

On the Partial Synchronization of Iterative Methods

A. V. Dmitriev and S. M. Ermakov

St. Petersburg State University, St. Petersburg, 199034 Russia

e-mail: alx.dmitriev@gmail.com

Received December 27, 2015

Abstract—The rapidly growing field of parallel computing systems promotes the study of parallel algorithms, with the Monte Carlo method and asynchronous iterations being among the most valuable types. These algorithms have a number of advantages. There is no need for a global time in a parallel system (no need for synchronization), and all computational resources are efficiently loaded (the minimum processor idle time). The method of partial synchronization of iterations for systems of equations was proposed by the authors earlier. In this article, this method is generalized to include the case of nonlinear equations of the form $x = F(x)$, where x is an unknown column vector of length n , and F is an operator from \mathbb{R}^n into \mathbb{R}^n . We consider operators that do not satisfy conditions that are sufficient for the convergence of asynchronous iterations, with simple iterations still converging. In this case, one can specify such an incidence of the operator and such properties of the parallel system that asynchronous iterations fail to converge. Partial synchronization is one of the effective ways to solve this problem. An algorithm is proposed that guarantees the convergence of asynchronous iterations and the Monte Carlo method for the above class of operators. The rate of convergence of the algorithm is estimated. The results can prove useful for solving high-dimensional problems on multiprocessor computational systems.

Keywords: Monte Carlo methods, asynchronous methods, asynchronous iterations, parallel algorithms, statistical modeling.

DOI: 10.3103/S1063454116030055

INTRODUCTION

The rapid development of computers makes the problems of synchronization of calculations rather topical. For many computational systems, it may be advantageous to use algorithms that are slower than the optimal ones but do not involve any synchronization of computations. In this article, we consider partial-synchronization algorithms for iterations of a wide class of operators. The linear case was considered by the authors earlier in [1].

1. ASYNCHRONOUS ITERATIONS

Let us consider the problem of finding a fixed point of

$$x = F(x), \quad (1)$$

where $x = (x_1, x_2, \dots, x_n)^T$ is a column vector of unknowns and $F = (f_1(x), f_2(x), \dots, f_n(x))^T$ is an operator from \mathbb{R}^n into \mathbb{R}^n . Since we will consider a sequence of vectors along with their elements, it is convenient to introduce the following notation. The components of vectors from \mathbb{R}^n will be denoted as x_i , $i = 1, \dots, n$, while the sequence of vectors from \mathbb{R}^n will be represented as $x(j)$, $j = 0, 1, \dots$. Using this notation, we can write the method of simple iterations that will be used for searching the fixed point of (1) in the form

$$x(k+1) = F(x(k)), \quad k = 0, 1, \dots, \quad (2)$$

for a certain initial $x(0)$.

Conditions under which the process in Eq. (2) converges to a fixed point of the operator F can be found, for example, in [2]. Of all the possible conditions, let us specially consider the following class of operators and a theorem that is related to this class.

Definition 1. A mapping $F: D \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ is called a contraction mapping on $D_0 \subset D$ if there exists such $\alpha < 1$ that $\|F(x) - F(y)\| < \alpha\|x - y\|$ holds true for all $x, y \in D_0$.

Theorem 1. *Let the mapping $F: D \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ be contracting on a closed set $D_0 \subset D$ and $F(D_0) \subset D_0$ be true. Then, F has a unique fixed point $x^* \in D_0$, and for any initial $x(0) \in D_0$ the sequence $\{x(k)\}$ that is defined by Eq. (2) converges to x^* .*

We will adhere to the rather general definition of asynchronous iterations that was given in [3].

Let X_1, X_2, \dots, X_n be the given sets and X be their Cartesian product:

$$X = X_1 \times X_2 \times \dots \times X_n.$$

Accordingly, an element $x \in X$ has the structure

$$x = (x_1, x_2, \dots, x_n),$$

where x_i belong to X_i , $i = 1, \dots, n$. Let the functions $f_i: X \rightarrow X_i$ be given and the function $F: X \rightarrow X$ be represented in the form

$$F(x) = (f_1(x), f_2(x), \dots, f_n(x)), \quad \forall x \in X.$$

The problem now consists in finding a fixed point of the operator F , i.e., finding such $x^* \in X$ that $x^* = F(x^*)$ is true or, written in a component-wise notation,

$$x_i^* = f_i(x^*), \quad i = 1, \dots, n.$$

Now let us define an asynchronous version of the method of simple iterations, which will be called asynchronous iterations here and below, as follows:

$$x_i := f_i(x), \quad i = 1, \dots, n.$$

Let us assume that the set $T = \{0, 1, 2, \dots\}$ is the set of time moments when one or several components x_i of the vector x are updated by a certain processor of a distributed computer system. The classification of parallel computing systems is available, for example, in [3]. Let T^i denote the set of time moments when x_i are updated.

