On the Convergence of Parallel Asynchronous Block-Iterative Computations

E. Kaszkurewicz and A. Bhaya*
Nucleo de Processamento Paralelo
COPPE
Universidade Federal do Rio de Janeiro
Caixa Postal 68504
21945 Rio de Janeiro, RJ Brasil

and

D. D. Šiljak**
EECS Department
Santa Clara University
Santa Clara, California 95053

Submitted by David H. Carlson

ABSTRACT

This paper considers the convergence problem of parallel asynchronous block-iterative computation schemes. A new mathematical state-space model for a class of nonlinear time-varying parallel iterative schemes is proposed. Using this model, which generalizes several models of the Chazan-Miranker type, together with large-scale systems and Liapunov techniques, it is shown that the well-known quasidominance condition on a certain aggregated matrix guarantees exponential convergence of this class of methods.

1. INTRODUCTION

The recent emergence of vector, array, and parallel processors with different high-speed low-cost alternative structures is provoking significant

*Research supported by the project "Computação Paralela," FINEP Grant No. 5 2 88 06 22 00.
**Research supported by NSF Grant ECS-8813273

© Elsevier Science Publishing Co., Inc., 1990
655 Avenue of the Americas, New York, NY 10010 0024-3795/90/$3.50

changes in languages, programming techniques, and algorithms in general. In scientific computing in particular, the different ways that numerical methods are conceived or adapted to these new machines and architectures are stimulating new theoretical and applied research. Complexity, convergence, load balancing, and comparative speedup are examples of such problems. It is worthy of remark that many old ideas and techniques that were abandoned in the past, due to the lack of sufficient computing power, are now being resuscitated. Examples are the ideas of diakoptics, hierarchical methods, and iterative methods for the solution of large sets of algebraic equations.

More specifically, in order to solve the classical linear equation $Ax = b$ when the dimension of A is of the order of many thousands of variables, Gaussian elimination combined with sparsity techniques is currently considered to be one of the most efficient methods for sequential machines. However, iterative techniques such as the Jacobi, Gauss-Seidel, SOR, and conjugate-gradient, which exhibit slow convergence in sequential machines, are performing quite well in parallel processors.

Typically, for large problems, iterative methods appear to be more effective than direct solvers. Most direct solution techniques do not have an inherently parallel structure and require a considerable amount of data communication among processors when implemented in parallel. Iterative methods, on the other hand, require less data communication and lend themselves to concurrent implementation [see references in Nour-Omid et al. (1987)], this being partly due to the inherent parallelism frequently associated with such methods (e.g., the Jacobi iterative method is essentially parallel). These facts justify the reevaluation of the classical iterative methods implemented in parallel processors, particularly with regard to convergence conditions as well as speed of convergence.

A parallel (or distributed) implementation of an iterative algorithm is one in which the computational load is shared by several processors while coordination is maintained by information exchange via communication links. Broadly speaking, there are two modes of operation of such implementations: synchronous and asynchronous [for a more detailed discussion of synchronism versus asynchronism, see Kung (1976)]. In the first, the point of departure is some iterative algorithm which is guaranteed to converge to the correct solution under the usual circumstances of sequential centralized computation in a single processor. The computational load of the typical iteration is then divided in some way (for example, pointwise or blockwise) among the available processors, and it is assumed that the processors exchange all necessary information regarding the outcomes of the current iteration before a new iteration can begin. Such synchronous parallel algorithms have two obvious implementation disadvantages: the need to program initiation of the algorithm and synchronization of the iterations, which is a nontrivial task for a large-scale computation; and the fact that the speed of computation is limited to that of the slowest processor, so that the faster processors spend considerable amounts of time in an idle or wait status.

The second mode of operation is asynchronous: computation and communication are carried out at the various processors completely independently of the progress in other processors. This extreme model of complete independence of each processor was introduced by Chazan and Miranker (1969) (for point iterative schemes) as a generalization of free steering, introduced by Ostrowski (1955), and was motivated by the numerical simulations of Rosenfeld (1967). Many researchers subsequently proposed various schemes —some hybrid, some nonlinear—ending with Baudet (1978) (see references therein), who surveyed earlier results and proposed a generalization of the Chazan-Miranker convergence criterion to the nonlinear case. The next important development, due to El Tarazi (1982) and Talukdar et al. (1983), was the introduction of a descriptive model for a class of asynchronous algorithms and the enunciation and proof of the first correct convergence conditions in the nonlinear case. Subsequently, Lubachevsky and Mitra (1986) introduced a variation and generalization of the Chazan-Miranker model and gave a convergence condition for a specific application: asynchronous computation of the stationary distribution of a Markov chain. Finally, Bru et al. (1988) introduced two models of parallel chaotic (asynchronous) linear iteration methods which are generalizations (block versions) of the Chazan-Miranker model; they also gave convergence conditions for asynchronous parallel iterative multisplitting algorithms.

In this paper a new mathematical state-space model for a class of nonlinear time-varying parallel asynchronous block-iterative schemes is proposed. This model, which generalizes several of those mentioned above, permits the application of Liapunov techniques to obtain sufficient conditions for the exponential convergence of the class of algorithms studied. Liapunov techniques have not been used in the analysis of convergence of asynchronous iterative numerical methods. In the synchronous/sequential context, the generalized distance functions used by Polak (1971), Poljak (1982), etc. can be regarded as Liapunov functions, as has been pointed out by Bertsekas (1983).

Traditionally, in order to treat convergence problems of iterative and block-iterative methods, such as the Jacobi, Gauss-Seidel, and SOR, that are linear and time-invariant, techniques based on the Perron-Frobenius and Gerschgorin theorems [Varga (1962), Young (1971)] and H-matrices [Robert (1969)] have been used. When disturbances caused by finite arithmetic and variable delays introduced by asynchronisms and memory access times etc. are considered, the models become nonlinear and time-varying. We show that the above mentioned techniques, combined with Liapunov techniques, can still be used to find convergence conditions for a class of nonlinearities.

