Variable Quasi-Bregman Monotone Sequences

Quang Van Nguyen

Sorbonne Universités – UPMC Univ. Paris 06
UMR 7598, Laboratoire Jacques-Louis Lions
F-75005 Paris, France
quangnv@ljll.math.upmc.fr

Abstract

We introduce a notion of variable quasi-Bregman monotone sequence which unifies the notion of variable metric quasi-Fejér monotone sequences and that of Bregman monotone sequences. The results are applied to analyze the asymptotic behavior of proximal iterations based on variable Bregman distance and of algorithms for solving convex feasibility problems in reflexive real Banach spaces.

Key words. Banach space, Bregman distance, Bregman projection, convex feasibility problem, Fejér monotone sequence, Legendre function, proximal iterations

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 1.1 [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

$$\partial f\colon \mathcal{X} \to 2^{\mathcal{X}^*} : x \mapsto \big\{x^* \in \mathcal{X}^* \;\big|\; (\forall y \in \mathcal{X})\, \langle y - x, x^*\rangle + f(x) \leqslant f(y)\big\}, \tag{1.1}$$

be Moreau subdifferential of f. The *Bregman distance* associated with f is

$$\begin{aligned} D^f\colon \mathcal{X}\times\mathcal{X} \to\; & [0,+\infty] \\ (x,y) \mapsto\; & \begin{cases} f(x) - f(y) - \langle x - y, \nabla f(y)\rangle, & \text{if } y \in \operatorname{int}\operatorname{dom} f; \\ +\infty, & \text{otherwise}. \end{cases} \end{aligned} \tag{1.2}$$

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

$$\begin{aligned} \operatorname{prox}^f_\varphi\colon \operatorname{int}\operatorname{dom} f &\to \operatorname{dom}\varphi \cap \operatorname{int}\operatorname{dom} f \\ y &\mapsto \underset{x\in\mathcal{X}}{\operatorname{argmin}}\, \varphi(x) + D^f(x,y). \end{aligned} \tag{1.3}$$

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

$$\begin{aligned} P^f_C\colon \operatorname{int}\operatorname{dom} f &\to C \cap \operatorname{int}\operatorname{dom} f \\ y &\mapsto \underset{x\in C}{\operatorname{argmin}}\, D^f(x,y), \end{aligned} \tag{1.4}$$

and the *D^f-distance* to C is the function

$$\begin{aligned} D^f_C\colon \mathcal{X} &\to [0,+\infty] \\ y &\mapsto \inf D^f(C,y). \end{aligned} \tag{1.5}$$

The paper is organized as follows. In Section 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 3. Applications to a variable Bregman proximal point algorithm and to the convex feasibility problem are considered in Section 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 $\operatorname{dom} M = \big\{x \in \mathcal{X} \;\big|\; Mx \neq \varnothing\big\}$, the range of M is $\operatorname{ran} M = \big\{y \in \mathcal{X} \;\big|\; (\exists\, x \in \mathcal{X})\; y \in Mx\big\}$, the graph of M is $\operatorname{gra} M = \big\{(x,y) \in \mathcal{X}\times\mathcal{X} \;\big|\; y \in Mx\big\}$, and the set of fixed points of M is $\operatorname{Fix} M = \big\{x \in \mathcal{X} \;\big|\; x \in Mx\big\}$. 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 $\operatorname{dom} f = \big\{x \in \mathcal{X} \;\big|\; f(x) < +\infty\big\} \neq \varnothing$. Let $f \in \Gamma_0(\mathcal{X})$. The set of global minimizers of a function f is denoted by $\operatorname{Argmin} f$. In addition, if f is Gâteaux differentiable on $\operatorname{int}\operatorname{dom} f \neq \varnothing$ then

$$\begin{aligned}\hat{f}\colon \mathcal{X} &\to \left]-\infty,+\infty\right]\\ x &\mapsto \begin{cases} f(x), & \text{if } x \in \operatorname{int}\operatorname{dom} f;\\ +\infty, & \text{otherwise.}\end{cases}\end{aligned} \tag{1.6}$$

Finally, $\ell^1_+(\mathbb{N})$ is the set of all summable sequences in $\left[0,+\infty\right[$.

2 Variable Bregman monotonicity

Definition 2.1 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

$$\mathcal{F}(f) = \big\{g \in \Gamma_0(\mathcal{X}) \;\big|\; g \text{ is Gâteaux differentiable on } \operatorname{dom} g = \operatorname{int}\operatorname{dom} f\big\}. \tag{2.1}$$

Moreover, if g_1 and g_2 are in $\mathcal{F}(f)$, then

$$g_1 \succcurlyeq g_2 \quad\Leftrightarrow\quad (\forall x \in \operatorname{dom} f)(\forall y \in \operatorname{int}\operatorname{dom} f) \quad D^{g_1}(x,y) \geqslant D^{g_2}(x,y). \tag{2.2}$$

For every $\alpha \in \left[0,+\infty\right[$, set

$$\mathcal{P}_\alpha(f) = \big\{g \in \mathcal{F}(f) \;\big|\; g \succcurlyeq \alpha f\big\}. \tag{2.3}$$

Remark 2.2 In Definition 2.1, suppose that $\mathcal{X}$ is a Hilbert space and let $\alpha \in \left]0,+\infty\right[$. 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

$$(\forall U_1 \in \mathcal{S}(\mathcal{X}))(\forall U_2 \in \mathcal{S}(\mathcal{X})) \quad U_1 \succcurlyeq U_2 \quad\Leftrightarrow\quad (\forall x \in \mathcal{X}) \quad \big\langle x, U_1 x\big\rangle \geqslant \big\langle x, U_2 x\big\rangle. \tag{2.4}$$

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

$$\begin{aligned}\|\cdot\|^2/2 \in \mathcal{P}_\alpha(f) &\Leftrightarrow\ (\forall x \in \mathcal{X})(\forall y \in \mathcal{X})\ \|x-y\|^2/2 \geqslant \alpha D^f(x,y)\\ &\Leftrightarrow\ (\forall x \in \mathcal{X})(\forall y \in \mathcal{X})\ \|x-y\|^2/(2\alpha) \geqslant f(x) - f(y) - \langle x-y, \nabla f(y)\rangle\\ &\Leftrightarrow\ (\forall x \in \mathcal{X})(\forall y \in \mathcal{X})\ f(x) \leqslant f(y) + \langle x-y, \nabla f(y)\rangle + \|x-y\|^2/(2\alpha).\end{aligned} \tag{2.5}$$

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,

$$\begin{aligned}(\forall x \in \mathcal{X})(\forall y \in \mathcal{X})\quad D^g(x,y) &= \langle x, Vx\rangle/2 - \langle y, Vy\rangle/2 - \langle x-y, Vy\rangle\\ &= \langle x-y, Vx - Vy\rangle/2\\ &\geqslant \alpha\langle x-y, Ux - Uy\rangle/2\\ &= \alpha D^f(x,y).\end{aligned} \tag{2.6}$$

□

Example 2.3 Let $\mathcal{X}$ be a reflexive real Banach space, let $f \in \Gamma_0(\mathcal{X})$ be Gâteaux differentiable on $\operatorname{int}\operatorname{dom} f \neq \varnothing$, let $\alpha \in [0, +\infty[$, and let $g \in \Gamma_0(\mathcal{X})$ be Gâteaux differentiable on $\operatorname{dom} g = \operatorname{int}\operatorname{dom} f$. 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. We first note that $\operatorname{dom} h = \operatorname{int}\operatorname{dom} f$. Since f and g are Gâteaux differentiable on $\operatorname{int}\operatorname{dom} f$ by [15, Proposition 3.3], $h = g - \alpha f$ is likewise. Furthermore,

$$(\forall x \in \operatorname{dom} f)(\forall y \in \operatorname{int}\operatorname{dom} f)\quad D^g(x,y) - \alpha D^f(x,y) = D^h(x,y) \geqslant 0. \tag{2.7}$$

□

The following definition brings together the notions of Bregman monotone sequences [4] and of variable metric Fejér monotone sequences [9].

