Numerical Algorithms **25:** 347–361, 2000.

Convergence of some asynchronous nonlinear multisplitting methods

Daniel B. Szyld * and Jian-Jun Xu

Department of Mathematics, Temple University, 1805 N. Broad Street, Philadelphia, PA 19122-6094, USA
E-mail: {szyld;jxu}@math.temple.edu

Received 30 June 1999; accepted in revised form 8 May 2000

Dedicated to Richard S. Varga on the occasion of his seventieth birthday

Frommer's nonlinear multisplitting methods for solving nonlinear systems of equations are extended to the asynchronous setting. Block methods are extended to include overlap as well. Several specific cases are discussed. Sufficient conditions to guarantee their local convergence are given. A numerical example is presented illustrating the performance of the new approach.

Keywords: nonlinear multisplittings, asynchronous parallel methods

AMS subject classification: 65H10, 65Y05

1. Introduction

Consider the parallel solution of nonlinear systems of equations of the form

$$F(x) = 0, \quad \text{where } F = (f_1, \dots, f_n)^{\mathrm{T}} \tag{1}$$

is a nonlinear operator from $\mathbb{R}^n$ into itself.

If the system (1) is linearized, e.g., using Newton's method, each linear system can be solved in parallel using some kind of block iterative method. Block methods can be studied using the concept of multisplittings [18], and its application to Newton's method was undertaken by White [26,27]. Frommer [13] applied the concept of multisplittings directly to the nonlinear equation (1). We review his approach in section 2; see also [2, 17,23]. In these approaches, the parallel methods are synchronous in the sense that all processors have to wait at some synchronization point before proceeding to the next iteration.

In this paper we study asynchronous nonlinear multisplitting methods for the solution of (1), i.e., methods where no synchronization barrier is present; see, e.g., [4,6,16] for some general discussion on asynchronous methods. We present here a framework which is different than that in the recent paper [1]. Our paper deals with local conver-

* This work was supported by National Science Foundation grants DMS-9625865 and DMS-9973219.

gence and, in a sense, is more specialized. In our special situation, our hypotheses are more general than those in [1]. We also emphasize the use of overlap, i.e., the process by which certain variables are updated by more than one processor.

This paper is organized as follows: after a brief review of the nonlinear multisplitting method of Frommer [13] in section 2, we present the asynchronous nonlinear multisplitting method and its variants in section 3. General local convergence theorems are given in section 4. In section 5, local convergence is shown when the Jacobian of the nonlinear operator $F'(x)$ at the solution x^* is either a monotone or an H-matrix. Finally, we give an illustrative numerical example.

2. Nonlinear multisplitting methods

In this section we give a brief review of Frommer's nonlinear multisplitting method.

Definition 2.1 [13]. Let D be a domain in $\mathbb{R}^n$. For $l = 1, \ldots, L$, let $\widehat{F}_l : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ be nonlinear maps such that $\widehat{F}_l(x, x) = F(x)$ for all $x \in D$. Moreover, for $l = 1, \ldots, L$, let $E_l \in \mathbb{R}^{n\times n}$ be nonnegative diagonal matrices such that

$$\sum_{l=1}^{L} E_l = I. \tag{2}$$

The collection of pairs $(\widehat{F}_l, E_l)$, $l = 1, \ldots, L$, is called a nonlinear multisplitting of F. The corresponding nonlinear multisplitting method (NM-method) for solving (1) is defined by the iteration

$$x^{k+1} = \sum_{l=1}^{L} E_l y_l^k, \quad k = 0, 1, \ldots, \tag{3}$$

where y_l^k is the solution of

$$\widehat{F}_l\bigl(x^k, y_l^k\bigr) = 0. \tag{4}$$

There are numerous examples of nonlinear splittings of the nonlinear system (1), and any collection of them can form a nonlinear multisplitting; see, e.g., [13,17], and the references given therein.

The framework provided by the nonlinear multisplitting method can be used to analyze block iterative methods, possibly with overlap; see, e.g., [3,12,28,29]. To see how this can be done, let L be a positive integer and for $l = 1, \ldots, L$, let S_l be nonempty subsets of $\{1, \ldots, n\}$ such that $\bigcup_{l=1}^{L} S_l = \{1, \ldots, n\}$. Let $n_l = |S_l|$, the cardinality of S_l. The sets S_l need not necessarily be pairwise disjoint, and in this case we would have $\sum_{l=1}^{L} n_l > n$. Let the diagonal matrices E_l be as in definition 2.1 and assume, in addition, that the ith diagonal entry of E_l is zero if $i \in \{1, \dots, n\} \setminus S_l$. For $l = 1, \dots, L$ define the projections $P_l : \mathbb{R}^n \to \mathbb{R}^n$ by

$$P_l(x)_i = \begin{cases} x_i, & \text{if } i \in S_l, \\ 0, & \text{otherwise,} \end{cases}$$

and $G_l(x, y) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ by

$$G_l(i)(x, y) = \begin{cases} F(i)\big((I - P_l)x + P_l y\big), & \text{if } i \in S_l, \\ F(i)(x_1, \dots, x_{i-1}, y_i, x_{i+1}, \dots, x_n), & \text{otherwise,} \end{cases} \tag{5}$$

where $F(i) = f_i$ as in (1), and similarly $G_l(i)(x, y)$ is the ith component of $G_l(x, y)$. Then (G_l, E_l), $l = 1, \dots, L$, is a nonlinear multisplitting of F, and it is called a nonlinear block-multisplitting based on the sets S_l, $l = 1, \dots, L$. Note that when computing y_l^k we are only interested in the components $y_l^k(i)$ such that $i \in S_l$, we actually solve the subsystem

$$G_l(i)\big(x^k, y_l^k\big) = 0, \quad i \in S_l,$$

with respect to $y_l^k(i)$, $i \in S_l$, i.e., we work with a system of dimension n_l, even if $y_l \in \mathbb{R}^n$.

Observe that when there is no overlap, i.e., when the sets S_l are pairwise disjoint, then $P_l = E_l$ and $I - P_l = \sum_{j \neq l} E_j$. When there is overlap these two formulas do not hold.

