ALGORITHMS AND CONVERGENCE RESULTS OF PROJECTION METHODS FOR INCONSISTENT FEASIBILITY PROBLEMS: A REVIEW

YAIR CENSOR AND MAROUN ZAKNOON

ABSTRACT. The convex feasibility problem (CFP) is to find a feasible point in the intersection of finitely many convex and closed sets. If the intersection is empty then the CFP is inconsistent and a feasible point does not exist. However, algorithmic research of inconsistent CFPs exists and is mainly focused on two directions. One is oriented toward defining other solution concepts that will apply, such as proximity function minimization wherein a proximity function measures in some way the total violation of all constraints. The second direction investigates the behavior of algorithms that are designed to solve a consistent CFP when applied to inconsistent problems. This direction is fueled by situations wherein one lacks a priori information about the consistency or inconsistency of the CFP or does not wish to invest computational resources to get hold of such knowledge prior to running his algorithm. In this paper we bring under one roof and telegraphically review some recent works on inconsistent CFPs.

1. INTRODUCTION

Inconsistent feasibility problems. Feasibility problems require to find a point in a given set C, any point, not a particular point such as, for example, one that optimizes some given function over C, which would constitute a problem of constrained optimization. Often times the set C is given as an intersection $C := \cap_{i=1}^{m} C_i$ of a finite family of sets $\{C_i\}_{i=1}^{m}$. The convex feasibility problem (CFP) is to find a feasible point $x^* \in C = \cap_{i=1}^{m} C_i$ when all sets C_i are convex and commonly also assumed to be closed. This prototypical problem underlies the modeling of real-world problems in the set theoretic estimation approach of Combettes [36] such as convex set theoretic image recovery [38] and many other fields, see, e.g., the pointers and references in Bauschke and Borwein [8, Section 1] and in Cegielski's book [19, Section 1.3]. In this approach, constraints of the real-world problem are

Date: February 21, 2018. Revised: March 19, 2018. Revised: April 25, 2018.

2010 *Mathematics Subject Classification.* 65K10, 65K15, 90C25.

Key words and phrases. Projection methods, feasibility problems, inconsistent, cyclic convergence, proximity function, string-averaging projections, product space, hard constraints, De Pierro's conjecture, Bregman projections, steered sequential projections, alternating Bregman proximity operators, variational characterization, split feasibility problems, best approximation pairs, set enlargements, ART3, ARM, Intrepid, Valiant.

represented by the demand that a solution should belong to sets C_i, called constraint sets.

If $C \neq \emptyset$ does not hold then the CFP is inconsistent and a feasible point does not exist. However, algorithmic research of inconsistent CFPs exists and is mainly focused on two directions. One is oriented toward defining solution concepts other than $x^* \in C = \cap_{i=1}^{m} C_i$ that will apply, such as proximity function minimization wherein a proximity function measures in some way the total violation of all constraints. The second direction investigates the behavior of algorithms that are designed to solve a consistent CFP when applied to inconsistent problems. The latter direction is fueled by situations wherein one lacks a priori information about the consistency or inconsistency of the CFP or does not wish to invest computational resources to get hold of such knowledge prior to running his algorithm. The next paragraphs on projection methods are quoted from the introduction of Censor and Cegielski [24].

Projection methods. 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 as we mean it here. Here *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 (or approximate projections) onto the given individual sets are easier to perform than projections onto other sets (intersections, image sets under some transformation, etc.) that are derived from the given family of 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 (or approximate 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 in some significant fields of applications.

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, methods start to stutter or 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 structures are either sequential or simultaneous or in-between, such as in the block-iterative projection (BIP) methods or in the more recent string-averaging projection (SAP) methods. An advantage of projection methods is that they work with initial data and do not require transformation of, or other operations on, the sets describing the problem.

Purpose of the paper. We present an effort to bring under one roof and telegraphically review some recent works on inconsistent CFPs. This should be helpful to researchers, veterans or newcomers, by directing them to some of the existing resources. The vast amount of research papers in the field of projection methods makes it sometimes difficult to master even within a specific sub-area. On the other hand, projection methods send branches both into fields of applications wherein real-world problems are solved and into theoretical areas in mathematics such as, but not only, fixed point theory and variational inequalities. Researchers in each of these, seemingly perpendicular, directions might benefit from this review.

A word about notations. We entertained the thought to unify all notations but quickly understood that the game is not worth the candle[1]. With notations left as they appear in the original publications it will make it easier for a reader when choosing to consult the original papers.

An apology. Oversight and lack of knowledge are human traits which we are not innocent of. Therefore, we apologize for omissions and other negligence and lacunas in this paper. We kindly ask our readers to communicate to us any additional items and informations that fit the structure and spirit of the paper and we will gladly consider those for inclusion in future revisions, extensions and updates of the paper that we will post on arXiv.

Organization of the paper. Section 2 contains our review divided into 17 subsections. Each subsection is focused on, and is centered around, one or two historical or recent works. We use these "lead" references to organize the subsections chronologically from older to recent works. An author index at the end of the paper will help locate the results reviewed here.

2. Algorithms and Convergence results of Projection Methods for Inconsistent Feasibility Problems

2.1. **1959: Composition of projections onto two disjoint convex sets.** Cheney and Goldstein [35] showed that if K_1 and K_2 are two closed and convex subsets of a Hilbert space, and P_i are the corresponding orthogonal projections onto K_i, where $i = 1, 2$, then every fixed point of the composition $Q := P_1P_2$ is a point of K_1 closest to K_2. Moreover, they showed that if one of the sets is compact or if one of the sets is finite-dimensional and the distance is attained then a fixed point of Q will be obtained by iterations of Q. In particular, if both sets are polytopes in a finite-dimensional Euclidean space, the distance between two sets is attained, and consequently a fixed point of Q will be obtained by iterations of Q. Their results are in the following three theorems.

[1] Meaning that what we would get from this undertaking is not worth the effort we would have to put into it.

Theorem 2.1. [35, Theorem 2] *Let K_1 and K_2 be two closed convex sets in Hilbert space. Let P_i denote the proximity map for K_i. Any fixed point of P_1P_2 is a point of K_1 nearest K_2, and conversely.*

Theorem 2.2. [35, Theorem 4] *Let K_1 and K_2 be two closed convex sets in Hilbert space and Q the composition P_1P_2 of their proximity maps. Convergence of $Q^n x$ to a fixed point of Q is assured when either (a) one set is compact, or (b) one set is finite-dimensional and the distance between the sets is attained.*

Theorem 2.3. [35, Theorem 5] *In a finite-dimensional Euclidean space, the distance between two polytopes is attained, a polytope being the intersection of a finite family of half-spaces.*

In this connection, see also Theorems 4.1 and 4.2 in the paper by Kopecká and Reich [53].

2.2. **1967: Cyclic convergence of sequential projections onto m sets with empty intersection.** Gubin, Polyak and Raik [49] consider m closed convex subsets, $C_1, C_2, \cdots, C_m$, of a normed space E, and studied the behavior of the sequence generated according the rule

$$x^{n+1} = P_{i(n)}x^n, \text{ where } i(n) := n \pmod m + 1 \tag{2.1}$$

with x^0 arbitrary. They showed that if one of the sets is bounded, the subsequences

$$\left\{x^{mn+1}\right\}_{n\in\mathbb{N}}, \left\{x^{mn+2}\right\}_{n\in\mathbb{N}}, \cdots, \left\{x^{mn+m}\right\}_{n\in\mathbb{N}} \tag{2.2}$$

converge weakly to cluster points $\overline{x}_1, \overline{x}_2, \cdots, \overline{x}_m$, respectively, that constitute a cycle, i.e.,

$$\overline{x}_2 = P_2\overline{x}_1,\ \overline{x}_3 = P_3\overline{x}_2,\ \cdots, \overline{x}_m = P_m\overline{x}_{m-1},\ \overline{x}_1 = P_1\overline{x}_m. \tag{2.3}$$

They proved the following (slightly paraphrased here) version of [49, Theorem 2].

Theorem 2.4. *Let all C_i, $i = 1, 2, \cdots, m$, be closed, convex and nonempty subsets of E and at least one of them (for explicitness, C_1) be bounded. Then it is possible to find points $\overline{x}_i \in C_i$, $i = 1, 2, \cdots, m$, such that $P_{i+1}(\overline{x}_i) = \overline{x}_{i+1}$, $i = 1, 2, \cdots, m-1$, $P_1(\overline{x}_m) = \overline{x}_1$, while in the method (2.1) we have $x^{km+i+1} - x^{km+i} \to \overline{x}_{i+1} - \overline{x}_i$, and x^{km+i} weakly converges to $\overline{x}_i$ as $k \to \infty$. If, in addition, any of the following conditions is satisfied,*
(a) all C_i with the possible exception of one ($C_{\overline{i}}$), are uniformly convex with the common function $\delta(\tau)$;
(b) E is finite-dimensional;
(c) all C_i are (closed) half-spaces;
then the convergence will be strong. If all C_i, apart from possibly one, are also strongly convex, and $\cap_{i=1}^m C_i$ is empty, the sequence x^{km+i} converges to $\overline{x}_i$ at the rate of a geometrical progression.

