Monotonic Sequences and Rates of Convergence of Asynchronized Iterative Methods

L. Elsner*
Fakultät für Mathematik
Universität Bielefeld
Postfach 8640
4800-Bielefeld 1, Germany

and

M. Neumann†
Department of Mathematics
University of Connecticut
Storrs, Connecticut 06269-3009

Submitted by Hans Schneider

ABSTRACT

In a recent paper B. Vemmer and the authors investigated the effect of varying the number of processors on the rate of convergence of the asynchronized parallel block Jacobi method associated with monotone matrices. It was found that, under certain simplifying assumptions, increasing the number of processors in relation to the number of blocks (or, what comes to the same in more general settings, the number of iteration operators) slows down the convergence. One interpretation for these results given in that paper was that increasing the number of processors means that when the current global approximation is updated by a local approximation from one of the processors, that local approximation was computed from a "much" *earlier* global approximation received from the host node. Hence the slowdown in the rate of

---

*Research supported in part by Sonderforschungsbereich 343 *Diskrete Strukturen in der Mathematik*.

†Research supported in part by U.S. Air Force Grant AFOSR-88-0047 and by NSF Grants DMS-8901860 and DMS-9007030. Most of the work of this author was carried out while visiting the Sonderforschungsbereich 343 *Diskrete Strukturen in der Mathematik*, Fakultät für Mathematik at the Universität Bielefeld.

655 Avenue of the Americas, New York, NY 10010 0024-3795/93/$6.00

convergence. The principal purpose of this paper is to remove some of the simplifying assumptions that were made in the above-mentioned paper and to prove that many of the results there hold under much more general conditions. Our present assumptions do not yield a fixed iteration matrix which models the process as was the case previously. This means that different tools have to be developed to establish results comparing the rate of convergence of two asynchronized processes.

---

1. INTRODUCTION

In a recent paper [3] B. Vemmer and the authors investigated the effect of varying the number of processors on the rate of convergence of the asynchronized parallel block Jacobi method. It was found that, under certain simplifying assumptions, increasing the number of processors in relation to the number of blocks (or, what comes to the same in more general settings, the number of iteration operators) slows down the convergence. One interpretation for these results given in that paper was that increasing the number of processors means that when the current global approximation is updated by a local approximation from one of the processors, that local approximation was computed from a "much" *earlier* global approximation received from the host node. Hence the slowdown in the rate of convergence. The principal purpose of this paper is to remove some of the simplifying assumptions that were made in [3] and to prove that many of the results hold under much more general conditions.

The asynchronized model which we have in mind is as follows: Let

$$A = M_l - N_l, \qquad l = 1, \dots, m,$$

be regular splittings of the $n \times n$ monotone matrix A, that is, $M^{-1} \geqslant 0$, $N \geqslant 0$, and $A^{-1} \geqslant 0$; let E_l, $l = 1, \dots, m$, be nonnegative diagonal matrices whose sum is the identity; let $\{j_i\}_{i=1}^{\infty}$ denote a sequence of integers satisfying $1 \leqslant j_i \leqslant m$; let $\{r_i\}_{i=1}^{\infty}$ be a sequence of positive integers; and consider the problem of solving the linear system

$$Ax = b.$$

Then the *asynchronized* model is given by

$$x^{(i+r_i)} = \left(I - E_{j_i}\right)x^{(i+r_i-1)} + E_{j_i}M_{j_i}^{-1}\left(N_{j_i}x^{(i)} + b\right), \qquad i = 1, 2, \dots \quad (1.1)$$

A possible realization of the model can be achieved with a parallel machine with k processors and a host node in the following way: At time i a processor, call it for now the *subject* processor, which has just completed a previous task is assigned the task which is defined by the number j_i (i.e., defined by E_{j_i}, M_{j_i}, and N_{j_i}). This processor then locally calculates $u = E_{j_i} M_{j_i}^{-1}(N_{j_i} x + b)$. The number $r_i - 1 \geqslant 0$ then represents the number of similar tasks completed by other processors after time i and before the subject processor completes the computation of u. When the local computation is complete, the sum $x^{(i+r_i)} = (I - E_{j_i})x^{(i+r_i-1)} + u$ is formed by the host node, and the subject processor is assigned task $i + r_i$. We comment that it is possible to interpret (1.1) as saying that at time $i + r_i$ the host node is updated by a local approximation which was computed from a global approximation of r_i units of time ago.

The principal assumption that was made in [3] was that among all splittings, the amount of work per splitting is equal, so that it makes sense to suppose that the sequence $\{j_i\}_{i=1}^{\infty}$ is cyclic and, furthermore, that the number of processors updating the host processor before any given processor is ready with its local update is a constant equal to the number of processors less one. In this case our model takes on the following form:

$$x^{(i+k)} = \left(I - E_{j_i}\right)x^{(i+k-1)} + E_{j_i} M_{j_i}^{-1}\left(N_{j_i} x^{(i)} + b\right), \qquad i = 1, 2 \dots . \tag{1.2}$$

Using an idea of Mathias Pott, it was shown in [3] that if the weighting matrices E_l, $l = 1, \dots, m$, satisfy $E_l E_{l'} = \delta_{l,l'} E_l$, then with

$$B = \sum_{l=1}^{m} E_l M_l^{-1} N_l$$

and with $c = \sum_{l=1}^{m} E_l M_l^{-1} b$, the iteration (1.1) is equivalent to

$$x^{(i)} = \left(I - E_{j_i}\right)x^{(i-1)} + E_{j_i}\left(Bx^{(i-r_i)} + c\right). \tag{1.3}$$

This iteration is simpler to analyze because it involves working with just one operator. Let

$$e^{(i)} = x^{(i)} - A^{-1}b, \qquad i = 0, 1, \dots,$$

where $x^{(i)}$ is given in (1.1). It follows from [2] that $\lim_{i\to\infty} e^{(i)} = 0$. We define the rate of convergence of (1.3) as

$$\mathscr{R}(\{j_i\},\{r_i\}) := \sup_{x^{(0)}\in R^n} \limsup_{i\to\infty} \|e^{(i)}\|^{1/i}, \tag{1.4}$$

where $\|\cdot\|$ is any vector norm on R^n.