3. Asynchronous nonlinear multisplitting methods

Consider a nonlinear multisplitting $(\widehat{F}_l, E_l)$ of F as in definition 2.1, and consider the case where the parallel computer has L processors. Each processor performs iterations corresponding to one of the splittings (4) by collecting the needed vectors computed by the other processors, as in (3). The nonlinear multisplitting methods given in section 2 are essentially synchronous in the sense that each processor has to wait for the others to complete their concurrent tasks in order to get the needed data for the next iteration. In practice, one needs to add a synchronization mechanism to ensure that the algorithm is carried out correctly. The time necessary to carry out the synchronization mechanism, as well as the time a processor must wait until its data is ready to continue the calculation, adds overhead to the computation. The amount of this overhead depends of course on the different size and difficulty of each nonlinear system (4). If there is a perfect load balance, the synchronization would not cause much overhead. By removing this synchronization and letting each processor continue its iteration according to the information available at the moment, we obtain an asynchronous nonlinear multisplitting method.

Denote the (approximate) solution of $\widehat{F}_l(x, y_l) = 0$ by $y_l = T_l(x)$, $l = 1, \dots, L$. A computational model for the asynchronous nonlinear multisplitting method in the case of distributed memory machines, can be written as the following pseudocode for the lth processor, where an initial guess $x_j = x$, $j = 1, \dots, L$, is assumed.

Computational model 3.1 (Asynchronous nonlinear multisplitting).

Until convergence do:
  use the vectors x_j (or $E_j x_j$), $j = 1, \ldots, L$, $j \neq l$, which have been received from the other processors to accumulate $x := \sum_{j=1}^{L} E_j x_j$;
  compute $x_l := T_l(x)$;
  send x_l (or $E_l x_l$) to all other processors.

There are several alternative termination criteria. For discussions of these, we refer the reader to [10,20] (and also [7]).

In order to analyze the convergence of the asynchronous nonlinear multisplitting (ANM) method we use a mathematical model describing the computational model 3.1. To that end, consider a counter k, which is updated every time a new vector is computed by some processor and let $x_l^0 = x^0$, $l = 1, \ldots, L$, be the initial guess; see, e.g., [4,6,16,24], and the references given therein. Then we write

$$x_l^k = \begin{cases} T_l\left(\sum_{j=1}^{L} E_j x_j^{r(j,k)}\right), & \text{if } l \in J_k, \\ x_l^{k-1}, & \text{otherwise}, \end{cases} \tag{6}$$

for $k = 1, 2, \ldots,$ where the sets J_k and the sequence $r(j,k)$ ($j = 1, \ldots, L$, $k = 1, 2, \ldots$) satisfy the following standard conditions:

(i) $r(j,k) \leqslant k - 1$, for all $j = 1, \ldots, L$, $k = 1, 2, \ldots,$

(ii) $\lim_{k \to \infty} r(j,k) = \infty$, for $j = 1, \ldots, L$,

(iii) each J_k, $k = 1, 2, \ldots,$ is a subset of $\{1, \ldots, L\}$, and the set $\{k\colon l \in J_k\}$ is unbounded for each $l = 1, \ldots, L$.

Remark 3.2. By using the computational model 3.1 to analyze the way the nonlinear multisplitting method works, we obtain the model (6), where T_l operates on the vector $\sum_{j=1}^{L} E_j x_j^{r(j,k)}$ if $l \in J_k$. This is different than the usual multisplitting approach; see [18], and also [8].

Nonlinear multisplitting methods for the systems (1) require the solutions of the nonlinear systems (4) at each step. The exact solutions of these nonlinear systems are generally not available. Instead, some iterative method is used to approximate their solution. In this paper we consider four such iterative methods: Newton's method, Newton–SOR method, Nonlinear SOR method, and SOR–Newton method; see, e.g., [19] for descriptions of these. Each of these produces a different variant of the ANM method (6). We end this section by briefly describing each of them. Again, an initial vector $x_l = x$, $l = 1, \ldots, L$, is assumed.

Assume that $\widehat{F}_l$ is differentiable, $l = 1, \ldots, L$. Let $\partial_1 \widehat{F}_l(x,y)$ and $\partial_2 \widehat{F}_l(x,y)$ denote the partial derivative of $\widehat{F}_l(x,y)$ with respect to x and y, respectively. Let

$\partial_2 F_l(x, x) = D_l(x) - V_l(x) - U_l(x)$, where $D_l(x)$, $V_l(x)$ and $U_l(x)$ are diagonal, strictly lower, and strictly upper triangular matrices, respectively. Applying one step of Newton's method to solve the nonlinear systems in (4) yields the ANM-Newton method (6) with T_l given by

$$T_l(x) = x - \partial_2 \widehat{F}_l(x, x)^{-1} F(x).$$

Similarly, applying one step of the Newton–SOR method yields the ANM–Newton–SOR method (6) with T_l given by

$$T_l(x) = x - A_l(x)^{-1} F(x),$$

where $A_l(x) = (1/\omega_l)(D_l(x) - \omega_l V_l(x))$, and $\omega_l > 0$.

Let $y_l = T_l(x)$ denote the approximation to the solution of $\widehat{F}_l(x, y_l) = 0$ by one step of the nonlinear SOR method. Then we get the ANM–SOR method (6). For $j = 1, \ldots, n$, the component $y_l(j)$ of y_l is calculated successively in the following way: Solve $\widehat{F}_l(j)(x, \bar{y}_l) = 0$ with respect to $\bar{y}_l(j)$, where

$$\bar{y}_l = \bigl(y_l(1), \ldots, y_l(j-1), \bar{y}_l(j), x_l(j+1), \ldots, x_l(n)\bigr)$$

and set $y_l(j) = x(j) - \omega_l(x(j) - \bar{y}_l(j))$, with the relaxation parameter $\omega_l > 0$.

Finally, if $y_l = T_l(x)$ denotes the approximation to the solution of $\widehat{F}_l(x, y_l) = 0$ by one step of the SOR–Newton method, we obtain the ANM–SOR–Newton method (6). For $j = 1, \ldots, n$, the components $y_l(j)$ of y_l are computed successively according to

$$y_l(j) = x(j) - \frac{\omega_l \widehat{F}_l(j)(x, \bar{y}_l)}{\alpha(j)},$$

where $\bar{y}_l = (y_l(1), \ldots, y_l(j-1), x(j), \ldots, x(n))$, $\alpha(j)$ denotes the value of the partial derivative of $\widehat{F}_l(j)(x, y)$ with respect to $y(j)$ evaluated at $(x, \bar{y})$, and the relaxation parameter $\omega_l > 0$.

4. Local convergence

In this section we give local convergence theorems for the ANM method and the four variants presented in section 3. We begin by recalling the following definition.

