A GENERALIZED BLOCK-ITERATIVE PROJECTION METHOD FOR THE COMMON FIXED POINT PROBLEM INDUCED BY CUTTERS

YAIR CENSOR, DANIEL REEM, AND MAROUN ZAKNOON

ABSTRACT. The block-iterative projections (BIP) method of Aharoni and Censor [Block-iterative projection methods for parallel computation of solutions to convex feasibility problems, Linear Algebra and its Applications 120, (1989), 165–175] is an iterative process for finding asymptotically a point in the nonempty intersection of a family of closed convex subsets. It employs orthogonal projections onto the individual subsets in an algorithmic regime that uses "blocks" of operators and has great flexibility in constructing specific algorithms from it. We extend this algorithmic scheme to handle a family of continuous cutter operators and to find a common fixed point of them. Since the family of continuous cutters includes several important specific operators, our generalized scheme, which ensures global convergence and retains the flexibility of BIP, can handle, in particular, metric (orthogonal) projectors and continuous subgradient projections, which are very important in applications. We also allow a certain kind of adaptive perturbations to be included, and along the way we derive a perturbed Fejér monotonicity lemma which is of independent interest.

1. Introduction

1.1. **Background and contributions.** Given a finite family of $m \in \mathbb{N}$ nonempty closed convex subsets $C_1, C_2, \ldots, C_m$ of the n-dimensional Euclidean space $\mathbb{R}^n$, the *convex feasibility problem* (CFP) is to find a point in their intersection $C := \cap_{j=1}^{m} C_j$, assuming that the intersection is nonempty. This well-known problem has applications in many theoretical and real-world scenarios, such as image reconstruction from projections, data compression, radiation therapy treatment planning, signal processing, sensor network source localization, the solution of systems of linear or nonlinear inequalities induced by convex functions (since the solution of a system of inequalities is a point in the intersection of the level-sets of the functions which induce these inequalities), as well as in many other areas, as can be seen in, e.g., Cegielski's book [**12**, p. 23]; see also [**30**] and some of the references therein for the application of the CFP for solving optimization problems. Additional details about

Date: April 20, 2022.
2020 *Mathematics Subject Classification.* 47H10, 90C31, 49K40, 90C30, 90C59.
Key words and phrases. block iterative projection (BIP), common fixed point problem, cutter, perturbation, weight function.

the CFP, including various algorithmic schemes for solving it and related references can be found in [**1**, **4**, **8**, **12**, **19**, **23**, **24**].

One of the methods for solving the CFP, which is of special importance to our paper, is the BIP (Block-Iterative Projections) method of Aharoni and Censor [**1**]. In a nutshell, each iteration of BIP is a relaxed convex combination of the orthogonal projections onto the given subsets $C_1, C_2, \ldots, C_m$, where the combination's weights themselves are dynamic, namely they may depend on the iteration index and can vary from iteration to iteration. This method is rather flexible since particular instances of it are fully sequential iterations with repetitive controls, fully simultaneous iterations, and block iterative iterations.

A more general problem is the *common fixed point problem* (CFPP) of finding a point in the intersection of the fixed point sets of a finite family of operators $T_i : \mathbb{R}^n \to \mathbb{R}^n$, $i \in \{1, 2, \ldots, m\}$. This problem reduces to the CFP when for each $i \in \{1, 2, \ldots, m\}$, the operator T_i is the orthogonal projection onto C_i. Many methods have been devised to solve the CFPP, under various settings: see, for example, [**12**, **20**, **39**, **41**] and some of the references therein, as well as Algorithm **3.1** and Examples **3.2**–**3.4** below.

In this paper we consider the CFPP in the case where all operators T_i are continuous cutters. The class of continuous cutters is quite wide and includes, among others, subgradient projections of differentiable convex functions having nonempty zero-level-sets, resolvents of maximally monotone operators, and orthogonal projections onto nonempty, closed and convex subsets of the space.

Cutters were introduced by Bauschke and Combettes in [**3**] and by Combettes in [**26**] (not yet under the name "cutters" though). More details regarding this class of operators, as well as a short history and other names of it, can be found in Section **2** below. Our goal is to solve the CFPP asymptotically, namely to construct an iterative sequence which converges to a point in the common fixed point set of the given family of cutters.

We solve the CFPP using Algorithm **3.1** which is introduced below. This algorithm is a generalization of the BIP method, mentioned above, where the generalization is expressed in the use of continuous cutters instead of orthogonal projections and in the permission of certain adaptive perturbations to appear in the iterative scheme (see Section **4** for other variants of the BIP method). As can be seen in both the formulation of Algorithm **3.1** and in Examples **3.2**–**3.5** below, Algorithm **3.1** retains the flexibility of BIP not only because the users have freedom regarding the relaxation parameters and the weight functions, but also because Algorithm **3.1** allows, as particular cases, fully sequential iterations with a repetitive control, fully simultaneous iterations, and block iterative iterations. Consequently, Algorithm **3.1** can be adapted naturally to both serial and parallel computational architectures.

We show in Theorem **5.1** below that the iterative sequence generated by Algorithm **3.1** always converges globally to a generalized solution of the CFPP, namely to a point in a set which contains the common fixed point set of the given family of cutters. As we explain after this theorem (in Remark **5.7**), under mild conditions both sets coincide, and so, in a wide class of scenarios, the iterative sequence generated by our algorithmic scheme converges globally to a common fixed point of the given family of cutters. Along the way we obtain a result of independent interest, namely the apparently new Lemma **5.2** below, related to Fejér-monotonicity in a perturbed form; this lemma shows that cutters, as well as relaxed versions of them, are not only quasi-nonexpansive, but rather their quasi-nonexpansivity is preserved under small perturbations of a certain type.

As mentioned above, Algorithm **3.1** allows perturbations of a certain kind, but still converges (globally). In other words, our algorithmic scheme exhibits a certain kind of resiliency, namely, it is "perturbation resilient". The perturbations in Algorithm **3.1** may appear as a result of noise, computational errors, and so on. These perturbations may also be generated actively by the user as part of the application of the "superiorization methodology" (SM). In the SM, in contrast to the case in which the perturbations are unknown to the user (frequently only their magnitude can be estimated), the goal is to harness the permissible perturbations in order to obtain solutions, or generalized solutions, which are superior with respect to some given cost function, over (generalized) solutions which would be obtained without the generated perturbations. More details regarding the superiorization methodology, in its classical form, can be found in the initial papers [**17**, **27**] and the survey papers [**14**,**31**]; a re-examination of this methodology, as well as a significant extension of its scope and an extensive list of related references, can be found in [**38**, Section 4]; a continuously updated bibliographical list of references related to the superiorization methodology can be found on-line in [**15**].

A final word about potential computational advantages. This paper is a theoretical work. Comparative computational performance of BIP-for-cutters algorithms proposed and studied here can really be made only with exhaustive testing of the many possible specific variants permitted by the general schemes and their various user-chosen parameters. The computational advantages of the BIP algorithmic structures have been shown in the past for algorithms that use orthogonal projections rather than other cutter operators in many publications. For example, the work on proton computed tomography (pCT) in [**35**] employs very efficiently a parallel code that uses a version of a block-iterative algorithm called "diagonally-relaxed orthogonal projections" (DROP), presented in [**18**]. See, e.g., also the recent paper on stochastic block projection algorithms by Necoara [**37**]. It is plausible to hypothesize that since the BIP algorithmic structure and the cutter operators [**3**], see also [**6**], have been demonstrated to be computationally useful separately, then so might very well be their combination in the BIP-for-cutters scheme presented here. Admittedly, such practical questions should be resolved in future experimental works, preferably within the context of a significant real-world application.

1.2. **Paper layout.** Section **2** presents the notation and basic definitions used throughout the paper. Section **3** presents the generalized BIP method, namely Algorithm **3.1**, and further elaborates on it. A few variants of the BIP method are discussed in Section **4**, where we compare them and their associated convergence results to Algorithm **3.1** and Theorem **5.1**. The convergence theorem (Theorem **5.1**) and its proof appear in Section **5**.

2. Notation and basic definitions

Given $n \in \mathbb{N}$, where $\mathbb{N}$ is the set of positive integers, let $X := \mathbb{R}^n$ endowed with the Euclidean inner product $\langle \cdot, \cdot \rangle$ and the corresponding Euclidean norm $\|\cdot\|$. We denote by $d(x, C)$ the distance between $x \in X$ and a nonempty set $C \subseteq X$, namely, $d(x, C) := \inf\{\|x - c\| \mid c \in C\}$, and denote by $B[x, r]$ the closed ball with center x and radius $r \in [0, \infty]$ (of course, $B[x, r] = X$ if $r = \infty$ and $B[x, r] = \{x\}$ if $r = 0$). The identity operator is denoted by Id, namely, $Id(x) = x$ for all $x \in X$. We use the convention that the sum over the empty set is zero. Finally, for each operator $U : X \to X$, the set $\mathrm{Fix}(U) := \{x \in X : U(x) = x\}$ stands for the set of all fixed points of U.

Given an operator $T : X \to X$ and a nonempty set $S \subseteq X$, we say that T is a separator of S provided

$$\langle x - T(x), q - T(x) \rangle \leq 0, \ \forall q \in S, x \in X. \tag{2.1}$$

In particular, if $\mathrm{Fix}(T)$ is nonempty and T is a separator of $S := \mathrm{Fix}(T)$, then we say that T is a *cutter*. Given $m \in \mathbb{N}$, denote $I := \{1, 2, \ldots, m\}$. A weight function with respect to I is a function $w : I \to [0, 1]$ which satisfies $\sum_{i \in I} w(i) = 1$. Given a weight function $w : I \to [0, 1]$ and a family $(T_i)_{i \in I}$ of cutters, let T_w be the operator $T_w : X \to X$ defined by $T_w(x) := \sum_{i \in I} w(i) T_i(x)$, for each $x \in X$.

The class of cutter operators was introduced in [**3**] and [**26**] under the name "the class $\mathcal{T}$". Other names appear in the literature, for instance "directed operators" [**21**, **40**]. The name "cutter" was first suggested in [**13**]. The reason behind this name is that for each point x in the space, which is not a fixed point of T, the operator T induces a hyperplane (the one which is orthogonal to the vector $T(x) - x$ and passes through $T(x)$) that "cuts" the space into two half-spaces: one of which contains the fixed point set of T and the other contains x. Various examples of cutters can be found in [**3**, **26**] and (explicitly or implicitly) in [**5**] and [**12**]. A relatively recent work on cutters is [**6**].

