Journal of Optimization Theory and Applications manuscript No.
(will be inserted by the editor)

On the Finite Convergence of a Projected Cutter Method

Heinz H. Bauschke · **Caifang Wang** ·
Xianfu Wang · **Jia Xu**

Received: date / Accepted: date

Abstract The subgradient projection iteration is a classical method for solving a convex inequality. Motivated by works of Polyak and of Crombez, we present and analyze a more general method for finding a fixed point of a cutter, provided that the fixed point set has nonempty interior. Our assumptions on the parameters are more general than existing ones. Various limiting examples and comparisons are provided.

Keywords Convex Function · Cutter · Fejér Monotone Sequence · Finite Convergence · Quasi Firmly Nonexpansive Mapping · Subgradient Projector

Mathematics Subject Classification (2000) 90C25 · 47H04 · 47H05 · 47H09 · 65K10

Heinz H. Bauschke (corresponding author)
Mathematics, University of British Columbia, Kelowna, B.C. V1V 1V7, Canada
E-mail: heinz.bauschke@ubc.ca

Caifang Wang
Department of Mathematics, Shanghai Maritime University, China
E-mail: cfwang@shmtu.edu.cn

Xianfu Wang
Mathematics, University of British Columbia, Kelowna, B.C. V1V 1V7, Canada
E-mail: shawn.wang@ubc.ca

Jia Xu
Mathematics, University of British Columbia, Kelowna, B.C. V1V 1V7, Canada
E-mail: jia.xu@ubc.ca

1 Introduction

Solving convex inequalities — or, more generally — the convex feasibility problem is a central problem in optimization (see [1] the references therein). One particular approach to this problem is to express the solution set as the fixed point set of a suitable algorithmic operator, e.g., the relaxation of a cutter [2]. Perhaps the most famous instance is Polyak's subgradient projector of a convex function.

If the operator is nonexpansive and asymptotically regular, then a result of Opial implies that the iterates of the operator converge weakly to a solution. In Chapter 5 of [2], weak convergence of the sequences generated by, e.g., the *alternating projection method*, the *extrapolated alternating projection method*, the *projected gradient method* and other similar methods is demonstrated; however, the question whether a fixed point can be reached within a *finite* number of steps is not fully answered. Censor et al. [3] studied the *modified cyclic subgradient projection* (MCSP) and provided finite convergence conditions while Crombez [4] focussed on cutter iterations where the radius of a ball lying inside the fixed point set is assumed to be known. Unfortunately, this radius is not necessarily known in practice. As a remedy, we propose a more general divergent-series condition.

We will obtain *finite convergence results* for more general algorithms provided a constraint qualification is satisfied. Our results complement and extend results by Crombez for cutters and by Polyak for subgradient projectors. Let us provide the reader with a some representative references for this area. For books dealing with cutters and subgradient projectors, see [2,5,6]. Key papers in this area are by Censor and co-workers [7,8], by Combettes and co-workers [9–11], by Polyak [1,12], and by Yamada and co-workers [13–17]. See also [18–21].

The paper itself is organized as follows. In Section 2, we collect various auxiliary results, that will facilitate the presentation of the main results in Section 3. Limiting examples are presented in Section 4. In Section 5, we compare to existing results.

Future research directions are discussed in Section 6. Finally, Section 7 concludes the paper. Notation is standard and follows, e.g., [22].

2 Auxiliary Results

Throughout this paper, we assume that X is a real Hilbert space with inner product $\langle\cdot,\cdot\rangle$ and induced norm $\|\cdot\|$.

2.1 Cutters

Definition 2.1
We call $T\colon X \to X$ a *cutter* iff $\operatorname{Fix} T := \{y \in X \mid y = Ty\} \neq \varnothing$ and furthermore we have $(\forall x \in X)(\forall y \in \operatorname{Fix} T)\ \langle y - Tx, x - Tx\rangle \leq 0$; equivalently,

$$(\forall x \in X)(\forall y \in \operatorname{Fix} T) \quad \|Tx - y\|^2 + \|x - Tx\|^2 \leq \|x - y\|^2.$$

Cutters are also known as *quasi firmly nonexpansive operators*.

Here we give the most important instance of a cutter, namely Polyak's subgradient projector [12].

Example 2.1
(Subgradient projector) Let $f\colon X \to \mathbb{R}$ be convex and continuous such that

$$\{x \in X \mid f(x) \leq 0\} \neq \varnothing,$$

and let $s\colon X \to X$ be a selection of ∂f, i.e., $(\forall x \in X)\ s(x) \in \partial f(x)$. Then the *associated subgradient projector*, defined by

$$(\forall x \in X) \quad G_f x := \begin{cases} x - \dfrac{f(x)}{\|s(x)\|^2} s(x), & \text{if } f(x) > 0, \\ x, & \text{otherwise}, \end{cases}$$

is a cutter.

Given $r \geq 0$, we follow Crombez [4] and define the operator $U_r\colon X \to X$ at $x \in X$ by

$$U_r x := \begin{cases} x + \dfrac{r + \|Tx - x\|}{\|Tx - x\|}(Tx - x) = Tx + \dfrac{r}{\|Tx - x\|}(Tx - x), & \text{if } x \neq Tx, \\ x, & \text{otherwise.} \end{cases} \tag{1}$$

When T is a subgradient projector, then U_r was also studied by Polyak [1]. Note that $\operatorname{Fix} U_r = \operatorname{Fix} T$.

We now collect some inequalities and identities that will facilitate the proofs of the main results. The inequality $\|U_r x - y\|^2 \leq \|Tx - y\|^2 - r^2$, which is a consequence of (ii) in the next lemma, was also observed by Crombez in [4, Lemma 2.3].

Lemma 2.1

Let $y \in \operatorname{Fix} T$, *let* $r \in \mathbb{R}_{++}$, *and suppose that* $\operatorname{ball}(y;r) \subseteq \operatorname{Fix} T$ *and that* $x \in X \smallsetminus \operatorname{Fix} T$. *Set*

$$\tau_x := \langle x - y, (x - Tx)/\|x - Tx\| \rangle - \big(r + \|x - Tx\|\big).$$

Then the following hold:

(i) $\tau_x \geq 0$.

(ii) $\|U_r x - y\|^2 = \|Tx - y\|^2 - r^2 - 2r\tau_x \leq \|Tx - y\|^2 - r^2$.

(iii) $\|U_r x - y\|^2 = \|x - y\|^2 - (r + \|x - Tx\|)^2 - 2\tau_x(r + \|x - Tx\|)$
$\leq \|x - y\|^2 - (r + \|x - Tx\|)^2 \leq \|x - y\|^2 - r^2 - \|x - Tx\|^2$.

