymbol{f}}\big(\boldsymbol{x}_n, (\nabla \boldsymbol{f} + \gamma_n \boldsymbol{A}_{\mathrm{i}(n)})^{-1}(\nabla \boldsymbol{f}(\boldsymbol{x}_n))\big). \tag{3.24}$$

Alors on retrouve les résultats de [4, Section 5.3]

(ii) Supposons que I soit fini et que

$$(\forall n \in \mathbb{N}) \quad \boldsymbol{C}_n = \bigcap_{i \in I} \Big\{ \boldsymbol{x} \in \boldsymbol{\mathcal{X}} \;\Big|\; D^{\boldsymbol{f}}\big(\boldsymbol{x}, (\nabla \boldsymbol{f} + \gamma_n \boldsymbol{A}_i)^{-1}(\nabla \boldsymbol{f}(\boldsymbol{x}_n) + \boldsymbol{b}_{i,n})\big) \leqslant D^{\boldsymbol{f}}\big(\boldsymbol{x}, \nabla \boldsymbol{f}^*(\nabla \boldsymbol{f}(\boldsymbol{x}_n) + \boldsymbol{b}^*_{i,n})\big) \Big\}, \tag{3.25}$$

où $(\gamma_n)_{n \in \mathbb{N}} \in \left]0, +\infty\right[^{\mathbb{N}}$ telle que $\inf_{n \in \mathbb{N}} \gamma_n > 0$ et $(\forall i \in I)$ $(\boldsymbol{b}^*_{i,n})_{n \in \mathbb{N}} \in (\boldsymbol{\mathcal{X}}^*)^{\mathbb{N}}$ telle que $\boldsymbol{b}^*_{i,n} \to 0$. Alors on récupère [8, Theorem 4.1].

(iii) Supposons que I soit fini et que

$$(\forall n \in \mathbb{N}) \quad \boldsymbol{C}_n = \Big\{ \boldsymbol{x} \in \boldsymbol{\mathcal{X}} \;\Big|\; D^{\boldsymbol{f}}\big(\boldsymbol{x}, (\nabla \boldsymbol{f} + \gamma_n \boldsymbol{A}_i)^{-1}(\nabla \boldsymbol{f}(\boldsymbol{x}_n))\big) \leqslant D^{\boldsymbol{f}}(\boldsymbol{x}, \boldsymbol{x}_n + \boldsymbol{b}_{i,n}) \Big\}, \tag{3.26}$$

où $(\gamma_n)_{n \in \mathbb{N}} \in \left]0, +\infty\right[^{\mathbb{N}}$ telle que $\inf_{n \in \mathbb{N}} \gamma_n > 0$ et $(\forall i \in I)$ $(\boldsymbol{b}_{i,n})_{n \in \mathbb{N}} \in \boldsymbol{\mathcal{X}}^{\mathbb{N}}$ telle que $\boldsymbol{b}_{i,n} \to 0$. Alors on retrouve [8, Theorem 4.2].

(iv) Supposons que $\boldsymbol{\mathcal{X}}$ soit hilbertien et que $\boldsymbol{f} = \|\cdot\|^2/2$. À chaque itération $n \in \mathbb{N}$, on prend $\mathrm{i}(n) \in I$, $\gamma_n \in \left]0, +\infty\right[$, et on pose

$$\boldsymbol{C}_n = H^{\boldsymbol{f}}\big(\boldsymbol{x}_n, (\mathrm{Id} + \gamma_n \boldsymbol{A}_{\mathrm{i}(n)})^{-1} \boldsymbol{x}_n\big). \tag{3.27}$$

Alors on récupère [3].

(v) Supposons que $\boldsymbol{\mathcal{X}}$ soit hilbertien, que $\boldsymbol{f} = \|\cdot\|^2/2$, et que $I = \{1\}$. Soient $\sigma \in \left[0, 1\right[$ et $(\mu_n)_{n \in \mathbb{N}} \in \left]0, +\infty\right[^{\mathbb{N}}$. À chaque itération $n \in \mathbb{N}$, on prend $\boldsymbol{y}_n = (\mathrm{Id} + \mu_n^{-1} \boldsymbol{A}_1)^{-1}(\boldsymbol{x}_n - \mu_n^{-1} \boldsymbol{e}_n)$ et $\boldsymbol{y}^*_n = \mu_n(\boldsymbol{x}_n - \boldsymbol{y}_n) - \boldsymbol{e}_n$, où $\boldsymbol{e}_n \in \boldsymbol{\mathcal{X}}$ tel que $\|\boldsymbol{e}_n\| \leqslant \sigma \max\{\|\boldsymbol{x}^*_n\|, \mu_n \|\boldsymbol{y}_n - \boldsymbol{x}_n\|\}$. De plus, posons

$$(\forall n \in \mathbb{N}) \quad \boldsymbol{C}_n = \big\{ \boldsymbol{x} \in \boldsymbol{\mathcal{X}} \mid \langle \boldsymbol{x} - \boldsymbol{y}_n, \boldsymbol{y}^*_n \rangle \leqslant 0 \big\}. \tag{3.28}$$

Alors on retrouve les résultats de [9].

Notons que les résultats ci-dessus imposent que les opérateurs monotones aient des zéros communs. Notons aussi que $\boldsymbol{C} \subset \bigcap_{n\in\mathbb{N}} \boldsymbol{C}_n$ et par le même raisonnement que dans la démonstration de [4, Proposition 3.3], on déduit que $\boldsymbol{C} \subset \bigcap_{n\in\mathbb{N}} H^{\boldsymbol{f}}(\boldsymbol{x}_0, \boldsymbol{x}_n)$.

Nous présentons deux cas particuliers.

Corollaire 3.8 *Soit $\boldsymbol{f} \in \Gamma_0(\boldsymbol{\mathcal{X}})$ une fonction de Legendre, soit $\boldsymbol{C}$ un sous-ensemble convexe fermé de $\overline{\operatorname{dom}}\,\boldsymbol{f}$ tel que $\boldsymbol{C} \cap \operatorname{int}\operatorname{dom}\boldsymbol{f} \neq \varnothing$, soit $\boldsymbol{x}_0 \in \operatorname{int}\operatorname{dom}\boldsymbol{f}$, et soit $\overline{\boldsymbol{x}} = P_{\boldsymbol{C}}^{\boldsymbol{f}}\boldsymbol{x}_0$. À chaque itération $n \in \mathbb{N}$, on prend $\boldsymbol{x}_{n+1/2} \in \operatorname{int}\operatorname{dom}\boldsymbol{f}$ et $\boldsymbol{b}_n \in \boldsymbol{\mathcal{X}}$ tels que $\overline{\boldsymbol{x}} \in H^{\boldsymbol{f}}(\boldsymbol{x}_n + \boldsymbol{b}_n, \boldsymbol{x}_{n+1/2})$, et on pose*

$$\boldsymbol{x}_{n+1} = P^{\boldsymbol{f}}_{H^{\boldsymbol{f}}(\boldsymbol{x}_0,\boldsymbol{x}_n)\cap H^{\boldsymbol{f}}(\boldsymbol{x}_n+\boldsymbol{b}_n,\boldsymbol{x}_{n+1/2})}\boldsymbol{x}_0. \tag{3.29}$$

Alors les propriétés suivantes sont satisfaites :

(i) $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ *est une suite bornée dans* $\operatorname{int}\operatorname{dom}\boldsymbol{f}$.

(ii) $\sum_{n\in\mathbb{N}} D^{\boldsymbol{f}}(\boldsymbol{x}_{n+1}, \boldsymbol{x}_n) < +\infty$.

(iii) $\sum_{n\in\mathbb{N}} D^{\boldsymbol{f}}(\boldsymbol{x}_{n+1/2}, \boldsymbol{x}_n + \boldsymbol{b}_n) < +\infty$.

(iv) $[\,\boldsymbol{x}_n \rightharpoonup \overline{\boldsymbol{x}}$ *et* $\boldsymbol{f}(\boldsymbol{x}_n) \to \boldsymbol{f}(\overline{\boldsymbol{x}})\,] \Leftrightarrow D^{\boldsymbol{f}}(\boldsymbol{x}_n, \overline{\boldsymbol{x}}) \to 0 \Leftrightarrow \mathfrak{W}(\boldsymbol{x}_n)_{n\in\mathbb{N}} \subset \boldsymbol{C}$.

Démonstration. Posons $(\forall n \in \mathbb{N})\ \boldsymbol{C}_n = H^{\boldsymbol{f}}(\boldsymbol{x}_n + \boldsymbol{b}_n, \boldsymbol{x}_{n+1/2})$. Alors le Corollaire est un cas particulier de la Proposition 3.5 avec $\boldsymbol{a}_n \equiv 0$. □

Corollaire 3.9 *Soit $\boldsymbol{f} \in \Gamma_0(\boldsymbol{\mathcal{X}})$ une fonction de Legendre, soit $\boldsymbol{C}$ un sous-ensemble convexe fermé de $\overline{\operatorname{dom}}\,\boldsymbol{f}$ tel que $\boldsymbol{C} \cap \operatorname{int}\operatorname{dom}\boldsymbol{f} \neq \varnothing$, soit $\boldsymbol{x}_0 \in \operatorname{int}\operatorname{dom}\boldsymbol{f}$, et soit $\overline{\boldsymbol{x}} = P_{\boldsymbol{C}}^{\boldsymbol{f}}\boldsymbol{x}_0$. À chaque itération $n \in \mathbb{N}$, on prend $\boldsymbol{x}_{n+1/2} \in \operatorname{int}\operatorname{dom}\boldsymbol{f}$ et $\boldsymbol{b}_n \in \boldsymbol{\mathcal{X}}$ tels que $\overline{\boldsymbol{x}} \in H^{\boldsymbol{f}}(\boldsymbol{x}_n, \boldsymbol{x}_{n+1/2} + \boldsymbol{b}_n)$, et on pose*

$$\boldsymbol{x}_{n+1} = P^{\boldsymbol{f}}_{H^{\boldsymbol{f}}(\boldsymbol{x}_0,\boldsymbol{x}_n)\cap H^{\boldsymbol{f}}(\boldsymbol{x}_n,\boldsymbol{x}_{n+1/2}+\boldsymbol{b}_n)}\boldsymbol{x}_0. \tag{3.30}$$

Alors les propriétés suivantes sont satisfaites :

(i) $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ *est une suite bornée dans* $\operatorname{int}\operatorname{dom}\boldsymbol{f}$.

(ii) $\sum_{n\in\mathbb{N}} D^{\boldsymbol{f}}(\boldsymbol{x}_{n+1}, \boldsymbol{x}_n) < +\infty$.

(iii) $\sum_{n\in\mathbb{N}} D^{\boldsymbol{f}}(\boldsymbol{x}_{n+1/2} + \boldsymbol{b}_n, \boldsymbol{x}_n) < +\infty$.

(iv) $[\,\boldsymbol{x}_n \rightharpoonup \overline{\boldsymbol{x}}$ *et* $\boldsymbol{f}(\boldsymbol{x}_n) \to \boldsymbol{f}(\overline{\boldsymbol{x}})\,] \Leftrightarrow D^{\boldsymbol{f}}(\boldsymbol{x}_n, \overline{\boldsymbol{x}}) \to 0 \Leftrightarrow \mathfrak{W}(\boldsymbol{x}_n)_{n\in\mathbb{N}} \subset \boldsymbol{C}$.

Démonstration. Posons $(\forall n \in \mathbb{N})\ \boldsymbol{C}_n = H^{\boldsymbol{f}}(\boldsymbol{x}_n, \boldsymbol{x}_{n+1/2} + \boldsymbol{b}_n)$. Alors le corollaire est un cas particulier de la Proposition 3.5 avec $\boldsymbol{a}_n \equiv 0$. □

3.3 Une application

Nous revisitons le cadre du Chapitre 2 avec une version perturbée du Théorème 2.24.

Problème 3.10 Soient $\mathcal{X}$ et $\mathcal{Y}$ des espaces de Banach réels réflexifs tels que $\mathcal{X} \neq \{0\}$ et $\mathcal{Y} \neq \{0\}$, soit $\boldsymbol{\mathcal{X}}$ l'espace vectoriel $\mathcal{X} \times \mathcal{Y}^*$ muni de la norme $(x, y^*) \mapsto \sqrt{\|x\|^2 + \|y^*\|^2}$, et soit $\boldsymbol{\mathcal{X}}^*$ le dual topologique de $\boldsymbol{\mathcal{X}}$, i.e., $\mathcal{X}^* \times \mathcal{Y}$ muni de la norme $(x^*, y) \mapsto \sqrt{\|x^*\|^2 + \|y\|^2}$. Soient $A\colon \mathcal{X} \to 2^{\mathcal{X}^*}$ et $B\colon \mathcal{Y} \to 2^{\mathcal{Y}^*}$ maximalement monotones et soit $L \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$. Considérons l'inclusion primale

$$\text{trouver } \ x \in \mathcal{X} \ \text{ tel que } \ 0 \in Ax + L^*BLx, \tag{3.31}$$

l'inclusion duale

$$\text{trouver } \ y^* \in \mathcal{Y}^* \ \text{ tel que } \ 0 \in -LA^{-1}(-L^*y^*) + B^{-1}y^*, \tag{3.32}$$

et l'ensemble de Kuhn-Tucker associé

$$\boldsymbol{Z} = \big\{(x, y^*) \in \boldsymbol{\mathcal{X}} \;\big|\; -L^*y^* \in Ax \quad \text{et} \quad Lx \in B^{-1}y^*\big\}. \tag{3.33}$$

Soient $f \in \Gamma_0(\mathcal{X})$ et $g \in \Gamma_0(\mathcal{Y})$ des fonctions de Legendre, soit $x_0 \in \operatorname{int}\operatorname{dom} f$, et soit $y_0^* \in \operatorname{int}\operatorname{dom} g^*$. Posons

$$\boldsymbol{f}\colon \boldsymbol{\mathcal{X}} \to \left]-\infty, +\infty\right] : (x, y^*) \mapsto f(x) + g^*(y^*), \tag{3.34}$$

et supposons que $\boldsymbol{Z} \cap \operatorname{int}\operatorname{dom} \boldsymbol{f} \neq \varnothing$. Le problème est de trouver la meilleure approximation au sens de Bregman $(\overline{x}, \overline{y}^*) = P_{\boldsymbol{Z}}^{\boldsymbol{f}}(x_0, y_0^*)$ de (x_0, y_0^*) dans $\boldsymbol{Z}$.

Pour énoncer nos résultats, nous rappelons la notion de coercivité d'un opérateur multivoque et la condition suivantes.

Définition 3.11 Soit $\mathcal{X}$ un espace de Banach réel réflexif tel que $\mathcal{X} \neq \{0\}$, soit $\mathcal{X}^*$ le dual topologique de $\mathcal{X}$, et soit $M\colon \mathcal{X} \to 2^{\mathcal{X}^*}$. Alors M est *coercif* si

$$(\exists\, z \in \operatorname{dom} M) \qquad \lim_{\|x\| \to +\infty} \inf \frac{\langle x - z, Mx \rangle}{\|x\|} = +\infty. \tag{3.35}$$

Condition 3.12 Soit $\mathcal{X}$ un espace de Banach réel réflexif et soit $f \in \Gamma_0(\mathcal{X})$ Gâteaux différentiable sur $\operatorname{int}\operatorname{dom} f \neq \varnothing$. Pour toute suite $(x_n)_{n\in\mathbb{N}}$ dans $\operatorname{int}\operatorname{dom} f$ et toute suite bornée $(y_n)_{n\in\mathbb{N}}$ dans $\operatorname{int}\operatorname{dom} f$,

$$D^f(x_n, y_n) \to 0 \quad \Rightarrow \quad x_n - y_n \to 0. \tag{3.36}$$

Le résultat principal dans cette section est le suivant.

Proposition 3.13 *Dans le Problème 3.10, soient $h \in \Gamma_0(\mathcal{X})$ et $j \in \Gamma_0(\mathcal{Y})$ des fonctions de Legendre telles que $\operatorname{int}\operatorname{dom} f \subset \operatorname{int}\operatorname{dom} h$, $L(\operatorname{int}\operatorname{dom} f) \subset \operatorname{int}\operatorname{dom} j$, et il existe ε et δ dans $]0,+\infty[$ tels que $\nabla h + \varepsilon A$ and $\nabla j + \delta B$ soient coercifs. Soit $\sigma \in [\max\{\varepsilon,\delta\},+\infty[$, soit $(c_n^*)_{n\in\mathbb{N}}$ une suite dans $\mathcal{X}^*$ telle que $c_n^* \to 0$, et soit $(d_n^*)_{n\in\mathbb{N}}$ une suite dans $\mathcal{Y}^*$ telle que $d_n^* \to 0$. On itère*

$$
\begin{array}{l}
\text{pour } n = 0,1,\ldots\\
\left\lfloor
\begin{array}{l}
(\gamma_n,\mu_n) \in [\varepsilon,\sigma]\times[\delta,\sigma]\\
a_n = (\nabla h + \gamma_n A)^{-1}\big(\nabla h(x_n) - \gamma_n L^* y_n^* + c_n^*\big)\\
a_n^* = \gamma_n^{-1}\big(\nabla h(x_n) - \nabla h(a_n) + c_n^*\big) - L^* y_n^*\\
b_n = (\nabla j + \mu_n B)^{-1}\big(\nabla j(Lx_n) + \mu_n y_n^* + d_n^*\big)\\
b_n^* = \mu_n^{-1}\big(\nabla j(Lx_n) - \nabla j(b_n) + d_n^*\big) + y_n^*\\
\boldsymbol{H}_n = \big\{(x,y^*) \in \boldsymbol{\mathcal{X}} \;\big|\; \langle x, a_n^* + L^* b_n^*\rangle + \langle b_n - La_n, y^*\rangle \leqslant \langle a_n, a_n^*\rangle + \langle b_n, b_n^*\rangle\big\}\\
(x_{n+1/2}, y_{n+1/2}^*) = P_{\boldsymbol{H}_n}^{\boldsymbol{f}}(x_n, y_n^*)\\
(x_{n+1}, y_{n+1}^*) = Q^{\boldsymbol{f}}\big((x_0,y_0^*),(x_n,y_n^*),(x_{n+1/2},y_{n+1/2}^*)\big).
\end{array}
\right.
\end{array}
\tag{3.37}
$$

Alors (3.37) engendre des suites infinies $(x_n)_{n\in\mathbb{N}}$ et $(y_n^*)_{n\in\mathbb{N}}$ et les propriétés suivantes sont satisfaites :

(i) $\sum_{n\in\mathbb{N}} D^f(x_{n+1},x_n) < +\infty$ *et* $\sum_{n\in\mathbb{N}} D^{g^*}(y_{n+1}^*,y_n^*) < +\infty$.

(ii) $\sum_{n\in\mathbb{N}} D^f(x_{n+1/2},x_n) < +\infty$ *et* $\sum_{n\in\mathbb{N}} D^{g^*}(y_{n+1/2}^*,y_n^*) < +\infty$.

(iii) *Supposons que f, g^*, h, et j satisfassent la Condition 3.12 et que ∇h et ∇j soient uniformément continus sur les sous-ensembles bornés de $\operatorname{int}\operatorname{dom} h$ et $\operatorname{int}\operatorname{dom} j$, respectivement. Alors $x_n \to \overline{x}$ et $y_n^* \to \overline{y}^*$.*

Démonstration. (i)–(ii) : Raisonner comme dans la démonstration du Théorème 2.24(i)&(ii).

(iii) : Notons que $(c_n^*)_{n\in\mathbb{N}}$ et $(d_n^*)_{n\in\mathbb{N}}$ sont bornées. Comme dans (2.91), (2.97), (2.98), et (2.99), on déduit que

$$
\begin{aligned}
&(x_n)_{n\in\mathbb{N}} \in (\operatorname{int}\operatorname{dom} f)^{\mathbb{N}},\ (y_n^*)_{n\in\mathbb{N}} \in (\operatorname{int}\operatorname{dom} g^*)^{\mathbb{N}},\ (a_n)_{n\in\mathbb{N}} \in (\operatorname{int}\operatorname{dom} h)^{\mathbb{N}},\\
&\qquad (b_n)_{n\in\mathbb{N}} \in (\operatorname{int}\operatorname{dom} j)^{\mathbb{N}},\ (a_n^*)_{n\in\mathbb{N}},\ \text{et } (b_n^*)_{n\in\mathbb{N}} \text{ sont bornées.}
\end{aligned}
\tag{3.38}
$$

D'une part, comme (ii) implique que $D^f(x_{n+1/2},x_n) \to 0$ et $D^{g^*}(y_{n+1/2}^*,y_n^*) \to 0$, et comme f et g^* vérifient la Condition 3.12, on obtient

$$
x_{n+1/2} - x_n \to 0 \quad \text{et} \quad y_{n+1/2}^* - y_n^* \to 0.
\tag{3.39}
$$

D'autre part, on déduit de (3.37) que $(\forall n \in \mathbb{N})$ $(x_{n+1/2}, y_{n+1/2}^*) \in \boldsymbol{H}_n$, et donc

$$
(\forall n \in \mathbb{N}) \quad \langle x_{n+1/2}, a_n^* + L^* b_n^*\rangle + \langle b_n - La_n, y_{n+1/2}^*\rangle \leqslant \langle a_n, a_n^*\rangle + \langle b_n, b_n^*\rangle.
\tag{3.40}
$$

Par conséquent,

$$
\begin{aligned}
(\forall n\in\mathbb{N})\quad & \|x_n-x_{n+1/2}\|\,\|a_n^*+L^*b_n^*\|+\|b_n-La_n\|\,\|y_n^*-y_{n+1/2}^*\|\\
&\geqslant \big\langle x_n-x_{n+1/2},a_n^*+L^*b_n^*\big\rangle+\big\langle b_n-La_n,y_n^*-y_{n+1/2}^*\big\rangle\\
&\geqslant \big\langle x_n,a_n^*+L^*b_n^*\big\rangle+\big\langle b_n-La_n,y_n^*\big\rangle-\big\langle a_n,a_n^*\big\rangle-\big\langle b_n,b_n^*\big\rangle\\
&=\big\langle x_n-a_n,a_n^*+L^*y_n^*\big\rangle+\big\langle Lx_n-b_n,b_n^*-y_n^*\big\rangle\\
&=\gamma_n^{-1}\big\langle x_n-a_n,\nabla h(x_n)-\nabla h(a_n)+c_n^*\big\rangle\\
&\quad+\mu_n^{-1}\big\langle Lx_n-b_n,\nabla j(Lx_n)-\nabla j(b_n)+d_n^*\big\rangle\\
&\geqslant \gamma_n^{-1}\big(D^h(x_n,a_n)+D^h(a_n,x_n)-\|c_n^*\|\,\|x_n-a_n\|\big)\\
&\quad+\mu_n^{-1}\big(D^j(Lx_n,b_n)+D^j(b_n,Lx_n)-\|d_n^*\|\,\|Lx_n-b_n\|\big)\\
&\geqslant \sigma^{-1}\big(D^h(x_n,a_n)+D^j(Lx_n,b_n)\big)\\
&\quad-\varepsilon^{-1}\|c_n^*\|\,\|x_n-a_n\|-\delta^{-1}\|d_n^*\|\,\|Lx_n-b_n\|.
\end{aligned}
\tag{3.41}
$$

Ainsi

$$
\begin{aligned}
(\forall n\in\mathbb{N})\quad & D^h(x_n,a_n)+D^j(Lx_n,b_n)\leqslant \sigma\|x_n-x_{n+1/2}\|\,\|a_n^*+L^*b_n^*\|\\
&+\sigma\|b_n-La_n\|\,\|y_n^*-y_{n+1/2}^*\|+\sigma\varepsilon^{-1}\|c_n^*\|\,\|x_n-a_n\|+\sigma\delta^{-1}\|d_n^*\|\,\|Lx_n-b_n\|.
\end{aligned}
\tag{3.42}
$$

En faisant tendre n vers $+\infty$ dans (3.42) et ensuite en utilisant (3.39), on obtient

$$
D^h(x_n,a_n)\to 0\quad\text{et}\quad D^j(Lx_n,b_n)\to 0.
\tag{3.43}
$$

Comme h et j satisfont la Condition 3.12, il résulte de (3.43) que

$$
x_n-a_n\to 0\quad\text{et}\quad Lx_n-b_n\to 0.
\tag{3.44}
$$

Par suite, puisque ∇h et ∇j sont uniformément continus sur les sous-ensembles bornés de $\operatorname{int}\operatorname{dom} h$ et $\operatorname{int}\operatorname{dom} j$, respectivement, et puisque $c_n^*\to 0$ et $d_n^*\to 0$, il vient

$$
\gamma_n^{-1}\big(\nabla h(x_n)-\nabla h(a_n)+c_n^*\big)\to 0\quad\text{et}\quad \mu_n^{-1}\big(\nabla j(Lx_n)-\nabla j(b_n)+d_n^*\big)\to 0.
\tag{3.45}
$$

En combinant avec (3.37), on obtient

$$
a_n^*+Ly_n^*\to 0\quad\text{et}\quad b_n^*-y_n^*\to 0,
\tag{3.46}
$$

ce qui implique que

$$
a_n^*+L^*b_n^*\to 0.
\tag{3.47}
$$

Maintenant, soit $\boldsymbol{x} = (x, y^*) \in \mathfrak{W}(\boldsymbol{x}_n)_{n\in\mathbb{N}}$, i.e., $\boldsymbol{x}_{k_n} \rightharpoonup \overline{\boldsymbol{x}}$. Alors $x_{k_n} \rightharpoonup x$ et $y^*_{k_n} \rightharpoonup y^*$. Par suite, on déduit de (3.44), (3.46), (3.47), et (3.37) que

$$\begin{cases} a_{k_n} \rightharpoonup x, \\ b^*_{k_n} \rightharpoonup y^*, \\ b_{k_n} - La_{k_n} \to 0, \\ a^*_{k_n} + L^* b^*_{k_n} \to 0, \\ (a_{k_n}, a^*_{k_n}) \in \operatorname{gra} A, \\ (b_{k_n}, b^*_{k_n}) \in \operatorname{gra} B. \end{cases} \tag{3.48}$$

Ainsi, la Proposition 2.11(iii) et (3.38) impliquent que $(x, y^*) \in \boldsymbol{Z} \cap \overline{\operatorname{dom} \boldsymbol{f}}$, i.e., $\mathfrak{W}(\boldsymbol{x}_n)_{n\in\mathbb{N}} \subset \boldsymbol{Z} \cap \overline{\operatorname{dom} \boldsymbol{f}}$. La convergence résulte donc de la Proposition 2.14(vii) et de (3.36). ☐

Remarque 3.14 Comparons le Théorème 2.24 et la Proposition 3.13. Dans les algorithmes (2.65) et (3.37), à chaque itération $n \in \mathbb{N}$, les opérateurs monotones A et B sont activés via les opérateurs $(\nabla h + \gamma_n A)^{-1}$ et $(\nabla j + \mu_n B)^{-1}$ pour construire $(a_n, a^*_n) \in \operatorname{gra} A$ et $(b_n, b^*_n) \in \operatorname{gra} B$. Dans l'algorithme (2.65), ces points sont déterminés par une résolution du problème

$$\text{trouver } (a, a^*) \in \operatorname{gra} A \quad \text{tel que} \quad 0 = \nabla h(a) - \nabla h(x_n) + \gamma_n L^* y^*_n + \gamma_n a^*, \tag{3.49}$$

et

$$\text{trouver } (b, b^*) \in \operatorname{gra} B \quad \text{tel que} \quad 0 = \nabla j(b) - \nabla j(Lx_n) + \mu_n y^*_n + \mu_n b^*. \tag{3.50}$$

Notons que (3.49) et (3.50) peuvent être difficiles à résoudre exactement. Par conséquent, les résolutions approchées sont considérées dans l'algorithme (3.37), i.e., on cherche une solution (a_n, a^*_n) du problème

$$\text{trouver } (a, a^*) \in \operatorname{gra} A \quad \text{tel que} \quad c^*_n = \nabla h(a) - \nabla h(x_n) + \gamma_n L^* y^*_n + \gamma_n a^*, \tag{3.51}$$

et une solution (b_n, b^*_n) du problème

$$\text{trouver } (b, b^*) \in \operatorname{gra} B \quad \text{tel que} \quad d^*_n = \nabla j(b) - \nabla j(Lx_n) + \mu_n y^*_n + \mu_n b^*. \tag{3.52}$$

Remarque 3.15 Dans la Proposition 3.13, supposons que $\mathcal{X}$ et $\mathcal{Y}$ soient hilbertiens, que $f = h = \|\cdot\|^2/2$, et que $g = j = \|\cdot\|^2/2$. Les suites d'erreurs $(c^*_n)_{n\in\mathbb{N}}$ et $(d^*_n)_{n\in\mathbb{N}}$ peuvent être choisies suivant l'une des règles suivantes.

(i) $(\forall n \in \mathbb{N}) \begin{cases} \|c^*_n\| \leqslant \theta \max\{\|x_n - a_n\|, \gamma_n \|a^*_n + L^* y^*_n\|\}; \\ \|d^*_n\| \leqslant \theta \max\{\|Lx_n - b_n\|, \mu_n \|b^*_n - y^*_n\|\}, \end{cases}$ où $\theta \in [0, 1[$.
Ce type d'erreurs est utilisé dans [9].

(ii) $(\forall n \in \mathbb{N})\ \gamma_n^{-2}\|c_n^*\|^2 + \mu_n^{-2}\|d_n^*\|^2 \leqslant \theta\big(\|x_n - a_n\|^2 + \|Lx_n - b_n\|^2\big)$ où $\theta \in [0, \sigma^{-1}[$. On récupère le résultat de [6] pour le cas de somme de deux opérateurs monotones et $L = \mathrm{Id}$.

Démonstration. Il suffit de montrer que $c_n^* \to 0$ et $d_n^* \to 0$.

(i) : D'abord, on montre que

$$(\forall n \in \mathbb{N}) \quad \big\langle x_n - a_n, a_n^* + L^*y_n^* \big\rangle \geqslant (1-\theta)\max\{\gamma_n^{-1}\|x_n - a_n\|^2, \gamma_n\|a_n^* + L^*y_n^*\|^2\}. \quad (3.53)$$

On fixe $n \in \mathbb{N}$ et on considère deux cas.

(a) Si $\|x_n - a_n\| \leqslant \gamma_n\|a_n^* + L^*y_n^*\|$ alors $\|c_n^*\| \leqslant \theta\gamma_n\|a_n^* + L^*b_n * \|$. Par suite, on déduit de (3.37) que

$$\begin{aligned}\big\langle x_n - a_n, a_n^* + L^*y_n^* \big\rangle &= \big\langle \gamma_n(a_n^* + L^*y_n^*) - c_n^*, a_n^* + L^*y_n^* \big\rangle \\ &= \gamma_n\|a_n^* + L^*y_n^*\|^2 - \big\langle c_n^*, a_n^* + L^*y_n^* \big\rangle \\ &\geqslant \gamma_n(1-\theta)\|a_n^* + L^*y_n^*\|^2 \\ &= (1-\theta)\max\{\gamma_n^{-1}\|x_n - a_n\|^2, \gamma_n\|a_n^* + L^*y_n^*\|^2\}. \quad (3.54)\end{aligned}$$

(b) Si $\|x_n - a_n\| \geqslant \gamma_n\|a_n^* + L^*y_n^*\|$ alors $\|c_n^*\| \leqslant \theta\|x_n - a_n\|$. Par suite, on déduit de (3.37) que

$$\begin{aligned}\big\langle x_n - a_n, a_n^* + L^*y_n^* \big\rangle &= \gamma_n^{-1}\big\langle x_n - a_n, x_n - a_n + c_n^* \big\rangle \\ &= \gamma_n^{-1}\|x_n - a_n\|^2 + \gamma_n^{-1}\big\langle x_n - a_n, c_n^* \big\rangle \\ &\geqslant \gamma_n^{-1}(1-\theta)\|x_n - a_n\|^2 \\ &= (1-\theta)\max\{\gamma_n^{-1}\|x_n - a_n\|^2, \gamma_n\|a_n^* + L^*y_n^*\|^2\}. \quad (3.55)\end{aligned}$$

Par le même raisonnement,

$$(\forall n \in \mathbb{N}) \quad \big\langle Lx_n - b_n, b_n^* - y_n^* \big\rangle \geqslant (1-\theta)\max\{\mu_n^{-1}\|Lx_n - b_n\|^2, \mu_n\|b_n^* - y_n^*\|^2\}. \quad (3.56)$$

Il résulte de (3.40), de (3.53), et de (3.56) que

$$\begin{aligned}(\forall n \in \mathbb{N}) \quad & \|x_n - x_{n+1/2}\|\,\|a_n^* + L^*b_n^*\| + \|b_n - La_n\|\,\|y_n^* - y_{n+1/2}^*\| \\ &\geqslant \big\langle x_n - x_{n+1/2}, a_n^* + L^*b_n^* \big\rangle + \big\langle b_n - La_n, y_n^* - y_{n+1/2}^* \big\rangle \\ &\geqslant \big\langle x_n, a_n^* + L^*b_n^* \big\rangle + \big\langle b_n - La_n, y_n^* \big\rangle - \big\langle a_n, a_n^* \big\rangle - \big\langle b_n, b_n^* \big\rangle \\ &= \big\langle x_n - a_n, a_n^* + L^*y_n^* \big\rangle + \big\langle Lx_n - b_n, b_n^* - y_n^* \big\rangle \\ &= (1-\theta)\max\{\gamma_n^{-1}\|x_n - a_n\|^2, \gamma_n\|a_n^* + L^*y_n^*\|^2\} \\ &\quad + (1-\theta)\max\{\mu_n^{-1}\|Lx_n - b_n\|^2, \mu_n\|b_n^* - y_n^*\|^2\}. \quad (3.57)\end{aligned}$$

Par suite, (3.38) et (3.39) impliquent que

$$\max\{\gamma_n^{-1}\|x_n - a_n\|^2, \gamma_n\|a_n^* + L^*y_n^*\|^2\} + \max\{\mu_n^{-1}\|Lx_n - b_n\|^2, \mu_n\|b_n^* - y_n^*\|^2\} \to 0. \quad (3.58)$$

Il en découle que

$$\begin{cases} \gamma_n^{-1}\|x_n - a_n\|^2 \to 0, \\ \gamma_n\|a_n^* + L^*y_n^*\|^2 \to 0, \\ \mu_n^{-1}\|Lx_n - b_n\|^2 \to 0, \\ \mu_n\|b_n^* - y_n^*\|^2 \to 0. \end{cases} \tag{3.59}$$

Par conséquent, $c_n^* \to 0$ et $d_n^* \to 0$.

(ii) : On déduit de (3.40), de (3.53), et de (3.56) que

$$\begin{aligned}
(\forall n \in \mathbb{N}) \quad & \|x_n - x_{n+1/2}\| \, \|a_n^* + L^*b_n^*\| + \|b_n - La_n\| \, \|y_n^* - y_{n+1/2}^*\| \\
& \geqslant \langle x_n - x_{n+1/2}, a_n^* + L^*b_n^* \rangle + \langle b_n - La_n, y_n^* - y_{n+1/2}^* \rangle \\
& \geqslant \langle x_n, a_n^* + L^*b_n^* \rangle + \langle b_n - La_n, y_n^* \rangle - \langle a_n, a_n^* \rangle - \langle b_n, b_n^* \rangle \\
& = \langle x_n - a_n, a_n^* + L^*y_n^* \rangle + \langle Lx_n - b_n, b_n^* - y_n^* \rangle \\
& = \gamma_n^{-1}\langle x_n - a_n, x_n - a_n + c_n^* \rangle + \mu_n^{-1}\langle Lx_n - b_n, Lx_n - b_n + d_n^* \rangle \\
& \geqslant \gamma_n^{-1}\|x_n - a_n\|^2 - \gamma_n^{-1}\|c_n^*\| \, \|x_n - a_n\| \\
& \quad + \mu_n^{-1}\|Lx_n - b_n\|^2 - \mu_n^{-1}\|d_n^*\| \, \|Lx_n - b_n\| \\
& \geqslant \gamma_n^{-1}\|x_n - a_n\|^2 + \mu_n^{-1}\|Lx_n - b_n\|^2 \\
& \quad - \left(\gamma_n^{-1}\|c_n^*\| \, \|x_n - a_n\| + \mu_n^{-1}\|d_n^*\| \, \|Lx_n - b_n\|\right) \\
& \geqslant \sigma^{-1}\left(\|x_n - a_n\|^2 + \|Lx_n - b_n\|^2\right) \\
& \quad - \sqrt{\gamma_n^{-2}\|c_n^*\|^2 + \mu_n^{-2}\|d_n^*\|^2} \, \sqrt{\|x_n - a_n\|^2 + \|Lx_n - b_n\|^2} \\
& \geqslant (\sigma^{-1} - \theta)\left(\|x_n - a_n\|^2 + \|Lx_n - b_n\|^2\right).
\end{aligned} \tag{3.60}$$

Ainsi $\|x_n - a_n\| \to 0$ et $\|Lx_n - b_n\| \to 0$. Par conséquent, $c_n^* \to 0$ et $d_n^* \to 0$. □

3.4 Bibliographie

[1] H. H. Bauschke, J. M. Borwein, and P. L. Combettes, Essential smoothness, essential strict convexity and Legendre functions in Banach spaces, *Commun. Contemp. Math.,* vol. 3, pp. 615–647, 2001.

[2] H. H. Bauschke, J. M. Borwein, and P. L. Combettes, Bregman monotone optimization algorithms, *SIAM J. Control Optim.,* vol. 42, pp. 596–636, 2003.

[3] H. H. Bauschke and P. L. Combettes, A weak-to-strong convergence principle for Fejér-monotone methods in Hilbert spaces, *Math. Oper. Res.,* vol. 26, pp. 248–264, 2001.

[4] H. H. Bauschke and P. L. Combettes, Construction of best Bregman approximations in reflexive Banach spaces, *Proc. Amer. Math. Soc.,* vol. 131, pp. 3757–3766, 2003.

[5] P. L. Combettes, Strong convergence of block-iterative outer approximation methods for convex optimization, *SIAM J. Control Optim.,* vol. 38, pp. 538–565, 2000.

[6] J. Eckstein and B. F. Svaiter, General projective splitting methods for sums of maximal monotone operators, *SIAM J. Control Optim.,* vol. 48, pp. 787–811, 2009.

[7] Y. Haugazeau, *Sur les Inéquations Variationnelles et la Minimisation de Fonctionnelles Convexes.* Thèse, Université de Paris, Paris, France, 1968.

[8] S. Reich and S. Sabach, Two strong convergence theorems for a proximal method in reflexive Banach spaces, *Numer. Funct. Anal. Optim.,* vol. 31, pp. 22–44, 2010.

[9] M. V. Solodov and B. F. Svaiter, Forcing strong convergence of proximal point iterations in a Hilbert space, *Math. Program.,* vol. 87, pp. 189–202, 2000.

Chapitre 4

Approximations bregmaniennes successives de l'ensemble de Kuhn-Tucker d'inclusions monotones

Dans le Théorème 2.5, (x_0, y_0^*) est un point de référence duquel on veut s'éloigner au minimum (par exemple parce que c'est la solution d'un problème voisin). Dans ce chapitre, on cherche un point de Kuhn-Tucker quelconque. À chaque itération, les opérateurs monotones sont activés via leurs résolvantes au sens des distances de Bregman pour construire des points dans leurs graphes. Ces points sont utilisés pour construire un demi-espace contenant l'ensemble de Kuhn-Tucker associé au système. La mise à jour primale-duale est ensuite obtenue via une projection relaxée, basée sur les distances de Bregman, de l'itérée courante sur ce demi-espace. Ce travail étend la méthode proposée dans [1] qui utilise la métrique classique dans les espaces hilbertiens.

4.1 Introduction et notations

Dans ce chapitre, nous considérons le problème suivant.

Problème 4.1 Soient $\mathcal{X}$ et $\mathcal{Y}$ des espaces de Banach réels réflexifs tels que $\mathcal{X} \neq \{0\}$ et $\mathcal{Y} \neq \{0\}$, soit $\boldsymbol{\mathcal{X}}$ l'espace vectoriel $\mathcal{X} \times \mathcal{Y}^*$ muni de la norme $(x, y^*) \mapsto \sqrt{\|x\|^2 + \|y^*\|^2}$, soit $\boldsymbol{\mathcal{X}}^*$ le dual topologique de $\boldsymbol{\mathcal{X}}$, i.e., $\mathcal{X}^* \times \mathcal{Y}$ muni de la norme $(x^*, y) \mapsto \sqrt{\|x^*\|^2 + \|y\|^2}$, soient $A\colon \mathcal{X} \to 2^{\mathcal{X}^*}$ et $B\colon \mathcal{Y} \to 2^{\mathcal{Y}^*}$ maximalement monotones, et soit $L \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$. Considérons l'inclusion primale

$$\text{trouver } x \in \mathcal{X} \text{ tel que } 0 \in Ax + L^*BLx, \tag{4.1}$$

l'inclusion duale

$$\text{trouver } y^* \in \mathcal{Y}^* \text{ tel que } 0 \in -LA^{-1}(-L^*y^*) + B^{-1}y^*, \tag{4.2}$$

et désignons par

$$\boldsymbol{Z} = \big\{(x, y^*) \in \boldsymbol{\mathcal{X}} \;\big|\; -L^*y^* \in Ax \quad \text{et} \quad Lx \in B^{-1}y^*\big\} \tag{4.3}$$

l'ensemble de Kuhn-Tucker y associé. Le problème est de trouver un élément de $\boldsymbol{Z}$. Les ensembles des solutions de (4.1) et de (4.2) sont notés $\mathscr{P}$ et $\mathscr{D}$, respectivement.

Nous avons montré, dans la Proposition 2.11 du Chapitre 2, que $\boldsymbol{Z}$ est un sous-ensemble convexe fermé de $\mathscr{P} \times \mathscr{D}$ et donc un point dans cet ensemble donne une solution primale-duale du Problème 4.1. De plus, une telle solution est obtenue dans le Problème 2.2 en cherchant dans $\boldsymbol{Z}$ la meilleure approximation basée sur une distance de Bregman d'un point de référence. Dans ce chapitre, nous proposons une méthode pour trouver un point a priori arbitraire dans $\boldsymbol{Z}$ en utilisant les projections successives sur les demi-espaces séparant l'itérée courante et $\boldsymbol{Z}$. Ce résultat unifie et étend les travaux de [4] où $L = \mathrm{Id}$ et les travaux de [1], où $\mathcal{X}$ et $\mathcal{Y}$ sont hilbertiens.

Notation 4.2 Les symboles $\rightharpoonup$ et $\to$ représentent la convergence faible et forte, respectivement. L'ensemble de tous points faiblement adhérents d'une suite $(x_n)_{n\in\mathbb{N}}$ est noté $\mathfrak{W}(x_n)_{n\in\mathbb{N}}$. Une fonction $f\colon \mathcal{X} \to \,]-\infty, +\infty]$ est coercive si $\lim_{\|x\|\to+\infty} f(x) = +\infty$ et supercoercive si $\lim_{\|x\|\to+\infty} f(x)/\|x\| = +\infty$. Soit $f \in \Gamma_0(\mathcal{X})$ une fonction *de Legendre* (i.e. f est *essentiellement lisse* au sens où ∂f est à la fois localement borné et univoque sur son domaine, et *essentiellement strictement convexe* au sens où ∂f^* est localement borné sur son domaine et f est strictement convexe sur les sous-ensembles convexes de $\operatorname{dom}\partial f$). Alors la distance de Bregman associée à f est

$$\begin{aligned} D^f\colon \mathcal{X} \times \mathcal{X} &\to [0, +\infty] \\ (x, y) &\mapsto \begin{cases} f(x) - f(y) - \big\langle x - y, \nabla f(y)\big\rangle, & \text{si } \; y \in \operatorname{int}\operatorname{dom} f; \\ +\infty, & \text{sinon.} \end{cases} \end{aligned} \tag{4.4}$$

De plus, soit $g \in \Gamma_0(\mathcal{X})$ Gâteaux différentiable sur $\operatorname{int}\operatorname{dom} g \supset \operatorname{int}\operatorname{dom} f$. Alors

$$f \succcurlyeq g \quad \Leftrightarrow \quad (\forall x \in \operatorname{dom} f)(\forall y \in \operatorname{int}\operatorname{dom} f) \quad D^f(x, y) \geqslant D^g(x, y). \tag{4.5}$$

Un opérateur $M\colon \mathcal{X} \to 2^{\mathcal{X}^*}$ est *coercif* si

$$(\exists\, z \in \operatorname{dom} M) \qquad \lim_{\|x\|\to+\infty} \inf \frac{\langle x - z, Mx\rangle}{\|x\|} = +\infty. \tag{4.6}$$

4.2 Projections successives basées sur une distance de Bregman

Soit $\boldsymbol{f} \in \Gamma_0(\boldsymbol{\mathcal{X}})$ une fonction de Legendre et soit $\boldsymbol{C} \subset \boldsymbol{\mathcal{X}}$ un sous-ensemble convexe fermé tel que $\boldsymbol{C} \cap \operatorname{int}\operatorname{dom}\boldsymbol{f} \neq \varnothing$. Alors le $D^{\boldsymbol{f}}$*-projecteur* sur $\boldsymbol{C}$ est

$$\begin{aligned} P_{\boldsymbol{C}}^{\boldsymbol{f}} \colon \operatorname{int}\operatorname{dom}\boldsymbol{f} &\to \boldsymbol{C} \cap \operatorname{int}\operatorname{dom}\boldsymbol{f} \\ \boldsymbol{y} &\mapsto \underset{\boldsymbol{x} \in \boldsymbol{C}}{\operatorname{argmin}}\, D^{\boldsymbol{f}}(\boldsymbol{x}, \boldsymbol{y}). \end{aligned} \tag{4.7}$$

Par ailleurs, [3, Proposition 3.32 et Corollary 3.35] affirment que $\operatorname{Fix} P_{\boldsymbol{C}}^{\boldsymbol{f}} = \boldsymbol{C} \cap \operatorname{int}\operatorname{dom}\boldsymbol{f}$ et que

$$(\forall \boldsymbol{x} \in \boldsymbol{C} \cap \operatorname{int}\operatorname{dom}\boldsymbol{f})(\forall \boldsymbol{y} \in \operatorname{int}\operatorname{dom}\boldsymbol{f}) \quad D^{\boldsymbol{f}}\big(\boldsymbol{x}, P_{\boldsymbol{C}}^{\boldsymbol{f}}\boldsymbol{y}\big) + D^{\boldsymbol{f}}\big(P_{\boldsymbol{C}}^{\boldsymbol{f}}\boldsymbol{y}, \boldsymbol{y}\big) \leqslant D^{\boldsymbol{f}}(\boldsymbol{x}, \boldsymbol{y}). \tag{4.8}$$

Exemple 4.3 Soit $\eta \in \left]0, +\infty\right[$, soit $\boldsymbol{s}^* \in \boldsymbol{\mathcal{X}}^* \setminus \{0\}$, soit $\boldsymbol{C} = \big\{\boldsymbol{x} \in \boldsymbol{\mathcal{X}} \;\big|\; \langle \boldsymbol{x}, \boldsymbol{s}^* \rangle \leqslant \eta\big\}$, et soit $\boldsymbol{f} \in \Gamma_0(\boldsymbol{\mathcal{X}})$ une fonction de Legendre telle que $\boldsymbol{C} \cap \operatorname{int}\operatorname{dom}\boldsymbol{f} \neq \varnothing$. Il résulte de (4.7) que $P_{\boldsymbol{C}}^{\boldsymbol{f}} \colon \operatorname{int}\operatorname{dom}\boldsymbol{f} \to \boldsymbol{C} \cap \operatorname{int}\operatorname{dom}\boldsymbol{f}$. Maintenant, soit $\boldsymbol{y} \in \operatorname{int}\operatorname{dom}\boldsymbol{f}$. Si $\boldsymbol{y} \in \boldsymbol{C}$ alors $P_{\boldsymbol{C}}^{\boldsymbol{f}}\boldsymbol{y} = \boldsymbol{y}$. Sinon, $\langle \boldsymbol{y}, \boldsymbol{s}^* \rangle > \eta$. Soit $\operatorname{bdry}\boldsymbol{C} = \big\{\boldsymbol{x} \in \boldsymbol{\mathcal{X}} \;\big|\; \langle \boldsymbol{x}, \boldsymbol{s}^* \rangle = \eta\big\}$ la frontière de $\boldsymbol{C}$. Comme $\boldsymbol{C} \cap \operatorname{int}\operatorname{dom}\boldsymbol{f} \neq \varnothing$ et $(\boldsymbol{\mathcal{X}} \setminus \boldsymbol{C}) \cap \operatorname{int}\operatorname{dom}\boldsymbol{f} = \{\boldsymbol{y}\} \neq \varnothing$, on obtient $(\operatorname{bdry}\boldsymbol{C}) \cap \operatorname{int}\operatorname{dom}\boldsymbol{f} \neq \varnothing$. On alors déduit de [3, Corollary 3.35(ii)] que

$$\begin{aligned} &\boldsymbol{z} = P_{\operatorname{bdry}\boldsymbol{C}}^{\boldsymbol{f}}\boldsymbol{y} \Leftrightarrow [\boldsymbol{z} \in \operatorname{bdry}\boldsymbol{C}] \text{ et } [\operatorname{bdry}\boldsymbol{C} \subset H^{\boldsymbol{f}}(\boldsymbol{y}, \boldsymbol{z})] \\ &\Leftrightarrow [\boldsymbol{z} \in \operatorname{bdry}\boldsymbol{C}] \text{ et } [(\exists \theta \in \mathbb{R})\ \theta \boldsymbol{s}^* = \nabla \boldsymbol{f}(\boldsymbol{y}) - \nabla \boldsymbol{f}(\boldsymbol{z})] \\ &\Leftrightarrow [(\exists \theta \in \mathbb{R})\ \big\langle \nabla \boldsymbol{f}^*(\nabla \boldsymbol{f}(\boldsymbol{y}) - \theta \boldsymbol{s}^*), \boldsymbol{s}^* \big\rangle = \eta] \text{ et } [\boldsymbol{z} = \nabla \boldsymbol{f}^*(\nabla \boldsymbol{f}(\boldsymbol{y}) - \theta \boldsymbol{s}^*)]. \end{aligned} \tag{4.9}$$

Ensuite, supposons que $\boldsymbol{z}$ et θ satisfassent (4.9). Alors, il résulte de la monotonie de $\nabla \boldsymbol{f}$ que

$$\theta\big(\langle \boldsymbol{y}, \boldsymbol{s}^* \rangle - \langle \boldsymbol{z}, \boldsymbol{s}^* \rangle\big) = \big\langle \boldsymbol{y} - \boldsymbol{z}, \nabla \boldsymbol{f}(\boldsymbol{y}) - \nabla \boldsymbol{f}(\boldsymbol{z}) \big\rangle \geqslant 0, \tag{4.10}$$

et comme $\langle \boldsymbol{y}, \boldsymbol{s}^* \rangle - \langle \boldsymbol{z}, \boldsymbol{s}^* \rangle > 0$, on obtient $\theta > 0$. De plus,

$$(\forall \boldsymbol{x} \in \boldsymbol{C}) \quad \big\langle \boldsymbol{x} - \boldsymbol{z}, \boldsymbol{f}(\boldsymbol{y}) - \nabla \boldsymbol{f}(\boldsymbol{z}) \big\rangle = \theta \,\langle \boldsymbol{x} - \boldsymbol{y}, \boldsymbol{s}^* \rangle = \theta\big(\langle \boldsymbol{x}, \boldsymbol{s}^* \rangle - \langle \boldsymbol{y}, \boldsymbol{s}^* \rangle\big) < 0, \tag{4.11}$$

et donc $\boldsymbol{z} \in \operatorname{bdry}\boldsymbol{C} \subset \boldsymbol{C} \subset H^{\boldsymbol{f}}(\boldsymbol{y}, \boldsymbol{z})$. Il donc résulte de [3, Corollary 3.35(ii)] que $P_{\boldsymbol{C}}^{\boldsymbol{f}}\boldsymbol{y} = \boldsymbol{z} = \nabla \boldsymbol{f}^*(\nabla \boldsymbol{f}(\boldsymbol{y}) - \theta \boldsymbol{s}^*)$, où $\theta \in \left]0, +\infty\right[$ vérifie $\big\langle \nabla \boldsymbol{f}^*(\nabla \boldsymbol{f}(\boldsymbol{y}) - \theta \boldsymbol{s}^*), \boldsymbol{s}^* \big\rangle = \eta$.

Le résultat suivant considère le problème de trouver un point dans un sous-ensemble convexe fermé en utilisant les projections successives basées sur une distance de Bregman sur des demi-espaces qui le contiennent. Ce résultat est démontré sous d'autres hypothèses dans [3, Section 4.2].

Proposition 4.4 *Soit $\boldsymbol{f} \in \Gamma_0(\boldsymbol{\mathcal{X}})$ une fonction de Legendre, soit $\boldsymbol{C}$ un sous-ensemble convexe fermé de $\boldsymbol{\mathcal{X}}$ tel que $\boldsymbol{C} \cap \operatorname{int}\operatorname{dom}\boldsymbol{f} \neq \varnothing$, soit $\varepsilon \in \left]0,1\right[$, et soit $\boldsymbol{x}_0 \in \operatorname{int}\operatorname{dom}\boldsymbol{f}$. À chaque itération $n \in \mathbb{N}$, on cherche un demi-espace affine fermé $\boldsymbol{H}_n \subset \boldsymbol{\mathcal{X}}$ tel que $\boldsymbol{C} \subset \boldsymbol{H}_n$, on prend $\lambda_n \in \left]\varepsilon, 1\right]$, et on itère*

$$\boldsymbol{x}_{n+1} = \nabla \boldsymbol{f}^*\Big((1-\lambda_n)\nabla \boldsymbol{f}(\boldsymbol{x}_n) + \lambda_n \nabla \boldsymbol{f}\big(P^{\boldsymbol{f}}_{\boldsymbol{H}_n}\boldsymbol{x}_n\big)\Big). \tag{4.12}$$

Supposons que $(\forall \boldsymbol{x} \in \operatorname{int}\operatorname{dom}\boldsymbol{f})$ $D^{\boldsymbol{f}}(\boldsymbol{x},\cdot)$ soit coercive. Alors les propriétés suivantes sont satisfaites :

(i) *Soit $\boldsymbol{x} \in \boldsymbol{C} \cap \operatorname{int}\operatorname{dom}\boldsymbol{f}$. Alors $(D^{\boldsymbol{f}}(\boldsymbol{x},\boldsymbol{x}_n))_{n\in\mathbb{N}}$ est bornée.*

(ii) *$(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ est une suite bornée dans $\operatorname{int}\operatorname{dom}\boldsymbol{f}$.*

(iii) *$\sum_{n\in\mathbb{N}} D^{\boldsymbol{f}}(P^{\boldsymbol{f}}_{\boldsymbol{H}_n}\boldsymbol{x}_n, \boldsymbol{x}_n) < +\infty$.*

(iv) *Supposons que pour tout $\boldsymbol{y}_1 \in \boldsymbol{\mathcal{X}}$ et pour tout $\boldsymbol{y}_2 \in \boldsymbol{\mathcal{X}}$,*

$$\begin{cases} \boldsymbol{y}_1 \in \mathfrak{W}(\boldsymbol{x}_n)_{n\in\mathbb{N}} \cap \boldsymbol{C} \\ \boldsymbol{y}_2 \in \mathfrak{W}(\boldsymbol{x}_n)_{n\in\mathbb{N}} \cap \boldsymbol{C} \\ \big(\langle \boldsymbol{y}_1 - \boldsymbol{y}_2, \nabla \boldsymbol{f}(\boldsymbol{x}_n)\rangle\big)_{n\in\mathbb{N}} \quad \textit{converge} \end{cases} \quad \Rightarrow \quad \boldsymbol{y}_1 = \boldsymbol{y}_2. \tag{4.13}$$

De plus, supposons que $\mathfrak{W}(\boldsymbol{x}_n)_{n\in\mathbb{N}} \subset \boldsymbol{C} \cap \operatorname{int}\operatorname{dom}\boldsymbol{f}$. Alors $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ converge vers un point dans $\boldsymbol{C} \cap \operatorname{int}\operatorname{dom}\boldsymbol{f}$.

Démonstration. Notons d'abord que

$$(\forall n \in \mathbb{N}) \quad \varnothing \neq \boldsymbol{C} \cap \operatorname{int}\operatorname{dom}\boldsymbol{f} \subset \boldsymbol{H}_n \cap \operatorname{int}\operatorname{dom}\boldsymbol{f}, \tag{4.14}$$

et il donc résulte de (4.7) que

$$(\forall n \in \mathbb{N}) \quad P^{\boldsymbol{f}}_{\boldsymbol{H}_n} \colon \operatorname{int}\operatorname{dom}\boldsymbol{f} \to \boldsymbol{H}_n \cap \operatorname{int}\operatorname{dom}\boldsymbol{f}. \tag{4.15}$$

Par ailleurs, d'après [2, Theorem 5.10], $\nabla \boldsymbol{f} \colon \operatorname{int}\operatorname{dom}\boldsymbol{f} \to \operatorname{int}\operatorname{dom}\boldsymbol{f}^*$ est bijective avec pour inverse $\nabla \boldsymbol{f}^* \colon \operatorname{int}\operatorname{dom}\boldsymbol{f}^* \to \operatorname{int}\operatorname{dom}\boldsymbol{f}$. Par suite, en utilisant le fait que $\boldsymbol{x}_0 \in \operatorname{int}\operatorname{dom}\boldsymbol{f}$, (4.12), (4.15), et le raisonnement par récurrence, on déduit que

$$(\boldsymbol{x}_n)_{n\in\mathbb{N}} \in (\operatorname{int}\operatorname{dom}\boldsymbol{f})^{\mathbb{N}} \quad \text{est bien définie.} \tag{4.16}$$

Ensuite,on déduit de (4.12), du Lemme 7.6, et de (4.8) que

$$\begin{aligned} &(\forall \boldsymbol{x} \in \boldsymbol{C} \cap \operatorname{dom}\boldsymbol{f})(\forall n \in \mathbb{N}) \\ &\qquad D^{\boldsymbol{f}}(\boldsymbol{x},\boldsymbol{x}_{n+1}) \leqslant (1-\lambda_n)D^{\boldsymbol{f}}(\boldsymbol{x},\boldsymbol{x}_n) + \lambda_n D^{\boldsymbol{f}}\big(\boldsymbol{x}, P^{\boldsymbol{f}}_{\boldsymbol{H}_n}\boldsymbol{x}_n\big) \\ &\qquad\qquad \leqslant (1-\lambda_n)D^{\boldsymbol{f}}(\boldsymbol{x},\boldsymbol{x}_n) + \lambda_n\big(D^{\boldsymbol{f}}(\boldsymbol{x},\boldsymbol{x}_n) - D^{\boldsymbol{f}}\big(P^{\boldsymbol{f}}_{\boldsymbol{H}_n}\boldsymbol{x}_n,\boldsymbol{x}_n\big)\big) \\ &\qquad\qquad = D^{\boldsymbol{f}}(\boldsymbol{x},\boldsymbol{x}_n) - \lambda_n D^{\boldsymbol{f}}\big(P^{\boldsymbol{f}}_{\boldsymbol{H}_n}\boldsymbol{x}_n,\boldsymbol{x}_n\big) \\ &\qquad\qquad \leqslant D^{\boldsymbol{f}}(\boldsymbol{x},\boldsymbol{x}_n). \end{aligned} \tag{4.17}$$

(i)–(ii) : Soit $\boldsymbol{x} \in \boldsymbol{C} \cap \operatorname{int} \operatorname{dom} \boldsymbol{f}$. Comme (4.17) implique que $(D^{\boldsymbol{f}}(\boldsymbol{x}, \boldsymbol{x}_n))_{n\in\mathbb{N}}$ est décroissante, elle converge. Ainsi $(D^{\boldsymbol{f}}(\boldsymbol{x}, \boldsymbol{x}_n))_{n\in\mathbb{N}}$ est bornée. Par suite, comme $D^{\boldsymbol{f}}(\boldsymbol{x}, \cdot)$ est coercive, $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ est bornée.

(iii) : Soit $\boldsymbol{x} \in \boldsymbol{C} \cap \operatorname{int} \operatorname{dom} \boldsymbol{f}$. On déduit de (4.17) que

$$\sum_{n\in\mathbb{N}} D^{\boldsymbol{f}}\big(P^{\boldsymbol{f}}_{\boldsymbol{H}_n}\boldsymbol{x}_n, \boldsymbol{x}_n\big) \leqslant \rho^{-1} \sum_{n\in\mathbb{N}} \big(D^{\boldsymbol{f}}(\boldsymbol{x}, \boldsymbol{x}_n) - D^{\boldsymbol{f}}(\boldsymbol{x}, \boldsymbol{x}_{n+1})\big) \leqslant \rho^{-1} D^{\boldsymbol{f}}(\boldsymbol{x}, \boldsymbol{x}_0). \tag{4.18}$$

(iv) : Comme $\boldsymbol{\mathcal{X}}$ est réflexif, on déduit de (ii) que $\mathfrak{W}(\boldsymbol{x}_n)_{n\in\mathbb{N}} \neq \varnothing$. Pour montrer la convergence, il suffit de montrer que $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ admet au plus un séquentiellement faiblement adhérent. On suppose que $\boldsymbol{y}_1 \in \mathfrak{W}(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ et $\boldsymbol{y}_2 \in \mathfrak{W}(\boldsymbol{x}_n)_{n\in\mathbb{N}}$. Alors $\boldsymbol{y}_1$ et $\boldsymbol{y}_2$ sont dans $\boldsymbol{C} \cap \operatorname{int} \operatorname{dom} \boldsymbol{f}$. Par suite, il résulte de (4.17) que $(D^{\boldsymbol{f}}(\boldsymbol{y}_1, \boldsymbol{x}_n))_{n\in\mathbb{N}}$ et $(D^{\boldsymbol{f}}(\boldsymbol{y}_2, \boldsymbol{x}_n))_{n\in\mathbb{N}}$ convergent. En outre, comme

$$(\forall n \in \mathbb{N}) \quad \big\langle \boldsymbol{y}_1 - \boldsymbol{y}_2, \nabla \boldsymbol{f}(\boldsymbol{x}_n)\big\rangle = \boldsymbol{f}(\boldsymbol{y}_1) - \boldsymbol{f}(\boldsymbol{y}_2) - D^{\boldsymbol{f}}(\boldsymbol{y}_1, \boldsymbol{y}_n) + D^{\boldsymbol{f}}(\boldsymbol{y}_2, \boldsymbol{x}_n), \tag{4.19}$$

on déduit que $(\langle \boldsymbol{y}_1 - \boldsymbol{y}_2, \nabla \boldsymbol{f}(\boldsymbol{x}_n)\rangle))_{n\in\mathbb{N}}$ converge et donc (4.13) implique que $\boldsymbol{y}_1 = \boldsymbol{y}_2$. $\square$

Remarque 4.5 La Proposition 4.4 requiert l'évaluation de la projection au sens d'une distance de Bregman sur des demi-espaces. Au vu de l'Exemple 4.3, ceci revient à résoudre une équation dans $]0, +\infty[$.