In particular, the following operators are continuous cutters:

(1) The subgradient projection of a (Fréchet) differentiable convex function $f : X \to \mathbb{R}$ whose zero-level-set $\{x \in X \mid f(x) \leq 0\}$ is nonempty. Here

$$T(x) := \begin{cases} x - \dfrac{f(x)}{\|\nabla f(x)\|^2} \nabla f(x), & \text{if } f(x) > 0, \\ x, & \text{if } f(x) \leq 0, \end{cases} \tag{2.2}$$

where $\nabla f(x)$ is, for each $x \in X$, the gradient of f at x. The subgradient inequality and the assumption that $\{x \in X \mid f(x) \leq 0\} \neq \emptyset$ imply that T is well-defined, namely, that $\nabla f(x) \neq 0$ if $f(x) > 0$. Since f is convex and differentiable, it is actually continuously differentiable [**33**, Remark 6.2.6,

p. 202]. Hence from [**5**, Proposition 29.41(ix), p. 553] it follows that T is continuous, from [**12**, Corollary 4.2.6, p. 146] it follows that T is a cutter, and from [**12**, Corollary 4.2.5, p. 145] it follows that the fixed point set of T is $\{x \in \mathbb{R}^n \mid f(x) \leq 0\}$. See also [**7**];

(2) Any firmly nonexpansive (FNE) operator, namely, $\|Tx - Ty\|^2 \leq \langle Tx - Ty, x - y\rangle$ for all $x, y \in X$. Indeed, it is well-known that every firmly nonexpansive is nonexpansive (see, for example, [**12**, Theorem 2.2.10(v) or Theorem 2.2.10(vi), p. 70]). Hence T is continuous, and from [**12**, Theorem 2.2.5] it follows that T is a cutter;

(3) An orthogonal projection on a nonempty, closed and convex subset C of the space. Indeed, it is well-known that any orthogonal projection is firmly nonexpansive (see, e.g., [**5**, Proposition 4.16, p. 70]), and so, as mentioned above, it is a continuous cutter. The fixed point set of T is C, as one can verify immediately. Several explicit expressions for T, in some particular cases where C has a simple form, appear in [**12**, Section 4.1];

(4) The resolvent of a maximally monotone operator, namely, $T := (Id + \gamma A)^{-1}$, where $Id : X \to X$ is the identity operator, $\gamma > 0$ and $A : X \to 2^X$ is a set-valued operator which is maximally monotone. The assertion follows from [**5**, Proposition 23.8(iii), p. 395], [**5**, Proposition 4.4(i),(v), p. 70] and an elementary calculation. It is worth noting that the fixed point set of T is the zero set of A, namely, the set $\{x \in X \mid 0 \in Ax\}$. This claim follows from an elementary calculation (see also [**5**, Proposition 23.38, p. 405]).

3. The generalized BIP method

Under the assumptions and notations of Section **2**, the generalized BIP algorithm is defined as follows:

Algorithm 3.1. The generalized BIP method for cutters

Input: A positive integer n, an arbitrary initialization point $x^0 \in X := \mathbb{R}^n$, two positive numbers τ_1 and τ_2 which satisfy $\tau_1 + \tau_2 \leq 2$, a positive integer m, an index set $I := \{1, 2, \ldots, m\}$, a family of cutters $(T_i)_{i \in I}$ defined on X, with fixed point sets $Q_i := Fix(T_i) = \{x \in X \mid T_i(x) = x\}$ and a nonempty common fixed point set $Q := \cap_{i \in I} Q_i$, a (generalized) real number $\sigma \in (0, \infty]$ with the property that $\sigma > d(x^0, Q)$, a sequence of relaxation parameters $(\lambda_k)_{k=0}^{\infty}$ which are positive numbers in the interval $[\tau_1, 2 - \tau_2]$, a sequence $(w_k)_{k=0}^{\infty}$ of weight functions with respect to I.

Iterative step: Given $k \in \mathbb{N} \cup \{0\}$ and the current iterate x^k, calculate the next iterate x^{k+1} by the iterative process

$$x^{k+1} := x^k + \lambda_k (T_{w_k}(x^k) - x^k) + e^k, \tag{3.1}$$

where the error term $e^k \in X$ has the form

$$e^k := \sum_{i \in I} w_k(i) e^{k,i}, \tag{3.2}$$

and, for all $i \in I$, the perturbation $e^{k,i}$ is any vector in X which satisfies

$$\|e^{k,i}\| \leq \frac{1}{2} \cdot \frac{\lambda_k(2-\lambda_k)\|T_i(x^k) - x^k\|^2}{\sqrt{\zeta_{k,i}} + \lambda_k \|T_i(x^k) - x^k\| + 2\sigma}, \tag{3.3}$$

where

$$\zeta_{k,i} := (\lambda_k \|T_i(x^k) - x^k\| + 2\sigma)^2 + \lambda_k(2-\lambda_k)\|T_i(x^k) - x^k\|^2. \tag{3.4}$$

Example 3.2. Algorithm **3.1** becomes fully sequential if for every $k \in \mathbb{N}$, one has $w_k(i) = 0$ for all $i \in I$ with the exception of one index $i_0(k)$ for which $w_k(i_0(k)) = 1$. In this case the index i_0 can be regarded as a control function that maps $\mathbb{N} \cup \{0\}$ to I by assigning to the given index $k \in \mathbb{N} \cup \{0\}$ the unique index $i_0(k) \in I$. If i_0 has the property that for all $j \in I$ there are infinitely many $k \in \mathbb{N}$ such that $i_0(k) = j$, namely $w_k(j) = 1$, then i_0 is the so-called *repetitive control*. Well-known particular cases of repetitive controls are cyclic and almost cyclic controls which are, in turn, also special cases of the class of *expanding controls* presented in [**16**].

In particular, Algorithm **3.1** generalizes the well-known method of successive orthogonal projections. Moreover, in the case of repetitive controls any $j \in I$ satisfies $\sum_{k=1}^{\infty} w_k(j) = \infty$, and, hence, Theorem **5.1** below ensures that the iterative sequence $(x^k)_{k \in \mathbb{N}}$ converges to a point in the intersection of the fixed point sets of the given family of operators $(T_i)_{i \in I}$.

Example 3.3. Algorithm **3.1** becomes fully simultaneous when $w_k(i) > 0$ for all $i \in I$ and $k \in \mathbb{N} \cup \{0\}$, since in this case at each iteration all the cutters $(T_i)_{i \in I}$ are considered. If, in addition, all the cutters are orthogonal projections onto given hyperplanes, then Algorithm **3.1** becomes a Cimmino-type algorithm for solving the linear system induced by these hyperplanes.

Example 3.4. Algorithm **3.1** becomes block-iterative in the classical sense if the following scenario occurs: first, one partitions the given index set $I := \{1, 2, \ldots, m\}$ into $\widetilde{m} \leq m$ "blocks", namely, into $\widetilde{m} \in \mathbb{N}$ disjoint and nonempty index subsets $I_1, I_2, \ldots, I_{\widetilde{m}}$ whose union is I; then one defines a control function over the block indices, namely a function $\widetilde{c} : \mathbb{N} \cup \{0\} \to \{1, 2, \ldots, \widetilde{m}\}$; then, for each $k \in \mathbb{N} \cup \{0\}$, one defines a weight function $w_k : I \to [0, 1]$ by $w_k(i) := 0$ if $i \notin I_{\widetilde{c}(k)}$, and $w_k(i)$ an arbitrary number in $[0, 1]$ if $i \in I_{\widetilde{c}(k)}$, with the additional condition that $\sum_{i \in I_{\widetilde{c}(k)}} w_k(i) = 1$.

For instance, suppose that $\widetilde{c}(k) = (k \mod \widetilde{m}) + 1$ for all $k \in \mathbb{N} \cup \{0\}$; suppose further that for each $\widetilde{j} \in \{1, 2, \ldots, \widetilde{m}\}$ there are $\alpha_{\widetilde{j}}$ elements in block number $\widetilde{j}$; if one defines $w_k(i) := 0$ when $i \notin I_{\widetilde{c}(k)}$ and $w_k(i) := 1/\alpha_{\widetilde{c}(k)}$ when $i \in I_{\widetilde{c}(k)}$, then this is a control which cycles periodically between the blocks; now, in order to construct x^{k+1}, one first observes that $\widetilde{c}(k) = \widetilde{j}$ for some $\widetilde{j} \in \{1, 2, \ldots, \widetilde{m}\}$, then one considers all the cutters in block number $\widetilde{j}$ and gives them an equal weight $1/\alpha_{\widetilde{j}}$, then one constructs the weighted sum T_{w_k} of the cutters in that block, and from T_{w_k} and **(3.1)** one obtains x^{k+1}.

Example 3.5. Algorithm 3.1 becomes block-iterative in the generalized sense if the following scenario occurs: one defines a block selection function $J : \mathbb{N} \cup \{0\} \to 2^I \setminus \{\emptyset\}$ which, at iteration number k, selects a block J_k, namely a nonempty subset of I; then, for each $k \in \mathbb{N} \cup \{0\}$, one defines a weight function $w_k : I \to [0,1]$ by $w_k(i) := 0$ if $i \notin J_k$, and $w_k(i)$ is an arbitrary number in $[0,1]$ if $i \in J_k$, with the additional condition that $\sum_{i\in J_k} w_k(i) = 1$. Under these assumptions, **(3.1)** becomes

$$x^{k+1} := x^k + \lambda_k \sum_{i\in J_k} w_k(T_i(x^k) - x^k) + \sum_{i\in J_k} w_k e^{k,i}. \tag{3.5}$$

Remark 3.6. The condition described in **(3.3)** is an adaptive one. It seems to be new, although it is inspired from other forms of adaptive error terms which appear in [**19**, Section 5], [**38**, Subsection 2.3]. It is unclear whether the sequence $(e^k)_{k=0}^{\infty}$ is summable, and hence convergence results which discuss **(3.1)** with summable errors cannot be used.

Remark 3.7. The (generalized) real number σ given in the input of Algorithm **3.1** poses a certain limitation on the error terms $e^{k,i}$ and e^k, for all $i \in I$ and $k \in \mathbb{N}\cup\{0\}$. Indeed, one has to be able to derive an estimate on how far is the solution set Q located from the initial iteration vector x^0 in order to have in hand an explicit σ. Such an explicit estimate can be derived sometimes.

For example, if one is able to show that Q is bounded, i.e., that it is strictly contained inside some ball $B[c^0, r]$, then the triangle inequality implies that any $\sigma \in (r + \|x^0 - c^0\|, \infty)$ satisfies $d(x^0, Q) < \sigma$. Such a case obviously occurs when, for instance, Q_i is bounded for some $i \in I$, or $Q_i \cap Q_j$ is bounded for some $i, j \in I$. Real-world scenarios in which Q is bounded occur, for example, in sensor network source localization problems in acoustics [**31**] and in wireless (electromagnetic) communication [**29**], since in both cases actually all the sets Q_i are bounded (they are discs).

As another example, consider the case of a consistent linear equation $Ax = y$, where $s \in \mathbb{N}$, $A \in \mathbb{R}^{s\times n}$ and $y \in \mathbb{R}^s$ are given and the desired solution $x \in \mathbb{R}^n$ should satisfy the additional constraint $\|x\|_1 \leq \varepsilon$ for some given $\varepsilon > 0$, where $\|x\|_1 := \sum_{i=1}^n |x_i|$ is the ℓ_1-norm of $x = (x_i)_{i=1}^n$. Such a problem has applications in signal processing [**11**]. Since $\|x\| \leq \|x\|_1$ always holds (where $\|\cdot\|$ is the Euclidean norm), one has $\|x\| \leq \varepsilon$; hence, from the triangle inequality, $d(x^0, Q) \leq \|x^0 - x\| \leq \|x^0\| + \|x\| \leq \|x^0\| + \varepsilon$; thus, any $\sigma > \|x^0\| + \varepsilon$ is good for the purpose of Algorithm **3.1**.

