ITERATIVE ALGORITHMS WITH SEMINORM-INDUCED OBLIQUE PROJECTIONS

YAIR CENSOR AND TOMMY ELFVING

A definition of oblique projections onto closed convex sets that use seminorms induced by diagonal matrices which may have zeros on the diagonal is introduced. Existence and uniqueness of such projections are secured via directional affinity of the sets with respect to the diagonal matrices involved. A block-iterative algorithmic scheme for solving the convex feasibility problem, employing seminorm-induced oblique projections, is constructed and its convergence for the consistent case is established. The fully simultaneous algorithm converges also in the inconsistent case to the minimum of a certain proximity function.

1. Introduction

The motivation of this paper comes from the recent work of Censor et al. [14, 15] who used generalized oblique projections for the solution of large and sparse systems of linear equations arising in the fully discretized approach to the problem of image reconstruction from projections, see, for example, Herman [22] and Censor [11]. In order to achieve significant acceleration in the algorithm's behavior, they had to use generalized oblique projections (as they called them) *which allow zeros on the diagonals of the weighting matrices* for the fully simultaneous projections algorithm that they used.

Since the work of [14, 15] is limited to linear equations, it is natural to inquire whether it can be extended, and if so, in what directions and to what extent? Is it possible to extend the definition of the generalized oblique projections of [14, 15] to cover convex sets which are not linear (i.e., not hyperplanes or half-spaces)? If there is a way to do so, can one employ these projections in iterative algorithmic schemes for solving the convex (not necessarily linear) feasibility problem? Under what conditions would such algorithms converge? How would a fully simultaneous iterative algorithm, which employs these projections for

the convex feasibility problem with general (not necessarily linear) convex sets, behave in the inconsistent case (i.e., when $\cap_{i=1}^{m} C_i = \varnothing$)?

In this paper, we provide some theoretical answers to these questions. We introduce (in Section 2) a new definition of oblique projections onto convex sets that use seminorms induced by diagonal matrices which may have zeros on the diagonal. In order to guarantee existence and uniqueness of such seminorm-induced oblique projections, we need to impose a certain relationship between the seminorm inducing matrix G and the convex set C onto which we project, namely, that the set will be directionally affine with respect to G. By this, we mean that it must be affine in the direction of every standard basis vector e^j of $\mathbb{R}^n$ whenever there is a zero in the jth location on the diagonal of G. With these seminorm-induced oblique projections, we construct a block-iterative algorithmic scheme for the solution of the convex feasibility problem of finding a point in the intersection of a finite family of closed convex sets $\{C_i\}_{i=1}^{m}$. The main feature of the new block-iterative scheme is the use of diagonal matrices for weighting the different sets which are acted upon in each block iteration. Using diagonal weighting matrices, instead of the traditionally used scalar weights, amounts to allowing the weights to change not only with the set index i and the iteration index k but also with the component index j. The block-iterative algorithmic scheme allows the blocks (i.e., the subfamilies) of sets C_i, which are acted upon in each iterative step, to vary in size and composition as iterations proceed. We study the convergence of the block-iterative algorithmic scheme for the consistent case $\cap_{i=1}^{m} C_i \neq \varnothing$ in Section 3. In Section 4, we investigate the fully simultaneous algorithm with seminorm-induced projections without assuming consistency of the convex feasibility problem. The result there is that the fully simultaneous algorithm generates sequences which converge to the minimum of a proximity function which measures the sum of the squares of the "seminorm-induced distances" to all sets of the convex feasibility problem.

The theory developed here ties up in several ways with existing results about iterative projection algorithms for the convex feasibility problem, see, for example, Bauschke and Borwein [5], Combettes [19] and [16, Chapter 5] for review and tutorial texts in this field. See also Reich [28] where a treatment of simultaneous algorithms involving nearest point and other projections is presented. Beside generalizing the work of [14, 15], the block-iterative algorithm with seminorm-induced oblique projections generalizes our recent work in [13]. It is also a generalization of the block-iterative projections (BIP) method of Aharoni and Censor [1] (see [16, Algorithm 5.6.1]) which uses orthogonal (least Euclidean distance) projections onto convex sets. It also generalizes the work of Eggermont et al. [20] and the work of Elfving [21]. For an extensive survey of linear least squares algorithms, including projection methods, see Björck [7]. Since classical oblique projections are a special case of seminorm-induced oblique projections, our work properly generalizes earlier work on oblique projection iterative algorithms. For example, oblique projections have been used in the past in

several contexts. Kayalar and Weinert [24] promote oblique projections for local processing in sensor arrays and credit Murray [26] and Lorch [25] for pioneering work on oblique projections. Behrens and Scharf [6] use oblique projections for signal processing applications. Oblique projections onto hyperplanes in a fully simultaneous (Cimmino) algorithm have been proposed and used by Arioli et al. [2]. They showed that using a Cimmino algorithm, in which all projections are oblique with respect to a given symmetric positive definite matrix G for a system $Ax = b$, is equivalent to applying a Cimmino algorithm with all orthogonal projections to the postconditioned system $AG^{-1/2}\tilde{x} = b$, where $\tilde{x} = G^{1/2}x$ and $G^{1/2}$, the symmetric square root of G, is the unique symmetric positive definite matrix obtained from $G = G^{1/2}G^{1/2}$.

Our work is related to the block-iterative multiprojection successive generalized projection (BIMSGP) method developed by Byrne in [9]. Specifically, a special case of our algorithmic scheme (Algorithm 3.2), in which the diagonal weight matrices are not allowed to change with the iteration index and for which only fixed blocks of constraints are permissible, is derivable from [9, Algorithm 4.2]. We return to this point at the end of Section 3.

The fully simultaneous projection algorithm with orthogonal projections was first proposed by Cimmino [17] for linear systems of equations and then by Auslender [3] for general convex sets, see the review paper of Bauschke and Borwein [5]. For general convex sets and orthogonal projections, Iusem and De Pierro [23] proved local convergence in the inconsistent case, while Combettes [18] showed global convergence in the inconsistent case by employing a product space formulation which extends the one of Pierra [27] and can handle the inconsistent case. Our work in Section 4 generalizes these previous algorithmic schemes to the case of seminorm-induced oblique projections. Our convergence analysis in Section 4 might also be deduced, with some additional efforts, from results in [10], but we prefer to prove it here along the self-contained lines of the proof developed in [15]; thereby, making it independent of the theory of Bregman functions and other technical tools used in [10].