Definition 4.1 [19]. Let $x^* \in G: D \subseteq \mathbb{R}^n \to \mathbb{R}^n$. Then x^* is a point of attraction of the iteration $x^{k+1} = G(x^k)$, $k = 0, 1, \ldots,$ if there is a neighborhood U of x^* such that $U \subseteq D$, and for any $x^0 \in U$ the iterates x^k remain in D and converge to x^*.

Similarly, let $F: D \subseteq \mathbb{R}^n \to \mathbb{R}^n$, such that $F(x^*) = 0$. We say that x^* is a point of attraction of the asynchronous iteration (6), if there is a neighborhood U of x^* such that $U \subseteq D$, and for any $x^0 \in U$, $x_l^0 = x^0$, $l = 1, \ldots, L$, the iterates $x^k = \sum_{j=1}^{L} E_j x_j^k$ remain in D and converge to x^*.

To analyze the convergence of (6) one could consider a sequence x^k of weighted vectors of the form $x^k = \sum_{j=1}^{L} E_j x_j^k$. Instead, we study the convergence of a sequence of corresponding augmented vectors $\hat{x}^k = ((x_1^k)^{\mathrm{T}}, \ldots, (x_L^k)^{\mathrm{T}})^{\mathrm{T}}$; cf. [2,14]. Because of the condition (2), the sequence of weighted vectors converges to the solution x^* if the sequence of the augmented vectors converges to $\hat{x}^* = (x^{*\mathrm{T}}, \ldots, x^{*\mathrm{T}})^{\mathrm{T}} \in \mathbb{R}^{nL}$.

Let $\rho(P)$ denote the spectral radius of the matrix P, and let the components of the vector $|v|$ be the absolute values of the components of the vector v. A similar definition applies to matrices. The following lemma is used in our proofs.

Lemma 4.2 [11]. Let $T = ((T_1)^{\mathrm{T}}, \ldots, (T_L)^{\mathrm{T}})^{\mathrm{T}} : D \subseteq \mathbb{R}^{nL} \to \mathbb{R}^{nL}$, $T(\hat{x}^*) = \hat{x}^*$. Assume that there exists a neighborhood $B(\hat{x}^*, \delta) = \{\hat{x}\colon \|\hat{x} - \hat{x}^*\| < \delta\} \subseteq D$, and a nonnegative matrix P with $\rho(P) < 1$ such that

$$\left|T(\hat{x}) - T(\hat{x}^*)\right| \leqslant P\left|\hat{x} - \hat{x}^*\right|, \quad \text{for } \hat{x} \in B(\hat{x}^*, \delta).$$

Then $\hat{x}^*$ is a point of attraction of the asynchronous iteration (6).

The following is the main result of the paper.

Theorem 4.3. Let $(\widehat{F}_l, E_l)$, $l = 1, \ldots, L$, be a nonlinear multisplitting of F. Suppose that x^* is a zero of F, and that for $l = 1, \ldots, L$, the function $\widehat{F}_l$ is continuously differentiable in a neighborhood of (x^*, x^*). Let $M_l = \partial_2\widehat{F}_l(x^*, x^*)$ and $N_l = -\partial_1\widehat{F}_l(x^*, x^*)$. Assume that M_l is nonsingular for $l = 1, \ldots, L$, and let

$$P = \operatorname{diag}\left(M_1^{-1}N_1, \ldots, M_L^{-1}N_L\right)\begin{bmatrix} E_1 & \ldots & E_L \\ \vdots & \ldots & \vdots \\ E_1 & \ldots & E_L \\ \vdots & \ldots & \vdots \\ E_1 & \ldots & E_L \end{bmatrix}, \tag{7}$$

where $\operatorname{diag}(M_1^{-1}N_1, \ldots, M_L^{-1}N_L) \in \mathbb{R}^{nL\times nL}$ is a block diagonal matrix. If

$$\rho\left(|P|\right) < 1, \tag{8}$$

then $\hat{x}^* = ((x^*)^{\mathrm{T}}, \ldots, (x^*)^{\mathrm{T}})^{\mathrm{T}}$ is a point of attraction of both the ANM method and the ANM–Newton method.

Proof. Let $\hat{x} = ((x_1)^{\mathrm{T}}, \ldots, (x_L)^{\mathrm{T}})^{\mathrm{T}} \in \mathbb{R}^{nL}$, and let $Q = (E_1, \ldots, E_L) \in \mathbb{R}^{n\times nL}$. We first prove the theorem for the ANM method. In this case, $T_l(Q\hat{x})$ is the exact solution y of $\widehat{F}_l(Q\hat{x}, y) = 0$. Since $\widehat{F}_l(Q\hat{x}^*, x^*) = F(x^*) = 0$, and $M_l = \partial_2\widehat{F}_l(x^*, x^*) = \partial_2\widehat{F}_l(Q\hat{x}^*, x^*)$ is nonsingular, we can apply the implicit function theorem (see, e.g., [19]) and thus for each $l = 1, \ldots, L$, there exist open neighborhoods $U_{l,1}$ of $\hat{x}^*$, $U_{l,2}$ of x^* such that the equation $\widehat{F}_l(Q\hat{x}, y) = 0$ has a unique solution $y = T_l(\hat{x}) \in U_{l,2}$, provided $\hat{x} \in U_{l,1}$. Moreover, T_l is continuously differentiable on $U_{l,1}$, and

$$\frac{\partial T_l(Q\hat{x})}{\partial\hat{x}} = -\partial_2\widehat{F}_l\left(Q\hat{x}, T_l(\hat{x})\right)^{-1}\partial_1\widehat{F}_l\left(Q\hat{x}, T_l(\hat{x})\right)Q.$$

Therefore we have that

$$\frac{\partial T_l(Q\hat{x}^*)}{\partial \hat{x}} = -\partial_2 \widehat{F}_l\bigl(x^*, x^*\bigr)^{-1} \partial_1 \widehat{F}_l\bigl(x^*, x^*\bigr) Q = M_l^{-1} N_l Q.$$