Remarque 4.6 Soient $\boldsymbol{\mathcal{X}}$ un espace de Banach réel réflexif, $\boldsymbol{f} \in \Gamma_0(\boldsymbol{\mathcal{X}})$ Gâteaux différentiable sur $\operatorname{int} \operatorname{dom} \boldsymbol{f} \neq \varnothing$, $(\boldsymbol{x}_n)_{n\in\mathbb{N}} \in (\operatorname{int} \operatorname{dom} \boldsymbol{f})^{\mathbb{N}}$, et $\boldsymbol{C} \subset \boldsymbol{\mathcal{X}}$. Supposons que

$$(\forall \boldsymbol{x} \in \boldsymbol{C} \cap \operatorname{dom} \boldsymbol{f})(\forall n \in \mathbb{N}) \quad D^{\boldsymbol{f}}(\boldsymbol{x}, \boldsymbol{x}_{n+1}) \leqslant D^{\boldsymbol{f}}(\boldsymbol{x}, \boldsymbol{x}_n). \tag{4.20}$$

De plus, supposons qu'une des conditions suivantes soit satisfaite :

(i) $\boldsymbol{C} \cap \overline{\operatorname{dom}} \boldsymbol{f}$ est un singleton.

(ii) $\boldsymbol{C} \subset \operatorname{int} \operatorname{dom} \boldsymbol{f}$ est convexe, $\boldsymbol{f}|_{\boldsymbol{C}}$ est strictement convexe, et $\nabla \boldsymbol{f}|_{\boldsymbol{C}}$ est faiblement séquentiellement continu.

(iii) $\boldsymbol{\mathcal{X}}$ est de dimension finie, $\boldsymbol{f}$ est une fonction de Legendre telle que $\operatorname{dom} \boldsymbol{f}^*$ est ouvert, et $\boldsymbol{C} \cap \operatorname{dom} \boldsymbol{f} \neq \varnothing$.

(iv) $\mathfrak{W}(\boldsymbol{x}_n)_{n\in\mathbb{N}} \subset \operatorname{int} \operatorname{dom} \boldsymbol{f}$, $\boldsymbol{f}|_{\operatorname{int} \operatorname{dom} \boldsymbol{f}}$ est strictement convexe, et $\nabla \boldsymbol{f}$ est faiblement séquentiellement continu.

(v) $\boldsymbol{C} \cap \operatorname{int} \operatorname{dom} \boldsymbol{f} \neq \varnothing$ et $\boldsymbol{f}$ est une fonction de Legendre telle que $\operatorname{dom} \boldsymbol{f}^*$ est ouvert et que $\nabla \boldsymbol{f}$ et $\nabla \boldsymbol{f}^*$ sont faiblement séquentiellement continus.

Alors (4.13) est vérifiée.

Démonstration. (i)–(iii) : [3, Examples 4.6, 4.7, 4.9].

(iv) : Soient $\boldsymbol{y}_1 \in \mathfrak{W}(\boldsymbol{x}_n)_{n\in\mathbb{N}} \cap \boldsymbol{C}$ et $\boldsymbol{y}_2 \in \mathfrak{W}(\boldsymbol{x}_n)_{n\in\mathbb{N}} \cap \boldsymbol{C}$, i.e., $\boldsymbol{x}_{k_n} \rightharpoonup \boldsymbol{y}_1$ et $\boldsymbol{x}_{l_n} \rightharpoonup \boldsymbol{y}_2$. Alors (4.20) implique que $(D^{\boldsymbol{f}}(\boldsymbol{y}_1, \boldsymbol{x}_n))_{n\in\mathbb{N}}$ et $(D^{\boldsymbol{f}}(\boldsymbol{y}_2, \boldsymbol{x}_n))_{n\in\mathbb{N}}$ sont convergentes. Comme

$$(\forall n \in \mathbb{N}) \quad \big\langle \boldsymbol{y}_1 - \boldsymbol{y}_2, \nabla \boldsymbol{f}(\boldsymbol{x}_n)\big\rangle = D^{\boldsymbol{f}}(\boldsymbol{y}_1, \boldsymbol{x}_n) - D^{\boldsymbol{f}}(\boldsymbol{y}_2, \boldsymbol{x}_n) - \boldsymbol{f}(\boldsymbol{y}_1) + \boldsymbol{f}(\boldsymbol{y}_2), \tag{4.21}$$

on déduit que $(\langle \boldsymbol{y}_1 - \boldsymbol{y}_2, \nabla \boldsymbol{f}(\boldsymbol{x}_n)\rangle)_{n\in\mathbb{N}}$ converge. Par conséquent, en passant à la limite dans les sous-suites $(\boldsymbol{x}_{k_n})_{n\in\mathbb{N}}$ et $(\boldsymbol{x}_{l_n})_{n\in\mathbb{N}}$, on obtient $\langle \boldsymbol{y}_1 - \boldsymbol{y}_2, \nabla \boldsymbol{f}(\boldsymbol{y}_1) - \nabla \boldsymbol{f}(\boldsymbol{y}_2)\rangle$, et ainsi, la monotone stricte de $\nabla \boldsymbol{f}$ implique que $\boldsymbol{y}_1 = \boldsymbol{y}_2$.

(v)⇒(iv) : Soient $\overline{\boldsymbol{x}} \in \mathfrak{W}(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ et $\boldsymbol{x} \in \boldsymbol{C} \cap \operatorname{int}\operatorname{dom}\boldsymbol{f}$. Alors il existe une suite strictement croissante $(k_n)_{n\in\mathbb{N}}$ dans $\mathbb{N}$ telle que $\boldsymbol{x}_{k_n} \rightharpoonup \overline{\boldsymbol{x}}$. D'après [2, Theorem 5.9], $\nabla \boldsymbol{f}(\boldsymbol{x}) \in \operatorname{int}\operatorname{dom}\boldsymbol{f}^*$ et il résulte de [2, Lemma 7.3(v)] que $D^{\boldsymbol{f}^*}(\cdot, \nabla \boldsymbol{f}(\boldsymbol{x}))$ est coercive. Comme (4.20) implique que $(D^{\boldsymbol{f}}(\boldsymbol{x}, \boldsymbol{x}_{k_n}))_{n\in\mathbb{N}}$ est bornée et comme [2, Lemma 7.3(vii)] affirme que

$$(\forall n \in \mathbb{N}) \quad D^{\boldsymbol{f}^*}(\nabla \boldsymbol{f}(\boldsymbol{x}_{k_n}), \nabla \boldsymbol{f}(\boldsymbol{x})) = D^{\boldsymbol{f}}(\boldsymbol{x}, \boldsymbol{x}_{k_n}), \tag{4.22}$$

on déduit que $(\nabla \boldsymbol{f}(\boldsymbol{x}_{k_n}))_{n\in\mathbb{N}}$ est bornée. Ensuite, soient $\overline{\boldsymbol{x}}^* \in \boldsymbol{\mathcal{X}}^*$ et $(p_{k_n})_{n\in\mathbb{N}}$ une suite strictement croissante dans $\mathbb{N}$ tels que $\nabla \boldsymbol{f}(\boldsymbol{x}_{p_{k_n}}) \rightharpoonup \overline{\boldsymbol{x}}^*$. Comme [2, Lemma 7.3(ii)] affirme que $D^{\boldsymbol{f}^*}(\cdot, \nabla \boldsymbol{f}(\boldsymbol{x}))$ est semi-continue inférieurement, on déduit de (4.22) que

$$D^{\boldsymbol{f}^*}(\overline{\boldsymbol{x}}^*, \nabla \boldsymbol{f}(\boldsymbol{x})) \leqslant \underline{\lim}\, D^{\boldsymbol{f}^*}\big(\nabla \boldsymbol{f}(\boldsymbol{x}_{p_{k_n}}), \nabla \boldsymbol{f}(\boldsymbol{x})\big) \leqslant \underline{\lim}\, D^{\boldsymbol{f}}(\boldsymbol{x}, \boldsymbol{x}_{p_{k_n}}) < +\infty, \tag{4.23}$$

ce qui montre que $\overline{\boldsymbol{x}}^* \in \operatorname{dom}\boldsymbol{f}^* = \operatorname{int}\operatorname{dom}\boldsymbol{f}^*$. Ainsi, d'après [2, Theorem 5.10], il existe $\overline{\boldsymbol{x}}_1 \in \operatorname{int}\operatorname{dom}\boldsymbol{f}$ tel que $\overline{\boldsymbol{x}}^* = \nabla \boldsymbol{f}(\overline{\boldsymbol{x}}_1)$. Puisque $\nabla \boldsymbol{f}^*$ est faiblement séquentiellement continu, $\overline{\boldsymbol{x}} \leftharpoonup \boldsymbol{x}_{p_{k_n}} = \nabla \boldsymbol{f}^*\big(\nabla \boldsymbol{f}(\boldsymbol{x}_{p_{k_n}})\big) \rightharpoonup \nabla \boldsymbol{f}^*(\overline{\boldsymbol{x}}^*) = \overline{\boldsymbol{x}}_1$. Par conséquent, $\overline{\boldsymbol{x}} = \overline{\boldsymbol{x}}_1 \in \operatorname{int}\operatorname{dom}\boldsymbol{f}$. □

4.3 Approximation d'un point de Kuhn-Tucker

Nous utilisons la condition suivante dans la suite.

Condition 4.7 Soit $f \in \Gamma_0(\mathcal{X})$ Gâteaux différentiable sur $\operatorname{int}\operatorname{dom} f \neq \varnothing$. Pour toute suite $(x_n)_{n\in\mathbb{N}}$ dans $\operatorname{int}\operatorname{dom} f$ et toute suite bornée $(y_n)_{n\in\mathbb{N}}$ dans $\operatorname{int}\operatorname{dom} f$,

$$D^f(x_n, y_n) \to 0 \quad \Rightarrow \quad x_n - y_n \to 0. \tag{4.24}$$

Nous déduisons de la Proposition 4.4 le résultat suivant pour résoudre le Problème 4.1.

Théorème 4.8 *Considérons le Problème 4.1. Soient $f \in \Gamma_0(\mathcal{X})$ et $g \in \Gamma_0(\mathcal{Y})$ des fonctions de Legendre. On pose*

$$\boldsymbol{f} \colon \boldsymbol{\mathcal{X}} \to \left]-\infty, +\infty\right] : (x, y^*) \mapsto f(x) + g^*(y^*), \tag{4.25}$$

et on suppose que $\boldsymbol{Z} \cap \operatorname{int}\operatorname{dom}\boldsymbol{f} \neq \varnothing$. De plus, on suppose que pour tout $x \in \operatorname{int}\operatorname{dom} f$ et pour tout $y^* \in \operatorname{int}\operatorname{dom} g^*$, $D^f(x, \cdot)$ et $D^{g^*}(y^*, \cdot)$ soient coercives. Soient $h \in \Gamma_0(\mathcal{X})$ et $j \in \Gamma_0(\mathcal{Y})$ des fonctions de Legendre telles que $\operatorname{int}\operatorname{dom} f \subset \operatorname{int}\operatorname{dom} h$, $L(\operatorname{int}\operatorname{dom} f) \subset$

$\operatorname{int}\operatorname{dom} j$*, et qu'il existe* ε *et* δ *dans* $]0,+\infty[$ *tels que* $\nabla h+\varepsilon A$ *et* $\nabla j+\delta B$ *soient coercifs. Soit* $\sigma \in [\max\{\varepsilon,\delta\},+\infty[$*, soit* $\rho \in]0,1[$*, soit* $(x_0,y_0^*) \in \operatorname{int}\operatorname{dom} f \times \operatorname{int}\operatorname{dom} g^*$*. On itère*

$$
\begin{array}{l}
\textit{pour } n = 0,1,\ldots \\
\left\lfloor
\begin{array}{l}
(\gamma_n,\mu_n) \in [\varepsilon,\sigma]\times[\delta,\sigma] \\
a_n = (\nabla h+\gamma_n A)^{-1}\big(\nabla h(x_n)-\gamma_n L^* y_n^*\big) \\
a_n^* = \gamma_n^{-1}\big(\nabla h(x_n)-\nabla h(a_n)\big) - L^* y_n^* \\
b_n = (\nabla j+\mu_n B)^{-1}\big(\nabla j(Lx_n)+\mu_n y_n^*\big) \\
b_n^* = \mu_n^{-1}\big(\nabla j(Lx_n)-\nabla j(b_n)\big) + y_n^* \\
\textit{si } \|a_n^*+L^*b_n^*\|^2+\|b_n-La_n\|^2 = 0 \\
\quad\left\lfloor
\begin{array}{l}
\overline{x} = a_n \\
\overline{y}^* = b_n^* \\
\textit{terminer.}
\end{array}
\right. \\
\boldsymbol{H}_n = \big\{(x,y^*) \in \boldsymbol{\mathcal{X}} \;\big|\; \langle x, a_n^*+L^*b_n^*\rangle + \langle b_n-La_n, y^*\rangle \leqslant \langle a_n,a_n^*\rangle + \langle b_n,b_n^*\rangle\big\} \\
(x_{n+1/2},y_{n+1/2}^*) = P^{\boldsymbol{f}}_{\boldsymbol{H}_n}(x_n,y_n^*) \\
\lambda_n \in [\rho,1] \\
x_{n+1} = \nabla f^*\big((1-\lambda_n)\nabla f(x_n)+\lambda_n\nabla f(x_{n+1/2})\big) \\
y_{n+1}^* = \nabla g\big((1-\lambda_n)\nabla g^*(y_n^*)+\lambda_n\nabla g^*(y_{n+1/2}^*)\big).
\end{array}
\right.
\end{array}
\tag{4.26}
$$

Alors (4.26) *converge vers une solution* $(\overline{x},\overline{y}^*) \in \boldsymbol{Z}$ *en un nombre fini d'itérations ou il engendre deux suites infinies* $(x_n)_{n\in\mathbb{N}}$ *et* $(y_n^*)_{n\in\mathbb{N}}$ *telles que les propriétés suivantes soient satisfaites :*

(i) $(x_n)_{n\in\mathbb{N}}$ *est une suite bornée dans* $\operatorname{int}\operatorname{dom} f$ *et* $(y_n^*)_{n\in\mathbb{N}}$ *est une suite bornée dans* $\operatorname{int}\operatorname{dom} g^*$.

(ii) $\sum_{n\in\mathbb{N}} D^f(x_{n+1/2},x_n) + \sum_{n\in\mathbb{N}} D^{g^*}(y_{n+1/2}^*,y_n^*) < +\infty$.

(iii) *Supposons que les conditions suivantes soient satisfaites :*

  (a) f*,* h*,* g^**, et* j *vérifient la Condition* 4.7.

  (b) ∇f *et* ∇g^* *sont faiblement séquentiellement continus.*

  (c) *Ou bien* ∇h *et* ∇j *sont uniformément continus sur les sous-ensembles bornés de* $\operatorname{int}\operatorname{dom} h$ *et* $\operatorname{int}\operatorname{dom} j$*, respectivement, ou bien* h^* *et* j^* *vérifient la Condition* 4.7 *et il existe* α *et* β *dans* $]0,+\infty[$ *tels que* $f \succcurlyeq \alpha h$ *et* $f \succcurlyeq \beta j\circ L$.

  (d) *Ou bien* $\mathscr{P} \subset \operatorname{int}\operatorname{dom} f$ *et* $\mathscr{D} \subset \operatorname{int}\operatorname{dom} g^*$ *ou bien* $\operatorname{dom} f^*$, $\operatorname{dom} g$ *sont ouverts et* ∇f^**,* ∇g *sont faiblement séquentiellement continus.*

  Alors $x_n \rightharpoonup \overline{x} \in \mathscr{P}$*,* $y_n \rightharpoonup \overline{y}^* \in \mathscr{D}$*, et* $(\overline{x},\overline{y}^*) \in \boldsymbol{Z}$.

Démonstration. Comme dans (2.76), les opérateurs $((\nabla h+\gamma_n A)^{-1})_{n\in\mathbb{N}}$ et $(\nabla j+\mu_n B)^{-1})_{n\in\mathbb{N}}$ sont bien définis, univoques, et coercifs. Tout d'abord, supposons qu'il existe $n\in\mathbb{N}$ tel que $La_n = b_n$ et $a_n^* = -L^*b_n^*$. Comme (4.26) implique que $(a_n,a_n^*) \in \operatorname{gra} A$ et $(b_n,b_n^*) \in \operatorname{gra} B$, on obtient $-L^*b_n^* = a_n^* \in Aa_n$ et $La_n = b_n \in B^{-1}b_n^*$, i.e., $(a_n,b_n^*) \in \boldsymbol{Z}$.

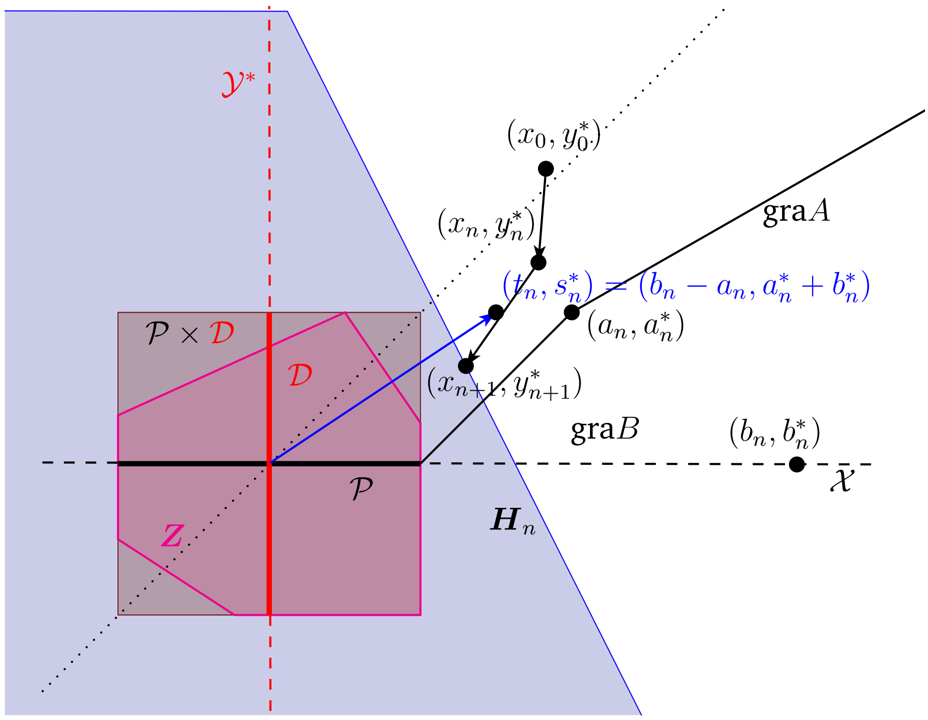

FIGURE 4.1 – Approximations successives des points de Kuhn-Tucker.

Ainsi l'algorithme se termine en $(\overline{x}, \overline{y}^*) = (a_n, b_n^*) \in \boldsymbol{Z}$. Maintenant, supposons que l'algorithme (4.26) engendre des suites infinies $(x_n)_{n\in\mathbb{N}}$ et $(y_n^*)_{n\in\mathbb{N}}$. Comme dans (2.70), $\boldsymbol{f}$ est une fonction de Legendre. En outre, il résulte de la Proposition 2.11(ii)(b) que

$$(\forall n \in \mathbb{N}) \quad \varnothing \neq \boldsymbol{Z} \subset \boldsymbol{H}_n. \tag{4.27}$$

Ensuite, posons

$$(\forall n \in \mathbb{N}) \quad \boldsymbol{x}_n = (x_n, y_n^*) \quad \text{et} \quad \boldsymbol{x}_{n+1/2} = (x_{n+1/2}, y_{n+1/2}^*). \tag{4.28}$$

Alors on récrit (4.26) comme suit

$$(\forall n \in \mathbb{N}) \quad \begin{cases} \boldsymbol{x}_{n+1/2} = P^{\boldsymbol{f}}_{\boldsymbol{H}_n} \boldsymbol{x}_n, \\ \boldsymbol{x}_{n+1} = \nabla \boldsymbol{f}^*\big((1-\lambda_n)\nabla \boldsymbol{f}(\boldsymbol{x}_n) + \lambda_n \nabla \boldsymbol{f}(\boldsymbol{x}_{n+1/2})\big). \end{cases} \tag{4.29}$$

On montre ensuite que $(\forall \boldsymbol{x} \in \operatorname{int}\operatorname{dom}\boldsymbol{f})$ $D^{\boldsymbol{f}}(\boldsymbol{x}, \cdot)$ est coercive. En fait, soit $\boldsymbol{x} = (x, y^*) \in \operatorname{int}\operatorname{dom}\boldsymbol{f}$ et soit $\eta \in \left]0, +\infty\right[$. Comme $\big\{\boldsymbol{z} = (u, v^*) \in \mathcal{X} \;\big|\; D^{\boldsymbol{f}}(\boldsymbol{x}, \boldsymbol{z}) \leqslant \eta\big\} \subset \big\{u \in \mathcal{X} \;\big|\; D^f(x, u) \leqslant \eta\big\} \times \big\{v^* \in \mathcal{Y}^* \;\big|\; D^{g^*}(y^*, v^*) \leqslant \eta\big\}$ et deux ensembles dans le membre de droite de cette inclusion sont bornés [5, Page 100], l'ensemble dans le membre de gauche de l'inclusion ci-dessus est borné, et ainsi $D^{\boldsymbol{f}}(\boldsymbol{x}, \cdot)$ est coercive.

(i) : Il résulte de la Proposition 4.4(ii) que $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ est une suite bornée dans $\operatorname{int}\operatorname{dom}\boldsymbol{f}$, ce qui implique que $(x_n)_{n\in\mathbb{N}}$ est une suite bornée dans $\operatorname{int}\operatorname{dom} f$ et $(y_n^*)_{n\in\mathbb{N}}$ est une suite bornée dans $\operatorname{int}\operatorname{dom} g^*$.

(ii) : La Proposition 4.4(iii) implique que

$$\sum_{n\in\mathbb{N}} D^f(x_{n+1/2}, x_n) + \sum_{n\in\mathbb{N}} D^f(y^*_{n+1/2}, y^*_n) = \sum_{n\in\mathbb{N}} D^{\boldsymbol{f}}(\boldsymbol{x}_{n+1/2}, \boldsymbol{x}_n) < +\infty. \tag{4.30}$$

(iii) : Premièrement, il résulte de (i) que $\mathfrak{W}(\boldsymbol{x}_n)_{n\in\mathbb{N}} \neq \varnothing$. Ensuite, on démontre que $\mathfrak{W}(\boldsymbol{x}_n)_{n\in\mathbb{N}} \subset \boldsymbol{Z}$. On suppose que $\boldsymbol{x} = (x, y^*) \in \mathfrak{W}(\boldsymbol{x}_n)_{n\in\mathbb{N}}$, i.e., $\boldsymbol{x}_{k_n} \rightharpoonup \boldsymbol{x}$. Alors

$$x_{k_n} \rightharpoonup x \quad \text{et} \quad y^*_{k_n} \rightharpoonup y^*. \tag{4.31}$$

On déduit de (ii) que

$$D^f(x_{k_n+1/2}, x_{k_n}) \to 0 \quad \text{et} \quad D^{g^*}(y^*_{k_n+1/2}, y^*_{k_n}) \to 0. \tag{4.32}$$

Ainsi, comme f et g^* vérifient la Condition 4.7, on obtient

$$x_{k_n+1/2} - x_{k_n} \to 0 \quad \text{et} \quad y^*_{k_n+1/2} - y^*_{k_n} \to 0. \tag{4.33}$$

Par ailleurs, on déduit de (4.26) et de (4.29) que

$$(\forall n \in \mathbb{N}) \quad \left\langle x_{k_n+1/2}, a^*_{k_n} + L^*b^*_{k_n} \right\rangle + \left\langle b_{k_n} - La_{k_n}, y^*_{k_n+1/2} \right\rangle \leqslant \left\langle a_{k_n}, a^*_{k_n} \right\rangle + \left\langle b_{k_n}, b^*_{k_n} \right\rangle, \tag{4.34}$$

et donc

$$\begin{aligned}
&(\forall n \in \mathbb{N}) \quad \sigma^{-1}\big(D^h(x_{k_n}, a_{k_n}) + D^j(Lx_{k_n}, b_{k_n})\big) \\
&\leqslant \sigma^{-1}\big(D^h(x_{k_n}, a_{k_n}) + D^h(a_{k_n}, x_{k_n}) + D^j(Lx_{k_n}, b_{k_n}n) + D^j(b_{k_n}, Lx_{k_n})\big) \\
&\leqslant \gamma_{k_n}^{-1} \left\langle x_{k_n} - a_{k_n}, \nabla h(x_{k_n}) - \nabla h(a_{k_n}) \right\rangle + \mu_{k_n}^{-1} \left\langle Lx_{k_n} - b_{k_n}, \nabla j(Lx_{k_n}) - \nabla j(b_{k_n}) \right\rangle \\
&= \left\langle x_{k_n} - a_{k_n}, a^*_{k_n} + L^*y^*_{k_n} \right\rangle + \left\langle Lx_{k_n} - b_{k_n}, b^*_{k_n} - y^*_{k_n} \right\rangle \\
&= \left\langle x_{k_n}, a^*_{k_n} + L^*b^*_{k_n} \right\rangle + \left\langle b_{k_n} - La_{k_n}, y^*_{k_n} \right\rangle - \left\langle a_{k_n}, a^*_{k_n} \right\rangle - \left\langle b_{k_n}, b^*_{k_n} \right\rangle \\
&\leqslant \left\langle x_{k_n} - x_{k_n+1/2}, a^*_{k_n} + L^*b^*_{k_n} \right\rangle + \left\langle b_{k_n} - La_{k_n}, y^*_{k_n} - y^*_{k_n+1/2} \right\rangle \\
&\leqslant \|x_{k_n} - x_{k_n+1/2}\| \, \|a^*_{k_n} + L^*b^*_{k_n}\| + \|b_{k_n} - La_{k_n}\| \, \|y^*_{k_n} - y^*_{k_n+1/2}\|.
\end{aligned} \tag{4.35}$$

On considère deux cas.

- Si ∇h et ∇j sont uniformément continus sur les sous-ensembles bornés de $\operatorname{int}\operatorname{dom} h$ et $\operatorname{int}\operatorname{dom} j$, respectivement, alors le Lemme 2.22 implique que $(\nabla h(x_{k_n}))_{n\in\mathbb{N}}$ et $(\nabla j(Lx_{k_n}))_{n\in\mathbb{N}}$ sont bornées et donc il résulte du Lemme 2.19 et de (4.26) que $(a_{k_n})_{n\in\mathbb{N}}$ et $(b_{k_n})_{n\in\mathbb{N}}$ sont bornées. Par suite, $(\nabla h(a_{k_n}))_{n\in\mathbb{N}}$ et $(\nabla j(b_{k_n}))_{n\in\mathbb{N}}$, et donc $(a^*_{k_n})_{n\in\mathbb{N}}$ et $(b^*_{k_n})_{n\in\mathbb{N}}$, sont bornées. Ainsi, on déduit de (4.35) que $D^h(x_{k_n}, a_{k_n}) \to 0$ et $D^j(Lx_{k_n}, b_{k_n}) \to 0$, et ainsi, comme h et j vérifient la Condition 4.7, $x_{k_n} - a_{k_n} \to 0$ et $Lx_{k_n} - b_{k_n} \to 0$. Ceci et (4.31) impliquent que

$$a_{k_n} \rightharpoonup x \quad \text{et} \quad La_{k_n} - b_{k_n} \to 0. \tag{4.36}$$

De plus, $\nabla h(a_{k_n}) - \nabla h(x_{k_n}) \to 0$ et $\nabla j(Lx_{k_n}) - \nabla j(b_{k_n}) \to 0$. Par suite, il résulte de (4.26) que $a^*_{k_n} + L^* y^*_{k_n} \to 0$ et $b^*_{k_n} - y^*_{k_n} \to 0$, et ainsi

$$b^*_{k_n} \rightharpoonup y^* \quad \text{et} \quad a^*_{k_n} + L^* b^*_{k_n} \to 0. \tag{4.37}$$

Par conséquent, on déduit de (4.36), de (4.37), et de la Proposition 2.11(iii) que $(x, y^*) \in \boldsymbol{Z}$.

- Supposons que h^* et j^* satisfassent la Condition 4.7 et qu'il existe α et β dans $]0, +\infty[$ tels que $f \succcurlyeq \alpha h$ et $f \succcurlyeq \beta j \circ L$. Soit $\overline{\boldsymbol{x}} = (\overline{x}, \overline{y}^*) \in \boldsymbol{Z} \cap \operatorname{int}\operatorname{dom} \boldsymbol{f}$. Comme la Proposition 4.4(i) implique que $\kappa = \sup_{n\in\mathbb{N}} D^{\boldsymbol{f}}(\overline{\boldsymbol{x}}, \boldsymbol{x}_{k_n}) < +\infty$, on obtient

$$(\forall n \in \mathbb{N}) \quad D^f(\overline{x}, x_{k_n}) \leqslant D^f(\overline{x}, x_{k_n}) + D^{g^*}(\overline{y}^*, y^*_{k_n}) = D^{\boldsymbol{f}}(\overline{\boldsymbol{x}}, \boldsymbol{x}_{k_n}) \leqslant \kappa. \tag{4.38}$$

Par ailleurs, on déduit du fait que $f \succcurlyeq \alpha h$ et de [2, Lemma 7.3(vii)] que

$$(\forall n \in \mathbb{N}) \quad D^{h^*}\big(\nabla h(x_{k_n}), \nabla h(\overline{x})\big) = D^h(\overline{x}, x_{k_n}) \leqslant \alpha^{-1} D^f(\overline{x}, x_{k_n}) \leqslant \alpha^{-1}\kappa. \tag{4.39}$$

Comme $\nabla h(\overline{x}) \in \operatorname{int}\operatorname{dom} h^*$, $D^{h^*}(\cdot, \nabla h(\overline{x}))$ est coercive [6, Lemma 7.3(v)], et comme (4.39) implique que $(D^{h^*}(\nabla h(x_{k_n}), \nabla h(\overline{x})))_{n\in\mathbb{N}}$ est bornée, on déduit que $(\nabla h(x_{k_n}))_{n\in\mathbb{N}}$ est bornée. Ainsi, on déduit de (4.26), de (i), et du Lemme 2.19 que $(a_{k_n})_{n\in\mathbb{N}}$ est bornée. On montre pareillement que $(b_{k_n})_{n\in\mathbb{N}}$ est bornée. Comme la convexité de h implique que

$$\begin{aligned}(\forall n \in \mathbb{N}) \quad D^h(x, a_{k_n}) &= h(\overline{x}) - h(a_{k_n}) - \langle \overline{x} - a_{k_n}, \nabla h(a_{k_n})\rangle \\ &= h(\overline{x}) - h(x_{k_n}) - \langle \overline{x} - x_{k_n}, \nabla h(x_{k_n})\rangle \\ &\quad + \langle \overline{x} - a_{k_n}, \nabla h(x_{k_n}) - \nabla h(a_{k_n})\rangle \\ &\quad - \big(h(a_{k_n}) - h(x_{k_n}) - \langle a_{k_n} - x_{k_n}, \nabla h(x_{k_n})\rangle\big) \\ &\leqslant D^h(\overline{x}, x_{k_n}) + \langle \overline{x} - a_{k_n}, \nabla h(x_{k_n}) - \nabla h(a_{k_n})\rangle\,, \end{aligned} \tag{4.40}$$

on déduit de (4.26) que

$$\begin{aligned}(\forall n \in \mathbb{N}) \quad \sigma^{-1} D^h(\overline{x}, a_{k_n}) &\leqslant \gamma_{k_n}^{-1} D^h(\overline{x}, a_{k_n}) \\ &\leqslant \varepsilon^{-1} D^h(\overline{x}, x_{k_n}) + \big\langle \overline{x} - a_{k_n}, \gamma_{k_n}^{-1}\big(\nabla h(x_{k_n}) - \nabla h(a_{k_n})\big)\big\rangle \\ &= \varepsilon^{-1} D^h(\overline{x}, x_{k_n}) + \big\langle \overline{x} - a_{k_n}, a^*_{k_n} + L^* y^*_{k_n}\big\rangle \\ &\leqslant \varepsilon^{-1}\alpha^{-1}\kappa + \big\langle \overline{x} - a_{k_n}, a^*_{k_n} + L^*\overline{y}^*\big\rangle \\ &\quad + \big\langle L\overline{x} - La_{k_n}, y^*_{k_n} - \overline{y}^*\big\rangle\,. \end{aligned} \tag{4.41}$$

De même,

$$\begin{aligned}(\forall n \in \mathbb{N}) \quad \sigma^{-1} D^j(L\overline{x}, b_{k_n}) &\leqslant \delta^{-1}\beta^{-1}\kappa + \big\langle L\overline{x} - b_{k_n}, b^*_{k_n} - \overline{y}^*\big\rangle \\ &\quad + \big\langle L\overline{x} - b_{k_n}, \overline{y}^* - y^*_{k_n}\big\rangle\,. \end{aligned} \tag{4.42}$$

Comme (4.27) implique que

$$
\begin{aligned}
\overline{\boldsymbol{x}} \in \boldsymbol{Z} &\subset \bigcap_{n\in\mathbb{N}} \boldsymbol{H}_{k_n}\\
&= \bigcap_{n\in\mathbb{N}} \big\{(x,y^*) \in \boldsymbol{\mathcal{X}} \;\big|\; \big\langle x-a_{k_n}, a^*_{k_n}+L^*y^*\big\rangle + \big\langle Lx-b_{k_n}, b^*_{k_n}-y^*\big\rangle \leqslant 0\big\},
\end{aligned}
\tag{4.43}
$$

on déduit de (4.41) et de (4.42) que

$$
\begin{aligned}
&(\forall n\in\mathbb{N})\quad \sigma^{-1}\big(D^h(\overline{x},a_{k_n})+D^j(L\overline{x},b_{k_n})\big)\\
&\leqslant (\varepsilon^{-1}\alpha^{-1}+\delta^{-1}\beta^{-1})\kappa + \big\langle L\overline{x}-La_{k_n}, y^*_{k_n}-\overline{y}^*\big\rangle + \big\langle L\overline{x}-b_{k_n}, \overline{y}^*-y^*_{k_n}\big\rangle\\
&= (\varepsilon^{-1}\alpha^{-1}+\delta^{-1}\beta^{-1})\kappa + \big\langle b_{k_n}-La_{k_n}, y^*_{k_n}-\overline{y}^*\big\rangle\\
&\leqslant (\varepsilon^{-1}\alpha^{-1}+\delta^{-1}\beta^{-1})\kappa + (\|b_{k_n}\|+\|L\|\,\|a_{k_n}\|)(\|y^*_{k_n}\|+\|\overline{y}^*\|).
\end{aligned}
\tag{4.44}
$$

Ainsi, comme $(a_{k_n})_{n\in\mathbb{N}}$, $(b_{k_n})_{n\in\mathbb{N}}$, et $(y^*_{k_n})_{n\in\mathbb{N}}$ sont bornées, $(D^h(\overline{x},a_{k_n}))_{n\in\mathbb{N}}$ et $(D^j(L\overline{x},b_{k_n}))_{n\in\mathbb{N}}$ sont également bornées. Par suite, il résulte de [2, Lemma 7.3(vii)] que $(D^{h^*}(\nabla h(a_{k_n}),\nabla h(\overline{x})))_{n\in\mathbb{N}}$ et $(D^{j^*}\nabla j(b_{k_n}),\nabla j(L\overline{x})))_{n\in\mathbb{N}}$ sont bornées. Puisque [2, Lemma 7.3(v)] affirme que $D^{h^*}(\cdot,\nabla h(\overline{x}))$ et $D^{j^*}(\cdot,\nabla j(L\overline{x}))$ sont coercives, on déduit que $(\nabla h(a_{k_n}))_{n\in\mathbb{N}}$ et $(\nabla j(b_{k_n}))_{n\in\mathbb{N}}$ sont bornées. Il résulte donc de (4.26) et (i) que $(a^*_{k_n})_{n\in\mathbb{N}}$ et $(b^*_{k_n})_{n\in\mathbb{N}}$ sont bornées. Par conséquent, on déduit de (4.33) et de (4.35) que

$$
D^h\big(x_{k_n},a_{k_n}\big)\to 0 \quad \text{et} \quad D^j\big(Lx_{k_n},b_{k_n}\big)\to 0, \tag{4.45}
$$

et puisque h et j vérifient la Condition 4.7, $x_{k_n}-a_{k_n}\to 0$ et $Lx_{k_n}-b_{k_n}\to 0$, ce qui implique que

$$
La_{k_n}-b_{k_n}\to 0 \quad \text{et} \quad a_{k_n}\rightharpoonup x. \tag{4.46}
$$

En outre, comme [2, Lemma 7.3(vii)] affirme que

$$
(\forall n\in\mathbb{N})\quad D^{h^*}\big(\nabla h(a_{k_n}),\nabla h(x_{k_n})\big) = D^h(x_{k_n},a_{k_n}) \tag{4.47}
$$

et

$$
(\forall n\in\mathbb{N})\quad D^{j^*}\big(\nabla j(b_{k_n}),\nabla j(Lx_{k_n})\big) = D^j(Lx_{k_n},b_{k_n}), \tag{4.48}
$$

on déduit de (4.45) que

$$
D^{h^*}\big(\nabla h(a_{k_n}),\nabla h(x_{k_n})\big)\to 0 \quad \text{et} \quad D^{j^*}\big(\nabla j(b_{k_n}),\nabla j(Lx_{k_n})\big)\to 0. \tag{4.49}
$$

Par suite, comme h^* et j^* vérifient la Condition 4.7, $\nabla h(a_{k_n})-\nabla h(x_{k_n})\to 0$ et $\nabla j(b_{k_n})-\nabla j(Lx_{k_n})\to 0$. Cela et (4.26) impliquent que $a^*_{k_n}+L^*y^*_{k_n}\to 0$ et $b^*_{k_n}-y^*_{k_n}\to 0$, et ainsi

$$
a^*_{k_n}+L^*b^*_{k_n}\to 0 \quad \text{et} \quad b^*_{k_n}\to y^*. \tag{4.50}
$$

Par conséquent, on déduit de (4.46), de (4.50), et de la Proposition 2.11(iii) que $(x,y^*)\in\boldsymbol{Z}$.

Finalement, au vu de la Proposition 4.4(iv), il suffit de montrer que (4.13) est satisfaite. On considère deux cas.

- Si $\mathscr{P} \subset \operatorname{int}\operatorname{dom} f$ et $\mathscr{D} \subset \operatorname{int}\operatorname{dom} g^*$ alors on déduit de la Proposition 2.11(i) que $\mathfrak{W}(\boldsymbol{x}_n)_{n\in\mathbb{N}} \subset \boldsymbol{Z} \subset \mathscr{P} \times \mathscr{D} \subset \operatorname{int}\operatorname{dom}\boldsymbol{f}$, et il donc résulte de la Remarque 4.6(ii) que (4.13) est satisfaite.
- Si $\operatorname{dom} f^*$ et $\operatorname{dom} g$ sont ouverts et que ∇f^* et ∇g sont faiblement séquentiellement continus alors $\operatorname{dom}\boldsymbol{f}^*$ est ouvert et $\nabla\boldsymbol{f}^*$ est faiblement séquentiellement continu. Ainsi, il résulte de la Remarque 4.6(v) que (4.13) est satisfaite.

□

Remarque 4.9 On reprend les notations dans le Théorème 4.8. Pour tout $n \in \mathbb{N}$, soit $\theta \in \,]0,+\infty[$ une solution de l'équation

$$\left\langle \nabla f^*\big(\nabla f(x_n) + \theta(a_n^* + L^*b_n^*)\big), a_n^* + L^*b_n^*\right\rangle \\ + \left\langle b_n - La_n, \nabla g\big(\nabla g^*(y_n^*) + \theta(b_n - La_n)\big)\right\rangle = \langle a_n, a_n^*\rangle + \langle b_n, b_n^*\rangle\,. \tag{4.51}$$

Alors, au vu de l'Exemple 4.3, on a

$$\begin{cases} x_{n+1/2} = \nabla f^*\big(\nabla f(x_n) + \theta(a_n^* + L^*b_n^*)\big);\\ y_{n+1/2} = \nabla g^*\big(\nabla g^*(y_n^*) + \theta(b_n - La_n)\big). \end{cases} \tag{4.52}$$

Remarque 4.10 Dans le Problème 4.1, supposons que $\mathcal{X}$ et $\mathcal{Y}$ soient uniformément convexes et uniformément lisses et que leurs applications de dualité soient faiblement séquentiellement continues. Alors, dans le Théorème 4.8, on peut choisir $f = \|\cdot\|^{s_1}/s_1$, $h = \|\cdot\|^{s_2}/s_2$, $j = \|\cdot\|^{s_3}/s_3$, et $g = \|\cdot\|^{s_4}/s_4$ où $(s_1, s_2, s_3, s_4) \in \,]1,+\infty[^4$. Les espaces possédant ces propriétés contiennent les espaces hilbertiens et les espaces $\ell^p(\mathbb{N})$ avec $p \in \,]1,+\infty[$.

Remarque 4.11

(i) Dans le Problème 4.1, supposons que $\mathcal{X}$ et $\mathcal{Y}$ soient hilbertiens. De plus, dans le Théorème 4.8, on prend $h = f = \|\cdot\|^2/2$ et $j = g = \|\cdot\|^2/2$. On note que, pour tout $\varepsilon \in \,]0,+\infty[$, $\nabla h + \varepsilon A = \operatorname{Id} + \varepsilon A$ et $\nabla j + \varepsilon B = \operatorname{Id} + \varepsilon B$ sont fortement monotones et donc coercifs d'après le Lemme 2.17(ii). Alors, on retrouve le cadre de [1].

(ii) Dans le Problème 4.1, supposons que $\mathcal{X} = \mathcal{Y}$ et $L = \operatorname{Id}$. Dans le Théorème 4.8, on prend $h = j = f \colon \mathcal{X} \to \mathbb{R}$ et $g^* \colon \mathcal{X}^* \to \mathbb{R}$ et on suppose en plus que h, f, j, et g^* sont supercoercives et uniformément convexes sur les sous-ensembles bornés de $\mathcal{X}$ et $\mathcal{X}^*$, respectivement. Alors, on récupère le cadre de [4].

Dans le cadre d'espaces de dimension finie, en reprenant la démonstration de la Proposition 2.30, la convergence de l'algorithme (4.59) est obtenue sous des hypothèses plus générales.

Proposition 4.12 *Dans le Problème 4.1, supposons que $\mathcal{X}$ et $\mathcal{Y}$ soient de dimension finie. Soient $f \in \Gamma_0(\mathcal{X})$ et $h \in \Gamma_0(\mathcal{X})$ des fonctions de Legendre telles que $\operatorname{int}\operatorname{dom} f \subset \operatorname{int}\operatorname{dom} h$ et $\nabla h + \varepsilon A$ est coercif pour un certain $\varepsilon \in \left]0,+\infty\right[$, soient $g \in \Gamma_0(\mathcal{Y})$ et $j \in \Gamma_0(\mathcal{Y})$ des fonctions de Legendre telles que $L(\operatorname{int}\operatorname{dom} f) \subset \operatorname{int}\operatorname{dom} j$ et $\nabla j + \delta B$ est coercif pour un certain $\delta \in \left]0,+\infty\right[$, soit $\rho \in \left]0,1\right[$, soit $\sigma \in \left[\max\{\varepsilon,\delta\},+\infty\right[$, soit $x_0 \in \operatorname{int}\operatorname{dom} f$, soit $y_0^* \in \operatorname{int}\operatorname{dom} g^*$, et on exécute l'algorithme (4.59). Supposons que $\operatorname{dom} f^*$ et $\operatorname{dom} g$ soient ouverts. Alors $(x_n, y_n^*) \to (\overline{x}, \overline{y}^*) \in \boldsymbol{Z}$.*

Démonstration. D'abord, puisque $\operatorname{dom} f^*$ et $\operatorname{dom} g$ sont ouverts, pour tout $x \in \operatorname{int}\operatorname{dom} f$ et tout $y^* \in \operatorname{int}\operatorname{dom} g^*$, $D^f(x,\cdot)$ et $D^{g^*}(y^*,\cdot)$ sont coercives [2, Lemma 7.3(ix)]. Ensuite, la Proposition 4.4(i) signifie que $(x_n)_{n\in\mathbb{N}} \in (\operatorname{int}\operatorname{dom} f)^{\mathbb{N}}$ et $(y_n^*)_{n\in\mathbb{N}} \in (\operatorname{int}\operatorname{dom} g^*)^{\mathbb{N}}$ sont bornées. Ensuite, soit $\overline{\boldsymbol{x}} = (\overline{x}, \overline{y}^*) \in \boldsymbol{Z} \cap \operatorname{int}\operatorname{dom}\boldsymbol{f}$. D'une part, la Proposition 4.4(i) implique que

$$\big(D^f(\overline{x}, x_n)\big)_{n\in\mathbb{N}} \quad \text{et} \quad \big(D^{g^*}(\overline{y}^*, y_n^*)\big)_{n\in\mathbb{N}} \quad \text{sont bornées.} \tag{4.53}$$

D'autre part, on déduit de (4.29) et de [3, Proposition 3.3(i) et Corollary 3.35(i)] que

$$(\forall n \in \mathbb{N}) \quad D^f(\overline{x}, x_{n+1/2}) + D^{g^*}(\overline{y}^*, y_{n+1/2}^*) = D^{\boldsymbol{f}}\big(\overline{\boldsymbol{x}}, \boldsymbol{x}_{n+1/2}\big) \leqslant D^{\boldsymbol{f}}\big(\overline{\boldsymbol{x}}, \boldsymbol{x}_n\big). \tag{4.54}$$

En utilisant (4.53), (4.54), et le même raisonnement de la Proposition 2.30, on obtient $(x, y^*) \in \boldsymbol{Z}$. Finalement, la convergence est donc un corollaire de la Proposition 4.4(iv) et la Remarque 4.6(v). ☐

Nous appliquons à présent le résultat du Théorème 4.8 au problème de minimisation convexe multivarié.

Problème 4.13 Soient m et p des nombres entiers strictement positifs, soient $(\mathcal{X}_i)_{1\leqslant i\leqslant m}$ et $(\mathcal{Y}_k)_{1\leqslant k\leqslant p}$ des espaces de Banach réels réflexifs, et soit $\boldsymbol{\mathcal{X}}$ l'espace vectoriel $\left(\bigtimes_{i=1}^m \mathcal{X}_i\right) \times \left(\bigtimes_{k=1}^p \mathcal{Y}_k^*\right)$ muni de la norme

$$(x, y^*) = \big((x_i)_{1\leqslant i\leqslant m}, (y_k^*)_{1\leqslant k\leqslant p}\big) \mapsto \sqrt{\sum_{i=1}^m \|x_i\|^2 + \sum_{k=1}^p \|y_k^*\|^2}. \tag{4.55}$$

Pour tout $i \in \{1, \ldots, m\}$ et pour tout $k \in \{1, \ldots, p\}$, soit $\varphi_i \in \Gamma_0(\mathcal{X}_i)$, soit $\psi_k \in \Gamma_0(\mathcal{Y}_k)$, et soit $L_{ki} \in \mathcal{B}\left(\mathcal{X}_i, \mathcal{Y}_k\right)$. Considérons le problème primal

$$\underset{x_1\in\mathcal{X}_1,\ldots,\,x_m\in\mathcal{X}_m}{\text{minimiser}} \quad \sum_{i=1}^m \varphi_i(x_i) + \sum_{k=1}^p \psi_k\bigg(\sum_{i=1}^m L_{ki} x_i\bigg), \tag{4.56}$$

le problème dual

$$\underset{y_1^*\in\mathcal{Y}_1^*,\ldots,\,y_p^*\in\mathcal{Y}_p^*}{\text{minimiser}} \quad \sum_{i=1}^m \varphi_i^*\bigg(-\sum_{k=1}^p L_{ki}^* y_k^*\bigg) + \sum_{k=1}^p \psi_k^*(y_k^*), \tag{4.57}$$

et désignons par

$$\boldsymbol{Z} = \left\{ (x, y^*) \in \boldsymbol{\mathcal{X}} \;\middle|\; (\forall i \in \{1, \ldots, m\}) \quad -\sum_{k=1}^{p} L_{ki}^* y_k^* \in \partial \varphi_i(x_i) \text{ et } (\forall k \in \{1, \ldots, p\}) \quad \sum_{i=1}^{m} L_{ki} x_i \in \partial \psi_k^*(y_k^*) \right\} \tag{4.58}$$

l'ensemble de Kuhn-Tucker associé. Le problème est de trouver un point dans $\boldsymbol{Z}$.

Proposition 4.14 *Considérons le Problème 4.13. Pour tout $i \in \{1, \ldots, m\}$, soient $f_i \in \Gamma_0(\mathcal{X}_i)$ et $h_i \in \Gamma_0(\mathcal{X}_i)$ des fonctions de Legendre telles que* $\operatorname{int} \operatorname{dom} f_i \subset \operatorname{int} \operatorname{dom} h_i$ *et $h_i + \varepsilon_i \varphi_i$ est supercoercive pour un certain $\varepsilon_i \in [0, +\infty[$. Pour tout $k \in \{1, \ldots, p\}$, soient $g_k^* \in \Gamma_0(\mathcal{Y}_k)$ et $j_k \in \Gamma_0(\mathcal{Y}_k)$ des fonctions de Legendre telles que* $\sum_{i=1}^{m} L_{ki}(\operatorname{int} \operatorname{dom} f_i) \subset \operatorname{int} \operatorname{dom} j_k$ *et $j_k + \delta_k \psi_k$ est supercoercive pour un certain $\delta_k \in [0, +\infty[$. Posons $\varepsilon = \max_{1 \leqslant i \leqslant m} \varepsilon_i$, et $\delta = \max_{1 \leqslant k \leqslant p} \delta_k$, et $\boldsymbol{f} \colon \boldsymbol{\mathcal{X}} \to \left]-\infty, +\infty\right] : (x, y^*) \mapsto \sum_{i=1}^{m} f_i(x_i) + \sum_{k=1}^{p} g_k^*(y_k^*)$, et supposons que* $\boldsymbol{Z} \cap \operatorname{int} \operatorname{dom} \boldsymbol{f} \neq \varnothing$. *Soit $\rho \in \left]0, 1\right[$, soit $\sigma \in \left[\max\{\varepsilon, \delta\}, +\infty\right[$, soit $x_0 = (x_{i,0})_{1 \leqslant i \leqslant m} \in$*

$\bigtimes_{i=1}^{m} \operatorname{int}\operatorname{dom} f_i$, *soit* $y_0^* = (y_{k,0}^*)_{1\leqslant k\leqslant p} \in \bigtimes_{k=1}^{p} \operatorname{int}\operatorname{dom} g_k^*$, *et on itère*

$$
\begin{array}{l}
\textit{pour } n = 0, 1, \ldots \\
\left\lfloor
\begin{array}{l}
(\gamma_n, \mu_n) \in \left]\varepsilon, \sigma\right] \times \left]\delta, \sigma\right] \\
\textit{pour } i = 1, \ldots, m \\
\left\lfloor
\begin{array}{l}
a_{i,n} = (\nabla h + \gamma_n \partial\varphi_i)^{-1}\big(\nabla h_i(x_{i,n}) - \gamma_n \sum_{k=1}^{p} L_{ki}^* y_{k,n}^*\big) \\
a_{i,n}^* = \gamma_n^{-1}\big(\nabla h_i(x_{i,n}) - \nabla h_i(a_{i,n})\big) - \sum_{k=1}^{p} L_{ki}^* y_{k,n}^*
\end{array}
\right. \\
\textit{pour } k = 1, \ldots, p \\
\left\lfloor
\begin{array}{l}
b_{k,n} = (\nabla j + \mu_n \psi_k)^{-1}\big(\nabla j_k\big(\sum_{i=1}^{m} L_{ki} x_{i,n}\big) + \mu_n y_{k,n}^*\big) \\
b_{k,n}^* = \mu_n^{-1}\big(\nabla j_k\big(\sum_{i=1}^{m} L_{ki} x_{i,n}\big) - \nabla j_k(b_{k,n})\big) + y_{k,n}^* \\
t_{k,n} = b_{k,n} - \sum_{i=1}^{m} L_{ki} a_{i,n}
\end{array}
\right. \\
\textit{pour } i = 1, \ldots, m \\
\left\lfloor
\begin{array}{l}
s_{i,n}^* = a_{i,n}^* + \sum_{k=1}^{p} L_{ki}^* b_{k,n}^*
\end{array}
\right. \\
\textit{si } \sum_{i=1}^{m} \|s_{i,n}^*\|^2 + \sum_{k=1}^{p} \|t_{k,n}\|^2 = 0 \\
\left\lfloor
\begin{array}{l}
\textit{pour } i = 1, \ldots, m \\
\left\lfloor \; \overline{x}_i = a_{i,n} \right. \\
\textit{pour } k = 1, \ldots, p \\
\left\lfloor \; \overline{y}_k^* = b_{k,n}^* \right. \\
\textit{terminer.}
\end{array}
\right. \\
\eta_n = \sum_{i=1}^{m} \langle a_{i,n}, a_{i,n}^* \rangle + \sum_{k=1}^{p} \langle b_{k,n}, b_{k,n}^* \rangle \\
\boldsymbol{H}_n = \left\{ (x, y^*) \in \boldsymbol{\mathcal{X}} \;\middle|\; \sum_{i=1}^{m} \langle x_i, s_{i,n}^* \rangle + \sum_{k=1}^{p} \langle t_{k,n}, y_k^* \rangle \leqslant \eta_n \right\} \\
\big((x_{i,n+1/2})_{1\leqslant i\leqslant m}, (y_{k,n+1/2}^*)_{1\leqslant k\leqslant p}\big) = P_{\boldsymbol{H}_n}^{\boldsymbol{f}}\big((x_{i,n})_{1\leqslant i\leqslant m}, (y_{k,n})_{1\leqslant k\leqslant p}\big) \\
\lambda_n \in [\rho, 1] \\
\textit{pour } i = 1, \ldots, m \\
\left\lfloor \; x_{i,n+1} = \nabla f_i^*\big((1-\lambda_n)\nabla f_i(x_{i,n}) + \lambda_n \nabla f_i(x_{i,n+1/2})\big) \right. \\
\textit{pour } k = 1, \ldots, p \\
\left\lfloor \; y_{k,n+1}^* = \nabla g_k\big((1-\lambda_n)\nabla g_k^*(u_{k,n}^*) + \lambda_n \nabla g_k^*(y_{k,n+1/2}^*)\big). \right.
\end{array}
\right.
\end{array}
\tag{4.59}
$$

Supposons que les conditions suivantes soient satisfaites :

(i) *Pour tout* $i \in \{1, \ldots, m\}$, f_i *et* h_i *vérifient la Condition 4.7,* ∇h_i *est uniformément continu sur les sous-ensembles bornés de* $\operatorname{int}\operatorname{dom} h_i$, ∇f_i *est faiblement séquentiellement continu, et* $(\forall x_i \in \operatorname{int}\operatorname{dom} f_i)$ $D^{f_i}(x_i, \cdot)$ *est coercive.*

(ii) *Pour tout* $k \in \{1, \ldots, p\}$, g_k^* *et* j_k *vérifient la Condition 4.7,* ∇j_k *est uniformément continu sur les sous-ensembles bornés de* $\operatorname{int}\operatorname{dom} j_k$, ∇g_k^* *est faiblement séquentiellement continu, et* $(\forall y_k^* \in \operatorname{int}\operatorname{dom} g_k^*)$ $D^{g_k^*}(y_k^*, \cdot)$ *est coercive.*

(iii) *Ou bien* $\boldsymbol{Z} \subset \operatorname{int}\operatorname{dom}\boldsymbol{f}$ *ou bien pour tout* $i \in \{1, \ldots, m\}$ *et pour tout* $k \in \{1, \ldots, p\}$, $\operatorname{dom} f_i^*$ *et* $\operatorname{dom} g_k$ *sont ouverts et* ∇f_i^*, ∇g_k *sont faiblement séquentiellement continus.*

Alors il existe $\big((\overline{x}_i)_{1\leqslant i\leqslant m}, (\overline{y}_k)_{1\leqslant k\leqslant p}\big) \in \boldsymbol{Z}$ *tel que*

$$(\forall i \in \{1,\ldots,m\}) \quad x_{i,n} \to \overline{x}_i \quad \textit{et} \quad (\forall k \in \{1,\ldots,p\}) \quad y^*_{k,n} \to \overline{y}^*_k. \tag{4.60}$$

Démonstration. D'abord, on désigne par $\mathcal{X}$ et $\mathcal{Y}$ les espaces vectoriels $\bigtimes_{i=1}^m \mathcal{X}_i$ et $\bigtimes_{k=1}^p \mathcal{Y}_k$ munis des normes $x = (x_i)_{1\leqslant i\leqslant m} \mapsto \sqrt{\sum_{i=1}^m \|x_i\|^2}$ et $y = (y_k)_{1\leqslant k\leqslant p} \mapsto \sqrt{\sum_{k=1}^p \|y_k\|^2}$, respectivement. Alors $\mathcal{X}^*$ est l'espace vectoriel $\bigtimes_{i=1}^m \mathcal{X}_i^*$ muni de la norme $x^* \mapsto \sqrt{\sum_{i=1}^m \|x_i^*\|^2}$ et $\mathcal{Y}^*$ est l'espace vectoriel $\bigtimes_{k=1}^p \mathcal{Y}_k^*$ muni de la norme $y^* \mapsto \sqrt{\sum_{k=1}^p \|y_k^*\|^2}$. Ensuite, on introduit les opérateurs suivants

$$\begin{cases} A\colon \mathcal{X} \to 2^{\mathcal{X}^*} \colon x \mapsto \bigtimes_{i=1}^m \partial\varphi_i(x_i) \\ B\colon \mathcal{Y} \to 2^{\mathcal{Y}^*} \colon y \mapsto \bigtimes_{k=1}^p \partial\psi_k(y_k) \\ L\colon \mathcal{X} \to \mathcal{Y} \colon x \mapsto \big(\sum_{i=1}^m L_{ki}x_i\big)_{1\leqslant k\leqslant p}. \end{cases} \tag{4.61}$$

Notons que A et B sont maximalement monotones [5, Theorem 3.1.11] et que

$$\begin{cases} A^{-1}\colon \mathcal{X}^* \to 2^{\mathcal{X}} \colon x^* \mapsto \bigtimes_{i=1}^m \partial\varphi_i^*(x_i^*) \\ B^{-1}\colon \mathcal{Y}^* \to 2^{\mathcal{Y}} \colon y^* \mapsto \bigtimes_{k=1}^p \partial\psi_k^*(y_k^*) \\ L^*\colon \mathcal{Y}^* \to \mathcal{X}^* \colon y^* \mapsto (\sum_{k=1}^p L_{ki}^* y_k^*)_{1\leqslant i\leqslant m}. \end{cases} \tag{4.62}$$

Par suite, le Problème 4.13 est un cas particulier du Problème 4.1. Ensuite, on introduit les fonctions suivantes

$$\begin{cases} f\colon \mathcal{X} \to \left]-\infty,+\infty\right] \colon x \mapsto \sum_{i=1}^m f_i(x_i) \\ h\colon \mathcal{X} \to \left]-\infty,+\infty\right] \colon x \mapsto \sum_{i=1}^m h_i(x_i) \\ \varphi\colon \mathcal{X} \to \left]-\infty,+\infty\right] \colon x \mapsto \sum_{i=1}^m \varphi_i(x_i) \\ g\colon \mathcal{Y} \to \left]-\infty,+\infty\right] \colon y \mapsto \sum_{k=1}^p g_k(y_k) \\ j\colon \mathcal{Y} \to \left]-\infty,+\infty\right] \colon y \mapsto \sum_{k=1}^p j_k(y_k). \end{cases} \tag{4.63}$$

Alors f, g, h, et j sont de Legendre, $\operatorname{int}\operatorname{dom} f = \bigtimes_{i=1}^m \operatorname{int}\operatorname{dom} f_i \subset \bigtimes_{i=1}^m \operatorname{int}\operatorname{dom} h_i = \operatorname{int}\operatorname{dom} h$, et $L(\operatorname{int}\operatorname{dom} f) = \bigtimes_{k=1}^p \sum_{i=1}^m L_{ki}(\operatorname{int}\operatorname{dom} f_i) \subset \bigtimes_{k=1}^p \operatorname{int}\operatorname{dom} j_k = \operatorname{int}\operatorname{dom} j$. Par le même raisonnement que le Théorème 4.8, on déduit que, pour tout $x \in \operatorname{int}\operatorname{dom} f$ et pour tout $y^* \in \operatorname{int}\operatorname{dom} g^*$, $D^f(x,\cdot)$ et $D^{g^*}(y^*,\cdot)$ sont coercives. Par ailleurs, pour tout $i \in \{1,\ldots,m\}$, comme $h_i + \varepsilon\varphi_i$ est supercoercive, $(h_i + \varepsilon\varphi_i)^*$ est majorée sur tout sous-ensemble borné de $\mathcal{X}_i^*$ [2, Theorem 3.3]. Ainsi $(h + \varepsilon\varphi)^* \colon x^* \mapsto \sum_{i=1}^m (h_i + \varepsilon\varphi_i)^*(x_i^*)$ est majorée sur tout sous-ensemble borné de $\mathcal{X}^*$, et il résulte de [2, Theorem 3.3] que $h+\varepsilon\varphi$ est supercoercive. Par conséquent, comme $\varnothing \neq \operatorname{dom} A \cap \operatorname{int}\operatorname{dom} f \subset \operatorname{dom}\varphi \cap \operatorname{int}\operatorname{dom} f$, on déduit de [5, Theorem 2.8.3] et du Lemme 2.17(iii) que $\nabla h + \varepsilon A = \nabla h + \varepsilon\partial\varphi = \partial(h + \varepsilon\varphi)$ est coercif. De même façon, $\nabla j + \delta B$ est coercif. Maintenant, on pose, pour tout $n \in \mathbb{N}$, $a_n = (a_{i,n})_{1 \leqslant i \leqslant m}$, $a_n^* = (a_{i,n}^*)_{1 \leqslant i \leqslant m}$, $b_n = (b_{k,n})_{1 \leqslant k \leqslant p}$, et $b_n^* = (b_{k,n}^*)_{1 \leqslant k \leqslant p}$. Alors, pour tout $n \in \mathbb{N}$, on a

$$\begin{aligned}
&(\forall i \in \{1,\ldots,m\}) \quad a_{i,n} = (\nabla h_i + \gamma_n \partial \varphi_i)^{-1}\Bigg(\nabla h_i(x_{i,n}) - \gamma_n \sum_{k=1}^{p} L_{ki}^* y_{k,n}^*\Bigg) \\
&\quad \Leftrightarrow (\forall i \in \{1,\ldots,m\}) \quad \nabla h_i(x_{i,n}) - \gamma_n \sum_{k=1}^{p} L_{ki}^* y_{k,n}^* \in \nabla h_i(a_{i,n}) + \gamma_n \partial \varphi_i(a_{i,n}) \\
&\quad \Leftrightarrow \nabla h(x_n) - \gamma_n L^* y_n^* \in \nabla h(a_n) + \gamma_n A a_n \\
&\quad \Leftrightarrow a_n = (\nabla h + \gamma_n A)^{-1}\big(\nabla h(x_n) - \gamma_n L^* y_n^*\big).
\end{aligned} \tag{4.64}$$

On montre de même que

$$(\forall n \in \mathbb{N}) \quad b_n = (\nabla j + \mu_n B)^{-1}\big(\nabla j(Lx_n) + \mu_n y_n^*\big). \tag{4.65}$$

Ainsi (4.59) est un cas particulier de (4.26). Par ailleurs, on déduit de (4.63) et des hypothèses que f, g^*, h, et j vérifient la Condition 4.7, que ∇h et ∇j sont uniformément continus sur les sous-ensembles bornés de $\operatorname{int}\operatorname{dom} h$ et $\operatorname{int}\operatorname{dom} j$, respectivement, et que ∇f et ∇g^* sont faiblement séquentiellement continus. Par conséquent, l'assertion est un corollaire du Théorème 4.8. $\square$

4.4 Bibliographie

[1] A. Alotaibi, P. L. Combettes, and N. Shahzad, Solving coupled composite monotone inclusions by successive Fejér approximations of their Kuhn-Tucker set, *SIAM J. Optim.*, vol. 24, pp. 2076–2095, 2014.