2. Seminorm-induced oblique projections

Let $G = \text{diag}(g_1, g_2, \ldots, g_n)$ be a diagonal $n \times n$ matrix with $g_j \geq 0$, for all $j = 1, 2, \ldots, n$, and $G \neq 0$, that is, at least one element of G is different from zero. Denote the index set of positive diagonal elements of G by $J = J(G) = \{j \mid g_j > 0, \ 1 \leq j \leq n\}$ and its cardinality by $|J|$. We consider the restriction $R^{|J|}$ of $\mathbb{R}^n$ to only those components that belong to J. For convenience and notational simplicity, assume, without loss of generality, that the zero entries on the diagonal of G (if any) all come last. If $z = (z_j) \in \mathbb{R}^n$, we denote its restriction to $R^{|J|}$ by $z_{[J]}$. For any nonempty closed convex set $C \subseteq \mathbb{R}^n$, the restriction of C to $R^{|J|}$ is the set

$$C_{[J]} := \{z_{[J]} \in R^{|J|} \mid z \in C\}. \tag{2.1}$$

Further, denote by $G_{[J]}$ the $|J| \times |J|$ matrix obtained from G by deleting all zeros from its diagonal. The function $|x|_G$ is defined, for every $x \in \mathbb{R}^n$, by

$$|x|_G^2 := \langle x, Gx \rangle, \tag{2.2}$$

where $\langle \cdot, \cdot \rangle$ is the standard inner product in $\mathbb{R}^n$. If $J = \{1, 2, \ldots, n\}$, then $|x|_G = \|x\|_G$ is the (classical) Hilbert space ellipsoidal norm. Observe that $|x|_G$ is a *vector seminorm* (see, e.g., Taylor [29]) because it may be equal to zero for $x \neq 0$ if G has at least one $g_j = 0$. For this reason, a projection onto a convex set with respect to a seminorm needs to be carefully defined. Our next definitions and proposition accommodate this possibility.

Definition 2.1 (seminorm-induced oblique projection). Let $C \subseteq \mathbb{R}^n$ be a closed convex set and let $G \neq 0$ be an $n \times n$ nonnegative diagonal matrix. For any $y \in \mathbb{R}^n$, the seminorm-induced oblique projection of y onto C with respect to G, denoted by $P_C^G(y)$, is defined as a point which has the following properties: (i) $P_C^G(y) \in C$, and (ii) $P_C^G(y) = y^*$, where

$$y_j^* = \begin{cases} z_j, & \text{if } j \in J, \\ y_j, & \text{if } j \notin J, \end{cases} \tag{2.3}$$

$$z_{[J]} = (z_j)_{j \in J} = \operatorname{argmin}\{\|x - y_{[J]}\|_{G_{[J]}}^2 \mid x \in C_{[J]}\}. \tag{2.4}$$

The presence of zeros on the diagonal of G whenever $j \in J$ implies that

$$\left(\operatorname{argmin}\left\{|x - y|_G^2 \mid x \in C\right\}\right)_{[J]} = z_{[J]}. \tag{2.5}$$

A seminorm-induced oblique projection reduces to an ellipsoidal oblique projection of y onto C if all diagonal elements of G are positive. This is a special case of a Bregman projection of y onto C according to the Bregman function $f(x) = \|x\|_G^2 = \langle x, Gx \rangle$, consult, for example, Censor and Zenios [16, Chapter 2] for background material and further references on Bregman functions, distances, and projections originating from Bregman [8].

On the other hand, a general seminorm-induced oblique projection onto a set $C \subseteq \mathbb{R}^n$ can be identified with an appropriately defined ellipsoidal oblique projection onto a certain *cylindrical set* generated from C, as we show next. Let G be a diagonal $n \times n$ nonnegative nonzero matrix with index set of positive diagonal elements J and $|J| < n$. Define the $n \times n$ diagonal positive matrix $\hat{G}$ by

$$\hat{g}_j = \begin{cases} g_j, & \text{if } j \in J, \\ 1, & \text{if } j \notin J, \end{cases} \tag{2.6}$$

and consider the cylindrical set $C_{[J]} \times R^{n \setminus |J|}$.

PROPOSITION 2.2. *Given C, G, J, and $\hat{G}$ as above and a point $y \in \mathbb{R}^n$, the following holds:*

$$y^* = P_C^G(y) = P_{C_{[J]} \times R^{n \setminus |J|}}^{\hat{G}}(y). \tag{2.7}$$

Proof. Denote the complement of J in $\{1, 2, \dots, n\}$ by $\bar{J}$ so that any $u \in \mathbb{R}^n$ can be represented as $u = (u_{[J]}, u_{[\bar{J}]})$. Then,

$$
\begin{aligned}
& P_{C_{[J]} \times R^{n \setminus |J|}}^{\hat{G}}(y) \\
&\quad = \operatorname{argmin} \left\{ \|u - y\|_{\hat{G}}^2 \mid u \in C_{[J]} \times R^{n \setminus |J|} \right\} \\
&\quad = \operatorname{argmin} \left\{ \|u_{[J]} - y_{[J]}\|_{G_{[J]}}^2 + \|u_{[\bar{J}]} - y_{[\bar{J}]}\|_I^2 \mid u_{[J]} \in C_{[J]},\ u_{[\bar{J}]} \in R^{n \setminus |J|} \right\},
\end{aligned}
\tag{2.8}
$$

where I is the $(n - |J|) \times (n - |J|)$ unit matrix. This concludes the proof in view of (2.3) and (2.4). $\qquad\square$

In order to secure existence and uniqueness of seminorm-induced oblique projections, we need to impose a "cylindricity" relationship between the seminorm inducing matrix G and the convex set C onto which we project, as we do in the next definitions.

Definition 2.3 (directional affinity of sets). A nonempty closed convex set $C \subseteq \mathbb{R}^n$ is said to be affine in the direction $d \in \mathbb{R}^n$ if (i) together with any two distinct points of the set whose difference vector is a scalar multiple of d, the line through these points belongs to the set, and (ii) there is at least one pair of points in the set which fulfills (i).

Definition 2.4 (directional affinity with respect to G). Let $C \subseteq \mathbb{R}^n$ be a nonempty closed convex set and let G be a nonnegative diagonal $n \times n$ matrix $G = \operatorname{diag}(g_1, g_2, \dots, g_n)$ with $g_j \geq 0$, for all $j = 1, 2, \dots, n$, and $G \neq 0$. If C is affine in the direction of every standard basis direction vector e^j of $\mathbb{R}^n$ for which $j \notin J$, then C is called directionally affine with respect to G.

PROPOSITION 2.5. *Let $C \subseteq \mathbb{R}^n$ be a nonempty closed convex set and let $G = \operatorname{diag}(g_1, g_2, \dots, g_n)$ be a nonnegative diagonal $n \times n$ matrix with $g_j \geq 0$, for all $j = 1, 2, \dots, n$, and $G \neq 0$. If C is directionally affine with respect to G, then for any $y \in \mathbb{R}^n$, there is a unique seminorm-induced oblique projection $P_C^G(y)$ of y onto C with respect to G.*

Proof. By standard results on (classical) oblique projections onto closed convex sets (use, e.g., Censor and Zenios [16, Lemma 2.1.2] since $\| \cdot \|_{G_{[J]}}^2$ is a Bregman function), it follows that there exists a unique $z_{[J]} \in C_{[J]}$ that solves (2.4). The directional affinity of C with respect to G allows us to define $y_j^* = y_j$, as we did in (2.3). $\qquad\square$

The next examples demonstrate these notions.

