Sparse string-averaging and split common fixed points

Yair Censor and Alexander Segal
({yair,asegal}@math.haifa.ac.il)

Department of Mathematics, University of Haifa
Mt. Carmel, Haifa 31905, Israel

May 30, 2008. Revised: October 7, 2008.

Abstract

We review the common fixed point problem for the class of directed operators. This class is important because many commonly used nonlinear operators in convex optimization belong to it. We present our recent definition of sparseness of a family of operators and discuss a string-averaging algorithmic scheme that favorably handles the common fixed points problem when the family of operators is sparse. We also review some recent results on the multiple operators split common fixed point problem which requires to find a common fixed point of a family of operators in one space whose image under a linear transformation is a common fixed point of another family of operators in the image space.

1   Introduction

In this paper we review some recent iterative projection methods for the common fixed point problem for the class of *directed operators*. This class is important because many commonly used nonlinear operators in convex optimization belong to it. We present our recent definition of *sparseness of a family of operators* and discuss a *string-averaging* algorithmic scheme

that favorably handles the common fixed point problem when the family of operators is sparse. For the convex feasibility problem a new subgradient projections algorithmic scheme is obtained. We also review some recent results on the *multiple split common fixed point problem* which requires to find a common fixed point of a family of operators in one space whose image under a linear transformation is a common fixed point of another family of operators in the image space. The presentation is admittedly biased towards our own work but contains also pointers to other works in the literature.

Projection algorithms employ projections onto convex sets in various ways. This class of algorithms has witnessed great progress in recent years and its member algorithms have been applied with success to fully discretized models of problems in image reconstruction and image processing, see, e.g., Stark and Yang [66], Censor and Zenios [33]. Our aim in this paper is to introduce the reader to certain algorithmic structures and specific algorithms inspired by projection methods and used for solving the sparse common fixed point problem and the split common fixed point problem.

Given a finite family of operators $\{T_i\}_{i=1}^m$ acting on the Euclidean space R^n with fixed points sets $\operatorname{Fix} T_i \neq \emptyset$, $i = 1, 2, \ldots, m$, the *common fixed point problem* is to find a point

$$x^* \in \cap_{i=1}^m \operatorname{Fix} T_i, \tag{1}$$

In this paper we focus on the common fixed point problem for sparse directed operators. We use the term *directed operators* for operators in the $\mathcal{T}$-class of operators as defined and investigated by Bauschke and Combettes in [6] and by Combettes in [41]. The first topic that we review here is the behavior of iterative algorithmic schemes when we have sparse operators and, for that purpose, we give a definition of sparseness of a family of operators. The algorithms that are in use to find a common fixed point can be, from their structural view point, sequential, when only one operator at a time is used in each iteration, or simultaneous (parallel), when all operators in the given family are used in each iteration. There are algorithmic schemes which encompass sequential and simultaneous properties. These are the, so called, *string-averaging* [24] and *block-iterative projections* (BIP) [1] schemes, see also [33]. It turns out that the sequential and the simultaneous algorithms are special cases of both the string-averaging and of the BIP algorithmic schemes.

In [30] we proposed and studied a string-averaging algorithmic scheme that enables *component-wise* weighting. Its origins lie in [26] where a simul-

taneous projection algorithm, called *component averaging* (CAV), for systems of linear equations that uses component-wise weighting was proposed. Such weighting enables, as shown and demonstrated experimentally on problems of image reconstruction from projections in [26], significant and valuable acceleration of the early algorithmic iterations due to the high sparsity of the system matrix appearing there. A block-iterative version of CAV, named BICAV, was introduced later in [27]. Full mathematical analyses of these methods, as well as their companion algorithms for linear inequalities, were presented by Censor and Elfving [23] and by Jiang and Wang [57]

The second topic that we review concerns the *multiple operators split common fixed point problem.* The multiple-sets split feasibility problem requires to find a point closest to a family of closed convex sets in one space such that its image under a linear transformation will be closest to another family of closed convex sets in the image space. It serves as a model for inverse problems where constraints are imposed on the solutions in the domain of a linear operator as well as in the operator's range. It generalizes the convex feasibility problem and the two-sets split feasibility problem. Formally, given nonempty closed convex sets $C_i \subseteq R^n$, $i = 1, 2, \ldots, t$, in the n-dimensional Euclidean space R^n, and nonempty closed convex sets $Q_j \subseteq R^m$, $j = 1, 2, \ldots, r$, and an $m \times n$ real matrix A, the *multiple-sets split feasibility problem* (MSSFP) is

$$\text{find a vector } x^* \in C := \cap_{i=1}^t C_i \text{ such that } Ax^* \in Q := \cap_{i=1}^r Q_j. \qquad (2)$$

Such MSSFPs, formulated in [25], arise in the field of intensity-modulated radiation therapy (IMRT) when one attempts to describe physical dose constraints and equivalent uniform dose (EUD) constraints within a single model, see [21]. This generalizes, of course, the convex feasibility problem.

The *convex feasibility problem* (CFP) is to find a point x^* in the intersection C of m closed convex subsets $C_1, C_2, \ldots, C_m \subseteq R^n$. Each C_i is expressed as

$$C_i = \{x \in R^n \mid f_i(x) \leq 0\}, \qquad (3)$$

where $f_i : R^n \to R$ is a convex function, so the CFP requires a solution of the system of convex inequalities

$$f_i(x) \leq 0, \quad i = 1, 2, \ldots, m. \qquad (4)$$

The convex feasibility problem is a special case of the common fixed point problem, where the directed operators are the subgradient projectors relative

to f_i (see, Example 4 and Lemma 5 below). It is a fundamental problem in many areas of mathematics and the physical sciences, see, e.g., Combettes [36, 40] and references therein. It has been used to model significant real-world problems in image reconstruction from projections, see, e.g., Herman [54], in radiation therapy treatment planning, see Censor, Altschuler and Powlis [20] and Censor [19], and in crystallography, see Marks, Sinkler and Landree [60], to name but a few, and has been used under additional names such as *set theoretic estimation* or the *feasible set approach*. A common approach to such problems is to use projection algorithms, see, e.g., Bauschke and Borwein [4], which employ *orthogonal projections* (nearest point mappings) onto the individual sets C_i.

1.1 Projection methods: Advantages and earlier work

The reason why feasibility problems of various kinds are looked at from the viewpoint of projection methods can be appreciated by the following brief comments regarding projection methods in general. Projections onto sets are used in a wide variety of methods in optimization theory but not every method that uses projections really belongs to the class of projection methods. *Projection methods* are iterative algorithms that use projections onto sets while relying on the general principle that when a family of (usually closed and convex) sets is present then projections onto the given individual sets are easier to perform then projections onto other sets (intersections, image sets under some transformation, etc.) that are derived from the given individual sets.

A projection algorithm reaches its goal, related to the whole family of sets, by performing projections onto the individual sets. Projection algorithms employ projections onto convex sets in various ways. They may use different kinds of projections and, sometimes, even use different projections within the same algorithm. They serve to solve a variety of problems which are either of the feasibility or the optimization types. They have different algorithmic structures, of which some are particularly suitable for parallel computing, and they demonstrate nice convergence properties and/or good initial behavior patterns.

