On the Convergence of Asynchronous Paracontractions with Application to Tomographic Reconstruction from Incomplete Data

L. Elsner
Fakultät für Mathematik
Universität Bielefeld
Postfach 8640
D-4800 Bielefeld, Federal Republic of Germany

and

I. Koltracht* and M. Neumann[†]
Department of Mathematics
University of Connecticut
Storrs, Connecticut 06268

Submitted by Yair Censor

ABSTRACT

Convergence of iterative processes in C^k of the form

$$x_{i+r_i} = \alpha_{j_i} x_{i+r_i-1} + \left(1 - \alpha_{j_i}\right) P_{j_i} x_i,$$

where $j_i \in \{1, 2, \ldots, n\}$, $i = 1, 2, \ldots$, is analyzed. It is shown that if the matrices $P_1, \ldots, P_n$ are paracontracting in the same smooth, strictly convex norm and if the sequence $\{j_i\}_{i=1}^{\infty}$ has certain regularity properties, then the above iterates converge. This result implies the convergence of a parallel asynchronous implementation of the algebraic reconstruction technique (ART) algorithm often used in tomographic reconstruction from incomplete data.

INTRODUCTION

This paper is motivated by two recent developments: (1) the asynchronous parallel implementation of iterative algorithms for solving nonsingular systems whose coefficient matrix is inverse positive, (see Bru, Elsner, and Neumann [3]), and (2) the close relationship between SOR and the algebraic reconstruction technique (ART) (see Natterer [14], Nelson and Neumann [15], and Elsner, Koltracht, and Lancaster [5]). These developments have led us to seek parallel implementation of the ART algorithm.

Projection methods for solving linear systems of equations have been known for quite some time and probably originated in a 1937 paper by Kaczmarz [10]. Renewed interest in these methods in the early 1970s was spurred by their successful use in computed tomography; see Herman [9] and references therein. In the tomographic literature these methods came to be known collectively as the *algebraic reconstruction technique* (ART), a term which we adopt here.

In applications of computed tomography to situations where the imaging data from any direction is available, e.g. in X-ray scanners, techniques based on the inverse Radon transform are currently preferred. However, in the absence of complete projection data, as in geophysical cross-hole tomography for example (see Dines and Little [2]), the ART algorithms remain useful. This is because these algorithms can be used for arbitrary systems of linear equations $Rx = f$ which are inconsistent, underdetermined, and of very large size; see [9, 14, 5] for example.

The description of a general 2-D tomographic problem with incomplete data and its solution using the ART algorithm are given in Section 1. The analysis of the convergence of the parallel asynchronous implementation of the ART algorithm is based on our main results, which are presented in two theorems in Section 2.

A matrix $P \in C^{k,k}$ is called *paracontracting* with respect to some vector norm $\|\cdot\|$ if

$$Px \neq x \quad \Leftrightarrow \quad \|Px\| < \|x\|.$$

Let $P_1, P_2, \ldots, P_n$ be paracontracting matrices with respect to the same norm $\|\cdot\|$ and for each $i = 1, 2, \ldots, n$ let

$$M_i := N(I - P_i)$$

be the nullspace of $I - P_i$. Consider the sequence of vectors $\{x_i\}_{i=1}^{\infty} \in C^k$

generated by the iterative process

$$x_i = P_{j_i} x_{i-1}, \quad i \geqslant 1, \ 1 \leqslant j_i \leqslant n, \qquad x_0 \in C^k.$$

Our first theorem establishes the convergence of this sequence to a limit y which belongs to the subspace $M = \bigcap_{i \in J} M_i$, where J is the set of integers t appearing infinitely often in the sequence $\{j_i\}_{i=1}^{\infty}$.

The limit y can be characterized more precisely if the norm $\|\cdot\|$ is smooth and strictly convex and each of the integers $1, 2, \ldots, n$ appears infinitely often in the sequence $\{j_i\}_{i=1}^{\infty}$. (Such a sequence is called *admissible*.) If the norm in which $P_1, P_2, \ldots, P_n$ are paracontracting is smooth, then $P_1^*, P_2^*, \ldots, P_n^*$ are paracontracting in the dual norm. In this case the subspaces M and

$$M^c = \text{span}\{R(I - P_i) \,|\, i = 1, \ldots, n\},$$

where $R(I - P_i)$ denotes the range of $I - P_i$, are complementary, and y is the projection of x_0 onto M along M^c.

This generalizes results of Amemiya and Ando [1] and Youla [17] for the 2-norm. The generalization is necessary for the analysis of the parallel asynchronous process of the form

$$x_{i+r_i} = \alpha_{j_i} x_{i+r_i-1} + (1 - \alpha_{j_i}) P_{j_i} x_i,$$

where $\alpha_1, \alpha_2, \ldots, \alpha_n$ are in $(0, 1)$. The convergence of this process is established in Theorem 2 under the additional assumption that the sequence $\{j_i\}_{i=1}^{\infty}$ is also *regulated*, that is, there exists an integer T such that during the computation cycle of length T each integer $1, 2, \ldots, n$ appears at least once, namely,

$$\{1, 2, \ldots, n\} \subset \{j_k, j_{k+1}, \ldots, j_{k+T-1}\}$$

for any k.

