Perturbed Projections and Subgradient Projections for the Multiple-Sets Split Feasibility Problem

Yair Censor (yair@math.haifa.ac.il)
Avi Motova (avimotova@walla.com)
and
Alexander Segal (asegal@math.haifa.ac.il)

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

March 15, 2006. Revised: May 1, 2006.

Abstract

We study the multiple-sets split feasibility problem 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. By casting the problem into an equivalent problem in a suitable product space we are able to present a simultaneous subgradients projections algorithm that generates convergent sequences of iterates in the feasible case. We further derive and analyze a perturbed projection method for the multiple-sets split feasibility problem and, additionally, furnish alternative proofs to two known results.

1 Introduction

1.1 The multiple-sets split feasibility problem

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

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

Such MSSFPs, formulated in [14], 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 [12]. In the present paper we (i) cast the MSSFP into an equivalent problem in a suitable product space, (ii) formulate a simultaneous subgradient algorithm for solving the MSSFP and study its convergence, and (iii) propose a perturbed projection algorithm for the MSSFP. En route we give alternative proofs of two earlier results, shading further light on the problem. The real-world application of IMRT inspired us to investigate this problem but we present in this report only theoretical results of algorithmic developments and convergence theorems. Our experimental computational work in [12], which involves nonlinear convex sets, shows the practical viability of this class of algorithms.

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 [13] and was called the *split feasibility problem* (SFP). They used their simultaneous multiprojections algorithm (see also [17, Subsection 5.9.2]) to obtain iterative algorithms for the SFP. Their algorithms, as well as others, see, e.g., Byrne [5], involve matrix inversion at each iterative step, which is time-consuming, particularly if the dimensions are large. Therefore, Byrne [6] devised the CQ-algorithm with the iterative step:

$$x^{k+1} = P_C \left(x^k + \gamma A^T (P_Q - I)(Ax^k) \right), \tag{2}$$

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 MSSFP, posed and studied in [14], was handled, for both the feasible and the infeasible cases, with a proximity function minimization approach. If the MSSFP 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.

In Section 2 we formulate a simultaneous subgradient projections algorithm for the MSSFP. Such projections are actually not projections onto the convex sets of the problem but onto half-spaces determined by the subgradient of the function that defines the convex set, calculated at the current (available) iterate. The algorithm is inherently parallel, and hence, suitable for implementation on multiple-processors machines. The use of such subgradient projections to replace projections onto convex sets in various projection algorithms was done before by several authors, see Censor and Lent [16], Bauschke and Borwein [2, Section 7] and references therein, and Yang [28] for the two-sets-SFP. We analyze the algorithm in a suitable product space framework. This same framework enables also an alternative proof for the convergence theorem in [14].

In Section 3 we formulate a perturbed projections algorithm for the MSSFP that allows to do orthogonal projections onto a sequence of supersets of the original sets of the problem instead of projections onto the latter. This development is based on results of Santos and Scheimberg [25] and includes earlier findings of Zhao and Yang [31] as a special case.

Finally, in the Appendix at the end of the paper, we supply an alternative proof, based on properties of averaged operators, of Yang's convergence result [28, Theorem 1]. We moved this proof to the appendix because it currently contains a nonempty-interior assumption that is not present in Yang's result. We conjecture that this extra assumption can be removed – but do not know at this time how to do so. Additional recent developments on the split feasibility problem appear in Cegielski [9, 10], where the linear split feasibility problem is studied, in Qu and Xiu [23] who modify algorithms by adopting Armijo-like searches, in Yang [29] where algorithms that do not depend on the calculation of the spectral radius of the matrix $A^T A$ are derived, in Yang and Zhao [30] where the algorithms for the SFP are shown to be special instances of more general algorithms designed to solve a variational inequalities problem (VIP), and in Byrne and Censor [8].

1.2 Projection methods and their advantage

The reason why the MSSFP is looked at from the viewpoint of projection methods can be appreciated from 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 that is 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. 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 [26], Bauschke and Borwein [2] and Censor and Zenios [17].

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 [11]) while simultaneous algorithmic structures favor parallel computing platforms, see, e.g., Censor, Gordon and Gordon [15]. 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 works of Crombez [18, 19], the connection with variational inequalities, see, e.g., Aslam Noor [21], Yamada's [27] 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 [3] and references therein.