Apart from theoretical interest, the main advantage of projection methods, which makes them successful in real-world applications, is computational. They commonly have the ability to handle huge-size problems of dimensions beyond which other, more sophisticated currently available, meth-

ods cease to be efficient. This is so because the building bricks of a projection algorithm are the projections onto the given individual sets (assumed and actually easy to perform) and the algorithmic structure is either sequential or simultaneous (or in-between). Sequential algorithmic structures cater for the row-action approach (see Censor [18]) while simultaneous algorithmic structures favor parallel computing platforms, see, e.g., Censor, Gordon and Gordon [26]. The field of projection methods is vast and we can only mention here a few recent works that can give the reader some good starting points. Such a list includes, among many others, the paper of Lakshminarayanan and Lent [58] on the SIRT method, the works of Crombez [43, 46], the connection with variational inequalities, see, e.g., Noor [62], Yamada's [68] which is motivated by real-world problems of signal processing, and the many contributions of Bauschke and Combettes, see, e.g., Bauschke, Combettes and Kruk [7] and references therein. Consult Bauschke and Borwein [4] and Censor and Zenios [33, Chapter 5] for a tutorial review and a book chapter, respectively. Systems of linear equations, linear inequalities, or convex inequalities are all encompassed by the convex feasibility problem which has broad applicability in many areas of mathematics and the physical and engineering sciences. These include, among others, optimization theory (see, e.g., Eremin [51], Censor and Lent [28] and Chinneck [34]), approximation theory (see, e.g., Deutsch [47] and references therein) and image reconstruction from projections in computerized tomography (see, e.g., Herman [54, 55], Censor [18]).

2   Directed operators

We recall the definitions and results on directed operators and their properties as they appear in Bauschke and Combettes [6, Proposition 2.4] and Combettes [41], which are also sources for further references on the subject. Let $\langle x, y \rangle$ and $\|x\|$ be the Euclidean inner product and norm, respectively, in R^n.

Given $x, y \in R^n$ we denote the half-space

$$H(x, y) := \left\{ u \in R^n \mid \langle u - y, x - y \rangle \leq 0 \right\}. \tag{5}$$

Definition 1 *An operator* $T : R^n \to R^n$ *is called* `directed` *if*

$$\operatorname{Fix} T \subseteq H(x, T(x)), \ \textit{for all } x \in R^n, \tag{6}$$

or, equivalently,

$$\text{if } z \in \text{Fix}\, T \text{ then } \langle T(x) - x, T(x) - z \rangle \leq 0, \text{ for all } x \in R^n. \qquad (7)$$

The class of directed operators is the $\mathcal{T}$-class of operators of Bauschke and Combettes [6] who defined directed operators (although without using this name) and showed (see [6, Proposition 2.4]) (i) that the set of all fixed points of a directed operator T with nonempty $\text{Fix}\, T$ is closed and convex because

$$\text{Fix}\, T = \bigcap_{x\,\in\,R^n} H(x, T(x)), \qquad (8)$$

and (ii) that the following holds

$$\text{If } T \in \mathcal{T} \text{ then } I + \lambda(T - I) \in \mathcal{T}, \text{ for all } \lambda \in [0,1], \qquad (9)$$

where I is the identity operator. The localization of fixed points is discussed in [53, pages 43-44]. In particular, it is shown there that a firmly nonexpansive operator, namely, an operator $N : R^n \to R^n$ that fulfils

$$\|N(x) - N(y)\|^2 \leq \langle N(x) - N(y), x - y \rangle, \text{ for all } x, y \in R^n, \qquad (10)$$

satisfies (8) and is, therefore, a directed operator. The class of directed operators, includes additionally, according to [**?**, Proposition 2.3], among others, the resolvents of a maximal monotone operators, the orthogonal projections and the subgradient projectors (see Example 4 below). Note that every directed operator belongs to the class of operators $\mathcal{F}^0$, defined by Crombez [45, p. 161],

$$\mathcal{F}^0 := \{T : R^n \to R^n \mid \|Tx - q\| \leq \|x - q\| \text{ for all } q \in \text{Fix}\, T \text{ and } x \in R^n\}, \qquad (11)$$

whose elements are called elsewhere quasi-nonexpansive or paracontracting operators.

Definition 2 *An operator $T : R^n \to R^n$ is said to be* `closed` *at $y \in R^n$ if for every $\overline{x} \in R^n$ and every sequence $\{x^k\}_{k=0}^{\infty}$ in R^n, such that, $\lim_{k\to\infty} x^k = \overline{x}$ and $\lim_{k\to\infty} T(x^k) = y$, we have $T(\overline{x}) = y$.*

For instance, the orthogonal projection onto a closed convex set is everywhere a closed operator, due to its continuity.

Remark 3 *[41] If $T : R^n \rightarrow R^n$ is nonexpansive, then $T - I$ is closed on R^n.*

The next example and lemma recall the notion of the subgradient projector.

Example 4 *Let $f : R^n \rightarrow R$ be a convex function such that the level-set $F := \{x \in R^n \mid f(x) \leq 0\}$ is nonempty. The operator*

$$\Pi_F(y) := \begin{cases} y - \dfrac{f(y)}{\|q\|^2}q, & if \quad f(y) > 0, \\ y, & if \quad f(y) \leq 0, \end{cases} \tag{12}$$

where q is a selection from the subdifferential set $\partial f(y)$ of f at y, is called a `subgradient projector relative to` *f. See, e.g., [6, Proposition 2.3(iv)].*

Lemma 5 *Let $f : R^n \rightarrow R$ be a convex function, let $y \in R^n$ and assume that the level-set $F \neq \emptyset$. For any $q \in \partial f(y)$, define the closed convex set*

$$L = L_f(y, q) := \{x \in R^n \mid f(y) + \langle q, x - y \rangle \leq 0\}. \tag{13}$$

Then the following hold:

(i) $F \subseteq L$. If $q \neq 0$ then L is a half-space, otherwise $L = R^n$.

(ii) Denoting by $P_L(y)$ the orthogonal projection of y onto L,

$$P_L(y) = \Pi_F(y). \tag{14}$$

(iii) $P_L - I$ is closed at 0.

Consider a finite family $T_i : R^n \rightarrow R^n$, $i = 1, 2, \ldots, m$, of operators. In sequential algorithms for solving the common fixed point problem the order by which the operators are chosen for the iterations is determined by a *control sequence* of indices $\{i(k)\}_{k=0}^{\infty}$, see, e.g., [33, Definition 5.1.1].

Definition 6 *(i) **Cyclic control.** A control sequence is* `cyclic` *if $i(k) = k \bmod m + 1$, where m is the number of operators in the common fixed point problem.*

*(ii) **Almost cyclic control.** $\{i(k)\}_{k=0}^{\infty}$ is* `almost cyclic` *on $\{1, 2, \ldots, m\}$, if $1 \leq i(k) \leq m$ for all $k \geq 0$, and there exists an integer $c \geq m$ (called the* `almost cyclicality constant`*), such that, for all $k \geq 0$, $\{1, 2, \ldots, m\} \subseteq \{i(k+1), i(k+2), \ldots, i(k+c)\}$.*

The notions "cyclic" and "almost cyclic" are sometimes also called "periodic" and "quasi-periodic", respectively, see, e.g., [48].

Given a finite family $T_i : R^n \to R^n$, $i = 1, 2, \ldots, m$, of directed operators with a nonempty intersection of their fixed points sets, such that $T_i - I$ are closed at 0, for every $i \in \{1, 2, \ldots, m\}$. The following algorithm for finding a common fixed point of such a family is a special case of [41, Algorithm 6.1]. We will use it in the sequel.