[2] H. H. Bauschke, J. M. Borwein, and P. L. Combettes, Essential smoothness, essential strict convexity and Legendre functions in Banach spaces, *Commun. Contemp. Math.*, vol. 3, pp. 615–647, 2001.

[3] H. H. Bauschke, J. M. Borwein, and P. L. Combettes, Bregman monotone optimization algorithms, *SIAM J. Control Optim.*, vol. 42, pp. 596–636, 2003.

[4] A. N. Iusem and B. F. Svaiter, Splitting methods for finding zeroes of the sum of maximal monotone operators in Banach spaces, *J. Nonlinear Convex Anal.*, vol. 15, pp. 379–397, 2014.

[5] C. Zălinescu, *Convex Analysis in General Vector Spaces*. World Scientific Publishing, River Edge, NJ, 2002.

Chapitre 5

Suites quasi-monotones au sens de Bregman

Nous introduisons une notion de suites quasi-monotones au sens de Bregman qui unifie la notion de suites quasi-fejériennes à métrique variable et celle de suites monotones au sens de Bregman. Les résultats sont appliqués à l'analyse du comportement asymptotique d'itérations proximales basées sur des distances de Bregman variables et d'algorithmes pour résoudre le problème d'admissibilité convexe dans des espaces de Banach réels réflexifs.

5.1 Description et résultats principaux

Nos résultats sont basés sur la notion suivante de distance de Bregman.

Définition 5.1 [3, 4] Soit $\mathcal{X}$ un espace de Banach réel réflexif et soit $f \in \Gamma_0(\mathcal{X})$ Gâteaux différentiable sur $\operatorname{int}\operatorname{dom} f \neq \varnothing$. La distance de Bregman associée à f est

$$
\begin{aligned}
D^f \colon \mathcal{X} \times \mathcal{X} &\to [0, +\infty] \\
(x, y) &\mapsto \begin{cases} f(x) - f(y) - \langle x - y, \nabla f(y) \rangle, & \text{si } \ y \in \operatorname{int}\operatorname{dom} f; \\ +\infty, & \text{sinon.} \end{cases}
\end{aligned}
\tag{5.1}
$$

De plus, f est *de Legendre* si elle est *essentiellement lisse* au sens où ∂f est à la fois localement borné et univoque sur son domaine, et *essentiellement strictement convexe* au sens où ∂f^* est localement borné sur son domaine et f est strictement convexe sur les sous-ensembles convexes de $\operatorname{dom}\partial f$. Soit $\varphi \in \Gamma_0(\mathcal{X})$ bornée inférieurement telle que $\operatorname{dom}\varphi \cap \operatorname{int}\operatorname{dom} f \neq \varnothing$. L'opérateur D^f*-proximal* de φ est

$$
\begin{aligned}
\operatorname{prox}^f_\varphi \colon \operatorname{int}\operatorname{dom} f &\to \operatorname{dom}\varphi \cap \operatorname{int}\operatorname{dom} f \\
y &\mapsto \underset{x \in \mathcal{X}}{\operatorname{argmin}}\, \varphi(x) + D^f(x, y).
\end{aligned}
\tag{5.2}
$$

Soit C un sous-ensemble convexe fermé de $\mathcal{X}$ tel que $C \cap \operatorname{int} \operatorname{dom} f \neq \varnothing$. Le *$D^f$-projecteur* sur C est

$$P_C^f \colon \operatorname{int} \operatorname{dom} f \to C \cap \operatorname{int} \operatorname{dom} f \qquad y \mapsto \underset{x \in C}{\operatorname{argmin}}\, D^f(x, y), \tag{5.3}$$

et la *D^f-distance* à C est

$$D_C^f \colon \mathcal{X} \to [0, +\infty] \qquad y \mapsto \inf D^f(C, y). \tag{5.4}$$

Définition 5.2 Soit $\mathcal{X}$ un espace de Banach réel réflexif et soit $f \in \Gamma_0(\mathcal{X})$ Gâteaux différentiable sur $\operatorname{int} \operatorname{dom} f \neq \varnothing$. Alors

$$\mathcal{F}(f) = \big\{ g \in \Gamma_0(\mathcal{X}) \;\big|\; g \text{ est Gâteaux différentiable sur } \operatorname{int} \operatorname{dom} g = \operatorname{int} \operatorname{dom} f \big\}. \tag{5.5}$$

De plus, si g_1 et g_2 appartiennent $\mathcal{F}(f)$, alors

$$g_1 \succcurlyeq g_2 \quad \Leftrightarrow \quad (\forall x \in \operatorname{dom} f)(\forall y \in \operatorname{int} \operatorname{dom} f) \quad D^{g_1}(x, y) \geqslant D^{g_2}(x, y). \tag{5.6}$$

Pour tout $\alpha \in [0, +\infty[$, posons

$$\mathcal{P}_\alpha(f) = \big\{ g \in \mathcal{F}(f) \;\big|\; g \succcurlyeq \alpha f \big\}. \tag{5.7}$$

Définition 5.3 Soit $\mathcal{X}$ un espace de Banach réel réflexif, soit $f \in \Gamma_0(\mathcal{X})$ Gâteaux différentiable sur $\operatorname{int} \operatorname{dom} f \neq \varnothing$, soit $(f_n)_{n\in\mathbb{N}}$ dans $\mathcal{F}(f)$, soit $(x_n)_{n\in\mathbb{N}} \in (\operatorname{int} \operatorname{dom} f)^{\mathbb{N}}$, et soit $C \subset \mathcal{X}$ tel que $C \cap \operatorname{dom} f \neq \varnothing$. Alors $(x_n)_{n\in\mathbb{N}}$ est :

(i) *quasi-monotone au sens de Bregman* par rapport à C relativement à $(f_n)_{n\in\mathbb{N}}$ si

$$\begin{aligned} &(\exists (\eta_n)_{n\in\mathbb{N}} \in \ell_+^1(\mathbb{N}))(\forall x \in C \cap \operatorname{dom} f)(\exists (\varepsilon_n)_{n\in\mathbb{N}} \in \ell_+^1(\mathbb{N}))(\forall n \in \mathbb{N}) \\ &\qquad D^{f_{n+1}}(x, x_{n+1}) \leqslant (1 + \eta_n) D^{f_n}(x, x_n) + \varepsilon_n; \end{aligned} \tag{5.8}$$

(ii) *stationnairement quasi-monotone au sens de Bregman* par rapport à C relativement à $(f_n)_{n\in\mathbb{N}}$ si

$$\begin{aligned} &(\exists (\varepsilon_n)_{n\in\mathbb{N}} \in \ell_+^1(\mathbb{N}))(\exists (\eta_n)_{n\in\mathbb{N}} \in \ell_+^1(\mathbb{N}))(\forall x \in C \cap \operatorname{dom} f)(\forall n \in \mathbb{N}) \\ &\qquad D^{f_{n+1}}(x, x_{n+1}) \leqslant (1 + \eta_n) D^{f_n}(x, x_n) + \varepsilon_n. \end{aligned} \tag{5.9}$$

Remarque 5.4

(i) Dans la Définition 5.3, supposons que $(\forall n \in \mathbb{N})$ $f_n = f$ et $\eta_n = \varepsilon_n = 0$. Alors on récupère la notion d'une suite monotone au sens de Bregman définie dans [4].

(ii) Dans la Définition 5.3, supposons que $\mathcal{X}$ soit hilbertien, que $f = \|\cdot\|^2/2$, et que $(\forall n \in \mathbb{N})$ $f_n = \langle \cdot, U_n \cdot \rangle /2$ où $(U_n)_{n\in\mathbb{N}}$ sont des opérateurs auto-adjoints dans $\mathcal{B}(\mathcal{X})$ tels que $(\forall n \in \mathbb{N})$ $\langle \cdot, U_n \cdot \rangle \geqslant \alpha \|\cdot\|^2$ avec $\alpha \in [0, +\infty[$. Alors on récupère [9, Definition 2.1] pour $\phi = |\cdot|^2/2$.

Nous présentons quelques propriétés élémentaires des suites quasi-monotones au sens de Bregman dans la proposition suivante.

Proposition 5.5 *Soit $\mathcal{X}$ un espace de Banach réel réflexif, soit $f \in \Gamma_0(\mathcal{X})$ Gâteaux différentiable sur $\operatorname{int}\operatorname{dom} f \neq \varnothing$, soit $\alpha \in \left]0,+\infty\right[$, soit $(f_n)_{n\in\mathbb{N}}$ dans $\mathcal{P}_\alpha(f)$, soit $(x_n)_{n\in\mathbb{N}} \in (\operatorname{int}\operatorname{dom} f)^{\mathbb{N}}$, soit $C \subset \mathcal{X}$ tel que $C \cap \operatorname{int}\operatorname{dom} f \neq \varnothing$, et soit $x \in C \cap \operatorname{int}\operatorname{dom} f$. Supposons que $(x_n)_{n\in\mathbb{N}}$ soit quasi-monotone au sens de Bregman par rapport à C relativement à $(f_n)_{n\in\mathbb{N}}$. Alors les propriétés suivantes sont satisfaites :*

(i) *$(D^{f_n}(x,x_n))_{n\in\mathbb{N}}$ converge.*

(ii) *Supposons que $D^f(x,\cdot)$ soit coercive. Alors $(x_n)_{n\in\mathbb{N}}$ est bornée.*

Nous avons le résultat suivant pour la convergence faible des suites quasi-monotones au sens de Bregman.

Proposition 5.6 *Soit $\mathcal{X}$ un espace de Banach réel réflexif, soit $f \in \Gamma_0(\mathcal{X})$ Gâteaux différentiable sur $\operatorname{int}\operatorname{dom} f \neq \varnothing$, soit $(x_n)_{n\in\mathbb{N}} \in (\operatorname{int}\operatorname{dom} f)^{\mathbb{N}}$, soit $C \subset \mathcal{X}$ tel que $C \cap \operatorname{int}\operatorname{dom} f \neq \varnothing$, soit $(\eta_n)_{n\in\mathbb{N}} \in \ell_+^1(\mathbb{N})$, soit $\alpha \in \left]0,+\infty\right[$, et soit $(f_n)_{n\in\mathbb{N}}$ dans $\mathcal{P}_\alpha(f)$ telle que $(\forall n \in \mathbb{N})$ $(1+\eta_n)f_n \succcurlyeq f_{n+1}$. Supposons que $(x_n)_{n\in\mathbb{N}}$ soit quasi-monotone au sens de Bregman par rapport à C relativement à $(f_n)_{n\in\mathbb{N}}$, qu'il existe $g \in \mathcal{F}(f)$ telle que pour tout $n \in \mathbb{N}$, $g \succcurlyeq f_n$, et que, pour tout $y_1 \in \mathcal{X}$ et pour tout $y_2 \in \mathcal{X}$,*

$$\begin{cases} y_1 \in \mathfrak{W}(x_n)_{n\in\mathbb{N}} \cap C \\ y_2 \in \mathfrak{W}(x_n)_{n\in\mathbb{N}} \cap C \\ \big(\langle y_1 - y_2, \nabla f_n(x_n)\rangle\big)_{n\in\mathbb{N}} \quad \textit{converge} \end{cases} \quad \Rightarrow \quad y_1 = y_2. \tag{5.10}$$

De plus, supposons que $(\forall x \in \operatorname{int}\operatorname{dom} f)$ $D^f(x,\cdot)$ soit coercive. Alors $(x_n)_{n\in\mathbb{N}}$ converge faiblement vers un point dans $C \cap \operatorname{int}\operatorname{dom} f$ si et seulement si $\mathfrak{W}(x_n)_{n\in\mathbb{N}} \subset C \cap \operatorname{int}\operatorname{dom} f$.

Pour énoncer et démontrer notre résultat de convergence forte, nous avons besoin de la condition suivante.

Condition 5.7 [4, Condition 4.4] Soit $\mathcal{X}$ un espace de Banach réel réflexif et soit $f \in \Gamma_0(\mathcal{X})$ Gâteaux différentiable sur $\operatorname{int}\operatorname{dom} f \neq \varnothing$. Pour toutes suites bornées $(x_n)_{n\in\mathbb{N}}$ et $(y_n)_{n\in\mathbb{N}}$ dans $\operatorname{int}\operatorname{dom} f$,

$$D^f(x_n,y_n) \to 0 \quad \Rightarrow \quad x_n - y_n \to 0. \tag{5.11}$$

Proposition 5.8 *Soit $\mathcal{X}$ un espace de Banach réel réflexif, soit $f \in \Gamma_0(\mathcal{X})$ une fonction de Legendre, soit $\alpha \in \left]0,+\infty\right[$, soit $(f_n)_{n\in\mathbb{N}}$ dans $\mathcal{P}_\alpha(f)$, soit $(x_n)_{n\in\mathbb{N}} \in (\operatorname{int}\operatorname{dom} f)^{\mathbb{N}}$, et soit C un sous-ensemble convexe fermé de $\mathcal{X}$ tel que $C \cap \operatorname{int}\operatorname{dom} f \neq \varnothing$. Supposons que $(x_n)_{n\in\mathbb{N}}$ soit stationnairement quasi-monotone au sens de Bregman par rapport à C relativement à $(f_n)_{n\in\mathbb{N}}$, que f satisfasse la Condition 5.7, et que $(\forall x \in \operatorname{int}\operatorname{dom} f)$ $D^f(x,\cdot)$ soit coercive. De plus, supposons qu'il existe $\beta \in \left]0,+\infty\right[$ tel que $(\forall n \in \mathbb{N})$ $\beta f \succcurlyeq f_n$. Alors $(x_n)_{n\in\mathbb{N}}$ converge fortement vers un point dans $C \cap \overline{\operatorname{dom}} f$ si et seulement si $\varliminf D_C^f(x_n) = 0$.*

Le résultat suivant concerne la convergence d'un algorithme du point proximal qui utilise des distances de Bregman différentes à chaque itération.

Théorème 5.9 *Soit $\mathcal{X}$ un espace de Banach réel réflexif, soit $\varphi \in \Gamma_0(\mathcal{X})$, soit $f \in \Gamma_0(\mathcal{X})$ une fonction de Legendre telle que $\operatorname{Argmin}\varphi \cap \operatorname{int}\operatorname{dom} f \neq \varnothing$, soit $(\eta_n)_{n\in\mathbb{N}} \in \ell^1_+(\mathbb{N})$, soit $\alpha \in \left]0, +\infty\right[$, et soient $(f_n)_{n\in\mathbb{N}}$ des fonctions de Legendre dans $\mathcal{P}_\alpha(f)$ telles que*

$$(\forall n \in \mathbb{N}) \quad (1+\eta_n) f_n \succcurlyeq f_{n+1}. \tag{5.12}$$

Soit $x_0 \in \operatorname{int}\operatorname{dom} f$ et soit $(\gamma_n)_{n\in\mathbb{N}} \in \left]0, +\infty\right[^{\mathbb{N}}$ telle que $\inf_{n\in\mathbb{N}} \gamma_n > 0$. Posons

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = \operatorname{prox}^{f_n}_{\gamma_n\varphi} x_n. \tag{5.13}$$

Alors les propriétés suivantes sont satisfaites :

(i) *$(x_n)_{n\in\mathbb{N}}$ est stationnairement quasi-monotone au sens de Bregman par rapport à $\operatorname{Argmin}\varphi$ relativement à $(f_n)_{n\in\mathbb{N}}$.*

(ii) *$D^f(x_{n+1}, x_n) \to 0$ et $(x_n)_{n\in\mathbb{N}}$ est une suite minimisante de φ.*

(iii) *Supposons que pour tout $x \in \operatorname{int}\operatorname{dom} f$, $D^f(x, \cdot)$ soit coercive, et qu'une des conditions suivantes soit satisfaite.*

(a) *$\operatorname{Argmin}\varphi \cap \overline{\operatorname{dom}} f$ est au plus un singleton.*

(b) *Ou bien $\operatorname{Argmin}\varphi \subset \operatorname{int}\operatorname{dom} f$ ou bien $\operatorname{dom} f^*$ est ouvert et ∇f^* est faiblement séquentiellement continu, il existe $g \in \mathcal{F}(f)$ telle que, pour tout $n \in \mathbb{N}$, $g \succcurlyeq f_n$, et, pour tout $y_1 \in \mathcal{X}$ et pour tout $y_2 \in \mathcal{X}$,*

$$\begin{cases} y_1 \in \mathfrak{W}(x_n)_{n\in\mathbb{N}} \\ y_2 \in \mathfrak{W}(x_n)_{n\in\mathbb{N}} \\ \big(\langle y_1 - y_2, \nabla f_n(x_n) \rangle \big)_{n\in\mathbb{N}} \quad \textit{converge} \end{cases} \Rightarrow \quad y_1 = y_2. \tag{5.14}$$

Alors il existe $\overline{x} \in \operatorname{Argmin}\varphi$ tel que $x_n \rightharpoonup \overline{x}$.

(iv) *Supposons que f satisfasse la Condition 5.7 et que $(\forall x \in \operatorname{int}\operatorname{dom} f)$ $D^f(x, \cdot)$ soit coercive. De plus, supposons que $\underline{\lim}\, D^f_{\operatorname{Argmin}\varphi}(x_n) = 0$ et qu'il existe $\beta \in \left]0, +\infty\right[$ tel que $(\forall n \in \mathbb{N})$ $\beta f \succcurlyeq f_n$. Alors il existe $\overline{x} \in \operatorname{Argmin}\varphi$ tel que $x_n \to \overline{x}$.*

Remarque 5.10 Dans le Théorème 5.9, supposons que $(\forall n \in \mathbb{N})$ $f_n = f$, $\gamma_n = \gamma$, et $\eta_n = 0$. Alors on retrouve l'algorithme du point proximal avec des distances de Bregman de [4] :

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = \operatorname{prox}^{f}_{\gamma\varphi} x_n. \tag{5.15}$$

Nous rappelons la notion suivante.

Définition 5.11 [4, Definition 3.1] Soit $\mathcal{X}$ un espace de Banach réel et soit $f \in \Gamma_0(\mathcal{X})$ Gâteaux différentiable sur $\operatorname{int} \operatorname{dom} f \neq \varnothing$. Pour tout $x \in \operatorname{int} \operatorname{dom} f$ et pour tout $y \in \operatorname{int} \operatorname{dom} f$, posons

$$H^f(x,y) = \big\{ z \in \mathcal{X} \;\big|\; \langle z - y, \nabla f(x) - \nabla f(y) \rangle \leqslant 0 \big\}. \tag{5.16}$$

Alors

$$\mathfrak{B}(f) = \Big\{ T \colon \mathcal{X} \to 2^{\mathcal{X}} \;\Big|\; \operatorname{ran} T \subset \operatorname{dom} T = \operatorname{int} \operatorname{dom} f \\ \text{et } (\forall (x,y) \in \operatorname{gra} T)\ \operatorname{Fix} T \subset H^f(x,y) \Big\}. \tag{5.17}$$

Enfin, nous proposons une méthode pour résoudre le problème d'admissibilité convexe, i.e., le problème de trouver un point dans l'intersection d'une famille d'ensembles convexes fermés dans un espace de Banach réel réflexif.

Théorème 5.12 *Soit $\mathcal{X}$ un espace de Banach réel réflexif, soit I un ensemble totalement ordonné au plus dénombrable, soit $(C_i)_{i \in I}$ une famille de sous-ensembles convexes fermés de $\mathcal{X}$ telle que $C = \bigcap_{i \in I} C_i \neq \varnothing$, soit $f \in \Gamma_0(\mathcal{X})$ Gâteaux différentiable sur $\operatorname{int} \operatorname{dom} f$, soit $(\eta_n)_{n \in \mathbb{N}} \in \ell^1_+(\mathbb{N})$, soit $\alpha \in \left]0, +\infty\right[$, soient $(f_n)_{n \in \mathbb{N}}$ des fonctions de Legendre dans $\mathcal{P}_\alpha(f)$ telles que*

$$(\forall n \in \mathbb{N}) \quad (1 + \eta_n) f_n \succcurlyeq f_{n+1}. \tag{5.18}$$

Soit $\mathrm{i} \colon \mathbb{N} \to I$ telle que

$$(\forall j \in I)(\exists M_j \in \mathbb{N} \backslash \{0\})(\forall n \in \mathbb{N}) \quad j \in \{\mathrm{i}(n), \ldots, \mathrm{i}(n + M_j - 1)\}. \tag{5.19}$$

Pour tout $i \in I$, soit $(T_{i,n})_{n \in \mathbb{N}}$ une suite d'opérateurs tels que

$$(\forall n \in \mathbb{N}) \quad T_{i,n} \in \mathfrak{B}(f_n), \quad C_i \cap \operatorname{Fix} T_{i,n} \neq \varnothing, \quad \text{et} \quad C_i \subset \overline{\operatorname{Fix}}\, T_{i,n}. \tag{5.20}$$

Soit $x_0 \in \operatorname{int} \operatorname{dom} f$ et posons

$$(\forall n \in \mathbb{N}) \quad x_{n+1} \in T_{\mathrm{i}(n),n} x_n. \tag{5.21}$$

Supposons que f satisfasse la Condition 5.7 et que $(\forall x \in \operatorname{int} \operatorname{dom} f)$ $D^f(x, \cdot)$ soit coercive. Alors, il existe $\overline{x} \in C$ tel que les propriétés suivantes sont satisfaites :

(i) *Supposons qu'il existe $g \in \mathcal{F}(f)$ telle que, pour tout $n \in \mathbb{N}$, $g \succcurlyeq f_n$, que, pour tout $y_1 \in \mathcal{X}$ et pour tout $y_2 \in \mathcal{X}$,*

$$\begin{cases} y_1 \in \mathfrak{W}(x_n)_{n \in \mathbb{N}} \cap C \\ y_2 \in \mathfrak{W}(x_n)_{n \in \mathbb{N}} \cap C \\ \big(\langle y_1 - y_2, \nabla f_n(x_n) \rangle \big)_{n \in \mathbb{N}} \quad \textit{converge} \end{cases} \quad \Rightarrow \quad y_1 = y_2, \tag{5.22}$$

et que, pour toute suite strictement croissante $(l_n)_{n\in\mathbb{N}}$ *dans* $\mathbb{N}$*, pour tout* $x \in \mathcal{X}$*, et pour tout* $j \in I$*,*

$$\begin{cases} x_{l_n} \rightharpoonup x \\ y_{l_n} \in T_{j,l_n} x_{l_n} \\ y_{l_n} - x_{l_n} \to 0 \\ (\forall n \in \mathbb{N})\ j = \mathrm{i}(l_n) \end{cases} \quad \Rightarrow \quad x \in C_j. \tag{5.23}$$

De plus, supposons que $\mathfrak{W}(x_n)_{n\in\mathbb{N}} \subset \operatorname{int}\operatorname{dom} f$*. Alors* $x_n \rightharpoonup \overline{x}$*.*

(ii) *Supposons que* f *soit de Legendre, que* $\underline{\lim}\, D_C^f(x_n) = 0$*, et qu'il existe* $\beta \in \left]0, +\infty\right[$ *tel que* $(\forall n \in \mathbb{N})\ \beta f \succcurlyeq f_n$*. Alors* $x_n \to \overline{x}$*.*

Remarque 5.13

(i) Dans le Théorème 5.12, supposons que $(\forall n \in \mathbb{N})\ f_n = f$ et $\eta_n = 0$. Alors on retrouve les résultats de [4, Section 4.2].

(ii) Dans le Théorème 5.12, supposons que $\mathcal{X}$ soit hilbertien, que $f = \|\cdot\|^2/2$, et que $(\forall n \in \mathbb{N})\ f_n = \langle \cdot, U_n \cdot \rangle /2$, où $(U_n)_{n\in\mathbb{N}}$ est une suite d'opérateurs auto-adjoints dans $\mathcal{B}(\mathcal{X})$ tels que $\sup_{n\in\mathbb{N}} \|U_n\| < +\infty$ et

$$(\forall n \in \mathbb{N})(\forall x \in \mathcal{X}) \quad \langle x, U_n x\rangle \geqslant \alpha\|x\|^2 \quad \text{et} \quad (1+\eta_n)\langle x, U_n x\rangle \geqslant \langle x, U_{n+1} x\rangle . \tag{5.24}$$

Alors on retrouve la version sans erreurs de [9, Theorem 5.1] avec $(\forall n \in \mathbb{N})$ $\lambda_n = 1$.

Le corollaire suivant est une application du Théorème 5.12 à la méthode de projections périodiques avec des distances de Bregman variables.

Corollaire 5.14 *Soit* $\mathcal{X}$ *un espace de Banach réel réflexif, soit* m *un entier strictement positif, soit* $(C_i)_{1\leqslant i\leqslant m}$ *une famille de sous-ensembles convexes fermés de* $\mathcal{X}$ *telle que* $C = \bigcap_{i=1}^m C_i \neq \varnothing$*, soit* $f \in \Gamma_0(\mathcal{X})$ *Gâteaux différentiable sur* $\operatorname{int}\operatorname{dom} f$ *telle que* $C \cap \operatorname{int}\operatorname{dom} f \neq \varnothing$*, soit* $(\eta_n)_{n\in\mathbb{N}} \in \ell_+^1(\mathbb{N})$*, soit* $\alpha \in \left]0, +\infty\right[$*, et soient* $(f_n)_{n\in\mathbb{N}}$ *des fonctions de Legendre dans* $\mathcal{P}_\alpha(f)$ *telles que*

$$(\forall n \in \mathbb{N})(\forall i \in \{1, \ldots, m\}) \quad (1+\eta_n) f_n \succcurlyeq f_{n+1}. \tag{5.25}$$

Soit $x_0 \in \operatorname{int}\operatorname{dom} f$ *et posons*

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = P^{f_n}_{C_{1+\mathrm{rem}(n,m)}} x_n, \tag{5.26}$$

où $\mathrm{rem}(\cdot, m)$ *est le reste de la division par* m*. Supposons que* f *satisfasse la Condition 5.7, et que* $(\forall x \in \operatorname{int}\operatorname{dom} f)\ D^f(x, \cdot)$ *soit coercive. Alors il existe* $\overline{x} \in C$ *tel que les propriétés suivantes soient satisfaites :*

(i) *Supposons qu'il existe* $g \in \mathcal{F}(f)$ *telle que, pour tout* $n \in \mathbb{N}$, $g \succcurlyeq f_n$, *et que, pour tout* $y_1 \in \mathcal{X}$ *et pour tout* $y_2 \in \mathcal{X}$,

$$\begin{cases} y_1 \in \mathfrak{W}(x_n)_{n\in\mathbb{N}} \cap C \\ y_2 \in \mathfrak{W}(x_n)_{n\in\mathbb{N}} \cap C \\ \big(\langle y_1 - y_2, \nabla f_n(x_n)\rangle\big)_{n\in\mathbb{N}} \quad \textit{converge} \end{cases} \qquad \Rightarrow \qquad y_1 = y_2. \tag{5.27}$$

De plus, supposons que $\mathfrak{W}(x_n)_{n\in\mathbb{N}} \subset \operatorname{int}\operatorname{dom} f$. *Alors* $x_n \rightharpoonup \overline{x}$.

(ii) *Supposons que* f *soit de Legendre, que* $\underline{\lim}\, D_C^f(x_n) = 0$, *et qu'il existe* $\beta \in \left]0, +\infty\right[$ *tel que* $(\forall n \in \mathbb{N})\ \beta f \succcurlyeq f_n$. *Alors* $x_n \to \overline{x}$.

5.2 Article en anglais

VARIABLE QUASI-BREGMAN MONOTONE SEQUENCES [1]

5.2.1 Introduction

The concept of Fejér monotonicity and its variants plays an important role in the convergence analysis of many fixed point and optimization algorithms in Hilbert spaces [1, 5, 7, 8, 11, 17]. A recent development in this area is the extension of the notion of (quasi)-Fejér sequence to the case when the underlying metric is allowed to vary over the iterations [9]. Since Fejér monotonicity is of limited use outside of Hilbert spaces, the notion of Bregman monotonicity was introduced in [4] to provide a unifying framework for the convergence analysis of various algorithms for solving nonlinear problems. The main objective of the present paper is to unify the work of [9] on variable metric Fejér sequences and that of [4] on Bregman monotone sequences by introducing the notion of a variable quasi-Bregman monotone sequence and by investigating its asymptotic properties. We apply these results to a variable Bregman proximal point algorithm and to convex feasibility problems in Banach spaces. Our paper revolves around the following definitions.

Definition 5.15 [3, 4] Let $\mathcal{X}$ be a reflexive real Banach space, let $\mathcal{X}^*$ be the topological dual space of $\mathcal{X}$, let $\langle\cdot,\cdot\rangle$ be the duality pairing between $\mathcal{X}$ and $\mathcal{X}^*$, let $f\colon \mathcal{X} \to \left]-\infty, +\infty\right]$ be a lower semicontinuous convex function that is Gâteaux differentiable on $\operatorname{int}\operatorname{dom} f \neq \varnothing$, let $f^*\colon \mathcal{X}^* \to \left]-\infty, +\infty\right] : x^* \mapsto \sup_{x\in\mathcal{X}}(\langle x, x^*\rangle - f(x))$ be conjugate of f, and let

$$\partial f\colon \mathcal{X} \to 2^{\mathcal{X}^*} : x \mapsto \big\{x^* \in \mathcal{X}^* \;\big|\; (\forall y \in \mathcal{X})\ \langle y - x, x^*\rangle + f(x) \leqslant f(y)\big\}, \tag{5.28}$$

1. V. Q. Nguyen, Variable Quasi-Bregman monotone sequences, *submitted*.

be Moreau subdifferential of f. The *Bregman distance* associated with f is

$$D^f \colon \mathcal{X} \times \mathcal{X} \to [0,+\infty]$$
$$(x,y) \mapsto \begin{cases} f(x) - f(y) - \langle x-y, \nabla f(y) \rangle, & \text{if } y \in \operatorname{int} \operatorname{dom} f; \\ +\infty, & \text{otherwise.} \end{cases} \tag{5.29}$$

In addition, f is a *Legendre function* if it is *essentially smooth* in the sense that ∂f is both locally bounded and single-valued on its domain, and *essentially strictly convex* in the sense that ∂f^* is locally bounded on its domain and f is strictly convex on every convex subset of $\operatorname{dom} \partial f$. Let $\varphi \colon \mathcal{X} \to \left]-\infty,+\infty\right]$ be a lower semicontinuous convex function which is bounded from below and $\operatorname{dom} \varphi \cap \operatorname{int} \operatorname{dom} f \neq \varnothing$. The *$D^f$-proximal operator* of φ is

$$\begin{aligned} \operatorname{prox}^f_\varphi \colon \operatorname{int} \operatorname{dom} f &\to \operatorname{dom} \varphi \cap \operatorname{int} \operatorname{dom} f \\ y &\mapsto \underset{x \in \mathcal{X}}{\operatorname{argmin}}\, \varphi(x) + D^f(x,y). \end{aligned} \tag{5.30}$$

Let C be a closed convex subset of $\mathcal{X}$ such that $C \cap \operatorname{int} \operatorname{dom} f \neq \varnothing$. The *Bregman projector* onto C induced by f is

$$\begin{aligned} P^f_C \colon \operatorname{int} \operatorname{dom} f &\to C \cap \operatorname{int} \operatorname{dom} f \\ y &\mapsto \underset{x \in C}{\operatorname{argmin}}\, D^f(x,y), \end{aligned} \tag{5.31}$$

and the *D^f-distance* to C is the function

$$\begin{aligned} D^f_C \colon \mathcal{X} &\to [0,+\infty] \\ y &\mapsto \inf D^f(C,y). \end{aligned} \tag{5.32}$$

The paper is organized as follows. In Section 5.2.2, we introduce the notion of a variable quasi-Bregman monotone sequence and investigate its asymptotic properties. Basic results on D^f-proximal operators are reviewed in Section 5.2.3. Applications to a variable Bregman proximal point algorithm and to the convex feasibility problem are considered in Section 5.2.4.

Notation and background. The norm of a Banach space is denoted by $\|\cdot\|$. The symbols $\rightharpoonup$ and $\to$ represent respectively weak and strong convergence. The set of weak sequential cluster points of a sequence $(x_n)_{n\in\mathbb{N}}$ is denoted by $\mathfrak{W}(x_n)_{n\in\mathbb{N}}$. Let $M \colon \mathcal{X} \to 2^{\mathcal{X}}$. The domain of M is $\operatorname{dom} M = \{x \in \mathcal{X} \mid Mx \neq \varnothing\}$, the range of M is $\operatorname{ran} M = \{y \in \mathcal{X} \mid (\exists x \in \mathcal{X})\, y \in Mx\}$, the graph of M is $\operatorname{gra} M = \{(x,y) \in \mathcal{X} \times \mathcal{X} \mid y \in Mx\}$, and the set of fixed points of M is $\operatorname{Fix} M = \{x \in \mathcal{X} \mid x \in Mx\}$. A function $f \colon \mathcal{X} \to \left]-\infty,+\infty\right]$ is coercive if $\lim_{\|x\|\to+\infty} f(x) = +\infty$. Denote by $\Gamma_0(\mathcal{X})$ the class of all lower semicontinuous convex functions $f \colon \mathcal{X} \to \left]-\infty,+\infty\right]$ such that $\operatorname{dom} f = \{x \in \mathcal{X} \mid f(x) < +\infty\} \neq \varnothing$. Let $f \in \Gamma_0(\mathcal{X})$. The set of global minimizers of f is denoted by $\operatorname{Argmin} f$. Finally, $\ell^1_+(\mathbb{N})$ is the set of all summable sequences in $\left[0,+\infty\right[$.

5.2.2 Variable Bregman monotonicity

Definition 5.16 Let $\mathcal{X}$ be a reflexive real Banach space and let $f \in \Gamma_0(\mathcal{X})$ be Gâteaux differentiable on $\operatorname{int}\operatorname{dom} f \neq \varnothing$. Then

$$\mathcal{F}(f) = \big\{ g \in \Gamma_0(\mathcal{X}) \;\big|\; g \text{ is Gâteaux differentiable on } \operatorname{int}\operatorname{dom} g = \operatorname{int}\operatorname{dom} f \big\}. \quad (5.33)$$

Moreover, if g_1 and g_2 are in $\mathcal{F}(f)$, then

$$g_1 \succcurlyeq g_2 \quad \Leftrightarrow \quad (\forall x \in \operatorname{dom} f)(\forall y \in \operatorname{int}\operatorname{dom} f) \quad D^{g_1}(x,y) \geqslant D^{g_2}(x,y). \quad (5.34)$$

For every $\alpha \in [0,+\infty[$, set

$$\mathcal{P}_\alpha(f) = \big\{ g \in \mathcal{F}(f) \;\big|\; g \succcurlyeq \alpha f \big\}. \quad (5.35)$$

Remark 5.17 In Definition 5.16, suppose that $\mathcal{X}$ is a Hilbert space and let $\alpha \in]0,+\infty[$. Then the following hold :

(i) Suppose that f is Fréchet differentiable on $\mathcal{X}$. Then $\|\cdot\|^2/2 \in \mathcal{P}_\alpha(f)$ if and only if ∇f is α^{-1}-Lipschitz continuous.

(ii) Let $\mathcal{S}(\mathcal{X})$ be the space of self-adjoint bounded linear operators from $\mathcal{X}$ to $\mathcal{X}$. The Loewner partial ordering on $\mathcal{S}(\mathcal{X})$ is defined by

$$(\forall U_1 \in \mathcal{S}(\mathcal{X}))(\forall U_2 \in \mathcal{S}(\mathcal{X})) \quad U_1 \succcurlyeq U_2 \quad \Leftrightarrow \quad (\forall x \in \mathcal{X}) \quad \big\langle x, U_1 x \big\rangle \geqslant \big\langle x, U_2 x \big\rangle. \quad (5.36)$$

Set $\mathcal{P}_\alpha(\mathcal{X}) = \big\{ U \in \mathcal{S}(\mathcal{X}) \;\big|\; U \succcurlyeq \alpha\,\mathrm{Id} \big\}$. Let $U \in \mathcal{S}(\mathcal{X})$ and $V \in \mathcal{S}(\mathcal{X})$ be such that $V \succcurlyeq \alpha U$. Suppose that $f \colon x \mapsto \langle x, Ux\rangle/2$ and $g \colon x \mapsto \langle x, Vx\rangle/2$. Then $g \in \mathcal{P}_\alpha(f)$.

Proof. (i) : First, since f is Fréchet differentiable, $\partial f = \nabla f$ [5, Proposition 17.26] and hence, by [5, Corollary 16.24], $(\nabla f)^{-1} = (\partial f)^{-1} = \partial f^*$. Now, we have

$$\begin{aligned}
\|\cdot\|^2/2 \in \mathcal{P}_\alpha(f) &\Leftrightarrow (\forall x \in \mathcal{X})(\forall y \in \mathcal{X})\ \|x-y\|^2/2 \geqslant \alpha D^f(x,y)\\
&\Leftrightarrow (\forall x \in \mathcal{X})(\forall y \in \mathcal{X})\ \|x-y\|^2/(2\alpha) \geqslant f(x) - f(y) - \langle x-y, \nabla f(y)\rangle\\
&\Leftrightarrow (\forall x \in \mathcal{X})(\forall y \in \mathcal{X})\ f(x) \leqslant f(y) + \langle x-y, \nabla f(y)\rangle + \|x-y\|^2/(2\alpha).
\end{aligned} \quad (5.37)$$

The assertion therefore follows by invoking [5, Theorem 18.15].

(ii) : We observe that f and g are Gâteaux differentiable on $\mathcal{X}$ with $\nabla f = U$ and $\nabla g = V$. Consequently,

$$\begin{aligned}
(\forall x \in \mathcal{X})(\forall y \in \mathcal{X}) \quad D^g(x,y) &= \langle x, Vx\rangle/2 - \langle y, Vy\rangle/2 - \langle x-y, Vy\rangle\\
&= \langle x-y, Vx-Vy\rangle/2\\
&\geqslant \alpha\langle x-y, Ux-Uy\rangle/2\\
&= \alpha D^f(x,y).
\end{aligned} \quad (5.38)$$

∎

Example 5.18 Let $\mathcal{X}$ be a reflexive real Banach space, let $f \in \Gamma_0(\mathcal{X})$ be Gâteaux differentiable on $\operatorname{int}\operatorname{dom} f \neq \varnothing$, let $\alpha \in [0, +\infty[$, and let $g \in \Gamma_0(\mathcal{X})$ be Gâteaux differentiable on $\operatorname{int}\operatorname{dom} g = \operatorname{int}\operatorname{dom} f$. Suppose that and $g - \alpha f$ is convex (which means that g is more convex than αf in the terminology of J. J. Moreau [14]). Then $g \in \mathcal{P}_\alpha(f)$.

Proof. Since f and g are Gâteaux differentiable on $\operatorname{int}\operatorname{dom} f$, $h = g - \alpha f$ is likewise. Furthermore,

$$(\forall x \in \operatorname{dom} f)(\forall y \in \operatorname{int}\operatorname{dom} f) \quad D^g(x, y) - \alpha D^f(x, y) = D^h(x, y) \geqslant 0. \tag{5.39}$$

□

The following definition brings together the notions of Bregman monotone sequences [4] and of variable metric Fejér monotone sequences [9].

Definition 5.19 Let $\mathcal{X}$ be a reflexive real Banach space, let $f \in \Gamma_0(\mathcal{X})$ be Gâteaux differentiable on $\operatorname{int}\operatorname{dom} f \neq \varnothing$, let $(f_n)_{n\in\mathbb{N}}$ be in $\mathcal{F}(f)$, let $(x_n)_{n\in\mathbb{N}} \in (\operatorname{int}\operatorname{dom} f)^{\mathbb{N}}$, and let $C \subset \mathcal{X}$ be such that $C \cap \operatorname{dom} f \neq \varnothing$. Then $(x_n)_{n\in\mathbb{N}}$ is :

(i) *quasi-Bregman monotone* with respect to C relative to $(f_n)_{n\in\mathbb{N}}$ if

$$\begin{aligned}(\exists(\eta_n)_{n\in\mathbb{N}} \in \ell^1_+(\mathbb{N}))(\forall x \in C \cap \operatorname{dom} f)(\exists(\varepsilon_n)_{n\in\mathbb{N}} \in \ell^1_+(\mathbb{N}))(\forall n \in \mathbb{N})\\ D^{f_{n+1}}(x, x_{n+1}) \leqslant (1 + \eta_n) D^{f_n}(x, x_n) + \varepsilon_n;\end{aligned} \tag{5.40}$$

(ii) *stationarily quasi-Bregman monotone* with respect to C relative to $(f_n)_{n\in\mathbb{N}}$ if

$$\begin{aligned}(\exists(\varepsilon_n)_{n\in\mathbb{N}} \in \ell^1_+(\mathbb{N}))(\exists(\eta_n)_{n\in\mathbb{N}} \in \ell^1_+(\mathbb{N}))(\forall x \in C \cap \operatorname{dom} f)(\forall n \in \mathbb{N})\\ D^{f_{n+1}}(x, x_{n+1}) \leqslant (1 + \eta_n) D^{f_n}(x, x_n) + \varepsilon_n.\end{aligned} \tag{5.41}$$

Remark 5.20

(i) In Definition 5.19, suppose that $(\forall n \in \mathbb{N})$ $f_n = f$ and $\eta_n = \varepsilon_n = 0$. Then we recover the notion of a Bregman monotone sequence defined in [4].

(ii) In Definition 5.19, suppose that $\mathcal{X}$ is a Hilbert space, that $f = \|\cdot\|^2/2$, and that $(\forall n \in \mathbb{N})$ $f_n \colon x \mapsto \langle x, U_n x\rangle/2$, where $(U_n)_{n\in\mathbb{N}}$ are operators in $\mathcal{P}_\alpha(\mathcal{X})$ for some $\alpha \in [0, +\infty[$. Then we recover [9, Definition 2.1] with $\phi = |\cdot|^2/2$.

Here are some basic properties of quasi-Bregman monotone sequences.

Proposition 5.21 *Let $\mathcal{X}$ be a reflexive real Banach space, let $f \in \Gamma_0(\mathcal{X})$ be Gâteaux differentiable on $\operatorname{int}\operatorname{dom} f \neq \varnothing$, let $\alpha \in \,]0, +\infty[$, let $(f_n)_{n\in\mathbb{N}}$ be in $\mathcal{P}_\alpha(f)$, let $(x_n)_{n\in\mathbb{N}} \in (\operatorname{int}\operatorname{dom} f)^{\mathbb{N}}$, let $C \subset \mathcal{X}$ be such that $C \cap \operatorname{int}\operatorname{dom} f \neq \varnothing$, and let $x \in C \cap \operatorname{int}\operatorname{dom} f$. Suppose that $(x_n)_{n\in\mathbb{N}}$ is quasi-Bregman monotone with respect to C relative to $(f_n)_{n\in\mathbb{N}}$. Then the following hold :*

(i) $(D^{f_n}(x, x_n))_{n\in\mathbb{N}}$ *converges.*

(ii) *Suppose that* $D^f(x,\cdot)$ *is coercive. Then* $(x_n)_{n\in\mathbb{N}}$ *is bounded.*

Proof. (i) : Let us set $(\forall n\in\mathbb{N})\ \xi_n = D^{f_n}(x,x_n)$. Since $(x_n)_{n\in\mathbb{N}}$ is quasi-Bregman monotone with respect to C relative to $(f_n)_{n\in\mathbb{N}}$, there exist $(\eta_n)_{n\in\mathbb{N}} \in \ell_+^1(\mathbb{N})$ and $(\varepsilon_n)_{n\in\mathbb{N}} \in \ell_+^1(\mathbb{N})$ such that

$$(\forall n\in\mathbb{N})\quad \xi_{n+1} \leqslant (1+\eta_n)\xi_n + \varepsilon_n. \tag{5.42}$$

It therefore follows from [16, Lemma 2.2.2] that $(\xi_n)_{n\in\mathbb{N}}$ converges, i.e., $(D^{f_n}(x,x_n))_{n\in\mathbb{N}}$ converges.

(ii) : Since $(f_n)_{n\in\mathbb{N}}$ is in $\mathcal{P}_\alpha(f)$, we deduce that

$$(\forall n\in\mathbb{N})\quad D^f(x,x_n) \leqslant \alpha^{-1} D^{f_n}(x,x_n). \tag{5.43}$$

Therefore, since (i) implies that $(D^{f_n}(x,x_n))_{n\in\mathbb{N}}$ is bounded, $(D^f(x,x_n))_{n\in\mathbb{N}}$ is bounded. In turn, since $D^f(x,\cdot)$ is coercive, $(x_n)_{n\in\mathbb{N}}$ is bounded. ☐

The following result concerns the weak convergence of quasi-Bregman monotone sequences.

Proposition 5.22 *Let* $\mathcal{X}$ *be a reflexive real Banach space, let* $f \in \Gamma_0(\mathcal{X})$ *be Gâteaux differentiable on* $\operatorname{int}\operatorname{dom} f \neq \varnothing$*, let* $(x_n)_{n\in\mathbb{N}} \in (\operatorname{int}\operatorname{dom} f)^{\mathbb{N}}$*, let* $C \subset \mathcal{X}$ *be such that* $C \cap \operatorname{int}\operatorname{dom} f \neq \varnothing$*, let* $(\eta_n)_{n\in\mathbb{N}} \in \ell_+^1(\mathbb{N})$*, let* $\alpha \in \left]0,+\infty\right[$*, and let* $(f_n)_{n\in\mathbb{N}}$ *in* $\mathcal{P}_\alpha(f)$ *be such that* $(\forall n\in\mathbb{N})\ (1+\eta_n)f_n \succcurlyeq f_{n+1}$*. Suppose that* $(x_n)_{n\in\mathbb{N}}$ *is quasi-Bregman monotone with respect to* C *relative to* $(f_n)_{n\in\mathbb{N}}$*, that there exists* $g \in \mathcal{F}(f)$ *such that for every* $n\in\mathbb{N}$*,* $g \succcurlyeq f_n$*, and that, for every* $y_1 \in \mathcal{X}$ *and every* $y_2 \in \mathcal{X}$*,*

$$\begin{cases} y_1 \in \mathfrak{W}(x_n)_{n\in\mathbb{N}} \cap C \\ y_2 \in \mathfrak{W}(x_n)_{n\in\mathbb{N}} \cap C \\ \big(\langle y_1 - y_2, \nabla f_n(x_n)\rangle\big)_{n\in\mathbb{N}} \quad \text{converges} \end{cases} \quad\Rightarrow\quad y_1 = y_2. \tag{5.44}$$

Moreover, suppose that $(\forall x \in \operatorname{int}\operatorname{dom} f)\ D^f(x,\cdot)$ *is coercive. Then* $(x_n)_{n\in\mathbb{N}}$ *converges weakly to a point in* $C \cap \operatorname{int}\operatorname{dom} f$ *if and only if* $\mathfrak{W}(x_n)_{n\in\mathbb{N}} \subset C \cap \operatorname{int}\operatorname{dom} f$.

Proof. Necessity is clear. To show sufficiency, suppose that every weak sequential cluster point of $(x_n)_{n\in\mathbb{N}}$ is in $C\cap\operatorname{int}\operatorname{dom} f$ and let y_1 and y_2 be two such points. First, it follows from Proposition 5.21(i) that

$$\big(D^{f_n}(y_1,x_n)\big)_{n\in\mathbb{N}} \quad\text{and}\quad \big(D^{f_n}(y_2,x_n)\big)_{n\in\mathbb{N}} \quad \text{are convergent.} \tag{5.45}$$

Next, let us define the following functions

$$\phi\colon\ [0,1] \to \mathbb{R}\colon t \mapsto \big\langle y_1 - y_2, \nabla g(y_2 + t(y_1-y_2)) - \nabla g(y_2)\big\rangle, \tag{5.46}$$

and

$$(\forall n\in\mathbb{N})\quad \phi_n\colon\ [0,1] \to \mathbb{R}\colon t \mapsto \big\langle y_1 - y_2, \nabla f_n(y_2 + t(y_1-y_2)) - \nabla f_n(y_2)\big\rangle. \tag{5.47}$$

Then

$$\int_0^1 \phi(t)dt = g(y_1) - g(y_2) \quad \text{and} \quad (\forall n \in \mathbb{N}) \quad \int_0^1 \phi_n(t)dt = f_n(y_1) - f_n(y_2). \tag{5.48}$$

For every $n \in \mathbb{N}$, since $(1+\eta_n)f_n \succcurlyeq f_{n+1}$, for every $t \in \left]0,1\right]$, we have

$$\begin{aligned}
\phi_{n+1}(t) &= \big\langle y_1 - y_2, \nabla f_{n+1}(y_2 + t(y_1 - y_2)) - \nabla f_{n+1}(y_2)\big\rangle \\
&= t^{-1}\big\langle y_2 + t(y_1 - y_2) - y_2, \nabla f_{n+1}(y_2 + t(y_1 - y_2)) - \nabla f_{n+1}(y_2)\big\rangle \\
&= t^{-1}\big(D^{f_{n+1}}\big(y_2 + t(y_1 - y_2), y_2\big) + D^{f_{n+1}}\big(y_2, y_2 + t(y_1 - y_2)\big)\big) \\
&\leqslant (1+\eta_n)t^{-1}\big(D^{f_n}\big(y_2 + t(y_1 - y_2), y_2\big) + D^{f_n}\big(y_2, y_2 + t(y_1 - y_2)\big)\big) \\
&= (1+\eta_n)t^{-1}\big\langle y_2 + t(y_1 - y_2) - y_2, \nabla f_n(y_2 + t(y_1 - y_2)) - \nabla f_n(y_2)\big\rangle \\
&= (1+\eta_n)\big\langle y_1 - y_2, \nabla f_n(y_2 + t(y_1 - y_2)) - \nabla f_n(y_2)\big\rangle \\
&= (1+\eta_n)\phi_n(t).
\end{aligned} \tag{5.49}$$

Consequently,

$$(\forall n \in \mathbb{N})(\forall t \in \left]0,1\right]) \quad 0 \leqslant \phi_{n+1}(t) \leqslant (1+\eta_n)\phi_n(t). \tag{5.50}$$

It is clear that (5.50) is valid for $t = 0$ since in this case, all terms are equal to 0. In turn, we deduce from [16, Lemma 2.2.2] that

$$(\phi_n)_{n\in\mathbb{N}} \quad \text{converges pointwise.} \tag{5.51}$$

On the other hand, for every $n \in \mathbb{N}$, since $g \succcurlyeq f_n$, the same argument as above shows that

$$(\forall t \in [0,1]) \quad 0 \leqslant \phi_n(t) \leqslant \phi(t). \tag{5.52}$$

By invoking (5.51), (5.52), and Lebesgue's dominated convergence theorem, we obtain that

$$\left(\int_0^1 \phi_n(t)dt\right)_{n\in\mathbb{N}} \quad \text{converges,} \tag{5.53}$$

which implies that

$$\Big(f_n(y_1) - f_n(y_2)\Big)_{n\in\mathbb{N}} \quad \text{converges.} \tag{5.54}$$

We also observe that

$$(\forall n \in \mathbb{N}) \quad D^{f_n}(y_1, x_n) - D^{f_n}(y_2, x_n) = f_n(y_1) - f_n(y_2) - \langle y_1 - y_2, \nabla f_n(x_n)\rangle, \tag{5.55}$$

and hence, it follows from (5.45) and (5.54) that

$$\big(\langle y_1 - y_2, \nabla f_n(x_n)\rangle\big)_{n\in\mathbb{N}} \quad \text{converges.} \tag{5.56}$$

In turn, (5.44) forces $y_1 = y_2$. Since Proposition 5.21(ii) asserts that $(x_n)_{n\in\mathbb{N}}$ is bounded and since $\mathcal{X}$ is reflexive, we conclude that $x_n \rightharpoonup y_1 \in C \cap \operatorname{int}\operatorname{dom} f$. □

Example 5.23 Let $\mathcal{X}$ be a reflexive real Banach space, let $f \in \Gamma_0(\mathcal{X})$ be Gâteaux differentiable on $\operatorname{int}\operatorname{dom} f \neq \varnothing$, let $(f_n)_{n\in\mathbb{N}}$ be in $\mathfrak{F}(f)$, let $(x_n)_{n\in\mathbb{N}} \in (\operatorname{int}\operatorname{dom} f)^{\mathbb{N}}$, and let $C \subset \mathcal{X}$. Suppose that $C \cap \overline{\operatorname{dom}} f$ is a singleton. Then (5.44) is satisfied.

Proof. Since $(x_n)_{n\in\mathbb{N}} \in (\operatorname{int}\operatorname{dom} f)^{\mathbb{N}}$, $\mathfrak{W}(x_n)_{n\in\mathbb{N}} \subset \overline{\operatorname{dom}} f$, and therefore, $\mathfrak{W}(x_n)_{n\in\mathbb{N}} \cap C$ is at most a singleton. □

Example 5.24 Let $\mathcal{X}$ be a reflexive real Banach space, let $f \in \Gamma_0(\mathcal{X})$ be Gâteaux differentiable on $\operatorname{int}\operatorname{dom} f \neq \varnothing$, let $(x_n)_{n\in\mathbb{N}} \in (\operatorname{int}\operatorname{dom} f)^{\mathbb{N}}$, let $C \subset \operatorname{int}\operatorname{dom} f$, and set $(\forall n \in \mathbb{N})\ f_n = f$. Suppose that $f|_{\operatorname{int}\operatorname{dom} f}$ is strictly convex and that ∇f is weakly sequentially continuous. Then (5.44) is satisfied.

Proof. Suppose that $y_1 \in \mathfrak{W}(x_n)_{n\in\mathbb{N}} \cap C$ and $y_2 \in \mathfrak{W}(x_n)_{n\in\mathbb{N}} \cap C$ are such that $(\langle y_1 - y_2, \nabla f_n(x_n)\rangle)_{n\in\mathbb{N}}$ converges and $y_1 \neq y_2$. Take strictly increasing sequences $(k_n)_{n\in\mathbb{N}}$ and $(l_n)_{n\in\mathbb{N}}$ in $\mathbb{N}$ such that $x_{k_n} \rightharpoonup y_1$ and $x_{l_n} \rightharpoonup y_2$. Since ∇f is weakly sequentially continuous, by taking the limit in (5.44) along subsequences $(x_{k_n})_{n\in\mathbb{N}}$ and $(x_{l_n})_{n\in\mathbb{N}}$, we get

$$\langle y_1 - y_2, \nabla f(y_1) - \nabla f(y_2)\rangle = 0 \tag{5.57}$$

Since $f|_{\operatorname{int}\operatorname{dom} f}$ is strictly convex, ∇f is strictly monotone [19, Theorem 2.4.4(ii)], i.e.,

$$\langle y_1 - y_2, \nabla f(y_1) - \nabla f(y_2)\rangle > 0, \tag{5.58}$$

and we reach a contradiction. □

Example 5.25 Let $\mathcal{X}$ be a real Hilbert space, let $f = \|\cdot\|^2/2$, let $C \subset \mathcal{X}$, let $(x_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{X}$, let $\alpha \in \left]0, +\infty\right[$, let U and $(U_n)_{n\in\mathbb{N}}$ be self-adjoint linear operators from $\mathcal{X}$ in $\mathcal{X}$ such that $U_n \to U$ pointwise, and set $(\forall n \in \mathbb{N})\ f_n = \langle \cdot, U_n \cdot\rangle/2$. Suppose that $\langle \cdot, U\cdot\rangle \geqslant \alpha\|\cdot\|^2$. Then (5.44) is satisfied.

Proof. It is easy to see that, for every $n \in \mathbb{N}$, f_n is Gâteaux differentiable on $\mathcal{X}$ with $\nabla f_n = U_n$. Suppose that $y_1 \in \mathfrak{W}(x_n)_{n\in\mathbb{N}} \cap C$ and $y_2 \in \mathfrak{W}(x_n)_{n\in\mathbb{N}} \cap C$ are such that $(\langle y_1 - y_2, \nabla f_n(x_n)\rangle)_{n\in\mathbb{N}}$ converges. Take strictly increasing sequences $(k_n)_{n\in\mathbb{N}}$ and $(l_n)_{n\in\mathbb{N}}$ in $\mathbb{N}$ such that $x_{k_n} \rightharpoonup y_1$ and $x_{l_n} \rightharpoonup y_2$. We have

$$\big\langle y_1 - y_2, \nabla f_{k_n}(x_{k_n})\big\rangle = \big\langle y_1 - y_2, U_{k_n} x_{k_n}\big\rangle = \big\langle U_{k_n} y_1 - U_{k_n} y_2, x_{k_n}\big\rangle \to \big\langle U y_1 - U y_2, y_1\big\rangle, \tag{5.59}$$

and

$$\big\langle y_1 - y_2, \nabla f_{l_n}(x_{l_n})\big\rangle = \big\langle y_1 - y_2, U_{l_n} x_{l_n}\big\rangle = \big\langle U_{l_n} y_1 - U_{l_n} y_2, x_{l_n}\big\rangle \to \big\langle U y_1 - U y_2, y_2\big\rangle, \tag{5.60}$$

and hence, $0 = \big\langle U y_1 - U y_2, y_1 - y_2\big\rangle \geqslant \alpha\|y_1 - y_2\|^2$, and therefore, $y_1 = y_2$. □

The following condition will be used subsequently (see [4, Examples 4.10, 5.11, and 5.13] for special cases).

Condition 5.26 [4, Condition 4.4] Let $\mathcal{X}$ be a reflexive real Banach space and let $f \in \Gamma_0(\mathcal{X})$ be Gâteaux differentiable on $\operatorname{int} \operatorname{dom} f \neq \varnothing$. For every bounded sequences $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ in $\operatorname{int} \operatorname{dom} f$,

$$D^f(x_n, y_n) \to 0 \quad \Rightarrow \quad x_n - y_n \to 0. \tag{5.61}$$

We now present a characterization of the strong convergence of stationarily quasi-Bregman monotone sequences.