We may assume that $U_{l,1} = U$ for $l = 1, \ldots, L$. Then $T = ((T_1)^{\mathrm{T}}, \ldots, (T_L)^{\mathrm{T}})^{\mathrm{T}}$ is defined on U with $T(\hat{x}^*) = \hat{x}^*$. The map T is continuously differentiable on U and

$$T'(\hat{x}^*) = \left(\left(\frac{\partial T_1(Q\hat{x}^*)}{\partial \hat{x}}\right)^{\mathrm{T}}, \ldots, \left(\frac{\partial T_L(Q\hat{x}^*)}{\partial \hat{x}}\right)^{\mathrm{T}}\right)^{\mathrm{T}} = \begin{bmatrix} M_1^{-1} N_1 Q \\ \vdots \\ M_L^{-1} N_L Q \end{bmatrix} = P.$$

Because of the differentiability of T, we have that

$$T\bigl(\hat{x}\bigr) - \hat{x}^* = T\bigl(\hat{x}\bigr) - T\bigl(\hat{x}^*\bigr) = T'\bigl(\hat{x}^*\bigr)\bigl(\hat{x} - \hat{x}^*\bigr) + \mathrm{o}\bigl(\|\hat{x} - \hat{x}^*\|\bigr),$$

where we can assume that the vector norm is the 1-norm. Thus, for any $\varepsilon > 0$, we can choose U small enough such that

$$\bigl|T\bigl(\hat{x}\bigr) - \hat{x}^*\bigr| \leqslant \bigl|T'\bigl(\hat{x}^*\bigr)\bigr|\,\bigl|\hat{x} - \hat{x}^*\bigr| + \varepsilon e\|\hat{x} - \hat{x}^*\| \leqslant \bigl(|P| + \varepsilon E\bigr)\bigl|\hat{x} - \hat{x}^*\bigr|,$$

for $\hat{x} \in U$, where $e \in \mathbb{R}^n$ and $E \in \mathbb{R}^{n\times n}$ have all entries equal to 1. Since $\rho(|P|) < 1$, by the continuity of the eigenvalues with respect to the entries of the matrix, we can choose ε small enough such that $\rho(|P| + \varepsilon E) < 1$. Thus by lemma 4.2, $\hat{x}^*$ is a point of attraction of the ANM method.

For the ANM–Newton method, we have $T_l(Q\hat{x}) = Q\hat{x} - \partial_2\widehat{F}_l(Q\hat{x}, Q\hat{x})^{-1}F(Q\hat{x})$. Since $\partial_2\widehat{F}_l(Q\hat{x}^*, Q\hat{x}^*) = \partial_2\widehat{F}_l(x^*, x^*)$ is nonsingular, for $l = 1, \ldots, L$, there is a neighborhood U_l of $\hat{x}^*$ such that $\partial_2\widehat{F}_l(Q\hat{x}, Q\hat{x})$ is nonsingular for $\hat{x} \in U_l$. Again we can assume that $U_l = U$ for $l = 1, \ldots, L$. Then $T = (T_1^{\mathrm{T}}, \ldots, T_L^{\mathrm{T}})^{\mathrm{T}}$ is well-defined on U, and moreover [19]

$$\begin{aligned}\frac{\partial T_l(Q\hat{x}^*)}{\partial \hat{x}} &= Q - \partial_2\widehat{F}_l\bigl(Q\hat{x}^*, Q\hat{x}^*\bigr)^{-1}F'\bigl(Q\hat{x}^*\bigr)Q \\ &= \bigl(I - \partial_2\widehat{F}_l\bigl(x^*, x^*\bigr)^{-1}F'\bigl(x^*\bigr)\bigr)Q = \bigl(I - M_l^{-1}(M_l - N_l)\bigr)Q = M_l^{-1}N_l Q.\end{aligned}$$

Thus we have $T'(\hat{x}^*) = P$ again, and the rest of the proof follows as that for the ANM method. □

We make several statements that relate to condition (8). The condition $\rho(|P|) < 1$ in theorem 4.3 can not be relaxed. Consider the nonsingular linear case, i.e., $F(x) = Ax - b$, with $A \in \mathbb{R}^{n\times n}$ nonsingular, and let $\widehat{F}_l(x, y) = M_l x - N_l y$ be a splitting of A. Since $F'(x) = A$, condition (8) is necessary and sufficient for the convergence of the asynchronous multisplitting method for nonsingular linear systems $F(x) = 0$; see, e.g., [6,9].

On the other hand, condition (8) can be obtained by knowing certain properties of the splittings as the following lemma indicates. To that end, consider a weighted max norm in $\mathbb{R}^n$ defined as $\|x\|_w = \max_{i=1}^n(\|x\|_i/w_i)$, where $w = (w_1, \ldots, w_n)$ is a positive vector, i.e., $w_i > 0$ for $i = 1, \ldots, n$.

Lemma 4.4. If there exists a weighted max norm $\|\cdot\|_w$ such that

$$\left\| M_l^{-1} N_l \right\|_w < 1, \quad l = 1, \dots, L, \tag{9}$$

and P is given by (7), then $\rho(|P|) < 1$.

Lemma 4.4 can be proved by directly checking that $\||P|\|_{\widehat{w}} < 1$, where $\widehat{w} = (w^{\mathrm{T}}, \dots, w^{\mathrm{T}})^{\mathrm{T}} \in \mathbb{R}^{nL}$. We conclude our comments on condition (8) by relating the spectral radius of $|P| \in \mathbb{R}^{nL \times nL}$ to that of an $n \times n$ matrix.

Lemma 4.5. Let P be defined as in (7). Then $\rho(|P|) < 1$ if and only if $\rho(R) < 1$, where $R = \sum_{l=1}^{L} E_l |M_l^{-1} N_l| \in \mathbb{R}^{n \times n}$.

Proof. Let

$$\widehat{P} = \begin{bmatrix} E_1 & \dots & E_L \\ \vdots & \dots & \vdots \\ E_1 & \dots & E_L \\ \vdots & \dots & \vdots \\ E_1 & \dots & E_L \end{bmatrix} \operatorname{diag}\bigl(M_1^{-1} N_1, \dots, M_L^{-1} N_L\bigr),$$

and since for any pair of square matrices $B,\ C,\ \rho(BC) = \rho(CB)$, from (7) we have that $\rho(|P|) = \rho(|\widehat{P}|)$. Thus, $\rho(|P|) = \rho(|\widehat{P}|) < 1$ implies that there exists a positive vector $v = [v_1^{\mathrm{T}}, \dots, v_L^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{R}^{nL}$ (with each $v_l \in \mathbb{R}^n$), such that $\sum_{l=1}^{L} E_L |M_l^{-1} N_l| v_l < v_j$ for all $j = 1, \dots, L$. Consider the positive vector $w \in \mathbb{R}^n$ whose components are taken as the minimum of the components of the v_j, i.e., $w = \min\{v_1, \dots, v_L\}$. For this w we have that $Rw < w$, and thus $\rho(R) \leqslant \|R\|_w < 1$. For the proof in the other direction, $\rho(R) < 1$ implies the existence of a positive vector $w \in \mathbb{R}^n$ such that $Rw < w$. Let $v = [w^{\mathrm{T}}, \dots, w^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{R}^{nL}$. Then we have $|\widehat{P}| v < v$ and the lemma follows. □

The following theorem states the local convergence of the other variants of the ANM method described in section 3. Its proof is analogous to that in theorem 4.3, and in some instances, by including the same arguments used in [19] for the proofs of local convergence of the nonlinear SOR and the SOR–Newton methods.