Example 2.6. Let the convex set be represented as $C = \{x \in \mathbb{R}^n \mid h(x) \leq 0\}$, where $h : \mathbb{R}^n \to R$ is a convex function, and assume that $\partial h / \partial x_j \equiv 0$, for all $j \notin J$. Then, $h(x)$ does not depend on the values that the jth component of x is taking as long as $j \notin J$ and Proposition 2.5, together with (2.3), guarantees that $P_C^G(y) \in C$.

Example 2.7. Let the convex set be a hyperplane $H = \{x \in \mathbb{R}^n \mid \langle a, x \rangle = b\}$ with $a = (a_j) \in \mathbb{R}^n$ and $b \in R$ given and assume that in the matrix G it is true that $g_j = 0$ if and only if $a_j = 0$. Then, there is a closed-form formula for the seminorm-induced oblique projection $P_H^G(y)$ of a point $y \in \mathbb{R}^n$ onto H with respect to G, given by Censor et al. in [15], which has the following form:

$$
\left(P_H^G(y)\right)_j := \begin{cases} y_j + \dfrac{b - \langle a, y \rangle}{\sum_{\substack{l=1 \\ g_l \neq 0}}^n (a_l^2 / g_l)} \cdot \dfrac{a_j}{g_j}, & \text{if } g_j \neq 0, \\[1em] y_j, & \text{if } g_j = 0. \end{cases}
\tag{2.9}
$$

It is not difficult to verify that $P_H^G(y) \in H$ and that it is the appropriate expression for the seminorm-induced oblique projection. This example is in fact a special case of the previous example. It has been analyzed and used in practice in [14, 15].

Postulating a relationship between the seminorm inducing matrix G and the convex set onto which we project, namely, that the set must be a *directionally affine set*, limits the scope of our theory of seminorm-induced oblique projections onto (general) convex sets. We do not know if it is possible to relax or lift this condition. In the linear case of Example 2.7, this condition was termed *sparsity pattern orientation* (SPO) by Censor et al. [15, Definition 3.1], and was taken as an advantage for creating a new accelerated fully simultaneous projection method called the *component averaging* (CAV) method.

The next proposition generalizes a classical result. We prove it here along the same lines as the proof of a similar result for Bregman generalized distances (see Bregman [8]) given in [16, Theorem 2.4.1].

PROPOSITION 2.8. *Let $Q \subseteq \mathbb{R}^n$ be a nonempty closed convex set and let $G \neq 0$ be a nonnegative diagonal matrix such that Q is diagonally affine with respect to G. If $z \in Q$ is any given point, then for any $y \in \mathbb{R}^n$, the following inequality holds:*

$$
\left| P_Q^G(y) - y \right|_G^2 \leq \left| z - y \right|_G^2 - \left| z - P_Q^G(y) \right|_G^2.
\tag{2.10}
$$

Proof. By expanding the function

$$
E(u) := \left| u - y \right|_G^2 - \left| u - P_Q^G(y) \right|_G^2,
\tag{2.11}
$$

according to (2.2), we find that $E(u) = \langle u, \alpha \rangle + \beta$ for some $\alpha \in \mathbb{R}^n$ and $\beta \in R$, which are independent of u; thus, $E(u)$ is convex. Denoting $u_\lambda = \lambda z + (1 - \lambda)P_Q^G(y)$, for any $0 \leq \lambda \leq 1$, we know that $u_\lambda \in Q$ by Proposition 2.5. Due to the convexity of $E(u)$, we then obtain (observe that due to the linearity of $E(u)$, the inequalities in the next two formulae are actually equalities)

$$E(u_\lambda) \leq \lambda \left(\left| z - y \right|_G^2 - \left| z - P_Q^G(y) \right|_G^2 \right) + (1 - \lambda) \left| P_Q^G(y) - y \right|_G^2. \qquad (2.12)$$

For $\lambda > 0$, this gives

$$\begin{aligned}
&\left| z - y \right|_G^2 - \left| z - P_Q^G(y) \right|_G^2 - \left| P_Q^G(y) - y \right|_G^2 \\
&\geq \frac{1}{\lambda} \left(\left| u_\lambda - y \right|_G^2 - \left| P_Q^G(y) - y \right|_G^2 \right) - \frac{1}{\lambda} \left| u_\lambda - P_Q^G(y) \right|_G^2.
\end{aligned} \qquad (2.13)$$

The first term on the right-hand side of (2.13) is nonnegative because of the minimization property (2.5) of $P_Q^G(y)$ and the second term tends to zero as $\lambda \to 0^+$. Therefore, for small enough positive values of λ, the right-hand side of (2.13) is nonnegative. $\qquad \square$

3. The block-iterative algorithmic scheme

Let $C_i \subseteq \mathbb{R}^n$, $i = 1, 2, \ldots, m$, be closed convex sets such that $C := \cap_{i=1}^m C_i \neq \varnothing$. For each $k = 0, 1, 2, \ldots$, we define a vector $\{G_i^k\}_{i=1}^m$, each of whose m components is an $n \times n$ diagonal matrix

$$G_i^k = \operatorname{diag}\left(g_{i1}^k, g_{i2}^k, \ldots, g_{in}^k\right) \qquad (3.1)$$

with $g_{ij}^k \geq 0$, for all $i = 1, 2, \ldots, m$, all $j = 1, 2, \ldots, n$, and all $k \geq 0$. Denote by $B(k)$ the index set

$$B(k) := \{i \mid 1 \leq i \leq m, \ G_i^k \neq 0\}. \qquad (3.2)$$

Definition 3.1 (a fair sequence of vectors of diagonal matrices). Let $\{\{G_i^k\}_{i=1}^m\}_{k \geq 0}$ be an infinite sequence of vectors of nonnegative diagonal $n \times n$ matrices.

(i) If there exists an $\epsilon > 0$ such that, for every $i \in B(k)$, the diagonal elements g_{ij}^k are either $g_{ij}^k = 0$ or $g_{ij}^k \geq \epsilon > 0$, for all $j = 1, 2, \ldots, n$, and for all $k \geq 0$; and

(ii) if $\sum_{i=1}^m G_i^k = I$, for all $k \geq 0$, where I denotes the unit matrix;

then $\{\{G_i^k\}_{i=1}^m\}_{k \geq 0}$ is called a fair sequence of vectors of diagonal matrices.