Proof

(i): Set $z := y + r(x - Tx)/\|x - Tx\|$. Then $z \in \operatorname{ball}(y;r) \subseteq \operatorname{Fix} T$. Since T is a cutter, we obtain

$$\begin{aligned} 0 &\geq \langle z - Tx, x - Tx \rangle \\ &= \langle y + r(x - Tx)/\|x - Tx\| - Tx, x - Tx \rangle \\ &= \langle y - Tx, x - Tx \rangle + r\|x - Tx\| \\ &= \langle y - x, x - Tx \rangle + \|x - Tx\|^2 + r\|x - Tx\|. \end{aligned}$$

Rearranging and dividing by $\|x - Tx\|$ yields

$$\langle x - y, (x - Tx)/\|x - Tx\| \rangle \geq r + \|x - Tx\|$$

and hence $\tau_x \geq 0$.

(ii): Using (1), we derive the identity from

$$\begin{aligned}
\|U_r x - y\|^2 &= \big\|x + (\|x-Tx\|+r)/\|x-Tx\|(Tx-x) - y\big\|^2 \\
&= \big\|(Tx-y) + r(Tx-x)/\|Tx-x\|\big\|^2 \\
&= \|Tx-y\|^2 + r^2 + 2r\left\langle (Tx-x)+(x-y), (Tx-x)/\|Tx-x\|\right\rangle \\
&= \|Tx-y\|^2 + r^2 + 2r\|x-Tx\| - 2r\left\langle x-y, (x-Tx)/\|x-Tx\|\right\rangle \\
&= \|Tx-y\|^2 - r^2 - 2r\tau_x.
\end{aligned}$$

The inequality follows immediately from (i).

(iii): Using (ii), we obtain

$$\begin{aligned}
\|U_r x - y\|^2 &= \|(x-y)+(Tx-x)\|^2 - r^2 - 2r\tau_x \\
&= \|x-y\|^2 + \|x-Tx\|^2 + 2\left\langle x-y, Tx-x\right\rangle - r^2 - 2r\tau_x \\
&= \|x-y\|^2 - \|x-Tx\|^2 - 2(\tau_x + r)\|x-Tx\| - r^2 - 2r\tau_x \\
&= \|x-y\|^2 - (r+\|x-Tx\|)^2 - 2\tau_x(r+\|x-Tx\|).
\end{aligned}$$

The inequalities now follow from (i). □

We note in passing that U_r itself is not necessarily a cutter:

Example 2.2

(U_r need not be a cutter) Suppose that $X = \mathbb{R}$ and that T is the subgradient projector associated with the function $f\colon \mathbb{R} \to \mathbb{R}\colon x \mapsto x^2 - 1$. Then $\operatorname{Fix} T = [-1, 1]$. Let

$$r \in \mathbb{R}_+ := \big\{\xi \in \mathbb{R} \;\big|\; \xi \geq 0\big\}.$$

Then

$$(\forall x \in \mathbb{R} \smallsetminus \operatorname{Fix} T) \quad U_r x = \frac{x}{2} + \frac{1}{2x} - r\operatorname{sgn}(x).$$

Choosing $y := 1 \in \operatorname{Fix} T$ and $x := y + \varepsilon \notin \operatorname{Fix} T$, where $\varepsilon \in \mathbb{R}_{++}$, we may check that U_r is not a cutter[1] when ε is sufficiently small and $r > 0$.

[1] In fact, U_r is not even a relaxed cutter in the sense of [2, Definition 2.1.30].

We now obtain the following result concerning a relaxed version[2] of U_r. Item (v) also follows from [2, Corollary 2.4.3].

Corollary 2.1

Let $y \in \operatorname{Fix} T$*, let* $r \in \mathbb{R}_{++}$*, let* $\eta \in \mathbb{R}_{+}$*, and suppose that* $\operatorname{ball}(y;r) \subseteq \operatorname{Fix} T$ *and that* $x \in X \smallsetminus \operatorname{Fix} T$*. Set*

$$U_{r,\eta}x := x + \eta \frac{r + \|x - Tx\|}{\|Tx - x\|}(Tx - x).$$

Then the following hold[3]*:*

(i) $U_{r,\eta}x = (1-\eta)x + \eta U_r x$.
(ii) $\|U_{r,\eta}x - y\|^2 = \eta\|U_r x - y\|^2 + (1-\eta)\|x-y\|^2 - \eta(1-\eta)\|x - U_r x\|^2$.
(iii) $\|U_r x - x\| = r + \|x - Tx\|$.
(iv) $\|U_r x - y\|^2 \le \|x-y\|^2 - (r + \|x - Tx\|)^2 = \|x-y\|^2 - \|x - U_r x\|^2$.
(v) $\|U_{r,\eta}x - y\|^2 \le \|x-y\|^2 - \eta(2-\eta)(r + \|x - Tx\|)^2$
$= \|x-y\|^2 - \eta^{-1}(2-\eta)\|x - U_{r,\eta}x\|^2$.

Proof

(i): This is a simple verification.

(ii): Using (i), we obtain $\|U_{r,\eta}x - y\|^2 = \|(1-\eta)(x-y) + \eta(U_r x - y)\|^2$. Now use [22, Corollary 2.14] to obtain the identity.

(iii): This is immediate from (1).

(iv): Combine (iii) with Lemma 2.1(iii).

(v): Combine (i)–(iv). □

2.2 Quasi Projectors

Definition 2.2

(Quasi projector) We call $Q\colon X \to X$ a *quasi projector* of C iff $\operatorname{ran} Q = \operatorname{Fix} Q = C$ and $(\forall x \in X)(\forall c \in C)\ \|Qx - c\| \le \|x - c\|$.

[2] $U_{r,\eta}$ can also be called a generalized relaxation of T with relaxation parameter η; see [2, Definition 2.4.1] .

[3] We note that item (iv) can also be deduced from [2, (2.27)] with $\lambda = (r + \|x - Tx\|)/\|x - Tx\|$, $z = y$, and $\delta = r$ in [2, Proposition 2.1.41]. This observation, as well as a similar one for (v), is due to a referee.

Example 2.3

(Projectors are quasi projectors) P_C is a quasi projector of C. More generally[4], if $R\colon X \to X$ is quasi nonexpansive, i.e., $(\forall x \in X)(\forall y \in \operatorname{Fix} R)$ $\|Rx - y\| \leq \|x - y\|$ and $C \subseteq \operatorname{Fix} R$, then $P_C \circ R$ is a quasi projector of C.