Anyway, if one is unable to estimate $d(x^0, Q)$ from above, then one may be forced to assume that $\sigma = \infty$, which means that all the error terms vanish.

Remark 3.8. Algorithm **3.1**, as described above, continues forever. Of course, in practice one needs some terminating condition in order to obtain an output. One such a criterion can be to stop the iterative process at some large iteration, say $k = 10^6$, and to take the corresponding point x^k as the output. Another criterion can be to check, in each iteration, the distance $d(x^k, Q_i)$ for each $i \in I$, assuming these distances can be evaluated, and to stop the process when $\max\{d(x^k, Q_i) : i \in I\} \leq \widehat{\epsilon}$ for some predetermined $\widehat{\epsilon} \geq 0$ (the case $\widehat{\epsilon} := 0$ is of interest only when one can prove convergence to $\cap_{i\in I} Q_i$ after finitely many iterations). A third criterion, at least in the case where all the sets Q_i are zero-level-sets of some functions f_i, is to evaluate, in each iteration, $f_i(x^k)$ for all $i \in I$, and to stop the process when $\max\{f_i(x^k) : i \in I\} \leq \bar{\epsilon}$ for some predetermined $\bar{\epsilon} \geq 0$. Other terminating conditions can be given.

4. Variants of BIP

Over the years other variations of the BIP method have appeared. We discuss the ones which we are aware of in this section, where we also make a few comparisons between them and our method and convergence result. We focus on variants in which the considered operators are cutters in general and not just particular cases of them such as orthogonal projections or firmly nonexpansive operators. For the sake of completeness, we also mention briefly, in the last item of the list, variants of BIP of this latter type, as well as corresponding convergence results.

In what follows $\widehat{I} := \{i \in I \mid \sum_{k=0}^{\infty} w_k(i) = \infty\}$ and $\widehat{Q} := \cap_{i\in\widehat{I}} Q_i$, with the convention that $\widehat{Q} := X$ if $\widehat{I} = \emptyset$. Here is our list.

(1) The original BIP method appears in [**1**, Algorithm 1]. There the space is finite-dimensional, the finitely many cutters are orthogonal projections onto given nonempty, closed and convex subsets and no perturbations are allowed. Our proof is inspired by [**1**, Theorem 1], but because of the different settings, there are several significant differences between our proof and the proof which appears in [**1**]; for instance, we need Lemma **5.2** and also need to perform a careful analysis in Lemma **5.5** as a result of the appearance of perturbations.

(2) In [**40**, Chapter 2], and in the unpublished technical report [**22**] (albeit some modifications are needed there), appears a version of [**1**, Algorithm 1] in which the index set I is finite, the space X is a finite-dimensional Euclidean space, the operators $(T_i)_{i\in I}$ are continuous cutters, perturbations are not allowed (namely, they vanish), and the relaxation parameters λ_k satisfy the condition $\lambda_k \in [\tau_1, (2 - \tau_2)L(x^k, w_k)]$ for all $k \in \mathbb{N} \cup \{0\}$, where, for all $k \in \mathbb{N} \cup \{0\}$,

$$L(x^k, w_k) := \begin{cases} 1, & \text{if } x^k = T_{w_k}(x^k), \\ \displaystyle\sum_{i\in I} \frac{w_k(i)\|T_i(x^k) - x^k\|^2}{\|T_{w_k}(x^k) - x^k\|^2}, & \text{otherwise.} \end{cases} \tag{4.1}$$

From the convexity of the square norm, it follows that $L(x^k, w_k) \geq 1$ for all $k \in \mathbb{N} \cup \{0\}$. It is also assumed that $Q \neq \emptyset$ and $\widehat{I} = I$ (hence $Q = \widehat{Q}$).

The first convergence theorem is [**40**, Theorem 2.4.11] (essentially [**22**, Theorem 20]), which says that if the interior of Q is nonempty, then the algorithmic sequence converges to a point in Q. The second convergence theorem is [**40**, Theorem 2.4.12] (essentially [**22**, Theorem 21]), which says that if merely $\lambda_k \in [\tau_1, 2-\tau_2]$ for all $k \in \mathbb{N} \cup \{0\}$, then the algorithmic sequence converges to a point in Q. The third convergence theorem is [**40**, Theorem 2.5.3] (essentially [**22**, Theorem 24]), which says that if Q_i is strictly convex for all $i \in I$, then the algorithmic sequence converges to a point in Q. The fourth convergence theorem is [**40**, Theorem 2.5.4] (essentially [**22**, Theorem 25]), in which it is assumed that the sequence of weight functions $(w_k)_{k=0}^{\infty}$ is fair (see Remark **5.8** below), and there is some fixed positive number ξ such that for all $i \in I$ and all $k \in \mathbb{N} \cup \{0\}$, if $w_k(i) > 0$, then actually $w_k(i) > \xi$; under these assumptions the theorem says that the algorithmic sequence converges to a point in Q.

The technique used in [**40**, Chapter 2] and [**22**] for establishing the convergence results has several similarities to the technique used here, but there are also some differences, partly because the settings are not identical. Examples of differences are the use of the Pierra's product-space formulation in [**22**, **40**] and not here, the use of Lemma **5.2** here (which is a new lemma not used elsewhere), the need to handle extrapolations in [**22**, **40**] and perturbations here, etc.

(3) Algorithm 6.1 in [**26**] is a general variant of [**1**, Algorithm 1], in which the setting is a real Hilbert space, certain kind of perturbations are allowed (essentially summable), cutters are used instead of just orthogonal projections, and one allows an infinite index set I where in each iteration k the sum is over a nonempty and finite subset I_k of I (namely, this algorithmic scheme is block-iterative in the sense of Example **3.5**, but with the modification that the range of the selection function is not $2^I \setminus \{\emptyset\}$ but rather the set of all nonempty and finite subsets of I). In [**26**, Theorem 6.6] it is proved that the sequence converges weakly to the feasible set, and under stronger assumptions strong convergence holds.

On the other hand, a stronger assumption is assumed there, namely, [**26**, Algorithm 6.1, Part 4] which says that there is a fixed positive number δ_1 such that in each iteration one of the weights, which corresponds to an index j, is at least as large as δ_1, and at this same index j another technical condition holds (that is, $\|T_{j,k}x^k - x^k\| = \max_{i \in I_k} \|T_{i,k}x^k - x^k\|$, where $T_{i,k}$ is the i-th operator at iteration k and where i is taken from the index set I_k).

Moreover, in the relevant convergence result [**26**, Theorem 6.6] one assumes that the control sequence $(I_k)_{k=0}^{\infty}$ is admissible (which is a general condition, but weaker than a repetitive control). Neither in [**1**, Algorithm 1 and Theorem 1] nor in Algorithm **3.1** and Theorem **5.1** here these assumptions are imposed. Furthermore, the convergence in [**26**, Theorem 6.6] is to the feasible set Q rather than to $\widehat{Q}$ as in [**1**, Theorem 1] and in Theorem **5.1** below (the equality $\widehat{Q} = Q$ holds under mild conditions, which in particular hold under the assumptions in [**26**, Theorem 6.6], but in general $Q \subset \widehat{Q}$).

(4) The setting in [**12**, Theorem 5.8.15] and [**13**, Theorem 9.27] (both results are essentially identical) is a possibly infinite-dimensional real Hilbert space and not necessarily continuous cutters and the sum in each iteration is over a nonempty subset J_k of the finite index set I; the cutters T_i^k in the sum are dynamic, namely they depend on both the iteration index k and the sum index i; however, these cutters should satisfy certain conditions, such as the existence of a fixed and finite family $U_i, i \in I$ of cutters with a nonempty common fixed point set Q such that $\bigcap_{i \in J_k} Q_i^k \supseteq Q$ (where Q_i^k is the fixed point set of T_i^k) and also that $U_i - Id$ is demi-closed at 0 for all $i \in I$.

Under further assumptions, such as approximate regularity of the weight functions, it is shown that the algorithmic sequence converges weakly to Q, and if the space is finite-dimensional and one assumes less (semi-regularity of the weight functions), then the algorithmic sequence converges to Q. The convergence result of Aharoni and Censor [**1**, Algorithm 1], is essentially obtained as a consequence of [**12**, Theorem 5.8.15] or [**13**, Theorem 9.27], and is illustrated, respectively, in [**12**, Example 5.8.18] and [**13**, Example 9.30], for the special case of orthogonal projections onto nonempty, closed and convex subsets of the space.

It is worthwhile to note that no perturbations are allowed in [**12**, Theorem 5.8.15], [**12**, Example 5.8.18], [**13**, Theorem 9.27] and [**13**, Example 9.30], and while it seems that the method of [**12**, Example 5.8.18] and [**13**, Theorem 9.27] can be generalized to other cutters (by modifying the arguments in [**13**, Example 26(e)]), it does not seem that it can be generalized to the perturbations that we consider in Algorithm **3.1**, since the method of [**12**, Example 5.8.18] is heavily based on a certain nonnegative (usually positive) lower bound on $\|x^{k+1}-q\|-\|x^k-q\|$, and this lower bound is eliminated when the perturbations that we consider in Algorithm **3.1** appear.

We also note that one can find in both [**12**] and [**13**] other results which are closely related to [**12**, Theorem 5.8.15] and [**13**, Theorem 9.27], such as [**12**, Theorem 5.10.2] and [**13**, Theorem 9.35] (admissible step sizes), and [**12**, Theorem 5.8.25] (for orthogonal projections), where in all of these cases no perturbations appear.

(5) The setting in [**39**, Theorem 4.1 and Theorem 4.5] is a possibly infinite-dimensional real Hilbert space, not necessarily continuous cutters, but ones which should satisfy other conditions, such as the Opial's demi-closedness principle (for the weak convergence case); the algorithmic scheme allows strings and not just convex combinations and relaxations as in Algorithm **3.1** above. On the other hand, in both [**39**, Theorem 4.1] and [**39**, Theorem 4.5] there is a certain restriction on the control, namely, Condition (ii) there which is something in the spirit of an almost cyclic control, and the whole convergence is to the common fixed point set, while in our Theorem **5.1** such a restriction does not exist and the convergence is not necessarily to a point in the common fixed point set, but rather to a point located in the possibly larger set $\widehat{Q}$.

As an illustration to this last point, consider the case of strings of length one and weights which, in each iteration, vanish with the exception of one place in which they are equal to 1. This is the case of a fully sequential algorithmic scheme, and Condition (ii) in [**39**, Theorem 4.1] is the classical almost cyclic control. On the other hand, in our Theorem **5.1** one allows the control to be repetitive, that is, more general, as explained in Example **3.2** above. No perturbations are allowed in [**39**, Theorem 4.1], while in [**39**, Theorem 4.5] summable perturbations are allowed but the operators must be firmly nonepxansive, and hence (see Section **2**) must be continuous cutters.