2 A simultaneous subgradient algorithm for the MSSFP

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 [28] 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. We use a product space formulation of the MSSFP. Assume, without loss of generality, that the sets C_i and Q_j are expressed as

$$C_i = \{x \in R^n \mid c_i(x) \leq 0\} \quad \text{and } Q_j = \{y \in R^m \mid q_j(y) \leq 0\}, \tag{3}$$

where $c_i : R^n \to R$, and $q_j : R^m \to R$ are convex functions for all $i = 1, 2, \dots, t$, and all $i = 1, 2, \dots, r$, respectively. For convenience reasons only we introduce yet another set as follows.

Definition 1 *[14] 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).*

We define the product spaces $\mathbf{V} = R^n$ and $\mathbf{W} = R^S$, where $S = tn + rm$ with r, t, n and m as in the CMSSFP and adopt the notational convention that the product spaces and all objects in them are represented in boldface type. Define the product set

$$\mathbf{Q} := \left(\prod_{i=1}^{t} \sqrt{\alpha_i} C_i \right) \times \left(\prod_{j=1}^{r} \sqrt{\beta_j} Q_j \right), \tag{4}$$

and the block-matrix

$$\mathbf{A} := \left(\sqrt{\alpha_1} I, \sqrt{\alpha_2} I, \ldots, \sqrt{\alpha_t} I, \sqrt{\beta_1} A^T, \sqrt{\beta_2} A^T, \ldots, \sqrt{\beta_r} A^T \right)^T, \tag{5}$$

where $\alpha_i > 0$, for $i = 1, 2, \ldots, t$, and $\beta_j > 0$, for $j = 1, 2, \ldots, r$, are arbitrary. This yields a two-sets split feasibility problem, with the sets $\Omega \subseteq \mathbf{V}$ and $\mathbf{Q} \subseteq \mathbf{W}$ and the matrix $\mathbf{A}$, whose solution solves the original CMSSFP. We represent the norm in $\mathbf{W}$ by $|||\cdot|||$, meaning that if $\mathbf{w} \in \mathbf{W}$ has the form $\mathbf{w} = (y^1, y^2, \ldots, y^t, z^1, z^2, \ldots, z^r)$ then $|||\mathbf{w}|||^2 = \sum_{i=1}^{t} \|y^i\|^2 + \sum_{j=1}^{r} \|z^j\|^2$. Projections in the product space $\mathbf{W}$ can be calculated by the following lemma.

Lemma 2 *Let $\mathbf{M} = \Pi_{l=1}^{s} M_l$ be a product, of s convex subsets of R^n, in a product space $\mathbf{U} = R^{ns}$ and let $\mathbf{y} \in \mathbf{U}$ have the form $\mathbf{y} = (y^1, y^2, \ldots, y^s)$. Then*

$$\mathbf{P_M}(\mathbf{y}) = (P_{M_1}(y^1), P_{M_2}(y^2), \ldots, P_{M_s}(y^s)). \tag{6}$$

Proof. See Pierra [22, Lemma 1.1(i)] or [17, Lemma 5.9.2]. ■

For the relaxed CQ-algorithm the sets C and Q in the two-sets split feasibility problem are given by

$$C = \{x \in R^n \mid c(x) \leq 0\} \quad \text{and} \quad Q = \{x \in R^m \mid q(x) \leq 0\}, \tag{7}$$

where $c : R^n \to R$ and $q : R^m \to R$ are convex functions, whose subdifferential sets are denoted by ∂c and ∂q, respectively.

Algorithm 3 *[28]* ***The relaxed CQ-algorithm***

Initialization: *Let x^0 be arbitrary.*

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

$$x^{k+1} = P_{C_k} \left(x^k + \gamma A^T (P_{Q_k} - I)(Ax^k) \right). \tag{8}$$

Here $\gamma \in (0, 2/\lambda)$, where λ is the spectral radius of $A^T A$,

$$C_k = \left\{ x \in R^n \mid c(x^k) + \left\langle \xi^k, x - x^k \right\rangle \leq 0 \right\}, \tag{9}$$

where ξ^k is a subgradient of c at the point x^k, i.e., $\xi^k \in \partial c(x^k)$, and

$$Q_k = \left\{ x \in R^m \mid q(x^k) + \left\langle \eta^k, y - Ax^k \right\rangle \leq 0 \right\}, \tag{10}$$

where $\eta^k \in \partial q(Ax^k)$.

The following convergence result was established by Yang.