Proposition 5.27 *Let $\mathcal{X}$ be a reflexive real Banach space, let $f \in \Gamma_0(\mathcal{X})$ be a Legendre function, let $\alpha \in \left]0, +\infty\right[$, let $(f_n)_{n\in\mathbb{N}}$ be in $\mathcal{P}_\alpha(f)$, let $(x_n)_{n\in\mathbb{N}} \in (\operatorname{int} \operatorname{dom} f)^{\mathbb{N}}$, and let C be a closed convex subset of $\mathcal{X}$ such that $C \cap \operatorname{int} \operatorname{dom} f \neq \varnothing$. Suppose that $(x_n)_{n\in\mathbb{N}}$ is stationarily quasi Bregman monotone with respect to C relative to $(f_n)_{n\in\mathbb{N}}$, that f satisfies Condition 5.26, and that $(\forall x \in \operatorname{int} \operatorname{dom} f)$ $D^f(x, \cdot)$ is coercive. In addition, suppose that there exists $\beta \in \left]0, +\infty\right[$ such that $(\forall n \in \mathbb{N})$ $\beta f \succcurlyeq f_n$. Then $(x_n)_{n\in\mathbb{N}}$ converges strongly to a point in $C \cap \overline{\operatorname{dom}} f$ if and only if $\underline{\lim}\, D_C^f(x_n) = 0$.*

Proof. To show the necessity, suppose that $x_n \to \overline{x} \in C \cap \overline{\operatorname{dom}} f$ and take $x \in C \cap \operatorname{int} \operatorname{dom} f$. Since Proposition 5.21(i) states that $(D^{f_n}(x, x_n))_{n\in\mathbb{N}}$ is bounded and since

$$(\forall n \in \mathbb{N}) \quad D^f(x, x_n) \leqslant D^{f_n}(x, x_n), \tag{5.62}$$

we deduce that $(D^f(x, x_n))_{n\in\mathbb{N}}$ is bounded. However, by [3, Lemma 7.3(vii)],

$$(\forall n \in \mathbb{N}) \quad D^{f^*}\big(\nabla f(x_n), \nabla f(x)\big) = D^f(x, x_n). \tag{5.63}$$

Therefore $(D^{f^*}(\nabla f(x_n), \nabla f(x)))_{n\in\mathbb{N}}$ is bounded. In turn, since $D^{f^*}(\cdot, \nabla f(x))$ is coercive [3, Lemma 7.3(v)], we get $(\nabla f(x_n))_{n\in\mathbb{N}}$ is bounded and hence $\langle \overline{x} - x_n, \nabla f(x_n)\rangle \to 0$. Since

$$\begin{aligned}(\forall n \in \mathbb{N}) \quad D_C^f(x_n) &= \inf D^f(C, x_n)\\ &\leqslant \inf D^f(C \cap \overline{\operatorname{dom}} f, x_n)\\ &\leqslant D^f(\overline{x}, x_n)\\ &= f(\overline{x}) - f(x_n) - \big\langle \overline{x} - x_n, \nabla f(x_n)\big\rangle,\end{aligned} \tag{5.64}$$

we obtain

$$\underline{\lim}\, D_C^f(x_n) \leqslant f(\overline{x}) - \overline{\lim}\, f(x_n) - \lim \big\langle \overline{x} - x_n, \nabla f(x_n)\big\rangle = f(\overline{x}) - \overline{\lim}\, f(x_n). \tag{5.65}$$

Since f is lower semicontinuous,

$$f(\overline{x}) \leqslant \underline{\lim}\, f(x_n) \leqslant \overline{\lim}\, f(x_n). \tag{5.66}$$

Altogether, (5.65) and (5.66) yield

$$\underline{\lim}\, D_C^f(x_n) \to 0. \tag{5.67}$$

We now show the sufficiency. First, since f is Legendre and $C \cap \operatorname{int} \operatorname{dom} f \neq \varnothing$, (5.31) yields

$$P_C^f \colon \operatorname{int} \operatorname{dom} f \to C \cap \operatorname{int} \operatorname{dom} f. \tag{5.68}$$

Next, we set

$$(\forall n \in \mathbb{N}) \quad \varrho_n = D_C^f(x_n) \quad \text{and} \quad \zeta_n = \inf_{x \in C \cap \operatorname{dom} f} D^{f_n}(x, x_n). \tag{5.69}$$

Then $\underline{\lim}\, \varrho_n = 0$. For every $n \in \mathbb{N}$, since $\beta f \succcurlyeq f_n \succcurlyeq \alpha f$, we obtain

$$(\forall x \in C \cap \operatorname{dom} f) \quad 0 \leqslant \alpha D^f(x, x_n) \leqslant D^{f_n}(x, x_n) \leqslant \beta D^f(x, x_n). \tag{5.70}$$

In the above inequalities, after taking the infimum over $x \in C \cap \operatorname{dom} f$, we get

$$(\forall n \in \mathbb{N}) \quad 0 \leqslant \alpha \varrho_n \leqslant \zeta_n \leqslant \beta \varrho_n \tag{5.71}$$

and therefore,

$$0 \leqslant \alpha\, \underline{\lim}\, \varrho_n \leqslant \underline{\lim}\, \zeta_n \leqslant \beta\, \underline{\lim}\, \varrho_n = 0. \tag{5.72}$$

On the other hand, since $(x_n)_{n \in \mathbb{N}}$ is stationarily quasi Bregman monotone with respect to C relative to $(f_n)_{n \in \mathbb{N}}$, there exist $(\eta_n)_{n \in \mathbb{N}} \in \ell_+^1(\mathbb{N})$ and $(\varepsilon_n)_{n \in \mathbb{N}} \in \ell_+^1(\mathbb{N})$ such that

$$(\forall x \in C \cap \operatorname{dom} f)(\forall n \in \mathbb{N}) \quad D^{f_{n+1}}(x, x_{n+1}) \leqslant (1 + \eta_n) D^{f_n}(x, x_n) + \varepsilon_n. \tag{5.73}$$

Taking the infimum in (5.73) over $C \cap \operatorname{dom} f$ yields

$$(\forall n \in \mathbb{N}) \quad \zeta_{n+1} \leqslant (1 + \eta_n)\zeta_n + \varepsilon_n. \tag{5.74}$$

It therefore follows from [16, Lemma 2.2.2] that $(\zeta_n)_{n \in \mathbb{N}}$ converges, and thus, we deduce from (5.72) that $\zeta_n \to 0$. Appealing to (5.71), we get $\varrho_n \to 0$, i.e.,

$$D^f\big(P_C^f x_n, x_n\big) \to 0. \tag{5.75}$$

Now let $x \in C \cap \operatorname{int} \operatorname{dom} f$. Then $x \in \operatorname{Fix} P_C^f$ [4, Proposition 3.22(ii)(b)] and it follows from Proposition 5.21(i) that $(D^{f_n}(x, x_n))_{n \in \mathbb{N}}$ is bounded, and hence, $(D^f(x, x_n))_{n \in \mathbb{N}}$ is likewise. In turn, since [4, Proposition 3.3(i) and Theorem 3.34] yield

$$(\forall n \in \mathbb{N}) \quad D^f\big(x, P_C^f x_n\big) \leqslant D^f(x, x_n), \tag{5.76}$$

we deduce that $(D^f(x, P_C^f x_n))_{n \in \mathbb{N}}$ is bounded, and hence, since $D^f(x, \cdot)$ is coercive, we obtain that

$$\big(P_C^f x_n\big)_{n \in \mathbb{N}} \in (\operatorname{int} \operatorname{dom} f)^{\mathbb{N}} \ \text{ is bounded}. \tag{5.77}$$

Therefore, since f satisfies Condition 5.26, it follows from (5.75) that

$$P_C^f x_n - x_n \to 0. \tag{5.78}$$

Since (5.68) entails that

$$(\forall n \in \mathbb{N}) \quad P_C^f x_n \in C \cap \operatorname{int} \operatorname{dom} f = \operatorname{Fix} P_C^f, \tag{5.79}$$

we obtain

$$(\forall n \in \mathbb{N}) \quad 0 \leqslant d_C(x_n) = \inf_{x \in C} \|x - x_n\| \leqslant \|P_C^f x_n - x_n\|. \tag{5.80}$$

Altogether, (5.78) and (5.80) imply that

$$d_C(x_n) \to 0. \tag{5.81}$$

Set $\tau = \prod_{k \in \mathbb{N}}(1 + \eta_k)$. Then $\tau < +\infty$ [12, Theorem 3.7.3]. By invoking (5.79) and [4, Proposition 3.3(i) and Theorem 3.34], we get

$$\begin{aligned}(\forall n \in \mathbb{N})(\forall m \in \mathbb{N}) \quad D^f\big(P_C^f x_n, P_C^f x_{m+n}\big) &\leqslant D^f\big(P_C^f x_n, x_{m+n}\big)\\ &\leqslant \alpha^{-1} D^{f_{m+n}}\big(P_C^f x_n, x_{m+n}\big)\\ &\leqslant \tau\alpha^{-1}\left(D^{f_n}\big(P_C^f x_n, x_n\big) + \sum_{k=n}^{n+m-1} \varepsilon_k\right)\\ &\leqslant \tau\alpha^{-1}\left(\beta D^f\big(P_C^f x_n, x_n\big) + \sum_{k \geqslant n} \varepsilon_k\right)\\ &= \tau\alpha^{-1}\left(\beta \varrho_n + \sum_{k \geqslant n} \varepsilon_k\right).\end{aligned} \tag{5.82}$$

After taking the limit as $n \to +\infty$ and $m \to +\infty$ in (5.82), we obtain

$$D^f\big(P_C^f x_{m+n}, P_C^f x_n\big) \to 0, \tag{5.83}$$

and thus (5.77) yield

$$P_C^f x_{m+n} - P_C^f x_n \to 0. \tag{5.84}$$

However,

$$(\forall n \in \mathbb{N})(\forall m \in \mathbb{N}) \quad \|x_{m+n} - x_n\| \leqslant \|x_{m+n} - P_C^f x_{m+n}\| + \|P_C^f x_{m+n} - P_C^f x_n\| + \|P_C^f x_n - x_n\|. \tag{5.85}$$

After taking the limit as $n \to +\infty$ and $m \to +\infty$ in (5.85) then using (5.78) and (5.84), we get

$$\|x_{n+m} - x_n\| \to 0. \tag{5.86}$$

Thus, $(x_n)_{n \in \mathbb{N}}$ is a Cauchy sequence in $\mathcal{X}$, and hence, there exists $\overline{x} \in \mathcal{X}$ such that $x_n \to \overline{x}$. By (5.81) and the continuity of d_C [5, Example 1.47], we obtain $d_C(\overline{x}) = 0$ and, since C is closed, $\overline{x} \in C$. Because $(x_n)_{n \in \mathbb{N}}$ is in $\operatorname{int} \operatorname{dom} f$, we conclude that $\overline{x} \in \overline{\operatorname{dom}} f$. □

Remark 5.28 In Proposition 5.27, suppose that $\mathcal{X}$ is a Hilbert space, that $f = \|\cdot\|^2/2$, and that $(\forall n \in \mathbb{N})$ $f_n \colon x \mapsto \langle x, U_n x\rangle/2$, where $(U_n)_{n\in\mathbb{N}}$ are operators in $\mathcal{P}_\alpha(\mathcal{X})$ such that $\sup_{n\in\mathbb{N}} \|U_n\| < +\infty$. Then we recover [9, Theorem 3.4] with $\phi = |\cdot|^2/2$.

5.2.3 Bregman distance-based proximity operators

Many algorithms in optimization in a real Hilbert space $\mathcal{H}$ are based on Moreau's proximity operator [13] of a function $\varphi \in \Gamma_0(\mathcal{H})$

$$\operatorname{prox}_\varphi \colon \mathcal{H} \to \mathcal{H} \colon x \mapsto \operatorname{argmin}\big(\varphi + \|\cdot - x\|^2/2\big). \tag{5.87}$$

Because the quadratic term in (5.87) is difficult to manipulate in Banach spaces since its gradient is nonlinear, alternative notions based on Bregman distances have been used (see [4] and the references therein). This leads to the notion of D^f-proximal operators. In this section, we investigate some their basic properties.

Lemma 5.29 [4, Section 3] *Let $\mathcal{X}$ be a reflexive real Banach space, let $\varphi \in \Gamma_0(\mathcal{X})$ be bounded from below, and let $f \in \Gamma_0(\mathcal{X})$ be a Legendre function such that* $\operatorname{dom}\varphi \cap \operatorname{int}\operatorname{dom} f \neq \varnothing$. *Then the following hold :*

(i) $\operatorname{prox}^f_\varphi$ *is single-valued on its domain.*

(ii) $\operatorname{ran}\operatorname{prox}^f_\varphi \subset \operatorname{dom}\operatorname{prox}^f_\varphi = \operatorname{int}\operatorname{dom} f$.

(iii) $\operatorname{prox}^f_\varphi = (\nabla f + \partial\varphi)^{-1} \circ \nabla f$.

(iv) $\operatorname{Fix}\operatorname{prox}^f_\varphi = \operatorname{Argmin}\varphi \cap \operatorname{int}\operatorname{dom} f$.

(v) *Let* $x \in \operatorname{Argmin}\varphi \cap \operatorname{int}\operatorname{dom} f$, *let* $y \in \operatorname{int}\operatorname{dom} f$, *and let* $v = \operatorname{prox}^f_\varphi y$. *Then*

$$D^f(x, v) + D^f(v, y) \leqslant D^f(x, y). \tag{5.88}$$

The following result in an extension of [5, Proposition 23.30].

Proposition 5.30 *Let m be a strictly positive integer, let $(\mathcal{X}_i)_{1\leqslant i\leqslant m}$ be reflexive real Banach spaces, and let $\mathcal{X}$ be the vector product space $\bigtimes_{i=1}^m \mathcal{X}_i$ equipped with the norm $x = (x_i)_{1\leqslant i\leqslant m} \mapsto \sqrt{\sum_{i=1}^m \|x_i\|^2}$. For every $i \in \{1, \ldots, m\}$, let $\varphi_i \in \Gamma_0(\mathcal{X}_i)$ be bounded from below and let $f_i \in \Gamma_0(\mathcal{X}_i)$ be a Legendre function such that* $\operatorname{dom}\varphi_i \cap \operatorname{int}\operatorname{dom} f_i \neq \varnothing$. *Set $f \colon \mathcal{X} \to \left]-\infty, +\infty\right] \colon x \mapsto \sum_{i=1}^m f_i(x_i)$ and $\varphi \colon \mathcal{X} \to \left]-\infty, +\infty\right] \colon x \mapsto \sum_{i=1}^m \varphi_i(x_i)$. Then*

$$\Big(\forall x \in \bigtimes_{i=1}^m \operatorname{int}\operatorname{dom} f_i\Big) \quad \operatorname{prox}^f_\varphi x = \big(\operatorname{prox}^{f_i}_{\varphi_i} x_i\big)_{1\leqslant i\leqslant m}. \tag{5.89}$$

Proof. First, we observe that $\mathcal{X}^*$ is the vector product space $\bigtimes_{i=1}^m \mathcal{X}_i^*$ equipped with the norm $x^* = (x_i^*)_{1\leqslant i\leqslant m} \mapsto \sqrt{\sum_{i=1}^m \|x_i^*\|^2}$. Since, for every $i \in \{1, \ldots, m\}$, φ_i is bounded from below, so is φ. Next, we derive from the definition of f that $\operatorname{dom} f = \bigtimes_{i=1}^{m} \operatorname{dom} f_i$ and that

$$\partial f \colon \mathcal{X} \to 2^{\mathcal{X}^*} \colon (x_i)_{1 \leqslant i \leqslant m} \mapsto \bigtimes_{i=1}^{m} \partial f_i(x_i). \tag{5.90}$$

Thus, ∂f is single-valued on

$$\operatorname{dom} \partial f = \bigtimes_{i=1}^{m} \operatorname{dom} \partial f_i = \bigtimes_{i=1}^{m} \operatorname{int} \operatorname{dom} f_i = \operatorname{int} \Big(\bigtimes_{i=1}^{m} \operatorname{dom} f_i \Big) = \operatorname{int} \operatorname{dom} f. \tag{5.91}$$

Likewise, since

$$f^* \colon \mathcal{X}^* \to \left]-\infty, +\infty\right] \colon (x_i^*)_{1 \leqslant i \leqslant m} \mapsto \sum_{i=1}^{m} f_i^*(x_i^*), \tag{5.92}$$

we deduce that ∂f^* is single-valued on $\operatorname{dom} \partial f^* = \operatorname{int} \operatorname{dom} f^*$. Consequently, [3, Theorems 5.4 and 5.6] assert that f is a Legendre function. In addition,

$$\operatorname{dom} \varphi \cap \operatorname{int} \operatorname{dom} f = \Big(\bigtimes_{i=1}^{m} \operatorname{dom} \varphi_i \Big) \cap \Big(\bigtimes_{i=1}^{m} \operatorname{int} \operatorname{dom} f_i \Big) = \bigtimes_{i=1}^{m} (\operatorname{dom} \varphi_i \cap \operatorname{int} \operatorname{dom} f_i) \neq \varnothing. \tag{5.93}$$

Now Lemma 5.29 asserts that $\operatorname{prox}_{\varphi}^{f} \colon \operatorname{int} \operatorname{dom} f \to \operatorname{dom} \varphi \cap \operatorname{int} \operatorname{dom} f$. For the remainder of the proof, let $x \in \operatorname{int} \operatorname{dom} f$, set $p = \operatorname{prox}_{\varphi}^{f} x$, and set $q = (\operatorname{prox}_{\varphi_i}^{f_i} x_i)_{1 \leqslant i \leqslant m}$. Since Lemma 5.29(iii) yields $\nabla f(x) - \nabla f(p) \in \partial \varphi(p)$, we deduce from (5.28) that

$$(\forall z \in \operatorname{dom} \varphi) \quad \langle z - p, \nabla f(x) - \nabla f(p) \rangle + \varphi(p) \leqslant \varphi(z). \tag{5.94}$$

Setting $z = q$ in (5.94) yields

$$\langle q - p, \nabla f(x) - \nabla f(p) \rangle + \varphi(p) \leqslant \varphi(q). \tag{5.95}$$

For every $i \in \{1, \ldots, m\}$, set $q_i = \operatorname{prox}_{\varphi_i}^{f_i} x_i$. The same characterization as in (5.94) yields

$$(\forall i \in \{1, \ldots, m\})(\forall z_i \in \operatorname{dom} \varphi_i) \quad \langle z_i - q_i, \nabla f_i(x_i) - \nabla f_i(q_i) \rangle + \varphi_i(q_i) \leqslant \varphi_i(z_i). \tag{5.96}$$

By summing these inequalities over $i \in \{1, \ldots, m\}$, we obtain

$$(\forall z \in \operatorname{dom} \varphi) \quad \langle z - q, \nabla f(x) - \nabla f(q) \rangle + \varphi(q) \leqslant \varphi(z). \tag{5.97}$$

Upon setting $z = p$ in (5.97), we get

$$\langle p - q, \nabla f(x) - \nabla f(q) \rangle + \varphi(q) \leqslant \varphi(p). \tag{5.98}$$

Adding (5.95) and (5.98) yields

$$\langle p - q, \nabla f(p) - \nabla f(q) \rangle \leqslant 0. \tag{5.99}$$

Suppose that $p \neq q$. Since f is essentially strictly convex, f is strictly convex on every convex subset of $\operatorname{dom} \partial f$. In particular, since $\operatorname{int} \operatorname{dom} f \subset \operatorname{dom} \partial f$, $f|_{\operatorname{int} \operatorname{dom} f}$ is strictly convex. Hence, by [19, Theorem 2.4.4(ii)], ∇f is strictly monotone, i.e.,

$$\langle p - q, \nabla f(p) - \nabla f(q) \rangle > 0, \tag{5.100}$$

and we reach a contradiction. Consequently, $p = q$ which proves the claim. □

Let us note that, even in Euclidean spaces, it may be easier to evaluate $\operatorname{prox}^f_\varphi$ than Moreau's usual proximity operator $\operatorname{prox}_\varphi$, which is based on $f = \|\cdot\|^2/2$. We provide illustrations of such instances in the standard Euclidean space $\mathbb{R}^m$.

Example 5.31 Let $\gamma \in \,]0, +\infty[$, let $\phi \in \Gamma_0(\mathbb{R})$ be such that $\operatorname{dom} \phi \cap \,]0, +\infty[\, \neq \varnothing$, and let ϑ be Boltzmann-Shannon entropy, i.e.,

$$\vartheta \colon \xi \mapsto \begin{cases} \xi \ln \xi - \xi, & \text{if } \ \xi \in \,]0, +\infty[\,; \\ 0, & \text{if } \ \xi = 0; \\ +\infty, & \text{otherwise.} \end{cases} \tag{5.101}$$

Set $\varphi \colon (\xi_i)_{1 \leqslant i \leqslant m} \mapsto \sum_{i=1}^m \phi(\xi_i)$ and $f \colon (\xi_i)_{1 \leqslant i \leqslant m} \mapsto \sum_{i=1}^m \vartheta(\xi_i)$. Note that f is a Legendre function [2, Theorem 5.12 and Example 6.5] and hence, Lemma 5.29 asserts that $\operatorname{dom} \operatorname{prox}^f_{\gamma\varphi} = \,]0, +\infty[^m$. Let $(\xi_i)_{1 \leqslant i \leqslant m} \in \,]0, +\infty[^m$, set $(\eta_i)_{1 \leqslant i \leqslant m} = \operatorname{prox}^f_{\gamma\varphi}(\xi_i)_{1 \leqslant i \leqslant m}$, let W be the Lambert function [10], i.e., the inverse of $\xi \mapsto \xi e^\xi$ on $[0, +\infty[$, and let $i \in \{1, \ldots, m\}$. Then η_i can be computed as follows.

(i) Let $\omega \in \mathbb{R}$ and suppose that

$$\phi \colon \xi \mapsto \begin{cases} \xi \ln \xi - \omega\xi, & \text{if } \ \xi \in \,]0, +\infty[\,; \\ 0, & \text{if } \ \xi = 0; \\ +\infty, & \text{otherwise.} \end{cases} \tag{5.102}$$

Then $\eta_i = \xi_i^{(\omega - 1)/(\gamma + 1)}$.

(ii) Let $p \in [1, +\infty[$ and suppose that either $\phi = |\cdot|^p/p$ or

$$\phi \colon \xi \mapsto \begin{cases} \xi^p/p, & \text{if } \ \xi \in [0, +\infty[\,; \\ +\infty, & \text{otherwise.} \end{cases} \tag{5.103}$$

Then

$$\eta_i = \begin{cases} \left(\dfrac{W(\gamma(p-1)\xi_i^{p-1})}{\gamma(p-1)} \right)^{\frac{1}{p-1}}, & \text{if } \ p \in \,]1, +\infty[\,; \\ \xi_i e^{-\gamma}, & \text{if } \ p = 1. \end{cases} \tag{5.104}$$

(iii) Let $p \in [1,+\infty[$ and suppose that

$$\phi\colon \xi \mapsto \begin{cases} \xi^{-p}/p, & \text{if } \xi \in \left]0,+\infty\right[;\\ +\infty, & \text{otherwise.}\end{cases} \tag{5.105}$$

Then

$$\eta_i = \left(\frac{W(\gamma(p+1)\xi_i^{-p-1})}{\gamma(p+1)}\right)^{\frac{-1}{p+1}}. \tag{5.106}$$

(iv) Let $p \in \left]0,1\right[$ and suppose that

$$\phi\colon \xi \mapsto \begin{cases} -\xi^{p}/p, & \text{if } \xi \in \left[0,+\infty\right[;\\ +\infty, & \text{otherwise.}\end{cases} \tag{5.107}$$

Then

$$\eta_i = \left(\frac{W(\gamma(1-p)\xi_i^{p-1})}{\gamma(1-p)}\right)^{\frac{1}{p-1}}. \tag{5.108}$$

Example 5.32 Let $\phi \in \Gamma_0(\mathbb{R})$ be such that $\operatorname{dom}\phi \cap \left]0,1\right[\neq \varnothing$ and let ϑ be Fermi-Dirac entropy, i.e.,

$$\vartheta\colon \xi \mapsto \begin{cases} \xi\ln\xi - (1-\xi)\ln(1-\xi), & \text{if } \xi \in \left]0,1\right[;\\ 0 & \text{if } \xi \in \{0,1\};\\ +\infty, & \text{otherwise.}\end{cases} \tag{5.109}$$

Set $\varphi\colon (\xi_i)_{1\leqslant i\leqslant m} \mapsto \sum_{i=1}^m \phi(\xi_i)$ and $f\colon (\xi_i)_{1\leqslant i\leqslant m} \mapsto \sum_{i=1}^m \vartheta(\xi_i)$. Note that f is a Legendre function [2, Theorem 5.12 and Example 6.5] and hence, Lemma 5.29 asserts that $\operatorname{dom}\operatorname{prox}_\varphi^f = \left]0,1\right[^m$. Let $(\xi_i)_{1\leqslant i\leqslant m} \in \left]0,1\right[^m$, set $(\eta_i)_{1\leqslant i\leqslant m} = \operatorname{prox}_\varphi^f(\xi_i)_{1\leqslant i\leqslant m}$, and let $i \in \{1,\ldots,m\}$. Then η_i can be computed as follows.

(i) Let $\omega \in \mathbb{R}$ and suppose that

$$\phi\colon \xi \mapsto \begin{cases} \xi\ln\xi - \omega\xi, & \text{if } \xi \in \left]0,+\infty\right[;\\ 0, & \text{if } \xi = 0;\\ +\infty, & \text{otherwise.}\end{cases} \tag{5.110}$$

Then $\eta_i = e^{\omega}(2-2\xi_i)^{-1}(-\xi_i + \sqrt{4\xi_i - 3\xi_i^2})$.

(ii) Suppose that

$$\phi\colon \xi \mapsto \begin{cases} (1-\xi)\ln(1-\xi)+\xi, & \text{if } \xi \in \left]-\infty,1\right[;\\ 1 & \text{if } \xi = 1;\\ +\infty, & \text{otherwise.}\end{cases} \tag{5.111}$$

Then $\eta_i = 1/2 + \xi_i^{-1}/2 - \sqrt{\xi_i^{-2}/4 + \xi_i^{-1}/2 - 3/4}$.

Example 5.33 Let $\phi \in \Gamma_0(\mathbb{R})$ be such that $\operatorname{dom}\phi \cap \left]0,+\infty\right[\neq \varnothing$ and let ϑ be Burg entropy, i.e.,

$$\vartheta\colon \xi \mapsto \begin{cases} -\ln \xi, & \text{if } \xi \in \left]0,+\infty\right[;\\ +\infty, & \text{otherwise.}\end{cases} \tag{5.112}$$

Set $\varphi\colon (\xi_i)_{1\leqslant i\leqslant m} \mapsto \sum_{i=1}^m \phi(\xi_i)$ and $f\colon (\xi_i)_{1\leqslant i\leqslant m} \mapsto \sum_{i=1}^m \vartheta(\xi_i)$. Note that f is a Legendre function [2, Theorem 5.12 and Example 6.5] and hence, Lemma 5.29 asserts that $\operatorname{dom}\operatorname{prox}_\varphi^f = \left]0,+\infty\right[^m$. Let $(\xi_i)_{1\leqslant i\leqslant m} \in \left]0,+\infty\right[^m$, set $(\eta_i)_{1\leqslant i\leqslant m} = \operatorname{prox}_\varphi^f(\xi_i)_{1\leqslant i\leqslant m}$, and let $i \in \{1,\ldots,m\}$. Then η_i can be computed as follows.

(i) Let $\gamma \in \left]0,+\infty\right[$ and suppose that $\phi = \gamma\vartheta$. Then $\eta_i = (1+\gamma)\xi_i$.

(ii) Let $(\gamma,\alpha) \in \left[0,+\infty\right[^2$, let $\omega \in \mathbb{R}$, and suppose that

$$\phi\colon \xi \mapsto \begin{cases} -\gamma\ln \xi + \omega\xi + \alpha\xi^{-1}, & \text{if} \quad \xi \in \left]0,+\infty\right[;\\ +\infty, & \text{otherwise.}\end{cases} \tag{5.113}$$

Then $\eta_i = (2+2\omega\xi_i)^{-1}\big((\gamma+1)\xi_i + \sqrt{(\gamma+1)^2\xi_i + 4\alpha\xi_i(1+\omega\xi_i)}\big)$.

(iii) Let $(\gamma,\alpha) \in \left[0,+\infty\right[^2$, let $p \in \left[1,+\infty\right[$, and suppose that

$$\phi\colon \xi \mapsto \begin{cases} -\gamma\ln \xi + \alpha\xi^p, & \text{if} \quad \xi \in \left]0,+\infty\right[;\\ +\infty, & \text{otherwise.}\end{cases} \tag{5.114}$$

Then η_i is the strictly positive solution of $p\alpha\xi_i\eta^p + \rho = (\gamma+1)\xi_i$.

(iv) Let $\alpha \in \left[0,+\infty\right[$, let $p \in \left[1,+\infty\right[$, and suppose that

$$\phi\colon \xi \mapsto \begin{cases} \alpha\xi^{-p}, & \text{if} \quad \xi \in \left]0,+\infty\right[;\\ +\infty, & \text{otherwise.}\end{cases} \tag{5.115}$$

Then η_i is the strictly positive solution of $p\eta^{p+1} - \xi_i\eta^p = \alpha p\xi_i$.

Example 5.34 Let $f\colon (\xi_i)_{1\leqslant i\leqslant m} \mapsto \sum_{i=1}^m \vartheta(\xi_i)$, where ϑ is Hellinger-like function, i.e.,

$$\vartheta\colon \xi \mapsto \begin{cases} -\sqrt{1-\xi^2}, & \text{if } \xi \in [-1,1];\\ +\infty, & \text{otherwise,}\end{cases} \tag{5.116}$$

let $\gamma \in \left]0,+\infty\right[$, and let $\varphi = f$. Note that f is a Legendre function [2, Theorem 5.12 and Example 6.5] and hence, Lemma 5.29 asserts that $\operatorname{dom}\operatorname{prox}_{\gamma\varphi}^f = \left]-1,1\right[^m$. Let $(\xi_i)_{1\leqslant i\leqslant m} \in \left]-1,1\right[^m$ and set $(\eta_i)_{1\leqslant i\leqslant m} = \operatorname{prox}_{\gamma\varphi}^f(\xi_i)_{1\leqslant i\leqslant m}$. Then $(\forall i \in \{1,\ldots,m\})$ $\eta_i = \xi_i/\sqrt{(\gamma+1)^2 + (\gamma^2+2\gamma+2)\xi_i^2}$.

5.2.4 Applications

5.2.4.1 Variable Bregman proximal point algorithm

The convex minimization problem, i.e., the problem of minimizing a convex function, can be solved by proximal point algorithm (see [5, 9] for Hilbertian setting and [4] for Banach space setting). In this section, we develop a proximal point algorithm which employs different Bregman distances at each iteration. This provides a unified framework for existing proximal point algorithms.

Theorem 5.35 *Let $\mathcal{X}$ be a reflexive real Banach space, let $\varphi \in \Gamma_0(\mathcal{X})$, let $f \in \Gamma_0(\mathcal{X})$ be a Legendre function such that $\operatorname{Argmin}\varphi \cap \operatorname{int}\operatorname{dom} f \neq \varnothing$, let $(\eta_n)_{n\in\mathbb{N}} \in \ell^1_+(\mathbb{N})$, let $\alpha \in]0,+\infty[$, and let $(f_n)_{n\in\mathbb{N}}$ be Legendre functions in $\mathcal{P}_\alpha(f)$ such that*

$$(\forall n \in \mathbb{N}) \quad (1+\eta_n)f_n \succcurlyeq f_{n+1}. \tag{5.117}$$

Let $x_0 \in \operatorname{int}\operatorname{dom} f$, let $(\gamma_n)_{n\in\mathbb{N}} \in]0,+\infty[^{\mathbb{N}}$ be such that $\gamma = \inf_{n\in\mathbb{N}} \gamma_n > 0$, and iterate

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = \operatorname{prox}^{f_n}_{\gamma_n\varphi} x_n. \tag{5.118}$$

Then the following hold :

(i) *$(x_n)_{n\in\mathbb{N}}$ is stationarily Bregman monotone with respect to $\operatorname{Argmin}\varphi$ relative to $(f_n)_{n\in\mathbb{N}}$.*

(ii) *$(x_n)_{n\in\mathbb{N}}$ is a minimizing sequence of φ.*

(iii) *Suppose that, for every $x \in \operatorname{int}\operatorname{dom} f$, $D^f(x,\cdot)$ is coercive, and that one of the following holds :*

   (a) *$\operatorname{Argmin}\varphi \cap \overline{\operatorname{dom}} f$ is a singleton.*

   (b) *Either $\operatorname{Argmin}\varphi \subset \operatorname{int}\operatorname{dom} f$ or $\operatorname{dom} f^*$ is open and ∇f^* is weakly sequentially continuous, there exists $g \in \mathcal{F}(f)$ such that, for every $n \in \mathbb{N}$, $g \succcurlyeq f_n$, and, for every $y_1 \in \mathcal{X}$ and every $y_2 \in \mathcal{X}$,*

$$\begin{cases} y_1 \in \mathfrak{W}(x_n)_{n\in\mathbb{N}} \\ y_2 \in \mathfrak{W}(x_n)_{n\in\mathbb{N}} \\ \big(\langle y_1 - y_2, \nabla f_n(x_n)\rangle\big)_{n\in\mathbb{N}} \quad \text{converges} \end{cases} \Rightarrow \quad y_1 = y_2. \tag{5.119}$$

   Then there exists $\overline{x} \in \operatorname{Argmin}\varphi$ such that $x_n \rightharpoonup \overline{x}$.

(iv) *Suppose that that f satisfies Condition 5.26 and that $(\forall x \in \operatorname{int}\operatorname{dom} f)$ $D^f(x,\cdot)$ is coercive. Furthermore, assume that $\underline{\lim}\, D^f_{\operatorname{Argmin}\varphi}(x_n) = 0$ and that there exists $\beta \in]0,+\infty[$ such that $(\forall n \in \mathbb{N})$ $\beta f \succcurlyeq f_n$. Then there exists $\overline{x} \in \operatorname{Argmin}\varphi$ such that $x_n \to \overline{x}$.*

Proof. First, for every $n \in \mathbb{N}$, since $\varnothing \neq \operatorname{Argmin} \varphi \cap \operatorname{int} \operatorname{dom} f \subset \operatorname{dom} \varphi \cap \operatorname{int} \operatorname{dom} f = \operatorname{dom} \varphi \cap \operatorname{int} \operatorname{dom} f_n$, Lemma 5.29 asserts that

$$\operatorname{prox}^{f_n}_{\gamma_n \varphi} \colon \operatorname{int} \operatorname{dom} f_n \to \operatorname{dom} \partial\varphi \cap \operatorname{int} \operatorname{dom} f_n \tag{5.120}$$

is well-defined and single-valued. Note that $x_0 \in \operatorname{int} \operatorname{dom} f$. Suppose that $x_n \in \operatorname{int} \operatorname{dom} f$ for some $n \in \mathbb{N}$. Then $x_n \in \operatorname{int} \operatorname{dom} f_n$, and hence, we deduce from (5.120) that $x_{n+1} \in \operatorname{dom} \partial\varphi \cap \operatorname{int} \operatorname{dom} f_n \subset \operatorname{int} \operatorname{dom} f$. By reasoning by induction, we conclude that

$$(x_n)_{n\in\mathbb{N}} \in \big(\operatorname{int} \operatorname{dom} f\big)^{\mathbb{N}} \quad \text{is well-defined.} \tag{5.121}$$

(i) : We first derive from (5.118) and Lemma 5.29(iii) that

$$(\forall n \in \mathbb{N}) \quad \nabla f_n(x_n) - \nabla f_n(x_{n+1}) \in \gamma_n \partial\varphi(x_{n+1}). \tag{5.122}$$

Next, by invoking (5.28) and (5.122), we get

$$(\forall x \in \operatorname{dom} \varphi \cap \operatorname{dom} f)(\forall n \in \mathbb{N}) \quad \gamma_n^{-1}\big\langle x - x_{n+1}, \nabla f_n(x_n) - \nabla f_n(x_{n+1})\big\rangle + \varphi(x_{n+1}) \leqslant \varphi(x). \tag{5.123}$$

It therefore follows from [3, Proposition 2.3(ii)] that

$$\begin{aligned}(\forall x \in \operatorname{dom} \varphi \cap \operatorname{dom} f)(\forall n \in \mathbb{N}) \quad \gamma_n^{-1}\big(D^{f_n}(x, x_{n+1}) + D^{f_n}(x_{n+1}, x_n) - D^{f_n}(x, x_n)\big)& \\ + \varphi(x_{n+1}) \leqslant \varphi(x),&\end{aligned} \tag{5.124}$$

and, in particular,

$$(\forall x \in \operatorname{Argmin} \varphi \cap \operatorname{dom} f)(\forall n \in \mathbb{N}) \quad D^{f_n}(x, x_{n+1}) \leqslant D^{f_n}(x, x_n) - D^{f_n}(x_{n+1}, x_n). \tag{5.125}$$

Since (5.117) yields

$$(\forall x \in \operatorname{Argmin} \varphi \cap \operatorname{dom} f)(\forall n \in \mathbb{N}) \quad D^{f_{n+1}}(x, x_{n+1}) \leqslant (1 + \eta_n) D^{f_n}(x, x_{n+1}), \tag{5.126}$$

it follows from (5.125) that

$$\begin{aligned}(\forall x \in \operatorname{Argmin} \varphi \cap \operatorname{dom} f)(\forall n \in \mathbb{N}) \quad D^{f_{n+1}}(x, x_{n+1}) &\leqslant (1 + \eta_n) D^{f_n}(x, x_n) \\ &\quad - (1 + \eta_n) D^{f_n}(x_{n+1}, x_n).\end{aligned} \tag{5.127}$$

In particular,

$$(\forall x \in \operatorname{Argmin} \varphi \cap \operatorname{dom} f)(\forall n \in \mathbb{N}) \quad D^{f_{n+1}}(x, x_{n+1}) \leqslant (1 + \eta_n) D^{f_n}(x, x_n). \tag{5.128}$$

This shows that $(x_n)_{n\in\mathbb{N}}$ is stationarily Bregman monotone with respect to $\operatorname{Argmin} \varphi$ relative to $(f_n)_{n\in\mathbb{N}}$.

(ii) : Let $x \in \operatorname{Argmin} \varphi \cap \operatorname{int} \operatorname{dom} f$. It follows from (i) and Proposition 5.21(i) that

$$\big(D^{f_n}(x, x_n)\big)_{n\in\mathbb{N}} \quad \text{converges} \tag{5.129}$$

and, since (5.127) yields

$$\begin{aligned}(\forall n \in \mathbb{N}) \quad D^{f_n}(x_{n+1}, x_n) &\leqslant (1+\eta_n) D^{f_n}(x_{n+1}, x_n) \\ &\leqslant (1+\eta_n) D^{f_n}(x, x_n) - D^{f_{n+1}}(x, x_{n+1}),\end{aligned} \tag{5.130}$$

we deduce that

$$D^{f_n}(x_{n+1}, x_n) \to 0. \tag{5.131}$$

On the other hand, since $(f_n)_{n\in\mathbb{N}}$ is in $\mathcal{P}_\alpha(f)$, we obtain

$$(\forall n \in \mathbb{N}) \quad \alpha D^f(x_{n+1}, x_n) \leqslant D^{f_n}(x_{n+1}, x_n). \tag{5.132}$$

Altogether, (5.131) and (5.132) yield

$$D^f(x_{n+1}, x_n) \to 0. \tag{5.133}$$

We also deduce from (5.124) that

$$(\forall n \in \mathbb{N}) \quad \varphi(x_{n+1}) \leqslant \gamma_n^{-1}\big(D^{f_n}(x_n, x_{n+1}) + D^{f_n}(x_{n+1}, x_n)\big) + \varphi(x_{n+1}) \leqslant \varphi(x_n). \tag{5.134}$$

This shows that $(\varphi(x_n))_{n\in\mathbb{N}}$ is decreasing, and hence, since it is bounded from below by $\inf \varphi(\mathcal{X})$, it converges. We now derive from (5.124) and (5.126) that

$$\begin{aligned}(\forall n \in \mathbb{N}) \quad &\frac{1}{\gamma}\left(\frac{1}{1+\eta_n} D^{f_{n+1}}(x, x_{n+1}) + D^{f_n}(x_{n+1}, x_n) - D^{f_n}(x, x_n)\right) + \varphi(x_{n+1}) \\ &\leqslant \frac{1}{\gamma_n}\left(\frac{1}{1+\eta_n} D^{f_{n+1}}(x, x_{n+1}) + D^{f_n}(x_{n+1}, x_n) - D^{f_n}(x, x_n)\right) + \varphi(x_{n+1}) \\ &\leqslant \varphi(x).\end{aligned} \tag{5.135}$$

Hence, by using (5.129) and (5.131) after letting $n \to +\infty$ in (5.135), we get

$$\inf \varphi(\mathcal{X}) \leqslant \lim \varphi(x_n) \leqslant \varphi(x) = \inf \varphi(\mathcal{X}). \tag{5.136}$$

In turn, $\varphi(x_n) \to \inf \varphi(\mathcal{X})$, i.e., $(x_n)_{n\in\mathbb{N}}$ is therefore a minimizing sequence of φ.

(iii) : We show actually that $\mathfrak{W}(x_n)_{n\in\mathbb{N}} \subset \operatorname{Argmin} \varphi$. To this end, suppose that $x \in \mathfrak{W}(x_n)_{n\in\mathbb{N}}$, i.e., $x_{k_n} \rightharpoonup x$. Since φ is lower semicontinuous and convex, it is weakly lower semicontinuous [19, Theorem 2.2.1], and hence,

$$\inf \varphi(\mathcal{X}) \leqslant \varphi(x) \leqslant \varliminf \varphi(x_{k_n}) = \inf \varphi(\mathcal{X}). \tag{5.137}$$

In turn, $\varphi(x) = \inf \varphi(\mathcal{X})$, i.e., $x \in \operatorname{Argmin} \varphi$.

(iii)(a) : Since $\mathcal{X}$ is reflexive, we derive from (i) and Proposition 5.21(ii) that $\mathfrak{W}(x_n)_{n\in\mathbb{N}} \neq \varnothing$. Let us fix $\overline{x} \in \mathfrak{W}(x_n)_{n\in\mathbb{N}}$. Since (5.121) yields $\mathfrak{W}(x_n)_{n\in\mathbb{N}} \subset \operatorname{Argmin} \varphi \cap \operatorname{dom} f$, we get $\mathfrak{W}(x_n)_{n\in\mathbb{N}} = \{\overline{x}\}$. In turn, $x_n \rightharpoonup \overline{x}$.

(iii)(b) : We shall show that $\mathfrak{W}(x_n)_{n\in\mathbb{N}} \subset \operatorname{int}\operatorname{dom} f$. To this end, let $\overline{x} \in \mathfrak{W}(x_n)_{n\in\mathbb{N}}$, i.e., $x_{k_n} \rightharpoonup \overline{x}$. If $\operatorname{Argmin} \varphi \subset \operatorname{int}\operatorname{dom} f$ then $\overline{x} \in \operatorname{Argmin} \varphi \subset \operatorname{int}\operatorname{dom} f$. Now suppose that $\operatorname{dom} f^*$ is open and ∇f^* is weakly sequentially continuous. Let $x \in \operatorname{Argmin} \varphi \cap \operatorname{int}\operatorname{dom} f$. Then $\nabla f(x) \in \operatorname{int}\operatorname{dom} f^*$ [3, Theorem 5.9] and it follows from [3, Lemma 7.3(v)] that $D^{f^*}(\cdot, \nabla f(x))$ is coercive. Since $(D^f(x, x_{k_n}))_{n\in\mathbb{N}}$ is bounded and since [3, Lemma 7.3(vii)] asserts that

$$(\forall n \in \mathbb{N}) \quad D^{f^*}(\nabla f(x_{k_n}), \nabla f(x)) = D^f(x, x_{k_n}), \tag{5.138}$$

we deduce that $(\nabla f(x_{k_n}))_{n\in\mathbb{N}}$ is bounded. Take $\overline{x}^* \in \mathcal{X}^*$ and a strictly increasing sequence $(p_{k_n})_{n\in\mathbb{N}}$ in $\mathbb{N}$ such that $\nabla f(x_{p_{k_n}}) \rightharpoonup \overline{x}^*$. Since [3, Lemma 7.3(ii)] states that $D^{f^*}(\cdot, \nabla f(x))$ is a proper lower semicontinuous convex function, we derive from (5.138) that

$$D^{f^*}(\overline{x}^*, \nabla f(x)) \leqslant \underline{\lim}\, D^{f^*}\big(\nabla f(x_{p_{k_n}}), \nabla f(x)\big) \leqslant \underline{\lim}\, D^f(x, x_{p_{k_n}}) < +\infty, \tag{5.139}$$

which shows that $\overline{x}^* \in \operatorname{dom} f^* = \operatorname{int}\operatorname{dom} f^*$ and thus, by [3, Theorem 5.10], there exists $\overline{x}_1 \in \operatorname{int}\operatorname{dom} f$ such that $\overline{x}^* = \nabla f(\overline{x}_1)$. Since ∇f^* is weakly sequentially continuous, we get

$$\overline{x} \leftharpoondown x_{p_{k_n}} = \nabla f^*\big(\nabla f(x_{p_{k_n}})\big) \rightharpoonup \nabla f^*(\overline{x}^*) = \overline{x}_1. \tag{5.140}$$

In turn, $\overline{x} = \overline{x}_1 \in \operatorname{int}\operatorname{dom} f$. Finally, the claim follows from Proposition 5.22.

(iv) : Since $\varphi \in \Gamma_0(\mathcal{X})$, $\operatorname{Argmin} \varphi$ is convex and closed, and the assertion therefore follows from Proposition 5.27. ☐

Remark 5.36 In Theorem 5.35, suppose that $(\forall n \in \mathbb{N})\ f_n = f$, $\gamma_n = \gamma$, and $\eta_n = 0$. Then (5.118) reduces to the Bregman proximal iterations [4]

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = \operatorname{prox}^f_{\gamma\varphi} x_n. \tag{5.141}$$

5.2.4.2 An application to the convex feasibility problem

In this section, we apply the asymptotic analysis of variable Bregman monotone sequences to study the convex feasibility problem, i.e., the generic problem of finding a point in the intersection of a family of closed convex sets. We first recall the following results.

Lemma 5.37 [4, Definition 3.1 and Proposition 3.3] *Let $\mathcal{X}$ be a reflexive real Banach space, let $f \in \Gamma_0(\mathcal{X})$ be Gâteaux differentiable on* $\operatorname{int}\operatorname{dom} f \neq \varnothing$*, set*

$$(\forall (x,y) \in (\operatorname{int}\operatorname{dom} f)^2) \quad H^f(x,y) = \left\{z \in \mathcal{X} \;\middle|\; \left\langle z-y, \nabla f(x) - \nabla f(y)\right\rangle \leqslant 0\right\}$$
$$= \left\{z \in \mathcal{X} \;\middle|\; D^f(z,y) + D^f(y,x) \leqslant D^f(z,x)\right\} \tag{5.142}$$

and

$$\mathfrak{B}(f) = \Big\{T\colon \mathcal{X} \to 2^{\mathcal{X}} \;\Big|\; \operatorname{ran} T \subset \operatorname{dom} T = \operatorname{int}\operatorname{dom} f$$
$$\text{and } (\forall (x,y) \in \operatorname{gra} T)\ \operatorname{Fix} T \subset H^f(x,y)\Big\}. \tag{5.143}$$

Let $T \in \mathfrak{B}(f)$ be such that $\operatorname{Fix} T \neq \varnothing$*. Suppose that $f|_{\operatorname{int}\operatorname{dom} f}$ is strictly convex. Then the following hold :*

(i) $\operatorname{Fix} T$ *is convex.*

(ii) $(\forall x \in \overline{\operatorname{Fix}}\, T)(\forall (y,v) \in \operatorname{gra} T)\ D^f(x,v) + D^f(v,y) \leqslant D^f(x,y)$.

The class of operators $\mathfrak{B}$ includes types of fundamental operators in Bregman optimization (see [4] for more discussions). We illustrate our result in Section 5.2.2 through an application to the problem of finding a common point of a family of closed convex subsets with nonempty intersection.

Theorem 5.38 *Let $\mathcal{X}$ be a reflexive real Banach space, let I be a totally ordered at most countable index set, let $(C_i)_{i\in I}$ be a family of closed convex subsets of $\mathcal{X}$ such that $C = \bigcap_{i\in I} C_i \neq \varnothing$, let $f \in \Gamma_0(\mathcal{X})$ be Gâteaux differentiable on* $\operatorname{int}\operatorname{dom} f \neq \varnothing$*, let $(\eta_n)_{n\in\mathbb{N}} \in \ell^1_+(\mathbb{N})$, let $\alpha \in \left]0,+\infty\right[$, and let $(f_n)_{n\in\mathbb{N}}$ be Legendre functions in $\mathcal{P}_\alpha(f)$ such that*

$$(\forall n \in \mathbb{N}) \quad (1+\eta_n) f_n \succcurlyeq f_{n+1}. \tag{5.144}$$

Let $\mathrm{i}\colon \mathbb{N} \to I$ be such that

$$(\forall j \in I)(\exists M_j \in \mathbb{N}\backslash\{0\})(\forall n \in \mathbb{N}) \quad j \in \{\mathrm{i}(n), \ldots, \mathrm{i}(n+M_j-1)\}. \tag{5.145}$$

For every $i \in I$, let $(T_{i,n})_{n\in\mathbb{N}}$ be a sequence of operators such that

$$(\forall n \in \mathbb{N}) \quad T_{i,n} \in \mathfrak{B}(f_n), \quad C_i \cap \operatorname{Fix} T_{i,n} \neq \varnothing, \quad \text{and} \quad C_i \subset \overline{\operatorname{Fix}}\, T_{i,n}. \tag{5.146}$$

Let $x_0 \in$ $\operatorname{int}\operatorname{dom} f$ *and iterate*

$$(\forall n \in \mathbb{N}) \quad x_{n+1} \in T_{\mathrm{i}(n),n} x_n. \tag{5.147}$$

Suppose that f satisfies Condition 5.26 and that $(\forall x \in \operatorname{int}\operatorname{dom} f)$ $D^f(x,\cdot)$ *is coercive. Then there exists $\overline{x} \in C$ such that the following hold :*

(i) *Suppose that there exists* $g \in \mathcal{F}(f)$ *that, for every* $n \in \mathbb{N}$, $g \succcurlyeq f_n$, *and that, for every* $y_1 \in \mathcal{X}$ *and every* $y_2 \in \mathcal{X}$,

$$\begin{cases} y_1 \in \mathfrak{W}(x_n)_{n\in\mathbb{N}} \cap C \\ y_2 \in \mathfrak{W}(x_n)_{n\in\mathbb{N}} \cap C \\ \big(\langle y_1 - y_2, \nabla f_n(x_n)\rangle\big)_{n\in\mathbb{N}} \quad \textit{converges} \end{cases} \quad \Rightarrow \quad y_1 = y_2, \tag{5.148}$$

and that, for every strictly increasing sequence $(l_n)_{n\in\mathbb{N}}$ *in* $\mathbb{N}$, *every* $x \in \mathcal{X}$, *and every* $j \in I$,

$$\begin{cases} x_{l_n} \rightharpoonup x \\ y_{l_n} \in T_{j,l_n} x_{l_n} \\ y_{l_n} - x_{l_n} \to 0 \\ (\forall n \in \mathbb{N})\ j = \mathrm{i}(l_n) \end{cases} \quad \Rightarrow \quad x \in C_j. \tag{5.149}$$

In addition, assume that $\mathfrak{W}(x_n)_{n\in\mathbb{N}} \subset \operatorname{int\,dom} f$. *Then* $x_n \rightharpoonup \overline{x}$.

(ii) *Suppose that* f *is Legendre, that* $\underline{\lim}\, D_C^f(x_n) = 0$, *and that there exists* $\beta \in \left]0, +\infty\right[$ *such that* $(\forall n \in \mathbb{N})\ \beta f \succcurlyeq f_n$. *Then* $x_n \to \overline{x}$.

Proof. For every $n \in \mathbb{N}$ and every $i \in I$, we observe that $\operatorname{ran} T_{i,n} \subset \operatorname{dom} T_{i,n} = \operatorname{int\,dom} f_n = \operatorname{int\,dom} f$. Hence, it follows from (5.146) and (5.147) that $(x_n)_{n\in\mathbb{N}}$ is a well-define sequence in $\operatorname{int\,dom} f$. We now derive from (5.142), (5.146), and (5.147) that

$$(\forall x \in C \cap \operatorname{dom} f)(\forall n \in \mathbb{N}) \quad D^{f_n}(x, x_{n+1}) + D^{f_n}(x_{n+1}, x_n) \leqslant D^{f_n}(x, x_n). \tag{5.150}$$

Since (5.144) yields

$$(\forall x \in C \cap \operatorname{dom} f)(\forall n \in \mathbb{N}) \quad D^{f_{n+1}}(x, x_{n+1}) \leqslant (1 + \eta_n) D^{f_n}(x, x_{n+1}), \tag{5.151}$$

we deduce that

$$(\forall x \in C \cap \operatorname{dom} f)(\forall n \in \mathbb{N}) \quad D^{f_{n+1}}(x, x_{n+1}) \leqslant (1+\eta_n) D^{f_n}(x, x_n) - (1+\eta_n) D^{f_n}(x_{n+1}, x_n). \tag{5.152}$$

In particular,

$$(\forall x \in C \cap \operatorname{dom} f)(\forall n \in \mathbb{N}) \quad D^{f_{n+1}}(x, x_{n+1}) \leqslant (1 + \eta_n) D^{f_n}(x, x_n), \tag{5.153}$$

which shows that $(x_n)_{n\in\mathbb{N}}$ is stationarily Bregman monotone with respect to C relative to $(f_n)_{n\in\mathbb{N}}$. In addition, we derive from (5.146) that $(\forall i \in \{1, \ldots, m\})\ C_i \cap \operatorname{int\,dom} f \neq \varnothing$. Hence, $C \cap \operatorname{int\,dom} f \neq \varnothing$.

(i) : In view of Proposition 5.22, it suffices to show that $\mathfrak{W}(x_n)_{n\in\mathbb{N}} \subset C \cap \operatorname{int\,dom} f$. To this end, let $\overline{x} \in \mathfrak{W}(x_n)_{n\in\mathbb{N}}$, let $(k_n)_{n\in\mathbb{N}}$ be a strictly increasing sequence in $\mathbb{N}$ such that $x_{k_n} \rightharpoonup \overline{x}$, let $j \in I$, and let $x \in C \cap \operatorname{int\,dom} f$. By (5.145), there exists a strictly increasing sequence $(l_n)_{n\in\mathbb{N}}$ in $\mathbb{N}$ such that

$$(\forall n \in \mathbb{N}) \quad \begin{cases} k_n \leqslant l_n \leqslant k_n + M_j - 1 < k_{n+1} \leqslant l_{n+1}, \\ j = \mathrm{i}(l_n). \end{cases} \tag{5.154}$$

Since $D^f(x,\cdot)$ is coercive, it follows from Proposition 5.21 that $(x_n)_{\in\mathbb{N}}$ is bounded and $(D^{f_n}(x_{n+1},x_n))_{n\in\mathbb{N}}$ converges. In turn, since (5.152) yields

$$\begin{aligned}(\forall n \in \mathbb{N}) \quad D^{f_n}(x_{n+1},x_n) &\leqslant (1+\eta_n)D^{f_n}(x_{n+1},x_n) \\ &\leqslant (1+\eta_n)D^{f_n}(x,x_n) - D^{f_{n+1}}(x,x_{n+1}),\end{aligned} \tag{5.155}$$

we deduce that

$$D^{f_n}(x_{n+1},x_n) \to 0. \tag{5.156}$$

However, since

$$(\forall n \in \mathbb{N}) \quad \alpha D^f(x_{n+1},x_n) \leqslant D^{f_n}(x_{n+1},x_n), \tag{5.157}$$

it follows from (5.156) that

$$D^f(x_{n+1},x_n) \to 0 \tag{5.158}$$

and hence, since f satisfies Condition 5.26,

$$x_{n+1} - x_n \to 0. \tag{5.159}$$

Altogether, (5.154) and (5.159) imply that

$$\|x_{l_n} - x_{k_n}\| \leqslant \sum_{m=k_n}^{k_n+M_j-2} \|x_{m+1} - x_n\| \leqslant (M_j - 1) \max_{k_n \leqslant m \leqslant k_n+M_j-2} \|x_{m+1} - x_m\| \to 0, \tag{5.160}$$

and therefore

$$x_{l_n} \rightharpoonup \overline{x}. \tag{5.161}$$

Now let $(\forall n \in \mathbb{N})$ $y_{l_n} \in T_{j,l_n} x_{l_n}$. We deduce from (5.154) and (5.159) that

$$y_{l_n} - x_{l_n} \to 0. \tag{5.162}$$

By invoking successively (5.149), (5.161), and (5.162), we get $\overline{x} \in C_j$, and hence, $\overline{x} \in C$. Consequently, $\mathfrak{W}(x_n)_{n\in\mathbb{N}} \subset C \cap \operatorname{int\,dom} f$.

(ii) : Since C is closed, the assertion follows from Proposition 5.27. □

Remark 5.39

(i) In Theorem 5.38, suppose that $(\forall n \in \mathbb{N})\ f_n = f$ and $\eta_n = 0$. Then we recover the framework of [4, Section 4.2].

(ii) In Theorem 5.38, suppose that $\mathcal{X}$ is a Hilbert space, that $f = \|\cdot\|^2/2$, and that $(\forall n \in \mathbb{N})\ f_n \colon x \mapsto \langle x, U_n x\rangle/2$, where $(U_n)_{n\in\mathbb{N}}$ are operators in $\mathcal{P}_\alpha(\mathcal{X})$ such that $\sup_{n\in\mathbb{N}} \|U_n\| < +\infty$ and $(\forall n \in \mathbb{N})\ (1+\eta_n)U_n \succcurlyeq U_{n+1}$. Then we recover the version of [9, Theorem 5.1(i) and (iii)] without errors and $(\forall n \in \mathbb{N})\ \lambda_n = 1$.

Our last result concerns a periodic projection method that uses different Bregman distances at each iteration.

Corollary 5.40 *Let $\mathcal{X}$ be a reflexive real Banach space, let m be a strictly positive integer, let $(C_i)_{1\leqslant i\leqslant m}$ be a family of closed convex subsets of $\mathcal{X}$ such that $C = \bigcap_{i=1}^m C_i \neq \varnothing$, let $f \in \Gamma_0(\mathcal{X})$ be Gâteaux differentiable on $\operatorname{int}\operatorname{dom} f$ such that $C \cap \operatorname{int}\operatorname{dom} f \neq \varnothing$, let $(\eta_n)_{n\in\mathbb{N}} \in \ell_+^1(\mathbb{N})$, let $\alpha \in \left]0,+\infty\right[$, and let $(f_n)_{n\in\mathbb{N}}$ be Legendre functions in $\mathcal{P}_\alpha(f)$ such that*

$$(\forall n \in \mathbb{N}) \quad (1+\eta_n) f_n \succcurlyeq f_{n+1}. \tag{5.163}$$

Let $x_0 \in \operatorname{int}\operatorname{dom} f$ and iterate

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = P^{f_n}_{C_{1+\operatorname{rem}(n,m)}} x_n, \tag{5.164}$$

where $\operatorname{rem}(\cdot, m)$ is the remainder of the division by m. Suppose that f satisfies Condition 5.26 and that $(\forall x \in \operatorname{int}\operatorname{dom} f)\ D^f(x,\cdot)$ is coercive. Then there exists $\overline{x} \in C$ such that the following hold :

(i) *Suppose that there exists $g \in \mathcal{F}(f)$ such that, for every $n \in \mathbb{N}$, $g \succcurlyeq f_n$, and that, for every $y_1 \in \mathcal{X}$ and every $y_2 \in \mathcal{X}$,*

$$\begin{cases} y_1 \in \mathfrak{W}(x_n)_{n\in\mathbb{N}} \cap C \\ y_2 \in \mathfrak{W}(x_n)_{n\in\mathbb{N}} \cap C \\ \big(\langle y_1 - y_2, \nabla f_n(x_n)\rangle\big)_{n\in\mathbb{N}} \quad \text{converges} \end{cases} \quad \Rightarrow \quad y_1 = y_2. \tag{5.165}$$

In addition, suppose that $\mathfrak{W}(x_n)_{n\in\mathbb{N}} \subset \operatorname{int}\operatorname{dom} f$. Then $x_n \rightharpoonup \overline{x}$.

(ii) *Suppose that f is Legendre, that $\underline{\lim}\, D^f_C(x_n) = 0$, and that there exists $\beta \in \left]0,+\infty\right[$ such that $(\forall n \in \mathbb{N})\ \beta f \succcurlyeq f_n$. Then $x_n \to \overline{x}$.*

Proof. First, we see that the function $\mathrm{i}\colon \mathbb{N} \to \{1,\ldots,m\}\colon n \mapsto 1 + \operatorname{rem}(n,m)$ satisfies (5.145), where $(\forall j \in \{1,\ldots,m\})\ M_j = m$. Now set

$$(\forall i \in \{1,\ldots,m\})(\forall n \in \mathbb{N}) \quad T_{i,n} = P^{f_n}_{C_i}. \tag{5.166}$$

Then, by [4, Theorem 3.34], for every $n \in \mathbb{N}$ and every $i \in \{1,\ldots,m\}$, we have

$$T_{i,n} \in \mathfrak{B}(f_n) \quad \text{and} \quad C_i \cap \overline{\operatorname{dom}} f \cap \operatorname{Fix} T_{i,n} = C_i \cap \operatorname{int}\operatorname{dom} f \supset C \cap \operatorname{int}\operatorname{dom} f \neq \varnothing. \tag{5.167}$$

In addition, it follows from [4, Lemma 3.2] that

$$(\forall n \in \mathbb{N})(\forall i \in \{1,\ldots,m\}) \quad C_i \cap \overline{\operatorname{dom} f} = \overline{C_i \cap \operatorname{int}\operatorname{dom} f} = \overline{C_i \cap \operatorname{int}\operatorname{dom} f_n} = \overline{\operatorname{Fix}} T_{i,n}. \tag{5.168}$$

Therefore, (5.164) is a particular case of (5.147). We shall actually apply Proposition 5.38 with the family $(C_i \cap \overline{\operatorname{dom} f})_{1\leqslant i\leqslant m}$.

(i) : Let us fix $j \in \{1,\ldots,m\}$ and suppose that

$$x_{l_n} \rightharpoonup x, \quad T_{j,l_n}x_{l_n} - x_{l_n} \to 0, \quad \text{and} \quad (\forall n \in \mathbb{N}) \quad j = \mathrm{i}(l_n). \tag{5.169}$$

Then $C_j \ni P^{f_{l_n}}_{C_j} x_{l_n} = T_{j,l_n}x_{l_n} \rightharpoonup x$, and hence, $x \in C_j$ since C_j is weakly closed [18, Corollary 4.5]. Moreover, since $(x_n)_{n\in\mathbb{N}}$ is in $\operatorname{int}\operatorname{dom} f$, $x \in \overline{\operatorname{dom} f}$ and hence $x \in C_j \cap \overline{\operatorname{dom} f}$. This shows that (5.149) is satisfied. Consequently, the assertion follows from Proposition 5.38(i).