The specific result which was proved in [3] is as follows: Suppose that B is partitioned into the $m \times m$ block matrix

$$B = \begin{pmatrix} B_{11} & B_{12} & \cdots & B_{1m} \\ B_{21} & B_{22} & \cdots & B_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ B_{m1} & B_{m2} & \cdots & B_{mm} \end{pmatrix} \tag{1.5}$$

and the E_l's are chosen such that

$$E_l = \begin{pmatrix} 0 & & & & & \\ & \ddots & & & & \\ & & 0 & & & \\ & & & I_{n_l} & & \\ & & & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 0 \end{pmatrix}, \qquad l = 1,\ldots,m, \tag{1.6}$$

where n_l is the dimension of the lth diagonal block of B. Then on taking the sequence $\{j_i\}_{i=1}^{\infty}$ to be cyclic [viz., $j_i = i(\mathrm{mod}\ m) + 1\ \forall i \geqslant 1$], the sequence $\{s_i\}_{i=1}^{\infty}$ with $s_i = k\ \forall i \geqslant 1$, and the sequence $\{r_i\}_{i=1}^{\infty}$ with $r_i = k'\ \forall i \geqslant 1$ and with $k' \geqslant k$, we have that

$$\mathscr{R}(\{j_i\},\{s_i\}) \leqslant \mathscr{R}(\{j_i\},\{r_i\}). \tag{1.7}$$

Because of the assumption in [3] that the sequences $\{s_i\}_{i=1}$ and $\{r_i\}_{i=1}^{\infty}$ are constant, it was possible to associate an iteration matrix with each of the two processes and it was possible to prove (1.7) by comparing the spectral radii of the iteration matrices. In this paper we wish to raise the question to what extent the inequality in (1.7) continues to be true when we only assume that

$$1 \leqslant s_i \leqslant r_i \leqslant T, \qquad i = 1,2,\ldots. \tag{1.8}$$

An equivalent formulation of our question is as follows: Suppose we have two asynchronized parallel iteration processes such that one of the processes always uses results of local iterations which were based on more recent global approximations to update the current global approximation; then does it follow that this process has a better rate of convergence than the process which always uses updates based on local computations which always use older global approximations? We shall call the iteration process associated with the s_i's the *more frequent updating process* and we shall call the process associated with the r_i's the *more infrequent updating process*.

To consider our question it is useful to embed the asynchronized process (1.3), which takes place in the n-dimensional space, in the nT-dimensional space as follows: For the positive integers T and m given above and for integers $1 \leqslant t \leqslant T$ and $1 \leqslant s \leqslant m$, consider the $T \times T$ block matrix whose blocks are specified by

$$\left(A_s^{(t)}\right)_{k,l} = \begin{cases} I - E_s & \text{when} \quad k = 1, l = 1, \\ E_s B & \text{when} \quad k = 1, l = t, \\ I & \text{when} \quad k = l + 1, l = 1, \dots, T-1, \\ 0 & \text{otherwise}. \end{cases} \tag{1.9}$$

Displayed, $A_s^{(t)}$ has the form

$$A_s^{(t)} = \begin{pmatrix} \overbrace{I - E_s & 0 & \cdots & 0 & E_s B}^{t \text{ blocks}} & 0 & \cdots & 0 \\ I & 0 & \cdots & 0 & 0 & & & \vdots \\ 0 & I & \ddots & \vdots & \vdots & & & \vdots \\ \vdots & \ddots & \ddots & 0 & 0 & & & \vdots \\ 0 & \cdots & 0 & I & 0 & & & \vdots \\ 0 & \cdots & 0 & 0 & I & \ddots & & \vdots \\ \vdots & & & & \vdots & \ddots & 0 & \vdots \\ 0 & \cdot & \cdot & \cdot & 0 & \cdots & I & 0 \end{pmatrix}.$$

It is readily seen that $A_s^{(t)}$ is the map from R^{nT} to R^{nT} given by

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_T \end{pmatrix} \to \begin{pmatrix} (I - E_s) x_1 + E_s B x_t \\ x_1 \\ \vdots \\ x_{T-1} \end{pmatrix}, \tag{1.10}$$

where $x_i \in R^n$, $i = 1, \dots, T$.

Using the operators on R^{nT} introduced above, the asynchronized iterative process given in (1.3) can be described as a sequential iteration process in the nT-dimensional space as follows:

$$z^{(i)} = A_{j_i}^{(r_i)} z^{(i-1)}, \tag{1.11}$$

where $z^{(i)} = [(e^{(i)})^t \ \ldots \ (e^{(i-T+1)t}]^t$, in which a plain superscript t refers to the transpose.