Definition 2.4 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

$$\begin{aligned}&(\exists(\eta_n)_{n\in\mathbb{N}} \in \ell_+^1(\mathbb{N}))(\forall x \in C \cap \operatorname{dom} f)(\exists(\varepsilon_n)_{n\in\mathbb{N}} \in \ell_+^1(\mathbb{N}))(\forall n \in \mathbb{N})\\ &\qquad D^{f_{n+1}}(x, x_{n+1}) \leqslant (1+\eta_n) D^{f_n}(x, x_n) + \varepsilon_n;\end{aligned} \tag{2.8}$$

(ii) *stationarily quasi-Bregman monotone* with respect to C relative to $(f_n)_{n\in\mathbb{N}}$ if

$$(\exists(\varepsilon_n)_{n\in\mathbb{N}} \in \ell_+^1(\mathbb{N}))(\exists(\eta_n)_{n\in\mathbb{N}} \in \ell_+^1(\mathbb{N}))(\forall x \in C \cap \operatorname{dom} f)(\forall n \in \mathbb{N})$$
$$D^{f_{n+1}}(x, x_{n+1}) \leqslant (1+\eta_n)D^{f_n}(x, x_n) + \varepsilon_n. \quad (2.9)$$

Remark 2.5

(i) In Definition 2.4, suppose that $(\forall n \in \mathbb{N})$ $f_n = \hat{f}$ and $\eta_n = \varepsilon_n = 0$. Then we recover the notion of a Bregman monotone sequence defined in [4].

(ii) In Definition 2.4, 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 2.6 *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

$$(\forall n \in \mathbb{N}) \quad \xi_{n+1} \leqslant (1+\eta_n)\xi_n + \varepsilon_n. \quad (2.10)$$

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

$$(\forall n \in \mathbb{N}) \quad D^f(x, x_n) \leqslant \alpha^{-1} D^{f_n}(x, x_n). \quad (2.11)$$

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 2.7 *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 $\alpha \in\,]0, +\infty[$, 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, for every* $x_1 \in \mathcal{X}$ *and every* $x_2 \in \mathcal{X}$*,*

$$\begin{cases} x_1 \in \mathfrak{W}(x_n)_{n\in\mathbb{N}} \cap C \\ x_2 \in \mathfrak{W}(x_n)_{n\in\mathbb{N}} \cap C \\ \big(\langle x_1 - x_2, \nabla f_n(x_n)\rangle\big)_{n\in\mathbb{N}} \quad \textit{converges} \end{cases} \qquad \Rightarrow \qquad x_1 = x_2. \tag{2.12}$$

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 x_1 and x_2 be two such points. First, it follows from Proposition 2.6(i) that

$$\big(D^{f_n}(x_1, x_n)\big)_{n\in\mathbb{N}} \quad \text{and} \quad \big(D^{f_n}(x_2, x_n)\big)_{n\in\mathbb{N}} \quad \text{are convergent.} \tag{2.13}$$

Next, let us define the following functions

$$\phi\colon\ [0,1] \to \mathbb{R}\colon t \mapsto \big\langle x_1 - x_2, \nabla g(x_2 + t(x_1 - x_2)) - \nabla g(x_2)\big\rangle, \tag{2.14}$$

and

$$(\forall n \in \mathbb{N}) \quad \phi_n\colon\ [0,1] \to \mathbb{R}\colon t \mapsto \big\langle x_1 - x_2, \nabla f_n(x_2 + t(x_1 - x_2)) - \nabla f_n(x_2)\big\rangle. \tag{2.15}$$

Then

$$\int_0^1 \phi(t)dt = g(x_1) - g(x_2) \quad \text{and} \quad (\forall n \in \mathbb{N}) \quad \int_0^1 \phi_n(t)dt = f_n(x_1) - f_n(x_2). \tag{2.16}$$

For every $n \in \mathbb{N}$, since $(1+\eta_n)f_n \succcurlyeq f_{n+1}$, for every $t \in \,]0,1]$), we have

$$\begin{aligned} \phi_{n+1}(t) &= \big\langle x_1 - x_2, \nabla f_{n+1}(x_2 + t(x_1 - x_2)) - \nabla f_{n+1}(x_2)\big\rangle \\ &= t^{-1}\big\langle x_2 + t(x_1 - x_2) - x_2, \nabla f_{n+1}(x_2 + t(x_1 - x_2)) - \nabla f_{n+1}(x_2)\big\rangle \\ &= t^{-1}\big(D^{f_{n+1}}\big(x_2 + t(x_1 - x_2), x_2\big) + D^{f_{n+1}}\big(x_2, x_2 + t(x_1 - x_2)\big)\big) \\ &\leqslant (1+\eta_n)t^{-1}\big(D^{f_n}\big(x_2 + t(x_1 - x_2), x_2\big) + D^{f_n}\big(x_2, x_2 + t(x_1 - x_2)\big)\big) \\ &= (1+\eta_n)t^{-1}\big\langle x_2 + t(x_1 - x_2) - x_2, \nabla f_n(x_2 + t(x_1 - x_2)) - \nabla f_n(x_2)\big\rangle \\ &= (1+\eta_n)\big\langle x_1 - x_2, \nabla f_n(x_2 + t(x_1 - x_2)) - \nabla f_n(x_2)\big\rangle \\ &= (1+\eta_n)\phi_n(t). \end{aligned} \tag{2.17}$$

Consequently,

$$(\forall n \in \mathbb{N})(\forall t \in \,]0,1]) \quad 0 \leqslant \phi_{n+1}(t) \leqslant (1+\eta_n)\phi_n(t). \tag{2.18}$$

It is clear that (2.18) 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

$$(\phi_n)_{n\in\mathbb{N}} \quad \text{converges pointwise.} \tag{2.19}$$

On the other hand, for every $n \in \mathbb{N}$, since $g \succcurlyeq f_n$, the same argument as above shows that

$$(\forall t \in [0,1]) \quad 0 \leqslant \phi_n(t) \leqslant \phi(t). \tag{2.20}$$

By invoking (2.19), (2.20), and Lebesgue's dominated convergence theorem, we obtain that

$$\left(\int_0^1 \phi_n(t) dt \right)_{n \in \mathbb{N}} \quad \text{converges,} \tag{2.21}$$

which implies that

$$\Big(f_n(x_1) - f_n(x_2) \Big)_{n \in \mathbb{N}} \quad \text{converges.} \tag{2.22}$$

We also observe that

$$(\forall n \in \mathbb{N}) \quad D^{f_n}(x_1, x_n) - D^{f_n}(x_2, x_n) = f_n(x_1) - f_n(x_2) - \langle x_1 - x_2, \nabla f_n(x_n) \rangle, \tag{2.23}$$

and hence, it follows from (2.13) and (2.22) that

$$\big(\langle x_1 - x_2, \nabla f_n(x_n) \rangle \big)_{n \in \mathbb{N}} \quad \text{converges.} \tag{2.24}$$

In turn, (2.12) forces $x_1 = x_2$. Since Proposition 2.6(ii) asserts that $(x_n)_{n \in \mathbb{N}}$ is bounded and since $\mathcal{X}$ is reflexive, we conclude that $x_n \rightharpoonup x_1 \in C \cap \operatorname{int} \operatorname{dom} f$. ☐

Example 2.8 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 $C \subset \mathcal{X}$. Suppose that $C \cap \overline{\operatorname{dom}}\, f$ is a singleton. Then (2.12) 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 2.9 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 $(\forall n \in \mathbb{N})$ $f_n = \hat{f}$. Suppose that $f|_{\operatorname{int} \operatorname{dom} f}$ is strictly convex and that ∇f is weakly sequentially continuous. Then (2.12) is satisfied.