(6) The setting in [**36**, Theorem 3.1, Theorem 3.2] is a possibly infinite-dimensional real Hilbert space and cutters which are not necessarily continuous, and the sum in each iteration is over a nonempty subset I_k of the finite index set I (namely, this algorithmic scheme is block-iterative in the generalized sense of Example **3.5**). The cutters, however, should satisfy other conditions which we do not impose in Algorithm **3.1**, such as having a representation to their fixed point sets as the zero-level-sets of well-behaved proximity functions.

Additional conditions which are not imposed in Algorithm **3.1** but are imposed in [**36**] are that the upper bound τ_2 on the relaxation parameters should be at least 1, and that there is a positive number $\omega^- \in [0,1]$ such that $w_k(i) \geqq \omega^-$ for all $k \in \mathbb{N} \cup \{0\}$ and all $i \in I$, namely the algorithmic scheme in [**36**] is fully simultaneous with a strictly positive lower bound on the weights. In addition, one needs to impose there an assumption (Condition (ii)) which is essentially an almost cyclic control and no perturbations are allowed there. Under these and additional assumptions, the authors of [**36**] derive weak, strong and linear convergence of the iterative algorithmic scheme to a common fixed point of the given cutters.

(7) Other variants of BIP, for more restricted cutters or for other types of operators, as well as associated convergence results, appear in the following publications. [**4**, The algorithmic scheme on p. 378, Theorem 3.20, Corollary 3.22, Corollary 3.24, Corollary 3.25]: Here finite and infinite-dimensional real Hilbert spaces with firmly nonexpansive operators are considered (many other convergence results for orthogonal projections in [**4**, Sections 4–6]); [**28**, Algorithm (2.2), Theorems 1, 2]: The finite-dimensional case with projections onto separating hyperplanes; [**9**, Theorem 4.4]: This is an almost simultaneous BIP method for orthogonal projections in a finite-dimensional Euclidean space; [**10**, The method of (1), Theorem 1]: This is an almost simultaneous BIP method for orthogonal projections in an infinite-dimensional Hilbert space; [**2**, The method of (8), Theorem 4.1]: Firmly nonexpansive mappings in finite-dimensional strictly convex normed spaces; [**25**, Algorithm 6.5, Theorem 6.4]: A modified BIP method with orthogonal projections in Hilbert spaces; [**34**, Theorem 4.1 and some theorems in Section 5]: a modified fully simultaneous BIP method with generalized nonexpansive mappings in smooth and uniformly convex Banach spaces.

5. The global convergence theorem and its proof

In this section we formulate and prove our global convergence theorem concerning Algorithm **3.1**.

Theorem 5.1. *Under the notations and assumptions of Sections 2 and 3, assume that T_i is continuous for all $i \in I$. Denote $\widehat{I} := \{i \in I \mid \sum_{k=0}^{\infty} w_k(i) = \infty\}$ and $\widehat{Q} := \cap_{i \in \widehat{I}} Q_i$, with the convention that $\widehat{Q} := X$ if $\widehat{I} = \emptyset$. Then any sequence defined in (3.1) converges to a point in $\widehat{Q} \cap B[x^0, 2\sigma]$. In particular, if $\widehat{I} = I$, then the sequence defined in (3.1) converges to a point in $Q \cap B[x^0, 2\sigma]$.*

The proof of Theorem 5.1 is based on several claims which are formulated and proved below. Before proceeding with these claims, we need a further notation: given $x \in X$, $q \in X$, $\lambda \in [0,2]$, $\theta \in [0,\infty)$, and $i \in I$, we denote by $E_\theta(x,q,\lambda,i)$ the set of all $e \in X$ which satisfy

$$\|e\| \leq \frac{\theta \cdot \lambda(2-\lambda)\|T_i(x)-x\|^2}{\sqrt{\zeta} + \lambda\|T_i(x)-x\| + \|x-q\|} \tag{5.1}$$

with

$$\zeta := (\lambda\|T_i(x)-x\| + \|x-q\|)^2 + \lambda(2-\lambda)\|T_i(x)-x\|^2, \tag{5.2}$$

where both sides of (5.1) mean zero if the denominator of the fraction on the right-hand side, and hence also the numerator, vanish. In addition, given a weight function $w : I \to [0,1]$, we denote by $E_\theta(x,q,\lambda,w)$ the set of all $e \in X$ which satisfy $e = \sum_{i \in I} w(i)e^i$, where $e^i \in E_\theta(x,q,\lambda,i)$ for all $i \in I$.

We start with the following apparently new lemma, which seems to be of independent interest. Since it enables us to prove the (essentially) Fejér monotonicity of the sequence (Lemma 5.6 below), and since a certain perturbation appears in it (the term e), Lemma 5.2 can be thought of as establishing a Fejér monotonicity phenomenon in a perturbed form. As explained in Remark 5.3 below, Lemma 5.2 actually shows that the operator $Id + \lambda(T - Id)$ is not only quasi-nonexpansive, but rather that its quasi-nonexpansiveness is stable under small perturbations, and that this phenomenon holds in a general setting.

Lemma 5.2. (Fejér monotonicity in a perturbed form): *Suppose that $T : X \to X$ is a separator of a nonempty subset $S \subseteq X$. Given $x \in X$, $q \in S$ and $\lambda \in [0,2]$, if $e \in X$ satisfies (5.1) with $\theta := 1$ and T instead of T_i, and if*

$$y := x + \lambda(T(x) - x) + e, \tag{5.3}$$

then

$$\|y - q\| \leq \|x - q\|. \tag{5.4}$$

If, in addition, $x \neq T(x)$, $0 < \lambda < 2$ and $e \in X$ satisfies (5.1) with strict inequality, then

$$\|y - q\| < \|x - q\|. \tag{5.5}$$

Proof. As a result of (5.3), (2.1), the Cauchy-Schwarz inequality and the assumption that $\lambda \in [0,2]$, we have

$$\begin{aligned}
\|y-q\|^2 &= \|x-q\|^2 + \|\lambda(T(x)-x)+e\|^2 + 2\langle x-q, \lambda(T(x)-x)+e\rangle \\
&= \|x-q\|^2+\lambda^2\|T(x)-x\|^2+\|e\|^2+2\lambda\langle T(x)-x,e\rangle+2\lambda\langle x-q,T(x)-x\rangle+2\langle x-q,e\rangle \\
&= \|x-q\|^2+\lambda^2\|T(x)-x\|^2+2\lambda\langle T(x)-q,T(x)-x\rangle-2\lambda\langle T(x)-x,T(x)-x\rangle \\
&\quad + 2\lambda\langle T(x)-x,e\rangle+2\langle x-q,e\rangle+\|e\|^2 \\
&= \|e\|^2+2\lambda\langle T(x)-x,e\rangle+2\langle x-q,e\rangle+\|x-q\|^2-\lambda(2-\lambda)\|T(x)-x\|^2 \\
&\quad + 2\lambda\langle x-T(x),q-T(x)\rangle \\
&\le \|e\|^2+2(\lambda\|T(x)-x\|+\|x-q\|)\|e\|-\lambda(2-\lambda)\|T(x)-x\|^2+\|x-q\|^2 \\
&\le \|x-q\|^2. \qquad (5.6)
\end{aligned}$$

To derive the last inequality in **(5.6)** we used the following simple facts: (i) from elementary analysis and algebra, given two nonnegative numbers α_1 and α_2, the inequality $t^2+2\alpha_1 t-\alpha_2\le 0$ for nonnegative t holds whenever $t\in[0,\sqrt{\alpha_1^2+\alpha_2}-\alpha_1]$; (ii) the simple identity $\sqrt{\alpha_1^2+\alpha_2}-\alpha_1=\alpha_2/(\sqrt{\alpha_1^2+\alpha_2}+\alpha_1)$ holds; (iii) the equation **(5.1)** actually says that $t\le\alpha_2/(\sqrt{\alpha_1^2+\alpha_2}+\alpha_1)$ for $t:=\|e\|$, $\alpha_1:=\lambda\|T(x)-x\|+\|x-q\|$ and $\alpha_2:=\lambda(2-\lambda)\|T(x)-x\|^2$; (iv) the last inequality in **(5.6)** can be written as $t^2+2\alpha_1 t-\alpha_2+\|x-q\|^2\le\|x-q\|^2$.

Finally, if $x\neq T(x)$ and $0<\lambda<2$, then the number on the right-hand side of **(5.1)** is positive. Hence there are vectors $e\in X$ fulfilling the strict version of **(5.1)**, namely, these vectors are all the ones whose magnitudes are smaller than the right-hand side of **(5.1)**. Let e be such a vector. This means that in the notation of the previous paragraph, $0\le t<\alpha_2/(\sqrt{\alpha_1^2+\alpha_2}+\alpha_1)=\sqrt{\alpha_1^2+\alpha_2}-\alpha_1$ and $\alpha_2>0$. These inequalities and elementary properties of quadratic inequalities imply that $t^2+2\alpha_1 t-\alpha_2<0$. Thus, the last inequality in **(5.6)** is strict. □

Remark 5.3. Lemma **5.2** is rather general, since the space X can be an arbitrary real inner product space and not necessarily a finite-dimensional Euclidean space, and the operator T there is not necessarily a cutter and not necessarily continuous. Moreover, this lemma shows that the operator T which appears there, and also a relaxed version of it, exhibit a certain stability property.

Indeed, we recall that an operator $T: X\to X$ with a nonempty fixed point set $Fix(T)$ is called quasi-nonexpansive if $\|Tx-q\|\le\|x-q\|$ for all $x\in X$ and $q\in Fix(T)$. Now, if we take $e:=0$ and $\lambda:=1$ in Lemma **5.2** then it follows that any cutter T is quasi-nonexpansive. Furthermore, Lemma **5.2** shows that the operator $T_\lambda:=Id+\lambda(T-Id)$ is quasi-nonexpansive for every $\lambda\in[0,2]$, where Id is the identity operator, and, as a matter of fact, the property of being quasi-nonexpansive holds true even if we translate T_λ by a vector e which satisfies **(5.1)** with $\theta:=1$. In other words, the property of T_λ being quasi-nonexpansive is stable under certain small perturbations.

Lemma 5.4. *Let $z \in X$ be given, and denote $I_z := \{i \in I \mid z \notin Q_i\}$. Suppose that G is a nonempty and compact subset of X. Then there exists an $\eta \in [0, \infty)$ such that for all $x \in G$, all $\lambda \in [0, 2]$, all weight functions $w : I \to [0, 1]$ and all $e \in E_1(x, z, \lambda, w)$, one has*

$$\|x + \lambda(T_w(x) - x) + e - z\| \leq \|x - z\| + \eta \sum_{i \in I_z} w(i). \tag{5.7}$$

Proof. Define for $I_z \neq \emptyset$

$$\eta := \sup\{\|x+\lambda(T_i(x)-x)+\widetilde{e}-z\| \mid x \in G, \lambda \in [0,2], i \in I_z, \widetilde{e} \in E_1(x, z, \lambda, i)\}, \tag{5.8}$$

and for $I_z = \emptyset$ define

$$\eta := 0. \tag{5.9}$$

If **(5.9)** holds, then obviously $\eta \in [0, \infty)$. Next we show that $\eta \in [0, \infty)$ also when **(5.8)** holds, from which it will follow that **(5.7)** holds regardless if $I_z \neq \emptyset$ or not.