Algorithm 7 *Almost Cyclic Sequential Algorithm (ACSA) for solving common fixed point problem*

Initialization: $x^0 \in R^n$ *is an arbitrary starting point.*

Iterative Step: *Given* x^k, *compute* x^{k+1} *by*

$$x^{k+1} = x^k + \lambda_k(T_{i(k)}\left(x^k\right) - x^k). \tag{15}$$

Control: $\{i(k)\}_{k=0}^{\infty}$ *is almost cyclic on* $\{1, 2, \ldots, m\}$.

Relaxation parameters: $\{\lambda_k\}_{k=0}^{\infty}$ *are confined to the interval* $[\varepsilon, 2 - \varepsilon]$, *for some fixed user-chosen* $\varepsilon > 0$.

3   The string-averaging algorithmic scheme

We review here different modifications of the string-averaging paradigm, adapted to handle the convex feasibility problem and the common fixed point problem. The string-averaging algorithmic scheme has attracted attention recently and further work on it has been reported since its presentation in [24]. In that paper the string-averaging algorithmic scheme for the solution of convex feasibility problem was proposed and a scheme employing Bregman projections was analyzed with the aid of an extended product space formalism.

To define string-averaging let the *string* S_p, for $p = 1, 2, \ldots, t$, be a finite, nonempty ordered subset of elements taken from $\{1, 2, \ldots, m\}$ of the form

$$S_p := \left\{i_1^p, i_2^p, \ldots, i_{\gamma(p)}^p\right\}. \tag{16}$$

The length $\gamma(p)$ of the string S_p is the number of its elements. We do not require that the strings $\{S_p\}_{p=1}^t$ should be disjoint. Suppose that there is a set $Q \subseteq R^n$ such that there are operators $V_1, V_2, \ldots, V_m$ mapping Q into Q and an operator V which maps Q^t into Q. Then the string-averaging prototypical scheme is as follow.

Algorithm 8 *The string averaging prototypical algorithmic scheme [24]*

    Initialization: $x^0 \in Q$ is an arbitrary starting point.
    Iterative Step: Given the current iterate x^k,
    (i) calculate, for all $p = 1, 2, \ldots, t$,

$$M_p(x^k) := V_{i^p_{\gamma(p)}} \ldots V_{i^p_2} V_{i^p_1}(x^k), \tag{17}$$

    (ii) and then calculate,

$$x^{k+1} = V(M_1(x^k), M_2(x^k), \ldots, M_t(x^k)). \tag{18}$$

For every $p = 1, 2, \ldots, t$, this algorithmic scheme applies to x^k successively the operators whose indices belong to the p-th string. This can be done in parallel for all strings and then the operator V maps all end-points onto the next iterate x^{k+1}. This is indeed an algorithm provided that the operators $\{V_i\}_{i=1}^m$ and V all have algorithmic implementations. In this framework we get a sequential algorithm by the choice $t = 1$ and $S_1 = \{1, 2, \ldots, m\}$ and a simultaneous algorithm by the choice $t = m$ and $S_p = \{p\}$, $p = 1, 2, \ldots, t$.

We may demonstrate the underlying idea of the string-averaging prototypical algorithmic scheme with the aid of Figure 1. For simplicity, we take the convex sets to be hyperplanes, denoted by H_1, H_2, H_3, H_4, H_5, and H_6, and assume all operators V_i to be orthogonal projections onto the hyperplanes. The operator V is taken as a convex combination

$$V(x^1, x^2, \ldots, x^t) = \sum_{p=1}^t \omega_p x^p, \tag{19}$$

with $\omega_p > 0$, for all $p = 1, 2, \ldots, t$, and $\sum_{p=1}^t \omega_p = 1$.

Figure 1(a) depicts the purely sequential algorithmic structure. This is the so-called POCS (Projections Onto Convex Sets) algorithm which coincides, for the case of hyperplanes, with the Kaczmarz algorithm, see, e.g., Algorithms 5.2.1 and 5.4.3, respectively, in [33]. The fully simultaneous algorithmic structure appears in Figure 1(b). With orthogonal reflections instead of orthogonal projections it was first proposed, by Cimmino [35], for solving linear equations, see also Benzi [9]. Here the current iterate x^k is projected on all sets simultaneously and the next iterate x^{k+1} is a convex combination of the projected points. In Figure 1(c) we show how averaging of *successive*

projections (as opposed to averaging of parallel projections in Figure 1(b)) works. In this case $t = m$ and $S_p = (1, 2, \ldots, p)$, for $p = 1, 2, \ldots, t$. This scheme, appearing in Bauschke and Borwein [4], inspired our formulation of the general string-averaging prototypical algorithmic scheme whose action is demonstrated in Figure 1(d). In this example it averages, via convex combinations, the end-points obtained from strings of sequential projections and in this figure the strings are $S_1 = (1, 3, 5, 6)$, $S_2 = (2)$, $S_3 = (6, 4)$. Such schemes offer a variety of options for steering the iterates towards a solution of the convex feasibility problem. It is an *inherently parallel* scheme in that its mathematical formulation is parallel (like the fully simultaneous method mentioned above). We use this term to contrast such algorithms with others which are sequential in their mathematical formulation but can, sometimes, be implemented in a parallel fashion based on appropriate model decomposition (i.e., depending on the structure of the underlying problem). Being inherently parallel, this algorithmic scheme enables flexibility in the actual manner of implementation on a parallel machine.
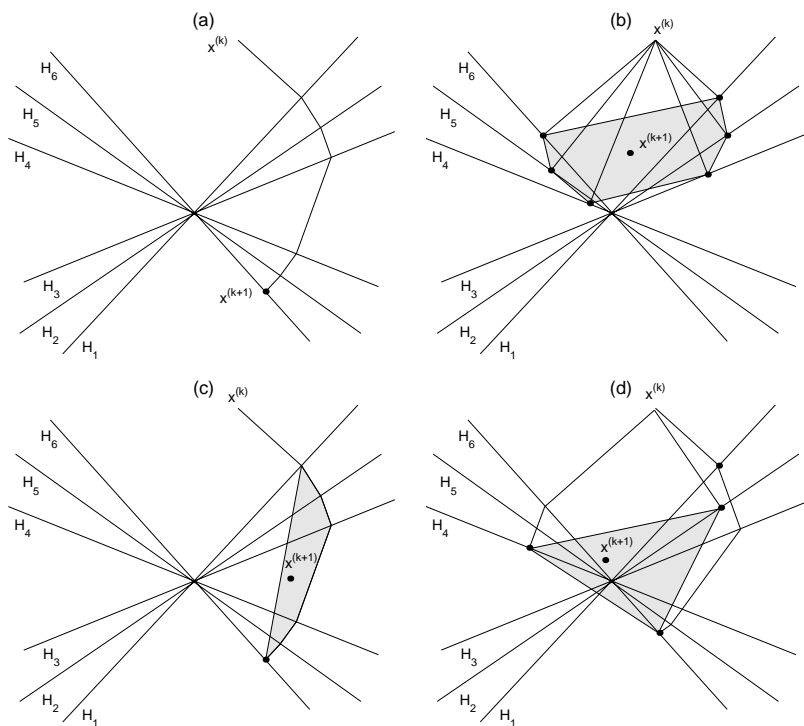

Figure 1. (a) Sequential projections. (b) Fully simultaneous projections.

(c) Averaging of sequential projections. (d) String-averaging. (Reproduced from Censor, Elfving and Herman [24]).