Theorem 2 is used in Section 3 to prove the convergence of a parallel asynchronous implementation of the ART algorithm for the solution of $Rx = f$ in the case when the system $Rx = f$ is consistent. The limit y is the minimum norm solution of such a system. The minimum norm least squares

solution in the inconsistent case can be found, following Miller and Neumann [13], as a part of the minimum norm solution of the consistent system

$$\begin{pmatrix} R & I \\ 0 & R^T \end{pmatrix} \begin{pmatrix} \hat{g} \\ s \end{pmatrix} = \begin{pmatrix} f \\ 0 \end{pmatrix}.$$

1. TOMOGRAPHIC RECONSTRUCTION FROM INCOMPLETE DATA

In this section we describe the problem of reconstructing certain properties of a medium from measurements of its response to probing signals taken on its boundary. Consider the standard two dimensional problem with limited access to boundary. It is assumed that the medium is confined to a rectangle as shown in Figure 1, and that one has access to any two or three sides of the rectangle.

The rectangle is divided into NM rectangular pixels, and it is further assumed that the property of interest of the medium does not change within each pixel and is quantitatively characterized by the unknown values $x_1, x_2, \ldots, x_{NM}$. A probing signal can be transmitted from one side of the rectangle, and the response of the medium is measured at different locations on other accessible sides of the medium as shown in Figure 2. It is also

x_1	x_2	· · ·	x_M
x_{M+1}	x_{M+2}	· · ·	x_{2M}
· ·	· ·	· · ·	· ·
$x_{M(N-1)+1}$	$x_{M(N-1)+2}$	· · ·	x_{MN}

Fig. 1.   Rectangular medium with unknown densities constant in each pixel.
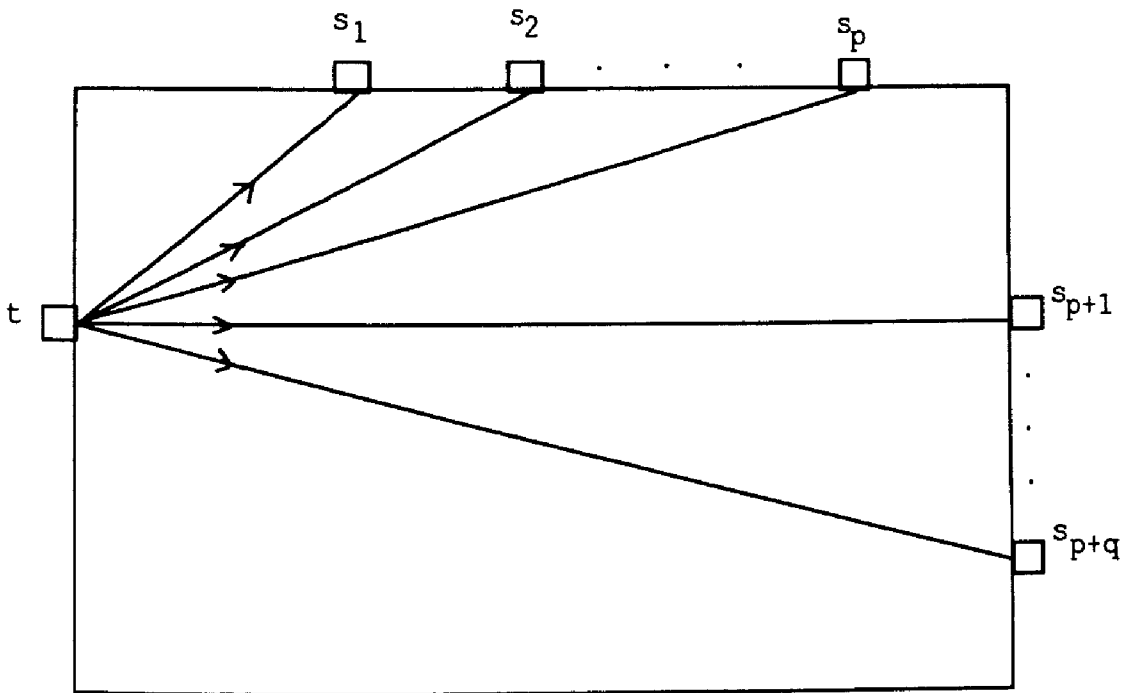
FIG. 2. Response of the medium to a signal transmitted from location t received at locations $s_1, s_2, \ldots, s_{p+q}$.

assumed that the signal travels along straight lines and that the measurement at the receiver, s_i, represents the integral (along the line segment connecting the transmitter t and the receiver s_i) of the piecewise constant function defined by its unknown values $x_1, x_2, \ldots, x_{NM}$.

For example let us consider the geophysical cross-hole tomography model as shown in Figure 3. Electromagnetic or acoustic signals sent from one well are received by a string of geophones in the second well. The travel times of the signals are measured. If v denotes the velocity of the signal in earth, then the travel time of the signal from the transmitter to the geophone equals

$$\int_L \frac{d\sigma}{v},$$
(1)

where L is the line segment connecting the transmitter and the geophone. For more details about the cross-hole tomography models see [2], McMechan [12], and Koltracht, Lancaster, and Smith [11].