Since G is a compact set and since for all $i \in I$ the real function $f_i(x) := 2\|T_i(x) - x\|$ is continuous on G as a result of the continuity of the norm and the assumption on T_i, it follows from the the well-known Weierstrass Theorem (that is, the Extreme Value Theorem of calculus) that f_i is bounded from above on G. Let $\mu_i > 0$ be any such upper bound. Elementary algebra shows that any $\widetilde{e}$ which satisfies **(5.1)** (with $\theta := 1$ and $\widetilde{e}$ instead of e) also satisfies $\|\widetilde{e}\| \leq f_i(x)$, and hence $\|\widetilde{e}\| \leq \mu_i$ whenever $(x, \lambda) \in G \times [0, 2]$ and $\widetilde{e} \in E_1(x, z, \lambda, i)$.

Since

$$g_i(x, \lambda, \widetilde{e}) := \|x + \lambda(T_i(x) - x) + \widetilde{e} - z\| \tag{5.10}$$

is continuous on $G \times [0, 2] \times B[0, \mu_i]$ for all $i \in I$ as a result of the continuity of the norm and the assumption on T_i, the Weierstrass Theorem ensures that g_i is bounded from above on $G \times [0, 2] \times B[0, \mu_i]$. Since I and hence I_z are finite, we conclude from the previous assertions and the definition of η that

$$0 \leq \eta \leq \max_{i \in I_z} \sup\{g_i(x, \lambda, \widetilde{e}) \mid (x, \lambda, \widetilde{e}) \in G \times [0, 2] \times B[0, \mu_i]\} < \infty, \tag{5.11}$$

and so $\eta \in [0, \infty)$.

Now let $x \in G$, $\lambda \in [0, 2]$ be arbitrary, let $w : I \to [0, 1]$ be an arbitrary weight function and let $e \in E_1(x, z, \lambda, w)$ be arbitrary. Since for all $i \notin I_z$ one has $z \in Q_i$, it follows from Lemma **5.2** (with $T := T_i$, $S := Q_i$, $q := z$ and e^i instead of e) that

$$\|x + \lambda(T_i(x) - x) + e^i - z\| \leq \|x - z\|. \tag{5.12}$$

These inequalities and the triangle inequality, together with the definition of η, as well as the convention that a sum over the empty set is zero, ensure that

$$
\begin{aligned}
&\|x+\lambda(T_w(x)-x)+e-z\| \\
&=\left\|\sum_{i\in I_z} w(i)\big(x+\lambda(T_i(x)-x)+e^i-z\big)+\sum_{i\notin I_z} w(i)\big(x+\lambda(T_i(x)-x)+e^i-z\big)\right\| \\
&\le \sum_{i\in I_z} w(i)\left\|x+\lambda(T_i(x)-x)+e^i-z\right\|+\sum_{i\notin I_z} w(i)\left\|x+\lambda(T_i(x)-x)+e^i-z\right\| \\
&\qquad\le \left(\sum_{i\in I_z} w(i)\right)\eta+\left(1-\sum_{i\in I_z} w(i)\right)\|x-z\| \\
&\qquad= \|x-z\|+\left(\sum_{i\in I_z} w(i)\right)(\eta-\|x-z\|)\le \|x-z\|+\eta\sum_{i\in I_z} w(i). \qquad (5.13)
\end{aligned}
$$

□

Lemma 5.5. *Let $q\in Q$ be fixed and suppose that $C\subseteq X$ is a nonempty and compact subset. Denote $I_C:=\{i\in I\mid C\cap Q_i=\emptyset\}$. Then there exists $\beta>0$ such that for all $x\in C$, all $\lambda\in[\tau_1,2-\tau_2]$, all weight functions $w:I\to[0,1]$ and all $e\in E_{0.5}(x,q,\lambda,w)$,*

$$\|x+\lambda(T_w(x)-x)+e-q\|\le\|x-q\|-\beta\sum_{i\in I_C} w(i). \qquad (5.14)$$

Proof. If $I_C=\emptyset$ then **(5.14)** holds with $\beta=1$ which is positive. Hence, from now on we assume that $I_C\neq\emptyset$. For each $i\in I_C$ denote

$$
\begin{aligned}
\beta_i:=\inf\{&\|x-q\|-\|x+\lambda(T_i(x)-x)+\widetilde{e}-q\|\mid \\
&x\in C,\lambda\in[\tau_1,2-\tau_2],\widetilde{e}\in E_{0.5}(x,q,\lambda,i)\}, \qquad (5.15)
\end{aligned}
$$

and let

$$\beta:=\min\{\beta_i\mid i\in I_C\}. \qquad (5.16)$$

We show next that $\beta>0$, and that **(5.14)** holds regardless whether $I_C=\emptyset$ or not.

The definition of β_i and Lemma **5.2** imply that $\beta_i\in[0,\infty)$ for every $i\in I_C$. Since I and hence I_C are finite, it suffices to show that $\beta_i>0$ for all $i\in I_C$ in order to conclude that $\beta>0$. Given $i\in I_C$, the definition of β_i implies that for each $\ell\in\mathbb{N}$ there exists a triplet $(x_{\ell,i},\lambda_{\ell,i},\widetilde{e}_{\ell,i})\in C\times[\tau_1,2-\tau_2]\times E_{0.5}(x_{\ell,i},q,\lambda_{\ell,i},i)$ such that

$$\beta_i\le\|x_{\ell,i}-q\|-\|x_{\ell,i}+\lambda_{\ell,i}(T_i(x_{\ell,i})-x_{\ell,i})+\widetilde{e}_{\ell,i}-q\|<\beta_i+\frac{1}{\ell}. \qquad (5.17)$$

Because of the compactness of $C\times[\tau_1,2-\tau_2]$, there exists an infinite set N_1 of natural numbers, and a pair $(x(i),\lambda(i))\in C\times[\tau_1,2-\tau_2]$, such that

$$(x(i),\lambda(i))=\lim_{\ell\to\infty,\ell\in N_1}(x_{\ell,i},\lambda_{\ell,i}). \qquad (5.18)$$

Let

$$\Lambda_i := \inf\{\|\widetilde{e}_{\ell,i}\| \mid \ell \in N_1\} \tag{5.19}$$

and let N_2 be an infinite subset of N_1 such that the subsequence $(\|\widetilde{e}_{\ell,i}\|)_{\ell\in N_2}$ converges to Λ_i.

We claim that the subsequence $(\widetilde{e}_{\ell,i})_{\ell\in N_2}$ has a convergent subsequence. Indeed, since C is a compact set and since for all $i \in I$ the real function $f_i(x) := \|T_i(x)-x\|$ is continuous on C as a result of the continuity of the norm and the assumption on T_i, it follows from the Weierstrass Theorem that f_i is bounded from above on C. Let $\mu_i > 0$ be an arbitrary upper bound on f_i over C.

Elementary algebra shows that any $\widetilde{e}$ which satisfies **(5.1)** (with $\theta := 0.5$ and with $\widetilde{e}$ instead of e) also satisfies $\|\widetilde{e}\| \leq f_i(x)$, and hence $\|\widetilde{e}\| \leq \mu_i$ whenever $(x, \lambda, \widetilde{e}) \in C \times [\tau_1, 2-\tau_2] \times E_{0.5}(x, q, \lambda, i)$.

Since $\widetilde{e}_{\ell,i} \in E_{0.5}(x_{\ell,i}, q, \lambda_{\ell,i}, i)$ for all $\ell \in \mathbb{N}$, it follows that $\|\widetilde{e}_{\ell,i}\| \leq \mu_i$ for all $\ell \in \mathbb{N}$ and, in particular, for all $\ell \in N_2$. Because the ball $B[0, \mu_i]$ is compact and the sequence $(\widetilde{e}_{\ell,i})_{\ell\in N_2}$ is contained in this ball, it indeed has a convergent subsequence which converges to some vector $\widetilde{e}(i)$ which belongs to this ball, namely, there exists an infinite subset N_3 of N_2 such that $\lim_{\ell\to\infty,\ell\in N_3} \widetilde{e}_{\ell,i} = \widetilde{e}(i)$.

Now, when we combine this fact, together with the continuity of the norm, the fact that $N_3 \subseteq N_2$ and the definition of Λ_i, we obtain

$$\Lambda_i = \lim_{\ell\to\infty,\ell\in N_3} \|\widetilde{e}_{\ell,i}\| = \|\widetilde{e}(i)\|. \tag{5.20}$$

In addition, denote by $h_i(x, \lambda)$ the function on the right-hand side of **(5.1)**, with $\theta := 1/2$ and h_i defined on $C \times [\tau_1, 2-\tau_2]$. Then $\|\widetilde{e}_{\ell,i}\| \leq h_i(x_{\ell,i}, \lambda_{\ell,i})$ for all $\ell \in \mathbb{N}$, and, in particular, for all $\ell \in N_3$, because $\widetilde{e}_{\ell,i} \in E_{0.5}(x_{\ell,i}, q, \lambda_{\ell,i}, i)$ for all $\ell \in \mathbb{N}$.

We recall that the right-hand side of **(5.1)** vanishes if its denominator, and hence its numerator, vanish. Thus, it is not clear that h_i is continuous at (x, λ) for which the denominator in the definition of $h_i(x, \lambda)$ vanishes, but from the continuity of the norm and of T_i it is clear that h_i is continuous at (x, λ) whenever the above-mentioned denominator does not vanish. Anyway, since the definition of I_C ensures that $C \cap Q_i = \emptyset$ for all $i \in I_C$, and therefore $x(i) \notin Q_i$, namely, $x(i) \neq T_i(x(i))$ for all $i \in I_C$, it follows that for all $i \in I_C$ the denominator in the definition of $h_i(x, \lambda)$ does not vanish at $(x(i), \lambda(i))$; hence h_i is continuous at $(x(i), \lambda(i))$ for all $i \in I_C$. This fact and the limit $(x(i), \lambda(i)) = \lim_{\ell\to\infty,n\in N_3}(x_{\ell,i}, \lambda_{\ell,i})$, yield

$$\|\widetilde{e}(i)\| = \lim_{\ell\to\infty,\ell\in N_3} \|\widetilde{e}_{\ell,i}\| \leq \lim_{\ell\to\infty,\ell\in N_3} h_i(x_{\ell,i}, \lambda_{\ell,i}) = h_i(x(i), \lambda(i)). \tag{5.21}$$

Hence, $\widetilde{e}(i) \in E_{0.5}(x(i), q, \lambda(i), i)$. Thus, if we let $\theta := 1$ in **(5.1)**, we see that $\widetilde{e}$ satisfies **(5.1)** with strict inequality, where in **(5.1)** we let $T := T_i$, $\lambda := \lambda(i)$, $x := x(i)$ and $\widetilde{e}(i)$ instead of e. Since $0 < \tau_1 \leq \lambda(i) \leq 2-\tau_2 < 2$ and $x(i) \neq T_i(x(i))$, and since $q \in Q \subseteq Q_i$, we conclude from Lemma **5.2** (in which $T := T_i$, $S := Q_i$, $\widetilde{e}(i)$ is instead of e, $\lambda := \lambda(i)$, $x := x(i)$), from **(5.17)**, and from the continuity of

T_i and the norm, that for each $i \in I_C$,

$$\beta_i = \lim_{\ell\to\infty, \ell \in N_3} \Big[\|x_{\ell,i} - q\| - \|x_{\ell,i} + \lambda_{\ell,i}(T_i(x_{\ell,i}) - x_{\ell,i}) + \widetilde{e}_{\ell,i} - q\|\Big]$$
$$= \|x(i) - q\| - \|x(i) + \lambda(i)(T_i(x(i)) - x(i)) + \widetilde{e}(i) - q\| > 0. \quad (5.22)$$

Finally, since $\beta = \min\{\beta_i, i \in I_C\}$ and since I, and hence I_C, are finite, it follows that $\beta = \beta_j$ for some $j \in I_C$, and so indeed $\beta > 0$ also in the case where $I_C \neq \emptyset$, as claimed.