At the extremes of the "spectrum" of possible specific algorithms, derivable from the string averaging prototypical algorithmic scheme, are the generically sequential method, which uses one set at a time, and the fully simultaneous algorithm, which employs all sets at each iteration. For results on the behavior of the fully simultaneous algorithm with orthogonal projections in the inconsistent case see, e.g., Combettes [39] or Iusem and De Pierro [56]. The "block-iterative projections" (BIP) scheme of Aharoni and Censor [1] also has the sequential and the fully simultaneous methods as its extremes in terms of block structures (see also Butnariu and Censor [16], Bauschke and Borwein [4], Bauschke, Borwein and Lewis [5], Elfving [50], Eggermont, Herman and Lent [49] and, recently, Aleyner ans Reich [2]). The question whether there are any other relationships between the BIP and the string-averaging prototypical algorithmic schemes is of theoretical interest and is still open. However, the string-averaging prototypical algorithmic structure gives users a tool to design many new inherently parallel computational schemes.

The behavior of the string-averaging algorithmic scheme, with orthogonal projections, in the inconsistent case when the intersection $Q = \cap_{i=1}^{m} Q_i$ is empty was studied by Censor and Tom in [32]. They defined the *projection along the string S_p operator* as the composition of orthogonal projections onto sets indexed by S_p, that is,

$$V_p := P_{i_{\gamma(t)}^p} \cdots P_{i_2^p} P_{i_1^p}, \text{ for } p = 1, 2, \ldots, t, \tag{20}$$

and, given a positive weight vector $\omega \in R^t$, they used as the algorithmic operator V the following

$$V = \sum_{p=1}^{t} \omega_p V_p. \tag{21}$$

Using this V the following string-averaging algorithm is obtained.

Algorithm 9
  Initialization: $x^0 \in R^n$ *is an arbitrary starting point.*
  Iterative Step: *Given x^k, use (20) and (21) to compute x^{k+1}*

$$x^{k+1} = V(x^k). \tag{22}$$

11

Theorem 10 *[32] Let $Q_1, Q_2, \ldots, Q_m$, be nonempty closed convex subsets of R^n. If for at least one $x^0 \in R^n$ the sequence $\left\{x^k\right\}_{k \geq 0}$, generated by the string-averaging algorithm (Algorithm 9 with V as in (21)), is bounded then any sequence $\left\{x^k\right\}_{k \geq 0}$, generated by the string averaging algorithm (Algorithm 9 with V as in (21)), converges for any $x^0 \in R^n$.*

The convergence of this string-averaging method in the inconsistent case was proved using translation of the algorithm into a fully sequential algorithm in the product space.

In Bauschke, Matoušková and Reich [8] string-averaging was studied in Hilbert space. In Crombez [42] the string-averaging algorithmic paradigm is used to find common fixed points of strict paracontractive operators in the $m-$dimensional Euclidean space. Given a finite set of strict paracontractive continuous operators having a nonempty set of common fixed points, finite strings of sequential iterations of them are formed, leading to a corresponding set of general paracontractions having the same common set of fixed points. By suitably averaging this set of operators, a fully parallel and a block-iterative algorithm can be obtained, both with a variable relaxation parameter. In Crombez [44] string-averaging is used to produce a asynchronous parallel algorithm that leads to a common fixed point of directed operators in Hilbert space. The assumption in this paper is that there exists a common fixed point that is an interior point. Averaging strings of different length of sequential iterations leads to an asynchronous parallel method which reaches a common fixed point in a finite number of steps. In Bilbao-Castro, Carazo, García and Fernández [12], an implementation of the string-averaging method to electron microscopy is reported. Butnariu, Davidi, Herman and Kazantsev [17] call a certain class of string-averaging methods the Amalgamated Projection Method and show that the algorithms in this class converge to solutions of the consistent convex feasibility problem, and that their convergence is stable under summable perturbations. A variant of this approach was proposed to approximate the minimum of a convex functional subject to convex constraints. The iterative procedure studied in Butnariu, Reich and Zaslavski [11, Sections 6 and 7] is also a particular case of the string-averaging method. In Rhee [63] the string-averaging scheme is applied to a problem in approximation theory.

3.1 String-averaging for the sparse common fixed point problem

The notion of sparseness is well understood and used for matrices and, from there, the road to sparseness of the Jacobian (or generalized Jacobian) matrix as an indicator of sparseness of nonlinear operators is short, see, e.g., Betts and Frank [10]. Our definition of sparseness of operators does not require differentiability (or subdifferentiability) and generalizes those previous notions.

In our algorithmic scheme, designed to efficiently handle sparsity, we assume that a finite family of directed operators (see Definition 1) $\{T_i\}_{i=1}^m$ is given with $\cap_{i=1}^m \operatorname{Fix} T_i \neq \emptyset$ such that $T_i - I$ are closed at 0, for every $i \in \{1, 2, \ldots, m\}$. After applying the operators $\{T_i\}_{i=1}^m$ along strings, the endpoints are averaged not by taking a plain convex combination but by doing a, so called, *component-averaging* step. The component averaging principle, introduced for linear systems in [26], [27], is a useful tool for handling sparseness in the linear case.

To define sparseness of the set of operators $\{T_i\}_{i=1}^m$ we need to speak about zeros of the vectors $x - T_i(x)$.

Definition 11 *Let $T : R^n \to R^n$ be a directed operator. If $(x - T(x))_j = 0$, for all $x \notin \operatorname{Fix} T$ then j is called a **void** of T and we write $j = \operatorname{void} T$.*

For every $i \in \{1, 2, \ldots, m\}$ define the following sets

$$Z_i := \{(i,j) \mid 1 \leq j \leq n, \ j = \operatorname{void} T_i\}, \tag{23}$$

i.e., Z_i contains all the pairs (i,j), such that $(x - T_i(x))_j = 0$, for all $x \notin \operatorname{Fix} T_i$.

Definition 12 *The family of directed operators $\{T_i\}_{i=1}^m$ will be called **sparse** if the set $Z := \cup_{i=1}^m Z_i$ is nonempty and contains many elements.*

Remark 13 *The word "many" in Definition 12 is meant to say that the more pairs (i,j) are contained in Z the higher is the sparseness of the family. It is of some interest to note that sparseness of matrices was considered as early as in 1971. Wilkinson [67, p. 191] refers to it by saying: "We shall refer to a matrix as dense if the percentage of zero elements or its distribution is such as to make it uneconomic to take advantage of their presence". Obviously, denseness is meant here as an opposite of sparseness.*

Denote by I_j, $1 \leq j \leq n$, the set of indices of strings that contain an index of an operator T_i for which $(i, j) \notin Z_i$, i.e.,

$$I_j := \{p \mid 1 \leq p \leq t, (i, j) \notin Z_i \text{ for some } i \in S_p\} \qquad (24)$$

and let $s_j = |I_j|$ (the cardinality of I_j). Equivalently,

$$I_j = \{p \mid 1 \leq p \leq t, \; j \neq \text{void} \, T_i \text{ for some } i \in S_p\}. \qquad (25)$$

Definition 14 *[52, Definition 1] The* `component-wise string-averaging` `operator relative to the family of strings` $S := \{S_1, S_2, \ldots, S_t\}$ *is a mapping* $CA_S : R^{n \times t} \to R^n$*, defined as follows. For* $x^1, x^2, \ldots, x^t \in R^n$*,*

$$\left(CA_S(x^1, x^2, \ldots, x^t)\right)_j := (1/s_j) \sum_{p \in I_j} x_j^p, \quad \text{for all } \; 1 \leq j \leq n, \qquad (26)$$

where x_j^p *is the j-th component of* x^p*, for* $1 \leq p \leq t$*.*

Our new scheme performs sequential steps within each of the strings of the family S and merges the resulting end-points by the component-wise string-averaging operator (26) as follows.