With regard to this definition, it should be noted that condition (i) is stronger than the condition used by Aharoni and Censor [1] regarding the weights they used in their BIP method. Their scalar weights in BIP can be obtained from the system $\{\{G_i^k\}_{i=1}^m\}_{k\geq 0}$, in Definition 3.1, by letting $G_i^k = w_i^k I$, where $w^k = (w_i^k)_{i=1}^m$ is their vector of weights used in the kth iteration. The (weaker) condition that they use is that "for all $i = 1, 2, \ldots, m$, the series $\sum_{k=0}^\infty w_i^k = +\infty$." The purpose of either condition (i) in Definition 3.1 or the condition of [1] is to guarantee that none of sets C_i is gradually ignored by ever diminishing weights. We do not know whether our convergence result, presented below, can be strengthened by using a condition similar to that of [1].

The algorithm that we propose and study here for solving the convex feasibility problem of finding a point $x^* \in C = \cap_{i=1}^m C_i$ is formulated as follows.

Algorithm 3.2. Initialization: $x^0 \in \mathbb{R}^n$ is arbitrary.
Iterative step: given x^k calculate

$$x^{k+1} = x^k + \lambda_k \left(\sum_{i=1}^m \left(G_i^k P_{C_i}^{G_i^k}(x^k) \right) - x^k \right), \qquad (3.3)$$

where $\{\lambda_k\}_{k\geq 0}$ are relaxation parameters, $\{\{G_i^k\}_{i=1}^m\}_{k\geq 0}$ is a sequence of vectors of diagonal matrices, and $P_{C_i}^{G_i^k}(x^k)$ is the seminorm-induced oblique projection of x^k onto C_i with respect to G_i^k.

This algorithmic scheme includes, as a special case, the sequential seminorm-induced oblique projections algorithm that is obtained from (3.3) by choosing at the kth iteration, for $i = 1, 2, \ldots, m$,

$$G_i^k = \begin{cases} G_{i(k)}^k \neq 0, & \text{if } i = i(k), \\ 0, & \text{if } i \neq i(k), \end{cases} \qquad (3.4)$$

where $\{i(k)\}_{k\geq 0}$ is a *control sequence* which governs the sequential progress such as a *cyclic control* in which $i(k) = k(\mod m) + 1$. If all G_i^k are nonzero matrices at every iterative step, then a fully simultaneous algorithm with seminorm-induced oblique projections is obtained from our algorithmic scheme. In addition to the consistent case result for such a fully simultaneous algorithm, we present in Section 4 also an inconsistent case analysis. In between these two extremes, the algorithmic scheme allows variable blocks of constraints to be acted upon, as iterations proceed. At the kth iteration, only those sets C_i for which $G_i^k \neq 0$ will be included in the block. The blocks may vary in size (i.e., number of sets included) and in composition (which sets are included) as long as the conditions of the convergence theorem, given below, are met.

We consider first the case of unity relaxation, that is, $\lambda_k = 1$, for all $k \geq 0$, and define the operators

$$T_k(z) := \sum_{i=1}^{m} G_i^k P_{C_i}^{G_i^k}(z) \tag{3.5}$$

so that (3.3) is now

$$x^{k+1} = T_k(x^k), \quad k \geq 0. \tag{3.6}$$

Our goal is to prove the following convergence result.

THEOREM 3.3. *If the following assumptions hold:*

(i) $C_i \subseteq \mathbb{R}^n$, $i = 1, 2, \ldots, m$, *are nonempty closed convex sets,*
(ii) *the convex feasibility problem is consistent, that is, $C = \cap_{i=1}^{m} C_i \neq \varnothing$,*
(iii) $\{\{G_i^k\}_{i=1}^{m}\}_{k \geq 0}$ *is a fair sequence of vectors of diagonal matrices,*
(iv) *every index $i = 1, 2, \ldots, m$ appears in infinitely many sets $B(k)$,*
(v) *for every $i = 1, 2, \ldots, m$, the set C_i is directionally affine with respect to G_i^k, for all $k \geq 0$,*

then any sequence $\{x^k\}_{k \geq 0}$, generated by Algorithm 3.2, with $\lambda_k = 1$ for all $k \geq 0$, converges to a point $x^* \in C$.

We define, for all $k \geq 0$, the functions

$$g_k(x) := \sum_{i=1}^{m} \left| P_{C_i}^{G_i^k}(x^k) - x \right|_{G_i^k}^2 \tag{3.7}$$

for which the following auxiliary result holds.

PROPOSITION 3.4. *Under the assumptions of Theorem 3.3, if $\{x^k\}_{k \geq 0}$ is any sequence, generated by Algorithm 3.2, with $\lambda_k = 1$, for all $k \geq 0$, then, for any $x \in \mathbb{R}^n$ and for every $k \geq 0$,*

$$g_k(x) = g_k(x^{k+1}) + \|x - x^{k+1}\|_2^2. \tag{3.8}$$

Proof. The proof follows the same lines as the proof of Lemma 4.1 in Censor et al. [15]. By (2.2), (3.6), and (3.7), we have

$$g_k(x) - g_k(x^{k+1}) = \sum_{i \in B(k)} \left(\left| P_{C_i}^{G_i^k}(x^k) - x \right|_{G_i^k}^2 - \left| P_{C_i}^{G_i^k}(x^k) - T_k(x^k) \right|_{G_i^k}^2 \right)$$

$$= \sum_{i \in B(k)} \left(\left| x \right|_{G_i^k}^2 - \left| T_k(x^k) \right|_{G_i^k}^2 + 2 \left\langle G_i^k P_{C_i}^{G_i^k}(x^k), T_k(x^k) - x \right\rangle \right). \tag{3.9}$$

Using (3.5) and the fact that, due to $\sum_{i=1}^{m} G_i^k = \sum_{i \in B(k)} G_i^k = I$, we have, for any $y \in \mathbb{R}^n$,

$$\sum_{i \in B(k)} \left| y \right|_{G_i^k}^2 = \|y\|_2^2, \tag{3.10}$$

the result follows.   □

Definition 3.5. A sequence $\{x^k\}_{k \geq 0}$ is called Fejér-monotone with respect to a nonempty set $\Omega \subseteq \mathbb{R}^n$ if, for any $x \in \Omega$,

$$\|x - x^{k+1}\|_2 \leq \|x - x^k\|_2, \quad \forall k \geq 0. \tag{3.11}$$

See, for example, [16, Definition 5.3.1]. It is easy to verify that any Fejér-monotone sequence is bounded. We have now all tools in place to prove the convergence theorem.