The definition of β implies that

$$\|x + \lambda(T_i(x) - x) + \widetilde{e} - q\| \leq \|x - q\| - \beta \quad (5.23)$$

for all $i \in I_C$ and all $(x, \lambda, \widetilde{e}) \in C \times [\tau_1, 2 - \tau_2] \times E_{0.5}(x, q, \lambda, i)$, with an empty inequality when $I_C = \emptyset$. In addition, since $q \in Q \subseteq Q_i$ for each $i \in I$, we infer from Lemma **5.2** that

$$\|x + \lambda(T_i(x) - x) + \widetilde{e} - q\| \leq \|x - q\| \quad (5.24)$$

for every $i \in I$ and all $(x, \lambda, \widetilde{e}) \in C \times [\tau_1, 2-\tau_2] \times E_{0.5}(x, q, \lambda, i)$, and, in particular, for every $i \notin I_C$ (with an empty inequality if $I \backslash I_C = \emptyset$).

It follows from these inequalities and the triangle inequality, together with the convention that the sum over the empty set is zero, that for all $x \in C$, all $\lambda \in [\tau_1, 2 - \tau_2]$, all weight functions $w : I \to [0, 1]$ and all $e \in E_{0.5}(x, q, \lambda, w)$,

$$\|x + \lambda(T_w(x) - x) + e - q\| = \left\| \sum_{i \in I} w(i)\big(x + \lambda(T_i(x) - x) + e^i - q\big) \right\|$$
$$= \left\| \sum_{i \in I_C} w(i) \left(x + \lambda(T_i(x) - x) + e^i - q\right) + \sum_{i \notin I_C} w(i) \left(x + \lambda(T_i(x) - x) + e^i - q\right) \right\|$$
$$\leq \left(\sum_{i \in I_C} w(i)\right) (\|x - q\| - \beta) + \left(1 - \sum_{i \in I_C} w(i)\right) \|x - q\|$$
$$= \|x - q\| - \beta \sum_{i \in I_C} w(i). \quad (5.25)$$

Consequently, we established **(5.14)**, as required. □

The next lemma establishes the Fejér monotonicity of sequences generated by Algorithm **3.1**.

Lemma 5.6. *For each $q \in Q \cap B[x^0, 2\sigma]$, $k \in \mathbb{N} \cup \{0\}$ and each e^k which satisfies **(3.3)**, one has*

$$\|x^{k+1} - q\| \leq \|x^k - q\|, \quad (5.26)$$

and this inequality is strict if there exists some $i \in I$ such that both $x^k \notin Q_i$ and $w_k(i) > 0$.

Proof. We prove the assertion using induction on k. First observe that by the choice of σ in Algorithm **3.1**, one has $d(x^0, Q) < \sigma$. Therefore, there is some $\widetilde{q} \in Q$ such that $\|x^0 - \widetilde{q}\| < \sigma$. Hence, $Q \cap B[x^0, \sigma] \neq \emptyset$ and, since $B[x^0, \sigma] \subseteq B[x^0, 2\sigma]$, also $Q \cap B[x^0, 2\sigma] \neq \emptyset$.

Now let $q \in Q \cap B[x^0, 2\sigma]$ be arbitrary. Suppose that $k = 0$. Since $\|x^0 - q\| \leq 2\sigma$ and $e^{0,i}$ satisfies **(3.3)** for each $i \in I$, it follows that $e^{0,i}$ satisfies **(5.1)** (with $e^{0,i}$ instead of e and with $\theta := 1/2$, and hence also with $\theta := 1$) for each $i \in I$.

This fact, the assumption that $q \in Q \subseteq Q_i$ for each $i \in I$, and the notation $y^{0,i} := x^0 + \lambda_0(T_i(x^0) - x^0) + e^{0,i}$ imply, using Lemma **5.2** (in which $x := x^0$, $S := Q_i$, $T := T_i$, $\lambda := \lambda_0$, $e := e^{0,i}$), that

$$\|y^{0,i} - q\| \leq \|x^0 - q\|. \tag{5.27}$$

Since $0 < \tau_1 \leq \lambda_0$ and $2 - \lambda_0 \geq \tau_2 > 0$, if $x^0 \notin Q_i$ then **(5.27)** is strict, again from Lemma **5.2**. These considerations, **(3.1)** and the triangle inequality imply that

$$\begin{aligned} \|x^1 - q\| &= \|x^0 + \lambda_0(T_{w_0}(x^0) - x^0) + e^0 - q\| \\ &= \left\| \sum_{i \in I} w_0(i) \left(x^0 + \lambda_0(T_i(x^0) - x^0) + e^{0,i} - q \right) \right\| = \left\| \sum_{i \in I} w_0(i) \left(y^{0,i} - q \right) \right\| \\ &\leq \sum_{i \in I} w_0(i) \left\| y^{0,i} - q \right\| \leq \sum_{i \in I} w_0(i) \|x^0 - q\| = \|x^0 - q\|, \end{aligned} \tag{5.28}$$

and this inequality is strict if there exists some $i \in I$ such that $x^0 \notin Q_i$ and $w_0(i) > 0$. In other words, **(5.26)** holds true for the case $k = 0$.

Suppose now that the assertion holds for all nonnegative integers up to $k \in \mathbb{N} \cup \{0\}$. We want to show that it holds for $k+1$ as well. The induction hypothesis implies that $\|x^k - q\| \leq \ldots \leq \|x^1 - q\| \leq \|x^0 - q\| \leq 2\sigma$. Since $e^{k,i}$ satisfies **(3.3)** for each $i \in I$, it follows that $e^{k,i}$ satisfies **(5.1)** (with $e^{k,i}$ instead of e and with $\theta := 1/2$, and hence also with $\theta := 1$) for each $i \in I$. This fact, the assumption that $q \in Q \subseteq Q_i$ for each $i \in I$, and the notation $y^{k,i} := x^k + \lambda_k(T_i(x^k) - x^k) + e^{k,i}$ imply, using Lemma **5.2** (in which $x := x^k$, $S := Q_i$, $T := T_i$, $\lambda := \lambda_k$, $e := e^{k,i}$), that

$$\|y^{k,i} - q\| \leq \|x^k - q\|. \tag{5.29}$$

Since $0 < \tau_1 \leq \lambda_k$ and $2 - \lambda_k \geq \tau_2 > 0$, if $x^k \notin Q_i$ then **(5.29)** is strict, again from Lemma **5.2**. These considerations, **(3.1)** and the triangle inequality imply that

$$\begin{aligned} \|x^{k+1} - q\| &= \|x^k + \lambda_k(T_{w_k}(x^k) - x^k) + e^k - q\| \\ &= \left\| \sum_{i \in I} w_k(i) \left(x^k + \lambda_k(T_i(x^k) - x^k) + e^{k,i} - q \right) \right\| = \left\| \sum_{i \in I} w_k(i) \left(y^{k,i} - q \right) \right\| \\ &\leq \sum_{i \in I} w_k(i) \left\| y^{k,i} - q \right\| \leq \sum_{i \in I} w_k(i) \|x^k - q\| = \|x^k - q\|, \end{aligned} \tag{5.30}$$

and this inequality is strict if there exists some $i \in I$ such that both $x^k \notin Q_i$ and $w_k(i) > 0$. In other words, **(5.26)** holds true also for $k+1$, as required. □

Proof of Theorem 5.1 . The proof is divided into several steps. In the following steps we fix some $q \in Q \cap B[x^0, \sigma]$. This is possible since $Q \cap B[x^0, \sigma] \neq \emptyset$ by the assumption that $d(x^0, Q) < \sigma$.

Step 1: $(x^k)_{k=0}^\infty$ **has a convergent subsequence**: Indeed, Lemma **5.6** ensures that $\|x^{k+1} - q\| \leq \|x^k - q\| \leq \ldots \leq \|x^0 - q\|$ for all $k \in \mathbb{N} \cup \{0\}$. Hence, x^k is in the compact ball $B[q, \|x^0 - q\|]$ for all $k \in \mathbb{N} \cup \{0\}$, so that $(x^k)_{k=0}^\infty$ must have a convergent subsequence.

Step 2: $(x^k)_{k=0}^\infty$ **has at most one accumulation point**: Step 1 ensures that $(x^k)_{k=0}^\infty$ has at least one accumulation point. Assume to the contrary that it has two different accumulation points u and u'. Then $\delta := \|u - u'\| > 0$. As explained in Step 1 above, the sequence $(\|x^k - q\|)_{k=0}^\infty$ is decreasing, and it is bounded from below by 0. Thus, $\lim_{k\to\infty} \|x^k - q\|$ exists. Since u is an accumulation point of $(x^k)_{k=0}^\infty$, it follows from the norm continuity that $\lim_{k\to\infty} \|x^k - q\| = \|u - q\|$ and also that, for all $k \in \mathbb{N} \cup \{0\}$,

$$\|u - q\| \leq \|x^k - q\|. \tag{5.31}$$

Step 2.1: $u \in Q \cap B[x^0, 2\sigma]$: We first show that $u \in B[x^0, 2\sigma]$. Indeed, **(5.31)** implies that $\|u - q\| \leq \|x^0 - q\|$, and since $\|x^0 - q\| \leq \sigma$ by the choice of q, we have

$$\|u - x^0\| \leq \|u - q\| + \|q - x^0\| \leq 2\|x^0 - q\| \leq 2\sigma. \tag{5.32}$$

Hence, $u \in B[x^0, 2\sigma]$.