The usual spectral-radius convergence condition for sequential linear stationary or parallel synchronous schemes does not guarantee convergence when asynchronous parallel computation schemes are used. A well-known and stronger convergence condition for classical block-iterative schemes such as the Jacobi and Gauss-Seidel methods to solve $Ax = b$ is the diagonal dominance of the matrix A. It is shown in this paper, using the Liapunov approach, that even when the block-iterative scheme is nonlinear, time-varying, and implemented in parallel using asynchronous communication, a related condition (quasidominance of an aggregate error matrix) still guarantees the convergence of this scheme

2. MATHEMATICAL MODELS FOR BLOCK-ITERATIVE PARALLEL COMPUTATION

It is well known [Ortega and Rheinboldt (1970)] that the problem of solving the equation

$$f(x) = 0, \tag{1}$$

where $f: \mathbb{R}^n \to \mathbb{R}^n$ is sufficiently smooth, can be transformed into the problem of finding the fixed points of an iteration that is written as follows:

$$x^{(k+1)} = g(x^{(k)}, k), \quad k = 0,1,2,\ldots, \qquad \text{and } x^{(0)} \text{ given}, \tag{2}$$

where $g: \mathbb{R}^n \times N^* \to \mathbb{R}^n$. Finding fixed points x^* that satisfy $x^* = g(x^*, k)$, $k = 0,1,2,\ldots$, is equivalent to solving (1). Note also that the classical iterative methods to solve the linear equation $Ax = b$, such as the Jacobi, Gauss-Seidel, and SOR, correspond to the choice of $f(x) = Ax - b$ in (1) and appropriate choice of g in (2).

Since we are concerned with the parallel asynchronous computation of solutions of large systems of equations, it is natural to consider block-partitioned versions of (2). Finally, we view (2) as a (large-scale) discrete-time dynamic system: thus convergence of the iterative process to the solution of (1) is implied by the asymptotic stability of the discrete-time system. In order to be able to use the powerful arsenal of Liapunov techniques to solve the latter problem, we propose a state-space model for a general block-iterative asynchronous computation scheme.

Consider, first, the general (synchronous) block-iterative scheme described by the equation

$$z_i^{(k+1)} = \sum_{j=1}^{n} H_{ij}(z^{(k)}, k) z_j^{(k)}, \qquad k = 0,1,2,\dots, \quad i = 1,2,\dots,n, \quad (3)$$

where $z_i^{(k)} \in \mathbb{R}^{n_i}$ $\forall k$, $N := \sum_{i=1}^{n} n_i$, $z^{(k)T} := (z_1^{(k)T}, \dots, z_n^{(k)T}) \in \mathbb{R}^{1\times N}$, and $H_{ij}(z^{(k)}, k) \in \mathbb{R}^{n_i \times n_j}$ $\forall k$.

Clearly, if $z^{(k)} := x^{(k)} - x^*$ [where x^* is the fixed point of (2) and the solution of (1)], then (3) may be thought of as the error equation of a time-varying block-iterative process, and convergence of this process is implied by the asymptotic stability of the zero solution of the discrete-time system described by (3). The term synchronous is used in the following sense: the evaluation of each of the n subvectors $z_i^{(k+1)}$ is assigned, at time k, to one of n processors, and the exchange of subvectors between processors is synchronized to occur at the same time instant (clock pulse), before a new iteration can begin.

When the communication between processors is not synchronous, the model (3) is no longer valid. There are various ways to model asynchronous computations, and one simple general model is the so-called ULA/D (use the latest available/data), terminology introduced in Talukdar et al. (1983). A simple way to represent this model is as follows:

$$x_i^{(k+1)} = g_i(x^{(*)}, k), \quad (4)$$

where $x^{(*)}$ is the state vector composed of the latest available subvectors, $x_i^{(*)}$, $i = 1,\dots,n$.

To arrive at our general state-space description of a class of models of the type (4), we first generalize (3) to[1]

$$x_i^{(k+1)} = \sum_{j=1}^{n} H_{ij}\left(x_1^{(e_1^i(k))}, \dots, x_n^{(e_n^i(k))}, e_1^i(k), \dots, e_n^i(k)\right) x_j^{(e_j^i(k))}$$

$$\text{for} \quad i = 1,2,\dots,n, \quad k = 0,1,2,\dots, \quad (5)$$

[1]To avoid conflict of notation, we henceforth replace z of (3) by x.

where $e_j^i(k) \in \{k, k-1, \dots, k-d\}$ for some integer $d > 0$ and for $i, j = 1, \dots, n$, and for all k.

REMARK 1. Note that the nonlinear time-varying discrete-time system (5) or, equivalently, (6) (see below) admits the zero solution $x_i^{(k)} = 0$ for all i, k. Thus, when we speak of convergence of an asynchronous iterative process we are speaking of the asymptotic (or exponential) stability of the zero solution of (5) or (6). For precise definitions of stability, see Kalman and Bertram (1960) and Ortega (1973).

REMARK 2. In model (5), at time instant k, the ith processor receives information from the jth processor with a time-varying delay of $k - e_j^i(k)$ units. The values of the individual delay terms $e_j^i(k)$ are determined by individual processor computation times, communication delays, memory-access delays, and software-access delays—all of which are, in practice, time-varying and bounded, and thus we are making the following:

ASSUMPTION 1. *The time delays $k - e_j^i(k)$ are bounded above by the positive integer d for all i, j, k.*

The model (5) can be written in standard state-space form by stacking the variables $x_i^{(k)}$ and their delayed versions $x_i^{(k-l)}$, $l = 1, \dots, d$, in a state vector. To do this, we introduce the following notation:

$$
\begin{aligned}
\overline{H}_{ij}^{(k)} &:= H_{ij}\left(x_1^{(e_1^i(k))}, x_2^{(e_2^i(k))}, \dots, x_n^{(e_n^i(k))}, e_1^i(k), \dots, e_n^i(k)\right),\\
x_{i,l}^{(k)} &:= x_i^{(k-l)}, \qquad l = 0, \dots, d, \quad i = 1, \dots, n,\\
\bar{x}_i^{(k)T} &:= \left(x_i^{(k)T}, x_{i,1}^{(k)T}, \dots, x_{i,d-1}^{(k)T}, x_{i,d}^{(k)T}\right),\\
\mathscr{x}^{(k)T} &:= \left(\bar{x}_1^{(k)T}, \bar{x}_2^{(k)T}, \dots, \bar{x}_n^{(k)T}\right),
\end{aligned}
\tag{6}
$$

and, finally, switching functions $\psi_{ij}(\bar{x}_j^{(k)}, k)$ that have the following property: for each triple i, j, k, $\psi_{ij}(\bar{x}_j^{(k)}, k)$ takes the value of exactly one of the variables in the set

$$\left\{x_j^{(k-d)}, x_j^{(k-d-1)}, \dots, x_j^{(k)}\right\},$$

which is alternatively written as

$$\left\{x_{j,d}^{(k)}, x_{j,d-1}^{(k)}, \dots, x_j^{(k)}\right\}$$

in a form appropriate for state-space representation, using the notation introduced above. Equation (5) can now be written in state-space form as follows:

$$\begin{aligned} x_i^{(k+1)} &= \sum_{j=1}^{n} \overline{H}_{ij}^{(k)} \psi_{ij}\left(\bar{x}_j^{(k)}, k\right), \\ x_{i,1}^{(k+1)} &= x_i^{(k)}, \\ x_{i,2}^{(k+1)} &= x_{i,1}^{(k)}, \\ &\vdots \\ x_{i,d}^{(k+1)} &= x_{i,d-1}^{(k)} \qquad i = 1,2,\dots,n, \end{aligned} \tag{7}$$

with initial conditions $x_{i,l}^{(0)}$, $l = 0,1,\dots,d$, appropriately chosen.