It can be shown (see [23, Proposition 3.4.4]) that when C is an affine subspace, then the only quasi projector of C is the projector. However, we will now see that for certain cones there are quasi projectors different from projectors.

Proposition 2.1

(Reflector of an obtuse cone) (See [24, Lemma 2.1].) *Suppose that C is an obtuse cone, i.e., $\mathbb{R}_+ C = C$ and $C^{\ominus} := \big\{x \in X \ \big|\ \sup \langle C, x\rangle = 0\big\} \subseteq -C$. Then the reflector $R_C := 2P_C - \mathrm{Id}$ is nonexpansive and* $\operatorname{ran} R_C = \operatorname{Fix} C = C$.

Corollary 2.2

Suppose that C is an obtuse cone and let $\lambda : X \to [1,2]$. Then

$$Q\colon X \to X : x \mapsto \big(1 - \lambda(x)\big)x + \lambda(x)P_C x$$

is a quasi projector of C.

Proof

Since, for every $x \in X$, we have $Q(x) \in [P_C x, R_C x]$ and the result thus follows from Proposition 2.1. □

Example 2.4

Suppose $X = \mathbb{R}^d$ and $C = \mathbb{R}^d_+$. Then R_C is a quasi projector.

Proof Because $C^{\ominus} = -C$, this follows from Corollary 2.2 with $\lambda(x) \equiv 2$. □

Remark 2.1

A quasi projector need not be continuous because we may choose λ in Proposition 2.1 discontinuously.

---

[4] This observation is a due to a referee.

2.3 Fejér Monotone Sequences

Recall that a sequence $(x_n)_{n\in\mathbb{N}}$ in X is Fejér monotone with respect to a nonempty subset S of X iff

$$(\forall s\in S)(\forall n\in\mathbb{N})\quad \|x_{n+1}-s\|\leq\|x_n-s\|.$$

Clearly, every Fejér monotone sequence is bounded.

We will require the following key result.

Fact 2.1

(Raik) Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in X that is Fejér monotone with respect to a subset S of X. If $\operatorname{int} S\neq\varnothing$, then $(x_n)_{n\in\mathbb{N}}$ converges strongly to some point in X and $\sum_{n\in\mathbb{N}}\|x_n-x_{n+1}\|<+\infty$.

Proof See [25] or e.g., [22, Proposition 5.10]. □

2.4 Differentiability

Lemma 2.2

Suppose that X is finite-dimensional, let $f\colon X\to\mathbb{R}$ be convex and Fréchet differentiable such that $\inf f(X)<0$. Then for every $\rho\in\mathbb{R}_{++}$, we have

$$\inf\big\{\|\nabla f(x)\|\;\big|\;x\in\operatorname{ball}(0;\rho)\cap f^{-1}(\mathbb{R}_{++})\big\}>0.$$

Proof

Let $\rho\in\mathbb{R}_{++}$ and assume to the contrary that the conclusion fails. Then there exists a sequence $(x_n)_{n\in\mathbb{N}}$ in $\operatorname{ball}(0;\rho)\cap f^{-1}(\mathbb{R}_{++})$ and a point $x\in\operatorname{ball}(0;\rho)$ such that $x_n\to x$ and $\nabla f(x_n)\to 0$. It follows that $f(x)\geq 0$ and $\nabla f(x)=0$, which is clearly absurd. □

3 Finitely Convergent Cutter Methods

From now on, we assume that

$$(r_n)_{n\in\mathbb{N}} \text{ is a sequence in } \mathbb{R}_{++} \text{ such that } r_n\to 0, \tag{2}$$

that

$$(\eta_n)_{n\in\mathbb{N}} \text{ is a sequence in }]0,2],$$

and that

$$Q_C \text{ is a quasi projector of } C.$$

We further assume that $x_0 \in C$ and that $(x_n)_{n\in\mathbb{N}}$ is generated by

$$(\forall n \in \mathbb{N}) \quad x_{n+1} := \begin{cases} Q_C\big(x_n + \eta_n(U_{r_n}x_n - x_n)\big), & \text{if } x_n \notin \operatorname{Fix} T, \\ x_n, & \text{otherwise.} \end{cases} \tag{3}$$

Note that $(x_n)_{n\in\mathbb{N}}$ lies in C. Also observe that if x_n lies in $\operatorname{Fix} T$, then so does x_{n+1}.

We are now ready for our first main result.

Theorem 3.1

Suppose that $\operatorname{int}(C \cap \operatorname{Fix} T) \neq \varnothing$ *and that* $\sum_{n\in\mathbb{N}} \eta_n r_n = +\infty$. *Then* $(x_n)_{n\in\mathbb{N}}$ *lies eventually in* $C \cap \operatorname{Fix} T$.

Proof

We argue by contradiction. If the conclusion is false, then *no* term of the sequence in $(x_n)_{n\in\mathbb{N}}$ lies in $\operatorname{Fix} T$, i.e., $(x_n)_{n\in\mathbb{N}}$ lies in $X \smallsetminus \operatorname{Fix} T$. By assumption, there exist $z \in C \cap \operatorname{Fix} T$ and $r \in \mathbb{R}_{++}$ and such that $\operatorname{ball}(z;2r) \subseteq C \cap \operatorname{Fix} T$. Hence

$$\big(\forall y \in \operatorname{ball}(z;r)\big) \quad \operatorname{ball}(y;r) \subseteq C \cap \operatorname{Fix} T. \tag{4}$$

Since $r_n \to 0$, there exists $m \in \mathbb{N}$ such that $n \geq m$ implies $r_n \leq r$. Now let $n \geq m$ and $y \in \operatorname{ball}(z;r)$. Using the assumption that Q_C is a quasi projector of C, that $y \in C$, (4), and Corollary 2.1, we obtain

$$\begin{aligned} \|x_{n+1} - y\| &= \big\|Q_C\big(x_n + \eta_n(U_{r_n}x_n - x_n)\big) - y\big\| \\ &\leq \|x_n + \eta_n(U_{r_n}x_n - x_n) - y\| \\ &\leq \|x_n - y\|. \end{aligned}$$

Hence the sequence

$$\big(x_m, x_m + \eta_m(U_{r_m}x_m - x_m), x_{m+1}, x_{m+1} + \eta_{m+1}(U_{r_{m+1}}x_{m+1} - x_{m+1}), x_{m+2}, \ldots\big)$$

is Fejér monotone with respect to $\operatorname{ball}(z;r)$. It follows from Fact 2.1 and Corollary 2.1(iii) that

$$+\infty > \sum_{n\geq m} \eta_n \|x_n - U_{r_n}x_n\| = \sum_{n\geq m} \eta_n\big(r_n + \|x_n - Tx_n\|\big) \geq \sum_{n\geq m} \eta_n r_n,$$

which is absurd because $\sum_{n\in\mathbb{N}} \eta_n r_n = +\infty$. □

We now present our second main result. Compared to Theorem 3.1, we have a less restrictive assumption on $(\operatorname{Fix} T, C)$ but a more restrictive one on the parameters (r_n, η_n). The proof of Theorem 3.2 is more or less implicit in the works by Crombez [4] and Polyak [1]; see Remark 3.1 and Remark 3.2.