Algorithm 15
   Initialization: $x^0 \in R^n$ *is an arbitrary starting point and define an integer constant N, such that $N \geq m$.*
   Iterative step: *Given x^k, compute x^{k+1} as follows:*
   (i) For every $1 \leq p \leq t$ (possibly in parallel): Execute a finite number, not exceeding N, of iterative steps of the form (15), on the operators $\{T_i\}_{i \in S_p}$ of the p-th string and denote the resulting end-points by $\{\overline{x}^p\}_{p=1}^t$.
   (ii) Apply
$$x^{k+1} = CA_S(\overline{x}^1, \overline{x}^2, \ldots, \overline{x}^t). \qquad (27)$$

Theorem 16 *Let $\{T_i\}_{i=1}^m$ be a family of directed operators with $\cap_{i=1}^m \text{Fix} \, T_i \neq \emptyset$ such that $T_i - I$ are closed at 0, for every $i \in \{1, 2, \ldots, m\}$. Any sequence $\{x^k\}_{k=0}^\infty$, generated by the Algorithm 15, converges to a solution of (1).*

In a recent paper by Gordon and Gordon [52] a new parallel "Component-Averaged Row Projections (CARP)" method for the solution of large sparse linear systems was introduced. It proceeds by dividing the equations into

nonempty, not necessarily disjoint, sets (strings), performing Kaczmarz (row-action) projections within the strings, and merging the results by component-averaging operations to form the next iterate. As shown in [52], using orthogonal projections onto convex sets, this method and its convergence proof also apply to the consistent nonlinear CFP.

In contrast, when applied to a CFP, Algorithm 15 gives rise to a method which is structurally similar to CARP but uses subgradient projections instead of orthogonal projections. This is, of course, a development that might be very useful for CFPs with nonlinear convex sets for which each orthogonal projection mandates an inner-loop of distance minimization.

Sparseness of the nonlinear system (4) can be defined in compliance with Definitions 11 and 12 by speaking about zeros of the subgradients of the functions f_i and to do so we use the next definition.

Definition 17 *Let $f_i : R^n \to R$, $i = 1, 2, \ldots, m$, be convex functions. For any $x \in R^n$, the $m \times n$ matrix $Q(x) = (q_{ij})_{i=1, j=1}^{m, n}$ is called* **a generalized Jacobian of the family of functions** $\{f_i\}_{i=1}^{m}$ **at the point** x *if $q_{ij} \equiv q_j^i$, for all i and all j, for some $q^i = (q_j^i)_{j=1}^{n}$ such that $q^i \in \partial f_i(x)$.*

This definition coincides in our case with Clarke's generalized Jacobian, see [37] and [38]. A generalized Jacobian $Q(x)$ of the functions in (4) is not unique because of the possibility to fill it up with different subgradients from each subdifferential set. In case all f_i are differentiable the generalized Jacobian reduces to the usual Jacobian.

Define for every $i \in \{1, 2, \ldots, m\}$ the following sets

$$Z_i := \{(i, j) \mid 1 \leq j \leq n, \ f_i(x) \text{ is independent of } x_j \text{ for all } x \in R^n\}. \quad (28)$$

A mapping $F : R^n \to R^m$ given by $F(x) = \{f_i(x)\}_{i=1}^{m}$ will be called *sparse* if some of its component functions f_i do not depend on some of their variables x_j which means that $Z = \cup_{i=1}^{m} Z_i \neq \emptyset$. The more pairs (i, j) are contained in Z the higher is the sparseness of the mapping F.

Recall the cyclic subgradient projections (CSP) method for the CFP (studied in [28]) which is a special version of the ACSA algorithm (Algorithm 7).

Algorithm 18 *Cyclic Subgradient Projections (CSP)*
    Initialization: $x^0 \in R^n$ *is arbitrary.*

Iterative step:

$$x^{k+1} := \begin{cases} x^k - \lambda_k \dfrac{f_{i(k)}(x^k)}{\|q^k\|^2} q^k, & \text{if } f_{i(k)}(x^k) > 0, \\ x^k, & \text{if } f_{i(k)}(x^k) \leq 0, \end{cases} \qquad (29)$$

where $q^k \in \partial f_{i(k)}(x^k)$ is a subgradient of $f_{i(k)}$ at the point x^k.

Relaxation parameters: *$\{\lambda_k\}_{k=0}^{\infty}$ are confined to the interval $[\varepsilon, 2 - \varepsilon]$, for some fixed user-chosen $\varepsilon > 0$.*

Control: *Almost cyclic on $\{1, 2, \ldots, m\}$.*

According to our scheme the algorithm for solving the CFP performs CSP steps within the strings and merges the results by the $CA_S(x_1, x_2, \ldots, x_t)$ component-averaging operation.

Algorithm 19

Initialization: *$x^0 \in R^n$ is arbitrary and define an integer constant N, such that $N \geq m$.*

Iterative step: *Given x^k, compute x^{k+1} via:*

(i) For every $1 \leq p \leq t$ (possibly in parallel): Execute a finite number, not exceeding N, of CSP steps on the inequalities of the p-th string S_p and denote the resulting point by $\{\overline{x}^p\}_{p=1}^t$.

(ii) Apply

$$x^{k+1} = CA_S(\overline{x}^1, \overline{x}^2, \ldots, \overline{x}^n). \qquad (30)$$

4   The split common fixed point problem for directed operators

In this section we review the *multiple split common fixed point problem* (MSCFPP) which requires to find a common fixed point of a family of operators in one space such that its image under a linear transformation is a common fixed point of another family of operators in the image space.

Problem 20 *Given operators $U_i : R^N \rightarrow R^N$, $i = 1, 2, \ldots, p$, and $T_j : R^M \rightarrow R^M$, $j = 1, 2, \ldots, r$, with fixed points sets $\text{Fix}\, U_i$, $i = 1, 2, \ldots, p$ and $\text{Fix}\, T_j$, $j = 1, 2, \ldots, r$, respectively. The* `multiple split common fixed point problem (MSCFPP)` *is*

find a vector $x^* \in C := \cap_{i=1}^p \text{Fix}\, U_i$ such that $Ax^* \in Q := \cap_{i=1}^r \text{Fix}\, T_j$. (31)

The MSCFPP generalizes the multiple-sets split feasibility problem (MSSFP) (2). It serves as a model for inverse problems where constraints are imposed on the solutions in the domain of a linear operator as well as in the operator's range. MSSFP in its turn generalizes the convex feasibility problem and the two-sets split feasibility problem. Such MSSFPs, formulated in [25], arise in the field of intensity-modulated radiation therapy (IMRT) when one attempts to describe physical dose constraints and equivalent uniform dose (EUD) constraints within a single model, see [21].

The problem with only a single set C in R^N and a single set Q in R^M was introduced by Censor and Elfving [22] and was called the *split feasibility problem* (SFP). They used their simultaneous multiprojections algorithm (see also [33, Subsection 5.9.2]) to obtain iterative algorithms to solve the SFP. Their algorithms, as well as others, see, e.g., Byrne [14], involve matrix inversion at each iterative step. Calculating inverses of matrices is very time-consuming, particularly if the dimensions are large. Therefore, a new algorithm for solving the SFP was devised by Byrne [15], called the CQ-algorithm, with the following iterative step

$$x^{k+1} = P_C \left(x^k + \gamma A^t (P_Q - I) A x^k \right), \qquad (32)$$

where x^k and x^{k+1} are the current and the next iteration vectors, respectively, $\gamma \in (0, 2/L)$ where L is the largest eigenvalue of the matrix $A^t A$ (t stands for matrix transposition), I is the unit matrix or operator and P_C and P_Q denote the orthogonal projections onto C and Q, respectively.

