Pubblicazioni del Centro De Giorgi
Proceedings of the Interdisciplinary Workshop on Mathematical Methods in Biomedical Imaging and Intensity-Modulated Radiation Therapy (IMRT), Pisa, Italy, October 2007.

Iterative Projection Methods in Biomedical Inverse Problems

Yair Censor and Alexander Segal

ABSTRACT. The convex or quasiconvex feasibility problem and the split feasibility problem in the Euclidean space have many applications in various fields of science and technology, particularly in problems of image reconstruction from projections, in solving the fully discretized inverse problem in radiation therapy treatment planning, and in other image processing problems. Solving systems of linear equalities and/or inequalities is one of them. The class of methods, generally called *Projection Methods,* has witnessed great progress in recent years and its member algorithms have been applied with success to fully discretized models of inverse problems in image reconstruction and image processing, and in intensity-modulated radiation therapy. We introduce the reader to this field by reviewing algorithmic structures and specific algorithms for the convex feasibility problem, the quasiconvex feasibility problem and the split feasibility problem.

Department of Mathematics, University of Haifa, Mt. Carmel, Haifa 31905, Israel
Department of Mathematics, University of Haifa, Mt. Carmel, Haifa 31905, Israel
[1]

Key words and phrases. Projection algorithms, block-iterative, Bregman projections, convex feasibility, split feasibility, string-averaging, quasiconvex.

[1]We thank Arnold Lent for his comments on an earlier version of the paper. Parts of the paper are adapted from [**28**]. This work was supported by grant No. 2003275 of the United States-Israel Binational Science Foundation (BSF) and by a National Institutes of Health (NIH) grant No. HL70472.

1. Introduction

In this paper we focus our attention on iterative projection methods that have been found useful in biomedical inverse problems. We review some basic results that are by now well-known and discuss also some more recent developments. The presentation is admittedly biased towards our own work but contains also many 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 [**87**], Censor and Zenios [**47**]. Our aim in this paper is to introduce the reader to this field by reviewing algorithmic structures and specific algorithms for the convex feasibility problem, the quasiconvex feasibility problem and the split feasibility problem.

The *convex feasibility problem* is to find a point (any point) in the nonempty intersection $C := \cap_{i=1}^{m} C_i \neq \emptyset$ of a family of closed convex subsets $C_i \subseteq R^n$, $1 \leq i \leq m,$ of the n-dimensional Euclidean space R^n. It is a fundamental problem in many areas of mathematics and the physical sciences, see, e.g., Combettes [**50, 52**] and references therein. It has been used to model significant real-world problems in image reconstruction from projections, see, e.g., Herman [**69**], in radiation therapy treatment planning, see Censor, Altschuler and Powlis [**29**] and Censor [**27**], and in crystallography, see Marks, Sinkler and Landree [**78**], 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* (i.e., nearest point mappings) onto the individual sets C_i. The orthogonal projection $P_\Omega(z)$ of a point $z \in R^n$ onto a closed convex set $\Omega \subseteq R^n$ is defined by

$$P_\Omega(z) := \operatorname{argmin}\{\| z - x \|_2 \mid x \in \Omega\}, \tag{1.1}$$

where $\| \cdot \|_2$ is the Euclidean norm in R^n. Frequently a *relaxation parameter* is introduced so that

$$P_{\Omega,\lambda}(z) := (1-\lambda)z + \lambda P_\Omega(z) \tag{1.2}$$

is the *relaxed projection* of z onto Ω with relaxation λ.

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 $C_i \subseteq R^n$, $i = 1, 2, \dots, t$, in the n-dimensional Euclidean space R^n, and nonempty closed convex sets $Q_j \subseteq R^m$, $j = 1, 2, \dots, r$, and an $m \times n$ real matrix A, the *multiple-sets split feasibility problem* (MSSFP) is

$$\text{find a vector } x^* \in C := \cap_{i=1}^t C_i \text{ such that } Ax^* \in Q := \cap_{i=1}^r Q_j. \quad (1.3)$$

Such MSSFPs, formulated in [**35**], 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 [**30**].

The *quasiconvex feasibility problem* is a different generalization of the convex feasibility problem. Let $f_1(x), f_2(x), \dots, f_m(x)$ be continuous quasiconvex functions defined on R^n. The quasiconvex feasibility problem is to find a point x^*, such that $f_i(x^*) \leq 0$ for $i = 1, 2, \dots, m$. The notion *quasiconvex feasibility problem* was introduced by Goffin, Luo and Ye in [**66**], where they used cutting planes algorithms but only the differentiable case was considered there.

Another problem that is related to the convex feasibility problem is the *best approximation problem* of finding the projection of a given point $y \in R^n$ onto the non-empty intersection $C := \cap_{i=1}^m C_i \neq \emptyset$ of a family of closed convex subsets $C_i \subseteq R^n$, $1 \leq i \leq m$, see, e.g., Deutsch's book [**61**]. The convex sets $\{C_i\}_{i=1}^m$ commonly represent mathematical constraints obtained from the modeling of the real-world problem. In the convex feasibility approach any point in the intersection is an acceptable solution to the real-world problem whereas the best approximation formulation is usually appropriate if some point $y \in R^n$ has been obtained from modeling and computational efforts which initially did not take into account the constraints represented by the sets $\{C_i\}_{i=1}^m$ and now one wishes to incorporate them by seeking a point in the intersection of the convex sets which is closest to the point y.

Iterative projection algorithms for finding a projection of a point onto the intersection of sets are more complicated then algorithms for finding just any feasible point in the intersection. This is so because they must have, in their iterative steps, some built-in "memory" mechanism to remember the original point whose projection is sought after. The sequential or parallel algorithms of Dykstra in Bregman, Censor and Reich [**15**], of Haugazeau in Bauschke and Combettes [**7**], of Bauschke [**3**] and others, and their modifications, employ different such memory mechanisms.

We will not deal with these algorithms here although many of them share the same algorithmic structural features described below.

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, that we made in earlier publications, 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, methods 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 [**26**]) while simultaneous algorithmic structures favor parallel computing platforms, see, e.g., Censor, Gordon and Gordon [**37**]. 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 [**76**] on the SIRT method, the works of Crombez [**56, 58**], the connection with variational inequalities, see, e.g., Aslam Noor [**80**], Yamada's [**89**] 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 [**8**] and references therein. Consult Bauschke and Borwein [**4**] and Censor and Zenios [**47**, 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 [**65**], Censor and Lent [**41**] and Chinneck [**48**]), approximation theory (see, e.g., Deutsch [**61**] and references therein) and image reconstruction from projections in computerized tomography (see, e.g., Herman [**69, 70**], Censor [**26**]).

2. Bregman projections

Bregman projections onto closed convex sets were introduced by Censor and Lent [**40**], based on Bregman's seminal paper [**14**], and were subsequently used in a plethora of research works as a tool for building sequential and parallel feasibility and optimization algorithms, see, e.g., Censor and Elfving [**31**], Censor and Reich [**44**], Censor and Zenios [**47**], De Pierro and Iusem [**60**], Kiwiel [**73, 74**], Bauschke and Borwein [**5**] and references therein, to name but a few.