It is reasonable to expect that in a distributed computer system a processor that updates the component x_i does not have up-to-date information regarding the other components of the vector x . Therefore, the use of outdated information is possible in the asynchronous case. This fact can be written as

$$x_i(t+1) = f_i(x_1(\tau_1^i(t)), \dots, x_n(\tau_n^i(t))), \quad \forall t \in T^i, \quad (3)$$

where $\tau_i^j(t)$ are time moments that satisfy the inequality

$$0 \leq \tau_i^j(t) \leq t, \quad \forall t \in T.$$

For all the moments $t \notin T^i$ we take that x_i is not updated:

$$x_i(t+1) = x_i(t), \quad \forall t \notin T^i. \quad (4)$$

The elements of the set T should be considered as the indices of the sequence of real-time moments when updating occurs. Processors that do not update the x_i component do not have to know the set T_i , as this is not required for calculating the iterations in (3) and (4). Hence, there is no need to have a global time in the system. The difference $(t - \tau_i^j(t))$ between the current time and the time $\tau_i^j(t)$ when the processor that updates x_i received information about the component x_j for the last time can be considered as a delay in information transfer. In a situation like this, it is convenient to consider the computational process in the following manner. At the time moment $t \in T^i$, the processor that has finished its previous calculations and is ready for new ones receives the values $x_1(\tau_1^i(t)), \dots, x_n(\tau_n^i(t))$ by means of a certain mechanism and updates x_i by the formula in (3). In doing so, the processor has absolutely no need to know the values of $t, \tau_1^i(t), \dots, \tau_n^i(t)$ or the elements of the sets $T^j, j = 1, \dots, n$.

It is worth noting that such iterative methods for solving systems of linear equations as the Jacobi or Gauss–Seidel methods are special cases of the iteration in (3).

For the iterations (3) and (4) to be called asynchronous, certain conditions need to be imposed on the sets T_i and the time moments $\tau_i^j(t)$.

The sets T^i are infinite and for any sequence $\{t_k\} \subseteq T^i$ that tends to infinity we have $\lim_{k \rightarrow \infty} \tau_i^j(t_k) = \infty$ for $j = 1, i, n$.

This supposition guarantees that each component will be updated an infinite number of times, while the outdated information will be eventually withdrawn from processing. In what follows, we will assume that the above supposition holds true.

After the iterations in (3) and (4) and certain assumptions about these iterations have been introduced, a question arises of what the conditions are under which these iterations converge to a fixed point of the operator F . Conditions that are sufficient for this (see [3]) are provided by the following theorem.

Theorem 2. *Let a sequence of nonempty sets $\{X(k)\}$ exist such that the following conditions are satisfied:*

- $\dots \subset X(k+1) \subset X(k) \subset \dots \subset X(0) \subset X$;
- $F(x) \in X(k+1)$ for any k and $\forall x \in X(k)$; moreover, if the sequence $\{y(k)\}$ is such that $y(k) \in X(k)$ is fulfilled for $k = 0, 1, \dots$, then each limiting point $\{y(k)\}$ is a fixed point of the operator F ;
- for any $k \in \{0, 1, \dots\}$ there exist sets $X_i(k) \in X_i$ such that $X(k) = X_1(k) \times X_2(k) \times \dots \times X_n(k)$ is valid;
- the initial approximation $x(0)$ belongs to $X(0)$.

Then, each limiting point of the sequence $\{x(t)\}$ that is defined by the asynchronous iterations is a fixed point of the operator F .

It should be noted that the first and second conditions in the theorem implicate that the synchronous iterations $x := F(x)$ that start from a certain initial x that belongs to $X(0)$ converge to a fixed point of the operator F . The third condition implies that if we take an arbitrary element $X(k)$ and perform its permutation, we will again obtain an element that belongs to the set $X(k)$.