The CQ-algorithm converges to a solution of the SFP, for any starting vector $x^0 \in R^N$, whenever the SFP has a solution. When the SFP has no solutions, the CQ-algorithm converges to a minimizer of $\|P_Q(Ac) - Ac\|$, over all $c \in C$, whenever such a minimizer exists. A block-iterative CQ-algorithm, called the BICQ-method, is also available in [15]. The multiple-sets split feasibility problem, posed and studied in [25], was handled, for both the feasible and the infeasible cases, with a proximity function minimization approach where the proximity function $p(x)$ is

$$p(x) = (1/2) \sum_{i=1}^{t} \alpha_i \|P_{C_i}(x) - x\|^2 + (1/2) \sum_{j=1}^{r} \beta_j \|P_{Q_j}(Ax) - Ax\|^2, \quad (33)$$

The algorithm for solving MSSFP presented there generalizes Byrne's CQ-algorithm [15] and involves orthogonal projections onto $C_i \subseteq R^N$, $i = 1, 2, \ldots, p$,

and $Q_j \subseteq R^M$, $j = 1, 2, \ldots, r$, which were assumed to be easily calculated, and has the following iterative step:

$$x^{k+1} = x^k + \gamma \left(\sum_{i=1}^{p} \alpha_i \left(P_{C_i}(x^k) - x^k \right) + \sum_{j=1}^{r} \beta_j A^t \left(P_{Q_j}(Ax^k) - Ax^k \right) \right), \quad (34)$$

where x^k and x^{k+1} are the current and the next iteration vectors, respectively, $\alpha_i > 0$, $i = 1, 2, \ldots, p$, $\beta_j > 0$, $j = 1, 2, \ldots, r$, $\gamma \in (0, 2/L)$, $L = \sum_{i=1}^{p} \alpha_i + \lambda \sum_{j=1}^{r} \beta_j$ and λ is the spectral radius of the matrix $A^t A$. Masad and Reich [61] is a recent sequel to [25] where they prove weak and strong convergence theorems for an algorithm that solves the multiple-set split convex feasibility problem in Hilbert space.

4.1 A subgradient projection method

In some cases, notably when the convex sets are not linear, computation of the orthogonal projections calls for the solution of a separate minimization problem for each projection. In such cases the efficiency of methods that use orthogonal projections might be seriously reduced. Yang [69] proposed a *relaxed CQ-algorithm* where orthogonal projections onto convex sets are replaced by subgradient projections. The latter are orthogonal projections onto, well-defined and easily derived, half-spaces that contain the convex sets, and are, therefore, easily executed. In [29] the following simultaneous subgradient algorithm for the multiple-sets split feasibility problem was introduced. Assume, without loss of generality, that the sets C_i and Q_j are expressed as

$$C_i = \{x \in R^n \mid c_i(x) \leq 0\} \quad \text{and} \quad Q_j = \{y \in R^m \mid q_j(y) \leq 0\}, \quad (35)$$

where $c_i : R^n \rightarrow R$, and $q_j : R^m \rightarrow R$ are convex functions for all $i = 1, 2, \ldots, p$, and all $j = 1, 2, \ldots, r$, respectively.

Algorithm 21
  Initialization: *Let x^0 be arbitrary.*

Iterative step: *For $k \geq 0$ let*

$$x^{k+1} = x^k + \gamma \left(\sum_{i=1}^{p} \alpha_i \left(P_{C_{i,k}}(x^k) - x^k \right) \right.$$

$$\left. + \sum_{j=1}^{r} \beta_j A^t \left(P_{Q_{j,k}}(Ax^k) - Ax^k \right) \right). \qquad (36)$$

Here $\gamma \in (0, 2/L)$, with $L = \sum_{i=1}^{p} \alpha_i + \lambda \sum_{j=1}^{r} \beta_j$, where λ is the spectral radius of $A^t A$, the constants $\alpha_i > 0$, for $i = 1, 2, \ldots, p$, and $\beta_j > 0$, for $j = 1, 2, \ldots, r$, are arbitrary, and

$$C_{i,k} = \left\{ x \in R^n \mid c_i(x^k) + \left\langle \xi^{i,k}, x - x^k \right\rangle \leq 0 \right\}, \qquad (37)$$

where $\xi^{i,k} \in \partial c_i(x^k)$ is a subgradient of c_i at the point x^k, and

$$Q_{j,k} = \left\{ x \in R^m \mid q_j(x^k) + \left\langle \eta^{j,k}, y - Ax^k \right\rangle \leq 0 \right\}, \qquad (38)$$

where $\eta^{j,k} \in \partial q_j(Ax^k)$.

4.2 A parallel algorithm for the multiple split common fixed point problem

In [31] Censor and Segal employed a product space formulation to derive and analyze a simultaneous algorithm for Problem 20 and obtained the following algorithm.

Algorithm 22
    Initialization: *Let x^0 be arbitrary.*
    Iterative step: *For $k \geq 0$ let*

$$x^{k+1} = x^k + \gamma \left(\sum_{i=1}^{p} \alpha_i \left(U_i(x^k) - x^k \right) + \sum_{j=1}^{r} \beta_j A^t \left(T_j(Ax^k) - Ax^k \right) \right). \qquad (39)$$

Here $\gamma \in (0, 2/L)$, with $L = \sum_{i=1}^{p} \alpha_i + \lambda \sum_{j=1}^{r} \beta_j$, where λ is the largest eigenvalue of the matrix $A^t A$.

The following convergence result was obtained.

Theorem 23 *Let $U_i : R^N \to R^N$, $i = 1, 2, \ldots, p$, and $T_j : R^N \to R^N$, $j = 1, 2, \ldots, r$, be directed operators with fixed points sets C_i, $i = 1, 2, \ldots, p$ and Q_j, $j = 1, 2, \ldots, r$, respectively, and let A be an $M \times N$ real matrix. Assume that $(U_i - I)$, $i = 1, 2, \ldots, p$ and $(T_j - I)$, $j = 1, 2, \ldots, r$, are closed at 0. If $\Gamma \neq \varnothing$ then every sequence, generated by Algorithm 22, converges to $x^* \in \Gamma$.*

Since the orthogonal projection P is a directed operator and $P - I$ is closed at 0, the algorithm from [25] with iterative step (34) is a special case of our Algorithm 22. The Algorithm 21 is also a special case of our Algorithm 22 (see Example 4 and Lemma 5).

4.3 A perturbed projection method

In this subsection we survey another method for the multiple-sets split feasibility problem. This method [29] is based on Santos and Scheimberg [65] who suggested replacing each nonempty closed convex set of the convex feasibility problem by a convergent sequence of supersets. If such supersets can be constructed with reasonable efforts and if projecting onto them is simpler than projecting onto the original convex sets then a perturbed algorithm becomes useful. The following notion of convergence of sequences of sets in R^n is called *Mosco-convergence*. See, e.g., [4, Lemma 4.2], where further useful references are given and the convergence of the corresponding sequence of orthogonal projections onto the sets is discussed. In Salinetti and Wets [64] one can learn about the relation with set convergence with respect to the Hausdorff metric. The notion of Mosco-convergence was also used in [59].