A *Bregman projection* of a point $z \in R^n$ onto a closed convex set $\Omega \subseteq R^n$ with respect to a, suitably defined, *Bregman function* f (see, e.g., [**47**, Definition 2.1.1]) is denoted by $P_\Omega^f(z)$. It is formally defined as

$$P_\Omega^f(z) := \operatorname{argmin}\{D_f(x,z) \mid x \in \Omega \cap \mathrm{cl}S\} \tag{2.1}$$

where cl$S$ is the closure of the open convex set S, which is the *zone* of f, and $D_f(x,z)$ is the so-called *Bregman distance*, defined by

$$D_f(x,z) := f(x) - f(z) - \langle \nabla f(z), x - z\rangle, \tag{2.2}$$

for all pairs $(x,z) \in \mathrm{cl}S \times S$, where $\langle \cdot,\cdot\rangle$ is the standard inner product in R^n. If $\Omega \cap \mathrm{cl}S \neq \emptyset$, then (2.1) defines a unique $P_\Omega^f(z) \in \mathrm{cl}S$, for every $z \in S$ [**47**, Lemma 2.1.2]. If, in addition, $P_\Omega^f(z) \in S$, for every $z \in S$, then f is called *zone consistent* with respect to Ω. If f is a Bregman/Legendre function (see Bauschke and Borwein [**5**, Theorem 3.14]) and $S = \operatorname{int}(\operatorname{dom} f)$, then f is zone consistent with respect to any closed convex set Ω such that $\Omega \cap \mathrm{cl}S \neq \emptyset$.

Orthogonal projections are a special case of Bregman projections, obtained from (2.1) by choosing [**47**, Example 2.1.1] $f(x) = (1/2)\|x\|^2$ and $S = R^n$. Bregman generalized distances and generalized projections are instrumental in several areas of mathematical optimization theory.

They were used, among others, in special-purpose minimization methods, in the proximal point minimization method, and for stochastic feasibility problems. These generalized distances and projections were also defined in non-Hilbertian Banach spaces, where, in the absence of orthogonal projections, they can lead to simpler formulas for projections, see, e.g., Butnariu and Iusem [**19**] and references therein.

Bregman's method for minimizing a convex function (with certain properties) subject to linear inequality constraints employs Bregman projections onto the half-spaces represented by the constraints [**40, 60**]. Recently the extension of this minimization method to nonlinear convex constraints has been identified with the Dykstra projection algorithm for finding the projection of a point onto an intersection of closed convex sets, see Bregman, Censor and Reich [**15**]. It looks as if there might be no point in using non-orthogonal projections for solving the convex feasibility problem in R^n since they are generally not easier to compute. But this is not always the case. Shamir and co-workers [**75, 77**] have used the multiprojection method of Censor and Elfving [**31**] to solve filter design problems in image restoration and image recovery posed as convex feasibility problems. They took advantage of that algorithm's flexibility to employ Bregman projections with respect to *different* Bregman functions within the same algorithmic run. Another example is the seminal paper by Csiszár and Tusnády [**59**], where the central procedure uses alternating entropy projections onto convex sets. In their "alternating minimization procedure," they alternate between minimizing over the first and second arguments of the Kullback-Leibler divergence. This divergence is nothing but the generalized Bregman distance obtained by using the negative of Shannon's entropy as the underlying Bregman function. Recent studies about Bregman projections (Kiwiel [**74**]), Bregman/Legendre projections (Bauschke and Borwein [**5**]), and averaged entropic projections (Butnariu, Censor and Reich [**17**]) – and their uses for convex feasibility problems in R^n discussed therein – attest to the continued theoretical and practical interest in employing Bregman projections in projection methods for convex feasibility problems.

3. Algorithmic structures

Projection algorithmic schemes for the convex feasibility problem and for the best approximation problem are, in general, either *sequential* or *simultaneous* or *block-iterative* (see, e.g., Censor and Zenios [**47**] for a classification of projection algorithms into such classes, and the review paper of Bauschke and Borwein [**4**] for a variety of specific algorithms of these kinds). In the following subsections we explain and demonstrate these structures along with the more recent *string averaging* structure. The philosophy behind these algorithms is that it is easier to calculate projections onto the individual sets C_i then onto the whole intersection of sets. Thus, these algorithms call for projections onto individual sets as they proceed sequentially, simultaneously or in the block-iterative or the string-averaging algorithmic modes.

3.1. Sequential projections. The well-known "Projections Onto Convex Sets" (POCS) algorithm for the convex feasibility problem is a *sequential* projection algorithm, see Bregman [**13**], Gubin, Polyak and Raik [**68**], Youla [**91**] and the review papers by Combettes [**50, 52**]. Starting from an arbitrary initial point $x^0 \in R^n$, the POCS algorithm's iterative step is

$$x^{k+1} = x^k + \lambda_k (P_{C_{i(k)}}(x^k) - x^k), \tag{3.1}$$

where $\{\lambda_k\}_{k\geq 0}$ are relaxation parameters and $\{i(k)\}_{k\geq 0}$ is a *control sequence*, $1 \leq i(k) \leq m$, for all $k \geq 0$, which determines the individual set $C_{i(k)}$ onto which the current iterate x^k is projected.

DEFINITION 3.1. **(Control sequences)**

(1) **Almost cyclic control.** A control sequence $\{i(k)\}_{k\geq 0}$ 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 $\sigma \geq m$ (called the *almost cyclicality constant*) such that, for all $k \geq 0$, $\{1, 2, \ldots, m\} \subseteq \{i(k+1), i(k+2), \ldots, i(k+\sigma)\}$. An almost cyclic control with $\sigma = m$ is called *cyclic*.
(2) **Most violated constraint control.** This sequence $\{i(k)\}_{k\geq 0}$ is obtained by determining which constraint is most violated by the iterate x^k. If $C_i = \{x \in R^n \mid f_i(x) \leq 0\}$, are the sets in the feasibility problem then $i(k)$ is the *most violated constraint control index* if $f_{i(k)}(x^k) > 0$ and

$$f_{i(k)}(x^k) = \max\{f_i(x^k) \mid i = 1, 2, \ldots, m\}. \tag{3.2}$$

Other controls are also available, e.g., [**47**, Definition 5.1.1]. Bregman's projection algorithm [**47, 14**], allowed originally only unrelaxed projections, i.e., its iterative step is of the form

$$x^{k+1} = P^f_{C_{i(k)}}(x^k), \quad \text{for all } k \geq 0. \tag{3.3}$$

This has been extended by Censor and Herman [**39**]. For the Bregman function $f(x) = (1/2)\|x\|^2$ with zone $S = R^n$ and for unity relaxation ($\lambda_k = 1$, for all $k \geq 0$), (3.3) coincides with (3.1).

3.2. The string averaging algorithmic structure. This prototypical algorithmic scheme was proposed by Censor, Elfving and Herman [**34**]. For $t = 1, 2, \ldots, M$, let the *string* I_t be an ordered subset of $\{1, 2, \ldots, m\}$ of the form

$$I_t = (i_1^t, i_2^t, \ldots, i_{m(t)}^t), \tag{3.4}$$

with $m(t)$ denoting the number of elements in I_t. Suppose that there is a set $S \subseteq R^n$ such that there are operators $R_1, R_2, \ldots, R_m$ mapping S into S and an operator R which maps $S^M = S \times S \times \cdots \times S$ (M times) into S. Initializing the algorithm at an arbitrary $x^0 \in S$, the iterative step of the string averaging prototypical algorithmic scheme is as follows. Given the current iterate x^k, calculate, for all $t = 1, 2, \ldots, M$,

$$T_t(x^k) = R_{i_{m(t)}^t}(\ldots(R_{i_2^t}(R_{i_1^t}(x^k))), \tag{3.5}$$

and then calculate

$$x^{k+1} = R(T_1(x^k), T_2(x^k), \ldots, T_M(x^k)). \tag{3.6}$$

For every $t = 1, 2, \ldots, M$, this prototypical algorithmic scheme applies to x^k successively the operators whose indices belong to the t-th string. This can be done in parallel for all strings and then the operator R maps all end-points onto the next iterate x^{k+1}. This is indeed an algorithm provided that the operators $\{R_i\}_{i=1}^m$ and R all have algorithmic implementations. In this framework we get a *sequential* algorithm by the choice $M = 1$ and $I_1 = (1, 2, \ldots, m)$ and a *simultaneous* algorithm by the choice $M = m$ and $I_t = (t),\ t = 1, 2, \ldots, M$.