Now we show that $u \in Q$. Assume to the contrary that $u \notin Q$. Then $I_u := \{i \in I \mid u \notin Q_i\}$ is nonempty and is finite since I is finite. Since Q_i is closed for each $i \in I$, it follows that $d(u, Q_i) > 0$ for every $i \in I_u$. Let ρ be any positive number which satisfies $\rho < \min\{\delta/2, d(u, Q_i) \mid i \in I_u\}$. Let C be the closed ball with radius ρ and center u. The choice of ρ implies that $C \cap Q_i = \emptyset$ for all $i \in I_u$. Let $I_C := \{i \in I \mid C \cap Q_i = \emptyset\}$.

From the previous line $I_u \subseteq I_C$. On the other hand, it must be that $I_C \subseteq I_u$ since if $i \in I_C \backslash I_u$, then both $u \in Q_i$ (since $i \notin I_u$) and $u \notin Q_i$ (since $u \in C$ and C is disjoint to Q_i according to the definition of I_C), a contradiction. Hence, $I_C = I_u$.

Now let η be as in Lemma **5.4**, where there $z := u$, and let G be the closed ball of radius $\|x^0 - q\|$ around q (or the closed ball with radius $\|x^0 - q\| + \|q\|$ around the origin). Let β be as in Lemma **5.5**. Define

$$\varepsilon := \frac{\rho\beta}{\eta + \beta}. \tag{5.33}$$

Since $\varepsilon > 0$ and u is an accumulation point of $(x^k)_{k=0}^\infty$, there exists an index $k \in \mathbb{N} \cup \{0\}$ sufficiently large such that

$$\|x^k - u\| < \varepsilon. \tag{5.34}$$

Since $\varepsilon < \rho$, it follows that $x^k \in C$. Since u' is also an accumulation point and $\rho < (1/2)\delta = (1/2)\|u - u'\|$, the set of all $\widetilde{k} \in \mathbb{N} \cup \{0\}$ such that $\widetilde{k} > k$ and $x^{\widetilde{k}} \notin C$ is nonempty (in fact, infinite). Let k' be the smallest element in this set. Then $k' > k$, and any $\widetilde{k} \in [k, k'-1] \cap (\mathbb{N} \cup \{0\})$ has the property that $x^{\widetilde{k}} \in C$. Hence, we can apply Lemma **5.5** repeatedly with $x := x^{\widetilde{k}}$ where $\widetilde{k} \in [k, k'-1] \cap (\mathbb{N} \cup \{0\})$, and by using **(3.1)**, **(5.31)**, **(5.34)**, the triangle inequality and the choice of k, it follows that

$$\begin{aligned}\|u - q\| \leq \|x^{k'} - q\| &= \|x^{k'-1} + \lambda_{k'-1}(T_{w_{k'-1}}(x^{k'-1}) - x^{k'-1}) + e^{k'-1} - q\| \\ &\leq \|x^{k'-1} - q\| - \beta \sum_{i \in I_C} w_{k'-1}(i) \leq \cdots \leq \\ \leq \|x^k - q\| - \beta \sum_{t=k}^{k'-1} \sum_{i \in I_C} w_t(i) &\leq \|x^k - u\| + \|u - q\| - \beta \sum_{t=k}^{k'-1} \sum_{i \in I_C} w_t(i) \\ &< \varepsilon + \|u - q\| - \beta \sum_{t=k}^{k'-1} \sum_{i \in I_C} w_t(i).\end{aligned} \tag{5.35}$$

As a result,

$$\sum_{t=k}^{k'-1} \sum_{i \in I_C} w_t(i) < \frac{\varepsilon}{\beta}. \tag{5.36}$$

In addition, according to the proof of Step 1 and the definition of the ball G (near **(5.33)**), we have $x^t \in G$ for all $t \in \mathbb{N} \cup \{0\}$; in particular, $x^t \in G$ for all $t \in \{k, k+1, \ldots, k'-1\}$. Thus, we can apply Lemma **5.4** repeatedly with $w := w_t$, $x := x^t$, $\lambda := \lambda_t$, $t \in \{k, k+1, \ldots, k'-1\}$ and $z := u$. By using **(3.1)**, **(5.33)**, **(5.34)**, **(5.36)** and the equality $I_C = I_u$, it follows that

$$\|x^{k'} - u\| \leq \|x^k - u\| + \eta \sum_{t=k}^{k'-1} \sum_{i \in I_C} w_t(i) < \varepsilon + \eta \sum_{t=k}^{k'-1} \sum_{i \in I_C} w_t(i) \leq \varepsilon + \eta \frac{\varepsilon}{\beta} = \rho. \tag{5.37}$$

Hence $x^{k'} \in C$, a contradiction to the choice of k'. This contradiction shows that our previous assumption that $u \notin Q$ is invalid. Therefore, $u \in Q$.

Step 2.2: $(x^k)_{k=0}^\infty$ converges to u, a contradiction: So far we have shown that $u \in Q \cap B[x^0, 2\sigma]$ under the assumption that the sequence $(x^k)_{k=0}^\infty$ has at least two distinct accumulation points u and u'. As a result, we can use Lemma **5.6** (where the q there is replaced by u) to conclude that $(\|x^k - u\|)_{k=0}^\infty$ is a decreasing sequence. Since u is an accumulation point of $(x^k)_{k=0}^\infty$, it follows that $(\|x^k - u\|)_{k=0}^\infty$ has a subsequence which converges to 0, and we conclude that $\lim_{k\to\infty}\|x^k-u\|=0$, namely, $(x^k)_{k=0}^\infty$ converges to u. This is a contradiction to the assumption that $(x^k)_{k=0}^\infty$ has two distinct accumulation points.

Step 2.3: $(x^k)_{k=0}^\infty$ converges: The previous step shows that the assumption that $(x^k)_{k=0}^\infty$ has more than one accumulation point is invalid. Since, according to Step 1, $(x^k)_{k=0}^\infty$ has at least one accumulation point, it follows that this sequence has exactly one accumulation point, namely it converges. Denote by x^∞ its limit.

Step 3: $x^\infty\in Q\cap B[x^0,2\sigma]$: Indeed, from Lemma **5.6**, the triangle inequality and the choice of q it follows that for all $k\in\mathbb{N}\cup\{0\}$, we have $\|x^\infty-x^0\|\leq\|x^\infty-q\|+\|q-x^0\|\leq\|x^\infty-x^k\|+\|x^k-q\|+\sigma\leq\|x^\infty-x^k\|+\|x^0-q\|+\sigma\leq\|x^\infty-x^k\|+2\sigma$. Since $\lim_{k\to\infty}\|x^\infty-x^k\|=0$, by letting $k\to\infty$ in the previous inequality, we have $\|x^\infty-x^0\|\leq 2\sigma$, namely $x^\infty\in B[x^0,2\sigma]$.

Now, if $\widehat{I}=\emptyset$, then $\widehat{Q}=X$ and, therefore, $x^\infty\in\widehat{Q}$. Suppose now that $\widehat{I}\neq\emptyset$ and assume to the contrary that $x^\infty\notin\widehat{Q}$. Before going further with the proof, it is noteworthy to say that we cannot use the conclusion of Step 2.1 (with x^∞ instead of u) since this step was based on the false assumption that $(x^k)_{k=0}^\infty$ has at least two different accumulation points.

Returning to our goal, the definition of $\widehat{Q}$ and the assumption that $x^\infty\notin\widehat{Q}$ imply that there exists an index $j\in\widehat{I}$ such that $x^\infty\notin Q_j$. Therefore, $d(x^\infty,Q_j)>0$. As a result, if we denote by B^∞ the closed ball of radius $(1/2)d(x^\infty,Q_j)$ and center x^∞, and denote $C:=B^\infty$ and $I_C:=\{i\in I\mid Q_i\cap C=\emptyset\}$, then we have $j\in I_C$.

Since $(x^k)_{k=0}^\infty$ converges to x^∞, there exists an index $\widetilde{k}\in\mathbb{N}\cup\{0\}$ such that $x^k\in C$ for all integers $k\geq\widetilde{k}$. Consequently, by fixing some $k>\widetilde{k}$ and applying Lemma **5.5** repeatedly with $x:=x^t$, $\lambda:=\lambda_t$, $w:=w_t$ and $e:=e^t$, $t\in[\widetilde{k},k-1]\cap(\mathbb{N}\cup\{0\})$, we conclude that for all $k>\widetilde{k}$,

$$\|x^k-q\|\leq\|x^{k-1}-q\|-\beta\sum_{i\in I_C}w_{k-1}(i)\leq\|x^{\widetilde{k}}-q\|-\beta\sum_{t=\widetilde{k}}^{k-1}\sum_{i\in I_C}w_t(i), \tag{5.38}$$

where $\beta>0$ is the number from Lemma **5.5** with respect to the set C. Since $j\in I_C$, we have $\sum_{t=\widetilde{k}}^{k-1}w_t(j)\leq\sum_{t=\widetilde{k}}^{k-1}\sum_{i\in I_C}w_t(i)$. As a result of this inequality and inequality **(5.38)**, we obtain that for all $k>\widetilde{k}$,

$$\sum_{t=\widetilde{k}}^{k-1}w_t(j)\leq\sum_{t=\widetilde{k}}^{k-1}\sum_{i\in I_C}w_{k-1}(i)\leq\frac{1}{\beta}(\|x^{\widetilde{k}}-q\|-\|x^k-q\|)\leq\frac{\|x^{\widetilde{k}}-q\|}{\beta}. \tag{5.39}$$

By letting $k\to\infty$ we conclude that $\sum_{t=\widetilde{k}}^{\infty}w_t(j)\leq(1/\beta)\|x^{\widetilde{k}}-q\|<\infty$. This is a contradiction since $j\in\widehat{I}$ and hence $\sum_{t=\widetilde{k}}^{\infty}w_t(j)=\infty$. Therefore, the assumption $x^\infty\notin\widehat{Q}$ cannot hold, namely, $x^\infty\in\widehat{Q}$, as required. □

Remark 5.7. A simple condition which ensures that a sequence $(x^k)_{k=0}^{\infty}$ generated by Algorithm **3.1** converges to a point located in the common fixed point set Q is that $Q = \widehat{Q}$; this condition holds if $I = \widehat{I}$. In other words, we simply need to make sure, in advance, that $\sum_{k=0}^{\infty} w_k(i) = \infty$ for each index $i \in I$.

This is a rather mild condition. Indeed, it holds in the case of Example **3.2** when the control is repetitive. It also holds in the case of Example **3.3** when all the weights are equal to $1/m$, or when $w_k(i) = 1/(mk+m)$ for each $k \in \mathbb{N}\cup\{0\}$ and each $i \in I$, with the exception of one index $i_k \in I$ for which $w_k(i_k) = (mk-1)/(mk+m)$ (there is no restriction at all on i_k, and yet $\sum_{k=0}^{\infty} w_k(i) = \infty$ for all $i \in I$; indeed, fix some $i \in I$ and let $k \geq 2$ be arbitrary; either $i \neq i_k$ and then $w_k(i) = 1/(mk+m)$, or $i = i_k$ and then $w_k(i) = (mk-1)/(mk+m) \geq 1/(mk+m)$; hence, $\sum_{k=2}^{\infty} w_k(i) \geq \sum_{k=2}^{\infty}(1/(mk+m)) = \infty$; thus, also $\sum_{k=0}^{\infty} w_k(i) = \infty$).