Theorem 4 *[28, Theorem 1] If the solution set of the [two-sets] SFP is nonempty then any sequence $\left\{x^k\right\}_{k=0}^{\infty}$, generated by Algorithm 3, converges to a solution of the SFP.*

Applying Algorithm 3 to the two-sets split feasibility problem in the product space setting with $C = R^n$ and $\mathbf{Q}$ of (4), for sets given as in (3), we obtain the following new simultaneous subgradient algorithm for the MSSFP.

Algorithm 5

Initialization: *Let x^0 be arbitrary.*

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

$$x^{k+1} = x^k + \gamma \left(\sum_{i=1}^{t} \alpha_i \left(P_{C_{i,k}}(x^k) - x^k \right) + \sum_{j=1}^{r} \beta_j A^T \left(P_{Q_{j,k}}(Ax^k) - Ax^k \right) \right). \tag{11}$$

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 $A^T A$, and

$$C_{i,k} = \left\{ x \in R^n \mid c_i(x^k) + \left\langle \xi^{i,k}, x - x^k \right\rangle \leq 0 \right\}, \tag{12}$$

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

$$Q_{j,k} = \left\{ x \in R^m \mid q_j(x^k) + \left\langle \eta^{j,k}, y - Ax^k \right\rangle \leq 0 \right\}, \tag{13}$$

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

Theorem 6 *If the MSSFP has a nonempty solution set then any sequence $\left\{x^k\right\}_{k=0}^{\infty}$, generated by Algorithm 5, converges to a solution of MSSFP.*

Proof. Applying Theorem 4 to the two-sets split feasibility problem in the product space setting with $C = R^n$ and $\mathbf{Q}$ of (4), for sets given as in (3), the proof follows. ■

Algorithm 3 can be applied to the MSSFP in the consistent case in a direct manner by defining

$$c(x) := \sup\{c_i(x) \mid i = 1, 2, \ldots, t\} \tag{14}$$

and

$$q(y) := \sup\{q_j(y) \mid j = 1, 2, \ldots, r\}. \tag{15}$$

But the resulting algorithm will be inferior to our Algorithm 5 because it will have slow practical convergence due to the need to compare $r + t$ constraint violations at each iterative step.

Next we make a different use of the product space formulation to derive an alternative proof of convergence for the projection algorithm presented in [14]. Applying the CQ-algorithm (2) to the two-sets split feasibility problem, with the sets $\Omega \subseteq \mathbf{V}$ and $\mathbf{Q} \subseteq \mathbf{W}$ and the matrix $\mathbf{A}$, we take an arbitrary $x^0 \in \mathbf{V}$ and use the iterative process

$$x^{k+1} = P_\Omega\left(x^k + \gamma \mathbf{A}^T(\mathbf{P_Q} - \mathbf{I})\mathbf{A}x^k\right),\ k \geq 0. \tag{16}$$

By Byrne's convergence theorem [6, Theorem 2.1], any sequence $\{x^k\}_{k=0}^\infty$, generated in this manner, converges to

$$\operatorname{argmin}\{(1/2)|||\mathbf{P_Q}(\mathbf{A}x) - \mathbf{A}x|||^2 \mid x \in \Omega\}, \tag{17}$$

assuming such a minimum exists. Translating the iterative step (16), using the relation

$$\mathbf{P_Q}(\mathbf{A}x) = \left(\sqrt{\alpha_1}P_{C_1}x, \ldots, \sqrt{\alpha_t}P_{C_t}x, \sqrt{\beta_1}P_{Q_1}(Ax), \ldots, \sqrt{\beta_r}P_{Q_r}(Ax)\right)^T, \tag{18}$$

which follows from Lemma 2, we rewrite (17) as

$$\operatorname{argmin}\{p(x) \mid x \in \Omega\} \tag{19}$$

where the proximity function $p(x)$ is

$$p(x) = (1/2)\sum_{i=1}^t \alpha_i \left\|P_{C_i}(x) - x\right\|^2 + (1/2)\sum_{j=1}^r \beta_j \left\|P_{Q_j}(Ax) - Ax\right\|^2, \tag{20}$$

and obtain the following algorithm.