Let us confine ourselves to considering operators of the form $F: \mathbb{R}^n \rightarrow \mathbb{R}^n$. Let us consider the following norm on \mathbb{R}^n :

$$\|x\|_{\omega} = \max_i \frac{|x_i|}{\omega_i},$$

where we denoted $\omega = (\omega_1, \omega_2, \dots, \omega_n)$ and $\omega_i > 0, i = 1, \dots, n$.

Now, if we consider contraction mappings with a contraction parameter $\alpha < 1$ and suppose $X_i = \mathbb{R}, i = 1, \dots, n$ as well as $X = \mathbb{R}^n$ in the definition of asynchronous operations, then, according to the theorem, we need to construct a sequence of sets $\{X(k)\}$ in order to show that the asynchronous operations converge to a fixed point x^* of the operator F . Let us define these sets in the following way:

$$X(k) = \{x \in \mathbb{R}^n \mid \|x - x^*\|_{\omega} \leq \alpha^k \|x(0) - x^*\|_{\omega}\}.$$

It is easy to check that the conditions of the theorem are satisfied. Hereinafter, we will use the following norms:

$$\|x\| = \max_{1 \leq i \leq n} |x_i| \quad \text{for the vector} \quad x = (x_1, \dots, x_n)^T; \quad (5)$$

$$\|A\| = \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}| \quad \text{for the matrix} \quad A = \|a_{ij}\|_{i,j=1}^n. \quad (6)$$

2. SYSTEM OF LINEAR EQUATIONS

Let us consider the case

$$F(x) = Ax + b,$$

where $A = \|a_{ij}\|_{i,j=1}^n$ is a given matrix, $b \in \mathbb{R}^n$ is a given vector of right-hand sides. Then, we seek an x^* such that the following equality is true:

$$x^* = Ax^* + b.$$

The asynchronous iterations in (3) and (4) will have the following form for a system of linear equations:

$$\begin{aligned} x_i(t+1) &= \sum_{j=1}^n a_{i,j} x_j(\tau_j^i(t)) + b_i, \quad t \in T^i, \\ x_i(t+1) &= x_i(t), \quad t \notin T^i. \end{aligned} \quad (7)$$

Chazan and Miranker [4] showed that for a system of linear equations, chaotic relaxations, which are a special case of asynchronous iterations, converge to a solution of the system if and only if the inequality $\lambda_1(|A|) > 1$ holds true. Bertsekas and Tsitsiklis [3] generalized this result to include the case of asynchronous iterations.

Theorem 3. *Let a matrix A be such that the matrix $I - A$ is invertible. The following statements are equivalent:*

1. $\lambda_1(|A|) < 1$;

2. *for any initial $x(0)$, for any $b \in \mathbb{R}^n$, for any sets T^i that satisfy the conditions from the definition of asynchronous iterations, and for any choice of variables $\tau_j^i(t)$ such that $t - 2 < \tau_j^i(t) < t$, the sequence that is generated by the asynchronous iterations in (7) converges to $(I - A)^{-1}b$.*

Therefore, if an ordinary (synchronous) process converges, that is, the inequalities $|\lambda_1(A)| < 1$ and $|\lambda_1(A)| > 1$ hold true, then asynchronous iterations, at least, of a certain form, necessarily diverge. The situation can be remedied by performing a certain number of synchronous iterations that reduce the error after a certain group of asynchronous ones (partial synchronization).

In what follows, we will consider an algorithm that uses l ordinary iterations after every m asynchronous ones. It is evident that there exists such a value of m that the combined iterative process converges even if slower, generally speaking, than a completely synchronized process. Therefore, our approach only makes sense if the asynchronous iterations are much cheaper than the synchronous ones.

The estimate of the gain that is made substantially depends on the form of the matrix A and, in a general case, can be rather crude. A numerical experiment looks most likely to be efficient here. Nevertheless, we will prove a lemma that indicates the limits of growth of errors in the asynchronous case at $\lambda_1(|A|) > 1$.