Straight away let us give an example showing that without some restriction, the desired result on the better rate of convergence of the sequence which updates more frequently is not true generally: Let $n = 1$, $T = 3$, $E_1 = I = (1)$, $j_i = 1\ \forall i \geqslant 1$, and let $B = (\rho)$, $\rho < 1$. Next let

$$s_1 = 1, \qquad s_2 = 3, \qquad s_3 = 1, \qquad s_4 = 2, \qquad \text{and} \quad s_5 = 2$$

and

$$r_1 = 1, \qquad r_2 = 3, \qquad r_3 = 2, \qquad r_4 = 3, \qquad \text{and} \quad r_5 = 2.$$

For $i \geqslant 6$ let $s_i = s_{i(\mathrm{mod}\, 5)}$ and $r_i = r_{i(\mathrm{mod}\, 5)}$. For $w_0 \in R^3$ set

$$z^{(i)} = A_{j_i}^{(s_i)} \cdots A_{j_1}^{(s_1)} w_0, \qquad \forall i \geqslant 1,$$

and

$$y^{(i)} = A_{j_i}^{(r_i)} \cdots A_{j_1}^{(r_1)} w_0, \qquad \forall i \geqslant 1.$$

Then it can be ascertained that

$$z^{(5k)} = \begin{pmatrix} 0 & \rho^3 & 0 \\ 0 & \rho^2 & 0 \\ 0 & \rho^2 & 0 \end{pmatrix}^k w_0, \qquad \forall k \geqslant 1$$

and

$$y^{(5k)} = \begin{pmatrix} \rho^3 & 0 & 0 \\ \rho^2 & 0 & 0 \\ \rho^2 & 0 & 0 \end{pmatrix}^k w_0, \qquad \forall k \geqslant 1.$$

Thus

$$\mathscr{R}(\{j_i\},\{s_i\}) = \rho^{2/5} > \rho^{3/5} = \mathscr{R}(\{j_i\},\{r_i\}).$$

In Section 2 we develop our main results. We introduce subcones of R_+^{nT}, the nT-dimensional nonnegative orthant, with several features. One of the main characteristics of these cones is that their T subvectors of dimension n form a nondecreasing sequence of vectors as the index of the subvector increases. We use these cones to show the main result of this paper (Theorem 1). An example of the implications of this theorem is that if, in addition to (1.8),

$$s_{i+1} \leqslant s_i + 1, \qquad \forall i \geqslant 1, \tag{1.12}$$

then (1.7) always holds. What the result means is this: When $s_{i_0+1} > s_{i_0} + 1$ for some $i_0 \geqslant 1$, then the more frequent updating process uses an older global approximation to compute the $(i_0 + 1)$th iteration then the approximation it has used to compute the immediately preceeding i_0th iterate. Therefore condition (1.12) says that *when the more frequent updating process does not "suddenly" use an older approximation in computing some iterate than it has used in computing the previous one, the rate of convergence of the more frequent updating process is at least as good as the rate of convergence of the more infrequent updating process.* In subsequent results we associate with both the more frequent and infrequent iteration processes auxiliary processes which, regardless of the condition (1.12), are always such that the auxiliary process associated with the more frequent updates has a rate of convergence at least as favorable as the auxiliary process associated with the more infrequent updates. We conclude the paper with some upper bounds on the rate of convergence of (1.13).

For more background material on nonnegative matrices and on iterative methods with nonnegative iteration matrices see Berman and Plemmons [1] and Varga [4].

2. MAIN RESULTS

To study the rate of convergence of the iterative scheme (1.11) we introduce subcones of R_+^{nT} as follows: For each $k = 1,\ldots,T$ define

$$\mathscr{K}_T^{(k)} = \left\{ \left[(x_1)^t \quad \ldots \quad (x_T)^t \right]^t \in R_+^{nT} : x_1 \leqslant \cdots \leqslant x_T \text{ and } Bx_k \leqslant x_1 \right\}. \tag{2.1}$$

We begin by considering some basic properties of the $\mathscr{K}_T^{(k)}$'s.

LEMMA 1.

(i) $\mathscr{K}_T^{(T)} \subseteq \mathscr{K}_T^{(T-1)} \subseteq \cdots \subseteq \mathscr{K}_T^{(1)}$.
(ii) $A_s^{(t)} \mathscr{K}_T^{(t)} \subseteq \mathscr{K}_T^{(t+1)}$ $\forall 1 \leqslant t \leqslant T-1$ *and* $\forall 1 \leqslant s \leqslant m$.

Proof. The proof of (i) is an easy consequence of the definition of the cones $\mathscr{K}_T^{(t)}$, $t = 1, \dots, T$. To prove (ii) let $x = [(x_1)^t \ \dots \ (x_T)^t]^t \in \mathscr{K}_T^{(t)}$. Put $y = A_s^{(t)} x$, and partition y in conformity with x. Then, using (1.10), we have that

$$y_1 = (I - E_s)x_1 + E_s B x_t \leqslant (I - E_s)x_1 + E_s x_1 = x_1 = y_2 \leqslant y_3 \leqslant \cdots \leqslant y_T$$

and

$$B y_{t+1} = B x_t \leqslant (I - E_s)x_1 + E_s B x_t = y_1.$$

Hence $y \in \mathscr{K}_T^{(t+1)}$. ∎

With Lemma 1 in hand we can now make the following observation concerning the iteration in (1.11):