Algorithm 7 *[14, Algorithm 1]*
Initialization: *Let x^0 be arbitrary.*
Iterative step: *For $k \geq 0$ let*

$$x^{k+1} = P_\Omega \left(x^k + \gamma \left(\sum_{i=1}^{t} \alpha_i \left(P_{C_i}(x^k) - x^k \right) + \sum_{j=1}^{r} \beta_j A^T \left(P_{Q_j}(Ax^k) - Ax^k \right) \right) \right), \tag{21}$$

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 $A^T A$.

Our alternative convergence proof, based on the formulations presented above, now follows.

Theorem 8 *[14, Theorem 3] Let C_i, $i = 1, 2, \ldots, t$ and Q_j, $j = 1, 2, \ldots, r$ be nonempty closed convex sets in R^n and R^m, respectively, let $\Omega \subseteq R^n$ be a nonempty closed convex set, and let A be an $m \times n$ real matrix. If α_i, $i = 1, 2, \ldots, t$ and β_j, $j = 1, 2, \ldots, r$ are some positive scalars, and $\gamma \in (0, 2/L)$ where $L = \sum_{i=1}^{t} \alpha_i + \lambda \sum_{j=1}^{r} \beta_j$ then any sequence $\{x^k\}_{k=0}^{\infty}$, generated by Algorithm 7, converges to a minimizer of the function (20) if such a minimizer exists.*

Proof. Applying Byrne's result [6, Theorem 2.1] to the problem (17) with the algorithm (16) the proof follows. ■

3 A perturbed projection method

In this section we derive a perturbed projection method for the MSSFP. Our work is based on Santos and Scheimberg [25] 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. Here we devise an algorithm for the CMSSFP, based on Algorithm 7. We will need the following definitions.

Definition 9 *(i) Let N and $\{N_k\}_{k=0}^{\infty}$, be operators on R^n. If for all $x \in R^n$, $\lim_{k\to\infty} \|N_k(x) - N(x)\| = 0$ then we say that $\{N_k\}_{k=0}^{\infty}$ converges to N.*

(ii) Given any $\rho \geq 0$, the ρ-distance between two operators N_1 and N_2 on R^n is defined by

$$D_\rho(N_1, N_2) := \sup\{\|N_1(x) - N_2(x)\| \mid \|x\| \leq \rho\}. \tag{22}$$

The following notion of convergence of sequences of sets in R^n is called *Mosco-convergence* (see, e.g., [2]).

Definition 10 *Let C and $\{C_k\}_{k=0}^\infty$ be a subset and a sequence of subsets of R^n, respectively. The sequence $\{C_k\}_{k=0}^\infty$ is said to be* `Mosco-convergent` *to C, denoted by $C_k \xrightarrow{M} C$, if*

(i) for every $x \in C$, there exists a sequence $\{x^k\}_{k=0}^\infty$ with $x^k \in C_k$ for all $k = 0, 1, 2, \ldots,$ such that, $\lim_{k\to\infty} x^k = x$, and

(ii) for every subsequence $\left\{x^{k_j}\right\}_{j=0}^\infty$ 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 C and C_k, for $k = 0, 1, 2, \ldots,$ belong to *NCCS(R^n).* If the sequence $\{C_k\}_{k=0}^\infty$ converges to C in the Mosco sense, then the sequence of projections $\{P_{C_k}\}_{k=0}^\infty$ converges to P_C (see, e.g., [2, Lemma 4.2]).

Definition 11 *Let C_1 and C_2 belong to NCCS(R^n). The ρ-distance between C_1 and C_2 is defined by*

$$d_\rho(C_1, C_2) := \sup\{\|P_{C_1}(x) - P_{C_2}(x)\| \mid \|x\| \leq \rho\}. \tag{23}$$

The following theorem which generalizes the Krasnoselskii–Mann (KM) theorem (see, e.g., [31, Theorem 2.1]) is necessary for our convergence analysis.

Theorem 12 *Let N and N_k, for $k = 0, 1, 2, \ldots,$ be nonexpansive operators on a Hilbert space H, with $\lim_{k\to\infty} N_k = N$ and let $\{\varepsilon_k\}_{k=0}^\infty$ be a sequence in $(0, 1)$ satisfying*

$$\sum_{k=0}^\infty \varepsilon_k(1 - \varepsilon_k) = +\infty. \tag{24}$$

Then the sequence $\left\{x^k\right\}_{k=0}^\infty$, defined by the iterative step

$$x^{k+1} = (1 - \varepsilon_k)x^k + \varepsilon_k N_k(x^k) \tag{25}$$

converges weakly to a fixed point of N, provided that $\sum_{k=0}^{\infty} \varepsilon_k D_\rho(N_k, N) < +\infty$ for any $\rho > 0$, whenever such fixed points exist.