Theorem 3.2

Suppose that $C \cap \operatorname{int} \operatorname{Fix} T \neq \varnothing$ and that $\sum_{n\in\mathbb{N}} \eta_n(2-\eta_n)r_n^2 = +\infty$. Then $(x_n)_{n\in\mathbb{N}}$ lies eventually in $C \cap \operatorname{Fix} T$.

Proof

Similarly to the proof of Theorem 3.1, we argue by contradiction and assume the conclusion is false. Then $(x_n)_{n\in\mathbb{N}}$ must lie in $X \smallsetminus \operatorname{Fix} T$. By assumption, there exist $y \in \operatorname{Fix} T$ and $r \in \mathbb{R}_{++}$ such that $\operatorname{ball}(y;r) \subseteq \operatorname{Fix} T$. Because $r_n \to 0$, there exists $m \in \mathbb{N}$ such that $n \geq m$ implies $r_n \leq r$. Let $n \geq m$. Using also the assumption that Q_C is a quasi projector of C and Corollary 2.1(v), we deduce that

$$\begin{aligned}
\|x_{n+1} - y\|^2 &= \big\|Q_C\big(x_n + \eta_n(U_{r_n}x_n - x_n)\big) - y\big\|^2 \\
&\leq \|x_n + \eta_n(U_{r_n}x_n - x_n) - y\|^2 \\
&\leq \|x_n - y\|^2 - \eta_n(2-\eta_n)\big(r_n + \|x_n - Tx_n\|\big)^2 \\
&\leq \|x_n - y\|^2 - \eta_n(2-\eta_n)r_n^2.
\end{aligned}$$

This implies

$$\|x_m - y\|^2 \geq \sum_{n\geq m}\big(\|x_n - y\|^2 - \|x_{n+1} - y\|^2\big) \geq \sum_{n\geq m} \eta_n(2-\eta_n)r_n^2 = +\infty,$$

which contradicts our assumption on the parameters. □

Theorem 3.1 and Theorem 3.2 have various applications. Since every resolvent of a maximally monotone operator is firmly nonexpansive and hence a cutter, we obtain the following result.

Corollary 3.1

Let $A\colon X \rightrightarrows X$ be maximally monotone, suppose that $Q_C = P_C$, that $T = (\mathrm{Id} + A)^{-1}$, and that one of following holds:

(i) $\operatorname{int}(C \cap A^{-1}0) \neq \varnothing$ *and* $\sum_{n\in\mathbb{N}} \eta_n r_n = +\infty$.
(ii) $C \cap \operatorname{int} A^{-1}0 \neq \varnothing$ *and* $\sum_{n\in\mathbb{N}} \eta_n (2-\eta_n) r_n^2 = +\infty$.

Then $(x_n)_{n\in\mathbb{N}}$ lies eventually in $C \cap A^{-1}0$.

Corollary 3.1 applies in particular to finding a constrained critical point of a convex function. When specializing further to a normal cone operator, we obtain the following result.

Example 3.1

(Convex feasibility) Let D be a nonempty, closed and convex subset of X, and suppose that $Q_C = P_C$, that $T = P_D$, and that one of the following holds:

(i) $\operatorname{int}(C \cap D) \neq \varnothing$ and $\sum_{n\in\mathbb{N}} r_n = +\infty$.
(ii) $C \cap \operatorname{int} D \neq \varnothing$ and $\sum_{n\in\mathbb{N}} r_n^2 = +\infty$.

Then the sequence $(x_n)_{n\in\mathbb{N}}$, generated by

$$(\forall n \in \mathbb{N}) \quad x_{n+1} := P_C\left(P_D x_n + r_n \frac{P_D x_n - x_n}{\|P_D x_n - x_n\|}\right)$$

if $x_n \notin D$ and $x_{n+1} := x_n$ if $x_n \in D$, lies eventually in $C \cap D$.

Remark 3.1

(Relationship to Polyak's work) In [1], B.T. Polyak considers random algorithms for solving constrained systems of convex inequalities. Suppose that only one consistent constrained convex inequality is considered. Hence the cutters used are all subgradient projectors (see Example 2.1). Then his algorithm coincides with the one considered in this section and thus is comparable. We note that our Theorem 3.1 is more flexible because Polyak requires $\sum_{n\in\mathbb{N}} r_n^2 = +\infty$ (see [1, Theorem 1 and Section 4.2]) provided that $0 < \inf_{n\in\mathbb{N}} \eta_n \leq \sup_{n\in\mathbb{N}} \eta_n < 2$ while we require only $\sum_{n\in\mathbb{N}} r_n = +\infty$ in this case. Regarding our Theorem 3.2, we note that our proof essentially follows his proof which actually works for cutters — not just subgradient projectors — and under a less restrictive constraint qualification.

Remark 3.2

(Relationship to Crombez's work) In [4], G. Crombez considers asynchronous parallel algorithms for finding a point in the intersection of the fixed point sets of finitely many cutters — without the constraint set C. Again, we consider the case when we are dealing with only one cutter. Then Crombez's convergence result (see [4, Theorem 2.7]) is similar to Theorem 3.2; however, he requires that the radius r of some ball contained in $\operatorname{Fix} T$ be known which may not always be realistic in practical applications.

We will continue our comparison in Section 5. While it is not too difficult to extend Theorem 3.1 and Theorem 3.2 to deal with finitely many cutters, we have opted here for simplicity rather than maximal generality. Instead, we focus in the next section on limiting examples.

We conclude this section with a comment on the proximal point algorithm.