2.3. **1983: The limits of the cyclic subsequences approach a single point as relaxation goes to zero.** Censor, Eggermont and Gordon [26] investigate the behavior of Kaczmarz's method with relaxation for inconsistent systems. They show that when the relaxation parameter goes to zero, the limits of the cyclic subsequences (See Theorem 2.4 here) generated by the method approach a weighted least squares solution of the system. This point minimizes the sum of the squares of the Euclidean distances to the hyperplanes of the system. If the starting point is chosen properly, then the limits approach the minimum norm weighted least squares solution. The proof is given for a block-Kaczmarz method.

Consider the linear system of equations $Ax = b$, where $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, a_i is the i'th row of the matrix A, and b_i is the i'th component the column vector b. Kaczmarz's algorithm [52] employs the iterative process

$$x^{k+1} = x^k + \frac{b_i - \left\langle a_i, x^k \right\rangle}{\left\| a_i \right\|^2} a_i, \tag{2.4}$$

where $x^0 \in \mathbb{R}^n$ is arbitrary, for solving the system $Ax = b$. Eggermont, Herman, and Lent [46] rewrote the above m and n as $m = LM$, $n = N$,with any natural numbers L, M, N, partitioned A and b as

$$A = \begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_M \end{pmatrix}, \quad b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_M \end{pmatrix}, \tag{2.5}$$

with $A_i \in \mathbb{R}^{L \times N}$ and $b_i \in \mathbb{R}^L$, and proposed the following block-Kaczmarz method for solving a linear system of the above form.

$$x^0 \in \mathbb{R}^N \text{ is arbitrary, } x^{k+1} = x^k + \lambda A_i^T \left(b_i - A_i x^k \right), \ i = k \pmod{M} + 1, \tag{2.6}$$

with relaxation $\lambda \in (0, 2)$.

The effect of strong underrelaxation on the limits of the cyclic subsequences generated by the block-Kaczmarz algorithm (2.6), investigated in [26] is included in the following theorem.

Theorem 2.5. [26, Theorem 1] *For all λ small enough,*

$$x^* (\lambda) = \lim_{k \to \infty} x^{kM} \tag{2.7}$$

exists, and

$$\lim_{\lambda \to 0} x^* (\lambda) = A^\dagger b + \left(Id - A^\dagger A \right) x^0. \tag{2.8}$$

(Where $A^\dagger$ is the Moore-Penrsose inverse of the matrix A.)

This applies to every subsequence $\left\{ x^{kM+\ell} \right\}_{k \geq 0}$, $\ell \in \{0, 1, 2, \cdots, M-1\}$. In Censor, Eggermont and Gordon [26, Page 91], it is shown that algorithm (2.4) is a special case of algorithm (2.6), with $L = 1$, and that the equality

(2.8) means that $\lim_{\lambda \to 0} x^*(\lambda)$ is a least squares solution of the system $Ax = b$. A relevant remark concerning the behavior of the Cimmino method in the inconsistent case appears in the Remark on pages 286–287 of the 1983 paper by Reich [59].

2.4. **1993 and 1994: Alternating projection algorithms for two sets.** Bauschke and Borwein [6] investigated the convergence of the von Neumann's alternating projection method for two arbitrary closed convex nonempty subsets A, B of a Hilbert space H. Finding a point in $A \cap B$, or if $A \cap B$ is empty a good substitute for it, is a basic problem in various areas of mathematics.

Defining the distance between two nonempty subsets M, N by $d(M, N) := \inf\{\|m - n\| \mid m \in M,\ n \in N\}$, denoting $E := \{a \in A \mid d(a, B) = d(A, B)\}$ and $F := \{b \in B \mid d(b, A) = d(B, A)\}$, one notes that if $A \cap B \neq \varnothing$ then $E = F = A \cap B$. The projection of any point x onto a closed convex nonempty subset C is denoted by $P_C x$. The von Neumann algorithm for finding a point in $A \cap B$ is as follows: Given a starting point $x \in X \subseteq H$, define, for every integer $n \geq 1$, the terms of the sequences (a_n), (b_n) by

$$b_0 := x, \quad a_n := P_A b_{n-1}, \quad b_n := P_B a_n. \tag{2.9}$$

von Neumann proved that both sequences converge to $P_{A \cap B}(x)$ in norm when A, B are closed subspaces.

Assuming [6, Page 201] that A, B are closed affine subspaces, say $A = a + K$, $B = b + L$ for vectors $a, b \in X$ and closed subspaces K, L. The angle between K and L is denoted by $\gamma(K, L)$ Bauschke and Borwein proved the following.

Theorem 2.6. [6, Theorem 4.11] *If $K + L$ is closed, then the von Neumann sequences converge linearly with rate* $\cos \gamma(K, L)$ *independent of the starting point. In particular, this happens whenever one of the following conditions holds: (i) K or L has finite dimension, (ii) K or L has finite codimension.*

In [7] Bauschke and Borwein analyzed Dykstra's algorithm for two arbitrary closed convex sets in a Hilbert space X. They greatly expanded on the Cheney and Goldstein papers (infinite dimensions, characterizations, etc.) See Subsections 2.1 here and the recent work of Kopecká and Reich [54] in Subsection 2.14 here.

2.5. **1994: Least-squares solutions of inconsistent signal feasibility problems in a product space.** Combettes's [37] presents parallel projection methods to find least-squares solutions to inconsistent convex set theoretic signal synthesis problems. The problem of finding a signal that minimizes a weighted average of the squares of the distances to constraint sets is reformulated in a product space, where it is equivalent to that of finding a point that lies in a particular subspace and at minimum distance from the Cartesian product of the original sets. A solution is obtained in the product space via methods of alternating projections which naturally lead to methods of parallel projections in the original space. The convergence properties of the proposed methods are analyzed and signal synthesis applications are demonstrated.

The, possibly inconsistent, feasibility problem: Find $a^* \in \cap_{i=1}^m S_i$, where the S_is are closed and convex subsets of a Hilbert space Ξ, is replaced by the unconstrained weighted least-squares minimization problem

$$\min \ \{ \ \Phi(a) \mid a \in \Xi\}, \tag{2.10}$$

where $\Phi(a) := \frac{1}{2}\sum_{i=1}^m w_i d(a, S_i)^2$, $d(a, S_i) := \inf\{d(a,b) \mid b \in S_i\}$, and $(w_i)_{1\leq i\leq m}$ are strictly convex weights, i.e., $\sum_{i=1}^m w_i = 1$ and $\forall i \in \{1, \cdots, m\}$ $\omega_i > 0$. In other words, the goal is to solve

$$\text{Find } a^* \in G := \{a \in \Xi \mid \ \Phi(a) \leq \Phi(b) \ \text{ for all } b \in \Xi\}. \tag{2.11}$$

In the Cartesian product space Ξ^m, with the scalar product $\langle\langle \mathbf{a}, \mathbf{b}\rangle\rangle := \sum_{i=1}^m w_i \left\langle a^{(i)}, b^{(i)} \right\rangle$ for all $\mathbf{a} := \left(a^{(1)}, a^{(2)}, \cdots, a^{(m)}\right) \in \Xi^m$ and $\mathbf{b} :=(b^{(1)}, b^{(2)}, \cdots, b^{(m)}) \in \Xi^m$ the problem (2.11) is reformulated as

$$\text{Find } \mathbf{a}^* \in \mathbf{G} := \{\mathbf{a} \in \mathbf{D} \mid d(\mathbf{a}, \mathbf{S}) = d(\mathbf{D}, \mathbf{S})\} = \{\mathbf{a} \in \mathbf{D} \mid P_{\mathbf{D}}(P_{\mathbf{S}}(\mathbf{a})) = \mathbf{a}\}, \tag{2.12}$$

where $\mathbf{D} = \{(a, a, \cdots, a) \in \Xi^m \mid a \in \Xi\}$, $\mathbf{S} = S_1 \times S_2 \cdots \times S_m$, and $P_{\mathbf{D}}$, $P_{\mathbf{S}}$ are the orthogonal projections onto the sets $\mathbf{D}$ and $\mathbf{S}$, respectively. All quantities related to the product space are written in boldface symbols. Solving this problem using two methods for finding a fixed point of the composition $P_{\mathbf{D}} \circ P_{\mathbf{S}}$ and translating back the results to the original space Ξ, the following two convergence results are obtained.

Theorem 2.7. [37, Theorem 4] *Suppose that one of the S_is is bounded. Then, for any $a_0 \in \Xi$, every sequence of iterates $(a_n)_{n\geq 0}$ defined by*

$$a_{n+1} = a_n + \lambda_n \left(\sum_{i=1}^m w_i P_i(a_n) - a_n\right), \tag{2.13}$$

where $(\lambda_n)_{n\geq 0} \subseteq [\varepsilon, 2-\varepsilon]$ with $0 < \varepsilon < 1$, converges weakly to a point in G.