Thus, if the section of earth is discretized as shown in Figure 1 and the velocity is assumed to be constant in each pixel, the unknowns $x_1, x_2, \ldots, x_{NM}$ represent the attenuation (the reciprocal of the velocity) of the signal in
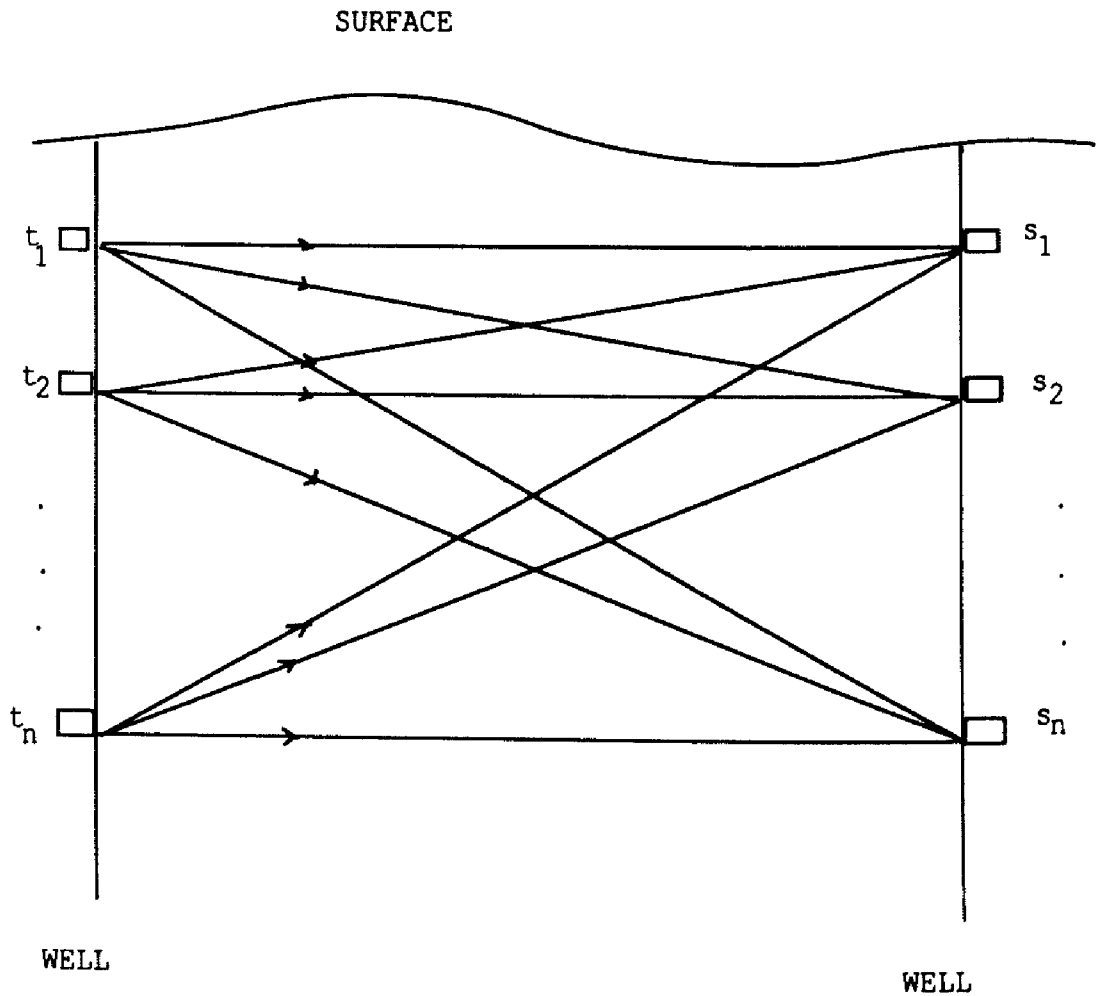
SURFACE

FIG. 3. Geophysical cross-hole tomography model.

corresponding pixels. It is now clear that each line integral (1) can be written as follows:

$$f_i = \int_{L_i} \frac{d\sigma}{v} = \sum_{j=1}^{NM} r_{ij} x_j, \qquad i = 1, 2, \ldots, n, \qquad (2)$$

where the index $i = 1, 2, \ldots, n$ corresponds to some ordering of the transmitter-geophone pairs, r_{ij} is the intersection length of the ith line segment with pixel number j, and f_i is the measured travel time of the signal.

The system of equations (2) can be rewritten in matrix form as follows:

$$Rx = f, \qquad (3)$$

where $R \in R^{n,NM}, x \in R^{NM}$, and $f \in R^n$. It is clear that the matrix R is very

large; for example, if $N = M = 100$ then R is 10,000 by 10,000. It is also sparse, as only $O(N + M)$ pixels have nonzero intersection with any given line segment. Moreover, the system (3) is in general inconsistent (there is no solution) and underdetermined (the nullspace dimension can be quite large; see [12, 11]).

A technique frequently used in geophysical tomography for approximate solution of (3) is the algebraic reconstruction technique (ART). This is an iterative method which makes proper use of the sparseness of R, and of the fact that the matrix R can be generated row by row whenever necessary. The convergence properties of ART have been investigated; see [9, 14, 15, 5] and Hanke and Niethammer [8].

Let us briefly describe this technique. Let

$$R = \begin{pmatrix} R_1^T \\ R_2^T \\ \vdots \\ R_n^T \end{pmatrix}$$

be the row representation of R. Let P_i denote the following matrices:

$$P_i = I - \omega \frac{R_i R_i^T}{R_i^T R_i}, \qquad i = 1, 2, \ldots, n,$$

where $\omega \in (0, 2)$. Consider the iterative process

$$x_i = P_{j_i} x_{i-1} + \omega \frac{f_{j_i}}{R_{j_i}^T R_{j_i}} R_{j_i}, \qquad j_i = \begin{cases} i \bmod n, & i \neq kn, \\ n, & i = kn, \end{cases} \qquad i = 1, 2, \ldots, \quad (4)$$

where x_0 is any vector in R^{NM}. Then the cyclic iterates $x_n, x_{2n}, x_{3n}, \ldots$ converge, and if the system (3) is consistent, they converge to its minimum norm solution plus the component of x_0 in the nullspace of R. If the system is inconsistent, then the cycle limit converges to the minimum norm least squares solution of (3) with normalized rows (plus the component of x_0 in the nullspace of R) when $\omega \to 0$. In general the distance from the limit to this minimum norm least squares solution is proportional to the distance from f to the column space of R (for more details see [5] for example).