EXAMPLE 1. To fix ideas and explain the choice of initial conditions, consider the following simple linear time-invariant example (see Appendix 1).

Suppose that the system (2) is given by

$$\begin{aligned} x_1^{(k+1)} &= \tfrac{1}{3} x_1^{(k)}, \\ x_2^{(k+1)} &= \tfrac{3}{4} x_1^{(k)} + \tfrac{1}{2} x_2^{(k)} \end{aligned} \tag{8}$$

with initial conditions $x_1^{(0)} = \frac{3}{4}$, $x_2^{(0)} = 0$. Suppose also that, in the parallel implementation, the processor P1 assigned to the evaluation of $x_1^{(\cdot)}$ is twice as slow as P2 (which evaluates $x_2^{(\cdot)}$), so that P1 updates once for every two P2 updates. This can be modeled as follows:

$$\begin{aligned} x_1^{(k+1)} &= \tfrac{1}{3}x_1^{(k-2)}, \\ x_2^{(k+1)} &= \tfrac{3}{4}x_1^{(k)} + \tfrac{1}{2}x_2^{(k)}, \end{aligned} \tag{9}$$

which can be written in state-space form (7) as follows:

$$\begin{array}{ll} & \text{Initial condition} \\ x_1^{(k+1)} = \tfrac{1}{3}x_{1,2}^{(k)}, & x_{1,2}^{(0)} = \tfrac{9}{4}, \\ x_{1,2}^{(k+1)} = x_{1,1}^{(k)}, & x_{1,1}^{(0)} = \tfrac{9}{4}, \\ x_{1,1}^{(k+1)} = x_1^{(k)}, & x_1^{(0)} = \tfrac{3}{4}, \\ x_2^{(k+1)} = \tfrac{3}{4}x_1^{(k)} + \tfrac{1}{2}x_2^{(k)}, & x_2^{(0)} = 0. \end{array} \tag{10}$$

It can be easily verified that with the initial conditions specified in (10), for $k = 1,2,3$, $(x_1^{(k)}, x_2^{(k)})$ is equal to $(\frac{3}{4}, \frac{9}{16})$, $(\frac{3}{4}, \frac{27}{32})$, $(\frac{1}{4}, \frac{63}{64})$ respectively. As expected, the value of $x_1^{(k)}$ changes once for every two updates of $x_2^{(k)}$.

REMARK 3. Our model, Equation (7), being block and nonlinear, specializes to the linear point Chazan-Miranker model with uniformly bounded delays. Inasmuch as it is state-space model, it is similar to the linear state-space models of Lubachevsky and Mitra (1986, Equation (4-9), p. 140) and Bru et al. (1988, Equation (2.19), p. 184), but differs from both in the way the delayed variables are handled. Talukdar et al. (1983) consider the class of nonlinear L_∞-contractions, which is a slightly different class of nonlinear functions from the one considered by us (see Appendix 1).

3. SUFFICIENT CONDITIONS FOR CONVERGENCE OF ASYNCHRONOUS BLOCK-ITERATIVE PROCESSES

In this section we use Liapunov techniques to find sufficient conditions for stability of the zero solution $x_{i,l}^{(k)} = 0$ $\forall i, l, k$ of (7). Note that the asynchronous model (7) includes the synchronous model as a special case [the switching functions $\psi_{ij}(\bar{x}_j^{(k)}, k\ )$ simply choose the variables $x_j^{(k)}$ for all i, j, k]. As a result, the only general necessary conditions that can be obtained are those necessary for the convergence of the corresponding synchronous iterative computation.[2]

Before stating the main result we need some definitions and a lemma on nonnegative matrices that is a compilation of various results that have been stated explicitly or implicitly in the abundant literature on such matrices. We use the standard definitions and properties of nonnegative matrices and the infinity norm $\|\cdot\|_\infty$ [see, for example, Varga (1962), Berman and Plemmons (1979)].

DEFINITION 1. An $n\times n$ nonnegative matrix $F=(f_{ij})$ is called quasidominant if there exist positive real numbers $d_1,\dots,d_n$ such that

$$d_i f_{ii} > \sum_{j\neq i} d_j f_{ij}, \qquad i=1,\dots,n.$$

REMARK 4. If F is not nonnegative, the f_{ij}'s are replaced by their moduli $|f_{ij}|$ in Moylan's (1977) definition of quasidominance. There are other slightly different notions of quasidominance, and different terminology (for instance, F is sometimes said to possess a quasidominant diagonal), but we will always use Definition 1.

DEFINITION 2. A matrix H is said to belong to class $\mathcal{D}$, written $H\in\mathcal{D}$, if, for some positive diagonal P, the matrix H^TPH-P is negative definite.

DEFINITION 3. Let A be a square matrix which can be expressed in the form $A=rI-B$ for some $B\geqslant 0$ and $r>\rho(B)$. Then A is called a nonsingular M-matrix.

[2]Chazan and Miranker (1969) also include synchronous computation as a special case of their chaotic (asynchronous) relaxation and, in the linear case, give a necessary condition for the convergence of asynchronous computations: their condition $[\rho(|H|)<1]$ is stronger than (i.e. implies) the corresponding condition for synchronous computations $[\rho(H)<1]$, and its necessity is interpreted as follows: if it is violated, then there exists a diverging *asynchronous* iteration sequence.

We can now state

LEMMA 1. *For a nonnegative matrix H, the following are equivalent:*

(L1) *There exists a positive diagonal matrix D such that* $\|D^{-1}HD\|_\infty < 1$.
(L2) *The spectral radius,* $\rho(H)$, *is strictly less than unity.*
(L3) $I - H$ *is quasidominant.*
(L4) $H \in \mathcal{D}$.
(L5) $I - H$ *is a nonsingular M-matrix.*

Proof. See Appendix 2. ■

REMARK 5. There are many equivalent characterizations of M-matrices [Poole and Boullion (1974), Plemmons (1977), Berman and Plemmons (1979)]: thus condition (L5) above may be the easiest to check.