Assuming that $(\alpha_n)_{n\geq 0}$ fulfills

$$\lim_{n\to+\infty} \alpha_n = 1, \ \sum_{n\geq 0} (1-\alpha_n) = +\infty \text{ and } \lim_{n\to+\infty} (\alpha_{n+1} - \alpha_n)(1-\alpha_{n+1})^{-2} = 0 \tag{2.14}$$

the next theorem holds.

Theorem 2.8. [37, Theorem 5] *Suppose that one of the S_is is bounded. Then, for any $a_0 \in \Xi$, every sequence of iterates $(a_n)_{n\geq 0}$ defined by*

$$a_{n+1} = (1-\alpha_n) a_0 + \alpha_n \left(\lambda \sum_{i=1}^m w_i P_i(a_n) - (1-\lambda) a_n\right), \tag{2.15}$$

where $(\alpha_n)_{n\geq 0}$ is as in (2.14) and $0 < \lambda \leq 2$, converges strongly to the projection of a_0 onto G.

2.6. **1995 and 2003: The method of cyclic projections for closed convex sets in Hilbert space.** Bauschke, Borwein and Lewis [10] consider closed convex nonempty sets $C_1, C_2, \cdots, C_N$ in a real Hilbert space H, with corresponding projections $P_1, P_2, \cdots, P_N$ and systematically study composition of projections in the inconsistent case. For an arbitrary starting point $x^0 \in H$ the method of cyclic projections generates N sequences $(x_i^n)_n$ by

$$\begin{array}{l} x_1^1 := P_1 x^0,\ x_2^1 := P_2 x_1^1,\ \cdots,\ x_N^1 := P_N x_{N-1}^1, \\ x_1^2 := P_1 x_N^1,\ x_2^2 := P_2 x_1^2,\ \cdots,\ x_N^2 := P_N x_{N-1}^2, \\ x_1^3 := P_1 x_N^2, \qquad \cdots\cdots . \end{array} \tag{2.16}$$

They collected these sequences cyclically in one sequence $(x^0, x_1^1, x_2^1, \cdots, x_N^1, x_1^2, x_2^2, \cdots)$ to which they referred as the orbit generated by x^0 or the orbit with starting point x^0. They further defined the composite projections operators

$$Q_1 := P_1 P_N P_{N-1} \cdots P_2,\ Q_2 := P_2 P_1 P_N \cdots P_3,\ \cdots,\ Q_N := P_N P_{N-1} \cdots P_1 \tag{2.17}$$

which allows to write more concisely

$$x_i^n := Q_i^{n-1} x_i^1, \text{ for all } n \geq 1 \text{ and for every } i; \tag{2.18}$$

after setting $P_0 := P_N$, $P_{N+1} := P_1$, $x_0^n := x_N^{n-1}$, and $x_{N+1}^n := x_1^{n+1}$, they reached

$$x_{i+1}^n = P_{i+1} x_i^n, \text{ for all } n \geq 1 \text{ and every } i. \tag{2.19}$$

When appropriate, they similarly identify $i = 0$ with $i = N$ and $i = N + 1$ with $i = 1$.

They gave a dichotomy result on orbits which roughly says that if each Q_i is fixed point free then, the orbit has no bounded subsequence; otherwise, each subsequence (x_i^n) converges weakly to some fixed point of Q_i. Two central questions were posed:

1. When does each Q_i have a fixed point?
2. If each Q_i has a fixed point, when do the subsequences (x_i^n) converge in norm (or even linearly)?

Concerning Question 1, They provide sufficient conditions for the existence of fixed points or approximate fixed points (that is $\inf_{x \in H} \|x - Q_i x\| = 0$, for each i). It follows that while fixed points of Q_i need not exist for nonintersecting closed affine subspaces, approximate fixed points must.

In respect to Question 2, a variety of conditions guaranteeing norm convergence (in the presence of fixed points for each Q_i) is offered: one of the sets C_i has to be (boundedly) compact or all sets are or convex polyhedra, or affine subspaces. In the affine subspace case each sequence (x_i^n) converges to the fixed point of Q_i nearest to x^0. Moreover the convergence is linear, whenever the angle of the N-tuple of the associated closed subspaces is positive.

In subsequent work, Bauschke [5] showed that the composition of finitely many projections $P_N, P_{N-1}, \cdots, P_1$ is asymptotically regular, i.e.,

$$(P_N, P_{N-1}, \cdots, P_1)^k x - (P_N, P_{N-1}, \cdots, P_1)^{k+1} x \to 0, \text{ for every } x \in X. \tag{2.20}$$

thus proving the so-called "zero displacement conjecture" of Bauschke, Borwein and Lewis [10].

2.7. **1999: Hard-constrained inconsistent signal feasibility problems.** Combettes and Bondon [39] consider the problem of synthesizing feasible signals in a Hilbert space $\mathcal{H}$, with inconsistent convex constraints that are divided into two parts, the hard constraints and the soft constraints. They look for a point in $\mathcal{H}$ which satisfies the hard constraints imperatively and minimizes the violation of the soft constraints.

Denote by Γ the class of all lower semicontinuous proper convex functions from $\mathcal{H}$ into $]-\infty, +\infty]$. Given $g \in \Gamma$ and $\alpha \in \mathbb{R}$, the closed and convex set $lev_{\leq\alpha} g := \{x \in \mathcal{H} \mid g(x) \leq \alpha\}$ is the lower level set of g at height α, and the nonempty convex set $dom g := \{x \in \mathcal{H} \mid g(x) < +\infty\}$ is its domain. The goal of a convex set theoretic signal synthesis (design or estimation) problem in $\mathcal{H}$ is to produce a signal x^* that satisfies convex constraints, say,

$$\text{find } x^* \in S = \cap_{i\in I} S_i, \text{ where } (\forall i \in I) \ S_i = lev_{\leq 0} g_i, \tag{2.21}$$

where I is a finite index set, and $(g_i)_{i\in I} \subset \Gamma$.

Let $I^\blacktriangle \subset I$ denote the, possibly empty, hard constraints index set, $I^\triangle = I \setminus I^\blacktriangle$ the nonempty soft constraints index set, $S^\blacktriangle = \cap_{i\in I^\blacktriangle} S_i$ the hard feasibility set and, by convention, $S^\blacktriangle = \mathcal{H}$ if $I^\blacktriangle = \varnothing$. $S^\triangle = \cap_{i\in I^\triangle} S_i$, $D^\triangle = \cap_{i\in I^\triangle} dom g_i$, and assume that $S^\blacktriangle \cap D^\triangle \neq \varnothing$. $\mathcal{F}$ is the class of all increasing convex functions from $[0, +\infty[$ into $[0, +\infty[$ that vanish (only) at 0; every $f \in \mathcal{F}$ is extended to the argument $+\infty$ by setting $f(+\infty) = +\infty$. For every $g \in \Gamma$, $g^+ = \max\{0, g\}$. The amount of violation of the soft constraints $(g_i(x) \leq 0)_{i\in I^\triangle}$ is measured by an objective function $\Phi^\triangle : \mathcal{H} \to [0, +\infty]$ of the general form

$$\Phi^\triangle := \sum_{i\in I^\triangle} f_i \circ g_i^+, \text{ where } (f_i)_{i\in I^\triangle} \subset \mathcal{F}. \tag{2.22}$$

The hard-constrained signal feasibility problem is to minimize the objective $\Phi^\triangle$ of (2.22) over the hard feasibility set $S^\blacktriangle$. Setting $\alpha^* = \inf_{x\in S^\blacktriangle} \Phi^\triangle(x)$, the problem reads

$$\text{find } x^* \in G := \left\{x \in S^\blacktriangle \mid \Phi^\triangle(x) = \alpha^*\right\}. \tag{2.23}$$

Combettes and Bondon [39] supply convergence theorems for the following processes under various conditions ($P^\blacktriangle$ denotes the projector onto $S^\blacktriangle$):

$$(\forall n \in \mathbb{N}) \ x_{n+1} = P^\blacktriangle\left(x_n - \gamma\nabla\Phi^\triangle(x_n)\right); \tag{2.24}$$

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = (1-\lambda_n) x_n + \lambda_n P^\blacktriangle\left(x_n - \gamma\nabla\Phi^\triangle(x_n)\right); \tag{2.25}$$

and

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = (1-\lambda_n)\, r + \lambda_n P^{\blacktriangle}\left(x_n - \gamma \nabla \Phi^{\triangle}(x_n)\right), \tag{2.26}$$

for a fixed given $r \in \mathcal{H}$.