The rate of convergence of ART is governed by the relaxation parameter ω. However, the optimal choice of ω remains an open question. It is

important therefore to investigate ways of improving the convergence of the ART algorithm other than the search for an optimal relaxation parameter. At present a plausible approach to this problem is the use of parallel computation.

2. MAIN RESULTS

Let $\|\cdot\|$ denote a norm in C^k. As in Nelson and Neumann [15], a matrix $P \in C^{k,k}$ is called *paracontracting* (*with respect to* $\|\cdot\|$) if

$$Px \neq x \quad \Leftrightarrow \quad \|Px\| < \|x\|. \tag{5}$$

We denote by $\mathcal{N}(\|\cdot\|)$ the set of all $k \times k$ paracontracting matrices. In [15] it is shown that if $P \in \mathcal{N}(\|\cdot\|)$ then $\exists \lim_{n \to \infty} P^n$. It is well known (e.g. Berman and Plemmons [4]) that the existence of this limit is equivalent to the following conditions being satisfied: the subspaces $N(I - P)$ and $R(I - P)$ are complementary subspaces in C^k and all eigenvalues of P other than 1 lie in the interior of the unit circle. Here $N(\cdot)$ and $R(\cdot)$ denote the nullspace and the range, respectively. It is further shown in [15] that if $P, Q \in \mathcal{N}(\|\cdot\|)$, then $PQ \in \mathcal{N}(\|\cdot\|)$ and

$$N(I - PQ) = N(I - P) \cap N(I - Q). \tag{6}$$

In what follows assume that $P_1, \ldots, P_n$ are given matrices in $\mathcal{N}(\|\cdot\|)$. Let $M_i = N(I - P_i)$, $i = 1, \ldots, n$, and let $M = \bigcap_{i=1}^{n} M_i$. We first wish to study the convergence of sequences of vectors defined by

$$x_i = P_{j_i} x_{i-1}, \qquad i = 1, 2, \ldots, \tag{7}$$

where x_0 is an arbitrary vector and $\{j_i\}_{i=1}^{\infty}$ is a sequence of integers such that $1 \leqslant j_i \leqslant n$ for all $i \geqslant 1$. Such a sequence of integers is called *admissible* if each one of the integers $1, \ldots, n$ appears in it infinitely often, and it is called *regulated* if there exists an integer $T > 0$ such that

$$\{1, 2, \ldots, n\} \subset \{j_k, \ldots, j_{k+T-1}\} \qquad \forall k \geqslant 1.$$

We call T a *computational cycle* (*of the sequence*).

Our first main result is the following:

THEOREM 1. *Let* $\|\cdot\|$ *be a vector norm on* C^k, *and suppose that* $P_i \in \mathcal{N}(\|\cdot\|)$, $i = 1, \ldots, n$. *Let* $\{j_i\}_{i=1}^{\infty}$ *with* $1 \leqslant j_i \leqslant n$, $i = 1, 2, \ldots$, *be a sequence of integers, and denote by* J *the set of all integers which appear infinitely often in* $\{j_i\}_{i=1}^{\infty}$. *Then for any* $x_0 \in C^k$ *the sequence of vectors* $x_i = P_{j_i} x_{i-1}$, $i \geqslant 1$, *has a limit* $y \in \bigcap_{i \in J} N(I - P_i)$. *If, in addition, the sequence* $\{j_i\}_{i=1}^{\infty}$ *is admissible and* $\|\cdot\|$ *is smooth, then* $y = P_{M,M^c} x_0$, *where* $M = \bigcap_{i=1}^{n} N(I - P_i)$, $M^c = \text{span}\{R(I - P_i) \mid 1 \leqslant i \leqslant n\}$, *and* P_{M,M^c} *is the projection on* M *along* M^c.

To prove Theorem 1 we need several lemmata.

LEMMA 1. *If* $P_i \in C^{k,k}$, $i = 1, \ldots, n$, *and* M^c *is the subspace given in Theorem 1, then* $P_i(M^c) \subseteq M^c$, $i = 1, \ldots, n$.

Proof. Suppose $\xi \in M^c$. Then since $(I - P_i)\xi \in M^c$, we must have that $P_i \xi \in M^c$, $i = 1, \ldots, n$. ∎

LEMMA 2. *Let* $\|\cdot\|_\alpha$ *and* $\|\cdot\|_\beta$ *be vector norms on* C^k, *and suppose that* $P_i \in \mathcal{N}(\|\cdot\|_\alpha)$ *and* $P_i^* \in \mathcal{N}(\|\cdot\|_\beta)$, $i = 1, \ldots, n$. *Then the subspaces* M *and* M^c *given in Theorem 1 are complementary.*

Proof. First, by (6), $M = N(I - P_1 \cdots P_n)$ and $\bigcap_{i=1}^{n} N(I - P_i^*) = N(I - P_n^* \cdots P_1^*)$. Second, $M^c = [\bigcap_{i=1}^{n} N(I - P_i^*)]^\perp = R(I - P_1 \cdots P_n)$. Since $P_1 \cdots P_n \in \mathcal{N}(\|\cdot\|_\alpha)$, the conclusion now follows. ∎

LEMMA 3. *Let* $\|\cdot\|$ *be a smooth norm on* C^k. *If* $P \in \mathcal{N}(\|\cdot\|)$, *then* $P^* \in \mathcal{N}(\|\cdot\|_D)$.