We 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 R_i to be orthogonal projections onto the hyperplanes. The operator R is taken as a convex combination

$$R(x^1, x^2, \ldots, x^M) = \sum_{t=1}^{M} \omega_t x^t, \tag{3.7}$$

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

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 [**47**]. The fully simultaneous algorithmic structure appears in Figure 1(b). With orthogonal reflections instead of orthogonal projections it was first proposed, by Cimmino [**49**], for solving linear equations, see also Benzi [**10**]. 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 $M = m$ and $I_t = (1, 2, \ldots, t)$, for $t = 1, 2, \ldots, M$. This scheme, appearing in Bauschke and Borwein [**4**], inspired the 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 $I_1 = (1, 3, 5, 6)$, $I_2 = (2)$, $I_3 = (6, 4)$. 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.

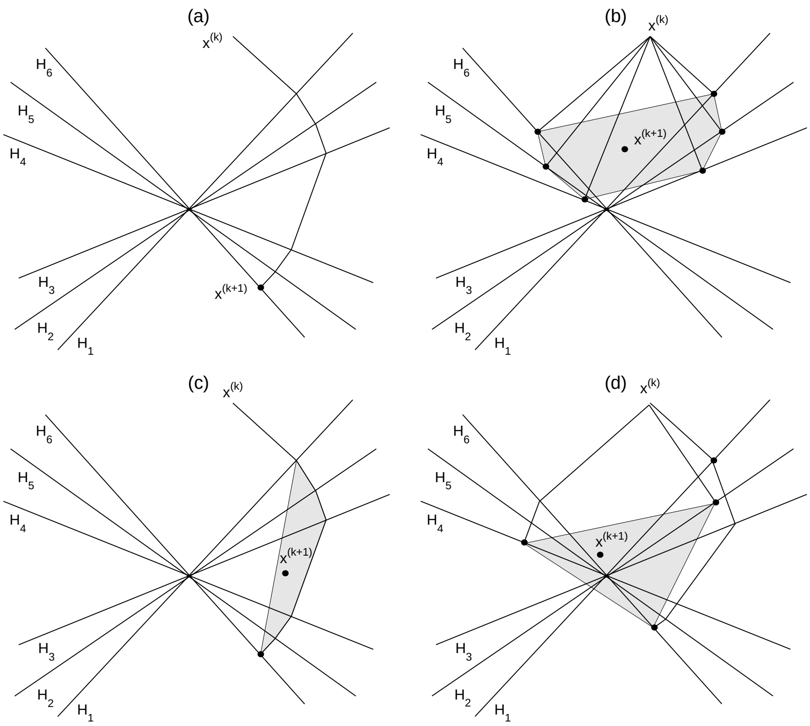

Figure 1. (a) Sequential projections. (b) Fully simultaneous projections. (c) Averaging of sequential projections. (d) String averaging. (Reproduced from Censor, Elfving and Herman [**34**]).

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 [**51**] or Iusem and De Pierro [**71**]. 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 [**6**], Elfving [**63**] and Eggermont, Herman and Lent [**62**]). 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 $C = \cap_{i=1}^m C_i$ is empty was studied by Censor and Tom in [**46**]. They defined projection along the string I_t operator as the composition of orthogonal projections onto sets indexed by I_t, that is,

$$T_t := P_{i^t_{m(t)}} \cdots P_{i^t_2} P_{i^t_1}, \text{ for } t = 1, 2, \ldots, M, \tag{3.8}$$

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

$$R = \sum_{t=1}^{S} \omega_t T_t. \tag{3.9}$$

Using this R the following string averaging algorithm is obtained.

ALGORITHM 3.2.
Initialization: *$x^0 \in V$ is an arbitrary starting point.*
Iterative Step: *Given x^k, use (3.8) and (3.9) to compute x^{k+1}*

$$x^{k+1} = R(x^k). \tag{3.10}$$

THEOREM 3.3. [**46**] *Let $C_1, C_2, \ldots, C_m$, be nonempty closed convex subsets of R^n. If for at least one $x^0 \in R^n$ the sequence $\{x^k\}_{k\geq 0}$, generated by the string averaging algorithm (Algorithm 3.2 with R as in*

(3.9)), is bounded then any sequence $\{x^k\}_{k\geq 0}$, *generated by the string averaging algorithm (Algorithm 3.2 with* R *as in (3.9)), converges for any* $x^0 \in R^n$.

Varying and iteration dependent relaxation parameters and string constructions could be interesting future extensions. The practical performance of specific algorithms needs also to be evaluated in applications and on parallel machines. The string averaging prototypical algorithmic scheme has attracted attention recently and further work on it has been reported since its presentation in [**34**]. In Bauschke, Matoušková and Reich [**9**] string averaging was studied in Hilbert space. In Crombez [**55, 57**] the string averaging algorithmic paradigm is used to find common fixed points of certain paracontractive operators in Hilbert space. In Bilbao-Castro, Carazo, García and Fernández [**11**] an implementation of the string averaging method to electron microscopy is reported. Butnariu, Davidi, Herman and Kazantsev [**18**] call a certain class of string averaging methods the Amalgamated Projection Method and show its stable behavior under summable perturbations. In Rhee [**83**] a string averaging scheme is applied to a problem in approximation theory.

3.3. The block-iterative algorithmic scheme with underrelaxed Bregman projections. In this subsection we briefly review the *block-iterative algorithmic scheme* with underrelaxed Bregman projections for the solution of the convex feasibility problem proposed by Censor and Herman [**39**]. By *block-iterative* we mean that, at the k-th iteration, the next iterate x^{k+1} is generated from the current iterate x^k by using a subset (called a block) of the family of sets $\{C_i\}_{i=1}^m$ of the convex feasibility problem [**47**, Section 1.1.3]. We use the term *algorithmic scheme* to emphasize that different specific algorithms may be derived by different choices of Bregman functions, and by various block structures. For example, if all blocks consist of a single set C_i, then our scheme gives rise to a sequential *row-action* [**26**] type algorithm. Taking the other extreme, if we let every block contain all sets, then we obtain a fully simultaneous algorithm. Such a block-iterative scheme for the convex feasibility problem was first proposed by Aharoni and Censor [**1**], using orthogonal projections onto convex sets. That block-iterative projections (BIP) method generalizes the sequential POCS method. The block-iterative scheme, described below, extends Aharoni and Censor's BIP method by employing underrelaxed Bregman projections which contain the underrelaxed orthogonal projections as a special case. The underrelaxed Bregman projection with Bregman function f and relaxation parameter $\lambda \in [0,1]$ of a point z onto a closed convex set Ω, denoted by $P^f_{\Omega,\lambda}(z)$, is given by

$$\nabla f(P^f_{\Omega,\lambda}(z)) = (1-\lambda)\nabla f(z) + \lambda \nabla f(P^f_{\Omega}(z)). \tag{3.11}$$