In order to state our main result, we first define the aggregate matrix $H = (h_{ij})$ as follows: h_{ij} is defined to be the supremum of $\|\overline{H}_{ij}^{(k)}\|_\infty$ over all possible values of $x_j^{(e_j^i(k))}$ and $e_j^{i(k)}$, where [see Equation (6)]

$$\overline{H}_{ij}^{(k)} := H_{ij}\left(x_1^{(e_1^i(k))}, x_2^{(e_2^i(k))}, \dots, x_n^{(e_n^i(k))}, e_1^i(k), \dots, e_n^i(k)\right).$$

With a slight abuse of notation, we can write

$$h_{ij} := \sup_{x,\,k} \|\overline{H}_{ij}^{(k)}\|_\infty. \tag{11}$$

We now state our main result:

THEOREM 1. *The zero solution of* (7) *is globally exponentially stable if there exist n positive real numbers d_i, $i = 1, \dots, n$, such that*

$$\max_{1 \leqslant i \leqslant n} \left\{ d_i^{-1} \sum_{j=1}^{n} d_j h_{ij} \right\} < 1. \tag{12}$$

REMARK 6. (12) is condition (L1) of Lemma 1. In the proof of the theorem, we will use (12) because it turns out to be the most convenient of the five equivalent conditions of Lemma 1, given our choice of Liapunov function.

REMARK 7. (12) implies that $\{\bar{H}_{ij}^{(k)}\}$ is a (norm-) bounded sequence for all k. We will use this fact in the proof. Furthermore, (12) implies that the zero solution is the only global equilibrium of (7) (see end of Appendix 1).

NOTATION FOR PROOF. $\mathbf{n} := \{1,\dots,n\}$, $\mathbf{d} := \{0,1,\dots,d\}$, and $(\mathbf{d}-1) := \{0,1,\dots,d-1\}$.

Proof. Consider the Liapunov-function candidate

$$V(k) := \max_{i\in\mathbf{n},\, l\in\mathbf{d}} \left\{ d_i^{-1}\|x_i^{(k)}\|_\infty, d_i^{-1}\|x_{i,l}^{(k)}\|_\infty \right\}. \tag{13}$$

Then, from (7),

$$\begin{aligned}
V(k+1) &= \max_{i\in\mathbf{n},\, l\in(\mathbf{d}-1)} \left\{ d_i^{-1}\left\| \sum_{j=1}^{n} \bar{H}_{ij}^{(k)}\psi_{ij}\left(\bar{x}_j^{(k)},k\right)\right\|_\infty, d_i^{-1}\|x_{i,l}^{(k)}\|_\infty \right\} \\
&\leqslant \max_{i\in\mathbf{n},\, l\in\mathbf{d}} \left\{ \sum_{j=1}^{n} d_j d_i^{-1}\|\bar{H}_{ij}^{(k)}\|_\infty d_j^{-1}\left\|\psi_{ij}\left(\bar{x}_j^{(k)},k\right)\right\|_\infty, d_i^{-1}\|x_{i,l}^{(k)}\|_\infty \right\} \\
&\leqslant \max_{i\in\mathbf{n},\, l\in\mathbf{d}} \left\{ \max_{i\in\mathbf{n}}\left\{ d_i^{-1}\sum_{j=1}^{n} d_j\|\bar{H}_{ij}^{(k)}\|_\infty \right\} \right. \\
&\qquad \left. \times \max_{j\in\mathbf{n}}\left\{ d_j^{-1}\left\|\psi_{ij}\left(\bar{x}_j^{(k)},k\right)\right\|_\infty\right\}, d_i^{-1}\|x_{i,l}^{(k)}\|_\infty \right\} \\
&\leqslant \max_{i\in\mathbf{n},\, l\in\mathbf{d}} \left\{ \max_{j\in\mathbf{n}}\left\{ d_j^{-1}\left\|\psi_{ij}\left(\bar{x}_j^{(k)},k\right)\right\|_\infty\right\}, d_i^{-1}\|x_{i,l}^{(k)}\|_\infty \right\} \\
&\leqslant \max_{i\in\mathbf{n},\, l\in\mathbf{d}} \left\{ d_i^{-1}\|x_i^{(k)}\|_\infty, d_i^{-1}\|x_{i,l}^{(k)}\|_\infty \right\} = V(k),
\end{aligned} \tag{14}$$

i.e., $V(k+1) \leqslant V(k)$.

Note that the penultimate majorization follows from (11) and (12). The last majorization (14) is a consequence of the fact that the range of $\psi_{ij}(\bar{x}_j^{(k)},k)$ is the set $\{x_{j,d}^{(k)}, x_{j,d-1}^{(k)},\dots,x_j^{(k)}\}$.

From the state-space equation (7), using Assumption 1 (namely the assumption that there is an upper bound of d time units on all delays), it follows that the function $V(k)$ defined in (13) can stay constant for at most d

steps, after which it must decrease [by (12)]. In other words, there exists $\delta \in (0,1)$ such that

$$V(k) - V(k+d+1) \geqslant \delta V(k). \tag{15}$$

Equation (15) implies that there is a subsequence $\{V(k_i)\}$ of $\{V(k)\}$ with

$$k_i \leqslant i(d+1) \tag{16}$$

which decays exponentially fast; i.e., for some $a > 0$, $b \in (0,1)$

$$V(k_i) \leqslant ab^i. \tag{17}$$

Let $\mathbf{x}^{(k)}$ be the state vector of the time-varying system (7), i.e.

$$\mathbf{x}^{(k)T} := \left(x_1^{(k)T}, x_{1,1}^{(k)T}, x_{1,2}^{(k)T}, \dots, x_{1,d}^{(k)T}, x_2^{(k)T}, \dots, x_{2,d}^{(k)T}, \dots, x_n^{(k)T}, \dots, x_{n,d}^{(k)T}\right).$$

From the definition (13) of $V(k)$, it follows that if we choose ε as $\min_i\{d_i^{-1}\} = (\max_i\{d_i\})^{-1} > 0$, then

$$\varepsilon \|\mathbf{x}^{(k)}\|_\infty \leqslant V(k) \qquad \forall k. \tag{18}$$

(17) now implies that the subsequence $\{\|\mathbf{x}^{(k_i)}\|_\infty\}$ of $\{\|\mathbf{x}^{(k)}\|_\infty\}$ decays at least at the same rate:

$$\|\mathbf{x}^{(k_i)}\|_\infty \leqslant \tilde{a} b^i \qquad \text{for some} \quad \tilde{a} > 0. \tag{19}$$