Let $x(t)$, $t = 0, 1, 2, \dots$ be the sequence of asynchronous iterations for the system

$$x = Ax + b,$$

and \tilde{x} be a solution of the system. Then, $x(t)$ can be represented in the form $x(t) = \tilde{x} + \Delta x(t)$, where $\Delta x(t)$ is the sequence of asynchronous iterations for the system

$$x = Ax.$$

Lemma 1. *If $|\lambda_1(A)| < 1$ and $\lambda_1(|A|) > 1$ for a matrix A , then the inequality*

$$\|\Delta x(k)\| \leq \|A\|^k \|\Delta x(0)\| \quad (8)$$

holds true for $k = 0, 1, 2, \dots$

Proof. Let us prove the lemma by induction. For $k = 0$ we have

$$\|\Delta x(0)\| = \|A\|^0 \|\Delta x(0)\|,$$

and, hence, the base case has been proved. Now, let Eq. (8) be true for all $k \leq m$. Let us show that Eq. (8) is then satisfied for $k = m + 1$. In accordance with the definition of asynchronous iterations, if we have $m \in T^i$, then

$$\Delta x_i(m+1) = \sum_{j=1}^n a_{i,j} \Delta x_j(\tau_j^i(m)),$$

and if we have $m \notin T^i$, then

$$\Delta x_i(m+1) = \Delta x_i(m).$$

Let $\hat{\Delta x}$ be such a vector of length $2n$ that the equalities $\Delta \hat{x}_i = \Delta x_i(m)$, $\Delta \hat{x}_{n+i} = 0$ and $\Delta \hat{x}_i = 0$, $\Delta \hat{x}_{n+i} = \Delta x_i(m)$ hold for $m \in T^i$ and $m \notin T^i$, respectively. Let us denote the vector $(\Delta x_1(\tau_1^i(m)), \dots, \Delta x_n(\tau_n^i(m)))^T$ as $\widetilde{\Delta x}$. For $\Delta \hat{x}$, $\widetilde{\Delta x}$, and $\Delta x(m)$ we have the equality

$$\Delta \hat{x} = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix} \begin{pmatrix} \widetilde{\Delta x} \\ \Delta x(m) \end{pmatrix},$$

where A_1 and A_2 are $n \times n$ matrices. For $m \in T_j$ the j th row of the matrix A_1 is equal to the j th row of the matrix A , and for $m \notin T_j$ the j th row of the matrix A_1 consists of zeros.

The matrix A_2 is a diagonal matrix for which the j th diagonal element equals zero if $m \in T_j$ and unity otherwise.

Let us note that the relations $\|\Delta\hat{x}\| = \|\Delta x(m+1)\|$, $\|A_1\| \leq \|A\|$ are valid and, by virtue of $\|A\| \geq \lambda_1(|A|) > 1$, the inequality $\|A_2\| \leq \|A\|$ is fulfilled. Then, the following inequality holds:

$$\|\Delta x(m+1)\| = \left\| \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix} \begin{pmatrix} \widetilde{\Delta x} \\ \Delta x(m) \end{pmatrix} \right\| \leq \|A\| \left\| \begin{pmatrix} \widetilde{\Delta x} \\ \Delta x(m) \end{pmatrix} \right\|.$$

Let $\left\| \begin{pmatrix} \widetilde{\Delta x} \\ \Delta x(m) \end{pmatrix} \right\| = |\Delta x_{j^*}(s')|$ for a certain natural $s' < m$ and $1 \leq j^* \leq n$. As we have $s' < m$, the following will hold true by virtue of the induction hypothesis

$$\left\| \begin{pmatrix} \widetilde{\Delta x} \\ \Delta x(m) \end{pmatrix} \right\| \leq \|\Delta x(s')\| \leq \|A\|^{s'} \|\Delta x(0)\| \leq \|A\|^m \|\Delta x(0)\|,$$

and, hence,

$$\|\Delta x(m+1)\| \leq \|A\|^{m+1} \|\Delta x(0)\|.$$

Thus the lemma has been proven. \square

Now, it is easy to prove the following theorem.