(ii) : We have

$$(\forall n \in \mathbb{N}) \qquad \inf_{x\in C\cap\overline{\operatorname{dom} f}} D^f(x,x_n) \leqslant \inf_{x\in C\cap\operatorname{dom} f} D^f(x,x_n) = D^f_C(x_n), \tag{5.170}$$

and hence, $\underline{\lim}\, D_{C\cap\overline{\operatorname{dom} f}}(x_n) = 0$. The claim therefore follows from Proposition 5.38(ii). ◻

Acknowledgment. I would like to thank my doctoral advisor Professor Patrick L. Combettes for bringing this problem to my attention and for helpful discussions.

5.3 Bibliographie

[1] H. H. Bauschke and J. M. Borwein, On projection algorithms for solving convex feasibility problems, *SIAM Rev.*, vol. 38, pp. 367–426, 1996.

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

[3] H. H. Bauschke, J. M. Borwein, and P. L. Combettes, Essential smoothness, essential strict convexity, and Legendre functions in Banach spaces, *Commun. Contemp. Math.*, vol. 3, pp. 615–647, 2001.

[4] H. H. Bauschke, J. M. Borwein, and P. L. Combettes, Bregman monotone optimization algorithms, *SIAM J. Control Optim.*, vol. 42, pp. 596–636, 2003.

[5] H. H. Bauschke and P. L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. Springer, New York, 2011.

[6] L. M. Bregman, The relaxation method of finding the common point of convex sets and its application to the solution of problems in convex programming, *USSR Comput. Math. Math. Phys.*, vol 7, pp. 200–217, 1967.

[7] P. L. Combettes, Quasi-Fejérian analysis of some optimization algorithms, in : *Inherently Parallel Algorithms for Feasibility and Optimization*, (D. Butnariu, Y. Censor, and S. Reich, eds.), pp. 115–152. Elsevier, New York, 2001.

[8] P. L. Combettes, Fejér monotonicity in convex optimization, in : *Encyclopedia of Optimization,* 2nd ed. (C. A. Floudas and P. M. Pardalos, eds.), pp. 1016–1024. Springer, New York, 2009.

[9] P. L. Combettes and B. C. Vũ, Variable metric quasi-Fejér monotonicity, *Nonlinear Anal.,* vol. 78, pp. 17–31, 2013.

[10] R. M. Corless, G. H. Gonnet, D. E. G. Hare, D. J. Jeffrey, and D. E. Knuth, On the Lambert W function, *Adv. Comput. Math.,* vol. 5, pp. 329–359, 1996.

[11] Yu. M. Ermol'ev, On convergence of random quasi-Fejér sequences, *Cybernetics,* vol. 7, pp. 655–656, 1971.

[12] K. Knopp, *Infinite Sequences and Series.* Dover, Inc., New York, 1956.

[13] J. J. Moreau, Fonctions convexes duales et points proximaux dans un espace hilbertien, *C. R. Acad. Sci. Paris Sér. A Math.,* vol. 255, pp. 2897–2899, 1962.

[14] J. J. Moreau, Proximité et dualité dans un espace hilbertien, *Bull. Soc. Math. France,* vol. 93, pp. 273–299, 1965.

[15] R. R. Phelps, *Convex Functions, Monotone Operators and Differentiability,* 2nd ed., Lecture Notes in Math. 1364, Springer-Verlag, Berlin, 1993.

[16] B. T. Polyak, *Introduction to Optimization.* Optimization Software Inc., New York, 1987.

[17] E. Raik, Fejér type methods in Hilbert space, *Eesti NSV Tead. Akad. Toimetised Füüs.-Mat.,* vol. 16, pp. 286–293, 1967.

[18] S. Simons, *From Hahn-Banach to Monotonicity,* 2nd ed., Lecture Notes in Math. 1693, Springer-Verlag, New York, 2008.

[19] C. Zălinescu, *Convex Analysis in General Vector Spaces.* World Scientific Publishing, River Edge, NJ, 2002.

Chapitre 6

Algorithme explicite-implicite avec des distances de Bregman

Nous proposons un algorithme explicite-implicite basé sur des distances de Bregman pour les problèmes de minimisation convexe composite dans des espaces de Banach réel réflexifs. La convergence est établie en utilisant la notion de suites quasi-monotones au sens de Bregman étudiée dans le Chapitre 5. De nombreux exemples sont présentés, y compris dans des espaces euclidiens, où des algorithmes nouveaux sont obtenus. Des applications numériques en traitement d'images sont proposées.

6.1 Description et résultats principaux

Nous considérons dans ce chapitre le problème suivant.

Problème 6.1 Soient $\mathcal{X}$ et $\mathcal{Y}$ des espaces de Banach réels réflexifs, soit $\varphi \in \Gamma_0(\mathcal{X})$, soit $\psi \in \Gamma_0(\mathcal{Y})$ une fonction Gâteaux différentiable sur $\operatorname{int}\operatorname{dom}\psi \neq \varnothing$, et soit $L \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$. Le problème est de

$$\underset{x \in \mathcal{X}}{\text{minimiser}} \;\; \varphi(x) + \psi(Lx). \tag{6.1}$$

L'ensemble des solutions de (6.1) est noté $\mathcal{S}$.

Nos résultats sont basés sur les définitions suivantes.

Définition 6.2 [6] Soit $\mathcal{X}$ un espace de Banach réel réflexif et soit $f \in \Gamma_0(\mathcal{X})$ Gâteaux différentiable sur $\operatorname{int}\operatorname{dom} f \neq \varnothing$. La distance de Bregman associée à f est

$$\begin{aligned} D^f \colon \mathcal{X} \times \mathcal{X} &\to [0, +\infty] \\ (x, y) &\mapsto \begin{cases} f(x) - f(y) - \big\langle x - y, \nabla f(y)\big\rangle, & \text{si } \; y \in \operatorname{int}\operatorname{dom} f; \\ +\infty, & \text{sinon}. \end{cases} \end{aligned} \tag{6.2}$$

Soit $\varphi \in \Gamma_0(\mathcal{X})$. Alors l'opérateur *$f$-proximal* est

$$\begin{aligned}\operatorname{Prox}_\varphi^f \colon \mathcal{X}^* &\to 2^{\mathcal{X}} \\ x^* &\mapsto \big\{x \in \mathcal{X} \;\big|\; \varphi(x) + f(x) - \langle x, x^*\rangle = \min\big(\varphi + f - x^*\big)(\mathcal{X}) < +\infty\big\}.\end{aligned} \tag{6.3}$$

De plus, f est *de Legendre* si elle est *essentiellement lisse* au sens où ∂f est à la fois localement borné et univoque sur son domaine, et *essentiellement strictement convexe* au sens où ∂f^* est localement borné sur son domaine et f est strictement convexe sur les sous-ensembles convexes de $\operatorname{dom} \partial f$. Soit C un sous-ensemble convexe fermé de $\mathcal{X}$ tel que $C \cap \operatorname{int} \operatorname{dom} f \neq \varnothing$. Le *$D^f$-projecteur* sur C est

$$\begin{aligned}P_C^f \colon \operatorname{int}\operatorname{dom} f &\to C \cap \operatorname{int}\operatorname{dom} f \\ y &\mapsto \underset{x\in C}{\operatorname{argmin}}\, D^f(x,y),\end{aligned} \tag{6.4}$$

et la *D^f-distance* à C est

$$\begin{aligned}D_C^f \colon \mathcal{X} &\to [0,+\infty] \\ y &\mapsto \inf D^f(C,y).\end{aligned} \tag{6.5}$$

Nous rappelons la définition et la condition suivantes utilisées dans le Chapitre 2.

Définition 6.3 Soit $\mathcal{X}$ un espace de Banach réel réflexif et soit $f \in \Gamma_0(\mathcal{X})$ Gâteaux différentiable sur $\operatorname{int}\operatorname{dom} f \neq \varnothing$. Alors

$$\mathcal{F}(f) = \big\{g \in \Gamma_0(\mathcal{X}) \;\big|\; g \text{ est Gâteaux différentiable sur } \operatorname{dom} g = \operatorname{int}\operatorname{dom} f\big\}. \tag{6.6}$$

De plus, si g_1 et g_2 appartiennent $\mathcal{F}(f)$, alors

$$g_1 \succcurlyeq g_2 \quad \Leftrightarrow \quad (\forall x \in \operatorname{dom} f)(\forall y \in \operatorname{int}\operatorname{dom} f) \quad D^{g_1}(x,y) \geqslant D^{g_2}(x,y). \tag{6.7}$$

Pour tout $\alpha \in [0,+\infty[$, posons

$$\mathcal{P}_\alpha(f) = \big\{g \in \mathcal{F}(f) \;\big|\; g \succcurlyeq \alpha f\big\}. \tag{6.8}$$

Notation 6.4 Soit $\mathcal{X}$ un espace de Banach réel réflexif et soit $f \in \Gamma_0(\mathcal{X})$ Gâteaux différentiable sur $\operatorname{int}\operatorname{dom} f \neq \varnothing$. Alors

$$\begin{aligned}\hat{f} \colon \mathcal{X} &\to \left]-\infty,+\infty\right] \\ x &\mapsto \begin{cases} f(x), & \text{si } x \in \operatorname{int}\operatorname{dom} f; \\ +\infty, & \text{sinon.}\end{cases}\end{aligned} \tag{6.9}$$

Condition 6.5 [4, Condition 4.4] Soit $\mathcal{X}$ un espace de Banach réel réflexif et soit $f \in \Gamma_0(\mathcal{X})$ Gâteaux différentiable sur $\operatorname{int}\operatorname{dom} f \neq \varnothing$. Pour toutes suites bornées $(x_n)_{n\in\mathbb{N}}$ et $(y_n)_{n\in\mathbb{N}}$ dans $\operatorname{int}\operatorname{dom} f$,

$$D^f(x_n,y_n) \to 0 \quad \Rightarrow \quad x_n - y_n \to 0. \tag{6.10}$$

Nous introduisons un algorithme explicite-implicite basé sur les distances de Bregman variables pour résoudre le Problème 6.1.

Théorème 6.6 *Considérons le Problème 6.1. Soit $f \in \Gamma_0(\mathcal{X})$ une fonction de Legendre telle que $\mathcal{S} \cap \operatorname{int}\operatorname{dom} f \neq \varnothing$, $L(\operatorname{int}\operatorname{dom} f) \subset \operatorname{int}\operatorname{dom} \psi$, et $f \succcurlyeq \beta\psi \circ L$ pour un certain $\beta \in \left]0,+\infty\right[$. Soient $(\eta_n)_{n\in\mathbb{N}} \in \ell_+^1(\mathbb{N})$, $\alpha \in \left]0,+\infty\right[$, et $(f_n)_{n\in\mathbb{N}}$ des fonctions de Legendre dans $\mathcal{P}_\alpha(f)$ telles que*

$$(\forall n \in \mathbb{N}) \quad (1+\eta_n)f_n \succcurlyeq f_{n+1}. \tag{6.11}$$

Supposons que ou bien $-L^*(\operatorname{ran}\nabla\psi) \subset \operatorname{dom}\varphi^*$ ou bien $(\forall n \in \mathbb{N})$ f_n soit cofinie. Soient $\varepsilon \in \left]0, \alpha\beta/(\alpha\beta+1)\right[$ et $(\gamma_n)_{n\in\mathbb{N}}$ une suite dans $\mathbb{R}$ telle que

$$(\forall n \in \mathbb{N}) \quad \varepsilon \leqslant \gamma_n \leqslant \alpha\beta(1-\varepsilon) \quad \text{et} \quad (1+\eta_n)\gamma_n - \gamma_{n+1} \leqslant \alpha\beta\eta_n. \tag{6.12}$$

Soit $x_0 \in \operatorname{int}\operatorname{dom} f$ et posons

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = \operatorname{Prox}_{\gamma_n\varphi}^{f_n}\big(\nabla f_n(x_n) - \gamma_n L^*\nabla\psi(Lx_n)\big). \tag{6.13}$$

De plus, supposons que $(\forall x \in \operatorname{int}\operatorname{dom} f)$ $D^f(x,\cdot)$ soit coercive. Alors $(x_n)_{n\in\mathbb{N}}$ est une suite bornée dans $\operatorname{int}\operatorname{dom} f$ et $\mathfrak{W}(x_n)_{n\in\mathbb{N}} \subset \mathcal{S}$. De plus, il existe $\overline{x} \in \mathcal{S}$ tel que les propriétés suivantes sont satisfaites :

(i) *Supposons que $\mathcal{S} \cap \overline{\operatorname{dom}} f$ soit un singleton. Alors $x_n \rightharpoonup \overline{x}$.*

(ii) *Supposons qu'il existe $g \in \mathcal{F}(f)$ telle que, pour tout $n \in \mathbb{N}$, $g \succcurlyeq f_n$, et que, pour tout $y_1 \in \mathcal{X}$ et tout $y_2 \in \mathcal{X}$,*

$$\begin{cases} y_1 \in \mathfrak{W}(x_n)_{n\in\mathbb{N}} \\ y_2 \in \mathfrak{W}(x_n)_{n\in\mathbb{N}} \\ \big(\langle y_1 - y_2, \nabla f_n(x_n) - \gamma_n L^*\nabla\psi(Lx_n)\rangle\big)_{n\in\mathbb{N}} \quad \text{converge} \end{cases} \Rightarrow \quad y_1 = y_2. \tag{6.14}$$

De plus, supposons qu'une des conditions suivantes soit satisfaite.

(a) $\mathcal{S} \subset \operatorname{int}\operatorname{dom} f$.

(b) $\operatorname{dom} f^*$ *est ouvert et* ∇f^* *est faiblement séquentiellement continu.*

Alors $x_n \rightharpoonup \overline{x}$.

(iii) *Supposons que f satisfasse la Condition 6.5 et qu'une des conditions suivantes soit vérifiée.*

(a) *φ est uniformément convexe en $\overline{x}$.*

(b) *ψ est uniformément convexe en $L\overline{x}$ et il existe $\kappa \in \left]0,+\infty\right[$ tel que $(\forall x \in \mathcal{X})$ $\|Lx\| \geqslant \kappa\|x\|$.*

(c) $\varliminf D_{\mathcal{S}}^f(x_n) = 0$ *et il existe $\mu \in \left]0,+\infty\right[$ tel que $(\forall n \in \mathbb{N})$ $\mu f \succcurlyeq f_n$.*

Alors $x_n \to \overline{x}$.

Le théorème suivant constitue la première version de l'algorithme explicite-implicite dans les espaces de Banach réels réflexifs.

Théorème 6.7 *Considérons le Problème 6.1. Soit f une fonction de Legendre telle que $\mathcal{S} \cap \operatorname{int}\operatorname{dom} f \neq \varnothing$, $L(\operatorname{int}\operatorname{dom} f) \subset \operatorname{int}\operatorname{dom} \psi$, et $f \succcurlyeq \beta\psi \circ L$ pour un certain $\beta \in \left]0, +\infty\right[$. Supposons que ou bien f soit cofinie ou bien $-L^*(\operatorname{ran}\nabla\psi) \subset \operatorname{dom}\varphi^*$, et que $(\forall x \in \operatorname{int}\operatorname{dom} f)$ $D^f(x, \cdot)$ soit coercive. Soient $\varepsilon \in \left]0, \beta/(\beta+1)\right[$, $(\eta_n)_{n\in\mathbb{N}} \in \ell_+^1(\mathbb{N})$, et $(\gamma_n)_{n\in\mathbb{N}}$ une suite dans $\mathbb{R}$ telle que*

$$(\forall n \in \mathbb{N}) \quad \varepsilon \leqslant \gamma_n \leqslant \beta(1-\varepsilon) \quad \text{et} \quad (1+\eta_n)\gamma_n - \gamma_{n+1} \leqslant \beta\eta_n. \tag{6.15}$$

De plus, soit $x_0 \in \operatorname{int}\operatorname{dom} f$ et posons

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = \operatorname{Prox}^f_{\gamma_n\varphi}\big(\nabla f(x_n) - \gamma_n L^*\nabla\psi(Lx_n)\big). \tag{6.16}$$

Alors il existe $\overline{x} \in \mathcal{S}$ tel que les propriétés suivantes sont satisfaites :

(i) *Supposons qu'une des conditions suivantes soit satisfaite.*
   (a) *$\mathcal{S} \cap \overline{\operatorname{dom}}\, f$ est un singleton.*
   (b) *∇f et $\nabla\psi$ sont faiblement séquentiellement continus et $\mathcal{S} \subset \operatorname{int}\operatorname{dom} f$.*
   (c) *$\operatorname{dom} f^*$ est ouvert et ∇f, ∇f^*, et $\nabla\psi$ sont faiblement séquentiellement continus.*

   Alors $x_n \rightharpoonup \overline{x}$.

(ii) *Supposons que f satisfasse la Condition 6.5 et qu'une des conditions suivantes soit satisfaite.*
   (a) *φ est uniformément convexe en $\overline{x}$.*
   (b) *ψ est uniformément convexe en $L\overline{x}$ et il existe $\kappa \in \left]0, +\infty\right[$ tel que $(\forall x \in \mathcal{X})$ $\|Lx\| \geqslant \kappa\|x\|$.*
   (c) *$\underline{\lim}\, D^f_{\mathcal{S}}(x_n) = 0$.*

   Alors $x_n \to \overline{x}$.

Nous obtenons le corollaire suivant qui concerne les espaces de dimension finie.

Corollaire 6.8 *Dans le Problème 6.1, supposons que $\mathcal{X}$ et $\mathcal{Y}$ soient de dimension finie. Soit $f \in \Gamma_0(\mathcal{X})$ une fonction de Legendre telle que $\mathcal{S} \cap \operatorname{int}\operatorname{dom} f \neq \varnothing$, $L(\operatorname{int}\operatorname{dom} f) \subset \operatorname{int}\operatorname{dom} \psi$, $f \succcurlyeq \beta\psi \circ L$ pour un certain $\beta \in \left]0, +\infty\right[$, et $\operatorname{dom} f^*$ est ouvert. Supposons que ou bien $-L^*(\operatorname{ran}\nabla\psi) \subset \operatorname{dom}\varphi^*$ ou bien f soit cofinie. Soient $\varepsilon \in \left]0, \beta/(\beta+1)\right[$, $(\eta_n)_{n\in\mathbb{N}} \in \ell_+^1(\mathbb{N})$, et $(\gamma_n)_{n\in\mathbb{N}}$ une suite dans $\mathbb{R}$ telle que*

$$(\forall n \in \mathbb{N}) \quad \varepsilon \leqslant \gamma_n \leqslant \beta(1-\varepsilon) \quad \text{et} \quad (1+\eta_n)\gamma_n - \gamma_{n+1} \leqslant \beta\eta_n. \tag{6.17}$$

De plus, soit $x_0 \in \operatorname{int}\operatorname{dom} f$ et posons

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = \operatorname{Prox}^f_{\gamma_n\varphi}\big(\nabla f(x_n) - \gamma_n L^*\nabla\psi(Lx_n)\big). \tag{6.18}$$

Alors il existe $\overline{x} \in \mathcal{S}$ tel que $x_n \to \overline{x}$.

Notre cadre permet de résoudre le problème multivarié suivant.

Problème 6.9 Soient m et p des entiers strictement positifs, soient $(\mathcal{X}_i)_{1\leqslant i\leqslant m}$ et $(\mathcal{Y}_k)_{1\leqslant k\leqslant p}$ des espaces de Banach réels réflexifs. Pour tout $i\in\{1,\ldots,m\}$ et pour tout $k\in\{1,\ldots,p\}$, soit $\varphi_i\in\Gamma_0(\mathcal{X}_i)$, soit $\psi_k\in\Gamma_0(\mathcal{Y}_k)$ Gâteaux différentiable sur $\operatorname{int}\operatorname{dom}\psi_k\neq\varnothing$, et soit $L_{ik}\in\mathcal{B}\,(\mathcal{X}_i,\mathcal{Y}_k)$. Le problème est de

$$\underset{x_1\in\mathcal{X}_1,\ldots,x_m\in\mathcal{X}_m}{\text{minimiser}}\quad \sum_{i=1}^m\varphi_i(x_i)+\sum_{k=1}^p\psi_k\left(\sum_{i=1}^m L_{ik}x_i\right). \tag{6.19}$$

L'ensemble des solutions de (6.19) est noté $\mathcal{S}$.

Proposition 6.10 *Considérons le Problème 6.9. Pour tout $k\in\{1,\ldots,p\}$, supposons qu'il existe $\sigma_k\in\,]0,+\infty[$ tel que pxour tout $(y_{ik})_{1\leqslant i\leqslant m}$ et $(v_{ik})_{1\leqslant i\leqslant m}$ dans $\operatorname{int}\operatorname{dom}\psi_k$ satisfaisant $\sum_{i=1}^m y_{ik}\in\operatorname{int}\operatorname{dom}\psi_k$ et $\sum_{i=1}^m v_{ik}\in\operatorname{int}\operatorname{dom}\psi_k$, on a*

$$D^{\psi_k}\left(\sum_{i=1}^m y_{ik},\sum_{i=1}^m v_{ik}\right)\leqslant\sigma_k\sum_{i=1}^m D^{\psi_k}(y_{ik},v_{ik}). \tag{6.20}$$

Pour tout $i\in\{1,\ldots,m\}$, soit $f_i\in\Gamma_0(\mathcal{X}_i)$ une fonction de Legendre telle que $(\forall x_i\in\operatorname{int}\operatorname{dom}f_i)$ $D^{f_i}(x_i,\cdot)$ est coercive. Pour tout $k\in\{1,\ldots,p\}$, supposons que $\sum_{i=1}^m L_{ik}(\operatorname{int}\operatorname{dom}f_i)\subset\operatorname{int}\operatorname{dom}\psi_k$, que pour tout $i\in\{1,\ldots,m\}$, il existe $\beta_{ik}\in\,]0,+\infty[$ tel que $f_i\succcurlyeq\beta_{ik}\psi_k\circ L_{ik}$, et posons $\beta_k=\min_{1\leqslant i\leqslant m}\beta_{ik}$. De plus, supposons que $\mathcal{S}\cap\bigtimes_{i=1}^m\operatorname{int}\operatorname{dom}f_i\neq\varnothing$ et que ou bien $(\forall i\in\{1,\ldots,m\})$ f_i soit cofinie ou bien $(\forall i\in\{1,\ldots,m\})$ φ_i soit cofinie. Soient $\varepsilon\in\big]0,1/\big(1+\sum_{k=1}^p\sigma_k\beta_k^{-1}\big)\big[$, $(\eta_n)_{n\in\mathbb{N}}\in\ell_+^1(\mathbb{N})$, et $(\gamma_n)_{n\in\mathbb{N}}$ une suite dans $\mathbb{R}$ telle que

$$(\forall n\in\mathbb{N})\quad\varepsilon\leqslant\gamma_n\leqslant\frac{1-\varepsilon}{\sum_{k=1}^p\sigma_k\beta_k^{-1}}\quad\text{et}\quad(1+\eta_n)\gamma_n-\gamma_{n+1}\leqslant\frac{\eta_n}{\sum_{k=1}^p\sigma_k\beta_k^{-1}}. \tag{6.21}$$

De plus, soit $(x_{i,0})_{1\leqslant i\leqslant m}\in\bigtimes_{i=1}^m\operatorname{int}\operatorname{dom}f_i$ et posons

$$\begin{array}{l}\textit{pour } n=0,1,\ldots\\ \left\lfloor\begin{array}{l}\textit{pour } i=1,\ldots,m\\ \left\lfloor\; x_{i,n+1}=\operatorname{Prox}^{f_i}_{\gamma_n\varphi_i}\big(\nabla f_i(x_{i,n})-\gamma_n\sum_{k=1}^p L_{ik}^*\nabla\psi_k\big(\sum_{j=1}^m L_{jk}x_{j,n}\big)\big).\right.\end{array}\right.\end{array} \tag{6.22}$$

Alors il existe $(\overline{x}_i)_{1\leqslant i\leqslant m}\in\mathcal{S}$ tel que les propriétés suivantes sont satisfaites :

(i) *Supposons que $\mathcal{S}\cap\bigtimes_{i=1}^m\overline{\operatorname{dom}}\,f_i$ soit un singleton. Alors $(\forall i\in\{1,\ldots,m\})$ $x_{i,n}\rightharpoonup\overline{x}_i$.*

(ii) *Pour tout $i\in\{1,\ldots,m\}$ et pour tout $k\in\{1,\ldots,p\}$, supposons que ∇f_i et $\nabla\psi_k$ soient faiblement séquentiellement continus, et qu'une des conditions suivantes soit satisfaite.*

  (a) $\operatorname{dom}\varphi_i\subset\operatorname{int}\operatorname{dom}f_i$.

(b) $\operatorname{dom} f_i^*$ *est ouvert et* ∇f_i^* *est faiblement séquentiellement continu.*

Alors $(\forall i \in \{1,\ldots,m\})\ x_{i,n} \rightharpoonup \overline{x}_i$.

Nous faisons des comparaisons avec les algorithmes existants pour résoudre le Problème 6.1.

- Le cadre des espaces de Banach : les méthodes des Chapitres 2, 3, et 4 de la thèse demandent plus de calculs : il faut évaluer deux opérateurs proximaux, des opérateurs linéaires L et L^*, la projection sur le demi-espace affine et le calcul de $\boldsymbol{Q}^f$, alors que l'algorithme explicite-implicite (6.16) demande les calculs d'un seul opérateur proximal, des opérateurs linéaires L et L^*, et des gradients de f et ψ.
- Le cadre d'espaces euclidiens :
  - La méthode explicite-implicite classique n'est pas applicable si ψ n'est pas à valeurs rélles, ou $\nabla\psi$ n'est pas lipschitzien. De plus, l'opérateur proximal au sens de Bregman peut être plus facile à évaluer.
  - L'algorithme de Douglas-Rachford n'est pas applicable puisqu'on ne sait pas calculer $\operatorname{prox}_{\psi\circ L}$ en général.
- Avec l'algorithme monotone+antisymétrique de [12] :

L'algorithme (6.16)	L'algorithme de [12]
pour $n = 0,1,\ldots$ $\left\lfloor \begin{array}{l} y_n^* = \nabla f(x_n) - \gamma_n L^*\nabla\psi(Lx_n) \\ x_{n+1} = \operatorname{Prox}^f_{\gamma_n\varphi} y_n^*. \end{array} \right.$	pour $n = 0,1,\ldots$ $\left\lfloor \begin{array}{l} v_{1,n} = x_n - \gamma_n L^* y_n^* \\ v_{2,n}^* = y_n^* + \gamma_n L x_n \\ p_{1,n} = \operatorname{prox}_{\gamma_n\varphi} v_{1,n} \\ p_{2,n}^* = \operatorname{prox}_{\gamma_n\psi} v_{2,n}^* \\ q_{1,n} = p_{1,n} - \gamma_n L^* p_{2,n}^* \\ q_{2,n}^* = p_{2,n}^* + \gamma_n L p_{1,n} \\ x_{n+1} = x_n - v_{1,n} + q_{1,n} \\ y_{n+1}^* = y_n^* - v_{2,n}^* + q_{2,n}^*. \end{array} \right.$
Bilan : • Il faut évaluer un opérateur proximal, deux gradients et une fois les opérateurs linéaires L et L^*. • Peu de stockage en mémoire : une seule variable à stocker.	**Bilan :** • Il faut évaluer deux opérateurs proximaux et deux fois les opérateurs linéaires L et L^*. • Stockage en mémoire important : 8 variables à stocker.

6.2 Article en anglais

FORWARD-BACKWARD SPLITTING WITH BREGMAN DISTANCES [1]

6.2.1 Introduction

We consider the following composite convex minimization problem.

Problem 6.11 Let $\mathcal{X}$ and $\mathcal{Y}$ be reflexive real Banach spaces, let $\varphi \in \Gamma_0(\mathcal{X})$, let $\psi \in \Gamma_0(\mathcal{Y})$ be Gâteaux differentiable on $\operatorname{int}\operatorname{dom}\psi \neq \varnothing$, and let $L\colon \mathcal{X} \to \mathcal{Y}$ be a bounded linear operator. The problem is to

$$\underset{x\in\mathcal{X}}{\text{minimize}}\ \ \varphi(x) + \psi(Lx). \tag{6.23}$$

The set of solutions to (6.23) is denoted by $\mathcal{S}$.

A particular instance of (6.23) is when $\psi = D^g(\cdot, r)$, where $g \in \Gamma_0(\mathcal{Y})$ is Gâteaux differentiable on $\operatorname{int}\operatorname{dom} g \ni r$, i.e.,

$$\underset{x\in\mathcal{X}}{\text{minimize}}\ \ \varphi(x) + D^g(Lx, r). \tag{6.24}$$

This model provides a framework for many problems arising in applied mathematics. For instance, when $\mathcal{X}$ and $\mathcal{Y}$ are Euclidean spaces and g is Boltzmann-Shannon entropy, it captures many problems in information theory and signal recovery [10]. Besides, the nearness matrix problem [21] and the log-determinant minimization problem [15] can be also regarded as special cases of (6.24).

An objective is constructing effective splitting methods, i.e, the methods that activate each function in the model separately, to solve Problem 6.11 (see [19] for more discussions). It was shown in [19] that if $\mathcal{X}$ and $\mathcal{Y}$ are Hilbert spaces and if ψ possess a β^{-1}-Lipschitz continuous gradient for some $\beta \in \left]0, +\infty\right[$, then Problem 6.11 can be solved by the standard forward-backward algorithm

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = \operatorname{prox}_{\gamma_n\varphi}\big(x_n - \gamma_n L^*(\nabla\psi(Lx_n))\big), \quad \text{where} \quad 0 < \gamma_n < 2\beta. \tag{6.25}$$

Here, $(\operatorname{prox}_{\gamma_n\varphi})_{n\in\mathbb{N}}$ are the Moreau proximity operators [25]. However, many problems in applications do not conform to these hypotheses, for example when $\mathcal{X}$ and $\mathcal{Y}$ are Euclidean spaces and ψ is Boltzmann-Shannon entropy. This type of functions appears in many problems in image and signal processing, in statistics, and in machine learning [2, 14, 15, 22, 23, 24]. Another difficulty in the implementation of (6.25) is that the operators $(\operatorname{prox}_{\gamma_n\varphi})_{n\in\mathbb{N}}$ are not always easy to evaluate. The objective of the present

1. V. Q. Nguyen, Forward-backward splitting with Bregman distances, *submitted*.

paper is to propose a version of the forward-backward splitting algorithm to solve Problem 6.11, which is so far limited to Hilbert spaces, in the general framework of reflexive real Banach spaces. This algorithm, which employs Bregman distance-based proximity operators, provides new algorithms in the framework of Euclidean spaces, which are, in some instances, more favorable than the standard forward-backward splitting algorithm. This framework can be applied in the case when ψ is not everywhere differentiable and in some instances, it requires less efforts in the computation of proximity operators than the classical framework. This paper revolves around the following definitions.

Definition 6.12 [5, 6] Let $\mathcal{X}$ be a reflexive real Banach space, let $\mathcal{X}^*$ be the topological dual space of $\mathcal{X}$, let $\langle\cdot,\cdot\rangle$ be the duality pairing between $\mathcal{X}$ and $\mathcal{X}^*$, let $f\colon \mathcal{X}\to\,]-\infty,+\infty]$ be a lower semicontinuous convex function that is Gâteaux differentiable on $\operatorname{int}\operatorname{dom} f\neq\varnothing$, let $f^*\colon\mathcal{X}^*\to\,]-\infty,+\infty]\colon x^*\mapsto\sup_{x\in\mathcal{X}}(\langle x,x^*\rangle-f(x))$ be conjugate of f, and let

$$\partial f\colon\mathcal{X}\to 2^{\mathcal{X}^*}\colon x\mapsto\big\{x^*\in\mathcal{X}^*\;\big|\;(\forall y\in\mathcal{X})\;\langle y-x,x^*\rangle+f(x)\leqslant f(y)\big\},\tag{6.26}$$

be the Moreau subdifferential of f. The *Bregman distance* associated with f is

$$\begin{aligned} D^f\colon\mathcal{X}\times\mathcal{X}&\to[0,+\infty]\\ (x,y)&\mapsto\begin{cases} f(x)-f(y)-\big\langle x-y,\nabla f(y)\big\rangle, & \text{if } y\in\operatorname{int}\operatorname{dom} f;\\ +\infty, & \text{otherwise.}\end{cases}\end{aligned}\tag{6.27}$$

In addition, f is a *Legendre function* if it is *essentially smooth* in the sense that ∂f is both locally bounded and single-valued on its domain, and *essentially strictly convex* in the sense that ∂f^* is locally bounded on its domain and f is strictly convex on every convex subset of $\operatorname{dom}\partial f$. Let C be a closed convex subset of $\mathcal{X}$ such that $C\cap\operatorname{int}\operatorname{dom} f\neq\varnothing$. The *Bregman projector* onto C induced by f is

$$\begin{aligned} P_C^f\colon \operatorname{int}\operatorname{dom} f&\to C\cap\operatorname{int}\operatorname{dom} f\\ y&\mapsto\underset{x\in C}{\operatorname{argmin}}\,D^f(x,y),\end{aligned}\tag{6.28}$$

and the D^f*-distance* to C is the function

$$\begin{aligned} D_C^f\colon\mathcal{X}&\to[0,+\infty]\\ y&\mapsto\inf D^f(C,y).\end{aligned}\tag{6.29}$$

The paper is organized as follows. In Section 6.2.2, we provide some preliminary results. We present the forward-backward splitting algorithm in reflexive Banach spaces in Section 6.2.3. Section 6.2.4 is devoted to an application of our result to multivariate minimization problem together with examples.

Notation and background. The norm of a Banach space is denoted by $\|\cdot\|$. The symbols $\rightharpoonup$ and $\to$ represent respectively weak and strong convergence. The set of weak sequential cluster points of a sequence $(x_n)_{n\in\mathbb{N}}$ is denoted by $\mathfrak{W}(x_n)_{n\in\mathbb{N}}$. Let $M\colon \mathcal{X} \to 2^{\mathcal{X}^*}$. The domain of M is $\operatorname{dom} M = \{x \in \mathcal{X} \mid Mx \neq \varnothing\}$ and the range of M is $\operatorname{ran} M = \{x^* \in \mathcal{X}^* \mid (\exists\, x \in \mathcal{X})\, x^* \in Mx\}$. Let $f\colon \mathcal{X} \to \left]-\infty,+\infty\right]$. Then f is cofinite if $\operatorname{dom} f^* = \mathcal{X}^*$, is coercive if $\lim_{\|x\|\to+\infty} f(x) = +\infty$, is supercoercive if $\lim_{\|x\|\to+\infty} f(x)/\|x\| = +\infty$, and is uniformly convex at $x \in \operatorname{dom} f$ if there exists an increasing function $\phi\colon \left[0,+\infty\right[\to [0,+\infty]$ that vanishes only at 0 such that

$$(\forall y \in \operatorname{dom} f)(\forall \alpha \in \left]0,1\right[) \quad f(\alpha x+(1-\alpha)y)+\alpha(1-\alpha)\phi(\|x-y\|) \leqslant \alpha f(x)+(1-\alpha)f(y). \tag{6.30}$$

Denote by $\Gamma_0(\mathcal{X})$ the class of all lower semicontinuous convex functions $f\colon \mathcal{X} \to \left]-\infty,+\infty\right]$ such that $\operatorname{dom} f = \{x \in \mathcal{X} \mid f(x) < +\infty\} \neq \varnothing$. Let $f \in \Gamma_0(\mathcal{X})$. The set of global minimizers of f is denoted by $\operatorname{Argmin} f$. Finally, $\ell_+^1(\mathbb{N})$ is the set of all summable sequences in $\left[0,+\infty\right[$.

6.2.2 Preminarily results

First, we recall the following definitions and results.

Definition 6.13 [26] Let $\mathcal{X}$ be a reflexive real Banach space and let $f \in \Gamma_0(\mathcal{X})$ be Gâteaux differentiable on $\operatorname{int}\operatorname{dom} f \neq \varnothing$. Then

$$\mathcal{F}(f) = \{g \in \Gamma_0(\mathcal{X}) \mid g \text{ is Gâteaux differentiable on } \operatorname{int}\operatorname{dom} g = \operatorname{int}\operatorname{dom} f\}. \tag{6.31}$$

Moreover, if g_1 and g_2 are in $\mathcal{F}(f)$, then

$$g_1 \succcurlyeq g_2 \quad \Leftrightarrow \quad (\forall x \in \operatorname{dom} f)(\forall y \in \operatorname{int}\operatorname{dom} f) \quad D^{g_1}(x,y) \geqslant D^{g_2}(x,y). \tag{6.32}$$

For every $\alpha \in \left[0,+\infty\right[$, set

$$\mathcal{P}_\alpha(f) = \{g \in \mathcal{F}(f) \mid g \succcurlyeq \alpha f\}. \tag{6.33}$$

Definition 6.14 [26] Let $\mathcal{X}$ be a reflexive real Banach space, let $f \in \Gamma_0(\mathcal{X})$ be Gâteaux differentiable on $\operatorname{int}\operatorname{dom} f \neq \varnothing$, let $(f_n)_{n\in\mathbb{N}}$ be in $\mathcal{F}(f)$, let $(x_n)_{n\in\mathbb{N}} \in (\operatorname{int}\operatorname{dom} f)^{\mathbb{N}}$, and let $C \subset \mathcal{X}$ be such that $C \cap \operatorname{dom} f \neq \varnothing$. Then $(x_n)_{n\in\mathbb{N}}$ is :

(i) *quasi-Bregman monotone* with respect to C relative to $(f_n)_{n\in\mathbb{N}}$ if

$$\begin{aligned}(\exists(\eta_n)_{n\in\mathbb{N}} \in \ell_+^1(\mathbb{N}))(\forall x \in C \cap \operatorname{dom} f)(\exists(\varepsilon_n)_{n\in\mathbb{N}} \in \ell_+^1(\mathbb{N}))(\forall n \in \mathbb{N})\\ D^{f_{n+1}}(x,x_{n+1}) \leqslant (1+\eta_n)D^{f_n}(x,x_n)+\varepsilon_n;\end{aligned} \tag{6.34}$$

(ii) *stationarily quasi-Bregman monotone* with respect to C relative to $(f_n)_{n\in\mathbb{N}}$ if

$$\begin{aligned}(\exists(\varepsilon_n)_{n\in\mathbb{N}} \in \ell_+^1(\mathbb{N}))(\exists(\eta_n)_{n\in\mathbb{N}} \in \ell_+^1(\mathbb{N}))(\forall x \in C \cap \operatorname{dom} f)(\forall n \in \mathbb{N})\\ D^{f_{n+1}}(x,x_{n+1}) \leqslant (1+\eta_n)D^{f_n}(x,x_n)+\varepsilon_n.\end{aligned} \tag{6.35}$$

Condition 6.15 [6, Condition 4.4] Let $\mathcal{X}$ be a reflexive real Banach space and let $f \in \Gamma_0(\mathcal{X})$ be Gâteaux differentiable on $\operatorname{int}\operatorname{dom} f \neq \varnothing$. For every bounded sequences $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ in $\operatorname{int}\operatorname{dom} f$,

$$D^f(x_n, y_n) \to 0 \quad \Rightarrow \quad x_n - y_n \to 0. \tag{6.36}$$

Proposition 6.16 [26] *Let $\mathcal{X}$ be a reflexive real Banach space, let $f \in \Gamma_0(\mathcal{X})$ be Gâteaux differentiable on $\operatorname{int}\operatorname{dom} f \neq \varnothing$, let $\alpha \in \left]0, +\infty\right[$, let $(f_n)_{n\in\mathbb{N}}$ be in $\mathcal{P}_\alpha(f)$, let $(x_n)_{n\in\mathbb{N}} \in (\operatorname{int}\operatorname{dom} f)^{\mathbb{N}}$, let $C \subset \mathcal{X}$ be such that $C \cap \operatorname{int}\operatorname{dom} f \neq \varnothing$, and let $x \in C \cap \operatorname{int}\operatorname{dom} f$. Suppose that $(x_n)_{n\in\mathbb{N}}$ is quasi-Bregman monotone with respect to C relative to $(f_n)_{n\in\mathbb{N}}$. Then the following hold.*

(i) $(D^{f_n}(x, x_n))_{n\in\mathbb{N}}$ *converges.*

(ii) *Suppose that $D^f(x, \cdot)$ is coercive. Then $(x_n)_{n\in\mathbb{N}}$ is bounded.*

Proposition 6.17 [26] *Let $\mathcal{X}$ be a reflexive real Banach space, let $f \in \Gamma_0(\mathcal{X})$ be Gâteaux differentiable on $\operatorname{int}\operatorname{dom} f \neq \varnothing$, let $(x_n)_{n\in\mathbb{N}} \in (\operatorname{int}\operatorname{dom} f)^{\mathbb{N}}$, let $C \subset \mathcal{X}$ be such that $C \cap \operatorname{int}\operatorname{dom} f \neq \varnothing$, let $(\eta_n)_{n\in\mathbb{N}} \in \ell_+^1(\mathbb{N})$, let $\alpha \in \left]0, +\infty\right[$, and let $(f_n)_{n\in\mathbb{N}}$ in $\mathcal{P}_\alpha(f)$ be such that $(\forall n \in \mathbb{N})\ (1 + \eta_n) f_n \succcurlyeq f_{n+1}$. Suppose that $(x_n)_{n\in\mathbb{N}}$ is quasi-Bregman monotone with respect to C relative to $(f_n)_{n\in\mathbb{N}}$, that there exists $g \in \mathcal{F}(f)$ such that for every $n \in \mathbb{N}$, $g \succcurlyeq f_n$, and that, for every $y_1 \in \mathcal{X}$ and every $y_2 \in \mathcal{X}$,*

$$\begin{cases} y_1 \in \mathfrak{W}(x_n)_{n\in\mathbb{N}} \cap C \\ y_2 \in \mathfrak{W}(x_n)_{n\in\mathbb{N}} \cap C \\ \big(\langle y_1 - y_2, \nabla f_n(x_n)\rangle\big)_{n\in\mathbb{N}} \quad \text{converges} \end{cases} \quad \Rightarrow \quad y_1 = y_2. \tag{6.37}$$

Moreover, suppose that $(\forall x \in \operatorname{int}\operatorname{dom} f)$ $D^f(x, \cdot)$ is coercive. Then $(x_n)_{n\in\mathbb{N}}$ converges weakly to a point in $C \cap \operatorname{int}\operatorname{dom} f$ if and only if $\mathfrak{W}(x_n)_{n\in\mathbb{N}} \subset C \cap \operatorname{int}\operatorname{dom} f$.

Proposition 6.18 [26] *Let $\mathcal{X}$ be a reflexive real Banach space, let $f \in \Gamma_0(\mathcal{X})$ be a Legendre function, let $\alpha \in \left]0, +\infty\right[$, let $(f_n)_{n\in\mathbb{N}}$ be in $\mathcal{P}_\alpha(f)$, let $(x_n)_{n\in\mathbb{N}} \in (\operatorname{int}\operatorname{dom} f)^{\mathbb{N}}$, and let C be a closed convex subset of $\mathcal{X}$ such that $C \cap \operatorname{int}\operatorname{dom} f \neq \varnothing$. Suppose that $(x_n)_{n\in\mathbb{N}}$ is stationarily quasi-Bregman monotone with respect to C relative to $(f_n)_{n\in\mathbb{N}}$, that f satisfies Condition 6.15, and that $(\forall x \in \operatorname{int}\operatorname{dom} f)$ $D^f(x, \cdot)$ is coercive. In addition, suppose that there exists $\beta \in \left]0, +\infty\right[$ such that $(\forall n \in \mathbb{N})\ \beta f \succcurlyeq f_n$. Then $(x_n)_{n\in\mathbb{N}}$ converges strongly to a point in $C \cap \overline{\operatorname{dom}} f$ if and only if $\underline{\lim}\, D_C^f(x_n) = 0$.*

We discuss some basic properties of a type of Bregman distance-based proximity operators in the following proposition.

Proposition 6.19 *Let $\mathcal{X}$ be a reflexive real Banach space, let $f \in \Gamma_0(\mathcal{X})$ be Gâteaux differentiable on $\operatorname{int}\operatorname{dom} f \neq \varnothing$, let $\varphi \in \Gamma_0(\mathcal{X})$, and let*

$$\begin{aligned} \operatorname{Prox}_\varphi^f \colon \mathcal{X}^* &\to 2^{\mathcal{X}} \\ x^* &\mapsto \big\{x \in \mathcal{X} \;\big|\; \varphi(x) + f(x) - \langle x, x^*\rangle = \min\big(\varphi + f - x^*\big)(\mathcal{X}) < +\infty\big\} \end{aligned} \tag{6.38}$$

be f-proximity operator *of φ. Then the following hold.*

(i) $\operatorname{ran}\operatorname{Prox}_\varphi^f \subset \operatorname{dom} f \cap \operatorname{dom}\varphi$ *and* $\operatorname{Prox}_\varphi^f = (\partial(f+\varphi))^{-1}$.

(ii) *Suppose that* $\operatorname{dom}\varphi \cap \operatorname{int}\operatorname{dom} f \neq \varnothing$ *and that* $\operatorname{dom}\partial f \cap \operatorname{dom}\partial\varphi \subset \operatorname{int}\operatorname{dom} f$. *Then the following hold.*

  (a) $\operatorname{ran}\operatorname{Prox}_\varphi^f \subset \operatorname{int}\operatorname{dom} f$ *and* $\operatorname{Prox}_\varphi^f = (\nabla f + \partial\varphi)^{-1}$.

  (b) $\operatorname{int}(\operatorname{dom} f^* + \operatorname{dom}\varphi^*) \subset \operatorname{dom}\operatorname{Prox}_\varphi^f$.

  (c) *Suppose that* $f|_{\operatorname{int}\operatorname{dom} f}$ *is strictly convex. Then* $\operatorname{Prox}_\varphi^f$ *is single-valued on its domain.*

Proof. Let us fix $x^* \in \mathcal{X}^*$ and define $f_{x^*} \colon \mathcal{X} \to \left]-\infty,+\infty\right] : x \mapsto f(x) - \langle x, x^*\rangle + f^*(x^*)$. Then $\operatorname{dom} f_{x^*} = \operatorname{dom} f$ and $\varphi + f_{x^*} \in \Gamma_0(\mathcal{X})$. Moreover, $\partial(\varphi + f_{x^*}) = \partial(\varphi + f) - x^*$.

(i) : By definition, $\operatorname{ran}\operatorname{Prox}_\varphi^f \subset \operatorname{dom} f \cap \operatorname{dom}\varphi$. Suppose that $\operatorname{dom} f \cap \operatorname{dom}\varphi \neq \varnothing$ and let $x \in \operatorname{dom} f \cap \operatorname{dom}\varphi$. Then

$$\begin{aligned} x \in \operatorname{Prox}_\varphi^f x^* &\Leftrightarrow 0 \in \partial(\varphi + f_{x^*})(x) \\ &\Leftrightarrow 0 \in \partial(\varphi + f)(x) - x^* \\ &\Leftrightarrow x^* \in \partial(\varphi + f)(x) \\ &\Leftrightarrow x \in \big(\partial(\varphi + f)\big)^{-1}(x^*). \end{aligned} \tag{6.39}$$

(ii) : Suppose that $x^* \in \operatorname{int}(\operatorname{dom} f^* + \operatorname{dom}\varphi^*)$. Since $\operatorname{dom}\varphi \cap \operatorname{int}\operatorname{dom} f \neq \varnothing$, it follows from [1, Theorem 1.1] and [29, Theorem 2.1.3(ix)] that

$$x^* \in \operatorname{int}(\operatorname{dom} f^* + \operatorname{dom}\varphi^*) = \operatorname{int}\operatorname{dom}(f+\varphi)^*. \tag{6.40}$$

(ii)(a) : Since $\operatorname{dom}\varphi \cap \operatorname{int}\operatorname{dom} f \neq \varnothing$, $\partial(\varphi + f) = \partial\varphi + \partial f$ by [1, Corollary 2.1], and hence (i) yields

$$\operatorname{ran}\operatorname{Prox}_\varphi^f = \operatorname{dom}\partial(f+\varphi) = \operatorname{dom}(\partial f + \partial\varphi) = \operatorname{dom}\partial f \cap \operatorname{dom}\partial\varphi \subset \operatorname{int}\operatorname{dom} f. \tag{6.41}$$

In turn, $\operatorname{ran}\operatorname{Prox}_\varphi^f \subset \operatorname{dom}\varphi \cap \operatorname{int}\operatorname{dom} f$. We now prove that $\operatorname{Prox}_\varphi^f = (\nabla f + \partial\varphi)^{-1}$. Note that $\operatorname{dom}(\nabla f + \partial\varphi) \subset \operatorname{dom}\varphi \cap \operatorname{int}\operatorname{dom} f$. Let $x \in \operatorname{dom}\varphi \cap \operatorname{int}\operatorname{dom} f$. Then $\partial(f+\varphi)(x) = \partial f(x) + \partial\varphi(x) = \nabla f(x) + \partial\varphi(x)$ and therefore,

$$x \in \operatorname{Prox}_\varphi^f x^* \Leftrightarrow x^* \in \partial(f+\varphi)(x) = \nabla f(x) + \partial\varphi(x) \Leftrightarrow x \in (\nabla f + \partial\varphi)^{-1}(x^*). \tag{6.42}$$

(ii)(b) : We derive from (6.40) and [5, Fact 3.1] that $\varphi + f_{x^*}$ is coercive. Hence, by [29, Theorem 2.5.1], $\varphi + f_{x^*}$ admits at least one minimizer, i.e., $x^* \in \operatorname{dom}\operatorname{Prox}_\varphi^f$.

(ii)(c) : Since $f|_{\operatorname{int}\operatorname{dom} f}$ is strictly convex, so is $(\varphi + f_{x^*})|_{\operatorname{int}\operatorname{dom} f}$ and thus, in view of (ii)(b), $\varphi + f_{x^*}$ admits a unique minimizer on $\operatorname{int}\operatorname{dom} f$. However, since

$$\operatorname{Argmin}(\varphi + f_{x^*}) = \operatorname{ran}\operatorname{Prox}_\varphi^f \subset \operatorname{int}\operatorname{dom} f, \tag{6.43}$$

it follows that $\varphi + f_{x^*}$ admits a unique minimizer and that $\operatorname{Prox}_\varphi^f$ is therefore single-valued. □

Proposition 6.20 *Let m be a strictly positive integer, let $(\mathcal{X}_i)_{1\leqslant i\leqslant m}$ be reflexive real Banach spaces, and let $\mathcal{X}$ be the vector product space $\bigtimes_{i=1}^{m}\mathcal{X}_i$ equipped with the norm $x = (x_i)_{1\leqslant i\leqslant m} \mapsto \sqrt{\sum_{i=1}^m \|x_i\|^2}$. For every $i \in \{1,\ldots,m\}$, let $f_i \in \Gamma_0(\mathcal{X}_i)$ be a Legendre function and let $\varphi_i \in \Gamma_0(\mathcal{X}_i)$ be such that $\operatorname{dom}\varphi_i \cap \operatorname{int}\operatorname{dom} f_i \neq \varnothing$. Set $f\colon \mathcal{X} \to \left]-\infty,+\infty\right] : x \mapsto \sum_{i=1}^m f_i(x_i)$ and $\varphi\colon \mathcal{X} \to \left]-\infty,+\infty\right] : x \mapsto \sum_{i=1}^m \varphi_i(x_i)$. Then*

$$\Big(\forall x^* = (x_i^*)_{1\leqslant i\leqslant m} \in \bigtimes_{i=1}^{m}\operatorname{int}(\operatorname{dom} f_i^* + \operatorname{dom}\varphi_i^*)\Big) \quad \operatorname{Prox}_{\varphi}^{f} x^* = \big(\operatorname{Prox}_{\varphi_i}^{f_i} x_i^*\big)_{1\leqslant i\leqslant m}. \tag{6.44}$$

Proof. First, we observe that $\mathcal{X}^*$ is the vector product space $\bigtimes_{i=1}^{m}\mathcal{X}_i^*$ equipped with the norm $x^* = (x_i^*)_{1\leqslant i\leqslant m} \mapsto \sqrt{\sum_{i=1}^m \|x_i^*\|^2}$. Next, we derive from the definition of f that $\operatorname{dom} f = \bigtimes_{i=1}^{m}\operatorname{dom} f_i$ and that

$$\partial f\colon \mathcal{X} \to 2^{\mathcal{X}^*} : (x_i)_{1\leqslant i\leqslant m} \mapsto \bigtimes_{i=1}^{m} \partial f_i(x_i). \tag{6.45}$$

Thus, ∂f is single-valued on

$$\operatorname{dom}\partial f = \bigtimes_{i=1}^{m}\operatorname{dom}\partial f_i = \bigtimes_{i=1}^{m}\operatorname{int}\operatorname{dom} f_i = \operatorname{int}\Big(\bigtimes_{i=1}^{m}\operatorname{dom} f_i\Big) = \operatorname{int}\operatorname{dom} f. \tag{6.46}$$

Likewise, since

$$f^*\colon \mathcal{X}^* \to \left]-\infty,+\infty\right] : (x_i^*)_{1\leqslant i\leqslant m} \mapsto \sum_{i=1}^{m} f_i^*(x_i^*), \tag{6.47}$$

we deduce that ∂f^* is single-valued on $\operatorname{dom}\partial f^* = \operatorname{int}\operatorname{dom} f^*$. Consequently, [5, Theorems 5.4 and 5.6] assert that

$$f \text{ is a Legendre function.} \tag{6.48}$$

In addition,

$$\operatorname{dom}\varphi\cap\operatorname{int}\operatorname{dom} f = \Big(\bigtimes_{i=1}^{m}\operatorname{dom}\varphi_i\Big)\cap\Big(\bigtimes_{i=1}^{m}\operatorname{int}\operatorname{dom} f_i\Big) = \bigtimes_{i=1}^{m}(\operatorname{dom}\varphi_i\cap\operatorname{int}\operatorname{dom} f_i) \neq \varnothing. \tag{6.49}$$

Hence, Proposition 6.19(ii)(b)&(ii)(c) assert that $\operatorname{int}(\operatorname{dom} f^* + \operatorname{dom}\varphi^*) \subset \operatorname{dom}\operatorname{Prox}_{\varphi}^{f}$ and $\operatorname{Prox}_{\varphi}^{f}$ is single-valued on its domain. Now set $x = \operatorname{Prox}_{\varphi}^{f} x^*$ and $q = (\operatorname{Prox}_{\varphi_i}^{f_i} x_i^*)_{1\leqslant i\leqslant m}$. We derive from Proposition 6.19(ii)(a) that

$$x = \operatorname{Prox}_{\varphi}^{f} x^* \Leftrightarrow x = (\nabla f + \partial\varphi)^{-1}(x^*) \Leftrightarrow x^* - \nabla f(x) \in \partial\varphi(x). \tag{6.50}$$

Consequently, by invoking (6.26), we get

$$(\forall z \in \operatorname{dom}\varphi) \quad \big\langle z - x, x^* - \nabla f(x)\big\rangle + \varphi(x) \leqslant \varphi(z). \tag{6.51}$$

Upon setting $z = q$ in (6.51), we obtain

$$\big\langle q - x, x^* - \nabla f(x) \big\rangle + \varphi(x) \leqslant \varphi(q). \tag{6.52}$$

For every $i \in \{1, \ldots, m\}$, let us set $q_i = \operatorname{Prox}^{f_i}_{\varphi_i} x_i^*$. The same characterization as in (6.51) yields

$$(\forall i \in \{1, \ldots, m\})(\forall z_i \in \operatorname{dom} \varphi_i) \quad \big\langle z_i - q_i, x_i^* - \nabla f_i(q_i) \big\rangle + \varphi_i(q_i) \leqslant \varphi_i(z_i). \tag{6.53}$$

By summing these inequalities over $i \in \{1, \ldots, m\}$, we obtain

$$(\forall z \in \operatorname{dom} \varphi) \quad \big\langle z - q, x^* - \nabla f(q) \big\rangle + \varphi(q) \leqslant \varphi(z). \tag{6.54}$$

Upon setting $z = x$ in (6.54), we get

$$\big\langle x - q, \nabla f(x) - \nabla f(q) \big\rangle + \varphi(q) \leqslant \varphi(x). \tag{6.55}$$

Adding (6.52) and (6.55) yields

$$\big\langle x - q, \nabla f(x) - \nabla f(q) \big\rangle \leqslant 0. \tag{6.56}$$

Now suppose that $x \neq q$. Since $f|_{\operatorname{int} \operatorname{dom} f}$ is strictly convex, it follows from [29, Theorem 2.4.4(ii)] that ∇f is strictly monotone, i.e.,

$$\big\langle x - q, \nabla f(x) - \nabla f(q) \big\rangle > 0, \tag{6.57}$$

and we reach a contradiction. □

In Hilbert spaces, the operator defined in (6.38) reduces to the Moreau's usual proximity operator $\operatorname{prox}_\varphi$ [25] if $f = \|\cdot\|^2/2$. We provide illustrations of instances in the standard Euclidean space $\mathbb{R}^m$ in which $\operatorname{Prox}^f_\varphi$ is easier to evaluate than $\operatorname{prox}_\varphi$.

Example 6.21 Let $\gamma \in \left]0, +\infty\right[$, let $\phi \in \Gamma_0(\mathbb{R})$ be such that $\operatorname{dom} \phi \cap \left]0, +\infty\right[\neq \varnothing$, and let ϑ be Boltzmann-Shannon entropy, i.e.,

$$\vartheta \colon \xi \mapsto \begin{cases} \xi \ln \xi - \xi, & \text{if } \xi \in \left]0, +\infty\right[; \\ 0, & \text{if } \xi = 0; \\ +\infty, & \text{otherwise.} \end{cases} \tag{6.58}$$

Set $\varphi \colon (\xi_i)_{1 \leqslant i \leqslant m} \mapsto \sum_{i=1}^m \phi(\xi_i)$ and $f \colon (\xi_i)_{1 \leqslant i \leqslant m} \mapsto \sum_{i=1}^m \vartheta(\xi_i)$. Note that f is a supercoercive Legendre function [4, Sections 5 and 6], and hence, Proposition 6.19(ii)(b) asserts that $\operatorname{dom} \operatorname{Prox}^f_\varphi = \mathbb{R}^m$. Let $(\xi_i)_{1 \leqslant i \leqslant m} \in \mathbb{R}^m$, set $(\eta_i)_{1 \leqslant i \leqslant m} = \operatorname{Prox}^f_{\gamma\varphi}(\xi_i)_{1 \leqslant i \leqslant m}$, let W be the Lambert function [20], i.e., the inverse of $\xi \mapsto \xi e^\xi$ on $\left[0, +\infty\right[$, and let $i \in \{1, \ldots, m\}$. Then η_i can be computed as follows.

(i) Let $\omega \in \mathbb{R}$ and suppose that

$$\phi\colon \xi \mapsto \begin{cases} \xi \ln \xi - \omega\xi, & \text{if } \ \xi \in \left]0,+\infty\right[; \\ 0, & \text{if } \ \xi = 0; \\ +\infty, & \text{otherwise.} \end{cases} \tag{6.59}$$

Then $\eta_i = e^{(\xi_i+\omega-1)/(\gamma+1)}$.

(ii) Let $p \in \left[1,+\infty\right[$ and suppose that either $\phi = |\cdot|^p/p$ or

$$\phi\colon \xi \mapsto \begin{cases} \xi^p/p, & \text{if } \ \xi \in \left[0,+\infty\right[; \\ +\infty, & \text{otherwise.} \end{cases} \tag{6.60}$$

Then

$$\eta_i = \begin{cases} \left(\dfrac{W(\gamma(p-1)e^{(p-1)\xi_i})}{\gamma(p-1)}\right)^{\frac{1}{p-1}}, & \text{if } \ p \in \left]1,+\infty\right[; \\ e^{\xi_i - \gamma}, & \text{if } \ p = 1. \end{cases} \tag{6.61}$$

(iii) Let $p \in \left[1,+\infty\right[$ and suppose that

$$\phi\colon \xi \mapsto \begin{cases} \xi^{-p}/p, & \text{if } \ \xi \in \left]0,+\infty\right[; \\ +\infty, & \text{otherwise.} \end{cases} \tag{6.62}$$

Then

$$\eta_i = \left(\frac{W(\gamma(p+1)e^{-(p+1)\xi_i})}{\gamma(p+1)}\right)^{\frac{-1}{p+1}}. \tag{6.63}$$

(iv) Let $p \in \left]0,1\right[$ and suppose that

$$\phi\colon \xi \mapsto \begin{cases} -\xi^p/p, & \text{if } \ \xi \in \left[0,+\infty\right[; \\ +\infty, & \text{otherwise.} \end{cases} \tag{6.64}$$

Then

$$\eta_i = \left(\frac{W(\gamma(1-p)e^{(p-1)\xi_i})}{\gamma(1-p)}\right)^{\frac{1}{p-1}}. \tag{6.65}$$

Example 6.22 Let $\phi \in \Gamma_0(\mathbb{R})$ be such that $\operatorname{dom}\phi \cap \left]0,1\right[\neq \varnothing$ and let ϑ be Fermi-Dirac entropy, i.e.,

$$\vartheta\colon \xi \mapsto \begin{cases} \xi \ln \xi - (1-\xi)\ln(1-\xi), & \text{if } \ \xi \in \left]0,1\right[; \\ 0 & \text{if } \ \xi \in \{0,1\}; \\ +\infty, & \text{otherwise.} \end{cases} \tag{6.66}$$

Set $\varphi\colon (\xi_i)_{1\leqslant i\leqslant m} \mapsto \sum_{i=1}^m \phi(\xi_i)$ and $f\colon (\xi_i)_{1\leqslant i\leqslant m} \mapsto \sum_{i=1}^m \vartheta(\xi_i)$. Note that f is a cofinite Legendre function [4, Sections 5 and 6], and hence Proposition 6.19(ii)(b) asserts that $\operatorname{dom}\operatorname{Prox}_\varphi^f = \mathbb{R}^m$. Let $(\xi_i)_{1\leqslant i\leqslant m} \in \mathbb{R}^m$, set $(\eta_i)_{1\leqslant i\leqslant m} = \operatorname{Prox}_\varphi^f(\xi_i)_{1\leqslant i\leqslant m}$, and let $i \in \{1,\ldots,m\}$. Then η_i can be computed as follows.