LEMMA 2. *Let* $\{j_i\}_{i=1}^{\infty}$ *be a sequence such that*

$$1 \leqslant j_i \leqslant m, \qquad i = 1, 2, \dots,$$

and let $\{t_i\}_{i=1}^{\infty}$ *be a sequence such that*

$$1 \leqslant t_i \leqslant T, \qquad i = 1, 2, \dots,$$

and

$$t_{i+1} \leqslant t_i + 1, \qquad i = 1, 2 \dots . \tag{2.2}$$

Then, beginning with a vector $z^{(0)} = x \in \mathscr{K}_T^{(t_1)}$, *the sequence of iterates generated by*

$$z^{(i)} = \mathbf{A}_{j_i}^{(t_i)} z^{(i-1)}, \qquad i = 1, 2, \dots,$$

has the following properties:

$$z^{(i)} \in \mathscr{K}_T^{(t_{i+1})}, \qquad i = 1, 2, \dots, \tag{2.3}$$

and

$$z^{(i)} \leqslant z^{(i-1)}, \qquad i = 1, 2 \dots . \tag{2.4}$$

Proof. The proof is by induction, the case $i = 1$ following immediately from the assumption that $z^{(0)} \in \mathscr{K}_T^{(t_1)}$. Suppose then that the result holds for $k = i - 1$, and we shall prove the result is true for $k = i$. Assume therefore that $z^{(i-1)} \in \mathscr{K}_T^{(t_i)}$. Then, by Lemma 1 (ii), $z^{(i)} = A_{j_i}^{(t_i)} z^{(i-1)} \in \mathscr{K}_T^{(t_i+1)}$, so that, because of (2.2) and by Lemma 1 (i), $z^{(i)} \in \mathscr{K}_T^{(t_{i+1})}$. Thus we have proved (2.3). To show (2.4) note first that from the fact that $z^{(i)} = A_{j_i}^{(t_i)} z^{(i-1)}$ we have that $(z^{(i)})_1 = (I - E_{j_i})(z^{(i-1)})_1 + E_{j_i} B(z^{(i-1)})_{t_i} \leqslant (z^{(i)})_2 = (z^{(i-1)})_1$. Next, for $j = 2, \dots, T$, the inequalities $(z^{(i)})_j \leqslant (z^{(i-1)})_j$ follow from (1.10) and because $z^{(i-1)} \in \mathscr{K}_T^{(t_i)}$. ∎

In Section 1 we gave an example of two asynchronized processes, one always with more frequent updates than the other, such that the process with more infrequent updates has faster convergence. We therefore raise the question: Under what additional conditions on the time lags in the process with the more frequent updates are we guaranteed a better rate of convergence than that of the process with more infrequent updates? Our main result leading to subsequent conclusions in this direction is the following:

THEOREM 1. *Let $\{j_i\}_{i=1}^{\infty}$ be a sequence such that $1 \leqslant j_i \leqslant m$, and let $\{s_i\}_{i=1}^{\infty}$, $\{t_i\}_{i=1}^{\infty}$, and $\{r_i\}_{i=1}^{\infty}$ be sequences such that*

$$1 \leqslant s_i \leqslant t_i \leqslant r_i \leqslant T, \qquad i = 1, 2, \dots, \tag{2.5}$$

and such that the elements of the sequence $\{t_i\}_{i=1}^{\infty}$ satisfy (2.2). For any $x^{(0)} \leqslant z^{(0)} \leqslant y^{(0)}$, with $z^{(0)} \in K_T^{(r_1)}$ define the iterative sequences

$$x^{(i)} = A_{j_i}^{(s_i)} x^{(i-1)}, \quad z^{(i)} = A_{j_i}^{(t_i)} z^{(i-1)}, \text{ and } y^{(i)} = A_{j_i}^{(r_i)} y^{(i-1)}, \qquad i = 1, 2 \dots .$$

Then

$$x^{(i)} \leqslant z^{(i)} \leqslant y^{(i)}, \qquad i = 1, 2 \dots . \tag{2.6}$$

Proof. Note first that as $z^{(0)} \in \mathscr{K}_T^{(r_1)}$, it is also in $\mathscr{K}_T^{(t_1)}$. Thus, according to Lemma 2, (2.3) and (2.4) hold for the sequence $\{z^{(i)}\}_{i=1}^{\infty}$. We prove (2.6) by induction, the case $i = 1$ being trivially true from the fact that $s_1 \leqslant t_1 \leqslant r_1$ and (1.10). Assume that

$$x^{(i-1)} \leqslant z^{(i-1)}.$$

To show that $x^{(i)} \leqslant z^{(i)}$ it suffices to show that $(x^{(i)})_1 \leqslant (z^{(i)})_1$. But, using the inductive hypothesis and the facts that $s_i \leqslant t_i$ and $z^{(i-1)} \in \mathscr{X}_T^{(t_i)}$, we have on close inspection that

$$\begin{aligned}(z^{(i)})_1 - (x^{(i)})_1 &= (I - E_{j_i})\left[(z^{(i-1)})_1 - (x^{(i-1)})_1\right] \\ &\quad + E_{j_i}B\left[(z^{(i-1)})_{t_i} - (x^{(i-1)})_{s_i}\right] \\ &= (I - E_{j_i})\underbrace{\left[(z^{(i-1)})_1 - (x^{(i-1)})_1\right]}_{\geqslant 0 \text{ by induction}} \\ &\quad + E_{j_i}B\underbrace{\left[(z^{(i-1)})_{t_i} - (z^{(i-1)})_{s_i}\right]}_{\geqslant 0 \text{ because } z^{(i-1)} \in \mathscr{X}_T^{(t_i)}} \\ &\quad + E_{j_i}B\underbrace{\left[(z^{(i-1)})_{s_i} - (x^{(i-1)})_{s_i}\right]}_{\geqslant 0 \text{ by induction}} \geqslant 0.\end{aligned}$$