Proof. It suffices to show that $\|P^*\xi\|_D = \|\xi\|_D$ implies $P^*\xi = \xi$. There exists η with $\|\eta\| = 1$ such that

$$\|P^*\xi\|_D = \Re((P^*\xi)^*\eta) = \Re(\xi^* P\eta),$$

where $\Re$ denotes the real part. As

$$\|\xi\|_D = \Re(\xi^* P\eta) \leqslant \|\xi\|_D \|P\eta\|,$$

we have $\|P\eta\| = \|\eta\|$ and hence $P\eta = \eta$. Thus there exist two vectors $\zeta_1 = \xi$

and $\zeta_2 = P^*\xi$ such that

$$\Re(\xi_i^*\eta) = \|\xi_i\|_D\|\eta\|, \qquad i = 1,2,$$

and so, due to the smoothness of $\|\cdot\|$, $\zeta_1 = \zeta_2$. Hence $P^*\xi = \xi$. ∎

REMARK. The matrix $P = \mathrm{diag}(1, \frac{1}{2})$ is in $\mathcal{N}(\|\cdot\|_1)$ but not in $\mathcal{N}(\|\cdot\|_\infty)$. This illustrates that in general we may not drop the assumption that $\|\cdot\|$ is smooth in Lemma 3.

LEMMA 4. *Let $\|\cdot\|$ be a norm on C^k, and suppose that $P_i \in \mathcal{N}(\|\cdot\|)$, $i = 1, \ldots, n$. Then for any $x_0 \in C^k$ and for any admissible sequence $\{j_i\}_{i=1}^\infty$, the iteration $x_i = P_{j_i} x_{i-1}$, $i = 1, 2, \ldots$, converges to a limit in $M = \bigcap_{i=1}^n M_i$.*

Proof. Let $\{x_i\}_{i=1}^\infty$ be a sequence as above. Since $\|x_i\| \geq \|x_{i+1}\|$, $i = 1, 2, \ldots$, the sequence is bounded and hence has an accumulation point, say y. Suppose $y \notin M$. Then there exists an $1 \leq r \leq n$ such that after possible reordering the P_i's, $P_i y = y$ for $i < r$ and $P_i y \neq y$ for $i \geq r$. Let $\{x_{p_i}\}_{i=1}^\infty$ be a subsequence of $\{x_i\}_{i=1}^\infty$ such that $x_{p_i} \to y$ as $i \to \infty$. Construct now a subsequence $\{y_i\}_{i=1}^\infty$ as follows: As $\{j_i\}_{i=1}^\infty$ is admissible, for each $i \geq 1$ there exists a smallest integer $q_i \geq p_i$ such that $j_{q_i+1} \geq r$. Observe that, as $x_{q_i} - y = P_{\alpha_1} P_{\alpha_2} \cdots P_{\alpha_t}(x_{p_i} - y)$ for some $1 \leq \alpha_i < r$, $0 \leq i \leq t$, we have $\|x_{q_i} - y\| \leq \|x_{p_i} - y\|$ for all i, showing that $x_{q_i} \to y$ as $i \to \infty$. Now at least one of the numbers $r, \ldots, n$ must occur infinitely often among the numbers j_{q_i+1}, $i \geq 1$. Assume that it is r. Consider the subsequence of $\{x_{q_i}\}_{i=1}^\infty$ for which $j_{q_{i+1}} = r$. Call it $\{y_i\}_{i=1}^\infty$. Then $\{y_i\}_{i=1}^\infty$ and $\{P_r y_i\}_{i=1}^\infty$ are subsequences of $\{x_i\}_{i=1}^\infty$. Hence

$$\|y\| = \lim_{i \to \infty} \|y_i\| = \lim_{i \to \infty} \|P_r y_i\| = \|P_r y\|,$$

so that $y = P_r y$. This contradicts $P_i y \neq y$ for $i \geq r$. Hence $y \in M$.

We next show that $x_i \to y$ as $i \to \infty$. Let $\epsilon > 0$. Then, there exists an integer $1 \leq j$ such that $\|x_{p_j} - y\| < \epsilon$. Then for some $1 \leq s \leq n$,

$$\|x_{p_{j+1}} - y\| = \|P_s x_{p_j} - P_s y\| \leq \|x_{p_j} - y\| < \epsilon,$$

and we see that $\|x_t - y\| < \epsilon$ for all $t \geq j$. This concludes the proof. ∎

Proof of Theorem 1. For some integer $S \geq 0$ the sequence $\{j_{S+i}\}_{i=1}^\infty$ is admissible with respect to J. Hence, on applying Lemma 4 to J (instead of $\{1, \ldots, n\}$), it follows that the sequence $\{x_i\}_{i=1}^\infty$ has a limit $y \in \bigcap_{i \in J} N(I - P_i)$.