Defining the proximity function $\Phi^{\triangle} := \frac{1}{2}\sum_{i\in I^{\triangle}} w_i d(\cdot, S_i)^2$, with weights $(w_i)_{i\in I^{\triangle}} \subset]0,1]$, $\sum_{i\in I^{\triangle}} w_i = 1$, and denoting by P_i the projector onto S_i, the authors prove that, under certain conditions, the iterative process

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = (1-\lambda_n)\, x_n + \lambda_n P^{\blacktriangle}\left((1-\gamma)\, x_n + \gamma \sum_{i\in I^{\triangle}} w_i P_i(x_n)\right) \tag{2.27}$$

generates sequences that converge, weakly or strongly, to a solution of the hard-constrained signal feasibility problem, i.e., to a point in G of (2.23) above. Also, they show that the iterative process

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = \frac{n}{n+1} P^{\blacktriangle}\left((1-\gamma)\, x_n + \gamma \sum_{i\in I^{\triangle}} w_i P_i(x_n)\right) \tag{2.28}$$

generates sequences that converge strongly to $P_G(0)$. In all these iterative processes there appears $P^{\blacktriangle}$, the projector onto $S^{\blacktriangle}$, which can potentially hinder practical applications if this projection is not simple to calculate.

2.8. **2001: De Pierro's conjecture.** Bauschke and Edwards [13] describe De Pierro's conjecture as follows. Suppose we are given finitely many nonempty closed convex sets in a real Hilbert space and their associated projections. For suitable arrangements of the sets, it is known that the sequence obtained by iterating the composition of the underrelaxed projections is weakly convergent. The question arises how these weak limits vary as the underrelaxation parameter tends to zero. In 2001, De Pierro conjectured [43] that the weak limits approach the least squares solution nearest to the starting point of the sequence. In fact, the result by Censor, Eggermont, and Gordon [26] described here in Subsection 2.3, implies De Pierro's conjecture for affine subspaces in Euclidean space.

De Pierro's conjecture [43, Section 3, Conjecture II] is succinctly formulated in Bauschke and Edwards [13, Conjecture 1.6] as follows. For a convex feasibility problem with N sets $(C_i)_{i=1}^N$ in a Hilbert space X, define, fort every $\lambda \in]0,1]$, the composition of underrelaxed projections

$$Q_\lambda := ((1-\lambda)Id + \lambda P_{C_N}) \cdots ((1-\lambda)Id + \lambda P_{C_2})((1-\lambda)Id + \lambda P_{C_1}), \tag{2.29}$$

where Id is the identity and P_{C_i} is the projection onto C_i. The corresponding sets of fixed points are defined by

$$F_\lambda := \operatorname{Fix} Q_\lambda := \{x \in X \mid x = Q_\lambda(x)\} \tag{2.30}$$

and the aim is to understand the behavior of the sequence $(Q_\lambda^n(x))_{n\in\mathbb{N}}$ in terms of $\lambda \in]0,1]$, for an arbitrary $x \in X$.

Conjecture 2.9. (De Pierro) [13, Conjecture 1.6]. Suppose that $F_\lambda \neq \emptyset$ for every $\lambda \in]0,1]$. Denoting, for all $x \in X$ and all $\lambda \in]0,1]$ the limits $x_\lambda = \text{weak}\lim_{n\to+\infty} Q_\lambda^n(x)$, De Pierro conjectured that $\lim_{\lambda\to 0^+} x_\lambda = P_{\mathcal{L}}(x)$, where $\mathcal{L}$ is the set of least squares solutions of the convex feasibility problem, i.e.,

$$\mathcal{L} := \left\{ x \in X \mid \sum_{i=1}^{N} \|x - P_{C_i}(x)\|^2 = \inf_{y\in X} \sum_{i=1}^{N} \|y - P_{C_i}(y)\|^2 \right\}. \tag{2.31}$$

Bauschke and Edwards proved this conjecture for families of closed affine subspaces satisfying a metric regularity condition. Baillon, Combettes and Cominetti [4, Theorem 3.3] proved the conjecture under a mild geometrical condition. Recently, Cominetti, Roshchina and Williamson [40, Theorem 1] proved that this conjecture is false in general by constructing a system of three compact convex sets in $\mathbb{R}^3$ for which the least squares solution exists but the conjecture fails to hold.

2.9. **2001: Proximity function minimization using multiple Bregman projections.** Motivated by the geometric alternating minimization approach of Csiszár and Tusnády [41] and the product space formulation of Pierra [57], Byrne and Censor [18] derive a new simultaneous multiprojection algorithm that employs generalized projections of Bregman [17], see also, e.g., Censor and Lent [30] and Bauschke and Borwein [9], to solve the convex feasibility problem (CFP) or, in the inconsistent case, to minimize a proximity function that measures the average distance from a point to all convex sets. For background material on Bregman functions and Bregman distances and projections see, e.g., the book of Censor and Zenios [34], [9], Solodov and Svaiter [61], Eckstein [44], [45], to name but a few. Byrne and Censor [18] assume that the Bregman distances involved are jointly convex, so that the proximity function itself is convex. When the intersection of the convex sets is empty, but the closure $\mathrm{cl}F(x)$ of the proximity function $F(x)$, defined by $\mathrm{cl}F(x) := \liminf_{y\to x} F(y)$, has a unique global minimizer, the sequence of iterates converges to this unique minimizer. Special cases of this algorithm include the "Expectation Maximization Maximum Likelihood" (EMML) method in emission tomography and a new convergence result for an algorithm that solves the split feasibility problem.

Let C_i, $i = 1, 2, \ldots, I$, be closed convex sets in the J-dimensional Euclidean space $\mathbb{R}^J$ and let C be their intersection. Let S be an open convex subset of $\mathbb{R}^J$ and f a Bregman function from the closure $\overline{S}$ of S into $\mathbb{R}$; see, e.g., [34, Chapter 2]. For a Bregman function $f(x)$, the Bregman distance D_f is defined by

$$D_f(z,x) := f(z) - f(x) - \langle \nabla f(x), z - x\rangle, \tag{2.32}$$

where $\nabla f(x)$ is the gradient of f at x. If the function f has the form $f(x) = \sum_{j=1}^{J} g_j(x_j)$, with the g_j scalar Bregman functions, then f and the associated $D_f(z,x)$ are called *separable*. With $g_j(t) = t^2$, for all j, the function $f(x) = \sum_{j=1}^{J} g_j(x_j) = \sum_{j=1}^{J} x_j^2$ is a separable Bregman function and $D_f(z,x)$ is the squared Euclidean distance between z and x. For each i, denote by $P_{C_i}^f(x)$ the *Bregman projection* of $x \in S$ onto C_i with respect to the *Bregman function* f, i.e., $D_f(P_{C_i}^f(x), x) \leq D_f(z,x)$, for all $z \in C_i \cap \overline{S}$. In [18, Eq. (1.2)] the proximity function $F(x)$ is of the form

$$F(x) = \sum_{i=1}^{I} D_{f_i}(P_{C_i}^{f_i}(x), x), \tag{2.33}$$

where the D_{f_i} are Bregman distances derived from possibly distinct, possibly nonseparable Bregman functions f_i with zones S_{f_i}. The function F is defined, for all x in the open convex set $U := \cap_{i=1}^{I} S_{f_i}$, which is assumed nonempty. The proximity function $F(x)$ of (2.33) is extended to all of $\mathbb{R}^J$ by defining $F(x) = +\infty$, for all $x \notin U$ and its closure $\mathrm{cl}F$ is as defined above. They proved convergence of their iterative method whenever $\mathrm{cl}F$ has a unique minimizer or when the set $C \cap \overline{U}$ is nonempty. The following algorithm is proposed.

Algorithm 2.10. [18, Algorithm 4.1].
Initialization: $x^0 \in U$ is arbitrary.
Iterative Step: Given x^k find, for all $i = 1, 2, \ldots, I$, the projections $P_{C_i}^{f_i}(x^k)$ and calculate x^{k+1} from

$$\sum_{i=1}^{I} \nabla^2 f_i(x^{k+1}) x^{k+1} = \sum_{i=1}^{I} \nabla^2 f_i(x^{k+1}) P_{C_i}^{f_i}(x^k), \tag{2.34}$$

where $\nabla^2 f_i(x^{k+1})$ denotes the Hessian matrix (of second partial derivatives) of the function f_i at x^{k+1}.

Let $F(x)$ be defined for $x \in U$ by (2.33) and for other $x \in \mathbb{R}^J$ let it be equal to $+\infty$ and let the set of minimizers of $\mathrm{cl}F$ over $\mathbb{R}^J$ be denoted by Φ. Denote $\Gamma := \inf\{\mathrm{cl}F(x) \mid x \in \mathbb{R}^J\}$ and consider the following assumptions.

Assumption A1: (*Zone Consistency*) For every $i = 1, 2, \ldots, I$, if $x^k \in S_{f_i}$ then $P_{C_i}^{f_i}(x^k) \in S_{f_i}$.

Assumption A2: For every $k = 1, 2, \ldots,$ the function $F_k(x) := \sum_{i=1}^{I} D_{f_i}(P_{C_i}^{f_i}(x^k), x)$ has a unique minimizer within U.

Assumption A3: If $\mathrm{cl}F(x) = 0$ for some x then x is in $C \cap \bar{U}$.