The proof that $z^{(i)} \leqslant y^{(i)}$ for all $i \geqslant 1$ follows similarly. ∎

A natural question to ask at this point is: Under which conditions on the sequences $\{s_i\}_{i=1}^{\infty}$ and $\{r_i\}_{i=1}^{\infty}$ with $s_i \leqslant r_i$, $i = 1, 2, \dots,$ does there exist a sequence $\{t_i\}_{i=1}^{\infty}$ satisfying (2.5) and (2.2)? We can prove the following characterization:

THEOREM 2. *Let $\{s_i\}_{i=1}^{\infty}$ and $\{r_i\}_{i=1}^{\infty}$ be two given sequences of real numbers. Then the following are equivalent:*

(i) *There exists a sequence $\{\tau_i\}_{i=1}^{\infty}$ satisfying $s_i \leqslant t_i \leqslant r_i$ and $t_{i+1} \leqslant t_i + 1$ $\forall i \geqslant 1$.*

(ii) *For $i \geqslant 1$ and for $0 \leqslant k < i$, one has $s_i \leqslant r_{i-k} + k$.*

In this case, the sequence $\{\tau_i\}_{i=1}^{\infty}$ defined either by

$$\tau_i = \min_{0 \leqslant k \leqslant i-1} (r_{i-k} + k) \tag{2.7}$$

or, recursively, by

$$\tau_1 = r_1 \quad \textit{and} \quad \tau_{i+1} = \min_{i \geqslant 1} \{r_{i+1}, \tau_i + 1\} \tag{2.8}$$

satisfies the same requirements that the sequence $\{t_i\}_{i=1}^{\infty}$ satisfies in (i), namely, $s_i \leqslant \tau_i \leqslant r_i$, $\tau_{i+1} \leqslant \tau_i + 1$, and it also satisfies $t_i \leqslant \tau_i$ $\forall i \geqslant 1$.

Proof. If the conditions of (i) hold, then for any $i \geqslant 1$ and for any $0 \leqslant k \leqslant i-1$, we have $s_i \leqslant t_i \leqslant t_{i-1}+1 \leqslant \cdots \leqslant t_{i-k}+k \leqslant r_{i-k}+k$, and so (ii) holds.

Suppose now that (ii) holds, and define the sequence $\{u_i\}_{i=1}^{\infty}$ as in (2.7), viz., $u_i = \min_{0\leqslant k \leqslant i-1}(r_{i-k}+k)$, $i \geqslant 1$. Then $s_i \leqslant u_i \leqslant r_i$ $\forall i \geqslant 1$, and, as $u_{i-1} = r_{i-s-1}+s$ for some $0 \leqslant s \leqslant i-2$, one has

$$u_{i-1}+1 \geqslant \min_{0\leqslant k\leqslant i-1}\{r_{i-k}+k\} = u_i.$$

Thus the sequence $\{u_i\}_{i=1}^{\infty}$ satisfies all the requirements of (i).

Let $\{t_i\}_{i=1}^{\infty}$ satisfy $s_i \leqslant t_i \leqslant r_i$ and $t_{i+1} \leqslant t_i + 1$ for all $i \geqslant 1$. Then, as above, $t_i \leqslant r_{i-k}+k$ for $1 \leqslant k < i$, and hence $t_i \leqslant u_i$. To show that $u_i = \tau_i$, where τ_i is given by (2.8), we see at once that this sequence satisfies (2.2). Hence $\tau_i \leqslant u_i$. To show that $u_i \leqslant \tau_i$ we proceed by induction. From $u_i \leqslant \tau_i$ we have $u_{i+1} \leqslant u_i + 1 \leqslant \tau_i + 1$ and, as $u_{i+1} \leqslant r_{i+1}$, we get that $u_{i+1} \leqslant \tau_{i+1}$. Hence $u_i = \tau_i$ for all $i \geqslant 1$. ∎

REMARK. The preceding result can be interpreted in the following way. The sequence $\{\tau_i\}_{i=1}^{\infty}$ is the maximal sequence satisfying (2.2) not exceeding the sequence $\{r_i\}_{i=1}^{\infty}$. One can also give the minimal sequence above $\{s_i\}_{i=1}^{\infty}$ satisfying (2.2), namely,

$$v_i = \sup_{j\geqslant 0}\{s_{i+j}-j\} \qquad \forall i \geqslant 1,$$

which is well defined. The condition $r_i \geqslant v_i$, $i \geqslant 1$, is the same as that in Theorem 2 (ii) and can be phrased in the more symmetric manner

$$i \geqslant j \quad\Rightarrow\quad s_i - r_j \leqslant i-j. \tag{2.9}$$

We shall now assume that the sequence $\{j_i\}_{i=1}^{\infty}$ is, in the language of Bru, Elsner, and Neumann [2], a *regulated sequence on m integers*. This means that each of the integers $1,\dots,m$ appears at least once every S consecutive elements, viz.,

$$\{1,2,\dots,m\} = \{j_i,\dots,j_{i+S-1}\} \qquad \forall i \geqslant 1.$$

For the iteration process (1.3) with B given in (1.5), the condition means that in the computation of the iterates, each one of the block rows of the block Jacobi iteration matrix B of (1.5) is used at least once every S consecutive iterations. We comment that for this reason S was called in [2] the *computation cycle of the asynchronized process*. Under the condition of regularity it will be shown in Section 3 that for any vector $z^{(0)} \in R^{nT}$ and any sequence $\{t_i\}_{i=1}^{\infty}$ with $1 \leqslant t_i \leqslant T$, the sequence of vectors given by