(i) Let $\omega \in \mathbb{R}$ and suppose that

$$\phi\colon \xi \mapsto \begin{cases} \xi\ln\xi - \omega\xi, & \text{if } \xi \in \left]0,+\infty\right[; \\ 0, & \text{if } \xi = 0; \\ +\infty, & \text{otherwise.} \end{cases} \tag{6.67}$$

Then $\eta_i = -e^{\xi_i+\omega-1}/2 + \sqrt{e^{2(\xi_i+\omega-1)}/4 + e^{\xi_i+\omega-1}}$.

(ii) Suppose that

$$\phi\colon \xi \mapsto \begin{cases} (1-\xi)\ln(1-\xi)+\xi, & \text{if } \xi \in \left]-\infty,1\right[; \\ 1, & \text{if } \xi = 1; \\ +\infty, & \text{otherwise.} \end{cases} \tag{6.68}$$

Then $\eta_i = 1 + e^{-\xi_i}/2 - \sqrt{e^{-\xi_i} + e^{-2\xi_i}/4}$.

Example 6.23 Let $f\colon (\xi_i)_{1\leqslant i\leqslant m} \mapsto \sum_{i=1}^m \vartheta(\xi_i)$, where ϑ is Hellinger-like function, i.e.,

$$\vartheta\colon \xi \mapsto \begin{cases} -\sqrt{1-\xi^2}, & \text{if } \xi \in [-1,1]; \\ +\infty, & \text{otherwise,} \end{cases} \tag{6.69}$$

let $\gamma \in \left]0,+\infty\right[$, and let $\varphi = f$. Since f is a cofinite Legendre function [4, Sections 5 and 6], Proposition 6.19(ii)(b) asserts that $\operatorname{dom}\operatorname{Prox}_{\gamma\varphi}^f = \mathbb{R}^m$. Let $(\xi_i)_{1\leqslant i\leqslant m} \in \mathbb{R}^m$, and set $(\eta_i)_{1\leqslant i\leqslant m} = \operatorname{Prox}_{\gamma\varphi}^f(\xi_i)_{1\leqslant i\leqslant m}$. Then $(\forall i \in \{1,\ldots,m\})\ \eta_i = \xi_i/\sqrt{(\gamma+1)^2+\xi_i^2}$.

Example 6.24 Let $\gamma \in \left]0,+\infty\right[$, let $\phi \in \Gamma_0(\mathbb{R})$ be such that $\operatorname{dom}\phi \cap \left]0,+\infty\right[\neq \varnothing$, and let ϑ be Burg entropy, i.e.,

$$\vartheta\colon \xi \mapsto \begin{cases} -\ln\xi, & \text{if } \xi \in \left]0,+\infty\right[; \\ +\infty, & \text{otherwise.} \end{cases} \tag{6.70}$$

Set $\varphi\colon (\xi_i)_{1\leqslant i\leqslant m} \mapsto \sum_{i=1}^m \phi(\xi_i)$ and $f\colon (\xi_i)_{1\leqslant i\leqslant m} \mapsto \sum_{i=1}^m \vartheta(\xi_i)$, let $(\xi_i)_{1\leqslant i\leqslant m} \in \mathbb{R}^m$, and set $(\eta_i)_{1\leqslant i\leqslant m} = \operatorname{Prox}_\varphi^f(\xi_i)_{1\leqslant i\leqslant m}$. Let $i \in \{1,\ldots,m\}$. Then η_i can be computed as follows.

(i) Suppose that $\phi = \vartheta$ and $\xi_i \in \left]-\infty,0\right]$. Then $\eta_i = -(1+\gamma)^{-1}\xi_i$.

(ii) Suppose that $\phi\colon \xi \mapsto \alpha|\xi|$ and $\xi_i \in \left]-\infty,\gamma\alpha\right]$. Then $\eta_i = (\gamma\alpha - \xi_i)^{-1}$.

The following result will be used subsequently.

Lemma 6.25 *Let $\mathcal{X}$ be a reflexive real Banach space, let $f \in \Gamma_0(\mathcal{X})$ be a Legendre function, let $x \in \operatorname{int\,dom} f$, and let $(x_n)_{n\in\mathbb{N}} \in (\operatorname{int\,dom} f)^{\mathbb{N}}$. Suppose that $(D^f(x, x_n))_{n\in\mathbb{N}}$ is bounded, that $\operatorname{dom} f^*$ is open, and that ∇f^* is weakly sequentially continuous. Then $\mathfrak{W}(x_n)_{n\in\mathbb{N}} \subset \operatorname{int\,dom} f$.*

Proof. [26, Proof of Theorem 4.1]. □

6.2.3 Forward-backward splitting in Banach spaces

The first result in this section is a version of the forward-backward splitting algorithm in reflexive real Banach spaces which employs different Bregman distance-based proximity operators over the iterations.

Theorem 6.26 *Consider the setting of Problem 6.11 and let $f \in \Gamma_0(\mathcal{X})$ be a Legendre function such that $\mathcal{S} \cap \operatorname{int\,dom} f \neq \varnothing$, $L(\operatorname{int\,dom} f) \subset \operatorname{int\,dom} \psi$, and $f \succcurlyeq \beta\psi \circ L$ for some $\beta \in \left]0, +\infty\right[$. Let $(\eta_n)_{n\in\mathbb{N}} \in \ell_+^1(\mathbb{N})$, let $\alpha \in \left]0, +\infty\right[$, and let $(f_n)_{n\in\mathbb{N}}$ be Legendre functions in $\mathcal{P}_\alpha(f)$ such that*

$$(\forall n \in \mathbb{N}) \quad (1 + \eta_n) f_n \succcurlyeq f_{n+1}. \tag{6.71}$$

Suppose that either $-L^*(\operatorname{ran} \nabla\psi) \subset \operatorname{dom} \varphi^*$ or $(\forall n \in \mathbb{N})$ f_n is cofinite. Let $\varepsilon \in \left]0, \alpha\beta/(\alpha\beta + 1)\right[$ and let $(\gamma_n)_{n\in\mathbb{N}}$ be a sequence in $\mathbb{R}$ such that

$$(\forall n \in \mathbb{N}) \quad \varepsilon \leqslant \gamma_n \leqslant \alpha\beta(1 - \varepsilon) \quad \text{and} \quad (1 + \eta_n)\gamma_n - \gamma_{n+1} \leqslant \alpha\beta\eta_n. \tag{6.72}$$

Furthermore, let $x_0 \in \operatorname{int\,dom} f$ and iterate

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = \operatorname{Prox}^{f_n}_{\gamma_n\varphi}\big(\nabla f_n(x_n) - \gamma_n L^*\nabla\psi(Lx_n)\big). \tag{6.73}$$

Suppose in addition that $(\forall x \in \operatorname{int\,dom} f)$ $D^f(x, \cdot)$ is coercive. Then $(x_n)_{n\in\mathbb{N}}$ is a bounded sequence in $\operatorname{int\,dom} f$ and $\mathfrak{W}(x_n)_{n\in\mathbb{N}} \subset \mathcal{S}$. Moreover, there exists $\overline{x} \in \mathcal{S}$ such that the following hold.

(i) *Suppose that $\mathcal{S} \cap \overline{\operatorname{dom}} f$ is a singleton. Then $x_n \rightharpoonup \overline{x}$.*

(ii) *Suppose that there exists $g \in \mathcal{F}(f)$ such that for every $n \in \mathbb{N}$, $g \succcurlyeq f_n$, and that, for every $y_1 \in \mathcal{X}$ and every $y_2 \in \mathcal{X}$,*

$$\begin{cases} y_1 \in \mathfrak{W}(x_n)_{n\in\mathbb{N}} \\ y_2 \in \mathfrak{W}(x_n)_{n\in\mathbb{N}} \\ \big(\langle y_1 - y_2, \nabla f_n(x_n) - \gamma_n L^*\nabla\psi(Lx_n)\rangle\big)_{n\in\mathbb{N}} \quad \text{converges} \end{cases} \Rightarrow \quad y_1 = y_2. \tag{6.74}$$

In addition, suppose that one of the following holds.

(a) $\mathcal{S} \subset \operatorname{int\,dom} f$.

(b) $\operatorname{dom} f^*$ *is open and* ∇f^* *is weakly sequentially continuous.*

Then $x_n \rightharpoonup \overline{x}$.

(iii) *Suppose that* f *satisfies Condition 6.15 and that one of the following holds.*

(a) φ *is uniformly convex at* $\overline{x}$.

(b) ψ *is uniformly convex at* $L\overline{x}$ *and there exists* $\kappa \in \left]0,+\infty\right[$ *such that* $(\forall x \in \mathcal{X})$ $\|Lx\| \geqslant \kappa\|x\|$.

(c) $\underline{\lim}\, D^f_{\mathrm{S}}(x_n) = 0$ *and there exists* $\mu \in \left]0,+\infty\right[$ *such that* $(\forall n \in \mathbb{N})$ $\mu f \succcurlyeq f_n$.

Then $x_n \to \overline{x}$.

Proof. We first derive from Proposition 6.19(ii)(c) that the operators $(\operatorname{Prox}^f_{\gamma_n\varphi})_{n\in\mathbb{N}}$ are single-valued on their domains. We also note that $x_0 \in \operatorname{int}\operatorname{dom} f$. Suppose that $x_n \in \operatorname{int}\operatorname{dom} f$ for some $n \in \mathbb{N}$. If f_n is cofinite then Proposition 6.19(ii)(b) yields

$$\nabla f_n(x_n) - \gamma_n L^*\nabla\psi(Lx_n) \in \mathcal{X}^* = \operatorname{dom}\operatorname{Prox}^{f_n}_{\gamma_n\varphi}. \tag{6.75}$$

Otherwise,

$$\begin{aligned}\nabla f_n(x_n) - \gamma_n L^*\nabla\psi(Lx_n) &\in \operatorname{int}\operatorname{dom} f_n^* + \gamma_n \operatorname{dom}\varphi^* = \operatorname{int}(\operatorname{int}\operatorname{dom} f_n^* + \gamma_n\operatorname{dom}\varphi^*)\\ &\subset \operatorname{int}(\operatorname{dom} f_n^* + \gamma_n\operatorname{dom}\varphi^*) = \operatorname{int}(\operatorname{dom} f_n^* + \operatorname{dom}(\gamma_n\varphi^*)).\end{aligned} \tag{6.76}$$

Since $\operatorname{int}(\operatorname{dom} f_n^* + \operatorname{dom}(\gamma_n\varphi^*)) \subset \operatorname{dom}\operatorname{Prox}^f_{\gamma_n\varphi}$ by Proposition 6.19(ii)(b), we deduce from (6.73), (6.75), (6.76), and Proposition 6.19(ii)(a) that x_{n+1} is a well-defined element in $\operatorname{ran}\operatorname{Prox}^{f_n}_{\gamma\varphi} = \operatorname{dom}\partial\varphi \cap \operatorname{int}\operatorname{dom} f_n = \operatorname{dom}\partial\varphi\cap\operatorname{int}\operatorname{dom} f \subset \operatorname{int}\operatorname{dom} f$. By reasoning by induction, we conclude that

$$(x_n)_{n\in\mathbb{N}} \in (\operatorname{int}\operatorname{dom} f)^{\mathbb{N}} \quad \text{is well-defined.} \tag{6.77}$$

Next, let us set $\Phi = \varphi + \psi \circ L$ and

$$(\forall n \in \mathbb{N}) \quad g_n \colon \mathcal{X} \to \left]-\infty,+\infty\right] \\ x \mapsto \begin{cases} f_n(x) - \gamma_n\psi(Lx), & \text{if } x \in \operatorname{int}\operatorname{dom} f;\\ +\infty, & \text{otherwise.}\end{cases} \tag{6.78}$$

Since $L(\operatorname{int}\operatorname{dom} f) \subset \operatorname{int}\operatorname{dom}\psi$, it follows from (6.78) that $(\forall n \in \mathbb{N})$ g_n is Gâteaux differentiable on $\operatorname{dom} g_n = \operatorname{int}\operatorname{dom} g_n = \operatorname{int}\operatorname{dom} f$. Since ψ is continuous on $\operatorname{int}\operatorname{dom}\psi \supset L(\operatorname{int}\operatorname{dom} f)$ and the functions $(f_n)_{n\in\mathbb{N}}$ are continuous on $\operatorname{int}\operatorname{dom} f$ [27, Proposition 3.3], we deduce that $(\forall n \in \mathbb{N})$ g_n is continuous on $\operatorname{dom} g_n$. In addition,

$$(\forall n \in \mathbb{N}) \quad g_n - \varepsilon\alpha f = (1-\varepsilon)(f_n - \alpha\beta\psi\circ L) + \varepsilon(f_n - \alpha f) + \big(\alpha\beta(1-\varepsilon) - \gamma_n\big)\psi\circ L. \tag{6.79}$$

Note that $f \succcurlyeq \beta\psi\circ L$ and $(\forall n \in \mathbb{N})$ $f_n \succcurlyeq \alpha f$. Hence, (6.79) yields

$$(\forall n \in \mathbb{N}) \quad f_n \succcurlyeq \alpha\beta\psi\circ L, \tag{6.80}$$

and hence, we deduce from (6.72) and (6.79) that $(\forall n \in \mathbb{N})$ $g_n \succcurlyeq \varepsilon\alpha f$. In turn,

$$(\forall n \in \mathbb{N})(\forall x \in \operatorname{dom} g_n)(\forall y \in \operatorname{dom} g_n) \quad \big\langle x - y, \nabla g_n(x) - \nabla g_n(y)\big\rangle \\ = D^{g_n}(x,y) + D^{g_n}(y,x) \geqslant \varepsilon\alpha\big(D^f(x,y) + D^f(y,x)\big) \geqslant 0, \quad (6.81)$$

and it therefore follows from [29, Theorem 2.1.11] that $(\forall n \in \mathbb{N})$ g_n is convex. Consequently,

$$(\forall n \in \mathbb{N}) \quad g_n \in \mathcal{P}_{\varepsilon\alpha}(f). \quad (6.82)$$

Set $\omega = 1 + 1/\varepsilon$. Then

$$(\forall n \in \mathbb{N}) \quad (1+\omega\eta_n)g_n - g_{n+1} = (1+\omega\eta_n)(f_n - \gamma_n\psi \circ L) - (f_{n+1} - \gamma_{n+1}\psi \circ L) \\ = (1+\eta_n)f_n - f_{n+1} + \eta_n\varepsilon^{-1}\big(f_n - (\gamma_n + \varepsilon\alpha\beta)\psi \circ L\big) \\ + \big(\alpha\beta\eta_n + \gamma_{n+1} - (1+\eta_n)\gamma_n\big)\psi \circ L. \quad (6.83)$$

We thus derive from (6.72) and (6.80) that

$$(\forall n \in \mathbb{N}) \quad (1+\omega\eta_n)g_n \succcurlyeq g_{n+1}. \quad (6.84)$$

By invoking (6.73) and Proposition 6.19(ii)(a), we get

$$(\forall n \in \mathbb{N}) \quad \nabla f_n(x_n) - \gamma_n L^*\nabla\psi(Lx_n) \in \nabla f_n(x_{n+1}) + \gamma_n\partial\varphi(x_{n+1}), \quad (6.85)$$

and therefore,

$$(\forall n \in \mathbb{N}) \quad \nabla f_n(x_n) - \gamma_n L^*\nabla\psi(Lx_n) \in \nabla f_n(x_{n+1}) - \gamma_n L^*\nabla\psi(Lx_{n+1}) \\ + \gamma_n\big(\partial\varphi(x_{n+1}) + L^*\nabla\psi(Lx_{n+1})\big). \quad (6.86)$$

Since [29, Theorem 2.4.2(vii)–(viii)] yield

$$(\forall n \in \mathbb{N}) \quad \partial\varphi(x_{n+1}) + L^*\nabla\psi(Lx_{n+1}) \subset \partial\varphi(x_{n+1}) + L^*\big(\partial\psi(Lx_{n+1})\big) \\ \subset \partial(\varphi + \psi \circ L)(x_{n+1}) = \partial\Phi(x_{n+1}), \quad (6.87)$$

we deduce from (6.86) that

$$(\forall n \in \mathbb{N}) \quad \nabla g_n(x_n) - \nabla g_n(x_{n+1}) \in \gamma_n\partial\Phi(x_{n+1}). \quad (6.88)$$

By appealing to (6.26) and (6.88), we get

$$(\forall x \in \operatorname{dom}\Phi \cap \operatorname{dom} f)(\forall n \in \mathbb{N}) \quad \gamma_n^{-1}\big\langle x - x_{n+1}, \nabla g_n(x_n) - \nabla g_n(x_{n+1})\big\rangle + \Phi(x_{n+1}) \leqslant \Phi(x), \quad (6.89)$$

and hence, by [6, Proposition 2.3(ii)],

$$(\forall x \in \operatorname{dom}\Phi \cap \operatorname{dom} f)(\forall n \in \mathbb{N}) \quad \gamma_n^{-1}\big(D^{g_n}(x, x_{n+1}) + D^{g_n}(x_{n+1}, x_n) - D^{g_n}(x, x_n)\big) \\ + \Phi(x_{n+1}) \leqslant \Phi(x). \quad (6.90)$$

In particular,

$$(\forall x \in \mathcal{S} \cap \operatorname{dom} f)(\forall n \in \mathbb{N}) \quad D^{g_n}(x, x_{n+1}) + D^{g_n}(x_{n+1}, x_n) - D^{g_n}(x, x_n) \leqslant 0. \tag{6.91}$$

By using (6.84), we deduce from (6.91) that

$$(\forall x \in \mathcal{S} \cap \operatorname{dom} f)(\forall n \in \mathbb{N}) \quad D^{g_{n+1}}(x, x_{n+1}) + (1 + \omega\eta_n) D^{g_n}(x_{n+1}, x_n) \\ \leqslant (1 + \omega\eta_n) D^{g_n}(x, x_n), \tag{6.92}$$

and therefore,

$$(\forall x \in \mathcal{S} \cap \operatorname{dom} f)(\forall n \in \mathbb{N}) \quad D^{g_{n+1}}(x, x_{n+1}) \leqslant (1 + \omega\eta_n) D^{g_n}(x, x_n). \tag{6.93}$$

This shows that $(x_n)_{n\in\mathbb{N}}$ is stationarily quasi-Bregman monotone with respect to $\mathcal{S}$ relative to $(g_n)_{n\in\mathbb{N}}$. Hence, we deduce from Proposition 6.16(ii) that

$$(x_n)_{n\in\mathbb{N}} \in (\operatorname{int}\operatorname{dom} f)^{\mathbb{N}} \quad \text{is bounded} \tag{6.94}$$

and, since $\mathcal{X}$ is reflexive,

$$\mathfrak{W}(x_n)_{n\in\mathbb{N}} \neq \varnothing. \tag{6.95}$$

In addition, we derive from (6.93) and Proposition 6.16(i) that

$$(\forall x \in \mathcal{S} \cap \operatorname{int}\operatorname{dom} f) \quad \big(D^{g_n}(x, x_n)\big)_{n\in\mathbb{N}} \quad \text{converges,} \tag{6.96}$$

and thus, since (6.92) yields

$$\begin{aligned}(\forall x \in \mathcal{S} \cap \operatorname{int}\operatorname{dom} f)(\forall n \in \mathbb{N}) \quad 0 &\leqslant D^{g_n}(x_{n+1}, x_n) \\ &\leqslant (1 + \omega\eta_n) D^{f_n}(x_{n+1}, x_n) \\ &\leqslant (1 + \omega\eta_n) D^{f_n}(x, x_n) - D^{f_{n+1}}(x, x_{n+1}),\end{aligned} \tag{6.97}$$

we obtain

$$D^{g_n}(x_{n+1}, x_n) \to 0. \tag{6.98}$$

On the other hand, it follows from (6.82) that

$$(\forall n \in \mathbb{N}) \quad \varepsilon\alpha D^f(x_{n+1}, x_n) \leqslant D^{g_n}(x_{n+1}, x_n), \tag{6.99}$$

and hence, (6.98) yields

$$D^f(x_{n+1}, x_n) \to 0. \tag{6.100}$$

Now, it follows from (6.90) that

$$(\forall n \in \mathbb{N}) \quad \Phi(x_{n+1}) \leqslant \gamma_n^{-1}\big(D^{g_n}(x_n, x_{n+1}) + D^{g_n}(x_{n+1}, x_n)\big) + \Phi(x_{n+1}) \leqslant \Phi(x_n), \tag{6.101}$$

which shows that $(\Phi(x_n))_{n\in\mathbb{N}}$ is decreasing and hence, since it is bounded from below by $\inf \Phi(\mathcal{X})$, it is convergent. However, (6.90) and (6.93) yield

$$
\begin{aligned}
&(\forall x \in \mathcal{S} \cap \operatorname{int}\operatorname{dom} f)(\forall n \in \mathbb{N})\\
&\varepsilon^{-1}\left(\frac{1}{1+\omega\eta_n} D^{g_{n+1}}(x, x_{n+1}) + D^{g_n}(x_{n+1}, x_n) - D^{g_n}(x, x_n)\right) + \Phi(x_{n+1})\\
&\leqslant \gamma_n^{-1}\left(\frac{1}{1+\omega\eta_n} D^{g_{n+1}}(x, x_{n+1}) + D^{g_n}(x_{n+1}, x_n) - D^{g_n}(x, x_n)\right) + \Phi(x_{n+1})\\
&\leqslant \Phi(x).
\end{aligned} \tag{6.102}
$$

Since $\eta_n \to 0$, by taking the limit in (6.102) and then using (6.96) and (6.98), we get

$$
\inf \Phi(\mathcal{X}) \leqslant \lim \Phi(x_n) \leqslant \inf \Phi(\mathcal{X}), \tag{6.103}
$$

and thus,

$$
\Phi(x_n) \to \inf \Phi(\mathcal{X}). \tag{6.104}
$$

We now show that

$$
\mathfrak{W}(x_n)_{n\in\mathbb{N}} \subset \mathcal{S}. \tag{6.105}
$$

To this end, suppose that $x \in \mathfrak{W}(x_n)_{n\in\mathbb{N}}$, i.e., $x_{k_n} \rightharpoonup x$. Since Φ is weakly lower semicontinuous [29, Theorem 2.2.1], by (6.104),

$$
\inf \Phi(\mathcal{X}) \leqslant \Phi(x) \leqslant \underline{\lim}\, \Phi(x_{k_n}) = \lim \Phi(x_n) = \inf \Phi(\mathcal{X}). \tag{6.106}
$$

This yields $\Phi(x) = \inf \Phi(\mathcal{X})$, i.e., $x \in \operatorname{Argmin} \Phi = \mathcal{S}$.

(i) : Let $\overline{x} \in \mathfrak{W}(x_n)_{n\in\mathbb{N}}$. Since (6.94) and (6.105) imply that $\mathfrak{W}(x_n)_{n\in\mathbb{N}} \subset \mathcal{S} \cap \overline{\operatorname{dom}} f$, we obtain $\mathfrak{W}(x_n)_{n\in\mathbb{N}} = \{\overline{x}\}$, and in turn, (6.95) yields $x_n \rightharpoonup \overline{x}$.

(ii) : In view of (6.105) and Proposition 6.17, it suffices to show that $\mathfrak{W}(x_n)_{n\in\mathbb{N}} \subset \operatorname{int}\operatorname{dom} f$.

(ii)(a) : We have $\mathfrak{W}(x_n)_{n\in\mathbb{N}} \subset \mathcal{S} \subset \operatorname{int}\operatorname{dom} f$.

(ii)(b) : This follows from Lemma 6.25.

(iii) : Let $\overline{x} \in \mathcal{S} \cap \operatorname{int}\operatorname{dom} f$. Since f satisfies Condition 6.15, (6.100) yields

$$
x_{n+1} - x_n \to 0. \tag{6.107}
$$

Now set

$$
(\forall n \in \mathbb{N}) \quad z_n = x_{n+1} \quad \text{and} \quad z_n^* = \gamma_n^{-1}\big(\nabla g_n(x_n) - \nabla g_n(z_n)\big). \tag{6.108}
$$

Then (6.88) and (6.107) imply that

$$(\forall n \in \mathbb{N}) \quad z_n^* \in \partial\Phi(z_n) \quad \text{and} \quad z_n - x_n \to 0. \tag{6.109}$$

Since (6.92) yields

$$\begin{aligned}(\forall n \in \mathbb{N}) \quad D^{g_{n+1}}(\overline{x}, x_{n+1}) &= D^{g_{n+1}}(\overline{x}, z_n)\\ &\leqslant (1+\omega\eta_n)D^{g_n}(\overline{x}, z_n)\\ &= (1+\omega\eta_n)D^{g_n}(\overline{x}, x_{n+1})\\ &\leqslant (1+\omega\eta_n)D^{g_n}(\overline{x}, x_n),\end{aligned} \tag{6.110}$$

we deduce that

$$(\forall n \in \mathbb{N}) \quad (1+\omega\eta_n)^{-1}D^{g_{n+1}}(\overline{x}, x_{n+1}) \leqslant D^{g_n}(\overline{x}, z_n) \leqslant D^{g_n}(\overline{x}, x_n). \tag{6.111}$$

Altogether, (6.96) and (6.111) yield

$$D^{g_n}(\overline{x}, z_n) - D^{g_n}(\overline{x}, x_n) \to 0. \tag{6.112}$$

In (6.89), by setting $x = \overline{x}$, we get

$$\begin{aligned}(\forall n \in \mathbb{N}) \quad 0 &\leqslant \gamma_n \langle z_n - \overline{x}, z_n^*\rangle\\ &= \big\langle z_n - \overline{x}, \nabla g_n(x_n) - \nabla g_n(z_n)\big\rangle\\ &= D^{g_n}(\overline{x}, x_n) - D^{g_n}(\overline{x}, z_n) - D^{g_n}(z_n, x_n)\\ &\leqslant D^{g_n}(\overline{x}, x_n) - D^{g_n}(\overline{x}, z_n).\end{aligned} \tag{6.113}$$

By taking to the limit in (6.113) and using (6.112), we get

$$\langle z_n - \overline{x}, z_n^*\rangle \to 0. \tag{6.114}$$

(iii)(a) : In this case $\mathcal{S} = \{\overline{x}\}$. Since φ is uniformly convex at $\overline{x}$, Φ is likewise and hence, there exists an increasing function $\phi\colon\ [0,+\infty[\ \to [0,+\infty]$ that vanishes only at 0 such that

$$(\forall n \in \mathbb{N})(\forall \tau \in\ ]0,1[) \quad \Phi(\tau\overline{x}+(1-\tau)z_n)+\tau(1-\tau)\phi(\|z_n-\overline{x}\|) \leqslant \tau\Phi(\overline{x})+(1-\tau)\Phi(z_n). \tag{6.115}$$

It therefore follows from [29, Page 201] that $\partial\Phi$ is uniformly monotone at $\overline{x}$ and its modulus of convexity is ϕ, i.e,

$$(\forall n \in \mathbb{N}) \quad \langle z_n - \overline{x}, z_n^*\rangle \geqslant \phi(\|z_n - \overline{x}\|) \geqslant 0. \tag{6.116}$$

Altogether, (6.114) and (6.116) yield $\phi(\|z_n-\overline{x}\|) \to 0$, and thus, $z_n \to \overline{x}$. In turn, (6.109) yields $x_n \to \overline{x}$.

(iii)(b) : By the same argument as in (iii)(a), $\mathcal{S} = \{\overline{x}\}$ and there exists an increasing function $\phi\colon\ [0,+\infty[\to [0,+\infty]$ that vanishes only at 0 such that

$$(\forall n \in \mathbb{N}) \quad \big\langle z_n - \overline{x}, \nabla\psi(Lz_n) - \nabla\psi(L\overline{x})\big\rangle \geqslant \phi(\|Lz_n - L\overline{x}\|). \tag{6.117}$$

In turn, it follows from (6.87) and [29, Theorem 2.4.2(iv)] that

$$(\forall n \in \mathbb{N}) \quad \langle z_n - \overline{x}, z_n^*\rangle \geqslant \phi(\|Lz_n - L\overline{x}\|). \tag{6.118}$$

This yields $\phi(\|Lz_n - L\overline{x}\|) \to 0$, and hence, $Lz_n \to L\overline{x}$. Since

$$(\forall n \in \mathbb{N}) \quad \|Lz_n - L\overline{x}\| \geqslant \kappa\|z_n - \overline{x}\|, \tag{6.119}$$

we obtain $z_n \to \overline{x}$ and in turn, (6.109) yields $x_n \to \overline{x}$.

(iii)(c) : First, we observe that $\mathcal{S}$ is closed and convex since $\Phi \in \Gamma_0(\mathcal{X})$. Next, for every $n \in \mathbb{N}$, since $\mu f \succcurlyeq f_n$, we derive from (6.78) that $\mu f \succcurlyeq g_n$. Finally, the strong convergence follows from Proposition 6.18. □

The following corollary of Theorem 6.26 appears to be the first version of the forward-backward algorithm outside of Hilbert spaces.

Theorem 6.27 *Consider the setting of Problem 6.11 and let $f \in \Gamma_0(\mathcal{X})$ be a Legendre function such that $\mathcal{S} \cap \operatorname{int}\operatorname{dom} f \neq \varnothing$, $L(\operatorname{int}\operatorname{dom} f) \subset \operatorname{int}\operatorname{dom}\psi$, and $f \succcurlyeq \beta\psi \circ L$ for some $\beta \in\,]0,+\infty[$. Suppose that either f is cofinite or $-L^*(\operatorname{ran}\nabla\psi) \subset \operatorname{dom}\varphi^*$, and that $(\forall x \in \operatorname{int}\operatorname{dom} f)$ $D^f(x,\cdot)$ is coercive. Let $\varepsilon \in\,]0, \beta/(\beta+1)[$, let $(\eta_n)_{n\in\mathbb{N}} \in \ell_+^1(\mathbb{N})$, and let $(\gamma_n)_{n\in\mathbb{N}}$ be a sequence in $\mathbb{R}$ such that*

$$(\forall n \in \mathbb{N}) \quad \varepsilon \leqslant \gamma_n \leqslant \beta(1-\varepsilon) \quad \textit{and} \quad (1+\eta_n)\gamma_n - \gamma_{n+1} \leqslant \beta\eta_n. \tag{6.120}$$

Furthermore, let $x_0 \in \operatorname{int}\operatorname{dom} f$ and iterate

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = \operatorname{Prox}^f_{\gamma_n\varphi}\big(\nabla f(x_n) - \gamma_n L^*\nabla\psi(Lx_n)\big). \tag{6.121}$$

Then there exists $\overline{x} \in \mathcal{S}$ such that the following hold.

(i) *Suppose that one of the following holds.*

 (a) *$\mathcal{S} \cap \overline{\operatorname{dom}} f$ is a singleton.*

 (b) *∇f and $\nabla\psi$ are weakly sequentially continuous and $\mathcal{S} \subset \operatorname{int}\operatorname{dom} f$.*

 (c) *$\operatorname{dom} f^*$ is open and ∇f, ∇f^*, and $\nabla\psi$ are weakly sequentially continuous.*

 Then $x_n \rightharpoonup \overline{x}$.

(ii) *Suppose that f satisfies Condition 6.15 and that one of the following holds.*

 (a) *φ is uniformly convex at $\overline{x}$.*

 (b) *ψ is uniformly convex at $L\overline{x}$ and there exists $\kappa \in\,]0,+\infty[$ such that $(\forall x \in \mathcal{X})$ $\|Lx\| \geqslant \kappa\|x\|$.*

(c) $\underline{\lim}\, D_{\mathcal{S}}^{f}(x_n) = 0$.

Then $x_n \to \overline{x}$.

Proof. Set $(\forall n \in \mathbb{N})\ f_n = f$. Then

$$(\forall n \in \mathbb{N}) \quad \begin{cases} f_n \in \mathcal{P}_1(f), \\ f \succcurlyeq f_n, \\ (1+\eta_n) f_n \succcurlyeq f_{n+1}. \end{cases} \tag{6.122}$$

(i)(a) : This is a corollary of Theorem 6.26(i).

(i)(b)–(i)(c) : Firstly, the proof of Theorem 6.26(ii)(a)–(ii)(b) shows that $\mathfrak{W}(x_n)_{n\in\mathbb{N}} \subset \operatorname{int} \operatorname{dom} f$. Next, in view of Theorem 6.26(ii), it suffices to show that (6.74) holds. To this end, suppose that y_1 and y_2 are two weak sequential cluster points of $(x_n)_{n\in\mathbb{N}}$ such that

$$\big(\langle y_1 - y_2, \nabla f(x_n) - \gamma_n L^* \nabla\psi(Lx_n)\rangle\big)_{n\in\mathbb{N}} \quad \text{converges.} \tag{6.123}$$

Then, there exist two strictly increasing sequences $(k_n)_{n\in\mathbb{N}}$ and $(l_n)_{n\in\mathbb{N}}$ in $\mathbb{N}$ such that $x_{k_n} \rightharpoonup y_1$ and $x_{l_n} \rightharpoonup y_2$. We derive from (6.120) and [28, Lemma 2.2.2] that there exists $\theta \in [\varepsilon, \beta(1-\varepsilon)]$ such that $\gamma_n \to \theta$. Since ∇f and $\nabla\psi$ are weakly sequentially continuous, after taking the limit in (6.123) along the subsequences $(x_{k_n})_{n\in\mathbb{N}}$ and $(x_{l_n})_{n\in\mathbb{N}}$, respectively, we get

$$\langle y_1 - y_2, \nabla f(y_1) - \theta L^* \nabla\psi(Ly_1)\rangle = \langle y_1 - y_2, \nabla f(y_2) - \theta L^* \nabla\psi(Ly_2)\rangle. \tag{6.124}$$

Let us define

$$\begin{aligned} h\colon \mathcal{X} &\to \left]-\infty, +\infty\right] \\ x &\mapsto \begin{cases} f(x) - \theta\psi(Lx), & \text{if } x \in \operatorname{int}\operatorname{dom} f; \\ +\infty, & \text{otherwise.} \end{cases} \end{aligned} \tag{6.125}$$

Then h is Gâteaux differentiable on $\operatorname{int}\operatorname{dom} h = \operatorname{int}\operatorname{dom} f$ and (6.124) yields

$$\langle y_1 - y_2, \nabla h(y_1) - \nabla h(y_2)\rangle = 0. \tag{6.126}$$

On the other hand,

$$h - \varepsilon f = f - \theta\psi\circ L - \varepsilon f = (1-\varepsilon)(f - \beta\psi\circ L) + \big(\beta(1-\varepsilon) - \theta\big)\psi\circ L. \tag{6.127}$$

In turn, since $f \succcurlyeq \beta\psi\circ L$ and $\theta \leqslant \beta(1-\varepsilon)$, we obtain $h \succcurlyeq \varepsilon f$, and hence,

$$D^h(y_1, y_2) \geqslant \varepsilon D^f(y_1, y_2) \quad \text{and} \quad D^h(y_2, y_1) \geqslant \varepsilon D^f(y_2, y_1). \tag{6.128}$$

Therefore, (6.126) yields

$$
\begin{aligned}
0 &= \big\langle y_1 - y_2, \nabla h(y_1) - \nabla h(y_2)\big\rangle \\
&= D^h(y_1, y_2) + D^h(y_2, y_1) \\
&\geqslant \varepsilon\big(D^f(y_1, y_2) + D^f(y_2, y_1)\big) \\
&= \varepsilon\big\langle y_1 - y_2, \nabla f(y_1) - \nabla f(y_2)\big\rangle.
\end{aligned} \tag{6.129}
$$

Suppose that $y_1 \neq y_2$. Since $f|_{\operatorname{int\,dom} f}$ is strictly convex, ∇f is strictly monotone [29, Theorem 2.4.4(ii)], i.e.,

$$
\big\langle y_1 - y_2, \nabla f(y_1) - \nabla f(y_2)\big\rangle > 0 \tag{6.130}
$$

and we reach a contradiction.

(ii) : The conclusions follow from (6.122) and Theorem 6.26(iii). □

Remark 6.28 In Problem 6.11, suppose that $L = \operatorname{Id}$. We rewrite algorithm (6.121) as follow

$$
(\forall n \in \mathbb{N}) \quad x_{n+1} = \underset{x\in\mathcal{X}}{\operatorname{argmin}} \ \Big(\varphi(x) + \big\langle x - x_n, \nabla\psi(x_n)\big\rangle + \psi(x_n) + \gamma_n^{-1} D^f(x, x_n)\Big). \tag{6.131}
$$

Another method to solve Problem 6.11 was proposed in [11]. In that method, instead of solving (6.131), the authors solve

$$
(\forall n \in \mathbb{N}) \quad x_{n+1} = \underset{x\in\mathcal{X}}{\operatorname{argmin}} \ \Big(\varphi(x) + \big\langle x - x_n, \nabla\psi(x_n)\big\rangle + \psi(x_n) + \gamma_n^{-1}\|x - x_n\|^p\Big), \tag{6.132}
$$

for some $1 < p \leqslant 2$. The weak convergence is established under the assumptions that Problem 6.11 admits a unique solution, $\nabla\psi$ is $(p-1)$-Hölder continuous with constant β, and $0 < \inf_{n\in\mathbb{N}} \gamma_n \leqslant \sup_{n\in\mathbb{N}} \gamma_n \leqslant (1-\delta)/\beta$, where $0 < \delta < 1$. The high nonlinearity of the regularization in (6.132) compared to (6.131) makes the numerical implementation of this method difficult in general. Furthermore, since (6.132) yields

$$
(\forall n \in \mathbb{N}) \quad 0 \in \partial\varphi(x_{n+1}) + \nabla\psi(x_n) + \gamma_n^{-1}\partial\big(\|x_{n+1} - x_n\|^p\big), \tag{6.133}
$$

and since $(\forall n \in \mathbb{N})$ $\partial\big(\|x_{n+1} - x_n\|^p\big)$ is not separable, this method is not a splitting method.

Remark 6.29 We can reformulate Problem 6.11 as the following joint minimization problem

$$
\underset{(x,y)\in V}{\operatorname{minimize}} \ \varphi(x) + \psi(y), \tag{6.134}
$$

where $V = \operatorname{gra} L = \big\{(x, y) \in \mathcal{X} \times \mathcal{Y} \ \big|\ y = Lx\big\}$. This constrained problem is equivalent to the following unconstrained problem

$$
\underset{(x,y)\in\mathcal{X}\times\mathcal{Y}}{\operatorname{minimize}} \ \varphi(x) + \psi(y) + \iota_V(x, y). \tag{6.135}
$$

In [9], a different coupling term between the variables x and y was considered and the problem considered there was

$$\underset{(x,y)\in\mathcal{X}\times\mathcal{Y}}{\text{minimize}}\;\varphi(x)+\psi(y)+D^f(x,y),\tag{6.136}$$

in the Euclidean spaces. Their method activates φ and ψ via their so-called left and right Bregman proximity operators alternatively (see also [7] for the projection setting). This method does not require the smoothness of ψ but it requires the computation of Bregman distance-based proximity operator of ψ.

Next, we provide a particular instance of Theorem 6.27 in finite-dimensional spaces.

Corollary 6.30 *In the setting of Problem 6.11, suppose that $\mathcal{X}$ and $\mathcal{Y}$ are finite-dimensional. Let $f\in\Gamma_0(\mathcal{X})$ be a Legendre function such that $\mathcal{S}\cap\operatorname{int}\operatorname{dom}f\neq\varnothing$, $L(\operatorname{int}\operatorname{dom}f)\subset\operatorname{int}\operatorname{dom}\psi$, $f\succcurlyeq\beta\psi\circ L$ for some $\beta\in\left]0,+\infty\right[$, and $\operatorname{dom}f^*$ is open. Suppose that either f is cofinite or $-L^*(\operatorname{ran}\nabla\psi)\subset\operatorname{dom}\varphi^*$. Let $\varepsilon\in\left]0,\beta/(\beta+1)\right[$, let $(\eta_n)_{n\in\mathbb{N}}\in\ell_+^1(\mathbb{N})$, and let $(\gamma_n)_{n\in\mathbb{N}}$ be a sequence in $\mathbb{R}$ such that*

$$(\forall n\in\mathbb{N})\quad \varepsilon\leqslant\gamma_n\leqslant\beta(1-\varepsilon)\quad\textit{and}\quad(1+\eta_n)\gamma_n-\gamma_{n+1}\leqslant\beta\eta_n.\tag{6.137}$$

Furthermore, let $x_0\in\operatorname{int}\operatorname{dom}f$ and iterate

$$(\forall n\in\mathbb{N})\quad x_{n+1}=\operatorname{Prox}^f_{\gamma_n\varphi}\big(\nabla f(x_n)-\gamma_nL^*\nabla\psi(Lx_n)\big).\tag{6.138}$$

Then there exists $\overline{x}\in\mathcal{S}$ such that $x_n\to\overline{x}$.

Proof. Since $\operatorname{dom}f^*$ is open, [5, Lemma 7.3(ix)] asserts that $(\forall x\in\operatorname{int}\operatorname{dom}f)$ $D^f(x,\cdot)$ is coercive. Hence, the claim follows from Theorem 6.27(i)(c). □

Remark 6.31 We provide some special cases of Problem 6.11 and Theorem 6.27.

(i) Let I and K be totally ordered countable index sets. In Problem 6.11, suppose that $\mathcal{X}$ and $\mathcal{Y}$ are separable Hilbert spaces, and that $\psi\colon y\mapsto\sum_{k\in K}|\langle y-r,y_k\rangle|^2/2$, where $r\in\mathcal{Y}$ and $(y_k)_{k\in K}$ is a frame in $\mathcal{Y}$, i.e.,

$$(\exists(\mu,\nu)\in\left]0,+\infty\right[^2)(\forall y\in\mathcal{Y})\quad \mu\|y\|^2\leqslant\sum_{k\in K}|\langle y,y_k\rangle|^2\leqslant\nu\|y\|^2.\tag{6.139}$$

Then in Theorem 6.27, we can choose $f\colon x\mapsto\sum_{i\in I}|\langle x,x_i\rangle|^2/2$, where $(x_i)_{i\in I}$ is a frame in $\mathcal{X}$, i.e.,

$$(\exists(\alpha,\beta)\in\left]0,+\infty\right[^2)(\forall x\in\mathcal{X})\quad \alpha\|x\|^2\leqslant\sum_{i\in I}|\langle x,x_i\rangle|^2\leqslant\beta\|x\|^2.\tag{6.140}$$

It follows from [17, Corollary 1] that f and ψ are Legendre functions and that ∇f and $\nabla\psi$ are weakly sequentially continuous. Now let x and z be in $\mathcal{X}$. Then

$$\begin{aligned} D^{\psi}(Lx,Lz) &= \sum_{k\in K} |\langle Lx-Lz, y_k\rangle|^2/2 \leqslant \nu\|Lx-Lz\|^2/2 \\ &\leqslant \nu\|L\|^2\|x-z\|^2/2 \leqslant \nu|L\|^2\alpha^{-1}\sum_{i\in I}|\langle x-z,x_i\rangle|^2/2 \\ &= \nu\|L\|^2\alpha^{-1}D^f(x,z), \end{aligned} \tag{6.141}$$

which implies that $f \succcurlyeq \alpha\nu^{-1}\|L\|^{-2}\psi\circ L$ and in addition, $D^f(x,\cdot)$ is coercive.

(ii) Let p and q be in $]1,+\infty[$ and set $p^* = p/(p-1)$ and $q^* = q(p-1)$. In Problem 6.11, suppose that $\mathcal{X} = \ell^p(\mathbb{N})$ and $\mathcal{Y} = \ell^q(\mathbb{N})$, that $r \in \ell^q(\mathbb{N})$, that $\psi\colon y \mapsto \|y\|^q/q - \langle y-r, \|r\|^{q-2}r\rangle - \|r\|^q/q$, and that there exists $\kappa \in\,]0,+\infty[$ such that $(\forall x\in\ell^p(\mathbb{N}))$ $\|Lx\| \geqslant \kappa\|x\|$. It follows from [16, Theorem 4.7] that $\ell^p(\mathbb{N})$ is uniformly convex and hence, strictly convex. Therefore, ψ is strictly convex and supercoercive. The property of L implies that $\psi\circ L$ is strictly convex and supercoercive, and $\varphi+\psi\circ L$ is likewise. In turn, Problem 6.11 admits a unique solution. Let $f\in\Gamma_0(\ell^p(\mathbb{N}))$ be a cofinite Legendre function such that $f \succcurlyeq \beta\|L\cdot\|^p$ for some $\beta\in\,]0,+\infty[$, let $x\in\ell^p(\mathbb{N})$, and set

$$\phi\colon [0,+\infty[\to[0,+\infty]\colon t\mapsto \inf_{\|z-x\|=t}\left(\|x\|^p - p\left\langle x-z, \|z\|^{p-2}z\right\rangle - \|z\|^p\right). \tag{6.142}$$

Then

$$(\forall z\in\ell^p(\mathbb{N}))\quad D^f(x,z) \geqslant p\beta D^{\psi}(Lx,Lz) \geqslant \beta\phi(\|Lz-Lx\|). \tag{6.143}$$

If $p\in[2,+\infty[$ then we derive from [13, Lemma 1.4.10] that

$$(\forall z\in\ell^p(\mathbb{N}))\quad \phi(\|Lz-Lx\|) \geqslant 2^{1-p}\|Lz-Lx\|^p \geqslant 2^{1-p}\kappa^p\|z-x\|^p. \tag{6.144}$$

Thus, it follows from (6.143) that

$$(\forall z\in\ell^p(\mathbb{N}))\quad D^f(x,z) \geqslant 2^{1-p}\kappa^p\beta\|z-x\|^p, \tag{6.145}$$

and $D^f(x,\cdot)$ is therefore coercive. If $p\in\,]1,2[$ then we derive from [13, Lemma 1.4.8] that

$$(\forall z\in\ell^p(\mathbb{N}))\quad \phi(\|Lz-Lx\|) \geqslant (\|Lz-Lx\|+\|Lz\|)^p - \|Lz\|^p - p\|Lz\|^{p-1}\|Lz-Lx\|, \tag{6.146}$$

and hence, (6.143) yields

$$(\forall z\in\ell^p(\mathbb{N}))\ D^f(x,z) \geqslant \beta\left((\|Lz\|+\|Lz-Lx\|)^p - \|Lz\|^p - p\|Lz\|^{p-1}\|Lz-Lx\|\right). \tag{6.147}$$

On the other hand, since

$$\lim_{\|Lz\|\to+\infty} \frac{(\|Lz\| + \|Lz - Lx\|)^p - \|Lz\|^p - p\|Lz\|^{p-1}\|Lz - Lx\|}{(2^p - 1 - p)\|Lz\|^p} = 1, \tag{6.148}$$

and since $2^p - 1 - p > 0$, it follows from and (6.147) and the property of L that

$$\lim_{\|z\|\to+\infty} D^f(x, z) = +\infty, \tag{6.149}$$

and $D^f(x, \cdot)$ is therefore coercive. Consequently, Theorem 6.27(i)(a) can be applied.

6.2.4 Application to multivariate minimization

We propose a variant of the forward-backward algorithm to solve the following multivariate minimization problem.

Problem 6.32 Let m and p be strictly positive integers, let $(\mathcal{X}_i)_{1\leqslant i\leqslant m}$ and $(\mathcal{Y}_k)_{1\leqslant k\leqslant p}$ be reflexive real Banach spaces. For every $i \in \{1, \dots, m\}$ and every $k \in \{1, \dots, p\}$, let $\varphi_i \in \Gamma_0(\mathcal{X}_i)$, let $\psi_k \in \Gamma_0(\mathcal{Y}_k)$ be Gâteaux differentiable on $\operatorname{int}\operatorname{dom}\psi_k \neq \varnothing$, and let $L_{ik} \colon \mathcal{X}_i \to \mathcal{Y}_k$ be linear and bounded. The problem is to

$$\underset{x_1\in\mathcal{X}_1,\dots,x_m\in\mathcal{X}_m}{\text{minimize}} \quad \sum_{i=1}^m \varphi_i(x_i) + \sum_{k=1}^p \psi_k\left(\sum_{i=1}^m L_{ik}x_i\right). \tag{6.150}$$

Denote by $\mathcal{S}$ the set of solutions to (6.150).

We derive from Theorem 6.27 the following result.

Proposition 6.33 *Consider the setting of Problem 6.32. For every $k \in \{1, \dots, p\}$, suppose that there exists $\sigma_k \in \left]0, +\infty\right[$ such that for every $(y_{ik})_{1\leqslant i\leqslant m} \in \operatorname{int}\operatorname{dom}\psi_k$ and every $(v_{ik})_{1\leqslant i\leqslant m} \in \operatorname{int}\operatorname{dom}\psi_k$ satisfying $\sum_{i=1}^m y_{ik} \in \operatorname{int}\operatorname{dom}\psi_k$ and $\sum_{i=1}^m v_{ik} \in \operatorname{int}\operatorname{dom}\psi_k$, one has*

$$D^{\psi_k}\left(\sum_{i=1}^m y_{ik}, \sum_{i=1}^m v_{ik}\right) \leqslant \sigma_k \sum_{i=1}^m D^{\psi_k}(y_{ik}, v_{ik}). \tag{6.151}$$

For every $i \in \{1, \dots, m\}$, let $f_i \in \Gamma_0(\mathcal{X}_i)$ be a Legendre function such that $(\forall x_i \in \operatorname{int}\operatorname{dom} f_i)$ $D^{f_i}(x_i, \cdot)$ is coercive. For every $k \in \{1, \dots, p\}$, suppose that $\sum_{i=1}^m L_{ik}(\operatorname{int}\operatorname{dom} f_i) \subset \operatorname{int}\operatorname{dom}\psi_k$, that, for every $i \in \{1, \dots, m\}$, there exists $\beta_{ik} \in \left]0, +\infty\right[$ such that $f_i \succcurlyeq \beta_{ik}\psi_k \circ L_{ik}$, and set $\beta_k = \min_{1\leqslant i\leqslant m} \beta_{ik}$. In addition, suppose that $\mathcal{S} \cap \bigtimes_{i=1}^m \operatorname{int}\operatorname{dom} f_i \neq \varnothing$ and that either $(\forall i \in \{1, \dots, m\})$ f_i is cofinite or $(\forall i \in \{1, \dots, m\})$ φ_i is cofinite. Let

$\varepsilon \in \left]0, 1/\left(1+\sum_{k=1}^p \sigma_k\beta_k^{-1}\right)\right[$, *let* $(\eta_n)_{n\in\mathbb{N}} \in \ell_+^1(\mathbb{N})$, *and let* $(\gamma_n)_{n\in\mathbb{N}}$ *be a sequence in* $\mathbb{R}$ *such that*

$$(\forall n \in \mathbb{N}) \quad \varepsilon \leqslant \gamma_n \leqslant \frac{1-\varepsilon}{\sum_{k=1}^p \sigma_k\beta_k^{-1}} \quad \text{and} \quad (1+\eta_n)\gamma_n - \gamma_{n+1} \leqslant \frac{\eta_n}{\sum_{k=1}^p \sigma_k\beta_k^{-1}}. \tag{6.152}$$

Furthermore, let $(x_{i,0})_{1\leqslant i\leqslant m} \in \bigtimes_{i=1}^m \operatorname{int}\operatorname{dom} f_i$ *and iterate*

$$\begin{array}{l} \text{for } n = 0, 1, \ldots \\ \left\lfloor \begin{array}{l} \text{for } i = 1, \ldots, m \\ \left\lfloor \; x_{i,n+1} = \operatorname{Prox}_{\gamma_n\varphi_i}^{f_i}\left(\nabla f_i(x_{i,n}) - \gamma_n \sum_{k=1}^p L_{ik}^*\nabla\psi_k\left(\sum_{j=1}^m L_{jk}x_{j,n}\right)\right). \right. \end{array} \right. \end{array} \tag{6.153}$$

Then there exists $(\overline{x}_i)_{1\leqslant i\leqslant m} \in \mathcal{S}$ *such that the following hold.*

(i) *Suppose that* $\mathcal{S} \cap \bigtimes_{i=1}^m \overline{\operatorname{dom} f_i}$ *is a singleton. Then* $(\forall i \in \{1,\ldots,m\})$ $x_{i,n} \rightharpoonup \overline{x}_i$.

(ii) *For every* $i \in \{1,\ldots,m\}$ *and every* $k \in \{1,\ldots,p\}$, *suppose that* ∇f_i *and* $\nabla\psi_k$ *are weakly sequentially continuous, and that one of the following holds.*

  (a) $\operatorname{dom}\varphi_i \subset \operatorname{int}\operatorname{dom} f_i$.

  (b) $\operatorname{dom} f_i^*$ *is open and* ∇f_i^* *is weakly sequentially continuous.*

  Then $(\forall i \in \{1,\ldots,m\})$ $x_{i,n} \rightharpoonup \overline{x}_i$.

Proof. Denote by $\mathcal{X}$ and $\mathcal{Y}$ the standard vector product spaces $\bigtimes_{i=1}^m \mathcal{X}_i$ and $\bigtimes_{k=1}^p \mathcal{Y}_k$ equipped with the norms $x = (x_i)_{1\leqslant i\leqslant m} \mapsto \sqrt{\sum_{i=1}^m \|x_i\|^2}$ and $y = (y_k)_{1\leqslant k\leqslant p} \mapsto \sqrt{\sum_{k=1}^p \|y_k\|^2}$, respectively. Then $\mathcal{X}^*$ is the vector product space $\bigtimes_{i=1}^m \mathcal{X}_i^*$ equipped with the norm $x^* \mapsto \sqrt{\sum_{i=1}^m \|x_i^*\|^2}$ and $\mathcal{Y}^*$ is the vector product space $\bigtimes_{k=1}^p \mathcal{Y}_k^*$ equipped with the norm $y^* \mapsto \sqrt{\sum_{k=1}^p \|y_k^*\|^2}$. Let us introduce the functions and operator

$$\begin{cases} \varphi\colon \mathcal{X} \to \left]-\infty,+\infty\right] : x \mapsto \sum_{i=1}^m \varphi_i(x_i) \\ f\colon \mathcal{X} \to \left]-\infty,+\infty\right] : x \mapsto \sum_{i=1}^m f_i(x_i) \\ \psi\colon \mathcal{Y} \mapsto \left]-\infty,+\infty\right] : y \mapsto \sum_{k=1}^p \psi_k(y_k) \\ L\colon \mathcal{X} \to \mathcal{Y}\colon x \mapsto \left(\sum_{i=1}^m L_{ik}x_i\right)_{1\leqslant k\leqslant p}. \end{cases} \tag{6.154}$$

Then ψ is Gâteaux differentiable on $\operatorname{int}\operatorname{dom}\psi = \bigtimes_{k=1}^p \operatorname{int}\operatorname{dom}\psi_k$ and Problem 6.32 is a special case of Problem 6.11. Since (6.154) yields $\operatorname{dom} f^* = \bigtimes_{i=1}^m \operatorname{dom} f_i^*$ and $\operatorname{dom}\varphi^* = \bigtimes_{i=1}^m \operatorname{dom}\varphi_i^*$, we deduce from our assumptions that either f is cofinite or φ is cofinite. As in (6.48) and (6.49), f is a Legendre function and $\operatorname{dom}\varphi \cap \operatorname{int}\operatorname{dom} f \neq \varnothing$. In addition,

$$L\big(\operatorname{int}\operatorname{dom} f\big) = \bigtimes_{k=1}^p \sum_{i=1}^m L_{ki}(\operatorname{int}\operatorname{dom} f_i) \subset \bigtimes_{k=1}^p \operatorname{int}\operatorname{dom}\psi_k = \operatorname{int}\operatorname{dom}\psi. \tag{6.155}$$

Now let $x \in \operatorname{int}\operatorname{dom} f$. First, to show that $D^f(x,\cdot)$ is coercive, we fix $\rho \in \mathbb{R}$. On the one hand,

$$\big\{z = (z_i)_{1\leqslant i\leqslant m} \in \mathcal{X} \;\big|\; D^f(x,z) \leqslant \rho\big\} \subset \bigtimes_{i=1}^{m} \big\{z_i \in \mathcal{X}_i \;\big|\; D^{f_i}(x_i,z_i) \leqslant \rho\big\}. \tag{6.156}$$

On the other hand, for every $i \in \{1,\ldots,m\}$, since $D^{f_i}(x_i,\cdot)$ is coercive, we deduce that

$$\big\{z_i \in \mathcal{X}_i \;\big|\; D^{f_i}(x_i,z_i) \leqslant \rho\big\} \quad \text{is bounded.} \tag{6.157}$$

Hence (6.156) implies that $\big\{z \in \mathcal{X} \;\big|\; D^f(x,z) \leqslant \rho\big\}$ is bounded and $D^f(x,\cdot)$ is therefore coercive. Next, set $\beta = 1/\sum_{k=1}^{p} \sigma_k \beta_k^{-1}$. We shall show that $f \succcurlyeq \beta\psi \circ L$. To this end, fix $z = (z_i)_{1\leqslant i\leqslant m} \in \operatorname{int}\operatorname{dom} f$. We have

$$\begin{aligned}
D^\psi(Lx, Lz) &= \sum_{k=1}^{p} D^{\psi_k}\left(\sum_{i=1}^{m} L_{ik}x_i, \sum_{i=1}^{m} L_{ik}z_i\right) \\
&\leqslant \sum_{k=1}^{p}\sum_{i=1}^{m} \sigma_k D^{\psi_k}(L_{ik}x_i, L_{ik}z_i) \\
&\leqslant \sum_{k=1}^{p}\sum_{i=1}^{m} \sigma_k \beta_{ik}^{-1} D^{f_i}(x_i, z_i) \\
&\leqslant \sum_{k=1}^{p} \sigma_k \beta_k^{-1} D^f(x,z).
\end{aligned} \tag{6.158}$$

Now let us set $(\forall n \in \mathbb{N})\ x_n = (x_{i,n})_{1\leqslant i\leqslant m}$. By virtue of Proposition 6.20, (6.153) is a particular case of (6.121).

(i) : Since $\mathcal{S} \cap \overline{\operatorname{dom}} f$ is a singleton, the claim follows from Theorem 6.27(i)(a).

(ii) : Our assumptions on $(f_i)_{1\leqslant i\leqslant m}$ and $(\psi_k)_{1\leqslant k\leqslant p}$ imply that ∇f and $\nabla\psi$ are weakly sequentially continuous.

(ii)(a) : Since $\mathcal{S} \subset \bigtimes_{i=1}^{m} \operatorname{dom}\varphi_i \subset \bigtimes_{i=1}^{m} \operatorname{int}\operatorname{dom} f_i = \operatorname{int}\operatorname{dom} f$, the claim follows from Theorem 6.27(i)(b).

(ii)(b) : Since, for every $i \in \{1,\ldots,m\}$, $\operatorname{dom} f_i^*$ is open and ∇f_i^* is weakly sequentially continuous, we deduce that $\operatorname{dom} f^*$ is open and ∇f^* is weakly sequentially continuous. The assertion therefore follows from Theorem 6.27(i)(c). □

Example 6.34 In Problem 6.32, suppose that $m = 1$, that $\mathcal{X}_1$ and $(\mathcal{Y}_k)_{1\leqslant k\leqslant p}$ are Hilbert spaces, and that, for every $k \in \{1,\ldots,p\}$, $\varphi_k = \omega_k\|\cdot - r_k\|^2/2$, where $(\omega_k)_{1\leqslant k\leqslant p} \in \left]0,+\infty\right[^p$ and let $(r_k)_{1\leqslant k\leqslant p} \in \bigtimes_{k=1}^{p} \mathcal{Y}_k$. Then the weak convergence result in [18, Proposition 6.3] without errors is a particular instance of Proposition 6.33 with $f_1 = \|\cdot\|^2/2$.

Example 6.35 Let m and p be strictly positive integers. For every $i \in \{1,\ldots,m\}$ and every $k \in \{1,\ldots,p\}$, let $\omega_{ik} \in \left]0,+\infty\right[$, let $\varrho_k \in \left]0,+\infty\right[$, and let $\varphi_i \in \Gamma_0(\mathbb{R})$ be cofinite. The problem is to

$$\underset{(\xi_1,\ldots,\xi_m)\in\left]0,+\infty\right[^m}{\text{minimize}} \quad \sum_{i=1}^{m}\varphi_i(\xi_i) + \sum_{k=1}^{p}\left(-\ln\frac{\sum_{i=1}^m \omega_{ik}\xi_i}{\varrho_k} + \frac{\sum_{i=1}^m \omega_{ik}\xi_i}{\varrho_k} - 1\right). \tag{6.159}$$

Denote by $\mathcal{S}$ the set of solutions to (6.159) and suppose that $\mathcal{S}\cap\left]0,+\infty\right[^m \neq \varnothing$. Let

$$\vartheta\colon \mathbb{R}\to\left]-\infty,+\infty\right] : \xi \mapsto \begin{cases} -\ln\xi, & \text{if} \quad \xi>0;\\ +\infty, & \text{otherwise}\end{cases} \tag{6.160}$$

be Burg entropy, let $\varepsilon \in \left]0,1/(1+p)\right[$, let $(\eta_n)_{n\in\mathbb{N}} \in \ell_+^1(\mathbb{N})$, and let $(\gamma_n)_{n\in\mathbb{N}}$ be a sequence in $\mathbb{R}$ such that

$$(\forall n\in\mathbb{N}) \quad \varepsilon \leqslant \gamma_n \leqslant p^{-1}(1-\varepsilon) \quad \text{and} \quad (1+\eta_n)\gamma_n - \gamma_{n+1} \leqslant p^{-1}\eta_n. \tag{6.161}$$

Let $(\xi_{i,0})_{1\leqslant i\leqslant m} \in \left]0,+\infty\right[^m$ and iterate

$$\begin{array}{l}\text{for } n=0,1,\ldots\\ \left\lfloor \begin{array}{l}\text{for } i=1,\ldots,m\\ \left\lfloor \xi_{i,n+1} = \operatorname{Prox}^{\vartheta}_{\gamma_n\varphi_i}\left(\dfrac{-1}{\xi_{i,n}} - \gamma_n\displaystyle\sum_{k=1}^{p}\omega_{ik}\left(\frac{-1}{\sum_{j=1}^m \omega_{jk}\xi_{j,n}} + \frac{1}{\varrho_k}\right)\right).\right.\end{array}\right.\end{array} \tag{6.162}$$

Then there exists $(\overline{\xi}_i)_{1\leqslant i\leqslant m} \in \mathcal{S}$ such that $(\forall i\in\{1,\ldots,m\})$ $\xi_{i,n}\to\overline{\xi}_i$.