Appealing to the definition of a convex combination with respect to a Bregman function f, as defined by Censor and Reich [**44**, Definiton 4.1], the natural formula for a block-iterative step using underrelaxed Bregman projections is

$$\nabla f(x^{k+1}) = \sum_{i=1}^{m} v_i^k \nabla f(P^f_{C_i,\lambda_i^k}(x^k)), \tag{3.12}$$

where x^k is the k-th iterate, $\lambda_i^k \in [0,1]$ is the relaxation parameter used in the underrelaxed Bregman projection onto the set C_i during the k-th iterative step and the v_i^k are the weights of the convex combination for the k-th iterative step (i.e., $v_i^k \geq 0$ for $1 \leq i \leq m$ and $\sum_{i=1}^m v_i^k = 1$). Substituting (3.11) into (3.12), defining $w_i^k := v_i^k \lambda_i^k$, for $1 \leq i \leq m$, and introducing

$$w_{m+1}^k := 1 - \sum_{i=1}^{m} w_i^k \ \text{ and } \ C_{m+1} := R^n, \tag{3.13}$$

we get the following alternative formulation of the block-iterative step (3.12)

$$\nabla f(x^{k+1}) = \sum_{i=1}^{m+1} w_i^k \nabla f(P^f_{C_i}(x^k)), \tag{3.14}$$

with $w_i^k \geq 0$ for $1 \leq i \leq m+1$ and $\sum_{i=1}^{m+1} w_i^k = 1$. The block-iterative nature of this scheme stems from the fact that for every iteration index k some of the parameters w_i^k can be set to zero. The set of those indices i for which $w_i^k \neq 0$ at the k-th iteration defines the "block" of active constraints at this iteration. These index sets might vary dynamically from iteration to iteration as long as some technical conditions are observed [**39**].

Many other block-iterative algorithms were studied by Byrne [**20, 21, 22, 23**] in reference to image reconstruction from projections, where such algorithmic schemes are sometimes termed *ordered subset methods.* A rich source is Byrne's recent book [**25**]. See also the work of Combettes [**53**] and Section 6 of his paper on quasi-Fejérian methods [**54**].

4. Component averaging

In [**37**] a *CAV (Component averaging)* method for solving systems of linear equations was introduced. In these methods the sparsity of the matrix is explicitly used when constructing the iteration formula. Using this new scaling, considerable improvement was observed compared to traditionally scaled iteration methods.

In Cimmino's simultaneous projections method [**49**], see also, e.g., Censor and Zenios [**47**, Algorithm 5.6.1] with relaxation parameters and with equal weights $w_i = 1/m$, the next iterate x^{k+1} is the average of the orthogonal projections of x^k onto the hyperplanes H_i defined by the i-th row of the linear system $Ax = b$ and has, for every component $j = 1, 2, \dots, n$, the form

$$x_j^{k+1} = x_j^k + \frac{\lambda_k}{m} \sum_{i=1}^{m} \frac{b_i - \langle a^i, x^k \rangle}{\|a^i\|_2^2} a_j^i \,, \tag{4.1}$$

where a^i is the i-th column of the transpose A^T of A and b_i is the i-th component of the vector b and λ_k are relaxation parameters. When the $m \times n$ system matrix $A = (a_j^i)$ is sparse, only a relatively small number of the elements $\{a_j^1, a_j^2, \dots, a_j^m\}$ of the j-th column of A are nonzero, but in (4.1) the sum of their contributions is divided by the relatively large m. This observation led to the replacement of the factor $1/m$ in (4.1) by a factor that depends only on the *nonzero* elements in the set $\{a_j^1, a_j^2, \dots, a_j^m\}$. For each $j = 1, 2, \dots, n$, denote by s_j the number of nonzero elements of column j of the matrix A, and replace (4.1) by

$$x_j^{k+1} = x_j^k + \frac{\lambda_k}{s_j} \sum_{i=1}^{m} \frac{b_i - \langle a^i, x^k \rangle}{\|a^i\|_2^2} a_j^i \,. \tag{4.2}$$

Certainly, if A is sparse then the s_j values will be much smaller than m. The iterative step (4.1) is a special case of

$$x^{k+1} = x^k + \lambda_k \sum_{i=1}^{m} w_i \frac{b_i - \langle a^i, x^k \rangle}{\|a^i\|_2^2} a^i \,, \tag{4.3}$$

where the fixed weights $\{w_i\}_{i=1}^m$ must be positive for all i and $\sum_{i=1}^m w_i = 1$. The attempt to use $1/s_j$ as weights in (4.2) does not fit into the scheme (4.3), unless one can prove convergence of the iterates of a fully simultaneous iterative scheme with component-dependent (i.e., j-dependent) weights of the form

$$x_j^{k+1} = x_j^k + \lambda_k \sum_{i=1}^{m} w_{ij} \frac{b_i - \langle a^i, x^k \rangle}{\|a^i\|_2^2} a_j^i \,, \tag{4.4}$$

for all $j = 1, 2, \dots, n$. To formalize this consider a set $\{G_i\}_{i=1}^m$ of real diagonal $n \times n$ matrices $G_i = \mathrm{diag}(g_{i1}, g_{i2}, \dots, g_{in})$ with $g_{ij} \geq 0$, for all $i = 1, 2, \dots, m$ and $j = 1, 2, \dots, n$, such that $\sum_{i=1}^m G_i = I$, where

I is the unit matrix. Referring to the sparsity pattern of A one needs the following definition [**37**].

DEFINITION 4.1. A family $\{G_i\}_{i=1}^m$ of real diagonal $n \times n$ matrices with all diagonal elements $g_{ij} \geq 0$ and such that $\sum_{i=1}^m G_i = I$ is called *sparsity pattern oriented* (SPO, for short) *with respect to an* $m \times n$ *matrix* A if, for every $i = 1, 2, \ldots, m$, $g_{ij} = 0$ if and only if $a_j^i = 0$.

The Component Averaging (CAV) algorithm combines three features: (*i*) Each orthogonal projection onto H_i in is replaced by a *generalized oblique projection with respect to* G_i, denoted below by $P_{H_i}^{G_i}$. (*ii*) The scalar weights $\{w_i\}$ in (4.3) are replaced by the diagonal weighting matrices $\{G_i\}$. (*iii*) The actual weights are set to be inversely proportional to the number of nonzero elements in each column, as motivated by the discussion preceding Equation (4.2). The iterative step resulting from the first two features has the form

$$x^{k+1} \; = \; x^k + \lambda_k \sum_{i=1}^m G_i \left(P_{H_i}^{G_i}(x^k) - x^k \right). \tag{4.5}$$

Recent work by Censor, Elfving, Herman and Nikazad [**36**] shows that component averaging is valid (i.e., generates convergent iterative sequences) even when orthogonal projections are used and not generalized oblique ones as described above.

4.1. Seminorm-induced oblique projections for sparse nonlinear convex feasibility problems. The component averaging ideas can be extended to a convex feasibility problem with nonlinear convex sets. An attempt to answer this question was made in [**33**]. However, when applying seminorm-induced oblique projections in a simultaneous algorithmic scheme for general (not necessarily linear) convex sets, the approach used in [**33**] mandated a certain relationship between the matrix G and the (nonlinear) convex set Q onto which the seminorm-induced projection is made, namely, that the set will be *directionally affine with respect to* G, see [**33**, Definitions 2.3 and 2.4]. In spite of the actual generalization obtained in this way, its scope is limited due to this extra condition.

4.2. BICAV: Block-iterative component averaging. A recent member of the powerful family of block-iterative projection algorithms is the *block-iterative component averaging* (BICAV) algorithm of Censor, Gordon and Gordon [**38**] which was applied to a problem of image reconstruction from projections. The BICAV algorithm is a block-iterative companion to the [**37**].