Definition 24 *Let C and $\{C_k\}_{k=0}^{\infty}$ be a subset and a sequence of subsets of R^n, respectively. The sequence $\{C_k\}_{k=0}^{\infty}$ is said to be Mosco-convergent to C, denoted by $C_k \xrightarrow{M} C$, if*
(i) for every $x \in C$, there exists a sequence $\{x^k\}_{k=0}^{\infty}$ with $x^k \in C_k$ for all $k = 0, 1, 2, \ldots$, such that, $\lim_{k \to \infty} x^k = x$, and
(ii) for every subsequence $\{x^{k_j}\}_{j=0}^{\infty}$ with $x^{k_j} \in C_{k_j}$ for all $j = 0, 1, 2, \ldots$, such that $\lim_{j \to \infty} x^{k_j} = x$ one has $x \in C$.

Using the notation $NCCS(R^n)$ for the family of nonempty closed convex subsets of R^n, let Ω_k and Ω be sets in $NCCS(R^n)$, such that, $\Omega_k \xrightarrow{M} \Omega$ as $k \to \infty$. Let C_i and $C_{i,k}$ be sets in $NCCS(R^n)$, for $i = 1, 2, \ldots, t$ and Q_j

and $Q_{j,k}$ be sets in NCCS(R^m), for $j = 1, 2, \ldots, r$, such that, $C_{i,k} \xrightarrow{M} C_i$, and $Q_{j,k} \xrightarrow{M} Q_j$ as $k \to \infty$. Define the operators

$$N(x) := P_\Omega \left\{ x + s \left(\sum_{i=1}^t \alpha_i (P_{C_i}(x) - x) \right. \right.$$
$$\left. \left. + \sum_{j=1}^r \beta_j A^T (P_{Q_j}(Ax) - Ax) \right) \right\}, \tag{40}$$

$$N_k(x) := P_{\Omega_k} \left\{ x + s \left(\sum_{i=1}^t \alpha_i (P_{C_{i,k}}(x) - x) \right. \right.$$
$$\left. \left. + \sum_{j=1}^r \beta_j A^T (P_{Q_{j,k}}(Ax) - Ax) \right) \right\}, \tag{41}$$

and let $\{\varepsilon_k\}_{k=0}^\infty$ be a sequence in $(0, 1)$ satisfying

$$\sum_{k=0}^\infty \varepsilon_k (1 - \varepsilon_k) = +\infty. \tag{42}$$

Then the following algorithm for the CMSSFP generates, under reasonable conditions (see, [29]), convergent iteration sequences.

Algorithm 25 *The perturbed projection algorithm for CMSSFP*
   Initialization: Let $x^0 \in R^n$ be arbitrary.
   Iterative step: For $k \geq 0$, given the current iterate x^k, calculate the next iterate x^{k+1} by

$$x^{k+1} = (1 - \varepsilon_k) x^k + \varepsilon_k N_k(x^k), \tag{43}$$

where N_k and ε_k are as defined above.

References

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

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

[3] H.H. Bauschke, The approximation of fixed points of compositions of nonexpansive mappings in Hilbert space, *Journal of Mathematical Analysis and Applications* **202** (1996), 150–159.

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

[5] H.H. Bauschke, J.M. Borwein and A.S. Lewis, The method of cyclic projections for closed convex sets in Hilbert space, *Contemporary Mathematics* **204** (1997), 1–38.

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

[7] H.H. Bauschke, P.L. Combettes and S.G. Kruk, Extrapolation algorithm for affine-convex feasibility problems, *Numerical Algorithms* **41** (2006), 239–274, .

[8] H.H. Bauschke, E. Matoušková and S. Reich, Projection and proximal point methods: convergence results and counterexamples, *Nonlinear Analysis: Theory, Methods and Applications* **56** (2004), 715–738.

[9] M. Benzi, Gianfranco Cimmino's contributions to numerical mathematics, *Seminario di Analisi Matematica*, Dipartimento di Matematica dell'Università di Bologna, Ciclo di Conferenze in Ricordo di Gianfranco Cimmino, Marzo-Maggio 2004. Tecnoprint, Bologna, Italy (2005), pp. 87–109.

[10] J.T. Bettes and P.D. Frank, A sparse nonlinear optimization algorithm, *Journal of Optimization Theory and Applications* **82** (1994), 519–541.

[11] D. Butnariu, S. Reich and A.J. Zaslavski, Stable convergence theorems for infinite products and powers of nonexpansive mappings, *Numerical Functional Analysis and Optimization* **29** (2008), 304–323.

[12] J.R. Bilbao-Castro, J.M. Carazo, I. García and J.J. Fernández, Parallel iterative reconstruction methods for structure determination of biological specimens by electron microscopy, *Proceedings of The International Conference on Image Processing (ICIP)* (2003), vol. **1**, pp. I565–I568.

[13] F.E. Browder, Convergence theorems for sequences of nonlinear operators in Banach spaces, *Mathematische Zeitschrift* **100** (1967), 201–225.

[14] C.L. Byrne, Bregman-Legendre multidistance projection algorithms for convex feasibility and optimization, *Inherently Parallel Algorithms in Feasibility and Optimization and Their Applications*, D. Butnariu, Y. Censor, and S. Reich, eds., Elsevier Science Publishers, Amsterdam, The Netherlands, 2001, pp. 87–99.

[15] C. Byrne, Iterative oblique projection onto convex sets and the split feasibility problem, *Inverse Problems* **18** (2002), 441–453.

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

[17] D. Butnariu, R. Davidi, G.T. Herman and I.G. Kazansev, Stable convergence behavior under summable perturbations of a class of projection methods for convex feasibility and optimization problems, *IEEE Journal of Selected Topics in Signal Processing* **1** (2007), 540–547.

[18] Y. Censor, Row-action methods for huge and sparse systems and their applications, *SIAM Review* **23** (1981), 444–466.

[19] Y. Censor, Mathematical optimization for the inverse problem of intensity modulated radiation therapy, in: *Intensity-Modulated Radiation Therapy: The State of the Art*, J.R. Palta and T.R. Mackie, eds., American Association of Physicists in Medicine, Medical Physics Monograph No. 29, Medical Physics Publishing, Madison, Wisconsin, USA, 2003, pp. 25–49.

[20] Y. Censor, M.D. Altschuler and W.D. Powlis, On the use of Cimmino's simultaneous projections method for computing a solution of the inverse problem in radiation therapy treatment planning, *Inverse Problems* **4** (1988), 607–623.

[21] Y. Censor, T. Bortfeld, B. Martin and A. Trofimov, A unified approach for inversion problems in intensity-modulated radiation therapy, *Physics in Medicine and Biology* **51** (2006), 2353–2365.

[22] Y. Censor and T. Elfving, A multiprojection algorithm using Bregman projections in a product space, *Numerical Algorithms* **8** (1994), 221–239.

[23] Y. Censor and T. Elfving, Block-iterative algorithms with diagonally scaled oblique projections for the linear feasibility problem, *SIAM Journal on Matrix Analysis and Applications* **24** (2002), 40–58.

[24] Y. Censor, T. Elfving, G.T. Herman, Averaging strings of sequential iterations for convex feasibility problems, in: *Inherently Parallel Algorithms in Feasibility and Optimization and their Applications*, D. Butnariu, Y. Censor and S. Reich (Editors), Elsevier, Amsterdam, 2001, 101–113.

[25] Y. Censor, T. Elfving, N. Kopf and T. Bortfeld, The multiple-sets split feasibility problem and its applications for inverse problems, *Inverse Problems* **21** (2005), 2071–2084.