$$z^{(i)} = A_{j_i}^{(t_i)} z^{(i-1)}, \qquad i = 1, 2, \dots, \tag{2.10}$$

satisfies

$$\lim_{i \to \infty} z^{(i)} = 0.$$

Let us introduce the rate of convergence of this iterative sequence as follows: Given a vector norm $\|\cdot\|$ on R^{nT}, the rate of convergence of (2.10) is given by

$$\tilde{\mathscr{R}}(\{j_i\}, \{t_i\}) = \sup_{z^{(0)} \in R^{nT}} \limsup_{i \to \infty} \|z^{(i)}\|^{1/i}. \tag{2.11}$$

The rate of convergence of the asynchronized iteration (1.3) given in (1.4) and the rate of convergence of the iteration (2.10) given in (2.11) can be analyzed to show that

$$\tilde{\mathscr{R}}(\{j_i\}, \{t_i\}) = \mathscr{R}(\{j_i\}, \{t_i\}). \tag{2.12}$$

To conclude that under the conditions of Theorem 1 the rate of convergence of the more frequent updating process is at least as favorable as that of the more infrequent updating process, we can assume, without loss of generality, that our nonnegative matrix B (with $\beta := \rho(B) < 1$) is irreducible. Otherwise consider $B_\epsilon = B + \epsilon J$, where J is the $n \times n$ matrix of all 1's, and let $\epsilon \downarrow 0$. Let x be a positive Perron vector of B, and construct the positive nT-vector

$$\tilde{x} = [\, \beta x^t \quad \underbrace{x^t \quad \cdots \quad x^t}_{T-1 \text{ times}} \,]^t. \tag{2.13}$$

Then $\tilde{x}$ induces the monotonic vector norm $\|\cdot\|_{\tilde{x}}$ given by

$$\|z\|_{\tilde{x}} = \inf\{\alpha > 0 \mid -\alpha\tilde{x} \leqslant z \leqslant \alpha\tilde{x}\}, \qquad z \in R^{nT}.$$

In particular it follows that for any vector $z \in R^{nT}$,

$$|z| \leqslant \|z\|_{\tilde{x}}\tilde{x},$$

where $|z|$ denotes the nonnegative nT-vector whose entries are the absolute values of the corresponding entries of z. Observe that $\tilde{x} \in \mathscr{K}_T^{(t)}$ $\forall 1 \leqslant t \leqslant T$.

Now let $\omega^{(i)} = A_{j_i}^{(t_i)} \cdots A_{j_1}^{(t_1)}\tilde{x}$ $\forall i \geqslant 1$. Then it can be readily ascertained that for $\tilde{\mathscr{R}}(\{j_i\},\{t_i\})$ given in (2.11),

$$\tilde{\mathscr{R}}(\{j_i\},\{t_i\}) = \limsup_{i\to\infty}\|\omega^{(i)}\|_{\tilde{x}}^{1/i}. \tag{2.14}$$

As a consequence of Theorem 1 we obtain:

THEOREM 3. *Let* $\{j_i\}_{i=1}^{\infty}$ *be a regulated sequence on m integers, and let* $\{s_i\}_{i=1}^{\infty}$ *and* $\{r_i\}_{i=1}^{\infty}$ *be two sequences satisfying* (2.9). *Then*

$$\tilde{\mathscr{R}}(\{j_i\},\{s_i\}) \leqslant \tilde{\mathscr{R}}(\{j_i\},\{r_i\}).$$

The results of [3] which we described in the introduction are a special case of Theorem 3 which we state here as follows:

COROLLARY 1 (Elsner, Neumann, and Vemmer [3, Theorem 1]). *Let* $\{j_i\}_{i=1}^{\infty}$ *be the sequence defined by* $j_i = i(\text{mod } m) + 1$ $\forall i \geqslant 1$, *and let* $\{s_i\}_{i=1}^{\infty}$ *and* $\{r_i\}_{i=1}^{\infty}$ *be sequences such that*

$$k = s_i \leqslant r_i = k' \qquad \forall i \geqslant 1.$$

Then

$$\tilde{\mathscr{R}}(\{j_i\},\{s_i\}) \leqslant \tilde{\mathscr{R}}(\{j_i\},\{r_i\}).$$

Prior to Theorem 1 we explained that without the main condition—(2.2)—on which it relies, the inference that the asynchronized process which updates more frequently has at least as favorable a rate of convergence as the asynchronized process which updates less frequently is generally untrue. Therefore it is of interest, if only a theoretical one, to note that from both processes auxiliary iteration schemes can be derived such that regardless of whether (2.2) holds, the auxiliary iteration schemes derived from the more frequent updating process have at least as favorable convergence rates as the auxiliary iteration schemes derived from the more infrequent updating process. Here we shall give examples of two different auxiliary iterations.