The basic idea of the BICAV algorithm is to break up the system $Ax = b$ into "blocks" of equations and treat each block according to the CAV methodology, passing cyclically over all the blocks. Throughout the following, T will be the number of blocks and, for $t = 1, 2, \ldots, T$, let the block of indices $B_t \subseteq \{1, 2, \ldots, m\}$, be an ordered subset of the form $B_t = \{i_1^t, i_2^t, \ldots, i_{m(t)}^t\}$, where $m(t)$ is the number of elements in B_t, such that every element of $\{1, 2, \ldots, m\}$ appears in at least one of the sets B_t. For $t = 1, 2, \ldots, T$, let A_t denote the matrix formed by taking all the rows of A whose indices belong to the block of indices B_t, i.e.,

$$A_t := \begin{pmatrix} a^{i_1^t} \\ a^{i_2^t} \\ \vdots \\ a^{i_{m(t)}^t} \end{pmatrix}, \ t = 1, 2, \ldots, T. \tag{4.6}$$

The iterative step of the BICAV algorithm, developed and experimentally tested in [**38**], uses, for every block index $t = 1, 2, \ldots, T$, generalized oblique projections with respect to a family $\{G_i^t\}_{i=1}^m$ of diagonal matrices which are SPO with respect to A_t. The same family is also used to perform the diagonal weighting. The resulting iterative step has the form

$$x^{k+1} \ = \ x^k + \lambda_k \sum_{i \in B_{t(k)}} G_i^{t(k)} \left(P_{H_i}^{G_i^{t(k)}}(x^k) - x^k \right), \tag{4.7}$$

where $\{t(k)\}_{k \geq 0}$ is a *control sequence* according to which the $t(k)$-th block is chosen by the algorithm to be acted upon at the k-th iteration, thus, $1 \leq t(k) \leq T$, for all $k \geq 0$. The real numbers $\{\lambda_k\}_{k \geq 0}$ are user-chosen *relaxation parameters*. Finally, in order to achieve the acceleration, the diagonal matrices $\{G_i^t\}_{i=1}^m$ are constructed with respect to each A_t. Let s_j^t be the number of nonzero elements $a_j^i \neq 0$ in the j-th column of A_t and define

$$g_{ij}^t := \begin{cases} \dfrac{1}{s_j^t}\,, & \text{if } a_j^i \neq 0, \\ \\ 0\,, & \text{if } a_j^i = 0. \end{cases} \tag{4.8}$$

It is easy to verify that, for each $t = 1, 2, \ldots, T$, $\sum_{i=1}^m G_i^t = I$ holds for these matrices. With these particular SPO families of G_i^t's one obtains the following block-iterative algorithm:

ALGORITHM 4.2. ***BICAV***
Initialization: *$x^0 \in R^n$ is arbitrary.*

Iterative Step: *Given x^k, compute x^{k+1} by using, for $j = 1, 2, \ldots, n$, the formula:*

$$x_j^{k+1} \; = \; x_j^k + \lambda_k \sum_{i \in B_{t(k)}} \frac{b_i - \langle a^i, x^k \rangle}{\sum_{l=1}^n s_l^{t(k)} (a_l^i)^2} a_j^i \, , \tag{4.9}$$

where λ_k are relaxation parameters, $\{s_l^t\}_{l=1}^n$ are as defined above, and the control sequence is cyclic, i.e., $t(k) = k \bmod T + 1$, for all $k \geq 0$.

Full mathematical analysis of these methods, as well as their companion algorithms for linear inequalities, was presented by Censor and Elfving [**32**] and by Jiang and Wang [**72**]. Our recent [**36**] extends this by presenting the diagonally-relaxed orthogonal projections (DROP) algorithmic scheme. DROP is a block-iterative scheme which allows component averaging without having to resort to sparsity pattern oriented oblique projections $P_{H_i}^{G_i^{t(k)}}$ mentioned above.

5. Subgradient projections and perturbed projections for the multiple-sets split feasibility problem

In this section we review the *multiple-sets split feasibility problem* (1.3) that 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. 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 [**31**] and was called the *split feasibility problem* (SFP). They used their simultaneous multiprojections algorithm (see also [**47**]) to obtain iterative algorithms for the SFP. Their algorithms, as well as others, see, e.g., Byrne [**23**], involve matrix inversion at each iterative step, which is time-consuming, particularly if the dimensions are large. Therefore, Byrne [**24**] devised the CQ-algorithm with the iterative step:

$$x^{k+1} = P_C \left(x^k + \gamma A^T (P_Q - I)(Ax^k) \right), \tag{5.1}$$

where x^k and x^{k+1} are the current and the next iteration vectors, respectively, $\gamma \in (0, 2/\lambda)$ where λ is the spectral radius (in our case, 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 two-sets-SFP, for any starting vector $x^0 \in R^n$, whenever the two-sets-SFP has a solution. When the two-sets-SFP has no solutions, the CQ-algorithm converges to a minimizer of $\|P_Q(Ax) - Ax\|$ over all $x \in C$, whenever such a minimizer exists. The multiple-sets split feasibility problem, posed and studied in [**35**], was handled, for both the feasible and the infeasible cases, with a proximity function minimization approach where the proximity function $p(x)$ is

$$p(x) = (1/2) \sum_{i=1}^{t} \alpha_i \|P_{C_i}(x) - x\|^2 + (1/2) \sum_{j=1}^{r} \beta_j \left\|P_{Q_j}(Ax) - Ax\right\|^2, \tag{5.2}$$

For convenience reasons yet another set was introduced as follows.

DEFINITION 5.1. [**35**] Given an additional closed convex set $\Omega \subseteq R^n$, the *constrained multiple-sets split feasibility problem* (CMSSFP) is to find an $x^* \in \Omega$ such that x^* solves (1.3).

If the CMSSFP problem is consistent then unconstrained minimization of the proximity function yields the value 0, otherwise, in the inconsistent case, it finds a point which is least violating the feasibility by being "closest" to all sets, as "measured" by the proximity function.

ALGORITHM 5.2. [**35**, Algorithm 1]
Initialization: *Let x^0 be arbitrary.*
Iterative step: *For $k \geq 0$ let*

$$x^{k+1} = P_\Omega \left(x^k + \gamma \left(\sum_{i=1}^{t} \alpha_i \left(P_{C_i}(x^k) - x^k \right) + \sum_{j=1}^{r} \beta_j A^T \left(P_{Q_j}(Ax^k) - Ax^k \right) \right) \right), \tag{5.3}$$

where $\gamma \in (0, 2/L)$, $L = \sum_{i=1}^{t} \alpha_i + \lambda \sum_{j=1}^{r} \beta_j$ and λ is the spectral radius of the matrix $A^T A$.

5.1. A subgradient projection method. In some cases, notably when the convex sets are not linear, the exact computation of the orthogonal projections calls for the solution of a separate optimization problem for each projection. In such cases the efficiency of methods that use orthogonal projections is seriously reduced. Yang [**90**] 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 [**43**] 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

$$C_i = \{x \in R^n \mid c_i(x) \leq 0\} \text{ and } Q_j = \{y \in R^m \mid q_j(y) \leq 0\}, \quad (5.4)$$

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

ALGORITHM 5.3.
Initialization: *Let x^0 be arbitrary.*
Iterative step: *For $k \geq 0$ let*

$$x^{k+1} = x^k + \gamma \left(\sum_{i=1}^{t} \alpha_i \left(P_{C_{i,k}}(x^k) - x^k \right) + \sum_{j=1}^{r} \beta_j A^T \left(P_{Q_{j,k}}(Ax^k) - Ax^k \right) \right). \quad (5.5)$$

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

$$C_{i,k} = \left\{ x \in R^n \mid c_i(x^k) + \left\langle \xi^{i,k}, x - x^k \right\rangle \leq 0 \right\}, \quad (5.6)$$

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

$$Q_{j,k} = \left\{ x \in R^m \mid q_j(x^k) + \left\langle \eta^{j,k}, y - Ax^k \right\rangle \leq 0 \right\}, \quad (5.7)$$

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

5.2. A perturbed projection method. In this subsection we survey a perturbed projection method for the multiple-sets split feasibility problem. This method [**43**] is based on Santos and Scheimberg [**84**] 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 then 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**]). This notion was also used in [**42**].