Remark 3.3

(Proximal point algorithm) Suppose that A is a maximally monotone operator on X (see, e.g., [22] for relevant background information) such that $Z := A^{-1}0 \neq \varnothing$. Then its resolvent $J_A := (\mathrm{Id} + A)^{-1}$ is firmly nonexpansive — hence a cutter — with $\operatorname{Fix} J_A = Z$. Let $y_0 \in X$ and set $(\forall n \in \mathbb{N})\ y_{n+1} := J_A y_n$. Then $(y_n)_{n\in\mathbb{N}}$, the sequence generated by the proximal point algorithm, converges weakly to a point in Z. If

$$(\exists \bar{x} \in X) \quad 0 \in \operatorname{int} A\bar{x}, \tag{5}$$

then the convergence is finite (see [26, Theorem 3]). On the other hand, our algorithms impose that $\operatorname{int} \operatorname{Fix} T \neq \varnothing$, i.e.,

$$(\exists \bar{x} \in X) \quad \bar{x} \in \operatorname{int} A^{-1}0. \tag{6}$$

(Note that (5) and (6) are independent: If A is $\partial\|\cdot\|$, then $0 \in \operatorname{int} A0$ yet $\operatorname{int} A^{-1}0 = \varnothing$. And if $A = \nabla d^2_{\mathrm{ball}(0;1)}$, then $0 \in \operatorname{int} A^{-1}0$ while $A = 2(\mathrm{Id} - P_{\mathrm{ball}(0;1)})$ is single-valued.)

4 Limiting Examples

In this section, we collect several examples that illustrate the boundaries of the theory. We start by showing that the conclusion of Theorem 3.1 and Theorem 3.2 both may fail to hold if the divergent-series condition is not satisfied.

Example 4.1

(Divergent-series condition is important) Suppose that $X = C = \mathbb{R}$, that

$$f\colon \mathbb{R} \to \mathbb{R}\colon x \mapsto x^2 - 1,$$

and that $T = G_f$ is the subgradient projector associated with f. Suppose that $x_0 > 1$, set $r_{-1} := x_0 - 1 > 0$ and $(\forall n \in \mathbb{N})\ r_n := r_{n-1}^2/(4(1 + r_{n-1}))$. Then $(r_n)_{n\in\mathbb{N}}$ lies in $\mathbb{R}_{++}$, $r_n \to 0$, and $\sum_{n\in\mathbb{N}} r_n < +\infty$ and hence $\sum_{n\in\mathbb{N}} r_n^2 < +\infty$. However, the sequence $(x_n)_{n\in\mathbb{N}}$ generated by (3) lies in $]1,+\infty[$ and hence does not converge finitely to a point in $\operatorname{Fix} T = [-1,1]$. Furthermore, the iteration $(\forall n \in \mathbb{N})\ y_{n+1} = Ty_n$ converges to some point in $\operatorname{Fix} T$, but not finitely when $y_0 \notin \operatorname{Fix} T$.

Proof

Clearly $\operatorname{Fix} T = [-1,1]$. Observe that $(\forall n \in \mathbb{N})\ 0 < r_n \leq (1/4)r_{n-1} \leq (1/4)^{n+1} r_{-1}$. It follows that $r_n \to 0$ and that $\sum_{n\in\mathbb{N}} r_n$ and $\sum_{n\in\mathbb{N}} r_n^2$ are both convergent series. Now suppose that $r_{n-1} = x_n - 1 > 0$ for some $n \in \mathbb{N}$. It then follows from Example 2.2 that

$$x_{n+1} = \frac{x_n}{2} + \frac{1}{2x_n} - r_n = \frac{(x_n-1)^2}{2x_n} + 1 - r_n = \frac{r_{n-1}^2}{2(1+r_{n-1})} + 1 - r_n = r_n + 1.$$

Hence, by induction, $(\forall n \in \mathbb{N})\ x_n = 1 + r_{n-1}$ and therefore $x_n \uparrow 1$.

As for the sequence $(y_n)_{n\in\mathbb{N}}$, it is follows from Polyak's seminal work (see [12]) that $(y_n)_{n\in\mathbb{N}}$ converges to some point in $\operatorname{Fix} T$. However, by e.g., [20, Proposition 9.9], $(y_n)_{n\in\mathbb{N}}$ lies outside $\operatorname{Fix} T$ whenever y_0 does. ☐

The next example illustrates that we cannot expect finite convergence if the interior of $\operatorname{Fix} T$ is empty, in the context of Theorem 3.1 and Theorem 3.2.

Example 4.2

(Nonempty-interior condition is important) Suppose that $X = C = \mathbb{R}$, and that $T = G_f$

is the subgradient projector associated with $f\colon \mathbb{R}\to\mathbb{R}\colon x\mapsto x^2$. Then $\operatorname{Fix} T=\{0\}$ and hence $\operatorname{int}\operatorname{Fix} T=\varnothing$. Set $x_0:=1/2$, and set $(\forall n\in\mathbb{N})$ $w_n:=(n+1)^{-1/2}$ and $r_n=w_n$ if $U_{w_n}x_n\neq 0$ and $r_n=2w_n$ if $U_{w_n}x_n=0$. Then $r_n\to 0$ and $\sum_{n\in\mathbb{N}} r_n^2=+\infty$. The sequence $(x_n)_{n\in\mathbb{N}}$ generated by (3) converges to 0 but not finitely.

Proof

The statements concerning $(r_n)_{n\in\mathbb{N}}$ are clear. It follows readily from the definition that $(\forall x\in\mathbb{R})(\forall r\in\mathbb{R}_+)$ $Tx=x/2$ and $U_r x=x/2-r\operatorname{sgn}(x)$. Since $x_0=1/2$, $w_0=1$, $U_1x_0=-3/4\neq 0$, and $r_0=w_0=1$, it follows that $0<|x_0/2|<r_0$. We now show that for every $n\in\mathbb{N}$,

$$0<|x_n/2|<r_n. \tag{7}$$

This is clear for $n=0$. Now assume (7) holds for some $n\in\mathbb{N}$.

Case 1: $|x_n|=2w_n$.

Then $U_{w_n}x_n=x_n/2-\operatorname{sgn}(x_n)w_n=0$. Hence $r_n=2w_n$ and thus

$$x_{n+1}=U_{r_n}x_n=x_n/2-2w_n\operatorname{sgn}(x_n)=\operatorname{sgn}(x_n)w_n-2w_n\operatorname{sgn}(x_n)=-\operatorname{sgn}(x_n)w_n.$$

Thus $0<|x_{n+1}/2|=w_n/2=1/(2\sqrt{n+1})<1/\sqrt{n+2}=w_{n+1}\leq r_{n+1}$, which yields (7) with n replaced by $n+1$.

Case 2: $|x_n|\neq 2w_n$.

Then $U_{w_n}x_n=x_n/2-\operatorname{sgn}(x_n)w_n\neq 0$. Hence $r_n=w_n$ and thus

$$x_{n+1}=U_{r_n}x_n=x_n/2-r_n\operatorname{sgn}(x_n).$$

It follows that $|x_{n+1}|=r_n-|x_n/2|>0$. Hence $0<|x_{n+1}/2|$ and also

$$|x_{n+1}|<r_n=w_n<2w_{n+1}\leq 2r_{n+1}.$$

Again, this is (7) with n replaced by $n+1$.