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

$$N(x) := P_\Omega \left\{ x + s \left(\sum_{i=1}^{t} \alpha_i (P_{C_i}(x) - x) + \sum_{j=1}^{r} \beta_j A^T (P_{Q_j}(Ax) - Ax) \right) \right\}, \tag{26}$$

$$N_k(x) := P_{\Omega_k} \left\{ x + s \left(\sum_{i=1}^{t} \alpha_i (P_{C_{i,k}}(x) - x) + \sum_{j=1}^{r} \beta_j A^T (P_{Q_{j,k}}(Ax) - Ax) \right) \right\}. \tag{27}$$

From [14, Theorem 2] we know that the operator

$$\sum_{i=1}^{t} \alpha_i (P_{C_i} - I) + \sum_{j=1}^{r} \beta_j A^T (P_{Q_j} - I) A \tag{28}$$

is Lipschitz continuous with Lipschitz constant $L = \sum_{i=1}^{t} \alpha_i + \lambda \sum_{j=1}^{r} \beta_j$, where λ is the spectral radius of $A^T A$. Therefore, it is ν-inverse strongly monotone (ν-*ism*) with $\nu = 1/L$ (see (40) in the Appendix), and so are the operators $\sum_{i=1}^{t} \alpha_i (P_{C_{i,k}} - I) + \sum_{j=1}^{r} \beta_j A^T (P_{Q_{j,k}} - I) A$, for $k = 0, 1, 2, \dots$. Combining these facts with [31, Proposition 2.1], and using Definition 16 in the Appendix, we obtain the following conclusion.

Lemma 13 *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 $i = 1, 2, \dots, t$ and Q_j and $Q_{j,k}$ be sets in NCCS(R^m), for $j = 1, 2, \dots, r$, such that $C_{i,k} \xrightarrow{M} C_i$, and $Q_{j,k} \xrightarrow{M} Q_j$ as $k \to \infty$. Then the operators N and N_k, defined in (26) and*

(27), are nonexpansive operators for $0 < s < 2/L$*, where* $L = \sum_{i=1}^{t} \alpha_i + \lambda \sum_{j=1}^{r} \beta_j$ *and* λ *is the spectral radius of* $A^T A$*. Moreover, the operator sequence* $\{N_k\}_{k=0}^{\infty}$ *converges to* N*.*

Algorithm 14 ***The perturbed projection algorithm for the 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*

$$x^{k+1} = (1 - \varepsilon_k)x^k + \varepsilon_k N_k(x^k), \tag{29}$$

where N_k *and* ε_k *are as defined above.*

The next theorem provides a convergence result for this algorithm.

Theorem 15 *If the assumptions of Lemma 13 are satisfied and* $\varepsilon_k \in (0, 1)$ *for* $k = 0, 1, 2, \dots$ *, then any sequence* $\left\{x^k\right\}_{k=0}^{\infty}$*, generated by Algorithm 14, converges to a fixed point of* N*, provided that such fixed point exists and that*

$$\sum_{k=0}^{\infty} \varepsilon_k \left\{ d_{\bar{\rho}}(\Omega_k, \Omega) + s \left(\sum_{i=1}^{t} \alpha_i d_{\bar{\rho}}(C_{i,k}, C_i) + \lambda^{1/2} \sum_{j=1}^{r} \beta_j d_{\bar{\rho}}(Q_{j,k}, Q_j) \right) \right\} < \infty \tag{30}$$

for any $\bar{\rho} > 0$*, and any* $\{\varepsilon_k\}_{k=0}^{\infty}$ *for which* $\sum_{k=0}^{\infty} \varepsilon_k(1 - \varepsilon_k) = +\infty$*.*

Proof. Denote

$$y^k := x + s \left(\sum_{i=1}^{t} \alpha_i (P_{C_{i,k}}(x) - x) + \sum_{j=1}^{r} \beta_j A^T (P_{Q_{j,k}}(Ax) - Ax) \right) \tag{31}$$

and

$$y := x + s \left(\sum_{i=1}^{t} \alpha_i (P_{C_i}(x) - x) + \sum_{j=1}^{r} \beta_j A^T (P_{Q_j}(Ax) - Ax) \right). \tag{32}$$