DEFINITION 5.4. 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 $C_k \overset{M}{\to} C$, 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

$$N(x) := P_{\Omega}\left\{x + s\left(\sum_{i=1}^{t} \alpha_i (P_{C_i}(x) - x)\right.\right.$$
$$\left.\left. + \sum_{j=1}^{r} \beta_j A^T (P_{Q_j}(Ax) - Ax)\right)\right\}, \tag{5.8}$$

$$N_k(x) := P_{\Omega_k}\left\{x + s\left(\sum_{i=1}^{t} \alpha_i (P_{C_{i,k}}(x) - x)\right.\right.$$
$$\left.\left. + \sum_{j=1}^{r} \beta_j A^T (P_{Q_{j,k}}(Ax) - Ax)\right)\right\}, \tag{5.9}$$

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

$$\sum_{k=0}^{\infty} \varepsilon_k (1 - \varepsilon_k) = +\infty. \tag{5.10}$$

Then the following algorithm for the CMSSFP generates, under reasonable conditions [**43**], convergent iteration sequences.

ALGORITHM 5.5. ***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*

$$x^{k+1} = (1 - \varepsilon_k) x^k + \varepsilon_k N_k(x^k), \tag{5.11}$$

where N_k and ε_k are as defined above.

6. Algorithms for the quasiconvex feasibility problem

Since a quasiconvex feasibility problem (QFP) is a generalization of the convex feasibility problem, it is natural to ask whether the algorithmic schemes used for solution of the convex feasibility problem can be utilized for solving a QFP. In [**45**] Censor and Segal investigated the possibilities of modifying and adapting some of these algorithmic schemes so that they become applicable to the QFP. In particular, the following algorithmic schemes were considered: the *cyclic subgradient projections* (CSP) (Censor and Lent [**41**]), *parallel subgradient projections* (PSP) (Santos [**85, 86**]) and *Eremin's algorithmic scheme* [**64**]. The common idea of all these algorithms is to employ projections of different types, with respect to the individual level sets of the functions, to generate a sequence of points that converges to a solution. When the functions on the left-hand side of the inequalities are quasiconvex the situation is more complicated because such functions lack separation properties that convex functions have. Given a function f and a point z, the subdifferential of f at z is defined by

$$\partial f(z) := \{t \in R^n \mid \langle t, x - z\rangle \leq f(x) - f(z), \text{ for all } x \in R^n\}. \quad (6.1)$$

Sometimes it is called the *Fenchel-Moreau (FM) subdifferential*. Straightforward generalizations of the aforementioned algorithms are not possible because the subdifferential of Fenchel-Moreau might be empty at some points, thus, inapplicable to quasiconvex functions.

We first recall the notion of a quasiconvex function.

DEFINITION 6.1. Let $f : C \to R$, where C is a nonempty convex set in R^n. The function f is said to be *quasiconvex* if, for all $x, y \in C$, the following inequality holds

$$f(\theta x + (1-\theta)y) \leq \max\{f(x), f(y)\}, \text{ for all } \theta \in (0,1). \quad (6.2)$$

Quasiconvexity has a geometrical interpretation. For any $a \in R$ the *level* (respectively, *strict level*) *set of* f, *corresponding to* a, is the set

$$lev_f(a) = \{x \in R^n \mid f(x) \leq a\}, \quad (6.3)$$

respectively,

$$lev_f^<(a) = \{x \in R^n \mid f(x) < a\}. \quad (6.4)$$

Indeed f is quasiconvex if and only if its level sets $lev_f(a)$ are convex for all $a \in R$ which, in turn, is true if and only if its strict level sets $lev_f^<(a)$ are convex for all $a \in R$. Convex functions have convex level sets, and, therefore, are quasiconvex, but the converse is not true (e.g., the function $\log x$ on $(0, +\infty)$). Applications of quasiconvex functions which are not convex can be found in approximation theory (fractional programming), see, e.g., Bajona-Xandri and Martinez-Legaz [**2**], Boncompte and Martinez-Legaz [**12**], Stancu-Minasian [**88**], location theory, see, e.g., Gromicho [**67**], microeconomic theory (utility functions), see, e.g., Mas-Colell, Whinston and Green [**79**].

For generalization of gradient methods to nondifferentiable quasiconvex functions a broader notion than the FM-subdifferential is needed because the FM-subdifferential might be empty even for a differentiable nonconvex function on R^n, e.g., the real-valued single variable function $y = x^3$ at $x = 0$. For functions that are not convex, concave and are not differentiable, several notions of subdifferentials have been proposed in the literature. In the last thirty years there have been several attempts to define an appropriate notion of subdifferential for quasiconvex functions. One of them that is used in [**45**] is the star-subdifferential which is defined as follows.

DEFINITION 6.2. Given a function f and a point z, the-*star subdifferential of* f *at* z, is defined by

$$\partial^\star f(z) := \begin{cases} \{t \in R^n \backslash \{0\} \mid \langle t, x-z \rangle > 0 \Longrightarrow f(x) \geq f(z)\}, & z \notin \Gamma, \\ R^n, & z \in \Gamma, \end{cases} \tag{6.5}$$

where Γ is the set of minimizers of f.

If f is quasiconvex on R^n and finite at z, then $\partial^\star f(z) \neq \emptyset$, see, e.g., the review paper of Penot [**81**, Proposition 22]. Note that (6.5) is equivalent to

$$\partial^\star f(z) = \{t \in R^n \backslash \{0\} \mid f(x) < f(z) \Longrightarrow \langle t, x-z \rangle \leq 0\}. \tag{6.6}$$

Plastria [**82**] introduced and explored properties of his lower subdifferential.

DEFINITION 6.3. Given a function f and a point z, the *Plastria (P) lower subdifferential of* f *at* z (denoted in [**82**] as $\partial^- f$), is defined by

$$\partial^P f(z) = \{t \in R^n \mid f(x) < f(z) \Longrightarrow \langle x-z, t \rangle \leq f(x) - f(z)\}. \tag{6.7}$$

A function f is called *lower subdifferentiable* (*lsd*) on $K \subseteq R^n$ if it admits at least one P-lower subgradient at each point. Every convex function is *lsd*, since $\partial f(z) \subseteq \partial f^P(z)$, but not conversely.

Consider a family of sets

$$D_i = \{x \in R^n \mid f_i(x) \leq 0\} \text{ for } i = 1, 2, \ldots, m, \tag{6.8}$$

where all f_i are continuous and quasiconvex and let

$$D = \cap_{i=1}^m D_i \tag{6.9}$$

represent a quasiconvex feasibility problem. The algorithms presented in [**45**] deal with quasiconvex functions satisfying a Hölder condition.

DEFINITION 6.4. A function $f : R^n \to R$ is said to satisfy the *Hölder condition with degree β at a point z on a set $C \subseteq R^n$* if there exists a number $L < \infty$ and a $\beta \in (0, 1]$ such that

$$|f(y) - f(z)| \leq L \|y - z\|^{\beta}, \quad \text{for all } y \in C. \tag{6.10}$$

A Hölder condition can be verified by estimating the growth behavior of a function. Note that if a function satisfies a Hölder condition then it is uniformly continuous and, therefore, continuous. The Hölder condition with degree 1 is called the *Lipschitz condition*.