It follows now by induction that (7) holds for every $n\in\mathbb{N}$. □

We now illustrate that when $\operatorname{Fix} T=\varnothing$, then $(x_n)_{n\in\mathbb{N}}$ may fail to converge.

Example 4.3

Suppose that $X=C=\mathbb{R}$, that $f\colon\mathbb{R}\to\mathbb{R}\colon x\mapsto x^2+1$, and that $T=G_f$ is the subgradient projector associated with f. Let $y_0\in\mathbb{R}$ and suppose that $(\forall n\in\mathbb{N})$ $y_{n+1}:=Ty_n$.

Then $(y_n)_{n\in\mathbb{N}}$ is either not well defined or it diverges. Suppose that $x_0 > 1/\sqrt{3}$, set $k_0 := x_0 - 1/\sqrt{3} > 0$ and $(\forall n\in\mathbb{N})\ k_{n+1} := \sqrt{(n+1)/(n+2)}k_n$. Suppose that

$$(\forall n\in\mathbb{N})\quad r_n := \frac{1}{2}\left(\sqrt{3} + 2k_{n+1} + k_n - \frac{1}{k_n + 1/\sqrt{3}}\right).$$

Then $r_n \downarrow 0$ and $\sum_{n\in\mathbb{N}} r_n^2 = +\infty$. Moreover, the sequence $(x_n)_{n\in\mathbb{N}}$ generated by (3) diverges.

Proof

Clearly, $\operatorname{Fix} T = \varnothing$ and one checks that

$$(\forall r\in\mathbb{R}_+)(\forall x\in\mathbb{R}\smallsetminus\{0\})\quad U_r x = \frac{x}{2} - \frac{1}{2x} - r\operatorname{sgn}(x). \tag{8}$$

If some $y_n = 0$, then the sequence $(y_n)_{n\in\mathbb{N}}$ is not well defined.

Case 1: $(\exists n\in\mathbb{N})\ y_n = 1/\sqrt{3}$.

Then

$$x_{n+1} = Tx_n = U_0 x_n = x_n/2 - 1/(2x_n) = -1/\sqrt{3} = -x_n$$

and similarly $x_{n+2} = -x_{n+1} = x_n$. Hence the sequence eventually oscillates between $1/\sqrt{3}$ and $-1/\sqrt{3}$.

Case 2: $(\exists n\in\mathbb{N})\ |y_n| = 1$.

Then $y_{n+1} = 0$ and the sequence is not well defined.

Case 3: $(\forall n\in\mathbb{N})\ |y_n| \notin \{1, 1/\sqrt{3}\}$.

Using the Arithmetic Mean–Geometric Mean inequality, we obtain

$$|y_{n+1} - y_n| = \left|\frac{y_n}{2} - \frac{1}{2y_n} - y_n\right| = \frac{1}{2}\left|y_n + \frac{1}{y_n}\right| = \frac{1}{2}\left(|y_n| + \frac{1}{|y_n|}\right) \geq 1$$

for every $n\in\mathbb{N}$. Therefore, $(y_n)_{n\in\mathbb{N}}$ is divergent or not well defined.

We now turn to the sequence $(x_n)_{n\in\mathbb{N}}$. Observe that

$$0 < k_n = \sqrt{n/(n+1)}k_{n-1} = \cdots = k_0/\sqrt{n+1} \downarrow 0$$

and hence $(k_n)_{n\in\mathbb{N}}$ is strictly decreasing. It follows that $r_n \downarrow 0$ and that

$$r_n > (2k_{n+1} + k_n)/2 > 3k_{n+1}/2 = 3k_0/(2\sqrt{n+2}).$$

Thus, $\sum_{n\in\mathbb{N}} r_n^2 = +\infty$. Next, (8) yields

$$\begin{aligned} x_1 &= \frac{x_0}{2} - \frac{1}{2x_0} - r_0 \\ &= \frac{k_0 + 1/\sqrt{3}}{2} - \frac{1}{2\big(k_0 + 1/\sqrt{3}\big)} - \frac{1}{2}\left(\sqrt{3} + 2k_1 + k_0 - \frac{1}{k_0 + 1/\sqrt{3}}\right) \\ &= -\frac{1}{\sqrt{3}} - k_1. \end{aligned}$$

Hence $x_1 < 0$ and we then see analogously that $x_2 = 1/\sqrt{3} + k_2 > 0$. We inductively obtain

$$(\forall n \in \mathbb{N}) \quad 0 < x_{2n} = \frac{1}{\sqrt{3}} + k_{2n} \ \text{ and } \ 0 > x_{2n+1} = -\frac{1}{\sqrt{3}} - k_{2n+1}.$$

It follows that $(-1)^n x_n \to 1/\sqrt{3}$; therefore, $(x_n)_{n\in\mathbb{N}}$ is divergent. □

5 Comparison

In this section, we assume for notational simplicity[5] that

$$f\colon X \to \mathbb{R} \text{ is convex and Fréchet differentiable with } \big\{x \in X \;\big|\; f(x) \leq 0\big\} \neq \varnothing$$

and that

$$T = G_f\colon X \to X\colon x \mapsto \begin{cases} x - \dfrac{f(x)}{\|\nabla f(x)\|^2}\nabla f(x), & \text{if } f(x) > 0, \\ x, & \text{otherwise} \end{cases}$$

is the associated subgradient projector (see Example 2.1). Then (1) turns into

$$U_r x = \begin{cases} x - \dfrac{f(x) + r\|\nabla f(x)\|}{\|\nabla f(x)\|^2}\nabla f(x), & \text{if } f(x) > 0, \\ x, & \text{otherwise} \end{cases}$$

and (3) into

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = \begin{cases} Q_C\left(x_n - \eta_n \dfrac{f(x_n) + r_n\|\nabla f(x_n)\|}{\|\nabla f(x_n)\|^2}\nabla f(x_n)\right), & \text{if } f(x_n) > 0, \\ x_n, & \text{otherwise.} \end{cases} \tag{9}$$

[5] If we replace Fréchet differentiability by mere continuity, then we may consider a selection of the subdifferential operator ∂f instead.