For any $x \in R^n$ with $\|x\| \leq \rho$, $\rho > 0$, we have

$$\begin{aligned} \|N_k(x) - N(x)\| &= \left\|P_{\Omega_k}(y^k) - P_\Omega(y)\right\| \\ &\leq \left\|P_{\Omega_k}(y^k) - P_{\Omega_k}(y)\right\| + \|P_{\Omega_k}(y) - P_\Omega(y)\|. \end{aligned} \tag{33}$$

By the nonexpansiveness of orthogonal projections, we obtain

$$\left\|P_{\Omega_k}(y^k) - P_{\Omega_k}(y)\right\| \leq s \left\|y^k - y\right\|. \tag{34}$$

Substituting (31) and (32) into (33) and using (34) we get

$$\begin{aligned}\|N_k(x) - N(x)\| \leq s & \left(\sum_{i=1}^{t} \alpha_i \left\|P_{C_{i,k}}(x) - P_{C_i}(x)\right\|\right. \\ & \left. + \sum_{j=1}^{r} \beta_j \left\|A^T(P_{Q_{j,k}}(Ax) - P_{Q_j}(Ax))\right\|\right) \\ & + \|P_{\Omega_k}(y) - P_{\Omega}(y)\|. \end{aligned} \tag{35}$$

Applying to the norm in the second summand above the well-known relation $\langle Bx, x\rangle \leq \rho(B)\|x\|^2$, which holds for any matrix B and its spectral radius $\rho(B)$, we obtain

$$\begin{aligned}\|N_k(x) - N(x)\| \leq s & \left(\sum_{i=1}^{t} \alpha_i \left\|P_{C_{i,k}}(x) - P_{C_i}(x)\right\|\right. \\ & \left. + \lambda^{1/2}\sum_{j=1}^{r} \beta_j \left\|P_{Q_{j,k}}(Ax) - P_{Q_j}(Ax)\right\|\right) \\ & + \|P_{\Omega_k}(y) - P_{\Omega}(y)\|. \end{aligned} \tag{36}$$

Finally, from (22) and (23) we obtain

$$D_\rho(N_k, N) \leq d_{\bar{\rho}}(\Omega_k, \Omega) + s\left(\sum_{i=1}^{t} \alpha_i d_{\bar{\rho}}(C_{i,k}, C_i) + \lambda^{1/2}\sum_{j=1}^{r} \beta_j d_{\bar{\rho}}(Q_{j,k}, Q_j)\right), \tag{37}$$

where $\bar{\rho} \geq \max(\|y\|, \|Ax\|, \|y\|)$. Therefore, from Theorem 12 and (30) the result follows. ■

The last theorem shows that any sequence generated by Algorithm 14 converges to a minimizer of the function (20) over the set Ω, provided that such minimizers exist, just as in the case of Algorithm 7. Zhao and Yang in [31] developed a perturbed projections method for the SFP based on the CQ-algorithm. Their method can be viewed as a special case of Algorithm 14.

Acknowledgments. This work was supported by grant No. 2003275 of the United States-Israel Binational Science Foundation (BSF), by a National Institutes of Health (NIH) grant No. HL70472 and by grant No. 522/04 of the Israel Science Foundation (ISF) and was partially done at the Center for Computational Mathematics and Scientific Computation (CCMSC) in the University of Haifa.

Appendix: Applying averaged operators

In this Appendix we furnish an alternative and shorter proof of Yang's convergence result [28, Theorem 1], that uses properties of averaged operators. To the best of our knowledge, the term "averaged mapping" to describe operators of the form $T = (1-\alpha)I + \alpha N$, was first used by Reich and co-workers [1, 4], see also Reich [24]. However, in this route we are forced to make an additional assumption on the two-sets-SFP that is not present in Yang's work and which we are unable to get rid of at this time. The additional assumption is that the solution set Θ of the two-sets-SFP must have a nonempty interior, i.e.,

$$\operatorname{int}\Theta := \operatorname{int}\{x \in C \mid Ax \in Q\} \neq \emptyset. \tag{38}$$

In the sequel we use definitions and results on averaged operators and their properties as they appear in Bauschke and Borwein [2] and in Byrne [7], which are also sources for references on the subject.