Proof. For every $i \in \{1,\ldots,m\}$ and every $k\in\{1,\ldots,p\}$, let us set $\mathcal{X}_i = \mathbb{R}$, $\mathcal{Y}_k = \mathbb{R}$, $\psi_k = D^{\vartheta}(\cdot,\varrho_k)$, and $L_{ik}\colon \xi_i \mapsto \omega_{ik}\xi_i$. Then (6.159) is a particular case of (6.150). Since ψ is not differentiable on $\mathbb{R}^p$, the standard forward-backward algorithm is inapplicable. We show that the problem can be solved by using Proposition 6.33. First, let $(\xi_i)_{1\leqslant i\leqslant m}$ and $(\eta_i)_{1\leqslant i\leqslant m}$ be in $\left]0,+\infty\right[^m$, and consider

$$\phi\colon \mathbb{R}\to\left]-\infty,+\infty\right] : \xi\mapsto \begin{cases} -\ln\xi+\xi-1, & \text{if} \quad \xi\in\left]0,+\infty\right[;\\ +\infty, & \text{otherwise.}\end{cases} \tag{6.163}$$

We see that ϕ is convex and positive. Thus,

$$\phi\left(\frac{\sum_{i=1}^m\xi_i}{\sum_{i=1}^m\eta_i}\right) = \phi\left(\sum_{i=1}^m \frac{\eta_i}{\sum_{j=1}^m\eta_j}\frac{\xi_i}{\eta_i}\right) \leqslant \sum_{i=1}^m\frac{\eta_i}{\sum_{j=1}^m\eta_j}\phi\left(\frac{\xi_i}{\eta_i}\right) \leqslant \sum_{i=1}^m\phi\left(\frac{\xi_i}{\eta_i}\right), \tag{6.164}$$

and hence,

$$-\ln\frac{\sum_{i=1}^m\xi_i}{\sum_{i=1}^m\eta_i} + \frac{\sum_{i=1}^m\xi_i}{\sum_{i=1}^m\eta_i} - 1 \leqslant \sum_{i=1}^m\left(-\ln\frac{\xi_i}{\eta_i}+\frac{\xi_i}{\eta_i}-1\right). \tag{6.165}$$

In turn,

$$D^{\vartheta}\left(\sum_{i=1}^{m}\xi_i,\sum_{i=1}^{m}\eta_i\right)\leqslant\sum_{i=1}^{m}D^{\vartheta}(\xi_i,\eta_i). \tag{6.166}$$

This shows that (6.151) is satisfied with $(\forall k\in\{1,\ldots,p\})\ \sigma_k=1$. Next, let us set $(\forall i\in\{1,\ldots,m\})\ f_i=\vartheta$. Fix $i\in\{1,\ldots,m\}$ and $k\in\{1,\ldots,p\}$, and let ξ_i and η_i be in $]0,+\infty[$. Then

$$D^{\psi_k}(L_{ik}\xi_i,L_{ik}\eta_i)=D^{\vartheta}(\omega_{ik}\xi_i,\omega_{ik}\eta_i)=D^{\vartheta}(\xi_i,\eta_i)=D^{f_i}(\xi_i,\eta_i), \tag{6.167}$$

which implies that $f_i\succcurlyeq\psi_k\circ L_{ik}$. In addition, since $\operatorname{dom}f_i^*=\,]-\infty,0[$ is open, [5, Lemma 7.3(ix)] asserts that $D^{f_i}(\xi_i,\cdot)$ is coercive. We therefore deduce the convergence result from Proposition 6.33(ii)(b). ☐

Example 6.36 Let m and p be strictly positive integers. For every $i\in\{1,\ldots,m\}$ and every $k\in\{1,\ldots,p\}$, let $\omega_{ik}\in\,]0,+\infty[$, let $\varrho_k\in\,]0,+\infty[$, and let $\varphi_i\in\Gamma_0(\mathbb{R})$. The problem is to

$$\underset{(\xi_1,\ldots,\xi_m)\in[0,+\infty[^m}{\text{minimize}}\quad\sum_{i=1}^{m}\varphi_i(\xi_i)+\sum_{k=1}^{p}\left(\left(\sum_{i=1}^{m}\omega_{ik}\xi_i\right)\ln\frac{\sum_{i=1}^{m}\omega_{ik}\xi_i}{\varrho_k}-\sum_{i=1}^{m}\omega_{ik}\xi_i+\varrho_k\right). \tag{6.168}$$

Denote by $\mathcal{S}$ the set of solutions to (6.168) and suppose that $\mathcal{S}\cap\,]0,+\infty[^m\neq\varnothing$. Let

$$\vartheta\colon\mathbb{R}\to\,]-\infty,+\infty]:\xi\mapsto\begin{cases}\xi\ln\xi-\xi, & \text{if }\xi\in\,]0,+\infty[;\\ 0, & \text{if }\xi=0;\\ +\infty, & \text{otherwise}\end{cases} \tag{6.169}$$

be Boltzmann-Shannon entropy, let $\beta=\max\limits_{1\leqslant k\leqslant p}\max\limits_{1\leqslant i\leqslant m}\omega_{ik}$, let $\varepsilon\in\,]0,1/(1+\beta)[$, let $(\eta_n)_{n\in\mathbb{N}}\in\ell_+^1(\mathbb{N})$, and let $(\gamma_n)_{n\in\mathbb{N}}$ be a sequence in $\mathbb{R}$ such that

$$(\forall n\in\mathbb{N})\quad\varepsilon\leqslant\gamma_n\leqslant(p\beta)^{-1}(1-\varepsilon)\quad\text{and}\quad(1+\eta_n)\gamma_n-\gamma_{n+1}\leqslant(p\beta)^{-1}\eta_n. \tag{6.170}$$

Let $(\xi_{i,0})_{1\leqslant i\leqslant m}\in\,]0,+\infty[^m$ and iterate

$$\begin{array}{l}\text{for } n=0,1,\ldots\\ \left\lfloor\begin{array}{l}\text{for } i=1,\ldots,m\\ \left\lfloor\ \xi_{i,n+1}=\operatorname{Prox}^{\vartheta}_{\gamma_n\varphi_i}\Big(\ln\xi_{i,n}-\gamma_n\sum\limits_{k=1}^{p}\omega_{ik}\Big(\ln\Big(\sum_{j=1}^{m}\omega_{jk}\xi_{j,n}\Big)-\ln\varrho_k\Big)\Big).\right.\end{array}\right.\end{array} \tag{6.171}$$

Then there exists $(\overline{\xi}_i)_{1\leqslant i\leqslant m}\in\mathcal{S}$ such that $(\forall i\in\{1,\ldots,m\})\ \xi_{i,n}\to\overline{\xi}_i$.

Proof. For every $i\in\{1,\ldots,m\}$ and every $k\in\{1,\ldots,p\}$, let us set $\mathcal{X}_i=\mathbb{R}$, $\mathcal{Y}_k=\mathbb{R}$, $\psi_k=D^{\vartheta}(\cdot,\varrho_k)$, and $L_{ik}\colon\xi_i\mapsto\omega_{ik}\xi_i$. Then (6.168) is a particular case of (6.150). We cannot apply the standard forward-backward algorithm here since ψ is not differentiable on $\mathbb{R}^p$. We shall verify the assumptions of Proposition 6.33. First, let $(\xi_i)_{1\leqslant i\leqslant m}$ and $(\eta_i)_{1\leqslant i\leqslant m}$ be in $]0,+\infty[^m$. Since

$$\phi\colon \mathbb{R}\to\,]-\infty,+\infty]:\xi\mapsto\begin{cases}\xi\ln\xi, & \text{if}\quad \xi\in\,]0,+\infty[\,;\\ 0, & \text{if}\quad \xi=0;\\ +\infty, & \text{otherwise}\end{cases}\tag{6.172}$$

is convex, we have

$$\phi\left(\frac{\sum_{i=1}^m\xi_i}{\sum_{i=1}^m\eta_i}\right)=\phi\left(\sum_{i=1}^m\frac{\eta_i}{\sum_{j=1}^m\eta_j}\frac{\xi_i}{\eta_i}\right)\leqslant\sum_{i=1}^m\frac{\eta_i}{\sum_{j=1}^m\eta_j}\phi\left(\frac{\xi_i}{\eta_i}\right),\tag{6.173}$$

and hence,

$$\frac{\sum_{i=1}^m\xi_i}{\sum_{i=1}^m\eta_i}\ln\frac{\sum_{i=1}^m\xi_i}{\sum_{i=1}^m\eta_i}\leqslant\sum_{i=1}^m\frac{\eta_i}{\sum_{j=1}^m\eta_j}\frac{\xi_i}{\eta_i}\ln\frac{\xi_i}{\eta_i}=\frac{\sum_{i=1}^m\xi_i\ln\dfrac{\xi_i}{\eta_i}}{\sum_{i=1}^m\eta_i}.\tag{6.174}$$

In turn,

$$\left(\sum_{i=1}^m\xi_i\right)\ln\frac{\sum_{i=1}^m\xi_i}{\sum_{i=1}^m\eta_i}\leqslant\sum_{i=1}^m\xi_i\ln\frac{\xi_i}{\eta_i},\tag{6.175}$$

which implies that

$$\begin{aligned}D^\vartheta\left(\sum_{i=1}^m\xi_i,\sum_{i=1}^m\eta_i\right)&=\left(\sum_{i=1}^m\xi_i\right)\ln\frac{\sum_{i=1}^m\xi_i}{\sum_{i=1}^m\eta_i}-\sum_{i=1}^m\xi_i+\sum_{i=1}^m\eta_i\\ &\leqslant\sum_{i=1}^m\left(\xi_i\ln\frac{\xi_i}{\eta_i}-\xi_i+\eta_i\right)\\ &=\sum_{i=1}^m D^\vartheta(\xi_i,\eta_i).\end{aligned}\tag{6.176}$$

This shows that (6.151) is satisfied with $(\forall k\in\{1,\ldots,p\})$ $\sigma_k=1$. Next, let us set $(\forall i\in\{1,\ldots,m\})$ $f_i=\vartheta$. Fix $i\in\{1,\ldots,m\}$ and $k\in\{1,\ldots,p\}$, and let ξ_i and η_i be in $]0,+\infty[$. Then

$$D^{\psi_k}(L_{ik}\xi_i,L_{ik}\eta_i)=D^\vartheta(\omega_{ik}\xi_i,\omega_{ik}\eta_i)=\omega_{ik}D^\vartheta(\xi_i,\eta_i)\leqslant\beta D^\vartheta(\xi_i,\eta_i),\tag{6.177}$$

which implies that $f_i\succcurlyeq\beta^{-1}\psi_k\circ L_{ik}$. In addition, since f_i is supercoercive, f_i is cofinite and [5, Lemma 7.3(viii)] asserts that $D^{f_i}(\xi_i,\cdot)$ is coercive. Therefore, the claim follows from Proposition 6.33(ii)(b). □

Remark 6.37 The Bregman distance associated with Burg entropy, i.e., the Itakura-Saito divergence, is used in linear regression [3, Section 3]. The Bregman distance associated with Boltzmann-Shannon entropy, i.e., the Kullback-Leibler divergence, is used in information theory [3, Section 3] and image processing [14].

Acknowledgment. I would like to thank my doctoral advisor Professor Patrick L. Combettes for bringing this problem to my attention and for helpful discussions.

6.3 Résultats numériques

Dans cette section, nous illustrons une application de l'algorithme (6.121) en restauration d'images. Pour chaque simulation, l'algorithme est exécuté avec $n = 5000$ itérations afin de trouver une solution approchée fiable x_{5000}. Ensuite, chaque algorithme est relancé avec le test d'arrêt relatif $\|x_n - x_{5000}\|/\|x_n - x_0\| \leqslant 10^{-5}$ et $D^f(x_{5000}, x_n)/D^f(x_{5000}, x_0) \leqslant 10^{-5}$ où f est une fonction indiquée dans chaque simulation.

Soient $N \in \mathbb{N}^*$ et $(\alpha, \beta) \in \left]0, +\infty\right[^2$. Pour toute image $x \in \mathbb{R}^{N\times N}$, $\mathscr{P}_\alpha(x) \in \mathbb{R}^{N\times N}$ est l'image bruitée à partir de x par un bruit de Poisson de paramètre α, et $\mathscr{G}_{\alpha,\beta}(x) \in \mathbb{R}^{N\times N}$ est l'image bruitée à partir de x par un bruit Γ-distribué de paramètres α et β. Nous considérons les modèles de dégradation suivants

$$r = \mathscr{P}_\alpha(L\overline{x}) \tag{6.178}$$

et

$$r = \mathscr{G}_{\alpha,\beta}(L\overline{x}), \tag{6.179}$$

où $\overline{x} \in \mathbb{R}^{N\times N}$ est l'image originale dégradée par un opérateur de convolution L et contaminée par un bruit de Poisson (le modèle (6.178)) ou un bruit Γ-distribué (le modèle (6.179)). Le vecteur r représente l'image dégradée. Le modèle que nous utilisons est [24]

$$\underset{x\in\mathbb{R}^{N\times N}}{\text{minimiser}} \;\; \tau\|x\|^2 + D^g(Lx, r), \tag{6.180}$$

où $\tau \in \left]0, +\infty\right[$, g est l'entropie de Boltzmann-Shannon définie sur $\mathbb{R}^{N\times N}$, et L est un opérateur de convolution défini par $L\colon x \mapsto h * x$ où h est un noyau uniforme de taille $m \times m$. Notons que $D^g(L\cdot, r)$ n'est pas différentiable partout et donc on ne peut pas appliquer l'algorithme explicite-implicite classique [19].

6.3.1 Bruit de Poisson

Le modèle (6.180) apparaît dans le problème de restauration d'images avec des bruits de Poisson, par exemple, [10]. Nous utilisons un bruit de Poisson de paramètre $\alpha = 0,95$. Le niveau de bruit est donné par le rapport image floutée-sur-bruit

$$20 \log_{10} \frac{\|L\overline{x}\|}{\|\overline{y} - L\overline{x}\|} = 25,88 \text{ dB}. \tag{6.181}$$

Nous utilisons l'algorithme (6.121) en choisissant pour f l'entropie de Boltzmann-Shannon pour le modèle (6.180) avec $N = 256$, $m = 13$, et $\tau = 1/1000$. Nous obtenons les résultats des figures 6.3, 6.4, et 6.5.

FIGURE 6.1 – Image originale.

FIGURE 6.2 – Image dégradée.

FIGURE 6.3 – Image restaurée.

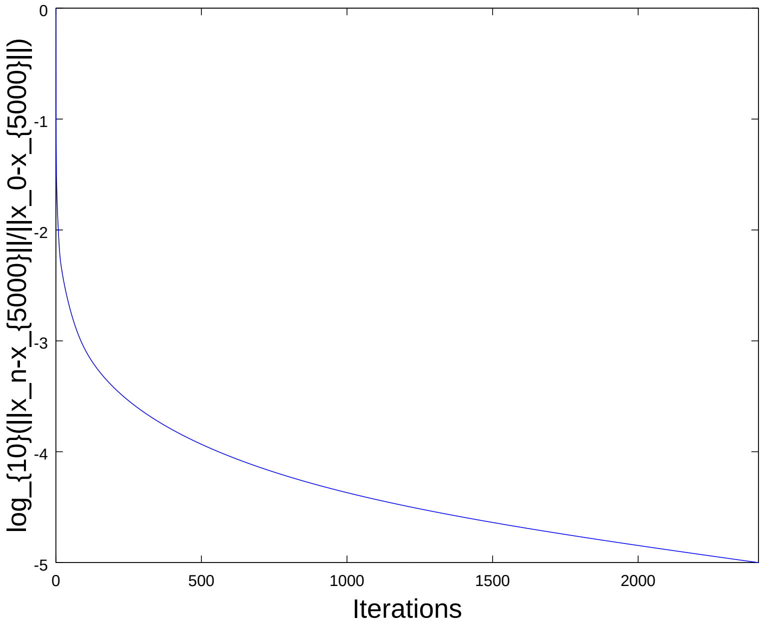

FIGURE 6.4 – Distance euclidienne à la limite normalisée $\|x_n - x_\infty\|/\|x_0 - x_\infty\|$.

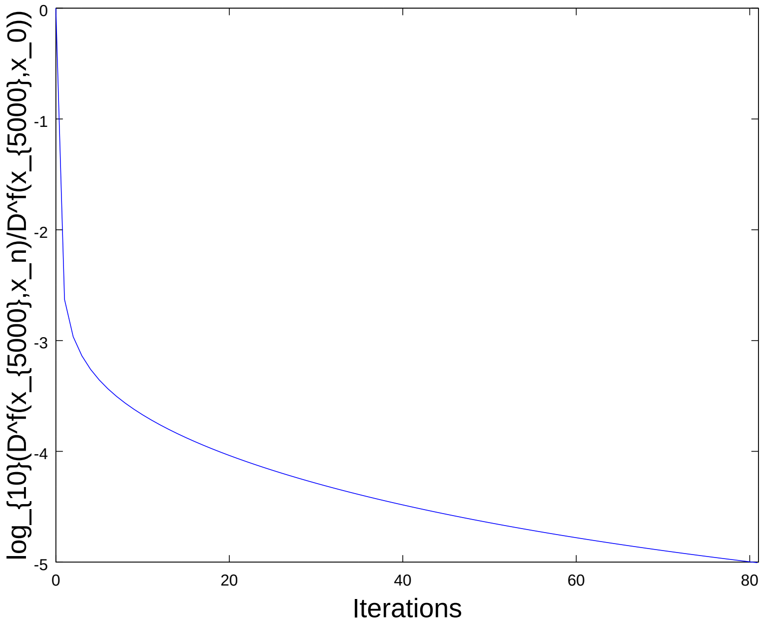

FIGURE 6.5 – Distance de Bregman à la limite normalisée $D^f(x_\infty, x_n)/D^f(x_\infty, x_0)$.

6.3.2 Bruit Γ-distribué

Le modèle (6.180) apparaît aussi dans le problème de restauration d'images avec des bruits Γ-distribués, par exemple, [28]. Nous utilisons un bruit Γ-distribué de paramètres $\alpha = \beta = 0,95$. Le niveau est donné par le rapport image floutée-sur-bruit :

$$20 \log_{10} \frac{\|L\overline{x}\|}{\|\overline{y} - L\overline{x}\|} = 25,8 \text{ dB}. \tag{6.182}$$

Nous utilisons l'algorithme (6.121) en choisissant pour f l'entropie de Boltzmann-Shannon pour le modèle (6.180) avec $N = 256$, $m = 13$, et $\tau = 1/1000$. Nous obtenons les résultats des figures 6.7, 6.8, et 6.9.

FIGURE 6.6 – Image dégradée.

FIGURE 6.7 – Image restaurée.

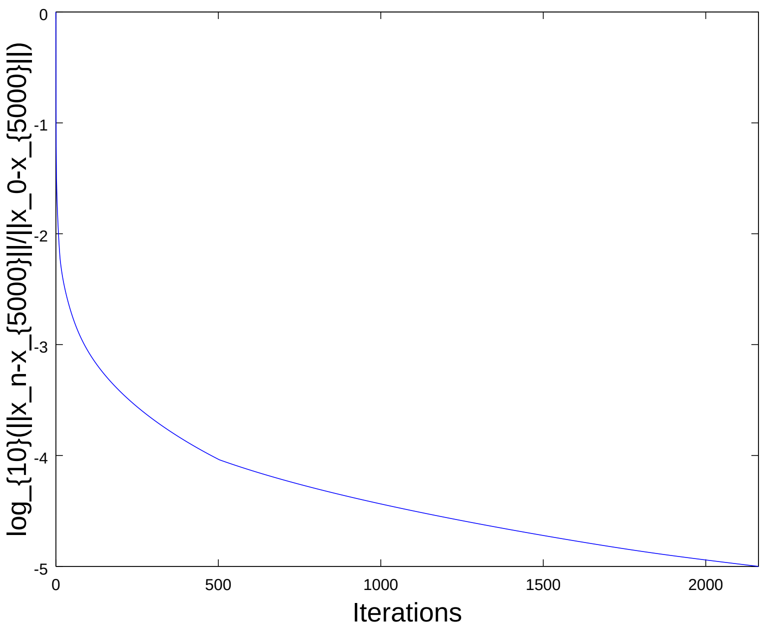

FIGURE 6.8 – Distance euclidienne à la limite normalisée $\|x_n - x_\infty\|/\|x_0 - x_\infty\|$.

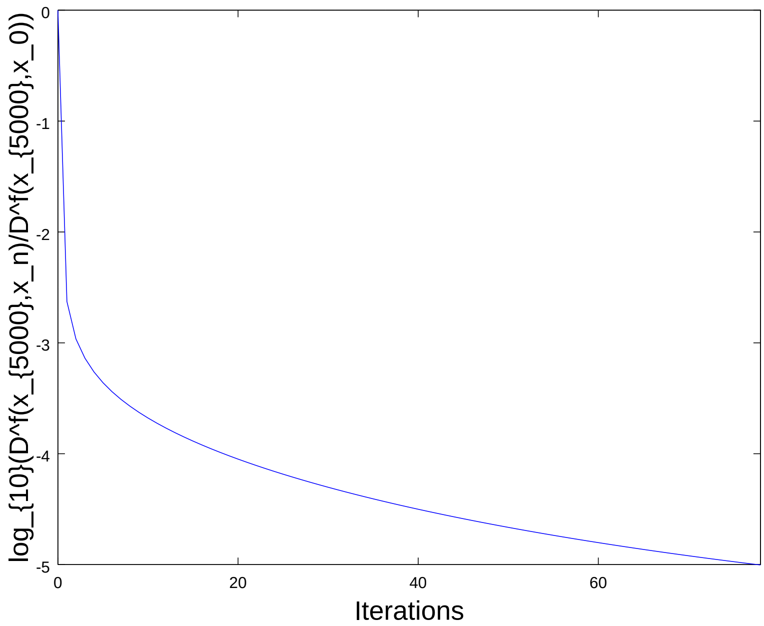

FIGURE 6.9 – Distance de Bregman à la limite normalisée $D^f(x_\infty, x_n)/D^f(x_\infty, x_0)$.

6.4 Bibliographie

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

[2] A. Banerjee, S. Basu, and S. Merugu, Multi-way clustering on relation graphs, in : *Proceedings of the 7th SIAM International Conference on Data Mining*, (C. Apte , D. Skillicorn, B. Liu, and S. Parthasarathy, eds.), pp. 145–156. SIAM, Philadelphia, 2007.

[3] M. Basseville, Divergence measures for statistical data processing : An annotated bibliography, *Signal Process.*, vol. 93, pp. 621–633, 2013.

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

[5] H. H. Bauschke, J. M. Borwein, and P. L. Combettes, Essential smoothness, essential strict convexity, and Legendre functions in Banach spaces, *Commun. Contemp. Math.*, vol. 3, pp. 615–647, 2001.

[6] H. H. Bauschke, J. M. Borwein, and P. L. Combettes, Bregman monotone optimization algorithms, *SIAM J. Control Optim.*, vol. 42, pp. 596–636, 2003.

[7] H. H. Bauschke and P. L. Combettes, Iterating Bregman retractions, *SIAM J. Optim.*, vol. 13, pp. 1159–1173, 2003.

[8] H. H. Bauschke and P. L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces.* Springer, New York, 2011.

[9] H. H. Bauschke, P. L. Combettes, and D. Noll, Joint minimization with alternating Bregman proximity operators, *Pac. J. Optim.,* vol. 2, pp. 401–424, 2006.

[10] M. Bertero, P. Boccacci, G. Desiderà, and G. Vicidomini, Image deblurring with Poisson data : from cells to galaxies, *Inverse Problems,* vol. 25, Article ID 123006, 26 pp., 2009.

[11] K. Bredies, A forward-backward splitting algorithm for the minimization of non-smooth convex functionals in Banach space, *Inverse Problems,* vol 25, Article ID 015005, 20 pp., 2009.

[12] L. M. Briceño-Arias and P. L. Combettes, A monotone+skew splitting model for composite monotone inclusions in duality, *SIAM J. Optim.*, vol. 21, pp. 1230–1250, 2011.

[13] D. Butnariu and A. N. Iusem, *Totally Convex Functions for Fixed Points Computation and Infinite Dimensional Optimization.* Kluwer Academic Publishers, Dordrecht, 2000.

[14] C. L. Byrne, Iterative image reconstruction algorithms based on cross-entropy minimization, *IEEE Trans. Image Process.,* vol. 2, pp. 96–103, 1993.

[15] V. Chandrasekaran, P. A. Parrilo, and A. S. Willsky, Latent variable graphical model selection via convex optimization, *Ann. Statist.,* vol. 40, pp. 1935–1967, 2012.

[16] I. Cioranescu, *Geometry of Banach spaces, Duality Mappings and Nonlinear Problems.* Kluwer, Dordrecht, 1990.

[17] P. L. Combettes and N. N. Reyes, Moreau's decomposition in Banach spaces, *Math. Program.,* vol. 139, pp. 103–114, 2013.

[18] P. L. Combettes and B. C. Vũ, Variable metric quasi-Fejér monotonicity, *Nonlinear Anal.,* vol. 78, pp. 17–31, 2013.

[19] P. L. Combettes and V. R. Wajs, Signal recovery by proximal forward-backward splitting, *Multiscale Model. Simul.,* vol. 4, pp. 1168–1200, 2005.

[20] R. M. Corless, G. H. Gonnet, D. E. G. Hare, D. J. Jeffrey, and D. E. Knuth, On the Lambert W function, *Adv. Comput. Math.,* vol. 5, pp. 329–359, 1996.

[21] I. S. Dhillon and J. A. Tropp, Matrix nearness problems with Bregman divergences, *SIAM J. Matrix Anal. Appl.,* vol. 29, pp. 1120–1146, 2007.

[22] J. Kivinen and M. K. Warmuth, Relative loss bounds for multidimensional regression problems, *Mach. Learn.,* vol. 45, pp. 301–329, 2001.

[23] H. Lantéri, M. Roche, and C. Aime, Penalized maximum likelihood image restoration with positivity constraints : multiplicative algorithms, *Inverse Problems,* vol. 18, pp. 1397–1419, 2002.

[24] J. Markham and J. A. Conchello, Fast maximal-likelihood image-restoration algorithms for three dimensional fluorescence microscopy, *J. Opt. Soc. Am. A,* vol. 18, pp. 1062–1071, 2001.

[25] J. J. Moreau, Fonctions convexes duales et points proximaux dans un espace hilbertien, *C. R. Acad. Sci. Paris Sér. A Math.,* vol. 255, pp. 2897–2899, 1962.

[26] Q. V. Nguyen, Variable quasi-Bregman monotone sequences, 2015.
http://arxiv.org/abs/1505.04460

[27] R. R. Phelps, *Convex Functions, Monotone Operators and Differentiability, 2nd ed.,* Lecture Notes in Math. 1364, Springer-Verlag, Berlin, 1993.

[28] B. T. Polyak, *Introduction to Optimization.* Optimization Software Inc., New York, 1987.

[29] C. Zălinescu, *Convex Analysis in General Vector Spaces.* World Scientific Publishing, River Edge, NJ, 2002.

Chapitre 7

Contractivité et cocoercivité relatives aux distances de Bregman

Nous étudions la notion de contraction basée sur les distances de Bregman dans des espaces de Banach réels et contruisons des algorithmes pour trouver des points fixes de ces contractions. Nous étudions le lien entre la notion de cocoercivité classique et celle basée sur les distances de Bregman. La convergence d'une version de l'algorithme explicite-implicite pour résoudre des inclusions monotones est démontrée en utilisant la notion de suites quasi-monotones au sens de Bregman étudiées dans le Chapitre 5.

7.1 Résultats techniques

D'abord, nous rappelons les notions de convexité et de lissité d'un espace de Banach.

Définition 7.1 [5] Soit $\mathcal{X}$ un espace de Banach réel. Le module de convexité de $\mathcal{X}$ est

$$\begin{aligned}\delta_{\mathcal{X}}\colon\ ]0,2] &\to [0,+\infty[\\ \varepsilon &\mapsto \inf\left\{1-\left\|\frac{x+y}{2}\right\| \;\middle|\; (x,y)\in\mathcal{X}^2,\ \|x\|=\|y\|=1,\ \|x-y\|\geqslant\varepsilon\right\},\end{aligned}\tag{7.1}$$

et le module de lissité de $\mathcal{X}$ est

$$\begin{aligned}\rho_{\mathcal{X}}\colon\ [0,+\infty[&\to [0,+\infty[\\ \tau &\mapsto \sup\left\{\frac{1}{2}(\|x+y\|+\|x-y\|)-1 \;\middle|\; (x,y)\in\mathcal{X}^2,\ \|x\|=1,\ \|y\|\leqslant\tau\right\}.\end{aligned}\tag{7.2}$$

On dit que $\mathcal{X}$ est *uniformément convexe* si $\delta_{\mathcal{X}}$ s'annule seulement en zero, et *uniformément lisse* si $\lim_{\tau\to 0}\rho_{\mathcal{X}}(\tau)/\tau=0$.

Exemple 7.2 [16, Theorem II.4.7] Soient $p\in\ ]1,+\infty[$ et $(\Omega,\mathcal{F},\mu)$ un espace mesuré. Alors $L^p(\Omega,\mathcal{F},\mu)$ est uniformément convexe et uniformément lisse.

Voici quelques rappels sur l'application de dualité d'un espace de Banach.

Définition 7.3 [16, Definition 4.1] Soient $\mathcal{X}$ un espace de Banach réel réflexif, $\mathcal{X}^*$ le dual topologique de $\mathcal{X}$, et $\phi\colon\ [0,+\infty[\ \to [0,+\infty[$ une fonction croissante telle que $\phi(0) = 0$ et $\lim_{t\to+\infty}\phi(t) = +\infty$. L'application de dualité Δ_ϕ de $\mathcal{X}$ par rapport à ϕ est

$$\Delta_\phi\colon \mathcal{X} \to 2^{\mathcal{X}^*}\colon x \mapsto \big\{x^* \in \mathcal{X}^* \;\big|\; \langle x, x^*\rangle = \|x\|\,\|x^*\| \text{ et } \|x^*\| = \phi(\|x\|)\big\}. \tag{7.3}$$

Exemple 7.4 [16, Propositions 4.8(i) et 4.14] L'application de dualité d'un espace hilbertien réel par rapport à la fonction $[0,+\infty[\ \ni t \mapsto t$ et l'application de dualité de $\ell^p(\mathbb{N})$ par rapport à la fonction $[0,+\infty[\ \ni t \mapsto t^r$, où $(p,r) \in\ ]1,+\infty[^2$, sont faiblement séquentiellement continues.

Lemme 7.5 [2, Section 5] *Soient $\mathcal{X}$ un espace de Banach réel réflexif, $\mathcal{X}^*$ le dual topologique de $\mathcal{X}$, et $f \in \Gamma_0(\mathcal{X})$. Alors f est* de Legendre *(i.e., f est* essentiellement lisse *au sens où ∂f est à la fois localement borné et univoque sur son domaine, et* essentiellement strictement convexe *au sens où ∂f^* est localement borné sur son domaine et f est strictement convexe sur les sous-ensembles convexes de* $\operatorname{dom}\partial f$*) si et seulement si les conditions suivantes sont satisfaites :*

(i) *∂f est univoque sur* $\operatorname{dom}\partial f = \operatorname{int}\operatorname{dom} f \neq \varnothing$.

(ii) *∂f^* est univoque sur* $\operatorname{dom}\partial f^* = \operatorname{int}\operatorname{dom} f^* \neq \varnothing$.

De plus, f^* est de Legendre et $\nabla f\colon$ $\operatorname{int}\operatorname{dom} f \to \operatorname{int}\operatorname{dom} f^*$ est bijective avec pour inverse $\nabla f^*\colon \operatorname{int}\operatorname{dom} f^* \to \operatorname{int}\operatorname{dom} f$.

Le lemme suivant sera utilisé dans la suite.

Lemme 7.6 *Soient $\mathcal{X}$ un espace de Banach réel réflexif, $f \in \Gamma_0(\mathcal{X})$ une fonction de Legendre, $x \in \operatorname{dom} f$, $y \in \operatorname{int}\operatorname{dom} f$, $z \in \operatorname{int}\operatorname{dom} f$, et $\lambda \in [0,1]$. Alors*

$$D^f\big(x, \nabla f^*(\lambda\nabla f(y) + (1-\lambda)\nabla f(z)\big) \leqslant \lambda D^f(x,y) + (1-\lambda)D^f(x,z). \tag{7.4}$$

Démonstration. Notons que $u = \nabla f^*(\lambda\nabla f(y) + (1-\lambda)f(z)) \in \operatorname{int}\operatorname{dom} f$ d'après le Lemme 7.5. Maintenant, la convexité de f^* implique que

$$\begin{aligned}
&D^f(x,u) = f(x) - f(u) - \big\langle x-u, \nabla f(u)\big\rangle\\
&= f(x) - \big\langle x, \nabla f(u)\big\rangle + f^*(\nabla f(u))\\
&= f(x) - \big\langle x, \lambda\nabla f(y) + (1-\lambda)\nabla f(z)\big\rangle + f^*\big(\lambda\nabla f(y) + (1-\lambda)\nabla f(z)\big)\\
&\leqslant \lambda\big((f(x) - \big\langle x, \nabla f(y)\big\rangle + f^*(\nabla f(y))\big) + (1-\lambda)\big(f(x) - \big\langle x, \nabla f(z)\big\rangle + f^*(\nabla f(z))\big)\\
&= \lambda D^f(x,y) + (1-\lambda)D^f(x,z).
\end{aligned} \tag{7.5}$$

□

Nous aurons également besoin de la condition suivante.

Condition 7.7 [3, Condition 4.4] Soit $\mathcal{X}$ un espace de Banach réel réflexif et soit $f \in \Gamma_0(\mathcal{X})$ Gâteaux différentiable sur $\operatorname{int} \operatorname{dom} f \neq \varnothing$. Pour toutes suites bornées $(x_n)_{n\in\mathbb{N}}$ et $(y_n)_{n\in\mathbb{N}}$ dans $\operatorname{int} \operatorname{dom} f$,

$$D^f(x_n, y_n) \to 0 \quad \Rightarrow \quad x_n - y_n \to 0. \tag{7.6}$$

7.2 Contractivité et résolvante relatives aux distances de Bregman

Une contraction T d'un espace de Banach $\mathcal{X}$ dans lui-même est une application lipschitzienne de constante 1, i.e.,

$$(\forall x \in \mathcal{X})(\forall y \in \mathcal{X}) \quad \|Tx - Ty\| \leqslant \|x - y\|. \tag{7.7}$$

En outre, T est une contraction ferme [12] si

$$(\forall x \in \mathcal{X})(\forall y \in \mathcal{X})(\forall \alpha \in \,]0, +\infty[) \quad \|Tx-Ty\| \leqslant \|\alpha(x-y)+(1-\alpha)(Tx-Ty)\|. \tag{7.8}$$

Dans un espace de Banach, il est difficile à manipuler les inégalités (7.7) et (7.8). Leurs variantes basées sur les distances de Bregman sont étudiées dans un nombre de travaux (voir [3, 14] et leurs bibliographies). Un tel exemple est la classe $\mathfrak{B}$.

Définition 7.8 [14] Soient $\mathcal{X}$ un espace de Banach réel, $f \in \Gamma_0(\mathcal{X})$ Gâteaux différentiable sur $\operatorname{int} \operatorname{dom} f \neq \varnothing$, et $T\colon \mathcal{X} \to \mathcal{X}$ tel que $\operatorname{ran} T \cup \operatorname{dom} T \subset \operatorname{int} \operatorname{dom} f$. Alors

$$\widehat{\operatorname{Fix}}\, T = \big\{x \in \overline{\operatorname{dom}}\, f \mid \exists (x_n)_{n\in\mathbb{N}} \in (\operatorname{int} \operatorname{dom} f)^{\mathbb{N}} \text{ telle que } x_n \rightharpoonup x \text{ et } x_n - Tx_n \to 0\big\}. \tag{7.9}$$

Définition 7.9 [3, Definition 3.1] et [14] Soient $\mathcal{X}$ un espace de Banach réel et $f \in \Gamma_0(\mathcal{X})$ Gâteaux différentiable sur $\operatorname{int} \operatorname{dom} f \neq \varnothing$. On pose, pour tous x et y dans $\operatorname{int} \operatorname{dom} f$,

$$H^f(x, y) = \big\{z \in \mathcal{X} \mid \langle z - y, \nabla f(x) - \nabla f(y)\rangle \leqslant 0\big\}. \tag{7.10}$$

Alors

$$\mathfrak{B}(f) = \Big\{T\colon \mathcal{X} \to 2^{\mathcal{X}} \;\Big|\; \operatorname{ran} T \subset \operatorname{dom} T = \operatorname{int} \operatorname{dom} f \\ \text{et } (\forall (x, y) \in \operatorname{gra} T)\ \operatorname{Fix} T \subset H^f(x, y)\Big\}, \tag{7.11}$$

et

$$\widehat{\mathfrak{B}}(f) = \Big\{T\colon \mathcal{X} \to 2^{\mathcal{X}} \;\Big|\; \operatorname{ran} T \subset \operatorname{dom} T = \operatorname{int} \operatorname{dom} f \\ \text{et } (\forall (x, y) \in \operatorname{gra} T)\ \widehat{\operatorname{Fix}}\, T \subset H^f(x, y)\Big\}. \tag{7.12}$$

Définition 7.10 [3, Definition 3.4] Soient $\mathcal{X}$ un espace de Banach réel et $f \in \Gamma_0(\mathcal{X})$ Gâteaux différentiable sur $\operatorname{int}\operatorname{dom} f \neq \varnothing$. Un opérateur $T\colon \mathcal{X} \to 2^{\mathcal{X}}$ tel que $\operatorname{ran} T \cup \operatorname{dom} T \subset \operatorname{int}\operatorname{dom} f$ est D^f-ferme si

$$(\forall (x_1, y_1) \in \operatorname{gra} T)(\forall (x_2, y_2) \in \operatorname{gra} T) \quad \big\langle y_1 - y_2, \nabla f(y_1) - \nabla f(y_2)\big\rangle \leqslant \big\langle y_1 - y_2, \nabla f(x_1) - \nabla f(x_2)\big\rangle. \tag{7.13}$$

Remarque 7.11 Soient $\mathcal{X}$ un espace de Banach réel réflexif, $f \in \Gamma_0(\mathcal{X})$ une fonction de Legendre, $T \in \mathfrak{B}(f)$, $\alpha \in\,]0,1[$, et $T_\alpha = \nabla f^*\big((1-\alpha)\nabla f + \alpha \nabla f \circ T\big)$. Alors $T_\alpha \in \mathfrak{B}(f)$.

Démonstration. D'abord, $\operatorname{ran} T_\alpha \subset \operatorname{ran} \nabla f^* = \operatorname{int}\operatorname{dom} f = \operatorname{dom} T = \operatorname{dom} T_\alpha$ et $\operatorname{Fix} T_\alpha = \operatorname{Fix} T$. Ensuite, soit $z \in \operatorname{Fix} T_\alpha$ et soit $(x, y_\alpha) \in \operatorname{gra} T_\alpha$. Alors $y \in \operatorname{Fix} T$ et il existe $y \in Tx$ tel que $\nabla f(x) - \nabla f(y_\alpha) = \alpha(\nabla f(x) - \nabla f(y))$. Puisque ∇f est monotone,

$$\begin{aligned}\big\langle y - y_\alpha, \nabla f(x) - \nabla f(y)\big\rangle &= (1-\alpha)\big\langle y - y_\alpha, \nabla f(y) - \nabla f(x)\big\rangle/(\alpha - 1)\\ &= \big\langle y - y_\alpha, \nabla f(y) - \nabla f(y_\alpha)\big\rangle/(\alpha-1)\\ &\leqslant 0.\end{aligned} \tag{7.14}$$

Par conséquent,

$$\begin{aligned}&\big\langle z - y_\alpha, \nabla f(x) - \nabla f(y_\alpha)\big\rangle\\ &= \big\langle z - y, \nabla f(x) - \nabla f(y_\alpha)\big\rangle + \big\langle y - y_\alpha, \nabla f(x) - \nabla f(y_\alpha)\big\rangle\\ &= \alpha\big\langle z - y, \nabla f(x) - \nabla f(y)\big\rangle + \alpha\big\langle y - y_\alpha, \nabla f(x) - \nabla f(y)\big\rangle\\ &\leqslant 0,\end{aligned} \tag{7.15}$$

ce qui implique l'assertion. □

Définition 7.12 [3, Definition 3.7] Soient $\mathcal{X}$ un espace de Banach réel, $\mathcal{X}^*$ le dual topologique de $\mathcal{X}$, $f \in \Gamma_0(\mathcal{X})$ Gâteaux différentiable sur $\operatorname{int}\operatorname{dom} f \neq \varnothing$, et $A\colon \mathcal{X} \to 2^{\mathcal{X}^*}$. La D^f-résolvante de A est

$$R_A^f = (\nabla f + A)^{-1} \circ \nabla f\colon \mathcal{X} \to 2^{\mathcal{X}} \tag{7.16}$$

et la f-résolvante de A est

$$J_A^f = (\nabla f + A)^{-1}\colon \mathcal{X}^* \to 2^{\mathcal{X}}. \tag{7.17}$$

Lemme 7.13 [3, Propostion 3.8] *Soient $\mathcal{X}$ un espace de Banach réel, $\mathcal{X}^*$ le dual topologique de $\mathcal{X}$, $f \in \Gamma_0(\mathcal{X})$ Gâteaux différentiable sur* $\operatorname{int}\operatorname{dom} f \neq \varnothing$*, et $A\colon \mathcal{X} \to 2^{\mathcal{X}^*}$. Les propriétés suivantes sont satisfaites :*

(i) $\operatorname{dom} R_A^f \subset \operatorname{int}\operatorname{dom} f$, $\operatorname{ran} R_A^f \subset \operatorname{int}\operatorname{dom} f$, *et* $\operatorname{Fix} R_A^f = \operatorname{zer} A \cap \operatorname{int}\operatorname{dom} f$.

(ii) *Supposons que A soit monotone. Alors :*

(a) *R_A^f est D^f-ferme.*

(b) *Supposons que* $f|_{\operatorname{int}\operatorname{dom} f}$ *soit strictement convexe. Alors* R_A^f *est univoque sur son domaine.*

(c) *Supposons que* $\operatorname{ran}\nabla f \subset \operatorname{ran}(\nabla f + A)$. *Alors* $R_A^f \in \mathfrak{B}(f)$.

Exemple 7.14 Soient m un nombre entier strictement positif, $(\mathcal{X}_i, \|\cdot\|_i)_{1\leqslant i\leqslant m}$ des espaces de Banach réels réflexifs, $\boldsymbol{\mathcal{X}}$ l'espace vectoriel $\bigtimes_{i=1}^m \mathcal{X}_i$ muni de la norme $\boldsymbol{x} = (x_i)_{1\leqslant i\leqslant m} \mapsto \sqrt{\sum_{i=1}^m \|x_i\|^2}$, et $\boldsymbol{\mathcal{X}}^*$ le dual topologique de $\boldsymbol{\mathcal{X}}$, i.e., l'espace vectoriel $\bigtimes_{i=1}^m \mathcal{X}_i^*$ muni de la norme $\boldsymbol{x}^* = (x_i^*)_{1\leqslant i\leqslant m} \mapsto \sqrt{\sum_{i=1}^m \|x_i^*\|^2}$. Pour tout $i \in \{1,\ldots,m\}$, soit $f_i \in \Gamma_0(\mathcal{X}_i)$ une fonction de Legendre et soit $A_i \colon \mathcal{X}_i \to 2^{\mathcal{X}_i^*}$ un opérateur monotone tel que $\operatorname{ran}\nabla f_i \subset \operatorname{ran}(\nabla f_i + A_i)$. Posons

$$\boldsymbol{f} \colon \boldsymbol{\mathcal{X}} \to \left]-\infty,+\infty\right] : \boldsymbol{x} \mapsto \sum_{i=1}^m f_i(x_i) \quad \text{et} \quad \boldsymbol{A} \colon \boldsymbol{\mathcal{X}} \to 2^{\boldsymbol{\mathcal{X}}^*} : \boldsymbol{x} \mapsto \bigtimes_{i=1}^m A_i x_i. \tag{7.18}$$

Alors $R_{\boldsymbol{A}}^{\boldsymbol{f}} = \bigtimes_{i=1}^m R_{A_i}^{f_i}$.

Démonstration. Le Lemme 7.13(ii) implique que, pour tout $i \in \{1,\ldots,m\}$, l'opérateur

$$R_{A_i}^{f_i} \colon \operatorname{int}\operatorname{dom} f_i \to \operatorname{dom} A_i \cap \operatorname{int}\operatorname{dom} f_i \tag{7.19}$$

est bien défini et univoque. D'une part, pour tout $i \in \{1,\ldots,m\}$, comme f_i est de Legendre, il résulte du Lemme 7.5 que ∂f_i est univoque sur $\operatorname{dom}\partial f_i = \operatorname{int}\operatorname{dom} f_i$. D'autre part, on déduit de (7.18) que $\operatorname{dom}\boldsymbol{f} = \bigtimes_{i=1}^m \operatorname{dom} f_i$ et que $\partial\boldsymbol{f} \colon \boldsymbol{\mathcal{X}} \to 2^{\boldsymbol{\mathcal{X}}^*} : \boldsymbol{x} \mapsto \bigtimes_{i=1}^m \partial f_i(x_i)$. Ainsi $\partial\boldsymbol{f}$ est univoque sur $\operatorname{dom}\partial\boldsymbol{f} = \bigtimes_{i=1}^m \operatorname{dom}\partial f_i = \bigtimes_{i=1}^m \operatorname{int}\operatorname{dom} f_i = \operatorname{int}\big(\bigtimes_{i=1}^m \operatorname{dom} f_i\big) = \operatorname{int}\operatorname{dom}\boldsymbol{f}$. De même façon,

$$\boldsymbol{f}^* \colon \boldsymbol{\mathcal{X}}^* \to \left]-\infty,+\infty\right] : \boldsymbol{x}^* \mapsto \sum_{i=1}^m f_i^*(x_i^*), \tag{7.20}$$

et $\partial\boldsymbol{f}^*$ est univoque sur $\operatorname{dom}\partial\boldsymbol{f}^* = \operatorname{int}\operatorname{dom}\boldsymbol{f}^*$. Par conséquent, le Lemme 7.5 affirme que $\boldsymbol{f}$ est une fonction de Legendre. Par ailleurs,

$$\operatorname{ran}\nabla\boldsymbol{f} = \bigtimes_{i=1}^m \operatorname{ran}\nabla f_i \subset \bigtimes_{i=1}^m \operatorname{ran}(\nabla f_i + A_i) = \operatorname{ran}(\nabla\boldsymbol{f} + \boldsymbol{A}). \tag{7.21}$$

Notons que, pour tout $i \in \{1,\ldots,m\}$, A_i est monotone, et donc, il résulte de (7.18) que $\boldsymbol{A}$ est monotone. Ainsi, d'après (7.21) et le Lemme 7.13(ii),

$$R_{\boldsymbol{A}}^{\boldsymbol{f}} \colon \operatorname{int}\operatorname{dom}\boldsymbol{f} \to \operatorname{dom}\boldsymbol{A} \cap \operatorname{int}\operatorname{dom}\boldsymbol{f}. \tag{7.22}$$

Maintenant, soit $\boldsymbol{x} \in \operatorname{int}\operatorname{dom}\boldsymbol{f}$, soit $\boldsymbol{p} = R_{\boldsymbol{A}}^{\boldsymbol{f}}\boldsymbol{x}$, et soit $(p_i)_{1\leqslant i\leqslant m} \in \boldsymbol{\mathcal{X}}$. Alors

$$\begin{aligned}
(\forall i \in \{1,\ldots,m\}) \quad p_i = R_{A_i}^{f_i} x_i &\Leftrightarrow (\forall i \in \{1,\ldots,m\}) \quad \nabla f_i(x_i) - \nabla f_i(p_i) \in A_i p_i \\
&\Leftrightarrow (\nabla f_i(x_i))_{1\leqslant i\leqslant m} - (\nabla f_i(p_i))_{1\leqslant i\leqslant m} \in \bigtimes_{i=1}^m A_i p_i \\
&\Leftrightarrow \nabla\boldsymbol{f}(\boldsymbol{x}) - \nabla\boldsymbol{f}((p_i)_{1\leqslant i\leqslant m}) \in \boldsymbol{A}(p_i)_{1\leqslant i\leqslant m} \\
&\Leftrightarrow (p_i)_{1\leqslant i\leqslant m} = R_{\boldsymbol{A}}^{\boldsymbol{f}}\boldsymbol{x} = \boldsymbol{p}.
\end{aligned} \tag{7.23}$$

□

Lemme 7.15 *Soient $\mathcal{X}$ un espace de Banach réel réflexif, $\mathcal{X}^*$ le dual topologique de $\mathcal{X}$, $f \in \Gamma_0(\mathcal{X})$ une fonction de Legendre cofinie, et $A\colon \mathcal{X} \to 2^{\mathcal{X}^*}$ maximalement monotone tel que $\operatorname{dom} A \cap \operatorname{int} \operatorname{dom} f \neq \varnothing$. Alors les propriétés suivantes sont satisfaites :*

(i) $J_A^f\colon \mathcal{X}^* \to \operatorname{dom} A \cap \operatorname{int} \operatorname{dom} f$ *est univoque.*

(ii) *Soient $z^* \in \mathcal{X}^*$ et $B = z^* + A$. Alors $J_B^f = J_A^f(\cdot - z^*)$.*

Démonstration. (i) : D'abord, comme f est une fonction de Legendre cofinie, $\operatorname{ran} \nabla f = \operatorname{int} \operatorname{dom} f^* = \mathcal{X}^*$, et donc, [8, Proposition 2.6 et Lemma 2.7] impliquent que $\operatorname{dom} J_A^f = \mathcal{X}^*$. Ensuite, comme $f|_{\operatorname{int} \operatorname{dom} f}$ est strictement convexe, ∇f est strictement monotone [17, Theorem 2.4.4(ii)] et $\nabla f + A$ l'est aussi. Si (x^*, x_1) et (x^*, x_2) sont deux éléments dans $\operatorname{gra} J_A^f$ tels que $x_1 \neq x_2$ alors (x_1, x^*) et (x_2, x^*) appartiennent à $\operatorname{gra}(\nabla f + A)$ et la monotonie stricte de $\nabla f + A$ implique que $0 < \langle x_1 - x_2, x^* - x^* \rangle = 0$, qui est absurde. Cela montre que $J_A^f\colon \mathcal{X}^* \to \operatorname{dom} A \cap \operatorname{int} \operatorname{dom} f$ est univoque.

(ii) : Par le même raisonnement que précédemment, $J_B^f\colon \mathcal{X}^* \to \operatorname{dom} B \cap \operatorname{int} \operatorname{dom} f$ est bien défini. Soient $x \in \mathcal{X}$ et $x^* \in \mathcal{X}^*$. Alors

$$x = J_B^f x^* \Leftrightarrow x^* \in \nabla f(x) + Bx \Leftrightarrow x^* - z^* \in \nabla f(x) + Ax \Leftrightarrow x = J_A^f(x^* - z^*). \quad (7.24)$$

□

Définition 7.16 [1, Definition 3.58] Soit $\mathcal{X}$ un espace de Banach réel réflexif et soit $\mathcal{X}^*$ le dual topologique de $\mathcal{X}$. La suite d'opérateurs $(A_n)_{n\in\mathbb{N}}$ de $\mathcal{X}$ dans $2^{\mathcal{X}^*}$ converge graphiquement vers l'opérateur $A\colon \mathcal{X} \to 2^{\mathcal{X}^*}$ (noté $A_n \xrightarrow{G} A$) si :

(i) Pour tout $(x, x^*) \in \operatorname{gra} A$, il existe une suite $(x_n, x_n^*)_{n\in\mathbb{N}}$ dans $\mathcal{X} \times \mathcal{X}^*$ telle que $(x_n, x_n^*) \to (x, x^*)$ et $(\forall n \in \mathbb{N})\ (x_n, x_n^*) \in \operatorname{gra} A_n$.

(ii) Soit $(x_n, x_n^*)_{n\in\mathbb{N}}$ une suite dans $\mathcal{X} \times \mathcal{X}^*$ telle que $(\forall n \in \mathbb{N})\ (x_n, x_n^*) \in \operatorname{gra} A_n$. Si $(k_n)_{n\in\mathbb{N}}$ est une suite strictement croissante dans $\mathbb{N}$ telle que $(x_{k_n}, x_{k_n}^*) \to (x, x^*)$ alors $(x, x^*) \in \operatorname{gra} A$.

Nous démontrons le résultat suivant qui étend [1, Proposition 3.60] pour le cas hilbertien.

Proposition 7.17 *Soit $\mathcal{X}$ un espace de Banach réel réflexif, soit $\mathcal{X}^*$ le dual topologique de $\mathcal{X}$, et soit $f \in \Gamma_0(\mathcal{X})$ une fonction de Legendre supercoercive, soient A et $(A_n)_{n\in\mathbb{N}}$ des opérateurs maximalement monotones de $\mathcal{X}$ dans $2^{\mathcal{X}^*}$ tels que $\operatorname{dom} A \cup (\cup_{n\in\mathbb{N}} \operatorname{dom} A_n) \subset \operatorname{int} \operatorname{dom} f$, et soit $(x_n)_{n\in\mathbb{N}} \in (\operatorname{int} \operatorname{dom} f)^{\mathbb{N}}$. Supposons que ∇f soit uniformément continu sur les sous-ensembles bornés de $\operatorname{int} \operatorname{dom} f$, que $A_n \xrightarrow{G} A$, et que $x_n \to x \in \operatorname{int} \operatorname{dom} f$. Alors $R_{A_n}^f x_n \rightharpoonup R_A^f x$.*

Démonstration. Comme f est une fonction de Legendre supercoercive, ∇f est maximalement monotone et $\operatorname{ran}\nabla f = \mathcal{X}^*$ [16, Theorem 18.7] and [2, Theorem 5.6]. On alors déduit du [3, Theorem 3.13] et du Lemme 7.13 que les opérateurs R^f_A et $(R^f_{A_n})_{n\in\mathbb{N}}$ sont bien définis et univoques sur $\operatorname{int}\operatorname{dom} f$. Posons

$$(\forall n \in \mathbb{N}) \quad y_n = R^f_{A_n} x_n \quad \text{et} \quad y_n^* = \nabla f(x_n) - \nabla f(y_n). \tag{7.25}$$

Alors $(\forall n \in \mathbb{N})$ $(y_n, y_n^*) \in \operatorname{gra} A_n$. Ensuite, soit $(u, u^*) \in \operatorname{gra} A$. Comme $A_n \xrightarrow{G} A$, il existe une suite $(u_n, u_n^*)_{n\in\mathbb{N}}$ dans $\mathcal{X} \times \mathcal{X}^*$ telle que $(u_n, u_n^*) \to (u, u^*)$ et $(\forall n \in \mathbb{N})$ $(u_n, u_n^*) \in \operatorname{gra} A_n$. Comme $x_n \to x$, $u_n \to u$, et f et ∇f sont continus sur $\operatorname{int}\operatorname{dom} f$, on déduit que

$$\rho = \sup_{n\in\mathbb{N}}\{\|x_n\|, \|u_n\|, \|u_n^*\|, |f(x_n)|, |f(u_n)|, \|\nabla f(x_n)\|, \|\nabla f(u_n)\|\} < +\infty. \tag{7.26}$$

Comme les opérateurs $(A_n)_{n\in\mathbb{N}}$ sont monotones,

$$(\forall n \in \mathbb{N}) \quad 0 \leqslant \big\langle y_n - u_n, \nabla f(x_n) - \nabla f(y_n) - u_n^* \big\rangle, \tag{7.27}$$

et ainsi

$$\begin{aligned}(\forall n \in \mathbb{N}) \quad & f(y_n) - f(u_n) - \big\langle y_n - u_n, \nabla f(x_n) \big\rangle \\ & = D^f(y_n, x_n) - D^f(u_n, x_n) \leqslant D^f(u_n, y_n) + D^f(y_n, x_n) - D^f(u_n, x_n) \\ & = \big\langle u_n - y_n, \nabla f(x_n) - \nabla f(y_n) \big\rangle \leqslant \big\langle u_n - y_n, u_n^* \big\rangle \leqslant \rho\|y_n\| + \rho^2. \end{aligned} \tag{7.28}$$

Par suite,

$$(\forall n \in \mathbb{N}) \quad f(y_n) \leqslant \rho + 2\rho\|y_n\| + \rho^2. \tag{7.29}$$

Si $(y_n)_{n\in\mathbb{N}}$ n'est pas bornée alors il existe une suite strictement croissante $(k_n)_{n\in\mathbb{N}}$ dans $\mathbb{N}$ telle que $\|y_{k_n}\| \uparrow +\infty$ et il résulte de (7.29) que

$$+\infty \leftarrow \frac{f(y_{k_n})}{\|y_{k_n}\|} \leqslant \frac{\rho + \rho^2}{\|y_0\|} + 2\rho, \tag{7.30}$$

qui est absurde, i.e., $(y_n)_{n\in\mathbb{N}}$ est bornée. Par suite, il résulte de (7.25) et le Lemme 2.22 que $(y_n^*)_{n\in\mathbb{N}}$ est bornée. Soit $y \in \mathcal{X}$, soit $y^* \in \mathcal{X}^*$, et soit $(k_n)_{n\in\mathbb{N}}$ une suite strictement croissante dans $\mathbb{N}$ tels que $y_{k_n} \rightharpoonup y$ et $y_{k_n}^* \rightharpoonup y^*$. Ensuite, on déduit de la monotonie de ∇f que

$$(\forall n \in \mathbb{N})(\forall m \in \mathbb{N}) \quad \langle y_n - y_m, \nabla f(y_n) - \nabla f(y_m) \rangle \geqslant 0, \tag{7.31}$$

et en combinant avec (7.25), on obtient

$$(\forall n \in \mathbb{N})(\forall m \in \mathbb{N}) \quad \langle y_n - y_m, y_n^* - y_m^* \rangle \leqslant \big\langle y_n - y_m, \nabla f(x_n) - \nabla f(x_m) \big\rangle. \tag{7.32}$$

Ainsi

$$(\forall n \in \mathbb{N})(\forall m \in \mathbb{N}) \quad \langle y_n, y_n^* \rangle + \langle y_m, y_m^* \rangle \leqslant \langle y_m, y_n^* \rangle + \langle y_n, y_m^* \rangle + \langle y_n - y_m, \nabla f(x_n) - \nabla f(x_m) \rangle. \quad (7.33)$$

En fixant $m \in \mathbb{N}$ et en passant à la limite dans (7.33) quand $n \to +\infty$ par rapport à la sous-suite $(y_{k_n})_{n\in\mathbb{N}}$, on obtient

$$(\forall m \in \mathbb{N}) \quad \overline{\lim} \langle y_{k_n}, y_{k_n}^* \rangle + \langle y_m, y_m^* \rangle \leqslant \langle y_m, y^* \rangle + \langle y, y_m^* \rangle + \langle y - y_m, \nabla f(x) - \nabla f(x_m) \rangle. \quad (7.34)$$

Ensuite, en passant à la limite dans (7.34) quand $m \to +\infty$ par rapport à la sous-suite $(y_{k_m})_{n\in\mathbb{N}}$, on obtient

$$\overline{\lim} \langle y_{k_n}, y_{k_n}^* \rangle \leqslant \langle y, y^* \rangle . \quad (7.35)$$

Par ailleurs, comme les opérateurs $(A_n)_{n\in\mathbb{N}}$ sont monotones,

$$(\forall n \in \mathbb{N}) \quad \langle y_n - u_n, y_n^* - u_n^* \rangle \geqslant 0, \quad (7.36)$$

et donc,

$$(\forall n \in \mathbb{N}) \quad \langle y_n, u_n^* \rangle + \langle u_n, y_n^* \rangle \leqslant \langle y_n, y_n^* \rangle + \langle u_n, u_n^* \rangle . \quad (7.37)$$

Ainsi, en combinant avec (7.35), on obtient

$$\begin{aligned} \langle \overline{y}, u^* \rangle + \langle u, \overline{y}^* \rangle &= \lim \langle y_{k_n}, u_{k_n}^* \rangle + \lim \langle u_{k_n}, y_{k_n}^* \rangle \\ &= \lim \left(\langle y_{k_n}, u_{k_n}^* \rangle + \langle u_{k_n}, y_{k_n}^* \rangle \right) \\ &\leqslant \overline{\lim} \left(\langle y_{k_n}, y_{k_n}^* \rangle + \langle u_{k_n}, u_{k_n}^* \rangle \right) \\ &\leqslant \overline{\lim} \langle y_{k_n}, y_{k_n}^* \rangle + \overline{\lim} \langle u_{k_n}, u_{k_n}^* \rangle \\ &\leqslant \langle y, y^* \rangle + \langle u, u^* \rangle . \end{aligned} \quad (7.38)$$

Par conséquent, $\langle y - u, y^* - u^* \rangle \geqslant 0$ et comme (u, u^*) est arbitraire dans gra A et A est maximalement monotone, on en déduit que $(y, y^*) \in$ gra A. En remplaçant (u, u^*) par (y, y^*) et utilisant le même raisonnement que précédemment, on obtient

$$\begin{aligned} \underline{\lim} \langle y_{k_n}, y_{k_n}^* \rangle &\geqslant \underline{\lim} \left(\langle y_{k_n}, u_{k_n}^* \rangle + \langle u_{k_n}, y_{k_n}^* \rangle - \langle u_{k_n}, u_{k_n}^* \rangle \right) \\ &= \lim \left(\langle y_{k_n}, u_{k_n}^* \rangle + \langle u_{k_n}, y_{k_n}^* \rangle - \langle u_{k_n}, u_{k_n}^* \rangle \right) \\ &= \langle y, y^* \rangle . \end{aligned} \quad (7.39)$$

Par suite, il résulte de (7.35) que

$$\langle y_{k_n}, y_{k_n}^* \rangle \to \langle y, y^* \rangle . \quad (7.40)$$

Ensuite, on définit $B = \nabla f - \nabla f(x)$ et $(\forall n \in \mathbb{N})$ $B_n = \nabla f - \nabla f(x_n)$. Alors B est maximalement monotone et $(\forall n \in \mathbb{N})$ B_n est monotone et $(y_n, -y_n^*) \in \operatorname{gra} B_n$. Soit $(v, v^*) \in \operatorname{gra} B$, soit $(v_n)_{n\in\mathbb{N}}$ une suite dans $\mathcal{X}$ telle que $v_n \to v$, et on pose $(\forall n \in \mathbb{N})$ $v_n^* = \nabla f(v_n) - \nabla f(x_n)$. Alors $v_n^* \to \nabla f(v) - \nabla f(x) = v^*$ et $(\forall n \in \mathbb{N})$ $(v_n, v_n^*) \in \operatorname{gra} B_n$. Puisque les opérateurs $(B_n)_{n\in\mathbb{N}}$ sont monotones,

$$(\forall n \in \mathbb{N}) \quad \langle v_n - y_n, v_n^* + y_n^* \rangle \geqslant 0, \tag{7.41}$$

et ainsi

$$(\forall n \in \mathbb{N}) \quad \langle v_n, v_n^* \rangle + \langle v_n, y_n^* \rangle - \langle y_n, v_n^* \rangle - \langle y_n, y_n^* \rangle \geqslant 0. \tag{7.42}$$

En particulier,

$$(\forall n \in \mathbb{N}) \quad \left\langle v_{k_n}, v_{k_n}^* \right\rangle + \left\langle v_{k_n}, y_{k_n}^* \right\rangle - \left\langle y_{k_n}, v_{k_n}^* \right\rangle - \left\langle y_{k_n}, y_{k_n}^* \right\rangle \geqslant 0. \tag{7.43}$$

En utilisant (7.40) et en faisant tendre n vers $+\infty$ dans (7.43), on obtient $\langle v - y, v^* + y^* \rangle \geqslant 0$, ce qui implique que $(y, -y^*) \in \operatorname{gra} B$. Par suite,

$$\nabla f(x) = \nabla f(y) + \nabla f(x) - \nabla f(y) = \nabla f(y) + y^* \in \nabla f(y) + Ay, \tag{7.44}$$

et donc, $y = R_A^f x$. Cela implique que $y_n \rightharpoonup y$. □

Corollaire 7.18 *Soit $\mathcal{X}$ un espace de Banach réel réflexif, soit $\mathcal{X}^*$ le dual topologique de $\mathcal{X}$, et soit $f \in \Gamma_0(\mathcal{X})$ une fonction de Legendre supercoercive, soit $(x_n)_{n\in\mathbb{N}} \in (\operatorname{int}\operatorname{dom} f)^{\mathbb{N}}$ telle que $x_n \to x \in \operatorname{int}\operatorname{dom} f$, et soient $(\varphi_n)_{n\in\mathbb{N}}$ une suite de fonctions dans $\Gamma_0(\mathcal{X})$ convergeant au sens de Mosco vers $\varphi \in \Gamma_0(\mathcal{X})$ (noté $\varphi_n \xrightarrow{M} \varphi$), i.e.,*

(i) *Pour toute suite $(z_n)_{n\in\mathbb{N}} \in \mathcal{X}^{\mathbb{N}}$ telle que $z_n \rightharpoonup x \in \mathcal{X}$ et pour toute suite strictement croissante $(k_n)_{n\in\mathbb{N}}$ dans $\mathbb{N}$, on a $\underline{\lim}\, \varphi_{k_n}(z_{k_n}) \geqslant \varphi(x)$.*

(ii) *Pour tout $z \in \mathcal{X}$, il existe $(z_n)_{n\in\mathbb{N}} \in \mathcal{X}^{\mathbb{N}}$ telle que $z_n \to z$ et $\varphi_n(z_n) \to \varphi(z)$.*

Supposons que $\operatorname{dom}\partial\varphi \cup (\cup_{n\in\mathbb{N}} \operatorname{dom}\partial\varphi_n) \subset \operatorname{int}\operatorname{dom} f$ et que ∇f soit uniformément continu sur les sous-ensembles bornés de $\operatorname{int}\operatorname{dom} f$. Alors $\operatorname{prox}_{\varphi_n}^f x_n \rightharpoonup \operatorname{prox}_{\varphi}^f x$. En particulier, soient C et $(C_n)_{n\in\mathbb{N}}$ des ensembles convexes fermés non vides de $\mathcal{X}$ tels que $\iota_{C_n} \xrightarrow{M} \iota_C$ et $C \cup (\cup_{n\in\mathbb{N}} C_n) \subset \operatorname{int}\operatorname{dom} f$. Alors $P_{C_n}^f x_n \rightharpoonup P_C^f x$.