Theorem 2.11. [18, Theorem 4.1] *Let Assumptions **A1**, **A2** and **A3** hold and assume that the distances D_{f_i} are jointly convex, for all $i = 1, 2, \cdots, I$, In addition, assume that the set Φ is nonempty. If $\mathrm{cl}F$ has a unique minimizer then any sequence $\{x^k\}$, generated by Algorithm 2.10, converges to this minimizer. If Φ is not a singleton but $\Gamma = \inf\{\mathrm{cl}F(x) \mid x \in \mathbb{R}^J\} = 0$, then the intersection C of the sets C_i is nonempty and $\{x^k\}$ converges to a solution of the CFP.*

2.10. **2003: String-averaging projection schemes for inconsistent convex feasibility problems.** Censor and Tom [33] study iterative projection algorithms for the convex feasibility problem of finding a point in the intersection of finitely many nonempty, closed and convex subsets in the Euclidean space. They propose (without proof) an algorithmic scheme which generalizes both the string-averaging projections (SAP) and the block-iterative projections (BIP) methods with fixed strings or blocks, respectively, and prove convergence of the string-averaging method in the inconsistent case by translating it into a fully sequential algorithm in the product space.

They consider the successive projections iterative process

$$\begin{aligned} &x^0 \in V \text{ is an arbitrary starting point,} \\ &x^{k+1} = P_{k(\operatorname{mod} m)+1}\left(x^k\right) \text{ for all } k \geq 0, \end{aligned} \tag{2.35}$$

and offer an extension of Gubin Polyak and Raik's Theorem 2.4 above by replacing the demand that one of the sets of the CFP is bounded by a weaker condition, as follows (V stands for the Euclidean space).

Theorem 2.12. [33, Theorem 4.4] *Let $C_1, C_2, \cdots, C_m$ be nonempty closed convex subsets of V. If for at least one set (for explicitness, say C_1) the cyclic subsequence (of points in C_1) $\left\{x^{km+1}\right\}_{k\geq 0}$ of a sequence $\left\{x^k\right\}_{k\geq 0}$, generated by (2.35), is bounded for at least one $x^0 \in \mathbb{R}^n$ then there exist points $x^{*,i} \in C_i$, $i = 1, 2, \cdots, m$, such that $P_{i+1}\left(x^{*,i}\right) = x^{*,i+1}$, $i = 1, 2, \cdots, (m-1)$, and $P_1\left(x^{*,m}\right) = x^{*,1}$, and for $i = 1, 2, \cdots, m$, we have*

$$\lim_{k\to\infty} x^{km+i+1} - x^{km+i} = x^{*,i+1} - x^{*,i}, \tag{2.36}$$

$$\lim_{k\to\infty} x^{km+i} = x^{*,i}, \tag{2.37}$$

where $\left\{x^k\right\}_{k\geq 0}$ is any sequence generated by (2.35).

The SAP method has its origins in [28], for a recent work about it see, e.g., Reich and Zalas [60].

As in Censor and Tom [33, Page 545]), each string I_t is a finite nonempty subset of $\{1, 2, \cdots, m\}$, for $t = 1, 2, \cdots, S$, of the form $I_t = \left(i_1^t, i_2^t, \cdots, i_{\gamma(I_t)}^t\right)$, where the length of the string I_t, denoted by $\gamma(I_t)$, is the number of elements in I_t. The projection along the string I_t operator is defined as the composition of projections onto the sets indexed by I_t, that is, $T_t := P_{i_{\gamma(I_t)}^t} \cdots P_{i_2^t} P_{i_1^t}$ for $t = 1, 2, \cdots, S$. Given a positive weight vector $\omega \in \mathbb{R}^S$, i.e., $\omega_t > 0$, $t = 1, 2, ..., S$, and $\sum_{t=1}^S \omega_t = 1$,define the algorithmic operator

$$T = \sum_{t=1}^{S} \omega_t T_t, \tag{2.38}$$

yielding the SAP method that employs the iterative process

$$x^0 \in V \text{ is an arbitrary starting point, } x^{k+1} = T\left(x^k\right) \text{ for all } k \geq 0. \tag{2.39}$$

The convergence of the string-averaging method in the possibly inconsistent case is included in the following theorem.

Theorem 2.13. [33, Theorem 5.2] *Let $C_1, C_2, \cdots, C_m$, be nonempty closed convex subsets of V. If for at least one $x^0 \in V$ the sequence $\left\{x^k\right\}_{k\geq 0}$, generated by the string-averaging algorithm (Algorithm 2.39) with T as in (2.38)), is bounded then it converges for any $x^0 \in V$.*

2.11. **2004: Steered sequential projections for the inconsistent convex feasibility problem.** Censor, De Pierro and Zaknoon [25] study a steered sequential gradient algorithm which minimizes the sum of convex functions by proceeding cyclically in the directions of the negative gradients of the functions and using steered step-sizes. They apply this algorithm to the convex feasibility problem by minimizing a proximity function which measures the sum of the Bregman distances to the members of the family of convex sets. The resulting algorithm is a new steered sequential Bregman projection method which generates sequences that converge, if they are bounded, regardless of whether the convex feasibility problem is or is not consistent (i.e., feasible). For orthogonal projections and affine sets the boundedness condition is always fulfilled.

The steering parameters in the algorithm form a sequence $\{\sigma_k\}_{k\geq 0}$ of real positive numbers that must have the following properties: $\lim_{k\to\infty}\sigma_k = 0$, $\lim_{k\to\infty}(\sigma_{k+1}/\sigma_k) = 1$, and $\sum_{k=0}^{\infty}\sigma_k = +\infty$. If instead of $\lim_{k\to\infty}(\sigma_{k+1}/\sigma_k) = 1$ one uses $\lim_{k\to\infty}\sigma_{km+j}/\sigma_{km} = 1$, for all $1 \leq j \leq m-1$, then the parameters are called m-steering parameters. For minimization of a function $g(x) := \sum_{i=0}^{m-1} g_i(x)$ where $\{g_i\}_{i=0}^{m-1}$ is a family of convex functions from $\mathbb{R}^n$ into $\mathbb{R}$ which have continuous derivatives everywhere the cyclic gradient method is as follows

Algorithm 2.14. [25, Algorithm 5] (**The m-steered cyclic gradient method).**

Initialization: $x^0 \in \mathbb{R}^n$ is arbitrary.

Iterative Step: Given x^k calculate the next iterate x^{k+1} by

$$x^{k+1} = x^k - \sigma_k \nabla g_{i(k)}(x^k). \tag{2.40}$$

Control Sequence: $\{i(k)\}_{k\geq 0}$ is a cyclic control sequence, i.e., $i(k) = k \bmod m$.

Steering Parameters: The sequence $\{\sigma_k\}_{k\geq 0}$ is m-steering.

The following convergence result holds.

Theorem 2.15. [25, Theorem 6] *Let $\{g_i\}_{i=0}^{m-1}$ be a family of functions $g_i : \mathbb{R}^n \to \mathbb{R}$ which are convex and continuously differentiable everywhere, let $g(x) := \sum_{i=0}^{m-1} g_i(x)$ and assume that g has an unconstrained minimum. If $\left\{x^k\right\}_{k\geq 0}$ is a bounded sequence, generated by Algorithm 2.14, then the sequence $\left\{g\left(x^k\right)\right\}_{k\geq 0}$ converges to the minimum of g. If, in addition, g has a unique minimizer then the sequence $\left\{x^k\right\}_{k\geq 0}$ converges to this minimizer.*

Applying Algorithm 2.14 with Bregman distance functions and using Theorem 2.15 yields a convergence result for sequential Bregman projections onto convex sets in the inconsistent case.

Algorithm 2.16. [25, Algorithm 14] **(Steered cyclic Bregman projections).**

Initialization: $x^0 \in S$ is arbitrary.

Iterative Step: Given x^k calculate the next iterate x^{k+1} by

$$x^{k+1} = x^k + \sigma_k \nabla^2 f(x^k) \left(P^f_{Q_{i(k)}}(x^k) - x^k \right). \tag{2.41}$$

Control Sequence: $\{i(k)\}_{k\geq 0}$ is a cyclic control sequence, i.e., $i(k) = k \bmod m$.

Steering Parameters: The sequence $\{\sigma_k\}_{k\geq 0}$ is an m-steering sequence.

For the Bregman function $f(x) := (1/2)\|x\|^2$, the algorithm's iterative process takes the form

$$x^{k+1} = x^k + \sigma_k \left(P_{i(k)}(x^k) - x^k \right). \tag{2.42}$$

Another Bregman function is $f(x) = -\operatorname{ent} x$, where $\operatorname{ent} x$ is Shannon's entropy function which maps the nonnegative orthant $\mathbb{R}^n_+$ into $\mathbb{R}$ by $\operatorname{ent} x := -\sum_{j=1}^n x_j \log x_j$, where "log" denotes the natural logarithms and, by definition, $0 \log 0 = 0$. The steered cyclic entropy projections method that is obtained uses the iterative process

$$x^{k+1} = x^k + \sigma_k \begin{pmatrix} \frac{1}{x_1} & 0 & \cdots & 0 \\ 0 & \frac{1}{x_2} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \frac{1}{x_n} \end{pmatrix} \left(P^f_{i(k)}(x^k) - x^k \right) \tag{2.43}$$

and its convergence along the lines described above is obtained.