Theorem 4.6. Let the hypotheses of theorem 4.3 hold. In addition, let $\partial_2 F_l(x^*, x^*) = M_l = D_l - V_l - U_l$, where $D_l,\ V_l$ and U_l are diagonal, strictly lower, and strictly upper triangular matrices, respectively. Assume that all matrices D_l are nonsingular. Define the matrices

$$A_l = \frac{1}{\omega_l}(D_l - \omega_l V_l), \qquad B_l = \frac{1}{\omega_l}\bigl[(1 - \omega_l) D_l + \omega_l U_l\bigr], \quad \omega_l > 0 \tag{10}$$

and

$$P_1 = \operatorname{diag}\bigl(A_1^{-1}(B_1+N_1),\ldots,A_L^{-1}(B_L+N_L)\bigr)\begin{bmatrix} E_1 & \ldots & E_L \\ \vdots & \ldots & \vdots \\ E_1 & \ldots & E_L \\ \vdots & \ldots & \vdots \\ E_1 & \ldots & E_L \end{bmatrix}. \tag{11}$$

Assume that $\rho(|P_1|) < 1$. Then $\hat{x}^*$ is a point of attraction of the ANM–Newton–SOR method, the ANM–nonlinear-SOR method, and the ANM–SOR–Newton method.

We end this section with a comment on the relationship between some of the results of [1] and those in this paper. In [1], the convergence of multisplitting methods for nonlinear fixed point problems was studied. The basic assumption in [1] is that all fixed point mappings T_l, $l = 1,\ldots,L$, are contractive at the fixed point x^* with the same weighted maximum norm $\|\cdot\|_w$. When T_l is differentiable, this means $\|T_l'(x^*)\|_w < 1$, $l = 1,\ldots,L$. Under the setting of theorem 4.3, this is equivalent to $\|M_l^{-1}N_l\|_w < 1$, $l = 1,\ldots,L$; cf. lemma 4.4. As the following example indicates, there are cases in which condition (8) holds, but (9) does not. This implies that our convergence results are more general.

Example 4.7. Consider the linear case $F(x) = Ax - b$,

$$A = \begin{bmatrix} 0.5 & -1 \\ 0 & 0.5 \end{bmatrix} = M_1 - N_1 = M_2 - N_2,$$

$$M_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad M_2 = \begin{bmatrix} -3 & -2 \\ 2 & 1 \end{bmatrix}, \qquad E_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \qquad E_2 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

Then it is easy to check that $\rho(|M_l^{-1}N_l|) < 1$, $l = 1, 2$, and that

$$\rho\bigl(|P|\bigr) = \rho\left(\begin{bmatrix} |M_1^{-1}N_1| & 0 \\ 0 & |M_2^{-1}N_2| \end{bmatrix} \begin{bmatrix} E_1 & E_2 \\ E_1 & E_2 \end{bmatrix} \right) > 1.$$

Thus the condition (9) in lemma 4.4 cannot hold. But we can find different weighting matrices for which $\rho(|P|) < 1$, for example,

$$E_1 = \begin{bmatrix} 0.1 & 0 \\ 0 & 0.1 \end{bmatrix}, \qquad E_2 = \begin{bmatrix} 0.9 & 0 \\ 0 & 0.9 \end{bmatrix}.$$

5. Monotone matrix and H-matrix cases

In this section we present specific instances where we can guarantee, under certain general assumptions, that the matrices P of (7) and P_1 of (11) satisfy the hypotheses of theorems 4.3 and 4.6. These assumptions are drawn from the theory of regular splittings and other type of splittings of monotone and H-matrices; see the classic texts [5,19,25], and the paper [15] and references therein.

A nonsingular matrix A is monotone if $A^{-1} \geqslant 0$. A splitting $A = M - N$ called regular if $M^{-1} \geqslant 0$, $N \geqslant 0$, it is called weak regular if $M^{-1} \geqslant 0$, $M^{-1}N \geqslant 0$.

We consider here two distinctive splittings of $F'(x^*)$. Since $F(x) = \widehat{F}_l(x, x)$, for $l = 1, \dots, L$, we have $F'(x^*) = \partial_1 \widehat{F}_l(x^*, x^*) + \partial_2 \widehat{F}(x^*, x^*) = M_l - N_l$, providing us with the first of these splitting. We also consider the matrices defined in (10) which provide the splitting $F'(x^*) = A_l - (B_l + N_l)$. We consider here the cases when $F'(x^*)$ is either monotone or it is an H-matrix.

We use the well known fact that if for a positive vector $v \in \mathbb{R}^n$, a nonnegative matrix A satisfies $Av < v$, then $\|A\|_v < 1$. We also use the fact that if a matrix is invertible none of its rows can be zero rows.

Theorem 5.1. Let the hypotheses of theorem 4.3 hold. In addition, let $F'(x^*)$ be monotone.

(a) If (M_l, N_l) is a weak regular splitting of $F'(x^*)$, for $l = 1, ..., L$, then P as defined in (7) is nonnegative with spectral radius less than 1, and thus $\hat{x}^*$ is a point of attraction of both the ANM method and the ANM–Newton method.

(b) If (M_l, N_l) is a regular splitting of $F'(x^*)$ and (A_l, B_l) is a weak regular splitting of M_l, for $l = 1, \dots, L$, then P_1 as defined in (11) is nonnegative with spectral radius less than 1, and thus $\hat{x}^*$ is a point of attraction of the ANM–Newton–SOR method, the ANM–SOR method, and the ANM–SOR–Newton method.