Theorem 4. *If an iterative process consists of a sequence of groups of m asynchronous iterations that are followed by l synchronous ones then the process converges for large enough values of l , the convergence rate being no slower than $\lambda_\varepsilon(\|A\|/\lambda_\varepsilon)^{m/(m+l)}$ for an arbitrary λ_ε that satisfies the inequality $|\lambda_1(A)| < \lambda_\varepsilon < 1$.*

Proof. It follows from Lemma 2 that the error may grow by no more than $\|A\|^m$ times after m asynchronous iterations. As we have $|\lambda_1(A)| < 1$, then, for $\forall \varepsilon > 0$ such that $\varepsilon < 1 - |\lambda_1(A)|$, there exists such l_0 that for $\forall l \geq l_0$ the inequality $\|A\|^l < (|\lambda_1(A)| + \varepsilon)^l < 1$ holds true. Let us denote $|\lambda_1(A)| + \varepsilon$ as λ_ε . It can be easily seen that the inequality $|\lambda_1(A)| < \lambda_\varepsilon < 1$ is fulfilled. The norm of the error after $m + l$ iterations (m asynchronous iterations and l synchronous ones) can be estimated as follows:

$$\|\Delta x(m+l)\| \leq \|A\|^l \|A\|^m \|\Delta x(0)\| < \lambda_\varepsilon^l \|A\|^m \|\Delta x(0)\|.$$

For a large enough l we have $\lambda_\varepsilon^l \|A\|^m$, which proves the first part of the theorem.

We can see that after $m + l$ iterations the error reduces $\lambda_\varepsilon^l \|A\|^m$ times. We would have had such a result for a geometric convergence with a parameter r if $r^{m+l} = \lambda_\varepsilon^l \|A\|^m$ was valid. That is, we obtain

$$r = (\lambda_\varepsilon^l \|A\|^m)^{\frac{1}{m+l}} = \lambda_\varepsilon \left(\frac{\|A\|}{\lambda_\varepsilon} \right)^{\frac{m}{m+l}},$$

which proves the second part of the theorem. \square

3. SYSTEM OF NONLINEAR EQUATIONS

Definition 2. An operator F from \mathbb{R}^n into \mathbb{R}^n is called a Lipschitz operator (also an n -Lipschitz operator [2]) on $D \subseteq \mathbb{R}^n$, if there exists a nonnegative matrix A such that the following inequality holds true:

$$|F(x) - F(y)| \leq A|x - y|, \quad \forall x, y \in D, \quad (9)$$

where the modulus operation is applied component-wise and the inequality holds for all the components. The matrix A is called the Lipschitz matrix of the operator F .

Definition 3. An operator F from \mathbb{R}^n into \mathbb{R}^n is called a contraction operator (also an n -contraction operator [2]) on $D \subseteq \mathbb{R}^n$, if it is Lipschitz on D and if its Lipschitz matrix A satisfies $\lambda_1(A) < 1$, where $\lambda_1(A)$ is the first eigenvalue of the matrix A .

The following theorem is valid [2] for contraction operators.

Theorem 5. *If F is a contraction operator on a closed set $D \subseteq \mathbb{R}^n$ and if $F(D) \subseteq D$, then arbitrary asynchronous iterations converge to the unique solution of the system in Eq. (1).*

Therefore, if an ordinary (synchronous) process $x_{k+1} = F(x_k)$, $k = 0, 1, \dots$ converges and the operator F does not satisfy the conditions of Theorem 5, then asynchronous operations, at least, of a certain form, necessarily diverge.

An example of the operator F for which asynchronous iterations diverge is easy to construct. It suffices to take a matrix for which $|\lambda_1(A)| < 1 - \delta$, $\delta > 0$, $|\lambda_1(A)| > 1$. Then, it can be easily seen that the operator $A + G$, where G is an operator with a norm that is less than δ , will possess the required properties, with the divergence rate of the synchronous process not necessarily being geometric. This situation is rather similar to the linear case. In fact, let us perform l ordinary iterations after every m asynchronous ones (the geometric convergence of the ordinary iterations is assumed).

It is evident that a value of m exists for which the combined process converges but at a rate that is, generally speaking, lower than the rate for the completely synchronized process. Therefore, our approach makes sense in this case, too, if the asynchronous iterations are much cheaper than the synchronous ones.

The estimate of the gain that is made substantially depends on the type of divergence and, in a general case, can be rather crude. A numerical experiment looks most likely to be efficient here. Let us prove a lemma that indicates the limits of growth of errors in the asynchronous case at $\lambda_1(|A|) > 1$.