2.12. **2006: Alternating Bregman proximity operators.** It all began with Bregman's paper [17] which was actually Lev Bregman's Ph.D. work. This paper had no follow up in the literature until 14 years later in Censor and Lent [30]. For a bibliographic brief review on Bregman functions, distances and projections consult page 1233 (in the notes and references section) of the book by Facchinei and Pang [47]. See Subsection 2.9 here for additional pointers and read the recent excellent report of Reem, Reich and De Pierro [58].

In an attempt to apply compositions of Bregman projections to two disjoint convex sets Bauschke, Combettes and Noll [12] investigated the proximity properties of Bregman distances. This investigation lead to the introduction of a new type of proximity operator which complements the usual Bregman proximity operator.

The lack of symmetry inherent to the Bregman distance $D(x,y)$ prompted the authors to consider two single-valued operators defined on U, namely,

$$\begin{aligned} &\overleftarrow{prox}_{\varphi} : y \to \operatorname*{arg\,min}_{x\in U} \varphi(x) + D(x,y), \\ &\overrightarrow{prox}_{\psi} : x \to \operatorname*{arg\,min}_{y\in U} \psi(y) + D(x,y), \end{aligned} \tag{2.44}$$

where a variaty of assumptions (consult Bauschke, Combettes and Noll [12]) apply to the functions $\varphi(x)$, $\psi(y)$ and $D(x,y)$. They proposed the following iterative process [12, Equation (13)]

$$\begin{cases} \text{fix } x_0 \in U \text{ and set} \\ (\forall n \in \mathbb{N}) \quad y_n = \overrightarrow{prox}_{\psi}(x_n) \text{ and } x_{n+1} = \overleftarrow{prox}_{\varphi}(y_n), \end{cases} \tag{2.45}$$

for which they proved convergence that yielded the following corollary.

Corollary 2.17. [12, Corollary 4.7] *Let A and B be closed convex sets in $\mathbb{R}^J$ such that $A \cap U \neq \varnothing$ and $B \cap U \neq \varnothing$. Suppose that the solution set S of the problem*

$$\textit{minimize } D \textit{ over } (A \times B) \cap (U \times U) \tag{2.46}$$

is nonempty. Then the sequence $((x_n, y_n))_{n\in\mathbb{N}}$ generated by the alternating left-right projections algorithm

$$\begin{cases} \textit{fix } x_0 \in U \textit{ and set} \\ (\forall n \in \mathbb{N}) \quad y_n = \overrightarrow{P}_B(x_n) \textit{ and } x_{n+1} = \overleftarrow{P}_A(y_n), \end{cases} \tag{2.47}$$

converges to a point in S.

2.13. **2012: There is no variational characterization of the cycles in the method of periodic projections.** Baillon, Combettes and Cominetti [3] studied the behavior of the sequences generated by periodic projections onto $m \geq 3$ closed convex subsets of a Hilbert space $\mathcal{H}$. For an ordered family of nonempty closed convex subsets $C_1, C_2, \cdots, C_m$ of $\mathcal{H}$ with associated projection operators $P_1, P_2, \cdots, P_m$ consider sequences defined by the following rule: Choose any $x_0 \in \mathcal{H}$. For every $n = 0, 1, 2, 3, \cdots$ perform

$$\begin{cases} x_{mn+1} = P_m x_{mn} \\ x_{mn+2} = P_{m-1} x_{mn+1} \\ \vdots \\ x_{mn+m} = P_1 x_{mn+m-1}. \end{cases} \tag{2.48}$$

They define the set of cycles associated with the given m closed convex subsets by:

$$cyc(C_1, C_2, \cdots, C_m) = \left\{ \begin{array}{c} (\overline{y}_1, \overline{y}_2, \cdots, \overline{y}_m) \in \mathcal{H}^m \text{ such that} \\ \overline{y}_1 = P_1\overline{y}_2, \cdots, \ \overline{y}_{m-1} = P_{m-1}\overline{y}_m, \ \overline{y}_m = P_m\overline{y}_1 \end{array} \right\} \tag{2.49}$$

and asked if there exist a function $\Phi : \mathcal{H}^m \to \mathbb{R}$ such that, for every ordered family of nonempty closed convex subsets $(C_1, C_2, \cdots, C_m)$ of $\mathcal{H}$,

$cyc\,(C_1, C_2, \cdots, C_m)$ can be characterized as the solution set of a minimization problem of Φ? They proved the negative answer to this question by the following theorem.

Theorem 2.18. [3, Theorem 2.3] *Suppose that* $\dim \mathcal{H} \geq 2$ *and let* m *be an integer at least equal to* 3. *There exists no function* $\Phi : \mathcal{H}^m \to \mathbb{R}$ *such that, for every ordered family of nonempty closed convex subsets* $(C_1, C_2, \cdots, C_m)$ *of* $\mathcal{H}$, $cyc\,(C_1, C_2, \cdots, C_m)$ *is the set of solutions to the variational problem: minimize*$\{\Phi(y_1, y_2, ..., y_m) \mid y_1 \in C_1,\ y_2 \in C_2,\ \cdots, y_m \in C_m\}$.

This shows, in particular, that the Cheney and Goldstein [35] (see Subsection 2.1 above) result of minimizing the distance between two disjoint sets cannot be extended to more than two sets.

2.14. **2012: Alternating projections onto two sets with empty or nonempty intersection.** Kopecká and Reich [54] used tools from nonexpansive operators theory to prove that alternating projections onto two closed and convex subsets of a real Hilbert space H, generate two subsequences such that the sequence of distances between them converges to the distance between the two sets. Formally, let $P_1 : H \to S_1$ and $P_2 : H \to S_2$ be the orthogonal projections of H onto S_1 and S_2, respectively, and denote the distance between them by $d\,(S_1, S_2)$. Define the sequence $\{x_n \mid n \in \mathbb{N}\}$ by

$$x_{2n+1} = P_1 x_{2n} \text{ and } x_{2n+2} = P_2 x_{2n+1}, \tag{2.50}$$

then the following theorem holds

Theorem 2.19. [54, Theorem 1.4] *Let* S_1 *and* S_2 *be two nonempty, closed and convex subsets of a real Hilbert space* $(H, \langle\cdot,\cdot\rangle)$, *with induced norm* $\|\cdot\|$, *and let* $P_1 : H \to S_1$ *and* $P_2 : H \to S_2$ *be the corresponding nearest point projections of* H *onto* S_1 *and* S_2, *respectively. Let the sequence* $\{x_n \mid n \in \mathbb{N}\}$ *be defined by (2.50) then*

$$\lim_{n\to\infty} \|x_{2n+2} - x_{2n+1}\| = \lim_{n\to\infty} \|x_{2n+1} - x_{2n}\| = d\,(S_1, S_2)\,. \tag{2.51}$$

2.15. **2016: Inconsistent split feasibility problems.** The split common fixed point problem (SCFPP), first proposed in Censor and Segal [32], 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. This generalizes the convex feasibility problem (CFP), the two-sets split feasibility problem (SFP) and the multiple sets split feasibility problem (MSSFP).

Problem 2.20. The split common fixed point problem.
Given operators $U_i : R^N \to R^N$, $i = 1, 2, \ldots, p$, and $T_j : R^M \to R^M$, $j = 1, 2, \ldots, r$, with nonempty fixed points sets C_i, $i = 1, 2, \ldots, p$ and Q_j, $j = 1, 2, \ldots, r$, respectively. The `split common fixed point problem (SCFPP)` is

$$\text{find a vector } x^* \in C := \cap_{i=1}^{p} C_i \text{ such that } Ax^* \in Q := \cap_{i=1}^{r} Q_j. \tag{2.52}$$

Such problems 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 Censor, Bortfeld, Martin and Trofimov [23]. The problem with only a single pair of sets C in R^N and Q in R^M was first introduced by Censor and Elfving [27] and was called the split feasibility problem (SFP). They used their simultaneous multiprojections algorithm (see also Censor and Zenios [34, Subsection 5.9.2]) to obtain iterative algorithms to solve the SFP.

Iiduka [51] discusses the multiple-set split feasibility problem (MSFP)

$$\text{Find } x^* \in C := \cap_{i\in\mathcal{I}} C^{(i)} \text{ such that } Ax^* \in Q := \cap_{j\in\mathcal{J}} Q^{(j)}, \tag{2.53}$$

where $C^{(i)} \subseteq \mathbb{R}^N$ for all $i \in \mathcal{I} := \{1, 2, \cdots, I\}$ and $Q^{(j)} \subseteq \mathbb{R}^M$ for all $j \in \mathcal{J} := \{1, 2, \cdots, J\}$ are nonempty, closed and convex, and $A \in \mathbb{R}^{M\times N}$ is a matrix. The author [51, Page 187] introduces an inconsistent split feasibility problem (IMSFP) in which the CFP in $\mathbb{R}^N$ is "hard" (called there "absolute") and the CFP in $\mathbb{R}^M$ is "soft" (called there "subsidiary") as meant in Subsection 2.7 here. The infeasibility imposed by assuming that $\left(\cap_{i\in\mathcal{I}} C^{(i)}\right) \cap \left(\cap_{j\in\mathcal{J}} D^{(j)}\right) = \emptyset$ where $D^{(j)} := \left\{x \in \mathbb{R}^N \mid Ax \in Q^{(j)}\right\}$.