Proof. Suppose that $x_1 \in \mathfrak{W}(x_n)_{n \in \mathbb{N}} \cap C$ and $x_2 \in \mathfrak{W}(x_n)_{n \in \mathbb{N}} \cap C$ are such that $(\langle x_1 - x_2, \nabla f_n(x_n) \rangle)_{n \in \mathbb{N}}$ converges and $x_1 \neq x_2$. Take strictly increasing sequences $(k_n)_{n \in \mathbb{N}}$ and $(l_n)_{n \in \mathbb{N}}$ in $\mathbb{N}$ such that $x_{k_n} \rightharpoonup x_1$ and $x_{l_n} \rightharpoonup x_2$. Since ∇f is weakly sequentially continuous, by taking the limit in (2.12) along subsequences $(x_{k_n})_{n \in \mathbb{N}}$ and $(x_{l_n})_{n \in \mathbb{N}}$, we get

$$\langle x_1 - x_2, \nabla f(x_1) - \nabla f(x_2) \rangle = 0 \tag{2.25}$$

Since $f|_{\operatorname{int} \operatorname{dom} f}$ is strictly convex, ∇f is strictly monotone [19, Theorem 2.4.4(ii)], i.e.,

$$\langle x_1 - x_2, \nabla f(x_1) - \nabla f(x_2) \rangle > 0, \tag{2.26}$$

and we reach a contradiction. ☐

Example 2.10 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 (2.12) 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 $x_1 \in \mathfrak{W}(x_n)_{n\in\mathbb{N}} \cap C$ and $x_2 \in \mathfrak{W}(x_n)_{n\in\mathbb{N}} \cap C$ are such that $(\langle x_1 - x_2, \nabla f_n(x_n)\rangle)_{n\in\mathbb{N}}$ converges. Take strictly increasing sequences $(k_n)_{n\in\mathbb{N}}$ and $(l_n)_{n\in\mathbb{N}}$ in $\mathbb{N}$ such that $x_{k_n} \rightharpoonup x_1$ and $x_{l_n} \rightharpoonup x_2$. We have

$$\big\langle x_1 - x_2, \nabla f_{k_n}(x_{k_n})\big\rangle = \big\langle x_1 - x_2, U_{k_n}x_{k_n}\big\rangle = \big\langle U_{k_n}x_1 - U_{k_n}x_2, x_{k_n}\big\rangle \to \big\langle Ux_1 - Ux_2, x_1\big\rangle, \quad (2.27)$$

and

$$\big\langle x_1 - x_2, \nabla f_{l_n}(x_{l_n})\big\rangle = \big\langle x_1 - x_2, U_{l_n}x_{l_n}\big\rangle = \big\langle U_{l_n}x_1 - U_{l_n}x_2, x_{l_n}\big\rangle \to \big\langle Ux_1 - Ux_2, x_2\big\rangle, \quad (2.28)$$

and hence, $0 = \big\langle Ux_1 - Ux_2, x_1 - x_2\big\rangle \geqslant \alpha\|x_1 - x_2\|^2$, and therefore, $x_1 = x_2$. □

The following condition will be used subsequently (see [4, Examples 4.10, 5.11, and 5.13] for special cases).

Condition 2.11 [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 $\operatorname{int}\operatorname{dom} f$,

$$D^f(x_n, y_n) \to 0 \quad \Rightarrow \quad x_n - y_n \to 0. \quad (2.29)$$

We now present a characterization of the strong convergence of stationarily quasi-Bregman monotone sequences.

Proposition 2.12 *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 2.11, 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 $(\forall n \in \mathbb{N})\ \beta\hat{f} \succcurlyeq f_n$. Then $(x_n)_{n\in\mathbb{N}}$ converges strongly to a point in $C \cap \overline{\operatorname{dom}} f$ if and only if $\underline{\lim}\, D^f_C(x_n) = 0$.*

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 2.6(i) states that $(D^{f_n}(x, x_n))_{n\in\mathbb{N}}$ is bounded and since

$$(\forall n \in \mathbb{N}) \quad D^f(x, x_n) \leqslant D^{f_n}(x, x_n), \quad (2.30)$$

we deduce that $(D^f(x, x_n))_{n\in\mathbb{N}}$ is bounded. However, by [3, Lemma 7.3(vii)],

$$(\forall n \in \mathbb{N}) \quad D^{f^*}\big(\nabla f(x_n), \nabla f(x)\big) = D^f(x, x_n). \quad (2.31)$$

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

$$(\forall n \in \mathbb{N}) \quad \begin{aligned} D_C^f(x_n) &= \inf D^f(C, x_n) \\ &\leqslant \inf D^f(C \cap \overline{\text{dom}}\, f, x_n) \\ &\leqslant D^f(\overline{x}, x_n) \\ &= f(\overline{x}) - f(x_n) - \big\langle \overline{x} - x_n, \nabla f(x_n)\big\rangle, \end{aligned} \tag{2.32}$$

we obtain

$$\underline{\lim}\, D_C^f(x_n) \leqslant f(\overline{x}) - \overline{\lim}\, f(x_n) - \lim \big\langle \overline{x} - x_n, \nabla f(x_n)\big\rangle = f(\overline{x}) - \overline{\lim}\, f(x_n). \tag{2.33}$$

Since f is lower semicontinuous,

$$f(\overline{x}) \leqslant \underline{\lim}\, f(x_n) \leqslant \overline{\lim}\, f(x_n). \tag{2.34}$$

Altogether, (2.33) and (2.34) yield

$$\underline{\lim}\, D_C^f(x_n) \to 0. \tag{2.35}$$

We now show the sufficiency. First, since f is Legendre and $C \cap \text{int}\,\text{dom} f \neq \varnothing$, (1.4) yields

$$P_C^f \colon \text{int}\,\text{dom} f \to C \cap \text{int}\,\text{dom} f. \tag{2.36}$$

Next, we set

$$(\forall n \in \mathbb{N}) \quad \varrho_n = D_C^f(x_n) \quad \text{and} \quad \zeta_n = \inf_{x \in C \cap \text{dom}\, f} D^{f_n}(x, x_n). \tag{2.37}$$

Then $\underline{\lim}\, \varrho_n = 0$. For every $n \in \mathbb{N}$, since $\beta \hat{f} \succcurlyeq f_n \succcurlyeq \alpha f$, we obtain

$$(\forall x \in C \cap \text{dom}\, f) \quad 0 \leqslant \alpha D^f(x, x_n) \leqslant D^{f_n}(x, x_n) \leqslant \beta D^f(x, x_n). \tag{2.38}$$

In the above inequalities, after taking the infimum over $x \in C \cap \text{dom}\, f$, we get

$$(\forall n \in \mathbb{N}) \quad 0 \leqslant \alpha \varrho_n \leqslant \zeta_n \leqslant \beta \varrho_n \tag{2.39}$$

and therefore,

$$0 \leqslant \alpha\, \underline{\lim}\, \varrho_n \leqslant \underline{\lim}\, \zeta_n \leqslant \beta\, \underline{\lim}\, \varrho_n = 0. \tag{2.40}$$