DEFINITION 1. Let $\{j_i\}_{i=1}^{\infty}$ be a regulated sequence on m integers, and let $\{t_i\}_{i=1}^{\infty}$ be sequences such that $1 \leqslant t_i \leqslant T$ $\forall i \geqslant 1$. Define the sequence $\{t_i'\}_{i=1}^{\infty}$ by

$$t_1' = t_1 \quad \text{and} \quad t_i' = \min(t_i, t_{i-1}' + 1), \quad i = 2, 3, \ldots .$$

Let $z^{(0)} \in \mathscr{K}_T^{(t_1)}$. The *first auxiliary rate of convergence* of the iterative scheme $z^{(i)} = A_{j_i}^{(t_i)} z^{(i-1)}$ at $z^{(0)}$ is given by

$$\hat{\mathscr{R}}(\{j_i\}, \{t_i\}, z^{(0)}) = \limsup_{i \to \infty} \|\zeta^{(i)}\|^{1/i},$$

where $\zeta^{(0)} = z^{(0)}$ and

$$\zeta^{(i)} = A_{j_i}^{(t_i')} \zeta^{(i-1)}, \qquad i = 1, 2, \ldots .$$

The following is an immediate corollary to Theorem 2:

COROLLARY 2. *Let $\{j_i\}_{i=1}^{\infty}$ be a regulated sequence on m integers, and let $\{s_i\}_{i=1}^{\infty}$ and $\{r_i\}_{i=1}^{\infty}$ be sequences such that $1 \leqslant s_i \leqslant r_i \leqslant T$. Then for any $z^{(0)} \in \mathscr{K}_T^{(r_1)}$,*

$$\hat{\mathscr{R}}(\{j_i\}, \{s_i\}, z^{(0)}) \leqslant \hat{R}(\{j_i\}, \{r_i\}, z^{(0)}).$$

Proof. From the sequences $\{s_i\}_{i=1}^{\infty}$ and $\{r_i\}_{i=1}^{\infty}$ obtain the sequences $\{s_i'\}_{i=1}^{\infty}$ and $\{r_i'\}_{i=1}^{\infty}$ in exactly the same manner in which the sequence $\{t_i'\}_{i=1}^{\infty}$ was obtained from the sequence $\{t_i\}_{i=1}^{\infty}$ in Definition 1. The result now follows by Theorem 1. ∎

REMARK. Observe that the auxiliary rate of convergence as defined above is local in the sense that it is defined at a vector in the appropriate cone. The relations between the true rates of convergence of the chaotic schemes determined by the sequences of time lags $\{s_i\}_{i=1}^{\infty}$ and $\{r_i\}_{i=1}^{\infty}$ and the local rates given in the above corollary are that $\hat{\mathscr{R}}(\{j_i\}, \{s_i\}, z^{(0)}) \leqslant \mathscr{R}(\{j_i\}, \{s_i\})$ and $\hat{\mathscr{R}}(\{j_i\}, \{r_i\}, z^{(0)}) \leqslant \tilde{\mathscr{R}}(\{j_i\}, \{s_i\})$.

There is a second possibility for defining the auxiliary rate of convergence.

DEFINITION 2. Let $\{j_i\}_{i=1}^{\infty}$ and $\{t_i\}_{i=1}^{\infty}$ be sequences such that $1 \leqslant j_i \leqslant m$ and $1 \leqslant s_i \leqslant T$, $i = 1, 2 \dots$. For $\eta^{(0)} = z^{(0)} \in \mathscr{K}_T^{(s_1)}$ define the sequence

$$\eta^{(i)} = \begin{cases} A_{j_i}^{(t_i)}\eta^{(i-1)} & \text{if} \quad t_i \leqslant t_{i-1} + 1, \\ A_{j_i}^{(t_i)}A_{j_{i-1}}^{(T-1)} \cdots A_{j_{i-1}}^{(t_{i-1}+1)}\eta^{(i-1)} & \text{if} \quad t_i > t_{i-1} + 1. \end{cases} \tag{2.15}$$

The *second auxiliary rate of convergence* of the iterative scheme $z^{(i)} = A_{j_i}^{(t_i)} z^{(i-1)}$ at $z^{(0)}$ is given by

$$\bar{\mathscr{R}}(\{j_i\}, \{t_i\}, z^{(0)}) = \limsup_{i \to \infty} \|\eta^{(i)}\|^{1/i}.$$

Using Lemma 2 one can ascertain upon inspection that for the second auxiliary rate of convergence one has the following comparison result:

COROLLARY 3. *Let $\{j_i\}_{i=1}^{\infty}$ be a regulated sequence on m integers, and let $\{s_i\}_{i=1}^{\infty}$ and $\{r_i\}_{i=1}^{\infty}$ be sequences such that $1 \leqslant s_i \leqslant r_i \leqslant T$ $\forall i \geqslant 1$. Then for any $z^{(0)} \in \mathscr{K}_T^{(r_1)}$,*

$$\bar{\mathscr{R}}(\{j_i\}, \{s_i\}, z^{(0)}) \leqslant \bar{\mathscr{R}}(\{j_i\}, \{r_i\}, z^{(0)}).$$

This last corollary has the following implication: For any $1 \leqslant t \leqslant T$ and $1 \leqslant s \leqslant m$ consider in conjunction with our usual operator $A_s^{(t)}$ the operator

$$B_s^{(t)} = \begin{cases} A_s^{(T-1)}A_s^{(T-2)} \cdots A_s^{(t+1)} & \text{if} \quad t < T-1, \\ I & \text{otherwise}. \end{cases}$$

Then we have the following outcome:

COROLLARY 4. *Let $\{j_i\}_{i=1}^{\infty}$ be a sequence such that $1 \leqslant j_i \leqslant m$ $\forall i \geqslant 1$, and let $\{s_i\}_{i=1}^{\infty}$ and $\{r_i\}_{i=1}^{\infty}$ be sequences such that $1 \leqslant s_i \leqslant r_i \leqslant T$ $\forall i \geqslant 1$. Then*

$$\limsup_{i \to \infty} \left\| \left(B_{j_i}^{(s_i)} A_{j_i}^{(s_i)} \right) \cdots \left(B_{j_1}^{(s_1)} A_{j_1}^{(s_1)} \right) \right\|^{1/m_i}$$

$$\leqslant \limsup_{i \to \infty} \left\| \left(B_{j_i}^{(r_i)} A_{j_i}^{(r_i)} \right) \cdots \left(B_{j_1}^{(r_1)} A_{j_1}^{(r_1)} \right) \right\|^{1/n_i},$$

where

$$m_i = \sum_{k=1}^{i} (T - s_k) \text{ and } n_i = \sum_{k=1}^{i} (T - r_k), \qquad i = 1,2\dots.$$

3. AN UPPER BOUND ON THE RATE OF CONVERGENCE

In this section we develop an upper bound for the rate of convergence of the iteration (1.3). The proof of this upper bound is an adaptation of the convergence proof given in [2, Theorem 2.2]. We remark that the assumptions and the results in that paper and those given here are not quite comparable.

THEOREM 4. *Consider the iteration* (1.3), *where* $\{j_i\}_{i=1}^{\infty}$ *is a regulated sequence on m integers with computation cycle S and*

$$1 \leqslant r_i \leqslant T, \qquad i = 1,2\dots. \tag{3.1}$$

Then

$$\mathscr{R}(\{j_i\},\{r_i\}) \leqslant \beta^{\frac{1}{T+S-1}}, \tag{3.2}$$

where $\beta = \rho(B) < 1$.

Proof. Without loss of generality we can assume that B is irreducible and hence has a positive eigenvector x corresponding to β. Consider the vector $\tilde{x}$ given in (2.13), and define the vectors ω_ν by

$$\omega_0 = \tilde{x} \quad \text{and} \quad \omega_\nu = A_{j_\nu}^{(r_\nu)}\omega_{\nu-1}, \quad \nu = 1,2\dots.$$

We claim that

$$\omega_{T+S-1} \leqslant \beta\tilde{x}. \tag{3.3}$$

To see this, partition the ω_ν's in conformity with the vector $\tilde{x}$. Thus $\omega_\nu = [(\omega_\nu)_1^t \ \dots \ (\omega_\nu)_T^t]^t$. We now prove by induction that

$$(\omega_\nu)_s \leqslant \beta x, \ s = 1,\dots,\nu+1, \text{ and } (\omega_\nu)_s \leqslant x, \ s \leqslant T, \qquad \nu \geqslant 0. \tag{3.4}$$

This is obvious for $\nu = 0$. Let $\nu > 0$ and let (3.4) be satisfied for all indices less than ν. Then

$$\begin{aligned}(\omega_\nu)_1 &= (I - E_{j_\nu})(\omega_{\nu-1})_1 + E_{j_\nu} B(\omega_{\nu-1})_{r_\nu} \\ &\leqslant (I - E_{j_\nu})\beta x + E_{j_\nu} Bx = \beta x \leqslant x,\end{aligned}$$

while because $(\omega_\nu)_{s+1} = (\omega_{\nu-1})_s$, we have the remaining inequalities of (3.4). Now let $\mu \geqslant T$. Then $(\omega_{\mu-1})_{r_\mu} \leqslant \beta x$ by (3.4), and hence

$$(\omega_\mu)_1 \leqslant (I - E_{j_\mu})(\omega_{\mu-1})_1 + E_{j_\mu} \beta^2 x. \tag{3.5}$$

Let $1 \leqslant k \leqslant n$. If for some ν, $T \leqslant \nu \leqslant \mu$, we have $(E_{j_\nu})_{k,k} = 1$, then we can conclude by (3.5) that $((\omega_\mu)_1)_k \leqslant \beta^2 x_k$. Hence by the definition of the computational cycle,

$$(\omega_{T+S-1})_1 \leqslant \beta^2 x,$$

as $\{j_T, \ldots, j_{T+S-1}\} = \{1, \ldots, m\}$. By (3.4), $(\omega_{T+S-1})_s \leqslant \beta x$ for all remaining indices s, and hence (3.3) is true. Finally, (3.2) follows from (3.3), (2.14), and (2.12). ∎

We remark that for $S = 1$ and $r_i = T$, $i \geqslant 1$, the bound given in (3.2) is $\beta^{1/T}$ and is therefore sharp. We finally remark that if we consider two sequences of time lags $\{s_i\}_{i=1}^\infty$ and $\{r_i\}_{i=1}^\infty$ such that $s_i \leqslant r_i$, $i \geqslant 1$, then while the upper bound for the convergence rate of the more frequently updating sequence does not exceed the corresponding bound for the more infrequently updating sequence, the counterexample of Section 1 shows that this does not always reflect the exact situation.

REFERENCES

1 A. Berman and R. J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*, Academic, New York, 1979.
2 R. Bru, L. Elsner, and M. Neumann, Models of parallel chaotic iteration methods, *Linear Algebra Appl.* 103:175–192 (1988).
3 L. Elsner, M. Neumann, and B. Vemmer, The effect of the number of processors on the convergence rate of the parallel block Jacobi method, *Linear Algebra Appl.* 154–156:311–330 (1991).
4 R. S. Varga, *Matrix Iterative Analysis*, Prentice-Hall, Englewood Cliffs, N.J., 1962.

Received 3 December 1991; final manuscript accepted 30 December 1991