Proof. (a) Let $e = (1, \dots, 1)^{\mathrm{T}} \in \mathbb{R}^n$, then $v = F'(x^*)^{-1}e > 0$. Thus $M_l^{-1}N_l v = M_l^{-1}(M_l - F'(x^*))v = v - M_l^{-1}e < v$, and since $M_l^{-1}N_l,\ l = 1, \dots, L$, are nonnegative, then (9) holds for the weighted max norm defined by v, and thus $\rho(|P|) = \rho(P) < 1$, by lemma 4.4.

(b) Since for $l = 1, \dots, L$,

$$A_l^{-1}(B_l + N_l) = A_l^{-1}(A_l - M_l + N_l) = I - A_l^{-1}(M_l - N_l) = I - A_l^{-1}F'\big(x^*\big),$$

the assertion follows using the same argument as in (a). □

A nonsingular matrix is called an M-matrix, if it is monotone and its off-diagonal entries are nonpositive. Given a matrix $A = (a_{i,j}) \in \mathbb{R}^{n\times n}$, its comparison matrix $\langle A\rangle = (\alpha_{i,j})$ is defined by $\alpha_{i,i} = |a_{i,i}|$, $\alpha_{i,j} = -|a_{i,j}|,\ i \neq j$. A matrix A is said to be an H-matrix if $\langle A\rangle$ is an M-matrix. A splitting $A = M - N$ is called an H-splitting if $\langle M\rangle - |N|$ is an M-matrix, it is called an H-compatible splitting if $\langle A\rangle = \langle M\rangle - |N|$. Two useful results which we use are that if A is an H-matrix, then $|A^{-1}| \leqslant \langle A\rangle^{-1}$, and if $A = M - N$ is an H-splitting, then $\rho(M^{-1}N) \leqslant \rho(\langle M\rangle^{-1}|N|) < 1$; see, e.g., [15] and references therein.

Theorem 5.2. Let the hypotheses of theorem 4.3 hold. In addition, let $F'(x^*)$ be an H-matrix. If (M_l, N_l), $l = 1, \ldots, L$, are H-compatible splittings of $F'(x^*)$, then $\rho(|P|) < 1$, where P is defined by (7), and thus $\hat{x}^*$ is a point of attraction of both the ANM method and the ANM–Newton method.

Proof. Since we have H-splittings, then $|M_l^{-1}N_l| \leqslant \langle M_l\rangle^{-1}|N_l|$, and since the comparison matrix $\langle F'(x^*)\rangle$ is monotone, the splittings $(\langle M_l\rangle, |N_l|)$, $l = 1, \ldots, L$, are regular splittings of $\langle F'(x^*)\rangle$. Following the arguments in the proof of theorem 5.1(a), we have now a vector $u = \langle F'(x^*)\rangle^{-1}e > 0$, such that $|M_l^{-1}N_l|u \leqslant \langle M_l\rangle^{-1}|N_l|u < u$, and the theorem follows in the same fashion. □

Theorem 5.3. Let the hypotheses of theorem 4.3 hold. In addition, let $F'(x^*)$ be an H-matrix. Let D, $-V$, and $-U$ be the diagonal, the strict lower and the strict upper triangular part of the comparison matrix $\langle F'(x^*)\rangle$, respectively. Assume further that (M_l, N_l) are H-compatible splittings of $F'(x^*)$, and that M_l and $F'(x^*)$ have the same diagonal part, i.e., that $D_l = D$, $l = 1, \ldots, L$. If $0 < \omega_l < 2/(1 + \rho(J))$, where $J = |D|^{-1}(|U| + |V|)$, then $\rho(|P_1|) < 1$, where P_1 is given by (11). Thus $\hat{x}^*$ is a point of attraction of the ANM–Newton–SOR method, the ANM–SOR method, and the ANM–SOR–Newton method.

Proof. Since $F'(x^*)$ and M_l have the same diagonal part,

$$\left\langle F'\left(x^*\right)\right\rangle = \langle M_l\rangle - |N_l| = |D| - |V_l| - |U_l| - |N_l| = |D| - \left(|U| + |V|\right),$$

and thus

$$\begin{aligned}\left|A_l^{-1}(B_l + N_l)\right| &= \left|(D - \omega_l V_l)^{-1}\left((1-\omega_l)D + \omega_l U_l + \omega_l N_l\right)\right| \\ &\leqslant \langle D - w_l V_l\rangle^{-1}\left(|1-\omega_l|\,|D| + \omega_l|U_l| + \omega_l|N_l|\right) \\ &= \left(|D| - \omega|V_l|\right)^{-1}\left[\left(|D| - \omega_l|V_l|\right) + \left(|1-\omega_l| - 1\right)|D| \right.\\ &\quad \left. + \,\omega_l\left(|U| + |V|\right)\right] \\ &= I - \left(|D| - \omega_l|V_l|\right)^{-1}|D|\left[\left(1 - |1-\omega_l|\right)I - \omega_l J\right].\end{aligned}$$

Since $F'(x^*)$ is an H-matrix, $\rho(J) < 1$. Let $J_\varepsilon = J + \varepsilon ee^{\mathrm{T}}$, where $e = (1, \ldots, 1)^{\mathrm{T}} \in \mathbb{R}^n$, $\varepsilon > 0$. Then $J_\varepsilon > 0$, and by the continuity of eigenvalues with respect to the entries, we can choose ε small enough such that $\rho_\varepsilon = \rho(J_\varepsilon) < 1$, and $0 < \omega_l < 2/(1 + \rho_\varepsilon)$, for $l = 1, \ldots, L$. Thus, there exists a vector $v_\varepsilon > 0$, such that $J_\varepsilon v_\varepsilon = \rho_\varepsilon v_\varepsilon$. Then $0 < 1 - |1-\omega_l| - \omega_l\rho_\varepsilon < 1$, and we have for $l = 1, \ldots, L$,

$$\begin{aligned}\left|A_l^{-1}(B_l + N_l)\right|v_\varepsilon &\leqslant v_\varepsilon - \left(|D| - \omega_l|V_l|\right)^{-1}|D|\left[\left(1 - |1-\omega_l|\right)I - \omega_l J_\varepsilon\right]v_\varepsilon \\ &\leqslant v_\varepsilon - \left(|D| - \omega_l|V_l|\right)^{-1}|D|\left[\left(1 - |1-\omega_l|\right) - \omega_\varepsilon\rho_\varepsilon\right]v_\varepsilon < v_\varepsilon.\end{aligned}$$