Suppose now that $\|\cdot\|$ is smooth. By Lemmata 2 and 3 the subspaces M and M^c are complementary. Decompose x_0 into $x_0 = x_M + x_{M^c}$, where $x_M \in M$ and $x_{M^c} \in M^c$. Assume, in addition, that $\{j_i\}_{i=1}^{\infty}$ is an admissible sequence. Consider the sequence of iterates $x_i = P_{j_i} x_i$, $i = 1, 2, \ldots$, with x_0 replaced by x_{M^c}. By Lemma 4 its limit is in M. However, by Lemma 1, the limit must lie in M^c and hence it must be zero. The final conclusion of the theorem is now obvious. ∎

Motivated by [3], we next consider a modification of the iteration (7) which can be implemented using parallel processors in an asynchronized manner. Let $\{j_i\}_{i=1}^{\infty}$ be a regulated sequence and consider the iteration

$$x_{i+r_i} = \alpha_{j_i} x_{i+r_i-1} + \left(1 - \alpha_{j_i}\right) P_{j_i} x_i, \tag{8}$$

where $P_1, \ldots, P_n$ are the $k \times k$ matrices which were introduced at the beginning of this section. Here $\alpha_1, \ldots, \alpha_n$ are numbers from $(0, 1)$, and r_i, $i = 1, 2, \ldots$, are integers satisfying $1 \leqslant r_i \leqslant T$, T being the *computational cycle*. The underlying model of computation is this: We have n processors $\pi_1, \ldots, \pi_n$. At time i processor π_{j_i} retrieves the global approximation x_i, which resides in some shared memory, and computes a local iteration $P_{j_i} x_i$. If the global approximation in the shared memory has been updated $r_i - 1$ times while processor j_i computes its local iteration, then the global approximation is updated as in (8), yielding the approximation at time $i + r_i$.

The second main result of this paper gives conditions under which the sequence (8) converges.

THEOREM 2. *Let $\|\cdot\|$ be a smooth and strictly convex vector norm on C^k and suppose that $P_i \in \mathcal{N}(\|\cdot\|)$, $i = 1, \ldots, n$. Let $\{j_i\}_{i=1}^{\infty}$ be a regulated sequence of integers, $1 \leqslant j_i \leqslant n$, with a computational cycle T. For each $i = 1, 2, \ldots$, let r_i be the smallest positive integer s such that $j_{i+s} = j_i$. For a given vector $z \in C^k$ consider the sequence defined in the following way:*

$$x_s = \begin{cases} z & s \leqslant T \\ \alpha_{j_i} x_{s-1} + \left(1 - \alpha_{j_i}\right) P_{j_i} x_i, & s = i + r_i > T. \end{cases} \tag{9}$$

Then

$$\lim_{s \to \infty} x_s = P_{M, M^c} z, \tag{10}$$

where $M = \bigcap_{j=1}^{n} N(I - P_j)$ and $M^c = \mathrm{span}\{R(I - P_j) | 1 \leqslant j \leqslant n\}$.

Proof. Consider the kn-vector ξ_i, $i \geqslant T$, partitioned into n k-subvectors as follows:

$$\xi_i = \begin{pmatrix} (\xi_i)_1 \\ \vdots \\ (\xi_i)_n \end{pmatrix},$$

where, for $m = 1, \ldots, n$ we have $(\xi_i)_m = x_t$, and t is the largest integer not greater than i satisfying $j_t = m$. Let $\mu > 1$. Then

$$\xi_\mu = B_\mu \xi_{\mu - 1}, \tag{11}$$

where B_μ is the $kn \times kn$ matrix given in a block form $((B_\mu)_{s,t})_{s,t=1}^n$, where

$$(B_\mu)_{s,t} = \begin{cases} \delta_{s,t} I & s \neq j_\mu \text{ or } s = j_\mu \text{ and } t \neq j_\mu, j_{\mu-1}, \\ (1 - \alpha_{j_\mu}) P_{j_\mu} & s = t = j_\mu, \\ \alpha_{j_\mu} I & s = j_\mu \text{ and } t = j_{\mu-1}. \end{cases}$$

Next, define a norm on C^{nk} by

$$||| \eta ||| = \left\| \begin{pmatrix} \eta_i \\ \vdots \\ \eta_n \end{pmatrix} \right\| = \max_{1 \leqslant m \leqslant n} \|\eta_m\|, \qquad \eta_m \in C^k, \quad m = 1, \ldots, n.$$

Now for $\mu > T + 1$,

$$(B_\mu \eta)_{j_\mu} = (1 - \alpha_{j_\mu}) P_{j_\mu} \eta_{j_\mu} + \alpha_{j_\mu} \eta_{j_{\mu-1}} \tag{12}$$

and

$$(B_\mu \eta)_{n_u} = \eta_\nu, \qquad \nu \neq j_\mu. \tag{13}$$

Because $P_i \in \mathscr{N}(\|\cdot\|)$, $i = 1, \ldots, n$, it easily follows that $||| B_\mu \eta ||| \leqslant ||| \eta |||$. For $\mu > T + 1$ define the matrices

$$C_\mu = B_{\mu + 2T - 1} B_{\mu + 2T - 2} \cdots B_\mu. \tag{14}$$

We claim that $C_\mu \in \mathcal{N}(|||\cdot|||)$. It suffices to show that $|||C_\mu\eta||| = |||\eta|||$ implies that $C_\mu\eta = \eta$. Consider $B_\mu\eta$. Either $||(B_\mu\eta)_{j_\mu}|| < |||\eta|||$ or, by (12) and the strict convexity of $||\cdot||$, we have that $(B_\mu\eta)_{j_\mu} = \eta_{j_\mu} = \eta_{j_{\mu-1}}$. Proceeding in this manner with $B_{\mu+1}, \ldots, B_{\mu+T}$, we infer, using the regularity of the sequence $\{j_i\}_{i=1}^\infty$, that there is either a $\nu \leqslant T$ such that

$$\left\| \left(B_{\mu+\nu} B_{\mu+\nu-1} \cdots B_\mu\eta \right)_{j_{\mu+\nu}} \right\| < |||\eta||| \tag{15}$$

or

$$\eta_1 = \eta_2 = \cdots = \eta_n \quad \text{and} \quad \eta_i = P_i\eta_i, \quad i = 1, \ldots, n. \tag{16}$$