Démonstration. Il résulte de [1, Theorem 3.36] que $\partial\varphi_n \xrightarrow{G} \partial\varphi$ et donc, l'assertion est un corollaire de la Proposition 7.17. □

7.3 Cocoercivité relative aux distances de Bregman

La notion de cocoercivité joue un rôle important dans l'algorithme explicite-implicite dans les espaces hilbertiens [11]. Un opérateur B d'un espace hilbertien réel dans lui-même est β-cocoercif pour un certain $\beta \in\,]0, +\infty[$ si $B^{-1} - \beta\,\mathrm{Id}$ est monotone.

Définition 7.19 [9, Definition 3.3] Soient $\mathcal{X}$ un espace de Banach réel réflexif, $\mathcal{X}^*$ le dual topologique de $\mathcal{X}$, $f \in \Gamma_0(\mathcal{X})$ une fonction de Legendre, $B\colon \mathcal{X} \to \mathcal{X}^*$ tel que $\operatorname{dom} B \cap \operatorname{int} \operatorname{dom} f \neq \varnothing$, et $\beta \in \left]0, +\infty\right[$. Alors B est *d'inverse fortement monotone* relativement à f si

$$(\forall x_1 \in \operatorname{dom} B \cap \operatorname{int} \operatorname{dom} f)(\forall x_2 \in \operatorname{dom} B \cap \operatorname{int} \operatorname{dom} f) \\ \big\langle \nabla f^*\big(\nabla f(x_1) - Bx_1\big) - \nabla f^*\big(\nabla f(x_2) - Bx_2\big), Bx_1 - Bx_2\big\rangle \geqslant 0. \quad (7.45)$$

De plus, B est *$\beta-$cocoercif* relativement à f si βB est d'inverse fortement monotone relativement à f.

Remarque 7.20 Soient $\mathcal{X}$ un espace de Banach réel rélfexif, $\mathcal{X}^*$ le dual topologique de $\mathcal{X}$, $f \in \Gamma_0(\mathcal{X})$ une fonction de Legendre, $B\colon \mathcal{X} \to \mathcal{X}^*$ tel que $\operatorname{dom} B \cap \operatorname{int} \operatorname{dom} f \neq \varnothing$, et $\beta \in \left]0, +\infty\right[$. Alors B est $\beta-$cocoercif relativement à f si et seulement si $T = \nabla f^* \circ (\nabla f - \beta B)$ est D^f-ferme.

Démonstration. D'abord, $\operatorname{dom} T \subset \operatorname{int} \operatorname{dom} f$ et on déduit du Lemme 7.5 que $\operatorname{ran} T \subset \operatorname{ran} \nabla f^* = \operatorname{int} \operatorname{dom} f$. Ensuite, soient $(x_1, y_1) \in \operatorname{gra} T$ et $(x_2, y_2) \in \operatorname{gra} T$. Alors $Bx_1 = \beta^{-1}(\nabla f(x_1) - \nabla f(y_1))$ et $Bx_2 = \beta^{-1}(\nabla f(x_2) - \nabla f(y_2))$. Ainsi l'inégalité (7.45) est équivalente à l'inégalité

$$\big\langle y_1 - y_2, \big(\nabla f(x_1) - \nabla f(y_1)\big) - \big((\nabla f(x_2) - \nabla f(y_2)\big)\big\rangle \geqslant 0, \quad (7.46)$$

qui est elle-même équivalente au fait que T est D^f-ferme. □

Nous étudions du lien entre la notion de cocoercivité classique et la notion dans la Définition 7.19.

Remarque 7.21 Soient $\mathcal{X}$ un espace de Banach réel réflexif, $\mathcal{X}^*$ le dual topologique de $\mathcal{X}$, $f \in \Gamma_0(\mathcal{X})$ une fonction de Legendre, $B\colon \mathcal{X} \to \mathcal{X}^*$ tel que $\operatorname{dom} B \cap \operatorname{int} \operatorname{dom} f \neq \varnothing$, et $\beta \in \left]0, +\infty\right[$.

(i) Supposons que $\mathcal{X}$ soit hilbertien et que $f = \|\cdot\|^2/2$. Alors la cocoercivité au sens classique et la cocoercivité définie dans la Définition 7.19 sont identiques.

(ii) La cocoercivité au sens classique n'implique pas la cocoercivité au sens de la Définition 7.19. Par exemple, supposons que $\mathcal{X} = \mathbb{R}$, que $B = \operatorname{Id}$, et que $f = |\cdot|^4/4$. Alors $\nabla f\colon \mathbb{R} \to \mathbb{R}\colon \xi \mapsto \xi^3$ et $\nabla f^*\colon \mathbb{R} \to \mathbb{R}\colon \xi \mapsto \sqrt[3]{\xi}$. Il est clair que B est $(1/2)$-cocoercif. Si B est $\beta-$cocoercif relativement à f pour un certain $\beta \in \left]0, +\infty\right[$, alors

$$(\forall \xi_1 \in \mathbb{R})(\forall \xi_2 \in \mathbb{R}) \quad (\xi_1 - \xi_2)\big(\sqrt[3]{\xi_1^3 - \beta\xi_1} - \sqrt[3]{\xi_2^3 - \beta\xi_2}\big) \geqslant 0, \quad (7.47)$$

ce qui est impossible si l'on prend $0 < \xi_1 < \sqrt{\beta}$ et $\xi_2 = 0$.

(iii) La cocoercivité au sens de la Définition 7.19 n'implique pas la cocoercivité au sens classique. Par exemple, supposons que $\mathcal{X} = \mathbb{R}$, que $\beta \in \left]0,1\right[$, que f soit l'entropie de Boltzmann-Shannon, et que

$$B\colon \mathbb{R} \to \mathbb{R}\colon \xi \mapsto \begin{cases} \ln \xi, & \text{si } \xi > 0;\\ \varnothing, & \text{sinon.} \end{cases} \tag{7.48}$$

Alors $\nabla f^*\colon \mathbb{R} \to \mathbb{R}\colon \xi \mapsto e^{\xi}$. Comme $\ln$ est croissante, pour tout $(\xi_1,\xi_2) \in \left]0,+\infty\right[^2$, on a

$$\begin{aligned}
&\big(B\xi_1 - B\xi_2\big)\big(\nabla f^*\big(\nabla f(\xi_1) - \beta B\xi_1\big) - \nabla f^*\big(\nabla f(\xi_2) - \beta B\xi_2\big)\big)\\
&= (1-\beta)^{-1}(1-\beta)(\ln \xi_1 - \ln \xi_2)(e^{(1-\beta)\ln \xi_1} - e^{(1-\beta)\ln \xi_2})\\
&= (1-\beta)^{-1}(\ln \xi_1^{1-\beta} - \ln \xi_2^{1-\beta})(\xi_1^{1-\beta} - \xi_2^{1-\beta})\\
&\geqslant 0,
\end{aligned} \tag{7.49}$$

ce qui montre que B est $\beta-$cocoercif relativement à f. Si B est β-cocoercif alors

$$(\forall (\xi_1,\xi_2) \in \left]0,+\infty\right[^2) \quad \beta|\ln \xi_1 - \ln \xi_2|^2 \leqslant \big(\ln \xi_1 - \ln \xi_2\big)\big(\xi_1 - \xi_2\big). \tag{7.50}$$

En prenant $\xi_2 = 1$ et $\xi_1 = 2^{-1/(\beta \ln 2)}$ dans (7.50), on obtient $1 \leqslant 1 - 2^{-\frac{1}{\beta \ln 2}}$, qui est impossible. Cela implique que B n'est pas β-cocoercif.

Exemple 7.22 Soient $\mathcal{X}$ un espace de Banach réel réflexif, $f \in \Gamma_0(\mathcal{X})$ une fonction de Legendre telle que Argmin $f \neq \varnothing$, $\beta \in \left]0,1\right[$, $x_1 \in \operatorname{int dom} f$, et $x_2 \in \operatorname{int dom} f$. Alors

$$\begin{cases} x_1^* = \nabla f(x_1) - \beta \nabla f(x_1) = (1-\beta)\nabla f(x_1) + \beta 0 \in \operatorname{int dom} f^*,\\ x_2^* = \nabla f(x_2) - \beta \nabla f(x_2) = (1-\beta)\nabla f(x_2) + \beta 0 \in \operatorname{int dom} f^*. \end{cases} \tag{7.51}$$

Ainsi la monotonie de ∇f^* implique que

$$\begin{aligned}
&\big\langle \beta \nabla f(x_1) - \beta \nabla f(x_2), \nabla f^*(x_1^*) - \nabla f^*(x_2^*)\big\rangle\\
&\qquad = \frac{\beta}{1-\beta}\big\langle x_1^* - x_2^*, \nabla f^*(x_1^*) - \nabla f^*(x_2^*)\big\rangle \geqslant 0,
\end{aligned} \tag{7.52}$$

ce qui également montre que ∇f est $\beta-$cocoercif relativement à f.

Nous présentons une caractérisation d'un opérateur cocoercif au sens de Bregman dans $\mathbb{R}$ dans la proposition suivante.

Proposition 7.23 *Soient $\psi \in \Gamma_0(\mathbb{R})$ une fonction de Legendre deux fois différentiable sur* $\operatorname{int dom} \psi$, *$\beta \in \left]0,+\infty\right[$, et $B\colon \mathbb{R} \to \mathbb{R}$ un opérateur monotone. Supposons que B soit différentiable sur* $\operatorname{dom} B \cap \operatorname{int dom} \psi \neq \varnothing$ *et que*

$$(\forall \xi \in \operatorname{dom} B \cap \operatorname{int dom} \psi) \quad \beta B'(\xi) \leqslant \psi''(\xi). \tag{7.53}$$

Alors B est $\beta-$cocoercif relativement à ψ.

Démonstration. Comme ψ est de Legendre, $\psi'\colon \operatorname{int}\operatorname{dom}\psi \to \operatorname{int}\operatorname{dom}\psi^*$ est une bijection. Puisque $\psi|_{\operatorname{int}\operatorname{dom}\psi}$ est strictement convexe et deux fois différentiable, $\psi'' > 0$ [17, Theorem 2.1.10] et donc, il résulte du Théorème d'inversion que $(\psi^*)'' = ((\psi')^{-1})' = 1/\psi''$. Ensuite, posons $T = (\psi^*)' \circ (\psi' - \beta B)$ et supposons que ξ_1 et ξ_2 soient deux éléments génériques dans $\operatorname{dom} B \cap \operatorname{int}\operatorname{dom}\psi$. Comme $T' = (\psi'' - \beta B')/(\psi'' \circ T)$, (7.53) implique que $T' \geqslant 0$ sur $\operatorname{dom} B \cap \operatorname{int}\operatorname{dom}\psi$, i.e., T est croissant sur $\operatorname{dom} B \cap \operatorname{int}\operatorname{dom}\psi$. Par suite,

$$\big(T(\xi_1) - T(\xi_2)\big)(\xi_1 - \xi_2) \geqslant 0. \tag{7.54}$$

Comme B est monotone, $(B(\xi_1) - B(\xi_2))(\xi_1 - \xi_2) \geqslant 0$, et il résulte de (7.54) que

$$(\xi_1 - \xi_2)^2\big(T(\xi_1) - T(\xi_2)\big)\big(B(\xi_1) - B(\xi_2)\big) \geqslant 0. \tag{7.55}$$

Par conséquent, $(T(\xi_1) - T(\xi_2))((B(\xi_1) - B(\xi_2)) \geqslant 0$, ce qui montre que B est β-D^ψ-cocoercif. □

Exemple 7.24 Soient $\beta \in \,]0, +\infty[$, $\psi \in \Gamma_0(\mathbb{R})$ une fonction de Legendre, et $B\colon \mathbb{R} \to \mathbb{R}$ monotone et différentiable. Alors B est $\beta-$cocoercif relativement à ψ dans chaque cas suivant :

(i) ψ est l'entropie de Boltzmann-Shannon et $(\forall \xi \in \,]0, +\infty[)\ \beta B'(\xi) \leqslant \xi^{-1}$.

(ii) ψ est l'entropie de Fermi-Dirac et $(\forall \xi \in \,]0, 1[)\ \beta B'(\xi) \leqslant \xi^{-1}(1-\xi)^{-1}$.

(iii) ψ est l'entropie de Burg et $(\forall \xi \in \,]0, +\infty[)\ \beta B'(\xi) \leqslant \xi^{-1}$.

(iv) ψ est la fonction de type Hellinger et $(\forall \xi \in \,]-1, 1[)\ \beta B'(\xi) \leqslant (1-\xi^2)^{-3/2}$.

(v) ψ est la fonction exponentielle et $(\forall \xi \in \mathbb{R})\ \beta B'(\xi) \leqslant e^{\xi}$.

Démonstration. Proposition 7.23. □

7.4 Construction de points fixes de contractions

Dans cette section, nous construisons les approximations successives du point fixe d'une famille de contractions pour des distances de Bregman. Nous reprenons les Définitions 5.2 et 5.3 du Chapitre 5.

Proposition 7.25 *Soient $\mathcal{X}$ un espace de Banach réel réflexif, I un ensemble totalement ordonné au plus dénombrable, $(C_i)_{i\in I}$ une famille de sous-ensembles de $\mathcal{X}$ tels que $C = \bigcap_{i\in I} C_i \neq \varnothing$, $f \in \Gamma_0(\mathcal{X})$ une fonction de Legendre telle que $C \cap \operatorname{int}\operatorname{dom} f \neq \varnothing$, $\alpha \in \,]0, +\infty[$, $(\eta_n)_{n\in\mathbb{N}} \in \ell^1_+(\mathbb{N})$, et $(f_n)_{n\in\mathbb{N}}$ des fonctions de Legendre dans $\mathcal{P}_\alpha(f)$ telles que*

$$(\forall n \in \mathbb{N}) \quad (1+\eta_n) f_n \succcurlyeq f_{n+1}. \tag{7.56}$$

Soit $\mathrm{i}\colon \mathbb{N} \to I$ *une application telle que*

$$(\forall j \in I)(\exists M_j \in \mathbb{N}\backslash\{0\})(\forall n \in \mathbb{N}) \quad j \in \{\mathrm{i}(n), \ldots, \mathrm{i}(n+M_j-1)\}. \tag{7.57}$$

Pour tout $i \in I$*, soit* $(T_{i,n})_{n\in\mathbb{N}}$ *une suite d'opérateurs tels que*

$$(\forall n \in \mathbb{N}) \quad T_{i,n} \in \widehat{\mathfrak{B}}(f_n) \text{ et } C_i = \widehat{\operatorname{Fix}} T_{i,n}. \tag{7.58}$$

Soient $\varepsilon \in \left]0,1\right[$*,* $(\lambda_n)_{n\in\mathbb{N}} \in \left]\varepsilon, 1\right]^{\mathbb{N}}$*,* $x_0 \in \operatorname{int}\operatorname{dom} f$*, et posons*

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = \nabla f_n^* \left(\nabla f_n(x_n) + \lambda_n \left(\nabla f_n(T_{\mathrm{i}(n),n} x_n) - \nabla f_n(x_n)\right)\right). \tag{7.59}$$

Supposons que f *satisfasse la Condition 7.7 et que* $(\forall x \in \operatorname{int}\operatorname{dom} f)$ $D^f(x,\cdot)$ *soit coercive. Alors les propriétés suivantes sont satisfaites :*

(i) *Soit* $x \in C \cap \operatorname{int}\operatorname{dom} f$*. Alors* $(D^f(x, x_n))_{n\in\mathbb{N}}$ *est bornée.*

(ii) $(x_n)_{n\in\mathbb{N}}$ *est une suite bornée dans* $\operatorname{int}\operatorname{dom} f$*.*

(iii) $\sum_{n\in\mathbb{N}} D^f(T_{\mathrm{i}(n),n} x_n, x_n) < +\infty$*.*

(iv) *Supposons qu'il existe* $g \in \mathcal{F}(f)$ *telle que, pour tout* $n \in \mathbb{N}$*,* $g \succcurlyeq f_n$*, et que, pour tout* $y_1 \in \mathcal{X}$ *et pour tout* $y_2 \in \mathcal{X}$*,*

$$\begin{cases} y_1 \in \mathfrak{W}(x_n)_{n\in\mathbb{N}} \cap C \\ y_2 \in \mathfrak{W}(x_n)_{n\in\mathbb{N}} \cap C \\ \left(\langle y_1 - y_2, \nabla f_n(x_n)\rangle\right)_{n\in\mathbb{N}} \quad \textit{converge} \end{cases} \quad \Rightarrow \quad y_1 = y_2. \tag{7.60}$$

De plus, supposons que $\mathfrak{W}(x_n)_{n\in\mathbb{N}} \subset \operatorname{int}\operatorname{dom} f$*. Alors il existe* $\overline{x} \in C$ *tel que* $x_n \rightharpoonup \overline{x}$*.*

(v) *Supposons que* C *soit convexe et fermé, que* $\underline{\lim}\, D_C^f(x_n) = 0$*, et qu'il existe* $\beta \in \left]0, +\infty\right[$ *tel que* $(\forall n \in \mathbb{N})$ $\beta f \succcurlyeq f_n$*. Alors il existe* $\overline{x} \in C$ *tel que* $x_n \to \overline{x}$*.*

Démonstration. Notons que $x_0 \in \operatorname{int}\operatorname{dom} f$. Supposons que $x_n \in \operatorname{int}\operatorname{dom} f$ pour un certain $n \in \mathbb{N}$. Alors $(\forall i \in I)$ $T_{i,n}x_n \in \operatorname{ran} T_{i,n} \subset \operatorname{int}\operatorname{dom} f_n$. Puisque $\nabla f_n\colon \operatorname{int}\operatorname{dom} f_n \to \operatorname{int}\operatorname{dom} f_n^*$ est bijective avec pour inverse est $\nabla f_n^*\colon \operatorname{int}\operatorname{dom} f_n^* \to \operatorname{int}\operatorname{dom} f_n$ d'après le Lemme 7.5, $\nabla f_n(x_n) + \lambda_n(\nabla f_n(T_{\mathrm{i}(n),n}x_n) - \nabla f_n(x_n)) \in \operatorname{int}\operatorname{dom} f_n^*$, et il résulte de (7.59) que $x_{n+1} \in \operatorname{int}\operatorname{dom} f_n = \operatorname{int}\operatorname{dom} f$. En raisonnant par récurrence,

$$(x_n)_{n\in\mathbb{N}} \in (\operatorname{int}\operatorname{dom} f)^{\mathbb{N}} \quad \text{est bien définie.} \tag{7.61}$$

Ensuite, on déduit de (7.59) et du Lemme 7.6 que

$$(\forall x \in C \cap \operatorname{dom} f)(\forall n \in \mathbb{N}) \quad D^{f_n}(x, x_{n+1}) \leqslant (1-\lambda_n) D^{f_n}(x, x_n) + \lambda_n D^{f_n}(x, T_{\mathrm{i}(n),n}x_n). \tag{7.62}$$

Comme (7.58) implique que

$$(\forall x \in C \cap \operatorname{dom} f)(\forall n \in \mathbb{N}) \quad \left\langle x - T_{\mathrm{i}(n),n}x_n, \nabla f_n(x_n) - \nabla f_n\big(T_{\mathrm{i}(n),n}x_n\big)\right\rangle, \tag{7.63}$$

on obtient

$$(\forall x \in C \cap \operatorname{dom} f)(\forall n \in \mathbb{N}) \quad D^{f_n}(x, T_{\mathrm{i}(n),n}x_n) \leqslant D^{f_n}(x, x_n) - D^{f_n}(T_{\mathrm{i}(n),n}x_n, x_n). \tag{7.64}$$

D'autre part, il résulte de (7.56) que

$$(\forall x \in C \cap \operatorname{dom} f)(\forall n \in \mathbb{N}) \quad D^{f_{n+1}}(x, x_{n+1}) \leqslant (1+\eta_n) D^{f_n}(x, x_{n+1}), \tag{7.65}$$

et donc, en combinant avec (7.62) et (7.64), on obtient

$$\begin{aligned}(\forall x \in C \cap \operatorname{dom} f)(\forall n \in \mathbb{N}) \quad D^{f_{n+1}}(x, x_{n+1}) &\leqslant (1+\eta_n) D^{f_n}(x, x_n) \\ &\quad - (1+\eta_n)\lambda_n D^{f_n}(T_{\mathrm{i}(n),n}x_n, x_n).\end{aligned} \tag{7.66}$$

Cela implique que $(x_n)_{n\in\mathbb{N}}$ est stationnairement quasi-monotone au sens de Bregman par rapport à C relativement à $(f_n)_{n\in\mathbb{N}}$. Donc, il résulte de la Proposition 5.5(i) que

$$(\forall x \in C \cap \operatorname{int}\operatorname{dom} f) \quad \big(D^{f_n}(x, x_n)\big)_{n\in\mathbb{N}} \quad \text{est convergente.} \tag{7.67}$$

(i) : On note que $(\forall n \in \mathbb{N})$ $D^f(x, x_n) \leqslant \alpha^{-1} D^{f_n}(x, x_n)$. Par suite, on déduit que (7.67) que $(D^f(x, x_n))_{n\in\mathbb{N}}$ est bornée.

(ii) : Proposition 5.5(ii).

(iii) : Soit $x \in C \cap \operatorname{int}\operatorname{dom} f$. Il résulte de (7.66) et de (7.67) que

$$\begin{aligned}\sum_{n\in\mathbb{N}} D^f(T_{\mathrm{i}(n),n}x_n, x_n) &\leqslant \alpha^{-1} \sum_{n\in\mathbb{N}} D^{f_n}(T_{\mathrm{i}(n),n}x_n, x_n) \\ &\leqslant \alpha^{-1}\varepsilon^{-1} \sum_{n\in\mathbb{N}} (1+\eta_n)\lambda_n D^{f_n}(T_{\mathrm{i}(n),n}x_n, x_n) \\ &\leqslant \alpha^{-1}\varepsilon^{-1} \sum_{n\in\mathbb{N}} \big((1+\eta_n) D^{f_n}(x, x_n) - D^{f_{n+1}}(x, x_{n+1})\big) \\ &\leqslant \alpha^{-1}\varepsilon^{-1} \sup_{n\in\mathbb{N}} D^{f_n}(x, x_n)\Big(1 + \sum_{n\in\mathbb{N}} \eta_n\Big) \\ &< +\infty.\end{aligned} \tag{7.68}$$

(iv) : Soit $x \in C \cap \operatorname{int}\operatorname{dom} f$. On déduit de (7.64) et de (7.67) que $(D^{f_n}(x, T_{\mathrm{i}(n),n}x_n))_{n\in\mathbb{N}}$ est bornée, et donc, comme

$$(\forall n \in \mathbb{N}) \quad D^f(x, T_{\mathrm{i}(n),n}x_n) \leqslant \alpha^{-1} D^{f_n}(x, T_{\mathrm{i}(n),n}x_n), \tag{7.69}$$

on déduit que $(D^f(x, T_{\mathrm{i}(n),n}x_n))_{n\in\mathbb{N}}$ est bornée. Par suite, $(T_{\mathrm{i}(n),n}x_n)_{n\in\mathbb{N}}$ est bornée. D'autre part, on déduit de (iii) que $D^f(T_{\mathrm{i}(n),n}x_n, x_n) \to 0$, et comme f satisfait la Condition 7.7,

$$T_{\mathrm{i}(n),n}x_n - x_n \to 0. \tag{7.70}$$

Comme (7.59) implique que

$$(\forall n \in \mathbb{N}) \quad \nabla f_n(x_{n+1}) - \nabla f_n(T_{\mathrm{i}(n),n}x_n) = (1-\lambda_n)\big(\nabla f_n(x_n) - \nabla f_n(T_{\mathrm{i}(n),n}x_n)\big), \tag{7.71}$$

on obtient

$$\begin{aligned}
(\forall n \in \mathbb{N}) \quad & D^{f_n}\big(T_{\mathrm{i}(n),n}x_n, x_{n+1}\big) + D^{f_n}\big(x_{n+1}, T_{\mathrm{i}(n),n}x_n\big) \\
&= \big\langle x_{n+1} - T_{\mathrm{i}(n),n}x_n, \nabla f_n(x_{n+1}) - \nabla f_n(T_{\mathrm{i}(n),n}x_n)\big\rangle \\
&= (1-\lambda_n)\big\langle x_{n+1} - T_{\mathrm{i}(n),n}x_n, \nabla f_n(x_n) - \nabla f_n(T_{\mathrm{i}(n),n}x_n)\big\rangle \\
&= (1-\lambda_n)\Big(D^{f_n}\big(x_{n+1}, T_{\mathrm{i}(n),n}x_n\big) + D^{f_n}\big(T_{\mathrm{i}(n),n}x_n, x_n\big) - D^{f_n}\big(x_{n+1}, x_n\big)\Big) \\
&\leqslant D^{f_n}\big(x_{n+1}, T_{\mathrm{i}(n),n}x_n\big) + D^{f_n}\big(T_{\mathrm{i}(n),n}x_n, x_n\big).
\end{aligned} \tag{7.72}$$

Ainsi

$$(\forall n \in \mathbb{N}) \quad D^f(T_{\mathrm{i}(n),n}x_n, x_{n+1}) \leqslant \alpha^{-1}D^{f_n}\big(T_{\mathrm{i}(n),n}x_n, x_{n+1}\big) \leqslant \alpha^{-1}D^{f_n}\big(T_{\mathrm{i}(n),n}x_n, x_n\big). \tag{7.73}$$

Comme (7.68) implique que $D^{f_n}(T_{\mathrm{i}(n),n}x_n, x_{n+1}) \to 0$, on déduit de (7.73) que $D^f(T_{\mathrm{i}(n),n}x_n, x_{n+1}) \to 0$. Par suite, $T_{\mathrm{i}(n),n}x_n - x_{n+1} \to 0$ et il résulte de (7.70) que

$$x_{n+1} - x_n \to 0. \tag{7.74}$$

Maintenant, soient $x \in \mathfrak{W}(x_n)_{n\in\mathbb{N}}$, $(k_n)_{n\in\mathbb{N}}$ une suite strictement croissante dans $\mathbb{N}$ telle que $x_{k_n} \rightharpoonup x$, et $j \in I$. D'après (7.57), il existe une suite strictement croissante $(l_n)_{n\in\mathbb{N}}$ dans $\mathbb{N}$ telle que

$$(\forall n \in \mathbb{N}) \quad \begin{cases} k_n \leqslant l_n \leqslant k_n + M_j - 1 < k_{n+1} \leqslant l_{n+1}, \\ j = \mathrm{i}(l_n). \end{cases} \tag{7.75}$$

Par conséquent,

$$\|x_{l_n} - x_{k_n}\| \leqslant \sum_{m=k_n}^{k_n+M_j-2} \|x_{m+1} - x_n\| \leqslant (M_j - 1) \max_{k_n \leqslant m \leqslant k_n+M_j-2} \|x_{m+1} - x_m\| \to 0, \tag{7.76}$$

et donc,

$$x_{l_n} \rightharpoonup x. \tag{7.77}$$

Par ailleurs, on dérive de (7.70) et (7.75) que

$$T_{j,l_n}x_{l_n} - x_{l_n} = T_{\mathrm{i}(l_n),l_n}x_{l_n} - x_{l_n} \to 0. \tag{7.78}$$

En invoquant (7.77) et (7.78), on obtient $x \in C_j$ et donc, $x \in C$, i.e., $\mathfrak{W}(x_n)_{n\in\mathbb{N}} \subset C$. Finalement, l'assertion résulte de la Proposition 5.6.

(v) : C'est un corollaire de la Proposition 5.8. ☐

Remarque 7.26 La Condition (7.57) est introduite dans [7, Definition 5]. Dans le cas où $I = \{1, \ldots, m\}$, on peut choisir $\mathrm{i} = 1 + \mathrm{rem}(\cdot, m)$ où $\mathrm{rem}(\cdot, m)$ est le reste de la division par m.

Remarque 7.27 Dans la Proposition 7.25, supposons que $I = \{1, \ldots, m\}$, que $f \in \Gamma_0(\mathcal{X})$ soit de Legendre et uniformément convexe, et que $(\forall n \in \mathbb{N})\ f_n = f$. Alors, on récupère [3, Section 4.2].

Remarque 7.28 Dans la Proposition 7.25, supposons que $I = \{1, \ldots, m\}$, que $f \colon \mathcal{X} \to \mathbb{R}$ soit une fonction de Legendre uniformément convexe dont le gradient est faiblement séquentiellement continu, et que $(\forall n \in \mathbb{N})\ f_n = f$. Alors, on récupère [13, Theorem 3.1].

Corollaire 7.29 *Soient $\mathcal{X}$ un espace de Banach réel réflexif, $f \in \Gamma_0(\mathcal{X})$ une fonction de Legendre, $(T_n)_{n\in\mathbb{N}}$ des opérateurs dans $\widehat{\mathfrak{B}}(f)$ telle que $C = \bigcap_{n\in\mathbb{N}} \widehat{\operatorname{Fix}}\, T_n \neq \varnothing$, $\varepsilon \in \left]0, 1\right[$, $(\lambda_n)_{n\in\mathbb{N}}$ une suite dans $\left]\varepsilon, 1\right]$, $x_0 \in \operatorname{int}\operatorname{dom} f$, et posons*

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = \nabla f^*\big((1 - \lambda_n)\nabla f(x_n) + \lambda_n \nabla f(T_n x_n)\big). \tag{7.79}$$

Supposons que $C \cap \operatorname{int}\operatorname{dom} f \neq \varnothing$, que f satisfasse la Condition 7.7, et que $(\forall x \in \operatorname{int}\operatorname{dom} f)$ $D^f(x, \cdot)$ soit coercive. Alors les propriétés suivantes sont satisfaites :

(i) *$(x_n)_{n\in\mathbb{N}}$ est une suite bornée dans $\operatorname{int}\operatorname{dom} f$.*

(ii) $\sum_{n\in\mathbb{N}} D^f(T_n x_n, x_n) < +\infty$.

(iii) *Supposons que, pour tout $y_1 \in \mathcal{X}$ et pour tout $y_2 \in \mathcal{X}$,*

$$\begin{cases} y_1 \in \mathfrak{W}(x_n)_{n\in\mathbb{N}} \cap C \\ y_2 \in \mathfrak{W}(x_n)_{n\in\mathbb{N}} \cap C \\ \big(\langle y_1 - y_2, \nabla f(x_n)\rangle\big)_{n\in\mathbb{N}} \quad \textit{converge} \end{cases} \quad \Rightarrow \quad y_1 = y_2. \tag{7.80}$$

De plus, supposons que $\mathfrak{W}(x_n)_{n\in\mathbb{N}} \subset \operatorname{int}\operatorname{dom} f$. Alors il existe $\overline{x} \in C$ tel que $x_n \rightharpoonup \overline{x}$.

(iv) *Supposons que C soit convexe et fermé et que $\underline{\lim}\, D^f_C(x_n) = 0$. Alors il existe $\overline{x} \in C$ tel que $x_n \to \overline{x}$.*

Démonstration. C'est un corollaire de la Proposition 7.25 avec $I = \{1\}$ et $(\forall n \in \mathbb{N})$ $f_n = f$. ☐

7.5 Algorithmes d'éclatement d'opérateurs monotones

Dans cette section, nous considérons le problème suivant.

Problème 7.30 Soient $\mathcal{X}$ un espace de Banach réel réflexif, $\mathcal{X}^*$ le dual topologique de $\mathcal{X}$, et $A \colon \mathcal{X} \to 2^{\mathcal{X}^*}$ et $B \colon \mathcal{X} \to 2^{\mathcal{X}^*}$ des opérateurs maximalement monotones tels que $Z = \operatorname{zer}(A + B) \neq \varnothing$. Le problème est de

$$\text{trouver} \quad x \in \mathcal{X} \quad \text{tel que} \quad 0 \in Ax + Bx. \tag{7.81}$$

Dans la proposition suivante, nous utilisons les notions dans les Définitions 5.2 et 5.3 du Chapitre 5.

Proposition 7.31 *Dans le Problème 7.30, supposons que* $B = \nabla g$ *où* $g \in \Gamma_0(\mathcal{X})$ *est une fonction essentiellement lisse. Soit* $f \in \Gamma_0(\mathcal{X})$ *une fonction de Legendre cofinie telle que* $Z \cap \operatorname{int} \operatorname{dom} f \neq \varnothing$, $\operatorname{int} \operatorname{dom} f \subset \operatorname{int} \operatorname{dom} g$, *et* $f \succcurlyeq \alpha g$ *pour un certain* $\beta \in \left]0, +\infty\right[$. *Soient* $\varepsilon \in \left]0, \beta/(\beta+1)\right[$, $(\eta_n)_{n\in\mathbb{N}} \in \ell_+^1(\mathbb{N})$, *et* $(\gamma_n)_{n\in\mathbb{N}}$ *une suite dans* $\mathbb{R}$ *telle que*

$$(\forall n \in \mathbb{N}) \quad \varepsilon \leqslant \gamma_n \leqslant \beta(1-\varepsilon) \quad \text{et} \quad (1+\eta_n)\gamma_n - \gamma_{n+1} \leqslant \beta\eta_n. \tag{7.82}$$

Soit $x_0 \in \operatorname{int} \operatorname{dom} f$ *et posons*

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = J^f_{\gamma_n A}\left(\nabla f(x_n) - \gamma_n \nabla g(x_n)\right). \tag{7.83}$$

Supposons que f *satisfasse la Condition 7.7 et que* $(\forall x \in \operatorname{int} \operatorname{dom} f)$ $D^f(x, \cdot)$ *soit coercive. Alors les propriétés suivantes sont satisfaites :*

(i) $(x_n)_{n\in\mathbb{N}}$ *est une suite bornée dans* $\operatorname{int} \operatorname{dom} f$.

(ii) *Supposons que* ∇f *soit uniformément continu sur les sous-ensembles bornés de* $\operatorname{int} \operatorname{dom} f$ *et qu'une des condition suivantes soit vérifiée :*

  (a) $Z \cap \overline{\operatorname{dom}} f$ *est un singleton.*

  (b) ∇f *et* ∇g *sont faiblement séquentiellement continus et* $Z \subset \operatorname{int} \operatorname{dom} f$.

  (c) $\operatorname{dom} f^*$ *est ouvert et* ∇f, ∇f^*, *et* ∇g *sont faiblement séquentiellement continus.*

  Alors il existe $\overline{x} \in Z$ *tel que* $x_n \rightharpoonup \overline{x}$.

(iii) *Supposons qu'une des conditions suivantes soit vérifiée :*

  (a) A *est uniformément monotone en* $\overline{x}$.

  (b) g *est uniformément convexe en* $\overline{x}$.

  (c) $\varliminf D^f_Z(x_n) = 0$.

  Alors il existe $\overline{x} \in Z$ *tel que* $x_n \to \overline{x}$.

Démonstration. Posons $M = A + \nabla g$. Comme g est essentiellement lisse, ∇g est maximalement monotone [16, Theorem 18.7] et [2, Theorem 5.6]. Puisque $\varnothing \neq Z \cap \operatorname{int} \operatorname{dom} f \subset \operatorname{dom} A \cap \operatorname{int} \operatorname{dom} f \subset \operatorname{dom} A \cap \operatorname{int} \operatorname{dom} g$, M est maximalement monotone d'après [16, Theorem 24.1(a)] et donc, $Z = \operatorname{zer} M$ est convexe et fermé [6, Lemma 1.1(a)]. Ensuite, le Lemme 7.15 implique que les opérateurs $(J^f_{\gamma_n A})_{n\in\mathbb{N}}$ sont bien définis et univoques de $\mathcal{X}^*$ dans $\operatorname{int} \operatorname{dom} f$. Ainsi, on déduit de (7.83) que $(x_n)_{n\in\mathbb{N}}$ est une suite bien définie dans $\operatorname{int} \operatorname{dom} f$. Ensuite, posons

$$(\forall n \in \mathbb{N}) \quad f_n \colon \mathcal{X} \to \left]-\infty, +\infty\right] \\ x \mapsto \begin{cases} f(x) - \gamma_n g(x), & \text{si } x \in \operatorname{int} \operatorname{dom} f; \\ +\infty, & \text{sinon.} \end{cases} \tag{7.84}$$

Notons que

$$(\forall n\in\mathbb{N})\quad (1+\eta_n)f_n-f_{n+1}=(1+\eta_n)(f-\gamma_n g)-(f-\gamma_{n+1}g)\\ =\eta_n(f-\beta g)+\big(\beta\eta_n+\gamma_{n+1}-(1+\eta_n)\gamma_n\big)g. \tag{7.85}$$

et

$$(\forall n\in\mathbb{N})\quad f_n-\varepsilon f=f-\gamma_n g-\varepsilon f=(1-\varepsilon)(f-\beta g)+\big(\beta(1-\varepsilon)-\gamma_n\big)g. \tag{7.86}$$

Par suite, en utilisant la notation de la Définition 5.2, on déduit de (7.82) que

$$(\forall n\in\mathbb{N})\quad (1+\eta_n)f_n\succcurlyeq f_{n+1}\quad\text{et}\quad f_n\in\mathcal{P}_\varepsilon(f). \tag{7.87}$$

D'autre part, on déduit de (7.83) que

$$(\forall n\in\mathbb{N})\quad \gamma_n Mx_{n+1}=\gamma_n\big(Ax_{n+1}+\nabla g(x_{n+1})\big)\\ \ni\nabla f(x_n)-\gamma_n\nabla g(x_n)-\nabla f(x_{n+1})+\gamma_n\nabla g(x_{n+1})\\ =\nabla f_n(x_n)-\nabla f_n(x_{n+1}). \tag{7.88}$$

Maintenant, (7.88) et la monotonie de M impliquent que

$$(\forall x\in Z\cap\operatorname{dom}f)(\forall n\in\mathbb{N})\quad \langle x-x_{n+1},\nabla f_n(x_n)-\nabla f_n(x_{n+1})\rangle\leqslant 0, \tag{7.89}$$

et ainsi

$$(\forall x\in Z\cap\operatorname{dom}f)(\forall n\in\mathbb{N})\quad D^{f_n}(x,x_{n+1})+D^{f_n}(x_{n+1},x_n)\leqslant D^{f_n}(x,x_n). \tag{7.90}$$

Par ailleurs, on déduit de (7.87) que

$$(\forall x\in Z\cap\operatorname{dom}f)(\forall n\in\mathbb{N})\quad (1+\eta_n)D^{f_n}(x,x_{n+1})\geqslant D^{f_{n+1}}(x,x_{n+1}). \tag{7.91}$$

Ainsi, on déduit de (7.90) que

$$(\forall x\in Z\cap\operatorname{dom}f)(\forall n\in\mathbb{N})\quad D^{f_{n+1}}(x,x_{n+1})\leqslant(1+\eta_n)D^{f_n}(x,x_n)-(1+\eta_n)D^{f_n}(x_{n+1},x_n). \tag{7.92}$$

Cela signifie que $(x_n)_{n\in\mathbb{N}}$ est stationnairement quasi-monotone au sens de Bregman par rapport à Z relativement à $(f_n)_{n\in\mathbb{N}}$ d'après la Définition 5.3. Donc, il résulte de la Proposition 5.5(i) que

$$(\forall x\in Z\cap\operatorname{int}\operatorname{dom}f)\quad \big(D^{f_n}(x,x_n)\big)_{n\in\mathbb{N}}\quad\text{est convergente.} \tag{7.93}$$

(i) : Proposition 5.5(ii).

(ii) : Comme (7.92) implique que

$$(\forall x\in Z\cap\operatorname{int}\operatorname{dom}f)(\forall n\in\mathbb{N})\quad D^{f_n}(x_{n+1},x_n)\leqslant(1+\eta_n)D^{f_n}(x_{n+1},x_n)\\ \leqslant(1+\eta_n)D^{f_n}(x,x_n)-D^{f_{n+1}}(x,x_{n+1}), \tag{7.94}$$

on déduit de (7.93) que $D^{f_n}(x_{n+1},x_n)\to 0$, et ainsi

$$D^f(x_{n+1},x_n)\to 0. \tag{7.95}$$

Comme f satisfait la Condition 7.7, on déduit de (7.95) que

$$x_{n+1}-x_n\to 0. \tag{7.96}$$

Maintenant, on prend $x\in\mathfrak{W}(x_n)_{n\in\mathbb{N}}$ et $(k_n)_{n\in\mathbb{N}}$ une suite strictement croissante dans $\mathbb{N}$ telle que $x_{k_n}\rightharpoonup x$. Posons

$$(\forall n\in\mathbb{N})\quad x_n^*=\gamma_n^{-1}(\nabla f(x_{n+1})-\nabla f(x_n))\in Mx_{n+1}. \tag{7.97}$$

Alors $x_{k_n+1}\rightharpoonup x$. De plus, comme ∇f est uniformément continu sur les sous-ensembles bornés de $\operatorname{int}\operatorname{dom} f$, on déduit de (7.96) que $x_{k_n}^*\to 0$. Puisque $\operatorname{gra} M$ est séquentiellement fermé dans $\mathcal{X}^{\text{faible}}\times\mathcal{X}^{*\text{fort}}$ [6, Lemma 1.2], $(x,0)\in\operatorname{gra} M$, i.e., $x\in\operatorname{zer} M=Z$.

(ii)(a) : On a $\varnothing\neq\mathfrak{W}(x_n)_{n\in\mathbb{N}}\subset Z\cap\overline{\operatorname{dom}}\, f$ et donc, $\mathfrak{W}(x_n)_{n\in\mathbb{N}}=\{\overline{x}\}$. Par suite, $x_n\rightharpoonup\overline{x}$.

(ii)(b)-(ii)(c) : Si $Z\subset\operatorname{int}\operatorname{dom} f$ alors $\mathfrak{W}(x_n)_{n\in\mathbb{N}}\subset Z\cap\operatorname{int}\operatorname{dom} f$. Si $\operatorname{dom} f^*$ est ouvert et ∇f^* est faiblement séquentiellement continu alors $\mathfrak{W}(x_n)_{n\in\mathbb{N}}\subset\operatorname{int}\operatorname{dom} f$ d'après le Lemme 6.25. Au vu de la Proposition 5.6, il suffit de montrer que (5.10) est satisfaite. Supposons que $y_1\in\mathfrak{W}(x_n)_{n\in\mathbb{N}}\cap Z$ et $y_2\in\mathfrak{W}(x_n)_{n\in\mathbb{N}}\cap Z$ tels que

$$\big\langle y_1-y_2,\nabla f_n(x_n)\big\rangle\to l. \tag{7.98}$$

Alors il existe des suites strictement croissantes $(k_n)_{n\in\mathbb{N}}$ et $(l_n)_{n\in\mathbb{N}}$ dans $\mathbb{N}$ telles que $x_{k_n}\rightharpoonup y_1$ et $x_{l_n}\rightharpoonup y_2$. On alors déduit de (7.82) et de [15, Lemma 2.2.2] qu'il existe $\theta\in[\varepsilon,\beta(1-\varepsilon)]$ tel que $\gamma_n\to\theta$. Comme ∇f et ∇g sont faiblement séquentiellement continus et comme $(\forall n\in\mathbb{N})$ $f_n=f-\gamma_n g$, en passant à la limite dans (7.98) par rapport aux sous-suites $(x_{k_n})_{n\in\mathbb{N}}$ et $(x_{l_n})_{n\in\mathbb{N}}$, on obtient

$$l=\big\langle y_1-y_2,\nabla f(y_1)-\theta g(y_1)\big\rangle=\big\langle y_1-y_2,\nabla f(y_2)-\theta g(y_2)\big\rangle. \tag{7.99}$$

Ensuite, posons

$$\begin{aligned} h\colon \mathcal{X}&\to\,]-\infty,+\infty]\\ x&\mapsto\begin{cases} f(x)-\theta g(x), & \text{si } x\in\operatorname{int}\operatorname{dom} f;\\ +\infty, & \text{sinon.}\end{cases}\end{aligned} \tag{7.100}$$

Alors h est Gâteaux différentiable sur $\operatorname{int}\operatorname{dom} h=\operatorname{int}\operatorname{dom} f$ et (7.99) implique que

$$\big\langle y_1-y_2,\nabla h(y_1)-\nabla h(y_2)\big\rangle=0. \tag{7.101}$$

D'autre part,

$$h-\varepsilon f=f-\theta g-\varepsilon f=(1-\varepsilon)(f-\beta g)+\big(\beta(1-\varepsilon)-\theta\big)g. \tag{7.102}$$

Ainsi, comme $f \succcurlyeq \beta g$ et $\theta \leqslant \beta(1-\varepsilon)$, on obtient $h \succcurlyeq \varepsilon f$ et donc,

$$D^h(y_1,y_2) \geqslant \varepsilon D^f(y_1,y_2) \quad \text{et} \quad D^h(y_2,y_1) \geqslant \varepsilon D^f(y_2,y_1). \tag{7.103}$$

Par suite, il résulte de (7.101) que

$$\begin{aligned} 0 &= \big\langle y_1 - y_2, \nabla h(y_1) - \nabla h(y_2)\big\rangle \\ &= D^h(y_1,y_2) + D^h(y_2,y_1) \\ &\geqslant \varepsilon\big(D^f(y_1,y_2) + D^f(y_2,y_1)\big) \\ &= \varepsilon\big\langle y_1 - y_2, \nabla f(y_1) - \nabla f(y_2)\big\rangle, \end{aligned} \tag{7.104}$$

et la monotonie de ∇f implique que

$$\big\langle y_1 - y_2, \nabla f(y_1) - \nabla f(y_2)\big\rangle = 0. \tag{7.105}$$

Comme $f|_{\operatorname{int dom} f}$ est strictement convexe, ∇f est strictement monotone [17, Theorem 2.4.4(ii)], et ainsi (7.105) implique que $y_1 = y_2$.

(iii)(a)-(iii)(b) : Soit $\overline{x} \in Z \cap \operatorname{int dom} f$. Si A est uniformément monotone en $\overline{x}$ alors M l'est aussi. Si g est uniformément convexe en $\overline{x}$ alors ∂g est uniformément monotone en $\overline{x}$ [17, Page 201] et ainsi M l'est aussi. Soit ϕ le module de convexité uniforme de M en $\overline{x}$. Il résulte de (7.97) que

$$\langle x_{n+1} - \overline{x}, x_n^*\rangle \geqslant \phi\big(\|x_{n+1} - \overline{x}\|\big). \tag{7.106}$$

Comme (7.90) implique que

$$(\forall n \in \mathbb{N}) \quad D^{f_{n+1}}(\overline{x}, x_{n+1}) \leqslant (1+\eta_n) D^{f_n}(\overline{x}, x_{n+1}) \leqslant (1+\eta_n) D^{f_n}(\overline{x}, x_n), \tag{7.107}$$

on obtient

$$(\forall n \in \mathbb{N}) \quad (1+\eta_n)^{-1} D^{f_{n+1}}(\overline{x}, x_{n+1}) \leqslant D^{f_n}(\overline{x}, x_{n+1}) \leqslant D^{f_n}(\overline{x}, x_n). \tag{7.108}$$

Par suite,

$$D^{f_n}(\overline{x}, x_{n+1}) - D^{f_n}(\overline{x}, x_n) \to 0. \tag{7.109}$$

Donc, il résulte de (7.106) que

$$\begin{aligned} (\forall n \in \mathbb{N}) \quad \varepsilon\phi\big(\|x_{n+1} - \overline{x}\|\big) &\leqslant \gamma_n \,\langle x_{n+1} - \overline{x}, x_n^*\rangle \\ &= \big\langle x_{n+1} - \overline{x}, \nabla f_n(x_n) - \nabla f_n(x_{n+1})\big\rangle \\ &= D^{f_n}(\overline{x}, x_n) - D^{f_n}(\overline{x}, x_{n+1}) - D^{f_n}(x_{n+1}, x_n) \\ &\leqslant D^{f_n}(\overline{x}, x_n) - D^{f_n}(\overline{x}, x_{n+1}). \end{aligned} \tag{7.110}$$

En passant à la limite dans (7.110) et en utilisant (7.109), on obtient

$$\phi\big(\|x_{n+1} - \overline{x}\|\big) \to 0. \tag{7.111}$$

Par suite, $\|x_{n+1} - \overline{x}\| \to 0$ et ainsi $x_n \to \overline{x}$ d'après (7.96).

(iii)(c) : Proposition 5.8. ☐

7.6 Bibliographie

[1] H. Attouch, *Variational Convergence for Functions and Operators.* Pitman, 1986.

[2] H. H. Bauschke, J. M. Borwein, and P. L. Combettes, Essential smoothness, essential strict convexity and Legendre functions in Banach spaces, *Commun. Contemp. Math.,* vol. 3, pp. 615–647, 2001.

[3] H. H. Bauschke, J. M. Borwein, and P. L. Combettes, Bregman monotone optimization algorithms, *SIAM J. Control Optim.,* vol. 42, pp. 596–636, 2003.

[4] H. H. Bauschke and P. L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces.* Springer, New York, 2011.

[5] Y. Benyanmini and J. Lindenstrauss, *Geometric Nonlinear Functional Analysis. Vol. 1.* American Mathematical Society, Providence, RI, 2000.

[6] F. E. Browder, Multi-valued monotone nonlinear mappings and duality mappings in Banach spaces, *Trans. Amer. Math. Soc.,* vol. 118, pp. 338–351, 1965.

[7] F. E. Browder, Convergence theorems for sequences of nonlinear operators in Banach spaces, *Math. Z.,* vol. 100, pp. 201–225, 1967.

[8] R. S. Burachik and S. Scheimberg, A proximal point method for the variational inequality problem in Banach spaces, *SIAM J. Control Optim.,* vol. 39, pp. 1633–1649, 2001.

[9] D. Butnariu and G. Kassay, A proximal-projection method for finding zeros of set-valued operator, *SIAM J. Control Optim.,* vol. 47, pp. 1096–2131, 2008.

[10] I. Cioranescu, *Geometry of Banach Spaces, Duality Mappings and Nonlinear Problems.* Kluwer Academic Publishers Group, Dordrecht, 1990.

[11] P. L. Combettes, Solving monotone inclusions via compositions of nonexpansive averaged operators, *Optimization,* vol. 53, pp. 475–504, 2004.

[12] K. Goebel and S. Reich, *Uniform Convexity, Hyperbolic Geometry, and Nonexpansive Mappings.* Marcel Dekker, Inc., New York, 1984.

[13] M. B. Lee and S. H. Park, Convergence of sequential parafirmly nonexpansive mappings in reflexive Banach spaces, *J. Optim. Theory Appl.,* vol. 123, pp. 549–571, 2004.

[14] V. Martín-Márquez, S. Reich, and S. Sabach, Iterative methods for approximating fixed points of Bregman nonexpansive operators, *Discrete Contin. Dyn. Syst. Ser. S,* vol. 6, pp. 1043–1063, 2013.

[15] B. T. Polyak, *Introduction to Optimization.* Optimization Software Inc., New York, 1987.

[16] S. Simons, *From Hahn-Banach to Monotonicity,* Lecture Notes in Math. 1693, Springer-Verlag, New York, 2008.

[17] C. Zălinescu, *Convex Analysis in General Vector Spaces.* World Scientific Publishing, River Edge, NJ, 2002.

Chapitre 8

Conclusions et perspectives

8.1 Conclusions

Nous avons proposé les premières méthodes pour résoudre des inclusions monotones composites dans les espaces de Banach. De plus, nous avons analysé des cas particuliers d'un algorithme de type d'Haugazeau/Bregman en présence d'erreurs.

Nous avons introduit la notion de suites quasi-monotones au sens de Bregman qui nous permet de proposer et de démontrer la convergence d'un algorithme du point proximal avec des distances de Bregman variables. Nous l'avons appliqué pour résoudre le problème d'admissibilité convexe.

Nous avons proposé la première version de l'algorithme explicite-implicite pour résoudre le problème de minimisation convexe composite dans des espaces de Banach réels réflexifs. Même dans des espaces hilbertiens et euclidiens, notre cadre donne des algorithmes nouveaux qui, dans certains cas, s'avèrent plus avantageux numériquement que les algorithmes basés sur la distance hilbertienne. Nous avons appliqué les résultats obtenus à divers schémas itératifs de construction de zéros et d'optimisation. Une application en restauration d'images a été présentée pour illustrer notre algorithme.

8.2 Perspectives

Voici quelques problèmes qui sont naturellement suggérés par la thèse.

- Pour montrer la convergence de l'algorithme explicite-implicite, nous l'avons décrit sous la forme d'un algorithme du point proximal par rapport aux distances de Bregman variables et donc, nous avons demandé que $\nabla\psi$ soit faiblement séquentiellement continu. Une approche qui pourrait relâcher cette hypothèse dans un espace de dimension infinie serait intéressante.

- Une construction d'un algorithme explicite-implicite pour résoudre des inclusions monotones générales dans des espaces de Banach est encore une question ouverte.
- Des méthodes d'éclatement basées sur la distance hilbertienne ont été appliquées aux problèmes d'optimisation stochastique [4, 5, 6]. Les extensions de nos résultats dans ces problèmes seraient intéressantes.
- Soit $M\colon \mathcal{X} \to 2^{\mathcal{X}^*}$ un opérateur maximalement monotone tel que zer $M \neq \varnothing$ et soit $f \in \Gamma_0(\mathcal{X})$ une fonction de Legendre. Il est montré dans [2] que sous des hypothèses convenables sur f, la suite engendrée par l'algorithme

$$x_0 \in \operatorname{int}\operatorname{dom} f \quad \text{et} \quad (\forall n \in \mathbb{N}) \quad x_{n+1} = Q^f\big(x_0, x_n, R^f_{\gamma_n M} x_n\big) \tag{8.1}$$

converge fortement vers $P^f_{\operatorname{zer} M} x_0$. Cependant, le calcul de la D^f-résolvante en général donne un résultat avec une erreur et donc, le comportement de la variante ci-après de (8.1)

$$x_0 \in \operatorname{int}\operatorname{dom} f \quad \text{et} \quad (\forall n \in \mathbb{N}) \quad x_{n+1} = Q^f\big(x_0, x_n, R^f_{\gamma_n M} x_n + a_n\big), \tag{8.2}$$

où $(a_n)_{n\in\mathbb{N}}$ représente les erreurs produites à la suite du calcul de la D^f-résolvante, est intéressant. Lorsque $P^f_{\operatorname{zer} M} x_0$ reste dans $H^f(x_n, R^f_{\gamma_n M} x_n + a_n)$ à chaque itérée $n \in \mathbb{N}$, alors $D^f(P^f_{\operatorname{zer} M} x_0, x_n) \to 0$. Plus général, cette convergence est préservée si $(\forall n \in \mathbb{N})$ $P^f_{\operatorname{zer} M} x_0 + b_n \in H^f(x_0, x_n)$ où $b_n \to 0$. Néanmoins, cette condition est difficile à vérifier en pratique. Le comportement de $(x_n)_{n\in\mathbb{N}}$ dans le cas général est encore une question ouverte.
- En considérant une inclusion primale-duale composite comme une somme de deux opérateurs monotones dans l'espace vectoriel produit et en utilisant les méthodes existantes (la méthode explicite-implicite-explicite [3, 8] et la méthode de l'inverse partiel [7]), les méthodes nouvelles sont obtenues [1, 3]. Une généralisation de l'inverse partiel dans les espaces de Banach permet d'étendre le résultat de [1] pour des inclusions monotones composites dans ces espaces. Cependant, l'existence d'une telle généralisation est une question ouverte.
- Les algorithmes opérants par blocs de coordonnées [4] fournissent un outil pour résoudre des problèmes d'optimisation de grande taille. Leurs versions basées sur les distances de Bregman seraient intéressantes à développer.

Paris, le 17 juillet 2015

8.3 Bibliographie

[1] M. A. Alghamdi, A. Alotaibi, P. L. Combettes, and N. Shahzad, A primal-dual method of partial inverses for composite inclusions, *Optim. Lett.,* vol. 8, pp. 2271–2284, 2014.

[2] H. H. Bauschke and P. L. Combettes, Construction of best Bregman approximations in reflexive Banach spaces, *Proc. Amer. Math. Soc.,* vol. 131, pp. 3757–3766, 2003.

[3] L. M. Briceño-Arias and P. L. Combettes, A monotone+skew splitting model for composite monotone inclusions in duality, *SIAM J. Optim.*, vol. 21, pp. 1230–1250, 2011.

[4] P. L. Combettes and J. C. Pesquet, Stochastic quasi-Fejér block-coordinate fixed point iterations with random sweeping, *SIAM J. Optim.,* to appear.

[5] L. Rosasco, S. Villa, and B. C. Vũ, Convergence of stochastic proximal gradient algorithm, 2014.
http://arxiv.org/abs/1403.5074

[6] L. Rosasco, S. Villa, and B. C. Vũ, A stochastic forward-backward splitting method for solving monotone inclusions in Hilbert spaces, 2014.
http://arxiv.org/abs/1403.7999

[7] J. E. Spingarn, Partial inverse of a monotone operator, *Appl. Math. Optim.*, vol. 10, pp. 247–265, 1983.

[8] P. Tseng, A modified forward-backward splitting method for maximal monotone mappings, *SIAM J. Control Optim.,* vol. 38, pp. 431–446, 2000.