For user-chosen weights $\left(w^{(j)}\right)_{j\in\mathcal{J}} \subset (0, 1)$ satisfying $\sum_{j\in\mathcal{J}} w^{(j)} = 1$ Iiduka employs, for all $x \in \mathbb{R}^N$, the proximity function

$$f_D(x) := \frac{1}{2} \sum_{j\in\mathcal{J}} w^{(j)} \left\| P_{Q^{(j)}}(Ax) - Ax \right\|^2, \tag{2.54}$$

where $P_{Q^{(j)}}$ is the metric projection onto $Q^{(j)}$, and represents the IMSFP as the constrained minimization problem

$$\text{Find } x^* \text{ such that } f_D(x^*) = \min\left\{f_D(x) \mid x^* \in C := \cap_{i\in\mathcal{I}} C^{(i)}\right\}. \tag{2.55}$$

Actually, Iiduka uses nonexpansive mappings $T^{(i)} : \mathbb{R}^N \to \mathbb{R}^N$ and defines $C^{(i)} := Fix\left(T^{(i)}\right)$. He proposes a sophisticated projections method for solving (2.55) and formulates conditions under which it converges.

2.16. **2017: Best approximation pairs relative to two closed convex sets.** Given two disjoint closed convex sets, say C and Q, a best approximation pair relative to them is a pair of points, one in each set, attaining the minimum distance between the sets. Cheney and Goldstein [35] showed that alternating projections onto the two sets, starting from an arbitrary point, generate a sequence whose two interlaced subsequences converge to a best approximation pair. See Subsections 2.1 and 2.14 here. While Cheney and Goldstein considered only orthogonal projections onto the sets, related results, using the averaged alternating reflections (AAR) method and applying it to not necessarily convex sets, appeared in Bauschke, Combettes and Luke [11] and [55]. Remaining solely on theoretical ground, Achiya Dax [42] explores the duality relations that characterize least norm problems. He presents a new Minimum Norm Duality (MND) theorem, that considers the distance between two convex sets. Roughly speaking, it says that the shortest distance between the two sets is equal to the maximal "separation" between the sets, where the term "separation" refers to the distance between a pair of parallel hyperplanes that separates the two sets.

The problem of best approximation pair relative to two sets cannot be extended to more than two sets in view of Baillon, Combettes and Cominetti's result [3], see Subsection 2.13 here. From a practical point of view, the best approximation pair relative to two sets obviously furnishes a solution to the hard and soft constrained inconsistent feasibility problem, see Subsection 2.7 here. However, the algorithmic approach is hindered by the need to perform projections onto the two sets C and Q if they are not "simple to project onto".

Aharoni, Censor and Jiang [2] propose, for the polyhedral sets case, a process based on projections onto the half-spaces defining the two polyhedra, which are more negotiable than projections on the polyhedra themselves. A central component in their proposed process is the Halpern–Lions–Wittmann–Bauschke (HLWB[2]) algorithm for approaching the projection of a given point onto a convex set.

The HLWB algorithm is applied alternatingly to the two polyhedra. Its application is divided into sweeps — in the odd numbered sweeps we project successively onto half-spaces defining the polyhedron C, and in even numbered sweeps onto half-spaces defining the polyhedron Q. A critical point is that the number of successive projections onto each set's half-spaces increases from sweep to sweep. The proof of convergence of the algorithm is rather standard in the case that the best approximation pair is unique. The non-uniqueness case, however, poses some difficulties and its proof is more involved.

2.17. **2018: Replacing inconsistent sets with set enlargements: ART3, ARM, Intrepid, Valiant.** Searching for a solution to a system of linear equations is a convex feasibility problem and has led to many different iterative methods. When the system of linear equations is inconsistent, due to modeling or measurements inaccuracies, it has been suggested to replace it by a system of pairs of opposing linear inequalities creating nonempty hyperslabs. Applying projection methods to this problem can be done by using any iterative method for linear inequalities, such as the method of Agmon [1] and Motzkin and Schoenberg [56] (AMS). However, in order to improve computational efficiency, Goffin [48] proposed to replace projections onto the hyperslabs by a strategy of projecting onto the original hyperplane (from which the hyperslab was created) when the current iterate is "far away" from the hyperslab, and reflecting into the hyperslab's boundary when the current iterate is "close to the hyperslab" while keeping the iterate unchanged if it is already inside the hyperslab.

[2]This acronym was dubbed in [22].

In [50] Herman suggested to implement Goffin's strategy by using an additional enveloping hyperslab in order to determine the "far" and the "close" distance of points from the hyperplane, resulting in his "Algebraic Reconstruction Technique 3" (ART3) algorithm. In [20] Censor also embraced the idea of hyperslabs, and defined an algorithmic operator that implemented Goffin's strategy in a continuous manner, resulting in the Automatic Relaxation Method (ARM). For applications and additional details see Censor and Herman [29] and [21].

A fundamental question that remained open since then was whether the hyperslabs approach to handle linear equations and Goffin's principle can be applied to general convex sets. This question was recently studied by Bauschke, Iorio and Koch in [14], see also [15] and Bauschke, Koch and Phan [16] for further details and interesting applications. They defined convex sets enlargements instead of hyperslabs and used them to generalize the algorithmic operator that appeared in Herman [50]. They defined an operator which they called the "intrepid projector", intended to generalize the ART3 algorithm of [50] to convex sets. Motivated by [14], Censor and Mansour [31] present a new operator, called the "valiant operator", that enables to implement the algorithmic principle embodied in the ARM of [20] to general convex feasibility problems. Both ART3 and ARM seek a feasible point in the intersections of the hyperslabs and so their generalizations to the convex case seek feasibility of appropriate enlargement sets that define the extended problem.

Acknowledgement. We dedicate this paper to Adi Ben-Israel, our scientific father and grandfather, respectively. We thank Heinz Bauschke for calling our attention to some relevant references. This work was supported by Research Grant No. 2013003 of the United States-Israel Binational Science Foundation (BSF).

Author Index

Agmon, 19
Aharoni, 19

Baillon, 11, 16, 19
Bauschke, 1, 6, 8–11, 15, 16, 18–20
Bondon, 9
Bortfeld, 18
Borwein, 1, 6, 8, 9, 11
Bregman, 11, 12, 14, 15
Byrne, 11

Cegielski, 1, 2
Censor, 2, 5, 10, 11, 13–15, 17–20
Cheney, 3, 6, 17, 18
Combettes, 1, 6, 9, 11, 15, 16, 18, 19
Cominetti, 11, 16, 19
Csiszár, 11

Dax, 18
De Pierro, 10, 11, 14, 15
Dykstra, 6

Eckstein, 11
Edwards, 10, 11
Eggermont, 5, 10
Elfving, 18

Facchinei, 15

Goffin, 19, 20
Goldstein, 3, 6, 17, 18
Gordon, 5, 10
Gubin, 4, 13

Halpern, 19
Herman, 5, 20

Iiduka, 18
Iorio, 20

Jiang, 19

Kaczmarz, 5
Kadec/Klee, 8
Koch, 20
Kopecká, 4, 6, 17

Lent, 5, 11, 15
Lewis, 8, 9
Lions, 19
Luke, 18

Mansour, 20
Martin, 18
Moore-Penrsose, 5
Motzkin, 19

Neumann, 6
Noll, 15, 16

Pang, 15
Phan, 20
Pierra, 11
Polyak, 4, 13

Raik, 4, 13
Reem, 15
Reich, 4, 6, 13, 15, 17
Roshchina, 11

Schoenberg, 19
Segal, 17
Solodov, 11
Svaiter, 11

Tom, 13
Trofimov, 18
Tusnády, 11

von Neumann, 6

Williamson, 11
Wittmann, 19

Zaknoon, 14
Zalas, 13
Zenios, 11, 18

References

[1] S. Agmon. The relaxation method for linear inequalities. *Canadian Journal of Mathematics*, 6:382–392, 1954.

[2] R. Aharoni, Y. Censor, and Z. Jiang. Finding a best approximation pair of points for two polyhedra. Technical report, https://arxiv.org/abs/1707.09639, July 2017.

[3] J.B. Baillon, P.L. Combettes, and R. Cominetti. There is no variational characterization of the cycles in the method of periodic projections. *Jornal of Functional Analysis*, 262:400–408, 2012.

[4] J.B. Baillon, P.L. Combettes, and R. Cominetti. Asymptotic behavior of compositions of underrelaxed nonexpansive operators. *Journal of Dynamics and Games*, 1:331–346, 2014.

[5] H.H. Bauschke. The composition of projections onto closed convex sets in hilbert space is asymptotically regular. *Proceedings of the American Mathematical Society*, 131:141–146, 2003.