Definition 16 *An operator $N : R^n \to R^n$ is called* **`nonexpansive`** *(abbreviated, ne) if $\|N(x) - N(y)\| \leq \|x - y\|$, for all x and y in R^n.*

Definition 17 *Given an ne operator N, let $T := (1-\alpha)I + \alpha N$ for some $\alpha \in (0,1)$. The operator T is called* **`averaged operator`** *(abbreviated, av).*

Lemma 18 *If A and B are av then $T := AB$ is av.*

Any operator T is related to its complement $G = I - T$ by

$$\|x - y\|^2 - \|T(x) - T(y)\|^2 = 2\langle G(x) - G(y), x - y\rangle - \|G(x) - G(y)\|^2. \tag{39}$$

An operator G is called *ν−inverse strongly monotone (ν-ism)* (see, e.g., [20]) if there is a $\nu > 0$, such that,

$$\langle G(x) - G(y), x - y\rangle \geq \nu \|G(x) - G(y)\|^2. \tag{40}$$

From (39) follows that N is *ne* if and only if its complement $G = I - N$ is a $(1/2)$-*ism*. It is also true that if G is a ν-*ism* and $\gamma > 0$ than the operator γG is a (ν/γ)-*ism*.

Lemma 19 *[7, Lemma 2.1] An operator T is av if and only if its complement $G = I - T$ is a ν-ism for some $\nu > 1/2$.*

Theorem 20 *If (38) holds then any sequence $\{x^k\}_{k=0}^{\infty}$, generated by Algorithm 3, converges to a solution of the SFP.*

Proof. Define the following operators, which depend on the sequence $\{x^k\}_{k=0}^{\infty}$,

$$T_k(x) := P_{C_k}\left(x + \gamma A^T(P_{Q_k} - I)(Ax)\right). \tag{41}$$

From [7, Lemma 8.1] follows that if A is an $m \times n$ matrix whose spectral radius is λ and if Φ is a nonempty closed convex set, then for every $\gamma \in (0, 2/\lambda)$ the operator $I + \gamma A^T(P_\Phi - I)A$ is *av*. Therefore, Lemma 18 shows that every operator T_k is *av* with

$$\Theta \subseteq \operatorname{Fix}(T_k), \tag{42}$$

for all $k = 0, 1, \ldots$, where Fix stands for the fixed points set. Since T_k is *av*, there is an *ne* operator N_k such that, for some $\alpha \in (0,1)$, $T_k = (1-\alpha)I + \alpha N_k$. Taking a $z \in \Theta$ and using (42) we have $T_k(z) = z$. Denoting $G_k = I - T_k$ we get, from (39),

$$\begin{aligned} \left\|z - x^k\right\|^2 - \left\|T_k(z) - x^{k+1}\right\|^2 &= 2\left\langle G_k(z) - G_k(x^k), z - x^k\right\rangle \\ &\quad - \left\|G_k(z) - G_k(x^k)\right\|^2. \end{aligned} \tag{43}$$

Lemma 19 guarantees that G_k is a $(1/2)\alpha$-*ism*, thus, we have

$$\left\|z - x^k\right\|^2 - \left\|z - x^{k+1}\right\|^2 \geq ((1/\alpha) - 1)\left\|x^k - x^{k+1}\right\|^2. \tag{44}$$

Therefore, a sequence $\{x^k\}_{k=0}^{\infty}$, generated by Algorithm 3, is Fejér-monotone with respect to Θ. Then, from the assumption (38) and [2, Lemma 2.16] follows that $\{x^k\}_{k=0}^{\infty}$ converges to some $x^* \in R^n$.

To complete the proof, we show that $x^* \in \Theta$ by verifying that $c(x^*) = 0$ and $q(Ax^*) = 0$, where c and q are defined in (14) and (15), respectively. To see that $q(Ax^*) = 0$, rewrite (44) as

$$\left\|z - x^k\right\|^2 - \left\|z - x^{k+1}\right\|^2 \geq \left(\lambda\gamma^2 - 2\gamma\right) \left\|(P_{Q_k} - I)(Ax^k)\right\|^2. \tag{45}$$

The left-hand side of the last inequality tends to 0 as $k \to \infty$, hence

$$\lim_{k\to\infty} \left\|(P_{Q_k} - I)(Ax^k)\right\| = \lim_{k\to\infty} \frac{\left|q(Ax^k)\right|}{\left\|\eta^k\right\|} = 0. \tag{46}$$