In the algorithmic setting of Section 3, Polyak uses $\eta \equiv \eta_n \in]0,2,[$ (e.g. $\eta = 1.8$; see [1, Section 4.3]). In the present setting, his framework requires $\sum_{n\in\mathbb{N}} r_n^2 = +\infty$. When $C = X$, one also has the following similar yet different update formula

$$(\forall n \in \mathbb{N}) \quad y_{n+1} = \begin{cases} y_n - \eta_n \dfrac{f(y_n)+\varepsilon_n}{\|\nabla f(y_n)\|^2}\nabla f(y_n), & \text{if } f(y_n) > 0, \\ y_n, & \text{otherwise,} \end{cases} \tag{10}$$

where $0 < \inf_{n\in\mathbb{N}} \eta_n \le \sup_{n\in\mathbb{N}} \eta_n < 2$ and $(\varepsilon_n)_{n\in\mathbb{N}}$ is a strictly decreasing sequence in $\mathbb{R}_{++}$ with $\sum_{n\in\mathbb{N}} \varepsilon_n = +\infty$. In this setting, this is also known as the *Modified Cyclic Subgradient Projection Algorithm (MCSPA)*, which finds its historical roots in works by Fukushima [27], by De Pierro and Iusem [28], and by Censor and Lent [7]; see also [3,29–31] for related works. Note that MCSPA requires the existence of a *Slater point*, i.e., $\inf f(X) < 0$, which is more restrictive than our assumptions (consider, e.g., the squared distance to the unit ball). Let us now link the assumption on the parameters of the MCSPA (10) to (9).

Proposition 5.1

Suppose that $X = C$ *is finite-dimensional,* $\inf f(X) < 0$, $\eta_n \equiv 1$, $\sum_{n\in\mathbb{N}} r_n = +\infty$ *(recall* (2)*), and* $(\forall n \in \mathbb{N})$ $\varepsilon_n = r_n\|\nabla f(x_n)\| > 0$. *Then* $\varepsilon_n \to 0$ *and* $\sum_{n\in\mathbb{N}} \varepsilon_n = +\infty$.

Proof

Corollary 2.1(iv) implies that $(x_n)_{n\in\mathbb{N}}$ is bounded. Because ∇f is continuous, we obtain that $\sigma := \sup_{n\in\mathbb{N}} \|\nabla f(x_n)\| < +\infty$. By Lemma 2.2, there exists $\alpha \in \mathbb{R}_{++}$ such that if $f(x_n) > 0$, then $\|\nabla f(x_n)\| \ge \alpha$. Hence

$$(\forall n \in \mathbb{N}) \quad f(x_n) > 0 \;\Rightarrow\; 0 < \alpha r_n \le \|\nabla f(x_n)\| r_n = \varepsilon_n \le \sigma r_n,$$

and therefore $\sum_{n\in\mathbb{N}} \varepsilon_n = +\infty$. □

The following example shows that our assumptions are independent of those on the MCSPA.

Example 5.1

Suppose that $X = C = \mathbb{R}$, that $f\colon \mathbb{R} \to \mathbb{R}\colon x \mapsto x^2 - 1$, that $r_n = (n+1)^{-1}$ if n is even and $r_n = n^{-1/2}$ if n is odd, and that $\eta_n \equiv 1$. Clearly, $r_n \to 0$ and $\sum_{n\in\mathbb{N}} r_n^2 = +\infty$. However, $(\varepsilon_n)_{n\in\mathbb{N}} := (r_n|f'(x_n)|)_{n\in\mathbb{N}}$ is not strictly decreasing.

Proof

The sequence $(x_n)_{n\in\mathbb{N}}$ is bounded. Suppose that $f(x_n) > 0$ for some $n \in \mathbb{N}$. By Example 2.2,

$$x_{n+1} = U_{r_n} x_n = \frac{x_n}{2} + \frac{1}{2x_n} - r_n \operatorname{sgn}(x_n). \tag{11}$$

Assume that n is even, say $n = 2m$, where $m \geq 2$, and that $1 < x_{2m} < (2m+1)/2$. Then $x_{2m} > 2x_{2m}/\sqrt{2m+1}$ and

$$\varepsilon_{2m} = r_{2m}|f'(x_{2m})| = 2r_{2m}x_{2m} = \frac{2x_{2m}}{2m+1}.$$

Hence, using (11),

$$x_{2m+1} = \frac{x_{2m}}{2} + \frac{1}{2x_{2m}} - r_{2m} > \frac{x_{2m}}{2} + \frac{1}{2m+1} - \frac{1}{2m+1} = \frac{x_{2m}}{2},$$

and therefore

$$2x_{2m+1} > x_{2m} > \frac{2x_{2m}}{\sqrt{2m+1}}.$$

Thus $\varepsilon_{2m+1} = r_{2m+1}|f'(x_{2m+1})| = 2r_{2m+1}x_{2m+1}$. It follows that

$$\varepsilon_{2m+1} = \frac{2x_{2m+1}}{\sqrt{2m+1}} > \frac{2x_{2m}}{2m+1} = \varepsilon_{2m}$$

and the proof is complete. □

6 Perspectives

Suppose that $X = \mathbb{R}$ and that $f\colon X \to \mathbb{R}\colon x \mapsto x^2 - 1$. Let T be the subgradient projector associated with f and assume that $C = X$. We chose 100 randomly chosen starting points in the interval $[1, 10^6]$. In the following table, we record the performance of the algorithms; here (r_n, η_n) signals that (9) was used, while ε_n points to (10) with $\eta_n \equiv 1$. Mean and median refer to the number of iterations until the current iterate was 10^{-6} feasible.

Now let us instead consider $f\colon X \to \mathbb{R}\colon x \mapsto 100x^2 - 1$. The corresponding data are in the following table.

We observe that the performance of the algorithms clearly depends on the step lengths r_n and ε_n, on the relaxation parameter η_n, and on the underlying objective function f; however, *the precise nature of this dependence is rather unclear*. It would thus

Algorithm for x^2-1	Mean	Median
$(r_n,\eta_n)=\big(1/(n+1),1\big)$	11.49	13
$(r_n,\eta_n)=\big(1/(n+1),2\big)$	2	2
$(r_n,\eta_n)=\big(1/\sqrt{n+1},1\big)$	10.83	12
$(r_n,\eta_n)=\big(1/\sqrt{n+1},2\big)$	2	2
$\varepsilon_n=1/(n+1)$	11.81	13
$\varepsilon_n=1/\sqrt{n+1}$	12.19	13

Algorithm for $100x^2-1$	Mean	Median
$(r_n,\eta_n)=\big(1/(n+1),1\big)$	13.29	14
$(r_n,\eta_n)=\big(1/(n+1),2\big)$	12	12
$(r_n,\eta_n)=\big(1/\sqrt{n+1},1\big)$	17.52	19
$(r_n,\eta_n)=\big(1/\sqrt{n+1},2\big)$	105	105
$\varepsilon_n=1/(n+1)$	15.27	16
$\varepsilon_n=1/\sqrt{n+1}$	15.76	17

be interesting to perform numerical experiments on a wide variety of problems and parameter choices with the goal to *obtain guidelines in the choice of algorithms and parameters* for the user.