Proof of Theorem 3.3. From (3.8), we have, for every $k \geq 0$,

$$\|x - x^k\|_2^2 - \|x - x^{k+1}\|_2^2 = g_k(x^{k+1}) - g_k(x) + \|x - x^k\|_2^2. \tag{3.12}$$

Using (3.7) and (3.10), the right-hand side of (3.12) may be rewritten, for some $x = \hat{x} \in C$, as

$$\begin{aligned}
g_k(x^{k+1}) - g_k(\hat{x}) + \|\hat{x} - x^k\|_2^2 &= \sum_{i \in B(k)} \left| P_{C_i}^{G_i^k}(x^k) - x^{k+1} \right|_{G_i^k}^2 \\
&\quad + \sum_{i \in B(k)} \left(\left| \hat{x} - x^k \right|_{G_i^k}^2 - \left| \hat{x} - P_{C_i}^{G_i^k}(x^k) \right|_{G_i^k}^2 \right).
\end{aligned} \tag{3.13}$$

The first sum on the right-hand side of (3.13) is nonnegative and each term in the second sum of (3.13) is nonnegative by application of Proposition 2.8 with $G = G_i^k$, $Q = C_i$, $z = \hat{x}$, and $y = x^k$. Thus, we obtain from (3.12) that, for every $\hat{x} \in C$,

$$\|\hat{x} - x^{k+1}\|_2^2 \leq \|\hat{x} - x^k\|_2^2, \quad \forall k \geq 0, \tag{3.14}$$

that is, Fejér-monotonicity of $\{x^k\}_{k \geq 0}$ with respect to C which implies its boundedness and, by monotonicity and nonnegativity, we obtain that the limit

$$\lim_{k \to \infty} \|\hat{x} - x^k\|_2^2 = \theta \tag{3.15}$$

exists. Therefore, the left-hand side of (3.12) tends to zero, as $k \to \infty$, which implies, by (3.13), that

$$\lim_{k \to \infty} \left(\sum_{i \in B(k)} \left| P_{C_i}^{G_i^k}(x^k) - x^{k+1} \right|_{G_i^k}^2 \right) = 0 \tag{3.16}$$

and that, for any $\hat{x} \in C$,

$$\lim_{k \to \infty} \left(\sum_{i \in B(k)} \left(\left| \hat{x} - x^k \right|_{G_i^k}^2 - \left| \hat{x} - P_{C_i}^{G_i^k}(x^k) \right|_{G_i^k}^2 \right) \right) = 0. \tag{3.17}$$

These limits imply the next two limits. By nonnegativity of the seminorms, (3.16) implies that, for all $i = 1, 2, \ldots, m$,

$$\lim_{k \to \infty} \left| P_{C_i}^{G_i^k}(x^k) - x^{k+1} \right|_{G_i^k}^2 = 0, \tag{3.18}$$

and the nonnegativity of the summands in (3.17), noted earlier, implies that, for all $i = 1, 2, \ldots, m$ and any $\hat{x} \in C$,

$$\lim_{k \to \infty} \left(\left| \hat{x} - x^k \right|_{G_i^k}^2 - \left| \hat{x} - P_{C_i}^{G_i^k}(x^k) \right|_{G_i^k}^2 \right) = 0. \tag{3.19}$$

Since $\{x^k\}_{k \geq 0}$ is bounded and hence has a cluster point, the following two observations imply the required convergence result: (i) if there exists a cluster point x^* of $\{x^k\}_{k \geq 0}$ in C, then $\{x^k\}_{k \geq 0}$ has only one cluster point, and (ii) every cluster point of $\{x^k\}_{k \geq 0}$ must be in C.

To prove (i), let $x^* \in C$ be a cluster point and assume that x^{**} is another cluster point, that is,

$$\lim_{k \to \infty, k \in K_1} x^k = x^*, \qquad \lim_{k \to \infty, k \in K_2} x^k = x^{**}, \tag{3.20}$$

with $K_1 \subseteq N_0$, $K_2 \subseteq N_0$, and $N_0 = \{0, 1, 2, \ldots\}$. Since $x^* \in C$, (3.15) applies with $\hat{x}$ replaced by x^*, and (3.20) shows that

$$\lim_{k \to \infty} \|x^* - x^k\|_2^2 = 0. \tag{3.21}$$

Then, taking the limit of

$$0 \leq \|x^* - x^{**}\|_2 \leq \|x^* - x^k\|_2 + \|x^{**} - x^k\|_2, \tag{3.22}$$

for $k \to \infty$, $k \in K_2$, and using (3.20) and (3.21), for the second and first summands on the right-hand side of (3.22), respectively, implies that $x^* = x^{**}$.

Next, we prove (ii). Let $\lim_{l \to \infty} x^{k_l} = x^*$ and assume that $x^* \notin C$. Defining

$$
\begin{aligned}
I_{\text{in}} &:= \{i \mid 1 \leq i \leq m,\ x^* \in C_i\}, \\
I_{\text{out}} &:= \{i \mid 1 \leq i \leq m,\ x^* \notin C_i\},
\end{aligned}
\tag{3.23}
$$

this means that $I_{\text{out}} \neq \varnothing$. Since, by assumption, every index i, $1 \leq i \leq m$, appears in infinitely many index sets $B(k)$, we may assume, without loss of generality (passing to a subsequence if necessary), that for every $l = 1, 2, \ldots,$

$$
B(k_l) \cup B(k_l + 1) \cup \cdots \cup B(k_{l+1} - 1) = \{1, 2, \ldots, m\}.
\tag{3.24}
$$

For every l, $l = 1, 2, \ldots,$ let μ_l be the smallest element in the set

$$
\{k_l, k_l + 1, \ldots, k_{l+1} - 1\}
\tag{3.25}
$$

such that

$$
B(\mu_l) \cap I_{\text{out}} \neq \varnothing
\tag{3.26}
$$

(such an element exists by (3.24) and since $I_{\text{out}} \neq \varnothing$). We want to show that the subsequence $\{x^{\mu_l}\}_{l \geq 0}$ also converges to x^*. By definition, $k_l \leq \mu_l$, for all $l = 1, 2, \ldots$. If $k_l < \mu_l$, then

$$
B(s) \subseteq I_{\text{in}}, \quad \text{for } s = k_l, k_l + 1, \ldots, \mu_l - 1,
\tag{3.27}
$$

and since $x^* \in \cap_{i \in I_{\text{in}}} C_i$, we have, from an appropriate variant of (3.14),

$$
0 \leq \|x^* - x^{\mu_l}\|_2^2 \leq \cdots \leq \|x^* - x^{k_l+1}\|_2^2 \leq \|x^* - x^{k_l}\|_2^2.
\tag{3.28}
$$

Letting $l \to \infty$ in (3.28), we obtain $\lim_{l \to \infty} x^{\mu_l} = x^*$. From (3.26), it follows that there exists an index $t \in I_{\text{out}}$ such that $t \in B(\mu_l)$ for infinitely many indices l. Removing from the sequence $\{\mu_l\}_{l \geq 0}$ all elements μ_l for which $t \notin B(\mu_l)$, we end up with a new infinite sequence $\{\mu_l\}_{l \geq 0}$ such that

$$
t \in B(\mu_l) \cap I_{\text{out}}, \quad \forall l = 1, 2, \ldots.
\tag{3.29}
$$