Another example is the one given in Example **3.4** for the control which cycles periodically between the blocks and gives equal weights to the elements in a specific block. Many more examples can be given.

Remark 5.8. One of the assumptions which is stated in [**1**, Algorithm 1, p. 168 and Theorem 1, p. 171] is that the sequence $(w_k)_{k=0}^{\infty}$ of weight functions is *fair*, that is, for every $i \in I$ there exist infinitely many iteration indices $k \in \mathbb{N} \cup \{0\}$ such that $w_k(i) > 0$. However, this assumption is never used during the proof of the main convergence theorem [**1**, Theorem 1]. In our proof above it is not used as well, and hence we did not even mention it before the proof. In other words, this assumption is unnecessary.

Remark 5.9. We want to say a few words regarding possible extensions of this work and the difficulties that one is expected to face when trying to do so.

One possible extension is to infinite-dimensional spaces. The main difficulty here is the lack of sequential compactness, as can be seen in: Step 1 in the proof of Theorem **5.1** (the existence of a convergent subsequence), Step 2 in the proof of Theorem **5.1** (the existence of accumulation points), the proof of Lemma **5.5** (the existence of accumulation points, Weierstrass Theorem) and the proof of Lemma **5.4** (Weierstrass Theorem).

Another possible extension is to cutters which are not necessarily continuous. The difficulty here is mainly in the proofs of Lemma **5.4** (Weierstrass Theorem for the functions g_i from **(5.10)**) and Lemma **5.5** (because of **(5.15)** and **(5.22)**), but the difficulty in Lemma **5.4** (and only there) can be overcome if one assumes in advance that each cutter maps bounded sets to bounded sets, since Lemma **5.4** is applied (in Step 2.1 of Theorem **5.1**) to closed balls.

A third possible extension is to cutters which are not necessarily defined on the whole space, such as subgradient projections of convex functions which are defined on subsets of the space. Here the whole algorithmic scheme **(3.1)** becomes undefined, but if the subset on which the cutter is defined is closed and convex, then one may overcome the problem (at least for the well-definedness of the scheme) by projecting the right-hand side of **(3.1)** on this subset.

Acknowledgements

The authors thank the referees for their comments, which helped improve the paper. The work of the first and second authors was supported by the ISF-NSFC joint research plan, Grant Number 2874/19.

Data availability

Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

References

[1] Aharoni, R., and Censor, Y. Block-iterative projection methods for parallel computation of solutions to convex feasibility problems. *Linear Algebra and its Applications 120* (1989), 165–175.

[2] Aleyner, A., and Reich, S. Block-iterative algorithms for solving convex feasibility problems in Hilbert and in Banach spaces. *Journal of Mathematical Analysis and Applications 343* (2008), 427–435.

[3] Bauschke, H., and Combettes, P. A weak-to-strong convergence principle for Fejér-monotone methods in Hilbert spaces. *Mathematics of Operations Research 26* (2001), 248–264.

[4] Bauschke, H. H., and Borwein, J. M. On projection algorithms for solving convex feasibility problems. *SIAM Review 38* (1996), 367–426.

[5] Bauschke, H. H., and Combettes, P. L. *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, 2nd edition. Springer International Publishing, Cham, Switzerland, 2017.

[6] Bauschke, H. H., Wang, C., Wang, X., and Xu, J. On the finite convergence of a projected cutter method. *Journal of Optimization Theory and Applications 165* (2015), 901–916.

[7] Bauschke, H.H., Wang, C., Wang, X., and Xu, J. On subgradient projectors, *SIAM Journal on Optimization 25* (2015), 1064–1082.

[8] Bregman, L. M. The relaxation method of finding the common point of convex sets and its application to the solution of problems in convex programming. *Computational Mathematics and Mathematical Physics. 7* (1967), 200–217.

[9] Butnariu, D., and Censor, Y. On the behavior of a block-iterative projection method for solving convex feasibility problems, *International Journal of Computer Mathematics 34 (1990)* , 79–94

[10] Butnariu, D., and Censor, Y. Strong convergence of almost simultaneous block-iterative projection methods in Hilbert spaces. *Journal of Computational and Applied Mathematics 53* (1994), 33–42.

[11] Carmi, A., Censor, Y., and Gurfil, P. Convex feasibility modeling and projection methods for sparse signal recovery. *Journal of Computational and Applied Mathematics 236* (2012), 4318–4335.

[12] Cegielski, A. *Iterative Methods for Fixed Point Problems in Hilbert Spaces*, Springer-Verlag, Berlin, Heidelberg, Germany, 2012.

[13] Cegielski, A., and Censor, Y. Opial-type theorems and the common fixed point problem. In: *Fixed-Point Algorithms for Inverse Problems in Science and Engineering* (New York, NY, 2011), H. H. Bauschke, R. S. Burachik, P. L. Combettes, V. Elser, D. R. Luke, and H. Wolkowicz, Editors, Springer New York, NY, USA, pp. 155–183.

[14] Censor, Y. Weak and strong superiorization: between feasibility-seeking and minimization. *Analele Stiintifice ale Universitatii Ovidius Constanta-Seria Matematica. 23* (2015), 41–54.
[15] Censor, Y. Superiorization and perturbation resilience of algorithms: A continuously updated bibliography, 2022. **http://math.haifa.ac.il/yair/bib-superiorization-censor.html**, with 150 references, as of March 1, 2022.
[16] Censor, Y., CHEN, W. and PAJOOHESH, H. Finite convergence of a subgradient projections method with expanding controls. *Applied Mathematics and Optimization 64* (2011), 273–285.
[17] Censor, Y., Davidi., R., and Herman, G. T. Perturbation resilience and superiorization of iterative algorithms. *Inverse Problems 26* (2010), 065008 (12 pages).
[18] Censor, Y. , Elfving, T. , Herman, G. T. , and Nikazad, T. On diagonally relaxed orthogonal projection methods, *SIAM Journal on Scientic Computing 30* (2008), 473–504.
[19] Censor, Y., and Reem, D. Zero-convex functions, perturbation resilience, and subgradient projections for feasibility-seeking methods. *Mathematical Programming (Series A) 152* (2015), 339–380.
[20] Censor, Y., and Segal, A. On the string averaging method for sparse common fixed-point problems. *International Transactions in Operational Research 16* (2009), 481–494.
[21] Censor, Y., and Segal, A. The split common fixed point problem for directed operators. *Journal of Convex Analysis 16* (2009), 587–600.
[22] Censor, Y., and Zaknoon, M. Methods for finding a common fixed point of a collection of directed operators. Unpublished Technical Report, 2003.
[23] Censor, Y., and Zenios, A. S. *Parallel Optimization: Theory, Algorithms, and Applications*. Oxford University Press, New York, NY, USA, 1997. With a foreword by George B. Dantzig.
[24] Combettes, P. The convex feasibility problem in image recovery. *Advances in Imaging and Electron Physics 95* (1996), 155–270.
[25] Combettes, P. L. Strong convergence of block-iterative outer approximation methods for convex optimization. *SIAM Journal on Control and Optimization 38* (2000), 538–565.
[26] Combettes, P. L. Quasi-Fejérian analysis of some optimization algorithms. In: *Inherently Parallel Algorithms in Feasibility and Optimization and their Applications*, Elsevier Science Publishers, Amsterdam, The Netherlands (2001), D. Butnariu, Y. Censor, and S. Reich, Editors., pp. 115–152.
[27] Davidi, R., Herman, G. T., and Censor, Y. Perturbation-resilient block-iterative projection methods with application to image reconstruction from projections. *International Transactions in Operational Research, 16* (2009), 505–524.
[28] Flåm, S. D., and Zowe, J. Relaxed outer projections, weighted averages and convex feasibility. *BIT 30* (1990), 289–300.
[29] Gholami, M. R., Wymeersch, H., Ström, E. G., and Rydström, M. Wireless network positioning as a convex feasibility problem. *EURASIP Journal on Wireless Communications and Networking 2011* (2011), 161 (15 pages).
[30] Gibali, A., Küfer, KH., Reem, D. and Süss, P. A generalized projection-based scheme for solving convex constrained optimization problems. *Comput Optim Appl 70*, 737–762 (2018).
[31] Herman, G. T. Superiorization for image analysis. In: *Combinatorial Image Analysis*, *Lecture Notes in Computer Science*. Springer, vol. 8466, 2014, pp. 1–7.
[32] Hero, A., and Blatt, D. Sensor network source localization via projection onto convex sets (POCS). In *Proceedings. (ICASSP '05). IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005.* (2005), vol. 3, pp. 689–692.

[33] Hiriart-Urruty, J.-B., and Lemaréchal, C. *Fundamentals of Convex Analysis* (Abridged version of "Convex analysis and Minimization Algorithms. I, II, Springer, 1993"), Grundlehren Text Editions, Springer-Verlag, Berlin, 2001.

[34] Ibaraki, T., and Takahashi, W. Block iterative methods for a finite family of generalized nonexpansive mappings in Banach spaces. *Numerical Functional Analysis and Optimization 29* (2008), 362–375

[35] Karonis, N. T., Dun, K. L., Ordoñez, C. E. , Erdely, B., Uram, T. D., Olson, E. C., Coutrakon, G., and Papka, M. E. Distributed and hardware accelerated computing for clinical medical imaging using proton computed tomography (pCT), *Journal of Parallel and Distributed Computing 73* (2013), 1605–1612.

[36] Kolobov, V. I., Reich., S., Zalas, R. Weak, strong, and linear convergence of a double-layer fixed point algorithm. *SIAM Journal on Optimization 27* (2017), 1431–1458.

[37] Necoara, I. Stochastic block projection algorithms with extrapolation for convex feasibility problems, *Optimization Methods and Software, to appear* (published online: 2022, DOI: 10.1080/10556788.2021.1998492)

[38] Reem, D., and De Pierro, A. R. A new convergence analysis and perturbation resilience of some accelerated proximal forward-backward algorithms with errors. *Inverse Problems 33* (2017), 044001 (28pp).

[39] Reich, S., and Zalas, R. A modular string averaging procedure for solving the common fixed point problem for quasi-nonexpansive mappings in Hilbert space. *Numerical Algorithms 72* (2016), 297–323.

[40] Zaknoon, M. *Algorithmic Developments for the Convex Feasibility Problem.* PhD thesis, University of Haifa, Haifa, Israel, 2003.

[41] Zaslavski, AJ. *Approximate solutions of common fixed point problems.* Springer Optimization and Its Applications, Springer, Cham, 2016.

Department of Mathematics, University of Haifa, Mt. Carmel, 3498838 Haifa, Israel.
Email address: (Yair Censor) yair@math.haifa.ac.il

The Center for Mathematics and Scientific Computation (CMSC), University of Haifa, Mt. Carmel, Haifa, 3498838, Israel.
Email address: (Daniel Reem) dream@math.haifa.ac.il

Department of Mathematics, The Arab Academic College for Education, 22 HaHashmal Street, Haifa, 32623, Israel
Email address: (Maroun Zaknoon) zaknoon@arabcol.ac.il