Denote by $g^+(x)$ the positive part $g^+(x) := \max\{0, g(x)\}$. Next an iterative algorithm for solving the QFP is presented. Denote by $S(0,1) = \{z \in R^n \mid \|z\| = 1\}$ the unit sphere.

ALGORITHM 6.5.
Initialization: *$x^0 \in R^n$ is arbitrary.*
Iterative step: *Given the current iterate x^k, calculate the next iterate x^{k+1} by*

$$x^{k+1} = x^k - \lambda_k \left(\frac{f^+_{i(k)}(x^k)}{L_{i(k)}} \right)^{1/\beta_{i(k)}} t^k, \tag{6.11}$$

where $t^k \in S(0,1) \cap \partial^\star f_{i(k)}(x^k)$ and $\beta_{i(k)}$ and $L_{i(k)}$ are the Hölder constant and degree, respectively, of $f_{i(k)}$.
Relaxation parameters: *$\{\lambda_k\}_{k=0}^{\infty}$ are confined to the interval $\varepsilon_1 \leq \lambda_k \leq 2 - \varepsilon_2$, for all $k \geq 0$, with some arbitrarily small $\varepsilon_1, \varepsilon_2 > 0$.*
Control: *Most violated constraint control or almost cyclic control (see, e.g.,* [**47**, Definition 5.1.1]*).*

The convergence of this algorithm is secured by following theorem.

THEOREM 6.6. [**45**] *Let the following assumptions hold: (i) the functions $f_i(x)$ are quasiconvex on R^n, (ii) the problem (6.9) is consistent, i.e., $D \neq \emptyset$, and (iii) the functions f_i satisfy, for every i, Hölder conditions with constants L_i and degrees β_i, for all $x \in D$, respectively, on R^n. Under these assumptions any sequence $\{x^k\}_{k=0}^{\infty}$, generated by Algorithm 6.5, converges to a solution of the problem* (6.9).

A companion parallel algorithm for solving the QFP can be formulated as follows.

ALGORITHM 6.7.
Initialization: *$x^0 \in R^n$ is arbitrary.*

Iterative step: *Given the current iterate x^k, calculate the next iterate x^{k+1} by*

$$x^{k+1} = x^k - \lambda_k \sum_{i=1}^{m} \alpha_i \left(\frac{f_i^+(x^k)}{L_i} \right)^{1/\beta_i} t^{i,k}, \tag{6.12}$$

where $t^{i,k} \in S(0,1) \cap \partial^\star f_i(x^k)$, and $0 < \alpha_i < 1$, for all i, and $\sum_{i=1}^{m} \alpha_i = 1$. The β_i and L_i are the Hölder constants and degrees, respectively, of f_i.

Relaxation parameters: *$\{\lambda_k\}_{k=0}^{\infty}$ are confined to the interval $\varepsilon_1 \le \lambda_k \le 2 - \varepsilon_2$, for all $k \ge 0$ with some arbitrary small $\varepsilon_1, \varepsilon_2 > 0$.*

THEOREM 6.8. [**45**] *Under the assumptions of Theorem 6.6, any sequence $\{x^k\}_{k=0}^{\infty}$, generated by Algorithm 6.7, converges to a solution of the problem (6.9).*

6.1. Algorithms for solving systems of inequalities with quasiconvex Lipschitz continuous functions in the left-hand side. Eremin's algorithms for the convex feasibility problem can be generalized for solving systems of inequalities with quasiconvex Lipschitz continuous functions $\{f_i\}_{i=1}^{m}$ on the left-hand side. Assume that $\{K_i\}_{i=1}^{m}$ is a set of real positive numbers and let $I(x) = \{j \mid \max\{f_i(x) \mid i = 1, 2, \ldots, m\} = f_j(x)\}$ and $s(x) = \{i \mid f_i(x) > 0\}$. The following definition was given by Eremin [**64**].

DEFINITION 6.9. Let $D \subseteq R^n$ be a closed convex set, let $d(x)$ be a continuous real-valued function, defined on R^n, that satisfies $\{x \mid d(x) \le 0\} = D$. Let $e(x)$ be a vector-valued function that is defined and nowhere equal to zero on $R^n \backslash D$. Assume also that $e(x)$ is bounded on any bounded set. Such a pair of functions $d(x)$ and $e(x)$ is said to have the *d-e property* if for arbitrary $z \notin D$ the half-space

$$\Omega = \{x \in R^n \mid \langle e(z), x - z\rangle + d(z) \le 0\} \tag{6.13}$$

contains D.

ALGORITHM 6.10. (***Eremin's algorithmic scheme***)
Initialization: *$x^0 \in R^n$ is arbitrary.*
Iterative step: *Given x^k, calculate the next iterate x^{k+1} from*

$$x^{k+1} = \begin{cases} x^k - \lambda_k \dfrac{d(x^k)}{\| e(x^k) \|^2} e(x^k), & \text{if } d(x^k) > 0, \\ x^k, & \text{if } d(x^k) \le 0, \end{cases} \tag{6.14}$$

where the pair $d(x)$ and $e(x)$ are user-chosen functions that have the d-e property.

Relaxation parameters: *$\{\lambda_k\}_{k=0}^{\infty}$ are confined to the interval $\varepsilon_1 \leq \lambda_k \leq 2-\varepsilon_2$, for all $k \geq 0$ with some arbitrary small $\varepsilon_1, \varepsilon_2 > 0$.*

While Eremin discussed this algorithmic scheme only for convex and differentiable functions, in [**45**] the scope of convergence was extended, as the following theorem shows.

THEOREM 6.11. *Let the following assumptions hold*

(i) the functions $f_i(x)$ are quasiconvex and Lipschitz continuous with Lipschitz constants L_i on R^n, for all $i \in \{1,2,\ldots,m\}$,

(ii) the problem (6.9) is consistent, i.e., $D \neq \emptyset$,

Then any sequence $\{x^k\}_{k=0}^{\infty}$, generated by Algorithm 6.10, converges to a point $x^* \in D$, if the pairs of functions $d(x)$ and $e(x)$ are chosen by one of the following methods.

Method 1:

$$d(x) = f_j(x) \ \text{ and } \ e(x) = L_j \frac{t^j}{\parallel t^j \parallel}, \tag{6.15}$$

where $t^j \in \partial^P f_j(x)$ and j is any index from $I(x)$.

Method 2:

$$d(x) = \begin{cases} \sum_{i\in s(x)} K_i f_i(x), & \text{if } s(x) \neq \emptyset, \\ 0, & \text{if } s(x) = \emptyset, \end{cases} \tag{6.16}$$

and

$$e(x) = \sum_{i\in s(x)} K_i L_i \frac{t^i}{\parallel t^i \parallel}, \tag{6.17}$$

where $t^i \in \partial^P f_i(x)$ for all $i \in \{1,2,\ldots,m\}$.

Method 3:

$$d(x) = \begin{cases} \sum_{i\in s(x)} f_i^2(x), & \text{if } s(x) \neq \emptyset, \\ 0, & \text{if } s(x) = \emptyset, \end{cases} \tag{6.18}$$

and

$$e(x) = \sum_{i\in s(x)} L_i f_i(x) \frac{t^i}{\parallel t^i \parallel}, \tag{6.19}$$

where $t^i \in \partial^P f_i(x)$ for all $i \in \{1,2,\ldots,m\}$.