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

$$(\forall x \in C \cap \text{dom}\, f)(\forall n \in \mathbb{N}) \quad D^{f_{n+1}}(x, x_{n+1}) \leqslant (1 + \eta_n) D^{f_n}(x, x_n) + \varepsilon_n. \tag{2.41}$$

Taking the infimum in (2.41) over $C \cap \text{dom}\, f$ yields

$$(\forall n \in \mathbb{N}) \quad \zeta_{n+1} \leqslant (1 + \eta_n)\zeta_n + \varepsilon_n. \tag{2.42}$$

It therefore follows from [16, Lemma 2.2.2] that $(\zeta_n)_{n\in\mathbb{N}}$ converges, and thus, we deduce from (2.40) that $\zeta_n \to 0$. Appealing to (2.39), we get $\varrho_n \to 0$, i.e.,

$$D^f\big(P_C^f x_n, x_n\big) \to 0. \tag{2.43}$$

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 2.6(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

$$(\forall n \in \mathbb{N}) \quad D^f\big(x, P_C^f x_n\big) \leqslant D^f(x, x_n), \tag{2.44}$$

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

$$\big(P_C^f x_n\big)_{n\in\mathbb{N}} \in (\operatorname{int}\operatorname{dom} f)^{\mathbb{N}} \ \text{ is bounded.} \tag{2.45}$$

Therefore, since f satisfies Condition 2.11, it follows from (2.43) that

$$P_C^f x_n - x_n \to 0. \tag{2.46}$$

Since (2.36) entails that

$$(\forall n \in \mathbb{N}) \quad P_C^f x_n \in C \cap \operatorname{int}\operatorname{dom} f = \operatorname{Fix} P_C^f, \tag{2.47}$$

we obtain

$$(\forall n \in \mathbb{N}) \quad 0 \leqslant d_C(x_n) = \inf_{x\in C} \|x - x_n\| \leqslant \|P_C^f x_n - x_n\|. \tag{2.48}$$

Altogether, (2.46) and (2.48) imply that

$$d_C(x_n) \to 0. \tag{2.49}$$

Set $\tau = \prod_{k\in\mathbb{N}}(1+\eta_k)$. Then $\tau < +\infty$ [12, Theorem 3.7.3]. By invoking (2.47) and [4, Proposition 3.3(i) and Theorem 3.34], we get

$$\begin{aligned}(\forall n \in \mathbb{N})(\forall m \in \mathbb{N}) \quad D^f\big(P_C^f x_n, P_C^f x_{m+n}\big) &\leqslant D^f\big(P_C^f x_n, x_{m+n}\big)\\ &\leqslant \alpha^{-1} D^{f_{m+n}}\big(P_C^f x_n, x_{m+n}\big)\\ &\leqslant \tau\alpha^{-1}\left(D^{f_n}\big(P_C^f x_n, x_n\big) + \sum_{k=n}^{n+m-1}\varepsilon_k\right)\\ &\leqslant \tau\alpha^{-1}\left(\beta D^f\big(P_C^f x_n, x_n\big) + \sum_{k\geqslant n}\varepsilon_k\right)\\ &= \tau\alpha^{-1}\left(\beta\varrho_n + \sum_{k\geqslant n}\varepsilon_k\right).\end{aligned} \tag{2.50}$$

After taking the limit as $n \to +\infty$ and $m \to +\infty$ in (2.50), we obtain

$$D^f\big(P_C^f x_{m+n}, P_C^f x_n\big) \to 0, \tag{2.51}$$

and thus (2.45) yield

$$P_C^f x_{m+n} - P_C^f x_n \to 0. \tag{2.52}$$

However,

$$(\forall n \in \mathbb{N})(\forall m \in \mathbb{N}) \quad \|x_{m+n} - x_n\| \leqslant \|x_{m+n} - P_C^f x_{m+n}\| + \|P_C^f x_{m+n} - P_C^f x_n\| + \|P_C^f x_n - x_n\|. \tag{2.53}$$

After taking the limit as $n \to +\infty$ and $m \to +\infty$ in (2.53) then using (2.46) and (2.52), we get

$$\|x_{n+m} - x_n\| \to 0. \tag{2.54}$$

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 (2.49) 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 2.13 In Proposition 2.12, 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 $\phi = |\cdot|^2/2$.

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 $\varphi \in \Gamma_0(\mathcal{H})$

$$\operatorname{prox}_\varphi \colon \mathcal{H} \to \mathcal{H} \colon x \mapsto \operatorname{argmin}\big(\varphi + \|\cdot - x\|^2/2\big). \tag{3.1}$$

Because the quadratic term in (3.1) 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 3.1 [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) $\operatorname{prox}_\varphi^f$ *is single-valued on its domain.*

(ii) $\operatorname{ran}\operatorname{prox}_\varphi^f \subset \operatorname{dom}\operatorname{prox}_\varphi^f = \operatorname{int}\operatorname{dom} f$.

(iii) $\operatorname{prox}_\varphi^f = (\nabla f + \partial\varphi)^{-1} \circ \nabla f$.

(iv) $\operatorname{Fix}\operatorname{prox}_\varphi^f = \operatorname{Argmin}\varphi \cap \operatorname{int}\operatorname{dom} f$.

(v) *Let* $x \in \operatorname{Argmin}\varphi \cap \operatorname{int}\operatorname{dom} f$, *let* $y \in \operatorname{int}\operatorname{dom} f$, *and let* $v = \operatorname{prox}_\varphi^f y$. *Then*

$$D^f(x, v) + D^f(v, y) \leqslant D^f(x, y). \tag{3.2}$$

The following result in an extension of [5, Proposition 23.30].

Proposition 3.2 *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 $f\colon \mathcal{X} \to \left]-\infty,+\infty\right] : x \mapsto \sum_{i=1}^m f_i(x_i)$ and $\varphi\colon \mathcal{X} \to \left]-\infty,+\infty\right] : x \mapsto \sum_{i=1}^m \varphi_i(x_i)$. Then*

$$\Big(\forall x \in \bigtimes_{i=1}^m \operatorname{int}\operatorname{dom} f_i\Big) \quad \operatorname{prox}_{\varphi}^{f} x = \big(\operatorname{prox}_{\varphi_i}^{f_i} x_i\big)_{1\leqslant i\leqslant m}. \tag{3.3}$$

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

$$\partial f\colon \mathcal{X} \to 2^{\mathcal{X}^*} : (x_i)_{1\leqslant i\leqslant m} \mapsto \bigtimes_{i=1}^m \partial f_i(x_i). \tag{3.4}$$

Thus, ∂f is single-valued on

$$\operatorname{dom}\partial f = \bigtimes_{i=1}^m \operatorname{dom}\partial f_i = \bigtimes_{i=1}^m \operatorname{int}\operatorname{dom} f_i = \operatorname{int}\Big(\bigtimes_{i=1}^m \operatorname{dom} f_i\Big) = \operatorname{int}\operatorname{dom} f. \tag{3.5}$$

Likewise, since

$$f^*\colon \mathcal{X}^* \to \left]-\infty,+\infty\right] : (x_i^*)_{1\leqslant i\leqslant m} \mapsto \sum_{i=1}^m f_i^*(x_i^*), \tag{3.6}$$