If (16) holds, then $C_\mu\eta = \eta$, while if (15) holds, then we can deduce that $|||C_\mu\eta||| < |||\eta|||$, contradicting the assumption that $|||C_\mu\eta||| = |||\eta|||$. Hence $C_\mu \in \mathcal{N}(|||\cdot|||)$.

Because there are only a finite number of distinct matrices among the C_μ's, the first part of Theorem 1 applies and we see that the sequence $\{\xi_{2\nu T}\}_{\nu=1}^\infty$ has a limit

$$\xi = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}$$

and $\xi = C_r\xi$ for some $r > T + 1$. Hence, by (16), $\xi_1, \xi_2, \ldots, \xi_n =: y \in M$. As $B_\mu\xi = \xi$ for all $\mu > T + 1$ and $|||B_\mu\zeta||| \leqslant |||\zeta|||$ for all $\zeta \in C^{kn}$, we conclude that $\xi_\mu \to \xi$ as $\mu \to \infty$, and it follows that $x_\nu \to y$ as $\nu \to \infty$.

To prove that (10) holds it suffices to let $z \in M^c$ and observe that, by Lemma 1 and (9), all iterates and hence also their limit y lie in M^c. Since $||\cdot||$ is smooth, M and M^c are complementary by Lemmata 2 and 3. Thus $y \in M \cap M^c = \{0\}$. This completes the proof. ∎

3. PARALLEL IMPLEMENTATION OF THE ART ALGORITHM

In this section the general results of Section 2 are applied to the ART algorithm given in (4). First we remark that the matrices

$$P_i = I - \omega \frac{R_i R_i^T}{R_i^T R_i}, \quad i = 1, 2, \ldots, n,$$

are paracontracting in the 2-norm. Indeed, the spectrum of each of them consists of 1 and $1 - \omega$ only. The subspace M^c as defined in Section 2 is now the row space of the matrix R, that is,

$$M^c = \text{span}(R_1, \ldots, R_n),$$

and the complementary space M is the nullspace of R,

$$M = \ker R.$$

Clearly in this case $M = (M^c)^\perp$

The algorithm (4) can be implemented on a parallel architecture in an asynchronous manner as described in Section 2. Specifically, assume that there are n processors $\pi_1, \ldots, \pi_n$ with local memory, which are independent of each other. Each processor π_i is connected to a shared memory. Before the iteration (4) starts, the local memory of each processor is supplied with the ith row R_i of the matrix R (or the code which allows the processor to generate this row), the weight $\alpha_i \in (0, 1)$, the relaxation parameter ω, the ith coordinate f_i of f and the initial vector x_0. The shared memory contains the initial vector x_0. Each processor π_i executes an identical code:

(a) Retrieve the current vector from the shared memory, say y.

(b) Compute the convex combination of y with the current vector in the local memory, say z:

$$z = \alpha_i y + (1 - \alpha_i)z.$$

Store z in the local memory and in the shared memory.

(c) Perform the ith step of the ART algorithm on z, namely, compute

$$z := P_i z + \frac{\omega f_i}{R_i^T R_i} R_i.$$

(d) Go to (a).

It is further assumed that the communication time between local and shared memories is negligibly small relative to the updating time in steps (b) and (c) and that no two processors access the shared memory at the same time.

It is clear that the process we just described can be expressed as follows:

$$x_{i+r_i} = \alpha_{j_i} x_{i+r_i-1} + (1 - \alpha_{j_i})\left(P_{j_i} x_i + \frac{\omega f_{j_i}}{R_{j_i}^T R_{j_i}} R_{j_i} \right),$$

where $\{x_k\}$ is the sequence of vectors subsequently stored in the shared memory and r_i is the time for the processor π_i to perform steps (b) and (c). Moreover, the sequence $\{j_i\}_{i=1}^\infty$ is admissible. On choosing a computational cycle T to be greater than any of the updating times for each processor, we see that this sequence is also regulated. First consider the consistent case.

THEOREM 3.   Let $R_1, \ldots, R_n \in C^k$ be the rows of an $n \times k$ matrix R; $\omega \in (0,2)$; $\alpha_i \in (0,1)$, $i = 1, \ldots, n$; and $x_0 = x_M + x_{M^c}$ be any vector in C^n, where $x_M \in M$ and $x_{M^c} \in M^c$. Suppose the equation

$$Rx = f \qquad (17)$$

has a solution. For $i = 1, 2, \ldots$ define

$$x_{i+r_i} = \alpha_{j_i} x_{i+r_i-1} + (1 - \alpha_{j_i})\left(P_{j_i} x_i + \frac{\omega f_{j_i}}{R_{j_i}^T R_{j_i}} R_{j_i} \right), \qquad (18)$$

where

$$P_i = I - \omega \frac{R_i R_i^T}{R_i^T R_i}, \qquad i = 1, 2, \ldots, n, \qquad (19)$$

and where R_i and j_i are as described above. Then

$$\lim_{k \to \infty} x_k = \hat{x} + x_M, \qquad (20)$$

where $\hat{x}$ is the unique minimum norm solution of (17). In particular, if x_0 is in the row space of R (e.g. $x_0 = 0$), then the limit is the minimum norm solution of (17).