It remains to show that, as a result, $\{\|\mathbf{x}^{(k)}\|_\infty\}$ also decays exponentially fast. For arbitrary k, there exists a greatest k_i with

$$k_i \leqslant k < k_i + d + 1. \tag{20}$$

Note that (7) can be written in matrix form as follows:

$$\mathbf{x}^{(k+1)} = \underline{H}(\mathbf{x}^{(k)}, k)\mathbf{x}^{(k)}, \tag{21}$$

where $H(\mathbf{x}^{(k)}, k)$ is a state-dependent time-varying matrix defined by the

$\overline{H}_{ij}^{(k)}$'s and ψ_{ij}'s of (7). From (20) and (21) we conclude that for arbitrary k, $x^{(k)}$ can be written as

$$x^{(k)} = H(x^{(k-1)}, k-1)H(x^{(k-2)}, k-2)\cdots H(x^{(k_i)}, k_i)x^{(k_i)},$$
$$\text{with} \quad 0 \leqslant k - k_i \leqslant d. \tag{22}$$

Since the $\overline{H}_{ij}^{(k)}$'s are bounded for all k (Remark 7), it follows from (19) and (22) that we can find a constant $a' > 0$, independent of $x^{(k_i)}$, such that

$$\begin{aligned}
\|x^{(k)}\|_\infty &\leqslant a'b^i \\
&\leqslant a'b^{k_i(d+1)^{-1}} && [\text{by } (16), \text{ since } 0 < b < 1] \\
&= \frac{a'}{(b')^d}\left(b^{(d+1)^{-1}}\right)^d b^{k_i(d+1)^{-1}} && \left[\text{where } b' = b^{(d+1)^{-1}} < 1\right] \\
&= a^* b^{(d+1)^{-1}(d+k_i)} && \left[\text{where } a^* = a'/(b')^d\right] \\
&\leqslant a^*(b')^k && [\text{by } (20)],
\end{aligned} \tag{23}$$

and this completes the proof. ∎

REMARK 8. From (23) it follows that as $d \to \infty$ one has $b' \to 1$ and $a^* \to a'$, so that convergence could be slower with larger delays.

REMARK 9. The proof uses the technique introduced by Anderson and Moore (1981) to prove their extended Liapunov lemma. However, their proof used a quadratic Liapunov function for linear systems: we use a max function for our nonlinear system, as in Šiljak (1978, p. 259). This is discussed further below.

4. DISCUSSION OF THEOREM 1

The above result can be interpreted from different points of view: considering (3) as a class of nonlinear time-varying systems, Theorem 1 can be viewed as a generalization of the result by Šiljak (1978, Theorem 4.12), and this generalization is twofold: first, Theorem 1 is a block version of the result; second, and more important, the condition (12) assures exponential stability of the system (3) even when time-varying delays are present in the interconnections between the subsystems represented by $x_i^{(k+1)} = \overline{H}_{ii}^{(k)} x_i^{(k)}$: this can be interpreted as saying that the condition (12) is robust.

On the other hand, within the perspective of chaotic or asynchronous iterative methods, and in view of Lemma 1, Theorem 1 generalizes Chazan and Miranker (1969) in the following aspects: first, the model (7) is somewhat more general than their model for chaotic relaxation; second, the condition (12) is the block version of $\rho(|H|) < 1$; third, the model (7) includes a class of nonlinearities and time-varying characteristics in the different blocks, whereas their model is restricted to linear iterative point methods.

For the class of linear, synchronous, block-iterative methods, the results derived by Robert (1969) (specifically Theorem 5 for the block Jacobi method) can be obtained using our Theorem 1 and Lemma 1 (see also end of Appendix 3). Also, for the particular case of constant blocks and no delays, we can compare condition (L3) of Lemma 1, which is equivalent to (12), with the results obtained by Okuguchi (1978, Theorem 5), which states that $\rho(H) < 1$ given that (i) all $h_{ii} \geqslant 0$, (ii) all diagonal blocks of $I - H$ are M-matrices, and (iii)

$$\sum_{j \neq i}^{n} \left\| (H_{ii} - I)^{-1} \right\|_{\infty} \|H_{ij}\|_{\infty} \frac{d_j}{d_i} < 1, \qquad i = 1, \dots, n. \tag{24}$$

Observe that, in the time-invariant delayless case, (12) is equivalent to

$$\sum_{j \neq i}^{n} (1 - \|H_{ii}\|_{\infty})^{-1} \|H_{ij}\|_{\infty} \frac{d_j}{d_i} < 1, \qquad i = 1, \dots, n, \tag{25}$$

and this condition implies that $\rho(H) < 1$ (using Lemma 1) without requiring the $I - H_{ii}$, $i = 1, \dots, n$, to be M-matrices.

Note, however, that whenever $\|H_{ii}\|_{\infty} < 1$, the Banach lemma states that

$$(1 + \|H_{ii}\|_{\infty})^{-1} \leqslant \left\| (I - H_{ii})^{-1} \right\|_{\infty} \leqslant (1 - \|H_{ii}\|_{\infty})^{-1}, \tag{26}$$

so that, under this condition, (25) implies (24). However, for Okuguchi's result to hold we need the additional conditions (i) and (ii) cited above.

Finally, we point out that Ostrowski (1955) introduced the notion of H-matrices, i.e. matrices which are diagonally similar to a strictly diagonally dominant matrix [see Fiedler and Pták (1967)]. He showed that free-steering iterative methods, in which the order of performing the iterations is arbitrary, converge when the iteration matrix is an H-matrix. Since our model permits free steering, delays, and a class of nonlinearities, we have generalized his result to this larger class of iterative methods, provided that a certain aggregate matrix is an H-matrix.

5. CONCLUDING REMARKS

In this paper we have proposed a state-space model for a class of nonlinear block-iterative parallel asynchronous methods. This model is sufficiently general to encompass many of the different models proposed earlier and is suited to the derivation of convergence conditions using Liapunov techniques. This approach has not been utilized so far in this context; some of its advantages are the following:

(a) clarification of the relationship between some of the different existing models and the corresponding convergence conditions;
(b) convergence conditions can be expressed in terms of the "aggregate" matrix that is associated to the block partition: this simplifies the testing of the condition for large-scale systems, because no matrix inversions are required and the infinity norm (maximum row sum) is easy to evaluate;
(c) the classical block and point iterative convergence conditions can be derived as particular cases of our result;
(d) explicit derivation of estimates for the rates of convergence is possible in the synchronous case (Appendix 3), and a clear relationship is obtained between the delays and the convergence rate in the asynchronous case.

We also point out that the techniques used above can be easily modified to handle the problem of iterative methods with overlapping blocks in the spirit of Ohta and Šiljak (1985).