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,

$$\operatorname{dom}\varphi \cap \operatorname{int}\operatorname{dom} f = \Big(\bigtimes_{i=1}^m \operatorname{dom}\varphi_i\Big) \cap \Big(\bigtimes_{i=1}^m \operatorname{int}\operatorname{dom} f_i\Big) = \bigtimes_{i=1}^m (\operatorname{dom}\varphi_i \cap \operatorname{int}\operatorname{dom} f_i) \neq \varnothing. \tag{3.7}$$

Now Lemma 3.1 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 3.1(iii) yields $\nabla f(x) - \nabla f(p) \in \partial\varphi(p)$, we deduce from (1.1) that

$$(\forall z \in \operatorname{dom}\varphi) \quad \langle z - p, \nabla f(x) - \nabla f(p)\rangle + \varphi(p) \leqslant \varphi(z). \tag{3.8}$$

Setting $z = q$ in (3.8) yields

$$\langle q - p, \nabla f(x) - \nabla f(p)\rangle + \varphi(p) \leqslant \varphi(q). \tag{3.9}$$

For every $i \in \{1,\ldots,m\}$, set $q_i = \operatorname{prox}_{\varphi_i}^{f_i} x_i$. The same characterization as in (3.8) yields

$$(\forall i \in \{1,\ldots,m\})(\forall z_i \in \operatorname{dom}\varphi_i) \quad \langle z_i - q_i, \nabla f_i(x_i) - \nabla f_i(q_i)\rangle + \varphi_i(q_i) \leqslant \varphi_i(z_i). \tag{3.10}$$

By summing these inequalities over $i \in \{1, \dots, m\}$, we obtain

$$(\forall z \in \operatorname{dom} \varphi) \quad \langle z - q, \nabla f(x) - \nabla f(q) \rangle + \varphi(q) \leqslant \varphi(z). \tag{3.11}$$

Upon setting $z = p$ in (3.11), we get

$$\langle p - q, \nabla f(x) - \nabla f(q) \rangle + \varphi(q) \leqslant \varphi(p). \tag{3.12}$$

Adding (3.9) and (3.12) yields

$$\langle p - q, \nabla f(p) - \nabla f(q) \rangle \leqslant 0. \tag{3.13}$$

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.,

$$\langle p - q, \nabla f(p) - \nabla f(q) \rangle > 0, \tag{3.14}$$

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 3.3 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.,

$$\vartheta \colon \xi \mapsto \begin{cases} \xi \ln \xi - \xi, & \text{if } \xi \in \left]0, +\infty\right[; \\ 0, & \text{if } \xi = 0; \\ +\infty, & \text{otherwise.} \end{cases} \tag{3.15}$$

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 3.1 asserts that $\operatorname{dom} \operatorname{prox}^f_{\gamma\varphi} = \left]0, +\infty\right[^m$. Let $(\xi_i)_{1 \leqslant i \leqslant m} \in \left]0, +\infty\right[^m$, set $(\eta_i)_{1 \leqslant i \leqslant m} = \operatorname{prox}^f_{\gamma\varphi} (\xi_i)_{1 \leqslant i \leqslant m}$, let W be the Lambert function [10], i.e., the inverse of $\xi \mapsto \xi e^\xi$ on $\left[0, +\infty\right[$, and let $i \in \{1, \dots, m\}$. Then η_i can be computed as follows.

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

$$\phi \colon \xi \mapsto \begin{cases} \xi \ln \xi - \omega \xi, & \text{if } \xi \in \left]0, +\infty\right[; \\ 0, & \text{if } \xi = 0; \\ +\infty, & \text{otherwise.} \end{cases} \tag{3.16}$$

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

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

$$\phi \colon \xi \mapsto \begin{cases} \xi^p/p, & \text{if } \xi \in \left[0, +\infty\right[; \\ +\infty, & \text{otherwise.} \end{cases} \tag{3.17}$$

Then

$$\eta_i = \begin{cases} \left(\dfrac{W(\gamma(p-1)\xi_i^{p-1})}{\gamma(p-1)}\right)^{\frac{1}{p-1}}, & \text{if } \ p \in \left]1,+\infty\right[\,; \\ \xi_i e^{-\gamma}, & \text{if } \ p = 1. \end{cases} \tag{3.18}$$

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

$$\phi\colon \xi \mapsto \begin{cases} \xi^{-p}/p, & \text{if } \ \xi \in \left]0,+\infty\right[\,; \\ +\infty, & \text{otherwise.} \end{cases} \tag{3.19}$$

Then

$$\eta_i = \left(\frac{W(\gamma(p+1)\xi_i^{-p-1})}{\gamma(p+1)}\right)^{\frac{-1}{p+1}}. \tag{3.20}$$

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

$$\phi\colon \xi \mapsto \begin{cases} -\xi^{p}/p, & \text{if } \ \xi \in \left[0,+\infty\right[\,; \\ +\infty, & \text{otherwise.} \end{cases} \tag{3.21}$$

Then

$$\eta_i = \left(\frac{W(\gamma(1-p)\xi_i^{p-1})}{\gamma(1-p)}\right)^{\frac{1}{p-1}}. \tag{3.22}$$

Example 3.4 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.,

$$\vartheta\colon \xi \mapsto \begin{cases} \xi\ln\xi - (1-\xi)\ln(1-\xi), & \text{if } \ \xi \in \left]0,1\right[\,; \\ 0 & \text{if } \ \xi \in \{0,1\}; \\ +\infty, & \text{otherwise.} \end{cases} \tag{3.23}$$

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 3.1 asserts that $\operatorname{dom}\operatorname{prox}^f_\varphi = \left]0,1\right[^m$. Let $(\xi_i)_{1\leqslant i\leqslant m} \in \left]0,1\right[^m$, set $(\eta_i)_{1\leqslant i\leqslant m} = \operatorname{prox}^f_\varphi(\xi_i)_{1\leqslant i\leqslant m}$, and let $i \in \{1,\ldots,m\}$. Then η_i can be computed as follows.

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

$$\phi\colon \xi \mapsto \begin{cases} \xi\ln\xi - \omega\xi, & \text{if } \ \xi \in \left]0,+\infty\right[\,; \\ 0, & \text{if } \ \xi = 0; \\ +\infty, & \text{otherwise.} \end{cases} \tag{3.24}$$

Then $\eta_i = e^{\omega}(2-2\xi_i)^{-1}(-\xi_i + \sqrt{4\xi_i - 3\xi_i^2})$.

(ii) Suppose that

$$\phi\colon \xi \mapsto \begin{cases} (1-\xi)\ln(1-\xi)+\xi, & \text{if } \ \xi \in \left]-\infty,1\right[;\\ 1 & \text{if } \ \xi = 1;\\ +\infty, & \text{otherwise.}\end{cases} \tag{3.25}$$

Then $\eta_i = 1/2 + \xi_i^{-1}/2 - \sqrt{\xi_i^{-2}/4 + \xi_i^{-1}/2 - 3/4}$.

Example 3.5 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.,

$$\vartheta\colon \xi \mapsto \begin{cases} -\ln\xi, & \text{if } \xi \in \left]0,+\infty\right[;\\ +\infty, & \text{otherwise.}\end{cases} \tag{3.26}$$

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 3.1 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

$$\phi\colon \xi \mapsto \begin{cases} -\gamma\ln\xi + \omega\xi + \alpha\xi^{-1}, & \text{if} \quad \xi \in \left]0,+\infty\right[;\\ +\infty, & \text{otherwise.}\end{cases} \tag{3.27}$$

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

$$\phi\colon \xi \mapsto \begin{cases} -\gamma\ln\xi + \alpha\xi^p, & \text{if} \quad \xi \in \left]0,+\infty\right[;\\ +\infty, & \text{otherwise.}\end{cases} \tag{3.28}$$

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

$$\phi\colon \xi \mapsto \begin{cases} \alpha\xi^{-p}, & \text{if} \quad \xi \in \left]0,+\infty\right[;\\ +\infty, & \text{otherwise.}\end{cases} \tag{3.29}$$

Then η_i is the strictly positive solution of $p\eta^{p+1} - \xi_i\eta^p = \alpha p\xi_i$.