Write (3.19) for that t, and use for it Proposition 2.8 with $G = G_t^k$, $Q = C_t$, $z = \hat{x}$, and $y = x^k$ to obtain

$$\lim_{k \to \infty} \left| P_{C_t}^{G_t^k}(x^k) - x^k \right|_{G_t^k}^2 = 0. \tag{3.30}$$

Pick a $G_t^k \neq 0$ (there must be infinitely many such nonzero matrices for any t), denote $w^{t,k} = P_{C_t}^{G_t^k}(x^k) - x^k$, and write

$$\left| w^{t,k} \right|_{G_t^k}^2 = \langle w^{t,k}, G_t^k w^{t,k} \rangle = \sum_{j=1}^n g_{tj}^k (w_j^{t,k})^2. \tag{3.31}$$

Let $J = J(G_t^k)$. By (2.2), Definitions 2.1 and 3.1, and (3.30), we obtain

$$\lim_{k \to \infty} \|w^{t,k}\|_2^2 = 0 \tag{3.32}$$

because when $g_{tj}^k = 0$ then, by Definition 2.1, $w_j^{t,k} = 0$ so that, using Definition 3.1,

$$\sum_{j=1}^n g_{tj}^k (w_j^{t,k})^2 \geq \epsilon \sum_{j \in J} (w_j^{t,k})^2 = \epsilon \|w^{t,k}\|_2^2. \tag{3.33}$$

Therefore, passing to the subsequence,

$$\lim_{l \to \infty} P_{C_t}^{G_t^{\mu_l}}(x^{\mu_l}) = x^* \tag{3.34}$$

which implies $x^* \in C_t$, contradicting (3.29). □

Once convergence of Algorithm 3.2 has been proved for unity relaxation, $\lambda_k = 1$, for all $k \geq 0$, it is possible to introduce *underrelaxation* parameters in the following manner.

THEOREM 3.6. *Under the assumptions of Theorem 3.3, any sequence $\{x^k\}_{k \geq 0}$, generated by Algorithm 3.2 with relaxation parameters*

$$0 < \epsilon \leq \lambda_k \leq 1, \quad \forall k \geq 0, \tag{3.35}$$

for some arbitrarily small but fixed ϵ, converges to a point $x^* \in C$.

Proof. The idea of this proof was developed during discussions with Charles Byrne. Define the $m + 1$ diagonal $n \times n$ matrices

$$\Gamma_i^k := \lambda_k G_i^k, \quad \forall i = 1, 2, \ldots, m, \ \forall k \geq 0,$$

$$\Gamma_{m+1}^k := I - \sum_{i=1}^m \Gamma_i^k, \quad \forall k \geq 0, \tag{3.36}$$

and let $C_{m+1} = \mathbb{R}^n$. Then (3.3) takes the form

$$x^{k+1} = \sum_{i=1}^{m+1} \Gamma_i^k P_{C_i}^{\Gamma_i^k}(x^k) \tag{3.37}$$

because, for any $z \in \mathbb{R}^n$,

$$P_{C_i}^{G_i^k}(z) = P_{C_i}^{\Gamma_i^k}(z) \tag{3.38}$$

since $|x|_{\lambda G} = \lambda |x|_G^2$. Also, the convex feasibility problem obviously remains unchanged, and $\{\{\Gamma_i^k\}_{i=1}^{m+1}\}_{k \geq 0}$ is a fair sequence according to Definition 3.1. Thus, Theorem 3.3 applies and the result follows.   □

Our Algorithm 3.2 could be derived from Byrne's [9, Algorithm 4.2] for only the special case of fixed blocks and a constant sequence of vectors of diagonal matrices, that is, for

$$G_i^k = G_i, \quad \text{for } i = 1, 2, \dots, m, \ \forall k \geq 0. \tag{3.39}$$

It is, therefore, natural to ask whether the BIMSGP method of [9] can be extended to accommodate diagonal weighting matrices which vary with the iteration index k, but we do not address this question here. The BIMSGP method is a multiprojection scheme and, as such, it also contains, as a special case, the forerunner [12] of multiprojection methods.

4. Analysis of the fully simultaneous algorithm in the inconsistent case

The fully simultaneous algorithm with seminorm-induced oblique projections for the convex feasibility problem is obtained from Algorithm 3.2 by imposing the condition that all diagonal matrices of (3.1) are nonzero matrices which do not change as iterations proceed. So, we assume throughout this section that, for all $k \geq 0$,

$$G_i^k = G_i \neq 0, \quad \text{for } i = 1, 2, \dots, m. \tag{4.1}$$

A separate treatment is applied in this section, which allows us to prove that, in this case, Algorithm 3.2 generates sequences $\{x^k\}_{k \geq 0}$ which always converge, regardless of the initial point $x^0 \in \mathbb{R}^n$ and independently from the consistency $C \neq \varnothing$ or inconsistency $C = \varnothing$ of the underlying system $C := \cap_{i=1}^m C_i$. Moreover, it will always converge to a minimizer of a certain *proximity function* assuming that this function has minimizers. Our analysis follows the same pattern of proof of the CAV algorithm given in [15]. We treat first the case $\lambda_k = 1$, for all $k \geq 0$, and then expand the result to underrelaxation.

We define a proximity function

$$F(x) := \sum_{i=1}^{m} \left| P_{C_i}^{G_i}(x) - x \right|_{G_i}^{2} \tag{4.2}$$

and an operator

$$T(z) := \sum_{i=1}^{m} G_i P_{C_i}^{G_i}(z). \tag{4.3}$$

The first step towards establishing convergence is the next proposition.