APPENDIX 1. AN EXAMPLE OF TALUKDAR ET AL. (1983)

Talukdar et al. (1983, p. 4-16, Equation (4-22)), gave the following example of a nonlinear operator $G:\mathbb{R}^2 \to \mathbb{R}^2$:

$$G: X \mapsto \begin{cases} AX & \forall X \in D_0 := \{X : \|X\|_2 \leqslant 1\}, \\ (X^TX, X^TX)^T & \forall X \notin D_0 \end{cases} \tag{A1.1}$$

with

$$A = \begin{bmatrix} \frac{1}{3} & 0 \\ \frac{3}{4} & \frac{1}{2} \end{bmatrix} \quad \text{and} \quad X \in \mathbb{R}^2,$$

which is a vectorial norm A-contraction $[\rho(A) < 1]$ in D_0, but can generate a diverging asynchronous iteration sequence *with initial condition in* D_0, thus providing a counterexample to Theorem 1, p. 231 of Baudet (1978). We used the *linear* part of this example to illustrate the choice of initial conditions for a given asynchronous iteration [see (10)] and now return to it to illustrate the differences between our Theorem 1 and Theorem 4-2, p. 4-12 in Talukdar et al. (1983). We consider only the linear part, with A and D_0 as in (A1.1) above:

$$G_l : X \to AX : \mathbb{R}^2 \to \mathbb{R}^2 \qquad \forall X \in D_0. \tag{A1.2}$$

First, note that there does not exist L, $0 \leqslant L < 1$, s.t.

$$\|G_l(X) - G_l(Y)\|_\infty \leqslant L\|X - Y\|_\infty \qquad \forall X, Y \in D_0. \tag{A1.3}$$

For example, it is enough to choose $X = (1/\sqrt{2}, 0)^T$, $Y = (0, -1/\sqrt{2})^T$, and it can be easily verified that the left-hand side of (A1.3) is equal to 0.884, while the right-hand side is equal to 0.707, so that (A1.3) cannot be satisfied for any $L \in [0, 1)$. Thus G_l does not satisfy the sufficient condition of Talukdar et al. (1983): namely, it is not an L_∞-contraction, and so we cannot decide whether or not G_l can generate diverging asynchronous iteration sequences.

Observe now that, if we choose $D = \text{diag}(1,3)$, then $\|D^{-1}G_l D\|_\infty = 0.75 < 1$, so that G_l satisfies the sufficient condition (12) of Theorem 1 and hence generates only convergent asynchronous iteration sequences [i.e., (7) has an exponentially stable zero solution]. This can be interpreted as follows.

If we consider the iterative system $x^{(k+1)} = G_l x^{(k)}$ and let $Dy^{(k)} = x^{(k)}$ [$D = \text{diag}(1,3)$ as above], then $y^{(k+1)} = D^{-1}G_l Dy^{(k)}$. Introducing delays as in Example 1, Equation (9), for the iterative system in the new (y) coordinates, we can write the state-space equations, which have the same form as (10) (replacing $x_{i,l}$ by $y_{i,l}$) except that the last equation becomes $y_2^{(k+1)} = \frac{1}{4}y_1^{(k)} + \frac{1}{2}y_2^{(k)}$, and it is easy to see that, in the new basis, the Liapunov function defined in (13) is just the L_∞-norm of the state vector $y := (y_1, y_{1,1}, y_{1,2}, y_2)$. Clearly, this observation is not specific to this example and indeed is valid in the general case of Theorem 1: equivalently, following Kalman and Bertram (1960, Example 2, p. 398), we may say that if $H(x^{(k)}) := (H_{ij}(x^{(k)}, k))$ [see (3)], then $H(x^{(k)})x^{(k)}$ is a contraction using the norm $\|x\| := \max_i\{d_i^{-1}\|x_i\|_\infty\}$, in the sense that $\|H(x^{(k)})x^{(k)}\| < \|x^{(k)}\|$. This also implies that the zero solution is the only equilibrium of (3).

APPENDIX 2. PROOF OF LEMMA 1

For ease of reference, we reproduce Lemma 1 below and provide one possible scheme of proof, citing references to the literature for the various implications and equivalences. A version of this lemma [without (L4)] appears in Ohta and Šiljak (1985).

LEMMA 1. *For a nonnegative matrix $H = (h_{ij})$, the following are equivalent*:

(L1) *There exists a positive diagonal D such that $\|D^{-1}HD\|_\infty < 1$.*
(L2) *The spectral radius, $\rho(H)$, is strictly less than unity.*
(L3) *$I - H$ is quasidominant.*
(L4) *$H \in \mathscr{D} := \{H : \exists$ positive diagonal P s.t. $H^T PH - P$ is negative definite$\}$.*
(L5) *$I - H$ is a nonsingular M-matrix.*

Scheme of proof:

$$\begin{array}{ccccc}
(\mathrm{L4}) & \Leftarrow & (\mathrm{L3}) & \Leftrightarrow & (\mathrm{L1}) \\
\Downarrow & & \Updownarrow & & \\
(\mathrm{L2}) & \Rightarrow & (\mathrm{L5}) & &
\end{array}$$

Proof. (L1) $\Leftrightarrow$ (L3): By calculation. From Definition 1, Section 3, $I - H$ is quasidominant iff there exist $d_i > 0$, $i = 1, \ldots, n$, such that

$$d_i(1 - h_{ii}) > \sum_{j \neq i} d_j h_{ij} \qquad \text{for} \quad i = 1, \ldots, n \tag{A2.1}$$

$$\Leftrightarrow \qquad 1 > h_{ii} + \sum_{j \neq i} \frac{d_j}{d_i} h_{ij} \qquad \text{for} \quad i = 1, \ldots, n$$

$$\Leftrightarrow \qquad 1 > \sum_{j=1}^{n} \frac{d_j}{d_i} h_{ij} \qquad \text{for} \quad i = 1, \ldots, n$$

$$\Leftrightarrow \qquad \max_{1 \leq i \leq n} \sum_{j=1}^{n} \frac{d_j}{d_i} h_{ij} = \|D^{-1}HD\|_\infty < 1,$$

$$\text{where} \quad D := \mathrm{diag}(d_1, \ldots, d_n).$$

Note that in (A2.1) the easily deduced fact that $h_{ii} < 1$ for all i is used.

(L3) $\Rightarrow$ (L4) is proved in Moylan (1977, Theorem 4, p. 57); see also Moylan and Hill (1978, Appendix, Theorem A.2).

(L4) $\Rightarrow$ (L2) is a classical result of Liapunov theory: see, for example, Kalman and Bertram (1960).