Example 3.6 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.,

$$\vartheta\colon \xi \mapsto \begin{cases} -\sqrt{1-\xi^2}, & \text{if } \ \xi \in [-1,1];\\ +\infty, & \text{otherwise,}\end{cases} \tag{3.30}$$

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 3.1 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 $(\forall i \in \{1,\ldots,m\})\ \eta_i = \xi_i/\sqrt{(\gamma+1)^2 + (\gamma^2+2\gamma+2)\xi_i^2}$.

4 Applications

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 4.1 *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 $\alpha \in \left]0,+\infty\right[$, and let $(f_n)_{n\in\mathbb{N}}$ be Legendre functions in $\mathcal{P}_\alpha(f)$ such that*

$$(\forall n \in \mathbb{N}) \quad (1+\eta_n)f_n \succcurlyeq f_{n+1}. \tag{4.1}$$

Let $x_0 \in \operatorname{int}\operatorname{dom} f$, let $(\gamma_n)_{n\in\mathbb{N}} \in \left]0,+\infty\right[^{\mathbb{N}}$ be such that $\gamma = \inf_{n\in\mathbb{N}}\gamma_n > 0$, and iterate

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = \operatorname{prox}^{f_n}_{\gamma_n\varphi} x_n. \tag{4.2}$$

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 $x_1 \in \mathcal{X}$ and every $x_2 \in \mathcal{X}$,*

$$\begin{cases} x_1 \in \mathfrak{W}(x_n)_{n\in\mathbb{N}} \\ x_2 \in \mathfrak{W}(x_n)_{n\in\mathbb{N}} \\ \big(\langle x_1 - x_2, \nabla f_n(x_n)\rangle\big)_{n\in\mathbb{N}} \quad \textit{converges} \end{cases} \quad \Rightarrow \quad x_1 = x_2. \tag{4.3}$$

Then there exists $\overline{x} \in \operatorname{Argmin}\varphi$ such that $x_n \rightharpoonup \overline{x}$.

(iv) *Suppose that that f satisfies Condition 2.11 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 $\beta \in \left]0,+\infty\right[$ such that $(\forall n \in \mathbb{N})$ $\beta\hat{f} \succcurlyeq f_n$. 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 3.1 asserts that

$$\operatorname{prox}^{f_n}_{\gamma_n\varphi} \colon \operatorname{int}\operatorname{dom} f_n \to \operatorname{dom}\partial\varphi \cap \operatorname{int}\operatorname{dom} f_n \tag{4.4}$$

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 (4.4) 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

$$(x_n)_{n\in\mathbb{N}} \in \big(\operatorname{int}\operatorname{dom} f\big)^{\mathbb{N}} \quad \text{is well-defined.} \tag{4.5}$$

(i): We first derive from (4.2) and Lemma 3.1(iii) that

$$(\forall n \in \mathbb{N}) \quad \nabla f_n(x_n) - \nabla f_n(x_{n+1}) \in \gamma_n \partial\varphi(x_{n+1}). \tag{4.6}$$

Next, by invoking (1.1) and (4.6), we get

$$(\forall x \in \operatorname{dom}\varphi \cap \operatorname{dom} f)(\forall n \in \mathbb{N}) \quad \gamma_n^{-1}\big\langle x - x_{n+1}, \nabla f_n(x_n) - \nabla f_n(x_{n+1})\big\rangle + \varphi(x_{n+1}) \leqslant \varphi(x). \tag{4.7}$$

It therefore follows from [3, Proposition 2.3(ii)] that

$$\begin{aligned}(\forall x \in \operatorname{dom}\varphi \cap \operatorname{dom} f)(\forall n \in \mathbb{N}) \quad \gamma_n^{-1}\big(D^{f_n}(x, x_{n+1}) + D^{f_n}(x_{n+1}, x_n) - D^{f_n}(x, x_n)\big)&\\ + \varphi(x_{n+1}) \leqslant \varphi(x),&\end{aligned} \tag{4.8}$$

and, in particular,

$$(\forall x \in \operatorname{Argmin}\varphi \cap \operatorname{dom} f)(\forall n \in \mathbb{N}) \quad D^{f_n}(x, x_{n+1}) \leqslant D^{f_n}(x, x_n) - D^{f_n}(x_{n+1}, x_n). \tag{4.9}$$

Since (4.1) yields

$$(\forall x \in \operatorname{Argmin}\varphi \cap \operatorname{dom} f)(\forall n \in \mathbb{N}) \quad D^{f_{n+1}}(x, x_{n+1}) \leqslant (1+\eta_n)D^{f_n}(x, x_{n+1}), \tag{4.10}$$

it follows from (4.9) that

$$\begin{aligned}(\forall x \in \operatorname{Argmin}\varphi \cap \operatorname{dom} f)(\forall n \in \mathbb{N}) \quad D^{f_{n+1}}(x, x_{n+1}) \leqslant (1+\eta_n)D^{f_n}(x, x_n)&\\ - (1+\eta_n)D^{f_n}(x_{n+1}, x_n).&\end{aligned} \tag{4.11}$$

In particular,

$$(\forall x \in \operatorname{Argmin}\varphi \cap \operatorname{dom} f)(\forall n \in \mathbb{N}) \quad D^{f_{n+1}}(x, x_{n+1}) \leqslant (1+\eta_n)D^{f_n}(x, x_n). \tag{4.12}$$

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 2.6(i) that

$$\big(D^{f_n}(x, x_n)\big)_{n\in\mathbb{N}} \quad \text{converges} \tag{4.13}$$

and, since (4.11) yields

$$\begin{aligned}(\forall n \in \mathbb{N}) \quad D^{f_n}(x_{n+1}, x_n) &\leqslant (1+\eta_n)D^{f_n}(x_{n+1}, x_n)\\ &\leqslant (1+\eta_n)D^{f_n}(x, x_n) - D^{f_{n+1}}(x, x_{n+1}),\end{aligned} \tag{4.14}$$

we deduce that

$$D^{f_n}(x_{n+1},x_n)\to 0. \tag{4.15}$$

On the other hand, since $(f_n)_{n\in\mathbb{N}}$ is in $\mathcal{P}_\alpha(f)$, we obtain

$$(\forall n\in\mathbb{N})\quad \alpha D^f(x_{n+1},x_n)\leqslant D^{f_n}(x_{n+1},x_n). \tag{4.16}$$

Altogether, (4.15) and (4.16) yield

$$D^f(x_{n+1},x_n)\to 0. \tag{4.17}$$

We also deduce from (4.8) that

$$(\forall n\in\mathbb{N})\quad \varphi(x_{n+1})\leqslant \gamma_n^{-1}\big(D^{f_n}(x_n,x_{n+1})+D^{f_n}(x_{n+1},x_n)\big)+\varphi(x_{n+1})\leqslant\varphi(x_n). \tag{4.18}$$