Let $x(t)$, $t = 0, 1, 2, \dots$ be a sequence of asynchronous iterations for the system in (1) and \tilde{x} be a solution of the system. Let F be a Lipschitz operator on $D \in \mathbb{R}^n$ with a Lipschitz matrix A , with F not being a contraction operator, i.e., the inequality $\lambda_1(|A|) > 1$ holds true. Without loss of generality, we can take that $\tilde{x} = F(\tilde{x}) = 0$ when we consider the operator $F(x + \tilde{x}) - \tilde{x}$. Assuming $y = \tilde{x}$ in the inequality in Eq. (9), the Lipschitz condition for the operator F takes the form $|F(x)| \leq A|x|$, $\forall x \in D$.

Lemma 2. *Under the above assumptions with regard to the operator F , the following inequality is fulfilled*

$$\|x(k)\| \leq \|A\|^k \|x(0)\| \quad (10)$$

for $k = 0, 1, 2, \dots$

Proof. Let us prove the lemma by induction. For k we have

$$\|x(0)\| = \|A\|^0 \|x(0)\|,$$

and, hence, the basis has been proved. Now, let (10) be fulfilled for all $k \leq m$. Let us show that (10) then also holds true for $k = m + 1$. According to the definition of asynchronous iterations, if $m \in T$ then we have

$$x_i(m+1) = f_i(x_1(\tau_1^i(m)), \dots, x_n(\tau_n^i(m))),$$

while if $m \notin T$ then we obtain

$$x_i(m+1) = x_i(m).$$

If we have $x_i(m+1) = x_i(m)$, then the following relationship holds by virtue of the induction hypothesis:

$$|x_i(m+1)| \leq \|x(m)\| \leq \|A\|^m \|x(0)\| \leq \|A\|^{m+1} \|x(0)\|.$$

If we have $x_i(m+1) = f_i(x_1(\tau_1^i(m)), \dots, x_n(\tau_n^i(m)))$, then we have

$$\begin{aligned} |x_i(m+1)| &= |f_i(x_1(\tau_1^i(m)), \dots, x_n(\tau_n^i(m)))| \leq \sum_{j=1}^n a_{ij} |x_j(\tau_j^i(m))| \\ &\leq \|A\| \|(x_1(\tau_1^i(m)), \dots, x_n(\tau_n^i(m)))\| \end{aligned}$$

due to the operator F being Lipschitz. From the definition of asynchronous iterations we have $\tau_j^i(m) \leq m$, $j = 1, \dots, n$ and the following inequality follows from the induction hypothesis:

$$|x_i(m+1)| \leq \|A\| \|(x_1(s_1(m)), \dots, x_n(s_n(m)))\| \leq \|A\| \|A\|^m \|x(0)\| \leq \|A\|^{m+1} \|x(0)\|.$$

Hence, we have $|x_i(m+1)| \leq \|A\|^{m+1} \|x(0)\|$ for any i , and the lemma has been proven.

Now, it is easy to prove the following theorem for the class of the nonlinear operators that were described above. \square

Theorem 6. *If an iterative process consists of a sequence of groups of m asynchronous iterations followed by l synchronous ones, then for a large enough value of l the process converges at a rate that is not slower than $r(\|A\|/r)^{m/m+l}$, where $r < 1$ is the parameter of geometric convergence of the iterations in Eq. (2) to the solution of Eq. (1).*

CONCLUSIONS

Thus, for a certain class of computational systems the method of partial synchronization may prove to be a useful tool in the nonlinear case as well. It is apparent that other linearization methods, for example, Newton's method, also allow partial synchronization. However, the question of the rate of convergence requires special study in this case.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research, grant no. 14-01-00271-a.

REFERENCES

1. A. V. Dmitriev and S. M. Ermakov, "Monte Carlo method and asynchronous iterations," *Vestn. St. Peterb. Univ., Ser. 1: Math., Mekh., Astron.* **59** (4), 517–528 (2014).
2. J. M. Ortega and W. C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables* (Academic, San Diego, 1970; Mir, Moscow, 1975).
3. D. P. Bertsekas and J. N. Tsitsiklis, *Parallel and Distributed Computation: Numerical Methods* (Prentice-Hall, Upper Saddle River, NJ, 1989).
4. D. Chazan and W. Miranker, "Chaotic relaxation," *Linear Algebra Appl.* **2** (2), 199–222 (1969).

Translated by V. Potapchouck