Therefore, using the same arguments as in theorem 5.1, $\rho(|P_1|) < 1$ and the proof is complete. □

We conclude the section by exhibiting (overlapping) nonlinear multisplittings of the form (5) for which theorems 5.1–5.3 apply. Assume that the function F is continuously differentiable in a neighborhood of a zero x^*, and that $F'(x^*)$ is an H-matrix. Let (G_l, E_l), $l = 1, \ldots, L$, be as in (5), based on the (overlapping) sets S_l. Define the matrices M_l, N_l, $l = 1, \ldots, L$ by

$$(M_l)_{ij} = \begin{cases} F'(x^*)_{ij}, & \text{if } i, j \in S_l, \\ F'(x^*)_{ii}, & \text{if } i \in \{1, \ldots, n\} \setminus S_l, \\ 0, & \text{otherwise.} \end{cases} \qquad N_l = M_l - F'(x^*).$$

Then, our theorems apply since we have $M_l = \partial_2 G_l(x^*, x^*)$, $N_l = -\partial_1 G_l(x^*, x^*)$, and for $l = 1, \ldots, L$, (M_l, N_l) are H-compatible splittings of $F'(x^*)$.

6. A numerical example

Let us consider the following semilinear elliptic partial differential equation used, e.g., in [13,21,27].

$$\begin{cases} -\left(K^1 u_x\right)_x - \left(K^2 u_y\right)_y = -g\mathrm{e}^u, & (x, y) \in \Omega, \\ u = x^2 + y^2, & (x, y) \in \partial\Omega, \end{cases} \tag{12}$$

where

$$K^1 = 1 + x^2 + y^2, \qquad K^2 = 1 + \mathrm{e}^x + \mathrm{e}^y,$$
$$g(x, y) = 2\left(2 + 3x^2 + y^2 + \mathrm{e}^x + (1 + y)\mathrm{e}^y\right)\mathrm{e}^{-x^2-y^2},$$

and $\Omega = \Omega_1 \cup \Omega_2$, $\Omega_1 = (0, 1) \times (0, \frac{2}{3})$, $\Omega_2 = (0, \frac{1}{3}) \times (0, 1)$; see figure 1. Problem (12) has the unique solution $u(x, y) = x^2 + y^2$. We use standard finite differences with step length $h = 1/(3m)$ to discretize (12). Let Ω_h denote the set of all the grid points, then at each grid point (ih, jh) we obtain the equation

$$a_{i,j}u_{i-1,j} + b_{i,j}u_{i+1,j} + c_{i,j}u_{i,j+1} + d_{i,j}u_{i,j-1} + f_{i,j}u_{i,j} + h^2 g_{i,j}\mathrm{e}^{u_{i,j}} = 0, \tag{13}$$

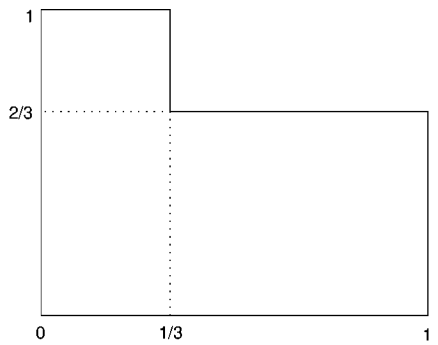

Figure 1. The L-shaped domain $\Omega = \Omega_1 \cup \Omega_2$.

where

$$
\begin{aligned}
a_{i,j} &= -K^1\left(\left(i-\frac{1}{2}\right)h,\, jh\right), & b_{i,j} &= -K^1\left(\left(i+\frac{1}{2}\right)h,\, jh\right),\\
c_{i,j} &= -K^2\left(ih,\, \left(j-\frac{1}{2}\right)h\right), & d_{i,j} &= -K^2\left(ih,\, \left(j+\frac{1}{2}\right)h\right),\\
f_{i,j} &= -(a_{i,j}+b_{i,j}+c_{i,j}+d_{i,j}), & g_{i,j} &= g(ih, jh).
\end{aligned}
$$

In (13), $u_{i,j}$ is the numerical approximation to $u(ih, jh)$, u being the exact solution of (12). The values for $u_{i,j}$ are unknown unless (ih, jh) is on the boundary of Ω. Thus (13) describes a system of $7m^2 - 6m + 1$ unknowns. The unknowns and the equations in (13) are numbered in the natural (lexicographical) order. In our calculation we considered three cases: $m = 4$, 8 and 12. All iterations were started with an initial guess having all components equal to 1. Since we know the exact solution, we used an artificial termination criterion, namely to stop when $\|x^k - v\|_2 \leqslant h^2$, where $\|\cdot\|_2$ denotes the Euclidean norm, x^k denotes the kth iterate and v is the vector representing the exact solution of (12), i.e., $v_{i,j} = (ih)^2 + (jh)^2$, $(ih, jh) \in \Omega_h$. Note that, as the system grows (h smaller) this stopping criteria produces more accurate solutions. Of course, in real applications, other stopping criteria are used; see the references [7,10,20] already mentioned.

Let us rewrite (13) as $F(x) = 0$. Let $\widehat{F}_l, E^l$, $l = 1, 2$ be the block-multisplitting belonging to the following two sets (see figure 1):

$$
\begin{aligned}
S_1 &= \bigl\{(i,j),\ 1 \leqslant i \leqslant m,\ 1 \leqslant j \leqslant 2m\bigr\} \cup \bigl\{(i,j),\ 1 \leqslant i < m,\\
&\qquad\qquad 2m+1 \leqslant j \leqslant 3m-1\bigr\},\\
S_2 &= \bigl\{(i,j),\ 1 \leqslant i \leqslant m,\ 1 \leqslant j \leqslant 2m\bigr\} \cup \bigl\{(i,j),\ m+1 \leqslant i \leqslant 3m-1,\\
&\qquad\qquad 1 \leqslant j \leqslant 2m-1\bigr\}.
\end{aligned}
$$

A diagonal entry of E_l ($l = 1, 2$) is 0.5 if the corresponding index is in $S_1 \cap S_2$, it is 1 if the corresponding index is in $S_l \setminus S_1 \cap S_2$, otherwise it is 0. The asynchronous block-Newton method is used. We implemented the asynchronous algorithm, its synchronous counterpart, and the sequential algorithm (i.e., sequential block-Jacobi method with overlap using the same code as the synchronous version with one processor) by using the Synergy software [22] on two DEC alpha 3000/300 computers connected by 10 Mbit 10-Base T Ethernet.