Another avenue for future research is to *construct a broad framework* that encompasses the present as well as previous related finite convergence results (see references in Section 5).

7 Conclusions

We have obtained new and more general finite convergence results for a class of algorithms based on cutters. A key tool was Raik's result on Fejér monotone sequences (Fact 2.1).

Acknowledgements The authors thank two anonymous referees for careful reading, constructive comments, and for bringing additional references to our attention. The authors also thank Jeffrey Pang for helpful discussions and for pointing out additional references.

HHB was partially supported by a Discovery Grant and an Accelerator Supplement of the Natural Sciences and Engineering Research Council of Canada (NSERC) and by the Canada Research Chair Program.

CW was partially supported by a grant from Shanghai Municipal Commission for Science and Technology (13ZR1455500). XW was partially supported by a Discovery Grant of NSERC. JX was partially supported by NSERC grants of HHB and XW.

References

1. Polyak, B.T.: Random algorithms for solving convex inequalities. In Butnariu, D., Censor, Y., Reich, S. (eds.): Inherently Parallel Algorithms in Feasibility and Optimization and their Applications, pp. 409–422, Elsevier Science Publishers, Amsterdam, The Netherlands (2001)
2. Cegielski, A.: Iterative Methods for Fixed Point Problems in Hilbert Spaces, Lecture Notes in Mathematics 2057, Springer-Verlag, Berlin, Heidelberg, Germany (2012)
3. Censor, Y., Chen, W., Pajoohesh, H.: Finite convergence of a subgradient projection method with expanding controls, Applied Mathematics and Optimization 64, 273–285 (2011)
4. Crombez, G.: Finding common fixed points of a class of paracontractions, Acta Mathematica Hungarica 103, 233–241 (2004)
5. Censor, Y., Zenios, S.A.: Parallel Optimization, Oxford University Press (1997)
6. Polyak, B.T.: Introduction to Optimization, Optimization Software, New York, NY, USA (1987)
7. Censor, Y., Lent, A.: Cyclic subgradient projections, Mathematical Programming 24, 233–235 (1982)
8. Censor, Y., Segal, A.: Sparse string-averaging and split common fixed points. In: Leizarowitz, A., Mordukhovich, B.S., Shafrir, I., Zaslavski, A.J. (eds): Nonlinear Analysis and Optimization I Contemporary Mathematics, vol. 513, pp. 125–142 (2010)
9. Combettes, P.L.: The foundations of set theoretic estimation, Proceedings of the IEEE 81, 182–208 (1993)
10. Combettes, P.L.: Convex set theoretic image recovery by extrapolated iterations of parallel subgradient projections, IEEE Transactions on Image Processing 6, 493–506 (1997)

11. Combettes, P.L., Luo, J.: An adaptive level set method for nondifferentiable constrained image recovery, IEEE Transactions on Image Processing 11, 1295–1304 (2002)

12. Polyak, B.T.: Minimization of unsmooth functionals, U.S.S.R. Computational Mathematics and Mathematical Physics 9, 14–29 (1969) (The original version appeared in Akademija Nauk SSSR. Žurnal Vyčislitel' noĭ Matematiki i Matematičeskoĭ Fiziki 9 (1969), 509–521.)

13. Slavakis, K., Yamada, I.: The adaptive projected subgradient method constrained by families of quasi-nonexpansive mappings and its application to online learning, SIAM Journal on Optimization 23, 126–152 (2013)

14. Yamada, I., Ogura, N.: Adaptive projected subgradient method for asymptotic minimization of sequence of nonnegative convex functions, Numerical Functional Analysis and Optimization 25, 593–617 (2004)

15. Yamada, I., Ogura, N.: Hybrid steepest descent method for variational inequality problem over the fixed point set of certain quasi-nonexpansive mappings, Numerical Functional Analysis and Optimization 25, 619–655 (2004)

16. Yamada, I., Slavakis, K., Yamada, K.: An efficient robust adaptive filtering algorithm based on parallel subgradient projection techniques, IEEE Transactions on Signal Processing 50, 1091–1101 (2002)

17. Yamagishi, M., Yamada, I.: A deep monotone approximation operator based on the best quadratic lower bound of convex functions, IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences E91–A, 1858–1866 (2008)

18. Bauschke, H.H., Borwein, J.M.: On projection algorithms for solving convex feasibility problems, SIAM Review 38, 367–426 (1996)

19. Bauschke, H.H., Combettes, P.L.: A weak-to-strong convergence principle for Fejér-monotone methods in Hilbert space, Mathematics of Operations Research 26, 248–264 (2001)

20. Bauschke, H.H., Wang, C., Wang, X., Xu, J.: On subgradient projectors, `http://arxiv.org/abs/1403.7135v1` (March 2014)

21. Pauwels, B.: Subgradient projection operators, `http://arxiv.org/abs/1403.7237v1` (March 2014)

22. Bauschke, H.H., Combettes, P.L.: Convex Analysis and Monotone Operator Theory in Hilbert Spaces, Springer (2011)

23. Bauschke, H.H.: Projection Algorithms and Monotone Operators, PhD thesis, Simon Fraser University, Burnaby, BC, Canada, August 1996.

24. Bauschke, H.H., Kruk, S.G.: Reflection-projection method for convex feasibility problems with an obtuse cone, Journal of Optimization Theory and Applications 120, 503–531 (2004)

25. Raik, E.: A class of iterative methods with Fejér-monotone sequences, Eesti NSV Teaduste Akadeemia Toimetised. Füüsika-Matemaatika 18, 22–26 (1969)

26. Rockafellar, R.T.: Monotone operators and the proximal point algorithm, SIAM Journal on Control and Optimization 14, 877–898 (1976)

27. Fukushima, M.: A finite convergent algorithm for convex inequalities, IEEE Transactions on Automatic Control 27, 1126–1127 (1982)

28. De Pierro, A.R., Iusem, A.N.: A finitely convergent "row-action" method for the convex feasibility problem, Applied Mathematics and Optimization 17, 225–235 (1988)

29. Iusem, A.N., Moledo, L.: A finitely convergent method of simultaneous subgradient projections for the convex feasibility problem, Matemática Aplicada e Computacional 5, 169–184 (1986)

30. Iusem, A.N., Moledo, L.: On finitely convergent iterative methods for the convex feasibility problem, Boletim da Sociedade Brasileira de Matemática 18, 11–18 (1987)

31. Pang, C.H.J.: Finitely convergent algorithm for nonconvex inequality problems, `http://arxiv.org/abs/1405.7280v1` (May 2014)