PROPOSITION 4.1. *Under assumptions (i), (iii), (iv), and (v) of Theorem 3.3, for every sequence $\{x^k\}_{k \geq 0}$, generated by Algorithm 3.2, with $\lambda_k = 1$ for all $k \geq 0$, and assuming (4.1), the sequence $\{F(x^k)\}_{k \geq 0}$, with F defined by (4.2), is decreasing and $\lim_{k \to \infty} \|x^{k+1} - x^k\|_2 = 0$.*

Proof. From (4.2) and Proposition 3.4 with (4.1), we obtain

$$F(x^k) = \sum_{i=1}^{m} \left| P_{C_i}^{G_i}(x^k) - T(x^k) \right|_{G_i}^{2} + \|T(x^k) - x^k\|_2^2. \tag{4.4}$$

The operator $T(z)$ of (4.3) is similar to (3.5) only with no dependence on k, thus using (3.3) with $\lambda_k = 1$ for all $k \geq 0$, along with (4.1), we have that $x^{k+1} = T(x^k)$. Since, by (2.5), $P_{C_i}^{G_i}(x^{k+1})$ minimizes $\left| x - x^{k+1} \right|_{G_i}^{2}$ over all $x \in C_i$, we can continue (4.4) to obtain, for all $k \geq 0$,

$$\begin{aligned} F(x^k) &\geq \sum_{i=1}^{m} \left| P_{C_i}^{G_i}(x^{k+1}) - x^{k+1} \right|_{G_i}^{2} + \|x^{k+1} - x^k\|_2^2 \\ &= F(x^{k+1}) + \|x^{k+1} - x^k\|_2^2 \\ &\geq F(x^{k+1}). \end{aligned} \tag{4.5}$$

The monotonicity and nonnegativity of $\{F(x^k)\}_{k \geq 0}$ guarantee that the limit $\lim_{k \to \infty} F(x^k)$ exists and thus, (4.5) implies $\lim_{k \to \infty} \|x^{k+1} - x^k\|_2 = 0$. $\square$

We denote the set of minimizers of the proximity function F by

$$\Phi := \{\hat{x} \in \mathbb{R}^n \mid F(\hat{x}) \leq F(x), \ \forall x \in \mathbb{R}^n\}. \tag{4.6}$$

The next proposition establishes the Fejér-monotonicity with respect to Φ of the iterates.

PROPOSITION 4.2. *Under the conditions of Proposition 4.1, if $\Phi \neq \varnothing$, then any sequence $\{x^k\}_{k \geq 0}$, generated by Algorithm 3.2, is Fejér-monotone with respect to Φ.*

Proof. Take any $\hat{x} \in \Phi$ and use (4.5) with $x^0 = \hat{x}$ and $x^1 = T(\hat{x})$, then we get

$$F(\hat{x}) \geq F(T(\hat{x})) + \left\| T(\hat{x}) - \hat{x}_2^2 \right\|_2. \tag{4.7}$$

Since $\hat{x}$ is a minimizer of F, $F(\hat{x}) \leq F(T(\hat{x}))$ and so (4.7) shows that $T(\hat{x}) = \hat{x}$. Using Proposition 2.8 with $y = x^k$, $G = G_i$, $Q = C_i$, and $z = P_{C_i}^{G_i}(\hat{x})$, we obtain, for $i = 1, 2, \ldots, m$, that for any sequence $\{x^k\}_{k \geq 0}$, generated by Algorithm 3.2 with $\lambda_k = 1$ for all $k \geq 0$, and assuming (4.1),

$$\left\| P_{C_i}^{G_i}(\hat{x}) - x^k \right\|_{G_i}^2 \geq \left\| P_{C_i}^{G_i}(x^k) - x^k \right\|_{G_i}^2 + \left\| P_{C_i}^{G_i}(\hat{x}) - P_{C_i}^{G_i}(x^k) \right\|_{G_i}^2. \tag{4.8}$$

Summing up all these inequalities, using Proposition 3.4 with (4.1) for the resulting left-hand side, and using (4.2), we obtain

$$\sum_{i=1}^{m} \left\| P_{C_i}^{G_i}(\hat{x}) - T(\hat{x}) \right\|_{G_i}^2 + \left\| T(\hat{x}) - x^k \right\|_2^2 \geq F(x^k) + \Gamma(x^k), \tag{4.9}$$

where

$$\Gamma(v) := \sum_{i=1}^{m} \left\| P_{C_i}^{G_i}(\hat{x}) - P_{C_i}^{G_i}(v) \right\|_{G_i}^2. \tag{4.10}$$

Using (4.2) again and the fact that $T(\hat{x}) = \hat{x}$, (4.9) can be rewritten as

$$\left\| \hat{x} - x^k \right\|_2^2 \geq F(x^k) - F(\hat{x}) + \Gamma(x^k). \tag{4.11}$$

Since $\hat{x}$ minimizes F, we have $F(x^k) - F(\hat{x}) \geq 0$. Denoting

$$\gamma^i(v) := P_{C_i}^{G_i}(\hat{x}) - P_{C_i}^{G_i}(v), \tag{4.12}$$

we have, by (4.3),

$$\sum_{i=1}^{m} G_i \gamma^i(v) = \sum_{i=1}^{m} G_i P_{C_i}^{G_i}(\hat{x}) - \sum_{i=1}^{m} G_i P_{C_i}^{G_i}(v) = T(\hat{x}) - T(v) = \hat{x} - T(v). \tag{4.13}$$

Then, from (4.10) and (2.2), we get

$$
\begin{aligned}
\Gamma(v) &= \sum_{i=1}^{m} \left\langle \gamma^i(v), G_i \gamma^i(v) \right\rangle \\
&\geq \left\langle \sum_{i=1}^{m} G_i \gamma^i(v), \sum_{i=1}^{m} G_i \gamma^i(v) \right\rangle \\
&= \left\| \hat{x} - T(v) \right\|_2^2
\end{aligned}
\tag{4.14}
$$

which, along with (4.11) and $x^{k+1} = T(x^k)$, proves the required result. The inequality in (4.14) follows from convexity considerations: due to $\sum_{i=1}^{m} G_i = I$, we have, for every $j = 1, 2, \ldots, n$, that $\sum_{i=1}^{m} g_{ij} = 1$ while $g_{ij} \geq 0$, for all i and all j. Convexity of the real-valued function $f(x) = x^2$ then implies that, for every $j = 1, 2, \ldots, n$,

$$
\sum_{i=1}^{m} g_{ij} \left(\gamma_j^i(v) \right)^2 \geq \sum_{i=1}^{m} \left(g_{ij} \gamma_j^i(v) \right)^2.
\tag{4.15}
$$

Summing up all these inequalities over j yields the inequality in (4.14).   □

Now, we are ready to prove the convergence theorem.