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 (4.8) and (4.10) that

$$\begin{aligned}(\forall n\in\mathbb{N})\quad &\frac{1}{\gamma}\left(\frac{1}{1+\eta_n}D^{f_{n+1}}(x,x_{n+1})+D^{f_n}(x_{n+1},x_n)-D^{f_n}(x,x_n)\right)+\varphi(x_{n+1})\\ &\leqslant\frac{1}{\gamma_n}\left(\frac{1}{1+\eta_n}D^{f_{n+1}}(x,x_{n+1})+D^{f_n}(x_{n+1},x_n)-D^{f_n}(x,x_n)\right)+\varphi(x_{n+1})\\ &\leqslant\varphi(x).\end{aligned} \tag{4.19}$$

Hence, by using (4.13) and (4.15) after letting $n\to+\infty$ in (4.19), we get

$$\inf\varphi(\mathcal{X})\leqslant\lim\varphi(x_n)\leqslant\varphi(x)=\inf\varphi(\mathcal{X}). \tag{4.20}$$

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,

$$\inf\varphi(\mathcal{X})\leqslant\varphi(x)\leqslant\underline{\lim}\,\varphi(x_{k_n})=\inf\varphi(\mathcal{X}). \tag{4.21}$$

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 2.6(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 (4.5) yields $\mathfrak{W}(x_n)_{n\in\mathbb{N}}\subset\operatorname{Argmin}\varphi\cap\overline{\operatorname{dom}}\,f$, 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

$$(\forall n\in\mathbb{N})\quad D^{f^*}(\nabla f(x_{k_n}),\nabla f(x))=D^f(x,x_{k_n}), \tag{4.22}$$

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 (4.22) that

$$D^{f^*}(\overline{x}^*, \nabla f(x)) \leqslant \underline{\lim}\, D^{f^*}\big(\nabla f(x_{p_{k_n}}), \nabla f(x)\big) \leqslant \underline{\lim}\, D^f(x, x_{p_{k_n}}) < +\infty, \tag{4.23}$$

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

$$\overline{x} \leftharpoonup x_{p_{k_n}} = \nabla f^*\big(\nabla f(x_{p_{k_n}})\big) \rightharpoonup \nabla f^*(\overline{x}^*) = \overline{x}_1. \tag{4.24}$$

In turn, $\overline{x} = \overline{x}_1 \in \operatorname{int}\operatorname{dom} f$. Finally, the claim follows from Proposition 2.7.

(iv): Since $\varphi \in \Gamma_0(\mathcal{X})$, $\operatorname{Argmin}\varphi$ is convex and closed, and the assertion therefore follows from Proposition 2.12. □

Remark 4.2 In Theorem 4.1, suppose that $(\forall n \in \mathbb{N})$ $f_n = \hat{f}$, $\gamma_n = \gamma$, and $\eta_n = 0$. Then (4.2) reduces to the Bregman proximal iterations [4]

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = \operatorname{prox}^f_{\gamma\varphi} x_n. \tag{4.25}$$

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 4.3 [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*

$$\begin{aligned}(\forall (x,y) \in (\operatorname{int}\operatorname{dom} f)^2) \quad H^f(x,y) &= \big\{z \in \mathcal{X} \;\big|\; \big\langle z - y, \nabla f(x) - \nabla f(y)\big\rangle \leqslant 0\big\}\\ &= \big\{z \in \mathcal{X} \;\big|\; D^f(z,y) + D^f(y,x) \leqslant D^f(z,x)\big\}\end{aligned} \tag{4.26}$$

and

$$\mathfrak{B}(f) = \Big\{T\colon \mathcal{X} \to 2^{\mathcal{X}} \;\Big|\; \operatorname{ran} T \subset \operatorname{dom} T = \operatorname{int}\operatorname{dom} f \text{ and } (\forall (x,y) \in \operatorname{gra} T)\ \operatorname{Fix} T \subset H^f(x,y)\Big\}. \tag{4.27}$$

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 2 through an application to the problem of finding a common point of a family of closed convex subsets with nonempty intersection.

Theorem 4.4 *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*

$$(\forall n \in \mathbb{N}) \quad (1+\eta_n) f_n \succcurlyeq f_{n+1}. \tag{4.28}$$

Let $\mathrm{i}\colon \mathbb{N} \to I$ be such that

$$(\forall j \in I)(\exists M_j \in \mathbb{N}\backslash\{0\})(\forall n \in \mathbb{N}) \quad j \in \{\mathrm{i}(n), \ldots, \mathrm{i}(n+M_j-1)\}. \tag{4.29}$$

For every $i \in I$, let $(T_{i,n})_{n\in\mathbb{N}}$ be a sequence of operators such that

$$(\forall n \in \mathbb{N}) \quad T_{i,n} \in \mathfrak{B}(f_n), \quad C_i \cap \operatorname{Fix} T_{i,n} \neq \varnothing, \quad \text{and} \quad C_i \subset \overline{\operatorname{Fix}}\, T_{i,n}. \tag{4.30}$$

Let $x_0 \in \operatorname{int}\operatorname{dom} f$ and iterate

$$(\forall n \in \mathbb{N}) \quad x_{n+1} \in T_{\mathrm{i}(n),n} x_n. \tag{4.31}$$

Suppose that f satisfies Condition 2.11 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, for every $x_1 \in \mathcal{X}$ and every $x_2 \in \mathcal{X}$,*

$$\begin{cases} x_1 \in \mathfrak{W}(x_n)_{n\in\mathbb{N}} \cap C \\ x_2 \in \mathfrak{W}(x_n)_{n\in\mathbb{N}} \cap C \\ \big(\langle x_1 - x_2, \nabla f_n(x_n)\rangle\big)_{n\in\mathbb{N}} \quad \textit{converges} \end{cases} \quad \Rightarrow \quad x_1 = x_2, \tag{4.32}$$

and that, for every strictly increasing sequence $(l_n)_{n\in\mathbb{N}}$ in $\mathbb{N}$, every $x \in \mathcal{X}$, and every $j \in I$,

$$\begin{cases} x_{l_n} \rightharpoonup x \\ y_{l_n} \in T_{j,l_n} x_{l_n} \\ y_{l_n} - x_{l_n} \to 0 \\ (\forall n \in \mathbb{N})\ j = \mathrm{i}(l_n) \end{cases} \quad \Rightarrow \quad x \in C_j. \tag{4.33}$$

In addition, assume 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 $\varliminf D^f_C(x_n) = 0$, and that there exists $\beta \in \left]0,+\infty\right[$ such that $(\forall n \in \mathbb{N})$ $\beta\hat{f} \succcurlyeq f_n$. Then $x_n \to \overline{x}$.*

Proof. For every $n \in \mathbb{N}$ and every $i \in I$, we observe that $\operatorname{ran} T_{i,n} \subset \operatorname{dom} T_{i,n} = \operatorname{int} \operatorname{dom} f_n = \operatorname{int} \operatorname{dom} f$. Hence, it follows from (4.30) and (4.31) that $(x_n)_{n\in\mathbb{N}}$ is a well-define sequence in $\operatorname{int} \operatorname{dom} f$. We now derive from (4.26), (4.30), and (4.31) that

$$(\forall x \in C \cap \operatorname{dom} f)(\forall n \in \mathbb{N}) \quad D^{f_n}(x, x_{n+1}) + D^{f_n}(x_{n+1}, x_n) \leqslant D^{f_n}(x, x_n). \tag{4.34}$$

Since (4.28) yields

$$(\forall x \in C \cap \operatorname{dom} f)(\forall n \in \mathbb{N}) \quad D^{f_{n+1}}(x, x_{n+1}) \leqslant (1 + \eta_n) D^{f_n}(x, x_{n+1}), \tag{4.35}$$

we deduce that

$$(\forall x \in C \cap \operatorname{dom} f)(\forall n \in \mathbb{N}) \quad D^{f_{n+1}}(x, x_{n+1}) \leqslant (1+\eta_n) D^{f_n}(x, x_n) - (1+\eta_n) D^{f_n}(x_{n+1}, x_n). \tag{4.36}$$

In particular,

$$(\forall x \in C \cap \operatorname{dom} f)(\forall n \in \mathbb{N}) \quad D^{f_{n+1}}(x, x_{n+1}) \leqslant (1 + \eta_n) D^{f_n}(x, x_n), \tag{4.37}$$

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 (4.30) that $(\forall i \in \{1, \ldots, m\})\ C_i \cap \operatorname{int} \operatorname{dom} f \neq \varnothing$. Hence, $C \cap \operatorname{int} \operatorname{dom} f \neq \varnothing$.