[26] Y. Censor, D. Gordon and R. Gordon, Component averaging: An efficient iterative parallel algorithm for large and sparse unstructured problems, *Parallel Computing* **27** (2001), 777–808.

[27] Y. Censor, D. Gordon and R. Gordon, BICAV: A block-iterative, parallel algorithm for sparse systems with pixel-related weighting, *IEEE Transactions on Medical Imaging* **20** (2001), 1050–1060.

[28] Y. Censor and A. Lent, Cyclic subgradient projections, *Mathematical Programming* **24** (1982), 233–235.

[29] Y. Censor, A. Motova, A. Segal, Perturbed projections and subgradient projections for the multiple-sets split feasibility problem, *Journal of Mathematical Analysis an Applications* **327** (2007), 1244–1256.

[30] Y. Censor and A. Segal, On the string averaging method for sparse common fixed points problems, *International Transactions in Operational Research*, accepted for publication.

[31] Y. Censor and A. Segal, The Multiple Split Common Fixed Points Problem for Directed Operators, Technical Report, February 2008, submitted for publication.

[32] Y. Censor and E. Tom, Convergence of string-averaging projection schemes for inconsistent convex feasibility problems, *Optimization Methods and Software* **18** (2003), 543–554.

[33] Y. Censor and S.A. Zenios, *Parallel Optimization: Theory, Algorithms, and Applications*, Oxford University Press, New York, NY, USA, 1997.

[34] J.W. Chinneck, The constraint consensus method for finding approximately feasible points in nonlinear programs, *INFORMS Journal on Computing* **16** (2004), 255–265.

[35] G. Cimmino, Calcolo approssimato per le soluzioni dei sistemi di equazioni lineari, La Ricerca Scientifica **XVI** Series II, Anno IX, **1** (1938), 326–333.

[36] P.L. Combettes, The foundations of set-theoretic estimation, *Proceedings of the IEEE* **81** (1993), 182–208.

[37] F.H. Clarke, Generalized gradients and applications, *Transactions of the American Mathematical Society* **205** (1975), 247–262.

[38] F.H. Clarke, On the inverse function theorem, *Pacific Journal of Mathematics* **64** (1976), 97–102.

[39] P.L. Combettes, Inconsistent signal feasibility problems: least-squares solutions in a product space, *IEEE Transactions on Signal Processing* **SP-42** (1994), 2955–2966.

[40] P.L. Combettes, The convex feasibility problem in image recovery, *Advances in Imaging and Electron Physics* **95** (1996), 155–270.

[41] P.L. Combettes, Quasi-Fejérian analysis of some optimization algorithms, in: *Inherently Parallel Algorithms in Feasibility and Optimization and their Applications*, D. Butnariu, Y. Censor and S. Reich (Editors), Elsevier, Amsterdam, 2001, 115–152.

[42] G. Crombez, Finding common fixed points of strict paracontractions by averaging strings of sequential iterations, *Journal of Nonlinear and Convex Analysis* **3** (2002), 345–351.

[43] G. Crombez, Non-monotoneous parallel iteration for solving convex feasibility problems, *Kybernetika* **39** (2003), 547–560.

[44] G. Crombez, Finding common fixed points of a class of paracontractions, *Acta. Math. Hungar.* **103** (2004), 233–241.

[45] G. Crombez, A geometrical look at iterative methods for operators with fixed points, *Numerical Functional Analysis and Optimization* **26** (2005), 157–175.

[46] G. Crombez, A sequential iteration algorithm with non-monotoneous behaviour in the method of projections onto convex sets, *Czechoslovak Mathematical Journal* **56** (2006), 491–506.

[47] F. Deutsch, *Best Approximation in Inner product Spaces*, Springer-Verlag, New York, NY, USA, 2001.

[48] J.M. Dye and S. Reich, Unrestricted iterations of nonexpansive mappings in Hilbert space, *Nonlinear Analysis* **18** (1992), 199–207.

[49] P.P.B. Eggermont, G.T. Herman and A. Lent, Iterative algorithms for large partitioned linear systems, with applications to image reconstruction, *Linear Algebra and Its Applications* **40** (1981), 37–67.

[50] T. Elfving, Block-iterative methods for consistent and inconsistent linear equations, *Numerische Mathematik* **35** (1980), 1–12.

[51] I.I. Eremin, Fejér mappings and convex programming, *Siberian Mathematical Journal* **10** (1969), 762–772.

[52] D. Gordon and R. Gordon, Component-averaged row projections: A robust, block-parallel scheme for sparse linear systems, *SIAM Journal of Scientific Computing* **27** (2005), 1092–1117.

[53] K. Goebel and S. Reich, *Uniform Convexity, Hyperbolic Geometry, and Nonexpansive Mappings*, Marsel Dekker, New York and Basel, 1984.

[54] G.T. Herman, *Image Reconstruction From Projections: The Fundamentals of Computerized Tomography*, Academic Press, New York, NY, USA, 1980.

[55] G.T. Herman and L.B. Meyer, Algebraic reconstruction techniques can be made computationally efficient, *IEEE Transactions on Medical Imaging* **12** (1993), 600–609.

[56] A.N. Iusem and A.R. De Pierro, Convergence results for an accelerated nonlinear Cimmino algorithm, *Numerische Mathematik* **49** (1986), 367–378.

[57] M. Jiang and G. Wang, Convergence studies on iterative algorithms for image reconstruction, *IEEE Transactions on Medical Imaging* **22** (2003), 569–579.

[58] A.V. Lakshminarayanan and A. Lent, Methods of least squares and SIRT in reconstruction, *Journal of Theoretical Biology* **76** (1979), 267–295.

[59] A. Lent and Y. Censor, The primal-dual algorithm as a constraint-set-manipulation device, *Mathematical Programming* **50** (1991), 343–357.

[60] L.D. Marks, W. Sinkler and E. Landree, A feasible set approach to the crystallographic phase problem, *Acta Crystallographica* **A55** (1999), 601–612.

[61] E. Masad and S. Reich, A note on the multiple-set split convex feasibility problem in Hilbert space, *Journal of Nonlinear Convex Analysis* **8** (2007), 367–371.

[62] M.A. Noor, Some developments in general variational inequalities, *Applied Mathematics and Computation* **152** (2004), 197–277.

[63] H. Rhee, An application of the string averaging method to one-sided best simultaneous approximation, *J. Korea Soc. Math. Educ. Ser. B: Pure Appl. Math.* **10** (2003), 49–56.

[64] G. Salinetti and R.J.-B. Wets, The convergence of sequences of convex sets in finite dimensions, *SIAM Review* **21** (1979), 18–33. Addendum: *SIAM Review* **22** (1980) 86.

[65] P.S.M. Santos and S. Scheimberg, A projection algorithm for general variational inequalities with perturbed constraint sets, *Applied Mathematics and Computation* **181** (2006), 649–661.

[66] H. Stark and Y. Yang, *Vector Space Projections: A Numerical Approach to Signal and Image Processing, Neural Nets, and Optics,* John Wiley & Sons, New York, NY, USA, 1998.

[67] J.H. Wilkinson, Introduction to Part II, The Algebraic Eigenvalue Problem, *Handbook for Automatic Computation*, Vol II, Linear Algebra, J.H. Wilkinson and C. Reinsch (Editors), Springer-Verlag, 1971.

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

[69] Q. Yang, The relaxed CQ algorithm solving the split feasibility problem, *Inverse Problems* **20** (2004), 1261–1266.