The required number of iteration steps for convergence for each block, and the elapsed time (in seconds) are reported in table 1. As is to be expected the asynchronous methods take more iterations, but in a system as small as 401 variables with only two processors, its convergence is faster than the synchronous counterpart. In the case of $n = 937$, it is considerably faster.

Table 1
Elapsed time and number of iterations.

m	n	Method	Elapsed time	Iter. block 1	Iter. block 2
4	89	asynchronous	1.09	6	4
		synchronous	0.85	3	3
		sequential	0.62	3	3
8	401	asynchronous	40.22	10	6
		synchronous	65.51	4	4
		sequential	56.17	4	4
12	937	asynchronous	595.67	12	9
		synchronous	957.20	5	5
		sequential	917.13	5	5

Acknowledgements

We would like to thank Professor Yuan Shi for making available the Synergy software, and to Anh-Dung Tran for his help during the parallel computations which led to the results in table 1. We thank Andreas Frommer, Violeta Migallón, and José Penadés, for reading an earlier version of the paper, pointing out some errors, and making useful comments, all of which helped improve our presentation. We also thank one of the referees for a careful reading of the manuscript, for several insightful observations, and for the suggestions that led to lemma 4.5.

References

[1] J. Bahi, J.-C. Miellou and K. Rhofir, Asynchronous multisplitting methods for nonlinear fixed point problems, Numer. Algorithms 15 (1997) 315–345.

[2] Z.-Z. Bai, On the convergence of parallel chaotic nonlinear multisplitting Newton-type methods, J. Comput. Appl. Math. 80 (1997) 317–334.

[3] Z.-Z. Bai, V. Migallón, J. Penadés and D.B. Szyld, Block and asynchronous two-stage methods for mildly nonlinear systems, Numer. Math. 82 (1999) 1–20.

[4] G.M. Baudet, Asynchronous iterative methods for multiprocessors, J. Assoc. Comput. Mach. 25 (1978) 226–244.

[5] A. Berman and R.J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*, 3rd ed. (Academic Press, New York, 1979); reprinted (SIAM, Philadelphia, 1994).

[6] D.P. Bertsekas and J.N. Tsitsiklis, *Parallel and Distributed Computation* (Prentice-Hall, Englewood Cliffs, NJ, 1989).

[7] K. Blathras, D.B. Szyld and Y. Shi, Timing models and local stopping criteria for asynchronous iterative algorithms, J. Parallel Distributed Comput. 58 (1999) 446–465.

[8] R. Bru, V. Migallón, J. Penadés and D.B. Szyld, Parallel, synchronous and asynchronous two-stage multisplitting methods, Electronic Trans. Numer. Anal. 3 (1995) 24–38.

[9] D. Chazan and W. Miranker, Chaotic relaxation, Linear Algebra Appl. 2 (1969) 199–222.

[10] D. El Baz, A method of terminating asynchronous iterative algorithms on message passing systems, Parallel Algorithms Appl. 9 (1996) 153–158.

[11] M.N. El Tarazi, Some convergence results for asynchronous algorithms, Numer. Math. 39 (1982) 325–340.

[12] D.J. Evans and D.-R. Wang, An asynchronous algorithm for solving a class of a nonlinear simultaneous equations, Parallel Computing 17 (1991) 165–180.

[13] A. Frommer, Parallel nonlinear multisplitting methods, Numer. Math. 56 (1989) 269–282.

[14] A. Frommer, H. Schwandt and D.B. Szyld, Asynchronous weighted additive Schwarz methods, Electronic Trans. Numer. Anal. 5 (1997) 48–61.

[15] A. Frommer and D.B. Szyld, H-splittings and two-stage iterative methods, Numer. Math. 63 (1992) 345–356.

[16] A. Frommer and D.B. Szyld, On asynchronous iterations, J. Comput. Appl. Math. 123 (2000) 201–216.

[17] W. Mönch, Convergence of parallel nonlinear regular and weak regular multisplitting methods, Technical Report 96-10, Fakultät für Mathematik und Informatik, Technische Universität Bergakademie Freiberg (1996).

[18] D.P. O'Leary and R.E. White, Multi-splittings of matrices and parallel solution of linear systems, SIAM J. Algebraic Discrete Methods 6 (1985) 630–640.

[19] J.M. Ortega and W.C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables* (Academic Press, New York, 1970).

[20] S.A. Savari and D.P. Bertserkas, Finite termination of asynchronous iterative algorithms, Parallel Computing 22 (1996) 39–56.

[21] A.H. Sherman, On Newton-iterative methods for the solution of systems of nonlinear equations, SIAM J. Numer. Anal. 15 (1978) 755–771.

[22] Y. Shi, The SAG distributed programming model and its applications to scientific computing, in: *Proc. of the 5th SIAM Conf. on Parallel Processing for Scientific Computing*, eds. J. Dongarra, K. Kennedy, P. Messina, D. Sorensen and R. Voigt (SIAM, Philadelphia, PA, 1991); see also `http://www.cis.temple.edu/~shi/synergy`.

[23] X. Sun, A generalization of parallel multisplitting methods for a nonlinear problem, Semiannual Journal of Nanjing University, Mathematics (Nan Jing Da Xue Xue Bao Shu Xue Ban Yue Kan) 1 (1992).

[24] D.B. Szyld, Different models of parallel asynchronous iterations with overlapping blocks, Comput. Appl. Math. 17 (1998) 101–115.

[25] R.S. Varga, *Matrix Iterative Analysis* (Prentice-Hall, Englewood Cliffs, NJ, 1962); 2nd ed., revised and expanded: (Springer, Berlin, 2000).

[26] R.E. White, A nonlinear parallel algorithm with application to the Stefan problem, SIAM J. Numer. Anal. 23 (1986) 639–652.

[27] R.E. White, Parallel algorithms for nonlinear problems, SIAM J. Algebraic Discrete Methods 7 (1986) 137–149.

[28] J.-J. Xu, Convergence of partially asynchronous block quasi-Newton methods for nonlinear systems of equations, J. Comput. Appl. Math. 103 (1999) 307–321.

[29] J.-J. Xu and H.-H. Wang, A class of asynchronous block methods for nonlinear systems of equations, in: *Proc. of the 7th SIAM Conf. on Parallel Processing for Scientific Computing*, eds. D.H. Bailey et al. (SIAM, Phialdeplphia, PA, 1995) pp. 108–109.