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] C. Bajona-Xandri and J.E. Martinez-Legaz, *Lower subdifferentiability in minimax fractional programming*, Optimization, **45** (1999), 1–12.

[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 and J.M. Borwein, *Legendre functions and the method of random Bregman projections*, Journal of Convex Analysis **4 (**1997), 27–67.

[6] 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.

[7] 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.

[8] H.H. Bauschke, P.L. Combettes and S.G. Kruk, *Extrapolation algorithm for affine-convex feasibility problems*, Numerical Algorithms, to appear (available at: http://people.ok.ubc.ca/bauschke/Research/38.pdf).

[9] 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.

[10] 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.

[11] 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.

[12] M. Boncompte and J.E. Martinez-Legaz, *Fractional programming by lower subdifferentiability techniques*, Journal of Optimization Theory and Applications **68** (1991), 95–116.

[13] L.M. Bregman, *The method of successive projections for finding a common point of convex sets*, Soviet Mathematics Doklady **6** (1965), 688–692.

[14] 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 (**1967), 200–217.

[15] L.M. Bregman, Y. Censor and S. Reich, *Dykstra's algorithm as the nonlinear extension of Bregman's optimization method*, Journal of Convex Analysis **6** (1999), 319–334.

[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, Y. Censor and S. Reich, *Iterative averaging of entropic projections for solving stochastic convex feasibility problems*, Computational Optimization and Applications **8 (**1997), 21–39.

[18] D. Butnariu, R. Davidi, G.T. Herman and I.G. Kazantsev, *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.

[19] D. Butnariu and A.N. Iusem, *Totally Convex Functions for Fixed Points Computation and Infinite Dimensional Optimization*, Kluwer Academic Publishers, Dordrecht, the Netherlands, 2000.

[20] C.L. Byrne, *Block-iterative methods for image reconstruction from projections*, IEEE Transactions on Image Processing **IP-5 (**1996), 792–794.

[21] C.L. Byrne, *Block-iterative interior point optimization methods for image reconstruction from limited data*, Inverse Problems **16 (**2000), 1405–1419.

[22] C.L. Byrne, *Notes on block-iterative or ordered subset methods for image reconstruction*, Technical Report, September 2000.

[23] 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.

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

[25] C.L. Byrne, *Applied Iterative Methods*, A.K. Peters Ltd., Wellesley, MA, USA, 2008.

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

[27] 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.

[28] Y. Censor, *On sequential and parallel projection algorithms for feasibility and optimization* (Keynote Paper), in: Visualization and Optimization Techniques, Y. Censor and M. Ding, eds., Proceedings of SPIE Volume 4553, SPIE-The Society for Optical Engineering, Bellingham, WA, USA, 2001, pp. 1–9.

[29] 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.

[30] 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.

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

[32] 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.

[33] Y. Censor and T. Elfving, *Iterative algorithms with seminorm-induced oblique projections*, Abstract and Applied Analysis **8** (2003), 387–406.

[34] Y. Censor, T. Elfving and 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, eds., Elsevier Science Publishers, Amsterdam, The Netherlands, 2001, pp. 101–114.

[35] 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.

[36] Y. Censor, T. Elfving, G.T. Herman and T. Nikazad, *On diagonally-relaxed orthogonal projection methods*, SIAM Journal on Scientific Computing **30** (2008), 473–504.

[37] 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.

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

[39] Y. Censor and G.T. Herman, *Block-iterative algorithms with underrelaxed Bregman projections*, SIAM Journal on Optimizations **13** (2002), 283–297.

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

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

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

[43] Y. Censor, A. Motova, A. Segal, *Perturbed Projections and Subgradient Projections for the Multiple-Sets Split Feasibility Problem*, Journal of Mathematical Analysis and Applications **327** (2007), 1244–1256.

[44] Y. Censor and S. Reich, *Iterations of paracontractions and firmly nonexpansive operators with applications to feasibility and optimization*, Optimization **37 (**1996), 323–339.

[45] Y. Censor and A. Segal, *Algorithms for the quasiconvex feasibility problem*, Journal for Computational and Applied Mathematics **185** (2006), 34–50.

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

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

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

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

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

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

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

[53] P.L. Combettes, *Strong convergence of block-iterative outer approximation methods for convex optimization*, SIAM Journal on Control and Optimization **38 (**2000), 538–565.

[54] 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, eds., Elsevier Science Publishers, Amsterdam, The Netherlands, 2001, pp. 115–152.

[55] 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.

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

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

[58] 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.

[59] I. Csiszár and G. Tusnády, *Information geometry and alternating minimization procedures,* Statistics and Decisions Supplement Issue **1** (1984), 205–237.

[60] A.R. De Pierro and A.N. Iusem, *A relaxed version of Bregman's method for convex programming*, Journal of Optimization Theory and Applications **51 (**1986**)**, 421–440.

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

[62] 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.

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

[64] I.I. Eremin, *On systems of inequalities with convex functions in the left sides*, American Mathematical Society Translations, **88** (1970), 67–83.

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

[66] J.-L. Goffin, Z.Q. Luo and Y. Ye, *On the complexity of a column generation algorithm for convex or quasiconvex feasibility problems*, in: Large Scale Optimization: State of the Art, W.W Hager, D.W. Hearn, and P.M. Pardalos, eds., Kluwer Academic Publishers, Dordrecht, The Netherlands, 1994, pp. 182–191.

[67] J. Gromicho, *Quasiconvex Optimization and Location Theory*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1998.

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

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

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

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

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

[73] K. Kiwiel, *Free–steering relaxation methods for problems with strictly convex costs and linear constraints*, Mathematics of Operations Research **22 (**1997), 326–349.

[74] K. Kiwiel, *Generalized Bregman projections in convex feasibility problems*, Journal of Optimization Theory and Applications **96 (**1998), 139–157.

[75] T. Kotzer, N. Cohen and J. Shamir, *A projection-based algorithm for consistent and inconsistent constraints*, SIAM Journal on Optimization **7 (**1997), 527–546.

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

[77] D. Lyszyk and J. Shamir, *Signal processing under uncertain conditions by parallel projections onto fuzzy sets*, Journal of the Optical Society of America **A16 (**1999), 1602–1611.

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

[79] A. Mas-Colell, M.D. Whinston and J.R. Green, *Microeconomic Theory*, Oxford University Press, New York, NY, USA, 1995.

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

[81] J.-P. Penot, *Are generalized derivatives useful for generalized convex functions?*, in: Generalized convexity, Generalized Monotonicity: Recent Results, J.-P. Crouzeix, J.E. Martinez-Legaz and M. Volle, eds., Kluwer Academic Publishers, Dordrecht, The Netherlands, 1998, pp. 3–59.

[82] F. Plastria, *Lower subdifferentiable functions and their minimization by cutting planes*, Journal of Optimization Theory and Applications **46** (1985), 37–53.

[83] 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.

[84] 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.

[85] L.T. Santos, *Metodos de Projecao do Subgradiente para o Problema de Factibilidade Convexa*, Tese de Mestrado, IMECC-UNICAMP, State University of Campinas, Campinas, SP, Brazil, 1985.

[86] L.T. Santos, *A parallel subgradient projections method for the convex feasibility problem*, Journal of Computational and Applied Mathematics **18** (1987), 307–320.

[87] 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.

[88] I.M. Stancu-Minasian, *Fractional Programming: Theory, Methods, and Applications*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1997.

[89] 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.

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

[91] D.C. Youla, *Mathematical theory of image restoration by the method of convex projections*, in: Image Recovery: Theory and Applications, H. Stark, ed., Academic Press, Orlando, Florida, USA, 1987, pp. 29–77.