The sequence $\left\{\left\|\eta^k\right\|\right\}_{k=0}^{\infty}$, where η^k are the subgradients defined in (10), is bounded by [2, Proposition 7.8], therefore, $q(Ax^*) = 0$. To see that $c(x^*) = 0$ denote

$$\varphi^k := \gamma A^T (P_{Q_k} - I)(Ax^k), \tag{47}$$

so that the iterative step (8) becomes

$$x^{k+1} = P_{C_k}\left(x^k + \varphi^k\right), \tag{48}$$

where C_k is the half-space (9). Calculating the projection we obtain

$$x^{k+1} = x^k + \varphi^k - \left(c(x^k + \varphi^k) + \left\langle \xi^k, \varphi^k \right\rangle\right) \left(\xi^k / \left\|\xi^k\right\|^2\right), \tag{49}$$

where ξ^k are the subgradients defined in (9). In (49) we have: (i) the sequence $\left\{\varphi^k\right\}_{k=0}^{\infty}$ tends to zero by (46) and (47), (ii) the Schwartz inequality shows that

$$\lim_{k\to\infty} \left\langle \xi^k, \varphi^k \right\rangle \left(\xi^k / \left\|\xi^k\right\|^2\right) = 0, \tag{50}$$

(iii) $\left\{x^k\right\}_{k=0}^{\infty}$ converges to x^*, as shown above, and (iv)

$$\lim_{k\to\infty} c(x^k + \varphi^k) = c(x^*). \tag{51}$$

Therefore, $c(x^*) = 0$. ■

References

[1] J.B. Baillon, R.E. Bruck and S. Reich, On the asymptotic behavior of nonexpansive mappings and semigroups in Banach spaces, *Houston Journal of Mathematics* **4** (1978), 1–9.

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

[3] 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).

[4] R.E. Bruck and S. Reich, Nonexpansive projections and resolvents of accretive operators in Banach spaces, *Houston Journal of Mathematics* **3** (1977), 459–470.

[5] C. Byrne, Bregman-Legendre multidistance projection algorithms for convex feasibility and optimization, in: D. Butnariu, Y. Censor and S. Reich (Editors), *Inherently Parallel Algorithms in Feasibility and Optimization and Their Applications*, Elsevier, Amsterdam, The Netherlands, 2001, pp. 87–100.

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

[7] C. Byrne, A unified treatment of some iterative algorithms in signal processing and image reconstruction, *Inverse Problems* **20** (2004), 103–120.

[8] 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** (2001), 77–98.

[9] A. Cegielski, Convergence of the projected surrogate constraints method for the linear split feasibility problems, Technical Report, September 14, 2004, Institute of Mathematics, University of Zielona Góra, Zielona Góra, Poland. *Journal of Convex Analysis* **14** (2007), to appear.

[10] A. Cegielski, Projection methods for the linear split feasibility problems, Technical Report, September 14, 2004, Institute of Mathematics, University of Zielona Góra, Zielona Góra, Poland.

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

[12] 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.

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

[14] 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.

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

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

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

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

[19] G. Crombez, A sequential iteration algorithm with non-monotoneous behaviour in the method of projections onto convex sets, *Czechoslovak Mathematical Journal*, to appear.

[20] E. Golshtein and N. Tretyakov, *Modified Lagrangians and Monotone Maps in Optimization*, John Wiley & Sons, New York, NY, USA, 1996.

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

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

[23] B. Qu and N. Xiu, A note on the CQ algorithm for the split feasibility problem, *Inverse Problems* **21** (2005), 1655–1665.

[24] S. Reich, Averaged mappings in the Hilbert ball, *Journal of Mathematical Analysis and Applications* **109** (1985), 199–206.

[25] P.S.M. Santos and S. Scheimberg, A projection algorithm for general variational inequalities with perturbed constraint sets, *Applied Mathematics and Computation* (2006), to appear.

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

[27] 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.

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

[29] Q. Yang, On variable-step relaxed projection algorithm for variational inequalities, *Journal of Mathematical Analysis and Applications* **302** (2005), 166–179.

[30] Q. Yang and J. Zhao, Generalized KM theorems and their applications, Technical Report, School of Mathematics and LPMC, Nankai University, 300071 Tianjin, P.R. China. Available at: http://www.math.nankai.edu.cn/keyan/pre/preprint06/06-21.pdf

[31] J. Zhao and Q. Yang, Several solution methods for the split feasibility problem, *Inverse Problems* **21** (2005), 1791–1799.