(i): In view of Proposition 2.7, it suffices to show that $\mathfrak{W}(x_n)_{n\in\mathbb{N}} \subset C \cap \operatorname{int} \operatorname{dom} f$. 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 $x \in C \cap \operatorname{int} \operatorname{dom} f$. By (4.29), there exists a strictly increasing sequence $(l_n)_{n\in\mathbb{N}}$ in $\mathbb{N}$ such that

$$(\forall n \in \mathbb{N}) \quad \begin{cases} k_n \leqslant l_n \leqslant k_n + M_j - 1 < k_{n+1} \leqslant l_{n+1}, \\ j = \mathrm{i}(l_n). \end{cases} \tag{4.38}$$

Since $D^f(x, \cdot)$ is coercive, it follows from Proposition 2.6 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 (4.36) yields

$$\begin{aligned} (\forall n \in \mathbb{N}) \quad D^{f_n}(x_{n+1}, x_n) &\leqslant (1 + \eta_n) D^{f_n}(x_{n+1}, x_n) \\ &\leqslant (1 + \eta_n) D^{f_n}(x, x_n) - D^{f_{n+1}}(x, x_{n+1}), \end{aligned} \tag{4.39}$$

we deduce that

$$D^{f_n}(x_{n+1}, x_n) \to 0. \tag{4.40}$$

However, since

$$(\forall n \in \mathbb{N}) \quad \alpha D^f(x_{n+1}, x_n) \leqslant D^{f_n}(x_{n+1}, x_n), \tag{4.41}$$

it follows from (4.40) that

$$D^f(x_{n+1}, x_n) \to 0 \tag{4.42}$$

and hence, since f satisfies Condition 2.11,

$$x_{n+1} - x_n \to 0. \tag{4.43}$$

Altogether, (4.38) and (4.43) imply that

$$\|x_{l_n} - x_{k_n}\| \leqslant \sum_{m=k_n}^{k_n+M_j-2} \|x_{m+1} - x_n\| \leqslant (M_j - 1) \max_{k_n \leqslant m \leqslant k_n+M_j-2} \|x_{m+1} - x_m\| \to 0, \tag{4.44}$$

and therefore

$$x_{l_n} \rightharpoonup \overline{x}. \tag{4.45}$$

Now let $(\forall n \in \mathbb{N})\ y_{l_n} \in T_{j,l_n} x_{l_n}$. We deduce from (4.38) and (4.43) that

$$y_{l_n} - x_{l_n} \to 0. \tag{4.46}$$

By invoking successively (4.33), (4.45), and (4.46), we get $\overline{x} \in C_j$, and hence, $\overline{x} \in C$. Consequently, $\mathfrak{W}(x_n)_{n\in\mathbb{N}} \subset C \cap \operatorname{int} \operatorname{dom} f$.

(ii): Since C is closed, the assertion follows from Proposition 2.12. ☐

Remark 4.5

- (i) In Theorem 4.4, suppose that $(\forall n \in \mathbb{N})\ f_n = \hat{f}$ and $\eta_n = 0$. Then we recover the framework of [4, Section 4.2].
- (ii) In Theorem 4.4, 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 4.6 *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* int 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*

$$(\forall n \in \mathbb{N}) \quad (1+\eta_n) f_n \succcurlyeq f_{n+1}. \tag{4.47}$$

Let $x_0 \in \operatorname{int}\operatorname{dom} f$ and iterate

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = P^{f_n}_{C_{1+\operatorname{rem}(n,m)}} x_n, \tag{4.48}$$

where $\operatorname{rem}(\cdot, m)$ is the remainder of the division by m. Suppose that f satisfies Condition 2.11 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, for every $x_1 \in \mathcal{X}$ and every $x_2 \in \mathcal{X}$,*

$$\begin{cases} x_1 \in \mathfrak{W}(x_n)_{n\in\mathbb{N}} \cap C \\ x_2 \in \mathfrak{W}(x_n)_{n\in\mathbb{N}} \cap C \\ \big(\langle x_1 - x_2, \nabla f_n(x_n)\rangle\big)_{n\in\mathbb{N}} \quad \textit{converges} \end{cases} \quad \Rightarrow \quad x_1 = x_2. \tag{4.49}$$

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 $\underline{\lim}\, D_C^f(x_n) = 0$, and that there exists $\beta \in \left]0,+\infty\right[$ such that $(\forall n \in \mathbb{N})\ \beta \hat{f} \succcurlyeq f_n$. 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 (4.29), where $(\forall j \in \{1,\ldots,m\})\ M_j = m$. Now set

$$(\forall i \in \{1,\ldots,m\})(\forall n \in \mathbb{N}) \quad T_{i,n} = P_{C_i}^{f_n}. \tag{4.50}$$

Then, by [4, Theorem 3.34], for every $n \in \mathbb{N}$ and every $i \in \{1,\ldots,m\}$, we have

$$T_{i,n} \in \mathfrak{B}(f_n) \quad \text{and} \quad C_i \cap \overline{\operatorname{dom}}\, f \cap \operatorname{Fix} T_{i,n} = C_i \cap \operatorname{int}\operatorname{dom} f \supset C \cap \operatorname{int}\operatorname{dom} f \neq \varnothing. \tag{4.51}$$

In addition, it follows from [4, Lemma 3.2] that

$$(\forall n \in \mathbb{N})(\forall i \in \{1,\ldots,m\}) \quad C_i \cap \overline{\operatorname{dom}}\, f = \overline{C_i \cap \operatorname{int}\operatorname{dom} f} = \overline{C_i \cap \operatorname{int}\operatorname{dom} f_n} = \overline{\operatorname{Fix}}\, T_{i,n}. \tag{4.52}$$

Therefore, (4.48) is a particular case of (4.31). We shall actually apply Proposition 4.4 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

$$x_{l_n} \rightharpoonup x, \quad T_{j,l_n} x_{l_n} - x_{l_n} \to 0, \quad \text{and} \quad (\forall n \in \mathbb{N}) \quad j = \mathrm{i}(l_n). \tag{4.53}$$

Then $C_j \ni P_{C_j}^{f_{l_n}} x_{l_n} = T_{j,l_n} x_{l_n} \rightharpoonup x$, 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 (4.33) is satisfied. Consequently, the assertion follows from Proposition 4.4(i).

(ii): We have

$$(\forall n \in \mathbb{N}) \qquad \inf_{x \in C \cap \overline{\operatorname{dom}}\, f} D^f(x, x_n) \leqslant \inf_{x \in C \cap \operatorname{dom} f} D^f(x, x_n) = D_C^f(x_n), \tag{4.54}$$

and hence, $\underline{\lim}\, D_{C\cap\overline{\operatorname{dom}}\, f}(x_n) = 0$. The claim therefore follows from Proposition 4.4(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.

References

[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.