Proof.   Since $\hat{x}$ is the unique minimum norm solution, we have $R\hat{x} = f$ and $\hat{x} \in \text{span}(R_1, \ldots, R_n)$; see Groetsch [6], for example. In particular $R_{j_i}^T \hat{x} =$

f_{j_i}. Thus (18) can be rewritten as

$$x_{i+r_i} = \alpha_{j_i} x_{i+r_i-1} + (1 - \alpha_{j_i}) \left(P_{j_i} x_i + \frac{\omega \hat{x}^T R_{j_i}}{R_{j_i}^T R_{j_i}} R_{j_i} \right).$$

Subtracting $\hat{x}$ from both sides of this equality and denoting $e_k = x_k - \hat{x}$, we get

$$e_{i+r_i} = \alpha_{j_i} e_{i+r_i-1} + (1 - \alpha_{j_i}) P_{j_i} e_i.$$

Since $\hat{x}$ is in the row space of R, it follows that the component of e_0 in the nullspace of R equals x_M. Thus it follows from Theorem 2 that

$$\lim_{k \to \infty} e_k = x_M,$$

and hence (20) holds.   ∎

In the inconsistent case we suggest an approach along similar lines to Miller and Neumann [13].

LEMMA 5. *The vector $\hat{x} \in R^k$ is the minimum norm least squares solution of (17) if and only if the vector $(\hat{x}, s)^T \in R^{k+n}$ is the minimum norm solution of the consistent system*

$$\begin{pmatrix} R & I \\ 0 & R^T \end{pmatrix} \begin{pmatrix} \hat{x} \\ s \end{pmatrix} = \begin{pmatrix} f \\ 0 \end{pmatrix}. \tag{21}$$

We can now apply the parallel ART algorithm to the augmented matrix of (21) and obtain the minimum norm solution of (21). The first k entries of this solution give the minimum norm least squares solution of the original equation (17). Even on a sequential architecture the additional computational effort required in the solution of (21) can be justified by the improved quality of the tomographic reconstruction given by the minimum norm least squares solution.

We comment that the solution of the augmented system on a parallel architecture amounts to the addition of k processors. Furthermore, also in the consistent case one cannot always expect that the number of processors

will match the numbers of rows, n. However, given p processors, one can partition the matrix R, or the augmented matrix of (21), into p submatrices

$$R = \begin{pmatrix} R_1^T \\ R_2^T \\ \vdots \\ R_p^T \end{pmatrix},$$

and let each of the processors work as a sequential processor on rows from the corresponding block of rows. Thus step (c) as described earlier in this section would become a sequence of identical updates for each row from the corresponding block.

REFERENCES

1   I. Amemiya and T. Ando, Convergence of random products of contractions in Hilbert Space, *Acta. Sci. Math.* (*Szeged*) 26(MR 32):239–244 (1965).

2   K. A. Dines and R. J. Little, Computerized geophysical tomography, *Proc. IEEE* 67:1065–1073 (1979).

3   R. Bru, L. Elsner, and M. Neumann, Models of parallel chaotic iterations methods, *Linear Algebra Appl.* 102:175–192 (1988).

4   A. Berman and R. J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*, Academic, New York, 1979.

5   L. Elsner, I. Koltracht, and P. Lancaster, Convergence properties of the ART and SOR algorithms, submitted for publication.

6   C. W. Groetsch, *Generalized Inverses of Linear Operators*, Marcel Dekker, New York, 1977.

7   I. Halperin, The product of projection operators, *Acta Sci. Math.* (*Szeged*) 23:96–99 (1962).

8   M. Hanke and W. Niethammer, On the acceleration of Kaczmarz's method for inconsistent linear systems, submitted for publication.

9   G. T. Herman, *Image Reconstruction from Projections*, Academic, 1980.

10 S. Kaczmarz, Angenäherte Auflösung von Systemen linearer Gleichungen, *Bull. Internat. Acad. Pol. Sci. Lett. A*, 1937, pp. 355–357.

11 I. Koltracht, P. Lancaster, and D. Smith, The structure of some tomography matrices arising in geophysics, submitted for publication.

12 G. A. McMechan, Seismic tomography in boreholes, *Geophys. J. Roy. Astron. Soc.* 74:601–612 (1983).

13 V. Miller and M. Neumann, A note on comparison theorems for nonnegative iteration matrices, *Numer. Math.* 47:427–434 (1986).

14 F. Natterer, *The Mathematics of Computerized Tomography*, Wiley, New York, 1986.

15   S. Nelson and M. Neumann, Generalization of the projection method with applications to SOR method for Hermitian positive semidefinite linear systems, *Numer. Math.* 51:123–141 (1987).

16   K. Tanabe, Projection method for solving a singular system of linear equations and its applications, *Numer. Math.* 17:203–214 (1971).

17   D. C. Youla, Mathematical theory of image reconstruction by the method of convex projections in *Image Recovery: Theory and Applications* (H. Stark, Ed.), Academic, 1987, Chapter 2.