(L2) $\Rightarrow$ (L5) follows from the definition of a nonsingular M-matrix (Definition 3, Section 3) with $r := 1 > \rho(H)$, $B := H \geqslant 0$, and $A := I - H$.

(L3) $\Leftrightarrow$ (L5): Moylan (1977, p. 56) makes the observation that if a matrix has the sign pattern required of an M-matrix, then it is quasidominant if and only if it is an M-matrix. In our case, since H is nonnegative, clearly $I - H$ has all off-diagonal entries nonpositive, and the equivalence is now obvious. ■

REMARK. (L1) $\Rightarrow$ (L2) can be proved directly from the properties of nonnegative matrices. From Varga (1962, p. 47, Exercise 2), we know that, if $H = (h_{ij})$ is an $n \times n$ nonnegative matrix, and d is any vector with positive components $d_1, \dots, d_n$, then

$$\rho(H) \leqslant \max_{1 \leqslant i \leqslant n} \left\{ d_i^{-1} \sum_{j=1}^{n} h_{ij} d_j \right\} = \|D^{-1}HD\|_\infty, \tag{A2.2}$$

where $D = \operatorname{diag}(d_1, \dots, d_n)$. From (A2.2) it follows directly that (L1) $\Rightarrow$ (L2). More interestingly, (L2) $\Rightarrow$ (L1) can also be proved directly for nonnegative *irreducible* matrices. For such matrices the inequality in (A2.2) is strict, and if P^* is the hyperoctant of strictly positive vectors $d > 0$ (i.e. vectors which have all components strictly positive), then it follows that

$$\rho(H) \leqslant \inf_{d \in P^*} \left\{ \max_{1 \leqslant i \leqslant n} \left(d_i^{-1} \sum_{j=1}^{n} h_{ij} d_j \right) \right\} = \inf_{D \in \mathscr{P}^*} \left\{ \|D^{-1}HD\|_\infty \right\}, \tag{A2.3}$$

where $\mathscr{P}^*$ is the set of positive diagonal matrices. Choosing the positive eigenvector $y \in P^*$ corresponding to the Perron-Frobenius eigenvalue, $\rho(H)$, of the irreducible matrix H shows that equality is valid in (A2.3).[3] From this we conclude [as in Varga (1962, Equation (2.22), p. 32)]

$$\rho(H) = \inf_{D \in \mathscr{P}^*} \|D^{-1}HD\|_\infty. \tag{A2.4}$$

By hypothesis (L2), $\rho(H)$ and hence the infimum in (A2.4) is strictly less than one, from which it follows that there exists $D_1 \in \mathscr{P}^*$ such that $\|D_1^{-1}HD_1\|_\infty < 1$, which is (L1).

[3]This is the step that breaks down if H is reducible; for such an H the Perron-Frobenius eigenvector need not be strictly positive (i.e. in P^*).

APPENDIX 3. THE SYNCHRONOUS CASE

The specialization of Theorem 1 to the case of synchronous iterations represented by the model (3) can be derived using a quadratic Liapunov function [we note, in passing, that this quadratic function does not seem to be suitable for handling the delays that appear in the model (7)]. Note also that we do not specify the vector and matrix norms used in this appendix, because the results are valid for any p-norm, $p \in [1, \infty]$.

Consider the system (3) [in this appendix we replace z of (3) by x], and let

$$h_{ij} = \sup_{x,k} \left\| H_{ij}(x^{(k)}, k) \right\|, \qquad H := (h_{ij}), \qquad \overline{H}^{(k)} := \left(H_{ij}(x^{(k)}, k) \right). \quad \text{(A3.1)}$$

We also introduce the following notation: Given $x^T = (x_1^T, \ldots, x_n^T)$, $x_i \in \mathbb{R}^{n_i}$, let $[\![x]\!]^T := (\|x_1\|, \ldots, \|x_n\|) \in \mathbb{R}^{1\times n}$. For $y^T = (y_1, \ldots, y_n) \in \mathbb{R}^{1\times n}$ we write $y \geqslant 0$ iff $y_i \geqslant 0$ $\forall i$; and for $y, z \in \mathbb{R}^n$, $y \geqslant z$ iff $y - z \geqslant 0$. Similarly, for $A = (a_{ij}) \in \mathbb{R}^{n\times n}$, $A \geqslant 0$ iff $a_{ij} \geqslant 0$ $\forall i, j$. For example, in (A3.1), $H \geqslant 0$, Finally, $\langle \cdot , \cdot \rangle$ is the standard inner product in $\mathbb{R}^n$.

Using (A3.1) and the notation above, we can write

$$[\![\overline{H}^{(k)} x^{(k)}]\!] \leqslant H [\![x^{(k)}]\!]. \quad \text{(A3.2)}$$

We are now ready to prove:

LEMMA A3.1. *The zero solution of* (3) *is globally exponentially stable if matrix $I - H$ is quasidominant* (*see Lemma* 1 *for other equivalent conditions on H*).

Proof. Let

$$V(x^{(k)}) = \sum_{i=1}^{n} d_i \|x_i^{(k)}\|^2 \quad \text{(A3.3)}$$

$$= \left\langle [\![x^{(k)}]\!], D[\![x^{(k)}]\!] \right\rangle, \qquad \text{where} \quad D = \operatorname{diag}(d_1, \ldots, d_n). \quad \text{(A3.4)}$$

Let $\Delta V(x^{(k)}) := V(x^{(k+1)}) - V(x^{(k)})$.

For the system (3),

$$V(x^{(k+1)}) = \left\langle [\![\overline{H}^{(k)} x^{(k)}]\!], D [\![\overline{H}^{(k)} x^{(k)}]\!] \right\rangle$$

$$\leqslant \left\langle H [\![x^{(k)}]\!], DH [\![x^{(k)}]\!] \right\rangle \qquad \text{[by (A3.2)]} \tag{A3.5}$$

$$\leqslant \rho(H)^2 V(x^{(k)}). \tag{A3.6}$$

Note that (A3.6) follows from the Perron-Frobenius theorem, since H is a nonnegative matrix and $[\![x^{(k)}]\!]$ a nonnegative vector for all k.