[6] H.H. Bauschke and J.M. Borwein. On the convergence of von neumann's alternating projection algorithm for two sets. *Set-Valued Analysis*, 1:185–212, 1993.

[7] H.H. Bauschke and J.M. Borwein. Dykstra's alternating projection algorithm for two sets. *Journal of Approximation Theory*, 79:418–443, 1994.

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

[9] H.H. Bauschke and J.M. Borwein. Legendre functions and the method of random bregman projections. *Journal of convex analysis*, 4:27–67, 1997.

[10] H.H. Bauschke, J.M. Borwein, and A.S. Lewis. The method of cyclic projections for closed convex sets in Hilbert space. In Y. Censor and S. Reich, editors, *Recent Developments in Optimization Theory and Nonlinear Analysis*, pages 1–38. American Mathematical Society, Providence, RI, USA, 1997.

[11] H.H. Bauschke, P.L. Combettes, and D.R. Luke. Finding best approximation pairs relative to two closed convex sets in hilbert spaces. *Journal of Approximation Theory*, 127:178–192, 2004.

[12] H.H. Bauschke, P.L. Combettes, and D. Noll. Joint minimization with alternating bregman proximity operators. *Pacific Journal of Optimization*, 2:401–424, 2006.

[13] H.H. Bauschke and M.R. Edwards. A conjecture by De Pierro is true for translates of regular subspaces. *Journal of Nonlinear and Convex Analysis*, 6:93–116, 2005.

[14] H.H. Bauschke, F. Iorio, and V. Koch. The method of cyclic intrepid projections: Convergence analysis and numerical experiments. In M. Wakayama, R.S. Anderssen, J. Cheng, Y. Fukumoto, R. McKibbin, K. Polthier, T. Takagi, and K.-C. Toh, editors, *The Impact of Applications on Mathematics*, pages 187–200. Springer, Japan, 2014.

[15] H.H. Bauschke and V. Koch. Projection methods: Swiss army knives for solving feasibility and best approximation problems with halfspaces. *Contemporary Mathematics*, 636:1–40, 2015.

[16] H.H. Bauschke, V. Koch, and H. Phan. Stadium norm and Douglas-Rachford splitting: a new approach to road design optimization. *Operations Research*, 64:201–218, 2016.

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

[18] C. Byrne and Y. Censor. Proximity function minimization using multiple Bregman projections, with applications to split feasibility and Kullback-Leibler distance minimization. *Annals of Operations Research*, 105:77–98, 2001.

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

[20] Y. Censor. An automatic relaxation method for solving interval linear inequalities. *Journal of Mathematical Analysis and Applications*, 106:19–25, 1985.

[21] Y. Censor. Parallel application of block-iterative methods in medical imaging and radiation therapy. *Mathematical Programming*, 42:307–325, 1988.

[22] Y. Censor. Computational acceleration of projection algorithms for the linear best approximation problem. *Linear Algebra and Its Applications*, 416:111–123, 2006.

[23] 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*, 51:2353–2365, 2006.

[24] Y. Censor and A. Cegielski. Projection methods: an annotated bibliography of books and reviews. *Optimization*, 64:2343–2358, 2015.

[25] Y. Censor, A.R. De Pierro, and M. Zaknoon. Steered sequential projections for the inconsistent convex feasibility problem. *Nonlinear Analysis*, 59:385–405, 2004.

[26] Y. Censor, P.P.B. Eggermont, and D. Gordon. Strong underrelaxation in Kaczmarz's method for inconsistent systems. *Numerische Mathematik*, 41:83–92, 1983.

[27] Y. Censor and T. Elfving. A multiprojection algorithm using bregman projections in a product space. *Numerical Algorithms*, 8:221–239, 1994.

[28] Y. Censor, T. Elfving, and G.T. Herman. Averaging strings of sequential iterations for convex feasibility problems. In D. Butnariu, Y. Censor, and S. Reich, editors, *Inherently Parallel Algorithms in Feasibility and Optimization and Their Applications*, pages 101–114. Elsevier Science Publishers, Amsterdam, The Netherlands, 2001.

[29] Y. Censor and G.T. Herman. On some optimization techniques in image reconstruction from projections. *Applied Numerical Mathematics*, 3:365–391, 1987.

[30] Y. Censor and A. Lent. An iterative row-action method for interval convex programming. *Journal of Optimization Theory and Applications*, 34:321–353, 1981.

[31] Y. Censor and R. Mansour. Convergence analysis of processes with valiant projection operators in Hilbert space. *Journal of Optimization Theory and Applications*, 176:35–56, 2018.

[32] Y. Censor and A. Segal. The split common fxed point problem for directed operators. *Journal of Convex Analysis*, 16:587–600, 2009.

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

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

[35] W. Cheney and A.A. Goldstein. Proximity maps for convex sets. *Proceeding of the American Mathematical Society*, 10:448–450, 1959.

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

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

[38] P.L. Combettes. Convex set theoretic image recovery by extrapolated iterations of parallel subgradient projections. *IEEE Transactions on Image Processing*, 6:493–506, 1997.

[39] P.L. Combettes and P. Bondon. Hard-constrained inconsistent signal feasibility problems. *IEEE Transactions on Signal Processing*, 47:2460–2468, 1999.

[40] R. Cominetti, V. Roshchina, and A. Williamson. A counterexample to De Pierro's conjecture on the convergence of under-relaxed cyclic projections. Technical report, https://arxiv.org/abs/1801.03216, 2018.

[41] Csiszár and Tusnády. Information geometry and alternating minimization procedures. *Statistics and Decisions, Supplement Issue*, 1:205–237, 1984.

[42] A. Dax. The distance between two convex sets. *Linear Algebra and its Applications*, 416:184–213, 2006.

[43] A.R. De Pierro. From parallel to sequential projection methods and vice versa in convex feasibility: Results and conjectures. In D. Butnariu, Y. Censor, and S. Reich, editors, *Inherently Parallel Algorithms in Feasibility and Optimization and Their Applications*, pages 187–201. Elsevier Science Publishers, Amsterdam, The Netherlands, 2001.

[44] J. Eckstein. Approximate iterations in Bregman-function-based proximal algorithms. *Mathematical Programming*, 83:113–123, 1998.

[45] J. Eckstein. A practical general approximation criterion for methods of multipliers based on Bregman distances. *Mathematical Programming*, Series A, 96:61–86, 2003.

[46] 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*, 40:37–67, 1981.

[47] F. Facchinei and J.S. Pang. *Finite-Dimensional Variational Inequalities and Complementarity Problems*, volume II. Springer-Verlag, New York, NY, USA, 2003.

[48] J.L. Goffin. On the finite convergence of the relaxation method for solving systems of inequalities. Technical report, Ph.D. Thesis, Report ORC 71-36, Operation Research Center, University of California, Berkeley, 1971.

[49] L.G. Gubin, B.T. Polyak, and E.V. Raik. The method of projections for finding the common point of convex sets. *USSR Computational Mathematics and Mathematical Physics*, 7:1–24, 1967.

[50] G.T. Herman. A relaxation method for reconstructing objects from noisy x-rays. *Mathematical Programming*, 8:1–19, 1975.

[51] H. Iiduka. Optimization for inconsistent split feasibility problems. *Numerical Functional Analysis and Optimization*, 37:186–205, 2016.

[52] S. Kaczmarz. Angenäherte Auflösung von Systemen linearer Gleichungen. *Bulletin International de l'Academié Polonaise des Sciences et des Lettres. Classe des Sciences Mathématiques et Naturelles. Série A, Sciences Mathématiques*, A 35:355–357, 1937.

[53] E. Kopecká and S. Reich. A note on the von Neumann alternating projections algorithm. *Journal of Nonlinear and Convex Analysis*, 5:379–386, 2004.

[54] E. Kopecká and S. Reich. A note on alternating projections in Hilbert space. *Journal of Fixed Point Theory and Applications*, 12:41–47, 2012.

[55] D.R. Luke. Finding best approximation pairs relative to a convex and prox-regular set in a Hilbert space. *SIAM Journal on Optimization*, 19:714–739, 2008.

[56] T.S. Motzkin and I.J. Schoenberg. The relaxation method for linear inequalities. *Canadian Journal of Mathematics*, 6:393–404, 1954.

[57] G. Pierra. Decomposition through formalization in a product space. *Mathematical Programming*, 28:96–115, 1984.

[58] D. Reem, S. Reich, and A. De Pierro. Re-examination of Bregman functions and new properties of their divergences. Technical report, https://arxiv.org/abs/1803.00641, March 2018.

[59] S. Reich. A limit theorem for projections. *Linear Multilinear Algebra*, 13:281–290, 1983.

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

[61] M.V. Solodov and B.F. Svaiter. An inexact hybrid generalized proximal point algorithm and some new results on the theory of bregman functions. *Mathematics of Operations Research*, 25:214–230, 2000.

(Y. Censor) Department of Mathematics, University of Haifa Mt. Carmel, Haifa 3498838, Israel
E-mail address: yair@math.haifa.ac.il

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