THEOREM 4.3. *Under assumptions (i), (iii), (iv), and (v) of Theorem 3.3, if the proximity function F has minimizers, that is, $\Phi \neq \varnothing$, then any sequence $\{x^k\}_{k \geq 0}$, generated by Algorithm 3.2 with $\lambda_k = 1$ for all $k \geq 0$, and assuming (4.1), converges to a minimizer of F.*

Proof. It follows from the Fejér-monotonicity, established in Proposition 4.2, that $\{x^k\}_{k \geq 0}$ is bounded, thus it has at least one cluster point. We show now that any cluster point of $\{x^k\}_{k \geq 0}$ is a minimizer of F. Let x^* be a cluster point of $\{x^k\}_{k \geq 0}$ and let $\hat{x} \in \Phi$. Using Proposition 2.8 with $y = x^*$, $G = G_i$, $Q = C_i$, and $z = P_{C_i}^{G_i}(\hat{x})$, for $i = 1, 2, \ldots, m$, summing up all inequalities, and using Proposition 3.4 with (4.1), (4.2), and (4.10), we obtain

$$
F(\hat{x}) + \left\| \hat{x} - x^* \right\|_2^2 \geq F(x^*) + \Gamma(x^*).
\tag{4.16}
$$

The fact that $\lim_{k \to \infty} \| x^{k+1} - x^k \|_2 = 0$ (see Proposition 4.1) guarantees that $T(x^*) = x^*$ because x^* is a cluster point. Thus, (4.16) and (4.14) show that $F(\hat{x}) \geq F(x^*)$. Finally, if x^* is a cluster point of $\{x^k\}_{k \geq 0}$, then $x^* \in \Phi$ and Proposition 4.2 guarantees that, for all $k \geq 0$,

$$
0 \leq \left\| x^* - x^{k+1} \right\|_2 \leq \left\| x^* - x^k \right\|_2.
\tag{4.17}
$$

Thus, $\lim_{k\to\infty}\|x^* - x^k\|_2$ exists and since x^* is a cluster point, this limit must be zero, proving that x^* is the limit of $\{x^k\}_{k\geq 0}$.   □

Now, we can introduce underrelaxation parameters as follows.

THEOREM 4.4. *Under assumptions (i), (iii), (iv), and (v) of Theorem 3.3, if the proximity function F has minimizers, that is, $\Phi \neq \varnothing$, then any sequence $\{x^k\}_{k\geq 0}$, generated by Algorithm 3.2 with $0 < \epsilon \leq \lambda_k \leq 1$ for all $k \geq 0$, for some arbitrarily small but fixed ϵ, and assuming (4.1), converges to a minimizer of F.*

Proof. Similar to the proof of Theorem 3.6.   □

The special case, when in Theorem 4.3 all matrices $G_i = (1/m)I, i = 1, 2, \ldots, m$, the proximity function F of (4.2) is the classical least-squares measure, and the projections $P_{C_i}^{G_i}$ are the orthogonal projections, was investigated by Iusem and De Pierro [23] and Combettes [18]. The case considered in [23] and in [18] is also covered by the results of Bauschke and Borwein [4, Theorem 6.2] where a discussion on when Φ is nonempty is also offered.

Acknowledgments

References

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

[2] M. Arioli, I. Duff, J. Noailles, and D. Ruiz, *A block projection method for sparse matrices*, SIAM J. Sci. Statist. Comput. **13** (1992), no. 1, 47–70.

[3] A. Auslender, *Optimisation. Méthodes Numériques*, Maîtrise de Mathématiques et Applications Fondamentales, Masson, Paris, 1976 (French).

[4] H. H. Bauschke and J. M. Borwein, *On the convergence of von Neumann's alternating projection algorithm for two sets*, Set-Valued Anal. **1** (1993), no. 2, 185–212.

[5] ———, *On projection algorithms for solving convex feasibility problems*, SIAM Rev. **38** (1996), no. 3, 367–426.

[6] R. T. Behrens and L. L. Scharf, *Signal processing applications of oblique projection operators*, IEEE Trans. Signal Processing **42** (1994), 1413–1424.

[7]   Å. Björck, *Numerical Methods for Least Squares Problems*, SIAM, Pennsylvania, 1996.

[8]   L. M. Bregman, *A relaxation method of finding a common point of convex sets and its application to the solution of problems in convex programming*, U.S.S.R. Comput. Math. and Math. Phys. **7** (1967), 200–217.

[9]   C. Byrne, *Iterative projection onto convex sets using multiple Bregman distances*, Inverse Problems **15** (1999), no. 5, 1295–1313.

[10] C. Byrne and Y. Censor, *Proximity function minimization using multiple Bregman projections, with applications to split feasibility and Kullback-Leibler distance minimization*, Ann. Oper. Res. **105** (2001), 77–98.

[11] Y. Censor, *Finite series-expansion reconstruction methods*, Proc. IEEE **71** (1983), 409–419.

[12] Y. Censor and T. Elfving, *A multiprojection algorithm using Bregman projections in a product space*, Numer. Algorithms **8** (1994), no. 2-4, 221–239.

[13] ———, *Block-iterative algorithms with diagonally scaled oblique projections for the linear feasibility problem*, SIAM J. Matrix Anal. Appl. **24** (2002), no. 1, 40–58.

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

[15] ———, *Component averaging: an efficient iterative parallel algorithm for large and sparse unstructured problems*, Parallel Comput. **27** (2001), no. 6, 777–808.

[16] Y. Censor and S. A. Zenios, *Parallel Optimization Theory, Algorithms, and Applications*, Numerical Mathematics and Scientific Computation, Oxford University Press, New York, 1997.

[17] G. Cimmino, *Calcolo approssimato per le soluzioni dei sistemi di equazioni lineari*, Ric. Sci. Progr. Tecn. Econom. Naz. **9** (1939), 326–333 (Italian).

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

[19] ———, *The convex feasibility problem in image recovery*, Advances in Imaging and Electron Physics, vol. 95, Academic Press, New York, 1996, pp. 155–270.

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

[21] T. Elfving, *Block-iterative methods for consistent and inconsistent linear equations*, Numer. Math. **35** (1980), no. 1, 1–12.

[22] G. T. Herman, *Image Reconstruction from Projections. The Fundamentals of Computerized Tomography*, Computer Science and Applied Mathematics, Academic Press, New York, 1980.

[23] A. N. Iusem and A. R. De Pierro, *Convergence results for an accelerated nonlinear Cimmino algorithm*, Numer. Math. **49** (1986), no. 4, 367–378.

[24] S. Kayalar and H. L. Weinert, *Oblique projections: formulas, algorithms, and error bounds*, Math. Control Signals Systems **2** (1989), no. 1, 33–45.

[25] E. R. Lorch, *On a calculus of operators in reflexive vector spaces*, Trans. Amer. Math. Soc. **45** (1939), no. 2, 217–234.

[26] F. J. Murray, *On complementary manifolds and projections in spaces L_p and l_p*, Trans. Amer. Math. Soc. **41** (1937), no. 1, 138–152.

[27] G. Pierra, *Decomposition through formalization in a product space*, Math. Programming **28** (1984), no. 1, 96–115.

[28] S. Reich, *A limit theorem for projections*, Linear and Multilinear Algebra **13** (1983), no. 3, 281–290.

[29]   A. E. Taylor, *Introduction to Functional Analysis*, John Wiley & Sons, New York, 1967.

Yair Censor: Department of Mathematics, University of Haifa, Mt. Carmel, Haifa 31905, Israel

E-mail address: yair@math.haifa.ac.il

Tommy Elfving: Mathematics Department, Linköping University, SE-581 83 Linköping, Sweden

E-mail address: toelf@mai.liu.se