From (A3.5) we can conclude that

$$\Delta V(x^{(k)}) \leqslant \left\langle [\![x^{(k)}]\!], (H^T DH - D) [\![x^{(k)}]\!] \right\rangle. \tag{A3.7}$$

The right-hand side of (A3.7) is always negative (for $x^{(\cdot)} \neq 0$) iff $H^T DH - D$ is negative definite (i.e. $H \in \mathscr{D}$), and by Lemma 1 (see Appendix 2), we know that a sufficient condition for this to happen is $I - H$ quasidominant. Finally, since $V(\cdot)$ is radially unbounded, by (A3.6) we are done: i.e., the system is globally exponentially stable with decay rate $\rho(H)^2 < 1$ (by Lemma 1). ∎

Remark. From (A3.6) we can conclude that

$$\Delta V(x^{(k)}) \leqslant -\left[1 - \rho(H)^2\right] V(x^{(k)}). \tag{A3.8}$$

The right-hand side of (A3.8) is always negative because $V(\cdot)$ is always positive (for $x^{(\cdot)} \neq 0$), and from Lemma 1, $I - H$ is quasidominant iff $\rho(H) < 1$, which implies that $1 - \rho(H)^2 > 0$. This is an alternative proof of Lemma A3.1.

Now, if we consider, for example, the classical block Jacobi method for solving $Ax = b$ with a conformal partition of $A = (A_{ij})$, the corresponding error equation is written in the form (3) as follows:

$$\begin{array}{ll} \text{for } i \neq j, & H_{ij}(x^{(k)}, k) = A_{ij}^{-1} A_{ij}; \\ \text{for } i = j, & \\ & H_{ij}(x^{(k)}, k) = 0. \end{array} \tag{A3.9}$$

Consequently, for this case, using (i) the definition of quasi-block-diagonal dominance [Okuguchi (1978, Definition 1)] and (ii) the equivalence of quasi-block-diagonal dominance of $A=(A_{ij})$ and the quasidominance of $I-H$, Lemma A3.1 implies that

$$\sum_{j\neq i}^{n} \|A_{ii}^{-1}A_{ij}\| \frac{d_j}{d_i} < 1 \qquad \text{for} \quad i=1,2,\ldots,n \tag{A3.10}$$

assures the convergence of the block Jacobi method. Furthermore, if in (A3.9) we assume

$$h_{ii}=0 \quad \text{and} \quad h_{ij}=\|A_{ii}^{-1}\|\cdot\|A_{ij}\|,$$

then

$$\sum_{j\neq i}^{n} \|A_{ii}^{-1}\|\cdot\|A_{ij}\| \frac{d_j}{d_i} < 1 \qquad \text{for} \quad i=1,\ldots,n \tag{A3.11}$$

also assures the convergence of the block Jacobi method (by Lemma A3.1).

REFERENCES

Anderson, B. D. O. and Moore, J. B. 1981. Detectability and stabilizability of time-varying discrete-time linear systems, *SIAM J. Control Optim.* 19(1):20–32.

Baudet, G. M. 1978. Asynchronous iterative methods for multiprocessors, *J. Assoc. Comput. Mach.* 25(2):226–244.

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

Bertsekas, D. P. 1983. Distributed asynchronous computation of fixed points, *Math. Programming* 27:107–120.

Bru, R., Elsner, L., and Neumann, M. 1988. Models of parallel chaotic iteration methods, *Linear Algebra Appl.* 103:175–192.

Chazan, D. and Miranker, W. 1969. Chaotic relaxation, *Linear Algebra Appl.* 2:199–222.

El Tarazi, M. N. 1982. Some convergence results for asynchronous algorithms, *Numer. Math.* 39:325–340.

Fiedler, M. and Pták, V. 1967. Diagonally dominant matrices, *Czechoslovak Math. J.* 17(92):420–433.

Kalman, R. E. and Bertram, J. E. 1960. Control system analysis and design via the "second method" of Lyapunov II: Discrete-time systems, *J. Basic Engrg. Trans. ASME* 82(2):394–400.

Kung, H. T. 1976. Synchronous and asynchronous parallel algorithms for multiprocessors, in *Algorithms and Complexity—New Directions and Recent Results* (J. Traub, Ed.), Academic, New York, pp. 153–200.

Lubachevsky, B. and Mitra, D. 1986. A chaotic asynchronous algorithm for computing the fixed point of a nonnegative matrix of unit spectral radius, *J. Assoc. Comput. Mach.* 33(1):130–150.

Moylan, P. J. 1977. Matrices with positive principal minors, *Linear Algebra Appl.* 17:53–58.

Moylan, P. J. and Hill, D. J. 1978. Stability criteria for large-scale systems, *IEEE Trans. Automat. Control* AC-23(2):143–149.

Nour-Omid B., Raefsky, A., and Lyzenga, G. 1987. Solving finite element equations on concurrent computers, in *Parallel Computations and Their Impact on Mechanics* (A. K. Noor, Ed.), ASME, New York.

Ohta, Y. and Šiljak, D. D. 1985. Overlapping block diagonal dominance and existence of Liapunov functions, *J. Math. Anal. Appl.* 112(2):396–410.

Okuguchi, K. 1978. Matrices with dominant diagonal blocks and economic theory, *J. Math. Econom.* 5:43–52.

Ortega, J. M. 1973. Stability of difference equations and convergence of iterative processes, *SIAM J. Numer. Anal.* 10(2):268–282.

Ortega, J. M. and Rheinboldt, W. C. 1970. *Iterative Solutions of Nonlinear Equations in Several Variables*, Academic, New York.

Ostrowski, A. 1955. Determinanten mit überwiegender Hauptdiagonale und die absolute Konvergenz von linearen Iterationsprozessen, *Comm. Math. Helv.* 30:175–210.

Plemmons, R. J. 1977. *M*-matrix characterizations. I. Nonsingular *M*-matrices, *Linear Algebra Appl.* 18:175–188.

Polak, E. 1971. *Computational Methods in Optimization: A Unified Approach*, Academic, New York.

Poljak, B. J. 1982. Convergence and convergence rate of iterative stochastic algorithms, *Automat. Remote Control* 12:83–94.

Poole, G. and Boullion, T. 1974. A survey on *M*-matrices, *SIAM Rev.* 16(4):419–427.

Robert, F. 1969. Bloc *H*-matrices et convergence des méthodes itératives classiques par blocs, *Linear Algebra Appl.* 2:223–265.

Rosenfeld, J. 1967. A Case Study on Programming for Parallel Processors, Research Report RC-1864, IBM Watson Research Center, Yorktown Heights.

Šiljak, D. D. 1978. *Large-Scale Dynamic Systems*, North-Holland, New York.

Talukdar, S. N., Pyo, S. S., and Mehrotra, R. 1983. Designing Algorithms and Assignments for Distributed Processing, Final Report, EPRI EL-3317 RP1764-3, Electric Power Research Inst.

Varga, R. 1962. *Matrix Iterative Analysis*, Prentice-Hall, Englewood Cliffs, N.J.

Young, D. M. 1971. *Iterative Solutions of Large Linear Systems*, Academic, New York.

Received 16 October 1988; final manuscript accepted 27 February 1989